3rd International Civil Engineering Conference (ICEC 2024)
Department of Civil Engineering
Federal University of Technology Minna

Development of Artificial Neural Network Models for Predicting Strength Properties of Tropical Clay Stabilized with Calcium Carbide Residue and Zeolite – A Review

*Mohammed, I. K., [1]Alhassan, M., [2]Alhaji, M. M., [3]Adejumo, T. E. and [4]Yusuf, A.

[1] Department of Civil Engineering Federal University of Technology, Minna
[2]Department of Civil Engineering Federal University of Technology, Minna
[3]Department of Civil Engineering Federal University of Technology, Minna
[4]Department of Civil Engineering Federal University of Technology, Minna
*Corresponding author email: kankoibraheem@gmail.com

ABSTRACT

The paper presents a literature review on the development of models for predicting strength properties of tropical clay stabilize with Calcium Carbide Residue **(**CCR) and zeolite. Application of Artificial Neural Networks (ANNs) in geotechnical analysis of tropical clay stabilised with CCR and zeolite, have been evaluated. Chemical treatment of expansive clays involves development of optimum binder mix proportions or 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 often involve large datasets and tedious mathematical procedures to correlate the variables and develop required models using traditional regression analysis. The findings from this study shows that ANNs is becoming well known in dealing with the problem of mathematical modelling involving nonlinear functions due to their robust data analysis and correlation capabilities, which has enabled them to be successfully applied, and with high performance, to the stabilisation process of clays. The study also shows that the supervised ANN model is well adapted to dealing with stabilisation of clays with high performance as indicated by high R^2 and low Mean Average Error (MAE), Root Mean Square Error (RMSE), and Mean Square Error (MSE) values. The Levenberg–Marquardt algorithm is effective in shortening the convergence time during model training.

Keywords: *Artificial neural networks; Calcium Carbide Residue; clays; predictive models; Stabilisation; Zeolite*

1.0 INTRODUCTION

The increasing population of the world, especially in the developing nations, has led to increasing demand for roadways, railways, housing facilities and other infrastructures. Soil with higher stability is required to bear the weight of these structures. Generally speaking, the stability of any civil engineering structure directly or indirectly depends on the stability of the bearing soil (Akuto, 2021; Suleiman *et al.* 2021; Balarabe and Mary, 2015). While some soils have the required stability to bear the weight of these structures, some other soils are deficient in this regard.

Deficient soils are regarded as soils which do not meet some or all the criteria required for their satisfactory performance in geotechnical applications. These could either be for base courses in pavement construction, embankment for dams, subsoil base, clay liners for containment of leachates and backfill for retaining walls (Alhassan and Mustapha, 2015). In the tropical region, these soils could be lateritic soils, tropical black clay, collapsible soils or any other tropical soil type (Alhassan and Mustapha, 2015).

Tropical Clay Soil (TCS) are expansive soils that principally occur in arid and semi-arid regions of the tropical zone, marked by dry and wet seasons, with low rainfall, poor drainage and exceedingly great heat (Nelson and Miller, 1992; Eberemu *et al,* 2012). Because of their unconventional behavior, they are problematic in geotechnical engineering applications because they exhibit large volume changes with respect to variation of moisture content (Eberemu *et al.,* 2012).

Tropical clay soil presents various challenges to engineers all over the world due to their characteristic of severe loss of strength and swelling with respect to changes in moisture regime. As a result, structures and highways constructed on them are subjected to severe deformations and frequent repairs, leading to high cost of maintenance. Various efforts have been made to stabilize tropical clay with cement, lime, admixtures and waste products, to make them meet the requirements for construction works (Balogun, 1991; Osinubi *et al.*; 2009; Suleiman *et al*, 2021).

Soil stabilization refers to the improvement of the bearing power of a deficient soil using compaction, proportioning and the addition of suitable stabilizers or appropriate admixtures (Alhassan and Mustapha, 2015). Soil stabilization methods include chemical, mechanical and physico-chemical methods to improve soil properties such as strength, stability and reduce swelling. Stabilization methods can generally be categorized as physical and chemical methods (Neeraj and Ahirwar, 2014). Physical stabilization involves improving soil structure through mechanical methods, while chemical stabilization uses additives like lime or cement to enhance soil strength and durability. Together, these methods enhance the resistance to erosion and increase load-bearing capacity of the soil. Locally available materials have also been experimented as soil stabilizing additives.

