PREDICTING CALIFORNIA BEARING CAPACITY VALUE OF STABILIZED POND ASH WITH LIME AND LIME SLURDE USING BIOGEOGRAPHY-BASED MULTI-LAYER PERCEPTRON NEURAL NETWORK

1 JIAMAN LI  
Shaanxi College of Communication Technology, Xi'an, Shaanxi, 710000, China

2 JUNDONG WU  
China Airport Construction Group Co., Ltd. Northwest Branch, Xi'an, Shaanxi, 710000, China

3 WEI HU  
Shaanxi Provincial Transport Planning Design And Research Institute, Xi'an, Shaanxi, 710000, China

ABSTRACT
In this study, a hybrid biogeography-based multi-layer perceptron neural network (BBO-MLP) with different number of hidden layers (one up to three) was developed for predicting the California bearing capacity (CBR) value of pond ash stabilized with lime and lime sludge. To this aim, model had five variables named maximum dry density, optimum moisture content, lime percentage, lime sludge percentage and curing period as inputs, and CBR as output variable. Regarding BBO-MLP models, BBO-MLP1 has the best results, which its $R^2$ stood at 0.9977, RMSE at 0.7397, MAE at 0.476, and PI at 0.0104. In all three developed models, the estimated CBR values specify acceptable agreement with experimental results, which represents the workability of proposed models for predicting the CBR values with high accuracy. Comparison of three developed models supply that BBO-MLP1 outperform others. Therefore, BBO-MLP1 could be recognized as proposed model.

Keywords: California bearing capacity; Pond Ash Stabilized; Lime; Lime Sludge; Hybrid Biogeography-Based Multi-Layer Perceptron Neural Network

INTRODUCTION
Fly ash, bottom ash and pond ash are produced from thermal power plants [1, 2]. Bottom ash produced from the boiler and FA stem from the electrostatic precipitators are blended water and produce slurry and disposed in ash ponds, so called pond ash [3]. Ash Class C contains high level of calcium content and react with water in any condition even in the presence or absence of other binder [4]. Class F have lower calcium and high percent of thermal power plant produces Class F ash [5]. Class F ash does not have sufficient strength to use as an independent building material. Class F ash improved with many additives such as lime, cement, and gypsum were studied by several researchers [6–9]. Another waste material is lime sludge, which produced from acetylene, sugar, sodium chromate, and water softening plants. In India, annually about 4.5 million tons of this material is generated [10], causes to dispose to be an environmental problem [11–13]. Because lime
sludge has calcium carbonate [14], it has utilized in various fields such as a replacement of cement in mixtures [10], sub-base in village road [15], and in building material [16].

In order to evaluate the soil subgrade either flexible or rigid pavements, California bearing ratio (CBR) is paramount importance parameter [17]. This test is carried out on compacted soil specimens in the laboratory, and is done on a ground surface in the field [18]. Beside this test is time-consuming, the result of this test could be untrustable because of the sample disturbance and poor quality of testing conditions. Hence, development of artificial-based methods could be beneficial to use in the different prediction process [19], especially CBR values [20].

Pavement subgrade soils were stabilized with two various waste marble powder (calcitic marble powder and dolomitic marble powder). CBR tests were applied to measure the improvement in their bearing capacity. According to the comparison based on the CBR values, the inclusion of an appropriate amount and type of waste marble powder positively affected the CBR of the subgrade soil samples regardless of curing time and number of freeze-thaw cycles. Additionally, both curing time and freezing-thawing cycles influenced the CBR values depending on the waste marble powder type and subgrade soil type [21].

Artificial neural network is a simulation of the human brain and widely were accepted in different civil engineering fields to specify the relationships between dependent and independent variables [22]. Therefore, neural network is enough accurate to predict the most geotechnical engineering fields [23]. Artificial neural network has been impressively used in various field of geotechnical engineering, including the compressive strength and Young's modulus of frozen of frozen sand [24, 25], bearing capacity of pile [26, 27], slope stability [28-30], and tunnels and underground openings [31, 32].

Although, huge number of artificial neural network application is done in geotechnical engineering, the number of studies to predict the CBR of soil are a few. Single and multiple regression were used to predict the CBR of soil using data gathered from highways of Turkey’s situated in various regions. The results show that artificial neural network was better than statistical models [33]. ANN and multiple regression models applied to predict the CBR of stabilized soil with lime and quarryby Sabat [34]. They considered quarry dust (%), lime (%), curing period (CP), optimum moisture content (OMC) and maximum dry density (MDD) as input variables. Both developed models predicted the CBR values exactly, while ANN outperform multiple regression.

