# IMPROVING THE PREDICTION FOR THE ULTIMATE AXIAL CAPACITY OF DRIVEN PILES IN ONTARIO SOILS WITH A GENETIC ALGORITHM

by

Markus Jesswein

B.Eng, Ryerson University, 2016

A thesis

presented to Ryerson University

in partial fulfillment of the

requirements for the degree of

Master of Applied Science

in the program of

Civil Engineering

Toronto, Ontario, Canada, 2018

© Markus Jesswein, 2018

# AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

IMPROVING THE PREDICTION FOR THE ULTIMATE AXIAL CAPACITY OF DRIVEN PILES IN ONTARIO SOILS WITH A GENETIC ALGORITHM

Markus Jesswein

Master of Applied Science, 2018

Department of Civil Engineering

Ryerson University

# ABSTRACT

A genetic algorithm (GA) was developed to improve predictions for the ultimate axial capacity of driven piles in Ontario soils. Challenges arise to accurately predict the ultimate capacity due to many influential factors, such as the ground conditions, installation method, and pile geometry. A total of 43 piles (H or pipe piles) were collected from the Ministry of Transportation of Ontario. Side and tip resistances were extracted from piles subjected to extension and compression load tests. The soil measurements and pile resistances were regressed with a statistical analysis and GA, and the developed relationships were compared to existing design methods. On average, existing design methods overestimated the capacity by a factor of 1.16 to 3.00. The proposed correlations were slightly conservative with the capacity but provided errors within ± 30 % of the measured side resistance. The new design methods from the GA offer substantial improvements for pile design.

# ACKNOWLEDGEMENTS

The presented research was made possible with generous funding and resources provided by the Ministry of Transportation of Ontario (MTO) and support by Arup Canada Inc. In addition, the National Sciences and Engineering Research Council of Canada (NSERC) granted generous funding and a master's scholarship.

A special thanks to Dr. Jinyuan Liu, my supervisor, for providing this unique opportunity and for offering recommendations, ideas, and life lessons. His time and guidance has truly led to the enhancement of this project and related applications and papers. In all, learning from his experience was enjoyable.

I would like to thank Mr. David Staseff and Ms. Minkyung Kwak from MTO for collaborating and sharing the database of pile load tests. This research completely depended on the available data, which took decades to collect.

I would like to express my gratitude to have the opportunity to collaborate with Ms. Mei Cheong for this research and during an internship at Arup. She offered many inspiring ideas for future research. Also, thank you to Mr. Colin McCreath, Mr. Evan Ma, and the many friendly people that I met and worked with at Arup.

Also, I appreciate the assistance from several people: Mr. Alexandre Almeida, Mr. Andy Lai, Mr. Filipe Batista, Ms. Rebecca Alain, Ms. Alexis Wei, Mr. Jamie Stein, Ms. Chantel Yung, Ms. Maeeda Khan, and Ms. Maribel Castro for their help with the collection of data. In the end, the development of the project was a collaborative and collective process.

I am grateful for my parents and their support during my entire education and life.

In addition, I wish the best and a prosperous life to everyone that was a part of the research team during my time as a MASc student.

# TABLE OF CONTENTS

Author's Declaration ..... ii
Abstract ..... iii
Acknowledgements ..... iv
List of Tables ..... viii
List of Figures ..... ix
1 Introduction ..... 1
1.1 Introduction ..... 1
1.2 Current Pile Design Methods ..... 1
1.3 Inconsistency of Glacial Deposits in Ontario ..... 1
1.4 Database of Pile Load Tests from MTO ..... 2
1.5 Research Objectives ..... 2
1.6 Research Methods ..... 3
1.7 Thesis Outline ..... 3
2 Design Methods for the Axial Capacity ..... 5
2.1 Introduction ..... 5
2.2 Design Methods with Standard Penetration Test (SPT) ..... 5
2.3 Design Methods for a Pile in Cohesive Soils ..... 7
2.3.1 $\alpha$ Method for Unit Side Resistance ..... 8
2.3.2 $\beta$ Method for Unit Side Resistance ..... 10
2.3.3 $\lambda$ Method for Unit Side Resistance ..... 11
2.3.4 Estimation of Unit Tip Resistance ..... 12
2.4 Design Methods for a Pile in Noncohesive Soils ..... 13
2.4.1 Estimation of Unit Side Resistance ..... 13
2.4.2 Estimation of Unit Tip Resistance ..... 14
2.5 Applications of Genetic Algorithms for Pile Capacity ..... 15
2.5.1 Potential Methods to Predict Pile Capacity ..... 15
2.5.2 Genetic Algorithms to Predict Pile Capacity ..... 16
2.6 Obtaining Soil Parameters ..... 18

2.6.1 Field Vane Shear Test (FVST) ..... 18
2.6.2 Standard Penetration Test (SPT) versus Undrained Shear Strength ($\boldsymbol{Cu}$) ..... 20
2.6.3 Friction Angle for Noncohesive Soils ..... 23
2.7 Other Influential Factors on Pile Capacity ..... 24
2.7.1 Pile Geometry and Installation ..... 24
2.7.2 Time Dependency and Setup Time ..... 25
2.8 Summary ..... 25
3 Statistical Analysis and Methodology ..... 28
3.1 Overview of Methodology ..... 28
3.2 Selected Sites and Pile Load Tests ..... 28
3.3 Resistances Interpreted from Pile Load Tests ..... 31
3.4 Statistical Analysis with Pile Resistances and Soil Parameters ..... 34
3.4.1 Evaluating Soil Strength Measurements ..... 34
3.4.2 Correlations with Side Resistance ..... 36
3.4.3 Correlations with Tip Resistance ..... 41
4 Predictions by A Genetic Algorithm ..... 46
4.1 Introduction ..... 46
4.1.1 Creation of Chromosomes ..... 48
4.1.2 Evaluation of Chromosomes ..... 49
4.1.3 Selection of Parents ..... 50
4.1.4 Crossover and Mutation ..... 50
4.1.5 Constant Refinement ..... 52
4.2 Correlations between Side Resistance and SPT N-Values ..... 52
4.3 Correlations between Side Resistance and Undrained Shear Strength ..... 58
4.4 Performance of Existing Design Methods ..... 61
4.4.1 Direct Methods with SPT ..... 61
4.4.2 Direct and Indirect Methods for Cohesive Soils ..... 63
4.4.3 Indirect Methods for Noncohesive Soils ..... 71
4.5 Summary and Discussions ..... 74
5 Conclusions and Future Research ..... 76
5.1 Summary ..... 76

5.2 Conclusions and Recommendations ..... 76
5.3 Limitations and Future Work ..... 78
Appendix ..... 79
Bibliography ..... 110

# LIST OF TABLES

Table 2-1: Existing Design Methods for Pile Side Resistance ..... 6
Table 2-2: Existing Design Methods for Pile Tip Resistance ..... 7
Table 2-3: Suggested A Coefficients for Various Soil Types (Sivrikaya and Toğrol, 2006) ..... 22
Table 3-1: Details on the Studied Piles ..... 29
Table 4-1: Settings for the Genetic Algorithm ..... 47
Table 4-2: Variables Included in the GA Analysis ..... 47
Table 4-3: Protection Instructions for Operators (Modified from Brameier & Banzhaf, 2007)... 50
Table 4-4: Percentage of Total Functions that Apply Each Variable ..... 54
Table 4-5: Percentage of Total Functions that Apply Each Variable ..... 59
Table 4-6: Descriptive Statistics of the Predicted to Measured Side Resistance Ratio for SPT Design Methods ..... 63
Table 4-7: Studied Design Methods for Side Resistance in Cohesive Soils ..... 64
Table 4-8: Studied Empirical Methods for Cohesive Soils ..... 64
Table 4-9: Descriptive Statistics of the Predicted to Measured Capacity Ratios by CGS (2006) 67
Table 4-10: Descriptive Statistics of the Predicted to Measured Capacity Ratios by Karlsrud *et al.* (2005) ..... 68
Table 4-11: Descriptive Statistics of the Predicted to Measured Capacity Ratios by Kolk and Van Der Velde (1996) ..... 69
Table 4-12: Descriptive Statistics of the Predicted to Measured Capacity Ratios by Patrizi and Burland (2001) ..... 69
Table 4-13: Descriptive Statistics of the Predicted to Measured Capacity Ratios by Proposed Design Methods for Piles in Cohesive Soils ..... 70
Table 4-14: Studied Empirical Methods for Noncohesive Soils ..... 71
Table 4-15: Descriptive Statistics of the Predicted to Measured Capacity Ratios by Berezantzev *et al.* (1961) ..... 73

# LIST OF FIGURES

Figure 2-1: Relationship of Adhesion Factor and Undrained Shear Strength (Kulhawy & Jackson, 1989) ..... 8
Figure 2-2: Relationship of Adhesion Factor to Strength Ratio (Karlsrud *et al.*, 2005)..... 10
Figure 2-3: Bearing Capacity Factor by Berezantzev *et al.* (1961) (Berezantzev *et al.*, 1961).... 14
Figure 2-4: Process of a Simple Genetic Algorithm ..... 17
Figure 2-5: Relationship of Estimating Correction Factors for FVST from Bjerrum (1972) (CGS, 2006) ..... 19
Figure 2-6: Relationship of Estimating Correction Factors for FVST from Aas *et al.* (1986) (CGS, 2006) ..... 19
Figure 2-7: Relationships between N-Values and $Cu$ (Kulhawy & Mayne, 1990) ..... 21
Figure 2-8: Relation of N-Values and Friction Angles by Peck *et al.* (1974) (Kulhawy & Mayne, 1990) ..... 23
Figure 3-1: Location of Studied Sites ..... 30
Figure 3-2: Soil Measurements at Site 22 ..... 31
Figure 3-3: Load, Displacement, and Time Measurements from Pile 4 at Site 22 ..... 32
Figure 3-4: Example of Load-Displacement Curves for Pile 4 at Site 22 ..... 33
Figure 3-5: Example of the De Beer Method for Pile 4 at Site 22 ..... 33
Figure 3-6: Comparison of UU Triaxial to UCS ..... 34
Figure 3-7: Comparison of UU Triaxial to FVST ..... 35
Figure 3-8: Influence of PI on FVST Correction Factor ..... 35
Figure 3-9: Influence of PI on Relation of $Cu$ by UU Triaxial and Corrected SPT N-Value ..... 36
Figure 3-10: Distribution of the Contribution from the Side Resistance to the Pile Capacity by Soil Type ..... 37
Figure 3-11: Relationship between Unit Side Resistance and Mean Effective Stress ..... 38
Figure 3-12: Relationship between Unit Side Resistance and Pile Embedment Length ..... 38
Figure 3-13: Relationship between Unit Side Resistance and Mean Corrected N-Value ..... 39
Figure 3-14: Relationship between Unit Side Resistance and Mean $Cu$ by UCS ..... 40
Figure 3-15: Relationship between Unit Side Resistance and Mean Liquidity Index ..... 40
Figure 3-16: Distribution of the Unit Tip Resistance by Soil Type ..... 41
Figure 3-17: Influence of Effective Stress on Unit Tip Resistance ..... 42
Figure 3-18: Influence of Pile Length on Unit Tip Resistance ..... 42
Figure 3-19: Influence of Soil Content on Unit Tip Resistance ..... 43
Figure 3-20: Relationship between Unit Tip Resistance and Uncorrected N-Value ..... 44
Figure 3-21: Relationship between Unit Tip Resistance and Corrected N-Value ..... 44
Figure 3-22: Relationship between Unit Tip Resistance and $Cu$ by UU Triaxial ..... 45
Figure 4-1: Process of the Developed Genetic Algorithm ..... 46
Figure 4-2: Example of Decoding the Chromosome with MEP ..... 49
Figure 4-3: Example of Uniform Crossover ..... 51

Figure 4-4: Example of Uniform Crossover with Brood Recombination..... 51
Figure 4-5: Fitness Performance with Pipe Piles..... 53
Figure 4-6: Fitness Performance with H Piles ..... 54
Figure 4-7: Comparison of the Measured and Predicted Side Resistance by the GA for Pipe Piles ..... 55
Figure 4-8: Comparison of the Measured and Predicted Side Resistance by the GA for H Piles 56
Figure 4-9: Pareto Front from GA Results for Pipe Piles..... 57
Figure 4-10: Pareto Front from GA Results for H Piles ..... 57
Figure 4-11: Fitness Performance with Piles in Cohesive Soils ..... 59
Figure 4-12: Comparison of the Measured and Predicted Side Resistance by the GA for Piles in Cohesive Soils..... 60
Figure 4-13: Pareto Front from GA Results for Piles in Cohesive Soils ..... 60
Figure 4-14: Comparison of Measured and Predicted Side Resistances by Shioi and Fukui (1982) ..... 61
Figure 4-15: Comparison of Measured and Predicted Side Resistances by Decourt (1982)..... 62
Figure 4-16: Comparison of Measured and Predicted Side Resistances by Brown (2001)..... 62
Figure 4-17: Comparison of Measured and Predicted Capacities by a) CGS (2006); b) Karlsrud *et al.* (2005); c) Kolk and Van der Velde (1996); and d) Patrizi and Burland (2001)... 66
Figure 4-18: Comparison of Measured and Predicted Capacities by the Proposed Method from the a) Statistical Analysis and b) GA ..... 67
Figure 4-19: Comparison of Measured and Predicted Capacities by the Berezantzev *et al.* (1961) Method ..... 72
Figure 4-20: Comparison of Measured and Predicted Capacities by the Proposed SPT Correlations from the GA ..... 73

# 1 INTRODUCTION

## 1.1 INTRODUCTION

The province of Ontario plans to invest $ 190 billion in public infrastructure by 2030 (Ontario, 2018). Many of the effected structures, specifically bridges and buildings, are supported by deep foundations, but accurately predicting the capacity of a pile is a challenge due to the influences from the installation method, pile geometry, and soil properties. In particular, glacial tills are famous for their inconsistent material properties and are commonly found in Ontario (Barnett, 1992). Since the soil conditions are rarely consistent at a site, designers experience challenges to characterize the ground conditions and evaluate the soil parameters. In addition, the standard penetration test (SPT) is commonly used in site investigations because it is a cheap, quick, and versatile measurement technique for gravel and cobble rich soils. However, this simple test is unreliable for several reasons: the test does not provide continuous data; the sampler can erroneously give high N-values with obstructions, such as gravels and boulders; and the accuracy of the results depend on the operator. For this investigation, the goal was to improve the predictability of the ultimate axial capacity of driven piles in Ontario. Greater accuracy and insight on foundation designs will deliver cost-savings and optimize infrastructure development within Ontario.

## 1.2 CURRENT PILE DESIGN METHODS

The axial capacity of a pile is generally composed of two mechanisms: the side resistance is the friction between the soil and wall of a pile, and the tip resistance is generated by the strength of the soil at the pile tip. Generally, the design methods are further divided into in-situ correlations, total stress methods, and effective stress methods. In this study, in-situ correlations directly relate the N-values, or blow counts, from SPT to the pile resistances. The total stress methods, such as the $\alpha$ Method, are intended for the short-term conditions of cohesive soils. The effective stress methods are usually called $\beta$ Methods and are for noncohesive soils or the long-term conditions of cohesive soils. Numerous design methods have been developed over the years, and they include a range of soil properties and design situations. After developing an approach with a machine learning technique, several of these existing methods were evaluated for their accuracy.

## 1.3 INCONSISTENCY OF GLACIAL DEPOSITS IN ONTARIO

Since glaciers once covered most of North America, glacial deposits will be heavily encountered within Ontario. Glacial till is sediment that was eroded, transported, and deposited by glaciers and usually contains an inconsistent mixture of particle sizes, mineral types, and weathered rock types

(Barnett, 1992). According to Barnett (1992), tills are often compact or overconsolidated (OC) due to the weight of the previously overlying glaciers. In Ontario, normally consolidated (NC) or lightly OC tills can be found, but they would exist by areas where drainage of meltwater occurred, such as the shores of the Great Lakes (Barnett, 1992). The different origins and processes that created tills lead to a variety of soil characteristics.

The formation of till in Ontario can help to predict the soil content. Precambrian igneous and metamorphic rocks composed a large portion of the Canadian Shield. If these igneous and metamorphic rocks eroded to form the till, the till can contain over 70 % sand and less than 5 % clay (Barnett, 1992). Tills likely range from non-plastic, especially from eroded Precambrian rocks (Barnett, 1992), to low plasticities due to the low silt content. However, tills from carbonate rocks, limestones, and dolostones can have a content of 40 to 50 % silt and 15 to 30 % clay (Barnett, 1992). Well graded sediments from carbonate rocks leads to compact to very compact tills (Barnett, 1992). Tills located deeper into the ground are expected to be more compact, and layers overlying bedrock are typically thin and contain boulders and gravels (Barnett, 1992). Fine-grained tills were likely formed by the addition of glaciolacustrine deposits and are composed of clay to a maximum of approximately 55 % (Barnett, 1992). After observing these formations, tills generally range in cohesive and noncohesive contents.

## 1.4 DATABASE OF PILE LOAD TESTS FROM MTO

Pile load tests are costly and time-consuming, but the full-scale tests are reliable and are commonly applied to verify designs in local ground conditions. For example, Briaud and Tucker (1988) analyzed a database with 91 driven piles, and they discovered design methods can over predict by 20 % on average and have a coefficient of variation (COV) between 40 to 45 %. A database of more than 100 piles was obtained over four decades by the Ministry of Transportation of Ontario (MTO). This database was used in this study, and it contained the results from tests with static axial loads. Axial loads were applied to piles by hydraulic jacks that acted against an anchored reaction frame, weighted box, or weighted platform. These pile load tests are invaluable to provide local or regional geotechnical parameters and are used in this study to develop new design methods for piles in Ontario.

## 1.5 RESEARCH OBJECTIVES

The main objectives of this research are the following:

1. Apply a machine learning method to improve predictions for the ultimate axial capacity of piles. Specifically, a genetic algorithm (GA) needs to regress several variables at once; predict the capacity of piles in varying soil conditions with multiple soil measurements at varying depths; and provide a practical equation for design. A GA was chosen for several reasons. First, it does not need knowledge of the problem domain (Negnevitsky, 2005) and,

therefore, is an efficient optimization method at handling several variables compared to traditional statistics. Instead, a GA requires a fitness function (Negnevitsky, 2005). The fitness function allows the GA to evaluate many potential solutions on a variety of criteria and to find a preferred solution (Negnevitsky, 2005). Artificial neural networks (ANN) are also efficient at finding relationships between data; however, due to the construction of the neural network, it is often challenging to obtain a simple, logical, or practical mathematical function to describe the determined relation. Typically, the relationships from an ANN are hidden, and the ANN by itself is the final deliverable. For this thesis, the goal includes providing a practical function, which may combine several variables in a nonlinear and interdependent form.

2. Evaluate the accuracy or existing design methods with pile load tests in Ontario.
3. Propose design methods for driven piles in the soil conditions within Ontario.

## 1.6 RESEARCH METHODS

The focus of this thesis is on the behaviour of driven piles with SPT N-values for various soil types and the undrained shear strength for cohesive soils. Influences were evaluated by statistical approaches, and nonlinear relations were determined with a GA. The accuracy of the developed design methods was compared to current design methods.

First, the following steps were performed to gather data on pile load tests in Ontario:

- Obtain a database of pile load tests and site investigations from MTO;
- Focus on pile tests and properties involving steel H piles and steel pipe piles with both extension and compression load tests and shear strength measurements from the soil.
- Obtain the side resistance, tip resistance, and ultimate capacity from the test results, such as the load-displacement curves; and
- Gather soil data at the testing sites.

Second, predictions for the capacity were improved and verified:

- Conduct a statistical analysis to correlate pile resistances and soil measurements;
- Develop a GA to predict the side resistance of piles with SPT N-values or undrained shear strength measurements; and
- Verify the regressions from the statistical analysis and GA by comparing predictions for the capacity to existing design methods.

## 1.7 THESIS OUTLINE

This thesis involves several aspects on the axial capacity of a driven pile. The formulas, relationships, and theories were explained for the methods that estimate soil parameters and pile

capacity. Both a statistical analysis and machine learning method were applied to develop new correlations. After, existing prediction methods were evaluated against correlations from the statistical analysis and GA.

Chapter One introduces the background, research objectives, and methods applied in this research.

Chapter Two contains a literature review and is composed of several parts. Design methods for the ultimate capacity are discussed in detail with influential factors, and the background to develop a GA is provided. Since different site investigation techniques can be applied, approaches to estimate the soil shear strength are included.

Chapter Three presents the methodology and results of a statistical analysis with soil conditions and pile resistances.

Chapter Four provides the GA that predicted the side resistance of piles. New design methods were proposed by the GA, and the accuracy of their predictions are compared to existing design methods.

Chapter Five summarizes the final discussions and conclusions of the investigation. The best design methods are proposed to determine the axial capacity, and recommendations for future research and GA development are included.

# 2 DESIGN METHODS FOR THE AXIAL CAPACITY

## 2.1 INTRODUCTION

As a compressive load is applied to a pile, the soil will resist the loading with two components: the side resistance ($Qs$) is the friction between in the soil and pile walls, and the tip resistance ($Qp$) is a compressive force by the soil at the pile tip. Thus, while subtracting the weight of the pile ($W_p$), the ultimate axial capacity ($Qu$) of a pile is expressed below:

$$Qu = Qs + Qp - W_p \qquad \text{Eq. 2-1}$$

The side resistance is dependent on the area of the pile walls ($A_s$), and this area is related to the pile perimeter ($P$) and the pile length ($L$). The tip resistance relies on cross-sectional area at the pile tip ($A_p$). Equation 2-1 is then modified:

$$Qu = \sum q_s\, A_s + q_p\, A_p - W_p \qquad \text{Eq. 2-2}$$

$$Qu = \sum q_s\, P\, \Delta L + q_p\, A_p - W_p \qquad \text{Eq. 2-3}$$

where $q_s$ is the unit side resistance and $q_p$ is the unit tip resistance. Over the decades, many methods have been proposed to estimate the unit tip and side resistance, and they include a range of soil properties and consider a variety of theories. The following sections explore the related concepts and design methods.

## 2.2 DESIGN METHODS WITH STANDARD PENETRATION TEST (SPT)

When in-situ tests, such as SPT, are applied at a site, the side resistance of a pile may be determined directly or indirectly with the test measurements. For SPT measurements, direct approaches rely on empirical correlations between the unit side resistance and average N-value ($\bar{N}$) along the pile length:

$$q_s = A\bar{N} + B \qquad \text{Eq. 2-4}$$

Also, the N-value at the pile base is directly related to the unit tip resistance:

$$q_p = K\, N \qquad \text{Eq. 2-5}$$

As shown in Equations 2-4 and 2-5, the correlations with SPT are usually linear and apply $A$, $B$, and $K$ as coefficients. Indirect approaches use SPT N-values to predict the shear strength parameters, and the resulting parameters calculate the side resistance with $\alpha$ or $\beta$ Methods.

Table 2-1: Existing Design Methods for Pile Side Resistance

| Soil Type | Reference | Equation for $q_s$ (kPa) | Remarks |
|---|---|---|---|
| Cohesive | Shioi & Fukui (1982) | $q_s = 9.8\,\bar{N}$ | |
| Noncohesive | Briaud *et al.* (1983) | $q_s = 21.45\,\bar{N}^{0.29}$ | |
| | Meyerhof (1956) | Large Displacement Piles: $q_s = 1.9\,\bar{N} \leq 100$ Low Displacement Piles: $q_s = 1.0\,\bar{N} \leq 100$ | |
| | Shariatmadari *et al.* (2008) | $q_s = 3.65\,\bar{N}$ | $\bar{N}$ is the geometric average |
| | Shioi & Fukui (1982) | $q_s = 1.9\,\bar{N}$ | |
| All | Aoki & Velloso (1975) | $q_s = Aa\bar{N}/3.5$ | $A$ ranges from 196 for clays to 980 for sands; $a$ ranges from 0.012 for medium sands to 0.04 for clays |
| | Brown (2001) | $q_s = 1.8\,\bar{N} + 25$ | $3 \leq N \leq 50$ |
| | Decourt (1982) | $q_s = 3.3\,\bar{N} + 9.8 \leq 170$ | $3 \leq N \leq 15$ |

As a simple approach, numerous empirical methods have been developed with SPT N-values over the years. As shown in Table 2-1, the empirical correlations differ on the soil conditions. Coefficient *A* is generally lower for sandy soils compared to clayey soils due to the cohesion, and most correlations were developed for either cohesive or noncohesive soils. Some references proposed design methods for all soils, but they are likely limited to the soil conditions within a local region, such as Brown (2001) and Decourt (1982). Many of the design methods were also developed with databases containing various pile types. For instance, Meyerhof (1956; 1976) collected pile load tests from several references on concrete, timber, and steel driven piles. Aoki & Velloso (1975) investigated Franki, precast concrete, and steel piles from Brazil, but the pile material will likely influence the side friction with the soil. In addition, Shariatmadari *et al.* (2008) recommended to apply the geometric average for the N-value over the common arithmetic mean. Unfortunately, the average soil conditions may not accurately represent the soil behaviour.

Table 2-2 contains the current correlations for the tip resistance. Sandy soils tend to have a higher coefficient with SPT N-values compared to silty or clayey soils. Since cohesive soils can vary in plasticity and compressibility, empirical methods are more popular with noncohesive soils. Yet, some sources offer a single linear regression for all soil types.

Table 2-2: Existing Design Methods for Pile Tip Resistance

| Soil Type | Reference | Equation for $q_p$ (kPa) | Remarks |
|---|---|---|---|
| Cohesive | Decourt (1982) | $q_p = 118\ N$ | |
| | Martin *et al.* (1987) | $q_p = 192\ N$ | |
| Noncohesive | Briaud *et al.* (1983) | $q_p = 1891\ N^{0.36}$ | |
| | Decourt (1982) | $q_p = 392\ N$ | |
| | Martin *et al.* (1987) | Silt/Sandy Silt: $q_p = 335\ N$<br>Sand: $q_p = 431\ N$ | |
| | Meyerhof (1976) | $q_p = 38\ (L/D)\ N \leq 383\ N$ | |
| | Shariatmadari *et al.* (2008) | $q_p = 385\ N$ | |
| | Thorburn & MacVicar (1971) | Glacial till and silt:<br>$q_p = 250\ N$ | |
| All | Aoki & Velloso (1975) | $q_p = K\ N\ /\ 1.75$ | $K$ ranges from 196 for clays to 980 for sands |
| | Brown (2001) | $q_p = 170\ N$ | $3 \leq N \leq 50$ |
| | Shioi & Fukui (1982) | If $L/D \geq 5$, $q_p = 287\ N$<br>If $L/D < 5$ (solid piles),<br>$q_p = (100 + 40\ L/D)N$ | $L/D$ is the slenderness ratio, where $D$ is the pile diameter or width |

## 2.3 DESIGN METHODS FOR A PILE IN COHESIVE SOILS

For cohesive soils, design methods are generally separated into total stress methods and effective stress methods, especially if the soil strength is measured with lab tests. Total stress methods, or $\alpha$ Methods, are intended for the short-term conditions of cohesive soils, and the unit side resistance is determined with an adhesion factor, $\alpha$, and the undrained shear strength, $C_u$:

$$q_s = \alpha C_u \qquad \text{Eq. 2-6}$$

Effective stress methods, also called β Methods, are for noncohesive soils or the long-term conditions of cohesive soils:

$$q_s = \beta \sigma' \qquad \text{Eq. 2-7}$$

where $\beta$ is the adhesion factor and $\sigma'$ is the vertical effective stress.

### 2.3.1 $\alpha$ Method for Unit Side Resistance

Tomlinson (1957) was one of the first for suggesting the $\alpha$ Method, and he developed his method by back analyzing the capacity of 56 driven piles with the average cohesion. For steel and concrete piles, the proposed relationships were associated with inconsistent results, but the approach is simple as the side resistance solely relies on the $C_u$. Kulhawy and Jackson (1989) studied a total of 109 concrete bored piles and obtained a relation for the adhesion factor that was very similar to Tomlinson (1957):

$$\alpha = 0.21 + 0.26\left(\frac{100}{C_u}\right) \leq 1 \qquad \text{Eq. 2-8}$$

The approach from Kulhawy and Jackson (1989), shown in Figure 2-1, was later adopted by the Canadian Foundation Engineering Manual (CFEM) (Canadian Geotechnical Society (CGS), 2006).

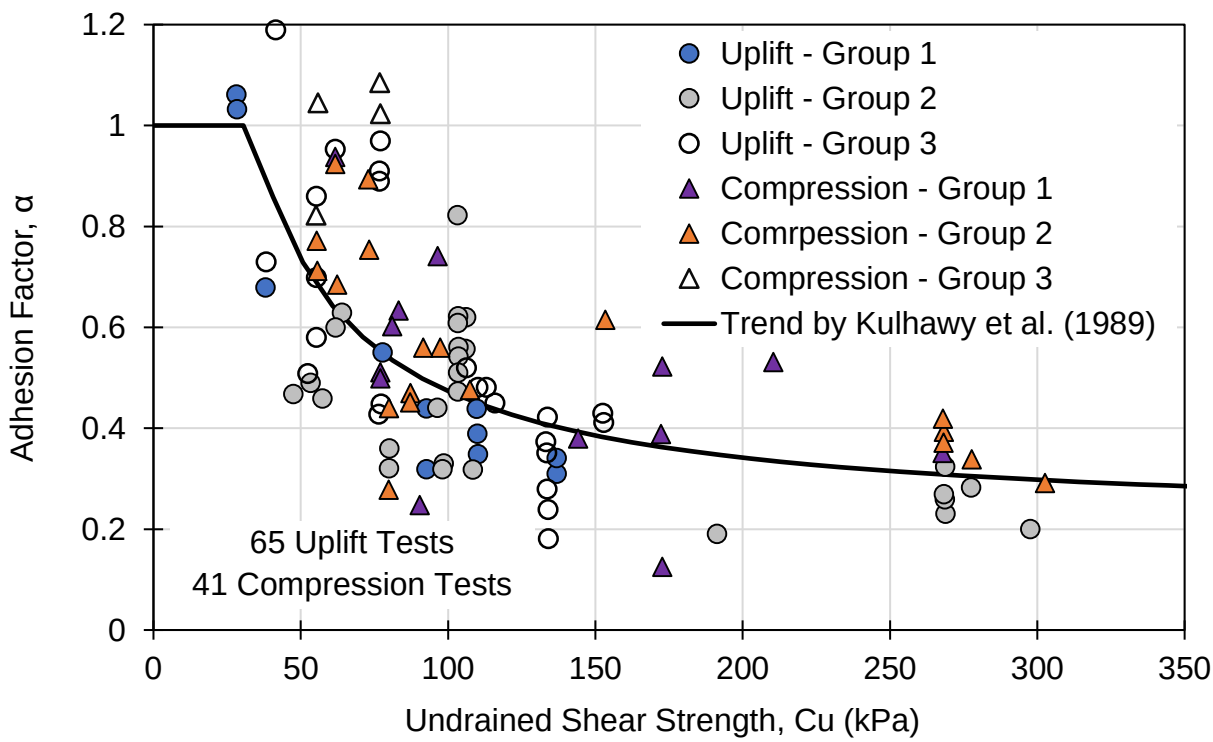



Figure 2-1: Relationship of Adhesion Factor and Undrained Shear Strength (Kulhawy & Jackson, 1989)

Many other references obtained better predictions by considering and including other factors. For example, Dennis and Olson (1983) applied a correction factor to the unit side resistance for the

pile length. Semple and Rigden (1986) went farther and included the soil strength ratio ($C_u/\sigma'$) and slenderness ratio (*L/D*), where *L* is the pile length and *D* is the pile diameter. Kolk and Van der Velde (1996) analyzed 26 pile load tests and proposed two equations for the adhesion factor after manipulating several variables with a trial-and-error approach. The first equation was intended for the average soil parameters:

$$\alpha = 0.5 \left(\frac{D}{L}\right)^{0.2} \left(\frac{\sigma'}{C_u}\right)^{0.3} \leq 1 \qquad \text{Eq. 2-9}$$

If the soil profile was divided into multiple layers, the second equation was recommended:

$$\alpha = 0.9 \left(\frac{L-z}{D}\right)^{-0.2} \left(\frac{\sigma'}{C_u}\right)^{0.3} \leq 1 \qquad \text{Eq. 2-10}$$

In the above equation, $z$ is the depth below the ground surface to the centre of a soil layer. The method provided reliable results for piles in OC clays, but the side resistance was generally overestimated for slender piles in NC clays (Kolk and Van der Velde, 1996). The plasticity index (*PI*) generated scattered relationships, and Kolk and Van der Velde (1996) suggested the *PI* does not have a significant influence on the pile capacity. Karlsrud *et al.* (2005) re-evaluated the influence of the soil plasticity with open-ended and closed-ended pile piles, and they noticed the adhesion factor decreased with a lower *PI*. In Figure 2-2, the proposed design method mainly focuses on NC clays and indicates a sensitive relation for low plastic clays. The *PI* is limited to a minimum of 10 because this was the lowest value found in the database. Karlsrud *et al.* (2005) also included several correction factors to calculate the adhesion factor: the testing technique to obtain $C_u$; duration of time between installing and testing a pile; and pile tip conditions. This collection of methods shows that the soil-pile interface, especially for clayey soils, may rely on many variables.

The unit side resistance, as proposed by Karlsrud *et al.* (2005), is determined by several cases. If $C_u/\sigma' < 0.25$, the following equations are applied:

$$q_s = \alpha\, C_u \qquad \text{Eq. 2-11}$$

$$\alpha = 0.32\,(PI - 10)^{0.3} \qquad \text{Eq. 2-12}$$

$\alpha$ ranges from 0.2 to 1. If $C_u/\sigma' > 1$, the unit side resistance is calculated by the following:

$$q_s = \alpha\, C_u\, F_{Tip} \qquad \text{Eq. 2-13}$$

$$\alpha = 0.5 \left(\frac{C_u}{\sigma'}\right)^{-0.3} \qquad \text{Eq. 2-14}$$

Here, $F_{Tip}$ is a modification factor that depends on the end conditions of a pipe pile. For open-ended piles, the factor is equal to one. For a closed pile, the factor is calculated:

$$F_{Tip} = 0.8 + 0.2\left(\frac{C_u}{\sigma'}\right)^{0.5} \qquad \text{Eq. 2-15}$$

$F_{Tip}$ ranges from 1 to 1.25. For 0.25 < $C_u$/σ'< 1, the adhesion factor is linearly interpolated between the other two conditions. In general, the method by Karlsrud *et al.* (2005) is complex as it depends on several variables and functions.

