Journal Pre-proof

Support vector machine (SVM) prediction of coefficients of curvature and uniformity of hybrid cement modified unsaturated soil with NQF inclusion

Kennedy C. Onyelowe, Chilakala B. Mahesh, Bandela Srikanth, Chidobere Nwa-David, Jesuborn Obimba-Wogu, Jamshid Shakeri

PII: S2666-7908(21)00250-0

DOI: https://doi.org/10.1016/j.clet.2021.100290

Reference: CLET 100290

To appear in: *Cleaner Engineering and Technology*

Received Date: 29 January 2021

Revised Date: 23 August 2021

Accepted Date: 5 October 2021

Support vector machine (SVM) prediction of coefficients of curvature and uniformity of hybrid cement modified unsaturated soil with NQF inclusion

Kennedy C. ONYELOWE[*1], Chilakala B. MAHESH[2], Bandela SRIKANTH[3], Chidobere NWA-DAVID[4], Jesuborn OBIMBA-WOGU[4], & Jamshid SHAKERI[5]

[1]**Senior Lecturer**, Department of Civil Engineering, Michael Okpara University of Agriculture, Umudike Nigeria & Department of Civil and Mechanical Engineering, Kampala International University, Kampala, Uganda. * kennedychibuzor@kiu.ac.ug, konyelowe@mouau.edu.ng, konyelowe@gmail.com

[2]**Bachelors Researcher**, Department of Computer Science and Engineering, Vishnu Institute of Technology, Bhimavaram, India

[3]**Assistant Professor,** Department of Civil Engineering, Vishnu Institute of Technology, Bhimavaram, Andhrapradesh, India

[4]**Graduate Assistants,** Department of Civil Engineering, Michael Okpara University of Agriculture, Umudike, Nigeria

[5]**Doctoral Candidate**, Department of Mining Engineering, Hamedan University of Technology, Hamadan, Iran

Abstract
Support vector machine (SVM) with its feature known as the statistical risk minimization (SRM) has been employed in the prediction of coefficient of curvature and uniformity on unsaturated lateritic soil treated with composites of hybrid cement and nanostructured quarry fines. This feature utilized by SVM is the advantage it exercises over other intelligent learning techniques. This prediction has become necessary due to the time and equipment needs required to regularly conduct laboratory experiments prior to earthwork designs and construction. It is important to note that earthwork projects involving unsaturated soils pose threats of failure due to volume changes during seasonal cycles of wetting and drying especially for hydraulically bound environments and substructures. With an intelligent prediction, these design and construction worries are overcome. The soil used in the current work has been classified as an A-7-6 group soil with highly plastic consistency. Multiple experiments were conducted to generate multitude of datasets for the hybrid cement, nanostructured quarry fines, clay content and activity and frictional angle, which were selected as the independent variables for the model to predict coefficients of curvature and uniformity as the dependent variables. In order to correlate the relationship between the input and output parameters and as well validate the SVM model, detailed statistical analysis including Pearson's coefficient of correlation (R) and determination (R^2) and error analysis were conducted. Based upon the statistical analysis, the parameters were observed to have good correlation and determination ranging between 0.97 and 0.99. It was also observed that SVM outclassed MLR more in predicting Cu then it did in predicting Cc. Finally, sensitivity analysis was carried out and it was found that the Cc value is dependent mostly on frictional angle while Cu is dependent most on the NQF.
Keywords: Support Vector Machine (SVM); Coefficient of Curvature; Coefficient of Uniformity; Model Performance Evaluation; Sensitivity Analysis; Unsaturated Soil; Waste Base Binders.

1. Introduction

Laterite is described as both a soil and a rock type rich in iron and aluminium and is commonly considered to have formed in hot and wet tropical areas. Nearly all laterites are of rusty-red coloration, because of high iron oxide content. An unsaturated lateritic soil refers to a laterite soil with no or moisture content below the saturation point as presented by K. C. Onyelowe et al. (2021a); which presented shrinkage parameters of modified compacted clayey soil for sustainable earthworks and K. C. Onyelowe et al. (2021b) in a presentation on swelling potential of clayey soil modified with rice husk ash activated by calcination for pavement underlay by using plasticity index arithmetical method. Lateritic soils (red earth) are very valuable in civil engineering, as they are used in various purposes during construction, this can also be credited to their availability. The engineering properties of soils need to be determined before they can be used as construction material. This made imperative, the testing of soils, on which a foundation or superstructure is to be laid. These tests would determine its geotechnical suitability as a construction material which was presented in a work by Belayhun (2013), who studied some of the engineering properties of soil found in Asela town, Addis-Ababa. Soil is graded as either well graded or poorly graded according to the unified soil classification system and other standard systems of classification. Soil gradation is determined by analysing the results of a sieve analysis or a hydrometer analysis following the stipulated conditions presented by Holtz & Kovacs (1981). The grain size distribution curve produced from the result of the laboratory test on the soil is used to determine the gradation of the soil that is tested. In determining the gradation of the soil particles, the coefficient of curvature and uniformity becomes one of the important parameters to analyze. Gradation of a soil can also be determined by calculating the coefficient of uniformity, C_u, and the coefficient of curvature, C_c, of the soil and comparing the calculated values with published gradation limits in the book "Foundation Design Principles and Practices" of Coduto (2000). The Coefficient of Curvature, C_c and Uniformity are parameters estimated using the gradation curve from the results gotten from sieve analysis. The parameters are used in the classification of soils. In calculating the coefficients of uniformity and curvature it requires grain diameters. The grain diameter can be found for each per cent of the soil passing a particular sieve. This means that if 40% of the sample is retained on the No. 200 sieve then there is 60% passing the No. 200 sieve as presented by Holtz & Kovacs (1981) in a book "An Introduction to Geotechnical Engineering". C_c is a shape parameter and is given by the relation:

$$C_c = \frac{(D_{30})^2}{D_{10}\, X D_{60}} \tag{1}$$

Where, D_{60} is the grain diameter at 60% passing and D_{30} is the grain diameter at 30% passing, and D_{10} is the grain diameter at 10% passing.

The Coefficient of Uniformity, C_u is a crude shape parameter and is calculated using the equation below:

$$C_u = \frac{D_{60}}{D_{10}} \tag{2}$$

Where D_{60} is the grain diameter at 60% passing, and D_{10} is the grain diameter at 10% passing.

Different soft computing techniques in hydraulics and environmental engineering applications have been used by various researchers. Recently, among soft computing techniques, Support Vector Machines (SVMs) have attracted greater interest. In an investigation carried out by (Singh et al.,2009), three soft computing techniques (SVR, GP and M5P tree) and two empirical models (MLR and Kostiakov model) were used to estimate the infiltration rate of the soil with different water qualities. Results obtained concluded that the results of SVM were more suitable as compared to the GP and MLR and also gave better prediction than Kostiakov model. A similar study carried out by Tiwari et al. (2017) shows that SVM works well than the other techniques, when the generalized regression neural network, MLR, M5P model tree and SVM, were used to predict the cumulative infiltration of soil. Different artificial intelligence methods for forecasting monthly discharge time series have been compared by Wang et al. (2009). Their research findings revealed that SVM model was able to obtain better forecasting accuracy in terms of the various evaluation measures during the both training and validation phases. According to Wei et al. (2000), predictions exhibit good agreement with actual soil water content measurements. In their study, they applied support vector machines to soil water content predictions and compared the results to other time series prediction methods in purple hilly area. Their results showed that the SVMs predictors perform better for soil water forecasting than ANN models, compared with other predictors. A study carried out by Besalatpour et al. (2012) shows that the support vector machines achieved greater accuracy in predicting both soil shear strength and soil aggregate stability properties comparing to traditional multiple-linear regression. They found that the coefficient of correlation (R) between the measured and predicted soil shear strength values using the support vector machine model was 0.98 while it was 0.52 using the multiple-linear regression model. Furthermore, a lower mean square error value of 0.06 obtained using the support vector machine model in prediction of soil shear strength as compared to the multiple-linear regression model. The results revealed that utilization of optimized support vector machine approach with simulated annealing algorithm in developing soil property prediction functions could be a suitable alternative to commonly used regression methods. Akshaya (2015) developed models for prediction of CBR of an expansive soil stabilized with lime and quarry dust at different curing periods, using artificial neural network (ANN) and support vector machine (SVM). The accuracy of the predictive models has been compared. It is found that both the ANN and the SVM models are very accurate in prediction of CBR of stabilized expansive soil and the performances of the SVM models are found to be better than that of the ANN models. Twarakavi et al. (2009) investigated the Development of Pedo-transfer Functions for Estimation of Soil Hydraulic Parameters using Support Vector Machines. They successfully applied SVM methodology to develop Pedo-transfer functions (PTFs) that used different input predictors to estimate soil hydraulic parameters. These PTFs utilize some or all of the following predictors: sand, silt, and clay percentages, bulk density and retention data at one or two points (at the field capacity and the wilting point). Bootstrapping was performed to allow for the estimation of uncertainty in the predictions. It was observed that an increase in the number of predictors resulted in improved predictions by PTFs. It was also observed that the predictions by the SVM-based PTFs showed considerable improvement over ROSETTA. Lamorski et al. (2008) used the Soil Profiles Bank of Polish Mineral Soils that includes hydraulic properties for 806 soil samples taken from 290 soil profiles, to see whether using the SVM to develop Pedotransfer functions (PTFs) may have some advantages compared with the ANN. This database was repeatedly randomly split into training and testing data sets, and both SVMs and ANNs were trained and tested for each split with bulk density, sand and clay as input variables, and water contents at 11 soil water potentials as the output variables. Results showed that the three-parameter SVMs performed mostly better than or with the same accuracy as the eleven-parameter ANNs. Since the advantage of SVM was more pronounced, they concluded that using SVM as a tool to develop PTF is worthwhile. A study carried out by Hai & Binh (2021), support vector machines (SVM) was applied to predict the shear strength of the soil using moisture content, clay content, void ratio, plastic limit, liquid limit and specific gravity, as input variables. Their research findings revealed that SVM model performed well for the prediction of soil shear strength (R=0.9 to 0.95) and the moisture content, liquid limit and plastic limit were found as the three most affecting features to the prediction of soil shear strength. Manju (2020), carried out research work on applying several machine learning approaches namely M5 model tree, artificial neural networks, support vector machines, random forest and Gaussian processes, for prediction of unconfined compressive strength of stabilized pond ashes. The research input dataset consists of some parameters of which uniformity coefficient and coefficient of curvature were considered as well as maximum dry density, optimum moisture content, lime, lime sludge, curing period and 7-day soaked California bearing ratio.