For several years, researchers have recognized the use of locally available materials which are available from industrial and agricultural wastes to improve the properties of expansive soils. This is aimed at reducing stabilization costs, related to conventional stabilizing agents such as cement and lime, as well as reducing the problem of CO_2 emission that is related to cement manufacturing process (Balarabe and Mary, 2015).

Calcium Carbide Residue (CCR) is a by-product from the production of acetylene gas (C_2H_2) used in oxyacetylene welding (Johnson *et al.;* 2023). It consists mainly of lime, caustic solid substances, and is white in appearance when pure.

Zeolites as a pozzolan are aqueous aluminum silicate containing alkali and alkaline earth elements. Their structure is made up of a framework of SiO_4 and AlO_4 tetrahedrons linked to each other's corners by sharing oxygen atoms (Capucto *et al.;* 2008). Zeolites have been recognized for 200 years, but only during the middle of the twentieth century did they attract the attention of scientists and engineers, who demonstrated their technological importance in several fields (Hahwu *et al*.; 2008).

1.1. Artificial Neural Network

Artificial Neural Networks (ANNs), or more simply neural networks, are new systems and computational methods for machine learning, knowledge demonstration, and finally the application of knowledge gained to maximize the responses of complex systems (Grekousis; 2019). ANN is a data processing model based on the way biological nervous systems, such as the brain, process data. They are focused on the neuronal structure of the mamalian cerebral cortex, but data is much smaller scale. Many artificial intelligence experts believe that artificial neural networks are the best and perhaps the only hope for designing an intelligent machine.

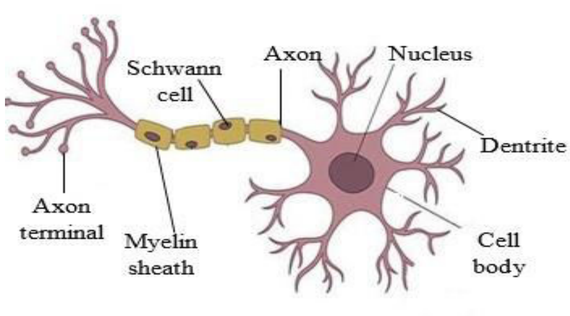
ANN, as a branch of artificial intelligence, is simply an automated optimisation system capable of learning the relationship and inter-dependencies between multiple input variables of a given system and modelling such relations (trends and patterns) in the form of mathematical functions for easy prediction (Stepniewska-Dziubinska *et al.;* 2018). 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 (Lal and Tripathy, 2012). 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 a number of hidden layers (shallow or deep networks), based on training algorithm (Back-Propagation Networks, Hopfield Networks, Self-Organizing Map Networks) *(*Chao, *et al*., 2018). This paper simplifies the underlying concepts of back propagation ANN models and explores its applicability in modelling 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

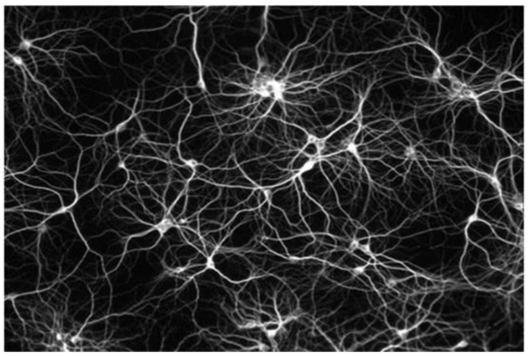
1.2.1 Neurons and Edges

ANN building blocks are 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 that 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 (Grekousis, 2019). This output of the weighted summation then becomes the input for the activation function- a linear or non-linear function (Nihat, 2009).

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 (Winston, 1992). Inputs parameters with greater importance are assigned a higher weight than those with lower importance (Zell, 2003). For instance, in a soil classification problem, the weights will be dependent on the contribution of the features in determining the class of soil (Cal, 1995). In a typical perceptron, as in Figure 2, the connection weights can be represented as Wj, which describes the importance of the connection.