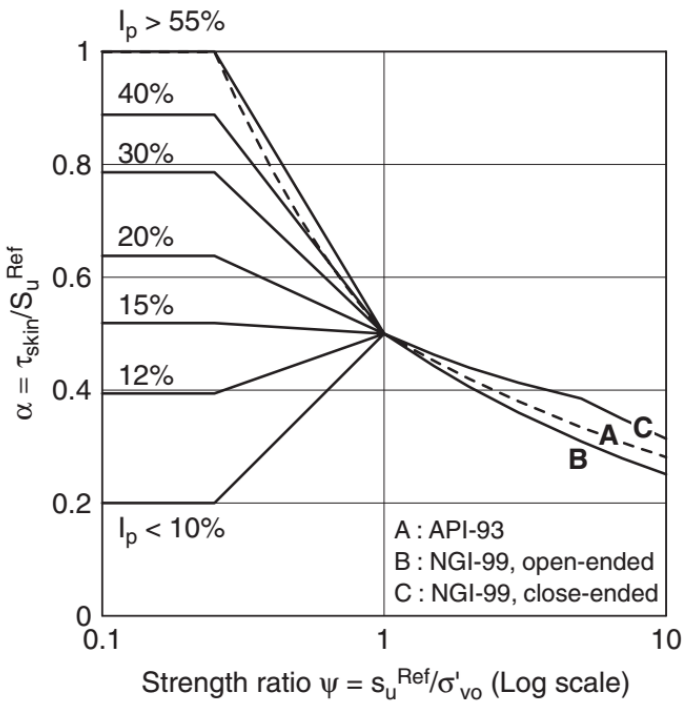



Figure 2-2: Relationship of Adhesion Factor to Strength Ratio (Karlsrud *et al*., 2005)

### 2.3.2 $\beta$ Method for Unit Side Resistance

As shown with Equation 2-7, $\beta$ Methods express the proportionality of the effective stress to the unit side resistance with a $\beta$ coefficient. The coefficient can be a function of the friction angle ($\emptyset'$) since the approach is intended for the drained conditions of soils (Burland, 1973):

$$q_s = K_s \tan(\emptyset')\, \sigma' \qquad \text{Eq. 2-16}$$

where $K_s$ is a coefficient related to the lateral earth pressure coefficient.

For NC cohesive soils, Burland (1973) substituted $K_s$ for the earth pressure at rest ($K_o$) to be conservative.

$$K_s \approx K_o = 1 - \sin(\emptyset') \qquad \text{Eq. 2-17}$$

Burland (1973) also suggested to predict the capacity of driven piles with the residual friction angle, $\emptyset'_r$, instead of the effective friction angle at critical or post-critical conditions. The following equation was the result of the modifications:

$$q_s = (1 - \sin(\emptyset'_r)\,)\,\tan(\emptyset'_r)\,\sigma' \qquad \text{Eq. 2-18}$$

Unfortunately, residual strength parameters may not be commonly collected in Ontario practice. The approach could also be very conservative for glacial tills, and Equation 2-16 was soon adjusted for OC clays by Meyerhof (1976):

$$q_s = (1 - \sin(\emptyset'))\tan(\emptyset')\sqrt{OCR}\,\sigma' \qquad \text{Eq. 2-19}$$

where OCR is the overconsolidation ratio of a clay.

Patrizi and Burland (2001) were inspired by the soil test results and correlation between the OCR with the strength ratio from Jamiolkowski *et al.* (1985):

$$(C_u/\sigma')_{OC} = (C_u/\sigma')_{NC}\,OCR^m \qquad \text{Eq. 2-20}$$

where m is a correlated exponent that can vary on the soil type and testing method to obtain $C_u$. Patrizi and Burland (2001) thought collecting the preconsolidation pressure with laboratory tests was a time-consuming process, and they offered a new empirical approach to predict the adhesion factor with the OCR. From several sources of literature, they collected a database of piles, and they obtained the average $\beta$ coefficient and strength ratios for each pile. For practical reasons, the two variables were then simply correlated with a linear regression:

$$\beta = 0.1 + 0.4\,C_u/\sigma' \qquad \text{Eq. 2-21}$$

They obtained accurate predictions with their new method with a database of pile load tests. The average and standard deviation for the predicted to measured capacity ratio was 0.97 and 0.29, respectively (Patrizi & Burland, 2001). Coincidentally, their design method can be simplified to resemble the $\lambda$ Method.

### 2.3.3 $\lambda$ Method for Unit Side Resistance

Vijayvergiya and Focht (1972) proposed the "$\lambda$ Method," which can be observed as a hybrid between total and effective stress methods.

$$q_s = \lambda(\overline{\sigma'} + \overline{C_u}) \qquad \text{Eq. 2-22}$$

The method applies the average undrained shear strength ($\overline{C_u}$) and effective stress ($\overline{\sigma'}$) along a pile, and the dimensionless coefficient, $\lambda$, depends on the pile length. Vijayvergiya and Focht (1972) predicted the capacity of several piles with their new method and compared to the results to predictions from other approaches, such as the $\alpha$ Method by Tomlinson (1957). Their proposed method obtained a higher level of accuracy; however, it is not applicable for heterogeneous soils or layered profiles.

### 2.3.4 Estimation of Unit Tip Resistance

Meyerhof (1976) recommended a limit for the tip resistance ($q_l$) if the pile passed a critical depth, and he proposed the following equation for homogeneous soils:

$$q_p = c' N_c + \sigma' N_q \leq q_l \qquad \text{Eq. 2-23}$$

$N_c$ and $N_q$ are bearing capacity factors and respectively correspond to the soil cohesion ($c'$) and effective stress ($\sigma'$) at the pile base. For undrained clayey soils, the tip resistance would depend on the soil cohesion, and the deformation and sensitivity properties of an undrained clay influenced the values of $N_c$. For example, $N_c$ can be 5 for a very sensitive, brittle, and NC clay, but an insensitive OC clay can have a $N_c$ of 10 (Meyerhof, 1976). Meyerhof (1976) also expressed that the tip resistance would increase with greater consolidation. Yet, if the soil sensitivity is not high, a value of 9 can be commonly applied for the cohesive bearing factor (Meyerhof, 1976).

Vesic (1977) offered a semi-empirical approach to obtain the bearing capacity factors. For drained soils, the tip resistance was suggested to be partly governed by the effective stress at the pile tip, and the effective stress was replaced by the mean normal ground stress, $\sigma_o'$:

$$q_p = c' N_c + \sigma_o{}' N_\sigma \qquad \text{Eq. 2-24}$$

The mean normal ground stress is related to the effective stress and lateral earth pressure at rest. $N_q$ was replaced by $N_\sigma$ as the modified bearing capacity factor. For undrained soils, which may be assumed to have a friction angle of zero, $N_c$ is estimated by the equation below:

$$N_c = \frac{4}{3}(\ln I_{rr} + 1) + \frac{\pi}{2} + 1 \qquad \text{Eq. 2-25}$$

$I_{rr}$ is the reduced rigidity index. For undrained or dense clays, $I_{rr}$ is approximately equal to the rigidity index, $I_r$:

$$I_r = \frac{E}{2(1+\nu)(c + \sigma' \tan \emptyset')} \qquad \text{Eq. 2-26}$$

$$I_{rr} \approx I_r = \frac{E}{3\, C_u} \qquad \text{Eq. 2-27}$$

$E$ is the modulus of elasticity, $\nu$ is the Poisson's ratio, and $c$ is the cohesion of the soil.

O'Neil and Reese (1999) later suggested the following relationship to approximate $I_r$ without $E$:

$$I_r = 3.62\, C_u - 33 \leq 300 \qquad \text{Eq. 2-28}$$

## 2.4 DESIGN METHODS FOR A PILE IN NONCOHESIVE SOILS

### 2.4.1 Estimation of Unit Side Resistance

Although Berezantzev *et al.* (1961) focused on the pile tip resistance, they also suggested that "conventional methods," such as the following, are practical to estimate the side resistance:

$$q_s = K_s\, \sigma' \tan \delta \qquad \text{Eq. 2-29}$$

Since frictional resistances vary between different pile materials and soils, effective stress methods for noncohesive soils typically reduce the friction angle. $\delta$ is commonly applied as a result and is called the friction angle of the soil-pile interface. The effective stress is usually the average value along the pile length if the soil is homogeneous or each soil layer if the soil is heterogeneous.

Coyle and Castello (1981) collected data from full-scale load tests and proposed a new approach to estimate the side resistance for piles in tension and compression. Average soil parameters were applied in the correlations. Particularly, $K_s$ was estimated from the average friction angle, pile diameter, and pile length, and the interface friction angle was assumed to be equal to 80 % of the friction angle. Coyle and Castello (1981) concluded that the slenderness ratio (*L/D*), also called "relative depth ratio," and friction angle were the most influential parameters. The effective stress and volume of displaced soil was likely related to the slenderness ratio, and friction angle is linked to the relative density or compressibility of the soil (Coyle and Castello, 1981). While comparing the predicted capacities to the slenderness ratio, Coyle and Castello (1981) had an error of ± 30 %, which is appropriate for the range of friction angles. Unfortunately, the slenderness ratios were limited to a small range with a maximum pile length of approximately 20 m.

The method of installation affects the coefficient of horizontal soil stress, $K_s$. Since this coefficient includes the lateral earth coefficient at rest, $K_o$, another coefficient is added: the coefficient of installation, $K_f$. Based on test results by Meyerhof (1976), Tomlinson and Woodward (2008) recommends $K_f$ to vary from 1 to 2 for driven piles. Other installation methods for deep foundations have lower ranges for the coefficient.

Tomlinson and Woodward (2008) provided approximate ranges of $\delta$ for noncohesive soils. For a smooth steel pile, the value of $\delta$ is between 50 to 70 % of the friction angle. For a rough steel pile, $\delta$ was recommended to be 70 to 90 % of the friction angle. Thus, a smooth steel pile will likely experience friction differently than a concrete pile with the soil.

### 2.4.2 Estimation of Unit Tip Resistance

Vesic (1967) conducted multiple tests on piles in sandy soils and offered the general concepts. Tests involved model cylindrical and prismatic piles of approximately 0.25 m to 2.8 m in length and large-scale instrumented piles in natural deposits. In both cases, piles were driven or jacked. Beneath the bases of the model piles, Vesic (1967) observed a dominance of punching shear failure at deep depths, regardless of the density of the sand. The tip resistance appeared to reach a limiting value after the pile reached depths of approximately 15 pile diameters, and this result showed evidence of the critical depth concept. The relative density and installation of the pile, overall, resulted to be the main factors. Vesic (1967) also called for more research with piles in silts.

Meyerhof (1976), similar to Vesic (1967), suggested limiting the tip resistance after the critical depth, but Meyerhof (1976) also offered modifications to the tip resistance for a pile driven into layered soils with varying stiffness. In general, he offered the following expression:

$$q_p = N_q \, \sigma' \leq 50 \, N_q \tan \emptyset' \qquad \text{Eq. 2-30}$$

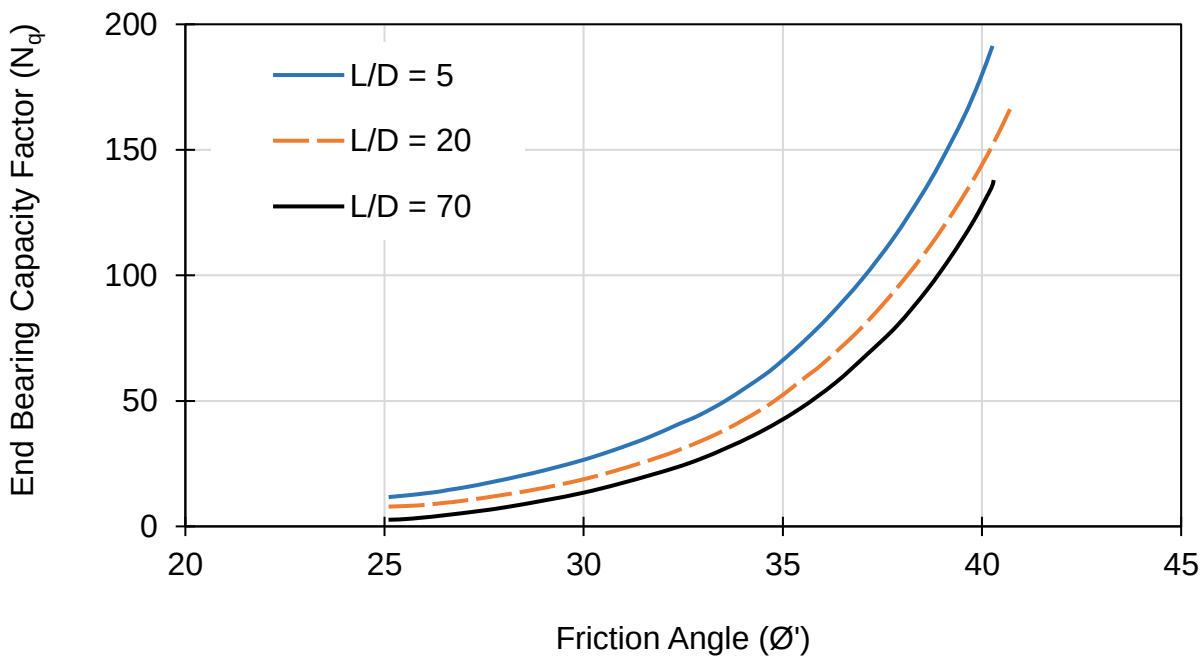



Figure 2-3: Bearing Capacity Factor by Berezantzev *et al.* (1961) (Berezantzev *et al.*, 1961)

Numerous bearing capacity factors were proposed over the years and usually varied based on the properties of particular noncohesive soils. Coyle and Castello (1981) offered a relationship with the slenderness ratio to predict the bearing capacity coefficient. The relation from Berezantzev *et*

*al.* (1961) was recommended by Tomlinson and Woodward (2008). The reason may be because the results are usually conservative compared to other bearing capacity factors. The bearing factors are estimated in Figure 2-3 with the friction angle and pile slenderness ratio.

## 2.5 APPLICATIONS OF GENETIC ALGORITHMS FOR PILE CAPACITY

### 2.5.1 Potential Methods to Predict Pile Capacity

Especially with a lack of fully-instrumented piles, previous design methods have commonly been developed empirically, by combining variables by trial and error, or by regressing the average soil parameters against the side resistance. Unfortunately, these processes simplify the soil conditions and lead to biased results. An alternative method is required to efficiently correlate data, especially if repetitive soil measurements and mixed soil profiles are involved.

More advanced statistical methods are available to handle grouped data or repetitive measurements. In particular, multilevel modeling is based on multilinear regression and divides variables and expresses relationships between several levels (Luke, 2004). Individuals or repetitive measurements are classified as a lower level, and they belong to groups, which exist as a higher level (Luke, 2004). For example, students of different genders, ethnicities, or ages belong as individuals to a school. The procedure to conduct multilevel modeling differs between macro-micro analyses or micro-macro analyses. For a macro-micro analysis, a dependent variable belongs to the individuals or the lowest level, which is level one, and the dependent variable is expressed by other level-one variables and higher-level variables (Croon and van Veldhoven, 2007). For a micro-macro analysis, the dependent variable is a higher-level variable, such as a group or level-two variable, and is expressed by the variables from the lower levels (Croon and van Veldhoven, 2007). Macro-micro analyses are common, but a limit amount of research has focused on the developments of micro-macro statistical methods (Croon and van Veldhoven, 2007; Lüdtke *et al.*, 2008). For this study, a micro-macro analysis is required to predict the overall pile capacity with several measurements for a single soil variable. Unfortunately, multilevel models are usually linear, and micro-macro analyses have not been developed to consider weighted observations. For instance, the unit side resistance would need to be multiplied, or weighted, by the side area of the pile to find the total side resistance as a load. As a result, the approach will be complex to implement or time-consuming to find nonlinear relationships.

Machine learning methods, such as ANN, are efficient at finding relationships with several variables. ANN were inspired by the operations of neurons within the brain of an organism, and they apply a system of algorithms to find underlying patterns within data (Negnevitsky, 2005). Neural networks do not require knowledge about the problem domain, but they require data to learn by example and from experience (Negnevitsky, 2005). Many variations of neural networks exist, and they have grown in popularity within geotechnical research. Shahin (2014) created a recurrent neural network to predict the settlement of drilled shafts with cone penetration tests,

while Ismail and Jeng (2011) predicted the pile settlement with a higher order neural network. The correlations from these networks usually provide significantly lower errors compared to traditional methods. However, it is difficult to obtain a simple or practical equation from this machine learning method. Das and Basudhar (2006) translated their network into a function, but it was long, complex, and impractical. Most networks are effective if the data was normalized before performing the analysis (Negnevitsky, 2005), but this process of normalization, as shown by Das and Basudhar (2006), would be contained in the final equation. In general, a developed relationship could be difficult to translate logically with engineering judgement, and the effectiveness of the ANN can be reduced as the network memorizes patterns within the data by a process called "overfitting" (Negnevitsky, 2005). The relationships from the ANN are usually hidden (Negnevitsky, 2005), and the developed ANN becomes the final deliverable or product. For these reasons, this is probably not the best approach to offer a new mathematical function in engineering practice for design.

GAs are another machine learning approach that is popular, and they are practical to apply if there is a lack of knowledge and if many possible solutions exist. The system relies on an objective or fitness function to evaluate the solutions on a particular set of criteria. Thus, for regression, a GA can rely on additional factors than solely the error between the measured and predicted results. The method can also consider nonlinear relationships between several variables and find a practical equation with symbolic regression. While developing a function, they choose whether or not to add a certain variable, and a GA can be developed to consider various mathematical operations to regress data. They have been customized for a variety of prediction problems, such as the settlement of tunnels (Gandomi & Alavi, 2012), strength parameters of soils (Levasseur *et al*., 2008), and capacity of piles with cone penetration tests (Alkroosh & Nikraz, 2011). Yet, little research has been conducted with GAs and the capacity of piles for soil deposits in Ontario.

### 2.5.2 Genetic Algorithms to Predict Pile Capacity

A GA is an optimization approach inspired by Darwin's theory of evolution (Banzhaf *et al.*, 1998). In nature, chromosomes give an organism its attributes to survive and succeed in an environment. Through reproduction, organisms can adapt and evolve to their environment. A GA represents the problem domain as a chromosome. In this case, a GA was developed to conduct symbolic regression and predict the side resistance. For symbolic regression, the genes of a chromosome represented the components of a function: a variable, constant, or operator.

As displayed in Figure 2-4, a simple GA is developed to search for a viable function through the following steps: chromosome creation, evaluation, selection, crossover, and mutation. First, multiple attempts for a problem are made at once in a trial, or generation, by generating a population of chromosomes with different attributes. The performance or fitness of a single chromosome is measured by an objective function. From the population of chromosomes, potential parents are selected for the creation of offspring. Typically, during selection, a preference is given

to chromosomes with a higher fitness, and the selection can be achieved by several approaches: tournament, rank, and roulette (Zhong *et al*., 2005). The population size remains constant throughout every generation, and the previous chromosomes, or at least a majority, are replaced by new offspring. The population then evolves through several generations by reproduction mechanisms, such as crossover and mutation.

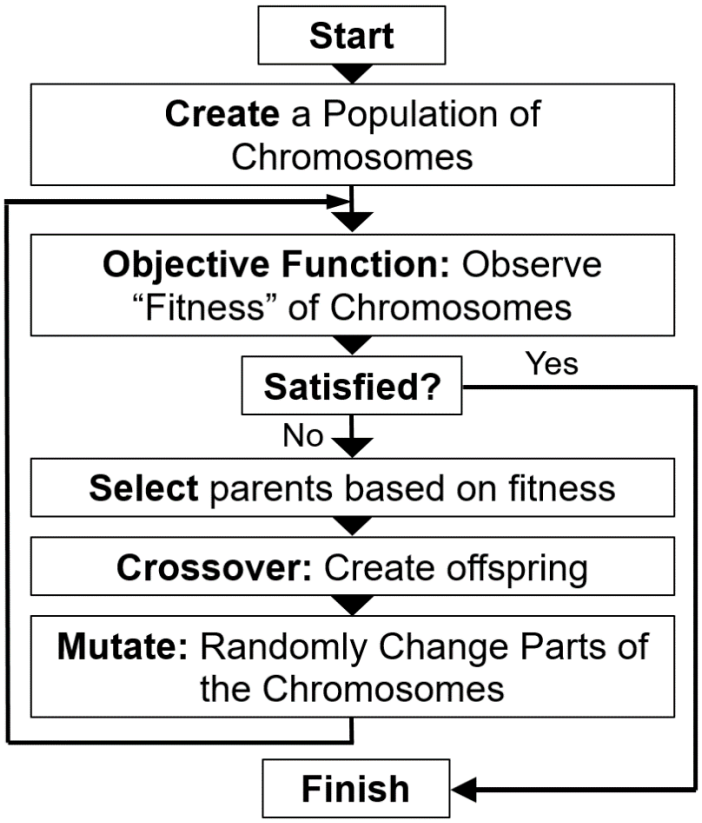



Figure 2-4: Process of a Simple Genetic Algorithm

A GA is a stochastic method. If a regression analysis was repeated with multiple trials that had the same initial conditions, the GA can complete the analysis with a similar level of fitness but provide a different solution. The method is stochastic because it relies on random sampling or events, but the randomness can be controlled by considering the diversity, volatility, and influence of events within the GA. The number of generations, chromosome population size, search time, and evolvability or search capabilities of the system is also important. The analysis may also be conducted with several trials. A GA is also data-driven and, depending on the quality of the data and complexity of a problem, may not find an exact solution or global minimum. However, since it does not require detailed information about the problem domain, it is an efficient optimization

approach, especially if aspects of a problem lack knowledge. The GA searches for a solution by randomly creating chromosomes, which represent a function or potential solution, and the chromosomes compete and evolve within the defined generations. The best candidates tend to last through several generations, and this situation creates challenges to define reliable termination criteria (Negnevitsky, 2005). The performance of the GA can be evaluated by plotting the fitness of the functions against the generations.

Various techniques have been proposed to encode chromosomes, but Oltean and Dumitrescu (2002) demonstrated that the Multi Expression Programming (MEP) technique was more efficient compared to other approaches, such as Gene Expression Programming (GEP). For a series of regression problems, MEP was able to achieve a higher fitness with fewer generations (Oltean and Dumitrescu, 2002). A MEP-based chromosome is simple to process, evaluate, and manipulate, especially for mutation and crossover (Oltean and Dumitrescu, 2002). Also, several solutions are contained in a chromosome, but the best solution typically represents the chromosome (Oltean and Dumitrescu, 2002).

## 2.6 OBTAINING SOIL PARAMETERS

Due to the natural variability of the soil, a range of soil parameters may be required to design the axial capacity of a pile. The parameters will usually differ on the soil content. Several tests may also be available to measure the same parameter, such as unconfined compression strength (UCS), unconsolidated undrained (UU) triaxial, and field vane shear test (FVST) for $C_u$. Errors easily propagate during the design of the pile since the design methods rely on the measurements. In order to reduce the variability, the approaches to obtain soil parameters were collected and evaluated on their practicality for the soil conditions in Ontario.

### 2.6.1 Field Vane Shear Test (FVST)

FVST values ($Cu_{FVST}$) are commonly corrected to obtain $C_u$ for design:

$$C_u = Cu_{FVST}\ \mu_{FVST} \qquad \text{Eq. 2-31}$$

For the equation above, CGS (2006) recommends two methods to determine the correction factor ($\mu_{FVST}$): one approach is by Bjerrum (1972; 1973), and the other is by Aas *et al.* (1986).

The correction by Bjerrum (1972; 1973) is likely to be popular in practice because it is simple. He related a correction factor to the *PI* of cohesive soils, as shown in Figure 2-5. The relation indicates that silty soils or non-plastic clays with a *PI* less than 20 will have a correction factor greater than one, but this factor is typically limited to a maximum of one to be conservative.

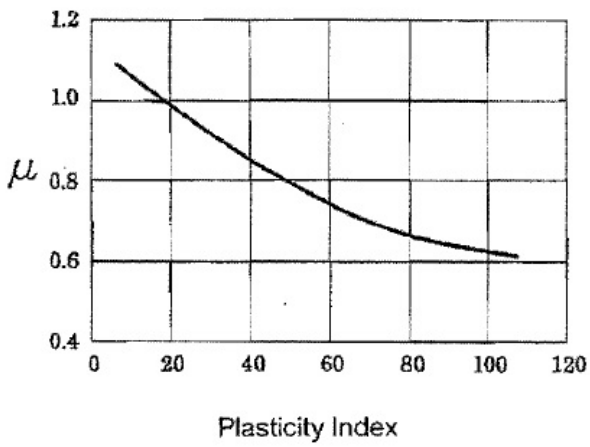
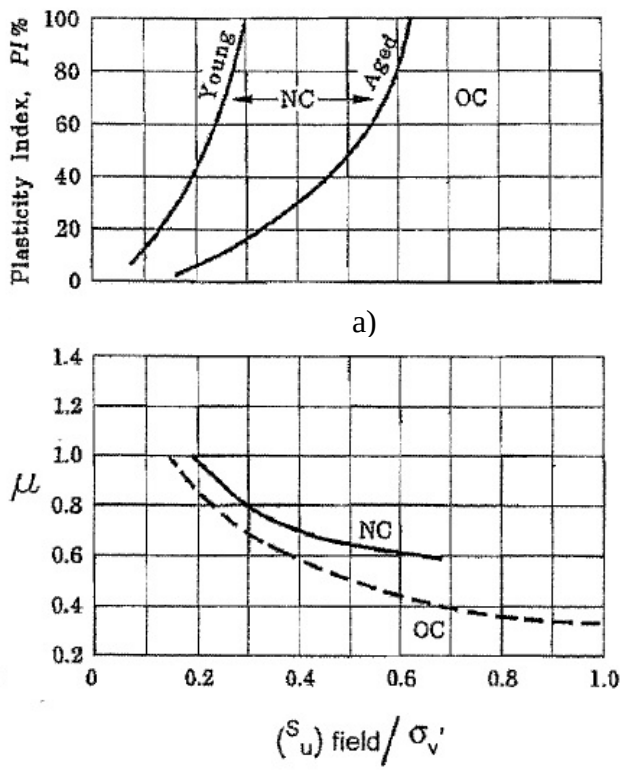



Figure 2-5: Relationship of Estimating Correction Factors for FVST from Bjerrum (1972) (CGS, 2006)




b)

Figure 2-6: Relationship of Estimating Correction Factors for FVST from Aas *et al.* (1986) (CGS, 2006)

Aas *et al.* (1986) modified the relationship to estimate the stress history and include the aging of the soil, as shown in Figure 2-6a and b. The *PI* and effective stress ($\sigma'$) are used to estimate if the soil is NC or OC in Figure 2-6a, and the corresponding NC or OC curve then indicates the value of the correction factor in Figure 2-6b. OC soils receive lower correction factors. Although Aas *et al.* (1986) included more variables, the parameters are an approximation, and the approach may not be applicable for desiccated clays with high a $C_u$ or heavily OC soils near the ground surface.

Several correction factors have developed over the years. Chandler (1988) observed several influences on soil strength measurements from the FVST: the *PI*; soil stress history; and vane insertion disturbance, dimensions, and shearing rate. Chandler (1988) also suggested that the correction factor from Bjerrum (1972) may be appropriate for inorganic, NC clays. Morris & Williams (1994) developed a correction factor for the liquid limit (*LL*) or *PI* with data from international references. The data included samples with a *PI* less than 20, and they suggested lower factors for these soils compared to Bjerrum (1972).

### 2.6.2 Standard Penetration Test (SPT) versus Undrained Shear Strength ($C_u$)

Numerous correlations have been determined between $C_u$ and field N-values (*N*) or corrected N-values ($N_{60}$). The relationships are usually linear:

$$Cu = A\,N + B \qquad \text{Eq. 2-32}$$

The variables $A$ and $B$ are correlation coefficients, but $B$ can be ignored since it is usually equal to zero.

As early as the 1950s, Sowers (1954) noticed the soil plasticity influenced the relation between SPT and $C_u$. He indicated that Coefficient $A$ generally ranges between 7.1 to 16.5, 4.7 to 9.5, and 2.4 to 4.7 respectively for highly plastic clays, medium plastic clays, and very low plastic clays and plastic silts. From his results, Coefficient $A$ is higher with a higher *PI*.

Stroud (1974) collected test data from insensitive clays and London Clay from a total of 70 boreholes at 18 different sites. He suggested $A$ is higher with a lower *PI*. For a *PI* from 35 to 65, Stroud (1974) found $A$ varied from 4 to 5, and Coefficient $A$ was greater than 6 for a *PI* less than 20. However, some of the investigated soils likely had a different mineral content or greater silt content than Sowers (1954). For a *PI* less than 15, Stroud (1974) also experienced difficulties to obtain *PI*.

Hara *et al.* (1974) developed his correlation with soil samples in Japan. $C_u$ was measured by unconfined compression tests (UCS), and N-values ranged from 1 to 100 because values less than 1 were rejected. The nonlinear relationship achieved a high fitness with a coefficient of correlation of 0.93.

$$Cu = 29.7\,N^{0.72} \qquad \text{Eq. 2-33}$$

The cohesive soil was mainly compressible as the *PI* and *LL* ranged between 25 to 60 and 50 to 100, respectively. The OCR was also measured and varied from 1 to 3.

As shown in Figure 2-7, Kulhawy and Mayne (1990) provided a series of linear correlations. The relations were developed in various locations around the world and with a range of soil conditions. While investigating a particular soil, the importance of applying the appropriate relationship can be observed.

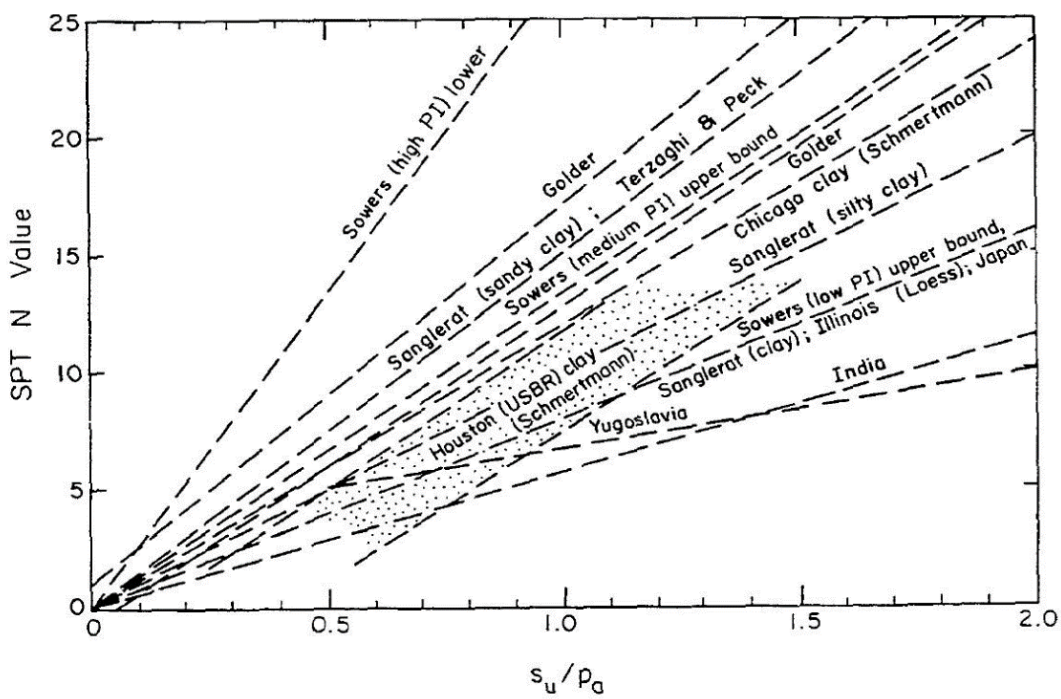



Figure 2-7: Relationships between N-Values and $C_u$ (Kulhawy & Mayne, 1990)

Sivrikaya and Toğrol (2006) considered different correlations for fine-grained soils with various *PI*. Table 2-3 displays the proposed ranges of $A$. Samples were collected by Shelby tubes, and UCS, unconsolidated undrained (UU) traxial tests, and FVST were conducted to gain $C_u$ at several regions across Turkey. Since the resulting $C_u$ differs from each testing method, it was logical to correlate the lab and in-situ tests individually to the N-values. Sivrikaya and Toğrol (2006) applied the correction factor by Aas *et al.* (1986) to shear strengths from FVST. For the different soil types, a relationship was attempted include N-values, $C_u$, and *PI* together; however, a weak correlation resulted.

Table 2-3: Suggested A Coefficients for Various Soil Types (Sivrikaya and Toğrol, 2006)

| Soil Type | Range of *A* Coefficient | Average *A* Coefficient |
|---|---|---|
| Fine-Grained Soil | 2 to 17.5 | 6.09 |
| CH | 2.25 to 17.5 | 7.52 |
| Clay | 2.12 to 17.5 | 6.38 |
| CL | 2.12 to 13 | 4.98 |
| ML | 2.68 to 6.67 | 4.22 |
| MH | 2 to 6.88 | 3.8 |

Sivrikaya and Toğrol (2006) had similar conclusions to Sowers (1954): Coefficient $A$ increased with a higher $PI$, and the coefficient was low for silts and non-plastic clays. However, Sivrikaya and Toğrol (2006) recommended to not apply in practice the correlations for highly plastic clays by Sowers (1954).

Nassaji and Kalantari (2011) offered multi-linear relationships with N-values, $C_u$, and Atterberg properties. SPT was conducted and soil samples were taken from 19 borehole logs, which were located at four regions around Terhan, Iran. From the collected soil conditions, the values of $PI$, $C_u$, and $N$ ranged from 5 to 28.5, 18 to 108 kPa, and 4 to 46, respectively. Four relationships were proposed for corrected and field N-values:

$$Cu = 1.6\,N + 15.4 \qquad \text{Eq. 2-34}$$

$$Cu = 2.1\,N_{60} + 17.6 \qquad \text{Eq. 2-35}$$

$$Cu = 1.5\,N - 0.1\,w_c - 0.9\,LL + 2.4\,PI + 21.1 \qquad \text{Eq. 2-36}$$

$$Cu = 2\,N_{60} - 0.4\,w_c - 1.1\,LL + 2.4\,PI + 33.3 \qquad \text{Eq. 2-37}$$

Although a large database was not used, the technique to correlate the data was suggested to offer a better approach to estimate $C_u$. Nassaji and Kalantari (2011) compared their proposed approaches to relations from two other references. The additional parameters included in Equations 2-36 and 2-37 were able to reduce the standard deviation and improve the correlation coefficient. For the multi-linear regressions, the variables are assumed to be independent, and the significance of the variables from a t-test were not provided. In the end, Nassaji and Kalantari (2011) recommended to apply the relationships to the intended areas. Soil conditions and in-situ testing may differ for other regions.

### 2.6.3 Friction Angle for Noncohesive Soils

A few relationships common relationships were collected to estimate the friction angle from SPT N-values. While predicting the capacity of pipe piles and H piles in sand, Wolff (1989) approximated the graphical relationship by Peck *et al.* (1974), which is shown in Figure 2-8, with the following expression:

$$\emptyset' = 27.1 + 0.3\, N_{60} - 0.00054\, {N_{60}}^2 \qquad \text{Eq. 2-38}$$

Although it is not indicated, Peck *et al.* (1974) probably directly correlated field N-values to friction angles. The application of corrected $N_{60}$ values may offer conservative results and provide unified or consistent results with varying SPT procedures. Wolff (1989) did correct the N-values for overburden pressures according to Liao and Whitman (1986) during the calculations of the pile capacities.

