

Intelligent Computing for Predicting Axial Capacity of Drilled Shafts

Mohamed A. Shahin¹ and Mark B. Jaksa²

¹Lecturer, Department of Civil Engineering, Curtin University of Technology, WA 6845, Australia; Phone: +61-8-9266 1822; Fax: +61-8-9266 2681; e-mail: m.shahin@curtin.edu.au

²Associate Professor, School of Civil, Environmental and Mining Engineering, University of Adelaide, SA 5005, Australia; Phone: +61-8-8303 4317; Fax: +61-8-8303 4359; e-mail: mark.jaksa@adelaide.edu.au

Abstract

In the last few decades, numerous methods have been developed for predicting the axial capacity of drilled shafts. Among the available methods, the cone penetration test (CPT) based models have been shown to give better predictions in many situations. This can be attributed to the fact that CPT-based methods have been developed in accordance with the results of the CPT tests, which have been found to yield more reliable soil properties, hence, more accurate axial capacity predictions of drilled shafts. In this paper, one of the most commonly used artificial intelligence techniques, i.e. artificial neural networks (ANNs), was utilized in an attempt to obtain more accurate axial capacity predictions for drilled shafts. The ANN model was developed using data collected from the literature that comprise CPT results and drilled shaft load tests of 94 case records. The predictions from the ANN model were compared with those obtained from three commonly used available CPT-based methods. The results indicate that the ANN-based model provides more accurate axial capacity predictions of drilled shafts and outperforms the available conventional methods.

Keywords: artificial neural networks, drilled shafts, modeling, axial capacity.

Introduction

The geotechnical literature has included many methods, both theoretical and experimental, to predict the ultimate capacity of drilled shafts. Due to the difficulty of obtaining undisturbed samples of soils, several methods have focused on correlations with in-situ tests, such as the cone penetration test (CPT), standard penetration test (SPT), dilatometer test and pressuremeter test. However, most available methods, by necessity, simplify the problem by incorporating several assumptions associated with the factors that affect the capacity of drilled shafts and thus these methods fail to achieve consistent success in relation to accurate capacity prediction of drilled shafts. In this respect, artificial neural networks (ANNs), which do not need incorporation of any assumptions or simplifications, are more efficient.

In recent years, ANNs have been found to solve many problems in the field of geotechnical engineering, and the authors have utilized successfully ANNs in different geotechnical engineering applications (e.g. Shahin and Jaksa 2006; Shahin et al. 2002b). ANNs have been also used by other researchers to predict the ultimate capacity of driven piles (e.g. Abu-Kiefa 1998; Lee and Lee 1996), however, their models were developed using a limited number of data cases and none of the models was based on the more accurate measures of soil properties from the CPT. More recently, Shahin (2008) has applied successfully ANNs with sufficient pile load tests and CPT data for prediction of ultimate capacity of driven piles. In this paper, the feasibility of using ANNs for developing a more accurate CPT-based prediction model for the ultimate capacity of drilled shafts is investigated. The predictive ability of the ANN model is examined by comparing its results with experimental data, and with those obtained from three commonly used drilled shaft capacity prediction methods. The model is translated into a simple design equation for routine use in practice.

Brief Overview of Artificial Neural Networks

The type of neural networks used in this study are multi-layer feed-forward that are trained with the back-propagation algorithm (Rumelhart et al. 1986). A comprehensive description of this type of neural networks is beyond the scope of this paper and can be found in many publications (e.g. Fausett 1994). The typical structure of a multi-layer feed-forward neural network consists of a number of processing elements (also called nodes or neurons) that are fully or partially linked via connection weights. These processing elements are usually arranged in layers: an input layer; an output layer; and one or more layers in between, called hidden layers. At each processing element, the weighted inputs are summed and a bias is added or subtracted. This combined input is then passed through a non-linear transfer function (e.g. *sigmoidal* function or *tanh* function) to produce the output of the processing element. Training of a multi-layer feed-forward neural network commences at the input layer, where the network is presented with an actual measured set of data (i.e. the training set) and the output of the network is obtained by utilizing a learning rule. The network output is compared with the desired output from which an error is calculated. This error is then used to adjust the connection weights so that the best input/output mapping is obtained. Once training has been accomplished successfully, the performance of the trained model has to be verified using an independent validation set.

