2024 Scienceline Publication

Journal of Civil Engineering and Urbanism

Volume 14, Issue 3S: 311-317; September 15, 2024 ISSN-2252-0430

DOI: https://dx.doi.org/10.54203/jceu.2024.34

The Use of Machine Learning Approach to Predict Pile Capacity in Non-Cohesive Soils

Lesego Palalane and Mahongo Dithinde

Department of Civil Engineering, University of Botswana, Gaborone, Botswana

Corresponding author's Email: dithinde@ub.ac.bw

RESEARCH ARTICLE
PII: S225204302400034-14
Received: June 25, 2024
Revised: September 02, 2024
Accepted: September 05, 2024

ABSTRACT

Existing theoretical and empirical pile design methods cannot accurately model the complex interaction between piles and soil. Consequently, there is a growing trend towards utilizing machine learning techniques to better capture the nonlinear soil-pile interaction. This paper aims to predict the capacity of bored piles in cohesionless soils using a machine learning approach. The machine learning algorithm was trained using a database of 18 bored pile cases in non-cohesive soils and validated with a separate dataset of 8 bored piles in cohesionless soil. Moreover, the performance of the machine learning method was compared with that of a traditional pile design method (i.e., SA-SPT method) in Southern Africa. The evaluation was based on the ratio of measured capacity to predicted capacity (Q_m/Q_p) statistics and the coefficient of determination (R^2). The results showed an R^2 of 0.89 for the machine learning method compared to 0.85 for the SA-SPT method, indicating the superior accuracy of the machine learning approach in predicting pile capacity.

Keywords: Machine learning, SPT-based pile methods, Load Bearing Capacity, Full-scale Load Test, Chin extrapolation method, Terzhagi's 10% criteria.

INTRODUCTION

Pile foundations are vital for supporting critical and complex structures due to their ability to accommodate varying soil conditions and withstand both vertical and lateral loads. The reliability of structures supported by piles hinges upon the performance and behavior of these foundational elements, as emphasized by Zhang et al. (2020).

Despite their importance, the interaction between soil and piles remains a highly complex and not fully understood phenomenon. Various theoretical and empirical methods, such as the Franki Africa method described by Johnson et al. (2001), have been developed to model soil-pile behavior. However, current procedures for pile design typically rely on semi-empirical approaches based on elastic theories and pile load test data.

Assessing pile capacity in non-cohesive soils, such as sand and gravel, poses significant challenges due to their granular composition and unpredictable behavior. Traditional pile design methods (e.g. Meyerhof, 1976, Decourt 1995, etc.) oversimplify the complex soil-pile interaction and generally neglect factors like soil gradation, particle shape, and compaction. Generally the soil properties are determined using the standard penetration test (SPT). However, SPT has a number of limitations as stated by Seed et al. (1985) and Skempton (1986). The main limitation is the measured SPT-N values are not well related to the pile loading process. Pile load tests are commonly used to verify nominal resistances, but they can be costly and time-consuming. Moreover, the determination of pile capacity based on load-settlement curves lacks a single standard methodology, leading to a wide range of results and making pile design somewhat of a subjective exercise, as noted by Horvitz et al. (1981) and Shariatmadari et al. (2008).

Given these challenges, there has been a recent trend toward leveraging machine learning techniques, as noted by Yago et al. (2021), to better model the intricate and nonlinear connections between piles and surrounding soil, reflecting their increasing adoption across various engineering applications. Shoaib and Abu-Farsakh (2023) emphasized the importance of carefully selecting input variables that influence the output, particularly in predicting ultimate pile capacity (Q_p).

MATERIALS AND METHODS

Pile load test data

The main input data in this study comprised of a static pile load tests database of 26 cases from Southern Africa (South Africa, Botswana, Lesotho, Mozambique, Zambia and Swaziland and Tanzania) along with the associated geotechnical data (soil profiles, field and laboratory test

To cite this paper: Palalane L and Dithinde M (2024). The Use of Machine Learning Approach to Predict Pile Capacity in Non-Cohesive Soils. *J. Civil Eng. Urban.*, 14 (3S): 311-317. DOI: https://dx.doi.org/10.54203/jceu.2024.34

results). The details of the test piles and associated geotechnical data are presented in Appendix 1.

