Statistical assessment of mobilized shear strength of stiff-fissured clays

Yuderka Trinidad Gonzalez

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Recommended Citation
https://lib.dr.iastate.edu/etd/18013

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Statistical assessment of mobilized shear strength of stiff-fissured clays

by

Yuderka Trinidad-González

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Civil Engineering (Geotechnical Engineering)

Program of Study Committee:
Vernon Schaefer, Co-major Professor
Derrick Rollins, Co-major Professor
Huaiqing Wu
R. Christopher Williams
Jeramy Ashlock
Cassandra Rutherford

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University
Ames, Iowa
2020

Copyright © Yuderka Trinidad Gonzalez, 2020. All rights reserved.
I dedicate this dissertation to God and my family.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>xi</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>xii</td>
</tr>
<tr>
<td>CHAPTER 1. GENERAL INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Scope of the Work</td>
<td>4</td>
</tr>
<tr>
<td>1.2 Research Objectives</td>
<td>5</td>
</tr>
<tr>
<td>1.3 Study Outline</td>
<td>5</td>
</tr>
<tr>
<td>References</td>
<td>6</td>
</tr>
<tr>
<td>CHAPTER 2. STATISTICAL ASSESSMENT OF FACTOR OF SAFETY FOR UNREINFORCED SLOPES</td>
<td>8</td>
</tr>
<tr>
<td>Abstract</td>
<td>8</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>9</td>
</tr>
<tr>
<td>2.2 Applied Methods</td>
<td>11</td>
</tr>
<tr>
<td>2.2.1 Slope Stability Analyses</td>
<td>11</td>
</tr>
<tr>
<td>2.2.2 Statistical Methods (DoE)</td>
<td>15</td>
</tr>
<tr>
<td>2.3 Results</td>
<td>22</td>
</tr>
<tr>
<td>2.3.1 Slope Stability Results</td>
<td>22</td>
</tr>
<tr>
<td>2.3.2 Significant Factors</td>
<td>23</td>
</tr>
<tr>
<td>2.3.3 Means FS Comparisons</td>
<td>27</td>
</tr>
<tr>
<td>2.3.4 Diagnostic Testing Results (Field validation)</td>
<td>28</td>
</tr>
<tr>
<td>2.4 Conclusions and Recommendations</td>
<td>31</td>
</tr>
<tr>
<td>References</td>
<td>34</td>
</tr>
<tr>
<td>CHAPTER 3. STATISTICAL ASSESSMENT OF F FOR PILE-REINFORCED SLOPES</td>
<td>37</td>
</tr>
<tr>
<td>Abstract</td>
<td>37</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>38</td>
</tr>
<tr>
<td>3.2 Applied Methods</td>
<td>42</td>
</tr>
<tr>
<td>3.2.1 Slope Stability Analyses</td>
<td>42</td>
</tr>
<tr>
<td>3.2.2 Statistical Methods</td>
<td>47</td>
</tr>
<tr>
<td>3.3 Results</td>
<td>53</td>
</tr>
<tr>
<td>3.3.1 Slope Stability Results</td>
<td>53</td>
</tr>
<tr>
<td>3.3.2 Significant Factors</td>
<td>53</td>
</tr>
<tr>
<td>3.3.3 FS Comparisons</td>
<td>60</td>
</tr>
<tr>
<td>3.4 Conclusions and Recommendations</td>
<td>62</td>
</tr>
<tr>
<td>References</td>
<td>65</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Fig. 1. (a) Stress-displacement curves for overconsolidated and fully softened (normally consolidated) test specimens, and (b) overconsolidated peak, fully softened peak, and residual strength envelopes after VandenBerge et al. (2013). .................. 3

Fig. 2. 1. A sketch of the sections and views of the generic slopes with the input properties...... 12

Fig. 2. 2. Mesh convergence study for (a) the smallest (h = 6, β = 60) and (b) the largest (h = 10, β = 20) section types for the 2DFE analyses. .................................................. 13

Fig. 2. 3. Mesh convergence study for (a) the smallest (h = 6, β = 60) and (b) the largest (h = 10, β = 20) section types for the 3DFE analyses. .................................................. 15

Fig. 2. 4. Probabilities tree diagram apply to the slope stability context......................... 21

Fig. 2. 5. Sample of results location of critical surface among methods for Run 4............... 23

Fig. 2. 6. Standardized effects (α = 0.05), mixed soils (c' and φ'). The red line indicates the t-statistic value to classify effects as statistically significant. .............................................. 25

Fig. 2. 7. Standardized effects (α = 0.05), frictional soils (φ'). The red line indicates the t-statistic value to classify effects as statistically significant. .............................................. 26

Fig. 2. 8. Standardized effects (α = 0.05), undrained (φ=0). The red line indicates the t-statistic value to classify effects as statistically significant. .............................................. 27

Fig. 2. 9. Graphical representation of FP and FN (a) stable slopes (b) failed slopes (Data from Sah et al. (1994)). .................................................................................. 29

Fig. 2. 10. Summary field condition vs. analytical methods from diagnostic testing........... 30

Fig. 3. 1. Graphical representation of geometry and inputs for pile-reinforced slopes for (a) 2DLE, (b) 2DFE, (c) 3DLE and (d) 3DFE................................................................. 42

Fig. 3. 2. Results from the mesh convergence study for (a) the smallest (h = 6, β = 60) and (b) the largest (h = 10, β = 20) section types for the 2DFE analyses. ......................... 44

Fig. 3. 3. Results from mesh convergence study for (a) the smallest (h = 6, β = 60) and (b) the largest (h = 10, β = 20) section types for the 3DFE analyses. ......................... 46

Fig. 3. 4. Example of the results compared in this study -- Run 4 from Table 3.2................ 53
Fig. 3. 5. Pareto charts of standardized effects at \( \alpha = 0.05 \) (the red line is a threshold to classify effects as significant or not). .............................................................. 55

Fig. 3. 6. Pareto chart of standardized effects from central composite analysis, \( c' - \phi' \) soil group, at \( \alpha = 0.05 \) (the red line is a threshold to classify effects as significant or not) .......................................................... 57

Fig. 3. 7. Response surface plots of the pile factors' effects in \( \Delta FS \) (a) purely cohesive group, (b) purely frictional group generated MATLAB (2019) .................................................. 58

Fig. 3. 8. Response surface plots of the pile factors' effects in \( \Delta FS \), \( c' - \phi' \) analysis generated by MATLAB (2019) ........................................................................................................ 59

Fig. 4. 1. Graphical representation of inputs for LE analyses ........................................................................................................ 72

Fig. 4. 2. Computer simulated FS distribution ......................................................................................................................... 73

Fig. 4. 3. Schematic illustration of ANN structure using JMP pro 14 (2019). ................................................................. 74

Fig. 4. 4. Observed vs. predicted plot for the best ANN model ................................................................................................. 77

Fig. 4. 5. Measured vs. predicted FS for all the models in this study (testing) ............................................................................... 80

Fig. 5. 1. Distribution of the FSS measured in the laboratory (studied dataset) ................................................................. 92

Fig. 5. 2. Schematic illustration of ANN structure using JMP-PRO 14 (JMP pro 14, 2019), where the first layer represents input variables such as LL, PI, and so on, the second layer shows the activation neurons of the hidden layer, and *FSS represents the response, secant FSS friction angle .................................................................................. 96

Fig. 5. 3. Prediction profiler, the influence of input factors in FSS ......................................................................................... 101

Fig. 5. 4. Shear/Normal function for BLC = Brown London Clay at Kingsbury Station for slope stability analysis Case 1. .................................................................................................................... 107

Fig. 5. 5. Comparison of slip plane presented by James (1970), (dashed-red line), and critical surface from slope stability analyses using various FSS correlations .................................................................................. 109

Fig. 5. 6. Shear/Normal function for Bentonitic Clay at Lesueur Landslide for slope stability analysis Case 2. ...................................................................................................................... 110

Fig. 5. 7. Comparison of slip plane presented by Abdel-Ghaffar (1990) citing Thomson (1971) (dashed red line), and critical surface from slope stability analyses using various FSS correlations .................................................................................. 112
Fig. 6.1. The power coefficient $b=0.86$ from Fig. 6.1 (c) with a small intercept (coefficient $c=1.31$ found by extrapolating the line in Fig. 6.1 (d) to the origin, for case history 1 (Watford By-Pass). ................................................................. 130

Fig. 6.2. Section of a generic slope with inputs ................................................................. 132

Fig. 6.3. Mesh convergence study results ........................................................................ 132

Fig. 6.4. Example of results from slope stability analyses for SRF=1 ............................... 133

Fig. 6.5. Summary of the procedure followed to develop the dataset for the proposed MRANN ................................................................................................................................. 134

Fig. 6.6. Schematic diagram of MRANN architecture (6–15–3) used ................................. 136

Fig. 6.7. Flow diagram of the procedure follow when developing the MRANN model ........ 139

Fig. 6.8. Superposition of critical surfaces between the field condition and the applied analytical approaches .............................................................................................................. 142
LIST OF TABLES

Table 2. 1. Soil and geometry properties considered factors for unreinforced slopes .................. 11
Table 2. 2. RCBD for unreinforced soil slopes configuration .................................................. 18
Table 2. 3. Soil properties from the literature review originally from Sah et al. (1994) .......... 20
Table 2. 4. Summary average coefficient of determination for all Models ................................ 23
Table 2. 5. Analysis of Variance Table RCBD unreinforced ..................................................... 27
Table 2. 6. Differences report pairwise comparison mean FS ................................................. 28
Table 3. 1. Values of the factors under study called “Input factors.” ......................................... 43
Table 3. 2. Randomized design table for DSD generated by MINITAB® .................................... 50
Table 3. 3. Randomized complete block design configuration ............................................... 51
Table 3. 4. Average R2 for the three soil groups ..................................................................... 54
Table 3. 5. Summary FS for each run for each analytical method from DSD ......................... 61
Table 3. 6. Analysis of Variance Table for the RCBD for all analyses ..................................... 61
Table 3. 7. Comparison means FS analytical methods by Tukey’s method for all analyses ...... 62
Table 4. 1. Soil and geometry properties for generating the multilevel full factorial design ...... 72
Table 4. 2. Summary of Statistics for performance evaluation ................................................. 77
Table 4. 3. Assessment of variance importance (input’s effect over the response) ................. 78
Table 4. 4. Range of inputs from the compared prediction models updated from Kostić et al. (2016) ........................................................................................................... 79
Table 4. 5. Data for comparison proposed model “TG model” with existing models and LE results ...................................................................................................................... 79
Table 4. 6. Summary of Statistics for performance evaluation, testing sets ............................. 81
Table 4. 7. Case histories used for the field accuracy assessment ........................................... 82
Table 4. 8. Confusion matrix for accuracy assessment of prediction models ........................... 83
Table 5. 1. Summary statistics of the range in input properties of the dataset used in this study. ................................................................. 91
Table 5. 2. Correlation matrix (Pearson correlation coefficients) for inputs and output. .......... 98
Table 5. 3. Summary of Statistics for performance evaluation (averaged five folds). .............. 99
Table 5. 4. Variable importance: Independent uniform inputs complete model raw data ANN. ............................................................................... 100
Table 5. 5. Performance statistics of the reduced final model ................................................. 104
Table 5. 6. Maximum and minimum constraints for the use of prediction tools based on the referenced tests. .............................................................................................................. 104
Table 5. 7. Comparison of factors of safety from back analysis and pre-failure analysis (design). .............................................................................................................................. 107
Table 5. 8. Comparison of factors of safety from back analysis and pre-failure analysis (design). .............................................................................................................................. 111
Table 6. 1. Some empirical correlations between FSS (output), index properties, and normal stress from previous studies................................................................. 124
Table 6. 2. Summary of case histories reported as first-time slides in stiff-fissured clays....... 125
Table 6. 3. Summary statistics of the range in input properties of the dataset used in this study with a sample size of 68 for each case................................................................. 127
Table 6. 4. Correlation matrix for the original dataset............................................................. 128
Table 6. 5. Performance results for the MRANN model.......................................................... 140
Table 6. 6. Variable importance ranking by total effects.......................................................... 141
Table 6. 7. Curve fitting coefficients for slope stability analysis testing case......................... 141
ACKNOWLEDGMENTS

First, I would like to give my special gratitude to my major professor, Dr. Vernon Schaefer, for his guidance, encouragement, and patience. I will always be grateful for the wonderful mentor I got. I also want to thank Dr. Derik Rollins for all his encouragement, commitment to helping me improve each of the publications presented herein. In addition, thanks to Dr. R. Chris Williams, for his guidance since the beginning of my program from my Master’s time. Additionally, I would like to thank Dr. Huaiqing Wu, Dr. Jeramy Ashlock, and Dr. Cassandra Rutherford for being part of my committee members and also for the enthusiasm and passion they always had when teaching.

I want to express my infinite gratitude to my mom Nancy González, the one that has supported me through thick and thin, and that has never stopped believing in me and my crazy endeavors. I also want to thank all the special people that supported and encouraged me to do my best during my studies, among these are; Paul (the sweetest and most encouraging of all), Henry, Kanika, Sinan, Conglin, Kathy, Keila, Parnian, Derya, Sam, Brody, Minas, and Diba. Last but not least, I would like to express my deep gratitude and love to my siblings, Nancy, Sandra, Anderson, Amel, and my father, Amado, who always cheered me up and cared about everything in my life.
ABSTRACT

Determination of the soil parameters for conducting stability analysis of first-time slides in stiff-fissured clays is still a significant challenge in the geotechnical field. Currently, instability analyses for these types of soils are mainly conducted by applying the fully softened shear strength (FSS) concept. However, this concept is only a practical approximation not applicable to all types of failures in first-time slides in stiff-fissured clays. This study presents the compilation of several advanced statistical assessments to evaluate the factor of safety (FS) of slopes to further focus on the specifics of slope instabilities in stiff-fissured clays. For this purpose, the integration of statistical design of experiments with analytical methods (limit equilibrium and finite element analyses) is performed to evaluate the factors and interactions affecting the FS of slopes for different soil configurations.

This evaluation allows for understanding of the controlling factors on FS for general soils not prone to softening. In addition, the use of statistical assessments is a novel application for the study of pile-reinforced slopes to determine the effects of reinforcement related factors on FS for optimal design purposes. Bayes' theorem is applied as diagnostic testing to determine the closeness of analytical tools to the field condition. This assessment delineates the accuracy of the current analytical methods and demonstrates the shortcomings of back-calculated parameters.

Regarding clays prone to softening, the prediction power of the existing correlations based on FSS concept is evaluated, generating a regression-based neural network (ANN) to simulate the laboratory determined FSS for the study and quantification of the relationships among the variables. Lastly, an original multiple response artificial neural network (MRANN) approach is undertaken to develop a prediction model that estimates three responses simultaneously. These three responses correspond to curve fitting parameters for a non-linear
failure envelope to characterize the mobilized shear strength of stiff-fissured clays. Several comparisons between the current methodology, the field condition, and the proposed modeling approach are provided to demonstrate the validity of the presented technique. The outcomes of this study provide three prediction tools that facilitate the application of research findings in the geotechnical practice without the need for special software for stages of preliminary design.

Other benefits from this study include contributions to research needs regarding the use and application of the FSS concept, contributions to optimize the selection of slope remediation techniques, contributions to the geotechnical profession by quantifying the difference between analytical methods, and contributions to design optimization for pile-reinforced slopes. Finally, the major contribution to advance the understanding of the mobilized shear strength of stiff-fissured clays resides in providing a general approach that could help overcome the applicability constraints of the FSS concept.
CHAPTER 1. GENERAL INTRODUCTION

Determination of the soil parameters for conducting stability analysis of stiff-fissured clays has been a major challenge in the geotechnical field for several years. Terzaghi (1936) was the first to point out the special considerations needed when performing slope stability analysis in stiff-fissured clays. Terzaghi separated clay soils into three categories regarding slope stability analyses; (1) soft, intact, joint-free, (2) stiff, intact, joint-free, and (3) stiff, fissured clays.

For first time slides, in either soft or stiff, joint-free clays, the mobilized shear falls very close to the peak strength from conventional laboratory test results (Duncan and Dunlop 1968). For the first time slides in stiff-fissured clays, however, the results of studies from several authors show that conventional methods of analysis do not lead to satisfactory results (Henkel and Skempton 1955; Henkel 1957; Skempton 1964; Skempton and LaRochelle 1965; Bjerrum 1967; Duncan and Dunlop 1968; Skempton 1977; Stark and Eid 1993; Mesri and Shahien 2002; Castellanos 2013). The later studies showed that the stress envelops for both residual and fully softened are stress-dependent, and curvature must be accounted for when performing the analyses.

