

# Comparative Soft Computing Approaches for Evaluating the Impact of Shearing Parameters on the Bearing Capacity of Shallow Foundations

Awaisu Shafiu Ibrahim<sup>1,5</sup>, Auwal Alhassan Musa<sup>2</sup>, Usman U. Aliyu<sup>1,3\*</sup>, Abubakar Usman Khalil<sup>4</sup>, Ahmad Idris<sup>5</sup>

- <sup>1</sup>Department of Civil Engineering, Federal University, Dutsin-ma, Katsina, Nigeria.
- <sup>2</sup>Mechanics and Civil Engineering Laboratory, University of Montpellier, France.
- <sup>3</sup>School of Engineering and Technology, Department of Civil Engineering, Sharda University, Knowledge Park III, Greater Noida, Uttar Pradesh 201310, India.
- <sup>4</sup>Department of Civil Engineering, Kano State Polytechnics, Kano, Nigeria.
- <sup>5</sup>Department of Civil Engineering, Bayero, University, Kano, Nigeria
- $*Corresponding author: \underline{uu287528@gmail.com}$

#### **Abstract**

The structural stability of foundations for most civil engineering infrastructures depends on the resistance properties of foundation soil to neutralize the possibilities of failure by shear. These resistance capacities are based on shear strength parameters. In contrast, the complexity of the analysis and its timeconsuming nature pose challenges when analyzing foundation soil. Thus, developing models for computing the bearing capacity of soil foundations from existing laboratory data and field results will help minimize the predicaments and shortcomings associated with analyzing soil foundations. Therefore, this paper correlates observed data of shear strength parameters with predicted values obtained from models developed using Artificial Neural Network (ANN) and Multiple Linear Regression (MLR). The performance indicators obtained from ANN for cohesion (R2, MSE, RMSE, R) are R2 = 0.9749, MSE = 17.0160, RMSE = 4.1250, R= 0.9874, and predicted values of angle of internal friction also show good correlations from the history of the indicators  $R^2 = 0.9499$ , MSE = 26.5235, RMSE = 5.1500, R = 0.9746. Likewise, MLR shows a promising correlation between dependent and independent parameters; the R<sup>2</sup> of cohesion and angle of internal friction were obtained to be 0.8789 and 0.91415. Other indicators are MSE = (82.3801 and 44.8042), RMSE = (9.0763 and 6.6935), R = (0.9374 and 0.9561). The models developed between these parameters present suitable and reliable outputs for determining the bearing capacity of foundation soil using shear strength parameters. In addition, higher precision was observed from the result given by the ANN.

Keywords: Bearing Capacity, Foundation, Shearing Parameters, Machine Learning, Prediction

#### 1. Introduction

The ability to resist failure by sliding within the inner portion of the mass of the soil is essentially among the geotechnical properties possessed by a soil to serve as foundation material. The shear resistance possessed by the foundation soil against failure through the surface of slippage due to imposed loading will influence the stability of any structure resting upon it [1]. Therefore, a deeper comprehension of the importance of shear strength properties of soil is highly essential for the analysis of the bearing capacity of soil foundation [2]. Furthermore, it is crucial for the stability analysis of slopes [3], the design of retaining walls and embankments for dams [4], tunneling excavation and linings [5], as well as for

providing sufficient resistance, traction, and tillage tools in agricultural purposes [6]. Mohr–Coulomb's theory offers a generalized approach to determining the shear strength of soils. The theory claims that the shear strength parameters of soils (i.e., angle of internal friction and cohesion), which constitute the applied stress, influence the shear strength of soil proportionally [7], [8]. The slope (angle of internal friction), which is usually expressed in degrees, and the intercept (cohesion), described in N/m2, are the main determinants of the tangent to the Mohr–Coulomb failure envelopes [9], [10]. Field and laboratory experiments are both applicable in obtaining shear strength parameters of soil [11], [12]. However, the most employed laboratory methods for determining shear strength parameters are the triaxial compression test and the direct shear test. According to Ersoy et al. [8] and Aridsson and Keller, [13], conventional approaches to obtaining shear strength properties in the field and laboratory require a substantial investment of time, funds, and labor. Therefore, to minimize the cumbersome attributes associated with both approaches, it is of utmost importance to develop and employ soft computation models, which, in the long run, will replicate traditional methods with high accuracy.

On the other hand, statistical analysis and tools are crucial for predicting the geotechnical properties of soils. Consequently, these tools for modeling predictions of specific parameters in geotechnical engineering have been utilized to model various engineering properties of soils, particularly in situations constrained by financial limitations, the unavailability of test equipment, and limited time for design purposes [7]. Correlations and empirical equations were found to be reliable tools for determining the essentials in preliminary geotechnical studies, where the challenges above were encountered. This is particularly alarming in some developing countries, where state-of-the-art testing equipment is currently insufficient, combined with the scarcity of adequately equipped and well-trained personnel needed for its operation. Over the last two decades, a broad and abrupt paradigm shift has occurred in research, spatially extending sparse and expensive soil measurements that previously focused on utilizing secondary information for the development of prediction models [12]. Empirical approaches are widely used in geotechnical engineering practice as tools for determining the engineering properties of soils. Aridsson and Keller, [13] and Musa and Dulawat, [14] claimed that distinct correlations exist between the index properties obtained from simple routine testing and the strength properties of soils, among others. For determining the shearing strength parameters of soil, numerous empirical approaches have been developed over the years. Some of these models developed are: [15] for clay and silt embankments, [16] for Barind soils, [17] for peaty soils, and [18] for mudrock. In addition, researchers focused on developing models based on sustained hypotheses before estimating the actual model parameters. This is based on their assumption about how the input and output variables of the models are related [7], [8], [19]–[22].

