TECHNICAL REPORT STANDARD PAGE

1. Title and Subtitle
Update the Pile Design by CPT Software to Incorporate Newly Developed Pile-CPT Methods and Other Design Features

2. Author(s)
Murad Abu-Farsakh, Mohsen Amirmojahedi, Md Ariful Mojumder and Mohammad Moontakim Shoaib

3. Performing Organization Name and Address
Louisiana Transportation Research Center
4101 Gourrier Avenue
Baton Rouge, LA 70808

4. Sponsoring Agency Name and Address
Louisiana Department of Transportation and Development
P.O. Box 94245
Baton Rouge, LA 70804-9245

5. Report No.
FHWA/LA.23/682

6. Report Date
September 2023

7. Performing Organization Code
LTRC Project Number: 17-2GT
SIO Number: DOTLT1000165

8. Type of Report and Period Covered Final Report
July 2017 – March 2023

9. No. of Pages
323

10. Supplementary Notes
Conducted in Cooperation with the U.S. Department of Transportation, Federal Highway Administration

11. Distribution Statement
Unrestricted. This document is available through the National Technical Information Service, Springfield, VA 21161.

12. Key Words
Ultimate pile capacity, Driven piles, Cone penetration test, Statistical analysis, Multidimensional unfolding, LRFD reliability analysis, Machine learning.

13. Abstract
This study presents the performance evaluation of 21 direct pile-CPT methods for estimating the ultimate load carrying capacity of square precast prestressed concrete (PPC) piles driven into Louisiana soils utilizing the cone penetration test (CPT) data. The investigated methods are: Schmertmann, De Ruiter and Beringen, Bustamante and Gianeselli (LCPC), Philipponnat, Price and Wardle, Zhou, Tumay and Fakhroo, UF (2007), probabilistic, Aoki and De Alencar, Penpile, NGI, ICP, UWA, CPT2000, Fugro, Purdue, German, Eurocode7, ERTC3, and Togliani direct pile-CPT methods. A search was conducted in the DOTD files to identify pile load test reports with CPT soundings adjacent to test piles. A database of 80 pile load tests that were loaded to failure, were identified, collected, and used in analysis. The measured ultimate load carrying capacity for each pile was interpreted from the pile load test using the Davisson and modified Davisson interpretation methods. The ultimate pile capacities estimated from the pile-CPT methods were compared with the measured ultimate pile capacities. In this study, three approaches were adopted to evaluate the performance of pile-CPT methods. In the first

approach, three statistical criteria were used: the best fit line of predicted (Q_p) versus measured (Q_m) capacity, arithmetic mean and standard deviation of Q_p/Q_m, and the cumulative probability of Q_p/Q_m. The results of this evaluation showed the following best-performed pile-CPT methods in order: LCPC, ERTC3, Probabilistic, UF, Philipponnat, De Ruiter and Beringen, CPT2000, UWA, and Schmertmann methods. The second approach used to evaluate the 21 pile-CPT methods is the MultiDimensional Unfolding (MDU), which showed similar ranking of top-performed pile-CPT methods. The third approach used for evaluating the pile-CPT methods was based on LRFD reliability analysis in terms of resistance factor and efficiency, and the results of evaluation are consistent with the previous two criteria.

The top-performed pile-CPT methods were further analyzed using the MDU analysis to evaluate the methods' similarity, and the results showed that the methods can be divided into three similar groups: Group 1: Philipponnat, UF, Probabilistic, LCPC, and De Ruiter methods; Group 2: Schmertmann and ERTC3 methods; and Group 3: UWA and CPT2000.

The collected pile load tests were further divided into four categories based on soil type and used to develop combined pile-CPT methods to estimate the ultimate pile capacity for each soil category. In addition, another combined pile-CPT method was developed for the general case for all piles without considering soil category. The evaluation results showed that the developed combined pile-CPT methods significantly improved the estimation of ultimate pile capacity.

Four machine learning (ML) techniques including the artificial neural network (ANN) and three tree-based techniques [decision tree (DT), random forest (RF), and gradient boosted tree (GBT)] were also used to develop models to estimate the ultimate pile capacity from CPT data. The comparison results between the ML models and selected direct pile-CPT methods demonstrated that the ANN and GBT models substantially outperform the top-performed pile-CPT methods in all evaluation criteria.

Project Review Committee

Each research project will have an advisory committee appointed by the LTRC Director. The Project Review Committee is responsible for assisting the LTRC Administrator or Manager in the development of acceptable research problem statements, requests for proposals, review of research proposals, oversight of approved research projects, and implementation of findings.

LTRC appreciates the dedication of the following Project Review Committee Members in guiding this research study to fruition.

LTRC Administrator/Manager

Zhongjie "Doc" Zhang, Ph.D., P.E.
Pavement and Geotechnical Research Manager

Members

Jesse Rauser, Chris Nickel, Naser Abu-Hejleh, Steven Saye, Francisco Gudiel, James Melton, Kristy Smith, Arturo Aguirre

Directorate Implementation Sponsor

Christopher P. Knotts, P.E.
DOTD Chief Engineer

Update the Pile Design by CPT Software to Incorporate Newly Developed Pile-CPT Methods and Other Design Features

By
Murad Y. Abu-Farsakh
Mohsen Amirmojahedi
Md Ariful Mojumder
Mohammad Moontakim Shoaib

Louisiana Transportation Research Center
4101 Gourrier Avenue
Baton Rouge, LA 70808

LTRC Project No. 17-2GT
SIO No. DOTLT1000165

conducted for
Louisiana Department of Transportation and Development
Louisiana Transportation Research Center

September 2023

Abstract

This study presents the performance evaluation of 21 direct pile-CPT methods for estimating the ultimate load carrying capacity of square precast prestressed concrete (PPC) piles driven into Louisiana soils utilizing the cone penetration test (CPT) data. The investigated methods are: Schmertmann, De Ruiter and Beringen, Bustamante and Gianeselli (LCPC), Philipponnat, Price and Wardle, Zhou, Tumay and Fakhroo, UF (2007), probabilistic, Aoki and De Alencar, Penpile, NGI, ICP, UWA, CPT2000, Fugro, Purdue, German, Eurocode7, ERTC3, and Togliani direct pile-CPT methods. A search was conducted in the DOTD files to identify pile load test reports with CPT soundings adjacent to test piles. A database of 80 pile load tests that were loaded to failure, were identified, collected, and used in analysis. The measured ultimate load carrying capacity for each pile was interpreted from the pile load test using the Davisson and modified Davisson interpretation methods. The ultimate pile capacities estimated from the pile-CPT methods were compared with the measured ultimate pile capacities. In this study, three approaches were adopted to evaluate the performance of pile-CPT methods. In the first approach, three statistical criteria were used: the best fit line of predicted (Q_p) versus measured (Q_m) capacity, arithmetic mean and standard deviation of Q_p/Q_m, and the cumulative probability of Q_p/Q_m. The results of this evaluation showed the following best-performed pile-CPT methods in order: LCPC, ERTC3, Probabilistic, UF, Philipponnat, De Ruiter and Beringen, CPT2000, UWA, and Schmertmann methods. The second approach used to evaluate the 21 pile-CPT methods is the MultiDimensional Unfolding (MDU), which showed similar ranking of top-performed pile-CPT methods. The third approach used for evaluating the pile-CPT methods was based on LRFD reliability analysis in terms of resistance factor and efficiency, and the results of evaluation are consistent with the previous two criteria.

The top-performed pile-CPT methods were further analyzed using the MDU analysis to evaluate the methods' similarity, and the results showed that the methods can be divided into three similar groups: Group 1: Philipponnat, UF, Probabilistic, LCPC, and De Ruiter methods; Group 2: Schmertmann and ERTC3 methods; and Group 3: UWA and CPT2000.

The collected pile load tests were further divided into four categories based on soil type and used to develop combined pile-CPT methods to estimate the ultimate pile capacity for each soil category. In addition, another combined pile-CPT method was developed for the general case for all piles without considering soil category. The evaluation results showed that the developed combined pile-CPT methods significantly improved the estimation of ultimate pile capacity.

Four machine learning (ML) techniques including the artificial neural network (ANN) and three tree-based techniques [decision tree (DT), random forest (RF), and gradient boosted tree (GBT)]

were also used to develop models to estimate the ultimate pile capacity from CPT data. The comparison results between the ML models and selected direct pile-CPT methods demonstrated that the ANN and GBT models substantially outperform the top-performed pile-CPT methods in all evaluation criteria.

Acknowledgments

This research project was funded by the Louisiana Department of Transportation and Development (DOTD) (SIO No. DOTLT1000165) and the Louisiana Transportation Research Center (LTRC Project No. 17-2GT). The help and support of Zhongjie Zhang, Ph.D., at LTRC is gratefully acknowledged.

Implementation Statement

This study aimed at evaluating the direct pile-CPT methods for estimating the ultimate axial capacity of square PPC piles driven into Louisiana soils and updating the Louisiana Pile Design from cone penetration test (LPD-CPT) software to include the top-performing methods. The findings of this study can be used effectively in the design and analysis of pile foundations as summarized below:

a) The six top-performed direct pile-CPT methods in addition to the modified Schmertmann method and the combined pile-CPT method were implemented in LPD-CPT for friendly use by DOTD engineers to better and more accurately design and analyze pile foundations. The user can input the applied load and pile size to calculate the required pile length.

b) The calibrated resistance factors provided for the different pile-CPT methods can be used in the load and resistance factor design (LRFD) of the pile foundation.

c) The developed machine learning (ML) models can be effectively implemented into a special code or LPD-CPT software to better estimate the ultimate axial capacity of PPC piles from pile characteristics, soil type, and CPT input data.

d) The effect of scour on the long-term ultimate pile capacity was implemented into the LPD-CPT software based on FHWA guidelines by considering the change on overburden pressure for sand layers.

e) The LPD-CPT software was updated to include batch analysis for the different bents in a bridge for given pile size, ground surface elevations, local scour elevation, and the load and resistance factor (LRFD).

Table of Contents

List of Tables

List of Figures

Introduction

Design engineers usually consider using deep foundations when the conditions of the upper soil layers are weak and unable to withstand and support the structural loads. Piles help transfer these loads deep in the ground through their interactions with the surrounding soils. Therefore, the safety and stability of pile-supported structures depend on the behavior of piles.

Most soil deposits in southern Louisiana are soft in nature. In addition, the high percentage of wetlands, marshes, swamps, bayous, rivers, and lakes makes it necessary to consider deep foundations in the design of transportation infrastructure. Therefore, pile foundations are frequently used by the Louisiana Department of Transportation and Development (DOTD) to support highway bridges and other transportation infrastructures. Square precast prestressed concrete piles (PPC) are the most common piles currently used in DOTD projects.

Piles are expensive structural members, and pile projects are always costly. Current DOTD practice of pile design is based on the static analysis (α-method) and sometimes in conjunction with the dynamic analysis using the Pile Driving Analyzer^{TM}. Soil properties are needed as input parameters for the static analysis. Therefore, it is necessary to conduct field and laboratory tests, which include soil boring, standard penetration test, unconfined compression test, soil classification, etc. Running these field and laboratory tests is expensive and time consuming. The cost of traditional soil boring and the associated laboratory tests in Louisiana is between $14,000 and $15,000, depending on the sampling depth and the laboratory tests involved.

Due to the uncertainties associated with pile design, load tests are usually conducted to verify the design loads and to evaluate the actual response of the pile under loading. Pile load tests are also expensive (the average cost of a pile load test in Louisiana including pile is about $50,000). Moreover, pile load tests are a verification tool for pile design and they cannot be a substitute for the engineering analysis of the pile behavior.

The use of in-situ tests, such as the cone penetration test (CPT), that are performed under existing stresses and boundary conditions in the field, can provide faster, and more accurate and reliable estimation of pile capacity than the traditional methods.

The CPT has been widely recognized as a preferred tool for site characterization and evaluation of soil properties. The test is a simple, fast, repeatable, and a cost-effective in-situ test that can provide continuous soundings of reliable soil measurements, especially when compared to traditional site characterization (borings and laboratory tests). During penetration, the CPT measures the tip resistance (q_c), sleeve friction (f_s), and also excess pore pressures (u) when the

piezocone (PCPT or CPTu) is used. The CPT technology can be effectively utilized for soil identification, classification, and for the evaluation of different soil properties, such as strength and deformation characteristics of the soil. Implementation of the CPT can drastically decrease the number of soil borings and reduce the cost and time required for subsurface characterization. Therefore, implementation of the CPT technology by DOTD in different engineering applications, such as the estimation of pile capacity, should be seriously considered.

Due to the similarity between the cone and the pile, the estimation of pile capacity utilizing the CPT test data is considered among the earliest applications of the CPT. The test can provide valuable and continuous information records with depth that can be interpreted for pile capacity. Therefore, the in situ characteristics of the soil are available to design engineers at a particular point. The pile design methods that utilize the CPT data proved to predict the pile capacity within an acceptable accuracy.

Several direct pile-CPT methods have been proposed in literature to estimate the pile capacity utilizing the CPT data, which correlate the uncorrected and corrected cone resistance, (q_c, q_t) and sleeve friction (f_s) to the ultimate pile capacity (Q_u) using some reduction factors due to scale effects, penetration rate, pile type, pile installation, etc. (e.g., [1], [2], [3], [4], [5], [6], [7], [8], [9]; [10], [11], [12], [13]). Several studies have been carried out by different researchers to evaluate the capability of the different direct pile-CPT methods for estimating the measured ultimate pile capacity from load tests (e.g., [14], [15], [16], [17]). The study carried out by Robertson et al. [14] on eight pile load tests showed that the pile capacities predicted using Schmertmann [1], De Ruiter and Beringen [2], and Bustamante and Gianeselli [3] methods fit the measured capacities better than other methods. Briaud and Tucker [15] evaluated six CPT methods using 98 pile load tests and concluded that the Bustamante and Gianeselli [3] method gave the best fit between measured and predicted pile capacities. Another study by Tand and Funegard [16] showed that the predicted capacities using the De Ruiter and Beringen [2] method showed the best fit to the measured capacities.

A previous study was conducted by Abu-Farsakh and Titi [17] using a database of 35 pile load tests to identify the most appropriate CPT methods to estimate the ultimate capacity of driven PPC friction piles in the state of Louisiana. Eight direct pile-CPT methods were evaluated based on their capability to predict the measured ultimate pile capacity. Based on this evaluation, the De Ruiter and Beringen [2] and Bustamante and Gianeselli [3] methods were identified as the best performance methods. These two methods, in addition, to the Schmertmann [1] method were implemented into a visual basic computer program (Louisiana Pile Design from CPT) for use by DOTD engineers in the analysis and design of friction piles.

Generally, pile design depends on soil conditions, pile characteristics, and driving and installation conditions. Local experience usually played an important role in design/analysis of piles. Therefore, it is essential to take advantage of the DOTD experience in the CPT technology to identify suitable CPT design methods. Implementation of the CPT (in conjunction with the currently used method) in the analysis/design of piles will foster confidence in the CPT technology. With time and experience, the role of the CPT can be increased while the role of traditional subsurface exploration is reduced.

This report presents the current research effort undertaken at the Louisiana Transportation Research Center (LTRC) to identify the most appropriate pile-CPT methods for estimating the ultimate axial load carrying capacity of piles driven into Louisiana soils. To achieve this goal, state projects that have both pile load tests and CPT soundings were identified and collected from DOTD files. Pile load test reports were selected based on selection criteria, compiled onto sheets, and analyzed. A database of 80 pile load tests and corresponding CPT test data were collected. The ultimate axial load carrying capacity for each pile was determined using the Davisson interpretation method [18]. The CPT soundings close to the test pile location were identified and used to predict the ultimate pile capacity. Twenty one direct pile-CPT methods were selected and evaluated in this study for their capabilities to estimate the ultimate pile capacity of PPC driven piles by CPT were selected. Detailed description of these methods are presented in Appendix A of this report. The ultimate pile load carrying capacities predicted by the CPT methods were compared with the ultimate capacities obtained from pile load tests using the Davisson method. Statistical analysis, MultiDimensional Unfolding (MDU), and Reliability analysis were used to investigate the performance of the 21 pile-CPT methods. In our study, it was shown that the estimation of top-ranked pile-CPT methods can be used in a combined method to yield an optimized method for predicting axial capacity of driven piles.

This report also explores the potential application of artificial intelligence (AI) and machine learning (ML) techniques to develop models for estimating the ultimate pile capacity utilizing the CPT data. This includes the artificial neural network (ANN), the most widely used AI method; and three tree-based ML methods: the decision tree (DT), random forest (RF), and gradient boosted tree (GBT). It is expected that the AI and ML techniques will resolve some of the shortcomings in traditional direct pile-CPT methods that involves assumptions and judgments in selecting the proper correlation coefficients between the CPT data and pile parameters for estimating the ultimate capacity of PPC driven piles. The results of AI and ML models were compared with the results of pile load tests as well as the results of top-performed direct pile-CPT methods to demonstrate its accuracy and bolster its reliability and feasibility.

Finally, the top-performed pile-CPT methods in addition to the developed combined pile-CPT method were implemented into the Louisiana Pile Design from cone penetration test (LPD-CPT) program for friendly use by Louisiana engineers to design PPC piles utilizing CPT technology.

Objectives

The primary objectives of this research project were:

- Collect all available pile load tests database from Louisiana DOTD and the corresponding CPT soundings and soil borings close to the test pile locations;
- Evaluate/rank the different direct pile-CPT methods for estimating the ultimate axial load carrying capacity of driven PPC piles from the cone penetration test (CPT) data for use in Louisiana soils;
- Select, modify and/or develop a new pile-CPT method for use in the design of piles driven in Louisiana soils;
- Identify the most appropriate pile-CPT methods for implementing into the LPD-CPT software;
- Re-calibrate the resistance factor (ϕ) for the selected pile-CPT methods;
- Update the Louisiana Pile Design-Cone Penetration Test (LPD-CPT) software to incorporate the newly selected pile-CPT prediction methods; and
- Update the "LPD-CPT" software to incorporate new features, such as the effect of scour on the long-term pile capacity, implement the calibrated resistance factors for the pile-CPT methods, implement pile setup empirical equations, and generate synthetic CPT profiles.

Scope

This research effort was focused on evaluating the capability of 21 direct pile-CPT methods for accurately estimating the ultimate axial load carrying capacity of driven piles utilizing CPT data. These methods are described in detail in Appendix A of this report. The predicted capacity was compared to the reference pile load capacity obtained from the pile load test using Davisson interpretation method.

The direct pile-CPT methods were used to investigate the load carrying capacity of square precast prestressed concrete (PPC) piles of different sizes driven into Louisiana soils. Other pile types such as timber piles and steel pipes were not covered in the current analyses.

To achieve the objective of this study, a total of 104 pile load tests database and corresponding CPT soundings and soil borings close to the test pile locations were initially collected from DOTD files. However, only 80 pile load tests that were loaded to failure during the load test were included in this study.

Different evaluation techniques were adopted in this study to evaluate and identify the best-performed direct pile-CPT methods for estimating the ultimate capacity of PPC piles: (a) an evaluation based on mathematical and statistical analysis; (b) an evaluation using MultiDimensional Unfolding; and (c) an evaluation based on reliability analysis.

A combined method from the best-performed Pile-CPT methods was developed based on contribution of sand layers to total ultimate capacity (kind of optimization). In addition, new pile-CPT methods were developed using the artificial intelligent and machine learning techniques [artificial neural networks (ANN), decision trees (DT), random forests (RF), gradient boosted tree (GBT)].

The top-performed eight pile-CPT methods in addition to the combined pile-CPT method were implemented into the Louisiana Pile Design from cone penetration test (LPD-CPT) program.

Literature Review

Axial Capacity of Piles

Piles are relatively long and generally slender structural foundation members that transmit superstructure loads to deep soil layers. In geotechnical engineering, piles usually serve as foundations when soil conditions are not suitable for the use of shallow foundations.

The behavior of the pile depends on many different factors, including pile characteristics, soil conditions and properties, installation method, and loading conditions. The performance of piles affects the serviceability of the structure they support.

Based on some factors including the mechanism of load transfer (friction and end-bearing piles), volume of soil displacement, and pile's material (concrete, steel, timber, etc.), different pile classification systems have been introduced in literature. These factors determine the pile's behavior that affect the serviceability of the supported structure. For example, the behavior of friction piles is mostly dependent on the pile-soil interface friction; while for end-bearing piles, most of the pile capacity comes from interaction between the pile's tip and the soil located in the tip area known as influence zone.

The prediction of pile load carrying capacity can be achieved using different methods such as pile load test, dynamic test, statnamic test, static analysis based on soil properties from laboratory tests, and static analysis utilizing the results of in situ tests such as the cone penetration test.

The ultimate axial load carrying capacity of the pile (Q_u) composed of the end-bearing capacity of the pile (Q_b) and the shaft friction capacity (Q_s). The general equation described in the literature is given by:

$$Q_u = Q_b + Q_s = q_b A_b + \sum_{i=1}^{n} f_i A_{si} \quad (1)$$

where, q_b is the unit tip bearing capacity, A_b is the area of the pile tip, f_i is the unit skin friction of the soil layer i, and A_{si} is the area of the pile shaft in the soil layer i. In sands, the end-bearing capacity (Q_b) dominates, while in soft clays the shaft friction capacity (Q_s) dominates. The design load carrying capacity (Q_d) of the pile can be calculated by:

$$Q_d = \frac{Q_u}{F.S.} \tag{2}$$

where, Q_u is the ultimate load carrying capacity and *F.S.* is the factor of safety.

Cone Penetration Test

The cone penetration test has been recognized as one of the most widely used in situ tests. Cone penetration test (CPT) was introduced by Swedish Railways in 1917. Dutch Mantle cone with 10 cm^2 area and 60° apex angle was introduced in 1936. The first electronic penetrometer was introduced in 1948. In 1953, a separated sleeve for measuring the sleeve friction resistance introduced by Begemann (friction cone penetrometer).

The cone penetration test consists of advancing a cylindrical rod with a conical tip into the soil and measuring the forces required to push this rod. The friction cone penetrometer measures two resistance forces during penetration: the total tip resistance (q_c), which is the soil resistance (within influence zone) to advance the cone tip, and the sleeve friction (f_s), which is the sleeve friction developed between the surrounding soil and the sleeve of the cone penetrometer. The ratio of sleeve friction to the tip resistance is known as the friction ratio (R_f), which is expressed in a percentage. A schematic of the electric cone penetrometer is depicted in Figure 1. The CPT resistance parameters (q_c, f_s) has been widely used to classify and identify soil strata and to evaluate the strength, stiffness, and the deformation characteristics of the soils.

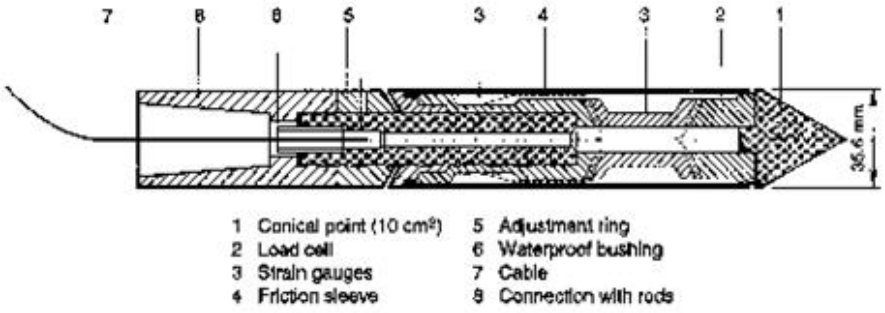
The cone penetration test (CPT) data has been used to estimate the ultimate axial pile load carrying capacity. Several methods are available in the literature to estimate the axial pile capacity utilizing the CPT data. These methods can be classified into two well-known approaches:

1) Direct approach, in which
   - The unit tip bearing capacity of the pile (q_b) is evaluated from the cone tip resistance (q_c) profile.
   - The unit skin friction of the pile (f) is evaluated from either the sleeve friction (f_s) profile or the cone tip resistance (q_c) profile.
2) Indirect approach: in which the CPT data (q_c and f_s) are first used to evaluate the soil strength parameters such as the undrained shear strength (S_u) and the angle of internal friction (φ). These parameters are then used to evaluate the unit end bearing capacity of the

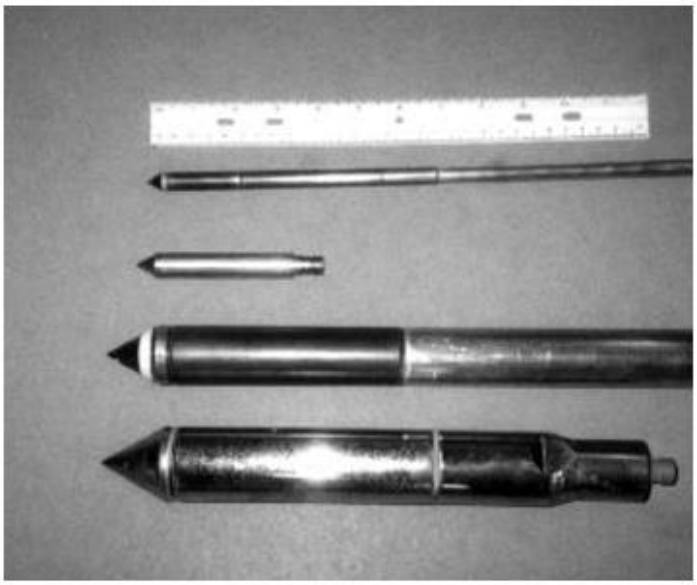
pile (q_b) and the unit skin friction of the pile (f) using formulas derived based on semi-empirical/theoretical methods.

In the current research, only the direct methods for estimating the ultimate pile capacity from CPT data are investigated (referred to as direct pile-CPT methods).

Figure 1. The electric cone penetrometer



(a) Schematic of the electric friction cone penetrometer



(b) The 1.27, 2, 10, and 15 cm^2 cone penetrometers

Determining Soil Type Using CPT

Most of the pile-CPT methods include different correlation equations for evaluating the unit end bearing capacity of the pile (q_b) and the unit skin friction of the pile (f) from the CPT data (q_c and f_s) in different soil types. Therefore, it is essential to evaluate the soil layering and determine the soil type for better calculating the pile capacity.

Soil classification and identification of soil stratigraphy can be achieved by analyzing the CPT data. Clayey soils usually show low cone tip resistance, high sleeve friction, and therefore high friction ratio, while sandy soils show high cone tip resistance, low sleeve friction, and low friction ratio. Many soil classification methods based on CPT employ the CPT data to identify the soil type from classification charts.

There are different soil classification methods proposed by different researchers such as Schmertmann [19], Douglas and Olsen [20], Robertson et al. [21], and Campanella et al. [22]. In this study, two soil soil-behavior type CPT classification methods were used for all pile-CPT methods: the probabilistic region estimation method for soil classification by Zhang and Tumay [23] and Robertson 2010 soil classification [24]. Description of these two CPT soil classification methods are presented below.

It should be noted that the detailed implementation of each soil classification method for each pile-CPT method has been explained at each pile-CPT description section.

Probabilistic Soil Classification

Zhang and Tumay [23] proposed the probabilistic region estimation method for soil behavior classification from CPT data, which is similar to the classical soil classification methods, where it is based on soil composition according to the Unified Soil Classification System (USCS). The method identifies three soil type behaviors: clayey, silty, and sandy soils. The probabilistic region estimation determines the probability of each soil behavior (clayey, silty, and sandy) at a certain depth.

In this method, a conformal mapping was introduced based on Douglas and Olsen [20] chart to transfer the CPT data to the soil classification index (U). The soil classification index, U, provides a soil profile over depth with the probability of belonging to different soil types, which more realistically and continuously reflects the in situ soil characterization, which includes the spatial variation of soil types. The conformal transformation is accomplished using the following equations:

$$x = 0.1539\, R_f + 0.8870 \log q_c - 3.35 \tag{3}$$

$$y = -0.2957\, R_f + 0.4617 \log q_c - 0.37 \tag{4}$$

The soil classification index (U) is obtained from:

$$U = \frac{(a_1x - a_2y + b_1)(c_1x - c_2y + d_1)}{(c_1x - c_2y + d_1)^2 + (c_2x + c_1y + d_2)^2} - \frac{(a_2x + a_1y + b_2)(c_2x + c_1y + d_2)}{(c_1x - c_2y + d_1)^2 + (c_2x + c_1y + d_2)^2} \quad (5)$$

where, a_1, a_2, b_1, b_2, c_1, c_2, d_1,and d_2 are -11.345, -3.795, 15.202, 5.085, -0.296, -0.759, -2.960, and 2.477, respectively.

A statistical correlation was then established between the U index and the compositional soil type given by the USCS classification system. A normal distribution of U was established for each reference USCS soil type (i.e., GP, SP, SM, SC, ML, CL, and CH). Each U value corresponds to several soil types with different probabilities. Boundary values were used to divide the U axis into seven regions, as shown in The equations for the different soil curves are given as follows:

For U < 0.14: $\text{sandy} = 0.0035$, $\text{silty} = 0.0184$, $\text{clayey} = 0.9781$

For U = 2.91: $\text{sandy} = 0.9771$, $\text{silty} = 0.0229$, $\text{clayey} = 0.000$

For -0.14 < U < 2.91:

$$\text{sandy} = 0.00132408 + 0.074195U + 0.0900763U^2 \quad (6)$$

$$\text{silty} = 0.147853 + 0.896769U - 0.499014U^2 \quad (7)$$

$$\text{clayey} = 0.848617 - 0.841851U + 0.275413U^2 \quad (8)$$

Figure 2. Soil types were further rearranged into three groups: sandy and gravelly soils (GP, SP, and SM), silty soils (SC and ML), and clayey soils (CL and CH). The original method gives constant probability of each soil type (represented by the step lines) regardless of the U value within the same region (R1 to R7 in The equations for the different soil curves are given as follows:

For U < 0.14: $\text{sandy} = 0.0035$, $\text{silty} = 0.0184$, $\text{clayey} = 0.9781$

For U = 2.91: $\text{sandy} = 0.9771, \text{silty} = 0.0229, \text{clayey} = 0.000$

For -0.14 < U < 2.91:

$$\text{sandy} = 0.00132408 + 0.074195\text{U} + 0.0900763\text{U}^2 \quad (6)$$

$$\text{silty} = 0.147853 + 0.896769\text{U} - 0.499014\text{U}^2 \quad (7)$$

$$\text{clayey} = 0.848617 - 0.841851\text{U} + 0.275413\text{U}^2 \quad (8)$$

Figure 2). This will allow for a sudden drop in the probabilities of U value across the border from one region to another. This method was further modified from origin to allow a smooth transition of probability (curved lines) with U values and hence to provide a continuous profile of the probability of soil constituents with depth [25].

The equations for the different soil curves are given as follows:

For U < 0.14: $\text{sandy} = 0.0035, \text{silty} = 0.0184, \text{clayey} = 0.9781$

For U = 2.91: $\text{sandy} = 0.9771, \text{silty} = 0.0229, \text{clayey} = 0.000$

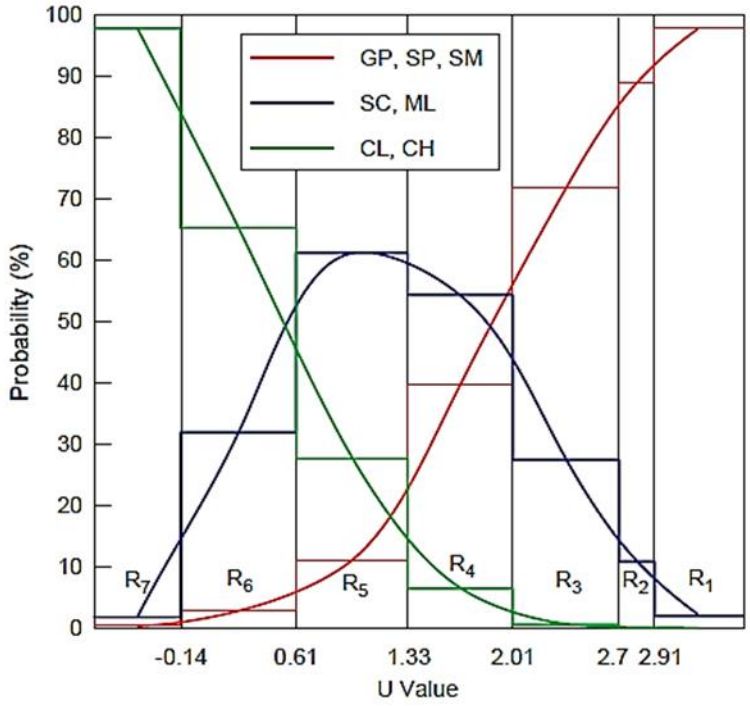
For -0.14 < U < 2.91:

$$\text{sandy} = 0.00132408 + 0.074195\text{U} + 0.0900763\text{U}^2 \quad (6)$$

$$\text{silty} = 0.147853 + 0.896769\text{U} - 0.499014\text{U}^2 \quad (7)$$

$$\text{clayey} = 0.848617 - 0.841851\text{U} + 0.275413\text{U}^2 \quad (8)$$

Figure 2. Regions' boundaries along the U-axis corresponding to probabilities of each soil group [25]



Robertson-2010 Soil Classification

Robertson [24] proposed a soil behavior type soil classification method that is presented in a chart of normalized cone tip resistance (q_c/P_a) versus friction ratio (R_f) space dividing the soil behavior into nine different soil behavior types (SBT). Here P_a is the atmospheric pressure. They used the soil behavior index that was proposed by Jefferies and Davies [26] and modified it to SBT index (I_{SBT}) as follows:

$$I_{SBT} = \sqrt{(3.47 - \log(q_c/P_{atm}))^2 + (1.22 + \log R_f)^2} \tag{9}$$

The SBT index has been used to divide the chart into the following 9 soil types:

1) Sensitive fine-grained
2) Clay-organic soil
3) Clays: clay to silty clay
4) Silt mixtures: clayey silt and silty clay
5) Sand mixtures: silty sand to sandy silt
6) Sand: clean sands to silty sands
7) Dense sand to gravelly sand
8) Stiff sand to clayey sand (overconsolidated or cemented)
9) Stiff fine-grained (overconsolidated or cemented)

The values of I_{SBT} for the soil type boundaries are shown in Figure 3.

In the chart, zones 1, 8, and 9 are defined as follows:

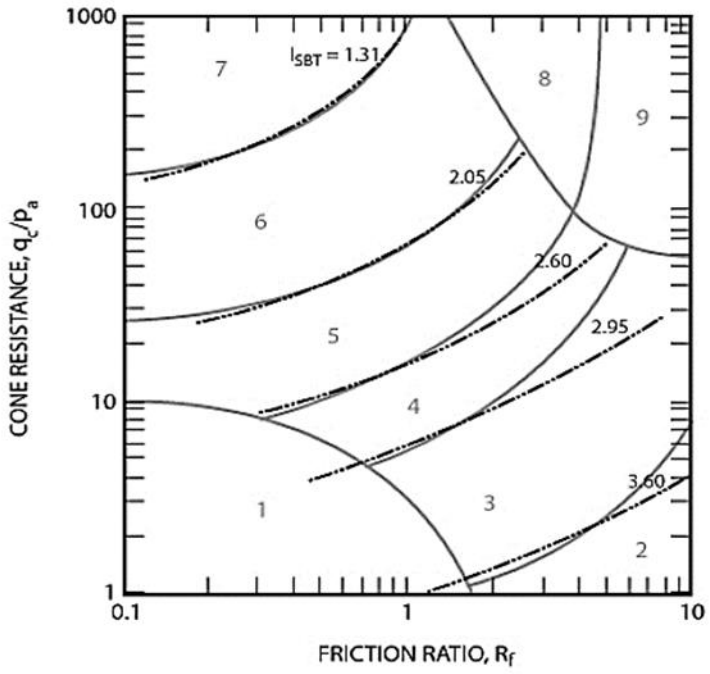
$$\text{Zone 1: } I_{z1} = q_c/P_a - 12\exp(-1.4R_f) < 0 \tag{10}$$

$$\text{Zone 8: } q_c/P_a - 5809.1\exp(-1.4R_f) > 56.86 \;\&\; R_f < 4.7 \tag{11}$$

$$\text{Zone 9: } q_c/P_a - 5809.1\exp(-1.4R_f) > 56.86 \;\&\; R_f > 4.7 \tag{12}$$

It should be noted here that no soil is located in zones 1, 8, and 9 in our study.

Figure 3. Robertson-2010: boundaries of I_{SBT} in the soil behavior type chart [24]



Direct Pile-CPT Methods

In this study, 21 direct pile-CPT methods were investigated. These methods are: Schmertmann [1], De Ruiter and Beringen [2], Bustamante and Gianeselli [3], Tumay and Fakhroo (cone-m) [4], Aoki and De Alencar [5], Price and Wardle [6], Philipponnat [7], Penpile [8], Probabilistic [10], NGI ([11], [27]), ICP [12], UWA ([28], [29]), CPT2000 [13], Fugro ([30], [31]), Purdue ([32], [33]), UF ([34], [35]), Togliani [36], and Zhou [37].

There are other pile-CPT methods available in literature that were not included for evaluation in this study. These methods are:

1) Methods using CPTu data: NGI-BRE (Almeida et al. [38] and Powel et al. [39]), Unicone [9], and Enhanced Unicone [40],
2) Methods limited to side resistance: KTRI [41],
3) Methods limited to clays: V-K [42], and
4) Methods for open-ended piles in sandy soils: UCD-05 [43], UCD-11 ([44], [45]), and HKU [46].

As stated earlier, the direct Pile-CPT methods evaluate the unit end bearing capacity of the pile (q_b) from the measured cone tip resistance (q_c) by averaging the cone tip resistance over an assumed influence zone. The unit shaft resistance (f) is either evaluated from the measured sleeve friction (f_s) in some methods or from the measured cone tip resistance (q_c) in other methods. It should be noted here that the cone tip resistance (q_c) is corrected for the pore water pressure and the probabilistic and Robertson-2010 CPT soil behavior classification methods were used to select the relevant equations and correlations for each soil type.

The detailed descriptions of the different direct pile-CPT methods are presented in Appendix A. Each pile-CPT method is introduced and the implementation of the probabilistic and Robertson-2010 CPT soil behavior classification methods into the different pile-CPT methods are explained.

Evaluation of Pile-CPT Methods

Several criteria have been used in literature to evaluate the direct pile-CPT methods for estimating the measured capacity from pile load tests. This section will summarize some of these evaluation criteria.

Briaud and Tucker [15] studied six direct pile-CPT methods using 98 pile load test database obtained from Mississippi State Highway Department. For statistical analyzing, the ratio of estimated to measured pile capacity (Q_p/Q_m) was investigated for different pile-CPT methods. The accuracy criteria of the method was determined by means of Q_p/Q_m close to 1. The precision criterion of the method was dependent on the standard deviation of Q_p/Q_m. For ranking the methods, they introduced a ranking index, RI, according to the following equation:

$$RI = |\mu(a)| + \sigma(a) \tag{13}$$

where, μ and σ are the mean and standard deviation of $\ln(Q_p/Q_m)$, respectively. They recommended using the log normal distribution. The method that overpredicts the pile capacity leads into lower values of RI, and therefore ranks better than the method that underpredicts the pile capacity. Based on their results, the LCPC, De Ruiter and Beringen, Penpile, Schmertmann, and Tumay and Fakhroo were set in order from the best to worst performance.

In 1999, a research study conducted in Louisiana Transportation Research Center (LTRC) by Abu-Farsakh and Titi [17] evaluated eight direct pile-CPT methods for estimating the ultimate pile capacity of 35 square precast prestressed concrete (PPC) driven friction piles by using the following four criteria:

1) The best fit line of estimated, Q_p versus measured pile capacity, Q_m with the corresponding coefficient of determination, R^2:
   The equation of best-fit line of estimated versus measured pile capacity with the corresponding coefficient of determination: linear regression is used to find a straight line between Q_m as the x values and Q_p as the y values. Forcing the regression line to pass through the origin leads to linear regression without the intercept term, $y = \beta x$, where the slope of best-fit line, β is found by the least-square approach in equation (14) as:

$$\beta = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} {x_i}^2} = \frac{\overline{xy}}{x^2} \tag{14}$$

   The coefficient of determination, R^2 is the proportion of the variance in the dependent variable, y from the independent variable, x. The following equation shows the most general definition of R^2:

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)^2} \tag{15}$$

   Where, $\hat{y}$ is the predicted values by the regression model and $\bar{y}$ is the mean of observed data (Q_p). R^2 ranges from 0 to 1 and shows how well Q_p values are replicated by the model. Accuracy and precision of a method can be estimated by having β and R^2 values close to 1, respectively.

2) The arithmetic mean and standard deviation of Q_p/Q_m:
   Mean and standard deviation are basic measures for accuracy and precision of CPT methods for predicting the pile capacity. Standard deviation should be understood in the

context of the mean of data. Coefficient of variation, CV is defined as the ratio of the standard deviation to mean and shows the extent of variation in relation to mean.

3) The 50% and 90% cumulative probability of Q_p/Q_m:
The concept is to arrange Q_p/Q_m values for each method in an ascending order and estimate the cumulative probability (P) using the following equation [47]:

$$P = \frac{i}{(n+1)} \tag{16}$$

The 50 and 90% cumulative probabilities are calculated as P_{50} and P_{90}, which provide an additional evaluation criteria to estimate the ability of CPT methods for predicting the axial capacity of piles. It should be noticed that P_{50} and P_{90} are representatives of median and 90 percentile of values of Q_p/Q_m, respectively. P_{50} values closer to 1 with a lower range of $P_{90} - P_{50}$ represent the best method.

4) The 20% accuracy level obtained from histogram and log normal distribution of Q_p/Q_m:
The value of Q_p/Q_m theoretically ranges between zero to infinity, with an ultimate value of 1. Therefore, log-normal distribution is better to catch the properties of Q_p/Q_m than normal distribution. The log-normal density is defined in the following equation:

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma_{ln}\,x}\exp\left[-\frac{1}{2}\left(\frac{\ln(x)-\mu_{ln}}{\sigma_{ln}}\right)^2\right] \tag{17}$$

Where, $x = Q_p/Q_m$, μ_{ln} and σ_{ln} are mean and standard deviation of $\ln(Q_p/Q_m)$, respectively. The histogram and log-normal distributions are used to calculate the ability of CPT methods to predict the pile capacity within a specified accuracy level. In their research, 20% accuracy has been chosen, which is the likelihood for Q_p values within 0.8 to $1.2Q_m$.

The ranking of each direct pile-CPT method was calculated in each criterion and summed up to determine the overall ranking index (RI = R1 + R2 + R3) of each method. Based on this ranking analysis, De Ruiter and Beringen, LCPC, Philipponnat, Schmertmann, Aoki and De Alencar, Price and Wardle, Tumay and Fakhroo, and Penpile methods showed the order of performance, respectively. Due to its rationality and simplicity, this evaluation approach has been adopted by other researchers to evaluate different direct pile-CPT method using different pile load test databases (e.g., [48]).

The load and resistance factor design (LRFD) is another approach that has been used for evaluating the different direct pile-CPT method. Bloomquist et al. [34] evaluated 14 direct Pile-CPT method using first-order second-moment (FOSM) resistance factor equation by Paikowsky [49] with correction for coefficient of variation of load by Styler [50] in LRFD equations. The values of bias parameter, $\lambda_R = R_m/R_n$ and resistance factor, $\phi = R_{design}/R_n$ define equation (18):

$$R_{design} = \left(\phi/\lambda_R\right) R_m \tag{18}$$

where, $\lambda_R = R_m/R_n$ is the bias parameter, R_m is the measured resistance referred to failure load defined by Davisson method from pile load test, R_n is the nominal resistance and R_{design} is the predicted design capacity be the method. The higher the value of efficiency,(ϕ/λ_R), the better the performance of the method is. Based on the analysis of 21 piles in Florida and 28 from Louisiana, LCPC and Philipponnat methods showed the best performance.

Another approach for evaluating different Pile-CPT methods is using fully-instrumented piles. Niazi and Mayne [40] used the 760-mm pipe pile driven in EURIPIDES project, instrumented by strain gauges, pore pressure cells, and toe load cells and found out that LCPC method underestimates the side resistance, while overestimates the tip resistance. Han et al. [51] studied the results of an instrumented closed-ended steel pipe pile driven in a multilayered soil and showed that the predictions by Purdue, ICP, UWA, NGI, and Fugro methods produce satisfactory estimates of the pile capacity; however, more field test data is needed for validation.

In this study, different evaluation approaches have been used for evaluating the ability of Pile-CPT methods for predicting the pile capacity.

Overview of Machine Learning (ML) Techniques

The application of artificial intelligence (AI) and machine learning (ML, a subset of AI) techniques by many industries have grown rapidly in recent years due to their powerful tools in predicting non-linear complex phenomenon and analyzing huge data sets. The AI and ML techniques usually use algorithms that function in an intelligent manner. They usually provide systems with the ability to learn and enhance from experience automatically without being specifically programmed. The interest of exploring the AI and ML in geotechnical engineering has been recently increased in civil and geotechnical engineering. The AI and ML techniques have recently demonstrated their high predictive ability to model complex civil and geotechnical

engineering problems (i.e., [52, 53]). The AI and ML techniques have an advantage over traditional regression modeling in terms of dealing with multiple outputs or responses while each regression model deals with only one response [54]. The AI and ML techniques involve algorithms that can accurately model complex mechanical behavior that include good generalization capability, universal function capability, resistance to noisy or missing data, and accommodation of multiple nonlinear variables for unknown interactions.

Several studies in literature have successfully applied the ML techniques for different geotechnical engineering applications, such as estimating different strength and deformation soil parameters, evaluating pile capacities, evaluating pile setup, and for liquefaction. Most of these applications use the artificial neural networks (ANN). Other researchers used the decision trees (DT), random forests (RF), gradient boosted tree (GBT), K-nearest neighbor (KNN) and vector machine (SVM) models for many geotechnical applications.

The soil-pile interaction is a complex phenomenon that requests advanced tools to model it. ML techniques involve algorithms that can capture complex non-linear relationships between variables. They can learn automatically from data and develop highly accurate generalized models without any prior simplifying assumptions about the relationships of interacting variables. Therefore, ML can be a promising alternative to better capture the soil-pile interaction and mitigate the assumptions and shortcomings of the direct pile-CPT methods. Several researchers have recently applied the ML techniques to estimate the ultimate pile capacity from CPT data. Shahin [55] applied the ANN technique, Kordjazi [56] used the SVM technique, and Alkroosh and Nikraz [57] utilized the gene expression programming (GEP) to predict the ultimate pile capacity from CPT data. Ghorbani et al. [58] explored the potential of adaptive neuro-fuzzy interface systems (ANFIS) in predicting the ultimate capacity of piles from CPT data. Harandizadeh et al. [59] developed a hybrid version of ANFIS, which is a combination of ANFIS and group method of data handling (GMDH) structure optimized by particle swarm optimization (PSO) algorithm called ANFIS-GMDH-PSO model. Ardalan et al. [60] built a prediction model for pile shaft resistance from CPT data using polynomial neural networks and genetic algorithm (GA). Baziar et al. [61] did the same using ANN. All these diverse ML methods have shown excellent performance in predicting the ultimate pile capacity that outperformed the conventional pile-CPT methods in most of those studies.

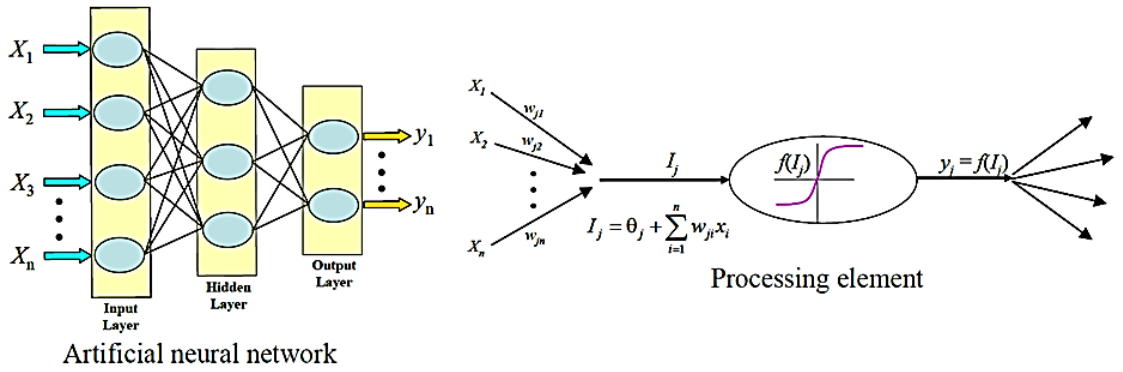
In this study, the ANN and three tree-based ML methods, the decision tree (DT), random forest (RF), and gradient boosted tree (GBT) were used to estimate the ultimate pile capacity from CPT data.

Artificial Neural Network (ANN)

The learning mechanism of the human brain, which is composed of very complex webs of interconnected neurons, is the primary inspiration towards the development of the artificial neural networks (ANNs). They intend to replicate the learning process of the human brain learning through mathematical algorithms using prior cases/instances. The ANNs can perform parallel computation for complex and massive data processing and knowledge representation.

Like the human brain, the primary element of ANN is neurons. They are also called nodes or processing elements. These processing elements are generally arranged in several layers consisting of an input layer (single layer), one or a few intermediate/hidden layers, and an output layer (single layer), as shown in Figure 4. The intermediate layers are also called hidden layers since they do not interact directly with the external environment. At least one neuron is present in each layer. The network is arranged in such a way that the output of one layer serves as the input for the following layer.

Figure 4. Typical structure ANN [53]



The neurons (or nodes) of each layer network is interconnected to other neurons through connection weights (Figure 4 b), which determine the strength of connections between the interconnected neurons. No connection between any two neurons should have a zero weight; whereas, a negative weight refers to a repressive relation. The received weighted inputs for an individual processing node are summed, aggregated, and scaled within a certain range to improve convergence property of ANN. The resultants are then propagated through a transfer function (e.g., step, linear, ramp, sigmoid logistic, or hyperbolic tangent) to generate the output of the processing node (Figure 4 b). The process, for any node j, is summarized using the following equations:

$$I_j = \theta_j + \sum_{i=1}^{n} w_{ji}\, x_i \quad (19)$$

$$y_j = f(I_j) \quad (20)$$

where, I_j^l = activation level of node j; w_{ji}^l = connection weight between nodes i and j; x_i^{l-1}= input from node i; i = 0, 1, …, n; θ_j^l=w_{j0} = bias for node j; y_j^l = output of node j; and $f(I_j)$ = transfer function. The hyperbolic tangent function (tanh) was used in this study, which is the hyperbolic analogue of the tan circular function. It is one of the most used functions for neural networks where the output ranges between -1 to +1. Ideally, tanh $(I_j) = (e^{I_j} - e^{-I_j}) / (e^{I_j} + e^{-I_j})$. The network is then propagated forward leading to final output, y_j. It is then compared with the target output, y_t and error, E, of the network is then calculated as, $E= ½ \sum(y_t - y_j)^2$.

The backpropagation algorithm is the prime algorithm used for training the ANN models [62]. The prime operation in backpropagation is searching for an error surface for point(s) with minimum error using a form of steepest descent. At each time step, the error gradient guides to a certain direction in the weight space, which reduces the local error drastically. The ANN backpropagation procedure can be described using the following steps [63]:

1) The input parameters are labeled as $x_1^0, x_1^0, x_3^0 \ldots . x_m^o$.
2) The connection weights can then be assigned as w_{ji}^l where l=0,1,2….l.
3) The forward network will then be propagated forward using equation 19 and 20:

   $$I_j^l = \theta_j^l + \sum_{n=1}^{i} w_{ji}^l x_i^{l-1}$$

   $$y_j^l = f(I_j)$$

   where f(.) is the activation function (e.g. logistic sigmoid).
4) For each j^{th} node belonged to output layer (l=l), calculate the correction factor δ:

   $$\delta_j^l = (y_t - y_j^l) y_j^l (1 - y_j^l) \quad (21)$$

5) Then update connection weights, w_{ji}^l, using the following equation:

   $$\Delta w_{ji}^{l\,(current)} = \eta \delta_j^l x_l^{l-1} + \mu \Delta w_{ji}^{l\,(previous)} \quad (22)$$

The above equation resembles the delta-rule ($\Delta w_{ji}^l = \eta \delta_j^l x_l^{l-1}$), where μ is the momentum rate ($0< \mu < 1$). This equation is also known as the generalized delta rule [62]. The update of bias can be done as follow:

$$\Delta\theta_{ji}^{l\,(current)} = \eta \delta_j^l + \mu \Delta\theta_{ji}^{l\,(previous)} \quad (23)$$

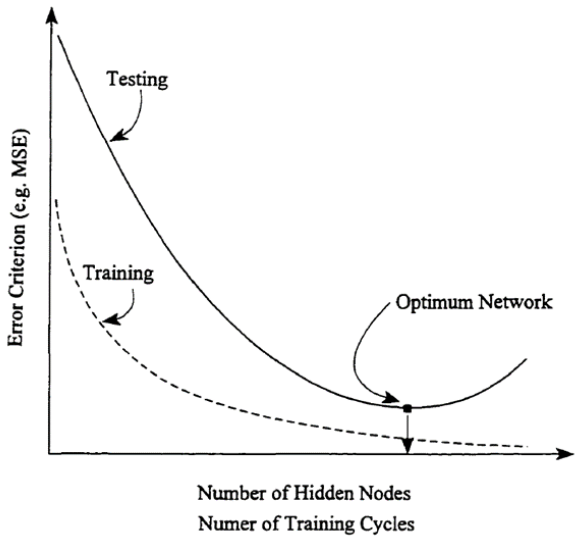
6) Similarly, for the case of hidden layers:

   $$\delta_j^l = y_j^l (1 - y_j^l) \sum_{k=1}^{r} \left(\frac{\partial E^l}{\partial y_k^{l-1}}\right) \left(\frac{\partial y_k^{l-1}}{\partial I_k^{l-1}}\right) \left(\frac{\partial I_k^{l-1}}{\partial y_{ji}^l}\right) \quad (24)$$

7) The weights and biases will be updated using equations 4 and 5, respectively.
8) Finally, the steps 1-7 are iterated until the output error is within acceptable tolerance.

The number of training cycles required for a better performance of the model is determined iteratively. A long training can result in overtraining or overfitting along with near-zero error on predicting training data. The generalization of test data degrades significantly in such situations (Figure 5). In the beginning, for a small number of training epochs, the error of the test-sets continues to decrease like the training examples. However, as the network loses its capability to generalize on test data, the error starts to increase after each epoch. The onset of an increase in the error of the test sets data resembles the optimum number of training cycles. When there are a limited number of training examples available, a sufficiently large test set is usually difficult to arrange. In such a case, Hecht-Nielsen [64] suggested the network to be trained on all available data and the training process is to be stopped when the error on training data is at the onset of stabilization.

Figure 5. Evolution of error for training and test data as a function of network size and number of training cycles [63]



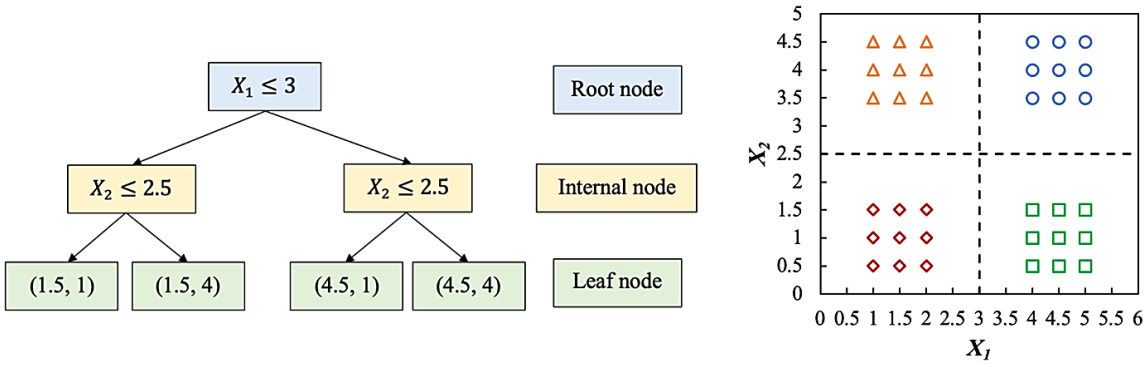
Tree-Based Machine Learning

The well-known decision tree (DT) is a classical, non-parametric supervised type of tree-based machine learning (ML) algorithm that can be used to solve non-linear problems. In general, using the DT algorithm alone can lead to a weak learner, which suffers from overfitting, i.e., inducing low bias and high variance. Hence, it can produce poor prediction accuracy. However, combining the DT algorithm with an ML ensemble technique usually results in significant improvements in the prediction accuracy while capturing highly non-linear complex relationships. In this study, two well-known techniques were explored along with the basic DT in which multitude of decision trees are either constructed in parallel (Random Forest) or constructed sequentially (Gradient Boosted Tree).

Decision Tree

The decision tree (DT), also known as regression tree, was proposed by Breiman et al. [65] to solve non-linear regression problems. The DT consists of three primary components: a root node representing the entire data set, internal nodes that split over each input feature, and several leaves or terminal nodes representing the outputs. Figure 6 presents the typical structure of a decision tree. The DT divides the input feature space into discrete, non-overlapping zones and predicts for each of them. For simplified illustration, consider Figure 6 with two input features, X_1 and X_2. For these two features, the dataset can be separated into four distinct areas or terminal regions. The four regions have the following mean values: (1.5, 1), (1.5, 4), (4.5, 4), and (4.5, 1). The predicted value of a new test sample is the average of the training observations in the region where the sample falls. In real process, more than two features are usually considered. The process of selecting the predictor space and division of that space, meaning tree's growth, occurs following a recursive binary splitting [66]. When a predetermined stopping criterion is met, the growth ceases. The minimum number of samples necessary to split an internal node can be predefined. The lowest number of samples required at a leaf node is an additional significant stopping criterion. These three values are essential hyperparameters for a DT model that will be fine-tuned later during the model-building process.

Figure 6. Structure of decision tree [66]

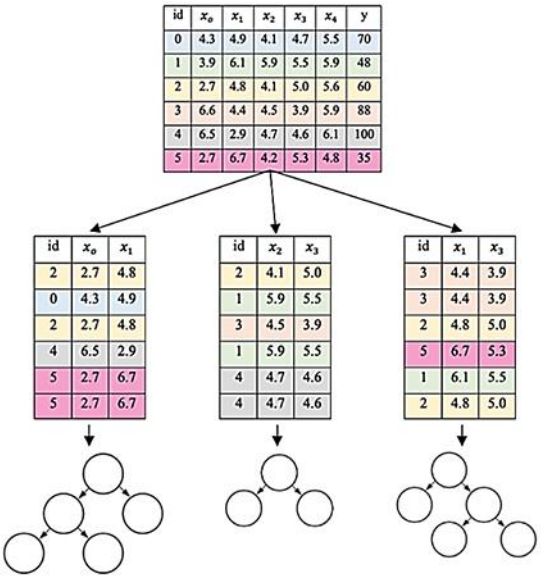


Random Forest

The Random Forest (RF) is an ML ensemble method developed by Breiman [67] to mitigate the limitations of individual DTs by building a certain number of them parallelly and introducing randomness to each weak tree. The RF generally follows two processes: bootstrap aggregation and selection of random feature. Using Figure 7 as a simplified example, the original dataset contains six entries of training examples from id_0 to id_5. Each training example has four features (x_0 to x_4) and one continuous output y. First, the RF algorithm creates N (N = 3 in this example) number of bootstrap datasets of the same size by randomly selecting training examples from the original dataset. This ensures each bootstrap dataset is independent of its peers, and it does not

depend on previously selected samples. Additionally, each bootstrap dataset is made with randomly selected features, as in the case of our example, the first bootstrap dataset comprises two (x_o, x_1) randomly selected features out of four. This random feature selection helps to reduce correlation between individual trees by reducing variance and tackle the overfitting problem that plagues individual DT model. Finally, each N number of bootstrap dataset is used to create N numbers of individual decision trees as explained, which ensures that the individual trees are as independent as possible. The RF model is then trained. While predicting using testing data samples, each input feature value is passed through its corresponding trees, and the independent output generated by each individual tree is averaged to get the final output from the RF model. The number of maximum features that randomly selected while bootstrapping is one crucial hyperparameter of the RF model. Another hyperparameter is the number of trees to build. In the RF model, each individual tree has its own hyperparameters. These hyperparameters are optimized for a given predictive modeling problem to get an optimum RF model.

Figure 7. Structure of Random Forest [67]



Gradient Boosted Tree

The Gradient Boosted Tree (GBT) proposed by Friedman [68] is another ML ensemble technique in which a sequence of weak DTs is constructed in an iterative fashion (sequentially). The GBT algorithm is described as follows:

Given the input training data, $\{(x_i, y_i)\}_{i=1}^{n}$, and a predefined differential loss function, $L(y_i, F(x))$:

1) Initialize a base model with a constant value:

$$F_0(x) = arg\,min_\gamma \sum_{i=1}^{n} L(y_i, \gamma) \quad (25)$$

where, y_i is the observed value and γ represents the predicted value, that minimizes the loss function $\sum_{i=1}^{n} L(y_i, \gamma)$

2) For m = 1 to M:

a. Compute the pseudo-residuals:

$$r_{im} = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)} \text{ for } i=1,\ldots\ldots,n \quad (26)$$

where, $F(x_i)$ is the previous model, and M is the total number of trees.

b. Fit a new regression tree to the r_{im} residuals and divide the input feature space into terminal regions R_{jm} for $j = 1,\ldots\ldots,J_m$. J_m represents the number of leaf nodes.

c. For $j = 1,\ldots\ldots, J_m$ compute:

$$\gamma_{jm} = arg\, min_{\gamma} \sum_{x_i \in R_{ij}} L(y_i, F_{m-1}(x_i) + \gamma) \quad (27)$$

d. Update $F_m(x) = F_{m-1}(x) + \eta \sum_{j=1}^{J_m} \gamma_m I(x \in R_{jm})$

where, η represents learning rate and $\sum_{j=1}^{J_m} \gamma_m I(x \in R_{jm})$ is the currently added tree.

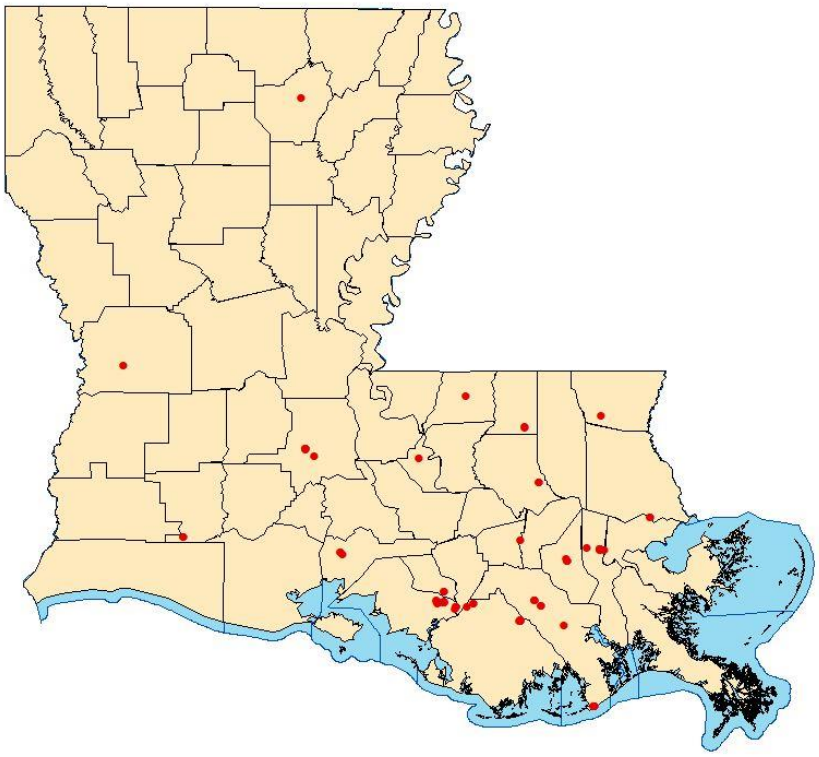
3) Output $F_M(x)$

In summary, the GBT algorithm adds new DTs sequentially to reduce the residual errors in prediction from the existing sequence of trees. This can rapidly reduce the error and eventually overfit the training data. Therefore, a weighting factor is usually applied for the corrections by new trees when added in the sequence, called the learning rate (η), which provides a regularization effect and enhances the training process. This learning rate is one of the significant hyperparameter of the GBT model, which will be fine-tuned later during the model development process.

Methodology

The main objective of this study was to evaluate the ability of different Pile-CPT methods for estimating the axial capacity of square PPC piles driven into Louisiana soils. For this reason, a database of 104 precast prestressed concrete (PPC) pile load test cases were collected from sites within state of Louisiana. Amongst those, only 80 PPC piles that were loaded to failure and their corresponding pile load tests, CPT tests, and soil borings data were collected. These piles are located in 34 project sites in Louisiana as shown in Figure 8.

Figure 8. Louisiana state map with location of analyzed piles



The collected pile load tests, soil properties and CPT data were analyzed. This section described the methodology of collecting, compiling, and analyzing the data.

Collection and Evaluation of Pile Load Test Reports

The information about the projects, soil stratifications, pile characteristics, load test data, CPT profiles, etc. were collected, processed, and transferred to different tables and graphs. All these data plus some available data about pile driving and dynamic test results were stored in digital format, so different analysis would be possible in future, regarding analyzing the reliability of different pile driving methods such as the Engineering News (EN) Formula, modified EN, Gates,

modified Gates, and Dynamic methods [Wave Equation Analysis (WEAP) and Case Pile Wave Analysis Program (CAPWAP)].

Pile load test reports were collected from available files at DOTD headquarters in Baton Rouge. These reports were studied carefully to examine their suitability to be included in this study. The main criteria for suitability of a project were availability of CPT soundings, site locations, and subsurface explorations. The characteristics of the square precast prestressed concrete piles obtained from DOTD are presented in Appendix B.

Compilation and Analysis of Pile Load Test Reports

The information about the projects, soil stratifications, pile characteristics, load test data, and CPT profiles were compiled. The graphs representing the summary of geotechnical data, including the soil stratigraphy, laboratory tests, and in-situ tests for the state projects are shown in Appendix C.

The following data and information were collected, compiled, and analyzed for each pile load test report.

Site Data

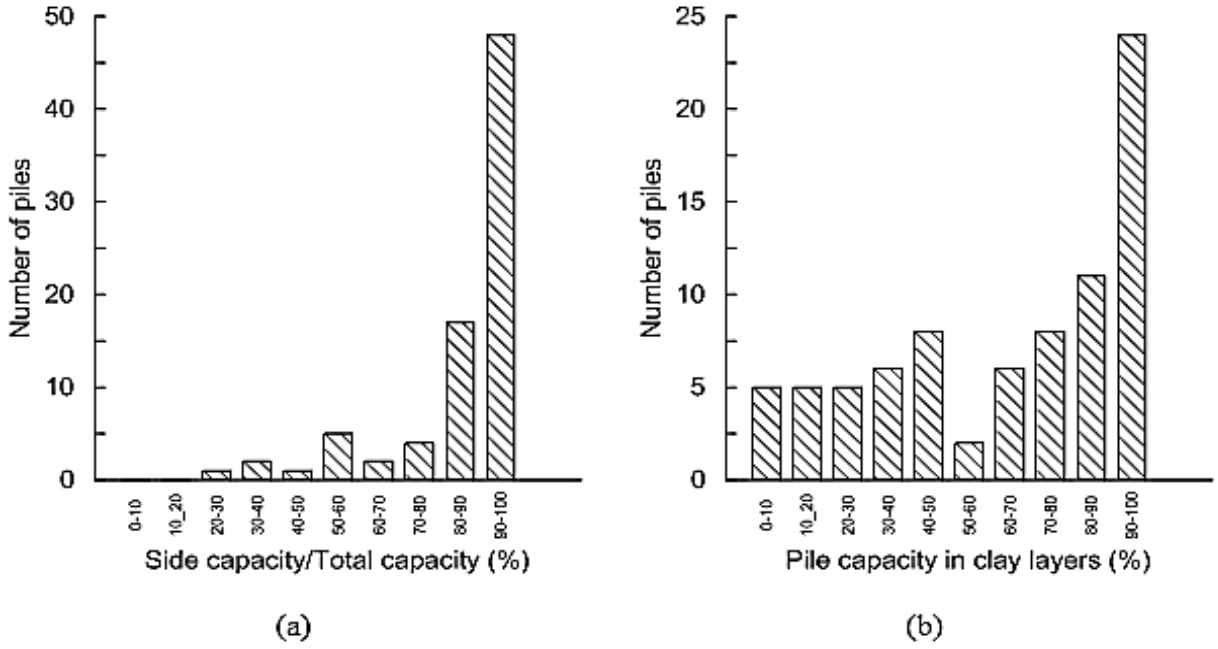
The site data provides the necessary information to identify the location of the project. The site identification used herein is the Louisiana state project number. For example, the site ID 260-05-0020 is the state project number 260-05-0020 (Tickfaw River Bridge and approaches on State Route LA-22). The project ID, location, and parish are available in Appendix B.

Soil Data

The soil data consist of information on the soil boring location, soil stratigraphy, and laboratory testing (shear strength, physical properties, etc.) for each soil layer. From soil stratification, the predominant soil type was identified. Appendix C shows the boring data for each pile studied in this project. Boring data near to the pile locations have been used in DRIVEN software (using the α-method and Nordlund method for clayey and sandy soils, respectively), which shows that most of the pile capacity driven in Louisiana soil is due to side resistance. It can be seen in Figure 9(a) shows that more than 70% of the pile capacity for 69 piles out of the 80 piles comes from the side resistance. Only four piles have a tip resistance that is more than 50% of the total pile capacity. This means that most of the piles in this study can be regarded as friction piles. The proportion of pile capacity in clay layers to the total pile capacity (defined as clay contribution) has been used to characterize the dominant soil for the pile database. As shown in Figure 9(b),

piles were driven into different sandy, clayey, and layered soils, as presented in percent in clayey soils.

Figure 9. Pile properties based on the soil type



Foundation Data

Foundation data consist of pile characteristics (pile ID, diameter, total length, embedded length), installation data (location of the pile, date of driving, driving record, hammer type, etc.) and pile load test (date of loading, applied load with time, pile head movement, pile failure under testing, etc.). All the piles studied in this report are driven square precast prestressed concrete piles. Appendix B represents information about the diameter, length, and embedment length of the piles, hammer type, and dates of driving and loading.

CPT Data

The cone penetration soundings information includes test location (station number), date, cone tip resistance, and sleeve friction profiles with depth. In most of the cases, the collected CPT soundings were not available as a digital data; therefore, the CPT soundings were scanned and digitized using the WebPlotDigitizer program. CPT graphs for each pile is shown in Appendix C.

Analysis of Ultimate Capacity of Piles from Load Test

Quick load test procedures as described in ASTM D1143 [69] were performed on the 80 different piles at approximately after 14 days of driving to obtain the load-settlement curve. Based on this procedure, the load was applied on the pile head in increments ranging from 10 to 15 percent of the design load and maintained for five minutes. The load was increased up to two to three times the design load or until pile failure. The load settlement curves for all the pile are shown in Appendix D.

The ultimate load capacity of the piles was determined based on the Davisson method [18]. Davisson failure criterion defines pile capacity as the load causes the pile top deflection equal to the calculated elastic compression plus 0.15 in. plus 1/120 of the pile's width/diameter. For piles with diameters more than 24 in., based on Florida Department of Transportation (FDOT) specification 2010, section 455 the criterion is modified to calculated elastic compression plus 1/30 of the pile's width/diameter [70].

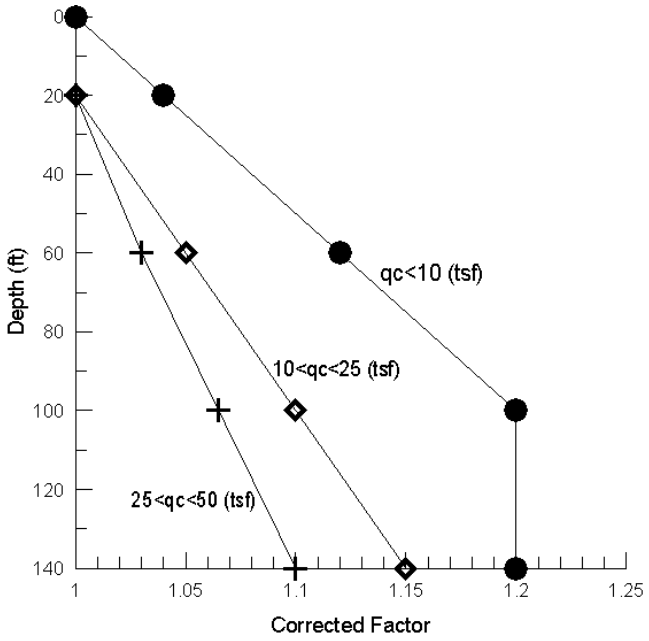
Correcting the Cone Tip Resistance

In order to improve the quality of CPT results, the cone tip resistance (q_c) should be corrected (q_t) due to the presence of porewater pressure acting behind the cone shoulder as follows ([71], [72]):

$$q_t = q_c + (1 - a)u_2 \quad (28)$$

where, a is the net area ratio for the cone (0.59 for CPT used in this research). In pile load test cases where the porewater pressures (u_2) are available, i.e., when piezocone PCPT (or CPTu) tests are used, the above equation was used directly to evaluate the corrected cone tip resistance (q_t). However, in many cases, only the CPT data (q_c, f_s with no u_2 measurement) using friction cone tests are available. To be able to correct the cone tip resistance when u_2 is not available, a database was collected from all available site locations in Louisiana with u_2 measurements. A comparison and statistical analysis was made between all collected q_c and q_t, which led to a correction factor that depends on the measured cone tip resistance (q_c) and depth, as shown in Figure 10. Details of collected database and correlating between the q_t and q_c are available in another study [73]. The correction factors obtained from Figure 10 were used to evaluate the q_t for the pile load test cases when the u_2 measurements are not available.

Figure 10. Correction factor for tip resistance with depth

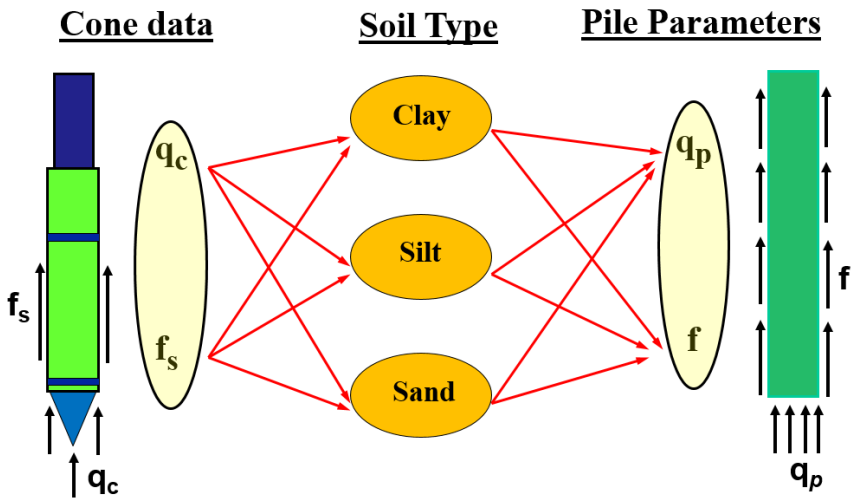


Incorporating CPT Soil Behavior Classification Methods

Almost all the pile-CPT methods require the determination of soil type and soil classification profile with depth along the pile in order to select the proper correlation parameters needed to evaluate the unit side capacity (f) for each soil layer and the unit end bearing capacity (q_p) of the pile from the measured CPT data (q_t, f_s). Figure 11 presents a simple schematic diagram of the procedure for estimating the piles′ unit capacities (q_p, f).

In this study, two CPT soil behavior type classification methods were selected: the probabilistic region estimation method [23] and the Robertson-2010 [24] classification method. Based on the CPT data, the probabilistic method determines the probability of soil behavior (clayey, sandy, and silty); while the Robertson-2010 presents a chart dividing the soil behavior into 9 different soil types. The details of each CPT soil classification method were explained earlier. The details of implementing the two CPT soil classification methods for selecting the correlation parameters for each pile-CPT method were discussed earlier, separately.

Figure 11. Effect of soil type on estimating pile capacity from CPT data



Development of Machine Learning (ML) Models

In this study, the artificial neural network (ANN) and three tree-based machine learning (ML) techniques [the decision tree (DT), random forest (RF), and gradient boosted tree (GBT)] were used to develop ML models to estimate the ultimate capacity of piles from CPT data. To develop any ML model, several parameters need to be identified and addressed. This include database compilation, selection of model inputs and outputs, data division and pre-processing, network architecture, hyperparameter optimization such as optimization of connection weights for ANN, training process, testing, stopping criteria, and validation of the ML models [55]. The collected 80 pile load test data were used in this study to train (calibrate), verify, and validate the ML models. The personal computer-based software Neural Designer was used in this work to simulate the ANN models; while the open source ML library for the Python programming language called scikit-learn or sklearn was used to simulate the three tree-based ML models [74].

Database Compilation

The database of 80 pile load tests of square PPC piles of varying widths and lengths that were collected from 34 project sites in Louisiana were used to develop the ML models. The lengths of piles range from 42 to 210 ft., while the pile widths range from 14 to 30 in. The load-settlement curve for each pile load test was interpreted to determine the measured ultimate pile capacity based on Davisson's offset limit method [18]. The associated CPT test data that were conducted close to each test pile were used to develop ML models based on CPT data.

Selection of Model Input Parameters

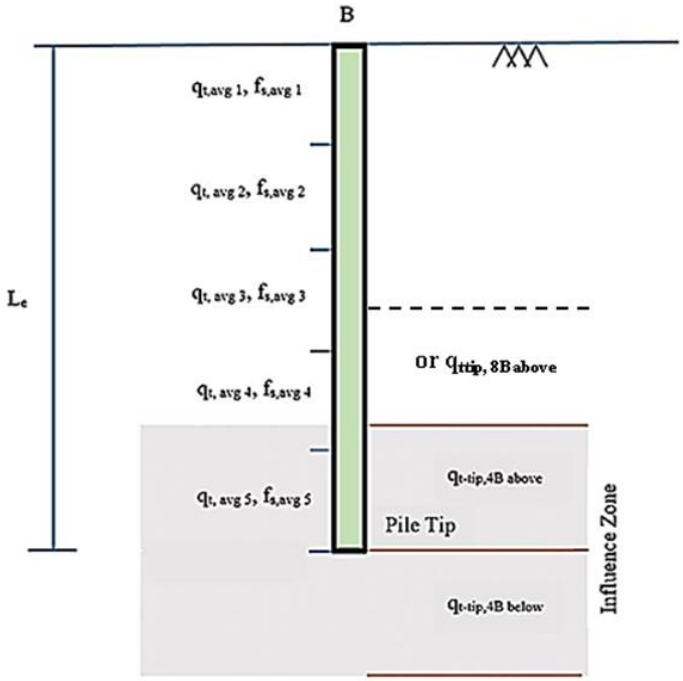
The quality and accuracy of any developed ML model depends on the proper selection of input variables that influence the output prediction of the ultimate pile capacity (Q_p). It is well-known that the ultimate capacity of piles depends on the pile characteristics such as pile material, geometry, and tip condition, pile installation method, testing procedure, and the soil properties. However, since all the tested piles in the compiled database are square PPC driven piles with closed end, many of these factors can be disregarded. In this study, the embedded length of pile (L_e) and pile width (B) along with the corresponding CPT data profile (corrected tip resistance, q_t, and sleeve friction, f_s) that represent the soil characteristics were considered as input parameters. The variation of soil properties along the pile shaft was sub-divided into five equal segments along the embedded pile length. The average values of corrected tip resistance ($q_{t,avg}$) and sleeve friction ($f_{s,avg}$) were calculated for each the five soil segments. The influence zone for the end bearing capacity was considered to range from 4B below the pile toe to 4B (or 8B for ANN only) above the pile toe, where B is the pile width. The corresponding averages of corrected tip resistance ($q_{t\text{-}tip}$) were determined for the upper and lower zones separately. Consequently, a total of 14 input parameters were used to develop the ML models as shown in Figure 12. Accordingly, the final selection of ML input parameters in this study were: (1) pile embedment depth (L_e), (2) pile width (B), (3) $q_{t,\ avg\ 1}$, (4) $q_{t,\ avg\ 2}$, (5) $q_{t,\ avg\ 3}$, (6) $qt_{,\ avg\ 4}$, (7) $q_{t,\ avg\ 5}$, (8) $f_{s,\ avg\ 1}$, (9) $f_{s,\ avg\ 2}$, (10) $f_{s,\ avg\ 3}$, (11) $f_{s,\ avg\ 4}$, (12) $f_{s,\ avg\ 5}$, (13) $q_{t\text{-}tip,\ 4B/8B\ above}$, and (14) $q_{t\text{-}tip,\ 4B\ below}$. These inputs parameters were used in three different combinations to determine the ML models that yield the best performance in terms of estimating the measured ultimate pile capacity. They only vary in the way the side resistance of piles were calibrated, i,e., using $q_{t,\ avg}$ alone, using $f_{s,\ avg}$ alone, or both $q_{t,\ avg}$ and $f_{s,\ avg}$. For each case, either $q_{t\text{-}tip,\ 4B\ above}$ or $q_{t\text{-}tip,\ 8B\ above}$ were used in ANN models, while only $q_{t\text{-}tip,\ 4B\ above}$ was used in tree-based ML models. Hence, a total of six different types of input parameter sets were considered in this study to obtain the best performed ANN models as shown in Table 1. Meanwhile only three input models were considered for the tree-based ML models.

Data Division and Pre-processing

Usually the available database is randomly divided into two subsets: a training set (to build the ML model) and another set for testing and validating the performance of ML model. Hammerstrom [75] suggested to consider two-thirds of the database for model training and the remaining one-third for testing and validation. Stone [76] proposed a modification of the above data division, which is known as cross-validation. In this technique, the data set is divided into three subsets: training, testing, and validation. The training set was used to improve the model network connection, the testing set to test the performance of ML model at different stages of

training, and validation set to determine the performance of the trained network. Shahin et al. [77] showed that there is no distinct relationship between the proportion of data for training, testing and validation, and the model performance. However, they obtained the best result using a combination of 70% of data for training and 30% of data for testing.

Figure 12. Selected input parameters and influence zone



In developing ANN modes, the database of 80 pile load tests was randomly divided into 70% for training, 10% for testing, and 20% for validation. However, in developing the tree-based ML Models, the database was randomly divided into 80% for training and 20% for testing. The training subset was used to train the ML models, while the testing subset was used to evaluate the accuracy and generalization ability of the trained ML models. It should be noted here that the ML models have difficulty in extrapolating beyond the range of the range of the training data. Therefore, all the existing patterns available in the dataset need to be included in the training set to develop a good ML model. If the extreme data points are excluded from the training dataset, then the validation data will test the models' extrapolation capability instead of its interpolation ability. Consequently, the model may not perform well as the ML models perform best when they do not extrapolate beyond the limit of the training data [24]. Therefore, care was taken to avoid this scenario. In this study, data division was carried out randomly through trial and error until the statistical properties (mean, standard deviation, range) of subsets are close to each other as possible with the minimum and maximum values included in the training subset [77].

Table 1. ANN model types used in this study

ANN model type	Input parameters
Type 1	(1) Pile embedment depth, L, (2) Pile width, D, (3) $q_{t,\ avg\ 1}$, (4) $q_{t,\ avg\ 2}$, (5) $q_{t,\ avg\ 3}$, (6) $q_{t,\ avg\ 4}$, (7) $q_{t,\ avg\ 5}$, (8) $q_{t\text{-}tip,\ 4B\ above}$, (9) $q_{t\text{-}tip,\ 4B\ below}$
Type 2	(1) Pile embedment depth, L, (2) Pile width, D, (3) $q_{t,\ avg\ 1}$, (4) $q_{t,\ avg\ 2}$, (5) $q_{t,\ avg\ 3}$, (6) $q_{t,\ avg\ 4}$, (7) $q_{t,\ avg\ 5}$, (8) $q_{t\text{-}tip,\ 8B\ above}$, (9) $q_{t\text{-}tip,\ 4B\ below}$
Type 3	(1) Pile embedment depth, L, (2) Pile width, D (3) $f_{s,\ avg\ 1}$, (4) $f_{s,\ avg\ 2}$, (5) $f_{s,\ avg\ 3}$, (6) $f_{s,\ avg\ 4}$, (7) $f_{s,\ avg\ 5}$, (8) $q_{t\text{-}tip,\ 4B\ above}$, (9) $q_{t\text{-}tip,\ 4B\ below}$
Type 4	(1) Pile embedment depth, L, (2) Pile width, D (3) $f_{s,\ avg\ 1}$, (4) $f_{s,\ avg\ 2}$, (5) $f_{s,\ avg\ 3}$, (6) $f_{s,\ avg\ 4}$, (7) $f_{s,\ avg\ 5}$, (8) $q_{t\text{-}tip,\ 8B\ above}$, (9) $q_{t\text{-}tip,\ 4B\ below}$
Type 5	(1) Pile embedment depth, L, (2) Pile width, D, (3) $q_{t,\ avg\ 1}$, (4) $q_{t,\ avg\ 2}$, (5) $q_{t,\ avg\ 3}$, (6) $q_{t,\ avg\ 4}$, (7) $q_{t,\ avg\ 5}$, (8) $f_{s,\ avg\ 1}$, (9) $f_{s,\ avg\ 2}$, (10) $f_{s,\ avg\ 3}$, (11) $f_{s,\ avg\ 4}$, (12) $f_{s,\ avg\ 5}$, (13) $q_{t\text{-}tip,\ 4B\ above}$, (14) $q_{t\text{-}tip,\ 4B\ below}$
Type 6	(1) Pile embedment depth, L, (2) Pile width, D, (3) $q_{t,\ avg\ 1}$, (4) $q_{t,\ avg\ 2}$, (5) $q_{t,\ avg\ 3}$, (6) $q_{t,\ avg\ 4}$, (7) $q_{t,\ avg\ 5}$, (8) $f_{s,\ avg\ 1}$, (9) $f_{s,\ avg\ 2}$, (10) $f_{s,\ avg\ 3}$, (11) $f_{s,\ avg\ 4}$, (12) $f_{s,\ avg\ 5}$, (13) $q_{t\text{-}tip,\ 8B\ above}$, (14) $q_{t\text{-}tip,\ 4B\ below}$

Training of ML Models

Training of an ANN model refers to the process of initializing a network through the deployment of initial values and then optimizing the connection weights in order to obtain a global minimum instead of a local one. A very widely used method to obtain the optimum weights is the back-propagation algorithm or the gradient descent method. However, the convergence is sometimes slower and requires lots of iterations in this method. Hence, a faster Quasi-Newton method was also used in this work to get the optimum ANN. The number of training cycles required for a better performance of the model is determined iteratively. A maximum of 1000 iterations was allowed in the Neural Designer software.

The tree-based ML models (DT, RF, and GBT) were trained and assessed independently to identify the model that provides the optimum or near optimum performance. Each model possesses some external tunable parameters called hyperparameters. These parameters regulate the learning process and must be set before the training process begins. Table 2 illustrates the significant hyperparameters of DT, RF, and GBT models tuned, as denoted in scikit-learn [74].

Table 2. Hyperparameters of tree-based models

Hyperparameters	Variable type	Range	Applicable models
learning_rate	Continuous	0 - ∞	GBT
n_estimators	Discrete	1 - ∞	RF, GBT
max_features	Continuous	0 - 1	RF, GBT
max_depth	Discrete	1 - ∞	DT, RF, GBT
min_samples_split	Discrete	2 - ∞	DT, RF, GBT
min_samples_leaf	Discrete	1 - ∞	DT, RF, GBT

Since both the RF and GBT models are built with DTs as base learners, they have some common hyperparameters (max_depth, min_samples_split, min_samples_leaf). In addition, the RF and GBT have two crucial hyperparameters to be tuned: the number of decision trees to be combined, denoted as n_estimators, and the number of random features to consider while building trees, denoted as max_features. The GBT model has an additional hyperparameter called the learning_rate. It should be noted here that the three tree-based models have more additional hyperparameters. Only the main hyperparameters that significantly impact the performance of ML models based on literature are explored in this study.

It is necessary to evaluate multiple combinations of hyperparameter settings since one set may perform well in one case but poorly in another case. The process of determining the optimal combination of hyperparameter settings for a certain problem is known as hyperparameter optimization. In previous studies, the produced ML models were optimized manually through a trial-and-error procedure that is tedious and computationally expensive. In this study, a more effective method termed random search is employed for hyperparameter optimization. The random search procedure begins by specifying a finite number of possible values for each hyperparameter, creating a hyperparameter search space. Then, the search algorithm selects random combinations of hyperparameters from the search space. Each hyperparameter combination represents a distinct candidate model. The performance of each candidate model is then determined using cross-validation procedure.

Stopping Criteria

Stopping criteria is important to determine when to stop the training process. There several approaches that can be used to decide when to stop training. Training can be stopped when a fixed number of training records are presented, when sufficiently small value of the training error is obtained or when changes in the training error is insignificant. However, these approaches may lead to premature model stopping or over-training. In this study, the cross-validation method was implemented to solve this issue. The testing set judges the capability of the model to be

generalized, through evaluating the performance of the model at different stages of the training process. Once the testing process is completed, we move to next validation set to assess the model performance.

Validation of ML Models

After completing the training phase and locating the optimal ML model, the model needs to be validated to ensure its ability to be generalized in a robust way within the limits of training data. A separate data set that was not utilized in the training phase is usually used to validate the ML model in terms of accurately predicting the measured ultimate pile capacity, Q_m. The satisfactory performance in this phase indicates the model's robustness. At this perspective, the coefficient of determination, R^2, the root mean squared error, RMSE, mean bias factor, λ, and the coefficient of variance, COV, of the measured over predicted ultimate pile capacity (Q_p/Q_m), are the prime criteria that were used to evaluate the performance of ML models. The coefficient of correlation, r, can also be used to obtain the relative correlation and goodness of fit between the measured and predicted values. Smith [78]suggested the following guide for |r| values:

|r| ≥ 0.8 strong correlation exists between two sets of variables;

0.2 < |r| < 0.8 correlation exists between the two sets of variables; and

|r| ≤ 0.2 weak correlation exists between the two sets of variables.

However, the RMSE is considered the most popular measure of error due to its advantage of giving greater attenuation towards large errors rather than the smaller ones. The parameters can be calculated as follow:

$$r = \frac{n(\sum Q_p Q_m) - (\sum Q_p)(\sum Q_m)}{\sqrt{[n(\sum Q_p{}^2) - (\sum Q_p)^2][n(\sum Q_m{}^2) - (\sum Q_m)^2]}} \quad (29)$$

$$RMSE = \sqrt[2]{\frac{1}{n}\sum_{i=1}^{n}(Q_p - Q_m)^2} \quad (30)$$

$$\lambda = \frac{Q_m}{Q_p} \quad (31)$$

$$COV\ (\%) = \left(\frac{Standard\ Deviation}{\bar{\lambda}}\right) \times 100 \quad (32)$$

Where, n = number of samples or observations, Q_p = predicted pile capacity, and Q_m = measured pile capacity.

Analysis of Results

The main objective of this study was to evaluate the ability of different direct pile-CPT methods for estimating the ultimate load carrying capacity of square PPC piles driven in Louisiana soils. A total of 80 pile load test database were used in this study to evaluate the performance of 21 direct pile-CPT design methods. The measured ultimate capacity for each pile was determined from the load deformation curve based on Davisson interpretation criteria [18]. In addition, the predicted ultimate pile capacity for each pile was also determined for each the 21 pile-CPT method. The comparison between the measured and estimated pile capacities for all pile-CPT methods are available in Appendix E.

In our analysis, the following tasks have been executed and discussed in this section:

1) Utilize sensitivity analysis of the pile-CPT methods to selected CPT soil classification method;
2) Evaluate pile-CPT methods based on mathematical and statistical criteria;
3) Evaluate pile-CPT methods using MultiDimensional Unfolding;
4) Evaluate pile-CPT methods based on efficiency from reliability analysis;
5) Develop combined methods from the best performed pile-CPT methods for different soil conditions based on contribution of sand layers to total ultimate capacity (kind of optimizing the prediction accuracy); and
6) Develop machine (ML) learning models using the artificial neural network (ANN) and three tree-based ML methods, the decision tree (DT), random forest (RF), and gradient boosted tree (GBT) to estimate the ultimate pile capacity from CPT data.

Sensitivity of Pile-CPT Methods to Selected CPT Soil Classification Method

Seventeen pile-CPT methods including: LCPC, Schmertmann, De Ruiter, Philipponnat, UF, probabilistic, Aoki, Penpile, NGI, ICP, UWA, CPT2000, Fugro, Purdue, German, Eurocode7, and ERTC3 are dependent on the soil type for estimating the ultimate pile capacity. This means that in order to use CPT for calculating the ultimate pile capacity, it is necessary to classify the soil and evaluate soil type with depth for proper selection of correlation parameters between the CPT data (q_t, f_s) and pile unit side and unit end bearing capacities (q_b, f).

For example, Figure 13 compares the predicted ultimate capacity profiles, $\mathbf{Q_p}$, for a pile in Gibson-Raceland highway site obtained from the UWA pile-CPT method using both the

probabilistic region estimation and Robertson-2010 CPT soil classification methods. It can be seen that the values of $\mathbf{Q_p/Q_m}$ for the UWA pile-CPT method are not much different for probabilistic and Robertson (2010) soil classification methods. However, in some cases there is a significant difference.

As a part of this research study, the sensitivity of the 21 pile-CPT methods to the selection of CPT soil classification methods were analyzed. As presented in Appendix E, for each pile-CPT method, the values of predicted pile capacities, Q_p, were calculated using the probabilistic region estimation and Robertson (2010) CPT soil classifications, separately.

In order to quantify the difference between pile capacity predictions using the two CPT soil classifications, the value of diff (%) is defined in Equation (33), which represents the percentage of increase in Q_p/Q_m in case of using probabilistic soil classification as compared to Robertson-2010 soil classification.

$$\text{diff (\%)} = \left[\left(Q_p/Q_m\right)_{\text{probabilistic}} - \left(Q_p/Q_m\right)_{\text{Robertson}}\right] \times 100 \tag{33}$$

Statistical analysis (using SAS/STATTM software) was used to test the null hypothesis of diff (%) equal to zero for different methods. The null hypothesis was rejected in all pile-CPT methods except for the LCPC, Schmertmann, and Aoki methods, which means that the selection of CPT soil classification has a significant result on the ability of the methods for estimating the ultimate pile capacity. The statistical results for diff (%) of the 17 pile-CPT methods are shown in Figure 14 and described in Table 3.

Figure 13. CPT data and boring log for tp4 at Gibson-Raceland highway project in Terrebonne parish (pile 42) (a) profile of cone tip resistance & friction ratio (b) soil classification from boring (c) CPT soil classification using Zhang and Tumay (1999) (d) soil classification using Robertson (2010) (e) estimated pile capacity from UWA direct pile-CPT method using different soil classifications.

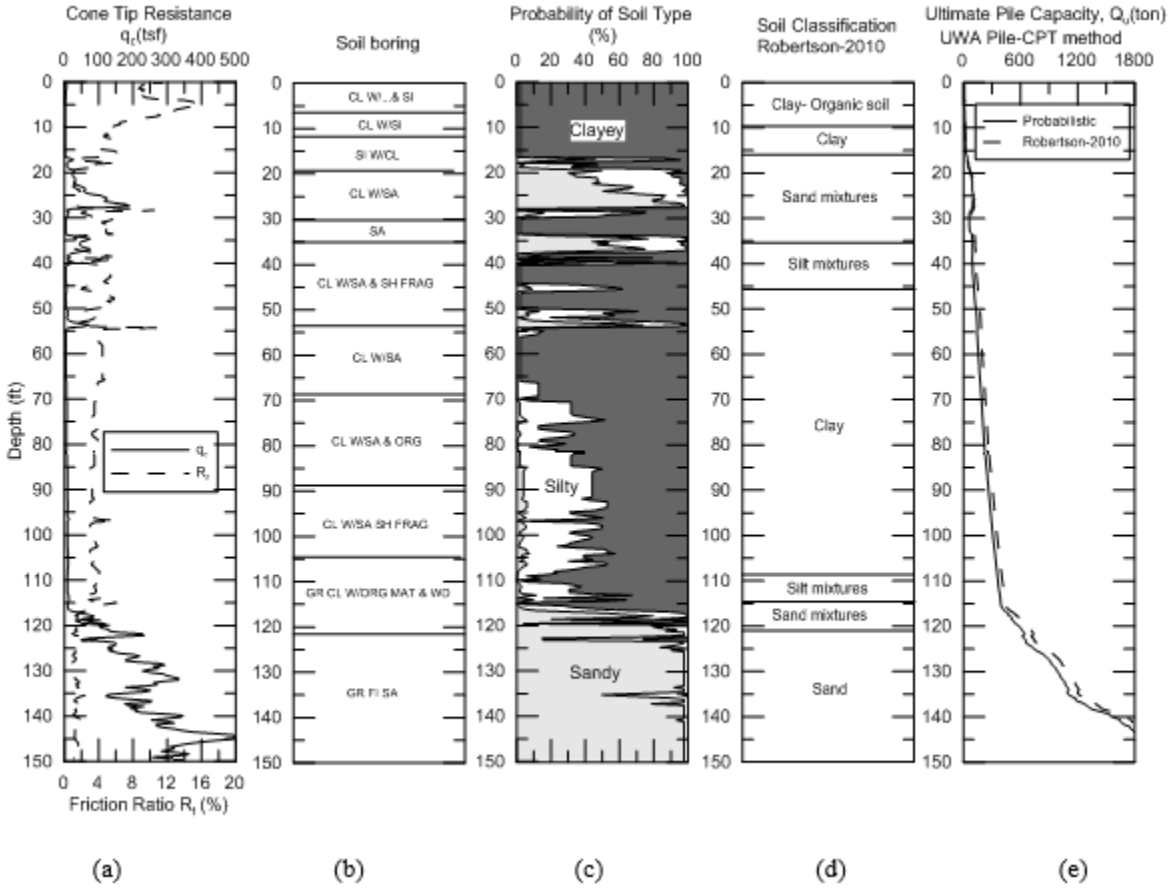


Figure 14. Box plots of the diff (%) in pile capacities for all pile-CPT methods between using either of the two CPT methods

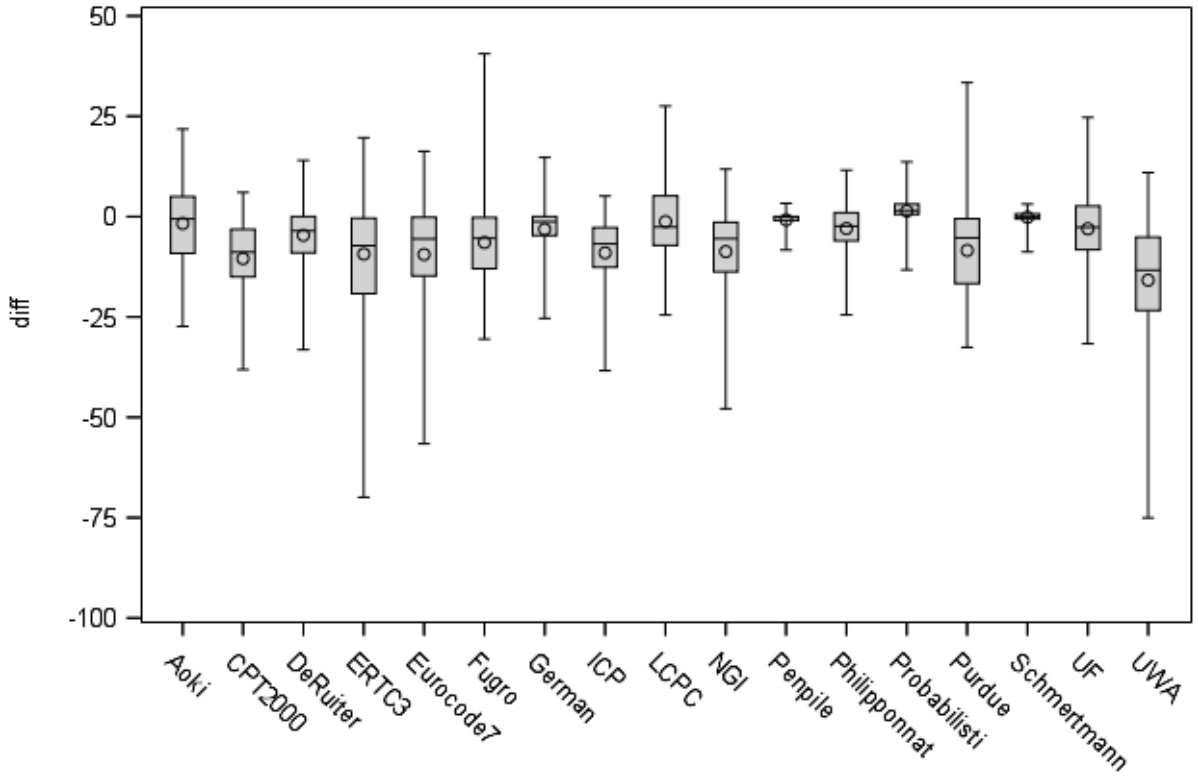


Table 3. Mean, standard deviation, max and min values of diff (%) for different pile-CPT methods

CPT methods		LCPC	Schmertmann	De Ruiter	Philipponnat	UF	Probabilistic	Aoki	Penpile	NGI
diff (%)	mean	-1.24	-0.18	-4.63	-2.97	-3.02	1.38	-1.69	-0.82	-8.73
	SD	9.03	1.91	7.68	6.21	10.00	3.52	9.31	1.74	10.09
	max	27.54	3.12	14.01	11.57	24.72	13.65	21.81	3.33	11.83
	min	-24.50	-8.78	-33.17	-24.49	-31.68	-13.27	-27.35	-8.32	-47.88

CPT methods		ICP	UWA	CPT2000	Fugro	Purdue	Eurocode7	ERTC3	German
diff (%)	mean	-9.03	-15.85	-10.47	-6.44	-8.43	-9.44	-9.32	-3.16
	SD	9.04	15.44	10.57	11.31	11.59	13.58	14.15	6.65
	max	5.12	10.97	6.03	40.57	33.41	16.29	19.65	14.75
	min	-38.39	-75.05	-38.12	-30.52	-32.56	-56.60	-69.92	-25.35

The lowest mean values of diff (%) are for Schmertmann and Penpile methods, which shows that on average these methods are less dependent on the selection of soil classification method. On the other hand, UWA, Eurocode7, ERTC3, CPT2000, ICP, NGI, and Purdue methods have the highest mean values for the diff (%), which implies that these methods show the significance of the selection of soil classification on the difference on predicted pile capacities. Analysis of standard deviations for the 21 pile-CPT methods shows that Penpile, Schmertmann, and probabilistic methods have the lowest values of standard deviation, which implies that these methods are less sensitive to the selection of soil classification method. On the other hand, UWA, ERTcC3, Eurocode7, Purdue, Fugro, NGI, and CPT2000 methods have higher values of

standard deviation, and therefore their sensitivity to the selection of soil classification method is higher. The max and min (extreme) values of diff (%) in Table 3 represent the range of diff (%). The lowest range is for Penpile, probabilistic, and Schmertmann methods, which is within ±13%. The UWA, ERTC3, and Eurocode7 methods have the highest range of -75% (UWA) to +20% (ERTC3), which means that using the probabilistic region soil classification might estimate 75% less or 20% higher ultimate pile capacities than using Robertson-2010 soil classification. The range of diff (%) for the other pile-CPT methods are between -48% (NGI) to +30% (Fugro).

It should be noted here that the value of diff (%) in predicted ultimate pile capacities also depends on the engineering judgment in implementing the soil classification methods for evaluating the different correlation parameters between CPT data (q_t, f_s) and pile data (q_b, f) for each pile-CPT method. For the purpose of evaluating the pile-CPT methods, the average values of predicted the ultimate pile capacity, Q_p, obtained from using either the probabilistic region estimation or the Robertson-2010 CPT soil classification methods was adopted in this study.

Evaluation of Pile-CPT Methods Based on Mathematical and Statistical Analyses

In this part of study, an evaluation scheme using three different criteria based on mathematical and statistical analysis was considered in order to rank the performance of the different pile-CPT methods for estimating the ultimate axial capacity of driven piles in Louisiana soils. These criteria are: (1) the equation of the best fit line of predicted versus measured capacity Q_p/Q_m and the corresponding coefficient of determinations, R^2; (2) the arithmetic mean and standard deviation for Q_p/Q_m; and (3) Q_p/Q_m at 50% and 90% cumulative probability (P50 and P90). Another criterion reported in a previous study by Titi and Abu-Farsakh [79] was based on the 20% accuracy level for Q_p/Q_m obtained from histogram and log-normal distribution. This criterion seems to represent information about the accuracy and precision of the methods as the others. For this reason, it was decided not to include this criterion in our evaluation analysis. It should be noticed that the log-normal distribution of Q_p/Q_m will be addressed later in the reliability analysis and evaluation for calibrating the LRFD resistance factor, ϕ.

The plot of predicted (Q_p) versus measured (Q_m) ultimate capacity and the cumulative probability plots for all pile-CPT methods are presented in Appendix F. A rank index (RI) was used in this part of study to quantify the overall performance of all methods. The rank index is the sum of the ranks from the different criteria, RI= R1+R2+R3. The lower the rank index RI,

the better the performance of the method. The performance of the prediction methods based on the three different criteria is discussed below.

Inspecting the results of Q_p versus Q_m plots in Appendix F shows that the LCPC, ERTC3, UF, and Philipponnat methods have the best fit equation $Q_{fit} = 1.03Q_m$ with R^2=0.74~0.82. The results also summarized in Table 4. These methods tend to overpredict the measured pile capacity by an average of 3%. Therefore, these methods are ranked number one according to this criterion and is given *R1 = 1* (*R1* is the rank based on this criterion). The Probabilistic and De Ruiter and Beringen methods with $Q_{fit} = (0.97 - 0.98)Q_m$ (R^2=0.77~0.78) tend to underpredict the measured capacity by 2-3% and therefore are ranked next (*R1=4*). Also, the German method has $Q_{fit} = 1.03Q_m$, but R^2 = 0.67; which is low and therefore it is ranked as *R1=4*. According to this criterion, Aoki and De Alencar, Price and Wardle, and Penpile methods tend to underpredict the measured ultimate pile capacity; while the other methods tend to overpredict the measured ultimate pile capacity. The Togliani method showed the worse performance with $Q_{fit} = 1.70Q_m$ (R^2 = 0.81) and therefore was given *R1* = 9.

In the second criterion, the arithmetic mean (μ) and standard deviation (σ) of the ratio Q_p/Q_m values for each method were calculated. The best method is the one that gives a mean value closer to one with a lower coefficient of variation (COV), which is the measure of scatter in the data around the mean. According to this criterion, the Probabilistic and UF methods rank number one (*R2* = 1) with $\mu(Q_p/Q_m) = 1.03$ and 1.04 and *COV* = 0.33 and 0.35, respectively. They are followed by the LCPC, ERTC3, Philipponnat, and De Ruiter and Beringen methods (*R2* = 3). De Ruiter and Beringen, Aoki and De Alencar, Price and Wardle, and Penpile have $\mu(Q_p/Q_m) < 1$, which means that these methods on average are underpredicting the measured pile capacity. On the other hand, other methods have $\mu(Q_p/Q_m) < 1$, which means that these methods on average are overpredicting the measured pile capacity.

The cumulative probability curves (Appendix F) were used to determine the 50 percent and 90 percent cumulative probability values (P_{50} and P_{90}), which are also summarized in Table 4. The pile capacity prediction method with P_{50} value closer to one and with lower P_{50} - P_{90} range is considered the best. Based on this criterion, the ERTC3, LCPC and Probabilistic methods with $P_{50} \approx 1.0$ and P_{90} = 1.41~1.45 rank number one (*R3*=1) followed by UF, Philipponnat, CPT2000, and De Ruiter and Beringen with *R3*=4. The Togliani method has worst P_{50} and P_{90} values and therefore ranks as the worst method.

In order to evaluate the overall performance of the different prediction methods, all criteria were considered in a form of an index. The Rank Index (*RI*) is the algebraic sum of the ranks obtained using the three criteria. Considering LCPC method, the *RI* equals to five as evaluated from

RI=R1+R2+R3. The Rank Index values for all other methods are presented in Table 4. Inspection of Table 4 demonstrates that Bustamante and Gianeselli (LCPC/LCP) method ranks number one along with ERTC3 method. These two methods showed the best performance according to the evaluation criteria and therefore considered the best methods. The Probabilistic and UF methods rank number three, followed by Philipponnat, De Ruiter and Beringen, CPT2000, UWA, and Schmertmann methods. The Zhou and Togliani methods showed the worst performance among all methods.

Evaluation of Pile-CPT Methods Using MultiDimensional Unfolding

The MultiDimensional Unfolding (MDU) is an approach used in this study, which displays the ranking data in a two-dimensional space. This approach helps us to find out the typical ranking of the pile-CPT methods, the extent of agreement between the piles, existence of outliers among the piles and pile-CPT methods, and the similarity between the different pile-CPT methods.

In this section, MultiDimensional Scaling (MDS) is described. Then, some examples are explained to be solved using MDS. Then, MultiDimensional Unfolding (MDU) as an MDS technique is described. Finally, MDU is used for ranking Pile-CPT methods.

MultiDimensional Scaling (MDS): Basics

The MultiDimensional scaling (MDS) is a technique for showing similarity between the objects in a low-dimensional space. A symmetric $n \times n$ matrix, Δ known as dissimilarity matrix with elements δ_{ij} is the input matrix for MDS. A very simple example for application of MDS is that distances between some cities in the US is given as the input matrix and the result is a two-dimensional locations of these points reflecting the US map [80]. Finding the location of the points in an Euclidean space of dimension p, without any additional transformation is the traditional way of performing MDS, referred as classical scaling [81].

The objective of performing MDS is finding the configuration matrix, X somehow that distances between points, d_{ij} be as close as possible to values of δ_{ij}. The values of d_{ij} is defined in a p-dimensional space as follows:

$$d_{ij}(X) = \sqrt{\sum_{s=1}^{p} \left(x_{is} - x_{js}\right)^2} \tag{34}$$

Table 4. Ranking of pile-CPT methods based on multiple criteria

	Best fit calculation			Arithmetic calculations Qp/Qm			Cumulative probability			Overall rank	
Pile capacity method	Q_{fit}/Q_m	R^2	R1	Mean	COV	R2	Q_p/Q_m at P_{50}	Q_p/Q_m at P_{90}	R3	RI	Final rank
LCPC	**1.03**	**0.74**	**1**	**1.07**	**0.39**	**3**	**0.99**	**1.45**	**1**	**5**	**1**
ERTC3	**1.04**	**0.73**	**1**	**1.08**	**0.35**	**3**	**1.01**	**1.41**	**1**	**5**	**1**
Probabilistic	**0.97**	**0.78**	**4**	**1.03**	**0.33**	**1**	**0.99**	**1.42**	**1**	**6**	**3**
UF	**1.03**	**0.82**	**1**	**1.04**	**0.35**	**1**	**0.95**	**1.45**	**4**	**6**	**3**
Philipponnat	**1.03**	**0.79**	**1**	**1.02**	**0.37**	**3**	**0.93**	**1.42**	**4**	**8**	**5**
De Ruiter	**0.98**	**0.77**	**4**	**0.95**	**0.36**	**3**	**0.87**	**1.24**	**4**	**11**	**6**
CPT2000	**1.17**	**0.79**	**6**	**1.11**	**0.34**	**6**	**1.08**	**1.56**	**4**	**16**	**7**
UWA	**1.19**	**0.82**	**6**	**1.17**	**0.31**	**6**	**1.09**	**1.60**	**4**	**16**	**7**
Schmertmann	**1.20**	**0.77**	**6**	**1.21**	**0.35**	**6**	**1.18**	**1.58**	**9**	**21**	**9**
German	**1.03**	**0.67**	**4**	**1.02**	**0.44**	**9**	**0.88**	**1.52**	**9**	**22**	**10**
Eurocode7	1.22	0.74	6	1.17	0.48	10	1.02	1.87	9	25	11
Price and Wardle	0.84	0.79	9	0.83	0.34	9	0.78	1.21	9	27	12
Static	1.16	0.60	9	1.26	0.40	10	1.17	1.71	9	28	13
NGI05	1.28	0.72	11	1.24	0.45	10	1.10	1.96	14	35	14
Tumay Fakhroo	1.29	0.69	11	1.36	0.35	10	1.26	2.02	15	36	15
Fugro	1.44	0.75	13	1.34	0.45	10	1.15	2.14	15	38	16
Purdue	1.45	0.60	13	1.29	0.56	14	1.02	2.36	19	38	18
Aoki	0.83	0.64	9	0.77	0.51	16	0.65	1.27	15	40	18
ICP	1.49	0.74	13	1.33	0.45	10	1.22	2.12	19	42	19
Penpile	0.54	0.85	13	0.59	0.28	15	0.57	0.77	15	43	20
Zhou	1.49	0.85	13	1.68	0.28	17	1.60	2.20	21	51	21
Togliani	1.70	0.81	18	1.83	0.30	18	1.79	2.45	22	58	22

Instead of classical scaling, the Stress Majorization of a COmplicated Function (SMACOF) approach can be used for solving the MDS problem, which offers more flexibility [82], [83]. SMACOF uses Kruskal's stress, $\sigma(X)$ [58] as the target criterion, which is defined using the following equation:

$$\sigma(x) = \sum_{i=1}^{n}\sum_{j=1}^{m} \omega_{ij}\left(\delta_{ij} - d_{ij}(X)\right)^2 \tag{35}$$

where ω_{ij}, is as:

$$\sum_{i=1}^{n}\sum_{j=1}^{m} \omega_{ij}{\delta_{ij}}^2 = n(n-1)/2 \tag{36}$$

The values of ω_{ij} represent the matrix of weights, w, which is a symmetric, non-negative, and hollow matrix. The w matrix can be used for imposing missing values, as $\omega_{ij} = 1$ if δ_{ij} is known and $\omega_{ij} = 0$ if δ_{ij} is missing. Other kinds of weighting structures are also available.

Details of SMACOF solution is available for MDS problems by De Leeuw and Mair and Borg and Groenen [84], [85].

As an example for MDS, below matrix, Δ ($n \times n$ matrix) shows distances (in miles) between 10 cities in Louisiana:

Table 5. Example of MDS showing distances between 10 Louisiana cities

City	New Orleans	Baton Rouge	Shreveport	Lafayette	Lake Charles	monroe	Alexandria	Slidell	Arcadia	Houma
New Orleans	0	75.14	281.1	118.13	189.23	214.42	169.73	28.36	246.29	46.02
Baton Rouge	75.14	0	208.38	52.14	122.19	152.59	95.47	84.79	177.47	65.48
Shreveport	281.1	208.38	0	189.14	162.08	95.13	113.67	281.27	48.43	270.72
Lafayette	118.13	52.14	189.14	0	71.57	158.18	79.34	133.81	169.38	89.26
Lake Charles	189.23	122.19	162.08	71.57	0	170.69	87.95	205.34	161.6	155.99
Monroe	214.42	152.59	95.13	158.18	170.69	0	85.05	207.21	46.79	217.93
Alexandria	169.73	95.47	113.67	79.34	87.95	85.05	0	173.75	90.04	157.06
Slidell	28.36	84.79	281.27	133.81	205.34	207.21	173.75	0	243.03	73.3
Arcadia	246.29	177.47	48.43	169.38	161.6	46.79	90.04	243.03	0	242.32
Houma	46.02	65.48	270.72	89.26	155.99	217.93	157.06	73.3	242.32	0

Solving the MDS to produce two-dimensional X, is as follows:

1) Torgerson's transformation: Create a double-center version of Δ, designated to Δ^*, where the row sums, the column sums, and overall sum of the cell entries in the matrix are zero. For dissimilarity δ_{ij} in matrix Δ, the corresponding δ_{ij}^* in matrix Δ^* is obtained as:

$$\delta_{ij}^{*} = -0.5\left(\delta_{ij}^{2} - \delta_{i.}^{2} - \delta_{.j}^{2} + \delta_{..}^{2}\right) \quad (37)$$

Δ^{*} should be factorized to obtain the matrix of point coordinated, X (an n × p matrix, p=2 for two-dimensional configuration):

$$\Delta^{*} = XX' \quad (38)$$

2) Carry out the factoring process by performing an eigendecomposition on Δ^{*}:

$$\Delta^{*} = V\Lambda^{2}V' \quad (39)$$

Where, V is the n × q matrix of eigenvectors, Λ^{2} is the q × q diagonal matrix of eigenvalues, and q is the rank of Δ^{*} (usually equal to n).

3) Create X from the first p eigenvectors (V_m) and the first p eigenvalues (V_m^2):

$$X = V_m\Lambda_m \quad (40)$$

X, contains point coordinates such that the interpoint distances have a least-squares fit to the entries in Δ.

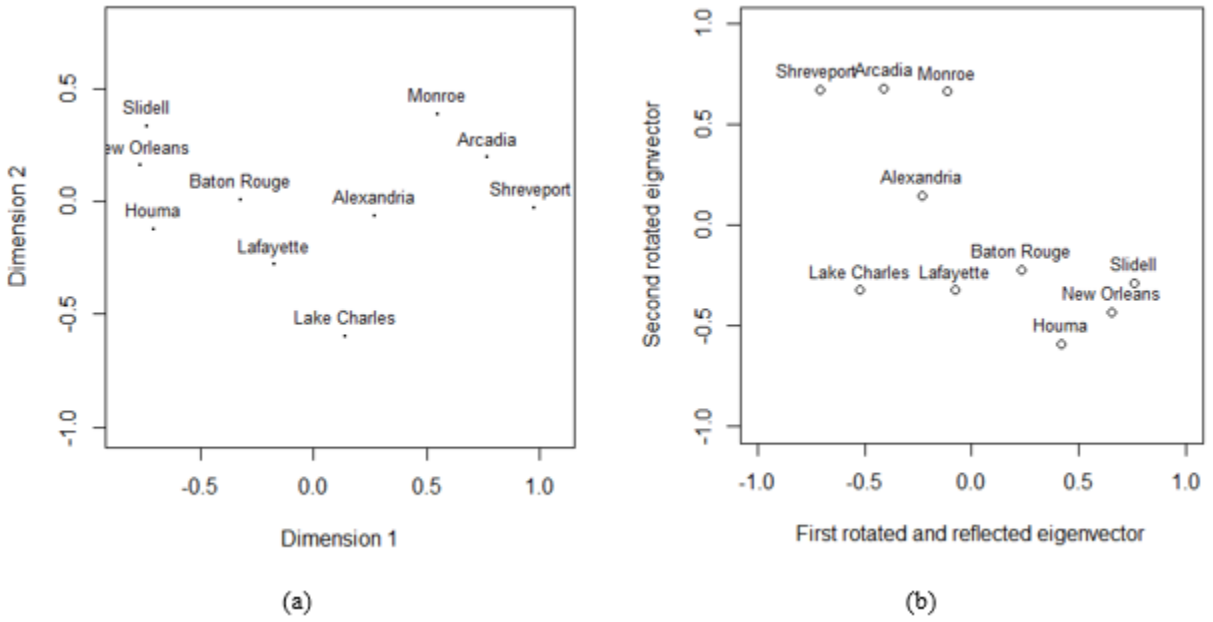
For this example, we used the below code in RStudio software as shown below:

```
getwd()
library(smacof)
library(stringi)
Delta0 <- read.csv(file= "dist.csv")
#View(Delta0)
Delta1 <- Delta0[1:10,2:11]
rownames(Delta1) <- Delta0[,"City"]
#View(Delta1)
fit.LA <- mds(Delta1)
fit.LA$conf
op <- par(mfrow = c(1,2))
plot(fit.LA)
theta <- 315*pi/180 ## degrees to radians
rot <- matrix(c(cos(theta), sin(theta), -sin(theta), cos(theta)), ncol = 2)
fit.LA2 <- fit.LA$conf %*% rot ## rotated configurations
xmirror <- matrix(c(-1, 0, 0, 1), ncol = 2)
fit.LA2 <- fit.LA2 %*% xmirror ## mirror configurations
fit.LA2
plot(fit.LA2 , xlim = c(-1.0, 1.0), ylim = c(-1.0, 1.0),
     main = "LA CITIES", xlab="First rotated and reflected eigenvector", ylab
="Second rotated eignvector")
text(fit.LA2, row.names(fit.LA2), cex=0.8, pos=3, col="black")
```

The two-dimensional configuration, X for dissimilarity matrix, Δ is obtained in Figure 15. Figure 15 (a) presents the original solution of MDS, which by appropriate rotating and reflecting, as shown in Figure 15 (b), the location of the cities in Louisiana can be obtained.

It should be noted here that, in this metric MDS solution, the relative distance between cities is obtained and appropriate rescaling is necessary to obtain the original map.

Figure 15. Graph of eigenvectors (a) MDS solution (b) rotated and first eigenvector reflected



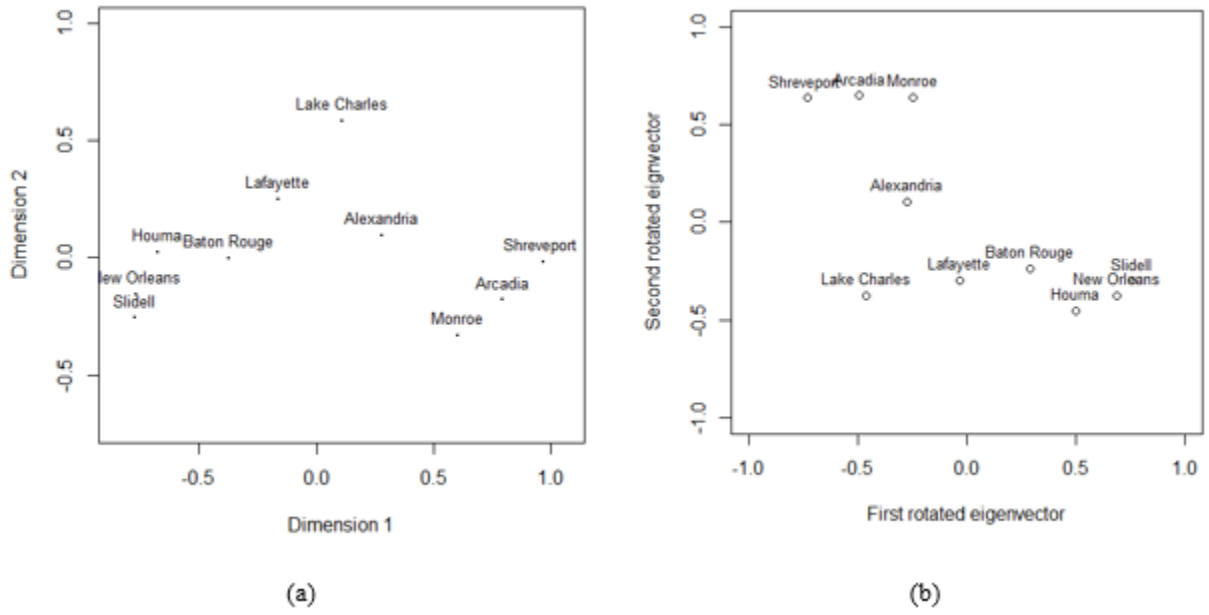
If we use the matrix of rank-ordered distances between the 10 Louisiana cities (ordinal data), we will get the following:

Table 6. Matrix of rank-ordered distances between the 10 Louisiana

City	New Orleans	Baton Rouge	Shreveport	Lafayette	Lake Charles	monroe	Alexandria	Slidell	Arcadia	Houma
New Orleans	0	9	44	19	34	38	29	1	42	2
Baton Rouge	9	0	37	5	20	22	17	11	32	6
Shreveport	44	37	0	33	27	16	18	45	4	43
Lafayette	19	5	33	0	7	25	10	21	28	14
Lake Charles	34	20	27	7	0	30	13	35	26	23
Monroe	38	22	16	25	30	0	12	36	3	39
Alexandria	29	17	18	10	13	12	0	31	15	24
Slidell	1	11	45	21	35	36	31	0	41	8
Arcadia	42	32	4	28	26	3	15	41	0	40
Houma	2	6	43	14	23	39	24	8	40	0

However, it seems to be less information in Δ matrix, the two-dimensional configuration still provides excellent fit, as shown in Figure 16.

