Proc. of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Jinsong Huang, D.V. Griffiths, Shui-Hua Jiang, Anna Giacomini and Richard Kelly

Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-5182-7_00-005-cd

ISGSR 2022

Applications of Artificial Intelligence in Geotechnical Engineering

Mark Jaksa[1]

[1]School of Civil, Environmental and Mining Engineering, University of Adelaide, Australia.
E-mail: mark.jaksa@adelaide.edu.au

Abstract: Since artificial intelligence (AI) was first applied to geotechnical engineering in the early-1990s, in the form of artificial neural networks (ANNs), it has shown great promise in providing superior estimates of a wide range of geotechnical behaviors and applications, when compared against traditional techniques. This paper examines two artificial intelligence techniques, namely ANNs and genetic programming (GP), that have been applied successfully by the author and his co-workers to the settlement of shallow and pile foundations and predicting the performance of rolling dynamic compaction. The paper also examines the benefits and limitations of AI in the context of geotechnical engineering, and the process for developing optimal models.

Keywords: Machine learning;artificial neural networks; genetic programming; shallow foundations; pile foundations; ground improvement.

1 Introduction

In 1956, a group of young scientists first imagined and explored the proposition that “every aspect of learning or any other feature of intelligence can, in principle, be so precisely described that a machine can be made to simulate it” (Crevier 1993). Artificial intelligence (AI) was thus born, and since that time, it has pervaded almost every aspect of human endeavor. Earlier, McCulloch and Pitts (1943) proposed a mathematical framework for artificial neural networks (ANNs), which seek to mimic the learning behavior of the brain. In the late 1980s, research in ANNs gained considerable momentum with the work of Rumelhart et al. (1986) and McClelland and Rumelhart(1988).ANNs werefirst applied to geotechnical engineering in the early 1990s,by Sterling and Lee (1992) and Gribb and Gribb (1994), as reported by Baghbani et al. (2022), and also by Goh (1994). Since then, AI, ANNs and machine learning have been applied to almost every aspect of geotechnical engineering. This paper focuses on machine learning, which is a subset of AI, and in particular on ANNs and genetic programing (GP). Since ANNs and GP are data-driven methods, they do not need any prior knowledge of the nature of the relationship(s) between the input and output variables, and this has made them particularly amenable to solving non-linear, complex problems, and consequentlyare well-suited to geotechnical engineering applications.This paper presentsthe application of ANNs and GP to a number of geotechnical engineering problems in order to demonstrate their efficacy.The processesfor obtaining optimal models will be discussed, as well as the benefits and limitations of ANNs and GP.

1.1 Artificial neural networks (ANNs)

Artificial neural networks (ANNs) arenumerical surrogates of the human brain, which is a biological neural network. The latter consists of billions of neurons, or nerve cells, which are connected via synapses. The ANN equivalent is termed a *node* or a *processing element* which,like neurons, are interconnected. A typical ANN structure is shown in Figure 1. On the left-hand-side of the figureis a series of input nodes, or input variables, which together form the *input layer*. On the right-hand-side is the *output layer*, which consists of one or more output nodes. An ANN also incorporates one or more hidden layers, which comprise a number of hidden layer nodes. Figure 1 shows a feedforward network, as the connections between nodes are in the forward direction only, whereas feedback networks pass information both in the forward and reverse directions.

As humans learn, the connections between the neurons (i.e. the neural pathways) strengthen. In the same way, as ANNs are presented with more data, they ‘learn’ and the weights (which are similar to the coefficients in traditional statistics) of the connections between the nodes are optimized, so that they yield the most accurate predictions of the output variable(s).

During the model development phase, the data are subdivided into three separate data sets, training, testing and validation. The training set usually contains the majority of the data, and it is this data set that the ANN uses to ‘learn’ the relationship(s) between the input parameters and the output(s); in other words, to optimize the connection weights. The testing data set is used to ‘test’ the ANN to ensure that it does not overfit the data. This is important, because the objective of developing most ANN models is to obtain one that is general enough to provide reliable predictions over a wide range of input values. Finally, the validation data set is one that is hidden from the ANN software during model development. It is only after the ANN model has been optimized, using only the training and testing data sets, that the validation set is presented to the ANN model to assess its accuracy.

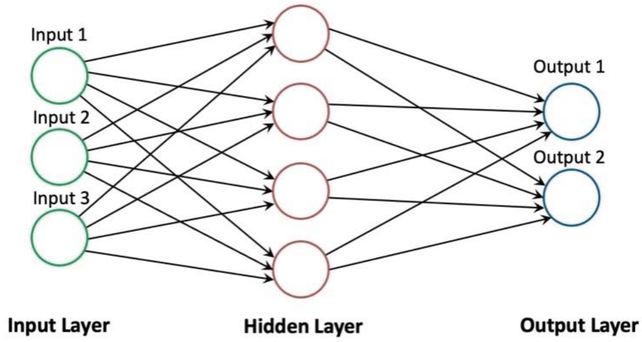


Figure 1. Typical ANN structure.

1.2 Genetic programing (GP)

Genetic programing (GP) is one of a number of approaches based upon evolutionary algorithms that mimic Darwin's evolution theory in relation to optimizing a solution to a predefined problem(Koza 1992).In GP, the individuals in a population are represented by computer 'programs' of variable size and shape that are hierarchically composed of a set of functions and terminals fitted to a particular problem domain (Koza 1992). The function set may consist of arithmetic functions (+, –, ×, /), mathematical functions (sin, *cos*, *ln*), Boolean logic operators (AND, OR, NOT), logical expressions (IF or THEN), iterative functions (DO, CONTINUE, UNTIL) and/or other user-defined functions (Sette and Boullart 2001). The terminal set typically consists of input variables attached to the problem domain and pre-specified, or randomly generated, numeric constants.

2 Geotechnical Engineering Applications

The processes of establishing and assessing the efficacy of ANN and GP models are best appreciated by applying them to various geotechnical engineering problems. My foray into AI began with ANNsmore than two decades ago in 1999, along with my first PhD student, Mohamed Shahin, who is now a professor at Curtin University, in Western Australia. Together, among other topics, we explored the efficacy of applying ANNs to predicting the settlement of shallow foundations on sands. This work is presented below in §2.1. In §2.2, the work undertaken together with Dr. FereydoonPooyaNejadis presented,where ANNs were used to predict the load-settlement behavior of a wide range of piles in ground with varying properties. Finally, in §2.3, more recent work will be explored involving the use of both ANNs and GP to predict the performance of the 4-sided impact rollerin varying ground conditions.