However, these presumed models may not have the essential attributes needed. This gives rise to the advent of soft computing methods in developing robust models, where the data trend is allowed to evolve, and an appropriate model is becoming widely accepted [23], [24]. Little contribution has been made concerning the use of modelling tools for the prediction of shear strength parameters for shallow foundations. Instead, most of the available research focused on the effect of these soil parameters on construction activities. Studies by [25] and [26] have significantly contributed to establishing a practical correlation between soil geotechnical properties through in-situ testing, resulting in a notable decrease in construction project expenses. Moreover, their innovative approach includes the development of an Artificial Neural Network ANN model, which outperforms traditional regression methods in accurately predicting parameters such as the angle of friction and cohesion. Despite the general correlation between high plasticity index and lower effective friction angle in clays, some may exhibit higher effective friction angles, which contradicts this relationship [13]. While certain clays exhibit significant undrained shear strength that escalates with depth, they inherently possess a liquidity index exceeding 100%, meaning their natural water content surpasses the liquid limit [27]. This also contradicts the typical associations that link undrained shear strength to liquid limit, as they assume that clays of this type have minimal undrained strength [27], [28]. Clay index properties are commonly employed in traditional methods for evaluating geotechnical parameters.

For instance, [29] utilized the liquidity index for assessing undrained shear strength. This provides an avenue for employing the plasticity index to approximate the effective angle of friction, which nowadays is becoming a common practice in geotechnical engineering. However, these correlations exhibit significant limitations, which a novel method introduced by [30] bridges by establishing a connection between shear wave velocity and small-strain shear modulus, as well as index properties of saturated soft and firm clays. The method offers reliable predictions of the strength and stiffness of clays based on their index properties. At the same time, research by [31] indicates a strong correlation between undrained shear strength and liquid limit, plastic limit, bulk density, dry density, natural moisture content, and plasticity index. However, no significant correlation was found between specific gravity and the liquidity index. In addition, a recent study conducted by [32] suggests that granular soil, when used as a geofoam material, exhibits promising characteristics in handling shear stresses under various loading conditions compared to cohesive soil. To sum it all up, ensuring the structural stability of civil engineering foundations, regardless of the soil type, is critical, and it largely depends on the shear strength parameters of the foundation soil to prevent shear failure. However, analyzing these parameters is complex and time-consuming, posing significant challenges. Addressing these discrepancies by developing reliable models to predict the bearing capacity of soil foundations using lab and field data is quite meaningful. Thus, this research addresses this need by correlating observed shear strength data of shallow foundations with predictions from Artificial Neural Network (ANN) and Multiple Linear Regression (MLR) models, aiming to demonstrate the reliability and precision of these models in determining the bearing capacity of foundation soil, particularly for shallow foundations.

#### 2. Model Theory

### 2.1 Artificial Neural Network (ANN)

ANNs are computational models that mimic the human brain's method of processing information (Figure 1). They are composed of interconnected units called neurons, which are arranged into layers—comprising an input layer, one or more hidden layers, and an output layer [33]. Neurons in one layer are linked to those in the next through weighted connections. During training, data is passed through these layers, and each connection is associated with a specific weight and bias. Neurons process this input by combining the weighted values and using a mathematical activation function that introduces non-linearity, enabling the network to detect and learn from complex data relationships [34]. ANNs are capable of uncovering patterns in data and generating accurate predictions. They are widely used in tasks such as image classification and speech processing. After one-layer processes the data, the resulting information is passed to the next layer for further computation [35].

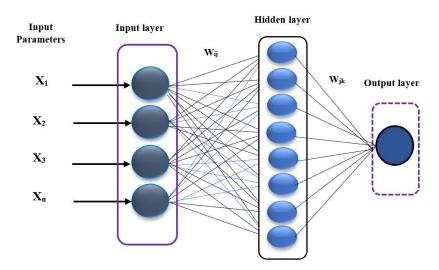


Figure 1: ANN model structure

## 2.2 Multiple Linear Regressions (MLR)

MLR is an extension of the standard regression approach that examines the relationship between a continuous dependent variable and two or more independent variables [36]. It assumes that the expected value of the dependent variable is a linear combination of the predictor variables (Figure 2). MLR is frequently applied in modelling and predicting the mechanical properties of construction materials [37]. Mathematically, the method fits a hyperplane within an n-dimensional space, where n corresponds to the number of independent variables. The model determines the unknown parameters by relating input (independent) variables to output (dependent) variables.

