Indonesian Geotechnical Journal, April 2023, Vol 2 (1). Page 45 - 52
DOI: https://doi.org/10.56144/igj.v2i1.33
E-ISSN:2829-7520
https://indonesian-geotechnical-journal.org/index.php/JGI/issue/archive

Prediction of SPT Value From Cone Penetration Test at Intake Water Project of Wolo River, Southeast Sulawesi

Uniadi Mangidi[1], Anafi Minmahddun[2,*], Edward Ngii[3], Rini Sriyani[4], Sulha[5]
[1]Department of Civil Engineering, Faculty of Engineering, Halu Oleo University, Kendari, Indonesia, 93232; u.mangidi@uho.ac.id
[2]Department of Civil Engineering, Faculty of Engineering, Halu Oleo University, Kendari, Indonesia, 93232; anafi.minmahddun@uho.ac.id
[3]Department of Civil Engineering, Faculty of Engineering, Halu Oleo University, Kendari, Indonesia, 93232; edward.ngii@uho.ac.id
[4]Department of Civil Engineering, Faculty of Engineering, Halu Oleo University, Kendari, Indonesia, 93232; rini.sriyani@uho.ac.id
[5]Department of Civil Engineering, Faculty of Engineering, Halu Oleo University, Kendari, Indonesia, 93232; sulha@uho.ac.id
*Correspondence: anafi.minmahddun@uho.ac.id

SUBMITTED 31 December 2022 **REVISED** 14 March 2023 **ACCEPTED** 17 March 2023

ABSTRACT Soil investigation plays an important role in construction design. Adequate information regarding soil types and parameters are needed for a safe and economical foundation. However, limited budget often leads to inadequate soil investigations, and as a result, unknown soil parameters are estimated through correlations from the limited investigations that have been carried out. Correlations from SPT and CPT results are often used for design, but their applicability from location to location is questionable. This paper discusses the correlation of N-SPT values to CPT values in Wolo District, Kolaka Regency. The correlation was performed using linear regression analysis to obtain relationships between N-SPT and CPT. The three correlations obtained are N-SPT = 5.64 + 0.012q_c, R^2 value of 0.082; N-SPT = 5.93 + 0.01 q_c – 0.095q_f, R^2 value of 0.125; N-SPT = 0.025q_c + 0.341q_f, R^2 value of 0.455.

KEYWORDS Correlation, N-SPT, CPT, Linear Regression

1 INTRODUCTION

Geotechnical investigation is one of the activities that has to be carried out in every construction work for foundation design and other geotechnical structures. In practice, it is often found that not all required parameters can be tested independently due to limited budget. Therefore, geotechnical engineers often use correlations to obtain the necessary parameters. The general correlations that are commonly used are: qc = 4N (Terzaghi and Peck, 1967), correlation of N-SPT data on shear strength parameters and soil density (Bowles, 1984), and correlation of various field tests to index and engineering properties of soil (Look, 2014).

The varying soil conditions from one location to another have encouraged other researchers to conduct research on site-specific correlation from field tests to other soil parameters. Research related to SPT-CPT correlations have been conducted in Tanzania, Brazil, Florida, United Arab Emirates, Egypt and Ghana, and they all gave different results. (Akca, 2003; Shahien and Albatal, 2014; Jarushi, S. Alkaabim and Paul Cosentino, 2015; Lingwanda, Larsson, and Nyaoro, 2015; dos Santos and Bicalho, 2017; Opuni *et al.*, 2017). Similar research have also been conducted in Indonesia, giving different results from one region to another (Arisandi, Apriyanti, and Fahriani, 2017; Tanuwijaya, Kawanda, and Wijaya, 2019; Haifani, 2021).

From the literature, it is clear that site-specific correlation is necessary. This study aims to investigate the SPT-CPT correlation in Kolaka area, Southeast Sulawesi. Further studies on other part of Sulawesi Island will be conducted in the future.

2 RESEARCH METHOD

This research was conducted in Wolo River Intake project, Kolaka Regency, Southeast Sulawesi Province. The analysis was carried out on 2 SPT data and 2 CPT data, the locations can be seen in Figure 1. In this study, data from the same depth are compared. BH-1 was paired to CPT-1 data and data from BH2 was paired to CPT-2 data. All data were then analyzed by linear regression analysis. In this study, the analysis was carried out in two scenarios: 1) the dependent variable is N-SPT and the independent variable is the value of q_c, and 2) the dependent variable is N-SPT and the independent variables are the value of q_c and q_f.