Determination of measured Pile Capacity

The static pile load test field records were further processed by plotting the load versus the head deflection to produce load-deflection curves. The load-deflection curves were then used to estimate the ultimate pile capacity or measured capacity (Q_m). However, majority of the test piles are working piles tested to a maximum load varying from one and half to two times the design load which limits the movement to which the pile head is subjected and requires extrapolation procedure to determine the ultimate capacity e.g. (e.g. Chin, 1970; Fleming, 1992; Decourt, 1999). On account of its popularity, Chin extrapolation method was adopted for this study. The values of measured capacities were compared with the predicted capacities obtained from Machine learning model and direct Southern African SPT method.

Development a machine learning model

The study used a Multiple Linear Regression (MLR) model developed using the SPSS Modeler to determine the predicted capacities. The SPSS Modeler randomly subdivided the 26 pile cases into (i) 18 dataset for model training and (ii) 8 dataset for model validation. The dataset for model training was structured into two main categories: independent variables and a dependent variable.

Independent variables encompassed factors believed to influence the pile's capacity, such as the length of the pile, area of the base and shaft as shown in Appendix 2. The dependent variable is the measured pile capacity obtained from pile load tests as previously described

The stream flow in Figure 1 shows how the dataset was partitioned into model training (70%) and model validation (30%) sets. After training, the MLR model was tested using the reserved testing set to assess its effectiveness in making accurate predictions on unseen data. The trained model was then applied to predict pile capacity by utilizing the adjusted coefficients in a linear equation, and the predicted pile capacity was analyzed and compared with the measured pile capacity.

Validation of the developed model

As previously mentioned, the developed Machine Learning model was validated using a randomly selected dataset of 8 piles cases. The regression coefficients derived from the 18 pile cases model training dataset were applied to predict the pile capacity of the eight (8) pile cases to validate the model's ability to extend its predictions to novel and unseen data.

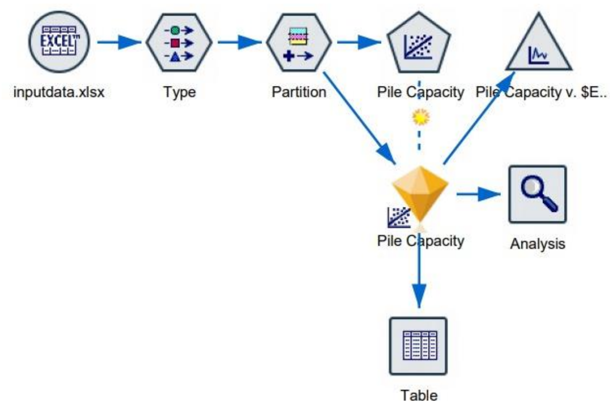

Figure 1. Stream flow diagram

Performance of the machine learning method

The performance of the MLR model was evaluated using statistical metrics, including coefficient of determination (R^2) which is a measure of fit between the measured capacity and predicted capacity from ML model. Furthermore, performance was evaluated on the basis of ML model uncertainty statistics (Mean, Standard Deviation, and Coefficient of Variation) which are a good indicator of the accuracy and precision of predicted capacities. The model uncertainty or model factor *(M)* was determined from Eq. 1.

$$M = \frac{Q_m}{Q_p} \quad [1]$$

Where Q_m = "capacity" interpreted from a load test, to represent the measured capacity; Q_p = capacity generally predicted using machine learning model, and *M* = model factor.

Comparison with the SA SPT method

For the Southern Africa SPT Method (SA-Method), the ultimate pile capacity for both shaft and base for all the 26 cases was computed using the following equations;

$$q_b = (N_1)_{60b} F_b \leq q_{max} \quad [2]$$

$$q_s = (N_1)_{60s} F_s \leq q_{max} \quad [3]$$

Furthermore, ultimate pile base and shaft capacity are determined by:

$$Q_b = q_b * A_b = q_b * (\frac{\pi d^2}{4}) \quad [4]$$

$$Q_s = q_s A_s = q_s(\pi dl) \quad [5]$$

$$Q_{ult} = Q_b + Q_s = q_b(\frac{\pi d^2}{4}) + q_s(\pi dl) \qquad [6]$$

Where Q_{ult} = ultimate pile capacity; Q_b = base pile capacity; Q_s = shaft pile capacity; q_b = base bearing pressure; q_s = shaft bearing pressure.