Figure 16. Graph of eigenvectors (a) MDS solution (b) rotated eigenvectors for ordinal data



As shown in Figure 16, the metric MDS seems to work for ordinal data, too. However, it imposes an implicit assumption about the relative sizes of differences between dissimilarities and using the concept of eigendecomposition is not appropriate. Therefore, a different strategy should be used for ordinal dissimilarities, which was described as SMACOF approach:

1. Initial configuration is created, randomly.
2. Distance between points is calculated, as d.
3. Optimal monotonic transformation of proximities is found, Based on this, optimal scaled data f(x) can be obtained.
4. Kruskal's stress, $\sigma(X)$ as shown in Equation (35) is determined.
5. By minimizing stress, new configuration of data points can be found.
6. If stress is small enough, terminate the loop; otherwise go to 2.

This approach is very useful to find the best configuration for rank-ordered dissimilarities among different objects.

Multidimensional Unfolding (MDU)

Different MDS techniques have been developed over the years [86]. The MDS can be divided into one-way and multi-way MDS. In multi-way MDS, different individuals (multiple judges and raters, repeated measurements) present dissimilarity for each pair of objects. Multi-mode MDS is the case when dissimilarities are qualitatively different (e.g., objects are rated based on different subjects). Each kind of MDS can be provided in metric and non-metric variants. If dissimilarities, δ_{ij}, are on ordinal scale, transformations of the dissimilarities, $\hat{d}_{ij} = f(\delta_{ij})$

(commonly known as disparities), can be defined to preserve the order as $\delta_{IJ} < \delta_{ij} \Rightarrow \hat{d}_{IJ} < \hat{d}_{ij}$, which is referred to nonmetric. Also, restrictions on configuration matrix, X can be applied as shown by De Leeuw and Heise [83].

As discussed in the previous section, the SMACOF is an MDS solving strategy that uses majorization to minimize stress. For the above extensions of classical MDS, some changes for the SMACOF solution is needed. For example, in the cases when K judges present dissimilarity matrices, additional algorithms like SMACOF routines for individual differences (also known as three-way SMACOF), such as INDSCAL (Individual Differences Scaling), IDIOSCAL, etc. can be used.

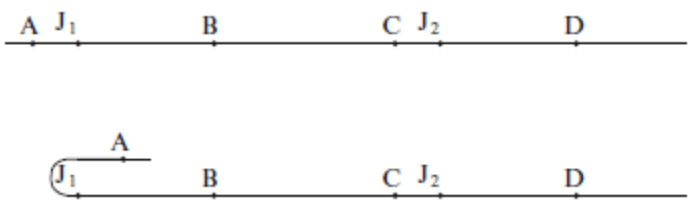
The case that we are going to use in this study is when we have n_1 judges rate n_2 objects. Hence, the dissimilarity matrix is not square. The basic idea is that objects and judges are going to be represented on the same scale.

A simple unidimensional example presented by Alvo and Philip [87] is shown in Figure 17, where judges and objects are shown on a line. For this example, the rankings given to four objects by two judges are as follows:

	First	Second	Third	Fourth
Judge J_1	A	B	C	D
Judge J_2	C	D	B	A

As seen in Figure 17, the objects and judges are placed on the line based on the rankings given to the objects. By folding the line at each judge point, the original ranking of the objects can be observed. For example, as shown in Figure 17, folding the line at point J_1 reveals that judge J_1 prefers objects A, B, C, and D, respectively.

Figure 17. Unidimensional unfolding



It can be seen that it is impossible to place judge 3 giving rankings as DABC for the previous example. The Multidimensional unfolding (MDU) is the MDS technique used for ranking data. The SMACOF approach was extended by De Leeuw and Mair [84] and smacof package in R [88] can be used for solving different MDU problems.

The MDS goodness of fit can be estimated using the standardized version of raw stress, called Kruskal's stress-1, which is somehow not dependent on the absolute values of dissimilarities.

$$\sigma_1(X) = \sqrt{\frac{\sum_{i<j} \omega_{ij}\left(\hat{d}_{ij} - d_{ij}(X)\right)^2}{\sum_{i<j} \omega_{ij} {d_{ij}}^2(X)}} = \sqrt{\frac{\sum_{i<j} \omega_{ij}\left(\hat{d}_{ij} - d_{ij}(X)\right)^2}{n(n-1)/2}} \tag{41}$$

For ordinal MDS, different stress-1 values of 0.2, 0.1, 0.05, 0.025, and 0 represent poor, fair, good, excellent, and perfect fit [3]. The goodness-of-fit of the results also can be estimated using the Shepard diagrams separately for the row and column dissimilarities.

Displaying the ranking data (preferably in two-dimensional space) help us to find the typical ranking of the objects, the extent of agreement between the judges, existence of outliers among the judges and objects, and the similarity between the objects.

MDU Results: 80 Piles

Each pile has been regarded as an individual (judge), which rates the methods based on the values of Q_p/Q_m. If the value of $\left(Q_p/Q_m\right)$ is one, the method's rank becomes one. So the ranking is based on the value of $\left[\text{abs}\left(Q_p/Q_m - 1\right)\right]$. The ranking of the pile-CPT methods has been obtained, which is shown in Table 7.

Table 7. Ranking of pile-CPT methods for each pile from 1 to 22

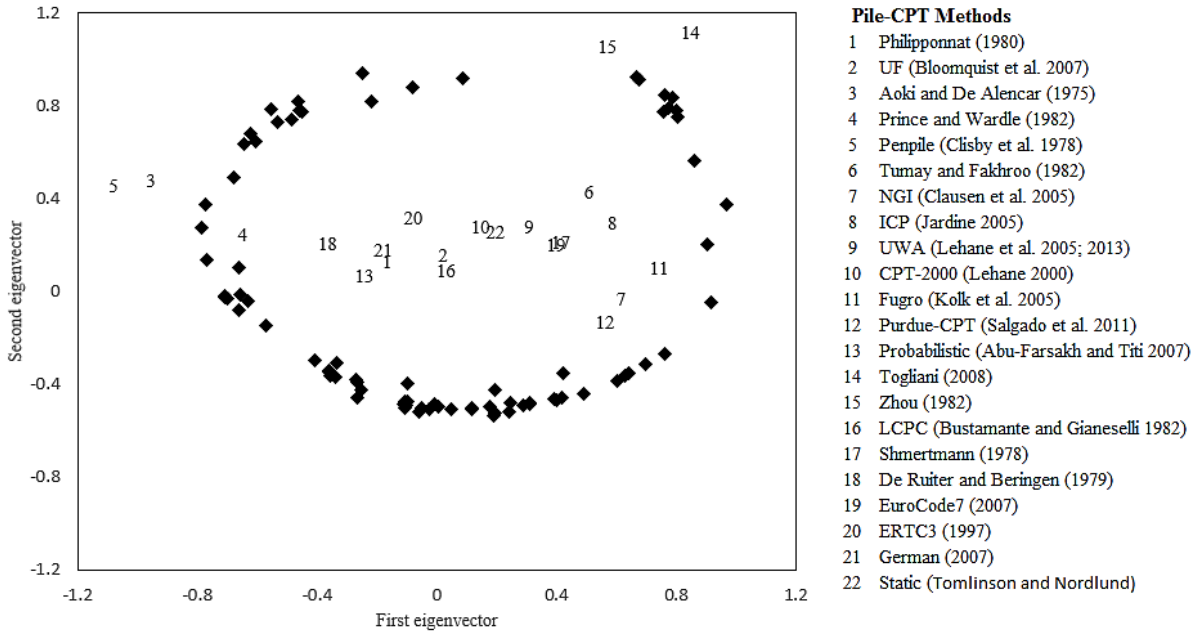
Pile / method	1	2	3	4	5	.....	80
1- Philipponnat	8	7	6	5	13	Rankings 1 to 22 for each Pile-CPT method	8
2- UF	6	13	4	2	14		6
3- Aoki	21	17	19	18	20		20
4- Price and Wardle	7	9	7	11	19		18
5- Penpile	22	19	20	17	21		22
6- Tumay Fakhroo	5	1	3	13	2		7
7- NGI05	12	6	5	9	7		10
8- ICP	17	15	17	19	10		15
.....	...	...	...	...	...		...
.....	...	...	...	...	...		...
22- Static	17	15	17	19	10		12

The below code in RStudio software was used for the MDU analysis in this study:

```
getwd()
library(smacof)
library(stringi)
CPTmethods <- read.csv(file= "methods.csv")
#View(CPTmethods)
CPTranking <- CPTmethods[,2:81]
#View(CPTranking)
rownames(CPTranking) <- CPTmethods[,"method"]
#View(CPTranking)
fit.CPT <- unfolding(CPTranking) ## 2D metric unfolding solution
fit.CPT$stress
fit.CPT[["conf.col"]]
fit.CPT[["conf.row"]]
plot(fit.CPT, label.conf.rows = list(label = TRUE, col=1), label.conf.columns
 = list(label = FALSE),col.rows = hcl(0), col.columns = hcl(240), pch = 10)
best <- sort(rowMeans(CPTranking, na.rm = TRUE))[1:12]
worst <- sort(rowMeans(CPTranking, na.rm = TRUE), decreasing = TRUE)[1:12]
bestworst <- names(c(best, worst))
text(fit.CPT$conf.row[bestworst,], labels = bestworst, cex = 0.8, pos = 3,
     col = hcl(0, l = 50))
```

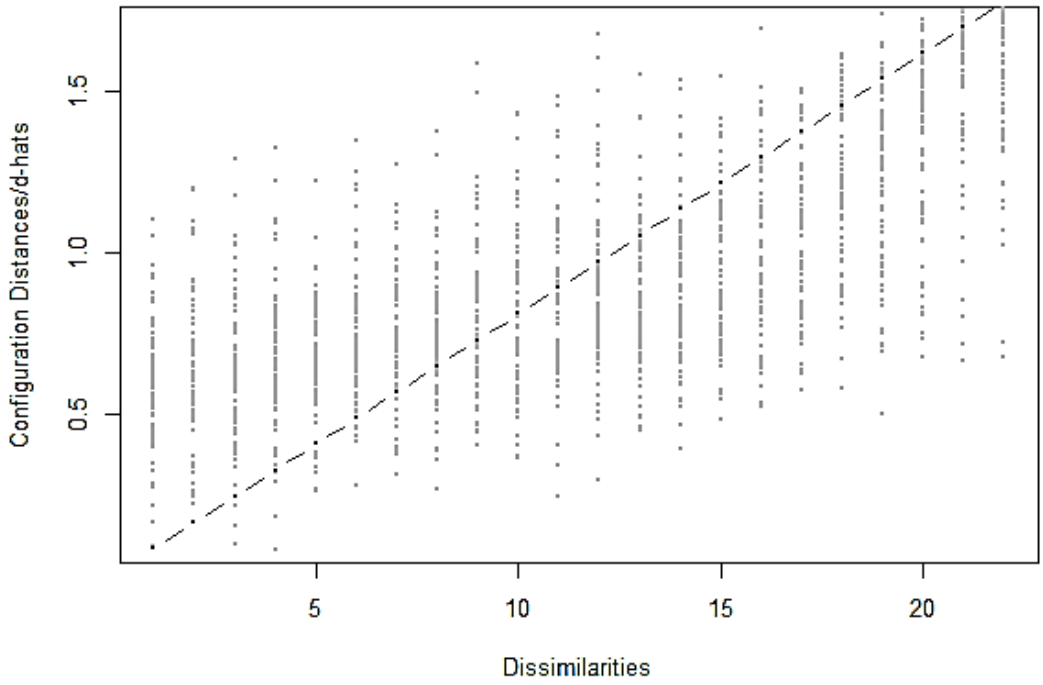
The results of MDU analysis for 21 pile-CPT methods in addition to the static analysis (Tomlinson and Nordlund) method for the 80 PPC piles are shown in Figure 18.

Figure 18. MDU results for the 21 pile-CPT methods and static analysis method (metric MDS)



It should be noted here that the results shown in Figure 18 were obtained by Metric MDS, which means that dissimilarities and distances are linearly related. The Shepard diagram for this metric MDS analysis is shown in Figure 19. The x-axis in Figure 19 represents dissimilarities, which refers to ranking of the methods and changes from 1 to 22. The y-axis represents the distances between points shown in Figure 18. For each ranking (dissimilarity), the 80 points are shown in Shepard diagram represent the distances of the 80 piles from the pile-CPT methods. For metric MDS, a linear relationship between these dissimilarities and distances exist, which is shown in the Shepard diagram. The value of σ_1 for this MDU analysis was obtained as 0.349. The Shepard diagram and σ_1 value suggest that metric MDU is poor and is not the best fit for the results. However, the two-dimensional configuration of points, as shown in Figure 19, is a useful tool for visualizing data.

Figure 19. Shepard diagram (metric MDS)



For the second MDU analysis, the ordinal MDS was used, and the results are shown in Figure 20. The value of σ_1 for this MDU analysis was obtained as 0.103, which is much lower than the metric analysis. However, as shown in Figure 20, most of the pile-CPT methods are located close to each other and the results are not useful for ranking the pile-CPT methods. The Shepard diagram for this analysis is shown in Figure 21, which shows that this analysis does not differentiate much between pile-CPT methods with rankings from 1 to 19. Therefore, the results of this analysis were not used for ranking pile-CPT methods.

Figure 20. MDU results for the 21 pile-CPT methods and static analysis method (ordinal MDS, primary)

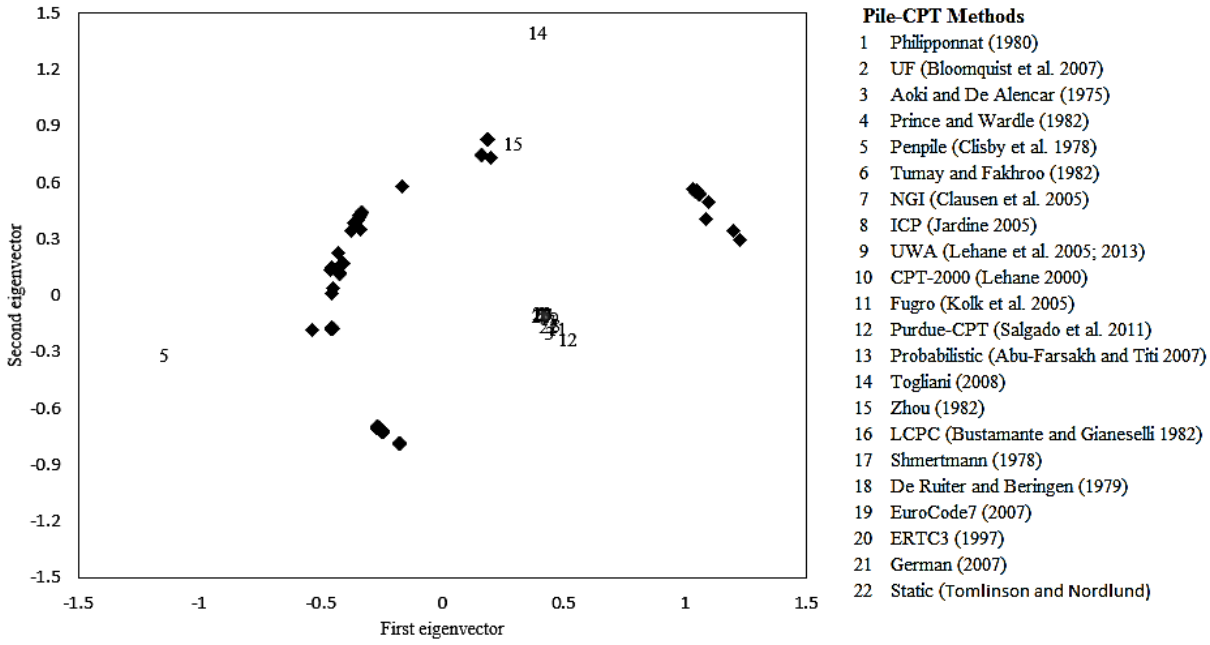
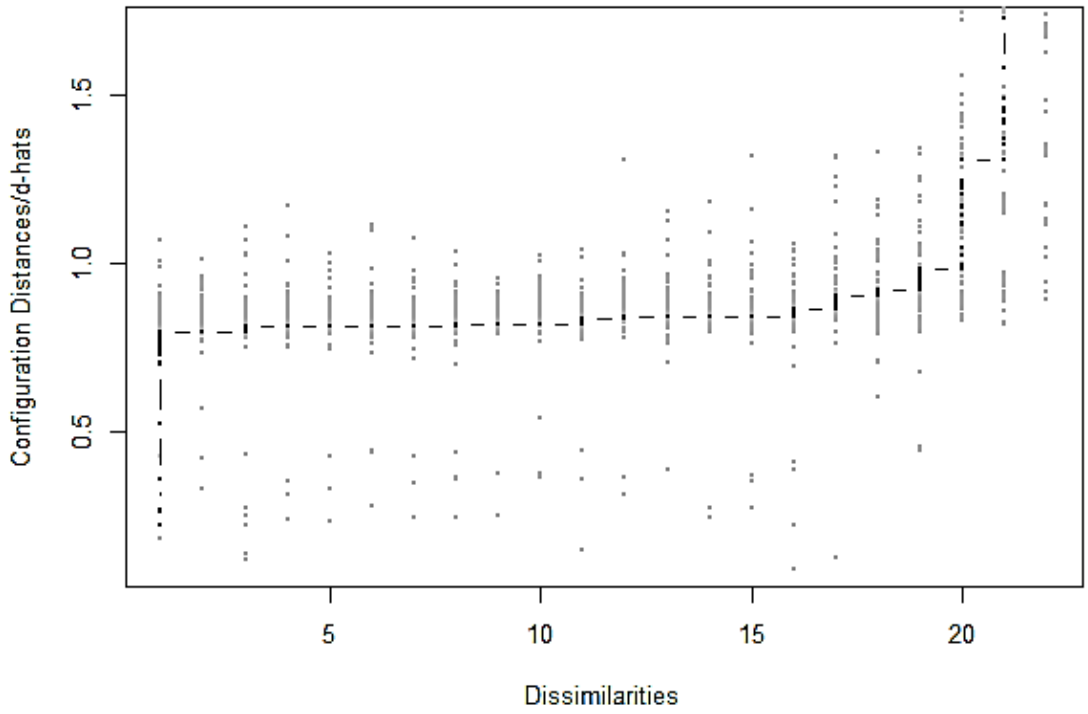


Figure 21. Shepard diagram (ordinal MDS, primary)



For the next MDU analysis, the ordinal secondary MDS was used, and the results are shown in Figure 22. The Shepard diagram for the ordinal secondary MDS analysis is shown in Figure 23. The value of σ_1 for this MDU analysis was obtained as 0.235.

Figure 22. MDU results the 21 pile-CPT methods and static analysis method (ordinal secondary MDS)

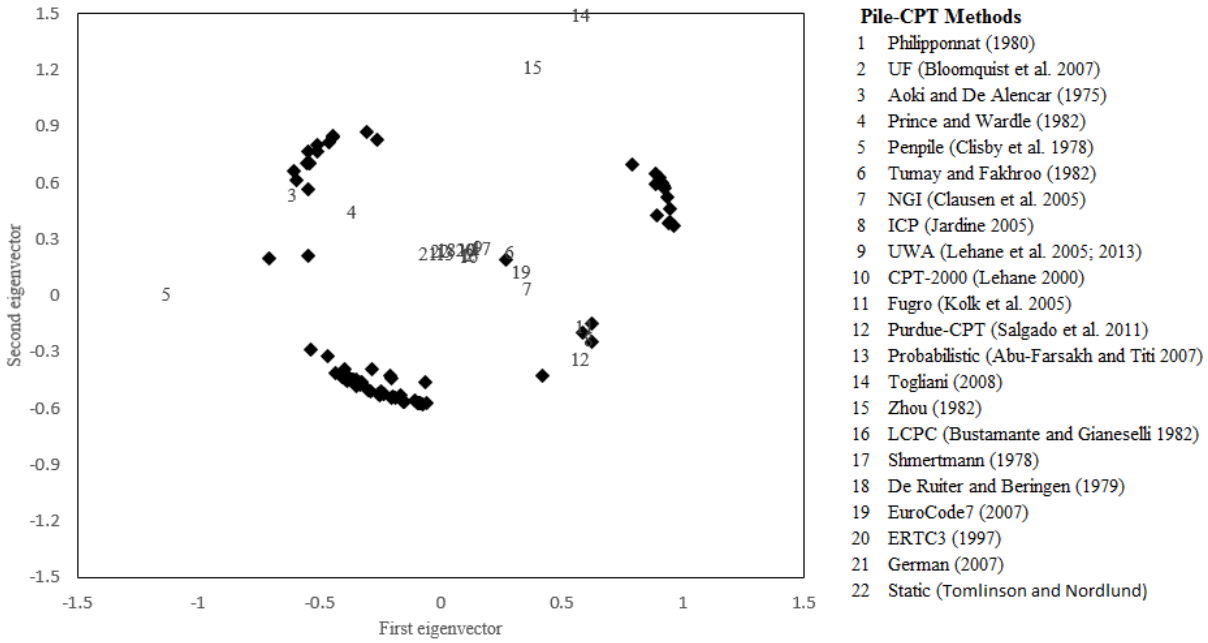
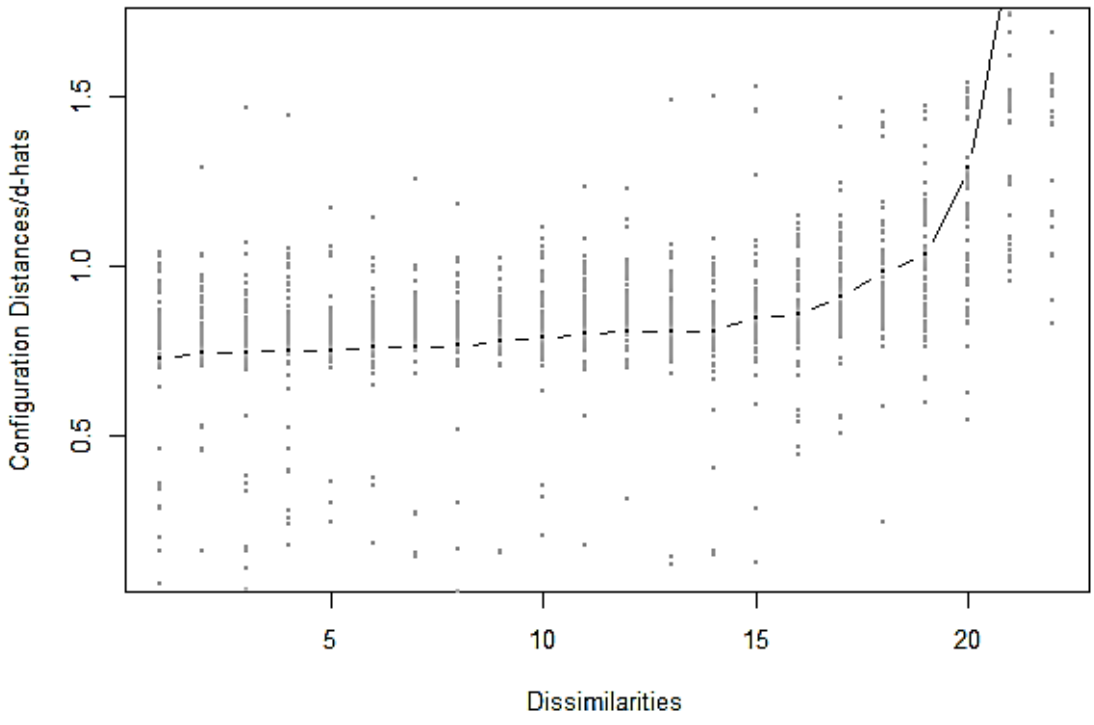


Figure 23. Shepard diagram (ordinal secondary MDS)



The other MDU analysis is monotone spline transform in MDU, and the results are presented in Figure 24. The Shepard diagram for this MDS analysis is shown in Figure 25. The value of σ_1 for this MDU analysis was obtained as 0.268.

Figure 24. MDU results for the 21 pile-CPT methods and static analysis method (monotonic spline MDS)

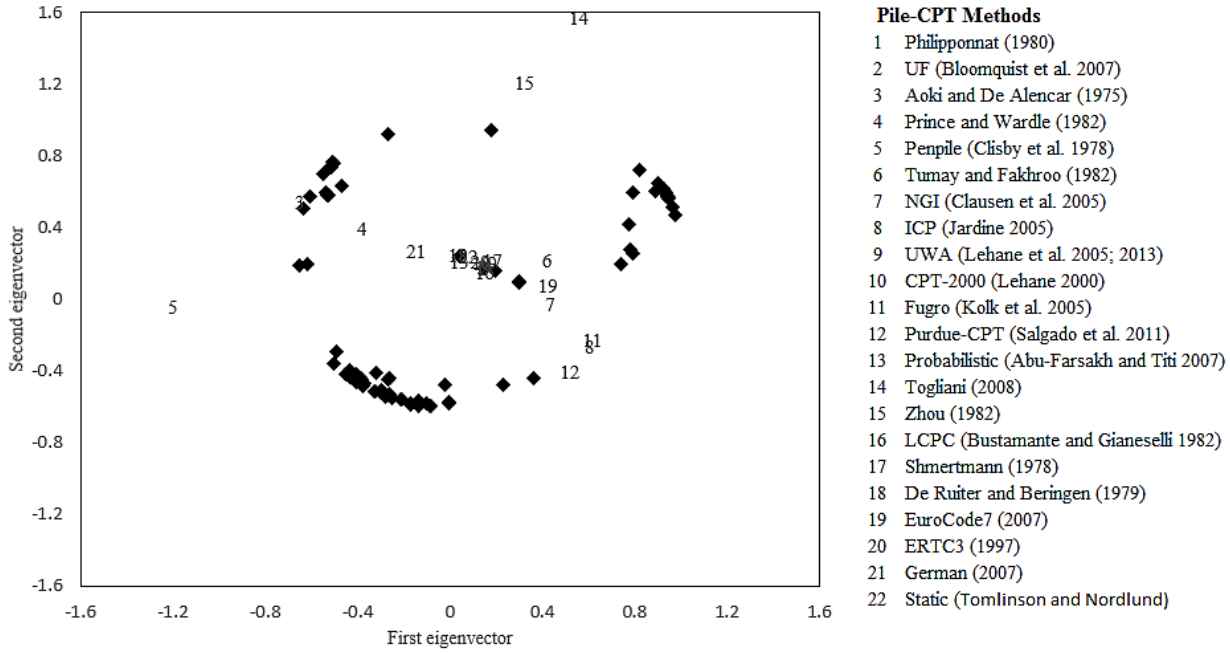
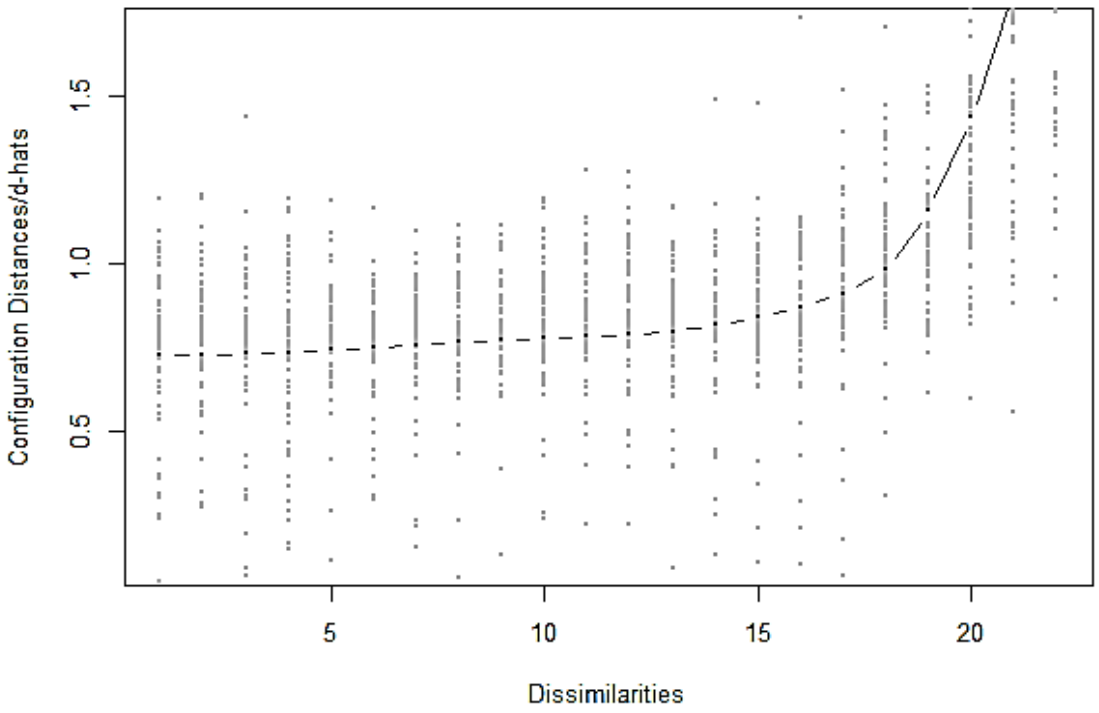


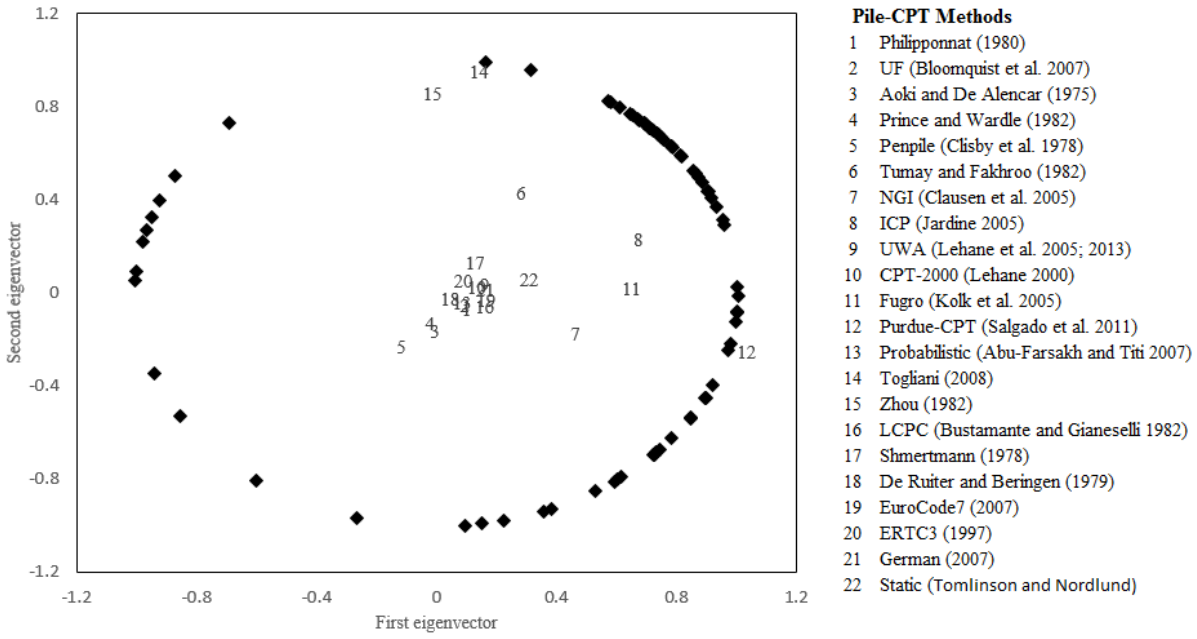
Figure 25. Shepard diagram (monotonic spline MDS)



It can be seen that monotonic-spline and ordinal secondary MDU results are similar. Therefore, ordinal secondary MDU, which has smaller σ_1 value, was used for ranking the methods.

Imposing circle restrictions on the MDU solution for values of $\left[\text{abs}\left(Q_p/Q_m - 1\right)\right]$ led into the results shown in Figure 26 with σ_1 value equal to 0.579.

Figure 26. MDU results for the 21 pile-CPT methods and static analysis method (circular restricted MDS)



Based on the ranking of the pile-CPT methods given by each pile, all the piles and pile-CPT methods were located in a two-dimensional space. The distance of each pile-CPT method with respect to the center (0, 0) coordinate represents how accurate that pile-CPT method in predicting the ultimate pile capacities. For metric, ordinal secondary, and circular restricted MDU results, the distance of each pile-CPT method to the center was calculated and presented in Table 8. Based on the distance from the center, the rank of each pile-CPT was calculated. The final ranking of the methods was obtained based on the summation of rankings for each MDU analysis type, known as RI2. Based on the MDU analysis, the Probabilistic, UF, Philipponnat, German, LCPC, De Ruiter and Beringen, ERTC3, CPT2000, and UWA methods show the lowest distance from the center of the piles, and therefore are ranked the best methods. As shown in Table 8, the results of MDU analysis are not much different from multiple criteria based on mathematical and statistical criteria described in a previous section.

Table 8. Ranking of pile-CPT methods based on MDU analysis

Pile capacity method	Metric MDU		Ordinal Secondary MDU		Circular restricted MDU		Overall rank	
	Distance	Ranking	Distance	Ranking	Distance	Ranking	RI2	Final rank
LCPC	**0.10**	**1**	**0.24**	**4**	**0.17**	**12**	**17**	**5**
ERTC3	**0.33**	**8**	**0.27**	**8**	**0.10**	**3**	**19**	**7**
Probabilistic	**0.25**	**4**	**0.23**	**1**	**0.09**	**2**	**7**	**1**
UF	**0.16**	**2**	**0.25**	**7**	**0.11**	**4**	**13**	**2**
Philipponnat	**0.21**	**3**	**0.24**	**5**	**0.12**	**5**	**13**	**2**
De Ruiter	**0.42**	**10**	**0.25**	**6**	**0.05**	**1**	**17**	**5**
CPT2000	**0.32**	**6**	**0.27**	**9**	**0.14**	**7**	**22**	**8**
UWA	**0.41**	**9**	**0.30**	**10**	**0.16**	**10**	**29**	**10**
Schmertmann	0.46	12	0.31	11	0.18	13	36	12
German	**0.26**	**5**	**0.23**	**2**	**0.16**	**8**	**15**	**4**
Eurocode7	0.45	11	0.35	12	0.17	11	34	11
Price and Wardle	0.70	17	0.58	15	0.13	6	38	13
Static	**0.33**	**7**	**0.24**	**3**	**0.31**	**15**	**25**	**9**
NGI05	0.62	14	0.36	13	0.49	16	43	14
Tumay Fakhroo	0.67	16	0.37	14	0.52	17	47	15
Fugro	0.75	18	0.61	16	0.65	18	52	18
Purdue	0.58	13	0.67	18	1.06	22	53	18
Aoki	1.07	19	0.81	19	0.16	9	47	15
ICP	0.66	15	0.65	17	0.71	19	51	17
Penpile	1.17	20	1.13	20	0.26	14	54	20
Zhou	1.20	21	1.27	21	0.86	20	62	21
Togliani	1.40	22	1.60	22	0.96	21	65	22

Evaluation of Pile-CPT Methods Using Reliability Analysis

Reliability-based calibration obtained from the principles of load and factor design (LRFD) can be used for evaluating the efficiency of the different pile-CPT methods. In this section, the following parts are described.

LRFD: Background

Under the working stress design (WSD), also known as allowable stress design (ASD), the design load, Q, is compared to resistance, or strength, R_n through a factor of safety, FS, which used to account for uncertainties in the applied loads and soil resistance. The magnitude of FS is dependent on the importance of the structure, the confidence level of the material properties, and design methodology. The equation is given as:

$$Q < Q_{all} = \frac{R_n}{FS} = \frac{Q_{ult}}{FS} \tag{42}$$

where, Q is design load, Q_{all} is allowable design load, R_n is resistance of the element or the structure, and Q_{ult} is the ultimate geotechnical pile resistance.

The bridge design specifications published by the American Association of Highway and Transportation Officials (AASHTO) in 1994 and 1998 ([89], [90]) have introduced the LRFD method to account for uncertainties associated with the estimated loads and resistances, separately. In 2007, the AASHTO mandated that all federal-funded new bridges shall be designed using the LRFD method [91].

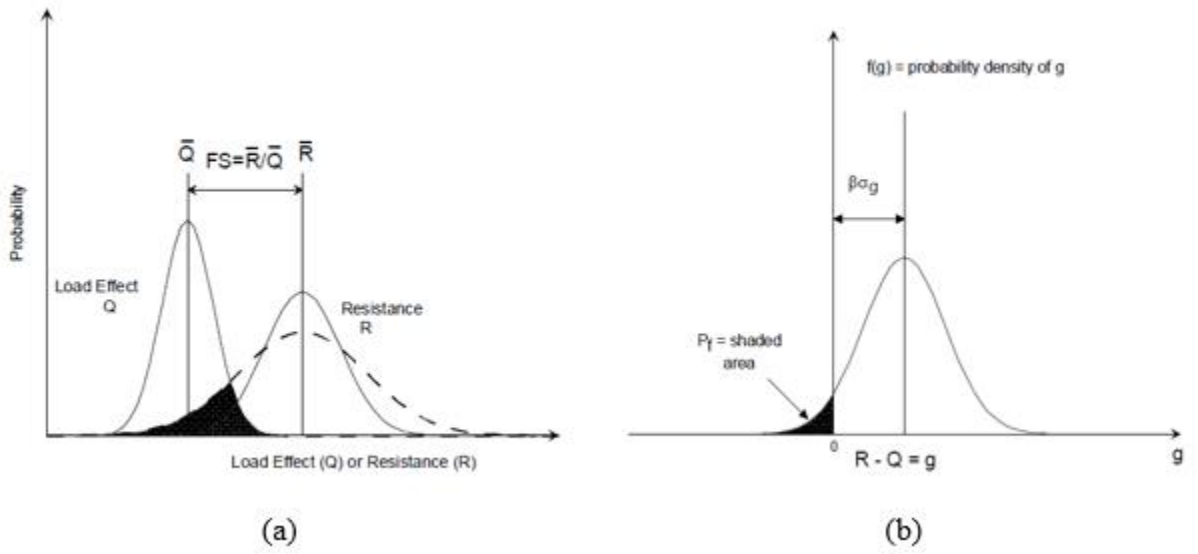
LRFD: Concept

The basic idea behind LRFD is shown in Figure 27. Here, the distributions of random load (Q) and resistance (R) values are shown as normal distributions. The performance limit state function for the state of the structural system can be described as follows:

$$g(R, Q) = R - Q \tag{43}$$

where, R is the resistance of a given structure, which is a random variable, and Q is the applied load, which is also a random variable.

Figure 27. LRFD concept of reliability



If $g \geq 0$, the structure is safe, and if $g < 0$, the structure is unsafe. The probability of failure is then defined as:

$$P_f = p[g(R, Q) < 0] = p[R < Q] \tag{44}$$

For a normal distribution of g values, the probability of failure can be equated explicitly to the value of reliability index $\beta = \mu_g/\sigma_g$, where μ_g is the mean value of g and σ_g is the standard deviation of g. The relationship between probability of failure and reliability index can be calculated using the following function.

$$P_f = 1 - \text{NORMDIST}(\beta) \tag{45}$$

In addition, if the load and resistance values are normally distributed and the limit state function is linear, then β can be determined from the following relation:

$$\beta = \frac{\mu_R - \mu_Q}{\sqrt{\sigma_R^2 + \sigma_Q^2}} \tag{46}$$

where, μ_R and μ_Q are the mean and σ_R, and σ_Q are the standard deviation of resistance and load, respectively.

If both the load and resistance distributions are lognormal and the limit state function is a product of random variables, then β can be calculated using a closed-form solution reported by Withiam et al. and Nowak as follows [92, 93]:

$$\beta = \frac{\ln\left[\mu_R/\mu_Q \sqrt{\left(1+\mathrm{COV}_Q^2\right)/\left(1+\mathrm{COV}_R^2\right)}\right]}{\sqrt{\ln\left[\left(1+\mathrm{COV}_Q^2\right)\left(1+\mathrm{COV}_R^2\right)\right]}} \tag{47}$$

where, μ_R is the mean value of the resistance R, and μ_Q is the mean value of the load Q; COV_R and COV_Q are the coefficients of variation for the resistance and load values, respectively.

LRFD: Calibration

The basic equation of LRFD is shown in equation (48) below. The idea is that the design resistance (which is the measured resistance decreased by the factor, ϕ, known as resistance factor) should be more than the summation of the design loads (which are measured loads Q_i, increased by load factors γ_i).

$$\phi R_n \geq \sum \eta \gamma_i Q_i \tag{48}$$

where, ϕ = resistance factor, R_n = nominal resistance, η = load modifier to account for effects of ductility, redundancy and operational importance. The value of η usually is 1.00. Q_i = Load effect, γ_i =Load factor.

Most of driven piles develop both skin and toe resistances, but the percentage of skin or toe resistance to total resistance is not constant. Therefore, it is not possible to provide a fixed correlation between the three resistance factors (skin, toe and total resistances). In this research only the resistance for total resistance was calibrated. Thus, it should be noted that the same resistance factors for skin and end bearing are assumed and the calibrated resistance factors are valid only for the ranges of pile dimensions (length and diameter) that employed in this study.

Consider the load combination of dead load and live load for AASHTO Strength I case, the performance limit equation is as follows:

$$\phi R_n = \gamma_D Q_D + \gamma_L Q_L \tag{49}$$

where, Q_D and Q_L are the dead load and live load, respectively, and γ_D and γ_L are the load factors for dead load and live load, respectively.

The load parameters in LRFD were studied extensively by different researchers, and the following LRFD parameters have been suggested, and specified by AASHTO LRFD Specifications (e.g., [92], [94]):

$$\gamma_L = 1.75 \qquad \gamma_{QL} = 1.15 \qquad COV_{QL} = 0.18$$
$$\gamma_D = 1.25 \qquad \gamma_{QD} = 1.08 \qquad COV_{QD} = 0.128 \tag{50}$$

where, γ_D and γ_L are the load factors for dead load and live load, respectively. λ_{QD} and λ_{QL} are the load bias factors (mean ratio of measured over predicted value) for the dead load and live load, respectively. COV_{QD} and COV_{QL} are the coefficient of variation values for the dead load and live load, respectively.

The Q_D/Q_L is the dead load to live load ratio which varies depending on the span length [4]. In this research, Q_D/Q_L of 3 is used.

Reliability based analyses using the First Order Second Moment (FOSM) method, the modified FOSM by Bloomquist et al. [5], the First Order Reliability Method (FORM), and the Monte Carlo (MC) simulation method were used to calibrate the resistance factors (ϕ) for the different pile-CPT design methods. A target reliability index (β_T) of 2.33 was selected for the calibration of resistance factors, similar to previous studies (e.g., [95]).

First Order Second Moment (FOSM) Method

In FOSM, the limit state function is linearized by expanding the Taylor series about the mean value of variable. Since only the mean and variance are used in the expansion, it is called first (mean) order second (variance) moment. For lognormal distribution of resistance and load statistics, Barker et al. suggested the following relation for calculating reliability index, β, as [96]:

$$\beta = \frac{\ln\left[\lambda_R FS\left(\frac{Q_{DL}/Q_{LL} + 1}{\lambda_{DL}\, Q_{DL}/Q_{LL} + \lambda_{LL}}\right)\sqrt{\frac{1 + COV_R^2 + COV_{DL}^2 + COV_{LL}^2}{1 + COV_R^2}}\right]}{\sqrt{\ln[(1 + COV_R^2)(1 + COV_{DL}^2 + COV_{LL}^2)]}} \tag{51}$$

For LRFD, this equation is modified by replacing the overall factor of safety (FS) by partial factor of safety and then rearranges to express relation for resistance factor (ϕ) as follows:

$$\phi = \frac{\lambda_R \left(\lambda_D \frac{Q_D}{Q_L} + \lambda_L\right) \sqrt{\frac{1 + COV_{QD}^2 + COV_{QL}^2}{1 + COV_R^2}}}{\left(\lambda_{QD} \frac{Q_D}{Q_L} + \lambda_{QL}\right) \exp\left(\beta_T \sqrt{\ln(1 + COV_R^2)\left(1 + COV_{QD}^2 + COV_{QL}^2\right)}\right)} \tag{52}$$

The resistance bias factor, λ_R and the resistance coefficient of variation, COV_R are important in estimating ϕ and is calculated based on the following equations:

$$\lambda_R = \frac{\sum \lambda_{Ri}}{N} \tag{53}$$

$$\sigma_R = \sqrt{\frac{\sum(\lambda_{Ri} - \lambda_R)^2}{N - 1}} \tag{54}$$

$$COV_R = \frac{\sigma_R}{\lambda_R} \quad \text{and} \quad \lambda_{Ri} = \frac{R_{mi}}{R_{ni}} \tag{55}$$

where, R_{mi} = measured resistance from a load test, and R_{ni} = predicted resistance from lab or field data.

Modified FOSM Method

Equation (52) used for FOSM has been shown to estimate ϕ values for about 10-15% less than the other methods. A modification to the equation was suggested by Styler [50] in the term COV(Q). Equation (52) assumes $COV(Q) = COV(Q_D) + COV(Q_L)$, but it has found that COV(Q) should be obtained by:

$$COV(Q) = \sqrt{\frac{\frac{Q_D}{Q_L}^2 \lambda_{QD}{}^2 COV_{QD}^2 + \lambda_{QL}{}^2 COV_{QL}^2}{\frac{Q_D}{Q_L}^2 \lambda_{QD}{}^2 + 2\frac{Q_D}{Q_L}\lambda_{QD}\lambda_{QL} + \lambda_{QL}{}^2}} \tag{56}$$

Using this modification in equation (55), the difference between the FOSM and other reliability methods becomes slight. The modified FOSM's resistance factor (ϕ) equation becomes:

$$\phi = \frac{\lambda_R \left(\lambda_D \frac{Q_D}{Q_L} + \lambda_L\right) \sqrt{\frac{1 + COV_Q^2}{1 + COV_R^2}}}{\left(\lambda_{QD} \frac{Q_D}{Q_L} + \lambda_{QL}\right) \exp\left(\beta_T \sqrt{\ln(1 + COV_R^2)\left(1 + COV_Q^2\right)}\right)} \tag{57}$$

First Order Reliability Moment (FORM) Method

Hasofer and Lind proposed a modified reliability index that did not exhibit the invariance problem [97]. The "correction" is to evaluate the limit state function at a point known as the "design point" instead of the mean values. The design point is a point on the failure surface g = 0. Since the design point is generally not known in advance, an iteration technique must be used to solve the reliability index. Detailed procedure regarding FORM can be found in Nowak and Collins [98]. The following steps describe the FORM using the Rackwitz-Fiessler method [99]:

1. Define limit state function, g(x1, x2, x3....).
   The limit state function for LRFD is developed as follows:

$$Q = Q_D + Q_L = \lambda_D Q_D + \lambda_L Q_L \tag{58}$$

$$\phi R = \gamma_D Q_D + \gamma_L Q_L \tag{59}$$

   From the above equations:

$$g(R.L) = \left(\frac{\gamma_D Q_D + \gamma_L Q_L}{\phi}\right)\lambda_R - (\lambda_D Q_D + \lambda_L Q_L) \tag{60}$$

   The specified live load to dead load ratio,(Q_L/Q_D) equation can be rearranged as:

$$g(R,Q) = \left(\frac{\gamma_D + \gamma_L \kappa}{\phi}\right)\lambda_R - (\lambda_D + \lambda_L \kappa) \tag{61}$$

where, $\kappa = {}^{Q_L}/_{Q_D}$

2. Assume an initial design point (x_i^*), which is usually the mean values are considered in most cases. Initial design values for dead load and live load (x_2 and x_3) assumed and the resistance (x_1) is determined by equating the limit state function to zero. For lognormal variables equivalent normal parameters are then determined as follows:

$$\mu_x^e = x^* - \sigma_x^e\left[\Phi^{-1}\left(F_x(x)\right)\right] \tag{62a}$$

$$\sigma_x^e = \frac{1}{f_x(x^*)}\varphi\left[\frac{x^* - \mu_x^e}{\sigma_x^e}\right] = \frac{1}{f_x(x^*)}\varphi\left[\Phi^{-1}\left(F_x(x^*)\right)\right] \quad (62b)$$

where, φ and Φ denotes the mass probability density function (PDF) and the cumulative distribution function (CDF) for normal distribution, respectively.

3. The reduced variable corresponding to the design point x* is found as:

$$z_i^* = \frac{x^* - \mu^e x_i}{\sigma^e x_i} \quad (63)$$

4. Partial derivatives of the limit state function is found at the design point, and vector G is defined as:

$G = \begin{Bmatrix} G_1 \\ G_2 \\ G_3 \end{Bmatrix}$, where, $G_i = -\frac{\partial g}{\partial z_i}$ at design point $= -\frac{\partial g}{\partial x_i} * (\sigma^e x_i)$ at design point. The values of β and α will then be determined as:

$$\beta = \frac{\{G\}^T\{z^*\}}{\sqrt{\{G\}^T\{G\}}} \quad (64)$$

where, $\{z^*\} = \begin{Bmatrix} z_1^* \\ z_2^* \\ z_3^* \end{Bmatrix}$

$$\alpha = \frac{\{G\}}{\sqrt{\{G\}^T\{G\}}} \quad (65)$$

5. The new design point is determined in the reduced variable as follows:

$$z_i^* = \alpha_i \beta \quad (66)$$

$$x^*{}_i = \mu_{xi}^e + z_i \sigma_{xi}^e \quad (67)$$

The new design point for resistance (x_1) is determined by inserting the new design values for loads (x_2 and x_3) into the g function. With new design points, steps from 1 to 5 are followed again iteratively. The process is repeated until β and the design point converges. In this study, the excel sheet was used to get the FOSM solution with the "Goal seek" function for given load and resistance statistics. Iterations for FORM is done using the "SOLVER" tool.

Monte Carlo simulation Method

For more complicated limit state functions, the application of the general statistical method for the calculation of the reliability index is either extremely difficult or impossible. Under this circumstance, Monte Carlo simulation provides the only feasible way to determine the reliability index or the probability of failure.

The Monte Carlo method is a technique by which a random number generator is used to extrapolate cumulative density function (CDF) values for each random variable. Extrapolation of CDF makes estimating β possible; otherwise, a limited quantity of data has restricted the reliable estimate of β. Once reliability index, β, is estimated, the probability of failure can be estimated by assuming the distribution of g(x). The steps of Monte Carlo simulation method are as follows:

1. Select a trial resistance factor (ϕ). Generate random numbers for each set of variables. Here there are three variables (resistance, dead load to live load, and bias factor), so three sets of random variables have to be generated independently for each case. The number of simulation points required is found using the following relation:

$$N = \frac{1 - P_{true}}{V_p^2 * (P_{true})} \tag{68}$$

   where, P_{true} is the lowest magnitude of probability that is to be determined using Monte Carlo simulation, and V_p is the desired coefficient of variation of the simulation result. For estimating probability as low as 10^{-2} and keeping variance under 10%, the number of points to be generated in Monte-Carlo simulation is 9900.

   For each lognormal variable, sample value x_i is estimated as:

$$x_i^* = \exp(\mu_{\ln x} + z_i \, \sigma_{\ln x}) \tag{69}$$

   where, $\sigma_{\ln x}^2 = \ln(V_x^2 + 1)$ and $\mu_{\ln x}^2 = \ln(\mu_x) - {}^1\!/_2 \, \sigma_{\ln x}^2$

   In the above expressions, μ_x and V_x are the arithmetic mean and variance of x; $\mu_{\ln x}$ and $\sigma_{\ln x}$ are equivalent lognormal mean and standard deviation of ln(x); and z_i is NORMSINV(RAND), the random standard normal variable generated using EXCEL function.

2. Define the limit state function (equation 48).

$$Q = \lambda_D Q_D + \lambda_L Q_L \tag{70}$$

From equations (49) and (70):

$$g(R,Q) = \left(\frac{\gamma_D Q_D + \gamma_L Q_L}{\varphi}\right)\lambda_R - (\lambda_D Q_D + \lambda_L Q_L) \tag{71}$$

where, equation (71) can be rearranged to:

$$g(R,Q) = \left(\frac{\gamma_D + \gamma_L \kappa}{\varphi}\right)\lambda_R - (\lambda_D + \lambda_L \kappa) \tag{72}$$

where, $\kappa = {Q_L}/{Q_D}$

3. Find the number of cases where $g(x_i) \leq 0$. The probability of failure is then defined as:

$$P_f = \frac{g \leq 0}{N} \tag{73}$$

and the reliability index β is estimated as:

$$\beta = \Phi^{-1}(P_f) \tag{74}$$

4. If the calculated reliability index (β) is different from the selected target reliability index (β_T), the trial resistance factor (ϕ) in step 1 should be changed and iteration needs to be done until $|\beta - \beta_T| <$ tolerance (0.01 in this study).

Results of LRFD Calibration and Efficiency of Pile-CPT Methods

Ranking of different pile-CPT methods can be determined by calculating the efficiency of each method and compare them. The predicted capacity from an individual pile-CPT method, R_n, is used to find the design capacity of the pile, as:

$$R_{design} = \phi R_n \tag{75}$$

However, the bias factor is:

$$\lambda_R = \frac{R_m}{R_n} \tag{76}$$

where, R_m is the measured capacity (using a criteria such as Davisson). Combining the equations:

$$R_{design} = \left({\phi}/{\lambda_R}\right) R_m \tag{77}$$

The term (ϕ/λ_R) identifies the percentage of the measured Davisson capacity that is available for design. Therefore, this term represents the efficiency of the pile-CPT methods. The higher the value of (ϕ/λ_R), the better the method is.

It should be noticed that for calculating λ_R for each pile the proportion of measured to predicted resistance, λ_{Ri} should be calculated and the average of these values should be considered as the resistance bias factor, λ_R. The target reliability, β_T of 2.33 was selected for driven piles, similar to previous studies (e.g., [95]).

The histogram and lognormal distribution of Q_m/Q_p for selected pile-CPT methods are presented in Figure 28. The resulting resistance factors (ϕ) using the different reliability methods are presented in Table 9. The ranking of all pile-CPT methods based on the mean of bias and efficiency were also judged and presented in Table 9. The closer the mean of bias (λ) to one and the higher efficiency leads to better ranking.

Comparing Table 9 with previous criteria (Table 4 for multiple criteria and Table 8 for MDU analysis), it is clear that the efficiency criteria based on LRFD is almost consistent with the previous criteria and LCPC, ERTC3, Probabilistic, UF, Philipponnat, De Ruiter, CPT2000, UWA, Schmertmann, German, and Eurocode7 are among the methods with highest rankings. The main difference between these rankings is seen for Zhou, Togliani, and Penpile methods. Based on the previous criteria, these methods were considered as the lowest pile-CPT methods. Zhou and Togliani methods overpredict the pile capacity with mean values for Q_p/Q_m as 1.68 and 1.83, respectively. This overprediction caused that the ability of these methods for estimating the pile capacity was considered weak. However, the standard deviations of these methods are low which compensates the weakness of these methods in overprediction aspect. In fact, LRFD criteria shows that Zhou and Togliani methods have low resistance factors of 0.38 and 0.35, and efficiency values, φ/λ_R of 0.60 and 0.59, respectively. The high values of efficiency for these methods suggest that they have to be modified to be considered for predicting the pile capacity. On the other hand, Penpile method is a method that underpredicts the pile capacity with mean value of 0.59 and, therefore, in previous criteria was considered as a method with low ranking. Using LRFD criterion, the resistance factor, φ for this method obtained 1.00 and, therefore, a high value of efficiency values, φ/λ_R as 0.54.

LRFD analysis suggest that using Zhou and Togliani method with low resistance factor of 0.35 and Penpile method with high resistance factor of 1.0 lead into acceptable predictions for the resistance for pile design. However, as seen for the other methods, the usual value for resistance factor, φ is in the range of 0.5-0.6. A very simple solution for making these methods consistent with the others is to modify them by multiplying a coefficient (equal to $1/\lambda_R$) to their equations.

Figure 28. Histogram and lognormal distribution of Q_m/Q_p for selected pile-CPT methods

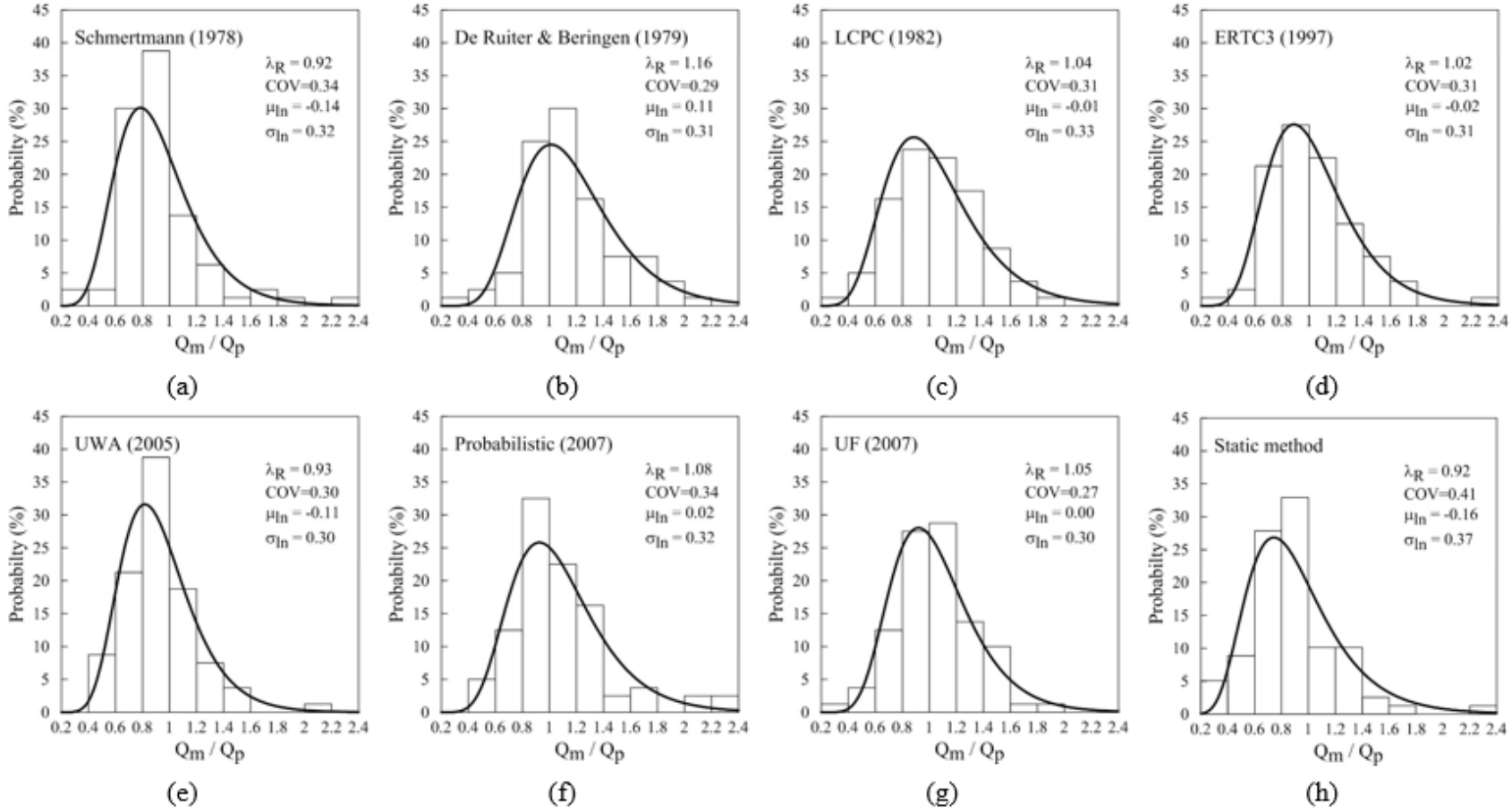


Table 9. Efficiency of Pile-CPT methods based on LRFD reliability criterion

Pile capacity method	Mean, λ_R	COV	σ	ϕ FOSM	ϕ FOSM modified	ϕ FORM	ϕ Monte Carlo	ϕ /λ_R	Final rank
LCPC	1.04	0.31	0.32	0.54	0.59	0.60	0.60	0.57	5
ERTC3	1.02	0.31	0.32	0.53	0.58	0.59	0.59	0.57	6
Probabilistic	1.08	0.34	0.37	0.52	0.57	0.58	0.57	0.53	7
UF	1.05	0.27	0.29	0.58	0.65	0.65	0.65	0.62	1
Philipponnat	1.09	0.30	0.33	0.57	0.64	0.64	0.64	0.59	3
De Ruiter	1.16	0.29	0.34	0.62	0.69	0.69	0.69	0.59	2
CPT2000	1.00	0.34	0.34	0.48	0.53	0.53	0.53	0.53	11
UWA	0.93	0.30	0.28	0.49	0.54	0.54	0.54	0.58	7
Schmertmann	0.92	0.34	0.31	0.44	0.48	0.49	0.48	0.53	14
German	1.14	0.37	0.43	0.51	0.56	0.56	0.56	0.49	12
Eurocode7	0.99	0.35	0.35	0.47	0.51	0.52	0.51	0.52	13
Price and Wardle	1.37	0.43	0.59	0.55	0.59	0.59	0.59	0.43	14
Static	0.91	0.42	0.38	0.38	0.41	0.41	0.41	0.45	19
NGI05	0.94	0.38	0.36	0.42	0.45	0.45	0.45	0.48	17
Tumay Fakhroo	0.81	0.30	0.25	0.42	0.47	0.47	0.47	0.58	10
Fugro	0.87	0.38	0.33	0.39	0.42	0.43	0.42	0.49	17
Purdue	0.94	0.38	0.35	0.42	0.46	0.46	0.46	0.49	16
Aoki	1.56	0.37	0.58	0.70	0.76	0.77	0.76	0.49	9
ICP	0.89	0.42	0.38	0.36	0.39	0.39	0.39	0.44	20
Penpile	1.86	0.33	0.62	0.91	1.00	1.00	1.00	0.54	4
Zhou	0.64	0.29	0.19	0.34	0.38	0.38	0.38	0.60	21
Togliani	0.59	0.30	0.18	0.31	0.35	0.35	0.35	0.59	22

Clustering of Pile-CPT Methods

Based on multiple statistics criteria, MDU analysis, and efficiency based on LRFD reliability analysis, the following pile-CPT methods showed acceptable performance in evaluating the ultimate capacity of driven piles in Louisiana soil: LCPC, ERTC3, Probabilistic, UF, Philipponnat, De Ruiter, CPT2000, UWA, and Schmertmann.

Table 10 shows the ranking of these methods based on each evaluation criterion:

Table 10. Ranking of top nine pile-CPT methods

Pile capacity method	Multiple criteria	MDU analysis	LRFD reliability analysis
LCPC	1	5	9
ERTC3	1	7	8
Probabilistic	3	1	11
UF	3	2	1
Philipponnat	5	2	5
De Ruiter	6	5	3
CPT2000	7	8	13
UWA	7	10	6
Schmertmann	9	12	12

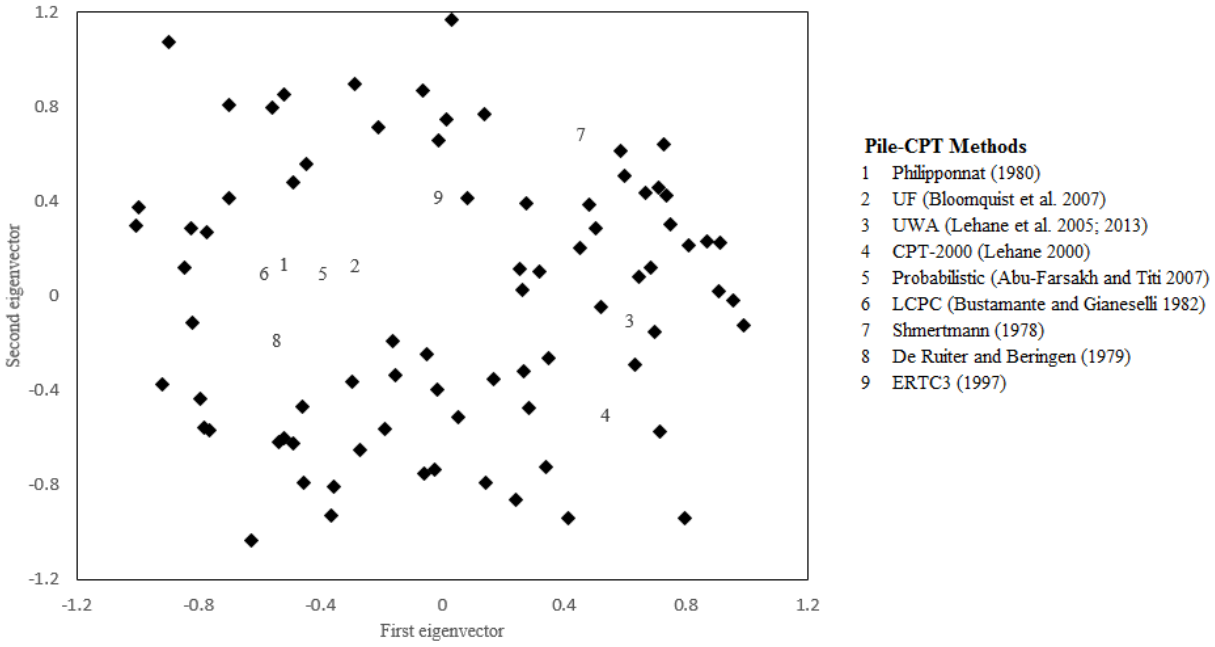
These selected pile-CPT methods can be shown in a two-dimensional configuration in Figure 29 (using MDU analysis as described in previous sections).

Displaying the ranking data in two-dimensional configuration enables us to identify if there is any similarity between the pile-CPT methods. In other words, if a pile-CPT method is close to a particular pile in

Figure 29, it means that the prediction of that pile-CPT method is close to the measured capacity of that particular pile. When two or three pile-CPT methods are close to each other, it means that those pile-CPT methods have similar predictions for the piles. For piles that are close to those pile-CPT methods, the predictions were accurate and for piles far from them, the predictions were different from measured capacities.

This concept shows us that we can use Figure 29 for dividing pile-CPT methods into different groups. For this purpose, K-means algorithm used in unsupervised machine learning was used.

Figure 29. MDU results for nine pile-CPT methods (Metric MDS)



K-means Clustering: Concept

The K-means algorithm is a method for automatically clustering similar data to each other. The idea is to start by guessing the initial centroids for each cluster, assigning data points to the closest centroids, re-computing the centroids based on these assignments, reassigning data points, and doing iterations to get the final groups.

The K-means algorithm can be explained through an example: If we want to cluster the 80 piles shown in

Figure 29 into three groups (K=3). The following algorithm is used:

1. Initially, three centroids as (-0.5, -0.5), (0, 0), and (0.5, 0.5) was chosen.
2. Closest points to the centroids are determined, and their index from 1 to 3 is determined as shown in Figure 30.
3. Based on the average of points in each index, the new location of centroids was determined as shown in Figure 31.
4. A loop from 2 to 4, until reaching to the minimum for the cost function.

$$J\left(c^{(1)}, \dots, c^{(m)}, \mu_1, \dots, \mu_K\right) = \frac{1}{m}\sum_{i=1}^{m}\left|x^{(i)} - \mu_{c(i)}\right|^2 \tag{78}$$

Figure 30. First iteration of K-means algorithm

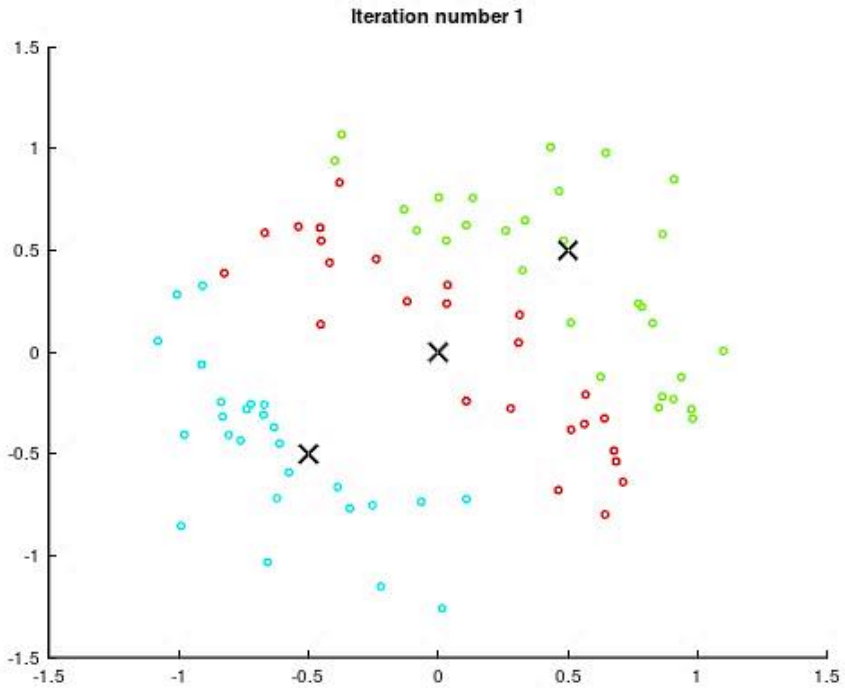
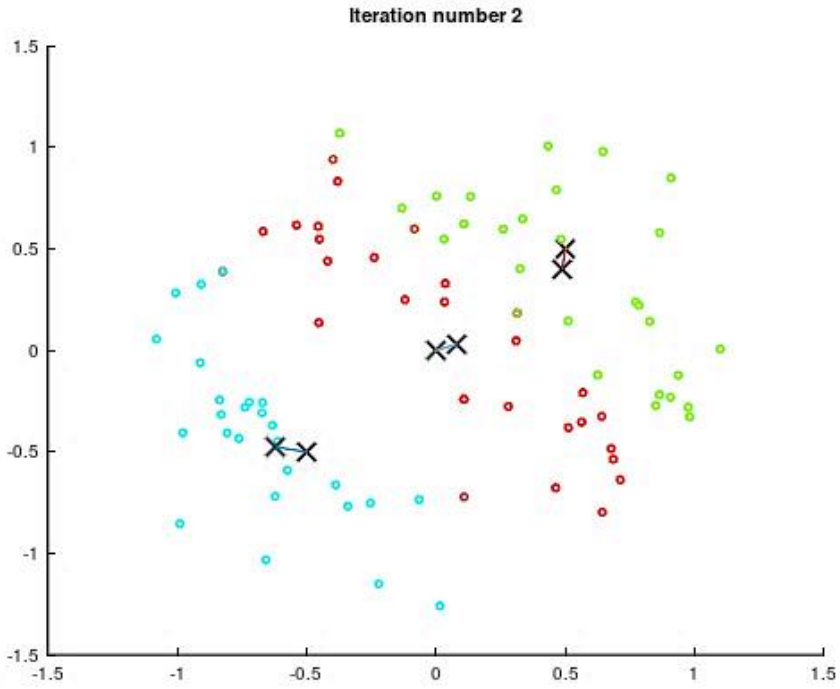


Figure 31. Second iteration of K-means algorithm

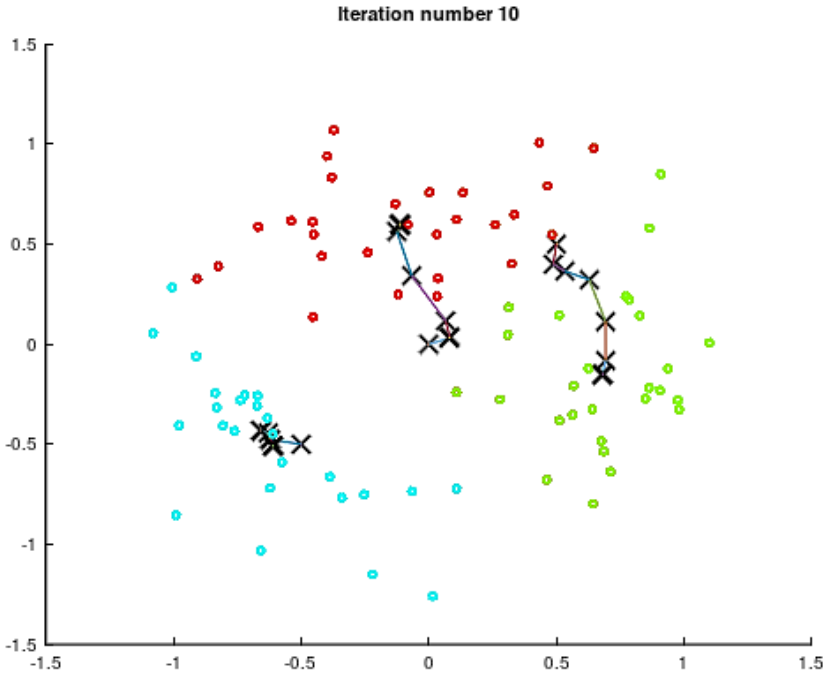

Iteration number 2
1.5
1
0.5
0
-0.5
-1
-1.5
-1.5
-1
-0.5
0
0.5
1
1.5

where, m is the number of data points (X), K is number of clusters, $c^{(i)}$ is the index of cluster (1,…,K) to which $x^{(i)}$ is assigned, and $\mu_{c(i)}$ is the coordinate of the centroid which example $x^{(i)}$ is assigned to it.

The purpose of K-means algorithm is to find the optimization of cost function (also known as distortion), using equation (79), and the results of clustering are presented in Figure 32:

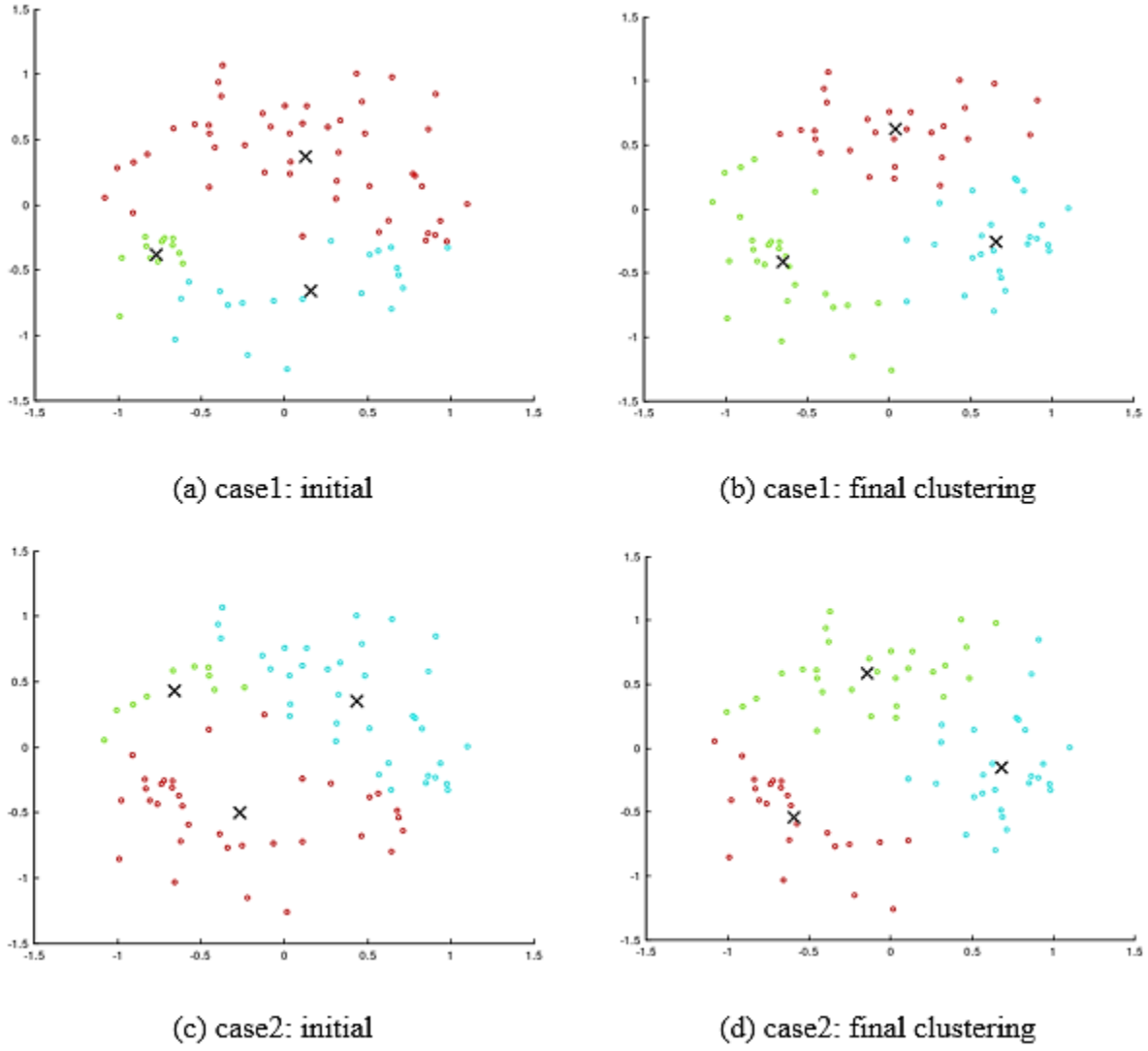
$$\min_{\substack{c^{(1)},\ldots,c^{(m)} \\ \mu_1,\ldots,\mu_K}} J\left(c^{(1)},\ldots,c^{(m)},\mu_1,\ldots,\mu_K\right) \tag{79}$$

Figure 32. Clustering piles into three groups after 10 iteration of K-means algorithm



5. It should be noted here that the minimum cost depends on the initial configuration of centroids. For different initial coordinates of centroids, the cost function was calculated and the minimum value of all different cost functions was chosen as the finial clustering of the piles (see Figure 33).

Figure 33. Effect of initial configuration on the final clustering



(a) case1: initial

(b) case1: final clustering

(c) case2: initial

(d) case2: final clustering

Clustering Pile-CPT methods

The K-means clustering was used for clustering LCPC, ERTC3, Probabilistic, UF, Philipponnat, De Ruiter, CPT2000, UWA, and Schmertmann pile-CPT methods, as shown in Figure 29. The number of clusters (K) was chosen to be 3.

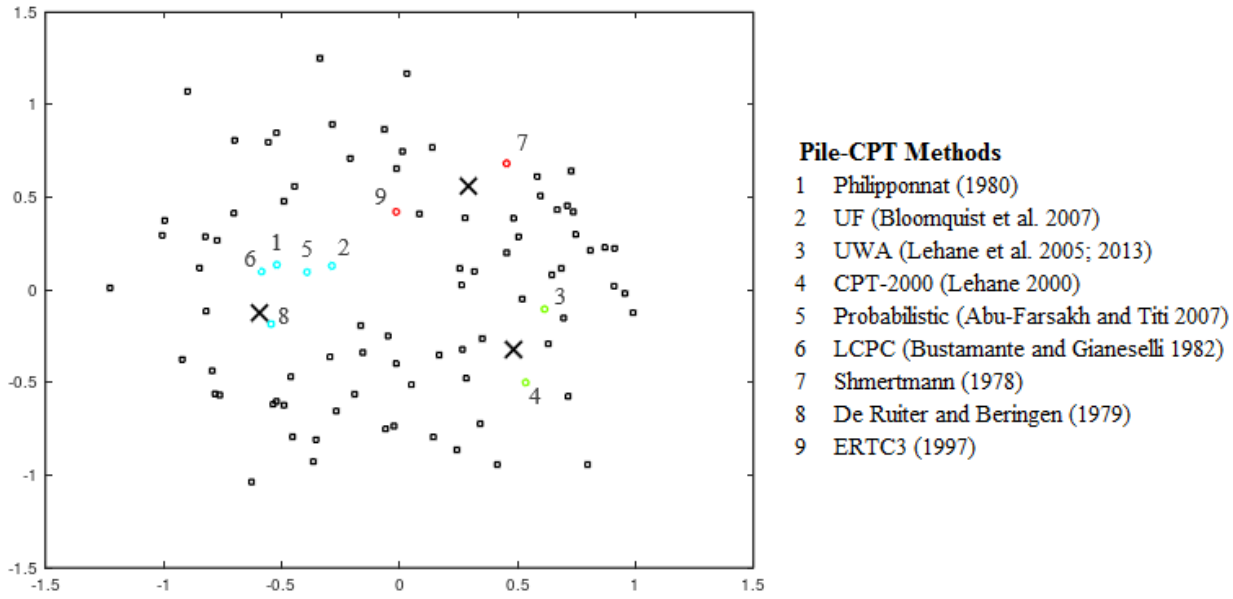
The optimization objective was set to be:

$$J\left(c^{(1)},\dots,c^{(m)},d^{(1)},\dots,d^{(n)},\mu_1,\dots,\mu_K\right)=\frac{1}{m}\sum_{i=1}^{m}\left|x^{(i)}-\mu_{c(i)}\right|^2+\frac{1}{n}\sum_{i=1}^{n}\left|y^{(i)}-\mu_{c(i)}\right|^2 \qquad (80)$$

$$\min_{\substack{c^{(1)},\ldots,c^{(m)} \\ d^{(1)},\ldots,d^{(n)} \\ \mu_1,\ldots,\mu_K}} J\left(c^{(1)},\ldots,c^{(m)},\mu_1,\ldots,\mu_K\right) \tag{81}$$

The final results for the different number of clusters are shown in Figure 34.

Figure 34. Clustering pile-CPT methods into three groups (J=0.36963)



Based on K-means analysis as shown in Figure 34, the following clustering of pile-CPT methods were obtained as follows:

Cluster 1: Philipponnat, UF, Probabilistic, LCPC, and De Ruiter

Cluster 2: Schmertmann and ERTC3

Cluster 3: CPT2000 and UWA

If geotechnical engineers are interested to get a better range of estimating the ultimate pile capacity by selecting three different pile-CPT methods, it is recommended to select one method from each cluster group.

Develop Combined Pile-CPT Methods

The evaluation of performance of pile-CPT methods showed that nine methods have the most performance in predictions the ultimate axial capacity of the piles. However, due to similarity

between Philipponnat and UF, and between CPT2000 and UWA, only UF and CPT2000 will be considered in developing the combined pile-CPT methods. Accordingly, the following seven pile-CPT methods will be considered here: Schmertmann, De Ruiter and Beringen, Bustamante and Gianeselli (LCPC), ERTC3 (European Regional Technical Committee 3), UWA, probabilistic, and UF methods. In this section, the piles are categorized based on the percentage of pile capacity obtained in sand layers and the log-normal distribution nature of the ratio of Q_p/Q_m is used to find a relationship for pile capacity based on the values of pile capacity obtained by different methods in different categories. This new method uses the combination of pile-CPT methods to overcome the concern of overprediction and high variability in sandy layers and helps engineers to have a tool for estimating the pile capacity in a more acceptable range.

Log-normal Distribution of Pile-CPT methods

The ratio of Q_p/Q_m for the investigated piles has a non-symmetrical distribution around the mean value. The log-normal distribution of the Q_p/Q_m ratio can be used for measuring the range distribution of pile capacity for different pile design methods. The following density function defines the log-normal distribution:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_{ln}x}\exp\left[-\frac{1}{2}\left(\frac{\ln(x)-\mu_{ln}}{\sigma_{ln}}\right)^2\right] \quad (82)$$

where. $x = Q_p/Q_m$; and μ_{ln} and σ_{ln} are the mean and standard deviation of $\ln(Q_p/Q_m)$, respectively. Figure 35 shows the histogram and log-normal probability distribution of the seven pile-CPT methods, in addition to the static analysis method.

Evaluation of Pile-CPT Methods based on Pile Category

In previous sections, it was shown that the pile-CPT methods are able to estimate the capacity of the piles in a reasonable range. Comparing the pile-CPT methods to static method showed that using CPT data for estimating the pile capacity is an efficient solution and more accurate predictions that enable engineers to have a better estimation for choosing the pile length and width. It was shown by researchers that ability of static analysis method for predicting the pile capacity for piles driven in sandy soils decreases considerably. Most of the time, static analysis method overpredicts the pile capacity of such piles. Based on the soil borings, the static analysis method uses the Nordlund method for estimating the tip and side capacity for sandy soil layers and α-method for clayey soil layers. In this study, the authors attempt to evaluate the performance of pile-CPT methods based on contributions of sand and clay layers to the ultimate

pile capacity. The percentage contribution of pile capacity in sandy layers to the total ultimate capacity of the entire piles was calculated and categorized into four groups. Table 11 summarizes the sand contribution for the pile load tests collected in our database.

Figure 35. Histogram and log-normal distribution of Q_p/Q_m for the different pile-CPT methods

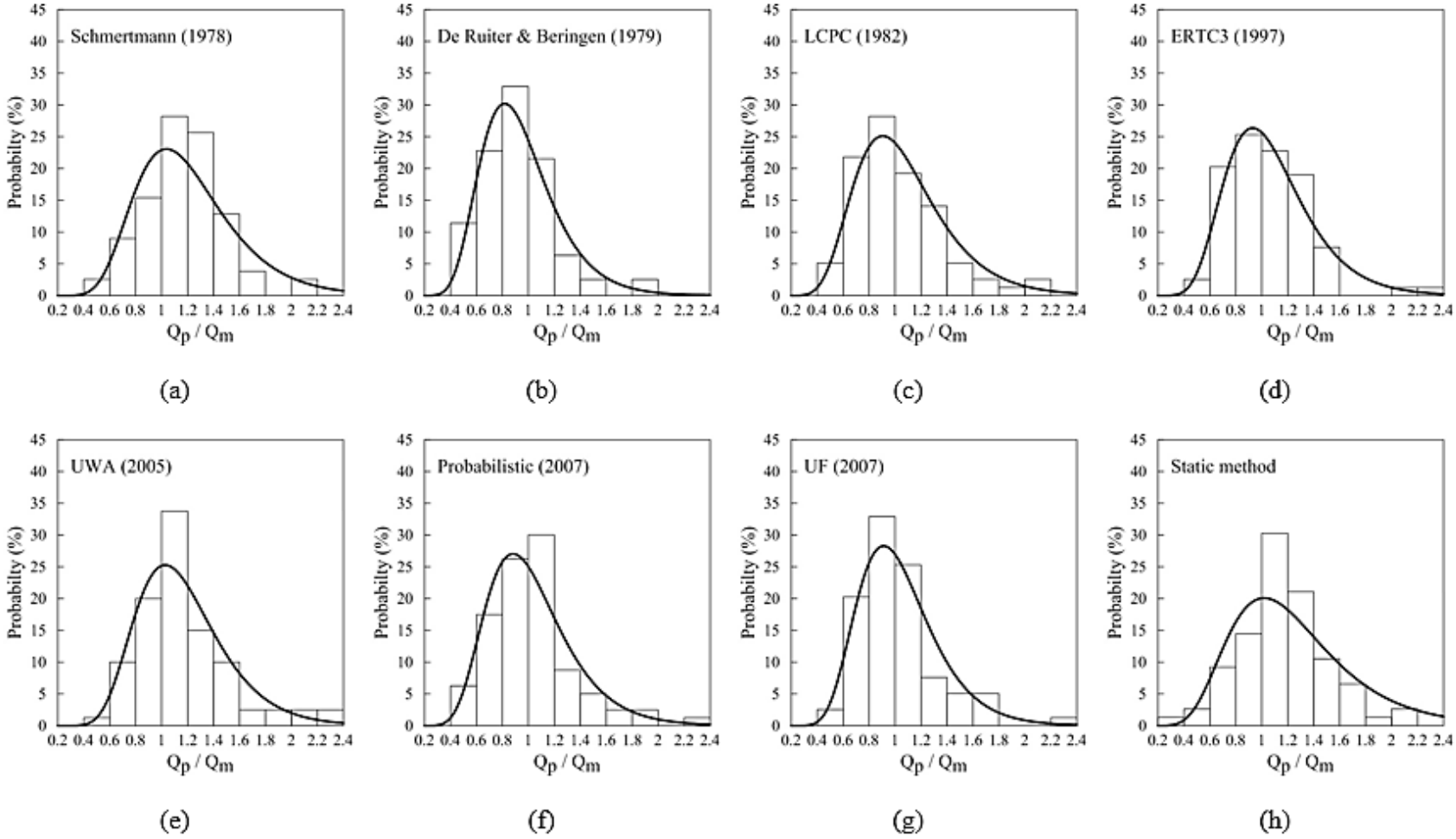


Table 11. Categories of the piles based on the percentage of contribution of sand layers to total ultimate capacity

Category	Sand contribution to the pile capacity (%)	Number of piles
1	0-25	43
2	25-50	9
3	50-75	17
4	75-100	11
All	0-100	80

The performance on each pile-CPT method is studied separately in each of five categories of piles based on sand contribution shown in Table 11. The first category included piles that had less than 25% of their capacity from the sand layers, while sand layers in the fifth category of piles contribute to more than 75% of the pile capacity. For static and pile-CPT methods arithmetic mean, μ and standard deviation, σ for each category were calculated separately. Also, the values of mean and standard deviation of $\ln(Q_p/Q_m)$ were calculated, which were used to identify the log-normal distribution of the density function in equation (82). Based on this distribution, 20% accuracy level was calculated that represents the probability of estimating pile capacity in the range of $0.8Q_m$ to $1.2Q_m$. Moreover, the values of $(Q_p/Q_m)_{min}$ and $(Q_p/Q_m)_{max}$ based on 95% confidential interval were determined. Assuming a normal distribution for $\ln(Q_p/Q_m)$, it can be discussed that with 95% confidence the values of $\ln(Q_p/Q_m)$ are located within $\mu_{ln} \pm 1.96\sigma_{ln}$. Hence, $(Q_p/Q_m)_{min}$ and $(Q_p/Q_m)_{max}$ are defined as:

$$(Q_p/Q_m)_{min} = \exp[\mu_{ln} - 1.96\sigma_{ln}] \tag{83}$$

$$(Q_p/Q_m)_{max} = \exp[\mu_{ln} + 1.96\sigma_{ln}] \tag{84}$$

Figure 36 presents the range values for our database categorized based on the contribution of sand layer to total pile capacity.

Figure 36. Evaluation of static and pile-CPT methods in different pile categories of Table 11

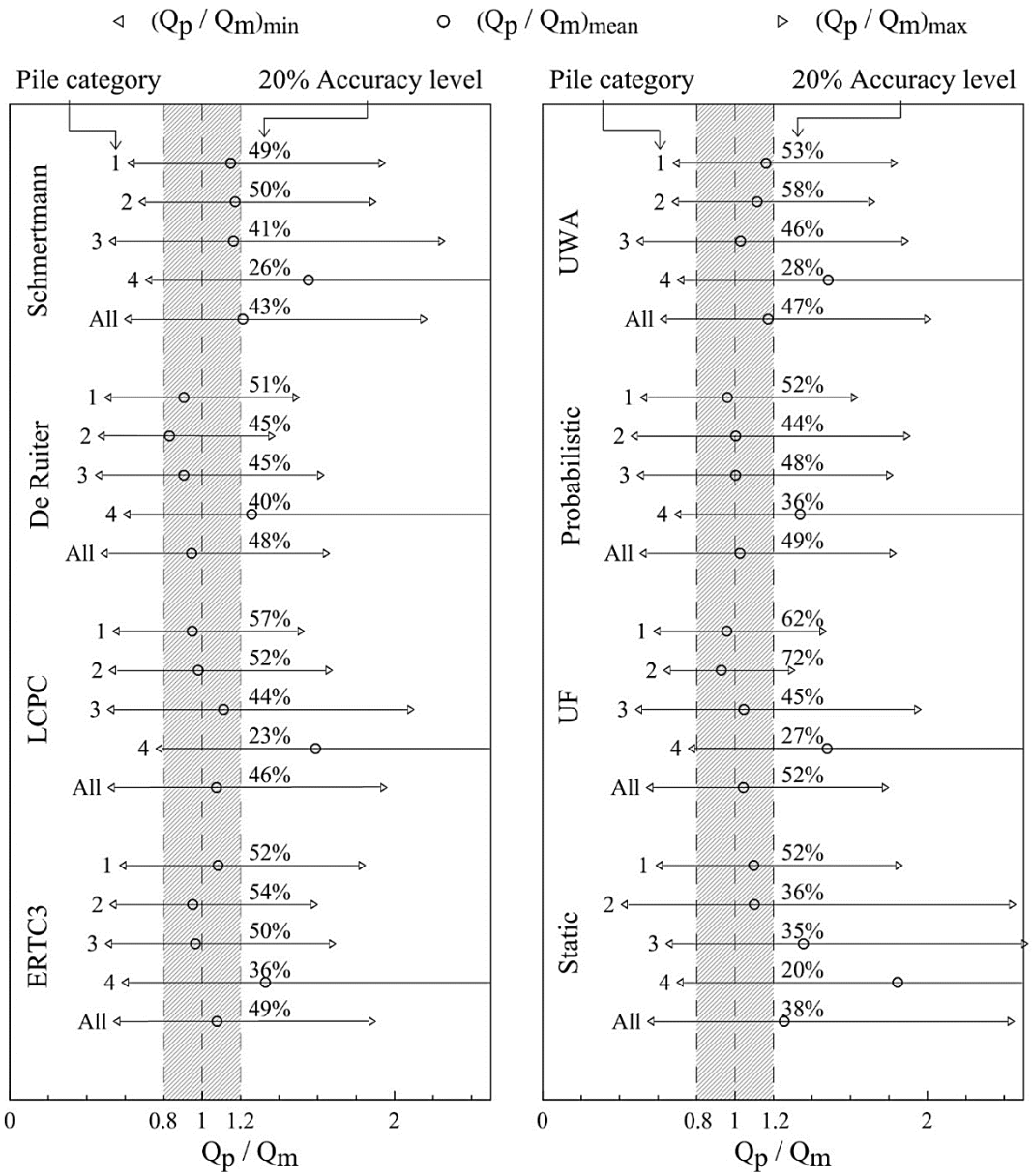


Figure 36 was determined by calculating arithmetic mean, μ and standard deviation, σ for each category, separately. For category 1, which is related to piles in soils where the contribution of sand layer to the total pile capacity was less than 25% (piles in clayey soils), De Ruiter, LCPC, probabilistic, and UF methods underpredicted, while Schmertmann, ERTC3, UWA, and static methods overpredicted the pile capacity. Schmertmann method with $\mu = 1.15$ and $\sigma = 0.34$ and UWA method with $\mu = 1.16$ and $\sigma = 0.30$ resulted in less accurate estimations for the measured pile capacity than other methods for this category. For category 4, where the contribution of sand layers to the total pile capacity was more than 75%, all methods overpredicted the measured pile capacity, and standard deviation, σ, was considerably higher than other categories. In this

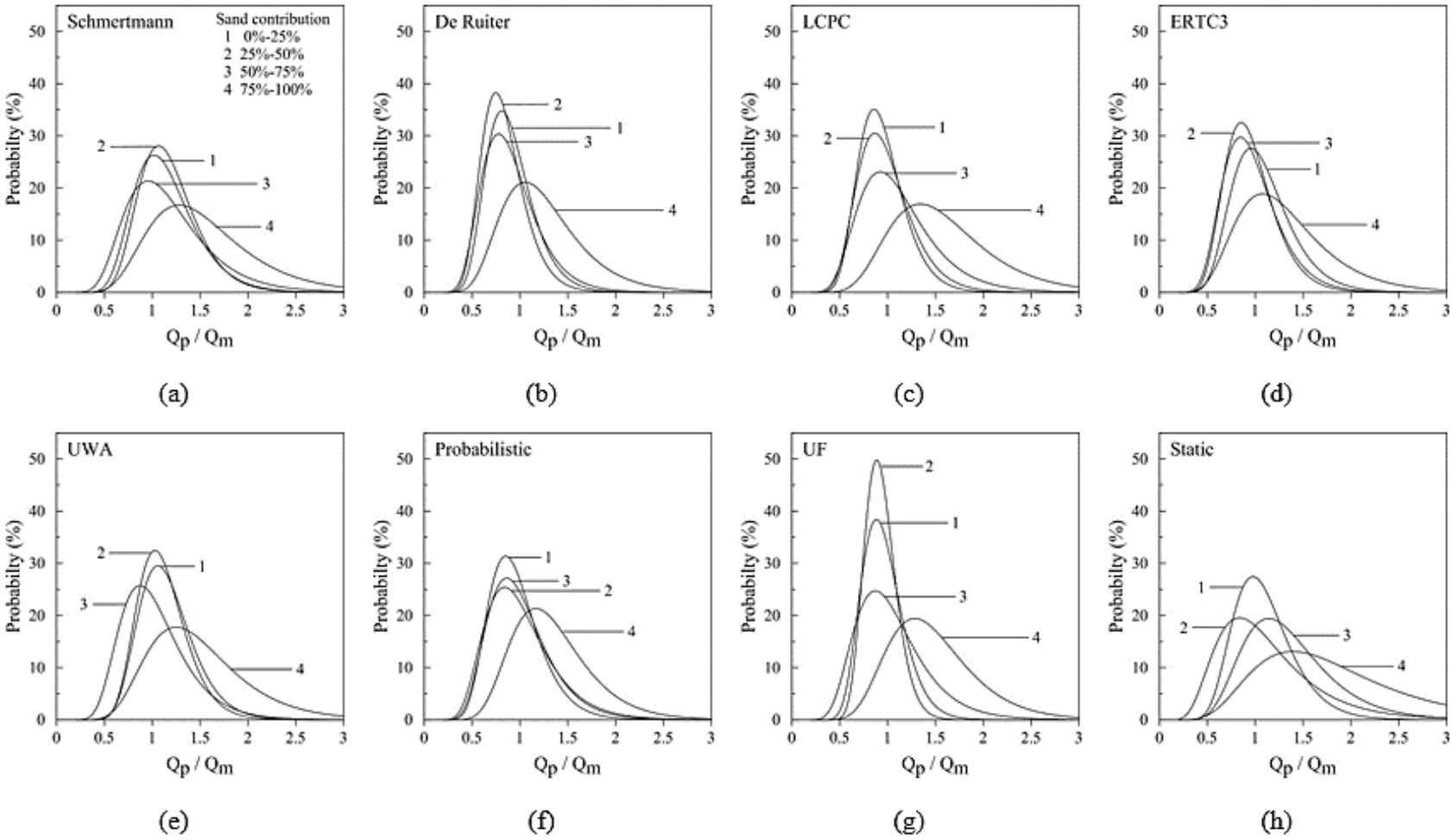
category, the arithmetic mean for De Ruiter method was the closest to one, which indicates more accuracy for this method for piles in category 4 than the other methods. The standard deviation for probabilistic method was less than the other methods in this category, indicating more precision for probabilistic method than the other methods.

The values of mean and standard deviation of $\ln(Q_p/Q_m)$ categorized based on the contribution of sand layer to total pile capacity (as defined in Table 11) were calculated to obtain log-normal distribution, which was used to calculate the 20% accuracy level as shown in Figure 36. For category 1, which was related to piles in soils where the contribution of sand layer to the total pile capacity was less than 25%, the probability of estimating the pile capacity using UF method within $0.8Q_m$ to $1.2Q_m$ was 62.48%, showing that UF method is the best method for piles in category 1. On the other hand, as shown in Figure 36, for piles in category 4, the accuracy level decreased considerably and the highest probability for estimating the pile capacity within $0.8Q_m$ to $1.2Q_m$ in category 4 was 40.21% obtained by De Ruiter method.

Figure 36 shows the values of $(Q_p/Q_m)_{min}$ and $(Q_p/Q_m)_{max}$. Value for $(Q_p/Q_m)_{min}$ more than 0.5 and values for $(Q_p/Q_m)_{max}$ less than 2 show that within 95% accuracy, the estimated pile capacity is more than half and lower than twice the measured pile capacity. For piles in category 1, UWA, Schmertmann, static, UF, and ERTC3 methods had the closest values of $(Q_p/Q_m)_{min}$ to 1, in order. On the other hand, the order of methods that had the closest values of $(Q_p/Q_m)_{max}$ to 1 was UF, De Ruiter, LCPC, and probabilistic. For piles in category 4, the order of the methods based on $(Q_p/Q_m)_{min}$ was LCPC, UF, Schmertmann, static, and UWA; while based on $(Q_p/Q_m)_{max}$ De Ruiter, probabilistic, UF, and ERTC3 methods showed the best performance in sequence.

Figure 37 presents the log-normal distributions of the seven pile-CPT methods in addition to static method based on each category and the static analysis method.

Figure 37. Log-normal distribution of (Q_p/Q_m) in different categories of Table 11



As shown in Figure 37, the log-normal distribution for each method is different in each category. It can be observed that static and pile-CPT methods overpredict the pile capacity in cases where the sand soil contribution in pile capacity is high. Also, increasing the sand contribution to the pile capacity causes more variability in the estimations of static and pile-CPT methods. In the next section, the properties of log-normal distribution will be discussed. Those properties will be used to combine pile-CPT methods and get a better estimation for the pile capacity based on the pile categories shown in Table 11.

Analytical calculations for log-normal distributions

It was shown that the value of (Q_p/Q_m) can be regarded as a log-normal distribution. Therefore, the value of $\ln(Q_p/Q_m)$ is a normal distribution of $(\mu_{ln}, \sigma_{ln}^2)$, where μ_{ln} and σ_{ln} are mean and standard deviation of the variable's natural logarithm, respectively. For a normal distributed variable, with 95% certainty, the values are located in $\mu_{ln} \pm 1.96\sigma_{ln}$, which gives us the values of $(Q_p/Q_m)_{max,min} = e^{\mu_{ln} \pm 1.96\sigma_{ln}}$. Also, it was shown that log-normal distribution of each method is different for each pile category. As shown in Figure 36 and Figure 37, by increasing the sand contribution, the accuracy of the methods decreases considerably.

Characteristics of log-normal distributions can be used for shifting the mean value to zero, decreasing the standard deviation, and combining the log-normal distributions.

For a normal distribution of $\ln(Q_p/Q_m)$, μ_{ln} can be shifted to zero by adding $\Delta = -\mu_{ln}$ to it, which causes the value of Q_p/Q_m to be changed to $e^{\Delta}(Q_p/Q_m)$. This change is shown in Figure 38 (a) and (b) where the mean of probability function number 1 has changed to zero in probability function number 2. In other words, if Q_p/Q_m is a log-normal distribution with μ_{ln} and σ_{ln} as its mean and standard deviation, $e^{-\mu_{ln}}(Q_p/Q_m)$ will be a log-normal distribution with 0 as its mean and no change in standard deviation.

Multiplying a normal distribution by λ causes the mean value to change from μ to $\lambda\mu$ and the standard deviation from σ to $\lambda\sigma$. For normal distribution of $\ln(Q_p/Q_m) + \Delta$, where the mean value is zero, multiplying by λ changes the standard deviation σ_{ln} to $\lambda\sigma_{ln}$, while the mean value remains zero. In other words, as shown in Figure 38 (c) and (d), $\lambda[\ln(Q_p/Q_m) + \Delta] = \ln([e^{\Delta}(Q_p/Q_m)]^{\lambda})$ is a normal distribution with mean and standard deviation equal to zero and $\lambda\sigma_{ln}$, respectively. Values of λ less than 1 cause the resultant normal distribution to have lower

standard deviation. Figure 38 (c) shows that probability function number 3, which is $\left[e^{\Delta}\left(Q_p/Q_m\right)\right]^{\lambda}$ is a log-normal distribution with $\mu = 0$ and standard deviation decreased to $\lambda\sigma_{ln}$.

Linear combination of normal distributions, X_i with means and standard deviations equal to μ_i and σ_i is a normal distribution, L, which has the following characteristics:

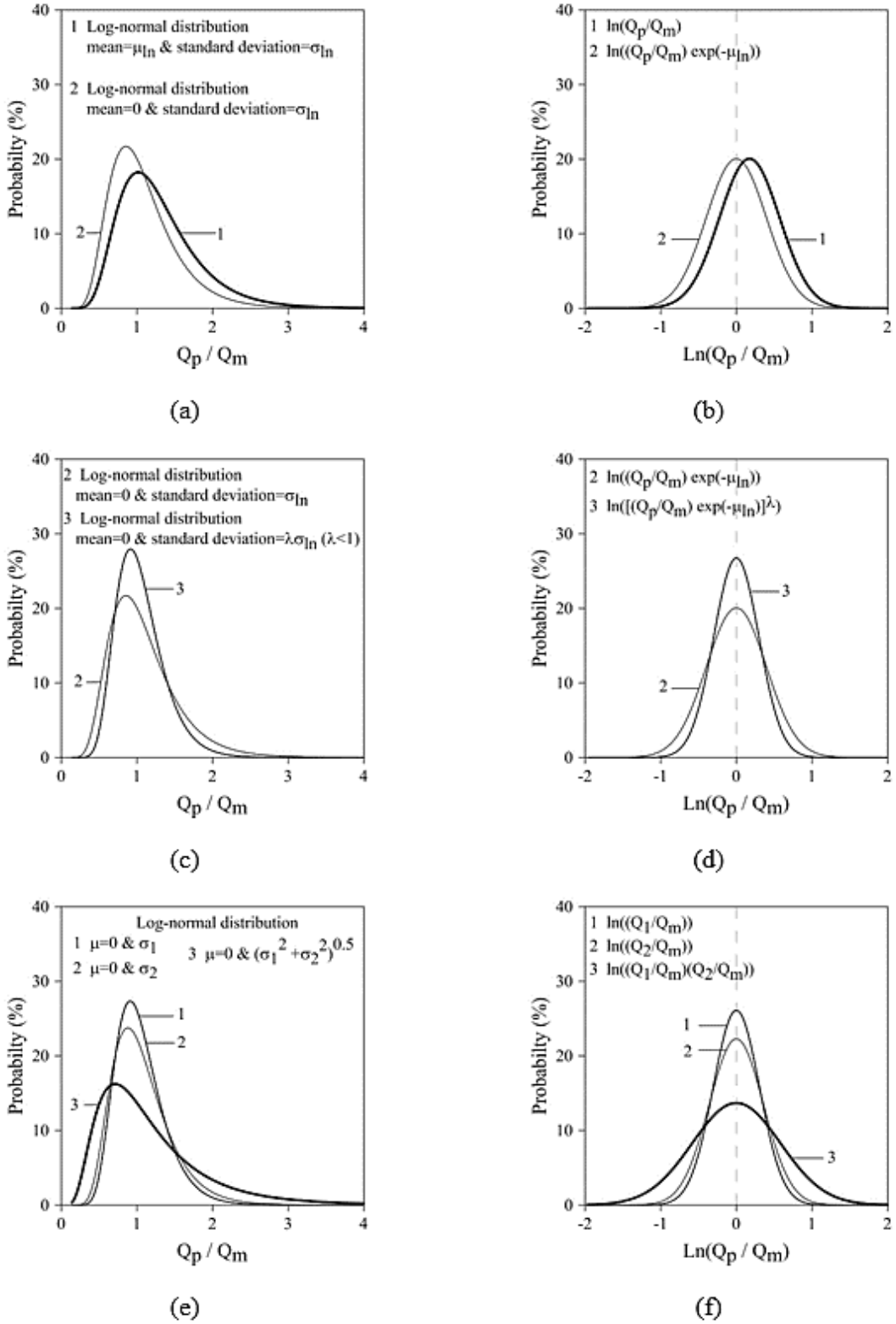
$$L = \sum c_i X_i \tag{85}$$

$$\mu_L = \sum c_i \mu_i \quad \& \quad \sigma_L^2 = \sum c_i^2 \sigma_i^2 \tag{86}$$

As shown in Figure 38 (f) combining two normal distributions of $\ln(X_1)$ and $\ln(X_2)$ with mean values equal to zero and standard deviations of σ_1 and σ_2, results in a normal distribution with $\mu = 0$ and $\sigma = ({\sigma_1}^2 + {\sigma_2}^2)^{0.5}$. Therefore, $\ln(X_1) + \ln(X_2) = \ln(X_1 X_2)$ is a normal distribution with mean and standard deviation equal to zero and $({\sigma_1}^2 + {\sigma_2}^2)^{0.5}$, respectively. In other words, $X_1 X_2$ is a log-normal distribution as shown by probability function number 3 in Figure 38 (e).

The summary of resultant changes in log-normal and normal distribution, due to adding $\Delta = -\mu_{ln}$, multiplying by λ, and linear combination of log-normal distributions is depicted in Figure 38.

Figure 38. Changes in log-normal and normal distribution by shifting the mean value to zero, decreasing the standard deviation, and combining the log-normal distributions



These properties will be used to shift the log-normal distributions to $\mu = 0$, decrease their standard deviations by multiplying by $\lambda < 1$, and finally adding them together somehow that the resultant standard deviation will be obtained.

Combined Pile-CPT Methods

Generally, by increasing the sand contribution in pile capacity, the pile-CPT methods tend to overpredict the pile capacity. In addition, the standard deviation of Q_p/Q_m for each method for piles in sandy soils is more than the other piles. It was shown in Figure 37 that for each category of piles presented in Table 11, the log-normal distribution of (Q_p/Q_m) for each pile-CPT method is different.

In previous section, it was shown that properties of log-normal distribution can be used to produce a log-normal distribution with desired properties. In this section, log-normal properties are used for combining predictions of pile-CPT methods of ultimate capacity of piles in each category of Table 11. In this section, combining predictions of pile-CPT methods for a pile in category 4 of Table 11 is explained. The same procedure can be done for piles in other categories.

As shown in Figure 36, the proportion of pile capacity predicted by Schmertmann Q_1, De Ruiter Q_2, LCPC Q_3, ERTC3 Q_4, UWA Q_5, probabilistic Q_6, and UF Q_7 methods to the measured capacity Q_m in category 4 (piles with 75% to 100% capacity due to sand layers) yields normal distributions for $\ln(Q_1/Q_m)$, $\ln(Q_2/Q_m)$,…, and $\ln(Q_7/Q_m)$ with means, μ_i equal to 0.37, 0.17, 0.41, 0.21, 0.34, 0.25, and 0.34 , respectively. The corresponding standard deviations, σ_i equal to 0.36, 0.34, 0.34, 0.37, 0.34, 0.31, and 0.31, respectively. As shown in Figure 38, adding $e^{-\mu_i}$ and multiplying by λ_i, makes the distributions of $\ln\left([e^{-0.37}(Q_1/Q_m)]^{\lambda_1}\right)$, $\ln\left([e^{-0.17}(Q_2/Q_m)]^{\lambda_2}\right)$,…, and $\ln\left([e^{-0.34}(Q_3/Q_m)]^{\lambda_7}\right)$ normal distributions with means equal to zero and standard deviations equal to $0.36\lambda_1$, $0.34\lambda_2$,…, and $0.31\lambda_7$, respectively.

The linear combination of these normal distributions is:

$$\begin{aligned}\ln\left([e^{-0.37}(Q_1/Q_m)]^{\lambda_1}\right) + \ln\left([e^{-0.17}(Q_2/Q_m)]^{\lambda_2}\right) + \cdots + \ln\left([e^{-0.34}(Q_3/Q_m)]^{\lambda_7}\right) \\ = \ln\left([e^{-0.40}(Q_1/Q_m)]^{\lambda_1} \times [e^{-0.21}(Q_2/Q_m)]^{\lambda_2} \times \ldots \right. \\ \left. \times [e^{-0.36}(Q_3/Q_m)]^{\lambda_7}\right)\end{aligned} \quad (87)$$

The standard deviation for the above normal distribution is:

$$\sigma_L = [(0.36\lambda_1)^2 + (0.34\lambda_2)^2 + \cdots + (0.31\lambda_7)^2]^{0.5} \quad (88)$$

Having the value of $\left(Q_p/Q_m\right)_{max,min} = e^{\mu_{ln} \pm 1.96\sigma_{ln}}$ equal to 0.5 and 2, respectively, means that Q_p is in the range of $[0.5Q_m, 2Q_m]$ with 95% confidence. This means that the value of σ_L should be equal to ${(\ln 2)}/{1.96} = 0.354$.

Equation (88) has seven unknown variables λ_1 to λ_7. For obtaining these values additional conditions have to be considered. For increasing the effect of normal distributions with less variance, the values of λ_i have been regarded as:

$$\lambda_i = k \times \left(\frac{1}{\sigma_i}\right)^2 \quad (89)$$

which results in $\lambda_1 = {k}/{0.36^2}$, $\lambda_2 = {k}/{0.34^2}$,…, and $\lambda_7 = {k}/{0.31^2}$.

Solving equation (88) for these values led to finding the values of λ_1 to λ_7 equal to 0.350, 0.389, 0.391, 0.329, 0.379, 0.476, and 0.474, respectively.

Substituting the values of λ_i in equation (87), they can be used to obtain the value of pile capacity, Q_u:

$$[e^{-\mu_1}(Q_1/Q_u)]^{\lambda_1} \times [e^{-\mu_2}(Q_2/Q_u)]^{\lambda_2} \times \ldots \times [e^{-\mu_7}(Q_7/Q_u)]^{\lambda_7} = 1 \quad (90)$$

Finally, the general equation for obtaining the value of pile capacity, Q_u, is given as:

$$Q_u = B\left[{Q_1}^{A_1} \times {Q_2}^{A_2} \times \ldots \times {Q_7}^{A_7}\right] \quad (91)$$

where, $B = \exp(\sum -\mu_i\lambda_i / \sum \lambda_i)$ and $A_i = {\lambda_i}/{\sum \lambda_i}$

The values of B and A_i for piles in category 4 were obtained as 0.741, 0.126, 0.140, 0.140, 0.118, 0.136, 0.171, and 0.170.

Knowing the value of $\sum A_i = 1$ made it possible to normalize equation (91) as shown in following equation:

$$\frac{Q_u}{Q_{avg}} = B\left(\frac{Q_1}{Q_{avg}}\right)^{A_1}\left(\frac{Q_2}{Q_{avg}}\right)^{A_2}\ldots\left(\frac{Q_7}{Q_{avg}}\right)^{A_7} \quad (92)$$

where, Q_{avg} is the average value of pile capacity from different methods.

Using the mentioned procedure, the values for B and A_i were obtained for all the categories in Table 11, where i=1 to 7 is related to Schmertmann, De Ruiter, LCPC, ERTC3, UWA, probabilistic, and UF methods, respectively. These values are shown in Table 12.

Table 12. Combined pile-CPT method parameters

Category	B	A_1	A_2	A_3	A_4	A_5	A_6	A_7
1	1.019	0.122	0.134	0.152	0.118	0.163	0.121	0.190
2	1.042	0.134	0.124	0.106	0.117	0.165	0.072	0.282
3	1.033	0.119	0.155	0.128	0.173	0.141	0.152	0.131
4	0.741	0.126	0.140	0.140	0.118	0.136	0.171	0.170
All	0.978	0.135	0.146	0.124	0.143	0.159	0.136	0.158

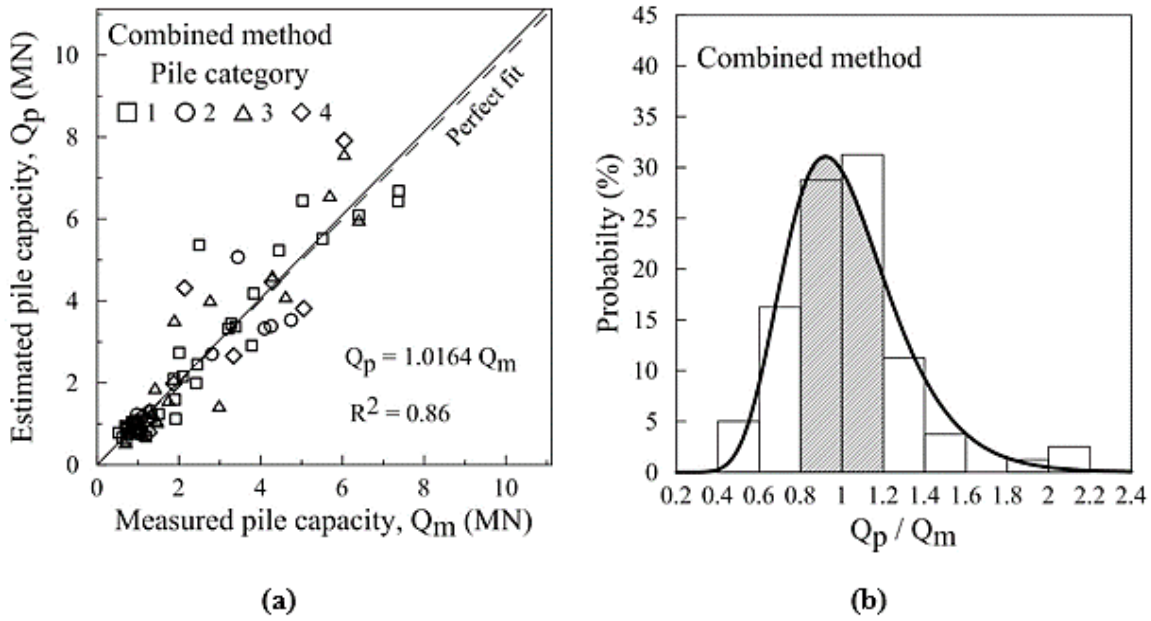
The proposed combined pile-CPT methods can be illustrated as follows:

1. Use the pile-CPT methods: Schmertmann, De Ruiter, LCPC, ERTC3, UWA, probabilistic, and UF to obtain pile capacities as Q_1, Q_2,…, and Q_7, respectively.
2. Determine the percentage of pile capacity in sand layers and categorize the pile based on this value in Table 11.
3. Find the values of B, A_1, A_2,…, and A_7 constants in Table 12 based on the pile category. The category of "All" is used, in case of unknown category for the pile.
4. Use equation (92) to evaluate the ultimate pile capacity, Q_u

Application of Combined Pile-CPT Method for Louisiana Pile Database

The procedure mentioned in the previous section was used to obtain the results of pile capacity for the piles in the Louisiana database using the combined pile-CPT method. Figure 39 shows the results of pile predictions obtained by this method and the comparison with the measured values.

Figure 39. (a) Estimated versus measured pile capacity (b) histogram and log-normal distribution of (Q_p/Q_m) for combined pile-CPT method after categorizing piles based on sand contribution

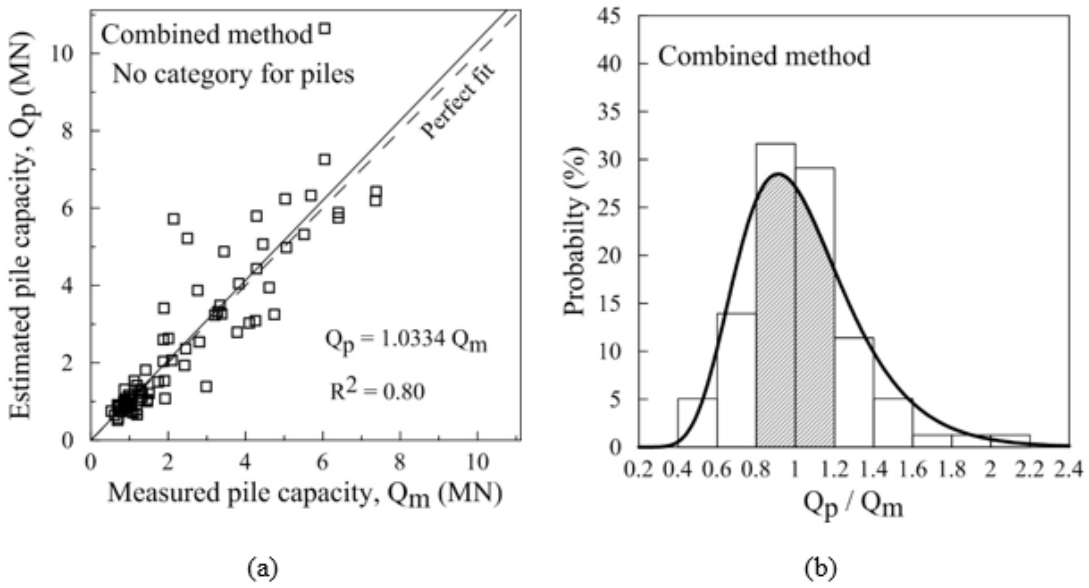


As shown in Figure 39 (a), the slope of the best-fit line for the combined method is 1.01. Moreover, the coefficient of determination, R^2 for combined method is 0.86, which is higher than static and pile-CPT methods. The values of μ_{ln} and σ_{ln} are -0.01 and 0.27, which are used to obtain the log-normal distribution for the combined method as shown in Figure 39 (b). The 20% accuracy level obtained from log-normal distribution is 55.98%, which shows a significant increase compared to accuracy levels of other methods. The above results indicate the accuracy and precision of the obtained pile capacities increases by combining pile-CPT methods.

For the case of using combined method without categorizing the piles, the obtained results are as shown in Figure 40.

Based on the results shown in Figure 40, a similar increase in accuracy and precision of predictions by combining pile-CPT methods is obtained. However, the results show that categorizing piles before combining them leads to more accuracy in pile prediction. The slope of the best-fit line in Figure 40 (a) is 1.03 with coefficient of determination equal to 0.8. The values of μ_{ln} and σ_{ln} equal to 0.0 and 0.29 were obtained, which are used for showing the log-normal distribution in Figure 40 (b), which has the 20% accuracy level of 51.94%.

Figure 40. (a) Estimated vs. measured pile capacity (b) histogram and log-normal distribution of (Q_p/Q_m) for combined pile-CPT method without categorizing piles



Ranking of different pile-CPT methods can be determined from determining the efficiency of them based on reliability analysis from LRFD. The results are presented in Table 13.

Based on the results of efficiency of the methods in Table 13, combining the pile-CPT methods with or without categorizing piles based on sand contribution shows improvement in predictions for pile capacity. The main advantage of categorizing piles and then combining them based on the factors in Table 12 can be seen in Figure 41, where log-normal distributions for piles in different categories of piles (Table 11) is shown, separately.

Table 13. Evaluating the performance of different pile-CPT methods based on LRFD reliability analysis

Pile capacity method	bias, λ_R	σ	COV	ϕ	ϕ/λ_R
Schmertmann	0.92	0.31	0.34	0.48	0.53
De Ruiter	1.16	0.34	0.29	0.69	0.59
LCPC	1.04	0.32	0.31	0.60	0.57
ERTC3	1.02	0.32	0.31	0.59	0.57
UWA	0.93	0.28	0.30	0.54	0.58
Probabilistic	1.08	0.37	0.34	0.57	0.53
UF	1.05	0.29	0.27	0.65	0.62
Combined method (with categorizing piles)	**1.05**	**0.28**	**0.27**	**0.65**	**0.62**
Combined method (without categorizing piles)	**1.05**	**0.31**	**0.29**	**0.62**	**0.60**

Figure 41. Log-normal distribution of Q_p/Q_m for the combined pile-CPT method (a) with categorizing piles (b) without categorizing piles based on sand contribution to the pile capacity as shown in Table 11

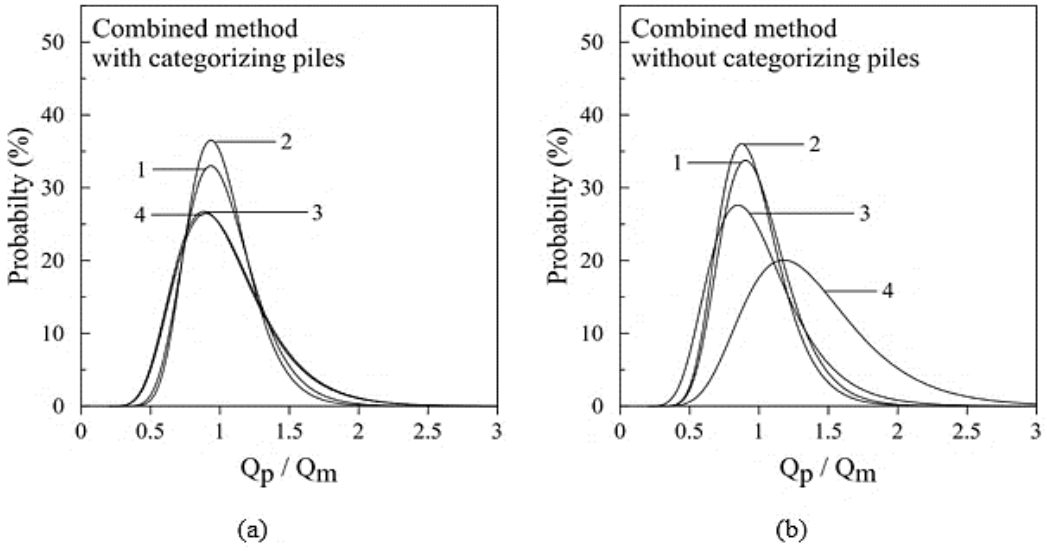


Figure 41 shows that combined pile-CPT method is a useful technique for increasing the accuracy of estimations for ultimate axial capacity of the piles. Using the combined pile-CPT method with categorizing piles shows significant improvement, especially in estimation for piles in sandy soils. Comparing Figure 41 (a) and (b) shows that, for piles that have less than 75% of their capacity from sand layers, no need for categorizing piles is needed.

Studying the combined pile-CPT method for different pile databases is recommended for evaluating its performance in an unbiased manner. The same procedure described in this paper can be used for combining different pile-CPT methods. Therefore, interested researchers can add more pile-CPT methods to the methods used in this study and obtain reliable values for ultimate axial capacity of piles.