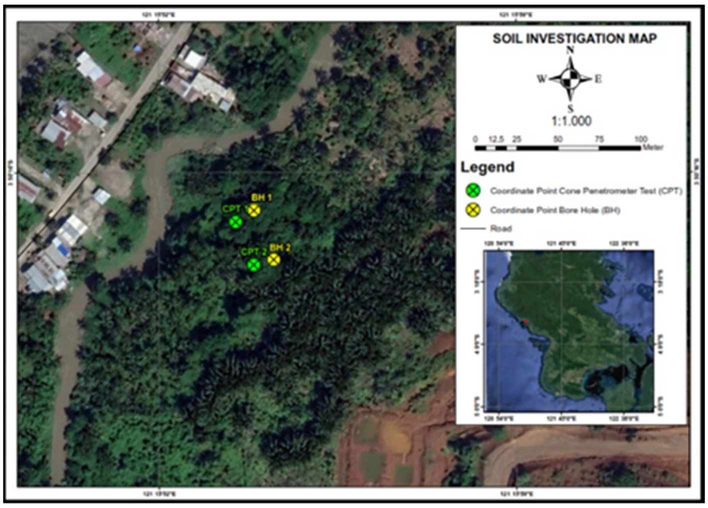

Figure 1. Soil investigation site

Regression analysis (linear and multiple) is used to determine the relationship of the independent variable on the dependent variable. In addition, linear regression is used to predict unknown data values using other related and known data values. Mathematically model the unknown or dependent variable and known or independent variable as a linear equation. Linear regression models are relatively simple and provide easy-to-interpret mathematical formulas to generate predictions (Ghozali, 2018). Multiple linear regression models which involve more than one independent variables can be seen in equation 1.

$$Y = b_0 + b_1 X_1 + b_2 X_2 + ... + b_n X_n \tag{1}$$

where,

$Y =$ dependent variable
$X_i =$ independent variable (I = 1, 2, …, n)
$b_0 =$ intercept
$b_i =$ regression coefficient (I = 1, 2, …, n)

to calculate b_0, b_1, b_2, …, b_k we use the Least Square Method which produces equation 2 as follows.

$$\begin{aligned} b_0 n + b_1 \sum X_1 + b_2 \sum X_2 + ... + b_n \sum X_n &= \sum Y \\ b_0 \sum X_1 + b_1 \sum X_1^2 + b_2 \sum X_1 X_2 + ... + b_n \sum X_1 X_n &= \sum X_1 Y \\ b_0 \sum X_2 + b_1 \sum X_2 X_1 + b_2 \sum X_2^2 + ... + b_n \sum X_2 X_n &= \sum X_2 Y \\ b_0 \sum X_n + b_1 \sum X_n X_1 + b_2 \sum X_n X_2 + ... + b_n \sum X_n^2 &= \sum X_n Y \end{aligned} \tag{2}$$

to simplify the calculation of b_0, b_1, b_2 the following matrices can be used:

$$\begin{bmatrix} n & \sum X_1 & \sum X_2 \\ \sum X_1 & \sum X_1^2 & \sum X_1X_2 \\ \sum X_2 & \sum X_2X_1 & \sum X_2^2 \end{bmatrix} \times \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \sum Y \\ \sum X_1Y \\ \sum X_2Y \end{bmatrix} \quad (3)$$

the value of the regression coefficient can then be determined by using the inverse matrix as follows:

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} n & \sum X_1 & \sum X_2 \\ \sum X_1 & \sum X_1^2 & \sum X_1X_2 \\ \sum X_2 & \sum X_2X_1 & \sum X_2^2 \end{bmatrix}^{-1} \times \begin{bmatrix} \sum Y \\ \sum X_1Y \\ \sum X_2Y \end{bmatrix} \quad (4)$$

The regression coefficient can be positive or negative. A positive value indicates that an increase in the independent variable also increases the dependent variable. On the other hand, a negative value indicates that an increase in the independent variable will decrease the value of the dependent variable. In addition to the regression coefficient, there are several indicators used to measure the relationship between two variables, namely the value of the correlation coefficient, the coefficient of determination, and the significance test (f-test and t-test).

The coefficient of correlation (R) is used to determine the relationship between the independent variables and the dependent variable. The value of R ranges from 0 to 1, the closer the value to 1, the stronger the relationship, conversely the closer the value to 0, the weaker the relationship. Guidelines to interpret how strong the relationship is from the correlation coefficient are given in Table 1.

