Ocean Systems Engineering, *Vol. 5, No. 1 (2015) 41-54*
DOI: http://dx.doi.org/10.12989/ose.2015.5.1.041

Prediction of uplift capacity of suction caisson in clay using extreme learning machine

Pradyut Kumar Muduli[1a], Sarat Kumar Das[*1], Pijush Samui[2b] and Rupashree Sahoo[1c]

[1]*Department of Civil Engineering, National Institute of Technology, Rourkela, Odisha, India*
[2]*Centre for Disaster Mitigation and Management, VIT University, Vellore-632014, India*

(Received June 29, 2013, Revised March 2, 2015, Accepted March 5, 2015)

Abstract. This study presents the development of predictive models for uplift capacity of suction caisson in clay using an artificial intelligence technique, extreme learning machine (ELM). Other artificial intelligence models like artificial neural network (ANN), support vector machine (SVM), relevance vector machine (RVM) models are also developed to compare the ELM model with above models and available numerical models in terms of different statistical criteria. A ranking system is presented to evaluate present models in identifying the 'best' model. Sensitivity analyses are made to identify important inputs contributing to the developed models.

Keywords: suction caisson; uplift capacity; extreme learning machine; support vector machine; artificial neural network; statistical performance criteria

1. Introduction

Caissons are generally used as deep foundations of transmission towers and anchors for the offshore facilities for less construction time and their suitability for both static and dynamic loads. The total uplift capacity of caisson depends upon passive suction under caisson-sealed cap, self weight of the caisson, frictional resistance along the soil- caisson interface, submerged weight of the soil plug inside the caisson and uplift soil (reverse end bearing) bearing pressure (Albert *et al.* 1987). Hence, suction caisson becomes more effective particularly in clayey soil. Various methods are in use to find the uplift capacity of suction caisson. Various studies using upper bound analysis (Clukey *et al.* 1995), finite element method (Whittle and Kavvadas 1994, El-Gharbawy and Olson 2000, Zdravkovic *et al.* 2001, Cao *et al.* 2001, 2002a, b), laboratory model (Goodman *et al.* 1961, Larsen 1989, Steensen-Bach 1992, Datta and Kumar 1996, Singh *et al.* 1996, Rao *et al.* 1997a, 1997b), centrifuge model (Clukey and Morrison 1993, Clukey *et al.* 1995) and prototype model tests (Hogervorst 1980, Tjelta *et al.* 1986, Dyvik *et al.* 1993, Cho *et al.* 2002) have been done to

*Corresponding author, Associate Professor, E-mail: saratdas@rediffmail.com
[a] Research Scholar, E-mail: pradyut.muduli@gmail.com
[b] Associate Professor, E-mail: pijush.phd.@gmail.com
[c] Research Scholar, rupashree.lucy@gmail.com

http://www.techno-press.org/?journal=ose&subpage=7 ISSN: 2093-6702 (Print), 2093-677X (Online)

understand the axial and lateral load capacity of suction caisson for static and cyclic load under different soil conditions. Though the finite element method (FEM) along with the laboratory and centrifuge tests are the most popular methods in predicting the uplift capacity of suction caisson, but soil properties are highly variable over short distances. Thus, developing a sufficiently accurate constitutive model for a detailed FEM analysis requires extensive site characterization effort. And also desired constitutive modeling of clayey soil is very difficult, even with considerable laboratory testing. Hence, although field tests are very expensive, various tests have been conducted to find out the feasibility of constructing suction caisson in various soil types (Cho *et al.* 2002). However, several issues and uncertainties related to uplift capacity estimation and failure mechanisms are still unresolved, for which accurate empirical models are required for prediction of the uplift capacity of suction caission.

Artificial intelligence techniques such as artificial neural networks (ANNs) and support vector machine (SVM) are considered as alternate statistical methods and found to be more efficient compared to statistical methods (Das and Basudhar 2006, Das *et al.* 2011). The ANN is found to be more efficient for prediction of ultimate load capacity of driving piles in cohesion-less soil (Goh 1996) and also for the prediction of lateral load capacity of pile (Das and Basudhar 2006)compared to traditional methods. The performance of SVM model was found to be better than that of the ANN model for prediction of frictional resistance of the pile in clay (Samui 2008).

Rahman *et al.* (2001) used artificial neural network (ANN) model to predict the uplift capacity of suction caisson in clay. The performance of ANN model was found to be better than that of the FEM model in terms of correlation coefficient (R). However, the only R value is not sufficient and its high value need not show good prediction (Das and Sivakugan 2010). It is very difficult to assess in terms of under or over prediction of the estimation only from R value. The most important problem associated with efficient implementation of ANN is generalization for some complex problems, and magnitude of weight is one of the reasons for poor generalization (Bartlett 1998).The methods like Bayesian regularization neural network (BRNN) (Das and Basduahr 2008) has been used to consider the magnitude of weights as the part of the error function. Hence, in the present study BRNN model was developed to compare the results of the ANN model as per Rahman *et al.* (2001). The error function in ANN is a nonlinear function; hence the chance of local minima cannot be avoided using traditional optimization algorithms, as the algorithms are the initial point dependent. As the error function in SVM is a convex function, traditional optimization problems can be effectively used to avoid the local minima. Though, SVM has better generalization (Samui 2008) compared to ANN, error parameter 'C' and 'e' are to be found out by trial and error. Using the above database, Muduli *et al.* (2013) observed that prediction model using genetic programming (GP) is more efficient compared to ANN and SVM model. It may be mentioned here that unlike ANN, GP is a 'grey box' model (Giustolisi *et al.* 2007), but the model parameters are found out by nontraditional optimization method, genetic algorithm and it is very difficult to explain the development of model.

