**PREDICTION OF SOIL’S COMPACTION CHARACTERISTICS USING ARTIFICIAL NEURAL NETWORKS (ANNs)**

Mohamed Alhag Emad Mohamed, University of Khartoum, Sudan, [mohamedemadalhaj@gmail.com](mailto:mohamedemadalhaj@gmail.com).

Alzobair Mohammed, University of Khartoum, Sudan, [alziber50@gmail.com](mailto:alziber50@gmail.com).

Ahmed Mohammed Elsharief, University of Khartoum, Sudan, [aelsharief@yahoo.com](mailto:aelsharief@yahoo.com).

**Abstract**

*In all geotechnical projects, the determination of the compaction properties of soil plays an important role. Soil is compacted to improve its engineering properties like increasing its density, reducing its permeability, and increasing its shear strength. The compaction test is a relatively tedious test that consumes a lot of time and resources. Classically modelling this phenomenon will require a multi-variable regression analysis for all the factors that influence it. Artificial neural networks (ANNs) provide a powerful tool to model such complex systems. This paper discusses the development and use of two artificial neural network models as tools to predict the optimum moisture content (OMC) and the maximum dry density (MDD) of soil compacted at a standard proctor test. The first one is developed with Matlab neural network tool (nntool) using Levenberg-Marquardt's algorithm. The second one is developed using* *stochastic gradient descent’s algorithm. Both models use atterberg's limits and particle sizes as inputs. The research shows that the two models are capable of predicting the maximum dry density and the optimum moisture content for various types of soils. This is indicated by the high coefficient of determination (R2) value and the low Mean Absolute Percentage Error (MAPE), Mean Square Error (MSE), and Root Mean Square Error (RMSE) of the two models.*

**Keywords:** artificial neural network, compaction characteristics, atterberg’s limits, sieve analysis.

**INTRODUCTION**

Compaction is the process in which soil particles are pushed closer to each other by reducing the volume of air in the soil. It improves its engineering properties. The compaction test consists of multiple points with calculated dry densities and moisture contents. From those points, a curve describing the relationship between moisture content and dry density is established. From this curve, the maximum dry density and its corresponding moisture content (known as the optimum moisture content) are found. The maximum dry density is the highest dry density achievable at a given compaction effort (energy). This takes several trials and it is both time and resources consuming. In a developing country like Sudan, engineers don't have the laboratory equipment to do this test in every city. These models will serve as a piece of technology that can overcome this shortage. This will significantly reduce the cost of engineering projects especially roads, and earth-fill dam construction which helps in the development of the country.

Developing a classic mathematical model to predict the outcome of a compaction test is not an easy problem. The compaction characteristics depend on many parameters. The degree of influence of those parameters varies with the soil type. For example: in clean coarse-grained soils the soil particle distribution will dominate the result of compaction. However, in fine-grained soils, the plasticity of clays presented by liquid limit (LL) and plastic limit (PL) will control the expected compaction results. This variation makes it difficult to come up with a single mathematical model that yields accurate results for all types of soils (clean coarse-grained soils, fine-grained soils, and in between). Artificial neural networks can overcome these difficulties.

**Artificial Neural Network**

Artificial Neural Network (ANN) is a new machine learning concept that processes the data in a stacked-layer structure starting from an input layer that contains input until reaching the last layer that generates the output. Every layer contains neurons (processing units). It's called a neural network because it mimics the human brain neurons in the way it processes data.

Neural networks are one of the supervised learning methods which mean it needs initial data to learn from.

**Neural Network Components**

Neural networks have necessary components:

***Input and Output***

Every artificial neural network must have input(s) and output(s). In this study, the inputs are soil's index properties and classification. The outputs are the maximum dry density and optimum moisture content.

***Weights***

Every neuron in the layers of the neural network has a value that’s called weight. It represents the importance of its corresponding input. During the training procedure, the model keeps updating and synchronizing these weights till the model’s prediction accuracy becomes as good as it can be. These weights are usually randomly initialized.

***Bias***

In addition to the weights, there's also the bias. This is a single neuron added to every layer in the network to take into account any bias in the data. It's added to the weights when predicting the output of the layer.

***Layers and neurons***

It’s very important to choose the number of layers in a network, and the number of neurons in every layer. It plays an important role in the accuracy of the model and the speed of training. It's chosen usually after a trial-and-error procedure.

