JENER Journal of Empirical and Non-Empirical Research
Volume 1, Issue 3, Page 328-333 | Article No. 38
DOI: 10.4898/jener.v2i1.a38
ISSN: 7895-9216
Journal homepage: https://jenerjournal.com

JENER Journal
JENER Journal of Empirical and Non-Empirical Research

Modeling of Shear Strength of Lateritic Soil using Python Programming Language

Ijeoma I. Nwajuaku[1]* & John C. Okonkwo[2]

[1,2]*Department of Civil Engineering, Nnamdi Azikiwe University, Awka, Nigeria.*
*Corresponding Author: Ijeoma I. Nwajuaku, Email: ii.nwajuakuijeoma@unizik.edu.ng

APA Citation and Referencing: Nwajuaku, I.I., & Okonkwo, J.C. (2026). Modeling of Shear Strength of Lateritic Soil using Python Programming Language. *JENER Journal of Empirical and Non-Empirical Research*, 2(1), 328-333

ARTICLE INFORMATION

Article history:
Published on 24th Jan 2026

Keywords:
Modeling
Shear strength
Lateritic soil
Index properties
Python programming language

ABSTRACT

This research aimed at modeling of shear strength of lateritic soil using python programming language. Shear strength is a property of soil that affects the stability of a soil mass under structural loads. Hence, adequate measurement is needed prior to any structural design. In this study, disturbed lateritic soil samples were collected at a depth of 1.5 m from ground surface and analyzed for different geotechnical properties such as specific gravity, particle size, liquid limit, plasticity index and shear strength. The values of the index properties of the soil were used to develop a multiple linear regression (MLR) model for shear strength prediction. The coefficients of the dependent variables of the MLR model were obtained under python framework. The results showed that the developed model achieved a high value of coefficient of determination (R^2 =1) and root mean square of 0.415. Therefore, the performance of the MLR showed that the shear strength of the lateritic soil in the study area can be computed successfully.

1. Introduction

Shear strength is a critical parameter in the field of structural engineering such as foundations and road pavement design (Ojuri, 2013). The determination of shear strength demand resources, time and labor both in-situ and in the laboratory (Roy & Dass, 2014; Ersoyl et al., 2013; Arvidsson & Keller, 2011). Hence, for effective handling of the problems of experimental study, engineering modeling and simulation are required. In research, correlations and empirical relationships are used when a geotechnical analysis is not achievable as a result of time or monetary limitations. This is a situation frequently seen in developing countries where there is inadequate state-of-art testing equipment and manpower required to run it. Currently, there are developed model uses simple secondary data to predict soil parameters (Bishop & McCartney, 2001). This involves empirical correlations for estimation of geotechnical properties of soils. However, despite their applications, they still have limitations in predicting shear strength for various soil mixtures (Goktepe, 2008).

Generally, lateritic soils can be defined as any red or reddish brown soil that obtained from the process of rock weathering with or without nodular concretions (Ojuri, 2013). In recent years, efforts have been made towards modeling of the strength properties of lateritic soils. Few studies (Sorensen & Okkels, 2013; Tafari et al., 2021; Eyael et al., 2024) have proved that important correlations exist between the index properties obtained from testing and the strength properties of soils. The index properties have vast applications in geotechnical engineering and have been successfully used by researchers (Osinubi & Bello 2011; Bello, 2011a; Bello, 2011b) in predicting soil behavior (Bello, 2011c).

Many empirical and polynomial models have been employed to determine shear strength parameters (Kiran & Lal, 2015). However, the emergence of machine learning provides a promising substitute to conventional method (Eyael et al., 2024). Studies by these researchers (Iyeke et al., 2016; Khanlari et al., 2012) have successfully predicted the shear strength parameters (internal friction angle 'Φ' and cohesion 'c') of soils using artificial neural network modeling technique. Hence, the basic novelty of this study is on the application of python programming language to develop a multiple linear regression (MLR) model for shear strength, using more than two index properties as independent variables.