(a)



(**b**)

Figure 1: (**a**, **b**) Biological Neural Network (Sharma, *et al.,* 2012)

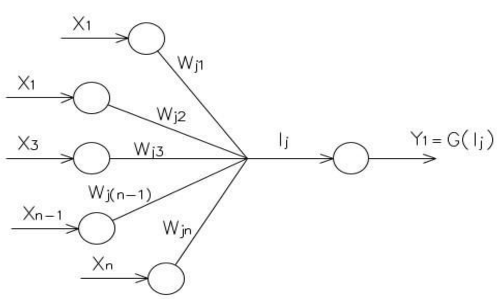


Figure 2: Typical Neuron showing associated edge weights

2.0 Materials and Method

Input data sets will be supplied to the neurons in the input layer and will be treated each with coefficients and constants known as weights and bias respectively to obtain a sum of weighted inputs and bias as given in Equation I

$$\rho = \sum_{i=1}^{n} xi.wij + b1 \qquad (1)$$

where (1)
P= weighted sum of inputand bias

Xi = input data i

Wij = weight associated with the input hidden layer and Equation 3 will be treated with a tangent sigmoid activation function given in Equation 2 to obtain the first layer output given in Equation 3

$$\beta = \frac{2}{1+e^{-2x}} - 1 \qquad 2$$

Where

B = activation function

$$\boldsymbol{\phi} = \beta \sum_{i=1}^{n} xi.wij + b1 \qquad 3$$

where;

Φ is the hidden layer output

The first (hidden) layer output *Φ* be supplied to the neuron in the output layer and will be further processed with new weights and bias. The weighted sum will be further treated with a linear activation function given in Equation 4 to obtain the overall model output given in Equation 5 termed as the case 1 model.

$$f(x) = x \qquad 4$$

$$\boldsymbol{\xi} = \mu\left(\sum_{i=1}^{n} wij\,.\boldsymbol{\phi} + b2\right) \qquad 5$$

Where;

$\boldsymbol{\xi}$ = Output of the entire case 1 ANN model
μ = Input activation function $f(x)$
wij = weight associated with output layer and
b2 = bias associated with the output layer
The models will be trained using back propagation algorithm. The sequence requires updating the connection weights and biases according to the learning capacity of the network. The iterative process last up until the network is able to identify the smallest error between the actual experimental result and the model output based on the parameters given.

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) (Naidu, et al., 2020). 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 (R^2), 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 the lowest error and highest R^2 (Eyo and Abbey, 2021). 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 a 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 the 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 an insignificant change in the output of the cost function (Wald, 1950). Based on the performance of various models with different hyperparameters, the best model topography is then selected.

2.2. Feed forward 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 the 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) (Ikizler *et al.,* 2009). 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. Additionally, 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 a lesser number of neurons in the succeeding layer even though one may be tempted to believe that a larger number of neurons will always result in a 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, such as the Deep Neural Network, the number of neurons has a lesser effect on the overall performance of the network than the number of layers and DNNs with more hidden layers have been 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 the amount of data (Sushama and Bindhu, 2015). Additionally, even though it is expected that a greater number of experimental datasets used in training the model will improve its performance, recent studies have shown that it might be advisable to use fewer experimental datasets of high quality than a large amount of experimental data which may be prone to errors (Stepniewska-Dziubinska *et al*., 2018). Moreover, 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 is a 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. Conversely, in the feedforward network, neurons in each layer are only connected to neurons in other layers as presented in Figure 4.

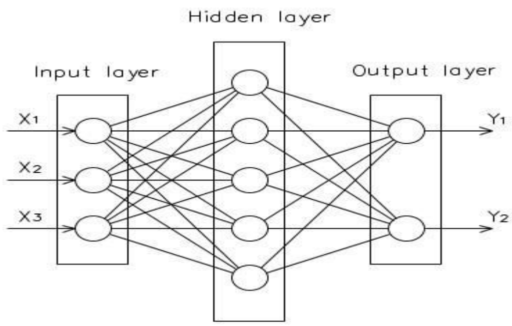
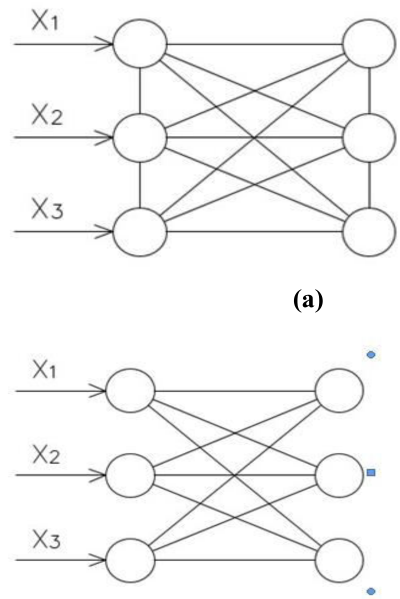