***Activation Function***

It’s a function that’s used to convert a list of inputs to an output. This function takes into account the weights of the neurons of the model.

**RELATED STUDIES**

A lot of research has been done to predict compaction characteristics. Their efforts revolved around using simpler inputs to give reasonably accurate outputs. This part of the paper focuses on some of those studies and discusses their methodology and results.

A model was developed by (Tenpe & Kaur, 2015) for predicting compaction parameters based on the liquid limit, plasticity index, soil particles less than 75 microns size, and soil particles greater than 75 microns size. The used dataset was divided into training, validation, and testing set. Other independent samples were used to evaluate the developed model.

A model was reported by (Jayan & N.Sankar, 2015) in the special issue for ICETTAS'15. They studied the use of an artificial neural network to predict compaction parameters. They used liquid limit, plastic limit, plasticity index, percentages of fines, sand, and gravel, and specific gravity as inputs. The dataset was divided into training, validation, and testing set. The validation was based on the mean square error, and mean absolute error.

Another model was developed by (Atalar, Adu-Parkoh, B.M.Das, & A.Ozyanki, 2019) to predict the compaction properties of soil. They compared multiple regression models and an artificial neural network model. They used liquid limit, plasticity index, and percentage of fines to estimate the maximum dry density, and the optimum moisture content.

An attempt was done by (Sivrikaya & Soycan, 2009) to predict compaction parameters of fine-grained soil. They investigated the use of plastic limit to predict optimum moister content and maximum dry density. They also investigated the use of optimum moisture content to estimate the maximum dry density.

In their published work by the Multidisciplinary Digital Publishing Institute (MDPI), (Jeremiah, Abbey, Booth, & Kashyap, 2021) provided a literature review for a vast collection of research related to geo-mechanical properties of stabilized clay. In these researches, artificial neural networks were used to predict several characteristics. Two models were reported in this research predicting optimum moisture content and maximum dry density. The first one used LL, PI, LS, clay-silt ratio, sand content, lime content, cement content, and asphalt content in percentage (Alavi, Gandomi, Mollahassani, Heshmati, & Rashed, 2010). The second one used Gs, Ls, free swell, D10, D30, D60, Cu, Cc, LL, and PL as inputs (Salahudeen, Ijimdiya, Eberemu, & Osinubi, 2018).

**METHODOLOGY**

**Data**

Data used in this research is taken from Building and Road Research Institute (BRRI). It consists of 232 data points from different parts of Sudan. The data covers various types of soil from expansive clays to clean coarse-grained soil. Some statistical parameters for the data are shown in table 1:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Max | Min | Mean | Standard deviation |
| LL | 80 | 18 | 34.6 | 12.25 |
| PL | 47 | 6 | 19.08 | 5.36 |
| PI | 48 | 3 | 15.47 | 7.88 |
| Fines (%) | 95 | 0.5 | 23.56 | 21.94 |
| Sand (%) | 84 | 2 | 18.74 | 12.59 |
| Gravel (%) | 97 | 0.1 | 57.77 | 25.34 |

***Table 1:*** *Statistical parameters of soil characteristics.*

***Data Types***

The data usually gets divided into three important sets:

*Training Set*

This is given to the model to learn from and adapt its weights to according to it. It’s usually the biggest part of the data so the model can learn about all kinds of data.

*Validation Set*

This is used to test the model's ability to predict against new data while training to make sure training is going in the right direction, or stop training and fix the model if it's not learning. It indicates overfitting.

*Testing Set*

This is used to get the real accuracy of the model by trying the model against new data that it hasn’t seen yet.

**How Does the Network Learn?**

When the input goes from the input layer to the first hidden layer, first the input is multiplied by the weights of the layer. The bias is added to the outcome and then it's transferred to the next layer after activating it using the activation function. This process continues from one layer to another till it reaches the output layer. The error is calculated using one of the error metrics. This error is used to update these weights to improve the accuracy of the network. The part of the network responsible for updating the weights is called the optimizer.

**Important Parameters to Choose**

***Learning Rate***

It's a rate used to decide the speed at which the model update its weights. Its value is very important in training. If it's very large the model may not reach its best accuracy, and if it's too small it may be very slow to train.