2.1 Shallow foundations using ANNs

Shahin et al. (2002b) examined the efficacy of applying ANNs to the settlement of shallow foundations on sand. The first step of the process of developing any ANN model is to compile a database. From the published literature, Shahin et al. (2002b) created a database which included189 separate shallow foundation load-displacement tests.The databaseincorporated 5 input variables (foundation width, B (m); net applied foundation pressure, q (kPa); average standard penetration test (SPT) number, N; foundation geometry ratio, L/B, where L(m) is the foundation length; and the foundation embedment ratio, D_f/B, where D_f(m) is the foundation embedment depth) and a single output variable, the measured settlement,S_m (mm). It is acknowledgedthat the SPT is not the most ideal measurementtechnique by which to represent the geotechnical characteristics of the sands in the model.However, when compiling databases, one must rely on the available data, otherwise one must perform the testing oneself, and generally the latter is a major endeavor, and is often prohibitive.

As mentioned above in §1.1, prior to ANN model development, the database needs to be subdivided into three separate data sets, training, testing and validation. For optimal ANN performance, it has been shown that it is best to ensure that the statistics of each of the three data sets, are as similar as possible (Shahin et al. 2004). This is to ensure that the three data sets are equivalent to the same overall, database population. As a consequence, the statistics of the input and output variables included in the database, for each of the three separate data sets, are summarized in Table 1.

The ANN model development was facilitated by the *Neuframe* software package (Neusciences 2000), which is unfortunately no longer available. However, interested readers who wish to develop ANN models of their own are encouraged to explore*MATLAB* (MathWorks 2022), which provides sophisticated ANN tools.

Determining the architecture and the internal settings for yielding optimal ANN performance is not automatic and is usually undertaken iteratively using a trial-and-error approach. It is beyond the scope of this paper to provide a detailed treatment of the entire ANN model development process and the internal parameters incorporated therein, but briefly, one most determine the appropriate number of hidden layers, and hidden layer nodes, epoch size, momentum term and learning rate, to yield optimal performance. Readers interested in learning more about these,and the optimization process,are referred to Shahin (2003). After several iterations, the model shown in Figure 2 was deemed optimal. As can be seen, only a single hidden layer incorporating two nodes, was adopted. As will be seen later, such a parsimonious model is particularly desirable as it can facilitate translating the ANN model into a tractable equation suitable for hand calculation.

Table 1. Summary of statistics for the training, testing and validation data sets (Shahin 2003).

Model variables and data sets	Statistical parameters				
	Mean	Std. dev.	Min.	Max.	Range
Footing width, B (m)					
Training	8.3	9.8	0.8	60.0	59.2
Testing	9.3	10.9	0.9	55.0	54.1
Validation	9.4	10.1	0.9	41.2	40.3
Footing net applied pressure, q(kPa)					
Training	188.4	129.0	18.3	697.0	678.7
Testing	183.2	118.7	25.0	584.0	559.0
Validation	187.9	114.6	33.0	575.0	542.0
Average SPT blow count, N					
Training	24.6	13.6	4.0	60.0	56.0
Testing	24.6	12.9	5.0	60.0	55.0
Validation	24.3	14.1	4.0	55.0	51.0
Footing geometry, L/B					
Training	2.1	1.7	1.0	10.6	9.6
Testing	2.1	1.9	1.0	9.9	8.9
Validation	2.1	1.8	1.0	8.1	7.1
Footing embedment ratio, D_f/B					
Training	0.52	0.57	0.0	3.4	3.4
Testing	0.49	0.52	0.0	3.0	3.0
Validation	0.59	0.64	0.0	3.0	3.0
Measured settlement, S_m (mm)					
Training	20.0	27.2	0.6	121.0	120.4
Testing	21.4	26.6	1.0	120.0	119.0
Validation	20.4	25.2	1.3	120.0	118.7

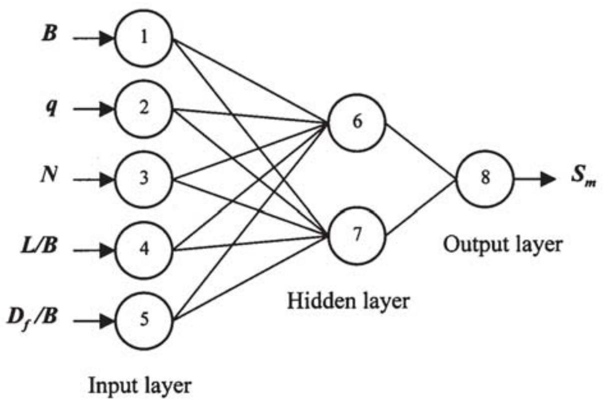


Figure 2. ANN structure for predicting settlements of shallow foundations on sand (Shahin 2003).

When assessing the performance of AI models, goodness-of-fit metrics are calculated, such as the coefficient of correlation, r, (which is given in any standard statistical textbook), the root mean squared error, $RMSE$(Eq. 1), and the mean absolute error, MAE(Eq. 2).

$$RMSE = \left\{ \frac{1}{n} \sum_{j=1}^{n} \left(y_j - d_j \right)^2 \right\}^{\frac{1}{2}} \tag{1}$$

$$MAE = \frac{1}{n} \sum_{j=1}^{n} \left| y_j - d_j \right| \tag{2}$$

where: y_j is the model (predicted) output; d_j is the desired (observed) output; and n is the number of data.

In order to assess the true predictive performance of the model, in the analyses that follow, only the validation set will be examined, recalling that the validation data are withheld from the model during training. When compared against the commonly used, traditional methods (Meyerhof 1965; Schultze and Sherif 1973; Schmertmann et al. 1978), as shown in Figure 3 and Table 2, the ANN performs extremely well.