Since Skempton's elucidation of the issues (1964, 1970), it has become well established that for overconsolidated clays, clays subjected to either progressive failure and/or softening mechanisms, the limiting strength falls in between the peak and the residual value for first-time slides. Skempton (1964) stated that the shear strength parameters determined by conventional tests do not essentially have a relationship with the values operative in the clay at the time of failure.
The factors controlling the limiting strength of overconsolidated clays were summarized by Skempton (1970) as follows:

1. The condition of the clay (intact vs. fissured)
   - Stiff-fissured clays have different limiting strength values than stiff-intact clays.

2. Sliding occurrence
   - First-time slides in stiff-fissured clays should consider strength values below peak values.
   - Pre-existing shear surfaces in slope indicate that residual values should be used for the stability analyses.

The studies performed in stiff-fissured London clay showed that the clay around slopes in stiff-fissured clays became weaker over time. The mechanism of softening acting upon this type of soils can be classified into two main softened groups: softening around cut slopes (first studied by Terzaghi (1936), and Skempton (1964)), and softening of compacted fills due to desiccation. The softening mechanisms occurring in natural soils may be due to: excavations, desiccation, and shrinkage, post-peak shear dilation, degradation by leaching of salts, drying, and slacking, drying and wetting, freezing-thawing, weathering, loss of initial pore water pressure, and pore water pressure increase. These mechanisms are generally simultaneous (more than one at the same time, not necessarily all).

To address the issue of mobilized shear in first-time slides in stiff-fissured clays, Skempton (1970), based on the critical state concept presented by Schofield and Wroth (1968), defined the fully softened shear strength (FSS). See Figure 1.1 for a graphical representation of the FSS concept.
Schofield and Wroth (1968) considered that the peak strength of a normally consolidated remolded clay occurs just before the critical state is reached. The critical state concept indicates that soil and granular materials, if continuously sheared until flowing as a frictional fluid, will reach a well-defined state in which shear distortion occurs without any further change in mean effective stresses. The conceptual models that represent the mechanical behavior of saturated remolded soils are based on the critical state concept. Hence, the FSS is a practical estimation, primarily based on laboratory tests conducted on London Clay, of the critical state as the drained peak shear strength of clay in its remolded normally consolidated state. The two governing equations of the critical state concept are given as:

\[ q = M p \]  
\[ \Gamma = \nu + \lambda \ln p \]

where \( M \) is a frictional constant, \( \Gamma \) and \( \lambda \) are soil material properties, \( q \) is the deviator stress needed to keep the soil flowing, \( p \) is the effective pressure, \( \nu \) is the specific volume occupied by unit volume of flowing particles.
Currently, slope stability analyses of first-time slides in stiff-fissured clays are mainly done by applying the FSS concept. However, some authors such as Potts et al. 1990; Stark and Eid 1997; Mesri and Shahien 2002; Skempton and Vaughan 2009 have found that residual strength can also play a part in the first time slides in clay fills and cuttings. Hence, the mobilized shear of first-time slides in stiff-fissured clays could be less than the FSS. Additional research is needed: to determine when is FSS equivalent to the mobilized shear, what limitations exist on FSS applications to design, and to develop a general framework for designing slopes in stiff-fissured clays not limited to FSS applicability.

1.1 Scope of the Work

This study encompasses evaluations of the general mechanism of slope stability to further focus on the specifics of slope instabilities in stiff-fissured clays. The integration of statistical design of experiments with numerical methods (limit equilibrium and finite element analyses) is applied to evaluate factors and interactions involved in slope stability mechanism for different soil configurations. This evaluation allows the understanding of the controlling factors in slope stability analysis for general soils not prone to softening. An evaluation to determine the closeness of analytical tools to the field condition is performed by diagnostic testing using Bayes’ theorem. In addition, a study of how reinforcing a slope changes the controlling factors, and the variability of the factor of safety is also performed. To evaluate the prediction power of the current correlations based on FSS concept, the basics of a regression-based neural network (ANN) are implemented. The ANN model is used to simulate the mechanism of laboratory determined FSS to study and quantify the hidden relationships among the variables. The applicability of the FSS concept is investigated by case history comparisons.
Finally, a multiple response artificial neural network (MRANN) approach is undertaken to develop a prediction model that estimates three responses simultaneously. These three responses correspond to curve fitting parameters for a non-linear general failure envelope.

1.2 Research Objectives

The main objectives of this research are (1) to statistically evaluate and determine the relationship between shear strength measured from laboratory tests and the mobilized shear in first-time slopes of stiff-fissured clays applying FSS concept, (2) to propose a general approach for slope design applicable to all types of stiff-fissured clays applying MRANN. The outcomes of this study will optimize the design and analysis of first-time slopes in stiff-fissured clays by providing an increased degree of reliability from the statistical analyses. Other benefits from this study include providing conclusions to numerous research needs regarding the use and application of the FSS concept, general slope stability matters, and pile-reinforced slopes.

1.3 Study Outline

This dissertation consists of seven chapters: a general introduction, five technical articles, and general conclusions. Chapter 1 provides a general introduction with background information related to this study. Chapter 2 studies the factors involved in the general slope stability mechanism by applying statistical design of experiments and Bayes theorem. This chapter demonstrates the importance of interactions when analyzing slopes and the pitfalls of using a threshold FS of 1 as boundary failure/stability criterion. Chapter 3 studies the factors involved in pile-reinforced slope stability analysis. This chapter demonstrates the importance of interactions when determining optimum pile location and design specific parameters related to the reinforcement of the slope. Chapter 4 presents a prediction tool for unreinforced slopes aiming to simplify slope stability analysis, decreasing the need for special software while providing the
reliability of the prediction tool when compared to the field condition. Chapter 5 introduces utilizing the basics of a regression-based neural network to study variable importance and the prediction power of the current prediction tools based on FSS. Chapter 6 present a novel multiple response approach intended to develop a general prediction model for the mobilized shear strength of stiff-fissured clays that can overcome FSS limitations. Finally, Chapter 7 summarizes the main conclusions and provides general recommendations for future research regarding the mobilized shear strength of stiff-fissured clays.

**Keywords:** Stiff-fissured clays, overconsolidated clay, fully softened strength, design of experiments, slope stability, limit equilibrium, numerical modeling in geotechnical engineering, multiple response artificial neural network, prediction tools.

**References**


CHAPTER 2. STATISTICAL ASSESSMENT OF FACTOR OF SAFETY FOR UNREINFORCED SLOPES

Yuderka Trinidad-Gonzalez\textsuperscript{1}, Vernon R. Schaefer\textsuperscript{2}, and Derrick K Rollins\textsuperscript{3}

Modified from a paper to be submitted to Computer-Aided Civil and Infrastructure Engineering

Abstract

This work provides a comprehensive statistical analysis of the factor of safety (FS) for homogenous unreinforced soil slopes. Statistical design of experiments (DoE) and other statistical techniques are combined with numerical methods to investigate the FS sensitivity to analysis technique and space dimensionality. For the purpose, limit Equilibrium (LE) and finite element (FE) analysis are performed in a fully randomized block design fashion for 2D and 3D spaces. From the results, statistically significant differences in the FS are found when performing 3D finite element analysis vs. limit equilibrium analyses in the 2D and 3D spaces, and 2D finite element analyses. From the analyses, a threshold FS of 1.20 is needed to achieve accuracy above 90\% for all current methods when compared to field conditions. All the analytical methods under study have a relatively low accuracy when a threshold value FS of 1 is used to indicate stability. In light of the findings, the reliability of back-calculated parameters should be carefully studied.

Keywords: Statistical design of experiments (DoE); Unreinforced slopes; Numerical methods, Factor of safety.

---

\textsuperscript{1}Graduate Research Assistant, Department of Civil, Construction and Environmental Engineering, Iowa State University, Ames, IA 50011 \texttt{yuderkat@iastate.edu}, Corresponding author
\textsuperscript{2}James M. Hoover Professor of Geotechnical Engineering, Department of Civil, Construction and Environmental Engineering, Iowa State University, Ames, IA 50011, \texttt{vern@iastate.edu}
\textsuperscript{3}Professor, Department of Chemical and Biological Engineering, Department of Statistics Iowa State University, Ames, IA 50011, \texttt{drollins@iastate.edu}
2.1 Introduction

Slope Stability can be defined as the safety of the earth masses against the movement. Hence, slope failure or mass wasting is the vertical and/or horizontal soil displacement downslope of the earth's mass, soil, rock, or debris (Turner and Schuster 1996). Slope stability analyses are conducted to evaluate if a slope is safe against such failure. If the slope is unstable, several approaches for remediation, such as geometry changes, reinforcement, or avoidance of the problem, are applied. If avoidance and/or geometry changes are not viable options, the slope may be strengthened.

Currently, there are several commonly applied numerical methods to evaluate instability, including the limit equilibrium (LE), the finite element (FE), and the finite difference (FD) methods. Several authors such as Memon (2018), Trinidad González (2017), Nian et al. (2008), Won et al. (2005), Carter et al. (2000), Griffiths and Lane (1999), Duncan (1998), Christian (1998), Duncan (1996), have presented comparisons between analytical techniques for unreinforced soil slopes. This comparison is primarily based on parametric studies and sensitivity analyses, performing one factor at time techniques.

Multi-factor sensitivity analyses have been performed through the application of orthogonal design, uniform design, artificial neural networks (ANN), and other methods, as cited by Luo et al. (2017). These sensitivity analyses primarily include a series of analytical assumptions and are performed based on a specific analytical method of analysis (e.g., either LE or FE).

Griffiths et al. (2010) presented results comparing probabilistic slope stability methods to study unreinforced homogenous slopes considering spatial variability. However, the analysis was limited to (one dimensional) 1D and 2D (two dimensional) random fields applied to LE and FE, respectively. Little information is found in the literature regarding the application of formal
statistical methodologies to address the differences in factors of safety (FS) when applying different analytical approaches under the same space dimensionality conditions (e.g., 2D or 3D). The advantages and disadvantages of using LE/FE/FD are widely reported in the literature largely based on listing the assumptions involved with each method, (Duncan 1996; Geo-Slope International 2010; Cai and Ugai 2000; Nian et al. 2008; Griffins and Lane 1999; Jeong et al. 2003; Won et al. 2005; among others). However, numerical comparisons on how the assumptions affect the FS of any given slope are not clear and are dependent on limited slope properties and geometries.

Travis et al. (2010 a,b) presented a meta-analysis of 301 slope failures to address slope stability risk assessment. The analyses used formal inferential statistics to analyze the effects of analytical methods, slope type, soil plasticity, and effective versus total stress analysis. However, the analytical methods studied are only based on 2DLE approaches (Direct, Complete, Force, and Bishop Methods). Azzouz et al. (1983) presented an FS comparison for 2D and 3D spaces (to correct measurements of the strength) for 18 slopes.

The summary of these cases was presented by Degroot and Baecher (1993) to account for model lack of fit when analyzing the overall variability (systematic, special, model) to analyze the James Bay Dyke. The conclusions suggested that modeling error increased FS by 7% for the single-staged dike.

Building in the previous work, this study provides supplementary analysis evaluating the changes in FS due to conduction slope stability analyses with analytical methods at different spaces. In addition, the relationship between analytical FS and field condition is further investigated through applying a diagnostic testing technique borrow from the medical and chemistry fields.
Hence, the main objectives of this study are:

- Studying the factors affecting $FS$ an Unreinforced slope (main effects and interactions).
- Determine whether there is a statistically significant difference between $FS$ of slopes analytically analyzed by FE and LE methods.
- Determine $FS$ sensitivity to modeling space (2D/3D).
- Determine which of the current analytical yield $FS$ closer to the field condition.

2.2 Applied Methods

2.2.1 Slope Stability Analyses

In this study, four different methods are evaluated, statistically, called “the treatments,” corresponding to 2DLE, 3DLE, 2DFE, and 3DFE condition. Readily available computer programs used for this analysis are from the Rocscience suite (Rocscience Inc.): Slide 2018, Slide3 2019, RS2 2019, and RS3 2019, corresponding to 2DLE, 3DLE, 2DFE, and 3DFE, respectively. A sketch of the sections and views of the generic slopes with the input properties is shown in Fig. 2.1.

All the input factors are presented in Table 2.1. For $\varphi' = 0$, the friction angle, and the pore water pressure coefficients are eliminated, and cohesion corresponds to the undrained cohesion.

Table 2.1. Soil and geometry properties considered factors for unreinforced slopes.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-1</td>
</tr>
<tr>
<td>Slope height (H, m)</td>
<td>6</td>
</tr>
<tr>
<td>Slope inclination ($\beta$, $^\circ$)</td>
<td>20</td>
</tr>
<tr>
<td>Cohesion (c', kPa)</td>
<td>10</td>
</tr>
<tr>
<td>Effective friction angle ($\varphi'$, $^\circ$)</td>
<td>15</td>
</tr>
<tr>
<td>Elastic modulus (E, MPa)</td>
<td>15</td>
</tr>
<tr>
<td>Poisson’s ratio ($\nu$)</td>
<td>0.35</td>
</tr>
<tr>
<td>Pore pressure coefficient ($r_u$)</td>
<td>0</td>
</tr>
<tr>
<td>Soil unit weight ($\gamma$, kN/m$^3$)</td>
<td>16</td>
</tr>
</tbody>
</table>
Fig. 2. 1. A sketch of the sections and views of the generic slopes with the input properties.

To compare $FS$ from LE analysis with the results from the FE analyses, the strength reduction technique is used. This is a procedure used in FE/FD in which $FS$ is obtained by weakening the soil in steps until the slope “fails” (Dawson et al. 2015, Griffins and Lane 1999, Pradel et al. 2010, Fu and Liao 2010). In this method, $FS$ is the ratio of the actual shear strength to the lowest shear strength of the soil material required to maintain the slope in equilibrium, also called strength reduction factor ($SRF$). For Mohr-Coulomb material the shear strength ($\tau$) reduced by $SRF$ is determined as

$$\frac{\tau}{SRF} = \frac{c'}{SRF} + \frac{\sigma'tan\phi'}{SRF}$$

(2.1)

$$c^* = \frac{c'}{SRF} \text{ and } \phi^* = arctan\left(\frac{tan\phi'}{SRF}\right)$$

(2.2)

where $c'$ and $\phi'$ are the drained Mohr-Coulomb shear strength parameters, and $c^*$ and $\phi^*$ are the reduced Mohr-Coulomb shear strength parameters. In the FE analyses, the conceptual mode of failure in 2D and 3D is a deformable, bounded material body with Mohr-
Coulomb, elastic-perfectly plastic, and non-associated flow rule assumptions (with an angle of dilatancy taken as zero). The sketch of a continuum model of a slope in 2D and the material body meshed is shown in Fig. 2.1 (c). The section represents a slope with a unit width in the z-direction. A 6-noded triangular element is used. The nodes in the mesh and the boundaries have two degrees of freedom, i.e., displacements $u$ and $w$, in the $x$- and $y$-directions, respectively.

The right and left boundaries of the mesh are fixed only in the horizontal direction ($u = 0$ at $x = 0$ and at $x = L$). No displacement is allowed at the base of the slope model ($u = w = 0$ at $y = 0$). For each slope geometry, the mesh is adjusted to guarantee the minimum number of elements to converge to the approximated solution. The results from a mesh convergence analysis for the largest ($h = 10$, $\beta = 20$) and smallest section types ($h = 6$, $\beta = 60$) are shown in Fig. 2.2. The number of elements used varies from 6,000-15,000 as the area of the sections varies.

Fig. 2. 2. Mesh convergence study for (a) the smallest ($h = 6$, $\beta = 60$) and (b) the largest ($h = 10$, $\beta = 20$) section types for the 2DFE analyses.
The sketch of a continuum model of a slope in 3D and the material body meshed is shown in Fig. 2.1 (d). A 10-noded tetrahedral element is used. The nodes in the mesh and the boundaries have three degrees of freedom, i.e., displacements $u$, $v$, and $w$, in the $x$-, $y$-, and $z$-directions, respectively. The boundary conditions are restrained displacement in the $xyz$-directions at the model ends of the planes the $yz$, $xy$, and $xz$ ($u = v = w = 0$), assuming a rigid contact with no possibility of movement at the end of the boundaries.

For the $y$-direction, the extruded face, findings from the literature from parametric studies for 3D unreinforced and pile-reinforced slopes suggested that the failure mechanism of 3D slopes is greatly affected by boundary conditions at those faces. Hence, to achieve a solution close to the plain strain solution, the ratio width/total height should be between 5 to 10 (Chugh (2003), Zhang et al. (2011), Gao et al. (2015)).

