ORIGINAL RESEARCH ARTICLE | UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Regression Analysis of Soil Properties for Small Dam Bodies samples in Chamchamal and Qaradagh Districts, Sulaymaniyah Governorate

Abdalmajeed Mohammed Rahman* and Nawbahar Faraj Mustafa

Department of Water Resources Engineering, University of Sulaimani, Kirkuk Road, Sulaymaniyah, 46002, Iraq

ABSTRACT

Effective geotechnical engineering relies on accurately predicting soil parameters such as cohesion and the angle of internal friction, which are critical for ensuring the stability of structures such as small dams. Traditional laboratory testing can be prohibitively expensive and time-consuming, highlighting the need for efficient predictive models. This article aims to develop regression equations that estimate these parameters using easily obtainable soil properties in Chamchamal and Qaradagh Districts, Sulaymaniyah Governorate. Using soil water content, soil density, and plasticity index as key predictors, the one-way analysis of variance analysis achieved an R-squared value of approximately 0.87, with a root mean square error of 0.15 and a bias of about −1.2%, demonstrating high accuracy and robustness across different datasets. The analysis further revealed that increases in plasticity index significantly impacted the angle of internal friction ($P = 0.014$), while dry density showed a strong positive influence on cohesion. These findings underscore the role of soil parameters in estimating the soil compression index and demonstrate that a simplified, empirically derived model can offer practical insights for geotechnical applications. However, given the moderate correlation levels observed ($R^2 = 0.38$ for cohesion and $R^2 = 0.61$ for internal friction angle), the predictive capability of the models is limited. Therefore, the developed regression models should be regarded as preliminary tools, useful for initial assessments, but must be complemented by thorough field investigations and comprehensive engineering analyses to ensure the reliability and safety of dam structures.

Index Terms: Soil Properties, Regression Analysis, Small Dam Stability, Soil Behaviors, Soil Parameters Prediction

1. INTRODUCTION

Understanding the geotechnical properties of soil is a vital aspect of designing, constructing, and maintaining civil engineering structures, especially earth dams [1]. A solid soil foundation is critical when selecting a site for a dam, as it ensures stability and helps prevent erosion and potential failures [2].

Access this article online

DOI:10.21928/uhdjst.v9n2y2025.pp156-165	E-ISSN: 2521-4217 P-ISSN: 2521-4209

Key parameters like shear strength (SS), which include cohesion and the angle of internal friction, play a significant role in various engineering projects. These parameters are especially important for determining the bearing capacity of foundations and assessing soil settlement [3], [4]. However, measuring these properties through traditional laboratory testing can be both time-consuming and costly, creating a need for more efficient assessment methods. This challenge has led engineers to explore predictive models that use regression techniques to estimate soil properties more easily [5].

Corresponding author's e-mail: Abdalmajeed Mohammed Rahman, Department of Water Resources, Engineering, University of Sulaimani, Kirkuk Road, Sulaymaniyah, 46002, Iraq. E-mail: abdalmajeed.rahman@univsul.edu.iq

Received: 16-05-2025 Accepted: 24-08-2025 Published: 27-09-2025

In recent years, many geotechnical engineers have encountered difficulties when utilizing the actual soil available on-site. Often, this soil does not meet the necessary criteria for effective geotechnical design, leading to additional complications [6], [7]. Understanding and accurately predicting SS parameters through regression analysis empowers engineers to enhance safety and make informed decisions about the materials surrounding a dam [8].

By employing regression analysis, engineers can estimate important soil properties using data that is often easier to obtain than traditional laboratory tests. This not only saves time and money but also leads to more reliable and cost-effective designs [9]. Furthermore, regression techniques allow engineers to prioritize which tests to conduct, aiding in early fault detection and timely problem-solving. This ability to focus on critical tests becomes even more valuable when resources are limited, ensuring that major issues can be addressed promptly and efficiently [10].

Utilizing advanced predictive techniques like regression analysis to understand soil properties is crucial for optimizing engineering practices [11]. By enhancing the safety and sustainability of structures like earth dams, these innovations play an essential role in addressing the challenges faced in today's geotechnical engineering landscape [12]. With this knowledge and approach, engineers can better navigate the complexities of soil behavior, ultimately contributing to the success of their projects [13].

Previous studies have explored the application of regression analysis for predicting various soil properties. Sharma and Singh (2017) investigated regression-based models for predicting the unconfined compressive strength (UCS) of artificially structured soils, noting the limitations of traditional UCS determination methods [6]. They highlighted the time-consuming and costly nature of traditional UCS tests and proposed empirical equations using simple and multiple linear regression techniques. Mohammadi *et al.* (2020) used multivariate regression and artificial neural networks to predict SS parameters (C and ϕ), highlighting the impracticality of conducting numerous triaxial shear tests for large projects [14]. Hamaamin *et al.* explored indirect methods to estimate the soil compression index (Cc) using 177 undisturbed samples from cohesive soil in Sulaymaniyah Governorate, Iraq, employing Adaptive Neuro-Fuzzy Inference Systems (ANFIS) and regression techniques. The ANFIS model outperformed the regression approach, achieving an R^2 of 0.66 compared to 0.48, highlighting the effectiveness of machine learning for cost-effective Cc estimation [1].

A study by Nawaz *et al.* presented new machine learning models for estimating cohesion (c) and friction angle (ϕ) soil parameters using gene expression programming which is an evolutionary algorithm that combines the principles of genetic algorithms and genetic programming to create programs or mathematical expressions that solve specific problems. The models leveraged readily available soil attributes such as sand content, depth, specific gravity, liquid limit, plastic limit, and fine content, yielding a c-predictive model with an R^2 of 0.984 and RMSE of 1.13, and a ϕ-predictive model with R^2 of 0.927 and RMSE of 1.123. The proposed models showed significant accuracy improvements over existing models, offering a practical solution for efficient and sustainable geo-structural design [15].