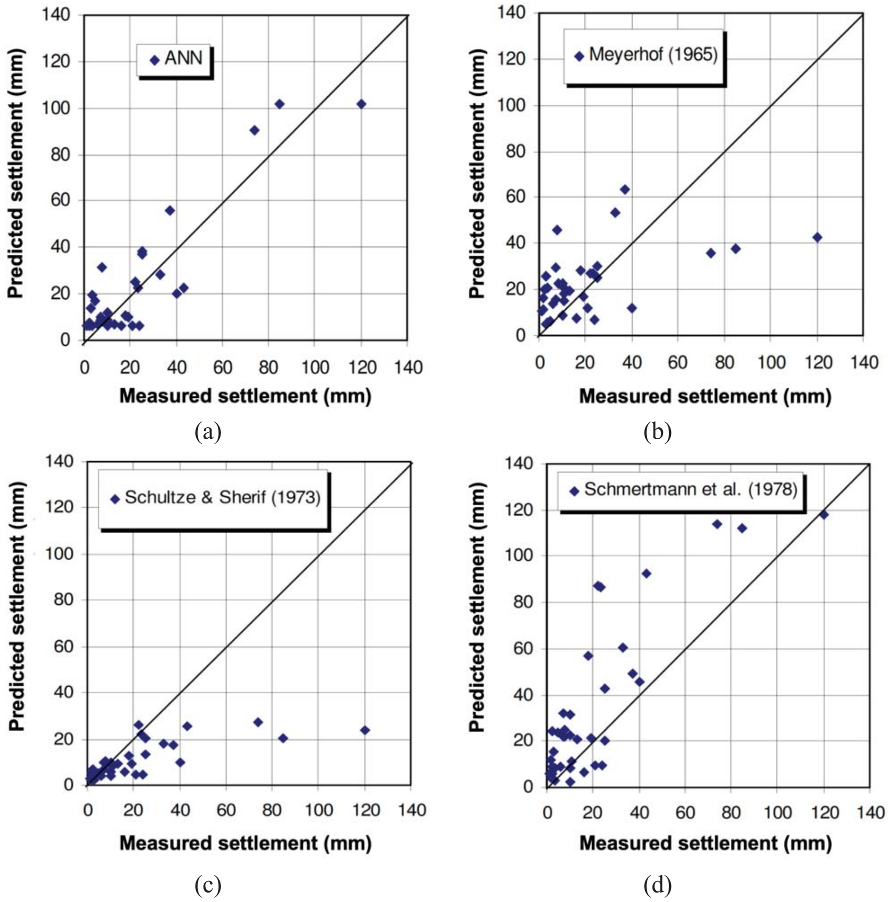


Figure 3. Predicted versus measured settlements of shallow foundations on sand for the: (a) ANN, (b) Meyerhof (1965), (c) Schultze and Sherif (1973), and (d) Schmertmann et al. (1978) methods (Shahin 2003).

Table 2. Summary of performance of ANN model and traditional methods(Shahin 2003).

Performance measure	ANN	Meyerhof (1965)	Schultze and Sherif (1973)	Schmertmann et al. (1978)
r	0.905	0.440	0.729	0.798
$RMSE$ (mm)	11.04	25.72	23.55	23.67
MAE (mm)	8.78	16.59	11.81	15.69

By interrogating the connection weights, as outlined by Shahin et al. (2002a), the following relationship is obtained from the optimal ANN model, to calculate the predicted settlement,S_p (mm):

$$S_p = 0.6 + \left[\frac{120.4}{1 + e^{(0.312 - 0.725\tanh x_1 + 2.984\tanh x_2)}}\right] \tag{3}$$

where:

$x_1 = 0.1 + 10^{-3}\left[3.8B + 0.7q + 4.1N - 1.8(L/B) + 19\left(D_f/B\right)\right]$

$x_2 = 10^{-3}\left[0.7 - 41B - 1.6q + 75N - 52(L/B) + 740\left(D_f/B\right)\right]$

Rezania and Javadi (2007) sought to improve the above ANN-based relationship, using genetic programing (GP). Selecting 173 cases of the 189 used by Shahin et al. (2002b), Rezania and Javadi (2007) proposed a more tractableand slightly more accurate model inEq. (4) than that presented in Eq. (3) above.

$$S_p = \frac{q(1.80B + 4.62) - 346.15D_f}{N^2} + \frac{11.22L - 11.11}{L} \tag{4}$$

Using Eq. (4), Rezania and Javadi (2007) reported values of $r = 0.978$; $RMSE = 6.86$ mm; and $MAE = 4.92$ mm, which for the ANN model in Eq. (3), compares against 0.960, 9.20 mm and 6.65 mm respectively, for the curtailed data set.

Adopting the process described by Shahin et al. (2002b), it is possible, using ANNs, to determine the sensitivity of the various parameters which contribute to the output. Shahin et al. (2002b) reported the parameters, ranked from highest to lowest relative importance, as: N, B, q, D_f/B, andL/B.

2.2 Pile foundations using ANNs

Pooya Nejad and Jaksa (2017) developed an ANN model to predict the axial load-settlement behavior of single piles embedded in a variety of ground types. The authors compiled a dataset containing 56 individual pile load tests. By digitizing these load-displacement curves, 499 pairs of loads and displacements were subsequently obtained. From these data, Pooya Nejad and Jaksa (2017) established a database consisting of the following 21 input parameters: (1) type of test (maintained load or constant rate penetration); (2) type of pile (concrete, steel, and composite); (3) type of installation (bored or driven); (4) end of pile (closed or open); (5) axial rigidity of the pile (EA); (6) cross-sectional area of the end of the pile (A_{tip}); (7) perimeter of the pile in contact with the soil (O); (8)length of the pile (L); (9) embedded length of the pile (L_{embed}), (10–19) the averaged cone penetration test (CPT)measurements of cone tip resistance (q_c) and sleeve friction (f_s) along the embedded length of the pile (q_{c_1}, f_{s_1}, q_{c_2}, f_{s_2}, q_{c_3}, f_{s_3}, q_{c_4}, f_{s_4}, q_{c_5},f_{s_5}); (20) cone tip resistance at the end of the pile ($q_{c_{\text{tip}}}$), which is calculated using the Bustamante and Gianeselli (1982) procedure;and (21) the applied load (P). The pile settlement (s_m) was the single output variable. The 10 CPT parameters ($q_{c_{1,2,\ldots5}}$, $f_{s_{1,2,\ldots5}}$) are obtained by dividing the embedded length of the pile into 5 layers of equal thickness. The values of q_c and f_s in each layer are then averaged to yield $q_{c_{1,2,\ldots5}}$ and$f_{s_{1,2,\ldots5}}$. The intention of this was to capture the fidelity of the vertical variability of the ground with an acceptable degree of accuracy,whilst at the same timeminimizing the number of soil layers.

