Journal of Rock Mechanics and Geotechnical Engineering 7 (2015) 584–592

Contents lists available at ScienceDirect

Journal of Rock Mechanics and Geotechnical Engineering

journal homepage: www.rockgeotech.org

Full length article

Regressive approach for predicting bearing capacity of bored piles from cone penetration test data

Iyad S. Alkroosh [a], Mohammad Bahadori [b], Hamid Nikraz [c], Alireza Bahadori [d,*]

[a] *Department of Civil Engineering, College of Engineering, Al-Qadissiya University, Al Diwaniyah, Iraq*
[b] *School of Soil Science and Engineering, University of Tehran, Tehran, Iran*
[c] *Department of Civil Engineering, Curtin University, Perth, Australia*
[d] *School of Environment, Science and Engineering, Southern Cross University, Lismore, New South Wales, Australia*

ARTICLE INFO

Article history:
Received 27 April 2015
Received in revised form
12 June 2015
Accepted 30 June 2015
Available online 5 August 2015

Keywords:
Bored piles
Cone penetration test (CPT)
Bearing capacity
Least square support vector machine (LSSVM)
Training
Validation

ABSTRACT

In this study, the least square support vector machine (LSSVM) algorithm was applied to predicting the bearing capacity of bored piles embedded in sand and mixed soils. Pile geometry and cone penetration test (CPT) results were used as input variables for prediction of pile bearing capacity. The data used were collected from the existing literature and consisted of 50 case records. The application of LSSVM was carried out by dividing the data into three sets: a training set for learning the problem and obtaining a relationship between input variables and pile bearing capacity, and testing and validation sets for evaluation of the predictive and generalization ability of the obtained relationship. The predictions of pile bearing capacity by LSSVM were evaluated by comparing with experimental data and with those by traditional CPT-based methods and the gene expression programming (GEP) model. It was found that the LSSVM performs well with coefficient of determination, mean, and standard deviation equivalent to 0.99, 1.03, and 0.08, respectively, for the testing set, and 1, 1.04, and 0.11, respectively, for the validation set. The low values of the calculated mean squared error and mean absolute error indicated that the LSSVM was accurate in predicting the pile bearing capacity. The results of comparison also showed that the proposed algorithm predicted the pile bearing capacity more accurately than the traditional methods including the GEP model.

1. Introduction

Bearing capacity is one of the most important factors that govern the design of pile foundations. Therefore, it has been the subject of interest for many researchers throughout the history of the geotechnical engineering profession. As a result, numerous theoretical and experimental procedures have been proposed to predict the pile behavior and bearing capacity. However, accurate evaluation of pile bearing capacity and certain interpretation of pile load transfer mechanism are still far from being accomplished due to the complexity of the problem.

The theoretical solutions, which employ the theory of bearing capacity to calculate the pile shaft and tip resistance, involve shortcomings resulting from considerable uncertainty over the factors that influence the bearing capacity. Among those factors are the effect of installation method, stress history and soil compressibility. For bored piles embedded in layered soil, the problem is more complex due to sensitivity of the factors that affect the behavior of the pile and the difficulty in quantifying those factors. For instance, the friction angle between pile and the surrounding soil cannot be exactly determined because of the effect of installation procedure and the difficulty in finding the real soil properties.

The experimental solutions that correlate the results of in-situ tests such as standard penetration test (SPT) or cone penetration test (CPT) with pile bearing capacity also involve setbacks. That may be attributed to that the SPT has substantially inherent variability and does not reflect soil compressibility (Abu-Kiefa, 1998). Moreover, the SPT results are affected by many factors, such as operator, drilling, hammer efficiency, and rate of blows. Hence, the accuracy of the proposed correlations between SPT data and pile bearing capacity is not assured. Although the correlation between pile capacity and CPT data can be a better alternative to the SPT correlation, comparative studies of the available CPT-based methods carried out by a number of researchers (e.g. Briaud, 1988; Roberston et al., 1988; Eslami, 1997; Abu-Farsakh and Titi, 2004; Cai et al., 2008) have shown that the capacity predictions can be very

* Corresponding author. Tel.: +61 2 6626 9347.
E-mail address: Alireza.bahadori@scu.edu.au (A. Bahadori).

Peer review under responsibility of Institute of Rock and Soil Mechanics, Chinese Academy of Sciences.

1674-7755 © 2015 Institute of Rock and Soil Mechanics, Chinese Academy of Sciences. Production and hosting by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jrmge.2015.06.011

different for the same case depending on the method employed. It is also found that these methods cannot provide consistent and accurate prediction of pile bearing capacity.

Considering the limitations of the proposed procedures for predicting pile bearing capacity and the limited success that they have achieved in terms of providing accurate prediction of pile bearing capacity, further research is required to overcome the complications associated with the problem. Artificial intelligence techniques may be better alternatives, due to the capability of being able to deal with complex and highly nonlinear functions, and employing the considerable capacity of computers to perform enormously iterated work. The modeling advantage of these techniques is their ability to capture the nonlinear and complex relationships between the targeted output and the factors affecting it, without having to assume a priori formula describing this relationship. A number of researchers (e.g. Teh et al., 1997; Abu-Kiefa, 1998; Das and Basudhar, 2006; Ardalan et al., 2009; Shahin, 2010; Ornek et al., 2012; Tarawneh, 2013) have successfully applied artificial neural network (ANN), which is a form of artificial intelligence, to solving engineering problems. Genetic programming (GP), which is another form of artificial intelligence, has been used successfully in solving engineering problems (Rezania and Javadi, 2007; Alavi et al., 2011; Alkroosh and Nikraz, 2011a, b; Gandomi, 2011; Gandomi and Alavi, 2012). Recently, an emerging algorithm, i.e. the least square support vector machine (LSSVM), which is a developed version of support vector machine (SVM), has been found successful in solving engineering problems (Das et al., 2011a, b; Samui and Kothari, 2011). This study investigates the feasibility of using the LSSVM to predict the bearing capacity of bored piles embedded in sand and mixed soils more accurately than the available methods.

2. Support vector machine (SVM)

The SVM is a method developed using the statistical learning concept (Suykens and Vandewalle, 1999). It has been widely used across the world (Cortes and Vapnik, 1995; Bazzani et al., 2001; Suykens et al., 2002; Amendolia et al., 2003; Baylar et al., 2009; Übeyli, 2010; Chen et al., 2011; Chamkalani et al., 2013; Rafiee-Taghanaki et al., 2013; Shokrollahi et al., 2013).