The ratio used in this study is 5. As in the 2DFE analyses, for the 3D analyses, the mesh is adjusted for each slope geometry. The results from the mesh convergence analysis for the largest ($h = 10, \beta = 20$), and smallest section types ($h = 6, \beta = 60$), are shown in Fig. 2.3. The number of elements used varies from 57,000-100,000, as the area of the extruded sections varies.

For the LE analyses, the boundaries are placed sufficiently far from the region where slope failure is expected to occur. For the 2DLE condition, the Spencer method with a non-circular mode of failure and auto refine search is used. For the 3DLE condition, a ratio width/total height of 5 is also used. The Spencer method with an ellipsoid surface and a cuckoo search is used.
Fig. 2. 3. Mesh convergence study for (a) the smallest (h = 6, β = 60) and (b) the largest (h = 10, β = 20) section types for the 3DFE analyses.

2.2.2 Statistical Methods (DoE)

Two techniques of statistical design of experiments (DoE) are used as synthetic data generators and as data evaluation tools (design and statistical analysis). In the context of this study, a synthetic data generator refers to generating data artificially by combining slope properties. DoE is a methodology used for planning and conducting experiments and later data evaluation. It was first applied by Ronald Fisher for conducting agricultural research around the 1920s (Durakovic 2018). Because of its multipurpose nature, DoE can be used in various conditions such as comparative designs, variable screening, transfer function identification, optimization, and robust design. The methodology is currently greatly applied in scientific research areas, mainly in the biomedical, engineering, and biochemistry fields (Durakovic 2018).
2.2.2.1 Factors Screening

A face-centered central composite design “CCD” is used to combine slopes and soil properties in a defined range of input space. CCD is a three-level surface technique part of the response surface methodology developed by Box and collaborators in the 1950s (Bezerra et al. 2008). The methodology is based on mathematical and statistical techniques for fitting empirical and semi-empirical model structures to experimental data. In this study, CCD is used to create an approximating function simpler than running the numerical models from LE and FE analyses. If the true response from slope stability mechanism is a function $y(x_1, x_2, ..., x_8)$, the approximated response from CCD is $\hat{y} (x_1, x_2, ..., x_8)$ (Wong 1985). The mathematical expression for symmetrical response surface designs is given by Eqs. 2.3 to 2.5;

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \sum_{j=1}^{k} \beta_{ij} x_i x_j + \varepsilon$$ (2.3)

$$\sum_{i=1}^{k} \beta_i x_i = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_8 x_8$$ (2.4)

$$\sum_{i=1}^{k} \sum_{j=1}^{k} \beta_{ij} x_i x_j = \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \cdots + \beta_{78} x_7 x_8$$ (2.5)

where $y$ is the response of the output variable in the input space of the $x$’s, $k$ is the number of input variables (in this case, $k = 8$), $\beta_0$ is the constant term, $\beta_i$ represents the linear coefficients for the $x_i$’s, $\beta_{ij}$ represents the coefficients with the $x_i x_j$ interaction terms, $\varepsilon$ is the random deviation term (giving the output, $y$, its stochastic nature) and is independently normal with mean 0 and variance $\sigma^2$. In this study, random errors are primarily due to reading inaccuracy and geometry rounding up. Therefore, it is expected that random errors are small and that there is some systematic error due to model miscalibration. A face-centered design CCD is selected
because of the efficacy and flexibility of CCDs and because the area of operability cannot exist for extremes outside of the studied region. For instance, outside of the specific range of studied inputs, negative values for the factors are not physically possible. The results from this analysis are compared with the findings from the literature to ensure repeatability and accuracy of the results.

A preliminary evaluation is performed to verify the findings presented by Griffiths and Lane (1999). These findings suggested that the FS of a slope is not significantly sensitive to soil elastic modulus (E) and the Poisson’s ratio (υ). The initial number of factors for the study is presented in Table 1, along with the levels for each of the factors. Each level represents changes in the slope properties (factors understudy). Statistically, each level is a defined value that the factor takes between the high (+1) and low levels (-1). To determine whether or not the model is useful, the model utility test (MUT) is conducted at the significance level $\alpha = 0.05$. Non-significant factors are removed in further analysis by a statistical model discrimination technique to retain in the model only the significant coefficients. The discrimination technique is a stepwise regression algorithm. The null hypothesis and the alternative hypothesis for MUT are given, respectively, by:

$$H_0 = \beta_1 = \beta_2 = \cdots = \beta_k = 0;$$

$$H_1: \beta_i \neq 0, \text{ for at least one } i \neq 0$$

The test statistic is

$$F = \frac{SSR/k}{SSE/[n - (k + 1)]}$$

where $k$ is the number of carriers (i.e., linear and interaction terms), $SSE$ and $SSR$ are error and regression sum of squares, respectively. This test rejects $H_0$ when the $P$-value $< \alpha = 0.05$. 
2.2.2.2 Randomized Complete Block Design

After the input relationships are determined, a Randomized Complete Block Design “RCBD” is conducted. In a RCBD, each block receives the treatment exactly once, and it is used to control variation within the blocks. The RCBD is an extension of the paired t-test when the factor of interest has more than two levels (Montgomery and Runger 2007). In the context of this study, each slope configuration is treated as a block (run is a blocking factor), analytical method is the factor under study, each analytical method (FE/FD in either 2D/3D space) is a “treatment,” and the response variable is FS.

Table 2. 2. RCBD for unreinforced soil slopes configuration.

<table>
<thead>
<tr>
<th>Block factor</th>
<th>Analytical Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Runs (Ri)</td>
<td>Treatment 1</td>
</tr>
<tr>
<td></td>
<td>Treatment 2</td>
</tr>
<tr>
<td></td>
<td>Treatment 3</td>
</tr>
<tr>
<td></td>
<td>Treatment 4</td>
</tr>
<tr>
<td>2DLE</td>
<td>2DFE</td>
</tr>
<tr>
<td>3DLE</td>
<td>3DFE</td>
</tr>
</tbody>
</table>

A run is a combination of slope properties. For instance, 6m slope, inclined at 30 degrees with a friction angle of 20 degrees, a cohesion of 35kPa, E = 40Mpa, and v = 0.35 is a run.

The use of a RCBD allows comparing treatment effects on the response while controlling the effects of the blocking factor (slope property changes). The model for a RCBD is given as:

\[ y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}, \quad i = 1, ..., a, j = 1, ..., b \]  

where \( y_{ij} \) is the response for ith treatment in the jth block, \( \mu \) is the overall mean, \( \tau_i \) the effect \( \mu_i - \mu \) of the ith treatment, \( B_j \) is the effect of the jth block, \( a \) is the number of treatments (4), \( b \) is the number of blocks, and \( \epsilon_{ij} \) are random deviations that are independent normal random variables with mean 0 and variance \( \sigma^2 \). The model is structured to achieve an unbiased estimator for the model coefficients, by using the restrictions \( \Sigma \tau_i = 0 \) and \( \Sigma \beta_j = 0 \).
The objective of this RCBD is to determine whether or not the mean FS for different analytical methods are significantly different, using synthetic data generated from the CCD design. The number of observations (the number of runs = number of blocks) is determined so that the following sizes of treatment effect \( \tau_i \) can be detected while controlling for type I, and type II errors of \( \alpha = 0.05 \) and \( \beta = 0.05 \) (95% power), respectively.

\[
\begin{align*}
\tau_1 &= -\frac{\sigma}{4}, \tau_2 = -\frac{\sigma}{4}, \tau_3 = \frac{\sigma}{4}, \tau_4 = \frac{\sigma}{4} \\
\Delta &= \sqrt{\frac{2}{\sigma^2} \left( \left( -\frac{\sigma}{4} \right)^2 + \left( -\frac{\sigma}{4} \right)^2 + \left( \frac{\sigma}{4} \right)^2 + \left( \frac{\sigma}{4} \right)^2 \right)} = 0.7
\end{align*}
\] (2.10)

where \( \sigma \) is standard deviation and \( \tau_i \) is \( i \)th treatment effect of the factor. The total number of observations per treatment is 72. Each block receives each treatment combination once. Thus, the total number of runs needed is 72. The significant differences among the treatment means are determined by applying Tukey’s honest significant difference (HSD) method below:

\[
|y_i - y_j| > HSD
\] (2.12)

\[
HSD = qa_{(a,(a-1)(b-1)}} \sqrt{\frac{MSError}{b}}
\] (2.13)

where, \( qa_{(a,(a-1)(b-1)}} \) is the upper \( \alpha \) percentile of the studentized range distribution, \( a \) is the number of treatments (4 analytical methods), \( b \) is the number of blocks (72 runs), \( MSError \) is the mean squared error found in the ANOVA table. In a pairwise comparison using Tukey’s method, two true means differences, \( \mu_i - \mu_j \), corresponding to \( y_i - y_j \), are concluded to be significantly different if their difference is larger than the HSD value, as shown by Eq. 10. 100(1 - \( \alpha \))% confidence intervals for \( \mu_i - \mu_j \) are also provided.
2.2.2.3 Application of Bayes’ theorem to diagnostic testing

To determine which analytical condition is closer to the field condition, a different approach is taken. The slope failure information found in the literature, originally presented by Sah et al. (1994) and subsequently used by authors such as Sakellariou and Ferentinou (2005), Samui and Kothari (2011), is verified for accuracy and organized for this analysis. From 68 cases, 39 (verified with the original sources) are selected to be modeled using LE and FE methods. The range of the properties of the slopes selected from the case histories is presented in Table 2.3.

Table 2.3. Soil properties from the literature review originally from Sah et al. (1994).

<table>
<thead>
<tr>
<th>Slope height (H) (m)</th>
<th>Slope inclination (β) (°)</th>
<th>Cohesion (c’ or c) (kPa)</th>
<th>Friction angle (φ’) (°)</th>
<th>Pore pressure coefficient (r_u)</th>
<th>Soil unit weight (γ) (kN/m²)</th>
<th>Slope Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.66-214</td>
<td>20-70</td>
<td>0-50.05</td>
<td>0-45</td>
<td>0-50</td>
<td>12-28.44</td>
<td>Failed/Stable</td>
</tr>
</tbody>
</table>

Because the factors of safety given as a number in the previous reports correspond to a certain analysis method (generally LE analysis), the condition of the slopes, whether stable or failed, is coded as 1 (stable) and 0 (failed). The numerical results from the analytical methods are coded to different threshold values of FS for stable: 1.0, 1.10, and 1.20. This coding is used because Travis et al. (2010b) found that certain slopes with FS >1 still failed. Bayes’ theorem is then applied to determine the probability that using a specific analytical method a slope classified as failed (by LE/FE method) is, in fact, failed, assessing the occurrence of false positives and false negatives. Considering failed as positive and stable as negative (the programs perform a test), false negative results in a dangerous and unreliable design, while false positive results in a conservative design.
A formal definition of Bayes’ theorem (Porkess, 2005) states that for two events A and B,

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$  \hspace{1cm} (2.14)

where $P(A|B)$ is the probability of event A occurring if event B has already occurred, and

$$P(B) = P(B|A)P(A) + P(B|A')P(A')$$  \hspace{1cm} (2.15)

The probability of event B occurring, where the accent symbol,’ means the complement of the event. This theorem has been widely applied as a research tool in several areas, including biomedical, engineering, medicine, and the general field called “Bayesian statistics.” A tree diagram of the application of Bayes’ theorem to the context of this study is presented in Fig. 2.4.

![Tree Diagram](image)

Fig. 2.4. Probabilities tree diagram apply to the slope stability context.

Events are defined as:
- $A = F$ = Slope actually failed (from the field information),
- $A' = S$ = Slope actually stable (from the field information)
- $B = PF$ = Analytical method classifying a slope as failed (+)
- $B' = PS$ = Analytical method classifying a slope as stable (-)
Applying Bayes’ theorem to diagnostic testing as presented by Diamond (1999) (where each analytical method corresponds to a diagnostic test):

- **TP** = True positive, slopes classified as failed that are actually failed.
- **TN** = True negative, slopes classified as stable that are actually stable.
- **FP** = False positive, slopes classified as failed that are actually stable.
- **FN** = False negative, slopes classified as stable that are actually failed.

Sensitivity in this context is defined as the proportion of slopes that actually failed that are correctly classified as failed.

\[
S_s = \frac{TP}{FN + TP} = P(PF/F) \tag{2.16}
\]

Specificity is then the proportion of actual stable slopes correctly identified as stable.

\[
S_p = \frac{TN}{FP + TN} = P(PS/S) \tag{2.17}
\]

Accuracy is the overall probability that the slope is correctly classified.

\[
A_{cc} = \frac{TP + TN}{FP + TN + FN + TP} \tag{2.18}
\]

### 2.3 Results

#### 2.3.1 Slope Stability Results

The slope stability analyses are conducted following the CCD design for 2DLE, 2DFE, 3DLE, and 3DFE. An example of results compared in this study is presented in Fig. 2.5. The results correspond to Run four from Table A.1. These results are provided to illustrate the degree of agreement of critical surface location among the comparisons.
Fig. 2. 5. Sample of results location of critical surface among methods for Run 4.

2.3.2 Significant Factors

After the assessment of model adequacy (including normality, constant variance, independence), a summary of the average of all coefficients of determination $R^2$ (the proportion of explained variation) is presented in Table 2.4.

Table 2.4. Summary average coefficient of determination for all Models.

<table>
<thead>
<tr>
<th>Model</th>
<th>S</th>
<th>R-sq</th>
<th>R-sq(adj)</th>
<th>R-sq(pred)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average all methods</td>
<td>0.04</td>
<td>99.9%</td>
<td>99.8%</td>
<td>99.6%</td>
</tr>
</tbody>
</table>

A general classification of soils into three main groups is used as: purely cohesive (c), purely frictional ($\phi'$), and mixed soils ($c'$ and $\phi'$), the last group being the most common. The analysis of the significant factors is performed separately under each soil condition, and the results are summarized in the Pareto charts presented in Figures 2.6-2.8. A Pareto chart shows absolute values of standardized effects from larger to smallest. The reference line on the plot
indicates the $t$-statistic value that tests the null hypothesis (the important coefficients for the model). Hence, the red line is a threshold to classify effects as significant or not. The reference line indicates the difference in significance from the highest to the lowest standardized effect.

For the mixed soils, Fig. 2.6 shows that six main coefficients ($c', \beta, r_u, \phi', \gamma, H$), and their interaction effects have the most significant effects on the response. The Elastic modulus and the Poisson’s ratio at the significant level of study are not significant for determining the slope factor of safety. Nevertheless, changes in both $E$ and $\nu$ need to be considerably large in order to significantly affect the factor of safety. Increasing the soil strength parameters ($c', \phi'$) results in an increase in the factor of safety.

On the contrary, increasing slope geometry properties ($H, \beta$) result in decreasing the factor of safety. The negative effects of the inclination angle are amplified as the cohesion of a mixed soil decreases. In the same way, increasing $r_u$ and $\gamma$ negatively affect the factor of safety. The results from Fig. 2.6 are in agreement with the findings presented by Kostić et al. (2016) when analyzing mixed soil slopes. The cohesion and slope inclination angle are the coefficients with the largest positive and negative effects on the response, respectively.

Understanding the predominant effects is a useful tool for the optimum selection of remediation alternatives. For instance, in mixed soil slopes, one primary goal could be increasing the cohesion of the soil by means of soil improvement or soil reinforcement.

Decreasing or flattening the slope could be the secondary target to achieve a more efficient solution. Dewatering is also a crucial alternative to increase the FS. A combination of alternatives accounting for both economical and effects hierarchy could work as a tool for optimization, maximizing the FS while reducing the cost by selection efficient solutions.
The analysis is then conducted for the purely cohesive and the purely frictional soils.

The Pareto chart for $\varphi'$ soils, Fig. 2.7, indicates that the main coefficients that affect the factor of safety are the inclination angle, the friction angle, and the pore water coefficients.

Fig. 2. 6. Standardized effects ($\alpha = 0.05$), mixed soils ($c'$ and $\varphi'$). The red line indicates the t-statistic value to classify effects as statistically significant.
The effects of unit weight and slope height on FS are not significant for a purely frictional soil. For this condition, flattening the slope combined with dewatering or flattening the slope combined with reinforcing the soil are potential effective combinations for remediation alternatives. On the other hand, the results from Fig. 2.8 indicate that for a purely cohesive soil, the primary factors affecting the response are cohesion, slope height, unit weight, and angle of inclination. The negative effect of the slope height is amplified for purely cohesive soils. Increasing the cohesion and decreasing the height of the slope may represent a good alternative for maximizing the FS in these types of slopes.
2.3.3 Means FS Comparisons

The mean FS (the mean response) for all the treatments are pairwise compared using Tukey’s method after the RCBD is analyzed. The model adequacy is verified, and the analysis of variance table is presented in Table 2.5. The results show that the P-values for both factors are less than 0.0001, indicating that both factors are statistically significant at $\alpha = 0.05$.

