International Journal of Geotechnical Engineering

ISSN: 1938-6362 (Print) 1939-7879 (Online) Journal homepage: http://www.tandfonline.com/loi/yjge20

Prediction of lateral load capacity of piles using extreme learning machine

Pradyut Kumar Muduli, Sarat Kumar Das & Manas Ranjan Das

Prediction of lateral load capacity of piles using extreme learning machine

Pradyut Kumar Muduli[1], Sarat Kumar Das*[1] and Manas Ranjan Das[2]

This study presents the development of a predictive model for the lateral load capacity of pile in clay using an artificial intelligence technique, extreme learning machine (ELM). Other artificial intelligence models like artificial neural networks (ANN) (Bayesian regularization neural network (BRNN), differential evolution neural network (DENN)) are also developed to compare the ELM model with them and available empirical models in terms of different statistical criteria. A ranking system is presented to evaluate the present models for identifying the "best" model. Sensitivity analyses are made to identify important inputs contributing to the developed models.

Keywords: Pile load capacity, Statistical performance criteria, Artificial neural network, Extreme learning machine, Bayesian regularization neural network, Differential evolution neural network

Introduction

Estimation of the load carrying capacity of pile foundation is one of the most sought after research areas in geotechnical engineering. Static equilibrium and other dynamic equations (Poulos and Davis, 1980) are used to predict the axial load capacity of pile. The prediction of lateral load capacity of piles, used in tall and offshore structures is more complex and requires solution of non-linear differential equations. The elastic analysis adopting Winkler soil model (Poulos and Davis, 1980) is not suitable for the non-linear soil behavior. Matlock and Reese (1962) adopted elastic analysis using non-linear *p–y* curves. Portugal and Seco e Pinto (1993) used non-linear *p–y* curves and finite element method for the prediction of the behavior of laterally loaded piles. The above two methods are more accurate and widely used. However, spatial variability of soil is inevitable. Thus, developing a sufficiently accurate site model for FEM analysis requires extensive site characterization effort and desired constitutive modeling of clayey soil is also very difficult, even with considerable laboratory testing. Out of various available methods based on field data, the Hansen (1961) and Broms (1964) methods have become very popular for the above study, particularly for the preliminary estimate of pile load capacity. These methods are based on pile load test case histories and involve statistically derived empirical equations for the estimation of expected lateral load capacity.

Artificial intelligence techniques such as artificial neural networks (ANN) 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.*, 2011a). Artificial neural network method has been found to be efficient in predicting the pile load capacity in both cohesionless soil and clayey soil compared to traditional empirical methods (Goh, 1995; Chan *et al.*, 1995; Goh, 1996; Lee and Lee, 1996; Teh *et al.*, 1997; Abu-Kiefa, 1998). The performance of the SVM model was found to be better than that of the ANN model for the prediction of frictional resistance of pile in clay (Samui, 2008). Similar studies have also been made for the prediction of lateral load capacity of piles in clay using ANN (Das and Basudhar, 2006). Using the database as per Rao and Suresh (1996), Das and Basudhar (2006) observed that the ANN model is better compared to Broms' and Hansen's methods, based on various statistical performance criteria. Using the same dataset, Pal and Deswal (2010) developed Gaussian process regression (GPR) and SVM models. They observed that the GPR model is better compared to the SVM model. However, they have compared the GPR model with the ANN model of Das and Basudhar (2006) only in terms of the correlation coefficient (*R*). The correlation coefficient is a biased estimate (Das and Sivakugan, 2010) and it is difficult to assess the prediction of the model in terms of under prediction or over prediction based on the *R* value only.