If we have training samples with given data $x_i \in \mathbb{R}^n$ and result data $y_i \in \mathbb{R}$ with labels −1 and 1, respectively, the SVM estimates the function shown below (Suykens and Vandewalle, 1999; Suykens et al., 2002):

$$y = \boldsymbol{w}^{\mathrm{T}}\Phi(x) + b \tag{1}$$

where $\Phi(x)$ is the function that maps x, and $\boldsymbol{w}$ and b are the weight vector and bias variable. When the data are separable, we will have (Suykens and Vandewalle, 1999; Suykens et al., 2002):

$$\left.\begin{array}{ll} \boldsymbol{w}^{\mathrm{T}}\Phi(x_k) + b \geq 1 & (y_k = 1) \\ \boldsymbol{w}^{\mathrm{T}}\Phi(x_k) + b \leq 1 & (y_k = -1) \end{array}\right\} \tag{2}$$

Eq. (2) is nearly equal to (Das et al., 2011a, b; Chamkalani et al., 2013):

$$y_k\left[\boldsymbol{w}^{\mathrm{T}}\Phi(x_k) + b\right] \geq 1 \quad (k = 1,\ 2,\ \ldots,\ N) \tag{3}$$

The further development of linear SVM to non-independent case was also created by Cortes and Vapnik (1995). Simply, it is done by presenting extra variables into Eq. (3) (Suykens and Vandewalle, 1999; Suykens et al., 2002):

$$y_k\left[\boldsymbol{w}^{\mathrm{T}}\Phi(x_k) + b\right] \geq 1 - \zeta_k,\ \zeta_k \geq 0 \quad (k = 1,\ 2,\ \ldots,\ N) \tag{4}$$

where ζ_k is the deviation factor.

The optimal separating hyperplane is predicted using the vector $\boldsymbol{w}$ that minimizes the functional conditions using the constraints (Eq. (4)) (Suykens and Vandewalle, 1999; Suykens et al., 2002; Übeyli, 2010):

$$\Phi(\boldsymbol{w},\ \zeta_i) = \frac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} + \frac{C}{2}\sum_{i=1}^{N}\zeta_i^p \tag{5}$$

where p is the upper limit, and C is a coefficient.

In the SVM, optimal separating hyperplane is calculated using the quadratic method (Cortes and Vapnik, 1995):

$$\begin{aligned}\Phi(\boldsymbol{w},\ b,\ \alpha_i,\ \zeta_i,\ \beta_i) = {} & \frac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} + \frac{C}{2}\sum_{i=1}^{N}\zeta_i \\ & - \sum_{i=1}^{N}\alpha_i\left[y_i\left(\boldsymbol{w}^{\mathrm{T}}x_i + b\right) - 1 + \zeta_i\right] - \sum_{j=1}^{N}\beta_i\zeta_i a\end{aligned} \tag{6}$$

where a is the adjustable parameter, α_i and β_i are the Lagrange multipliers (Suykens and Vandewalle, 1999; Suykens et al., 2002).

In contrast to the SVM, the LSSVM is developed using minimization of cost equation (Suykens and Vandewalle, 1999; Suykens et al., 2002):

$$\Phi(\boldsymbol{w},\ \zeta_i) = \frac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} + \frac{C}{2}\sum_{i=1}^{N}\zeta_i^2 \tag{7}$$

$$y_i\left[\boldsymbol{w}^{\mathrm{T}}\Phi(x_i) + b\right] = 1 - \zeta_i \quad (i = 1,\ 2,\ \ldots,\ N) \tag{8}$$

To derive the dual problem for the nonlinear classification problem of LSSVM, the Lagrange function is defined as (Suykens and Vandewalle, 1999; Suykens et al., 2002):

$$\begin{aligned}L(w,\ b,\ \zeta,\ \alpha) = {} & \frac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} + \frac{C}{2}\sum_{i=1}^{N}\zeta_i^2 \\ & - \sum_{i=1}^{N}\alpha_i\left\{y_i\left[\boldsymbol{w}^{\mathrm{T}}\Phi(x_i) + b\right] - 1 + \zeta_i\right\}\end{aligned} \tag{9}$$

The conditions for optimality can be obtained as

$$\left.\begin{array}{l} \dfrac{\partial L}{\partial \boldsymbol{w}} = 0 \Rightarrow \boldsymbol{w} = \displaystyle\sum_{i=1}^{N}\alpha_i y_i \Phi(x_i) \\ \dfrac{\partial L}{\partial b} = 0 \Rightarrow \displaystyle\sum_{i=1}^{N}\alpha_i y_i = 0 \\ \dfrac{\partial L}{\partial \zeta_i} = 0 \Rightarrow \alpha_i = \gamma\zeta_i \quad (i = 1,\ 2,\ \ldots,\ N) \\ \dfrac{\partial L}{\partial \alpha_k} = 0 \Rightarrow y_i\left[\boldsymbol{w}^{\mathrm{T}}\Phi(x_i) + b\right] = 1 - \zeta_i \quad (i = 1,\ 2,\ \ldots,\ N) \end{array}\right\} \tag{10}$$

By defining $\boldsymbol{Z}^{\mathrm{T}} = [\Phi^{\mathrm{T}}(x_1)y_1,\ \Phi^{\mathrm{T}}(x_2)y_2,\ \ldots,\ \Phi^{\mathrm{T}}(x_N)y_N]$, $\boldsymbol{Y} = [y_1,\ y_2,\ \ldots,\ y_N]$, $\overrightarrow{1} = [1,\ 1,\ \ldots,\ 1]$, $\boldsymbol{\zeta} = [\zeta_1,\ \zeta_2,\ \ldots,\ \zeta_N]$, $\boldsymbol{\alpha} = [\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_N]$, Eq. (10) is finally converted into the below form (Minoux, 1986; Suykens and Vandewalle, 1999):

$$\begin{bmatrix} 0 & \vec{1}^{\mathrm{T}} \\ \vec{1} & \boldsymbol{\Omega} + \gamma^{-1}\boldsymbol{I}_N \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \boldsymbol{Y} \end{bmatrix} \tag{11}$$

where $\boldsymbol{I}_N$ is an $N \times N$ identity matrix, and $\boldsymbol{\Omega} \in \mathbb{R}^{N \times N}$ is the kernel matrix defined by

$$\Omega_{ij} = \Phi(x_i)\Phi(x_j) = K(x_i,\ x_j) \tag{12}$$

Most extensively used kernel functions are radial basis function (RBF) (Eq. (13)) and polynomial function (Eq. (14)) (Gunn, 1998; Müller et al., 2001):

$$K(x_i,\ x_j) = \exp\left(-\|x_i - x_j\|^2 \big/ \sigma^2\right) \tag{13}$$

$$K(x_i,\ x_j) = \left(1 + x_i^{\mathrm{T}} x_j \big/ C\right)^{\gamma} \tag{14}$$

where σ is an independent variable.

This work uses the RBF (Baylar et al., 2009; Xavier-de-Souza et al., 2009; Deng and Yeh, 2010; Chamkalani et al., 2013; Rafiee-Taghanaki et al., 2013; Shokrollahi et al., 2013) kernel, which is suitable for the LSSVM mathematical modeling.

