Journal of Sustainable Civil Engineering and Technology
e-ISSN: 2948-4294 | Volume 4 Issue 1 (March 2025), 1-6
https://joscetech.uitm.edu.my
https://doi.org/10.24191/jscet.v4i1.G_000010

Estimation on SPT Number based on the Prediction N-Value Model for Soft Soils

Raudhah Razali[1], Norazlan Khalid[2*], Mazidah Mukri[2], Ismacahyadibagus Mohamed Jais[2] and Rini Kusumawardani[3]

[1] Graduate Student, School of Civil Engineering, College of Engineering, Universiti Teknologi MARA, 40450, Selangor, Malaysia

[2] School of Civil Engineering, College of Engineering, Universiti Teknologi MARA, 40450, Selangor, Malaysia

[3] Civil Engineering Department, Faculty of Engineering, Universitas Negeri Semarang, Semarang, Indonesia

*Corresponding author: norazlan0481@uitm.edu.my

Received: 11 November 2024 / Accepted: 27 February 2025 / Published online: 31 March 2025

Abstract

This paper proposes a precise multiple linear regression model for N-Value to estimate the Standard Penetration Test (SPT) number for soft soil. The study used 76 data from soil properties for N-Value model development collected from site investigation report. Meanwhile, 47 datasets for N-Value model validation collected from other site investigation report and technical paper. The N-value model developed in this study has met the requirements and has been verified according to the statistical validity requirements. The results shows that the developed N-Value model has the significant accuracy linear regression with R2 = 68%. For the model accuracy it shows the small deviation discrepancy results for Mean Square Error (MSE) is 0.4, Mean Absolute Error (MAE) is 0.4, and Mean Absolute Percentage Error (MAPE) is 0.31 with a determination coefficient, R2 = 93.7%. Accordingly, the N-Value models developed in this study demonstrate a good prediction capability for SPT number. The proposed model will serve as a reference for civil engineers in estimating SPT number for provide information on the geotechnical engineering properties of soil, especially for the determination of bearing capacity and settlement.

Keywords: Standard penetration test, N-Value, Soft Soil

1. Introduction

Standard Penetration Test (SPT) presently the common in situ dynamic testing method is conducted inside a borehole by measuring the blow counts (N) and provide an information for the geotechnical engineering properties for subsurface soil (Ji et al., 2023) and (Ghali et al., 2020). The number of blows required to effect 300mm penetration (test drive) below an initial penetration of 150mm (seating drive), was recorded as the penetration resistance or N-value. The relative density of cohesionless soils and the consistency of cohesive soils have been defined based on the penetration resistance, N-value. Meanwhile the outcomes result for N-value can be utilized to assess the relative density, strength of soil, bearing capacity, settlement of granular soil, and to establish an approximate correlation with the strength of cohesive soil (Hong et al., 2022 & Marto et al., 2018). Therefore, the application of SPT in various domains of soil mechanics and foundations, and it's become an essential parameter within the field of geotechnical engineering (Hatta and Syed, 2015). Originally designed for coarse-grained soils, however correlations for the of N-values with various physical and mechanical soil properties have been developed by many researchers previously. These correlations have broadened its applicability to all soil types and facilitated the determination of several soil parameters (Arshid, 2021). In geotechnical engineering, the determination of the number of blows (N) in the SPT plays an important role, particularly when assessing soil conditions. The importance of in-situ testing increases due to the demanding nature of laboratory tests in terms of time and expenses. In fact, immediate in-situ data is often required, and delays in acquiring precise N values can result in setbacks in decision-making and project planning (Qi et al., 2006; Rosli et al., 2020). Therefore, this paper aims to develop an N-Value prediction model for soft soil with a strong correlation with soil properties that have been determined in this study such as plastic limit and bulk density. The SPT numbers (N-Value) for soft soil can be estimated using the developed N-Value model without conducting the SPT testing on site and useful for especially for bearing capacity for foundation.

2. Method

A total of 76 primary data sets for soil laboratory properties test results were collected from actual site investigation (SI) reports from several locations with soft soil distribution and were used for model development as shown in Table 1. The main dataset was collected from site investigation report such as N-value, liquid limit (LL), plastic limit (PL), plasticity index (PI), Sand, Silt, Clay, Bulk Density (BD) and Saturated Density (SD). Then the descriptive statistics data were analysed and used to illustrate the characteristics of collected dataset in a form of easy understanding as well as quantitative evaluation criteria. A detailed summary including mean, standard deviation, maximum, minimum, mode, skewness, and kurtosis.