The most important problem associated with efficient implementation of ANN is generalization for some complex problems. The developed model needs to be equally efficient for new data during testing or validation, which is called generalization. There are different methods for generalization like early stopping and cross validation (Basheer, 2001; Das and Basudhar, 2006). The magnitude

[1]Department of Civil Engineering, National Institute of Technology Rourkela, Odisha, India
[2]Department of Civil Engineering, ITER, SOA University, Bhubaneswar-751030, India

*Corresponding author, email saratdas@rediffmail.com

Received 28 November 2012; accepted 11 April 2013
DOI 10.1179/1938636213Z.00000000041

of weight is one of the reasons for poor generalization (Bartlett, 1998). The methods like Bayesian regularization neural network (BRNN) (Das and Basudhar, 2008) has been used to consider the magnitude of weights as a part of the error function. Another reason for poor generalization is the optimization of the error function of ANN. The error function associated with weights and sigmoid function is a highly non-linear optimization problem with many local minima (Das and Basudhar, 2008). As the characteristics of the traditional non-linear programing based optimization method are initial point dependent, the use of global optimization algorithms like genetic algorithm and simulated annealing are being widely used in training ANN model (Morshed and Kaluarachchi, 1998; Goh *et al.*, 2005). The training of the feed forward neural network using differential evolution (DE) optimization is known as differential evolution neural network (DENN) (Ilonen *et al.*, 2003; Das *et al.*, 2011b). Das *et al.* (2011b) observed that the performance of DENN is better than BRNN and the traditionally used Levenberg-Marquardt neural network (LMNN) for slope stability analysis.

In the recent past, a modified learning algorithm called extreme learning machine (ELM) was proposed by Huang *et al.* (2006a) for single hidden layer feed forward neural network (SLFN). This learning algorithm for SLFN is very fast and hence named ELM. In ELM the hidden nodes are randomly selected and output weights are computed analytically, hence the problem of local minima due to optimization algorithm is avoided. The ELM and its variants have been used for different large complex applications (Huang *et al.*, 2006a; Huang *et al.*, 2006b; Huang and Chen, 2008; Huang *et al.*, 2012) with success and are also found to be efficient compared to ANN and SVM (Huang *et al.*, 2006a). 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 networks (Huang *et al.*, 2006a). However, its use in geotechnical engineering is limited (Das and Muduli, 2011). The following reference citations have been changed to match the reference citations: "Das and Moduli, 2011" to "Das and Muduli, 2011", "Hunag et al. (2006a)" to "Huang *et al.* (2006a)", "Pal and Desiwal (2010)" to "Pal and Deswal (2010)". Like any other numerical method, it needs critical evaluation while application to a new problem. Thus, the efficacy of the model is to be compared with other artificial intelligence models like ANN and other empirical models in terms of different statistical performance criteria.

In the present study, lateral load capacity of piles in clay under un-drained condition has been developed using ELM and ANN (BRNN, DENN) with emphasis on the generalization of the developed models. Different statistical criteria like correlation coefficient (*R*), Nash–Sutcliff coefficient of efficiency (*E*), root mean square error (RMSE), average absolute error (AAE), and maximum absolute error (MAE) are used to compare the ELM model with existing empirical models (Broms' and Hansen's) and ANN (DENN, BRNN) models. A ranking system (Abu-Farsakh and Titi, 2004) using rank index (RI) has also been followed to compare the different models based on four criteria: (i) the best fit calculations (*R* and *E*) for the predicted load capacity (Q_p) and measured capacity (Q_m), (ii) arithmetic calculations (mean, μ, and standard deviation, σ) of the ratio, Q_p/Q_m, (iii) 50 and 90% cumulative probabilities (P_{50} and P_{90}) of the ratio, $Q_p/Q_{m.}$ and (iv) probability of lateral load capacity within 20% accuracy level in percentage using histogram and lognormal probability distribution of Q_p/Q_m.

Methodology

In the present study artificial intelligence techniques, ANN (DENN, BRNN) and ELM have been used. The ANN has been extensively used in geotechnical engineering and hence, the technique is discussed briefly for completeness, whereas details can be obtained in Das and Basudhar (2008). Although ELM has been used in various complex problems (Huang *et al.*, 2006a), it is not very common to geotechnical engineering professionals, and hence, is discussed in brief.

Artificial neural networks

In the present study, the ANN models are trained with DE and Bayesian regularization method and are defined as DENN and BRNN respectively. The use of DENN and BRNN are limited in geotechnical engineering (Das and Basudhar, 2006; Das and Basudhar, 2008; Goh *et al.*, 2005; Das *et al.*, 2011b). A brief description about Bayesian regularization and DENN is presented here for completeness.