RESULTS AND DISCUSSIONS

Machine learning results

The multiple regression equation derived from the SPSS Modeler was formulated in the form of Eq. 8, where the coefficients represent the contributions of each independent variable to the prediction of pile design capacity.

Qp = -2641.984+30.112L + 66.468Nb +45.619Ns + 3988.987Ab + 63.725As [8]

Where Q_p: Predicted Pile Capacity, L= Length, N_b = SPTN for base, Ns = SPT-N for shaft, A_b =Area of base and As = Area of shaft

The measured capacities (Q_m) obtained and the predicted capacities (Q_p) from the ML model are presented in Table 1. The ensuing model factors computed as per Eq.1 are also shown in Table 1.

To gain better insight, the measured capacities were plotted against the predicted capacities as shown in Figure 2.

It can be seen from Figure 2 that the R^2 is 0.90 which indicate an excellent fit between the predicted and measured pile capacities. On the basis of the R^2, it can be inferred that the Machine Learning method predicts pile capacity with high accuracy.

Further insight is shown by the summary statistics of the model factor which are also shown in Fig. 2. At μ = 0.99 and COV = 0.30, the ML pile capacity prediction method is relatively very good. The superiority of Machine Learning method in predicting pile capacity compared to traditional empirical methods has been other studies (e.g. Gomes et al 2021, Shoaib and Abu-Farsakh, 2023).

For example, Gomes et al reported that all machine learning techniques investigated obtained a root mean squared error (RMSE) below 710, surpassing Meyerhof's and Décourt-Quaresma's semi-empirical methods, which both obtained RMSE values close to 900.

Table 1. Q_m and Q_p for ML model

Case	Q_m	Q_p	M = Q_m/Q_p
1	1427	1647	0.866
2	3319	3226	1.029
3	3315	3265	1.015
4	1224	992	1.234
5	887	939	0.944
6	1587	939	1.690
7	1087	774	1.405
8	1361	939	1.449
9	990	884	1.120
10	4706	4543	1.036
11	1232	1647	0.748
12	435	1307	0.333
13	1136	1143	0.994
14	653	541	1.205
15	943	1064	0.886
16	1695	1623	1.044
17	1227	1397	0.878
18	617	969	0.637

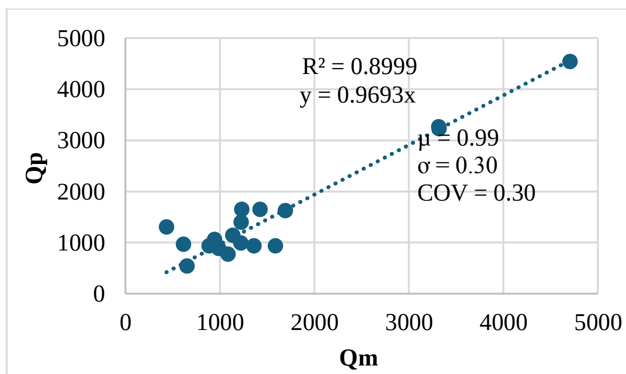

Figure 2- Scater plot of Qm Vs Qp

Validation results

The measured and predicted capacities for the 8 validation cases together with the model factors are presented in Table 2. Furthermore, a plot of the measured versus predicted capacities is presented in Figure 3. Further analysis of Figure 3 shows an R^2 of 0.74 which is good enough for geotechnical work. The Model factor statistics are even better at μ = 1.03 and COV = 0.24. This implies that the developed Machine Learning Model can be used to predict pile capacity for bored piles in non-cohesive soils outside the current database.