Develop Machine Learning Pile-CPT Models

As stated earlier, the Artificial Neural Network (ANN) and three tree-based ML techniques [the decision tree (DT), random forest (RF), and gradient boosted tree (GBT)] were used to develop ML models to estimate the ultimate pile capacity from CPT data. The ML input parameters include two pile parameters (pile embedment depth, L_e, and pile width, B), 10 CPT data for the five equal segments along pile shaft ($q_{t,\ avg\ 1}$ to $q_{t,\ avg\ 5}$, and $f_{s,\ avg\ 1}$ to $f_{s,\ avg\ 5}$), and 2 CPT data for the influence zone around pile tip [$q_{t\text{-}tip,\ 4B\ below}$ and ($q_{t\text{-}tip,\ 4B\ above}$ or $q_{t\text{-}tip,\ 8B\ above}$)], as shown in Figure 12. The results and analyses of these ML models will be presented in the following sections.

ANN Models

Numerous ANN models were tried using different numbers of hidden layers, different numbers of nodes per hidden layer, and different input parameters (see Table 1 for different model types) to determine the best-performed ANN model(s). The resulted best-performed ANN models in terms of estimating the measured ultimate pile capacity (for testing training, and validation) are summarized in Table 14. As discussed earlier, the performance of ANN models was evaluated based on the following criteria: coefficient of correlation, r, coefficient of determination, R^2, root mean squared error, RMSE, mean bias factor, λ, and the coefficient of variance, COV. The ANN models in Table 14 are designated based on their network structure. For example, the first and last number for model 9-4-1-1 represent the number of nodes in the input (9 parameters) and output (1 parameter), respectively. The intermediate numbers represent two hidden layers with 4 and 1 nodes, respectively.

Table 14 presents the top three best-performed ANN models obtained for each input type in Table 1 from hundreds of trial ANN models. The results show that the Type 4 ANN model 9-7-

7-1, which uses f_s to simulate the side capacity, and an influence zone of 4B below to 8B above pile toe for end bearing capacity, can be considered as the best ANN model in estimating the measured ultimate pile capacity based on validation phase. However, based on both testing and validation phases, the Type 5 ANN model 14-9-3-1, which considers both q_t and f_s to simulate the side capacity and an influence zone of 4B below to 4B above pile toe for end bearing capacity, outstands all the ANN models with the most stable and best-performed model in estimating the ultimate pile capacity. The comparison between the predicted and measured ultimate pile capacity for training, testing, and validation phases for ANN Type 4 model 9-7-7-1 and ANN Type 5 model 14-9-3-1 are presented in Figure 42a and Figure 42b, respectively. In general, the results in Table 14 show that using the combination of $q_{t,\,avg}$, and $f_{s,\,avg}$, to evaluate the pile's side capacity yields better ANN prediction models.

Figure 42. Predicted versus measured ultimate pile capacity for training, testing and validation phases: (a) ANN Type 4 model 9-7-7-1, and (b) ANN Type 5 model 14-9-3-1

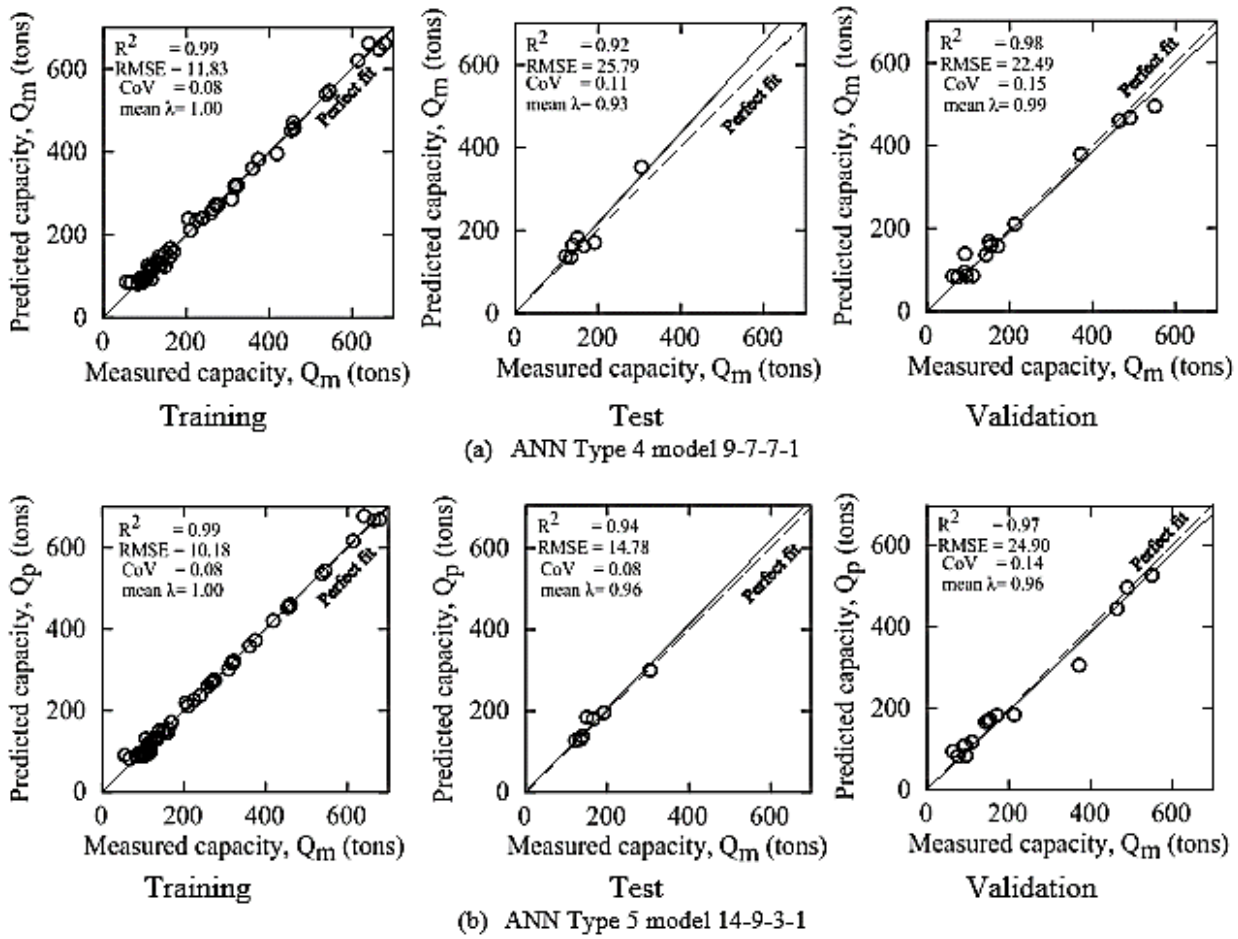


Table 14. Best-performed ANN models

ANN type	ANN model	Phase	r	R^2	RMSE	Mean λ	COV
Type 1	9-7-1	Training	0.99	0.97	28.45	1.00	0.18
		Testing	0.98	0.90	52.09	0.94	0.20
		validation	0.98	0.93	25.39	0.93	0.17
	9-6-1-1	Training	0.97	0.93	44.13	0.99	0.27
		Testing	0.89	0.75	48.34	1.10	0.28
		validation	0.96	0.92	29.85	0.93	0.23
	9-9-4-1	Training	0.99	0.99	10.24	0.99	0.08
		Testing	0.95	0.84	37.12	1.11	0.19
		validation	0.96	0.93	33.06	1.11	0.26
Type 2	9-4-1-1	Training	0.97	0.93	44.40	1.00	0.26
		Testing	0.90	0.78	41.82	1.11	0.27
		validation	0.95	0.87	37.06	0.91	0.27
	9-7-5-1	Training	0.98	0.97	27.42	1.00	0.18
		Testing	0.95	0.90	35.46	1.14	0.23
		validation	0.97	0.91	29.40	0.97	0.20
	9-8-1-1	Training	0.98	0.97	29.26	1.02	0.20
		Testing	0.96	0.91	45.66	1.17	0.15
		validation	0.98	0.95	33.84	0.90	0.23
Type 3	9-7-7-1	Training	0.99	0.99	11.65	1.00	0.08
		Testing	0.94	0.68	33.23	1.06	0.19
		validation	0.99	0.98	26.78	0.99	0.19
	9-7-5-1	Training	0.99	0.99	13.44	0.99	0.10
		Testing	0.99	0.98	28.80	0.98	0.17
		validation	0.98	0.96	32.62	0.95	0.18
	9-9-9-1	Training	0.99	0.99	8.76	1.00	0.07
		Testing	0.91	0.81	31.34	1.12	0.23
		validation	0.98	0.95	37.66	1.01	0.22
Type 4	9-7-7-1	Training	0.99	0.99	11.83	1.00	0.08
		Testing	0.96	0.92	25.79	0.93	0.11
		validation	0.99	0.98	22.49	0.99	0.15
	9-6-5-1	Training	0.99	0.99	18.05	0.99	0.13
		Testing	0.93	0.79	55.28	1.03	0.23
		validation	0.97	0.95	37.31	0.95	0.19
	9-17-1	Training	0.99	0.99	13.43	1.00	0.08
		Testing	0.95	0.77	45.39	1.20	0.23
		validation	0.99	0.98	33.46	1.08	0.17
Type 5	14-9-3-1	Training	0.99	0.99	10.18	1.00	0.08
		Testing	0.98	0.94	14.78	0.96	0.08
		validation	0.99	0.97	24.90	0.96	0.14
	14-7-1	Training	0.99	0.98	22.98	0.99	0.12
		Testing	0.90	0.82	30.28	1.10	0.19
		validation	0.98	0.93	38.08	0.93	0.20
	14-10-8-1	Training	0.99	0.99	7.44	1.00	0.06
		Testing	0.96	0.93	20.79	0.97	0.19
		validation	0.97	0.93	33.58	0.99	0.23
Type 6	14-9-4-1	Training	0.99	0.99	7.17	1.00	0.07
		Testing	0.97	0.94	30.27	0.96	0.23
		validation	0.99	0.98	29.65	0.97	0.14
	14-20-1	Training	0.99	0.99	7.57	0.99	0.07
		Testing	0.99	0.98	36.07	1.01	0.15
		validation	0.98	0.95	24.31	0.93	0.16
	14-25-1	Training	0.99	0.99	9.30	0.99	0.08
		Testing	0.84	0.67	49.16	1.16	0.27
		validation	0.98	0.97	0.39	1.03	0.22

Sensitivity Analyses of Input Parameters in ANN Models

Sensitivity analyses were conducted to investigate the effect of input parameters for the ANN Type 5 model 14-9-3-1 on the output ultimate pile capacity. Garson evaluated the relative importance of each parameter for a single-layered ANN model as follow [100]:

$$\text{Relative importance of a certain variable} = \frac{\sum_{j}^{n_h}\left(\frac{w_{ji}^{l} w_{yj}^{O}}{\sum_{1}^{n_v} w_{ji}^{l}}\right)}{\sum_{1}^{n_v}\left(\sum_{j}^{n_h}\left(\frac{w_{ji}^{l} w_{yj}^{O}}{\sum_{1}^{n_v} w_{ji}^{l}}\right)\right)} \quad (93)$$

where, w_{ji}^{l} = weighted connection between the ith and jth nodes, , w_{yj}^{O} = weighted connection between the jth node and the output layer, n_h = number of hidden nodes, n_v = number of variables. It is well known that both the pile embedment length, L_e, and pile width, B, are directly affecting the pile capacity. Therefore, the sensitivity analyses here will be focused on the relative importance of CPT input parameters. The results of sensitivity analysis are presented in Table 15, which shows that, apart from L_e and B, the average corrected cone tip resistance below pile toe, $q_{t\text{-}tip,\ 4B\ below}$, has the highest value of relative importance among the CPT input parameters. Meanwhile, the average sleeve friction along pile shaft, $f_{s\text{-}avg}$, has higher importance than $q_{t\text{-}avg}$. However, the results of sensitivity analysis (i.e., percent of relative importance) demonstrated that the four CPT input parameters are important in estimating the ultimate pile capacity using the ANN models.

Table 15. Relative importance of ANN input parameters

ANN Input Variables	The relative importance of input variables (%)
Embedment length of pile, L_e	14.8
Width of pile, B	14.1
$q_{t\text{-}tip,\ 4B\ above}$	19.0
$q_{t\text{-}tip,\ 4B\ below}$	22.8
$q_{t\text{-}avg}$ along the pile shaft	12.9
$f_{s\text{-}avg}$ along the pile shaft	16.4

Tree-based ML Models

Numerous tree-based ML models using DT, RF, and GBT techniques were first tried using the randomly 80% training subset data and different architectures to determine the best-performed models for estimating the measured ultimate pile capacity from CPT data. After locating the

optimal models by random search, a final evaluation is performed using the 20% subset testing data to ensure that these models can be robustly generalized. This implies that the ML models must not overfit the training data and have minimal bias and variance. Specifically, the optimal ML models should have low error rates on both the training and testing data. The entire random search process is repeated using different search spaces and different architectures to locate the most optimal ML models that satisfy the evaluation criteria. The coefficient of determination (R^2) and root mean squared error (RMSE) were used in this study to evaluate the accuracy and generalizability of the tree-based ML models.

The resulted optimal DT, RF, and GBT models in terms of estimating the measured ultimate pile capacity (for training and testing phases) based on R^2 and RMSE are summarized in Table 16. The comparisons between the predicted and measured ultimate pile capacity for the three ML models are presented in Figure 43. It can be seen that the DT model has the least performance as compared to other ML models in both the training and testing phases. The RMSE (=99.96) for testing subset is much higher than the RMSE (=51.6) for training subset, indicating high overfitting. The second best-performed model is the RF model, which has a testing R^2 and RMSE values of 0.94 and 43.23, respectively. The overfitting condition reduced considerably for RF model as compared to the DT model. However, the GBT model significantly outperforms both the DT and RF models and demonstrates to be the best-performed and generalized tree-based ML model in this study.

Table 16. Optimum tree-based ML models based on training and testing phases

Tree based ML models	Phase	R^2	RMSE (tons)
Decision Tree (DT)	Training	0.91	51.6
	Testing	0.66	99.96
Random Forest (RF)	Training	0.98	25.78
	Testing	0.94	43.23
Gradient Boosted Tree (GBT)	Training	0.99	16.9
	Testing	0.97	27.2
Artificial Neural Network *(26)*	Training	0.99	10.18
	Testing	0.97	24.9

Comparison between ML Models and Selected Direct Pile-CPT Methods

The best-performed ANN (9-7-7-1, 14-9-3-1) and tree-based (GBT) ML models developed in this study were compared with selected top-performed direct pile-CPT methods (LCPC, ERTC3,

probabilistic, and UF) using the same testing subset, which includes 20% of the entire 80-pile load test database. The results of this comparison in terms of the slope of the best-fit line (Q_{fit}/Q_m) of predicted (Q_p) versus measured (Q_m) ultimate pile capacity, the root mean squared error (RMSE), and the arithmetic mean (μ) and coefficient of variation (COV) of Q_p/Q_m, are presented in Table 17. The comparison clearly shows that the ANN and GBT models outperform the top-performed pile-CPT methods in all evaluation criteria. The values of RMSE seem to be much lower for the ML models than the top-performed pile-CPT methods. Thus, it can be concluded that, in general, the ML models perform substantially better than the conventional direct pile-CPT methods.

Figure 43. Predicted versus measured ultimate pile capacity for training and testing phases: (a) decision tree model, (b) random forest model, and (c) gradient boosted tree model

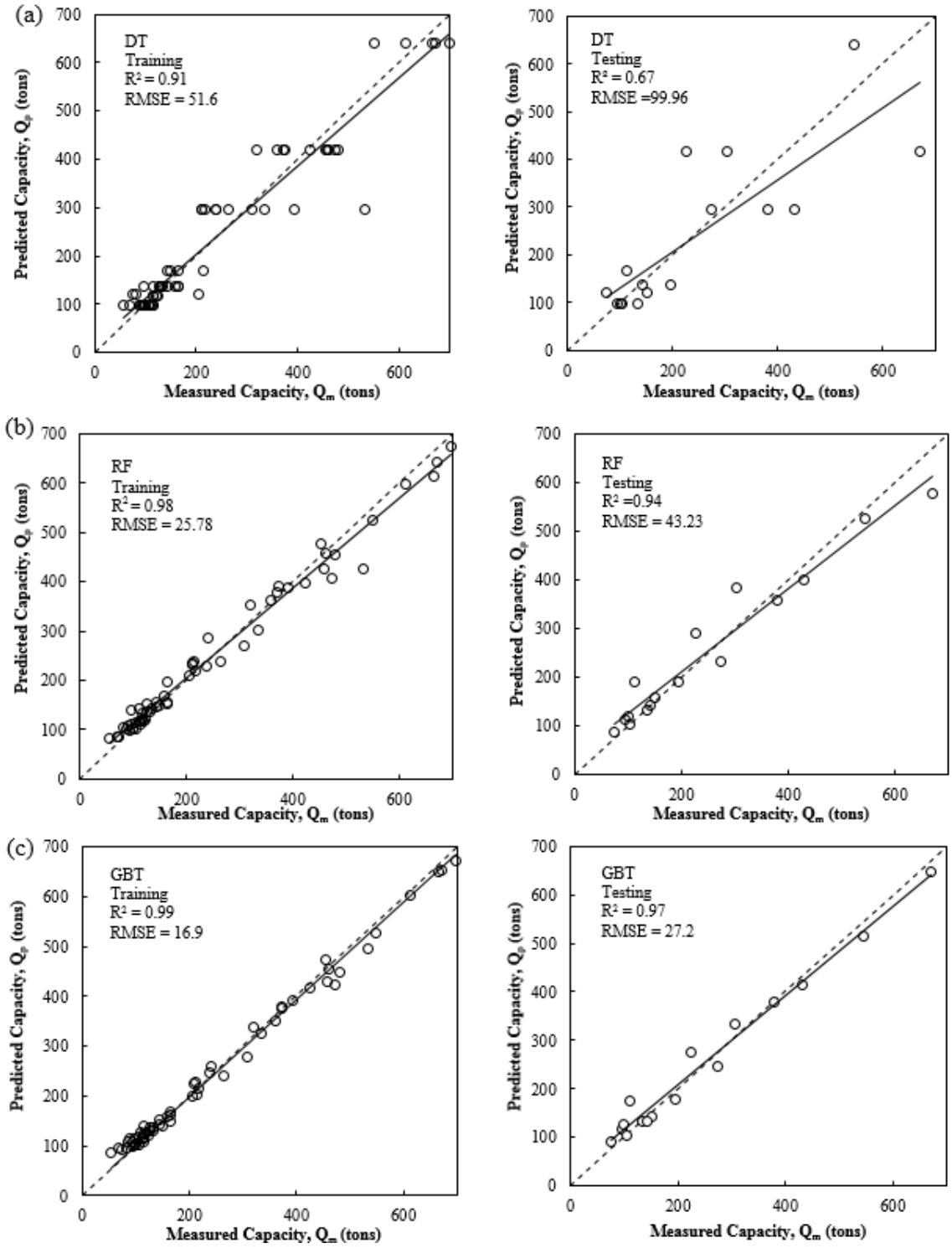


Table 17. Comparison between ML models and selected pile-CPT methods

		First criteria				Second criteria			Overall
Methods	Q_{fit}/Q_m	R^2	R_1	RMSE (tons)	R_2	Mean of Q_p/Q_m	COV	R_3	Final rank
ANN (9-7-7-1)	0.97	0.98	2	22.49	1	1.02	0.15	1	1
ANN (14-9-3-1)	0.97	0.97	3	24.90	2	1.05	0.14	2	2
GBT	0.98	0.97	1	27.20	3	1.07	0.17	3	2
UF	1.12	0.93	6	51.30	5	1.05	0.24	4	5
Probabilistic	0.99	0.91	4	35.29	4	1.06	0.23	5	4
LCPC	1.10	0.91	5	57.13	6	1.11	0.28	6	6
ERTC3	1.09	0.87	7	61.97	7	1.10	0.32	7	7

Update the Pile Design from CPT Software

The top-performed direct pile-CPT methods in this study were implemented in the updated version of the "Louisiana Pile Design from Cone Penetration Test (LPD-CPT)" software for use in the design and analysis of pile foundations. This included the LCPC, ERTC3, Probabilistic, Philipponnat, De Ruiter and Beringen, and Schmertmann methods. The combined pile-CPT method of the top-performed methods was also implemented in the LPD-CPT. In addition, the Schmertmann method was modified for better and more accurate estimating the ultimate pile capacity and implemented in the LPD-CPT software. The program performs the analysis on the CPT sounding, classify the soil profile based on probabilistic and Robertson 2010 CPT classification methods, and then uses the selected pile- CPT method to plot the estimated ultimate pile capacity with depth. The main features of the updated LPD-CPT software are presented in Figure 44 through Figure 52. The main menu of the software (input parameters and data) is presented in Figure 44, which include: (1) enter project information, (2) enter elevations and scour information, and (3) select input CPT data file and corresponding units and format.

Figure 45 Figure 45 illustrates the section and opens the input CPT data file. After selecting the CPT file, the software plots the profiles of CPT data in parallel with soil classification using the probabilistic and Robertson 2010 CPT classification methods (Figure 46). Following this, the user inputs the pile information (type and size) and elevations (Figure 47) and then selects the pile-CPT design method as shown in Figure 48. The software then plots the profile of ultimate pile capacity with depth for the selected pile-CPT method as shown in Figure 49. The software can plot the ultimate pile capacity with depth for three different scenarios: no pre-bore, pre-bore (casing), and long-term scour condition (based on FHWA guidelines). For the selected scenario, the software calculates the required pile length for the applied load based on selected pile-CPT method as shown in Figure 44. The input and output of the batching process analysis in the LPD-CPT software for the different bents in a bridge are presented in Figure 51 and Figure 52, respectively, for given pile size, ground surface elevations, local scour elevation, and load and resistance factor.

Figure 45. Section and open the input CPT data file

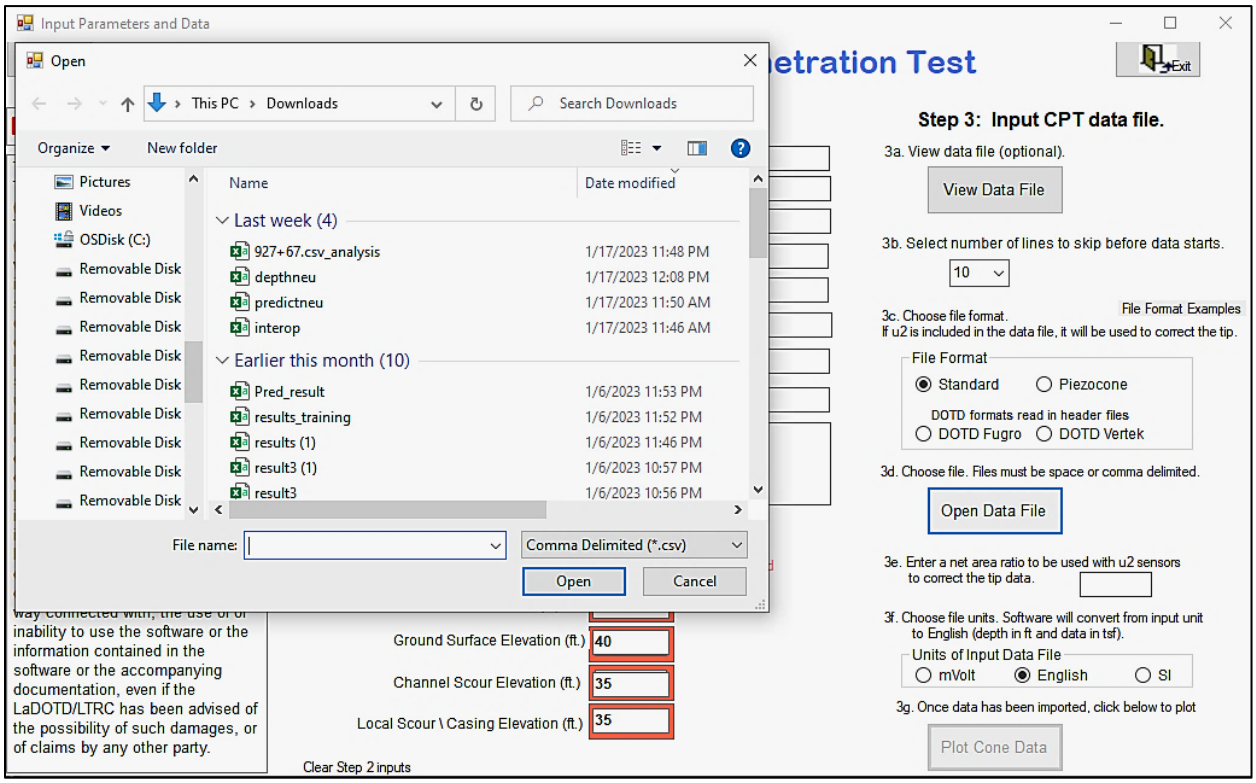


Figure 46. Plot the soil profile using the probabilistic and Robertson 2010 CPT classification methods

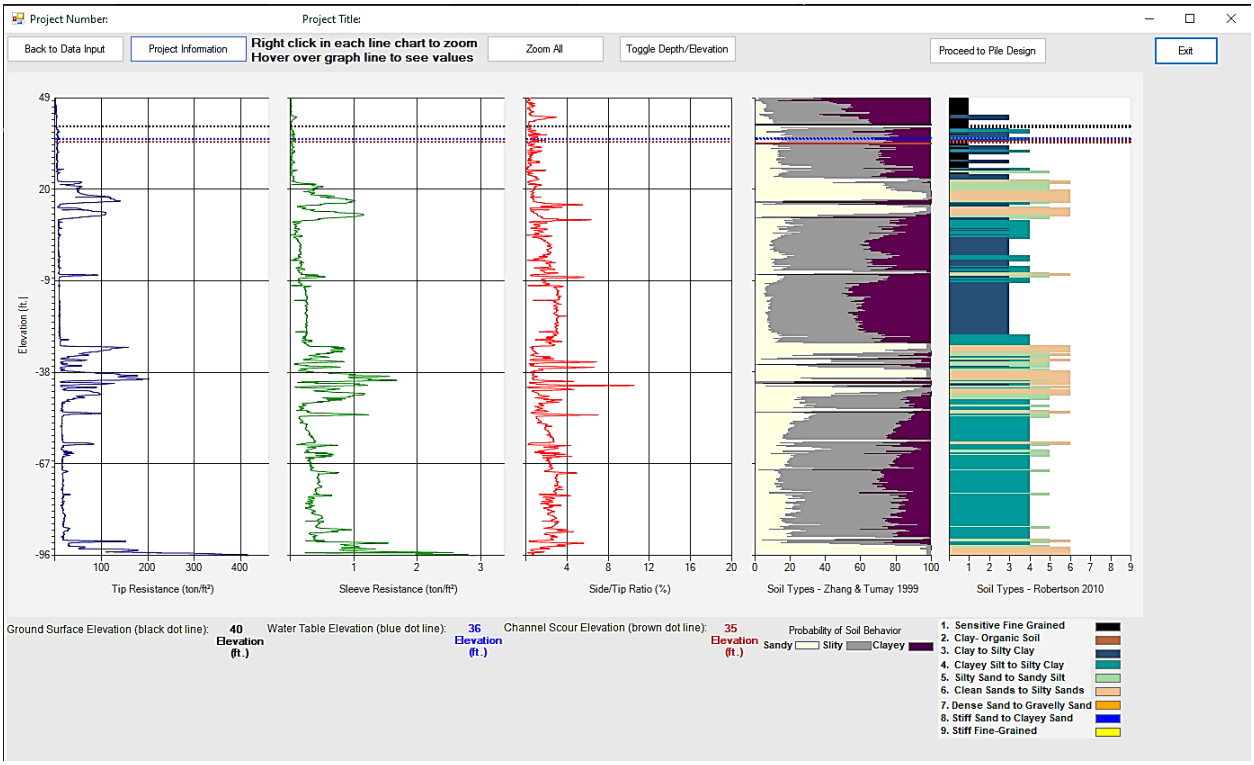


Figure 47. Input pile information and elevations

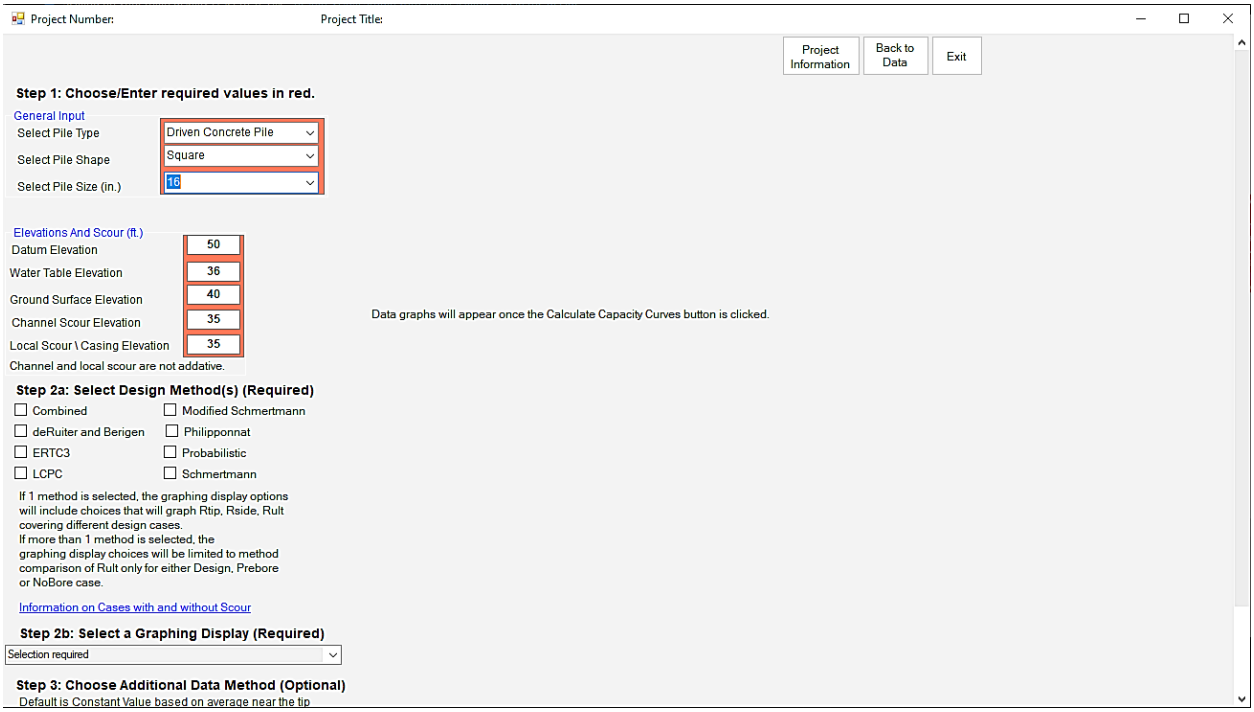


Figure 48. Section of pile design method

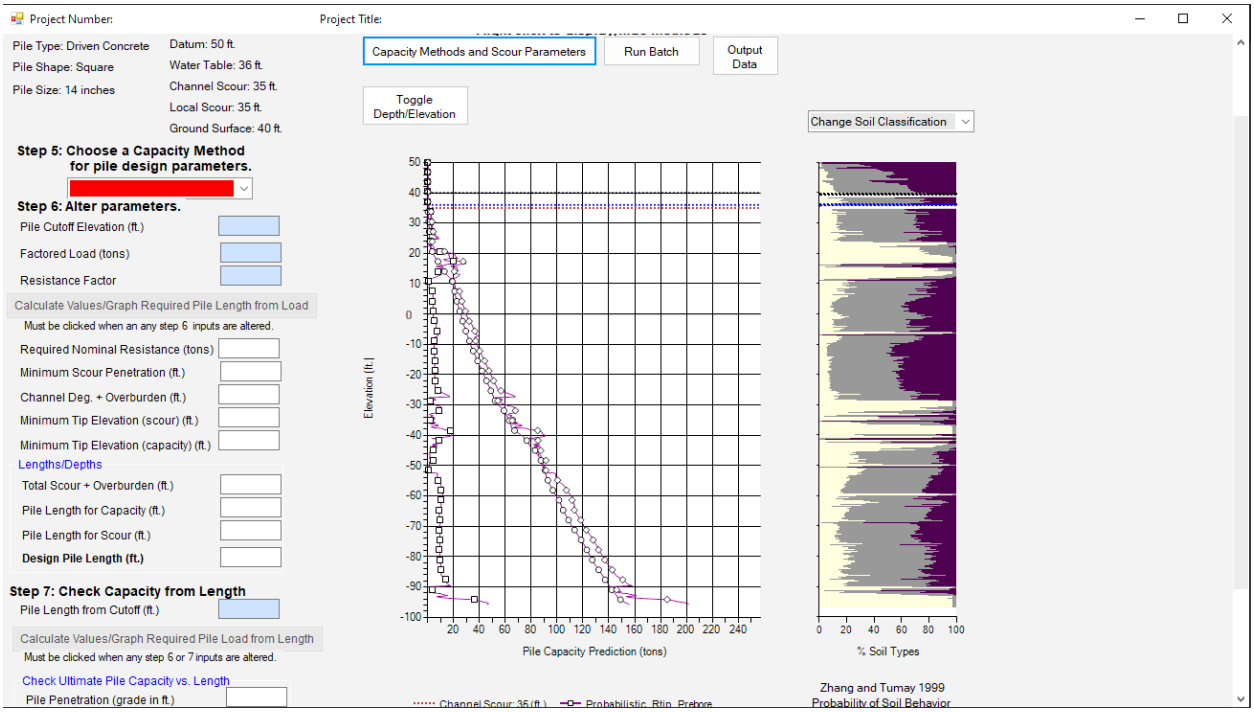


Figure 49. Plot the profile of ultimate pile capacity with depth for the selected pile-CPT method

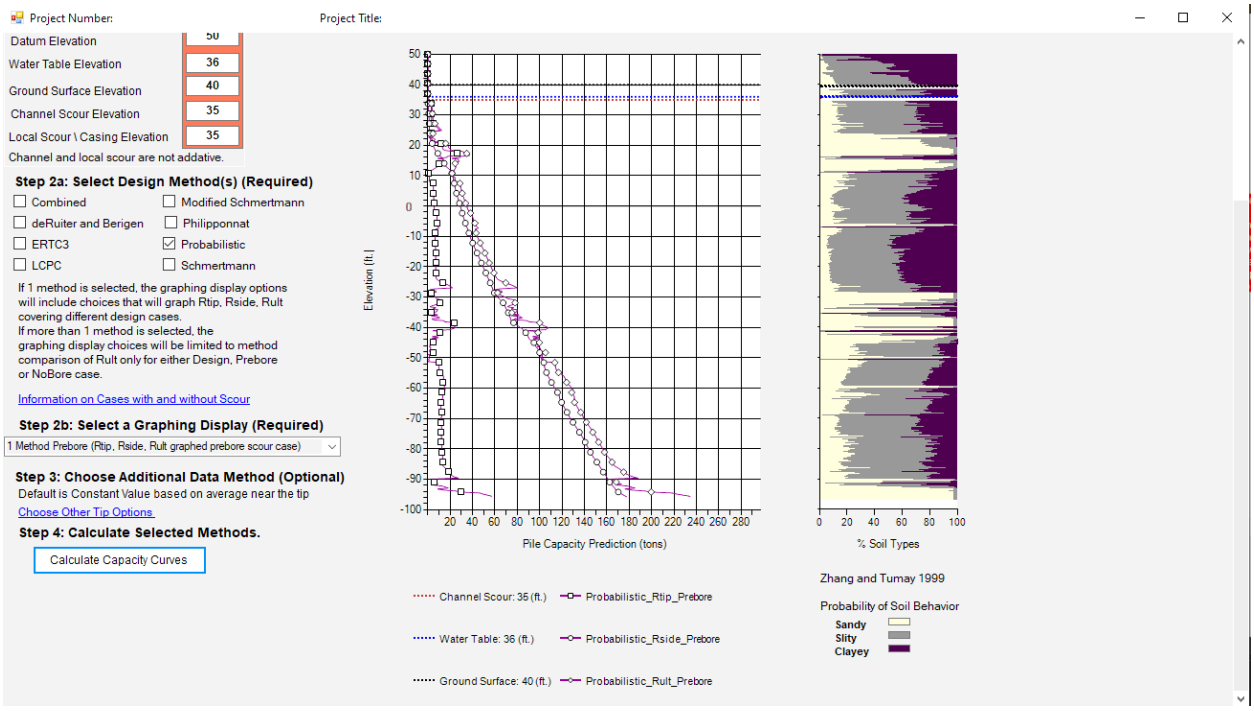


Figure 49. Calculate the required pile length from applied load based on selected pile-CPT method

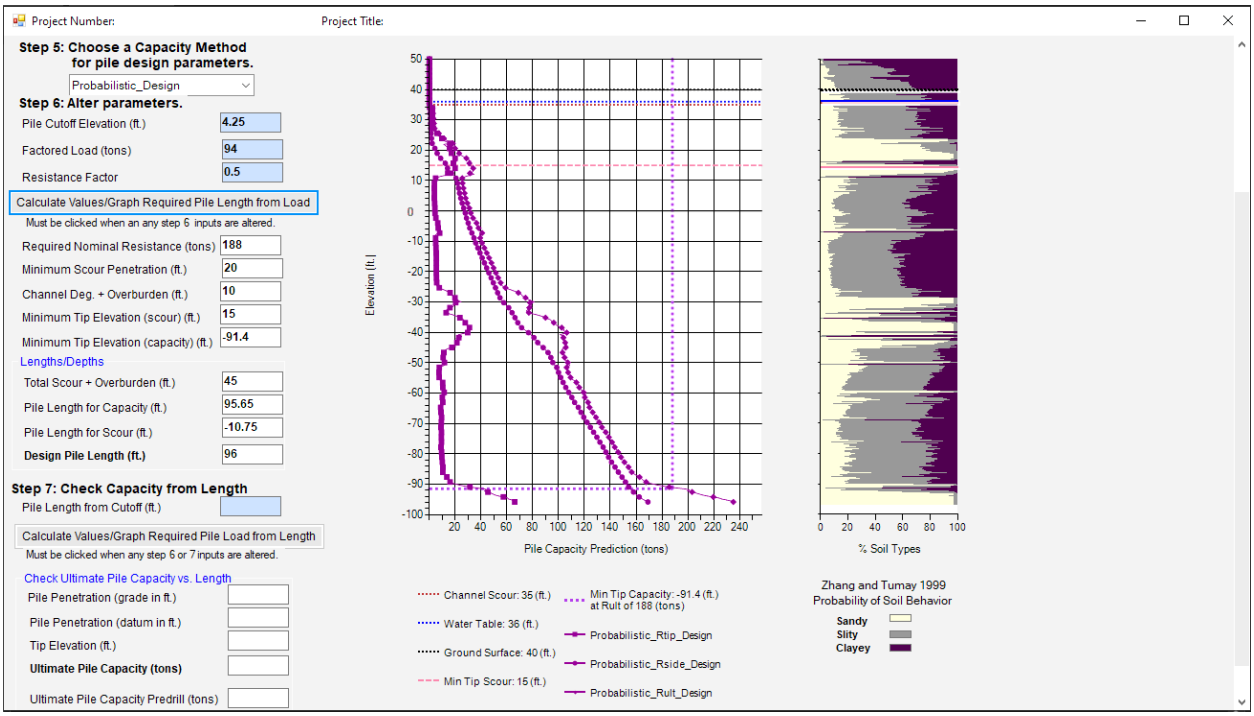


Figure 51. Input the parameters for the batch process analysis of the bridge

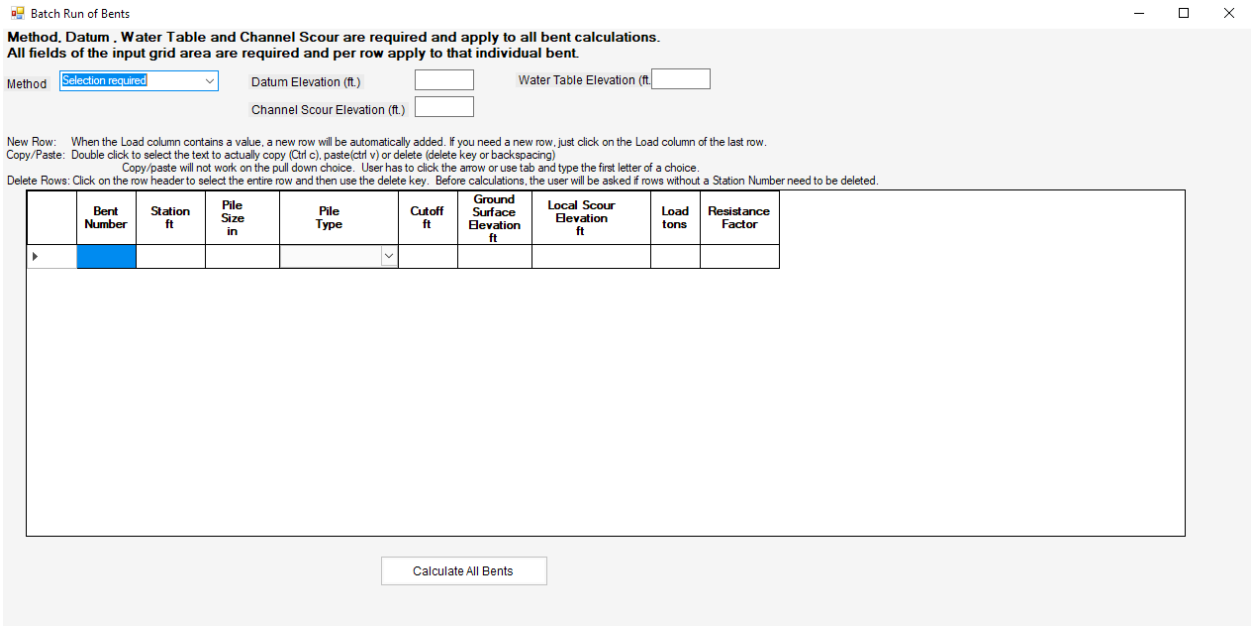
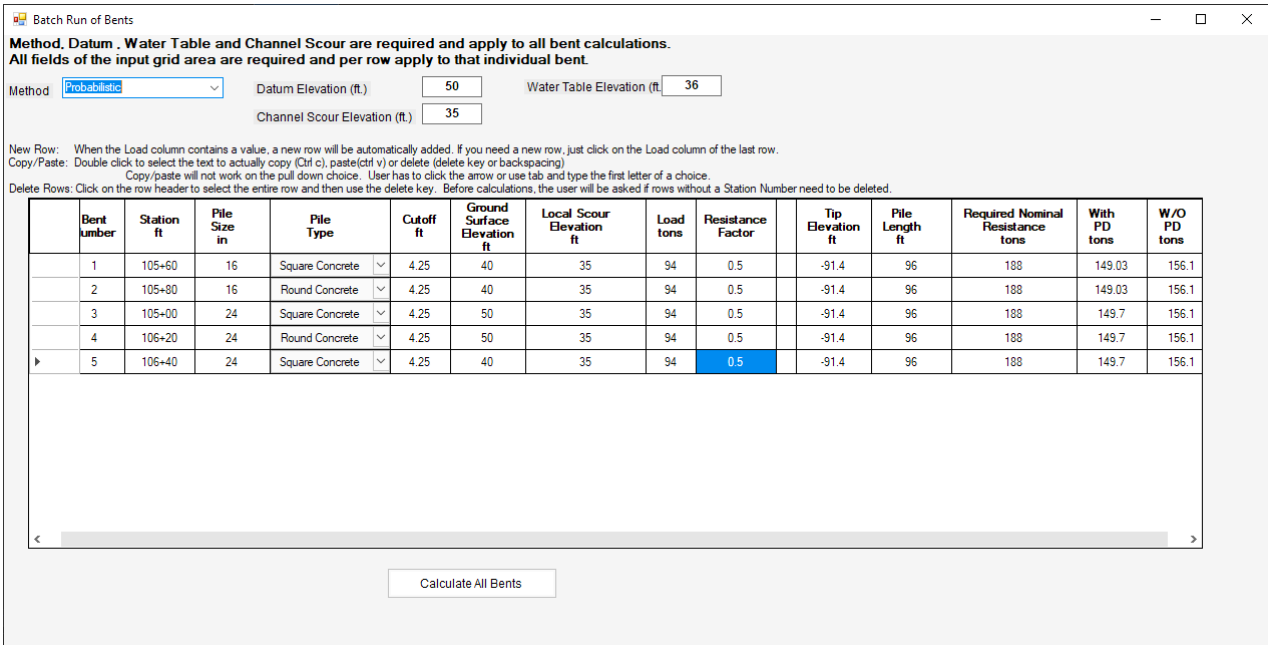


Figure 52. Output results of the batch process analysis of the bridge



Summary and Conclusions

This study focused on evaluating the direct pile-CPT methods for estimating the ultimate axial capacity of square PPC piles driven into Louisiana soils. A total of 21 direct CPT methods were included in this investigation. A database of 80 PPC pile load tests that were loaded to failure were used in this evaluation analysis. The measured ultimate load capacity (Q_m) for each pile was determined from the load-settlement curve using the Davisson and modified Davisson interpretation methods. The ultimate load capacity of each pile (Q_p) was also determined using the 21 direct pile-CPT methods, and the estimated values were compared with the measured pile capacities from static pile load tests.

Researchers used three approaches to evaluate the different pile-CPT methods. In the first approach, three statistical criteria (best fit line for Q_p versus Q_m, arithmetic mean and standard deviation of Q_p/Q_m, and the cumulative probability of Q_p/Q_m) were adopted to evaluate the performance of the 21 direct pile-CPT methods. These criteria were used to rank the CPT methods based on their performance. The final rank of each method was then determined from the Rank Index (RI). The results of this evaluation showed the following order of the best performed pile-CPT methods: LCPC, ERTC3, Probabilistic, UF, Philipponnat, De Ruiter and Beringen, CPT2000, UWA, and Schmertmann methods.

The second approach used for evaluating the pile-CPT methods is the MultiDimensional Unfolding (MDU), which is a technique for representing different objects and judges in a two-dimensional space. In this approach, the result of each pile load test was regarded as a judge that ranks the objects, which are the 21 pile-CPT methods, based on the value of Q_p/Q_m. The abilities of MDU analysis for showing the extent of agreement between the measured pile capacities and predicted capacities from the pile-CPT methods, the existence of outliers, and the similarity between the pile-CPT methods were described. The MDU analysis helps to find the typical ranking of the pile-CPT methods. Based on MDU analysis, the pile-CPT methods that were located close to the center of the measured pile capacities plot were considered as high performance methods, while those methods that were located far from the center were considered as low performance methods. Interestingly, the pile-CPT methods that showed the best performance according to MDU approach are the same methods that have the highest rankings based on statistical analysis using the three criteria.

The third approach used for evaluating pile-CPT methods is based on LRFD reliability analysis. The results of the LRFD analysis in terms of resistance factor and efficiency are consistent with the previous criteria in which the LCPC, ERTC3, Probabilistic, UF, Philipponnat, De Ruiter methods have the highest rankings.

For selecting the best pile-CPT methods, the nine top-performed methods were further analyzed using the MDU analysis to evaluate the methods' similarity. In this analysis, the pile-CPT methods that are located close to each other in the two-dimensional MDU space can be considered to be similar to each other. Accordingly, the top nine pile-CPT methods can be divided into three groups:

- Group 1: Philipponnat, UF, Probabilistic, LCPC, and De Ruiter methods:
- Group 2: Schmertmann and ERTC3 methods; and
- Group 3: UWA and CPT2000.

If geotechnical engineers are interested to get a better range of estimating the ultimate pile capacity by choosing three pile-CPT methods, it is recommended to select one method from each group.

The pile database was further divided into four different categories based on soil type for extra analysis and evaluation. In category 1, the piles have less than 25% of their capacity due to sand layers. In category 2, the sand layers contribute between 25% and 50% of the pile capacity; for piles in category 3, the sand layers contribute between 50% and 75% of their capacity. For the piles in category 4, the sand layers contribute more than 75% of their capacity. The evaluation of pile-CPT methods for each category of piles separately showed that the performance of pile-CPT methods is different at each category. This is mainly due to the fact that each pile-CPT method includes different equations for tip and side resistance of the piles in different soil types, so it is possible that a pile-CPT method has a more accurate equation for pile side resistance in a soil type but under-/over-predicts the pile capacity with less accuracy in other soil type. The performance of each pile-CPT method was studied in detail for each category. The general observation showed that increasing the sand contribution to the pile capacity causes overprediction in pile capacity for pile-CPT methods, and hence an increase in the standard deviation of Q_p/Q_m, which means less accuracy and reliability in estimations of pile-CPT methods in sandy soils.

The log-normal distribution of the ratio of predicted to measured pile capacity, Q_p/Q_m was used for adding, multiplying, and linear combination for normal distribution of $\ln(Q_p/Q_m)$ of different pile-CPT methods, and hence exploring better relationship and developing combined methods for estimating the pile capacity. The 80 pile load tests collected in Louisiana were used to develop combined pile-CPT methods to estimate the ultimate pile capacity for four soil categories based on the contribution of sandy layers to the total capacity. In addition, another combined pile-CPT method was developed for the general case for all piles without considering soil categories. Evaluating the developed combined pile-CPT methods demonstrated significant improvement in the accuracy of estimating the ultimate pile capacity. The main advantage of

using the combined pile-CPT methods is for piles in soil category 4 with more than 75% contribution of sandy layers to the total pile capacity.

Four machine learning (ML) techniques including the artificial neural network (ANN) and three tree-based techniques (decision tree, DT, random forest, RF, and gradient boosted tree, GBT) were also used to develop ML models to estimate the ultimate pile capacity from CPT data (q_t and f_s). The comparison results between the ML models and selected direct pile-CPT methods demonstrated that the ANN and GBT models substantially outperform the top-performed pile-CPT methods in all evaluation criteria. The values of the slope of best-fit line (Q_{fit}/Q_m) of predicted (Q_p) versus measured (Q_m) ultimate pile capacity and the arithmetic mean (μ) of Q_p/Q_m are much better, and the RMSE and coefficient of variation (COV) of Q_p/Q_m are much lower for the ML models than the top-performed pile-CPT methods.

Recommendations

Based on the results of this research study, the following recommendations are offered to DOTD engineers:

- It is highly recommended that the DOTD design engineers start using the top-performing direct pile-CPT methods that were implemented in the updated version of the "Louisiana Pile Design from Cone Penetration Test (LPD-CPT)" software for the design of piles in new bridges and other infrastructure.
- It is recommended to start using the modified Schmertmann method instead of the original Schmertmann method for design of piles in bridges and other infrastructures.
- It is recommended that the DOTD design engineers begin using the developed combined pile-CPT method that is implemented in the LPD-CPT software for the design of piles in new bridges and other infrastructure, and compare the results with the top-performed direct pile-CPT methods to build confidence on the accuracy of the combined methods.
- It is recommended to start exploring the potential benefits of using the machine learning (ML) models for estimating the ultimate pile capacity from CPT data, and compare the results with the top-performing direct pile-CPT methods to build confidence in using ML models.
- It is recommended to hold a workshop to train DOTD engineers on using the updated version of the LPD-CPT software for design of individual piles and pile foundations for bridge bents.
- It is recommended to continue evaluating the top-performing pile-CPT methods, the combined pile-CPT method, and the ML models as more pile load tests become available.
- It is recommended to consider updating the LPD-CPT software to include different pile types other than the PPC piles.
- It is recommended to start conducting a piezocone penetration test (PCPT) on all old and new bridges to create database to evaluate and/or develop pile-PCPT methods for estimate the ultimate pile capacity from PCPT data
- It is recommended that a new research study investigate the effect of scour on the long-term pile capacity utilizing the CPT/PCPT data.
- It is highly recommended to consider using the artificial intelligent (AI), supervised and unsupervised machine learning (ML), and deep learning (DL) techniques to develop advanced and more accurate models to estimate the ultimate pile capacity from CPT data, considering all important pile, soil, and CPT/PCPT parameters.

Acronyms, Abbreviations, and Symbols

Term	Description
AASHTO	American Association of State Highway and Transportation Officials
A_b	Area of the pile tip
AI	Artificial intelligence
ANFIS	Adaptive neuro-fuzzy interface systems
ANN	Artificial neural network
A_{si}	Area of the pile shaft in the soil layer i
B	Pile width
CAPWAP	Case Pile Wave Analysis Program
CDF	Cumulative density function
CL	Low plasticity clay
CH	High plasticity clay
COV	Coefficient of variation
COV_Q	Coefficients of variation for the load
COV_{QD}	Coefficients of variation for dead load
COV_{LD}	Coefficients of variation for live load
COV_R	Coefficients of variation for the resistance
CPT	Cone penetration test
CPTu	Cone penetration test with porewater measurement
DOTD	Department of Transportation and Development
DRIVEN	Driven Piles program
DT	Decision tree
E	Error of the network
EN	Engineering News
ERTC	European Technical Committee
f	Unit skin friction of the pile
f_i	Unit skin friction of soil layer i
f_s	Sleeve friction
FDOT	Florida Department of Transportation
FHWA	Federal Highway Administration

Term	Description
FORM	First order reliability method
FOSM	First order second-moment
F.S.	Factor of safety
$f_{s,\,avg}$	Average sleeve friction
g	Limit state function
GA	Genetic algorithm
GBT	Gradient boosted tree
GNDH	Group method of data handling
GEP	Gene expression programming
GP	Poorly graded gravel
ICP	Imperial College Pile
I_{SBT}	Soil behavior index
KNN	K-nearest neighbor
LCPC	Laboratoire Central des Ponts et Chaussées
L_e	Embedded length of pile
LPD-CPT	Louisiana Pile Design from cone penetration test
LRFD	Load and resistance factor design
LTRC	Louisiana Transportation Research Center
LPD-CPT	Louisiana Pile Design from cone penetration test
MDU	MultiDimensional Unfolding
MDS	MultiDimensional Scaling
ML	Machine learning
NGI	Norwegian Geotechnical Institute
NORMDIST	Normal distribution function
P	Cumulative probability
P_a	Atmospheric pressure
P_{50}	50% cumulative probability
P_{90}	90% cumulative probability
P_f	Probability of failure
PCPT	Piezocone penetration test
PPC	Precast prestressed concrete

Term	Description
PSO	Particle swarm optimization
P_{true}	Lowest magnitude of probability
Q	Design load
Q_{all}	Allowable design load
Q_{avg}	Average value of pile capacity
q_b	Unit tip bearing capacity
Q_b	End-bearing capacity
q_c	Cone tip resistance
Q_D	Dead load
Q_{fit}	Best fit of pile capacity
Q_i	Load effect
Q_L	Live load
Q_m	Measured pile capacity
Q_d	Design load carrying capacity
Q_p	Predicted pile capacity
Q_s	Shaft friction capacity
q_t	Corrected cone tip resistance
$q_{t,\ avg}$	Average cone tip resistance
$q_{t\text{-}tip,\ 4B\ above}$	Average cone tip resistance within 4 width above pile tip
$q_{t\text{-}tip,\ 8B}$	Average cone tip resistance within 8 width above pile tip
$q_{t\text{-}tip,\ 4B\ below}$	Average cone tip resistance within 4 width below pile tip
Q_u	Ultimate capacity
Q_{ult}	Ultimate geotechnical pile resistance
R	Resistance of a given structure
r	Coefficient of correlation
R^2	Coefficient of determination
R_{design}	Design pile resistance
RF	Random forest
RI	Ranking index
R_m	Measured pile resistance
RMSE	Root mean squared error

Term	Description
R_n	Nominal pile resistance
SBT	Soil behavior types
SC	Clayey sand
SM	Silty sand
SMACOF	Stress Majorization of a COmplicated Function
SP	Poorly graded sand
S_u	Undrained shear strength
SVM	Support vector machines
u_2	Porewater pressure
U	Soil classification index
UF	University of Florida
UWA	University of Western Australia
USCS	Unified Soil Classification System
V_m	Eigenvector
WEAP	Wave Equation Analysis
β	Reliability index
β_T	Target reliability index
Δ	Distance matrix
ϕ	Resistance factor
φ	angle of internal friction
δ_{ij}	Input matrix for MDS
λ	Multiplication factor
λ_{QD}	Load bias factor for dead load
λ_R	Resistance bias factor
λ_{QL}	Load bias factor for live load
γ_D	Load factor for dead load
γ_i	Load factor
γ_L	Load factor for live load
μ	Arithmetic mean
μ_{ln}	Mean of $\ln(Q_p/Q_m)$

Term	Description
μ_{linx}	Equivalent lognormal mean of ln(x)
μ_R	Mean value of the load
μ_R	Mean value of the resistance
η	Learning rate
σ	Standard deviation
σ_{linx}	Equivalent standard deviation of ln(x)
σ_{ln}	Standard deviation of $\ln(Q_p/Q_m)$
$\sigma(X)$	Kruskal's stress
ω_{ij}	Matrix of weights

References

[1] J. H. Schmertmann, "Guidelines for Cone Penetrometer Test, Performance and Design," U.S. Department of Transportation, Report No. FHWA TS-78-209, Washington, DC, 1978.

[2] J. De Ruiter and F. Beringen, "Pile foundations for large North Sea structures," *Marine Georesources & Geotechnology,* vol. 3, no. 3, pp. 267-314, 1979.

[3] M. Bustamante and L. Gianeselli, "Pile bearing capacity prediction by means of static penetrometer CPT," *Proceedings of the 2nd European symposium on penetration testing,* pp. 493-500, 1982.

[4] M. T. Tumay and M. Fakhroo, "Friction Pile Capacity Prediction in Cohesive Soils Using Electric Quasi-Static Penetration Tests," Louisiana Departmentof Transportation and Development, Interim Research Report No. 1, Baton Rouge, LA, 1982.

[5] N. Aoki and D. A. De Alencar, "An approximate method to estimate the bearing capacity of piles," in *Proceedings of the 5th Pan-American Conf. of Soil Mechanics and Foundation Engineering, No. 1, International Society of Soil Mechanics and Geotechnical Engineering*, Buenos Aires, Argentina, 1975.

[6] G. Price and W. I., "A comparison between cone penetration test results and the performance of small diameter instrumented piles in stiff clay," in *Proceedings of the 2nd European Symposium on Penetration Testing, , No. 2,* Amsterdam, the Netherlands, 1982.

[7] G. Philipponnat, "Méthode pratique de calcul d'un pieu isolé, à l'aide du pénétromètre statique," *Revue Francaise de Geotechnique,* vol. 10, pp. 55-64, 1980.

[8] M. Clisby, R. Scholtes, M. Corey, H. Cole, P. Teng and J. Webb, "An evaluation of pile bearing capacities," Mississippi State Highway Department, Volume 1, Final Report, 1978.

[9] A. Eslami and B. Fellenius, "Pile capacity by direct CPT and CPTu methods applied to 102 case histories," *Canadian Geotechnical Journal,* vol. 34, p. 886–904, 1997.

[10] M. Abu-Farsakh and H. Titi, "Probabilistic CPT method for estimating the ultimate capacity of friction piles," *Geotechnical Testing Journal,* vol. 30, no. 5, pp. 387-398, 2007.

[11] K. Karlsrud, C. Clausen and P. Aas, "Bearing capacity of driven piles in clay, the NGI approach," in *Proceedings of Int. Symp. on Frontiters in Offshore Geotehcnics,* Perth, Australia, 2005.

[12] R. Jardine, F. Chow, R. Overy and J. Standing, ICP design methods for driven piles in sands and clays, London, UK: Thomas Telford , 2005.

[13] B. Lehane, F. Chow, B. McCabe and R. Jardine, "Relationships between shaft capacity of driven piles and CPT end resistance," *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering,* vol. 143, no. 2, pp. 93-102, 2000.

[14] P. K. Robertson, R. G. Campanella, M. P. Davis and A. Sy, "Axial Capacity of Driven Piles in Deltaic Soils Using CPT," in *Proceedings of the 1st International Symposium on Penetration,* Orlando, FL, 1988.

[15] J. L. Briaud and L. M. Tucker, "Measured and Predicted Axial Response of 98 Piles," *Journal of Geotechnical Engineering,* vol. 114, no. 8, p. 984–1001, 1988.

[16] K. E. Tand and E. E. Funegard, "Pile Capacity in Stiff Clays - CPT Method," in *Proceedings, 12th International Conference on Soil Mechanics and Foundation Engineering*, Rio de Janeiro, Brazil, 1989.

[17] M. Y. Abu-Farsakh and H. H. Titi, "Assessment of Direct CPT Methods for Predicting the Ultimate Capacity of Friction Driven Piles," *Journal of Geotechnical and Geoenvironmental Engineering,* vol. 130, no. 9, pp. 935-944, 2007.

[18] M. T. Davisson, "High capacity piles," in *Proc., Lecture Series on Innovation in Foundation Construction, American Society of Civil Engineering*, New York, USA, 1972.

[19] J. H. Schmertmann, "Guidelines for cone penetration test: performance and design," U.S. Department of Transportation, Report No. FHWA-TS-78-209, Washington, D.C., 1978.

[20] B. Douglas and R. Olsen, "Soil classificaion using electric cone penetrometer," in *Proceedings of Conference on Cone Penetration Testing and Experience*, St. Louis, MO, 1981.

[21] P. Robertson, R. Campanella, D. Gillespie and J. Greig, "Use of piezometer cone data. Paper presented at the Use of in situ tests in geotechnical engineering.," in *In-Situ '86: Use of In-situ Testing in Geotechnical Engineering*, New York, NY, 1986.

[22] R. Campanella, P. Robertson, M. Davies and A. Sy, "Use of in-situ tests in pile design. Paper presented at the," in *Proceedings of the 12th international conference on soil mechanics and foundation engineering, ICSMFE*, Rio de Janeiro, Brazil, 1989.

[23] Z. Zhang and M. Tumay, "Statistical to fuzzy approach toward CPT soil classification," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 125, no. 3, pp. 179-186, 1999.

[24] P. Robertson, "Soil behaviour type from the CPT: an update," in *Proceedings of the 2nd international symposium on cone penetration testing, CPT 10*, Huntington Beach, CA, 2010.

[25] M. Y. Abu-Farsakh, Z. Zhang, M. Tumay and M. Morvant, "Computerized cone penetration test for soil classification: development of MS-Windows software," *Transportation Research Record*, vol. 2053, no. 1, pp. 47-64, 2008.

[26] M. G. Jefferies and M. P. Davies, "Use of CPTU to estimate equivalent SPT N 60," *Geotechnical Testing Journal*, vol. 16, no. 4, pp. 458-468, 1993.

[27] J. Clausen C., P. M. Aas and K. Karlsrud, "Bearing Capacity of Driven Piles in Sand, the NGI Approach," in *Proceedings of the ISFOG Conference*, Perth, WA, 2005.

[28] B. Lehane, J. Schneider and X. Xu, "The UWA-05 method for prediction of axial capacity of driven piles in sand," in *Frontiers in offshore geotechnics: ISFOG, 683-689.*, Perth, Australia, 2005.

[29] B. M. Lehane, Y. Li and R. Williams, "Shaft capacity of displacement piles in clay using the cone penetration test," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 139, no. 2, pp. 253-266, 2012.

[30] H. Kolk and E. der Velde, "A reliable method to determine friction capacity of piles driven into clays," in *Offshore Technology Conference*, Houston, Texas, 1996.

[31] B. Dijk and H. Kolk, "CPT-based design method for axial capacity of offshore piles in clays," in *Proceedings of the International Symposium on Frontiers in Offshore Geotechnics II*, London, UK, 2011.

[32] M. Randolph, "Science and empiricism in pile foundation design," *Géotechnique*, vol. 53, no. 10, pp. 847-876, 2003.

[33] R. Salgado, S. I. Woo and D. Kim, "Development of load and resistance factor design for ultimate and serviceability limit states of transportation structure foundations," Indiana Department of Transportation, Report No. FHWA/IN/JTRP-2011/03, Indianapolis, IN, 2011.

[34] D. Bloomquist, M. C. McVay and Z. Hu, "Updating Florida Department of Transportation's (FDOT) Pile/shaft Design Procedures Based on CPT & DPT Data," Department of Civil and Coastal Engineering, Report No. BD-545, PPWO #43, Gainesville, FL , 2007.

[35] Z. Hu, M. McVay, D. Bloomquist, D. Horhota and P. Lai, "New ultimate pile capacity prediction method based on cone penetration test (CPT)," *Canadian Geotechnical Journal*, vol. 49, no. 8, pp. 961-967, 2012.

[36] G. Togliani, "Pile capacity prediction for in situ tests. , Taylor and Francis Group, ,, 1187-1192.," in *Proceedings of geotechnical and geophysical site characterization*, London, UK, 2008.

[37] J. Zhou, Y. Xie, Z. Zuo, M. Luo and X. Tang, "Prediction of limit load of driven pile by CPT," in *Proceedings of the 2nd European symposium on penetration testing*, Amsterdam, Netherlands, 1982.

[38] M. S. Almeida, F. A. Danziger and T. Lunne, "Use of the piezocone test to predict the axial capacity of driven and jacked piles in clay," *Canadian Geotechnical Journal*, vol. 33, no. 1, pp. 23-41, 1996.

[39] J. Powell, T. Lunne and R. Frank, "Semi empirical design procedures for axial pile capacity in clays," in *15th International Conference on Soils Mechanics and Geotechnical Engineering*, Istanbul, Turkey, 2001.

[40] F. S. Niazi and P. W. Mayne, "CPTu-based enhanced UniCone method for pile capacity," *Engineering Geology*, vol. 212, pp. 21-34, 2016.

[41] K. Takesue, H. Sasao and T. Matsumoto, "Correlation between ultimate pile skin friction and CPT data," *Geotechnical Site Characterization*, vol. 2, pp. 1177-1182, 1998.

[42] B. Van Dijk and H. Kolk, "CPT-based design method for axial capacity of offshore piles in clays," in *Proceedings of the International Symposium on Frontiers in Offshore Geotechnics II*, Perth, Australia, 2011.

[43] K. G. Gavin and B. M. Lehane, "The shaft capacity of pipe piles in sand," *Canadian Geotechnical Journal,*, vol. 40, no. 1, pp. 36-45, 2003.

[44] D. Igoe, K. Gavin and B. O'Kelly, "Field tests using an instrumented model pipe pile in sand," in *7th Int. Conf on Physical Modelling in Geotechnics*, Zurich, Switzerland, 2010.

[45] D. J. Igoe, K. G. Gavin and B. C. O'Kelly, "Shaft capacity of open-ended piles in sand," *Journal of Geotechnical and Geoenvironmental Engineering,* vol. 137, no. 10, pp. 903-913, 2011.

[46] F. Yu and J. Yang, "Base capacity of open-ended steel pipe piles in sand," *Journal of Geotechnical and Geoenvironmental Engineering,* vol. 138, no. 9, pp. 1116-1128, 2011.

[47] J. H. Long and M. H. Wysockey, "Accuracy of methods for predicting axial capacity of deep foundations," in *Proceedings of the Analysis, Design, Construction, and Testing of Deep Foundations.*, 1999.

[48] A. Eslami, E. Aflaki and B. Hosseini, "Evaluating CPT and CPTu based pile bearing capacity estimation methods using Urmiyeh Lake Causeway piling records," *Scientia Iranica,* vol. 18, no. 5, pp. 1009-1019, 2011.

[49] S. G. Paikowsky, "Load and resistance factor design (LRFD) for deep foundations. Foundation Design Codes," in *Proceedings of IWS*, Kamakura, 2007.

[50] M. A. Styler, Development and implementation of the DIGGS format to perform LRFD resistance factor calibration of driven concrete piles in Florida, Gainesville: University of Florida, 2006.

[51] F. Han, M. Prezzi, R. Salgado and M. Zaheer, "Axial resistance of closed-ended steel-pipe piles driven in multilayered soil," *Journal of Geotechnical and Geoenvironmental Engineering,* vol. 143, no. 3, 2016.

[52] N. S. Juwaied, "Applications of artificial intelligence in geotechnical engineering," *ARPN Journal of Engineering and Applied Sciences*, vol. 13, no. 8, pp. 2764-2785, 2018.

[53] M. A. Shahin, M. B. Jaksa and H. Maier, "Recent Advances and Future Challenges for Artificial Neural Systems in Geotechnical Engineering Applications," *Advances in Artificial Neural Systems,* vol. 2009, pp. 1-9, 2009.

[54] P. H. S. and J. C. H., "A simple and fast algorithm for k-medoids clustering," *Expert Syst Appl.,* vol. 36, no. 2, p. 3336–41, 2009.

[55] M. A. Shahin, "Intelligent computing for modeling axial capacity of pile foundations," *Canadian Geotechnical Journal,* vol. 47, no. 2, pp. 230-243, 2010.

[56] A. Kordjazi, F. N. P. Pooya and M. B. Jaksa, "Prediction of ultimate axial load-carrying capacity of piles using a support vector machine based on CPT data," *Computers and Geotechnics,* vol. 55, pp. 91-102, 2014.

[57] I. Alkroosh and H. Nikraz, "Correlation of Pile Axial Capacity and CPT Data Using Gene Expression Programming," *Geotech Geol Eng,* vol. 29, p. 725–748, 2011.

[58] B. Ghorbani, E. Sadrossadat and J. Bolouri Bazaz, "Numerical ANFIS-Based Formulation for Prediction of the Ultimate Axial Load Bearing Capacity of Piles Through CPT Data," *Geotech Geol Eng,* vol. 36, p. 2057–2076, 2018.

[59] H. Harandizadeh, J. D. Armaghani and M. Khari, "A new development of ANFIS–GMDH optimized by PSO to predict pile bearing capacity based on experimental datasets," *Engineering with Computers,* vol. 37, p. 685–700, 2021.

[60] H. Ardalan, A. Eslami and N. Nariman-Zadeh, "Piles shaft capacity from CPT and CPTu data by polynomial neural networks and genetic algorithms," *Computers and Geotechnics,* vol. 36, pp. 616-625, 2009.

[61] M. H. Baziar, A. Kashkooli and A. Saeedi-Azizkandi, "Prediction of pile shaft resistance using cone penetration tests (CPTs)," *Computers and Geotechnics,* vol. 45, pp. 74-82, 2012.

[62] D. E. Rumelhart and J. L. McClleland, "Learning internal representation by error propagation," *MIT Press, Cambridge, MA,* vol. 1, 1986.

[63] I. A. Basheer, "Neuromechanistic-Based Modeling and Simulation of Constitutive Behavior of Fine-Grained Soils," PhD diss, Kansas State University, Manhattan, Kansas, 1998.

[64] R. Hecht-Nielson, "Kolmogrov's mapping neural network existence theorem," in *Proceedings of the 1st IEEE International Joint Conference on Neural Networks*, San Diego, CA, 1987.

[65] L. Breiman, J. Friedman, R. Olshen and Stone C., Classification and Regression Trees, Belmont, CA: Wadsworth, 1984.

[66] G. James, D. Witten, T. Hastie and R. Tibshirani, An Introduction to Statistical Learning with Applications in R, New York: Springer, 2013.

[67] L. Breiman, "Random Forests," *Machine Learning*, vol. 5–32, p. 45, 2001.

[68] J. H. Friedman, "Stochastic gradient boosting," *Computational statistics & data analysis*, vol. 38, no. 4, pp. .367-378, 2002.

[69] A. International, Standard Test Method for Deep Foundations Under Static Axial Compressive Load, ASTM International, 2013.

[70] D. Florida, Standard specifications for road and bridge construction, Tallahassee, FL. : Florida Department of Transportation, 2010.

[71] R. G. Campanella, D. G. Gillespie and P. K. Robertson, Pore pressures during cone penetration testing: Department of Civil Engineering,, University of British Columbia, 1981.

[72] T. Lunne, T. Eidsmoen, D. Gillespie and J. D. Howland, "Laboratory and field evaluation of cone penetrometers," in *Paper presented at the Use of In Situ Tests in Geotechnical Engineering*, New York, NY, 1986.

[73] M. I. Hossain, Evaluation of Undrained Shear Strength and Soil Classification from Cone Penetration Test, Baton Rouge: Louisiana State University, 2018.

[74] P. F., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research12*, pp. 2825-2830, 2011.

[75] D. Hammerstrom, "Working with neural networks," *IEEE Spectrum*, vol. 30, no. 7, pp. 46 - 53, 1993.

[76] M. Stone, "Cross-validatory choice and assessment of statistical predictions," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 36, no. 2, pp. 111-147, 1974.

[77] M. A. Shahin, H. R. Maier and M. B. Jaksa, "Data Division for Developing Neural Networks Applied to Geotechnical Engineering," *Journal of Computing in Civil Engineering*, vol. 18, no. 2, pp. 105-114, 2004.

[78] G. N. Smith, Probability and statistics in civil engineering: An introduction, Collins, London: Nichols Publishing Company, 1986.

[79] H. H. Titi and M. Y. Abu-Farsakh, "Evaluation of bearing capacity of piles from cone penetration test data," Report No. FHWA/LA.99/334, Louisiana Transportation Research Center, Baton Rouge, Louisiana, 1999.

[80] J. Kruskal and M. Wish, *Multidimensional scaling*, University Paper Series on Quantitative Applications in the Social Sciences, Series 07-011. In: Sage Publications, 1978.

[81] W. S. Torgerson, Torgerson, W. S. (1958). Theory and methods of scaling., New York: John Wiley, 1958.

[82] J. De Leeuw, Recent developments in statistics, In: Chapt. Applications of convex analysis to multidimensional scaling, 1977.

[83] J. De Leeuw and W. J. Heiser, "Multidimensional scaling with restrictions on the configuration," *Multivariate analysis*, vol. 5, pp. 501-522, 1980.

[84] J. De Leeuw and P. Mair, "Multidimensional scaling using majorization: SMACOF in R.," *Journal of Statistical Software*, vol. 31, no. 3, p. 1–30, 2009.

[85] I. Borg and P. J. Groenen, Modern multidimensional scaling: Theory and applications, Berlin, Germany: Springer, 2005.

[86] T. F. Cox and M. A. Cox, Multidimensional scaling, London, UK: Chapman & Hall/CRC, 2000.

[87] M. Alvo and L. Philip, Statistical methods for ranking data, Springer, 2014.

[88] R. C. TeamR, *A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna)*, AustriaURL http://www. R-project. org, 2007.

[89] T. Officials, "AASHTO LRFD Bridge Design Specifications: Customary US Units," American Association of State highway and transportation officials., Washington, DC, 1994.

[90] L. AASHTO, "Bridge Design Specifications," AASHTO, Washington, DC, 1998.

[91] L. AASHTO, "AASHTO LRFD bridge design specifications. Transportation (Amst).," American Association of State Highway and Transportation Officials, Washington, DC., 2007.

[92] A. S. Nowak, "Calibration of LRFD bridge design code," Transportation Research Board, Washington, DC, 1999.

[93] J. Withiam, E. Voytko, R. Barker, M. Duncan, B. Kelly, S. Musser and V. Elias, "Load and Resistance Factor Design (LRFD) of Highway Bridge Substructures," Federal Highway Administration, Washington, DC, 1998.

[94] M. McVay, R. Ellis Jr, B. Birgisson, G. Consolazio, S. Putcha and S. Lee, "Load and resistance factor design, cost, and risk: Designing a drilled-shaft load test program in Florida limestone," *Transportation Research Record*, vol. 1849, pp. 98-106, 2003.

[95] M. Y. Abu-Farsakh, S. Yoon and C. Tsai, "Calibration of resistance factors needed in the LRFD design of driven piles. Retrieved from," Louisiana Transportation Research Center, Baton Rouge, LA, 2009.

[96] R. M. Barker, J. M. Duncan, K. B. Rojiani, P. S. Ooi, C. K. Tan and S. Kim, "Manuals for the design of bridge foundations: Shallow foundations, driven piles, retaining walls and abutments, drilled shafts, estimating tolerable movements, and load factor design specifications and commentary. 1991.," NCHRP Report No. 343, Transportation Research Board, Washington, DC, 1991.

[97] A. Hasofer, "An Exact and Invariant First Order Reliability Format," *Journal of Engineering Mechanics,* vol. 100, no. 1, pp. 111-121, 1974.

[98] A. S. Nowak and K. R. Collins, Reliability of structures, Boca Raton, FL: CRC Press, 2012.

[99] R. Rackwitz and B. Flessler, "Structural reliability under combined random load sequences," *Computers & Structures,* vol. 9, no. 5, pp. 489-494, 1978.

[100] G. D. Garson, "Interpreting Neural Network Connection Weights," *AI Expert,* vol. 6, no. 4, pp. 46-51, 1991.

[101] J.-L. Briaud, L. M. Tucker, J. S. Anderson, D. Perdomo and H. M. Coyle, "Development of an improved Design Procedure for Single Piles in Clays and Sands," Mississippi State Highway Department, Report No. MSHDRD-RD-86-050-1, Jackson, MS, 1986.

[102] S. Lacasse, F. Nadim, K. H. Andersen, S. Knudsen, U. K. Eidsvig, G. L. Yetginer, T. Guttormsen and A. Eide, "Reliability of API, NGI, ICP and Fugro axial pile capacity calculation methods," in *Offshore Technology Conference*, Houston, Texas, 2013.

[103] P. K. Robertson and R. G. Campanella, "Interpretation of Cone Penetration Tests - Part II (Clay)," *Canadian Geotechnical Journal*, vol. 20, no. 4, pp. 735-745, 1983.

[104] H. Kolk, A. Baaijens and M. Senders, "Design criteria for pipe piles in silica sands. in," in *Proc., 1st Int. Symp. on Frontiers in Offshore Geotechnics*, Perth, Australia, 2005.

[105] G. Baldi, R. Bellotti, V. N. Ghionna, M. Jamiolkowski and D. C. LoPresti, "Modulus of Sands from CPT's and DMT's," in *Proceedings of the 12th ICSMFE conference*, Rio de Janeiro, Brazil, 1989.

[106] H. G. Kempfert and P. Becker, "Grundlagen und Ergebnisse der Ableitung von axialen Pfahlwiderständen aus Erfahrungs-werten für die EA-Pfähle," *Bautechnik*, vol. 87, no. 7, pp. 441-449, 2007.

[107] B. EN, "Eurocode 7: Geotechnical design-Part 1: General rules," British Standards, London, UK, 2004.

[108] BSI, "Eurocode 7- Geotechnical design—Part 2: Ground Investigation and testing," BSI London, London, UK, 2007.

[109] F. De Cock, "Design of Axially Loaded Piles: European Practice," in *Proceedings of an ERTC-3 Seminar*, Brussels, 1997.

Appendix A

Pile-CPT Methods

Schmertmann Method

Schmertmann proposed the following relationship to predict the unit end bearing capacity of the pile (q_b) from the cone tip resistance (q_c) [1]:

$$q_b = \frac{q_{c1} + q_{c2}}{2} \tag{A.1}$$

where, q_{c1} is the minimum of the average cone tip resistances of zones ranging from *0.7D to 4D* below the pile tip (where *D* is the pile diameter) and q_{c2} is the average of minimum cone tip resistances over a distance *8D* above the pile tip. To determine q_{c1}, the minimum path rule is used as illustrated in Figure 52. The described zone (from *8D* above to *0.7D-4D* below the pile tip) represents the failure surface, which is approximated by a logarithmic spiral. Schmertmann suggested an upper limit of 150 TSF (15 MPa) for the unit tip bearing capacity (q_b).

According to Schmertmann method, the unit skin friction of the pile (f) is given by:

$$f = \alpha_c f_s \tag{A.2}$$

where, α_c is a reduction factor, which varies from 0.2 to 1.25 for clayey soil, and f_s is the sleeve friction. Figure 53 depicts the variation of α_c with f_s for different pile types in clay.

For piles in sand, the friction capacity (Q_s) is obtained by:

$$Q_s = \alpha_s \left(\sum_{y=0}^{8D} \frac{y}{8D} f_s A_s + \sum_{y=8D}^{L} f_s A_s \right) \tag{A.3}$$

where, α_s is the correction factor for sand, which can be obtained from Figure 54, y is the depth at which side resistance is calculated, and L is the pile length.

Schmertmann suggested a limit of 1.2 TSF (120 kPa) on f.

Figure 52. Calculation of the average cone tip resistance in Schmertmann method [1]

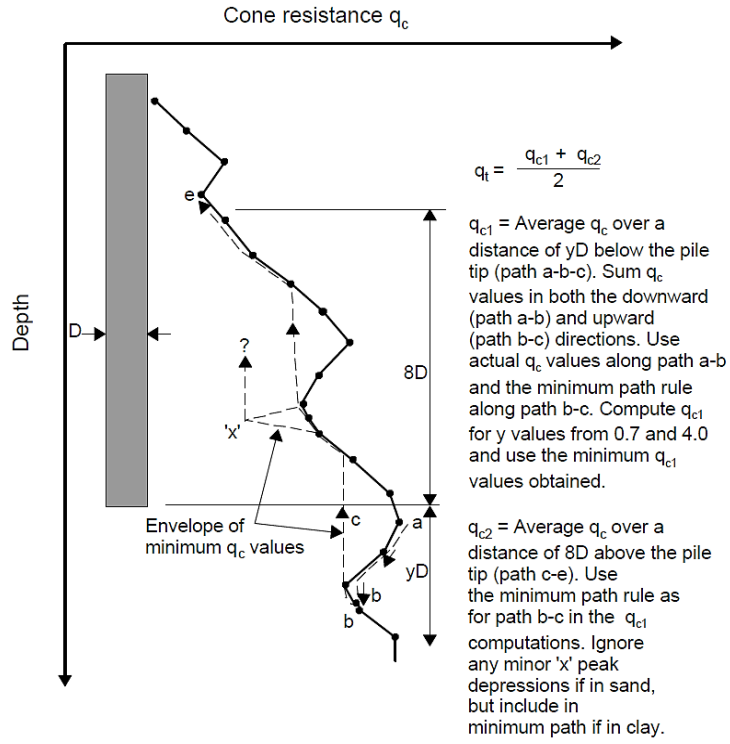


Figure 53. Penetration design curves for pile side friction in clay in Schmertmann method

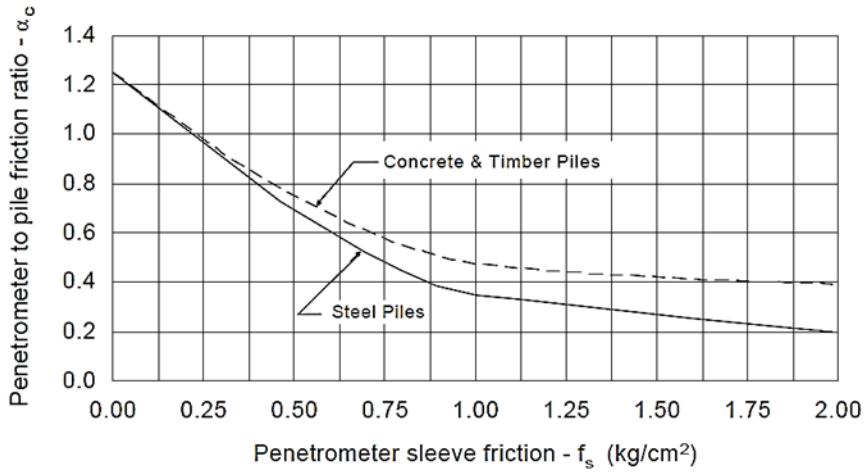
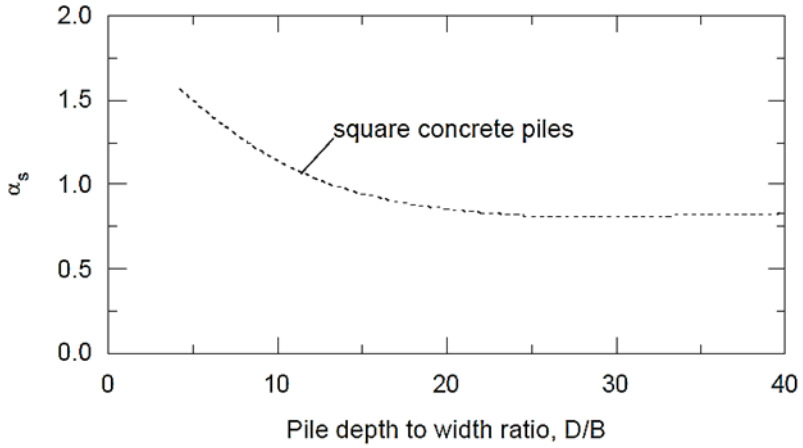


Figure 54. Penetration design curves for pile side friction in sand in Schmertmann method



For implementing Schmertmann method in the code, using the Probabilistic soil classification:

If $(\%\text{clay} + \%\text{silt}) > 0.8$

$$f = (\alpha_c) \times f_s$$

Otherwise:

$$f = \left((\%\text{clay} + \%\text{silt}) \times \alpha_c + \%\text{sand} \times \alpha_s\right) \times f_s$$

Using Robertson-2010 soil classification:

$$f = (\alpha_s \times \alpha 1 + \alpha_c \times \beta 1) \times f_s$$

where, α and β are obtained from the Table 18:

Table 18. α1 and β1 values for calculating unit side resistance for Schmertmann method

Soil index	α_1	□₁
1	0	1
2	0	1
3	0	1
4	0	1
5	2/3	1/3
6	1	0
7	1	0
8	2/3	1/3
9	0	1

De Ruiter and Beringen Method

This method is proposed by De Ruiter and Beringen and is based on the experience gained in the North Sea [2]. This method is also known as the European method and uses different procedures for clay and sand.

In clay, the undrained shear strength (S_u) for each soil layer is first evaluated from the cone tip resistance (q_c). Then, the unit tip bearing capacity and the unit skin friction are computed by applying suitable multiplying factors. The unit tip bearing capacity is given by:

$$q_{p-clay} = N_c S_{u(tip)}$$
$$S_{u(tip)} = \frac{q_{c(tip)}}{N_k} \quad \text{(A.4)}$$

where, N_c is the bearing capacity factor and N_c=9 is considered by this method. N_k is the cone factor that ranges from 15 to 20, depending on the local experience. $q_c(tip)$ is the average of cone tip resistances around the pile tip computed similar to Schmertmann method.

The unit skin friction is given by:

$$f_{clay} = \beta S_{u(side)} \quad \text{(A.5)}$$

where, β is the adhesion factor, β=1 for normally consolidated (NC) clay, and β =0.5 for overconsolidated (OC) clay. $S_u(side)$, the undrained shear strength for each soil layer along the pile shaft, is determined by:

$$S_{u(side)} = \frac{q_{c(side)}}{N_k} \quad \text{(A.6)}$$

where, $q_c(side)$ is the average cone tip resistance along the soil layer.

In the current study, the cone factor N_k=20 and the adhesion factor β =0.5 were adopted in the analysis, since these values gave better predicted ultimate pile capacity for the investigated piles.

In sand, the unit tip bearing capacity of the pile (q_{p-sand}) is calculated similar to Schmertmann method. The unit skin friction (f) for each soil layer along the pile shaft is given by:

$$f_{sand} = \min \begin{cases} f_s \\ \frac{q_{c(side)}}{300} \text{ (compression)} \\ \frac{q_{c(side)}}{400} \text{ (tension)} \end{cases} \quad \text{(A.7)}$$

De Ruiter and Beringen imposed limits on q_p and f in which $q_p \leq 150$ TSF (15 MPa) and $f \leq 1.2$ TSF (120 kPa).

Implementing Probabilistic soil classification into this method is as follows:

If $(\%clay + \%silt) > 0.8$

$$q_b = q_{b-clay}$$
$$f = f_{clay}$$

Otherwise:

$$q_b = (\%clay + \%silt) \times q_{b-clay} + \%sand \times q_{b-sand}$$
$$f = (\%sand + \%silt) \times f_{sand} + \%clay \times f_{clay}$$

For Robertson-2010 soil classification:

$$q_b = q_{b-sand} \times \alpha 1 + q_{b-clay} \times \beta 1$$
$$f = f_{sand} \times \alpha 1 + f_{clay} \times \beta 1$$

where, $\alpha 1$ and $\beta 1$ are obtained from Table 18.

Bustamante and Gianeselli Method (LCPC Method)

Bustamante and Gianeselli [3] proposed this method for the French Highway Department based on the analysis of 197 pile load tests with a variety of pile types and soil conditions. It is also known as the French method and the LCPC (Laboratoire Central des Ponts et Chaussees) method. In this method, both the unit tip bearing capacity (q_b) and the unit skin friction (f) of the pile are obtained from the cone tip resistance (q_c). The sleeve friction (f_s) is not used. The unit tip bearing capacity of the pile (q_b) is predicted from the following equation:

$$q_b = k_b q_{eq\,(tip)} \quad \text{(A.8)}$$

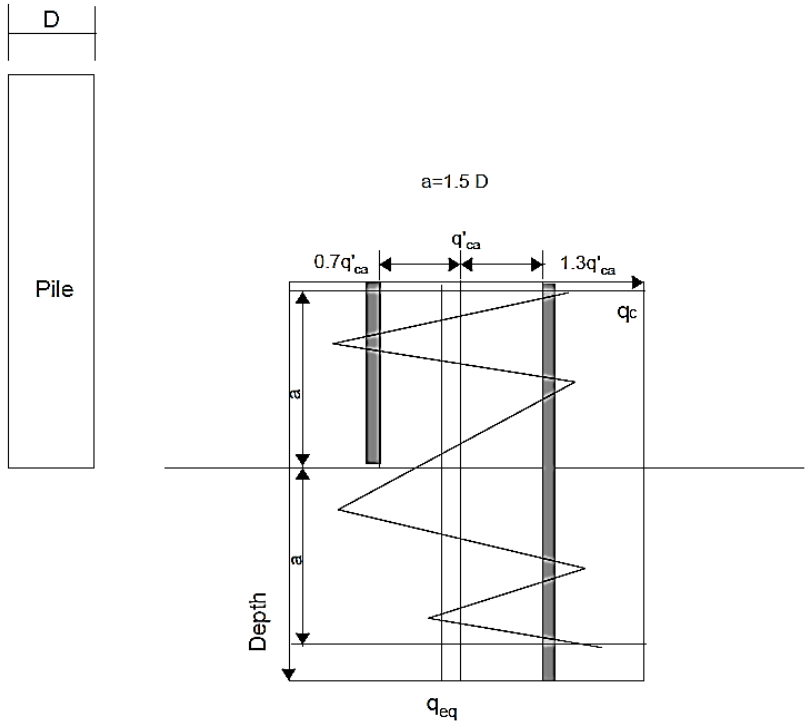
where, k_b is an empirical bearing capacity factor that varies from 0.15 to 0.60 depending on the soil type and pile installation procedure (Table 19) and $q_{eq}(tip)$ is the equivalent average cone tip resistance around the pile tip, which is obtained as follows:

1) Calculate the average tip resistance (q_{ca}) at the tip of the pile by averaging q_c values over a zone ranging from 1.5D below the pile tip to 1.5D above the pile tip (D is the pile diameter);
2) Eliminate q_c values in the zone that are higher than 1.3q_{ca} and those are lower than 0.7q_{ca} as shown in Figure 55; and
3) Calculate the equivalent average cone tip resistance ($q_{eq}(tip)$) by averaging the remaining cone tip resistance (q_c) values over the same zone (bordered by thick lines in Figure 55).

Table 19. LCPC bearing capacity factor (k_b)

Soil Type	Bored Piles	Driven Piles
Clay-Silt	0.375	0.60
Sand-Gravel	0.15	0.375
Chalk	0.20	0.40

Figure 55. Calculation of the equivalent average tip resistance for LCPC method (after Bustamante and Gianeselli [3])



The pile unit skin friction (*f*) in each soil layer is estimated from the equivalent cone tip resistance ($q_{eq}(side)$) of the soil layer:

$$f = \frac{q_{eq(side)}}{K_s} < f_{max} \quad (A.9)$$

where, K_s is obtained from Table 20; and the categories are determined from the Table 21.

Table 20. Values of K_s for LCPC method

Nature of the soil	q_c (10^5 Pa)	Coefficient K_s category			
		I		II	
		IA	IB	IIA	IIB
Soft clay and mud	<10	30	30	30	30
Moderately compact clay	10 to 50	40	80	40	80
Silt and loose sand	<50	60	150	60	120
Compact to stiff clay and compact silt	>50	60	120	60	120
soft chalk	<50	100	120	100	120
moderately compact sand and gravel	50 to 120	100	200	100	200
Weathered to fragmented chalk	>50	60	80	60	80
compact to very compact sand and gravel	>120	150	300	150	200

Table 21. Pile category for LCPC method

Pile Category	Type of the pile
IA	Plain bored piles, mud bored piles, hollow auger bored piles, case screwed piles
IB	Cased bored piles, driven cast piles
IIA	Driven precast piles, prestressed tubular piles, jacked concrete piles
IIB	Driven steel piles, jacked steel piles
IIIA	Driven grouted piles, driven rammed piles
IIIB	High pressure grouted piles (d>0.25 m), Type II micropiles

The following procedure explains how to determine the maximum unit skin friction (f_{max}), which is dependent on soil type, pile type, and installation procedure:

A. Based on the pile type, select the pile category from Table 22 (for example, pile category is 9 for square PPC piles),
B. For each soil layer, select the appropriate curve number (Table 23 and Table 24) based on soil type, equivalent cone tip resistance along the soil layer and ($q_{eq}(side)$), and pile category, use Table 23 for clay and silt and Figure 56 for sand and gravel,

C. From Figure 56, use the selected curve number and the equivalent cone tip resistance ($q_{eq}(side)$) to obtain the maximum unit skin friction (f_{max}), use Figure 56a for clay and silt and Figure 56b for sand and gravel.

Figure 56. Maximum friction curves for LCPC method (after Briaud [101])

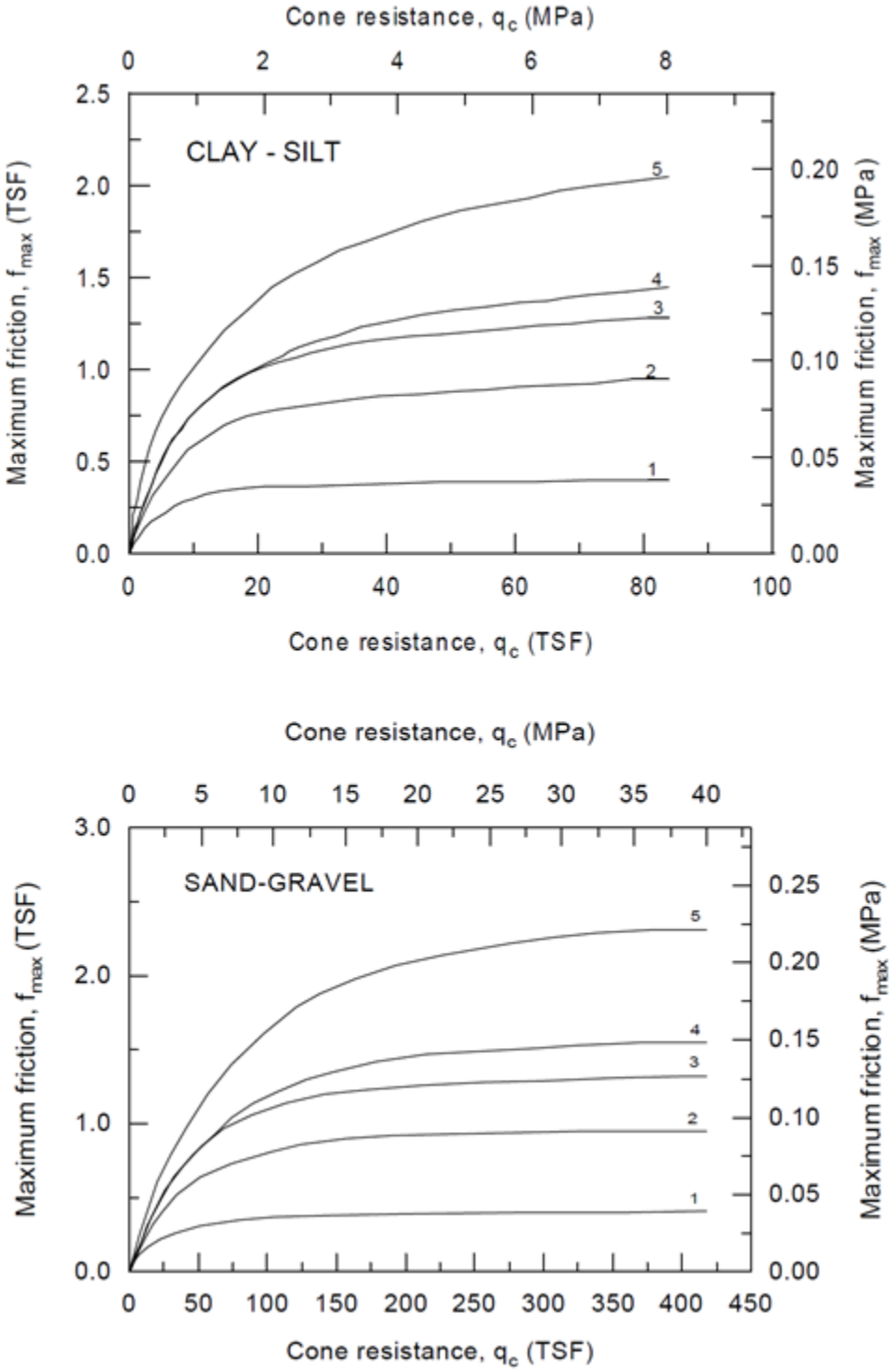


Table 22. Pile categories for LCPC method

1. FS Drilled shaft with no drilling mud	Installed without supporting the soil with drilling mud. Applicable only for cohesive soils above the water table.
2. FB Drilled shaft with drilling mud	Installed using mud to support the sides of the whole. Concrete is poured from the bottom up, displacing the mud.
3. FT Drilled shaft with casing (FTU)	Drilled within the confinement of a steel casing. As the casting is retrieved, concrete is poured in the hole.
4. FTC Drilled shaft, hollow auger (auger cast piles)	Installed using a hollow stem continuous auger having a length at least equal to the proposed pile length. The auger is extracted without turning while, simultaneously, concrete is injected through the auger stem.
5. FPU Pier	Hand excavated foundations. The drilling method requires the presence of workers at the bottom of the excavation. The sides are supported with retaining elements or casing.
6. FIG Micropile type1 (BIG)	Drilled pile with casting. Diameter less than 250 mm (10 inch). After the casting has been filled with concrete, the top of the casing is plugged. Pressure is applied inside the casting between the concrete and the plug. The casing is recovered by maintaining the pressure against the concrete.
7. VMO Screwed-in piles	Not applicable for cohesionless or soils below water table. A screw type tool is placed in front of a corrugated pipe which is pushed and screwed in place. The rotation is reversed for pulling out the casting while concrete is poured.
8. BE Driven piles, concrete coated	- pipe piles 150 mm (6 in.) To 500 mm (20 in.) External diameter - H piles - caissons made of 2, 3, or 4 sheet pile sections. The pile is driven with an oversized protecting shoe. As driving proceeds, concrete is injected through a hose near the oversized shoe producing a coating around the pile.
9. BBA Driven prefabricated piles	Reinforced or prestressed concrete piles installed by driving or vibrodriving.
10. BM Steel driven piles	Piles made of steel only and driven in place. - H piles - Pipe piles - any shape obtained by welding sheet-pile sections.
11. BPR Prestressed tube pile	Made of hollow cylinder elements of lightly reinforced concrete assembled together by prestressing before driving. Each element is generally 1.5 to 3 m (4-9 ft) long and 0.7 to 0.9 m (2-3 ft) in diameter; the thickness is approximately 0.15 m (6 in.). The piles are driven open ended.
12. BFR Driven pile, bottom concrete plug	Driving is achieved through the bottom concrete plug. The casting is pulled out while low slump concrete is compacted in it.
13. BMO Driven pile, molded.	A plugged tube is driven until the final position is reached. The tube is filled with medium slump concrete to the top and the tube is extracted.
14. VBA Concrete piles, pushed-in.	Pile is made of cylindrical concrete elements prefabricated or cast-in-place, 0.5 to 2.5 m (1.5 to 8 ft) long and 30 to 60 cm (1 to 2 ft) in diameter. The elements are pushed in by a hydraulic jack.
15. VME Steel piles, pushed-in	Piles made of steel only are pushed in by a hydraulic jack..
16. FIP Micropile type II	Drilled pile < 250 mm (10 in.) In diameter. The reinforcing cage is placed in the hole and concrete placed from bottom up.
17. BIP High pressure injected pile, large diameter	Diameter > 250 mm (10 in.). The injection system should be able to produce high pressures.

Table 23. Input parameters for clay and silt for LCPC method

CURVE #	q_c (ksf)	PILE TYPE (see Table 2)	COMMENTS ON INSERTION PROCEDURE
1	< 14.6	1-17	
	> 14.6	1,2	- very probable values when using tools without teeth or with oversized blades and where a remoulded layer of material can be deposited along the sides of the drilled hole. Use these values also for deep holes below the water table where the hole must be cleaned several times. Use these values also for cases when the relaxation of the sides of the hole is allowed due to incidents slowing or stopping the pouring of concrete. For all the previous conditions, experience shows, however, that q_s can be between curves 1 and 2; use an intermediate value of q_s is such value is warranted by a load test.
2	> 25.1	4, 5, 8, 9, 10, 11, 13, 14, 15	- for all steel piles , experience shows that, in plastic soils, q_s is often as low as curve 1; therefore, use curve 1 when no previous load test is available. For all driven concrete piles use curve 3 in low plasticity soils with sand or sand and gravel layers or containing boulders and when q_c>52.2 ksf.
	> 25.1	7	- use these values for soils where q_c<52.2 ksf and the rate of penetration is slow; otherwise use curve 1. Also for slow penetration, when q_c>93.9 ksf, use curve 3. - use curve 3 based on previous load test.
	> 25.1	6	- use these values when careful method of drilling with an auger equipped with teeth and immediate concrete pouring is used. In the case of constant supervision with cleaning and grooving of the borehole walls followed by immediate concrete pouring, for soils of q_c>93.9 ksf, curve 3 can be used.
	> 25.1	1, 2	
	> 25.1	3	- for dry holes. It is recommended to vibrate the concrete after taking out the casing. In the case of work below the water table, where pumping is required and frequent movement of the casing is necessary, use curve 1 unless load test results are available.
3	> 25.1 < 41.8	12	- usual conditions of execution as described in DTU 13.2
5	> 14.8	16, 17	- in the case of injection done selectively and repetitively at low flow rate it will be possible to use curve 5, if it is justified by previous load test.

Table 24. Input parameters for sand and gravel for LCPC method

CURVE #	q_c (ksf)	PILE TYPE (see Table 2)	COMMENTS ON INSERTION PROCEDURE
1	< 73.1	2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	
2	> 73.1	6, 7, 9, 10, 11, 12, 13, 14, 15	- for fine sands. Since steel piles can lead to very small values of q_s in such soils, use curve 1 unless higher values can be based on load test results. For concrete piles, use curve 2 for fine sands of q_c>156.6 ksf.
	>104.4	2, 3	- only for fine sands and bored piles which are less than 30 m (100 ft) long. For piles longer than 30 m (100 ft) in fine sand, q_s may vary between curves 1 and 2. Where no load test data is available, use curve 1.
	>104.4	4	- reserved for sands exhibiting some cohesion.
3	> 156.6	6, 7, 9, 10, 11, 13, 14, 15, 17	- for coarse gravelly sand or gravel only. For concrete piles, use curve 4 if it can be justified by a load test.
	> 156.6	2, 3	- for coarse gravelly sand or gravel and bored piles less than 30 m (100 ft) long.
			- for gravel where q_c>83.5 ksf, use curve 4
4	> 156.6	8, 12	- for coarse gravelly sand and gravel only.
5	> 104.4	16, 17	- use of values higher than curve 5 is acceptable if based on load test.

Implementing Probabilistic soil classification into LCPC method is described as follows:

$$K_b = 0.375 \times \%\text{sand} + 0.6 \times \%\text{silt} + 0.6 \times \%\text{clay}$$

If sand% + silt% <50%: case 1

$$K_s = 30 \text{ if } q_c < 10 \text{ TSF}$$
$$K_s = 40 \text{ if } q_c < 50 \text{ TSF}$$
$$K_s = 60 \text{ if } q_c \geq 50 \text{ TSF}$$

Otherwise: case 2

$$K_s = 60 \text{ if } q_c < 50 \text{ TSF}$$
$$K_s = 100 \text{ if } q_c < 120 \text{ TSF}$$
$$K_s = 150 \text{ if } q_c \geq 120 \text{ TSF}$$

For f_{max} the values obtained from the curves in Figure 56 used as $f_{max\text{-}clay}$ and $f_{max\text{-}sand}$

If $(\%clay + \%silt) > 0.8$

$$f_{max} = f_{max-clay}$$

Otherwise:

$$f_{max} = (\%clay + \%silt) \times f_{max-clay} + \%sand \times f_{max-sand}$$

For Robertson-2010 soil classification:

K_b and K_s are obtained from Table 25:

Table 25. K_b and K_s for LCPC method

Soil index	k_b	K_s
1	0.6	Case1
2	0.6	Case1
3	0.6	Case1
4	0.6	Case2
5	0.488	Case2
6	0.375	Case2
7	0.375	Case2
8	0.45	Case2
9	0.6	Case1

$$f_{max} = f_{max-sand} \times \alpha 1 + f_{max-clay} \times \beta 1$$

where, $\alpha 1$ and $\beta 1$ are obtained from Table 18.

Tumay and Fakhroo Method (Cone-m Method)

Tumay and Fakhroo proposed this method to predict the ultimate pile capacity of piles in clayey soils [4]. The unit end bearing capacity (q_b) is estimated using a procedure similar to Schmertmann's method as follows:

$$q_b = \frac{q_{c1} + q_{c2}}{4} + \frac{q_a}{2} \quad (A.10)$$

where, q_{c1} is the average of q_c values 4D below the pile tip, q_{c2} is the average of the minimum q_c values 4D below the pile tip, and q_a is the average of the minimum of q_c values 8D above the

pile tip. Tumay and Fakhroo suggested an upper limit of 150 TSF (15 MPa) for the unit pile end bearing capacity (q_b).

The unit skin friction (f) is given by the following expression:

$$f = mf_{sa} \tag{A.11}$$

Tumay and Fakhroo suggested that $f \leq 0.72$ TSF (72 kPa). The adhesion factor (m) is expressed as:

$$m = 0.5 + 9.5e^{-9f_{sa}} \tag{A.12}$$

where, $f_{sa}=F_t/L$ is the average local friction in TSF, and F_t is the total cone penetration friction determined for pile penetration length (L).

Aoki and De Alencar Method

Aoki and De Alencar Velloso proposed the following method to estimate the ultimate load carrying capacity of the pile from CPT data [5]. The unit end bearing capacity (q_b) is obtained from:

$$q_b = \frac{q_{ca(tip)}}{F_b} \tag{A.13}$$

where, $q_{ca}(tip)$ is the average cone tip resistance around the pile tip, and F_b is an empirical factor that depends on the pile type. The unit skin friction of the pile (f) is predicted by:

$$f = q_{c(side)} \frac{\alpha_s}{F_s} \tag{A.14}$$

where, $q_c(side)$ is the average cone tip resistance for each soil layer along the pile shaft, F_s is an empirical factor that depends on the pile type and α_s is an empirical factor that depends on the soil type. Factors F_b and F_s are given in Table 26. The values of the empirical factor α_s are presented in Table 27.

Table 26. Empirical factors F_b and F_s

Pile Type	F_b	F_s
Bored	3.5	7.0
Franki	2.5	5.0
Steel	1.75	3.5
Precast concrete	1.75	3.5

Table 27: Empirical factor α_s values for different soil types

Soil Type	α_s (%)	Soil Type	α_s (%)	Soil Type	α_s (%)
Sand	1.4	Sandy silt	2.2	Sandy clay	2.4
Silty sand	2.0	Sandy silt with clay	2.8	Sandy clay with silt	2.8
Silty sand with clay	2.4	Silt	3.0	Silt clay with sand	3.0
Clayey sand with silt	2.8	Clayey silt with sand	3.0	Silty clay	4.0
Clayey sand	3.0	Clayey silt	3.4	Clay	6.0

Upper limits were imposed on q_p and f as follows: $q_p \leq 150$ TSF (15 MPa) and $f \leq 1.2$ TSF (120 kPa).

α_s has been implemented using the following equations:

For probabilistic method:

If $(\%\text{clay}) > 0.8$: $\alpha_s = 6\%$

If $(\%\text{sand}) > 0.8$: $\alpha_s = 1.4\%$

Otherwise: $\alpha_s = (1.4 \times \%\text{sand} + 3 \times \%\text{silt} + 6 \times \%\text{clay})/100$

For Robertson-2010 method, the following values of α_s versus soil index are given in Table 28.

Table 28: Values of α_s versus soil index for Robertson-2010

Soil index	α_s (%)
1	6
2	6
3	5
4	3.7
5	2.1
6	1.7
7	1.4
8	2.5
9	4

Price and Wardle Method

Price and Wardle proposed the following relationship to evaluate the unit end bearing capacity (q_b) of the pile from the cone tip resistance [17]:

$$q_b = k_b q_c \tag{A.15}$$

where, k_b is a factor depends on the pile type (k_b = 0.35 for driven piles and 0.3 for jacked piles). (For influence zone, no specific hint has been introduced, therefore normal average 3D above and below the tip was used for q_c)

The unit skin friction (f) is obtained from:

$$f = k_s f_s \tag{A.16}$$

where, k_s is a factor depends on the pile type (k_s=0.53 for driven piles, 0.62 for jacked piles, and 0.49 for bored piles). Price and Wardle proposed the values for these factors based on analysis conducted on pile load tests in stiff clay (London clay).

Upper limits were imposed on q_p and f as follows: $q_p \leq 150$ TSF (15 MPa) and $f \leq 1.2$ TSF (120 kPa).

Pilipponnat Method

Philipponnat proposed the following expression to estimate the unit end bearing capacity of the pile (q_b) from the cone tip resistance (q_c) [7]:

$$q_b = k_b q_{ca} \tag{A.17}$$

where, k_b is a factor that depends on the soil type as shown in Table 29. The cone tip resistance (q_{ca}) is averaged as follows:

$$q_{ca} = \frac{q_{ca\,(A)} + q_{ca\,(B)}}{2} \tag{A.18}$$

where, $q_{ca(A)}$ is the average cone tip resistance within $3B$ (B is the pile width) above the pile tip and $q_{cb(B)}$ is the average cone tip resistance within $3B$ below the pile tip. Philipponnat recommended the removal of the extreme peaks (spikes) when the tip resistance profiles is irregular and imposed a condition in which $q_{ca(A)} \leq q_{cb(B)}$.

The unit skin friction of the pile (f) is determined by:

$$f = \frac{\alpha_s}{F_s} q_{cs} \tag{A.19}$$

where, q_{cs} is the average cone tip resistance for each soil layer along the pile shaft, F_s is a factor depends on the soil type as presented in Table 30. The factor α_s depends on the pile type where α_s equals to 1.25 for precast concrete driven piles. Philipponnat suggested an upper limit for the skin friction (f_{lim}), for precast concrete driven piles $f_{lim} \leq 1.2\ P_A$ (P_A is the atmospheric pressure).

Table 29. Bearing capacity factor (k_b)

Soil Type	k_b
Gravel	0.35
Sand	0.40
Silt	0.45
Clay	0.50

Table 30. Empirical factor F_s

Soil Type	F_s
Clay and calcareous clay	50
Silt, sandy clay, and clayey sand	60
Loose sand	100
Medium dense sand	150
Dense sand and gravel	200

K_b and F_s have been implemented using the following values:

For probabilistic method:

If $(\%\text{clay}) > 0.8$: $K_b = 0.5$ and $F_s = 50$

If $(\%\text{sand}) > 0.8$: $K_b = 0.4$ and $F_s = F_{s-sand}$

Otherwise:

$$K_b = 0.4 \times \%\text{sand} + 0.45 \times \%\text{silt} + 0.5 \times \%\text{clay}$$

$$F_s = F_{s-sand} \times \%\text{sand} + 60 \times \%\text{silt} + 60 \times \%\text{clay}$$

F_{s-sand} is 100, 150, or 200 depending on the value of D_r, which is described for UF method.

If $D_r < 0.4$, $F_{s-sand} = 100$

If $D_r > 0.7$, $F_{s-sand} = 200$

For other cases: $F_{s-sand} = 150$

For Robertson-2010 method, the values of k_p and F_s versus soil index are given in Table 31.

Table 31. Values of *kp* and F_s for Robertson-2010

Soil index	k_b	F_s
1	0.5	50
2	0.5	50
3	0.485	52
4	0.475	55
5	0.425	F_{s-sand}
6	0.4	F_{s-sand}
7	0.375	F_{s-sand}
8	0.43	60
9	0.45	55

Penpile Method

The penpile method was proposed by Clisby et al. for the Mississippi Department of Transportation [8]. The unit end bearing capacity of the pile (q_b) is determined from the following relationship:

$$q_b = \begin{cases} 0.25q_c \text{ for pile tip in clay} \\ 0.125q_c \text{ for pile tip in sand} \end{cases} \quad \text{(A.20)}$$

where, q_c is the average of three cone tip resistances close to the pile tip. In this study, normal average 3D above and below the tip has been used for the influence zone.

The unit skin friction of the pile shaft (f) is obtained from the following relationship:

$$f = \frac{f_s}{1.5 + 0.1f_s} \quad \text{(A.21)}$$

where, f is expressed in psi (lb/in^2) and f_s is the sleeve friction of the cone expressed in psi.

For implementing the unit tip resistance, using Probabilistic and Robertson-2010 soil classifications:

$$q_b = \alpha \times q_c$$

For probabilistic method:

If $(\%clay) > 0.8$: $\alpha = 0.25$

If $(\%sand) > 0.8$: $\alpha = 0.125$

Otherwise: $\alpha = 0.125 \times (\%sand) + 0.25 \times (\%clay + \%silt)$

For Robertson-2010 method, the values of α versus soil index are given in Table 32.

Table 32. Values of α for Robertson-2010

Soil index	α
1	0.25
2	0.25
3	0.25
4	0.25
5	0.167
6	0.125
7	0.125
8	0.167
9	0.25

No max limits for tip and side resistance have been proposed by this method.

NGI Method

NGI method is for Norwegian Geotechnical Institute established by Clausen et al., which is based on 121 individual pile tests from 47 different sites with clayey soil with depths from 5 to 110 meters and ${S_u}/{P_0}$ values from 0.2 to 10 and 85 individual piles from 30 different sites with sandy soil with depths from 5 to 40 meters and D_r from 30% to 90% ([11], [27]).

The unit end bearing capacity in sands for close end piles is obtained by:

$$q_{b-sand} = \frac{0.8\, q_c}{1 + {D_r}^2} \tag{A.22}$$

where:

$$D_r = 0.4 \ln \frac{q_c}{22\sqrt{\sigma'_{vo}\, \sigma_{atm}}} \tag{A.13}$$

For open end piles, q_b is determined by the minimum of the plugged and unplugged values:

$$q_{Plugged} = \frac{0.7\, q_c}{1 + 3\, {D_r}^2} \tag{A.24}$$

$$q_{Unplugged} = q_c A_r + \frac{12}{d}(1 - A_r) \int_0^L f dz \tag{A.25}$$

where:

$$A_r = \frac{d^2 - {d_i}^2}{d^2} \tag{A.26}$$

And f is the unit skin friction which the method gives the method for calculating it.

$$f_{sand} = \frac{z}{z_{tip}} \sigma_{atm} F_{D_r} F_{sig} F_{tip} F_{load} F_{mat} > 0.1 \sigma'_{vo} \tag{A.27}$$

where:

$$F_{D_r} = 2.1 (D_r - 0.1)^{1.7} \geq 0 \tag{A.28}$$

$$F_{sig} = (\sigma'_{vo}/\sigma_{atm})^{0.25} \tag{A.29}$$

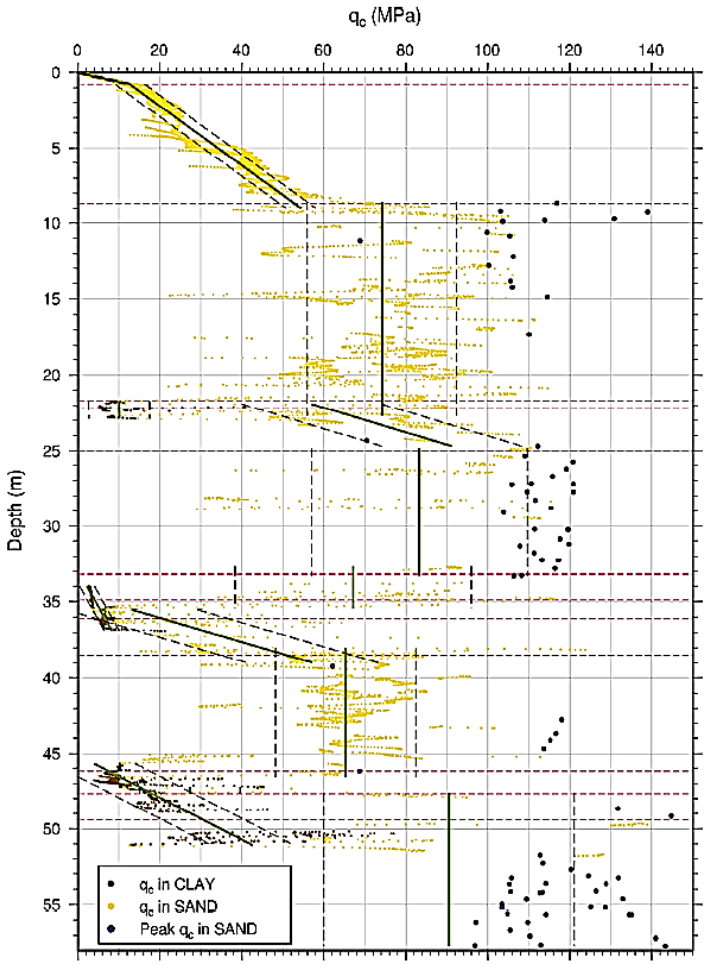
$$F_{tip} = 1.0 \quad \text{For OE \& 1.6 For CE} \tag{A.30}$$

$$F_{load} = 1.0 \text{ For tension \& 1.3 For compression} \tag{A.31}$$

$$F_{mat} = 1.0 \text{ For steel \& 1.2 For concrete} \tag{A.32}$$

A process for statistical treatment of CPTu data in dense sand has been illustrated by Lacasse et al. where the soil profile is divided into distinctive layers and the q_c-values considered to be constant or increasing with depth in each layer [102]. An example of this procedure is shown in Figure 58.

Figure 57. The mean of q_c-values with one standard deviation after filtering, re-sampling and merging (after Lacasse et al. [102])



The NGI method recommends the use of α method in API (American Petroleum Institute) with some corrections for estimating the unit side friction in clayey soils, which is basically counted as an indirect method. In NGI method, the unit tip resistance in clays is the same as α method. No specific recommendation about the influence zone and averaging the value of q_t has described in the method.

Karlsrud et al. has described the NGI approach for pile capacity in clays based on an indirect method [11]. S_u is determined from UU triaxial tests and shaft resistance measured 100 days after driving.

$$q_{b-clay} = 9S_u \tag{A.33}$$

The procedure presented in UWA method is suggested if the results of CPTu tests are not available to calculate q_t, which is the corrected cone resistance.

The unit shaft resistance for normally consolidated (NC) clays with $\left({S_u}/{\sigma'_{v0}}\right) < 0.25$ is:

$$f_{clay} = \alpha^{NC} S_u \tag{A.34}$$

where:

$$\alpha^{NC} = 0.32(PI - 10)^{0.3} \tag{A.35}$$

The upper and lower limits for α^{NC} are:

$$0.20 < \alpha^{NC} < 1.0 \tag{A.36}$$

For overconsolidated (OC) clays with $\left({S_u}/{\sigma'_{v0}}\right) > 1.0$, we have:

$$f_{clay} = \alpha S_u F_{tip} \tag{A.37}$$

where:

$$\alpha = 0.5\left({S_u}/{\sigma'_{v0}}\right)^{-0.3} \tag{A.38}$$

And for close end piles:

$$F_{tip} = 0.8 + 0.2\left({S_u}/{\sigma'_{v0}}\right)^{0.5} \tag{A.39}$$

The upper and lower limits for F_{tip} are:

$$1.0 < F_{tip} < 1.25 \tag{A.40}$$

For clays with $0.25 < \left({S_u}/{\sigma'_{v0}}\right) < 1.0$, an interpolation between the above values should be done.

In this study, for calculating α^{NC}, the value of PI was estimated by:

$PI = \max(\frac{N_{kT} - 7.636}{0.285}, 10)$ where $N_{kt} = \frac{q_c - \sigma_{vo}}{S_u}$

No upper limits have been proposed by this method.

ICP Method (MTD Method)

This method was developed at the Imperial College in London by Jardine et al. [12]. They call it the ICP method as an abbreviation to the Imperial College Pile.

The tests have been done in different sites as shown in the Table 33 and later on, data from other locations like Belfast and Mexico City has been added to the database.

Table 33. ICP sites

Site	Soil conditions
1. Canons Park	London Clay: stiff to very stiff, high plasticity, Eocene marine clay; high YSR
2. Cowden	Cowden till: stiff to very stiff, lean, glacial lodgement till; high YSR
3. Bothkennar	Carse clay: soft, high plasticity, moderately organic, Holocene shallow-marine/estuarine clay-silt, lightly cemented: moderate YSR
4. Labenne	Dune sand: loose to medium dense, medium-sized, Holocene; low YSR
5. Pentre	Glacio-lacustrine clay-silt and laminated clays: very soft to firm, low plasticity, low YSR
6. Dunkirk	Marine sand: dense to very dense, shelly medium-sized sand, Flandrian: low YSR

Note: Yield Stress Ratio (YSR) is the apparent OCR

The method presents the following procedures for calculating the bearing capacity of piles in sand, which is defined as the load for d/10 settlement. For the close-ended piles with circular sections:

$$q_{p-sand} = q_{c,avg}\left(1 - 0.5\log\frac{d}{d_{CPT}}\right) > (0.3q_{c,avg} \text{ for piles with } d > 0.90m) \quad (A.45)$$

It should be noted that ICP method was originally developed for tubular piles. The results of load tests on 16 square-piles and 16 H-piles showed that for non-circular piles the tip resistance is:

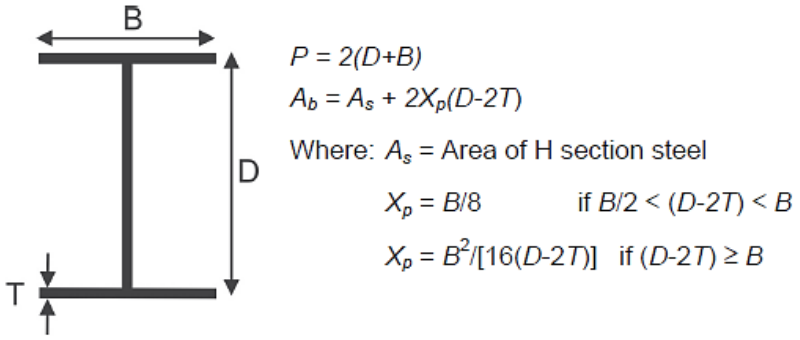
$$q_{b-sand} = 0.7\, q_{c,avg} \quad (A.46)$$

For H-piles the tip resistance is:

$$q_{b-sand} = q_{c,avg} \quad (A.47)$$

And the area for the H-piles (A_b) have been calculated as the procedure shown in Figure 59.

Figure 58. Perimeter and Area for calculating tip and shaft resistance of H-piles



The procedure for calculating $q_{c,avg}$ for ICP method is the same as LCPC method, which takes the average of q_c for 1.5D below and above the tip. For cases that the variation in q_c is extreme and the depth intervals between peak and toughs is greater than d/2, a $q_{c,avg}$ less than the average should be chosen because the base resistance will be dependent on localized failure within the weak layers. Also if weak layers exists at 8d below the pile's tip, the reduction in tip resistance should be considered. However ICP method has not showed how these reductions should be taken into account.