Bayesian regularization neural network

In case of back propagation neural network (BPNN) the error function considered for minimization is the mean square error (MSE). This may lead to over-fitting due to unbounded values of weights. The other method is called regularization, in which the performance function is changed by adding a term that consists of the MSE of weights and biases as given below

$$\mathrm{MSEREG} = \gamma \mathrm{MSE} + (1-\gamma)\,\mathrm{MSW} \tag{1}$$

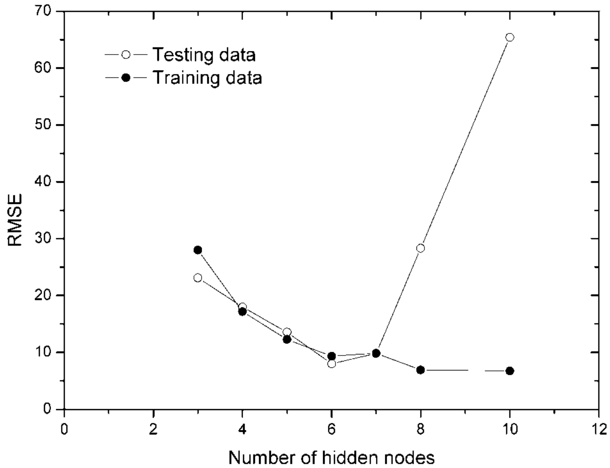
where MSE is the mean square error of the network, γ is the regularization parameter and

$$\mathrm{MSW} = \frac{1}{n}\sum_{\mathrm{j}=1}^{n} w_{\mathrm{j}}^2 \tag{2}$$

This performance function will cause the network to have smaller weights and biases, thereby forcing networks less likely to be over-fit. The optimal regularization parameter γ is determined through Bayesian framework (Demuth and Beale, 2000) as the low value of γ will not adequately fit the training data and its high value may result in over-fitting. The number of network parameters (weights and biases) being effectively used by the network can be found by the above algorithm. The effective number of parameters remains the same irrespective of the total number of parameters in the network.

Differential evolution neural network

The DE optimization is a population based heuristic global optimization method. Unlike other evolutionary



1 Generalization performance of ELM on a wide range of hidden nodes

optimization, in DE, the vectors in the current populations are randomly sampled and combined to create vectors for the next generation with real valued crossover factor and mutation factor. The detail of DENN is available in Ilonen *et al.* (2003).

Extreme learning machine

Huang *et al.* (2006a) 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 a linear system and the output weights (linking the hidden layer to output layer) can be analytically determined through inverse operation of the hidden layer output matrices. Mathematically the above concept can be described for the standard SLFN with L hidden nodes to predict the output variable as follows

$$o_{\mathrm{j}} = \sum_{i=1}^{L} \beta_{\mathrm{i}} f\left(w_{\mathrm{i}} \times x_{\mathrm{j}} + b_{\mathrm{i}}\right) \tag{3}$$

where $f(x)$=activation function, $w_{\mathrm{i}}=[w_{\mathrm{i1}}, w_{\mathrm{i2},\ldots}, w_{\mathrm{in}}]^{\mathrm{T}}$= weight vector connecting the i^{th} hidden node and the input nodes, $\beta_{\mathrm{i}}=[\beta_{\mathrm{i1}}, \beta_{\mathrm{i2}},\ldots, \beta_{\mathrm{im}}]^{\mathrm{T}}$=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], o_{j}=predicted normalized output variable corresponding to the j^{th} input node in the range [0, 1], j=1,…, N, N=number of arbitrary distinct samples (x_{i}, t_{i}), $x_{\mathrm{i}}=[x_{\mathrm{i1}}, x_{\mathrm{i2}},\ldots,x_{\mathrm{in}}]^{\mathrm{T}} \in \mathrm{R}^{\mathrm{n}}$ and $\mathrm{t_i}=[t_{\mathrm{i1}}, t_{\mathrm{i2}},\ldots,t_{\mathrm{im}}]^{\mathrm{T}} \in \mathrm{R}^{\mathrm{m}}$.