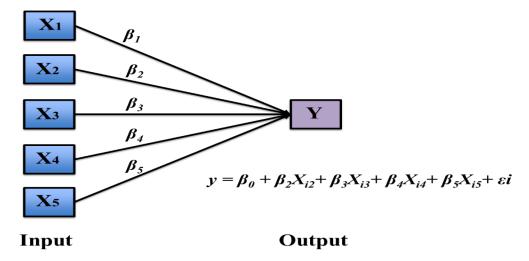


Figure 2: MLR model structure

## 3. Methodology

#### 3.1 Data Collection

The research utilized secondary data sourced from the Department of Civil Engineering Library at Kano University of Science and Technology (KUST) in Wudil, Nigeria. This data, extracted from seven prior research studies, includes soil index properties and shear strength parameters (specifically cohesion (c) and angle of internal friction  $(\phi)$ . A total of 28 datasets were compiled according to depth increments. The information was gathered using the soil parameters, including Atterberg's limit, particle size distribution, specific gravity, and shear parameters.

## 3.2 Development of Models Using MLR Technique And ANN

The identified input variables were used to construct two distinct predictive models: an MLR model and an ANN model. The MLR model was developed using Microsoft Excel 2010 Pro, leveraging its regression analysis tool to establish a mathematical relationship between the dependent and independent variables (Figure 3). Conversely, the ANN model was created using MATLAB, which provided advanced computational capabilities for training, validating, and testing the network. The MLR technique assumes that a linear relationship exists between the predictor variables and the target variable, allowing for the formulation of a deterministic equation, as shown in Figure 2. For the ANN model, a feedforward backpropagation network architecture was adopted, consisting of an input layer corresponding to the selected independent variables, one or more hidden layers with a specified number of neurons, and an output layer representing the predicted shearing strength parameters. The network was trained using a supervised learning approach, with the dataset divided into training, validation, and testing subsets to ensure generalization capability. Nonlinear transfer functions (e.g., tansig or logsig) were applied in the hidden layers. In contrast, a linear transfer function (purelin) was used at the output layer to capture both linear and nonlinear relationships. The development of these models aimed to compare the predictive performance of a traditional statistical approach (MLR) with that of a datadriven intelligent system (ANN). This comparative analysis was expected to provide insights into the

adequacy of each modeling technique in accurately forecasting shearing strength parameters based on the defined input variables.

#### 3.3 Data Processing

Multiple scatter plots are shown in a matrix style in a scatter matrix chart, a sort of data visualization. A visual summary of the pairwise associations between all variables in the dataset is provided by the complete matrix (Figure 4). In contrast, each scatter plot in the matrix illustrates the relationship between two variables in the dataset. A separate row represents each variable and column in a scatter matrix chart, and the scatter plots within the matrix illustrate the relationship between each pair of variables. The diagonal of the matrix displays a histogram or density plot for each variable, providing a visual representation of the distribution of values for each variable. The benefit of a scatter matrix chart is that it simplifies the identification of patterns, trends, and correlations by providing a rapid and thorough overview of the interactions between every variable in the dataset [1]. Because it enables the viewer to rapidly evaluate the relationships between all pairs of variables, rather than having to examine each pair separately, it is particularly beneficial when working with a large number of variables.

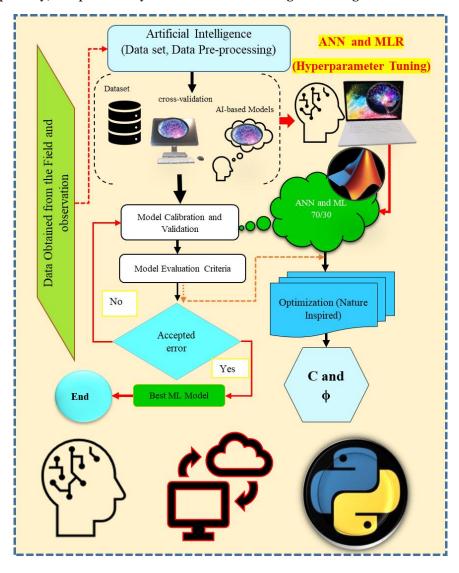


Figure 3: Methodology Flowchart

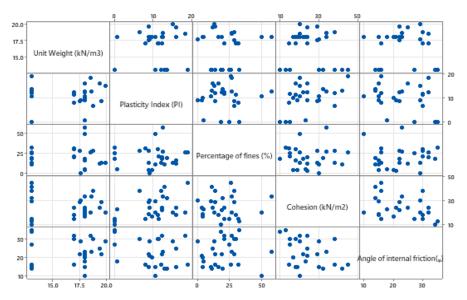


Figure 4: Data Visualization chart

# 3.4 Relative Statistical Measures of the Dataset

Interval plots enable rapid mean comparison and evaluation of potentially significant differences by graphically representing confidence intervals for various groups. While non-overlapping intervals may indicate a statistically significant difference, overlapping intervals suggest similar means. Narrower intervals imply a more accurate estimate, whereas wider intervals show a less precise estimate [25]. Figures 5a-5f below illustrate the relationships between the considered parameters. Table 1 describes the basic statistical variables.