Table 2.5. Analysis of Variance Table RCBD unreinforced.

<table>
<thead>
<tr>
<th>Source</th>
<th>*DF</th>
<th>Sum Squares</th>
<th>F Ratio</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Run</td>
<td>71</td>
<td>441.962</td>
<td>504.1849</td>
<td>&lt;.0001*</td>
</tr>
<tr>
<td>Method</td>
<td>3</td>
<td>2.12077</td>
<td>57.2579</td>
<td>&lt;.0001*</td>
</tr>
<tr>
<td>Error</td>
<td>213</td>
<td>2.62977</td>
<td>0.01235</td>
<td></td>
</tr>
<tr>
<td>C. Total</td>
<td>287</td>
<td>446.713</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Degrees of freedom

Fig. 2. 8. Standardized effects ($\alpha = 0.05$), undrained ($\varphi = 0$). The red line indicates the t-statistic value to classify effects as statistically significant.
The sum of squares error for runs is higher than the sum of squares error for method. This is expected because it was previously found that slope properties have a stronger influence on the FS. In this section, the factor under study is method of analysis, and by using slope as a blocking factor, the test sensitivity for the factor method increased. After verifying the significance of the method of analysis with a \( p\)-value \(< 0.05\), pairwise comparison is conducted. The results are summarized and presented in Table 2.6.

Table 2.6. Differences report pairwise comparison mean FS.

<table>
<thead>
<tr>
<th>Level</th>
<th>-</th>
<th>Difference</th>
<th>(*SE_{df})</th>
<th>Lower CL</th>
<th>Upper CL</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>3DFE</td>
<td>2DLE</td>
<td>0.21</td>
<td>0.0185</td>
<td>0.16</td>
<td>0.26</td>
<td>&lt;.0001*</td>
</tr>
<tr>
<td>3DFE</td>
<td>2DFE</td>
<td>0.18</td>
<td>0.0185</td>
<td>0.14</td>
<td>0.23</td>
<td>&lt;.0001*</td>
</tr>
<tr>
<td>3DFE</td>
<td>3DLE</td>
<td>0.18</td>
<td>0.0185</td>
<td>0.13</td>
<td>0.23</td>
<td>&lt;.0001*</td>
</tr>
<tr>
<td>3DLE</td>
<td>2DLE</td>
<td>0.02</td>
<td>0.0185</td>
<td>-0.02</td>
<td>0.07</td>
<td>0.5011</td>
</tr>
<tr>
<td>2DFE</td>
<td>2DLE</td>
<td>0.02</td>
<td>0.0185</td>
<td>-0.02</td>
<td>0.07</td>
<td>0.5904</td>
</tr>
<tr>
<td>3DLE</td>
<td>2DFE</td>
<td>0.002</td>
<td>0.0185</td>
<td>-0.04</td>
<td>0.05</td>
<td>0.9990</td>
</tr>
</tbody>
</table>

\(*SE_{df}\) = Standard error of the difference

From the results, at the significance level \( \alpha = 0.05\), the mean FS from 3DFE analysis is found statistically significant with the other three analytical methods (2DLE, 2DFE, 3DLE). Hence, it is expected that the FE (3D) analysis results in higher FS than the LE in 2D and 3D and FE (2D). The largest sample mean difference in FS is found to be around 0.21 between 2DLE (most commonly used) and 3DFE. In practice, this difference is not very significant if the slope is far from the failure or far from stability. However, for slopes around the threshold FS value, around one, 0.21 could be of key importance for decision-making.

2.3.4 Diagnostic Testing Results (Field validation)

The results from this section are useful to verify which method or methods is closer to the real field condition. A graphical representation of a confusion matrix is presented in Fig.2.9 (a and b).
Fig. 2.9 (a) presents the slopes actually stable from literature, the dashed line indicates the boundary between failure and stability, assuming one as the boundary FS. False positives (stable slopes classify as failed) are below the dashed line.

Sensitivity (Ss), Specificity (Sp), and Accuracy (A) are then determined for each case. Figure 9 (a) shows that for all methods, characterizing a stable slope as failed when the slope is stable represents a low percentage of the studied cases. On the contrary, from Fig. 9 (b), there is...
a considerable amount of data points above the dashed line into the false negative zone. The false-negative zone represents slopes that are actually failed but are classified by the analytical methods as stable. Designing for this condition results in an unsafe and unreliable design. A summary of the results is presented in Fig. 2.10.

Fig. 2.10. Summary field condition vs. analytical methods from diagnostic testing.

A minimum accuracy line of 90% is used for comparison purposes. The red arrow indicates increasing accuracy as the FS threshold changes. The results indicate that 2DLE methods have the highest sensitivity and accuracy for all the threshold values. Conversely, the differences between 2DLE methods and 2DFE and 3DLE decrease as the threshold value are increased. Assuming slopes with FS of 1.0 are stable and below FS of 1.0 are failed, all of the analytical methods seem to have relatively low accuracy in predicting the slope condition. This is related to the variability of soil properties at the time of the failure and the accuracy of the measured strength at the failure surfaces of the slopes registered in the literature. The highest
accuracy is achieved from 2DLE around 79.4%. The lowest accuracy, around 66.6%, is found when performing a 3DFE analysis. This could be related to the difficulty of creating 3D geometry and their high dependency on boundary conditions. These results are in agreement with the findings in the literature that 2DLE results tend to be more “conservative” (Duncan 1996). The results also indicate that from a classification point of view to minimized false negatives (slopes predicted as stable when in fact, they are failed), conservativism is needed. Using a threshold value FS of 1.2 significantly increases sensitivity, and the accuracy of the analytical methods, reducing the probability of false negatives that results in unreliable designs.

At this condition (using threshold FS of 1.2), LE 2D/3D showed the highest accuracy of around 95%. From this analysis, all the analytical methods can achieve accuracy around and beyond 90%, if a threshold value is set to 1.2 as a minimum. Hence, FS of 1.0 as a threshold between stability and failure is not appropriate. In this analysis, it is seldom found (1/39 cases) a stable slope classified as failed (the analytical methods predict as FS < 1.0). However, several actually failed slopes are found around FS of 1.01 to FS of 1.20 by the analytical methods. This may be due to the uncertainties and variabilities involved in the soil properties as inputs for the slope stability analyses. The exact strength value at the time of failure of a soil slope is unknown, and the soil variability can just be considered in the analyses by introducing stochastic random field, transient analysis, or progressive failure models. In addition, these findings may indicate that back-calculation analysis assuming FS of 1.0 is not necessarily reliable since the resulting soil strength slightly or significantly varied from one method to another.

2.4 Conclusions and Recommendations

In this paper, a comprehensive statistical analysis of the variability of the factor of safety for unreinforced slopes is conducted. The applied statistical tools provide formal insight into the slope stability mechanism. From the results, the following conclusions can be drawn:
Understanding the factors involved in the slope stability mechanism is crucial to determining an optimum remediation technique. In general, for all soil types used in this study, the results indicated that six primary factors \((c', \beta, \text{ru}, \varphi', \gamma, H)\) and their interactions have statistically significant effects on FS at \(\alpha = 0.05\). In addition, the findings statistically demonstrate that the elastic modulus and the Poisson’s ratio do not affect FS significantly.

Cohesion and friction angle of soils have different effects on the FS. In this study, three main groups of soil are evaluated: purely cohesive \((c)\), purely frictional \((\varphi')\), and mixed soils \((c'\text{ and } \varphi')\).

For \(\varphi'\) soils, the negative effect of the slope inclination angle and pore water pressure coefficient are amplified. This makes regrading and dewatering as primary candidates for remediation techniques.

For \(c\) soils, slope height, unit weight, and slope inclination angle are the most negative contributors to FS. The largest positive effect comes from cohesion. Increasing the cohesion and decreasing the height of the slope may represent a good alternative for maximizing the FS in slopes containing \(c\) soils.

For mixed soils, the cohesion and slope inclination angle are the coefficients with the largest positive and negative effects on the response, respectively. Candidates for remediation for mixed soil sloped could be increasing the cohesion of the soil by means of soil improvement or soil reinforcement, decreasing or flattening the slope, and dewatering. A combination of alternatives accounting for both economical and effects hierarchy could work as a tool for optimization, maximizing the FS while reducing the cost by selection efficient solutions.
• For the three soil groups, increasing the shear strength parameters of the soils increased the FS. This makes slope reinforcement an efficient technique to improve slope stability. Considering the effects of $c$ in the overall slope stability mechanism, for all types of soil, introducing (if the soil is cohesionless) or increasing cohesion results in higher stability.

• The comparison of mean FS by RCBD indicated that there is no statistically significant difference at $\alpha = 0.05$ for any method but the 3DFE method. The results exhibited that introducing 3D space in the slope stability analysis caused more variability in FS than the variabilities due to the method assumptions. The largest mean difference in FS is found to be around 0.21. In practice, this difference would not be very significant for slopes far from failure but could be important for slopes showing low FS. For decision-making, each individual slope condition must be carefully studied.

• The diagnostic testing results indicated that the analytical methods in this study had a relatively low accuracy when a threshold FS value of 1 is used to indicate stability/instability. The LE method is the one with the largest agreement with the field condition with a 78% accuracy. This suggested that considering a unit value for the threshold between failure and stability might not be analytically true. The properties determined at failure may include measurement error and inherent variability. To prevent unsafe design, analytically, a threshold FS value greater than 1.20 is recommended based on this study. In light of the findings, the reliability of back-calculated parameters should be carefully studied.
All in all, the current analytical methods for slope stability analysis are good estimating tools for simple homogeneous unreinforced slopes. Caution should be taken when performing slope stability analyses. The importance of truthfully representing the soil properties is highlighted by the low accuracy that the studied methods exhibited for slopes laying in the boundary between failure and stability. An agreement between analytical methods does not guarantee a 100% agreement with the real condition. A combination of methods and engineering judgment must be exercised for decision making.

References


CHAPTER 3. STATISTICAL ASSESSMENT OF F FOR PILE-REINFORCED SLOPES

Yuderkat Trinidad-Gonzalez1, Vernon R. Schaefer2, and Derrick K Rollins3

Modified from a paper accepted by the Journal of Geotechnical and Geoenvironmental Engineering - ASCE

Abstract

Statistical design of experiments is combined with coupled and uncoupled pile-reinforced slope stability analyses to draw conclusions on the factor of safety (FS) sensitivity to pile related properties (pile location, spacing, embedment depth, and diameter), and to analyze differences between analytical techniques and space dimensionality. This evaluation is done by performing 2D and 3D, limit equilibrium (LE), and finite element (FE) analysis in a fully randomized block design, a definitive screening design, and a central composite design fashion. Pile location, embedment depth, and spacing are the pile properties with significant effects on $FS$. The optimum pile location is not fixed and varies not only with changes in soil strength but also with variations in geometry properties because of the presence of interactions among the factors. The comparative analysis shows no statistically significant difference between the studied methods. However, the confidence intervals show $FS$ differences as high as 0.27, a variation of crucial importance for practical decision-making when $FS$ is around the stability/failure threshold of one. Recommendations for analysis and design of pile-reinforced slopes are given.

Keywords: Statistical design of experiments (DoE); Pile-reinforced slopes; Numerical methods, Factor of safety.

---

1Graduate Research Assistant, Department of Civil, Construction and Environmental Engineering, Iowa State University, Ames, IA 50011 yuderkat@iastate.edu, Corresponding author
2James M. Hoover Professor of Geotechnical Engineering, Department of Civil, Construction and Environmental Engineering, Iowa State University, Ames, IA 50011, vern@iastate.edu
3Professor, Department of Chemical and Biological Engineering, Department of Statistics Iowa State University, Ames, IA 50011, drollins@iastate.edu
3.1 Introduction

Pile elements have been found useful and economical to improve the stability of slopes (De Beer and Wallays (1970), Ito and Matsui (1975), Cai and Ugai (2000), Hassiotis et al. (1997), Lee et al. (1995), Poulos (1995), Chen and Poulos (1997)). However, evaluating the stability of pile-reinforced slopes remains a challenge in the geotechnical field because soil-structure interaction is complex. The basis of the analysis to assess the stability of pile-reinforced slopes is to determine the factor of safety ($FS$). The $FS$ is commonly defined as the ratio between resisting forces and the driving forces along the failure surface. When analyzing pile-reinforced slopes, an assessment of the piles must also be performed to verify the resistance of the pile to the loads developed by the soil movement (Hassiotis et al. 1997). The overall analysis can be done coupled (pile and slope are analyzed together) or uncoupled (stabilizing pile forces and slope stability analyzed separately).

Regarding coupled analysis, numerical methods such as finite element (FE) and finite difference (FD) methods allow for the simultaneous characterization of pile and slope. FE and FD approaches are considered more useful and accurate but also time-consuming, complicated, and costly when compared to limit equilibrium (LE) methods (Duncan 1996, Jeong et al. 2003, Nian et al. 2008). The use of numerical techniques results in performing a single analysis. No assumptions of the location and shape of the critical failure surface are needed, sensitivity analysis to determine design alternatives is easily done, and progressive failure and shear failure can be monitored. Some of the disadvantages of these methods are related to the complexity and learning curve for the application of the technique, the cost of the software packages compared to LE programs, the number of inputs, laboratory calibration for representing the soil properties, and the knowledge required to avoid misuse and misinterpretation of the results.
In contrast, procedures based on LE analyses (uncoupled) do not include the stabilizing pile responses on the analysis. As a result, an additional study of the stabilizing pile must be conducted. The general design approach for stabilizing piles follows a procedure presented by Viggiani (1981) in which the three main steps are: (1) evaluate the needed shear force to increase the slope $FS$; (2) assess the maximum shear force provided by each pile; and (3) select the number of piles and their optimum location. The first step is conducted by performing a stability analysis with a target $FS$. From the difference between the unreinforced $FS$ and the target value, the required stabilizing force is calculated. The second step is addressed by performing a lateral pile response analysis to verify the required stabilizing force available from a particular pile. The available methods can be categorized as pressure-based methods (De Beer and Wallays 1970, Ito and Matsui 1975, Hassiotis et al. 1997) and displacement-based methods (Poulos 1995, Chen and Poulos 1997).

The pressure-based method is built on analyzing piles subjected to lateral pressures as passive piles with a technique developed by Ito and Matsui (1975). This method was developed, assuming that the soil surrounding the pile undergoes two plastic states -- deformation and flow. The first stage satisfies the Mohr-Coulomb yield criterion, and in the second one, the ground is considered as a visco-plastic solid. Hence, plastic deformation is compared to hard soil layers, and plastic flow is compared to creep deformation of a soft layer.

The displacement-based method determines lateral pile responses to the lateral ground movements. Therefore, the lateral soil movement needs to be assessed for its application. The method can use either measured inclinometer data or analytical results from a FE analysis to simulate the soil-pile interaction mechanism. Poulos (1973) presented an approach to determine laterally loaded pile responses using a soil-pile interaction analysis in which the movement of the
soil through the pile is considered. Poulos (1973) used a simplified boundary element method where the pile was modeled as an elastic beam and the soil as an elastic continuum. The lateral displacements for each pile element are associated with its bending stiffness, and the horizontal soil-pile interaction stresses. Most of the available methods to analyze laterally loaded piles are based on the $p$-$y$ method. This method uses a FD technique for solving the nonlinear fourth-order differential equation of a beam-column system on an elastic foundation. This equation was initially presented by Hetenyi (1946) and subsequently modified by Poulos (1973) and Byrne et al. (1984) to analyze the passive case in the slope stability context. In the $p$-$y$ method, the soil resistance is simulated as non-linear springs, with $p$ representing the soil pressure, and $y$ the pile deflection. $p$-$y$ curves are the empirical representation of the soil-pile relationship that varies among soil types.

Procedures based on LE methods have been widely used due to their simplicity and easy application. When performing a displacement-based analysis, for an active slip, the direct measurement of accurate pile responses from the soil movement is possible. Nonetheless, the methods possess some drawbacks related to the assumptions needed to apply the techniques. The main disadvantages of the uncoupled method are: (1) two analyses are required (slope and pile); (2) an assumption is required for the shape and location of the failure surface; (3) the soil movement is considered as a rigid body block; (4) uniform location of shear stresses; (5) soil-structure interaction is assumed or not considered; and (6) the $p$-$y$ curves are not defined for all soil types and, in addition, they may need laboratory or field calibration to ensure reliable results.