The research revealed that GP models are in a position to predict the unconfined compressive strength of stabilized pond ashes with an excessive degree of accuracy.

However, prediction of unsaturated soil gradation parameters such as coefficients of curvature and uniformity using SVM is limited in literature. The objective of this paper is to develop SVM models to predict coefficients of curvature and uniformity of unsaturated soils.

2. Materials and Methods

2.1 Materials

The lateritic soil was collected from a borrow pit, which is a source of earthwork soil for decades now. Lumps were removed by soft tapping with a rubber pestle and sundried for three days in open air. The rice husk was collected from rice mills at Abakaliki, Nigeria where the people's main occupation is rice farming. The husk of rice is an agro-based solid waste disposed indiscriminately for lack of adequate disposal management program. It was combusted to generate rice husk ash (RHA) in a controlled incinerator to check the emission of CO_2 as a result in order to achieve environmentally friendly procedure. Furthermore, the rice husk ash was enhanced through activation by mixing it with 5% dosage of hydrated lime by weight of the ash in order to generate the composite binder called the hybrid cement (HC). Note, "*hydrated lime (Ca $(OH)_2$) is the quicklime combined chemically in water with 33% to 34% magnesium oxide (MgO), 46% to 48% of CaO, and 15% to 17% chemically combined with water. It is a crystal, non-flammable, odorless inorganic powder, which is soluble in water at ambient temperature. It has a melting point of 580°C, a boiling point of 2850°C, and a density of 2.21g/cm^3. Its density is less than that of quicklime (3.34g/cm^3) due to its more aqueous condition that creates pores in the structure of the solid. It is caustic with a pH of 12.8 and possesses pozzolanic characteristics, which makes it a good supplementary or alternative binder in civil engineering and earthworks".* Due to the properties outlined, hydrated lime serves as a good activation enhancer to ordinary ash materials utilized for soil stabilization. The HC is one of the major locally synthesized binders in this exercise with improved pozzolanic properties. In addition, quarry dust (QD) was also collected from a crushed rock site at Amasiri, Nigeria where aggregates are produced for construction works. The QD was further pulverized to fineness and sieved through a 200 nm sieve to generate the nanostructured quarry fines. This was used as a second binder in the stabilization phase of this work. The binding materials; HC and NQF were observed to meet the requirements for materials to be classified as pozzolanas (ASTM C618, 2019). Finally, the soil, HC and NQF were made ready and stored for use in the stabilization work.

2.2 Methods

2.2.1 Laboratory Methods

The requirements of the British Standard International (BSI) (BS1377, 1990) were observed in conducting general tests on the materials for the purpose of materials characterization and classification. Under the above conditions, the particle size analysis, compaction, Atterberg limits, specific gravity, and the angle of internal friction were conducted. In order to determine the oxides composition of the binder materials (RHA, HC and NQF), the XRF was carried out in accordance with the ASTM E1621-13 (2013), which provided the standard guide for elemental analysis by wavelength dispersion x-ray fluorescence spectrometry, and in order to achieve reliable and precise results, the samples were prepared and mixed to high homogeneity and the results were observed and recorded. Further, the treatment exercise was conducted in accordance with the requirements of the BS1924 (1990) and multiple data points were generated for varying dosages of HC and NQF between 0% and 12%.

2.2.2 Model Development

The multiple laboratory exercises gave rise to multiple datasets as presented in Appendix A, which were collected for the five independent variables otherwise called the predictors; HC-hybrid cement, NQF-nanostructured quarry fines, C-clay content, A_c-Clay activity, and $\emptyset^{\circ}$-frictional angle and the two dependent variables otherwise called the targets; C_c-coefficient of curvature, and C_u-coefficient of uniformity and produced a mathematical functional of functions as follows;

$$C_c \;\&\; C_u = f\,(HC, NQF, C, A_c, \emptyset^{\circ}) \tag{3}$$

The datasets of the input and outputs variables generated from multiple experiments were deployed to the multi linear regression (baseline regression), and support vector machine (SVM) to predict the outputs parameters as functions of the inputs.

2.2.3 Multi Linear Pearson's Regression

In relation to previous literatures, the current work employed Pearson's correlation coefficients presented in Figures 1-3 to measure the linear correlation between the model parameters as presented by Adler and Parmryd (2010); which studied quantifying colocalization by correlation: the Pearson correlation coefficient is superior to the Mander's overlap coefficient, Benesty et al. (2008); which studied on the importance of the Pearson correlation coefficient in noise reduction and Benesty et al. (2009), which presented Pearson correlation coefficient, in Noise reduction in speech processing, respectively with remarkable results. The use of HC and NQF influenced the values of Cc and Cu almost in a similar manner, which increased the values of the output parameters. It can be observed in the analysis in Figures 1-3 that the values of Cc and Cu depicted the strong positive linear relationship with the use of HC and NQF. The correlation regression values between the outputs and the inputs were observed to be above 0.97. This implied that the parameters have direct influence on the model outcome.

