Indian Geotech J (February 2019) 49(1):58–69
https://doi.org/10.1007/s40098-017-0288-9

ORIGINAL PAPER

Prediction of Liquefaction Susceptibility of Clean Sandy Soils Using Artificial Intelligence Techniques

Ayad Salih Sabbar[1] · Amin Chegenizadeh[1] · Hamid Nikraz[1]

Received: 14 March 2017 / Accepted: 2 November 2017 / Published online: 9 November 2017

Abstract The liquefaction susceptibility of sandy soil is generally characterised by some parameters in the static liquefaction potential evaluation. These parameters are usually measured by static laboratory tests on distributed and undistributed samples under different test conditions. This study performs the ANN and genetic programming to estimate the static liquefaction susceptibility of clean sand soils based on experimental results to predict and develop an equation for the ratio of q_{min}/q_{peak} which is considered as the static liquefaction criterion. The q_{min}/q_{peak} model is a function of the minimum and maximum void ratios, relative density, initial effective confining pressure, and some other parameters. The findings of this study demonstrated that a good agreement between ANN and symbolic regression in predicting the ratio of q_{min}/q_{peak} based on laboratory tests. The possible application of the proposed q_{min}/q_{peak} equation is restricted by some limitations. The outcomes of the present work can be used in the preliminary liquefaction assessment of clean sandy soils prior to the complementary experimental studies.

Keywords Static liquefaction · ANN · Symbolic regression · Genetic programming · q_{min}/q_{peak}

Abbreviations

AI	Artificial intelligent
ANN	Artificial neural network
B	Skempton's coefficient
C_u	Uniformity coefficient
CPT	Cone Penetration Test
D_r	Relative density
D_{50}	Mean grains size
e_{max}	Maximum void ratio
e_{min}	Minimum void ratio
e	Void ratio
GP	Genetic programming
MGGP	Multi-Gene Genetic Programming
I	Number of input variables
q_{min}	Minimum deviatoric stress
q_{peak}	Initial peak deviatoric stress
RMSE	Root mean square error
R^2	Coefficient of determination
SPT	Standard Penetration Test
α	The ratio of initial shear stresses to the initial effective confining pressure
σ'_{3c}	Initial effective confining pressure
σ_1	Axial stress
σ_{3c}	Confining pressure

✉ Amin Chegenizadeh
amin.chegenizadeh@curtin.edu.au

[1] Department of Civil Engineering, Curtin University, Perth, Australia

Introduction

When saturated cohesionless soils were subjected to the undrained static or cyclic loading, the pore water pressure developed abruptly, and the effective stress reduced dramatically leading to a loss in shear strength or liquefaction [1–5]. The static liquefaction has been considered one of the biggest disastrous failure mechanisms in saturated sandy soils because when it occurs, the resistance of the soil reduces, and the ability of a soil layer to sustain many of geotechnical applications such as foundations of buildings and bridges, earth dams, slopes, and embankments is reduced [3, 6]. The phenomenon of complete static liquefaction can be described by

$$\sigma'_{3c} = 0 \text{ and } \sigma_1 - \sigma_{3c} = 0 \quad (1)$$

In which σ'_{3c} is the initial effective confining pressure, and $(\sigma_1 - \sigma_{3c})$ is the principle stress difference. The static liquefaction has been examined in many of previous experimental studies and most of these work reported that the static liquefaction is profoundly dependent on numerous factors such as void ratio (e), effective confining pressure (σ'_{3c}), relative density (D_r), consolidation type which defined by ratio of initial shear stress to the initial effective confining pressure (α), the degree of saturation (Skempton's coefficient B), fine content and the sample preparation methods [4, 7–9]. For practical purposes, the liquefaction susceptibility of sandy soils is generally evaluated in two ways: (1) utilizing monotonic undrained tests on undisturbed and remoulded specimens and adapting the collected results to equivalent parameters such as state parameter, undrained brittleness index, liquefaction potential, pore pressure ratio, and the ratio of minimum of deviatoric stress to the initial peak deviatoric stress [3, 4, 6, 10, 11], and (2) using the empirical relations based on the correlation between the field tests and laboratory tests [12]. The liquefaction susceptibility has been investigated in many studies based on different parameters [3, 13, 14]. The liquefaction susceptibility of soils may also be evaluated by assessing the ratio of minimum deviatoric stress to the peak deviatoric stress (q_{min}/q_{peak}) [4]. The complete static liquefaction is associated with an q_{min}/q_{peak} ratio of zero and the stable behaviour with complete dilation is associated with an q_{min}/q_{peak} ratio of 1 [4]. The q_{min}/q_{peak} ratio can be used to compute the liquefaction potential which is defined by $(q_{peak} - q_{min})/q_{peak} = 1 - q_{min}/q_{peak}$ [15, 16]. Rahman and Lo [17] pointed out the ratio of q_{min}/q_{peak} can be correlated to the state parameter to normalize the amount of strain softening in samples.

There are many limitations for using the results of experimental tests on disturbed specimens directly in the field situations. One of these constraints is these tests do not take into account the all actual properties of natural soils; for example fabric, cementation, strain history, and overconsolidation [18]. Next, the experimental tests are often costly and time-consuming. In the same way, analytical methods such as finite element method which is used to analyse many geotechnical engineering problems are hindered because these techniques require a large number of parameters in order to obtain an accurate constitutive model for complex problems such as liquefaction [19]. Therefore, artificial intelligence (AI) approaches have been used to reduce these difficulties and to provide an easy technique for evaluating the complex issues in geotechnical engineering [20, 21]. Artificial neural network (ANN) is one of artificial intelligence approaches which has been used widely to solve varieties of complicated issues in civil engineering. ANNs are becoming more dependable than analytical techniques such as traditional empirical and statistical methods due to two reasons. The first one is ANNs can learn from input data given to it. The second reason is the ability of ANNs to recognise the precise practical relationships between input data even though the fundamental relationships are unrecognised. Research on the development of ANN models based on laboratory elements tests is still very limited. Young-Su and Byung-Tak [22] used data that have been collected from different published works including simple cyclic shear and undrained cyclic triaxial tests in estimating the liquefaction resistance ratio of sand by using the ANN model. Banimahd et al. [23] developed ANN models to predict the undrained stress–strain behaviour and excess pore water pressure of sandy soils containing nonplastic fines. Their results showed that the undrained behaviour of sandy soil was dominated by four major input variables, including fine content, fine shape, relative density, and effective confining pressure. However, many ANN models have been used in various research to estimate the seismic liquefaction susceptibility of soils based on standard penetration test (SPT), cone penetration test (CPT), and seismic records after some main earthquakes took place in different countries Goh [24, 25], Ural and Saka [26], Hanna et al. [21], Tung et al. [27], Baziar and Nilipour [28]. For example, Goh [24] stated that Neural Networks could be considered as feasible tools for soil liquefaction evaluation and it is simpler to apply when compared with other methods. Mughieda et al. [29] adopted the ANN to estimate the liquefaction susceptibility of soils by relying on a database of CPT. Contradictory to previous studies, Mughieda et al. [29] argued that the preprocessing, normalising or calibrating the data before evaluating the liquefaction potential is not necessary. Farrokhzad et al. [30] developed ANN models to predict seismic liquefaction potential of soils based on a dataset from field tests of 30 boreholes. Additionally, ANN has been adopted to assess many problems in the area of geotechnical engineering such as prediction of scours at bridge piers, unsaturated shear strength of soil, safety of a typical artificial slope subjected to earthquake forces, horizontal ground displacement generated by earthquakes, the maximum dry density and permeability of various types of soils, and residual friction angle of clay soils [31–36].