3. Application of LSSVM for predicting pile bearing capacity

3.1. Database used

The application of LSSVM in this work is based on pile load test and CPT data, reported in Alsamman (1995), Eslami (1997) and Alkroosh and Nikraz (2011a, b). The case records were collected from a wide spectrum of geographic locations all over the world. Most of the piles embedded in layered soil of sand and clay. The piles were tested under slow maintained compression loads. According to Ng et al. (2004), the piles were classified into small diameter piles (diameter $<$ 600 mm) and large diameter piles (diameter $>$ 600 mm).

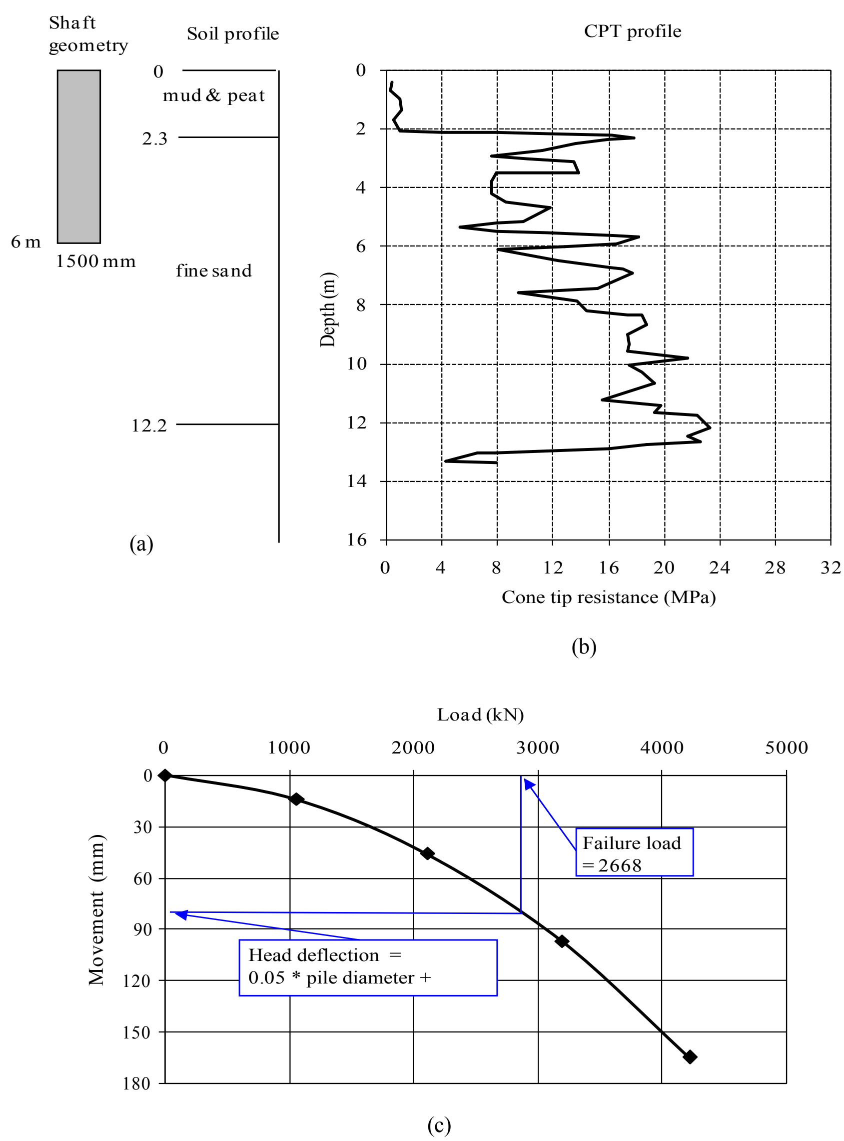

Fig. 1. Summary sheet for selected case record: (a) pile geometry and soil profile; (b) cone tip resistance profile; (c) load-movement plot (Alkroosh and Nikraz, 2011a, b).

3.2. Input and output variables

A proper estimation of bearing capacity of pile foundation requires the identification of the factors that influence the pile-soil interaction. An extensive study of relevant existing literature was carried out and concluded that pile geometry and soil properties are the significant factors that affect the pile bearing capacity.

(1) Pile geometry

All geotechnical engineering sources confirm that pile diameter and length have significant influence on bearing capacity of pile foundations. Therefore, these factors were selected to represent pile geometry for input of LSSVM model (Eslami, 1997; Teh et al., 1997; Abu-Kiefa, 1998).

(2) Soil properties

The soil properties that have principal influence on pile bearing capacity are mainly interpreted as the angle of internal friction, φ, relative density, D_r, earth pressure coefficient, K_0, and cohesion, c. Many researchers have suggested correlations between these factors and CPT data, and proposed solutions for prediction of pile bearing capacity. However, indirect correlations of CPT data with pile bearing capacity have been found to give inaccurate estimate of pile bearing capacity due to considerable uncertainty resulting from the intermediate steps and correlations. Consequently, direct correlations of CPT results with pile bearing capacity have been suggested and found to give better results.

Thus, in this work, the CPT results were directly used as input variables based on the suggestions provided by the current CPT-based methods (e.g. Alsamman, 1995; Eslami, 1997). According to the methods, the cone point resistance measurements are averaged within pile tip influence zone, which is a function of pile diameter, and used for calculating pile tip resistance. For small diameter piles, the influence zone was taken based on the work done by Eslami (1997), while for large diameter piles the influence zone was taken based on Alsamman (1995). The average cone point resistance along pile shaft was included as input, as most of the methods also suggest it to calculate pile shaft resistance.

In fact, the use of sleeve friction in calculation of bearing capacity is not recommended by number of commonly used CPT-based methods, such as Laboratoire Central des Pont et Chausees (LCPC), due to the unreliability of sleeve friction measurements (Alkroosh and Nikraz, 2011a, b).

The mechanical cone point resistance was converted to equivalent electric values using Eq. (15) which was developed by Kulhawy and Mayne (1990):

$$\left(\frac{q_c}{p_a}\right)_{\text{Electric}} = 0.47\left(\frac{q_c}{p_a}\right)^{1.19}_{\text{Mechanical}} \tag{15}$$

where q_c is the cone point resistance; p_a is the atmospheric pressure, which is 101 kPa.

3.3. Pile bearing capacity, Q_u

The interpreted failure load, Q_u, was considered as the pile bearing capacity and calculated according to Alsamman (1995). The failure load was defined as the measured axial load that corresponds to a displacement equal to 5% of pile diameter plus the elastic compression of the pile. Fig. 1 presents the definition of the failure load for a case record selected from the database.

3.4. LSSVM for predicting pile bearing capacity

The data were mapped into range $[-1, 1]$ using below equation:

$$x_n = 2\frac{x - x_{\min}}{x_{\max} - x_{\min}} - 1 \tag{16}$$

Table 1
Summary of database used in the application of LSSVM and the output results (Alkroosh and Nikraz, 2011a, b).

