Review

Results of Application of Artificial Neural Networks in Predicting Geo-Mechanical Properties of Stabilised Clays - A Review

Jeremiah J. Jeremiah 1, Samuel J. Abbey 2,\*, Colin A. Booth 2 , Anil Kashyap 2

1 Ph.D. Candidate Civil Engineering Cluster, Department of Geography and Environmental Management, Faculty of Environment and Technology, University of the West of England, Bristol BS16 1QY, UK 1; jeremiah.jeremiah@uwe.ac.uk

2 Faculty of Environment and Technology, University of the West of England, Bristol BS16 1QY, UK 2; samuel.abbey@uwe.ac.uk

\* Correspondence: samuel.abbey@uwe.ac.uk

|  |
| --- |
| **Citation:** Jeremiah, J. J.; Abbey, S. J.; Booth, C. A.; Kashyap, A. Application of Artificial Neural Networks in Predicting Geo-Mechanical Properties of Stabilised Clays. *Geotechnics* **2021**, *1*, Firstpage–Lastpage. https://doi.org/10.3390/xxxxx  Academic Editor: Firstname Lastname  Received: date  Accepted: date  Published: date  **Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.    **Copyright:** © 2021 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/). |

**Abstract:** This study presents a literature review on the use of artificial neural networks in the prediction of geo-mechanical properties of stabilised clays. In this paper, the application of ANN in geotechnical analysis of clay stabilised with cement, lime, geopolymers and by-product cementitious materials has been evaluated. Chemical treatment of expansive clays will involve the development of optimum binder mix proportions or the improvement of a specific soil property using additives. These procedures often generate large data requiring regression analysis in order to correlate experimental data and model the performance of the soil in the field. These analyses involve large dataset and tedious mathematical procedures to correlate the variables and develop required models using traditional regression analysis. The findings from this study show that ANN due to their robust data analysis and correlation capabilities are becoming well known in dealing with the problem of mathematical modeling involving nonlinear functions and have been successfully applied to the stabilisation of clays with high performance. The study also shows that supervised ANN model is well adapted to dealing with stabilisation of clays with high performance as indicated by high R2 and low MAE, RMSE and MSE values. The Levenberge-Marquardt algorithm is effective in shortening the convergence time during model training.

**Keywords:** ANN; Regression models; Stabilisation of clays; Predictive models; Artificial Neural Networks

1. Introduction

Expansive clays at shallow depths and other poor ground conditions are common problems faced in geotechnical engineering and these problematic soils pose enormous challenges to proposed infrastructural development [1], [2]. Oftentimes, the option of soil replacement is not economical and there is need to perform some form of ground improvement or stabilisation prior to construction [3]. Soil stabilisation may be done chemically or mechanically depending on the peculiarity of the problem at hand. However, the treatment of most expansive clays will involve the use of chemical stabilisation. Irrespective of the method used, it becomes imperative to investigate the performance of such stabilised soils using key parameters of the soil such as Unconfined Compressive Strength (UCS), Maximum Dry Density (MDD), Optimum Moisture Content (OMC), California Bearing Ratio (CBR), Plasticity index(PI), Liquid Limit (LL), Plastic Limit (PL), grain size, specific gravity GS etc. [4-8]. These parameters are most times influenced by several variables leading to a multi-variable problem for which researchers have used several statistical methods to analyze, investigate and model trends using laborious mathematical processes involving multivariable regression analyses.

Artificial Neural Network (ANN) is an emerging trend in dealing with ground improvement involving weak clays. ANN is a computing tool with its architectural concept drawn from that of the biological brain [9-11]. Like other artificial intelligence application or machine enabled activities, the application of ANN in geotechnical engineering with regards to soil stabilisation is to overcome the supposed complexities and shortcomings of traditional methods in other to save time and resources while obtaining more reliable results. This study explores the usefulness of ANN as a stand-alone machine learning tool in the stabilisation of expansive clays and evaluates its capability in the development of useful mathematical models for performance evaluation of key properties of stabilised clays. Previous experimental studies are reviewed with a view of assessing the state-of-art in terms of model development, training, and performance evaluation.

1.1. Artificial neural network

ANN as a branch of artificial intelligence is simply an automated optimisation system capable of learning the relationship and inter-dependencies between multiple input variables of a given system and modelling such relations (trends and patterns) in the form of mathematical functions for easy prediction [12]. ANN has been successfully used in the study of complex systems to identify patterns and model real-life problems relating to complex behaviours involving nonlinear functional relations. The capability of ANN to discover the mapping between several domains of data has drawn the interest of many researchers in geotechnical engineering [13]. ANNs are classified based on numerous criteria such as the learning condition (supervised and unsupervised networks), based on model topography (feedforward or recurring networks), based on number of hidden layers (shallow or deep networks), based on training algorithm (Back-Propagation Networks, Hopfield Networks, Self-Organizing Map Networks [14].This paper simplifies the underlying concepts of back propagation ANN models and explores its applicability in modeling the behaviour of stabilised clays viz a viz predicting the response of key soil parameters in other to clear the wide-spread complexities and misconceptions associated with the method and to encourage its use in soil stabilisation problems for more reliable solutions.

1.2. Components of artificial neural network

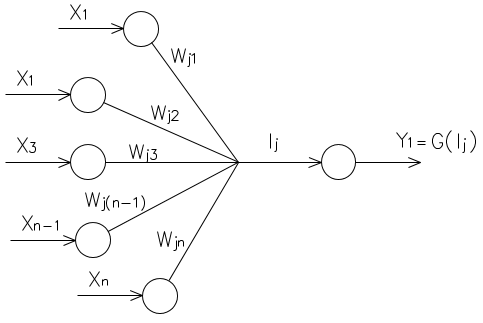
1.2.1. Neurons and edges

ANN building blocks is a collection of neurons (nodes) and links mimicking the biological neural network as shown in Figure 1. The neurons are linked to other neurons by edges and are connected to others so that results from preceding neurons might automatically become inputs for succeeding ones, thereby creating the network. These neurons are the data collection or processing points in the network. Here, signals (input) are processed and transferred to other neurons through the connecting links with each neuron generating a unique output which may become inputs to multiple neurons. In this current subject area of application, these inputs would be laboratory results of key soil parameters described as the dependent variables. The input value of a given neuron is simply obtained by computing the weighted sum of the inputs from connected neurons with the addition of a bias [15, 16]. This output of the weighted summation then becomes the input for the activation function- a linear or non-linear function [17], [18].

|  |  |
| --- | --- |
|  |  |
| (**a**) | (**b**) |

**Figure 1. (**a-b). Biological Neural Network [10].

Edges are the links or connections between neurons and convey signals with associated weights depending on the influence of the input from such a link on the output of a given neuron [19]. Inputs parameters with greater importance are assigned higher weight than those with lower importance [20]. For example, in a soil classification problem, the weights will be dependent on the contribution of the features in determining the class of a soil [21]. In a typical perceptron as in Figure 2, the connection weights can be represented as Wj which describes the importance of the connection.



**Figure 2.** Typical Neuron showing Associated Edge Weights.

2. Materials and Methods

The propagation function [20] computes the weighted summation of inputs from each preceding neuron with the addition of a bias term to create an activation which is utilized by the activation function in computing the output signal. For a typical case such as represented in Fig. 2, G(I\_j) is the activation function where I\_j is the propagation function defined in equation 1 [22, 23]. The activation function could be any of binary step function (equation 3), logistic function (equation 4), hyperbolic function (equation 5), Gaussian function (6) etc. [12].

|  |  |
| --- | --- |
| + | (1) |

is the associated weights of the connections.

are the input data

is the threshold or bias

n is the number of inputs

The connection weights Wji, are continuously adjusted for each training iteration step using the expression below [24].

|  |  |
| --- | --- |
| = | (2) |

In equation 2, is the change in the weight of a given connection, represent the learning rate, the error in a neuron’s output and the momentum respectively, while m is the iteration number. The momentum is used in factoring the weight correction and plays an important role in model training in the sense that too large values, causes the model to skip the optimum values of the weight causing oscillations. On the other hand, too small values will result in delayed convergence between predicted and experimental values [25].

An activation function such as the binary step function expresses the output of a given neuron as either of 0 or 1 depending on the input x and is given here as equation 3.

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

Binary function due to the gap in the output (only 0 and 1), is unsuitable for training algorithms which rely on the derivatives of the activation functions [26]. The logistic function also limits the output of a neuron to 0-1 and is expressed by [27].

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

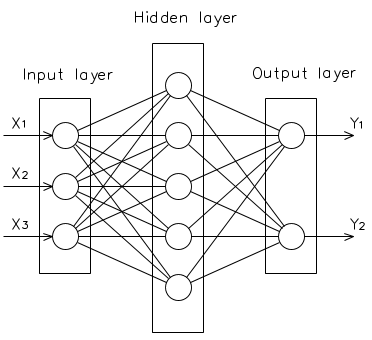
The hyperbolic tangent function as shown in equation (5), is also a commonly used activation function in ANN [28], and it limits the output of a neuron to -1 and 1. The gaussian function is another commonly used activation function that defines the normal distribution of a given random input variable about its mean, [29] as presented in equation (6).

|  |  |
| --- | --- |
|  | (5) |
|  | (6) |

Where σ and µ are the standard deviation and mean respectively.

