ARTIFICIAL NEURAL NETWORK APPLICATIONS IN GEOTECHNICAL ENGINEERING

Mohamed A. Shahin, Mark B. Jaksa and Holger R. Maier
Department of Civil and Environmental Engineering, Adelaide University

ABSTRACT

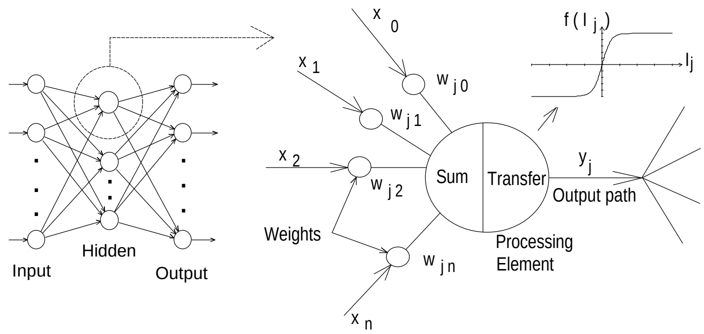
Over the last few years or so, the use of artificial neural networks (ANNs) has increased in many areas of engineering. In particular, ANNs have been applied to many geotechnical engineering problems and have demonstrated some degree of success. A review of the literature reveals that ANNs have been used successfully in pile capacity prediction, modelling soil behaviour, site characterisation, earth retaining structures, settlement of structures, slope stability, design of tunnels and underground openings, liquefaction, soil permeability and hydraulic conductivity, soil compaction, soil swelling and classification of soils. The objective of this paper is to provide a general view of some ANN applications for solving some types of geotechnical engineering problems. It is not intended to describe the ANNs modelling issues in geotechnical engineering. The paper also does not intend to cover every single application or scientific paper that found in the literature. For brevity, some works are selected to be described in some detail, while others are acknowledged for reference purposes. The paper then discusses the strengths and limitations of ANNs compared with the other modelling approaches.

1 INTRODUCTION

The engineering properties of soil and rock exhibit varied and uncertain behaviour due to the complex and imprecise physical processes associated with the formation of these materials (Jaksa 1995). This is in contrast to most other civil engineering materials, such as steel, concrete and timber, which exhibit far greater homogeneity and isotropy. In order to cope with the complexity of geotechnical behaviour, and the spatial variability of these materials, traditional forms of engineering design models are justifiably simplified. An alternative approach, which has been shown to have some degree of success, is based on the data alone to determine the structure and parameters of the model. The technique is known as artificial neural networks (ANNs) and is well suited to model complex problems where the relationship between the model variables is unknown (Hubick 1992). This paper is intended to be for readers in the field of geotechnical engineering who are not familiar with artificial neural networks. The paper aims to detail some features associated with ANNs through a review for some of their applications to-date in geotechnical engineering. It is hoped that this review may attract more geotechnical engineers to pay better attention to this promising tool. The paper starts with a brief overview of the structure and operation of the ANNs and gives a general overview of most ANN applications that have appeared in the geotechncial engineering literature. Finally, the paper discusses the relative success of ANNs in predicting various geotechnical engineering properties and behaviour.

2 OVERVIEW OF ARTIFICIAL NEURAL NETWORKS

Artificial neural networks (ANNs) are a form of artificial intelligence which attempt to mimic the behaviour of the human brain and nervous system. A comprehensive description of ANNs is beyond the scope of this paper. Many authors have described the structure and operation of ANNs (e.g. Hecht-Nielsen 1990; Maren et al. 1990; Zurada 1992; Fausett 1994; Ripley 1996). A typical structure of ANNs consists of a number of processing elements (PEs), or nodes, that are usually arranged in layers: an input layer, an output layer and one or more hidden layers (Figure 1).



The input from each PE in the previous layer (x_i) is multiplied by an adjustable connection weight (w_{ji}). At each PE, the weighted input signals are summed and a threshold value (q_j) is added. This combined input (I_j) is then passed through a non-linear transfer function (f(.)) to produce the output of the PE (y_j). The output of one PE provides the input to the PEs in the next layer. This process is summarised in Equations 1 and 2 and illustrated in Figure 1.

$$I_j = \sum w_{ji} x_i + q_j \qquad \text{summation} \qquad (1)$$

$$y_j = f(I_j) \qquad \text{transfer} \qquad (2)$$

The propagation of information in ANNs starts at the input layer where the input data are presented. The network adjusts its weights on the presentation of a training data set and uses a learning rule to find a set of weights that will produce the input/output mapping that has the smallest possible error. This process is called “learning” or “training”. Once the training phase of the model has been successfully accomplished, the performance of the trained model has to be validated using an independent testing set. Details of the ANN modelling process and development are beyond the scope of this paper and are given elsewhere (e.g. Moselhi et al. 1992; Flood and Kartam 1994; Maier and Dandy 2000).

As described above, ANNs learn from data examples presented to them and use these data to adjust their weights in an attempt to capture the relationship between the model input variables and the corresponding outputs. Consequently, ANNs do not need any prior knowledge about the nature of the relationship between the input/output variables, which is one of the benefits that ANNs have compared with most empirical and statistical methods.

The ANN modelling philosophy is similar to a number of conventional statistical models in the sense that both are attempting to capture the relationship between a historical set of model inputs and corresponding outputs. For example, suppose a set of x-values and corresponding y-values in 2 dimensional space, where $y = f(x)$. The objective is to find the unknown function f, which relates the input variable x to the output variable y. In a linear regression model, the function f can be obtained by changing the slope tanö and intercept â of the straight line in Figure 2, so that the error between the actual outputs and outputs of the straight line is minimised. The same principle is used in ANN models. ANNs can form the simple linear regression model by having one input, one output, no hidden layer nodes and a linear transfer function (Figure 3). The connection weight w in the ANN model is equivalent to the slope tanö and the threshold è is equivalent to the intercept â, in the linear regression model. ANNs adjust their weights by repeatedly presenting examples of the model inputs and outputs in order to minimise an error function between the historical outputs and the outputs predicted by the ANN model.

