Prediction of pile behavior using artificial neural networks based on standard penetration test data

F. Pooya Nejad
Ferdowsi University of Mashhad, Iran

M.B. Jaksa
The University of Adelaide, Australia

Abstract

This paper presents an artificial neural network (ANN) model for the prediction of non-linear behavior of vertically loaded piles based on the results of standard penetration test (SPT) data. The geotechnical literature has in-cluded many methods, both theoretical and experimental, to predict pile behavior. Most of the available methods simplify the problem by incorporating several assumptions associated with the factors that affect pile behavior. With respect to the design of pile foundations, accurate prediction of pile behavior is necessary to ensure appropriate structural and serviceability performance. Approximately, 1,000 data sets, obtained from the published literature, are used to develop the ANN model. In addition, the paper discusses the choice of input and internal network parameters which were examined to obtain the optimum model. Finally, the paper proposes a series of charts for predicting pile behavior that will be useful for pile design.

1 INTRODUCTION

Pile foundations are the part of the sub-structure and their main function is to support and transfer loads to some depth below the ground surface. The evaluation of the load-settlement performance of a single pile is one of the main aspects in the design of piled foundations. In addition, the behavior of a pile is influenced by several factors, such as the mechanical non-linear behavior of the soil, the characteristics of the pile itself, as well as its method of installation (Berardi & Bovolenta 2005). There are many techniques in the geotechnical literature, both theoretical and experimental, for predicting the settlement of piles. In recent years, artificial neural networks (ANNs) have been applied to many geotechnical engineering problems and have demonstrated some degree of success. In this paper, ANNs are used to predict the behavior of piles based on standard penetration test (SPT) data. The aim of the paper is to propose some ANN-based load-settlement charts for predicting pile behavior.

2 NEURAL NETWORK MODEL

The development of ANN models requires the determination of model inputs and outputs, division and pre-processing of the available data, the determination of appropriate network architecture, stopping criteria and model validation (Shahin et al. 2002). In this work, the PC-based computer software NEUFRAME version 4.0 (Neurosciences Corp. 2000) is used to simulate ANN operation. The data used to calibrate and validate the neural network model were obtained from the literature of pile load tests. Suitable case studies were those having pile load tests that include field measurements of full-scale pile settlements, as well as the corresponding information regarding the piles and soil characteristics. The database contains a total of 1,013 cases from 76 individual pile load tests. Details of the database are given by Pooya Nejad et al. (2009).

3 MODEL INPUTS AND OUTPUTS

In order to obtain accurate predictions of pile behavior (including settlement and capacity), an understanding of the factors affecting pile behavior is needed. Most traditional methods include the following fundamental parameters: pile geometry, pile material properties, soil properties and applied load (for prediction of settlement). Additional factors include the method of pile installation, the type of load test and whether the pile tip is closed or open. Since pile behavior depends on soil strength and compressibility and the SPT is one of the most commonly used tests in practice for quantifying such soil characteristics, the SPT blow count/300 mm (N) along the embedded length of the pile is used in this study. To account more accurately for the variability of soil properties along the shaft of the pile, the embedded length of the pile is divided into five segments of equal thickness, with each associated with an average of N over that segment. In addition, as suggested by Liao & Whitman (1986), for sand the value of N_{ave_j} for each subdivision is corrected for overburden pressure, as given below. This correction is not used for clays.

$$N_{\text{correct}_j} = C_N \times N_{\text{ave}j} \quad (1)$$

$$C_N = \sqrt{\frac{95.76}{\sigma_v'}} \quad (2)$$

where C_N = the adjustment for effective overburden pressure; and σ_v' = the effective overburden pressure (kPa).

Hence, the factors that are presented to the ANN as model input variables are the: (i) type of pile load test (maintained load or constant rate of penetration); (ii) pile material (concrete, steel, composite and plastic); (iii) method of installation (replacement or displacement); (iv) pile tip (closed or open); (v) axial rigidity of the pile (EA); (vi) cross-sectional area of the pile tip (A_{tip}); (vii) perimeter of the pile in contact with the soil (O); (viii) length of the pile (L); (ix) embedded length of the pile (L_{embed}); (x–xiv) the averaged and corrected SPT blow count/300 mm along the embedded length of the pile (N_1, N_2, N_3, N_4, N_5); (xv) the corrected SPT blow count/300 mm at the tip of the pile (N_{tip}); and (xvi) applied load (P). Pile settlement is the single output variable.