Development of Artificial Neural Network Model

In this work, the ANN model is developed using the software package *NEUFRAME* Version 4.0 (Neosciences 2000). The data used to calibrate and validate the ANN model are obtained from the literature and comprise a series of 94 in-situ drilled shaft load tests reported by Alsamman (1995). The tests were conducted on straight and belled drilled shafts located on sites of different soil types

and geotechnical conditions, ranging from cohesive clays to cohesionless sands. The tests include compression loading (for straight and belled shafts) and tension loading (for straight shafts only). The drilled shafts used have stem diameters ranging from 305 mm to 1798 mm and embedment lengths from 4.5 m to 27.4 m. A database given by Alsamman (1995) comprises details of the data derived from the tests.

The parameters that are considered to be significant in prediction of the ultimate capacity of drilled shafts include the shaft stem diameter, D_{stem} , shaft base diameter, D_{base} , embedment length, L , weighted average cone point resistance over shaft base failure zone, \bar{q}_{c-base} , weighted average cone point resistance over shaft length, $\bar{q}_{c-shaft}$. These parameters are presented to the ANN model as potential model input variables. The ultimate shaft capacity, Q_u , is the single model output variable. In this work, Q_u for shafts under compression is taken as the axial load measured at a displacement equal to 5% of shaft base diameter plus the elastic compression of the shaft (i.e. PL/EA , where: P is the applied load, L is the shaft length, A is the shaft cross-sectional area and E is the shaft elastic modulus). On the other hand, Q_u for shafts under tension is defined as the axial load at 12 mm (0.5") of displacement. The above criteria for determination of ultimate load are as suggested by Alsamman (1995) and recommended by Reese and O'Neill (1988). It should be noted that the shaft base failure zone over which \bar{q}_{c-base} is calculated is taken in accordance with Alsamman (1995) to be equal to one diameter depth beneath the shaft base. It should be also noted that the majority of the tests reported by Alsamman (1995) include mechanical rather than electric CPT data and thus, it was necessary to convert the mechanical CPT readings into equivalent electric CPT values as the electric CPT is the one that is commonly used nowadays. This is carried out using the following correlation proposed by Kulhawy and Mayne (1990):

$$\left(\frac{q_c}{p_a} \right)_{Electric} = 0.47 \left(\frac{q_c}{p_a} \right)_{Mechanical}^{1.19} \quad (1)$$

where: p_a is the atmospheric pressure, and p_a and q_c are in kPa. A comparison between the mechanical and electric CPT is beyond the scope of this paper and can be found in Kulhawy and Mayne (1990).

The next step in the development of the ANN model is dividing the available data into their subsets. As recommended by Masters (1993) and detailed by Shahin et al. (2004), the available data are randomly divided into two statistically consistent sets: a training set for model calibration and an independent validation set for model verification. In total, 76 case records of the available 94 cases are used for training and 18 cases for validation. It should be noted that, like all empirical models, ANNs perform best in interpolation rather than extrapolation (Masters 1993), thus, the extreme values of the available data should be included in the training set. The statistics of the data used for the training and validation sets are given in Table 1, which include the mean, standard deviation, minimum, maximum and range. Once the data have been divided into their subsets, the input and output variables are pre-

processed by scaling them between 0.0 and 1.0 to eliminate their dimension and to ensure that all variables receive equal attention during training.

Table 1. ANN input and output statistics

Model variables and data sets	Statistical parameters				
	Mean	Standard deviation	Minimum	Maximum	Range
Shaft stem diameter, D_{stem} (mm)					
Training set	617.0	371.9	304.8	1798.3	1493.5
Validation set	525.8	245.5	320.0	1100.3	780.3
Shaft base diameter, D_{base} (mm)					
Training set	741.7	417.9	304.8	2100.1	1795.3
Validation set	630.6	277.5	320.0	1149.1	829.1
Shaft embedment length, L (m)					
Training set	10.0	4.7	4.5	27.4	22.9
Validation set	9.0	4.4	5.8	24.2	18.3
Weighted average cone point resistance along base failure zone, \bar{q}_{c-base} (MPa)					
Training set	16.8	10.3	0.0	47.5	47.5
Validation set	18.2	11.7	0.0	39.5	39.5
Weighted average cone point resistance along shaft length, $\bar{q}_{c-shaft}$ (MPa)					
Training set	8.2	5.2	1.1	28.8	27.7
Validation set	10.1	5.3	2.5	21.5	19.0
Ultimate shaft capacity, Q_u (kN)					
Training set	2184.3	2161.5	355.8	9652.2	9296.3
Validation set	2075.0	2222.0	355.8	8824.8	8469.0