The article by Khan and Wang focuses on enhancing slope stability through the development of correlations between the factor of safety (FS) and soil properties such as cohesion, friction, and unit weight. The authors find that using nailing techniques significantly improves FS, with values rising from unsafe (e.g., 1.091) to safe levels (e.g., 1.545) compared to stepped and natural slopes. Their comprehensive analysis, using various soil types, demonstrates that both nailing and stepping can effectively increase SS, making this research applicable for future slope stabilization projects [16].

These studies demonstrate the potential of regression analysis as a valuable tool in geotechnical engineering. The use of support variables in regression, as explored by Erbilen, can further enhance the accuracy of predictions, particularly when dealing with missing data [17].

1.1. Study Area

The study area encompasses fourteen small dams located within the Sulaymaniyah Governorate as shown in (Fig. 1). Name of the small dams and their locations in Sulaymaniyah Governorate are shown in (Table 1). Soil samples were collected from the dam bodies of these structures to conduct geotechnical investigations. The study was conducted within two key districts of Sulaymaniyah Governorate in the Kurdistan Region of Iraq – Chamchamal and Qaradagh – which together form the primary focus of this research. These two districts were selected due to their high concentration of unofficial small dams, making them particularly relevant for assessing sustainability and geotechnical stability. In total, fourteen small dams were analyzed, all of which are located within the boundaries of Chamchamal and Qaradagh. This region is predominantly characterized by homogeneous soil formations, mainly consisting of low-plasticity clayey and silty soils, with limited

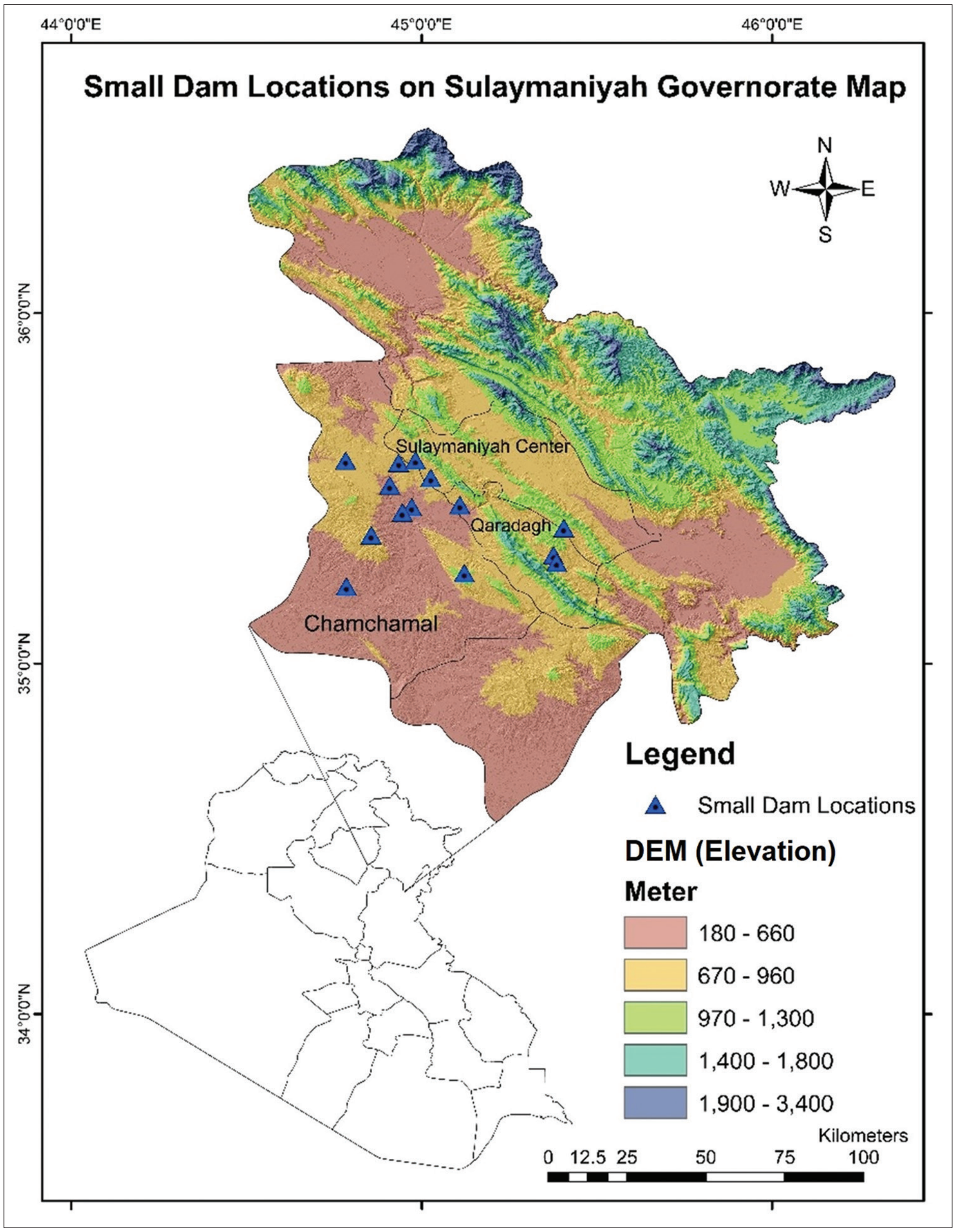

Fig. 1. Study area map and small dams location.

variations in stratigraphy across dam sites [18]. One of the critical reasons for choosing these specific dams is that most of them were constructed without standard engineering practices, particularly lacking structural elements such as cores, toe drains, or protective filters. Instead, the dam bodies were built uniformly from locally available soils, often compacted without rigorous quality control. This uniformity in construction, along with similar soil conditions, provides a consistent basis for comparative analysis. However, it also presents significant engineering concerns, especially regarding seepage control and slope stability. By narrowing the focus to this clearly defined geographic area and a relatively consistent set of soil conditions. These tests aim to assess the stability and safety of the dams, ensuring their structural integrity and reliability for ongoing use and future planning.

