



## **Statistical Analysis on Shear Strength Parameter from Index Properties of Fine-grained Soils**

**Tariku Tafari Bakala<sup>1\*</sup>, Emer Tucay Quezon<sup>2</sup> and Mohammed Yasin<sup>3</sup>**

<sup>1</sup>Civil Engineering Department, Assosa University, Assosa, Benishangul- Gumuz Region, Ethiopia.

<sup>2</sup>Civil Engineering and Construction Engineering and Management Streams, Ambo Institute of Technology, Ambo University, Ambo, Oromia Region, Ethiopia.

<sup>3</sup>Civil Engineering, Jimma Institute of Technology, Jimma University, Jimma, Oromia, Ethiopia.

### **Authors' contributions**

*This work was carried out in collaboration among all authors. Author TTB designed the study, performed the statistical analysis, wrote the protocol and wrote the first draft of the manuscript. Author ETQ managed the analyses of the study. Author MY managed the literature searches. All authors read and approved the final manuscript.*

### **Article Information**

DOI: 10.9734/JERR/2021/v20i417291

#### Editor(s):

(1) Dr. Guang Yih Sheu, Chang-Jung Christian University, Taiwan.

#### Reviewers:

(1) Jian-Hong Wu, National Cheng Kung University, Taiwan.

(2) Lim Hwee San, Universiti Sains Malaysia (USM), Malaysia.

Complete Peer review History: <http://www.sdiarticle4.com/review-history/66257>

**Original Research Article**

**Received 05 January 2021**

**Accepted 08 March 2021**

**Published 18 March 2021**

### **ABSTRACT**

Shear strength is the essential engineering property of soil required to analyze and design foundations, retaining walls, bridges, embankment, and related infrastructure. The laboratory equipment and field instruments are not sufficient in developing countries to obtain soil engineering properties, especially strength properties. Thus, Geotechnical engineers usually endeavor to develop statistical models that best fit a particular area and soil type, especially for analysis and design purposes. In this research, a Statistical Analysis on the Shear Strength parameter from the Index Properties of Fine-Grained Soils was studied. For predicting the undrained shear strength parameter, single linear regression (SLR) and multiple linear regressions (MLR) analyses were developed. To develop the intended statistical models for a study, SAS JMP Pro 13, SPSS v22, and Microsoft Excel-2013 software were introduced. The results of a study indicated that undrained shear strength( $C_u$ ) was significantly correlated with liquid limit(LL), plastic limit(PL), bulk density ( $\rho_{bulk}$ ), dry density( $\rho_{dry}$ ), natural moisture content(NMC), and plasticity index(PI). While it was not significantly correlated with a specific gravity (Gs) and liquidity index (LI) of study area soil.

\*Corresponding author: Email: [tarikutafari09@yahoo.com](mailto:tarikutafari09@yahoo.com), [tarikutafari0917@yahoo.com](mailto:tarikutafari0917@yahoo.com);

Finally, a strong Model of Cu with a coefficient of determination ( $R^2 = 0.806$ ), good significance level, and less Std. error was obtained from multiple linear regression (MLR) analysis. The developed model can figure undrained shear strength parameter and wide application in the construction industry to minimize the cost, effort, and time for laboratory tests of shear strength parameter of a study area.

*Keywords: Index properties; undrained shear strength; linear; statistical analysis.*

## 1. INTRODUCTION

One of the most significant engineering properties of soil is its capacity to respond sliding along internal surfaces within a mass. The stability of structures built on soil depends on the shearing resistance obtainable by the soil along probable surfaces of slippage [1]. It is quite essential that an engineer has to safeguard the structure is safe against shear catastrophe in the soil that supports it and does not undergo excessive settlement [2].

Undrained shear strength is one of the parameters to the bearing capacity of soil that could convey any structure rest on it. Some laboratory tests needed to obtain these parameters are costly and laborious, while other soil properties like index properties can be achieved quicker and cheaper. Index properties are the basis for distinguishing soils [3].

The undrained shear strength is used to estimate the short-term bearing capacity of fine-grained soils for foundations and estimate slopes' short-term stability. Besides, it compares the shear strength of soils from a site to create soil strength variability quickly and cost-effectively and use for stress-strain characteristics under fast (undrained) loading conditions [4-6]. In this research, undrained shear strength was performed by correlating with index properties of soils, which can be used to minimize cost, effort, and time for any geotechnical practice to analyze and design conditions of the study area.

Most of the time, for the need to analyze and design foundations, slope stability and other infrastructure engineers follow some techniques apart from doing an applicable investigation of subsoil incident, which may reason for the destruction of structures on it.

Experimental determination of the strength parameters used for design purposes is widespread, challenging to perform, and costly compared to index properties of soils.

Statistical modeling for a prediction in geotechnical engineering has been used to model different engineering properties of soils. This is pointed out that the prominence of statistical modeling for prediction in geotechnical practice is greatly vigorous.

It is often necessary for the geotechnical engineer to quickly characterize the soil and determine its engineering properties to estimate the soil's appropriateness for any industrial practices.

Consequently, statistical analysis is a crucial method to predict soils' engineering properties, especially for developing countries like Ethiopia, where there is a financial limitation, lack of test equipment, and limited time, which is used for design purposes.

However, index properties can be obtained simply with low-cost equipment when compared to strength properties equipment. This study was done in Agaro town, one of the rapidly growing southwestern parts of Ethiopia. A few soil tests were carried out in this town before, although it is known as a coffee production center wherein construction undertakings have been enormously creating.

In this study area, no laboratory equipment is available to test engineering properties of soils, especially strength parameters, for analysis and design purposes. From the related literature review, one author mentioned that not always possible to conduct the tests on every new situation. To cope with such problems, numerical solutions have been developed to estimate the shear strength parameters [7].

This study developed a statistical model that helped forecast the undrained shear strength from soils' index properties. For restraint of time, costs, and lack of testing equipment of shear strength, it is accessible for an engineer or consultants, contractors, clients, municipality, and researchers to get whatever they want for analysis and design purposes area from simple index properties test results.

## 2. MATERIALS AND METHODS

### 2.1 Description of Study Area

This study was accompanied in Agaro town, as indicated in Fig. 1, found in the Jimma Zone of the Oromia Region, Ethiopia. It sits at estimated latitude and longitude of 7°51'N 36°35'E and has an altitude of 1670 meters above sea level. It is 393 km from the capital, Addis Ababa, and 46 km from Jimma town.

### 2.2 Methods

For this work, various field visits (reconnaissance) were done to differentiate the study area's real situations. Furthermore, undrained shear strength and index properties of study area soil were performed in a Jimma Institute of Technology laboratory. Based on obtained test results, linear statistical analyses were done between undrained shear strength and index properties of study area soil.