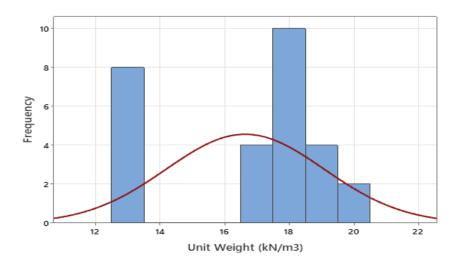


Figure 5a: Distribution of Unit weights

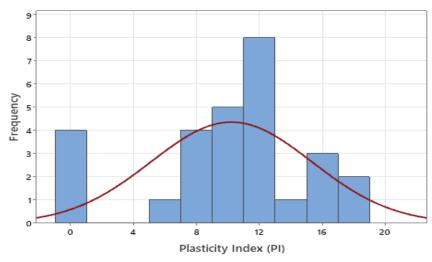
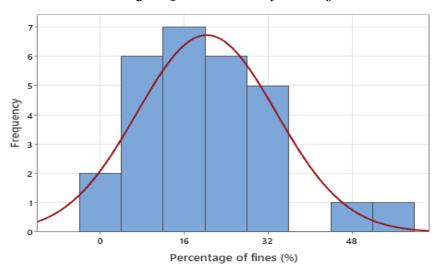


Figure 5b: Distribution of Plasticity Index



**Figure 5c:** Distribution of % fines

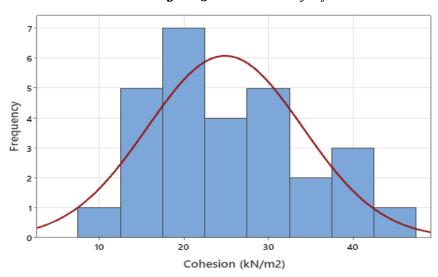


Figure 5d: Distribution of cohesion

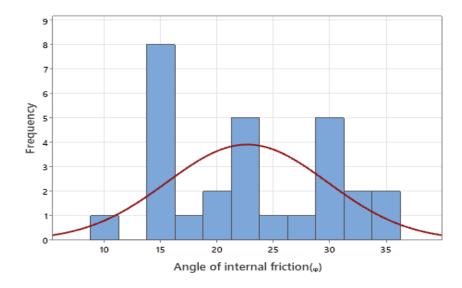


Figure 5e: Distribution of the Angle of internal friction

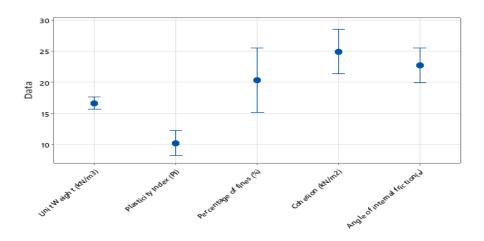


Figure 5f: Trend of all input parameters

 Table 1. Statistical properties of the dataset.

| <b>Unit Weight</b> | <b>Unit Weight</b> | Plasticity | Percentage of | Cohesion   | Angle of internal     |
|--------------------|--------------------|------------|---------------|------------|-----------------------|
| $(kN/m^3)$         | $(kN/m^3)$         | Index (PI) | fines (%)     | $(kN/m^2)$ | friction ( $\theta$ ) |
| Mean               | 16.6382            | 10.2221    | 20.3059       | 24.8821    | 22.6786               |
| Standard Error     | 0.46269            | 0.96835    | 2.5111        | 1.7369     | 1.3481                |
| Median             | 17.79              | 10.95      | 18.99         | 23.5       | 22                    |
| Mode               | 18                 | 0          | 28.8          | 24         | 30                    |
| Standard           | 2.4483             | 5.1240     | 13.2878       | 9.1908     | 7.1339                |
| Deviation          |                    |            |               |            |                       |
| Sample             | 5.9943             | 26.2559    | 176.5654      | 84.4712    | 50.8929               |
| Variance           |                    |            |               |            |                       |
| Kurtosis           | -1.1404            | 0.2960     | 1.3929        | -0.5323    | -1.2051               |
| Skewness           | -0.7004            | -0.6883    | 0.93764       | 0.51835    | 0.1091                |
| Range              | 6.94               | 19         | 57.468        | 35         | 25                    |
| Minimum            | 13                 | 0          | 0.132         | 10         | 10                    |

| Maximum | 19.94 | 19 | 57.6 | 45 | 35 |
|---------|-------|----|------|----|----|
|---------|-------|----|------|----|----|

**Table 2.** The Pearson correlation matrix

|                       | <b>Unit Weight</b> | Plasticity  | Percentage of | Cohesion    | Angle of internal   |
|-----------------------|--------------------|-------------|---------------|-------------|---------------------|
|                       | $(kN/m^3)$         | Index (PI)  | fines (%)     | $(kN/m^2)$  | friction $(\theta)$ |
| Unit Weight           | 1                  |             |               |             |                     |
| $(kN/m^3)$            |                    |             |               |             |                     |
| Plasticity Index      | 0.252954294        | 1           |               |             |                     |
| (PI)                  |                    |             |               |             |                     |
| Percentage of         | -0.014223516       | 0.047214755 | 1             |             |                     |
| fines (%)             |                    |             |               |             |                     |
| Cohesion              | -0.159279248       | 0.424170356 | 0.017043187   | 1           |                     |
| $(kN/m^2)$            |                    |             |               |             |                     |
| Angle of internal     | 0.121978781        | -           | 0.070481846   | -0.36471865 | 1                   |
| friction ( $\theta$ ) |                    | 0.265327026 |               |             |                     |