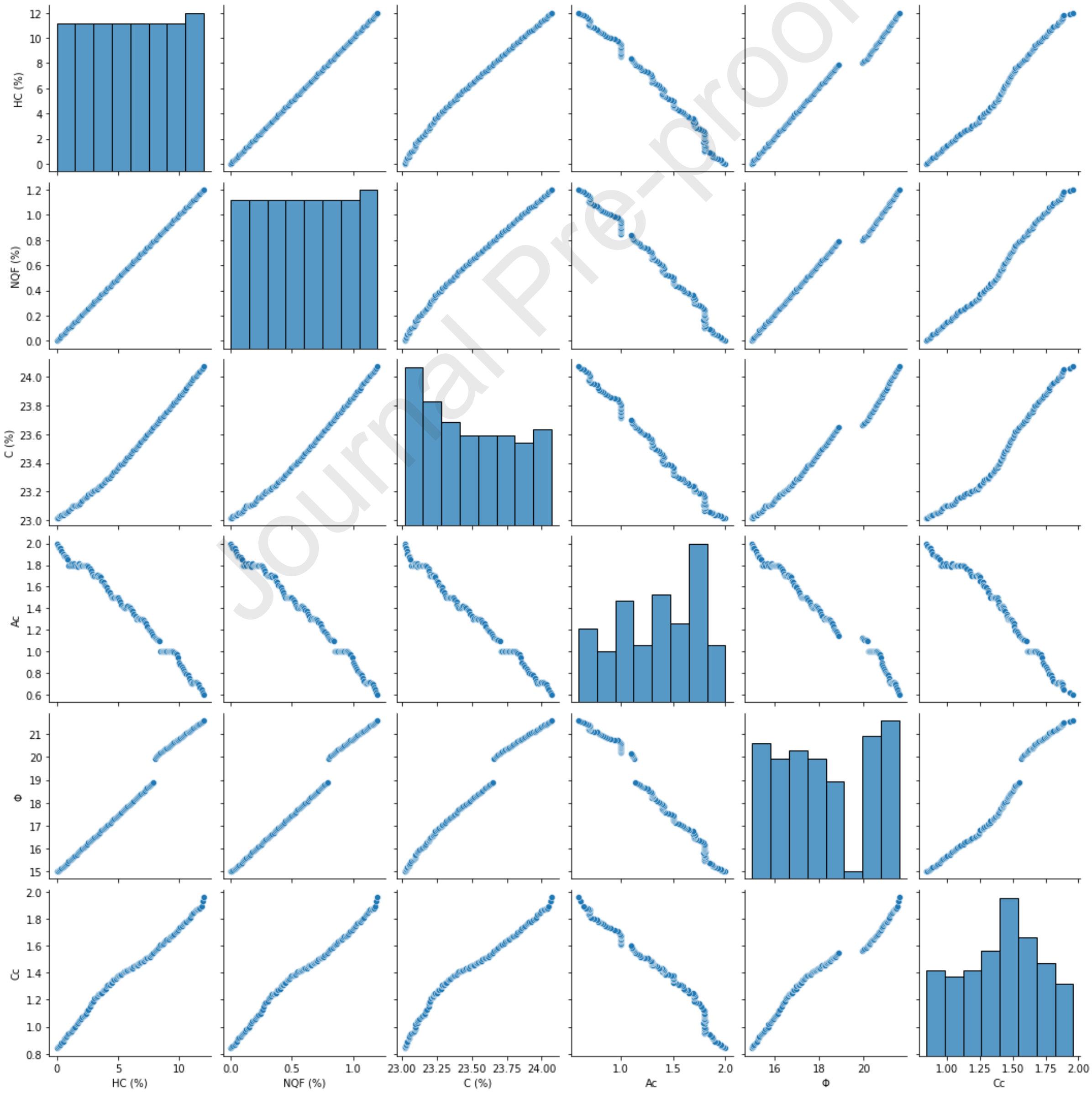

Figure 1: Pearson's schematic linear correlation regression of presented variables with coefficient of curvature

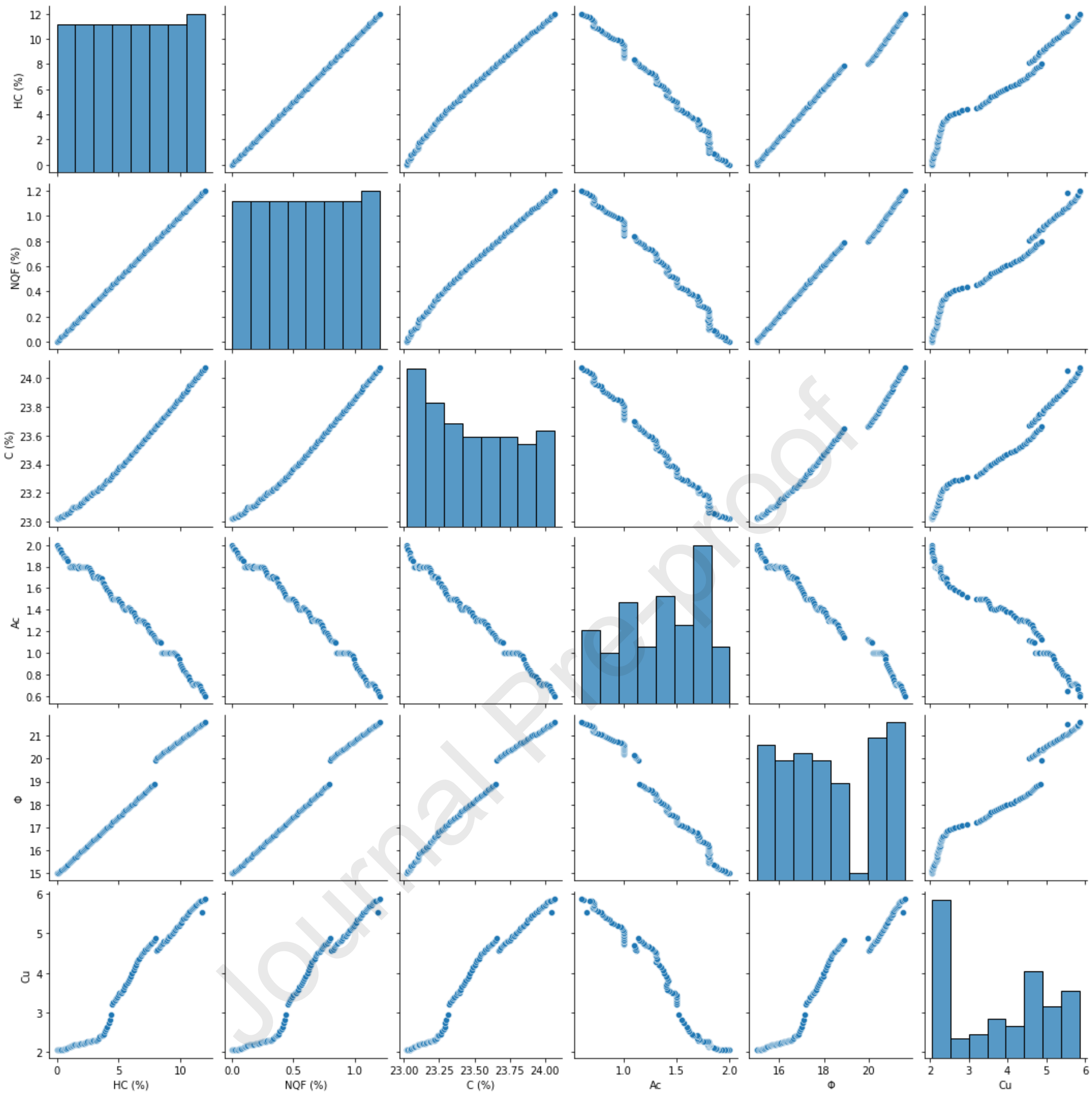

Figure 2: Pearson's schematic linear correlation regression of presented variables with coefficient of uniformity

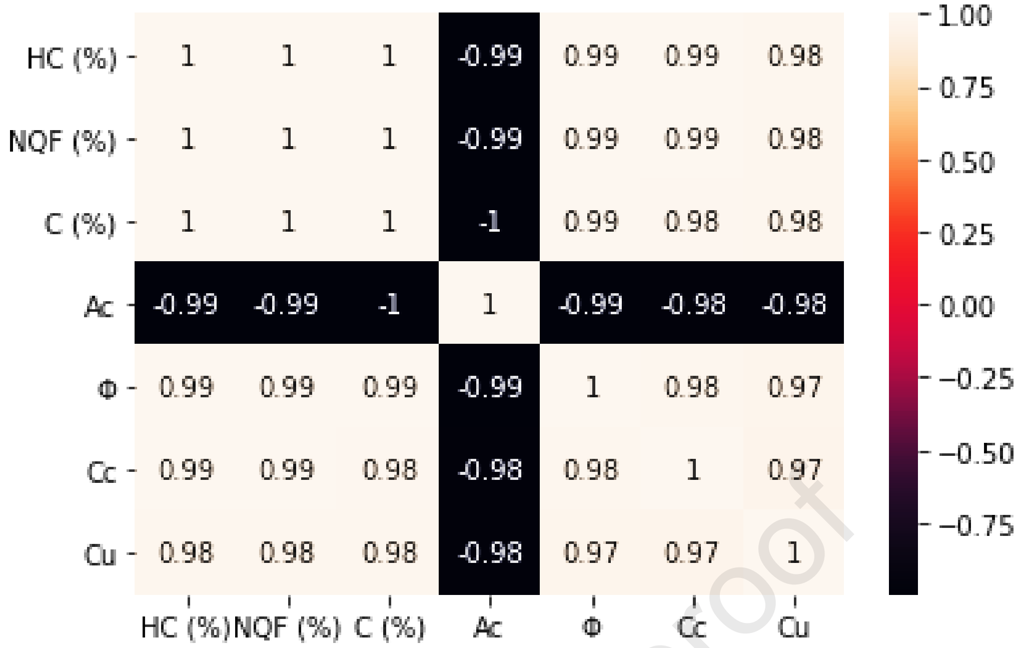

Figure 3: Pearson's linear correlation of presented input and output variables

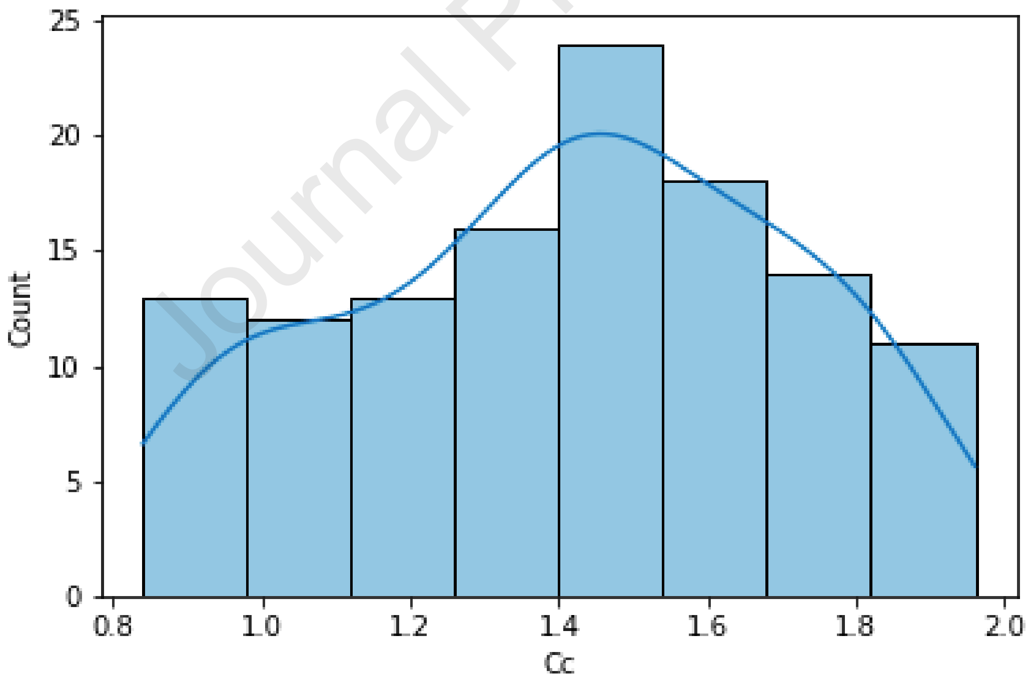

Figure 4: Frequency histogram of the variables against the coefficient of curvature