The location of test pits was selected to characterize the soil types found in the study area. It was conducted the excavations of test pits. From excavated test pits, thirty representative undisturbed and disturbed samples were taken by Shelby tube & plastic bags, respectively.

For determination of natural moisture content, undrained shear strength, and density tests, undisturbed samples were used. In contrast, disturbed samples were used to determine index properties tests such as specific gravity, liquid limit, plastic limit, plasticity index, and grain size analysis. In this study, the following laboratory

tests were performed to determine the index properties and undrained shear strength of study area soils.

#### 2.2.1 Natural moisture content (NMC)

Moisture contents of the soil samples were determined in the laboratory as per ASTM D 2216 standard. A set of samples were dried to a constant weight using oven-dry at a temperature of 105°C.

#### 2.2.2 Specific gravity (Gs)

The test was performed as per ASTM D854-58, Standard Test for Specific Gravity of Soil Solids by density bottle, procedure.

#### 2.2.3 Dry density ( $\rho_{dry}$ ) and bulk density ( $\rho_{bulk}$ )

The density of soil was determined according to ASTM D 2937 (a standard test for the density of soil in place by the drive cylinder method). This technique is achieved to determine the in-place density of undisturbed soil found by pushing or drilling a thin-walled cylinder.

#### 2.2.4 Grain size analysis

This test was performed as per ASTM D422 standard to determine the percentage of different grain sizes contained within a soil. The mechanical or sieve analysis is used to decide the distribution of the coarser, larger-sized particles, and the hydrometer analysis method was used to determine the distribution of the finer particles, respectively. For this study, both wet sieve analysis and hydrometer analysis was done.

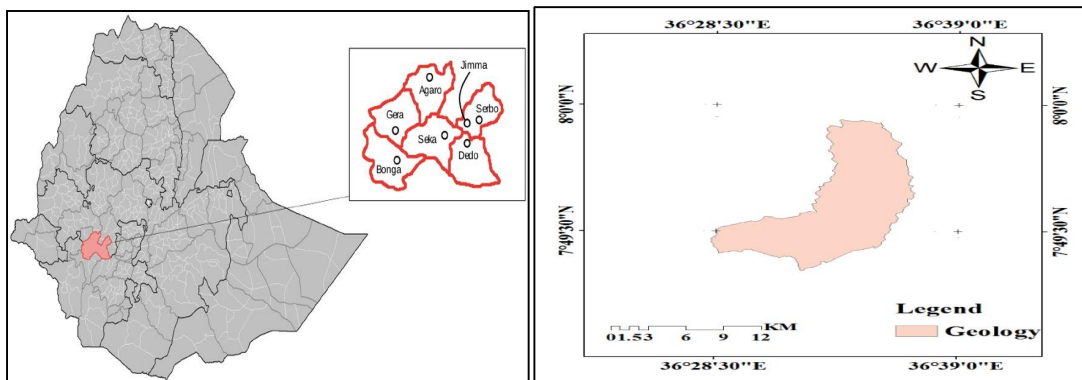


Fig. 1. Location of the study area

### 2.2.5 Atterberg limit's test

This test was performed as per ASTM D-4318 standard for Liquid Limit (LL), Plastic Limit (PL), and Plasticity Index (PI) of soils. The air-dried samples were organized by drying the specimen in the air. The quotas of the samples passing the No. 40 (0.425mm) sieve were used for the preparation of the sample for the test.

### 2.2.6 Undrained shear strength (Cu)

Undrained shear strength (Cu) was determined according to ASTM D 2166 standard. The test was conducted on undisturbed samples collected by Shelby tube sampler. These sample properties were used during the statistical analysis between the undrained shear strength and index properties of soils using the SAS JMP Pro 13, SPSS v22 software, and excel 2013 Microsoft. Moreover, statistical regression analyses of single and multiple models of test results were carried out. Statistical models by regression analysis were analyzed and developed to fit the test results. Beneath the discussions of the obtained results, the developed models' fitness was inspected in different ways.

## 3. RESULTS AND DISCUSSIONS

### 3.1 Regression Analysis and Modeling between the Response Variable and Predictors

#### 3.1.1 Scatter plot strategy

To develop correlations, the first step is creating a scatter plot of the data [8]. Correlation analysis

is a term used to denote the association or relationship between two (or more) quantitative variables. This analysis is fundamentally based on the assumption of a straight-line with the construction of a scatter plot or scatter diagram [a graphical of the data] with one variable on the X-axis and the other on the Y-axis [9]. In this study, the undrained shear strength (Cu) was taken as the predicted variable (dependent) while the predictors (independent) variables represented by the specific gravity, liquid limit, plastic limit, plasticity index, liquidity index, bulk density, dry density, and natural moisture content. Before the execution of the regression analysis using the test results, a scatter plot matrix was produced by applying the SAS JMP Pro.13 to study the relations developed between the dependent variable and the predictor variables by visualizing to determine the model that best outfits the test results as indicated on Fig. 2 and Fig. 3.

From scatter plots offered in Fig. 2 and Fig. 3, a visual method of displaying a relationship between variables is plotted in a two-dimensional coordinate system. The scatter plots' assessment indicated that a real indication of the points lie scattered arbitrarily as a straight or looks like a straight line, mainly for the liquid limit, plastic limit, natural moisture content, bulk density, dry density, and plasticity index. However, the remaining independent variables, such as specific gravity and liquidity index, by some extent, outliers away from the possible visual straight. The above scatter plots are indicated a linear response, and hence, a linear regression model expressed the association between the focus parameters.

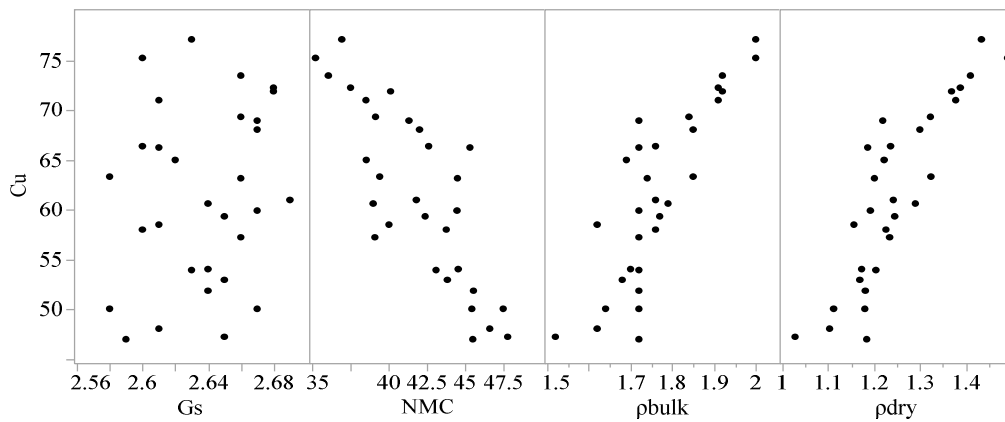


Fig. 2. Scatter plot diagram of undrained shear strength versus Gs, NMC,  $\rho_{bulk}$ ,  $\rho_{dry}$