Symbolic regression via genetic programming (GP) is another artificial intelligence approach that has been employed in many previous studies to develop new predictive relationships for numerous geotechnical engineering issues. GP is a process of developing computer programs to solve a problem. It depends on the evolutionary algorithms to provide a good approximate solution to problems by unexpectedly creating populations of computer applications presented by a tree structure. GP technique is distinguished from other AI techniques and statistical method by the possibility of predicting compact and explicit prediction model equations in terms of various model parameters [37]. Although the GP technique has been successfully used to predict some of the complex soil mechanics parameters, application of this technique in the evaluation of liquefaction potential is very limited [37–40]. Muduli and Das [37] used the multi-gene genetic programming (MGGP) to calculate the liquefaction potential of soil by relying on datasets from standard penetration tests (SPT). The results showed that the liquefaction potential model which was developed by MGGP was more accurate than models developed by using the support vector machine (SVM) and ANN when the same database was used. Muduli & Das [38] developed two different multi-gene genetic programming MGGP models to predict the liquefaction potential of soils regarding liquefaction index based on CPT database. Das and Muduli [40] investigated the liquefaction susceptibility of soil by using genetic programming based on CPT data collected after the 1999 Chi–Chi earthquake which took place in Taiwan. Javadi et al. [41] developed a new approach to evaluate liquefaction-induced lateral displacement of soil by using GP and based on SPT data. Jafarian et al. [42] have implemented GP to develop a predictive equation for strong ground motions induced by the earthquake. Jafarian et al. [18] used data of cyclic triaxial tests to develop the predicted equation of cyclic resistance ratio using GP. Gandomi and Alavi [39] have employed GP to analyse different geotechnical problems including soil liquefaction under earthquakes. Johari et al. [43] developed a predictive equation for soil–water characteristics curve using GP. Rezania and Javadi [44] used the GP to predict the settlement of shallow foundations.

As shown above, most studies in ANN and GP focused on developing models for prediction of seismic liquefaction potential based on in situ tests and seismic records. However, the studies on the prediction of static liquefaction susceptibility of clean sandy soil are rare. Moreover, using data-sets that include experimental parameters may provide another way to understand the static loading response of soils in an essential manner. Therefore, in present work, two types of artificial intelligence (AI) approaches are used to evaluate the static liquefaction susceptibility of sandy soils based on results of various undrained monotonic triaxial tests on the clean sand that were collected carefully from the previously published work with different initial characteristics. The first approach is ANN, which is used to predict the ratio of (q_{min}/q_{peak}) based on various combinations of input data. The second method is symbolic regression via genetic programming using the HeuristicLab software to correlate the ratio of (q_{min}/q_{peak}) to the initial soil parameters based on one of ANN models that led to the best estimation. For geotechnical engineering application, the outcomes of the present study can be used in initial investigations of the ratio of (q_{min}/q_{peak}) for clean sandy soils before final liquefaction evaluation.

Database of Undrained Monotonic Triaxial Tests

An input data involving the findings of both anisotropically and isotropically consolidated static triaxial tests on clean sand soils were collected by investigating previously presented research. Results of experimental studies conducted by Murthy et al. [45], Yamamuro and Lade [4], Jafarian et al. [3], Della et al. [46], Della et al. [47], Rahman [17], Verdugo & Ishihara [48], Belhouari et al. [49], and Yang and Wei [50] have been used in the database. Table 1 summarises the key characteristics of these laboratory tests. According to Table 1, most samples were prepared by using moist tamping method while few samples were prepared by using other sample preparation methods. The moist tamping method is widely employed in previous studies because various relative densities can be obtained using this method. Although the above tests were performed under different parameters, two significant parameters such as relative density and initial effective confining pressure were extensively varied in these tests. The static liquefaction database involves the results of 135 undrained static triaxial tests and correlating ratio of (q_{min}/q_{peak}) with different initial characteristics of clean sandy soils. The coefficient of uniformity (C_u), mean diameter (D_{50}), maximum void ratio (e_{max}), minimum void ratio (e_{min}), void ratio (e), relative density (D_r), initial confining pressure (σ'_{3c}), ratio of initial shear stress to the initial effective confining pressure (α), and the Skempton's coefficient B were selected in this study because they have been considered as the primary factors which might affect the static behaviour of sandy soil in previous research. The criterion for complete static liquefaction in this database, and accordingly in this study, is the ratio of (q_{min}/q_{peak}) of zero. However, the unity, and range between 0 and 1 of the ratio of (q_{min}/q_{peak}) associated with dilative behaviour and limited liquefaction respectively. Table 2 presents the statistical distribution of input parameters.

Artificial Neural Network (ANN) Models

Artificial neural network (ANN) is one of machine learning approaches constructed to imitate the human central nervous system especially the brain to model many complex

Table 1 Properties of sands that are used in the database

Sand type	Sample preparation method	C_u	D_{50}	Dr %	σ_{3c} (Mpa)	References
Ottawa and	Slurry deposition, moist tamping, and water pluviation	1.43	0.39	18–65	0.148–0.653	Murthy et al. [1]
Nevada and Ottawa	Funnel deposition and moist tamping	1.83	0.18	0–20	0.025–0.5	Yamamuro and Lade [4]
		2.317	0.205			
Babolsar	Moist tamping	1.8	0.24	8.5–68	0.04–0.41	Jafarian et al. [3]
Chlef	Wet deposition and funnel deposition	3.2	0.45	50	0.05–0.2	Della et al. [35]
Chlef	Wet deposition and funnel deposition	3.2	0.45	29, 80	0.05–0.2	Della et al. [2]
Sydney	Modified moist tamping	1.2	0.3	0–41	0.1–0.85	Rahman [3]
Toyoura	Wet tamping	1.7	0.17	18.5, 37	0.1–3	Verdugo and Ishihara [4]
Mostaganem	Dry funnel	1.704	0.32	15, 45	0.1–0.3	Belhouari et al. [5]
Toyoura	Moist tamping	1.392	0.216	0–34	0.1–0.5	Yang and Wei [39]
Fujian		1.532	0.397			