2. Literature Review

From related literature, all the studies focused primarily on predicting the soil shear stress parameters using either statistical method or machine learning techniques. Tafari et al. (2021) study applied statistical models of both single and multiple linear regressions to identify the relationships between and index properties of fine grained soil and shear strength. Several models were developed between shear strength (dependent variable) and one independent variable (as single linear regression) or two variables (multiple regressions). The independent variables are the specific gravity, liquid limit, plastic limit, plasticity index, liquidity index, bulk density, dry density, and natural moisture content. Finally, the best model with a coefficient of determination (R^2 = 0.806) was obtained from all the proposed models. In the machine learning application, ANN is generally adopted for model development. Contrary to most conventional statistical methods, this computational tool use only data to determine the structure of a model and its unknown parameters (Kiran & Lal 2015). It does assume a model form, as required in the parametric techniques. Kiran & Lal (2015) employed artificial neutral network (ANN) to develop a model from the index properties of a soil and predict the internal friction angle and cohesion. A regression analysis was performed using multivariable linear least square regression (LLSR) and non-linear least square regression (NLSR). The input parameters include water content, plasticity index, bulk density, sand, silt, and clay. The results showed that all the regression technique performed well in predicting cohesion, while ANN outperformed over the LLSR and NLSR indicating internal friction angle with high R value. Additionally, Eyael et al. (2024) evaluated ANN model for predicting soil stress parameters. The research utilizes soil index properties such as sand, fines, liquid limit, plastic limit, and plasticity index to produce separate ANN models for internal friction angle and cohesion. The models achieved success in predicting shear strength parameters, with correlation values of 0.99 and 0.98 for cohesion and internal friction angle respectively. These computational techniques aim to reduce the time, cost and effort involved in geotechnical engineering procedure (Murata, 1994; Paruelo & Fernando 1997). These studies have shown to be at the nexus of conventional geotechnical methods and modern computational techniques, which could transform the estimation and practical application of shear strength in real-world scenarios.

3. Methodology

3.1 Soil Sample Collection and Laboratory Test

The main materials used for this research are lateritic soil. The lateritic soil sample used was collected from a borrow pit located at Nibo town in Awka South local government area of Anambra State, Nigeria. The geographical location of the study area lies between latitude 6^{o} 09" 60" N and longitudes 7° 03' 00" E. The disturbed soil sample was collected at a depth of 1.5m and transported to the Civil Engineering Laboratory, Nnamdi Azikiwe University. The soil was air dried for five days to allow for partial elimination of natural water and then sieved with sieve no. 4 (4.75mm) to obtain the final soil samples for the tests. Four main geotechnical test; grain size analysis, specific gravity, atterberg limits, and unconfined compression were conducted on thirty (30) lateritic soil samples to determine their index properties and shear strengths. All tests were performed in accordance to BS1377 (BS, 1990) and BS5930 (BS, 1999).

3.2 Data Analysis

The purpose of this analysis is to develop a correlation between soil index properties and shear strength of soil samples used in the study. The software and tools used are python, scikit-learn and google collab. Google Colab is a free online version of Jupyter notebook hosted by Google.

3.2 Model Development

In this study, a multiple linear regression (MLR) model was employed to predict the shear strength of lateritic soils from index properties. The model is a used to establish a relationship between one dependent variable and two or more independent variables. It assumes that the dependent variable can be expressed as a linear combination of the independent variables and an error term.
Mathematically, the model can be expressed in equation (1) as:

$$SS = \beta_0 + \beta_{SG}.SG + \beta_{LL}.LL + \beta_{PL}.PL + \beta_{PI}.PI + \beta_{FSC}.FSC + \beta_{FC}.FC + \varepsilon \quad (1)$$

Where:

SS is shear strength of the lateritic soil (dependent variable),
β_0 is model intercept
$\beta_{1\ldots6}$ is regression coefficients (weights) representing the contribution of each predictor
SG, LL, PL, PI, FSC, FC are specific gravity, liquid limit, plastic limit, plasticity index, fine sand content and fines content respectively (independent variables)
ε is random error term

Linear regression is commonly used in geotechnical research for modeling soil strength parameters due to its ability to approximate complex physical relationships using simple, measurable properties. It also allows for the use of evaluation metrics such as coefficient of determination (R^2) and root mean square error (RMSE) to ascertain the predictive power of models.