In the recent past a modified learning algorithm called extreme learning machine (ELM) has been proposed by Huang *et al.* (2006) for single hidden layer feed forward neural network (SLFN). This learning algorithm for SLFN is very fast and hence named as extreme learning machine (ELM). In ELM the hidden nodes are randomly selected and output weights are computed analytically to avoid the problem of local optima. The ELM and its variants have been used for different large complex applications (Wang and Huang 2005, Huang *et al.* 2006a, Huang *et al.* 2006b, Huang and Chen 2008, Huang *et al.* 2010, Huang *et al.* 2012) with success and are found to be efficient compared to ANN and SVM (Huang *et al.* 2006). It has been shown that this new algorithm can produce good generalization performance and can learn faster than conventional learning algorithms of feed forward neural net works (Huang *et al.* 2006).However, its use in geotechnical engineering is limited (Das and Muduli 2011). Like any other numerical methods, it needs critical evaluation while applying to a new problem. Thus, the efficacy of the model is to be compared with other artificial intelligence models like ANN, SVM, relevance vector machine (RVM) and genetic programming (GP) models in terms of different statistical performance criteria.

In the present study prediction of uplift capacity of suction caisson in clay under undrained condition has been developed by using ELM, ANN, SVM and RVM models. Different statistical criteria like correlation coefficient (R), Nash-Sutcliff coefficient of efficiency (E), root mean square error (RMSE), the average absolute error (AAE), maximum absolute error (MAE) and normalized mean biased error (NMBE) are used to compare the developed ELM, ANN, SVM and RVM models with the FEM and GP models as available in the literature (Rahman *et al.* 2001, Muduli *et al.* 2013). A ranking system (Abu-Farsakh and Titi 2004) using rank index (RI) has also been followed to compare the performance of different models on the basis of four criteria. : (i) the best fit calculations (R and E) for predicted uplift capacity (Q_p) and measured capacity (Q_m), (ii) arithmetic calculations (mean, μ and standard deviation, σ) of the ratio, Q_p/Q_m (R2) (iii) 50% and 90% cumulative probability (P_{50} and P_{90}) of the ratio, $Q_p/Q_{m.}$ (R3) and (iv) probability of uplift capacity within ± 20% accuracy level in percentage using histogram and lognormal probability distribution of Q_p/Q_m (R4).

2. Methodology

In geotechnical engineering, the ANN has been extensively used with a few applications of SVM and RVM and hence these techniques are not discussed in detail. The application of ELM in geotechnical engineering is new and hence is elaborated as follows.

2.1 Extreme learning machine (ELM)

Huang *et al.* (2006) proved that the input weights and hidden layer biases of SLFN can be randomly assigned if the activation functions in the hidden layer are infinitely differentiable which is true as sigmoid activation function is generally used in ANN. The SLFN can simply be considered as linear system and the output weights (linking the hidden layer to output layer) can be analytically determined through the inverse operation of the hidden layer output matrices. Mathematically above concept can be described for the standard SLFN with L hidden nodes as follows

$$o_j = \sum_{i=1}^{L} \beta_i f\left(w_i \cdot x_j + b_i\right) \tag{1}$$

$$t_j = \sum_{i=1}^{L} \beta_i f\left(w_i \cdot x_j + b_i\right)$$

$$j = 1, \ldots, N \tag{2}$$

By approximating these N samples with zero error means that minimum norm least-squares solution of general linear system can be applied and presented as

$$\sum_{j=1}^{L} \left\| o_j - t_j \right\| = 0 \tag{3}$$

where $\|.\|$ is a norm in Euclidean space.

where, $f(x)$= activation function, $w_i=[w_{i1}, w_{i2}, \ldots, w_{in}]^T$ = weight vector connecting the i^{th} hidden node and the input nodes, $\beta_i=[\beta_{i1}, \beta_{i2}, \ldots, \beta_{im}]^T$ = weight vector connecting the i^{th} hidden node and the output nodes, b_i=the bias of the i^{th} hidden node, x_j = normalized input variable at j^{th} input node in the range [0,1], , j=1,…,N, N=number of arbitrary distinct training sample (x_i, t_i), $x_i=[x_{i1}, x_{i2},\ldots,x_{in}]^T \in R^n$ and $t_i=[t_{i1}, t_{i2},\ldots,t_{im}]^T \in R^m$,

As the output nodes are chosen linear and the above N equations can be presented compactly as

$$T = H\beta \tag{4}$$

where, H ($w_1,\ldots,w_L, b_1,\ldots,b_L, x_1,\ldots, x_N$)

$$= \begin{bmatrix} f(w_1.x_1 + b_1)\ldots f(w_L.x_1 + b_L) \\ . \\ . \\ . \\ f(w_1.x_N + b_1)\ldots f(w_L.x_N + b_L) \end{bmatrix}_{N\times L} \tag{5}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ . \\ . \\ . \\ \beta_L^T \end{bmatrix}_{L\times m} \qquad T = \begin{bmatrix} t_1^T \\ . \\ . \\ . \\ t_N^T \end{bmatrix}_{N\times m} \tag{6}$$

where, H is called hidden layer output matrix of the neural network , β is the output weight matrix and T is the output matrix. For a given training set,(x_i, t_i), activation function, f(x),and hidden node number, L, ELM algorithm can be summarized as