Of the 499 sets of variables, each containing the 22 parameters above, 79% (395 cases, 8,690 data) were used for training, 11% (55 cases, 1,210 data) for testing, and 10% (49 cases, 1,078 data) for validation. The ANN model development was again undertaken using the *Neuframe* software package (Neusciences 2000). In order to obtain the optimal model, various ANN architectures were explored, consisting of one to four hidden layers. It was found that the optimal single-layer model consisted of 10 nodes; the optimal two-layer model consisted of 8 nodes in both the first and second layers; the optimal 3-layer model contained12 nodes in the first,6 in the second, and 3 in the third layer; and the optimal 4-layer model contained 16, 11, 5 and 2 nodes in hidden layers 1 to 4, respectively. Overall, the two-layer model was found to provide the greatest accuracy, with$r = 0.991$ and$RMSE$ =3.21 mm, which is shown in Figure 4 in the form of a scatterplot.

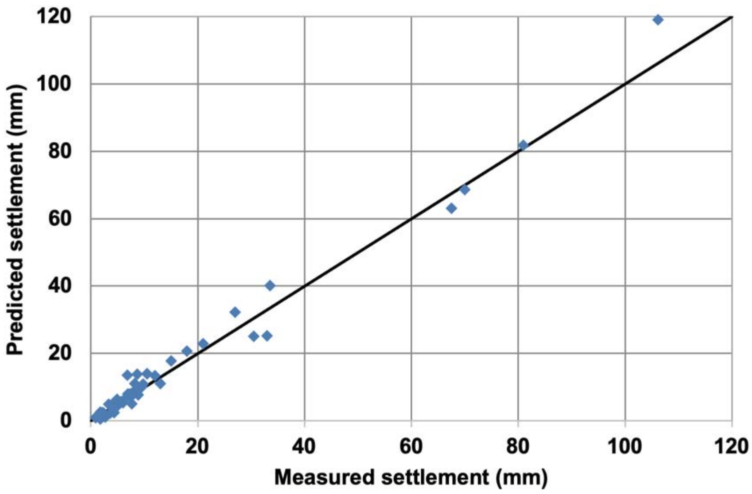
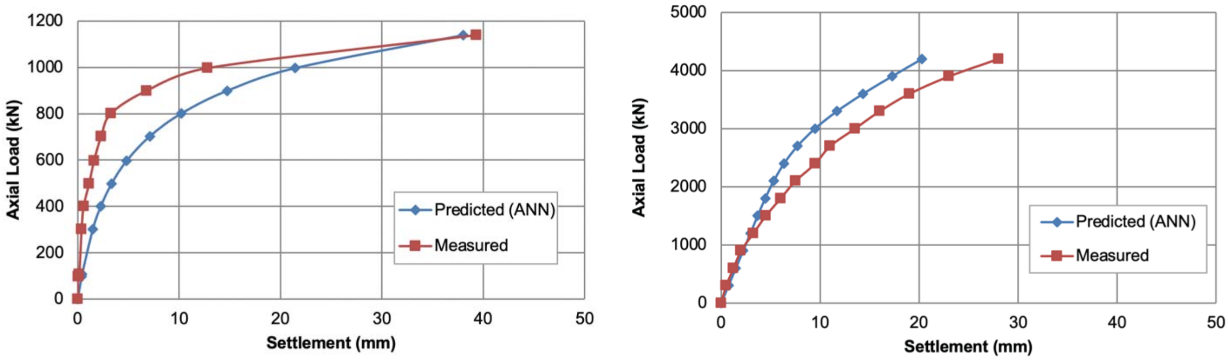


Figure 4. Predicted versus measured pile settlements (PooyaNejad and Jaksa 2017).

The ANN model was found to outperform significantly the non-AI methods of Vesic (1977) and Poulos and Davis (1980).In order to examine the validity of the ANN model further, it was tested against two completepile load tests that were not incorporated in the ANN model development. In other words, these data were not part of the 499 cases used to create the ANN model. The predicted load-settlement curves (shown in blue) are compared against those obtained from measurements (in red), as shown in Figure 5.



(a) (b)

Figure 5. Comparison of predicted load-settlement curves from two additional pile load tests: (a) A&M1-concrete pile (Briaudand Tucker 1988); (b) TWNTP4-steel pile (Yen et al. 1989).

It can be seen from Figure 5, that while not perfect, the ANN performs admirably. The goodness-of-fit results indicate that the model performs well for both the concrete pile, with $r = 0.956$ and$RMSE$ =4.39 mm, and the steel pile, with $r = 0.998$ and$RMSE$ =3.44 mm. For the concrete pile (Fig. 5a), an error of 3.2% is observed between the predicted (38.0 mm) and the measured (39.3 mm) settlements at the peak load of 1,140 kN. The steel pipe pile (Fig. 5b) performs less satisfactorily, with an error of 27.5% (7.7 mm) between the predicted (20.3 mm) and measured (28 mm) settlements associated with the peak load of 4,200 kN. Furthermore, it is evident from the predicted behavior that the ANN model has honored the non-linear shape of both of the load-settlement curves. Hence, it can be concluded from the above, that the ANN model provides a very good predictive capability of the pile behavior within the range of the database.

A sensitivity analysis performed on the input variables employed in the ANN model, found the relative importance as summarized in Table 3. As one might expect, the soil characteristics from the CPT, as well as the applied load, are the most critical parameters that influence pile settlement. Interestingly, the pile perimeter, O, which contributes to shaft adhesion, and the type of installation, whether bored or driven, are the least significant of the input parameters.

Table 3. Summary of relative importance (%) of the ANN input variables(Pooya Nejad and Jaksa 2017).

Input variables	Relative importance (%)
Soil properties ($q_{c_{1,2,\ldots5}}$, $f_{s_{1,2,\ldots5}}$, $q_{c_{\text{tip}}}$)	35.1
Applied load (P)	28.1
Embedded pile length (L_{embed})	5.6
End of pile (closed or open)	5.2
Pile length (L)	5.1
Type of pile	4.9
Type of test	4.4
Pile rigidity (EA)	3.8
Area of pile tip (A_{tip})	3.3
Type of installation	3.2
Perimeter of pile (O)	1.3

As the optimal ANN model incorporates two hidden layers, tractable relationships, like the ones given in Eqs. 3 and 4, are not readily available. To facilitate dissemination of the ANN model, Pooya Nejad and Jaksa (2017) provided a series of design charts and the complete pile load test database.

2.3 Rolling dynamic compaction using ANNs and GP

This section presents the application of AI to the rolling dynamic compaction (RDC)ground improvement technique.RDC consists of a non-circular module, which has either 3, 4 or 5 sides, that is towed by a tractor, as shown in Figure 6. As the module rolls, it tips about its corners and then falls impacting the ground with a combination of potential and kinetic energy. Whilst this technique has been adopted successfully in practice for more than 70 years, until recently, modest fundamental research has been carried out to model and predict its behavior. A significant challenge with RDC is predicting and optimizing its performance in a range of ground conditions. Recently, ANNs and GP have been adopted to examine their efficacy with respect to the 4-sided ‘impact roller’. These are described in turn below.