4 DATA DIVISION AND PRE-PROCESSING

In this study cross-validation, as suggested by Stone (1974) is implemented to divide the data are be into three sets: training, testing and validation. The training set is used to adjust the connection weights, whereas the testing set is used to check the performance of the model at various stages of training and to determine when to stop training to avoid overfitting. The validation set is used to estimate the performance of the trained network in the deployed environment. In total, 85.6% of the data (867 cases) are used for training and 14.4% (146 cases) are used for validation. The training data are further divided into 81% (701 cases) for the training set and 19% (166 cases) for the testing set. Since it is essential that the data used for training, testing, and validation represent the same population (Masters 1993), the statistical properties (e.g. mean, standard deviation and range) of the data subsets need to be similar (Shahin et al. 2004). In this study in order to achieve this, several random combinations of the training, testing and validation sets are examined until three statistically consistent data sets are obtained (Pooya Nejad et al. 2009).

In general, after dividing the available data into their subsets, the variables are pre-processed by scaling them to a suitable form and to eliminate their dimension, before presenting them to the ANN. The output variables also need to be scaled to be commensurate with the limits of the transfer functions used in the output layers (Shahin 2003). In this particular application, scaling of the input variables is unnecessary. Hence, in this study the output variables are scaled between 0.0 to 1.0, as the sigmoid transfer function is used in the output layer.

5 ANN MODEL ARCHITECTURE

Determining the network architecture is one of the most important and difficult tasks in ANN model development (Maier & Dandy 2000). It requires the selection of the optimum number of hidden layers and the number of nodes in each of these. There is, however, no unified theory for achieving this (Shahin 2003). The number of nodes in the input and output layers are restricted by the number of model inputs and outputs. A total of 16 input variables are included in this study and the output layer has a single node representing the measured value of settlement.

In this study, models incorporating a single and multiple hidden layers are examined. In order to determine the optimum network geometry, first ANNs with a single hidden layer are trained, followed by models with two, three and four hidden layers with different numbers of nodes in the hidden layers.

Both models (single layer and multi-hidden layers) have been trained with a sigmoidal and a hyperbolic tangent (tanh) transfer function for the hidden layers. In the single hidden layer model, a sigmoidal transfer function is adopted for the hidden and output layers, and for the multi-hidden layer models, a tanh transfer function is used for the hidden layers and a sigmoidal transfer function is adopted for the output layer.

6 TRAINING

Training, or learning, is the process of optimizing the connection weights. Its aim is to identify a global solution to what is typically a highly non-linear optimization problem (White 1989). The method most commonly used for finding the optimum weight combination of feed-forward neural networks is the back propagation algorithm (Rumelhart et al. 1986), which is based on first-order gradient descent. The advantage of this method is that it has the ability to escape local minima in the error surface and, thus, produces optimal or near optimal solutions (Shahin 2003). However, it also has a slow convergence rate. Consequently, in this study, the back-propagation algorithm is used for optimizing the connection weights.

The general strategy adopted for identifying the optimal parameters that control the training process is as follows. A number of trials are carried out using NEUFRAME's default parameters, i.e. a momentum of 0.8 and a learning rate of 0.2. The network that performs best is then retrained with the different combinations of momentum terms and learning rates in an attempt to improve model performance (Pooya Nejad et al. 2009).

7 STOPPING CRITERIA

Stopping criteria determine whether the model has been optimally or sub-optimally trained (Maier & Dandy 2000). Many approaches can be used to determine when to stop training. As mentioned previously, the cross-validation technique (Stone 1974) is used in this work, as it is considered that sufficient data are available to create training, testing and validation sets and it is the most valuable tool to ensure over-fitting does not occur (Smith 1993). The training set is used to adjust the connection weights, whereas the testing set measures the ability of the model to generalize and, using this set, the performance of the model is checked at many stages during the training process and training is stopped when the testing set error begins to increase (Shahin et al. 2002).