3.3Python Modeling Procedure

The modelling process using python framework involves systematic approach and the first step is the importation of essential python libraries. These includes libraries for data manipulation (pandas and NumPy), model training (scikit-learn's LinearRegression), data scaling (StandardScaler), and model evaluation (r^2 score and root mean square error). Each library was selected for its suitability in handling structured datasets and execution of effective regression analysis.

After importing the required libraries, the next step was loading the dataset. The experimental data; the index properties of lateritic soil and its measured shear strength were stored as a comma-separated values (CSV) file and read into python as a dataframe. This allowed the dataset to be treated as a structured table in which each column represented predictor variables such as specific gravity, liquid limit, plastic limit, plasticity index, fine sand content and fines content, while the response variable was represented by the shear strength.

Once the dataset was loaded, data cleaning and inspection were performed by removing the rows and columns that were empty to ensure that only valid observations were used. The summary statistics and previews of the dataset were generated to verify that all values were numeric in the expected columns. This step is crucial to prevent errors during model training and also confirm the quality of the input data.

Following the cleaning process, the predictor variables (X) and the response variable (y) were clearly defined. The predictor and response variables were the index properties and shear strength respectively. The dataset was then split into training and testing sets. The initial dataset of 70% was used for model training, while the remaining 30% was used for model testing.
Before fitting the model, the independent variables were standardised using the StandardScaler. This transformed each predictor to have a mean of zero and a standard deviation of one, thereby placing all variables on a comparable scale. Although linear regression does not strictly require standardisation, it improves the numerical stability of the computations and makes the resulting coefficients more interpretable in terms of relative influence.
The linear regression model was then fitted on the scaled training data. The model estimated the intercept and coefficients that minimised the squared differences between the predicted and observed shear strengths. After training, the model was applied to the scaled test data to generate predictions of shear strength.
Finally, the model's performance was evaluated using the coefficient of determination (R^2) and the root mean square error (RMSE). R^2 quantified the proportion of variance in shear strength explained by the selected index properties, while RMSE measured the average prediction error in the same units as the dependent variable.

4. Findings

The index properties and shear strength of soils samples used in the study are presented in Table 1.

Table 1: Index properties of lateritic soils

Sample Code	Specific Gravity, SG	Liquid Limit, LL(%)	Plastic Limit, PL(%)	Plasticity Index, PI (%)	% of Soil Retained on 75μm Sieve, FSC	% of Soil Passing 75μm Sieve, FC	Unconfined Compressive Strength, UCS (kPa)	Shear Strength, SS, UCS/2 (kPa)
S1	2.64	44.2	23.0	21.2	52.7	47.3	176	88
S2	2.68	44.8	23.3	21.5	52.1	47.9	170	85
S3	2.61	43.7	22.7	20.9	51.3	46.7	181	91
S4	2.64	44.1	23.0	21.2	51.8	47.2	177	88
S5	2.58	43.2	22.5	20.7	52.8	46.2	186	93
S6	2.64	44.1	22.9	21.1	52.9	47.1	177	89
S7	2.55	42.7	22.2	20.5	52.3	45.7	191	96
S8	2.58	43.2	22.5	20.7	53.8	46.2	186	93
S9	2.60	43.5	22.6	20.9	53.4	46.6	183	91
S10	2.61	43.7	22.7	21.0	53.2	46.8	181	90
S11	2.68	38.0	26.0	12.0	57.3	40.7	238	119
S12	2.71	38.5	26.3	12.2	50.8	41.2	233	117
S13	2.65	37.5	25.7	11.9	57.8	40.2	243	121
S14	2.67	37.9	23.6	14.3	59.4	40.6	239	119
S15	2.62	37.1	24.4	12.8	58.3	39.7	247	123
S16	2.67	37.9	24.9	13.0	58.5	40.5	239	120
S17	2.59	36.7	24.1	12.6	58.7	39.3	251	126
S18	2.62	37.1	24.4	12.8	59.3	39.7	247	123
S19	2.64	37.4	24.6	12.9	60.0	40.0	244	122
S20	2.65	37.6	24.7	12.9	59.8	40.2	242	121
S21	2.66	48.2	23.6	24.6	48.4	51.6	136	68
S22	2.69	48.8	23.9	24.9	47.8	52.2	129	65
S23	2.63	47.6	23.3	24.3	44.0	51.0	141	71
S24	2.65	48.1	23.5	24.5	47.5	51.5	137	68
S25	2.60	47.1	23.1	24.0	49.6	50.4	147	73
S26	2.65	48.0	23.5	24.5	48.6	51.4	137	69
S27	2.57	46.5	22.8	23.7	50.2	49.8	152	76
S28	2.60	47.1	23.1	24.0	49.6	50.4	147	73
S29	2.62	47.5	23.2	24.2	49.2	50.8	143	72
S30	2.63	47.7	23.3	24.3	49.0	51.0	141	71