### 3.1.2 Normality test

It is essential to check normality before proceeding with any applicable statistical procedures. If the assumption of normality is violated, interpretation and inference may not be reliable or valid. For N (number of samples, in this case, thirty) is less than 2,000 (for a small number of the sample), it is recommended to read the Shapiro-Wilk statistic that does not reject the null hypothesis of normality for  $p > 0.05$  [10]. Shapiro-Wilk test has a proper performance with a sample size of 7-2000. It is not possible to apply all the available tests to evaluate normality in any of the statistical software programs. However, it is possible to run the two commonly used Kolmogorov-Smirnov and Shapiro-Wilk tests with SPSS [11]. Based on Table 1, both predicted, and predictors data were normally distributed (i.e. p-value more than 0.05).

### 3.2 Regression Analysis

In this study, an effort was made to apply single linear regression model (SLR) and multiple

linear regression (MLR) models to describe the strength of (as correlation indicated in Table 2.) cohesive soil from soil index properties using a statistical approach. Multiple linear regression is a method of analysis for assessing the strength of the relationship between each set of independent variables and a single response variable. In contrast, when only a single explanatory variable is involved, it is generally referred to as simple linear regression [12].

To do this modeling, statistical Software SAS JMP Pro.13, SPSS V22 and Microsoft excel 2013 soft wares were used to study the significance of individual predictor variables as well as to get the best model. In view of that, the thirty-laboratory test results of the independent and dependent variables were used in the regression analysis to get the intended statistical model. To detect the influence of one variable on the other, a stepwise linear regression has been analyzed, and as a result, the respective correlation coefficients and level of significance were determined.

Table 1. Normality tests

	Kolmogorov-Smirnov			Shapiro-wilk		
	Statistic	df	Sig.	Statistic	Df	Sig.(p-value)
Gs	0.134	30	0.177	0.947	30	0.136
NMC	0.102	30	0.2	0.966	30	0.441
$\rho_{bulk}$	0.161	30	0.046	0.946	30	0.129
$\rho_{drv}$	0.173	30	0.022	0.959	30	0.296
LL	0.103	30	0.2	0.979	30	0.802
PL	0.137	30	0.156	0.937	30	0.074
PI	0.134	30	0.18	0.964	30	0.386
LI	0.131	30	0.2	0.966	30	0.438
Cu	0.095	30	0.2	0.962	30	0.349

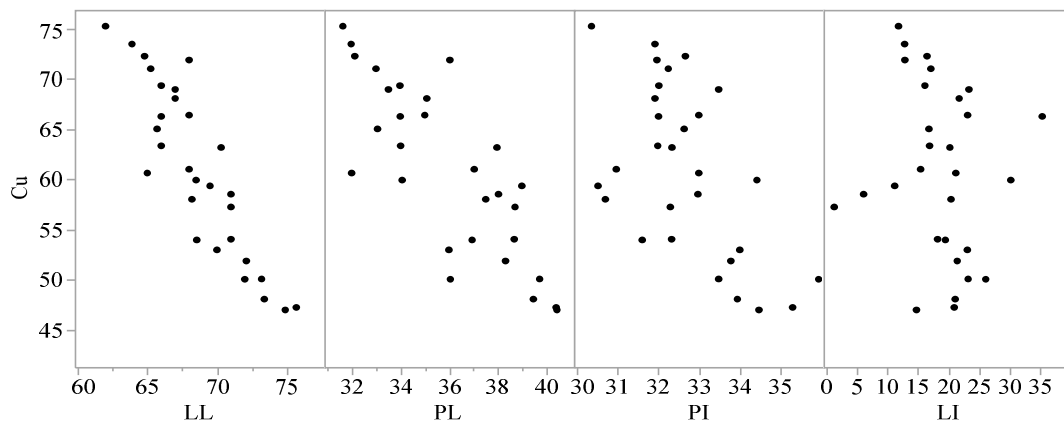


Fig. 3. Scatter plot diagram of undrained shear strength with LL, PL, PI, LI

**Table 2. Significance level ( $\alpha$ ) and Pearson correlation coefficient (R) in correlations**

		Gs	NMC	$\rho_{bulk}$	$\rho_{dry}$	LL	PL	PI	LI	Cu
Gs	R	1	-0.111	0.108	0.103	-0.076	-0.124	0.063	-0.036	0.187
	Sig.(2tail)		0.561	0.571	0.587	0.688	0.515	0.739	0.849	0.322
	N	30	30.000	30.000	30.000	30.000	30.000	30.000	30.000	30.000
NMC	R	-	1.000	-0.777	-0.879	0.832	0.756	0.547	0.500	-0.813
	Sig.(2tail)	0.111		0.000	0.000	0.000	0.000	0.002	0.001	0.000
	N	30	30.000	30.000	30.000	30.000	30.000	30.000	30.000	30.000
$\rho_{bulk}$	R	0.108	-0.777	1.000	0.981	-0.805	-0.713	-0.566	-0.260	0.835
	Sig.(2tail)	0.571	0.000		0.000	0.000	0.000	0.001	0.166	0.000
	N	30	30.000	30.000	30.000	30.000	30.000	30.000	30.000	30.000
$\rho_{dry}$	R	0.103	-0.879	0.981	1.000	-0.855	-0.762	-0.591	-0.356	0.866
	Sig.(2tail)	0.587	0.000	0.000		0.000	0.000	0.001	0.053	0.000
	N	30	30.000	30.000	30.000	30.000	30.000	30.000	30.000	30.000
LL	R	-	0.832	-0.805	-0.855	1.000	0.927	0.618	0.062	-0.897
	Sig.(2tail)	0.076		0.000	0.000		0.000	0.000	0.746	0.000
	N	30	30.000	30.000	30.000	30.000	30.000	30.000	30.000	30.000
PL	R	-	0.756	-0.713	-0.762	0.927	1.000	0.278	-0.113	-0.822
	Sig.(2tail)	0.124		0.000	0.000	0.000		0.137	0.554	0.000
	N	30	30.000	30.000	30.000	30.000	30.000	30.000	30.000	30.000
PI	R	0.063	0.547	-0.566	-0.591	0.618	0.278	1.000	0.393	-0.575
	Sig.(2tail)	0.739	0.002	0.001	0.001	0.000	0.137		0.031	0.001
	N	30	30.000	30.000	30.000	30.000	30.000	30.000	30.000	30.000
LI	R	-	0.562	-0.260	-0.356	0.062	-0.113	0.393	1.000	-0.171
	Sig.(2tail)	0.036		0.166	0.053	0.746	0.554	0.031		0.367
	N	30	30.000	30.000	30.000	30.000	30.000	30.000	30.000	30.000
Cu	R	0.187	-0.813	0.835	0.866	-0.897	-0.822	-0.575	-0.171	1.000
	Sig.(2tail)	0.322	0.000	0.000	0.000	0.000	0.000	0.001	0.367	
	N	30	30.000	30.000	30.000	30.000	30.000	30.000	30.000	30.000

