Scientific
Research
Publishing

Prediction of Strength Properties of Soft Soil Considering Simple Soil Parameters

Md. Janibul Hoque, Md. Bayezid, Ahnaf Rafi Sharan, Mozaher Ul Kabir, Tahsin Tareque

Department of Civil and Environment Engineering, Islamic University of Technology, Gazipur, Bangladesh
Email: ahnafrafi@iut-dhaka.edu

Abstract

Cohesion is an important soil strength parameter for the overall structure and quality of building foundations as well as slope stability. For a civil engineering project, cohesion (c) can be determined directly from mainly unconfined compression tests, direct shear tests, and triaxial tests of soil. However, it's quite challenging to collect soil samples as there are time and cost constraints, as well as a diversity of soil deposits. Hence, this research aims to demonstrate a simplified method in order to determine the strength parameter of cohesive soil. Here, we propose an alternative solution adopting statistical correlations and machine learning techniques to establish correlations between the liquid limit, plastic limit, moisture content and %fine of soil with the strength parameter. In laboratory testing, these parameters can be obtained easily and these tests are relatively simple, quick to perform and also comparatively inexpensive. Hence, several test results were used from 100 boreholes which were soft soil or silty clay-type soil. Using the collected *in-situ* and lab test results of soil samples, a Multiple Linear Regression (MLR), Random Forest Regression (RFR) and Machine Learning (ML) model will be developed to establish a relationship between cohesion and the available test results. In order to assess the performances of both models, several performance indicators like: correlation coefficient (R^2), mean squared error (MSE), root mean square error (RMSE), and mean average error (MAE) will be used. These correlation coefficients will be used to demonstrate the prediction capacity and accuracy of both models. It should be noted that this approach will substitute the conventional testing required for strength parameters, which is both expensive and time-consuming.

Keywords

1. Introduction

Assessing Soil Parameters is essential in geotechnical engineering and construction practices. They provide critical information about the soil's properties such as Bearing capacity and stability, slope stability analysis, foundation design etc. These properties of soil must be understood properly to provide critical failure criteria for the construction of superstructures as well as the foundation of constructions [1]. Important soil parameters for assessing geotechnical properties are SPT-N value, Dry Density, Moisture content, Particle size distribution (Cu, Cc), Liquid limit, Plastic Limit etc. [2]. The resistance property of soil can be measured by its SPT-N value during the soil penetration test [3]. Dry Density refers to the mass of soil per unit volume when it is completely in a dry state. Moisture content, expressed as a percentage of the dry weight of soil, indicates the amount of water present in the soil. Particle size distribution provides information about the distribution and composition of soil particles across different size ranges. Atterberg limits such as Liquid limit which is the moisture content at which fine-grained soil transitions from a liquid-like to a plastic state and Plastic limit which is the moisture content at which fine-grained soil transitions from a plastic state to a semi-solid state.

Cohesion is an important part of the physical distribution forces of soil among others such as shear strength which can be experimentally determined with the help of other soil properties [4]. Soil consistency can be predicted using the value of cohesion which is important in Soil physics and Soil mechanics. It is the cohesive force that acts between soil particles according to soil physics. Cohesion in soil mechanics is the degree of shear of soil at a point where the compressive stress is equal to zero. The stability and capability of a soil to adjust when facing overburden loads and loading from structures, are greatly impacted by the shear strength of soil [5]. This shear strength parameter is important in terms of soil stability which denotes how much shear stress a soil can take before sliding down [6]. Thus, it is important to use soil with higher cohesion value to construct retaining walls on inaccessible terrain [7]. The shear strength parameter, especially the cohesion value of soil is of prime importance in the case of different foundation designs. For, Rough rigid strip footing, there is a variance of bearing capacity for the spatial variation of cohesion, and for shallow foundations, bearing capacity is mainly based upon cohesion characteristics of soil [8] [9]. Bearing capacity is an important factor for designing mechanically stabilized earth walls. Detachment rate of soil particles also has relation with soil cohesion. An inverse correlation has been seen between soil detachment and soil cohesion properties [10]. The detachment rate of soil can be determined using the depth of surface runoff of an area. Subsoil reaction to heavy live loads on the road also changes with the cohesion of soil. Soil modeling is important for the sensitivity analysis of soil and cohesion is an important parameter for soil modeling [11].

The cohesion value of a soil specimen taken from different depths of a borehole can be determined primarily by performing the UCS test, Direct Shear test,

and Triaxial tests. These tests have some restraints as it is very time-consuming. If a soil sample that is undisturbed is preferred for these tests different places in Bangladesh have different properties of soil, so we have to perform different shear strength tests [12]. However, for large-scale construction, the cohesion value is preferred to be obtained more easily and rapidly for a huge number of datasets. To reduce the constraints and to obtain the shear strength parameters *i.e.*, cohesion value, much research has been carried out. Analytical approaches for predicting shear strength values were first studied considering the bearing capacity mechanism of failure at the cone tip and from direct shear failure along the penetrometer sleeve [13]. For predicting soil shear wave velocities (Vs) from a cone penetration test, Multiple linear regression was used in Christchurch, New Zealand [14]. Some constitutive models were also developed to further research the creep behavior of soil which gives information about the shear strength of soil [15]. Among these studies, establishing a correlation between soil parameters and the cohesion property of soil to predict the shear strength is the most efficient way without performing laboratory-based shear strength tests. Correlation analysis between the shear strength parameter and unconfined compressive strength, bulk, and dry unit weight has been shown [16]. Multiple regression models for assessing correlation and prediction of shear strength [17] were also studied the correlation between effective cohesion and plasticity index of clay has been also studied earlier [18]. Predicting software-based cohesion metrics with Machine learning models which were studied earlier was not quite satisfactory [19]. Therefore, obtaining the cohesion value of soil samples using Correlation with different soil parameters is more rapid and can give the cohesion value of a large number of datasets with much accuracy.