1. Input weight (w_i) and bias (b_i) for i=1,…, N is randomly assigned.
2. Hidden layer output matrix H is calculated as given by Eq. (5).
3. Then, the output weight (β) is calculated as presented in Eq. (7).

$$\beta = H^{-1}T \tag{7}$$

where H^{1} is the Moore-Penrose generalized inverse of matrix H (Rao and Mitra 1971). The resolution of a general linear system $Ax = y$, where A may be a singular and may not even be square, can be made simple by use of Moore-Penrose generalized inverse. A matrix B of order n x m is the Moore-Penrose generalized inverse of matrix A of order m x n, if

$$ABA = A;\ BAB = B;\ (AB)^T = AB;\ (BA)^T = BA$$

4. Once β is calculated the output can be predicted using Eq. (1).

The detailed methodology is presented in Huang *et al.* (2006). In the present study ELM model is developed using Matlab (Math Work Inc 2005). The four input variables used for the ELM model in this study are L/d (L is the embedded length of the caisson and d is the diameter of caisson), undrained shear strength of soil at the depth of the caisson tip (S_u), D/L (D is the depth of the load application point from the soil surface), inclined angle (θ) and load rate parameter (T_k). The output of the ELM model is Q_p. So, in this study, $x = [L/d, S_u, D/L, \theta, T_k]$ and $t = [Q_p]$

2.2 *Support vector machine (SVM)*

Support Vector Machine (SVM) has originated from the concept of statistical learning theory pioneered by Boser *et al.* (1992). More details can be found in literature on SVM (Boser *et al.* 1992, Vapnik 1998). This study uses the SVM as a regression technique by introducing a ε-insensitive loss function. Considering a set of training data $\{(x_1, y_1), \ldots, (x_l, y_l)\}$, $x \in R^n$, $y \in r$. Where x is the input, y is the output, R^n is the N-dimensional vector space and r is the one-dimensional vector space. The SVM has also been used different geotechnical engineering applications (Samui *et al.* 2008, Das *et al.* 2011).

In the present study the four input variables used for the SVM model in this study are L/d, S_u, D/L, θ and T_k. The output of the SVM model is Q_p. So, in this study, $x = [L/d, S_u, D/L, \theta, T_k]$ and $y = [Q_p]$

Following the methodology as described in Das *et al.* (2011) the final equation of SVM can be written as

$$f(x) = \sum_{i=1}^{nsv} \left(\alpha_i - \alpha_i^*\right) K\left(x_i . x_j\right) + b \tag{8}$$

where α_i, α_i^* are the Lagrangian Multipliers, nsv is the number of support vectors and K ($x_i.x_j$) is kernel function. This study uses radial basis function ($K(x_i, x) = \exp\left\{-\frac{(x_i - x)(x_i - x)^T}{2\sigma^2}\right\}$, where σ is width of radial basis function) as a kernel function.

2.3 *Relevance vector machine (RVM)*

The RVM, introduced by Tipping (2001), is a sparse linear model. Let $D = \{(x_i, t_i), i = 1, \ldots, N\}$ be a dataset of observed values. Where x_i=input, t_i=output, $x_i \in R^n$ and $t_i \in r$. In this study, as discussed above $x = [L/d, S_u, D/L, \theta, T_k]$ and $t = [Q_p]$.

$\mathbf{N}(y(x_n), \sigma^2)$ is the normal distribution with mean $y(x_n)$ and variance σ^2. y(x) can be expressed as a linearly weighted sum of M nonlinear fixed basis function,

$$\{\Phi_j(x) \mid j = 1, \ldots, M\}:$$

$$y(x; w) = \sum_{i=1}^{M} w_i \Phi_i(x) = \mathbf{\Phi} w \tag{9}$$

The likelihood of the complete data set can be written as given below following Berger (1985)

$$p\left(t \mid w,\sigma^2\right) = \left(2\pi\sigma^2\right)^{-N/2} \exp\left\{-\frac{1}{2\sigma^2}\left\|t-\Phi w\right\|^2\right\} \tag{10}$$

Where $w = \left(w_0,...,w_N\right)$ is a vector of weights, $t = \left(t_1...,t_N\right)^T$ is a vector of outputs and

$$\Phi^T = \begin{bmatrix} 1 & K(x_1,x_1) & K(x_1,x_2) & \cdots & K(x_1,x_n) \\ 1 & K(x_1,x_2) & K(x_2,x_2) & \cdots & K(x_2,x_n) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & K(x_n,x_1) & K(x_n,x_2) & \cdots & K(x_n,x_n) \end{bmatrix}$$, where $K(x_i,x_n)$ is a kernel function.

This article adopts radial basis function as kernel function.
The details abut RVM implementation has been described in Tipping (2001).

2.4 Artificial neural network (ANN)

In the present study the ANN model has been trained using Bayesian regularization method as described in Das and Basudhar (2008) and termed as BRNN.