The detailed algorithm is presented in Huang *et al.* (2006a). In the present study ELM model is developed using Matlab (MathWork Inc., 2005). The activation function used to develop present prediction model is a simple sigmoid function, which is described in equation (4) as follows

$$f(x) = \frac{1}{(1 + \mathrm{e}^{-x})} \tag{4}$$

Database and preprocessing

In the present study the database as per Rao and Suresh (1996) is used. This has also been used by Das and Basudhar (2006) and Pal and Deswal (2010) for the development of ANN, SVM, and GPR models respectively. The data consist of the diameter of pile (D), depth of pile embedment (L), eccentricity of load (e), and undrained shear strength of soil (S_{u}) as the inputs and measured lateral load capacity(Q_{m}) as the output. The

Table 1 Summary of database used for development of different models

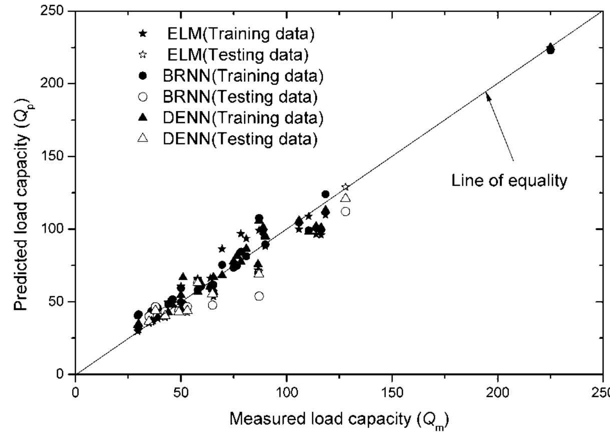
Model variables	Type	Maximum value	Minimum value	Mean	Standard deviation
D (mm)	Input	33·3	6·35	17·784	6·119
L (mm)	Input	300	130	278·892	52·823
e (mm)	Input	50	0	44·166	14·731
S_u (kN m^{-2})	Input	38·8	3·4	9·945	10·115
Q_m (N)	Output	225	29·5	72·816	36·889

Table 2 Comparison of statistical performances of different models

		Statistical performances				
Models		*R*	*E*	Average absolute error (AAE)	Maximum absolute error (MAE)	Root mean square error (RMSE)
Extreme learning machine (ELM)	Training	0·969	0·940	6·934	20·550	9·354
	Testing	0·965	0·920	6·491	16·070	7·983
Differential evolution neural network (DENN)	Training	0·980	0·959	5·647	18·705	7·667
	Testing	0·967	0·905	7·170	18·110	8·549
Bayesian regularization neural network (BRNN)	Training	0·975	0·949	6·609	20·680	8·582
	Testing	0·899	0·734	10·800	33·169	14·312
Hansen	Training	0·950	0·209	30·712	65·360	33·825
	Testing	0·919	0·119	23·650	49·480	26·066
Broms	Training	0·967	0·807	12·391	48·660	16·703
	Testing	0·985	0·574	12·082	46·380	18·127

Table 3 Evaluation of performance of different prediction models considered in this study

Pile capacity methods		Best fit calculations			Arithmetic calculations of Q_p/Q_m			Cumulative probability			±20% Accuracy (%)			Overall rank	
		R	E	$R1$	μ	σ	$R2$	Q_p/Q_m at P_{50}	Q_p/Q_m at P_{90}	$R3$	Log-normal	Histogram	$R4$	rank index (RI)	Final rank
Extreme learning machine (ELM)	Training	0·969	0·940	1	1·012	0·120	1	1·007	1·212	2	92	90	1	5	1
	Testing	0·965	0·920		0·954	0·128		0·924	1·165		84	100			
Differential evolution neural network (DENN)	Training	0·980	0·959	2	1·018	0·106	2	1·012	1·156	1	90	92	2	7	2
	Testing	0·967	0·905		0·948	0·125		0·945	1·161		84	88			
Bayesian regularization neural network (BRNN)	Training	0·975	0·949	3	1·042	0·143	3	1·005	1·238	3	86	86	3	12	3
	Testing	0·899	0·734		0·942	0·196		0·896	1·226		62	66			
Hansen	Training	0·950	0·209	5	0·580	0·111	5	0·542	0·741	5	0	0	5	20	5
	Testing	0·919	0·119		0·590	0·149		0·523	0·838		8	22			
Broms	Training	0·967	0·807	4	1·143	0·144	4	1·112	1·382	4	64	72	4	16	4
	Testing	0·985	0·574		1·166	0·136		1·140	1·392		50	66			