2.1. ANN Architecture

Deciding the ANN topography is a critical part of the model development and involves an iterative trial and error process (training) [17, 30-32]. In most studies, an initial model topography is assumed and trained while monitoring the performance of the model using predefined statistical measures such as coefficient of determination (R2), root mean square error (RMSE), mean average error (MAE), and mean square error (MSE). The hyperparameters are continuously modified and the model retrained until an optimum model architecture is obtained with lowest error and highest R2. [33-35, 11]. This training in simple terms, is “showing the network an example” of the problem using experimental input and output data. Many training algorithms exist; quasi-newton backpropagation, bayesian regularization backpropagation algorithm, gradient descent, Levenberg-Marquardt optimization etc. but the process is similar and begins with feeding the model with quality dataset and allowing the system to process this data in order to learn the relationship between the variables and hence generate weighted associations between the data within the network and predict result. The predicted result is then compared with the experimental result to evaluate the error which is then used to modify the weights of the connections by a reverse error minimization process using a chosen cost function. The process is repeated until there is insignificant change in the output of the cost function [36, 37]. Based on the performance of various models with different hyperparameters, the best model topography is then selected.



**Figure 3.** A Typical ANN Architecture Showing Neurons and Layers.

2.2. Feedforward and recurring networks

The ANN topography is such that the neurons are grouped into layers namely, input layers and output layers and in some cases, there is need to include hidden layers between the input and output layer making a multilayered perceptron neural network model in order to create sufficient degrees of freedom to avoid overfitting. The hidden layer could be made one layer (shallow networks) as shown in Figure 3 or multiple layers (deep neural network, DNN) [38]. In addition, the connections between layers could be in such a way that a neuron in one layer could be connected to all neurons in the succeeding layer and is said to be fully connected resulting in a larger number of neurons in the succeeding. Also, the models could be organized such that multiple neurons in a layer are connected to a single neuron of the succeeding layer. The latter condition is said to be a pooled connection and is synonymous with lesser number of neurons in the succeeding layer [39]. Even though one may be tempted to believe that a larger number of neurons will always result in better prediction using ANNs. However, the optimum number of neurons will depend on several factors such as the amount of data and the complexity of the relationship. In certain types of NNs, like the Deep Neural Network, the number of neurons has lesser effect on the overall performance of the network than the number of layers [12] and DNNs with more hidden layers have shown to yield more results than shallow networks (network architecture with a single hidden layer). However, the number of hidden layers to be used in each network will depend on the complexity of the mapping between input and output domain, the quality and amount of data [40]. Also, even though it is expected that a greater number of experimental dataset used in training the model will improve the performance of the model, recent studies have shown that it might be advisable to use less experimental dataset of high quality than a large amount of experimental data which may be prone to errors [12]. Moreso, the quality of the output is dependent on how the input database is utilized in the training. In terms of the way data is transferred from one layer of the network to another, one can generalize that there are two broad categories of ANN architecture - the recurring network and feedforward network. In the recurring network, there’s connection between neurons of a given layer and that of preceding and/or succeeding layers, forming a loop and allowing an input to be processed many times by the same neuron. Whereas, in the feedforward network, neurons in each layer are only connected to neurons in other layers.

|  |  |
| --- | --- |
|  | |
| (**a**) | (**b**) |

**Figure 4.** (a) Recurring Network (b) Feed Forward Network.

2.2.1. Data Preprocessing

The input dataset used to develop ANN models for geotechnical engineering problems will comprise of various range of input variables because the ANN neurons connection weights are representatives of the importance of the variables. The weights of the connections are influenced by the euclidean distance which for any given points y1(x11, x21, x31…xn1) and y2(x12, x22, x32…xn2) in a data space can be expressed as Equation 7. Therefore, in other to ensure that proper significance is attributed to the features, it is imperative to scale the features. Feature scaling can be achieved using methods such as standardization (Z-score normalization) or max-min normalization method. For standardization, the feature is expressed in its standardised form as in equation 7 and 8 below.

|  |  |
| --- | --- |
| = | (7) |
|  | (8) |

In Equation 7 and 8, is the standardised value of , is the mean and is the standard deviation. The input variables can also be scaled using the max-min normalisation to achieve the same purpose using the expression. Where is the normalised value of and are the minimum and maximum values respectively.

3. Selecting design parameters for ANN in soil stabilisation

The training of the neural network is actually the process of selecting or ‘designing’ the best network model parameters (hyperparameters). However, there is a start point where a first architecture is proposed. Selecting the right number of neurons in the hidden layer is critical as it influences the performance of the network. Too fewer number of neurons in the hidden layer can lead to underfitting. In this case, the training data presents a more complex problem than the network is modelled to handle. For some problems, increasing the number of neurons by introducing additional features can make learning easier for the model and resolve the problem. In some other cases, for example, in some geotechnical applications, the input parameters of interest may have been predetermined and measured and this option may not be feasible. Alternatively, increasing the number of hidden layers and neurons may be helpful but again, if the number of neurons become excess, this leads to overfitting. In this case, the model is training on less complex data than it is designed to analyze. The effect is that the model is unable to properly generalize on new data set outside the training data as the weights are not optimally adjusted. Also, an excessively high number of neurons in the hidden layer can extend training time and lead to poor training even with sizable database for training. The goal, therefore, is to find a balance. However, this is not a straight-forward process. A good step would be to make a good initialization of the network parameters and there are several ideas concerning how these parameters can be initialized. Amongst the huge suggestions which are available and of course effective under different conditions, one simple empirical rule for selecting the number of neurons in the hidden layer is to use the mean of the number of neurons in both input and output layers. Some other idea involves taken about 60-70% of the total input and output neurons. In general, the idea is to provide a reasonable start so that during the training, the model can be optimised or pruned, and redundant neurons can be removed based on the assigned weights while keeping track of the performance. However, for most regression analysis relating to stabilisation, one hidden layer has been found very effective. In rare cases, two hidden layers have been used but there are seldom cases where over two layers have been needed in developing reliable predictive models. As shown in the succeeding section, almost all the applications in soil stabilisation have utilised one hidden layer with one or two utilizing more than one hidden layer.

3.1. Training, validation, and testing

As mentioned in the earlier section, the training of most neural network applied to modeling geomechanical properties of stabilised soils are done under supervised condition. In this type of training, the supplied data is partitioned, and a part (training dataset) utilised in learning the relationship between the variables, thereby providing the initial weights. This sample is continuously fed to the network with a view to understanding the data rather than recognizing it. If the network learns progressively, it converges with reduced error after each iteration until a pre-defined error range is attained. The quality of the training dataset influences the convergence of the model [41]. A dataset may lack the necessary independent variables required for the model to understand the data and hence can lead to non-convergence. A very small sample space can lead to the network memorizing rather than learning. Hence, it is important that part of the data is separated to be used in evaluating the training. This is the validation dataset. The major aspect of developing a suitable model would then be to continuously monitor and tweak the number of neurons in the hidden layer, or number of hidden layers, modify the activation function or even the training algorithm [42]. The model validation utilizes the successive trial of the trained model on the validation dataset [43]. This is an unbiased evaluation of how well the model understands the training data. The final test of the model’s predictive ability is carried out on the test dataset which was never seen by the model. In some cases, the data is continuously partitioned into two as in cross-validation. The dataset are switched and utilised for training and validation in a crossed pattern.

The performance of models is usually evaluated by using statistical measures such as the coefficient of determination (R2), the mean absolute error (MAE), the root mean square error (RMSE), the mean square error (MSE) and other. The R2, MAE, RMSE and MSE expressions are defined in Equation 9, 10, 11, and 12 respectively.

|  |  |
| --- | --- |
| R2 = | (9) |
| MAE = | (10) |
| RMSE = | (11) |
| MSE = | (12) |

For which and are experimental and predicted values of a given dependent variable while and are the mean values of the predicted and experimental values.

3.2. Estimating the amount of training data

Determining the sample size required for successful model training is a vital step in successful model development. A common start off point is the ‘’rule of ten’’ which proposes that the training sample size is taken as not less than ten times the number of network parameters. The number of parameters maybe estimated as the number of edges or connections including biases neuron. It is expected that the performance of the model would improve with increasing sample size following a power function up to a point is reached where there is no significant increase in performance. Usually, in practical situations, the data set is split in ratio of 70%:30% or 80%:20% where the higher percentage is that of the training sample space. Although this split is only important with relatively low sample size. The underlying idea is to make available sufficient example data with which the network is trained and evaluated. Too few training data will result in a higher variance of the network parameters. Also, two few testing data will create higher variance during evaluation of the performance of the model.

4. Application of ANN in predicting the properties of stabilised clays

The variability of soils following uncontrollable and imprecise chemical and mechanical processes of their formation makes it very complex to predict their behaviour even in the natural state. Several mathematical and graphical models have been proposed for use in modeling the post-stabilisation behaviour of treated clays through prediction of various engineering parameters for the purpose of design and construction. These methods relay on laboratory results of a few samples with which in-situ post-stabilisation behaviour is to be predicted. However, in many cases, the post-stabilisation behaviour of any soil will be dependent on multiple variables such as curing time, curing duration, soil type, binder content, curing temperature, moisture content, compaction method and effort, plasticity etc. which are key and influence the dependent variables [44], [45]. The large variability of input parameters, the extensive laboratory experiments required and the unknown relationship between the variables put together makes it even more complex to predict the behaviour of the soils [46], [47]. In other to simplify the problem, many studies tend to concentrate on a set of few parameters and employ simple mathematical models to map the domains of input to out variables. An example of such simplified mathematical models is the multiple linear regression model which for a given set of features can be related to its dependent variable as stated in equation 13 below [23].

|  |  |
| --- | --- |
| = + +++ … + ε | (13) |

Where, ,, ,… are the coefficients and ε is the error.

