Indian Journal of Geo Marine Sciences
Vol.46 (11) , November 2017, pp. 2350-2356

Minimax probability machine regression and extreme learning machine applied to compression index of marine clay

Pijush Samui[1] & Dookie Kim[2]

[1] Department of Civil Engineering, National Institute of Technology Patna, Patna 800005, India
[2] Department of Civil Engineering, Kunsan National University, Kunsan, Jeonbuk, South Korea.

[Email: pijush.phd@gmail.com; kim2kie@gmail.com]

Received 06 August 2015 ; revised 02 November 2016

This article uses Minimax Probability Machine Regression (MPMR) and Extreme Learning Machine (ELM) for determination of Compression Index (C_c) of marine clay. MPMR is developed in a probabilistic framework. It maximizes the minimum probability of future predictions being within some bound of the true regression function. ELM is the advanced learning algorithm of single-hidden layer feed forward neural network. Natural moisture content (w_n), liquid limit (LL), void ratio (e) and plasticity index (PI) have been used as inputs of MPMR and ELM. The output of MPMR and ELM is C_c. The results of MPMR and ELM have been compared with the regression models. This study gives a powerful tool based on the developed MPMR for determination of C_c of marine clay.

[Keywords: Minimax Probability Machine Regression; Regression; Compression Index; Marine Clay; Extreme Learning Machine]

Introduction

Compression Index (C_c) is a key parameter for determination of settlement of marine clay. There are lots of correlation are available for determination of C_c of marine clay in the literatures[1-6].Every available correlation has some disadvantages. [7]successfully used regression models for determination of C_c based on natural moisture content (w_n), liquid limit(LL), dry density(γ_d), void ratio(e) and plasticity index(PI). The regression model uses least-square method for prediction. Least-square method is sensitive to the presence of outliers, and it performs poorly when the underlying distribution of the additive noise has a long tail.

This study employs Minimax Probability Machine Regression (MPMR) and Extreme Learning Machine (ELM) for prediction of C_c of marine clay based on e, w_n,LL, and PI. MPMR is developed based on Minimax Probability Machine Classification (MPMC)[8].It maximizes the minimum probability that future predicted outputs of the regression model will be within some bound of the true regression function. There are lots of applications of MPMR in the literatures[9-11] . ELM is developed by[12]. It is a single hidden layer forward network (SLFN). It has been successfully applied to solve different problems in engineering[13-15]. MPMR and ELM have been developed based on the database collected from the work of [7]. The dataset contains information about C_c, w_n, e, LL, and PI. The developed MPMR and ELM have been compared with the regression models. This article is organized as follows. Section 2 describes the methodology of MPMR. The details of ELM have been described in section 3. Section 4 gives the results and discussion. Major conclusions have been drawn in section 5.

This section will serve the details of MPMR for prediction of C_c. In MPMR, the relation between input(x) and output(y) is given by the following relation.

$$y = \sum_{i=1}^{N} \beta_i K(x_i, x) + b \quad (1)$$

where N is the number of datasets, $K(x_i,x)$ is kernel function and b and β_i are outputs from MPMR. For prediction of C_c of marine clay using single marine clay parameter, $x = [w_n \; or \; LL \; or \; PI \; or \; e_0]$ and $y = [C_c]$. For prediction of C_c of marine clay using multiple marine clay parameters, $x = [PI, LL, e_0]$ and $y = [C_c]$.

MPMR is developed by constructing a dichotomy classifier[16]. One data set is obtained by shifting all of the regression data +ε along the output variable axis. The other data is obtained by shifting all of the regression data - ε along the output variable axis. The classification boundary between these two classes is defined as a regression surface.

To develop the MPMR, the total dataset have been divided into the following two groups:

Training Dataset: This is used to construct the MPMR. This article uses 131 datasets out of 186 as a training dataset.

Testing Dataset: This is used to verify the developed MPMR. The remaining 55 datasets have been used as testing dataset.

The datasets are normalized between 0 and 1. 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) has been adopted as a kernel function. The program of MPMR has been developed by using MATLAB.