Several authors such as Memon (2018), Trinidad González (2017), Nian et al. (2008), Won et al. (2005), Jeong et al. (2003), Carter et al. (2000), Griffiths and Lane (1999), Duncan (1998), Christian (1998), Duncan (1996), Cheng et al. (2007), have presented comparisons
between analytical techniques for unreinforced and pile-reinforced slopes, based on parametric studies and sensitivity analyses mainly changing one factor at a time. The literature is limited regarding the application of formal statistical theories to address the differences in FS when applying different analytical approaches under the same space dimensionality conditions (e.g., 2D or 3D), when introducing piles as a reinforcing element.

Pioneering work by Cai and Ugai (2000) Jeong et al. (2003), Won et al. (2005), and Trinidad González (2017) conducted comprehensive comparative studies for coupled and uncoupled approaches. Building on this work, our goal is to evaluate interactions using formal statistical inference and data analysis.

Two significant differences between traditional sensitivity analysis and applying statistical design of experiments (DoE) can be highlighted. DoE techniques pre-evaluate data collection that arrange the levels of factors to minimize uncertainty and maximize information content. After data collection, DoE analysis allows for the factor characterization of main effects and interactions that have been commonly left out in a traditional sensitivity analysis. Hence, the main objectives of this study are:

- Studying the factors affecting FS in a pile-reinforced slope (main effects and interactions).
- Determining whether there is a statistically significant difference between the FS of slopes analyzed by FE and LE methods.
- Determining FS sensitivity to the modeling space (2D or 3D) for pile-reinforced slopes.
3.2 Applied Methods

3.2.1 Slope Stability Analyses

Four different methods of analysis are evaluated, statistically called “the treatments,” corresponding to 2DLE, 3DLE, 2DFE, and 3DFE conditions. Readily available computer programs used for this analysis are from the Rocscience suite (Rocscience Inc.): Slide 2018, Slide3 2019, RS2 2019, and RS3 2019, corresponding to 2DLE, 3DLE, 2DFE, and 3DFE, respectively. Geometry differences between programs are eliminated by using the import geometry function. A sketch of the sections and views of the generic slopes with the input properties is shown in Fig. 3.1. All the input factors are presented in Table 3.1. To compare FS from LE analysis with the results from the FE analyses, the strength reduction technique is used.

Fig. 3.1. Graphical representation of geometry and inputs for pile-reinforced slopes for (a) 2DLE, (b) 2DFE, (c) 3DLE and (d) 3DFE.
The shear strength reduction method is a procedure used in FE/FD in which $FS$ is obtained by weakening the soil in steps in an elastic-plastic FE or FD analysis until the slope “fails” (Dawson et al. 2015, Griffins and Lane 1999, Pradel et al. 2010, Fu and Liao 2010). In this method, $FS$ is considered to be the ratio of the actual shear strength to the lowest shear strength of a rock or soil material that is required to maintain the slope in equilibrium, also called strength reduction factor ($SRF$). Numerically, failure occurs when it is no longer possible to obtain a converged solution. For Mohr-Coulomb material the shear strength ($\tau$) reduced by $SRF$ is determined as

$$\frac{\tau}{SRF} = \frac{c'}{SRF} + \frac{\sigma'tan\phi'}{SRF}$$  

$$c^* = \frac{c'}{SRF} \text{ and } \phi^* = arctan\left(\frac{tan\phi'}{SRF}\right)$$

where $c'$ and $\phi'$ are the drained Mohr-Coulomb shear strength parameters, and $c^*$ and $\phi^*$ are the reduced Mohr-Coulomb shear strength parameters.

Table 3. 1. Values of the factors under study called “Input factors.”

<table>
<thead>
<tr>
<th>Factors</th>
<th>Factors abbreviation</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope angle (°)</td>
<td>$\beta$</td>
<td>-</td>
</tr>
<tr>
<td>Slope height (m)</td>
<td>$H$</td>
<td>20 40 60</td>
</tr>
<tr>
<td>Cohesion (kPa)</td>
<td>$c'$</td>
<td>6 8 10</td>
</tr>
<tr>
<td>Effective friction angle (°)</td>
<td>$\phi'$</td>
<td>10 22.5 35</td>
</tr>
<tr>
<td>Elastic modulus (MPa)</td>
<td>$E$</td>
<td>15 40 65</td>
</tr>
<tr>
<td>Poisson’s ratio</td>
<td>$\nu$</td>
<td>0.35 0.4 0.45</td>
</tr>
<tr>
<td>Pore pressure coefficient</td>
<td>$r_u$</td>
<td>0 0.3 0.6</td>
</tr>
<tr>
<td>Soil unit weight (kN/m³)</td>
<td>$\gamma$</td>
<td>14 17 20</td>
</tr>
<tr>
<td>Pile location (From toe to pile)</td>
<td>$PL$</td>
<td>0 0.5 1</td>
</tr>
<tr>
<td>Embedment depth (Fraction*H, m)</td>
<td>$ED$</td>
<td>0.5 1 1.5</td>
</tr>
<tr>
<td>Pile Diameter (m)</td>
<td>$Dia.$</td>
<td>0.5 1.5 2.5</td>
</tr>
<tr>
<td>Spacing (Constant*Diameter)</td>
<td>$S$</td>
<td>1 4 7</td>
</tr>
</tbody>
</table>

In the FE analyses, the conceptual mode of failure in 2D and 3D is a deformable, bounded material body with Mohr-Coulomb, elastic-perfectly plastic, and non-associated flow.
rule assumptions (with an angle of dilatancy taken as zero). The sketch of a continuum model of a slope in 2D and the material body meshed is shown in Fig. 3.1 (c). The section represents a slope with a unit width in the z-direction. A 6-noded triangular element is used. The nodes in the mesh and the boundaries have two degrees of freedom, i.e., displacements $u$ and $w$, in the $x$- and $y$-directions, respectively. The right and left boundaries of the mesh are fixed only in the horizontal direction ($u = 0$ at $x = 0$ and at $x = L$). No displacement is allowed at the base of the slope model ($u = w = 0$ at $y = 0$). For each slope geometry, the mesh is adjusted to guarantee the minimum number of elements to converge to the approximated solution. The results from a mesh convergence analysis for the largest ($h = 10$, $\beta = 20$) and smallest section types ($h = 6$, $\beta = 60$) are shown in Fig. 3.2. The number of elements used varies from 6,000-15,000 as the area of the sections varies.

![Graph 1](image1)

![Graph 2](image2)

Fig. 3.2. Results from the mesh convergence study for (a) the smallest ($h = 6$, $\beta = 60$) and (b) the largest ($h = 10$, $\beta = 20$) section types for the 2DFE analyses.
The sketch of a continuum model of a slope in 3D and the material body meshed is shown in Fig. 3.1 (d). A 10-noded tetrahedral element is used. The nodes in the mesh and the boundaries have three degrees of freedom, i.e., displacements \( u, v, \) and \( w, \) in the \( x, y, \) and \( z, \) directions, respectively. The boundary conditions are restrained displacement in the \( xyz, \) directions at the model ends of the planes the \( yz, xy, \) and \( xz (u = v = w = 0), \) assuming a rigid contact with no possibility of movement at the end of the boundaries. Preliminary evaluating extended cross-section models (6H) verify this assumption. The results indicated no differences between the critical SRFs from both 3H and 6H cross-sections. Therefore, the boundaries at 3H cross sections are sufficiently far that the prescribed boundary conditions do not affect the \( yz, \) and \( xz \) planes (Chugh 2003).

For the \( y, \) direction, the extruded face, findings from the literature from parametric studies for 3D unreinforced and pile-reinforced slopes suggested that the failure mechanism of 3D slopes is greatly affected by boundary conditions at those faces. Hence, to achieve a solution close to the plain strain solution, the ratio width/total height should be between 5 to 10 (Chugh (2003), Zhang et al. (2011), Gao et al. (2015)). The ratio used in this study is 5. As in the 2DFE analyses, for the 3D analyses, the mesh is adjusted for each slope geometry. The results from the mesh convergence analysis for the largest \( (h = 10, \beta = 20), \) and smallest section types \( (h = 6, \beta = 60), \) are shown in Fig. 3.3.

The number of elements used varies from 57,000-100,000, as the area of the extruded sections varies. For the uncoupled analyses, LE analyses, (2DLE, and 3DLE), are combined with a laterally loaded pile analysis using RSPile software (Rocscience Inc.). The boundaries are placed sufficiently far from the region where slope failure is expected to occur. For the 2DLE condition, the Spencer method with a non-circular mode of failure and auto refine search is used.
For the 3DLE condition, a ratio width/total height of 5 is also used. The Spencer method with an ellipsoid surface and a cuckoo search is used. To design the resisting pile, the lateral displacement is unconstrained until ultimate lateral resistance or soil cutoff is reached. To represent the non-linear soil-pile interaction, default $p$-$y$ curves corresponding to soft clay, stiff clay with free water, API Method for sand, and silt cemented ($c'$ and $\varphi'$) are used as the strength factors vary among combinations (Matlock (1970), Reese and Welch (1975), Reese et al. (1974)).

For both coupled and uncoupled methods, piles are modeled as structural elements of reinforced concrete with a fixed elastic modulus (20GPa), Poisson’s ratio (0.15), and steel reinforcement.

![Graph](image)

Fig. 3. Results from mesh convergence study for (a) the smallest ($h = 6, \beta = 60$) and (b) the largest ($h = 10, \beta = 20$) section types for the 3DFE analyses.
3.2.2 Statistical Methods

Several DoE techniques are used to generate synthetic data and as a data evaluation tool. In the context of this study, synthetic data refers to data generated artificially by combining slope properties to perform LE and FE analyses. DoE was first applied by Ronald Fisher for conducting agricultural research around the 1920s (Durakovic 2017). The initial aim of the application of DoE was to understand and to control the variability of the agricultural data gathered over more than 70 years (Niedz and Evens 2016). Because of its multipurpose nature, DoE has been used in various conditions such as comparative designs, variable screening, transfer function identification, optimization, and robust design.

The methodology is currently applied in scientific research, mainly in the biomedical, engineering, and biochemistry fields (Durakovic 2017). In geotechnical engineering, several probabilistic concepts are used to achieve rational risk assessment. However, the systematic application of DoE is still developing (Lye and Science, 2002; Lye, 2017). In the slope stability context, response surface theory to provide faster computational approximation (Wong 1985) and to evaluate the critical factors of slope instability mechanism of unreinforced slopes (Kostić et al. 2016; Jin-kui and Wei-wei 2018) have been applied. The authors have not found references to the application of DoE to the study of pile-reinforced slopes.

3.2.2.1 Factors Screening

The input factors (i.e., variables) along with their levels are presented in Table 1. Each level represents changes in the slope properties and the pile factors in this study. Statistically, + level represents the upper boundary of the factor’s value. The – level is the lower boundary, and the 0 represents the midpoint. Spacing is accounted for in the 2DLE analyses because the software used allows for the specification of out-of-plane spacing. For the 2DFE analyses, the pile properties (area and moment of inertia) are divided by spacing. For the condition $c' = 0$, the
cohesion is eliminated from the analyses. For $\phi' = 0$, the friction angle, and the pore water pressure coefficients are eliminated, and cohesion corresponds to the undrained cohesion.

A preliminary evaluation is carried out by generating a definitive screening design (DSD). DSDs are a tool from DoE, introduced by Jones and Nachtsheim (2011). This tree-level screening design is oriented to estimate unbiased main effects (effects from the factors) by the presence of active two-factor interactions (effects by a combination of factors). The response for this design follows the normal theory of linear model as presented by Jones and Nachtsheim (2011),

$$y_i = \beta_0 + \sum_{j=1}^{m} \beta_j x_i + \sum_{j=1}^{m-1} \sum_{k=j+1}^{m} \beta_{jk} x_{i,j,k} + \epsilon_i, \quad i = 1, \ldots, 2m + 1$$

where the parameters $\beta_0, \ldots, \beta_{mm}$ are unknown constants, and the $\epsilon_i$’s are independent and identically normally distributed random variables with mean 0 and variance $\sigma^2$, i.e., $N(0, \sigma^2)$.

In matrix form, we have $Y = X\beta + \epsilon$, where $X$ is the design matrix representing the inputs. The least-squares estimate of a parameter $\beta_j$ is denoted $\hat{\beta}_j$. To generate and analyze the DSD, the statistical software MINITAB® (Minitab 2019) is used.

The combinations for the DSD are presented in Table 3.2. A run is a combination of slope and pile properties. For instance, 6m slope, inclined at 30 degrees with a friction angle of 20 degrees, a cohesion of 35kPa, $E = 40$Mpa, and $\nu = 0.35$, pile located at 1 (the top), embedment depth is half the slope $H$, and diameter is 0.5, is a run. The run order is fully randomized. Following the results of the preliminary evaluation (presented in the Results section), soil elastic modulus ($E$) and the Poisson’s ratio ($\nu$) are considered constant values, and a more detailed evaluation oriented to the analysis of the effects on $FS$ of the pile-specific factors
is conducted. This analysis is done by generating a face-centered central composite design. The central composite design is a three-level response surface methodology developed by Box and collaborators in the 1950s (Bezerra et al. 2008). The method is based on mathematical and statistical techniques for fitting empirical and semi-empirical model structures to experimental data. In this study, a central composite design is used to create an approximating function simpler than running the numerical models from FE analyses. If the true response from slope stability mechanism is a function \(y(x_1, x_2, \ldots x_{10})\), the approximated response is given by \(\hat{y}(x_1, x_2, \ldots x_{10})\) (Wong 1985). The mathematical expression for symmetrical response surface designs is given by Eqs. 4 to 6:

\[
y = \beta_0 + \sum_{i=1}^{10} \beta_i x_i + \sum_{i=1}^{10} \sum_{j=1}^{9} \beta_{i,j} x_i x_j + \varepsilon
\]  

(3.4)

where

\[
\sum_{i=1}^{10} \beta_i x_i = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_{10} x_{10}
\]  

(3.5)

\[
\sum_{i=1}^{10} \sum_{j=1}^{9} \beta_{i,j} x_i x_j = \beta_{1,2} x_1 x_2 + \beta_{1,3} x_1 x_3 + \cdots + \beta_{10,9} x_1 x_{10}
\]  

(3.6)

and \(y\) is the response of the output variable in the input space of the \(x\)'s, \(\beta_0\) is the constant term, \(\beta_i\) represents the linear coefficients for the \(x_i\)'s, \(\beta_{i,j}\) is associated with the \(x_i x_j\) interaction. In this study, sources of errors are round-up errors, data entry, and lack of fit. The design table for the central composite design is provided as requested information due to the extent of the table.

To determine whether or not the two models are useful, a model utility test is conducted at the significance level \(\alpha = 0.05\). Non-significant factors are removed by a model discrimination technique to retain in the model only the significant coefficients.
Table 3.2. Randomized design table for DSD generated by MINITAB®.

<table>
<thead>
<tr>
<th>Run</th>
<th>Factors</th>
<th>β</th>
<th>H</th>
<th>φ'</th>
<th>c'</th>
<th>γ</th>
<th>E</th>
<th>v</th>
<th>PL</th>
<th>ED</th>
<th>Dia.</th>
<th>S</th>
<th>ru</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>0</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>+</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>0</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>24</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>

The discrimination technique is a stepwise regression algorithm. The null hypothesis and the alternative hypothesis for this test are given, respectively, by:

\[ H_0 = \beta_1 = \beta_2 = \cdots = \beta_k = 0; \quad (3.7) \]

\[ H_1: \beta_i \neq 0, \text{ for at least one } i \neq 0 \quad (3.8) \]

The test statistic is

\[ F = \frac{SSR/k}{SSE/[n - (k + 1)]}, \quad (3.9) \]

where \( k \) is the number of carriers (i.e., linear and interaction terms), \( SSE \) and \( SSR \) are error and regression sum of squares, respectively. This test rejects \( H_0 \) when the \( P \)-value < \( \alpha = 0.05 \).
3.2.2.2 Randomized Complete Block Design

After the study of the factors, a randomized complete block design (RCBD) is conducted. For this design, each block receives the treatment once. In this study, each slope configuration is treated as a block (run is a blocking factor), analytical method is the factor, each analytical method (FE/FD in either 2D/3D space) is a “treatment,” and the response variable is $FS$. The design configuration is presented in Table 3.3. JMP Pro 14 (2019) is used for the analysis.

Table 3.3. Randomized complete block design configuration.