(a) (b) (c)

Figure 6. Rolling dynamic compaction modules: (a) 3-sided (Landpac); (b) 4-sided (Broons); (c) 5-sided (Landpac).

2.3.1 4-sided impact roller – ANNs

The performance of RDC is generally assessed by means of in situ tests, such as the CPT and the dynamic cone penetration test (DCP). Other techniques are also used, such as surface settlements, in situ density measurements, and geophysical methods. Since the mid-1980s, Broons (www.broons.com) has acquired a significant database,from a wide range of projects involving the use of the 4-sided, 8-tonneimpact roller module (Fig. 6c), which includes many in situ tests in a variety of soil types and ground conditions.These data were exploited to develop both ANN and GP models.

Ranasinghe et al. (2017a) developed a database incorporating 2,048 DCP records from 12 separate projects. These data were subsequently sub-divided into the training (64%), testing (16%) and validation (20%) sets. The database consisted of 5 input variables: (1) soil type, eitherSand–clay, Clay–silt, Sand–none,or Sand–gravel (indicating primary and secondary soil components); (2) average depth below the ground surface, D(m); (3) initial number of module passes; (4) initial DCP count (blows/300 mm); and (5) the final number of roller passes. The single output variable was the final DCP count (blows/300 mm) at depth D after compaction. After trial-and-error, the optimal ANN model consisted of 8 input nodes, a single hidden layer containing 4 nodes, and the single output variable node, as shown in Figure 7.It is important to recall that the DCP suffers from many uncertainties, and as such, is not a precise measurement instrument. In addition, the soil type parameter is also somewhat imprecise. With this in mind, as can be seen from Figure 8, the optimal ANN model performs reasonably well, with an $r = 0.79$, $RMSE = 7.54$ (blows/300 mm) and $MAE = 5.59$ with respect to the validation set.

As the optimal model is parsimonious in nature, again a tractable equation can be derived, as:

$$DCP_{\text{final}} = \frac{102.5}{1 + \exp\left(2.113\,T_9 + 2.307\,T_{10} + 3.725\,T_{11} + 3.163\,T_{12} - 2.269\right)} - 8.25 \tag{5}$$

where:

$T_9 = [1 + \exp\left(3.128I_1 + 5.257I_2 - 1.216I_3 + 0.973I_4 + 0.99I_5 + 0.04I_6 - 0.177I_7 - 0.006I_8 + 7.424\right)]^{-1}$

$T_{10} = [1 + \exp\left(2.291I_1 + 2.225I_2 + 3.206I_3 - 7.35I_4 + 0.985I_5 - 0.028I_6 + 0.165I_7 - 0.048I_8 + 0.059\right)]^{-1}$

$T_{11} = [1 + \exp\left(-0.082I_1 - 1.678I_2 + 0.014I_3 + 1.869I_4 - 1.302I_5 - 0.031I_6 + 0.017I_7 + 0.012I_8 + 0.757\right)]^{-1}$

$T_{12} = [1 + \exp\left(-1.486I_1 + 0.743I_2 - 1.482I_3 + 0.301I_4 + 1.749I_5 + 0.023I_6 + 0.053I_7 + 0.002I_8 - 0.517\right)]^{-1}$

and: I_i are the input variables, such that: I_1 to I_4 are binary values (0 or 1) dependent on the soil type (e.g. for a Sand–clay: $I_1 = 1$, $I_2 = 0$, $I_3 = 0$ and $I_4 = 0$); $I_5 =$ average depth, D (m); $I_6 =$ the initial of number roller passes; $I_7 =$ the initial DCP count (blows/300 mm); and $I_8 =$ the final number of roller passes.

A sensitivity analysis performed on the optimal ANN model indicated the importance of the input variables as follows (ranked most to least important): (1) soil type; (2) initial DCP count; (3) average depth; (4) initial number of roller passes; and (5) the final number of roller passes.

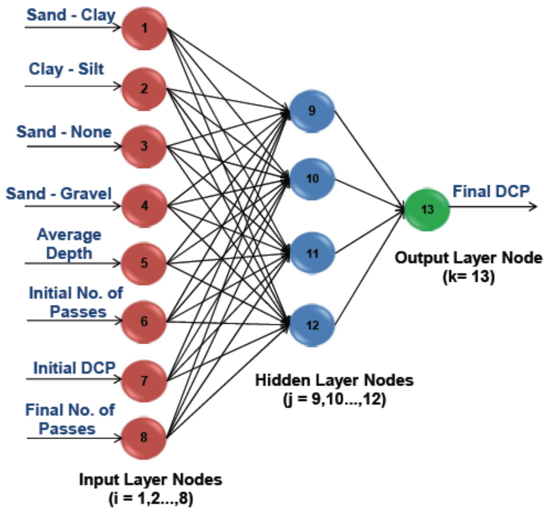


Figure 7. Optimal ANN model for 4-sided, 8t impact roller and dynamic cone penetration test (DCP) data (Ranasinghe et al. 2017a).