For open-ended piles if the below criteria is satisfied, it can be assumed that pile is plugged:

$$d_{inner}(\text{in meter}) < 0.02(D_r(\%) - 30) \tag{A.48}$$

$$\frac{d_i}{d_{CPT}} < \frac{0.083 q_{c,avg}}{P_{atm}} \tag{A.49}$$

The procedure for calculating the tip resistance for open-ended piles starts with categorizing the piles into the plugged and unplugged piles. For unplugged piles:

$$d_{inner}(\text{in meter}) \geq 0.02(D_r(\%) - 30) \text{ or } \frac{d_i}{d_{CPT}} \geq \frac{0.083 q_{c,avg}}{P_{atm}} \tag{A.50}$$

For unplugged piles:

$$Q_t = q_{c,avg} A_r \tag{A.51}$$

For plugged piles:

$$Q_t = q_{c,avg} \,.\, \max\left(0.5 - 0.25\log\frac{d}{d_{CPT}} \,,\, 0.15\right).\pi R^2{}_{outer} > q_{c,avg} A_r \tag{A.52}$$

where:

$$A_r = \pi(R^2{}_{outer} - R^2{}_{inner}) \tag{A.53}$$

The unit skin friction (f_{sand}) for close-ended circular piles is obtained by the following procedure:

For compression:

$$f_{sand} = (\sigma'_{rc} + \Delta\sigma'_{rd})\tan\delta'_f \tag{A.54}$$

For tension:

$$f_{sand} = (0.8\,\sigma'_{rc} + \Delta\sigma'_{rd})\tan\delta'_f \tag{A.55}$$

where:

σ'_{rc} is the radial effective stress acting on the shaft a few days after installation. $\Delta\sigma'_{rd}$ is the changes in radial effective stress developed during pile loading. In fact, based on ICP method, the radial effective stress acting on the shaft at failure is composed of σ'_{rc} and $\Delta\sigma'_{rd}$. δ'_f is the operational interface angle of friction.

$$\sigma'_{rc} = 0.029 q_c \left(\frac{\sigma'_{v0}}{P_a}\right)^{0.13} \times \left(\max\left[\left(\frac{h}{R}\right), 8\right]\right)^{-0.38} \tag{A.56}$$

$$\Delta\sigma'_{rd} = 2G\frac{\Delta r}{R} \tag{A.57}$$

$$G = q_c(0.0203 + 0.00125\eta - 1.216e^{-6}\eta^2)^{-1} \tag{A.58}$$

$$\eta = \left(\frac{q_c}{P_a}\right) / \left(\frac{\sigma'_{v0}}{P_a}\right)^{0.5} = q_c(P_a\sigma'_{v0})^{-0.5} \tag{A.59}$$

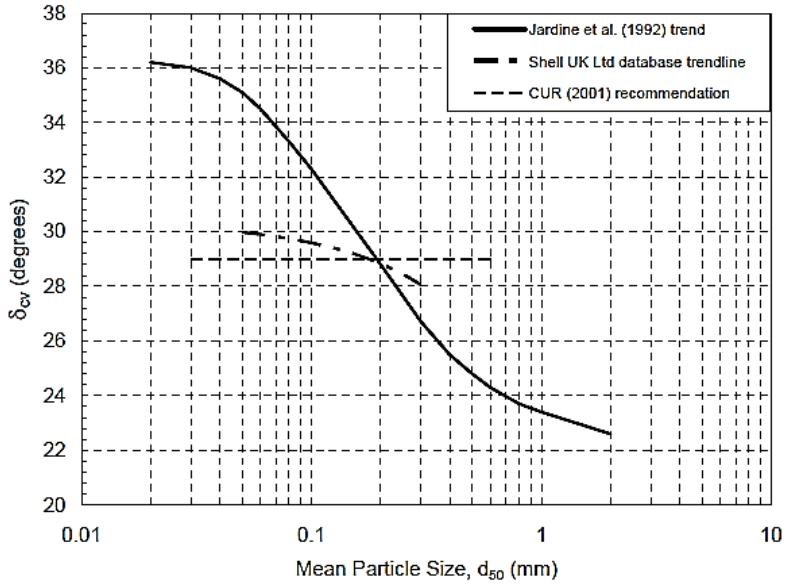
$$P_a = 100\ \text{kPa} \tag{A.60}$$

For lightly rusted steel pile:

$$\Delta r = 2R_{cla} \approx 0.02\text{mm} \tag{A.61}$$

δ'_f can be estimated by direct tests or from Figure 59 as a function of d_{50}

Figure 59. Interface friction angle δ'_f



For non-circular piles (square, rectangular, and H-shaped piles) the R is modified as:

$$R = \left(\frac{A_b}{\pi}\right)^{0.5} \tag{A.62}$$

where, A_b is the section area of the square and rectangular piles, and for H-piles is defined in Figure 59.

For open-ended tubular piles these equations should be modified by defining the value of R is as:

$$R = (R^2{}_{outer} - R^2{}_{inner})^{0.5} \tag{A.63}$$

And for tension:

$$f_{sand} = 0.9(0.8\,\sigma'_{rc} + \Delta\sigma'_{rd})\tan\delta'_f \tag{A.64}$$

The unit tip resistance of tubular piles in clays is:

For undrained loading:

$$q_{p-clay} = 0.8q_{c,avg} \tag{A.65}$$

For drained loading:

$$q_{p-clay} = 1.3q_{c,avg} \tag{A.66}$$

While for the non-circular piles the same rules in sands are applicable which means:

For close end square and rectangular piles:

$$q_{p-clay} = 0.7\, q_{c,avg} \tag{A.67}$$

For H-pile:

$$q_{p-clay} = q_{c,avg} \tag{A.68}$$

While the calculation of A_b for H-piles are shown in

Figure 58.

For calculating the unit tip resistance of open-ended tubular piles in clay, the plugged pile is:

$$\frac{d_{inner}}{d_{CPT}} + 0.45\frac{q_c}{P_a} < 36 \tag{A.69}$$

which gives:

$$q_{p-clay} = 0.4\ q_{c,avg} \text{ for undrained loading} \tag{A.70}$$

$$q_{p-clay} = 0.65\ q_{c,avg} \text{ for drained loading} \tag{A.71}$$

For unplugged piles:

$$q_{p-clay} = q_{c,avg} \text{ for undrained loading} \tag{A.72}$$

$$q_{p-clay} = 1.60\ q_{c,avg} \text{ for drained loading} \tag{A.73}$$

The unit skin friction for the close-ended piles in clay is determined by the following procedure:

$$f_{clay} = \left({}^{K_f}/_{K_c}\right)\sigma'_{rc}\tan\delta'_f \tag{A.74}$$

where, ${}^{K_f}/_{K_c}$ is the loading factor, which is constant and is equal to 0.8 regardless of the loading direction and drainage condition. K_c is the radial to vertical effective stress ration.

$$\sigma'_{rc} = K_c\sigma'_{v0} \tag{A.75}$$

$$K_c = \left[2.2 + 0.016\ YSR - 0.870\ \Delta I_{vy}\right]YSR^{0.42}\left(\max\left[{}^{h}/_{R},\ 8\right]\right)^{-0.20} \tag{A.76}$$

$$\Delta I_{vy} = \log_{10} S_t \tag{A.77}$$

In which the clay sensitivity S_t should be estimated based on the type of the clay.

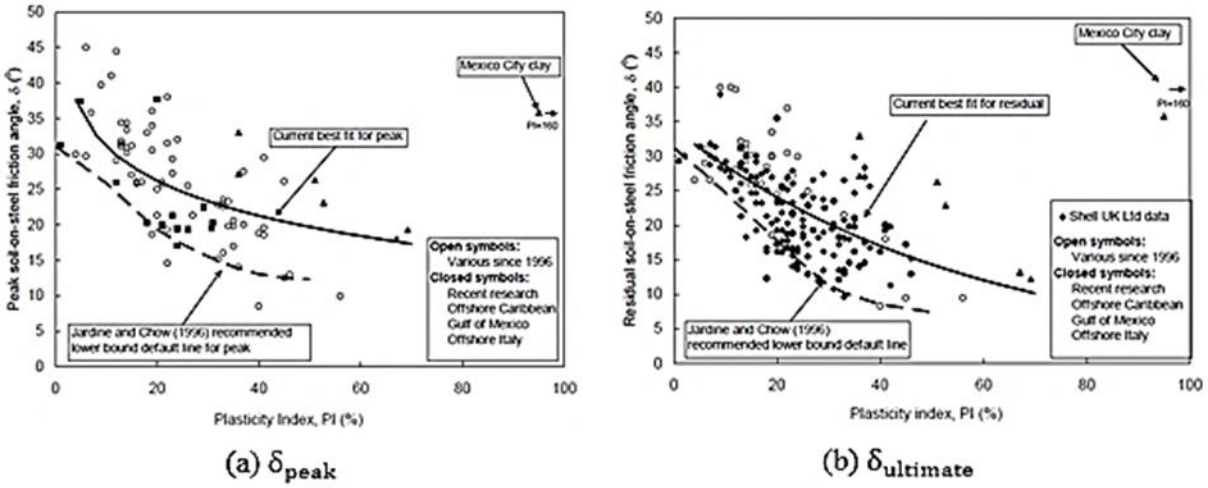
Or:

$$K_c = \left[2 - 0.625\ \Delta I_{vy}\right]YSR^{0.42}\left(\max\left[{}^{h}/_{R},\ 8\right]\right)^{-0.20} \tag{A.78}$$

Using the second one typically leads to lower values for K_c (around 4% less). YSR (also known as apparent OCR) is the clay's local yield stress ratio.

δ'_f is a value between the peak δ_{peak} and ultimate $\delta_{ultimate}$ interface angles of friction, which can be measured in interface ring shear tests. Figure 60 can be used for estimating δ_{peak} and $\delta_{ultimate}$ by using the value of PI.

Figure 60. Ring shear interface results for (a) δ_{peak} and (b) $\delta_{ultimate}$ in clays



The unit skin friction for OE piles in clay is calculated with the same modified R value as shown for the sands.

Implementing Probabilistic soil classification into this method is as follows:

$$q_b = 0.7\, q_{c,avg}$$

If $(\%clay) > 0.8$

$$f = f_{clay}$$

If $(\%sand) > 0.8$

$$f = f_{sand}$$

Otherwise:

$$f = (\%clay + \%silt) \times f_{clay} + \%sand \times f_{sand}$$

For implementing Robertson-2010 soil classification:

$$q_b = q_{p-sand} \times \alpha 1 + q_{p-clay} \times \beta 1$$
$$f = f_{sand} \times \alpha 1 + f_{clay} \times \beta 1$$

Where $\alpha 1$ and $\beta 1$ are obtained from Table 18.

It should be noticed that for calculating δ'_f for f_{sand}, Table 34 can be used.

Table 34. Values of δ'_f versus D_r

D_r	δ'_f
$D_r < 0.2$	15
0.2 < Dr <0.4	$\frac{15 \times \text{silt\%} + 20 * \text{sand\%}}{\text{silt\%} + \text{sand\%}}$
0.4 <Dr <0.6	$\frac{15 \times \text{silt\%} + 25 * \text{sand\%}}{\text{silt\%} + \text{sand\%}}$
0.6 <Dr <0.8	$\frac{20 \times \text{silt\%} + 30 * \text{sand\%}}{\text{silt\%} + \text{sand\%}}$
0.8 < Dr	$\frac{30 \times \text{silt\%} + 35 * \text{sand\%}}{\text{silt\%} + \text{sand\%}}$

The charts presented in Figure 62 show that the value of δ'_f can be estimated based on the Dr < 0.2 (very loose), 0.2 < Dr < 0.4 (loose), 0.4 < Dr < 0.6 (medium), 0.6 < Dr < 0.8 (dense), and 0.8 < Dr (very dense) and the sand% and silt%.

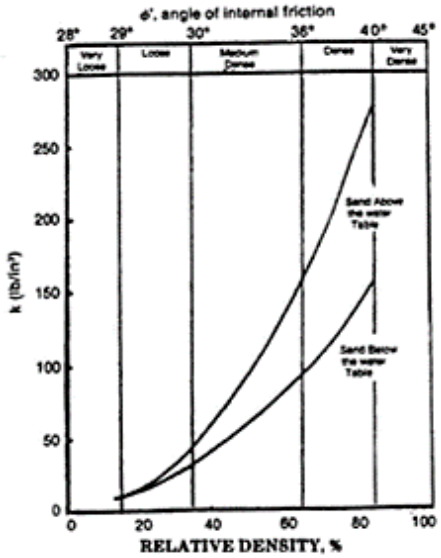
For implementing the above charts, the value of D_r was defined similar to UF method:

$$D_r = \frac{1}{2.41} \ln\left(\frac{95.76\, q_{c-avg}}{157(95.76 * \sigma'_{vo})^{0.55}}\right)$$

where, q_c and σ'_{vo} units are TSF. q_{c-avg} has been calculated for 1D above and below the location and σ'_{vo} has been calculated based on $\gamma_w = \frac{62.43}{2000}$TCF and $\gamma_{soil} = 1.75 \times \gamma_w$

Figure 61. Interface friction angle for silts, sands, and gravel

Density	Soil Description	Soil-Pile Friction angle, δ [degrees]	Limiting Skin Friction, $\tau_{f,lim}$ [kips/ft²]	[kPa]	N_t	Limiting Unit End Bearing, $q_{b,0.10in}$ [kips/ft²]	[kPa]
Very loose Loose Medium	Sand Sand-silt Silt	15	1	47.8	8	40	1.9
Loose Medium Dense	Sand Sand-silt Silt	20	1.4	67	12	60	2.9
Medium Dense	Sand Sand-silt	25	1.7	81.3	20	100	4.8
Dense Very Dense	Sand Sand-silt	30	2	95.7	40	200	9.6
Dense Very Dense	Gravel Sand	35	2.4	114.8	50	250	12



For calculating f_{clay}, S_t and YSR (OCR) can be estimated as flows:

Based on Lehane et al. [13]:

$$\mathrm{YSR} = 0.04427 \left(\frac{q_c}{\sigma'_{v0}}\right)^{1.667}$$

For YSR we observed that Lehane gives very high values so this equation was used:

$$\mathrm{YSR} = 0.152 \frac{q_t - \sigma'_{v0}}{\sigma_{v0}} \geq 1$$

Clay sensitivity, S_t was obtained based on Robertson and Campanella [103]:

$$S_t = \frac{10}{R_f(\%)}$$

Value of PI can be estimated similar to NGI:

$$N_{kt} = 0.285\ \mathrm{PI} + 7.636$$

where:

$$N_{kt} = \frac{q_c - \sigma_{vo}}{S_u}$$

So for clay soils, we estimate PI, and take the average of the below values for calculating δ'_f:

$$\delta_{peak} = 28.4 - 0.177 \times PI$$

$$\delta_{ultimate} = 29.49 - 0.306 \times PI$$

It should be noticed that no upper limits have been proposed by this method.

UWA and CPT-2000 Methods

Lehane et al. at the University of Western Australia (UWA) developed a new method after investigating ICP, NGI, and Fugro pile-CPT methods for close-ended and open-ended piles in siliceous sand [18]. Later on, this method was developed for piles in clay soils by Lehane et al. [29].

Different factors that influence the pile capacity are considered in this method including loading direction, soil dilation, soil displacement, friction fatigue, etc.

The unit end bearing capacity for close-ended piles in sands is:

$$q_{b-sand} = 0.6 q_{c,avg} \tag{A.79}$$

It should be noted that q_t is referred to the ultimate load for 0.1d settlement in the pile, where d is the diameter of the pile. The value of $q_{c,avg}$ is determined by Schmertmann method.

The unit tip resistance for open-ended piles in sands is:

$$q_{b-sand} = q_{c,avg}\left(0.15 + 0.45 A_{rb,eff}\right) \tag{A.80}$$

where:

$$A_{rb,eff} = 1 - FFR\left(\frac{d_i}{d}\right)^2 \tag{A.81}$$

Here FFR (final filling ratio) is the averaged IFR (incremental filling ratio) for the last 3d of the pile penetration, where d is the pile diameter. When FFR approaches to zero, the pile's behavior will be similar to close end pile and when it goes to 1, the pile is equivalent to a fully coring pile. FFR can be estimates as:

$$FFR = \min\left[\left(\frac{d_i\ (m)}{1.5\ (m)}\right)^{0.2}, 1\right] \tag{A.82}$$

The unit shaft friction in sands is:

$$f_{sand} = \frac{f_t}{f_c}(\sigma'_{rc} + \Delta\sigma'_{rd})\tan\delta'_f \tag{A.83}$$

where:

$$\sigma'_{rc} = 0.03 q_c A_{rs,eff}{}^{0.3} \left\{\max\left(\frac{h}{d}, 2\right)\right\}^{-0.5} \tag{A.84}$$

$$\frac{f_t}{f_c} = 1 \text{ For compression } \& \ \frac{f_t}{f_c} = 0.75 \text{ For tension} \tag{A.85}$$

$$A_{rs,eff} = 1 - IFR\left(\frac{d_i}{d}\right)^2 \tag{A.86}$$

$$IFR = \min\left[\left(\frac{d_i\ (m)}{1.5\ (m)}\right)^{0.2}, 1\right] \tag{A.87}$$

$\Delta\sigma'_{rd}$ and δ'_f are calculated as the same as ICP method with:

$$G = 185\ q_{c1N}{}^{-0.75} q_c \tag{A.88}$$

where:

$$q_{c1N} = \left(q_c/P_a\right) \Big/ \left(\sigma'_{v0}/P_a\right)^{0.5} \tag{A.89}$$

δ'_f is calculated as the modified ICP method.

The unit shaft resistance in clays can be determined from [19] which is known as CPT-2000 method:

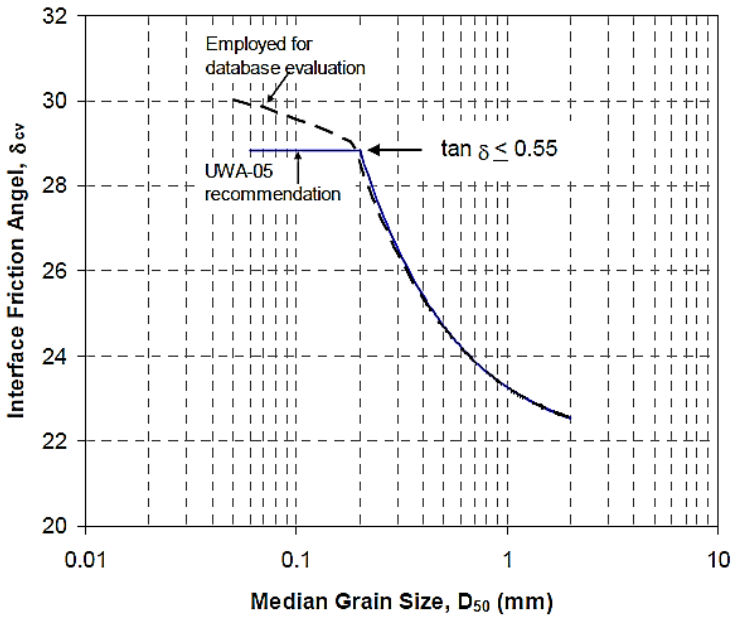
$$f_{clay} = \sigma'_{rf} \tan\delta'_f \tag{A.90}$$

$$\sigma'_{rf} = (0.3 + 0.3e^{-I_{vr}})\left(q_t/\sigma'_{v0}\right)^{0.6}\left[\max\left(\frac{h}{R}, 8\right)\right]^{-0.2} \text{ For } I_p \geq 35\% \tag{A.91}$$

$$\sigma'_{rf} = (0.45 + 0.15e^{-I_{vr}})\left(q_t/\sigma'_{v0}\right)^{0.6}\left[\max\left(\frac{h}{R}, 8\right)\right]^{-0.2} \quad \text{For } I_p < 35\% \tag{A.92}$$

The variation of Variation of δ'_f with D_{50} as modified from ICP-05 is presented in Figure 62.

Figure 62. Variation of δ'_f with D_{50} –modified from ICP-05



Lehane et al. examined 75 pile load tests [29]. For the cases that the corrected end resistance q_t was not available the following procedure was used:

1. For soft and firm clays (q_c <1 MPa):

$$q_{b-clay} = 1.14\, q_c \tag{A.93}$$

This value is estimated from the corrected cone resistance equation, assuming a pore-pressure ratio (B_q) of 0.60 (lightly consolidated clays) with cone end ratio (a) of 0.80.

2. For stiff clays (q_c >1 MPa):

$$q_{b-clay} = q_c \tag{A.94}$$

Based on recommendations of [20] the unit end bearing capacity for close-ended piles was assumed to be:

$$q_{b-clay} = 0.8 q_{t,avg} \tag{A.95}$$

And for open-ended piles:

$$q_{b-clay} = 0.4 q_{t,avg} \tag{A.96}$$

And the shaft capacity in tension and compression was assumed to be equal. The equivalent pile radius for square piles is:

$$R = {B}/{\pi^{0.5}} \tag{A.97}$$

Based on these assumptions, the following equations are suggested for calculating the unit shaft resistance (the average of these two equations is used for f_{clay})

$$f_{clay} = 0.055 q_t \left[\max\left(\frac{h}{R}, 1\right)\right]^{-0.2} \tag{A.98}$$

$$f_{clay} = \frac{0.23 q_t \left[\max\left(\frac{h}{R}, 1\right)\right]^{-0.2}}{\left(\frac{q_t}{\sigma'_{v0}}\right)^{0.15}} \tan\delta'_f \tag{A.99}$$

where, δ'_f is estimated based on ICP method. It should be noticed that CPT-2000 methods is not different from UWA method, only the equations for f_{clay} of CPT-2000 are simpler.

Implementing Probabilistic and Robertson-2010 soil classifications into these methods are the same as described for NGI method.

The only difference between UWA and CPT-2000 method is the equation for shaft resistance in clay. No upper limits was proposed by these methods.

Fugro Method

This method is based on the studies performed by Kolk and der Velde and Kolk et al. on 45 sites (24 open-ended and 21 close-ended piles) consisting sandy soils and 26 pile load tests in clayey soils ([30], [104]).

The unit end bearing capacity for sands is:

$$q_{b-sand} = 8.5 q_{c,avg} \left({P_0}/{q_{c,avg}}\right)^{0.5} {A_r}^{0.25} \tag{A.100}$$

where:

$$A_{rb} = 1 - \left(\frac{d_i}{d}\right)^2 \tag{A.101}$$

The arithmetic average of CPT q_c is taken over the influence zone defined as 1.5d above and below the pile's tip.

The unit friction resistance is estimated by considering a reduction near to the tip:

Compression loading, and ${h}/{R^*} \geq 4$:

$$f_{sand} = 0.08 q_c \left({\sigma'_{v0}}/{P_{atm}}\right)^{0.05} \left(\frac{h}{R^*}\right)^{-0.9} \tag{A.102}$$

Compression loading, and ${h}/{R^*} < 4$:

$$f_{sand} = 0.08 q_c \left({\sigma'_{v0}}/{P_{atm}}\right)^{0.05} 4^{-0.9} \left(\frac{h}{4R^*}\right) \tag{A.103}$$

Tension loading:

$$f_{sand} = 0.45 q_c \left({\sigma'_{v0}}/{P_{atm}}\right)^{0.15} 4^{-0.9} \left[\max\left(\frac{h}{R^*}, 8\right)\right]^{-0.85} \tag{A.104}$$

where:

$$R^* = \left(R^2 - {R_i}^2\right)^{0.5} \tag{A.105}$$

And for non-circular piles, the equivalent circular area is used to estimate R^*.

The procedure of Fugro method for the clay soils is as follows:

$$q_{b-clay} = 0.7(q_c - \sigma_{v0}) \tag{A.106}$$

$$f_{clay} = \alpha S_u \tag{A.107}$$

$$\alpha = 0.9 \left(\frac{L - z}{D}\right)^{-0.2} \left(\frac{S_u}{\sigma'_{v0}}\right)^{-0.3} \leq 1 \tag{A.108}$$

where, L is the pile's length and z is the depth.

Probabilistic and Robertson-2010 soil classifications are implemented as NGI and UWA methods. No upper limits was proposed by this method.

Purdue-CPT Method

The method is developed by Salgado et al. based on the following procedure described in [33]:

The unit end bearing capacity for sands:

$$q_{b-sand} = (1 - 0.0058D_r)q_{c,avg} \tag{A.109}$$

The unit skin resistance for sands:

$$f_{sand} = K\sigma'_{v0} \tan\delta'_f \tag{A.110}$$

where:

$$K = K_{min} + (K_{max} - K_{min}) \exp\left(-\beta \frac{h}{B}\right) \tag{A.111}$$

$$K_{min} = 0.2 \tag{A.112}$$

$$K_{max} = 0.02 \, {q_c}/{\sigma'_{v0}} \tag{A.113}$$

where, h is the distance from the depth being considered to the pile base.

$$\beta = 0.05 \tag{A.114}$$

The unit end bearing capacity for clays:

$$q_{b-clay} = 10S_u \tag{A.115}$$

The unit skin resistance for clays:

$$f_{clay} = \alpha S_u \tag{A.116}$$

$$\alpha = 1.28 \left(\frac{S_u}{\sigma'_{v0}}\right)^{-0.05} [A_1 + (1 - A_1)e^{\gamma}] \tag{A.117}$$

$$\gamma = -\left(\frac{\sigma'_{v0}}{P_A}\right)\left(\varphi_c - \varphi_{r,min}\right)^{A_2} \tag{A.118}$$

$$\text{For } \left(\varphi_c - \varphi_{r,min}\right) \leq 5°\text{: } A_1 = 0.75 \tag{A.119}$$

$$\text{For } \left(\varphi_c - \varphi_{r,min}\right) \geq 12°\text{: } A_1 = 0.43 \tag{A.120}$$

For $5 < \left(\varphi_c - \varphi_{r,min}\right) < 12$, A_1 is obtained by interpolation.

$$A_2 = 0.64 + 0.4 \ln\left(\frac{S_u}{\sigma'_{v0}}\right) \tag{A.121}$$

The following assumptions have been used for implementing this method:

The value of D_r is between 0 and 1 and average of two methods for determining it, has been used:

$$D_r = 0.4 \ln \frac{q_c}{22\sqrt{\sigma'_{vo}\,\sigma_{atm}}}$$

$$D_r = \frac{1}{2.41} \ln\left(\frac{95.76\, q_c}{157(95.76\sigma'_{vo})^{0.55}}\right)$$

where, q_c and σ'_{vo} units are TSF. q_c has been calculated for 1d above and below the location and

The value of S_u was determined from:

$$S_u = \frac{q_{c,avg}}{N_k}$$

where, Nk = 20 has been used, similar to De Ruiter and Beringen method [2].

φ_c and $\varphi_{r,min}$ are assumed to be equal to $1.25\delta_{peak}$ and $1.25\delta_{ultimate}$ calculated in ICP method.

Also another method for calculating α is used and the average of these two alpha values is used in calculating f_{clay}.

$$\alpha = 0.4\left[1 - 0.12 \ln\left(\frac{S_u}{P_{atm}}\right)\right] \tag{A.122}$$

Probabilistic and Robertson-2010 soil classifications are implemented as NGI, UWA, and Fugro methods. No upper limits was proposed by this method.

Probabilistic Method

Abu-Farsakh and Titi used the data from 35 square PPC piles (26 driven in clay, 9 driven in layered soil) for estimating the ultimate load obtained by the Butler-Hoy method [17]. It has been shown that using the Davisson method [18] for estimating the bearing capacity of piles from pile-load test will give very close results [17].

$$Q_{u(\text{Davisson})} = 1.02 \times Q_{u(\text{Butler}-\text{Hoy})} \tag{A.123}$$

They used the Zhang and Tumay method [21] for soil classification, which was illustrated in earlier section:

The unit end bearing capacity is:

$$q_b = K_b \, q_{ca} < 15 \text{ MPa} \tag{A.124}$$

where, q_{ca} is determined similar to Schmertmann method for the influence zone of 4D below and 8D above the pile's tip. Also, a weight function has been introduced to give more weight to readings near to the tip as shown in Figure 63.

It should be noted that using the q_t as the corrected value of q_c is also possible, assuming that:

$$q_c = 0.95 \, q_t \tag{A.125}$$

We have:

$$K_b = 0.3 \, \Pr(\text{sand}) + 0.4 \, \Pr(\text{silt}) + 0.5 \, \Pr(\text{clay}) \tag{A.126}$$

where, the values of $\Pr(\text{sand})$, $\Pr(\text{silt})$, and $\Pr(\text{clay})$ within the influence zone is determined by Zhang and Tumay CPT soil classification method [23].

For the unit skin friction:

$$f = K_s \, f_s \tag{A.127}$$

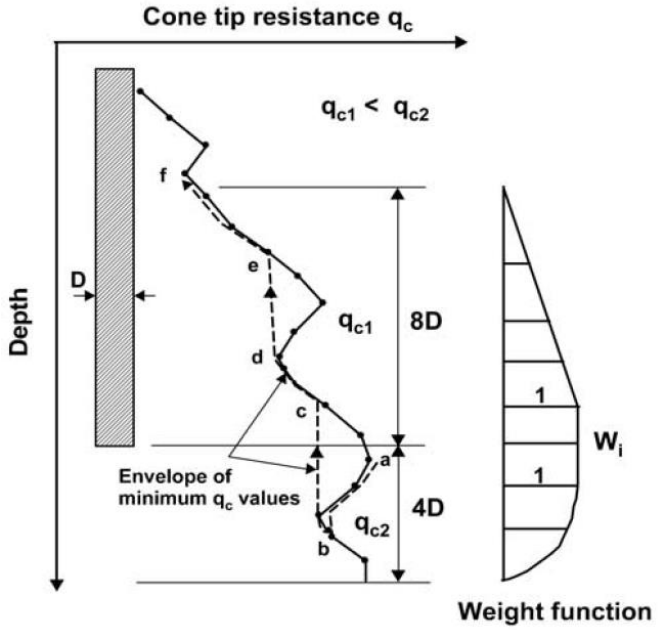
where:

$$K_s = K_{s(silt-clay)} \Pr(silt - clay) + K_{s(sand)} \Pr(sand) \quad (A.128)$$

$$K_{s(silt-clay)} = \frac{1}{(0.74 + 1.62\, f_s)} \quad (A.129)$$

$$K_{s(sand)} = \frac{1}{(0.61 + 0.7\, f_s)} \quad (A.139)$$

Figure 63. Calculation of q_{ca} using the weight function



Implementing Robertson-2010 soil classification in this method is as follows:

The value of K_b is determined from the below table:

Table 35. Values of K_p versus soil index for Robertson-10

Soil index	K_b
1	0.5
2	0.5
3	0.467
4	0.45
5	0.35
6	0.3
7	0.3
8	0.367
9	0.45

And:

$$f = (K_{sand} \times \alpha 1 + K_{silt-clay} \times \beta 1) \times f_s$$

where, $\alpha 1$ and $\beta 1$ are obtained from Table 18.

UF Method

Bloomquist et al. and Hu et al. at the University of Florida used 21 cases of pile load test in Florida with sandy soil and 28 from Louisiana with clayey soil to developed a new method, which is considered as a modification of the Philipponnat method ([34], [35], [7]).

According to UF method, the unit end bearing capacity is given as:

$$q_b = k_b q_{ca} \leq 150 \text{ TSF} \tag{A.140}$$

where, k_b is a factor that depends on the soil type as shown in Table 36.

Table 36. Bearing capacity factor

Soil Type	k_b
Well-cemented sand	0.1
Lightly cemented sand	0.15
Gravel	0.35
Sand	0.40
Silt	0.45
Clay	1

The soil classification chart for electronic friction cone developed by Robertson et al. has used for determining the soil category in UF method [21]. The chart is presented in Figure 64.

The soil cementation is determined based on SPT numbers where $q_c/N > 10$.

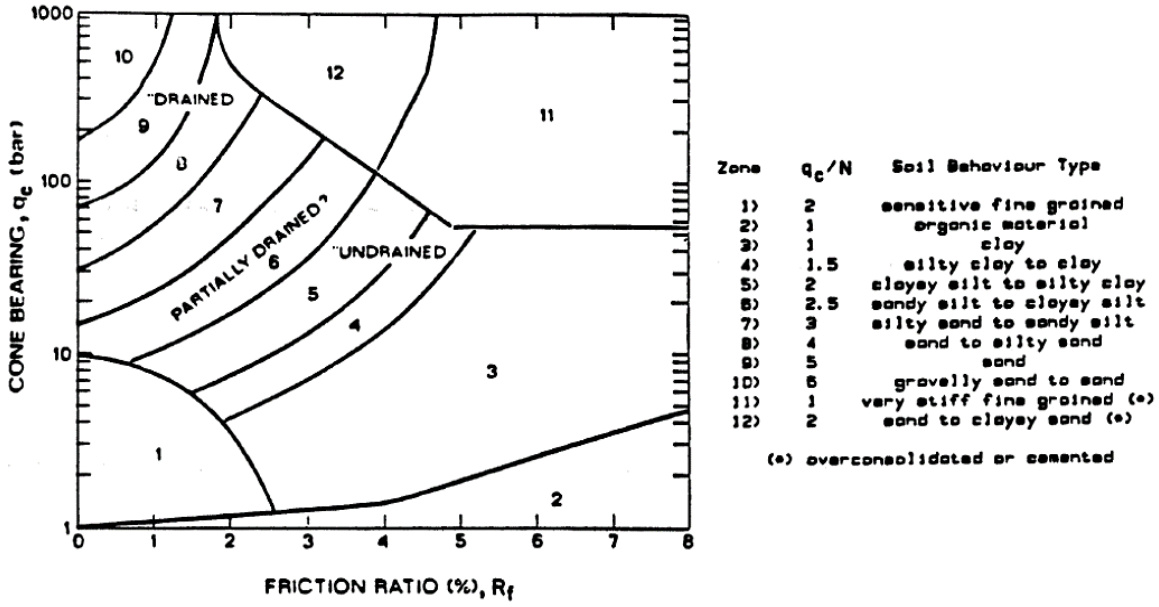
The value of q_{ca} is obtained by:

$$q_{ca} = \frac{q_{c1} + q_{c2}}{2} \tag{A.141}$$

where q_{c1} is average of q_c measure from pile's tip to the depth of 3D for sand and 1D for clay. q_{c2} is average of q_c measure from pile's tip to 8D above it. In cases that $q_{c2} > q_{c1}$, then:

$$q_{ca} = q_{c1} \tag{A.142}$$

Figure 64. Soil classification method used by the UF method



The unit skin friction of the pile (f) is determined by:

$$f = \frac{\alpha_s}{F_s} q_{cs} \leq 1.2 \text{ TSF} \tag{A.143}$$

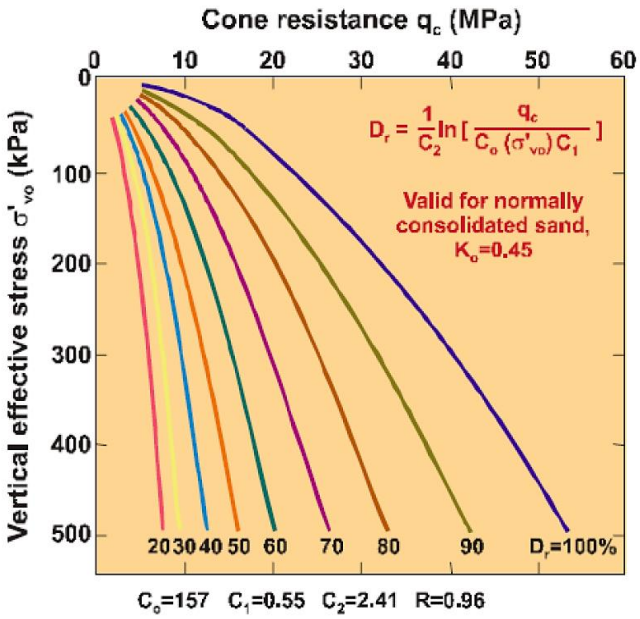
where, α_s is the same as the Philipponnat method and the values of F_s are shown in Table 37.

Table 37. Empirical factor F_s

Soil Type	F_s
Clay and calcareous clay	50
silt, sandy clay, and clayey sand	60
Loose sand	100
Medium dense sand	150
Dense sand and gravel	200
Lightly cemented sand	250
Well-cemented sand	300

Figure 65 from Baldi et al. references has been used for determining the relative density of sands, D_r, for evaluating the sand state (Loose: D_r<40%, Medium: 40%<D_r<70%, and Dense: 70%<D_r) [105].

Figure 65. Chart for determining the sand state in UF method [105].



For probabilistic soil classification:

If (%clay) > 0.8: $K_b = 1.0$

If (%sand) > 0.8: $K_b = 0.4$

$$K_b = 0.4 \times \%\text{sand} + 0.45 \times \%\text{silt} + 1 \times \%\text{clay}$$

For Robertsson-2010 soil classification, Table 38 can be used to estimate k_b.

Table 38. Values of k_b versus soil index for Robertson-10

Soil index	k_b
1	1
2	1
3	0.82
4	0.45
5	0.425
6	0.4
7	0.375
8	0.6
9	0.65

The values of α_s and F_s were implemented the same as in Philipponnat method.

Togliani Method

This method is developed by Togliani for cylindrical piles driven and bored in different soils [36]. The pile load tests were performed 30 days after the initial pile driving.

The unit end bearing capacity is given as:

$$q_b = k_3 q_c \tag{A.144}$$

where, q_c is the average 8D/4D above/below the pile's tip.

$$k_3 = \lambda + \left[0.01\left(\frac{L}{d}\right)\right] \tag{A.145}$$

where, L is length and d is diameter of the pile.

where, λ is 0.1 and 0.2 for bored and driven piles, respectively.

For pile unit side resistance:

$$f = k_1 {q_c}^{0.5} \tag{A.146}$$

where, the units are assumed to be in kPa.

$$k_1 = 1.2\left(0.8 + {R_f}/{8}\right) \text{ for } R_f < 1 \tag{A.147}$$

$$k_1 = 1.1(0.4 + \ln R_f) \text{ for } R_f \geq 1 \tag{A.148}$$

where,

$$R_f = \left({f_s}/{q_c}\right) 100 \tag{A.149}$$

No upper limits was proposed by this method.

Zhou et al. Method

This method which is developed by Zhou et al. predicts the limit load capacity instead of the ultimate load capacity [37]. Using the Schmertmann relationship we have:

$$(Q_t)_{limit} = 0.73\ (Q_t)_{ultimate} \tag{A.150}$$

The limit load is the point where the shaft resistance of the pile is fully mobilized, while the end resistance is only partially mobilized. If the point is not obvious using the data, they recommend using the load at a relative settlement of 0.4-0.5, which is the ratio of settlement to ultimate

settlement (punching failure). They argued that many authors take failure load of the testing pile as criterion for compression and when failure load could not be reached, Van der veen interpretation method is used. They mentioned that based on their experience, Van der veen interpretation involves unavoidable artificial error.

The unit end bearing capacity q_b of the pile

$$q_b = \alpha\, q_{ca} \tag{A.151}$$

where,

q_{ca} : The average CPT tip resistance 4D above and 4D below the pile's tip with condition of $q_{ca(A)} \leq q_{cb(B)}$.

α: A function of soil type and q_{ca}

Soil type I: $q_{ca} > 2$ Mpa and $f_{sa}/q_{ca} < 0.014$

Soil type II: Other soils other than soil type I

Soil type I:

$$\alpha = 0.71\, q_{ca}^{-0.25} \tag{A.152}$$

Soil type II:

$$\alpha = 1.07\, q_{ca}^{-0.35} \tag{A.153}$$

The unit skin friction f' of the pile is:

$$f' = \beta\, f_{sa} \tag{A.154}$$

f_{sa}: The average of f_s along the soil layer

β: The function of soil type and f_{sa}

Soil type I:

$$\beta = 0.23\, f_{sa}^{-0.45} \tag{A.155}$$

Soil type II:

$$\beta = 0.22\, f_{sa}^{-0.55} \tag{A.156}$$

where, units are in MPa. No upper limits was proposed by this method.

German Method

Kempfert and Becker references obtained empirical clause for tip and side resistance of the piles based on the CPT data, which are integrated into national German recommendations for piles "EA-Pfähle" (DIN 4014) [106]. Their results were based on about 1000 pile load tests on different piles (121 driven PPC piles, 98 driven steel piles, 70 driven cast-in-place Simplex piles, 300 driven cast-in-place Franki piles, 124 screwed Atlas piles, 52 screwed Fundex piles, and 38 micro-piles). Based on their analysis, lower and upper bounds for base and shaft resistances in noncohesive and cohesive soils were obtained.as shown in Figure 66 and Figure 67, respectively.

The German method charts were converted to below equations:

In sands:

$$\begin{aligned} q_{b-sand} &= \min(0.56q_c + 5.4, 0.1q_c + 79, 105)\ (\text{TSF}) \\ f_{sand} &= \min(0.007q_c, 0.004q_c + 0.48, 1.52)\ (\text{TSF}) \end{aligned} \tag{A.157}$$

In clays: q_{b-clay} value was obtained from the chart available in Eurocode 7 method as:

$$\begin{aligned} q_{b-clay} &= 7S_u + 1\ (\text{TSF}) \\ f_{clay} &= \min(0.15S_u + 0.15, 0.5)\ (\text{TSF}) \end{aligned} \tag{A.158}$$

Probabilistic and Robertson-2010 soil classifications are implemented as De Ruiter and Beringen method [2].

Figure 66. Upper and lower empirical values for (a) base resistance (b) shaft resistance in noncohesive soils [106]

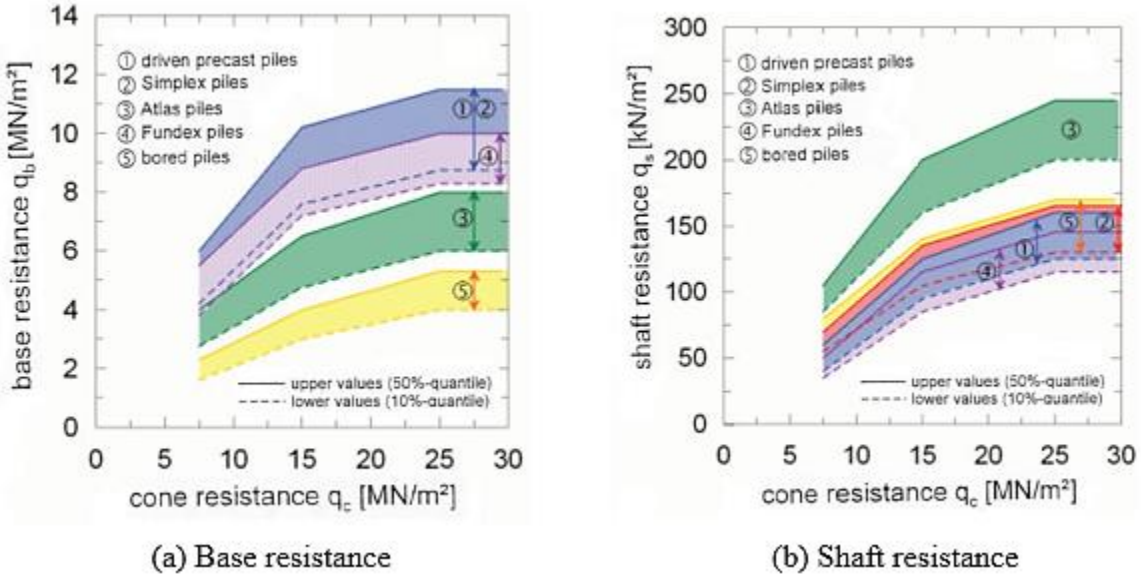
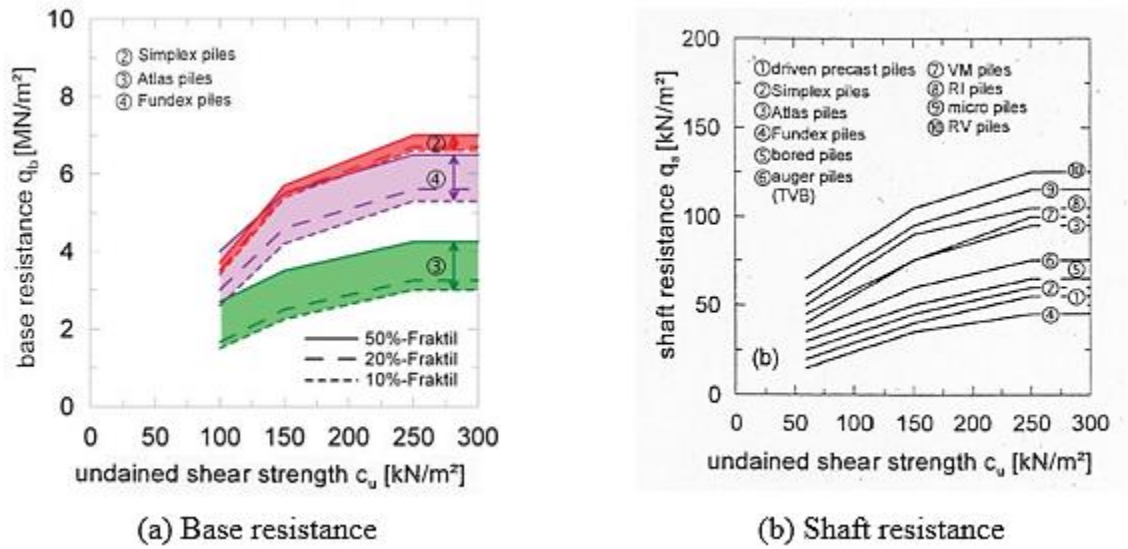


Figure 67. Upper and lower empirical values for (a) base resistance (b) shaft resistance in cohesive soils [106]



Eurocode 7 Method

Eurocode 7 standard presents different possible design approaches for pile design [107]:

1) Static load tests on similar piles

2) Dynamic load tests verified by static load results

3) Empirical or analytical calculation methods validated by static load tests

Eurocode 7 determines no specific procedure for pile design using CPT data. However, in the annex of the Eurocode 7-2 two examples of methods based on the direct implementation of the results of CPT are presented, which are based on methods of German and Dutch codes, respectively [108]. The first example, based on the German Standard (DIN 4014) presents Tables D.3 and D.4 in Annex D.6 of Eurocode 7-2, is restricted to coarse-grained soils. The second example described in Annex D.7 is more complex and takes into consideration a number of different factors. This example was presented in Annex B.4 of Eurocode 7-3; however, some changes in the values of α_s have been added as shown in tables.

The unit end bearing capacity of the pile (q_b) is obtained from:

$$q_b = \alpha_p \beta s\left(q_{c,avg}\right) \leq 15\text{MPa} \tag{A.159}$$

where, $q_{c,avg}$ is the determined similar to Schmertmann method, α_p is the pile class factor in Table 39, β is the factor taking the shape of the pile point as shown in Figure 68 and s is the factor account for the shape of the pile base shown in Figure 69.

For the square piles in this study, the values of α_p, β, and s are equal to 1.0, which makes the unit tip resistance equal to the Schmertmann method.

The unit skin friction of the pile (f) is given by:

$$f = \alpha_s q_c \tag{A.160}$$

where, α_s is a reduction factor based on Table 40 and Table 41.

Table 39. Maximum values for α_p and α_s for sands and gravely sands

Pile class or type	α_p	α_s [1]
Soil displacement type piles, diameter > 150 mm – driven prefabricated piles, – cast in place piles made by driving a steel tube with closed end. The steel pipe is reclaimed during concreting.	1,0	0,010
Soil replacement type piles, diameter > 150 mm – flight auger piles,	1,0	0,014
– bored piles (with drilling mud).	0,8 0,6	0,006 [2] 0,005

1) Values valid for fine to coarse sands. For very coarse sands a reduction factor of 0,75 is necessary; for gravel this reduction factor is 0,5.

2) This value is used in the case of applying the results of CPT s which were carried out before pile installation. When CPT s are used that have been carried out in the vicinity of the flight auger piles, α_s may be raised to 0,01.

Figure 68. Pile point shape β [108]

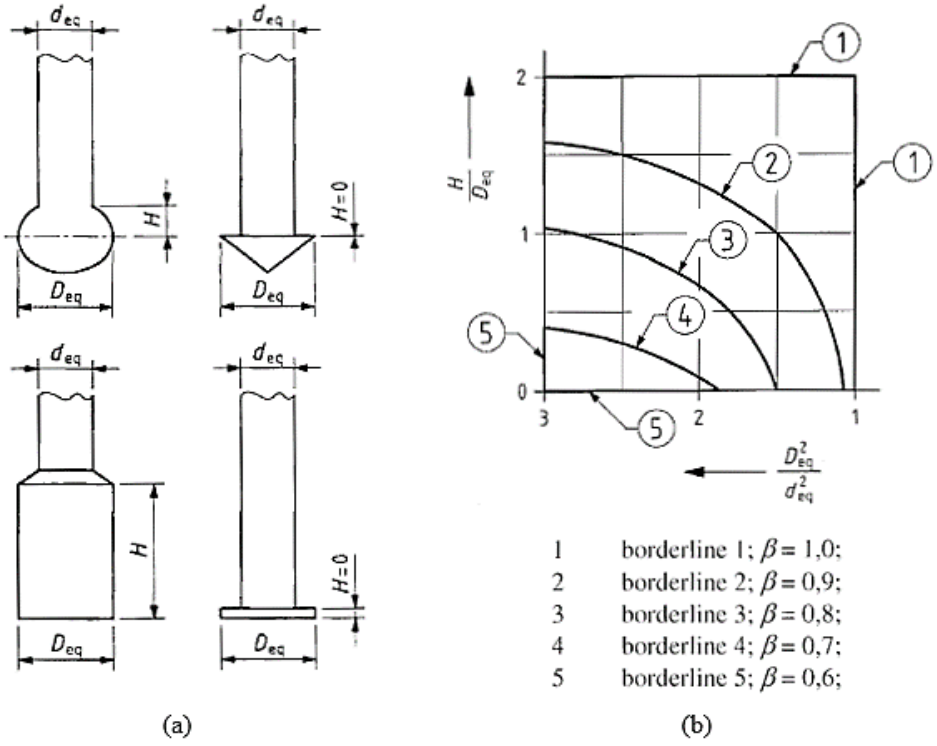


Figure 69. Values of s based on width a and b of the pile [108]

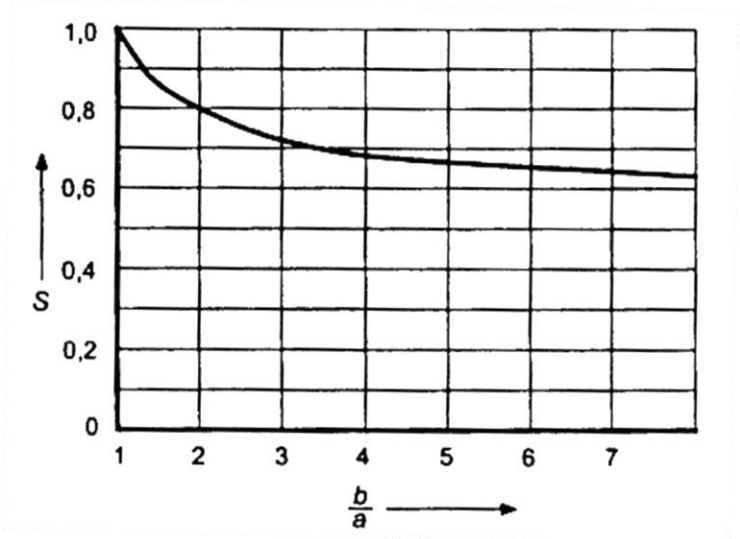


Table 40. Maximum α_s for clays, silt, and peat for Eurocode 7-1 [108]

Soil type	relative depth z/d_{eq}	α_s
clay/silt ($q_c \leq 1$ MN/m^2)	$5 < z/d_{eq} < 20$	0,025
clay/silt ($q_c \geq 1$ MN/m^2)	$z/d_{eq} \geq 20$	0,055
clay/silt ($q_c > 1$ MN/m^2)	not applicable	0,035
peat	not applicable	0
d_{eq} equivalent pile shaft diameter		

Values of α_S for clays and silts are updated in Eurocode 7-2 as:

Table 41. Maximum α_s for clays, silt, and peat for Eurocode 7-2 [108]

Soil type	q_c, MPa	α_s
clay	> 3	< 0,030
clay	<3	< 0,020
silt		< 0,025
peat		0

Implementing Probabilistic soil classification into this method is as follows:

If (%clay) > 0.8

$$f = \alpha_{clay}\, q_c$$

If (%sand) > 0.8

$$f = \alpha_{sand}\, q_c$$

Otherwise:

$$f = (\%clay \times \alpha_{clay} + \%silt \times \alpha_{silt} + \%sand \times \alpha_{sand}) \times q_c$$

For implementing Robertson-2010 soil classification:

$$f = \alpha \times q_c$$

where, α values are obtained from Table 42.

Table 42. Values α for calculating unit side resistance for EC 7-2 method

Soil index	α
1	0.02
2	0.02
3	α_{clay}
4	0.0225
5	0.015
6	0.01
7	0.0075
8	0.01
9	α_{clay}

ERTC3 Method

European Regional Technical Committee 3 "ERTC3" has the same process for calculation of unit base resistance as the Eurocode 7 method, but the determination of the shaft resistance is modified [109]. It can be calculated with the same process, but α_s values are different, as shown in Table 43.

Table 43. Values of α_s for ERTC3 method [109]

Non-cohesive	Gravel	0.003
	Sandy gravel	0.0045
	Fine sand	0.006
	Sandy silt	0.008
	Silt	0.01
Cohesive	$q_c >$ 2500 kPa	0.015
	1500 kPa $< q_c <$ 2500 kPa	0.025
	1000 kPa $< q_c <$ 1500 kPa	0.035
	500 kPa $< q_c <$ 1000 kPa	0.045
	$q_c <$ 500 kPa	0.055

Implementing the probabilistic soil classification into this method is as follows:

If (%clay) > 0.8

$$f = \alpha_{clay}\, q_c$$

If (%sand) > 0.8

$$f = \alpha_{sand}\, q_c$$

Otherwise:

$$f = \left(\%clay \times \alpha_{clay} + \%silt \times \alpha_{silt} + \%sand \times \alpha_{sand}\right) \times q_c$$

For implementing Robertson-2010 soil classification:

$$f = \alpha \times q_c$$

where, α values are obtained from Table 44, similar to Eurocode 7-2 method.

Table 44. Values of α for calculating unit side resistance for EC 7-2 method [109]

Soil index	α
1	0.055
2	α_{clay}
3	α_{clay}
4	$1/2 * (0.01 + \alpha_{clay})$
5	α_{sand}
6	α_{sand}
7	0.045
8	α_{sand}
9	α_{clay}

Appendix B

Summary of DOTD State Projects Investigated in this Study

No.	Project ID	Project Name	Parish	Pile ID	Diam. (inch)	Tip elevation (ft)	Ground elevation (ft)	Cut-off elevation (ft)	Pile length (ft)	Embedment length (ft)	Weight of Pile (Tons)	Hammer type	Date of driving	Date of loading
1	003-07-0019	BNSS Overpass - Jennings	Jefferson Davis Parish	1	24	-33.9	20.3	30	63.9	54.2	15.4	ICE 60S Single Acting Diesel	8/2/2000	8/21/2000
2	003-10-0011	Southern Pacific Railroad Overpass	Acadia	1	24	-28	21.5	57	85	49.5	20.5	Vulcan 512	9/8/1998	10/8/1998
3	003-10-0011	Southern Pacific Railroad Overpass	Acadia	3	24	-37	24	48	85	61	20.5	Vulcan 512	8/31/1998	10/1/1998
4	005-01-0056	Southern pacfic railroad overpass	St. Mary Parish	1	24	-79.5	5.5	10.5	90	85	21.7	Delmag 46-02 Single Acting Diesel - OED	4/12/1995	5/12/1995
5	005-01-0056	Southern pacfic railroad overpass	St. Mary Parish	2	14	-64	-0.3	10	74	63.7	7.6	ICE 42S Single Acting Diesel - OED	4/6/1995	5/1/1995
6	005-01-0056	Southern pacfic railroad overpass	St. Mary Parish	3	24	-80	7	12	92	87	22.2	Delmag 30-32 Single Acting Diesel - OED	4/18/1995	5/15/1995
7	047-02-0022	Bogue Chitto Bridge	Washigton	2	30	60.25	132.75	132.75	72.5	72.5	23.6	Vulcan 020 Single acting ECH	-	-
8	061-05-0044	Route LA 10	East Feliciana Parish	2	24	117.31	177.31	182.09	64.78	60	16.5	ICE 120S Single Acting Diesel	-	-
9	064-06-0036	Bayou Lafourche Bridge and Approaches	Lafourche Parish	1	24	-42	-3	16.91	58.91	39	14.2	Delmag 30-23 Single Acting Diesel	8/30/2000	9/13/2000
10	065-90-0024 855-04-0046	Houma I.C.W.W. Bridges	Terrebonne Parish	1	14	-80	0	0	80	80	8.2	Delmag 19-32 Single Acting Diesel	7/6/1993	7/20/1993
11	065-90-0024 855-04-0046	Houma I.C.W.W. Bridges	Terrebonne Parish	2	14	-70	0	3.48	73.48	70	7.5	Delmag 19-32 Single Acting Diesel	6/28/1993	7/12/1993
12	065-90-0024 855-04-0046	Houma I.C.W.W. Bridges	Terrebonne Parish	3	14	-80	0	14.5	94.5	80	9.6	Delmag 19-32 Single Acting Diesel	10/20/1993	11/4/1993
13	065-90-0024 855-04-0046	Houma I.C.W.W. Bridges	Terrebonne Parish	4	14	-80.5	0.47	11.5	92	80.97	9.4	Delmag 19-32 Single Acting Diesel	2/25/1994	3/11/1994
14	065-90-0024 855-04-0046	Houma I.C.W.W. Bridges	Terrebonne Parish	5	16	-70	1.5	2	72	71.5	9.6	Delmag 19-32 Single Acting Diesel	7/1/1993	7/15/1993
15	065-90-0024 855-04-0046	Houma I.C.W.W. Bridges	Terrebonne Parish	6	16	-98.7	0	11.3	110	98.7	14.7	Delmag 30-23 Single Acting Diesel	9/14/1993	9/29/1993
16	239-01-0080	ICWW Bridge Approaches (Louisa)	St. Mary Parish	3	14	-52	4.125	-1.87	50.13	50.13	5.1	Vulcan 010 (Air)	6/14/2001	7/5/2001

No.	Project ID	Project Name	Parish	Pile ID	Diam. (inch)	Tip elevation (ft)	Ground elevation (ft)	Cut-off elevation (ft)	Pile length (ft)	Embedment length (ft)	Weight of Pile (Tons)	Hammer type	Date of driving	Date of loading
17	239-01-0080	ICWW Bridge Approaches (Louisa)	St. Mary Parish	4	30	-83.5	1.325	8.36	91.86	84.825	43.1	Vulcan 025 (Air)	6/15/2001	7/2/2001
18	260-05-0020	Tichfaw River Bridge and Approaches	Livingston Parish	1	30	-60	6.3	13	73	66.3	23.8	Delmag 46-32 Single Acting Diesel - OED	2/25/1998	3/23/1998
19	260-05-0020	Tickfaw River Bridge	Livingston Parish	3	30	-70	2	11	81	72	26.4	Delmag 46-32 Single Acting Diesel - OED	3/26/1998	4/17/1998
20	262-06-0009 & 262-07-0012	Bridge #1 Tickfaw River	St. Helena and Tangipahoa	1	24	8.8	93.7	99.8	91	84.9	21.9	Delmag 46-13 Single Acting Diesel - OED	10/16/1990	11/5/1990
21	262-06-0009 & 262-07-0012	Bridge #1 Tickfaw River	St. Helena and Tangipahoa	2	24	-8.2	96.8	96.8	105	105	25.3	Delmag 46-13 Single Acting Diesel - OED	10/16/1990	11/12/1990
22	283-03-0052 (New Orleans)	West bank expressway	Jefferson	1	18	-120		0	120	120	20.3	-	-	-
23	424-04-0026 (New Iberia)	US 90 - Lewis Street Interchange (Overpass)	New Iberia	2	14	-49.5	5.5	10.5	60	55	6.1	Vulcan 010 (Air)	10/16/1997	11/14/1997
24	424-04-0027	US 90 Interchange at John Darnell Road	Iberia Parish	1	14	-30	10.4	22	52	40.4	5.3	ICE 60S Single Acting Diesel	5/4/2000	5/18/2000
25	424-04-0027	US 90 Interchange at John Darnell Road	Iberia Parish	2	14	-30.2	7.3	20.63	50.83	37.5	5.2	ICE 60S Single Acting Diesel	5/4/2000	5/19/2000
26	424-05-0078	Bayou Boeuf Bridge Main Span	St. Mary Parish	1	14	-70	7	10	80	77	8.2	Delmag 19-32 Double Acting Diesel	4/21/1992	5/5/1992
27	424-05-0078	Bayou Boeuf Bridge Main Span	St. Mary Parish	2	14	-70	5	10	80	75	8.2	Delmag 19-32 Double Acting Diesel	4/21/1992	5/5/1992
28	424-05-0078	Bayou Boeuf Bridge Main Span	St. Mary Parish	5	14	-80	0.5	5	85	80.5	8.7	Delmag 19-32 Single Acting Diesel	12/14/1992	12/29/1992
29	424-05-0081	Bayou Boeuf Bridge (West Approch)	St. Mary Parish	1	14	-90	0	6	96	90	9.8	Delmag 19-32 Single Acting Diesel	4/29/1994	5/9/1994
30	424-05-0081	Bayou Boeuf Bridge (West Approch)	St. Mary Parish	2	30	-112.5	0	3	115.5	112.5	37.6	Delmag 46-32 Single Acting Diesel - OED	6/20/1994	7/12/1994
31	424-05-0081	Bayou Boeuf Bridge (West Approch)	St. Mary Parish	3	14	-68	-3.5	6	74	64.5	7.6	Delmag 19-32 Single Acting Diesel	4/28/1994	5/12/1994
32	424-05-0081	Bayou Boeuf Bridge (West Approch)	St. Mary Parish	4	16	-69	0.85	11	80	69.85	10.7	Delmag 19-32 Single Acting Diesel	5/12/1994	5/27/1994
33	424-05-0087	Morgan City - Gibson Highway	St. Mary Parish	1	16	-68.5	1	16.5	85	69.5	11.3	ICE640 Double Acting Diesel	6/9/1992	6/25/1992
34	424-05-0087	Morgan City - Gibson Highway	St. Mary Parish	2	30	-88.5	1.2	3	91.5	89.7	42.9	Delmag 46-23 Single Acting Diesel - OED	1/16/1992	1/28/1992

No.	Project ID	Project Name	Parish	Pile ID	Diam. (inch)	Tip elevation (ft)	Ground elevation (ft)	Cut-off elevation (ft)	Pile length (ft)	Embedment length (ft)	Weight of Pile (Tons)	Hammer type	Date of driving	Date of loading
35	424-05-0087	Morgan City - Gibson Highway	St. Mary Parish	3	30	-100	4.05	7	107	104.05	50.2	Delmag 46-23 Single Acting Diesel - OED	2/10/1992	2/26/1992
36	424-05-0087	Morgan City - Gibson Highway	St. Mary Parish	4	30	-100	-0.7	6.38	106.38	99.3	49.9	Delmag 46-23 Single Acting Diesel - OED	2/19/1992	3/5/1992
37	424-05-0087	Morgan City - Gibson Highway	St. Mary Parish	5	30	-107	6	7	114	113	53.4	Delmag 46-23 Single Acting Diesel - OED	3/11/1992	3/25/1992
38	424-05-0087	Morgan City - Gibson Highway	St. Mary Parish	7	16	-75	2	10	85	77	11.3	ICE640 Double Acting Diesel	6/9/1992	6/30/1992
39	424-07-0008	GIBSON-CHACAHOULA (relocated US 90) Route LA 3052	TERREBONNE	1	30	-130	2.7	7	137	132.7	44.6	Delmag 46-23 Single Acting Diesel - OED	10/12/1990	11/13/1990
40	424-07-0008	GIBSON-CHACAHOULA (relocated US 90) Route LA 3052	TERREBONNE	3	30	-125	2.4	12	137	127.4	44.6	Delmag 46-23 Single Acting Diesel - OED	12/21/1990	1/12/1991
41	424-07-0009	GIBSON-RACELAND HIGHWAY	TERREBONNE	3	30	-116.7	5	26.3	143	121.7	46.5	Delmag 46-23 Single Acting Diesel - OED	5/15/1991	5/31/1991
42	424-07-0009	GIBSON-RACELAND HIGHWAY	TERREBONNE	4	30	-115	5	10	125	120	40.7	Delmag 46-23 Single Acting Diesel - OED	12/14/1990	1/18/1991
43	424-07-0009	GIBSON-RACELAND HIGHWAY	TERREBONNE	4A	30	-119	5	6	125	124	40.7	Delmag 46-23 Single Acting Diesel - OED	1/23/1991	2/7/1991
44	424-06-0005	Bayou Boeuf Bridge	Assumption Parish	1	14	-70	-2	5	75	68	7.7	ICE 640 Double Acting Diesel	5/18/1993	19936/16/19
45	424-06-0005	Bayou Boeuf Bridge	Assumption Parish	2	14	-76	-4	4	80	72	8.2	ICE 640 Double Acting Diesel	5/18/1993	6/1/1993
46	424-06-0005	Bayou Boeuf Bridge	Assumption Parish	3	14	-84.5	-7	0.5	85	77.5	8.7	ICE 640 Double Acting Diesel	10/1/1993	10/19/1993
47	424-06-0005	Bayou Boeuf Bridge	Assumption Parish	4	14	-85	-6	0	85	79	8.7	ICE 640 Double Acting Diesel	9/30/1993	10/26/1993
48	424-06-0005	Bayou Boeuf Bridge	Assumption Parish	5	14	-85	-6	0	85	79	8.7	ICE 640 Double Acting Diesel	9/30/1993	10/28/1993
49	424-07-0021	Bayou L'ourse	Terrebonne	1	30	-114.5	2	9.5	124	116.5	40.4	D46-02	-	-
50	434-01-0002	Mississippi River Bridge at Gramercy (West Approaches)	-	3	14	-46.45	17.55	40.55	87	64	28.3	ICE 42S Single Acting Diesel - OED	12/1/1992	1/5/1993
51	450-15-0085	I-10 Williams Boulevard Interchange	Jefferson	3-A	14	-64	11.5	23	87	75.5	8.9	-	-	-

No.	Project ID	Project Name	Parish	Pile ID	Diam. (inch)	Tip elevation (ft)	Ground elevation (ft)	Cut-off elevation (ft)	Pile length (ft)	Embedment length (ft)	Weight of Pile (Tons)	Hammer type	Date of driving	Date of loading
52	450-15-0100	Causeway Boulevard interchange (Phase I)	Jefferson	1	14	-45	-3	0	45	42	4.6	APE D30-42 (EOD)	2/10/2009	2/26/2009
53	450-15-0100	Causeway Boulevard interchange (Phase I)	Jefferson	2	14	-45	-3	0	45	42	4.6	APE D30-42 (EOD)	3/4/2009	3/19/2009
54	450-15-0100	Causeway Boulevard interchange (Phase I)	Jefferson	3	14	-80.87	-3	-0.87	80	77.87	8.2	APE D30-42 (EOD)	11/24/2009	12/8/2009
55	450-15-0103	Causeway Boulevard interchange (Phase II)	Jefferson	1	14	-43	-3	-1	42	40	4.3	APE D30-42 (EOD)	12/22/2009	1/5/2010
56	450-15-0103	Causeway Boulevard interchange (Phase II)	Jefferson	2	14	-84.2	-3	0	84.2	81.2	8.6	APE D30-42 (EOD)	-	-
57	450-15-0103	Causeway Boulevard interchange (Phase II)	Jefferson	5	14	-85.4	-3.5	0	85.4	81.9	8.7	APE D30-42 (EOD)	-	-
58	450-15-0103	Causeway Boulevard interchange (Phase II)	Jefferson	7	14	-68.23	-3	-1.23	67	65.23	6.8	APE D30-42 (EOD)	-	-
59	450-36-0002	Luling Bridge (North Approach)-US61	St. Charles	8	30	-110	2	10	120	112	39.1	Vulcan 010 (MOD)	2/4/1991	3/21/1991
60	455-05-0036	Sugarhouse Road	-	1	14	5	75	75	70	70	7.1	Vulcan06 Single Acting Diese	5/29/1990	6/26/1990 7/12/1990
61	455-05-0036	Sugarhouse Road	-	2	14	5	78	80	75	73	7.7	Vulcan06 Single Acting Diese	5/30/1990	6/21/1990 7/10/1990
62	455-05-0036	Sugarhouse Road	-	3	14	14	78	79	65	64	6.6	Vulcan06 Single Acting Diese	6/6/1990	6/21/1990 7/10/1990
63	713-48-0083	Bayou Milhomme Bridge & Approaches	St. Martin	1	24	-55	-20	17	72	35	17.4	Delamg D30-32 OED	7/24/2000	8/8/2000
64	713-48-0083	Bayou Milhomme Bridge & Approaches	St. Martin	2	24	-85	-20	11	96	65	23.2	Delamg D30-32 OED	7/24/2000	8/8/2000
65	742-06-0073 (742-37-0010)	North Eighteenth Street	Ouachit Parish	2	14	21.08	79.58	91.08	70	58.5	7.1	Delamg D19-32 OED	5/22/2000	6/6/2000

No.	Project ID	Project Name	Parish	Pile ID	Diam (inch)	Tip elevation (ft)	Ground elevation (ft)	Cut-off elevation (ft)	Pile length (ft)	Embedment length (ft)	Weight of Pile (Tons)	Hammer type	Date of driving	Date of loading
66	829-10-0013	Hwy 308 & Hwy 654	Terrebonne	1	24	-40	6	6	46	46	11.1	-	-	-
67	855-14-0003 (Houma)	Intercoastal waterway Bridge (Project Avenue Extension)		1	18	-95	0	0	95	95	16.0	-	-	-
68	SETUP	Bayou Boeuf US 90	Morgan City	3	30	-130	0	12	142	130	46.2	ICE I-62v2		
69	SETUP H.002071 (196-03-0030)	Bayou Lacassine	Jefferson Davis	1	30	-65	0	10	75	65	24.4	ICE I-62v2		
70	SETUP H.002071 (196-03-0030)	Bayou Lacassine	Jefferson Davis	3	30	-65	0	10	75	65	24.4	-		
71	SETUP	Bayou Zourie	Vernon	-	24	146.59	196.59	201.59	55	50	13.3	Vulcan 010		
72	SETUP	LA-01		2	16	-120	0	10	130	120	17.3	Vulcan 010		
73	SETUP	LA-01		3	30	-180	0	10	190	180	61.8	Vulcan 020		
74	SETUP	LA-01		4a	24	-150	0	10	160	150	38.6	Vulcan 020		
75	SETUP	LA-01		4b	24	-200	0	10	210	200	50.6	Vulcan 020		
76	SETUP	LA-01		5a	24	-139	0	6	145	139	35.0	Vulcan 020		
77	SETUP	LA-01		5b	24	-163	0	7	170	163	41.0	Vulcan 020		
78	008-01-0042	La 415 – La 983	West Baton Rouge	1	16	-52.66	-2	28.59	81.25	50.66	10.8	ICE 60S single acting Diesel	10/4/2000	10/19/2000
79	450-17-0025	Interstate Twin Span over Lake Pontchartrain	Orleans and St. Tammany	1	36	-119.3	0.6		141	119.9	66.0	Vulcan 30 single action air hammer	8/21/2000	9/11/2000
80	450-17-0025	Interstate Twin Span over Lake Pontchartrain	Orleans and St. Tammany	3	36	-130.17	0.4		141	130.57	66.0	Vulcan 30 single action air hammer	9/29/2000	0/31/2006