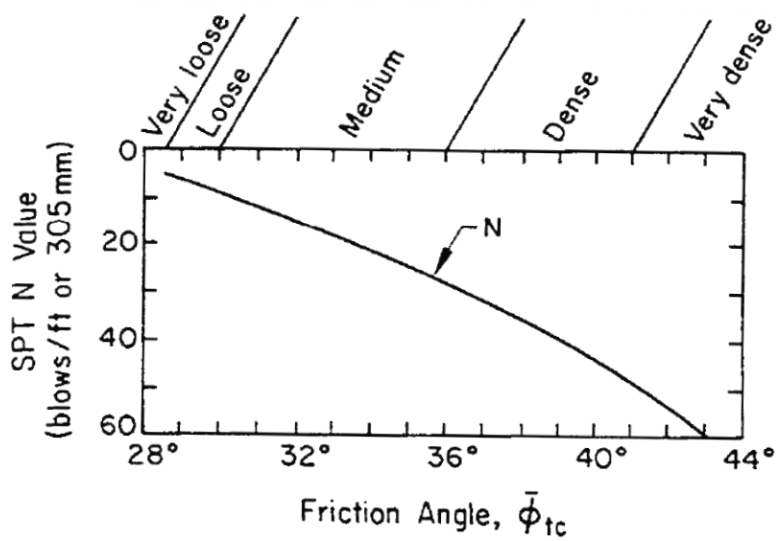



Figure 2-8: Relation of N-Values and Friction Angles by Peck *et al.* (1974) (Kulhawy & Mayne, 1990)

Kulhawy and Mayne (1990) presented a series of correlations to estimate soil parameters. Considering that the N-values are influenced by different overburden pressures, the suggested relationship modifies the friction angle with the effective stress:

$$\emptyset' = \arctan\left(\frac{(N_{60})_1}{12.2+20.3\left(\frac{\sigma'}{100}\right)}\right)^{0.34} \qquad \text{Eq. 2-39}$$

In the equation, $(N_{60})_1$ is the N-value corrected for both the hammer energy and overburden pressure in the noncohesive soil. Kulhawy and Mayne (1990) advised to not apply the method for depths less than 1 to 2 m and mentioned the relationship may offer conservative results.

Hatanaka and Uchida (1996) developed and compared correlations to other references, such as the Japan Road Association. A method was proposed after obtaining in-situ frozen samples of sandy soil and determining the friction angle with drained triaxial tests. The samples had friction angles from approximately 28 to 45 degrees, and the correlated N-values were corrected for the overburden pressure by the equation suggested by Liao and Whitman (1986). Hatanaka and Uchida (1996) in the end suggested an equation with the best fit for the investigated soil:

$$\emptyset' = \sqrt{20\,(N_{60})_1} + 20 \qquad \text{Eq. 2-40}$$

It was also concluded that the particle shape may affect the relationship.

## 2.7 Other Influential Factors on Pile Capacity

The design of a single pile is a very complex problem due to the interaction between the soil and pile. In addition to the soil conditions, there are many influential factors, including the pile geometry, pile material, and duration of time between installing and testing a pile. Since the design process can involve many variables, technical literature was reviewed for potential solutions.

### 2.7.1 Pile Geometry and Installation

Through this literature review, different references have investigated a range of pile types and materials. Many references tested or observed databases with pipe piles or solid piles. If H piles were studied, they were usually found in small numbers. Karlsrud *et al.* (2005) experienced a small difference between closed- and open-ended pipe piles. The difference was accounted by the $F_{Tip}$ coefficient, which modified the determined adhesion factor up to 25 %. Karlsrud (2012) later removed this factor in the predictions of the adhesion coefficient.

While comparing two piles with same length in the same soil of a site, Stermac *et al.* (1969) noticed a steel pile with an oversized based plate had a reduced axial capacity after a year. Yet, the differences in capacities were negligible for the short-term experiments (Stermac *et al.*, 1969).

The idea that the pile length or slenderness ratio are factors on the axial capacity has been commonly investigated. The pile length was linked to “whipping” effects, which is the lateral movement of a pile or soil during the driving process (Karlsrud, 2012; Semple & Rigden, 1986). Whipping may reduce the capacity as the soil is pushed away from the pile walls. Especially in a

cohesive OC soil, the soil will take a long time to flow and apply a pressure on the walls of the pile, or a permanent gap may even develop at the pile-soil interface (Semple & Rigden, 1986). This probably explains the reduction in capacity by Stermac *et al.* (1969) with oversized base plates. Tomlinson (1957) suspected inconsistencies for his design method were due to the development of gaps between the soil and upper section of the pile. In general, the whipping effects are expected to be the greatest towards the ground surface, where there is less lateral resistance by the soil (Karlsrud, 2012). Other considerations may include the flexibility of a pile and the dissipation of loads within a pile at greater depths. Many references, such as with Kolk and Van der Velde (1996), have incorporated adjustments for the slenderness or length in their proposed methods. On the other hand, Karlsrud (2012) did not observe a significant impact.

Another idea involving the pile length is the critical depth concept. Particularly in noncohesive soil, Vesic (1967) noticed the side resistance remained constant along a pile after a certain depth. This depth of maximum resistance is the critical depth, and Meyerhof (1976) suggested the critical depth was located at a distance of 15 to 20 diameters from the ground surface. Based on this concept, Meyerhof (1976) developed equations to estimate the tip and side resistance.

From a literature review by Poulos *et al.* (2001), the critical depth concept was explained. Other than testing errors, the reduction in the side resistance with a cohesive soil could be due to the change in OCR; an OC soil typically has higher in-situ lateral stresses near the ground surface, which affects the side resistance (Poulos *et al.*, 2001). The rate of increase by the side resistance will then follow the decreasing and eventually stabilized OCR with depth (Poulos *et al.*, 2001). Mainly in a noncohesive soils, Poulos *et al.* (2001) suggested that the confinement of the effective stress will likely gradually dissipate with greater depths, and limiting the side resistance will likely offer conservative estimates of the capacity.

### 2.7.2 Time Dependency and Setup Time

Particularly for cohesive soils, Stermac *et al.* (1969) showed that time can have a significant impact on the axial capacity of a pile in Toronto, Ontario. However, the setup time will not be considered in the scope of this investigation since most of the piles in this study had a setup time of 10 to 30 days. In addition, correction factors typically require advanced testing, such as the soil sensitivity, OCR, and coefficient of consolidation (Abu-Farakh *et al*., 2016; Karlsrud *et al*., 2005; Karlsrud, 2012). Kolk and Van der Velde (1996) found the influence of time to be difficult to isolate with empirical approaches to estimate the horizontal coefficient of consolidation.

## 2.8 SUMMARY

α Methods, such as the one by Kulhawy and Jackson (1989), may only rely on $C_u$, but many references have received improvements by including other variables. Kolk and Van der Velde (1996) had less inconsistent predictions by including the effect of the pile length to predict $\alpha$, and

Karlsrud *et al.* (2005) noticed soils with a low $PI$ had a significantly lower $\alpha$ compared to CGS (2006). Thus, while considering the availability of the data, the $C_u$, $L/D$, $\sigma'$, and $PI$ will be considered as potential variables for a new design method for cohesive soils.

The side and tip resistances for noncohesive soils are usually expressed as a function of the friction angle ($\emptyset'$) (Berezantzev *et al*., 1961), but some references experienced varying impacts with the pile length (Meyerhof, 1976; Vesic, 1977, Poulos *et al*., 2001). Meyerhof (1976) recommends limiting the side and tip resistance once a particular depth is reached. Poulos *et al.* (2001) stated the confinement likely dissipates gradually for longer piles. In this study, the side resistance will not be limited by the critical depth concept, but judgement is needed to handle the parameters, especially from in-situ measurements.

For SPT correlations, N-values are simply regressed to the pile resistances, and the results will vary by the soil type. In this study, heterogeneous soils were considered by the soil type. Other factors included in this analysis are the $L/D$ and $\sigma'$, which may be related to the pile installation effects and soil confinement.

Part of the problem of predicting the pile capacity involves the consistency of the soil profile, and in reality, the application of average parameters is not always appropriate. Many sources used statistical, semi-empirical, or trial-and-error approaches to develop their design methods, and they applied the average soil parameters from fairly homogenous soils. This situation is very common for SPT correlations, but it simplifies the soil conditions as extremes from the measurements are not considered. For heterogeneous soils, biased correlations will likely develop. Very few references, such as Kolk and Van der Velde (1996), proposed a method intended for layered soil conditions. A machine learning method, specifically a GA, was applied in this study because it is efficient at handling several variables and developing nonlinear relationships.

Artificial neural networks may be efficient at processing data and finding relationships, but the method can also create challenges to extract a practical function for design. A GA can evaluate a solution on several qualities, such as the accuracy and complexity of a function, and it was selected to perform symbolic regression. The focus was to correlate the side resistance of piles in heterogeneous soils to multiple measurements of SPT blowcounts or $C_u$. Oltean and Dumitrescu (2002) demonstrated that the MEP encoding technique was more efficient compared to GEP, and MEP was applied for this GA.

Input parameters and their interdependency are important as the uncertainties will propagate to the final result. The correction factor by Bjerrum (1972) was applied in this study due to its simplicity, and it was recommended by CGS (2006). However, the OCR and disturbance of a cohesive soil likely has a large influence on the measured values.

For the correlations between N-values and $C_u$, the typical trend is that a lower $PI$ will result in a lower Coefficient $A$ (Sowers, 1954; and Sivrikaya & Toğrol, 2006). Due to the silt content in tills, which usually results in low $PI$ values, the relationships by Sowers (1954), Sivrikaya and Toğrol

(2006), and Nassaji and Kalantari (2011) are probably the most applicable. While observing the results from Sivrikaya and Toğrol (2006), many methods presented by Kulhawy and Mayne (1990) (Figure 2-7) will overestimate the value of $C_u$. Hara *et al.* (1974) will also likely overestimate because it was intended for compressible soils. The many correlations suggest analyzing the conditions of a local soil and determining customized correlations for soils within Ontario.

In noncohesive soils, three methods were presented in the literature review. The correlation by Wolff (1989) was suggested to offer conservative values (Kulhawy and Mayne, 1990), but all of the methods are applicable in this study for indirect design methods.

In summary, many uncertainties exist to predict the ultimate capacity of piles. In addition to the soil strength and contents, predictions can be influenced by the soil measurement technique, installation method, and pile material and geometry. This study will focus on driven piles in Ontario soils and obtain an improved design method with a GA.

# 3 STATISTICAL ANALYSIS AND METHODOLOGY

## 3.1 OVERVIEW OF METHODOLOGY

For driven piles in heterogeneous soils, the goal was to improve the predictability of the ultimate capacity. New correlations for design were developed by following 3 steps. (1) Results were gathered from pile load tests and site investigations from a database by MTO. (2) For every pile, the side resistance, tip resistance, or ultimate capacity was extracted from load-displacement curves. The side resistance was approximated from the yielding load of extension load tests, and the tip resistance was assumed to be the difference in failure loads from the compression and extension tests. (3) A statistical analysis was performed to determine trends with soil strength measurements and pile resistances.

## 3.2 SELECTED SITES AND PILE LOAD TESTS

For this investigation, a database was compiled with the site conditions and pile test data from MTO (1993). The focus was on driven H piles and concrete-filled steel pipe piles that were subjected to static-axial-load tests. The tests followed standards similar to D1143-07 from the American Society for Testing and Materials (ASTM) (2007). H piles varied by designation, but the pipe piles had consistent diameter of 0.324 m. From the database, a total of 43 piles were selected for this study. Among them, 30 piles (15 pipe piles and 15 H piles) were tested with both tensile and compressive loads and were selected for this investigation. The side ($Qs$) and tip resistance ($Qp$) was approximated with results from both load tests. Typically, the compressive load test was performed first, and the duration between installing and testing a pile was approximately 10 to 30 days. Piles embedded in mainly noncohesive soils had a short setup time of about 10 days, while piles installed in cohesive soils usually had a longer setup time. Another 13 piles were also used to investigate the accuracy of existing design methods in Chapter 4. These piles were in cohesive or noncohesive soils and experienced compressive load tests. The ultimate capacity ($Qu$) was extracted from these test results, and the ultimate failure loads and pile properties are in Table 3-1. The embedment lengths ranged from 3 to 45 m, but most of the piles were founded to depths of 12 to 25 m. Additional information on the pile load tests is provided in Appendix A.

As shown in Figure 3-1, pile load tests were located in various regions of Ontario, but most of these test sites were located in Southern Ontario. Soils were mainly inorganic, silty, and compact or stiff. Glacial deposits were often found with sites that contained cohesive soils, and they usually surrounded the lower portion or base of the piles, especially if they overlaid bedrock. From the test sites, some loose sands and slightly organic clays were also found. Particularly for the statistical analysis with the soil measurement techniques, data on slightly sensitive clays was collected from two additional sites near Ottawa.

Table 3-1: Details on the Studied Piles

| Site No. | Pile No. | Pile Type[1] | Length[2] (m) | Embedded Soil Type[3] | $Qu$ (kN) | $Qs$ (kN) | $Qp$ (kN) | Notes |
|---|---|---|---|---|---|---|---|---|
| 11 | 1 | HP 310x79 | 26.82 | Sand to Silt | 523 | | | |
| 13 | 19 | 324 OD Pipe | 19.74 | Silty Sand to Sandy Silt | 1286 | | | 343 mm Ø shoe |
| 14 | 2 | 324 OD Pipe | 18.29 | Silty Clay | 298 | | | 343 mm Ø shoe |
| 21 | 4 | HP 370 x 108 | 21.50 | Silty Clay | 1661 | | | |
| 22 | 3 | 324 OD Pipe | 15.30 | Clayey Silt | 192 | 118 | 74 | 343 mm Ø shoe |
| 22 | 4 | 324 OD Pipe | 30.15 | Clayey Silt | 996 | 340 | 665 | 343 mm Ø shoe |
| 22 | 5 | 324 OD Pipe | 15.28 | Clayey Silt | 233 | 133 | 99 | 324 mm Ø shoe |
| 23 | 2 | 324 OD Pipe | 3.02 | Silty Clay | 449 | 209 | 239 | 343 mm Ø shoe |
| 23 | 3 | HP 310x110 | 3.05 | Silty Clay | 429 | 236 | 193 | |
| 24 | 2 | 324 OD Pipe | 15.39 | Sand | 638 | 372 | 265 | 343 mm Ø shoe |
| 24 | 3 | 324 OD Pipe | 22.40 | Sand | 709 | 401 | 308 | 343 mm Ø shoe |
| 24 | 4 | HP 310x79 | 22.40 | Sand | 1389 | 403 | 986 | |
| 24 | 5 | HP 310x79 | 15.39 | Sand | 715 | 263 | 451 | |
| 25 | 1 | 324 OD Pipe | 5.64 | Silty Clay | 343 | 241 | 102 | 343 mm Ø shoe |
| 25 | 4 | HP 310x79 | 18.44 | Silty Clay | 873 | 507 | 366 | |
| 25 | 5 | 324 OD Pipe | 18.35 | Silty Clay | 667 | 383 | 284 | 343 mm Ø shoe |
| 25 | 6 | 324 OD Pipe | 9.27 | Silty Clay | 472 | 320 | 151 | 343 mm Ø shoe |
| 25 | 9 | HP 310x79 | 9.39 | Silty Clay | 489 | 323 | 166 | |
| 28 | 1 | HP 310x79 | 6.10 | Clayey Silt | 499 | | | |
| 28 | 2 | HP 310x79 | 18.29 | Clayey Silt | 486 | 316 | 169 | |
| 28 | 3 | HP 310x79 | 12.19 | Clayey Silt | 563 | | | |
| 28 | 7 | 324 OD Pipe | 6.10 | Clayey Silt | 671 | 557 | 113 | 343 mm Ø shoe |
| 28 | 8 | 324 OD Pipe | 18.29 | Clayey Silt | 695 | 406 | 289 | 343 mm Ø shoe |
| 28 | 9 | 324 OD Pipe | 12.04 | Clayey Silt | 591 | 571 | 20 | 343 mm Ø shoe |
| 34 | 19 | 324 OD Pipe | 18.59 | Sandy Silt | 433 | | | 343 mm Ø shoe |
| 34 | 23 | 324 OD Pipe | 18.59 | Sandy Silt | 442 | | | 343 mm Ø shoe |
| 34 | 25 | 324 OD Pipe | 18.59 | Sandy Silt | 473 | | | 343 mm Ø shoe |
| 34 | 27 | 324 OD Pipe | 18.59 | Sandy Silt | 464 | | | 343 mm Ø shoe |
| 35 | 1 | HP 310x110 | 14.69 | Layered Clayey Silt and Silty Sand | 1609 | 506 | 1102 | |
| 35 | 4 | 324 OD Pipe | 14.69 | Layered Clayey Silt and Silty Sand | 1537 | 730 | 807 | 343 mm Ø shoe |
| 35 | 5 | HP 310x110 | 27.58 | Layered Clayey Silt and Silty Sand | 2744 | 1493 | 1251 | |
| 37 | 3 | HP 310x79 | 14.48 | Sand to Silty Sand | 1054 | 333 | 721 | |
| 37 | 4 | HP 310x79 | 38.94 | Sand to Silty Sand | | 1394 | | |
| 37 | 5 | HP 310x79 | 31.24 | Sand to Sandy Silt | 1633 | 420 | 1214 | |
| 37 | 6 | HP 310x110 | 14.48 | Sand to Silty Sand | 798 | 383 | 414 | |
| 37 | 7 | HP 310x110 | 45.29 | Sand to Silty Sand | 1943 | 1524 | 419 | |
| 37 | 8 | HP 310x110 | 30.92 | Sand to Silty Sand | 1525 | 699 | 826 | |
| 39 | 2 | HP 310x110 | 25.50 | Silty Sand; Layered Clay and Silt | 1307 | 614 | 693 | |
| 39 | 3 | 324 OD Pipe | 25.40 | Silty Sand; Layered Clay and Silt | 1170 | 470 | 700 | 343 mm Ø shoe |
| 40 | 2 | HP 310x110 | 24.50 | Layered Sand and Silty Clay | 1232 | 598 | 634 | |
| 40 | 3 | 324 OD Pipe | 17.20 | Sandy Silt to Sand | 1167 | 505 | 661 | 343 mm Ø shoe |
| 41 | 2 | HP 310x110 | 19.50 | Sand | | 1052 | | |
| 41 | 3 | 324 OD Pipe | 16.00 | Sand | | 664 | | 343 mm Ø shoe |

Notes: [1] Steel H pile designations are size (mm) by weight (kg/m). Steel pipe piles were filled with concrete before testing, and OD is the outside diameter (mm); [2] Embedment Length; [3] The dominating soil type, and classifications are according to MTO standards.

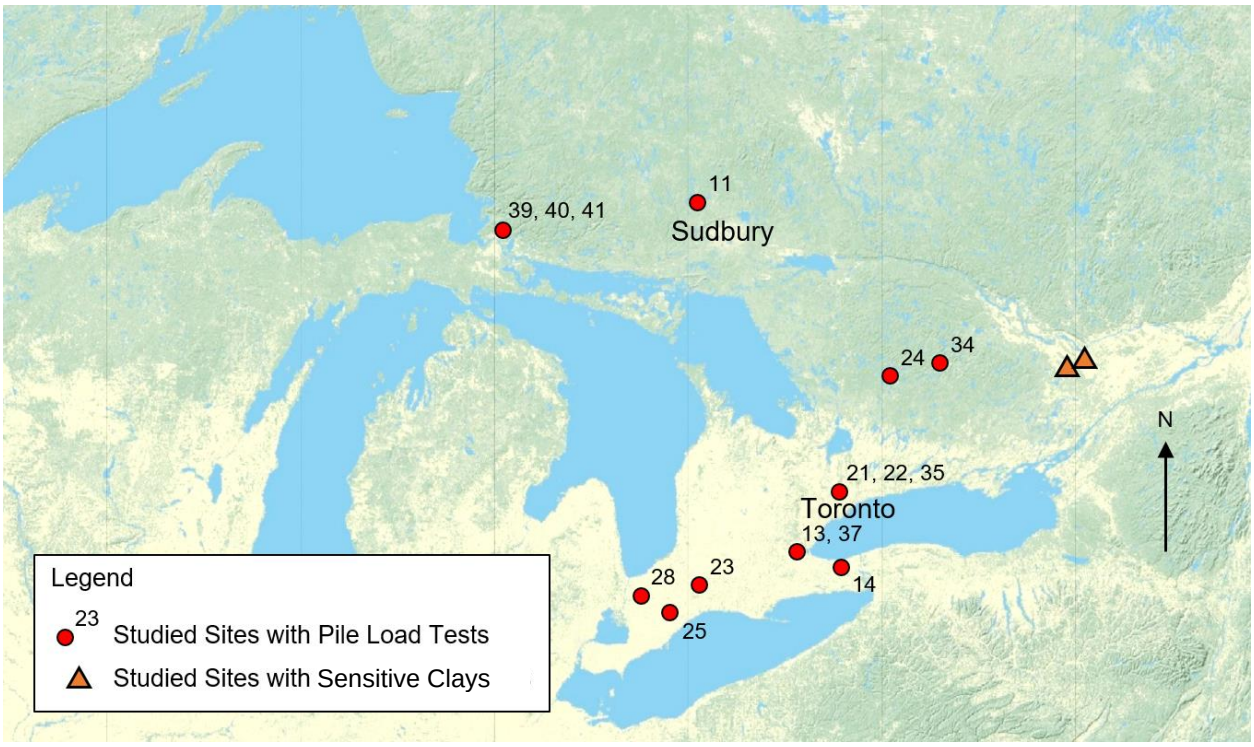



Figure 3-1: Location of Studied Sites

Soil conditions were provided by borehole logs, but every site varied in the extent and diversity of the field and lab tests. An example is shown in Figure 3-2 for Site 22, which dominated with cohesive soils. The figure was compiled with results from several borehole logs. Although the soil profiles were similar among the borings, the soil profile in the figure is based on the borehole that was closest to the piles. In general, measurements from the database included Atterberg limits, natural moisture contents, unit weights, and the composition of gravel, sand, silt, and clay according to the grain-size distribution. Soil classifications and SPT N-values were commonly recorded at different depths. Depending on the dominating soil type, N-values from the field were corrected according to CGS (2006) for 60% hammer efficiency ($N_{60}$) and the overburden pressure ($(N_1)_{60}$ ). The $C_u$ was measured with unconfined compression strength (UCS), unconsolidated undrained (UU) triaxial, and field vane shear (FVST) tests. $C_u$ commonly reached values greater than 100 kPa with UCS and UU. A variety of soil measurements were collected from the sites, but a majority of the results were from SPT and Atterberg limits.

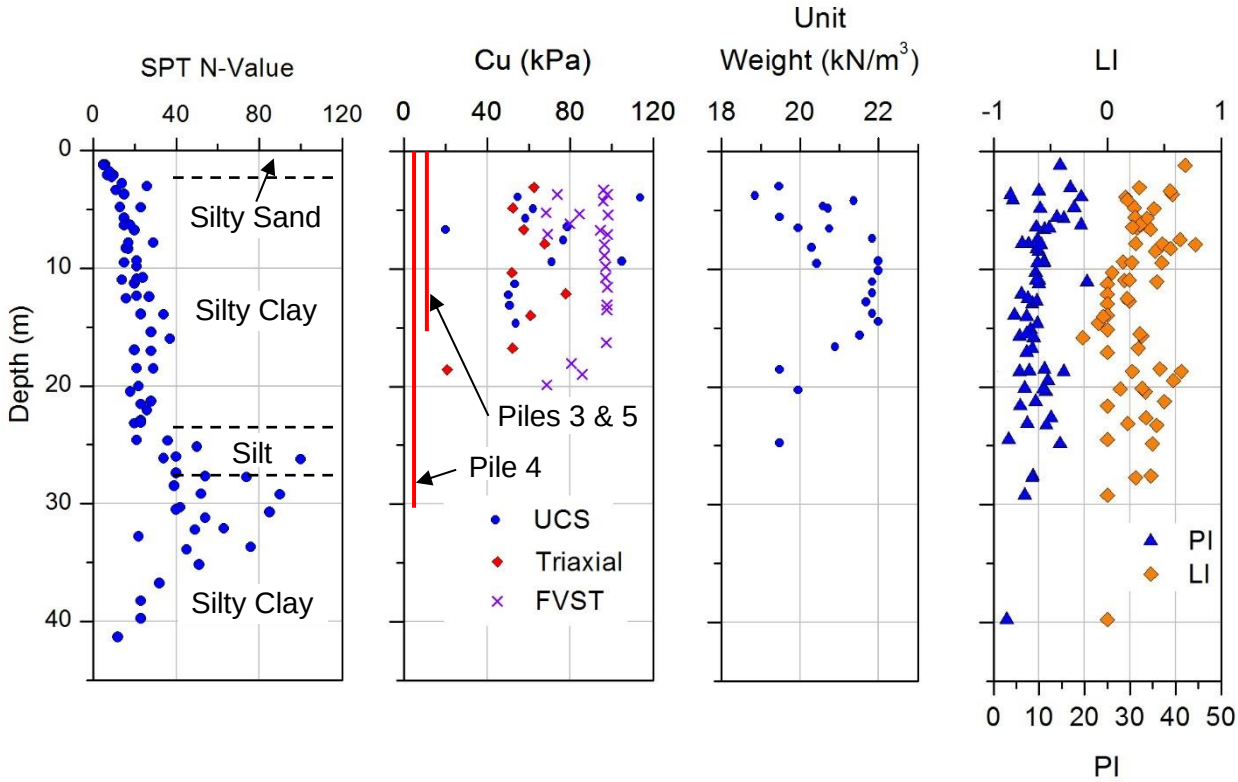



Figure 3-2: Soil Measurements at Site 22

## 3.3 RESISTANCES INTERPRETED FROM PILE LOAD TESTS

During the load tests, dial gauges measured the load-displacement response at the top of the piles. Figure 3-3 provides the load test data for Pile 4 from Site 22. In this figure, "Test No. 1" was the compression test on the pile, and "Test No. E1" was the extension test. Figure 3-4 has the extracted plots of the load-displacement curves, and they are commonly evaluated with graphical methods to determine the failure load. Although the Davisson Offset Method is popular, it is an empirical method and may inaccurately estimate the yielding point. The failure loads for this investigation were evaluated with the criteria by De Beer (Fellenius, 1980), which relies on the characteristics of the plot. De Beer suggested to plot both axes of the load-displacement curve on a log-scale, and the failure load is indicated in Figure 3-5 by the point on the curve with the greatest change in slope (Fellenius, 1980).

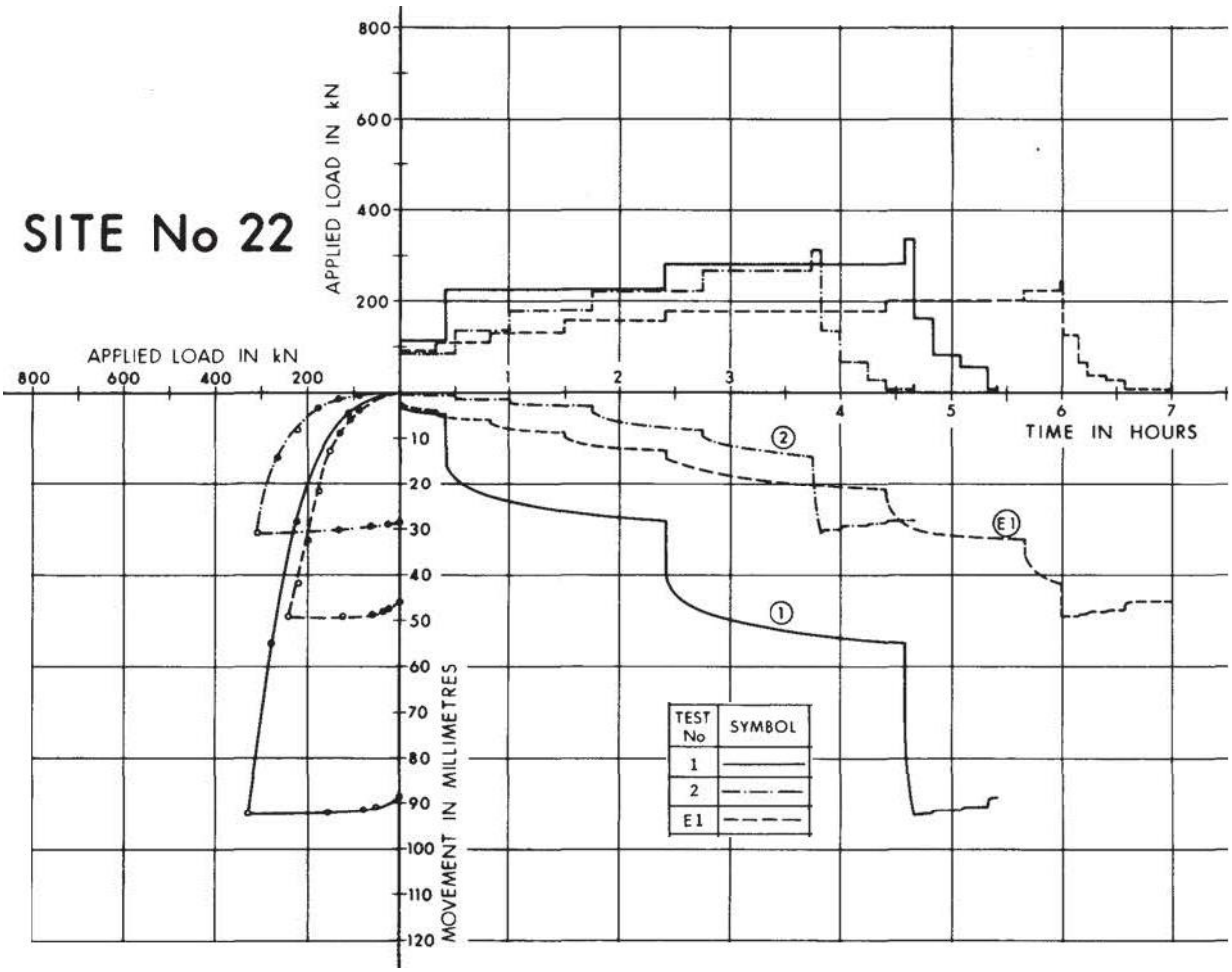



Figure 3-3: Load, Displacement, and Time Measurements from Pile 4 at Site 22

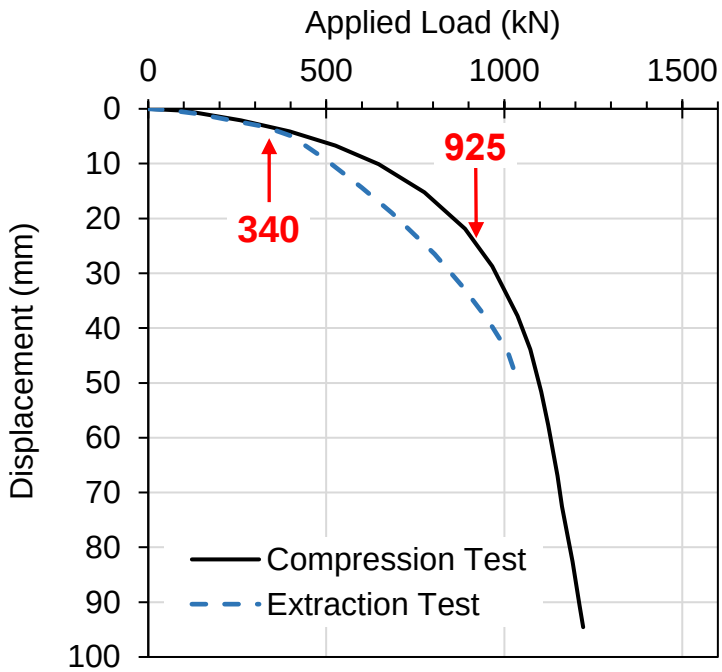



Figure 3-4: Example of Load-Displacement Curves for Pile 4 at Site 22

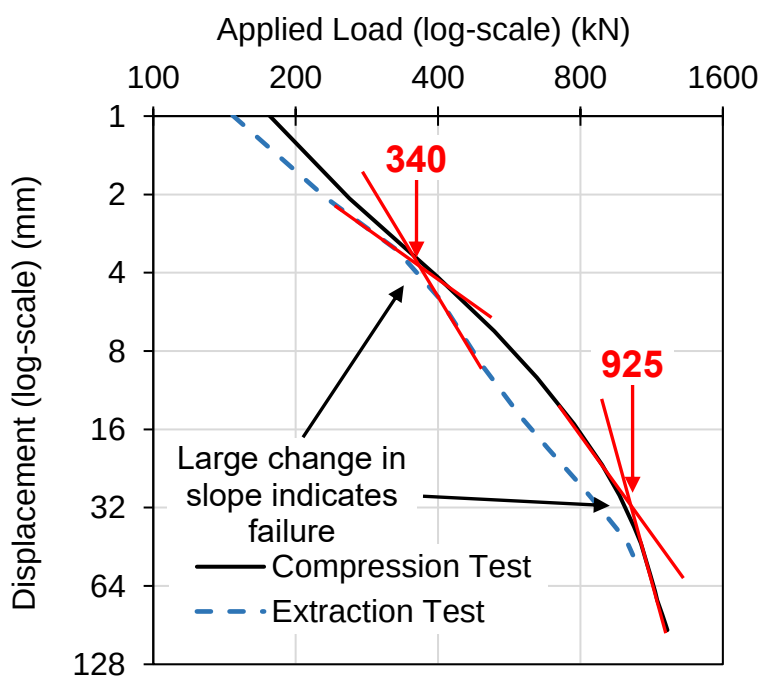



Figure 3-5: Example of the De Beer Method for Pile 4 at Site 22

## 3.4 STATISTICAL ANALYSIS WITH PILE RESISTANCES AND SOIL PARAMETERS

After collecting the pile resistances, a statistical analysis was conducted to find trends with the soil conditions. H piles were assumed to be fully plugged during this stage. Correlations were also obtained between various soil strength measurements because strength parameters may be predicted with SPT or FVST results.

### 3.4.1 Evaluating Soil Strength Measurements

Particularly for cohesive soils, a variety of tests can determine the soil strength, but each approach shears the soil differently. From every site, the soil measurements at similar depths were collected and compared. Figure 3-6 shows that results by UCS and UU triaxial were usually over 100 kPa, especially for silty clays and desiccated clays. Values over 100 kPa may not be reliable in design according to CGS (2006); yet, the two testing methods were consistent and similar. FVST measurements were not often collected in very stiff soils, but Figure 3-7 shows they were approximately 35 % higher on average than UU triaxial. Correction factors, such as the one by Bjerrum (1972), are usually applied. Unfortunately, as shown in Figure 3-8, a lot of variability was found with Ontario soils, and the *PI* may not sufficiently correct FVST results on its own.