The previous studies were seen to have been correlated to only some of the soil's parameters like plasticity index, cone resistance values [20], CPT-SPT values [21] etc. with other soil properties, especially the shear strength parameter of soil. If the correlation is established between several soil parameters such as SPT-N value, %sand and %fines, dry density, and moisture content, which are some important soil parameters that need to be tested before any construction upon the soil, with undrained cohesion (C_u) of soil, the prediction will be much more accurate. The correlation coefficient (R^2) value would be much higher with combined Multiple Linear Regression (MLR) of those soil parameters with cohesion.

Soil sample collection and analysis is the foremost step in soil characteristics examination and establishing a correlation between other soil properties. Most study of any type of soil prefers the undisturbed state of the soil sample as it represents the realistic behavior of soil under loading condition. It is preferable if a correlation of cohesion can be established with soil parameters using undisturbed soil samples to provide a more accurate assessment for further research using these correlations. The focus of our study is to correlate different soil parameters with shear strength (cohesion) using silty clay soil samples in undisturbed condition taken from boreholes of a specific region. This proposed study

of correlation can provide a better and accurate analogy of soil using important soil parameters compared to other correlation study performed before. Previously, a correlation of soil parameters was developed with existing CPT and SPT values for Bangladesh [22]. Akan *et al.* (2015) [23], Sharma *et al.* (2018) [24] and Ahmad *et al.* (2018) [25] attempted to develop correlations using multiple regression models to predict the unconfined compressive strength of soil. Zaman *et al.* (2016) [26] developed a correlation between consolidation properties of soil with liquid limit, in situ water content, void ratio, and plasticity index. Statistical models were also developed to predict the liquefaction and seismic hazard from SPT data [27] [28]. Moreover, several statistical approaches were used to develop models to conduct slope stability assessment of clayey sand hill tracts [29] [30] [31] and also the deformation of soil due to tunneling [32] [33] [34]. Later on, more advanced support vector machine (SVM) and artificial neural network (ANN) models were also used to predict the unconfined compressive strength by Tabarsa *et al.* (2021) [35]. The application of ANN models was also implemented in other fields of civil engineering, especially in transportation modeling, quality assessment and safety-related predictions [36] [37] [38] [39] [40]. Kabir *et al.* (2019) [41] also developed an ANN model to predict the bearing capacity of shallow foundations using relevant soil parameters that can be acquired from *in-situ* testing. These previous studies clarify the successful implementation of both statistical and advanced machine learning (ML) models in developing several prediction models and establishing correlations using data obtained from *in-situ* testing and relevant soil parameters.

Considering all these factors to establish a much more accurate correlation between shear strength (Undrained cohesion, C_u) and relevant soil parameters which are SPT-N value, %sand and %fines, dry density, and moisture content were taken into consideration from undisturbed soil samples of 100 Boreholes which were soft soil or silty clay type soil. They contained mostly fine particles and less amount of sand particles. Using these soil parameters, initially, Linear Regression (LR) was proposed with different variables according to their impact on cohesion value, and then Multiple Linear Regression (MLR) and Random Forest algorithms were also applied to develop correlation. All their performances were compared with the results obtained from the Machine Learning (ML) model. In order to assess the accuracy and performances of all the models, mean squared errors (MSE), Root mean square errors (RMSE), and mean average errors (MAE) were also calculated. By performing all these, this research aims to facilitate the prediction of the strength properties of soil to understand soil behavior.

2. Methodology

2.1. Study Area

Soil samples for our study were collected from various locations around Dhaka city, where we identified the existence of silty clay type soil. The soil samples

were primarily obtained from areas where the fine particles exceeded the other soil particle percentage by 8 - 9 times. We specifically focused on gathering undisturbed soil samples. We have collected samples from different depths of soil strata *i.e.*, 2.5 m, 5 m, 10 m, and 15 m below the surface. The gathered samples were placed in strong, labeled, sealed polythene bags before being taken to the lab for examination.

2.2. Material Collection

To drill into the boreholes of soils containing fine particles, the rotary method was performed with hydraulic rotation. The borings were advanced by rotation of the drilling bit which was pushed by hydraulic rotary pressure. Thin-walled Shelby tube has been used in the field for collecting undisturbed soil samples at an interval of 1.5 m. The borehole water level was measured when all the sediments were stabilized and the water in the borehole attained the condition of natural groundwater. After boring, the area has been backfilled properly.

2.3. Soil Characteristics

As we performed a sub-soil investigation, it was found that the percentage of fines was much higher than the sand or any other soil particle sizes. Sand is found in a very low percentage among the boreholes whereas the average percentage of clay and silt (fines) is (87% - 97%) which is much higher suggesting a clay-type soil sample (Table 1).

2.4. Laboratory Analysis

2.4.1. Particle Size Analysis
According to ASTM D422, Sieve analysis tests were carried out for the analysis of particle size distribution. Only the particles retained on a 0.075 mm sieve were used as oven-dry materials in order to analyze. It is calculated and given as a percentage of the sample mass how much soil was retained on each sieve. Fine materials underwent hydrometer examination.

2.4.2. Liquid Limit and Plastic Limit
Atterberg limits (mainly Liquid limit and Plastic limit) were determined according to ASTM D4318, of representative soil samples of cohesive soil. To determine these limits, we performed standard tests on soil samples *i.e.*, the Casagrande test for obtaining LL. These tests helped us understand the soil's consistency and its ability to undergo deformation under different moisture conditions.

2.4.3. Water Content
Water content was determined on an oven-drying basis (ASTM D2216). It was found by extensively drying moist or wet soil for 18 to 24 hours at a constant temperature of 105°C. The moisture or water content of a soil sample is calculated by splitting the mass of water by the mass of solid particles.