ANN application has shown good results in terms of prediction of engineering parameters of stabilised soils for various purposes. ANN’s ability to learn the relations between a wider and more complex set of experimental variables and map these variables to the target output domain using well-adjusted weights makes it a more reliable tool in the prediction of the in-situ post-stabilisation behaviour of treated clay soils. In addition, the ability of ANNs to simultaneously handle multiple dependent and independent variables using the same experimental dataset and its adaptability in finding correlations for highly non-linear data which are characteristic of many civil engineering problems makes it more advantageous over traditional regression analysis [48], [49]. A review of studies modeling various soil properties is presented in the subsequent sections. The review is in a bid to explore the input parameters considered in the study, the number of data utilised in model development, the training algorithm utilised, the hyperparameters of the model, the model performance and finally the results or predictive models developed.

4.1. Unconfined compressive strength

ANN has been employed in tracking and modeling the UCS of geopolymer stabilised clays [50]. In the study, ground granulated blast furnace slag (GGBS) and pulverised fly ash (PFA) were considered as binders for the improvement of the compressive strength of three clays with varying properties. The soil characteristics such as liquid limit (LL), plastic limits (PL) etc. were evaluated according to the Indian standards and in combination with key experimental variables were utilised as input data for the ANN model development. Eight input variables namely, LL, plasticity index (PI), GGBS content, PFA content, molarity of alkaline activator used (M), the ratio of the activator to binder, the ratio of sodium to aluminum in the activator-binder mixture (Na/Al), the ratio of silicon to aluminum in the activator-binder mixture (Si/Al), and 28 days UCS were considered. A multi-layered perceptron ANN model was chosen with 1 hidden layer to study the experimental data. The optimum architecture was selected by varying the number of neurons in the hidden layer while evaluating the performance of the model and was obtained as 1 hidden layer with 9 neurons. The ANN showed better ability to learn the relationship of the data set as shown in Figure 5.

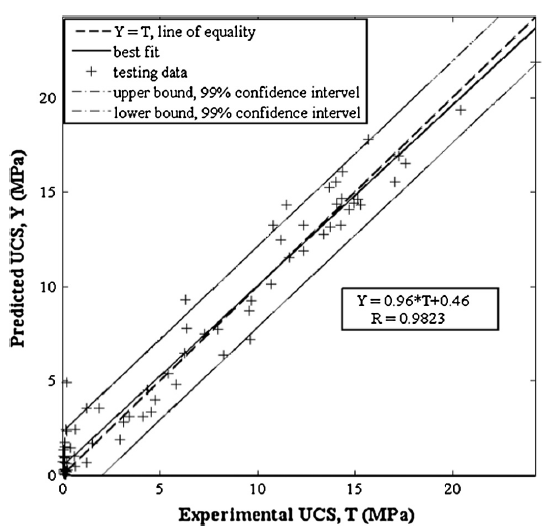


Figure 5. Experimental versus ANN Predicted UCS values [50]

From Figure 5, it is obvious that almost all data points are within the 99% confidence limit. The line of best fit is almost aligned with the line of equality at which all experimental observations and predicted values are the same. Table 1. compares the performance of the ANN model with that of a multivariable regression analysis (MVR) on the same data set. Its is obvious that ANN performed better at correlating the independent variables and their influence on the UCS of the stabilised soils.

**Table 1.** Statistical Evaluation of Performance of ANN and MVR models [50].

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Dataset | Statistical Parameter | | |
| **R2** | **MSE** | **MAE (%)** |
| ANN | Training data | 0.992 | 0.34 | 3.65 |
| Testing data | 0.964 | 1.50 | 8.34 |
| MVR | Training data | 0.828 | 7.24 | 19.20 |
| Testing data | 0.808 | 8.04 | 19.26 |

Expansive soils are known for their vulnerability to significant volume changes with moisture due to their high plasticity properties. Table 2 shows the classification of expansive clays according to the Building Research Establishment (BRE) based on a modified plasticity index.

Table 2. Classification of shrink-swell clays [51]

|  |  |
| --- | --- |
| Modified plasticity index | Volume change potential (VCP) |
| >60 | Very high |
| 40-60 | High |
| 20-40 | Medium |
| <20 | Low |

In another study, [52] employed a feedforward ANN model in studying the UCS of expansive clays stabilised with cement-kiln dust (CKD). The liquid limit, plastic limit and plasticity index of the clay was determined to be 48.2%, 27.2%, and 21% respectively. The expansive clay is characterised with a free swell of 80%. The natural and stabilised clay samples were subjected to compaction and UCS test in accordance with [53] and [54]. Three ANN models were developed with a total of eight input variables namely specific gravity (Gs), linear shrinkage (LS), coefficient of uniformity (CU), coefficient of gradation (CC), LL, PL, optimum moisture content (OMC), and maximum dry density (MDD) for 3 UCS outputs (7, 14 and 28 days). A total of 72 data sample data were utilised in model development. The ANN model topography consisted of one input layer with 8 neurons and one output neuron (the UCS value). However, the number of hidden layers was varied in order to determine the optimum number of neurons in the hidden layer. Figure 6 shows the performance of the trial models.

Figure 6. Selection of optimum number of neurons in hidden layer [52]

The optimum model was found to be of nine neurons in the hidden layer based on lowest MSE. [52] suggests that the MSE is a more reliable parameter for network selection when R values alone become insufficient for optimum network selection. The results of the analysis showed that the ANN model was able to predict the variation of UCS as a function of the predictor variables with a high correlation coefficient as seen in Figure 7.

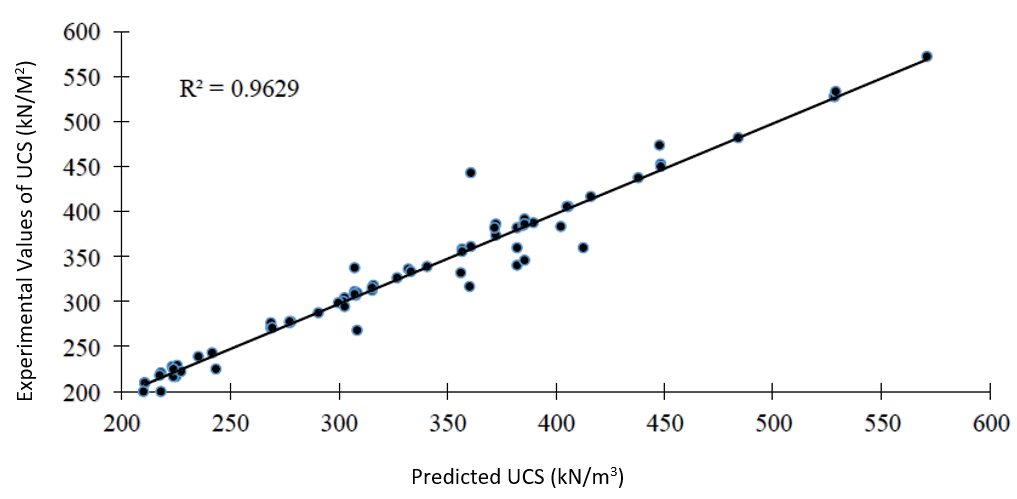


Figure 7. Observed Versus Predicted UCS [52]

The study reveals a high correlation between experimental values and ANN predicted values. The high correlation coefficient with low RMSE confirm the performance of the model.

Moreso, [55] conducted experimental investigation on the suitability and performance of a kaolin clay stabilised with fly ash (PFA), rice husk ash (RHA) and cement. The improvements in the soil was ascertained by considering increase in the UCS. The LL, PL, and PI of the unstabilised kaolin clay was determined as 43.3%, 19.5% and 23.8%. The input variables considered for correlation were clay content, RHA content, cement, PFA content, and curing duration. The number of neurons in the hidden layer was selected manually and varied while evaluating the performance of the model using the MSE. Figure 8. Shows the variation of MSE and the number of neurons in the hidden layer.

Figure 8. Selection of Optimum number of neurons in hidden layer [55]

As seen in Figure 8, the optimum model performance was found at 10 neurons in the hidden layer based on Levenberg- Marquardt algorithm. The results of the analysis using the optimized model showed a good correlation between the datasets. The predictive model developed from ANN is given in equation 14.

= (14)

Where is the dependent variable to be estimated and represent the UCS for a given combination of independent variables. A comparison of ANN performance with multivariable regression analysis further proves ANN’s advantage as shown in Table 3. ANN has been described as a black box since it is difficult to predict the way the predictive model will be selected. However, sensitivity analysis can be carried out in determining the influence of the input variables on the target variable.

**Table 3.** Statistical evaluation of performance of ANN and MVR models [55]

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Dataset | Statistical Parameter | | |
| R2 | MSE | RSME |
| ANN | Training data | 0.9813 | 0.0395 | 0.1987 |
| Testing data | 0.9714 | 51.34 | 7.1651 |
| MVR |  | 0.8870 | 68.7603 | 8.2921 |

In a separate study, [56] investigated the performance of stabilised soils with a view to predict the effects of dynamic disturbance on already stabilised soils. The study was to understand the effects of transportation and compaction on strength recovery of already setting stabilised soft clay. Hence, stabiised soil samples were agitated and compacted after curing for some days to simulate the disturbed sample. The UCS was utilised as performance evaluation parameter. The soils were treated with varying quantities of cement. A kaolin clay with liquid limit, plastic limit, and plasticity index of 77.5%, 30.3%, and 47.2% respectively. Ordinary Portland cement was used a binder for different ratios by mass of dry soil. The UCS test was conducted according to ASTM D2 166-06. It was required to estimate the influence of disturbance on the UCS in terms of an increase factor. ANN was employed in correlating the predictive variables and the dependent variable. A total of 80 experimental data points were generated from combinations of input variables namely, curing duration and compaction energy. Trial and error method was utilised in determining the network topography. However, the performance of one hidden layer was unsatisfactory due to high errors and low R2. The optimum network model was taken as two hidden layers with eight neurons each. The result of the analysis showed good understanding of the relationship and modeling of the increase factor with high coefficient of determination and low error. As seen in Figure 9, there is close approximation of the UCS by the ANN model. The prediction and experimental values are clustered around the line of equality for all three conditions of training, validation, and testing.