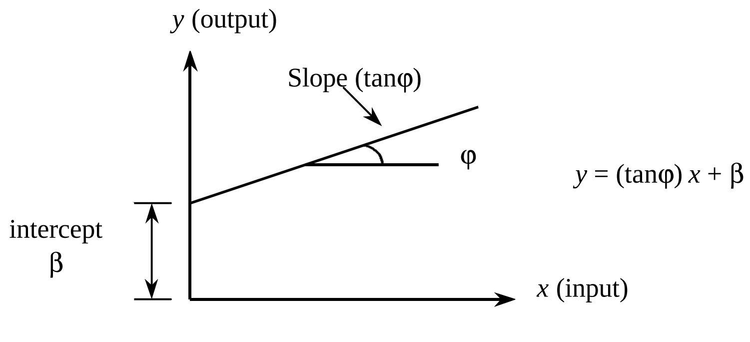


Figure 2 Linear regression model

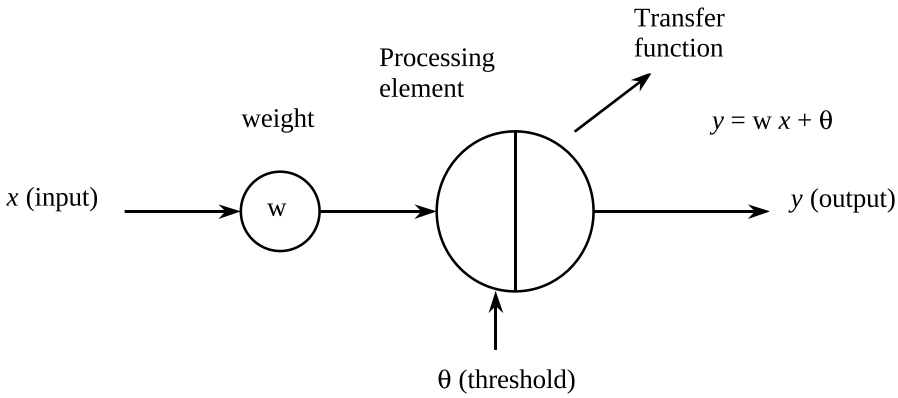


Figure 3 ANN representation of a linear regression model

If the relationship between *x* and *y* is non-linear, regression analysis can only be successfully applied if prior knowledge of the nature of the non-linearity exists. On the contrary, this prior knowledge of the nature of the non-linearity is not required for ANN models. In the ANN model, the degree of non-linearity can be also changed easily by changing the transfer function and the number of hidden layer nodes. In the real world, it is likely to encounter problems that are complex and highly non-linear. In such situations, traditional regression analysis is not adequate (Gardner, 1998). In contrast, ANNs can be used to deal with this complexity by changing the transfer function or network structure, and the type of non-linearity can be changed by varying the number of hidden layers and the number of nodes in each layer. In addition, ANN models can be upgraded from univariate to multivariate by increasing the number of input nodes.

3 ANN APPLICATIONS IN GEOTECHNICAL ENGINERING

3.1 PILE CAPACITY

The prediction of the load capacity, particularly those based on pile driving data, has been examined by several ANN researchers. Goh (1994a; 1995b) presented a neural network to predict the friction capacity of piles in clays. The neural network was trained with field data of actual case records. The model inputs were considered to be the pile length, the pile diameter, the mean effective stress and the undrained shear strength. The skin friction resistance was the only model output. The results obtained by utilising the neural network were compared with the results obtained by the method of Semple and Rigden (1986) and the â method (Burland 1973). The methods were compared using regression analysis as well as the error rate as shown in Table 1. It is evident from Table 1 that ANNs outperform the conventional methods. The study also pointed out that the main criticism of the ANN methodology is its inability to trace and explain the logic it uses to arrive at the outputs from the inputs.

Method	**Coefficient of correlation**		**Error rate (kPa)**	
	Training	**Testing**	**Training**	**Testing**
Neural network	0.985	0.956	1.016	1.194
Semple and Rigden (1986)	0.976	0.885	1.318	1.894
â method	0.731	0.704	4.824	3.096

Table 1 Summary of correlation coefficients and error rate for friction pile capacity (Goh 1995)

Goh (1995a; 1996b), soon after, developed another neural network to estimate the ultimate load capacity of driven piles in cohesionless soils. In this study, the data used were derived from the results of actual load tests on timber, precast concrete and steel piles driven into sandy soils. The inputs to the ANN model that were found to be more significant were the hammer weight, the hammer drop, the pile length, the pile weight, the pile cross sectional area, the pile set, the pile modulus of elasticity and the hammer type. The model output was the pile load capacity. When the model was examined with the testing set, it was observed that the neural network successfully modelled the pile load capacity. By examining the connection weights, it was observed that the more important input factors are the pile set, the hammer weight and the hammer type. The study compared the results obtained by the neural networks with the following common relationships: the Engineering News formula (Wellington 1892), the Hiley formula (Hiley 1922) and the Janbu formula (Janbu 1953). Regression analysis was carried out to obtain the coefficients of correlation of predicted versus measured results for neural networks and the traditional methods. Table 2 summarises the regression analysis results which indicate that the neural network predictions of the load capacity of driven piles were found to be better than these obtained using the other methods.

Method	Coefficient of correlation	
	Training data	Testing data
Neural network	0.96	0.97
Engineering News	0.69	0.61
Hiley	0.48	0.76
Janbu	0.82	0.89

Table 2 Summary of regression analysis results of pile capacity prediction (Goh 1995)

Chan et al. (1995) developed a neural network as an alternative to pile driving formulae. The network was trained with the same input parameters listed in the simplified Hiley formula (Broms and Lim 1988), including the elastic compression of the pile and soil, the pile set and the driving energy delivered to the pile. The model output considered was the pile capacity. The desired output value of the pile capacity that was used in the training process was estimated by using a commercial computer code called CAPWAP (Rausche et al. 1972) or the CASE method (Goble et al. 1975). The root mean square percentage error of the neural network was 13.5% for the training set, and 12.0% for the testing set.

Lee and Lee (1996) utilised neural networks to predict the ultimate bearing capacity of piles. The problem was simulated using data obtained from model pile load tests using a calibration chamber and results of in-situ pile load tests. For the simulation using the model pile load test data, the model inputs were the penetration depth ratio (i.e. penetration depth of pile/pile diameter), the mean normal stress of the calibration chamber and the number of blows. The ultimate bearing capacity was the model output. The prediction of the ANN model showed maximum error not more than 20% and average summed square error less than 15%. For the simulation using the in-situ pile load test data, five input variables were used representing the penetration depth ratio, the average standard penetration number along the pile shaft, the average standard penetration number near the pile tip, pile set and hammer energy. Two neural network models were developed. The results of these models were compared with Meyerhof's equation (Meyerhof 1976) based on the average standard penetration value. Figure 4 shows the plots of the testing set results of estimated versus measured pile bearing capacity obtained from the neural network models and Meyerhof's equation. The plots in Figure 4 show that the predicted values from the neural networks matched the measured values much better than those obtained from Meyerhof's equation.

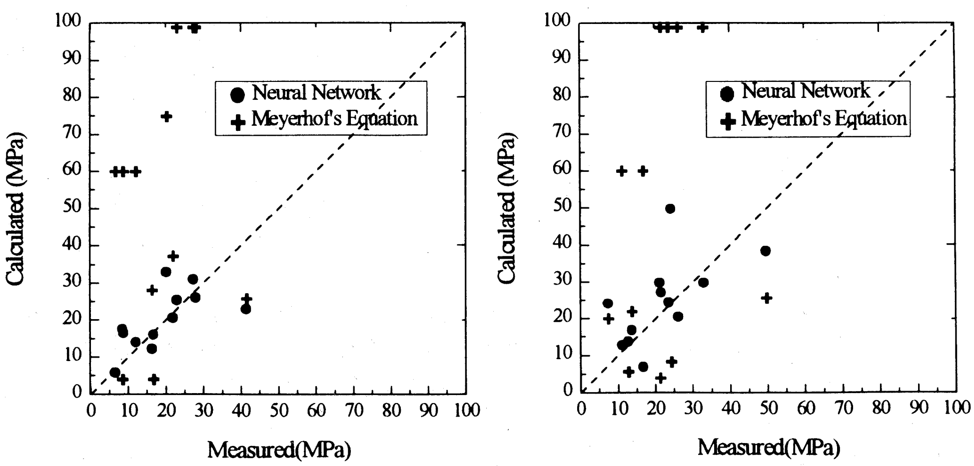


Figure 4 Testing results of predicted vs measured pile bearing capacity from in-situ pile load test (Lee and Lee 1996)