Table 1. Interpretation of the correlation coefficient value (Sujarweini, 2014)

R	Interpretation
0.0 – 0.199	Very low
0.2 – 0.399	Low
0.4 – 0.599	Moderate
0.6 – 0.799	Strong
0.8 – 1.00	Very strong

The coefficient of determination (R^2) in linear regression is used to determine the percentage contribution of the independent variables on the dependent variable (Y). If R^2 is equal to 0, there is not even the slightest percentage of contribution from the independent variables have on the dependent variable. On the other hand, if R^2 is equal to 1, the independent variables have 100% contribution on the dependent variable.

For example, if the analysis shows a value of R^2 equals to 0.772, or 77.2%, then the contribution of the independent variables to the dependent variable is 77.2%. In other words the variation of the independent variables used in the model can explain 77.2% of the variation in the dependent variable, while the remaining 22.8% is influenced or explained by other variables not included in the model (Santoso, 2013).

To find out whether the independent variables have a simultaneous or partial effect on the dependent variable, the f-test and t-test are carried out. Before conducting this test, the level of confidence has to be predetermined to find out the maximum error value (alpha). In this study, a confidence level of 95% was used, so the alpha value was 5%.

The f-test is used to determine whether the independent variables jointly have a significant effect on the dependent variable, or to find out whether the regression model can be used to predict the dependent variable or not. Significant means that the relationship that occurs can be generalized to the population. For example, in a case where the population is 50 and 18 soil data samples are taken, if the conclusion is significant, the 18 samples can be generalized to the population of 50 samples. The f-test is carried out by looking at the sig.F value in the ANOVA table at the output of the regression analysis. The sig.F value is then compared to the alpha value, if the sig.F value is less than alpha it can be concluded that there is an significant(?) influence between all the independent variables and the dependent variable, and vice versa.

The t-test is used to examine how significant is the effect of each independent variables on the dependent variable. It is extremely likely that one or more independent variables have no impact on the dependent variable. The t-test uses P-value obtained from the regression analysis. If an independent variable has a P-value lower than its alpha value (5%), the independent variable does not significantly impact the dependent variable. On the other hand, f its P-value is larger than its alpha value, the independent variable has significant impact on the dependent variable (Santoso, 2013).

The analysis is carried out in two scenarios, namely:

1. The dependent variable is N-SPT, and the independent variable is the value of q_c.
2. The dependent variable is N-SPT, and the independent variables are the value of q_c and q_f.

3 ANALYSIS RESULT

3.1 Data Collection

The data used in the analysis are based on soil investigations at BH-1, BH-2, CPT-1, and CPT-2 as shown in Figure 1. The results of soil investigations at these points are shown in Figure 2 and Figure 3. The investigations data at the same depth are then averaged, grouped and shown in Table 2.

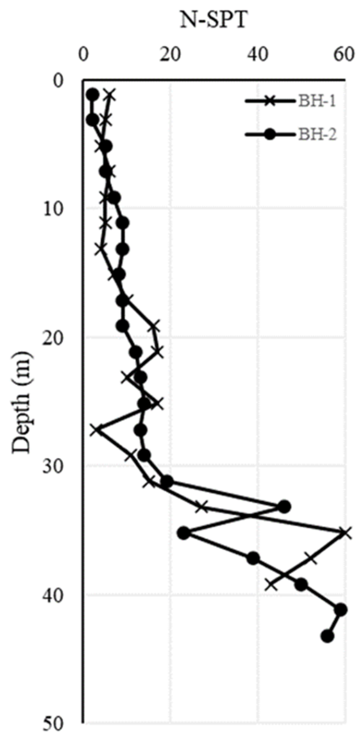

Figure 2. Results of N-SPT

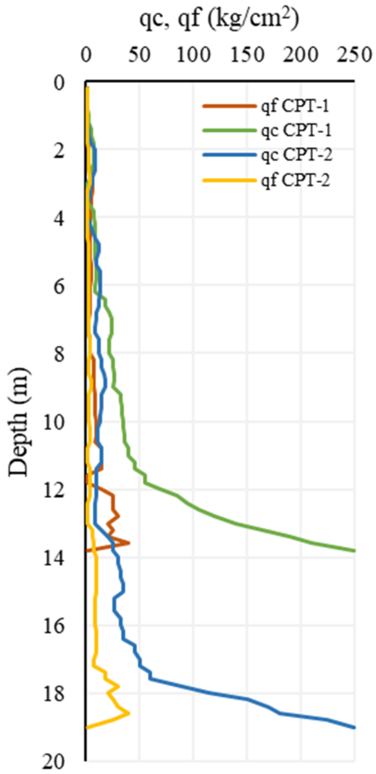

Figure 3. Results of CPT

Table 2. Grouping of N-SPT and CPT data based on the same depth

Depth	N-SPT	q_c (kg/cm^2)	q_f (kg/cm^2)
1.15	5	3	1
3.15	5	2	6
5.15	4	9	4
7.15	5	24	3
9.15	5	32	8
11.15	5	45	15
13.15	4	165	25
1.15	2	2	1
3.15	2	4	2
5.15	5	10	3
7.15	5	9	3
9.15	8	15	3
11.15	10	14	2
13.15	10	15	6
15.15	9	27	8
17.15	10	50	7
19.15	10	250	2