Appendix C

Summary of DOTD State Projects Investigated in this Study

Figure 70. 1) 003-07-0019 TP#1

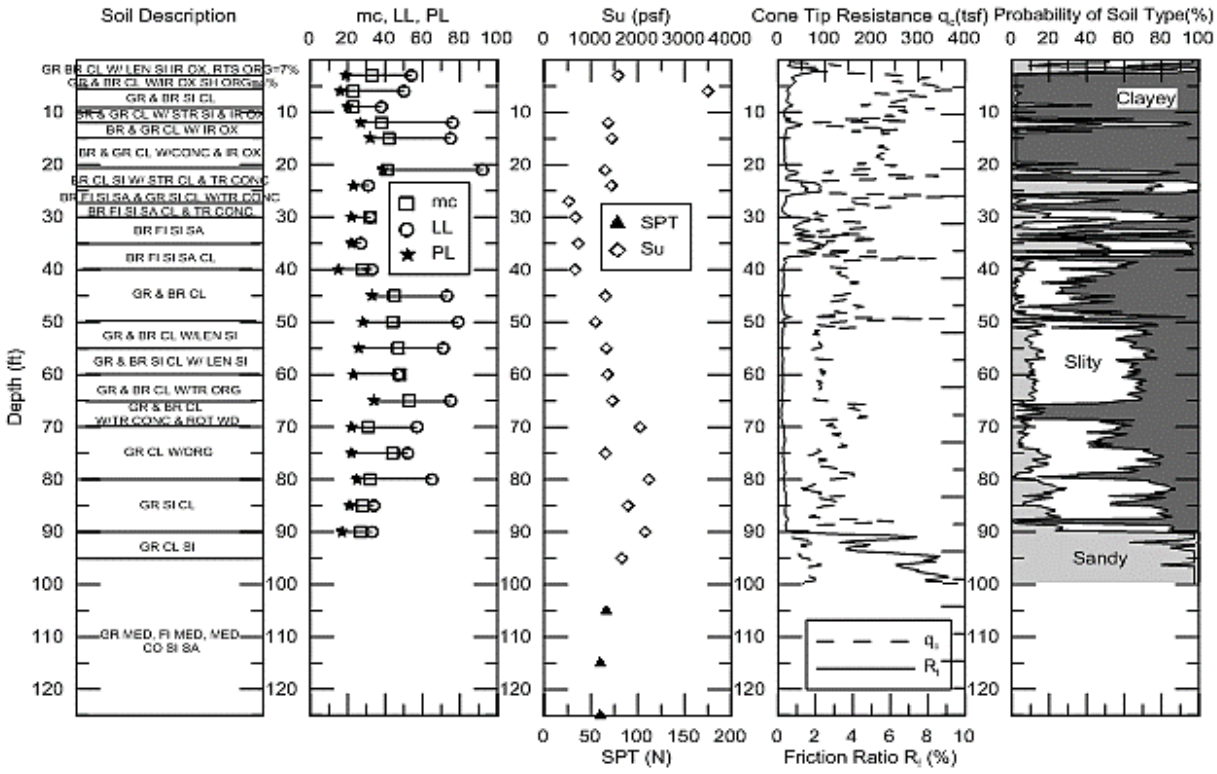


Figure 71. 2) 003-10-0011 TP#1

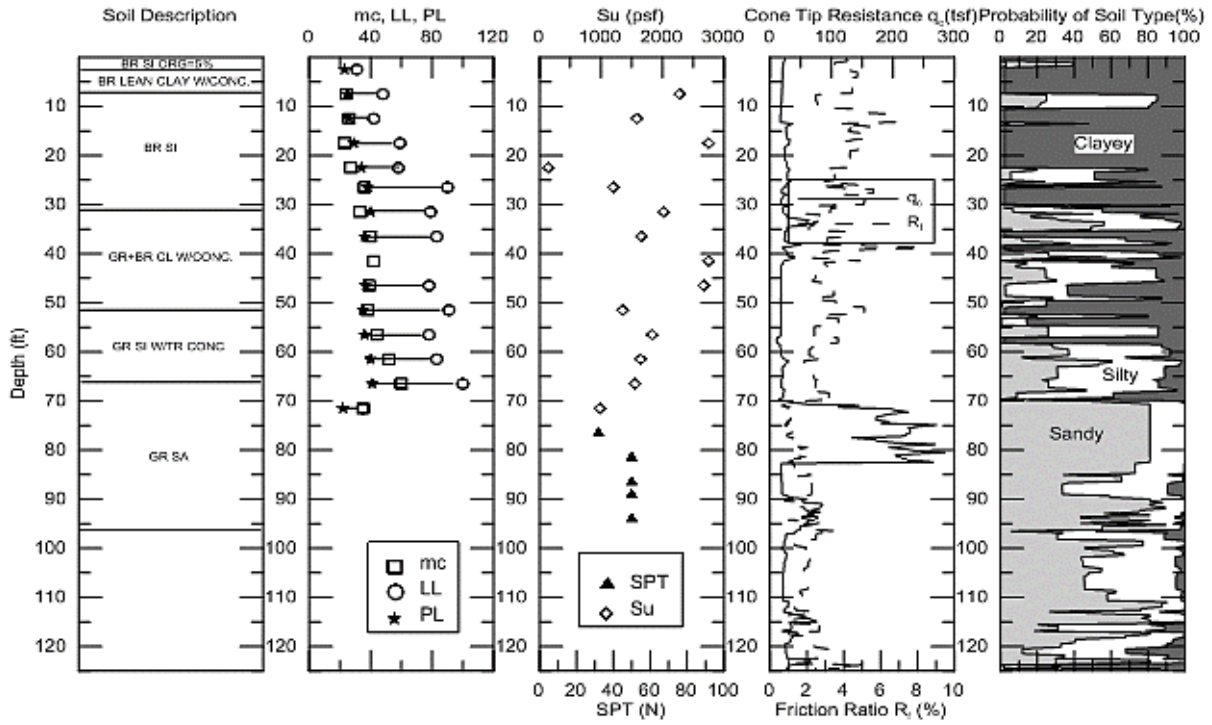


Figure 72. 3) 003-10-0011 TP#3

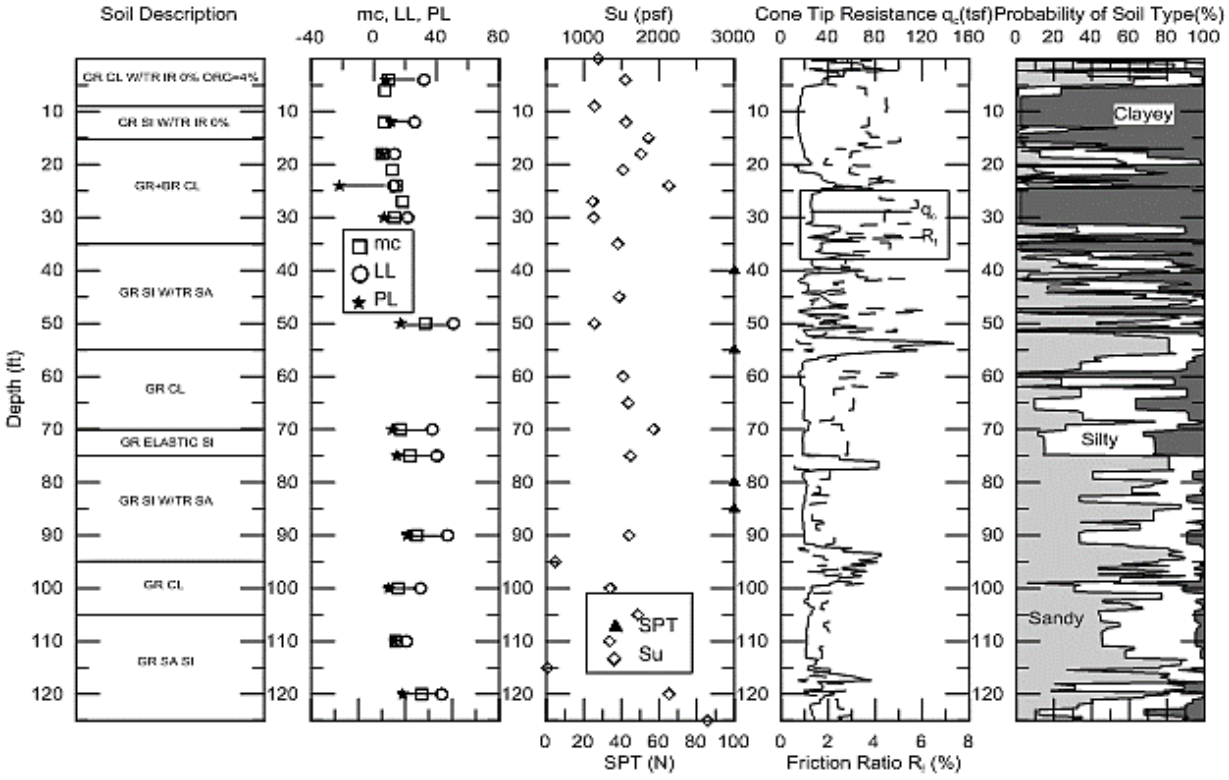


Figure 73. 4) 005-01-0056 TP#1

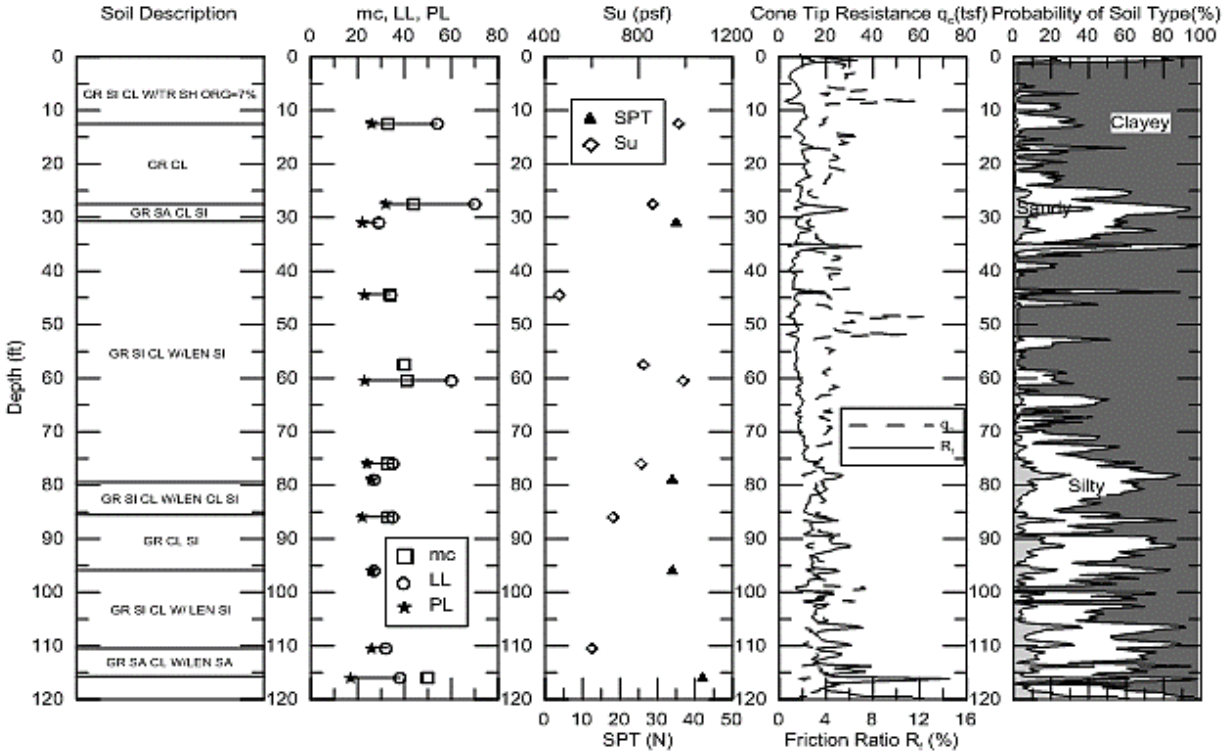


Figure 74. 5) 005-01-0056 TP#2

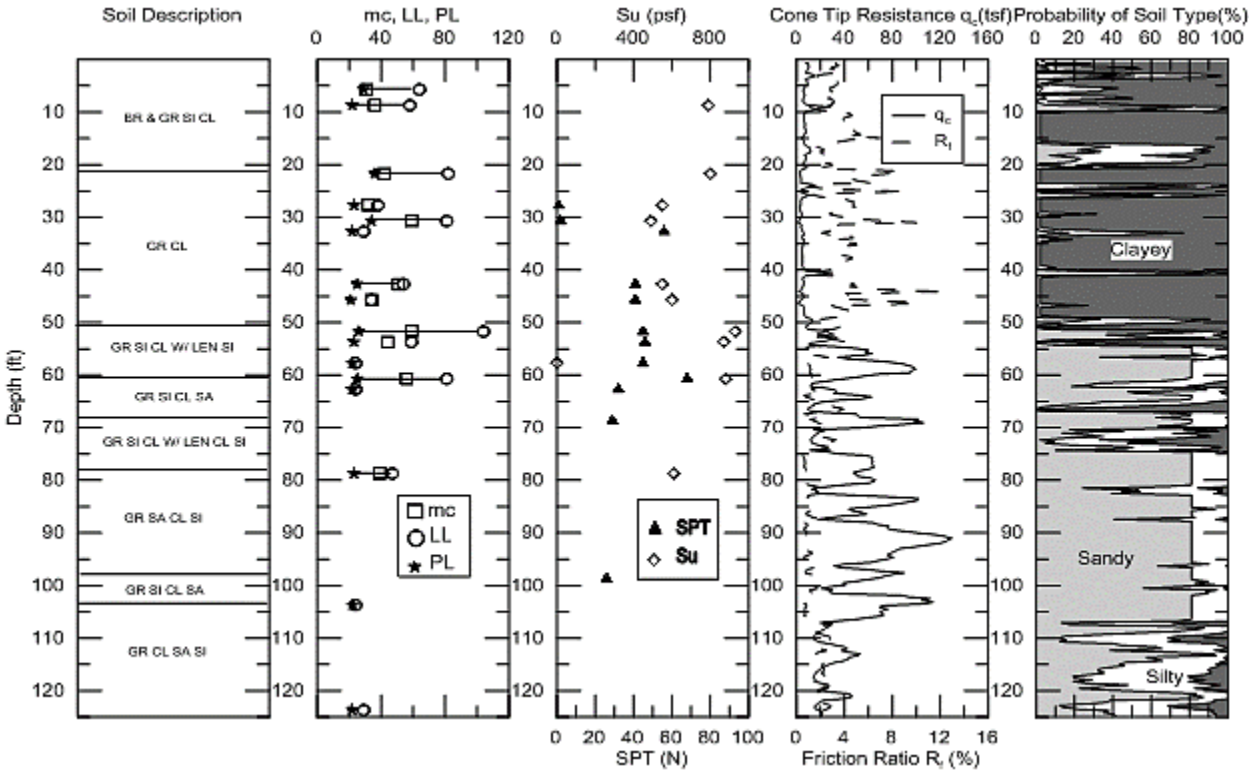


Figure 75. 6) 005-01-0056 TP#3

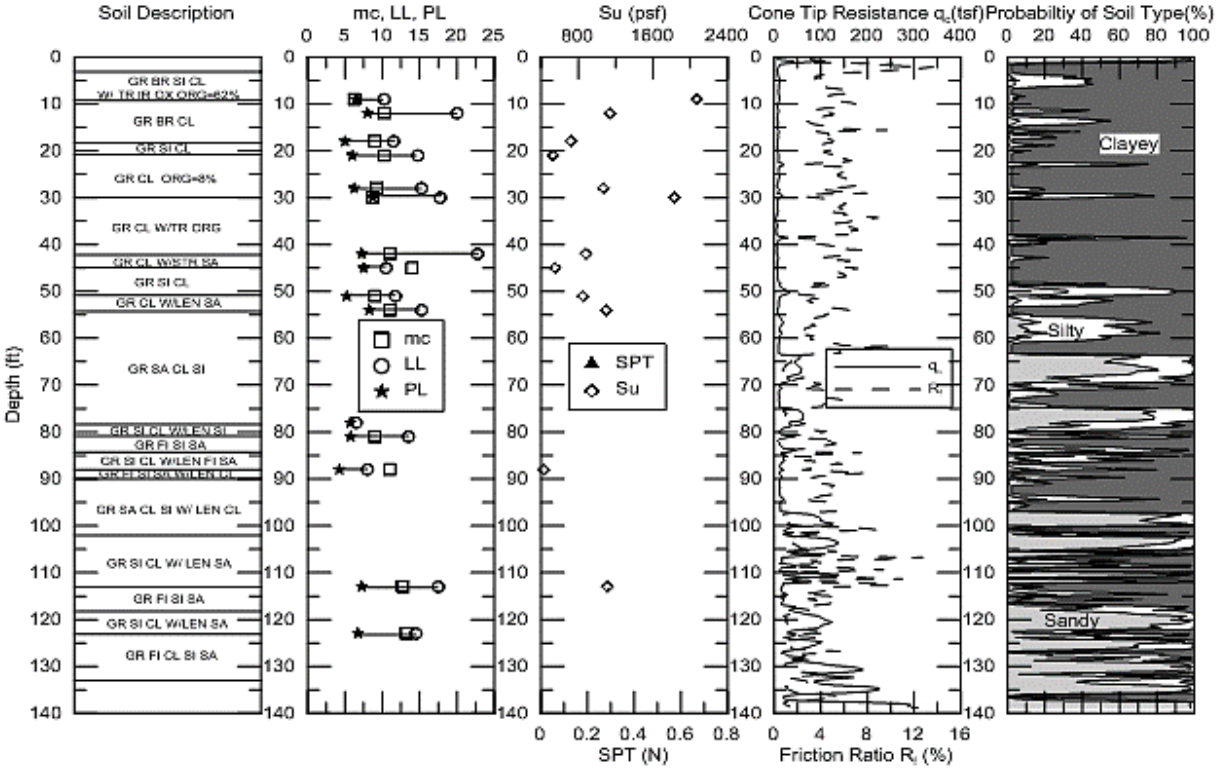


Figure 76. 7) 047-02-0022 TP#2

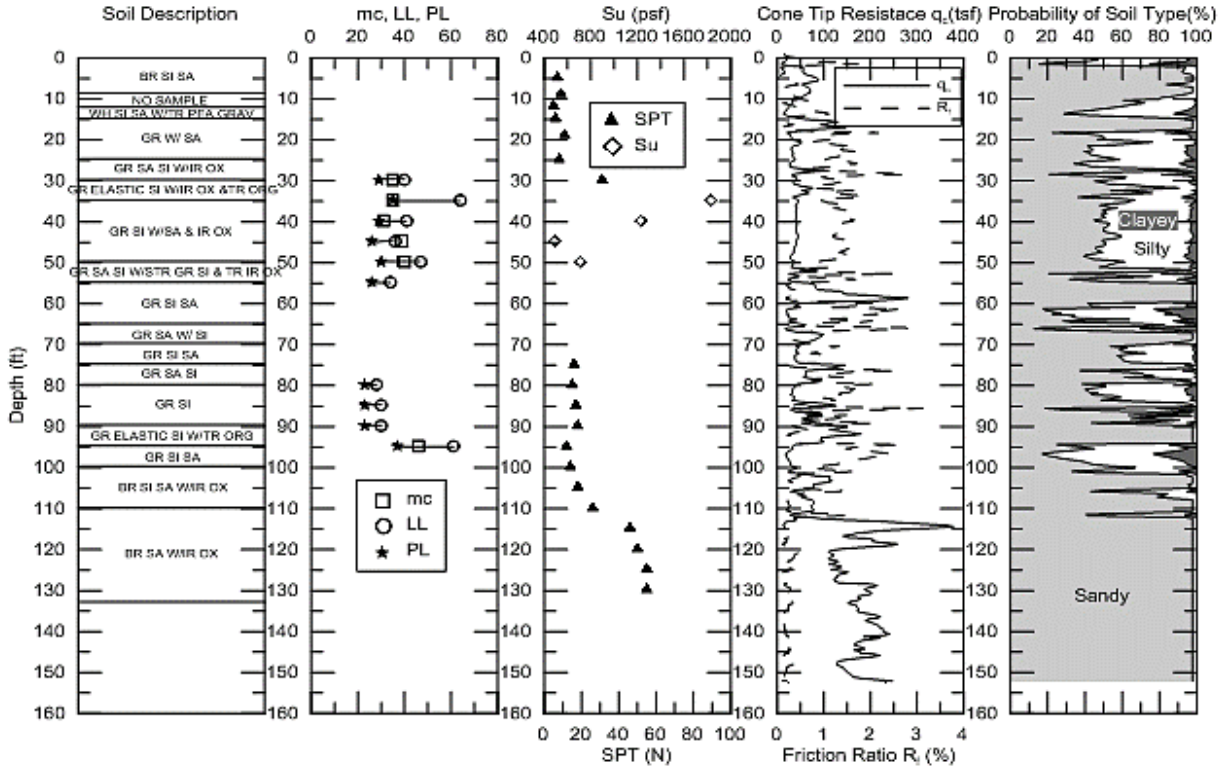


Figure 77. 8) 061-05-0044 TP#2

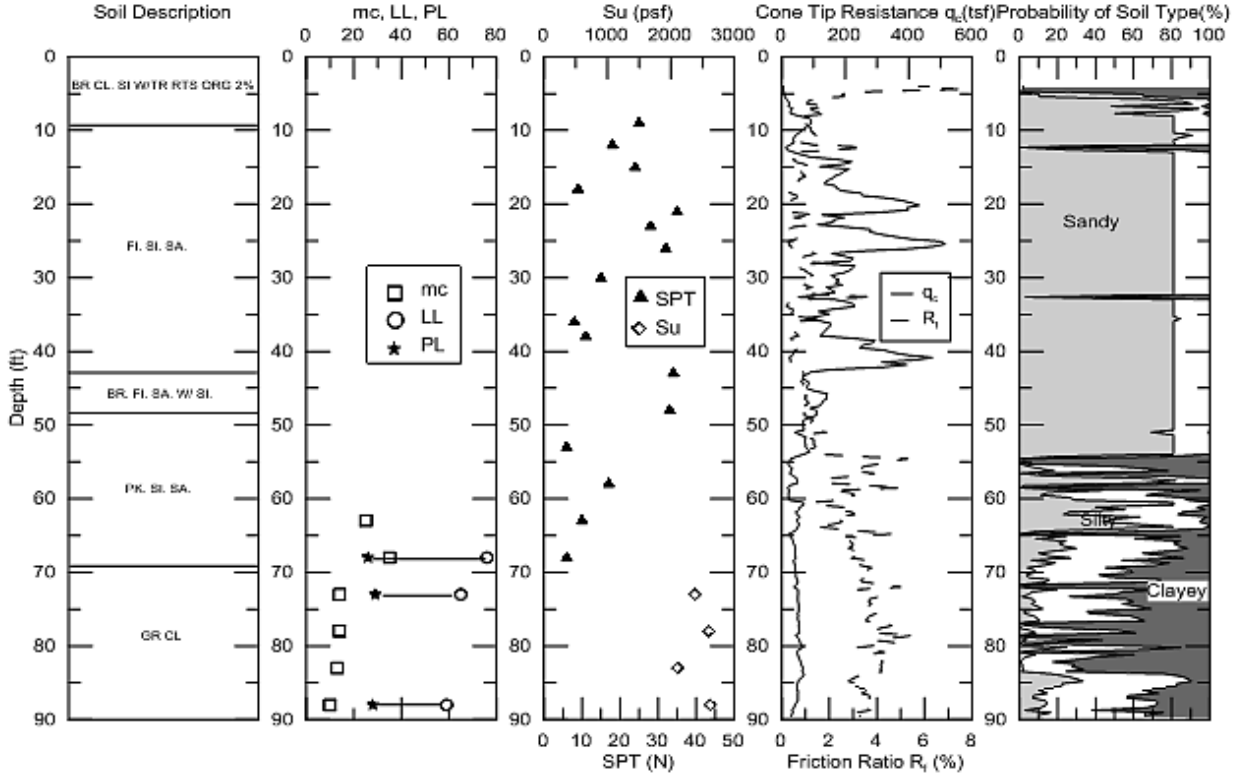


Figure 78. 9) 064-06-0036 TP#1

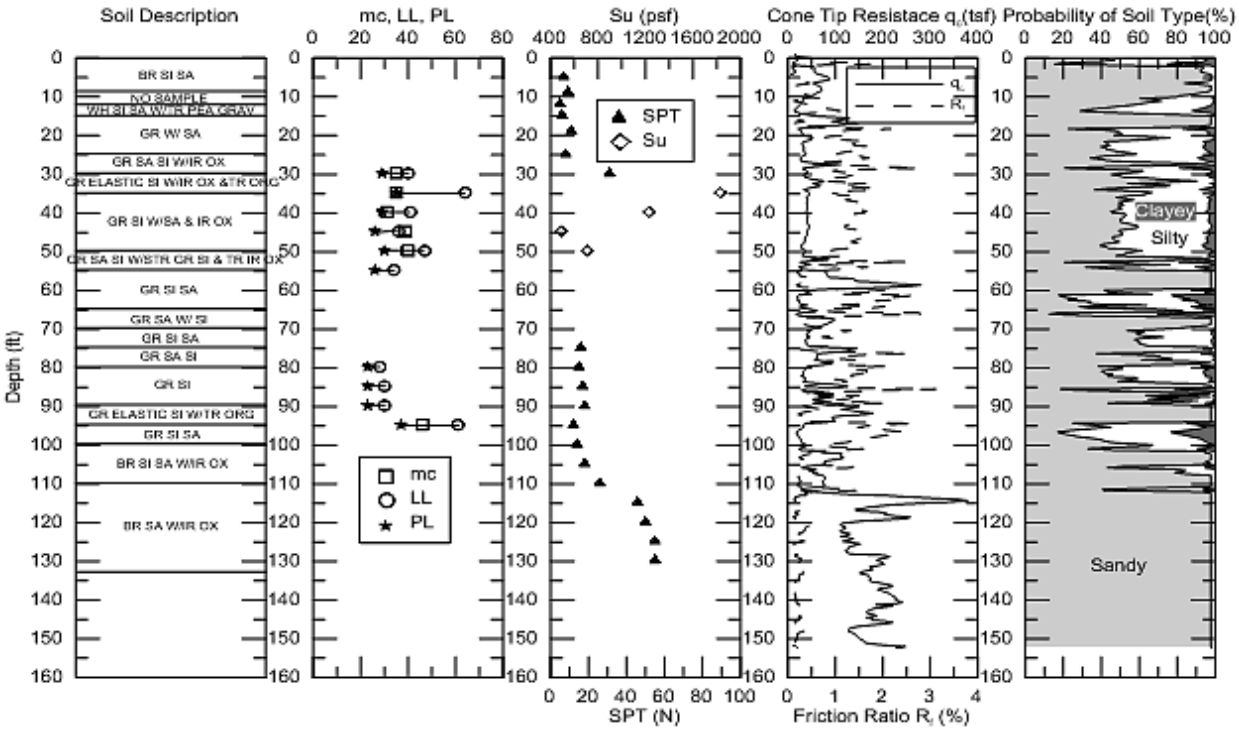


Figure 79. 10) 065-90-0024 TP#1

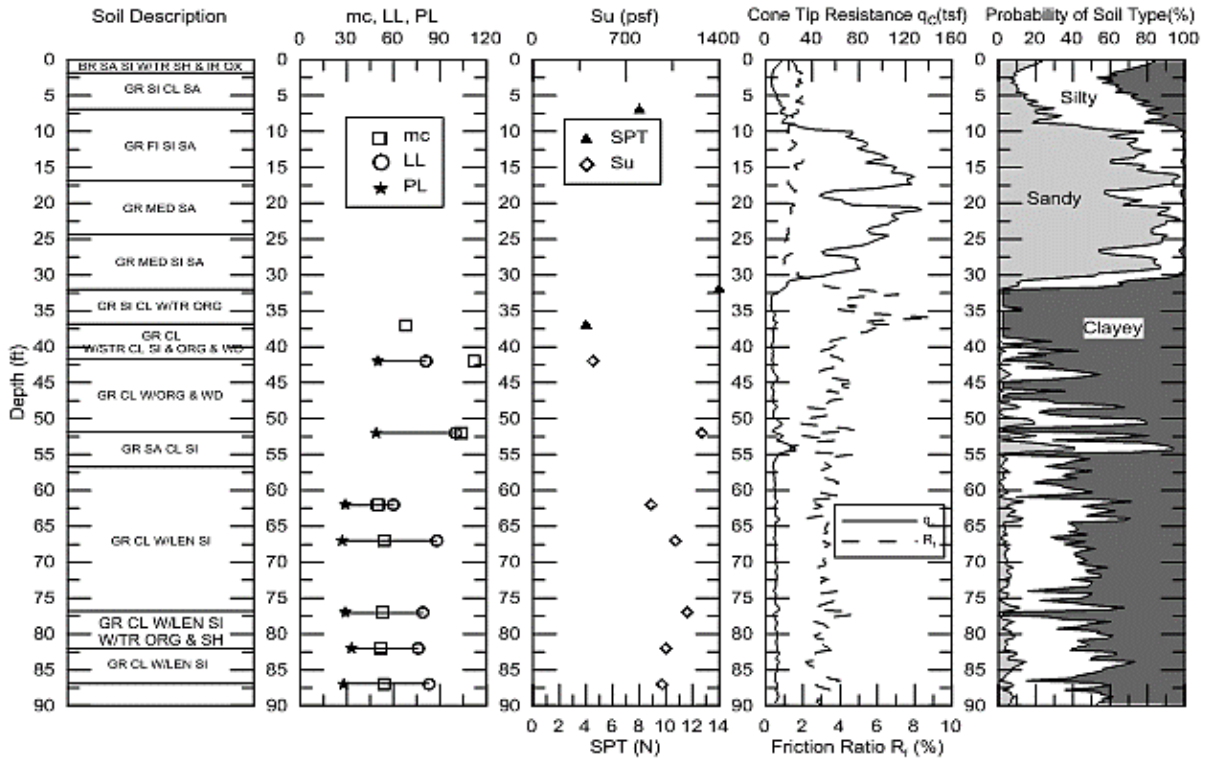


Figure 80. 11) 065-90-0024 TP#2

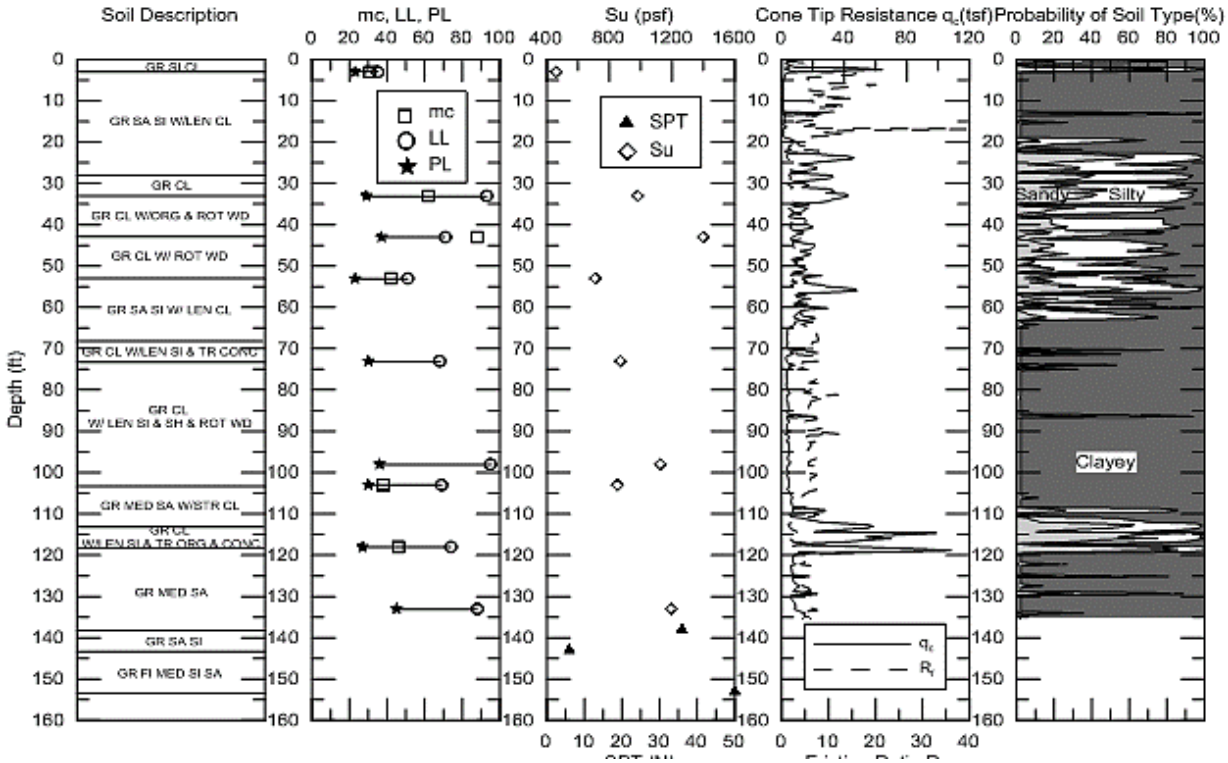


Figure 81. 12) 065-90-0024 TP#3

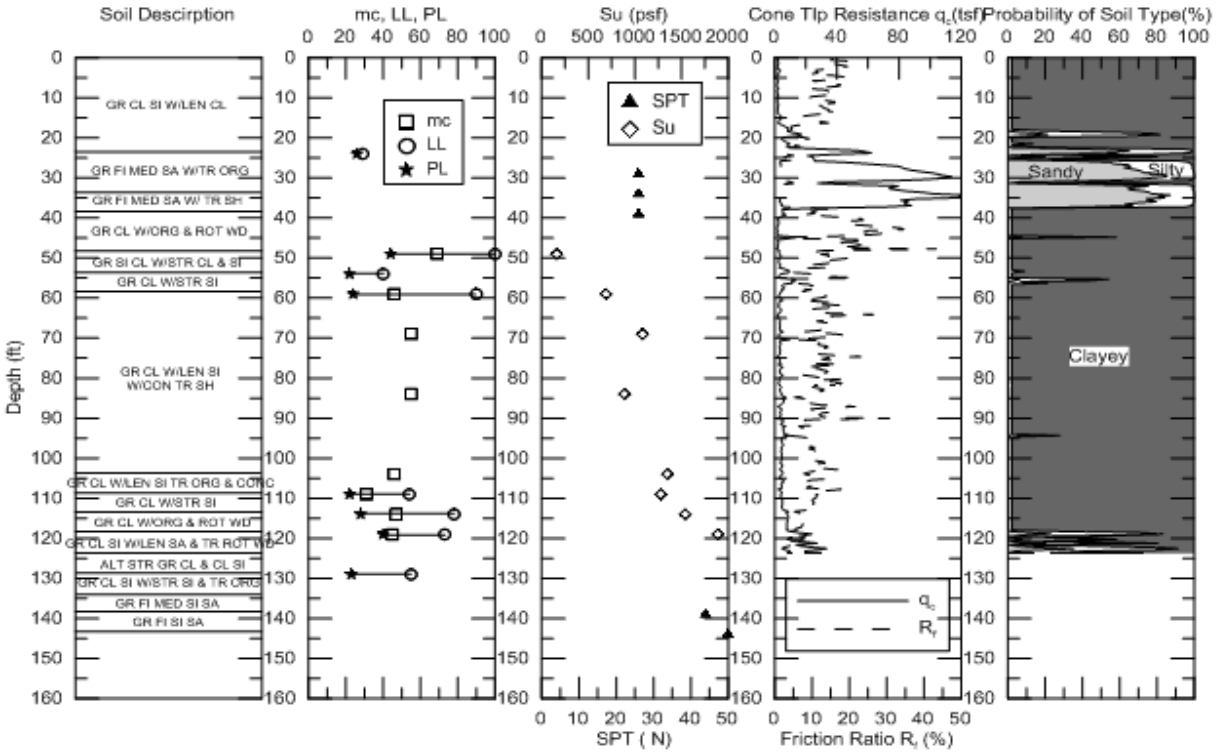


Figure 82 13) 065-90-0024 TP#4

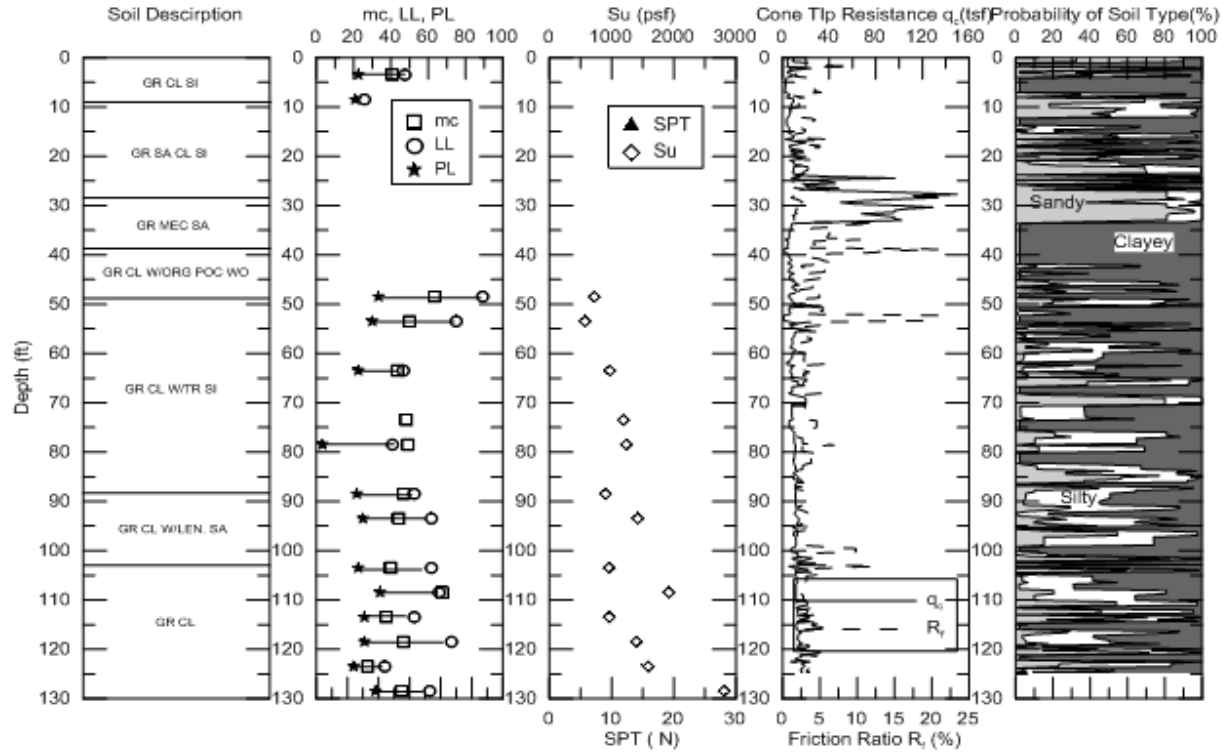


Figure 83. 14) 065-90-0024 TP#5

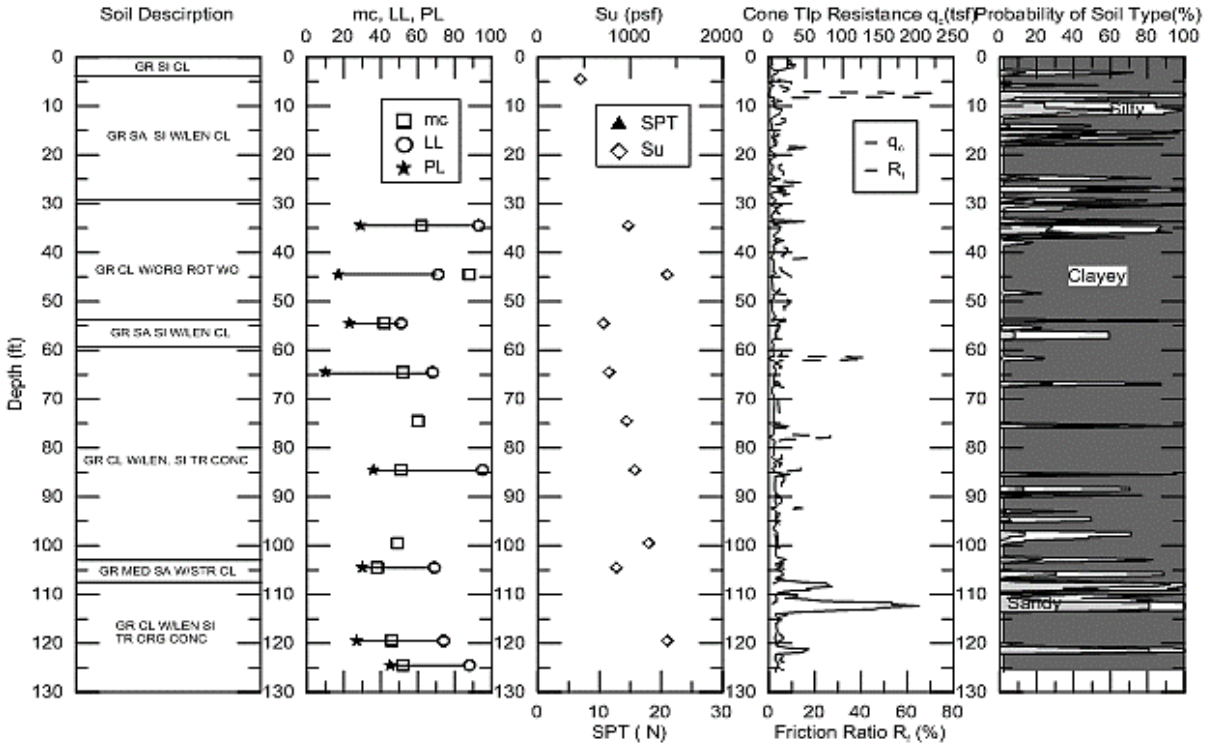


Figure 84. 15) 065-90-0024 TP#6

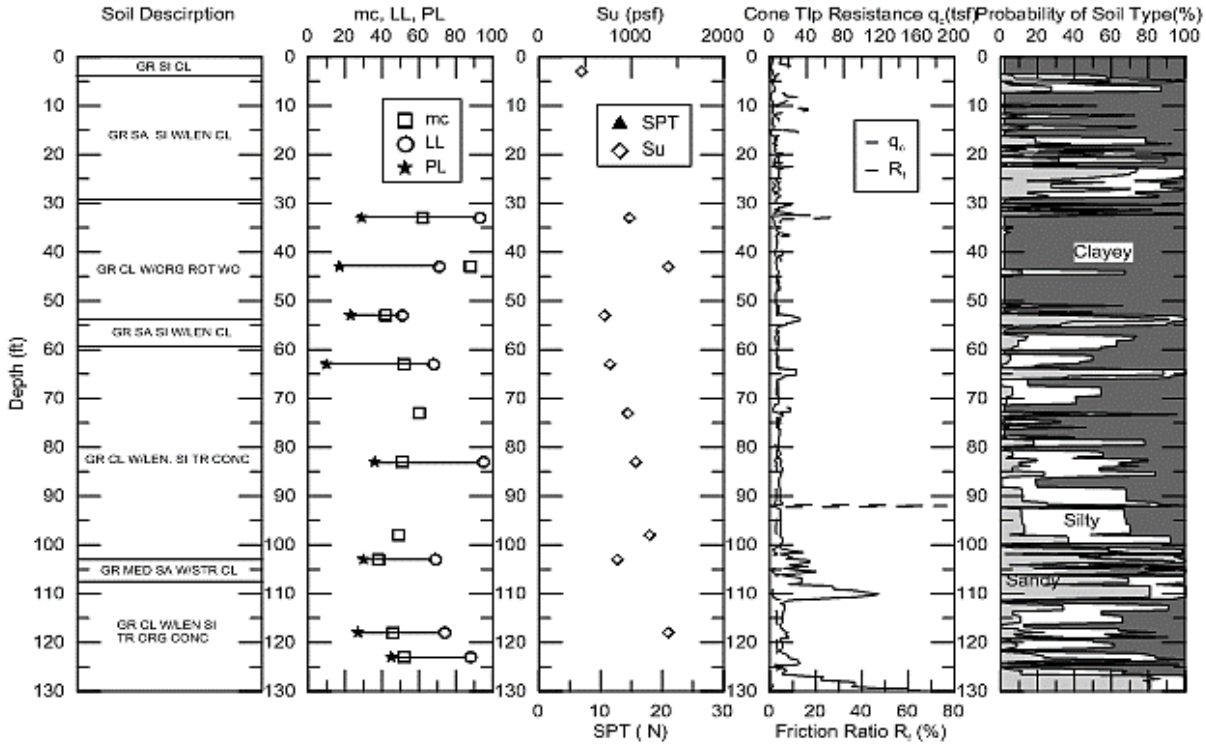


Figure 85. 16) 239-01-0080 TP#3

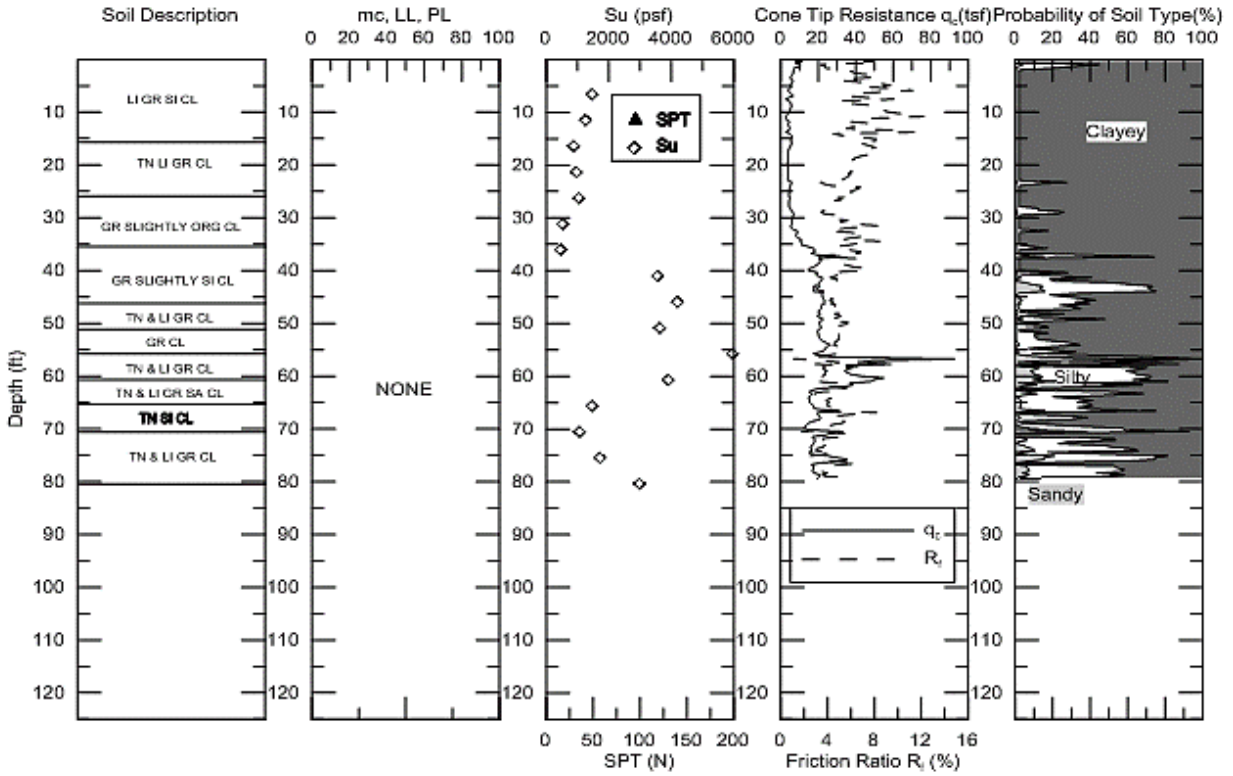


Figure 86. 17) 239-01-0080 TP#4

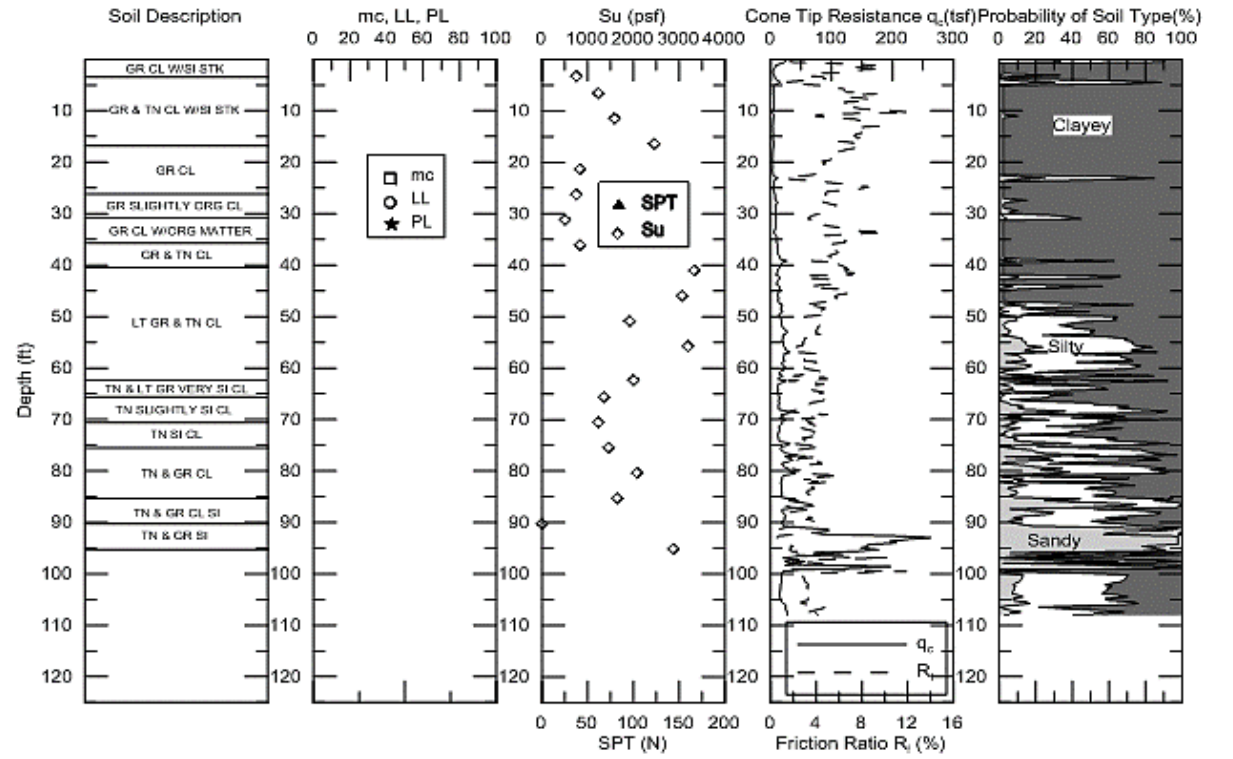


Figure 87. 18) 260-05-0020 TP#1

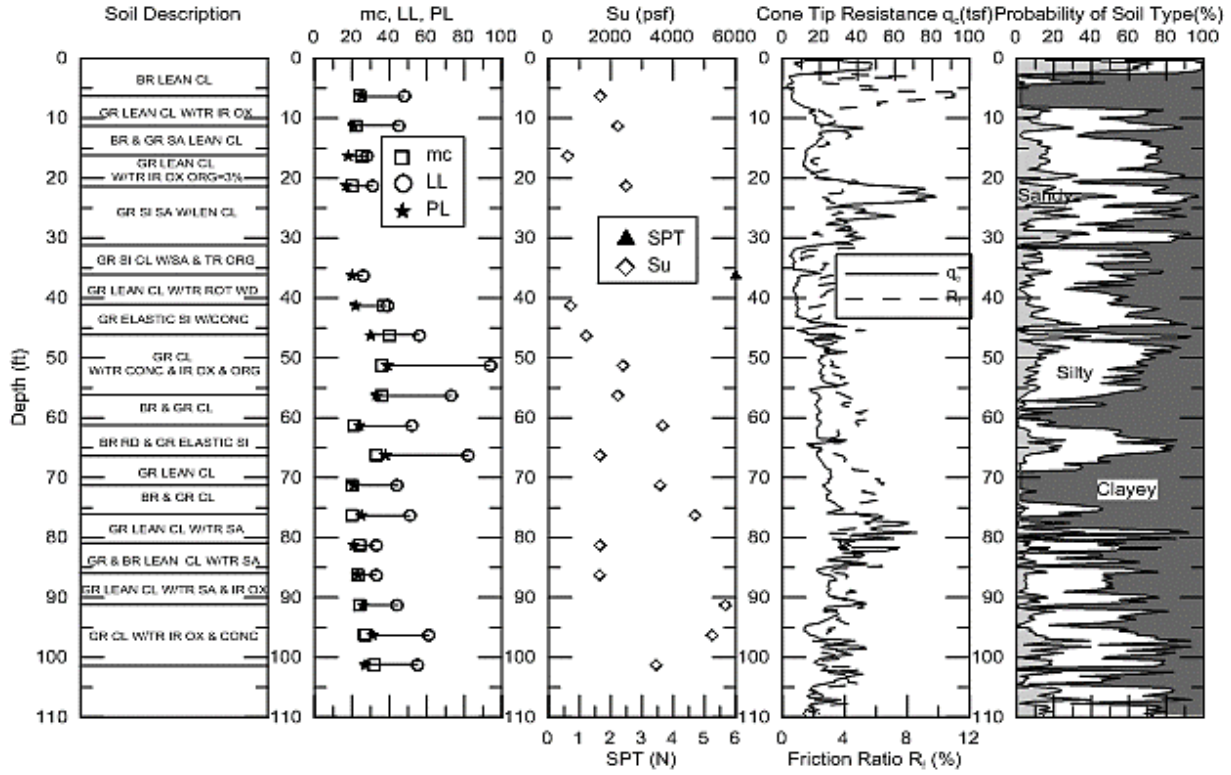


Figure 88. 19) 260-05-0020 TP#3

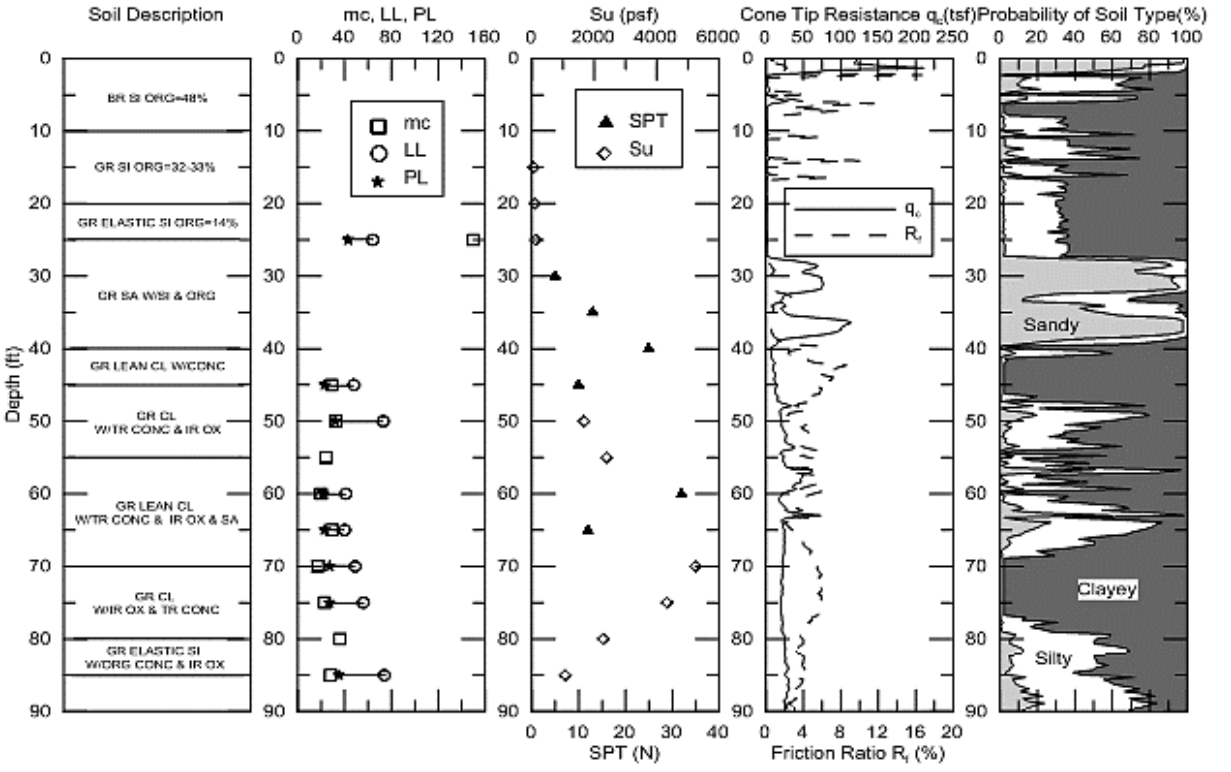


Figure 89. 20) 262-06-0009 TP#1

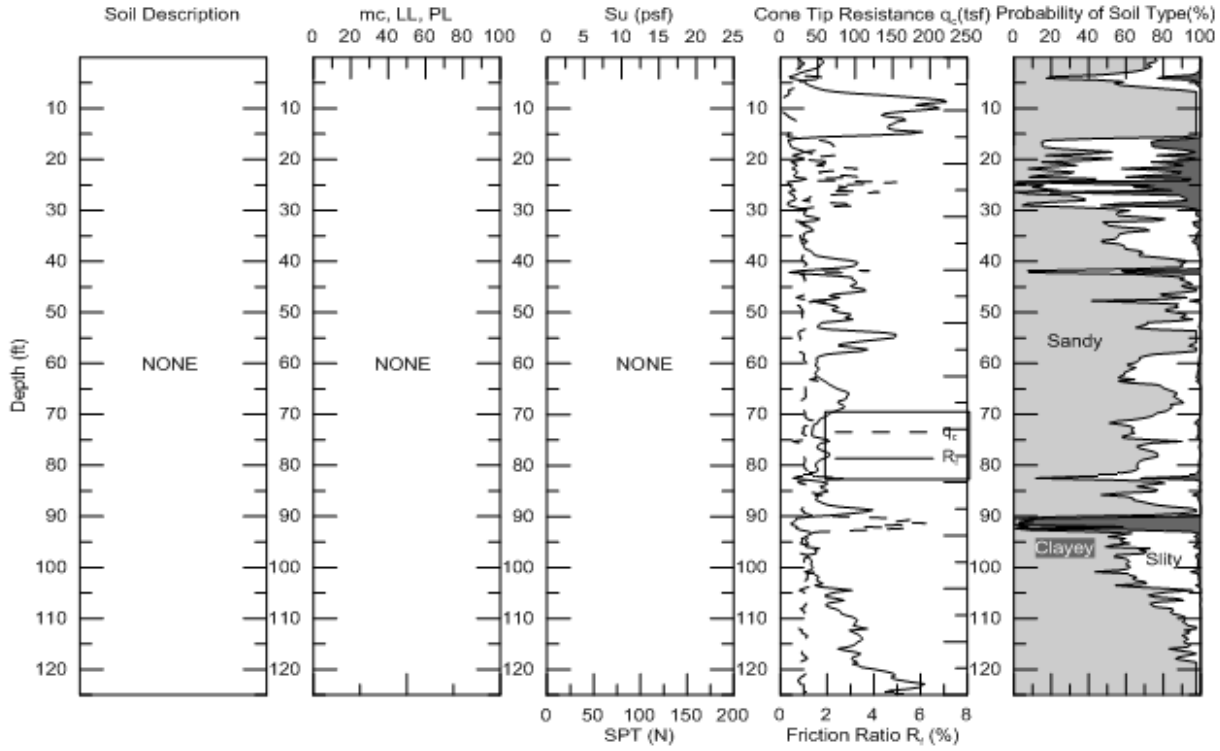


Figure 90. 21) 262-06-0009 TP#2

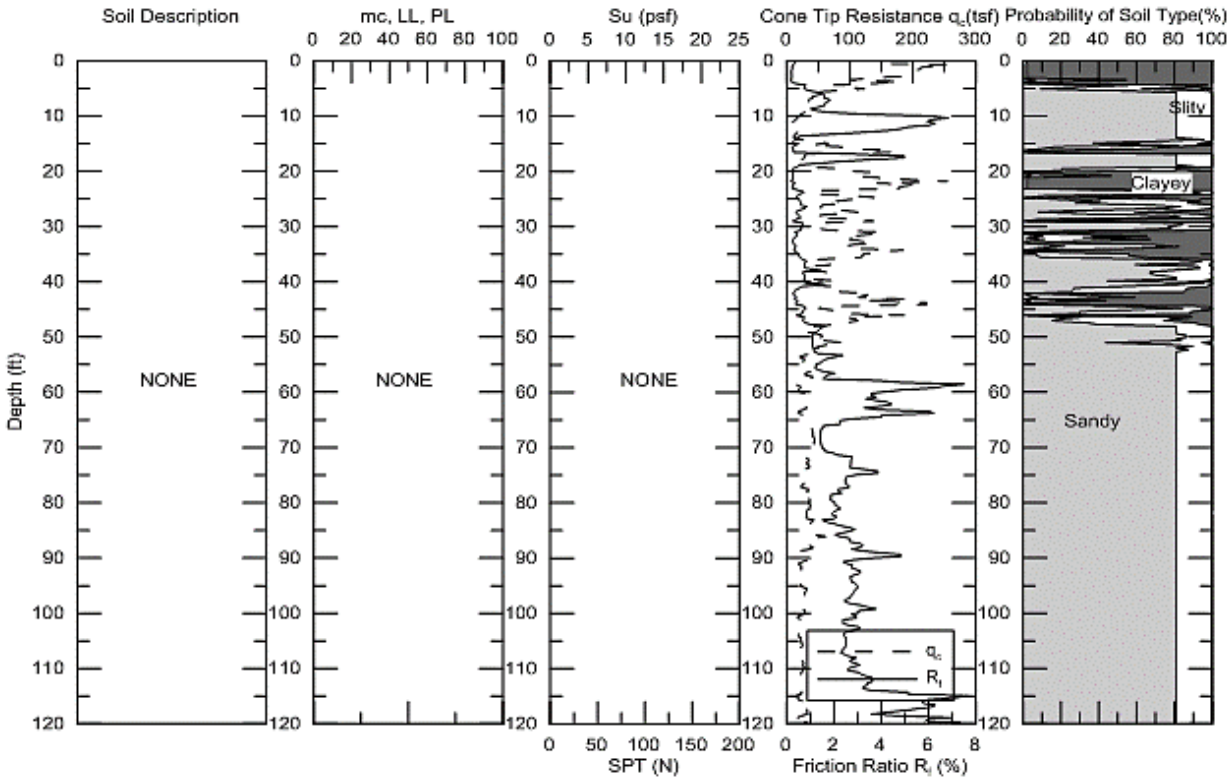


Figure 91. 22) 283-03-0052 TP#1

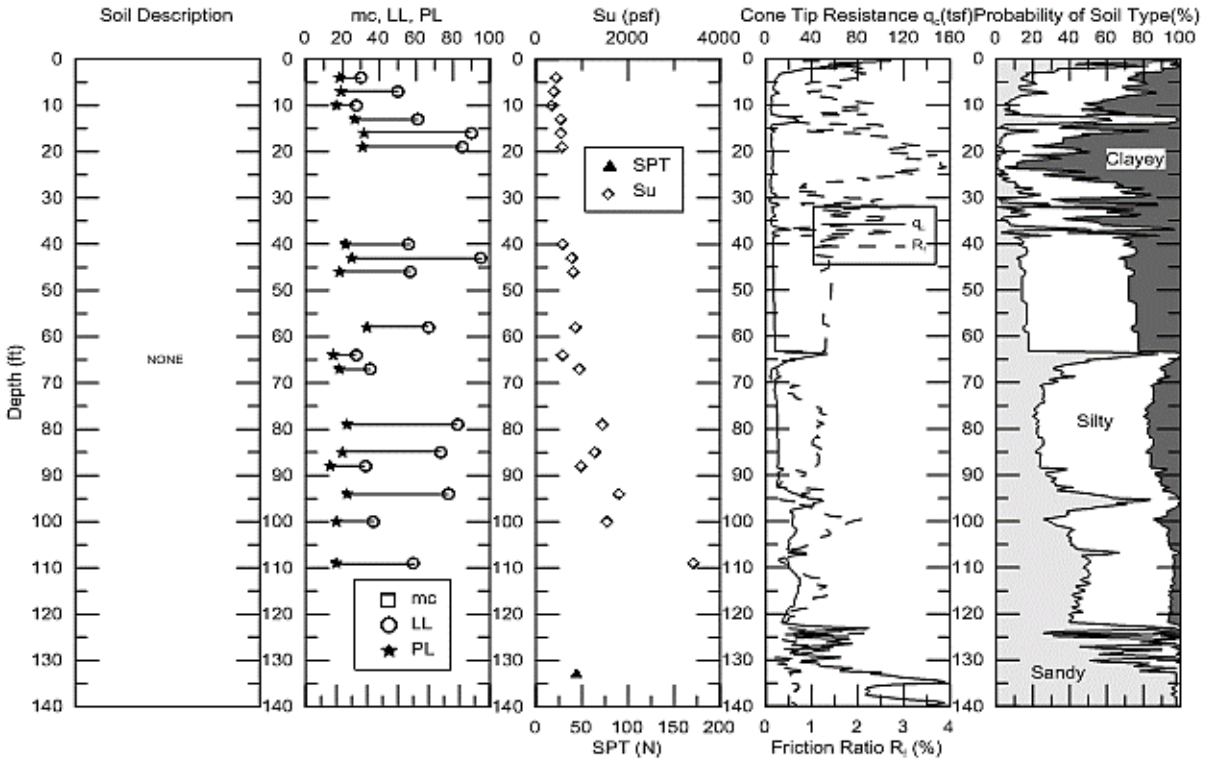


Figure 92. 23) 424-04-0026 TP#2

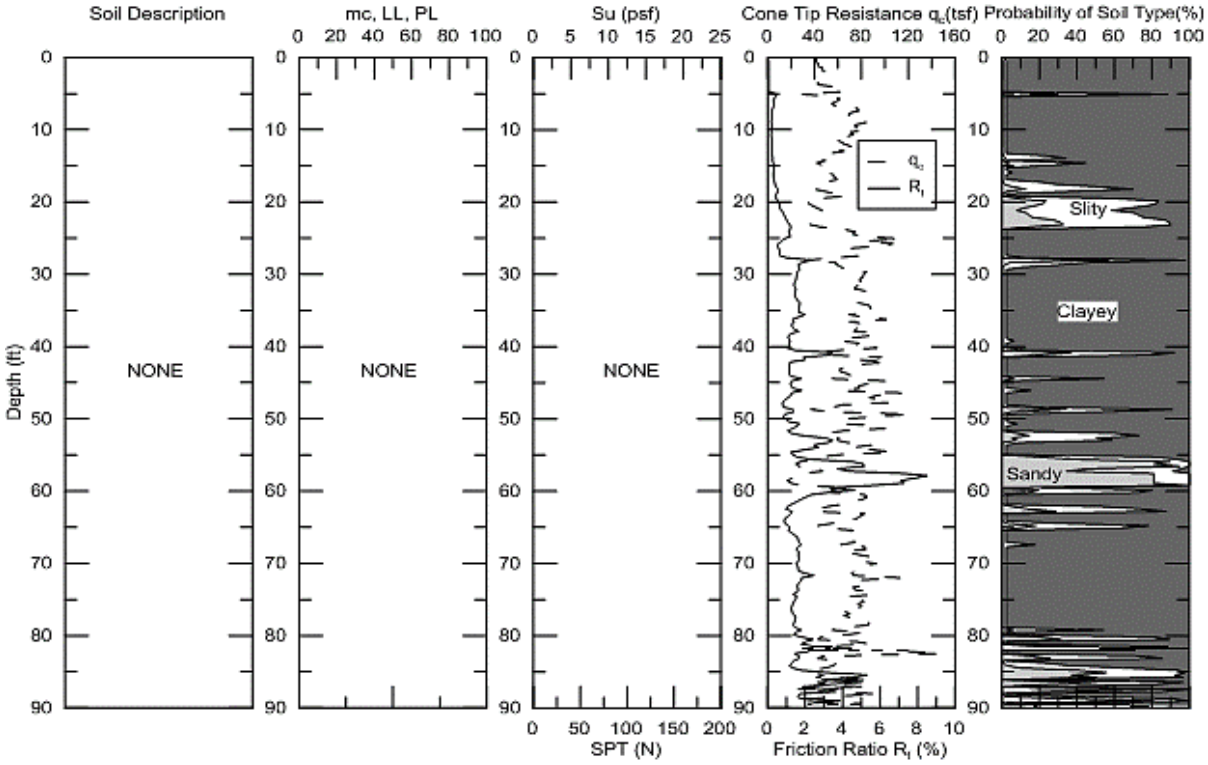


Figure 93. 24) 424-04-0027 TP#1

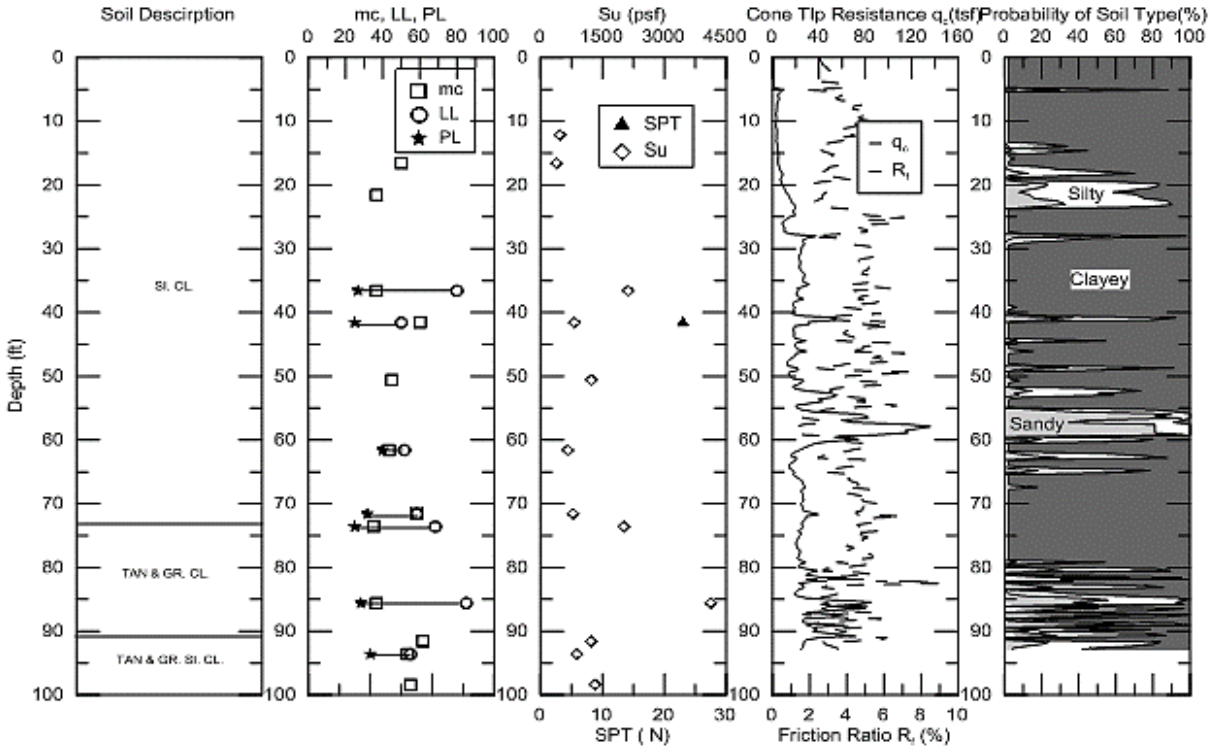


Figure 94. 25) 424-04-0027 TP#2

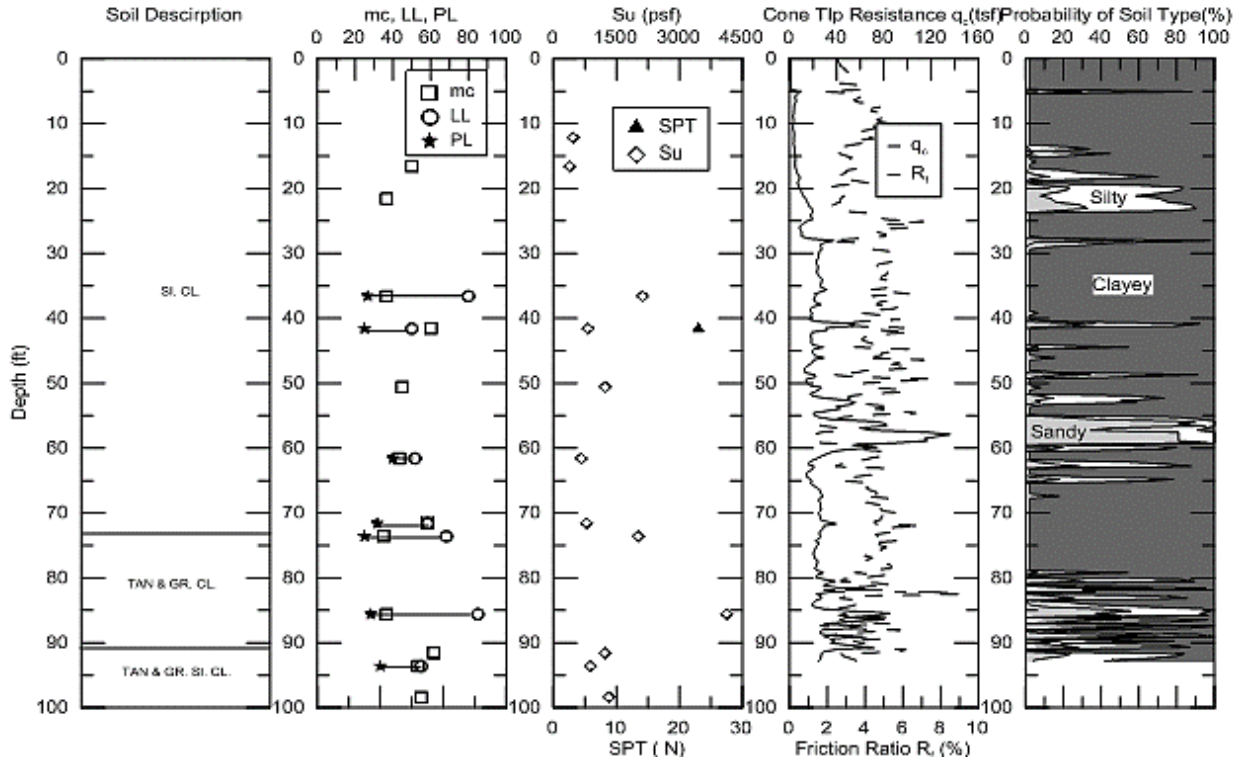


Figure 95. 26) 424-05-0078 TP#1

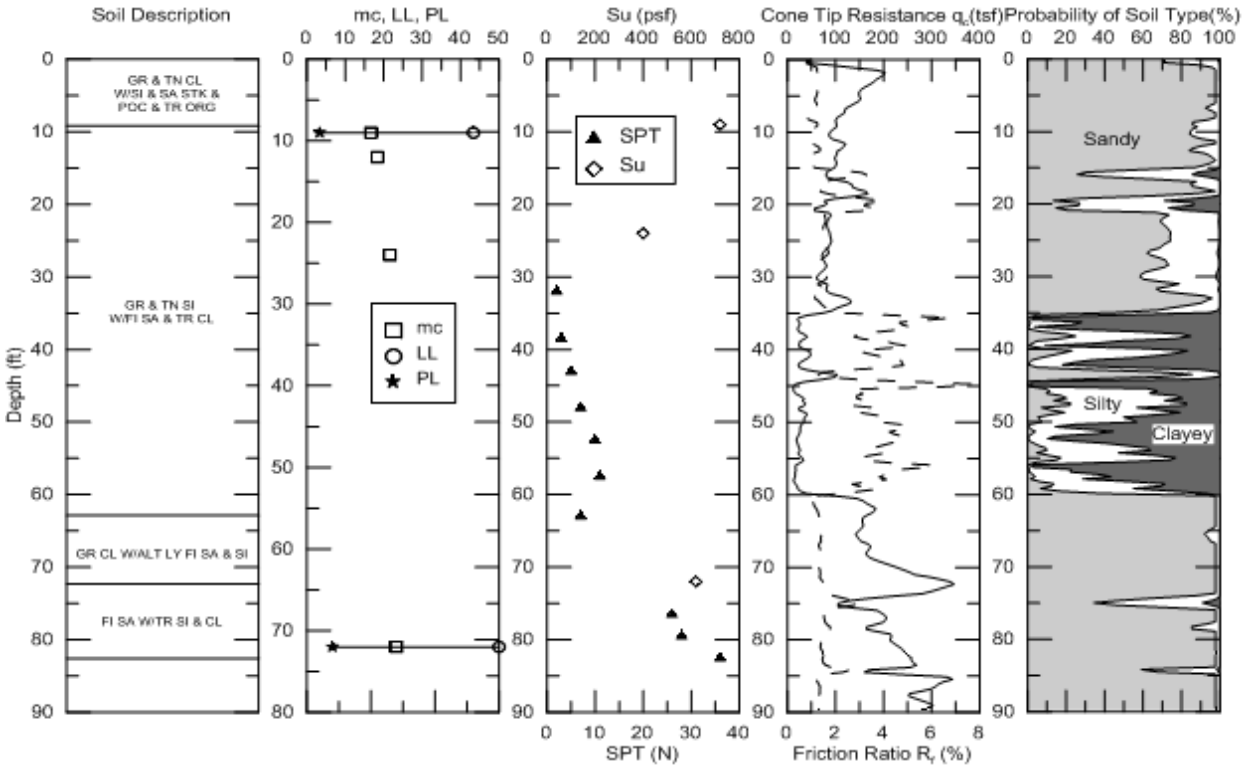


Figure 96. 27) 424-05-0078 TP#2

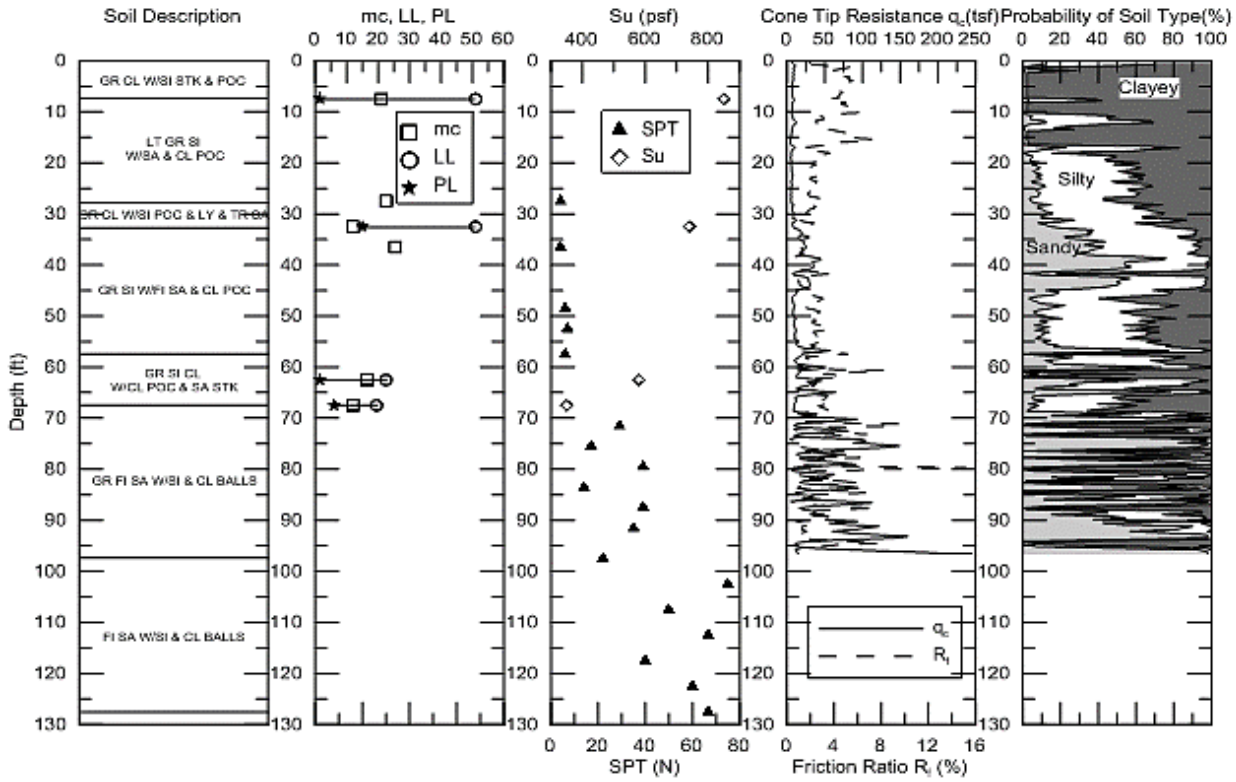


Figure 97. 28) 424-05-0078 TP#5

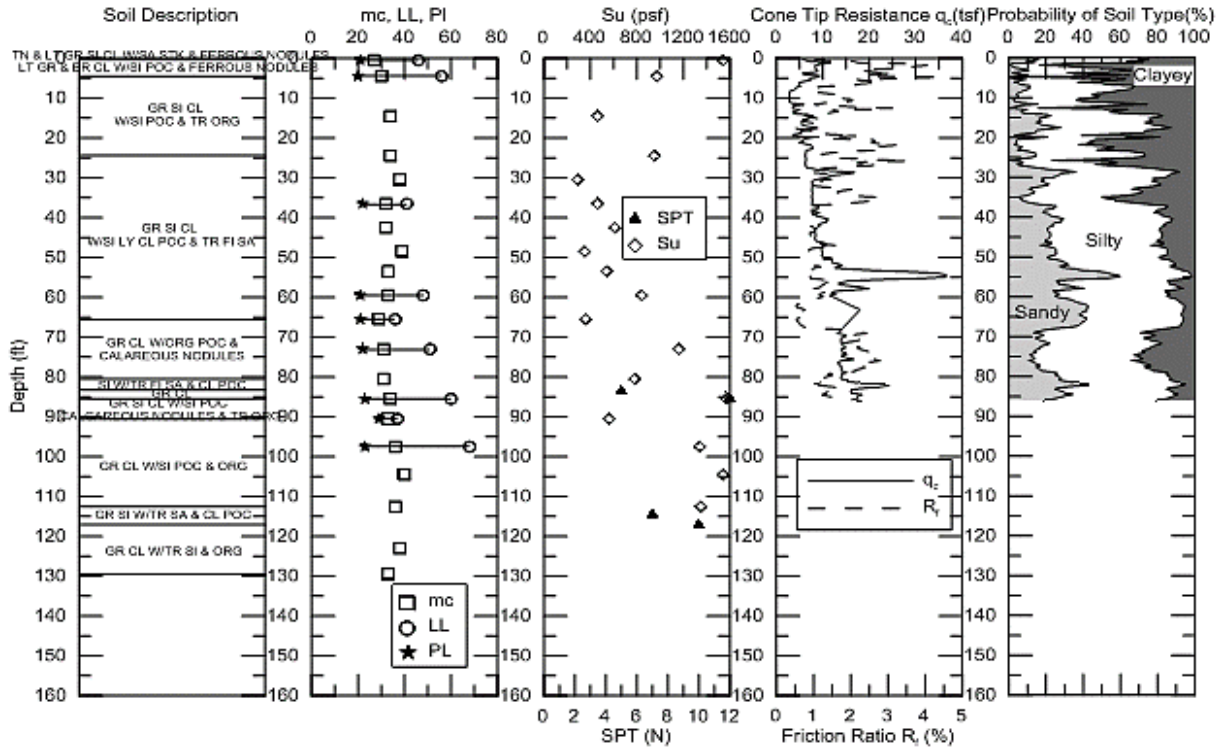


Figure 98. 29) 424-05-0081 TP#1

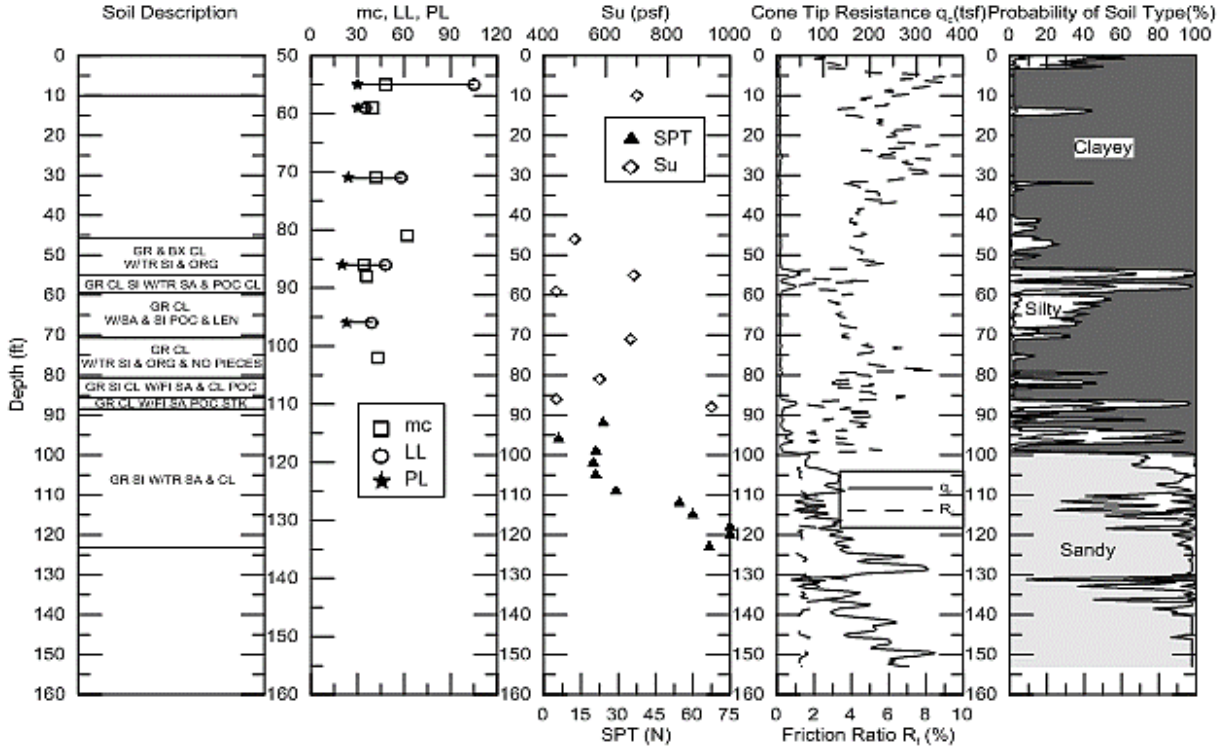


Figure 99. 30) 424-05-0081 TP#2

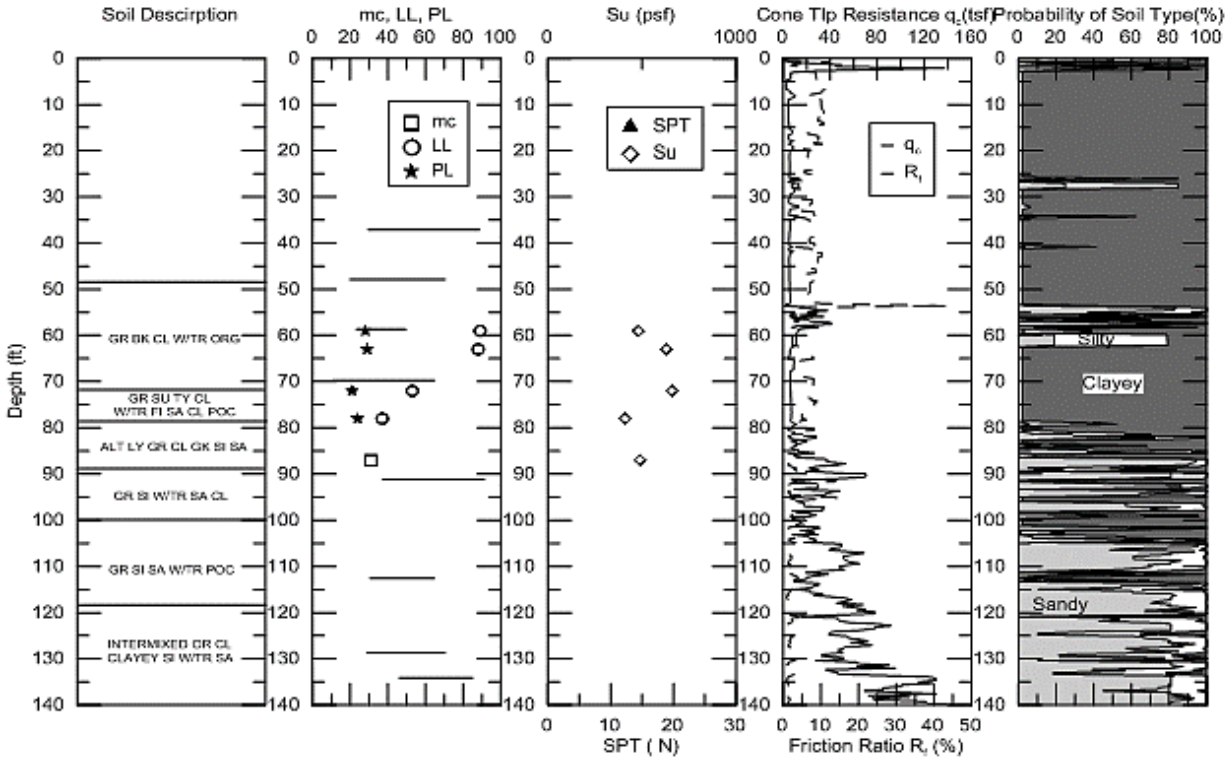


Figure 100. 31) 424-05-0081 TP#3

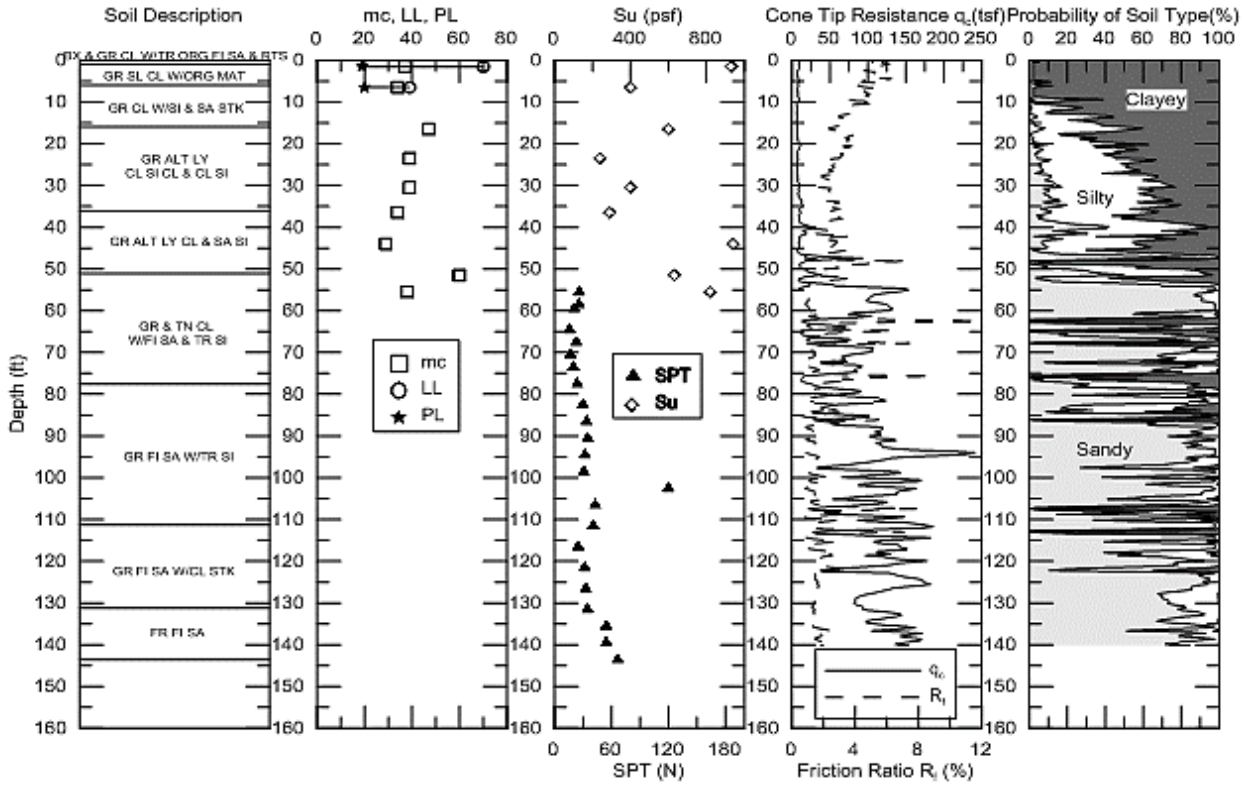


Figure 101. 32) 424-05-0081 TP#4

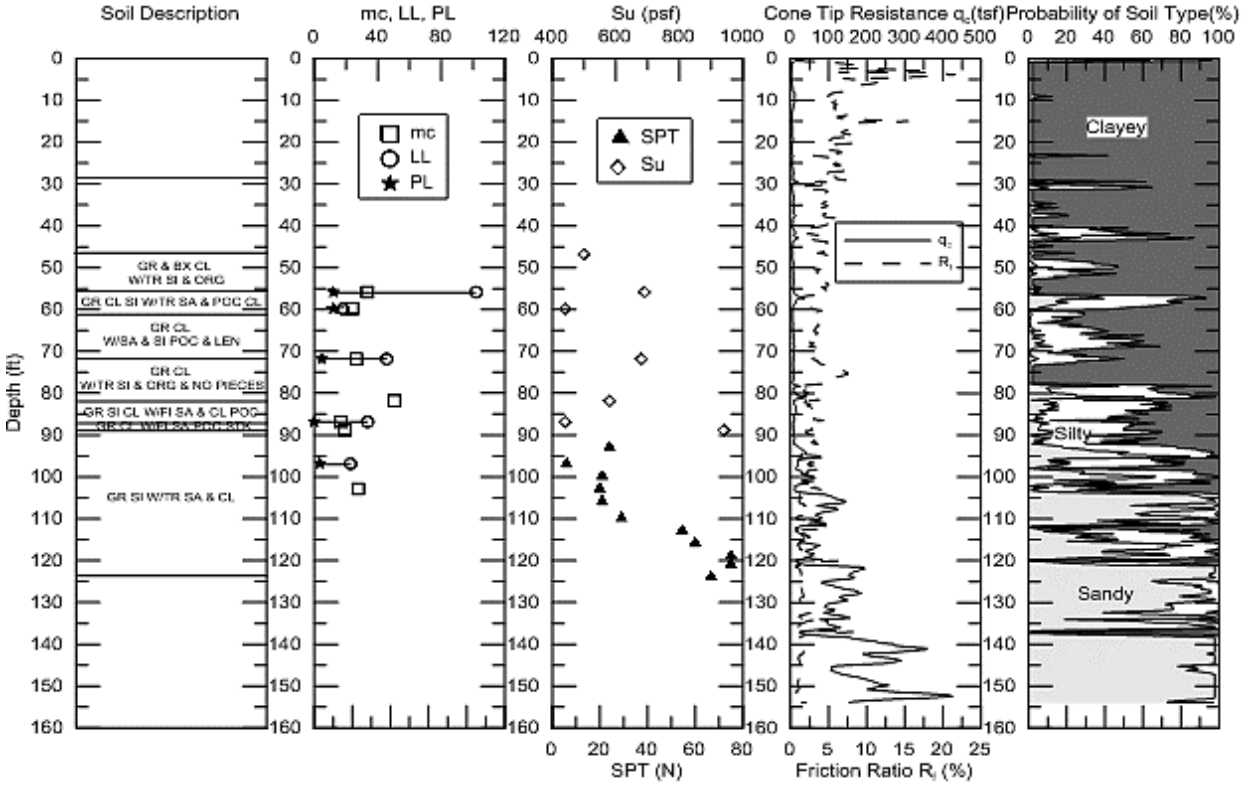


Figure 102. 33) 424-05-0087 TP#1

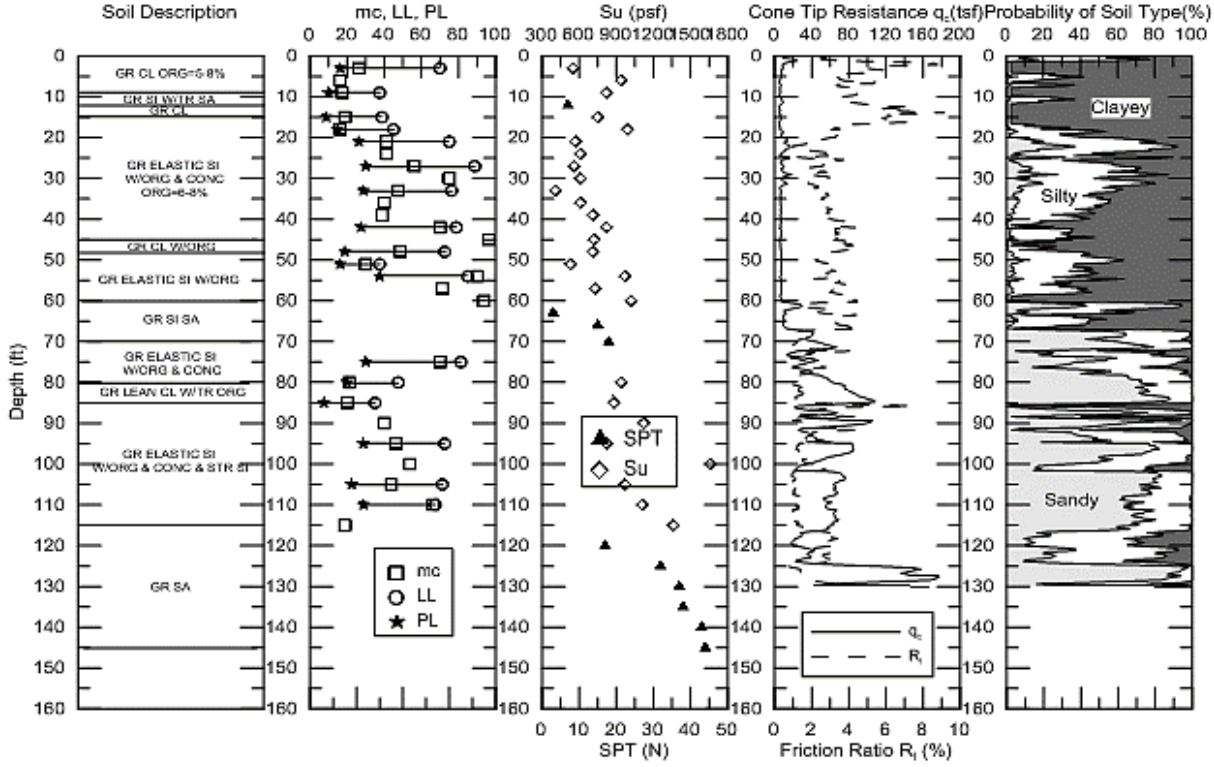


Figure 103. 34) 424-05-0087 TP#2

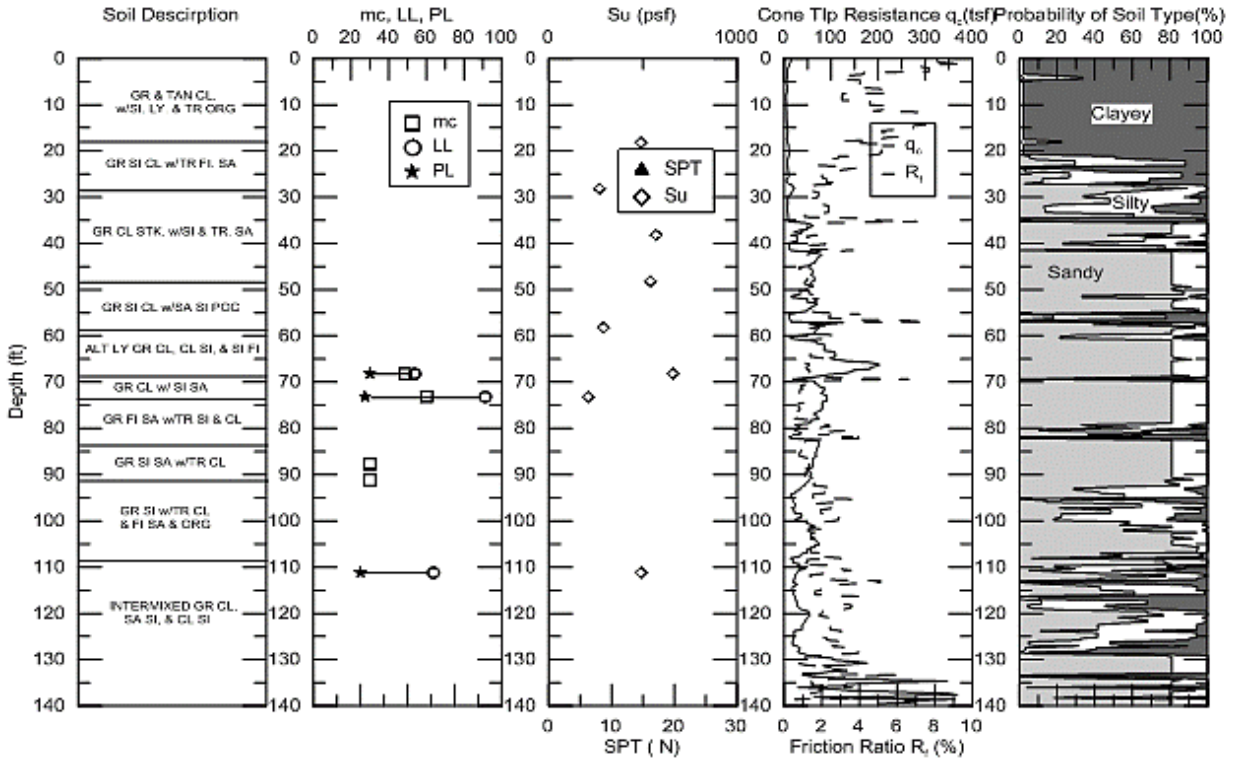


Figure 104. 35) 424-05-0087 TP#3

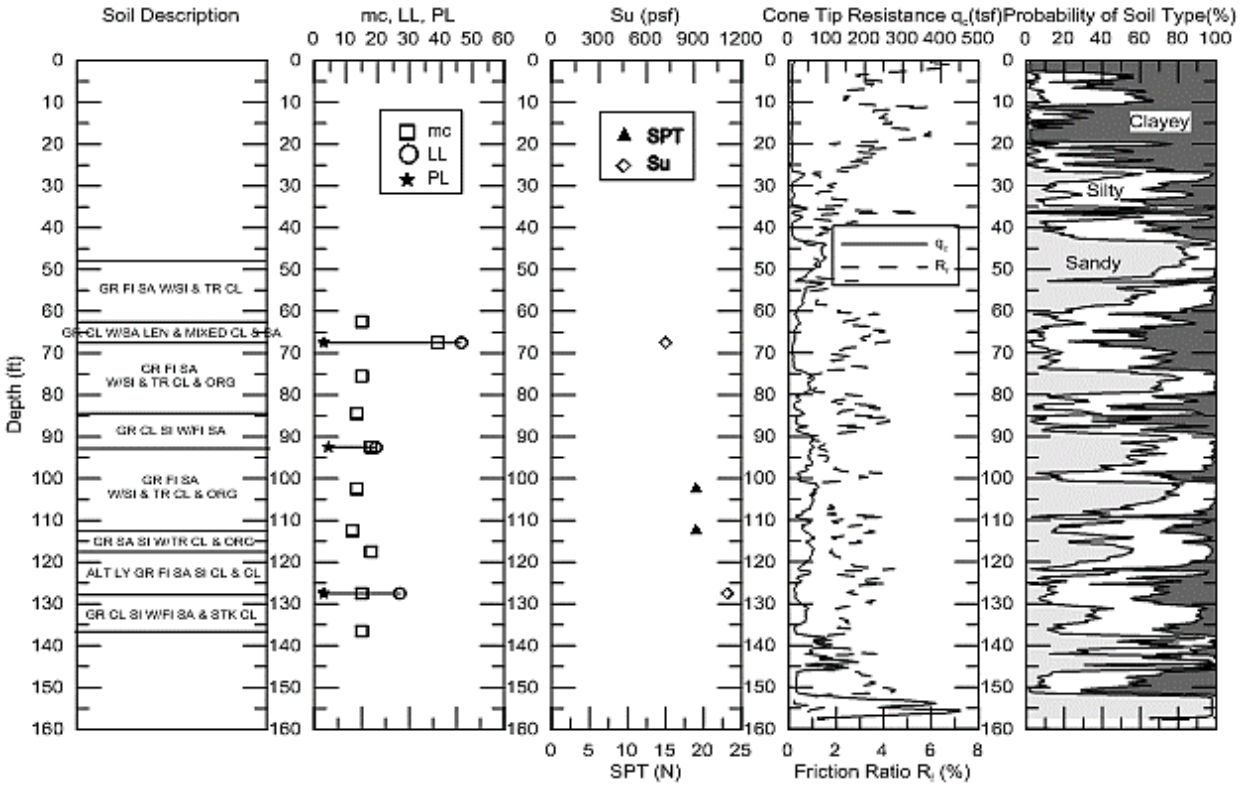


Figure 105. 36) 424-05-0087 TP#4

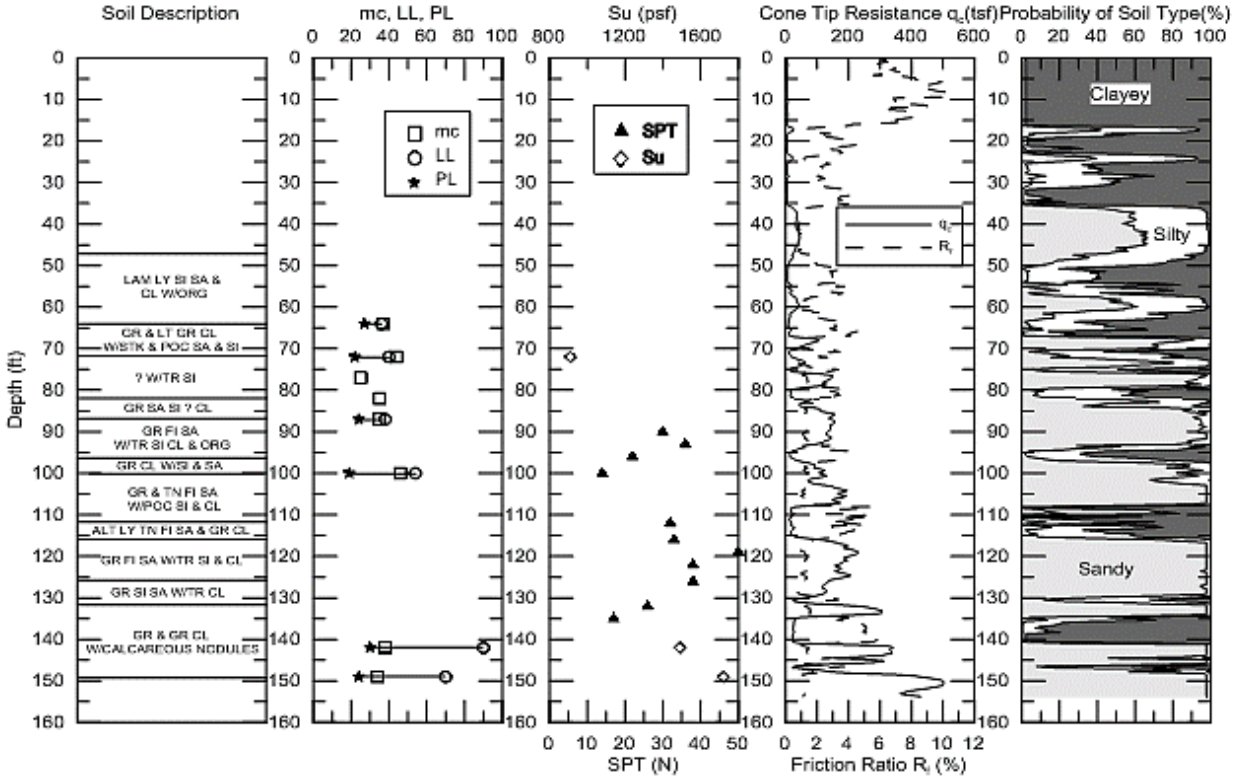


Figure 106. 37) 424-05-0087 TP#5

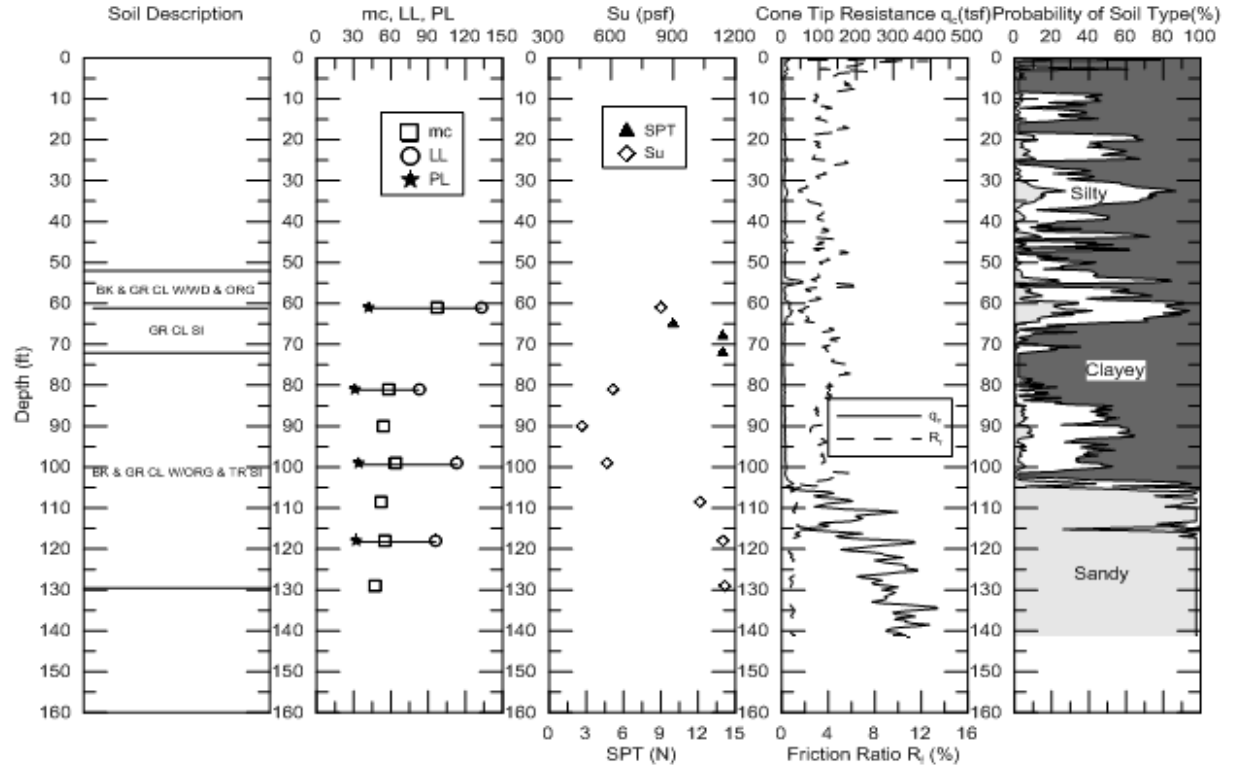


Figure 107. 38) 424-05-0087 TP#7

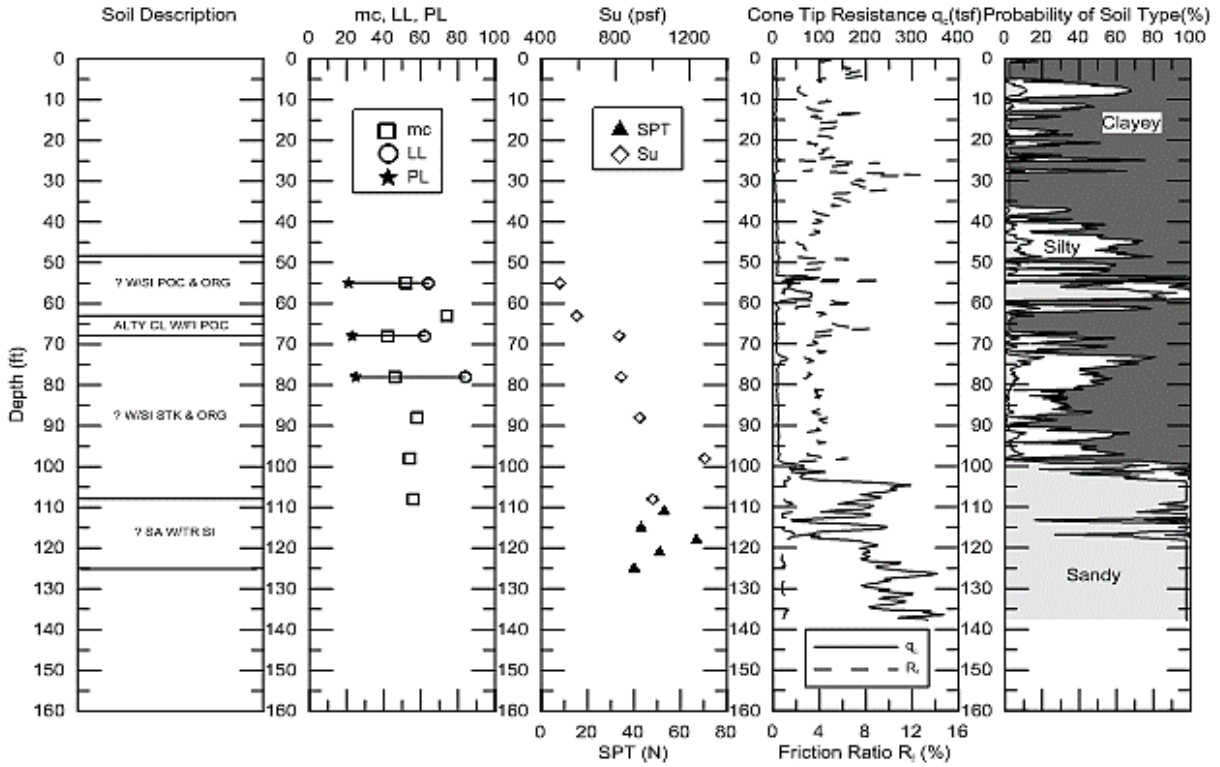


Figure 108. 39) 424-07-0008 TP#1

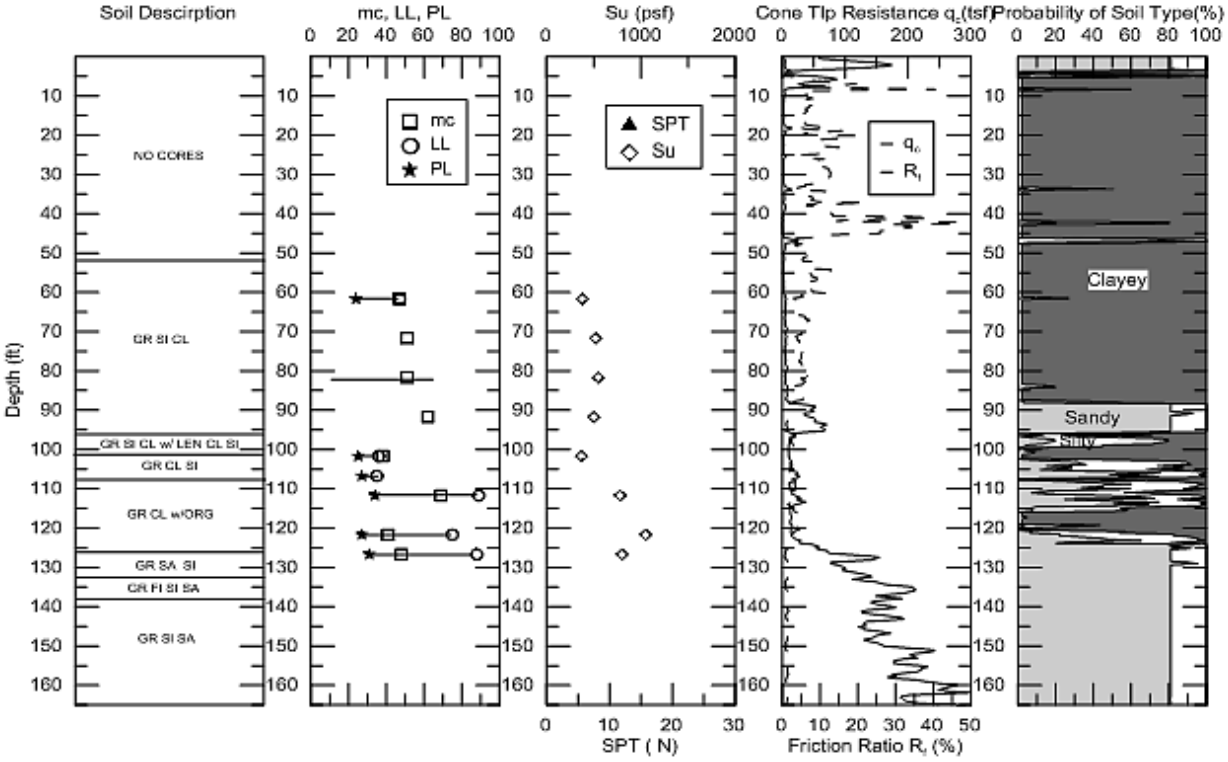


Figure 109. 40) 424-07-0008 TP#3

