

An Optimized Artificial Neural Network Structure to Predict Clay Sensitivity in a High Landslide Prone Area Using Piezocone Penetration Test (CPTu) Data: A Case Study in Southwest of Sweden

Abbas Abbaszadeh Shahri

Received: 8 September 2015 / Accepted: 1 January 2016 / Published online: 19 January 2016
© Springer International Publishing Switzerland 2016

Abstract Application of artificial neural networks (ANN) in various aspects of geotechnical engineering problems such as site characterization due to have difficulty to solve or interrupt through conventional approaches has demonstrated some degree of success. In the current paper a developed and optimized five layer feed-forward back-propagation neural network with 4-4-4-3-1 topology, network error of 0.00201 and $R^2 = 0.941$ under the conjugate gradient descent ANN training algorithm was introduced to predict the clay sensitivity parameter in a specified area in southwest of Sweden. The close relation of this parameter to occurred landslides in Sweden was the main reason why this study is focused on. For this purpose, the information of 70 piezocone penetration test (CPTu) points was used to model the variations of clay sensitivity and the influences of direct or indirect related parameters to CPTu has been taken into account and discussed in detail. Applied operation process to find the optimized ANN model using various training algorithms as well as

different activation functions was the main advantage of this paper. The performance and feasibility of proposed optimized model has been examined and evaluated using various statistical and analytical criteria as well as regression analyses and then compared to in situ field tests and laboratory investigation results. The sensitivity analysis of this study showed that the depth and pore pressure are the two most and cone tip resistance is the least effective factor on prediction of clay sensitivity.

Keywords Clay sensitivity · Landslide · Artificial neural network model · Piezocone penetration test

1 Introduction

Homogeneity and isotropy are two properties that are taken into account for most of the materials such as steel, concrete and timber in civil engineering design (Shahin et al. 2001; Park 2011). Although for the soils, it has been proved that due to complexity of geological formation which causes the imprecise physical processes, the geotechnical engineering properties of soil show varied and uncertain behavior (Jaksa 1995). Therefore developing analytical or empirical models in some simplified situations are feasible; however models that are more practical and less expensive than the analytical ones are of interest (Shahin et al. 2001; Park 2011). Inherent soil variability, loading, time and construction effects, human error, errors in soil boring,

A. Abbaszadeh Shahri (✉)
Department of Civil Engineering, College of Civil Engineering, Islamic Azad University, Roudehen Branch, Tehran, Iran
e-mail: a_abbaszadeh@iauh.ac.ir;
shahri@kth.se

A. Abbaszadeh Shahri
Department of Civil and Architectural Engineering, Royal Institute of Technology, 10044 Stockholm, Sweden

sampling, in situ and laboratory testing, characterization of the shear strength and stiffness of soils are some of the recognized uncertainty sources.

In this regard, Artificial Neural Networks (ANNs) as an alternate method is well suited to model complex problems where the relationship between the model variables is unknown (Hubick 1992). The ANNs are relatively crude electronic and computational models based on the neural structure of the brain consist of billions of highly interconnected neurons which provide a strong predicting and classification tool (e.g. Maulenkamp and Grima 1999; Zurada 1992; Fausett 1994).

Original ANNs as a sub system of Artificial Intelligence (AI) was introduced by McCulloch and Pitts (1943), and since then as an applicable tool have been used successfully for modeling of various fields of science and technology (e.g. Maier and Dandy 2000; Shahin et al. 2001; Zaheer and Bai 2003; Das 2005) in particular to almost all aspects of geotechnical engineering to solve complicated problems (e.g. Sayadi et al. 2013; Shahin et al. 2008; Basheer et al. 1996; Zhou and Wu 1994; Baziar and Ghorbani 2005; Goh 2002; Hanna et al. 2007; Kim and Kim 2006; Mayoraz et al. 1966; Fernandez-Steeger et al. 2002) as well as estimating geotechnical soil properties (e.g. Celik and Tan 2005; Lee et al. 2003; Yang and Rosenbaum 2002; Erzin 2007; Gribb and Gribb 1994; Sinha and Wang 2008; Cal 1995).

In the present paper applicability of ANNs in estimation of clay sensitivity (S_t) as one of the geotechnical soil properties has been notified. According to literature reviews the clay sensitivity parameter has a close relation to a unique type of high sensitive clay namely quick clay in Sweden (e.g. Nadim et al. 2008, Rankka et al. 2004; Rosenquist 1953; Torrance 1983; Lundström et al. 2009; Solheim et al. 2005; Abbaszadeh Shahri et al. 2015) and some other northern countries such as Norway, Canada, USA and Russia. Existence relation between this parameter with quick clay which is prone to slide and considered as the main responsible of occurred landslide in Sweden is the reason of focusing on this factor.

Reduction of clay shear strength to a very small fraction of its former value on remoulding at constant moisture content is called S_t and cyclic loading produced by wind, waves, ice and snow accumulation, earthquakes and other live loads cause cyclic stresses on foundations may lead to quick clay conditions and catastrophic failure. Terzaghi (1944) originally defined

the S_t in terms of unconfined compressive strength (UCS) however; the concept of S_t as a ratio between the undisturbed undrained shear strength (S_u) and the remaining strength after a so complete remoulding of the material that no further reduction can occur (disturbed undrained shear strength, S_{ur}) is commonly used to describe the possible loss of strength in clay due to remoulding (Åhnberg and Larsson 2012). The S_u can be assessed by in situ testing methods such as field shear vane and piezocone penetration test (PCPT or CPTu). The CPTu is a special type of Cone Penetration Test (CPT) which allows additional measurement of pore-water pressure (u) generated during the penetration as well as cone tip resistance (q_c), sleeve friction (f_s) and depth (e.g. Baligh et al. 1980; Tumay et al. 1981; Zuidberg et al. 1982; Lunne et al. 1997; Cai et al. 2010; Abbaszadeh Shahri et al. 2015). The CPTu not only provide valuable information on soil types but are also useful in deriving correlations with the engineering properties of soil for the purposes of analysis and design of foundations. In the recent years, the CPT and CPTu have been used as standard investigation tools, mainly to determine quickly the soil profile (through the friction ratio) as well as estimation of the S_u . In case of cohesive soils, the S_u is the most important quantity for geotechnical design in clay (Schmertmann 1975; Anagnostopoulou et al. 2003; Robertson 1999) and hence many empirical correlations have been developed to find a clear relationship between q_c and S_u from in situ tests such as CPTu and laboratory tests (e.g. Lunne and Kleven 1981; Jamiolkowski et al. 1982; Aas et al. 1986; Stark and Juhrend 1989; Mitchell and Brandon 1998; Lunne et al. 1986; La Rochelle et al. 1988; Rad and Lunne 1988). However, the accuracy of these correlations is poor, and their underlying theory is undependable (Kim et al. 2006).

The objective of the present paper is to evaluate the feasibility of different ANN algorithms to predict S_t using CPTu data. The quick propagation, conjugate gradient descent, quasi-Newton, limited memory quasi-Newton and Levenberg–Marquardt were the trained tested and developed ANN algorithms.

The optimized ANN model was selected using try and error method and tested by several statistical analyses criteria. The results showed that the conjugate gradient descent algorithm with minimum network root mean square error (RMSE) indicate better correlation with measured data. The performed sensitivity analysis in this study represented that depth

and pore pressure are the two most effective and cone tip resistance is the least effective factors on prediction of S_t .