8 MODEL VALIDATION

Once model training has been successfully accomplished, the performance of the trained model should be validated against data that have not been used in the learning process. This data set is known as the validation set. The purpose of the model validation phase is to ensure that the model has the ability to generalize within the limits set by the training data in a robust fashion, rather than simply having memorized the input-output relationships that are contained in the training data (Shahin et al. 2002). The coefficient of correlation, r, the root mean squared error, RMSE, and the mean absolute error, MAE, are the main criteria that are used to evaluate the prediction performance of ANN models. The RMSE is the most popular measure of error and it has the advantage that large errors receive much greater attention than small ones (Hecht-Nielsen 1990).

9 RESULTS

As stated previously, in this study two ANN models have been developed. The first model incorporates a single hidden layer and the second utilizes multiple hidden layers. In order to determine the optimum network geometry, ANNs are trained with a single hidden layer incorporating different numbers of hidden layer nodes, and then ANNs are trained with two, three and four hidden layers with different numbers of nodes in the hidden layers.

The results of the optimum networks for the single hidden layer and the multiple hidden layers are summarized in Table 1. It is observed that Model 7 is the optimum of the single layer models with 7 nodes in the hidden layer, and Model 14-6 is the optimum of the two hidden layer models, with 14 nodes in the first hidden layer and 6 nodes in the second. Model 13-8-3 is the optimum of the three hidden layer models, with 13 nodes in the first, 8 in the second and 3 in the third hidden layer. Finally, it can be seen that the best result is obtained by the four hidden layer model with 15-13-5-2 nodes in the four hidden layers.

The effect of the internal parameters controlling the back-propagation algorithm (i.e. momentum and learning rate) on model performance (model with 15-13-5-2 nodes in the four hidden layers) was examined by Pooya Nejad et al. (2009). The best prediction was obtained with a momentum of 0.6 and

Table 1 Results of optimum single and multiple hidden layer networks

Optimum single and multiple hidden layer networks	Correlation coefficient, r	RMSE (mm)
7 (single)		
Training	0.922	6.95
Testing	0.818	6.55
Validation	0.720	12.42
14-6		
Training	0.993	2.12
Testing	0.937	3.02
Validation	0.950	10.34
13-8-3		
Training	0.993	2.10
Testing	0.949	2.74
Validation	0.930	7.00
15-13-5-2		
Training	0.991	2.42
Testing	0.930	3.20
Validation	0.961	5.12

Table 2 Results of optimum model with the optimum internal parameters

Optimum multiple hidden layer network	Correlation coefficient, r	RMSE (mm)
15-13-5-2		
Training	0.993	2.17
Testing	0.958	2.47
Validation	0.972	4.49