As cited by Roy et al. [7], the procedure is commonly used to produce a parsimonious model that maximizes accuracy with an optionally reduced number of predictor variables.

Table 2. linear relationships shows that the correlation between  $C_u$  with liquid limit (LL), plastic limit (PL), natural moisture content (NMC), density ( $\rho_{bulk}, \rho_{dry}$ ) & plasticity index (PI) relatively stronger. However,  $C_u$  has a weak correlation with specific gravity (Gs) & Liquidity index (LI). In fact, the strength of fine-grained soil has a more significant association with soil consistency. Consequently, those parameters have resulted in a relatively strong correlation with the strength parameter ( $C_u$ ). This was due to the presence of more clay and silty in that soils. In this study, many alternative linear regression analyses that best fit the obtained test results were carried out.

### 3.2.1 Simple linear regression (SLR) analysis

#### 3.2.1.1 Model-A

Model between undrained shear strength ( $C_u$ ) and specific gravity (Gs).

The resulting Model, after correlating  $C_u$  with Gs is obtained from SAS JMP pro 13 & SPSS v22 outputs with its corresponding statistical parameters as on Fig. 4.

$$C_u = -75.786 + 52.014 * Gs, \text{ with } R^2 = 0.035, \alpha = 0.322 > 0.05, N = 30$$

The statistical output details showed that the relationship developed between Gs and  $C_u$  is insignificant (i.e.  $\alpha > 0.05$ ).

Furthermore, the relationship between correlation variables is weak ( $R^2 < 0.5$ ).

3.2.1.2 Model-B

Model between Cu and Natural moisture content (NMC)

$$Cu = 148.515 - 2.078 * NMC, \text{ with } R^2=0.662, \alpha = 0.00 < 0.05, N = 30$$

The statistical output details indicated that the relationship developed between NMC and Cu is significant ( $\alpha < 0.05$ ), and a good correlation happened concerning the correlating variables. Fig. 5 shows the graphical output of the model between Cu and Natural moisture content (NMC).

3.2.1.3 Model-C

Model between Cu and bulk density ( $\rho_{bulk}$ ) as indicated on Fig. 6.

$$Cu = -54.278 + 65.452 * \rho_{bulk}, \text{ with } R^2=.698, \alpha = 0.00 < 0.05, N = 30$$

The statistical output details indicated that the relationship developed between Cu and  $\rho_{bulk}$  is significant ( $\alpha < 0.05$ ), and a good relationship exists between the correlation variables.

3.2.1.4 Model-D

Model between Cu and dry density ( $\rho_{dry}$ )

$$Cu = -29.670 + 73.031 * \rho_{dry}, \text{ with } R^2=.750, \alpha = .000 < 0.05$$

The statistical output details specified that the relationship developed between  $\rho_{dry}$  and Cu is significant ( $\alpha < 0.05$ ) and good determinant factor.

3.2.1.5 Model-E

Model between Cu and Liquid limit (LL)