Table 2 The statistical distribution of each parameter in the database

Parameters	C_u	D_{50}	e_{max}	e_{min}	e	D_r %	σ_{3c} (Mpa)	B	α
Mean	1.85	0.31	0.88	0.55	0.78	30.72	0.34	0.94	0.03
Standard Error	0.05	0.01	0.01	0.00	0.01	1.75	0.03	0.01	0.01
Median	1.70	0.30	0.86	0.55	0.78	29.00	0.21	0.95	0.00
Mode	1.80	0.24	0.81	0.56	0.77	50.00	0.10	0.95	0.00
Standard deviation	0.63	0.10	0.09	0.05	0.10	20.32	0.38	0.07	0.09
Sample variance	0.40	0.01	0.01	0.00	0.01	412.91	0.15	0.00	0.01
Kurtosis	0.53	− 1.53	− 1.61	0.14	− 0.97	− 0.24	21.06	50.11	10.96
Skewness	1.37	0.14	0.30	0.60	− 0.02	0.56	3.83	− 6.34	3.46
Range	2.00	0.28	0.24	0.21	0.39	80.00	2.98	0.68	0.40
Minimum	1.20	0.17	0.78	0.48	0.58	0.00	0.03	0.32	0.00
Maximum	3.20	0.45	1.02	0.69	0.97	80.00	3.00	1.00	0.40

engineering issues. Further, ANNs have the ability to learn from data given to them, generalising the predicted inter-relationships for a future solution, and self-updating [51, 52]. Detailed information on ANN is reported in [53, 54]. In present work, nine parameters were utilised as input data of ANN models. Only one variable was used as output data. The process of training in a neural network includes adjusting the weights of parameters inside hidden layers until reaching the lowest difference between predicted output and actual output. In this study, the ANN models were generated using MATLAB's Neural Network Toolbox. The ANN model was based on Levenberg–Marquardt back-propagation algorithm, two layers feed-forward backprop, and seven hidden neurones. The number of hidden neurones has been chosen after many trails until reaching the minimum Root Mean Square Error (RMSE). All datasets were normalised using MATLAB's normalisation function which is a requirement of ANN modelling. The dataset in this study was split into three groups namely training, testing, and validation. The percentage of each group was 70, 15, and 15% for training, testing, and validation respectively. The training data is implemented to modify the connection weights. The testing dataset is adopted to avoid the overfitting, and the validation set is used to investigate the estimation ability of the model. The model can be considered an optimal model if it combines three conditions: (1) perfect performance in testing set, (2) a minimum number of hidden neurones, and (3) good performance in the training, testing, and validation sets. The performance of the ANN model was investigated by Root Mean Square Error (RMSE), and coefficient of determination (R^2) as shown below:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2} \qquad (2)$$

$$R^2 = \left[\frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}\right]^2 \qquad (3)$$

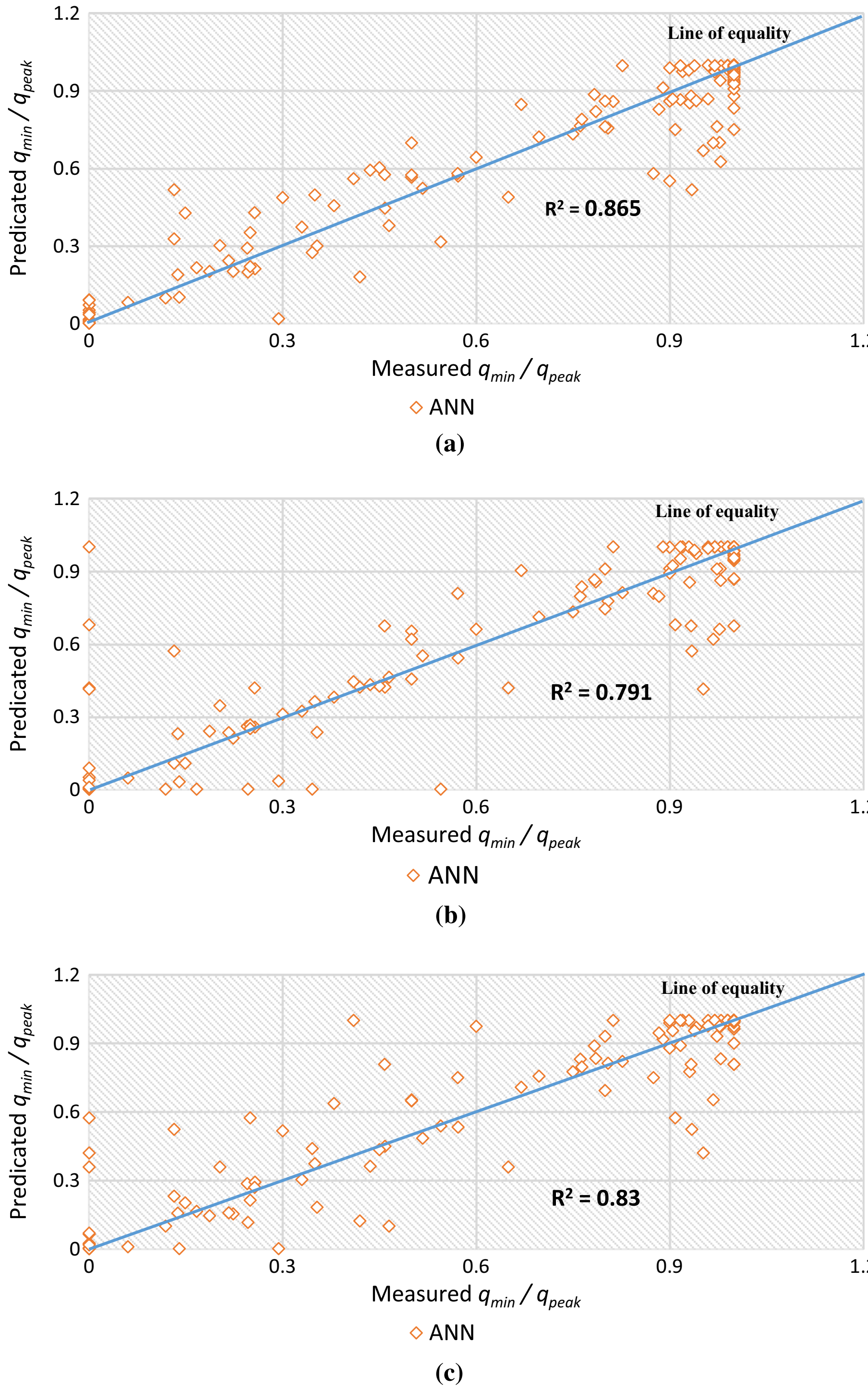

Fig. 1 Measured versus predicted q_{min}/q_{peak} for the ANN models: **a** model 1; **b** model 2 and **c** model 3

where x_i are the input data, y_i is model estimation, n is number of data points, $\bar{x}$ and $\bar{y}$ is the mean value of observed data model estimation respectively. The best performance of ANN models is reached through some steps of trial and error until the coefficient of determination R^2 around 90% is achieved. After completing the training and testing of the model, a sample data-set is used to test the accuracy of the model. Nine parameters are used to the ANN models as inputs. These include C_u, D_{50}, e_{max}, e_{min}, e, D_r, σ'_{3c}, α, and B. These input variables are chosen according to the previous experimental studies that pointed out the static behaviour of sandy soil is profoundly dependent on many factors such as physical properties, relative density, initial effective confining pressure, degree