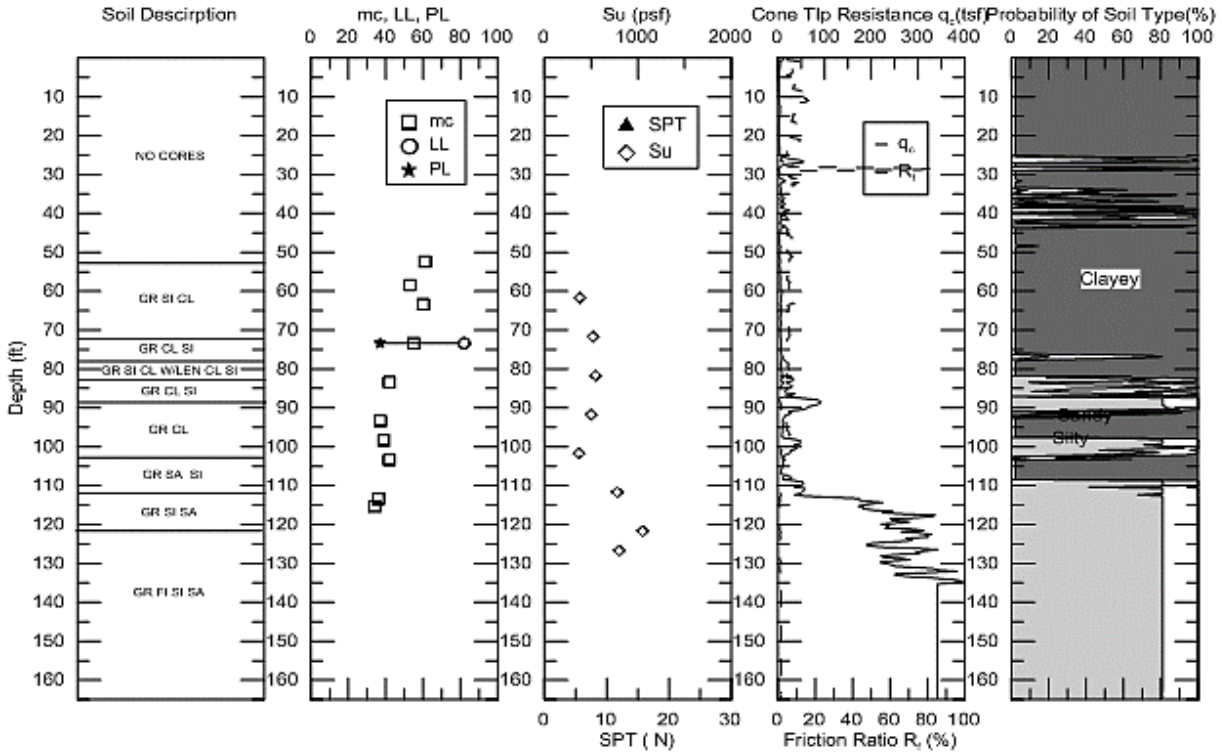


Figure 110. 41) 424-07-0009 TP#3

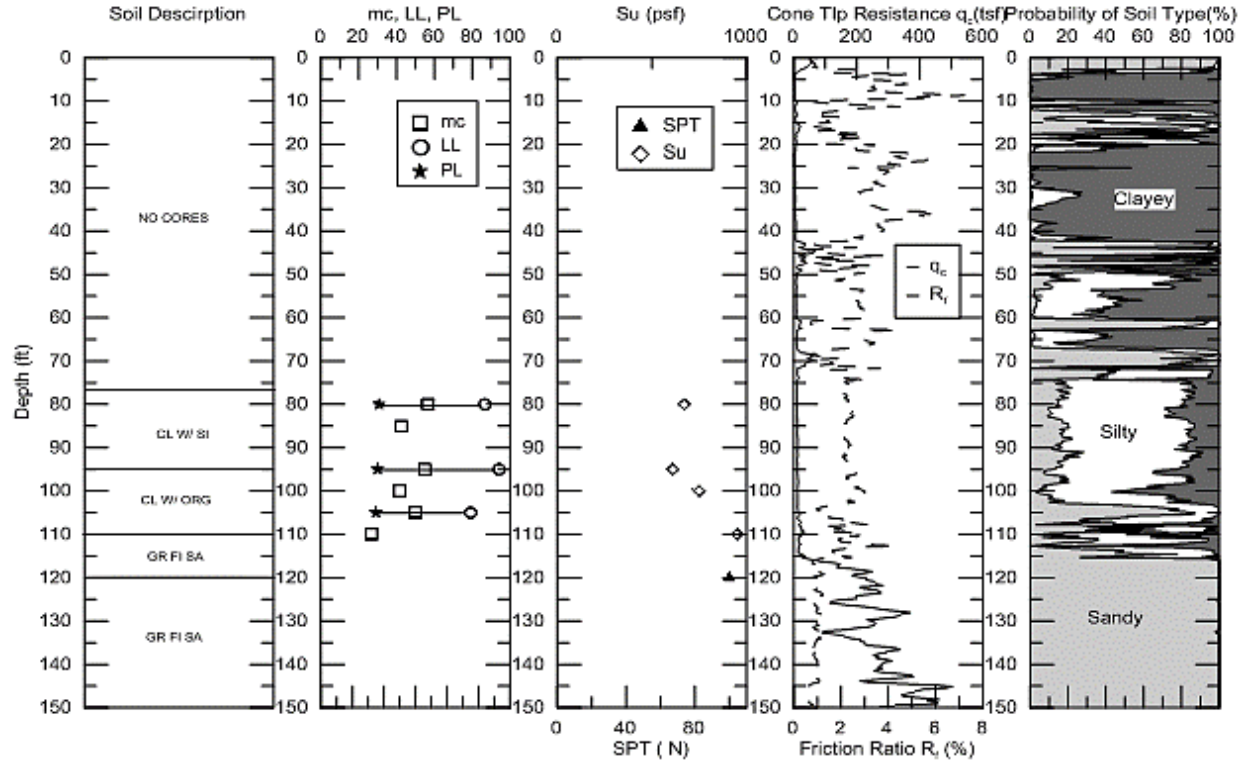


Figure 111. 42) 424-07-0009 TP#4

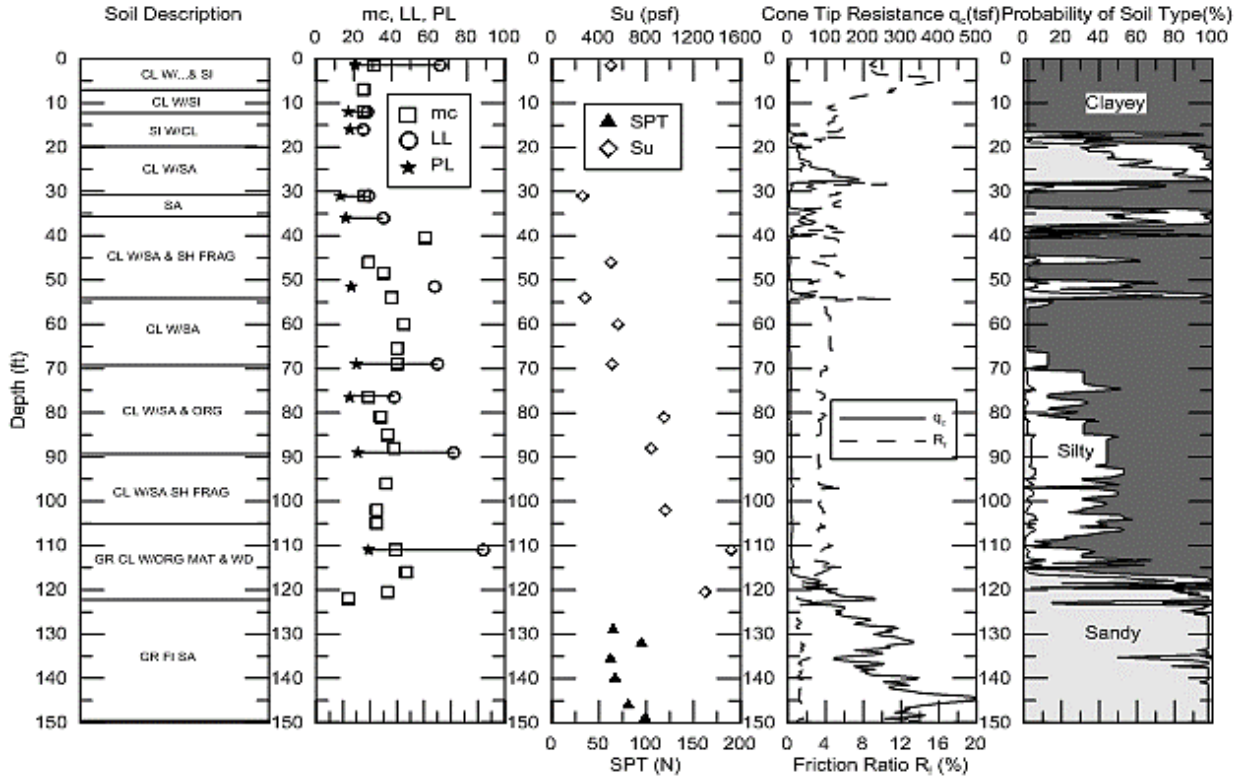


Figure 112. 43) 424-07-0009 TP#4A

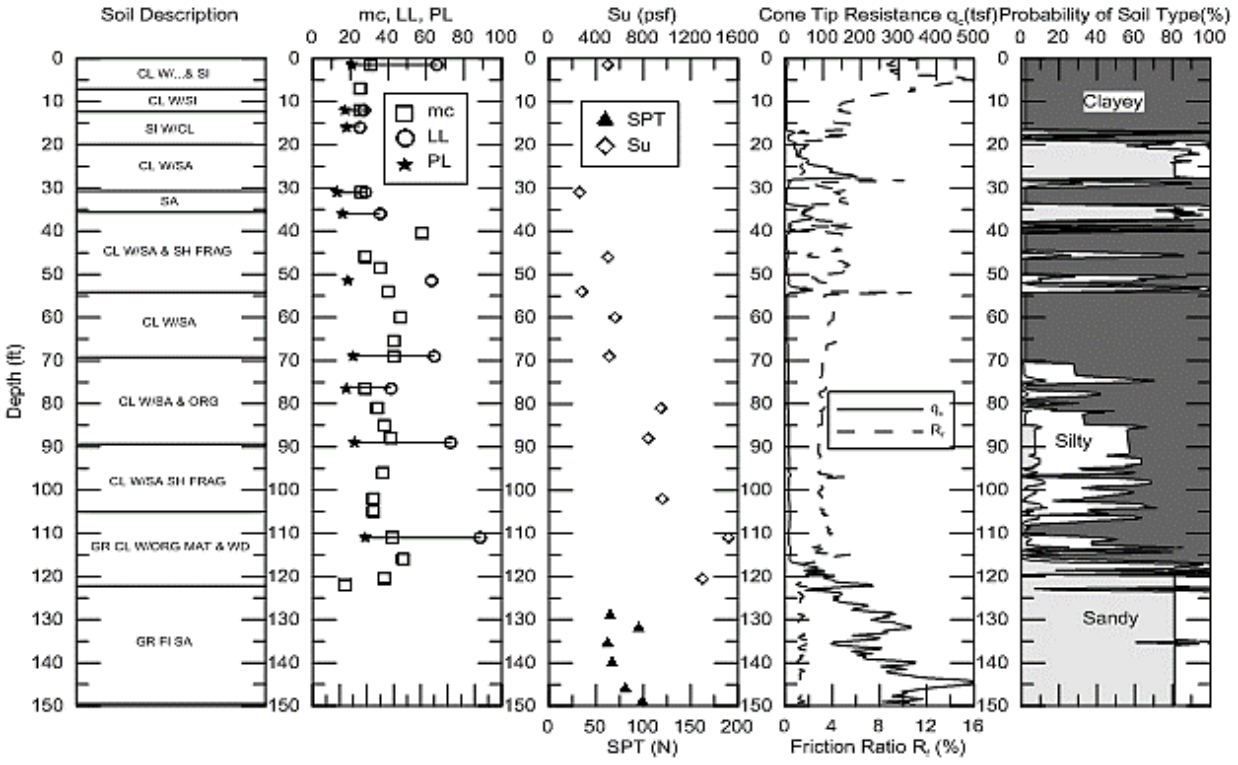


Figure 113. 44) 424-06-0005 TP#1

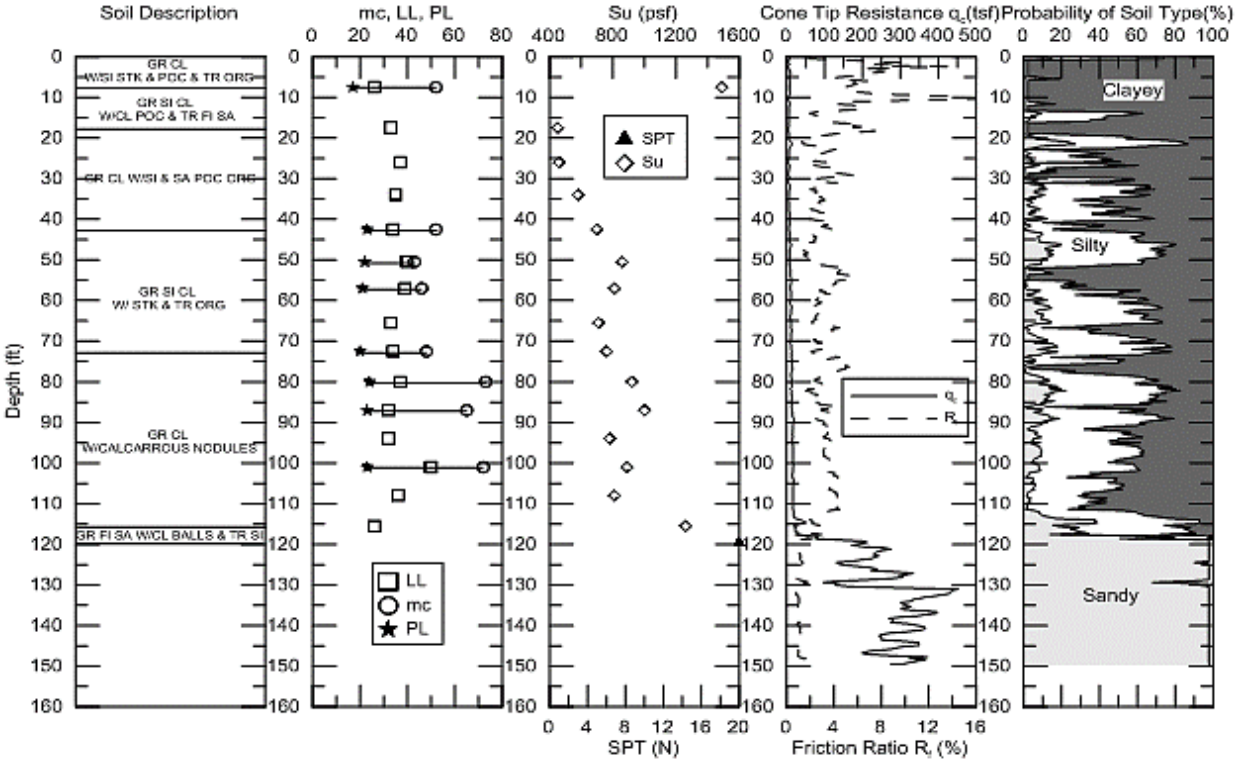


Figure 114. 45) 424-06-0005 TP#2

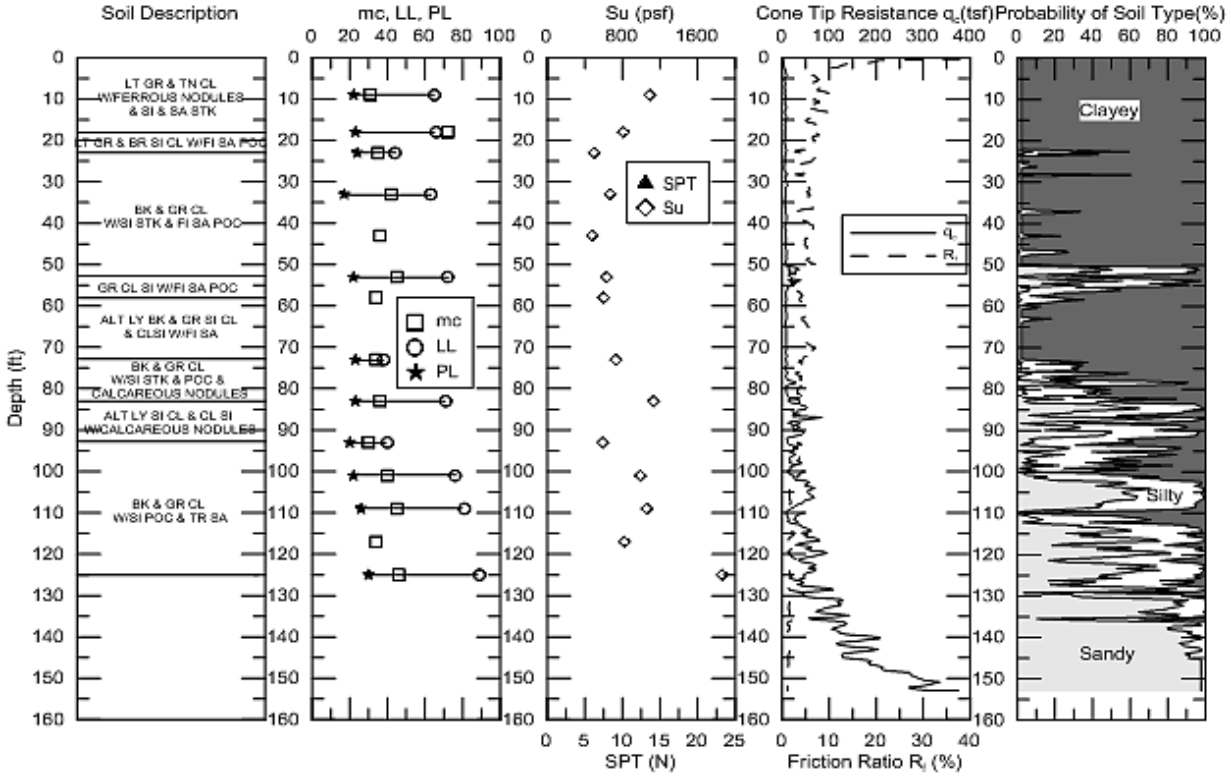


Figure 115. 46) 424-06-0005 TP#3

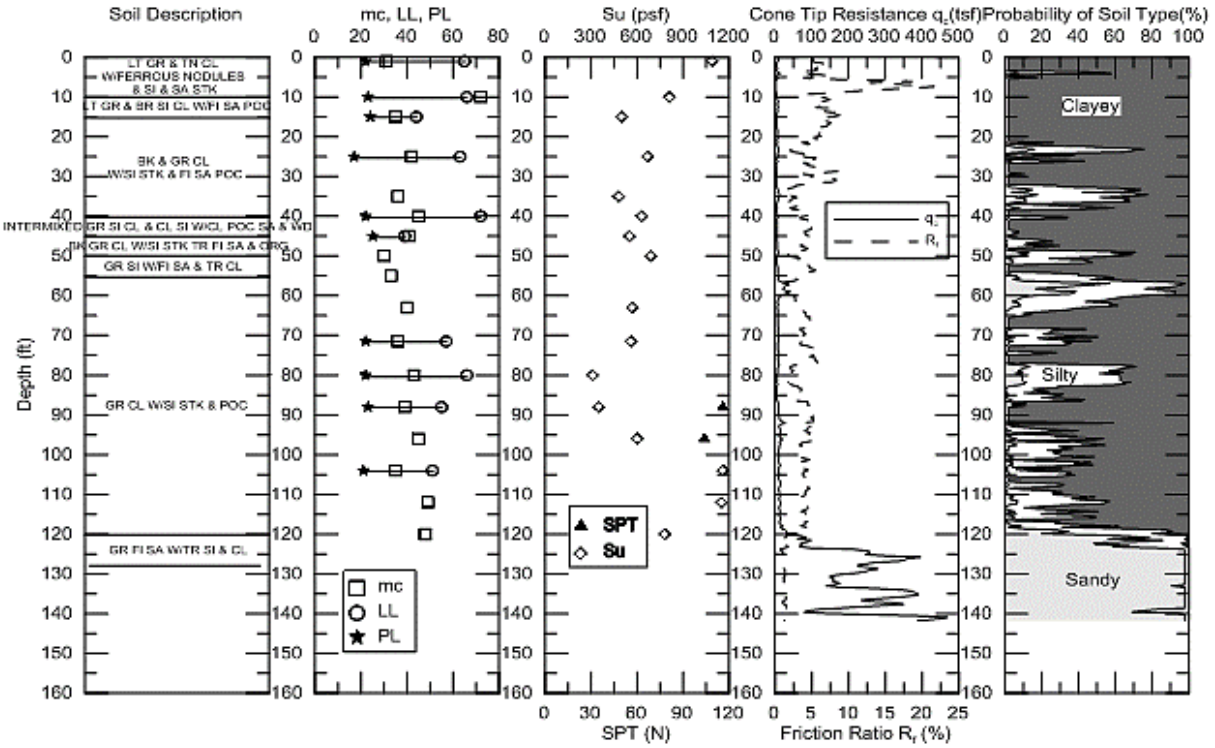


Figure 116. 47) 424-06-0005 TP#4

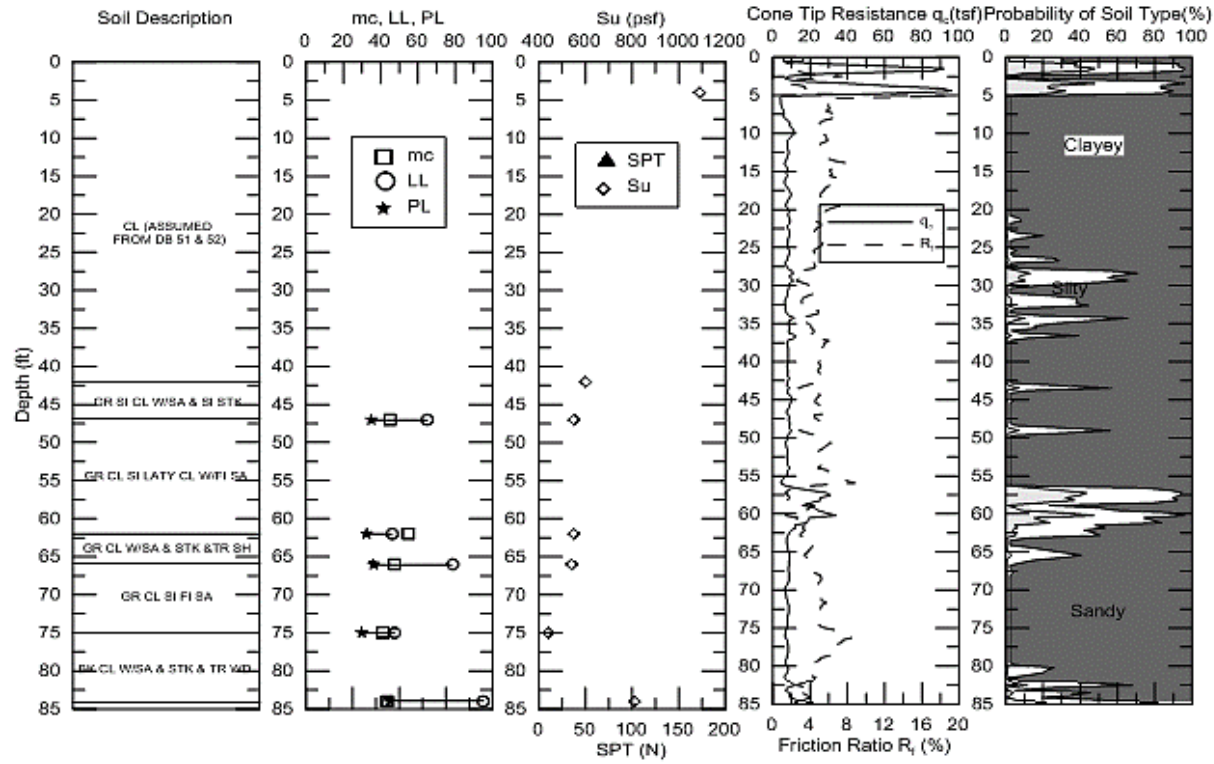


Figure 117. 48) 424-06-0005 TP#5

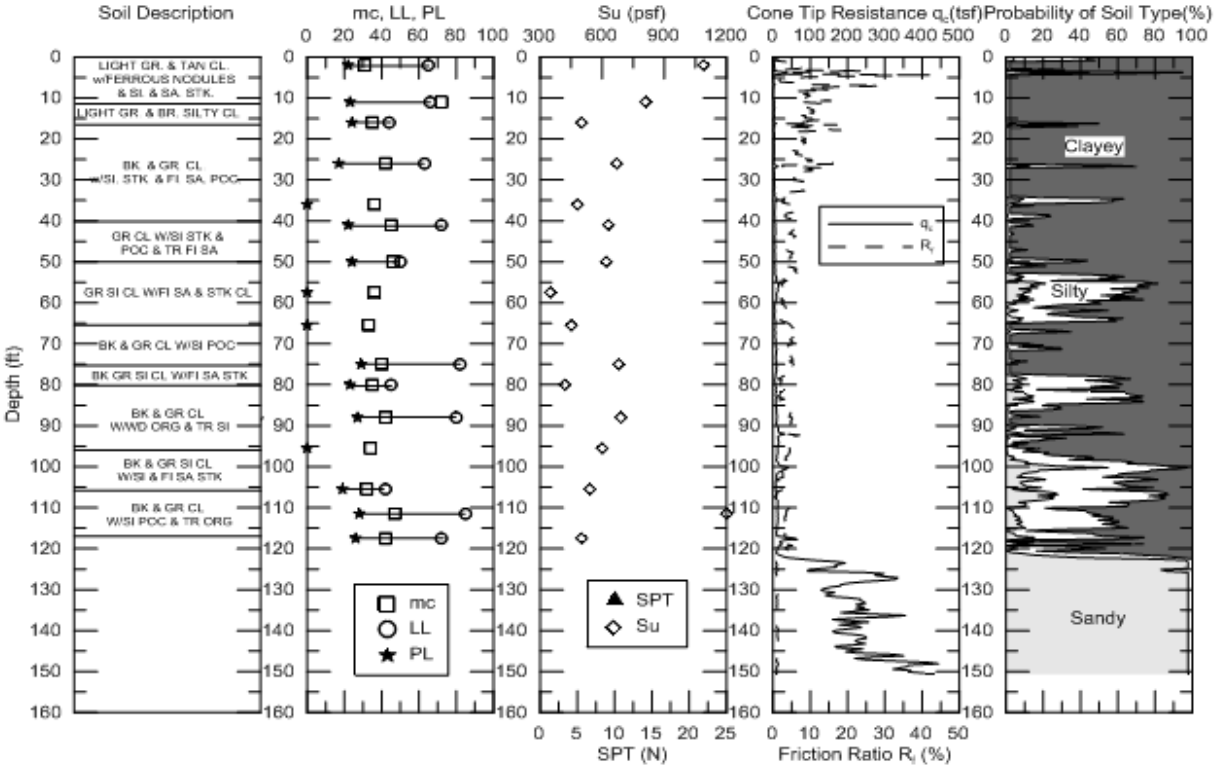


Figure 118. 49) 424-07-0021 TP#1

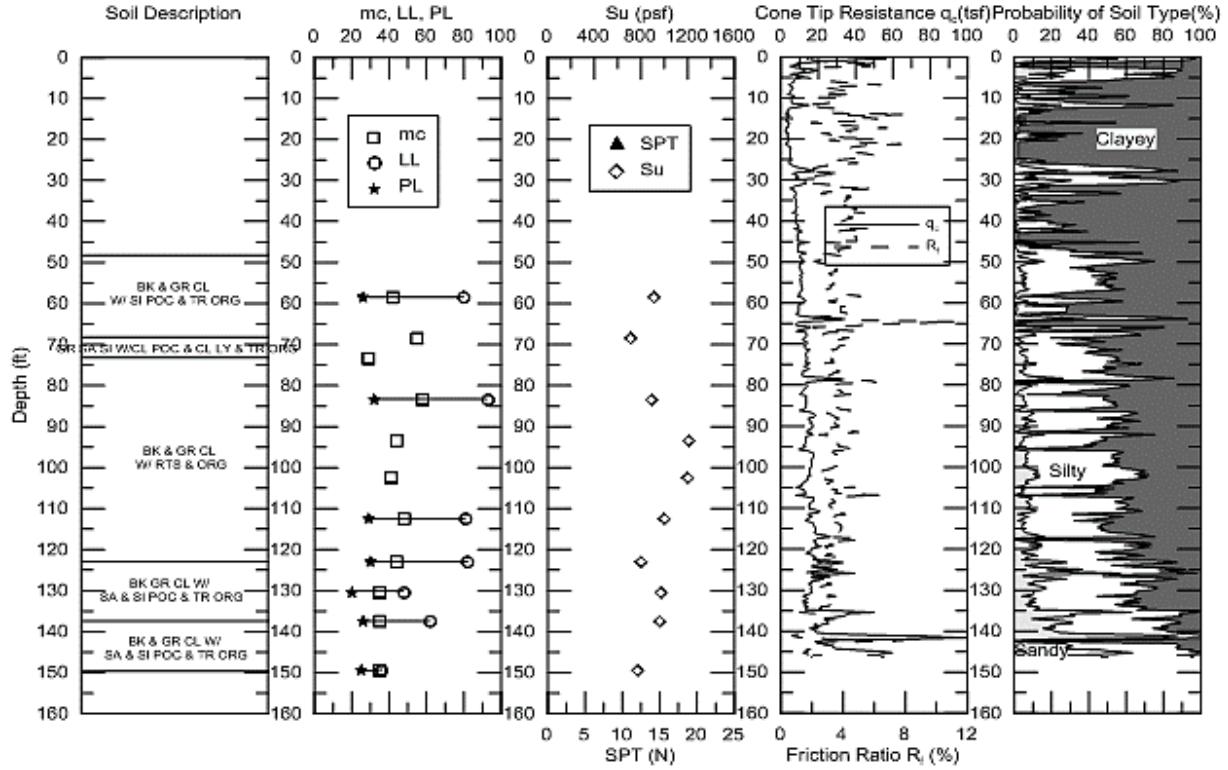


Figure 119. 50) 434-01-0002 TP#3

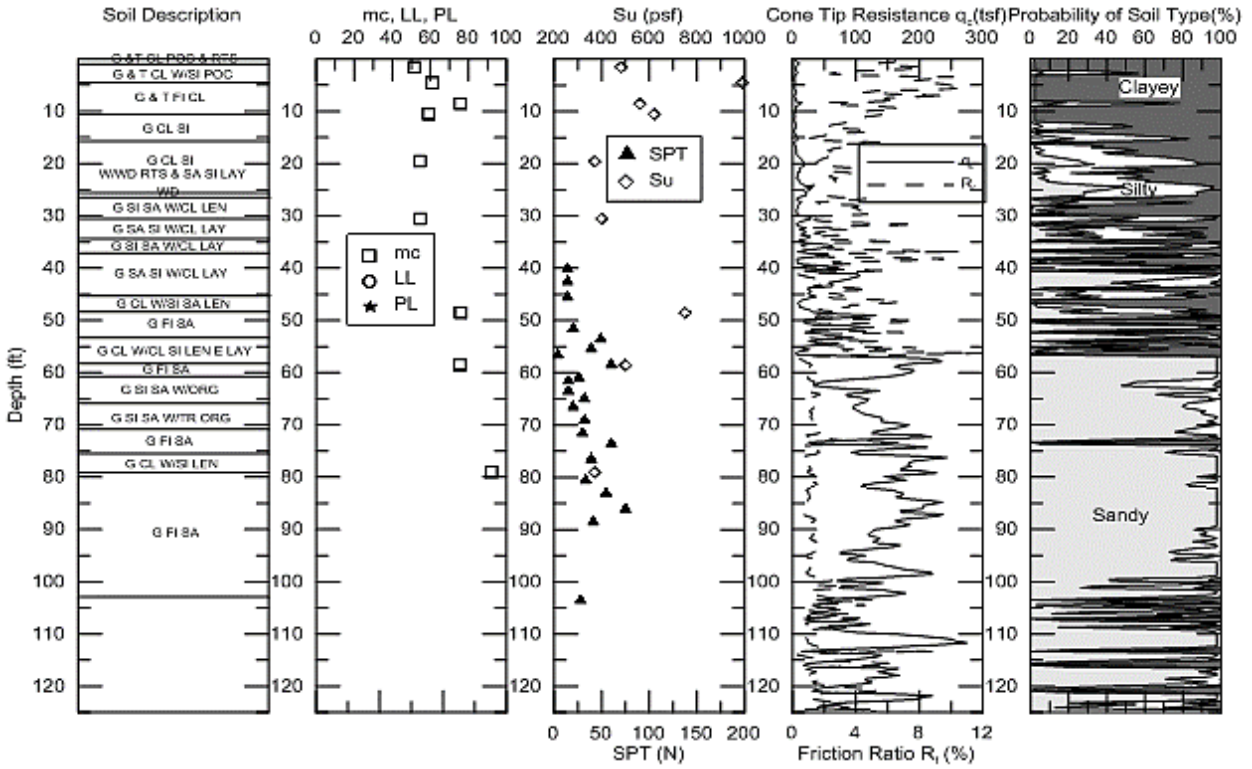


Figure 120. 51) 450-15-0085 TP#3A

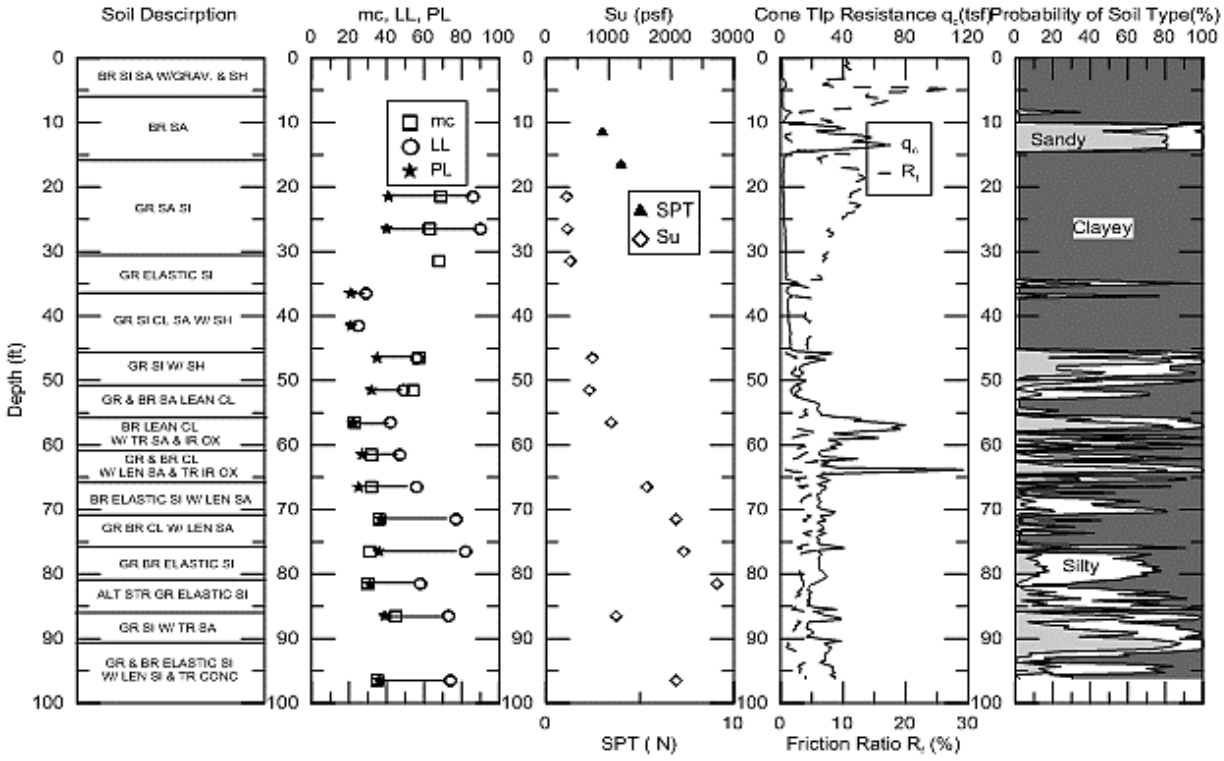


Figure 121. 52) 450-15-0100 TP#1

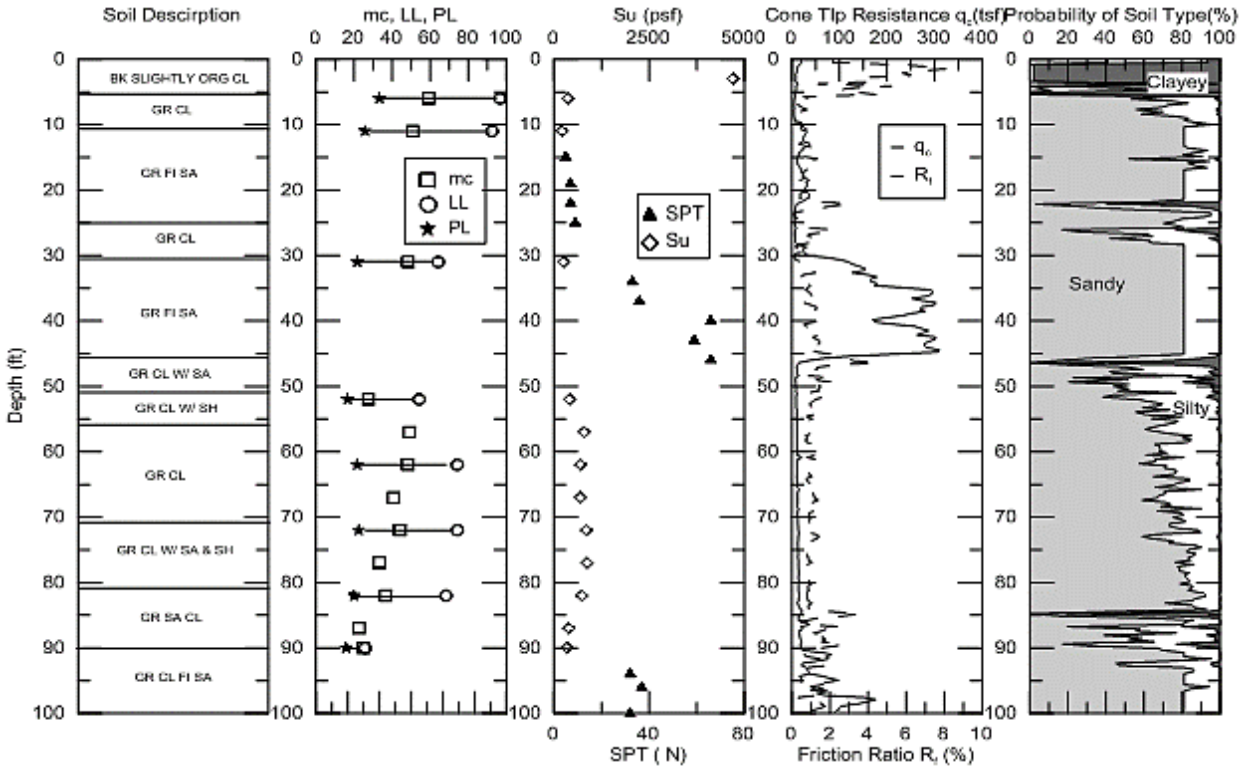


Figure 122. 53) 450-15-0100 TP#2

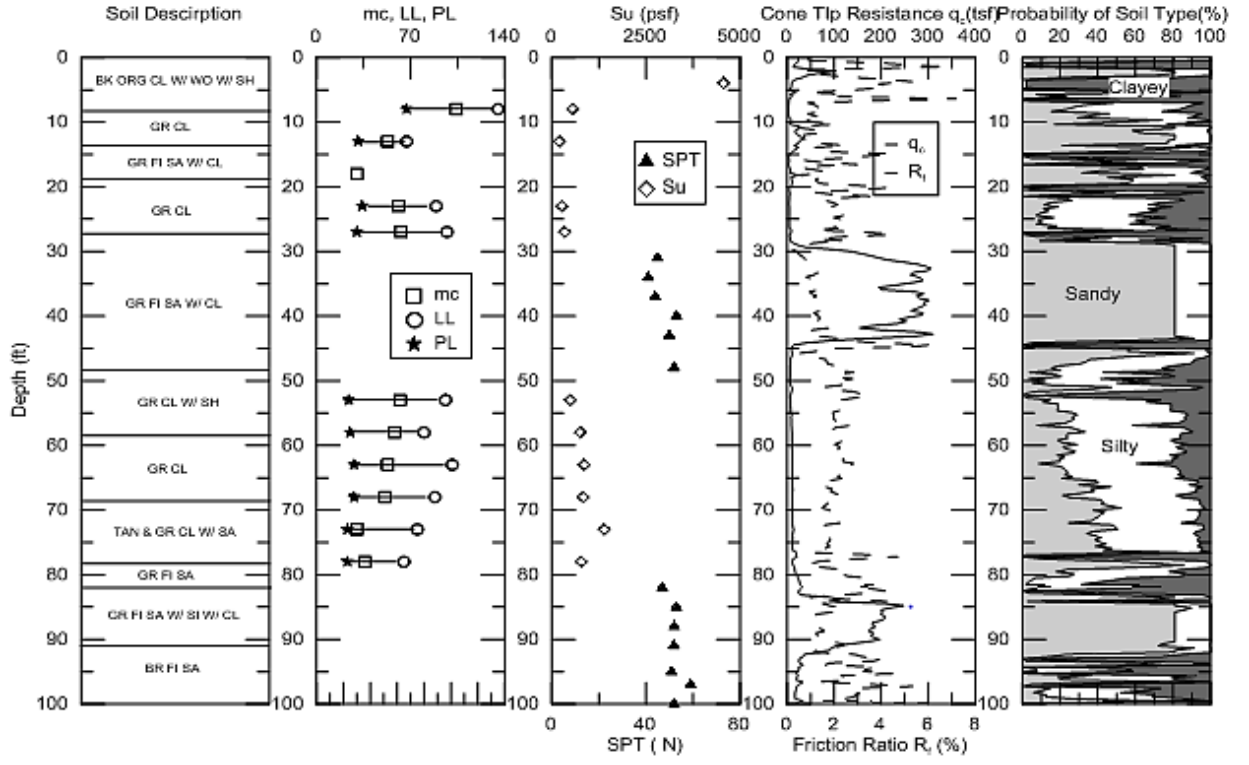


Figure 123. 54) 450-15-0100 TP#3

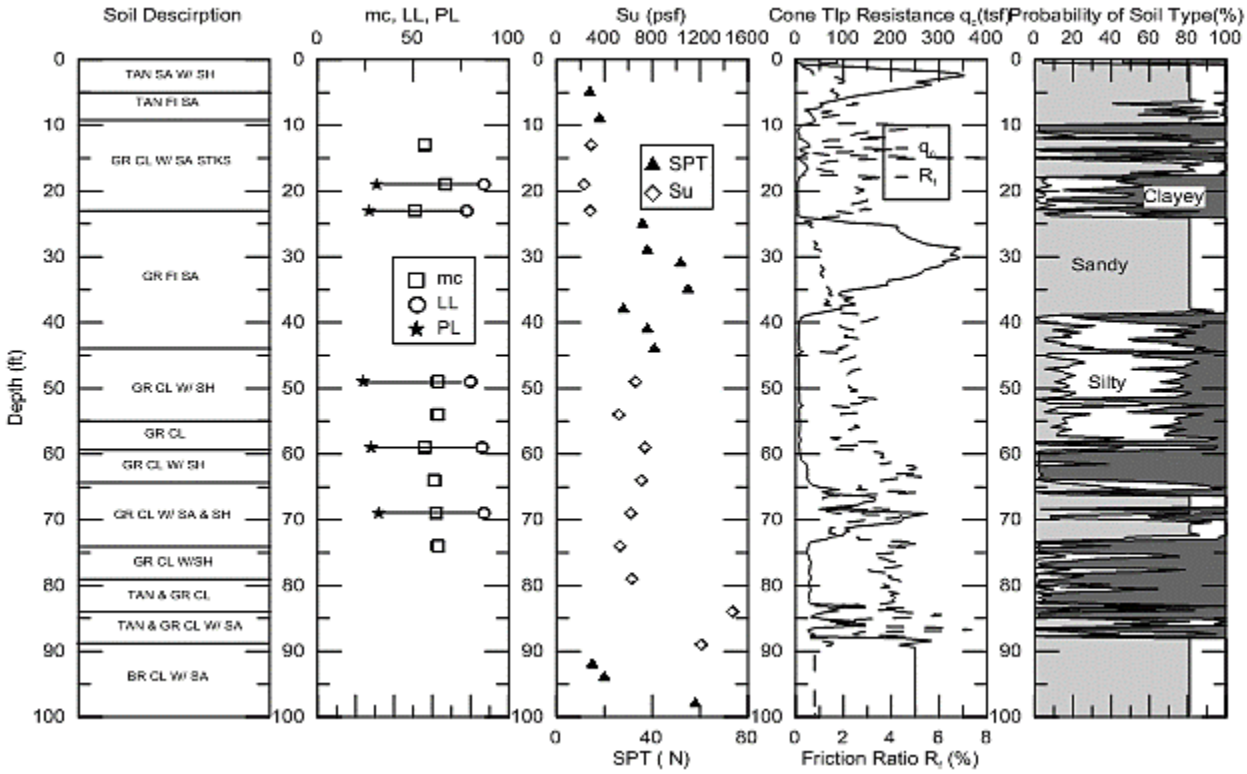


Figure 124. 55) 450-15-0103 TP#1

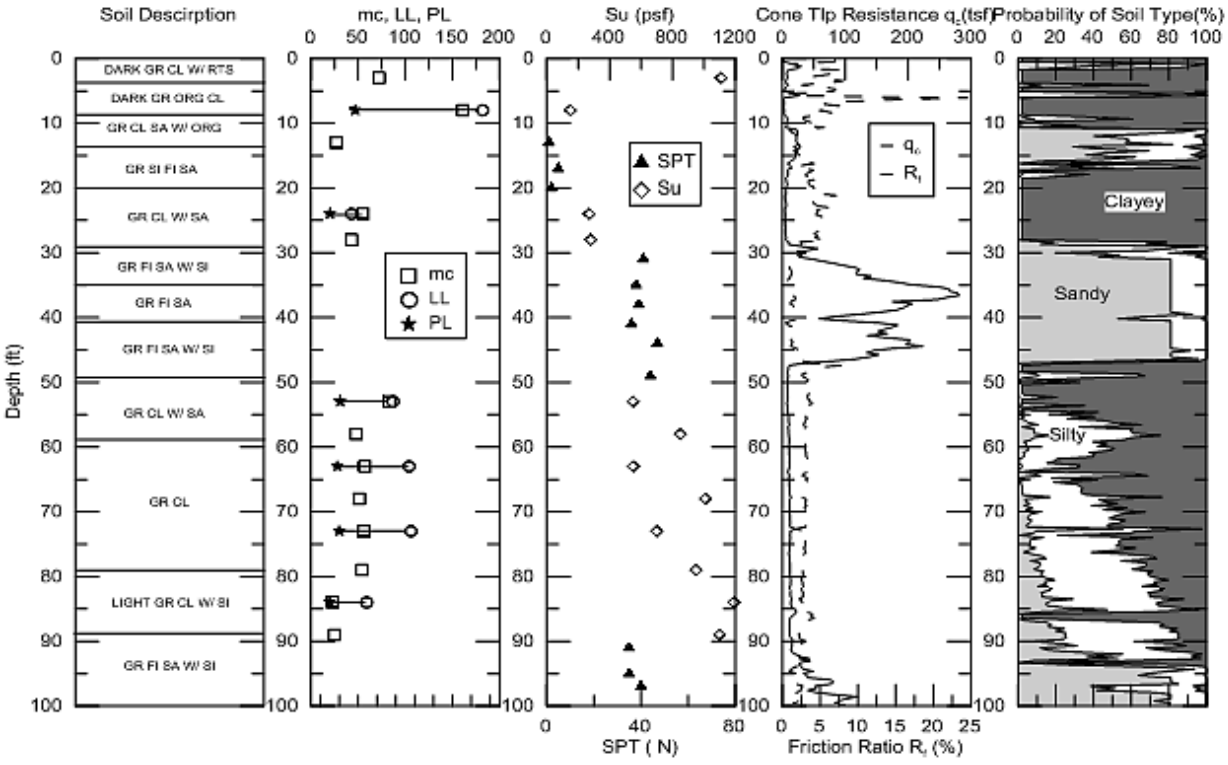


Figure 125. 56) 450-15-0103 TP#2

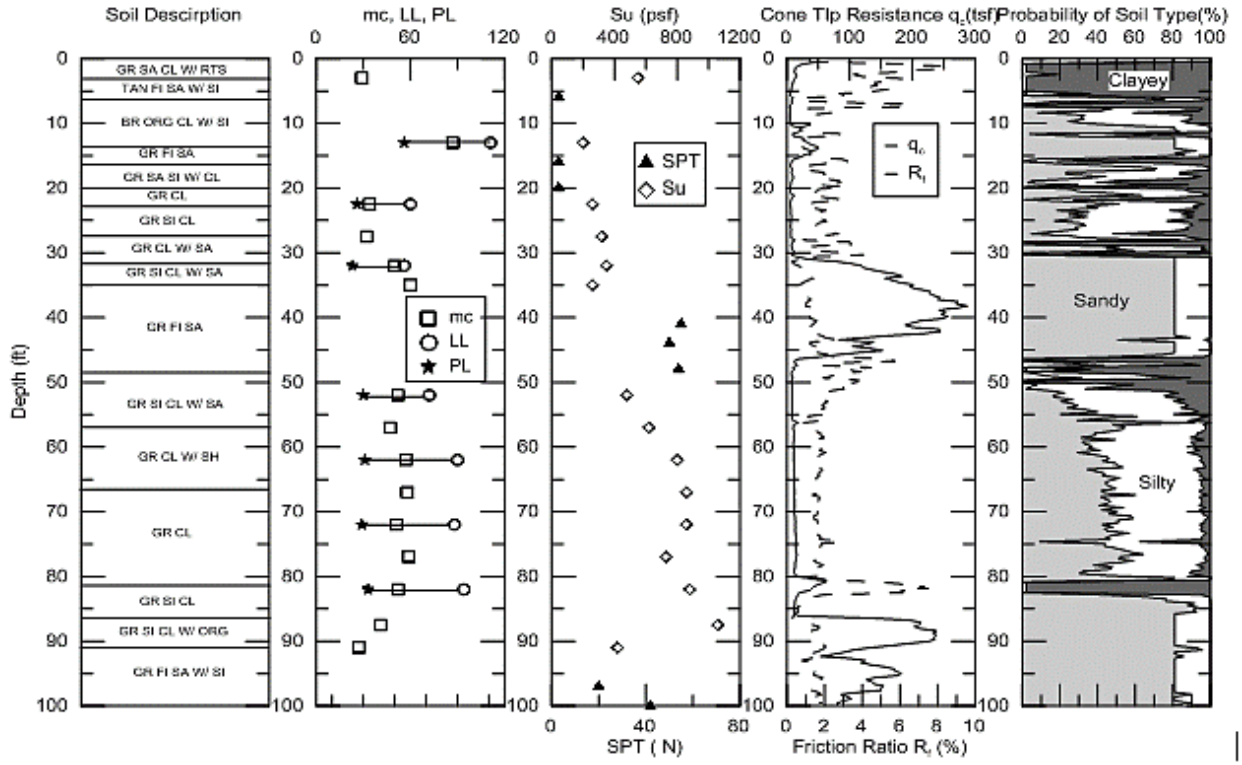


Figure 126. 57) 450-15-0103 TP#5

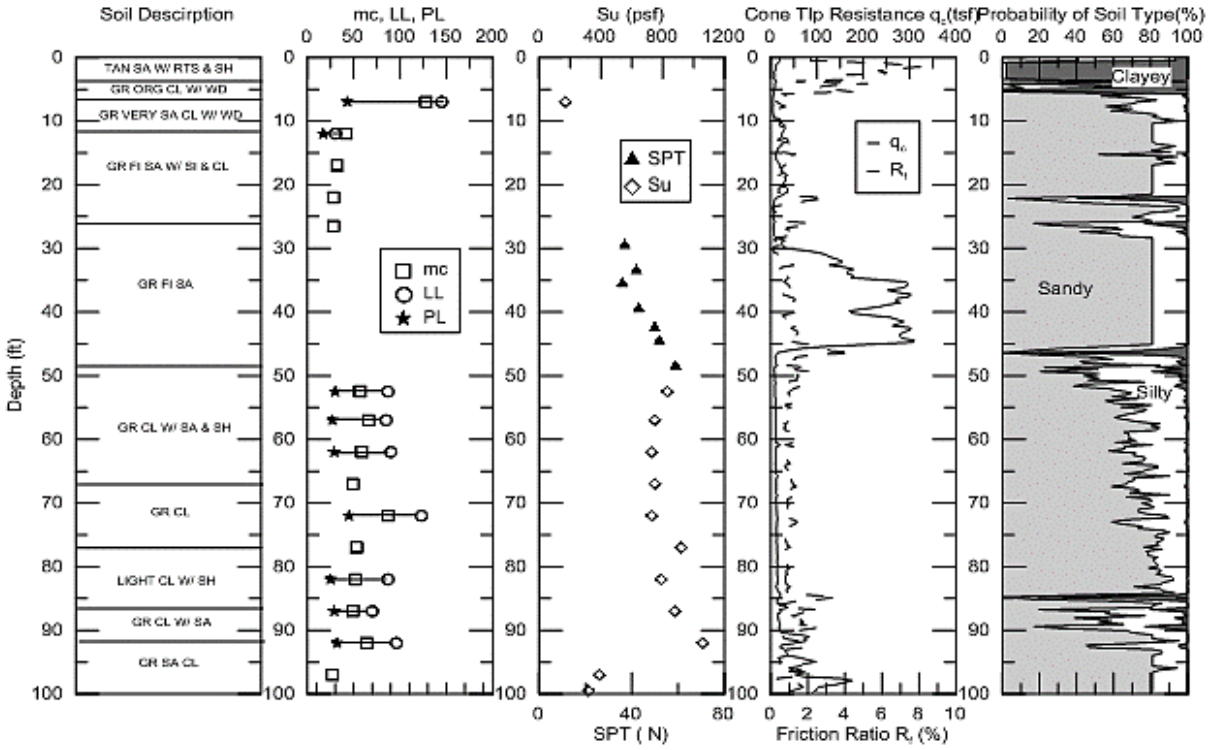


Figure 127. 58) 450-15-0103 TP#7

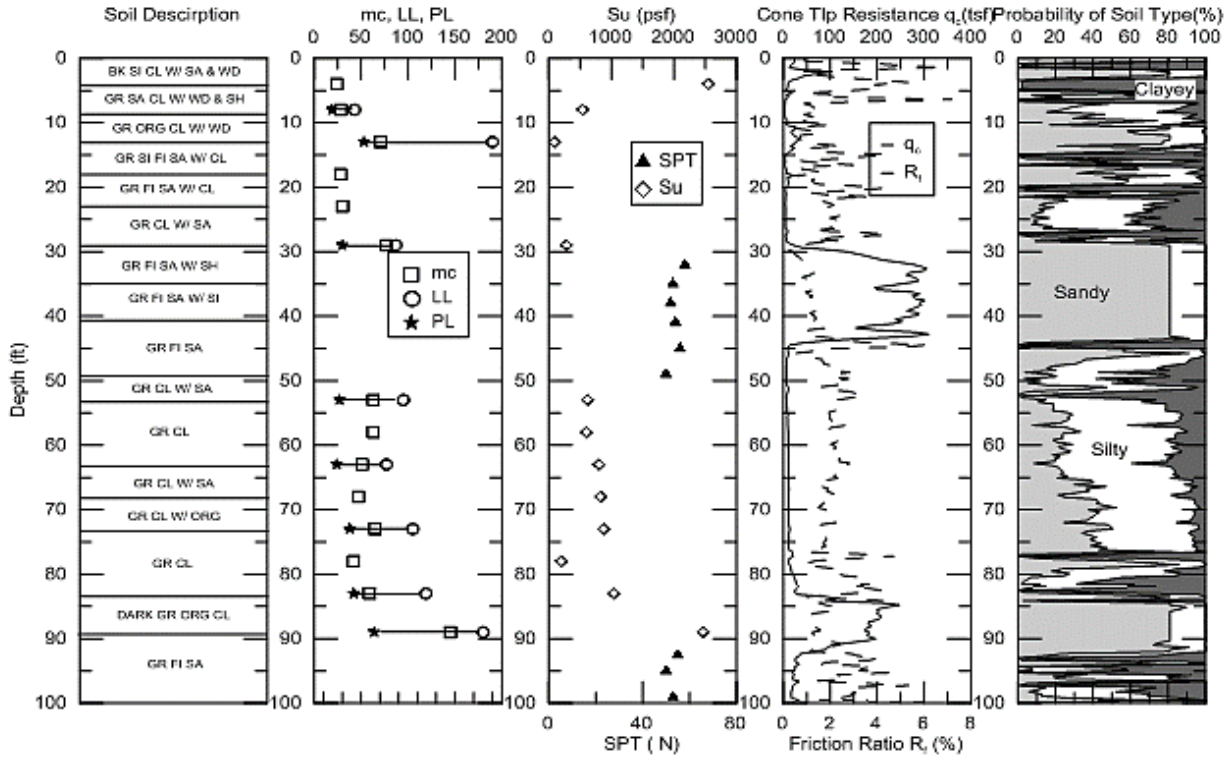


Figure 128. 59) 450-36-0002 TP#8

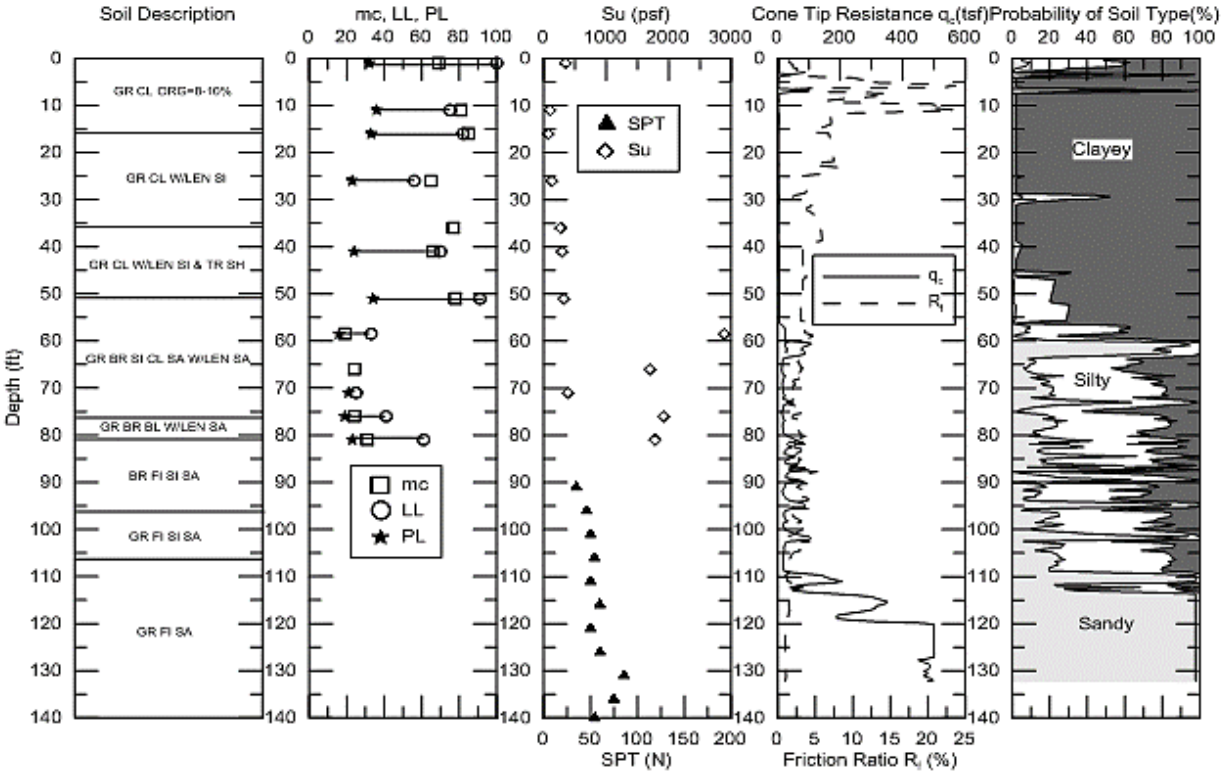


Figure 129. 60) 455-05-0036 TP#1

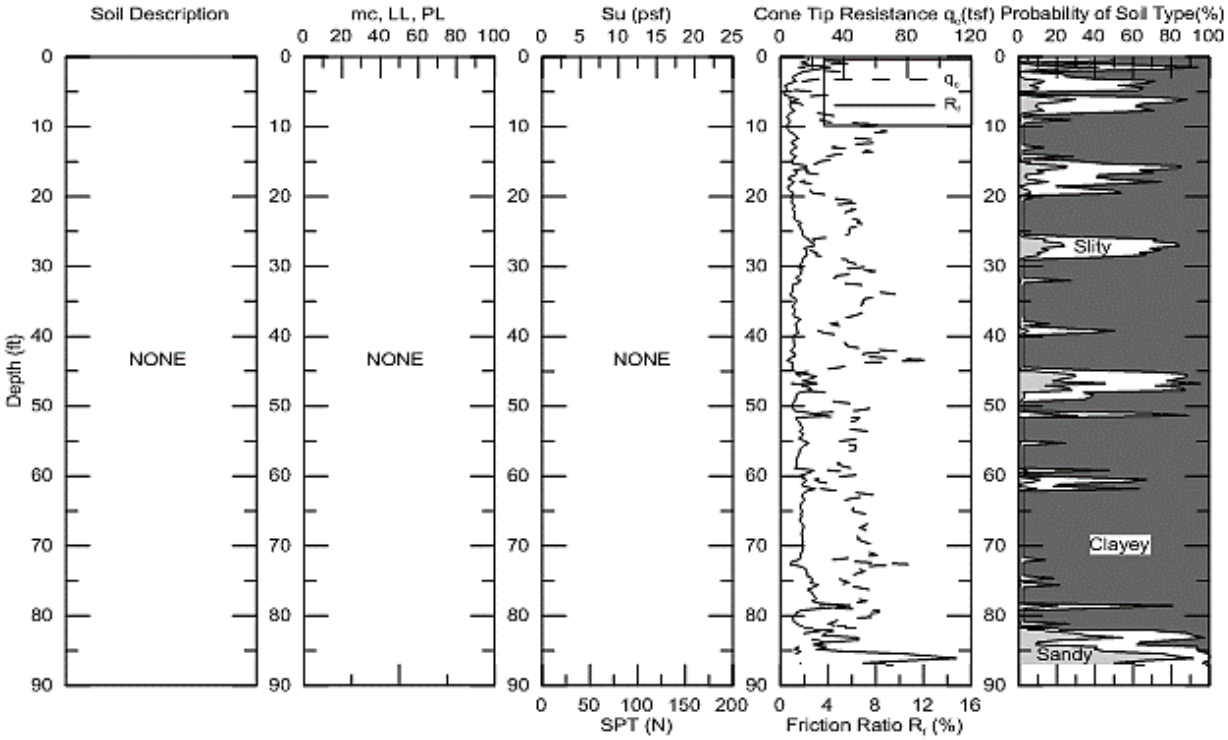


Figure 130. 61) 455-05-0036 TP#2

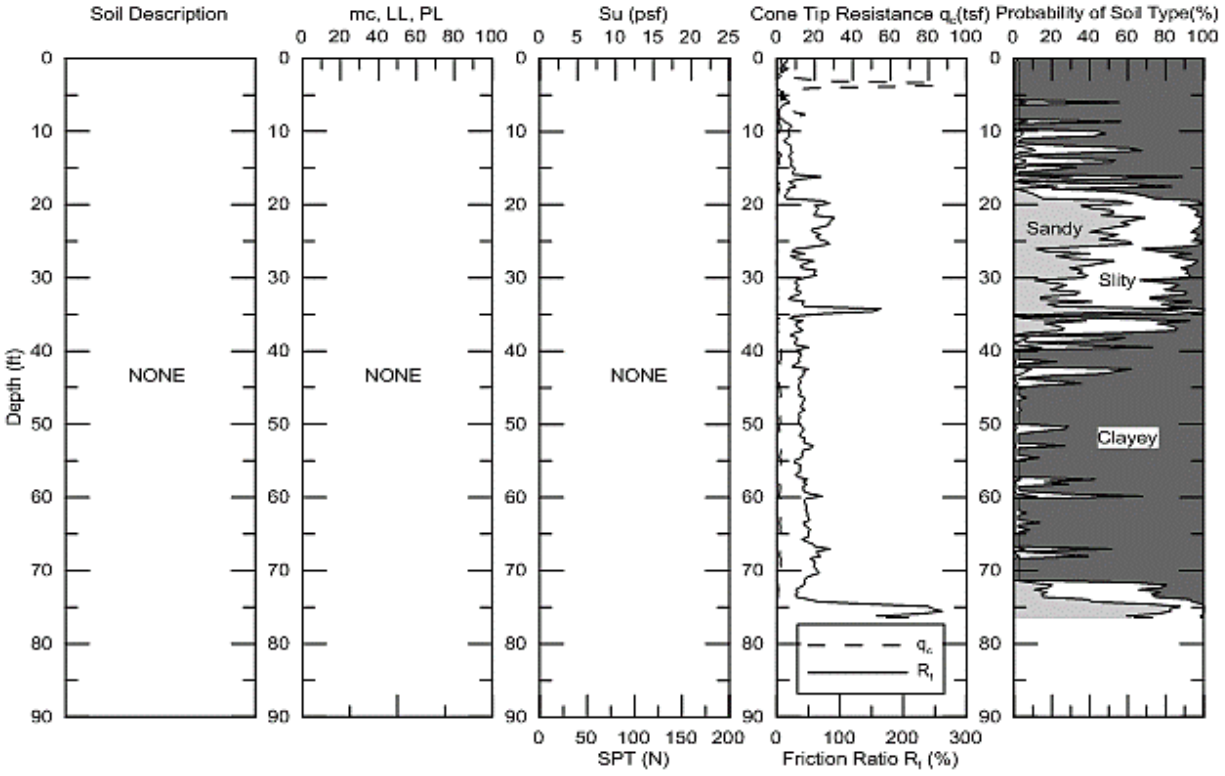


Figure 131. 62) 455-05-0036 TP#3

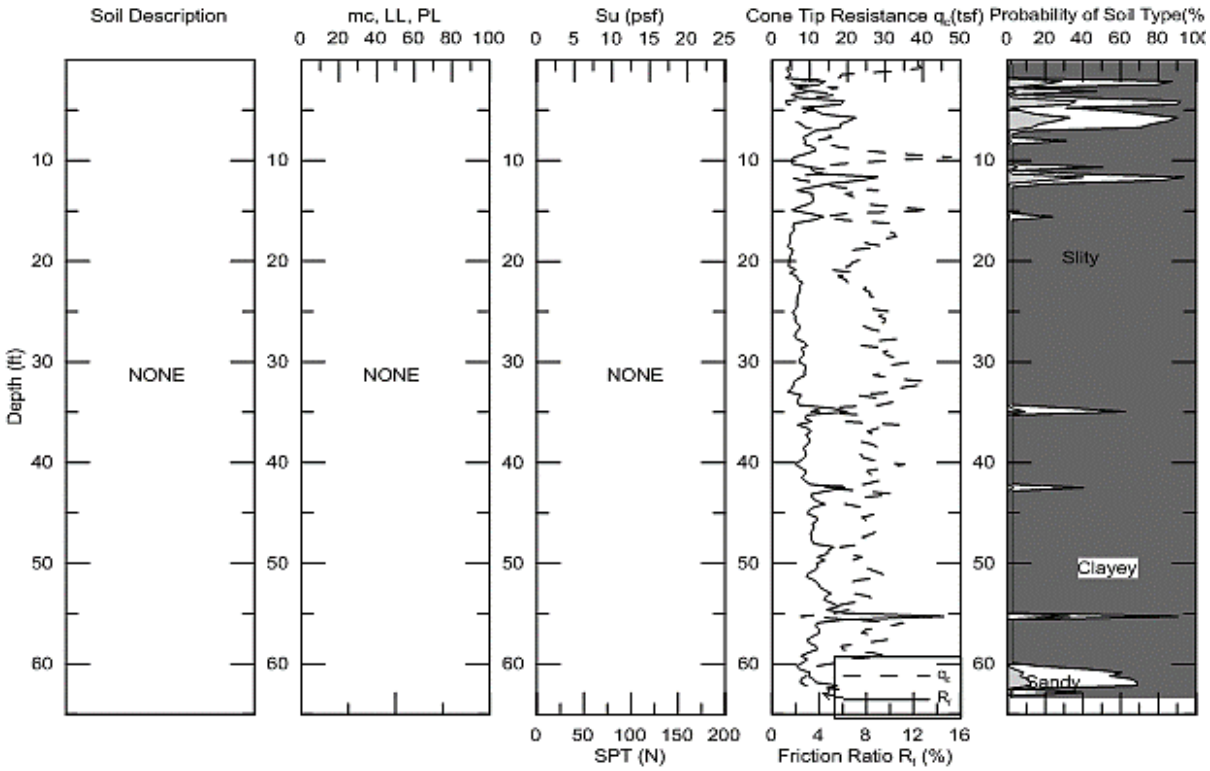


Figure 132 63) 713-48-0083 TP#1

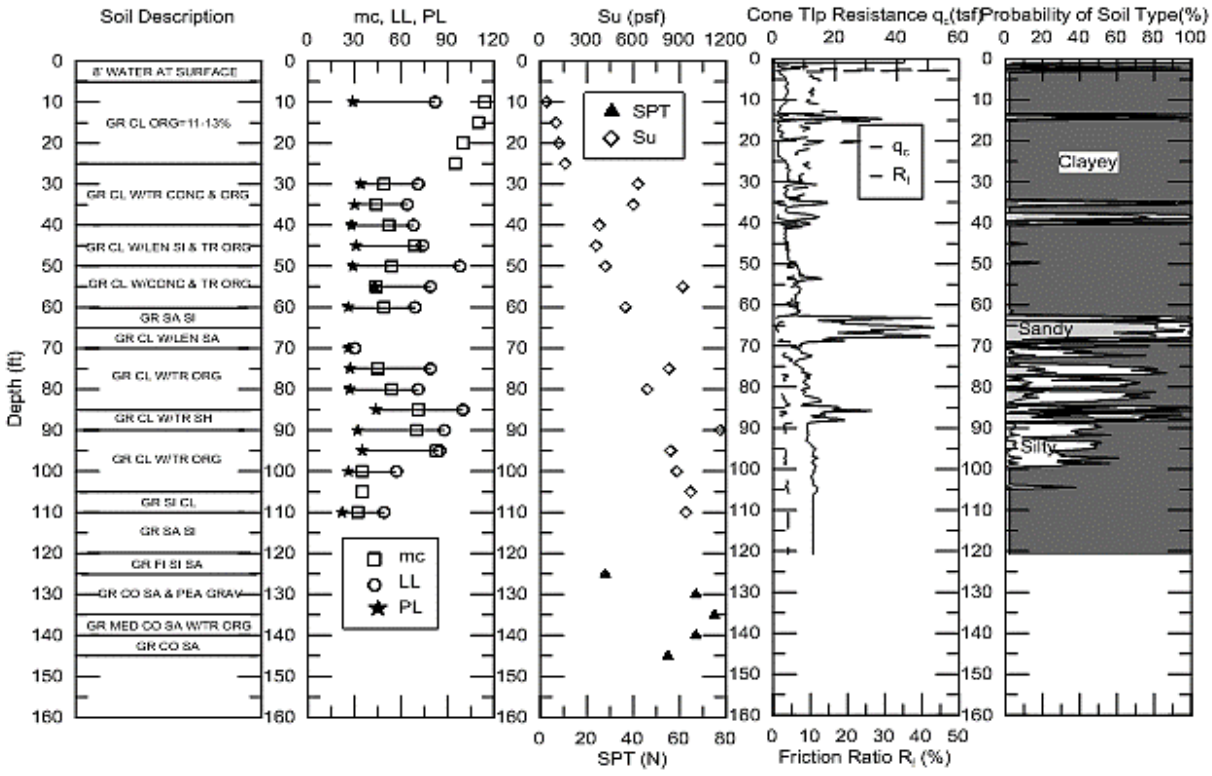


Figure 133. 64) 713-48-0083 TP#2

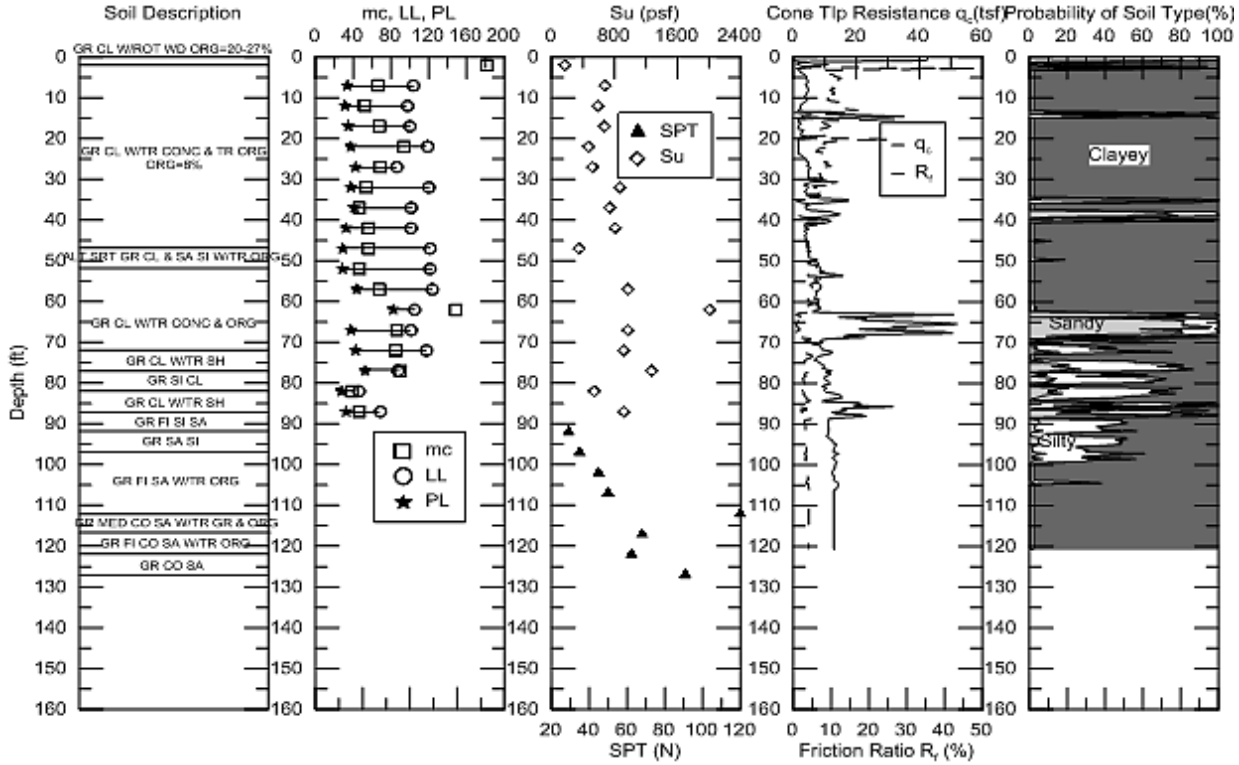


Figure 134. 65) 742-06-0073 TP#2

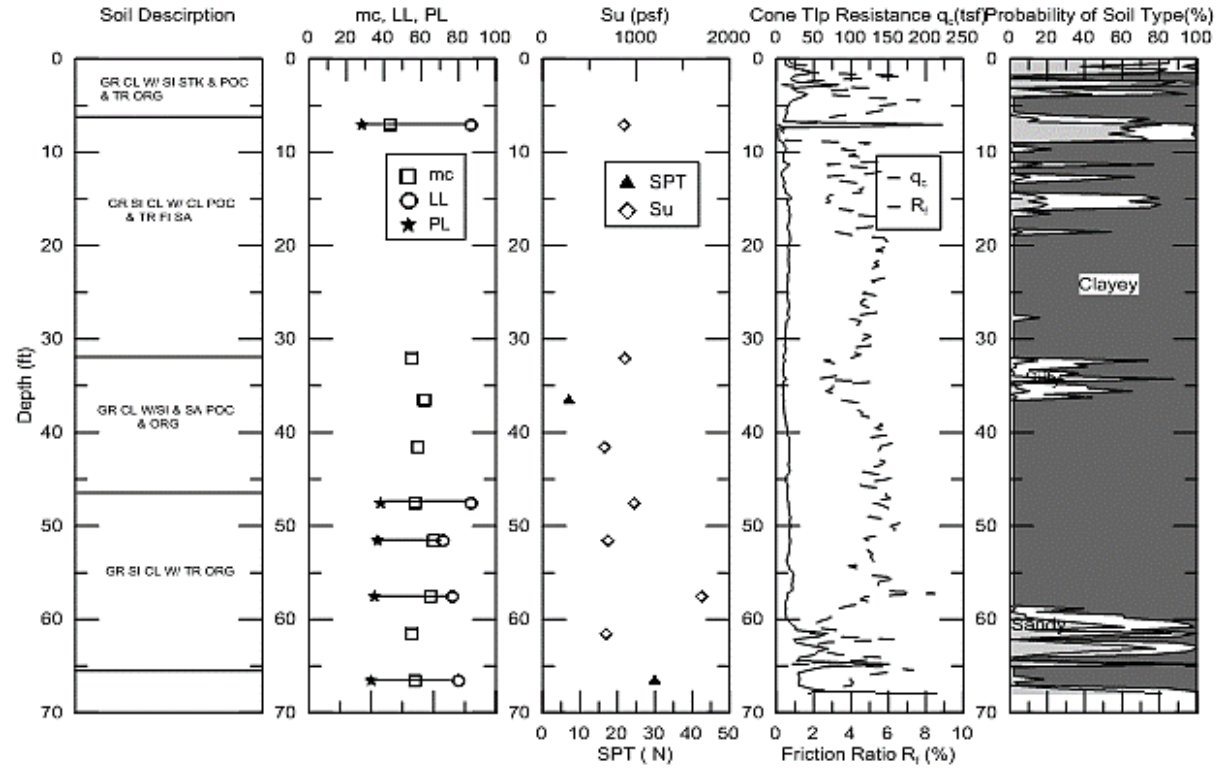


Figure 135. 66) 829-10-0013 TP#1

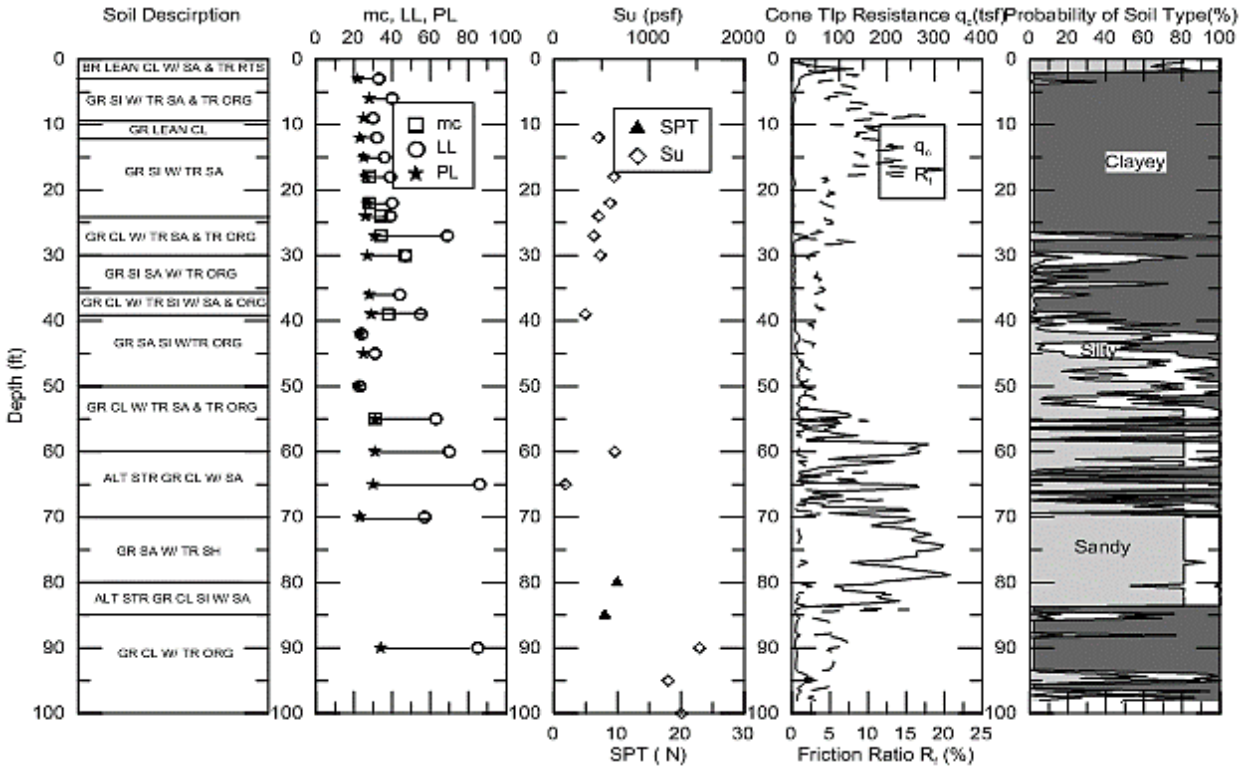


Figure 136. 67) 855-14-0013 TP#1

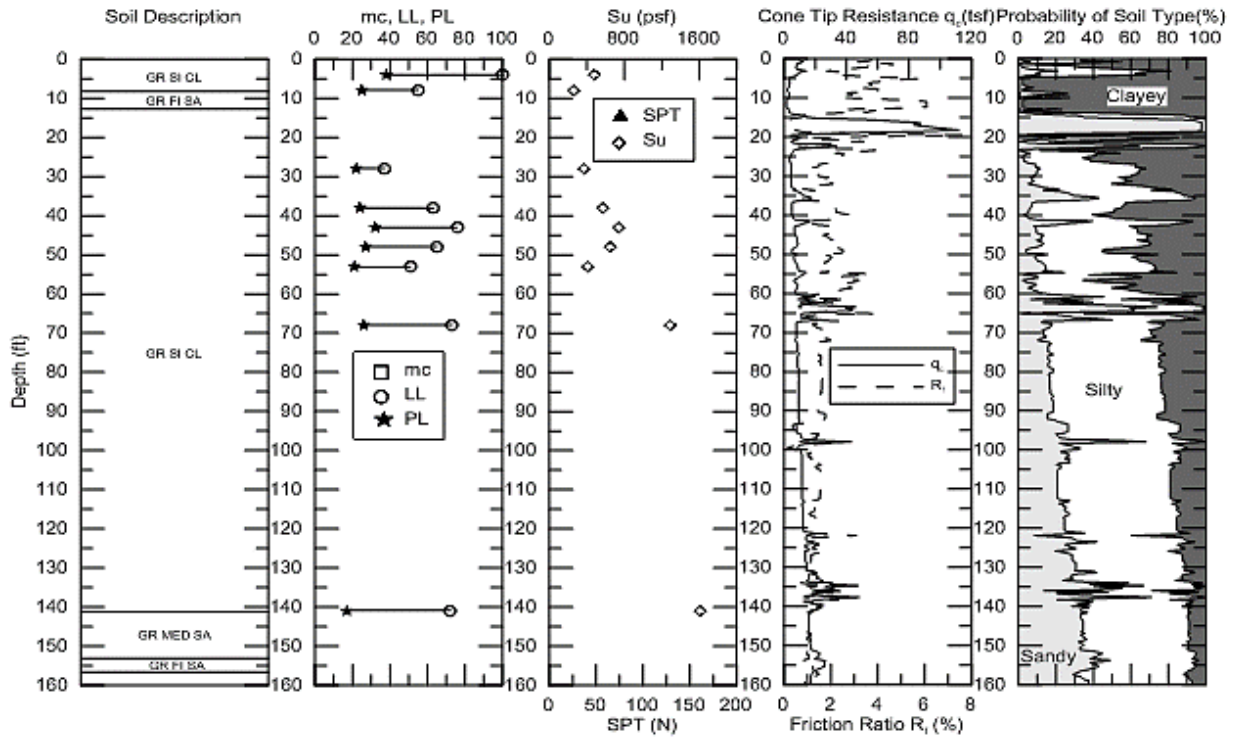


Figure 137. 68) Bayou Beouf TP#3

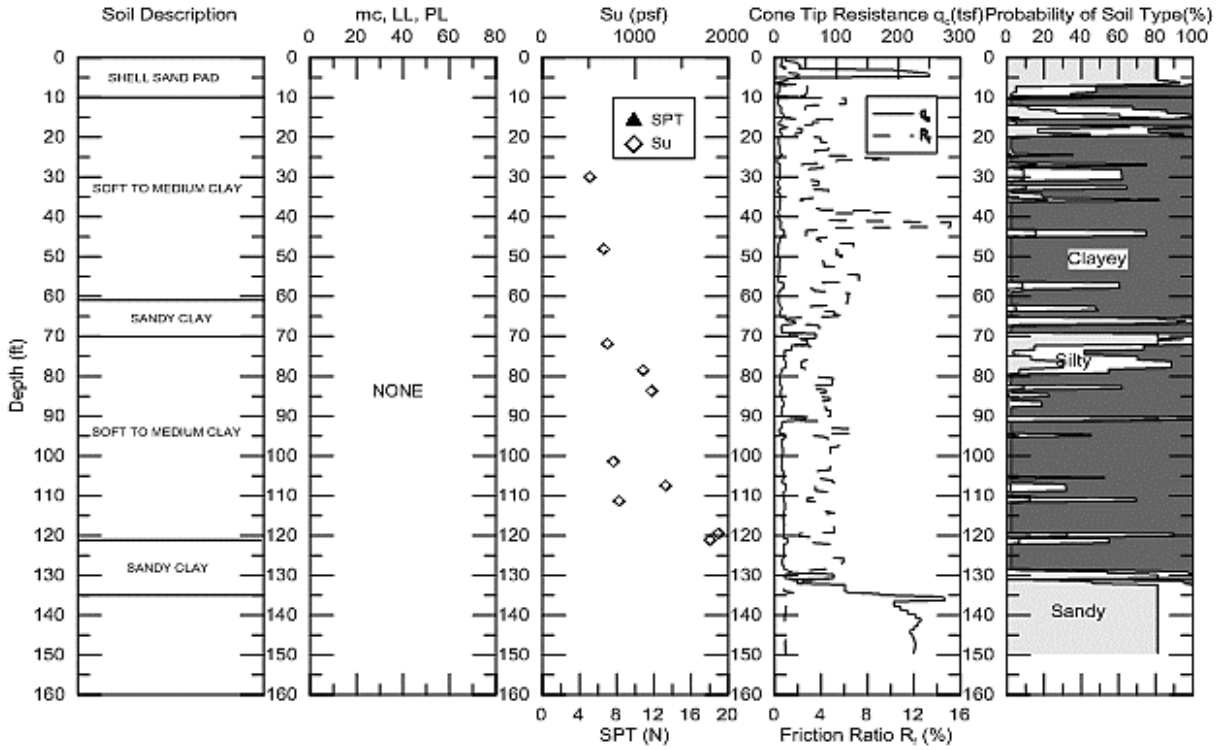


Figure 138. 69) Bayou Lacassine TP#1

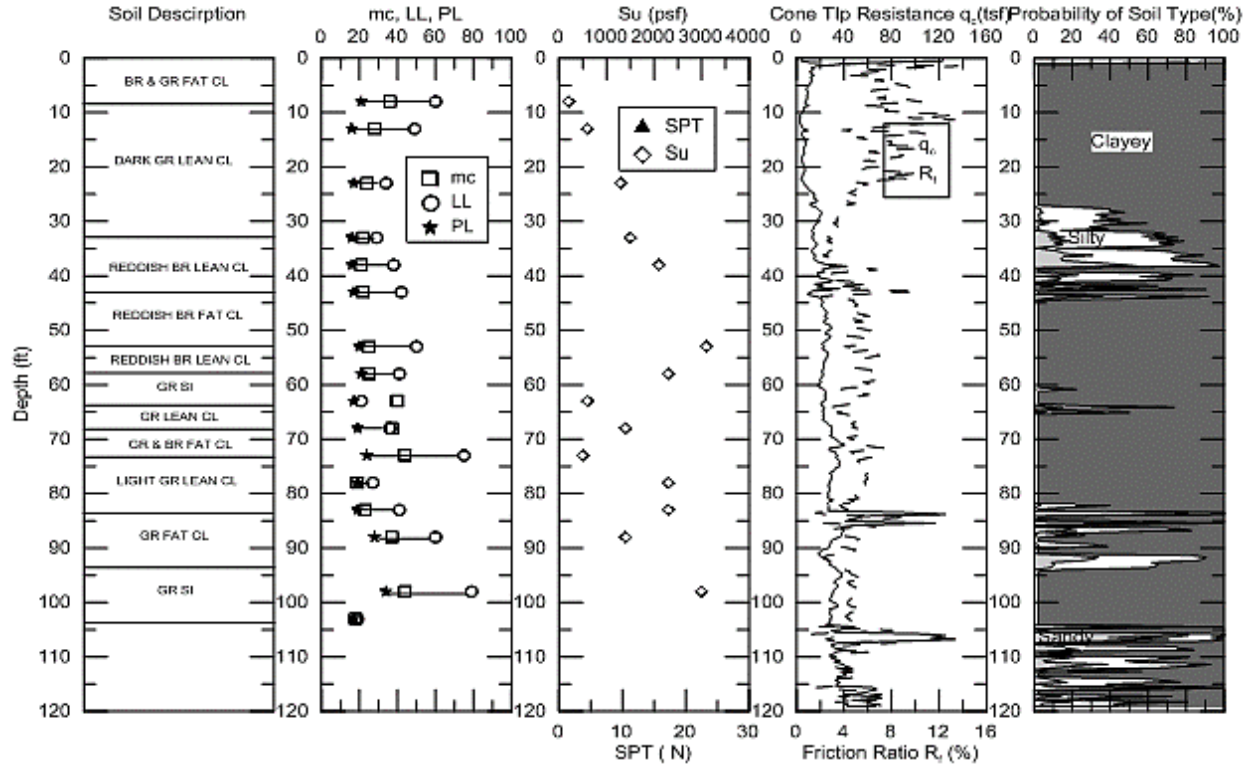


Figure 139. 70) Bayou Lacassine TP#3

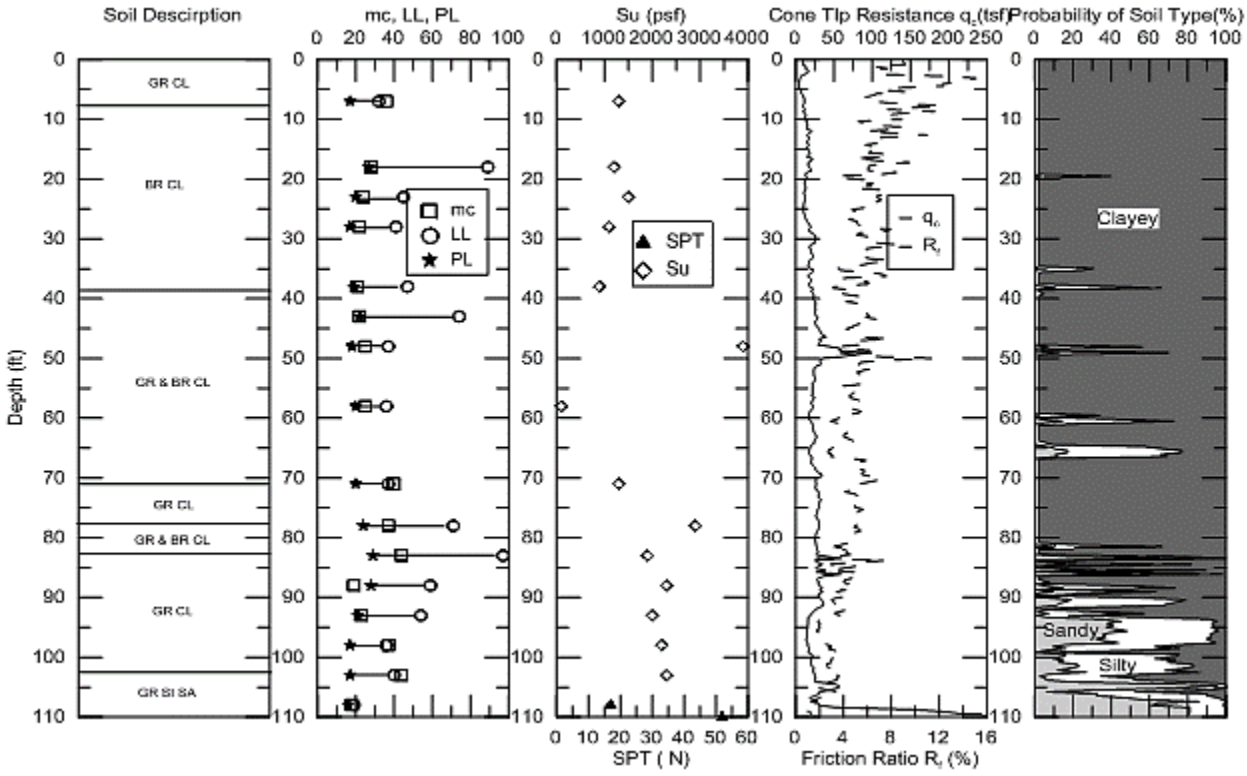


Figure 140. 71) Bayou Zourrie TP#1

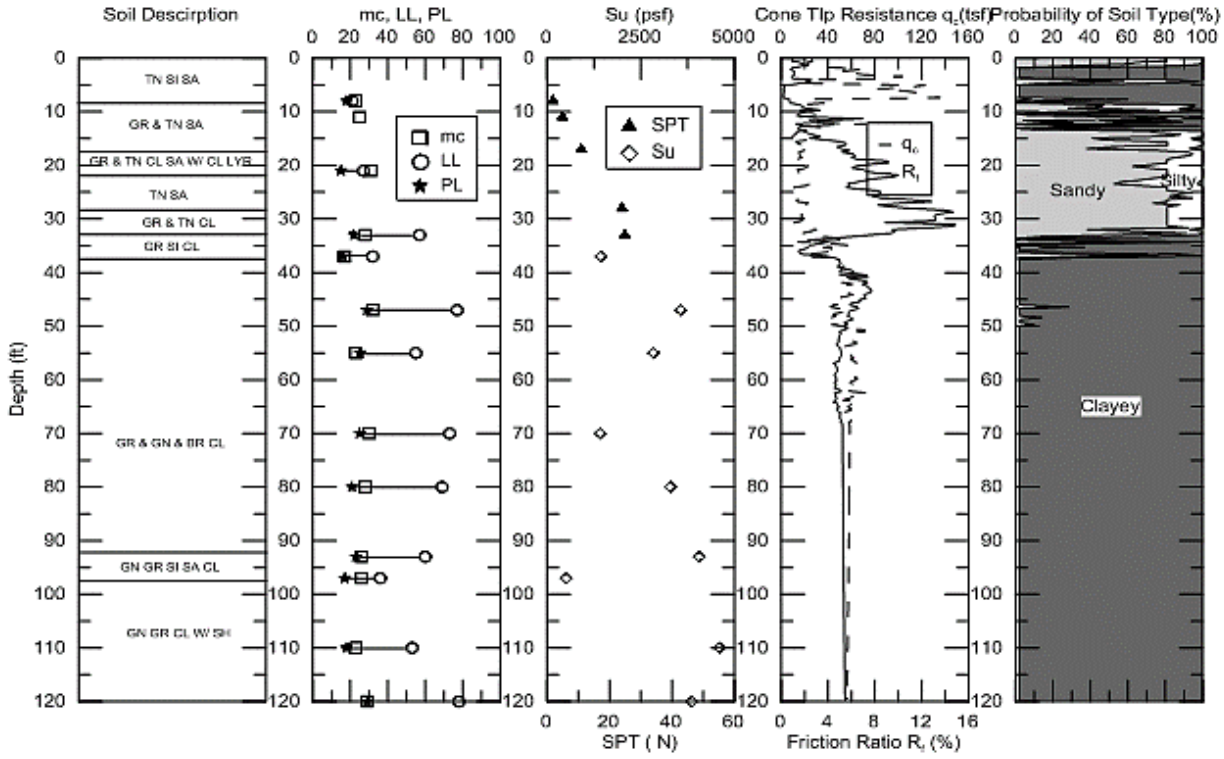


Figure 141. 72) LA-01 TP#2

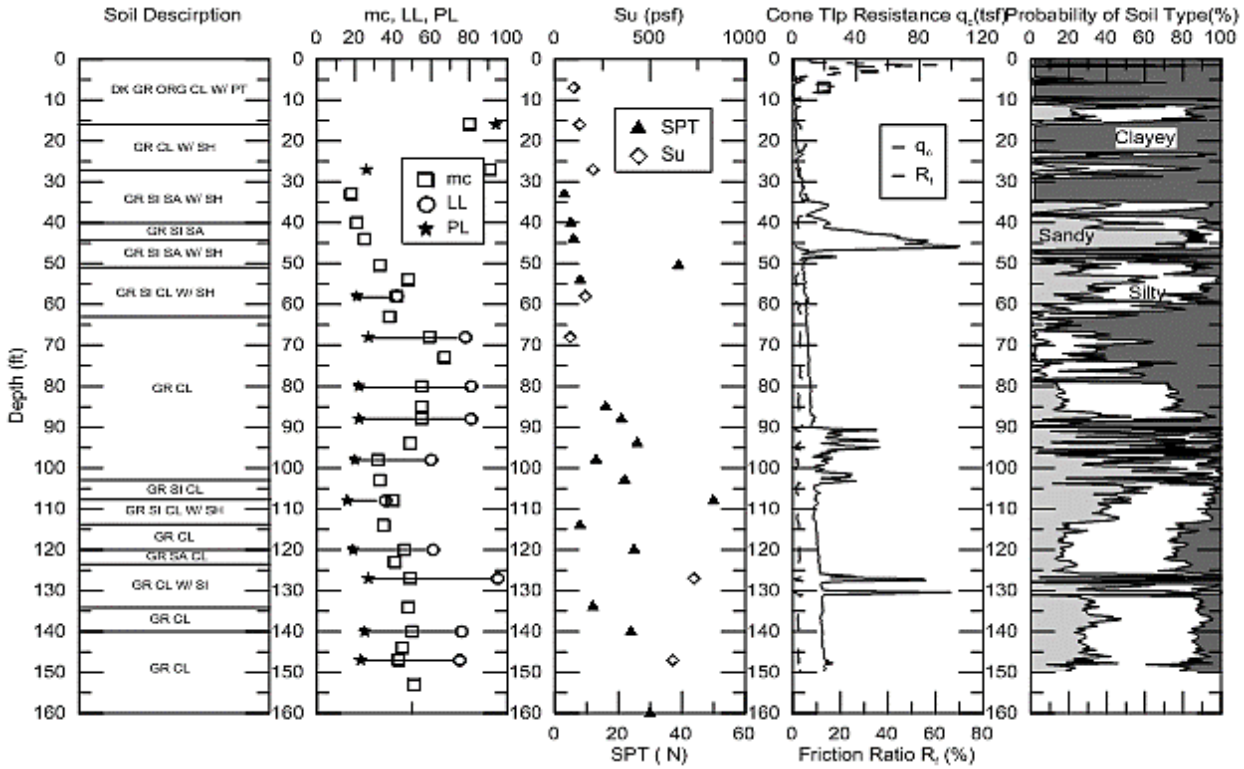


Figure 142. 73) LA-01 TP#3

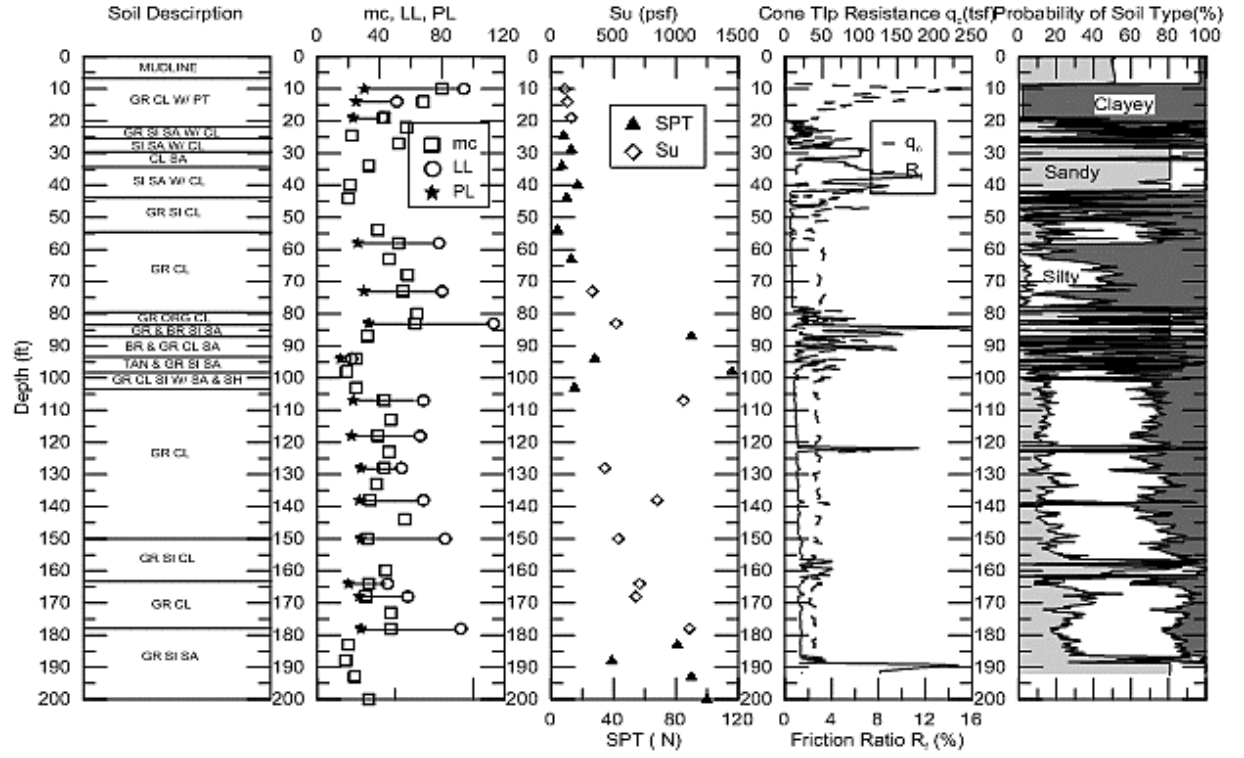


Figure 143. 74) LA-01 TP#4a

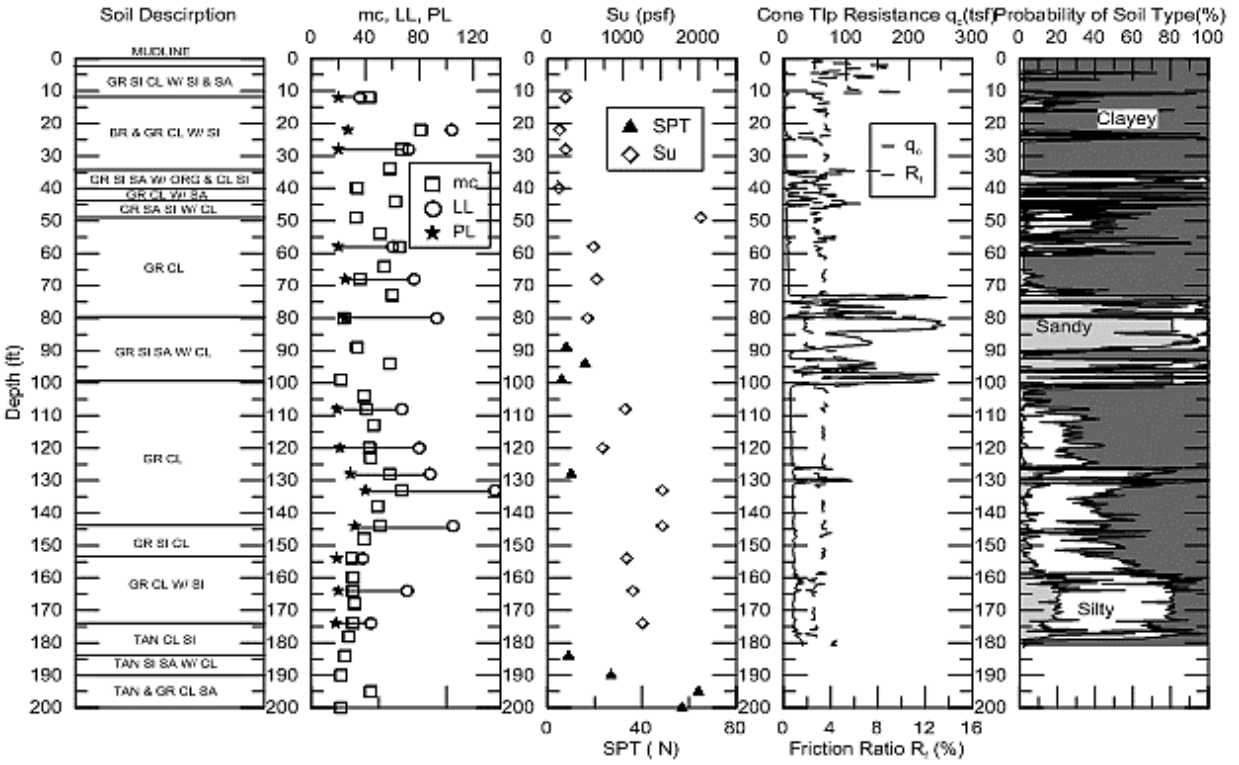


Figure 144. 75) LA-01 TP#4b

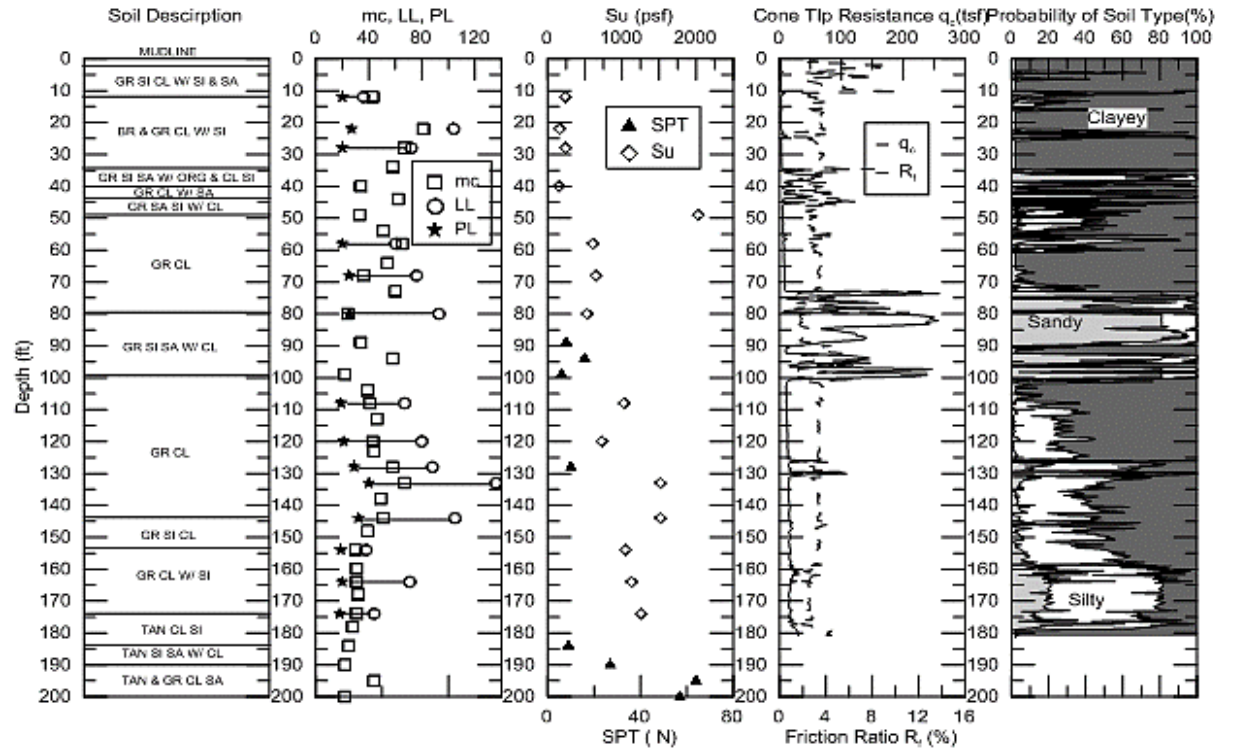


Figure 145. 76) LA-01 TP#5a

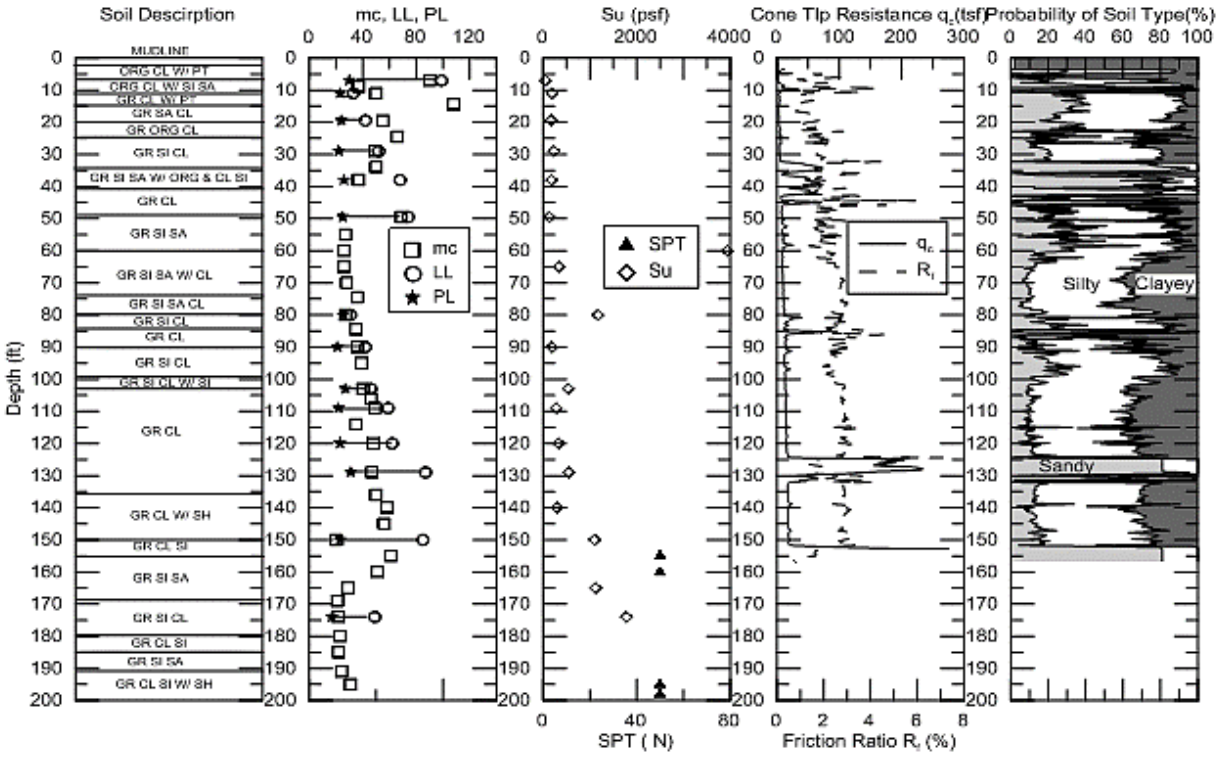


Figure 146. 77) LA-01 TP#5b

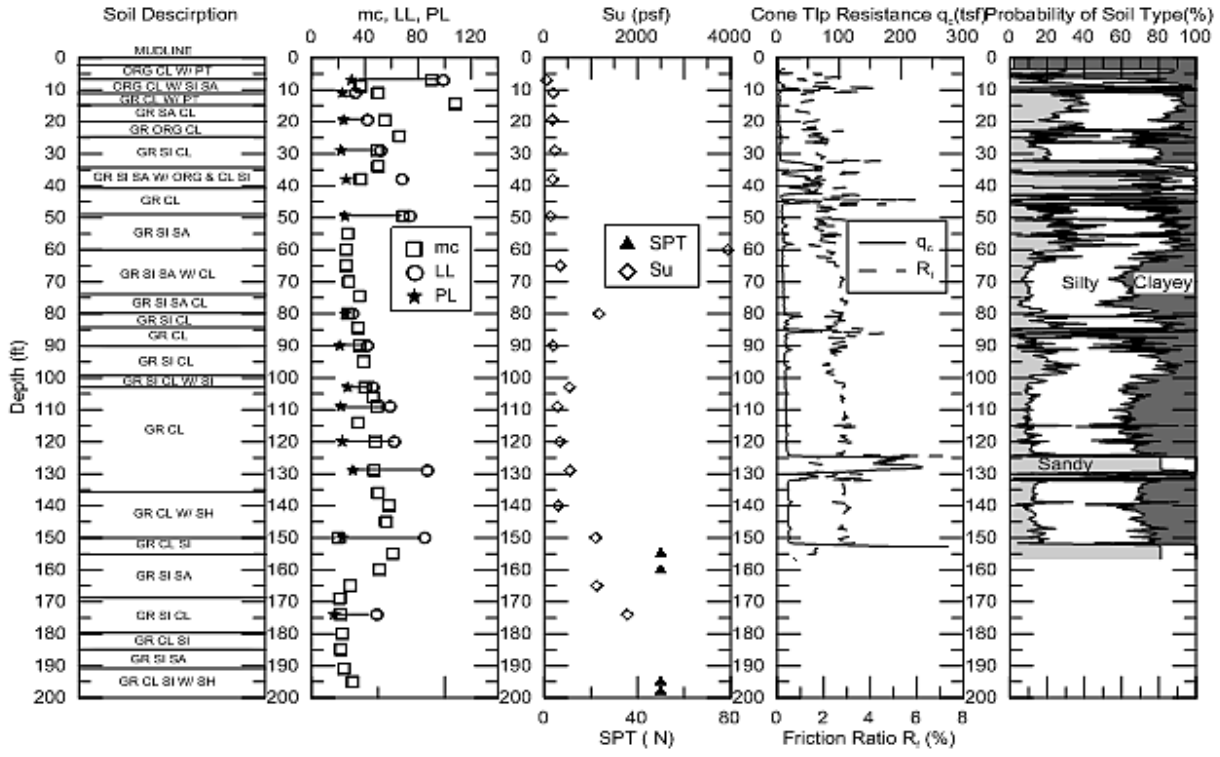


Figure 147. 78) 008-01-0042 TP#1

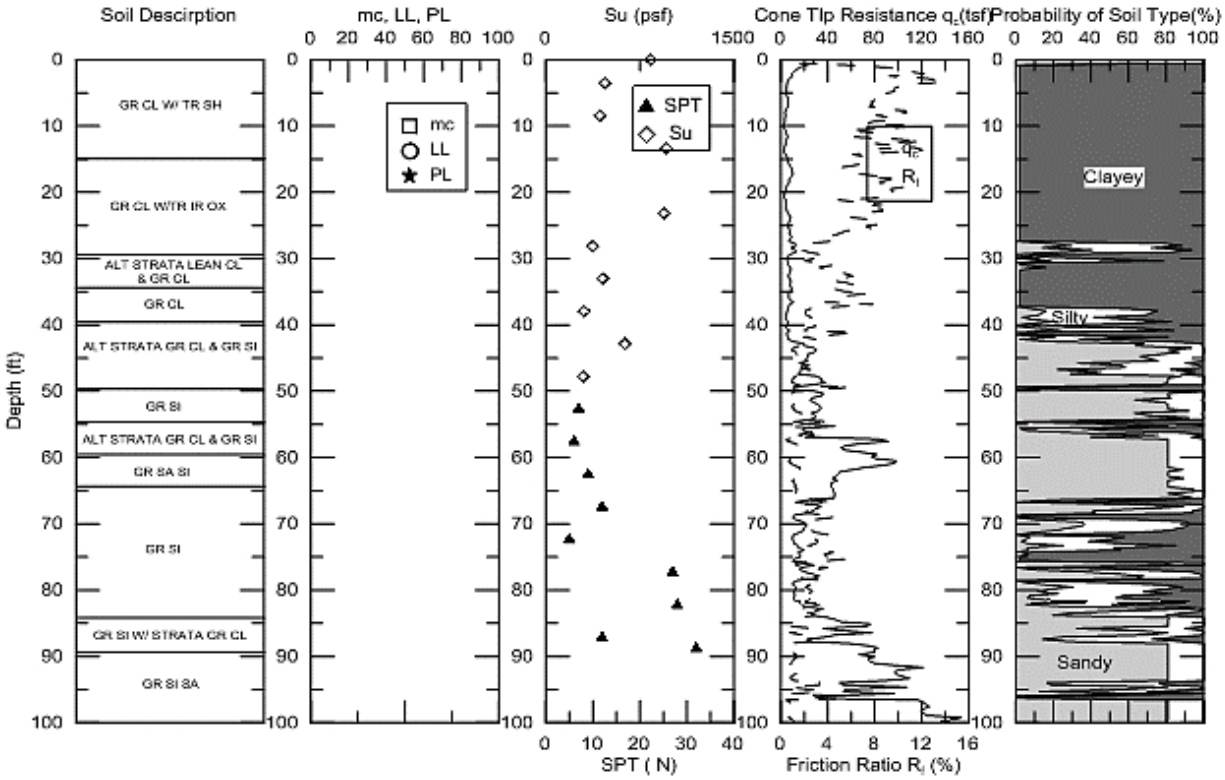


Figure 148. 79) 450-17-0025 TP#1

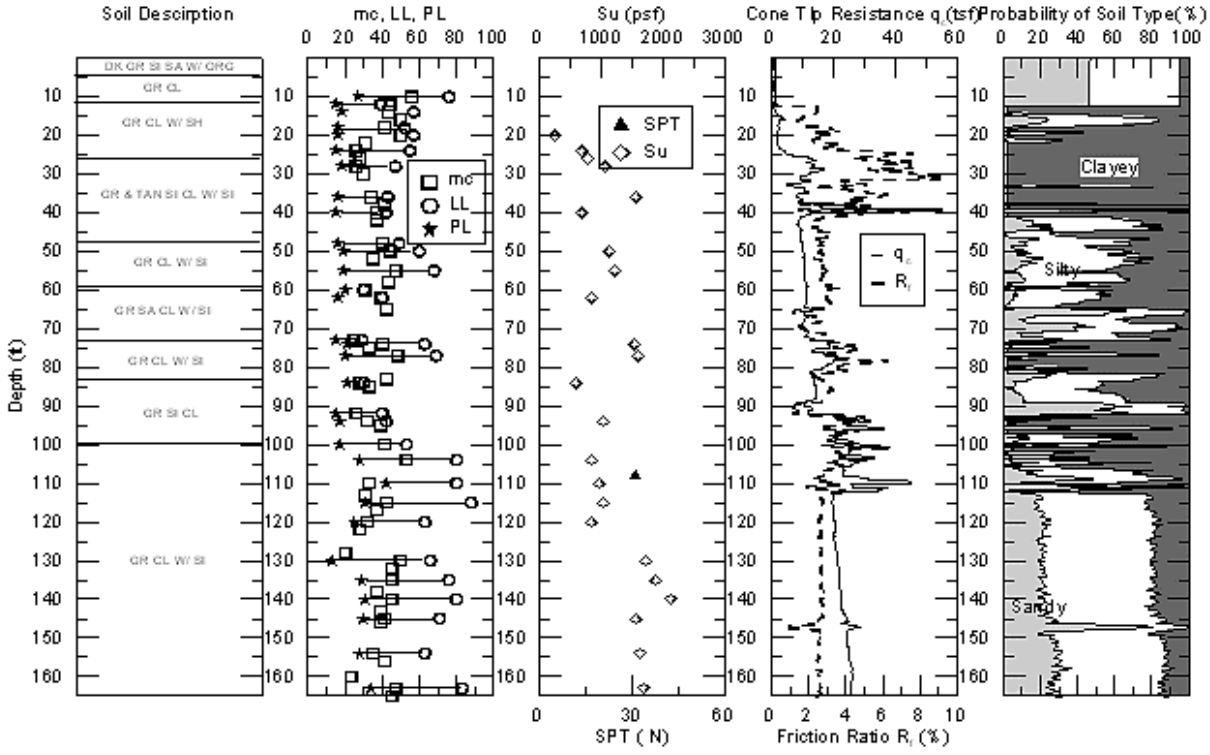
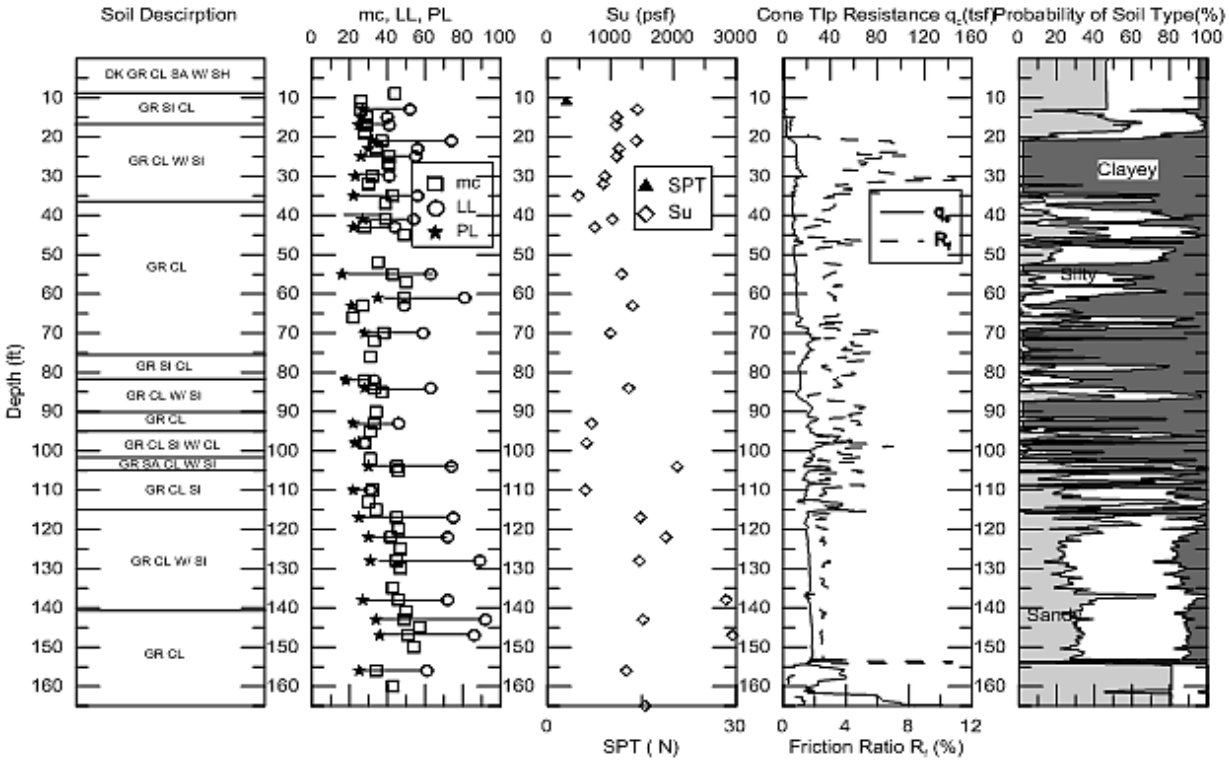