Abu-Kiefa (1998) introduced three neural network models (referred to the paper as GRNNM1, GRNNM2 and GRNNM3) to predict the capacity of driven piles in cohesionless soils. The first model was developed to estimate the total pile capacity. The second model was employed to estimate the tip pile capacity, whereas the final model was used to estimate the shaft pile capacity. Five variables were selected to be the model inputs in the first model. These inputs were the angle of shear resistance of the soil around the shaft, the angle of shear resistance at the tip of the pile, the effective overburden pressure at the tip of the pile, the pile length and the equivalent cross-sectional pile area. The model had one output representing the total pile capacity. In the model used to evaluate the pile tip capacity, the above variables were also used. The input variables used to predict the pile shaft capacity were four, representing the average standard penetration number around the shaft, the angle of shear resistance around the shaft, pile length and pile diameter. The results of the networks obtained in this study were compared with four other empirical techniques. These techniques were those proposed by Meyerhof (1976), Coyle and Castello (1981), the American Petroleum Institute (1984) and Randolph (1985). The results of the total pile capacity prediction demonstrated high coefficients of determination (0.95) for all data records obtained from the neural network model, while they ranged between 0.52 and 0.63 for the other methods. Figures 5 to 7 show the measured versus predicted values of all data records for the pile capacity, tip pile capacity and shaft pile capacity, respectively. It can be seen from these figures that the predictions of the neural networks produce less scatter than the predictions of all other methods, and thus provide the best prediction of pile load capacity, tip pile capacity and shaft pile capacity.

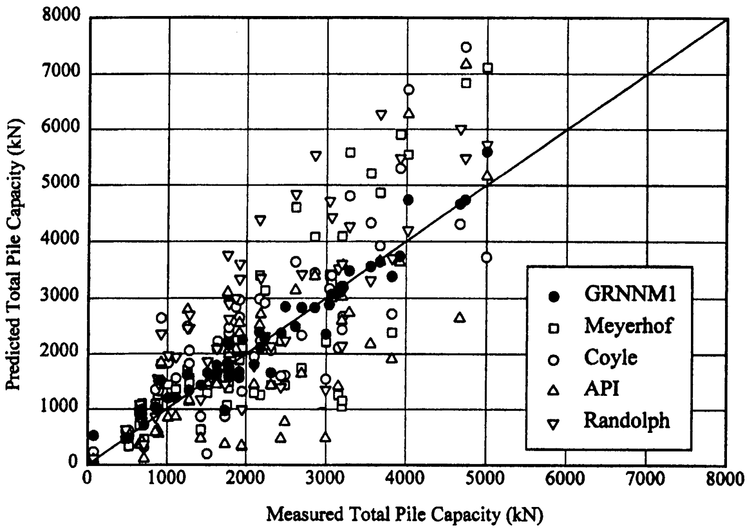


Figure 5 Comparison of predicted and measured total pile capacity (Abu-Kiefa 1998)

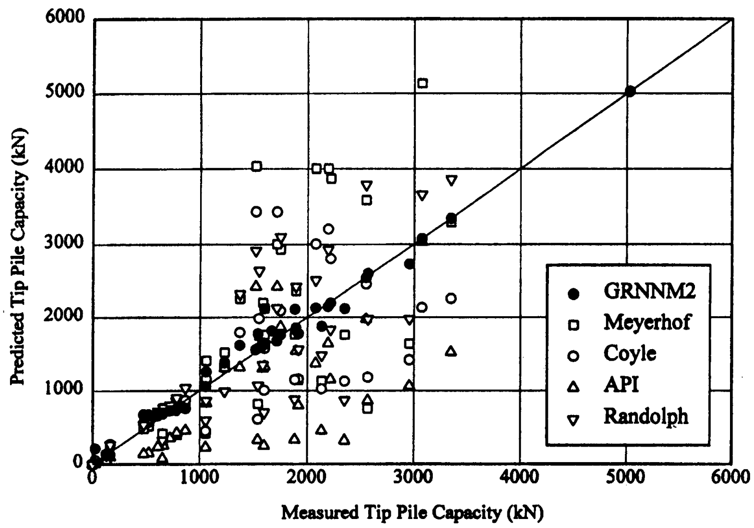


Figure 6 Comparison of predicted and measured tip pile capacity (Abu-Kiefa 1998)

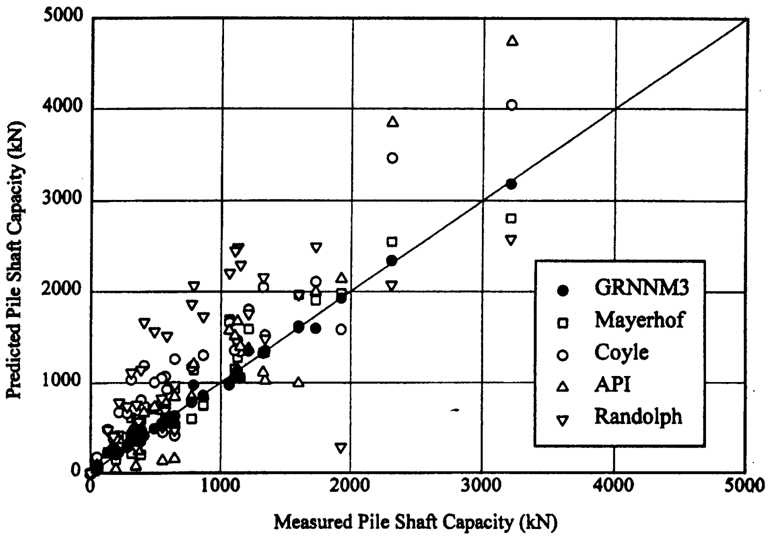


Figure 7 Comparison of predicted and measured shaft pile capacity (Abu-Kiefa 1998)

Teh et al. (1997) proposed a neural network for estimating the static pile capacity determined from dynamic stress-wave data for precast reinforced concrete piles with a square section. The networks were trained to associate the input stress-wave data with capacities derived from the commercial computer code CAPWAP (Rausche et al. 1972). The study was concerned with predicting the 'CAPWAP predicted capacity' rather than the true bearing capacity of the pile. The neural network learned the training data set almost perfectly for predicting the static total pile capacity with a root mean square error less than 0.0003. The trained neural network was assessed for its ability to generalise with a set of testing data. Good prediction was obtained for seven piles out of ten as shown in Figure 8.