Test No.	Dataset	D (mm)	L (m)	$\bar{q}_{c-tip}$ (MPa)	$\bar{q}_{c-shaft}$ (MPa)	$Q_{u(exp.)}$ (kN)	$Q_{u(pred.)}$ (kN)	$Q_{u(exp.)}/Q_{u(pred.)}$	Test No.	Dataset	D (mm)	L (m)	$\bar{q}_{c-tip}$ (MPa)	$\bar{q}_{c-shaft}$ (MPa)	$Q_{u(exp.)}$ (kN)	$Q_{u(pred.)}$ (kN)	$Q_{u(exp.)}/Q_{u(pred.)}$
1	Training	1100	13	16.2	4	2624	2677	0.98	26	Training	671	10.2	13.7	20.1	4697	4692	1
2	Training	421	5.8	22.9	11.8	912	824	1.11	27	Training	430	8.7	31.7	14.5	516	589	0.88
3	Training	320	10.2	22	7.2	712	764	0.93	28	Training	320	7.7	7.9	2.6	356	371	0.96
4	Testing	457	15.2	1.6	8.1	1423	1388	1.03	28	Testing	399	10	24.6	12.7	756	766	0.99
5	Training	393	6.5	10.1	12.8	738	707	1.04	30	Validation	600	12	21.4	10.8	2687	2466	1.09
6	Training	410	5.6	16.7	15.8	560	729	0.77	31	Testing	600	12	21.3	11.1	2406	2527	0.95
7	Training	320	10.2	14.6	4.5	832	795	1.05	32	Validation	1100	27	7	9.4	8207	8198	1
8	Testing	320	7.7	8.3	2.6	445	392	1.14	33	Training	320	7.7	8.2	2.6	391	386	1.01
9	Validation	403	9.2	13.1	10.3	1352	1098	1.23	34	Testing	400	9.4	2.4	1.4	480	489	0.98
10	Validation	814	24.2	6.5	9.6	5872	5882	1	35	Training	1085	25.1	32	9	7695	7693	1
11	Training	320	10.2	21.9	7.1	818	775	1.06	36	Training	350	15.8	5.1	5.5	840	865	0.97
12	Training	671	13	25.6	17.2	4270	4257	1	37	Training	500	10.2	14.7	3.2	1299	1255	1.03
13	Training	1000	9.5	29.3	5.1	2358	2368	1	38	Training	405	7.9	6.2	12.8	792	750	1.06
14	Training	1000	9	35.9	8.5	3692	3691	1	39	Training	1100	6	21	7.8	2469	2495	0.99
15	Training	840	24.4	47.6	9.2	9653	9644	1	40	Testing	631	18.3	30	11.7	1770	1797	0.98
16	Validation	600	7.2	10.9	7.6	1437	1375	1.05	41	Training	521	8.2	12.8	9.5	1263	1410	0.9
17	Validation	1100	9	15.4	5.4	3247	3173	1.02	42	Training	405	7	17.8	14.3	1294	1084	1.19
18	Training	500	10.2	8.9	2.2	1005	1019	0.99	43	Training	399	7.8	13.1	4.1	578	679	0.85
19	Training	329	6.2	20.7	10.6	605	735	0.82	44	Training	1500	6	10.4	8.5	2669	2679	1
20	Validation	408	5.8	17.6	8.2	765	705	1.09	45	Training	400	7.8	10.6	3.6	543	575	0.94
21	Training	521	8.2	12.9	9.6	1334	1422	0.94	46	Training	320	7.7	8.5	2.6	409	402	1.02
22	Training	1800	11.5	36.6	7.6	7651	7645	1	47	Training	762	16.8	5.9	5.2	3425	3413	1
23	Testing	405	8.4	33.4	11.5	1019	907	1.12	48	Training	430	8.7	26.8	11.7	627	821	0.76
24	Validation	405	10.4	8.9	11.3	1019	1207	0.84	49	Training	329	6.3	25.9	15.6	756	701	1.08
25	Training	399	7.8	12.8	4.4	667	657	1.02	50	Training	1078	13	31	19	8825	8820	1

Note: $\bar{q}_{c-shaft}$ is the weighted average cone point resistance in the pile shaft; $\bar{q}_{c-tip}$ is the weighted average cone point resistance within pile tip influence zone calculated as by Eslami (1997) for small diameter piles and as by Alsamman (1995) for large diameter piles; D is the pile diameter; L is the pile embedment depth; $Q_{u(pred.)}$ and $Q_{u(exp.)}$ are the LSSVM-predicted and experimental pile bearing capacity, respectively.

where x and x_n are the original and normalized desired variables, respectively; x_{min} and x_{max} are the extreme values of the variable x. This pre-processing procedure was applied to obtain the parameters of the LSSVM algorithm. Later, these values were changed to their original values.

In the next step, the database was divided into three subsets including the training, testing and validation sets. The division of database into three subsets was normally performed randomly but considering the statistical consistency of the data of the subsets. For this purpose, 70%, 15% and 15% of the main dataset were randomly selected for building the LSSVM models (i.e. training, testing and validation sets, respectively). The allocated data for each subset are shown in Table 1. The effect of the percent allocation of the three subsets from the database on the accuracy of the final model has been studied (Gharagheizi et al., 2011). As for the distribution of the data through the three subsets, we generally performed many distributions to avoid the local accumulations of the data in the feasible region of the problem. As a result, the acceptable distribution is the one with homogeneous accumulations of the data on the domain of the three subsets (Eslamimanesh et al., 2011).

4. Results and discussion

The ability of the LSSVM to predict the pile bearing capacity was evaluated primarily during the training phase by comparing the measured values of pile bearing capacity with the predicted values obtained by the LSSVM. When training phase was completed, it was expected that the proposed algorithm would correctly reproduce the targeted output values presuming the error is minimal. However, this was insufficient to conclude that the trained LSSVM can predict the pile bearing capacity accurately as this technique, in general, has a high tendency towards over-fitting. Over-fitting refers to the large error in predictions when new data are presented to the trained LSSVM (i.e. ability of the LSSVM to memorize rather than generalize the form of the relationship between input and output data). Generally, over-fitting is expected when data points in training sets are scanty (Das, 2013). To overcome this setback, a testing dataset, not included in training data, was used to examine the predictive ability of LSSVM. If it was found to give accurate predictions, further evaluation was made to verify its generalization ability. This was done by examining the predictions of LSSVM through the use of validation dataset which was neither included among training nor testing datasets. Once the training testing and validation steps were successfully completed, the trained LSSVM can be recommended as a design tool for this type of application.

The predictions of pile bearing capacity by the LSSVM as well as other CPT-based methods including Schmertmann (1978), Bustamante and Gianeselli (1982), Alsamman (1995), and Alkroosh and Nikraz (2011a, b) were evaluated. A brief description of the compared methods is provided in Table 2. The results of comparison are shown numerically in Table 3 and graphically in Figs. 2 and 3. Table 3 also presents the results of statistical analysis used to evaluate the accuracy of the proposed method. Five statistical indices including coefficient of determination, R, mean, μ, standard deviation, σ, mean absolute error, *MAE*, and mean squared error, *MSE*, were used in the analysis. The coefficient of determination is calculated by

$$R = r^2 = \left[\frac{\sum_{i=1}^{n}\left(Q_{mi}-\overline{Q}_m\right)\left(Q_{pi}-\overline{Q}_p\right)}{\sqrt{\sum_{i=1}^{n}\left(Q_{mi}-\overline{Q}_m\right)^2\sum_{i=1}^{n}\left(Q_{pi}-\overline{Q}_m\right)^2}}\right]^2 \tag{17}$$

where r is the coefficient of correlation; Q_{mi} and Q_{pi} are the measured and predicted pile bearing capacity of case i, respectively; $\overline{Q}_m$ and $\overline{Q}_p$ are the averages of measured and predicted pile bearing capacity, respectively; n is the number of observations.