<table>
<thead>
<tr>
<th>Block factor</th>
<th>Analytical Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Runs (Ri)</td>
<td>Treatment 1 2DLE</td>
</tr>
</tbody>
</table>

The use of this design allows comparing treatment effects on the response while controlling the impact of the blocking factor (slope property changes). The model is given as:

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}, \quad i = 1, \ldots, a, j = 1, \ldots, b$$

where $y_{ij}$ is the response for $i^{th}$ treatment in the $j^{th}$ block, $\mu$ is the overall mean, $\tau_i$ the effect ($\mu_i - \mu$) of the $i^{th}$ treatment, $\beta_j$ is the effect of the $j^{th}$ block, $a$ is the number of treatments (4), and $b$ is the number of blocks. The model is structured to achieve unbiased estimators for the model coefficients by using the restrictions $\sum \tau_i = 0$ and $\sum \beta_j = 0$.

The objective is to determine whether or not the mean $FS$ for different analytical methods are significantly different. The number of observations (the number of runs = number of blocks) is determined so that the following sizes of treatment effect ($\tau_i$) can be detected while controlling for type I and type II errors.
For this study with $\alpha = 0.05$ and $\beta = 0.05$ (95% power) and,

$$
\tau_1 = -\frac{\sigma}{2}, \tau_2 = -\frac{\sigma}{2}, \tau_3 = \frac{\sigma}{2}, \tau_4 = \frac{\sigma}{2},
$$

(3.11)

$$\Delta = \sigma = \sqrt{\frac{2}{\sigma^2}} \left( (-\frac{\sigma}{4})^2 + (-\frac{\sigma}{4})^2 + (\frac{\sigma}{4})^2 + (\frac{\sigma}{4})^2 \right) = 1.4
$$

(3.12)

where $\sigma$ is standard deviation and $\tau_i$ is $i^{th}$ treatment effect of the factor. The minimum number of observations is 23; the number used is 25. The significant differences among the treatment means are determined by applying Tukey’s honest significant difference (HSD) method as follows (Montgomery and Runger 2007):

$$
|y_i - y_j| > HSD
$$

(3.13)

$$
HSD = qa_{a,(a-1)(b-1)} \sqrt{\frac{MSE}{b}}
$$

(3.14)

where $q_{a,(a-1)(b-1)}$ is the 100(1 - $\alpha$)th percentile of the studentized range distribution, $a$ is the number of treatments (4 analytical methods), $b$ is the number of blocks (25 runs), and MSE is mean squared error.

Two mean differences, $\mu_i - \mu_j$, corresponding to $y_i - y_j$, are concluded to be significant if Eq. 13 is met. 100(1 - $\alpha$)% confidence intervals for $\mu_i - \mu_j$ are provided in the results section.

$$
y_1 - y_2 \pm t_{0.05, (\frac{a-1}{2} (a-1) (b-1))} \sqrt{\frac{2MSE}{b}}
$$

(3.15)
3.3 Results

3.3.1 Slope Stability Results

The slope stability analyses are conducted following the DSD and the central composite design for 2DLE, 2DFE, 3DLE, and 3DFE. An example of results compared in this study is presented in Fig. 3.4. The results correspond to Run 4 from Table 3.2.

Fig. 3.4. Example of the results compared in this study -- Run 4 from Table 3.2.

3.3.2 Significant Factors

A general classification of soils into groups is used. The classification corresponds to three main groups, purely cohesive (undrained analysis, \( c \)), purely frictional (\( \varphi' \)), and mixed soils (\( c' \) and \( \varphi' \)), the last type being the most common case. The preliminary analysis of the factors is performed separately for each soil group. After the assessment of the model adequacy (including normality, constant variance, independence), a summary of the average of all coefficients of determination \( R^2 \) (the proportion of explained variation) is presented in Table 3.4.
Table 3.4. Average R2 for the three soil groups.

<table>
<thead>
<tr>
<th>Model</th>
<th>*SE</th>
<th>$R^2$</th>
<th>$R^2$ (adj)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average for $c' = 0$</td>
<td>0.13</td>
<td>99.5%</td>
<td>98.4%</td>
</tr>
<tr>
<td>Average for $\varphi = 0$</td>
<td>0.22</td>
<td>97.04%</td>
<td>94.74%</td>
</tr>
<tr>
<td>Average $c'$ and $\varphi' = 0$</td>
<td>0.18</td>
<td>98.64%</td>
<td>97.29%</td>
</tr>
</tbody>
</table>

*SE = Standard error

The results are summarized in the Pareto charts presented in Fig. 3.5. A Pareto chart shows absolute values of standardized effects from largest to smallest. The reference line on the plot indicates the $t$-value (critical value) that tests the null hypothesis. Hence, the red line is a threshold to classify effects as significant or not, with the exception of the existence of two-factor interactions. If a two-factor interaction is concluded to be significant, then the corresponding factors are also concluded to be significant. For example, in Fig. 3.5 (a) the $c \times PL$ interaction is shown to be significant. Therefore, $c$ and $PL$ are significant (although the test for the main effect is not significant). The product of two factors represents two factor-interactions.

Non-significant terms are removed from the charts by the stepwise algorithm. For the three soil groups, at the significance level $\alpha = 0.05$, the factors with the larger effects on $FS$ are soil and geometry properties. Fig. 3.5 (a) indicates that in an undrained analysis, the pile location is the pile property found to be significant. For the purely frictional group, the results are presented in Fig. 3.5 (b). The embedment depth is the pile property that significantly affects the response. Fig. 3.5 (c) indicates that for the soil group where $c'$ and $\varphi'$ are considered for a drained analysis, the pile location, and embedment depth affect the response significantly.
Fig. 3. Pareto charts of standardized effects at $\alpha = 0.05$ (the red line is a threshold to classify effects as significant or not).

The preliminary analysis for all soil groups indicates that as previously presented by Griffiths and Lane (1999) in the analysis of unreinforced slopes, soil elastic modulus ($E$), and the Poisson’s ratio ($\nu$) do not affect $FS$ significantly. In other words, changes in both $E$ and $\nu$ need to
be considerably much larger to affect $FS$ significantly. The central composite design evaluation of the effects of pile properties is then performed, and $E$ and $v$ are held constant. Hence, only ten factors from Table 3.1 are evaluated for the $c'$ and $\varphi'$ soil group. $R^2$ is 0.99, and the SE is 0.15. The Pareto chart is presented in Fig. 3.6.

This analysis allows for a detailed evaluation of how the main effects and interactions affect $FS$. Results are in agreement with the preliminary analysis presented in Fig. 3.5 (c), where soil properties ($c'$, $\varphi'$, $\gamma$), and slope properties ($\beta$, $H$, $r_u$) have larger effects on the response. Nevertheless, it can be seen that all studied pile properties have some effect on $FS$. From these properties, pile location, embedment depth, and spacing appear to be the most significant.

For each soil group, the changes in $FS$ due to changes in the pile input factors are evaluated. Surfaces plots are generated considering $\Delta FS$ as the difference between the unreinforced $FS$ and the $FS$ after placing the pile. The surface plots for the purely cohesive and the purely fictional groups are presented in Fig. 3.7.

For an undrained analysis, Fig. 3.7 (a) indicates that for the lower level of cohesion (softer soils), all the locations of the pile within the slope yield the same $FS$. As the cohesion increases, $\Delta FS$ slightly increases when the pile is located from the toe towards the $\frac{3}{4}$ section of the slope. For the purely frictional group, Fig. 3.7 (b), all the locations of the pile within the slope yield similar $FS$, with only slightly higher $\Delta FS$ when the friction angle increases and the piles are located around the middle of the slope. The cohesive soils present deeper slip surfaces in which the pile around the toe represents the optimal condition. In contrast, the frictional materials and stiff clays (high cohesion values) exhibit shallower critical surfaces where $FS$ increases moving away from the toe. For both soil groups, the favorable location of the pile
within the slope matches the middle of the critical circle whose shape changes as the soil strength parameter vary.

Fig. 3.6. Pareto chart of standardized effects from central composite analysis, c’- φ’ soil group, at α = 0.05 (the red line is a threshold to classify effects as significant or not).
Fig. 3.7. Response surface plots of the pile factors' effects in ΔFS (a) purely cohesive group, (b) purely frictional group generated MATLAB (2019).

In addition, for both soil groups, increasing embedment depth and pile diameter increases $FS$. However, those effects over $FS$ are minor, as evidenced by the Pareto charts. The response surface plots for the $c' - \varphi'$ analysis are shown in Fig. 3.8. The results indicate that because of the interaction between $\beta$ and $PL$ shown in the Pareto chart (Fig. 3.6), the optimum pile location varies not only with variations in soil strength but also with $\beta$, as shown in Fig. 3.8 (c). The
lower the slope angle, the wider the slope, and for the range of input factors of this study, this generates surfaces whose middle coincide with the quarter to the middle of the slope, hence, placing the pile in those locations results in a larger $\Delta FS$. The steeper the slope, the optimum location moves from the middle to $\frac{3}{4}$ to the top of the slope, also depending on the slope material, as seen in Fig. 3.8 (a and b). Changes in $\Delta FS$ with $c'$ in the presence of $\phi'$ are shown in Fig. 3.8 (a), while changes in $\Delta FS$ with $\phi'$ in the presence of $c'$ are shown in Fig. 3.8 (b).

Increasing spacing, as shown in Fig. 3.8 (d), decreases $\Delta FS$.

Fig. 3.8. Response surface plots of the pile factors' effects in $\Delta FS$, $c'$- $\phi'$ analysis generated by MATLAB (2019).

The response surface plots presented herein are a schematic representation of the changes in $FS$ when holding some inputs constant and changing others. However, following the results from the Pareto charts, the relationship among the factors is dynamic. For optimal design, achieving the optimal combination of pile location, embedment depth, diameter, and spacing, all factors, and their interactions should be considered. In general, the pile location is the pile
property with the most significant interaction with the soil and slope geometry properties. Generally, in practice, for design purposes, a slope with specific soil type and geometry characteristics is designed to be reinforced with piles after $FS$ of the unreinforced condition needs to be increased. Hence, the optimization process using the results from this study can be used holding soil type constant and determining the combination of pile properties that yield the larger or target $FS$.

### 3.3.3 $FS$ Comparisons

The results from the DSD are presented in Table 3.5. These results are used for the $FS$ and $SRF$ comparisons. For the mean comparisons, the notation $FS$ is used for both $FS$ from LE and SRF from FE analyses. Model adequacy (including normality, constant variance, and independence) is verified. The analysis of variance table is presented in Table 3.6. The $R^2$ is 0.98. The results show that the $P$-value for method of analysis (i.e., 2DFE, 3DFE, and so on) is higher than $\alpha = 0.05$, indicating that the method of analysis is not statistically significant.

The larger variation on the model comes from “runs” (factor combinations), an indication of the stronger influence of slope properties on $FS$. $FS$ means for all the treatments are pairwise compared using Tukey’s method following equations (3.13) and (3.14). The results are presented in Table 3.7. They indicate that, at $\alpha = 0.05$, there is no significant difference between $FS$ means computed using the four evaluated methods (2DLE, 3DLE, 2DFE, 3DFE). The larger differences are found between 3DFE and both 2DFE and 2DLE with a 0.12 difference in $FS$. The $FS$ is close to $SRF$ from the coupled analysis. This may indicate that the current $p-y$ curves methodology used to introduce soil-pile interaction in a LE/lateral pile analysis captures the behavior of the studied soils. Although the results indicate no significant difference, the confidence intervals from Table 3.7 show differences in $FS$ as high as 0.27 among the methods.
Table 3.5. Summary FS for each run for each analytical method from DSD.

<table>
<thead>
<tr>
<th>Run</th>
<th>2DFE</th>
<th>3DFE</th>
<th>2DLE</th>
<th>3DLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.87</td>
<td>1.96</td>
<td>1.92</td>
<td>2.02</td>
</tr>
<tr>
<td>2</td>
<td>0.92</td>
<td>1.11</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>3</td>
<td>4.91</td>
<td>5.32</td>
<td>4.98</td>
<td>5.17</td>
</tr>
<tr>
<td>4</td>
<td>2.35</td>
<td>2.37</td>
<td>2.31</td>
<td>2.39</td>
</tr>
<tr>
<td>5</td>
<td>2.40</td>
<td>2.65</td>
<td>2.59</td>
<td>2.22</td>
</tr>
<tr>
<td>6</td>
<td>0.69</td>
<td>0.67</td>
<td>0.73</td>
<td>0.55</td>
</tr>
<tr>
<td>7</td>
<td>4.17</td>
<td>4.40</td>
<td>4.18</td>
<td>4.32</td>
</tr>
<tr>
<td>8</td>
<td>2.47</td>
<td>2.57</td>
<td>2.45</td>
<td>2.57</td>
</tr>
<tr>
<td>9</td>
<td>1.36</td>
<td>1.39</td>
<td>1.33</td>
<td>1.36</td>
</tr>
<tr>
<td>10</td>
<td>1.91</td>
<td>1.90</td>
<td>1.85</td>
<td>1.85</td>
</tr>
<tr>
<td>11</td>
<td>1.14</td>
<td>1.29</td>
<td>1.11</td>
<td>1.14</td>
</tr>
<tr>
<td>12</td>
<td>4.50</td>
<td>4.70</td>
<td>4.53</td>
<td>4.79</td>
</tr>
<tr>
<td>13</td>
<td>2.75</td>
<td>3.25</td>
<td>2.75</td>
<td>2.85</td>
</tr>
<tr>
<td>14</td>
<td>1.98</td>
<td>2.06</td>
<td>2.00</td>
<td>2.04</td>
</tr>
<tr>
<td>15</td>
<td>0.55</td>
<td>0.68</td>
<td>0.54</td>
<td>0.50</td>
</tr>
<tr>
<td>16</td>
<td>0.78</td>
<td>0.81</td>
<td>0.82</td>
<td>0.83</td>
</tr>
<tr>
<td>17</td>
<td>1.33</td>
<td>1.40</td>
<td>1.22</td>
<td>1.26</td>
</tr>
<tr>
<td>18</td>
<td>2.06</td>
<td>2.02</td>
<td>1.81</td>
<td>1.82</td>
</tr>
<tr>
<td>19</td>
<td>2.74</td>
<td>2.65</td>
<td>2.35</td>
<td>2.17</td>
</tr>
<tr>
<td>20</td>
<td>1.23</td>
<td>1.34</td>
<td>2.52</td>
<td>2.65</td>
</tr>
<tr>
<td>21</td>
<td>2.24</td>
<td>2.33</td>
<td>1.90</td>
<td>1.86</td>
</tr>
<tr>
<td>22</td>
<td>0.74</td>
<td>0.86</td>
<td>0.71</td>
<td>0.78</td>
</tr>
<tr>
<td>23</td>
<td>0.76</td>
<td>0.80</td>
<td>0.67</td>
<td>0.69</td>
</tr>
<tr>
<td>24</td>
<td>1.54</td>
<td>1.63</td>
<td>1.43</td>
<td>1.34</td>
</tr>
<tr>
<td>25</td>
<td>3.28</td>
<td>3.57</td>
<td>3.10</td>
<td>2.98</td>
</tr>
</tbody>
</table>

(Mean) $y_i$ 2.02 2.14 2.03 2.04

In practice, this difference may not be very significant if the slope is far from failure or far from stability. However, for slopes around the stability/failure threshold of one, this difference could be of key importance for decision-making.

Table 3.6. Analysis of Variance Table for the RCBD for all analyses.

<table>
<thead>
<tr>
<th>Source</th>
<th>*DF</th>
<th>SSE</th>
<th>F Ratio</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Runs</td>
<td>24</td>
<td>143.996</td>
<td>156.07</td>
<td>&lt;.0001*</td>
</tr>
<tr>
<td>Method</td>
<td>3</td>
<td>0.253</td>
<td>2.197</td>
<td>0.096</td>
</tr>
<tr>
<td>Error</td>
<td>72</td>
<td>2.767</td>
<td>0.038</td>
<td></td>
</tr>
<tr>
<td>C. Total</td>
<td>99</td>
<td>147.0.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Degrees of freedom
The mean difference between 2DLE and 2DFE is virtually zero (0.003 difference in \( FS \)), and it is smaller than the difference in \( FS \) between 2DLE and 3DLE (0.014). This may indicate that dimensionality (performing a 3D rather than a 2D analysis) causes slightly more differences within the same method (e.g., 2DFE and 3DFE) than the variability related to the assumptions from the methods (e.g., 2DLE and 2DFE).

Table 3. 7. Comparison means FS analytical methods by Tukey’s method for all analyses.