2 Study Area and Available Data

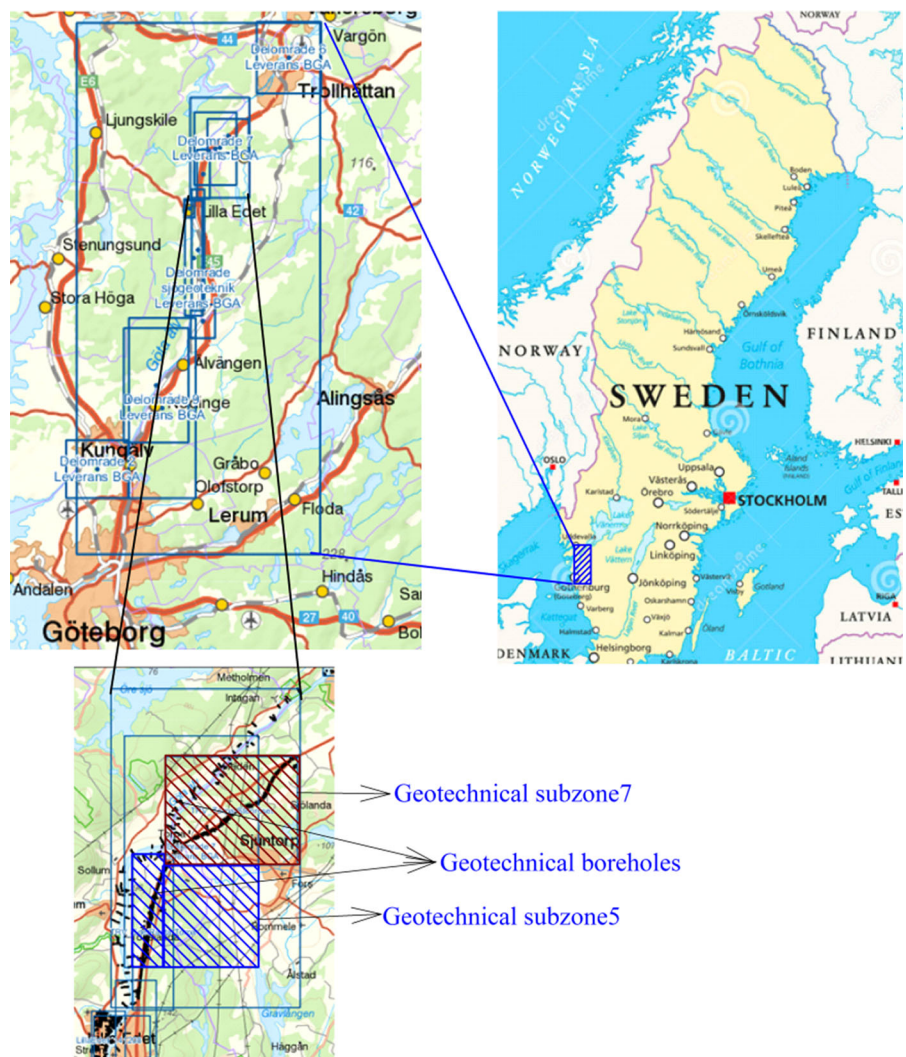
The selected area in this study with several subzones is around the Göta River near the Lilla Edet in the southwest of Sweden (Fig. 1) which has been recently studied by the Swedish Geotechnical Institute (SGI). A part of the mentioned area has been the subject of several geophysical investigations as well as CPTu analysis as referenced by Abbaszadeh Shahri et al. (2015).

Large number information for CPTu data and some laboratory test results are available in SGI website which this study used a total of 70 CPTu geotechnical test points and their available laboratory results belonging to subzones 5 and 7. The evidences of previous investigations indicate that the quick clay has been detected in glacial clay as layers or planes in clay sediments (Lindskog 1983; Klingberg 2010).

3 Estimation of Clay Sensitivity Using CPTu Data

To determine the S_t , the CPT or CPTu has been used to estimate the peak and minimum S_u of clay soils through empirical relations (Robertson 1999). However there

Fig. 1 Location of studied area in this paper (Swedish Geotechnical Institute (SGI) web site; <http://bga.swedgeo.se/bga/>)



are several methods for estimating the minimum S_u of soils from in situ tests (Worth 1984; Lunne et al. 1997; Seed and Harder 1990; Stark and Mesri 1992; Wride et al. 1999; Yoshimine et al. 1999; Ishihara 1993).

Typically, the peak S_u from CPTu data is estimated using Eq. (1) as proposed by Lunne et al. (1997).

$$S_{u(\text{peak})} = \frac{q_t - \sigma_v}{N_k} \quad (1)$$

where q_t total cone penetration resistance corrected for unequal end area effects, N_k empirical cone factor between 10 and 20 with an average of about 15, σ_v total overburden stress

Another possibility to calculate S_u is the use of effective cone resistance (q_E) which can be defined by Eq. (2).

$$S_u = \frac{q_E}{N_{ke}} = \frac{q_t - u_2}{N_{ke}} \quad (2)$$

where N_{ke} is the empirical cone factor (for the expression using q_E).

According to Lunne et al. (1997), the minimum (residual) S_u in clays often assumed to be equal to f_s , since the clay is almost fully remoulded as it passes the f_s and hence the clay sensitivity can be define by Eq. (3).

$$S_t = \frac{S_{u(\text{peak})}}{S_{u(\text{min})}} = \frac{q_t - \sigma_v}{f_s \times N_k} \quad (3)$$

The correction of f_s (Lunne et al. 1997) and q_c (Robertson et al. 1986) for pore water pressure can be obtained by Eqs. (4) and (5).

$$f_t = f_s - \frac{(u_2 \times A_{sb} - u_3 \times A_{st})}{A_s} \quad (4)$$

$$q_t = q_c + u_2(1 - a) \quad (5)$$

where f_t corrected sleeve friction, u_2 water pressure at the base of sleeve (measured pore pressure), A_{sb} cross section area of sleeve at the base, A_{st} cross section area of sleeve at the top, A_s surface area of sleeve, a ration between cone base un affected by the pore water pressure to total shoulder area.

The suggested normalized friction ratio (F_r) by Robertson (1990) is defined as Eq. (6) and hence the clay sensitivity can be estimated by Eq. (7) from CPTu data.

$$F_r = \frac{f_s}{q_t - \sigma_v} \times 100 \quad (6)$$

$$S_t = \frac{100}{F_r \times N_k} \xrightarrow{N_k=10-20} S_t = \frac{10}{F_r} \text{ to } S_t = \frac{5}{F_r} \quad (7)$$

where σ'_v is effective overburden stress.

In the case of soft clays the measured q_c values are relatively small and hence even minor errors can influence the measured values significantly (Rémai 2013). Therefore for very soft clays the use of excess pore water pressure may be better to find a reliable correlation (Eq. 8). Based on Robertson et al. (1986) and Robertson (1990), pore pressure ratio (B_q) and normalized cone resistance (q_{cnrm}) can be defined as Eqs. (9) and (10).

$$S_u = \frac{\Delta u}{N_{\Delta u}} = \frac{u_2 - u_0}{N_{\Delta u}} \quad (8)$$

$$B_q = \frac{u_2 - u_0}{q_t - \sigma_v} \quad (9)$$

$$q_{cnrm} = \frac{q_t - \sigma_v}{\sigma'_v} \quad (10)$$

where u_0 in situ pore pressure, σ_v and σ'_v total and effective overburden stress and $N_{\Delta u}$ is the empirical cone factor (Lunne et al. (1985): 4 and 10; Karlsrud et al. (1996): 6 and 8; Hong et al. (2010): 4 and 9). The q_{cnrm} and F_r in normally consolidated insensitive clays is around 2–6 and 5–10 % respectively (Robertson 1999).