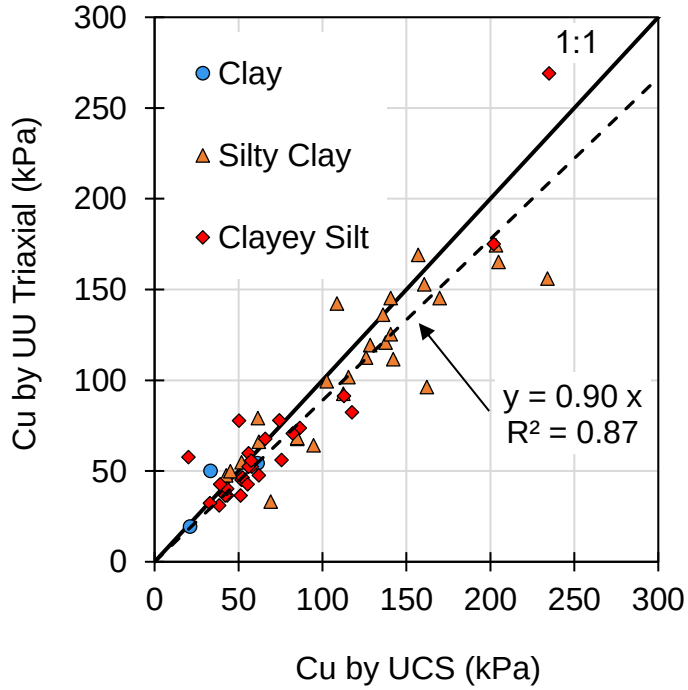



Figure 3-6: Comparison of UU Triaxial to UCS

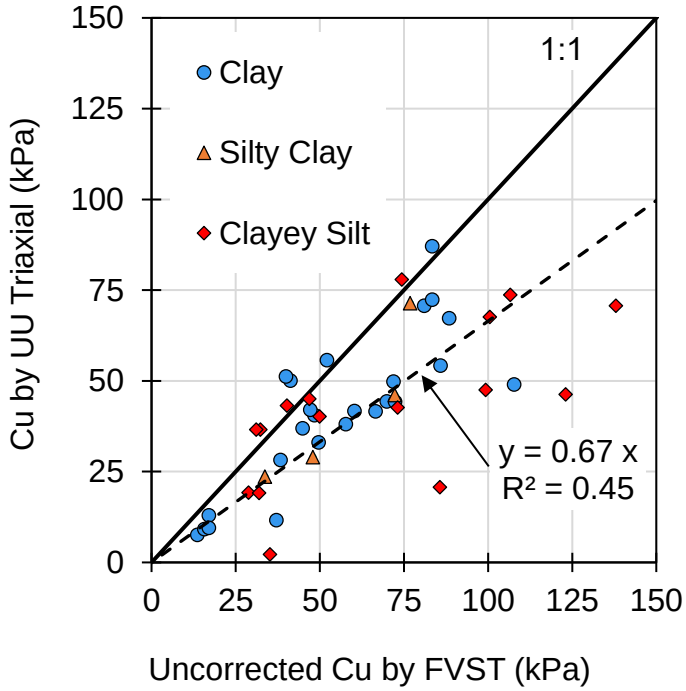



Figure 3-7: Comparison of UU Triaxial to FVST

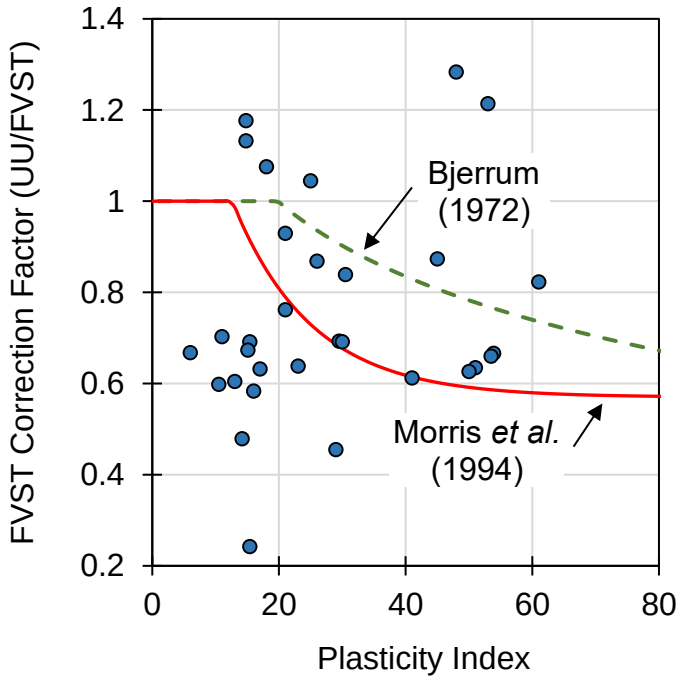



Figure 3-8: Influence of *PI* on FVST Correction Factor

Estimates for the undrained strengths by SPT can be improved with the *PI*, as shown in Figure 3-9, or water content. Ratios of $C_u$ to corrected N-values can be selected between 4 to 7 based on Kulhawy and Mayne (1990), and silty soils will likely use the upper portion of the range.

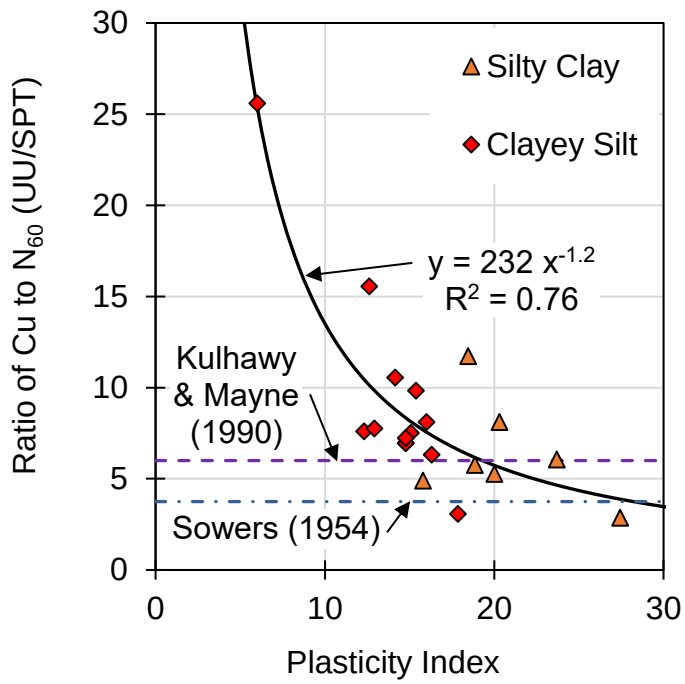



Figure 3-9: Influence of *PI* on Relation of $C_u$ by UU Triaxial and Corrected SPT N-Value

### 3.4.2 Correlations with Side Resistance

The contribution of the side resistance to the ultimate capacity was expressed as a ratio ($Qs/Qu$), and the distribution of the values are represented as a boxplot in Figure 3-10. A pile was classified to be embedded in clay, silt, or sand in the figure if at least 66 % of the pile length was covered by one soil type. Otherwise, the pile had "mixed" soil conditions. The whiskers indicate the minimum and maximum values that were obtained, and ends of the boxes are the 75 % and 25 % percentiles. The horizontal line that divides the boxes is the median. A ratio closer to 100 % indicates a friction pile, while an end bearing pile has a low ratio. From the limited size of sampled load tests, piles in sand tend to be end bearing piles and had slightly lower ratios than piles in clay and silt. Yet, few piles were close to a pure end bearing or frictional pile, and most of them were combination piles.

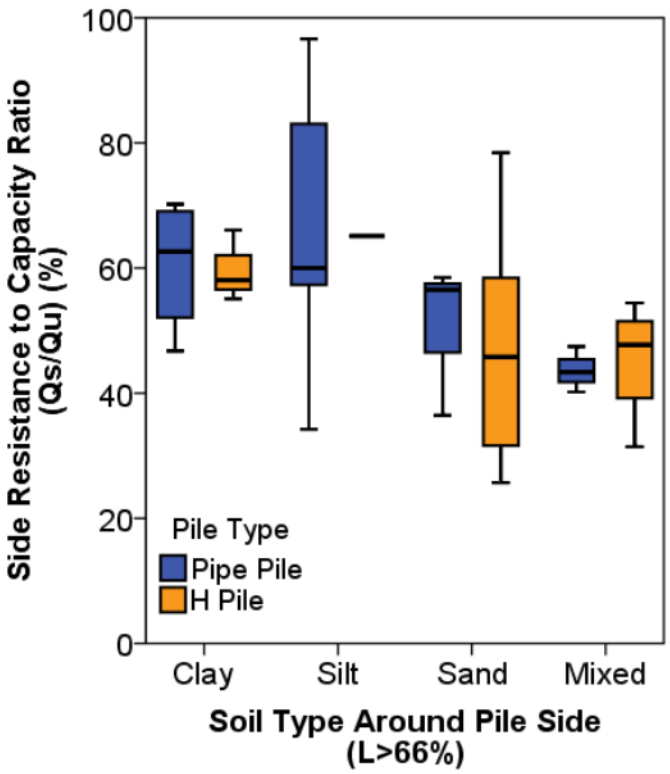



Figure 3-10: Distribution of the Contribution from the Side Resistance to the Pile Capacity by Soil Type

The side resistance is likely influenced by the horizontal confinement or lateral earth pressure of the soil; however, it is challenging to estimate relationships with this variable without sufficient lab testing or in-situ measurements. In Figures 3-11 and 3-12, the unit side resistance is compared to the mean effective stress and pile length. The figures demonstrate a general negative trend with the mean side resistance, but there are a lot of inconsistencies since other factors are involved. In addition, the average soil parameters simplify the soil conditions. As a note, the figures show that the piles in cohesive soils are generally installed at shallower depths compared to the piles in noncohesive soils.

In Figure 3-13, correlations with corrected SPT N-values and the side resistances were also separated by the soil conditions. A consistent trend was found with clayey soils. Shioi and Fukui (1982) likely had softer soils and provided a higher rate for the trend line. The methods for low and large displacement piles by Meyerhof (1976) will generally overestimate the side resistance. Since the piles in this database are fairly slender, they may be classified as low displacement piles if the method by Meyerhof (1976) is used; nonetheless, this existing design method will still overestimate, especially for very dense sands. Meyerhof (1976) developed his method with concrete and timber piles, which will likely have different frictional resistances than steel piles.

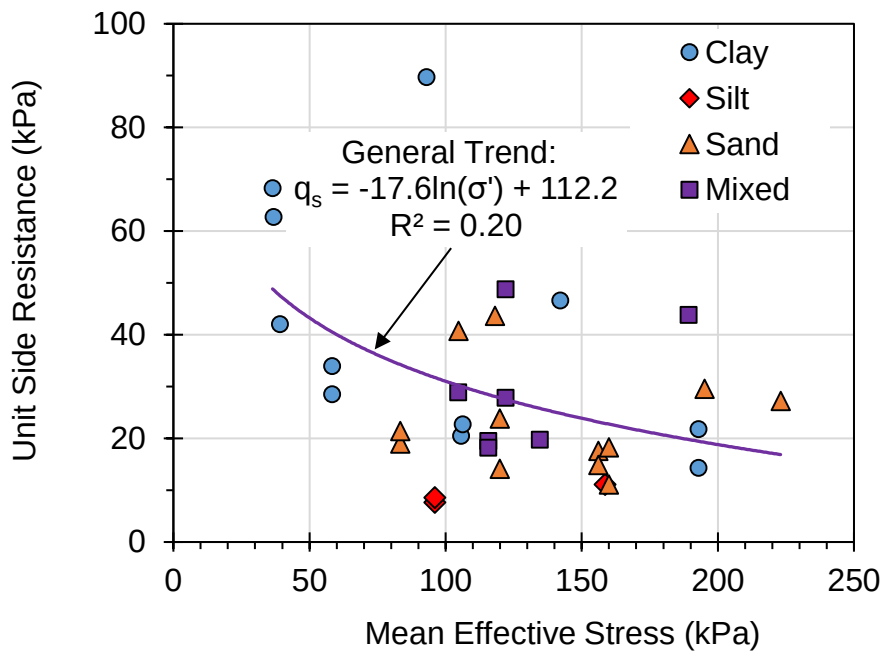



Figure 3-11: Relationship between Unit Side Resistance and Mean Effective Stress

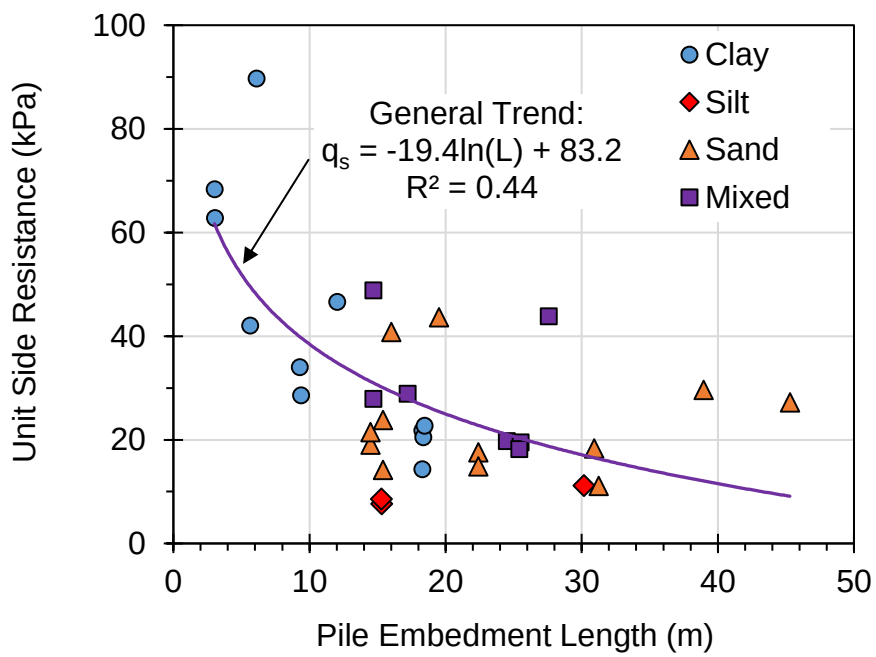



Figure 3-12: Relationship between Unit Side Resistance and Pile Embedment Length

For cohesive soils, Cu can sufficiently predict the side resistance. The relationship with Cu is similar between H piles and pipe piles, as shown in Figure 3-14, but H piles tend to have less variability. For pipe piles, the linear equation is below ($R^2 = 0.91$):

$$q_s = 0.19\ C_u \qquad \text{Eq. 3-1}$$

H piles have the following relation ($R^2 = 0.98$):

$$q_s = 0.17\ C_u \qquad \text{Eq. 3-2}$$

For the selected sample, the $\alpha$ Method recommended by CGS (2006) overestimates the side resistance. These results may change with a larger sample size.

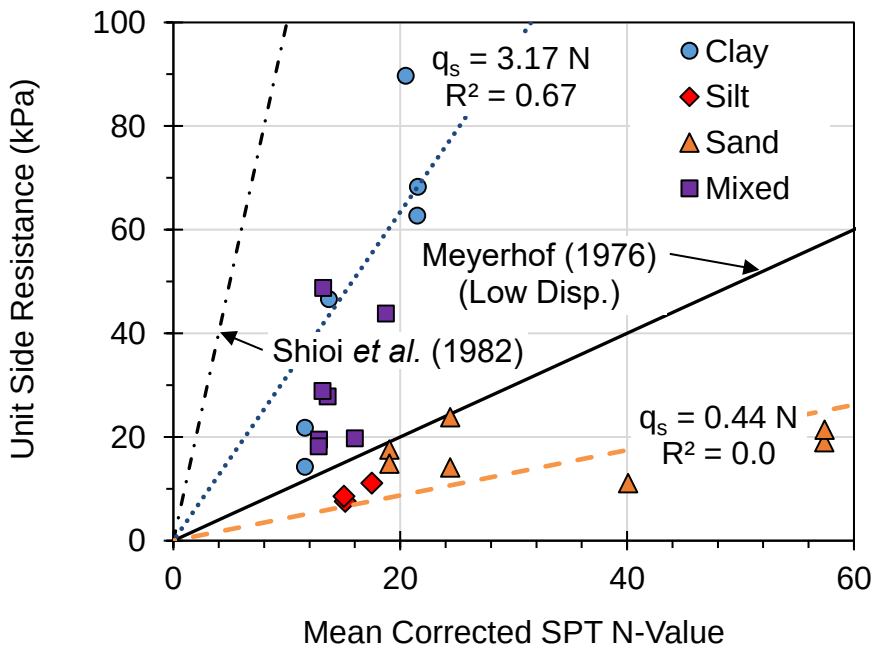



Figure 3-13: Relationship between Unit Side Resistance and Mean Corrected N-Value

In Figure 3-15, the side resistance for cohesive soils was highly correlated to the liquidity index (*LI*). Low *LI*s and high cohesions were found with desiccated clays, and weaker soils generally had higher indices.

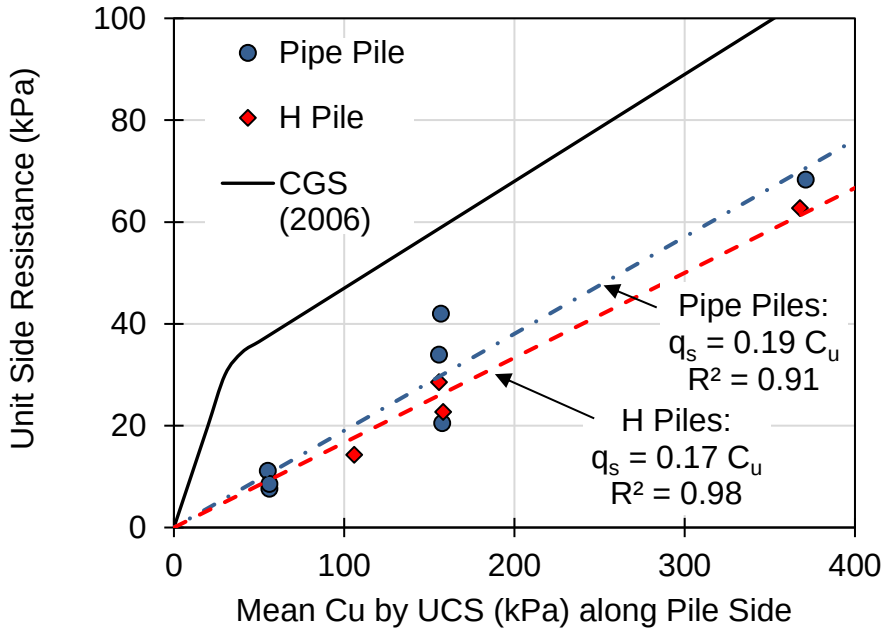



Figure 3-14: Relationship between Unit Side Resistance and Mean $C_u$ by UCS

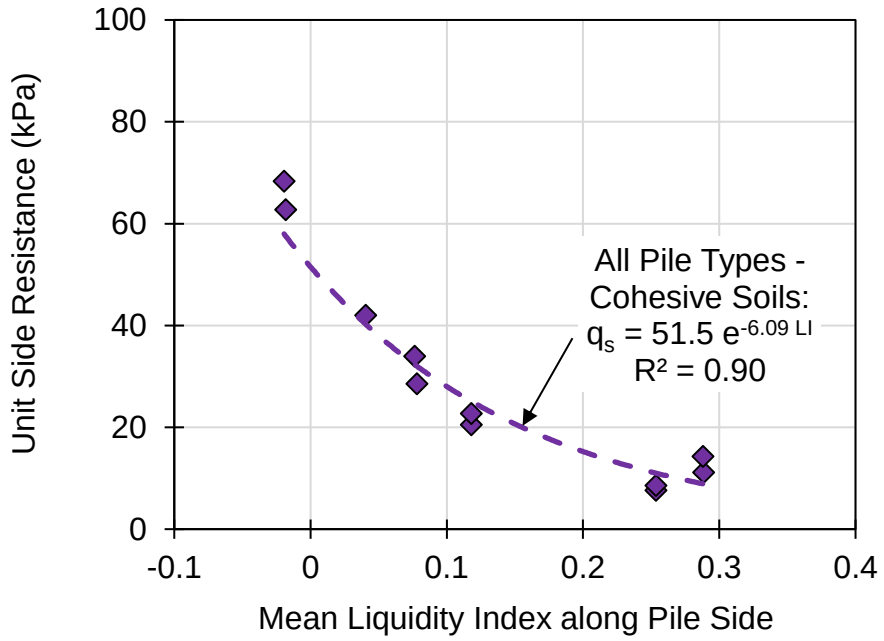



Figure 3-15: Relationship between Unit Side Resistance and Mean Liquidity Index

### 3.4.3 Correlations with Tip Resistance

The boxplot in Figure 3-16 displays the distributions of the unit tip resistances from the collected piles, and the statistical representations of the plot are the same as Figure 3-10. Clayey soils typically generated lower tip resistances than piles in sandy soils. Also, H piles tend to have a higher unit tip resistance for the same soil type, but the unit tip resistance highly depends on other soil properties, such as the strength and compressibility.

As shown in Figures 3-17 and 3-18, any influence by the effective stress or pile length appears to be weak or almost insignificant. Possibly due to installation effects or the soil strength, the pile length may have a positive trend with the unit tip resistance, especially with silty and sandy soils, but it the relationship is very weak.

The unit tip resistance was heavily influenced by the soil content, and this relationship was best displayed with the clay fraction in Figure 3-19. At the pile tip, all the soils had a high silt fraction of at least 33 %, and the silt fraction is indicated by the numbers in the figure. Sandy soils had much higher tip resistances than clays, but the soils likely became more well-graded or cohesive as the clay fraction increased. As a result, the soil strength and tip resistance was higher than soils with mainly silt and sand.

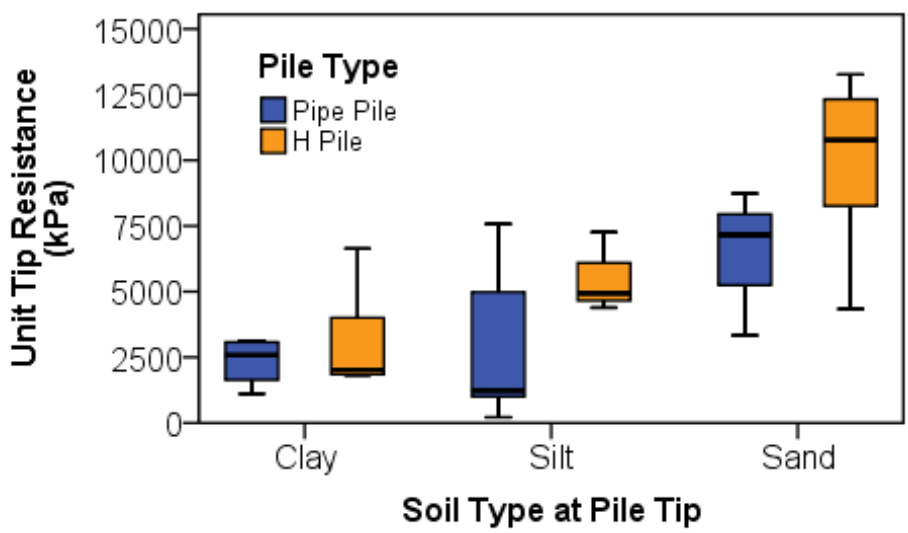



Figure 3-16: Distribution of the Unit Tip Resistance by Soil Type

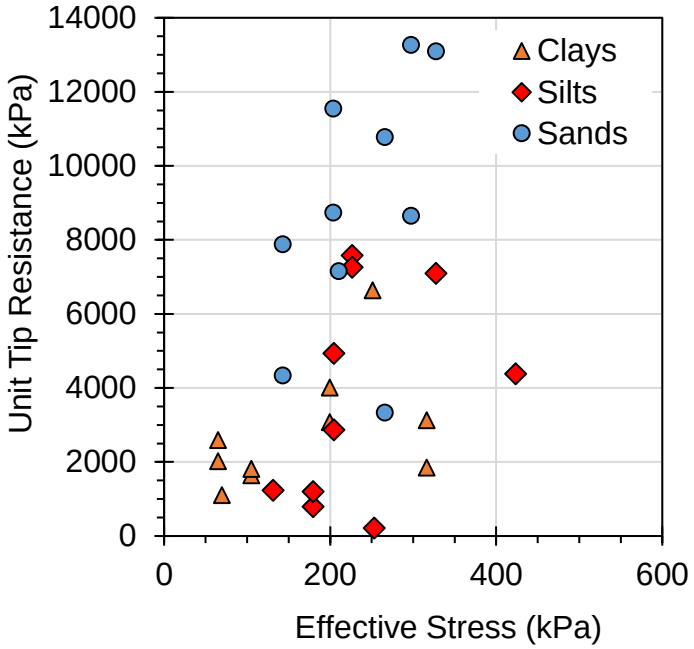



Figure 3-17: Influence of Effective Stress on Unit Tip Resistance

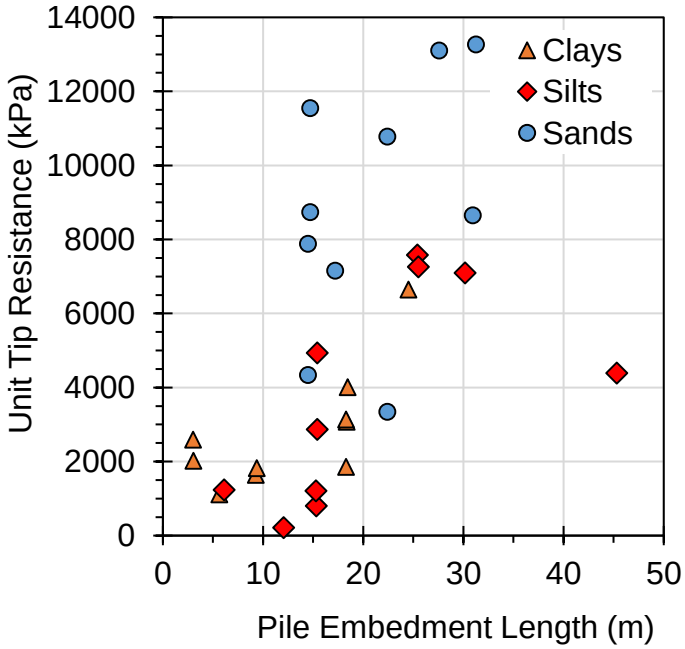



Figure 3-18: Influence of Pile Length on Unit Tip Resistance

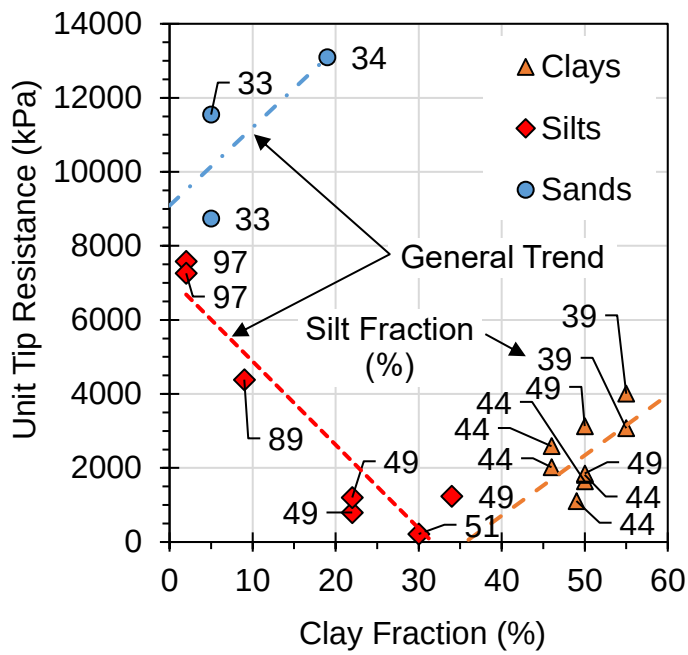



Figure 3-19: Influence of Soil Content on Unit Tip Resistance

The N-values measured closest to the pile tip were collected, and the measurements were usually within 0.5 m above or 1 m below the pile tip. The unit tip resistances ($q_p$) from the piles were plotted against the uncorrected ($N$) and corrected ($N_{corr}$) SPT N-values in Figures 3-20 and 3-21, respectively. $N_{corr}$ corresponds to $N_{60}$ for a cohesive soil or $(N_{60})_1$ for a noncohesive soil. Uncorrected N-values had a logarithmic trend with the tip resistance, while a linear relationship was provided by corrected values. The decrease in fitness between the figures was likely due to the overburden correction factor, which was for general soil conditions. H piles were assumed to be fully plugged and had a slightly higher unit tip resistance than pipe piles. The difference is probably due to the pile geometry. The end-bearing sands were stiffer on average than the clays and silts, but the soil type did not have a large impact on the trend. A common trend line can then represent the soil and pile types with corrected N-values ($R^2 = 0.47$):

$$q_p = 304\, N_{corr} \qquad \text{Eq. 3-3}$$

Since the soils were usually stiff and silty, Equation 3-3 is closer to the correlation by Meyerhof (1976) for sandy soils than the relation from Decourt (1982) for clayey soils.

In Figure 3-22, the $C_u$ cannot predict the tip resistance of the piles, and the poor correlation could be due to the varying silt content and disturbance of the soil samples. The common approach by Meyerhof (1976) tends to be a conservative estimate.

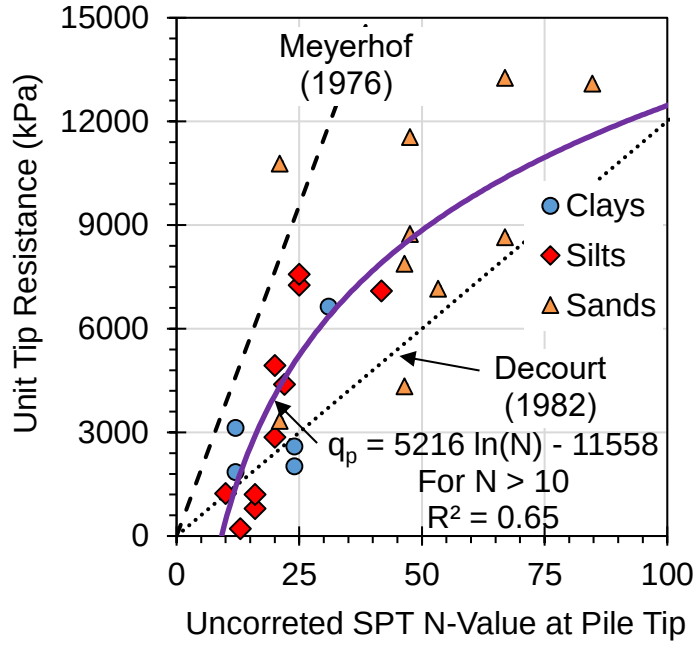



Figure 3-20: Relationship between Unit Tip Resistance and Uncorrected N-Value

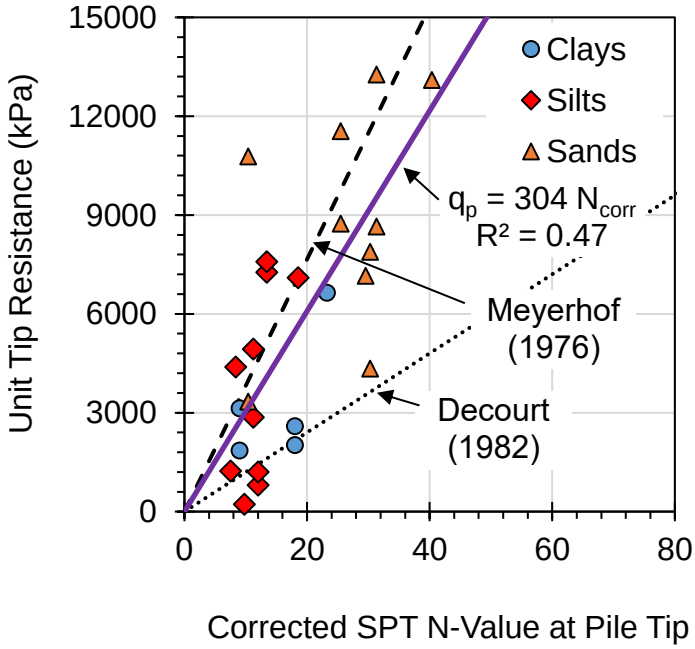



Figure 3-21: Relationship between Unit Tip Resistance and Corrected N-Value

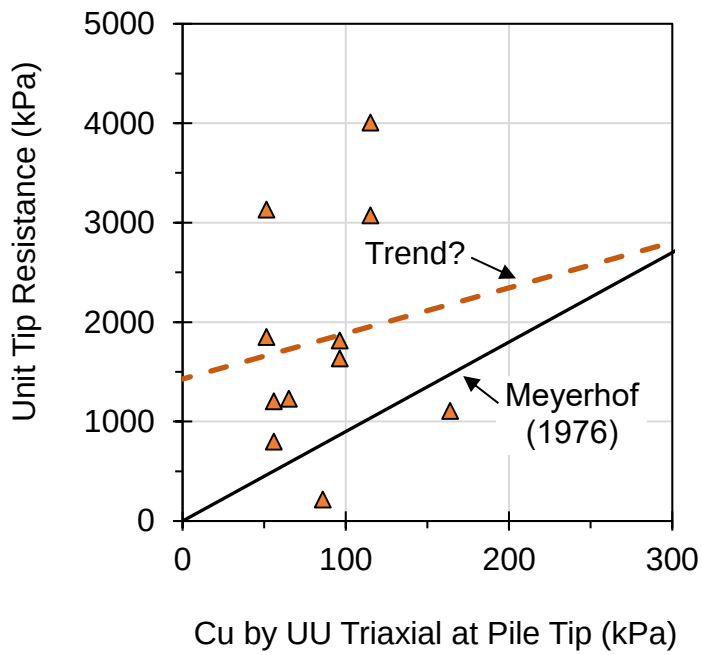



Figure 3-22: Relationship between Unit Tip Resistance and $C_u$ by UU Triaxial

# 4 PREDICTIONS BY A GENETIC ALGORITHM

## 4.1 INTRODUCTION

As displayed in Figure 4-1, the GA evolved the chromosomes with the following steps: creation, evaluation, selection, crossover, mutation, and constant refinement. A GA can be created with various programming languages, but the GA in this study was programmed with Matlab (Mathworks, 2017) because the software can easily manipulate matrices. The source code for this thesis is provided in Appendix B. Since many of the features were customized, toolboxes from Matlab were not used, but the GA was modified from C++ code from Oltean (2016). The GA encoded and evaluated the chromosomes with the Multi Expression Programming (MEP) technique. MEP was inspired by the activation of programs or programming code with integers and can efficiently encode or decode functions for symbolic regression compared to other techniques (Oltean & Dumitrescu, 2002). The settings of the GA are shown in Table 4-1.