Table 2. Q_m and Q_p for validation cases

Case	Q_m	Q_p	$m = Q_m/Q_p$
1	1055	1514	0.70
2	1453	2110	0.69
3	1483	1217	1.22
4	1799	1458	1.23
5	2375	2247	1.06
6	3094	2878	1.08
7	4606	3618	1.27
8	3154	2988	1.06

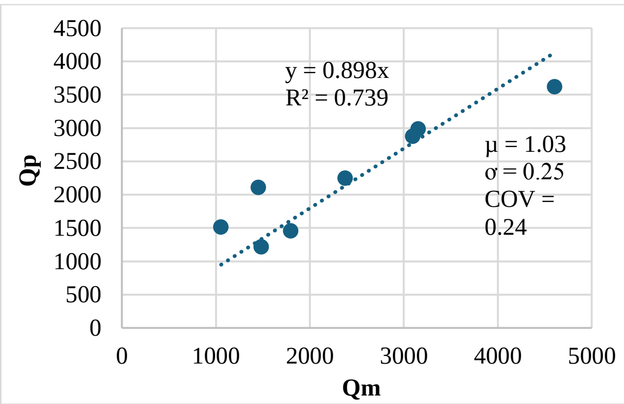

Figure 3: Scater plot of Qm Vs Qp for validation data

Results for SA SPT-based method

For comparison purposes, the measured capacities, predicted capacities and model factors of the SA- SPT based method are presented in Table 3. The R^2 is 0.84 (Figure 4) which is good but lower that that yielded by the ML method. However, the Model factor statistics at μ = 0.87 and COV = 0.25 is slightly better, especially the COV which is a measure of variability. The key point here is that the results of the ML method are comparable with that of the well-established SA- SPT based method.

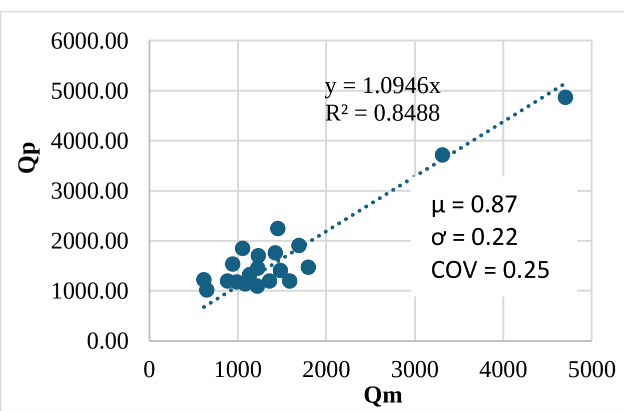

Figure 4: Scater plot of Qm Vs Qp for SA SPT method

Table 3: Q_m and Q_p for SA- SPT based method

Case	Measured pile capacity	Predicted pile capacity	$M = Q_m/Q_p$
1	1427	1754.97	0.813
2	3319	4583.09	0.724
3	3315	3713.36	0.893
4	1224	1088.5	1.124
5	887	1192.31	0.744
6	1587	1192.31	1.331
7	1087	1134.12	0.958
8	1361	1192.31	1.141
9	990	1172.91	0.844
10	4706	4860.7	0.968
11	1232	1694.94	0.727
12	435	1586.46	0.274
13	1136	1321.98	0.860
14	653	1014.73	0.643
15	943	1528.82	0.617
16	1695	1898.7	0.893
17	1227	1448.27	0.847
18	617	1216.49	0.507
19	1055	1842.94	0.572
20	1453	2242.7	0.648
21	1483	1403.7	1.056
22	1799	1467.71	1.225
23	2375	4721.78	0.503
24	3094	5721.56	0.541
25	4606	9545.19	0.483
26	3154	7954.32	0.396

CONCLUSIONS

A plot of Q_m versus Q_p for the Machine Learning approach yielded a high R^2 of 0.9, which is an indication that the method is very accurate in predicting pile capacity. Model validation results produced R^2 of 0.74 and model factor statistics of μ = 1.03 and COV = 0.24 which are very good. This implies that the developed Machine Learning Model can be used to predict pile capacity for bored piles in non-cohesive soils outside the current database. The results of the well-established SA-SPT method (R^2 of 0.84 and M- statistics of μ = 0.87 and COV = 0.25) is comparable to that of the ML method.