Robertson (2008) defined the Eq. (11) for estimating the S_t .

$$S_t = \frac{S_u}{S_{ur}} = \frac{7.1}{F_r} \quad (11)$$

where S_{ur} is remoulded S_u at the same water content of undisturbed S_u .

By increasing the S_t of the clays, the f_s and friction ratio will decrease (Robertson 1999). Based on this relation, in general $S_t > 50$ characterizes the quick clay, however as shown in Fig. 2 different classifications for sensitivity have been developed which show that the quick clay has been defined based on its geotechnical behavior rather than its composition. In Sweden quick clay are defined as clays with $S_t > 50$ and $S_{ur} < 0.4$ kPa (Rankka et al. 2004) whereas in Norway, they are defined as clays with S_t up to 30 and $S_{ur} < 0.5$ kPa (Lundström et al. 2009) and in Canada as clays with $S_{ur} < 1.0$ kPa and a liquidity index of at least 1.2 (Robitaille et al. 2002).

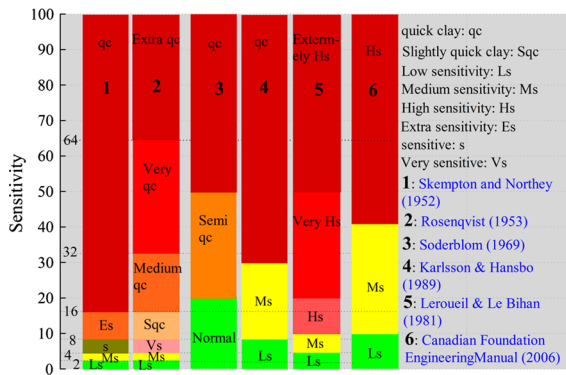


Fig. 2 Comparison of different classifications of S_t values (figure is provided by Abbaszadeh Shahri)

4 Applied Method

The efficiency of ANN in handling of highly non-linear relationships in data, even when the exact nature of such relationship is unknown can be considered as one of the major advantages of ANN. Therefore, due to the potential of mapping complex and non-linear relations between input and output variables of a system which are commonly used in non-linear engineering problems, the ANNs can be applied successfully in learning related classification, generalization, characterization and optimization functions.

The ability of ANNs for working with incomplete data, possess an error tolerance and show gradual convergence have been proved and hence they can easily form models for complex problems. Especially in the development of solutions for semi-structural or non-structural problems, ANN models can provide very successful results that are cheaper, faster and more adaptable than traditional methods. Hence, Networks are suitable approach to decompose a complex system into simpler elements or gathering simple elements to produce a complex one (Bar-Yam 1997).

Among of the various algorithms for training neural networks, the feed-forward back-propagation algorithm is the most efficient, simplest and also most general ones which are used for supervised training of multilayered neural networks. Back propagation works by approximating the non-linear relationship between the input and output by adjusting the weights values internally. It can further be generalized for the input that is not included in the training patterns (predictive abilities).

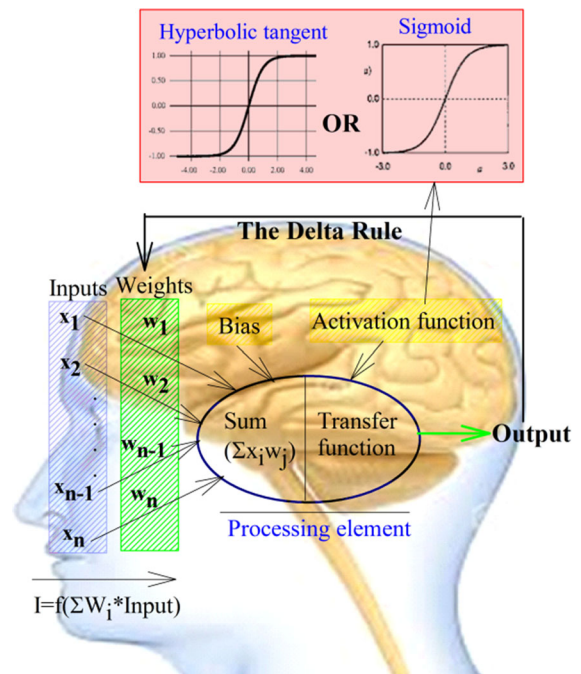


Fig. 3 Basic component of an artificial neuron and simplified procedure of Δ -rule (gradient descent)

In general, training and testing are common two stages of the back propagation network. Before taking new information, a network should be trained to provide the most efficient learning procedure for multilayer neural network. This fact that back-propagation algorithms are especially capable of solving predictive problems makes them very popular. The operations of the back propagation neural networks can be divided into feed forward and back propagation steps. In the feed forward step, an input pattern is applied to the input layer and its effect propagates, layer by layer, through the network until an output is produced.

As shown in Fig. 3, an artificial neuron as the basic element of a neural network consists of input, weights, bias, transfer function, activation function and output. Each neuron receives inputs x_1, x_2, \dots, x_n attached with a weight w_i which shows the connection strength of input for each connection which then multiply by the corresponding weight of the neuron connection.

A bias (b_i) can be defined as a type of connection weight with a constant nonzero value added to the summation of inputs and corresponding weights ‘u’, given in Eq. (12) (Cevik et al. 2011).

$$u_i = \sum_{j=1}^H w_{ij}x_j + b_i \quad (12)$$

The summation u_i is transformed using a scalar-to-scalar function called an activation or transfer function ($\varphi(u_i)$) yielding a value called the unit's "activation", given in Eq. (13).

$$Y_i = f(u_i) \quad (13)$$

In the output layer, the neuron computes the total weighted input x_j (Eq. 14) and then calculates the activity y_j using some function of the total weighted input (Fig. 3). The Fermi (Logistic or Sigmoid) and hyperbolic tangent functions (Eqs. 15 and 16) are very popular activation functions which this study considered both of them.

$$x_j = \sum_{i=1}^n y_i w_{ij} \quad (14)$$

$$y_i = \frac{1}{1 + e^{-x_j}} \textit{sigmoid} \quad (15)$$

$$y_i = \frac{e^{x_j} - e^{-x_j}}{e^{x_j} + e^{-x_j}} \textit{hyperbolic tangent} \quad (16)$$

The network's actual output value is then compared to the expected output, and an error signal is computed for each of the output nodes (Eq. 17).

$$E = \frac{1}{2} \sum_{i=1}^n (y_i - d_i)^2 \quad (17)$$

Since all the hidden nodes have, to some degree, contributed to the errors evident in the output layer, the output error signals are transmitted backwards from the output layer to each node in the hidden layer that immediately contributed to the output layer. This process is then repeated, layer by layer, until each node in the network has received an error signal that describes its relative contribution to the overall error.

Once the error signal for each node has been determined, the errors are then used by the nodes to update the values for each connection weights until the network converges to a state that allows all the training patterns to be encoded (Eq. 18).

$$w_{ji}(t+1) = w_{ji}(t) + \Delta w_{ji}(t+1) \quad (18)$$

where t is the iteration number between the output and hidden layers and Δw indicates the next value of the adaptation weights.

The back propagation algorithm looks for the minimum value of the error function in weight space using a technique called the Δ -rule or gradient descent (Fig. 3; Eq. 18) (Rojas 1996; Rumelhart et al. 1986). The weights that minimize the error function are then considered to be a solution to the learning problem.

5 Assessing the Optimized Network Architecture

In this paper, different types of networks using the MATLAB, NuMap7 (nonlinear regression/approximation networks) and NuClass 7 (nonlinear classification networks) which have developed in university of Texas at Arlington have been examined (training and testing) and developed to find an optimized ANN architecture model to predict the S_t .