Table 1. Main dataset for soil properties and N-Value from site investigation work report for model development

No	Site Investigation (SI) Report	No. Data
1	SI Work at No. 5 Kg. Sungai Serdang, Klang Selangor	15
2	SI Work at Ambang Botanic, Bandar Botanic, Klang.	18
3	SI Work at Universiti Sains Malaysia, Nibong Tebal, Pulau Pinang.	15
4	SI Report Bagan Datuk, Perak	12
5	SI Report for Cadangan membina Loji IWK at Mukim Kapar, Klang	16

The multiple linear regression (MLR) analysis was carried out that uses several explanatory variables to predict the outcome of a response variable and using statistical validation. The appropriate model was accomplished with coefficient of regression, R2 to develop a N-Value model. The confidence level of 95% or 0.05 was chosen and produced the significant level of P-Values = 0.05. The confidence level was based on the hypothesis testing for each testing. The main possibility independent soil properties parameters were evaluated using the linear correlation analysis. This hypothesis relies on the Pearson correlation and the significance of the P-value between the soil properties variables and N-value from SI report. The nearer Pearson correlation coefficients value to 1 shows the stronger correlation between variables and a zero correlation indicates there is no relationship between the variables. Meanwhile, the P-value < 0.05 means the correlation is significant (Mendenhall et al., 2006). The final predicted N-Value model efficiency criterion was verified using the comparison with measured data from the new 47 dataset collected from site investigation report and technical paper as shown in Table 2.

Table 2. Dataset for soil properties and N-Value from site investigation work report for validation

No	Site Investigation (SI) Report	No. Data
1	SI Report for Construction of Jakarta Mass Rapid Transit Project, Jakarta Indonesia.	20
2	SI Report for East Coast Road, Singapore	12
3	The compression behaviour of marine clay in Malaysia	15

Journal of Sustainable Civil Engineering and Technology
e-ISSN: 2948-4294 | Volume 4 Issue 1 (March 2025), 1-6
https://joscetech.uitm.edu.my
https://doi.org/10.24191/jscet.v4i1.G_000010

The N-Value model was verified to ensure the accuracy of the developed model. Therefore, the mean squared error (MSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) were used to quantify the performance of the developed model. It considered closer to the measured results if all entire produce small discrepancies values (Montgomery et al., 2012; Montgomery 2017). The relationship between the N-Value predicted models versus the N-Value measured data was carried out for model validation and the plot will show the scatter point closely between the predicted and measured value (Diah et al., 2012). Meanwhile, the hypothesis testing was carried out to find out whether the observations studied on the sample are within the confidence level. The T-Test was performed for the hypothesis testing and the rejection or acceptance of a tested hypothesis was determined at a confidence level of 95% or 0.95. This confidence level produced the significant level more than 5% or 0.05 which is given by the symbol of $P > 0.05$ (Khalid et al., 2024). Omar et al., (2018) stated that the model is considered reliable when it's fit to the three requirement such as coefficient of determination (R2), mean square error (MSE) and T-Test.This section details the procedures and techniques used to collect and analyze data. Subsections might include and not limited to:

3. Results and Discussion

The respectable and most significant N-Value model was developed on predicting the SPT value for soft soil expressed as shown in Equation 1. Two significant independent soil properties variables have been used are Plastic Limit (PL) and Bulk Density (BD).

$$\text{N-Value} = 39.73 - 1.39\,(PL/BD) \tag{1}$$

The relatively good regression model obtained in these two models was discussed detail in this paper. Table 3 shows the good correlation results between the soil properties for N-Value from SI report. It shows the most significant with strong correlation soil properties for N-Value model were PL and BD with the Pearson correlation result near to 1 and -1 which is -0.52 and 0.815 respectively with the P-Value < 0.05 for both variables.

Table 3. Correlation result

Model	Variable	Pearson Correlation	P-Value
N-Value	PL	-0.705	0.000
	BD	0.815	0.000

Prior to develop the N-value model for soft soil, the descriptive statistics of 76 data obtained to describe the characteristics of a sample in an easily understandable form as shown in Table 4. It shows the data gathered were normally scattered in a normal distribution with left-skewed and slightly leptokurtic due to the range of skewness and kurtosis was in the range of -2 and 2. George and Mallery (2010) stated that any score for skewness and kurtosis is in the range -2 and 2, it can be accepted as a normal requirement.

Table 4. Descriptive data results for SPT-Model

Variable	Mean	SE Mean	St.Dev	Median	Min	Max	Skewness	Kurtosis
N-Value	7.15	0.76	4.86	7.00	0.00	18.00	0.41	-0.80
PL	36.01	1.22	7.22	37.00	15.00	50.00	-0.71	1.03
BD	1.53	0.02	0.14	1.53	1.24	1.99	0.76	1.23

Table 5 shows the residual normality testing result for N-Value model based on the result of Anderson-Darling, Ryan-Joiner and Kolmogorov Smirnov. It provides valuable insights into the residual distributional characteristics of the dataset. The results show the entire of P-Value for normality test for N-Value model more than significance value (P-Value > 0.05). This indicates N-Value model considered in normal distribution curve. Therefore, there is no reason to deny the validity of the assumptions of these regressions.