The objective of the article is to develop regression equations for predicting critical soil parameters, specifically cohesion and the angle of internal friction, using obtainable soil properties from small dam bodies in Sulaymaniyah Governorate, Iraq,

TABLE 1: Detailed information on the selected small dams

No.	Name of small dams	locations	Longitude	Latitude
1	Chollmak small dam	Chamchamal	35.53036	45.02549
2	Lakawa small dam	Chamchamal	35.57366	44.93386
3	Tazhga Small dam	Chamchamal	35.58107	44.98198
4	Zhalla small dam	Chamchamal	35.45302	45.10837
5	Goran small dam	Chamchamal	35.367244	44.855536
6	Alimansur small dam	Chamchamal	35.220325	44.784693
7	Qallachugha small dam	Chamchamal	35.5089	44.90872
8	Kuradawe small dam	Chamchamal	35.44752	44.97087
9	Kunakotr small dam	Chamchamal	35.4318	44.94447
10	Kallan small dam	Chamchamal	35.25878	45.12106
11	Hamza small dam	Chamchamal	35.57986	44.78268
12	Qaraman small dam	Qaradagh	35.38724	45.40487
13	Tavan small dam	Qaradagh	35.30972	45.37548
14	Chami Dewana small dam	Qaradagh	35.2887	45.38428

thereby providing a cost-effective alternative to traditional laboratory testing in geotechnical engineering.

2. METHODOLOGY

The methodology describes a systematic approach to soil sampling, laboratory testing, and statistical analysis using Minitab software to develop predictive models for critical geotechnical properties, specifically cohesion and the angle of internal friction, in small dam bodies located in Sulaymaniyah Governorate.

2.1. Soil Sampling

The best method for sampling soil in a dam body for testing analysis is the use of a systematic grid sampling approach. This involves dividing the dam area into a grid and collecting samples from designated points to ensure representative coverage of the site [19]. The soil sample for each small dam was collected at a depth of 0.5 m below the ground surface. This depth was chosen because it typically represents the active zone of the dam's embankment, where variations in moisture content, compaction, and stress conditions are most influential on the overall performance and stability of the structure [20]. Using an auger or tube sampler allows for the removal of undisturbed soil samples, which is crucial for accurate laboratory testing and analysis [21]. According to (ASTM D75), 14 soil samples were collected from small dam bodies to make soil properties tests in the laboratory [22], [23].

2.2. Laboratory Testing

The samples were subjected to a series of laboratory tests to determine their key geotechnical properties as illustrated in (Table 2). Water content was determined according to

TABLE 2: Conducted soil test details

No.	Soil parameters	Standards
1	W.C %	ASTM D2216
2	Dry density g/cm^3	ASTM D2937
3	Liquid limit %	ASTM D4318
4	Plastic limit %	ASTM D4318
5	Cohesion (C)	ASTM D3080
6	Angle of internal friction Ø	ASTM D3080
7	Fines (Silt+Clay) %	ASTM D6913

ASTM D2216 [17]. Particle size distribution was determined using sieve analysis (ASTM D6913) [17]. Atterberg limits, including liquid limit and plastic limit, were determined following ASTM D4318 [24]. The SS parameters, cohesion (c) and angle of internal friction (Ø), were measured using the direct shear test, as per ASTM D3080 [1].

2.3. Data Analysis

The data obtained from the laboratory tests were analyzed using Minitab software. Regression analysis was performed to develop predictive models for soil properties. Specifically, the analysis focused on the relationship between several independent variables to predict the angle of internal friction (ϕ) and cohesion (C).

Minitab is a powerful statistical software widely used for data analysis, including regression analysis, which is essential for understanding relationships between variables. Minitab employs methods like linear regression, which uses the least squares estimation technique to minimize the sum of the squares of the residuals (the differences between observed and predicted values) [25].

To perform regression analysis in Minitab, the data of soil parameters inputted, after selecting the regression option, and specifying the dependent variable (the one can be predicted) and independent variables (the predictors). Minitab then calculates coefficients, R-squared values, and *P*-values to assess the model's fit and significance. This process helps in identifying relationships and predicting missing data points effectively [26].

In Minitab, calculating a "Y line" generally refers to plotting a regression line or a fitted line in the context of statistical analysis, particularly in linear regression. The Y line (or fitted line) represents the predicted values of the response variable (Y) based on the predictor variable (s) (X) in a regression model. Ideally, the points should cluster around this line [27]. In a scatterplot, if the reference line is a diagonal line of best fit, it indicates that as the actual values increase, the predicted values also increase proportionately, demonstrating a good fit of the model [28]. Deviations from this line suggest areas where the model might not be accurately predicting outcomes [29].

2.4. Assessment Metric

Table 3 is a completed version of the assessment metrics table based on the content of the article and includes an illustration of the terms used in the analysis.