<table>
<thead>
<tr>
<th>Level - Level</th>
<th>( y_i - y_j )</th>
<th>( *SE_{diff} )</th>
<th>Lower CL</th>
<th>Upper CL</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>3DFE 2DFE</td>
<td>0.122</td>
<td>0.055</td>
<td>-0.023</td>
<td>0.268</td>
<td>0.131</td>
</tr>
<tr>
<td>3DFE 2DLE</td>
<td>0.118</td>
<td>0.055</td>
<td>-0.027</td>
<td>0.264</td>
<td>0.149</td>
</tr>
<tr>
<td>3DFE 3DLE</td>
<td>0.104</td>
<td>0.055</td>
<td>-0.041</td>
<td>0.250</td>
<td>0.244</td>
</tr>
<tr>
<td>3DLE 2DFE</td>
<td>0.018</td>
<td>0.055</td>
<td>-0.127</td>
<td>0.163</td>
<td>0.988</td>
</tr>
<tr>
<td>3DLE 2DLE</td>
<td>0.014</td>
<td>0.055</td>
<td>-0.131</td>
<td>0.160</td>
<td>0.993</td>
</tr>
<tr>
<td>2DLE 2DFE</td>
<td>0.003</td>
<td>0.055</td>
<td>-0.142</td>
<td>0.149</td>
<td>0.999</td>
</tr>
</tbody>
</table>

\( *SE_{diff} \) = Standard error of the difference

Moreover, the differences in \( FS \) due to method of analysis seen to increase in the presence of the third dimension on the analyses (e.g., 2DLE and 2DFE difference is significantly smaller than 3DFE and 3DLE). This could be related to the higher dependency of 3D analysis to boundary conditions, side forced assumptions, and the difficulty of mesh convergence for large extruded areas as the computational time increases considerably, for 3DFE analyses. However, this conclusion is an assumption that may need to be further evaluated.

3.4 Conclusions and Recommendations

In this paper, a comprehensive statistical analysis of the factor of safety (\( FS \)) for pile-reinforced slopes is conducted. The methodology is based on mathematical and statistical techniques for fitting empirical model structures to synthetic data. The applied statistical tools provided formal insights on the \( FS \) (sensitivity to inputs and method of analysis) of pile-reinforced slopes limited to the assumptions, boundary conditions, and considered factors of this study.
Nevertheless, the methodology presented herein can be easily replicated and applied to different types of slopes, and the study of a different or broader range of factors. From the results, several conclusions and recommendations can be drawn:

- The main factors influencing $FS$ of a pile-reinforced slope are soil and geometry properties. By performing separate analysis as purely cohesive soils ($c$, undrained analysis), purely frictional ($\phi'$), and mixed soils ($c'-\phi'$), the main effects and interactions of the input related to the pile placement into the slope can be studied.
- For mixed soils ($\phi'$-$c'$), pile location, embedment depth, and spacing are the pile properties with significant effects on $FS$. The relationship pile location-$FS$ is modified by cohesion, friction angle, the slope angle, embedment depth, spacing, diameter, unit weight, pore water pressure, and slope height. The variation of these factors and their interactions modify the optimum pile location ranging from the middle to the top of the slope wherever the middle of the critical surface is located within the slope.
- For purely cohesive soils ($c$, undrained analysis), the optimum pile location is found around the toe, and as cohesion increases (the soil becomes stiffer), the optimum location moves away from the toe towards the $3/4$ point of the slope. However, the improvement in the $FS$ with changing the location of the pile is small.
- For purely frictional soils ($\phi'$), the pile location does not significantly change $FS$. Only slightly larger $FS$ is found around the middle to $3/4$ of the slope, but the changes in $FS$ are small.
For all the soil groups, if the piles are located in a favorable location, increasing the diameter and the embedment depth and decreasing spacing increases $FS$. However, for spacing selection, the decision comes to whether the pile row is intended to work as a group, single piles, or a pile wall (for close piles).

For designing a pile-reinforced slope, it is recommended to perform an interactive analysis accounting for the main effects and interactions of the factors. A preliminary design could be achieved following results from the Pareto, and the surface plots presented. An optimal combination of pile location, embedment depth, diameter, and spacing, while fixing the slope properties, can be analyzed by applying the basics of statistical design of experiments/regression techniques presented herein.

The comparison of $FS$ mean values from 2D and 3D, limit equilibrium (LE), and finite element (FE) analyses indicate no statistical difference at the 0.05 significance level, for any of the studied methods. However, the confidence intervals show differences in $FS$ as high as 0.27, a variation of crucial importance for practical decision-making when the $FS$ of the slope is around the stability/failure threshold of one.

The comparative analysis also shows that dimensionality (performing a 3D rather than a 2D analysis) causes slightly more differences in $FS$ within the same method (e.g., 2DFE and 3DFE) than the variability related to the assumptions from the methods (e.g., 2DLE versus 2DFE). In addition, the differences in $FS$ due to method of analysis seen to increase in the presence of the third dimension on the analyses.
All in all, the current analytical methods are powerful estimating tools for $FS$ of pile-reinforced slopes. However, caution should be taken when performing slope stability analyses. The agreement between analytical methods does not guarantee a strong agreement with the field condition. A combination of techniques and engineering judgment must be exercised.

References


CHAPTER 4. FACTOR OF SAFETY PREDICTION TOOL: UNREINFORCED SLOPES

Yuderka Trinidad-Gonzalez¹, Vernon R. Schaefer², and Derrick K Rollins³

Modified from a paper under review by the Geomechanics and Geoengineering Journal

Abstract

An analytical model for the prediction of the factor of safety (FS) of unreinforced slopes as a function of soil and geometry properties is presented. The proposed model is developed by combining statistical design of experiments (DOE), artificial neural network (ANN), and limit equilibrium (LE) analysis to generate suitable combinations of input factors and for data analysis. The aim of the proposed model is a simplification of performing slope stability analysis without the need for special software while providing the reliability of the prediction tool when compared to the field condition. The results indicate superior performance and higher accuracy in predicting FS; more specifically, the model gives a 93% accuracy using a 1.20 FS threshold and 100% accuracy using a 1.30 FS threshold for the studied cases. The proposed model is provided as an Excel spreadsheet algorithm to facilitate its use as a preliminary estimation tool.

Keywords: Algorithms, data processing, factor of safety, slope stability, artificial neural networks

¹Graduate Research Assistant, Department of Civil, Construction and Environmental Engineering, Iowa State University, Ames, IA 50011 yuderkat@iastate.edu, Corresponding author
²James M. Hoover Professor of Geotechnical Engineering, Department of Civil, Construction and Environmental Engineering, Iowa State University, Ames, IA 50011, vern@iastate.edu
³Professor, Department of Chemical and Biological Engineering, Department of Statistics Iowa State University, Ames, IA 50011, drollins@iastate.edu
4.1 Introduction

In many areas, slope instability is a significant threat disrupting infrastructures, causing casualties and economic losses. Therefore, in the geotechnical field, slope stability is a crucial analysis conducted to verify the safety of cuts, embankments, and/or natural slopes against failure. In this regard, estimating the factor of safety (FS) of a slope is one of the main tasks when performing slope stability analyses.

The current methods of analysis to perform slope stability include limit equilibrium (LE) and numerical methods (generally, finite element (FE) and finite difference (FD)). Several authors such as Memon (2018), Trinidad González (2017), Nian et al. (2008), Griffiths and Lane (1999), Duncan (1998), Christian (1998), Duncan (1996), and Cheng et al. (2007) have presented the basics of the analytical techniques and some comparisons among them. However, the use of special software is required to perform such analyses. For this reason, authors such as Jin-Kui and Wei-Wei 2018; Kostić et al. (2016); Liu et al. 2014; Erzin and Cetin 2013; Das et al. 2011; Samui and Kothari 2011; Zhao 2008 have implemented statistical tools, support vector machines, and genetic algorithms to develop mathematical expressions for the FS. These models aim to develop reliable estimation tools with a simpler application than traditional analytical methods.

Kostić et al. (2016) generated synthetic information from applying response surface techniques as a data generator and a solution using the LE approach (Spencer method). The results indicated that this model had higher prediction power (at that time) than the models generated from field information. However, the model space was very limited and constrained to slopes with heights between 6m-10m. The models developed with ANN methods used a database from the literature of slopes around the world gathered by Sah et al. (1994). These models presented by Sah et al. (1994), Sakellariou and Ferentinou (2005), Yang et al. (2004), Ahangar-Asr et al. (2010), Samui and Kothari (2011), and Manouchehrian et al. (2014), are limited to; (1)
the accuracy of information used for their development (the data has high variability); (2) the slope condition, mainly failed dry slopes, which limits the mechanism of failure to those types of slopes; and (3) the testing set is mainly from the same dataset.

The objectives of this study are: (1) to develop a prediction model for FS of unreinforced slopes with prediction power as high as the one presented by Kostić et al. (2016), but with an augmented input space to overcome, the constraints of the existing model; (2) to provide the correlation between the model and the field condition, which was not previously done for models developed with synthetic data; and (3) facilitate the model as a prediction tool, coded in Excel, for easy use as a preliminary estimation tool.

This manuscript is organized as follows. Applied methods are briefly described in the next section. The performance of the model, applications, and comparisons are presented in results and discussions. A general review of the findings and conclusions are given in the final section.

4.2 Applied Methods

In this work, statistical design of experiment (DOE), artificial neural networks (ANN), and limit equilibrium (LE) analysis are combined to generate a suitable combination of input factors and for data analysis. A multifactor (i.e., input), full factorial design is used as a synthetic data generator (combining slope geometry and soil properties). Six factors are selected for the analysis. The multifactor, full factorial, design generated 4,032 combinations of slope and soil properties in a complex, highly dimensional input space. Each factor and its levels (e.g., L1 is Level 1) are given in Table 4.1.
The use of a full factorial design allows for the investigation of all main effects and all orders of interaction effects of the factors. This is useful to study the effects of factors on the response at several levels without confounding (note that, confounding means perfectly correlated and indistinguishable) effects and limiting conclusions to a subset of factors and their interactions. The results of the 4,032 \((7 \times 4^3 \times 3^2)\) input combinations are then used to perform LE analyses to determine the FS for each condition. For the LE analyses, Slide 2018 (Rocscience Inc. 2018) and the Spencer method are used. The Spencer method satisfies all the requirements for static equilibrium (Duncan et al. 2014). Non-circular mode of failure with auto refine search is used as a search method. The boundary conditions are set sufficiently far not to influence the FS. A sketch of the section of a slope with the input properties for the LE analyses is shown in Fig. 4.1.

![Fig. 4.1: Graphical representation of inputs for LE analyses.](image-url)
4.3 Prediction Model

Once the slope stability analyses are conducted, a distribution analysis, for the response, as a pre-processing step, is conducted. The factors of safety from the slope stability analyses are the output, i.e., response $\hat{y}$ (in this study, the computer-simulated response). Fig. 4.2 indicates that the $FS$ closely follows an exponential distribution with parameters $\lambda$ of 1.54 and a scale $\sigma$ (standard deviation) of 1.54. Before generating the ANN, model normalization is done by a Johnson Su distribution, using maximum likelihood to transform the inputs closer to normality, thereby mitigating the input of outliers and skewed distributions.

![Factor of safety, (FS)](image)

Fig. 4. 2. Computer simulated FS distribution.

After pre-processing, the ANN prediction model is generated. ANN incorporates a large class of model structures and learning methods. ANN is a two-staged non-linear regression or classification model structure consisting of an input layer (the properties or variables), the hidden layers (or hidden units that relate the input and output from a linear combination of the inputs) and, the output layer (the fitted response or responses of the system) (Hastie et al. 2009). The ANN used in this study is a fully connected, multiple-layer perceptron with an input layer of six neurons, a hidden layer of eight neurons, and an output layer of one neuron. The fitting criterion is based on a penalizing maximum likelihood approach for the estimation of the model.
parameters. The Laplacian likelihood function is used, and a robust fitting is enabled to reduce the influence of outliers in the response variable. A cross-validation technique is used to test the effectiveness of the model and avoid overfitting (Ghasemi et al. 2019; Refaeilzadeh et al. 2009). The holdout procedure is applied to determine which data are randomly assigned to either training or validation sets. For the data classification, 1/3 of the dataset is assigned to the validation set. A visual representation of the ANN applied in this study is presented in Fig. 4.3.

Fig. 4.3. Schematic illustration of ANN structure using JMP pro 14 (2019).

The mathematical representation of the model is given by matrices that represent the inputs ($X$), an additional node representing the bias (not shown in Fig. 4.3) is added here, and the hidden layer nodes or neurons ($A$). The arrows represent the weights ($W$) as the contribution of the inputs to the activation nodes. $\beta_0$ is the intercept representing the bias at the output layer ($O$). Each input element, $x_i$, is multiplied by a connection weight $w_{ij}$. The weighted input signals are summed, and $b_A$, a bias value at the nodes, is added. The combined $j^{th}$ input is passed through the nonlinear activation function, $f(j)$, to produce the output of the hidden layer, $y_A$, or input of the
next layer. The weights are adjusted using $w_{Oj}$ by learning using the training data set. The sum of the adjusted weights is then added to the bias of the output layer $\beta_0$. A quasi-Newton method, BFGS (Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm) iterates for optimization of the penalty parameter in the training stage. Simultaneously, the BFGS algorithm monitors the likelihood function of the validation set (an independent set that is used to check model performance Ghasemi et al. 2019; Cheng and Titterington 1994). When the cross-validation no longer improves, the early stopping rule terminates the iteration. A general form of the model equation can be written as:

$$\hat{y} = \beta_0 + \sum_{j=1}^{n} \left[ w_{Oj} \cdot f_j \left( b_{Aj} + \sum_{i=1}^{n} w_{ij} x_i \right) \right]$$

(4.1)

where $\beta_0$ is the bias of the output layer, $w_{Oj}$ is the connection weight between neuron $j$ of the hidden layer ($j = 1$ to 8) and the output layer; $b_{Aj}$ is the bias at neuron $j$ of the hidden layer; $w_{ij}$ is the connection weights between the input variable $i$ (for $i = 1$ to 6) and neuron $j$ of the hidden layer; $x_i$ is input $i$, $f_j$ is the activation function at the hidden layer. The activation function used in this study is the identity transformation (Gaussian) that for the variable $j$ (neurons ($j = 1$ to 8)) is defined as:

$$f(j) = e^{(b_{Aj} + \sum_{i=1}^{n} w_{ij} x_i)}$$

(4.2)

4.4 Results and Discussions

4.4.1 Model performance

The general components of Eq. 1 (the estimates) are given as follows:

$$\hat{y} = 1.71 + 7.19x_{f_1} + 40.78x_{f_2} + -3.41x_{f_3} + 4.075x_{f_4}$$

$$+ 1.25x_{f_5} + -14.72x_{f_6} + 10.73x_{f_7} + 4.63x_{f_8}$$

(4.3)
The performance results of the final model are given based on the following statistics and presented in Table 2. The “average difference (AD)” as the estimate of the model bias

\[ AD = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i) \]  

(4.4)

where \( n \) is the number of input vectors, \( y_i \) is the \( i \)th measured response value, and \( \hat{y}_i \) is the \( i \)th fitted response value. The “average absolute difference (AAD)” as the averaged absolute distance between fitted and measured values, defined as

\[ AAD = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \]  

(4.5)

The correlation between the measured and fitted values, \( r_{fit} \), is defined as

\[ r_{fit} = \frac{n \sum_{i=1}^{n} y_i \hat{y}_i - (\sum_{i=1}^{n} y_i)(\sum_{i=1}^{n} \hat{y}_i)}{\sqrt{n \sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2} \sqrt{n \sum_{i=1}^{n} \hat{y}_i^2 - (\sum_{i=1}^{n} \hat{y}_i)^2}} \]  

(4.6)

The closer \( r_{fit} \) is to 1, the better the fit.

A generalized coefficient of determination, \( r_{Generalized}^2 \), that compares to a null log-likelihood that corresponds to a Laplacian log-likelihood using the sample median and mean absolute deviation as scaling parameters is given as

\[ r_{Generalized}^2 = 1 - e^{\frac{1}{n}(L_{\hat{\beta}} - L_{null})} \]

(4.7)

where \( L_{\hat{\beta}} \) is the negative loglikelihood of the set using the model parameters on the training data. According to the values presented in Table 4.2, the predicted values of FS have a high correlation with the measured values as well as low AD and AAD. Hence, the prediction tool modeled the response well.
Table 4.2. Summary of Statistics for performance evaluation.

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Training</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$AD$</td>
<td>0.007</td>
<td>0.01</td>
</tr>
<tr>
<td>$AAD$</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>$r_{fit}$</td>
<td>-</td>
<td>0.998</td>
</tr>
<tr>
<td>*$R^2$</td>
<td>0.997</td>
<td>-</td>
</tr>
</tbody>
</table>

*$R^2$ is given by Eq. 4.7.

The observed vs. predicted plots for both training and validation sets are presented in Fig. 4.4.