Figure 3: A Typical ANN Architecture Showing Neurons and Layers.



(a)

(b)

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

2.3 Data Preprocessing

The input dataset used to develop ANN models for geotechnical engineering problems comprises of various ranges 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, y_1 (x_{11}, x_{21}, x_{31}, . . . , x_{n1}) and y_2 (x_{12}, x_{22}, x_{32}, . . . , x_{n2}) in a data space can be expressed as Equation (6). Therefore, in order 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 the max-min normalization method. For standardization, the feature X_i is expressed in its standardized form as in Equations (6) and (7) below.

$$d_e = \sqrt{(x_{12} - x_{11})^2 + (x_{22} - x_{21})^2 + \ldots + (x_{n2} - x_{n1})^2} \quad 6$$

$$X_s = \frac{X_i - \mu}{\sigma}$$

$$X_{norm} = \frac{X_i - X_{min}}{} \quad 7$$

In Equations (1) and (2), X_s is the standardised value of X_i, μ is the mean and σ 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 X_{norm} is the normalised value and X_{min} and X_{max} are the minimum and maximum values, respectively.

3.0 Selecting Design Parameters for ANN in Soil Stabilisation

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 few a 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 becomes excess, this leads to overfitting. In this case, the model is training on less complex data than it is designed to analyse. 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. Additionally, an excessively high number of neurons in the hidden layer can extend training time and lead to poor training even with a sizable database for training. The goal, therefore, is to find a balance. However, this is not a straightforward 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 that 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 analyses 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 networks applied to modelling geomechanical properties of stabilised Clay is done under supervised conditions. In this type of training, the supplied data are partitioned, and a part (training dataset) is 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 (Murata*, et al.,* 1993). 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 the number of hidden layers, modify the activation function or even the training algorithm (Maind and Wankar, 2014). The model validation utilizes the successive trial of the trained model on the validation dataset (Gareth, 2013). 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 are continuously partitioned into two as in cross-validation. The dataset is 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 (R^2), the mean absolute error (MAE), the root mean square error (RMSE), the mean square error (MSE) and others. The R^2, MAE, RMSE and MSE expressions are defined in Equations (8), (9), (10) and (11), respectively.

$$R^2 = \left(\frac{\sum_1^n \left(y_{\exp(i)} - \overline{y}_{exp}\right)\left(y_{pre(i)} - \overline{y}_{pre}\right)}{\sum_1^n \left(y_{\exp(i)} - \overline{y}_{exp}\right)^2 \sum_1^n \left(y_{pre(i)} - \overline{y}_{pre}\right)^2} \right)^2 \quad 8$$

$$MAE = \frac{I}{n}\left(y_{\exp(i)} - y_{pre(i)}\right) \quad 9$$

$$RMSE = \sqrt{\frac{1}{n} \times \sum_1^n \left(y_{\exp(i)} - y_{pre(i)}\right)^2} \quad 10$$

$$MSE = \frac{1}{n} \times \sum_1^n \left(y_{\exp(i)} - y_{pre(i)}\right)^2 \quad 11$$

For which $y_{\exp(i)}$ and $y_{pre(i)}$ are experimental and predicted values of a given dependent variable, while y_{pre} and y_{exp} 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 may be estimated as the number of edges or connections, including biased neurons. 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 into the 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 a relatively low sample size. The underlying idea is to make available sufficient example data with which the network is trained and evaluated. Too little training data will result in a higher variance of the network parameters. Additionally, two few testing data will create higher variance during the evaluation of the performance of the model.