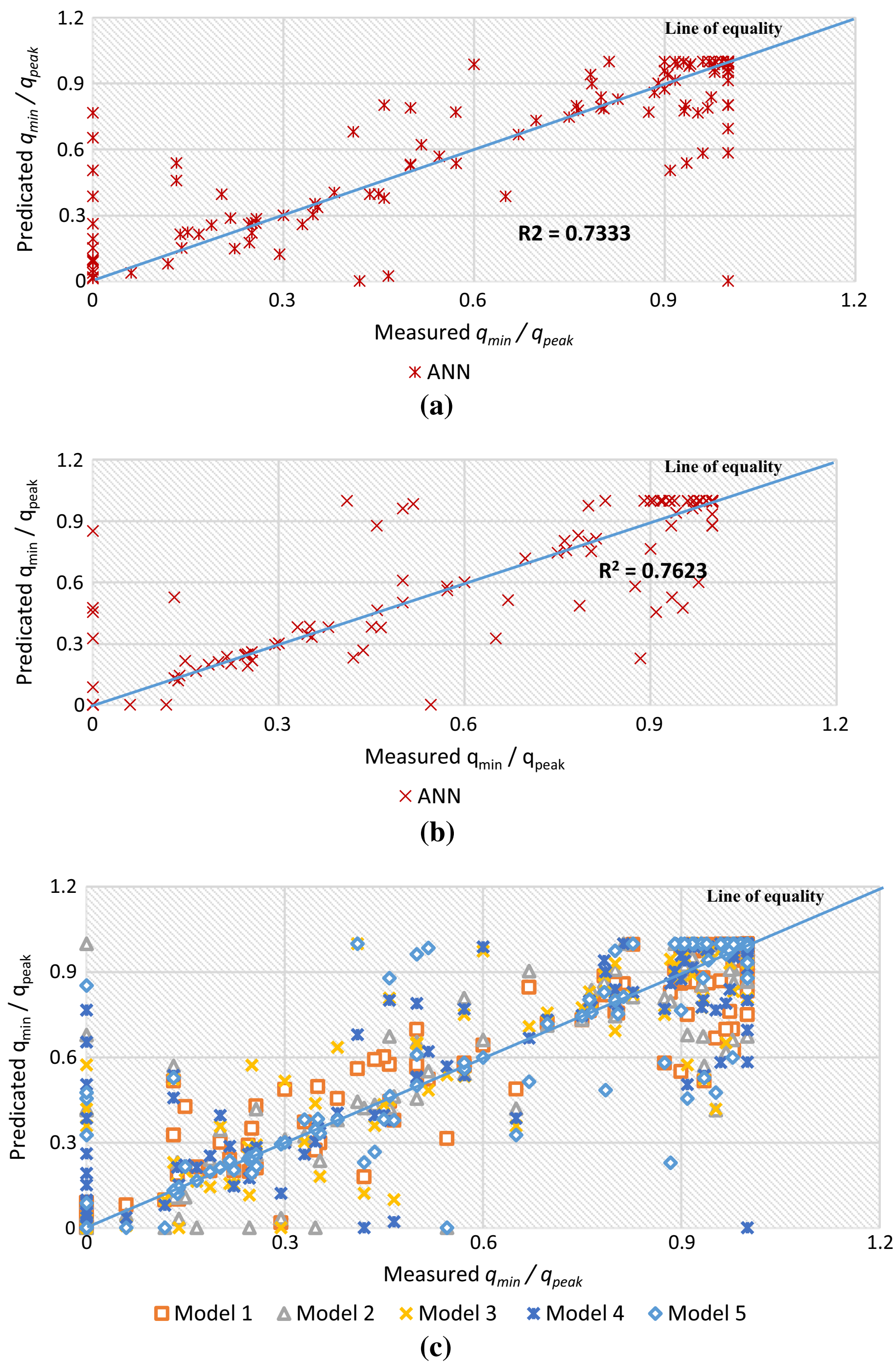

Fig. 2 Measured versus predicted q_{min}/q_{peak} for the ANN models: **a** model 4; **b** model 5 and **c** all ANN models

of saturation, and consolidation type. It is pertinent to mention here that the effect of sample preparation method was not considered in input data-set because the majority of samples used in the present work were deposited by the moist tamping technique as listed in Table 1. The ratio of q_{min}/q_{peak} serves to identify which of the input data is more superior in estimating the liquefaction susceptibility of sandy soil. The number of input variables is varied, and the five ANN models are investigated. Figures 1 and 2 compare the measured ratio of q_{min}/q_{peak} and ANN predicts the overall data-sets, whereas Table 3 presents the coefficient of determination (R^2) and Root Mean Square Error (RMSE) for overall data, training, testing, and validation of each model. Furthermore, it shows identity in statistical significance values for testing and validation in all models. It can be seen from Figs. 1, 2 and Table 3 that for model 1, when all inputs variables were used, it showed a good prediction with highest R^2 values 0.865, 0.846, and 0.864 for overall data, testing and training respectively. Lowest R^2 for overall data (0.733) and training set (0.722) were

Table 3 The performance and details of the ANN models

Model no.	Input parameters	No. hidden layer	Datasets	Performance	
				R^2	RMSE
1	(C_u), (D_{50}), (e_{max}), (e_{min}), (e), (D_r), (σ'_{3c}), (α),B	2	Overall data	0.865	0.0993
			Training	0.846	0.1000
			Testing	0.864	0.0905
			Validation	0.883	0.0993
2	(C_u), (D_{50}), (e_{max}), (e_{min}), (e), (D_r), (σ'_{3c}), B	2	Overall data	0.791	0.1247
			Training	0.789	0.1539
			Testing	0.828	0.1000
			Validation	0.828	0.1225
3	(C_u), (D_{50}), (e_{max}), (e_{min}), (e), (D_r), (σ'_{3c}), (α)	2	Overall data	0.830	0.1062
			Training	0.809	0.1283
			Testing	0.860	0.1534
			Validation	0.931	0.1062
4	(C_u), (D_{50}), (e_{max}), (e_{min}), (e), (D_r), (σ'_{3c})	2	Overall data	0.733	0.1430
			Training	0.722	0.1524
			Testing	0.748	0.2097
			Validation	0.783	0.1414
5	(C_u), (D_{50}), (e), (D_r), (σ'_{3c})	2	Overall data	0.762	0.1612
			Training	0.779	0.1414
			Testing	0.707	0.2302
			Validation	0.756	0.1612