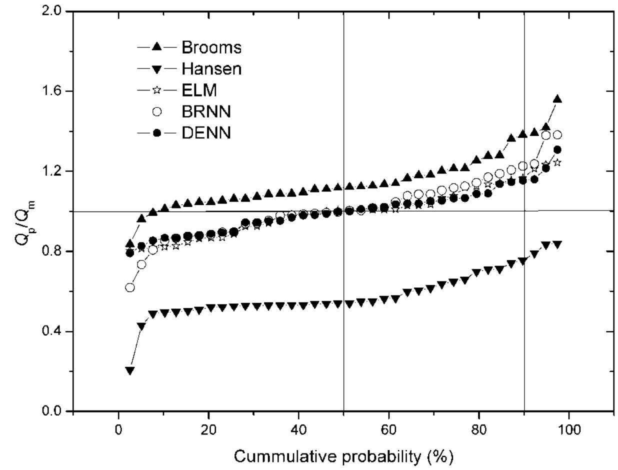
2 Comparisons of predicted and measured load capacity of piles by different models

summarized extract of the database is provided in the Table 1 in terms of maximum value, minimum value, mean, and standard deviation of all the input and output parameters. Out of the mentioned 38 data, 29 data are selected for training and the remaining nine data are used for testing the developed model as per Das and Basudhar (2006). The data are normalized in the range [0, 1] and [−1, 1] for the ELM model and the ANN (DENN, BRNN) models respectively to avoid the dimensional effect of input parameters.

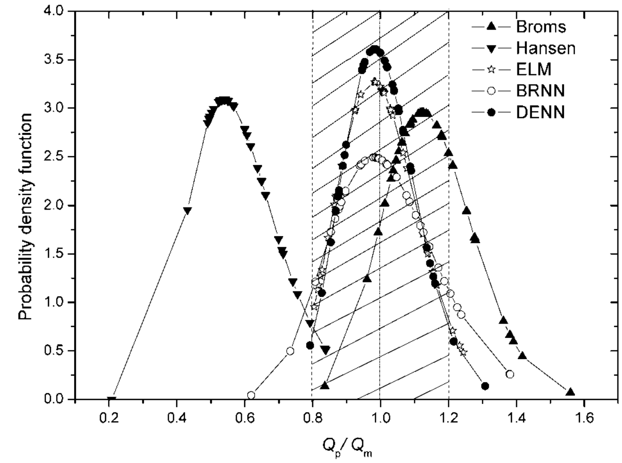
Results and discussion

The best ELM model has been developed with a six hidden nodes SLFN (architecture of 4-6-1) after several trials with different number of hidden nodes. Figure 1 shows the plot of RMSE value versus hidden nodes indicating that the generalization performance of the ELM is stable on a wide range of hidden nodes though the generalization performance gets worse when very few or very large number of hidden nodes are randomly generated. Figure 2 shows the performance of predicted and observed values of lateral load capacity of piles for ELM and other developed models. There is less scatter of data for the ELM model compared to the other models. Table 2 shows the statistical performance in terms of R, E, AAE, MAE, and RMSE for the ELM model along with the results of ANN (DENN, BRNN) and empirical (Brom's and Hansen's) models for both training and testing dataset. The developed ELM model shows good generalization in terms of close values of R and E for training and testing data. It also indicates the robustness of the ELM model as it outperforms all other models in terms of most of the statistical parameters under consideration.

While describing the 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. Abu-Farsakh and Titi (2004) and Das and Basudhar (2006) have used the mean (μ) and standard deviation (σ) of the ratio of the predicted pile load capacity (Q_p) to measured pile load capacity (Q_m) as important parameters



3 Cumulative probability plots of Q_p/Q_m for different models



4 Log normal distribution of Q_p/Q_m for different models

in evaluating different models. The best model is represented by the μ value close to 1·0 and σ close to 0. In present study the μ(1·012, 0·954) and σ(0·120, 0·128) of Q_p/Q_m for the ELM model are very close to those of the DENN (μ(1·018, 0·948), σ (0·106,0·125)) model, and better than BRNN (μ(1·042,0·942), σ(0·143,0·196) and other models as presented in Table 3. The other criteria like cumulative probability of Q_p/Q_m have also been considered for the evaluation of the performance of different models following Briaud and Tucker (1988) 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 implies 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 of Q_p/Q_m for the total observations. The model with P_{90} for Q_p/Q_m close to 1·0 is a better model.