4.0 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 modelling the post-stabilisation behaviour of treated clays through the prediction of various engineering parameters for the purpose of design and construction. These methods rely 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 (Lorenzo and Bergado, 2004). 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 (Fan *et al.,* 2018). 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 our variables. An example of such simplified mathematical models is the multiple linear regression model, which, for a given set of features x_{1n}, x_{2n}, x_{3n}, . . . , x_{mn}, can be related to its dependent variable as stated in Equation (12) below (Gunaydin *et al.,* 2010).

$$y= \beta0+\beta1x_1n+\beta2x_2n+\beta3x_3n+ . . . \beta1xmn +\varepsilon \quad (12)$$

where β_0, β_1, β_3, . . . β_n are the coefficients, and ε is the error.

ANN application has shown good results in terms of the 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 (Gardner, 1998). A review of studies modelling 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 modelling the UCS of geopolymer stabilised clays, by Mozumder *et al.,* (2015). 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 British 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, the molarity of alkaline activator used (M), the ratio of the activator to the binder, the ratio of sodium to aluminium in the activator–binder mixture (Na/Al), the ratio of silicon to aluminium in the activator–binder mixture (Si/Al), and 28 days UCS, were considered. A multi-layered perceptron ANN model was` chosen with one 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 one hidden layer with nine neurons. The ANN showed a better ability to learn the relationship of the data set, as shown in Figure 5.

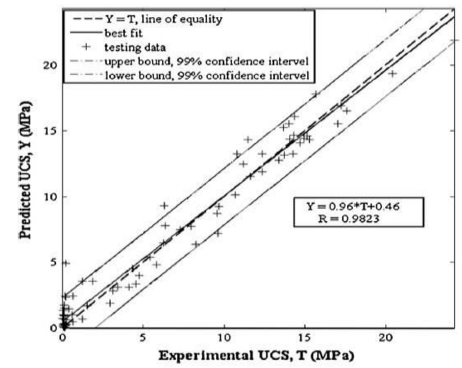


Figure 5: Experimental versus ANN-Predicted UCS values (Mozumder*, et al.,* 2015)

From Figure 5, it is obvious that almost all data points are within 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. It 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 the Performance of ANN and MVR models

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

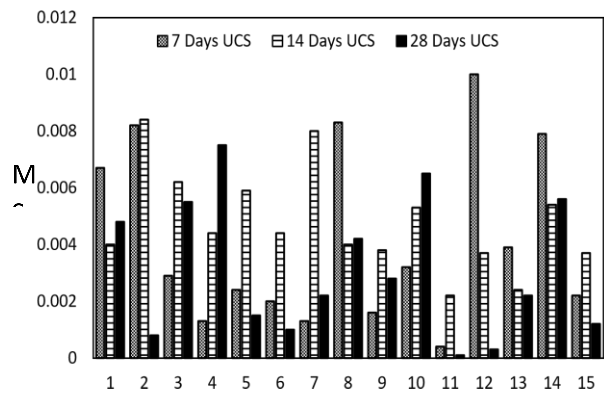
(Mozumder et al., 2015)

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 *(BRE. 1993)*

0 Modified Plasticity Index I_P	**Volume Change Potential (VCP)**
>60	Very high
40–60	High
20–40	Medium
<20	Low

In another study, Salahudeen *et al.* (2020) 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 were determined to be 48.2, 27.2 and 21%, respectively, with free swell of 80%. The natural and stabilised clay samples were subjected to compaction and UCS tests in accordance with BSI (BS 1377, 1990). Three ANN models were developed with a total of eight input variables, namely specific gravity (Gs), linear shrinkage (L_S), coefficient of uniformity (C_U), coefficient of gradation (C_C), LL, PL, Maximum Dry Density (MDD) and Optimum Moisture Content (OMC) for three UCS outputs (7, 14 and 28 days). A total of 72 sample data were utilised in model development. The ANN model topography consisted of one input layer with eight 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.



Number of Hidden Layers

Figure 6: The selection of optimum number of neurons in hidden layer (Salahudeen, *et al.,* 2020)

The optimum model was found to be of nine neurons in the hidden layer based on lowest MSE. The authors (Salahudeen *et al.,* 2020) suggest 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.