3. Database used

In the present study the databases as per Rahman *et al*. (2001) is used which contains information about L/d , S_u, D/L , θ, T_k and measured uplift capacity of caisson (Q_m). Out of the mentioned 62 data, 51 data are selected for training and remaining 11 data are used for testing the developed ELM and other models. The data was normalized in the range 0 to 1 to avoid the dimensional effect of input parameters.

4. Results and discussion

The best ELM model for the prediction of uplift capacity of suction caisson is obtained by varying the number of hidden layer neurons. Fig. 1 shows the plot of RMSE value versus hidden nodes and the best model was observed with a 20 hidden nodes SLFN (architecture of 5-20-1) after several trials with different number of hidden nodes. As it is important that the efficiency of models should be compared in terms of testing data than that with training data (Das and Basudhar 2008), in this study the comparison of the methods are done for the testing data only. Fig. 2 shows the performance of predicted and observed values of uplift capacity of suction caisson for ELM and other models for testing data. There is less scatter in the data for the ELM model compared to the other models. Table 1 shows the statistical performance in terms of R, E, AAE, MAE, RMSE and NMBE for the developed ANN, SVM , RVM and ELM models along with the results of FEM and ANN models of Rahman *et al.* (2001) and GP model of Muduli *et al.* (2013) for testing data set. It can be seen that though the R value as per FEM and ANN model of Rahman *et al.* (2001) are comparable, FEM model is found to better than ANN models based on other statistical criteria.

Hence, in the present study separate ANN (BRNN) model was developed to improve the performance of ANN based model. Based on the statistical parameters considered in the present study, the ELM model is found to be comparable with GP model of Muduli *et al.* (2013) and better than all other (ANN, SVM and RVM) models, which indicate the robustness of the model.

While describing prediction of pile load capacity based on cone penetration test (CPT) Briaud and Tucker (1988) have emphasized that other statistical criteria should be used along with the correlation coefficient (R). Abu-Farsakh and Titi (2004) and Das and Basudhar (2006) have used the mean (μ) and standard deviation (σ) of the ratio of predicted pile capacity (Q_p) to the measured pile capacity (Q_m) as important parameters in evaluating different models. The mean (μ) and standard deviation (σ) of Q_p/Q_m are important indicators of the accuracy and precision of the prediction method. Under ideal condition an accurate and precise method gives the mean value as 1.0 and the standard deviation to be 0. The μ value greater than 1.0 indicates over prediction and under prediction otherwise. The best model is represented by μ value close to 1.0 and σ close to 0. In present study for testing data the μ (1.069) and σ (0.241) of Q_p/Q_m for the ELM model are comparable with the GP [μ (1.026) and σ (0.191) and SVM [μ (0.956) and σ (0.248)] models but better than those of ANN and other models as presented in Table 2.The other criterion like cumulative probability of the ratio, Q_p/Q_m , has also been considered for the evaluation of different models following Abu-Farsakh and Titi (2004) and Das and Basudhar (2006). If the computed value of 50% cumulative probability (P_{50}) is less than unity, under prediction is implied; values greater than unity indicates over prediction. The 'best' model is corresponding to the P_{50} value close to unity. The 90% cumulative probability (P_{90}) reflects the variation in the ratio, Q_p /Q_m, for the total observations. The model with P_{90} close to 1.0 is a better model (Das and Basudhar 2006).

Fig. 3 shows the cumulative probability plots of Q_p /Q_m for different methods for the testing dataset. It can be seen that ELM model is found to be the best model (P_{50}=1.001) followed by ANN (0.960),GP (0.95) and FEM (1.050) as the values are close to unit. The corresponding SVM and RVM values are found to be 0.890 and 0.920 respectively showing under prediction. However based on the P_{90} value, ELM (1.220), GP (1.38) and SVM (1.380) models are found to be better than RVM (1.600), ANN (2.100) and FEM (2.100) models. The lognormal distributions of the Q_p /Q_m for different models of the testing data are shown in Fig. 4. Based on the plots it can be seen that ELM model is comparable with the GP model, but better than FEM, ANN, SVM and RVM models according to the criterion i.e., probability of uplift capacity within ± 20% accuracy level is concerned as the shaded area under the lognormal distribution plot of Q_p/Q_m for ELM is more than that of the other models except the GP model. The probability of uplift capacity within ± 20% accuracy level for all the models are also obtained from the histograms of Q_p/Q_m and presented in Table 2. As per the statistical criteria (R, E) (R1), arithmetic calculation of Q_p/Q_m (μ, σ) (R2), cumulative probability of Q_p/Q_m (P_{50}, P_{90}) (R3) and prediction of pile load capacity within 20% accuracy level(R4), a ranking system is developed by using RI for different models according to Abu-Farsakh and Titi (2004) and presented in Table 2. The overall performance of the various models under present study are evaluated using RI .The RI is the sum of the ranks of different models according to the above four criteria respectively (RI=R1+R2+R3+R4). Lower the value of RI indicates the better performance of the particular method. Based on the RI values the ELM (RI= 7) model is found to be at par with that of the GP model (RI=7) followed by, FEM (RI=16) and ANN (RI=16), SVM (RI=19) and RVM (RI=19) model.