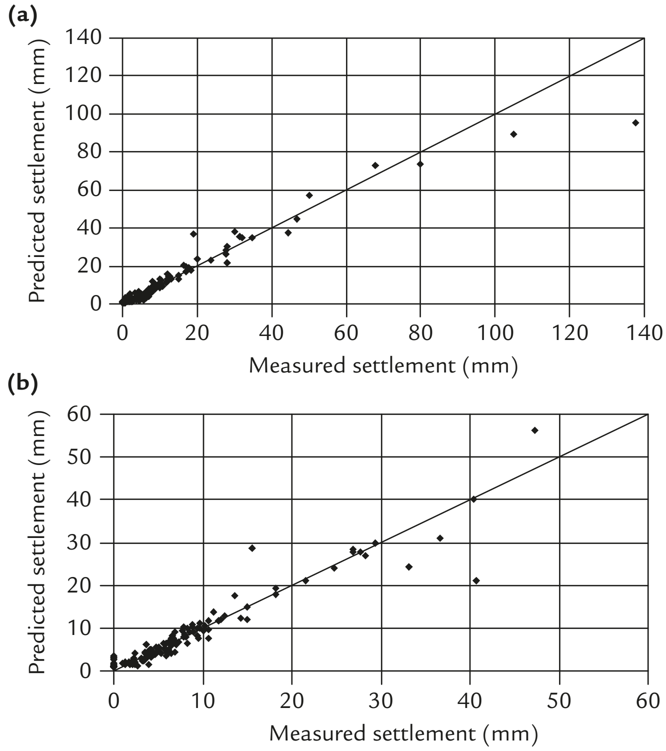


Figure 1 Measured versus predicted settlements for ANN models with 4 hidden layers with 15-13-5-2 hidden layer nodes: (a) Validation set; and (b) Testing set.

learning rate of 0.4. The results for the optimum model with optimal momentum and learning rate are shown in Table 2.

The relationship between the measured and predicted settlements for the validation and testing sets are shown in Figure 1. The results indicate that the model performs well, with an r and RMSE, respectively, of 0.972 and 4.49 mm for the validation set and 0.958 and 2.47 mm for the testing set.

10 PILE DESIGN CHARTS

By implementing the optimum model, it is possible to generate several load-settlement charts for various combinations of input parameters. Focusing on concrete piles, Figures 2 to 4 provide typical load-settlement curves for maintained pile load tests on piles of different lengths (L) and diameters (d) in soil with various SPT numbers (average N along the pile shaft).

It can be observed that in some of the load-settlement charts, immediately prior to plastic yielding, the settlement decreases for increasing loads. This is noticeable at pile lengths of 20 m and greater. This behavior is unrealistic and is likely the result of a lack of measured data in the training set. This inaccuracy is likely to improve as more data become available and the model is subsequently retrained.