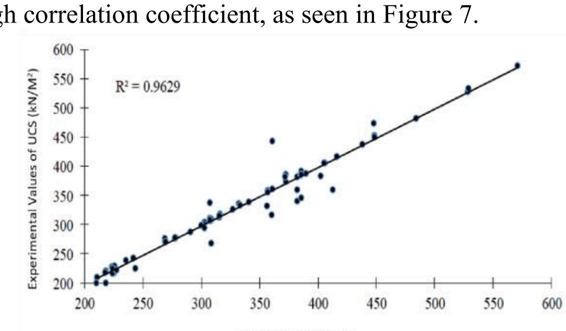


Figure 7: Observed Versus Predicted UCS (Salahudeen *et al.,* 2020)

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

Priyadarashee *et al.* (2020) conducted an experimental investigation on the suitability and performance of kaolin clay stabilised with Fly Ash (PFA), Rice Husk Ash (RHA) and cement. The improvements in the soil were ascertained by considering increase in the UCS. The LL, PL, and PI of the unstabilised kaolin clay were determined as 43.3, 19.5 and 23.8%. The input variables considered for the 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.

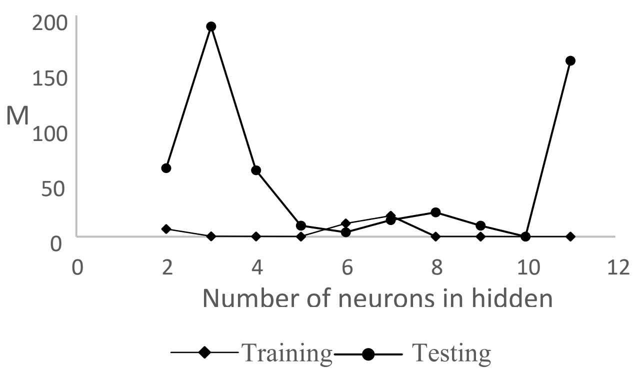


Figure 8: The selection of the optimum number of neurons in hidden layer (Ayeldeen *et al.,* 2016)

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 good correlation between the datasets. The predictive model developed from ANN is given in Equation (14), where UCS_m is the dependent variable to be estimated and represents 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.

Table 3, comparison of ANN performance with multivariable regression

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

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.

$$\frac{UCS_m}{UCS_{C7}} = -1.003 + 1.568\left(\frac{Clay}{100}\right)^2 + 2.477\left(\frac{PFA}{100}\right)^{0.5} + 2.287\left(\frac{RHA}{100}\right)^{0.5} + 6.731\left(\frac{Cement}{100}\right)^{0.5} + 0.017C \quad 13$$

The UCS of cement stabilised clays was investigated by Ngo *et al*. (2021) in order to find a relationship between the several key variables. Different types of cement were utilised to study the effects of cement type on the UCS.
The soil under investigation was collected at various depths below the 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 combinations of fourteen input variables, namely soil type, moisture content (MC), bulk density (We), the mass of cement (CM), sample diameter (DI), the length of the sample (L), the cross-sectional area of the sample (CA), the volume of the sample (SV), the depth of sample collection (D), the mass of the sample (MS), sample density (DS), curing condition (CD), curing time and cement type (CT). ANN training was done using the 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 a 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.

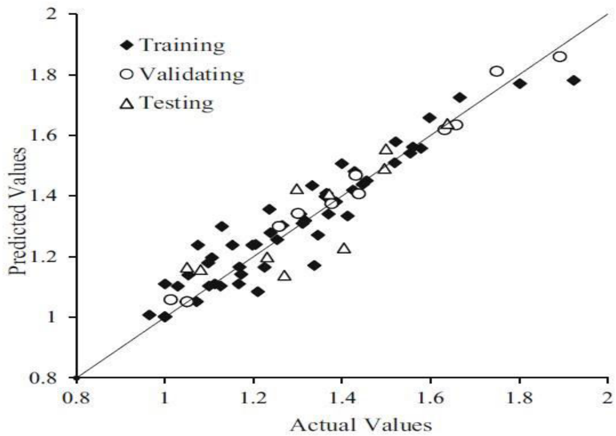


Figure 9: Predicted and experimental observations *(Ngo et al., 2021)*

To alleviate the difficulties associated with the need for continuous experimental determination of UCS, Sabat (2015) employed machine learning in analysing and modelling the performance of stabilised dredged sediments. A total of 51 experimental datasets were collated from existing literature for the 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 two 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 two 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 a lot of original research on stabilisation, there is usually limited experimental data. A cross-validation approach may be combined for a relatively lower sample size. The statistical evaluation of the performance of the model in training and testing is presented at the end of Section 4.