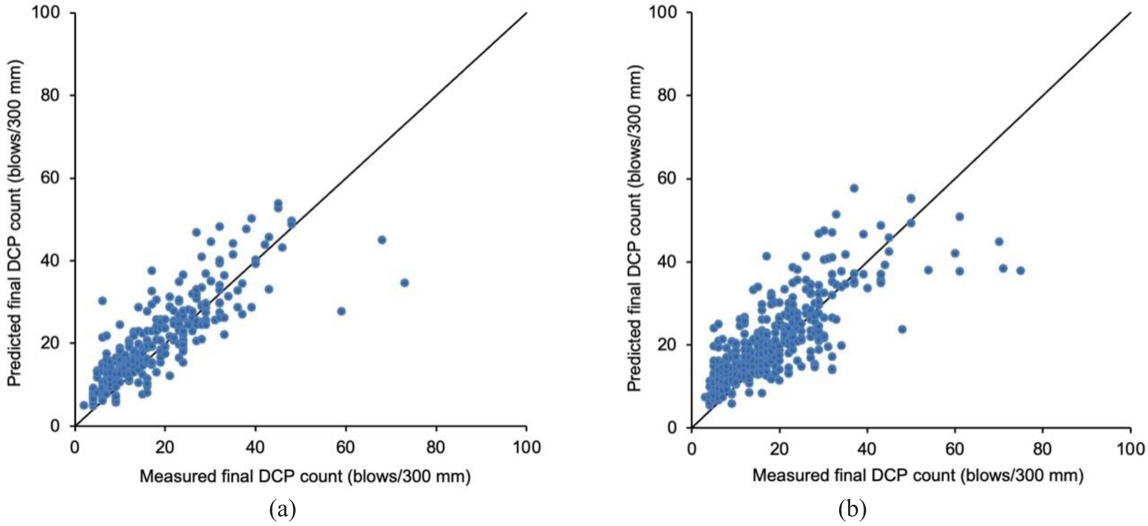


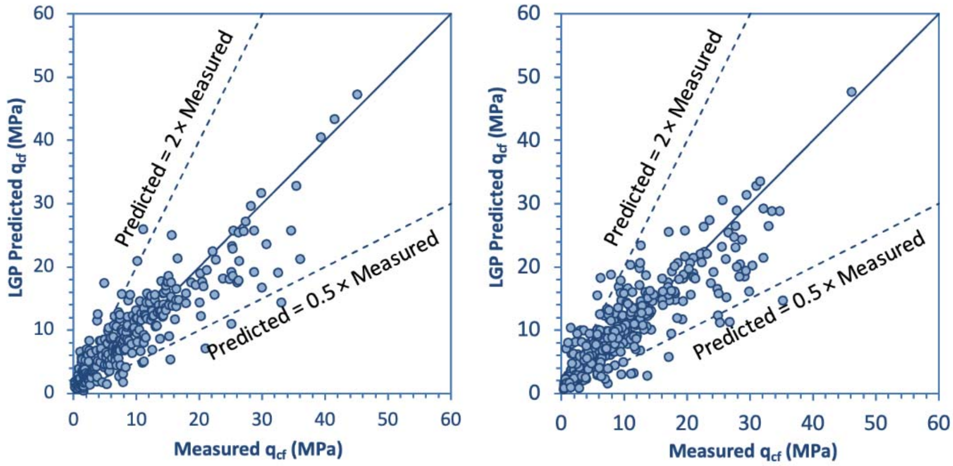
Figure 8. Performance of optimal ANN model with respect to the: (a) testing set; and (b) validation set (Ranasinghe et al. 2017a).

2.3.24-sided impact roller – GP

Ranasinghe et al. (2017b) used CPT measurements,from 1,977 data records from Broons' archive of RDC projects, to develop a genetic programing (GP) predictive model. The database consists of 103 CPT soundings, with CPT measurements up to 4 m in depth and sampled at 0.2 m vertical intervals. The GP model incorporated 4 input variables – the depth of the CPT measurement (D), the cone tip resistance (q_{ci}) and sleeve friction (f_{si}) prior to compaction, and the number of roller passes (P) – and a single output variable,the cone tip resistance after P roller passes (q_{cf}). Unlike the ANN model described in §2.3.1, the CPT model does not require a separate soil type input parameter, as this is provided indirectly through the well-established ratiobetweenf_{si}andq_{ci}, via the friction ratio.

In order to develop the GP model, the entire dataset wassubdivided into two sets: a training dataset (consisting of 1,755 records from 91 CPT soundings, i.e. 80% of the data) and a validation dataset (consisting of 222 records from 12 CPT soundings, i.e. 20% of the data). As with the ANN modeling described above, the training dataset was used to train and verify the GP model during the modeling phase, andthe validation dataset is withheld from the model, and deployed only after the model has been developed.

The commercial software suite *Discipulus* version 5.2 (Francone 2010) was used to develop and refine the GP model. Like *Neuframe* described above, *Discipulus*also seems to be no longer available. As can be seen from Figure 9, the GP model performs very well, as also indicated by $r = 0.87$, $RMSE = 4.03$ (MPa) and$MAE = 2.71$ with respect to the validation set. Using the same datasets, Ranasinghe et al. (2019a) developed an ANN model, and Ranasinghe et al. (2017b) showed that the GP model slightly outperformed its ANN counterpart.



(a) (b)

Figure 9. Performance of optimal GP model with respect to the: (a) testing set; and (b) validation set (Ranasinghe et al. 2017b).

In order to further validate the performance of the GP model and compare it against that of the ANN model by Ranasinghe et al. (2019a), Ranasinghe et al. (2017b) examined a new, additional dataset, unseen by either model. The results from a selection of the projects, involving complete CPT soundings, are shown in Figure 10. Whilst not perfect, the AI models perform reasonably well and provide useful for tools preliminary design purposes.

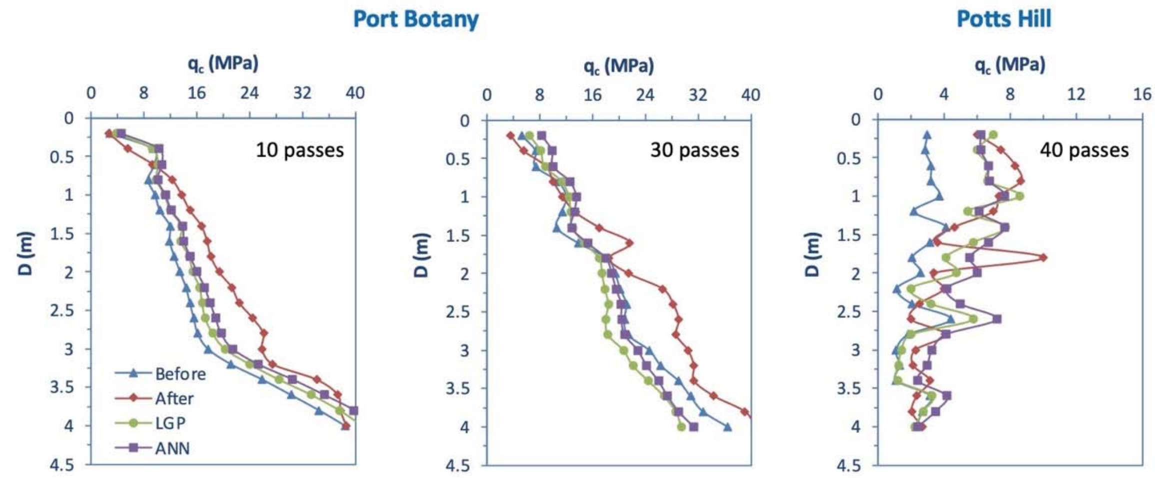


Figure 10. Performance of GP and ANN models on new, unseen data (Ranasinghe et al. 2017b).