Materials and Methods

ELM is developed by modifying single hidden-layer feed forward neural network (SLFN). In SLFN, the relation between input(x) and output(y) is given below:

$$\sum_{i=1}^{K} \beta_i g_i (w_i . x_j + b_j) = y_j \quad , \quad j=1,\ldots,N \quad (2)$$

where w_i is the weight vector connecting the i^{th} hidden neuron and the input neurons, β_i is the weight vector connecting the ith hidden neuron and the output neurons, b_i is the threshold of the ith hidden neuron, g_i is activation function, K is number of hidden nodes and N is the number of datasets. The above equation can be written in the following way.

$$H\beta = T \quad (3)$$

where $H = \{h_{ij}\}$ (i=1,…,N, j=1,…,K and $h_{ij} = g(w_j . x_i)$) is the hidden-layer output matrix, β $(\beta = [\beta_1, ..., \beta_K])$ is the matrix of output weights, and T $(T = y_1, y_2, ..., y_N)^T$ is the matrix of targets.The value of β is determined from the following expression.

$$\beta = H^{-1}T \quad (4)$$

Where H^{-1} is the Moore–Penrose generalized inverse[17] of H. The learning speed of ELM is increase by using Moore–Penrose generalized inverse method.

ELM adopts the same inputs, output, training dataset, testing dataset and normalization technique as used by the MPMR model. The program of ELM has been developed by using MATLAB.

Results and discussion

The performance of MPMR depends on the choice of proper value of ε and σ. The design values of ε and σ have been determined by trial and error approach. Table 1 shows the value of ε and σ for the different input variable.

The performance of training and testing dataset has been determined by trial and error approach. Coefficient of Correlation(R) has been adopted to asses the performance of MPMR. For a good model, the value of R should be close to one. The performances of training and testing dataset have been shown in figures 1,2,3,4, and 5.

Table 1-Performance of the developed MPMR models.

Input Variables	Design value of ε	Design value of σ	Training Performance(R)	Testing Performance(R)
LL	0.01	0.01	0.970	0.862
w_n	0.02	0.07	0.944	0.912
e	0.06	0.02	0.996	0.831
PI	0.04	0.08	0.953	0.898
e,LL,PI	0.05	0.03	0.989	0.980

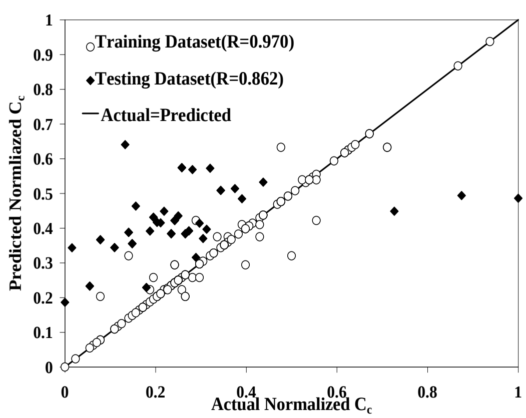

Fig.1- Performance of the MPMR model by using LL.