#### 3.5 Pearson Correlation Matrix

The Pearson correlation matrix is a widely used statistical tool that provides insight into the degree of linear association between pairs of variables in a dataset. Each element of the matrix represents the correlation coefficient (R) between two variables, with values ranging from -1 to +1. A value of +1 indicates a perfect positive correlation, meaning that as one variable increases, the other increases proportionally (Table 2). Conversely, a value of -1 1signifies a perfect negative correlation, where a rise in one variable results in a proportional decrease in the other. In the present study, the correlation matrix was computed to identify the strength and direction of associations between the input variables (such as soil properties, loading conditions, and shearing parameters) and the dependent variables related to bearing capacity. This analysis is instrumental in highlighting potential multicollinearity issues among independent variables, which may affect the performance of regression-based models such as MLR. Furthermore, the correlation matrix provides preliminary insights into the relative influence of input variables on the output, serving as a guide for model development. For instance, variables that exhibit strong positive or negative correlations with the shearing strength parameters are likely to play a significant role in predictive modeling. On the other hand, weak correlations suggest that such variables may have a limited or indirect effect. By employing the Pearson correlation matrix, this study ensures that the relationships among variables are systematically examined before advanced modeling with MLR and ANN techniques, thereby improving the reliability of the developed models.

#### 3.6 Hyper-parameter Tuning Process in ANN

In this study, the ANN was designed and optimized through systematic hyperparameter tuning to achieve the best predictive performance for estimating the shearing strength parameters. The adopted network architecture was a multilayer perceptron (MLP), consisting of an input layer, a single hidden layer, and an output layer. After several trials, a hidden layer configuration with nine neurons was selected because it produced the lowest absolute percentage error for both training and testing datasets [3]. Each neuron in the hidden layer was fully connected to the neurons in the preceding and subsequent layers through weighted connections. These weights served as adjustable parameters, continuously updated during training to minimize the error between predicted and actual outputs. The output layer neurons generated the final network predictions corresponding to the shearing strength parameters.

The Levenberg-Marquardt (LM) optimization algorithm was employed for training, as it combines the speed of the Gauss-Newton method with the stability of gradient descent. This algorithm is widely preferred for medium-sized networks due to its efficiency in minimizing the error function. To update the weights, the backpropagation algorithm was applied, which propagates the prediction error backward from the output to the hidden layer, iteratively adjusting the weights to reduce the overall error for each training pattern [3]. A sigmoid transfer function (log-sigmoid) was used in the hidden layer to introduce nonlinearity into the model, thereby enabling the ANN to capture complex, nonlinear relationships between input variables and outputs. In the output layer, a linear transfer function was adopted to produce continuous numerical predictions. Hyperparameters such as the learning rate and momentum rate were carefully tuned. The learning rate controlled the step size during weight updates, balancing the trade-off between convergence speed and stability. The momentum rate helped the model avoid local minima by incorporating a fraction of the previous weight update into the current adjustment, thus improving convergence efficiency. The network was trained iteratively, with performance monitored using training, validation, and testing datasets. Overfitting was controlled by observing validation error trends, and the final configuration was chosen to ensure both accuracy and generalization capability.

# 4. Results of the Models Developed

Figures 6a-6d present the correlations between observed and predicted values of shear strength parameters. The graph was plotted for training data, testing data, and validation datasets, particularly for the ANN. The best line fit was drawn to examine the model's effect on systematic and random error. The output was presented as a function of targeted results and R value, also expressed as derived from the R<sup>2</sup> value.

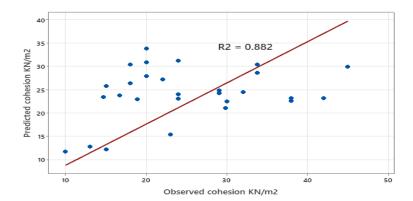


Fig. 6a: Correlation between observed and MLR estimated values of shear strength parameters

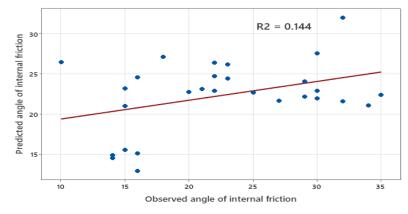


Figure 6b: Correlation between observed and MLR estimated values of the angle of internal friction

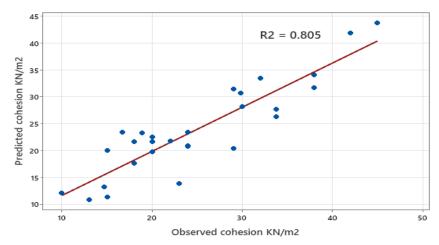


Figure 6c: Correlation between observed and ANN estimated values of shear strength parameters

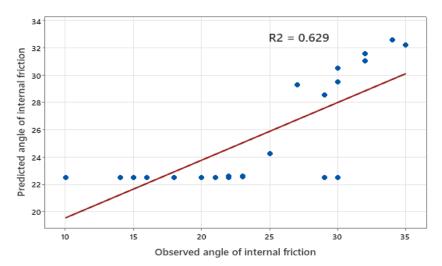


Figure 6d: Correlations between observed and ANN-predicted values of the angle of internal friction

#### 4.1 Comparison of the ANN Models

Figure 7a shows a comparison of the observed and predicted values of cohesion using an ANN. The chart indicates a strong relationship between the observed and predicted values, with some variations. Significant variation was observed in sample 10, where the expected value is more than twice the observed. Similarly, there is a higher decrease in predicted values for samples 21 and 22. In Fig. 7b, the observed and predicted values of the angle of internal friction using ANN are presented. As shown in the cohesion in Fig. 7a, the values of the internal friction angle exhibit a similar prediction capability to that of the ANN, albeit with some slight variation between the observed and predicted values. In sample 10, the predicted value is almost three times that of the observed. Additionally, Samples 9, 25, and 28 also exhibit a significant increase in predicted values, whereas Samples 21 and 22 display higher observed values compared to their predicted counterparts.