Application of different ANN algorithms including quick propagation, conjugate gradient descent, quasi-Newton, limited memory quasi-Newton and Levenberg–Marquardt is advantage of this paper. The logistic, hyperbolic tangent and linear functions were used for activation of hidden and output layers and the sum-of squares were employed as output errors function respectively. The operation to find the optimized network architecture based on try and error method was started with one hidden layer and logistic activation function. Three components including the number of neurons, training algorithm and activation functions were considered and then the process was executed in three different stages. In each stage two of the mentioned components are fixed and the other will change. Therefore, numerous structures using different training algorithms with various activation functions as well as number of neurons in hidden layers were generated and controlled. For example, in number of neuron 10, several structures such as 4-4-3-3-1, 4-3-4-3-1, 4-3-2-5-1 and 4-3-5-2-1 separately was controlled for all training algorithms using hyperbolic tangent, logistic and then linear function. Then the operation was repeated and tested for the same structure using both logistic and hyperbolic tangent in different hidden layers. The value of network correlations and minimum root mean square error (RMSE) were the criteria to select the optimized network structure model. For each tested model the network error using RMSE was calculated and as presented in Fig. 4, the minimum error was observed in number of neurons 11 which is correspond to 4-4-4-

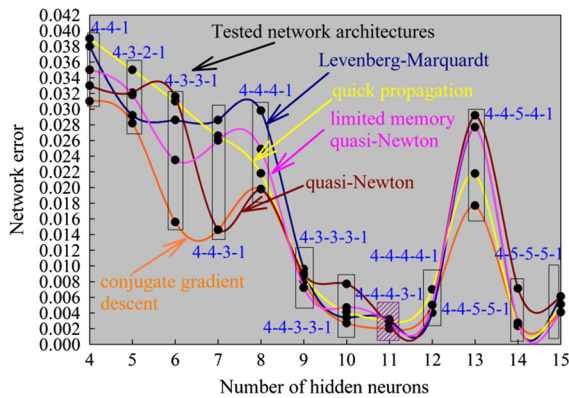


Fig. 4 Network performances of applied ANN algorithms for different number of neurons and some of the tested structure architectures

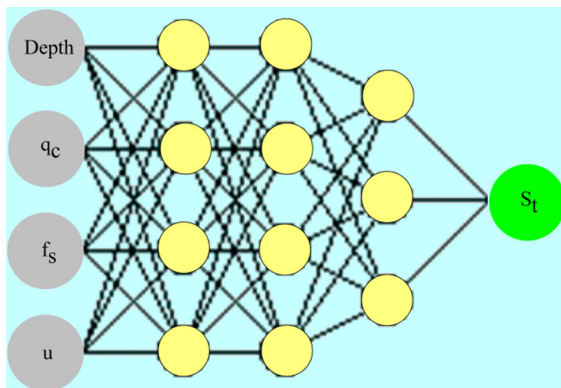


Fig. 5 Optimized network structure in this study

3-1 structure. This ANN structure model under training of conjugate gradient descent ANN algorithm and using the logistic activation function both in hidden and output layers was selected as the optimized structure model (Fig. 5). The characteristics of used database and ANN algorithms are given in Tables 1

Table 1 Characteristics of employed ANN architecture and datasets

Number of input neurons	4		
Number of output neurons	1		
Number of hidden layers	3 layers (layer 1: 4; layer 2: 4; layer 3: 4)		
Number of total dataset	2734		
Range of input data for ANN	Depth (m)	1.25 (min)	43.86 (max)
	Cone tip resistance (MPa)	0.137 (min)	2.661 (max)
	Sleeve friction (KN)	2.3125 (min)	1431.1 (max)
	pore pressure (KPa)	16.5 (min)	3345.09 (max)

and 2. The percentage of data for training, testing and validation with randomized selection were considered as 55, 20 and 25 % respectively.

In order to evaluate the results, the regression analyses between measured and predicted clay sensitivity values using the optimized ANN structure were performed (Fig. 6; Table 3). To compare the data scattering a 1:1 slope line has been used to show the dispersion of obtained data which can be used for interpretation of exact prediction and correlation (Fig. 7).

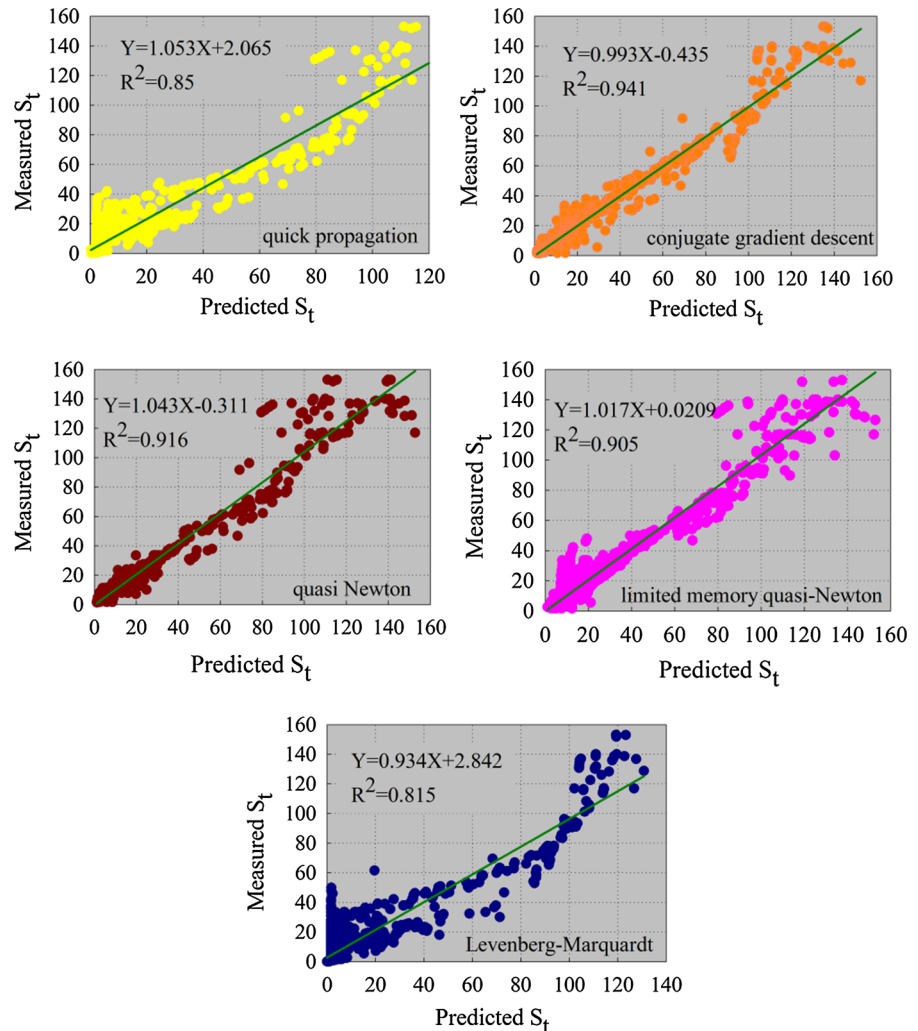
6 Discussion

In the recent years the ANN algorithms were modified many times and as a large number of training algorithms appeared that generally have improved characteristics as compared with the original. The quick propagation method (Fahlman 1988; Fahlman and Lebiere 1990) is a heuristic modification of the back propagation algorithm shows good results when working with most problems. The conjugate gradient descent method (Shewchuk 1994) ensures perfect training speed when working with up to 2000 training sets.