***The Train/Test/Validation Data Percentage***

It’s important to decide which percentage of the data is set to the test data and the validation data. It affects the data distribution. If test data was so small, it might not give you a good idea of the model’s accuracy. Both models in this paper use 80/10/10 percentages for training, validation, and testing respectively.

**Accuracy Metrics**

To know how our model is doing, we need to have some metrics to calculate the accuracy. We do this by calculating the error (the difference between the model output and the real data). Four common methods were used in this paper:

1. ***Mean Square Error (MSE):*** it’s calculated using:

|  |  |  |
| --- | --- | --- |
|  |  | **(1)** |

1. ***Root Mean Square Error (RMSE):*** it’s calculated using:

|  |  |  |
| --- | --- | --- |
|  |  | **(2)** |

1. ***Mean Absolute Percentage Error (MAPE):*** it’s calculated using:

|  |  |  |
| --- | --- | --- |
|  |  | **(3)** |

1. ***Coefficient of Determination (R2):*** it’s calculated using:

|  |  |  |
| --- | --- | --- |
|  |  | **(4)** |

Where:

Predicted value of a given output.

Observed value of the same output.

Mean of the observed values.

Number of samples.

**Levenberg-Marquardt’s****Model**

***Normalization of Data***

In artificial neural networks, normalization means putting the data to the same scale. This process can be done with a lot of equations most commonly used being the linear ones. In this model, both the inputs and the outputs are normalized. The normalization can increase the learning speed. It can also improve the accuracy of the model. The following equation is used for the normalization of inputs and outputs after trying various approaches:

|  |  |  |
| --- | --- | --- |
|  |  | **(5)** |

Where:

Xnorm Normalized value of the parameter.

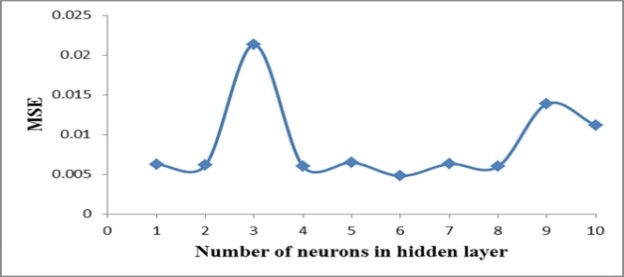
Xi Value of the parameter.

Xmin Minimum value of the parameter.

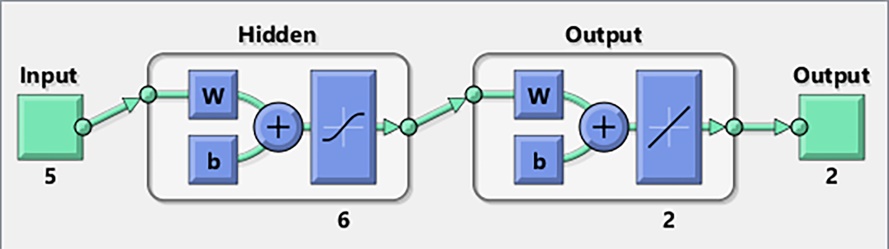
Xmax Maximum value of the parameter.

***Construction of the Artificial Neural Network***

A feed-forward backpropagation network is developed to predict the compaction properties. It consists of an input layer, hidden layer, and output layer. The inputs used in this model are liquid limit (LL), plasticity index (PI), and percentages of fines, sand, and gravel. The outputs are the optimum moisture content (O.M.C.) and the maximum dry density (M.D.D.). The number of neurons in the hidden layer plays an important role in the accuracy of the model. To determine the optimum number of neurons, a trial-and-error procedure was used. In this procedure, the architecture of the model was determined by using a varying number of neurons in the hidden layer. The numbers ranged from 1 to 10, and the (MSE) was used as a performance metric. The results of this procedure are shown in figure 1, and the final network’s architecture is shown in figure 2:



***Figure 1:*** *Optimization of the hidden layer.*

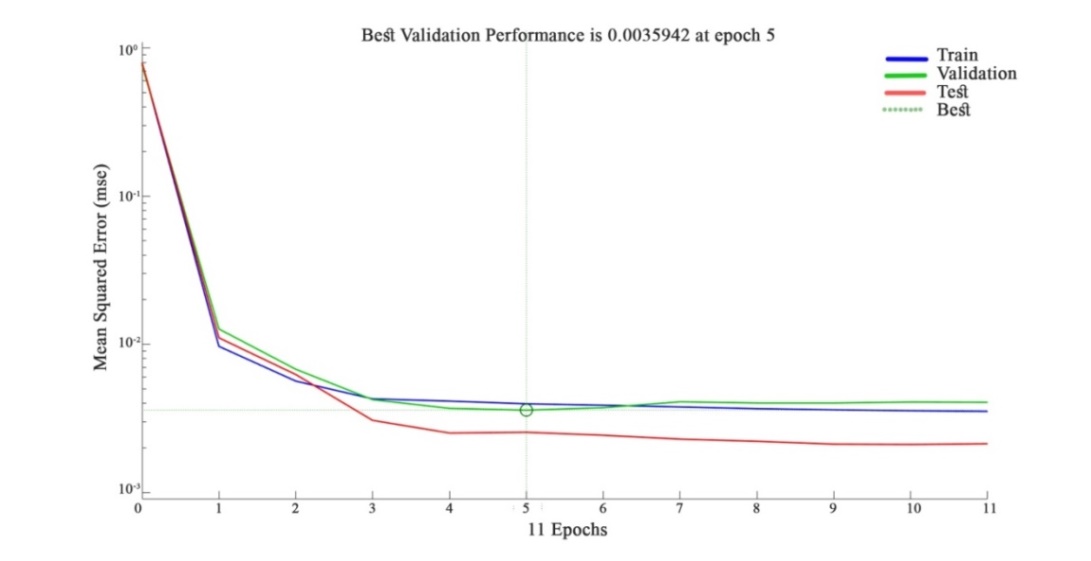


***Figure 2:****Network’s architecture from MATLAB.*

The hidden layer uses (TANSIG) transfer function. The output layer uses (PURELIN) transfer function. The network uses Levenberg-Marquardt's algorithm as a training function and gradient descent with momentum as a learning function.

***Training of The network***

The training is a crucial part of developing a functional model. Excessive training may lead to overfitting of the date. Overfitting is when the model gets too familiar with the dataset rather than developing a general repeatable relationship between the inputs and the outputs. An overfit model will have a nearly perfect result in training and very poor accuracy in testing rendering it useless in predicting future points. To overcome this problem, the weights and biases of the network corresponding to the most generalized network are used. To find this point, the accuracy of the validation is reported after every complete cycle of learning. The point with the minimum (MSE) is considered the optimum point. This procedure is shown in figure 3:



***Figure 3:*** *Training performance of the model.*

**Stochastic Gradient Descent’s****Model**

***Normalization of Data***

In this model, the inputs are normalized in a linear scale between 0 and 1 using the following equation:

|  |  |  |
| --- | --- | --- |
|  |  | **(6)** |

Where:

Xnorm Normalized value of the parameter.

Xi Value of the parameter.

Xmin Minimum value of the parameter.

Xmax Maximum value of the parameter.

***Construction of the Model***

This model uses the normalized liquid limit, plastic limit and plasticity index, classification, and a new parameter (Ψ) as inputs. The classification is partially based on the Unified Soil Classification System (USCC). In this method, the classification starts with determining the mass percentage of fines, sand, and gravel within the soil. If the fines percentage is more than 50% the soil is classified as fine and it is designated by its type based on the A-line (silt (M) or clay (C)) followed by a letter describing its plasticity. The plasticity's description is explained as follows (Knappett & Craig, 2012):

L: low plasticity (liquid limit < 35).

I: Intermediate plasticity (35 < liquid limit < 50).

H: High plasticity (50 < liquid limit < 70).

V: Very high plasticity (70 < liquid limit < 90).

E: Extreme plasticity (liquid limit >90).

When the percentage of fines is less than 50%, the soil is classified according to the coarse part of it, and the percentage of fines. This procedure is explained in table 2:

|  |  |
| --- | --- |
| Fines’ percentage | Classification method |
| Less than 5% | The soil is classified by its coarse-grained part (gravel (G) and sand (S)) by using two letters. The first one represents the dominant coarse-grain part. For example (GS) is sandy gravel. |
| Between 5% and 15% | The classification is a mix of 1 and 3. For example (GS-CH) is sandy gravel with some high plasticity clay. |
| More than 15% | The coarse part is ignored, and the classification is based on the fine part. |