Figure 149. 80) 450-17-0025 TP#3



Appendix D

Load-Settlement Curves of DOTD State Projects Investigated in this Study

Figure 150. 1) 003-07-0019 TP#1

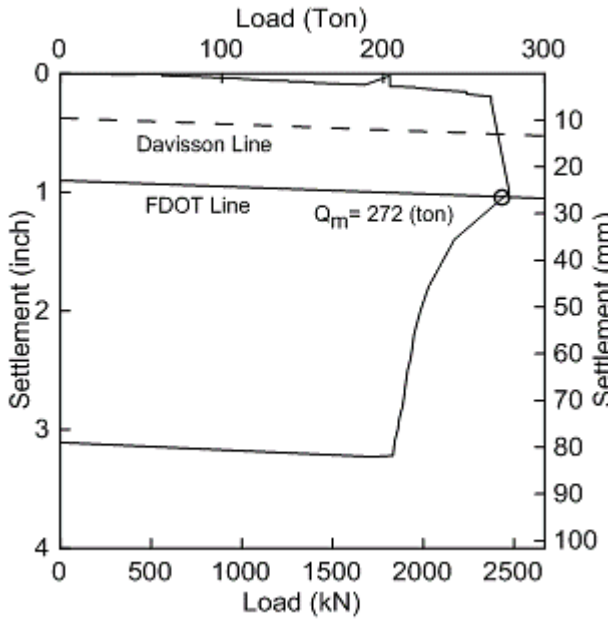


Figure 151. 2) 003-10-0011 TP#1

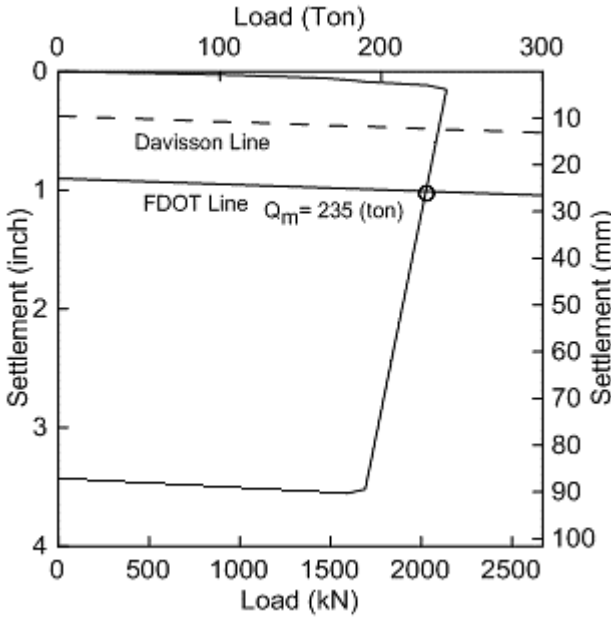


Figure 152. 3) 003-10-0011 TP#3

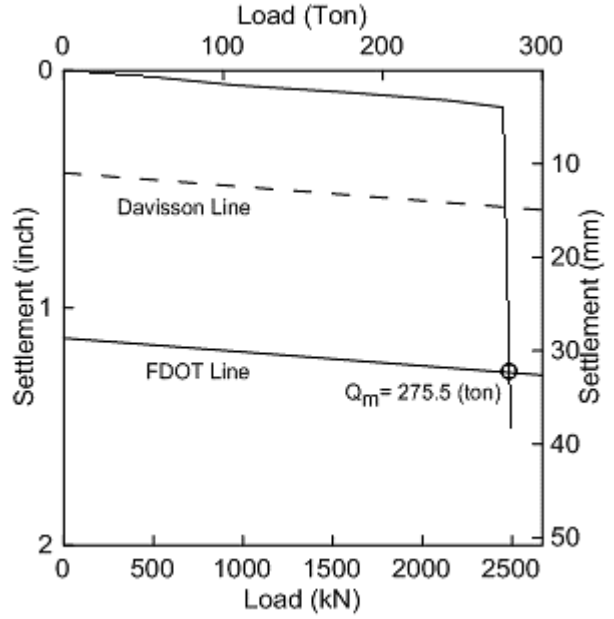


Figure 153. 4) 005-01-0056 TP#1

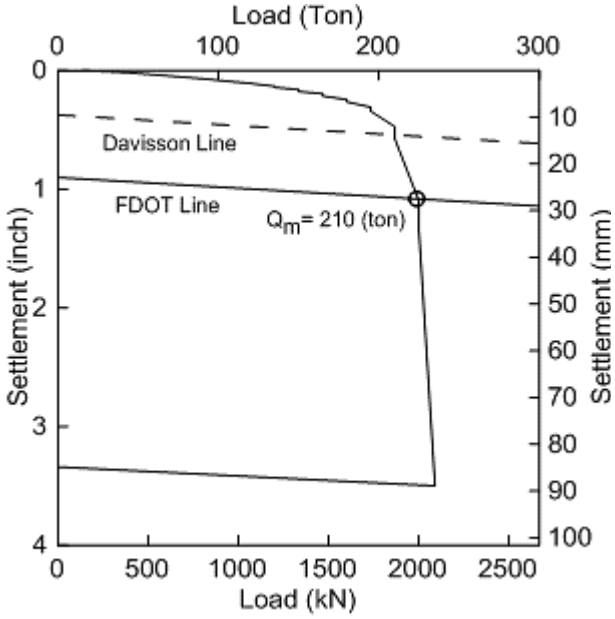


Figure 154. 5) 005-01-0056 TP#2

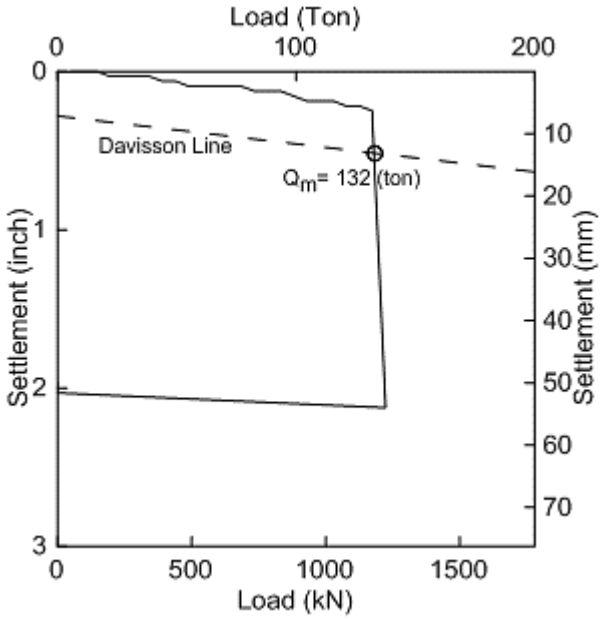


Figure 155. 6) 005-01-0056 TP#3

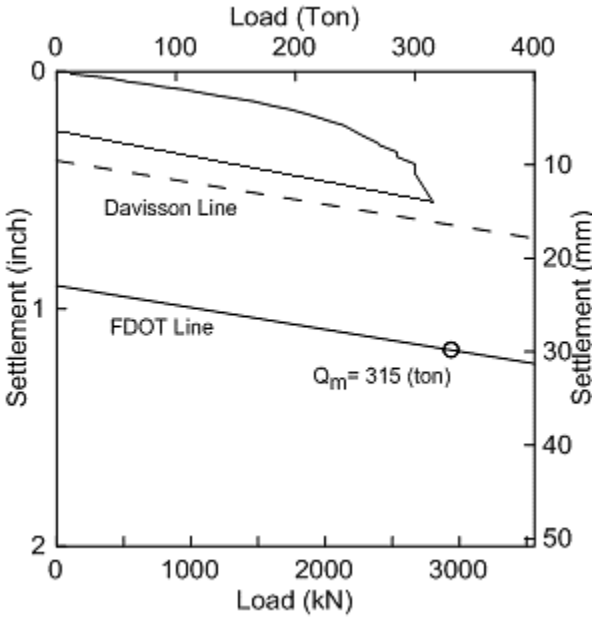


Figure 156. 7) 047-02-0022 TP#2

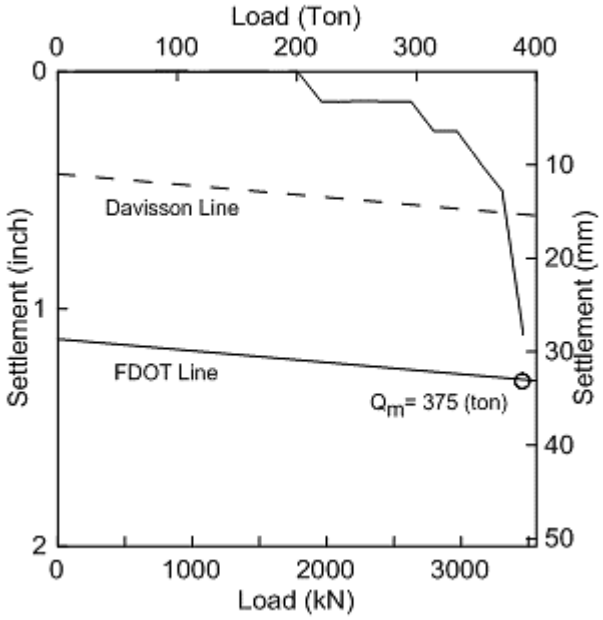


Figure 157. 8) 061-05-0044 TP#2

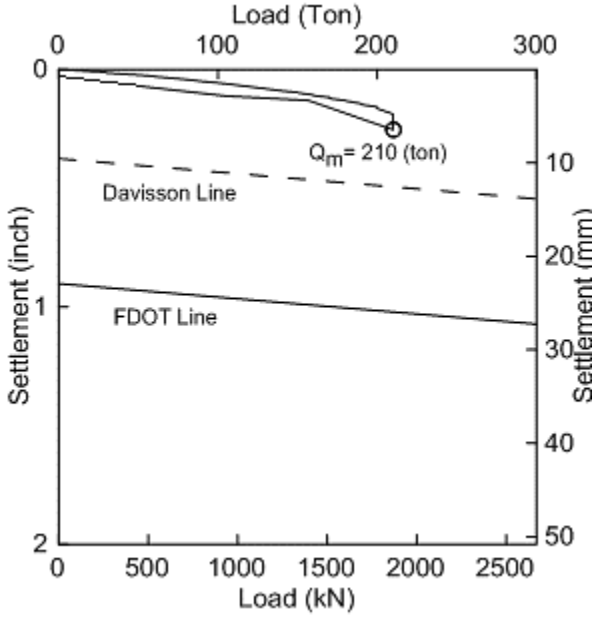


Figure 158. 9) 064-06-0036 TP#1

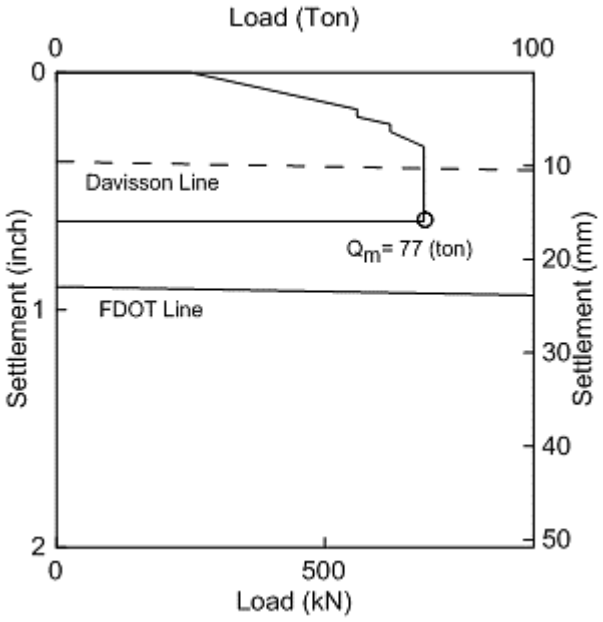


Figure 159. 10) 065-90-0024 TP#1

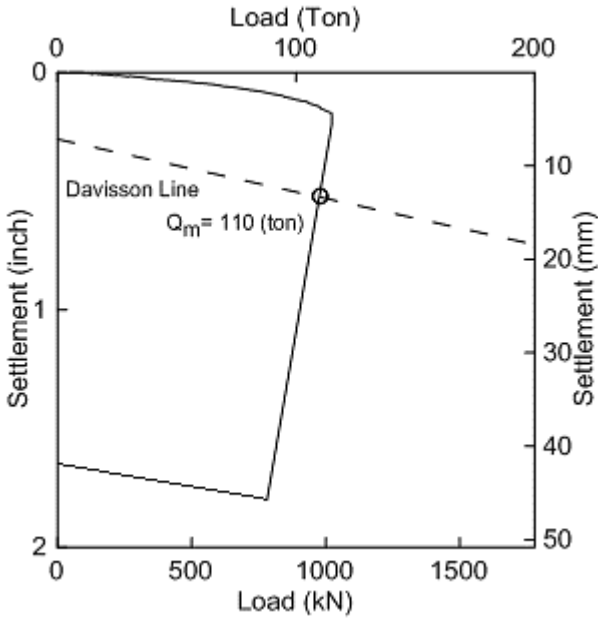


Figure 160. 11) 065-90-0024 TP#2

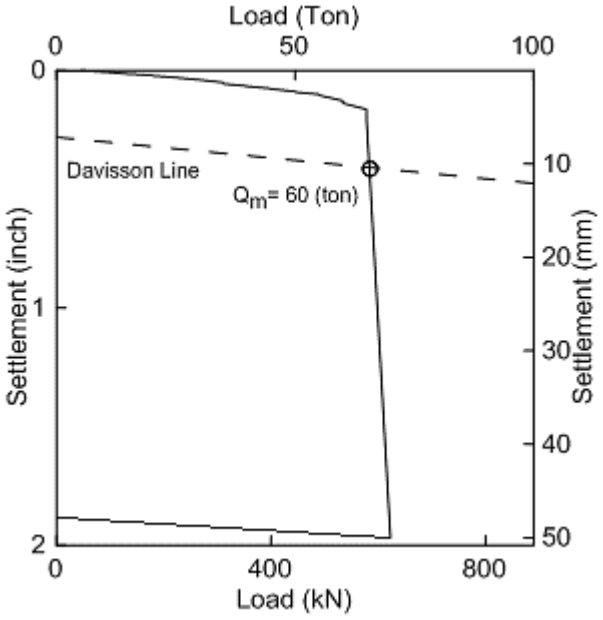


Figure 161. 12) 065-90-0024 TP#3

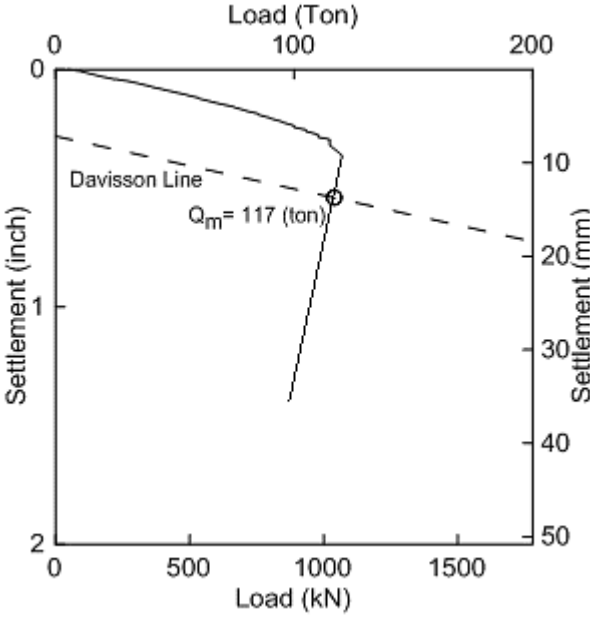


Figure 162. 13) 065-90-0024 TP#4

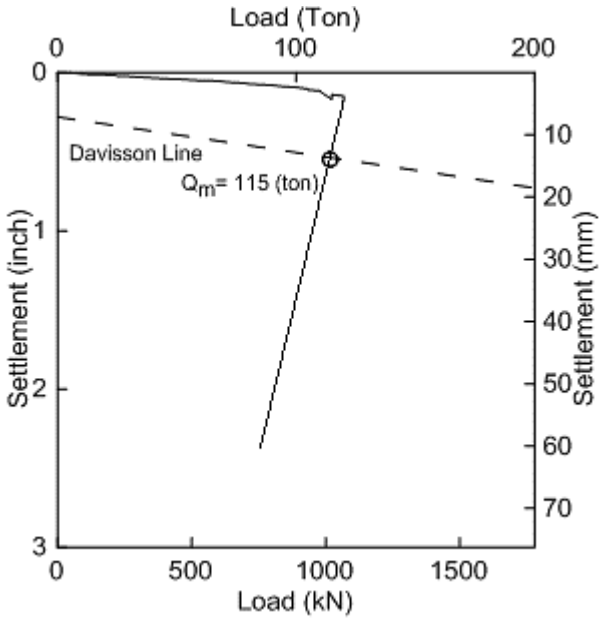


Figure 163. 14) 065-90-0024 TP#5

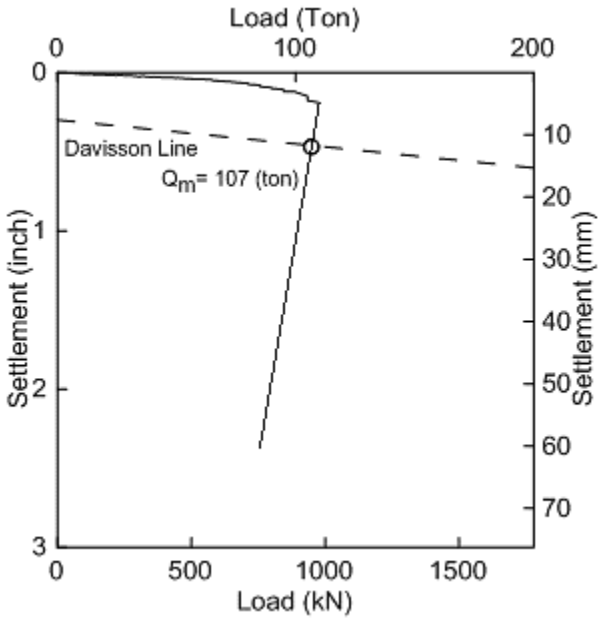


Figure 164. 15) 065-90-0024 TP#6

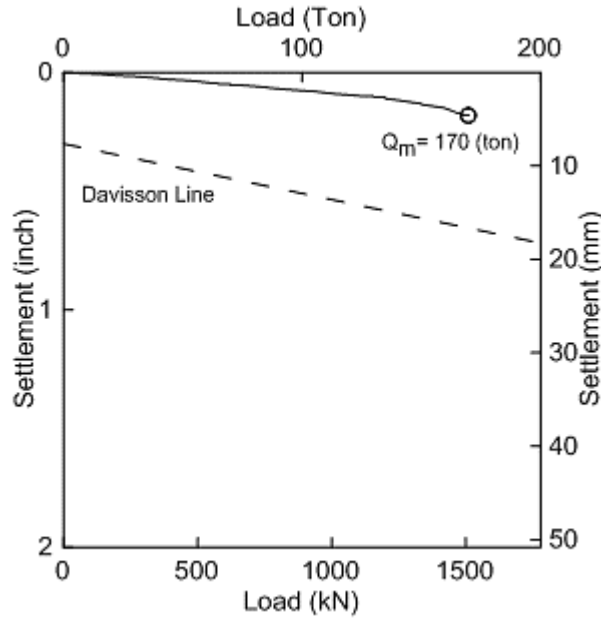


Figure 165. 16) 239-01-0080 TP#3

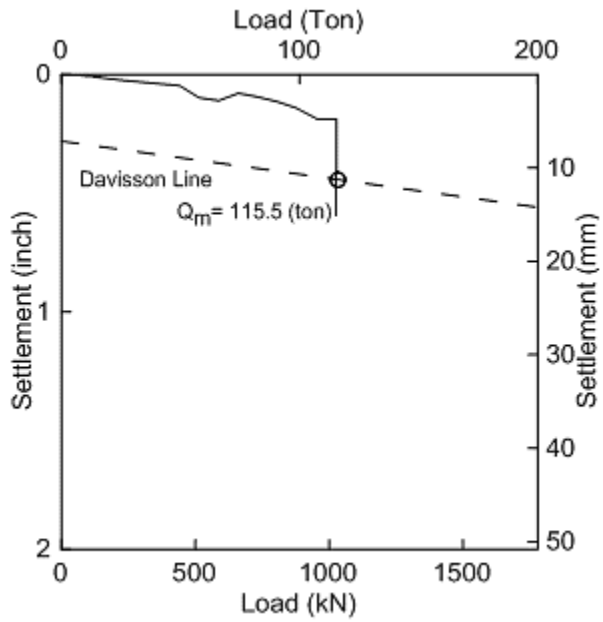


Figure 166. 17) 239-01-0080 TP#4

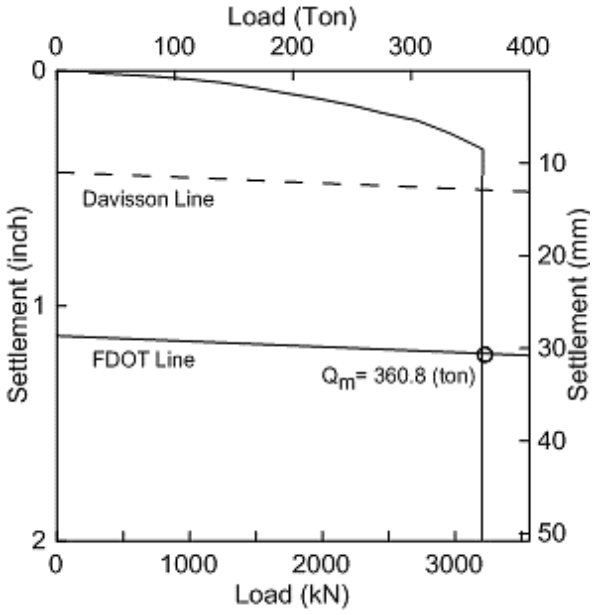


Figure 167. 18) 260-05-0020 TP#1

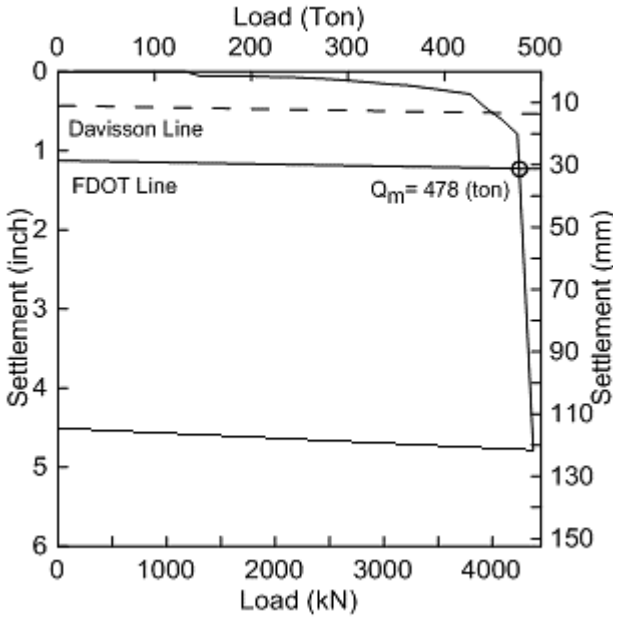


Figure 168. 19) 260-05-0020 TP#3

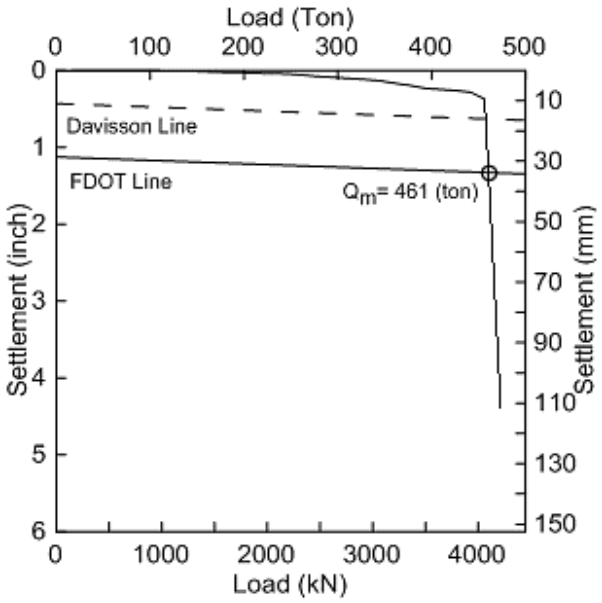


Figure 169. 20) 262-06-0009 TP#1

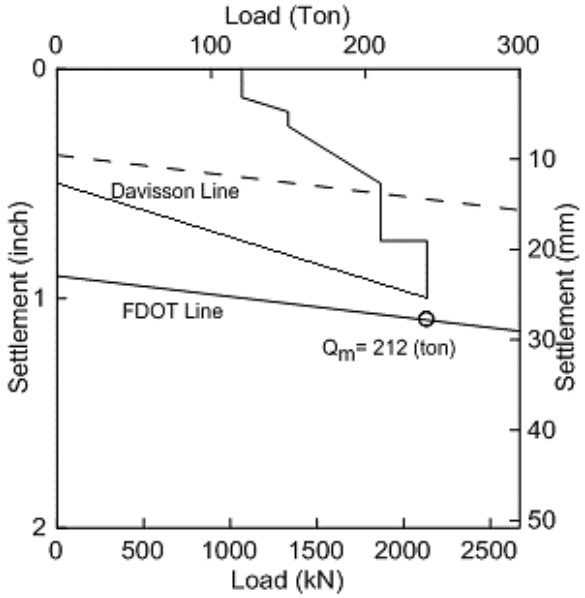


Figure 170. 21) 262-06-0009 TP#2

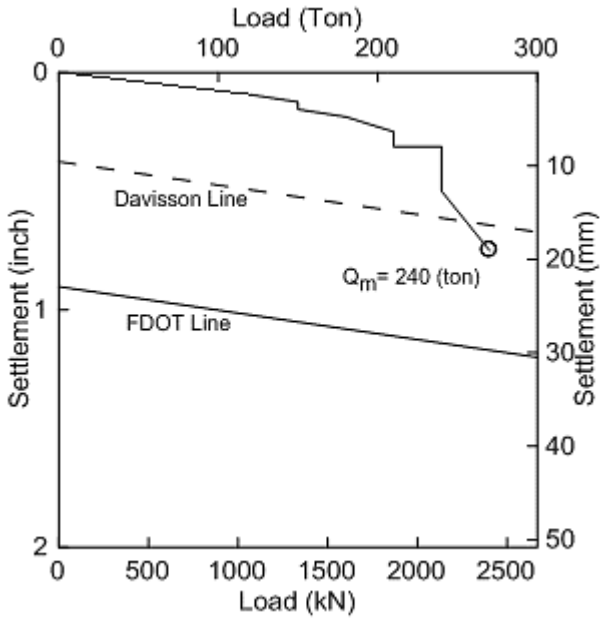


Figure 171. 22) 283-03-0052 TP#1

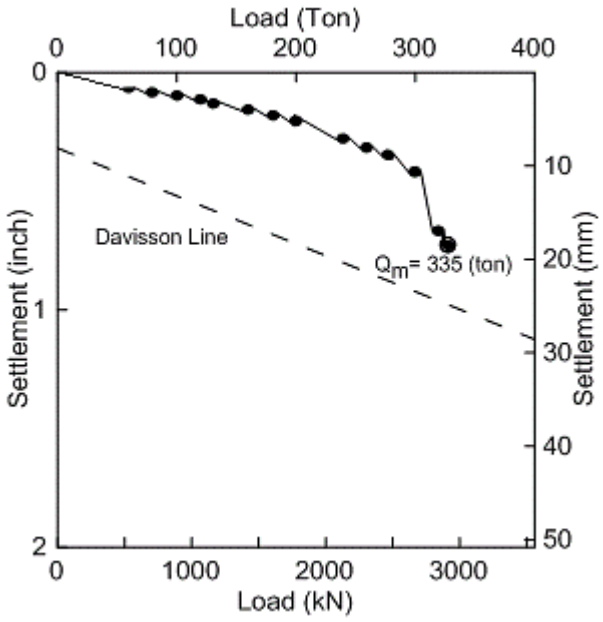


Figure 172. 23) 424-04-0026 TP#2

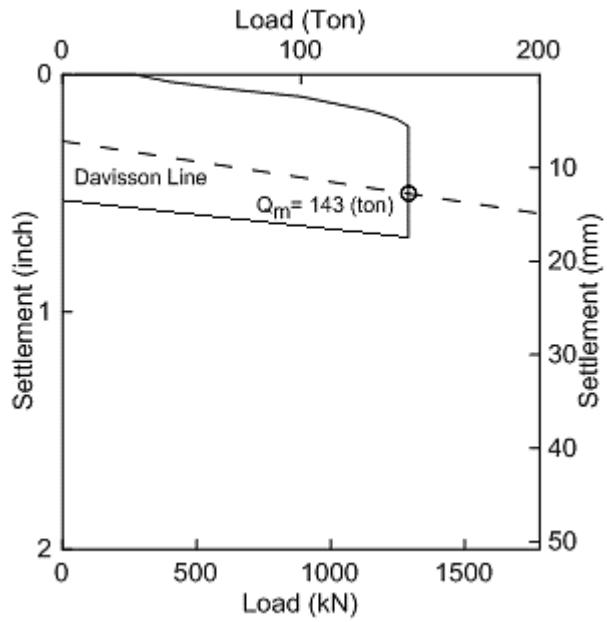


Figure 173. 24) 424-04-0027 TP#1

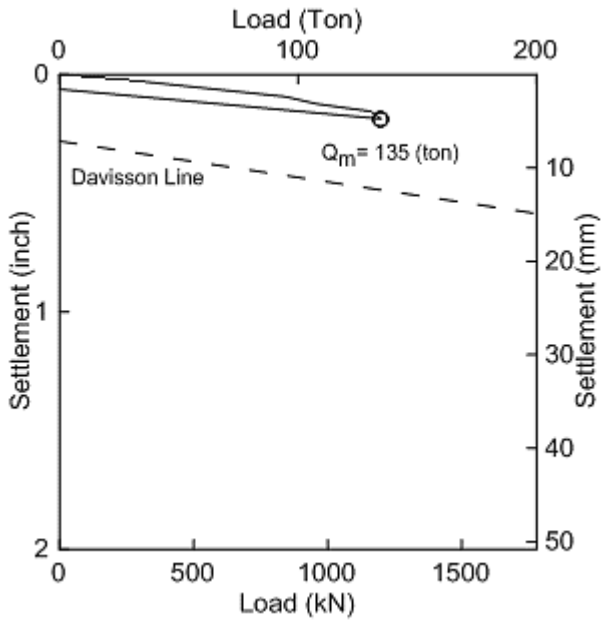


Figure 174. 25) 424-04-0027 TP#2

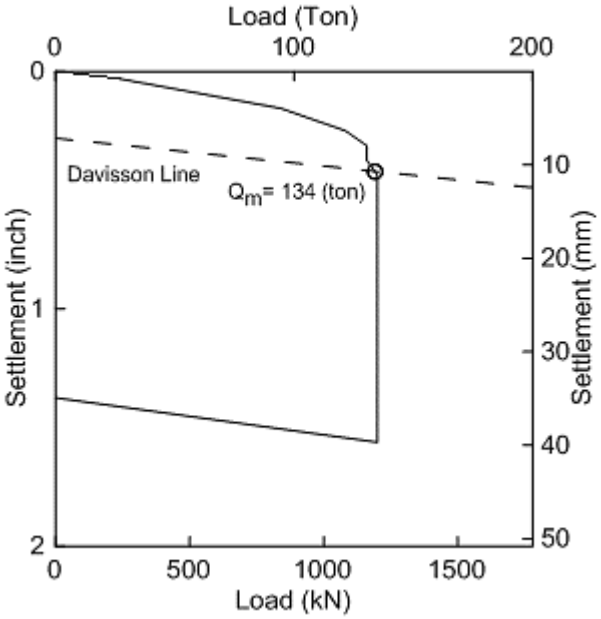


Figure 175. 26) 424-05-0078 TP#1

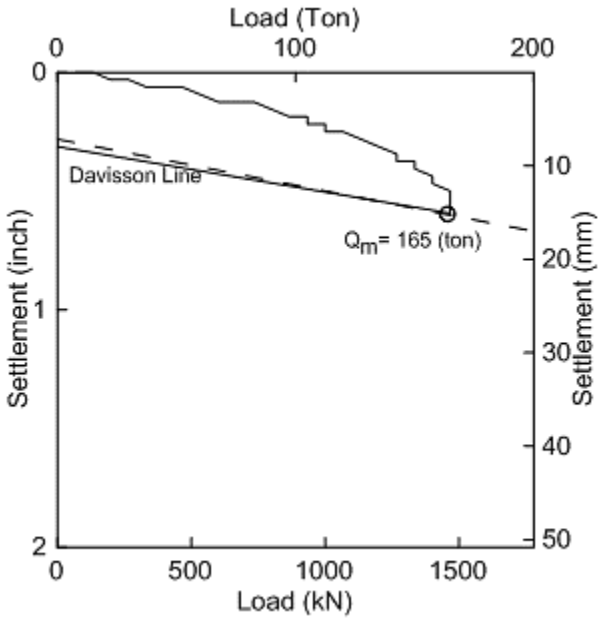


Figure 176. 27) 424-05-0078 TP#2

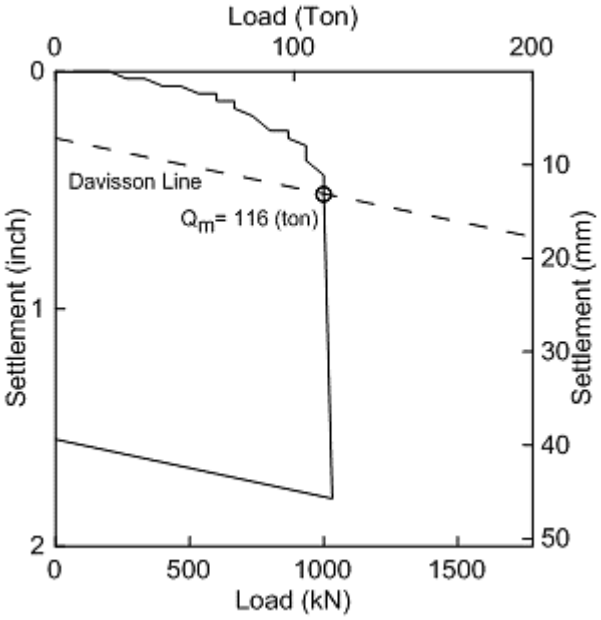


Figure 177. 28) 424-05-0078 TP#5

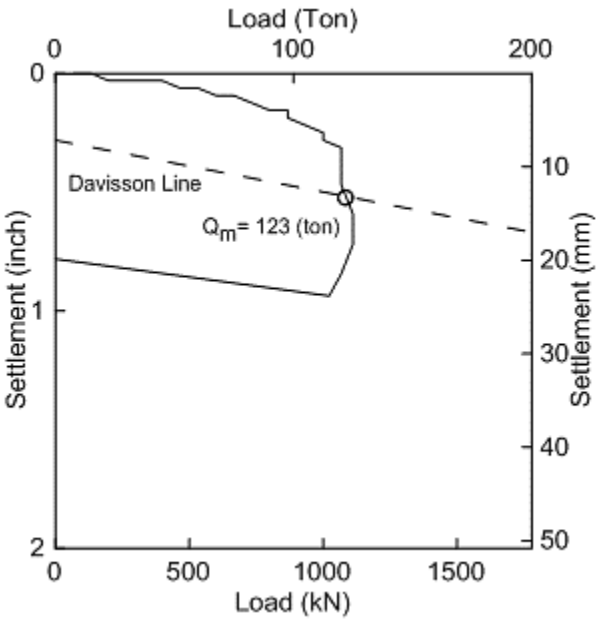


Figure 178. 29) 424-05-0081 TP#1

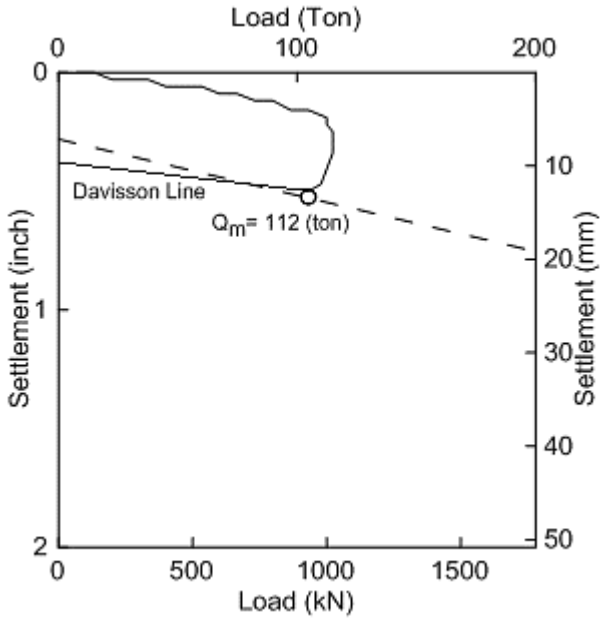


Figure 179. 30) 424-05-0081 TP#2

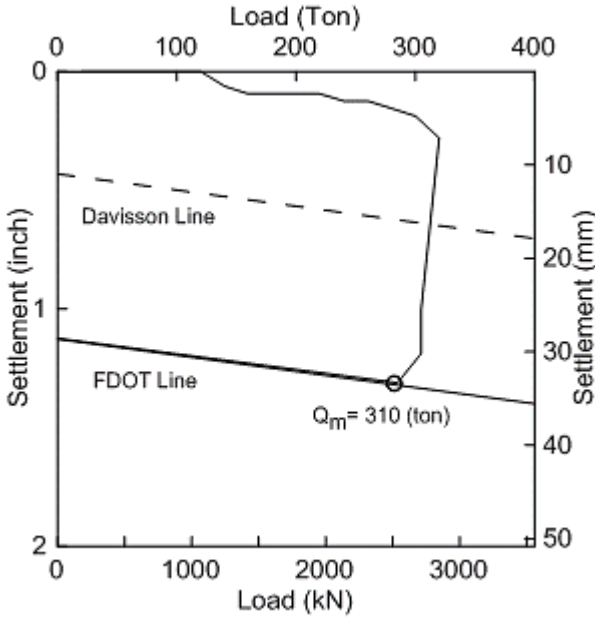


Figure 180. 31) 424-05-0081 TP#3

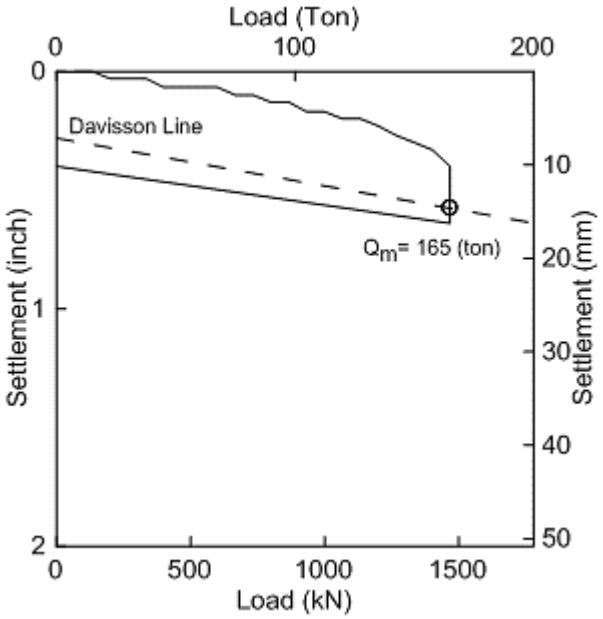


Figure 181. 32) 424-05-0081 TP#4

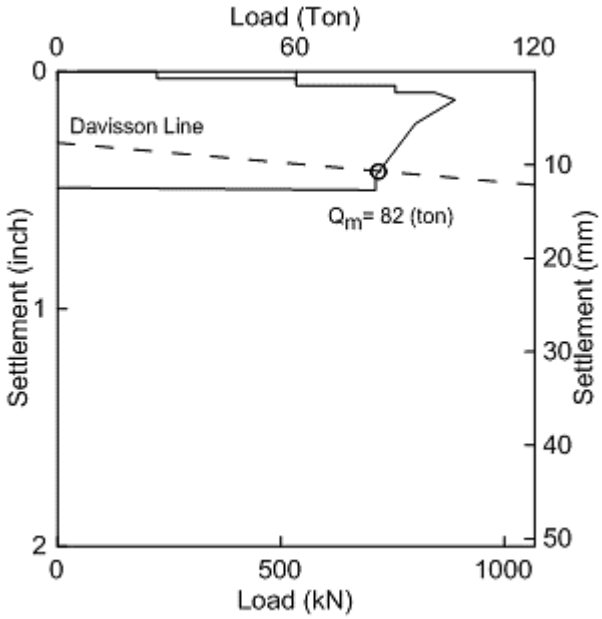


Figure 182. 33) 424-05-0087 TP#1

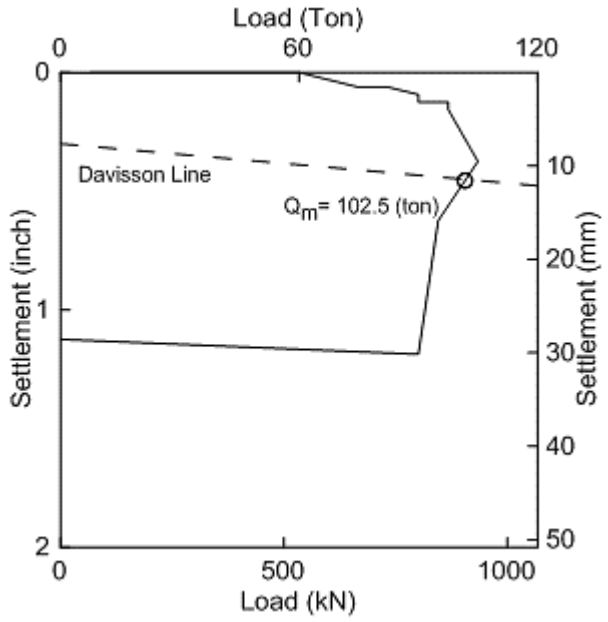


Figure 183. 34) 424-05-0087 TP#2

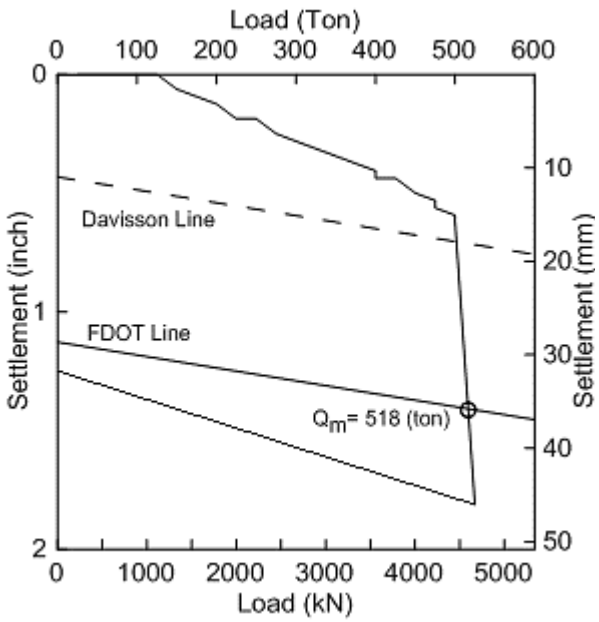


Figure 184. 35) 424-05-0087 TP#3

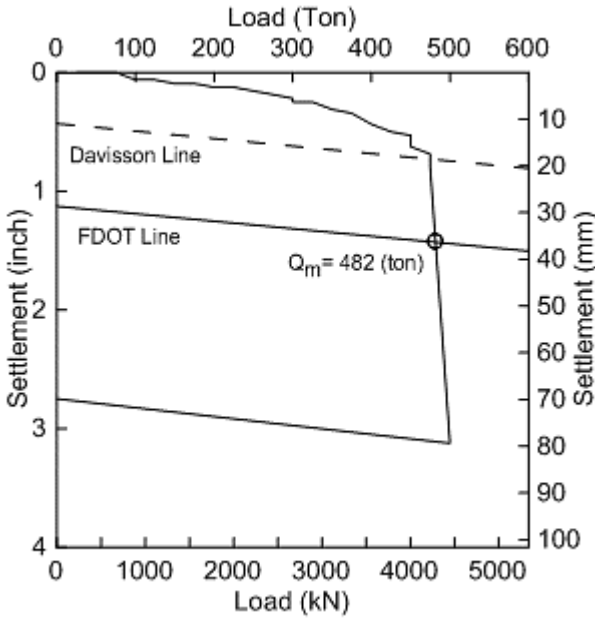


Figure 185. 36) 424-05-0087 TP#4

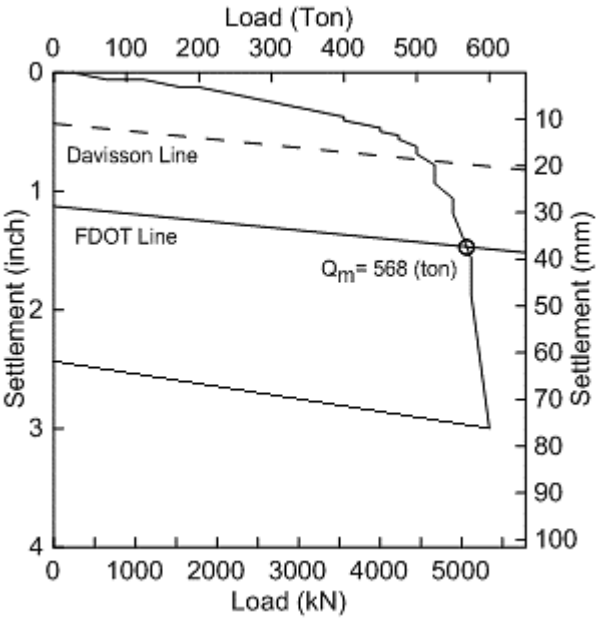


Figure 186. 37) 424-05-0087 TP#5

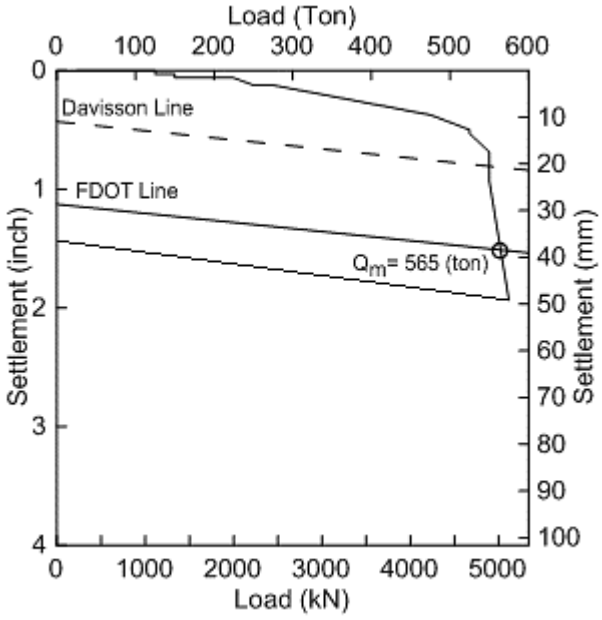


Figure 187. 38) 424-05-0087 TP#7

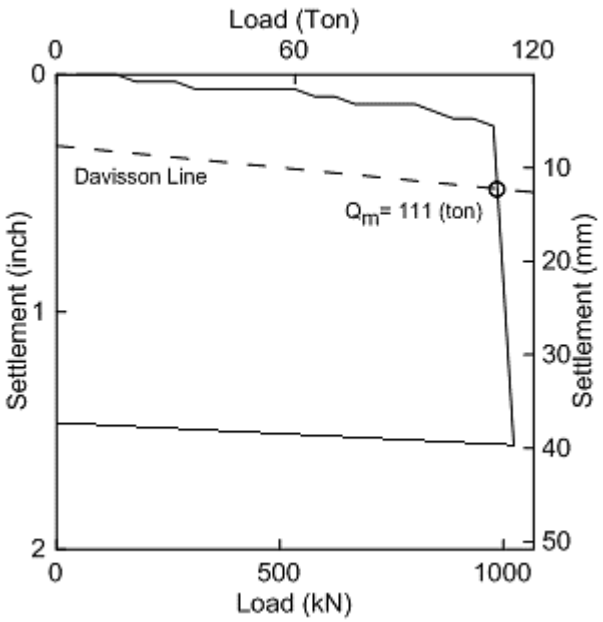


Figure 188. 39) 424-07-0008 TP#1

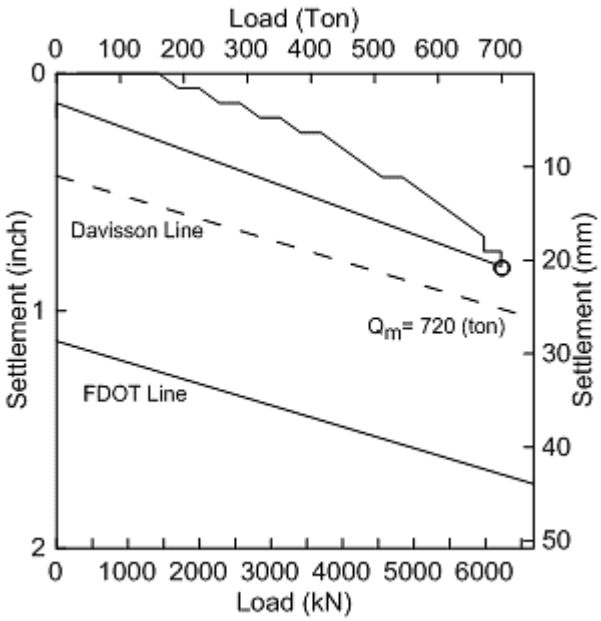


Figure 189. 40) 424-07-0008 TP#3

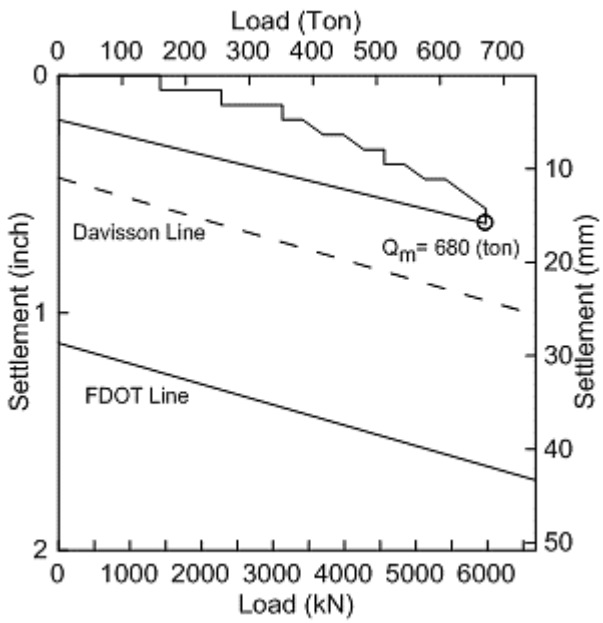


Figure 190. 41) 424-07-0009 TP#3

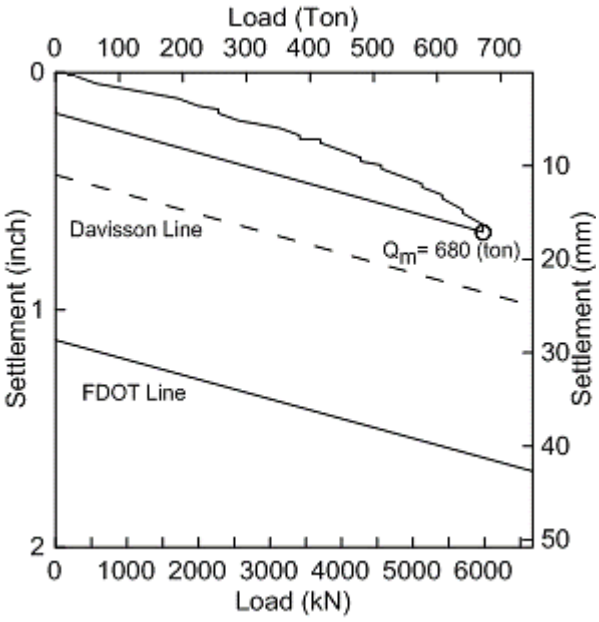


Figure 191. 42) 424-07-0009 TP#4

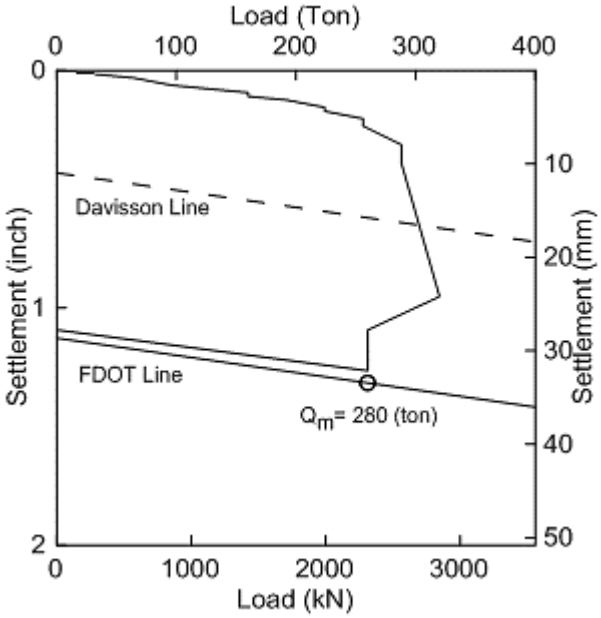


Figure 192. 43) 424-07-0009 TP#4A

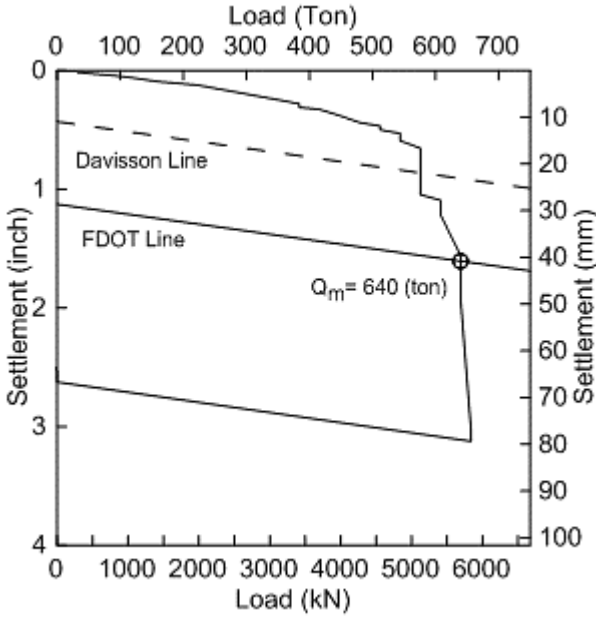


Figure 193. 44) 424-06-0005 TP#1

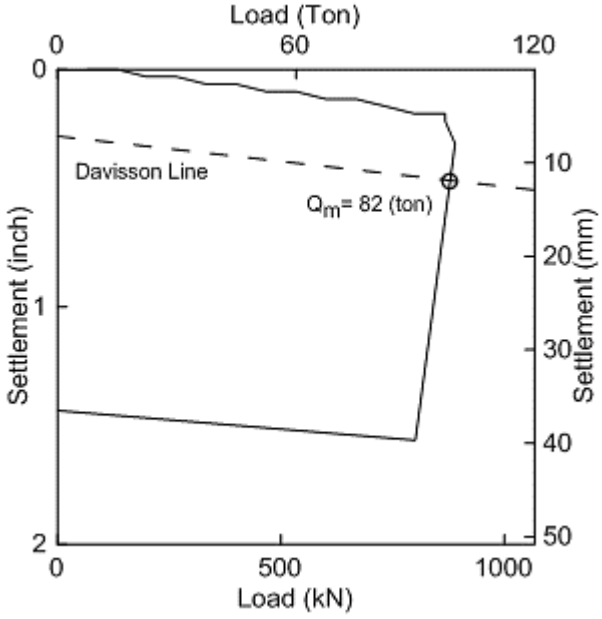


Figure 194. 45) 424-06-0005 TP#2

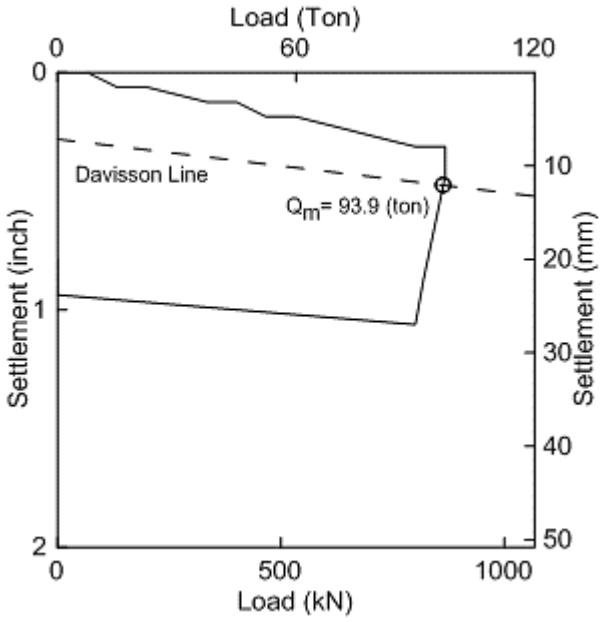


Figure 195. 46) 424-06-0005 TP#3

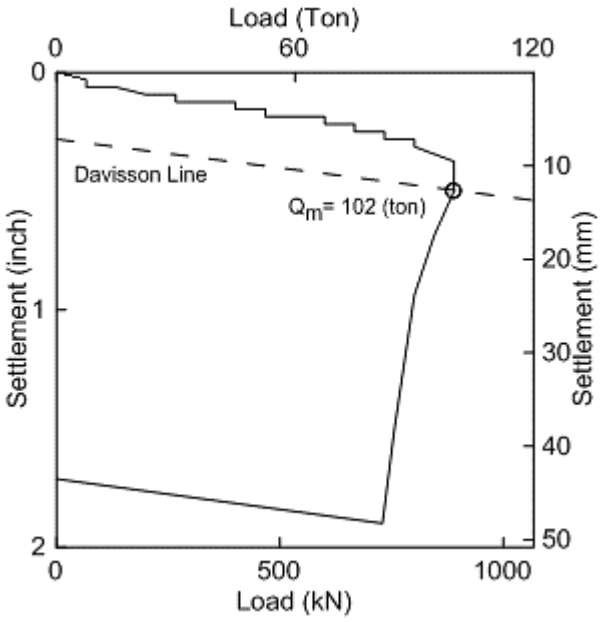


Figure 196. 47) 424-06-0005 TP#4

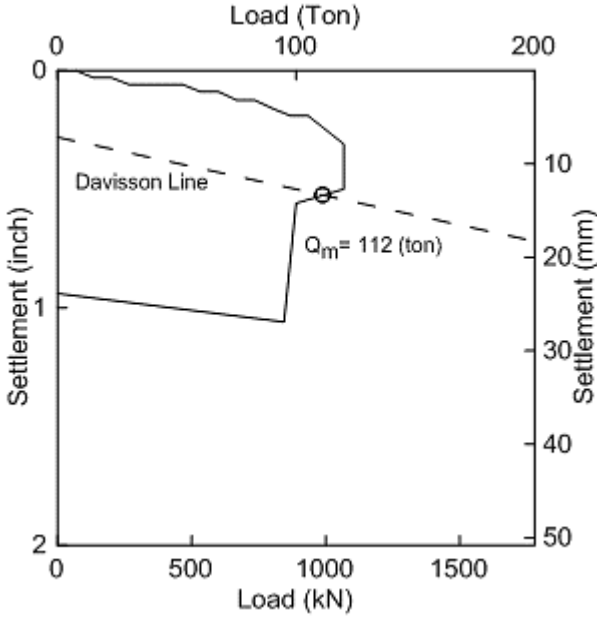


Figure 197. 48) 424-06-0005 TP#5

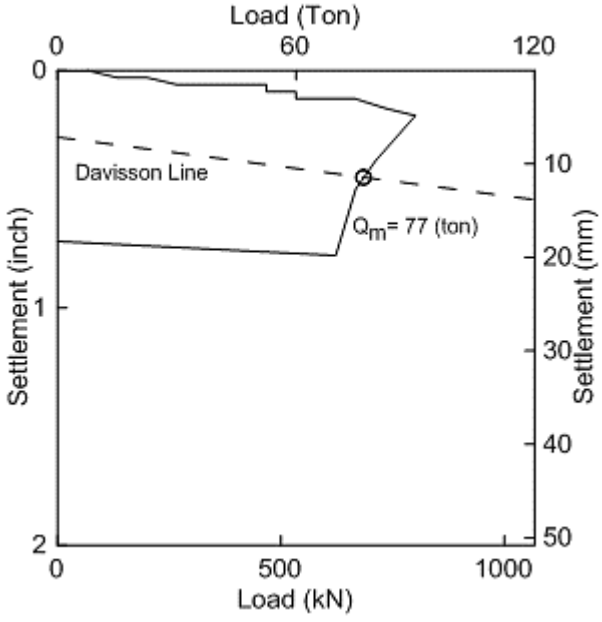


Figure 198. 49) 424-07-0021 TP#1

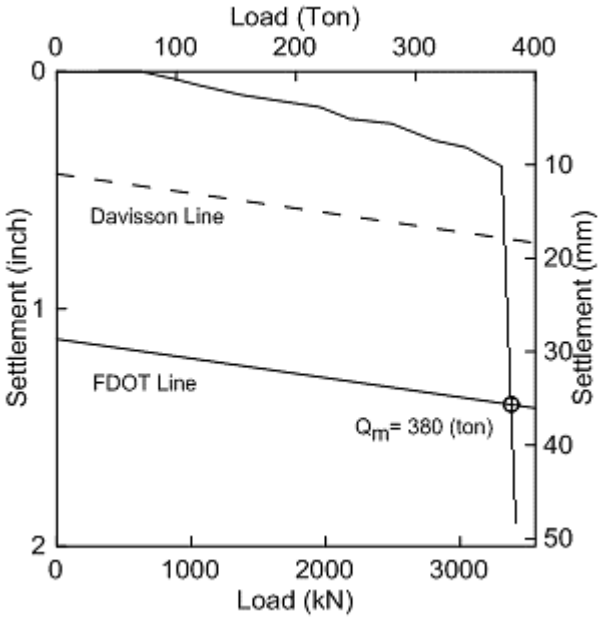


Figure 199. 50) 434-01-0002 TP#3

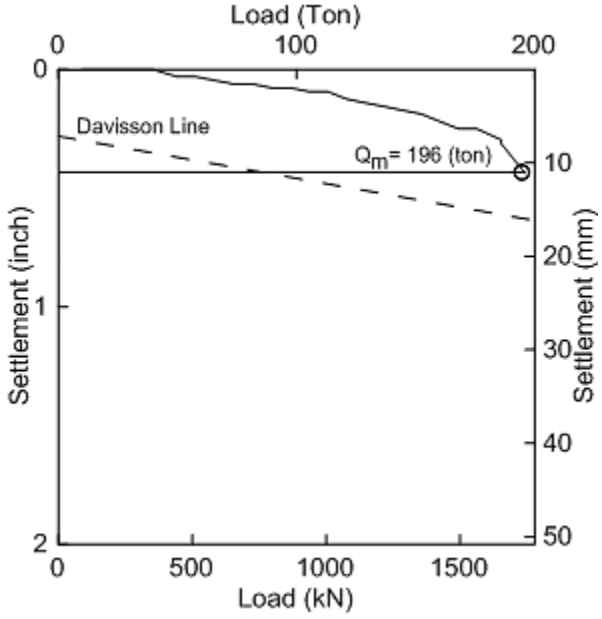


Figure 200. 51) 450-15-0085 TP#3A

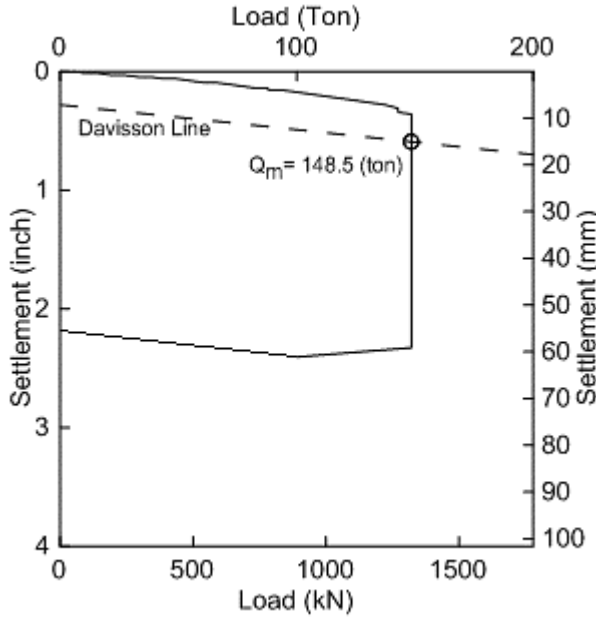


Figure 201. 52) 450-15-0100 TP#1

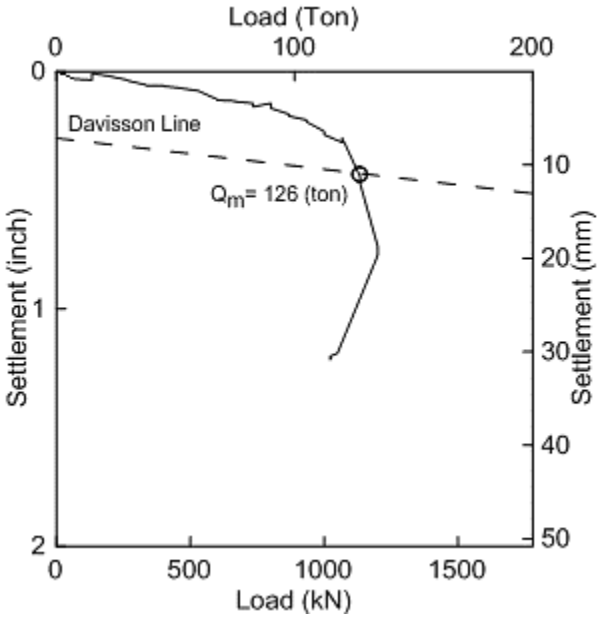


Figure 202. 53) 450-15-0100 TP#2

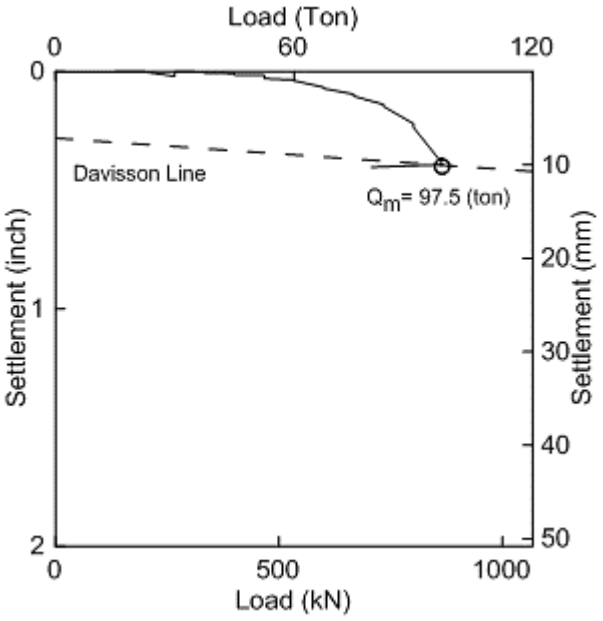


Figure 203. 54) 450-15-0100 TP#3

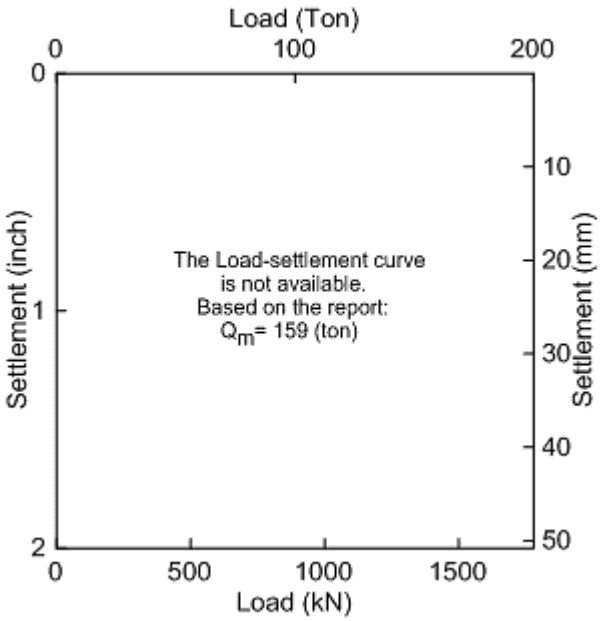


Figure 204. 55) 450-15-0103 TP#1

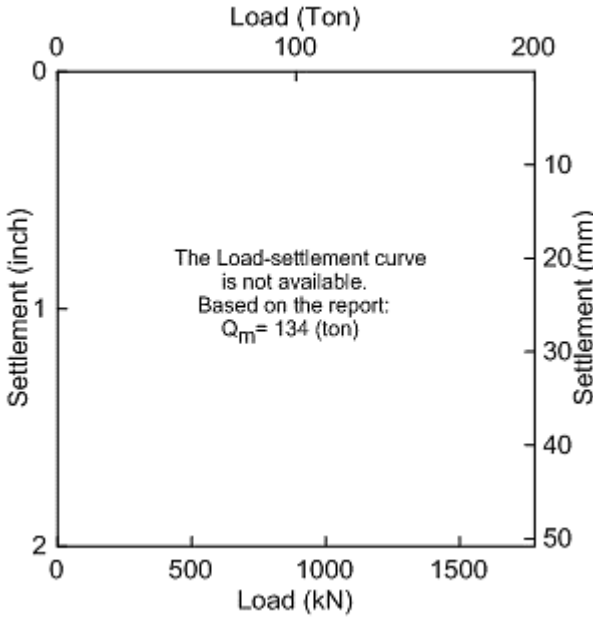


Figure 205. 56) 450-15-0103 TP#2

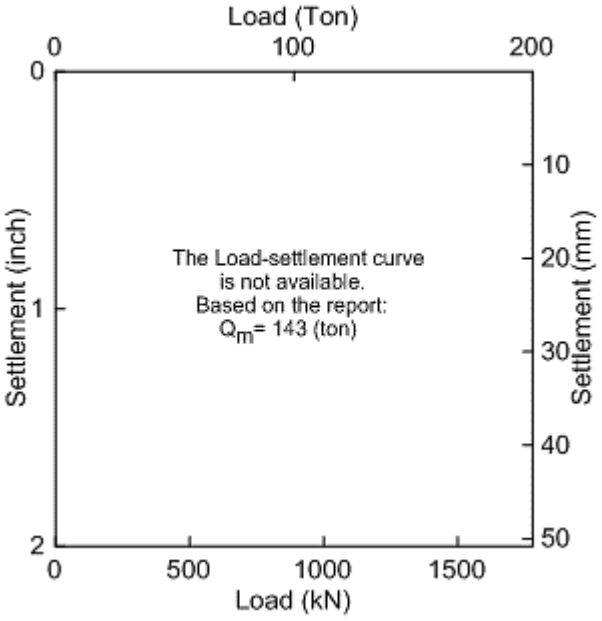


Figure 206. 57) 450-15-0103 TP#5

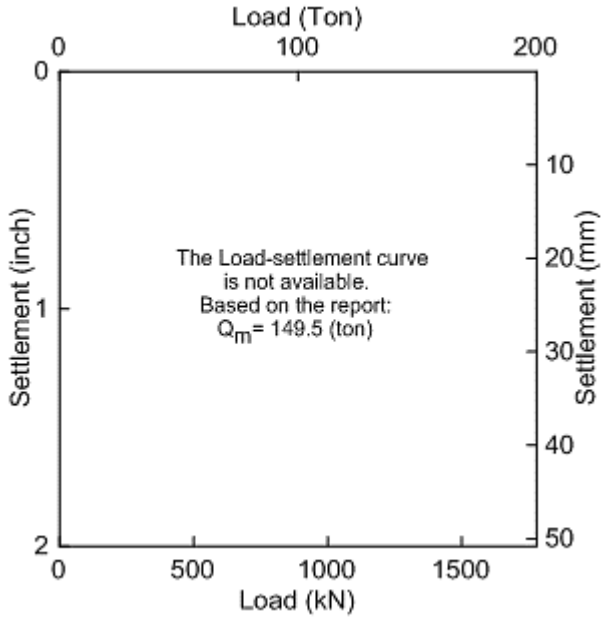


Figure 207. 58) 450-15-0103 TP#7

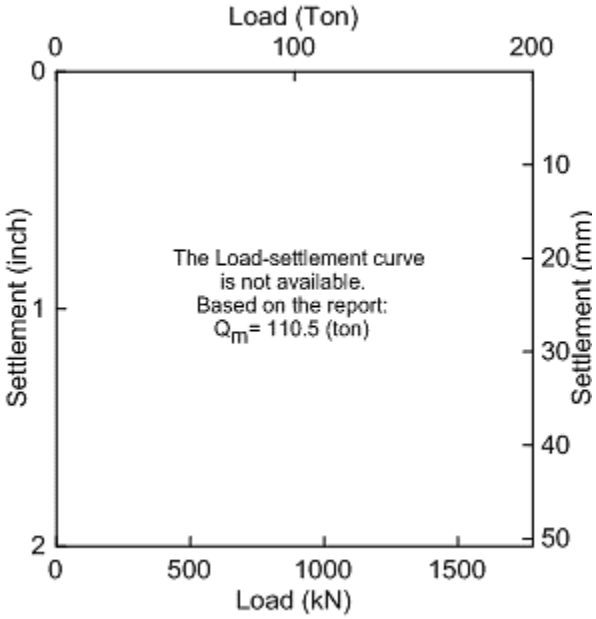


Figure 208. 59) 450-36-0002 TP#8

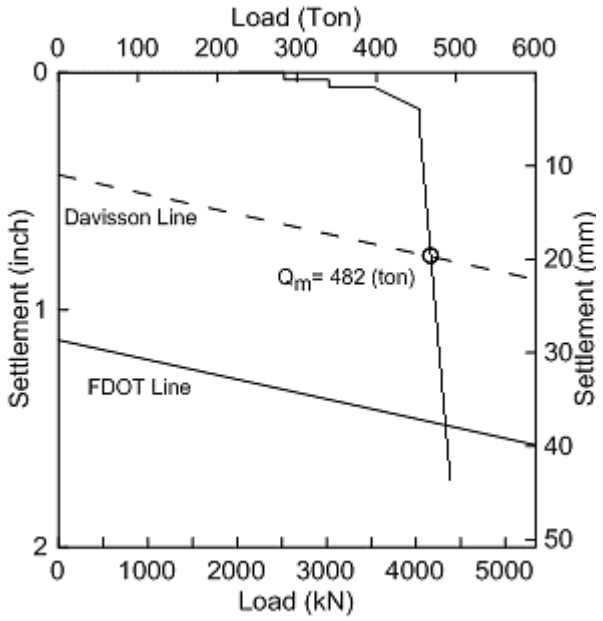


Figure 209. 60) 455-05-0036 TP#1

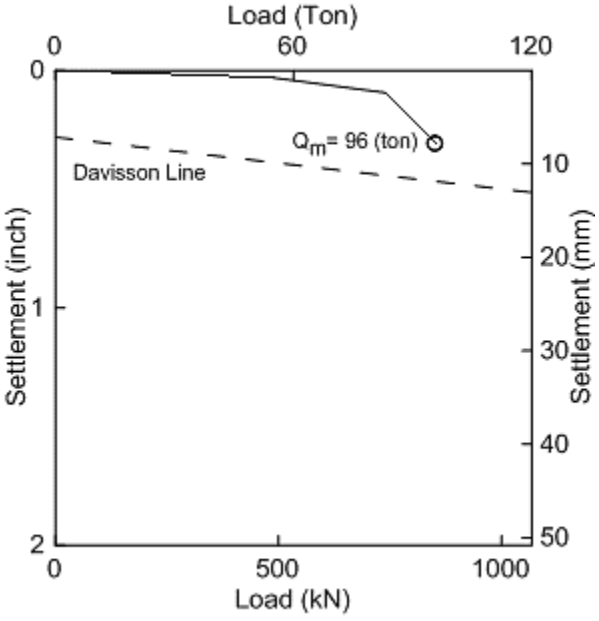


Figure 210. 61) 455-05-0036 TP#2

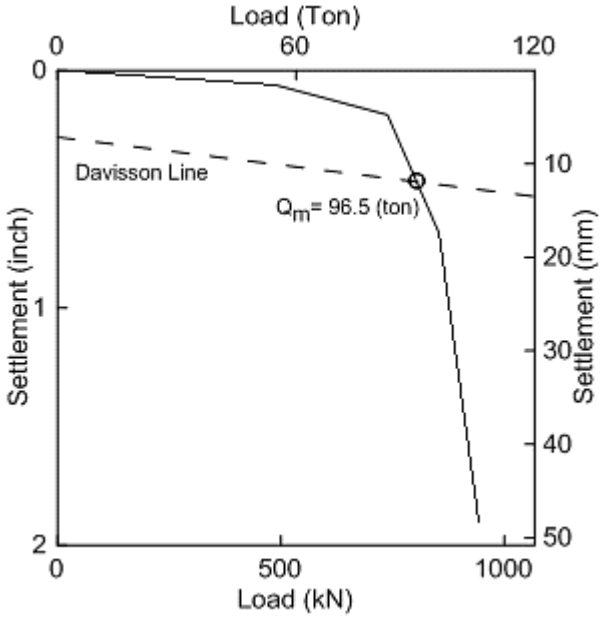


Figure 211. 62) 455-05-0036 TP#3

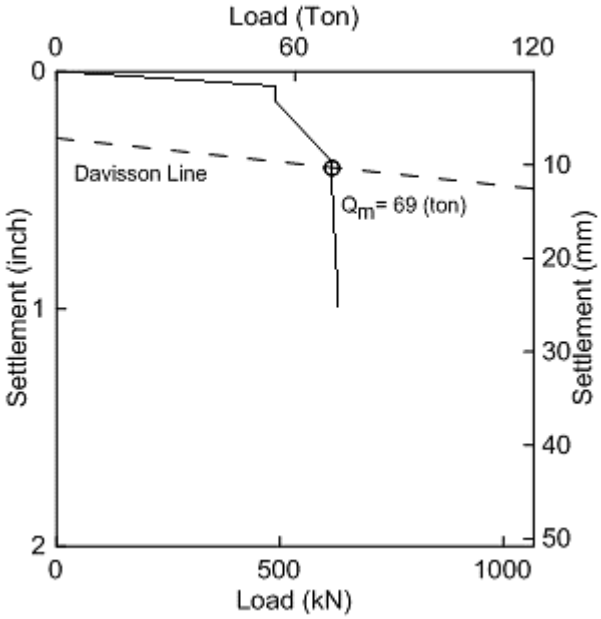


Figure 212. 63) 713-48-0083 TP#1

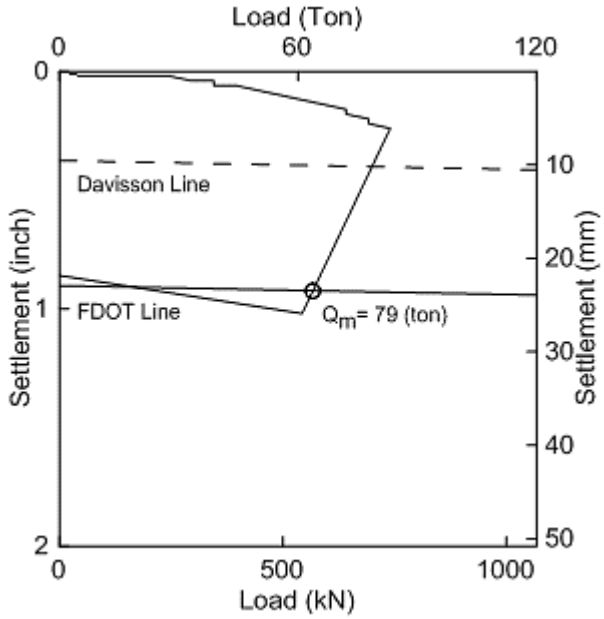


Figure 213. 64) 713-48-0083 TP#2

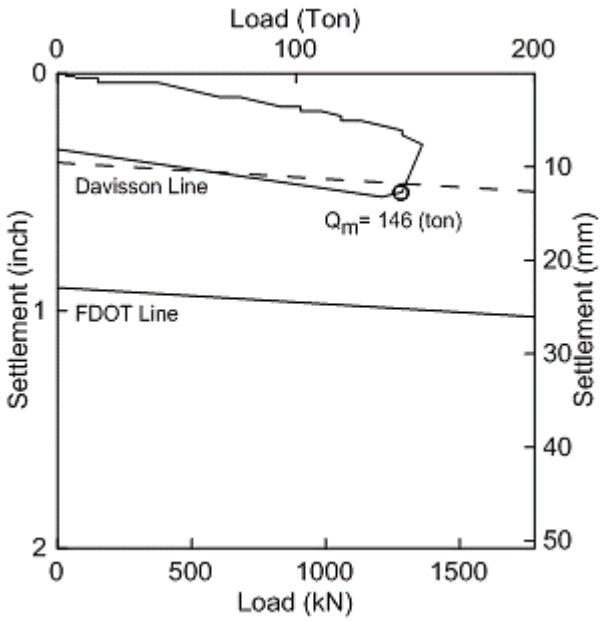


Figure 214. 65) 742-06-0073 TP#2

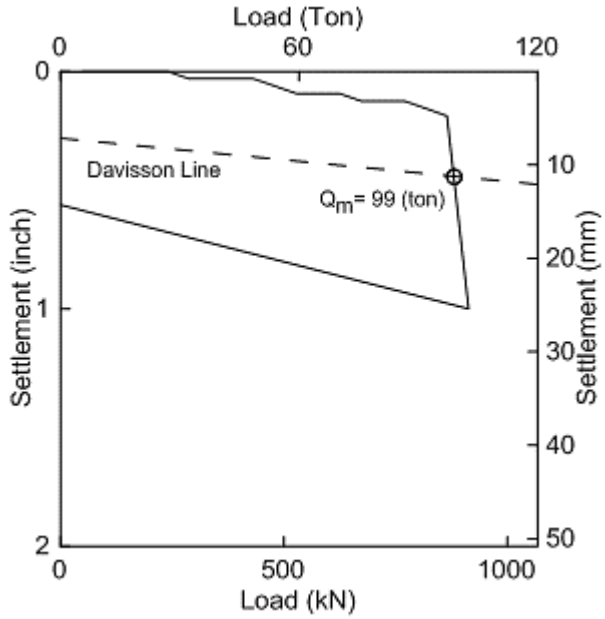


Figure 215. 66) 829-10-0013 TP#1

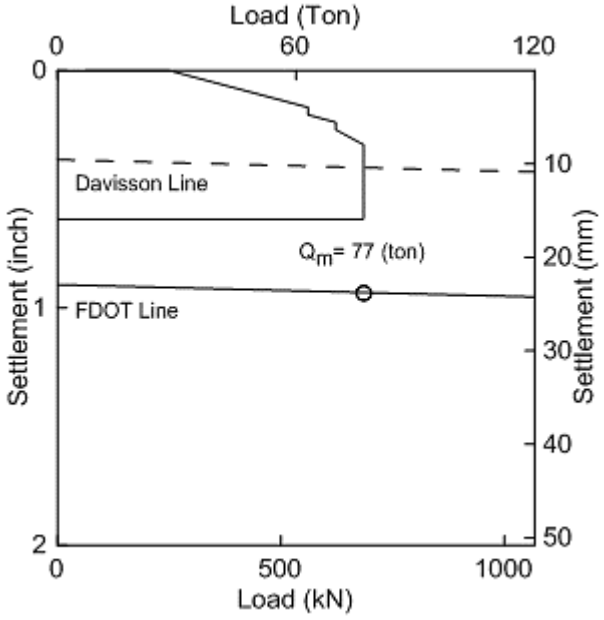


Figure 216. 67) 855-14-0013 TP#1

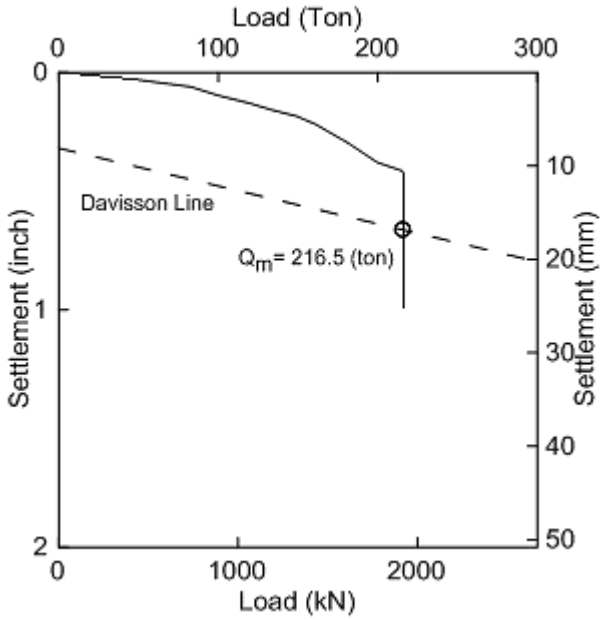


Figure 217. 68) Bayou Beouf TP#3

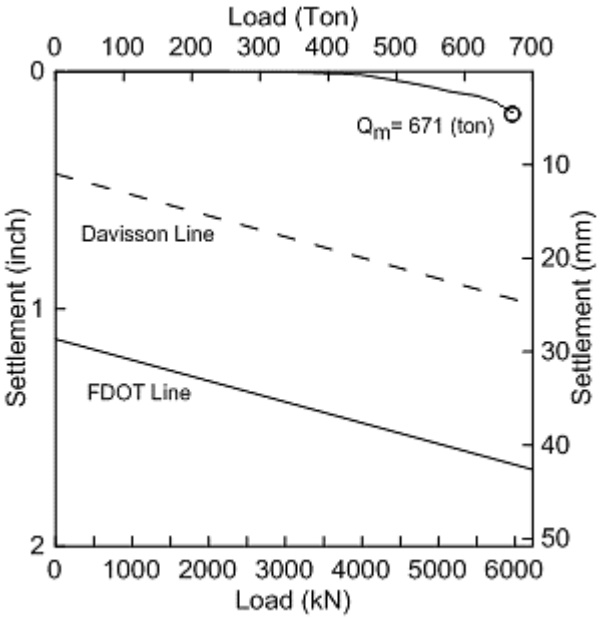


Figure 218. 69) Bayou Lacassine TP#1

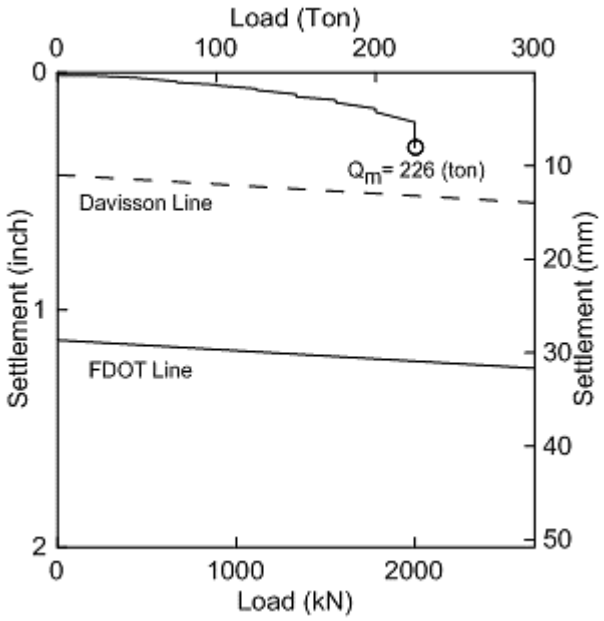


Figure 219. 70) Bayou Lacassine TP#3

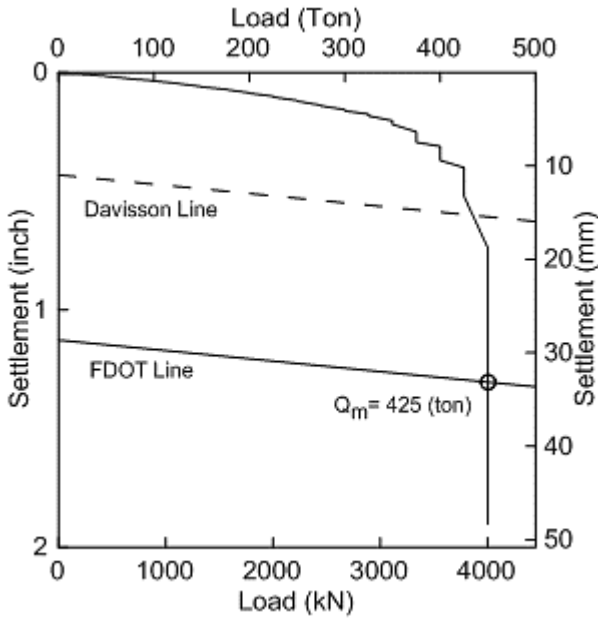


Figure 220. 71) Bayou Zourrie TP#1

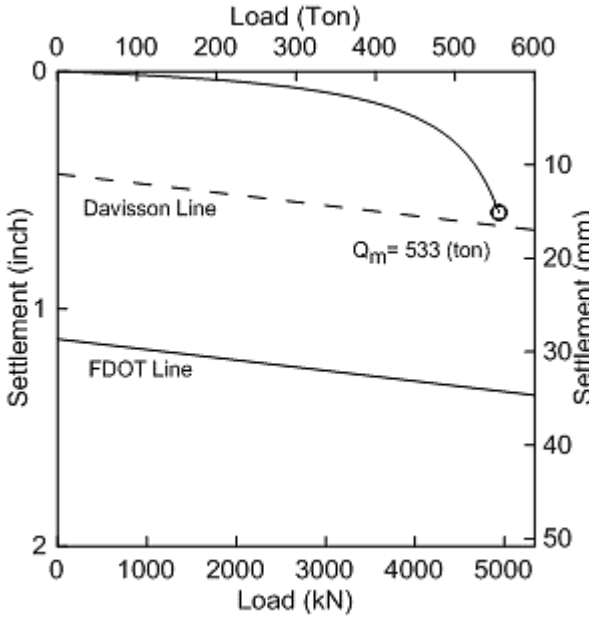


Figure 221. 72) LA-01 TP#2

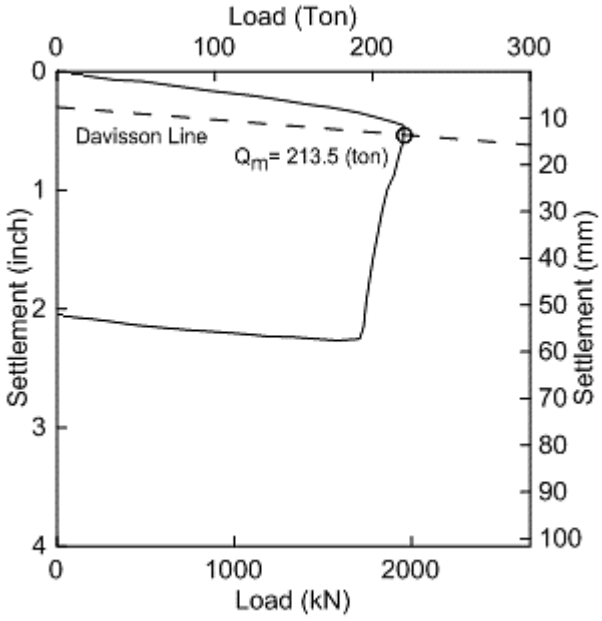


Figure 222. 73) LA-01 TP#3

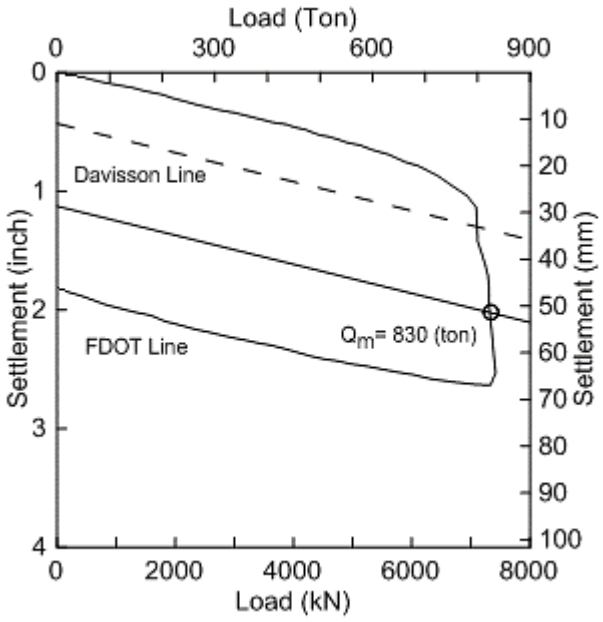


Figure 223. 74) LA-01 TP#4A

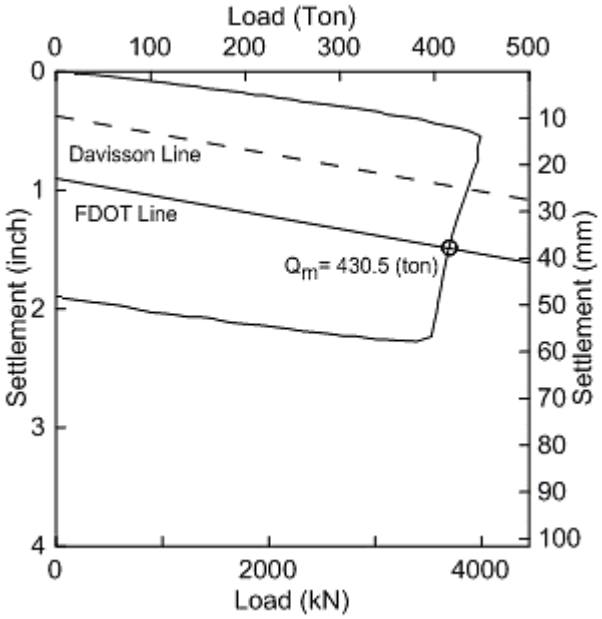


Figure 224. 75) LA-01 TP#4B

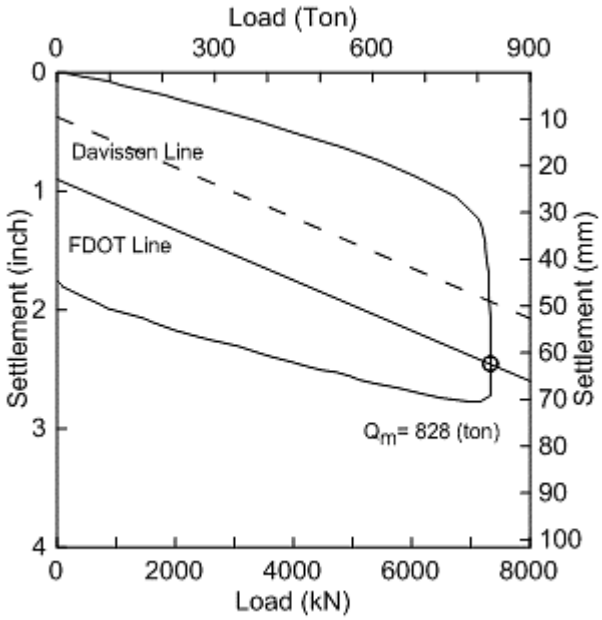


Figure 225. 76) LA-01 TP#5A

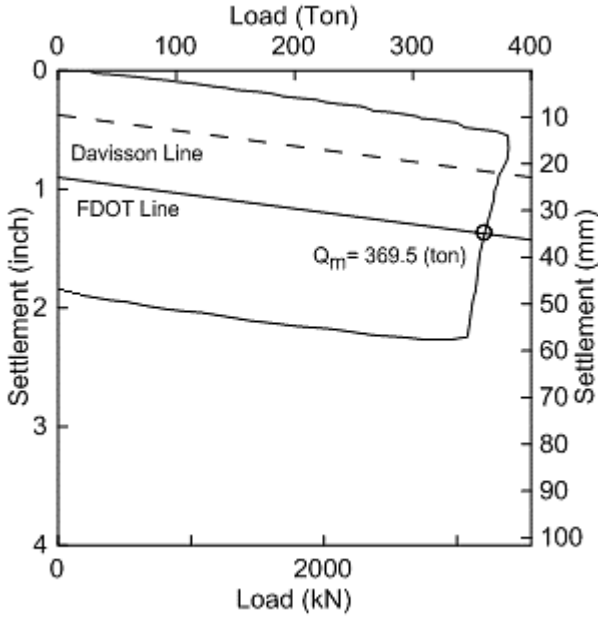


Figure 226. 77) LA-01 TP#5B

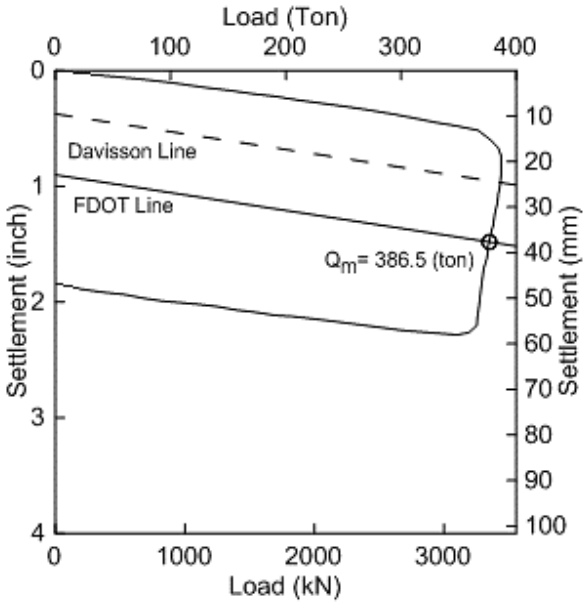


Figure 227. 78) 008-01-0042 TP#1

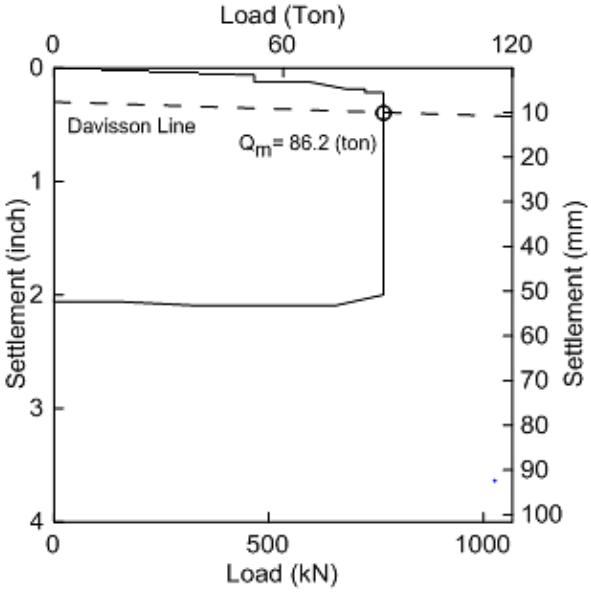


Figure 228. 79) 450-17-0025 TP#1

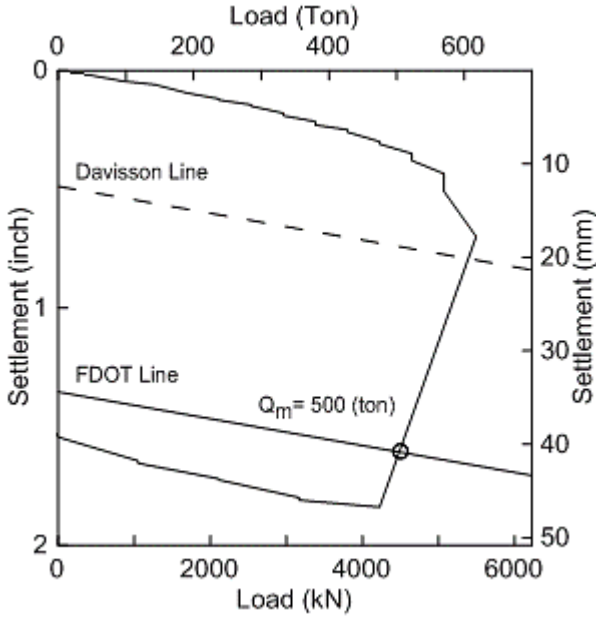
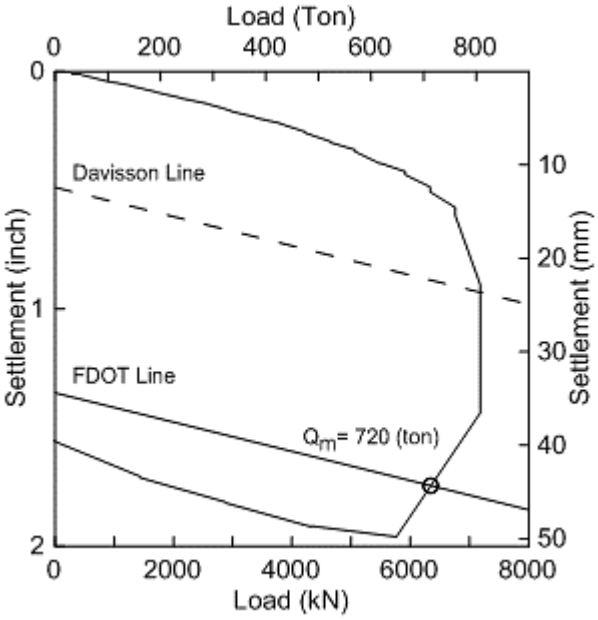


Figure 229. 80) 450-17-0025 TP#3



Appendix E

Comparison of Measured and Estimated Pile Capacities by Different Pile-CPT Methods

Figure 230. Comparison of measured and ultimate pile capacity predicted by LCPC method (piles 1-40)

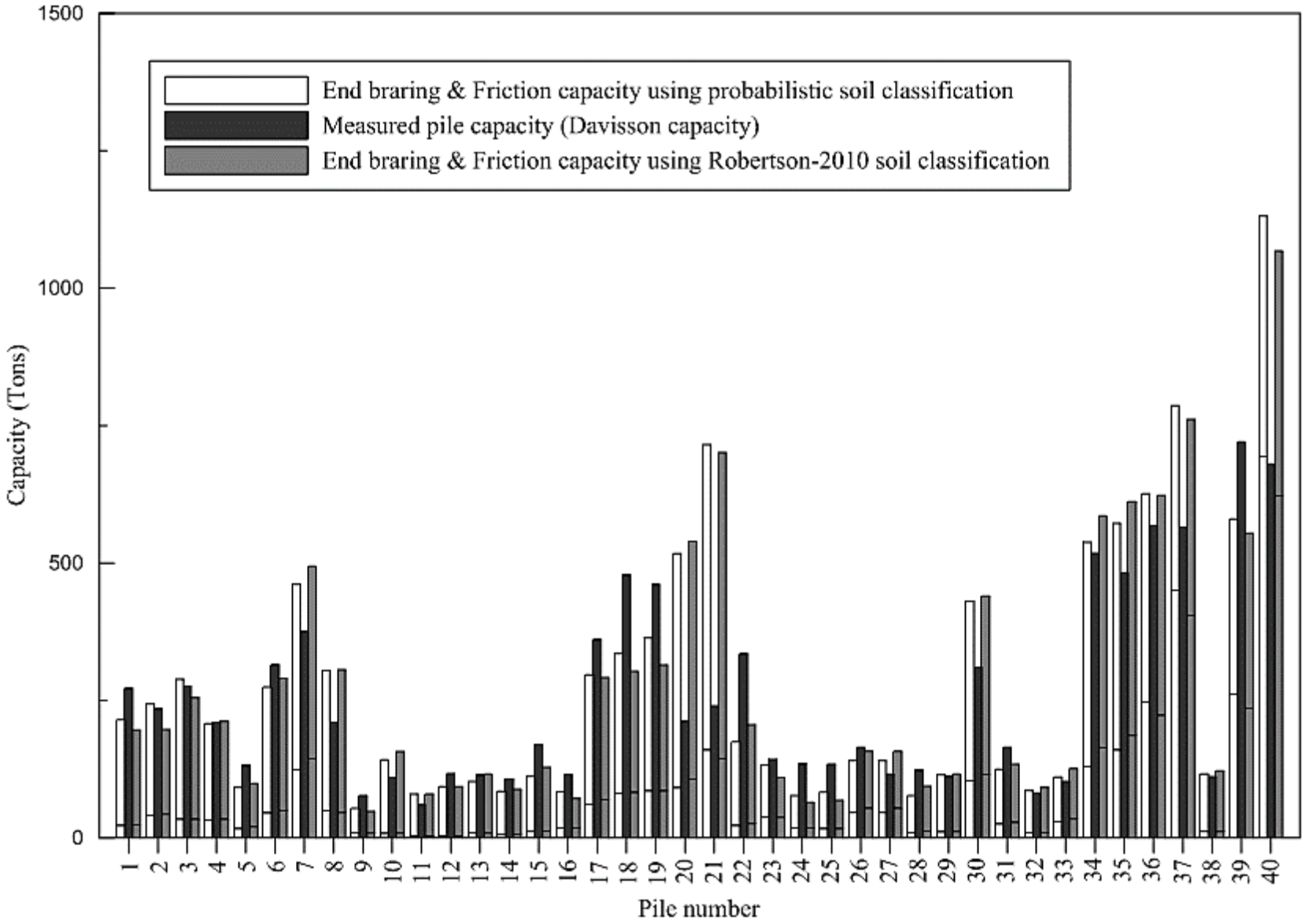


Figure 231. Comparison of measured and ultimate pile capacity predicted by LCPC method (piles 41-80)

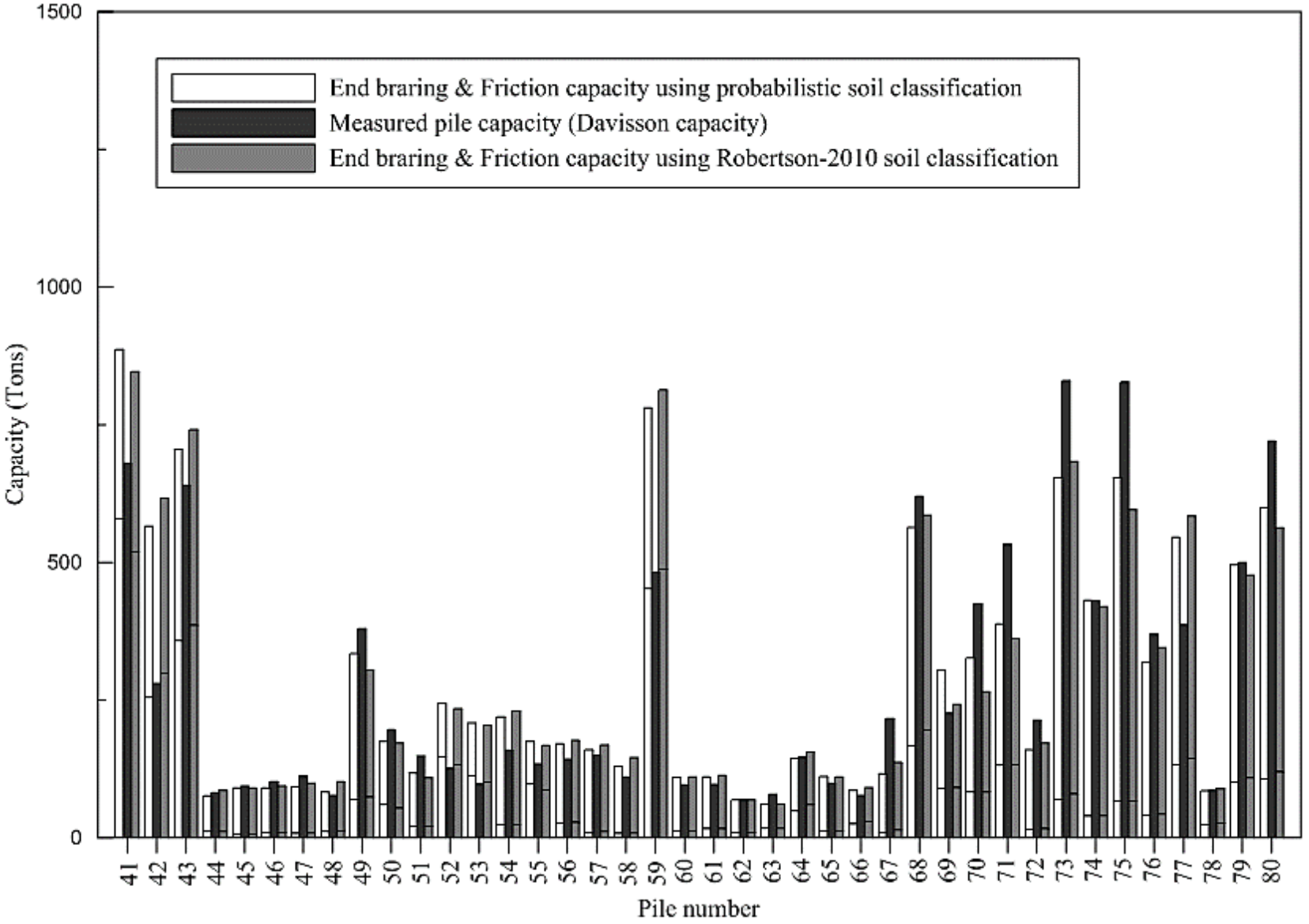


Figure 232. Comparison of measured and ultimate pile capacity predicted by Schmertmann method (piles 1-40)

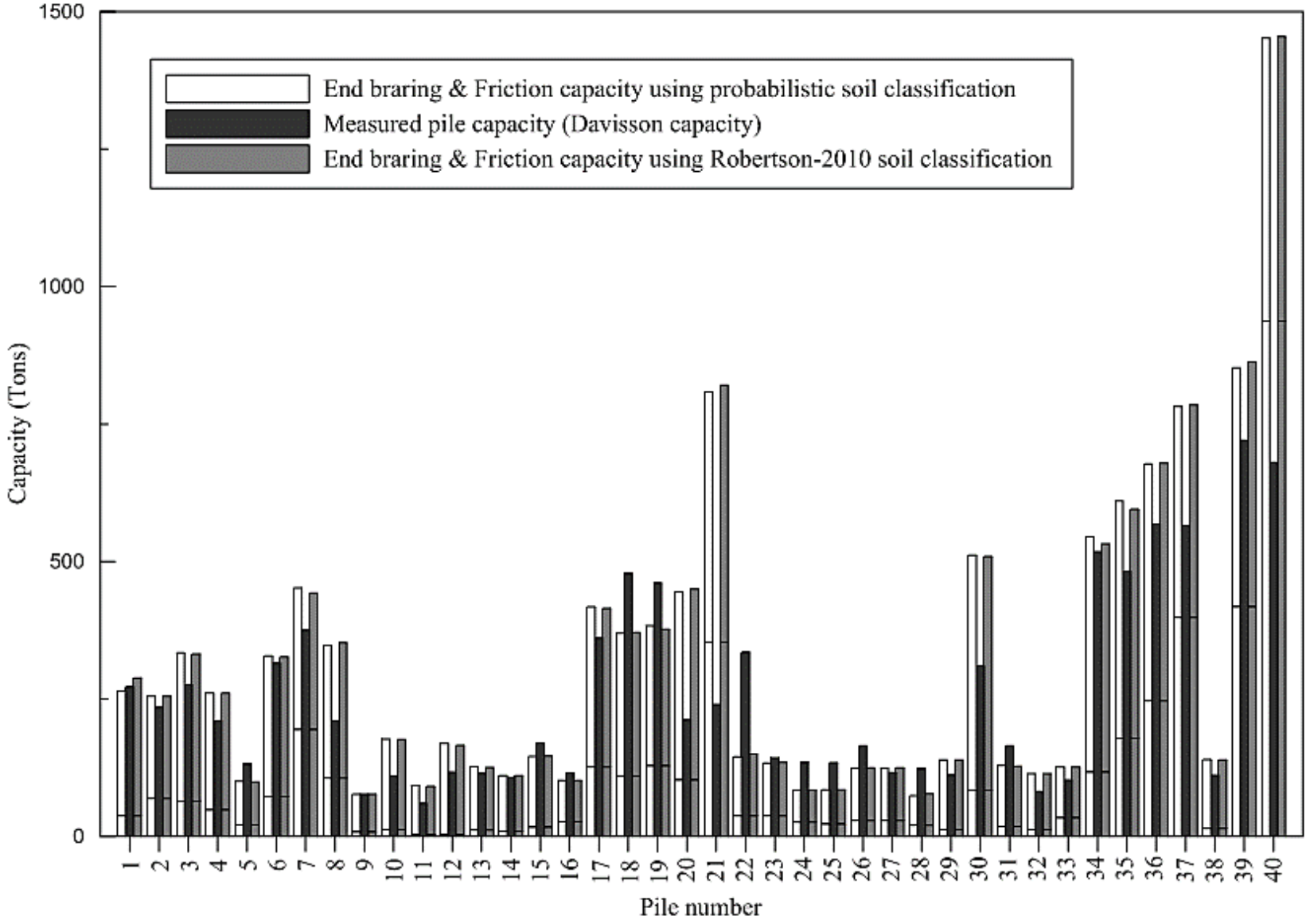


Figure 233. Comparison of measured and ultimate pile capacity predicted by Schmertmann method (piles 41-80)

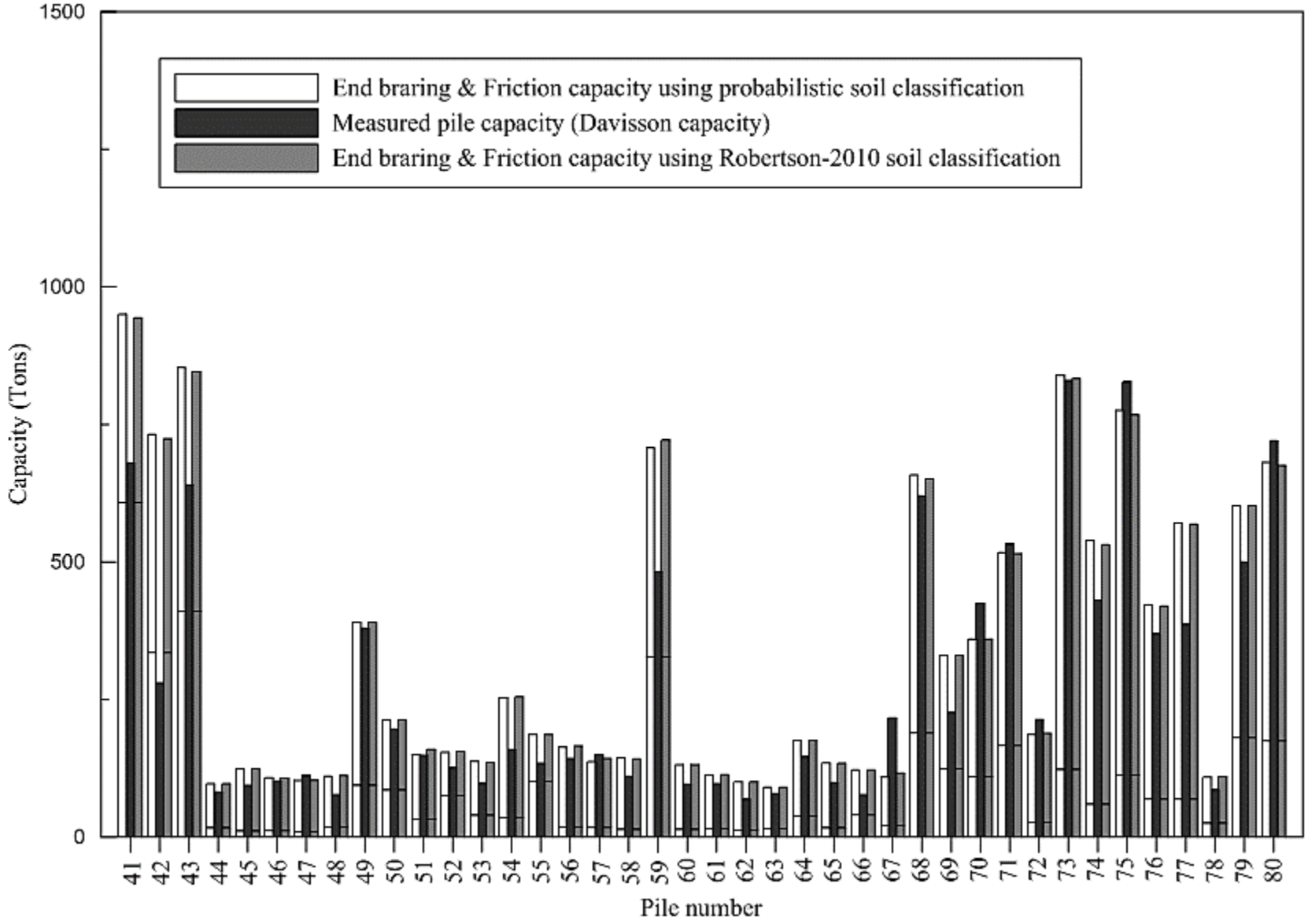


Figure 234. Comparison of measured and ultimate pile capacity predicted by De Ruiter method (piles 1-40)

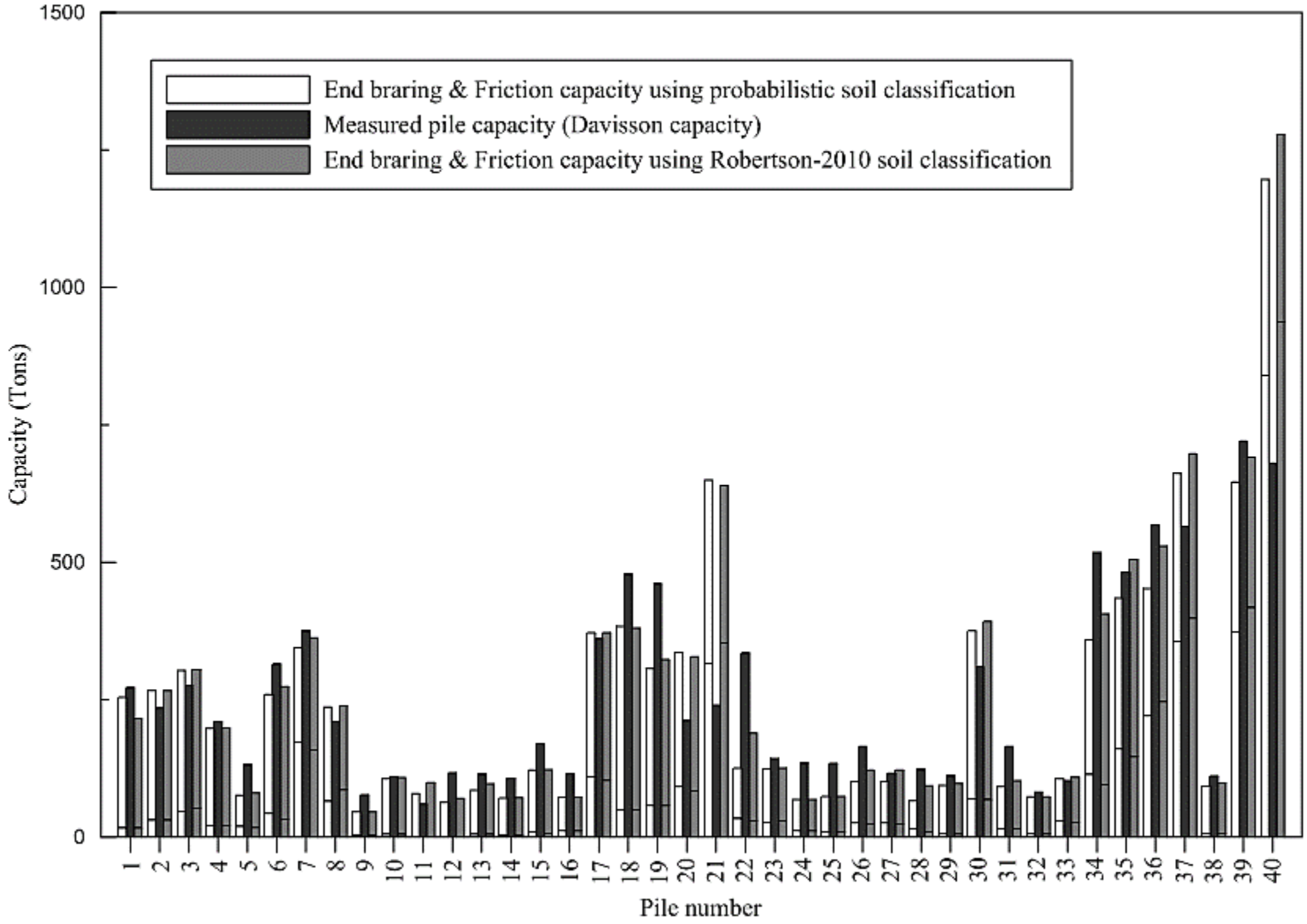


Figure 235. Comparison of measured and ultimate pile capacity predicted by De Ruiter method (piles 41-80)

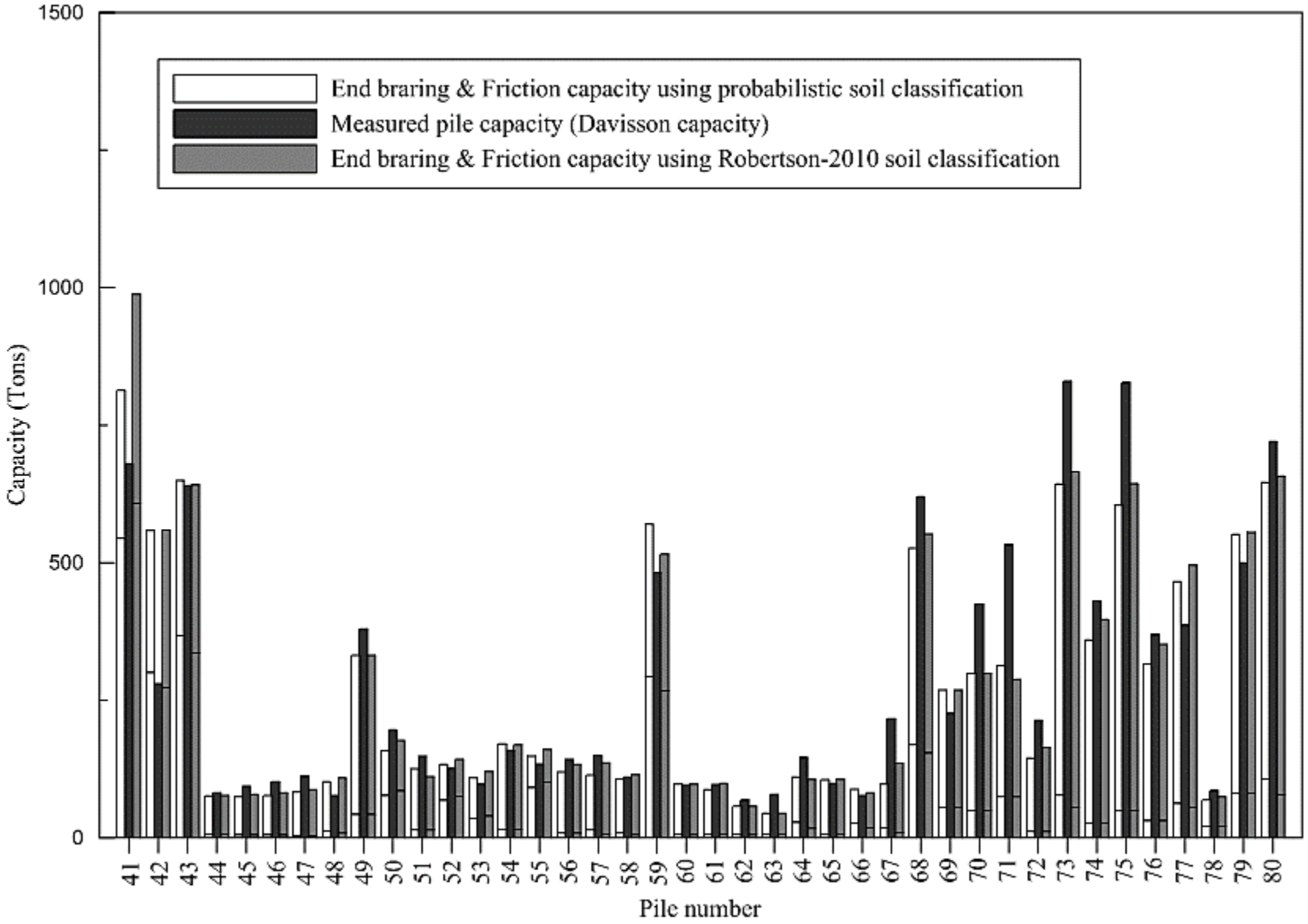


Figure 236. Comparison of measured and ultimate pile capacity predicted by Philipponnat method (piles 1-40)

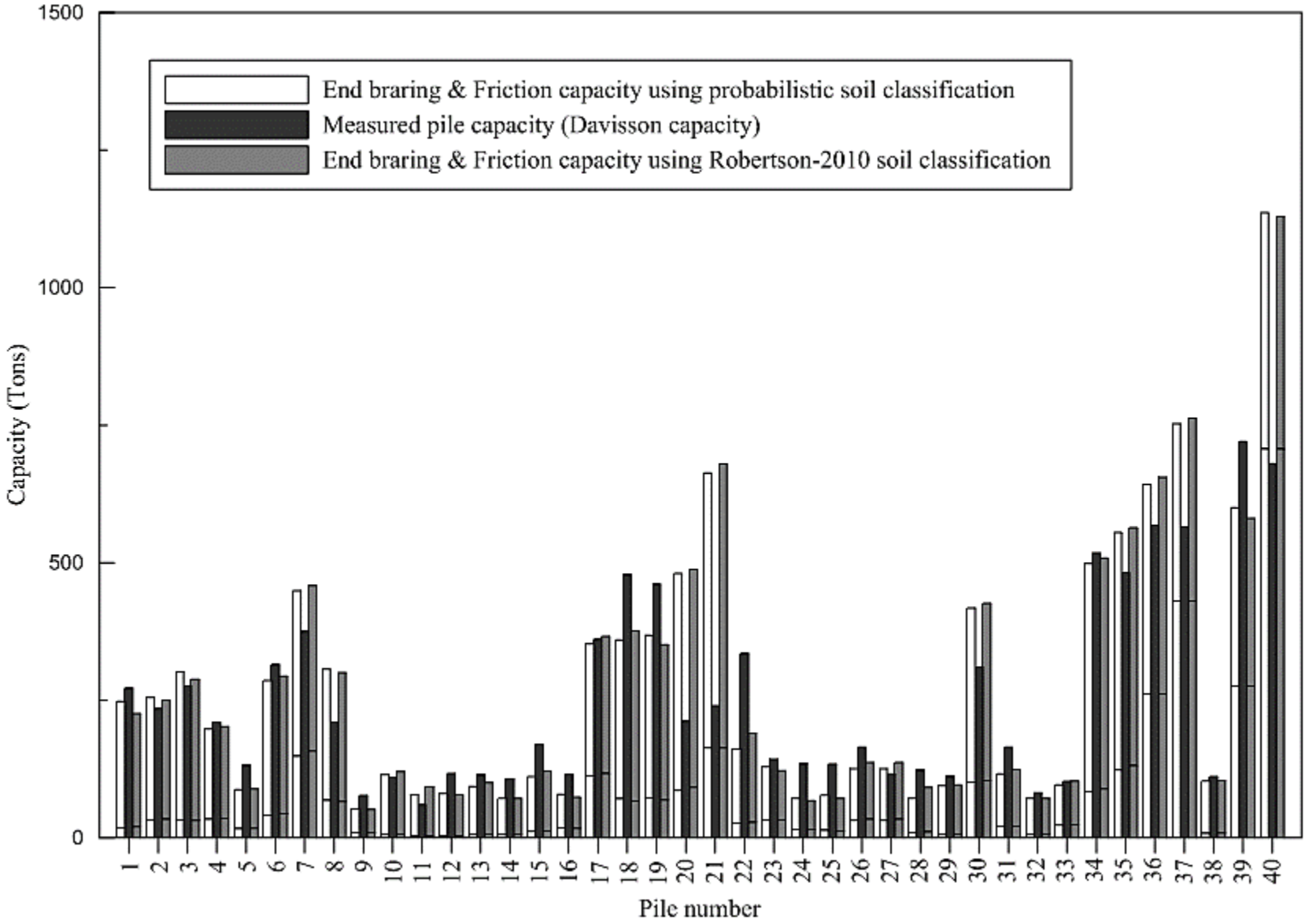


Figure 237. Comparison of measured and ultimate pile capacity predicted by Philipponnat method (piles 41-80)

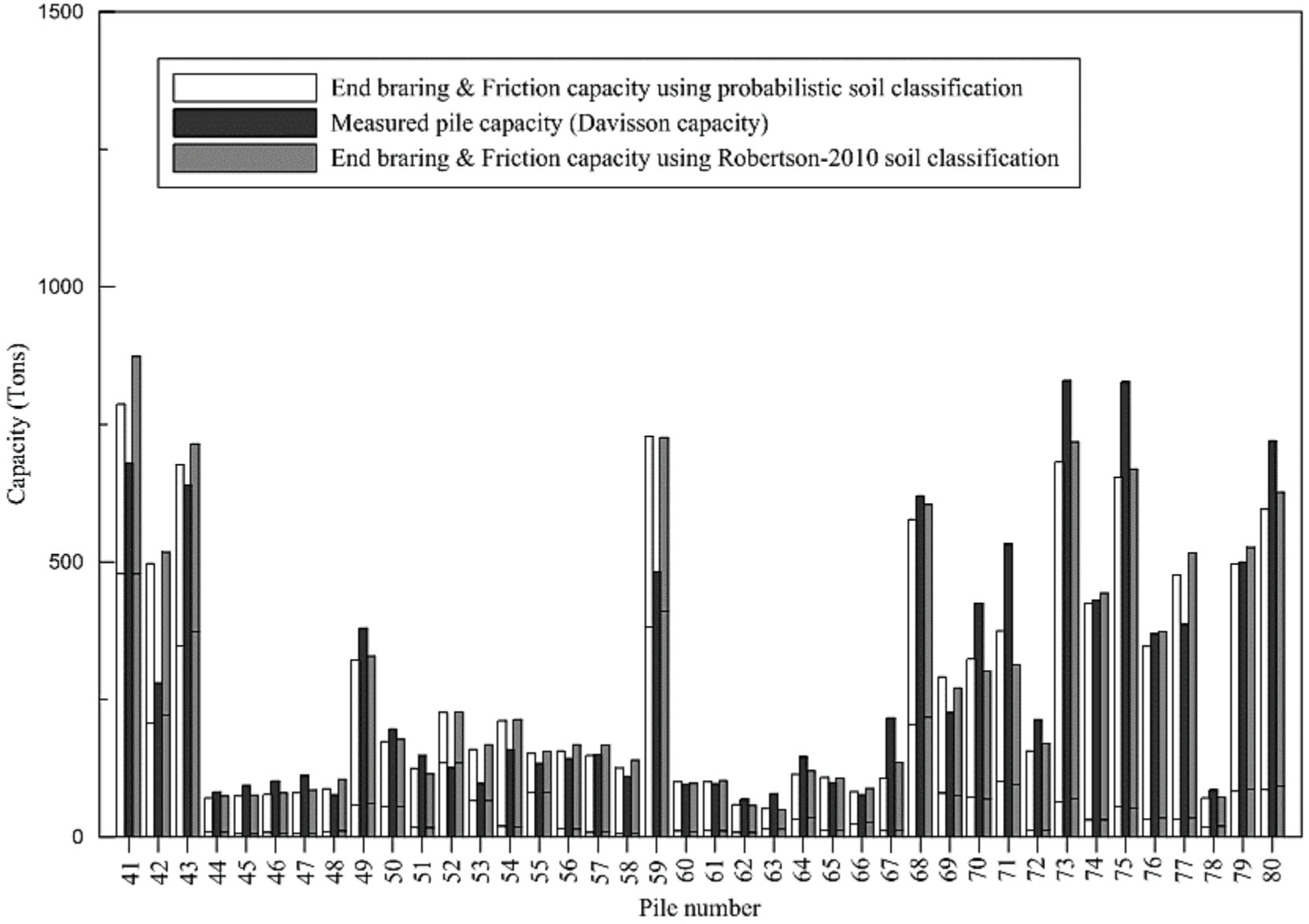


Figure 238. Comparison of measured and ultimate pile capacity predicted by Price and Wardle method (piles 1-40)

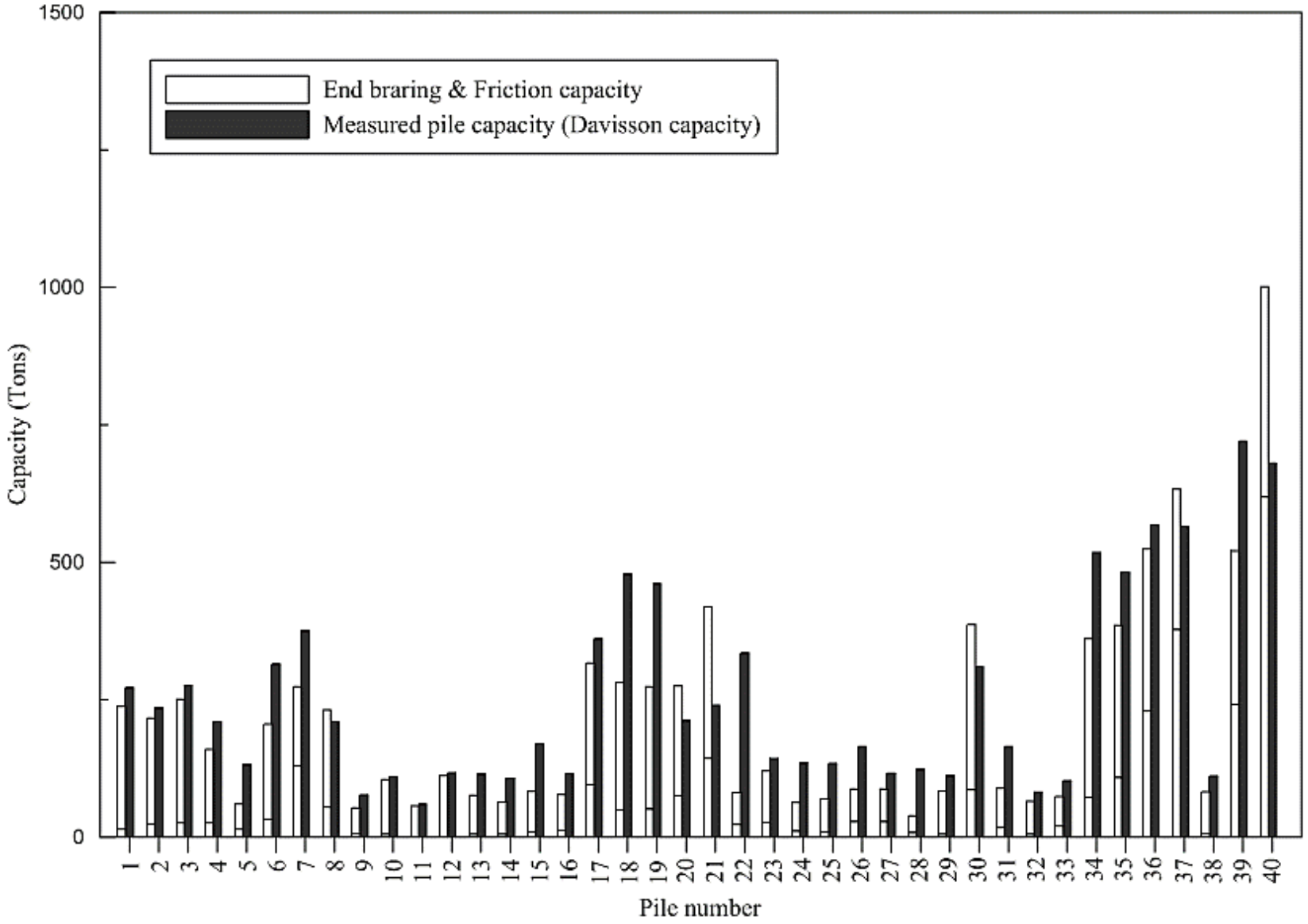


Figure 239. Comparison of measured and ultimate pile capacity predicted by Price and Wardle method (piles 41-80)

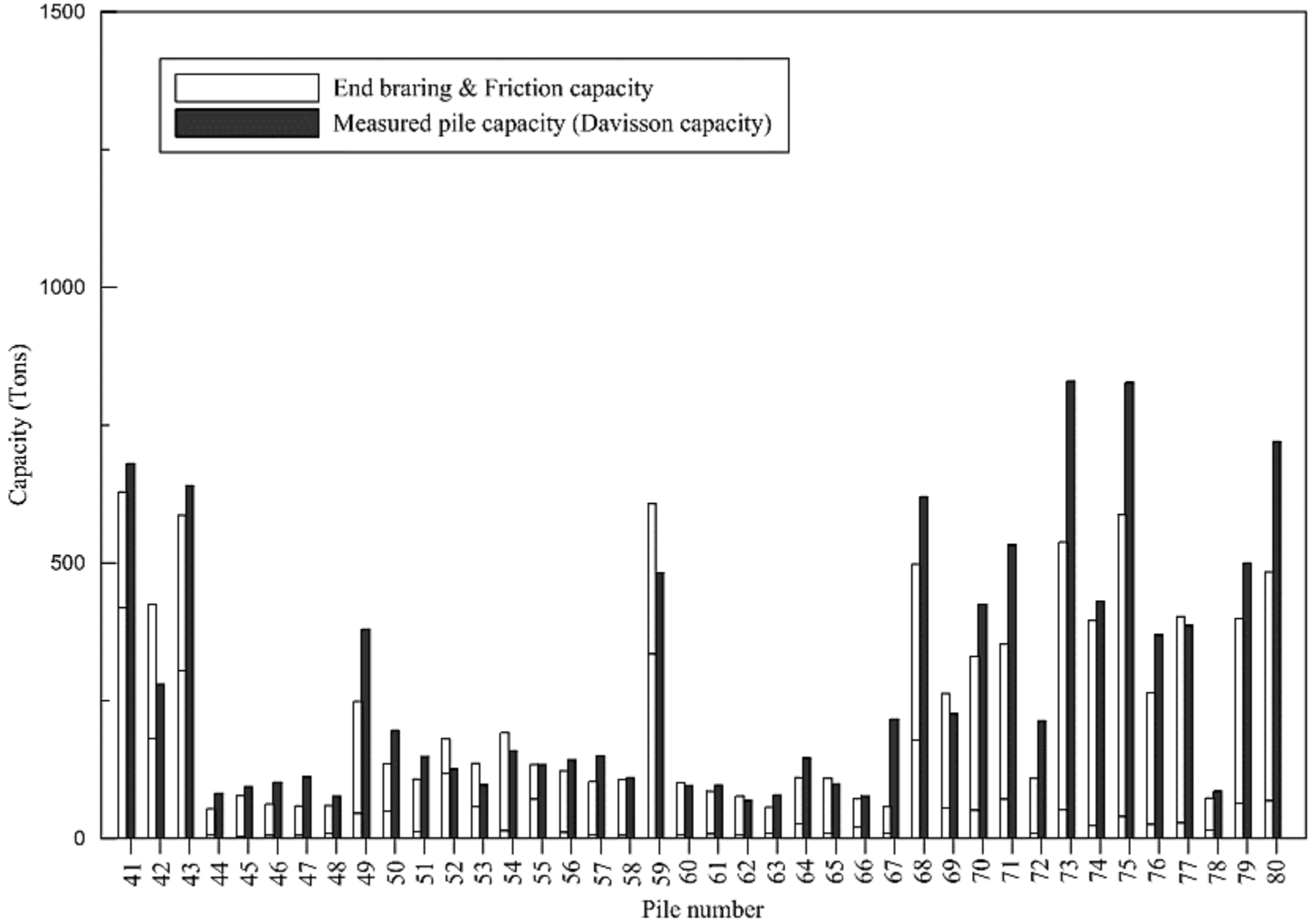


Figure 240. Comparison of measured and ultimate pile capacity predicted by Zhou method (piles 1-40)

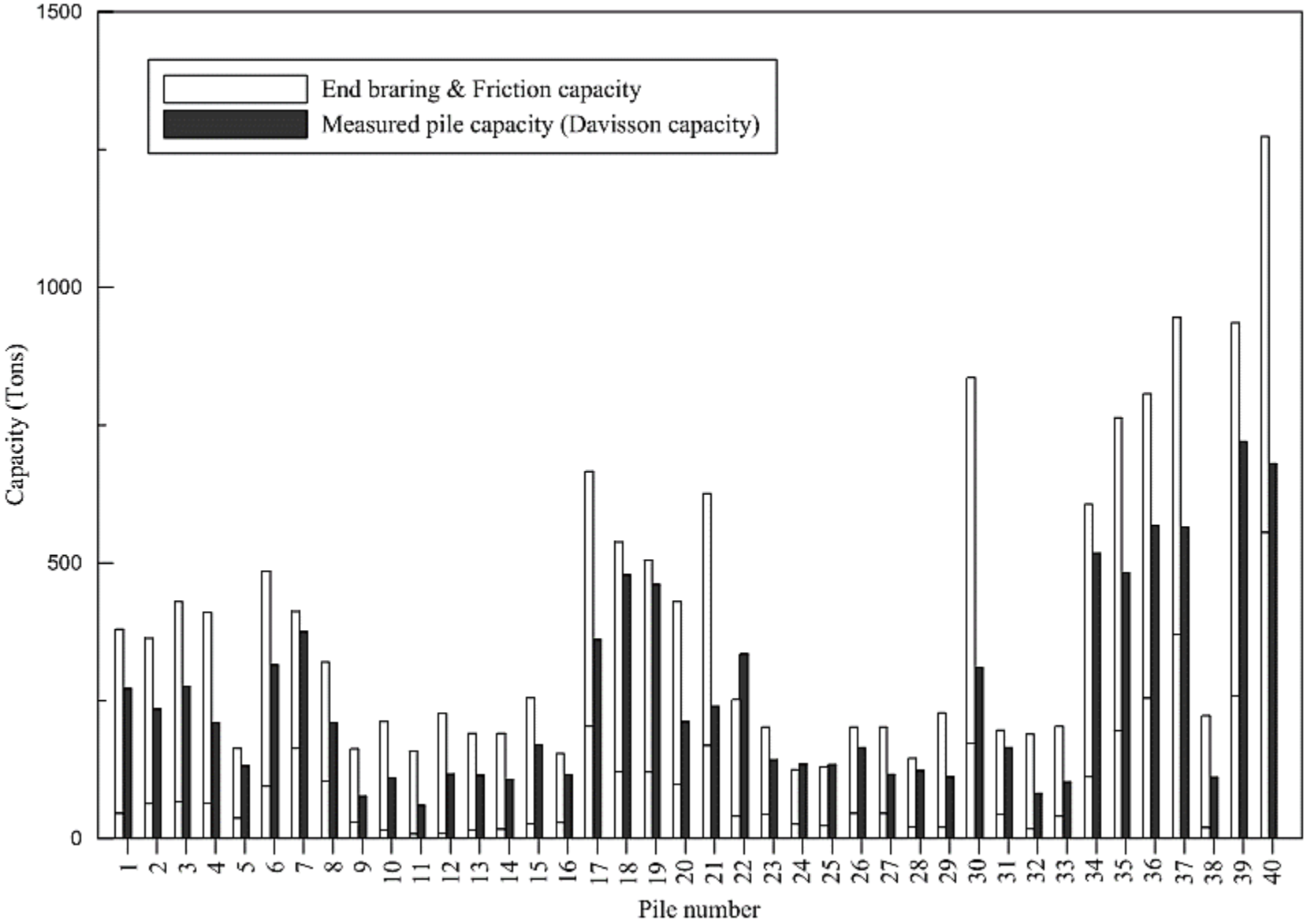


Figure 241. Comparison of measured and ultimate pile capacity predicted by Zhou method (piles 41-80)

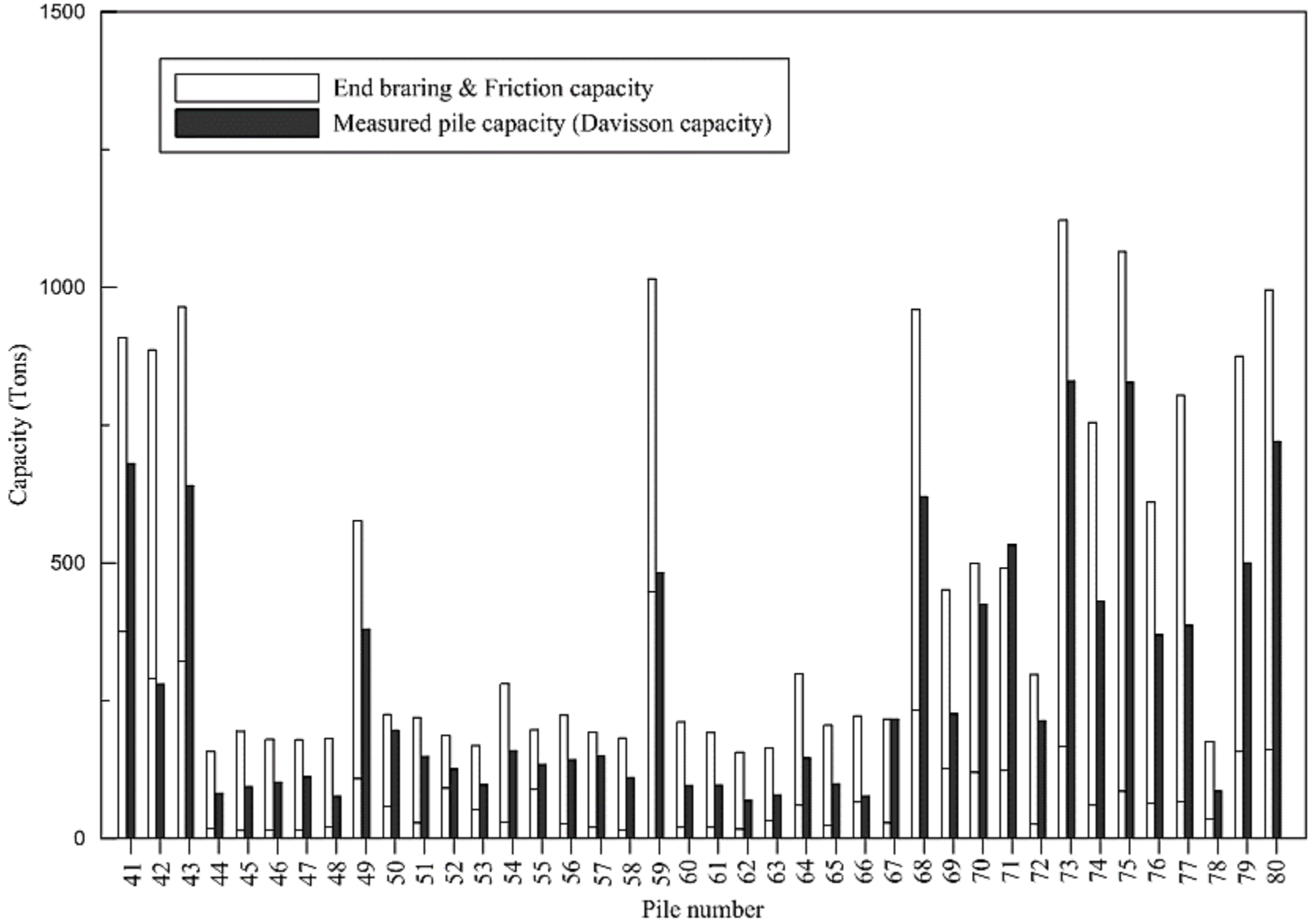


Figure 242. Comparison of measured and ultimate pile capacity predicted by Tumay and Fakhroo method (piles 1-40)

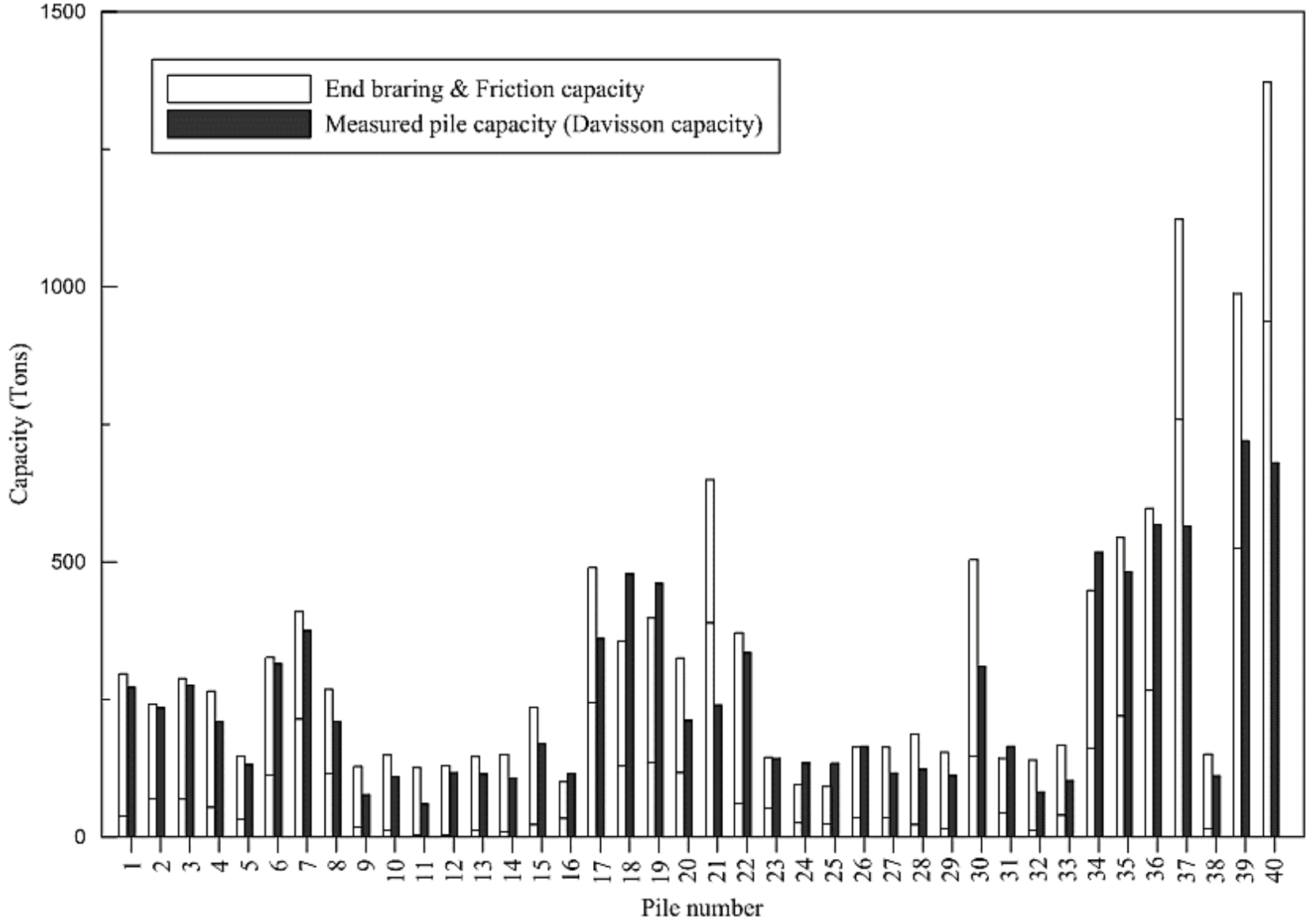


Figure 243. Comparison of measured and ultimate pile capacity predicted by Tumay and Fakhroo method (piles 41-80)

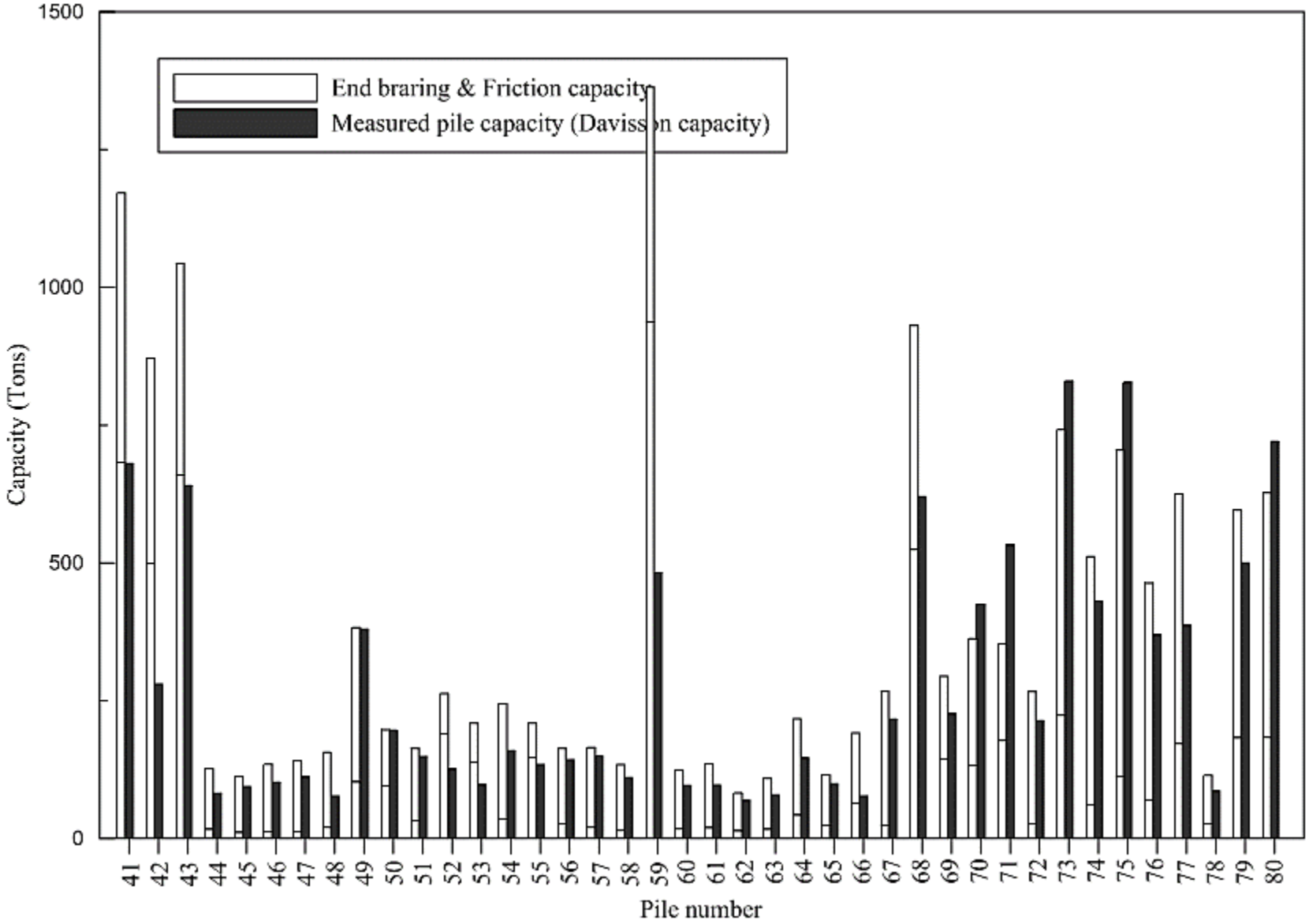


Figure 244. Comparison of measured and ultimate pile capacity predicted by UF method (piles 1-40)

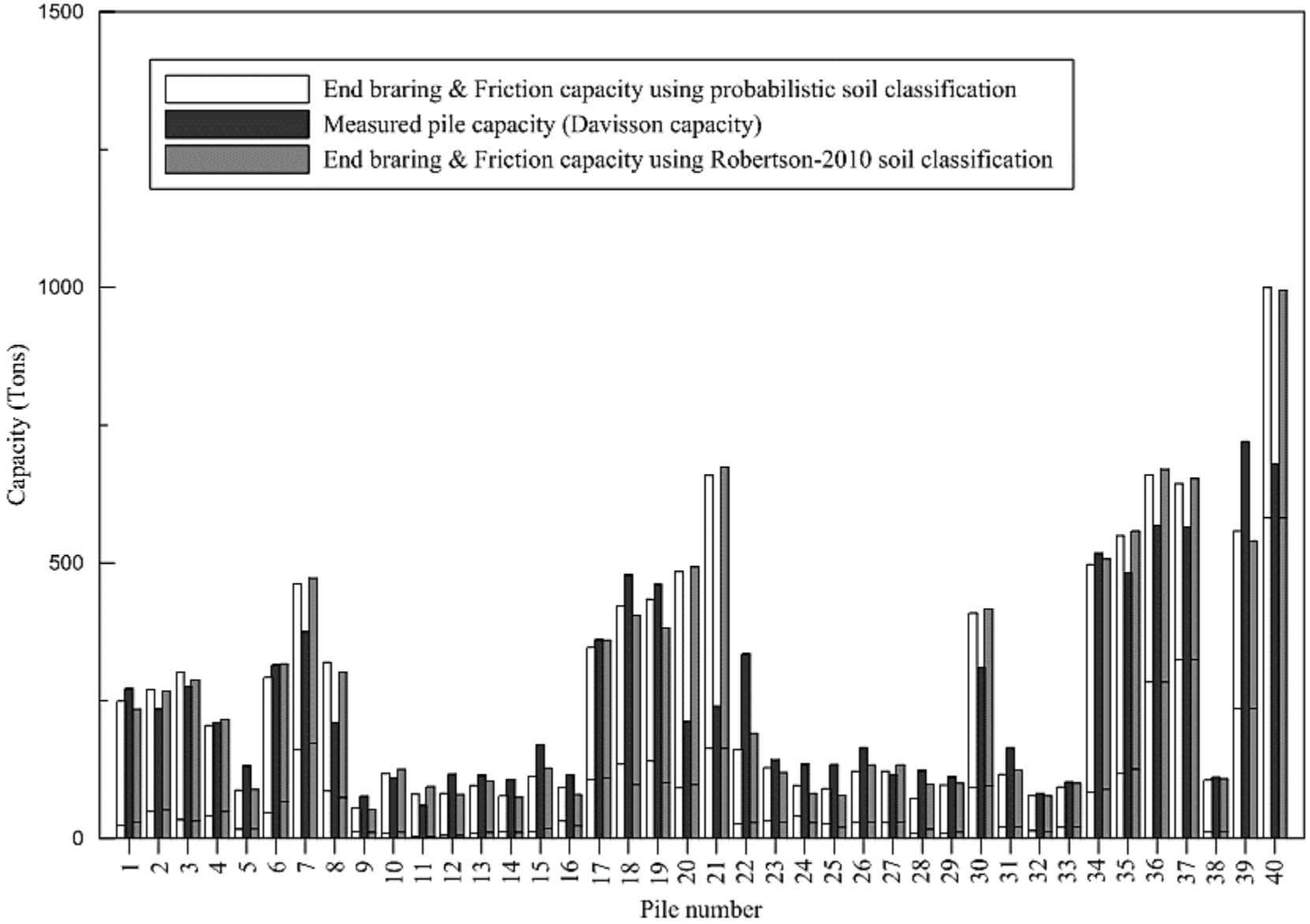


Figure 245. Comparison of measured and ultimate pile capacity predicted by UF method (piles 41-80)

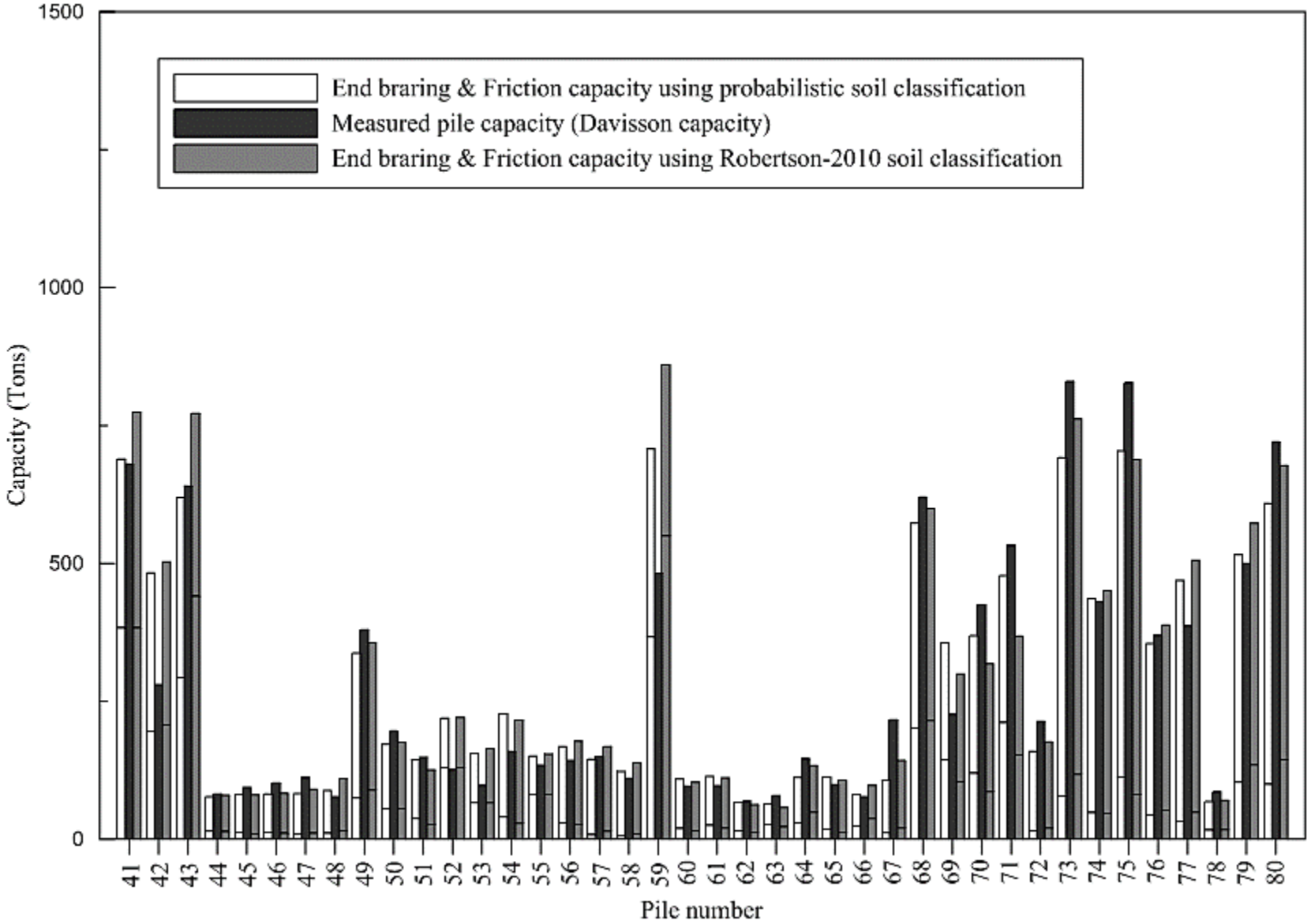


Figure 246. Comparison of measured and ultimate pile capacity predicted by Probabilistic method (piles 1-40)

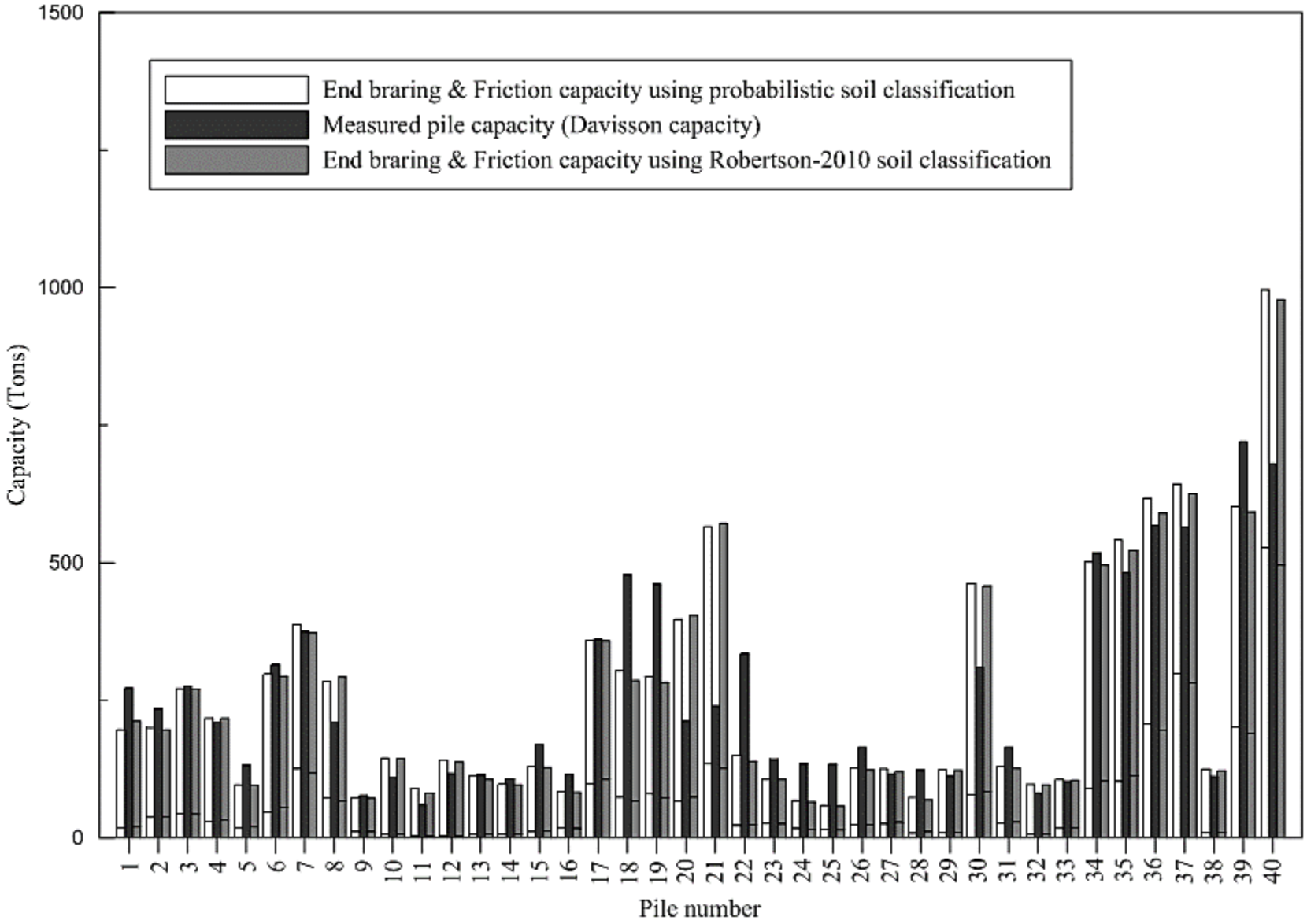


Figure 247. Comparison of measured and ultimate pile capacity predicted by Probabilistic method (piles 41-80)

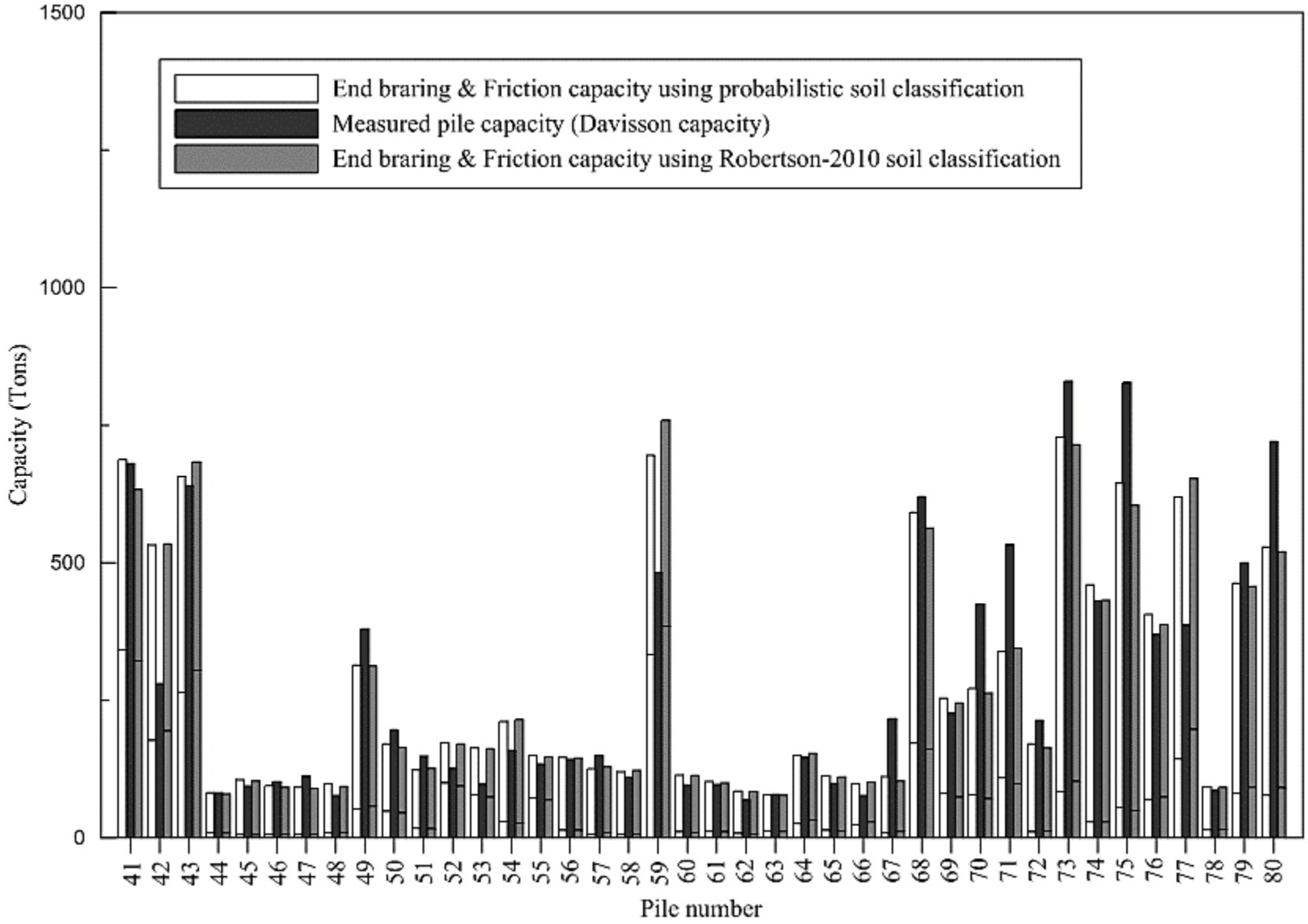


Figure 248. Comparison of measured and ultimate pile capacity predicted by Aoki method (piles 1-40)

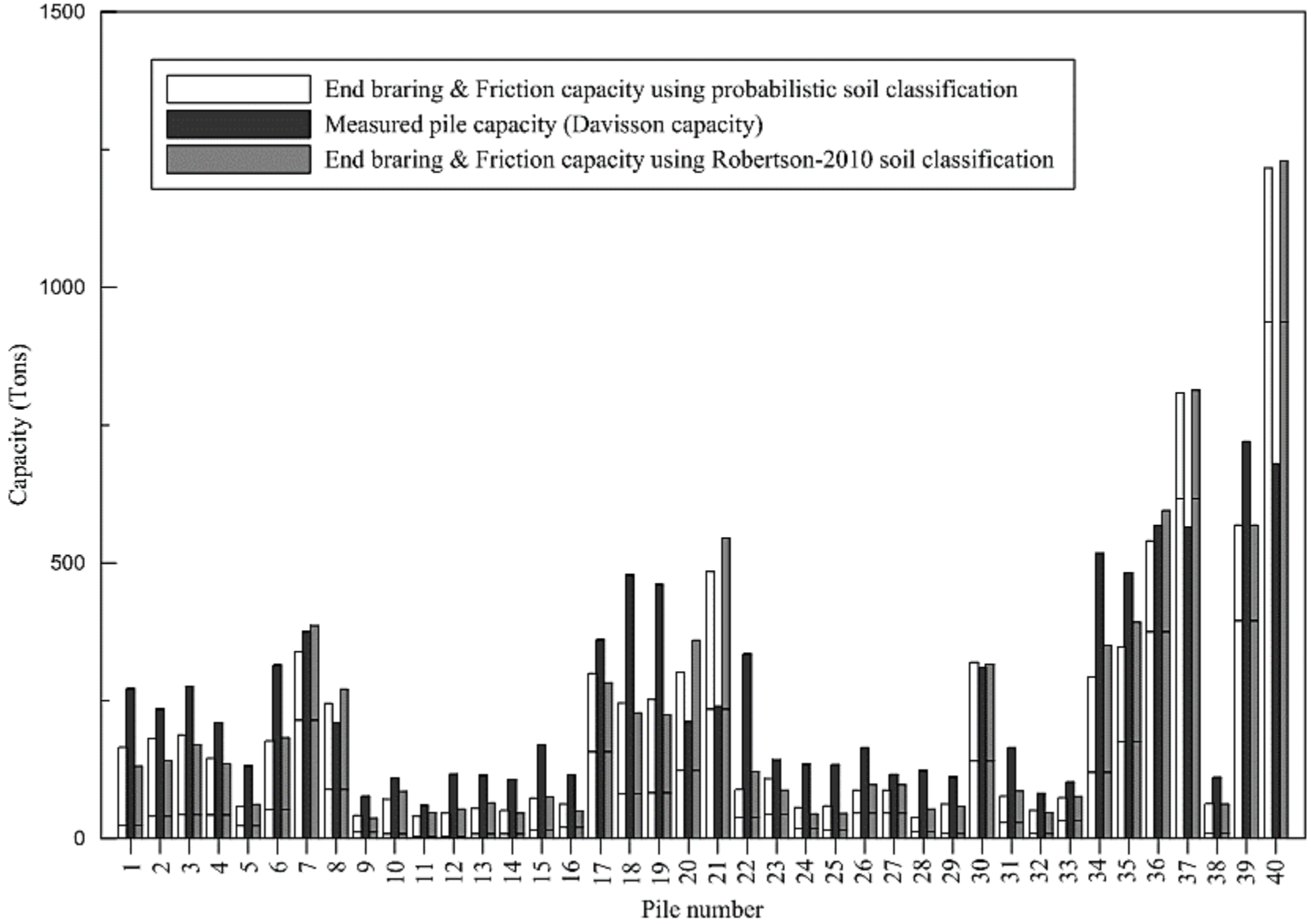


Figure 249. Comparison of measured and ultimate pile capacity predicted by Aoki method (piles 41-80)

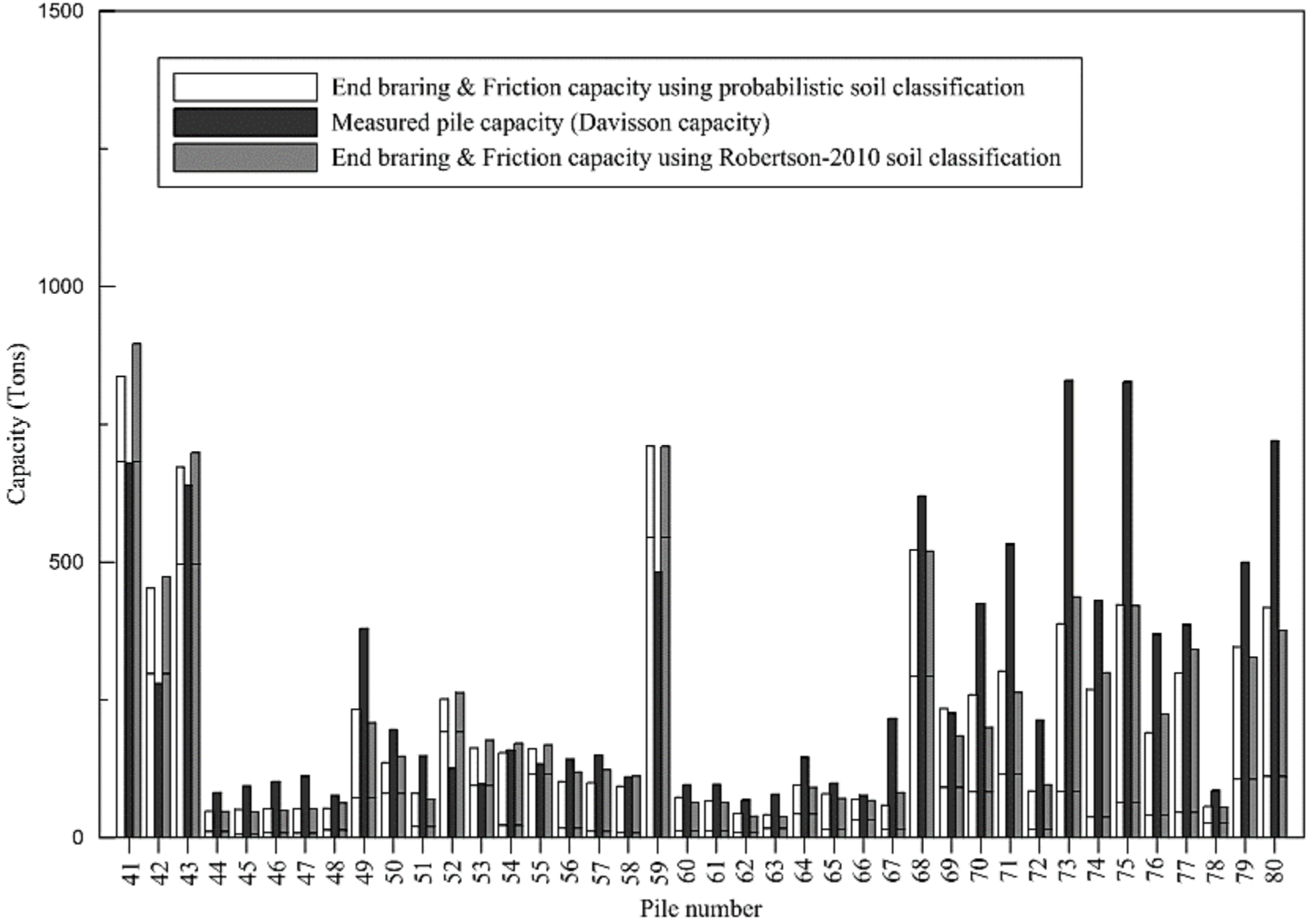


Figure 250. Comparison of measured and ultimate pile capacity predicted by Penpile method (piles 1-40)

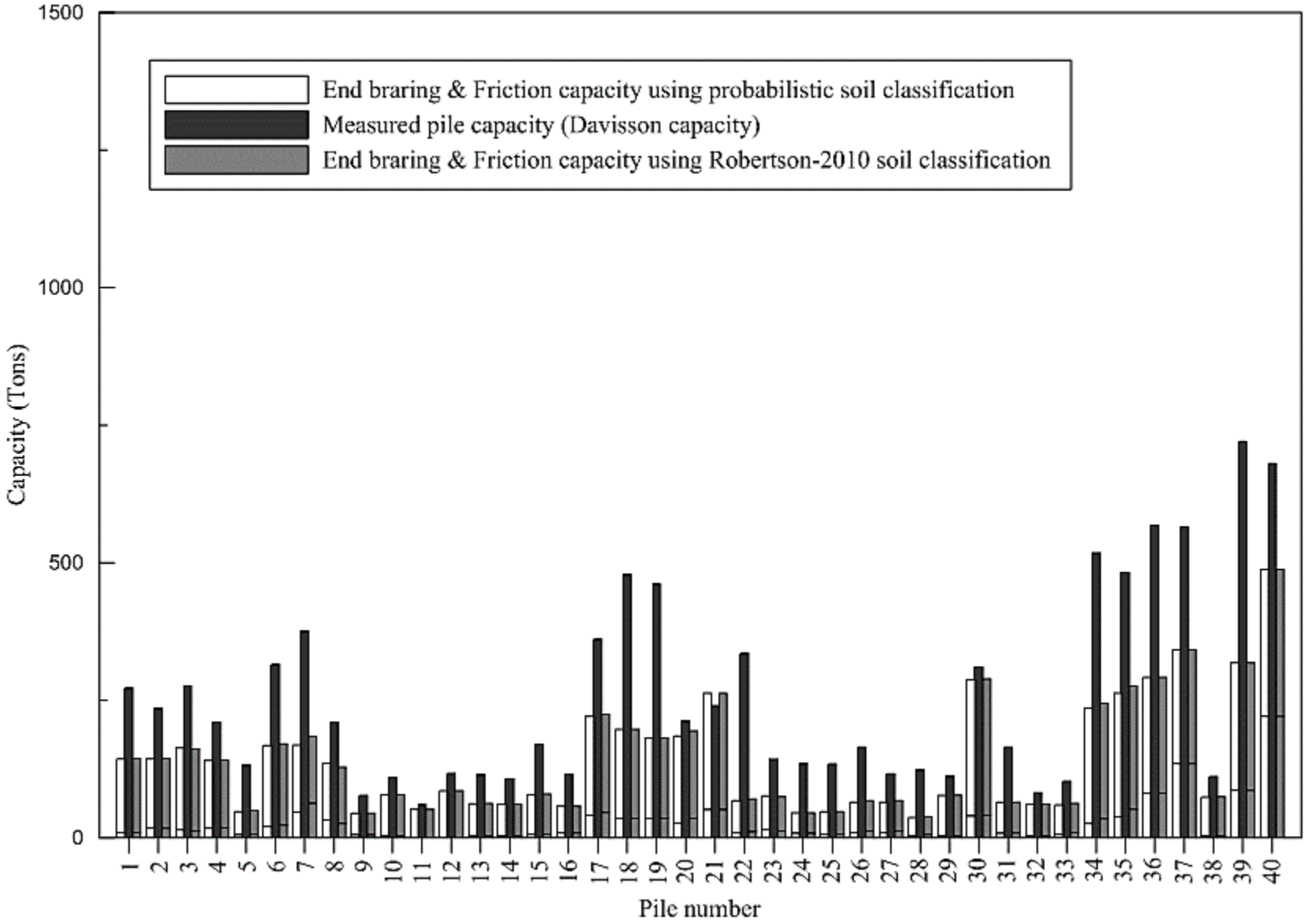


Figure 251. Comparison of measured and ultimate pile capacity predicted by Penpile method (piles 41-80)

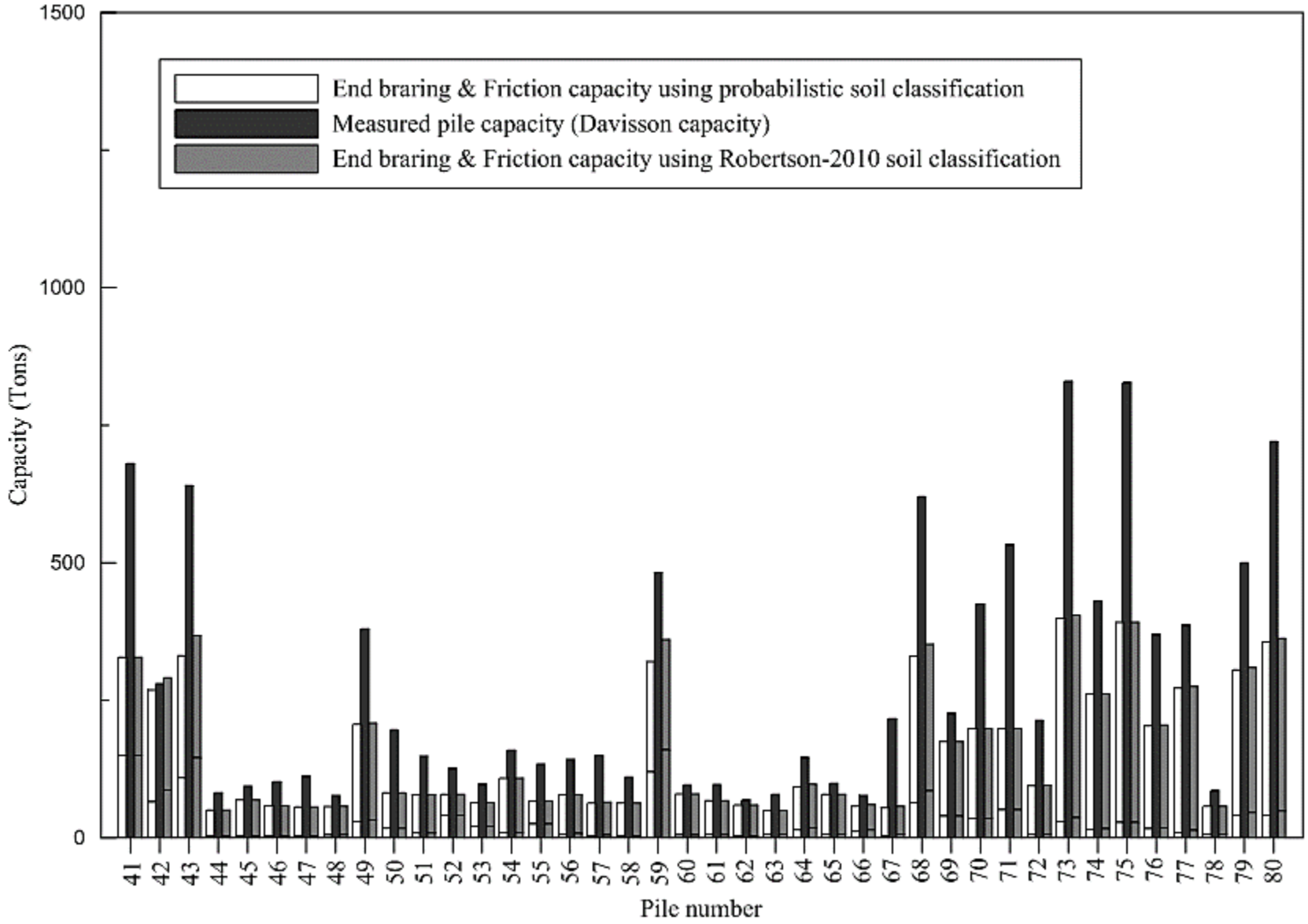


Figure 252. Comparison of measured and ultimate pile capacity predicted by NGI method (piles 1-40)