The Levenberg–Marquardt method as an advanced non-linear optimization and fastest available algorithm for multi-layer perceptrons illustrates excellent results when working with small training sets. However, the Levenberg–Marquardt can only be used on networks with a single output unit with small networks (a few hundred weights) because its memory requirements are proportional to the square of the number of weights in the network. It is only defined for the sum squared error function and therefore it is only appropriate for regression problems (Levenberg 1944; Marquardt 1963; Lourakis 2005; Nielson 1999; Transtrum and Sethna 2012).

Table 2 Properties of used ANN algorithms

ANN training algorithm	Network error	Number of iteration	Activation function	
			Hidden layer	Output
Quick propagation	0.00303	501	Hyperbolic tangent	Hyperbolic tangent
Conjugate gradient descent	0.00201	501	Logistic	Logistic
Quasi-Newton	0.00242	501	Hyperbolic tangent	Logistic
Limited memory quasi-Newton	0.00281	501	Logistic	Logistic
Levenberg–Marquardt	0.00318	501	Hyperbolic tangent	Logistic

Fig. 6 Regression results of the measured and predicted clay sensitivity using the employed ANN algorithms

On the basis of Newton's method, the quasi-Newton algorithm computes an approximate Hessian matrix during any iteration based on the gradients whereas the limited memory quasi Newton as a variation of quasi Newton avoids the need to store

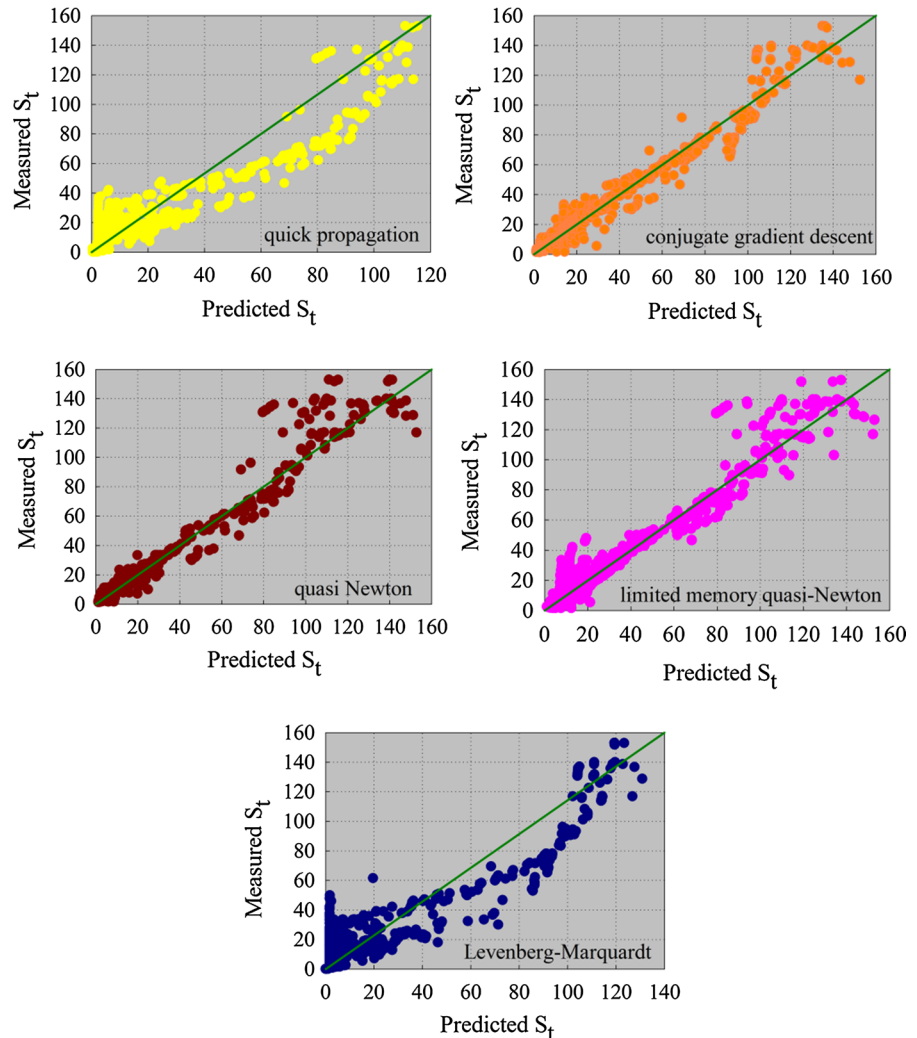
Hessian matrix and thus require less memory and can be used for bigger networks (Bertsekas 1995).

As mentioned above, due to number of data, the author didn't expect to get suitable results with Levenberg–Marquardt method regarding to other

Table 3 Obtained results of applied ANN models for prediction of clay sensitivity using CPTu data

	Train data		Test data		Validate data		All data	
	Correlation	R ²	Correlation	R ²	Correlation	R ²	Correlation	R ²
Quick propagation	0.930	0.843	0.933	0.865	0.938	0.887	0.931	0.855
Conjugate gradient descent	0.949	0.927	0.951	0.949	0.967	0.948	0.952	0.941
Quasi-Newton	0.947	0.922	0.948	0.912	0.931	0.909	0.945	0.916
Limited memory quasi-Newton	0.937	0.903	0.934	0.912	0.948	0.915	0.940	0.905
Levenberg–Marquardt	0.915	0.809	0.913	0.830	0.920	0.822	0.915	0.815

Fig. 7 Dispersion of measured and predicted S_t values regarding 1:1 slope line for the selected area



tested algorithms because it works well for small data sets, however the quick propagation method showed moderate adaptability (Fig. 6). Hence the competition for best applicable algorithm was between the

conjugate gradient descent, quasi Newton and limited memory quasi Newton methods.

According to obtained RMSE as selection criteria for optimized ANN model, the conjugate gradient

descent showed the minimum RMSE but based on the network type and number of used data sets, the author also expected to get reasonable results from quasi Newton and limited memory quasi Newton regarding to quick propagation and Levenberg–Marquardt as shown in Fig. 6. In Fig. 7, the located points on the 1:1 slope line indicate exact prediction and correlation which can demonstrate the accuracy of predicted S_t values using the tested ANN algorithms.

Moreover, in addition of using RMSE and coefficient of determination (R^2), the performance of selected ANN algorithms were tested by mean absolute percentage error (MAPE), variance absolute relative error (VARE), median absolute error (MEDAE) and variance account for (VAF) statistical criteria (Eqs. 19–22). Higher value of VAF and lower values of MAPE, VARE and MEDAE illustrate better network performance in prediction of S_t as presented in Table 4.

$$MAPE = \frac{1}{n} \times \left[\sum_{i=1}^n \left| \frac{t_i - x_i}{t_i} \right| \times 100 \right] \quad (19)$$

$$VARE = \frac{1}{n} \times \left[\sum_{i=1}^n \left(\left| \frac{t_i - x_i}{t_i} \right| - \text{mean} \left| \frac{t_i - x_i}{t_i} \right| \right)^2 \right] \times 100 \quad (20)$$

$$MEDAE = \text{median}(t_i - x_i) \quad (21)$$

$$VAF = \left[1 - \frac{\text{var}(t_i - x_i)}{\text{var}(t_i)} \right] \times 100 \quad (22)$$

where t_i and x_i are measured and predicted values.