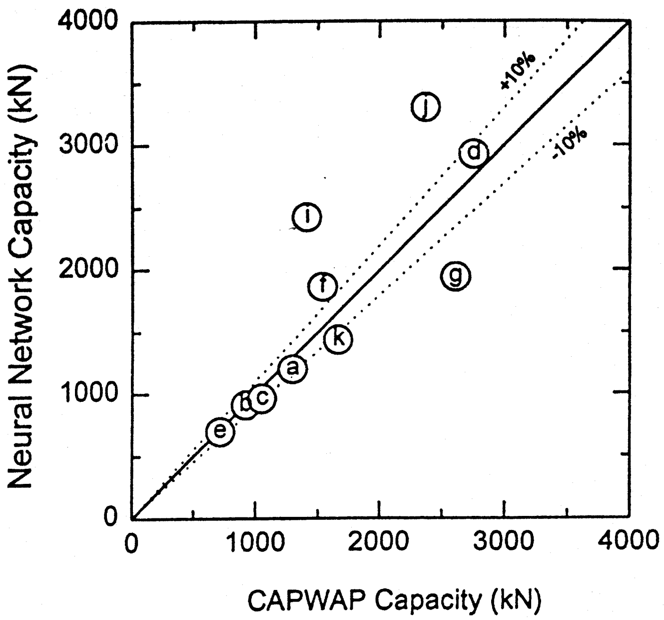


Figure 8 Static capacity predicted by CAPWAP and neural network for testing set (Teh et al. 1997)

Another application includes the prediction of axial and lateral load capacity of steel H-piles, steel piles and pre-stressed and reinforced concrete piles by Nawari et al. (1999).

3.2 SETTLEMENT OF FOUNDATIONS

The design of foundations is generally controlled by the criteria of bearing capacity and settlement; the latter often governing. The problem of estimating the settlement of foundations is very complex, uncertain and not yet entirely understood. This fact encouraged some researchers to apply the ANN technique to settlement prediction. Goh (1994a) developed a neural network for the prediction of settlement of a vertically loaded pile foundation in a homogeneous soil stratum. The input variables for the proposed neural network consisted of the ratio of the elastic modulus of the pile to the shear modulus of the soil, pile length, pile load, shear modulus of the soil, Poisson's ratio of the soil and radius of the pile. The output variable was the pile settlement. The desired output that was used for the ANN model training was obtained by means of finite element and integral equation analyses developed by Randolph and Wroth (1978). A comparison of the theoretical and predicted settlements for the training and testing sets is given in Figure 9. The results in Figure 9 show that the neural network was able to successfully model the settlement of pile foundations.

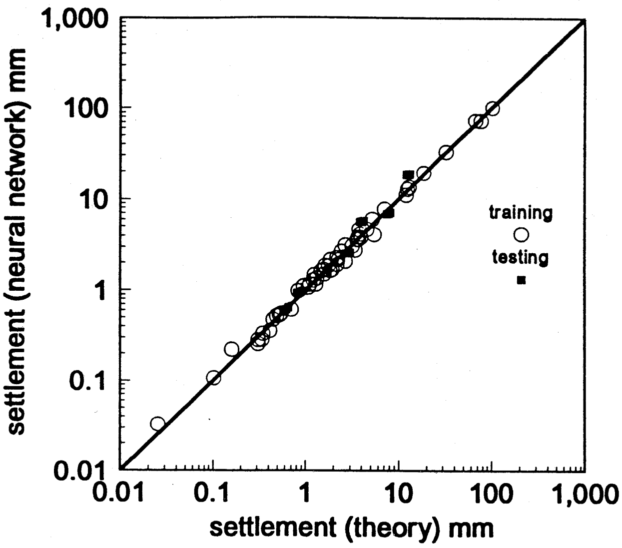


Figure 9 Comparison of theoretical settlements and neural network predictions (Goh 1994)

Sivakugan et al. (1998) explored the possibility of using neural networks to predict the settlement of shallow foundations on granular soils. A neural network was trained with five inputs representing the net applied pressure, average blow count from the standard penetration test, width of foundation, shape of foundation and depth of foundation. The output was the settlement of the foundation. The results obtained by the neural network were compared with methods proposed by Terzaghi and Peck (1967) and Schmertmann (1970). Based on the results obtained, it was shown that the traditional method of Terzaghi and Peck and Schmertmann's method overestimate the settlements by about 2.18 times and 3.39 times respectively as shown in Figure 10. In contrast, the predictions using the ANN model were very good (Figure 11).

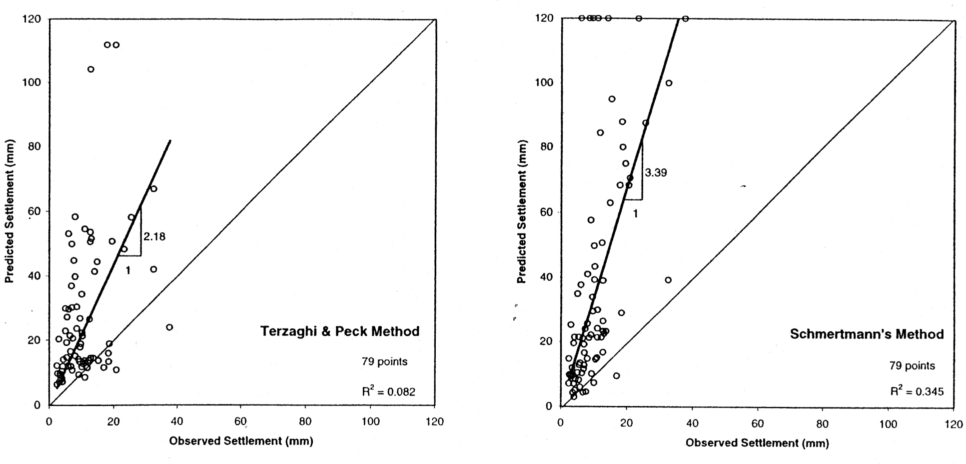


Figure 10 Settlements predicted using traditional methods (Sivakugan et al. 1998)

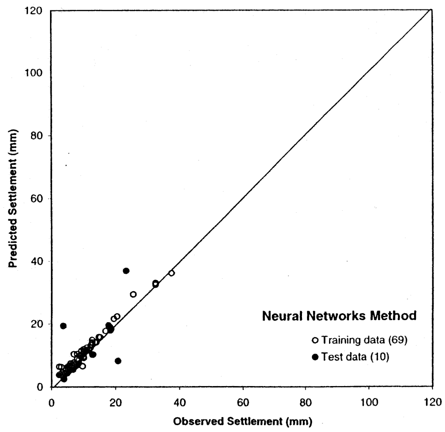


Figure 11 Settlement prediction using artificial neural network (Sivakugan et al. 1998)

Most recently, Shahin et al. (2000) carried out similar work for predicting the settlement of shallow foundations on cohesionless soils. In this work, 272 data records were used for modelling. The input variables considered to have the most significant impact on settlement prediction were the footing width, the footing length, the applied pressure of the footing and the soil compressibility. The results of the ANN were compared with three of the most commonly used traditional methods. These methods were Meyerhof (1965), Schultze and Sherif (1973) and Schmertmann et al. (1978). The results of the study confirmed those found by Sivakugan et al. (1998), in the sense that ANNs were able to predict the settlement well and outperform the traditional methods. As shown in Table 3, the ANN produced high coefficients of correlation, *r*, low root mean squared errors, RMSE, and low mean absolute errors, MAE, compared with the other methods.