obtained in model 4 when the variables α and B are eliminated while the lowest R^2 for testing set was obtained in model 5. In present study the efficiency of five models was evaluated in terms of testing data-sets as reported in Muduli and Das [38] and Das and Basudhar [36]. They stated that the performance of testing data-set should be used in evaluating the efficiency of different developed ANN models. Thus, it is found that the performance of model 1 with R^2 for testing set 0.864 was highest when compared to other models which indicates that good agreement between the measured and predicted the ratio of q_{min}/q_{peak}. However, model 5 showed the lowest R^2 for testing set 0.707 but with some difficulty in predicting the ratio of q_{min}/q_{peak}. The results in Fig. 1c and Table 3 also show a slight reduction in R^2 of testing and overall data-sets for model 3 when the Skempton's coefficient B eliminated comparing with model 1. This behaviour could be related to the almost tests were completely saturated, and the B values were more than 0.95. The significant reduction in R^2 (0.707) of testing data-set for model 5 indicates the high effect of e_{max}, e_{min}, and α on liquefaction susceptibility of sandy soils. The results of models 1–4 indicate the significant effect of C_u, D_{50}, e_{max}, e_{min}, D_r, σ'_{3c} on undrained static behaviour of sandy soils. A similar conclusion has been observed and reported by Young-Su and Byung-Tak [22] when they reported that using data that include C_u, D_{50}, e_{max}, e_{min}, D_r, and σ'_{3c} increased the ability of ANN model to capture the liquefaction resistance ratio of sandy soils. Banimahd et al. [23] also reported that the relative density, effective confining pressure, fines content, and fines shape has a profound effect on the ability of ANN models to predict the undrained static stress–strain behaviour and excess pore water pressure of sandy soils. Figure 2c compares the five ANN models implemented in present work, and as can be seen the results of five models are quiet close which exhibited a higher prediction performance for ANN models.

Genetic Programming Method

In present work, the HeuristicLab software was used to develop a functional relationship for the ratio of q_{min}/q_{peak} of clean sandy soils based on symbolic regression via genetic programming (GP). Symbolic regression can be defined as one of data mining approaches which is used to extract hidden, meaningful relationships using input data and weights. Symbolic regression depends on a tree-based genetic programming (GP) system to develop mathematical equations. The structure and operation of genetic programming (GP) have been described by numerous authors Koza [55], Johari et al. [43], [37, 38], and Rezania and Javadi [44]. HeuristicLab has been employed in some engineering problems. However, the utilisation of this

Table 4 Symbolic regression parameters

Parameters	Value
Population size	1000
Maximum of generation	75
Parent selection	Tournament (group size 7)
Replacement	1-Elitism
Crossover	Sub-tree-swapping
Mutation rate	15%
Fitness function	R^2 and RMSE
Function set	+, -, *,/, exp, ln
Terminal set	Constant, variable

software in soil mechanics and foundation engineering is still quite rare. HeuristicLab is an open source software based on heuristic and evolutionary algorithms and is generated by the Heuristic and Evolutionary Algorithms Laboratory (HEAL) since 2002 at the University of Applied Sciences Upper Austria. It works in C+ and is based on the Microsoft.Net framework [56]. Also, it has a high ability to provide a graphical user interface. HeuristicLab is supporting a broad range of algorithms such as genetic algorithm, Gaussian process regression and classification, neural network regression and classification…etc. Also, it supports many types of problems such as artificial art, classification, clusterin, symbolic regression, symbolic classification, etc. One of the features of HeuristicLab is the ability of software to simplify the complex model by trimming it to find a good agreement between the complexity and accuracy [56]. For more details about HeuristicLab, please see [56]. The input data were chosen based on the best results in ANN's modelling as described in part 3 which showed that model 1 with nine input variables exhibits the best prediction with the highest value of R^2. Nine parameters were used to the GP model as inputs. These include C_u, D_{50}, e_{max}, e_{min}, e, D_r, σ'_{3c}, α, and the B, while one parameter the ratio of q_{min}/q_{peak} is used as output. The input data were loaded into the software, thereafter a symbolic regression by GP was performed with variables set listed in Table 4.

The input data were divided into 67% for training and 33% for testing. The software approached the better model of the ratio of q_{min}/q_{peak} with the highest values of R^2 for training and testing, after a cycle of 75 generations. Therefore, the following equation was evolved to connect the ratio of q_{min}/q_{peak} of clean sandy soils to nine input parameters:

$$
\begin{aligned}
q_{min}/q_{peak} = {} & (7.314 * \mathrm{EXP}(19.980/((0.027 * -13.896 \\
& - (c_1 * e_{min} + c_2 * D_{50}))/((6.146 * -0.83/(1.100) \\
& - c_3 * C_u/(1.670))))/((6.146 * (13.875 * -13.283/(c_4 * e) \\
& + (\mathrm{EXP}(((c_5 * D_r + C_6 * e_{max}) + 5.780)) \\
& - (c_5 * D_r + c_6 * e_{max}))) - \mathrm{LN}(\mathrm{LN}(\mathrm{EXP}((c_5 * D_r + c_6 * e_{max}))) \\
& * (c_7 * D_r + c_8 * \sigma'_{3c})/((c_9 * D_{50} - 6.544)) * 18.994 \\
& * (c_{10} * D_r + -0.187))))) * -1.864 + 14.832)
\end{aligned}
\tag{4}
$$

The values of coefficients c_1–c_{10} are listed in Table 5.

The developed model (Eq. 4) was more sensitive to change in physical properties and initial state than other parameters. The same results were reported in ANN modelling where the performance of ANN model is affected by the change in C_u, D_{50}, e_{max}, e_{min}, D_r, and σ'_{3c}. This finding is consistent with findings of past studies by Young-Su and Byung-Tak [22] and Banimahd et al. [23]. It is also worth noting that the impact of C_u, D_{50}, e_{max}, e_{min}, D_r, and σ'_{3c} on the liquefaction susceptibility of sandy soil has been reported in many previous experimental studies [3, 4, 6]. The developed q_{min}/q_{peak} equation ignores some parameters such as ratio of initial shear stress to the initial effective confining pressure (α), and the Skempton's coefficient B. This can be related to these values are considered barely useful because experimental tests were almost fully saturated and isotropically consolidated with values around 1 and 0 for (B) and (α) respectively. Moreover, Eq. 4 does not take into account the effect of some field factors such as ageing, strain history, cementation, and stratification due to the difficulties in mimicking these conditions in experimental work. Figure 3 shows the measured values of the ratio of q_{min}/q_{peak} versus the equivalent values as predicted by Eq. (4). This Figure shows the data of training and testing sets are closely distributed around bisector line which indicates a good prediction ability for the developed model. The performance of the HeuristicLab model was examined using statistical precision parameters such as Root Mean Square Error (RMSE) and coefficient of determination (R^2). Table 6 presents R^2 and RMSE of the proposed model for the training and testing set. Figure 4 illustrates the plot of the normalized q_{min}/q_{peak} (i.e. the ratio of measured to the estimated q_{min}/q_{peak} values) versus the estimated q_{min}/q_{peak} values for all data-sets. The Figure shows that the almost

