Geomechanics and Engineering, *Vol. 20, No. 5 (2020) 385-397*
DOI: https://doi.org/10.12989/gae.2020.20.5.385

Application of artificial neural networks in settlement prediction of shallow foundations on sandy soils

Nguyen-Vu Luat[1a], Kihak Lee[1b] and Duc-Kien Thai[*2]

[1]Department of Architectural Engineering, Sejong University, 98 Gunja-dong, Gwangjin-gu, Seoul, 173-147, South Korea
[2]Department of Civil and Environmental Engineering, Sejong University, 98 Gunja-dong, Gwangjin-gu, Seoul, 173-147, South Korea

(Received August 27, 2019, Revised February 5, 2020, Accepted February 6, 2020)

Abstract. This paper presents an application of artificial neural networks (ANNs) in settlement prediction of a foundation on sandy soil. In order to train the ANN model, a wide experimental database about settlement of foundations acquired from available literatures was collected. The data used in the ANNs model were arranged using the following five-input parameters that covered both geometrical foundation and sandy soil properties: breadth of foundation B, length to width L/B, embedment ratio D_f/B, foundation net applied pressure q_{net}, and average SPT blow count N. The backpropagation algorithm was implemented to develop an explicit predicting formulation. The settlement results are compared with the results of previous studies. The accuracy of the proposed formula proves that the ANNs method has a huge potential for predicting the settlement of foundations on sandy soils.

Keywords: neural networks; sandy soils; shallow foundation; settlement prediction; back propagation

1. Introduction

The estimation of foundation settlement is a basic and critical subject matter of foundation engineering and is a common procedure carried out by geotechnical engineers. Due to the heterogeneous nature of soil, the process of predicting settlement has become more tedious and complex. In general, sandy soil have a higher differential settlement compared to cohesive soils since the former is less homogeneous than the latter. Moreover, the deformation behavior of shallow foundations obtaining their support from granular soils (i.e., sands, gravels) mainly governs the final design of structures which are built on these soil types. Therefore, predicting settlement is a crucial issue and is one of the greatest concerns in foundation design codes.

Numerous researches have been published over the past five decades investigating the correlation of predicted settlements with the measured settlement of shallow foundations on cohesionless soils. Some of them proposed new methods to estimate the settlement (Terzaghi and Peck 1968, Schmertmann 1970, Jorden 1977), while the others made an effort to present a comparison among various methods to assess whether or not any one distinct method provided an upper-level accuracy over the others (Maail 1987, Maugeri *et al.* 1998). The reliability of settlement estimation for shallow foundations on sandy soils also received substantial attention in the recent years (Berardi *et al.* 1991, Sivakugan and Johnson 2004). Das and Sivakugan (2007) produced an overview and indicated that the probabilistic design chart method proposed by Sivakugan and Johnson (2004), can be utilized for making an estimate of the probability that the actual settlement would be exceed 25 mm in the field.

Due to the difficulty of obtaining undisturbed samples, the majority of the available methods for the settlement prediction of shallow foundations on sandy soils relied upon in-situ tests, such as the pressuremeter test, plate load test, dilatometer test, drive cone test, cone penetration test (CPT) and standard penetration test (SPT) Meyerhof (1956, 1965). The obtained results were either used: to estimate a soil elastic modulus for elastic deformation analysis; to directly predict settlement based on an empirical relationship; or to estimate other soil properties (i.e., over consolidation ratio, relative density) for the estimation of settlement. However, most of these methods attempted to simplify the problem by assuming a linear response between load and deformation, and related it to the factors that can affect the settlement value. Therefore, the previous methods failed in obtaining a good solution to predict the settlement accurately. The aforesaid reviews between predicted and measured settlement of shallow foundations in sandy soil indicated that no particular method was superior to the others in all cases and the calculated results of settlement were inconsistent. Consequently, there remains a need for a more efficient method that can provide settlement prediction results with higher accuracy. To obtain that expectation, a new approach using the advantage of artificial intelligence technique is employed in this research.

ANN models are computing systems designed to mimic how human brain works and contain a number of interconnected processing elements or neurons. These elements are able to be trained to map a specified input to

∗Corresponding author, Professor
E-mail: thaiduckien@sejong.ac.kr
[a]Ph.D. Student
[b]Professor

http://www.techno-press.org/?journal=gae&subpage=7
ISSN: 2005-307X (Print), 2092-6219 (Online)

obtain an expected output. In contradiction to traditional computing models which follow an obvious step-by-step procedure based on predefined rules, ANNs can learn by available data as examples and adjust the connection strength between different neurons to acquire desired outputs. These advantages make ANNs capable of learning complicated relationships between multi-dimensional data. Therefore, ANNs are frequently being used in numerous fields of science and technology, as well as in expanding various engineering applications (Shahin *et al.* 2002, González and Zapico 2008, Tsompanakis *et al.* 2008). The fundamentals of neural computing procedures in geotechnical analysis and design were first implemented by Goh (1994), and an increasing number of articles have been published over the years. Most of these studies focused on soil properties (static and dynamic) (Çelik and Tan 2005, García *et al.* 2006, Kamatchi *et al.* 2010, Erzin and Ecemis 2017), predicting behavior of foundations (Das and Basudhar 2006, Chaudhary *et al.* 2007, Shahin 2010, Alkroosh and Nikraz 2011, Dinçer 2011, Gandomi and Alavi, 2012), predicting liquefaction (Chern and Lee 2008, Samu and Thallak 2011), estimating slope stability and landslides (Cho 2009, Das *et al.* 2011), and applying in the others of geomechanics field (Javadi and Rezania 2009, Shivapullaiah *et al.* 2009, Alavi *et al.* 2010, Erz and Cetin 2014, Shahrbanouzadeh *et al.* 2015, Fei *et al.* 2019).

This study focused on the application of an ANN to propose a simple method of predicting the settlement of shallow foundation on sandy soil using several parameters. For this purpose, 180 experimental data from different technical literatures were collected. First, a simple regression analysis was performed to show the relationship between the settlement and the five-main variables: the breadth of foundation B, the shape factor L/B, the embedment ratio D_f/B, the net-applied load pressure at footing base q_{net}, and the average SPT blow count N. Afterwards, ANN method was applied for predicting the settlement of foundation on sand using the aforementioned experiment results. Then, an explicit solution of settlement of shallow foundation in terms of the mentioned input parameters was proposed by using the backpropagation algorithm. To evaluate the accuracy of the developed ANN, some classic empirical formulas and other ANN model Shahin *et al.* (2002) were selected and the results obtained from the different methods were compared to the results from the developed ANN.

2. Overview of artificial neural networks (ANNs)

ANN is an information processing model that mimics the function of the human brain which is the ability to acquire, represent, and create a desired mapping of information from multi-dimensional data to another with a set of data representing that mapping. This model consists of an interconnected assembly of simple processing elements called neurons, which are linked to each other. These artificial neurons are modeled on the biological model which is comprised four main parts: dendrite, axon, synapse, and cell body as illustrated in Fig. 1. The signals are transferred to the dendrites from the other neurons

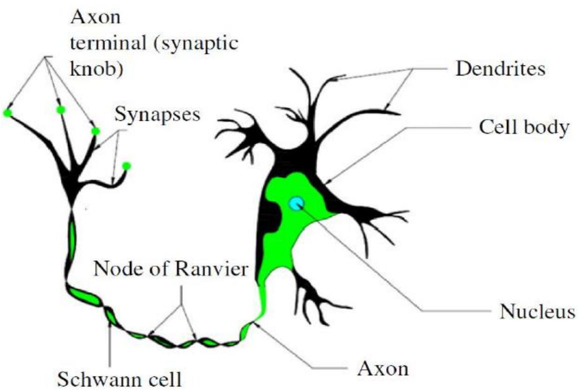


Fig. 1 The biological neurons

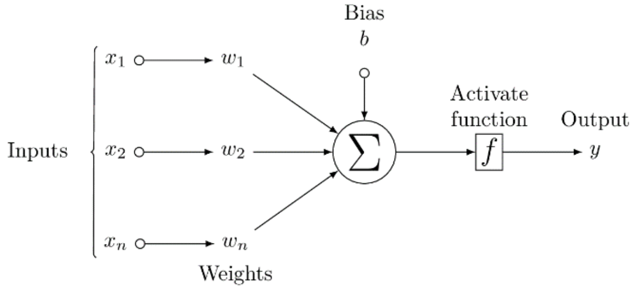


Fig. 2 The structure of artificial neuron

(cells). The axon of each neuron is of use to organize synaptic connections with other neurons. The cell body synthesizes the incoming signals from the dendrites'. The volumes of signal transferred relies on the synaptic strength of the connection, that means the synaptic integration, which determines whether a neuron becomes active or not. In order to become active, the input signals must reach a threshold level where the neuron will send a nerve impulse (spike) to its axon. In contrast, if the inputs are not sufficient to reach the required level, no excitation will occur.