Category	ANN	Meyerhof (1965)	Schultze & Sherif (1973)	Schmertmann et al. (1978)
Correlation, *r*	0.99	0.33	0.86	0.70
RMSE (mm)	3.9	27.0	23.8	45.2
MAE (mm)	2.6	20.8	11.1	29.5

Table 3 Comparison of predicted vs measured settlements (Shahin et al. 2000)

3.3 SOIL PROPERTIES AND BEHAVIOUR

Soil properties and behaviour is an area that has attracted many researchers to modelling using ANNs. Developing engineering correlations between various soil parameters is an issue discussed by Goh (1995a; 1995c). Goh used neural networks to model the correlation between the relative density and the cone resistance from cone penetration test (CPT), for both normally consolidated and over-consolidated sands. Laboratory data, based on calibration chamber tests, were used to successfully train and test the neural network model. The neural network model used the relative density and the mean effective stress of soils as inputs and the CPT cone resistance as a single output. The ANN model was found to give high coefficients of correlation of 0.97 and 0.91 for the training and testing data, respectively, which indicated that the neural network was successful in modelling the non-linear relationship between the CPT cone resistance and the other parameters. Many other studies have successfully used ANNs for modelling soil properties and behaviour, which, for brevity, are acknowledged for reference purposes in the following paragraphs.

Ellis et al. (1995) developed an ANN model for sands based on grain size distribution and stress history. Sidarta and Ghaboussi (1998) employed an ANN model within a finite element analysis to extract the geometerial constitutive behaviour from non-uniform material tests. Penumadu and Jean-Lou (1997) used neural networks for representing the behaviour of sand and clay soils. Ghaboussi and Sidarta (1998) used neural networks to model both the drained and undrained behaviour of sandy soil subjected to triaxial compression-type testing. Penumadu and Zhao (1999) also used ANNs to model the stress-strain and volume change behaviour of sand and gravel under drained triaxial compression test conditions. Zhu et al. (1998a; 1998b) used neural networks for modelling the shearing behaviour of a fine-grained residual soil, dune sand and Hawaiian volcanic soil. Cal (1995) used a neural network model to generate a quantitative soil classification from three main factors (plastic index, liquid limit and clay content). Najjar et al. (1996a) showed that neural network-based models can be used to accurately assess soil swelling, and that neural network models can provide significant improvements in prediction accuracy over statistical models. Romero and Pamukcu (1996) showed that neural networks are able to effectively characterise and estimate the shear modulus of granular materials. Agrawal et al. (1994); Gribb and Gribb (1994) and Najjar and Basheer (1996b) all used neural network approaches for estimating the permeability of clay liners. Basheer and Najjar (1995) and Najjar et al. (1996b) presented neural network approaches for soil compaction.

Other applications include modelling the mechanical behaviour of medium-to-fine sand (Ellis et al. 1992), modelling rate-dependent behaviour of clay soils (Penumadu et al. (1994), simulating the uniaxial stress-strain constitutive behaviour of fine-grained soils under both monotonic and cyclic loading (Basheer 1998; Basheer and Najjar 1998), characterising the undrained stress-strain response of Nevada sand subjected to both triaxial compression and extension stress paths (Najjar and Ali 1999; Najjar et al. 1999), predicitng the axial and volumetric stress-strain behaviour of sand during loading, unloading and reloading (Zhu and Zaman 1997), predicting the anisotropic stiffness of granular materials from standard repeated load triaxial tests (Tutumluer and Seyhan 1998).

3.4 LIQUEFACTION

Liquefaction is a phenomenon which occurs mainly in loose and saturated sands as a result of earthquakes. It causes the soil to lose its shear strength due to an increase in pore water pressure, often resulting in large amounts of damage to most civil engineering structures. Determination of liquefaction potential due to earthquakes is a complex geotechnical engineering problem. Goh (1994b) used neural networks to model the complex relationship between seismic and soil parameters in order to investigate liquefaction potential. The neural network used in this work was trained using case records from 13 earthquakes that occurred in Japan, United States and Pan-America during the period 1891–1980. The study used eight input variables and only one output variable. The input variables were the SPT-value, the fines content, the mean grain size, the total stress, the effective stress, the equivalent dynamic shear stress, the earthquake magnitude and the maximum horizontal acceleration at ground surface. The output was assigned a binary value of 1, for sites with extensive or moderate liquefaction, and a value of 0 for marginal or no liquefaction. The results obtained by the neural network model were compared with the method of Seed et al. (1985). The study showed that the neural network gave correct predictions in 95% of cases, whereas Seed et al. (1985) gave a success rate of 84%. Goh (1996a) also used neural networks to assess liquefaction potential from cone penetration test (CPT) resistance data. The data records were taken for sites of sand and silty sand deposits in Japan, China, United States and Romania, representing five earthquakes that occurred during the period 1964–1983. A similar neural network modelling strategy, as used in Goh (1994b), was used for this study and the results were compared with the method of Shibata and Teparaksa (1988). The neural network showed a 94% success rate, which is equivalent to the same number of error predictions as the conventional method by Shibata and Teparaksa (1988).

Two other works (Najjar and Ali 1998; Ural and Saka 1998) also used CPT data to evaluate soil liquefaction potential and resistance. Najjar and Ali (1998) used neural networks to characterise the soil liquefaction resistance utilising field data sets representing various earthquake sites from around the world. The ANN model that was developed in this work was generated to produce a liquefaction potential assessment chart that could be used by geotechnical engineers in liquefaction assessment tasks. Ural and Saka (1998) used neural networks to analyse liquefaction. Comparison between this approach and a simplified liquefaction procedure indicated a similar rate of success for the neural network approach as the conventional approach.

Other applications of ANNs for liquefaction prediction include the prediction of liquefaction resistance and potential (Juang and Chen 1999), investigation of the accuracy of liquefaction prediction of ANNs compared with fuzzy logic and statistical approaches (Ali and Najjar 1998) and assessment of liquefaction potential using standard penetration test results (Agrawal et al. 1997).

3.5 SITE CHARACTERISATION

Site characterisation is an area concerned with the analysis and interpretation of geotechnical site investigation data. Zhou and Wu (1994) used a neural network model to characterise the spatial distribution of rockhead elevations. The data used to train the model were taken from seismic refraction surveys on more than 11 km of transverse lines. The network used the spatial position (x- and y-coordinate) and the surface elevation as inputs, and was used to estimate the rockhead elevation at that location as the output. The trained network was tested to estimate the rockhead elevations for all locations within the area of investigation by producing a contour map. Results from the neural network model compared well with similar contour maps generated using kriging, with the additional benefit that neural networks do not make assumptions or simplify spatial variations.

A similar application relevant to ground water characterisation was described by Basheer et al. (1996). Basheer et al. (1996) indicated that neural networks can be used to map and logically predict the variation of soil permeability in order to identify landfill boundaries and construct a waste landfill. Rizzo et al. (1996) presented a new site characterisation method called SCANN (Site Characterisation using Artificial Neural Networks) that is based on the use of neural networks to map discrete spatially-distributed fields. Other applications were presented by Najjar and Basheer (1996a) and Rizzo and Dougherty (1994).

3.6 EARTH RETAINING STRUCTURES

Goh et al. (1995) developed a neural network model to provide initial estimates of maximum wall deflections for braced excavations in soft clay. The neural network was used to synthesise data derived from finite element studies on braced excavations in clay. The input parameters used in the model were the excavation width, soil thickness/excavation width ratio, wall stiffness, height of excavation, soil undrained shear strength, undrained soil modulus/shear strength ratio and soil unit weight. The maximum wall deflection was the only output. Using regression analysis, the scatter of the predicted neural network deflections relative to the deflections obtained using the finite element method were assessed. The results produced high coefficients of correlation for the training and testing data of 0.984 and 0.967, respectively. Some additional testing data from actual case records were also used to confirm the performance of the trained neural network model. The agreement of the neural network predicted and measured wall deflections was encouraging, as shown in Table 4. The study intended to use the neural network model as a time-saving and user-friendly alternative to the finite element method.