The optimal value of R is unity, which means that a perfect fit is achieved between predicted and measured values. Table 3 shows that the calculated $R = 1$ for the training and validation sets and $R = 0.99$ for the testing set indicate that the LSSVM can achieve an accurate correlation between predicted and measured pile bearing capacity. Table 3 also shows that the proposed method is more accurate than other methods.

Table 2
Some CPT methods versus the LSSVM.

Method	Procedure for calculating pile bearing capacity
Schmertmann (1978)	The end bearing capacity is obtained by $r_t = w(q_{c1}+q_{c2})/2$ where q_{c1} is the average of q_c after elimination of extreme values for a zone ranging from 0.7D to 4D below the pile tip, q_{c2} is the average of q_c for a zone extending to 8D above the pile tip, w is a coefficient ranging from 0.5 to 1 depending on overconsolidation ratio. The unit shaft resistance is determined by $r_s = Cq_c$ where C ranges from 0.0018 to 0.008.
Bustamante and Gianeselli (1982)	Pile bearing capacity is the sum of unit end resistance, q_p, and unit shaft friction, f_p: $q_p = k_{b1}q_{eq(tip)}, f_p = q_{eq(side)}/k_{s1}$ where k_{b1} ranges from 0.15 to 0.6 depending on soil type and installation procedure; q_{eq} is the equivalent average of q_c values of zone ranging from 1.5D below pile tip to 1.5D above pile tip; k_{s1} ranges from 30 to 150 depending on soil type, pile type and installation procedure.
Alsamman (1995)	For piles in cohesionless soil, the unit tip resistance is calculated by $r_t = 0.15\overline{q}_{c(tip)}$ for $\overline{q}_{c(tip)} \le 9.5$ MPa and $r_t = 1.44+0.075[\overline{q}_{c(tip)}-9.5]$ for $\overline{q}_{c(tip)} > 9.5$ MPa where r_t must not exceed 2.87 MPa. The unit side resistance is determined as follows: (1) In sand and silty sand, we have $r_s = 0.015\overline{q}_{c(shaft)}$ for $\overline{q}_{c(shaft)} \le 4.75$ MPa where r_s must not exceed 95 kPa. (2) In gravelly sand and gravel, we have $r_s = 0.02\overline{q}_{c(shaft)}$ for $\overline{q}_{c(shaft)} \le 4.75$ MPa and $r_s = 0.095+0.0025[\overline{q}_{c(shaft)}-4.75]$ for $\overline{q}_{c(shaft)} \ge 4.75$ MPa where r_s must not exceed 130 kPa.
Alkroosh and Nikraz (2011a, b)	In this method, gene expression programming is used to predict pile bearing capacity from CPT data. For bored piles, pile bearing capacity is estimated as $Q_u = L+\frac{D^{1.25}L\overline{q}_{c-shaft}}{250.8}+\frac{L}{265}(D-1171)-2\overline{q}_{c-tip}\overline{q}_{c-shaft}+\overline{q}_{c-shaft}+D+\overline{q}_{c-tip}^{1.33}L-52$

Table 3
Results of evaluating the performance of LSSVM.

Method	Dataset	Coefficient of determination, *R*	Mean, μ	Standard deviation, σ	Mean absolute error, *MAE*	Mean squared error, *MSE*
Proposed method	Training	1	0.98	0.09	50	5699
	Testing	0.99	1.03	0.08	51	4534
	Validation	1	1.04	0.11	110	20,315
Schmertmann (1978)	Training	0.65	1.43	0.61	766	2,601,053
	Testing	0.89	1.11	0.20	168	69,669
	Validation	0.87	1.14	0.34	1076	2,235,985
Bustamante and Gianeselli (1982)	Training	0.8	0.99	0.28	688	1,782,956
	Testing	0.88	0.97	0.19	213	71,214
	Validation	0.97	1.09	0.26	752	1,201,709
Alsamman (1995)	Training	0.89	1.03	0.25	462	789,032
	Testing	0.88	1.07	0.23	214	63,969
	Validation	0.97	1.14	0.23	370	251,119
GEP by Alkroosh and Nikraz (2011a, b)	Training	0.95	1.03	0.28	421	374,303
	Testing	0.99	0.96	0.09	51	3991
	Validation	0.99	1.16	0.15	419	243,574

Then the mean and standard deviation were respectively calculated by

$$\mu = \frac{1}{n}\sum_{1}^{n}\frac{Q_m}{Q_p} \tag{18}$$

$$\sigma = \sqrt{\frac{\sum_{1}^{n}\left(\frac{Q_m}{Q_p}-\mu\right)^2}{n-1}} \tag{19}$$

If the mean value is unity, the predicted values of pile bearing capacity are on average equivalent to the measured values. A mean value of less than unity indicates that the method tends to under-predict the pile bearing capacity. Conversely, more than unity is an indication of over-prediction. The optimum value of standard deviation is zero; the closer the standard deviation approaches to zero, the greater the accuracy is. As presented in Table 3, the calculated mean values of the proposed method are 0.98, 1.03 and 1.04 for the training, testing and validation sets, respectively, which indicate that the LSSVM possesses a high capability in predicting pile bearing capacity. The results also indicate that, on average, the proposed method may tend to over-predict the pile bearing capacity in testing and validation sets, whereas under-predict it in training set. In comparison with other methods, the calculated mean values indicate that the LSSVM provides better estimate for pile bearing capacity.

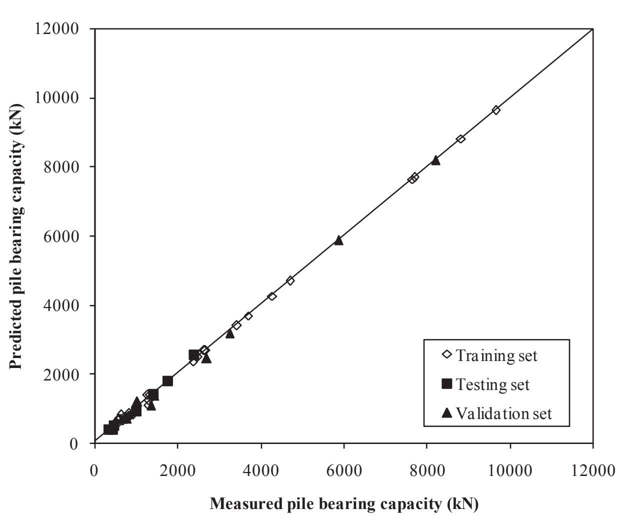

Fig. 2. Performance of the LSSVM model in training, testing and validation sets.