The following step in the development of the ANN model is the determination of model geometry (i.e. the number of hidden layers and corresponding number of hidden nodes in each layer). In this work, the optimal model geometry is obtained by utilizing a trial-and-error approach in which the ANN model is trained using one hidden layer with 1, 2, 3, ..., and 11 hidden layer nodes, respectively. It should be noted that a network with one hidden layer can approximate any continuous function provided that sufficient connection weights are used (Hornik et al. 1989). Consequently, one hidden layer is used in the current work. It should be also noted that 11 hidden nodes (i.e. $2I + 1$, where I is the number of input variables) is the upper limit needed to map any continuous function for a network with five inputs, as discussed by Caudill (1988). The optimal network parameters are obtained by training the ANN model with different combinations of learning rates (i.e. 0.05, 0.1, 0.2, 0.4, 0.6, 0.9 and 0.95) and momentum terms (i.e. 0.05, 0.1, 0.2, 0.4, 0.6, 0.9 and 0.95). To determine the criterion that should be used to terminate the training process, the scaled mean squared error between the actual and predicted values of pile capacities of the validation set is monitored until no significant improvement in the error occurs. This is achieved at approximately 5000 training cycles (epochs).

Among the 11 models developed that correspond to each number of hidden nodes, the optimum ANN model is selected so that the model has a minimum number of hidden layer nodes accompanied with good and consistent performance with respect to the training and validation sets. A model with three hidden layer nodes; a learning rate of 0.2; a momentum term of 0.8; *tanh* transfer function for the hidden layer nodes; and *sigmoidal* transfer function for the output layer node, is found to perform best. A schematic view of the structure of the optimum ANN model is shown in Figure 1.

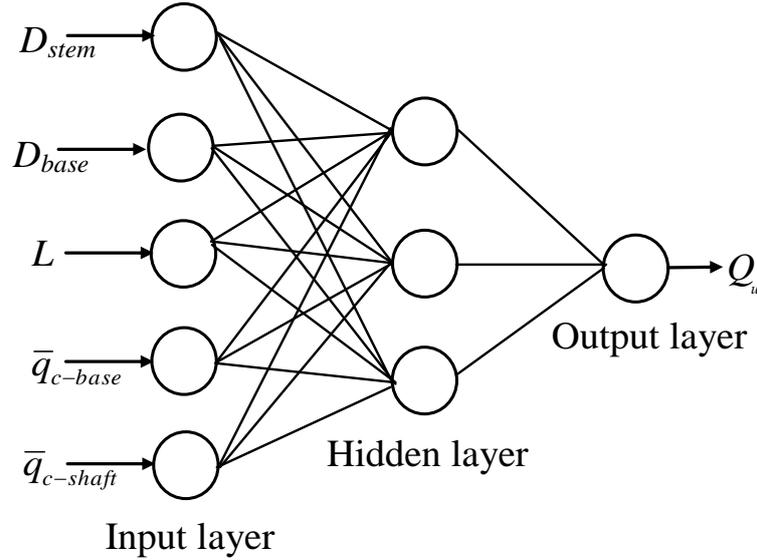


Figure 1. Schematic representation of the structure of optimum ANN model

The performance of the optimum ANN model is shown in Figure 2 (upper left graph). It can be seen that the ANN model has minimum scatter around the line of equality between the measured and predicted ultimate shaft capacities. The model also has high coefficients of correlation, r , of 0.97 and 0.96 for the training and validation sets, respectively. The above results indicate that the developed ANN performs well and can be used with confidence for the prediction of the ultimate capacity of drilled shafts.

To examine the accuracy of the ANN model against available methods, the model predictions are compared with those obtained from three commonly used CPT-based models including Schmertmann (1978), LCPC (Bustamante and Gianceselli 1982) and Alsamman (1995), and the results of comparison are shown graphically in Figure 2. In addition, statistical analyses including the coefficient of correlation and the

average prediction ratio (i.e. $\sum_{i=1}^n \left(\frac{Q_{u(predicted)}}{Q_{u(measured)}} \right) / n$; n is the number of case records)

from the ANN model and the three methods considered, in relation to the available 94 case records, are carried out and the results are given in Table 2. The results indicate that the ANN model outperforms the other CPT-based methods.