Case history	Measured wall deflection (mm)	Predicted wall deflection (mm)
Laveder (Singapore)	32	31
Laveder (Singapore)	36	28
Telecom (Singapore)	56-84	65
Vaterland 3	76	76
(NGI 1962)	114-140	107
San Francisco	20-60	59
(Mana 1977)	72-150	122

Table 4 Comparison of neural network predictions and field measurements (Goh 1995)

3.7 SLOPE STABILITY

Ni et al. (1996) proposed a methodology of combining fuzzy sets theory with artificial neural networks for evaluating the stability of slopes. In this approach, the input parameters were gradient, horizontal profile, vertical profile, location, height, geological origin, soil texture, depth of weathering, direction of slopes, vegetation, land use, maximum daily precipitation and maximum hour precipitation. The output was the slope failure potential. A number of hypothetical natural slopes were evaluated by both neural networks and an analytical model, and the results of the neural network approach were in a good agreement when compared with those obtained by the analytical model.

3.8 TUNNELS AND UNDERGROUND OPENINGS

Shi et al. (1998) presented a study of neural networks for predicting settlements of tunnels. A general neural network model was trained and tested using data from the 6.5 km Brasilia Tunnel, Brazil. The study identified many factors to be used as the model inputs and three settlement parameters as the model outputs. The input parameters were the length of excavation from drive start, the depth of soil cover above tunnel crown, the area of tunnel section, the delay for closing invert, the water level depth, the rate of advance of excavation, the construction method, the mean blow count from standard penetration test at tunnel crown level, the mean blow count from standard penetration test at tunnel spring-line level and the mean blow count from standard penetration test at tunnel inverted arch level. The three output parameters were the settlement at the face passage, the settlement at the invert closing and the final settlement after stabilisation. The results showed that the neural network model could not achieve a high level of accuracy. To improve the prediction accuracy, the study proposed a modular neural network model based on the concept of integrating multiple neural network modules in one system, with each module being constrained to operate at one specific situation of a complicated real world problem. The modular concept showed an improvement in terms of model convergence and prediction. The capability to improve the models developed in this work was later extended by Shi (2000) by applying input data transformation. This extended study indicated that distribution transformation of the input variables reduced the prediction error by more than 13%.

Lee and Sterling (1992) developed a neural network for identification of probable failure modes for underground openings from prior case history information. The study used the knowledge obtained by the neural network to generate a design tool of a tunnel. Sterling and Lee (1992) used the neural network as part of a knowledge-based expert system for assisting with tunnel design. Moon et al. (1995) also used ANNs, integrated with an expert system for the preliminary design of tunnels.

4 DISCUSSION AND CONCLUSIONS

It is evident from this review that ANNs have been applied successfully to many geotechnical engineering areas. This includes pile capacity prediction, settlement of foundations, soil properties and behaviour, liquefaction, site characterisation, earth retaining structures, slope stability and the design of tunnels and underground openings. Perhaps the most successful and well-established applications are the capacity prediction of driven piles, liquefaction and the prediction of soil properties and behaviour. Other applications (e.g. settlement of structures) need to be treated with caution until additional research has been conducted. There are also several areas in which the feasibility of ANNs has yet to be tested, such as bearing capacity prediction of shallow foundations, capacity of bored piles, design of sheet pile walls and dewatering, among others.

Based on the results of the studies reviewed in this paper, it is evident that ANNs perform better than, or as well as, the conventional methods used as a basis for comparison in many situations, whereas, they fail to perform well in a few.
In many situations in geotechnical engineering, it is possible to encounter some types of problems that are very complex and not well understood. For most mathematical models that attempt to solve such problems, the lack of physical understanding is usually supplemented by either simplifying the problem or incorporating several assumptions into the models. Mathematical models also rely on assuming the structure of the model in advance, which may be sub-optimal. Consequently, many mathematical models fail to simulate the complex behaviour of most geotechnical engineering problems. In contrast, ANNs are based on the data alone in which the model can be trained on input-output data pairs to determine the structure and parameters of the model. In this case, there is no need to simplify the problem nor incorporate any assumptions. Moreover, ANNs can always be updated to obtain better results by presenting new training examples as new data become available.

Despite their good performance in many situations, ANNs suffer from a number of shortcomings, notably, the lack of theory to help with their development, the fact that success in finding a good solution is not always guaranteed and their limited ability to explain the way they use the available information to arrive a solution. Consequently, there is a need to develop some guidelines, which can help in the design process of ANNs. There is also a need for more research to give a comprehensive explanation of how ANNs arrive at a prediction.

A further issue that needs to be given some attention in the future development of ANNs is to include treatment of uncertainties associated with geotechnical engineering parameters. Despite these uncertainties, ANN models that have been developed to date in the field of geotechnical engineering are essentially deterministic. Consequently, procedures that incorporate such uncertainties into ANN models are essential, as they provide more realistic solutions. Overall, despite the limitations of ANNs, they have a number of significant benefits that make them a powerful and practical tool for solving many problems in the field of geotechnical engineering.

5 REFERENCES

Abu-Kiefa, M. A. (1998). “General regression neural networks for driven piles in cohesionless soils.” *J. Geotech. & Geoenv. Engrg.*, ASCE, 124(12), 1177-1185.

Agrawal, G., Chameau, J. A., and Bourdeau, P. L. (1997). “Assessing the liquefaction susceptibility at a site based on information from penetration testing.” In: *Artificial neural networks for civil engineers: fundamentals and applications,* N. Kartam, I. Flood, and J. H. Garrett, eds., New York, 185-214.

Agrawal, G., Weeraratne, S., and Khilnani, K. (1994). “Estimating clay liner and cover permeability using computational neural networks.” *Proc., First Congress on Computing in Civil Engrg.*, Washington, June 20-22.

Ali, H. E., and Najjar, Y. M. (1998). “Neuronet-based approach for assessing liquefaction potential of soils.” *Transportation Research Record No. 1633*, 3-8.

American Petroleum Institute. (1984). *RP2A: Recommended practice for planning, designing and constructing fixed offshore platforms*, Washington, D.C.

Basheer, I. A. (1998). “Neuromechanistic-based modeling and simulation of constitutive behavior of fine-grained soils.” Ph.D. dissertation, Kansas State University, Manhattan, KS.

Basheer, I. A., and Najjar, Y. M. (1995). “A neural-network for soil compaction.” *Proc., 5th Int. Symp. Numerical Models in Geomechanics*, G. N. Pande and S. Pietruszczak, eds., Roterdam: Balkema, 435-440.

Basheer, I. A.,and Najjar, Y. M. (1998). “Modeling cyclic constitutive behavior by neural networks: Theoritical and real data.” *Proc., 12th Engineering Mechanics Conf.*, H. Murakami and J. E. Luco, eds, La Jolla, California, May 17-20, 952-955.

Basheer, I. A., Reddi, L. N., and Najjar, Y. M. (1996). “Site characterisation by neuronets: An application to the landfill siting problem.” *Ground Water*, 34, 610-617.

Broms, B. B., and Lim, P. C. (1988). “A simple pile driving formula based on stress-wave measurements.” *Proc., The 3rd Int. Conf. on the Application of Stress-Wave Theory to Piles*, B. H. Fellenius, ed., Vancouver, 591-600.

