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

Pijush Samui¹ & Dookie Kim²

¹Department of Civil Engineering, National Institute of Technology Patna, Patna 800005, India
²Department 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 (Cc) of marine clay. MPMR is developed in a probabilistic framework. It maximizes the minimum probability that future predicted outputs of the regression model will be within some bound of the true regression function. ELM is the advanced learning algorithm of single-hidden layer feedforward neural network. Natural moisture content (wn), 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 Cc. 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 Cc of marine clay.

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

Introduction
Compression Index (Cc) is a key parameter for determination of settlement of marine clay. There are lots of correlation are available for determination of Cc of marine clay in the literatures1-6. Every available correlation has some disadvantages. ⁷successfully used regression models for determination of Cc based on natural moisture content (wn), liquid limit (LL), dry density (γd), 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 Cc of marine clay based on e, wn, 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 literatures⁹-⁻¹¹. ELM is developed by¹². It is a single hidden layer forward network (SLFN). It has been successfully applied to solve different problems in engineering¹³⁻¹⁵. MPMR and ELM have been developed based on the database collected from the work of ⁷. The dataset contains information about Cc, wn, 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.

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

where $N$ is the number of datasets, $K(x_i, x)$ is kernel function and $b$ and $\beta_i$ are outputs from MPMR. For prediction of $C_c$ of marine clay using single marine clay parameter, $x = [w_n \text{ or } LL \text{ or } PI \text{ or } e_0]$ 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 $+\varepsilon$ along the output variable axis. The other data is obtained by shifting all of the regression data $-\varepsilon$ 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 $\sigma$ 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:

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

where $w_i$ is the weight vector connecting the $i^{th}$ hidden neuron and the input neurons, $\beta_i$ is the weight vector connecting the $i^{th}$ hidden neuron and the output neurons, $b_i$ is the threshold of the $i^{th}$ 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.

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

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

$$ \beta = H^{-1}T $$ (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 $\varepsilon$ and $\sigma$. The design values of $\varepsilon$ and $\sigma$ have been determined by trial and error approach. Table 1 shows the value of $\varepsilon$ and $\sigma$ 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 assess 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.

<table>
<thead>
<tr>
<th>Input Variables</th>
<th>Design value of e</th>
<th>Design value of σ</th>
<th>Training Performance(R)</th>
<th>Testing Performance(R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LL</td>
<td>0.01</td>
<td>0.01</td>
<td>0.970</td>
<td>0.862</td>
</tr>
<tr>
<td>wn</td>
<td>0.02</td>
<td>0.07</td>
<td>0.944</td>
<td>0.912</td>
</tr>
<tr>
<td>e</td>
<td>0.06</td>
<td>0.02</td>
<td>0.996</td>
<td>0.831</td>
</tr>
<tr>
<td>PI</td>
<td>0.04</td>
<td>0.08</td>
<td>0.953</td>
<td>0.898</td>
</tr>
<tr>
<td>e, LL, PI</td>
<td>0.05</td>
<td>0.03</td>
<td>0.989</td>
<td>0.980</td>
</tr>
</tbody>
</table>
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>
<thead>
<tr>
<th>Input Variables</th>
<th>Number of Hidden Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>LL</td>
<td>4</td>
</tr>
<tr>
<td>$w_n$</td>
<td>6</td>
</tr>
<tr>
<td>E</td>
<td>3</td>
</tr>
<tr>
<td>PI</td>
<td>4</td>
</tr>
<tr>
<td>e,LL,PI</td>
<td>7</td>
</tr>
</tbody>
</table>

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

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

Fig. 6- Performance of the ELM model by using LL.

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

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

Fig. 9- Performance of the ELM by using 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.