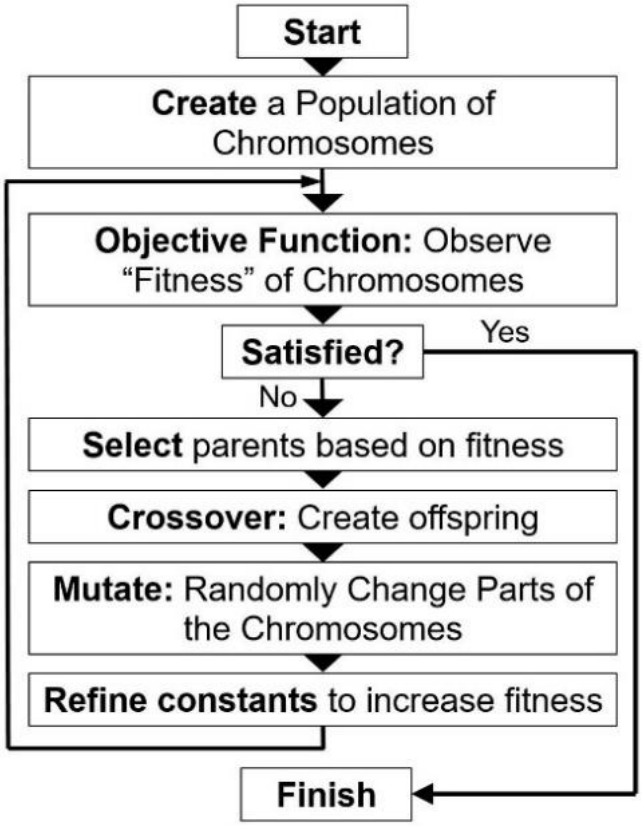



Figure 4-1: Process of the Developed Genetic Algorithm

Table 4-1: Settings for the Genetic Algorithm

| Parameter | Parameter Setting |
|---|---|
| Number of generations | 100 |
| Population Size | 2000 |
| Function Set | $+, -, \times, \div$, power, exponential, natural logarithmic, hyperbolic tangent |
| Chromosome length | 20 |
| Fitness function | Root Mean Squared Error (RMSE) |
| Mutation rate (%) | 10 |
| Crossover rate (%) | 90 |
| Crossover type | Uniform with brood recombination |
| Population size for brood crossover | 8 |
| Brood crossover rate (%) | 50 |
| Population size for brood constant refinement | 50 |
| Tournament selection size | 2 |
| Initial operator likelihood (%) | 40 |
| Initial variable likelihood (%) | 30 |
| Initial constant likelihood (%) | 30 |

Table 4-2: Variables Included in the GA Analysis

| Analysis | Variables[1] | Remarks |
|---|---|---|
| Unit side resistance and corrected SPT N-values | $N_{corr}, S_T, \sigma', (L-z)/D$ | Dataset was divided by pile type; piles were embedded in various soil types |
| Unit side resistance and $C_u$ | $C_u, LI, PI, (L-z)/D, \sigma'$ | All pile types were analyzed together; piles were mainly embedded in cohesive soils |

Notes: [1] $N_{corr}$ is the corrected SPT N-value: $N_{60}$ for cohesive soils and $(N_1)_{60}$ for noncohesive soils. $S_T$ is the soil type, where 1 is for noncohesive soils and 2 is for cohesive soils. $(L-z)/D$ is the modified slenderness ratio that is composed of the pile embedment length ($L$), depth to the centre of a pile segment ($z$), and maximum pile width or diameter ($D$).

The GA was intended to improve predictions for side resistance, and the side resistances were extracted from the piles subjected to extension load tests. Table 4-2 summaries the variables

included in the regressions, and regressions were made with two types of soil strength measurements: corrected SPT N-values and $C_u$. As a result, the GA analyzed two datasets with different soil types. Similar to Brown (2001), H piles were assumed to be fully plugged in both cases, but the analysis could as easily be conducted assuming unplugged conditions.

### 4.1.1 Creation of Chromosomes

For symbolic regression with the MEP technique, chromosomes were linear entities, or arrays, and represented a function for the unit side resistance of a pile. An example to decode a MEP chromosome is shown in Figure 4-2. Chromosomes are generally composed of genes. In this case, for the MEP technique, genes are a single component of a function or a row within the arrays, and these genes are divided into two components: activators and links.

In the figure, activators are entered into the first column of the chromosome and specify which variable, constant, or operator is applied for a function. Links are stored into the last two columns (columns 2 and 3) of the chromosome and indicate which row within the chromosome should be combined by the operators. The activators in the chromosome represent operators with negative integers, such as "-1" for addition and "-2" for multiplication. Positive integers designate the activation of variables and constants. In general, "1" and "2" respectively represent $a$ and $b$, which can be either a constant or variable. The first activator in the chromosome must be a constant or variable to prevent illogical errors during evaluation (Oltean & Dumitrescu, 2002). Next to each activator are the corresponding links, and the links indicate which results should be combined by the operator. The values of the links, which are in parentheses, are the locations, or row numbers of the chromosome, for the operators to be performed. During the evaluation stage, the result of every gene is stored, and the operators are applied to the results from the previous portions of the chromosome. The corresponding links of variables and constants are then ignored because these function entities are numerical values.

For this investigation, the GA was capable of power, natural logarithmic, exponential, and hyperbolic tangent operators in addition to simple arithmetic. These operators allow the GA to consider many common nonlinear relationships. More can be added, but if too many operators are included, the number of possible combinations for a GA to search can increase exponentially (Banzhaf *et al*., 1998). Other operators would also likely apply redundant relationships and may not offer additional benefits.

As the first stage of the GA, a population of chromosomes was created with a maximum length of 20 genes. In other words, the matrix representing the chromosome had a width of 20 rows. A probability was assigned for the likelihood of occurrence for the operators, constants, and variables. These probabilities generate a roulette wheel and allow the genes to be populated with a slight bias towards a particular functional component. In Table 4-1, the probabilities were fairly even since a small number of variables were analyzed. Among the possible operators and variables, the probability for selection was equal.

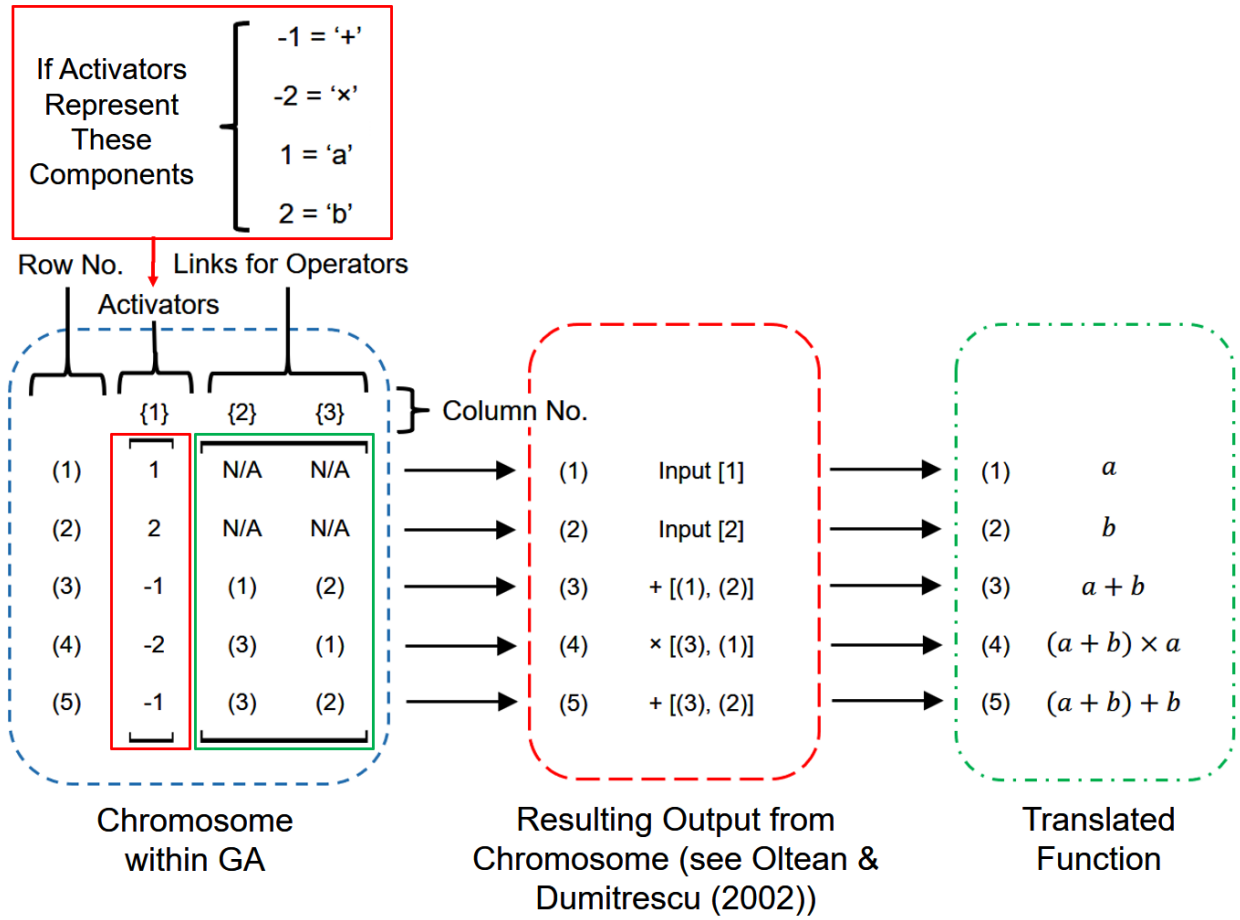



Figure 4-2: Example of Decoding the Chromosome with MEP

### 4.1.2 Evaluation of Chromosomes

The goal of the GA in this study was to find a function with the best fitness to represent the unit side resistance of a pile. During evaluation, the changing shear strength from the soil conditions was considered by dividing the piles into several segments. The GA predicted the unit side resistance for each layer, and the total side resistance of a pile was the summation of the side resistances from all the pile segments. During a load test, load-displacement responses were measured at the top of the piles. Thus, the fitness function compared the predicted total side resistance ($Qs_{pred}$) to the measured side resistance ($Qs$) with the root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum\left(Qs_{pred} - Qs\right)^2} \qquad \text{Eq. 4-1}$$

where *n* is the number of analyzed piles. A lower RMSE indicates a better fit between the measured and predicted values. Other fitness functions exist, such as the absolute error, but the RMSE is fairly simple and common for regression.

Illogical errors must be prevented during evaluation. Examples include dividing by zero or applying the logarithm to a negative value. Division operators may be simply protected by returning the numerator if a near-zero denominator is found (Banzhaf *et al.*, 1998), but Oltean and Dumitrescu (2002) suggested to mutate division into a variable or constant. Operators were protected as suggested by Brameier and Banzhaf (2007) and mutated as shown in Table 4-3.

Table 4-3: Protection Instructions for Operators (Modified from Brameier & Banzhaf, 2007)

| Operation | Protected Definition | |
|---|---|---|
| Division: $r_i = r_j/r_k$ | If $\lvert r_k \rvert \geq 10^{-6} : r_i = r_j/r_k$ | Else $r_i = \text{variable}$ |
| Power: $r_i = r_j^{r_k}$ | If $\lvert r_k \rvert \leq 20 : r_i = \lvert r_j \rvert^{r_k}$ | Else $r_i = \tanh(r_j)$ |
| Exponent: $r_i = \exp(r_j)$ | If $\lvert r_j \rvert \leq 32 : r_i = \exp(r_j)$ | Else $r_i = r_j + r_k$ |
| Natural Logarithm: $r_i = \ln(r_j)$ | If $\lvert r_j \rvert \geq 10^{-6} : r_i = \ln(\lvert r_j \rvert)$ | Else $r_i = \text{constant}$ |

### 4.1.3 Selection of Parents

Two parents were selected for mating using tournament selection. The process is simple and requires less computational effort compared to other selection methods (Zhong *et al*., 2005). For each parent, a tournament size, or group, of chromosomes were randomly sampled from the population. The chromosome from this group with the best fitness became a parent, and the selection process was repeated until a new population was created with the same size as the original population. As part of process for tournament selection, the offspring replace the chromosomes with the worst fitness in the original population.

### 4.1.4 Crossover and Mutation

Crossover was assigned a 90 % chance of occurrence. If crossover does not occur, the parents were copied and sent for mutation. Otherwise, uniform crossover was applied with brood recombination. Uniform crossover randomly distributes the genes of the parents to the offspring, as demonstrated in Figure 4-3. Brood recombination was inspired by organisms that have a litter of offspring (Tackett, 1994), and the process attempts to obtain the best attributes from the parents. As shown in Figure 4-4, crossover was repeated several times, and a subpopulation ($N_{sub}$) of offspring were created by the same two parents. The two offspring with the best fitness were selected and continued for mutation. If every pair of parents performed brood recombination, the computational

effort would increase significantly because the total number of created offspring would be $N_{sub}$ multiplied by the chromosome population size (Banzhaf *et al*., 1998). Thus, brood recombination was assigned a probability of occurrence.

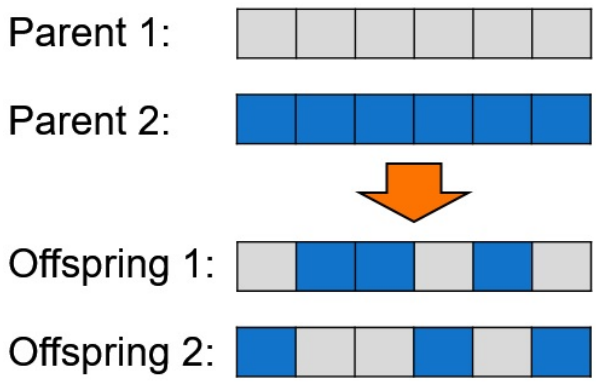



Figure 4-3: Example of Uniform Crossover

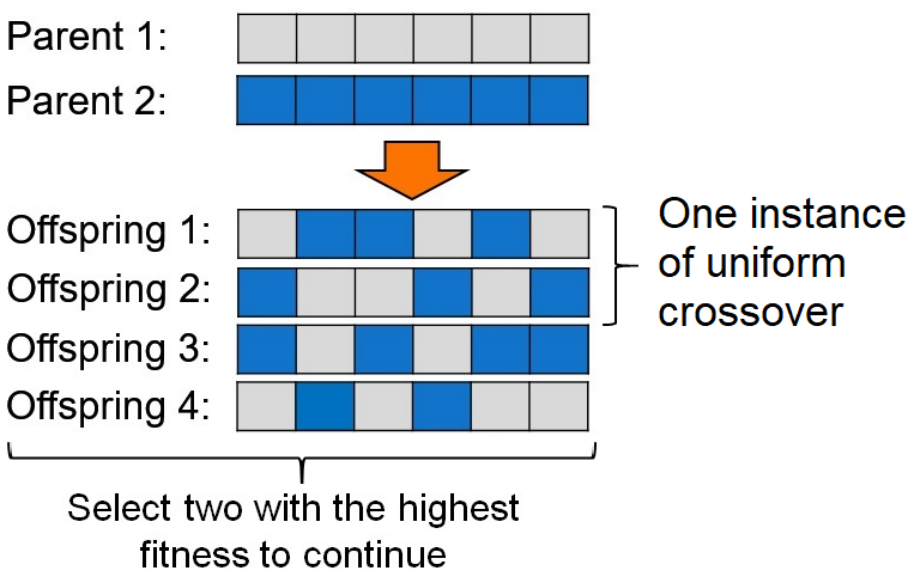



Figure 4-4: Example of Uniform Crossover with Brood Recombination

During mutation, a selected gene would be transformed randomly into a different component type. For example, a multiplication operator could become a constant, or a variable could become another variable. The selection is random, and the process is not completely constructive, like a

genetic disease. The chance of mutation was set to a low 10 %, but the process is necessary to add diversity and prevent cloning.

### 4.1.5 Constant Refinement

Constants in symbolic regression are either non-evolutional or evolutional (Banzhaf *et al*., 1998). Non-evolutional constants are kept the same throughout a generation, but their influence and values may be manipulated by operations from the GA. For evolutionary constants, the values are modified by optimization techniques, such as the Levenberg-Marquardt algorithm (Marquardt, 1963) or Nelder-Mead simplex method (Nelder & Mead, 1965). These methods are usually mathematically complex and iterative. They may also take numerous trials to terminate on a potential solution, especially with highly nonlinear relationships and multiple variables. Since the population of chromosomes may be large, the computational effort should be reasonably minimized. Brood recombination was applied by this GA as a simple approach to refine the constants. Values of the constants for every chromosome were randomly changed with 50 attempts, and the values that provided the best fitness were kept as the new constants.

## 4.2 CORRELATIONS BETWEEN SIDE RESISTANCE AND SPT N-VALUES

For this analysis, 23 piles (9 pipe piles and 14 H piles) with high-quality soil and test results were selected. Each pile was divided into 50 segments to consider the varying shear resistance along their length. The number of segments was determined by several conditions: N-values were frequently measured along the piles; piles were usually 15 to 25 m in length, but one reached 45 m long; and the length of each segment was considered by the corresponding side area ($A_s$) during calculations. Mainly for short piles with less data, the soil parameters were linearly interpolated. Wolff (1989) applied a similar approach while predicting the capacity. The variables in the dataset included the corrected SPT N-value ($N_{corr}$), soil type ($S_T$), effective stress ($\sigma'$), and pile slenderness ratio ($(L-z)/D$). $N_{corr}$ is $N_{60}$ for cohesive soils and $(N_1)_{60}$ for noncohesive soils. Due to obstructions in the ground, such as gravels and boulders, N-values occasionally reached magnitudes greater than 100, but they were limited to a maximum of 60 for the GA analyses. Due to the unreliability of high N-values, Brown (2001) recommended a similar limit for his design methods. The soil type was a binary variable equal to 1 for noncohesive soils and 2 for cohesive soils. Since a developed function would be intended for heterogeneous soil profiles, the slenderness ratio was modified as suggested by Kolk and Van der Velde (1996), and it was composed of the embedment length ($L$), depth to the centre of a pile segment ($z$), and maximum pile width or diameter ($D$). For SPT measurements, the dataset was divided for H piles and pipe piles, and the GA performed 5 trials with each pile type (a total of 10 runs) to regress the variables and test results. Since genetic programming is a stochastic method, especially with brood recombination, several trials were performed to independently create 2000 functions over the generations.

While correlating the side resistance with SPT N-values, the GA analyzed both pipe piles and H piles separately. The plots in Figures 4-5 and 4-6 show the average and lowest RMSE within the population of chromosomes for pipe and H piles. For the 5 trials, the stochastic nature of the GA can be observed, and the final fitness depends on the sampling of the population and occurrence of events throughout the generations. In Figure 4-5, the population of chromosomes developed similar characteristics in Trial 1 and had a higher RMSE than other trials. For Trial 4, the functions began to increase in size, or bloat, and the GA terminated with a low RMSE. The remaining trials maintained a balance in diversity and function length, and these trials terminated with a very similar fitness. In Figure 4-6, the fitness among the trials was consistent after 40 generations. Trial 5 quickly developed a low RMSE for the best function after the first generation, and the small differences in the average and best fitness after 35 generations may indicate cloning between functions. After 60 generations, better links or constants were probably found, and the functions began to diversify. In general, the analysis usually terminated with a similar RMSE for both pile types. As better constants were found and links were made, the brood recombination resulted in sudden drops in the best fitness throughout the generations.

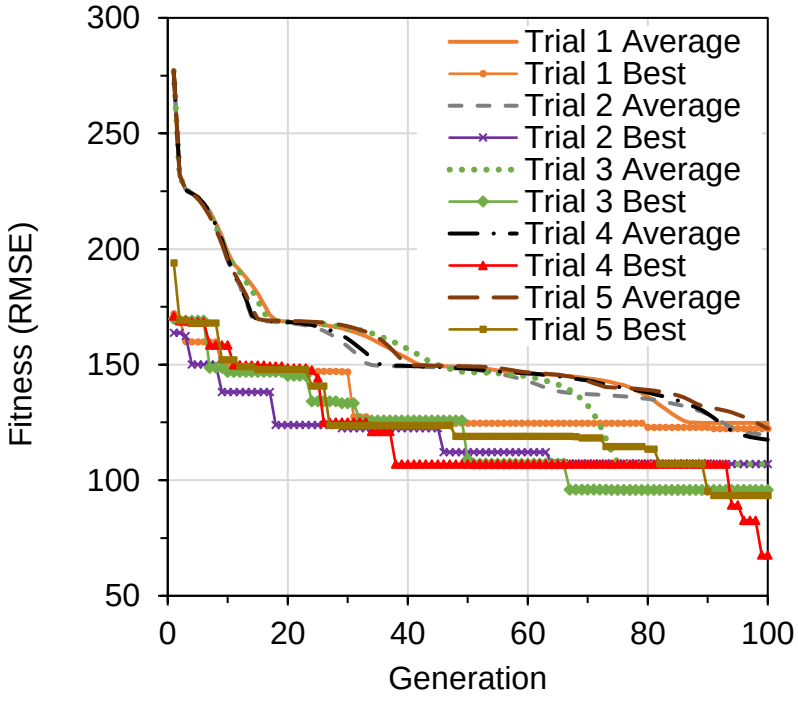



Figure 4-5: Fitness Performance with Pipe Piles

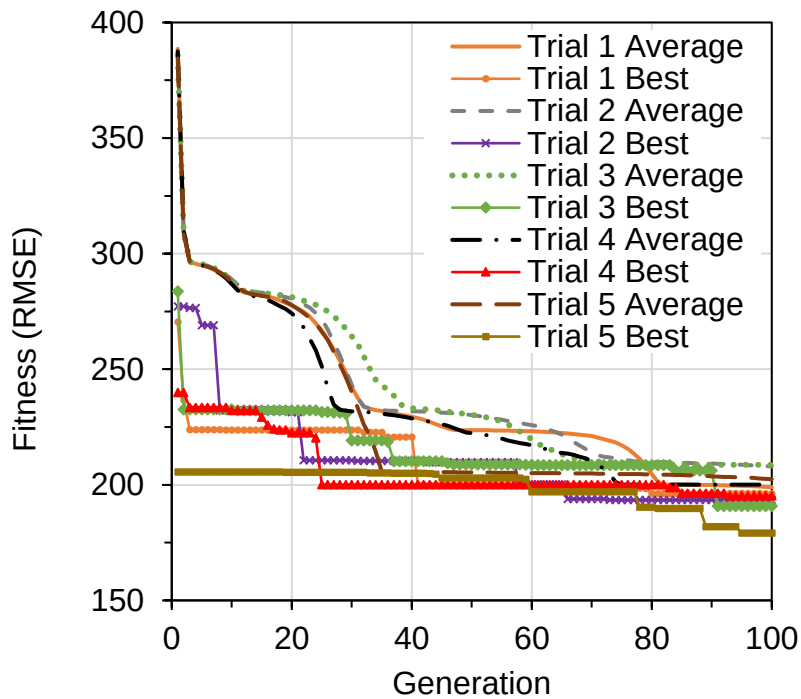



Figure 4-6: Fitness Performance with H Piles

From the final generation of the 5 trials, the chromosomes were pooled together to create a population of 10000 functions. Table 4-4 shows percentage of total functions that contained each variable at least once. The slenderness ratio was commonly found in 98.3 % and 100 % of the functions for pipe piles and H piles, respectively. Especially for H piles, the effective stress or installation effects provided the best correlations, and the corrected SPT N-value was less frequently found. It was used in 63.3 % and 42.6 % of the functions for pipe and H piles, respectively. The reduced frequency of the N-values for H piles could be related to the assumed plugging conditions.

Table 4-4: Percentage of Total Functions that Apply Each Variable

| Variable | Functions for Pipe Piles (%) | Functions for H Piles (%) |
|---|---|---|
| Corrected SPT N-Value, $N_{corr}$ | 63.3 | 42.6 |
| Soil Type, $S_T$ | 56.1 | 1.6 |
| Effective Stress, $\sigma'$ | 0.1 | 61.8 |
| Slenderness Ratio, $(L-z)/D$ | 98.3 | 100.0 |

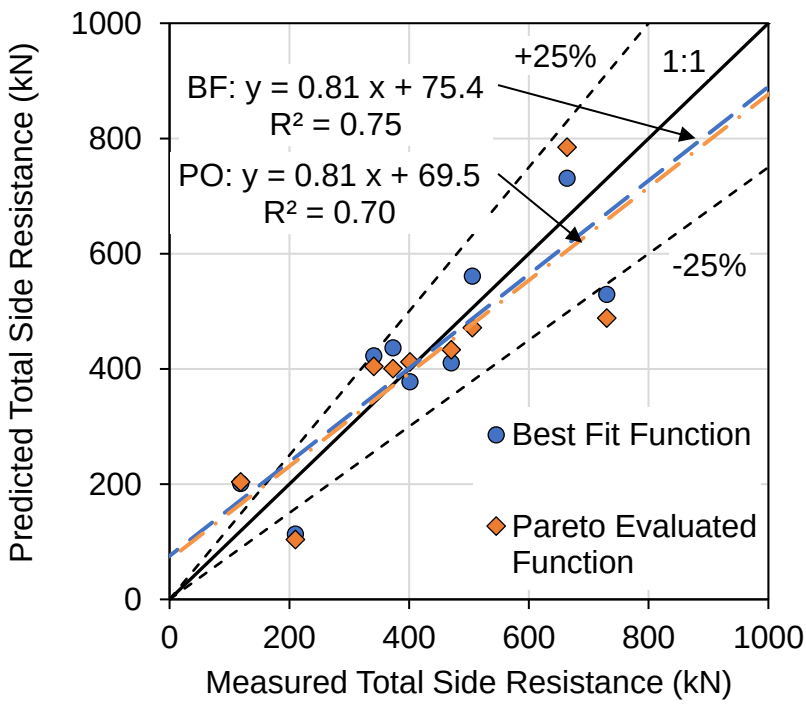



Figure 4-7: Comparison of the Measured and Predicted Side Resistance by the GA for Pipe Piles

It is logical for the SPT N-values to be included into any recommended function. Thus, among the functions that applied the N-value as a variable, the one with the best fitness was collected for each pile type. The RMSE by itself may be misleading as H piles had on average a higher unit side resistance than pipe piles, but H piles received slightly better correlations. Figures 4-7 and 4-8 show the best-fit (BF) functions for both pile types had a reasonably good $R^2$. For pipe piles, the $R^2$ was 0.75 for the BF function:

$$q_s = 4.8\,[N_{corr} + (L - z)/D]/\{S_T\,[(L - z)/D]^{0.85}\} \qquad \text{Eq. 4-2}$$

The $R^2$ was 0.88 for plugged H piles:

$$q_s = 0.12\,[\sigma' + 5 \cdot N_{corr} \cdot D/(L - z) + 8.5] \qquad \text{Eq. 4-3}$$

In these functions, $N_{corr}$ is the corrected SPT N-value; $L$ is the embedment length of a pile; $D$ is the diameter or width of a pile; $z$ is the depth to the centre of a pile segment or soil layer; $S_T$ is the soil type that is equal to 1 for noncohesive soils and 2 for cohesive soils; and $\sigma'$ is the effective stress. The BF function for pipe piles had consistent results with an average predicted to measured ratio ($Qs_{pred}/Qs$) and coefficient of variance (COV) of 1.04 and 30.0 %, respectively. An average ratio of 0.99 and COV of 29.8 % was provided by the BF function for H piles. Piles with cohesive

soils provided most of the inconsistencies, and the GA rarely considered the soil type since the sites dominated in noncohesive soils.

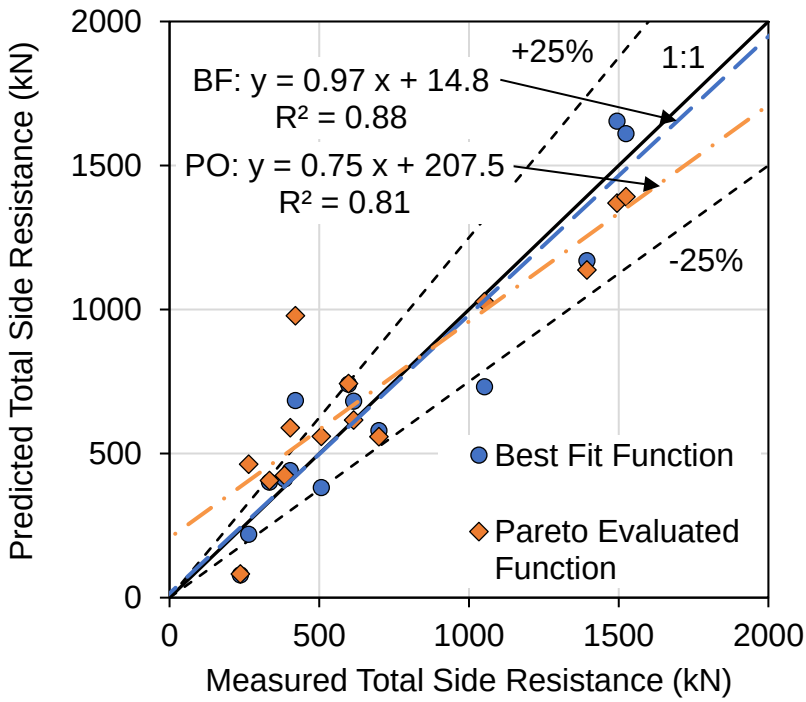



Figure 4-8: Comparison of the Measured and Predicted Side Resistance by the GA for H Piles

The BF functions may not always be practical, beneficial, or appropriate; as a response, another function was selected for each pile type by Pareto optimization (PO). From the final generation, functions were analyzed if they contained the SPT N-value as a variable, and these functions were graphically evaluated by their fitness and complexity. The complexity is the number of components within a function, and the Pareto front was created in Figures 4-9 and 4-10 by finding the best fitness for each complexity. In general, a lower complexity, or shorter function, likely results in a poorer fitness compared to a longer function, which can manipulate data with more operations. The orange square markers indicate points along the Pareto front, and the blue circle markers are the remaining results. Any point on the Pareto front can be a potential solution; thus, the preferred solution mainly relies on the tolerable error and judgement from the investigator (Smits and Kotanchek, 2005).

For pipe piles, the RMSE does not improve by much for a complexity between 9 and 13, but a shorter function results in a significantly higher RMSE. The function with 9 components was selected, had an $R^2$ of 0.70, and was simplified to the following:

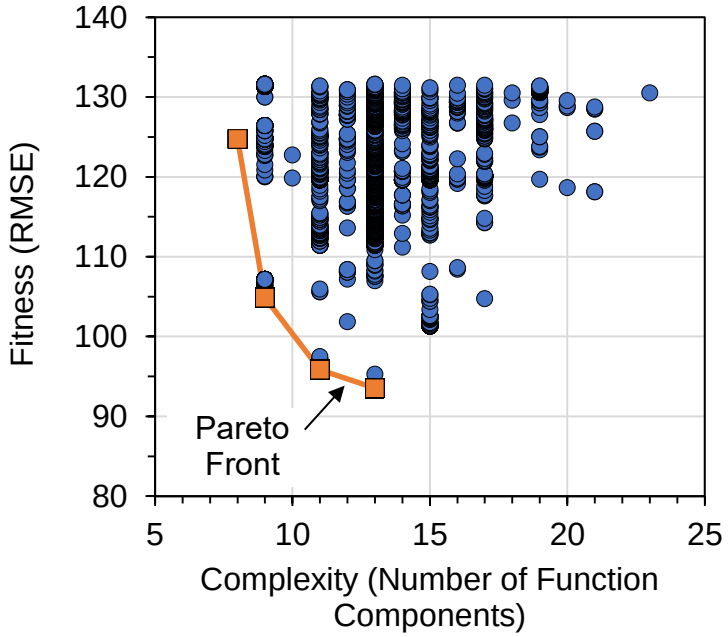



Figure 4-9: Pareto Front from GA Results for Pipe Piles

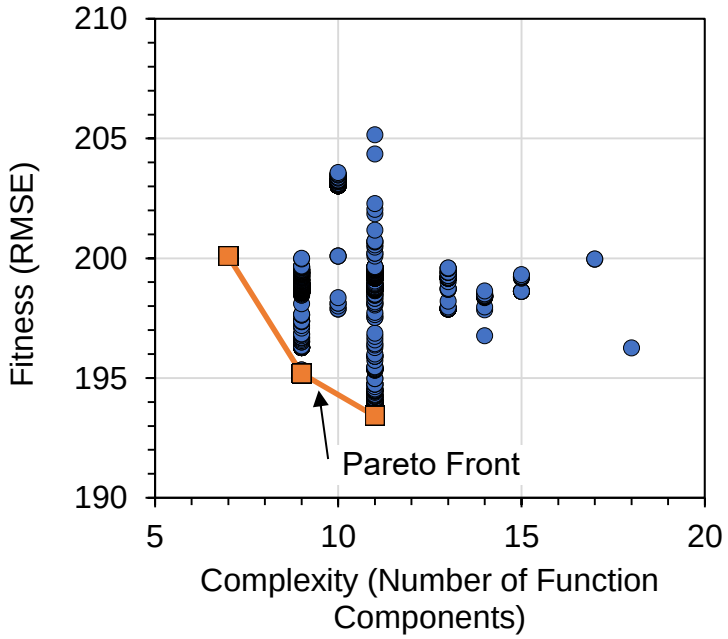



Figure 4-10: Pareto Front from GA Results for H Piles

$$q_s = [2.8 \cdot N_{corr} \cdot D/(L-z) + 4.5]/S_T \qquad \text{Eq. 4-4}$$

Since small percentage of the 10000 functions applied the SPT N-value, the Pareto front for H piles was short. The front was also linear with small improvements between complexities. The shortest function with a complexity of 7 was chosen, and the correlation was fairly strong with an $R^2$ of 0.81:

$$q_s = \sigma'/[(1.9/N_{corr}) + (L-z)/D] \qquad \text{Eq. 4-5}$$

The fitness of Equations 4-4 and 4-5 is displayed in Figures 4-7 and 4-8. The functions from the Pareto front have a small difference in the $R^2$ compared to the BF functions. The PO function for pipe piles tend to underestimate long piles or piles with high capacities, but the $Qs_{pred}/Qs$ is 1.02 on average. The COV is 32.0 %. For H piles, the PO function trends to overestimate shorter piles and has a higher average $Qs_{pred}/Qs$ of 1.14 compared to the BF function. The spread of predictions also increases with a COV of 39.9 %. Overall, the PO and BF functions offer good predictions.

## 4.3 Correlations between Side Resistance and Undrained Shear Strength

Nine piles were analyzed with $C_u$ measurements, and the piles were divided into 25 segments. The number of segments was chosen with the following conditions: $C_u$ measurements were less frequently measured compared to SPT N-values; and piles were fairly short with lengths usually between 10 m to 18 m. The soil parameters were linearly interpolated for piles with limited data. For this analysis, the variables included $C_u$, $LI$, $PI$, the slenderness ratio, and $\sigma'$. The $\sigma'$ was considered because many design methods apply the strength ratio ($C_u/\sigma'$). Since some sites only had UCS measurements, the values were converted to UU results by applying a factor shown in Figure 3-6. The number of piles with $C_u$ measurements was small; thus, the pile types were analyzed together. The correlation analysis by the GA was conducted with 5 trials.