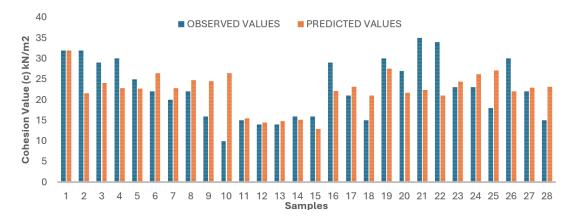
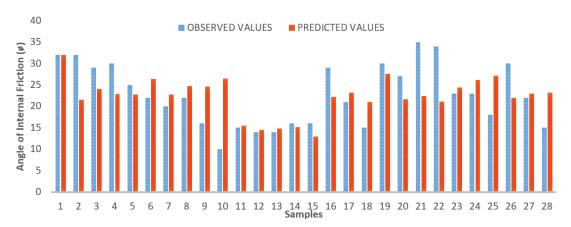


Figure 7a: Comparison between observed and predicted values of cohesion using ANN



**Figure 7b:** Comparison between observed and predicted values of the angle of internal friction from ANN

# 4.2 Comparison of the MLR Models

Figure 8a presents the relationship between observed and predicted values of cohesion using MLR. The model exhibits better performance, albeit with some significant variation, in samples 9, 10, 16, 21, 22, 25, 26, and 28. In samples 9, 10, 25, and 28, the predicted values are significantly higher than the observed, and vice versa in samples 16, 21, 22, 25, and 26. In Figure 8b, a comparison between the observed and predicted values of the angle of internal friction using MLR was presented. The model generally performs well, but there are noticeable differences in samples 9, 10, 16, 21, 22, 25, 26, and 28. Specifically, in samples 9, 10, 25, and 28, the predicted values are significantly higher than the observed ones, while in samples 16, 21, 22, 25, and 26, the predicted values are notably lower.



Figure 8a: Comparison between observed and predicted values of cohesion using MLR

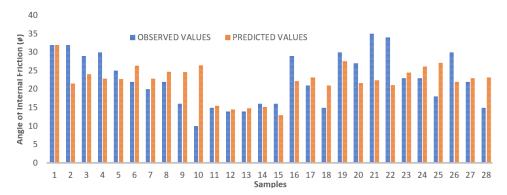


Fig.8b: Comparison between observed and predicted values of angle of internal friction using MLR

Table 3.0: Performance criteria of the models developed for both observed and predicted values

|   | ANN            |         |        |        |  |  |
|---|----------------|---------|--------|--------|--|--|
|   | $\mathbb{R}^2$ | MSE     | RMSE   | R      |  |  |
| A | 0.9749         | 17.0160 | 4.1250 | 0.9874 |  |  |
| В | 0.9499         | 26.5235 | 5.1500 | 0.9746 |  |  |
|   |                | M       | LR     |        |  |  |
|   | $\mathbb{R}^2$ | MSE     | RMSE   | R      |  |  |
| С | 0.8788         | 82.3801 | 9.0763 | 0.9374 |  |  |
| D | 0.9141         | 44.8042 | 6.6935 | 0.9561 |  |  |

Where A and B represent the performance criteria for the of actual and predicted cohesion and angle of internal friction obtained from ANN, while C and D represent the details of the performance criteria of actual and predicted cohesion and angle of internal friction obtained using MLR.

In general, the R² values of groups A and B were better than those of groups C and D, indicating that the ANN model provides a better prediction for the testing set than the training data compared to MLR, although the variation is tolerable. The combination of the transfer function component trans-sigmoid and the linear function yields better results. Again, the multiple linear regression shows a good correlation between the input and output parameters; the coefficients of correlation for cohesion and angle of internal friction, R², were 0.8789 and 0.91415, respectively. This proves that the linearity in the data was better. Moreover, the R² results show the robustness of the models and that the independent variables have excellent correlation with the dependent. To estimate the shear strength parameters, three additional evaluation criteria (MSE, RMSE, and R) were used to identify the best-fitting regression line for both MLR and ANN. The RMSE, being an error measure, has the added significance that minor errors are given less concern than significant errors. However, according to [38] and [39]. Other evaluation parameters should be used in addition to the error measure RMSE to achieve higher accuracy.