***Table 2:*** *Classification method for soils with a significant quantity of fines.*

The parameter (Ψ) takes into account the water required to reach (OMC). For example, soil with 80% fines will have significantly higher (OMC) than soil with 5% fines. This parameter has a high correlation with (OMC) and (MDD) and it is given by the following equation:

|  |  |  |
| --- | --- | --- |
|  |  | **(7)** |

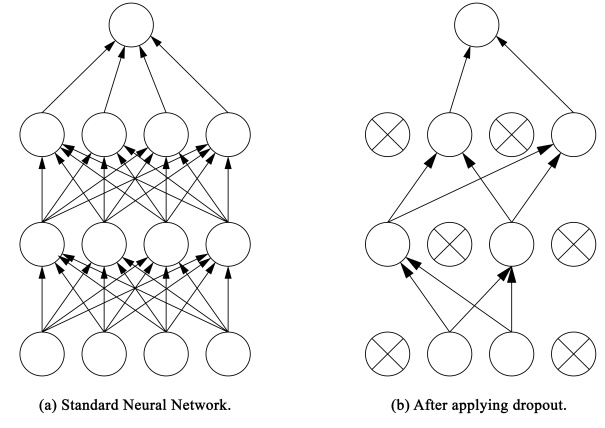
Where:

Fine % Percentage of fines.

Sand % Percentage of sand.

Gravel % Percentage of gravel.

In this model, we used stochastic gradient descent (SGD) as an optimizer and mean squared error as a loss function. We tried a large network to get better accuracy. The optimum network architecture was found by trial and error. The network has 6 hidden layers with the following number of neurons respectively: 160, 80, 40, 20, 10, and 5. The activation functions used are Rectified linear activation unit (Relu) and Sigmoid. The learning rate was selected to be 0.01. Two dropout layers were used in this model since it was overfitting. One was added after the 80 neurons ANN layer, and the second was added after the 20 neurons ANN layer. The dropout layer concept is shown in figure 4:



***Figure 4:*** *Dropout layer in a typical neural network.*

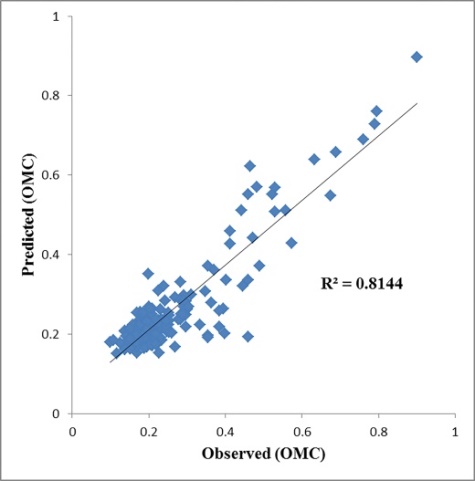
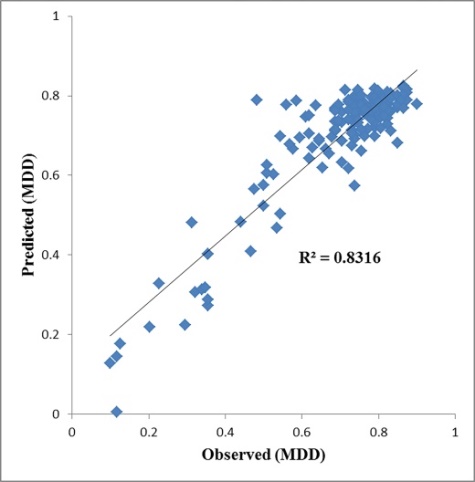
**RESULTS AND DISCUSSION**

**Levenberg-Marquardt’s Model**

This part of the paper discusses the results of the model developed using Levenberg-Marquardt’s algorithm. The results during training, validation, and testing are shown in table 3 and figures 5 to 7.