3.2 Analysis of Linear Regression

The results of linear regression analysis of the relationship between N-SPT and q_c with a 95% confidence level are listed in Tables 3, 4, and 5. The analysis obtained the equation N-SPT = 5.64 + 0.012q_c. The coefficient of correlation for the equation is 0.281, meaning relationship between variables is in the low category. The value of the coefficient of determination is 0.082 or 8.2%, which indicates that the independent variable only contributes to 8.2% of the dependent variable. The significance test from the t-test and f-test obtained a significance value of 0.26, which is greater than 5%. Since the significance value is greater than the alpha value, this indicates that there is no significant effect of q_c value on the N-SPT value.

Table 3. Linear regression equation of N-SPT and q_c

	Coeff.	Stand. Error	t Stat	P-value
Intercept	5.638	0.78	7.19	$3x10^{-6}$
q_c	0.012	0.01	1.16	0.26

Table 4. Coefficient of correlation and coefficient of determination

Regression Statistics	
Multiple R	0.281
R Square	0.082
Adjusted R Square	0.023

Table 5. ANOVA table

	df	SS	MS	F	sig.F
Regression	1	10.24	10.24	1.35	0.26
Residual	15	113.53	7.57		
Total	16	123.77			

The next analysis is the relationship between N-SPT on the values of q_c and q_f. The results of the analysis are listed in Tables 6, 7, and 8.

Table 6. Linear regression equation of N-SPT and q_c, q_f

	Coeff.	Stand..Error	t Stat	P-value
Intercept	5.925	0.80	7.39	$3x10^{-6}$
q_c	0.013	0.01	1.36	0.20
q_f	-0.095	0.10	-0.92	0.38

Table 7. Coefficient of correlation and coefficient of determination

Regression Statistics	
Multiple R	0.354
R Square	0.125
Adjusted R Square	0.001

Table 8. ANOVA table

	df	SS	MS	F	sig.F
Regression	2	10.76	5.38	1.00	0.39
Residual	14	75.00	5.36		
Total	16	85.76			

The regression equation from this second model is N-SPT = 5.93 + 0.01 q_c – 0.095q_f. This equation shows that the value of q_f is inversely proportional to the value of N-SPT. The coefficient of correlation and determination in the second model is slightly better than the previous model, namely 0.354 and 0.125 respectively, which means that 12.5% of the independent variables (q_c and q_f) affect the dependent variable (N-SPT). The sig.F value for this analysis is 0.39, which indicates that q_c and q_f also have no significant effect on the N-SPT.

The second model has slightly better correlation than the first model. However, the results of the significance test still show that there are no significant effect between the variables. Therefore, the third model was developed by setting the intercept value in the regression equation of the second model to zero. The results of the analysis of the third model are listed in Tables 9, 10, and 11.

Table 9. Linear regression equation of N-SPT and q_c, q_f

	Coeff.	Stand.Error	t Stat	P-value
Intercept	0.000			
q_c	0.025	0.02	1.25	0.23
q_f	0.341	0.18	1.87	0.08

Table 10. Coefficient of correlation and coefficient of determination

Regression Statistics	
Multiple R	0.674
R Square	0.455
Adjusted R Square	0.352

Table 11. ANOVA table

	df	SS	MS	F	sig.F
Regression	2	306.52	153.26	6.25	0.011
Residual	15	367.47	24.49		
Total	17	674			

the equation obtained from the third model is N-SPT = $0.025q_c + 0.341q_f$. This equation shows that both the q_c and q_f values are directly proportional to the N-SPT values. The coefficient of correlation and the coefficient of determination are 0.674 and 0.455 respectively, which means that the correlation between the independent and dependent variables is in the strong category and 45.5% of the independent variables contributes to the dependent variable. This correlation is much better than the previous two models.