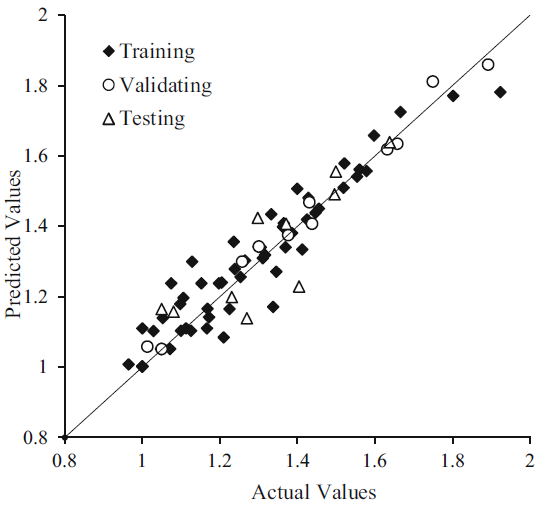


Figure 9. Predicted and experimental observations [56]

The UCS of cement stabilised clays was investigated by [57] in order to find a relationship between the several key variables. Different types of cements were utilised to study the effects of cement type on the UCS. The soil under investigation was collected at various depths below ground surface to also study the effect of confining pressure on the strength of the stabilised soils. Three machine learning algorithms were utilised. ANN was used in correlating the predictive variables. A total of 216 experimental data points were generated from various combination of fourteen input variables namely, soil type, moisture content (MC), bulk density (We), mass of cement (CM), sample diameter (DI), length of sample (L), cross sectional area of sample (CA), volume of sample (SV), depth of sample collection (D), mass of sample (MS), sample density (DS), curing condition (CD), curing time and cement type (CT). ANN training was done using Quasi-Newton method, Stochastic Gradient Descent, and Adam in order to select the optimal hyperparameters for the ANN topography. The dataset was divided in the ratio of 80% and 20% for training and testing. Performance evaluation of the developed ANN model showed good correlation. ANN performed better than the other machine learning algorithms. The high number of neurons in the hidden layer may be due to the multiple dimensions of the input vector in the sample dataset. The study showed that ANN could be utilised effectively for stabilisation problems, to track and model a wide range of independent variables.

To alleviate the difficulties associated with the need for continuous experimental determination of UCS, [58] employed machine learning in analysing and modeling the performance of stabilised dredged sediments. A total of 51 experimental datasets were collated from existing literature for development of the ANN model. The input predictive variables were percentage moisture content (MC), cement content, air foam content, and waste fishing net content. The ANN model to topography was initialized to 2 hidden layers while varying the number of neurons in order to find the optimum model. This approach has been taken in most soil stabilisation applications. The optimum ANN architecture was made up of 2 hidden layers with 12 and 10 neurons based on training using the Levenberg-Marquardt algorithm. From the results of the analysis, the model has been able to find the trend within the available training dataset. However, for certain values of the water content, the error was high. For the model to generalize better, a wider training dataset may still be needed. Previous studies have suggested a training dataset at least ten times the network parameters. In many original research on stabilisation, there are usually limited experimental data. A cross-validation approach maybe combined for relatively lower sample size. The statistical evaluation of the performance of the model in training and testing is shown in Table 6.

4.2. California Bearing Ratio

California bearing ratio (CBR) of soils is of interest in pavement design. Several geotechnical investigations are concerned with measuring the CBR of stabilised soils. The aim of some soil stabilisation is for improvements in CBR to meet some desired outcome. However, determination of CBR is expensive. Its desirable to be able to correlate and predict CBR from known soil characteristics. A study by [59] involved modeling the CBR of stabilised expansive soils. The expansive clays were treated with lime and quarry dust as binder. The quarry dust is a byproduct from the crushing of rock for aggregates. The characteristics of the expansive clay such as Atterberg properties, compaction and CBR were determined following the Indian standards. The stabilisation of the soil was carried out by addition of lime and quarry dust of different percentages to the soil and tested of CBR at 0, 7 and 28 days. A total of 49 data samples were generated for correlation from combinaton of four imput parameters namely, LL, PI, OMC, and MDD. To develop the required regression models, three ANN models were utilised based on different training algorithms using Differential evolution (DE), Bayesian regularization (BR) and Levenberg- Marquardt (LM) algorithm. The training data was taken as 70% of the 49 data samples while 30% was utilised for testing. Again, the limitations in availability of extensive experimental data is seen and may be due to cost implications. However, the Levenberg-Marquardt-trained ANN performed better based on the limited data sample. Also, the models showed signs of overfitting which may be due to insufficient training data. Table 6 shows the summary of model performances for the 28day CBR model.

Experimental study by [60] employed ANN in modeling soaked CBR of expansive stabilised clays. The clay was stabilised with varying proportions of coal ash, bagasse ash and groundnut shell ash (GSA) and geogrid layers. The soil classification was conducted in line with Indian Standards. The input variables used were type of ash, mix proportion, LL, PL, MDD, OMC and number of geogrid layers. The soil was stabilised with various combination of the different features to generate 210 data samples which was utilised in training and testing the ANN model. The optimum model architecture was found to be made of one hidden layer with 7 neurons. The model was trained with different algorithms in order to select the best algorithm based on R2 and MSE. The performance of the models based on training algorithm is given in Table 4.

Table 4. ANN performance using various training algorithms [60]

|  |  |  |
| --- | --- | --- |
| Training algorithm | R | MSE |
| Conscience bias learning function | 0.7693 | 7.08 |
| Gradient descent weight and bias learning function | 0.8991 | 3.98 |
| Gradient descent with momentum weight | 0.9163 | 6.66 |
| Levenberg-Marquardt function | 0.94317 | 0.49 |
| Hebb weight learning rule | 0.8761 | 2.4 |

The Levenberg-Marquardt training algorithm was selected as the best performing based on R and MSE values. The model performed well with good predictive ability. The results showed that the ANN model was capable of modeling the soaked CBR of the stabilised soils. The performance of black cotton soil stabilised with fly ash and geotextile was studied by [61] with a view to develop a regression model for predicting the CBR of stabilised black cotton soil. The soil used in the study was characterised according to Indian standards. The LL and PL of the soil was reported as 47% and 17% respectively. Also, the compaction characteristics were determined using standard proctor compaction tests with MDD and OMC given as 16.26kN/m3 and 15% respectively. The ANN models were developed using LL, PL, fly-ash content, OMC, MDD, and number of geotextile layers as input variables while soaked CBR was the target variable in the output neuron. Different training algorithms were employed while varying the number of neurons in the hidden layer to select the best model architecture. The performance of the models is presented here as shown in Table 5.

Table 5. ANN performance using various training algorithms [61]

|  |  |  |
| --- | --- | --- |
| Training algorithm | R | MSE |
| Quasi-Newton back propagation | 0.88712 | 1.083x10-4 |
| Bayesian regularisation back propagation | 0.85190 | 4.983x10-5 |
| Conjugate gradient back propagation with Powell-Beale restarts | 0.94122 | 3.776x10-7 |
| Conjugate gradient back propagation with Fletcher-Reeves updates | 0.81167 | 7.339x10-6 |
| Conjugate gradient back propagation with Polak-Ribiére updates | 0.85819 | 2.964x10-9 |
| Gradient descent back propagation | 0.94862 | 9.985x10-9 |
| Levenberg-Marquardt back propagation | 0.98695 | 8.0242x10-11 |
| One-step secant back propagation | 0.92335 | 1.388x1010 |
| Scaled conjugate gradient back propagation | 0.96904 | 1.946x10-6 |

The Levenberg-Marquardt function was adopted based on minimum error and highest R2. The models performed well with minimum error in relation to higher coefficient of determination. A regression analysis of the CBR of soils stabilised with lime and rice husk ash (RHA) was carried out by [62]. The LL and PL of the soil was reported as 34% and 20% with clay and silt content of 20% and 71%. Also, the MDD and OMC of the soil was reported as 17.7kN/m3 and 14.7% respectively. The CBR test was conducted in line with ASTM D1883 -07. Soaked CBR tests was conducted on 0, 7 and 28 days cured samples. The predictive variables considered in the study were the percentage of RHA, percentage of lime, curing duration, OMC, and MDD. The topography of ANN model was determined by trial and error through varying the number of neurons in the hidden layer. Optimum model topography was found to be 1 hidden layer and 12 neurons from the R value and MAE. The model performance as per the training data showed higher R and low MAE. The number of data utilised for training is deemed small. The performance of the model in generalizing new data may be affected due to overfitting. This again is a common factor confronting the application of ANN in soil stabilisation studies.

Furthermore, [63] investigated a predictive model for correlation of soaked and unsoaked CBR of stabilised black clay. The black clay is an expansive clay found in semi-arid regions of tropical and temperate climates. The stabilisation of the clay was done with cement kiln dust as binder. The characteristics of the soil and cement kiln dust were determined according to the British Standards [53] and [54]. Eight input parameters namely PL, LL, specific gravity (GS), linear shrinkage (LS), coefficient of uniformity (CU), coefficient of gradation (CC), OMC and MDD were utilised in developing 72 data samples for two ANN models; soaked and unsoaked CBR. The model architecture was determined by varying the number of neurons in the hidden layer from 1 to 20 while evaluating the MSE and R values. Figure 10 shows the performance of various model architectures.