Source: Research Data, 2026

From Table 1, the plasticity index ranges from 11.9 – 24.9%, the plastic limit (PL) between 22-26%, while the liquid limit (LL) is from 37 to 49%, indicating moderate to high plasticity. The PL range is typical of tropical lateritic soils, indicates a moderate degree of plasticity that is influenced by oxides of iron and aluminum. The soil with a PL closer to 26% indicates higher clay content, with increased cohesiveness. These results are in conformity to some extent with the findings of Ogbuagu & Okeke, (2019).They worked on the geotechnical properties of lateritic soils from Nimo and Nteje areas of Anambra State, Southeastern Nigeria; where liquid limit ranged from 35- 46%, plastic limit was from 18-25%, and plasticity index was between 17- 21%. However, 35% of liquid limit and 18% of plastic limit indicating the probable absence of expandable clay materials. As reported in (BS, 1990), soil samples with liquid limit less than 35% are classified as low plasticity soil, between 50 to 70% as high plasticity soil, between 70 to 90% as very high plasticity soil, while greater that 90% as an extremely high plasticity soil. In addition, these results agrees with that of Adeyemi and Oyeyemi (2000); where LL values was found to range from 25% to 80% and PI values from 5% to 50% in lateritic soils from southwestern Nigeria, attributing this variability to differences in clay content and iron oxide concentration. The lateritic soil samples were experimentally tested and found to have a specific gravity (SG) range of 2.5-2.7. The strength of a soil mass is directly proportional to its specific gravity (Ogbuchukwu et al., 2019). The specific gravity values are within the range recommended in (Kiran & Lal, 2015). Since the specific gravity value is on the higher side, the lateritic soil samples contains low clay fraction; which indicates high strength. Tests on the lateritic soil samples revealed a fine sand content range of 51-60%. This range indicates a significant portion of fine sand, influencing soil texture and drainage. Higher fine sand content (near 60%) enhances permeability and reduces cohesiveness, potentially increasing shear strength. Conversely, lower values (51%) suggest more fines, affecting stability. Unconfined compressive strength (UCS) tests on the lateritic soil samples yielded a range of 136-247 kPa. This range indicates moderate to high strength, influenced by soil composition and moisture. Higher UCS values (near 247 kPa) suggest denser, less plastic soils, while lower values (136 kPa) indicate higher clay content, reducing strength. The MLR between index properties and the shear strength of lateritic soil is expressed in equation (2) as:

$$SS = 94.09 - 0.05(SG) + 9.93(LL) - 4.04(PL) - 19.18(PI) - 0.29\ (FSC) - 14.72(FC) \quad (2)$$

$$RMSE = 0.415 \quad (3)$$

$$R^2 = 1.0 \quad (4)$$

The python visual interface showing the processing functions and the generated coefficients of the MLR model are shown in Figures 1 and 2 respectively.