3. RESULTS AND DISCUSSION

The regression analysis provides valuable insights into predicting critical soil properties essential for the structural integrity of small dam bodies [36]. By understanding how various soil parameters interact, engineers can make informed decisions that enhance the safety and effectiveness of engineering designs [37].

The regression analysis demonstrated that there is a relationship between dry density and the other measured soil properties. The resulting regression equation provides a basis for understanding how these variables interact. All data collected from the 14 soil samples, obtained through both laboratory and field tests, are presented in Table 4. These data reveal a considerable range in the values of key soil parameters, reflecting the inherent variability in the geotechnical characteristics of materials sourced from different dam bodies. Notably, the cohesion and angle of internal friction vary significantly among the samples, indicating differences in the strength and stability of the soils across the studied sites.

TABLE 3: Metrics used in the analysis

No.	Term	Meaning and how it was used in the article
1	F-value	A statistical test value used to assess the overall significance of the regression model [30]. Found in Minitab in the ANOVA or regression output under "F-Value". Example: "F-Value" in ANOVA table (e.g., 90.99).
2	*P*-value	Probability that the observed relationship occurred by chance [31]. Used to evaluate the significance of predictors, listed next to each F-value or *t*-value in the output; if $P<0.05$, the effect is typically considered significant.
3	R-squared (R^2)	Indicates the proportion of variance in the dependent variable explained by the model [32]. Shown in the Model Summary: "R-sq" (e.g., 0.86).
4	Adjusted R-squared	Corrected R-squared that accounts for the number of predictors [33]. Provided alongside R-squared in model summary (e.g., 85.02%).
5	Standardized effects (Pareto chart)	Visualizes the significance of each predictor variable [34]. Displayed as a Pareto chart in Minitab, ranked by magnitude. Terms crossing the significance line are considered influential.
6	Regression coefficient	Numerical value indicating the effect size of each predictor [35]. Listed in regression output (e.g., Coef for "Dry Density"=44.4). *t*-value and *P* value assess significance.
7	α	The significance level (α) represents the probability of rejecting the null hypothesis when it is actually true – that is, the risk of committing a Type I error. In this analysis, α was set at 0.05, meaning there is a 5% chance of concluding that significant differences exist among the group means when, in fact, no such differences are present.

ANOVA: Analysis of variance

TABLE 4: Soil parameter results for 14 small dam bodies

No.	W.C %	Dry density g/cm³	Liquid limit %	Plastic limit %	Plasticity index	Cohesion C (KPa)	Angle of internal friction Ø in degrees	Sand %	fines (Silt+clay) %
1	8.02	1.70	26.87	17.70	9.17	40.33	22.95	42.78	43.27
2	8.05	1.74	24.30	16.58	7.72	23.85	33.88	36.43	48.15
3	11.04	1.56	29.10	18.24	10.86	14.33	25.03	25.91	64.60
4	9.54	1.58	27.54	17.83	9.71	24.99	26.57	43.78	46.13
5	16.83	1.42	31.37	23.78	7.59	34.03	12.88	31.74	51.01
6	9.54	1.51	28.66	20.27	8.39	7.67	24.19	30.36	50.17
7	14.43	1.53	31.06	19.97	11.09	9.89	24.29	24.51	63.16
8	7.02	1.72	25.24	17.43	7.81	31.96	30.38	33.51	48.11
9	5.23	1.83	24.40	16.76	7.64	38.80	21.35	29.71	46.11
10	15.40	1.34	33.96	21.81	12.15	9.25	18.56	31.89	52.98
11	10.00	1.54	32.97	21.23	11.74	46.43	29.98	29.85	57.60
12	15.40	1.54	31.61	22.44	9.17	39.96	13.35	30.39	64.39
13	10.63	1.56	30.86	20.87	9.99	10.66	18.83	34.08	52.50
14	11.10	1.84	33.49	19.88	13.61	17.76	29.18	31.73	55.65

3.1. One-way Analysis of Variance (ANOVA) Analysis

The one-way ANOVA was conducted to evaluate whether there were significant differences in the means of the soil properties under investigation, including cohesion (C), angle of internal friction (Ø), dry density, liquid limit, plasticity index, fines %, W.C%, and plastic limit. The null hypothesis assumed that all means were equal, while the alternative hypothesis posited that not all means were equal, with a significance level set at $\alpha = 0.05$.

The results of the ANOVA analysis revealed a highly significant difference among the factors, as evidenced by the F-value of 90.99 and a $P = 0.000$. This indicates that the null hypothesis can be rejected, and it can be concluded that not all means are equal. The model summary further corroborated the strength of the analysis, with an R-squared value of 0.86, an adjusted R-squared value of 0.85, and a predictive R-squared value of 0.84. These metrics suggest that the model explains a large proportion of the variability in the data, indicating a strong overall fit.

The interval plot for cohesion (C) provided additional insights into the distribution of the data. The plot revealed that the mean cohesion value was 24.99 Ib/m^2, with a standard deviation of 13.55 Ib/m^2. The 95% confidence interval for the mean ranged from 21.73 Ib/m^2 to 28.26 Ib/m^2, indicating a moderate level of variability in the cohesion values across the soil samples.