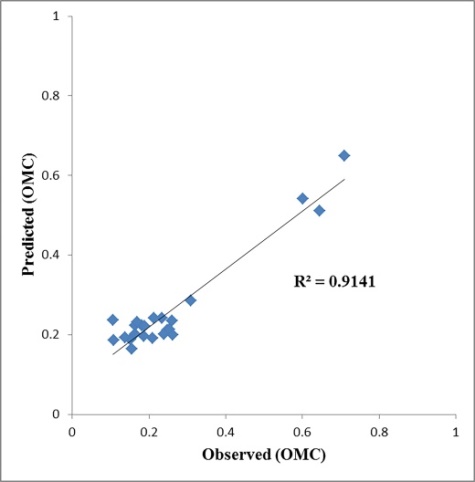
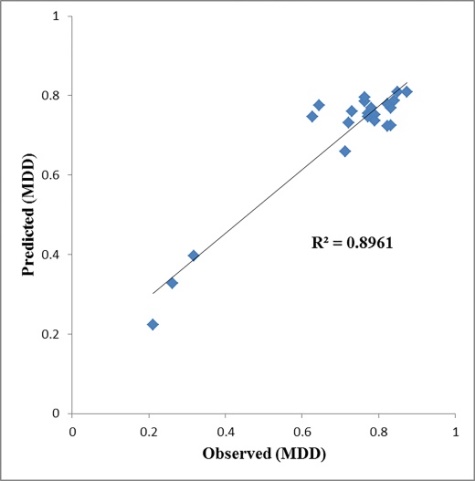
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Performance  metric | training | | validation | | testing | |
| OMC | MDD | OMC | MDD | OMC | MDD |
| MAPE (%) | 18.77 | 2.96 | 25.56 | 2.98 | 19.25 | 2.11 |
| MSE | 0.0035 | 0.0045 | 0.0033 | 0.0039 | 0.0029 | 0.0021 |
| RMSE | 0.0589 | 0.0668 | 0.0572 | 0.0626 | 0.0544 | 0.0463 |
| R2 | 0.8144 | 0.8316 | 0.9141 | 0.8961 | 0.9472 | 0.9511 |

***Table 3:*** *The results of Levenberg-Marquardt’s model.*

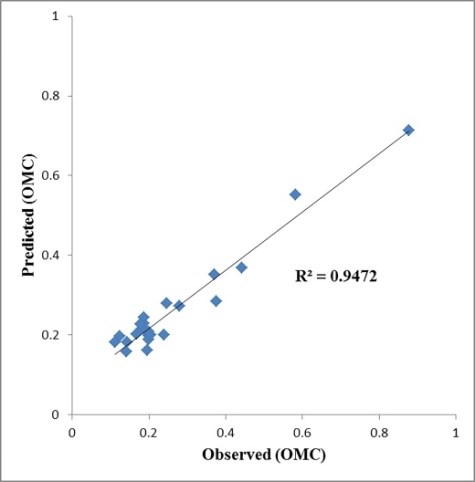
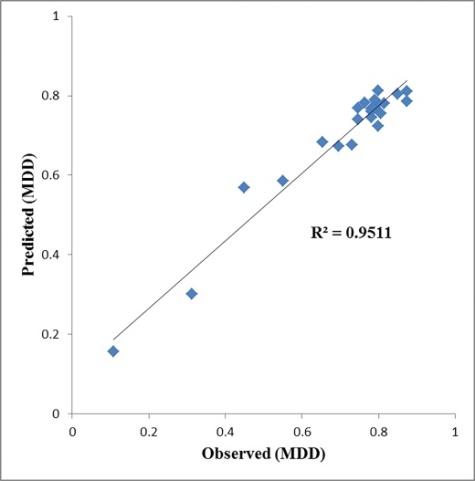
*(a) (b)*

***Figure 5:*** *Training results: (a) predicted vs. observed (OMC), (b) predicted vs. observed (MDD) for Levenberg-Marquardt’s model.*

*(a) (b)*

***Figure 6:*** *Validation results: (a) predicted vs. observed (OMC), (b) predicted vs. observed (MDD) for Levenberg-Marquardt’s model.*