Table 5 The coefficients of Eq. (4)

c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}
1.9215	0.7856	2.4800	1.0416	1.5366	0.2935	3.0886	1.7902	1.1763	0.2620

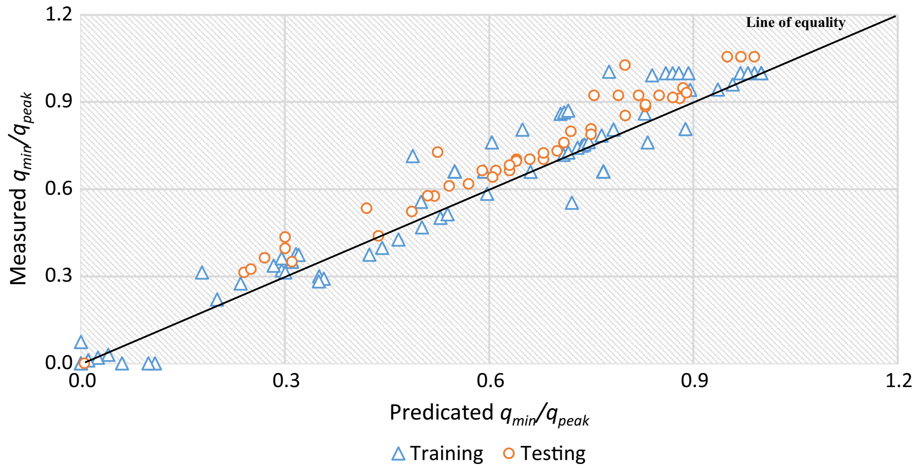

Fig. 3 Measured values of the ratio of q_{min}/q_{peak} versus those predicted by the developed model for training and testing datasets

Table 6 Performance of the q_{min}/q_{peak} model for the training and testing data sets

Dataset	Performance	
	R^2	RMSE
Training	0.868	0.12
Testing	0.842	0.17

normalized q_{min}/q_{peak} values are distributed around one which indicates a good agreement between measured and predicted values. Figure 5 demonstrates the tree of the developed model. In comparison, the classification accuracy of the ANN model1 was 0.846 and 0.864 for training and testing respectively. Similarly, the clasification accuracy for GP model was 0.868 and 0.842 for training and testing respectively. Thus, it is found that a good agreement exits between the two models in predicting the ratio of q_{min}/q_{peak}.

Parametric Study

The efficiency of newly proposed models in the prediction of static liquefaction susceptibility of sandy soils requires one to compare it with that of other modelling methods or experimental results. Thus, in present work, the methodical parametric study was implemented for the verification of the rate of success for Eq. 4 in the estimation of the ratio of q_{min}/q_{peak}, taking into consideration its physical meaning. In this parametric study, the findings of three undrained static triaxial tests were compared with the results of Eq. 4. These tests are part of authors published work which has been executed on very loose clean Perth sand samples [57]. Static undrained compression triaxial tests were performed on soil samples deposited by the moist tamping technique and isotropically consolidated under three different confining pressures; namely 100, 150, and 200 kPa. The sands used for the experiments was clean sand, and C_u, D_{50}, e_{max}, e_{min}, e, D_r, B, α were equal at 2.235, 0.35, 0.675, 0.544, 0.6615, 10%, 0.95 and zero respectively. The ratio of q_{min}/q_{peak} is calculated using GP model following Eq. 4 and experimental tests. As per the comparison presented in Fig. 6, a good agreement between experimental results and modelling results at confining pressure of 100 and 150 kPa. However, there was a slight difference at a confining pressure 200 kPa. This can be related to the developed equation which is more suitable in low confining pressure than higher confining pressures. Also, the applicability and validity of the developed equation are dependent upon the range of variables in input data which were collected from previous studies. Furthermore, the parametric study demonstrated that the ratio of q_{min}/q_{peak} increased with increasing confining pressure. This is reported in many experimental studies which showed that the liquefaction susceptibility of very loose samples decreased with increasing the relative density and confining pressure. Thus, it can be observed that the current GP model is equally efficient in predicting the ratio of q_{min}/q_{peak} when compared to experimental methods.

Summary and Conclusion

An artificial neural network and genetic programming models were developed for the prediction of static liquefaction susceptibility of sandy soils in terms of the ratio q_{min}/q_{peak} using data obtained from previously published work. The dataset included nine input parameters namely C_u, D_{50}, e_{max}, e_{min}, e, D_r, σ'_{3c}, α, and B and one target parameter called the ratio of q_{min}/q_{peak}. The ANN results demonstrate that the developed model using all nine parameters is able to efficiently capture the liquefaction sucepbtibility of soils with a coefficent of determination

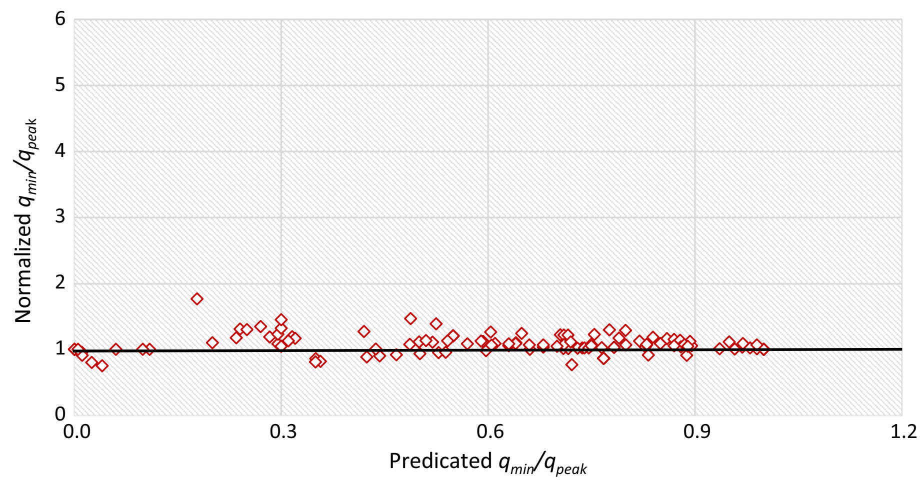

Fig. 4 The normalized ratio of q_{min}/q_{peak} versus those predicted by the developed model for all datasets