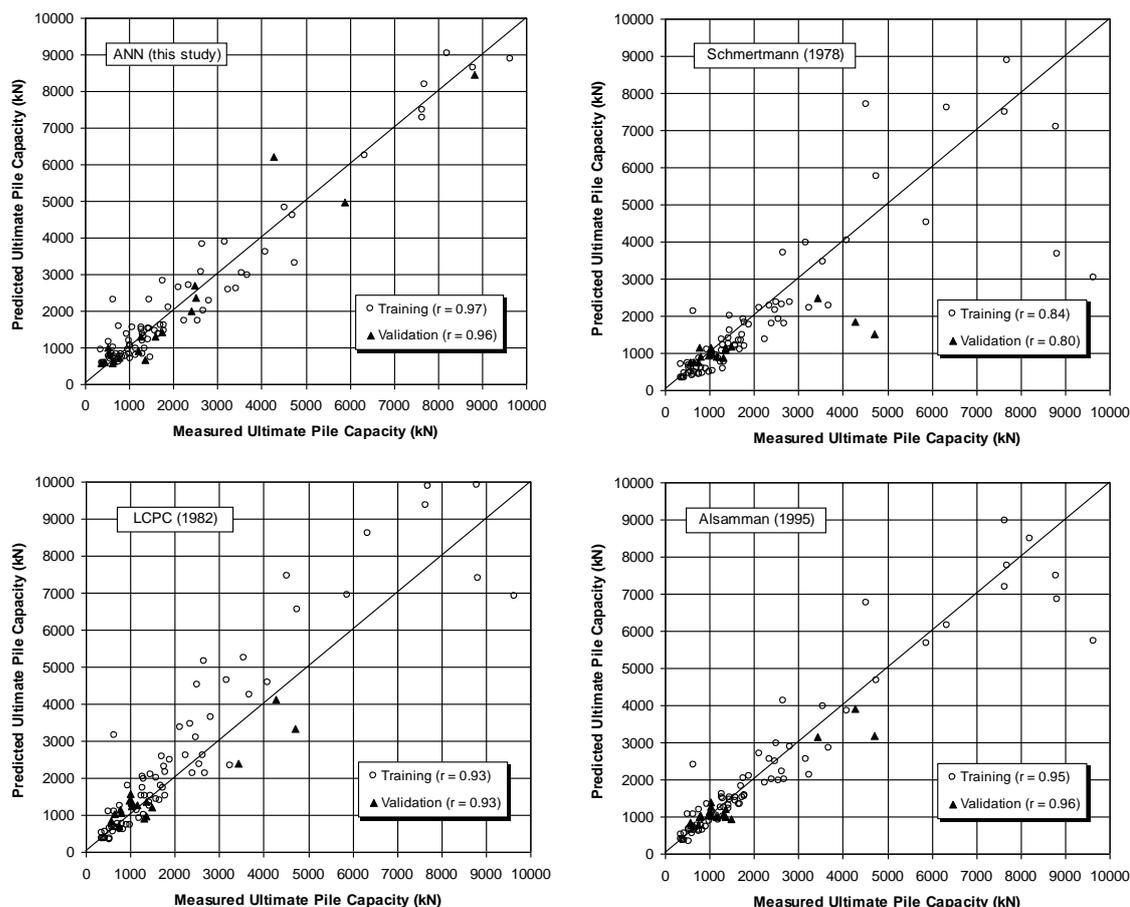


Figure 2. Performance of the optimum ANN model and other CPT-based methods in the training and validation sets

Table 2. Performance of the ANN model against available CPT-based methods

Performance measure	Method			
	ANN	Schmertmann (1978)	LCPC (1982)	Alsamman (1995)
Correlation coefficient, r	0.97	0.83	0.92	0.95
Average prediction ratio	1.01	1.23	0.93	1.03

ANN Model for Hand Calculations

In order to facilitate the use of the developed ANN model, it is translated into a relatively simple equation suitable for hand calculations. For brevity, detailed description of the procedure used to convert the ANN model into a simple equation is not given, but can be found in Shahin et al. (2002a). Based on interpretation of the developed neural network weights and biases, the predicted ultimate drilled shafts capacity can be expressed as follows:

$$Q_{u(ANN)} = 355.8 + \left[\frac{9296.3}{1 + e^{(-1.4 + 1.9 \tanh H_1 + 3.5 \tanh H_2 + 3.7 \tanh H_3)}} \right] \quad (2)$$

$$H_1 = 5.97 - 10^{-3} [1.16D_{stem} - 0.09D_{base} + 128.9L + 40.7\bar{q}_{c-base} + 95.3\bar{q}_{c-shaft}] \quad (3)$$

$$H_2 = 0.86 - 10^{-3} [0.61D_{stem} + 1.01D_{base} + 55.3L + 4.96\bar{q}_{c-base} + 6.18\bar{q}_{c-shaft}] \quad (4)$$

$$H_3 = 4.16 - 10^{-3} [0.35D_{stem} + 2.31D_{base} - 109L - 34.5\bar{q}_{c-base} - 20.8\bar{q}_{c-shaft}] \quad (5)$$

where; $Q_{u(ANN)}$ = ultimate drilled shaft capacity (kN), D_{stem} = shaft stem diameter (mm), D_{base} = shaft base diameter (mm), L = shaft embedment length (m), \bar{q}_{c-base} = weighted average cone point resistance over shaft base failure zone (MPa) and $\bar{q}_{c-shaft}$ = weighted average cone tip resistance along shaft embedment length (MPa).