4.1 Development of model equation

According to the Eq. (1) the ELM model equation can be written using the weights and biases of the trained model as provided in Table 3. The developed model equation can be used to predict the uplift capacity of caisson by the Geotechnical engineers with the help of a spreadsheet without going into the complexities of model development using ELM. Similarly the SVM and RVM models equations can be written based on the parameters obtained from Figs. 5 and 6 as per Eqs. (8) and (10), respectively.

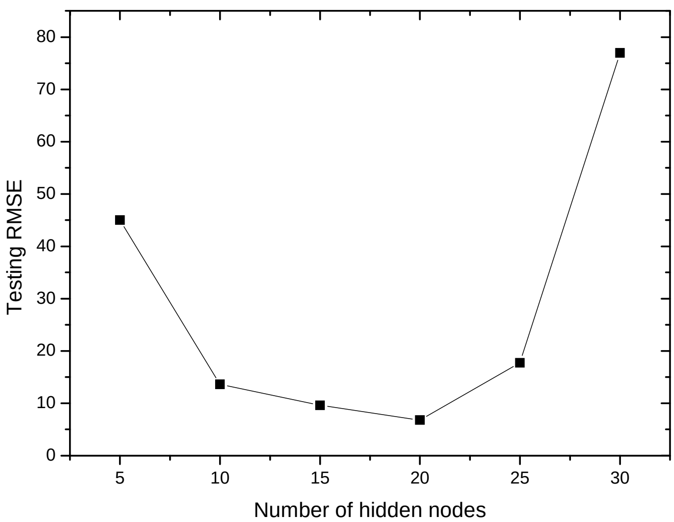

Fig. 1 The generalisation performance of ELM on a wide range of number of hidden nodes

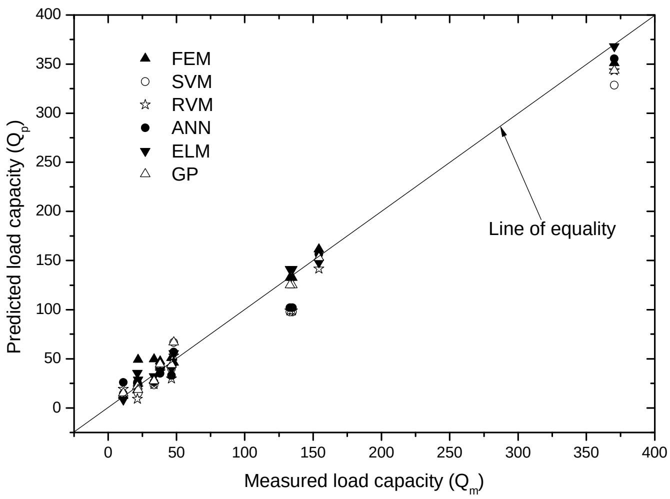

Fig. 2 Comparisons of predicted and measured uplift capacity of suction caisson by different models for testing data

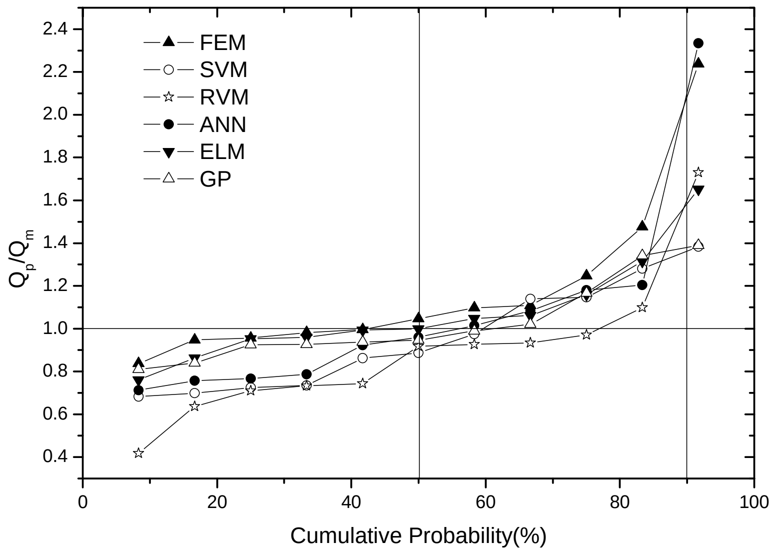

Fig. 3 Cumulative probability plots of Q_p/Q_m for different models for testing data