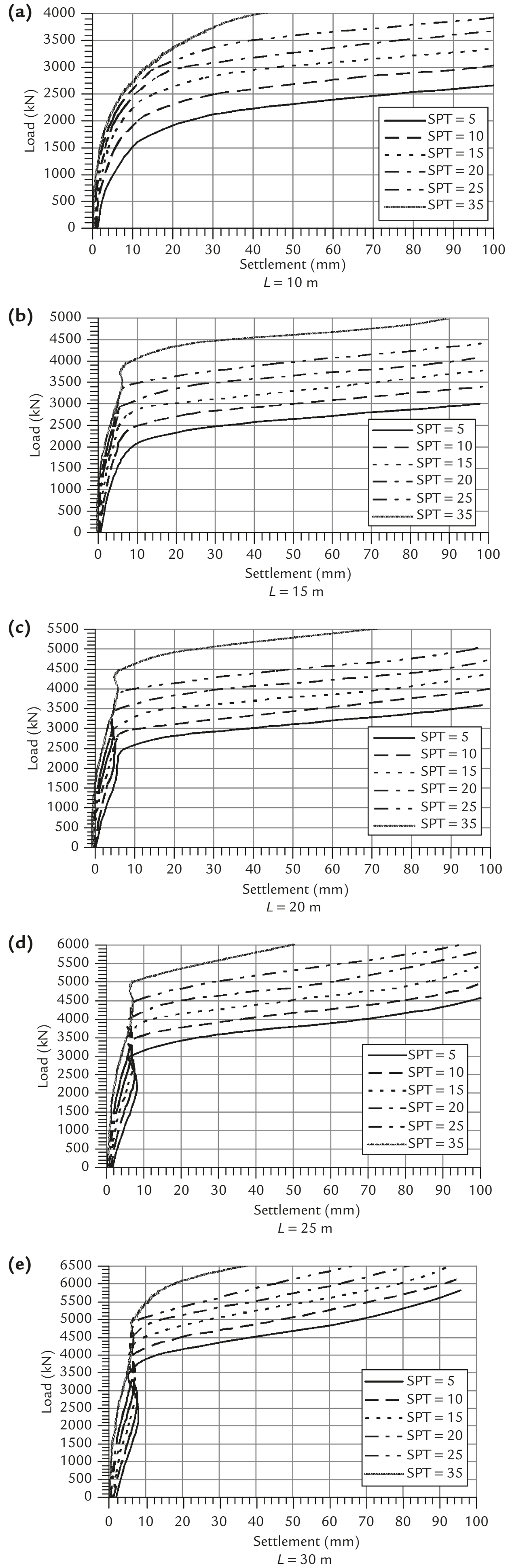
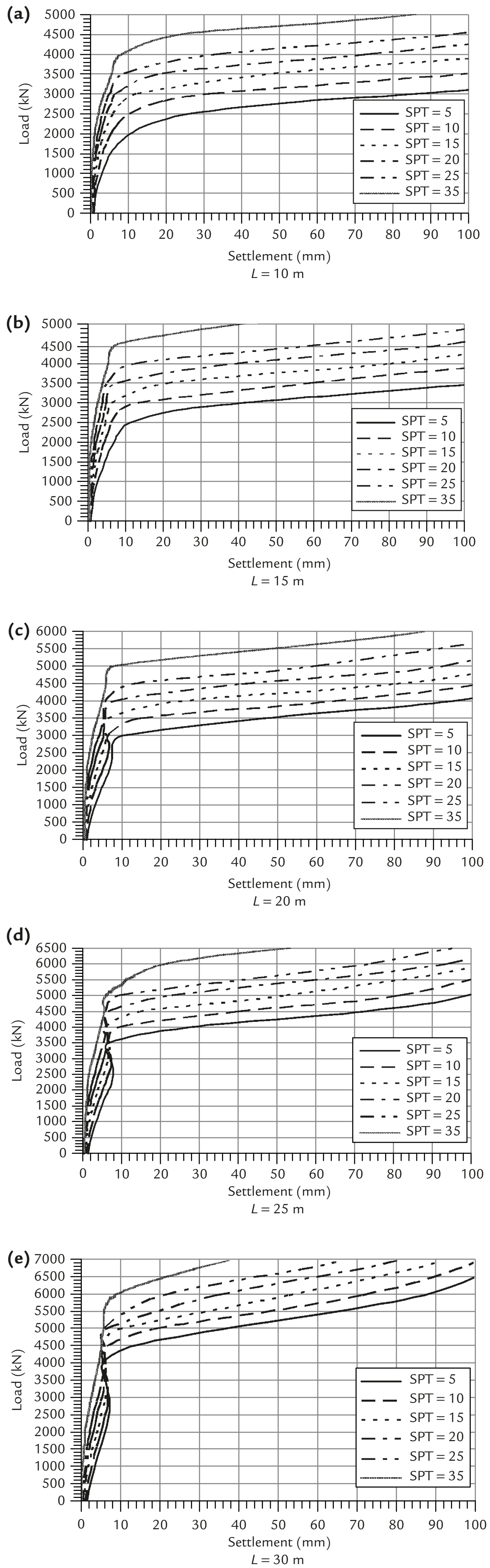


Figure 2 Load-settlement curves for maintained pile load tests on concrete piles 700 mm in diameter.