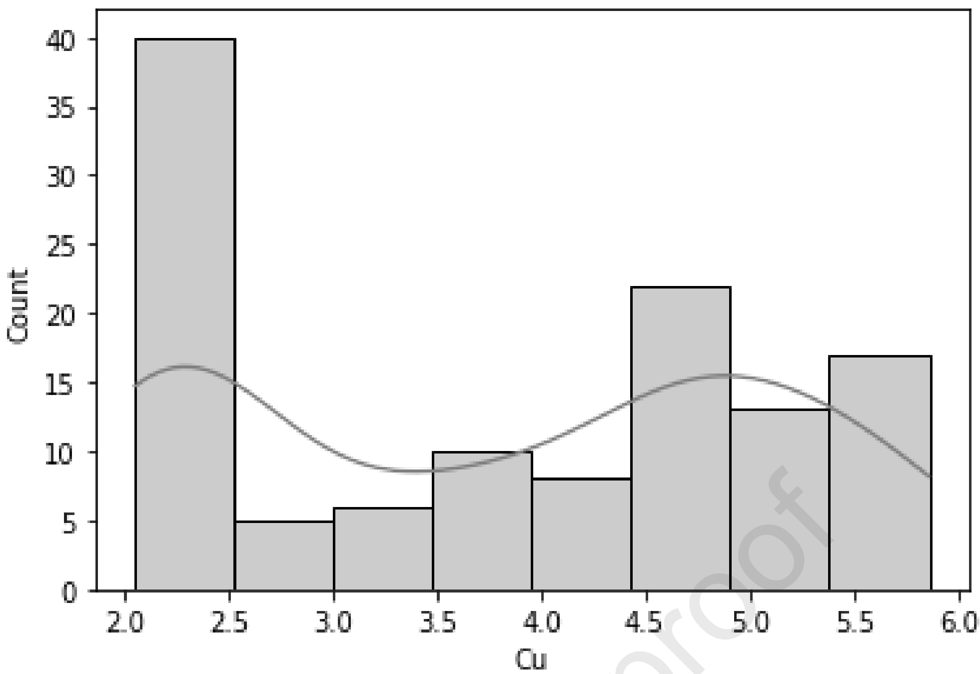

Figure 5: Frequency histogram of the variables against the coefficient of uniformity

3. Results and Discussion

3.1 Basic Materials Properties

The particle distribution of the soil is presented in Figure 6. The soil was observed to be a poorly graded soil with coefficients of curvature and uniformity of 0.84 and 2.05 respectively and classified as an A-7-6 group according to AASHTO method of soil classification. The soil was discovered to be highly plastic with high clay content. The angle of internal friction of the soil was observed to be 15°, with clay content and clay activity of 23.02% and 2 respectively (see Appendix A). It can be observed from Table 1, that the RHA, HC and NQF showed characteristics of pozzolanas with the combined oxide compositions of SiO_2, Al_2O_3 and Fe_2O_3 as more than 70%, according to ASTM C618 (2019), which provides the standard specification for pozzolanas. The behaviour was important in the stabilization operation where these materials in single and composite forms were blended with soil to trigger hydration, pozzolanic, calcination, and cation exchange reactions culminating to the behavioural changes observed in the treated soil with respected to the measured parameters.

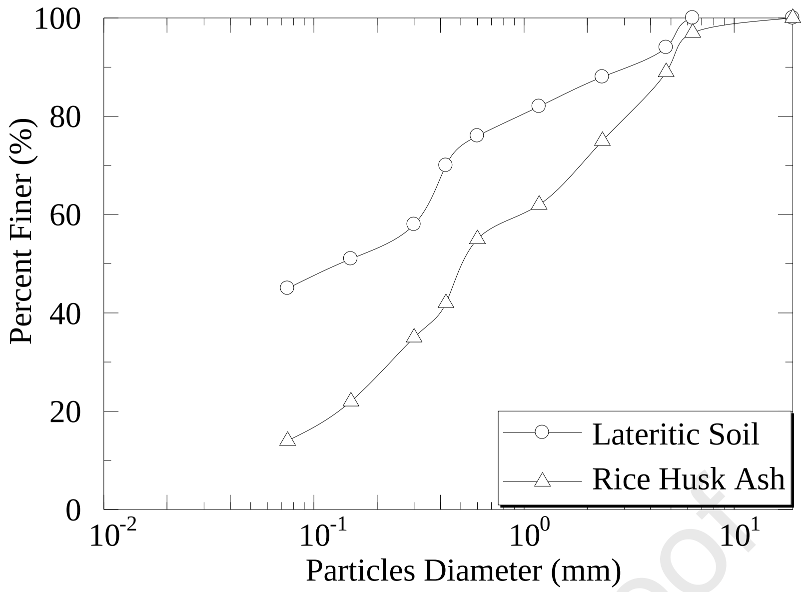

Figure 6: Particle size distribution curve of clayey soil and rice husk ash

Table 1. Chemical oxide composition of the additive materials

Materials	Oxides composition (content by weight, %)												
	SiO_2	Al_2O_3	CaO	Fe_2O_3	MgO	K_2O	Na_2O	TiO_2	LOI	P_2O_5	SO_3	IR	free CaO
clay soil	12.45	18.09	2.30	10.66	4.89	12.10	34.33	0.07	-	5.11	-	-	-
rice husk ash	56.48	22.72	5.56	3.77	4.65	2.76	0.01	3.17	0.88	-	-	-	-
HC	59.12	25.3	6.3	4.23	2.5	1.21	-	1.34	-	-	-	-	-
NQF	62.48	18.72	4.83	6.54	2.56	3.18	-	0.29	1.01	-	-	-	-

**IR is insoluble residue; LOI is loss on ignition,*

3.2 Support Vector Model Presentation and Validation

Figures 7-10 show the SVM models and the validation of the outcomes between the measured and predicted values. This was necessary to select the set of variables that would produce the best classification of parameters for the predicted Cc and Cu, which is achieved by adopting the hyper-plane technique. The hyper-planes are the unique features of the SVM model technique which separates the data into two parts and lies at the center. The support vectors are the two lines parallel and equidistant to the hyper-plane separated by a distance called the marginal distance. The data points which lie on the support vectors are called the supports and these points lie near to the hyper-plane. From the models as presented in the SVM presentation, there are four points, two on either sides of the hyper-plane in the predicted Cc model in Figure 9 while there is one vector point on the predicted values side and two vector points on the measured values side of the Cu model. Apart from this behaviour, the scatter orientation of the data points with respect to the hyper-plane orientation shows closer agreement between the predicted and measured values of the SVM model.