The fundamental element of a neural network is the artificial neuron as presented in Fig. 2. The scalar input x is multiplied by the scalar weight w and added by the other input, called bias b, which is multiplied by 1. This summation, denoted z given in Eq.(1), is then transformed using an active function f, which produces the scalar neuron output y.

$$z_i = \sum_{j=1}^{n} w_{ij} x_j + b_i \tag{1}$$

An output value for any single neuron can be expressed as:

$$Y_i = f(z_i) \tag{2}$$

There are a variety of ANN models (Hertz *et al.* 1992) but the multi-layer perceptrons (MLPs) are broadly used in engineering applications. In general, an MLPs consists of three main components: an input layer, one or several hidden layers, and an output layer. Each layer has one or more neurons. The output of one neuron provides the input to the other neurons in the next layer. The hidden layers

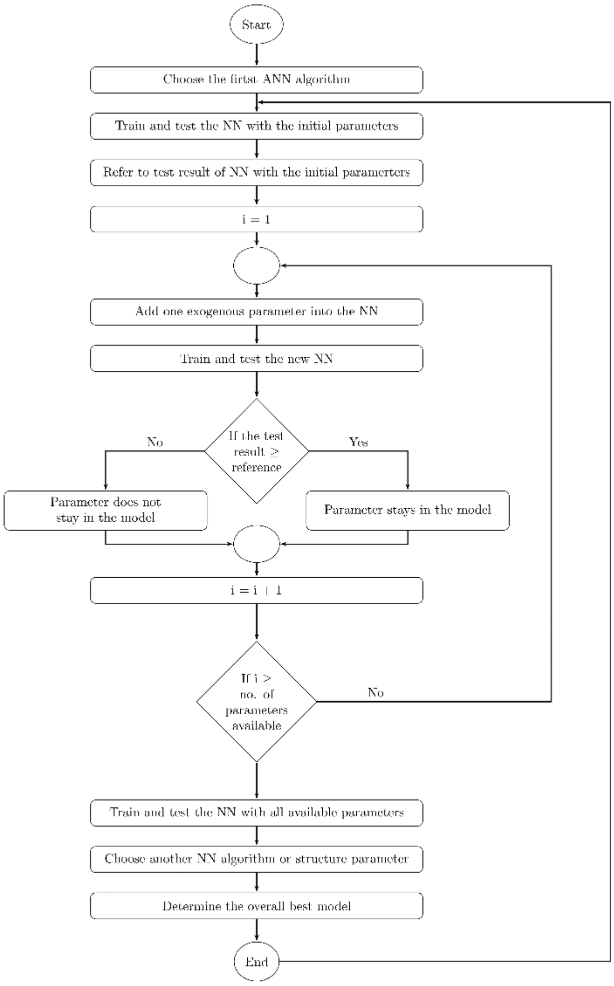


Fig. 3 The flowchart for selecting ANN structure

enable these networks to represent and calculate complex associations between input and output layers.

The number of hidden layers depend on the complexity of the problem. There are several methods to determine this number (Kaastra and Boyd 1996, Kanellopoulos and Wilkinson 1997, Alvarez Grima and Babuška 1999, Haque and Sudhakar 2002). Most of these methods are, however, not presented specifically. The neural network can "learn" or "train" from several learning algorithms. The most commonly used training algorithm engineering applications is the backpropagation algorithm which can be divided into three parts: (i) the feeding of data at input layers, (ii) the output calculation and its error backpropagation, and (iii) the weights modifications. This procedure can be implemented with various optimization strategies. The error of the ANN model's output and the actual label is transmitted backward and used to renew weights of the previous layers can be observed in Fig. 3.

In order to estimate whether the overall ANN is sound or not, several criteria can be considered such as: coefficient of determination R^2, root mean square error (*RMSE*), mean absolute error (*MAE*), minimal absolute error, and maximum absolute error. A well-trained model should show the result with R^2 close to 1 and small values of error terms. In this paper, the settlement prediction model for shallow foundation using ANN was built using *Python* code which is an open-source and powerful programming language with numerous available libraries (e.g., *keras, tensorflow, mxnet* and so on). The reliability of the model was estimated using the first three-aforementioned criteria including R^2, *RMSE* and *MAE*. These values were also utilized to compare with some empirical existing methods, which are presented in the following section.

3. Empirical classical method

There have been numerous researches which proposed prediction method for the settlement of shallow foundations on sandy soils. Three methods were chosen to compare and assess the reliability of the performance of the ANN model. The considered methods were performed the studies of Schultze and Sherif (1973), Meyerhof (1974), and Anagnostopoulos *et al.* (1991), all of which used data from SPT. These were chosen since these had been recently used in many design codes, and the ANN model had similar input variables.

3.1 Schultze and Sherif (1973)

Based on the results of a study by Schultze and Sherif (1973), wherein they measured settlements from 48 sites with SPT, they were able to suggest an empirical method to estimate the settlement of shallow foundations on sand. This settlement can be predicted from:

$$S_c = \frac{q_{net} F_c}{N^{0.87}\left(1 + 0.4\dfrac{D_f}{B}\right)} \tag{3}$$

3.2 Meyerholf (1974)

Meyerholf's most recent expressions for settlement were further modifications of his previous methods (Meyerhof 1956, 1965) and were generally considered to be conservative. In this case, when the adjustment for foundation embedment was taken into account, the settlement was given as:

$$S_c = \frac{1.33 q_{net}}{N}\left(1 - \frac{D_f}{B}\right) \quad \text{for } B \le 1.22\ m \tag{4}$$

$$S_c = \frac{1.33 q_{net}}{N}\left(1 - \frac{D_f}{B}\right) \quad \text{for } B > 1.22\ m \tag{5}$$

3.3 Anagnostopoulus (1991)

Anagnostopoulos *et al.* (1991) suggested another empirical method for grouping estimates according to stiffness, e.g., loose, medium or dense sand as well as small versus large footings. This method was based on a statistical evaluation of measured settlements and multiple regression analyses. of the case histories obtaining primarily from Schultze and Sherif (1973) and Burland and Burbidge (1985). Presumably one would estimate settlement by using the appropriate expression for both SPT blow count range and the appropriate footing width, and then average the two results to give a single settlement estimate. These formulae for predicting settlement can be expressed as:

$$S_c = \left(0.57\,q^{0.94}B^{0.90}\right)/N^{0.87} \quad \text{for } 0 < N \le 10 \tag{6}$$

$$S_c = \left(0.35\,q^{1.01}B^{0.69}\right)/N^{0.94} \quad \text{for } 10 < N \le 30 \tag{7}$$

$$S_c = \left(604\,q^{0.90}B^{0.76}\right)/N^{2.82} \quad \text{for } N > 30 \tag{8}$$

$$S_c = \left(1.90\,q^{0.77}B^{0.45}\right)/N^{1.08} \quad \text{for } B \le 3 \text{ m} \tag{9}$$

$$S_c = \left(1.64\,q^{1.02}B^{0.59}\right)/N^{1.37} \quad \text{for } B > 3 \text{ m} \tag{10}$$

where, in Eq. (3-10), S_c = settlement (*mm*); q_{net} = net applied pressure (kPa); B = footing width (m); D_f = depth of embedment (m); N = the blow count from SPT, were not corrected for overburden stress; σ'_c = pre-consolidation pressure (*kPa*); F_c = settlement coefficient (obtained from design chart).