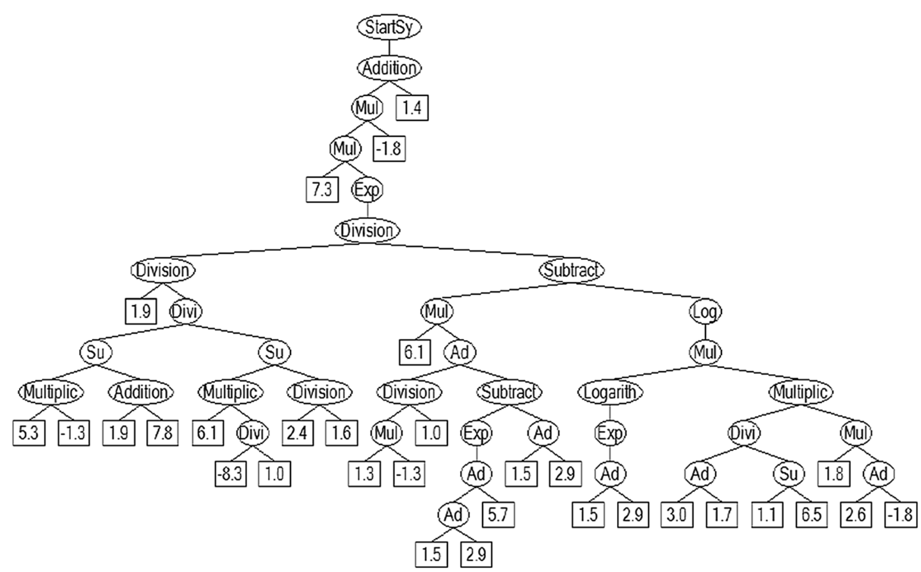

Fig. 5 The tree of the developed model

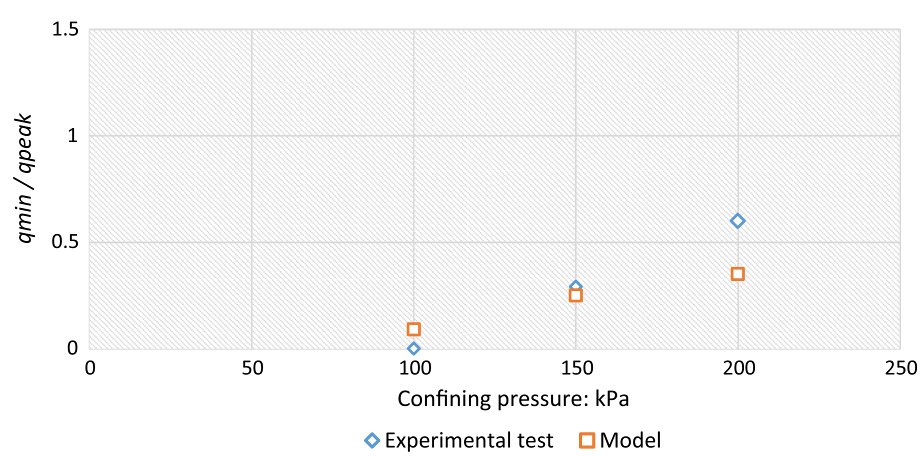

Fig. 6 The ratio of q_{min}/q_{peak} measured by experimental tests and the ratio of q_{min}/q_{peak} predicted by the developed model

(R^2) of 0.864 for testing set. However, the accuracy of ANN model is reduced to 0.707 for testing set when four parameters were eliminated (i.e., e_{max}, e_{min}, α, and B). A new equation for the prediction of liquefaction susceptibility of clean sandy soil in terms of the ratio of q_{min}/q_{peak} is proposed by using the symbolic regression via genetic programming.The results indicate a good agreement between ANN and GP approches in predicting the ratio of q_{min}/q_{peak}. Although the the ANN models and GP model showed a successful rate of predictiing the ratio of q_{min}/q_{peak}, the proposed models still contain some limitations. The limitations in present work might be related to some sources such as properties of database, the number of data, method of sample deposition, type of software, and type of regression analysis. Therefore, the findings of this work should be used carefully to account for the limitations presented above.

Acknowledgements The first author sincerely acknowledges the funding received from the Higher Committee for Education Development in the Republic of Iraq in the form of a scholarship for his PhD study.