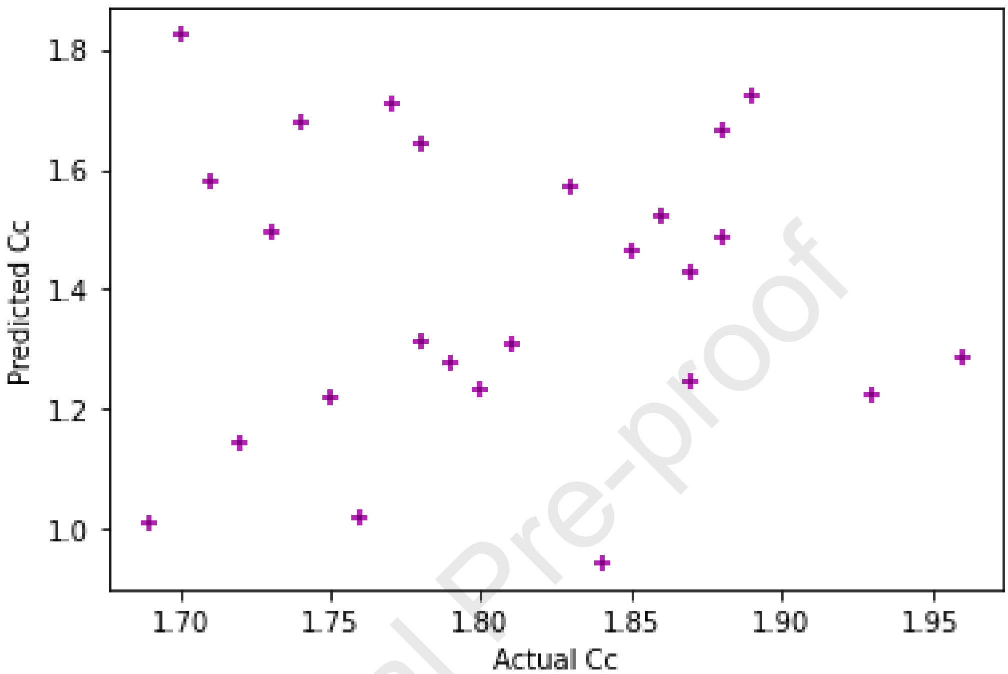

Figure 7: Validation of predicted coefficient of curvature

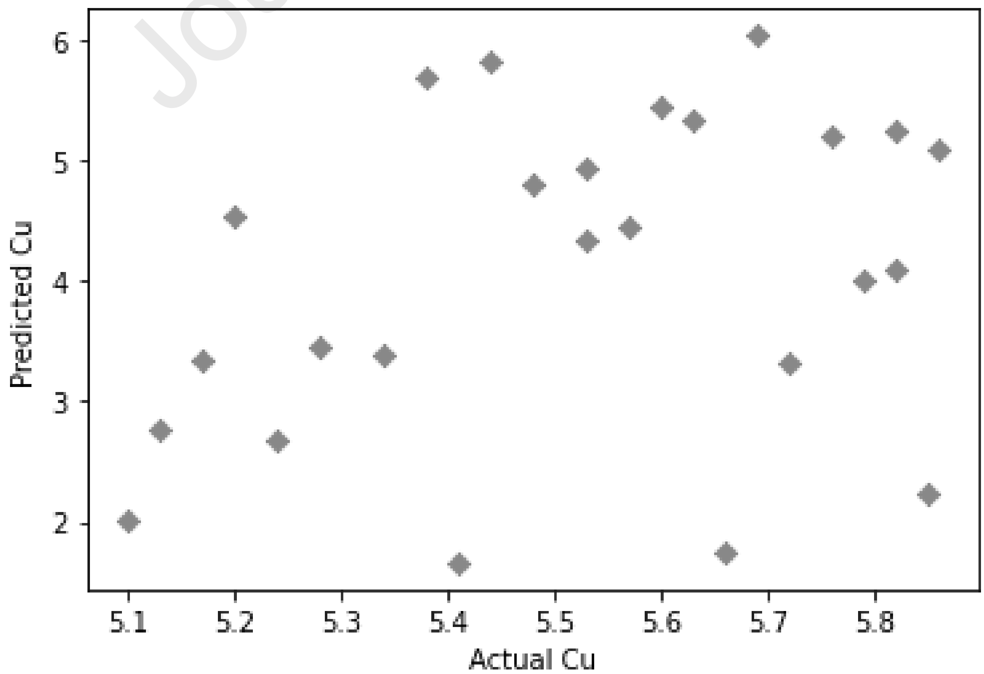

Figure 8: Validation of predicted coefficient of uniformity

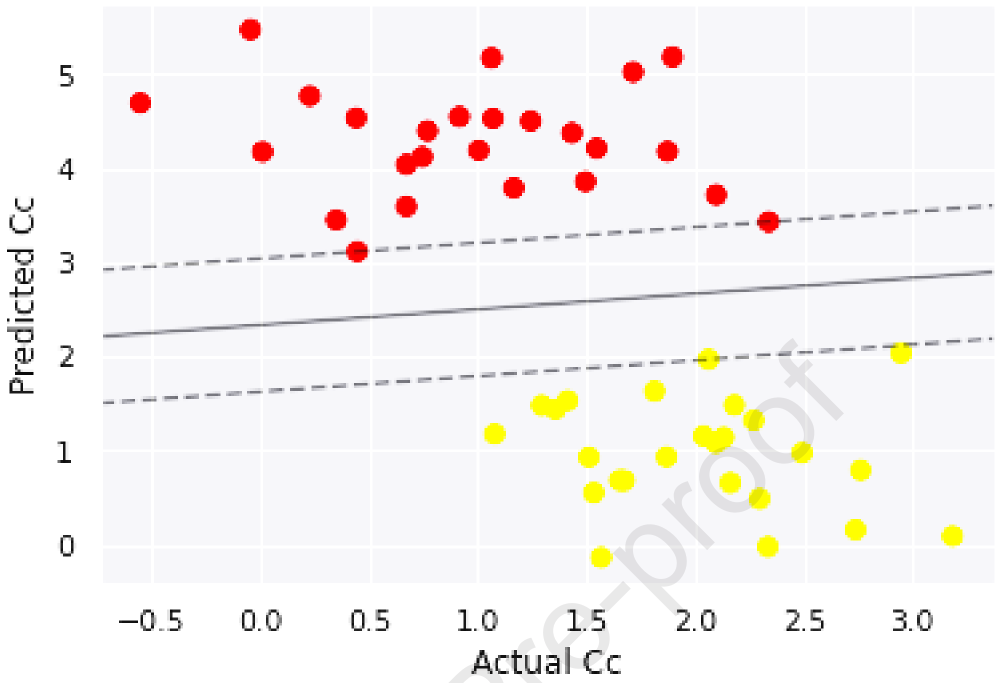

Figure 9: The support vector model of coefficient of curvature

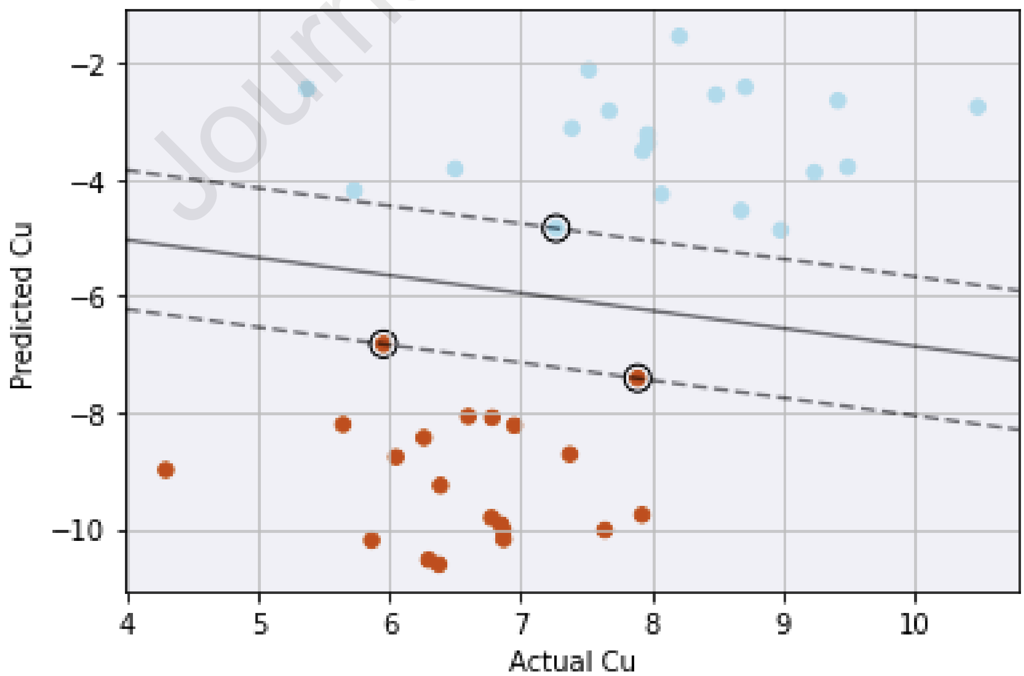

Figure 10: The support vector model of coefficient of uniformity

3.3 Performance Evaluation of Models

The evaluation of the performance outcome of the model has been presented in Table 2 and Figures 11 & 12. In order to achieve this, a variety of performance indices have been estimated, which include, mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2) to evaluate the performance of developed Cc and Cu models. The following equations (Eqs 4-7) were used to calculate these performance indices.

$$\mathrm{MSE} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}}(e_i - p_i)^2 \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(e_i - p_i)^2}{n}} \quad (5)$$

$$MAE = \frac{\sum_{i=1}^{n}|e_i - p_i|}{n} \quad (6)$$

$$R = \frac{\sum_{i=1}^{n}(e_i - \bar{e}_i)(p_i - \bar{p}_i)}{\sqrt{\sum_{i=1}^{n}(e_i - \bar{e}_i)^2 \sum_{i=1}^{n}(p_i - \bar{p}_i)^2}} \quad (7)$$

where e_i and p_i are the i number of experimental and predicted outputs, respectively; $\bar{e}_i$ and $\bar{p}_i$ are the average values of the experimental and predicted output values, respectively, and n is total the number of samples.