4. ANN models and numerical application

4.1 ANN models and input parameters

The principal performance of an ANN model is controlled on the NN structure and parameters setting. One of the most important and difficult tasks in developing NN is to figure out the most optimal structure. It can be done by using trial and error approach to determine the number of hidden layers and the number of neurons in hidden layers. The universal approximation theorem states that MLP can approximate any continuous function with a single hidden layer containing a finite number of neurons. One of the first versions of this theorem was proved by Cybenko (1989) for sigmoid functions. Consequently, the NN model in this study was built with one hidden layer.

The experimental data used in this study cosists of 180 experimental shallow foundation was reported by various studies and given in **Appendix A**. The five variables were used as input parameters for the ANN model: (1) B; (2) L/B; (3) D_f/B; (4) q_{net}; (5) N while the output was selected as the measured settlement of shallows foundation S_c. The uncorrected N-values were used in predicting settlement, however, if the sand was dense, saturated and very fine or silty, Terzaghi and Peck (1968) recommended that the blow count applied to submerged case should be corrected according to:

$$N_c = 15 + 0.5(N - 15) \tag{11}$$

If the soil was gravelly sand or sandy gravel, a correction for N was recommended by Burland and Burbidge (1985) as:

$$N_c = 1.25N \tag{12}$$

Table 1 Statistic of the database used for ANN

Parameter used	Minimum	Maximum	Mean (μ)	Std. (σ)
Input parameters				
N_c	4.0	60	24.58	13.53
B (m)	0.8	255	9.82	20.48
L/B	1.0	10.6	2.19	1.80
D_f/B	0	3.44	0.53	0.58
q_{net} (kPa)	18.32	1532	194.31	157.35
Output parameters				
S_c (mm)	3.3	103.40	20.20	26.10

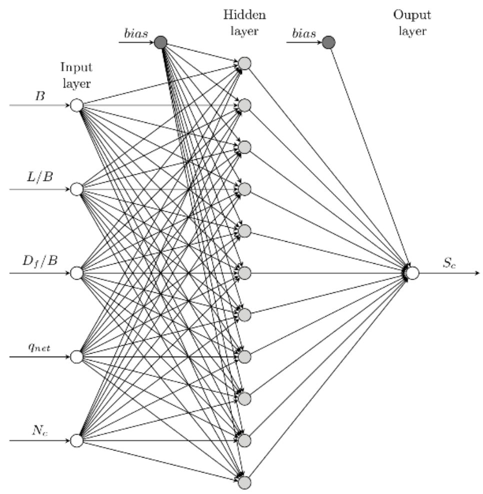


Fig. 4 The optimization process of selected NN

In total, 80% of the collected data (144 cases) were utilized for training and the remaining 20% for testing (36 cases). Statistical parameters of the data used for training and testing sets are presented in Table 1. Model input and output variables commonly have different dimensions and orders of magnitude. Thus, they need to be normalized to make training less sensitive to the scale of the input variables and to eliminate their dimensions (Luat *et al.* 2020). Moreover, normalization makes the problem better conditioned and prevents numerical difficulties during the calculation. Therefore, all variables are normalized to the range of [0, 1] using the min-max normalization method, which is expressed as follows:

$$x_{norm} = \frac{x - min(x)}{max(x) - min(x)} \tag{13}$$

where, x is the original value and x_{norm} is the normalized value.

4.2 Optimal ANN model

Although there is no rule for determining the number of nodes in a hidden layer in general, this number must be satisfactory for correcting the model of the problem, as well as should be small enough to ensure the simplification. In order to optimize the network, ANNs with multiple neurons in its hidden layer are trained in range 1 to 11 as discussed

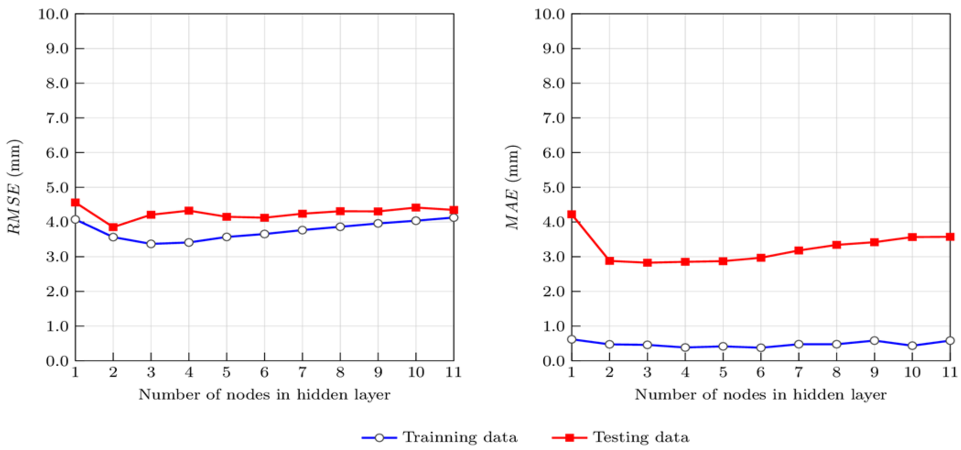


Fig. 5 The optimization process of selected NN

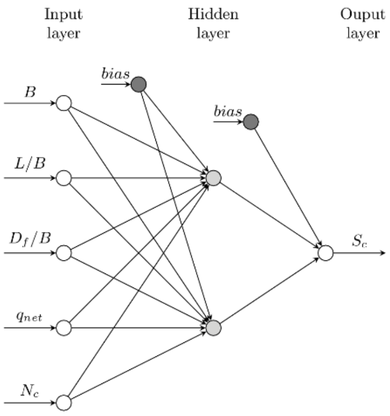
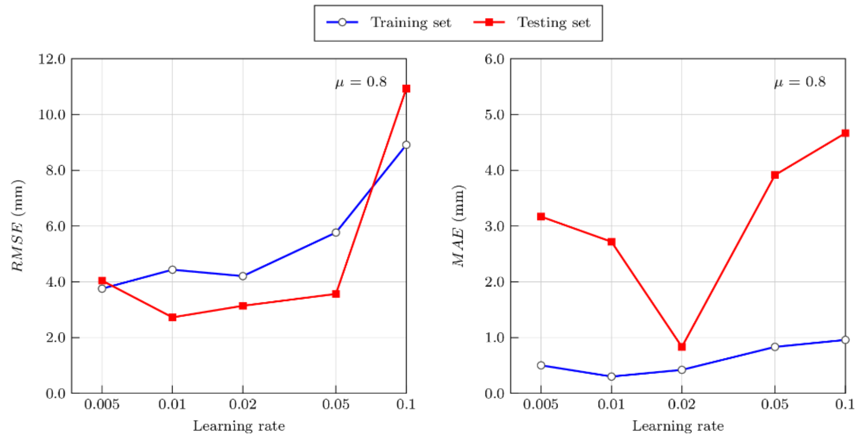