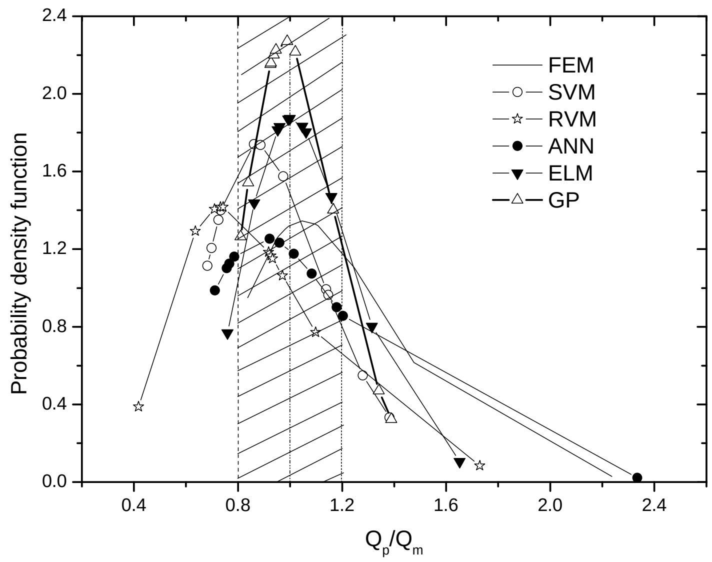

Fig. 4 Log normal distribution of Q_p/Q_m for different models for testing data

Table 1 Comparison of statistical performances of different models

Models	Statistical Performances					
	R	*E*	*AAE*	*MAE*	*RMSE*	*NMBE*
ELM	0.998	0.995	5.674	14.017	6.777	2.377
GP (Muduli *et al.* 2013)	0.997	0.988	8.065	27.055	11.155	-2.979
FEM (Rahman *et al.* 2001)	0.995	0.986	8.490	27.100	11.876	3.690
ANN (Rahman *et al.* 2001)	0.986	0.969	11.009	40.700	17.767	-4.430
ANN (BRNN)	0.991	0.975	12.204	32.820	16.031	-6.970
SVM	0.989	0.955	15.640	42.020	21.310	-11.060
RVM	0.992	0.964	14.960	35.980	19.040	-13.698

Table 2 Evaluation of performance of different prediction models considered in this study

Models	Best fit Calculations			Arithmetic calculations of *Qp/Qm*			Cumulative probability			± 20% Accuracy (%)			Overall rank	
	R	*E*	R1	Mean	σ	R2	Q_p/Q_m at P_{50}	Q_p/Q_m at P_{90}	R3	Log-normal	Histo-gram	R4	RI	Final rank
ELM	0.998	0.995	1	1.069	0.241	3	1.001	1.220	1	64	72	2	7	1
GP (Muduli *et al.* 2013)	0.997	0.988	2	1.026	0.191	1	0.950	1.380	3	72	82	1	7	1
ANN	0.991	0.975	4	1.065	0.454	4	0.960	2.100	2	48	46	6	16	2
FEM (Rahman *et al.*2001)	0.995	0.986	3	1.175	0.391	6	1.050	2.100	4	45	72	3	16	2
SVM	0.989	0.955	6	0.956	0.248	2	0.890	1.380	6	53	46	5	19	3
RVM	0.992	0.964	5	0.893	0.335	5	0.920	1.60	5	38	72	4	19	3

Table 3 Weights and biases of the developed ELM model

Hidden Neuron	Weights						Biases (Hidden nodes)
	Input parameters					Output	
No.	L/d	S_u(kPa)	log T_k		D/L	Q_p (kPa)	b_i
1	-0.568	-0.694	-0.476	0.967	-0.974	18.974	0.626
2	-0.843	-0.189	0.139	0.793	0.211	6.686	0.025
3	0.866	-0.375	-0.281	0.731	0.153	-9.221	0.062
4	0.206	0.388	-0.946	0.602	0.615	-48.882	0.130
5	-0.245	0.781	0.001	0.110	0.310	-33.822	0.451
6	0.330	-0.019	0.654	-0.162	0.756	62.278	0.672
7	0.584	0.612	-0.482	-0.746	0.805	-26.427	0.856
8	-0.333	-0.347	-0.908	0.309	-0.696	12.688	0.498
9	0.385	0.100	-0.507	0.728	-0.615	-14.383	0.049
10	-0.592	-0.222	0.321	-0.451	0.582	-31.634	0.314
11	0.917	0.794	-0.341	0.680	-0.879	7.679	0.642
12	0.424	0.352	0.319	-0.858	-0.220	37.711	0.786
13	-0.666	0.657	-0.974	-0.242	-0.400	33.908	0.289
14	-0.114	-0.780	0.436	-0.464	0.468	-103.001	0.498
15	0.266	-0.442	-0.218	-0.694	-0.792	-67.935	0.818
16	0.860	0.535	-0.933	0.262	0.585	12.145	0.595
17	0.059	-0.568	-0.188	-0.367	0.565	133.007	0.536
18	0.253	-0.932	0.433	0.918	0.065	-5.229	0.331
19	0.362	-0.127	0.843	-0.003	-0.493	34.817	0.412
20	0.846	0.874	0.968	0.477	-0.858	-24.673	0.794

Table 4 Sensitivity analysis of inputs for different models

Parameters	SVM		RVM		ANN				ELM				GP (Muduli *et al.* 2013)	
					Garson's algorithm		Connection weight approach		Garson's algorithm		Connection weight approach			
	Sensitivity analysis (%)	Ranking	Sensitivity analysis (%)	Ranking	Relative importance (%)	Ranking	S_j values	Ranking	Relative importance (%)	Ranking	S_j values	Ranking	Sensitivity analysis (%)	Ranking
L/d	5.020	2	6.250	2	17.380	4	-0.900	3	18.610	5	-9.320	3	-3.527	4
S_u	5.940	1	6.900	1	22.580	2	2.460	1	19.370	4	-10.650	2	109.195	1
T_k	4.780	3	4.980	3	13.550	5	-0.190	5	19.750	3	-7.080	4	-6.426	2
θ	0.920	5	2.450	4	18.310	3	0.810	4	20.630	2	-1.710	5	0.001	5
D/L	1.340	4	0.940	5	28.180	1	1.720	2	21.640	1	11.310	1	4.383	3