$$Cu = 220.604 - 2.323 * LL, R^2= 0.805, \alpha = 0.000 < 0.05$$

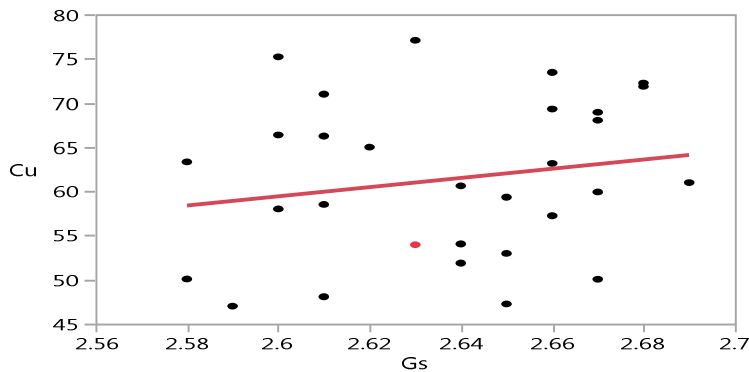


Fig. 4. Linear fit of Gs-Cu

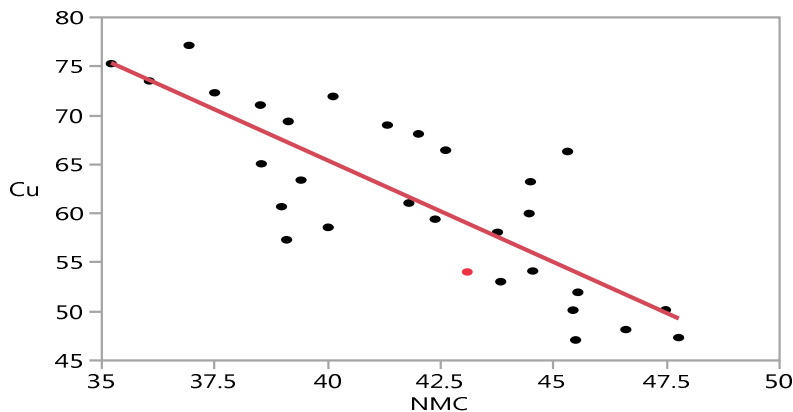


Fig. 5. Linear fit of NMC-Cu

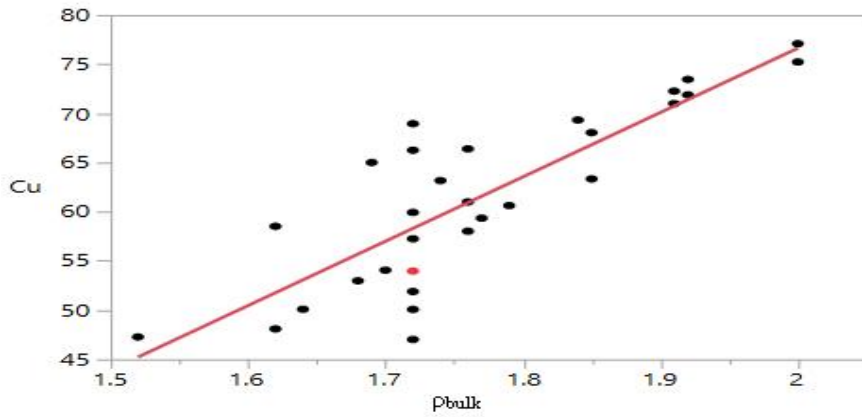


Fig. 6. Linear fit of  $\rho_{\text{bulk}}$ -Cu

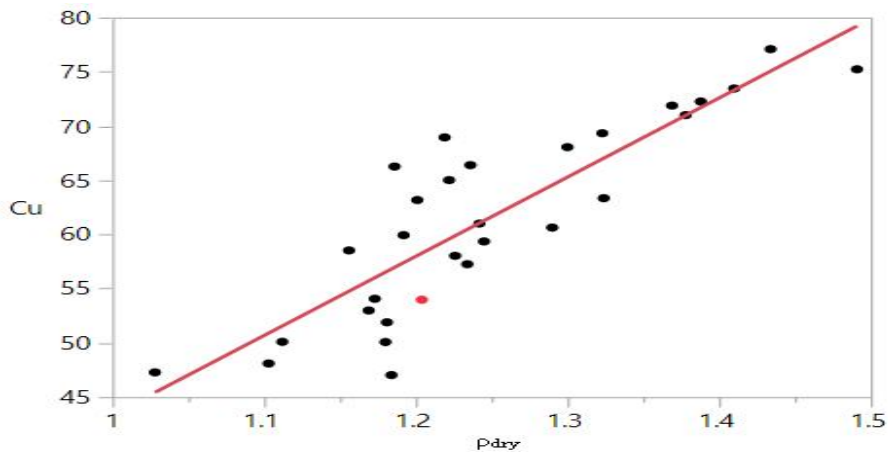


Fig. 7. Linear fit of  $\rho_{\text{dry}}$ -Cu

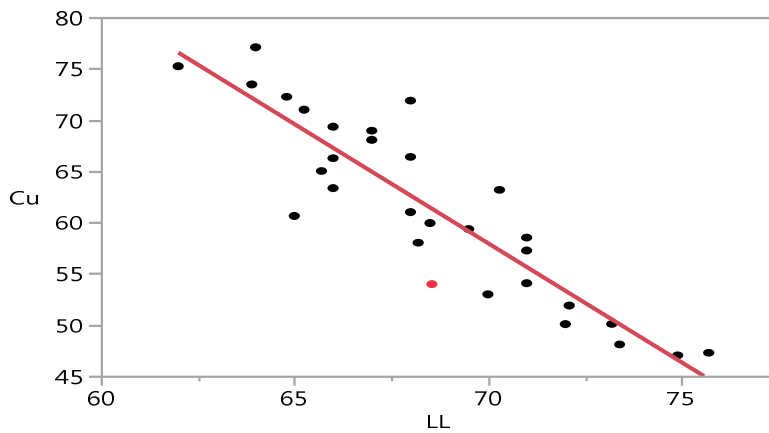


Fig. 8. Linear fit of LL-Cu

The details of the statistical output showed that the relationship developed between LL and Cu is significant ( $\alpha < 0.05$ ) and has a strong relationship.



### 3.2.1.6 Model-F

Model between Cu and Plastic Limit (PL)

$$\text{Cu} = 154.661 - 2.602 \cdot \text{PL}, R^2 = .676, \alpha = 0.000 < 0.05$$

The statistical output details indicated that the relationship developed between Cu and PL is significant ( $\alpha < 0.05$ ), and a strong correlation exists between the correlation variables.

### 3.2.1.7 Model-G

Model between Cu and Plasticity Index (PI)

$$\text{Cu} = 185.778 - 3.807 \cdot \text{PI}, R^2 = .330, \alpha = 0.001 < 0.05$$

The statistical output details indicated that the relationship developed between PI and Cu is significant ( $\alpha < 0.05$ ), but a weak relationship exists between the correlation variables.

### 3.2.1.8 Model-H

Model between Cu and liquidity Index (LI)

$$\text{Cu} = 65.606 - 0.22914 \cdot \text{LI}, R^2 = .029, \alpha = .367 > 0.05$$

The statistical output details indicated that the relationship developed between LI and Cu is insignificant ( $\alpha > 0.05$ ), and a weak relationship exists between the correlation variables. Table 3 illustrates that the developed single linear regression (SLR) models based on the level of the significance ( $\alpha$ ) and coefficient of determination ( $R^2$ ), Cu value has a strong relationship with LL,  $\rho_{\text{dry}}$ ,  $\rho_{\text{bulk}}$ , PL & NMC (i.e., from order 1 to 5). On the other hand, PI, Gs & LI (i.e., Orders from 6 to 8) indicated a weak relationship ( $R^2 < 0.5$ ) and insignificant level for Gs & LI (i.e.  $\alpha > 0.05$ ) with Cu. Those predictors were also good indicators to form better multiple linear regression analysis that could provide better models for the prediction of dependent variables of intensive area.

## 3.2.2 Multiple Linear Regression (MLR) analysis

Multiple Linear Regression (MLR) analysis is tried to model the relationship between two or more explanatory variables and a predicted variable by fitting an equation to experimental data. For this study, the stepwise regression analysis method of variable selection was

applied. For this section, the significance level and correlation coefficient of predictors on each other obtained from the single linear regression analysis and the scatter plot were used. For independent variables highly correlated (interdependent) to each other (i.e., Correlated at 0.50 or 0.60 and above), then one might decide to combine (aggregate) them into a composite variable or eliminate one or more of the highly correlated variables [13]. Spotting multicollinearity among a set of explanatory variables might not be easy. A useful approach is examining the variance inflation factors (VIFs) or the tolerances of the explanatory variables. Accordingly, VIFs above 10 or tolerances below 0.1 are seen as a cause of concern [12]. Moreover, Durbin-Watson used to examine multicollinearity of predictors with no concern for the value of 1 to 3. Hence, after going through several alternative groupings of predictors, a model which contains plastic limit (PL) and plasticity index (PI) with a good significance level (i.e.  $\alpha = 0$ ). The strongest determination coefficient ( $R^2 = 0.806$ ) is modeled and taken as the best model.