Table 1. Data collection of soil samples.

Number	Soil Sample	Wet basis	LL	PL	Fines	Undrained Shear Strength (qu), kPa
1	Undisturbed	34.66	45	11	89	91.68
2	Undisturbed	35.89	50	24	87	74.41
3	Undisturbed	27.03	46	19	91	150.9
4	Undisturbed	28.79	47	20	90	140.9
5	Undisturbed	29.66	58	30	84	282.5
6	Undisturbed	25.13	54	24	89	236.2
7	Undisturbed	35.22	53	26	89	33.02
8	Undisturbed	24.49	52	26	87	153.8
9	Undisturbed	20.42	54	28	86	0.65
10	Undisturbed	22.34	45	20	89	0.64
11	Undisturbed	24.68	45	19	83	224.3
12	Undisturbed	22.89	46	21	78	261.3
13	Undisturbed	23.63	52	26	82	345.9
14	Undisturbed	25.01	57	29	80	192.6
15	Undisturbed	27.66	52	26	80	284.7
16	Undisturbed	25.33	50	26	75	101.8
17	Undisturbed	45.68	77	42	86	136.9
18	Undisturbed	32.88	52	27	89	88.04
19	Undisturbed	29.93	60	32	85	249.4
20	Undisturbed	28.76	50	24	82	94.64
21	Undisturbed	28.76	48	24	86	157.3
22	Undisturbed	30.22	52	27	89	200.8
23	Undisturbed	41.69	45	20	90	17.09
24	Undisturbed	31.25	46	22	84	33.83
25	Undisturbed	21.47	45	22	86	238.5
26	Undisturbed	36.92	48	21	82	67.68
27	Undisturbed	22.12	51	27	79	230.9
28	Undisturbed	25.06	46	19	88	181.7
29	Undisturbed	25.72	54	28	84	203.6
30	Undisturbed	25.37	61	36	82	298.2
31	Undisturbed	21.24	57	32	89	153.8
32	Undisturbed	30.37	55	29	65	169.1
33	Undisturbed	44.15	56	25	82	110.7
34	Undisturbed	43.33	62	31	79	153.8
35	Undisturbed	46.39	70	36	71	96.04

Continued

36	Undisturbed	33.81	55	27	49	110.3
37	Undisturbed	30.09	56	30	84	176.9
38	Undisturbed	34.82	58	32	82	205.2
39	Undisturbed	32.79	56	28	86	202.3
40	Undisturbed	36.27	62	34	89	171.7
41	Undisturbed	18.80	55	18	97.1	327
42	Undisturbed	23.2	37	17	99.2	229
43	Undisturbed	23.2	33	16	95.9	153
44	Undisturbed	72.2	111	43	96.8	141
45	Undisturbed	22	55	18	96.5	219
46	Undisturbed	35.6	42	23	97.4	76
47	Undisturbed	75.9	144	23	76.7	249
48	Undisturbed	21.1	52	20	98.9	355
49	Undisturbed	28.3	56	22	98.8	151
50	Undisturbed	30.3	46	20	99.0	140
51	Undisturbed	25.5	57	22	98.6	99
52	Undisturbed	25.6	44	23	98.3	168
53	Undisturbed	23.2	30	15	49.7	124
54	Undisturbed	33.5	48	25	94.0	68
55	Undisturbed	20.2	38	16	86.0	260
56	Undisturbed	23.1	56	20	86.0	188
57	Undisturbed	37.1	57	21	93.1	37
58	Undisturbed	34.3	56	23	94.10	178
59	Undisturbed	31.1	51	21	97.80	47
60	Undisturbed	51	52	25	97.8	292
61	Undisturbed	20.8	38	17	95.5	194
62	Undisturbed	19.9	33	33	76.2	132
63	Undisturbed	88.9	101	46	93.4	45
64	Undisturbed	41.7	50	32	99.0	89
65	Undisturbed	30.5	52	21	99.5	128
66	Undisturbed	21.7	45	20	96.7	350
67	Undisturbed	24.7	61	18	92.9	117
68	Undisturbed	23.4	42	16	53.2	139
69	Undisturbed	61.2	63	31	98.0	9
70	Undisturbed	40.9	47	31	99.8	9
71	Undisturbed	21.9	33	18	96.3	351
72	Undisturbed	18.4	49	18	92.3	251
73	Undisturbed	28.7	46	24	97.5	23

Continued

74	Undisturbed	31.1	60	23	98.4	277
75	Undisturbed	49.4	95	40	98.7	23
76	Undisturbed	43.6	55	26	99.2	35
77	Undisturbed	41.4	46	29	99.50	65
78	Undisturbed	21.2	34	17	99.00	426
79	Undisturbed	33.6	40	29	96.9	117
80	Undisturbed	18.4	39	16	90.2	288
81	Undisturbed	19.5	41	17	97.4	350
82	Undisturbed	23.8	48	19	78.5	188
83	Undisturbed	24.2	30	17	48.0	80
84	Undisturbed	25.0	38	17	93.1	183
85	Undisturbed	22.0	34	17	48.8	128
86	Undisturbed	23.3	53	20	99.6	360
87	Undisturbed	21.3	58	21	89.0	450
88	Undisturbed	21.5	32	17	48.4	128
89	Undisturbed	23.7	56	22	92.6	381
90	Undisturbed	21.9	40	16	53.1	212
91	Undisturbed	25.3	57	20	91.3	486
92	Undisturbed	21.8	51	18	90.9	494
93	Undisturbed	24.2	47	18	66.2	392
94	Undisturbed	24.0	57	17	89.1	334
95	Undisturbed	23.9	54	19	82.9	478
96	Undisturbed	25.2	57	18	86.5	438
97	Undisturbed	23.1	50	18	76.6	381
98	Undisturbed	23.3	38	16	55.6	237
99	Undisturbed	22.6	34	15	47.1	141
100	Undisturbed	18.5	33	18	42.8	120

2.4.4. Dry and Bulk Unit Weight

We determined the dry unit weight of the soil by measuring the weight of a given volume of oven-dried soil according to ASTM D4531. This provided insights into the soil's density and compaction characteristics.