Figure 3 shows the cumulative probability plots of Q_p/Q_m for different methods for both training and testing data and the P_{50} and P_{90} values are presented in Table 2. The P_{50} values of ELM (1·007, 0·924), DENN (1·012, 0·945) and BRNN (1·005, 0·896) models for training and testing data are comparable whereas the Hansen's method (P_{50}=0·542, 0·523) under predicts the pile load capacity and Brom's method (P_{50}=1·112, 1·140) over-predicts the same. However, based on the P_{90} value, the ELM (1·212, 1·165) model is also found to very close to DENN (1·156, 1·161) models and better than other models. The lognormal distributions of Q_p/Q_m for different models are shown in Fig. 4. The ELM and DENN models are equally efficient as the shaded area under the lognormal distribution plots of Q_p/Q_m for both the models are almost equal and are better than other models in terms of probability of load capacity within 20% accuracy level.

As per the best fit calculations criterion (R, E), arithmetic calculations (μ, σ) of Q_p/Q_m, cumulative probabilities (P_{50} and P_{90}) of Q_p/Q_m, and prediction of pile load capacity within 20% accuracy level a ranking system is developed by using RI for different models according to Abu-Farsakh and Titi (2004) and presented in Table 3. Rank index is defined as the sum of ranks obtained by a particular model based on the four different evaluation criteria discussed above and can be mathematically presented as

$$RI = R1 + R2 + R3 + R4 \quad (5)$$

where $R1$=rank obtained from the best fit calculations criterion (R and E), $R2$=rank obtained considering the arithmetic mean (μ) and standard deviation(σ) of Q_p/Q_m, $R3$=rank obtained on the basis of cumulative probabilities (P_{50} and P_{90}) of Q_p/Q_m, $R4$=rank obtained from the consideration of the probability of the pile load capacity within 20% accuracy level based on the histogram and lognormal distribution of Q_p/Q_m. Lower value of RI indicates better performance of the particular method. Based on the RI values, the ELM (RI=5) model is found to be the "best" model followed by DENN (RI=7), BRNN (RI=12), Brooms (RI=16), and Hansen (RI=20) model.

Table 4 Weights and biases for developed ELM model

Hidden nodes	Input weights D	L	e	S_u	Input bias (b_i)	Output weights Q_p
1	0·102	−0·278	0·669	−0·024	0·297	29·617
2	−0·357	0·942	0·731	0·585	0·852	−24·884
3	0·800	0·111	−0·917	0·049	0·974	9·513
4	0·084	0·540	0·228	0·195	0·624	26·493
5	−0·074	0·360	−0·453	−0·560	0·892	−3·765
6	0·616	−0·590	0·493	−0·272	0·981	−28·891

The results of the present ELM model are also compared with the results of the SVM and GPR models presented by Pal and Deswal (2010). However, the SVM and the GPR results are available for the testing data only. The R value of the SVM and GPR methods are 0·92 and 0·98 respectively whereas the R value of the ELM model is 0·965. Similarly, the RMSE value for testing data of the ELM model is found to be 7·983 compared to 11·47 and 6·32 of the SVM and GPR models, respectively. Hence the present ELM model is also found to be better than the SVM model. The R and RMSE value of the ELM model is comparable with the GPR model. However, due to the absence of other statistical parameters, the performance of the GPR model based on the other criteria discussed above could not be found to make an elaborate comparison using RI with ELM.