*(a) (b)*

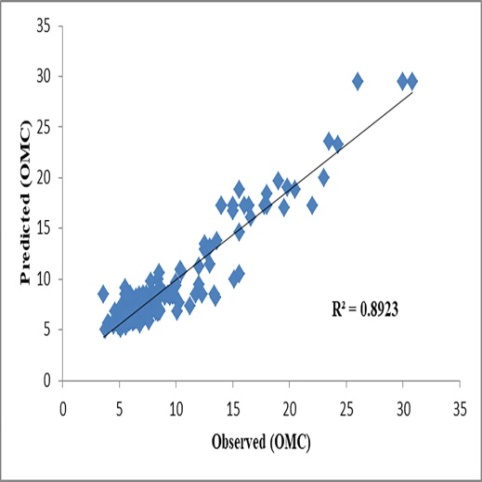
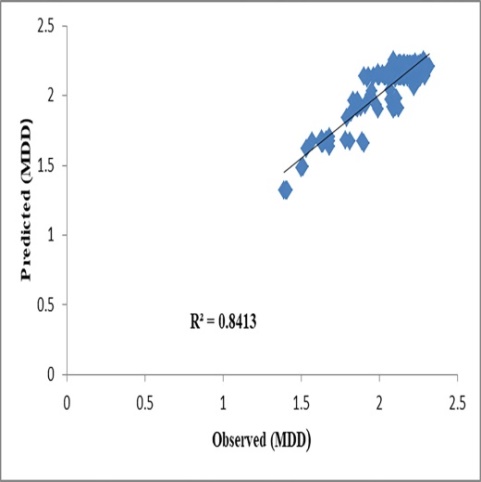
***Figure 7:*** *Testing results: (a) predicted vs. observed (OMC), (b) predicted vs. observed (MDD) for Levenberg-Marquardt’s model.*

**Stochastic Gradient Descent’s model**

This part of the paper discusses the results of the (ANN) model developed using stochasticgradientdescent. The results of training, validation, and testing are shown in table 4 and figures 8 to 10.

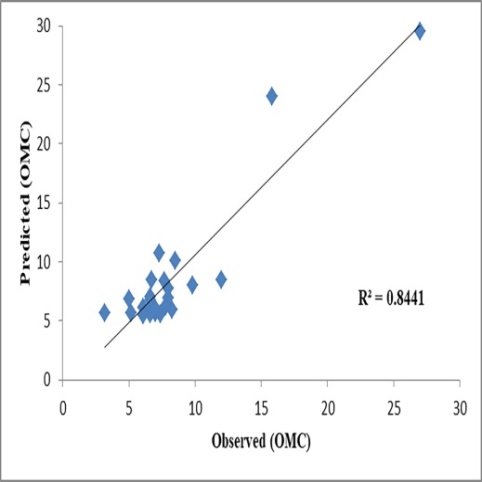
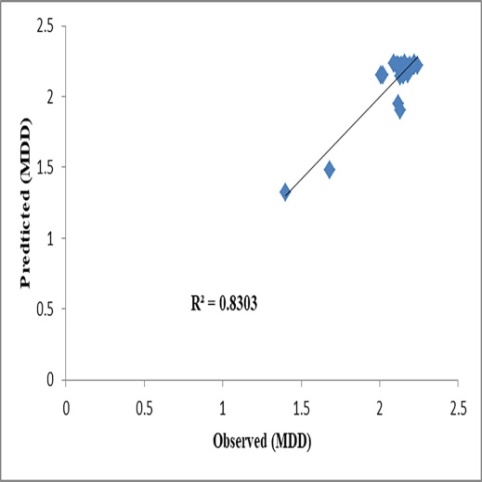
***Table 4:*** *The results of the stochastic gradient**descent**model.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Performance  metric | Training | | Validation | | Testing | |
| OMC | MDD | OMC | MDD | OMC | MDD |
| MAPE (%) | 13.97 | 3.05 | 19.96 | 3.99 | 20.98 | 2.55 |
| MSE | 2.5655 | 0.0064 | 5.8192 | 0.0105 | 5.824 | 0.0047 |
| RMSE | 1.6017 | 0.0799 | 2.4123 | 0.1022 | 2.4133 | 0.0688 |
| R2 | 0.8923 | 0.8413 | 0.8441 | 0.8303 | 0.8331 | 0.9055 |