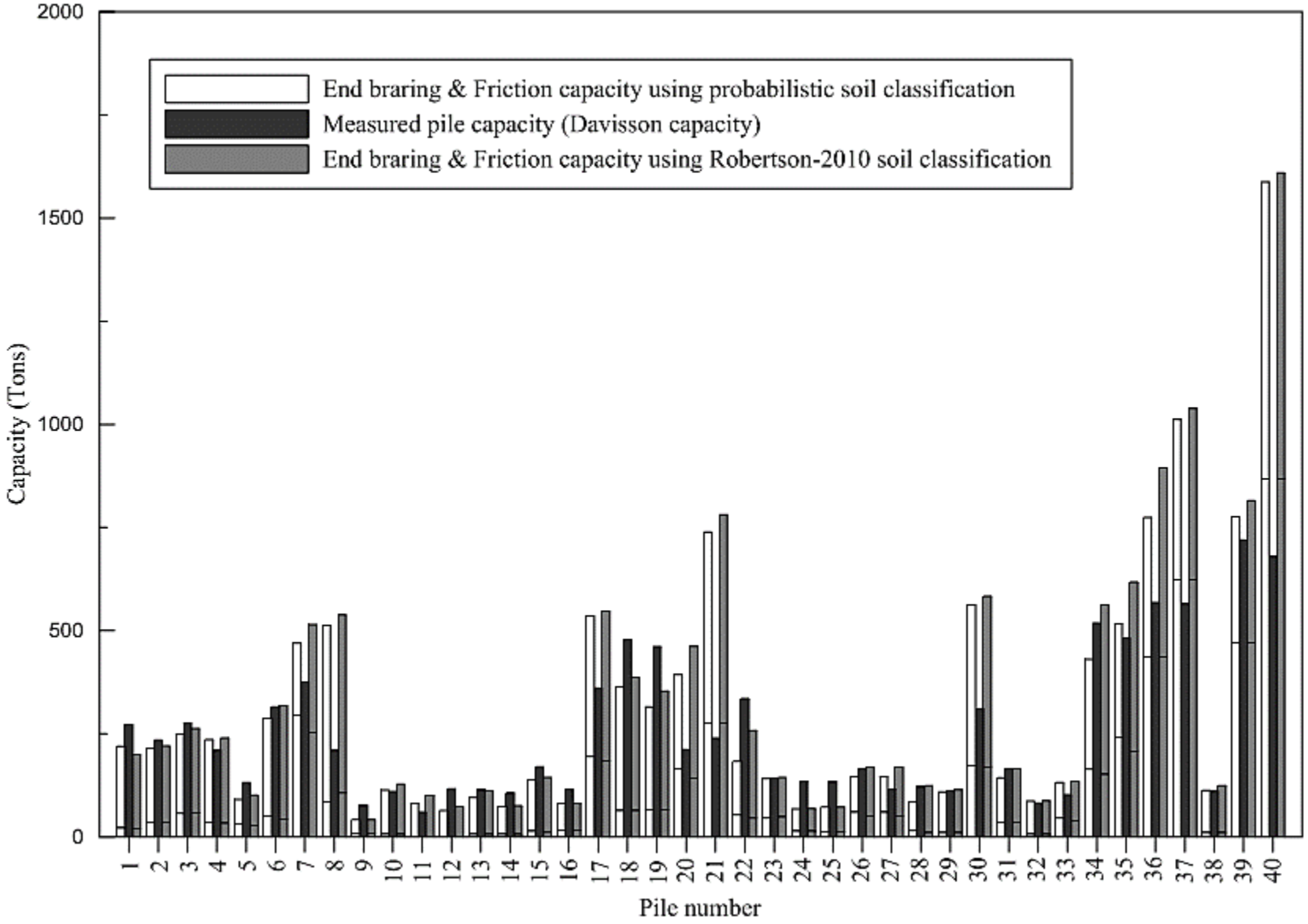


Figure 253. Comparison of measured and ultimate pile capacity predicted by NGI method (piles 41-80)

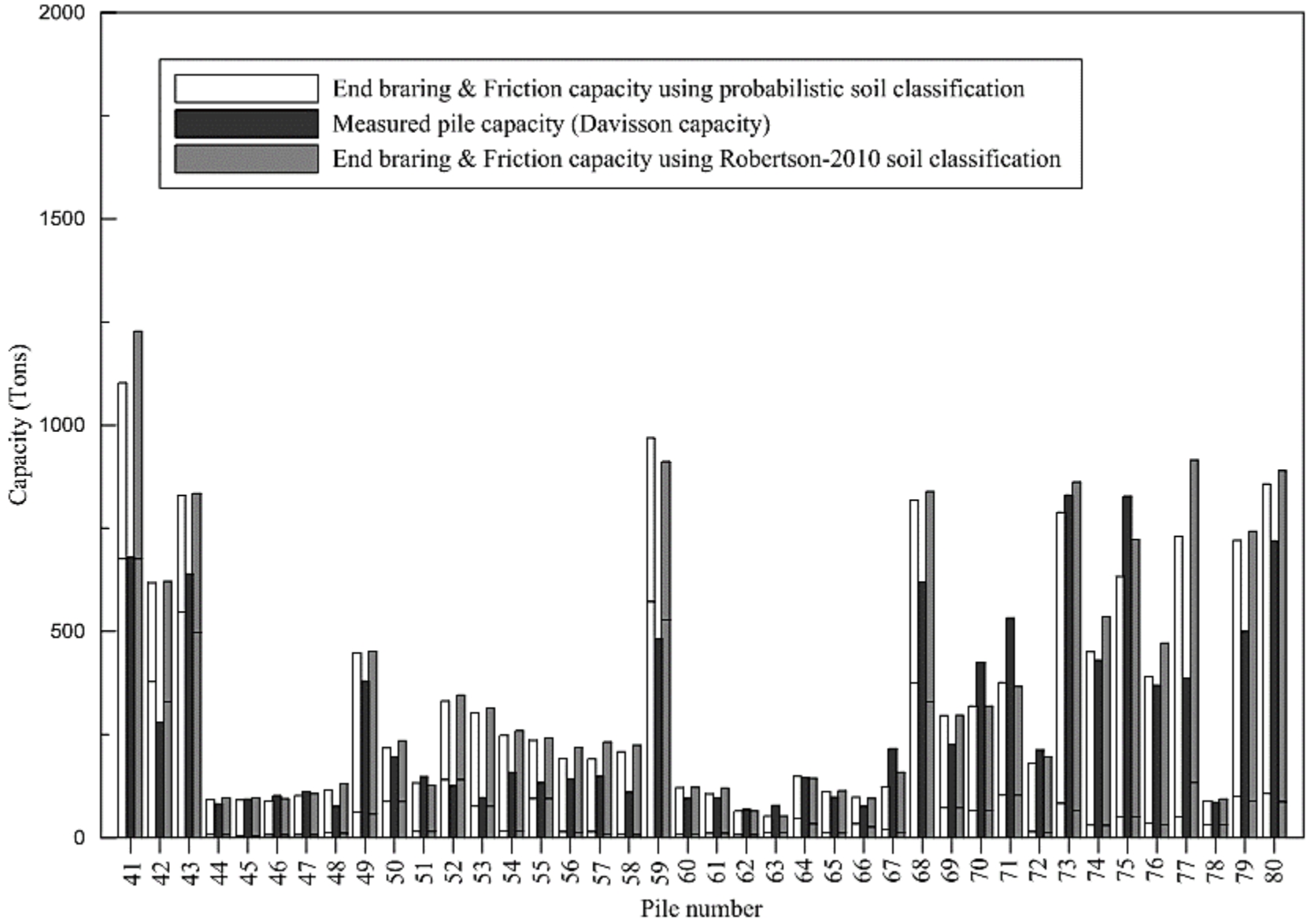


Figure 254. Comparison of measured and ultimate pile capacity predicted by ICP method (piles 1-40)

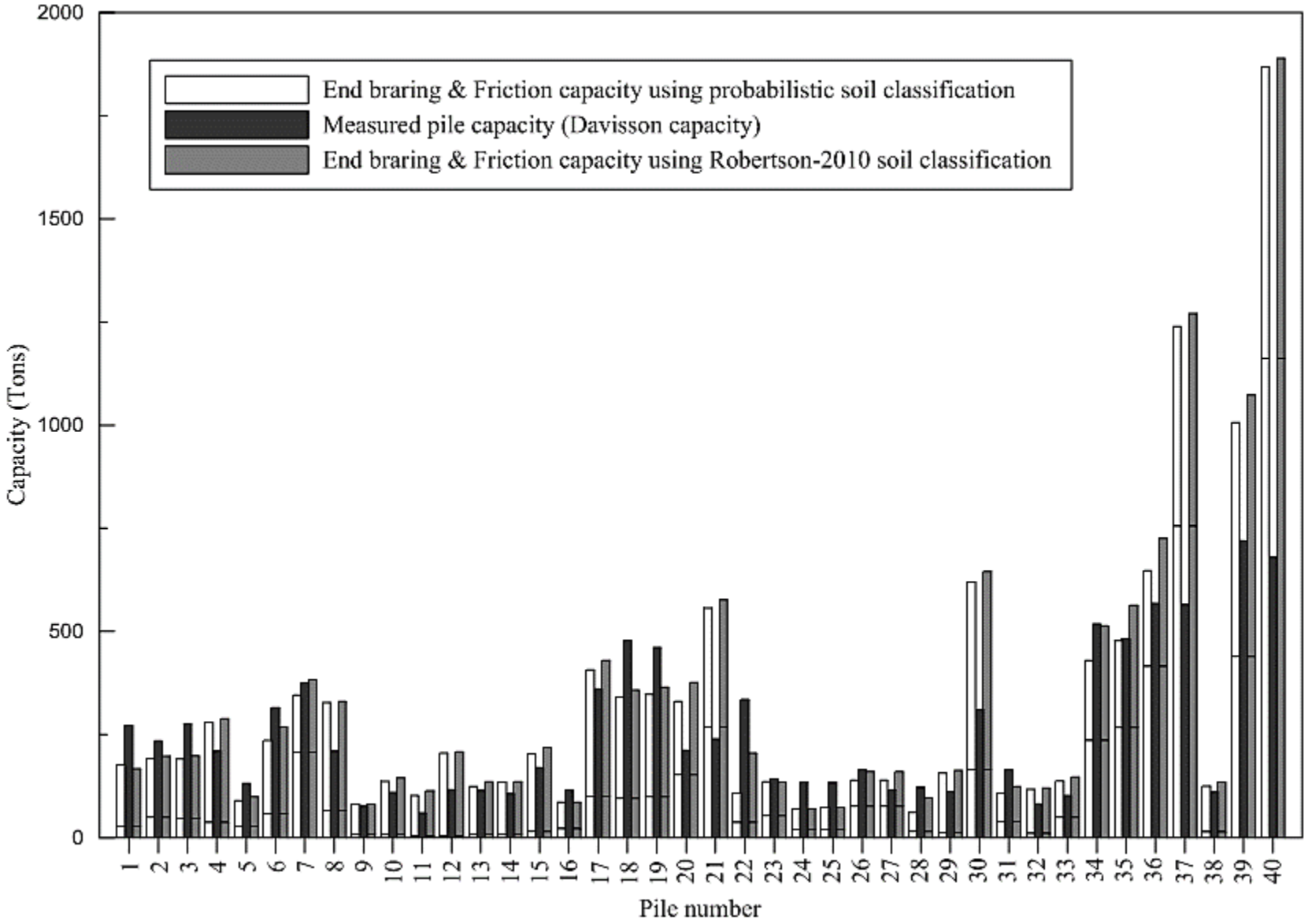


Figure 255. Comparison of measured and ultimate pile capacity predicted by ICP method (piles 41-80)

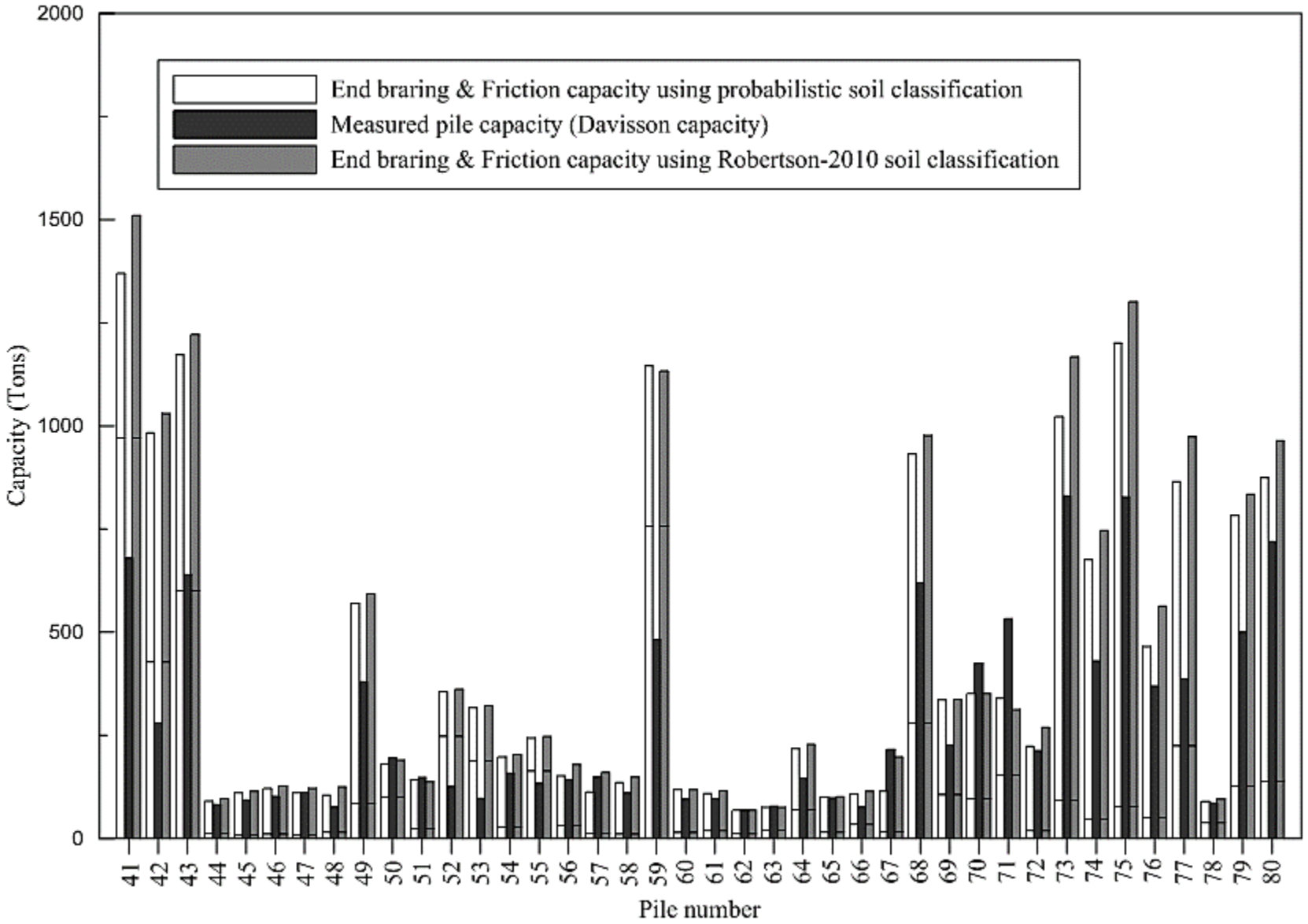


Figure 256. Comparison of measured and ultimate pile capacity predicted by UWA method (piles 1-40)

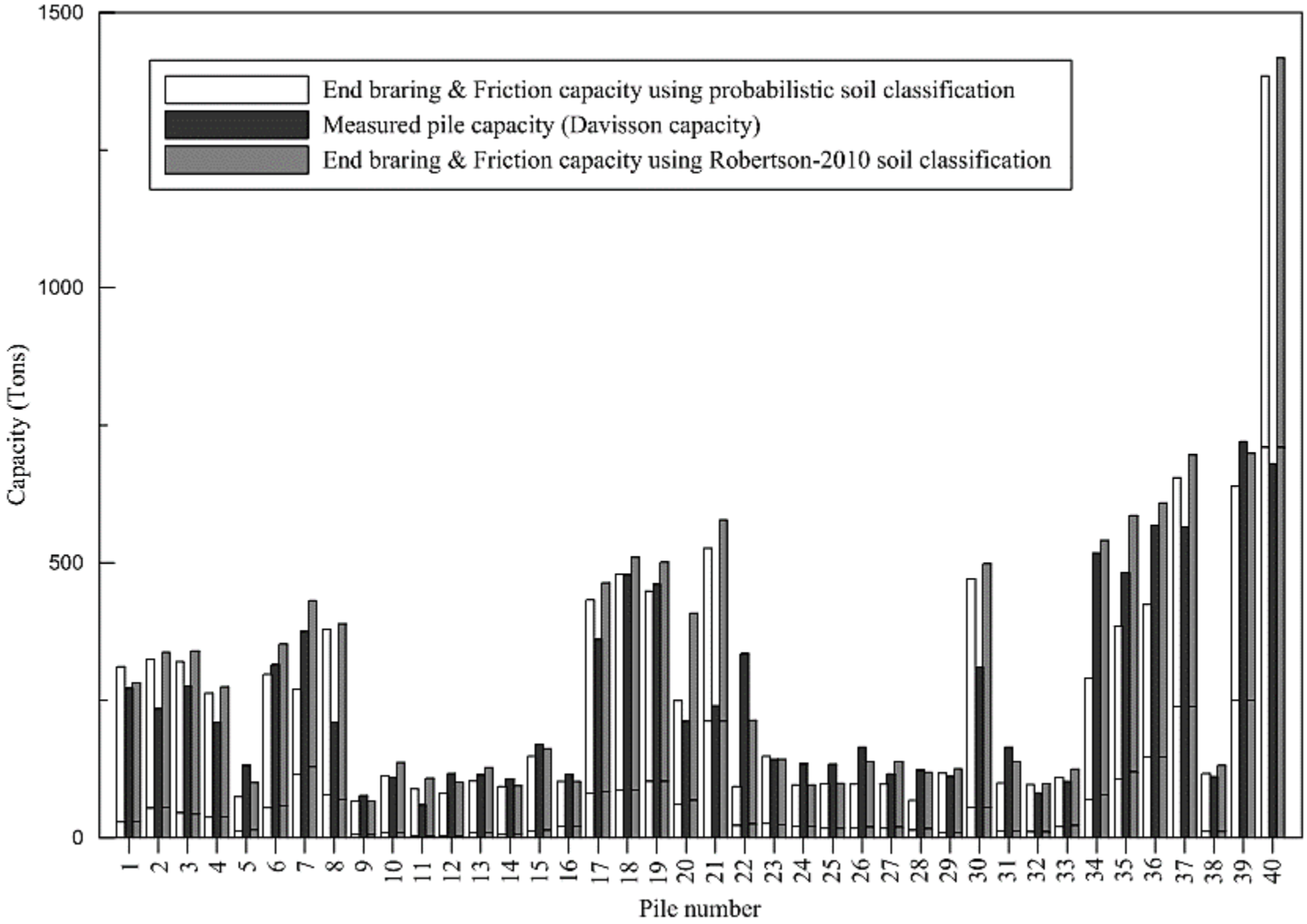


Figure 257. Comparison of measured and ultimate pile capacity predicted by UWA method (piles 41-80)

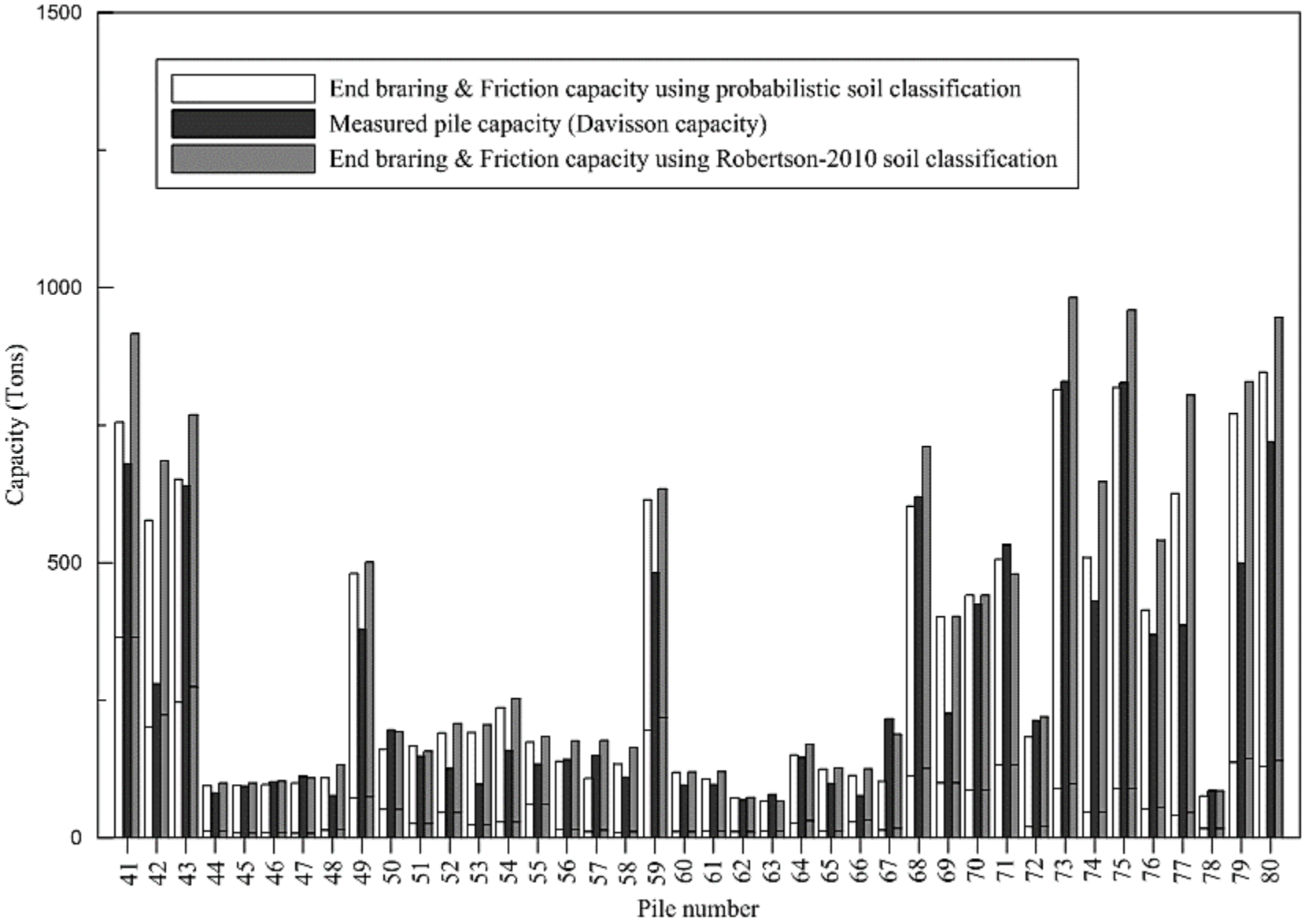


Figure 258. Comparison of measured and ultimate pile capacity predicted by CPT2000 method (piles 1-40)

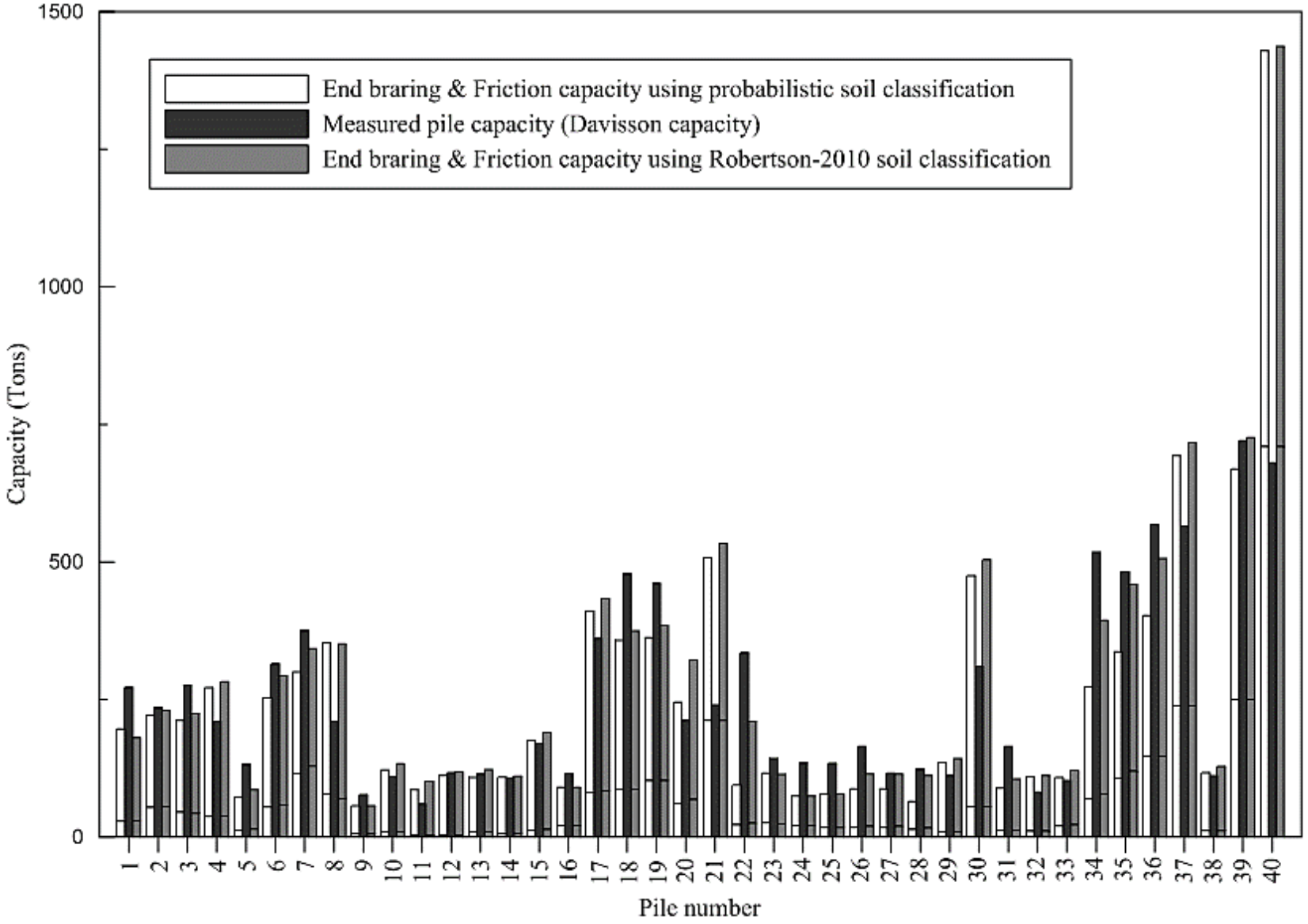


Figure 259. Comparison of measured and ultimate pile capacity predicted by CPT2000 method (piles 41-80)

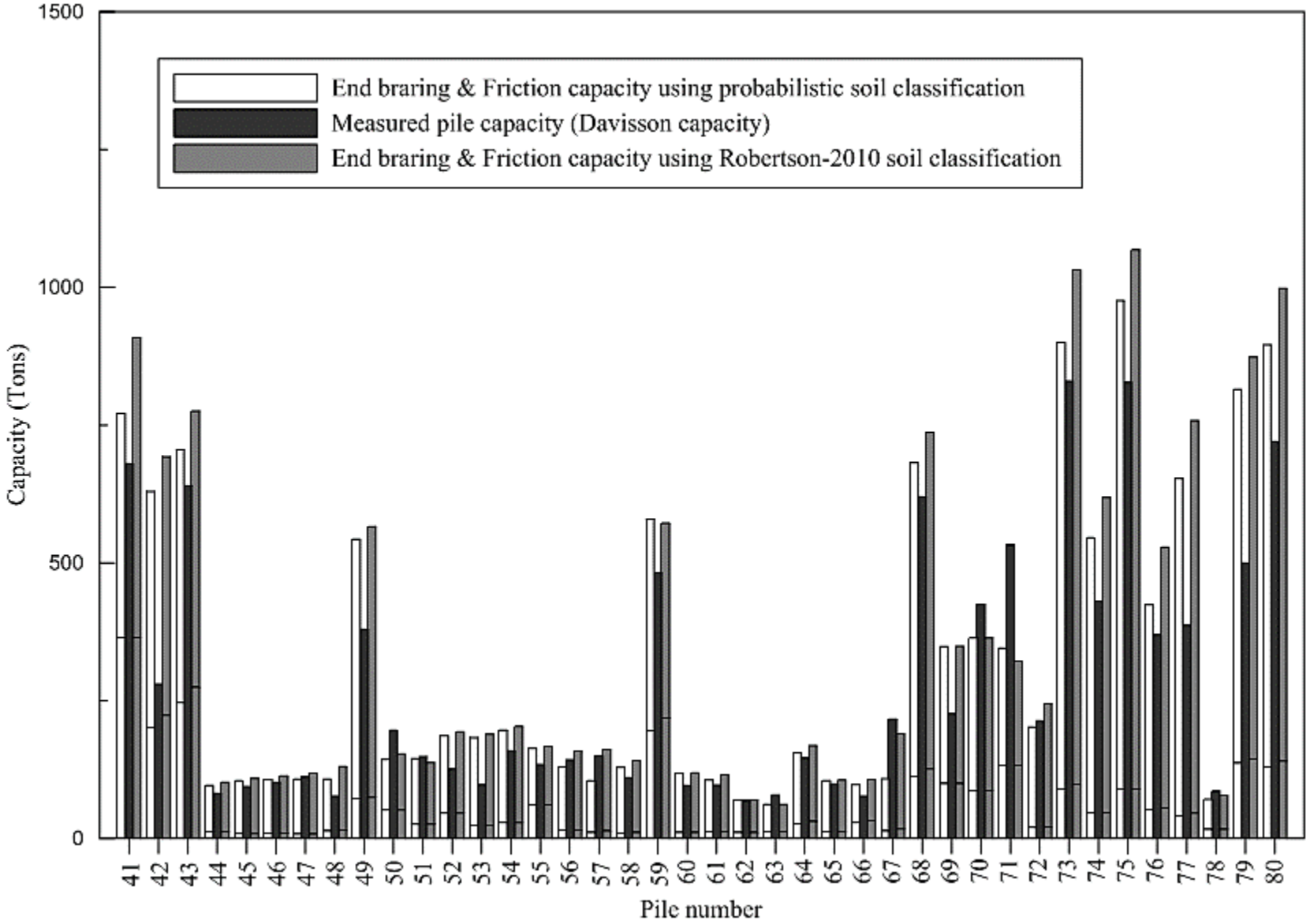


Figure 260. Comparison of measured and ultimate pile capacity predicted by Fugro method (piles 1-40)

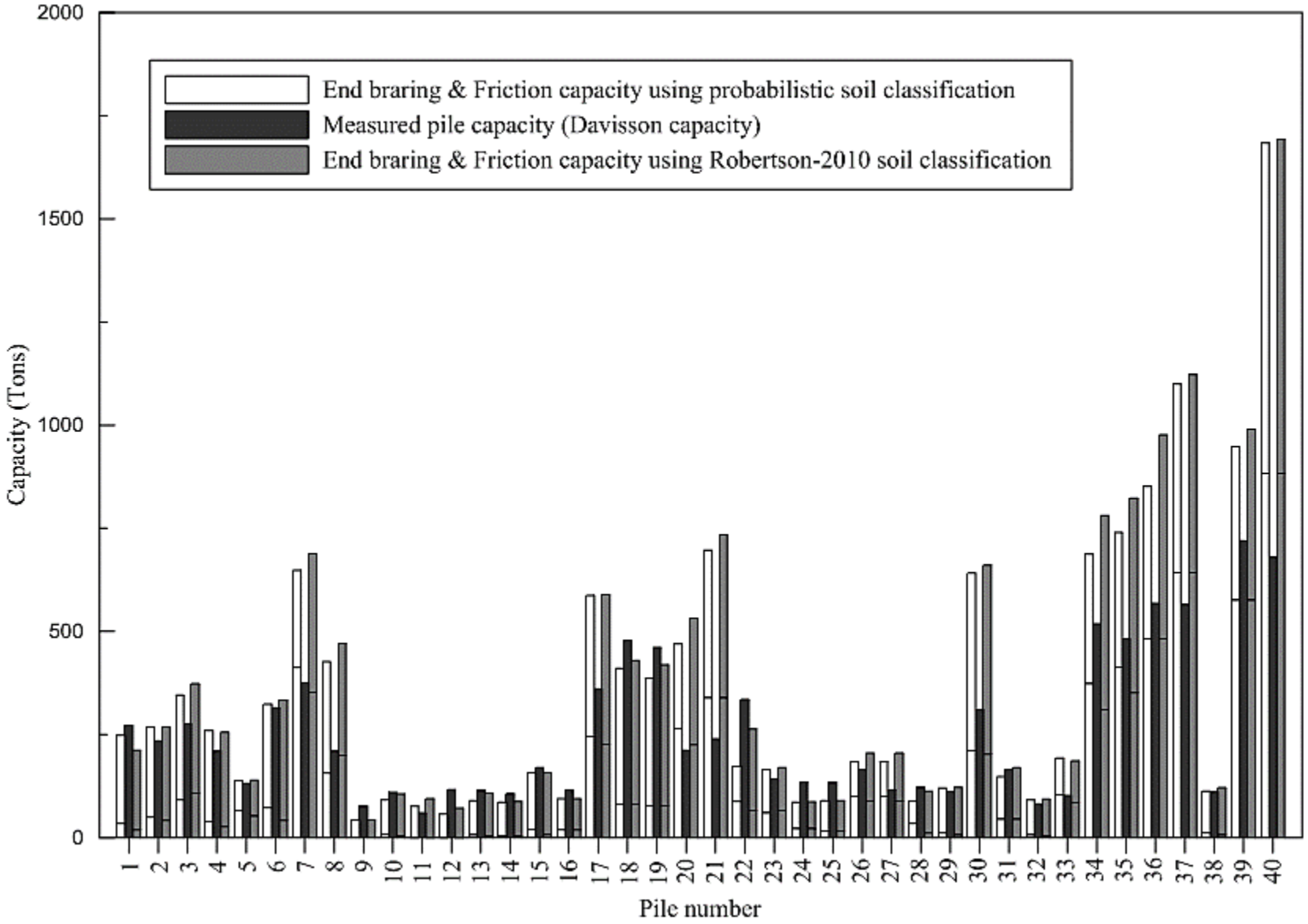


Figure 261. Comparison of measured and ultimate pile capacity predicted by Fugro method (piles 41-80)

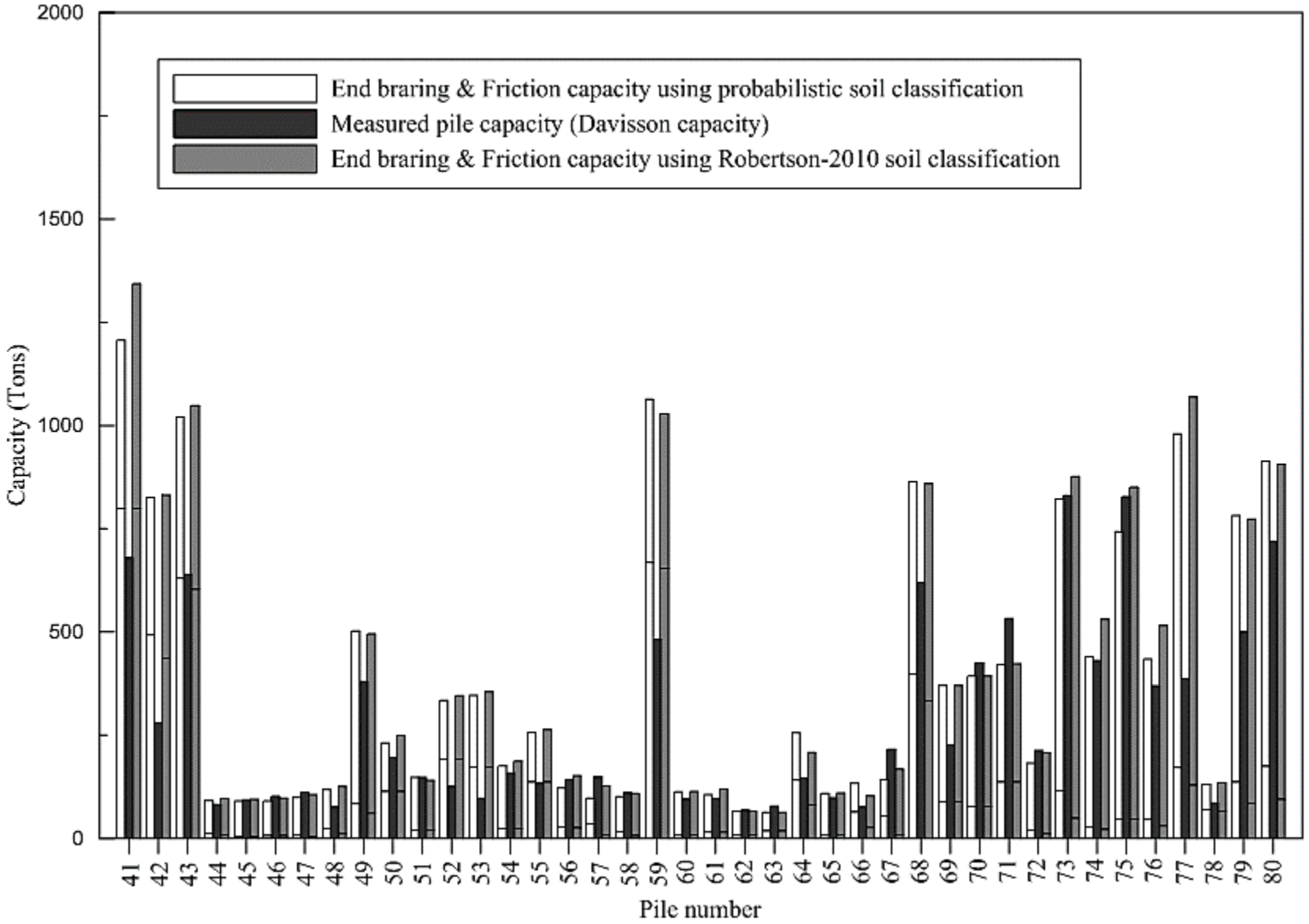


Figure 262. Comparison of measured and ultimate pile capacity predicted by Purdue method (piles 1-40)

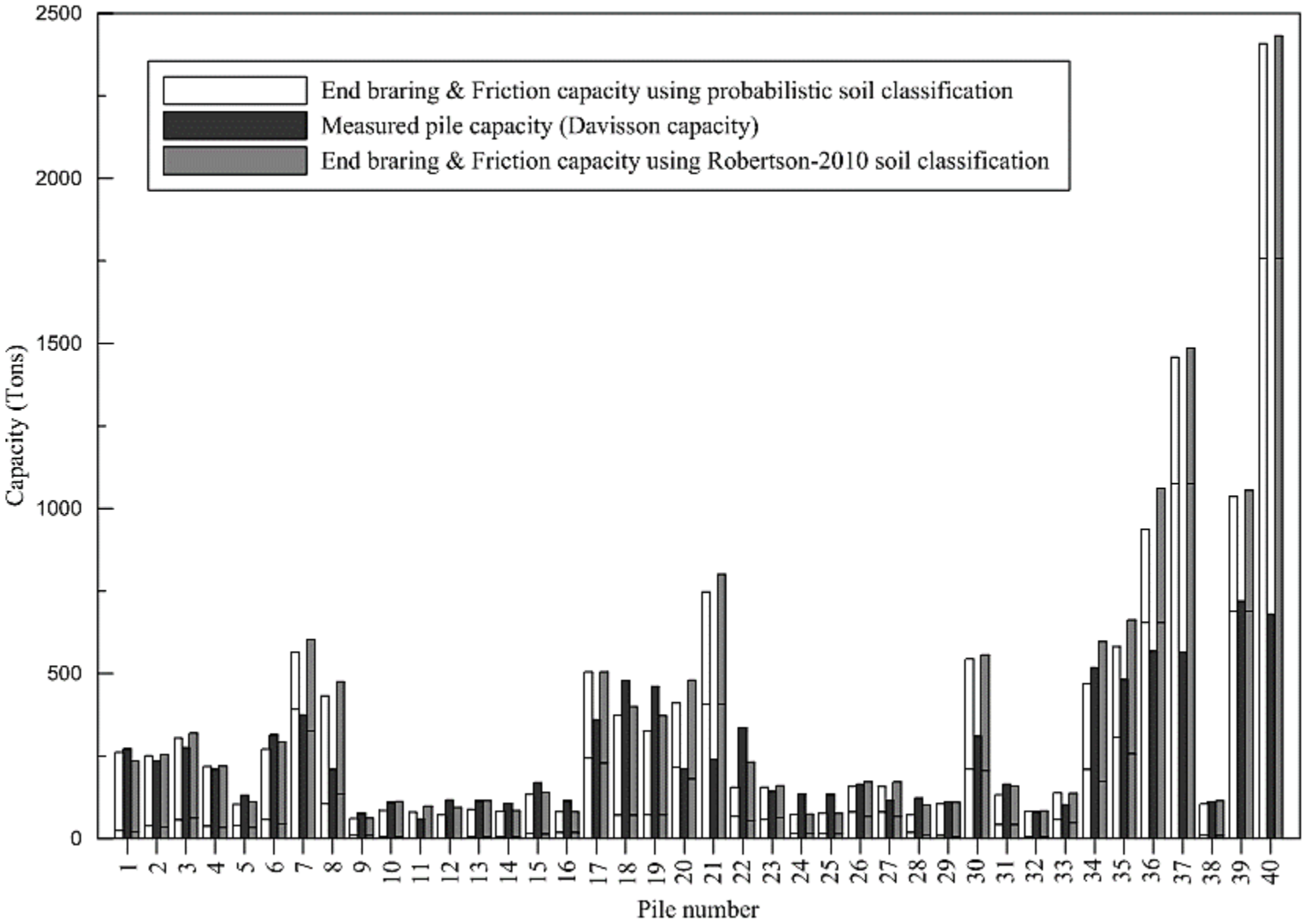


Figure 263. Comparison of measured and ultimate pile capacity predicted by Purdue method (piles 41-80)

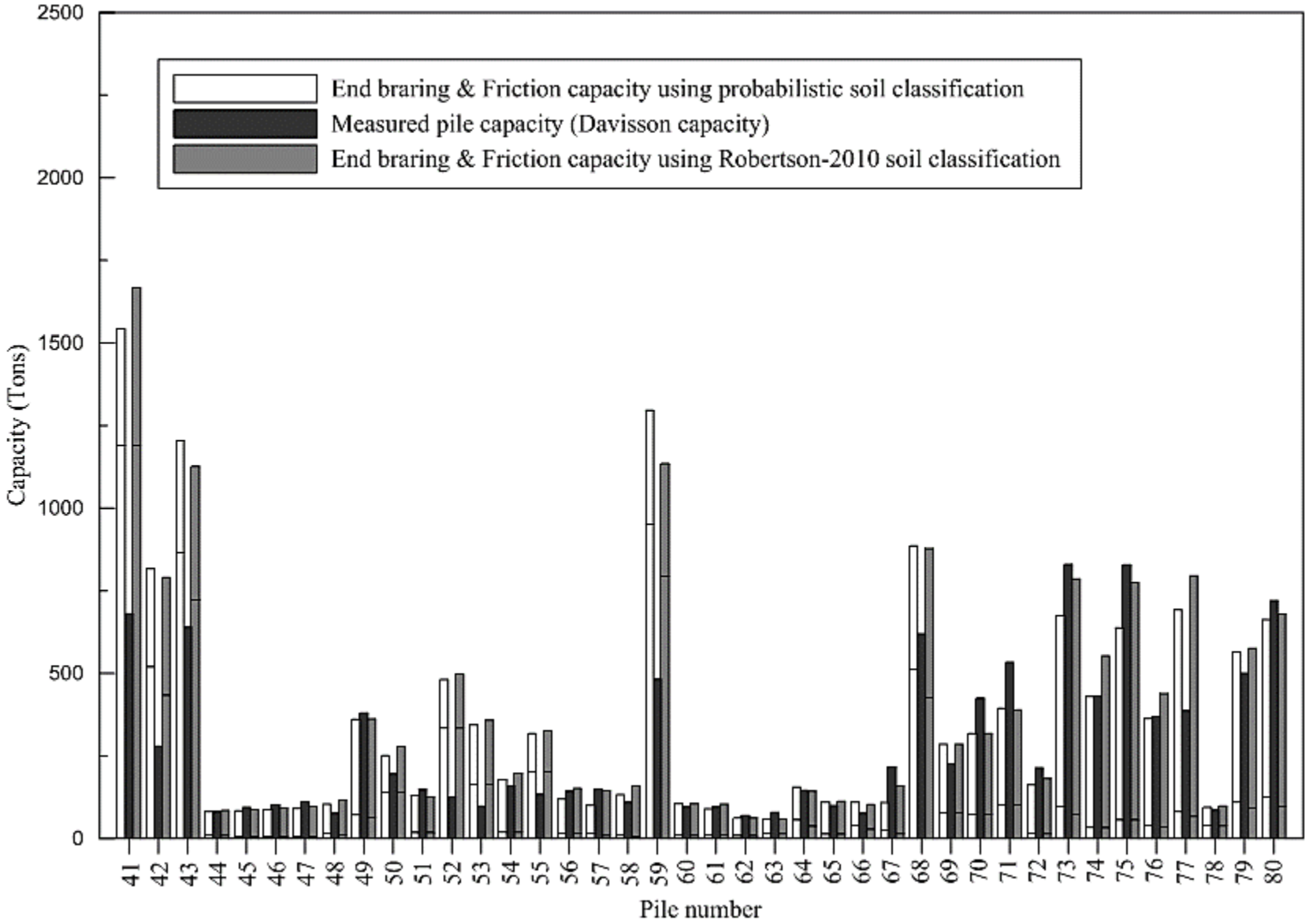


Figure 264. Comparison of measured and ultimate pile capacity predicted by Togliani method (piles 1-40)

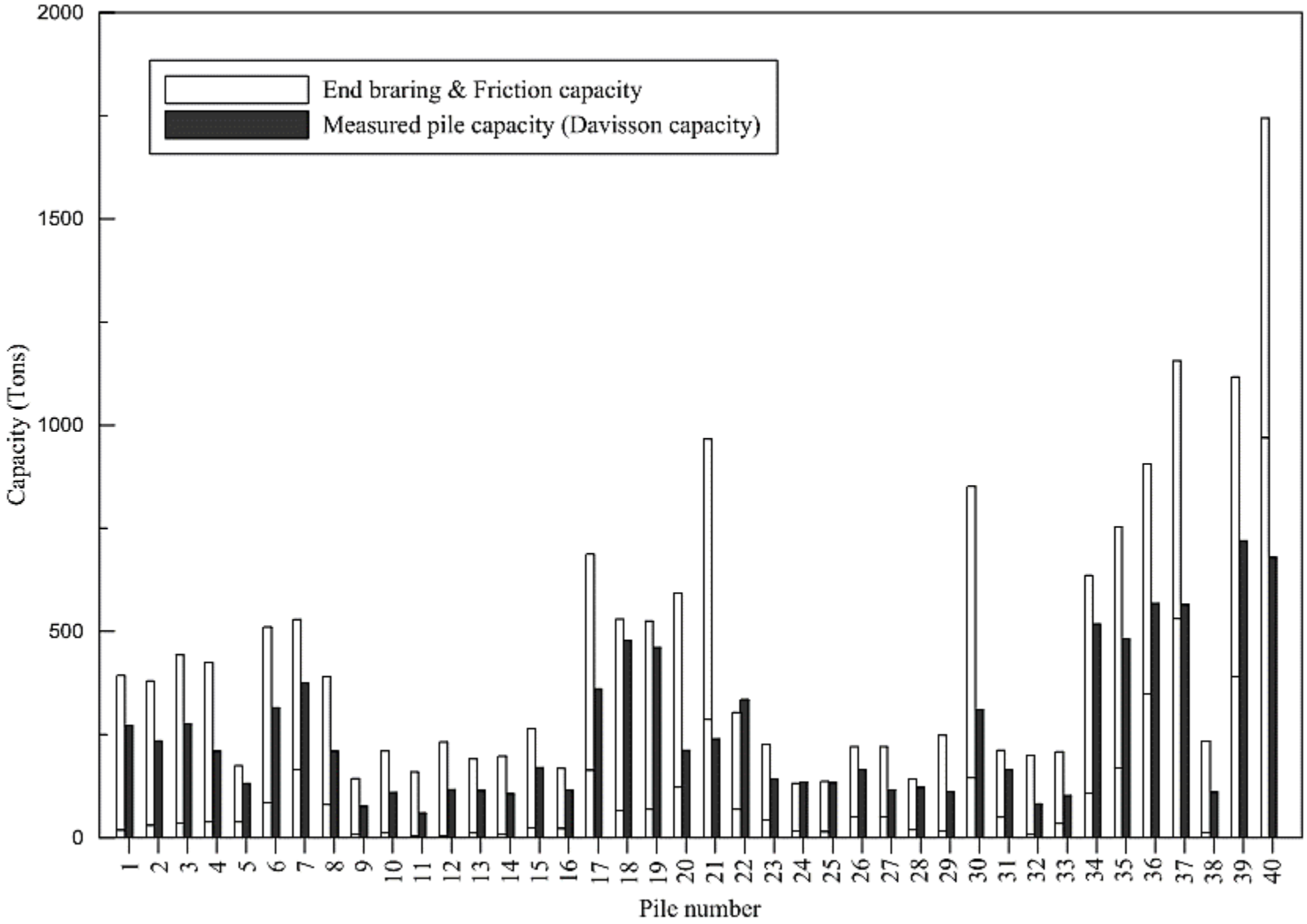


Figure 265. Comparison of measured and ultimate pile capacity predicted by Togliani method (piles 41-80)

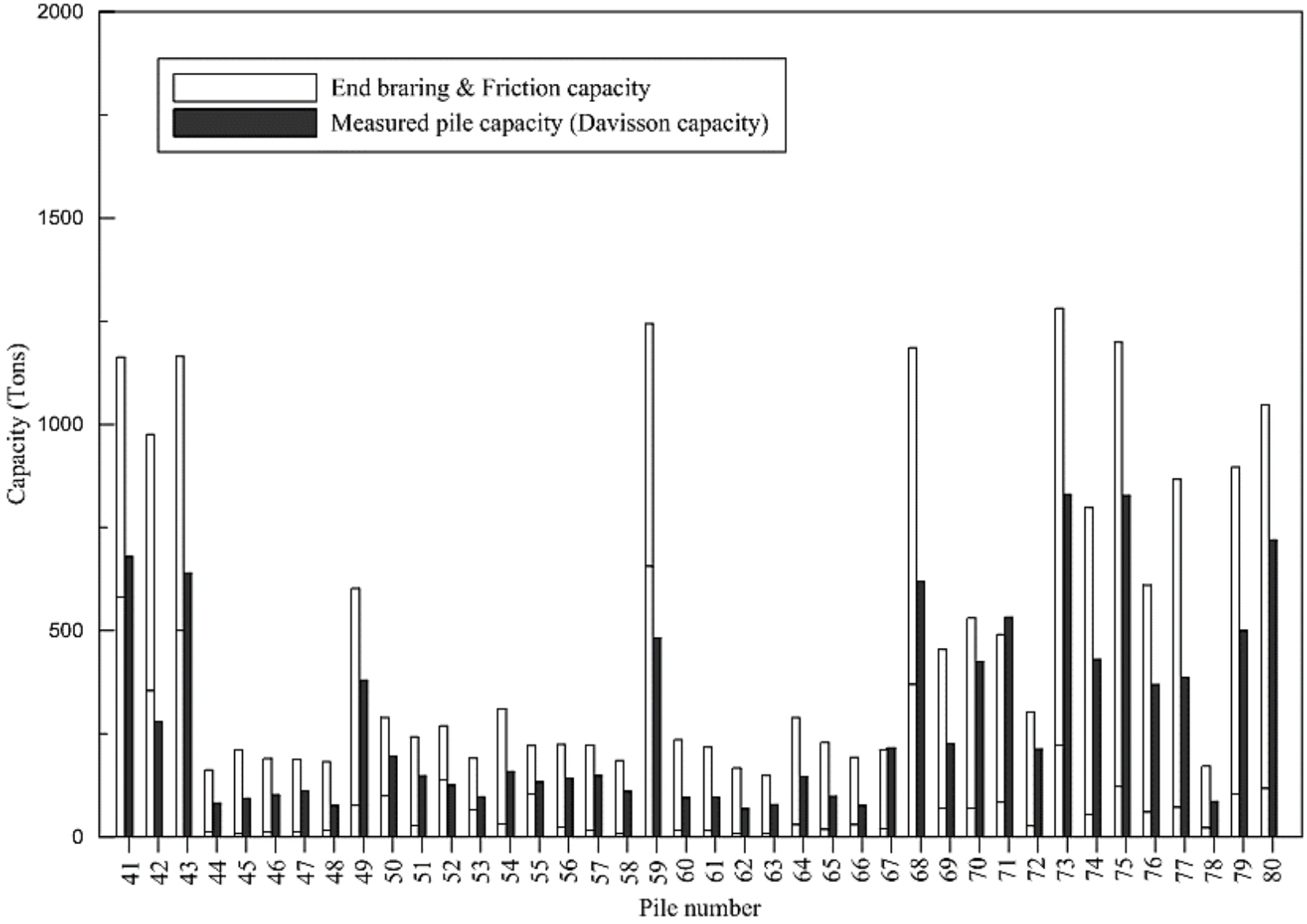


Figure 266. Comparison of measured and ultimate pile capacity predicted by German method (piles 1-40)

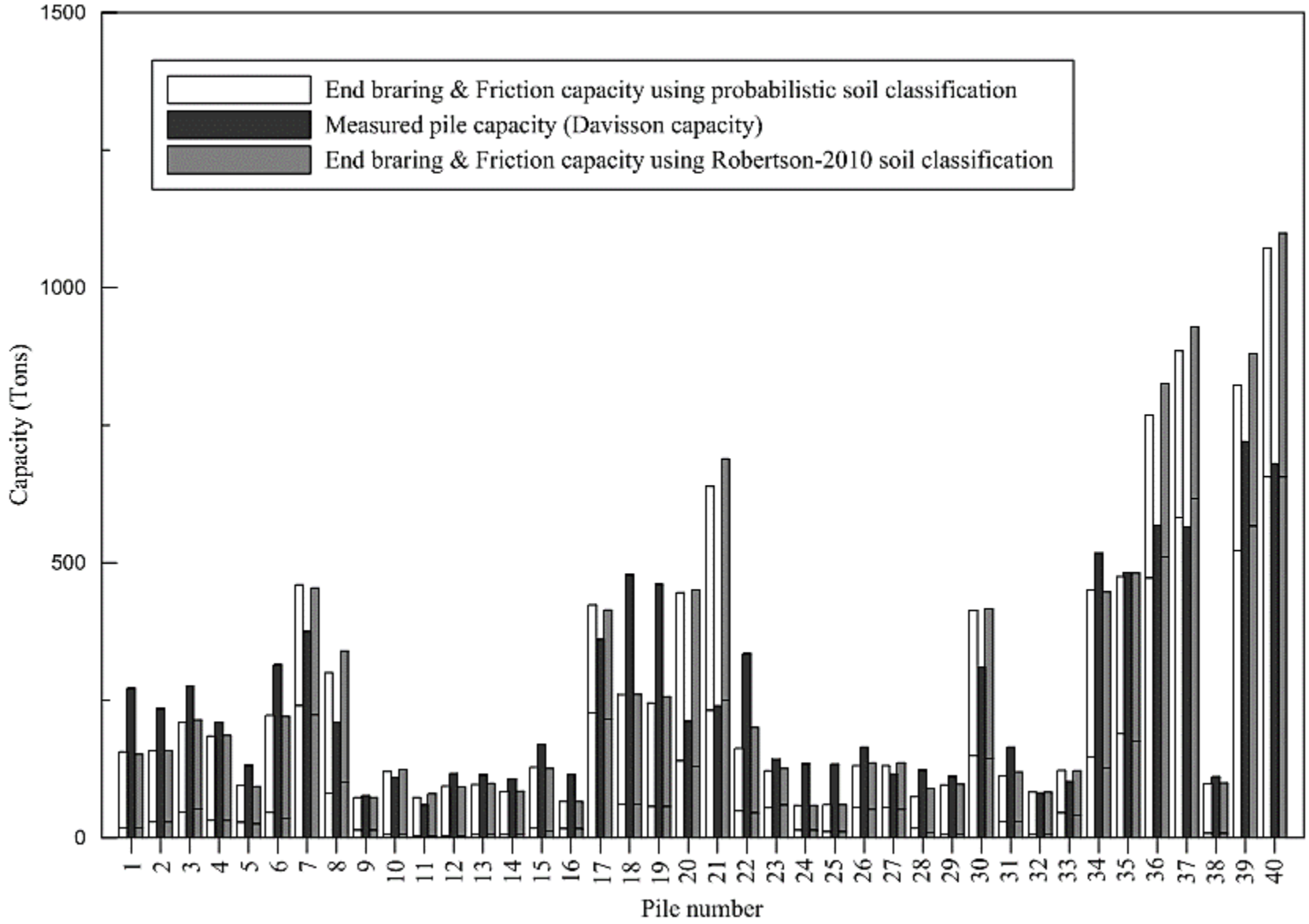


Figure 267. Comparison of measured and ultimate pile capacity predicted by German method (piles 41-80)

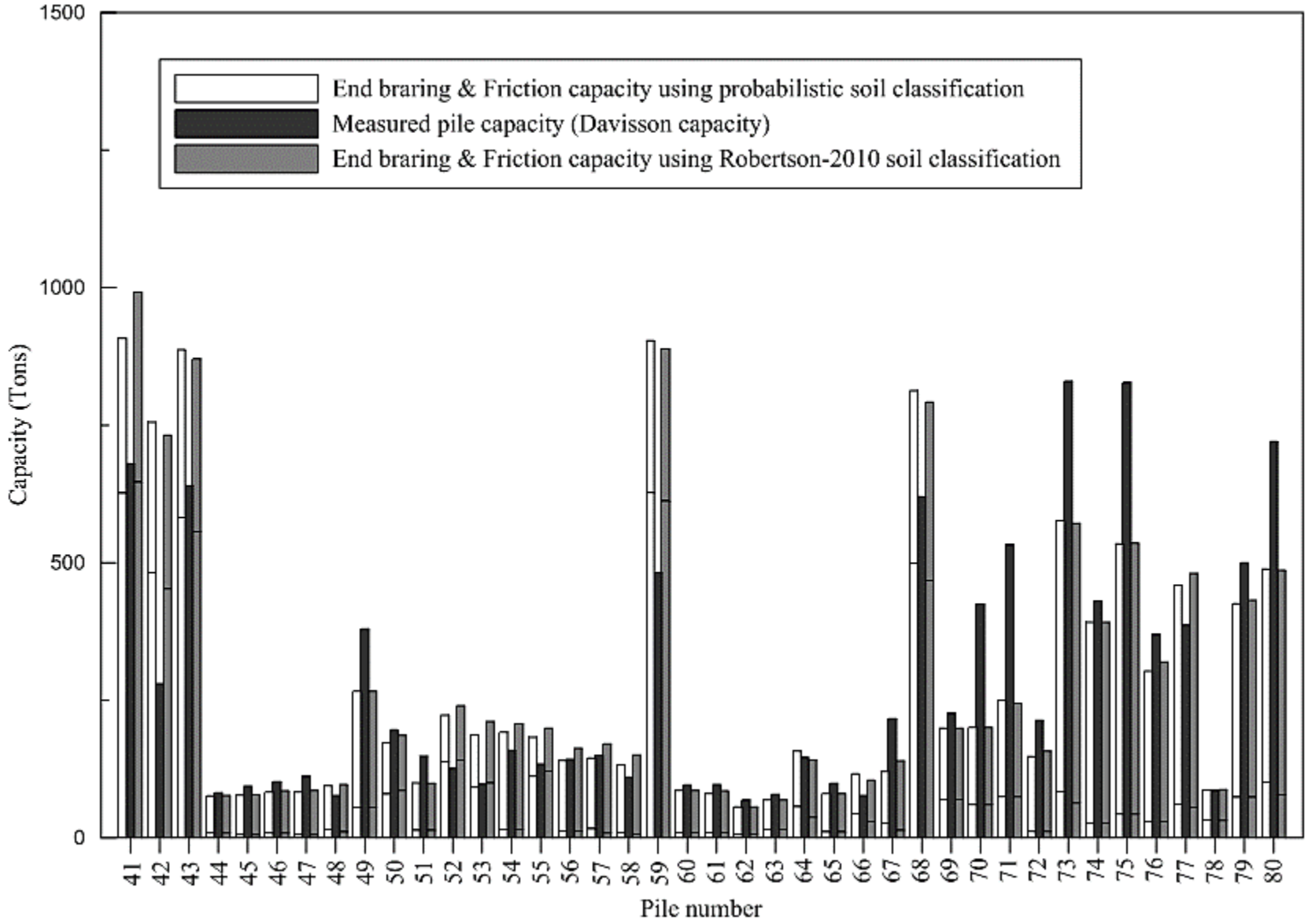


Figure 268. Comparison of measured and ultimate pile capacity predicted by Eurocode7 method (piles 1-40)

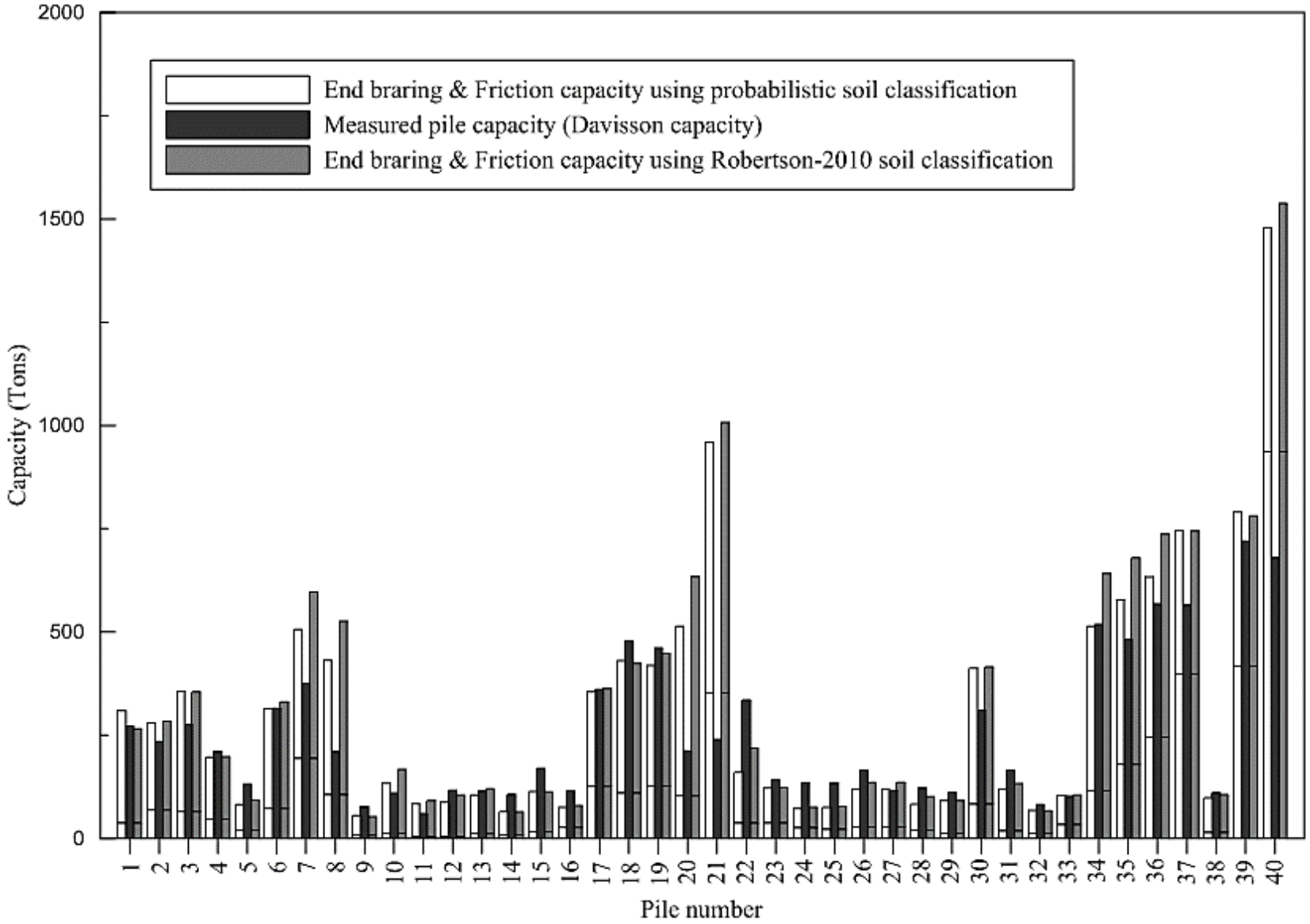


Figure 269. Comparison of measured and ultimate pile capacity predicted by Eurocode7 method (piles 41-80)

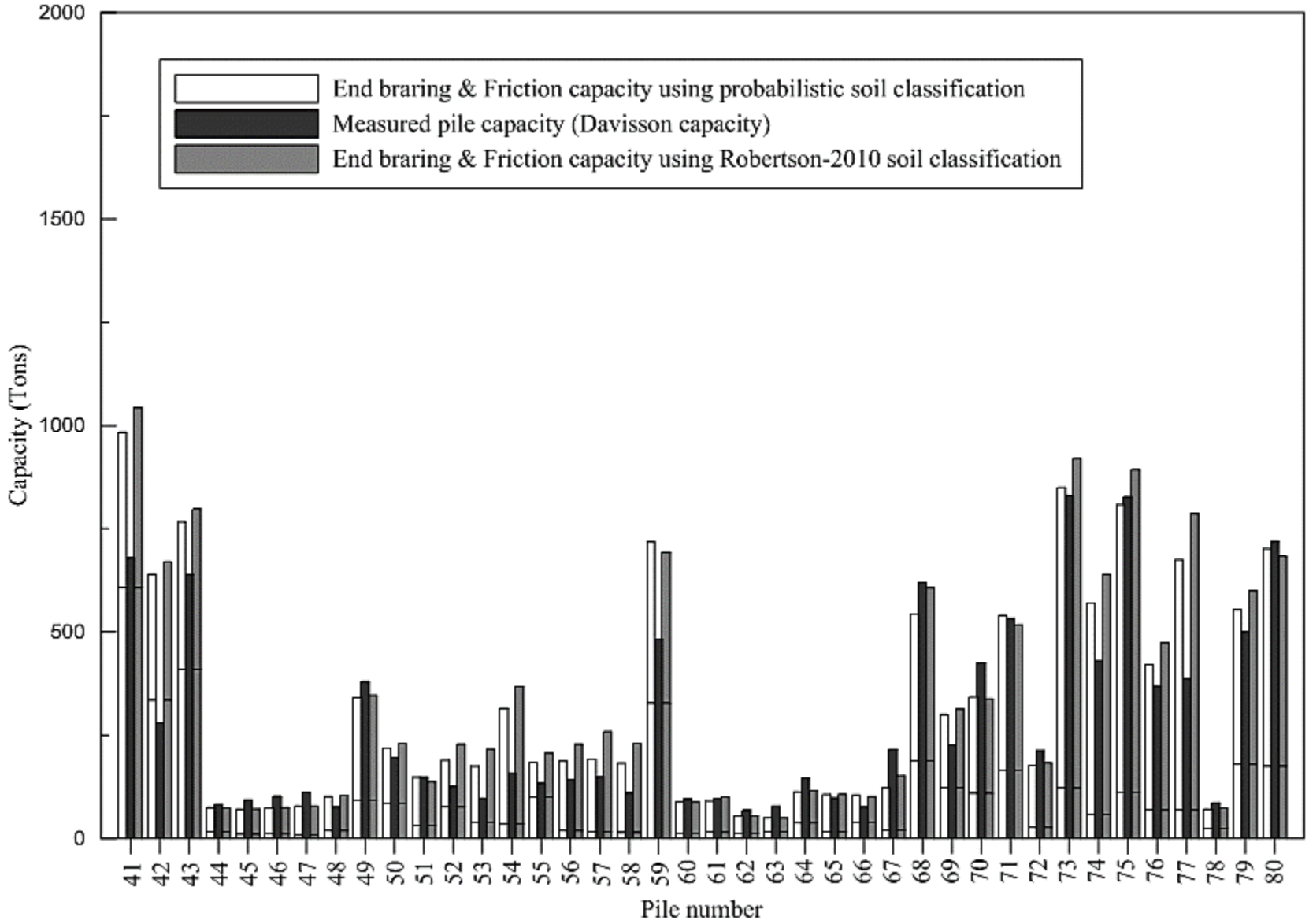


Figure 270. Comparison of measured and ultimate pile capacity predicted by ERTC3 method (piles 1-40)

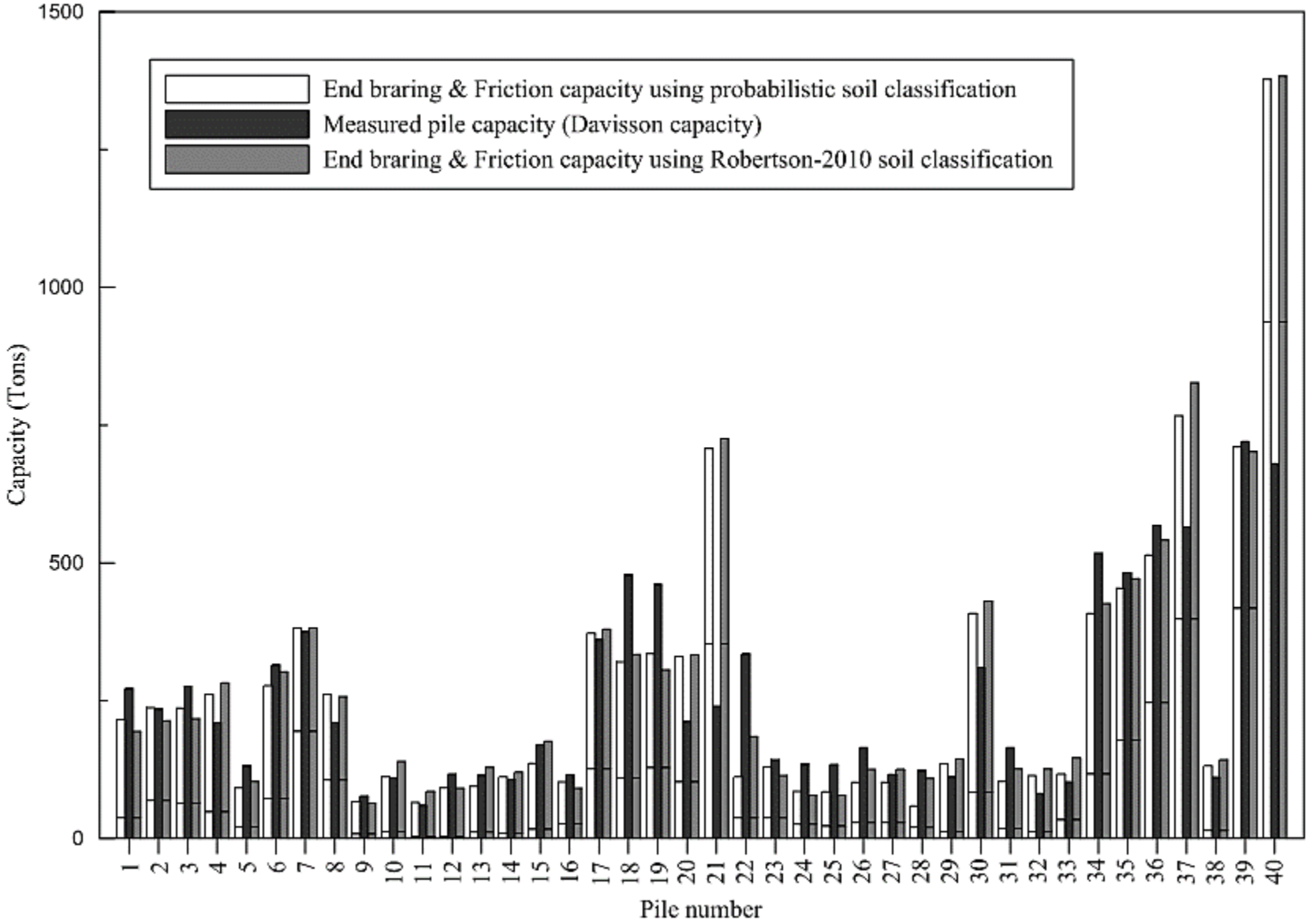
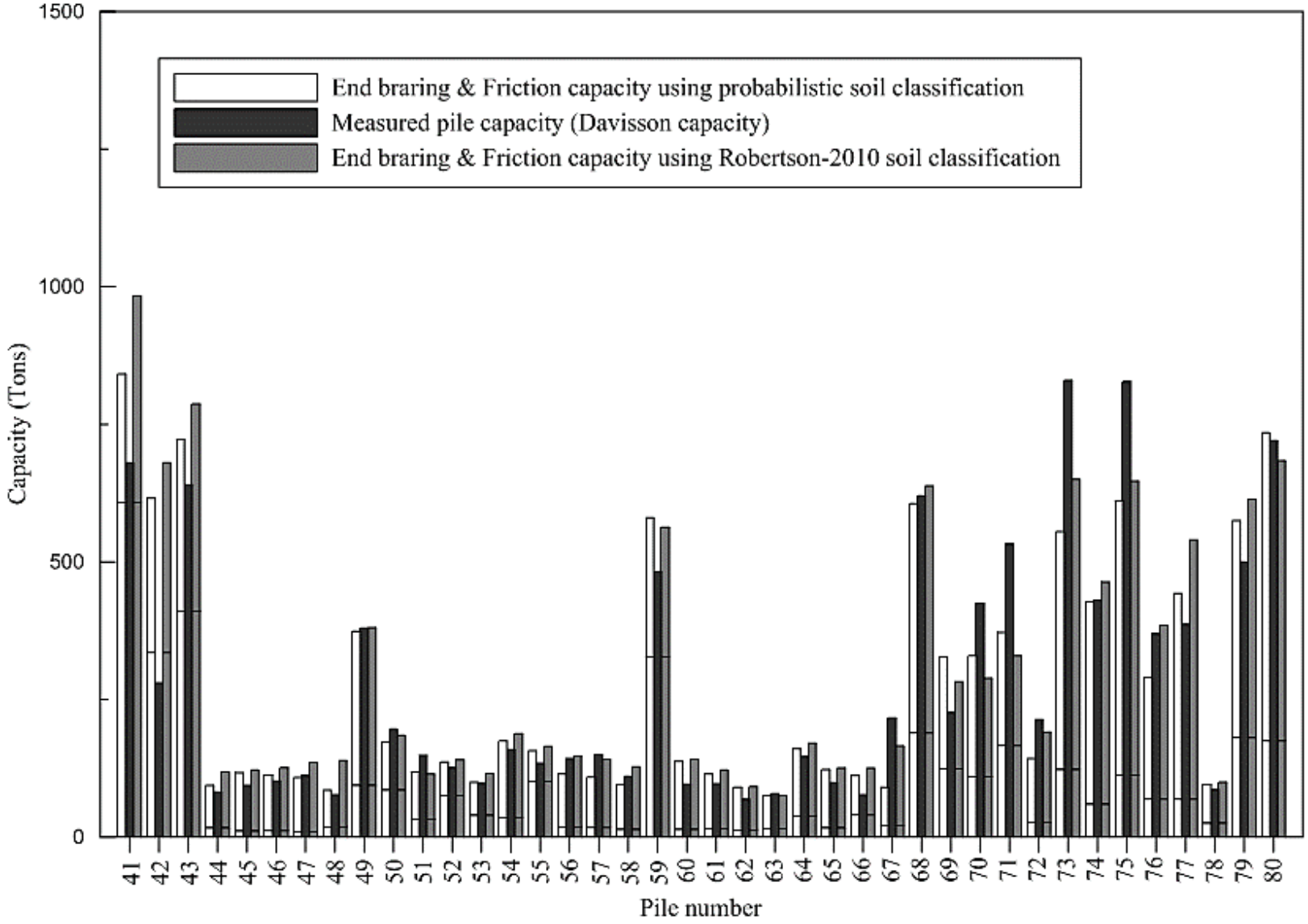


Figure 271. Comparison of measured and ultimate pile capacity predicted by ERTC3 method (piles 41-80)



Appendix F

Predicted versus Measured Ultimate Capacity and Cumulative Probability Plots

Figure 272. Predicted versus measured ultimate capacity and cumulative probability for Bustamante and Gianeselli (LCPC)

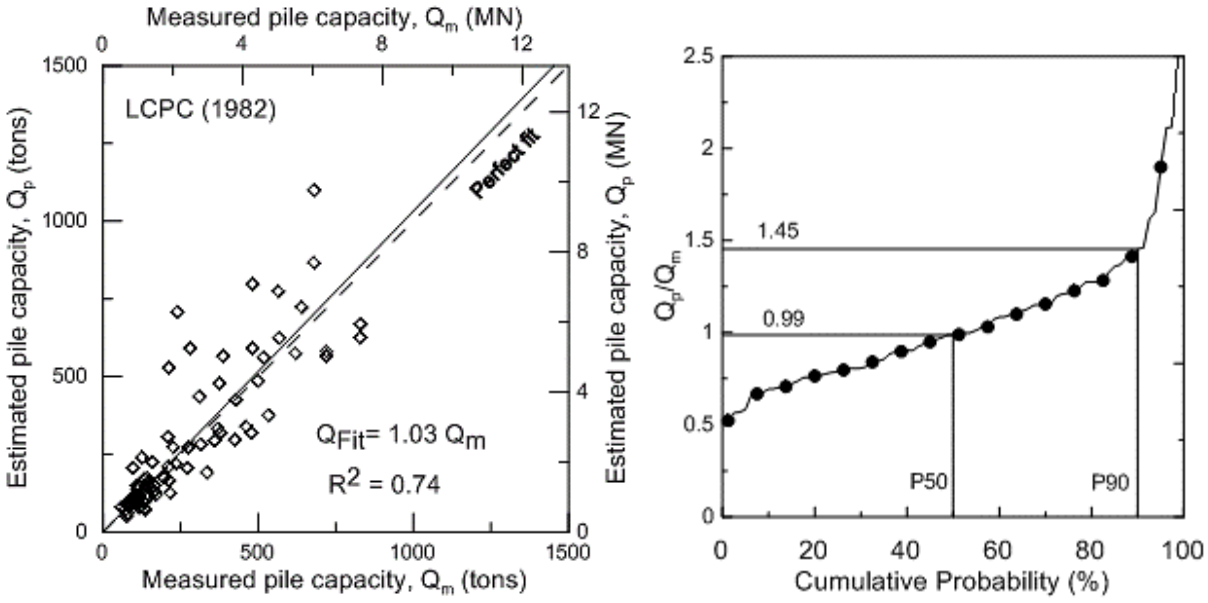


Figure 273. Predicted versus measured ultimate capacity and cumulative probability for Schmertmann

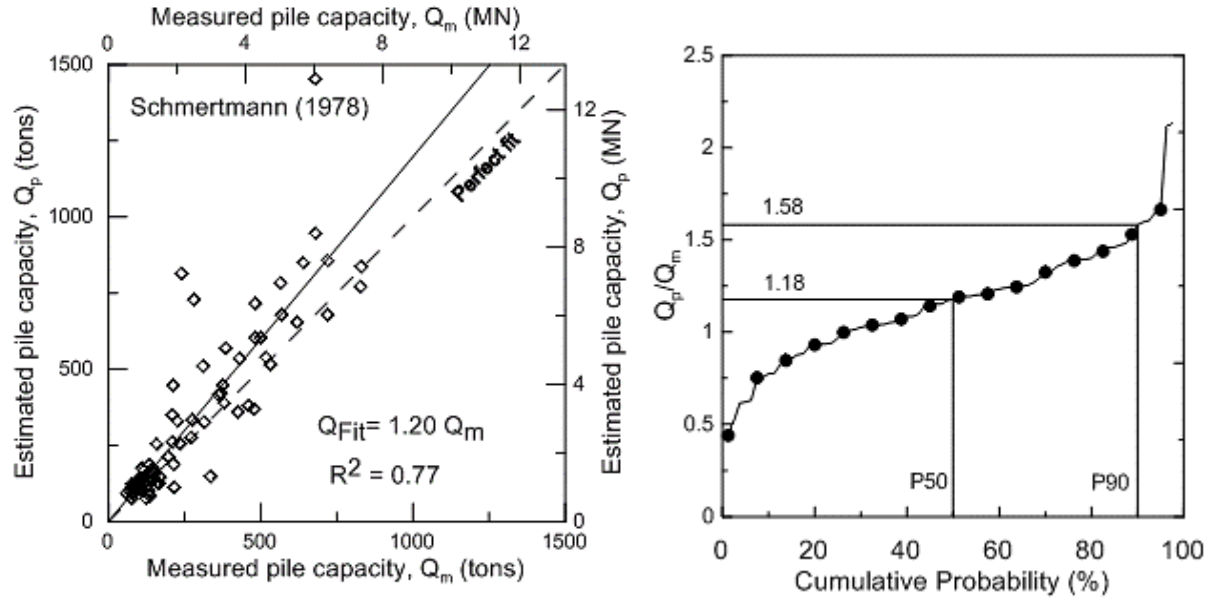


Figure 274. Predicted versus measured ultimate capacity and cumulative probability for De Ruiter and Beringen

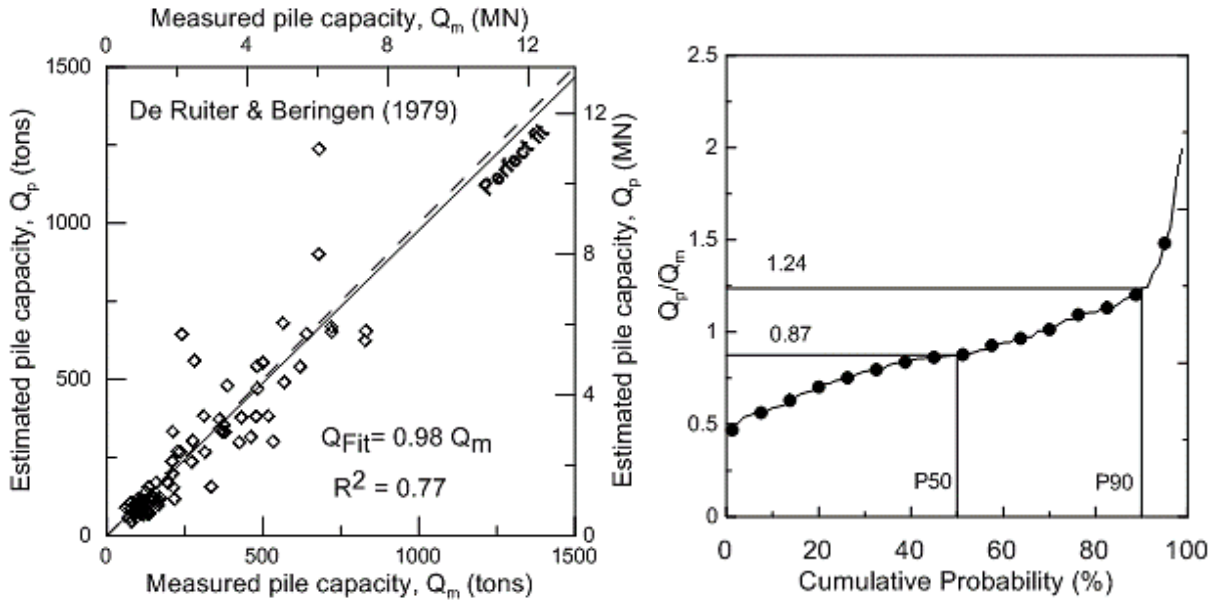


Figure 275. Predicted versus measured ultimate capacity and cumulative probability for Philipponnat

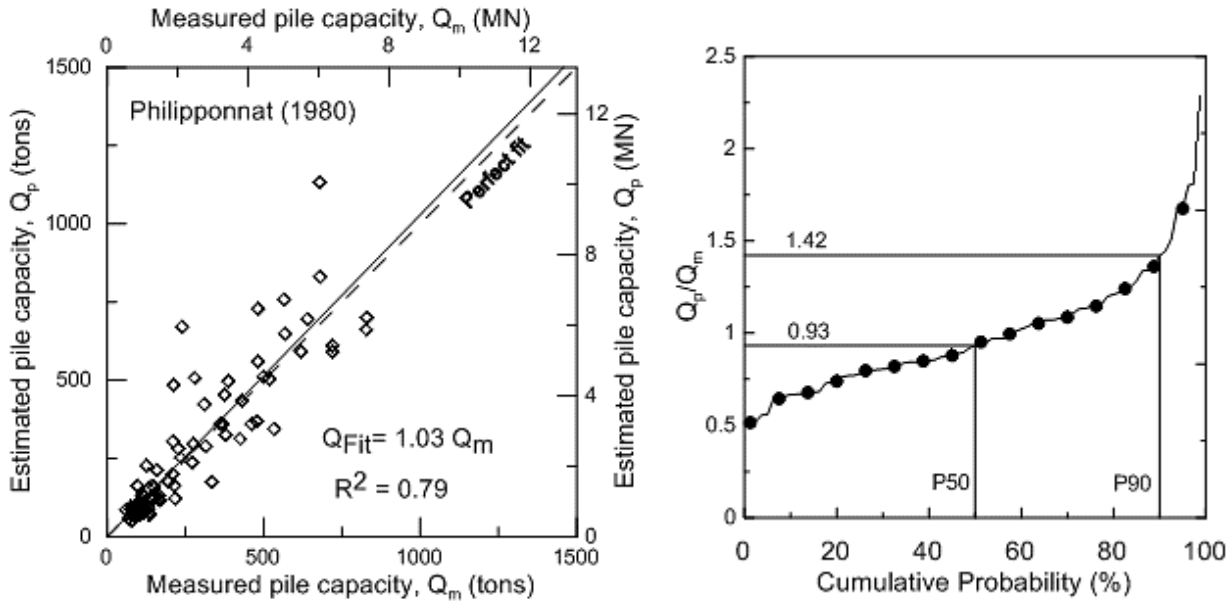


Figure 276. Predicted versus measured ultimate capacity and cumulative probability for Price and Wardle

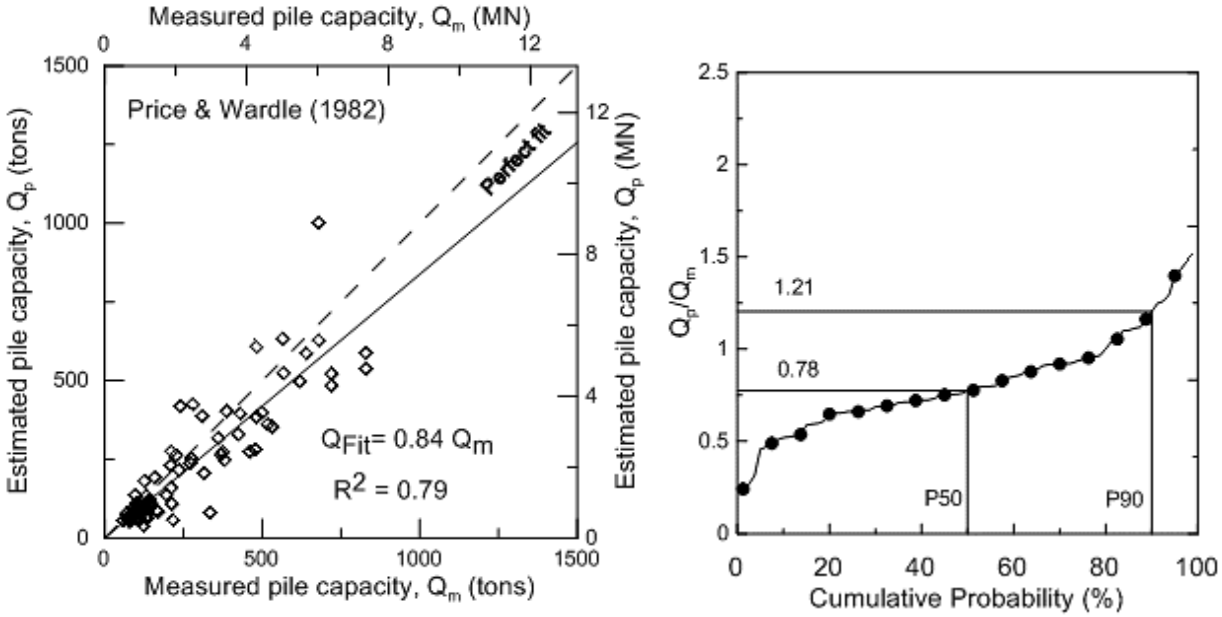


Figure 277. Predicted versus measured ultimate capacity and cumulative probability for Zhou

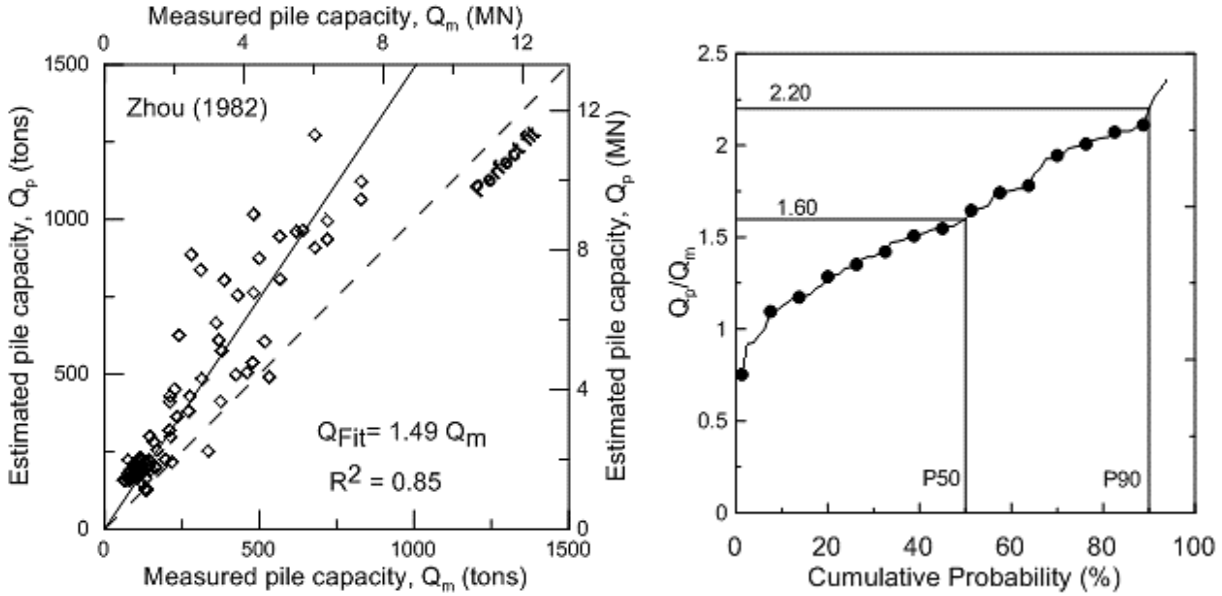


Figure 278. Predicted versus measured ultimate capacity and cumulative probability for Tumay and Fakhroo

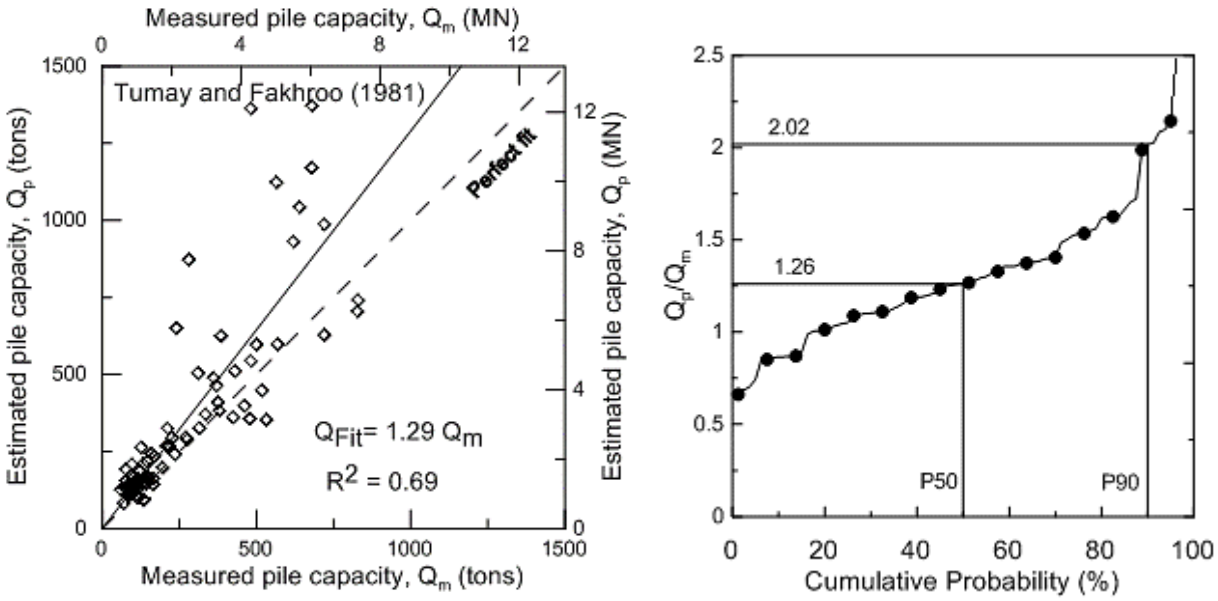


Figure 279. Predicted versus measured ultimate capacity and cumulative probability for UF

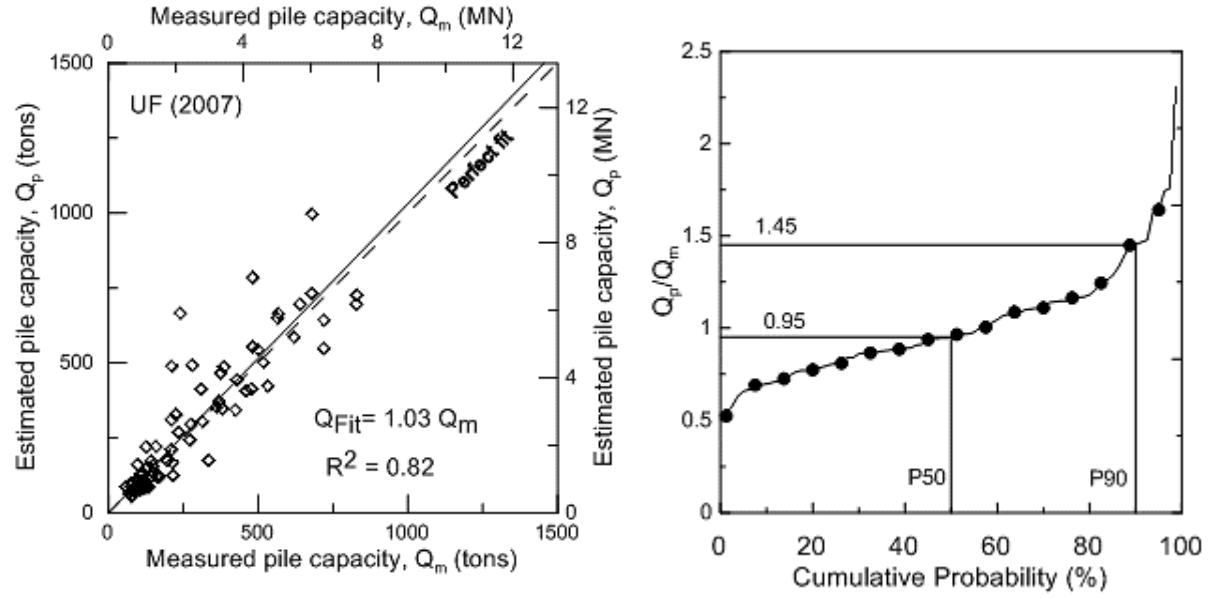


Figure 280. Predicted versus measured ultimate capacity and cumulative probability for Probabilistic

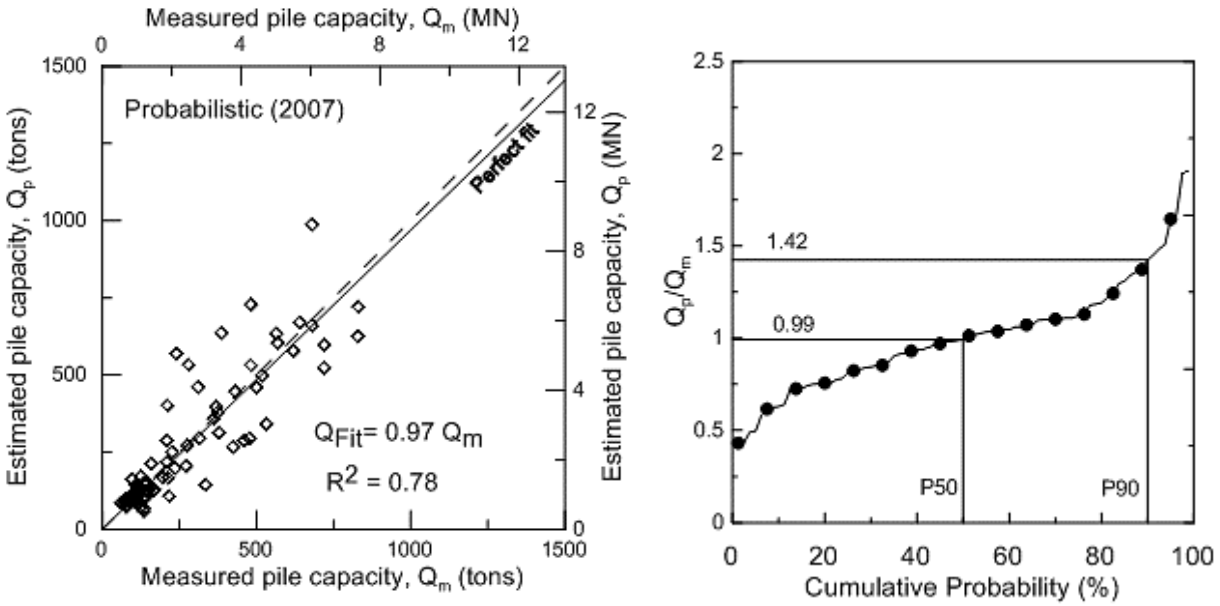


Figure 281. Predicted versus measured ultimate capacity and cumulative probability for Aoki and De Alencar

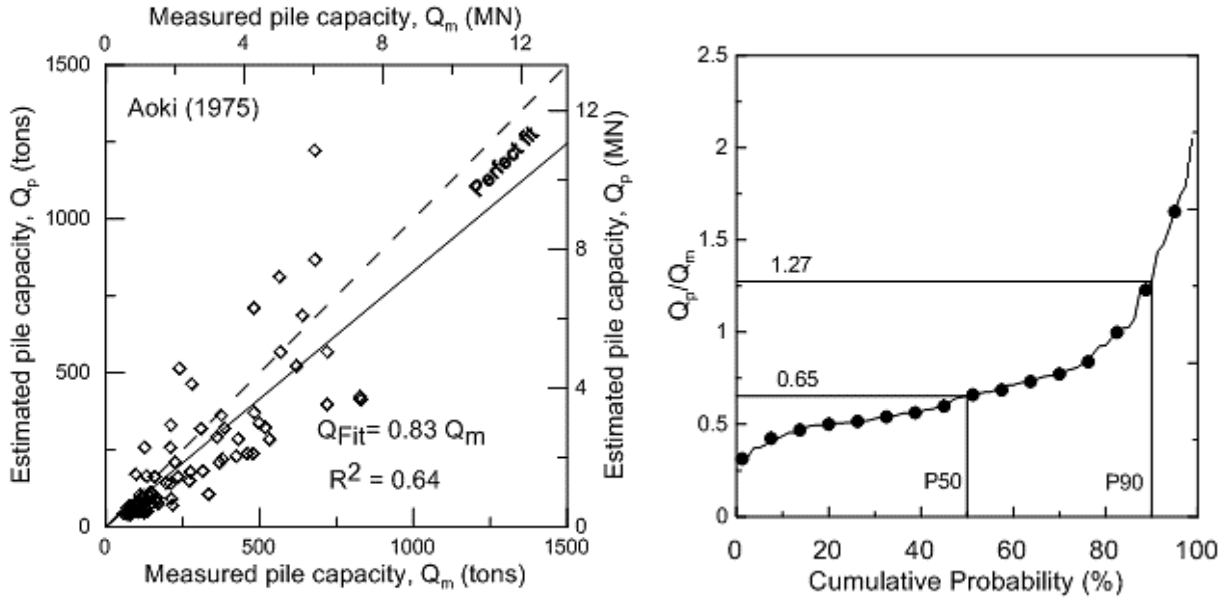


Figure 282. Predicted versus measured ultimate capacity and cumulative probability for Penpile

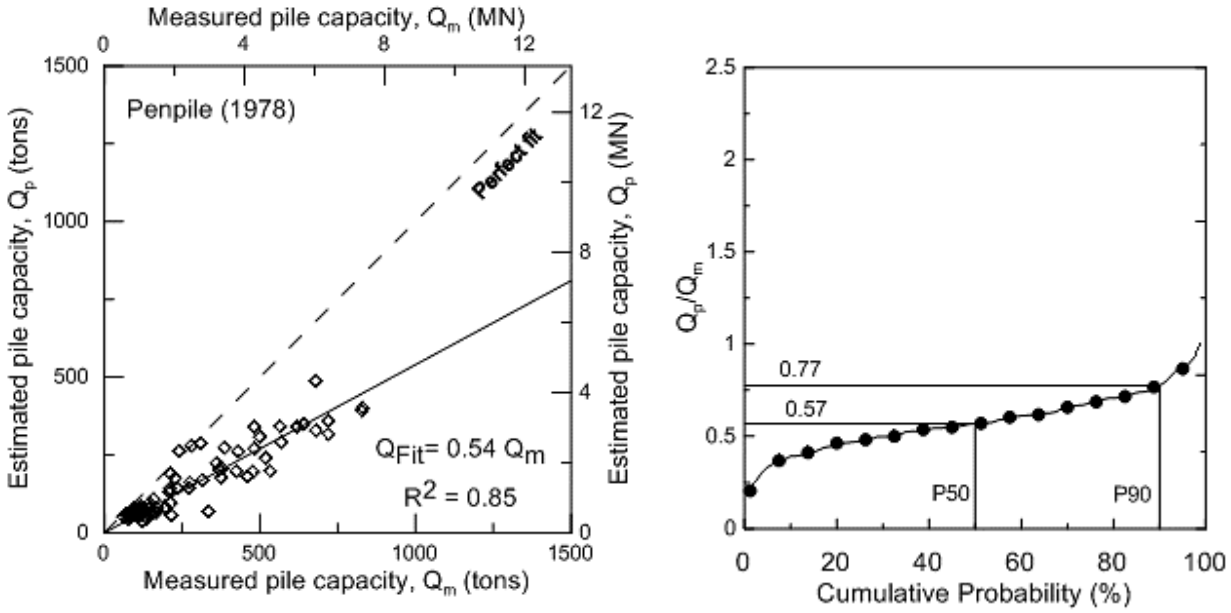


Figure 283. Predicted versus measured ultimate capacity and cumulative probability for NGI

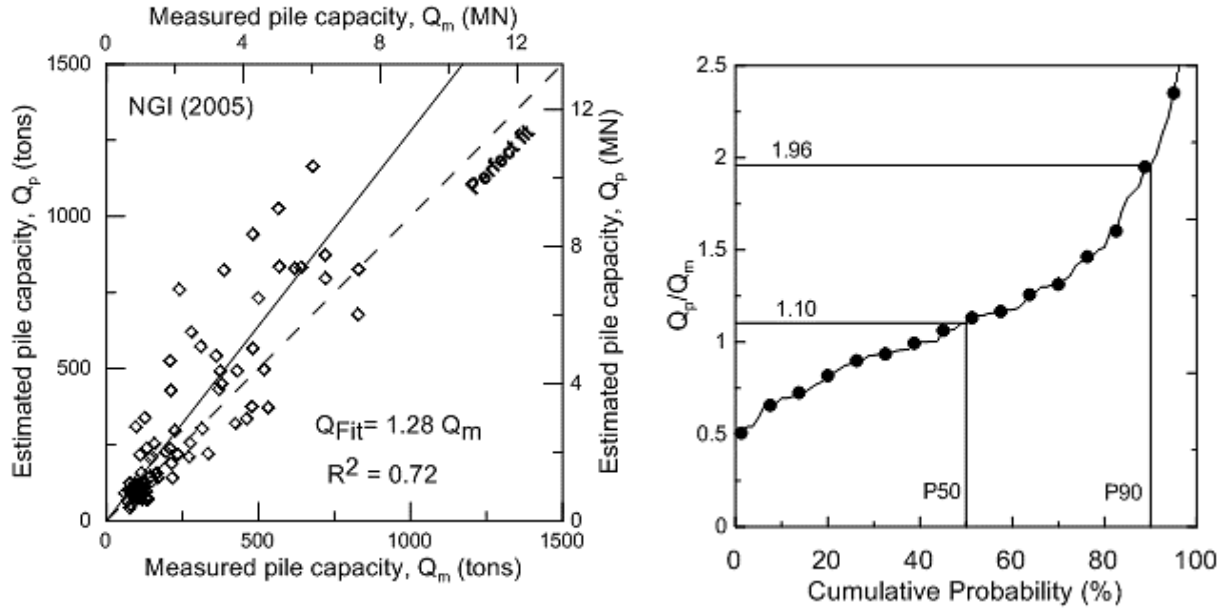


Figure 284. Predicted versus measured ultimate capacity and cumulative probability for ICP

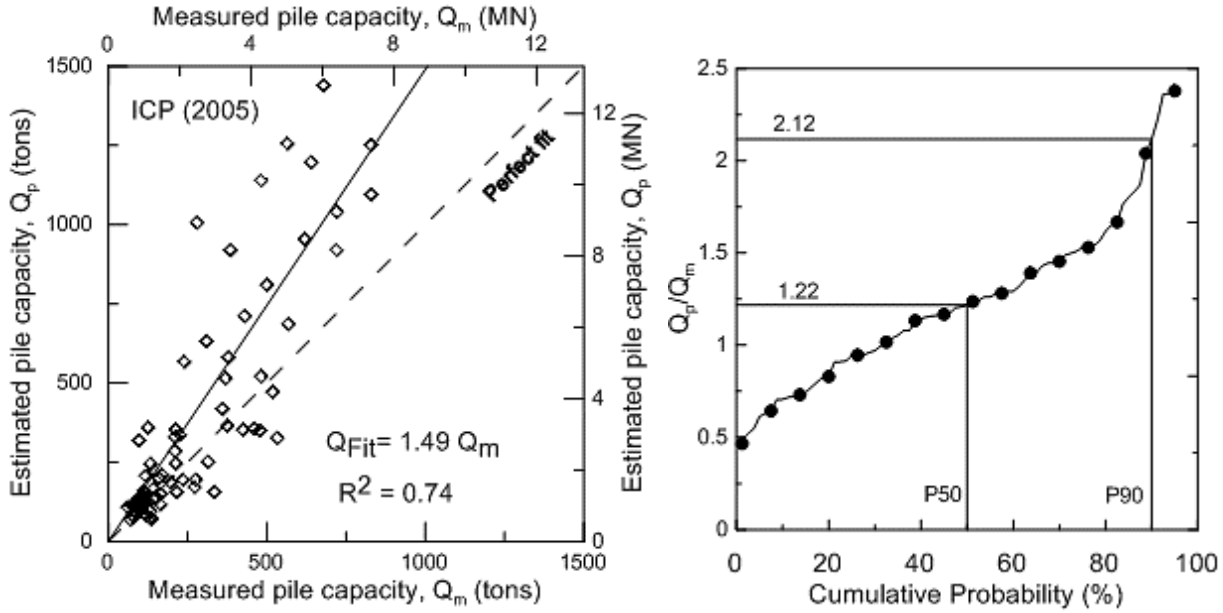


Figure 285. Predicted versus measured ultimate capacity and cumulative probability for UWA

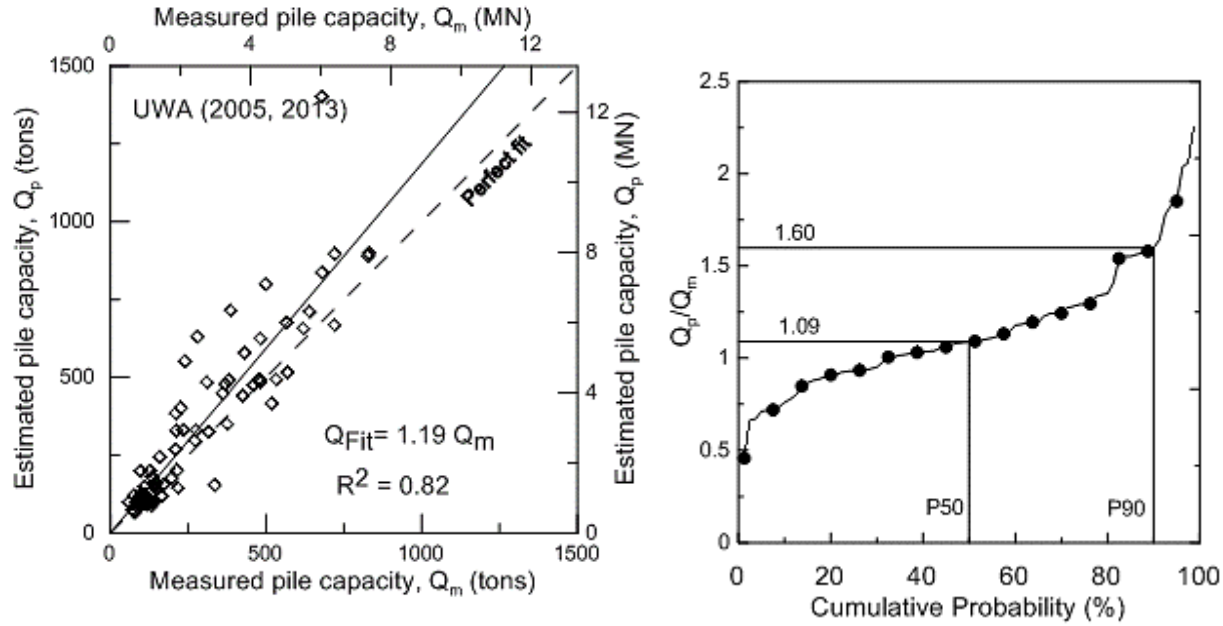


Figure 286. Predicted versus measured ultimate capacity and cumulative probability for CPT2000

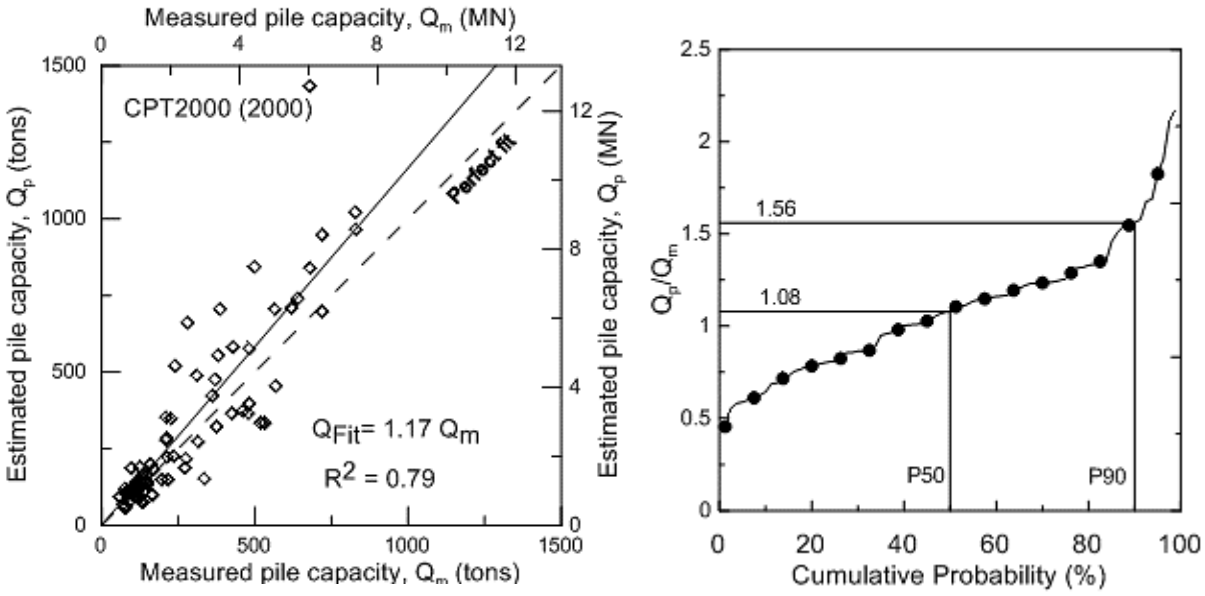


Figure 287. Predicted versus measured ultimate capacity and cumulative probability for Fugro

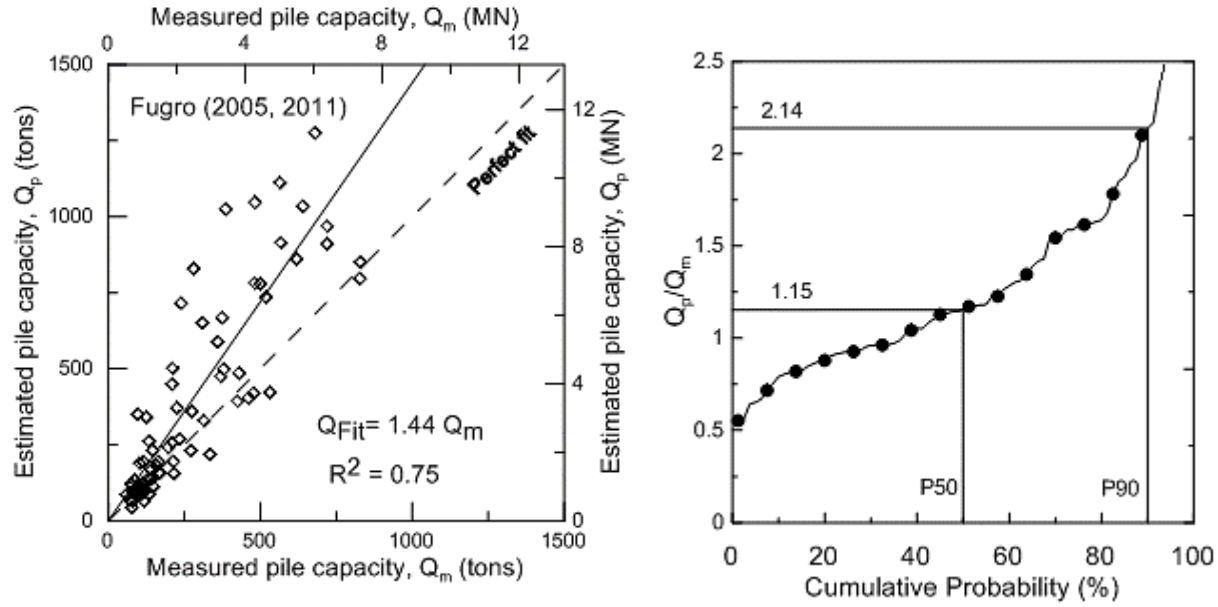


Figure 288. Predicted versus measured ultimate capacity and cumulative probability for Purdue

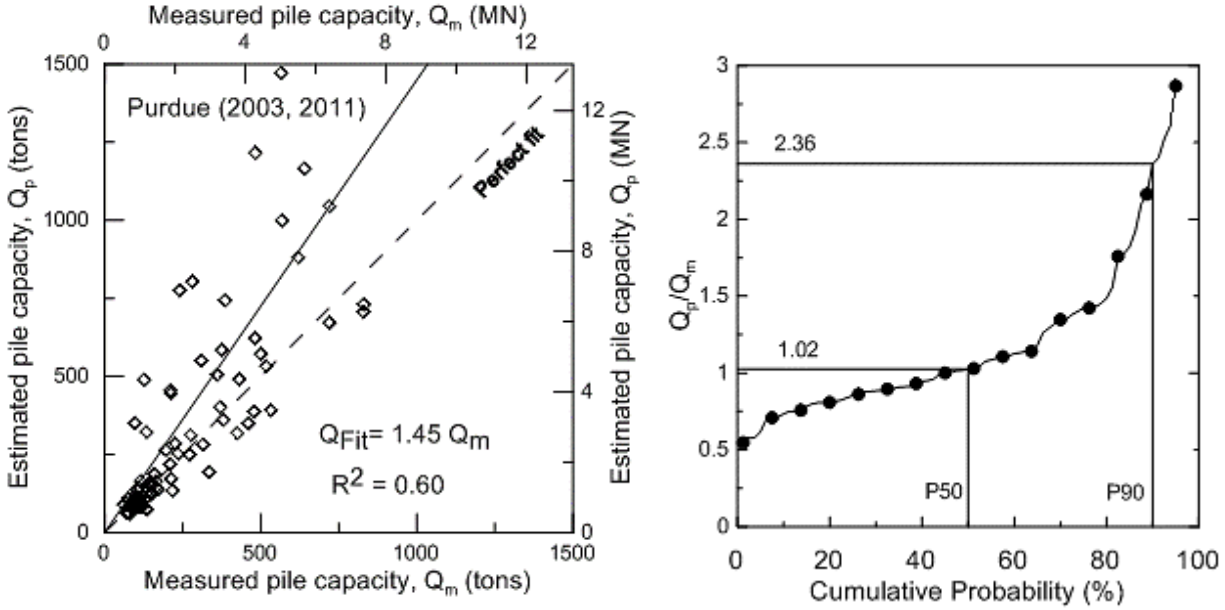


Figure 289. Predicted versus measured ultimate capacity and cumulative probability for Togliani

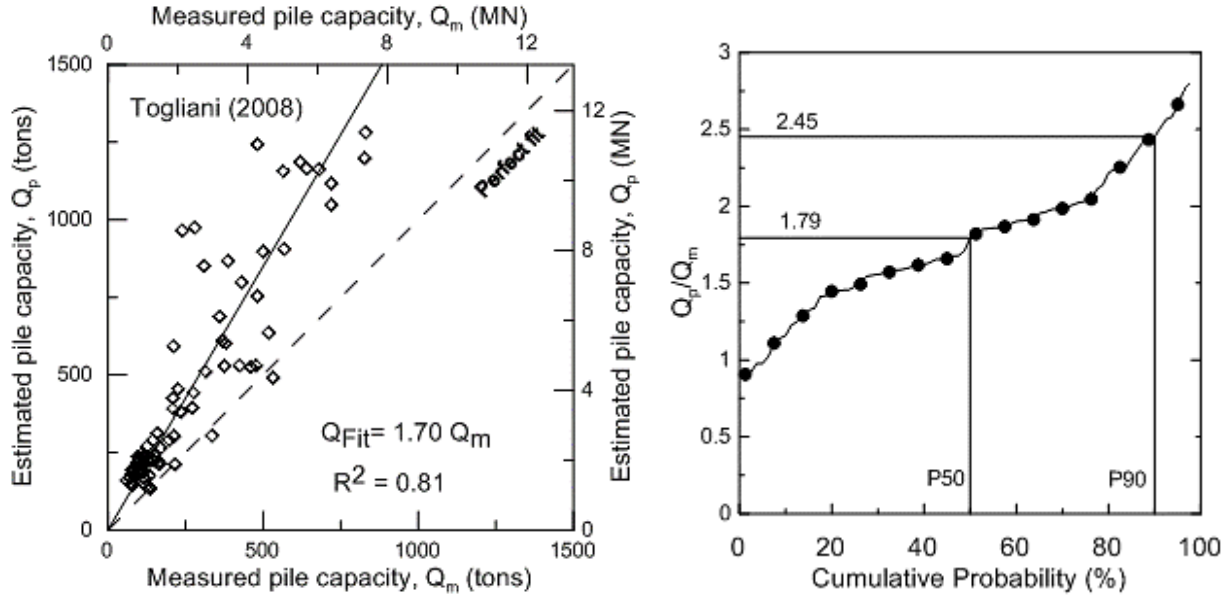


Figure 290. Predicted versus measured ultimate capacity and cumulative probability for ERTC3

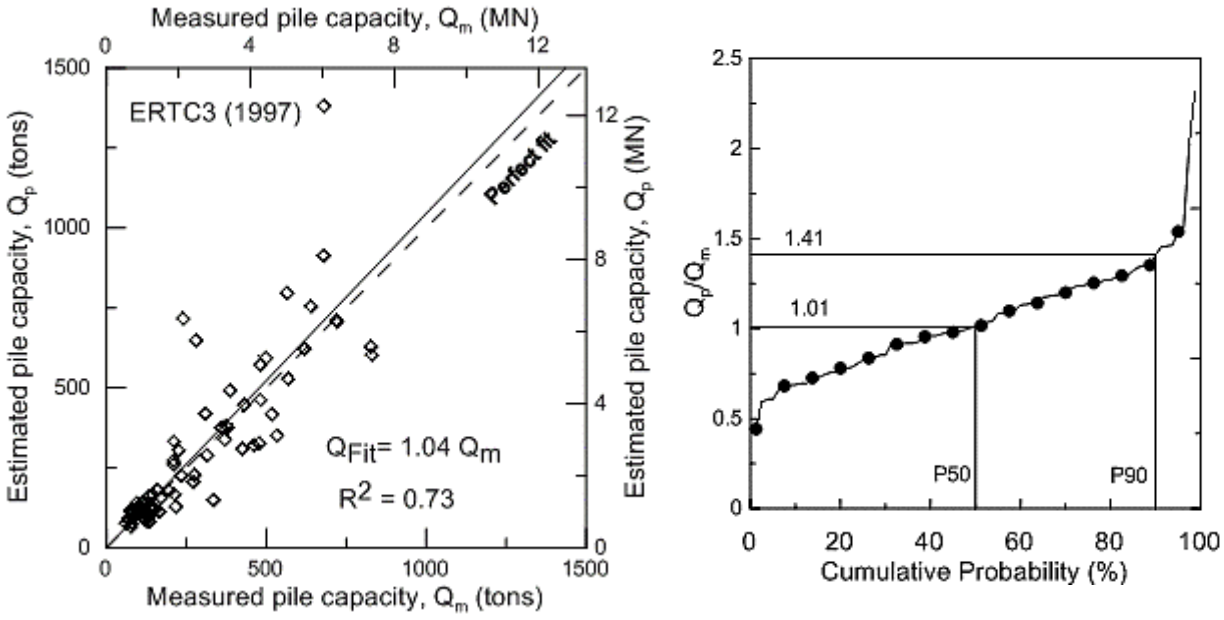


Figure 291. Predicted versus measured ultimate capacity and cumulative probability for Eurocode 7

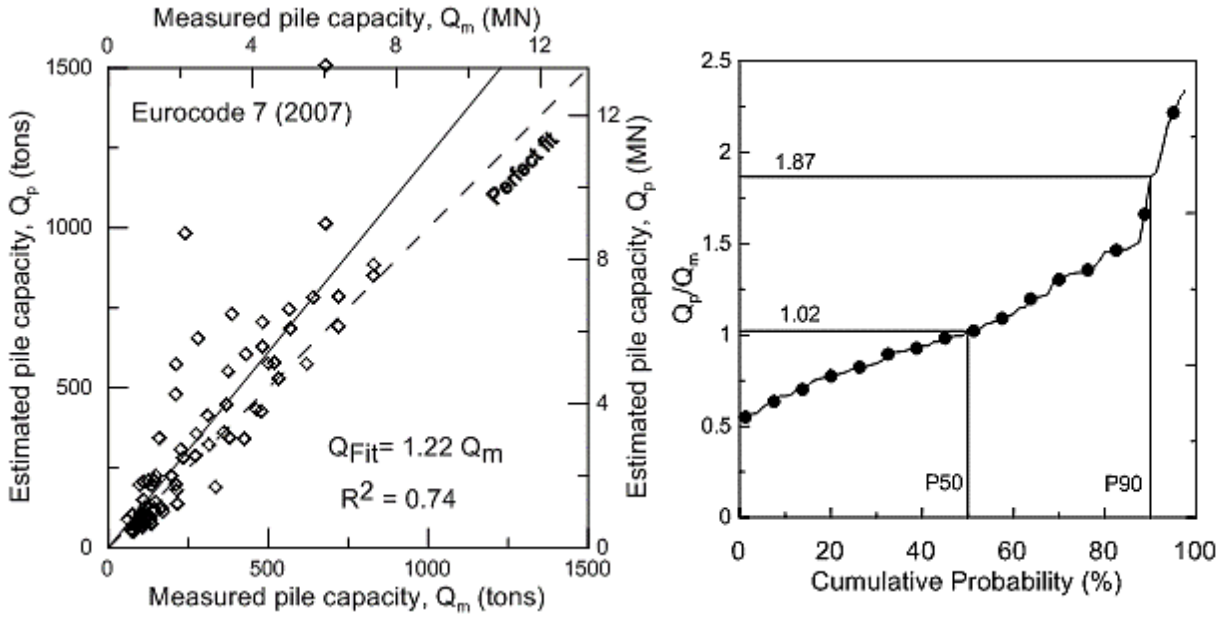


Figure 292. Predicted versus measured ultimate capacity and cumulative probability for German