The sensitivity analysis (Jong and Lee 2004; Eq. 21) as a method for determination of the effectiveness of each input parameter showed that depth with 47.932 % and pore pressure ratio with 42.76 % are the most and cone tip resistance with 37.59 % is

the least effective input parameters for S_t prediction using the selected ANN model (Fig. 8).

$$R_{ij} = \frac{\sum_{k=1}^m (x_{ik} \times x_{jk})}{\sqrt{\sum_{k=1}^m x_{ik}^2 \sum_{k=1}^m x_{jk}^2}}, \quad (23)$$

x_i and x_j : elements of data pairs

To represent how the obtained results of CPTu data and ANN algorithms can fit to real data; tow test points from different subzones of studied area randomly have been selected. At the first, by use of the CPTu data the predicted soil profile for each test point based on Robertson et al. (1986) was provided and then correlation between results of confirmed ANN architecture between CPTu results and measured data as a function of depth were executed and plotted (Fig. 9). The collected laboratory results of clay sensitivity show that 34.96 % data falls in the range of $20 < S_t < 30$, 33.21 % in the domain of $S_t > 32$ and 31.818 % in the range of $16 < S_t < 20$. The detailed results of sensitivity classifications (Fig. 2) are given in Table 5

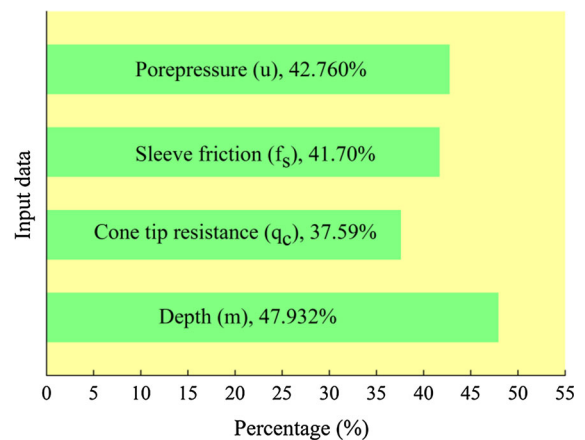


Fig. 8 Sensitivity analysis of input parameters in prediction of clay sensitivity

Table 4 Results of statistical criteria for tested ANN algorithms

	Quick propagation	Conjugate gradient descent	Quasi-Newton	Limited memory quasi-Newton	Levenberg–Marquardt
MAPE	8.783	6.476 (X)	6.89	8.320	9.237
RMSE	0.00303	0.00201 (X)	0.00242	0.00281	0.00318
VARE	9.30	5.84 (X)	6.53	7.08	10.64
MEDAE	0.049	0.029 (X)	0.032	0.039	0.051
VAF	94.76	97.28	97.31 (X)	96.65	90.80
R^2	0.85	0.941 (X)	0.916	0.905	0.815

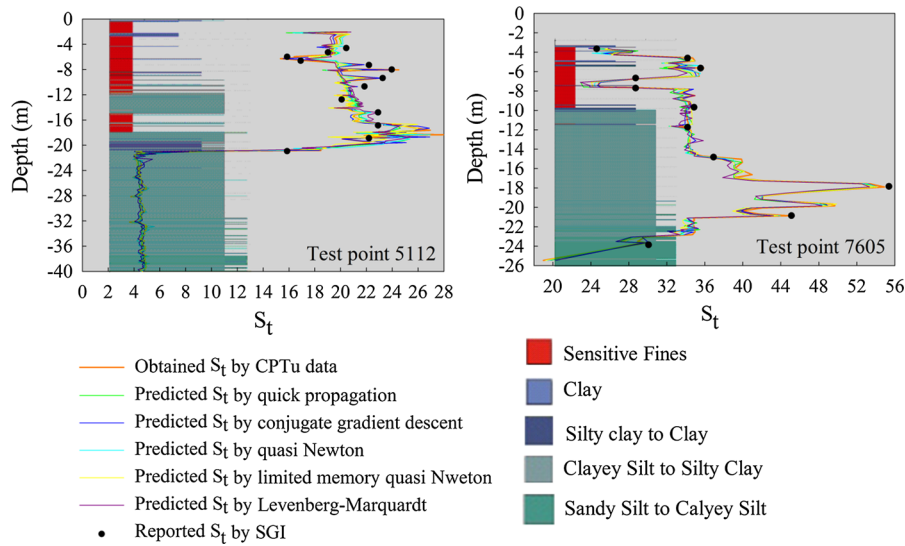


Fig. 9 Comparison of measured sensitivity with predicted values by CPTu data and various types of tested ANN algorithms. The predicted soil profile for each test point have been obtained using proposed method by Robertson et al. (1986)

Table 5 Comparison of available data in the study area with define range by previous researchers

Proposed classification	S _t values		
	16–20 (31.818 %)	20–30 (34.96 %)	>32 (33.21 %)
Skempton and Northey (1952)		Quick clay (100 %)	
Rosenquist (1953)	Medium quick clay (66.778 %)		Very quick clay
Soderblom (1969)	Medium sensitivity (66.778 %)		High sensitivity
Karlsson and Hansbo (1989)	Medium sensitivity (66.778 %)		High sensitivity
Le Bihan and Leroueil (1981)	High sensitivity	Very high sensitivity (68.17 %)	
Canadian Foun. Engng Manual (2006)	Medium sensitivity (66.778 %)		High sensitivity (>40)

which can be compared and correlated with the presented test point examples in Fig. 9 respectively.

It is proved that the ANN models can be varied from case to case. However, the proposed model in this study can be a very good initial guess to develop and adapt for another area. Due to flexibility of proposed ANN model in this study, it is recommended that without changing the model structure, use other training algorithms and then change the activation functions alternately. In final step, the number of neurons or neuron arrays can be considered.

7 Conclusions

ANNs can be applied for problems where the relationships may be quite dynamic or non-linear

and can provide an analytical alternative to conventional techniques which are often limited by assumptions. ANN can capture many kinds of relationships and allows to model phenomena which otherwise may have been very difficult or impossible to explain. This modeling capability, as well as the ability to learn from experience, has given ANNs superiority over most traditional modeling methods since there is no need for making assumptions about what the underlying rules that govern the problem in hand could be.

In the present paper several ANN algorithms with CPTu as inputs and clay sensitivity as output were tested and developed and the procedure to find the optimized ones were executed. The results of try and error method for testing numerous ANN structure architecture showed that increasing the number of hidden layers up to 4 in ANN will be able to improve

the results and in this condition the prediction of clay sensitivity by ANN can be reliable and reasonable and then the 4-4-4-3-1 structure due to minimum RMSE and higher correlation coefficient was confirmed as the developed optimized network structure. The proposed model for the selected area in this paper can be update and adapt for another area. However, the optimized topology can play an important role for initial guess to develop and making compatible for another area.

Depending on the nature of the application and the strength of the internal data patterns it can generally expect a network to train quite well and hence according to obtained results in Table 2 and Fig. 4, in this study the conjugate gradient descent training algorithm gives better results and therefore it can be applied as an alternative method for reasonable prediction of sensitivity and also the condition of data scattering using a 1:1 slope line to show the correlation were presented.

The sensitivity analysis showed that the depth and pore pressure are the two most and cone tip resistance is the least effective factor on estimation of clay sensitivity in this study.

A comparison between the proposed classifications by other researcher with our data to show adaptability percentage was executed (Table 5). The performed correlation between the laboratory and ANN results with predicted soil profile using the CPTu data represent a good condition whereas the nearly most of all measured sensitivity and those obtained by CPTu data and ANN algorithms fall in the part of soil profile which indicate the sensitive fine grained.