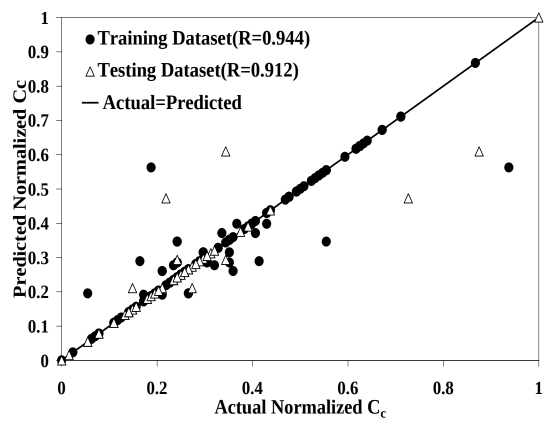

Fig.2- Performance of the MPMR model by

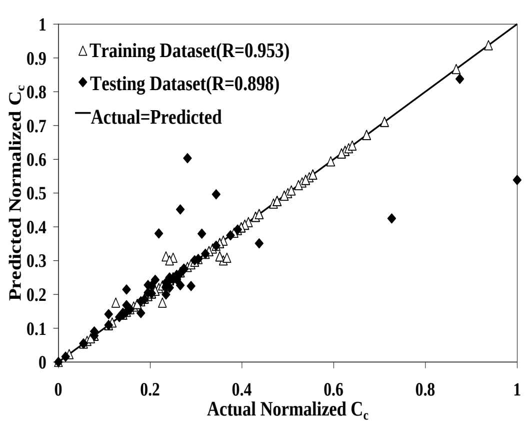

Fig.3- Performance of the MPMR by using e

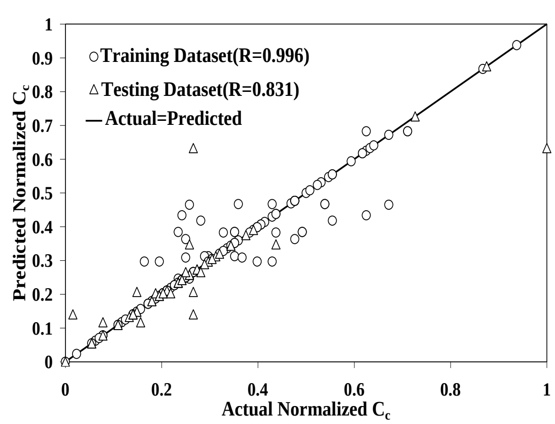

Fig. 4- Performance of the MPMR by using PI

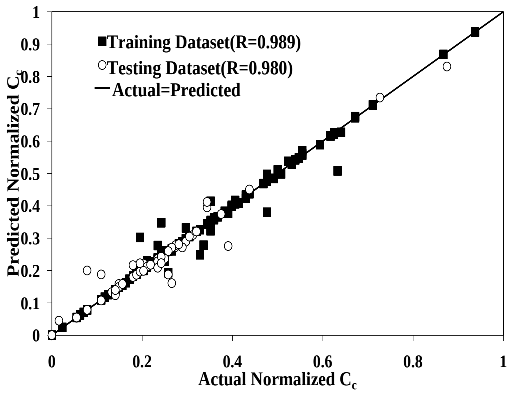

Fig. 5- Performance of the MPMR by using e,LL, and PI.

It is clear from figures that the value R is close to one for training as well as testing dataset. Therefore, the developed MPMR predicts C_c reasonably well.

For developing ELM, radial basis function has been adopted as activation function. The performance of ELM depends on the number of hidden nodes. Table 2 shows the number of hidden nodes for the different models.

The performance of ELM by considering different inputs has been depicted in figures 6,7,8,9 and 10.

Table 2. Number of hidden nodes for the different input variables.

Input Variables	Number of Hidden Nodes
LL	4
w_n	6
E	3
PI	4
e,LL,PI	7

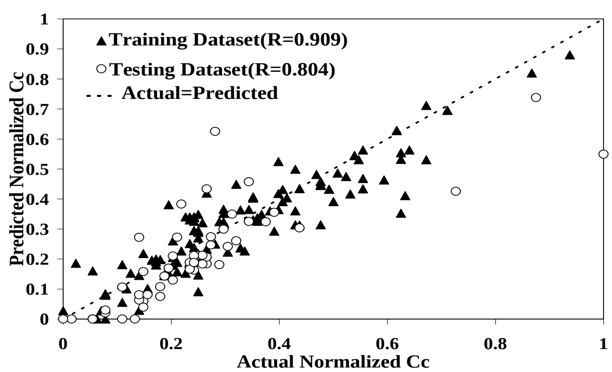

Fig.6- Performance of the ELM model by using

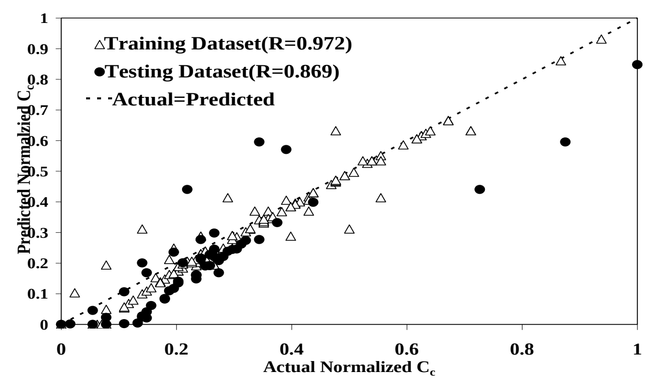

Fig. 8- Performance of the ELM by using e.