### 3.2.2.1 Model-1

Model of Cu with NMC and LI

The resulting Model after correlating Cu with NMC and LI is obtained from SAS JMP pro 13 & SPSS v22 outputs with its corresponding statistical parameters:  $\text{Cu} = 163.309 + .559 \cdot \text{LI} - 2.677 \cdot \text{NMC}$ ,  $R^2 = .781$ ,  $\alpha = 0.000 < 0.05$ , Tolerance = .685 > 0.2 & VIF = 1.461 < 10, Durbin-Watson = 2.629 ~ 2

The details of the statistical output of Model-1 indicated that the relationship developed between predictors with no concern for the value of 1 to 3. Hence, after going through several alternative groupings of predictors. This model contains plastic limit (PL) and plasticity index (PI) with a good significance level (i.e.  $\alpha = 0$ ), and the strongest determination coefficient ( $R^2 = 0.806$ ) is modeled and taken as the best model.

### 3.2.2.2 Model-2

Model of Cu with PL and LI

The resulting Model after correlating Cu with PL and LI is obtained from SAS JMP pro 13 & SPSS v22 outputs:  $\text{Cu} = 164.670 - 2.697 \cdot \text{PL} - .357 \cdot \text{LI}$ ,  $R^2 = .746$ ,  $\alpha = 0.000 < 0.05$ ,

Tolerance=.987 > 0.2 & VIF=1.013 < 10, Durbin-Watson=2.566~2. The details of the statistical out-put of Model-2 indicate that the relationship between Cu and PL and LI is significant ( $\alpha < 0.05$ ). Besides, the  $R^2$  value of the multiple regression analysis is improved than the  $R^2$  value of the individual parameters, i.e., PL and LI.

3.2.2.3 Model-3

Model of Cu with PL and PI

The details of the statistical output of Model 3 indicate that the relationship developed between Cu with PL and PI is significant ( $\alpha < 0.05$ ). The  $R^2$  value of the multiple regression analysis is improved than the  $R^2$  value of the individual parameters, i.e., PL, PI.  $Cu = 224.032 - 2.272 * PL - 2.485 * PI$ ,  $R^2 = .806$ ,  $\alpha = 0.000 < 0.05$ , Tolerance=.923 > 0.2 & VIF=1.084 < 10, Durbin-Watson=2.791~2

Based on Table 4, all models are good since all models are significant and the coefficient of determinations is strong. However, Model-3 is

the "best model" for the prediction of undrained shear strength (Cu) of the study area based on the relative correlation coefficient(R), determinant factor ( $R^2$ ) & significance level ( $\alpha$ ) of all developed models.

3.2.3 Validation of predicted value from the actual (measured) value of Cu

To check the validity of developed models, experimental values obtained from test samples (actual) should be compared with the predicted value [14]. Considering the acceptability of Model-3 as the best model, it can be used to approximate the study area's undrained shear strength parameter. Moreover, it is possible to understand from Fig.15. the relationship between the predicted and the measured (actual) value was strong.

In general, the scatter plot on Fig. 15 illustrated that the predicted Cu value scatters near the straight line, through which the actual and predicted Cu value is equal. However, there is a little bit of variation between the actual and the measured Cu.