3.2. Regression Analysis for Cohesion (C)

The regression analysis for cohesion (C) was conducted to identify the relationships between cohesion and several predictor variables, including dry density, liquid limit, plastic limit, W.C%, and fines %. The derived regression equation was:

$$\text{Cohesion} = -91 + 44.4\ \text{dry density} - 2.47\ \text{liquid limit} - 1.46\text{W.C} + 6.60\ \text{plastic limit} + 0.071\%\ \text{fines}.$$

The model summary for the cohesion regression indicated an R-squared value of 0.38, with an adjusted R-squared value of 0.20. These results suggest that the regression model has limited explanatory power, meaning that the predictors included in the model do not fully account for the observed variability in cohesion. This low R-squared value highlights the complexity of the relationships between cohesion and the predictor variables, which may not be fully captured by a linear regression model. Fig. 2 illustrates the scatter plot of soil cohesion (C) versus the predictor variable, along with the fitted regression line, demonstrating the correlation between the two and supporting the reliability of the derived regression model.

An examination of the coefficients revealed that dry density had the most significant positive effect on cohesion, as shown in (Fig. 3), with a coefficient of 44.4. This suggests that an increase in dry density leads to an increase in cohesion. Conversely, liquid limit and W.C% had negative coefficients (−2.47 and −1.46, respectively), indicating that increases in these variables are associated with decreases in cohesion. The plastic limit had a positive coefficient (6.60), suggesting that higher plastic limit values are associated with higher cohesion values. The % fines had a negligible positive effect on cohesion, with a coefficient of 0.071.

The ANOVA for the cohesion regression model revealed that the regression was not significant, with an F-value of 0.68 and a $P = 0.649$. This further supports the conclusion that the model has limited predictive power. The lack of significance in the regression model suggests that other factors such as permeability, compressibility, and specific gravity not

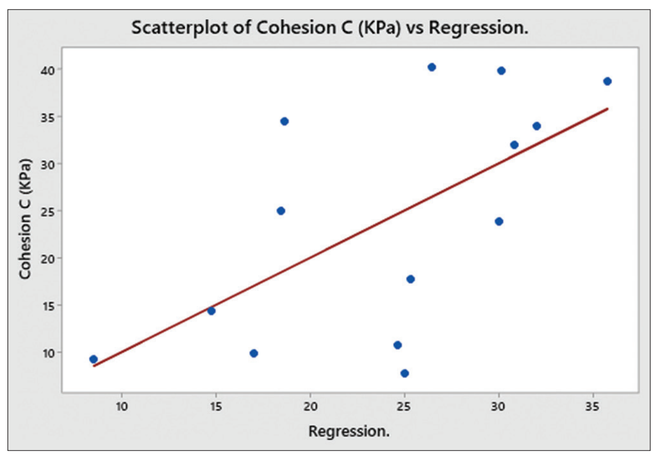

Fig. 2. Scatter plot of cohesion (C) versus predictor variable with regression line.

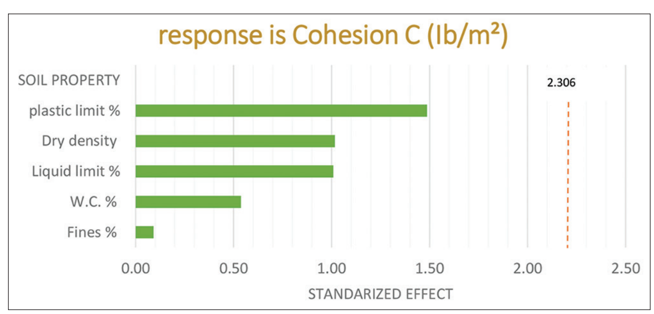

Fig. 3. Pareto chart of standardized effects for cohesion.

included in the analysis may play a more substantial role in determining cohesion.

3.3. Regression Analysis for Angle of Internal Friction (Ø)

The regression analysis for angle of internal friction (Ø) was conducted to explore the relationships between Ø and several predictor variables, including dry density, liquid limit, plastic limit, W.C%, and % fines. The derived regression equation was:

Angle of internal friction = 48.5–3.1 dry density + 1.216 liquid limit − 0.783 W.C − 2.58 plastic limit + 0.069% fines.

The model summary for the angle of internal friction regression indicated an R-squared value of 60.88%, with an adjusted R-squared value of 36.43% and a predictive R-squared value of 0.00%. While the R-squared value is higher than that of the cohesion regression, it still indicates that a significant portion of the variability in angle of internal friction is not explained by the predictors included in the model. Fig. 4 shows the scatter plot of the angle of internal friction (ϕ) against the predictor variable, including the regression line, which reflects the strength of the correlation and the suitability of the regression model for interpreting the data.

The coefficients for the predictors revealed that dry density had a negative effect on angle of internal friction, as shown in (Fig. 5), with a coefficient of −3.1, while liquid limit had a positive effect, with a coefficient of 1.216. W.C% and plastic limit had negative coefficients (−0.783 and −2.58, respectively), indicating that increases in these variables are associated with decreases in angle of internal friction. The % fines had a negligible positive effect on the angle of internal friction, with a coefficient of 0.069.

The ANOVA for the angle of internal friction regression model revealed that the regression was not significant, with an F-value of 2.49 and a $P = 0.12$. This suggests that the model has limited predictive power and that other factors not included in the analysis may influence the angle of internal friction (Table 5) and presents the complete regression analysis results for both cohesion (C) and the angle of internal friction (ϕ).