Fig. 6 The optimum NN structure

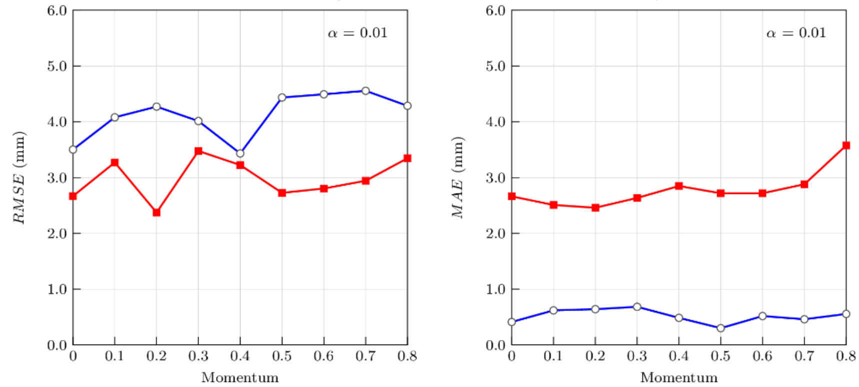
by Caudill (Caudill, 1988) for 1000 epochs, shown in Fig. 4. The performance of the NN is evaluated by utilizing the default sigmoid functions in the hidden layer units and a linear activation function in the output layer. An *RMSE* was used for the backpropagation with the efficient stochastic gradient decent (SGD) optimization algorithm while keeping the values of the learning rate and momentum factor equal to 0.01 and 0.8, respectively. The mini batches of data were shuffled for each epochs in order to serve the purpose of reducing variance and ensuring that NN model remained general and over-fitted less. Each result of the NN was tested over ten simulations with a different initializing random weights, then the average *RMSE* and *MAE* of training and testing set were obtained. It should be noted that in an ANN model, the performance on training data shows the learning process, while the obtained result of testing data demonstrates the model predictability. Therefore, the final model is chosen based on the testing data performance. As shown in Fig. 5, the two-node layer had the lowest error for the testing set in terms of *RMSE*, indicated the best performance. Moreover, the small distance of *RMSE* between the training and testing set showed a less over-fitting prediction. In terms of *MAE*, the smallest value yielded the three-node layer. However, it is noteworthy that one of the main tasks was to derive an explicit-predictive formula for practical design. The smaller the number of neurons is, the simpler the equation derives. For this reason, the ANN model with two neurons in the hidden layer was chosen as the optimal model. This selection was generally acceptable because the two-node layer had a difference of only 0.1 mm against the three hidden layer nodes in terms of *MAE*. In brief, the optimal NN structure was found to be 5-2-1, as displayed in Fig. 6.

In an attempt to develop an NN by combining transfer functions with optimization, some functions and optimization algorithm were used. Three functions including logistic sigmoid function, hyperbolic tangent (tanh) and rectified linear unit (ReLU) were applied to the activation functions in all units of hidden layers. The five backpropagation training algorithms used consisted of SGD, adaptive moment estimation (Adam) (Kingma and Ba 2015), Adadelta (Zeiler 2012), adaptive gradient algorithm (AdaGrad) (Duchi *et al.* 2011), and signum optimizer (Bernstein *et al.* 2018). These algorithms were governed by hyperparameters (e.g., learning rate - α, momentum - μ). Actually, all above algorithms have learning rate as a hyperparameter. Trial and error method was used to find out the best set of these factors. This meant that one factor was changed while the others remained constant. Based on the procedure of selecting the hyperparameters for SGD algorithm as shown in Fig. 7, the learning rate and momentum were chosen as 0.01 and 0.2 respectively. This process was also done for other algorithms.

The results showed the performance indices (*RMSE*, *MAE* and R^2) of the ANN with various activation functions and optimization algorithms. It was evident that overall, the sigmoid function performed the best in training and testing whereas the ReLU function had the lowest accuracy. Out of the five optimization algorithms, sigmoid function combined with signum algorithm presented the smallest *RMSE* of 3.22 mm for training set while for testing set, the algorithm that yielded the smallest value was Adam at 2.66 mm. This rule was also true in the case of tanh activation function with *RMSE* with 3.09 mm and 2.82 as training and testing data, respectively. The value of minimum *RMSE* object was altered based on the coefficient of determination R^2 for the testing results. The R^2 value reflected the input data participation in predicting the output value, and that the errors were decreased when predicting the settlement by using five input parameters. Table 2 also reports the



(a) Various learning rate values with momentum $\mu = 0.8$



(b) Various momentum values with learning rate $\alpha = 0.1$

Fig. 7 Performance of ANN model with sigmoid activation function

Table 2 The comparison of *RMSE, MAE,* R^2 for training and testing sets among various optimization algorithms and activation function

Opimization algorithm		SGD		Adam		Adadelta		Adagrad		Signum	
Activation function		Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
RMSE	Sigmoid	3.82	2.79	3.53	2.66	3.46	4.20	3.2	3.99	3.22	4.01
	Tanh	3.86	3.94	4.51	2.82	3.88	3.58	3.14	4.75	3.09	4.52
	ReLU	3.87	6.03	11.71	10.24	11.78	10.32	11.74	10.36	11.04	8.77
MAE	Sigmoid	0.30	2.60	0.28	2.54	0.35	2.73	0.38	3.00	0.39	3.09
	Tanh	0.46	4.04	0.34	2.94	0.39	2.81	0.36	2.97	0.33	3.05
	ReLU	0.49	5.17	1.35	8.78	1.44	9.12	1.45	9.09	1.38	7.75
R^2	Sigmoid	0.956	0.979	0.973	0.979	0.971	0.946	0.972	0.952	0.972	0.951
	Tanh	0.967	0.966	0.965	0.978	0.969	0.962	0.974	0.931	0.974	0.937
	ReLU	0.961	0.897	0.603	0.680	0.599	0.674	0.600	0.671	0.649	0.777

Table 3 Model parameters of the best ANN

Neuron	Weights						Bias	
	Input					Output		
	N_c	B	L/B	D_f/B	Q_{net}	S_c	Hidden layer	Output layer
1	-30.6090	21.5313	0.8972	1.1797	0.2728	79.4644	-3.2050	-4.1869
2	-21.5254	-7.7179	4.2135	4.5605	145.3265	13.8813	-4.1869	

comparison of R^2 with various combination of transfer functions and backpropagation algorithms. The combination of SGD and Adam algorithm with sigmoid function showed the best correlation with a value of 0.979. Thus, the optimal

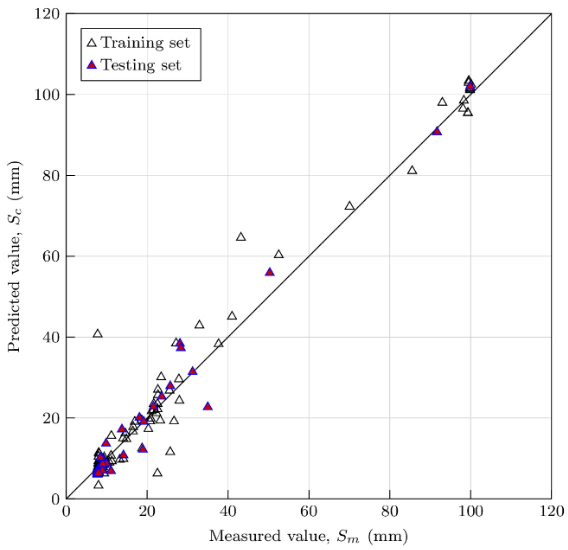


Fig. 8 NN prediction and actual values for training and testing set

ANN model chosen for predicting settlement was the sigmoid activation function associated with Adam optimization algorithm that turned out a learning rate of 0.01, which also presented the best MAE of 0.25 mm for training data and 2.54 mm for testing data. From the backpropagation algorithm results, the connection weights and biases of the final model are represented in Table 3. The relationship between the measured settlement and the predicted values which obtained from ANN model can be observed in Fig. 8. As seen in the figure, almost all of the data points showed good correlation between predicted and measured values.