The predictive ability of the LSSVM is also evaluated by calculating the error. Table 3 shows that the calculated *MSE* and *MAE* are low when using the proposed method to predict pile bearing capacity. The table also shows that the proposed method has the lowest error in comparison with other methods.

As shown in Fig. 2, the strong ability of LSSVM to predict the pile bearing capacity is obvious. All the points in training, testing and validations sets are situated on or very close to the line of equality, suggesting that the proposed method is accurate in predicting the pile bearing capacity. Fig. 3 shows that the LSSVM achieves very high correlation between predicted and measured values and it has the lowest scatter around the line of equality in comparison with other methods. From Fig. 3, it is most evident that the LSSVM provides best prediction of total pile bearing capacity.

In order to verify whether or not the results of the proposed method agree with those in the existing literature and experimental results, the sensitivity analysis was carried out. In this analysis, values of one input variable were allowed to change within the range of the training data whereas the values of other input variables were set constant. The data were then input in the developed LSSVM and the outputs were evaluated. The results, as presented in Fig. 4, have shown that the proportion between each input variable and the predicted pile bearing capacity is incremental, and the variations of pile diameter and average cone point resistance within pile tip influence zone and along pile shaft have great influence on pile bearing capacity. The variation of pile length slightly affects pile bearing capacity.

5. Conclusions

The results of this study indicate that the proposed LSSVM can accurately predict the bearing capacity of bored piles from CPT data. The statistical analysis of the results reveals that this technique can obtain coefficient of determination and mean close to unity for all training, testing and validation sets. It has also achieved standard deviation less than 0.2 and low error values.

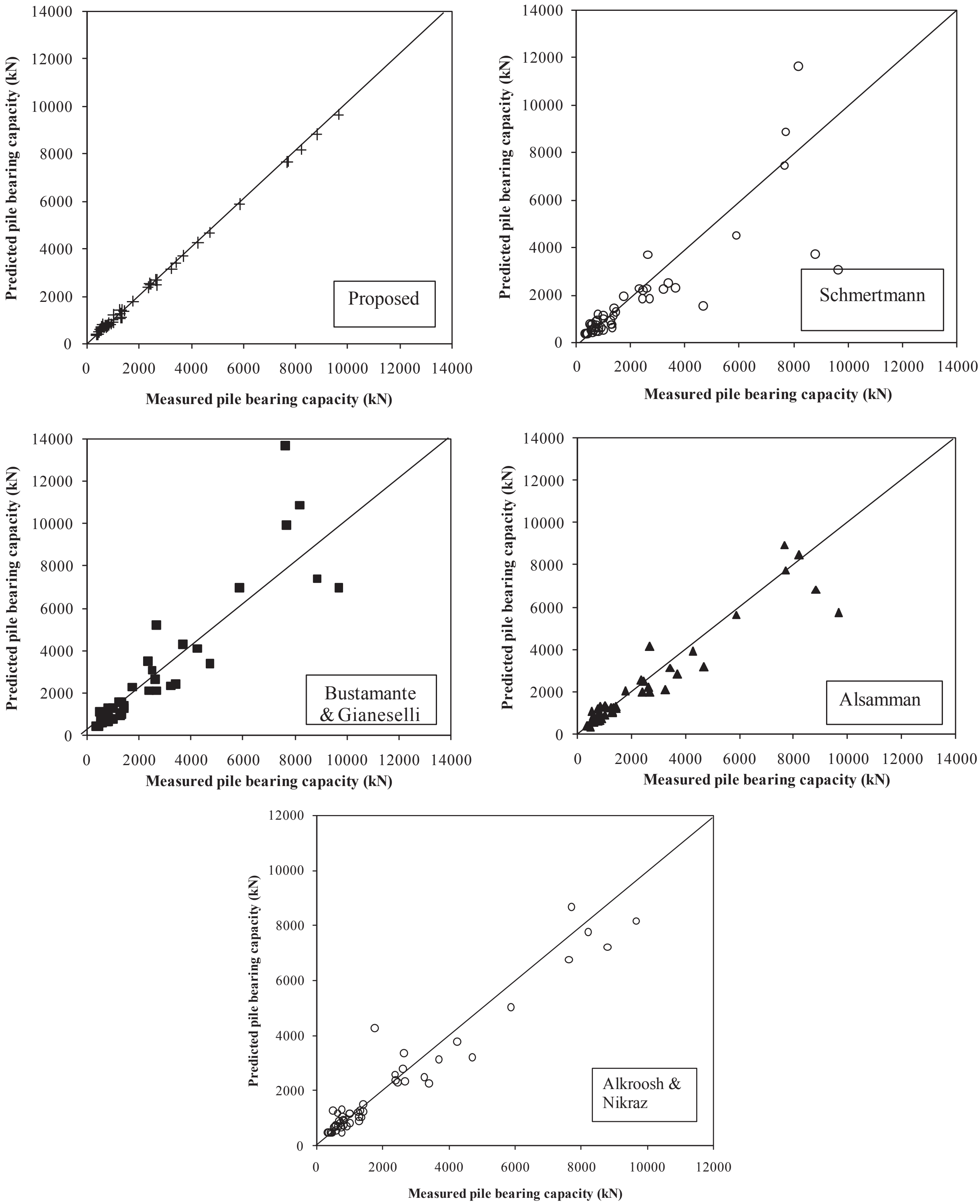

Fig. 3. Comparisons of the performance of the LSSVM model with CPT-based methods.

The comparisons of the predictive ability of LSSVM with traditional CPT-based methods as well as the GEP model have shown that this technique predicts pile bearing capacity better. The relationship of pile bearing capacity with each of pile diameter, length, average of cone point resistance is incremental. The average cone point resistance within pile tip zone is the most influential factor on pile bearing capacity. The output of this study suggests that LSSVM can be used as a design tool for