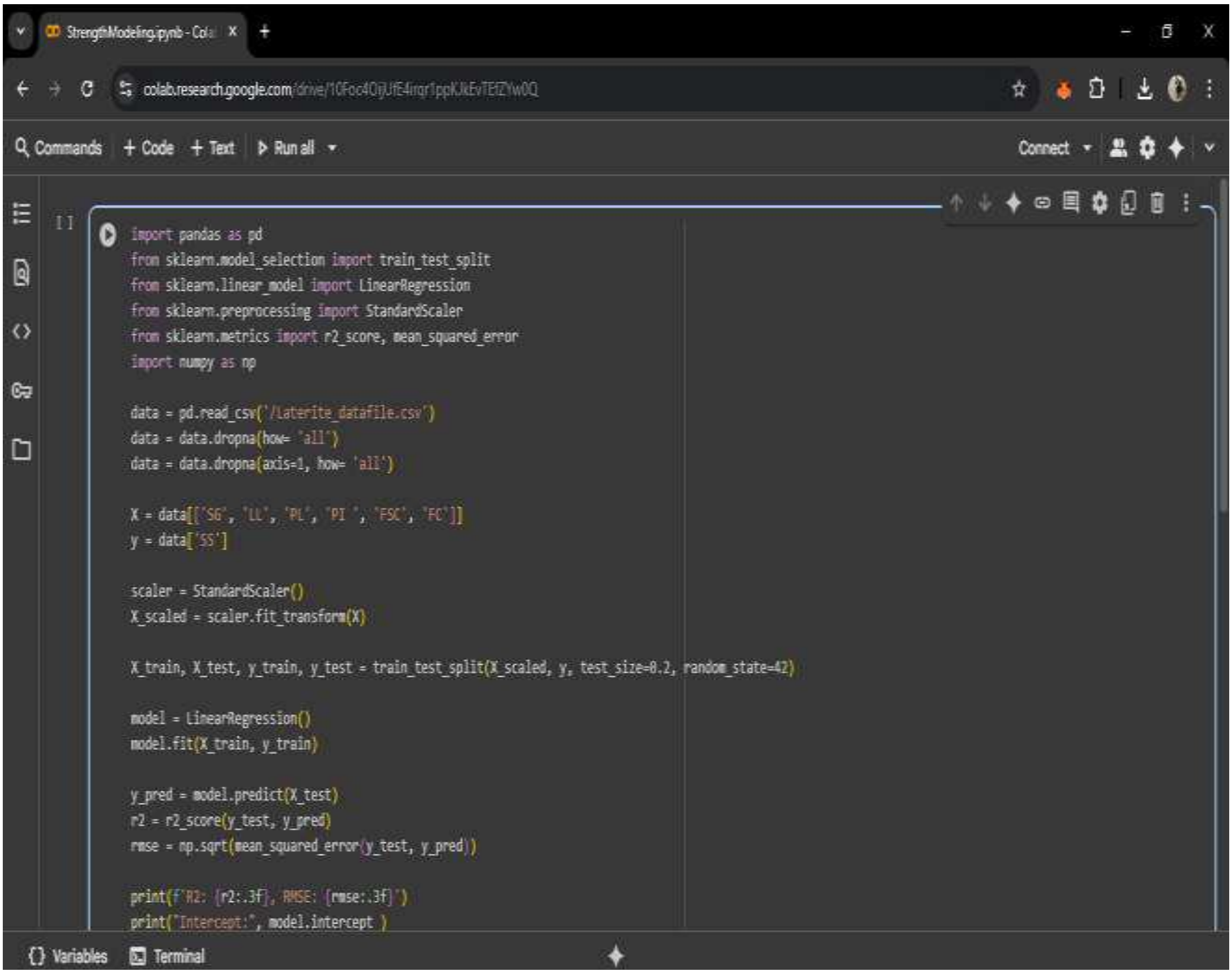

Figure 1: Python visual interface showing the processing functions. Source: Research Data, 202X