5. Explicit formulation for predicting settlement

The main object is to derive an explicit formulation of settlement prediction as a function of input variables. It should be noted that, before training, all input variables were normalized using Eq. (14). Hence, the connection weights and biases of the optimal ANN (Table 3) was re-scaled before deriving the close-formed solution formulation. Finally, the settlement of foundation was expressed in terms of the blow count from SPT *Nc*, footing width (*B*), foundation geometry ratio (*L/B*), embedment ratio (D_f/B) and net applied pressure (q_{net}) as follows:

$$S_c = 79.4644 \times sig(Z_1) + 13.8813 \times sig(Z_2) - 4.1869 \tag{14}$$

where, $sig(Z) = \frac{1}{1+e^{-Z}}$, and:

$$\begin{aligned} Z_1 = &-0.2883 \times N + 0.2028 \times B + 0.084 \times (L/B) \\ &+ 0.0111 \times (D_f/B) + 0.0026 \times q_{net} - 0.2119 \end{aligned} \tag{15}$$

$$\begin{aligned} Z_2 = &-0.2027 \times N + 0.0727 \times B + 0.0397 \times (L/B) \\ &+ 0.0429 \times (D_f/B) + 1.3686 \times q_{net} - 59.2171 \end{aligned} \tag{16}$$

It should be noted that the closed-form solution derived above is valid for the ranges of variables given in Table 1.

6. Performance evaluation of the proposed ANN model

For the purpose of examining the accuracy of the ANN model, the model was compared against a few traditional methods as presented in Section 3. In addition, the ANN model results proposed by Shahin *et al.* (2002) were also used to compare with the results obtained from the present work. Some considered indices needed for each method and their equations are summarized in Table 4. The performance of the ANN model and the four considered methods, in relation to the testing set, are also displayed graphically in Fig. 9. According to Table 4, the ANN model performed reasonably well over the full range of settlement predictions with a high coefficient of determination between the predicted and measured values of 0.973. Table 4 also presents the lowest *RMSE* and MAE obtained from ANN model of this study. The values of *RMSE* and MAE obtained using Adam optimization were 3.53 and 0.28 mm, respectively. By contrast, the *RMSE* and MAE of considered methods ranged from 11.04 to 25.72 mm and from 8.78 and 16.59 mm, respectively. As seen in Fig. 9, the ANN model demonstrated a better prediction with less scatter with respect to the best fit line than those obtained from other methods, when the range of testing set was analyze. For instance, the method of Schultze and Sherif (1973) had a tendency to under-predict the settlement of

shallow foundation, especially for larger (than 40 mm) settlement, with a lowest coefficient of determination of 0.16 as listed in Table 4.

Three field tests were selected for verification of Eq. (14) by estimating and comparing settlement of shallow foundation. A 3×3 m footing (Exp No. 1) was implemented at Riverside Campus of A & M Texas University by Briaud and Gibbens (1994). Two underlain shallow foundation of two buildings (D2 and E1 – Exp No.2 and 3) located in Mascali, Italy were examined by Maugeri *et al.* (1998). This comparison is presented in Table 5. It is evident that the proposed equation provides the best accurate results compared to other methods. Moreover, the most accurate calculation out of traditional methods are at least twice as much as the measured settlement.

Table 4 Criteria indices and proposed equation for predicting

Method	Performance indices			Proposed equation
	R^2 (mm)	*RMSE* (mm)	*MSE* (mm)	
Schultze & Sherif (1973)	0.16	25.72	16.59	Eq. (3)
Meyerhof (1974)	0.518	23.55	11.81	Eq. (4-5)
Anagnostopoulos (1991)	0.768	15.02	9.91	Eq. (6-10)
Shahin *et al.* (2002)	0.819	11.04	8.78	-
This study	0.979	2.66	3.53	Eq. (14)

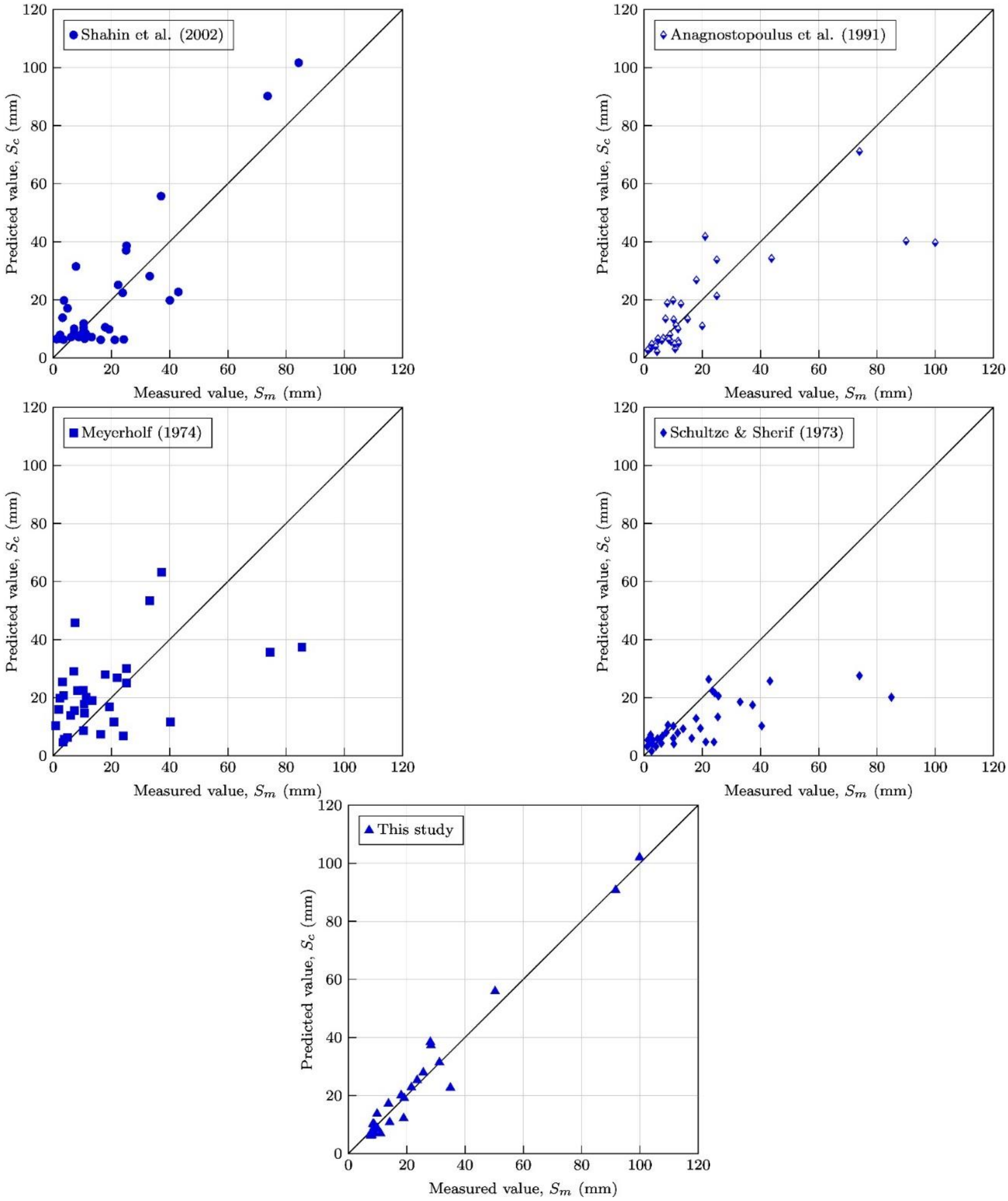


Fig. 9 Comparison measured versus predicted settlements between the proposed ANN model and available methods.

7. Conclusions

This paper presented an ANN model for the prediction of shallow foundation settlements on sandy soils. To propose this model, a database containing 180 experimental cases was collected from the available literature. The backpropagation algorithm was used to derive an explicit formula that could make sense in actual designs. Based on the discussion above and comparisons of the proposed model with available methods, some conclusions can be given, as follows:

• The optimal model for ANNs was associated with the number of hidden layers and the number of neurons in hidden layers which was found to be 5-2-1. The activation function used for the hidden layer was a logistic sigmoid function whereas the output layer was applied by a linear regression function.