2.5. Research Methodology

In the data analysis phase of the methodology, we employed linear regression to explore the relationship between liquid limit, plastic limit, moisture content, %fine, and the shear strength parameter. Specifically, we calculated the R-squared (R^2) value for each individual linear regression model. Next, we selected the linear

regression model with the highest R^2 value as a basis for conducting multiple linear regression. This approach allowed us to incorporate multiple independent variables and assess their combined impact on predicting the shear strength parameter. Also, for accurate prediction Random Forest regression model was used alongside the Multiple Linear Regression model. Multiple linear regressions (MLR) are the method of statistics in regression that is used to analyze the relationship between a single response variable (dependent variable) with two or more controlled variables (independent variables) [42]. The selection of this method for the research was based on the presence of more independent variables. The Random Forest method integrates more than one decision tree and integrates all of the results from every one of the decision trees. Then find the final result. And because the simple processed many times from each decision tree, the result will have excellent accuracy [43]. To further evaluate the predictive performance of our models, we employed machine learning techniques. We utilized methods such as Mean Absolute Error (MAE) and Mean Squared Error (MSE) to quantify the accuracy and precision of our predictions. By combining traditional statistical analysis through linear regression, multiple linear regression, and machine learning methods, we aimed to gain a comprehensive understanding of the factors influencing the shear strength parameter and improve our predictive capabilities.

Here, the MLP (Multi-Layer Perceptron) model was used. MLP is an artificial neural network model. It is a deep learning basic model and is widely used for classification, regression, and pattern recognition. Multiple layers of neurons form the MLP. Here the feedforward neural network was applied. Backpropagation updates neuron weights to minimize the discrepancy between predicted and actual results during MLP training. An MLP neuron receives inputs from the previous layer, applies an activation function to the weighted sum, and generates to the next layer. MLP hidden layers between the input and output layers allow the model to learn complicated representations and extract features from input data. Stochastic gradient descent is used to calculate the gradient of the loss function with respect to the model's parameters and update the weights. The number of hidden layers and neurons in each layer are hyperparameters that can be modified for better model performance.

3. Performance Evaluation

The model's performance is assessed by comparing simulation values to actual output. This study assessed goodness-of-fit or correlation statistics. Prior to the model's prediction efficiency assessment, the important soil parameters were identified to be used as the input parameters for the prediction of cohesion. For that, initially, Simple Linear Regression was performed for the soil parameters, and comparing their R^2 values decisions on input parameters were taken. Among them only four such soil parameters (moisture content, LL, PL and %fines) which had R^2 values comparatively better than other parameters.

Moreover, to compare accuracy between all models we summarized the R^2, MAE, MSE, RMSE value of all the models used in our research. As they reflect smaller errors between actual and predictive values, lower values of MSE, MAE and RMSE represent better model performance.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2 \tag{1}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2} \tag{2}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right| \tag{3}$$

where,

N = number of data points, y_i = Actual observations and $\hat{y}_i$ = Predicted observations.

4. Results and Discussions

The correlation between considered soil parameters and undrained shear strength are shown in Figures 1-4 with their respective R^2 values. Although their coefficient of determination, R^2 is very much lower, the impact of their characteristics delivers much more important factors deciding them as our base properties to establish proper correlation from which we can observer and predict further soil cohesion values without the mediocre and time-consuming laboratory tests and only from our correlation easily.

For every variable, we got our different slope (m) value, and the intercept (c) was the same for every variable. Which is shown in Table 2. Equation (4) can be used to form four different equations from which we can predict cohesion test values to compare them with our experimented cohesion values and for further evaluation.