So that the GP model can be disseminated, *Discipulus*incorporates afeature to export the final GP model as a C-program, and Ranasinghe et al. (2017b) included the C code for their GP–CPT model.

3 Benefits and Limitations of ANNs and GP

Like almost everything, artificial intelligence in geotechnical engineering has several benefits, but also a number of limitations. This section briefly explores both.

3.1 Benefits

As geotechnical engineering deals with materials which, by their very nature, can exhibit extreme variability, AIis particularly well-suited to modeling the complex behavior of these materials, and have generally demonstrated superior predictive performance when compared with traditional methods.As AI is data driven, it does not need prior knowledge about the nature of the relationship between model inputs and their corresponding outputs, as AIsolely use the data to capture this relationship. By examining 1,235 papers published since 1992 on AI applications to 9 sub-disciplines of geotechnical engineering, Baghbani et al. (2022) concluded that AI methods have yielded successful and promising results when solving geotechnical engineering problems. Shahin et al. (2009) concluded that AI is able to capture the subtle functional relationships among the presented data, even if the underlying relationships are unknown or the physical meaning is difficult to explain.Several researchers (e.g. Shahin 2003) stated that ANNs are able also to accommodate noisy data, to some degree, without degradation in predictive capability.

3.2 Limitations

The main criticisms that are levelled to AI are that they require large amounts of data, and they lack transparency and, hence, are often seen as black boxes. Further to these, Jaksa et al. (2008) and Shahin et al. (2009) discussed several issueswith ANNs that require attention in the future. These included: (i) ensuring that the developed models are robust; (ii) improving extrapolation ability; and (iii) dealing with uncertainty.

With regards to lack of transparency, this is often true, as the vast majority of papers presenting AI applications in geotechnical engineering do not provide access to the final AI model. Baghbani et al. (2022) stated that, despite the large number of studies undertaken, AI applications in geotechnical engineering have not been effectively transferred from research to practice. They suggest that one of the reasons for this is the lack of familiarity, and thereby confidence, within geotechnical engineering community to use AI based predictions. However, in the author's opinion, this is largely due to the lack of transparency. As has been discussed above, through the use of tractable relationships, such as those presented in Eq. (3)–(5), and the C code provided by Ranasinghe et al. (2017b), the AI models are transparent, relevant and accessible to the geotechnical engineering profession.

The issue of robustness is an important one and is often ignored in most of the published AI papers. Here, robustness is defined as the ability of the model to generalize over a range of data similar to that used for model training (Shahin et al. 2009). This is assessed by examining the predictions of the ranges of each of the input variables, as detailed by Shahin et al. (2005c). Authors of AI studies often only present measures of fitness, and subsequently conclude that their model is optimal. However, as Shahin et al. (2005c) highlight, this may not always be the case.

It is generally accepted that ANNs perform best when they are used to interpolate within the ranges of the input variables used for calibration (Tokar and Johnson 1999); that is, training. Whilst this is similar to other models, it is nevertheless an important limitation of ANNs, as it restricts their usefulness and applicability. Like other statistical techniques, machine learning models can be re-trained, and therefore improved, as new data become available.

Finally, a further limitation of ANNs is that the uncertainty in the predictions generated is seldom quantified (Maier and Dandy, 2000). Failure to account for such uncertainty makes it very difficult to assess the quality of the ANN predictions, which severely limits their efficacy. In an effort to address this, a few researchers have applied Bayesian techniques to ANN training. For example, Goh et al. (2005) observed that the integration of the Bayesian framework into the backpropagation algorithm enhanced the neural network prediction capabilities and provided assessment of the confidence associated with the network predictions. Shahin et al. (2005a, b) also incorporated uncertainty in the ANN process by developing a series of design charts expressing the reliability of settlement predictions for shallow foundations on sand.Since then, many researchers have applied Bayesian theory to AI (e.g. Hu et al. 2015, Tang et al. 2018, Hasanpour et al. 2020).

4 Conclusions

This paper haspresented the application of artificial neural networks (ANNs) and genetic programing (GP) to a number of geotechnical engineering problems and demonstrated their efficacy, as well as the processes for obtaining optimal models. Four examples have been presented: the prediction of the settlement of shallow foundations on sand using ANNs, the load-settlement behavior of pile foundations using ANNs, and the improvement in the ground due to rolling dynamic compaction using ANNs and GP. In each of these examples, the AI technique provided superior predictions to those of any other currently available method. In three of the four examples given, the optimal modelswere translated into, either a tractable equation, or as a C-language program, making them accessible for use in geotechnical engineering practice.

Lastly, the benefits and limitations of AI have been discussed. It is concluded that AI presents alternative and reliable methods for predicting geotechnical response in a wide range of ground conditions and applications.

Acknowledgments

This paper would not be possible but for the extensive work undertaken by the author's former PhD students: Prof. Mohamed Shahin, Dr. FereydoonPooyaNejad and Dr. Tharanga Ranasinghe. Their efforts are greatlyappreciated and acknowledged. The assistance of my colleague, Prof. Holger Maier, during our formative days with ANNs in the early 2000s, is also gratefully acknowledged.

References

Baghbani, A., Choudhury, T., Costa, S. and Reiner, J. (2022). Application of artificial intelligence in geotechnical engineering: A state-of-the-art review, *Earth-Science Reviews*, doi: 10.1016/j.earscirev.2022.103991.

Briaud, J.L. and Tucker, L.M. (1988). Measured and predicted axial capacity of 98 piles, *Journal Geotechnical Engineering*, 114(9), 984-1001.

Bustamante, M. and Gianeselli, L. (1982). Pile bearing capacity prediction by means of static penetrometer CPT. *Proc., 2nd European Symp. on Penetration Testing*, Amsterdam, 493-500.

Crevier, D. (1993). *AI: The Tumultuous Search for Artificial Intelligence*,BasicBooks, New York, USA.

Francone, F.D. (2010). *Discipulus with Notitia and solution analytics owner's manual*, Register Machine Learning Technologies Inc., Littleton, CO, USA.

Goh, A. (1994). Seismic Liquefaction Potential Assessed by Neural Networks. *Journal ofGeotechnical Engineering*, 120(9), 1467-1480.

Goh, A.T.C., Kulhawy, F.H. and Chua, C.G. (2005). Bayesian neural network analysis of undrained side resistance of drilled shafts, *Journal of Geotechnical and Geoenvironmental Engineering*, 131(1), 84-93.