The RMSE errors are squared implying that relatively a much larger weight is assigned to the larger errors. High R^2 values and low RMSE values achieve high degree of accuracy, which agrees with the results of Gandomi and Roke (2015); which assessed the ability of ANN and GP as predictive tools and those of Onyelowe et al. (2021a); which proposed predictive model results on Cc and Cu of treated unsaturated soil using the adaptive neuro-fuzzy inference system and Onyelowe and Shakeri (2021); which also researched on the prediction of Cc and Cu of treated unsaturated soil using ANN, GEP and LMR. The proposed models indicate that the MAE, MSE, and RMSE values are significantly lower for Cc in MLR model than they are in SVM but R^2 is higher. However, the values of R^2 by SVM model are observed to be more than 0.97, which is good performance accuracy. It can be observed that SVM outclassed MLR more in predicting Cu than it did in predicting Cc.

Table 2. Performance evaluation of predicted models using the error indices

Models Performance Codes	SVM		MLR	
	Cc	Cu	Cc	Cu
Mean absolute error (MAE)	0.0347	0.1508	0.0161	0.1967

Mean squared error (MSE)	0.0019	0.0469	0.0004	0.0587
Root mean squared error (RMSE)	0.0436	0.2165	0.0197	0.2422
R^2 Score	0.9707	0.9711	0.9954	0.9580

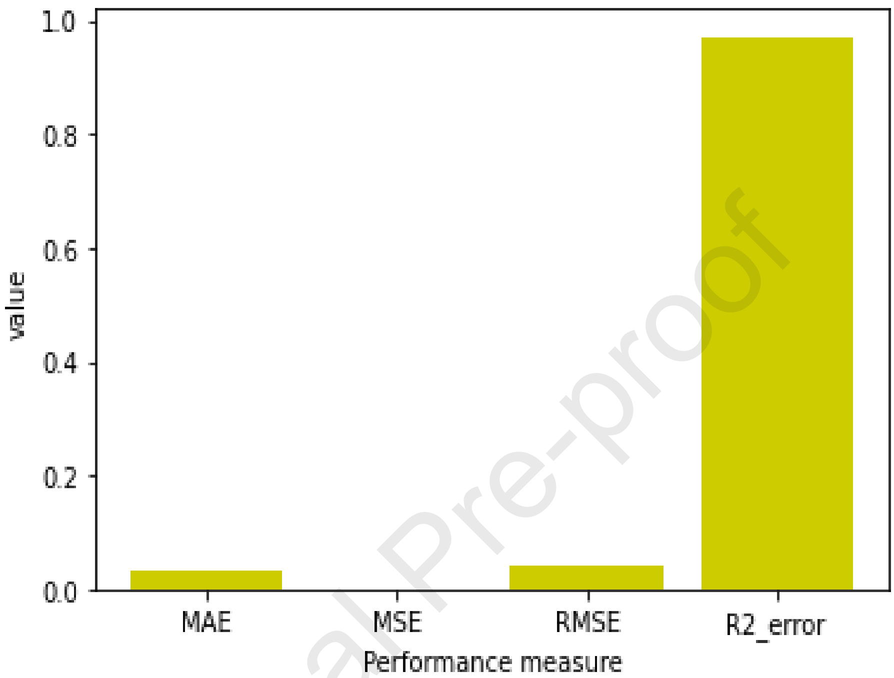

Figure 11: Performance evaluation histogram for the coefficient of curvature model

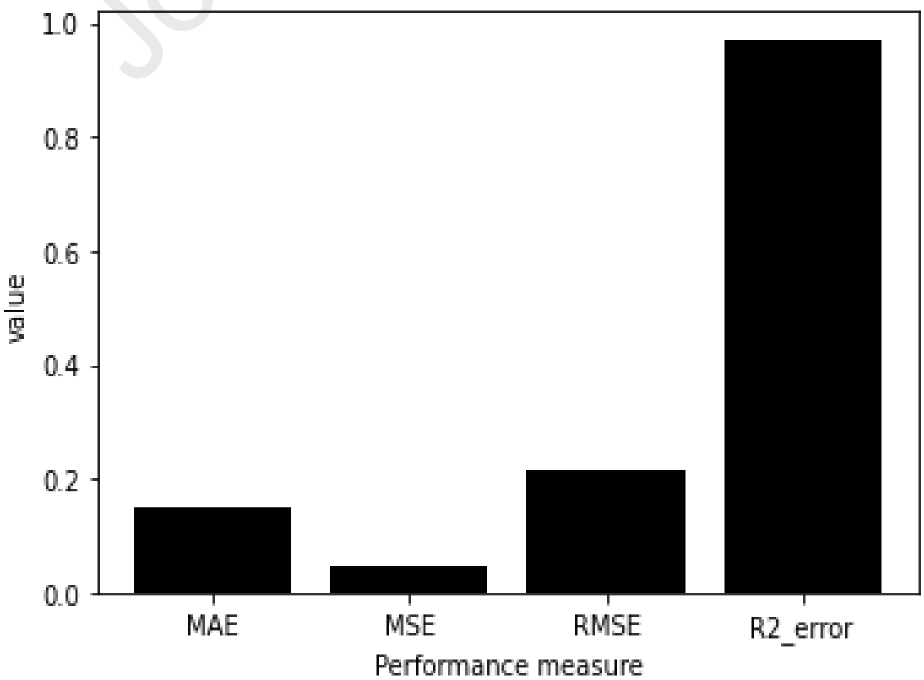

Figure 12: Performance evaluation histogram for the coefficient of uniformity model

3.4 Model Parametric and Sensitivity Analysis

The contributions of the input parameters to the behaviours of Cc and Cu have been depicted in Figures 13 & 14. It can be observed that the magnitude of input parameters linearly increase the output parameters. Hence the results of this study are in line with the literature as presented by Yadu et al. (2011), which researched on the comparison of fly ash and rice husk ash stabilized black cotton soil. Furthermore, a useful concept has been proposed to identify the significance of each independent factor (the input parameter) on the “dependent factors” (the output parameters) as presented in Figures 15 and 16. This is to hierarchically recognize the most sensitive factors influencing the Cc and Cu. In order to achieve this focus, sensitivity analysis was conducted to observe the relative influence of the parameter in the model system by the cosine amplitude method of Yang and Zang (1997) which provided reliable findings on hierarchical analysis for rock engineering using artificial neural networks. To apply this method, all of the data pairs were expressed in common X-space. And to undertake this technique, all data pairs should be utilized to build a data array X as proposed by previous research results by Faradonbeh et al. (2016); which worked on prediction of ground vibration due to quarry blasting based on gene expression programming: a new model for peak particle velocity prediction, Khandelwal et al. (2016); which researched on new model based on gene expression programming to estimate air flow in a single rock joint, and Sayevand et al. (2018); which worked on the development of imperialist competitive algorithm in predicting the particle size distribution after mine blasting, as well as Onyelowe et al. (2021b); which presented the sensitivity analysis of ANN in predicting treated soil erodibility and Onyelowe et al. (2021c); which successfully presented sensitivity determination of volumetric stability and erodibility of treated unsaturated soil using triple AI-based genetic programming, artificial neural network and evolutionary polynomial regression techniques. The sensitivity analysis outcome shows that frictional angle has greatest effect on the performance of the Cc SVM model while NQF showed the greatest effect on the performance of the Cu SVM model.