#### 5.0 Conclusion

This study investigated the predictive capacity of Artificial Neural Networks (ANN) and Multiple Linear Regression (MLR) models in estimating the shear strength parameters of soils using readily available soil index properties. The development of these models was grounded in actual experimental data, which served both training and validation, ensuring that the predictive outcomes remained realistic and reliable. The results demonstrated that both MLR and ANN were able to capture the relationships between soil index properties and shear strength parameters with a satisfactory degree of accuracy. While the MLR model provided a straightforward linear representation of the underlying relationships, the ANN model offered flexibility to account for nonlinear dependencies, thereby yielding improved precision in several cases. The close agreement observed between the predicted values and the experimental results underscores the robustness of both modeling approaches. A notable advantage of these models lies in their reliance on simple soil index properties, which are relatively straightforward, cost-effective, and easily determined in geotechnical practice. By using such data as input, the models significantly reduce the need for expensive and time-consuming laboratory shear strength tests, without

compromising the reliability of predictions. The graphical comparisons between predicted and experimental results further confirmed the consistency of the models, with both methods showing a substantial degree of similarity to observed behavior. The findings also highlight that numerical modeling offers a practical alternative for estimating shear strength parameters in scenarios where time, resources, or accessibility may limit direct experimental determination. In particular, the ANN model's capability to handle nonlinear patterns gives it an edge in predicting complex soil behavior, making it a valuable tool for practitioners and researchers. In conclusion, both MLR and ANN provide efficient, precise, and practical tools for predicting soil shear strength parameters. Their successful application in this study demonstrates their potential for broader adoption in geotechnical engineering design and analysis. By integrating such models into routine practice, engineers can enhance decisionmaking processes, optimize foundation design, and improve the reliability of geotechnical evaluations. Future studies may further refine these models by incorporating additional variables, larger datasets, and advanced learning techniques, thereby strengthening their predictive capability and generalization across diverse soil conditions.

**Competing Interests:** The authors declare that they have no competing interests.

Data Availability Statement: The supported data associated with this researcher is available upon request from the corresponding author.

#### References

- [1] R. Jain, P. K. Jain, and S. S. Bhadauria, "Computational approach to predict soil shear strength," Int. J. Eng. Sci. Technol., vol. 2, no. 8, pp. 3874–3885, 2010.
- V. K. Too, "3D Modelling for geotechnical investigation." University of Nairobi, 2012.
- F. Z. Rick LiuS, W. YanL, "Vane shear strength based stability analysis of slopes in unconsolidated soft [3] clay," Trans. Tianjin Univ., vol. 10, no. 3, pp. 195-200, 2004.
- A. Ghahramani, M. Ajdari, G. Habibagahi, H. Nowamooz, and F. Masrouri, "Shear strength behavior and [4] soil water retention curve of a dual porosity silt-bentonite mixture," Sci. Iran., vol. 17, no. 6, 2010.
- M. A. Tekinsoy, C. Kayadelen, M. S. Keskin, and M. Söylemez, "An equation for predicting shear strength [5] envelope with respect to matric suction," Comput. Geotech., vol. 31, no. 7, pp. 589-593, 2004.
- [6] D. G. Fredlund and S. K. Vanapalli, "2.7 Shear Strength of Unsaturated Soils," Methods Soil Anal. Part 4 Phys. Methods, vol. 5, pp. 329–361, 2002.
- S. R. Surendra Roy and G. D. Gurcharan Dass, "Statistical models for the prediction of shear strength [7] parameters at Sirsa, India.," 2014.
- H. Ersoy, M. B. Karsli, S. Cellek, B. Kul, I. Baykan, and R. L. Parsons, "Estimation of the soil strength [8] parameters in Tertiary volcanic regolith (NE Turkey) using analytical hierarchy process," J. Earth Syst. Sci., vol. 122, no. 6, pp. 1545-1555, 2013.
- V. N. S. Murthy, Geotechnical engineering: principles and practices of soil mechanics and foundation [9] engineering. CRC press, 2002.
- [10] M. I. Alsaleh, Numerical modeling of strain localization in granular materials using Cosserat theory enhanced with microfabric properties. Louisiana State University and Agricultural & Mechanical College, 2004.
- A. Mollahasani, A. H. Alavi, A. H. Gandomi, and A. Rashed, "Nonlinear neural-based modeling of soil [11] cohesion intercept," KSCE J. Civ. Eng., vol. 15, no. 5, pp. 831–840, 2011.
- T. F. A. Bishop and A. B. McBratney, "A comparison of prediction methods for the creation of field-extent [12] soil property maps," Geoderma, vol. 103, no. 1-2, pp. 149-160, 2001.
- J. Arvidsson and T. Keller, "Comparing penetrometer and shear vane measurements with measured and [13] predicted mouldboard plough draught in a range of Swedish soils," Soil Tillage Res., vol. 111, no. 2, pp. 219-223, 2011.
- A. A. MUSA and S. DULAWAT, "STRENGTH EVALUATION OF CLAYEY SOIL STABILIZED WITH [14]
- MARBLE SLURRY USED AS SUBGRADE SOIL MATERIAL," 2021. T. Masada, "Shear strength of clay and silt embankments.," Ohio Research Institute for Transportation [15] and the Environment, 2009.
- [16] S. A. Mofiz and M. M. Rahman, Shear strength behavior of barind soil on triaxial extension stress path tests. University of Canterbury. Civil and Natural Resources Engineering, 2010.
- S. Cola and G. Cortellazzo, "The shear strength behavior of two peaty soils," Geotech. Geol. Eng., vol. 23, [17] no. 6, pp. 679-695, 2005.
- A. Hajdarwish and A. Shakoor, "Predicting the shear strength parameters of mudrocks," in Geol Soc [18] Lond, 2006, vol. 2, p. 607.
- E. Köken and T. K. Koca, "A comparative study to estimate the mode I fracture toughness of rocks using [19] several soft computing techniques," Turkish J. Eng., vol. 7, no. 4, pp. 296-305, 2023.