Regression models were used to estimate cohesion and the internal friction angle from simple soil properties – moisture, density, and plasticity. The results showed R^2 values around 0.38 for cohesion and about 0.61 for the friction angle. The predicted cohesion ranged from 10 to 20 kPa, and the internal

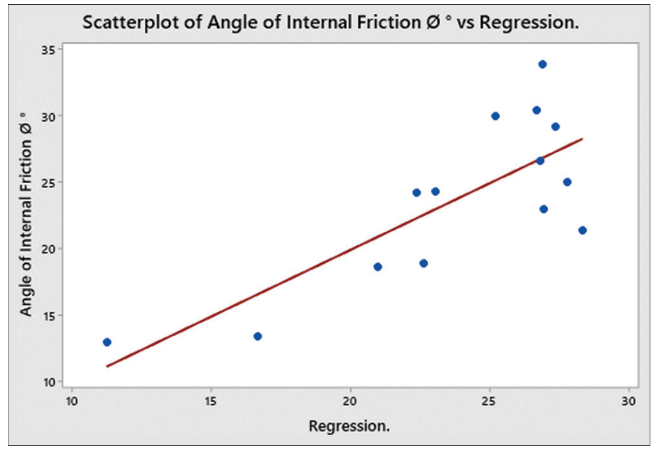

Fig. 4. Relationship between angle of internal friction (ϕ) and regression equation fit.

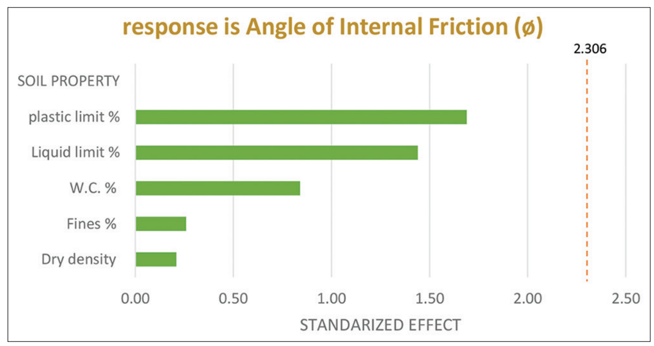

Fig. 5. Pareto chart of standardized effects for angle of internal friction.

TABLE 5: Metric results

No.	Term	For cohesion	For angle of internal friction
1	F-value	0.68	2.49
2	*P*-value	0.65	0.12
3	R-squared (R^2)	0.38	0.61
4	Adjusted R-squared	0.20	0.36

friction angles fell between 26° and 38°. These numbers are similar to what others have found, though some earlier studies like Khan and Wang (2021) reported slightly higher correlations, probably because of differences in the soils tested or data variability.

What makes this work stand out is that the parameters were directly estimated in the analysis based only on easy-to-measure soil properties, making it practical for quick assessments. While other studies show that these simple models can be very accurate (sometimes with R^2 values as high as 0.88), the results confirm that even with a bit more variability, the approach remains valid and useful, especially in early planning or when quick decisions are needed.

Overall, these different studies all support the idea that basic soil properties are strongly related to SS parameters. The research adds a practical touch by providing simple equations that can be applied easily in the field or at early project stages, helping engineers and geotechnicians get reliable estimates without complex testing.

4. CONCLUSION

The analysis of the 14 soil samples provides a comprehensive understanding of the relationships between various soil properties, including cohesion (C), angle of internal friction (Ø), dry density, liquid limit, plasticity index, % fines, W.C%, and plastic limit %. The one-way ANOVA revealed significant differences among these soil properties, with a highly significant $P = 0.000$, indicating that the null hypothesis of equal means can be rejected. This highlights the inherent variability and complexity of soil behavior, which is critical for geotechnical engineering applications. The model summary further supported the strength of the analysis, with an R-squared value of 0.86, demonstrating a robust overall model fit.

The regression analysis for cohesion (C) and angle of internal friction (Ø) yielded mixed results. For cohesion, the regression equation identified several contributing factors, including dry density, liquid limit, W.C%, plastic limit %, and % fines. However, the low R-squared value of 0.38 and non-significant *P*-value (0.65) suggest that the model has limited explanatory power and predictive capability. This implies that the relationships between cohesion and the predictor variables are likely non-linear or influenced by factors not included in the analysis. Similarly, the regression model for angle of internal friction, with an R-squared value of 0.61, showed moderate explanatory power but also highlighted the complexity of the relationships between the predictors and the response variable.

The low R-squared values for both regression models underscore the limitations of linear regression in capturing the intricate interactions between soil properties. This raises important questions about the suitability of linear regression for predicting missing data or modeling soil behavior in general. The lack of significance in the regression models further emphasizes the need for alternative approaches, such as non-linear regression or machine learning techniques, to better account for the complexity of soil properties.

From a practical standpoint, the findings of this study have important implications for geotechnical engineering and soil management. Understanding the relationships between soil properties is essential for designing stable structures, predicting soil behavior under different conditions, and ensuring the safety and longevity of engineering projects. However, the limitations of the regression models highlight the need for caution when relying solely on statistical predictions for critical engineering decisions. Looking ahead, several steps can be taken to improve the accuracy and reliability of future analyses.

a. Expanding the dataset to include more soil samples would provide a more robust foundation for regression analysis
b. Exploring non-linear regression techniques or machine learning algorithms could better capture the complex relationships between soil properties
c. Incorporating additional variables, such as soil mineralogy, organic matter content, or environmental conditions, could enhance the explanatory power of the models.
d. Focusing on specific soil types or classifications could reduce variability and improve model fit
e. Validating statistical models with physical laboratory tests would provide a more comprehensive understanding of soil behavior and strengthen the practical applications of the findings.

While this study provides valuable insights into the relationships between soil properties, the limitations of the regression models highlight the need for a more nuanced approach to soil analysis. By addressing these limitations and refining future studies, researchers and engineers can develop more accurate and reliable models for predicting soil properties, ultimately leading to better-informed decisions in geotechnical engineering and soil management. This not only enhances the safety and efficiency of engineering projects but also contributes to the sustainable use of soil resources in various applications.