Development of model equation

According to equation (3) the ELM model equation is presented in equation (11) using the weights and biases of the trained model as provided in Table 4. The developed model equation can be used by geotechnical engineers to predict the lateral load capacity with the help of a spreadsheet without going into the complexities of model development using ELM.

$$Q_p = \frac{29.617}{(1+\exp(-A_1))} - \frac{24.884}{(1+\exp(-A_2))} + \frac{9.513}{(1+\exp(-A_3))} + \frac{26.493}{(1+\exp(-A_4))} - \frac{3.765}{(1+\exp(-A_5))} - \frac{28.891}{(1+\exp(-A_6))} \tag{6}$$

$$A_1 = 0{\cdot}297 + 0{\cdot}102D - 0{\cdot}278L + 0{\cdot}669e - 0{\cdot}024s_u \tag{6a}$$

$$A_2 = 0{\cdot}852 - 0{\cdot}357D + 0{\cdot}942L + 0{\cdot}731e + 0{\cdot}585s_u \tag{6b}$$

$$A_3 = 0{\cdot}974 + 0{\cdot}8D + 0{\cdot}111L - 0{\cdot}917e + 0{\cdot}049s_u \tag{6c}$$

$$A_4 = 0{\cdot}624 + 0{\cdot}084D + 0{\cdot}54L + 0{\cdot}228e + 0{\cdot}195s_u \tag{6d}$$

$$A_5 = 0{\cdot}892 - 0{\cdot}074D + 0{\cdot}36L - 0{\cdot}453e - 0{\cdot}56s_u \tag{6e}$$

$$A_6 = 0{\cdot}981 + 0{\cdot}616D - 0{\cdot}59L + 0{\cdot}493e - 0{\cdot}272s_u \tag{6f}$$

Model equations for BRNN and DENN can be written using the obtained weights and biases following Das and Basudhar (2006, 2008). It may be mentioned here that Pal and Deswal (2010) have not presented anymodel equation to be used by professional engineers.

Sensitivity analysis

The sensitivity analysis is an important aspect of a developed model to find out important input parameters. In the present study sensitivity analyses were made as per Garson's algorithm and connection weight approach for ELM and ANN (DENN, BRNN) models as per Das and Basudhar (2006). Table 5 presents the above analyses and

Table 5 Sensitivity analysis of inputs as per different approaches

	Extreme learning machine (ELM)				Differential evolution neural network (DENN)				Bayesian regularization neural network (BRNN)			
	Garson's algorithm		Connection weight approach		Garson's algorithm		Connection weight approach		Garson's algorithm		Connection weight approach	
Parameters	Relative importance (%)	Ranking	S_j values	Ranking	Relative importance (%)	Ranking	S_j values	Ranking	Relative importance (%)	Ranking	S_j values	Ranking
D	19·05	3	4·24	2	18·48	4	−0·41	4	23·08	3	0·49	4
L	28·79	2	−0·65	3	33·17	1	4·61	3	20·31	2	2·06	2
e	35·91	1	−13·59	1	22·62	3	−10·98	1	14·36	4	−1·49	3
S_u	16·25	4	0·33	4	25·73	2	6·37	2	42·25	1	2·51	1

as per the "best" model, ELM, eccentricity of load (e) is the most important input parameter. Similar observation is also made by the DENN model (connection weight approach). Das and Basudhar (2006, 2008) have observed that connection weight approach is a better indicator of sensitivity analysis compared to Garson's algorithm. The other important inputs are D, L, and S_u.

Conclusions

Based on the results and discussion the following conclusions can be drawn:

1. Using a ranking method based on different statistical criteria (the best fit calculations (R and E) for predicted load capacity (Q_p) and measured capacity (Q_m); the mean and standard deviation of the ratio, Q_p/Q_m; the cumulative probabilities (P_{50} and P_{90}) for Q_p/Q_m. and prediction of load capacity within 20% accuracy level), it has been found that the developed ELM model is more efficient compared to ANN models and empirical models of Hansen and Broms.
2. A model equation is presented based on the ELM analysis and it can be helpful for the professionals to find the lateral load capacity of pile in clay using a spreadsheet.
3. Based on sensitivity analysis eccentricity is found to be most important parameter followed by diameter of pile, length of pile, and un-drained shear strength of clay.

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. Eng.*, **130**, (90), 935–944.

Abu-Kiefa, M. A. 1998. General regression neural networks for driven piles in cohesionless soils, *J. Geotech. Geoenviron. Eng.*, **124**, (12), 1177–1185.

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.