Figure 10. Mean square error versus number of neurons in hidden layer [63]

The optimum was determined using lowest MSE value. As seen in Figure 10, for soaked CBR, a model architecture with 8 neurons in the hidden layer was chosen while 17 neurons in the hidden was optimum for unsoaked CBR. The models performed effectively in correlating and predicting the CBR of the stabilised soils with high R and low MSE.

4.3. Permeability and resilient modulus

The permeability of soil is an important property of soil describing the ease with which water flows through a soil mass. In certain geotechnical applications such as earth embankment and dam designs, it is desirable to be able to model and predict the permeability of the soil. The study by [64] investigated the effects of lime and pozzolan on the permeability of dispersive clays. Dispersive clays are highly susceptible to erosion due to their high sodium ion content which often leads to structural dis-integration into finer particles easily washed away during seepage. The LL and PL of the unstabilised clay were determined as 34% and 17% while the MDD and OMC were 17.85kN/m3 and 16.5% respectively. The clays soil was stabilised with various combinations of lime and pozzolan. Permeability tests were conducted to develop a dataset of 69 samples for the regression analysis. A total of six independent variables were used in the development of the ANN model including percentage passing 0.005 mm size sieve, PI, MDD, lime percentage, pozzolan percentage, and curing time (Cd). To develop the predictive model, ANN with varying numbers of neurons in the hidden layer were tried. An optimum model with 9 neurons in the hidden layer was selected based low MSE and high R value as with other studies. In line with other studies, 70% of the data was used for training, 15% for testing, and 15% for validation. The performance of the model was satisfactory in terms of low MSE and high R values.

Resilient modulus of soil is also a useful parameter in pavement design. Few studies are available on resilient modulus of soils. Hence, predictive models for resilient modulus of stabilised clays is desirable. ANN has also been utilised in correlating and modeling resilient modulus (Mr) of stabilised clays [65]. The LL, PL, particle size, and compaction characteristics of four unstabilised soil samples were determined according to ASTM D4318, ASTM D4318, ASTM D422, and ASTM D698. The resilient moduli were determined by subjecting treated samples to cyclic load test according to AASHTO T307 standard method (AASHTO 2003). A total of 125 data samples were generated for ANN modeling. Eight inputs parameters namely, cement content, lime content, PI, silt content, PFA, OMC, MC, and clay content were considered. Trial and error method was used to determine the optimum model as 1 hidden layer and 9 neurons. The performance evaluation was based on R values. The model could predict the resilient modulus with high correlation. The predicted and experimental values were very close which was indicative of the accuracy of the developed predictive model. Figure 11 shows the correlation between predicted and experimental data.

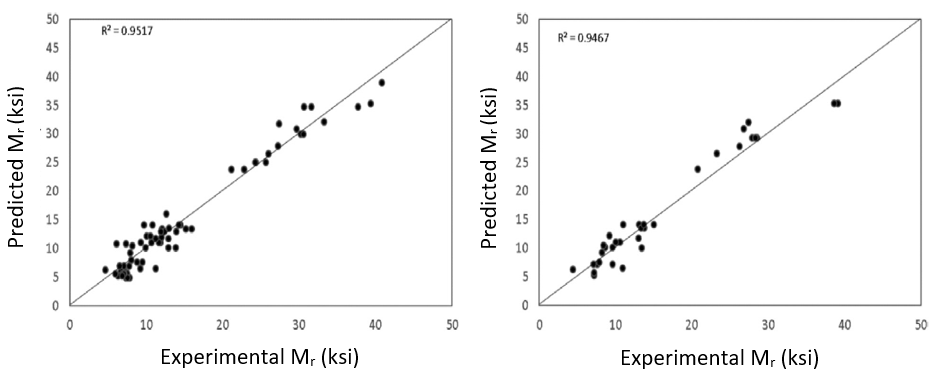


Figure 11. Experimental and predicted resilient modulus [65]

4.4. Plasticity index and compaction characteristics

ANN was also utilised by [66] in regression analysis to develop predictive models for MDD and OMC using eight input variables LL, PI, LS, clay-silt ratio, sand content, lime content, cement content, asphalt content in percentage. A total of 192 datasets were obtained from reported laboratory experiments and utilised in training based on gradient descent momentum algorithm for the development of the regression model for MDD and OMC predictions. The optimum model was obtained by trial and error using 52% of the data for training, 24% for testing, and 24% for validation. Performance evaluation of the models shows good correlation as with similar studies. Furthermore, [67] investigated the MDD and OMC of cement stabilised black cotton clay. The clay was treated with various percentages of cement kiln dust (CKD). The input parameters used were LL, PL, specific gravity (GS), linear shrinkage (LS), free swell, grain sizes (D10, D30 and D60), coefficient of uniformity (CU), and coefficient of gradation (CC). Two separate ANN models were developed for MDD and OMC using 5 and 7 neurons respectively in the hidden layers. The models were reported to have performed effectively and predicted the MDD and OMC with high correlation and low errors.

**Table 6.** Details of ANN Models and Performance.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | ANN Data | | | | Network Architecture | | | Model Performance | | | Reference |
| Soil Parameter investigated | **ANN input variables** | **No. of data used** | **No. of training data used** | **No. of testing data used** | **No. of validation data** | **No of hidden layers** | **No of neurons in hidden layer** | **Training algorithm utilised** | **Training** | **Validation** | **Testing** |  |
| UCS | 8 Inputs. LL, PI, GGBS content, PFA content, Molarity, activator/binder ratio, Na/Al ratio, Si/Al ratio | 283 | 70% | 15% | 15% | 1 | 9 | Bayesian regularizaton | R = 0.996  MSE = 0.34  MAPE(%) = 3.65 |  | R=0.982  MSE=1.50  MAPE(%)= 8.34 | [50] |
| UCS (7 days) | Gs, LS, CU, CC, LL, PL, OMC, MDD | 72 | 70% | 15% | 15% | 1 | 11 | - | R = 0.9782 | R = 0.996 | R = 0.9711 | [52] |
| UCS  (14 days) | Gs, LS, CU, CC, LL, PL, OMC, MDD | 72 | 70% | 15% | 15% | 1 | 11 | - | R = 0.9824 | R = 0.9843 | R = 0.9615 | [52] |
| UCS  (28 days) | Gs, LS, CU, CC, LL, PL, OMC, MDD | 72 | 70% | 15% | 15% | 1 | 11 | - | R = 0.9946 | R = 0.9992 | R = 0.9929 | [52] |
| UCS | Percentage of Clay, percentage of RHA, percentage of cement, percentage of PFA, curing duration(days) | 129 | 70% | 15% | 15% | 1 | 10 | Levenberg- Marquardt | R2= 0.9813,  MSE= 0.0395,  RMSE=0.1987 |  | R2= 0.9714,  MSE= 51.34,  RMSE =7.1651 | [55] |
| UCS | Curing duration, compaction energy. | 80 | 70% | 15% | 15% | 2 | 8 in each layer |  | R2= 0.924 | R2=0.908 | R2=0.889 | [56] |
| UCS | ST, MC, We, CM, DI, L, CA, SV, D, MS, DS, CD, curing time (days), CT | 216 | 80% | 20% |  | 1 | 50 | Adam | R=0.980 (Average),  MAE=115.29,  RMSE=231.2 |  | R=0.925 (Average),  MAE=292.2,  RMSE=419.82 | [57] |
| UCS | Moisture content, cement (percentage), air foam(percentage), and waste fishing net content (percentage) | 51 | 70% | 30% |  | 2 | 12 and 10 neurons | Levenberg- Marquardt | R=0.95,  MAE=3.9001,  RMSE=9.1948 |  | R=0.94,  MAE=8.6535,  RMSE=10.3390 | [58] |
| CBR  (28 days) | LL, PL, OMC, and MDD | 49 | 70% | 30% | - | - | - | Differential evolution | R=0.98 |  | R= 0.86 | [59] |
| CBR  (28 days) | LL, PL, OMC, and MDD | 49 | 70% | 30% | - | - | - | Levenberg- Marquardt | R=0.96 |  | R= 0.93 | [59] |
| CBR  (28 days) | LL, PL, OMC, and MDD | 49 | 70% | 30% | - | - | - | Bayesian regularization | R=0.96 |  | R= 0.93 | [59] |
| CBR | Type of ash, mix proportion (percentage), LL, PL, MDD, OMC and number of geogrid layers | 210 | - | - | - | 1 | 7 | Levenberg- Marquardt | R=0.94472 | R=0.93327 | R=0.94685 | [60] |
| CBR (28 days Soaked) | LL, PI, percentage of PFA, OMC, MDD, and no. of geotextile layers | - | - | - | - | 1 | - | Levenberg- Marquardt | R=0.99846 | R=0.98508 | R=0.92149 | [61] |
| CBR  (28days Soaked) | Percentage of RHA, percentage of lime, curing time (days), OMC and MDD | 48 | 70% | 15% | 15% | 1 | 12 | - | R=0.9948 | R=0.98909 | R=0.98895 | [62] |
| CBR (Soaked) | PL, LL, GS, LS, CU, CC, OMC and MDD | 72 | 70% | 15% | 15% | 1 | 8 | - | R=0.9988 | R=0.9996 | R=0.9976 | [63] |
| CBR (Unsoaked) | PL, LL, GS, LS, CU, CC, OMC and MDD | 72 | 70% | 15% | 15% | 1 | 17 | - | R=0.9912 | R=0.9993 | R=0.9806 | [63] |
| Coefficient of permeability (K) | percentage passing 0.005 mm, PI, MDD, lime percentage, pozzolan percentage, Cd | 69 | 70% | 15% | 15% | 1 | 9 |  | R=0.9968 | R=0.98883 | R=0.99405 | [64] |
| Resilient Modulus (Mr) | Percentage of cement, percentage of lime, PI, percentage of silt, percentage of PFA, OMC, MC, and clay | 125 | - | - | - | 1 | 9 | - | R=0.9517 | - | R=0.9467 | [65] |
| MDD | LL, PI, LS, clay-silt ratio, sand content, lime content, cement content, asphalt content in percentage | 192 | 52% | 24% | 24% | 1 | 18 | gradient  descent momentum | R2=0.9183  MSE=  0.28%  MAE=4.44% |  | R2=0.9101  MSE=0.26%  MAE=4.24% | [66] |
| OMC | LL, PI, LS, clay-silt ratio, sand content, lime content, cement content, asphalt content in percentage | 192 | 52% | 24% | 24% | 1 | 15 | gradient  descent momentum | R2=0.9025  MSE=  88.21%  MAE=118.37% |  | R2=0.8916  MSE=  89.57%  MAE=113.03% | [66] |
| MDD | GS, LS, free swell, D10, D30, D60, CU, CC, LL, PL | 90 | 70% | 15% | 15% | 1 | 7 |  | R=0.9946 | R=0.9715 | R=0.9754 | [67] |
| OMC | GS, LS, free swell, D10, D30, D60, CU, CC, LL, PL | 90 | 70% | 15% | 15% | 1 | 5 | - | R=0.9977 | R=0.9779 | R=0.8855 | [67] |