4.2 Sensitivity analysis

The sensitivity analysis is an important aspect of a developed model to find out important input parameters. In the present study sensitivity analysis was made as per Liong *et al.* (2000) for SVM and RVM models whereas Garson's algorithm and Connection weight approach were used for ANN and ELM models (Das and Basudhar 2006). As per Liong *et al.* (2000) the sensitivity (S_i) of each parameter in which one input parameter is varied at a time keeping the others constant and is expressed as

$$S_i = \frac{1}{N}\sum_{i}^{N}\left(\frac{\%\text{Change in output}}{\%\text{Change in input}}\right)\times 100 \tag{11}$$

where N= number of training data. Table 4 presents the above analysis for the present study and as per the "best" model (ELM), D/L is the most important parameter similar to the observation made in ANN analysis (Garson algorithm). The other important inputs are S_u, L/d, T_k and θ, is the least important input parameter.

5. Conclusions

The following conclusions can be drawn from the above studies:

(1) The proposed ELM model is found to be effective and efficient than GP, FEM, ANN SVM and RVM models in predicting the uplift capacity of suction caisson in clay as per the statistical performance of the different models.

(2) Using a ranking method based on different statistical criteria (the best fit calculations for predicted uplift capacity (Q_p) and measured capacity (Q_m), the mean and standard deviation of the ratio, Q_p/Q_m, the cumulative probability of Q_p/Q_m and probability of uplift capacity within 20% accuracy level based on histogram and lognormal distribution of Q_p/Q_m), the developed ELM model is found to be more efficient compared to SVM, RVM, FEM and ANN models, whereas it is at par with that of GP model (Muduli *et al.* 2013).

(3) A model equation is presented based on the ELM analysis and it can be helpful for the professionals to find out the uplift capacity of suction caisson in clay using a spreadsheet.

(4) Based on sensitivity analysis, D/*L* (*D* is the depth of the load application point from the soil surface and L is the embedded length of the caisson) has the most significant effect on predicted value of uplift capacity.

References

Abu-Farsakh, M.Y. and Titi, H.H. (2004), "Assessment of direct cone penetration test methods for predicting the ultimate capacity of friction driven piles", *J. Geotech. Geoenviron.*, **130**(90), 935-944.

Bartlett, P.L. (1998), "The sample complexity of pattern classification with neural networks; the size of the weights is more important than the size of network", *IEEE Trans. Inf .Theory*, **44**(2), 525-536.

Berger, J.O. (1985), *Statistical Decision Theory and Bayesian Analysis*, 2nd Ed., Springer, New York.

Boser, B.E., Guyon, I.M. and Vapnik, V. (1992), "A training algorithm for optimal margin classifiers", *Proceedings of the 5th Annual ACM Workshop on COLT*, (Ed., D. Haussler), Pittsburgh, PA, ACM Press.

Briaud, J.L. and Tucker, L.M. (1988), "Measured and predicted axial response of 98 piles", *J. Geotech. Eng*., **114**(9), 984-1001.

Cao, J., Phillips, R., Audibert, J.M.E. and Al-Khafazi, Z. (2002a), "Numerical analysis of the behavior of suction caissons in clay", *Proceedings of the 12th International Offshore and Polar Engineering Conference,* Kitakyushu, Japan.

Cao, J., Phillips, R., Popescu, R., Al-Khafaji, Z. and Audibert, J.M.E. (2002b), "Penetration resistance of suction caissons in clay", *Proceedings of the 12th International Offshore and Polar Engineering Conference,* Kitakyushu, Japan.

Cauble, D.F. (1996), *An experimental investigation of the behavior of a model suction caisson in a cohesive soil*, Ph.D. Dissertation, Massachusetts Institute of Technology, USA.

Cho, Y., Lee, T.H., Park, J.B., Kwag, D.J., Chung, E.S. and Bang, S. (2002), "Field tests on suction pile installation in sand", *Proceedings of the 21st International Conference on offshore Mechanics and Artic Engineering*, Oslo, Norway.

Clukey, E.C., Morrison, M.J., Gariner, J. and Corte, J.F. (1995), "The response of suction caisson in normally consolidated clays to cyclic TLP loading conditions", *Proceedings of the 27th annual Offshore Technology Conference,* Houston, Texas.

Clukey, E.C. and Morrison, M.J. (1993), *A centrifuge and analytical study to evaluate suction caissons for TLP applications in the Gulf of Mexico. In Design and performance of deep foundations: Piles and piers in soil and soft rock*, (Eds., P.P. Nelson, T.D. Smith and E.C. Clukey), ASCE Geotechnical Special Publication.

Das, S.K. and Basudhar, P.K. (2006), "Undrained lateral load capacity of piles in clay using artificial neural network", *Comput. Geotech*., **33**, 454-459.

Das, S.K. and Basudhar, P.K. (2008), "Prediction of residual friction angle of clays using artificial neural network", *Eng. Geology*, **100**(3-4), 142- 145.

Das, S.K. and Sivakugan, N. (2010), "Discussion of Intelligent computing for modeling axial capacity of pile foundations", *Can. Geotech. J.*, **37** (8), 928-930.