• The obtained results from the proposed ANN model performed a good agreement with experimental results where its measured errors were low. These values illustrated that the neural networks had a capability to estimate the settlement of shallow foundation upon the sandy soils ranging from 3.3 to 103.4 mm.

• It was found that the ANN model can be efficiently utilized to develop an empirical equation for predicting the settlement of shallow foundation on sandy soils using the standard penetration tests. Moreover, the derived formulation can be employed as a handy prediction tool with satisfactory predictability.

• The range of applicability of the derived equation is constrained by the data used in the model. Consequently, for cases in which the input variable values are beyond this range, the proposed ANN model should be used with caution.

• In spite of the current research was able to do good a prediction, neither a parametric study nor variable importance was conducted. Therefore, future work should focus on the sensitivity analysis to determine the most and least important input variables. Moreover, another soft-computing techniques, such as multivariate adaptive regression splines (Luat *et al.* 2020), gradient tree machine (Thai *et al.* 2019), etc. can be further explored with nature-inspired evolutionary algorithm such as Genetic Algorithm, where the influence of each parameter required to predict the settlement of shallow foundation can help avoid the manual trial and error procedures.

Acknowledgments

The research described in this paper was financially supported by Ministry of Land, Infrastructure and Transport of Korean Government (Grant 20CTAP-C143093-03).

References

Alavi, A.H., Gandomi, A.H., Mousavi, M. and Mollahasani, A. (2010), "High-precision modeling of uplift capacity of suction caissons using a hybrid computational method", *Geomech. Eng.*, **2**(4), 253-280. https://doi.org/10.12989/gae.2010.2.4.253.

Alkroosh, I. and Nikraz, H. (2011), "Simulating pile load-settlement behavior from CPT data using intelligent computing", *Centr. Eur. J. Eng.*, **1**(3), 295-305. https://doi.org/10.2478/s13531-011-0029-2.

Alvarez Grima, M. and Babuška, R. (1999), "Fuzzy model for the prediction of unconfined compressive strength of rock samples", *Int. J. Rock Mech. Min. Sci.*, **36**(3), 339-349. https://doi.org/10.1016/S0148-9062(99)00007-8.

Anagnostopoulos, A.G., Papadopoulos, B.P. and Kavvadas, M.J. (1991), "Direct estimation of settlements on sand, based on SPT results", *Proceeding of the 10th European Conference on Soil Mechacnics and Foundation Engineering*, Florence, Italy, May.

Berardi, R., Jamiolkowski, M. and Lancellotta, R. (1991), "Settlement of shallow foundations in sand selection of stiffness on the basis of penetration resistance", *Proceedings of the Geotechnical Engineering Congress*, Boulder, Colorado, U.S.A., June.

Bernstein, J., Wang, Y.X., Azizzadenesheli, K. and Anandkumar, A. (2018), "signSGD: Compressed optimisation for non-convex problems", *arXiv preprint arXiv:1802.04434*.

Briaud, J.L. and Gibbens, R.M. (1994), "Predicted and measured behavior of five spread footings on sand", *Geotech. Spec. Publ.*, **41**, 255.

Burbidge, M.C. (1982), "A case study review of settlement on granular soil", M.Sc. Thesis, Imperial College of Science and Technology, University of London, London, U.K.

Burland, J.B. and Burbidge, M.C. (1985), "Settlement of foundations on sand and gravel", *Proc. Inst. Civ. Eng.*, **78**(6), 1325-1381. https://doi.org/10.1680/iicep.1985.1058.

Çelik, S. and Tan, Ö. (2005), "Determination of preconsolidation pressure with artificial neural network", *Civ. Eng. Environ. Syst.*, **22**(4), 217-231. https://doi.org/10.1080/10286600500383923.

Chaudhary, S., Pendharkar, U. and Nagpal, A.K. (2007), "Bending moment prediction for continuouscomposite beams by neural networks", *Adv. Struct. Eng.*, **10**(4), 439-454. https://doi.org/10.1260/136943307783239390.

Chern, S.G. and Lee, C.Y. (2008), "CPT-based simplified liquefaction assessment by using fuzzy-neural network", *J. Mar. Sci. Technol.*, **16**(2), 139-148.

Cho, S.E. (2009), "Probabilistic stability analyses of slopes using the ANN-based response surface", *Comput. Geotech.*, **36**(5), 787-797. https://doi.org/10.1016/ j.compgeo.2009.01.003.

Cybenko, G. (1989), "Approximations by super positions of sigmoidal functions", *Math. Control. Signals Syst.*, **2**(4), 303-314. https://doi.org/10.1007/bf02551274.

Das, B. and Sivakugan, N. (2007), "Settlements of shallow foundations on granular soil — an overview", *Int. J. Geotech. Eng.*, **1**(1), 19-29. https://doi.org/10.3328/ 10.3328/IJGE.2007.01.01.19-29.

Das, S.K. and Basudhar, P.K. (2006), "Undrained lateral load capacity of piles in clay using artificial neural network", *Comput. Geotech.*, **33**(8), 454-459. https://doi.org/10.1016/j.compgeo.2006.08.006.

Das, S.K., Biswal, R.K., Sivakugan, N. and Das, B. (2011), "Classification of slopes and prediction of factor of safety using differential evolution neural networks", *Environ. Earth Sci.*, **64**(1), 201-210. https://doi.org/ 10.1007/s12665-010-0839-1.

Dinçer, İ. (2011), "Models to predict the deformation modulus and the coefficient of subgrade reaction for earth filling structures", *Adv. Eng. Softw.*, **42**(4), 160-171. https://doi.org/10.1016/j.advengsoft.2011.02.001.

Duchi, J., Hazan, E. and Singer, Y. (2011), "Adaptive subgradient methods for online learning and stochastic optimization", *J. Mach. Learn. Res.*, **12**, 2121-2159.

Erzin, Y. and Cetin, T. (2014), "The prediction of the critical factor of safety of homogeneous finite slopes subjected to earthquake forces using neural networks and multiple regressions", *Geomech. Eng.*, **6**(1), 1-15. https://doi.org/10.12989/gae.2014.6.1.001.

Erzin, Y. and Ecemis, N. (2017), "The use of neural networks for the prediction of cone penetration resistance of silty sands", *Neural Comput. Appl.*, **28**(1), 727-736. https://doi.org/10.1007/s00521-016-2371-z.

Fei, S., Tan, X., Wang, X., Du, L. and Sun, Z. (2019), "Evaluation of soil spatial variability by micro - structure simulation", *Geomech. Eng.*, **17**(6), 565-572. https://doi.org/10.12989/gae.2019.17.6.565.

Gandomi, A.H. and Alavi, A.H. (2012), "A new multi-gene genetic programming approach to nonlinear system modeling. Part I: materials and structural engineering problems", *Neural Comput. Appl.*, **21**(1), 171-187. https://doi.org/10.1007/s00521-011-0734-z.

García, S.R., Romo, M.P. and Figueroa-Nazuno, J. (2006), "Soil dynamic properties determination: A neurofuzzy system approach", *Control Intell. Syst.*, **34**(1), 2121-2159.

Goh, A. (1994), "Seismic liquefaction potential assessed by neural networks", *J. Geotech. Eng.*, **120**(9), 1467-1480. https://doi.org/10.1061/(ASCE)0733-9410(1994)120:9(1467).

González, M.P. and Zapico, J.L. (2008), "Seismic damage identification in buildings using neural networks and modal data", *Comput. Struct.*, **86**(3), 416-426. https://doi.org/10.1016/j.compstruc.2007.02.021.

Haque, M.E. and Sudhakar, K.V. (2002), "ANN back-propagation prediction model for fracture toughness in microalloy steel", *Int. J. Fatigue*, **24**(9), 1003-1010. https://doi.org/10.1016/S0142-

1123(01)00207-9.