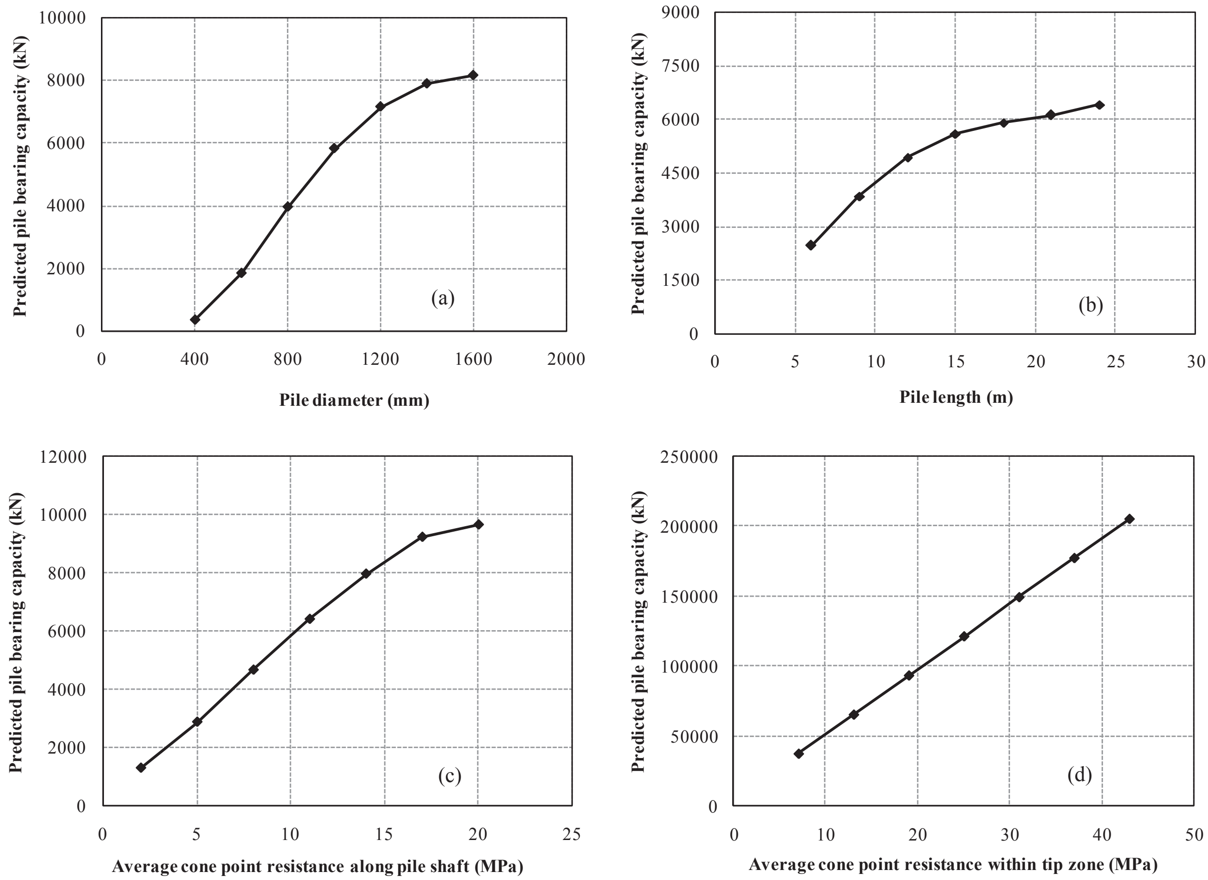

Fig. 4. Effects of variations of input variables on pile bearing capacity: (a) pile diameter; (b) pile length; (c) average cone point resistance along shaft; (d) average cone point resistance within tip zone.

predicting the bearing capacity of bored piles installed in sand and mixed soils.

Conflict of interest

The authors confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

References

Abu-Farsakh M, Titi H. Assessment of direct cone penetration test methods for predicting the ultimate capacity of friction driven piles. Journal of Geotechnical and Environmental Engineering 2004;130(9):935–44.

Abu-Kiefa MA. General regression neural networks for driven piles in cohesionless soils. Journal of Geotechnical and Geoenvironmental Engineering 1998;124(12):1177–85.

Alavi AH, Aminian P, Gandomi AH, Esmaeili MA. Genetic based modeling of uplift capacity of suction caissons. Expert Systems with Applications 2011;38(10):12608–18.

Alkroosh I, Nikraz H. Correlation of pile axial capacity and CPT data using gene expression programming. Geotechnical and Geological Journal 2011a;29(5):725–48.

Alkroosh I, Nikraz H. Predicting axial capacity of driven piles in cohesive soils using intelligent computing. Engineering Applications of Artificial Intelligence 2011b;25(3):618–27.

Alsamman O. The use of the CPT for calculating axial capacity of drilled shafts. PhD Thesis. Urbana, USA: University of Illinois at Urbana; 1995.

Amendolia SR, Cossu G, Ganadu ML, Golosio B, Masala GL, Mura GM. A comparative study of K-nearest neighbour, support vector machine and multi-layer perceptron for Thalassemia screening. Chemometrics and Intelligent Laboratory Systems 2003;69(1–2):13–20.

Ardalan H, Eslami A, Nariman-Zadeh N. Piles shaft capacity from CPT and CPTU data by polynomial neural networks and genetic algorithms. Computers and Geotechnics 2009;36(4):616–25.

Baylar A, Hanbay D, Batan M. Application of least square support vector machines in the prediction of aeration performance of plunging overfall jets from weirs. Expert Systems with Applications 2009;36(4):8368–74.

Bazzani A, Bevilacqua A, Bollini D, Brancaccio R, Campanini R, Lanconelli N, Riccardi A, Romani D. An SVM classifier to separate false signals from microcalcifications in digital mammograms. Physics in Medicine and Biology 2001;46(6):1651.

Briaud J. Evaluation of cone penetration test methods using 98 pile load tests. In: Proceedings of the 1st International Symposium on Penetration TestingVol. 2. Rotterdam, Netherlands: A.A. Balkema; 1988. p. 687–97.

Bustamante M, Gianeselli L. Pile bearing capacity prediction by means of static penetrometer CPT. In: Proceedings of the 2nd European Symposium on Penetration Testing, ESOPT-II, Amsterdam, The Netherland; 1982. p. 493–500.

Cai G, Liu S, Tong L, Du G. Assessment of direct CPT and CPTU methods for predicting the ultimate bearing capacity of single piles. Engineering Geology 2008;104(3–4):211–22.

Chamkalani A, Amani M, Kiani MA, Chamkalani R. Assessment of asphaltene deposition due to titration technique. Fluid Phase Equilibria 2013;339:72–80.

Chen TS, Chen J, Lin YC, Tsai YC, Kao YH, Wu KS. A novel knowledge protection technique based on support vector machine model for anti-classification. In: Electrical engineering and control. Berlin, Germany: Springer; 2011. p. 517–24.

Cortes C, Vapnik V. Support-vector networks. Machine Learning 1995;20(3):273–97.

Das S, Samui P, Kim D, Sivakugan N, Biswal R. Lateral displacement of liquefaction induced ground using least square support vector machine. International Journal of Geotechnical Earthquake Engineering 2011b;2(2):29–39.

Das S, Samui P, Kothari D. Site characterization model using machine learning. In: Machine tools: design, reliability and safety. New York, USA: Nova Science Publishers Inc.; 2011a. p. 175–85.

Das SK, Basudhar PK. Undrained lateral load capacity of piles in clay using artificial neural network. Computers and Geotechnics 2006;33(8):454–9.