Burland, J. B. (1973). “Shaft friction of piles in clay.” *Ground Engineering*, 6(3), 1-15.

Cal, Y. (1995). “Soil classification by neural-network.” *Advances in Engineering Software*, 22(2), 95-97.

Chan, W. T., Chow, y. K., and Liu, L. F. (1995). “Neural network: An alternative to pile driving formulas.” *J. Computers and Geotechnics*, 17, 135-156.

Coyle, H. M., and Castello, R. R. (1981). “New design correlations for piles in sand.” *J. Geotech. Engrg.*, ASCE, 107(7), 965-986.

Ellis, G. W., Yao, C., and Zhao, R. (1992). “Neural network modeling of the mechanical behavior of sand.” *Proc., Engineering Mechanics*, ASCE, 421-424.

Ellis, G. W., Yao, C., Zhao, R., and Penumadu, D. (1995). “Stress-strain modelling of sands using artificial neural networks.” *J. Geotech. Engrg.*, ASCE, 121(5), 429-435.

Fausett, L. V. (1994). *Fundamentals neural networks: Architecture, algorithms, and applications*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Flood, I., and Kartam, N. (1994). “Neural networks in civil engineering I: Principles and understanding.” *J. Computing in Civil Engrg*, ASCE, 8(2), 131-148.

Gardner, M. W., and Dorling, S. R. (1998). “Artificial neural networks (The multilayer perceptron) - A review of applications in the atmospheric sciences.” *Atmospheric Environment*, 32(14/15), 2627-2636.

Ghaboussi, J., and Sidarta, D. E. (1998). “New nested adaptive neural networks (NANN) for constitutive modeling.” *J. Computers and Geotechnics*, 22(1), 29-52.

Goble, G. G., Likins, G. E., and Rausche, F. (1975). “Bearing capacity of piles from dynamic measurements.” *Final Report*, Dept. of Civil Engineering, Case Western University.

Goh, A. T. C. (1994a). “Nonlinear modelling in geotechnical engineering using neural networks.” *Australian Civil Engineering Transactions*, CE36(4), 293-297.

Goh, A. T. C. (1994b). “Seismic liquefaction potential assessed by neural network.” *J. Geotech. & Geoenv. Engrg.*, ASCE, 120(9), 1467-1480.

Goh, A. T. C. (1995a). “Back-propagation neural networks for modeling complex systems.” *Artificial Intelligence in Engineering*, 9, 143-151.

Goh, A. T. C. (1995b). “Empirical design in geotechnics using neural networks.” *Geotechnique*, 45(4), 709-714.

Goh, A. T. C. (1995c). “Modeling soil correlations using neural networks.” *J. Computing in Civil Engrg.*, ASCE, 9(4), 275-278.

Goh, A. T. C. (1996a). “Neural-network modeling of CPT seismic liquefaction data.” *J. Geotech. Engrg.*, ASCE, 122(1), 70-73.

Goh, A. T. C. (1996b). “Pile driving records reanalyzed using neural networks.” *J. Geotech. Engrg.*, ASCE, 122(6), 492-495.

Goh, A. T. C., Wong, K. S., and Broms, B. B. (1995). “Estimation of lateral wall movements in braced excavation using neural networks.” *Canadian Geotech. J.*, 32, 1059-1064.

Gribb, M. M., and Gribb, G. W. (1994). “Use of neural networks for hydraulic conductivity determination in unsaturated soil.” *Proc., 2nd Int. Conf. Ground Water Ecology*, J. A. Stanford and H. M. Valett, eds., Bethesda MD: Amer, Water Resources Assoc., 155-163.

Hecht-Nielsen, R. (1990). *Neurocomputing*, Addison-Wesely Publishing Company.

Hiley, A. (1922). “The efficiency of the hammer blow, and its effects with reference to piling.” *Engineering*, 2, 673.

Hubick, K. T. (1992). *Artificial neural networks in Australia*, Department of Industry, Technology and Commerce, Commonwealth of Australia, Canberra.

Jaksa, M. B. (1995). “The influence of spatial variability on the geotechncial design properties of a stiff, overconsolidated clay,” PhD thesis, The University of Adelaide, Adelaide.

Janbu, N. (1953). “Une analyse energetique du battage des pieux a l'aide de parametres sans dimension.” *Norwegian Geotech. Inst.*, Oslo, Norway, 3, 63-64.

Juang, C. H., and Chen, C. J. (1999). "CPT-based liquefaction evaluation using artificial neural networks." *Computer-Aided Civil and Infrastructure Engineering*, 14(3), 221-229.
Lee, C., and Sterling, R. (1992). "Identifying probable failure modes for underground openings using a neural network." *Int. J. Rock Mechanics and Mining Science & Geomechanics Abstracts*, 29(1), 49-67.
Lee, I. M., and Lee, J. H. (1996). "Prediction of pile bearing capacity using artificial neural networks." *Computers and Geotechnics*, 18(3), 189-200.
Maier, H. R., and Dandy, G. C. (2000). "Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications." *Environmental Modelling & Software*, 15(2000), 101-124.
Mana, I. A. (1997). "Finite element analyses of deep excavation behavior in soft clay." PhD thesis, Stanford University, Stanford, Calif.
Maren, A., Harston, C., and Pap, R. (1990). *Handbook of neural computing applications*, Academic Press, Inc., San Diego, California.
Meyerhof, G. G. (1965). "Shallow foundations." *J. Soil Mech. & Found. Div.*, ASCE, 91(SM2), 21-31.
Meyrehof, G. G. (1976). "Bearing capacity and settlement of pile foundations." *J. Geotech. Engrg.*, ASCE, 102(3), 196-228.
Moon, H. K., Na, S. M., and Lee, C. W. (1995). "Artificial neural-network integrated with expert-system for preliminary design of tunnels and slopes." *Proc., 8th Int. Congress on Rock Mechanics*, T. Fujii, ed., Rotterdam: Balkema, 1 & 2, 901-905.
Moselhi, O., Hegazy, T., and Fazio, P. (1992). "Potential applications of neural networks in construction." *Can. J. Civil Engrg*, 19, 521-529.
Najjar, Y. M., and Ali, H. E. (1998). "CPT-based liquefaction potential assessment: A neuronet approach." *Geotechnical Special Publication*, ASCE, 1, 542-553.
Najjar, Y. M., and Ali, H. E. (1999). "Simulating the stress-strain behaviour of Nevada sand by ANN." *Proc., 5th U.S.National Congress on Computational Mechanics (USACM)*, Boulder, Colorado, August 4-6.
Najjar, Y. M., Ali, H. E., and Basheer, I. A. (1999). "On the use of neurons for simulating the stress-strain behavior of soils." *Proc., 7th Int. Symposium on Numerical Models in Geomechanics*, G. N. Pande, ed., Graz, Austria, NUMOG VII, September 1-3, 657-662.
Najjar, Y. M., and Basheer, I. A. (1996a). "Neural network approach for site characterization and uncertainty prediction." *Geotechnical Special Publication*, ASCE, 58(1), 134-148.
Najjar, Y. M., and Basheer, I. A. (1996b). "Utilizing computational neural networks for evaluating the permeability of compacted clay liners." *Geotechnical and Geological Engineering*, 14, 193-221.
Najjar, Y. M., Basheer, I. A., and McReynolds, R. (1996a). "Neural modeling of Kansan soil swelling." *Transportation Research Record No. 1526*, 14-19.
Najjar, Y. M., Basheer, I. A., and Naouss, W. A. (1996b). "On the identification of compaction characteristics by neuronets." *J. Computers and Geotechnics*, 18(3), 167-187.
Nawari, N. O., Liang, R., and Nusairat, J. (1999). "Artificial intelligence techniques for the design and analysis of deep foundations." *Electronic J. Geotech. Engrg*., http://geotech.civeng.okstate.edu/ejge/ppr9909/index.html.
NGI. (1962). "Measurements at a strutted excavation, Oslo subway, Vaterland, kn 1450." *Technical Report 8*, Norwegian Geotechncial Institute, Oslo, Norway.
Ni, S. H., Lu, P. C., and Juang, C. H. (1996). "A fuzzy neural network approach to evaluation of slope failure potential." *J. Microcomputers in Civil Engineering*, 11, 59-66.
Penumadu, D., and Jean-Lou, C. (1997). "Geomaterial modeling using artificial neural networks." *Artificial Neural Networks for Civil Engineers: Fundamentals and Applications*, ASCE, 160-184.
Penumadu, D., Jin-Nan, L., Chameau, J.-L., and Arumugam, S. (1994). "Rate dependent behavior of clays using neural networks." *Proc., 13th Conf. Int. soc. Soil Mech. & Found. Engrg.*, New Delhi, 4, 1445-1448.
Penumadu, D., and Zhao, R. (1999). "Triaxial compression behavior of sand and gravel using artificial neural networks (ANN)." *J. Computers and Geotechnics*, 24, 207-230.
Randolph, M. F. (1985). "Capacity of piles driven into dense sand." *Rep. Soils TR 171*, Engrg. Dept., Cambridge University, Cambridge, U. K.
Randolph, M. F., and Wroth, C. P. (1978). "Analysis of deformation of vertically loaded piles." *J. Geotech. Engrg.*, ASCE, 104(12), 1465-1488.
Rausche, F., Moses, F., and Goble, G. G. (1972). "Soil resistance predictions from pile dynamics." *J. Soil Mech. and Found. Div.*, ASCE, 98, 917-937.
Ripley, B. D. (1996). *Pattern recognition and neural networks*, Cambridge University Press.
Rizzo, D. M., and Dougherty, D. E. (1994). "Application of artificial neural networks for site characterization using hard and soft information." *Proc., 10th Int. Conf. Computational Methods in Water Resources*, A. Peters et al., eds., Dordrecht: Kluwer Academic, 12, 793-799.