5.0 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 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.

An artificial neural network is reliable and can be employed in modelling 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.

An artificial neural network should be developed with a relatively substantial dataset to regression models with good correlation. Many of the studies in regression analysis of stabilised clays have used relatively small data sets, although the models have performed well. The ability of the models to generalize can be improved with a 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 predictive models for the prediction of geotechnical characteristics of stabilised clays and in modelling the response of stabilised expansive clays. The Levenberg–Marquardt training algorithm has been reported to be the most used among the studies reviewed.

6.0 Reference

Akuto*, T. T. (2021).* Strength assessment of calcium carbide residue stabilized clay admixed with zeolite. Federal University of Technology, Minna Repository. Retrieved from https://repository.futminna.edu.ng

Alhassan, M. & Mustapha, A. M. (2015). Improvement of deficient soils in Nigeria using bagasse ash - A review, *proceeding of 17th International Conference on Civil and Building Engineering*, 1040-1049.

Ayeldeen, M.; Yuki, H.; Masaki, K.; Abdelazim, N. (**2016)** Unconfined Compressive Strength of Compacted Disturbed Cement-Stabilised Soft Clay. *Int. J. Geosynth. Ground Eng. 2*, 1–10.

Balarabe, W. I., & Mary, R. S. (2015). Soil Stabilization Using Calcium CarbideResidue and Coconut Shell Ash. *Journal of Basic and Alied EngineeringResearch,* 2 (12), 1039-1044.

Balogun, L. A. (1991). Effect of sand and salt additives on some geotechnical properties of lime stabilized black cotton soil. *Niger,* 26(4), 15-24.

BRE. (1993). *Low-Rise Buildings on Shrinkable Clay Soils: BRE Digest*; CRC: London, UK, 1993; Volumes 240–242.

BSI. *BS 1377: (1990) —Methods of Test. for Soils for Civil. Engineering Purposes*; British Standards Institute: Milton Keynes, UK, 1990.

Cal, Y. (1995). Soil classification by neural network. *Adv. Eng. Softw. 22*, 95–97.

Capucto, D., Liguori, B., & Colella, C. (2008). Some Advances in Understanding the Pozzolanic Activity of Zeolites: The Effect of Zeolite Structure. *Cement and Concrete Composite,* 30(5), 455-462.

Chao, Z.; Ma, G.; Zhang, Y.; Zhu, Y.; Hu, H. (**2018)** The application of artificial neural network in geotechnical engineering. International Conference on Civil and Hydraulic Engineering (IConCHE). *IOP Conf. Ser. Earth Environ. Sci. 189*, 22054.

Eberemu, A. O, Amadi, A. A. & Lawal, M. (2012). The geotechnical properties of black cotton soil treated with glass cullet. *Nigeria Journal of Engineering*, 2, 23-30.

Eyo, E.U.; Abbey, S.J. (2021) Machine learning regression and classification algorithms utilized for strength prediction of OPC/by- product materials improved soils. *Constr. Build. Mater. 284*.

Fan, J.; Wang, D.; Qian, D. (2018) Soil-cement mixture properties and design considerations for reinforced excavation. *J. Rock Mech. Geotech. Eng. 10*, 791–797.

Gardner, M.W.; (1998) Dorling, S.R. Artificial neural networks (The multilayer perceptron)—A review of applications in the atmospheric sciences. *Atmos. Environ. 32*, 2627–2636.

Gareth, J. (2013). *An Introduction to Statistical Learning: With Applications in R*; Springer Science and Business Media, New York, NY, USA; p. 176. ISBN 978-1461471370.

Gunaydin, O.; Gokoglu, A.; Fener, M. (2010). Prediction of artificial soil's unconfined compression strength test using statistical analyses and artificial neural networks. *Adv. Eng. Softw. 41*, 1115–1123

Ikizler, S.B.; Aytekin, M.; Vekli, M.; Kocabas, F. (2009). Prediction of swelling pressures of expansive soils using artificial neural networks. *Adv. Eng. Softw, 41*, 647–655.