- Not Reported

5. Discussions

Problematic clays are commonly encountered in geotechnical engineering practice and will demand extensive ground improvement program under practical conditions. Undoubtedly, chemical stabilisation involving the use of calcium and non-calcium-based binders have been successfully used to stabilise and improve the engineering properties of clays, making them fit for a specified purpose. The stabilisation process will often require series of detailed laboratory procedures including sample preparation, curing and testing. This is followed by the analysis of the experimental data, observation of trends, establishment of relationship between variables and development of predictive models [69-72]. In the case of large collection of data involving multiple variables, it is very tedious to learn the functional relations and inter-variable dependencies without the use of a more sophisticated tool or technique such as the concept of artificial intelligence.

[68] has observed that ANN’s application in soil stabilisation is a relatively new development. However, the present study shows that it is fast becoming a more reliable tool in the development of predictive models for the prediction of geotechnical properties of stabilised clays. The ability of ANN to learn the functional relations and inter-variable dependencies within large collections of independent variables gives it a major advantage over traditional regression analysis involving cumbersome mathematical procedures. Some studies have utilised a relatively large database of soil properties as seen in Table 3 and will require a longer analysis time using regression analysis. These traditional procedures become even more cumbersome if a change is made in any variable in the dataset as the analysis must be repeated. However, in using ANN such changes can easily be accommodated without any repetition and this gives ANN the advantage of fitting a closer curve on the experimental data for the parameter of interest as larger dataset encompasses a wider variation of soil responses. Another advantage of ANN is the possibility of modeling multiple dependent or target variables simultaneously. ANN can search the relationship between a given dataset and 2 or more dependent variable while assigning required coefficients for the various targets. As seen in Table 3, some studies have simultaneously modeled more than one dependent variable and obtained good predictive models for the dependent variables. This again is a time saving advantage.

From Table 3, it can be seen that the performance of the model is not directly proportional to the number of neurons in the hidden layer. However, sufficient number of neurons based on model training must be allowed for optimum performance and is majorly dependent on several parameters depending on the complexity of the problem investigated. ANN has been successfully utilized in performance evaluation and post-stabilisation behaviour of treated clays with high degree of reliability. From some of the studies presented in Table 3, statistical evaluation of ANN’s performance show superior modeling capabilities than the usual regression models. The high R2 and low MSE, MAE, RMSE of ANN regression analysis are pointers to it’s advantage. From the study, it was discovered that most of the ANN models utilized were backpropagation feed-forward networks with one hidden layer. Only two studies [56, 58] used two hidden layers for the analysis. Other studies have also pointed out that most regression analysis in soil stabilisation problems are easily solved using a simple architecture of one hidden layer.

In addition, the quality and size of the database used in training contributes to the robustness of the model and enhances its performance in terms of generalisation. As seen in Table 7, many of the studies have used a relatively small dataset for the analysis and this raises concerns if the models would be reliable under practical conditions.

Table 7. Number of Dataset Utilised for Model Development

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Network Architecture | |  |  | ANN Data | | |  | | | Reference |
| No of hidden layers | No of neurons in hidden layer | ANN input variables | No. of network parameters | No. of data used | | No. of training data used | | Remarks | |  |
| 1 | 9 | 8 | 91 | 283 | 198 | | | | Sufficient | [50] |
| 1 | 11 | 8 | 111 | 72 | 50 | | | | May overfit | [52] |
| 1 | 11 | 8 | 111 | 72 | 50 | | | | May overfit | [52] |
| 1 | 11 | 8 | 111 | 72 | 50 | | | | May overfit | [52] |
| 1 | 10 | 5 | 71 | 129 | 90 | | | | Sufficient | [55] |
| 2 | 8 in each layer | 2 | 105 | 80 | 56 | | | | May overfit | [56] |
| 1 | 50 | 14 | 801 | 216 | 172 | | | | May overfit | [57] |
| 2 | 12 and 10 | 4 | 201 | 51 | 36 | | | | May overfit | [58] |
| - | - | 4 | - | 49 | 34 | | | | No Remark | [59] |
| - | - | 4 | - | 49 | 34 | | | | No Remark | [59] |
| - | - | 4 | - | 49 | 34 | | | | No Remark | [59] |
| 1 | 7 | 7 | - | 210 | - | | | | No Remark | [60] |
| 1 | - | 6 | - | - | - | | | | No Remark | [61] |
| 1 | 12 | 5 | 85 | 48 | 34 | | | | May overfit | [62] |
| 1 | 8 | 8 | 81 | 72 | 50 | | | | May overfit | [63] |
| 1 | 17 | 8 | 171 | 72 | 50 | | | | May overfit | [63] |
| 1 | 9 | 6 | 73 | 69 | 48 | | | | May overfit | [64] |
| 1 | 9 | 8 | - | 125 | - | | | | No Remark | [65] |
| 1 | 18 | 8 | 181 | 192 | 100 | | | | May overfit | [66] |
| 1 | 15 | 8 | 151 | 192 | 100 | | | | May overfit | [66] |
| 1 | 7 | 10 | 85 | 90 | 63 | | | | May overfit | [67] |
| 1 | 5 | 10 | 61 | 90 | 63 | | | | Sufficient | [67] |

- Not Reported,

One major factor may be due to the high cost of conducting geotechnical experimentation. This factor may be more significant in other areas of geotechnical engineering other than soil stabilisation. However, as shown in Table 3, satisfactory performance has also been achieved with relatively small dataset within the available training dataset.

The capabilities of ANN in soil stabilisation is wide-ranged including determination of optimum mix design for a given additive, optimise soil parameters by adjusting dependent soil properties, predict the mineralogical composition of stabilised soils, classify soils based on established criteria, model and predict the performance of stabilised soils. In summary, it is obvious that ANN can be utilised in regression analysis for the development of useful models for modeling the geo-mechanical properties of stabilised clays and indeed in soil stabilisation as a whole.

6. Conclusions

* The advantages of the artificial neural over traditional regression analysis as applied to stabilisation have been highlighted in the foregoing sections. In a typical field stabilisation project, in order to improve on the properties of expansive clays, experimental data are usually generated from several field and laboratory tests to monitor and ascertain the progress made in terms of improvement. These procedures are expensive and time consuming and may be reduced to a minimum using ANN to predict the field response of the soils. In summary, the following conclusions are made.
* Artificial neural network is reliable and can be employed in modeling various properties of stabilised clays for easy prediction of soil response while eliminating the need for extensive experimental procedures.
* Backpropagation feedforward networks are the most used models in dealing with the problem of regression analysis for stabilisation of clays.
* Artificial neural network should be developed with relatively substantial dataset to regression models with good correlation. Many of the studies in regression analysis of stabilised clays have used relatively small data set although the models have performed well. The ability of the models to generalize can be improved with larger dataset which fields a wide range of possible soil behaviour for proper training of the model.
* Shallow networks made up of one hidden layer are the most used ANN architecture in developing predic-tive models for the prediction of geotechnical characteristics of stabilised clays, and in modeling the re-sponse of stabilised expansive clays. The Levenberg-Marquardt training algorithm has been reported to be the most used among the studies reviewed.

**Author Contributions:** Conceptualization, Abbey, S. J. and Jeremiah, J. J.; methodology, Jeremiah, J. J. and Abbey, S. J.; software, Kashyap, A.; validation, Jeremiah, J. J., Abbey, S. J.; formal analysis, Jeremiah, J. J. and Abbey, S. J.; investigation, Jeremiah, J. J. and Abbey, S. J.; resources, Abbey, S. J., Kashyap, A., and Booth, C.; data curation, Jeremiah, J. J., and Abbey, S. J.; writing—original draft preparation, Jeremiah, J. J and Abbey, S. J.; writing—review and editing, Jeremiah, J. J., Abbey, S. J., and Booth, C..; visualization, Abbey, S. J., Booth, C., and Kashyap, A.; supervision, Abbey, S. J., Booth, C., and Kashyap, A.; project administration, Booth, C., Kashyap, A., and Abbey, S. J.