For the 5 trials, the lowest RMSE terminated within small range of approximately 32 to 38, as shown in Figure 4-11. The 10000 functions from the final generation were collected and analyzed. The slenderness ratio was frequently found in 99.1 % of the functions, but the effective stress was the least common variable and was found in 12.5 % of functions. This may indicate a strong influence due to the installation effects, but the strength ratio, which can be empirically related to the OCR, may have a weak effect. Table 4-5 displays the usage of the variables within the functions.

Preference was given to the 65.9 % of functions that applied $C_u$, and the function with the best fitness was selected from this group:

$$q_s = |0.17(L-z)/D - C_u \cdot D/[0.17(L-z)] + 5.3|^{\tanh[0.17(L-z)/D]} \qquad \text{Eq. 4-6}$$

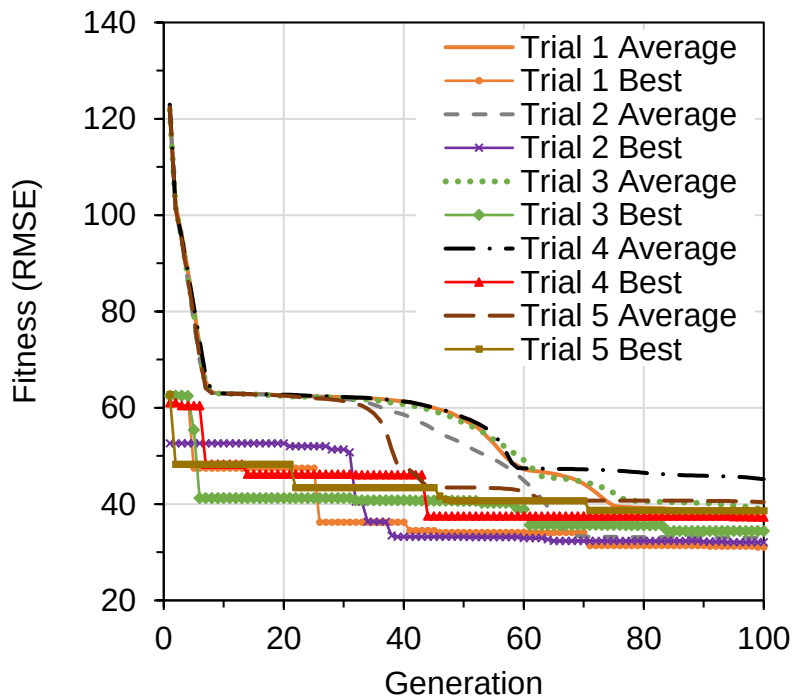



Figure 4-11: Fitness Performance with Piles in Cohesive Soils

Table 4-5: Percentage of Total Functions that Apply Each Variable

| Variable | Functions for Cohesive Soils (%) |
|---|---|
| Undrained Shear Strength, $C_u$ | 65.9 |
| Plasticity Index, $PI$ | 55.3 |
| Liquidity Index, $LI$ | 24.5 |
| Effective Stress, $\sigma'$ | 12.5 |
| Slenderness Ratio, $(L-z)/D$ | 99.1 |

where $L$ is the embedment length of a pile; $D$ is the diameter or width of a pile; $z$ is the depth to the centre of a pile segment or soil layer; and $C_u$ is the undrained shear strength from UU triaxial tests. This complex function with 16 components had a $R^2$ of 0.91, as shown in Figure 4-12. The fitness is very similar in Figure 4-13 for the various function complexities along the Pareto front, and the Pareto front is linear as a result. The shortest PO equation with a complexity of 7 had an $R^2$ of 0.92, which is the same as the BF function:

$$q_s = C_u^{(LI+0.90)} \cdot D/(L-z) \qquad \text{Eq. 4-7}$$

where $LI$ is the liquidity index of the cohesive soil.

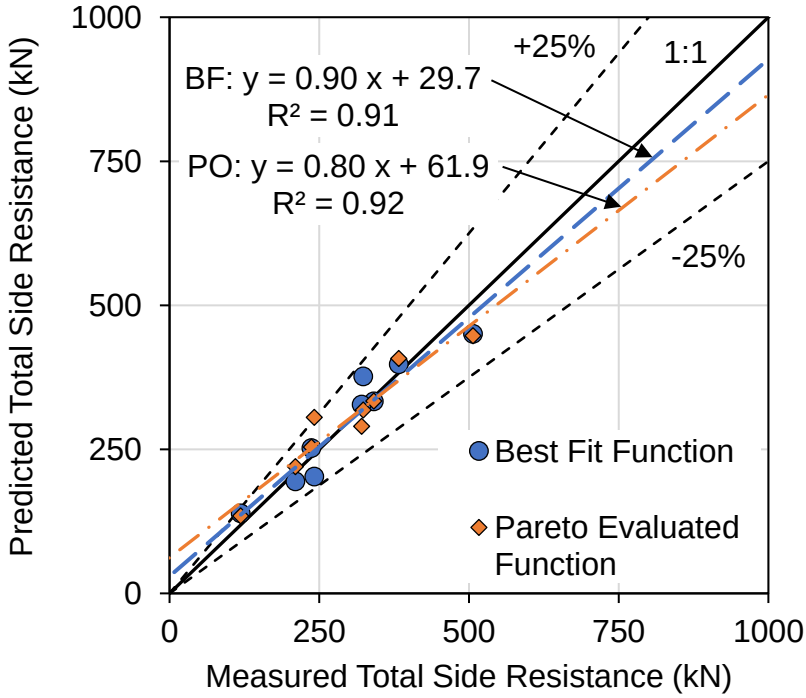



Figure 4-12: Comparison of the Measured and Predicted Side Resistance by the GA for Piles in Cohesive Soils

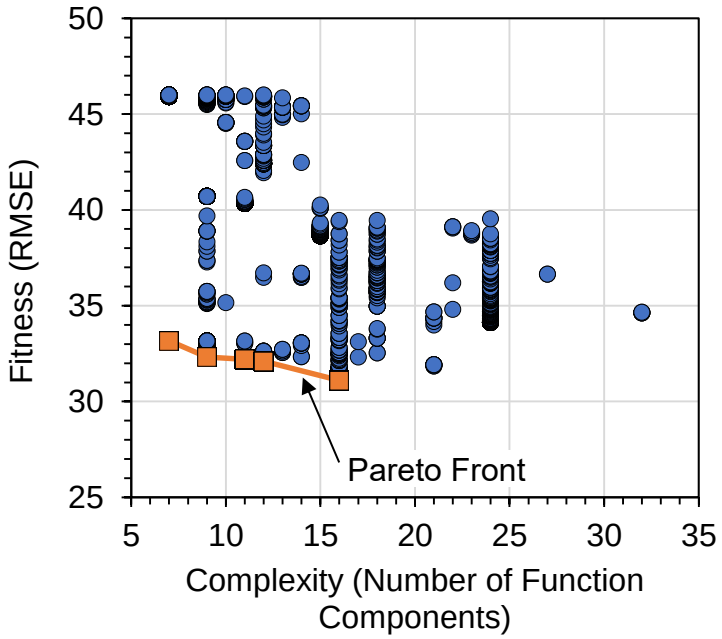



Figure 4-13: Pareto Front from GA Results for Piles in Cohesive Soils

## 4.4 PERFORMANCE OF EXISTING DESIGN METHODS

### 4.4.1 Direct Methods with SPT

The side resistance of the piles was calculated with design methods that were intended for both cohesive and noncohesive soils: Shioi and Fukui (1982), Decourt (1982), and Brown (2001). Shioi and Fukui (1982) provided separate design methods for cohesive and noncohesive soils, but Brown (2001) provided a general equation for all soil types. For the calculations, N-values were corrected and limited as suggested by the references, and H piles were assumed to be fully plugged as suggested by Brown (2001). The results of the predictions are provided in Figures 4-14 to 4-16.

The three design methods greatly overestimated the side resistance and gave erratic results. While comparing the predicted and measured values for pipe piles, a decent fitness could not be established with a logical linear relationship. The approach by Brown (2001) had the worst performance with an average $Qs_{pred}/Qs$ of 2.51 and 2.97 for pipe and H piles, respectively. The method by Decout (1982) gave the best results among the existing design methods, but predictions were 2.34 to 2.36 times higher on average than the measured values. Piles in clays or very stiff soils had the greatest over predictions. These design methods were developed with weaker soils, and they are not applicable or were intended for Ontario since stiff, silty, desiccated, or OC soils are commonly found.

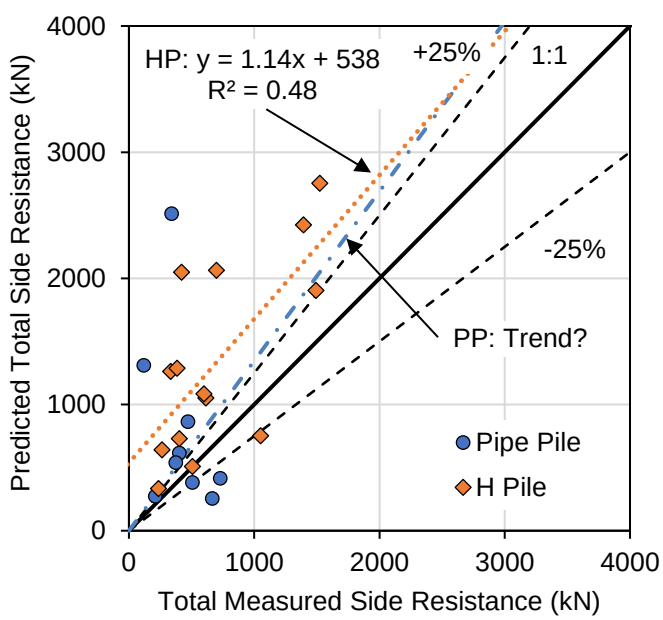



Figure 4-14: Comparison of Measured and Predicted Side Resistances by Shioi and Fukui (1982)

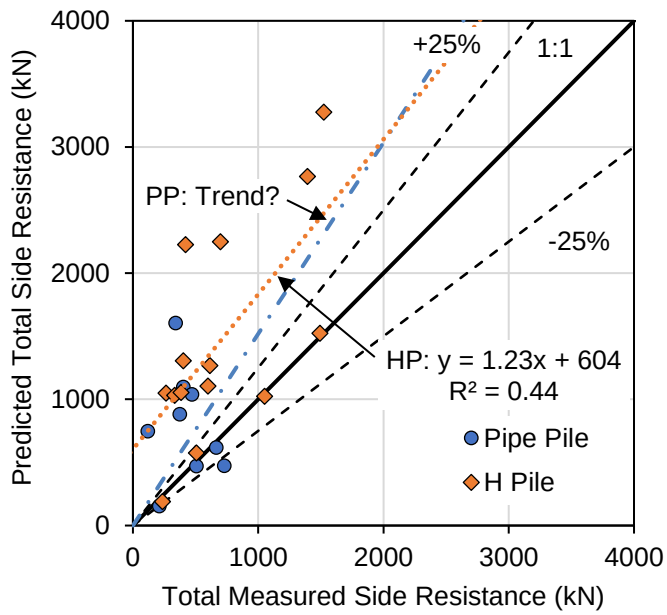



Figure 4-15: Comparison of Measured and Predicted Side Resistances by Decourt (1982)

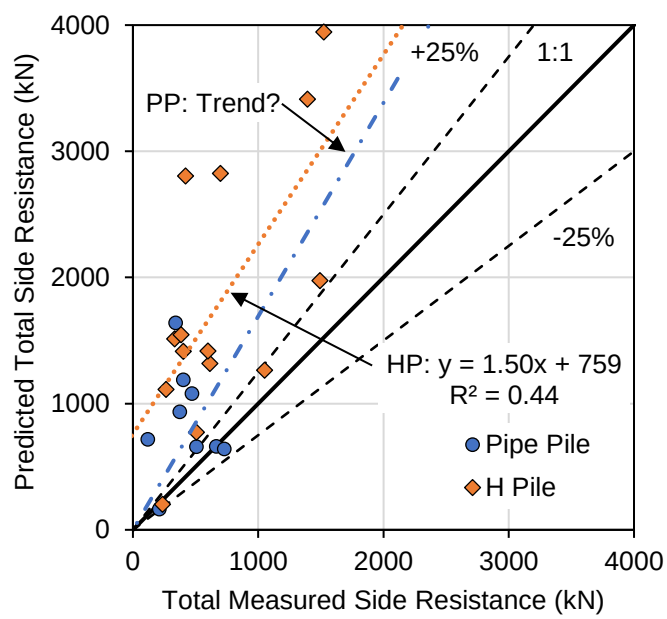



Figure 4-16: Comparison of Measured and Predicted Side Resistances by Brown (2001)

Table 4-6 summarizes the descriptive statistics for the $Qs_{pred}/Qs$ ratio and includes the mean, standard deviation ($\sigma_{SD}$), COV, and number of samples or pile tests (n). The table also contains

the results from the GA with the best-fit functions (Equations 4.2 and 4.3) and Pareto optimal functions (Equations 4.4 and 4.5). The predictions from the GA have a low bias as the mean $Qs_{pred}/Qs$ ranges from 0.99 to 1.10, and the standard deviation for the ratio is also much lower than the existing design methods with values ranging from 0.29 to 0.46. From existing methods, the inconsistent predictions demonstrate a need for locally developed design methods for steel piles in Ontario soils and heterogeneous soil profiles.

Table 4-6: Descriptive Statistics of the Predicted to Measured Side Resistance Ratio for SPT Design Methods

| Design Method | Pile Type | Mean | $\sigma_{SD}$ | COV (%) | n |
|---|---|---|---|---|---|
| Shioi & Fukui (1982) | H Piles | 2.19 | 1.12 | 51.2 | 14 |
| | Pipe Piles | 2.91 | 3.50 | 120 | 9 |
| | All | 2.47 | 2.38 | 96.3 | 23 |
| Decourt (1982) | H Piles | 2.34 | 1.28 | 54.5 | 14 |
| | Pipe Piles | 2.36 | 1.88 | 79.4 | 9 |
| | All | 2.35 | 1.54 | 65.5 | 23 |
| Brown (2001) | H Piles | 2.97 | 1.56 | 52.7 | 14 |
| | Pipe Piles | 2.51 | 1.75 | 69.6 | 9 |
| | All | 2.79 | 1.65 | 59.3 | 23 |
| GA: Best-Fit Function (Proposed) | H Piles | 0.99 | 0.29 | 29.8 | 14 |
| | Pipe Piles | 1.04 | 0.31 | 30.0 | 9 |
| | All | 1.01 | 0.30 | 30.0 | 23 |
| GA: Pareto Optimal Function (Proposed) | H Piles | 1.14 | 0.46 | 39.9 | 14 |
| | Pipe Piles | 1.02 | 0.33 | 32.0 | 9 |
| | All | 1.10 | 0.42 | 37.8 | 23 |

### 4.4.2 Direct and Indirect Methods for Cohesive Soils

For this study, piles were selected if at least 80 % of the length was covered by cohesive soils. The predicted ultimate capacity, $Qu_p$, was calculated for piles subjected to compression and tension. For compression load tests, the various design methods were combined for the side and tip resistance. Prediction methods were chosen that applied different concepts and parameters. At the same time, consideration was made for methods that are commonly used in practice and may not rely on many parameters. The side resistance was calculated by the $\alpha$ and $\beta$ Methods summarized in Table 4-7. The α Methods included those suggested by CGS (2006), Kolk and Van der Velde (1996), and Karlsrud *et al.* (2005). One β Method was applied in the analysis: Patrizi and Burland

(2001). The multiple and changing soil measurements were considered by dividing the soil profiles into smaller layers along a pile, and the analysis included the weight of the piles. In general, $Qu_p$ was calculated by Equation 2-1 if the pile was subjected to a compression load test; otherwise, the capacity of an extension load test was calculated by the design methods for the side resistance.

Table 4-7: Studied Design Methods for Side Resistance in Cohesive Soils

| Method Type | Reference | Condition | Equation for $\alpha$ or $\beta$ |
|---|---|---|---|
| $\alpha$ Method | CGS (2006) | | $\alpha = 0.21 + (26/C_u) \leq 1$ |
| | Karlsrud *et al.* (2005)[1] | Case 1 ($C_u/\sigma$' < 0.25): | $\alpha = 0.32\,(PI - 10)^{0.3} \leq 1$ |
| | | Case 2 ($C_u/\sigma$' > 1)[2]: | $\alpha = 0.5\,(C_u/\sigma')^{-0.3} F_{Tip}$ |
| | | Case 3: | Linear interpolate between Case 1 and 2 |
| | Kolk and Van der Velde (1996)[3] | | $\alpha = 0.9[(L - z)/D]^{-0.2}(\sigma'/C_u)^{0.3} \leq 1$ |
| $\beta$ Method | Patrizi and Burland (2001) | | $\beta = 0.1 + 0.4\,(C_u/\sigma')$ |

Notes: [1] The effect of increased soil strength with time was not included; [2] $F_{Tip}$ is a factor for the pile tip geometry; [3] $z$ is the depth that corresponds to the midpoint of a soil layer along a pile.

The tip resistance was predicted throughout the analysis by the common approach from Meyerhof (1976):

$$q_p = 9\,C_u \qquad \text{Eq. 4-8}$$

Usually, the soil parameters had to be estimated with empirical correlations, especially if SPT N-values were obtained. The correlations for $C_u$ in cohesive soils are typically linear relationships with SPT N-values, as shown in Table 4-8. Nassaji and Kalantari (2011) suggested a multilinear relationship would give better estimates with several additional variables, such as the water content ($w_c$) and liquid limit ($LL$).

Table 4-8: Studied Empirical Methods for Cohesive Soils

| Reference | Equation |
|---|---|
| Sowers (1954) | $Cu = 3.75\,N_{60}$ |
| Sivrikaya and Toğrol (2006) | $Cu = 4.22\,N_{60}$ |
| Nassaji and Kalantari (2011) | $Cu = 1.5\,N_{60} - 0.1\,w_c - 0.9\,LL + 2.4\,PI + 21.1$ |

In addition to the several design methods to calculate the side resistance, the capacity was also found by the various combinations to estimate the soil parameters. For example, the $\alpha$ Method recommended by CGS (2006) determines the side resistance and relies on $C_u$. However, $C_u$ can be measured by three methods: UCS, triaxial, and FVST. Plus, the three SPT correlations were applied. In all, the capacity can be calculated in six different ways for the one method to predict the side resistance.

Particularly for H piles, the lowest capacity was selected between fully plugged and unplugged conditions. The plugged condition governed for all the prediction methods, but in Figure 4-17, many inconsistencies were observed among the pile types with the existing design methods. Most of the variability was due to empirical SPT correlations and measurements by FVST. Design methods tend to overestimate the side resistance, especially for short piles. The $\alpha$ Method suggested by CGS (2006) solely relies on the $C_u$ and has the greatest scatter among the design methods. For UCS and UU measurements, the COV of the predicted to measured capacity ($Qu_p/Qu$) ratio was 56 % and 58 %, respectively. The other methods included more variables in the calculations and provided lower COVs that generally range from 42 % to 48 % for lab tests. The approach from Patrizi and Burland (2001) has the lowest COVs among the design methods, but it provides the greatest overestimations on average, which may be expected from a $\beta$ Method. The average $Qu_p/Qu$ ratio was 2.07 for UCS and 2.04 for UU measurements. The method by Karlsrud *et al*. (2005) considers the *PI* in its calculations, and it over predicts the least with an average $Qu_p/Qu$ ratio of 1.55 and 1.57 for UCS and UU triaxial, respectively. Most design methods were intended for soils with higher plasticities; thus, the performance of $\alpha$ Method from Karlsrud *et al*. (2005) may indicate the importance of the soil content.

For the existing design methods, Tables 4-9 to 4-12 provide details on the descriptive statistics for the $Qu_p/Qu$ ratio. Piles subjected to tension loads had large $Qs_p/Qu$ ratios compared to $Qu_p/Qu$ ratios for piles under compression. For example, the mean $Qs_p/Qu$ was 2.35 for the method recommended by CGS (2006) with UCS tests, but the mean $Qu_p/Qu$ for compression tests was 1.49 for UCS tests. This demonstrates that most design methods will overpredict the side resistance. Since the design method from Meyerhof (1976) conservatively predicts the tip resistance, overestimations for compression loads will be lower than tension loads for the same pile. Also, lower estimates were given for piles with higher capacities. These piles are usually long and are end-bearing on stiff soils.

The current methods were compared to predictions by Equation 3-2 from the statistical analysis and Equation 4-7 from the GA. The proposed equations were limited to the unit side resistance of 80 kPa, and the results are in Figure 4-18 and Table 4-13. The proposed methods were conservative, which is mainly due to predictions for the tip resistance. For UU measurements, the average $Qu_p/Qu$ ratio was 0.76 for both methods. Equation 3-2 from the statistical analysis was a conservative selection as the proposed equation was only developed with H piles. A higher slope for pipe piles was found. The equation from the GA underestimates because it provides a greater

significance towards the base of a pile due to the relationship with the slenderness ratio, and the unit side resistance was limited to a maximum amount. Yet, the proposed methods were very consistent. The statistical correlation had a COV of 29 % with UCS and UU values. From the GA, the COVs were 33 % and 29 % respectively for UCS and UU values. In the end, the proposed methods offer improvements compared to existing approaches.

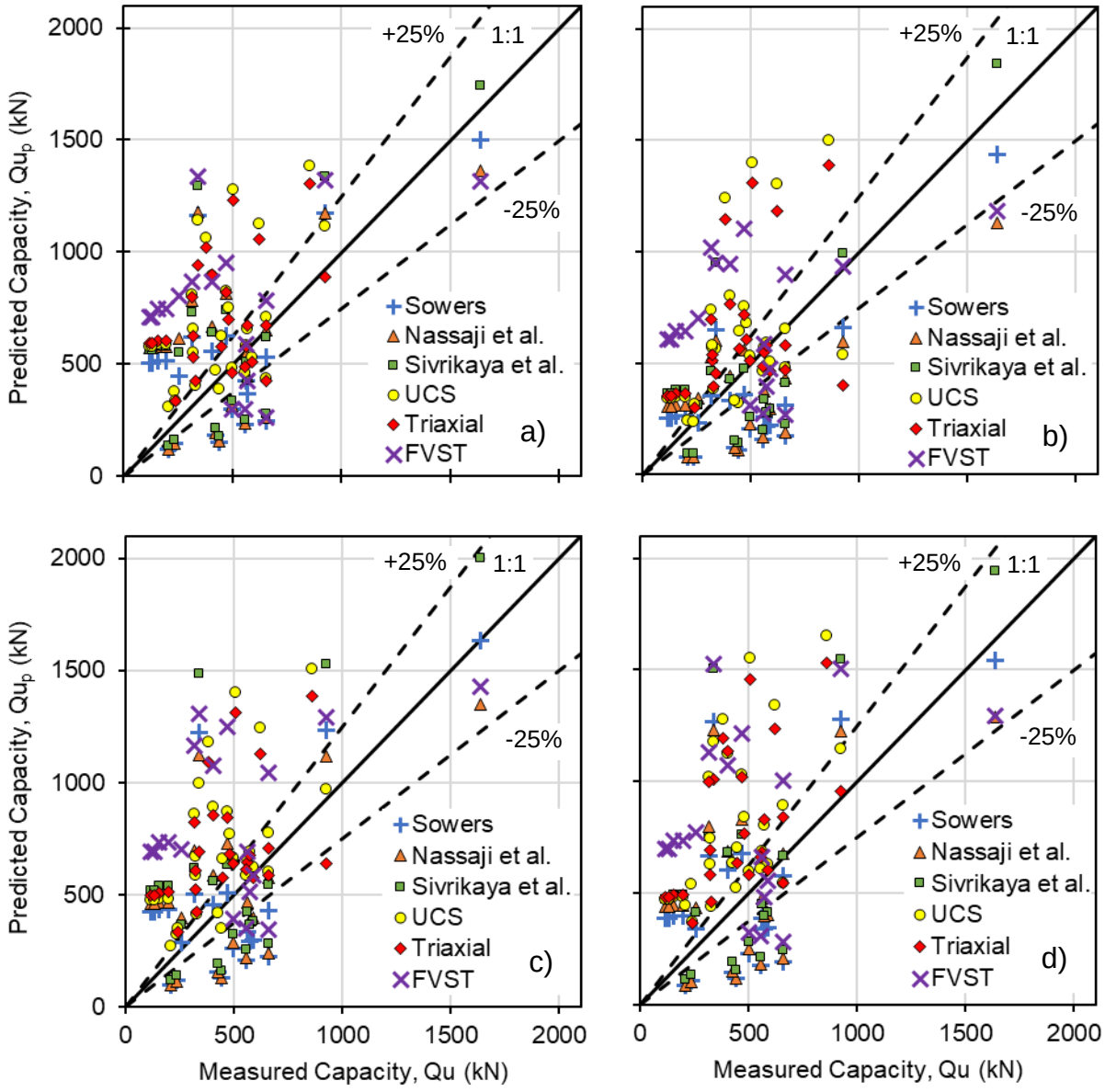



Figure 4-17: Comparison of Measured and Predicted Capacities by a) CGS (2006); b) Karlsrud *et al.* (2005); c) Kolk and Van der Velde (1996); and d) Patrizi and Burland (2001)

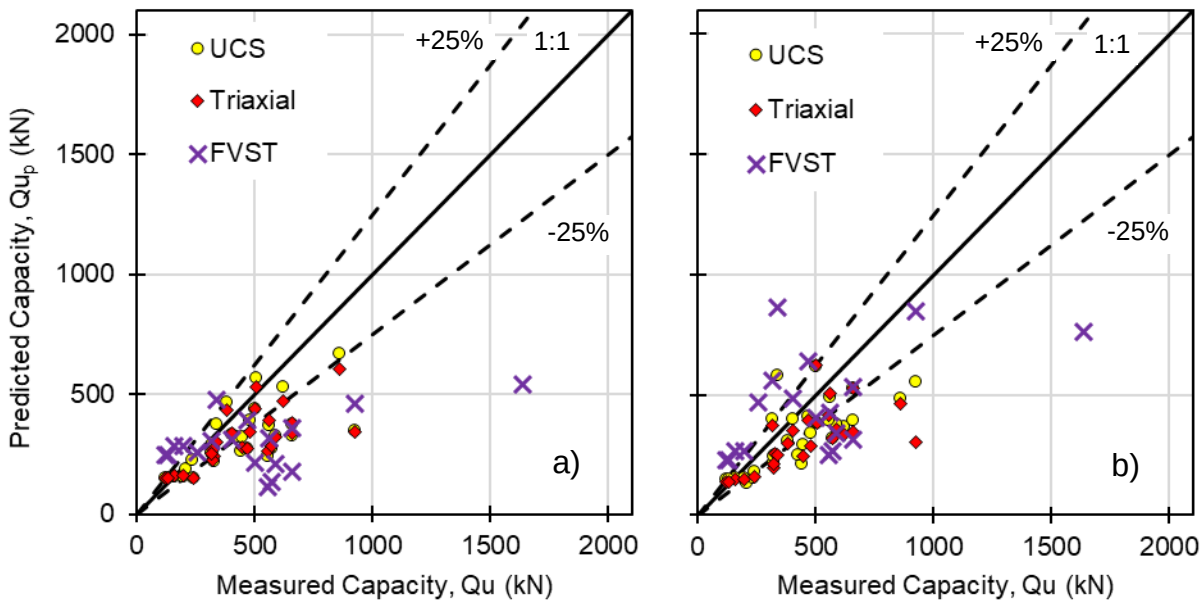



Figure 4-18: Comparison of Measured and Predicted Capacities by the Proposed Method from the a) Statistical Analysis and b) GA

Table 4-9: Descriptive Statistics of the Predicted to Measured Capacity Ratios by CGS (2006)

| Condition | Method to Obtain $C_u$ | Mean | $\sigma_{SD}$ | COV (%) | n |
|---|---|---|---|---|---|
| Tension Loads ($Qs_p/Qu$) | Sowers | 1.89 | 1.45 | 76.8 | 9 |
| | Nassaji *et al.* | 2.12 | 1.61 | 76.1 | 9 |
| | Sivrikaya *et al.* | 2.12 | 1.60 | 75.7 | 9 |
| | UCS | 2.35 | 1.14 | 48.6 | 14 |
| | Triaxial | 2.42 | 1.18 | 48.9 | 12 |
| | FVST | 3.05 | 1.96 | 64.3 | 7 |
| Compression Loads ($Qu_p/Qu$) | Sowers | 1.20 | 0.89 | 73.6 | 13 |
| | Nassaji *et al.* | 1.39 | 1.04 | 75.1 | 13 |
| | Sivrikaya *et al.* | 1.39 | 1.01 | 72.2 | 13 |
| | UCS | 1.49 | 0.78 | 52.5 | 16 |
| | Triaxial | 1.53 | 0.86 | 56.4 | 14 |
| | FVST | 1.92 | 1.41 | 73.6 | 11 |

Table 4-9: Descriptive Statistics of the Predicted to Measured Capacity Ratios by CGS (2006)

| Condition | Method to Obtain $C_u$ | Mean | $\sigma_{SD}$ | COV (%) | n |
|---|---|---|---|---|---|
| All (Capacity) ($Qu_p/Qu$) | Sowers | 1.50 | 1.21 | 80.9 | 22 |
| | Nassaji *et al.* | 1.70 | 1.37 | 80.4 | 22 |
| | Sivrikaya *et al.* | 1.70 | 1.35 | 78.9 | 22 |
| | UCS | 1.89 | 1.06 | 55.9 | 30 |
| | Triaxial | 1.94 | 1.12 | 57.6 | 26 |
| | FVST | 2.38 | 1.75 | 73.4 | 18 |

Table 4-10: Descriptive Statistics of the Predicted to Measured Capacity Ratios by Karlsrud *et al.* (2005)

| Condition | Method to Obtain $C_u$ | Mean | $\sigma_{SD}$ | COV (%) | n |
|---|---|---|---|---|---|
| Tension Loads ($Qs_p/Qu$) | Sowers | 1.03 | 0.72 | 70.2 | 9 |
| | Nassaji *et al.* | 1.21 | 0.84 | 69.5 | 9 |
| | Sivrikaya *et al.* | 1.42 | 1.07 | 75.2 | 9 |
| | UCS | 1.88 | 0.75 | 39.8 | 14 |
| | Triaxial | 1.92 | 0.72 | 37.4 | 12 |
| | FVST | 2.74 | 1.63 | 59.4 | 7 |
| Compression Loads ($Qu_p/Qu$) | Sowers | 0.68 | 0.42 | 61.9 | 13 |
| | Nassaji *et al.* | 0.81 | 0.52 | 64.6 | 13 |
| | Sivrikaya *et al.* | 0.93 | 0.62 | 67.2 | 13 |
| | UCS | 1.26 | 0.49 | 39.0 | 16 |
| | Triaxial | 1.27 | 0.50 | 39.7 | 14 |
| | FVST | 1.67 | 1.18 | 70.6 | 11 |
| All (Capacity) ($Qu_p/Qu$) | Sowers | 0.82 | 0.59 | 71.8 | 22 |
| | Nassaji *et al.* | 0.97 | 0.70 | 71.9 | 22 |
| | Sivrikaya *et al.* | 1.13 | 0.87 | 76.9 | 22 |
| | UCS | 1.55 | 0.70 | 45.0 | 30 |
| | Triaxial | 1.57 | 0.69 | 44.2 | 26 |
| | FVST | 2.09 | 1.47 | 70.3 | 18 |

Table 4-11: Descriptive Statistics of the Predicted to Measured Capacity Ratios by Kolk and Van Der Velde (1996)

| Condition | Method to Obtain $C_u$ | Mean | $\sigma_{SD}$ | COV (%) | n |
|---|---|---|---|---|---|
| Tension Loads ($Qs_p/Qu$) | Sowers | 1.65 | 1.33 | 80.1 | 9 |
| | Nassaji *et al.* | 1.80 | 1.35 | 74.7 | 9 |
| | Sivrikaya *et al.* | 2.02 | 1.61 | 79.8 | 9 |
| | UCS | 2.25 | 0.91 | 40.5 | 14 |
| | Triaxial | 2.27 | 0.93 | 41.1 | 12 |
| | FVST | 3.25 | 1.84 | 56.8 | 7 |
| Compression Loads ($Qu_p/Qu$) | Sowers | 0.98 | 0.72 | 73.5 | 13 |
| | Nassaji *et al.* | 1.11 | 0.78 | 70.3 | 13 |
| | Sivrikaya *et al.* | 1.22 | 0.90 | 73.2 | 13 |
| | UCS | 1.49 | 0.60 | 39.8 | 16 |
| | Triaxial | 1.51 | 0.67 | 44.2 | 14 |
| | FVST | 1.92 | 1.28 | 66.6 | 11 |
| All (Capacity) ($Qu_p/Qu$) | Sowers | 1.26 | 1.07 | 84.8 | 22 |
| | Nassaji *et al.* | 1.39 | 1.10 | 79.2 | 22 |
| | Sivrikaya *et al.* | 1.55 | 1.30 | 84.0 | 22 |
| | UCS | 1.85 | 0.85 | 45.9 | 30 |
| | Triaxial | 1.86 | 0.89 | 47.7 | 26 |
| | FVST | 2.44 | 1.66 | 67.9 | 18 |

Table 4-12: Descriptive Statistics of the Predicted to Measured Capacity Ratios by Patrizi and Burland (2001)

| Condition | Method to Obtain $C_u$ | Mean | $\sigma_{SD}$ | COV (%) | n |
|---|---|---|---|---|---|
| Tension Loads ($Qs_p/Qu$) | Sowers | 1.65 | 1.33 | 80.1 | 9 |
| | Nassaji *et al.* | 1.80 | 1.35 | 74.7 | 9 |
| | Sivrikaya *et al.* | 2.02 | 1.61 | 79.8 | 9 |
| | UCS | 2.25 | 0.91 | 40.5 | 14 |
| | Triaxial | 2.27 | 0.93 | 41.1 | 12 |
| | FVST | 3.25 | 1.84 | 56.8 | 7 |

Table 4-12: Descriptive Statistics of the Predicted to Measured Capacity Ratios by Patrizi and Burland (2001)

| Condition | Method to Obtain $C_u$ | Mean | $\sigma_{SD}$ | COV (%) | n |
|---|---|---|---|---|---|
| Compression Loads ($Qu_p/Qu$) | Sowers | 0.98 | 0.72 | 73.5 | 13 |
| | Nassaji *et al.* | 1.11 | 0.78 | 70.3 | 13 |
| | Sivrikaya *et al.* | 1.22 | 0.90 | 73.2 | 13 |
| | UCS | 1.49 | 0.60 | 39.8 | 16 |
| | Triaxial | 1.51 | 0.67 | 44.2 | 14 |
| | FVST | 1.92 | 1.28 | 66.6 | 11 |
| All (Capacity) ($Qu_p/Qu$) | Sowers | 1.26 | 1.07 | 84.8 | 22 |
| | Nassaji *et al.* | 1.39 | 1.10 | 79.2 | 22 |
| | Sivrikaya *et al.* | 1.55 | 1.30 | 84.0 | 22 |
| | UCS | 1.85 | 0.85 | 45.9 | 30 |
| | Triaxial | 1.86 | 0.89 | 47.7 | 26 |
| | FVST | 2.44 | 1.66 | 67.9 | 18 |

Table 4-13: Descriptive Statistics of the Predicted to Measured Capacity Ratios by Proposed Design Methods for Piles in Cohesive Soils

| Condition | Method to Obtain $C_u$ | Proposed from Statistics | | | | Proposed from GA | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Mean | $\sigma_{SD}$ | COV (%) | n | Mean | $\sigma_{SD}$ | COV (%) | n |
| Tension Loads ($Qs_p/Qu$) | UCS | 0.89 | 0.25 | 27.9 | 14 | 0.90 | 0.31 | 34.5 | 14 |
| | Triaxial | 0.85 | 0.25 | 29.1 | 12 | 0.81 | 0.20 | 24.4 | 12 |
| | FVST | 1.07 | 0.68 | 63.6 | 7 | 1.44 | 0.72 | 50.1 | 7 |
| Compression Loads ($Qu_p/Qu$) | UCS | 0.69 | 0.15 | 22.2 | 16 | 0.73 | 0.19 | 25.3 | 16 |
| | Triaxial | 0.68 | 0.16 | 23.8 | 14 | 0.71 | 0.22 | 31.5 | 14 |
| | FVST | 0.73 | 0.47 | 64.3 | 11 | 1.01 | 0.46 | 45.7 | 11 |
| All (Capacity) ($Qu_p/Qu$) | UCS | 0.79 | 0.23 | 29.0 | 30 | 0.81 | 0.26 | 32.7 | 30 |
| | Triaxial | 0.76 | 0.22 | 29.4 | 26 | 0.76 | 0.22 | 28.8 | 26 |
| | FVST | 0.86 | 0.59 | 67.7 | 18 | 1.17 | 0.61 | 52.2 | 18 |

### 4.4.3 Indirect Methods for Noncohesive Soils

Piles were selected if 80 % of the embedment length was in a noncohesive soil, and the capacities were predicted for a combination of 23 tension and compression load tests. The design method suggested by Berezantzev *et al.* (1961) was examined for both the tip and side resistance. The bearing capacity factor ($N_q$) for the unit tip resistance has a relationship with $\emptyset'$ and the pile slenderness ratio ($L/D$). The unit tip resistance was then calculated with the effective stress ($\sigma'$) at the pile tip (Berezantzev *et al.*, 1961):

$$q_p = N_q\, \sigma' \leq 6000 \text{ kPa} \qquad \text{Eq. 4-9}$$

The unit tip resistance was limited as suggested by Poulos *et al.* (2001) and Bowles (1996). The unit side resistance was determined as below (Tomlinson and Woodward, 2008):

$$q_s = K_f K_o\, \sigma'\, tan\, \delta \qquad \text{Eq. 4-10}$$

where $K_f$ is the installation coefficient, $K_o$ is the lateral earth coefficient at rest, $\sigma'$ is the effective stress, and $\delta$ is the friction angle of the soil-pile interface. As recommended by Tomlinson and Woodward (2008), $\delta$ was 60 % of $\emptyset'$, and $K_f$ was 1.5 for solid pipe piles and 1.0 for H piles.