Das SK. Artificial neural networks in geotechnical engineering: modeling and application issues. In: Yan X, Gandomi AH, Talatahati S, Alavi AH, editors. Metaheuristics in water; 2013. p. 231–70.
Deng S, Yeh TH. Applying least squares support vector machines to the airframe wing-box structural design cost estimation. Expert Systems with Applications 2010;37(12):8417–23.
Eslami A. Bearing capacity of piles from cone penetration data. PhD Thesis. Ottawa, Canada: University of Ottawa; 1997.
Eslamimanesh A, Gharagheizi F, Mohammadi AH, Richon D. Phase equilibrium modeling of structure H clathrate hydrates of methane+water "insoluble" hydrocarbon promoter using group contribution-support vector machine technique. Industrial and Engineering Chemistry Research 2011;50(20): 12807–14.
Gandomi AH, Alavi AH. A new multi-gene genetic programming approach to nonlinear system modeling. Part II: geotechnical and earthquake engineering problems. Neural Computing and Applications 2012;21(1):189–201.
Gandomi AH. Nonlinear genetic based models for prediction of flow number of asphalt mixtures. Journal of Materials in Civil Engineering 2011;23(3):248–63.
Gharagheizi F, Eslamimanesh A, Farjood F, Mohammadi AH, Richon D. Solubility parameters of nonelectrolyte organic compounds: determination using quantitative structure-property relationship strategy. Industrial and Engineering Chemistry Research 2011;50(19):11382–95.
Gunn SR. Support vector machines for classification and regression. Technical report. Southampton, UK: Faculty of Engineering, Science and Mathematics, School of Electronics and Computer Science, University of Southampton; 1998.
Kulhawy F, Mayne P. Manual on estimating soil properties for foundation design. Palo Alto, CA, USA: Electric Power Research Institute; 1990.
Minoux M. Mathematical programming: theory and algorithms. New York, USA: Wiley; 1986.
Müller KR, Mika S, Rätsch G, Tsuda K, Schölkopf B. An introduction to kernel-based learning algorithms. IEEE Transactions on Neural Networks 2001;12(2):181–201.
Ng C, Simons W, Menzies B. Soil-structure engineering of deep foundations, excavations and tunnels. London, UK: Thomas Telford Ltd.; 2004.
Ornek M, Laman M, Demir A, Yildiz A. Prediction of bearing capacity of circular footings on soft clay stabilized with granular soil. Soils and Foundations 2012;52(1):69–80.
Rafiee-Taghanaki S, Arabloo M, Chamkalani A, Amani M, Zargari MH, Adelzadeh MR. Implementation of SVM framework to estimate PVT properties of reservoir oil. Fluid Phase Equilibria 2013;346:25–32.
Rezania M, Javadi A. A new genetic programming model for predicting settlement of shallow foundations. Canadian Geotechnical Journal 2007;44(12):1462–73.
Roberston P, Campanella R, Davies M, Sy A. Axial capacity of driven piles in deltaic soils using CPT. In: Proceedings of the 1st International Symposium on Penetration TestingVol. 2. Rotterdam, Netherlands: A.A. Balkema; 1988. p. 919–28.
Samui P, Kothari DP. Utilization of a least square support vector machine (LSSVM) for slop stability analysis. Scientia Iranica 2011;18(1):53–8.
Schmertmann JH. Guidelines for cone penetration test, performance and design. Washington D.C., USA: US Department of Transportation; 1978.
Shahin M. Intelligent computing for modeling axial capacity of pile foundations. Canadian Geotechnical Journal 2010;47(2):230–43.
Shokrollahi A, Arabloo M, Gharagheizi F, Mohammadi AH. Intelligent model for prediction of CO_2–reservoir oil minimum miscibility pressure. Fuel 2013;112: 375–84.
Suykens JA, Van Gestel T, De Brabanter J, De Moor B, Vandewalle J. Least squares support vector machines. Singapore: World Scientific Publishing Company; 2002.
Suykens JA, Vandewalle J. Least squares support vector machine classifiers. Neural Processing Letters 1999;9(3):293–300.
Tarawneh B. Pipe pile setup: database and prediction model using artificial neural network. Soils and Foundations 2013;53(4):607–15.
Teh CI, Wong KS, Goh AT, Jaritngam S. Prediction of pile capacity using neural networks. Journal of Computing in Civil Engineering 1997;11(2):129–38.
Übeyli ED. Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals. Expert Systems with Applications 2010;37(1):233–9.
Xavier-de-Souza S, Suykens JA, Vandewalle J, Bolle D. Coupled simulated annealing. IEEE Transactions on Systems, Man, and Cybernetics. Part B: Cybernetics 2009;40(2):320–35.

Dr. Iyad Alkroosh began his career as a civil engineer after graduation from University of Technology Baghdad in 1987. He received his MSc and PhD degrees in civil engineering from Curtin University, Western Australia in 2000 and 2011, respectively. Currently he has been a lecturer at Department of Civil Engineering, College of Engineering, Al-Qadisiya University, Iraq. Since 2011, Dr. Alkroosh has been working as a consultant for the Scientific and Engineering Bureau, Al-Qadisiya University. His field of interest includes but not limited to modeling using artificial intelligence, load transfer behavior of piles, soil improvement, and structural design of shallow and deep foundations. Dr. Alkroosh is the author of 15 publications in international journals and conferences.

Mohammad Bahadori is a research student in the Department of Soil Science and Engineering at the University of Tehran, Iran. He received his BSc degree from Shahaid Chamran University, Ahwaz, Iran and his MSc degree from the University of Tehran, Iran.

Dr. Hamid Nikraz is a professor at the Department of Civil Engineering, School of Engineering, Curtin University, Australia. Prior to his academic career he had many years' experience working in industry in the Middle East, Europe and Australia. He has provided numerous consultancy services in his field to over 200 organizations nationally and internationally. He obtained his PhD at Curtin University of Technology in 1989. Dr. Nikraz is a Fellow Member of the Institution of Engineers, Australia. He has guided 75 PhD theses, 30 Master students and published over 500 publications in various journals, books and conference proceedings. Prof. Nikraz has initiated research works in the area of Geopolymers, Concrete, Soil Stabilisation and Pavement Engineering. Prof. Nikraz is also the reviewer of many national and international journals.

Alireza Bahadori, PhD, P.Eng, MIEAust., is a research staff member and lecturer in the School of Environment, Science and Engineering at Southern Cross University, Lismore, NSW, Australia. He received his PhD from Curtin University, Perth, Western Australia. During the past twenty years, Dr. Bahadori has held various process and petroleum engineering positions and was involved in many large-scale projects at National Iranian Oil Co. (NIOC), Petroleum Development Oman (PDO), and Clough AMEC Pty Ltd. He is the author of several articles and books. His books have been published by multiple major publishers, including Elsevier, Springer, John Wiley & Sons, and Taylor and Francis Group. Dr. Bahadori is the recipient of the highly competitive and prestigious Australian Government's Endeavour International Postgraduate Research Award as part of his research in oil and gas area. He also received a Top-Up Award from the State Government of Western Australia through Western Australia Energy Research Alliance in 2009. Dr. Bahadori serves as a member of the editorial board and reviewer for a large number of journals. He is member of the Institution of Engineers Australia as a professional engineer and IChemE, UK.