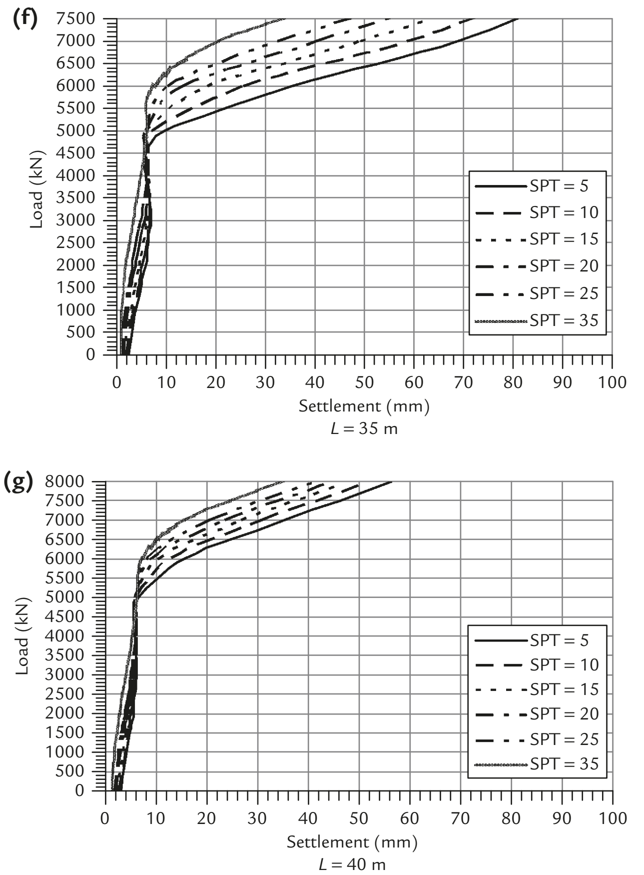


Figure 3 Load-settlement curves for maintained pile load tests on concrete piles 1,000 mm in diameter.

11 CONCLUSIONS

A back-propagation neural network has been used to examine the feasibility of ANNs to predict the load-settlement characteristics of piles. A database containing 1,013 case records of field measurements of pile settlements was used to develop and verify the model. The results indicate that back-propagation neural networks have the ability to predict the behavior of piles with an acceptable degree of accuracy ($r = 0.972$, RMSE = 4.49 mm) for settlements up to 185 mm. The ANN method has an additional advantage over conventional methods in that, once the model is trained, it can be used as an accurate and quick tool for estimating the behavior of piles.

From the optimal model several load-settlement charts for concrete piles with various lengths and diameters founded in soil with a range of SPT values (average of N along the pile) have been proposed to assist with pile design.

ACKNOWLEDGEMENTS

The authors would like to thank Messrs Richard Herraman and Peter Lai from the Departments of Transportation in Adelaide and Florida, respectively, and also Dr. Bryan McCabe from the Department