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 ($\rho$) and Variance Account Factor (VAF). The expression of RMSE, MAE, E, RSR, NMBE, $\rho$ and VAF is given below\(^1\):

\[
MAE = \frac{\sum_{i=1}^{n} |C_{cai} - C_{cpi}|}{N}
\]

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n} (C_{cai} - C_{cpi})^2}{N}}
\]

\[
\rho = \frac{RMSE}{C_{ca}} \times \frac{1}{R + 1}
\]

\[
E = 1 - \frac{\sum_{i=1}^{N} (C_{cai} - C_{cpi})^2}{\sum_{i=1}^{N} (C_{cai} - C_{ca})^2}
\]

\[
RSR = \frac{RMSE}{\sqrt{\frac{1}{N} \sum_{i=1}^{N} (C_{cai} - C_{ca})}}
\]

\[
NMBE(\%) = \frac{1}{N\sum_{i=1}^{N} (C_{cpi} - C_{cai})}
\]

\[
VAF = \left[1 - \left(\frac{\text{var}(C_{ca} - C_{cp})}{\text{var}(C_{ca})}\right)\right] \times 100
\]

Where $C_{ca}$ is actual $C_c$, $C_{cp}$ is predicted $C_c$, $C_{ca}$ is the mean of $C_{ca}$, $\text{var}$ is variance and $N$ is number of dataset. For an accurate model, the value of E and $\rho$ 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.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>MPMR(LL)</th>
<th>MPMR(w_o)</th>
<th>MPMR(e)</th>
<th>MPMR(PI)</th>
<th>MPMR(e,LL,PI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training</td>
<td>Testing</td>
<td>Training</td>
<td>Testing</td>
<td>Training</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.038</td>
<td>0.227</td>
<td>0.058</td>
<td>0.072</td>
<td>0.053</td>
</tr>
<tr>
<td>MAE</td>
<td>0.014</td>
<td>0.205</td>
<td>0.018</td>
<td>0.023</td>
<td>0.022</td>
</tr>
<tr>
<td>E</td>
<td>0.987</td>
<td>0.736</td>
<td>0.971</td>
<td>0.924</td>
<td>0.975</td>
</tr>
<tr>
<td>RSR</td>
<td>0.337</td>
<td>1.163</td>
<td>0.505</td>
<td>1.059</td>
<td>0.464</td>
</tr>
<tr>
<td>NMBE(%)</td>
<td>4.070</td>
<td>146.675</td>
<td>5.485</td>
<td>9.205</td>
<td>6.697</td>
</tr>
<tr>
<td>ρ</td>
<td>0.511</td>
<td>0.949</td>
<td>0.522</td>
<td>0.544</td>
<td>0.507</td>
</tr>
<tr>
<td>VAF</td>
<td>95.302</td>
<td>18.217</td>
<td>89.127</td>
<td>83.008</td>
<td>90.895</td>
</tr>
</tbody>
</table>

Table 4-Different parameters for the developed ELM.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>ELM(LL)</th>
<th>ELM(w_o)</th>
<th>ELM(e)</th>
<th>ELM(PI)</th>
<th>ELM(e,LL,PI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training</td>
<td>Testing</td>
<td>Training</td>
<td>Testing</td>
<td>Training</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.074</td>
<td>0.114</td>
<td>0.073</td>
<td>0.100</td>
<td>0.045</td>
</tr>
<tr>
<td>MAE</td>
<td>0.057</td>
<td>0.080</td>
<td>0.050</td>
<td>0.078</td>
<td>0.030</td>
</tr>
<tr>
<td>E</td>
<td>0.950</td>
<td>0.762</td>
<td>0.943</td>
<td>0.792</td>
<td>0.980</td>
</tr>
<tr>
<td>RSR</td>
<td>0.679</td>
<td>2.091</td>
<td>0.779</td>
<td>2.088</td>
<td>0.437</td>
</tr>
<tr>
<td>NMBE(%)</td>
<td>13.576</td>
<td>15.108</td>
<td>4.414</td>
<td>11.185</td>
<td>4.042</td>
</tr>
<tr>
<td>ρ</td>
<td>0.518</td>
<td>1.682</td>
<td>1.771</td>
<td>0.465</td>
<td>0.486</td>
</tr>
<tr>
<td>VAF</td>
<td>82.674</td>
<td>63.894</td>
<td>87.057</td>
<td>75.819</td>
<td>94.420</td>
</tr>
</tbody>
</table>

Conclusion
This article uses MPMR and ELM for prediction of C_c of marine clay based on e,w_o,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