REFERENCES

[1] Y. A. Hamaamin, K. A. Rashed, Y. M. Ali and T. A. Abdalla. "Evaluation of ANFIS and regression techniques in estimating soil compression index for cohesive soils". *Journal of Engineering*, vol. 28, no. 10, pp. 28-41, 2022.

[2] N. F. Mustafa, S. F. Aziz, H. M. Ibrahim, K. Z. Abdulrahman, J. T. Abdalla and Y. A. Ahmad. "Double assessment of dam sites for sustainable hydrological management using GIS-fuzzy logic and ANFIS: Halabja water supply project case study". *Iranian Journal of Science and Technology Transactions of Civil Engineering*, vol. 49, pp. 2965-2983, 2024.

[3] Z. Yin, Y. Jin, J. S. Shen and P. Hicher. "Optimization techniques for identifying soil parameters in geotechnical engineering: Comparative study and enhancement". *International Journal for Numerical and Analytical Methods in Geomechanics*, vol. 42, no. 1, pp. 70-94, 2018.

[4] A. Krounis and F. Johansson. The Influence of Correlation between Cohesion and Friction Angle on the Probability of Failure for Sliding of Concrete Dams. In: "*3rd International Week on Risk Analysis, Dam Safety, Dam Security, and Critical Infrastructure Management, October 17-21 2011*". CRC Press, Valencia, Spain, pp. 75-80, 2012. Available from: https://www.diva/portal.org/smash/record.jsf?pid=diva2:490455 [Last accessed on 2025 Apr 16].

[5] J. G. Egbe, D. E. Ewa, S. E. Ubi, G. B. Ikwa and O. O. Tumenayo. "Application of multilinear regression analysis in modeling of soil properties for geotechnical civil engineering works in Calabar South". *Nigerian Journal of Technology*, vol. 36, no. 4, pp. 1059-1065, 2017.

[6] L. K. Sharma and T. N. Singh. "Regression-based models for the prediction of unconfined compressive strength of artificially structured soil". *Engineering with Computers*, vol. 34, no. 1, pp. 175-186, 2017.

[7] K. K. Phoon, W. A. Prakoso, Y. Wang and J. Ching. *Uncertainty Representation of Geotechnical Design Parameters*. In: "*Reliability of Geotechnical Structures in ISO2394*". CRC Press, pp. 67-106, 2016. Available from: https://www.taylorfrancis.com/chapters/edit/10.1201/978131536417911/uncertainty/representation/geotechnical/design/parameters/kok/kwang/phoon/widjojo/prakoso/yu/wang/jianye-ching [Last accessed on 2025 Apr 01].

[8] S. Pul, A. Ghaffari, E. Öztekin, M. Hüsem and S. Demir. "Experimental determination of cohesion and internal friction angle on conventional concretes". *ACI Materials Journal,* vol. 114, no. 3, pp. 407-416, 2017.

[9] M. M. Nawaz, S. R. Khan, R. Farooq, M. N. Nawaz, J. Khan, M. A. U. R. Tariq, R. F. Tufail, R. Farooq and A. W. M. Ng. "Development of a cost-based design model for spread footings in cohesive soils". *Sustainability*, vol. 14, no. 9, p. 5699, 2022.

[10] H. Do, S. Mirarab, L. Tahvildari and G. Rothermel. An empirical study of the effect of time constraints on the cost-benefits of regression testing. In: "*Proceedings of the 16th ACM SIGSOFT International Symposium on Foundations of Software Engineering*". ACM, Atlanta Georgia, pp. 71-82, 2008.

[11] P. Abbasimaedeh. "Soil liquefaction in seismic events: Pioneering predictive models using machine learning and advanced regression techniques". *Environmental Earth Sciences*, vol. 83, no. 7, p. 189, 2024.

[12] D. Basu, A. Misra and A. J. Puppala. "Sustainability and geotechnical engineering: Perspectives and review". *Canadian Geotechnical Journal*, vol. 52, no. 1, pp. 96-113, 2015.

[13] A. A. Firoozi and A. A. Firoozi. Soil-Structure Interaction: Understanding and Mitigating Challenges. In: "*Challenges in Foundation Engineering-Case Studies and Best Practices*". IntechOpen, 2023. Available from: https://www.intechopen.com/chapters/87761 [Last accessed on 2025 Apr 27].

[14] M. Mohammadi, S. M. F. Aghda, M. Talkhablou and A. Cheshomi. "Prediction of the shear strength parameters from easily-available soil properties by means of multivariate regression and artificial neural network methods". *Geomechanics and Geoengineering*, vol. 17, no. 2, pp. 442-454, 2020.

[15] M. N. Nawaz, B. Alshameri, Z. Maqsood and W. Hassan. "Predictive modelling of cohesion and friction angle of soil using gene expression programming: A step towards smart and sustainable construction". *Neural Computing and Applications*, vol. 36, no. 18, pp. 10545-10566, 2024.

[16] M. I. Khan and S. Wang. "Slope stability analysis to develop correlations between different soil parameters and factor of safety using regression analysis". *Polish Journal of Environmental Studies*, vol. 30, pp. 4021-4030, 2021.

[17] M. Erbilen. "*Imputing Missing Values using Support Variables with Application to Barley Grain Yield*". 2019. Available from: https://i/rep.emu.edu.tr/xmlui/handle/11129/5338 [Last accessed on 2025 Apr 01].

[18] K. A. Rashed. "Study of raw materials for manufacturing perforated clay bricks (Chamchamal, and Qaradagh regions in Sulaimani/Iraq)". *Sulaimani Journal for Engineering Sciences*, vol. 4, no. 4, pp. 50-66, 2017.