Rizzo, D. M., Lillys, T. P., and Dougherty, D. E. (1996). "Comparisons of site characterization methods using mixed data." *Geotechnical Special Publication*, ASCE, 58(1), 157-179.

Romero, S., and Pamukcu, S. (1996). "Characterization of granular meterial by low strain dynamic excitation and ANN." *Geotechnical Special Publication*, ASTM-ASCE, 58(2), 1134-1148.

Schmertmann, J. H. (1970). "Static cone to compute static settlement over sand." *J. Soil Mech. & Found. Div.*, ASCE, 96(SM3), 7302-1043.

Schmertmann, J. H., Hartman, J. P., and Brown, P. B. (1978). "Improved strain influence factor diagrams." *J. Geotech. Engrg.*, ASCE, 104(GT8), 1131-1135.

Schultze, E., and Sherif, G. (1973). "Prediction of settlements from evaluated settlement observations for sand." *Proc., 8th Int. Conf. Soil Mech. & Found. Engrg.*, 1(3), 225-230.

Seed, H. B., Tokimatsu, H., Harder, L. F., and Chung, R. M. (1985). "Influence of SPT procedure in seismic liquefaction resistance evaluations." *J. Geotech. Engrg.*, ASCE, 111(12), 1425-1445.

Semple, R. M., and Rigden, W. J. (1986). "Shaft capacity of driven pipe piles in clay." *Ground Engineering*, 19(1), 11-17.

Shahin, M. A., Jaksa, M. B., and Maier, H. R. (2000). "Predicting the settlement of shallow foundations on cohesionless soils using back-propagation neural networks." *Research Report No. R 167*, The University of Adelaide, Adelaide.

Shi, J., Ortigao, J. A. R., and Bai, J. (1998). "Modular neural networks for predicting settlement during tunneling." *J. Geotech. & Geoenv. Engrg.*, ASCE, 124(5), 389-395.

Shi, J. J. (2000). "Reducing prediciton error by transforming input data for neural networks." *J. Computing in Civil Engrg.*, ASCE, 14(2), 109-116.

Shibata, T., and Teparaksa, W. (1988). "Evaluation of liquefaction potentials of soils using cone penetration tests." *Soils and Foundations*, 28(2), 49-60.

Sidarta, D. E., and Ghaboussi, J. (1998). "Constitutive modeling of geomaterials from non-uniform material tests." *J. Computers & Geomechanics*, 22(10), 53-71.

Sivakugan, N., Eckersley, J. D., and Li, H. (1998). "Settlement predictions using neural networks." *Australian Civil Engineering Transactions*, CE40, 49-52.

Sterling, R. L., and Lee, C. A. (1992). "A neural network - expert system hybird approach for tunnel design." *Proc., 33rd United-States Symp. Rock Mechanics*, J. R. Tillerson and W. R. Wawerisk, eds., Brookfield VT: Balkema, 501-510.

Teh, C. I., Wong, K. S., Goh, A. T. C., and Jaritngam, S. (1997). "Prediction of pile capacity using neural networks." *J. Computing in Civil Engineering*, ASCE, 11(2), 129-138.

Terzaghi, K., and Peck, R. B. (1967). *Soil mechanics in engineering practice*, John Wiley & Son, Inc., New York.

Tutumluer, E., and Seyhan, U. (1998). "Neural network modeling of anisotropic aggregate behavior from repeated load triaxial tests. " *Transportation Research Record 1615*, National Research Council, Washington, D.C.

Ural, D. N., and Saka, H. (1998). "Liquefaction assessment by neural networks." *Electronic Journal of Geotechnical Engrg.*, http://geotech.civen.okstate.edu/ejge/ppr9803/index.html.

Wellington, A. M. (1892). "The iron wharf at Fort Monroe, Va." *Transactions*, ASCE, 27, 129-137.

Zhou, Y., and Wu, X. (1994). "Use of neural networks in the analysis and interpretation of site investigation data." *Computer and Geotechnics*, 16, 105-122.

Zhu, J. H., and Zamman, M. M. (1997). "Neural network modeling for a cohesionless soil." *76th Meeting of the Transportation Research Board*, January, Washington, D.C.,

Zhu, J. H., Zaman, M. M., and Anderson, S. A. (1998a). "Modeling of soil behavior with a recurrent neural network." *Canadian Geotech. J.*, 35(5), 858-872.

Zhu, J. H., Zaman, M. M., and Anderson, S. A. (1998b). "Modelling of shearing behaviour of a residual soil with recurrent neural network." *Int. J. Numerical and Analytical Methods in Geomechanics*, 22(8), 671-687.

Zurada, J. M. (1992). *Introduction to artificial neural systems*, West Publishing Company, St. Paul.