References

- Aas G, Lacasse S, Lunne T, Hoeg K (1986) Use of in situ tests for foundation design on clay. Use of in situ tests in geotechnical engineering (GSP 6). ASCE, New York, pp 1–30
- Åhnberg H, Larsson R (2012) Strength degradation of clay due to cyclic loadings and enforced deformation. Report No 75, Swedish Geotechnical Institute (SGI), Linköping, Sweden
- Anagnostopoulou A, Koukis G, Sabatakakis N, Tsiambaos G (2003) Empirical correlations of soil parameters based on Cone Penetration Tests (CPT) for Greek soils. *Geotech Geol Eng* 21:377–387
- Baligh MM, Vivatrat V, Ladd CC (1980) Cone penetration in soil profiling. *J Geotech Eng* 112(7):727–745
- Bar-Yam Y (1997) Dynamics of complex systems. Addison-Wesley, Boston
- Basheer IA, Reddi LN, Najjar YM (1996) Site characterization by neuronets: an application to the landfill sitting problem. *Ground Water* 34:610–617
- Baziar MH, Ghorbani A (2005) Evaluation of lateral spreading using artificial neural networks. *Soil Dyn Earthq Eng* 25(1):1–9
- Bertsekas DP (1995) Nonlinear programming. Athena Scientific, Belmont
- Cai GJ, Liu SY, Tong LY (2010) Field evaluation of deformation characteristics of a lacustrine clay deposit using seismic piezocone tests. *Eng Geol* 116(3):251–260
- Cal Y (1995) Soil classification by neural-network. *Adv Eng Softw* 22(2):95–97
- Canadian Geotechnical Society (2006) Canadian foundation engineering manual, 4th edn. p 506
- Celik S, Tan O (2005) Determination of pre-consolidation pressure with artificial neural network. *Civil Eng Environ Syst* 22(4):217–231
- Cevik A, Sezer EA, Cabalar AF, Gokceoglu C (2011) Modeling of the uniaxial compressive strength of some clay-bearing rocks using neural network. *Appl Soft Comput* 11:2587–2594
- Das SK (2005) Applications of genetic algorithm and artificial neural network to some geotechnical engineering problems. Ph.D. Thesis, Indian Institute of Technology Kanpur, Kanpur, India
- Erzin Y (2007) Artificial neural networks approach for swell pressure versus soil suction behavior. *Can Geotech J* 44(10):1215–1223
- Fahlman SE (1988) Faster learning variations on back propagation: an empirical study. In: Sejnowski TJ, Hinton GE, Touretzky DS (eds) Connectionist models summer school. Proceedings of the 1988 connectionist summer school, San Mateo, USA
- Fahlman SE, Lebiere C (1990) The cascade-correlation learning architecture. In: Touretzky DS (ed) Advances in neural information processing systems 2. Morgan Kaufman, San Mateo, pp 524–532
- Fausett L (1994) Fundamentals of neural networks: architectures, and applications. Prentice-Hall, Englewood Cliffs
- Fernandez-Steegeer TM, Rohn J, Czurdza K (2002) Identification of landslide areas with neural nets for hazard analysis. In: Rybar J, Stemmerk J, Wagner P (eds) Landslides. Proceedings of the IECL, Prague, Cz. Rep. June 24–26, 2002. Balkema, The Netherlands, pp 163–168
- Goh AT (2002) Probabilistic neural network for evaluating seismic liquefaction potential. *Can Geotech J* 39(1): 219–232
- Gribb MM, Gribb GW (1994) Use of neural networks for hydraulic conductivity determination in unsaturated soil. In: Proceedings of the 2nd international conference on ground water ecology, Bethesda, pp 155–163
- 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
- Hong S, Lee M, Kim J, Lee W (2010) Evaluation of undrained shear strength of Busan clay using CPT, 2nd international symposium on Cone Penetration Testing, CPT'10. In: Proceedings of 2nd international symposium on Cone Penetration Testing, CPT'10, online, 2010. Paper No. 2–23

- Hubick KT (1992) Artificial neural networks in Australia. Department of Industry, Technology and Commerce, Commonwealth of Australia, Canberra
- Ishihara K (1993) Liquefaction and flow failure during earthquakes. *Geotechnique* 43(3):351–415
- Jaksa MB (1995) The influence of spatial variability on the geotechnical design properties of a stiff, overconsolidated clay. Ph.D. dissertation, The University of Adelaide, Adelaide
- Jamiolkowski M, Lancellotta R, Tordella L, Battaglio M (1982) Undrained Strength from CPT. In: Proceedings of 2nd European symposium on penetration testing, Amsterdam, pp 599–606
- Jong YH, Lee CI (2004) Influence of geological conditions on the powder factor for tunnel blasting. *Int J Rock Mech Min Sci* 41(Supplement 1):533–538
- Karlsrud K, Lunne T, Brattlieu K (1996) Improved CPTu correlations based on block samples. *Nordisk Geoteknikermote, Reykjavik*
- Karlsson R, Hansbo S (1989) Soil classification and identification. *Bygghörskningsrådet Doc D 8:1989*
- Kim Y, Kim B (2006) Use of artificial neural networks in the prediction of liquefaction resistance of sands. *J Geotech Geoenviron Eng* 132(11):1502–1504
- Kim KK, Prezzi M, Salgado R (2006) Interpretation of cone penetration tests in cohesive soils. Publication FHWA/IN/JTRP-2006/22. Joint transportation research program, Indiana Department of Transportation and school of Civil Engineering Purdue University, West Lafayette, Indiana. doi:[10.5703/1288284313387](https://doi.org/10.5703/1288284313387)
- Klingberg F (2010) Bottenförhållanden i Göta Älv: SGU-rapport 2010:7. Sveriges Geologiska Undersökning, Göteborg
- La Rochelle P, Zebdi PM, Leroueil S, Tavenas F, Virely D (1988) Piezocone tests in sensitive clays of eastern Canada. In: Proceedings of the international symposium on penetration testing, ISOPT-1, Orlando, 2, Balkema Pub., Rotterdam, pp 831–841
- Le Bihan JP, Leroueil S (1981) The fall cone and behavior of remoulded clay. Terratech Ltd, Research report, Montreal
- Lee SJ, Lee SR, Kim YS (2003) An approach to estimate unsaturated shear strength using artificial neural network and hyperbolic formulation. *Comput Geotech* 30(6):489–503
- Levenberg K (1944) A method for the solution of certain nonlinear problems in least squares. *Q Appl Math* 2:164–168
- Lindskog G (1983) Brief report of the investigation of the slope stability along the river in Göta River valley. Statens Geotekniska Institut, Linköping
- Lourakis MIA (2005) A brief description of the Levenberg-Marquardt algorithm implemented by levmar. Technical Report, Institute of Computer Science, Foundation for Research and Technology—Hellas
- Lundström K, Larsson R, Dahlin T (2009) Mapping of quick clay formations using geotechnical and geophysical methods. *Landslides* 6:1–15
- Lunne T, Kleven A (1981) Role of CPT in North Sea foundation engineering. In: Symposium on cone penetration engineering division, ASCE, pp 49–75
- Lunne T, Robertson PK, Powell JJM (1997) Cone penetration testing in geotechnical practice. Blackie Academic, EF Spon/Routledge Publ, New York, p 312
- Lunne T, Christoffersen H, Tjelta T (1985) Engineering use of piezocone data in North Sea clays. In: Proceedings of ICSMFE-11; San Francisco, 2, 1985, pp 907–912
- Lunne T, Eidsmoen T, Gillespie D, Howland JD (1986) Laboratory and field evaluation of cone penetrometer. In: Proceedings of in situ '86, use of in situ tests in geotechnical engineering. ASCE GSP 6, Blacksburg, Virginia, pp 714–729
- Maier HR, Dandy GC (2000) Neural networks for prediction and forecasting of water resource variables: a review of modeling issues and applications. *Environ Model softw* 15:101–123
- Marquardt DW (1963) An algorithm for least-squares estimation of nonlinear parameters. *SIAM J Appl Math* 11:431–441
- Maulenkamp F, Grima MA (1999) Application of neural networks for the prediction of the unconfined compressive strength (UCS) from equotip hardness. *Int J Rock Mech Min Sci* 36(1):29–39
- Mayoraz F, Cornu T, Vuillet L (1996) Using neural networks to predict slope movements. In: Proceedings of VII international symposium on landslides, Trondheim, June 1966, 1. Balkema, Rotterdam, pp 295–300
- McCulloch WS, Pitts WH (1943) A logical calculus of ideas immanent in nervous activity. *Bull Math Biophys* 5(4):115–133
- Mitchell JK, Brandon TL (1998) Analysis and Use of CPT in earthquake and environmental engineering. In: Keynote lecture, proceedings of ICS'98, vol 1, pp 69–97
- Nadim F, Pedersen SAS, Schmidt-Thomé P, Sigmundsson F, Engdahl M (2008) Natural hazards in Nordic Countries. *Episodes* 31(1):176–184
- Nielson B (1999) Damping parameter In Marquardt's method. Technical Report IMM-REP-1999-05, Dept. of Mathematical Modeling, Technical University Denmark
- Park HL (2011) Study for application of artificial neural networks in geotechnical problems. In: Hui CLP (ed) Artificial neural networks-application. InTech, Croatia, pp 303–336. doi:[10.5772/2052](https://doi.org/10.5772/2052). ISBN 978-953-307-188-6
- Rad NS, Lunne T (1988) Direct correlations between piezocone test results and undrained shear strength of clay. In: Proceedings of 1st international symposium on penetration testing, Orlando, vol 2, pp 911–917
- Rankka K, Andersson-Sköld Y, Hultén C, Larsson R, Leroux V, Dahlin T (2004) Quick clay in Sweden. Report 65. Swedish Geotechnical Institute, Linköping
- Rémai Z (2013) Correlation of undrained shear strength and CPT resistance. *Period Polytech Civil Eng* 57(1):39–44. doi:[10.3311/PPci.2140](https://doi.org/10.3311/PPci.2140)
- Robertson PK (1990) Soil classification using the cone penetration test. *Can Geotech J* 27(1):151–158
- Robertson PK (1999) Estimation of minimum undrained shear strength for flow liquefaction using the CPT. In: Seco e Pinto (ed) Earthquake geotechnical engineering. Balkema, Rotterdam
- Robertson PK (2008) Discussion of 'liquefaction of silts from CPTu. *Can Geotech J* 44:140–141
- Robertson PK, Campanella RG, Gillespie D, Greig J (1986) Use of piezometer cone data. In-situ '86 use of in-situ testing in geotechnical engineering, GSP 6, ASCE, Reston, VA, Specialty Publication, pp 1263–1280