Lal, B.; Tripathy, S.S. (2012). Prediction of dust concentration in open cast coal mine using artificial neural network. *Atmos. Pollut. Res. 3*, 211–218.

Lorenzo, G.A.; Bergado, D.T. *(*2004). Fundamental Parameters of Cement Admixed Clay—New Approach. *J. Geotech. Geoenvironmental Eng. 130*, 1042–1050.

Maind, S.B.; Wankar, P. (2014). Research Paper on Basic of Artificial Neural Network. *Int. J. Recent Innov. Trends Comput. Commun. 2*, 96–100.

Maind, S.B.; Wankar, P. (2014). Research Paper on Basic of Artificial Neural Network. *Int. J. Recent Innov. Trends Comput. Commun. 2*, 96–100.

Mozumder, R.A.; Laskar, A.I. (2015). Prediction of unconfined compressive strength of geopolymer-stabilised clayey soils using artificial neural network. *Comput. Geotech. 69*, 291–300.

Murata, N.; Yoshizawa, S.; Amari, S. (1993). Learning curves, model selection and complexity of neural networks. In *Advances in Neural Information Processing Systems 5*; Jose Hanson, S., Cowan, J.D., Lee Giles, C., Eds.; Morgan Kaufmann: San Mateo, CA, USA; pp. 607–614.

Naidu, K.; Ali, M.S.; 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, 15*, e0227494.

Neeraj, K. & Ahirwar, S. K. (2014). Performance analysis of black cotton Soil treated with calcium carbide Residue and stone dust. *International Journal of Engineering Research And Science and Technology,* 3 (4), 202-209.

Nelson, J. D. & Miller, D. J. (1992). *Expansive soils: Problems and practice in foundation and pavement engineering*. John Wiley and Sons. New York.

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 Stabilized Soil in Vietnam. *Appl. Sci. 11*, 1949.

Nihat, S.I. (2009). Estimation of swell index of fine-grained soils using regression equations and artificial neural networks. *Sci. Res. Essay*, *4*, 1047–1056.

Osinubi, K. J. Bafyau, V., & Eberemu, A. O. (2009). Bagasse ash stabilization of lateritic soil. In *Appropriate Technologies for Environmental Production in the developing World*,271-280.Springer, Dordrecht.

Salahudeen, A.B.; Sadeeq, J.A.; Badamasi, A.; Onyelowe, K. C. (2020). Prediction of Unconfined Compressive Strength of Treated Expansive Clay Using Back-Propagation Artificial Neural Networks. *Niger. J. Eng. Fac. Eng. Ahmadu Bello Univ. 27*, 45–58, ISSN 0794–4756.

Sharma, V.; Rai, S.; Anurag Dev, A. (2012). A Comprehensive Study of Artificial Neural Networks. *Int. J. Adv. Res. Comput. Sci. Softw. Eng. 2*, 278–284, ISSN 2277 128X

Snyman, J. (2005). Practical Mathematical Optimization: an Introduction to Basic Optimisation Theory and Classical and New Gradient-Based Algorithms; *Springer Science & Business Media*: New York, NY, USA; ISBN 978-0-387-24348-1.

Stepniewska-Dziubinska, M.M.; Zielenkiewicz, P.; Siedlecki, P. (2018). Development and evaluation of a deep learning model for protein–ligand binding affinity prediction. *Bioinformatics*, *34*, 3666–3674.

Sushama, K.; Bindhu, L. (2015). ANN Based Prediction of Shear Strength of Soils from their Index Properties. *Int. J. Earth Sci. Eng. 8*, 2195–2202, ISSN 0974 5904.

Wald, A. (1950). *Statistical Decision Functions*; John Wiley & Sons: New York, NY, USA.

Winston, P.H. (1992). *Artificial Intelligence*, 3rd ed.; Addison-Wesley Pub. Co.:Boston, MA, USA; ISBN 0-201-53377-4.

Zell, A. (2003). *Simulation of Neural Networks*, 1st ed.; Addison-Wesley Pub. Co.: Boston, MA, USA; ISBN 978-3-