Das, S.K. and Muduli, P.K. (2011), "Evaluation of liquefaction potential of soil using extreme learning machine, *Proceedings of the IACMAG*, 9-11 May, Melbourne, Austrailia.

Das, S.K., Samui, P. and Sabat, A.K. (2011), "Application of artificial intelligence to maximum dry density and unconfined compressive strength of cement stabilized soil", *Geotech. Geological J*., **29**(3), 329-342.

Das, S.K., Samui, P., Sabat, A.K. and Sitharam, T.G. (2010), "Prediction of swelling pressure of soil using artificial intelligence techniques", *Environ. Earth Sci*., **61**(2), 393-403.

Datta, M. and Kumar, P. (1996), "Suction beneath cylindrical anchors in soft clay", *Proceedings of the 6th International Offshore and Polar Engineering Conference,* Los Angeles, California.

Dyvik, R., Anderson, K.H., Hansen, S.B. and Christophersen, H.P. (1993), "Field tests on anchors in clay. I: description", *J. Geotech. Eng*., **119**(10), 1515-1531.

El-Gharbawy, S.L. and Olson, R.E. (2000), "Modeling of suction caisson foundations", *Proceedings of the 10th International Offshore and Polar Engineering Conference,* Seattle, USA

Goh, A.T.C. (1996), "Pile driving records reanalyzed using neural networks", *J. Geotech. Eng. - ASCE,* **122**(6), 492-495

Giustolisi, O., Doglioni, A., Savic, D.A. and Webb, B.W. (2007), "A multi-model approach to analysis of environmental phenomena", *Environ. Modell. Softw.*, **5**, 674-682.

Hogervorst, J.R. (1980), "Field trials with large diameter suction piles", *Proceedings of the 12th annual Offshore Technology Conference,* Houston, Texas.

Huang, G.B., Zhu, Q.Y. and Siew, C.K. (2006), "Extreme learning machine: Theory and applications", *J. Neurocomputing*, **70**(1-3), 489-501.

Huang, G.B., Zhou, H., Ding, X. and Zhang, R. (2012), "Extreme learning machine for regression and multiclass classification", *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics,* **42**(2), 513-529.

Larsen, P. (1989), "Suction anchors as an anchoring system for floating offshore constructions", *Proceedings of the 21st annual Offshore Technology Conference*, Houston.

Liong, S.Y., Lim W.H. and Paudyal, G.N. (2000), "River stage forecasting in Bangladesh: neural network approach", *J. Comput. Civ. Eng.*, **14**(1), 1-8.

MathWork Inc. (2005), *Matlab User's Manual. Version 6.5*, Natick (MA).

Muduli, P.K., Das, M.R., Samui, P. and Das, S.K. (2013), "Uplift capacity of suction caisson in clay using artificial intelligence techniques", *Marine Georesourc. Geotechnol.*, **31**(4), 375-390.

Rahman, M.S., Wang J., Deng, W. and Carter, J.P. (2001), "A neural network model for the uplift capacity of suction caissons", *Comput. Geotech.*, **28**, 269-287.

Rao, C.R. and Mitra. S.K. (1971), *Generalised inverse of matrices and its applications*, Wiley, New York.

Rao, S.N., Ravi R. and Ganapathy, C. (1997a), " Behavior of suction anchors in marine clays under TLP loading", *Proceedings of the 16th International Conference on Offshore Mechanics and Arctic Engineering,* Yokohama, Japan.

Rao, S.N., Ravi, R. and Ganapathy, C. (1997b), "Pullout behavior of model suction anchors in soft marine clays", *Proceedings of the 7th International Offshore and Polar Engineering Conference,* Honolulu.

Samui, P. (2008), "Prediction of friction capacity of driven piles in clay using the support vector machine", *Can. Geotech. J.*, **2**, 288-295.

Samui, P., Sitharam, T.G. and Kurup, P.U. (2008), "OCR Prediction Using Support Vector Machine Based on Piezocone Data", *J. Geotech. Geoenviron.*, **134**(6), 894-898.

Samui, P. (2011), "Prediction of pile bearing capacity using support vector machine", *Int. J. Geotech. Eng.*, **5**(1), 95-102.

Singh, B., Datta M. and Gulhati, S.K. (1996), "Pullout behavior of superpile anchors in soft clay under static loading", *Marine Georesourc. Geotechnol.*, **14**, 217-236.

Smola, A.J. and Scholkopf, B. (2004), "A tutorial on support vector regression", *Stat. Comput*., **14**, 199-222.

Steensen-Bach, J.O. (1992), "Recent model tests with suction piles in clay and sand", *Proceedings of the 24th annual Offshore Technology Conference*, Houston, Texas.

Tipping, M.E. (2001), "Sparse Bayesian learning and the relevance vector machine", *J. Machine Learning Res*., **1**, 211-244

Tjelta, T.I., Guttormsen, T.R. and Hermstad, J. (1986), "Large-scale penetration test at a deepwater site", *Proceedings of the 18th annual Offshore Technology Conference*, Houston, Texas.

Vapnik, V.N. (1998), *Statistical learning theory*, New York, Wiley.

Zdravkovic, L., Potts, D.M. and Jardine, R.J. (2001), "A parametric study of the pull-out capacity of bucket foundations in soft clay", *Geotechnique*, **51**(1), 55-67.