Simple linear equation,

$$y = mx + c \tag{4}$$

In our research study of the correlation between the shear strength parameter of soil, we applied python based Multiple Linear Regression model to correlate between multiple variables which are Moisture content (wet basis), Liquid Limit, Plastic Limit, %fines and predict simple linear equations to further assess the cohesion value of the soil sample. The correlation is predicted in Figure 5. Performing the MLR model on our dataset it was found that the result (R^2 value) was not satisfactory for the establishment of correlation which is why we introduced our dataset to the RFR (Random Forest Regression) model to address the limitation of low R^2 value. By RFR it is possible to perform a more comprehensive analysis of the data shown in Figure 6, resulting in a more accurate predictive model. Unlike MLR which depends on test datasets, RFR is a potent ensemble learning technique that makes predictions by combining various decision trees.
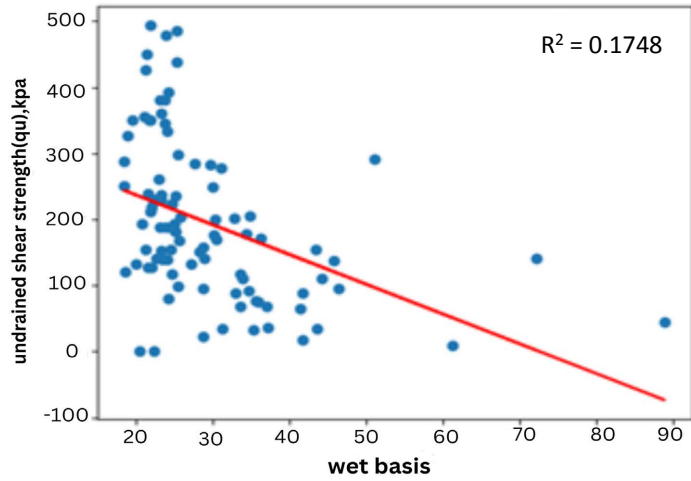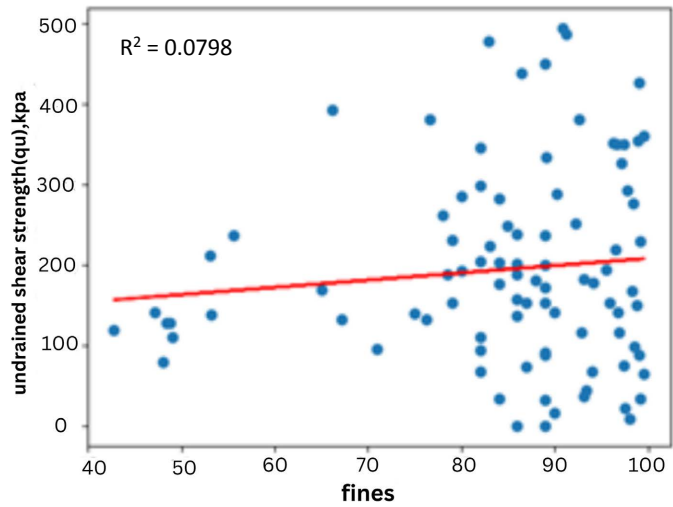
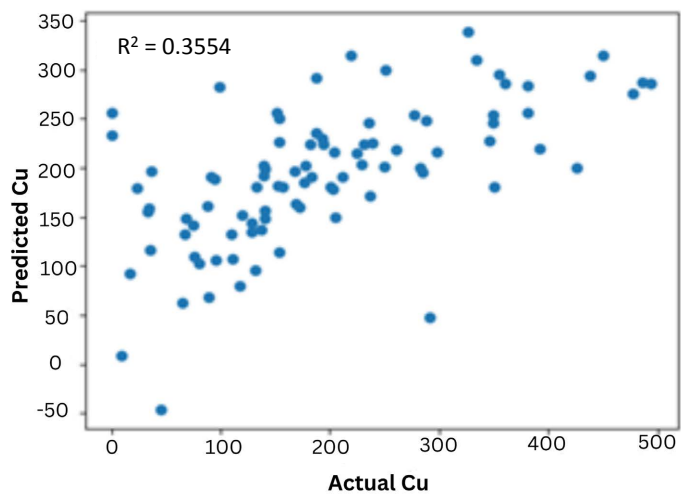**Figure 1.** Correlation with wet basis.
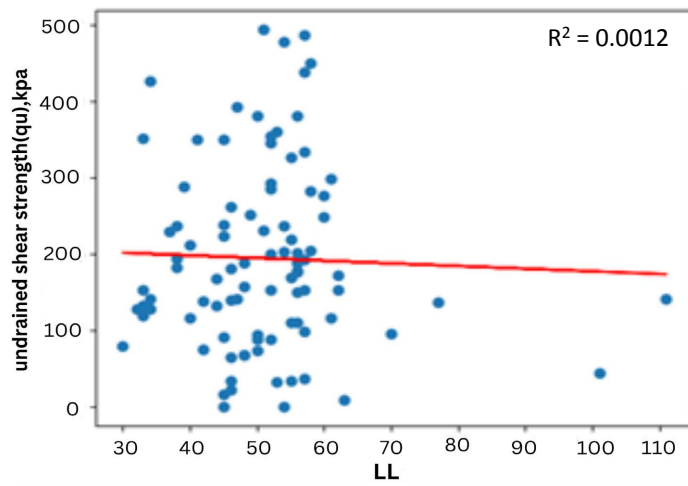
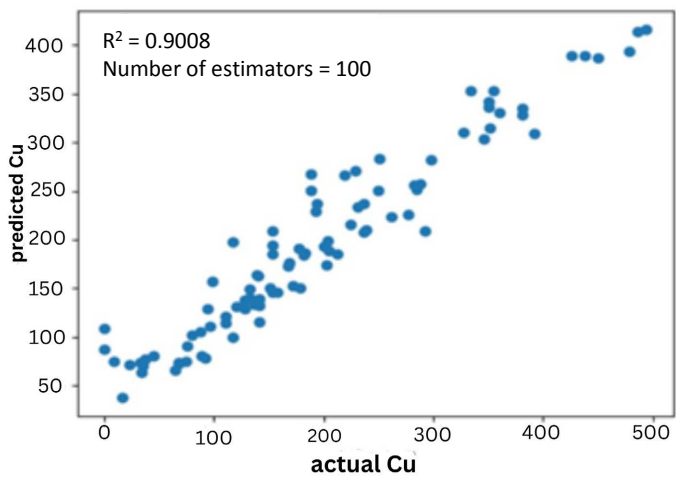**Figure 2.** Correlation with Liquid Limit.
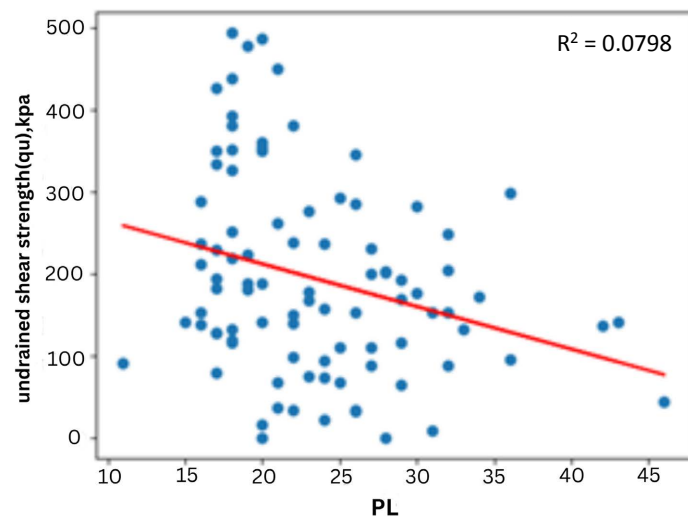
Figure 3. Correlation with Plastic Limit.