[19] M. R. Carter and E. G. Gregorich. "*Soil Sampling and Methods of Analysis*". CRC Press, 2007. Available from: https://www.taylorfrancis.com/books/mono/10.1201/9781420005271/soil/sampling/methods-analysis-gregorich-carter [Last accessed on 2025 Apr 14].

[20] D. Pennock, T. Yates and J. Braidek. Soil sampling designs. In: "*Soil Sampling and Methods of Analysis*". Canadian Society of Soil Science, Ontario, pp. 1-14, 2007.

[21] O. Fernández-Ugalde, A. Jones and R. G. Meuli. "Comparison of sampling with a spade and gouge auger for topsoil monitoring at the continental scale". *European Journal of Soil Science*, vol. 71, no. 2, pp. 137-150, 2020.

[22] T. Stark, S. Wilk and R. Swan. Sampling, Reconstituting, and Gradation Testing of Railroad Ballast In: "*Symposium on Railroad Ballast Testing and Properties*". ASTM International, pp. 135-143, 2018. Available from: https://asmedigitalcollection.asme.org/astm/ebooks/book/chapter/pdf/7398285/10_1520_stp160520170026.pdf [Last accessed on 2025 Apr 25].

[23] J. R. Boulding. "*Description and Sampling of Contaminated Soils: A Field Guide*". CRC Press, 2017. Available from: https://www.taylorfrancis.com/books/mono/10.1201/9781315140926/description/sampling-contaminated-soils-russell-boulding [Last accessed on 2025 Apr 27].

[24] A. R. Aquib, M. M. Uzzaman and M. F. Alam. "*Determining Engineering Properties of Fine-Grained Soils using Atterberg Limits*". Available from: https://www.researchgate.net/profile/ashkar/aquib/publication/372941330_determining_engineering_properties_of_fine/grained_soils_using_atterberg_limits/links/64d5f45bb684851d3d9e639e/determining/engineering/properties/of/fine/grained/soils/using/atterberg/limits.pdf [Last accessed on 2025 Apr 01].

[25] D. C. Montgomery and G. C. Runger. "*Applied Statistics and Probability for Engineers*". John Wiley and Sons, 2019. Available from: https://books.google.com/books?hl=en&lr=&id=nuo9eaaaqbaj&oi=fnd&pg=pa1&dq=montgomery,+d.+c.,+%26+runger,+g.+c.+(2010).+applied+statistics+and+probability+for+engineers.+wiley.&ots=jnbaw/rhip&sig=kwcr02cmpfzezjylw2pydnmdi2o [Last accessed on 2025 Apr 01].

[26] H. I. Okagbue, P. E. Oguntunde, E. C. Obasi and E. M. Akhmetshin. "Trends and usage pattern of SPSS and minitab software in scientific research". *Journal of Physics Conference Series*, vol. 1734, p. 012017, 2021.

[27] S. A. Lesik. "*Applied Statistical Inference with MINITAB®*". Chapman and Hall/CRC, 2018. Available from: https://www.taylorfrancis.com/books/mono/10.1201/9780429444951/applied/statistical/inference-minitab%c2%ae-second-edition-sally-lesik [Last accessed on 2025 Apr 16].

[28] R. S. Kenett and S. Zacks. "*Modern Industrial Statistics: With Applications in R, MINITAB, and JMP*". John Wiley and Sons,

2021. Available from: https://books.google.com/books?hl=en&lr=&id=iyeqeaaaqbaj&oi=fnd&pg=pr17&dq=minitab+y+line&ots=-cpvzlxfkx&sig=jqmw5a-6-rxp9ozynvftajokgia [Last accessed on 2025 Apr 27].

[29] W. C. Rinaman. Minitab. In: "*Revival: The Handbook of Software for Engineers and Scientists (1995)*". CRC Press, pp. 1028-1046, 2018. Available from: https://www.taylorfrancis.com/chapters/edit/10.1201/9780203710012/47/minitab/william/rinaman [Last accessed on 2025 Apr 16].

[30] E. Fragkiadaki, G. Evangelidis, N. Samaras and D. A. Dervos. "F-Value: Measuring an article's scientific impact". *Scientometrics*, vol. 86, no. 3, pp. 671-686, 2011.

[31] M. S. Thiese, B. Ronna and U. Ott. "P value interpretations and considerations". *Journal of Thoracic Disease*, vol. 8, no. 9, p. E928, 2016.

[32] J. Gao. "R-squared (R^2) - how much variation is explained". *Research Methods in Medicine and Health Sciences*, vol. 5, no. 4, pp. 104-109, 2024.

[33] J. Karch. "Improving on adjusted R-squared". *Collabra Psychology*, vol. 6, no. 1, p. 45, 2020.

[34] K. A. Hart, B. A. Steinfeldt and R. D. Braun. Formulation and applications of a probabilistic Pareto chart. In: *56th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. American Institute of Aeronautics and Astronautics, Kissimmee, Florida, 2015.

[35] P. Nieminen. "Application of standardized regression coefficient in meta-analysis". *BioMedInformatics*, vol. 2, no. 3, pp. 434-458, 2022.

[36] J. Hur and M. C. Jung. "The effects of soil properties on the turbidity of catchment soils from the Yongdam dam basin in Korea". *Environmental Geochemistry and Health*, vol. 31, no. 3, pp. 365-377, 2009.

[37] R. B. J. Brinkgreve. Selection of soil models and parameters for geotechnical engineering application. In: "*Soil Constitutive Models*". American Society of Civil Engineers, Austin, Texas, United States, pp. 69-98, 2005.