Gribb, M. and Gribb, G. (1994). Use of neural networks for hydraulic conductivity determination in unsaturated soil. *Proc., 2nd Int. Conf. on Ground Water Ecology*.

Hasanpour, R., Rostami, J., Schmitt, J., Ozcelik, Y. and Sohrabian, B. (2020). Prediction of TBM jamming risk in squeezing grounds using Bayesian and artificial neural networks, *Journal of Rock Mechanics and Geotechnical Engineering*, 12(1), 21-31.

Hu, J.-L., Tang, X.-W. and Qiu, J.-N. (2015). A Bayesian network approach for predicting seismic liquefaction based on interpretive structural modeling, *Georisk*, 9(3), 200-217.

Jaksa, M.B., Maier, H.R. and Shahin, M.A. (2008). Future challenges for artificial neural network modelling in geotechnical engineering. *Proc., 12th Int. Assoc. for Computer Methods and Advances in Geotech. Engrg. Conf.*, Goa, 1710-1719.

Koza, J.R. (1992). *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, MIT Press, Cambridge, USA.

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.

MathWorks (2022). MATLAB, https://matlab.mathworks.com, accessed 31July 2022.

McClelland, J.L. and Rumelhart, D.E. (1988). *Explorations in Parallel Distributed Processing*, MIT Press, Cambridge, USA.

McCulloch, W. and Pitts, W. (1943). A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5(4), 115-133.

Meyerhof, G.G. (1965). Shallow foundations, *Journal of Soil Mechanics and Foundations Division*, 91(SM2), 21-31.

Neusciences (2000). Neuframe Version 4.0, Neusciences Corp., Southampton, Hampshire.

Pooya Nejad, F. and Jaksa, M.B. (2017). Load-settlement behavior modeling of single piles using artificial neural networks and CPT data, *Computers and Geotechnics*, 89, 9-21.

Poulos, H.G. and Davis, E.H. (1980). *Pile Foundation Analysis and Design*, Wiley, New York.

Ranasinghe, R.A.T.M., Jaksa, M.B., Kuo, Y.L. and Pooya Nejad, F. (2017a). Application of artificial neural networks for predicting the impact of rolling dynamic compaction using dynamic cone penetrometer test results, *Journal of Rock Mechanics and Geotechnical Engineering*, 9(2), 340-349.

Ranasinghe, R.A.T.M., Jaksa, M.B., Pooya Nejad, F. and Kuo, Y.L. (2017b). Predicting the effectiveness of rolling dynamic compaction using genetic programming and cone penetration test data. *Proc. of Institution of Civil Engineers – Ground Improvement*, 170(4), 193-207.

Ranasinghe, R.A.T.M., Jaksa, M.B., Kuo, Y.L. and Pooya Nejad, F. (2019a). Prediction of the effectiveness of rolling dynamic compaction using artificial neural networks and cone penetration test data, *Chinese Journal of Rock Mechanics and Engineering*, 38(1), 153-170.

Ranasinghe, R.A.T.M., Jaksa, M.B., Kuo, Y.L. and Pooya Nejad, F. (2019b). Genetic programming for predictions of effectiveness of rolling dynamic compaction with dynamic cone penetrometer test results, *Journal of Rock Mechanics and Geotechnical Engineering*, 11(4), 815-823.

Rezania, M. and Javadi, A.A. (2007). A new genetic programming model for predicting settlement of shallow foundations, *Canadian Geotechnical Journal*, 44, 1462-1473.

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, USA.

Schmertmann, J.H., Hartman, J.P. and Brown, P.B. (1978). Improved strain influence factor diagrams, *Journal of Geotechnical Engineering*, 104(GT8), 1131-1135.

Schultze, E. and Sherif, G. (1973). Prediction of settlements from evaluated settlement observations for sand. *Proc., 8th Int. Conf. on Soil Mechanics and Foundation Engineering*, 1(3), Moscow, 225-230.

Sette, S. and Boullart, L. (2001). Genetic programming: principles and applications, *Engineering Applications of Artificial Intelligence*, 14(6), 727-736.

Shahin, M.A. (2003). *Use of Artificial Neural Networks for Predicting Settlement of Shallow Foundations on Cohesionless Soils*, PhD Thesis, School of Civil and Environmental Engineering, University of Adelaide, Australia.

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., Jaksa, M.B. and Maier, H.R. (2005a). Neural network based stochastic design charts for settlement prediction, *Canadian Geotechnical Journal*, 42(1), 110-120.

Shahin, M.A., Jaksa, M.B. and Maier, H.R. (2005b). Stochastic simulation of settlement prediction of shallow foundations based on a deterministic artificial neural network model. *Proc. Int. Congress on Modelling and Simulation, MODSIM 2005*, Melbourne, 73-78.

Shahin, M.A., Jaksa, M.B. and Maier, H.R. (2009). Recent advances and future challenges for artificial neural systems in geotechnical engineering applications, *Advances in Artificial Neural Systems*, Article ID 308239, doi: 10.1155/2009/308239.

Shahin, M.A., Maier, H.R. and Jaksa, M.B. (2002b). Predicting settlements of shallow foundations using neural networks,*Journal of Geotechnical and Geoenvironmental Engineering*, 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*, 18(2), 105-114.

Shahin, M.A., Maier, H.R. and Jaksa, M.B. (2005c). Investigation into the robustness of artificial neural network models for a case study in civil engineering. *Proc. Int. Congress on Modelling and Simulation, MODSIM 2005*, Melbourne, 79-83.

Sterling, R.L. and Lee, C.A. (1992). A neural network-expert system hybrid approach for tunnel design. *Proc., 33rd US Symposium on Rock Mechanics*, ARMA-92-0501.

Tang, X.-W., Bai, X., Hu, J.-L. and Qiu, J.-N. (2018). Assessment of liquefaction-induced hazards using Bayesian networks based on standard penetration test data, *Natural Hazards and Earth Systems Sciences* 18(5), 1451-1468.

Tokar, S.A. and Johnson, P.A. (1999). Rainfall-runoff modeling using artificial neural networks *Journal of Hydrologic Engineering*, 4(3), 232-239.

Vesic, A.S. (1977). Design of pile foundations. National Cooperative Highway Research Program, *Synthesis of Practice No. 42*, Transportation Research Board, Washington, DC.

Yen, T.L., Chin, C.T. and Wang, R.F. (1989). Interpretation of instrumented driven steel pipe piles. In, *Foundation Engineering: Current Principles and Practices*, ASCE, Reston, USA, 1293-1308.