Hertz, J., Krogh, A. and Palmer, R.G. (1992), *Introduction to the Theory of Neural Computation*, Addison-Wesley Pub. Co., Redwood City, California, U.S.A.

Javadi, A.A. and Rezania, M., (2009), “Applications of artificial intelligence and data mining techniques in soil modeling”, *Geomech. Eng.*, **1**(1), 53–74. https://doi.org/10.12989/gae.2009.1.1.053.

Jorden, E., (1977), “Settlement in sand-methods of calculating and factors affecting”, *Ground Eng.*, **10**(1), 30-67. https://doi.org/10.1139/t06-029.

Kaastra, I. and Boyd, M. (1996), “Designing a neural network for forecasting financial and economic time series”, *Neurocomputing*, **10**(3), 215-236. https://doi.org/ 10.1016/0925-2312(95)00039-9.

Kamatchi, P., Rajasankar, J., Ramana, G.V. and Nagpal, A.K. (2010), “A neural network based methodology to predict site-specific spectral acceleration values”, *Earthq. Eng. Eng. Vib.*, 9(4), 459-472. https://doi.org/10.1007/ s11803-010-0041-1.

Kanellopoulos, I. and Wilkinson, G.G. (1997), “Strategies and best practice for neural network image classification”, *Int. J. Remote Sens.*, **18**(4), 711-725. https://doi.org/ 10.1080/014311697218719.

Kingma, D. and Ba, J. (2015), “Adam: A method for stochastic optimization”, *arXiv preprint arXiv:1412.6980*.

Luat, N.V., Lee, J., Lee, D.H. and Lee, K. (2020), “GS - MARS method for predicting the ultimate load - carrying capacity of rectangular CFST columns under eccentric loading”, *Comput. Concrete*, **25**(1), 1-14. https://doi.org/10.12989/cac.2020.25.1.001.

Maail, S. (1987), “Comparion of methods of predicting foundation settlement of sand and gravel”, *Proceedings of the 9th Southeast Asian Geotechnical Conference*, Bangkok, Thailand, December.

Maugeri, M., Castelli, F., Massimino, M.R. and Verona, G. (1998), “Observed and computed settlements of twoshallow foundations on sand”, *J. Geotech. Geoenviron. Eng.*, **124**(7), 595-605. https://doi.org/10.1061/(ASCE)1090-0241(1998)124:7(595).

Meyerhof, G.G. (1956), “Penetration tests and bearing capacity of cohesionless soils”, *J. Soil Mech. Div.*, **82**(1), 1-12. https://doi.org/10.1061/(ASCE)1090-0241(2006)132:4(478).

Meyerhof, G.G. (1965), “Shallow foundations”, *J. Soil Mech. Found. Div.*, **91**(2), 21-32. https://doi.org/10.1061/(ASCE)1090-0241(2006)132:4(478).

Meyerhof, G.G. (1974), “General report: State-of-the-art of penetration testing in countries outside Europe”, *Proceedings of the 1st European Symposium on Penetration Testing*, Stockholm, Sweden.

Samu, P. and Thallak, S. (2011), “Determination of liquefaction susceptibility of soil based on field test and artificial intelligence”, *Int. J. Earth Sci.*, **4**(2), 216-222.

Schmertmann, J.H. (1970), “Static cone to compute static settlement over sand”, *J. Soil Mech. Found. Div.*, **96**(3), 1011-1043.

Schultze, E. and Sherif, G. (1973), “Prediction of settlements from evaluated settlement observation for sand”, *Proceedings of the 8th International Conference on Soil Mechanics and Foundation Engineering*, Moscow, Russia, August.

Shahin, M.A., Maier, H.R. and Jaksa, M.B. (2002), “Predicting settlement of shallow foundations using neural networks”, *J. Geotech. Geoenviron. Eng.*, **128**(9), 785-793. https://doi.org/10.1061/(ASCE)1090-0241(2002)128:9(785).

Shahin, Mohamed, A. (2010), “Intelligent computing for modeling axial capacity of pile foundations”, *Can. Geotech. J.*, **47**(2), 230-243. https://doi.org/10.1139/T09-094.

Shahrbanouzadeh, M., Barani, G.A. and Shojaee, S. (2015), “Analysis of flow through dam foundation by FEM and ANN models Case study: Shahid Abbaspour Dam”, *Geomech. Eng.*, **9**(4), 465-481. https://doi.org/10.12989/gae.2015.9.4.465.

Sivakugan, N. and Johnson, K. (2004), “Settlement predictions in granular soils: a probabilistic approach”, *Géotechnique*, **54**(7), 499-502. https://doi.org/10.1680/geot.2004.54.7.499.

Sivapullaiah, P.V., Guru Prasad, B. and Allam, M.M. (2009), “Modeling sulfuric acid induced swell in carbonate clays using artificial neural networks”, *Geomech. Eng.*, **1**(4), 307-321. https://doi.org/10.12989/gae.2009.1.4.307.

Terzaghi, K. and Peck, R.B. (1968), *Soil Mechanics in Engineering Practice*, John Wiley & Sons, New York, U.S.A.

Thai, D.K., Tu, T.M., Bui, T.Q. and Bui, T.T. (2019), “Gradient tree boosting machine learning on predicting the failure modes of the RC panels under impact loads”, *Eng. Comput.*, 1-12. https://doi.org/10.1007/s00366-019-00842-w.

Tsompanakis, Y., Lagaros, N.D. and Stavroulakis, G.E. (2008), “Soft computing techniques in parameter identification and probabilistic seismic analysis of structures”, *Adv. Eng. Softw.*, **39**(7), 612-624. https://doi.org/10.1016/j.advengsoft.2007.06.004.

Wahls, H.E. (1997), “Settlement of shallow foundations in sands-selection of stiffness on the basis of penetration resistance”, *Proceeding of the 3rd International Geotechnical Engineerign Conference*, Cairo, Egypt.

Zeiler, M.D. (2012), “ADADELTA: An adaptive learning rate method”, *arXiv preprint arXiv:1212.5701*.

CC

Appendix

Table A Experiment database used for training and testing

Reference	B (m)	q_{net} (kPa)	N_{SPT}	L/B	D_f/B	S_m (mm)
Burland and Burbidge (1982, 1985)	15	0.8	1	0.00	78	7
	50	2.1	1	0.71	697	2.3
	60	2.5	3.8	1.20	284	1
	25	1.2	10.6	0.25	250	10
	20	0.9	1	3.44	300	6.7
	45	1.2	1	0.50	150	0.6
	35	4.5	1.3	0.67	195	3.9
	35	5.5	2.9	0.52	93	6.5
	20	4.3	1.6	0.49	161	5
	12	4.5	6.8	0.60	91	11
	35	15.0	4.9	0.20	81	5.4
	20	4.9	1.6	0.47	188	15
	20	4.0	1.6	0.50	145	7.4
	13	1.5	1	0.80	77	2.1
	45	1.0	1	0.50	284	4.7
	40	3.3	1.7	0.90	304	11.6
	17	12.2	1	0.09	130	22
	21	6.7	1.6	0.51	113	5
	6	25.0	1	0.11	75	87
	7	1.2	1.6	0.17	199	13
	20	4.3	1	0.49	102	7.1
	30	21.7	5.3	0.14	148	19.8
	34	1.0	1	0.00	220	3.6
	16	2.5	1.1	0.00	245	11
	37	4.0	1	1.30	512	12.8
	13	1.5	1	0.80	77	1.3
	28	3.6	4.4	0.00	193	18
	6	14.5	1	0.07	74	75
	20	4.0	1.6	0.50	225	9.1
	20	6.4	1.6	0.50	150	14.5
	20	4.3	1.6	0.49	138	7.1
	20	4.9	1.6	0.47	123	6.6
	50	1.2	1	0.42	300	4.5
	20	4.9	1.6	0.47	107	3.6
	20	3.7	1	1.40	135	10.1
	6	22.4	3.8	0.04	64	70
	20	4.3	1	1.20	134	15.4
	20	4.9	1.6	0.47	97	4.3
	20	4.3	1.6	0.49	150	6.8
	40	1.0	1	0.00	294	5
	21	22.0	1	0.23	79	10.5
	6	25.0	1	0.09	75	87
	19	33.5	1	0.00	156	90