For further research, it is recommended that more pile load test data be collected for both training and validating the Machine Learning Model.

DECLARATIONS

Corresponding author

Correspondence and requests for materials should be addressed to Mahongo Dithinde; E-mail:

dithinde@ub.ac.bw; ORCID: https://orcid.org/0000-0002-4541-7438

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Author's contribution

Mahongo Dithinde has collected a lot of pile load test data across Southern Africa for various soils condition and pile types. Lesego Palalane applied Machine learning method to ultimate capacity of piles in non-cohesive soils extracted from the load test database availed by Dithinde.

Acknowledgements

The authors would like to Faculty IT Unit for assistance with SPSS modeler required for Machning Learning analysis for this study.

Consent to publish

Not applicable.

Competing interests

The authors declare no competing interests in this research and publication.

REFERENCES

Chin, F.K. 1970. Estimation of Pile Not Carried to Failure, Proceedings 2nd Southeast Asian Conference on Soil Engineering, Singapore, 81-90.

Decourt, L. (1999). Behaviour of foundations under working load conditions. Proceedings of the 11th Pan American Conference on Soil Mechanics and Geotechnical Engineering, Foz Dolguassu, Brazil, Vol. 4, pp. 453-488.

Decourt, L., (1995). Prediction of load-settlement relationships for foundations on the basis of the SPT T, Ciclo de Conferencias Internationale, Leonardo Zeevaert, UNAM, Mexico, 1995, pp. 85-104.

Fleming, W. G. K. (1992). A new method for single pile settlement prediction and analysis; Geotechnique, Vol. 42, No. 3, pp. 411-425. https://doi.org/10.1680/geot.1992.42.3.411

Johnson, K., Karunasena, W., Sivakugan, N., & Guazzo, A. (2001). Modeling Pile-Soil Interaction Using Contact Surfaces. In S. Valliappan & N. Khalili (Eds.), Computational Mechanics–New Frontiers for the New Millennium (pp. 375-380). Elsevier. ISBN 9780080439815. https://doi.org/10.1016/B978-0-08-043981-5.50058-4.

Horvitz G.E., Stettler D.R. and Crowser J.C. (1981. Comparison of predicted and observed pile capacity. In Proceedings of a session sponsored by the Geotechnical Engineering Division at the ASCE National Convention (pp. 413-433).

Meyerhof, G. G. (1976). Bearing capacity and settlement of pile foundations. Journal of Geotechnical Engineering, ASCE, Vol. 102, No. 3, pp. 196-228. https://doi.org/10.1061/AJGEB6.0000243.

Seed B., Tokimatsu K., Harder M and Chung. (1985). Influence of SPT Procedures in Soil Liquefaction Resistance Evaluations. Journal of Geotechnical Engineering, ASCE. 111(12): 1425-1445. https://doi.org/10.1061/(ASCE)0733-9410(1985)111:12(1425)

Shariatmadari, N., Eslami, A., & Karimpour-Fard, M. (2008). Bearing capacity of driven piles in sands from SPT–applied to 60 case histories. Iranian Journal of Science & Technology, Transaction B, Engineering, 32(B2), 125-140.

Skempton A. W. 1986. Standard Penetration Test Procedures and the Effects in Sands of Overburden Pressure, Relative Density, Particle Size, Ageing and Over consolidation. Geotechnique. 36(3): 25-447.

https://doi.org/10.1680/geot.1986.36.3.425.

Shoaib, M. M., & Abu-Farsakh, M. Y. (2023). Exploring Tree-Based Machine Learning Models to Estimate the Ultimate Pile Capacity From Cone Penetration Test Data. Transportation Research Record. Advance online publication. https://doi.org/10.1177/03611981231170128.

Yago, Gomes & Verri, Filipe & Ribeiro, Dimas. (2021). Use of machine learning techniques for predicting the bearing capacity of piles. Soils and Rocks. 44. 1-14. https://doi.org/10.28927/SR.2021.074921.