Figure 4. Correlation with %fines.

Figure 5. Actual C_u vs Predicted C_u (MLR).

Figure 6. Actual C_u vs Predicted C_u (RFR).

Table 2. Summary of different equations evaluated from MLR model.

Model	Result of the MLR model			
	Variables	Slope (m)	Intercept (c)	R^2
MLR	Wet basis	−7.3080	−139.453	0.354
	LL	6.0994		
	PL	−5.3190		
	%fines	0.9952		

Using MLP as a Machine Learning (ML) model to predict cohesion values, initially, the data sets were split into training and testing sets. 80% of the data was used to train the model. The rest 20% were used for testing and validation purposes. Batch size, number of hidden layers, and neurons in each layer are hyperparameters that were adjusted to obtain the best model performance. Table 3 represents the result of the ML model with varying hyperparameters. It was found that the R^2 value for the best-fitted model was quite like the Multiple Linear Regression model. From the results of the Machine learning model, it was clear that the model with batch size 8 and 10 hidden layers was able to forecast the unconfined strength more accurately. Table 4 represents the overall comparison of the all-model performances.

These models show how precisely we can correlate the soil parameters with the cohesion of soil and how accurately we will be predicting further cohesion values using those models. The R^2 values we got from different models were quite similar. But the R^2 value obtained from RFR model has quite higher credibility in predicting the unconfined compressive strength from the considered soil parameters. The soil sample which was collected for our study was in isotropic condition with standard climate change effect. If there happens to be any extreme weather conditions e.g. heavy rainfall, the property of soil will change according to the intensity of the weather condition so as the correlation this study has stated. When performing model framework, the laboratory results, and statistical measures were taken based on samples from similar region with minimal variation. Factors which transform soil into anisotropic condition e.g., seismic activities, can also alter local soil characteristics therefore impacting our observed correlation.

The result obtained from RFR model is very much satisfactory and accurate for the prediction of soil cohesion parameter. Thus, we can predict the cohesion values of clayey soil using these correlations which are quite accurate according to the result summary.

In our study, only silty clay-type soil sample data were used in our established model framework and to correlate them with other soil parameters. This framework would help us to study soil sample data of other regions with different soil characteristics resulting in the establishment of correlation using broad soil types.

Table 3. Results Obtained from ML model.

Model	Result of ML model					
	Batch Size	Hidden Layer	R^2	MAE	MSE	RMSE
ML	8	8	0.524	98.254	10742.451	10.3.645
		10	0.593	87.254	9278.621	96.325
	16	8	0.384	116.845	13984.647	118.256
		10	0.445	106.458	11847.277	108.845

Table 4. Result summary of all models.

Result Summary of models used				
Model Name	R^2	MAE	MSE	RMSE
MLR	0.355	95.811	1345.232	115.521
ML	0.593	87.254	9278.621	96.325
RF	0.901	64.513	6162.531	78.501

5. Conclusions

The outcome of this research offers many perspectives to understanding soil behavior. Initially, from this study of Dhaka soil, we formulated different prediction models and established correlations among soil parameters and shear strength of soil. Different approaches adopted were using Regression Models (Multiple Linear Regression model, RFR Model) and Machine Learning Model. Soil index properties, fines, and moisture content were also individually correlated with undrained cohesion where it showed very insignificant or almost no correlations. But while they were combined, it was found that the RFR model shows higher accuracy in predicting the undrained cohesion (c_u). Hence, using this best fitted prediction model RFR, it's possible to predict the unconfined compressive strength of soil with higher accuracy from plasticity properties and moisture content of soil. In the prediction of the unconfined compressive strength, it was clear that, though incorporating the machine learning model MLR, it yielded R^2 values which represent a lower capability to correlate the strength parameter of clayey soil. On the other hand, the RFR model showed much accuracy with a higher correlation value. The model performance indicators like: MAE, MSE, and RMSE were also less in the case of the RFR model which proves the higher accuracy of the model we used. So, this outcome can facilitate the prediction of unconfined compressive strength without performing any rigorous strength tests. As this study is in its preliminary state with only available data of 100 boreholes of a specific region, we could extend our study to different types of advanced soil including sandy soil of various field areas. Variation of data due to collection of soil samples from larger region and higher depth bore logs was not present in our study as this study was performed for a special soil type (silty clay) with data variation in narrow band. In future this study

could incorporate different machine learning methods such as LASSO Regression, Recursive Feature Elimination (RFE) and other Artificial Neural Network (ANN) models to correlate vast number of soil sample data and build strong correlation. Long term monitoring of different areas will enable the study of soil samples of numerous conditions and state with the help of correlation that this study has established. It can help us predict the soil cohesion with less deviation from the real field data taken from laboratory tests. These steps can be taken in future to implement our study to many civil engineering constructions to hasten future projects.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Omotoso, O.A., Ojo, O.J. and Adetolaju, E.T. (2012) Engineering Properties of in Sango Area, Ilorin, Nigeria. *Earth Science Research*, **1**, 71-81. https://doi.org/10.5539/esr.v1n2p71

[2] Surendra, R. and Sanjeev, K.B. (2017) Role of Geotechnical Properties of Soil on Civil Engineering Structures. *Resources and Environment*, **7**, 103-109.

[3] Sarker, D. and Abedin, M.Z. (2015) Applicability of Standard Penetration Test in Bangladesh and Graphical Representation of SPT-N Value. *International Journal of Science and Engineering Investigations*, **4**, 55-59.