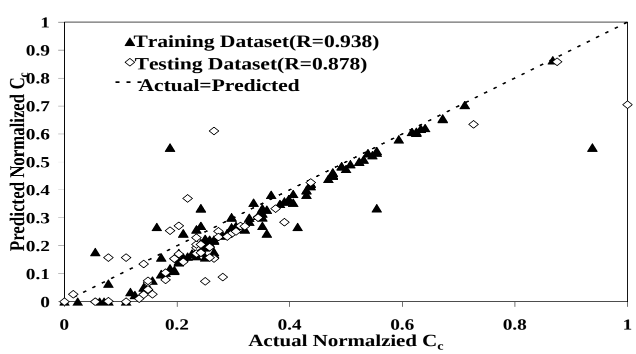

Fig.7-Performance of the ELM model by using w_n.

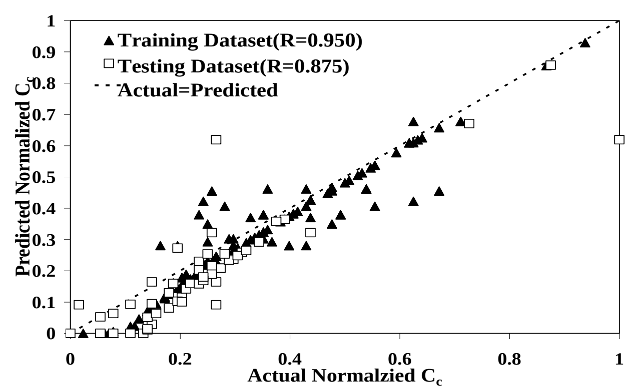

Fig.9- Performance of the ELM by using PI.

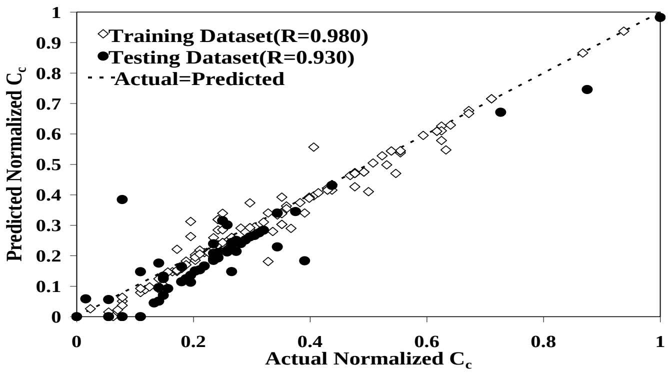

Fig.10- Performance of the ELM by using e,LL, and PI.

It is clear from figures 6,7,8,9 and 10 that the value of R is close to one. So, the developed ELM proves his capability for prediction of C_c. The developed MPMR has been compared with the regression models developed by [7]. Figure 11 shows the bar chart of R values of the different models. It is observed from figure 11 that the performance of MPMR is better than the regression and ELM models.

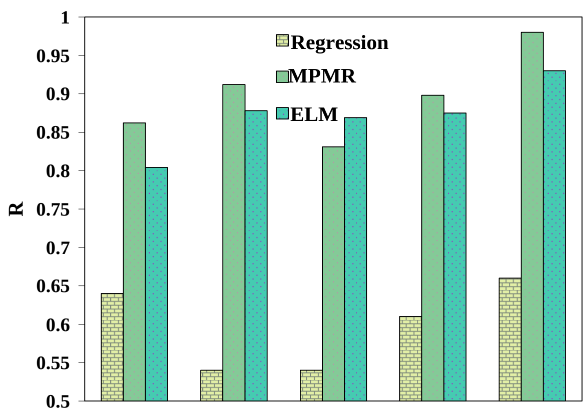

Figure 11. Bar chart of R values of the different models

The performance of ELM and MPMR has been assessed by using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Coefficient of Efficiency (E), Root Mean Square Error to Observation's Standard Deviation Ratio (RSR), Normalized Mean Bias Error (NMBE), Performance Index (ρ) and Variance Account Factor (VAF). The expression of RMSE, MAE, E, RSR, NMBE, ρ and VAF is given below[18-19]