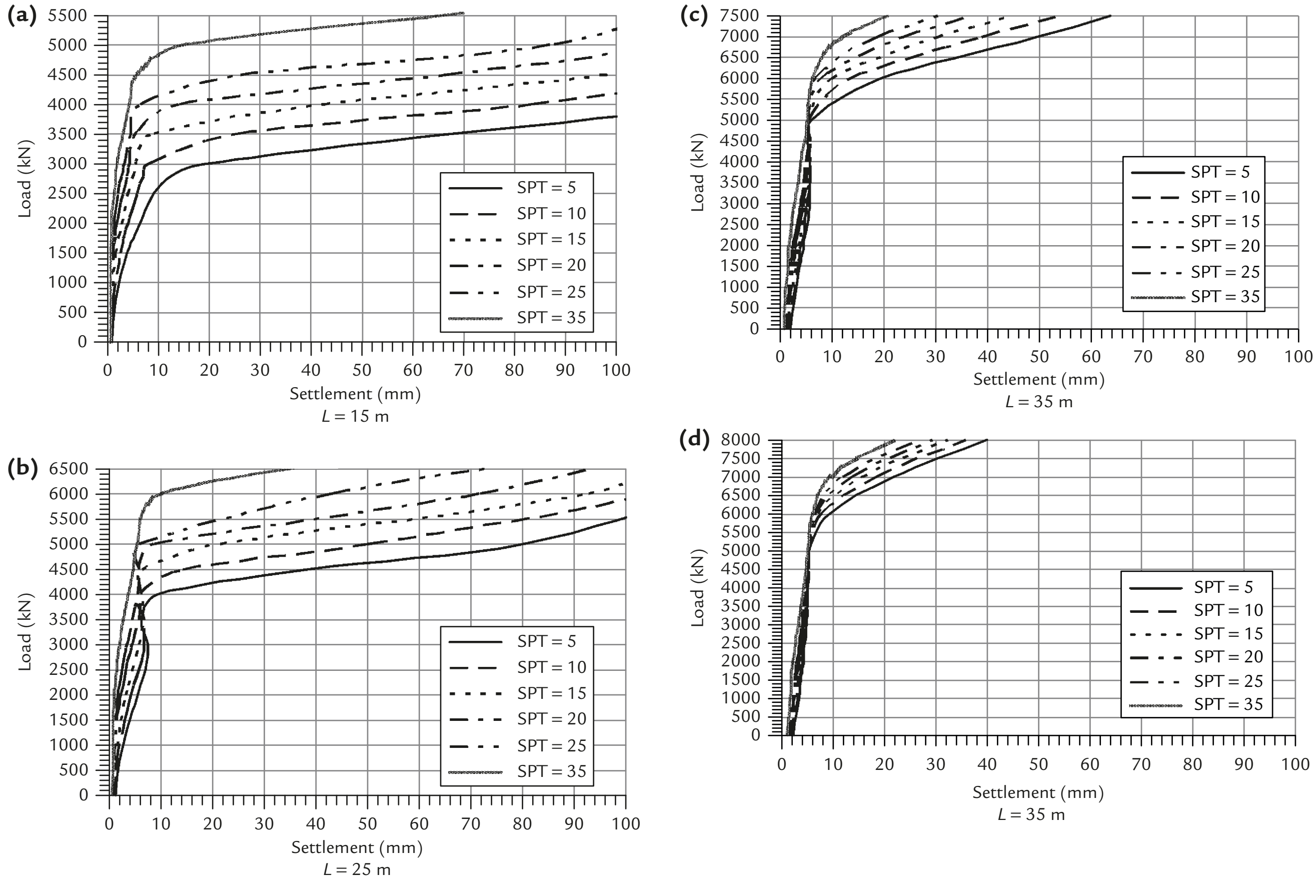


Figure 4 Load-settlement curves for maintained pile load tests on concrete piles 1,200 mm in diameter.

of Civil Engineering, National University of Galway, Ireland and Keller Ground Engineering, UK, for their assistance in providing access to some of their pile load test data.

REFERENCES

Berardi, R. & Bovolenta, R. 2005. Pile settlement evaluation using field stiffness non-linearity. *Geotechnical Engineering,* 158: 35–44.

Hecht-Nielson, R. 1990. *Neurocomputing,* Addison-Wesley, Reading, Mass.

Liao, S.S. & Whitman, R.V. 1986. Overburden correction factors for SPT in sand. *Journal of Geotechnical Engineering,* ASCE, 112(3): 373–377.

Maier, H.R. & Dandy, G.C. 2000. Applications of artificial neural networks to forecasting of surface water quality variables: Issues, applications and challenges. *Artificial Neural Networks in Hydrology*, Govindaraju R.S. and Rao A.R., eds., Kluwer, Dordrecht, The Netherlands, 287–309.

Masters, T. 1993. *Practical neural network recipes in C++,* Academic, San Diego.

Pooya Nejad, F., Jaksa, M.B., Kakhi, M. & McCabe, B.A. 2009. Prediction of pile settlement using artificial neural networks based on standard penetration test data, *Computers and Geotechnics,* 36: 1125–1133.

Rumelhart, D.E., Hinton, G.E. & Williams, R.J. 1986. Learning internal representation by error propagation. *Parallel distributed processing,* Vol. 1, Chap. 8, D.E. Rumelhart and J.L. McClelland, eds., MIT Press, Cambridge, Mass.

Shahin, M.A. 2003. Use of Artificial Neural Networks for Predicting Settlement of Shallow Foundations on Cohesionless Soil. PhD thesis, University of Adelaide.

Shahin, M.A., Maier, H.R. & Jaksa, M.B. 2002. Predicting settlements of shallow foundations using artificial neural networks. *J. Geotech. and Geoenv. Engrg.,* ASCE, 128(9): 785–793.

Shahin, M.A., Maier, H.R. & Jaksa, M.B. 2004. Data division for developing neural networks applied to geotechnical engineering. *J. Computing in Civil Engrg.,* ASCE, 18(2): 105–114.

Smith, M. 1993. *Neural networks for statistical modelling,* Van Nostrand-Reinhold, New York.

Stone, M. 1974. Cross-validatory choice and assessment of statistical predictions. *J. R. Stat. Soc.,* B 36, 111–147.