Two modeling techniques namely Random forest RF and M5P model tree are used to model, soaked CBR value of pond ash. Pond ash was stabilized with the help of LI and LS. CBR data generated from experimental program was used. Performance of models was measured using standard statistical parameters. Although, both the model's performance in predicting CBR value is satisfactory however from the
statistical parameters it is evident that RF method perform better in comparison to M5P model [35]

Recently, the use of geosynthetic-reinforced soil has become famous for constructing sustainable pavement structures. Its strength assessed by California bearing ratio (CBR). The principal goal of the paper was to find and evaluate the competency of the several intelligent methods such as artificial neural network, least squares regression, Gaussian processes regression, elastic net regularization regression, M-5 model trees, alternating model trees and random forest in estimating the CBR of reinforced soil [36].

In this study, hybrid biogeography-based multi-layer perceptron neural network (BBO-MLP) with different number of hidden layers were developed for predicting the CBR value of pond ash stabilized with lime and lime sludge. To this aim, model had five variables named maximum dry density (MDD), optimum moisture content (OMC), lime percentage (L), lime sludge percentage (LS) and curing period (CP) as inputs, and CBR as output variable. In order to evaluate the accuracy of developed model, four performance indices were considered.

2. Dataset and Methodologies

2.1 Description of the dataset
To evaluate the effect of lime and lime sludge on the CBR value in different ages, an experimental-based dataset containing 51 outstanding data records was collected from the published document (Table 1) [37]. Five different variables that can affect the value of the CBR were considered as input variables. These variables included: lime percentage (LI), lime sludge percentage (LS), curing period (CP), optimum moisture content (OMC) and maximum dry density (MDD). The statistics of the input and output variables used for developing the model are given in Table 2.

Table 1. Seventeen mix designs and CBR results

<table>
<thead>
<tr>
<th>Mix No.</th>
<th>LI (%)</th>
<th>LS (%)</th>
<th>MDD (g/cc)</th>
<th>OMC (%)</th>
<th>CBR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1.175</td>
<td>26.8</td>
<td>2.2</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1.194</td>
<td>25.3</td>
<td>10.1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>5</td>
<td>1.227</td>
<td>25.0</td>
<td>14.6</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>10</td>
<td>1.230</td>
<td>24.8</td>
<td>20.6</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>15</td>
<td>1.238</td>
<td>24.3</td>
<td>24.7</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>0</td>
<td>1.207</td>
<td>24.5</td>
<td>16.2</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>5</td>
<td>1.237</td>
<td>23.8</td>
<td>22.5</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>10</td>
<td>1.251</td>
<td>23.3</td>
<td>28.1</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>15</td>
<td>1.253</td>
<td>23.1</td>
<td>32.4</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>0</td>
<td>1.227</td>
<td>24.2</td>
<td>26.1</td>
</tr>
<tr>
<td>11</td>
<td>6</td>
<td>5</td>
<td>1.259</td>
<td>23.0</td>
<td>33.1</td>
</tr>
</tbody>
</table>
Table 2. The statistical values of the input and output variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>All data</th>
<th>LI</th>
<th>LS</th>
<th>MDD</th>
<th>OMC</th>
<th>CP</th>
<th>CBR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min</td>
<td>0.0</td>
<td>0.0</td>
<td>1.175</td>
<td>22.8</td>
<td>7.0</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>Max</td>
<td>8.0</td>
<td>15.0</td>
<td>1.259</td>
<td>26.8</td>
<td>45.0</td>
<td>63.8</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>4.706</td>
<td>7.059</td>
<td>1.233</td>
<td>24.112</td>
<td>26.667</td>
<td>35.722</td>
<td></td>
</tr>
<tr>
<td>St. deviation</td>
<td>2.468</td>
<td>5.703</td>
<td>0.022</td>
<td>1.008</td>
<td>15.542</td>
<td>15.438</td>
<td></td>
</tr>
</tbody>
</table>

2.2 Biogeography-based optimization

The biogeography-based optimization (BBO) method is one of the meta-heuristic optimization algorithms developed by Simon[38]. BBO algorithm is inspired by the geographical distribution, emigration, and immigration of species in an ecosystem. In this method, it is assumed that an ecosystem includes a restricted number of habitats. Many factors called suitability index variables (SIVs) affect the quality of habitat for species, including food, water resources, climate condition, etc. The quality of each habitat is represented by the habitat suitability index (HSI). When a habitat is saturated or has a large HIS, the species are inclined to emigrate from this habitat and immigrate to the low value of HIS. Each habitat depicted one possible answer for the problem, and its suitability indexes are the decision variables (DVs). In the minimization procedure, the solutions with lower objective values have higher values of HSIs. Two operators, “migration” and “mutation,” are used in this algorithm. The migration operator is used to explore the neighborhood of the available solutions; however, the mutation operator is utilized to find the new solutions and to assist the exploration.