[4] Temiz, C., Ari, F., Deviren Saygin, S., Arslan, S., Altay Unal, M. and Erpul, G. (2020) Soil Cohesion Development under Different Pore and Size Characteristics. 22*nd EGU General Assembly*, 4-8 May, 319-336. https://doi.org/10.5194/egusphere-egu2020-8325

[5] Roy, S. and Bhalla, S.K. (2017) Role of Geotechnical Properties of Soil on Civil Engineering Structures. *Resources and Environment*, **7**, 103-109.

[6] Yokoi, H. (2012) Relationship between Soil Cohesion and Shear Strength. *Soil Science and Plant Nutrition*, **14**, 89-93. https://doi.org/10.1080/00380768.1968.10432750

[7] Suljić, P.N. (2013) Importance of Soil Cohesion on the Stability of Retaining Reinforced Concrete Wall. 17*th International Research/Expert Conference*: *Trends in the Development of Machinery and Associated Technology*, Istanbul, 10-11 September 2013. https://www.researchgate.net/publication/331453309_Importance_of_soil_cohesion_on_the_stability_of_retaining_wall

[8] Griffiths, D.V., Fenton, G.A. and Manoharan, N. (2002) Bearing Capacity of Rough Rigid Strip Footing on Cohesive Soil: Probabilistic Study. *Journal of Geotechnical and Geoenvironmental Engineering*, **128**, 743-755. https://doi.org/10.1061/(ASCE)1090-0241(2002)128:9(743)

[9] Leshchinsky, B. (2015) Bearing Capacity of Footings Placed Adjacent to c'-ϕ' Slopes. *Journal of Geotechnical and Geoenvironmental Engineering*, **141**, Article ID: 04015022. https://doi.org/10.1061/(ASCE)GT.1943-5606.0001306

[10] Torri, D., Sfalanga, M. and Del Sette, M. (1987) Splash Detachment: Runoff Depth

and Soil Cohesion. *Catena*, **14**, 149-155.
https://doi.org/10.1016/S0341-8162(87)80013-9

[11] Poodt, M.P., Koolen, A.J. and Van der Linden, J.P. (2003) FEM Analysis of Subsoil Reaction on Heavy Wheel Loads with Emphasis on Soil Preconsolidation Stress and Cohesion. *Soil and Tillage Research*, **73**, 67-76.
https://doi.org/10.1016/S0167-1987(03)00100-4

[12] Ly, H.B. and Pham, B.T. (2020) Prediction of Shear Strength of Soil Using Direct Shear Test and Support Vector Machine Model. *The Open Construction & Building Technology Journal*, **14**, 41-50. https://doi.org/10.2174/1874836802014010041

[13] Motaghedi, H. and Eslami, A. (2014) Analytical Approach for Determination of Soil Shear Strength Parameters from CPT and CPTu Data. *Arabian Journal for Science and Engineering*, **39**, 4363-4376. https://doi.org/10.1007/s13369-014-1022-x

[14] McGann, C.R., Bradley, B.A., Taylor, M.L., Wotherspoon, L.M. and Cubrinovski, M. (2015) Development of an Empirical Correlation for Predicting Shear Wave Velocity of Christ Church Soils from Cone Penetration Test Data. *Soil Dynamics and Earthquake Engineering*, **75**, 66-75. https://doi.org/10.1016/j.soildyn.2015.03.023

[15] Oliveira, P.J.V., Correia, A.A. and Lemos, L.J. (2017) Numerical Prediction of the Reep Behaviour of an Unsterilized and a Chemically Stabilised Soft Soil. *Computers and Geotechnics*, **87**, 20-31. https://doi.org/10.1016/j.compgeo.2017.02.006

[16] Yimer, Y.M., Makesh, A.P. and Muhammed, S. (2021) Prediction of Undrained Shear Strength and Correlation in between Soil Parameters. *Journal of Physics: Conference Series*, **2040**, Article ID: 012024.
https://doi.org/10.1088/1742-6596/2040/1/012024

[17] Roy, S. and Dass, G. (2014) Statistical Models for the Prediction of Shear Strength Parameters at Sirsa, India. *International Journal of Civil & Structural Engineering*, **4**, 483-498.

[18] Tchakalova, B. and Ivanov, P. (2021) Correlation between Effective Cohesion and Plasticity Index of Clay. *Geologica Balcanica*, **51**, 45-49.
https://doi.org/10.52321/GeolBalc.51.3.45

[19] Haner Kırğıl, E.N. and Erçelebi Ayyıldız, T. (2023) Predicting Software Cohesion Metrics with Machine Learning Techniques. *Applied Sciences*, **13**, Article 3722.
https://doi.org/10.3390/app13063722

[20] Estikhamah, F. and Solin, D.P. (2021) Correlation between Cone Resistance Values and Cohesion Values in Cohesive Soils (Case Study in Gunung Anyar District). *E3S Web of Conferences*, **328**, Article No. 01005.
https://doi.org/10.1051/e3sconf/202132801005

[21] Hossain, M.M., Sultana, N. and Malo, R.C. (2020) Correlations between CPT, SPT and Soil Parameters for Khulna, Bangladesh. *Journal of Geotechnical Studies*, **5**, 27-32.

[22] Mominul, H.M., Ansary, M.A. and Abedin, M.J. (2014) Applicability of Existing CPT, SPT and Soil Parameter Correlations for Bangladesh. *3rd International Symposium on Cone Penetration Testing*, Las Vegas, 13-14 May 2014, 373-382.