References

1. Castro G (1969) Liquefaction of sands. Division of Engineering and Applied Mechanics, Harvard Univ, Cambridge
2. El Mohtar C, Bobet A, Drnevich V, Johnston C, Santagata M (2013) Pore pressure generation in sand with bentonite: from small strains to liquefaction. Géotechnique 64(2):108–117
3. Jafarian Y, Ghorbani A, Salamatpoor S, Salamatpoor S (2013) Monotonic triaxial experiments to evaluate steady-state and liquefaction susceptibility of Babolsar sand. J Zhejiang Univ Sci A 14(10):739–750
4. Yamamuro JA, Lade PV (1997) Static liquefaction of very loose sands. Can Geotech J 34(6):905–917
5. Madabhushi GS, Haigh SK (2012) How well do we understand earthquake induced liquefaction? Indian Geotech J 42(3):150–160
6. Salamatpoor S, Salamatpoor S (2014) Evaluation of babolsar sand behaviour by using static triaxial tests and comparison with case history. Open J Civ Eng 4:181–197
7. Yamamuro JA, Covert KM (2001) Monotonic and cyclic liquefaction of very loose sands with high silt content. J Geotech Geoenviron Eng 127(4):314–324
8. Ishihara K, Tsuchiya H, Huang Y, Kamada K (2001) Recent studies on liquefaction resistance of sand effect of saturation. In: Proceedings of 4th conference on recent advances in geotechnical earthquake engineering, keynote lecture
9. Wei L, Yang J (2014) On the role of grain shape in static liquefaction of sand–fines mixtures. Géotechnique 64(9):740–745
10. Sadrekarimi A (2014) Static liquefaction-triggering analysis considering soil dilatancy. Soils Found 54(5):955–966
11. Kramer SL, Seed HB (1988) Initiation of soil liquefaction under static loading conditions. J Geotech Eng 114(4):412–430
12. Stark TD, Mesri G (1992) Undrained shear strength of liquefied sands for stability analysis. J Geotech Eng 118(11):1727–1747
13. Konrad J-M (1988) Interpretation of flat plate dilatometer tests in sands in terms of the state parameter. Geotechnique 38(2):263–277
14. Bishop AW (1971) Shear strength parameters for undisturbed and remoulded soil specimens. In: Proceedings of the roscoe memorial symposium, Cambridge University, Cambridge, Mass, pp 3–58
15. Ishihara K (1993) Liquefaction and flow failure during earthquakes. Geotechnique 43(3):351–451
16. Lade PV, Yamamuro JA (2011) Evaluation of static liquefaction potential of silty sand slopes. Can Geotech J 48(2):247–264
17. Rahman MM, Lo S (2014) Undrained behavior of sand-fines mixtures and their state parameter. J Geotech Geoenviron Eng 140(7):04014036
18. Jafarian Y, Vakili R, Abdollahi AS (2013) Prediction of cyclic resistance ratio for silty sands and its applications in the simplified liquefaction analysis. Comput Geotech 52:54–62
19. Das SK (2013) 10 Artificial neural networks in geotechnical engineering: modeling and application issues. Metaheuristics in Water Geotech Transp Eng 45:231–267
20. Kim Y-S, Kim B-T (2008) Prediction of relative crest settlement of concrete-faced rockfill dams analyzed using an artificial neural network model. Comput Geotech 35(3):313–322
21. Hanna AM, Ural D, Saygili G (2007) Neural network model for liquefaction potential in soil deposits using Turkey and Taiwan earthquake data. Soil Dyn Earthq Eng 27(6):521–540
22. Young-Su K, Byung-Tak K (2006) Use of artificial neural networks in the prediction of liquefaction resistance of sands. J Geotech Geoenviron Eng 132(11):1502–1504
23. Banimahd M, Yasrobi S, Woodward P (2005) Artificial neural network for stress–strain behavior of sandy soils: knowledge based verification. Comput Geotech 32(5):377–386
24. Goh AT (1994) Seismic liquefaction potential assessed by neural networks. J Geotech Eng 120(9):1467–1480
25. Goh AT (1996) Neural-network modeling of CPT seismic liquefaction data. J Geotech Eng 122(1):70–73
26. Ural DN, Saka H (1998) Liquefaction assessment by neural networks. Electron J Geotech Eng (3)
27. Tung AT, Wang YY, Wong FS (1993) Assessment of liquefaction potential using neural networks. Soil Dyn Earthq Eng 12(6):325–335
28. Baziar M, Nilipour N (2003) Evaluation of liquefaction potential using neural-networks and CPT results. Soil Dyn Earthq Eng 23(7):631–636
29. Mughieda O, Bani-Hani K, Safieh B (2009) Liquefaction assessment by artificial neural networks based on CPT. Int J Geotech Eng 3(2):289–302
30. Farrokhzad F, Choobbasti A, Barari A (2010) Artificial neural network model for prediction of liquefaction potential in soil deposits. In: Fifth international conference on recent advances in geotechnical earthquake engineering and soil dynamics, San Diego, California, May 24–29 2010
31. Kaya A (2010) Artificial neural network study of observed pattern of scour depth around bridge piers. Comput Geotech 37(3):413–418
32. Lee S, Lee S, Kim Y (2003) An approach to estimate unsaturated shear strength using artificial neural network and hyperbolic formulation. Comput Geotech 30(6):489–503
33. Erzin Y, Cetin T (2012) The use of neural networks for the prediction of the critical factor of safety of an artificial slope subjected to earthquake forces. Sci Iran 19(2):188–194
34. Wang J, Rahman MS (1999) A neural network model for liquefaction-induced horizontal ground displacement. Soil Dyn Earthq Eng 18(8):555–568
35. Sinha SK, Wang MC (2008) Artificial neural network prediction models for soil compaction and permeability. Geotech Geol Eng 26(1):47–64
36. Das SK, Basudhar PK (2008) Prediction of residual friction angle of clays using artificial neural network. Eng Geol 100(3):142–145

37. Muduli P, Das S (2014) Evaluation of liquefaction potential of soil based on standard penetration test using multi-gene genetic programming model. Acta Geophys 62(3):529–543
38. Muduli PK, Das SK (2014) CPT-based seismic liquefaction potential evaluation using multi-gene genetic programming approach. Indian Geotech J 44(1):86–93
39. Gandomi AH, Alavi AH (2012) A new multi-gene genetic programming approach to non-linear system modeling. Part II: geotechnical and earthquake engineering problems. Neural Comput Appl 21(1):189–201
40. Das SK, Muduli PK (2011) Evaluation of liquefaction potential of soil using genetic programming. In: Proceedings of golden jubilee Indian geotechnical conference, pp 827–830, 15–17 Dec 2011, Kochi, India
41. Javadi AA, Rezania M, Nezhad MM (2006) Evaluation of liquefaction induced lateral displacements using genetic programming. Comput Geotech 33(4):222–233
42. Jafarian Y, Kermani E, Baziar MH (2010) Empirical predictive model for the v max/a max ratio of strong ground motions using genetic programming. Comput Geosci 36(12):1523–1531
43. Johari A, Habibagahi G, Ghahramani A (2006) Prediction of soil–water characteristic curve using genetic programming. J Geotech Geoenviron Eng 132(5):661–665
44. Rezania M, Javadi AA (2007) A new genetic programming model for predicting settlement of shallow foundations. Can Geotech J 44(12):1462–1473
45. Murthy T, Loukidis D, Carraro J, Prezzi M, Salgado R (2007) Undrained monotonic response of clean and silty sands. Géotechnique 57(3):273–288
46. Della N, Arab A, Belkhatir M (2011) Influence of specimen-reconstituting method on the undrained response of loose granular soil under static loading. Acta Mech Sin 27(5):796–802
47. Della N, Arab A, Belkhatir M, Missoum H (2009) Identification of the behavior of the Chlef sand to static liquefaction. Comptes Rendus Mécanique 337(5):282–290
48. Verdugo R, Ishihara K (1996) The steady state of sandy soils. Soils Found 36(2):81–91
49. Belhouari F, Bendani K, Missoum H, Derkaoui M (2015) Undrained static response of loose and medium dense silty sand of Mostaganem (Northern Algeria). Arab J Sci Eng 40(5):1327–1342
50. Yang J, Wei L (2012) Collapse of loose sand with the addition of fines: the role of particle shape. Geotechnique 62(12):1111–1125
51. Farrokhzad F, JanAliZadeh A, Barari A (2008) A prediction of slope stability using artificial neural network (Case study: Noabad, Mazandaran, Iran). In: Sixth international conference on case histories in geotechnical engineering, Arlington, Verginia, 11–16 August 2008
52. Elhag TM, Wang Y-M (2007) Risk assessment for bridge maintenance projects: neural networks versus regression techniques. J Comput Civ Eng 21(6):402–409
53. Flood I, Kartam N (1994) Neural networks in civil engineering. I: principles and understanding. J Comput Civ Eng 8(2):131–148
54. Hammerstrom D (1993) Working with neural networks. IEEE spectrum 30(7):46–53
55. Koza JR (1990) Genetic programming: a paradigm for genetically breeding populations of computer programs to solve problems. Department of Computer Science, Stanford University, Stanford
56. Wagner S, Kronberger G, Beham A, Kommenda M, Scheibenpflug A, Pitzer E, Vonolfen S, Kofler M, Winkler S, Dorfer V (2014) Architecture and design of the heuristiclab optimization environment. In: Advanced methods and applications in computational intelligence. Springer, pp 197–261
57. Sabbar AS, Chegenizadeh A, Nikraz H (2017) Static liquefaction of very loose sand–slag–bentonite mixtures. Soils Found 57(3):341–356. https://doi.org/10.1016/j.sandf.2017.05.003