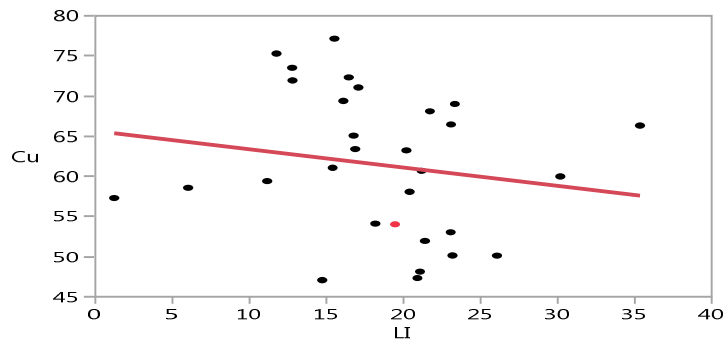


Fig. 9. Linear fit of LI-Cu

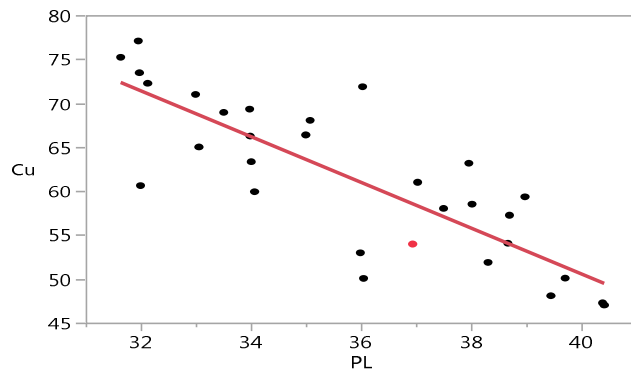


Fig. 10. Linear fit of PL-Cu

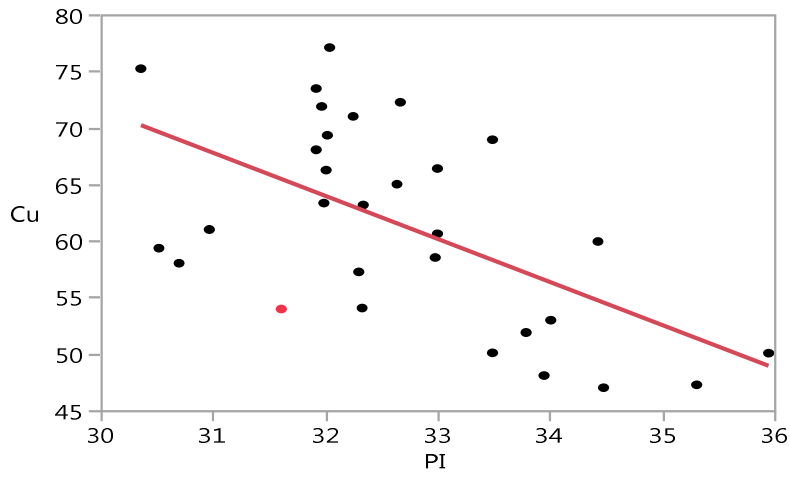


Fig. 11. Linear fit of PI-Cu

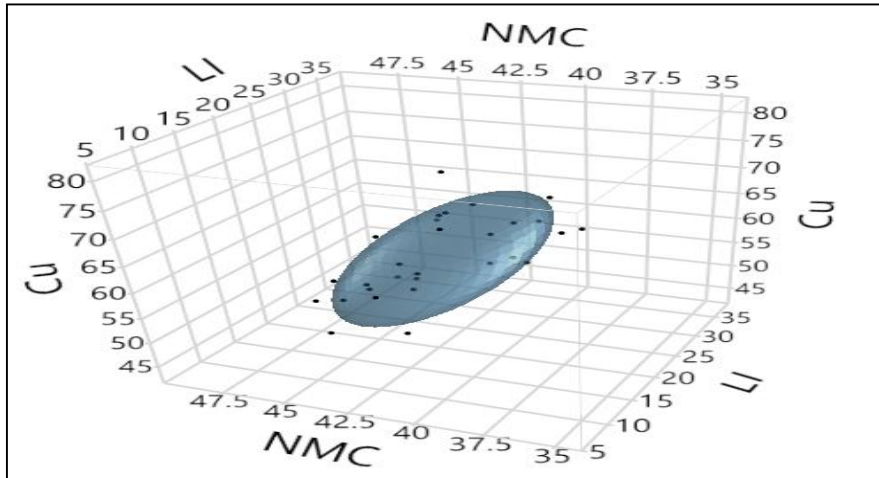


Fig. 12. Model-1 3D scatter plot

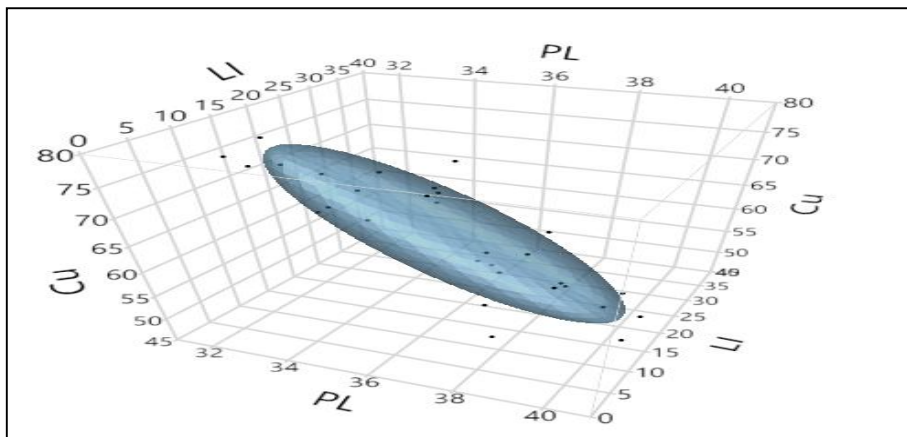


Fig. 13. Model-2 3D scatter plot

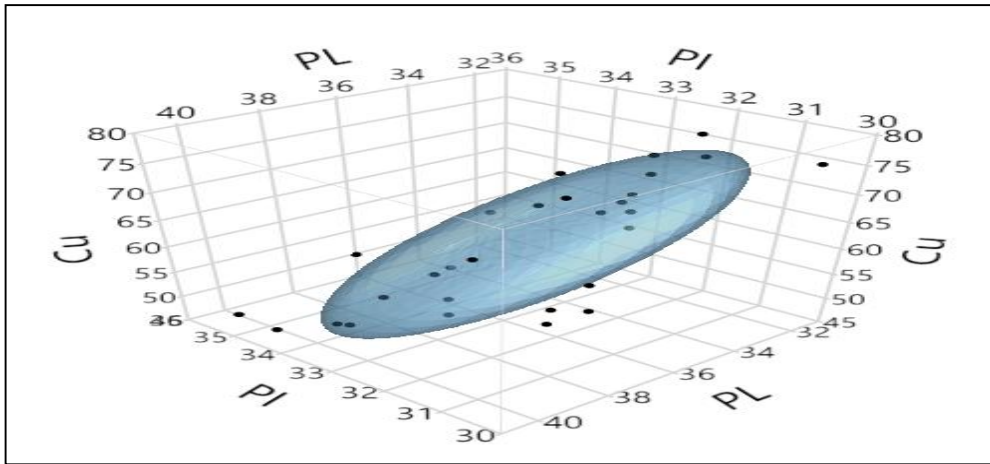


Fig. 14. Model-3 3D scatter plot

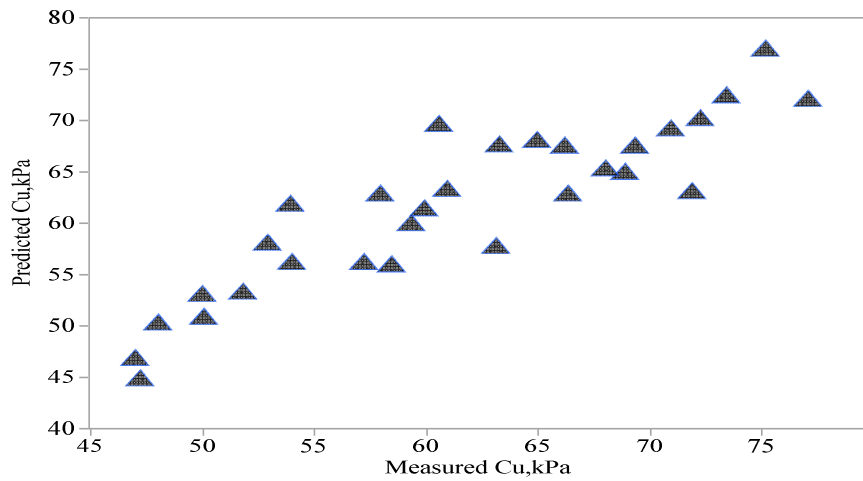


Fig. 15. Plots of predicted and actual values of undrained shear strength (Cu)

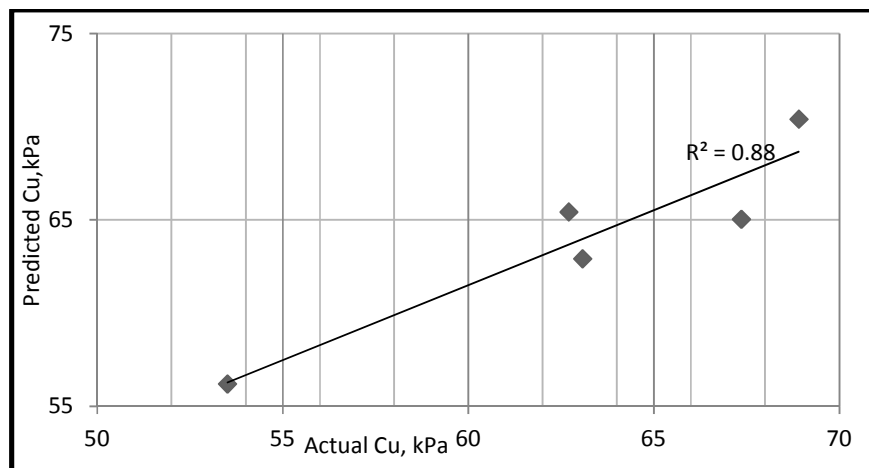


Fig. 16. Plots of predicted and actual values of undrained shear strength (Cu)