It is worthwhile to note that the prediction of Equation 2 is better when used for the ranges of values of D_{stem} , D_{base} , L , \bar{q}_{c-base} and $\bar{q}_{c-shaft}$ shown in Table 1. This is because ANNs work well for interpolation rather than extrapolation, as mentioned previously. However, ANNs can always be updated to obtain better predictions by presenting new training examples of wider ranges, as new data become available.

Conclusions

The results indicate that the ANN model was capable of accurately predicting the ultimate capacity of drilled shafts, with high coefficients of correlation, r , of 0.97 and 0.96, respectively, in the calibration and validation sets used for model development. The results also demonstrate that over the four CPT-based methods of drilled shaft capacity predictions used for comparison, the ANN model performs best followed by Alsamman (1995), LCPC (Bustamante and Gianceselli 1982) and the Schmertmann method (1978).

References

- Abu-Kiefa, M. A. (1998). "General regression neural networks for driven piles in cohesionless soils." *Journal of Geotechnical & Geoenvironmental Engineering*, ASCE, 124(12), 1177-1185.
- Alsamman, O. M. (1995). "The use of CPT for calculating axial capacity of drilled shafts," PhD Thesis, University of Illinois-Champaign, Urbana, Illinois.

- Bustamante, M., and Gianceselli, L. (1982). "Pile bearing capacity prediction by means of static penetrometer CPT." *Proceedings of the 2nd European Symposium on Penetration Testing*, Amsterdam, 493-500.
- Caudill, M. (1988). "Neural networks primer, Part III." *AI Expert*, 3(6), 53-59.
- Fausett, L. V. (1994). *Fundamentals neural networks: Architecture, algorithms, and applications*, Prentice-Hall, Englewood Cliffs, New Jersey.
- Hornik, K., Stinchcombe, M., and White, H. (1989). "Multilayer feedforward networks are universal approximators." *Neural Networks*, 2, 359-366.
- Kulhawy, F. H., and Mayne, P. W. (1990). "Manual on estimating soil properties for foundation design " *Report EL-6800*, Electric Power Research Institute, Palo Alto, CA.
- Lee, I. M., and Lee, J. H. (1996). "Prediction of pile bearing capacity using artificial neural networks." *Computers and Geotechnics*, 18(3), 189-200.
- Masters, T. (1993). *Practical neural network recipes in C++*, Academic Press, San Diego, California.
- Neosciences. (2000). *Neuframe Version 4.0*, Neosciences Corp., Southampton, Hampshire.
- Reese, L. C., and O'Neill, M. W. (1988). "Drilled shafts: Construction procedures and design methods." *FHWA-HI-88-042*, U. S. Department of Transportation, Dallas, Texas.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). "Learning internal representation by error propagation." *Parallel Distributed Processing*, D. E. Rumelhart and J. L. McClelland, eds., MIT Press, Cambridge.
- Schmertmann, J. H. (1978). "Guidelines for cone penetration test, performance and design." *Report No. FHWA-TS-78-209*, U. S. Department of Transportation, Washington, D. C.
- Shahin, M. A. (2008). "Modeling axial capacity of pile foundations by intelligent computing." *Proceedings of the Second BGA International Conference on Foundations, ICOF2008*, Dundee, Scotland, 283-294.
- Shahin, M. A., and Jaksa, M. B. (2006). "Pullout capacity of small ground anchors by direct cone penetration test methods and neural networks." *Canadian Geotechnical Journal*, 43(6), 626-637.
- Shahin, M. A., Jaksa, M. B., and Maier, H. R. (2002a). "Artificial neural network-based settlement prediction formula for shallow foundations on granular soils." *Australian Geomechanics*, 37(4), 45-52.
- Shahin, M. A., Maier, H. R., and Jaksa, M. B. (2002b). "Predicting settlement of shallow foundations using neural networks." *Journal of Geotechnical & Geoenvironmental Engineering*, ASCE, 128(9), 785-793.
- Shahin, M. A., Maier, H. R., and Jaksa, M. B. (2004). "Data division for developing neural networks applied to geotechnical engineering." *Journal of Computing in Civil Engineering*, ASCE, 18(2), 105-114.