The significance test of independent variables on the dependent variable also gave better results than the previous two models. The results of the f-test show that the independent variables (q_c and q_f) affect the dependent variable (N-SPT) which can be seen from the value of sig.F < 5%. The results of the partial significance test in Table 9 show that the variable q_f affects the N-SPT variable, but in this study the q_f variable does not significantly affect the N-SPT variable. This could be due to very limited data available. However, combining q_c and q_f gives better results in predicting the N-SPT value.

4 CONCLUSION

Based on the results of the regression analysis, the conclusions in this study are as follows:

1. The regression equation for the relationship between the N-SPT value and the CPT value is as follows.
   - N-SPT = $5.64 + 0.012q_c$, with coefficient of determination 0.082
   - N-SPT = $5.93 + 0.01\ q_c - 0.095q_f$, with coefficient of determination 0.125
   - N-SPT = $0.025q_c + 0.341q_f$, with coefficient of determination 0.455
2. The results of the significance test of the three models are as follows:
   - Model 1 has no significance both simultaneously and partially.
   - Model 2 has no significance both simultaneously and partially.
   - Model 3 has a significant simultaneous effect of the independent variables on the dependent variable, and the q_f variable is partially significant on the N-SPT variable.
3. Entering the values of q_c and q_f together give better results in predicting the value of N-SPT.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to Institution for Research and Community Service of Halu Oleo University for the financial assistance in this research. The authors also highly appreciate the support of the Laboratory of Soil Mechanics, Faculty of Engineering, Halu Oleo University, which has assisted in the data collection process.

REFERENCES

Akca, N., 2003. Correlation of SPT-CPT data from the United Arab Emirates. *Engineering Geology*, 67(3–4), pp. 219–231. https://doi.org/10.1016/S0013-7952(02)00181-3

Arisandi, D., Apriyanti, Y. and Fahriani, F., 2017. Analisis Korelasi Nilai N-SPT dengan Nilai Tahanan Konus Sondir dengan Menggunakan Metode Statistika. *Jurnal Fropil*, 5(1), pp. 1–10.

Bowles, J. E., 1984. *Physical and Geotechnical Properties of Soils*. 2nd edn. Auckland: McGraw-Hill.

Ghozali, I., 2018. *Aplikasi Analisis Multivariate dengan Program IBM SPSS 25*. Semarang: Badan Penerbit Universitas Diponegoro.

Haifani, A. M., 2021. Developing New Relationship between Standard Penetration, Cone Penetration Test and Sleeve Friction in Various Soil Type. *Journal of Hunan University Natural Sciences*, 48(4).

Jarushi, F., Alkaabim, S. and Cosentino, P., 2015. A New Correlation between SPT and CPT for Various Soils. *International Journal of Environmental, Ecological, Geological and Geophysical Engineering* , 9(2), pp. 101–107.

Lingwanda, M. I., Larsson, S. and Nyaoro, D. L., 2015. Correlations of SPT, CPT and DPL Data for Sandy Soil in Tanzania. *Geotechnical and Geological Engineering*, 33(5), pp. 1221–1233. https://doi.org/10.1007/s10706-015-9897-1

Look, B. G., 2014. *Handbook of Geotechnical Investigation and Design Tables*. 2nd edn. London: Taylor and Francis.

Opuni, K. O. *et al.*, 2017. Correlations of SPT and DCPT data for sandy soils in Ghana, *Lowland Technology International*, 19(2), pp. 145–150.

dos Santos, M. D. and Bicalho, K. V., 2017. Proposals of SPT-CPT and DPL-CPT correlations for sandy soils in Brazil. *Journal of Rock Mechanics and Geotechnical Engineering*, 9(6), pp. 1152–1158. https://doi.org/10.1016/j.jrmge.2017.08.001

Santoso, S., 2013. *Menguasai SPSS 21 di Era Informasi*. Jakarta: PT. Elek Media Komputindo.

Shahien, M. M. & Albatal, A. H., 2014. SPT-CPT Correlations for Nile Delta Silty Sand Deposits in Egypt. *Proceeding of 3rd International Symposium on Cone Penetration Testing*, pp. 2–43.

Sujarweini, V. W., 2014. *SPSS untuk penelitian*. Yogyakarta: Pustaka Baru Press.

Tanuwijaya, E., Kawanda, A. and Wijaya, H., 2019. Studi Korelasi Nilai Tahanan Konus Sondir Terhadap Parameter Tanah Pada Proyek Di Jakarta Barat. *JMTS: Jurnal Mitra Teknik Sipil*, 2(4), pp. 169–176. https://doi.org/10.24912/jmts.v2i4.6188

Terzaghi, K. and Peck, R. B., 1967. *Soil Mechanics in Engineering Practice*. 2nd edn. USA: John Wiley and Sons.