Both the side and tip resistance are dependent on $\emptyset'$. Since $\emptyset'$ was not measured, it was estimated with three empirical approaches in Table 4-14. For each pile, the ultimate capacity was calculated with each of the empirical methods. N-values were limited to 60 during the calculations due to the unreliability of extremely high N-values. Since stiff soils were encountered, judgement was applied to prevent unrealistically high friction angles from the correlations.

The capacity was calculated with both plugged and unplugged conditions, and the governing case was chosen as the predicted capacity. From the predictions, the fully plugged condition governed for piles subjected to tension loads, but the unplugged condition gave lower capacities for compression loads.

Table 4-14: Studied Empirical Methods for Noncohesive Soils

| Reference | Equation |
|---|---|
| Hatanaka & Uchida (1996) | $\emptyset' = \sqrt{20\,(N_{60})_1} + 20$ |
| Kulhawy & Mayne (1990) | $\emptyset' = \arctan\{(N_{60})_1/[12.2 + 20.3\,(\sigma'/100)\,]\}^{0.34}$ |
| Wolff (1989) | $\emptyset' = 27.1 + 0.3\,N_{60} - 0.00054\,{N_{60}}^2$ |

In Figure 4-19, the three empirical approaches offered similar predictions for the $\emptyset'$ and pile capacity. The COV of the $Qu_p/Qu$ ranged from 47 % to 49 %, but Wolff (1989) tends to provide

lower estimates. The average $Qu_p/Qu$ ratio was 1.16, 1.20, and 1.20 for Wolff (1989), Hatanaka and Uchida (1996), and Kulhawy and Mayne (1990), respectively. These indirect methods offered reasonably accurate predictions for noncohesive soils.

The results from the indirect approaches, which applied the method by Berezantzev *et al.* (1961), were compared to the proposed direct approaches from the GA. The unit side resistance was calculated with Equations 4-4 or 4-5 and was limited to 100 kPa. Since some piles were subjected to compressive loads, the tip resistance was calculated with Equation 3-3 from the statistical analysis. H piles were assumed to be fully plugged for these proposed equations.

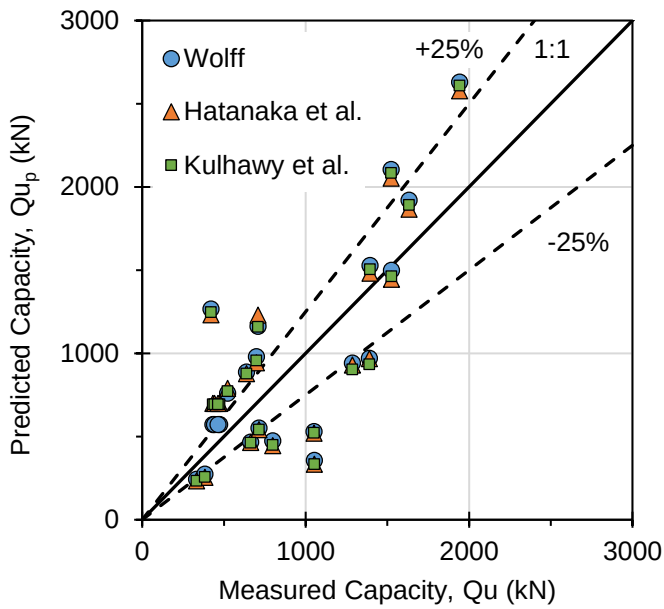



Figure 4-19: Comparison of Measured and Predicted Capacities by the Berezantzev *et al.* (1961) Method

In Figure 4-19, the method by Berezantzev *et al.* (1961) tended to over predict piles with higher capacities, but Figure 4-20 shows the proposed correlations had unbiased predictions for the same piles. For the proposed functions, the average $Qu_p/Qu$ ratio was 1.16. This result was similar to the existing methods, and overestimations mostly occurred with short H piles. However, the COV of the $Qu_p/Qu$ ratio was significantly reduced to 28 % from 47 % or 49 %. Details on the resulting $Qu_p/Qu$ ratios from the existing design methods are summarized in Table 4-15.

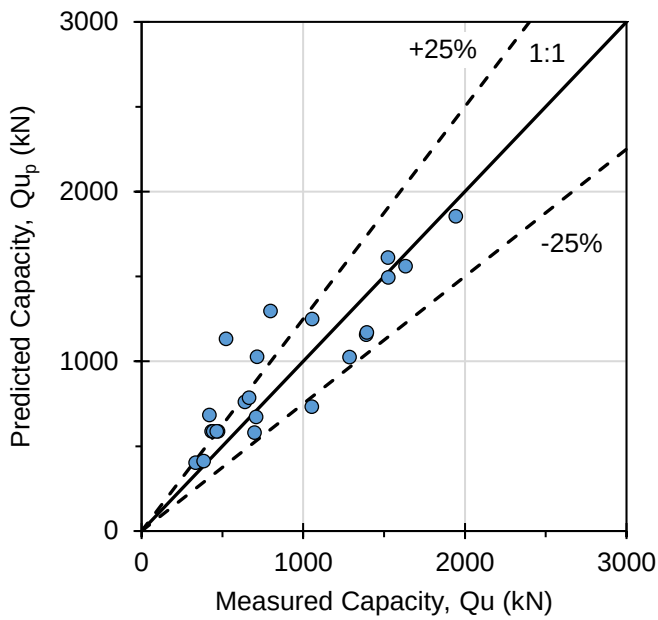



Figure 4-20: Comparison of Measured and Predicted Capacities by the Proposed SPT Correlations from the GA

Table 4-15: Descriptive Statistics of the Predicted to Measured Capacity Ratios by Berezantzev *et al.* (1961)

| Condition | Estimation Method for Friction Angle | Mean | $\sigma_{SD}$ | COV (%) | n |
|---|---|---|---|---|---|
| Tension Loads ($Qs_p/Qu$) | Wolff | 1.19 | 0.76 | 63.3 | 8 |
| | Hatanaka *et al.* | 1.16 | 0.74 | 63.6 | 8 |
| | Kulhawy *et al.* | 1.17 | 0.75 | 64.0 | 8 |
| Compression Loads ($Qu_p/Qu$) | Wolff | 1.14 | 0.40 | 34.9 | 15 |
| | Hatanaka *et al.* | 1.22 | 0.49 | 40.0 | 15 |
| | Kulhawy *et al.* | 1.21 | 0.48 | 39.5 | 15 |
| All (Capacity) ($Qu_p/Qu$) | Wolff | 1.16 | 0.55 | 47.4 | 23 |
| | Hatanaka *et al.* | 1.20 | 0.59 | 48.9 | 23 |
| | Kulhawy *et al.* | 1.20 | 0.59 | 49.0 | 23 |

## 4.5 SUMMARY AND DISCUSSIONS

A correlation study compared the results from soil testing techniques in cohesive soils. For the $C_u$, UCS measurements were 10 % higher than UU triaxial results. The *PI* by itself cannot provide decent relation between UU triaxial and FVST results. Unfortunately, FVST were rarely paired with other soil tests, and a reasonable correction factor was not established for FVST values. The suggested relation from Morris and William (1994) may be preferred over the one from Bjerrum (1972) since it is more conservative, but the correction factors were not intended for highly OC or silty soils.

The $C_u/N_{60}$ ratios likely have a saddle relationship with the plasticity. The ratios are sometimes high for very silty soils, but it decreases to a range of 2 to 5 for a *PI* between 20 to 30. Stroud (1974) obtained a similar trend. From this investigation, the highest *PI* was approximately 30, but other references, such as Sowers (1954), had some soils with a greater plasticity and indicated the ratios increase with a higher *PI*. However, other factors that were not investigated will influence of the $C_u/N_{60}$ ratios, such as the soil OCR or compressibility and testing errors. Especially for silty soils, the $C_u/N_{60}$ ratio may be inconsistent as a weak relation was found with *PI*.

The GA provided better results compared to existing design methods with SPT N-values. The existing design methods had erratic predictions, and they overestimated because they were mainly intended for soft or weak soils. In general, SPT measurements usually provided poor correlations for the side resistance, but the GA was efficient at obtaining correlations for the side resistance of piles with other variables. Since most of sites dominated in noncohesive soils, predictions for cohesive soils by the GA were usually inconsistent, and the GA rarely applied the soil type. Equation 4-4 assumes cohesive soils have lower side resistance, but other references, such as such as Shioi and Fukui (1982), received the opposite result. Due to the influence of the effective stress and soil type in Equations 4-4 and 4-5, these functions may be more appropriate for noncohesive soils or clays subjected to drained conditions. SPT N-values were not very influential according to the GA. For Equation 4-5, the magnitude of N-values mainly affects the side resistance towards the base of the pile, and a low percentage of functions at 42.6 % applied the N-values for H piles. Yet, almost all the functions by the GA contained the slenderness ratio. Since an inverse relationship was found with the variable, Equations 4-5 and 4-5 can indicate that the installation effects likely have a large impact on pile capacity. A higher side resistance towards the pile base could occur during pull-out because every pile had an over-sized base plate or reinforcement plate, but the over-sized based plate would contribute to the tip resistance during a compression test. Since most sites dominated in noncohesive soils, the frequency of the slenderness ratio could be related to the lateral earth pressure and confinement of the effective stress. In the end, it is difficult to specify the exact influence as additional soil testing and fully instrumented piles are required, but the GA attempts to extract the best correlation with the available data.

Design methods for cohesive soils considered a variety of variables but were developed empirically, semi-empirically, or by trial-and-error. All the design methods overestimated and had a lot of inconsistencies, especially with FVST measurements. The $\alpha$ Method by Karlsrud *et al.* (2005) over predicted with the lowest $Qu_{pred}/Qu$ ratio on average, but the $\beta$ Method by Patrizi and Burland (2001) had the lowest COV for the ratio. The GA offered better predictions for the side resistance and had the lowest COVs for the $Qu_{pred}/Qu$ ratio. 99.1 % of the functions from the GA contained the slenderness ratio, like Kolk and Van der Velde (1996). Both proposed methods from the statistical analysis and GA offer conservative estimates for the capacity, and this result is likely because the method by Meyerhof (1976) is conservative for the tip resistance. Karlsrud *et al*. (2005) indicated the *PI* impacts the side resistance, but the LI may have a stronger relationship, as shown in Figure 3-15. Although the Atterberg indices are obtained from remoulded soils, the LI may indicate the drainage conditions of the soil. For practical reasons, the side resistance can be sufficiently predicted with the $C_u$ alone, but the results may change with more pile load tests and varying soil conditions.

# 5 CONCLUSIONS AND FUTURE RESEARCH

## 5.1 SUMMARY

After conducting a statistical analysis with soil measurements and pile resistances, this investigation demonstrated the capability of a simple GA to improve predictions for the capacity of 43 piles. For heterogeneous soil conditions, the GA was tailored to find a function for the unit side resistance with SPT N-values or the undrained shear strength ($C_u$). Even though the sample size was small, the GA was given greater details on the soil measurements by dividing the piles into segments. The correlated functions were also refined with Pareto optimization, and their results were compared to several existing design methods. In all, predictions for the ultimate capacity were improved with the GA as existing methods generally over predicted and provided inconsistent results.

## 5.2 CONCLUSIONS AND RECOMMENDATIONS

Functions from the GA were selected from two different criteria: the best fitness and Pareto optimal. For practical reasons, the Pareto optimal functions are recommended for predicting the side resistance with corrected SPT measurements. Thus, the following function is proposed for capped pipe piles ($R^2 = 0.70$):

$$q_s = [2.8 \cdot N_{corr} \cdot D/(L - z) + 4.5]/S_T \leq 100 \text{ kPa} \qquad \text{Eq. 4-4}$$

The function below is recommended for fully plugged H piles ($R^2 = 0.81$):

$$q_s = \sigma'/[(1.9/N_{corr}) + (L - z)/D] \leq 100 \text{ kPa} \qquad \text{Eq. 4-5}$$

In these functions, $N_{corr}$ is the corrected SPT N-value; $L$ is the embedment length of a pile; $D$ is the diameter or width of a pile; $z$ is the depth to the centre of a pile segment or soil layer; $S_T$ is the soil type that is equal to 1 for noncohesive soils and 2 for cohesive soils; and $\sigma'$ is the effective stress. It is suggested, like Meyerhof (1976), to limit the unit side resistance to 100 kPa. From the studied piles, the measured unit side resistance did not surpass this value, and the corrected N-values were limited to 60 due to the unreliability of extremely high values.

Several variables influence the pile capacity. Equation 4-4 is directly proportionate to the SPT N-value and indicates a higher unit side resistance with stiffer soils. Equation 4-5 is not heavily influenced by the SPT N-value, which may be due to the plugging conditions or installation effects. The side resistance increases towards the pile base as both equations are inversely related to the slenderness ratio. Several reasons can explain the common appearance of this variable. For instance, the horizontal confinement of the soil, which was not experimentally measured, likely has a high influence on the side resistance, but it is difficult to specify without results from fully

instrumented piles or further soil measurements. Since noncohesive soils were commonly found, Equations 4-4 and 4-5 may be more appropriate for noncohesive soils or clays under drained conditions.

All the design methods tend to overestimate the capacity and side resistance of the piles in cohesive soils within Ontario. Thus, the correlation from the statistical analysis is recommended for the side resistance ($R^2 = 0.98$):

$$q_s = 0.17\ C_u \leq 80 \text{ kPa} \qquad \text{Eq. 3-2}$$

In the equation, $C_u$ is measured by UU triaxial, and the studied piles did not have a side resistance greater than 80 kPa. The equation is simple and received a similar fitness to the functions from the GA. Yet, for cohesive soils, the side resistance is likely influenced by the installation effects, lateral earth pressure, and soil content, which can be related and indicated by the slenderness ratio and soil plasticity.

If the soil strength is measured with SPT N-values, the pile and soil type had a small influence on the tip resistance. For cohesive and noncohesive soils, a single relationship can predict the unit tip resistance for both pipe piles and plugged H piles ($R^2 = 0.47$):

$$q_p = 304\ N_{corr} \leq 13000 \text{ kPa} \qquad \text{Eq. 3-3}$$

Piles in clayey and sandy soils did not have a unit tip resistance greater than 4000 kPa or 13000 kPa, respectively. The corrected N-values were less than 20 and 40 respectively for cohesive and noncohesive soils. The soil content indicates the strength of the soil structure and highly influences the tip resistance. For silty soils, the relation can be approximated with the plasticity index for cohesive soils. Yet, from the studied piles, the $C_u$ cannot reliably predict the tip resistance by itself, and the design method by Meyerhof (1976) will likely offer conservative estimates.

While comparing testing methods for the soil strength, SPT and UU triaxial measurements were proportionate for cohesive soils, and they are influenced by the water content and plasticity index. $C_u/N_{60}$ ratios can range from 3 to 7, but these ratios will likely have a lot of variation for silty soils.

An attempt was made to find a FVST correction factor, but several other variables likely influence the measurements, such as the silt content, soil stress history, sample disturbance, and shearing rate and dimensions of the field vane. The correction factor by Morris and William (1994) may be preferred over Bjerrum (1972). Due to the variability, applying a simple reduction factor of 0.6 to FVST results may be appropriate, but the inconsistency indicates a need for a new correction factor.

## 5.3 LIMITATIONS AND FUTURE WORK

The studied pile load tests offer valuable data that can be further investigated. The number of studied piles can be increased by adding results from piles subject to only compression load tests. For the test results, the GA would be modified to correlate with both extension and compression load tests. Since the GA feasibly predicted the ultimate capacity, it could be implemented to improve predictions for the service limit state. Also, a reliability analysis can provide the accuracy and variability of the proposed and existing design methods.

Genetic programming is a customizable approach to develop correlations, and a simple GA can be improved in multiple ways. A GA does not indicate the significance of a variable compared to a statistical hypothesis test, and categorical variables, such as the soil type, should be handled differently. For example, instead of applying a variable in a function, the GA can search for different constants for each category. Especially since the soil conditions change along a pile, the manipulation of repetitive observations needs to be controlled to prevent the creation of asymptotes. Interval awareness (Dick, 2017) can ensure the development of logical functions as it follows the numerical domain of the genes. The GA could also benefit from prior knowledge on the influence of variables. For example, if the effective stress increases the confinement of drained soils at greater depths, the GA should not assume a low effective stress results in a high side resistance. The current approach for constant refinement can be changed to reduce the computational effort. Optimization methods, such as the Nelder-Mead simplex method (Nelder & Mead, 1965), may be practical to refine constants while regressing for particular target variables, such as the tip resistance. However, the method may be inefficient with numerous variables and requires modification for repetitive measurements. The MEP technique with Pareto optimization is effective at finding practical expressions, but a function is typically mixed with unused genes in the chromosome. The efficiency of data processing and function searching can be increased with a hybrid system. For instance, an artificial neural network can learn to adjust the settings of the MEP-based GA and maintain diversity in the search space. Altogether, Banzhaf *et al*. (1998) suggested to improve a GA by focusing on 3 aspects: computational speed, evolvability, and search power.

The GA provided practical and accurate functions since it was not limited to the average conditions. Although the findings greatly rely on the extent of the site investigations and number of pile load tests, the performance of existing design methods demonstrated a need for locally-developed design methods in Ontario. The purpose of the GA was to obtain the most information from the available soil conditions, and it was more efficient at handling several variables compared to traditional regression. The GA outperformed the existing design methods in the end, and it likely could offer more accurate results with advanced soil testing, such as the cone penetration test, or results from fully instrumented piles. The Canadian government plans to make large investments in its future infrastructure, and machine learning techniques can help address uncertainty in geotechnical engineering and contribute towards efficient designs within Ontario.

# APPENDIX

# APPENDIX A: SUMMARY OF PILE LOAD TESTS

Table A-1: Descriptive Statistics of Variables Related to the Sides of the Piles

| Pile Type | Dominating Soil Type Around Side | | Ultimate Capacity (kN) | Percent Pile Length in Sand (%) | Percent Pile Length in Silt (%) | Percent Pile Length in Clay (%) |
|---|---|---|---|---|---|---|
| Pipe Piles | Cohesive | n | 12 | 12 | 12 | 12 |
| | | Mean | 564.500 | 5.121 | 48.334 | 46.543 |
| | | Std. Deviation | 298.3910 | 8.1441 | 47.2076 | 49.6045 |
| | Mixed | n | 2 | 2 | 2 | 2 |
| | | Mean | 1353.500 | 45.801 | 38.465 | 15.734 |
| | | Std. Deviation | 259.5082 | 15.9986 | .3896 | 15.6091 |
| | Noncohesive | n | 7 | 8 | 8 | 8 |
| | | Mean | 635.000 | 92.683 | 7.317 | .000 |
| | | Std. Deviation | 306.2744 | 11.9779 | 11.9779 | .0000 |
| | Total | n | 21 | 22 | 22 | 22 |
| | | Mean | 663.143 | 40.660 | 32.521 | 26.817 |
| | | Std. Deviation | 366.4064 | 43.0121 | 40.0428 | 42.5264 |
| H Piles | Cohesive | n | 8 | 8 | 8 | 8 |
| | | Mean | 779.000 | 5.884 | 39.100 | 55.016 |
| | | Std. Deviation | 448.7242 | 11.9352 | 49.8322 | 47.4328 |
| | Mixed | n | 3 | 3 | 3 | 3 |
| | | Mean | 1886.667 | 45.024 | 43.687 | 11.289 |
| | | Std. Deviation | 757.6717 | 12.2541 | 10.0118 | 13.3630 |
| | Noncohesive | n | 8 | 10 | 10 | 10 |
| | | Mean | 1197.500 | 87.641 | 12.359 | .000 |
| | | Std. Deviation | 501.0429 | 17.8178 | 17.8178 | .0000 |
| | Total | n | 19 | 21 | 21 | 21 |
| | | Mean | 1130.105 | 50.408 | 27.021 | 22.571 |
| | | Std. Deviation | 626.0372 | 41.2092 | 35.0648 | 38.7324 |
| Total | Cohesive | n | 20 | 20 | 20 | 20 |
| | | Mean | 650.300 | 5.426 | 44.640 | 49.932 |
| | | Std. Deviation | 370.6134 | 9.5409 | 47.1872 | 47.6613 |
| | Mixed | n | 5 | 5 | 5 | 5 |
| | | Mean | 1673.400 | 45.335 | 41.598 | 13.067 |
| | | Std. Deviation | 623.8183 | 11.8005 | 7.6380 | 12.4950 |
| | Noncohesive | n | 15 | 18 | 18 | 18 |
| | | Mean | 935.000 | 89.882 | 10.118 | .000 |
| | | Std. Deviation | 500.0987 | 15.2904 | 15.2904 | .0000 |
| | Total | n | 40 | 43 | 43 | 43 |
| | | Mean | 884.950 | 45.420 | 29.835 | 24.744 |
| | | Std. Deviation | 552.7267 | 41.9284 | 37.3490 | 40.2895 |

Table A-2: Descriptive Statistics of Variables Related to the Sides of the Piles

| Pile Type | Dominating Soil Type Around Side | | Length (m) | Setup Time (Days) | Mean Corrected SPT | Mean Wc (%) | Mean Effective Stress (kPa) |
|---|---|---|---|---|---|---|---|
| Pipe Piles | Cohesive | n | 12 | 12 | 9 | 12 | 12 |
| | | Mean | 14.078 | 17.75 | 11.975 | 21.778 | 100.684 |
| | | Std. Deviation | 7.4414 | 4.634 | 2.9940 | 5.9654 | 46.3669 |
| | Mixed | n | 2 | 2 | 2 | 2 | 2 |
| | | Mean | 20.045 | 26.50 | 8.583 | 22.693 | 118.848 |
| | | Std. Deviation | 7.5731 | 13.435 | 1.6067 | 6.3518 | 4.4905 |
| | Noncohesive | n | 8 | 8 | 8 | | 8 |
| | | Mean | 17.736 | 36.13 | 9.057 | | 90.495 |
| | | Std. Deviation | 2.2664 | 21.970 | 4.9003 | | 36.2283 |
| | Total | n | 22 | 22 | 19 | 14 | 22 |
| | | Mean | 15.950 | 25.23 | 10.389 | 21.909 | 98.630 |
| | | Std. Deviation | 6.1859 | 16.065 | 3.9842 | 5.7728 | 40.3874 |
| H Piles | Cohesive | n | 8 | 8 | 6 | 8 | 8 |
| | | Mean | 14.180 | 25.75 | 13.898 | 22.007 | 103.268 |
| | | Std. Deviation | 7.6661 | 12.349 | 5.7923 | 5.9514 | 50.1190 |
| | Mixed | n | 3 | 3 | 3 | 3 | 3 |
| | | Mean | 22.590 | 20.33 | 10.476 | 21.344 | 142.318 |
| | | Std. Deviation | 6.9202 | 12.858 | 3.4706 | 5.0595 | 40.7764 |
| | Noncohesive | n | 10 | 10 | 10 | 6 | 10 |
| | | Mean | 25.796 | 11.50 | 28.307 | 16.557 | 140.215 |
| | | Std. Deviation | 10.6691 | 5.874 | 12.5679 | 2.3077 | 46.7386 |
| | Total | n | 21 | 21 | 19 | 17 | 21 |
| | | Mean | 20.913 | 18.19 | 20.942 | 19.967 | 126.440 |
| | | Std. Deviation | 10.3469 | 11.461 | 12.4324 | 5.2111 | 48.7436 |
| Total | Cohesive | n | 20 | 20 | 15 | 20 | 20 |
| | | Mean | 14.119 | 20.95 | 12.744 | 21.870 | 101.718 |
| | | Std. Deviation | 7.3290 | 9.208 | 4.2492 | 5.8022 | 46.6026 |
| | Mixed | n | 5 | 5 | 5 | 5 | 5 |
| | | Mean | 21.572 | 22.80 | 9.719 | 21.884 | 132.930 |
| | | Std. Deviation | 6.3424 | 11.798 | 2.7826 | 4.8407 | 31.6490 |
| | Noncohesive | n | 18 | 18 | 18 | 6 | 18 |
| | | Mean | 22.214 | 22.44 | 19.752 | 16.557 | 118.117 |
| | | Std. Deviation | 8.9084 | 19.379 | 13.7982 | 2.3077 | 48.4070 |
| | Total | n | 43 | 43 | 38 | 31 | 43 |
| | | Mean | 18.374 | 21.79 | 15.665 | 20.844 | 112.212 |
| | | Std. Deviation | 8.7414 | 14.292 | 10.5597 | 5.4671 | 46.3122 |

Table A-3: Descriptive Statistics of Variables Related to the Pile Tips

| Pile Type | Soil Type at Tip | | Ultimate Capacity (kN) | Length (m) | Setup Time (Days) | Corrected SPT at Tip | Effective Stress at Tip (kPa) |
|---|---|---|---|---|---|---|---|
| Pipe Piles | Clay | n | 6 | 6 | 6 | 3 | 6 |
| | | Mean | 487.333 | 12.143 | 16.50 | 10.000 | 184.415 |
| | | Std. Deviation | 163.5783 | 7.0409 | 4.037 | 6.9687 | 128.3335 |
| | Silt | n | 8 | 8 | 8 | 8 | 8 |
| | | Mean | 722.125 | 17.425 | 20.38 | 16.060 | 214.418 |
| | | Std. Deviation | 403.6821 | 7.5725 | 8.017 | 6.4015 | 62.8364 |
| | Sand | n | 7 | 8 | 8 | 8 | 8 |
| | | Mean | 746.429 | 17.331 | 36.63 | 20.807 | 169.176 |
| | | Std. Deviation | 437.5553 | 2.2262 | 21.394 | 5.7089 | 65.4340 |
| | Total | n | 21 | 22 | 22 | 19 | 22 |
| | | Mean | 663.143 | 15.950 | 25.23 | 17.102 | 189.784 |
| | | Std. Deviation | 366.4064 | 6.1859 | 16.065 | 6.9998 | 84.0586 |
| H Piles | Clay | n | 6 | 6 | 6 | 4 | 6 |
| | | Mean | 861.667 | 15.858 | 23.17 | 30.750 | 193.860 |
| | | Std. Deviation | 498.6813 | 8.0652 | 13.467 | 30.2923 | 91.2241 |
| | Silt | n | 5 | 5 | 5 | 5 | 5 |
| | | Mean | 721.400 | 16.900 | 23.00 | 11.437 | 208.230 |
| | | Std. Deviation | 337.9657 | 8.4552 | 15.083 | 1.3436 | 47.8062 |
| | Sand | n | 8 | 10 | 10 | 10 | 10 |
| | | Mean | 1586.875 | 25.952 | 12.80 | 27.873 | 284.314 |
| | | Std. Deviation | 587.6883 | 10.7612 | 5.329 | 15.4880 | 85.6684 |
| | Total | n | 19 | 21 | 21 | 19 | 21 |
| | | Mean | 1130.105 | 20.913 | 18.19 | 24.153 | 240.355 |
| | | Std. Deviation | 626.0372 | 10.3469 | 11.461 | 18.3182 | 87.8230 |
| Total | Clay | n | 12 | 12 | 12 | 7 | 12 |
| | | Mean | 674.500 | 14.001 | 19.83 | 21.857 | 189.138 |
| | | Std. Deviation | 404.2480 | 7.4743 | 10.098 | 24.4544 | 106.2690 |
| | Silt | n | 13 | 13 | 13 | 13 | 13 |
| | | Mean | 721.846 | 17.223 | 21.38 | 14.282 | 212.038 |
| | | Std. Deviation | 364.8744 | 7.5730 | 10.728 | 5.4760 | 55.4515 |
| | Sand | n | 15 | 18 | 18 | 18 | 18 |
| | | Mean | 1194.667 | 22.121 | 23.39 | 24.733 | 233.142 |
| | | Std. Deviation | 665.6582 | 9.0982 | 18.759 | 12.3882 | 95.4686 |
| | Total | n | 40 | 43 | 43 | 38 | 43 |
| | | Mean | 884.950 | 18.374 | 21.79 | 20.628 | 214.481 |
| | | Std. Deviation | 552.7267 | 8.7414 | 14.292 | 14.1367 | 88.6564 |