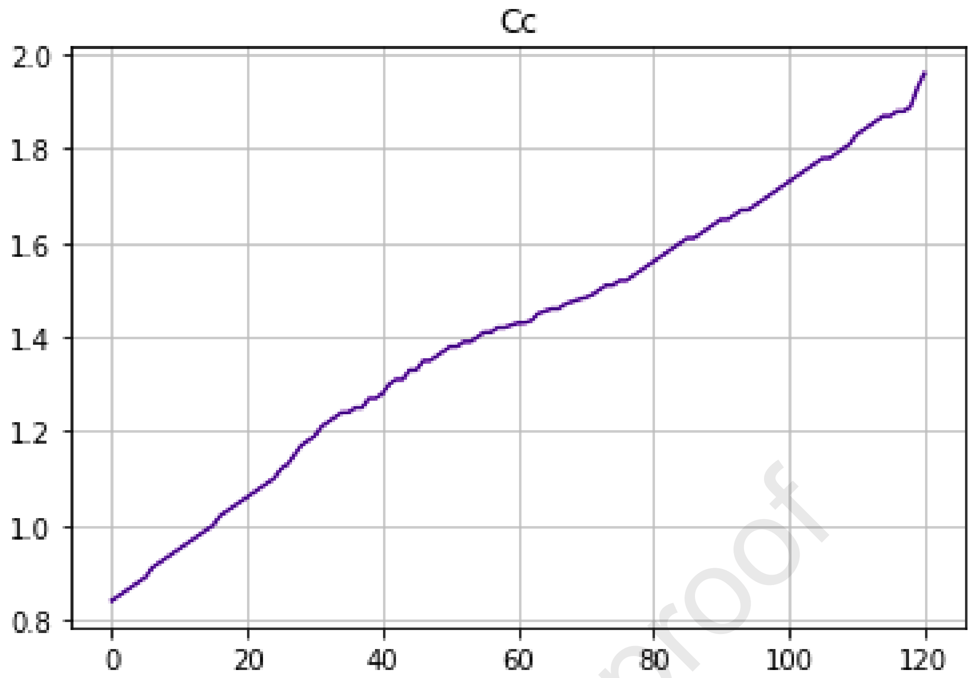

Figure 13: Parametric analysis of the coefficient of curvature model

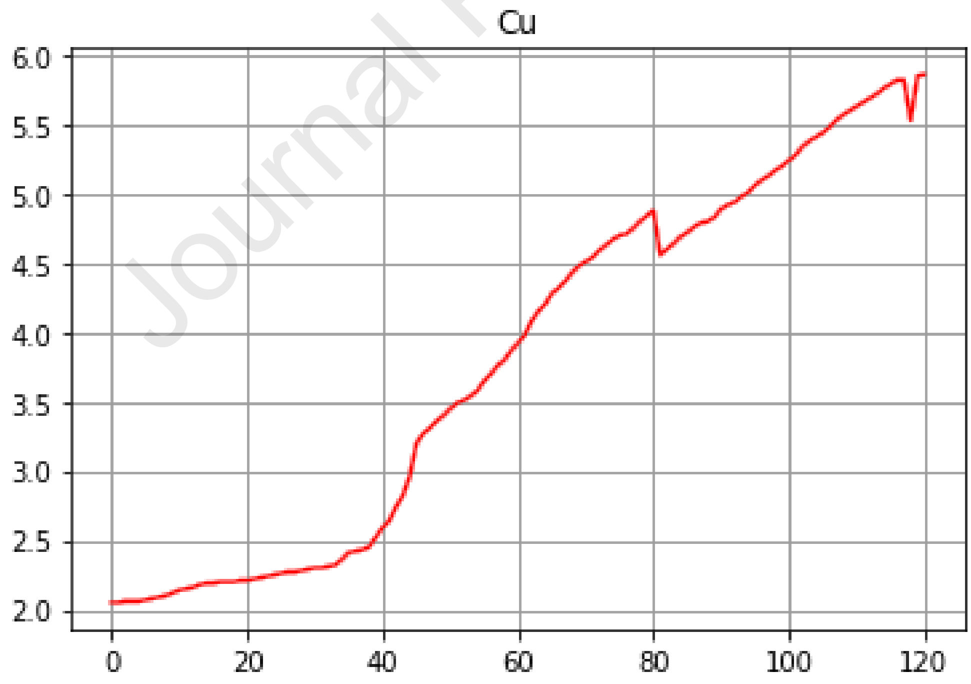

Figure 14: Parametric analysis of the coefficient of uniformity model

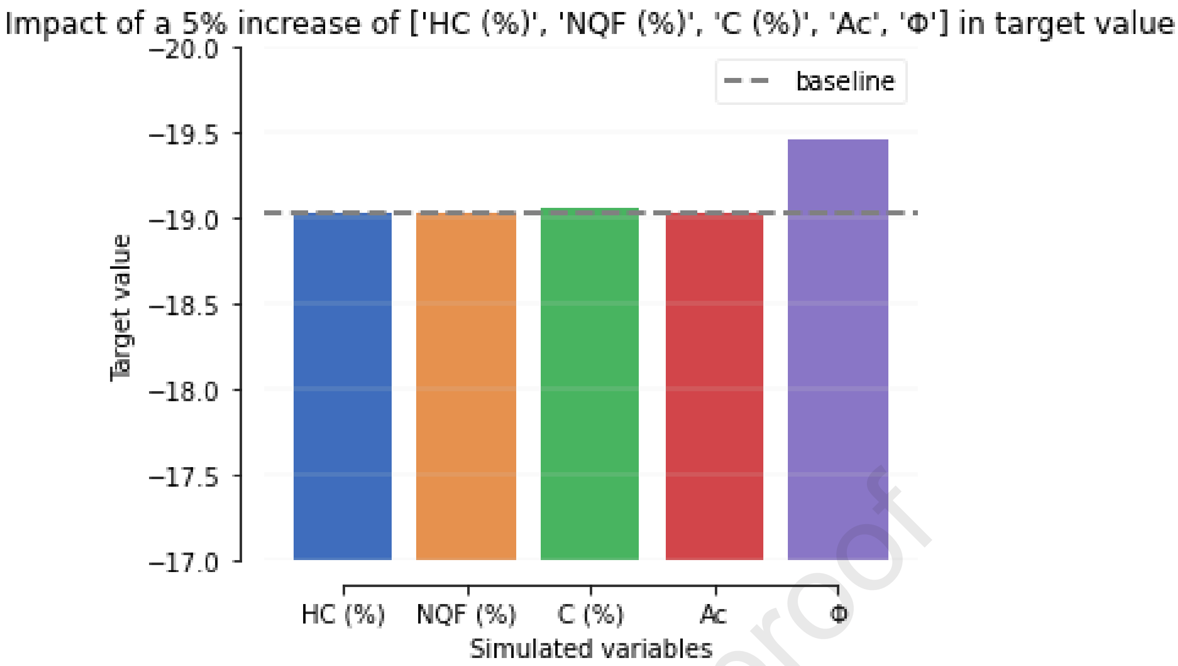

Figure 15: Sensitivity analysis of the variables against the coefficient of curvature model

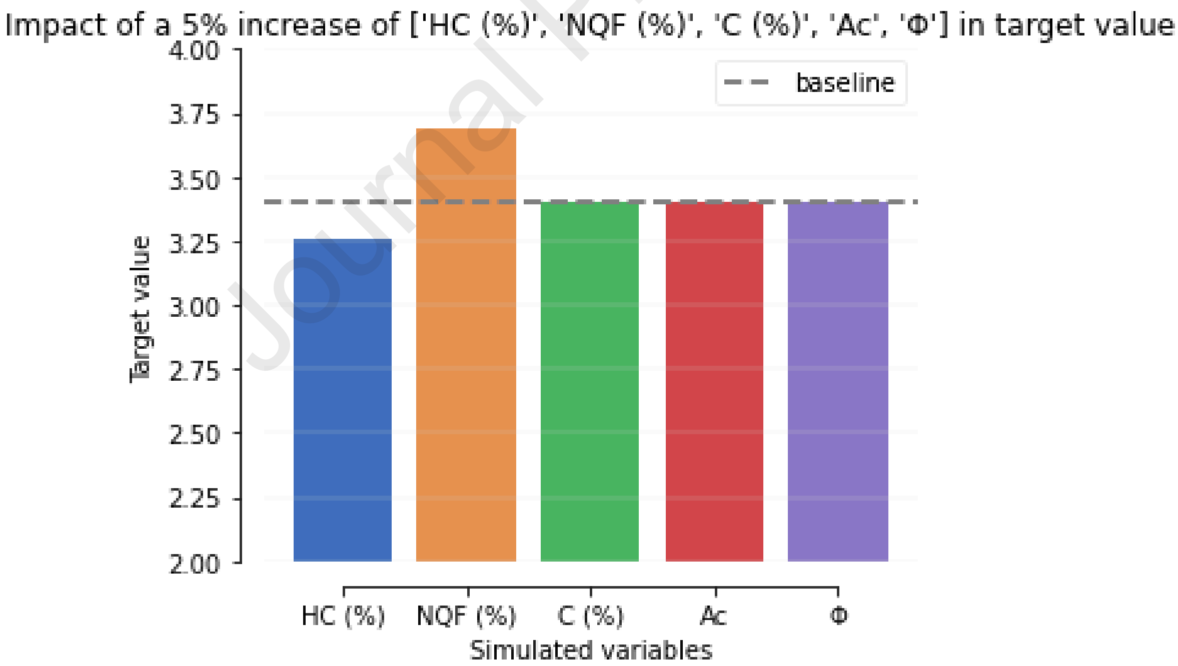

Figure 16: Sensitivity analysis of the variables against the coefficient of uniformity model

4. Conclusions

The support vector machine which is a supervised learning technique has been adopted in the prediction of Cc and Cu of an unsaturated soil treated with composites of hybrid cement (HC) and nanostructured quarry fines (NQF). It has demonstrated high performance in the prediction procedure and the following can be concluded;

- The geotechnics behavior of the soil mixed with HC and NQF in varying dosages showed a linear increase in the selected geotechnical properties of the unsaturated soil which were clay content, clay activity, friction angle, which were the predictors and of course the coefficients of curvature and uniformity which were the targets.
- The MLR model showed that the selected parameters were in goof correlation with each other and the outputs parameters.
- The SVM model generated hyper-planes that optimally separated the vector points after the input data have been transformed mathematically into a space of high-dimension.
- The SVM model outclassed the MLR more in the prediction of Cu than it did in the Cc prediction.
- The sensitivity analysis showed that the frictional angle has the greatest influence on the behavior of the predicted Cc while NQF was observed to have the greatest influence on the behavior of the predicted Cu model.
- Generally, SVM showed that it has the capability to forecast the gradation properties of unsaturated soil treated with local binder materials.