Basheer, I. A. 2001. Empirical modeling of the compaction curve of cohesive soil, *Can. Geotech. J.*, **38**, (1), 29–45.

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

Broms, B. B. 1964. Lateral resistance of piles in cohesive soils, *J. Soil Mech. Found. Eng.*, **90**, (2), 27–63.

Chan, W. T., Chow, Y. K. and Liu, L. F. 1995. Neural network: an alternative to pile driving formulas, *J. Comput. Geotech.*, **17**, 135–156.

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 artifical neural network, *Eng. Geology*, **100**, (3–4), 142–145.

Das, S. K. and Muduli, P. K. 2011. Evaluation of liquefaction potential of soil using extreme learning machine, Proc. 13th Int. Conf.of 'IACMAG 2011', Melbourne, Austrailia , 1, 548–552.

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

Das, S. K., Biswal, R. K., Sivakugan, N. and Das, B. (2011b). Classification of slopes and prediction of factor of safety using differential evolution neural networks, *Environ. Earth Sci.*, **64**, 201–210.

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.

Demuth, H. and Beale, M. 2000. *Neural network toolbox*, The MathWorks Inc, USA.

Goh, A. T. C. 1995. Empirical design in geotechnics using neural networks, *Geotechnique*, **45**, (4), 709–714.

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

Goh, A. T. C., Kulhawy, F. H. and Chua, C. G. 2005. Bayesian neural network analysis of undrained side resistance of drilled shafts, *J. Geotech. Geoenviron. Eng.*, **131**, (1), 84–93.

Hansen, B. 1961. The ultimate resistance of rigid piles against transversal force, *Copenhagen Danish Geotechnical Institute Bulletin*, **12**, 5–9.

Huang, G. B. and Chen, L. 2008. Enhanced random search based incremental extreme learning machine, *Neurocomputing*, **71**, 3460–3468.

Huang, G. B., Zhou, H., Ding, X. and Zhang, R. 2012. Extreme learning machine for regression and multiclass classification, *IEEE Trans. Syst. Man. Cybern. B Cybern.*, **42**, (2), 513–529.

Huang, G. B., Zhu, Q. Y. and Siew, C. K. (2006a). Extreme learning machine: theory and applications, *Neurocomputing*, 70, 489–501.

Huang, G. B., Chen, L. and Siew, C. K. (2006b). Universal approximation using incremental constructive feed forward networks with random hidden nodes, *IEEE Trans. Neural Netw.*, **17**, (4), 879–892.

Ilonen, J., Kamarainen, J. K. and Lampinen, J. 2003. Differential evolution training algorithm for feed-forward neural network, *Neural Process. Lett*., 17, 93–105.

Lee, I. M. and Lee, J. H. 1996. Prediction of pile bearing capacity using artificial neural networks, *Comput. Geotech.*, **18**, (3), 189–200.

MathWork Inc. 2005. *Matlab User's Manual. Version 6·5*, MathWork Inc., Natick, MA.

Matlock, H. and Reese, L. C. 1962. Generalized solutions for laterally loaded piles, *Trans. Am. Soc. Civ. Eng.*, **127**, 1220–48.

Morshed, J. and Kaluarachchi, J. J. 1998. Parameter estimation using artificial neural network and genetic algorithm for free-product migration and recovery, *Water Res. Res.*, **34**, (5), 1101–1113.

Pal, M. and Deswal, S. 2010. Modelling pile capacity using Gaussian process regression, *Comput. Geotech.*, **37**, 942–947.

Portugal, J. C. and Seco e Pinto, P. S. 1993. Analysis and design of pile under lateral loads, 11th Int. Geotechnical Seminar on 'Deep foundation on bored and auger piles', Belgium, 309–313.

Poulos, H. G. and Davis, E. H. 1980. Pile foundation analysis and design; Wiley, New York.

Rao, K. M. and Suresh, K. V. 1996. Measured and predicted response of laterally loaded piles, Proc. 6th Int. Conf. and Exhibition on 'Piling and deep foundations', India, 161–167.

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

Teh, C. I., Wong, K. S., Goh, A. T. C. and Jaritngam, S. 1997. Prediction of pile capacity using neural networks, *J. Comput. Civil Eng.*, **11**, (2), 129–138.