# APPENDIX B: GENETIC ALGORITHM SOURCE CODE

```
function [resfit, rescon, reschrom, resadd1, resadd2,
resgen]=startga(Data,PL,Goal)

% Genetic algorithm with MEP encoding
% Coded in Matlab by
% Markus Jesswein (markus.jesswein[at]ryerson.ca)

%% Start of program %%
% Use the above function and required inputs to run the GA

% Input:
% "Data" is a 3D array (i,j,k), where "i" is for each pile test, "j"
is for
% each pile segment, and "k" is each variable to be regressed
% "PL" (i x j) is the side area for each pile test
% "Goal" (i x 1) is the target variable (measured side resistance)


%Output:
% "Resfit" contains the fitness results of the chromosomes from the
last
% generation [Best fitness (RMSE), index (location of best fitness in
the
% chromosome), MAE, R2]
% "Rescon" has constants used by the chromosomes
% "Reschrom" contains the activators of the chromosomes
% (operator/constant/variable)
% "Resadd1" and "resadd2" contains the links of the chromosomes
% "Resgen" summarizes the fitness results for every generation
[generation
% No., best fitness, index of best fitness, highest fitness, mean
fitness]

% Settings:

saveopt = 1; % 1 = Print and save; 2 = Save only (do not print)

set.popsize = 200; %Chromosome population size: must be an even number
set.codelength = 20; %Length of chromosomes
set.numgen = 10; %Number of generations
set.mutprob = 0.1; %Mutation probability
set.crossprob = 0.9; %Crossover probability
set.broodact = 1; % Brood: On = 1 / Off = 0
set.broodprob = 0.5; %Brood probability during crossover
set.broodpop = 4; %Number of additional trials for brood (should be
even an even number -- offspring are coupled)

set.varprob = 0.3; %Probability of variable
set.operprob = 0.4; %Probability of operation
set.consprob = 1 - set.varprob - set.operprob; %Probability of
constant (remaining probability)
```

```
set.conint = 2; %How many constant sets (1-3)?
set.numcons = 4; %How many constants to use for each set?
set.conbrood = 10; %How many trials to find best constant for
chromosome?
set.consmin = -1; %Minimum value for constants (set 1)
set.consmax = 1; %Maximum value for constants (set 1)
set.consmin2 = 1; %Minimum value for constants (set 2)
set.consmax2 = 10; %Maximum value for constants (set 2)
set.consmin3 = 0; %Minimum value for constants (set 3)
set.consmax3 = 10; %Maximum value for constants (set 3)

set.toursize = 2; %Tournament size (minimum of 2)
set.numop = 8; %Range of operators to apply (4 = up to '/'; 5 = P; 6 =
E; 7 = Ln; 8= Tanh)

set.numrep = size(Data,2); %Number of repetitive variables (ex number
of pile segments)
set.numvar = size(Data,3); %Number of variables

[resfit, rescon, reschrom, resadd1, resadd2,
resgen]=gengares(set,saveopt,Data,PL,Goal);

end
```

```
function [resfit, rescon, reschrom, resadd1, resadd2,
genres]=gengares(set,save,Data,PL,Goal)

%Prints out all the chromosomes from the final generation



numrep = set.numrep;
numvar = set.numvar;
numsamp = size(Data,1);

% One population of chromosomes
% Newly created individuals will replace poorly performing individuals

% Begin creatation of chromosomes
% -1='+';-2='-';-3='*';-4='/';-5='Power';-6='Exp';-7='Ln';-8='Tanh'
% 1 to n = Var1 to VarN
% n+1 to n+n = Const1 to ConstN

codelen=set.codelength;
add1=zeros(set.popsize,codelen); % Link 1
add2=zeros(set.popsize,codelen); % Link 2
chrom=zeros(set.popsize,codelen); % Activators

const=zeros(set.popsize,set.numcons*set.conint);
fitres=zeros(set.popsize,4);
evalm=zeros(numsamp,codelen);

genres=zeros(set.numgen,5);


for i = 1:1:set.popsize %for every chrom

%Randomly create constants
%Set 1
const(i,1:set.numcons)= set.consmin+(set.consmax -
set.consmin).*rand(1,set.numcons);
if set.conint>=2 %Set 2
const(i,(set.numcons+1):2*set.numcons)= set.consmin2+(set.consmax2 -
set.consmin2).*rand(1,set.numcons);
end
if set.conint>=3 %Set 3
const(i,(2*set.numcons+1):3*set.numcons)= set.consmin3+(set.consmax3 -
set.consmin3).*rand(1,set.numcons);
end

%First position in chromosome should be a variable or constant
sum1=set.varprob+set.consprob;
P=sum1.*rand();

if P<=set.varprob
   chrom(i,1)= randi([1 numvar]);
```

```
else
   chrom(i,1)= randi([numvar+1 numvar+set.numcons*set.conint]);
end

add1(i,1)=1;
add2(i,1)=1;

%For rest of chromosome
for j=2:1:codelen

    P=rand();
    if P>sum1
        chrom(i,j)= (-1).*randi([1 set.numop]); %operator
    elseif P<=set.varprob
        chrom(i,j)= randi([1 numvar]); %variable
    else
        chrom(i,j)= randi([numvar+1 numvar+set.numcons*set.conint]);
%constant
    end

    add1(i,j)=randi([1, j-1]);
    add2(i,j)=randi([1, j-1]);
end

end


for i=1:1:set.popsize

%Evaluate Fitness--One chromosome at a time

[fitres(i,:),evalm,chrom(i,:)]=fitness(chrom(i,:),add1(i,:),add2(i,:),
const(i,:), codelen, numvar, numsamp,numrep,evalm,Data,PL,Goal);

%Refine Constants
[const(i,:),fitres(i,:),chrom(i,:)]=refconst(chrom(i,:),add1(i,:),add2
(i,:),const(i,:),set,numvar,fitres(i,:),evalm,Data,PL,Goal);

end

%Sort by ascending fitness
[fitres,chrom,add1,add2,const]=sortfit(fitres,chrom,add1,add2,const,co
delen);

genres(1,:)=[1 fitres(1,1) fitres(1,2) max(fitres(:,1))
mean(fitres(:,1))];

fprintf('Generation %d: best fitness = %f at index = %d. Max fitness =
%f\n',1,fitres(1,1),fitres(1,2),genres(1,4));


for g=2:1:set.numgen %For every generation
```

```
    for k=1:2:set.popsize %Two chromosomes at a time in the population

    %Choose parents -- Binary tournament

    r1=tournamentsel(chrom,fitres,set.popsize,set.toursize);
    r2=tournamentsel(chrom,fitres,set.popsize,set.toursize);

    %Crossover

    P=rand();
    if P<=set.crossprob

        P=rand();
        if P<=set.broodprob

            for i=1:1:(set.broodpop/2)

                if i==1


[offsp1,offsp2,offcon1,offcon2,offadd1A,offadd1B,offadd2A,offadd2B]=ra
nuniform_cross(chrom(r1,:),chrom(r2,:),const(r1,:),const(r2,:),add1(r1
,:),add2(r1,:),add1(r2,:),add2(r2,:),codelen);


[off_fit1,evalm,offsp1]=fitness(offsp1,offadd1A,offadd1B, offcon1,
codelen, numvar, numsamp,numrep,evalm,Data,PL,Goal);

[off_fit2,evalm,offsp2]=fitness(offsp2,offadd2A,offadd2B, offcon2,
codelen, numvar, numsamp,numrep,evalm,Data,PL,Goal);

                else


[offspt1,offspt2,offcont1,offcont2,offaddt1A,offaddt1B,offaddt2A,offad
dt2B]=ranuniform_cross(chrom(r1,:),chrom(r2,:),const(r1,:),const(r2,:)
,add1(r1,:),add2(r1,:),add1(r2,:),add2(r2,:),codelen);


[off_fitp1,evalm,offspt1]=fitness(offspt1,offaddt1A,offaddt1B,
offcont1, codelen, numvar, numsamp,numrep,evalm,Data,PL,Goal);

[off_fitp2,evalm,offspt2]=fitness(offspt2,offaddt2A,offaddt2B,
offcont2, codelen, numvar, numsamp,numrep,evalm,Data,PL,Goal);

                    if off_fitp1(1,1)<off_fit1(1,1)

                        off_fit1=off_fitp1;
                        offsp1=offspt1;
                        offcon1=offcont1;
                        offadd1A=offaddt1A;
```

```
                        offadd1B=offaddt1B;

                    end
                    if off_fitp2(1,1)<off_fit2(1,1)

                        off_fit2=off_fitp2;
                        offsp2=offspt2;
                        offcon2=offcont2;
                        offadd2A=offaddt2A;
                        offadd2B=offaddt2B;

                    end

                end

            end

        else

[offsp1,offsp2,offcon1,offcon2,offadd1A,offadd1B,offadd2A,offadd2B]=ra
nuniform_cross(chrom(r1,:),chrom(r2,:),const(r1,:),const(r2,:),add1(r1
,:),add2(r1,:),add1(r2,:),add2(r2,:),codelen);

        end

    else


[offsp1,offcon1,off_fit1,offadd1A,offadd1B]=copyindivid(set,chrom(r1,:
),const(r1,:),fitres(r1,:),add1(r1,:),add2(r1,:));

[offsp2,offcon2,off_fit2,offadd2A,offadd2B]=copyindivid(set,chrom(r2,:
),const(r2,:),fitres(r2,:),add1(r2,:),add2(r2,:));

    end

    %Mutation -- Mutate and calculate fitness


[offsp1,offadd1A,offadd1B,offcon1]=mutation(offsp1,offadd1A,offadd1B,o
ffcon1,set,numvar);
    [off_fit1,evalm1,offsp1]=fitness(offsp1,offadd1A,offadd1B,
offcon1, codelen, numvar, numsamp,numrep,evalm,Data,PL,Goal);


[offsp2,offadd2A,offadd2B,offcon2]=mutation(offsp2,offadd2A,offadd2B,o
ffcon2,set,numvar);
    [off_fit2,evalm2,offsp2]=fitness(offsp2,offadd2A,offadd2B,
offcon2, codelen, numvar, numsamp,numrep,evalm,Data,PL,Goal);

    %Refine Constants
```

```
[offcon1,off_fit1,offsp1]=refconst(offsp1,offadd1A,offadd1B,offcon1,se
t,numvar,off_fit1,evalm1,Data,PL,Goal);

[offcon2,off_fit2,offsp2]=refconst(offsp2,offadd2A,offadd2B,offcon2,se
t,numvar,off_fit2,evalm2,Data,PL,Goal);

    %Replace worst chromosome
    if off_fit1(1,1)<fitres(set.popsize,1)


[chrom(set.popsize,:),const(set.popsize,:),fitres(set.popsize,:),add1(
set.popsize,:),add2(set.popsize,:)]=copyindivid(set,offsp1,offcon1,off
_fit1,offadd1A,offadd1B);


[fitres,chrom,add1,add2,const]=sortfit(fitres,chrom,add1,add2,const,co
delen);

    end

    if off_fit2(1,1)<fitres(set.popsize,1)


[chrom(set.popsize,:),const(set.popsize,:),fitres(set.popsize,:),add1(
set.popsize,:),add2(set.popsize,:)]=copyindivid(set,offsp2,offcon2,off
_fit2,offadd2A,offadd2B);


[fitres,chrom,add1,add2,const]=sortfit(fitres,chrom,add1,add2,const,co
delen);

    end

    end

    genres(g,:)=[g fitres(1,1) fitres(1,2) max(fitres(:,1))
mean(fitres(:,1))];

    fprintf('Generation %d: best fitness = %f at index = %d. Max
fitness = %f\n',g,fitres(1,1),fitres(1,2),genres(g,4));


end

%Print results from final generation

[resfit,rescon,reschrom,resadd1,resadd2]=printres(fitres,chrom,const,a
dd1,add2,set,numvar,save);


end
```

```
function [fitresl,evalmt,chrom]=fitness(chrom,add1,add2, const,
codelen, numvar, numsamp,numrep,evalm,Data,PL,Goal)


%%% Using RMSE for fitness
%Includes R2 and MAE

evalmt=zeros(numsamp,codelen);

n=1;
while n<(numrep+1) %evaulate for every pile segment

if n==1
[evalm,chrom,c]=comp_evalm(n,chrom,add1,add2,const,codelen,numvar,nums
amp,evalm,Data);

evalmt=evalmt+evalm.*PL(:,n);

n=n+1;
else


[evalm,chrom,c]=comp_evalm(n,chrom,add1,add2,const,codelen,numvar,nums
amp,evalm,Data);

    if c==0

    evalmt=evalmt+evalm.*PL(:,n);
    n=n+1;
    else
    evalmt=zeros(numsamp,codelen);
    n=1;
    end
end

end


for i=1:1:codelen %read chrom from beginning (left to right)
%evaluate every part of the chromosome

   sumerr=0; %overall error for a chromosome
   tempm=0;
   sump=0;
   summ=0;
   sumpm=0;
   sump2=0;
   summ2=0;


   for j=1:1:numsamp %for every sample in dataset
```

```
        sumerr=sumerr+abs(evalmt(j,i)-Goal(j,1)); %absolute difference
between predicted and actual

        tempm=tempm+((evalmt(j,i)-Goal(j,1))^2);

        sump=sump+evalmt(j,i);
        sump2=sump2+((evalmt(j,i))^2);
        summ=summ+Goal(j,1);
        summ2=summ2+((Goal(j,1))^2);
        sumpm=sumpm+(evalmt(j,i)*Goal(j,1));


    end

    mse = sqrt((tempm/numsamp)); %root mean square error



    if i==1


       fitness=mse;

       index=1;
       mae = sumerr/numsamp; %mean absolute error

       R2 = ((numsamp*(sumpm)-(sump)*(summ))/(((numsamp*sump2-
sump^2)*(numsamp*summ2-summ^2))^(0.5)))^2; %R2

       if R2>1||R2<=0

          R2=0;
       end
    end

    if fitness>mse

        fitness=mse;

        index=i; %position in chromosome that gives best result (do not
need full chromosome)
        mae = sumerr/numsamp; %mean absolute error

        R2 = ((numsamp*(sumpm)-(sump)*(summ))/(((numsamp*sump2-
sump^2)*(numsamp*summ2-summ^2))^(0.5)))^2; %R2

        if R2>1||R2<=0

          R2=0;
        end
    end
```

```
end
fitresl=[fitness,index,mae,R2];
end
```

```
function
[evalm,chrom,c]=comp_evalm(n,chrom,add1,add2,const,codelen,numvar,nums
amp,evalm,Data)


%Keep intermediate results in case of errors
%In case of error (ex. dividing by Zero), the chromosome is mutated
evalm=zeros(numsamp,codelen);
numconst = size(const,2);

c=0; %track if change is needed (0 for No/1 for Yes)

for i=1:1:codelen %Read chromosome from beginning

switch chrom(1,i)

    case -1 %'+'

        for k=1:1:numsamp

            evalm(k,i)=evalm(k,add1(1,i)) + evalm(k,add2(1,i));

        end

    case -2 %'-'

        for k=1:1:numsamp

            evalm(k,i)=evalm(k,add1(1,i)) - evalm(k,add2(1,i)); %%%%

        end

    case -3 %'*'

        for k=1:1:numsamp

            evalm(k,i)=evalm(k,add1(1,i)) .* evalm(k,add2(1,i)); %%%%

        end

    case -4 %'/'

        for k=1:1:numsamp

            if abs(evalm(k,add2(1,i)))< 1.0e-06  %A small value
(prevent division by 0)

                chrom(1,i)= randi([1 numvar]); %change to division to
variable
                c=1; %track if change is needed
                for k=1:1:numsamp
```

```
                    evalm(k,i)= Data(k,n,chrom(1,i));

                end
                break

            else %normal operation

                for k=1:1:numsamp

                    evalm(k,i)=evalm(k,add1(1,i)) ./
evalm(k,add2(1,i)); %%%%

                end


            end

        end

    case -5 %Power (^) (abs() of base)

        for k=1:1:numsamp

            if abs(evalm(k,add2(1,i)))> 20  %A power greater than 20

                chrom(1,i)= -8; %change to tanh()

                c=1;
                for k=1:1:numsamp

                    evalm(k,i)=tanh(evalm(k,add1(1,i)));


                end
                break

            else %normal operation

                for k=1:1:numsamp

                    evalm(k,i)= abs(evalm(k,add1(1,i))) .^
evalm(k,add2(1,i)); %%%%

                end


            end

        end

    case -6 %Exponent (e)
```

```
        for k=1:1:numsamp

            if abs(evalm(k,add1(1,i)))> 32  %An exp() greater than 32

                chrom(1,i)= -1; %change to exp() to addition
                c=1;
                for k=1:1:numsamp

                    evalm(k,i)=evalm(k,add1(1,i)) +
evalm(k,add2(1,i));

                end
                break

            else %normal operation

                for k=1:1:numsamp

                    evalm(k,i)=exp(evalm(k,add1(1,i))); %%%%

                end


            end

        end

    case -7 %Ln(abs())

        for k=1:1:numsamp

            if abs(evalm(k,add1(1,i)))< 1.0e-06  %A small value
(prevent log of 0)

                chrom(1,i)= randi([numvar+1 numvar+numconst]); %change
to Ln to constant
                c=1; %track if change is needed
                for k=1:1:numsamp

                    evalm(k,i)= const(1,chrom(1,i)-numvar);

                end
                break

            else %normal operation

                for k=1:1:numsamp

                    evalm(k,i)=log(abs(evalm(k,add1(1,i)))); %%%%

                end
```

```
            end

        end
    case -8 %tanh
        for k=1:1:numsamp
            evalm(k,i)=tanh(evalm(k,add1(1,i)));
        end
    otherwise
        for k=1:1:numsamp
            if chrom(1,i)<= numvar
                evalm(k,i)= Data(k,n,chrom(1,i));
            else
                evalm(k,i)= const(1,chrom(1,i)-numvar);
            end
        end

end
end
end
```

```
function
[con1,fitfin,ofsp]=refconst(chrom,add1,add2,const,set,numvar,fitres,ev
alm,Data,PL,Goal)

numrep=set.numrep;
numsamp=size(Data,1);
codelen = size(chrom,2);
count=0;
actconst=zeros(1,set.numcons*set.conint);
fit=zeros(set.conbrood-1,4);

    for i=1:1:codelen %Count number of constants used
        if chrom(1,i)>numvar

        count=count+1;
        if actconst(1,chrom(1,i)-numvar)==0
           actconst(1,chrom(1,i)-numvar)=1;
        end

        end

    end

if count<1
   con1=const;
   fitfin=fitres;
   ofsp=chrom;

else

%create random constants
tempconst=zeros(set.conbrood-1,set.numcons*set.conint);

for i=1:1:(set.numcons*set.conint)

    if actconst(1,i)==1

        if set.conint>1
           if i<=set.numcons
               tempconst(:,i)=(set.consmin+(set.consmax -
set.consmin).*rand(set.conbrood-1,1));
           elseif i>set.numcons&&(i<=2*set.numcons)
               tempconst(:,i)=(set.consmin2+(set.consmax2 -
set.consmin2).*rand(set.conbrood-1,1));
           else
               tempconst(:,i)=(set.consmin3+(set.consmax3 -
set.consmin3).*rand(set.conbrood-1,1));
           end

           else
                tempconst(:,i)=(set.consmin+(set.consmax -
set.consmin).*rand(set.conbrood-1,1));
```

```
        end

    else
        tempconst(:,i)=const(1,i);
    end

end

fitfin=fitres;
con1=const;
ofsp=chrom;

for i=1:1:(set.conbrood-1)

    [fit(i,:),evalm,chrom]=fitness(chrom,add1,add2, tempconst(i,:),
codelen, numvar, numsamp,numrep,evalm,Data,PL,Goal);

    if fit(i,1)<fitfin(1,1)
    fitfin = fit(i,:);
    con1=tempconst(i,:);
    ofsp=chrom;

    end

end

end


end
```

```
function
[fit,chrom,a1,a2,ct]=sortfit(fitres,chrom,add1,add2,const,codelen)

max1=size(fitres,2);
max2=size(const,2);

temp=[fitres chrom add1 add2 const];

temp=sortrows(temp,1);

fit=temp(:,1:max1);
chrom=temp(:,(max1+1):(max1+codelen));
a1=temp(:,(max1+1+codelen):(max1+2*codelen));
a2=temp(:,(max1+1+2*codelen):(max1+3*codelen));
ct=temp(:,(max1+1+3*codelen):(max1+3*codelen+max2));

end
```

```
function [offsp1,offcon1,offit,offadd1,offadd2]=copyindivid(set, par1,
parcon1, fitres,add1,add2)

offsp1=par1;
offcon1=parcon1;
offit=fitres;
offadd1=add1;
offadd2=add2;

end
```

```
function p=tournamentsel(chrom,fitres,popsize,toursize)

p=randi([1 popsize],1,1);
for i=2:1:toursize

    t=randi([1 popsize],1,1);

    if fitres(p,1)<fitres(t,1)
        p=p;
    else
        p=t;
    end

end


end
```

```
function
[offsp1,offsp2,offcon1,offcon2,offadd1A,offadd1B,offadd2A,offadd2B]=ra
nuniform_cross(par1,par2,const1,const2,paradd1A,paradd1B,paradd2A,para
dd2B,codelen)

csize=size(const1);
offcon1=zeros(1,csize(2));
offcon2=zeros(1,csize(2));
offsp1=zeros(1,codelen);
offsp2=zeros(1,codelen);
offadd1A=zeros(1,codelen);
offadd1B=zeros(1,codelen);
offadd2A=zeros(1,codelen);
offadd2B=zeros(1,codelen);

P1=0.5;

for i=1:1:codelen
P=rand();
if P<=P1

    offsp1(1,i)=par1(1,i);
    offsp2(1,i)=par2(1,i);
    offadd1A(1,i)=paradd1A(1,i);
    offadd1B(1,i)=paradd1B(1,i);
    offadd2A(1,i)=paradd2A(1,i);
    offadd2B(1,i)=paradd2B(1,i);

else
    offsp1(1,i)=par2(1,i);
    offsp2(1,i)=par1(1,i);
    offadd1A(1,i)=paradd1B(1,i);
    offadd1B(1,i)=paradd1A(1,i);
    offadd2A(1,i)=paradd2B(1,i);
    offadd2B(1,i)=paradd2A(1,i);

end
end

%for constants

if csize(2)>0

    for i=1:1:csize(2)
    P=rand();
    if P<=P1

        offcon1(1,i)=const1(1,i);
        offcon2(1,i)=const2(1,i);

    else
```

```
        offcon1(1,i)=const2(1,i);
        offcon2(1,i)=const1(1,i);

    end
    end

end


end
```

```
function
[offsp1,add1,add2,const]=mutation(offsp1,add1,add2,const,set,numvar)

codelen=set.codelength;
%mutate based on probability for every chromosome entry
%start with first entry, which is either a constant or variable

P = rand();
if P<=set.mutprob
    sum = set.consprob+set.varprob;
    P=sum*rand();

    if P<=set.varprob
    offsp1(1,1)= randi([1 numvar]);
    else
    offsp1(1,1)= randi([numvar+1 numvar+set.numcons*set.conint]);

    end

end
%check other genes

for j=2:1:codelen

    P=rand();
    if P<=set.mutprob %mutate, but select what to mutate

    P=rand();

    if P<=set.operprob
        offsp1(1,j)= (-1).*randi([1 set.numop]); %operator
    elseif P<=set.operprob+set.varprob
        offsp1(1,j)= randi([1 numvar]); %variable
    else
        offsp1(1,j)= randi([numvar+1 numvar+set.numcons*set.conint]);
%constant
    end

    P=rand();
    if P<=set.mutprob
    add1(1,j)=randi([1, j-1]); %location 1
    end
    P=rand();
    if P<=set.mutprob
    add2(1,j)=randi([1, j-1]); %location 2
    end


    end

end
```

```
%constants
csize=size(const);
for i=1:1:csize(2)

    P=rand();
    if P<=set.mutprob

        if set.conint>1
           if i<=set.numcons
               const(1,i)=set.consmin+(set.consmax -
set.consmin).*rand(1,1);
           elseif i>set.numcons&&(i<=2*set.numcons)
               const(1,i)=set.consmin2+(set.consmax2 -
set.consmin2).*rand(1,1);
           else
               const(1,i)=set.consmin3+(set.consmax3 -
set.consmin3).*rand(1,1);
           end

        else
           const(1,i)=set.consmin+(set.consmax -
set.consmin).*rand(1,1);
        end

    end

end


end
```

```
function
[resfit,rescon,reschrom,resadd1,resadd2]=printres(fitres,chrom,const,a
dd1,add2,set,numvar,save)

%save and print results

codelen=set.codelength;
operstr='+-*/PELT'; %operations
P=Power,E=Exponent,L=Ln(abs()),T=Tanh()

if save==1

fprintf('The chromosome is: \n');

for i=1:1:(set.numcons*set.conint)

    fprintf('constant[%d] = %f\n', i, const(1,i));

end
for i=1:1:codelen

    if chrom(1,i)<0

        if abs(chrom(1,i))<=5
        fprintf('%d: %c %d %d\n', i, operstr(1,abs(chrom(1,i))),
add1(1,i),add2(1,i));
        else % expon, etc
        fprintf('%d: %c %d\n', i, operstr(1,abs(chrom(1,i))),
add1(1,i));

        end
    else

        if chrom(1,i)<=numvar

            fprintf('%d: input[%d]\n', i, chrom(1,i));

        else

            fprintf('%d: constant[%d]\n', i, chrom(1,i)-numvar);

        end

    end

end

fprintf('Best index = %d\n',fitres(1,2));
fprintf('Fitness = %f\n', fitres(1,1));

end
```

```
    resfit = fitres;
    rescon = const;
    reschrom = chrom;
    resadd1=add1;
    resadd2=add2;


end
```

# BIBLIOGRAPHY


Aas, G., Lacasse, S., Lunne, I, & Hoeg, K. (1986). Use of in situ tests for foundation design on clay. *In Situ 1986 Proceedings*, ASCE, 1-30.

Abu-Farakh, M. Y., Haque, Md. N., & Chen, Q. (2016). *Field instrumentation and testing to study set-up phenomenon of piles driven into Louisiana clayey soils* (Final Report 562). Baton Rouge, LA: Louisiana Transportation Research Center (LTRC).

Alkroosh, I. & Nikraz, H. (2011). Correlation of pile axial capacity and CPT data using gene expression programming. *Geotechnical and Geological Engineering, 29*(5), 725-748.

American Society for Testing and Materials (ASTM). (2007). Standard test methods for deep foundations under static axial compressive load (D1143-07). West Conshohocken, PA: ASTM.

Aoki, N. & Velloso, D. A. (1975). An approximate method to estimate the bearing capacity of piles. *Proceedings of the Fifth Pan-American Conference on Soil Mechanics and Foundation Engineering*, Buenos Aires, Argentina, 367-376.

Banzhaf, W., Nordin, P., Keller, R.E., & Francone, F.D. (1998). *Genetic programming: an introduction*. San Francisco, CA: Morgan Kaufmann Publishers Inc.

Barnett, P. J. (1992). Chapter 21: quaternary geology of Ontario. In P. C. Thurston, H. R. Williams, R. H. Sutcliffe, & G. M. Scott (Eds.), *Geology of Ontario: Ontario Geological Survey* (Special Vol. 4, Part 2, pp. 1011-1088). Sudbury, ON: Ontario Ministry of Northern Development and Mines.

Berezantzev, V.G., Khristoforov, V.S., & Golubkov, V. N. (1961). Load bearing capacity and deformation of piled foundations. *Proceedings of the 5th International Conference on Soil Mechanics and Foundation Engineering, 2*, 11-15.

Bjerrum, L. (1972). Embankments on soft ground. *Proceedings of the Specialty Conference on Earth and Earth-Supported Structures*, ASCE, *2*, 1-54.

Bjerrum, L. (1973). Problems of soil mechanics and construction on soft clays and structurally unstable soils (collapsible, expansive and others). *Proceedings of the 8th International Conference on Soil Mechanics and Foundation Engineering, 3*, 111-159.

Bowles, J. E. (1996). *Foundation analysis and design* (5th ed.). New York, NY: The McGraw-Hill Companies Inc.

Brameier, M. & Banzhaf, W. (2007). *Linear genetic programming*. New York, NY: Springer.

Briaud, J., & Tucker, L. M., (1988). Measured and predicted axial response of 98 piles. *Journal of Geotechnical Engineering*, ASCE, *114*(9), 984-1001.

Briaud, J-L., Tucker, L., Lytton, R. L., & Coyle, H. M. (1983). *Behavior of piles and pile groups in cohesionless soils* (FHWA/RD - 83-38). Washington, DC: Federal Highway Administration.

Brown, R. P. (2001). *Predicting the ultimate axial resistance of single driven piles* (Doctoral dissertation). Department of Civil Engineering, University of Texas, Austin, TX.

Burland, J. B. (1973). Shaft friction of piles in clay – a simple fundamental approach. *Ground Engineering*, *6*(3), 30-42.

Canadian Geotechnical Society (CGS). (2006). *Canadian Foundation Engineering Manual* (4th ed.). Richmond, BC: CGS.

Chandler, R. J. (1988). The in-situ measurement of the undrained shear strength of clays using the field vane. In A. F. Richards (Eds.), *Vane Shear Strength Testing in Soils: Field and Laboratory Studies* (ASTM STP 1014). Philadelphia, PA: American Society for Testing and Materials (ASTM), 13-44.

Coyle, H.M., & Castello, R.R. (1981). New design correlations for piles in sand. *Journal of the Geotechnical Engineering Division*, ASCE, *107*(7), 965-986.

Croon, M. A. & van Veldhoven, M. J. P. M. (2007). Predicting group-level outcome variables from variables measured at the individual level: a latent variable multilevel model. *Psychological Methods*, *12*(1), 45-57.

Das, S. K. & Basudhar, P. K. (2006). Undrained lateral load capacity of piles in clay using artificial neural network. *Computers and Geotechnics*, *33*, 454-459.

Decourt, L. (1982). Predictions of bearing capacity based exclusively on N values of the SPT. *Proceedings of the 2nd European Symposium on Penetration Testing*, Amsterdam, *1*, 29-34.

Dennis, N. D., & Olson, R.E. (1983). Axial capacity of steel pipe piles in clay. *Proceedings of the Conference on Geotechnical Practice in Offshore Engineering*, ASCE, 370-388.

Dick, G. (2017). Interval arithmetic and interval-aware operators for genetic programming. *Computing Research Repository* (*CoRR*). Retrieved from https://arxiv.org/pdf/1704.04998.pdf

Fellenius, B. H. (1980). The analysis of results from routine pile load tests. *Ground engineering, 13*(6), 19-31.

Gandomi, A. H. & Alavi, A. H. (2012). A new multi-gene genetic programming approach to non-linear system modeling. Part II: geotechnical and earthquake engineering problems. *Neural Computing and Applications, 21*(1), 189-201.

Hara, A., Ohta, T., Niwa, M., Tanaka, S., & Banno, T. (1974). Shear Modulus and Shear Strength of Cohesive Soils. *Soils and Foundations, 14*(3), 1-12.

Hatanaka, M. & Uchida, A. (1996). Empirical correlation between penetration resistance and internal friction angle of sandy soils. *Soils and Foundations, 36*(4), 1-9.

Ismail, A. & Jeng, D-S. (2011). Modelling load-settlement behaviour of piles using higher order neural network (HON-PILE model). *Engineering Applications of Artificial Intelligence, 24*, 813-821.

Jamiolkowski, M., Ladd, C. C., Germaine, J. T., & Lancellotta, R. (1985). New developments in field and laboratory testing soils. *Proceedings of the 11th International Conference on Soil Mechanics and Foundation Engineering, 1*, 57-154.

Karlsrud, K. (2012). *Prediction of load-displacement behaviour and capacity of axially loaded piles in clay based on analyses and interpretation of pile load test results* (Doctoral dissertation). Norwegian University of Science and Technology, Trondheim, Norway.

Karlsrud, K., Clausen, C. J. F., & Aas, P. M. (2005). Bearing capacity of driven piles in clay, the NGI approach. *Proceedings of the International Symposium on Frontiers in Offshore Geotechnics*, 775-782.

Kolk, H. J., & Van der Velde, E. (1996). A reliable method to determine friction capacity of pile driven into clays. *Proceedings of the 28th Offshore Technology Conference*, 337-346.

Kulhawy, F. H. & Jackson, C. S. (1989). Some observations on undrained side resistance of drilled shafts. In F. H. Kulhawy (Ed.), *Foundation Engineering: Current Principles and Practices* (Vol. 2, pp. 1011-1025). New York, NY: ASCE.

Kulhawy, F. H. & Mayne, P. W. (1990). *Manual on estimating soil properties for foundation design: report EL-6800*. Palo Alto, California: Electric Power Research Institute.

Levasseur, S., Malécot, Y., Boulon, M., & Flavigny, E. (2008). Soil parameter identification using a genetic algorithm. *International Journal for Numerical and Analytical Methods in Geomechanics, 32*, 189-213.

Liao, S. S. C. & Whitman, R. V. (1986). Overburden correction factors for the SPT in sand. *Journal of Geotechnical Engineering*, ASCE, *112*(3), 373-377.

Lüdtke, O., Marsh, H. W., Robitzch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: a new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, *13*(3), 203-229.

Luke, D. A. (2004). *Quantitative applications in the social sciences: multilevel modeling*. Thousand Oaks, CA: SAGE Publications Ltd.

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, *11*(2), 431-441.

Martin, R. E., Seli, J. J., Powell, G. W. & Bertoulin, M. (1987). Concrete pile design in Tidewater, Virginia. *Journal of Geotechnical Engineering*, ASCE, *113*(6), 568-585.

Mathworks. (2017). *Matlab* [Computer software]. Natick, MA: Matworks.

Meyerhof, G. G. (1956). Penetration tests and bearing capacity of cohesionless soils. *Journal of the Soil Mechanics and Foundations Division*, ASCE, *82*(1), 1-19.

Meyerhof, G. G. (1976). Bearing capacity and settlement of pile foundations. *Journal of the Geotechnical Engineering Division*, ASCE, *102*(3), 197-228.

Ministry of Transportation of Ontario (MTO). (1993). *Pile load and extraction tests: 1954 – 1992* (Rev. 1993). Toronto, ON: MTO.

Morris, P. H. & Williams, D. J. (1994). Effective stress vane shear strength correction factor correlations. *Canadian Geotechnical Journal*, *31*, 335-342.

Negnevitsky, M. (2005). *Artificial intelligence: a guide to intelligent systems* (2nd ed.). Harlow, England: Pearson Education Ltd.

Nelder, J. A. & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, *7*(4), 308-313.

Oltean, M. (2016). Multi Expression Programming [source code]. Retrieved from https://github.com/mepx/mep-basic-src

Oltean, M. & Dumitrescu, D. (2002). *Multi Expression Programming*. Cluj-Napoca, Romania: Babes-Bolyai University.

O'Neil, M. W. & Reese, L. C. (1999). *Drilled shafts: construction procedures and design methods* (FHWA-IF-99-025). Washington, DC: Federal Highway Administration (FHWA).

Ontario. (2018). Ministry of Infrastructure. Retrieved from https://www.ontario.ca/page/ministry-infrastructure

Patrizi, P., & Burland, J. B. (2001). Developments in the design of driven piles in clay in terms of effective stresses. *Rivista Italiana Di Geotechnica*, *3*, 35-49.

Peck, R. B., Hanson, W. E., & Thornburn, T. H. (1974). *Foundation Engineering*. New York, NY: Wiley and Sons Inc.

Poulos, H.G., Carter, J.P., & Small, J.C. (2001). Foundations and retaining structures -research and practice. *Proceedings of the 15th International Conference on Soil Mechanics and Geotechnical Engineering*, Istanbul, *4*, 2527-2606.

Semple, R. M., & Rigden, W. J. (1986). Shaft capacity of driven pipe piles in clay. *Ground Engineering*, 11-19.

Shahin, M. A. (2014). Load-settlement modeling of axially loaded drilled shafts using CPT-based recurrent neural networks. *International Journal of Geomechanics*, ASCE, *14*(6), 06014012-1-7.

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

Shioi, Y. & Fukui, J. (1982). Application of N-value to design foundations in Japan. *Proceedings of the 2nd European Symposium on Penetration Testing*, Amsterdam, *1*, 159-164.

Sivrikaya, O., & Toğrol, E. (2006). Determination of undrained shear strength of fine-grained soils by means of SPT and its applications in Turkey. *Engineering Geology*, *86*(1), 52-69.

Smits, G. F. & Kotanchek, M. (2005). Pareto-front exploitation in symbolic regression. In O'Reilly *et al*. (Eds.), *Genetic programming theory and practice II*. Boston, MA: Springer.

Stermac, A. G., Shelby, K. G., & Devata, M. (1969). Behaviour of various types of piles in stiff clay. *Proceedings of the 7th International Conference on Soil Mechanics and Foundation Engineering*, *2*, 239-245.

Stroud, M. A. (1974). The standard penetration test in insensitive clays and soft rocks, *Proceedings of the European Symposium on Penetration Testing*, Stockholm, *2*(2), 367-375.

Sowers, G. F. (1954). Modern Procedures for Underground Investigations. *Proceedings of the American Society of Civil Engineering*, *80*(435), 1-11.

Tackett, W. A. (1994). *Recombination, selection, and the genetic construction of computer programs* (Doctoral dissertation). Department of Electrical Engineering Systems, University of Southern California, Los Angeles, CA.

Tomlinson, M. J. (1957). The adhesion of piles driven in clay soils. *Proceedings of the 4th International Conference of Soil Mechanics, 2*, 66-71.

Tomlinson, M. J. & Woodward, J. (2008). *Pile design and construction practice* (5th ed.). New York, NY: Taylor and Francis.

Thorburn, S. & MacVicar, S. L. (1971). Pile load tests to failure in the Clyde alluvium. *Behaviour of piles*, London, UK: Institution of Civil Engineers, *1*(7), 53-54.

Vesic, A. S. (1967). *A study of bearing capacity of deep foundations, final report*. Atlanta, GA: Georgia Institute of Technology.

Vesic, A. S. (1977). *Design of pile foundations, synthesis of highway practice 42*. Washington, DC: Transportation Research Board.

Vijayvergiya, V. N., & Focht, J. A. Jr. (1972). A new way to predict the capacity of piles in clay. *Proceedings of the 4th Offshore Technology Conference, 2*, 865-874.

Wolff, T. F. (1989). Pile capacity prediction using parameter functions. *Predicted and Observed Axial Behavior of Piles: Results of a Pile Prediction Symposium* (ASCE Geotechnical Special Publication No. 23), 96-106.

Zhong, J. Hu, X., Gu, M., & Zhang, J. (2005). Comparison of performance between different selection strategies on simple genetic algorithms. *Proceedings of the International Conference on Computational Intelligence of Modelling, Control and Automation, and International Conference on Intelligent Agents, Web Technologies and Internet Commerce*, Vienna, Austria, 1115-1121.