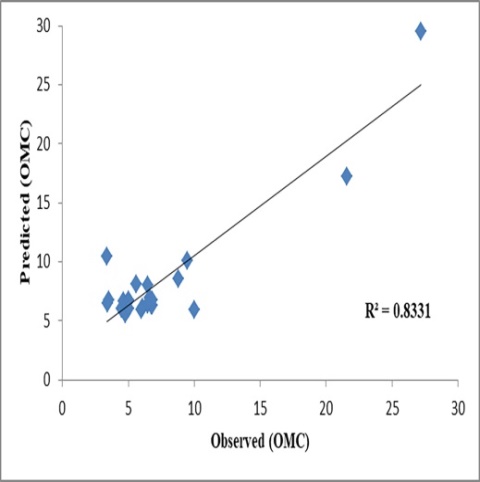
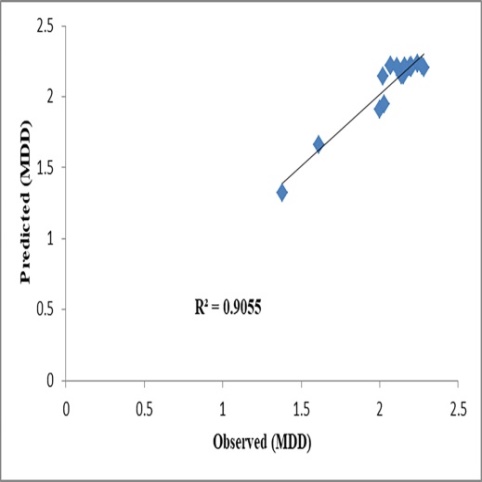
*(a) (b)*

***Figure 8:*** *Training results: (a) predicted vs. observed (OMC), (b) predicted vs. observed (MDD) for Stochastic Gradient Descent’s model.*

*(a) (b)*

***Figure 9:*** *Validation results: (a) predicted vs. observed (OMC), (b) predicted vs. observed (MDD) for Stochastic Gradient Descent’s model.*

*(a) (b****)***

***Figure 10:*** *Testing results: (a) predicted vs. observed (OMC), (b) predicted vs. observed (MDD) for Stochastic Gradient Descent’s model.*

The two models achieved high (R2), and low error indicated by the three metrics (MAPE, MSE, and RMSE) during training, validation, and testing.

**CONCLUSION**

This research showed that the developed models can predict the optimum moisture content and the maximum dry density at a standard proctor compaction effort for different types of soils. This is indicated by their small error and a high coefficient of determination.

The ability to predict the soil's compaction characteristics without the need for the compaction test has a lot of benefits. It can save time and resources during the geotechnical investigation stage. Such models can allow engineers around the country to decide in an instant without needing to transport soil samples several kilometres to be tested in another state. They can bring development to less developed areas in the form of large-scale projects like road construction, embankments for irrigation canals, earth-fill dams, etc…

The models are hence deemed to be satisfactory in predicting the compaction characteristics of soils at a standard proctor compaction effort.

**ACKNOWLEDGMENT**

The authors would like to thank all the technicians and engineers at the (BRRI). Their efforts and data helped us in completing this paper.

**LIST OF REFERENCES**

Alavi, A. H., Gandomi, A. H., Mollahassani, A., Heshmati, A. A., & Rashed, A. (2010). Modelling of maximum dry density and optimum moisture content of stabilized soil using artificial neural networks. *Journal of Plant Nutrition and Soil science*, 368-379.

Atalar, C., Adu-Parkoh, E., B.M.Das, & A.Ozyanki. (2019). Comparison of multiple regression, and artificial neural networks in estimating compaction characteristics. *European Conference on Soil Mechanics snd Geotechnical Engineering.* Reykjavik.

Jayan, J., & N.Sankar. (2015). Prediction of Compaction Parameters of Soil using Artificial Neural Network. *Asian Journal of Engineering and Technology, 3*(4), 368-375.

Jeremiah, J. J., Abbey, S. J., Booth, C. A., & Kashyap, A. (2021). Results of Application of Artificial Neural Networks in Predicting Geo-Mechanical Properties of Stabilised Clays-A Review. *geotechnics*, 147-171.

Knappett, J., & Craig, R. (2012). *Craig's Soil Mechanics, Eighth Edition.* Spon Press.

Salahudeen, A., Ijimdiya, T., Eberemu, A., & Osinubi, K. (2018). Artificial Neural Networks Prediction of Compaction Characteristics of Black Cotton Soil stabilized with Cement Kiln Dust. *Journal of Soft Computing in Civil Engineering*, 50-71.

Sivrikaya, O., & Soycan, Y. T. (2009). Estimation of compaction parameters of fine-grained soils using Artificial Neural Networks. *International Conference on New Developments in Soil Mechanics and Geotechnical Engineering*, (pp. 406-413). Nicosia.

Tenpe, A., & Kaur, S. (2015). Artificial Neural Network Modeling For Predicting Compaction Parameters Based on Index Properties of Soil. *International Journal of Science and Research, 4*(7), 1198-1202.