For habitats with the size of HS, the habitats are sorted from their cost function values. The suitability of the $i^{th}$ habitat (HSI$_i$) in the sorted population is defined as Eq. (1).
The emigration ($\mu_i$) and immigration ($\lambda_i$) values are computed using Eqs. (2)-(3).

$$H_S I_i = -i + H_S + 1 \quad (1)$$

$$\mu_i = \frac{H_S I_i}{H_S} \quad (2)$$

$$\lambda_i = 1 - \frac{H_S I_i}{H_S} \quad (3)$$

In Fig. 1, the migration curve of the BBO method is presented. In this study, the maximum values of the emigration and immigration rates are assumed to be one. Migration from the $j^{th}$ decision variable of $r^{th}$ habitat to the decision variable of $i^{th}$ habitat is formulated using Eq. (4).

$$DV_i^k = \alpha DV_i^j + (1 - \alpha)DV_j^j \quad (4)$$

Fig. 1. Migration curve of the BBO algorithm

2.3 Biogeography-based multi-layer perceptron neural network (BBO-MLP)
An artificial neural network is based on a mathematical simulation of a human brain procedure. ANN includes three basic sections: 1) transfer function, 2) the pattern of connection, and 3) learning function [39]. Mentioned principles are selected to train a specific network by altering its biases and weights regarding the problem nature and structure [41,42]. Each layer includes neurons with special mathematical relationships, but the output layer's neuron numbers depend on the target dataset.
The hidden layers’ neurons are responsible for discovering the detecting the features lying inside the transferred signals. Following this, the following process is performed in the hidden layer: Receiving of the neurons of input layer and applying their weighted summation, Adding the biases to the weighted summation, locating the results of the second step into brought up transfer function, and transferring the results to the output layer or to the next hidden layer.

Turning to hidden layers, in this study, the maximum number of each hidden layer is constant at 25, although the number of hidden layers is investigated. Likewise, the BBO algorithm is applied to the model with a different number of hidden layers to determine the neurons' optimized number in each hidden layer. In the ANN training procedure, many algorithms have been explored to get the weights and biases. To aim this, the back-propagation algorithm for MLP learning is used according to the previous successful use [39,40]. In this technique, the inputs nutrified signals are changed and weighted between various neurons of layers to obtain the acceptable output. The process of the hybrid BBO-MLP is clarified in Fig. 2.

![Fig. 2. The process of the hybrid BBO-MLP](image)

2.1 Performance evaluators
Different statistical evaluators were used to appraisal the performance of developed hybrid models for predicting the CS. Coefficient of determination ($R^2$), root mean
squared error (RMSE), mean absolute error (MAE) and performance index (PI) were used as precision measurements (Eqs. (5)-(8)):

\[ R^2 = \left( \frac{\sum_{p=1}^{P} (t_p - \bar{t})(y_p - \bar{y})}{\left( \sum_{p=1}^{P} (t_p - \bar{t})^2 \right)^{1/2} \left( \sum_{p=1}^{P} (y_p - \bar{y})^2 \right)^{1/2}} \right)^2 \]  

(5)

\[ RMSE = \sqrt{\frac{1}{P} \sum_{p=1}^{P} (y_p - t_p)^2} \]  

(6)

\[ MAE = \frac{1}{P} \sum_{p=1}^{P} |y_p - t_p| \]  

(7)

\[ PI = \frac{1}{|\bar{t}|} \frac{RMSE}{\sqrt{R^2 + 1}} \]  

(8)

- \( y_p \): the predicted values of the \( p^{th} \) pattern
- \( t_p \): the target values of the \( p^{th} \) pattern
- \( \bar{t} \): the averages of the target values
- \( \bar{y} \): the averages of the predicted values

3. Results and discussion

The results of the developed models for predicting CBR value of pond ash stabilized with lime and lime sludge are presented as follows. Comparing the measured records from experimental efforts with those predicted by BBO-MLP1, BBO-MLP2, and BBO-MLP3 models are supplied in Figs. 3 (a-c). It can be observed that the developed models have \( R^2 \) larger than 0.9962. It means that the correlation between measured and predicted values from integrated models are mostly the same so that it shows the highest accuracy. Besides, in order to compare the productivity of the applied models, a score-based mechanism is designed by considering various statistical parameters such as \( R^2 \), RMSE, MAE, and PI. The results are shown in Table 3.

Regarding BBO-MLP models, BBO-MLP1 has the highest score (12), with \( R^2 \) stood at 0.9977, RMSE at 0.7397, MAE at 0.476, and PI at 0.0104. All indices get worse by increasing the number of hidden layers from 1 (BBO-MLP1) to 3 (BBO-MLP3). For instance, RMSE increase from 0.7397 to 0.9904 by raising the number of hidden layers. Therefore, BBO-MLP1 outperform BBO-MLP2 (\( R^2 = 0.997 \), RMSE=0.8755,
and PI=0.0123) and BBO-MLP3 (R²=0.9962, RMSE=0.9904, and PI=0.0139). All in all, far apart from the BB optimization algorithm has the acceptable performance to find out the number of neurons for each hidden layer.