- G. O. Adunoye, "Fines content and angle of internal friction of a lateritic soil: an experimental study," [20] Am. J. Eng. Res., vol. 3, no. 3, pp. 16-21, 2014.
- [21] E. L. Bravo, M. H. Suárez, O. G. Cueto, E. Tijskens, and H. Ramon, "Determination of basics mechanical properties in a tropical clay soil as a function of dry bulk density and moisture," Rev. Ciencias Técnicas Agropecu., vol. 21, no. 3, pp. 5-11, 2012.
- G. O. Adunoye, "Study of relationship between fines content and cohesion of soil.," 2014. [22]
- [23] I. I. Aminu et al., "Ensemble Machine Learning Technique Based on Gaussian Algorithm for Stream Flow Modelling," vol. 1, no. July, pp. 1–17, 2025.
- S. I. Malami et al., "Implementation of soft-computing models for prediction of flexural strength of [24] pervious concrete hybridized with rice husk ash and calcium carbide waste," Model. Earth Syst. Environ., vol. 8, no. 2, pp. 1933-1947, 2022, doi: 10.1007/s40808-021-01195-4.
- B. Ghoreishi et al., "Assessment of geotechnical properties and determination of shear strength [25] parameters," Geotech. Geol. Eng., vol. 39, no. 1, pp. 461-478, 2021.
- [26] A. S. Ibrahim, A. A. Musa, A. Y. Abdulfatah, and A. Idris, "Developing soft-computing regression model for predicting soil bearing capacity using soil index properties," Model. Earth Syst. Environ., vol. 9, no. 1, pp. 1223-1232, 2023, doi: 10.1007/s40808-022-01541-0.
- H. Tanaka, "Re-examination of established relations between index properties and," in Coastal [27] Geotechnical Engineering in Practice, Volume 2: Proceedings of the International Symposium IS-Yokohama 2000, Yokohama, Japan, 20-22 September 2000, 2002, vol. 2, p. 3.
- P. La Rochelle, B. Trak, F. Tavenas, and M. Roy, "Failure of a test embankment on a sensitive Champlain [28] clay deposit," *Can. Geotech. J.*, vol. 11, no. 1, pp. 142–164, 1974.

  P. W. Mayne and F. H. Kulhawy, "Ko-OCR relationships in soil," *J. Geotech. Eng. Div.*, vol. 108, no. 6, pp.
- [29] 851-872, 1982.
- S. M. Ahmed, "Assessment of clay stiffness and strength parameters using index properties," J. Rock [30] Mech. Geotech. Eng., vol. 10, no. 3, pp. 579-593, 2018.
- T. Tafari, E. T. Quezon, and M. Yasin, "Statistical analysis on shear strength parameter from index [31] properties of fine-grained soils," Available~SSRN~3811717, 2021.
- [32] T. Karademir, "Counterface soil type and loading condition effects on granular/cohesive soil-Geofoam interface shear behavior," Turkish J. Eng., vol. 8, no. 1, pp. 76-91, 2024.
- C. Grosan and A. Abraham, "Artificial Neural Networks," vol. 17, pp. 281-323, Jan. 2011, doi: [33] 10.1007/978-3-642-21004-4\_12.
- U. Usman Aliyu, S. I. Malami, M. M. Jibril, I. A. Mahmoud, U. J. Muhammad, and J. Z. Yau, "Design of [34] machine learning model for predicting the compressive strength of fabric fiber-reinforced Portland cement," *Techno-computing J.*, vol. 1, no. 1, pp. 14–26, 2025, doi: 10.71170/tecoj.2025.1.1.pp14-26.
- R. Mahadeva et al., "Improvised grey wolf optimizer assisted artificial neural network (IGWO-ANN) [35] predictive models to accurately predict the permeate flux of desalination plants," Heliyon, vol. 10, no. 13, p. e34132, 2024, doi: 10.1016/j.heliyon.2024.e34132.
- L. V Casaló and J.-J. Escario, "Heterogeneity in the association between environmental attitudes and pro-[36] environmental behavior: A multilevel regression approach," *J. Clean. Prod.*, vol. 175, pp. 155–163, 2018. K. Stergiou, C. Ntakolia, P. Varytis, E. Koumoulos, P. Karlsson, and S. Moustakidis, "Enhancing property
- [37] prediction and process optimization in building materials through machine learning: A review," Comput. Mater. Sci., vol. 220, no. August 2022, p. 112031, 2023, doi: 10.1016/j.commatsci.2023.112031.
- V. Cherkassky, V. Krasnopolsky, D. P. Solomatine, and J. Valdes, "Computational intelligence in earth sciences and environmental applications: Issues and challenges," *Neural Networks*, vol. 19, no. 2, pp. [38] 113-121, 2006.
- A. S. Ibrahim, A. Idris, and A. Musa, "Predicting Ultimate Bearing Capacity from Shear Strength [39] Parameters Using Artificial Neural Network," 2022.