Journal of Sustainable Civil Engineering and Technology
e-ISSN: 2948-4294 | Volume 4 Issue 1 (March 2025), 1-6
https://joscetech.uitm.edu.my
https://doi.org/10.24191/jscet.v4i1.G_000010

Table 5. Residual normality result for N-Value Model

Model	P-Value		
	Anderson Darling	Ryan-Joiner	Kolmogorov Smirnov
N-Value	0.208	> 0.10	> 0.150

Table 6 shows the regression analysis result for N-Value model through the MLR analysis. It shows the result of P-Value < 0.05 was at the significant level and there is enough evidence to conclude that is the variables can be used to predict the N-Value. To obtain a good N-Value model with a result of P-Value < 0.05 and VIF < 5, the variable transformation must be performed by combining two correlated variables. In this study, the PL parameter has been combined with BD to (PL/BD) as one of the main parameters in the N-Value model. It shows the combined transformation of these parameters was influenced to the N-Value value and makes the relationship better and reflected with a high of coefficient of determination, R^2 (adj) value at 68% as shown in Table 4. It considered as a practical for making prediction of N-Value since it has given the R^2 and R^2(adj) more than 50%. The (R^2) describes the magnitude of the response variation that can be explained by the predictor The chosen for value of R^2(adj) because it is more sensitive to the several addition parameters in the model [Khalid et al., 2023]. Meanwhile, the positive sign to the variable indicates that it has a direct relationship with the N-Value model. For example, the higher value from one of the variables, the greater of N-Value.

Table 6. Regression analysis result for N-Value model

Model	Variables / Predictor	Coefficient	Std. Error Coefficient	P-Value	VIF	R^2 (%)	$R^2_{(adj)}$ (%)
N-Value	Constant	39.73	6.00	0.00	-	69.7	68.2
	PL/BD	-1.39	0.023	0.00	1.00		

The analysis of variance (ANOVA) was done to provide information about levels of variability within a regression model and form a basis for tests of significance. Table 7 shows the ANOVA result for N-Value model was at the significant level of P-Value < 0.05 and it can be presumed that the model has been created can be used for predicting the N-Value.

Table 7. Analysis of variance result for N-Value model

Model	Source	DF	P-Value
N-Value	Regression	1	0.00
	Residual Error	33	-
	Total	24	-

The evaluation and validation predicted N-Value model was carried out between measured N-Value and predicted N-Value was made using the T-Test as shown in Table 8. The result shows that the P-value is 0.011 for N-Value model which is more than $\alpha = 0.05$ (P-Value > 0.05). Since the predicted N-Value was not much different from the measured N-Value, the model has been determined and can be used to predict the SPT value (N-Value). The result shows small deviation for developed N-Value model based on the MSE, MAE, and MAPE from the comparison between N-Value predicted and N-Value measured.

Table 8. Summarize validation results for SPT-Model

Model	No. Data	T-Test (P-Value)	MSE	MAE	MAPE	R^2
N-Value	47	0.011	0.4	0.4	0.31	0.937

Figure 1 presents the connection graph between the predicted and measured for N-Value. It shows the reasonably good regression coefficients for R^2 value was 0.937 respectively with the small discrepancy MSE value at 0.00065. It can be seen the points were scattered closely around the 45-degree line. It can be determined that the predicted value matched the empirical value exactly; hence, the N-Value model can be accepted.

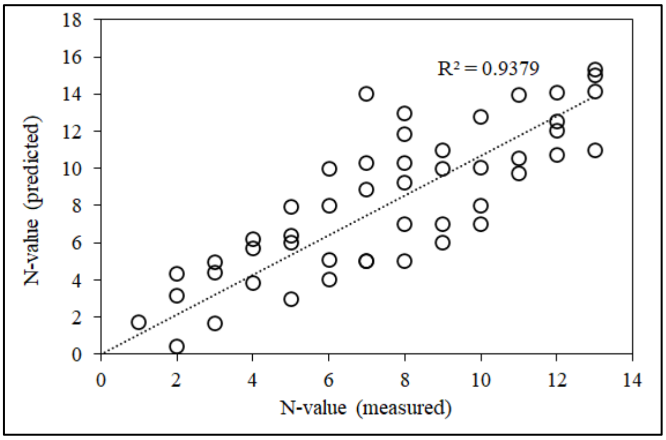

Figure 1. Fitted graph of N-Value predicted versus measured

4. Conclusion

The established N-Value model was developed from two main parameters, namely, plastic limit (PL) and bulk density (BD), based on the MLR analysis and may be useful in forecasting the N-Value or SPT number in the absence of fieldwork. It has been proven by statistical conditions; the N-value shows satisfactory results with a good regression coefficient value, $R^2 = 68\%$. Meanwhile, the effectiveness of the developed N-value model compared between measured and predicted was verified with the small deviation divergence of MSE, MPE, and MSE value with a higher coefficient of determination, $R^2 = 93.7\%$. Therefore, it is recommended that the proposed N-value model will be practically used for predicting the N-value or SPT number for a geotechnical design where there are financial and time limitations. Suggestions for improving the model evaluation by using various more comprehensive model methods by considering several other factors such as energy used. Summarize the key findings, their implications, and any limitations of the study. Also, suggest areas for future research.