- Robitaille D, Demers D, Potvin J, Pellerin F (2002) Mapping of landslide-prone areas in the Saguenay region, Quebec, Canada. In: *Instability-planning and management*. Tomas Telford, London
- Rojas R (1996) Neural networks—a systematic introduction, chapter 7, the back propagation algorithm. <http://www.inf.fu-berlin.de/~rojas/neural/chap7.p.s>
- Rosenquist IT (1953) Considerations on the sensitivity of Norwegian quick-clays. *Geotechnique* 3:195–200
- Rumelhart DE, Hinton GE, Williams RJ (1986) Chapter 8, learning internal representation by error propagation parallel distribution processing: exploration in the microstructure of cognition, vol 1. MIT Press, Cambridge
- Sayadi A, Monjezi M, Talebi N, Khandelwal M (2013) A comparative study on the application of various artificial neural networks to simultaneous prediction of rock fragmentation and backbreak. *J Rock Mech Geotech Eng* 5:318–324
- Schmertmann JH (1975) Measurement of in situ shear strength. In: *Proceedings of specialty conference on in situ measurement of soil properties*: ASCE, Raleigh, vol 2, pp 57–138
- Seed RB, Harder LF Jr (1990) SPT-based analysis of cyclic pore pressure generation and undrained residual strength. In: Duncan JM (ed) *Proceedings, seed memorial symposium*, BiTech Publishers, Vancouver, pp 351–376
- Shahin MA, Jaksa MB, Maier HR (2001) Artificial neural network applications in geotechnical engineering. *Aust Geomech* 36(1):49–62
- Shahin MA, Jaksa MB, Maier HR (2008) State of the art of artificial neural networks in geotechnical engineering. *Electron J Geotech Eng Bouquet* 08:1–26
- Shahri AA, Malehmir A, Juhlin C (2015) Soil classification analysis based on piezocone penetration test data—a case study from a quick-clay landslide site in southwestern Sweden. *Eng Geol* 189:32–47
- Shewchuk JR (1994) An introduction to the conjugate gradient method without the agonizing pain, 1 1/4 edn. School of Computer Science, Carnegie Mellon University, Pittsburgh
- Sinha SK, Wang MC (2008) Artificial neural network prediction models for soil compaction and permeability. *Geotech Eng J* 26(1):47–64
- Skempton AW, Northey RD (1952) The sensitivity of clays. *Geotechnique* 3:30–53
- Soderblom R (1969) Salt in Swedish clays and its importance for quick clay formation. In: *Swedish Geotechnical Institute, Proceedings*, vol 22
- Solheim A, Berg K, Forsberg CF, Bryn P (2005) The storegga slide complex: repetitive large scale sliding with similar cause and development. *Mar Pet Geol* 22:97–107
- Stark TD, Juhrend JE (1989) Undrained shear strength from cone penetration tests. In: *Proceedings of the 12th international conference on soil mechanics and foundation engineering*, Rio de Janeiro, vol 1, pp 327–330
- Stark TD, Mesri G (1992) Undrained shear strength of sands for stability analysis. *J Geotech Eng Div ASCE* 118(11):1727–1747
- Terzaghi K (1944) Ends and means in soil mechanics. *Eng J* 27:608–613
- Torrance JK (1983) Towards a general model of quick clay development. *Sedimentology* 30:547–555
- Transtrum MK, Sethna JP (2012) Improvements to the Levenberg-Marquardt algorithm for nonlinear least-squares minimization. Preprint submitted to *Journal of Computational Physics*, Cornell University Library. [arXiv:1201.5885](https://arxiv.org/abs/1201.5885)
- Tumay MT, Boggess RL, Acar Y (1981) Subsurface investigation with piezocone penetrometer. *ASCE GSP Cone Penetr Test Exp*, St Louis, pp 325–342
- Worth CP (1984) The interpretation of in situ soil tests. *Geotechnique* 34(4):449–489
- Wride CE, McRoberts EC, Robertson PK (1999) Reconsideration of case histories for estimating undrained shear strength in sandy soils. *Can Geotech J* 36:907–933
- Yang Y, Rosenbaum MS (2002) The artificial neural network as a tool for assessing geotechnical properties. *Geotech Eng J* 20(2):149–168
- Yoshimine M, Robertson PK, Wride CE (1999) Undrained shear strength of clean sands to trigger flow liquefaction. *Can Geotech J* 36:891–906
- Zaheer I, Bai CG (2003) Application of artificial neural network for water quality management. *Int J Lowland Technol* 5(2):10–15
- Zhou Y, Wu X (1994) Use of neural networks in the analysis and interpretation of site investigation data. *Comput Geotech* 16:105–122
- Zuidberg HM, Schaap LHJ, Beringen FL (1982) A penetrometer for simultaneously measuring of cone resistance, sleeve friction and dynamic pore pressure. In: *Proceedings of the second European symposium on penetration testing*, Amsterdam, vol 2, pp 963–970
- Zurada JM (1992) *Introduction to artificial neural systems*. West Publishing Company, St. Paul