[23] Akan, R., Keskin, S.N. and Uzundurukan, S. (2015) Multiple Regression Model for the Pre Diction of Unconfined Compressive Strength of Jet Grout Columns. *Procedia Earth and Planetary Science*, **15**, 299-303.
https://doi.org/10.1016/j.proeps.2015.08.072

[24] Sharma, L.K. and Singh, T.N. (2018) Regression-Based Models for the Prediction of Uncon Fined Compressive Strength of Artificially Structured Soil. *Engineering with*

Computers, **34**, 175-186. https://doi.org/10.1007/s00366-017-0528-8

[25] Ahmad, K.I., Abrar, M., Al Shafian, S. and Shahin, H.M. (2018) Correlation among the Soil Parameters of the Karnaphuli River Tunnel Project. *International Journal of GEOMATE*, **15**, 86-90. https://doi.org/10.21660/2018.48.7257

[26] Zaman, M.W., Hossain, R., Hossain, S. and Al Alam, A. (2016) A Study on Correlation between Consolidation Properties of Soil with Liquid Limit, *in situ* Water Content, Void Ratio and Plasticity Index. *The* 3*rd International Conference on Geotechnics for Sustainable Infrastructure Development*, Hanoi, 24-25 November 2016, 899-902.

[27] Islam, S., Shahin, H.M., Kabir, M.U. and Islam, M. (2023) Provision of Building Codes in the Context of Seismic Site Characterization and Liquefaction Susceptibility Assessment. *International Journal of GEOMATE*, **25**, 50-58. https://doi.org/10.21660/2023.108.3795

[28] Shahin, H.M., Kabir, M.U. and Islam, S. (2023) Seismic Hazard Analysis at Site Specific Condition: Case Study in Araihazar, Bangladesh. *IUT Journal of Engineering and Technology* (*JET*), **15**, 8-20.

[29] Kabir, M.U., Islam, M.S., Nazrul, F.B. and Shahin, H.M. (2023) Comparative Stability and Behavior Assessment of a Hill Slope on Clayey Sand Hill Tracts. *International Journal of Engineering Trends and Technology*, **71**, 11-24. https://doi.org/10.14445/22315381/IJETT-V71I1P202

[30] Zaman, M.W., Mita, K.S., Al Azad, A., Hossain, R. and Ul, M. (2019) A Numerical Study on Slope Stability Analysis by Finite Element Method Using Femtij-2D Application. *International Conference on Disaster Risk Management* (*ICDRM* 2019), Dhaka, 12-14 January 2019, page.

[31] Al Shafian, S., Talha, S.A., Hossain, M. and Shahin, H.M. (2021) Understanding the Effect of Degree of Saturation on Compressive Strength Behaviour for Reconstituted Clayey Soil of Dhaka. *The* 11*th International Conference on Geotechnique, Construction Materials and Environment*, Kyoto, 3-5 November 2021, 242-246.

[32] Al Shafian, S., Imtiyaz, M.N. and Emtiaz, M. (2023) Deformation Behaviour Analysis Using Finite Element Method during Deep Excavation in Dhaka City. *Proceedings of the* 8*th World Congress on Civil, Structural, and Environmental Engineering* (*CSEE'*23), Lisbon, 29-31 March 2023. https://doi.org/10.11159/icgre23.139

[33] Al Shafian, S., Imtiyaz, M.N. and Emtiaz, M. (2023) A Finite Element Approach to Investigate the Deformation Behaviour in Deep Excavation for TBM Launching Shaft. *International Journal of Civil Infrastructure*, **6**, 36-40. https://doi.org/10.11159/ijci.2023.005

[34] Kabir, M.U., Hossain, S.A., Alam, M.D. and Azim, M.D. (2017) Numerical Analyses of the Karnaphuli River Tunnel. Master's Thesis, Islamic University of Technology (IUT), Gazipur.

[35] Tabarsa, A., Latifi, N., Osouli, A. and Bagheri, Y. (2021) Unconfined Compressive Strength Prediction of Soils Stabilized Using Artificial Neural Networks and Support Vector Machines. *Frontiers of Structural and Civil Engineering*, **15**, 520-536. https://doi.org/10.1007/s11709-021-0689-9

[36] Imtiyaz, M.N. (2018) Estimating Future Crash Rates on Roads in Louisiana. Master's Thesis, Louisiana State University and Agricultural & Mechanical College, Baton Rouge. https://doi.org/10.31390/gradschool_theses.5493

[37] Wu, Z. and Emtiaz, M. (2022) Quality Management of Cracking Distress Survey in Flexible Pavements Using LTRC Digital Highway Data Vehicle. Louisiana State

University, Louisiana Transportation Research Center, No. FHWA/LA. 21/659.

[38] Emtiaz, M. (2020) Quality Assessment of Flexible Pavement Automated Cracking Data in Louisiana. Master's Thesis, Louisiana State University and Agricultural & Mechanical College, Baton Rouge.

[39] Majumder, M. (2020) An Approach to Counting Vehicles from Pre-Recorded Video Using Computer Algorithms. Master's Thesis, Louisiana State University and Agricultural & Mechanical College, Baton Rouge.
https://doi.org/10.31390/gradschool_theses.5231

[40] Majumder, M. and Wilmot, C. (2023) Automated Vehicle Counting from Pre-Recorded Video Using You Only Look Once (YOLO) Object Detection Model. *Journal of Imaging*, **9**, Article 131. https://doi.org/10.3390/jimaging9070131

[41] Kabir, M.U., Sakib, S.S., Rahman, I. and Shahin, H.M. (2019) Performance of ANN Model in Predicting the Bearing Capacity of Shallow Foundations. In: Sundaram, R., Shahu, J. and Havanagi, V., Eds., *Geotechnics for Transportation Infrastructure*, Springer, Singapore, 695-703. https://doi.org/10.1007/978-981-13-6713-7_55

[42] Md Ghania, I.M. and Ahmadb, S. (2010) Stepwise Multiple Regression Method to Forecast Fish Landing. *Procedia—Social and Behavioral Sciences*, **8**, 549-554.

[43] Guan, R.D. (2023) Predicting Forest Fire with Linear Regression and Random Forest. *Highlights in Science, Engineering and Technology*, **44**, 1-7.
https://doi.org/10.54097/hset.v44i.7159