**Funding:** This research received no external funding.

**Acknowledgments:** The lead author is thankful to the PhD supervisory team for their enormous support during this study.

**Conflicts of Interest:** The authors declare no conflict of interest.

References

1. Gillott, J. (1986) Some clay-related problems in engineering geology in North America. Clay Minerals, Vol. 21, Issue 3, Pp 261-278. doi:10.1180/claymin.1986.021.3.02
2. Das, A. and Roy, S. (2014) Effect of Expansive Soil on Foundation and Its Remedies. International Journal of Innovative Research in Science, Engineering and Technology, Vol. 3, Special Issue 6, Pp 92-95.
3. Balasingam, M. and Farid, S. (2008): Interpretation of Geotechnical Properties of Cement Treated Soil. Washington State Department of Transportation (WSDOF) Research Report. Pp 77-122
4. S.H. Chew, A.H.M. Kamruzzaman, F.H. Lee (2004) Physicochemical and engineering behavior of cement treated clays. J. Geotech. Geoenviron. Eng., Vol. 130, Issue 7. Pp. 696-706.
5. Abbey, S. J., Eyo, E. U., Oti, J., Amakye, S.Y., and Ngambi, S. (2020) Mechanical Properties and Microstructure of Fibre-Reinforced Clay Blended with By-Product Cementitious Materials. Geosciences, 10, 241. https://doi.org/10.3390/geosciences10060241.
6. Abbey, S. J., Ngambi, S., Olubanwo, A. O., and Tetteh, F. K. (2018) Strength and Hydraulic Conductivity of Cement and By – Product Cementitious Materials Improved Soil. International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 10, Pp. 8684-8694.
7. Eyo, E.U., Ng’ambi, S., Abbey, S.J., (2020a). Incorporation of a nanotechnology-based additive in cementitious products for clay stabilisation. J. Rock Mech. Geotech. Eng. https://doi.org/10.1016/j.jrmge.2019.12.018.
8. Eyo, E.U., Ng’ambi, S., Abbey, S.J., (2020b). Performance of clay stabilized by cementitious materials and inclusion of zeolite/alkaline metals-based additive. Transp. Geotech. 23,. https://doi.org/10.1016/j.trgeo.2020.100330 100330.
9. Dongare, A. D., Kharde, R. R., and Kachare, A. D. (2012) Introduction to Artificial Neural Network. International Journal of Engineering and Innovative Technology (IJEIT), Vol. 2, Issue 1, ISSN 2277-3754. Pp. 189-193.
10. Sharma, V., Rai, S., and Anurag Dev, A. (2012) A Comprehensive Study of Artificial Neural Networks. International Journal of Advanced Research in Computer Science and Software Engineering. Vol. 2, Issue 10, ISSN: 2277 128X. Pp. 278-284.
11. Doreswamy and Vastrad, M. C. (2013) Performance Analysis Of Neural Network Models For Oxazolines And Oxazoles Derivatives Descriptor Dataset. International Journal of Information Sciences and Techniques (IJIST) Vol.3, No.6, doi: 10.5121/ijist.2013.3601.
12. Stepniewska-Dziubinska, M., Zielenkiewicz, P., and Siedlecki, P. (2018) Development and evaluation of a deep learning model for protein–ligand binding affinity prediction, Bioinformatics, Volume 34, Issue 21, Pp 3666–3674, https://doi.org/10.1093/bioinformatics/bty374
13. Lal, B. and Tripathy, S. S. (2012): Prediction of dust concentration in open cast coal mine using artificial neural network. Atmospheric Pollution Research 3, Pp 211-218.
14. Chao, Z., Guotao, M., Zhang, Y., Zhu, Y., and Hu, H. (2018) The application of artificial neural network in geotechnical engineering. International Conference on Civil and Hydraulic Engineering (IConCHE). IOP Conf. Series: Earth and Environmental Science. Vol. 189, doi :10.1088/1755-1315/189/2/022054.
15. Nihat, S. I. (2009) Estimation of swell index of fine-grained soils using regression equations and artificial neural networks. Scientific Research and Essay, Vol.4, Issue 10, Pp. 1047-1056.
16. Rani, C. S., Vaddi, P. K., and Togati, N. V. V. K. (2013) Artificial Neural Networks (ANN) for Prediction of Engineering Properties of Soils. International Journal of Innovative Technology and Exploring Engineering (IJITEE). Volume-3, Issue-1, ISSN: 2278-3075. Pp. 123-130.
17. Kalantary, F. and Kordnaeij, A. (2012) Prediction of compression index using artificial neural network. Scientific Research and Essays Vol. 7, Issue 31, DOI: 10.5897/SRE12.297, Pp. 2835-2848.
18. Singh, V., Bano, S., Yadav, A. K., and Sabih, A. (2019) Feasibility of Artificial Neural Network in Civil Engineering. International Journal of Trend in Scientific Research and Development (IJTSRD). Vol. 3, Issue 3, Pp. 724-728.
19. Winston, P. H. (1992). Artificial intelligence (3rd ed.). Addison-Wesley Pub. Co. ISBN 0-201-53377-4.
20. Zell, A. (2003). Simulation of Neural Networks. 1st ed. Addison-Wesley. ISBN 978-3-89319-554-1. OCLC 249017987
21. Cal, Y. (1995) Soil Classification by Neural Network. Advances in Engineering Software, Vol. 22, Pp. 95-91.
22. Shahin, M. A., Jaksa, M. B. and Maier, H. R. (2002): Artificial Neural Networks-Based Settlement Prediction Formula for Shallow Foundations on Granular Soils. Austrialian Geomechanics. Pp 45-52.
23. Gunaydin, O., Gokoglu, A., and Fener, M. (2010) Prediction of artificial soil’s unconfined compression strength test using statistical analyses and artificial neural networks. Advances in Engineering Software, Vol. 41, Pp. 1115–1123. doi:10.1016/j.advengsoft.2010.06.008
24. Kerh, T., Hu, Y. G., Wu, C. H. (2003) Estimation of consolidation settlement caused by groundwater drawdown using artificial neural networks. Advances in Engineering Software, Vol.34 Pp. 559–568. doi:10.1016/S0965-9978(03)00053-X
25. Reuter, U., Sultana, A., and Reischlb, D. S. (2018) A comparative study of machine learning approaches for modeling concrete failure surfaces. Advances in Engineering Software, Vol. 116 Pp. 67–79. doi.org/10.1016/j.advengsoft.2017.11.006.
26. Snyman, J. (2005). Practical Mathematical Optimization: An Introduction to Basic Optimisation Theory and Classical and New Gradient-Based Algorithms. Springer Science & Business Media. ISBN 978-0-387-24348-1.
27. Hinkelmann, K. (n. d.) Neural Networks. University of Applied Sciences Northwestern Switzerland School of Business.[online](Accessed 26/4/2021) http://didattica.cs.unicam.it/lib/exe/fetch.php?media=didattica:magistrale:kebi:ay\_1718:ke-11\_neural\_networks.pdf
28. Weisstein, E. W. (n. d.) Hyperbolic Tangent. MathWorld - A Wolfram Web Resource. [Online] (Accessed on 26/4/2021) <https://mathworld.wolfram.com/HyperbolicTangent.html>
29. Squires, G. L. (2001) Practical Physics (4 ed.). Cambridge University Press. doi:10.1017/cbo9781139164498. ISBN 978-0-521-77940-1.
30. Mote, H. and Kumar, S. R. S. (2019) Use of Artificial Neural Network for Initial Design of Steel Structures. IOP Conf. Series: Materials Science and Engineering, Vol. 660, 012064 IOP Publishing doi:10.1088/1757-899X/660/1/012064.
31. Sildir, H., Aydin, E., and Kavzoglu, T. (2020) Design of Feedforward Neural Networks in the Classification of Hyperspectral Imagery Using Superstructural Optimization. Remote Sens., Vol. 12, Issue 956, doi:10.3390/rs12060956.
32. Naidu, K., Syukri, M. A., Abu Bakar, A. H., Tan, C. K., Arof, H., Mokhlis, H. (2020) Optimized artificial neural network to improve the accuracy of estimated fault impedances and distances for underground distribution system. PLoS ONE, Vol. 15, Issue 1, https://doi.org/10.1371/journal.pone.0227494. [Online](Accessed 30/4/2021)
33. Renotte, C., Wouwer, A. V., Bogaertst, P., and M. Remy, M. (2001) Neural Network Applications in Non-linear Modelling of (Bio)chemical Processes. Measurement + Control, Volume 34. Pp. 197-201.
34. Zhang, T., Fell, F., Liu, Z., Preusker, R., Fischer, J., and He, M. (2003) Evaluating the performance of artificial neural network techniques for pigment retrieval from ocean color in Case I waters. Journal of Geophysical Research, Vol. 108, No. C9, 3286, doi:10.1029/2002jc001638.
35. Doreswamy and Vastrad, M. C. (2013) Performance Analysis Of Neural Network Models For Oxazolines And Oxazoles Derivatives Descriptor Dataset. International Journal of Information Sciences and Techniques (IJIST) Vol.3, No.6, doi: 10.5121/ijist.2013.3601.
36. Schmidhuber J., (2015). "Deep learning in neural networks:An overview". Neural Networks 61:85-117.
37. Wald, A. (1950). Statistical Decision Functions. Wiley.
38. Ikizler, S. B., Aytekin, M., Vekli, M., and Kocabas, F. (2010) Prediction of swelling pressures of expansive soils using artificial neural networks. Advances in Engineering Software, Vol. 41, Pp. 647–655. doi:10.1016/j.advengsoft.2009.12.005