**Table 3. Summary of Single Linear Regression (SLR) models**

No.	Model name	SLR models from different variables	R <sup>2</sup>	Significance level, $\alpha$	Rank based on $\alpha$ and R <sup>2</sup>
1	Model-A	Cu = -75.786 + 52.014*Gs	0.035	0.322	7
2	Model-B	Cu = 148.515 - 2.078*NMC	0.662	0.00	5
3	Model-C	Cu=-54.278+ 65.452* $\rho_{bulk}$	0.698	0.00	3
4	Model-D	Cu = -29.670 + 73.031* $\rho_{dry}$	0.75	0.00	2
5	Model-E	Cu = 220.604 - 2.323*LL	0.805	0.00	1
6	Model-F	Cu = 154.661- 2.602*PL	0.676	0.00	4
7	Model-G	Cu = 185.778 - 3.807*PI	0.330	0.001	6
8	Model-H	Cu = 65.606 - 0.22914*LI	0.029	0.367	8

**Table 4. Summary of Multiple Linear Regression (MLR) models**

No.	Model name	MLR models from different variables	R	R <sup>2</sup>	$\alpha$	Rank based on R, $\alpha$ and R <sup>2</sup>
1	Model-1	Cu= 163.309+.559* LI -2.677*NMC	.884	0.781	0.0	2
2	Model-2	Cu= 164.670-2.697* PL-.357*LI	.864	0.746	0.0	3
3	Model-3	Cu= 224.032-2.272* PL-2.485*PI	.898	0.806	0.0	1

### 3.2.4 Validation of Predicted Value with additional test results

The validation of the developed model is led by using these test results of the study area. Subject to the relative correlation coefficient(R), determinant factor (R<sup>2</sup>) & significance level( $\alpha$ ), Model-3 (i.e. Cu= 224.032-2.272\* PL-2.485\*PI) is preferred among the different alternative models discussed & developed above. Consequently, from Fig.16 relation of measured (actual) and the predicted value of Cu is exhibited a little variation (R<sup>2</sup>=0.88).

## 4. CONCLUSION

It was studied the statistical analysis of undrained shear strength from index properties of Fine-Grained Soil. For this purpose, different models were developed to predict Cu value from Gs, NMC,  $\rho_{bulk}$ ,  $\rho_{dry}$ , LL, PL, and PI & LI. Consequently, the best Model from all with a better coefficient of determination (R<sup>2</sup> = 0.806), good significance level, and less Std. error were obtained from multiple linear regression (MLR) analysis as below:

Cu=224.032-2.272\*PL-2.485\*PI, R<sup>2</sup>=.806, p-value =0.000 < 0.05, Tolerance =.923> 0.2 & VIF=1.084< 10, Durbin-Watson=2.791~2.

From the study result, undrained shear strength parameters were significantly correlated with

liquid limit, plastic limit, bulk density, dry density, natural moisture content, and plasticity index. In contrast, it was not significantly correlated with this study area soil's specific gravity and liquidity index.

The research result would be beneficial for individuals, researchers, consultants, contractors, municipal and other government agencies involved in building construction and other structures in the study area. It is also suggested to do such statistical modeling on other areas of the country by different agencies and researchers.

Furthermore, it is desirable to conduct comparative statistical modeling between undisturbed and remolded soils to get Cu value from soil index properties.

## CONSENT

It is not applicable.

## ETHICAL APPROVAL

It is not applicable.

## ACKNOWLEDGEMENT

The authors would like to thank Assosa University, Jimma Institute of Technology, Jimma

University, and the Ethiopian Road Authority (ERA) for allowing this work to progress. This work may not be achieved without supportive sponsoring.

### COMPETING INTERESTS

Authors have declared that no conflicts of interest exist.

### REFERENCES

1. Rajeev, Jain, Pramod, Sharma and Sanja, Bhandari. Unconsolidated Undrained Shear Strength of Remolded Clays by Anns Technique. International Journal of Engineering Research and Technology (IJERT).2013;2(6):2827-2832.
2. Zumrawi, Magdi M. E. and Mohammed, Lina A. D. Correlation of Placement Conditions and Soil Intrinsic Properties with Shear Strength of Cohesive Soils. 7th Annual Conference for Postgraduate Studies and Scientific Research - Basic Sciences and Engineering Studies. Khartoum; 2016.
3. Teferra, Alemayehu and Leikun, Mesfin. Principle of Foundation Engineering. Addis Ababa: Addis Ababa university press; 1999.
4. Budhu, Muni. Soil Mechanics and Foundations. 3rd Ed. New York: John Wiley & Sons, Inc; 2011.
5. Gunaratne, Manjriker. The Foundation Engineering Handbook. New York: CRC Press; 2006.
6. Ethiopian Building Code Standard. Addis Ababa: Ministry of Urban Development and Construction; 2013.
7. Roy, Surendra and Das, Gurcharan. Statistical models for the prediction of shear strength parameters at Sirsa, India. International Journal of Civil and Structural Engineering. 2014;4(4):483-498.
8. Andualem, Gedeon. Developing Correlation between Dynamic Cone Penetration Index and Undrained Shear Strength of Soils Found in Debre Markos Town. Addis Ababa University, Addis Ababa Institute of Technology. Addis Ababa: A master's thesis presented to School of graduate studies, Addis Ababa University; 2015.
9. Adunoye, G.O. Fines Content and Angle of Internal Friction of a Lateritic Soil: An Experimental Study. American Journal of Engineering Research. 2014;3(3):16-21.
10. Park, Hun Myoung. Univariate Analysis and Normality Test Using SAS, STATA, and SPSS. The Trustees of Indiana University; 2006.
11. Molavi, H Vardanjani, et al.. Cleansing and preparation of data. J Oral Health Oral Epidemiol. 2016; 5(4):171-185.
12. Sabine Landau and Everitt, Brian S. A Handbook of Statistical Analyses using SPSS. New York: A CRC Press Company; 2004.
13. Leech, L. Nancy, Barrett, Karen C. and Morgan, George A. SPSS for Intermediate Statistics: Use and Interpretation. 2<sup>nd</sup> ed. London: Lawrence Erlbaum Associates, Inc; 2005.
14. Khalid, Usama, et al.. Prediction of Unconfined Compressive Strength from Index Properties of Soils. Sci. Int. (Lahore). 2015;27(5):4071-4075.

© 2021 Bakala et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Peer-review history:*  
*The peer review history for this paper can be accessed here:*  
<http://www.sdiarticle4.com/review-history/66257>