Acknowledgments

The authors would like to extend my gratitude to the School of Civil Engineering, Universiti Teknologi MARA for providing necessary facilities for this research.

Declaration of Conflicting Interests

All authors declare that they have no conflicts of interest.

References

Arshid, M.U. (2021). Knowledge Based Prediction of Standard Penetration Resistance of Soil Using Geotechnical Database. *Civil Engineering Journal*, Vol. 7 (Special Issue), 1–12. https://doi.org/10.28991/CEJ-SP2021-07-01

Diah, J. M., Ahmad, J. & Mukri, M. (2012). The Methodology on Statistical Analysis of Data Transformation for Model Development. *International Journal of Statistics and Applications.* 2(6): 94-100. Doi: https://doi.org/10.5923/j.statistics.20120206.01

George, D., & Mallery, M. (2010). *SPSS for Windows Step by Steps: A Simple Guide and Reference*. Boston: Pearson.

Ghali, M., Chekired, M., & Karray, M. (2020). Framework to improve the correlation of SPT-N and geotechnical parameters in sand. *Acta Geotechnica*, 15(3), 735–759. https://doi.org/10.1007/s11440-018-0745-3

Hatta, K. A., & Syed Osman, S. B. A. (2015). Correlation of Electrical Resistivity and SPT-N Value from Standard Penetration Test (SPT) of Sandy Soil. *Applied Mechanics and Materials*, Vol. 785, 702–706. https://doi.org/10.4028/www.scientific.net/amm.785.702

Hong, W. T., Kim, S. Y., & Lee, J. S. (2022). Evaluation of driving energy transferred to split spoon sampler for accuracy improvement of standard penetration test. Measurement: *Journal of the International Measurement Confederation*, 188. https://doi.org/10.1016/j.measurement.2021.110384

Ji, P., Yang, J., Zhang, N., Zhai, Q., Li, M., Wang, Y., Shen, T., & Satyanaga, A. (2023). Energy Measurement in Standard Penetration Tests. *Sustainability*, 15(6), 4763. https://doi.org/10.3390/su15064763

Khalid, N., Mukri, M., Zain, N.H.M., Zainuddin, A.N. & Kamarudin, F. (2024). Prediction model of hydraulic conductivity for sedimentary residual soil mixed bentonite as compacted clay liner. *IOP Conference Series: Earth and Environmental Science*, 1347 (1), art. no. 012048.

Khalid, N., Mukri, M., Zain, N.M., Razak, Z. & Jais, I.M. (2023). Multiple Linear Regression Modelling for The Compaction Characteristics of Sedimentary Soil Mixed Bentonite As Compacted Liner. *Jurnal Teknologi*, 85 (6), pp. 181-186.

Marto, A., Sakai, G., Suemasa, N., & Jamaludin, N. (2018). An Attempt of Screw Driving Sounding Test in Malaysia. *MATEC Web of Conferences*, 203, 04008. doi:10.1051/matecconf/20182030400

Mendenhall, W., Beaver, R.J. & Beaver B.M. (2006). *"Probability and Statistics."* 12th Edition. Thomson Brooks-Cole, The Thomson Corporation. U.S.

Montgomery, D.C. (2017). Design and analysis of experiments. John Wiley & Sons.

Montgomery, D.C., Peck, E.A. & Vining, G.G. (2012). *Introduction to Linear Regression Analysis* (5th ed.). New Jersey: John Wiley & Sons, Inc.

Omar, M., Shanableh, A., Arab, M., Hamad, K. & Tahmaz, A. (2018). Advanced Mathematical Models to Predict the Compaction Properties of Coarse-grained Soils from Various Physical Properties. *Geotech Geol Eng.* Doi: https://doi.org/10.1007/s10706-018-0547-2.

Qi, X., Bard, J.F., & Yu, G. (2006). Disruption management for machine scheduling: The case of SPT schedules. *International Journal of Production Economics*, 103(1), 166–184. https://doi.org/10.1016/j.ijpe.2005.05.02

Rosli, N., Saad, R., Rahman, N., & Ismail, N. A. (2020). Soft soils: A study on their electrical resistivity values and geotechnical properties (porosity, SPT and particle size distribution). *Warta Geologi*, 46(3), 186–190. https://doi.org/10.7186/wg463202002