39. Ciresan, D., Meier, U., Masci, J., Gambardella, L. M., Schmidhuber, J. (2011). Flexible, High Performance Convolutional Neural Networks for Image Classification. Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence. Vol. 2, Issue 2: 1237-1242. Retrieved 17 November 2013.
40. Sushama K. and Bindhu L. (2015): “ANN Based Prediction of Shear Strength of Soils from their Index Properties” International Journal of Earth Sciences and Engineering, ISSN 0974-5904, Volume 08, No. 05, pp.2195-2202.
41. Murata, N., Yoshizawa, S., and Amari, S. (1993) Learning curves, model selection and complexity of neural networks. Advances in Neural Information Processing Systems 5, S. Jose Hanson, J. D. Cowan, and C. Lee Giles, ed. San Mateo, CA: Morgan Kaufmann, Pp. 607-614.
42. Maind, S. B. and Wankar, P. (2014) Research Paper on Basic of Artificial Neural Network. International Journal on Recent and Innovation Trends in Computing and Communication. Vol. 2, Issue 1, Pp. 96-100.
43. Gareth, J. (2013) An Introduction to Statistical Learning: with Applications in R. Springer. p. 176. ISBN 978-1461471370.
44. Pathivada, S. (2005): Effects of Water-cement Ratio on Deep Mixing Treated Expansive Clay Characteristics. Arlington: The University of Texas at Arlington.
45. Lorenzo, G. and Bergado, D. (2004, October): Fundamental Parameters of Cement-Admixed Clay-New Approach. Journal of Geotechnical and Geoenvironmental Engineering, Pp. 1042-1050.
46. Turk, G., Logar, J., and Majes, B. (2001) Modelling Soil Behaviour in Uniaxial Strain Conditions by Neural Networks. Advances in Engineering Software, Vol. 32, Pp. 805–812.
47. Fan, J., Wang, D. and Qian, D. (2018): Soil-cement mixture properties and design considerations for reinforced excavation. Journal of Rock Mechanics and Geotechnical Engineering 10; 791-797
48. Shahin, M. A., Jaksa, M. B. and Maier, H. R. (2001): “Artificial Neural Networks Application in Geotechnical Engineering.” Austrialian Geomechanics. Pp49-62.
49. Gardner, M. W., and Dorling, S. R. (1998): “Artificial neural networks (The multilayer perceptron) - A review of applications in the atmospheric sciences.” Atmospheric Environment, 32(14/15), 2627-2636.
50. Mozumder, R. A. and Laskar, A. I. (2015) Prediction of unconfined compressive strength of geopolymer-stabilised clayey soils using artificial neural network. Computers and Geotechnics, vol 69, Pp 291-300.
51. BRE (1993) Low-rise buildings on shrinkable clay soils: BRE Digest, Vols. 240, 241.& 242. CRC, London.
52. Salahudeen, A. B., Sadeeq, J. A., Badamasi, A., and Onyelowe, K. C. (2020) Prediction of Unconfined Compressive Strength of Treated Expansive Clay Using Back-Propagation Artificial Neural Networks. Nigerian Journal of Engineering Faculty of Engineering Ahmadu Bello University, Samaru - Zaria, Nigeria. Vol. 27, No. 1, ISSN: 0794 – 4756. Pp. 45-58.
53. BSI (1990) BS 1377: 1990—Methods of Test for Soils for Civil Engineering Purposes. British Standards Institute, Milton Keynes.
54. BSI (1990) BS 1924: 1990—Stabilised Materials for Civil Engineering Purposes. British Standards Institute, Milton Keynes.
55. Priyadarshee, A., Chandra, S., Gupta, D., and Kumar, V. (2020) Neural Models for Unconfined Compressive Strength of Kaolin Clay Mixed with Pond Ash, Rice Husk Ash and Cement. Journal of Soft Computing in Civil Engineering. Vol. 4(2), Pp. 85-120. https://doi.org/10.22115/SCCE.2020.223774.1189 [online](Accessed 14/4/2021)
56. Ayeldeen, M., Yuki H., Masaki K. and Abdelazim N. (2016): Unconfined Compressive Strength of Compacted Disturbed Cement-Stabilised Soft Clay. Int. J. of Geosynth. and Ground Eng. 2:28
57. Ngo, H. T. T., Pham, T. A., Vu, H. L. T., Giap, L. V. (2021) Application of Artificial Intelligence to Determined Unconfined Compressive Strength of Cement-Stabilised Soil in Vietnam. Appl. Sci., Vol. 11, 1949. <https://doi.org/10.3390/app11041949>.
58. Tran, V. Q. (2021) Compressive Strength Prediction of Stabilised Dredged Sediments Using Artificial Neural Network. Advances in Civil Engineering. Vol. 2021, Article ID 6656084, <https://doi.org/10.1155/2021/6656084>.
59. Sabat, A. K. (2015): Prediction of California Bearing Ratio of a Stabilised Expansive Soil using Artificial Neural Network and Support Vector Machine. Electronic Journal of Geotechnical Engineering. Vol. 20, Bund. 3 Pp981-991
60. Rajakumar, C. and Babu, G. R. (2020) Experimental study and neural network modelling of expansive sub grade stabilised with industrial waste by-products and geogrid. Materials Today: Proceedings. https://doi.org/10.1016/j.matpr.2020.06.578 [online] (Accessed 12/4/2021)
61. Vinodhkumar, S. and Balaji, S. (2016) Artificial Neural Network Modelling and Economic Analysis of Black Cotton Soil Subgrade Stabilised with Flyash and Geotextile. International Journal of Earth Sciences and Engineering. Vol. 09, No. 01, ISSN 0974-5904, Pp.81-86.
62. Belal, A., Rahman, M. A., and Rafizul, I. M. (2016) Prediction Of California Bearing Ratio Of Stabilised Soil Using Artificial Neural Network. Proceedings of the 3rd International Conference on Civil Engineering for Sustainable Development (ICCESD 2016), KUET, Khulna, Bangladesh. ISBN: 978-984-34-0265-3, Pp. 977-987.
63. Salahudeen, A. B. and Sadeeq, J. A. (2019) California Bearing Ratio Prediction of Modified Black Clay Using Artificial Neural Networks. West Africa Built Environment Research (WABER) Conference 10th Anniversary Conference, Accra, Ghana. Proceedings Edited by Laryea, S. and Essah, E. ISBN: 978 – 9988 – 2– 6010 – 1 doi: <https://doi.org/10.33796/waberconference2019.19>
64. Vakili, A. H., Davoodi, S., Arab, A., Razip, M., and Bin Selamat, M. R. (2015). Use of Artificial Neural Network in Predicting Permeability of Dispersive Clay Treated With Lime and Pozzolan. International Journal of Scientific Research in Environmental Sciences 3(1), Pp. 0023-0037.
65. Hanandeha, S., Ardahb, A., and Abu-Farsakhc, M. (2020) Using artificial neural network and genetics algorithm to estimate the resilient modulus for stabilised subgrade and propose new empirical Formula. Transportation Geotechnics 24, 100358.
66. Alavi, A. H., Gandomi, A. H., Mollahassani, A., Heshmati, A. A., and Rashed, A.(2010) Modeling of maximum dry density and optimum moisture content of stabilised soil using artificial neural networks. J. Plant Nutr. Soil Sci., 173, Pp. 368–379.
67. Salahudeen, A. B., Ijimdiya, T. S., Eberemu, A. O., and Osinubi,j K. J. (2018) Artificial Neural Networks Prediction of Compaction Characteristics of Black Cotton Soil Stabilised with Cement Kiln Dust. J. Soft Comput. Civ. Eng. 2(3), Pp. 50–71.
68. Bahmeda, I. T., Harichanea, K., Ghricia, M., Boukhatema, B., Rebouha, R., and Gadouria, H. (2017): Prediction of geotechnical properties of clayey soils stabilised with lime using artificial neural networks (ANNs), International Journal of Geotechnical Engineering, DOI: 10.1080/19386362.2017.1329966
69. A.H.M. Kamruzzaman, S.H. Chew, F.H. Lee (2009) Structuration and destructuration behavior of cement-treated Singapore marine clay cement-treated Singapore marine clay. J. Geotech. Geoenviron. Eng., Vol. 135, Issue 4. Pp. 573-589.
70. Taskiran, T. (2010) Prediction of California bearing ratio (CBR) of fine-grained soils by AI methods. Advances in Engineering Software, Vol. 41, Pp. 886–892. doi:10.1016/j.advengsoft.2010.01.003
71. Radhakrishnan, G. Kumar, M. A. and Raju, G. V. R. P. (2016) Influence of Chloride Compounds on the Swelling and Strength Properties of Expansive Subgrade Soil. Journal of Mechanical and Civil Engineering (IOSR-JMCE), Vol. 13, Issue 1, Ver. IV. Pp. 57-60.
72. Ikeagwuani, C. C. and Nwonu, C. D. (2019) Emerging trends in expansive soil stabilisation: A review. Journal of Rock Mechanics and Geotechnical Engineering. vol. 11, Pp.423-440.
73. Eyo, E.U., Ngambi, S., and Abbey, S.J. (2018) Investigative study of behaviour of treated expansive soil using empirical correlations. In: International Foundation Congress and Equipment Expo 5–10 March. Orlando, Florida, pp. 373–384.
74. Eyo, E. U., Abbey, S. J., Ngambi, S., Ganjian, E., and Coakley, E. (2021) Incorporation of a nanotechnology-based product in cementitious binders for sustainable mitigation of sulphate-induced heaving of stabilised soils. Engineering Science and Technology, an International Journal 24, Pp. 436–448. https://doi.org/10.1016/j.jestch.2020.09.002