$$MAE = \frac{\sum_{i=1}^{n}\left|C_{cai} - C_{cpi}\right|}{N} \quad (5)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(C_{cai} - C_{cpi}\right)^2}{N}} \quad (6)$$

$$\rho = \frac{RMSE}{\overline{C}_{ca}} \times \frac{1}{R+1} \quad (7)$$

$$E = 1 - \frac{\sum_{i=1}^{N}\left(C_{cai} - C_{cpi}\right)^2}{\sum_{i=1}^{N}\left(C_{cai} - \overline{C}_{ca}\right)} \quad (8)$$

$$RSR = \frac{RMSE}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(C_{cai} - \overline{C}_{ca}\right)}} \quad (9)$$

$$NMBE(\%) = \frac{1/N\sum_{i=1}^{N}\left(C_{cpi} - C_{cai}\right)}{1/N\sum_{i=1}^{N}C_{cai}} \quad (10)$$

$$VAF = \left(1 - \left(\frac{\mathrm{var}\left(C_{ca} - C_{cp}\right)}{\mathrm{var}\left(C_{ca}\right)}\right)\right) \times 100 \quad (11)$$

Where C_{ca} is actual C_c, C_{cp} is predicted C_c, $\overline{C}_{ca}$ is the mean of C_{ca}, var is variance and N is number of dataset. For an accurate model, the value of E and ρ should be close to one and zero respectively. The value of RSR should be low for a good model. For an over prediction model, the value of NMBE will be positive. For perfect association between the actual and predicted values, the value of VAF is 100. Table 3 and 4 shows the values the above parameters of the MPMR and ELM respectively. All MPMR models are over prediction. Only one ELM model is under-prediction. The developed MPMR has control over future prediction. However, the ELM and regression models have no control over future prediction.

Table 3-Different parameters for the developed MPMR.

Parameters	MPMR(LL)		MPMR(w_n)		MPMR(e)		MPMR(PI)		MPMR(e,LL,PI)	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing
RMSE	0.038	0.227	0.058	0.072	0.053	0.077	0.015	0.097	0.025	0.035
MAE	0.014	0.205	0.018	0.023	0.022	0.026	0.004	0.042	0.010	0.018
E	0.987	0.736	0.971	0.924	0.975	0.913	0.998	0.868	0.994	0.981
RSR	0.337	1.163	0.505	1.059	0.464	1.127	0.131	1.353	0.226	0.541
NMBE(%)	4.070	146.675	5.485	9.205	6.697	10.339	1.132	16.616	2.876	7.087
ρ	0.511	0.949	0.522	0.544	0.507	0.572	0.513	2.037	0.504	0.510
VAF	95.302	18.217	89.127	83.008	90.895	80.669	99.309	69.189	97.922	96.138

Table 4-Different parameters for the developed ELM.

Parameters	ELM(LL)		ELM(w_n)		ELM(e)		ELM(PI)		ELM(e,LL,PI)	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing
RMSE	0.074	0.114	0.073	0.100	0.045	0.101	0.063	0.101	0.036	0.074
MAE	0.057	0.080	0.050	0.078	0.030	0.079	0.048	0.075	0.023	0.056
E	0.950	0.762	0.943	0.792	0.980	0.790	0.958	0.779	0.988	0.889
RSR	0.679	2.091	0.779	2.088	0.437	2.080	0.655	2.193	0.337	1.489
NMBE(%)	13.576	15.108	4.414	11.185	4.042	12.211	5.907	8.567	-4.176	7.816
ρ	0.518	1.682	1.771	0.465	0.486	0.471	0.478	0.457	0.494	0.463
VAF	82.674	63.894	87.057	75.819	94.420	74.598	89.517	76.242	96.078	86.292

Conclusion

This article uses MPMR and ELM for prediction of C_c of marine clay based on e,w_n,LL and PI. The datasets have been collected from the different points at east coast of South Korea. The developed MPMR proves his capability for prediction of C_c. It outperforms the regression and ELM models. The developed MPMR can be used as a quick tool for determination of C_c of marine clay. This study shows that developed MPMR is a reliable model for determination of C_c of marine clay.