An acceptable fit between measured and predicted CBR is obtainable from the time series plots presented in Figs. 4(a-c). As can be seen, in all applied models, the estimated CBR values in specify remarkable close agreement with experimental results.

Table 3. The result of created GWMLP models for predicting CBR

<table>
<thead>
<tr>
<th>Models</th>
<th>BBO-MLP1</th>
<th>BBO-MLP2</th>
<th>BBO-MLP3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of hidden layer (s)</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Number of hidden layers neurons</td>
<td>[6]</td>
<td>[4, 4]</td>
<td>[6, 11, 5]</td>
</tr>
<tr>
<td>Network results</td>
<td>All data (Ranking score)</td>
<td>RMSE</td>
<td>R²</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.7397 (3)</td>
<td>0.9977 (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.8755 (2)</td>
<td>0.997 (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9904 (1)</td>
<td>0.9962 (1)</td>
</tr>
<tr>
<td>TRS</td>
<td>12</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Rank</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
Fig. 3. The scatter plot of between observed and predicted CBR; (a) BBO-MLP1, (b) BBO-MLP2, (c) BBO-MLP3
2. Sensitivity analysis

An evaluation of the sensitivity of the hybrid models was conducted to assess the most determinative input parameters to compute the CBR. Various data were built by removing a single input parameter simultaneously, and reported the amounts of statistical performance criteria as $R^2$, and RMSE. The best model for the sensitivity analysis is chosen using the statistical performance criteria. In the present study, the BBO-MLP1 model is selected due to the remarkable performance. The results are as Table 4, which is shown that the CP is the most influential parameter for predicting the CBR using the mentioned model. From this perspective, other inputs have slight

Fig. 4. CBR prediction using models; (a) BBO-MLP1, (b) BBO-MLP2, (c) BBO-MLP3
impression on the performance of the model. It is worth considering that eliminating input variables may only cause a minimal performance loss for the model, but in the present study, because the analysis was based upon the results of experimental measurements so as to specify the impressions of stabilized materials, eliminating variable could decline the generalizability of the model. Taking into account the multicollinearity problem has not a significant impact on the fit of a model and also commonly does not impress remarkably on predictions, the present study does not prefer deleting any variable.

Table 4. Sensitivity analysis using the tree-based Random Forest model

<table>
<thead>
<tr>
<th>Inputs</th>
<th>Removed parameter</th>
<th>R²</th>
<th>RMSE</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDD, OMC, LI, LS, and CP</td>
<td>MDD</td>
<td>0.9977</td>
<td>0.7397</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>OMC</td>
<td>0.9956</td>
<td>1.03</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>LI</td>
<td>0.9936</td>
<td>1.2512</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>LS</td>
<td>0.9936</td>
<td>1.2784</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>CP</td>
<td>0.6648</td>
<td>7.9461</td>
<td>1</td>
</tr>
</tbody>
</table>

5. Conclusion

In this study, a hybrid biogeography-based multi-layer perceptron neural network (BBO-MLP) with different number of hidden layers (one up to three) was developed for predicting the California bearing capacity (CBR) value of pond ash stabilized with lime and lime sludge. To this aim, model had five variables named maximum dry density, optimum moisture content, lime percentage, lime sludge percentage and curing period as inputs, and CBR as output variable.

Regarding BBO-MLP models, BBO-MLP1 has the best results, which its $R^2$ stood at 0.9977, RMSE at 0.7397, MAE at 0.476, and PI at 0.0104. All indices worsen by decreasing the number of hidden layers. Therefore, BBO-MLP1 outperform BBO-MLP2 ($R^2=0.997$, RMSE=0.8755, and PI=0.0123) and BBO-MLP3 ($R^2=0.9962$, RMSE=0.9904, and PI=0.0139).

In all three developed models, the estimated CBR values specify acceptable agreement with experimental results, which represents the workability of proposed models for predicting the CBR values with high accuracy. Comparison of three developed models supply that BBO-MLP1 outperform others. Therefore, BBO-MLP1 could be recognized as proposed model.
Declarations

Funding: no funding supported

Conflicts of interest/Competing interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Availability of data and material: Data is available on the request from the author.

Code availability: Not applicable

Authors' contributions:
1. Jiaman Li
   Conceptualization, Investigation, Methodology, Software, Validation, Writing - review & editing.
2. Jundong Wu
   Formal analysis, Methodology, Investigation, Supervision.
3. Wei Hu
   Methodology, Data curation, Resources, Supervision.

Additional declarations: Not applicable

Ethics approval: Not applicable

Consent to participate: Not applicable

Consent for publication: Not applicable

References