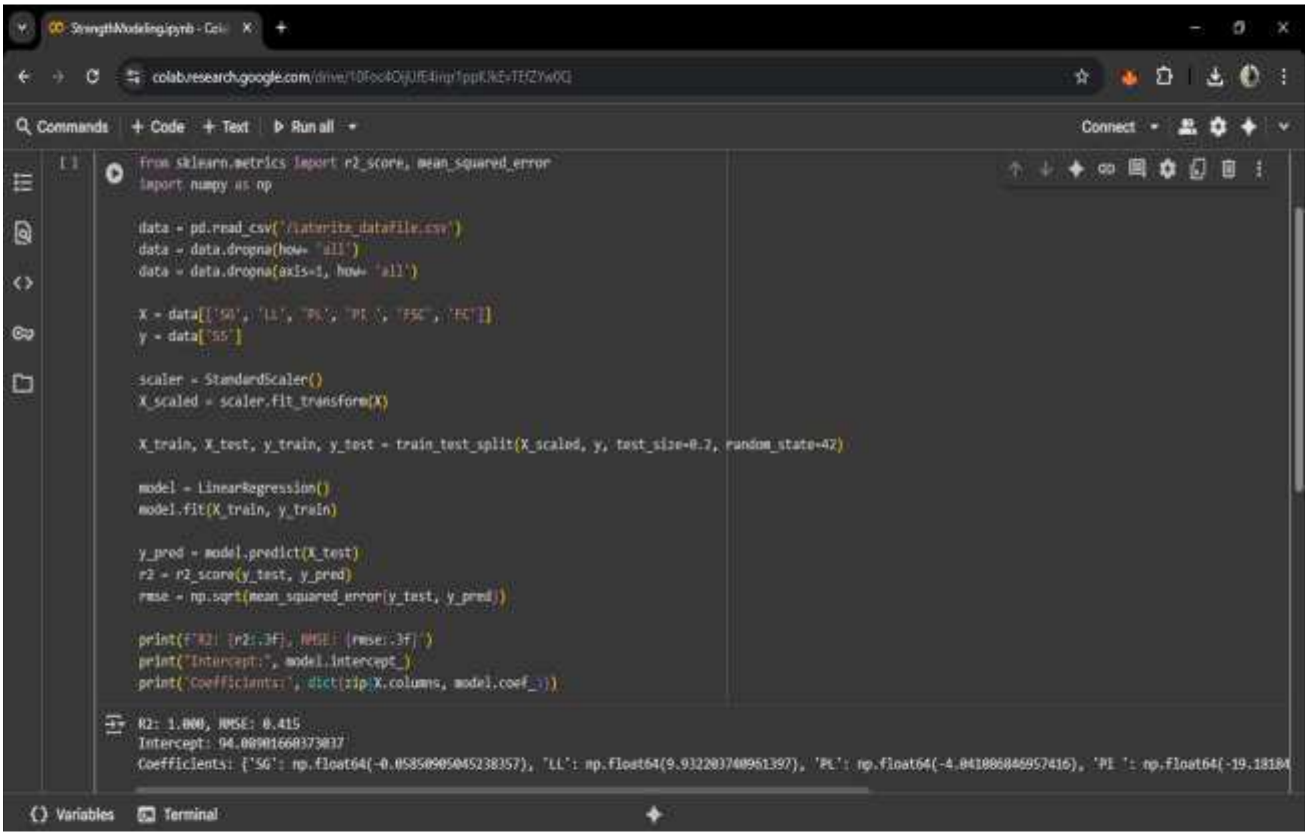

Figure 2: Results of the generated coefficients of the MLR model. Source: Research Data, 202X

5. Conclusion

The results demonstrated that the chosen model achieved remarkably high predictive accuracy ($R^2 = 1.0$). However, the variability in the geotechnical properties of lateritic soil necessitates the need for site-specific testing and modeling based on index properties. With further validation and refinement, the approach presented here could significantly reduce the reliance on time-consuming, laboratory testing and improve the efficiency of geotechnical investigations for infrastructure projects in regions dominated by lateritic soils. The resulted model provides engineers and practitioners with a powerful decision-making tool for faster and more economical preliminary assessments of soil strength

References

[1] Adeyemi, G. O., Oyeyemi, F. (2000). Geotechnical basis for failure of sections of the Lagos–Ibadan expressway, south western Nigeria. *Bulletin of engineering geology and the environment*, 59 (1), 39-45

[2] Arvidsson J., Keller, T.(2011). Comparing penetrometer and shear vane measurements with measured and predicted moldboard plow draft in a range of Swedish soils. *Soil and Tillage Research*, (111), 219-223.

[3] Bello A.A. (2011a). Analysis of Shear Strength of Compacted Lateritic Soils. *Pacific Journal of Science and Technology*. 12(1), 425-433.