Acknowledgement

Authors thank to Gil Lim Yoon and Sarat Das for providing the datasets.

References

1. Skempton, A. W., Notes on the compressibility of clays. *Quarterly Journal of the Geological Society of London*, 100 (2) 1944:119–135.
2. Terzaghi, K., and Peck, R. B., Soil mechanics in engineering practice. 2nd ed. New York: John Wiley & Sons Inc, 1967.
3. Lambe, T. W. and Whitman, R. V. *Soil mechanics*. New York: John Wiley & Sons Inc. 1969.
4. Mayne, P. W., Cam-clay predictions of undrained strength. *Journal of the Geotechnical Engineering Division*, 106 (11) 1980:1219–1242.
5. Nakase, A., Kamei, T. and Kusakabe, O. Constitute parameters estimated by plasticity index. *Journal of Geotechnical Engineering*, 114 (GT7) 1988:844–858.
6. Yin, J. H. Properties and behaviour of Hong Kong marine deposits with different clay contents. *Canadian Geotechnical Journal*, 36 (1999):1085–1095.
7. Yoon, G. L., Kim, B. T. and Jeon, S. S. Empirical correlations of compression index for marine clay from regression analysis. *Can Geotech J*, 41 (6) 2004:1213-1221.
8. Lanckriet, G. R. G, *et al.*, Minimax probability machine. *In*: T. G. Editor, S. Editor, & Z. Editor (Eds), Adva*nces in Neural Information Processing Systems 14*, MA. Cambridge: MIT Press. 2002.
9. Sun, J., et al., Modelling of a chaotic time series using a modified minimax probability machine regression. *Chines`e Journal of Physics*, 47 (4) 2009:491-501.
10. Yang, L., *et al*,Simultaneous feature selection and classification via Minimax Probability Machine. *International Journal of Computational Intelligence Systems*, 3 (6) 2010:754-760.
11. Zhou, Z., Wang, Z. and Sun, X., Face recognition based on optimal kernel minimax

probability machine. *Journal of Theoretical and Applied Information Technology*, 48 (3) 2013:1645-1651.

12. Huang, G. B., Zhu, Q. Y. and Siew, C. K., Extreme learning machine. Theory and applications. *Neurocomputing*, 70 (1) 2006.
13. Jiang, C., Deng, M., and Hayakawa, Y., Operator-based parallel compensation control for hysteresis using ELM-based PI model. *International Journal of Advanced Mechatronic Systems*, 4(2012):280-286.
14. Xie, P, et al, Feature extraction and recognition of motor imagery EEG based on EMD-multiscale entropy and extreme learning machine. *Chinese Journal of Biomedical Engineering*, 32 (6) 2013 :641-648.
15. Du, S., *et al*., A novel deformation prediction model for mine slope surface using meteorological factors based on kernel extreme learning machine. *International Journal of Engineering Research in Africa*, 12(2014):67-81.
16. Strohmann, T. R. and Grudic, G. Z.,A Formulation for minimax probability machine regression, *In* T. G. Editor, S. Editor, & Z. Editor (Eds.), *Advances in Neural Information Processing Systems (NIPS) 14*, MA. Cambridge: MIT Press, 2002.
17. Serre, D. *T. Matrices Theory and Applications*. Berlin: Springer. 2002.
18. Gandomi, A. H., and Roke, D. A., Intelligent formulation of structural engineering systems. In Seventh M.I.T. Conference on Computational Fluid and Solid Mechanics Focus: Multiphysics & Multiscale, Massachusetts Institute of Technology, Cambridge, MA, 2013.
19. Chandwani, V, Agrawal, V & Nagar, R 2014. Modeling Slump of Ready Mix Concrete Using Genetic Algorithms. *Assisted Training of Artificial Neural Networks Expert Systems with Applications*, 42(2):885-93.