Table A Continued

Reference	B (m)	q_{net} (kPa)	N_{SPT}	L/B	D_f/B	S_m (mm)
Burland and Burbidge (1982, 1985)	18	1.2	1	2.20	215	8.6
	6	22.4	3.8	0.04	75	92
	10	2.6	8.5	0.77	147	12
	40	4.5	1.5	0.67	304	18.3
	8	1.2	1	0.75	268	12.7
	35	1.5	1	0.40	150	2.1
	21	22.0	3.4	0.22	82	7.7
	26	1.2	1	2.20	215	1.5
	60	10.0	1	0.15	240	7
	25	1.4	1	2.10	230	3.9
	8	3.3	4.2	0.54	52	35
	50	2.1	1.1	1.40	584	4.6
	50	1.4	1	2.60	300	1.5
	25	1.6	7.9	0.25	250	9.3
	20	4.9	1.6	0.47	199	11.7
	35	23.6	1.14	0.13	167	15.4
	8	3.3	4.2	0.54	52	20
	60	2.5	3.8	1.20	284	3
	15	19.0	1	0.00	80	52
	30	22.9	1.4	0.13	165	20.4
	20	3.0	1.6	0.50	231	8.1
	20	3.7	1.6	0.49	290	11.2
	20	3.4	1.6	0.50	247	12.2
	22	7.0	1.6	0.50	177	8.3
	18	13.0	2.4	0.16	193	22
	25	1.2	1	0.00	320	2.8
	18	13.0	1.7	0.16	194	18.8
	45	1.0	1	0.50	564	4.4
	20	4.6	1.6	0.50	113	5.1
	20	3.7	1.6	0.49	139	7.4
	30	6.0	2.7	0.60	162	11
	6	0.9	1	1.00	113	6.4
	20	4.0	1.6	0.50	97	6.1
	50	1.5	1	0.40	150	1
	17	5.3	9.9	0.49	121	12
	25	1.0	1	3.00	196	6
	21	42.7	1	0.00	166	80
	5	20.0	1	0.15	85	81
	30	0.9	1	1.30	300	4
	12	3.5	1	0.43	25	3
	50	2.1	1.1	1.10	584	4.4
	27	24.4	1	0.00	120	14.3
	29	1.2	1	2.20	215	2.5
	40	3.6	1.8	0.83	304	13.3
	18	13.0	2.1	0.16	193	23.5

Table A Continued

Reference	B (m)	q_{net} (kPa)	N_{SPT}	L/B	D_f/B	S_m (mm)
Burland and Burbidge (1982, 1985)	26	14.5	1	0.24	255	18
	36	41.2	1	0.24	104	10
	30	6.0	2.7	0.47	162	10.5
	30	34.0	1.7	0.23	270	22
	4	3.3	4.4	0.30	99	37
	50	1.8	1.6	0.83	575	2.7
	45	1.0	1	0.50	339	6
	37	2.6	4.1	0.38	293	10.9
	20	4.6	1.6	0.50	166	8.1
	28	1.2	1	0.50	150	1.3
	20	6.1	1.6	0.49	161	10.2
	13	1.1	1	1.09	78	2
	25	1.8	1	1.70	230	3.4
	5	0.9	1	0.33	133	7.6
	20	0.9	1	1.33	300	2.7
	50	2.1	1.9	1.40	347	1.8
	6	14.5	4.4	0.77	74	74
	25	2.2	22	1.40	284	10.5
	20	18.3	1	0.02	41	4.8
	17	17.2	2.5	0.27	34	3.6
	20	17.6	4.8	0.61	218	26
	17	5.8	4.1	0.43	73	11.9
	10	4.4	5.5	0.57	93	8
	14	16.0	2.7	0.46	209	18.6
	34	33.0	1	0.16	191	43.8
	38	3.0	4.8	0.95	140	3
	15	14.0	1.6	0.18	18	4.2
	21	2.5	5.24	0.00	158	11.7
	12	3.8	3.2	0.39	90	15.5
	20	4.1	1	1.00	125	17.8
	18	2.5	1	0.30	576	25
	20	4.9	1.6	0.47	182	13.8
	18	3.0	1	0.29	500	25
	32	4.0	1.8	1.30	508	11.9
	20	4.9	1.7	0.31	112	7.4
	20	4.9	1.6	0.47	113	8.9
	20	4.3	1.6	0.49	134	10.2
	20	4.9	1.6	0.47	102	6.9
	42	6.0	2.7	0.60	215	4.1
	18	1.5	1	0.51	666	25
	20	5.5	1.6	0.47	139	9.4
	53	12.2	1	0.25	181	9.6
	7	3.3	4.4	0.61	99	37.1
	6	25.0	1	0.08	63	84
	20	3.7	1.6	0.49	252	16.5

Table A Continued

Reference	B (m)	q_{net} (kPa)	N_{SPT}	L/B	D_f/B	S_m (mm)
Burland and Burbidge (1982, 1985)	23	6.1	5	1.10	144	11.7
	42	6.0	2.7	0.47	158	7.9
	12	16.0	1.3	0.09	70	90
	34	3.4	6.7	0.00	81	10.7
	20	4.3	1.6	0.49	177	8.1
	18	3.0	1	0.25	500	25
	60	55.0	1.8	0.18	234	25
	11	9.0	8	0.50	115	25
	6	25.0	1	0.08	76	85
	20	3.7	1.6	0.49	225	7.4
	20	15.0	1.3	0.00	148	40
	20	15.2	1	0.02	33	2.8
	55	15.0	1.7	0.40	136	16.2
	18	6.4	1	0.23	101	7.1
	19	5.1	3.1	0.24	117	19.3
	8	2.3	1.1	1.02	400	43
	9	2.6	8.1	0.77	196	33
Wahls (1997)	58	5.2	3.7	0	127.8	17
	42	5.2	1	0.44	95.8	9.9
	22	4.9	1.8	0.3	118.7	6.4
	18	6.4	1.45	0.23	71.8	6.6
	16	16.2	1.6	0.29	139	15
	20	22.5	2.9	0.44	221	21
	42	5.1	4.6	0.35	114.9	5.8
	22	5.2	1	0.96	134	14.7
	39	6.6	2	0	168.1	15.5
	20	4.3	1.6	0.49	145	11
	44	5.2	3.7	0	1532	8.9
	49	4.9	2.8	0	161.4	7.1
	24	5	1.7	0.5	181.9	11.9
	20	3.4	1	1.5	129	11.5
	24	11	3	0.45	120	19.6
	20	3.7	1.59	0.49	215	15
	25	13.1	1.8	0.23	47.6	3.6
	24	8.5	1	0	102.5	16.3
	16	1	1	0	247.5	9.9
	20	3.7	1.59	0.49	215	6.4
	22	5.6	4.3	0.27	112	15.5
	24	4.6	5	0.43	112	11.2
	38	6.1	5	0.25	155.6	16.8
	39	4.6	4.5	0.59	85.7	21.1
	42	7	5.1	0.33	131.2	11.9
	20	3.7	1.6	0.49	279	8.6
	22	2.4	1.6	1.9	190	8.5
	43	4.6	3.5	0	111.1	23.9

Table A Continued

Reference	B (m)	q_{net} (kPa)	N_{SPT}	L/B	D_f/B	S_m (mm)
Wahls (1997)	50	3	3.3	1	230.8	21.1
	21	25.5	1	0.1	175	25