[4] Bello A.A. (2011b).Influence of Compaction Delay on CBR and UCS of Cement Stabilized Lateritic Soil. *Pacific Journal of Science and Technology*. 12(2), 87-98.

[5] Bello A.A., (2011c). The use of Standard Proctor for the determination of Shrinkage Properties of Reddish Brown Tropical soil. *Leonardo Journal of Science*. 19 (2), 51-62.

[6] Bishop, T. F. A. and McBratney, A.B.,(2001). A comparison of prediction methods for the creation of field-extent soil property maps. *Geoderma* (103), 149–160.

[7] British Standard (BS) 1377-part 2, (1990). Methods of testing soils for Civil Engineering purposes. British Standards Institution, London

[8] British Standard (BS) 5930 (1999) Code of practice for site investigations. British Standard Institute, London, UK.

[9] Ersoyl, H., Karsli, M. B., Cellek, S., Kul, B., Baykan, I., Parsons, R. (2013). Estimation of the soil strength parameters in tertiary volcanic regolith (NE Turkey) using analytical hierarchy process. *J. Earth Syst. Sci.*, 122 (6),1545–1555.

[10] Eyael T. H., Srikanth V., Mnqobi N., and Peace M. (2024).Advanced prediction of soil shear strength parameters using index properties and artificial neural network approach. *World Journal of Advanced Research and Reviews.* 21(01), 427–445

[11] Goktepe, A. B., Altun, S., Altintas, G., Tan, O. (2008). Shear strength estimation of plastic clays with statistical and neural approaches. in: *Building and Environment*, 43(5),849-860.

[12] Iyeke, S. D., Eze, E. O., Ehiorobo, J. O., Osuji, S. O. (2016). Estimation of shear strength parameters of lateritic soils using artificial neural network. *Nigerian Journal of Technology*, 35 (2), 260-269.

[13] Khanlari, G. R., Heidari, M. Momeni, A. A., Abdilor, Y. (2012) Prediction of shear strength parameters of soils using artificial neural networks and multivariate regression. *Engineering Geology*, (131), 11-18.
[14] Kiran, S., and, Lal, B. (2015). ANN based prediction of shear strength of soil from their index properties. *Int. J. Earth Sci Eng*, (8), 2195-2202.
[15] Murata, N., Shuji Y., Shun-ichi A. (1994) Network information criterion-determining the number of hidden units for an artificial neural network model. *IEEE transactions on neural networks*. 5 (6), 865-872.
[16] Ogbuagu, F. U., Okeke, C. A. U. (2019). Geotechnical properties of lateritic soil from Nimo and Nteje areas of Anambra State, Southeastern Nigeria. In *IOP Conference Series: Materials Science and Engineering*, 640 (1), 012078. IOP Publishing.
[17] Ogbuchukwu, P. O., Okeke, O. C., Ahiarakwem, C. A., Ozotta, O. O. (2019). Geotechnical properties of expansive soils in Awka and environs, Southeastern Nigeria, in relation to engineering problems. *Int J Appl Sci Res*, (2),79-94.
[18] Ojuri, O. O. (2013). Predictive shear strength models for tropical lateritic soils. *Journal of Engineering*, (1) 595626.
[19] Osinubi K.J., Bello A.A. (2011). Soil water characteristics curves for Reddish brown tropical soil. *Electronic Journal of Geotech. Engineering*. (16)1-25.
[20] Paruelo, J., Fernando T. (1997). Prediction of functional characteristics of ecosystems: a comparison of artificial neural networks and regression models. *Ecological Modelling*, 98 (2-3), 173-186.
[21] Roy, S. and Dass, G. (2014). Statistical models for the prediction of shear strength parameters at Sirsa, India. *International Journal of Civil and Structural Engineering*. 4 (4), 483-498.
[22] Sorensen, K. K. and Okkels, N. (2013). Correlation between drained shear strength and plasticity index of undisturbed over consolidated clays. *Proceedings of the 18th International Conference on Soil Mechanics and Geotechnical Engineering, Paris,* 1-6.
[23] Tafari, T., Emer T. Q., Mohammed Y. (2021). Statistical analysis on shear strength parameter from index properties of fine-grained soils. *Journal of Engineering Research and Reports*, 20(4),15-28.