Conflict of Interest Statement

The authors declare no conflicting interests in this work

Notations

HC	Hybrid cement (%)
NQF	Percentage nanostructured quarry fines (%)
C	Clay content (%)
A_C	Clay activity
C_C	Coefficient of curvature
C_u	Coefficient of uniformity
$\emptyset^{\circ}$	Frictional angle (°)
SVM	Support vector machine

MLR	Multilinear regression
MAE	Mean absolute error
MSE	Mean squared error
RMSE	Root mean squared error
R^2	Coefficient of determination
SRM	Statistical risk minimization
RHA	Rice husk ash
SiO_2	Silica
Al_2O_3	Alumina
Fe_2O_3	Ferrite
CaO	Calcium oxide
K_2O	Potassium oxide
Na_2O	Sodium oxide
Ti_2O	Titanium oxide

References

Adler, J. and I. Parmryd, Quantifying colocalization by correlation: the Pearson correlation coefficient is superior to the Mander's overlap coefficient. Cytometry Part A, 2010. **77**(8): p. 733-742.

American Standard for Testing and Materials (ASTM) C618, Specification for Pozzolanas. ASTM International, Philadelphia, 2019, USA.

American Standard for Testing and Materials (ASTM) E1621-13 (2013). Standard guide for elemental analysis by wavelength dispersion x-ray fluorescence spectrometry, ASTM International, West Conshohocken, PA. DOI: 10.1520/E1621-13

Akshaya K. S., (2015). Prediction of California Bearing Ratio of a Stabilized Expansive Soil using Artificial Neural Network and Support Vector Machine. EJGE Vol. 20(3).p 981-991

Belayhun, Yilma (2013) study of some of the engineering properties of soil found in Asela town. An unpublished M.Eng thesis submitted to the school of graduate studies, Addis Ababa.

Benesty, J., J. Chen, and Y. Huang, On the importance of the Pearson correlation coefficient in noise reduction. IEEE Transactions on Audio, Speech, and Language Processing, 2008. **16**(4): p. 757-765.

Benesty, J., et al., Pearson correlation coefficient, in Noise reduction in speech processing. 2009, Springer. p. 1-4.

Besalatpour A.,Hajabbasi M. A.,Ayoubi S., Gharipour A., & Jazi A.Y.,(2012). Prediction of soil physical properties by optimized support vector machines. International Agrophysics, Vol. 26, p.109-115 Doi: 10.2478/v10247-012-0017-7

BS 1377 - 2, 3, (1990). Methods of Testing Soils for Civil Engineering Purposes, British Standard Institute, London.

BS 1924, (1990). Methods of Tests for Stabilized Soil, British Standard Institute, London.

Coduto, Donald P. (2000). Foundation Design Principles and Practices (2nd Edition). 02nd ed. Upper Saddle River: Prentice Hall, 2000. Print. ISBN 0-13-589706-8

Faradonbeh, R. S., Armaghani, D. J., Abd Majid, M. Z., Tahir, M. M., Murlidhar, B. R., Monjezi, M., & Wong, H. M. (2016). Prediction of ground vibration due to quarry blasting based on gene expression programming: a new model for peak particle velocity prediction. International journal of environmental science and technology, 13(6), 1453-1464. DOI: 10.1007/s13762-016-0979-2

Gandomi, A. H. & Roke, D. A. 2015. Assessment of artificial neural network and genetic programming as predictive tools. Advances in Engineering Software, 88, 63-72. https://doi.org/10.1016/j.advengsoft.2015.05.007

Hai B. L., & Binh T. P., 2021. Prediction of Shear Strength of Soil Using Direct Shear Test and Support Vector Machine Model. The Open Construction & Building Technology Journal. Vol 15. **DOI:** 10.2174/1874836802014010041

Holtz, R. & Kovacs, W. (1981). An Introduction to Geotechnical Engineering, Prentice-Hall, Inc. ISBN 0-13-484394-0

K. C. Onyelowe, M. E. Onyia, D. Bui Van, A. A. Firoozi, O. A. Uche, S. Kumari, I. Oyagbola, T. Amhadi and L. Dao-Phuc (2021a). Shrinkage Parameters of Modified Compacted Clayey Soil for Sustainable Earthworks. Jurnal Kejuruteraan 33(1): 137-144 https://doi.org/10.17576/jkukm-2020-33(1)-13

K. C. Onyelowe, M. E. Onyia, D. Nguyen-Thi, D. Bui Van, E. Onukwugha, H. Baykara, I. I. Obianyo, L. Dao-Phuc, and H. U. Ugwu (2021b). Swelling Potential of Clayey Soil Modified with Rice Husk Ash Activated by Calcination for Pavement Underlay by Plasticity Index Method (PIM). Advances in Materials Science and Engineering, Vol. 2021, Article ID 6688519, https://doi.org/10.1155/2021/6688519

Khandelwal, M., Armaghani, D. J., Faradonbeh, R. S., Ranjith, P. G., & Ghoraba, S. (2016). A new model based on gene expression programming to estimate air flow in a single rock joint. Environmental Earth Sciences, 75(9), 739. https://doi.org/10.1007/s12665-016-5524-6

Lamorski, K., Pachepsky, Y., Sławiński, C., & Walczak, R. T., (2008). Using Support Vector Machines to Develop Pedotransfer Functions for Water Retention of Soils in Poland. Soil Science Society of America Journal, Vol 72(5). p.1243-1247 DOI:10.2136/sssaj2007.0280N

Manju S., 2020. Applying Several Machine Learning Approaches for Prediction of Unconfined Compressive Strength of Stabilized Pond Ashes. Neural Computing and Applications. Vol. 32(6), p. 9019-9028. DOI: 10.1007/s00521-019-04411-6

Onyelowe, K. C., J. Shakeri, H. A. Khoshalann, A. B. Salahudeen, E. E. Arinze, H. U. Ugwu (2021a). Application of ANFIS hybrids to predict coefficients of curvature and uniformity of treated unsaturated lateritic soil for sustainable earthworks. Cleaner Materials, Vol. 1(1), pp. 1-21. https://doi.org/10.1016/j.clema.2021.100005

Onyelowe, K. C., and Shakeri, J., (2021). Intelligent prediction of coefficients of curvature and uniformity of hybrid cement modified unsaturated soil with NQF inclusion, Cleaner Engineering and Technology, Vol. 4, (2021), https://doi.org/10.1016/j.clet.2021.100152.

Onyelowe, K. C., Gnananandarao, T., Nwa-David, C. (2021b). Sensitivity analysis and prediction of erodibility of treated unsaturated soil modified with nanostructured fines of quarry dust using novel artificial neural network. Nanotechnology for Environmental Engineering, 6 (2), 37. https://doi.org/10.1007/s41204-021-00131-2

Onyelowe, K. C., A. M. Ebid, L. I. Nwobia (2021c). Predictive models of volumetric stability (durability) and erodibility of lateritic soil treated with different nanotextured bio-ashes with application of loss of strength on immersion; GP, ANN and EPR performance study. Cleaner Materials, Vol. 1(1), pp. 1-10. https://doi.org/10.1016/j.clema.2021.100006

Sayevand, K., Arab, H., & Golzar, S. B. (2018). Development of imperialist competitive algorithm in predicting the particle size distribution after mine blasting. Engineering with Computers, 34(2), 329-338.

Singh B., Sihag P., & Deswal S., (2019). Modelling of the impact of water quality on the infiltration rate of the soil. Applied Water Science Vol.9(15) p.1-9 https://doi.org/10.1007/s13201-019-0892-1

Tenpe, A., & Patel, A. (2020). Utilization of support vector models and gene expression programming for soil strength modeling. Arabian Journal of Science and Engineering. DOI: 10.1007/s13369-020-04441-6

Tiwari N.K, Sihag P., Ranjan S., (2017) Modeling of infiltration of soil using adaptive neuro-fuzzy inference system (ANFIS). J Eng Technol Educ Vol. 11(1), p.13–21

Twarakavi N. K. C., Simunek J., & Schaap M. G., (2009). Development of Pedotransfer functions for estimation of soil hydraulic parameters using support vector machines. Soil Sci. Soc. Am. J., 73, 1443-1452. Doi:10.2136/sssaj2008.0021

Wang, W. C., Chau, K. W., Cheng, C. T., & Qiu, L., (2009). A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. J. Hydrol., 374, p. 294-306. DOI: 10.1016/j.jhydrol.2009.06.019

Wei, W., Xuan, W., Deti, X., & Hongbin, L. (2003). Soil Water Content Forecasting by Support Vector Machine in Purple Hilly Region. Journal of Soil and Water Conservation, P.224-230

Yadu, L., Tripathi, R. J., & Singh, D. (2011). Comparison of fly ash and rice husk ash stabilized black cotton soil. International Journal of Earth Sciences and Engineering, 4 (6), p. 42-45. **DOI:** 10.12691/materials-2-3-2

Yang, Y., & Zhang, Q. (1997). A hierarchical analysis for rock engineering using artificial neural networks. Rock Mechanics and Rock Engineering, 30(4), 207-222. DOI:10.1007/BF01045717