Zhang, W., Goh, A. T. C., Zhang, R., Li, Y., & Wei, N. (2020). Chapter 19 - Back-propagation neural network modeling on the load–settlement response of single piles. In P. Samui, D. T. Bui, S. Chakraborty, & R. C. Deo (Eds.), Handbook of Probabilistic Models (pp. 467-487). Butterworth-Heinemann. https://doi.org/10.1016/B978-0-12-816514-0.00019-9.

Appendix 1: Pile load tests database

Case	**Soil type**		**Pile type**	**Shaft dia. (mm)**	**Base dia.(mm)**	**SPT N-value**		**Length (m)**
	Base	**Shaft**				**base**	**Shaft**	
1	**Medium dense sandy gravel**	**Medium dense sandy gravel**	**Auger**	**430**	**430**	**30**	**20**	**8**
2	Gravel	Gravel	Auger	600	750	30	20	9
3	Gravel	Gravel	Auger	750	600	30	20	11
4	Sand	Sand	CFA	360	350	29	14	7.8
5	Sand	Sand	CFA	400	400	23	13	9.5
6	Sand	Sand	CFA	400	400	23	13	9.5
7	Sand	Sand	CFA	400	400	23	13	8
8	Sand	Sand	CFA	400	400	23	13	9.5
9	Sand	Sand	CFA	400	400	23	13	9
10	Sand	Sand	Auger	520	520	38	38	16.5
11	Sand	Sand	Auger	430	430	25	17	11.5
12	Dense sand	Sand	Auger	450	450	25	15	9
13	Sand	Sand	CFA	400	400	22	17	10
14	Sand	Sand	CFA	400	400	19	16	7
15	Sand	Sand	CFA	500	500	19	16	7.8
16	Sand	Sand	CFA	500	500	22	17	10
17	Sand	Sand	CFA	400	400	25	16	11
18	Sand	Sand	CFA	400	400	24	13	9.2
19	Sand	Sand	CFA	500	500	25	14	9.5
20	Sand	Sand	CFA	500	500	26	17	11.8
21	Sand	Sand	CFA	500	500	13	14	12.3
22	Sand	Sand	CFA	500	500	13	13	14.5
23	very dense sand	Sand	Franki	520	760	17	9	12
24	Dense gravel	medium sand	Franki	520	760	17	14	15
25	dense sand	medium sand	Franki	520	760	36	32	6
26	very dense sand	medium sand	Franki	520	760	30	20	8

Appendx 2: Input data used in MLR Model

Case No.	Length	SPT N-base	SPT N-shaft	Area-base	Area-shaft	Q_m
1	8	28.5	19	0.15	10.81	1427
2	9	28.5	19	0.44	16.96	3319
3	11	28.5	20	0.28	25.92	3315
4	7.8	27.55	13.3	0.1	8.82	1224
5	9.5	21.85	12.35	0.13	11.94	887
6	9.5	21.85	12.35	0.13	11.94	1587
7	8	21.85	12.35	0.13	10.05	1087
8	9.5	21.85	12.35	0.13	11.94	1361
9	9	21.85	12.35	0.13	11.31	990
10	16.5	36.1	38	0.21	26.95	4706
11	11.5	23.75	17	0.15	15.54	1232
12	9	23.75	14.25	0.16	12.72	435
13	10	20.9	17	0.13	12.57	1136
14	7	18.05	15.2	0.13	8.8	653
15	7.8	18.05	15.2	0.2	12.25	943
16	10	20.9	17	0.2	15.71	1695
17	11	23.75	16	0.13	13.82	1227
18	9.2	22.8	12.35	0.13	11.56	617
19	9.5	23.75	11.9	0.2	14.92	1055
20	11.8	24.7	17	0.2	18.54	1453
21	12.3	12.35	14	0.2	19.32	1483
22	14.5	12.35	13	0.2	22.78	1799
23	12	16.15	9	0.45	19.6	2375
24	15	16.15	14	0.45	24.5	3094
25	6	34.2	30.4	0.45	9.8	4606
26	8	28.5	19	0.45	13.07	3154