The input analysis indicated that the ones with the larger effect on the response are the inclination angle and the friction angle. These results are presented in Table 4.3. These results also indicate that interactions among the inputs also affect the response. This can be seen from the difference between the main effect and total effect in Table 4.3. The results are in partial agreement with the findings presented by Kostić et al. (2016) regarding the order of the factors affecting the slope stability mechanism.
Table 4.3. Assessment of variance importance (input’s effect over the response).

<table>
<thead>
<tr>
<th>Input</th>
<th>Main Effect</th>
<th>Total Effect</th>
<th>.2</th>
<th>.4</th>
<th>.6</th>
<th>.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.437</td>
<td>0.531</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi'$</td>
<td>0.115</td>
<td>0.214</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H (m)</td>
<td>0.131</td>
<td>0.178</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_u$</td>
<td>0.069</td>
<td>0.112</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c'$ (kPa)</td>
<td>0.057</td>
<td>0.101</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma$ (kN/m³)</td>
<td>0.01</td>
<td>0.024</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The minor discrepancies could be related to the wider ranges of the inputs used to develop the current model. However, an agreement is found pertaining to the large negative effect of the slope inclination angle for soils with $c'$ and $\phi'$. Besides, the effects of the unit weight of the soil from the literature are found to be similar to the lowest of the study factors affecting the response.

4.4.2 Application and Comparisons with Previous Models

Kostić et al. (2016) indicated that to demonstrate the need for a new prediction model, the prediction performance should be compared with the existing mathematical tools. In this study, the comparison follows (1) demonstration of superior prediction capabilities in a wider input space; and (2) demonstration of relatively higher accuracy when compared to the field condition.

The models presented by Kostić et al. (2016), Manouchehrian et al. (2014), Ahangar-Asr et al. (2010), Yang et al. (2004) are compared with the proposed model (TG Model). All but Kostić et al. (2016) models used the same data set of slopes presented by Sah et al. (1994). The Kostić et al. (2016) model was created with a synthetic data set. The selected models use the same number of inputs as the TG Model with different ranges. The constraints to avoid extrapolation are specified in Table 4.4.
Table 4.4. Range of inputs from the compared prediction models updated from Kostić et al. (2016).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>10-70</td>
<td>25-70</td>
<td>16-53</td>
<td>16-53</td>
<td>16-53</td>
</tr>
<tr>
<td>φ</td>
<td>10-46</td>
<td>10-50</td>
<td>0-45</td>
<td>0-45</td>
<td>0-45</td>
</tr>
<tr>
<td>H (m)</td>
<td>6-63</td>
<td>6-10</td>
<td>3.6-214</td>
<td>3.66-214</td>
<td>3.66-214</td>
</tr>
<tr>
<td>r_u</td>
<td>0-0.6</td>
<td>0-0.5</td>
<td>0-0.5</td>
<td>0-0.5</td>
<td>0.11-0.5</td>
</tr>
<tr>
<td>c (kPa)</td>
<td>5-50</td>
<td>0-50</td>
<td>0-50</td>
<td>0-150.05</td>
<td>0-150.05</td>
</tr>
<tr>
<td>γ (kN/m3)</td>
<td>12-26</td>
<td>16-20</td>
<td>12-28.44</td>
<td>12-28.44</td>
<td>12-28.44</td>
</tr>
</tbody>
</table>

Twenty random combinations of slopes are selected and modeled with LE analyses to determine the FS using the Spencer method. Table 4.5 shows the data for comparison. The summary of the performance from the comparison is presented in Table 4.6. Figure 4.5 summarizes the prediction accuracy of the models when compared to the analytical solution from LE analyses.

Table 4.5. Data for comparison proposed model “TG model” with existing models and LE results.

<table>
<thead>
<tr>
<th>H (m)</th>
<th>β</th>
<th>c (kPa)</th>
<th>Phi</th>
<th>r_u</th>
<th>γ (kN/m3)</th>
<th>FS (Spencer)</th>
<th>Model 2015</th>
<th>Model 2014</th>
<th>Model 2010</th>
<th>Model 2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.6</td>
<td>26</td>
<td>9</td>
<td>15</td>
<td>0.1</td>
<td>16.8</td>
<td>1.31</td>
<td>1.29</td>
<td>1.17</td>
<td>1.67</td>
<td>1.30</td>
</tr>
<tr>
<td>7.2</td>
<td>36</td>
<td>46</td>
<td>22</td>
<td>0.22</td>
<td>17.2</td>
<td>3.14</td>
<td>3.22</td>
<td>3.37</td>
<td>3.08</td>
<td>1.66</td>
</tr>
<tr>
<td>8.4</td>
<td>42</td>
<td>24</td>
<td>41</td>
<td>0.41</td>
<td>17</td>
<td>1.82</td>
<td>1.83</td>
<td>2.03</td>
<td>2.16</td>
<td>1.43</td>
</tr>
<tr>
<td>9.6</td>
<td>33</td>
<td>18</td>
<td>35</td>
<td>0.39</td>
<td>19.4</td>
<td>1.54</td>
<td>1.53</td>
<td>1.70</td>
<td>2.05</td>
<td>1.66</td>
</tr>
<tr>
<td>6.8</td>
<td>38</td>
<td>38</td>
<td>11</td>
<td>0.13</td>
<td>18.3</td>
<td>2.33</td>
<td>2.35</td>
<td>2.37</td>
<td>2.42</td>
<td>1.40</td>
</tr>
<tr>
<td>8.8</td>
<td>45</td>
<td>44</td>
<td>29</td>
<td>0.26</td>
<td>19.8</td>
<td>2.23</td>
<td>2.32</td>
<td>2.49</td>
<td>2.88</td>
<td>1.45</td>
</tr>
<tr>
<td>9.1</td>
<td>51</td>
<td>21</td>
<td>33</td>
<td>0.48</td>
<td>16.2</td>
<td>1.14</td>
<td>1.15</td>
<td>1.21</td>
<td>1.53</td>
<td>0.70</td>
</tr>
<tr>
<td>7</td>
<td>28</td>
<td>15</td>
<td>12</td>
<td>0.31</td>
<td>19.6</td>
<td>1.18</td>
<td>1.19</td>
<td>1.19</td>
<td>1.66</td>
<td>1.18</td>
</tr>
<tr>
<td>9</td>
<td>53</td>
<td>4</td>
<td>44</td>
<td>0.2</td>
<td>17.5</td>
<td>0.80</td>
<td>0.77</td>
<td>0.92</td>
<td>1.31</td>
<td>0.49</td>
</tr>
<tr>
<td>6.9</td>
<td>32</td>
<td>12</td>
<td>19</td>
<td>0.11</td>
<td>18.9</td>
<td>1.38</td>
<td>1.37</td>
<td>1.35</td>
<td>1.74</td>
<td>1.36</td>
</tr>
<tr>
<td>46</td>
<td>40</td>
<td>25</td>
<td>28</td>
<td>0.3</td>
<td>19</td>
<td>0.72</td>
<td>0.73</td>
<td>NA</td>
<td>1.84</td>
<td>1.24</td>
</tr>
<tr>
<td>46</td>
<td>40</td>
<td>25</td>
<td>28</td>
<td>0.3</td>
<td>19</td>
<td>0.72</td>
<td>0.73</td>
<td>NA</td>
<td>1.84</td>
<td>1.24</td>
</tr>
<tr>
<td>46</td>
<td>40</td>
<td>50</td>
<td>28</td>
<td>0.3</td>
<td>19</td>
<td>0.96</td>
<td>1.01</td>
<td>NA</td>
<td>2.88</td>
<td>1.50</td>
</tr>
<tr>
<td>46</td>
<td>40</td>
<td>25</td>
<td>28</td>
<td>0.3</td>
<td>26</td>
<td>0.64</td>
<td>0.66</td>
<td>NA</td>
<td>1.98</td>
<td>1.33</td>
</tr>
<tr>
<td>46</td>
<td>40</td>
<td>25</td>
<td>28</td>
<td>0.3</td>
<td>19</td>
<td>0.72</td>
<td>0.73</td>
<td>NA</td>
<td>1.84</td>
<td>1.24</td>
</tr>
<tr>
<td>46</td>
<td>70</td>
<td>25</td>
<td>28</td>
<td>0.3</td>
<td>19</td>
<td>0.20</td>
<td>0.20</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>46</td>
<td>40</td>
<td>25</td>
<td>28</td>
<td>0.3</td>
<td>19</td>
<td>0.72</td>
<td>0.73</td>
<td>NA</td>
<td>1.84</td>
<td>1.24</td>
</tr>
<tr>
<td>46</td>
<td>40</td>
<td>25</td>
<td>28</td>
<td>0.3</td>
<td>19</td>
<td>0.72</td>
<td>0.73</td>
<td>NA</td>
<td>1.84</td>
<td>1.24</td>
</tr>
<tr>
<td>46</td>
<td>10</td>
<td>25</td>
<td>28</td>
<td>0.3</td>
<td>19</td>
<td>2.65</td>
<td>2.61</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
Fig. 4.5. Measured vs. predicted FS for all the models in this study (testing).
Table 4.6. Summary of Statistics for performance evaluation, testing sets.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>0.62</td>
<td>0.63</td>
<td>0.48</td>
<td>0.49</td>
<td>0.50</td>
</tr>
<tr>
<td>AAD</td>
<td>0.34</td>
<td>0.37</td>
<td>0.42</td>
<td>0.34</td>
<td>0.26</td>
</tr>
<tr>
<td>$r_{fit}$</td>
<td>0.999</td>
<td>0.989</td>
<td>0.70</td>
<td>0.29</td>
<td>0.75</td>
</tr>
</tbody>
</table>

The results indicate that for the models created from the field database, the $r_{fit}$ value is considerably low compared to the models from synthetic database Kostić et al. (2016) and TG model. These results are also in agreement with the findings presented by Kostić et al. (2016). The low correlation between the models 2004, 2010, and 2014 with the testing set could be related to the high variability of the dataset used for the models’ development. Additionally, the majority of the slopes from the case histories are dry failed slopes. Hence, the mechanism of failure could be prone to that condition.

To determine the model accuracy compared to the field condition, 15 cases are selected from the ones presented by Sah et al. (1994). The cases are coding to zero when the slope is failed and one when the slope is stable to determine accuracy. A 1.20 And 1.30 are used as threshold FS between equilibrium and failure. Table 4.7 presents the information from case histories used for this assessment. The model presented by Kostić et al. (2016) is not evaluated because the majority of the slopes from the case histories are outside of the range of applicability of this model. Hence, even though the prediction capability of this model is high, it uses is very limited due to the narrow input space.

The results are summarized for thresholds FS of 1.2 and 1.3 and presented in Table 4.8. Where, TP is truly positive in this context, slopes classified as failed that are actually failed. TN is a true negative; in this context, slopes classified as stable that are actually stable.
FP is false positive; in this context, slopes classified as failed that are actually stable. FN is a false negative; in this context, slopes classified as stable that are actually failed. Accuracy is the overall probability that the slope is correctly classified.

\[ A_{cc} = \frac{TP + TN}{FP + TN + FN + TP} \]  

(4.8)

Table 4.7. Case histories used for the field accuracy assessment.

<table>
<thead>
<tr>
<th>H (m)</th>
<th>(\beta)</th>
<th>(c) (kPa)</th>
<th>(\phi)</th>
<th>(r_u)</th>
<th>(\gamma) (kN/m³)</th>
<th>Condition</th>
<th>TG Model 2014</th>
<th>Model 2010</th>
<th>Model 2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>30.5</td>
<td>20</td>
<td>14.36</td>
<td>25</td>
<td>0</td>
<td>18.84</td>
<td>Stable</td>
<td>1.74</td>
<td>2.07</td>
<td>1.72</td>
</tr>
<tr>
<td>30.5</td>
<td>20</td>
<td>57.46</td>
<td>20</td>
<td>0</td>
<td>18.84</td>
<td>Stable</td>
<td>2.09</td>
<td>3.74</td>
<td>1.98</td>
</tr>
<tr>
<td>40</td>
<td>30</td>
<td>16.28</td>
<td>26.5</td>
<td>0</td>
<td>20.6</td>
<td>Failed</td>
<td>0</td>
<td>1.16</td>
<td>1.83</td>
</tr>
<tr>
<td>10</td>
<td>30</td>
<td>10</td>
<td>35</td>
<td>0</td>
<td>22.4</td>
<td>Stable</td>
<td>1.86</td>
<td>2.01</td>
<td>1.85</td>
</tr>
<tr>
<td>20</td>
<td>30</td>
<td>10</td>
<td>30.34</td>
<td>0</td>
<td>21.4</td>
<td>Stable</td>
<td>1.4</td>
<td>1.78</td>
<td>1.67</td>
</tr>
<tr>
<td>50</td>
<td>45</td>
<td>20</td>
<td>36</td>
<td>0</td>
<td>22</td>
<td>Failed</td>
<td>0</td>
<td>1.08</td>
<td>1.76</td>
</tr>
<tr>
<td>10.67</td>
<td>22</td>
<td>24.9</td>
<td>13</td>
<td>0.35</td>
<td>20.41</td>
<td>Stable</td>
<td>1.41</td>
<td>2.17</td>
<td>1.35</td>
</tr>
<tr>
<td>12.19</td>
<td>22</td>
<td>11.97</td>
<td>20</td>
<td>0.41</td>
<td>19.63</td>
<td>Failed</td>
<td>0</td>
<td>1.06</td>
<td>1.72</td>
</tr>
<tr>
<td>10.67</td>
<td>25</td>
<td>15.32</td>
<td>30</td>
<td>0.38</td>
<td>18.84</td>
<td>Stable</td>
<td>1.57</td>
<td>2.06</td>
<td>1.71</td>
</tr>
<tr>
<td>50</td>
<td>45</td>
<td>20</td>
<td>36</td>
<td>0.25</td>
<td>20</td>
<td>Failed</td>
<td>0</td>
<td>0.75</td>
<td>1.7</td>
</tr>
<tr>
<td>12.8</td>
<td>28</td>
<td>8.62</td>
<td>32</td>
<td>0.49</td>
<td>21.82</td>
<td>Failed</td>
<td>0</td>
<td>0.85</td>
<td>1.49</td>
</tr>
<tr>
<td>50</td>
<td>45</td>
<td>20</td>
<td>36</td>
<td>0.5</td>
<td>20</td>
<td>Failed</td>
<td>0</td>
<td>0.39</td>
<td>1.31</td>
</tr>
<tr>
<td>30.5</td>
<td>20</td>
<td>14.36</td>
<td>25</td>
<td>0.45</td>
<td>18.84</td>
<td>Failed</td>
<td>0</td>
<td>1.02</td>
<td>1.75</td>
</tr>
<tr>
<td>20</td>
<td>45</td>
<td>24</td>
<td>30.15</td>
<td>0.12</td>
<td>18</td>
<td>Failed</td>
<td>0</td>
<td>1.27</td>
<td>1.97</td>
</tr>
<tr>
<td>10</td>
<td>45</td>
<td>10</td>
<td>35</td>
<td>0.4</td>
<td>22.4</td>
<td>Failed</td>
<td>0</td>
<td>0.7</td>
<td>1.42</td>
</tr>
</tbody>
</table>

The results from Table 8 indicate that the accuracy of the TG Model compares to using a LE software for performing slopes stability analyses for a 93% accuracy when compared to the field condition. If the threshold FS is increased to 1.3, the results give 100% accuracy for the cases analyzed in this study.

The methodology applied herein can be replicated to broaden the input space widening the applicability of this model. An enclosed solution limited to the presented case, homogenous, unreinforced slopes with the input space presented in Table 4 is facilitated in Excel for the user’s easy applicability (including constraints on the inputs to avoid extrapolations).
Table 4.8. Confusion matrix for accuracy assessment of prediction models.

<table>
<thead>
<tr>
<th></th>
<th>TH-1.2</th>
<th>TG Model</th>
<th>Model 2014</th>
<th>Model 2010</th>
<th>Model 2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP</td>
<td>8.00</td>
<td>0.00</td>
<td>3.00</td>
<td>4.00</td>
<td></td>
</tr>
<tr>
<td>TN</td>
<td>6.00</td>
<td>6.00</td>
<td>6.00</td>
<td>6.00</td>
<td></td>
</tr>
<tr>
<td>FP</td>
<td>0.00</td>
<td>9.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>FN</td>
<td>1.00</td>
<td>0.00</td>
<td>6.00</td>
<td>5.00</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.93</td>
<td>0.40</td>
<td>0.60</td>
<td>0.67</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>TH-1.3</th>
<th>TG Model</th>
<th>Model 2014</th>
<th>Model 2010</th>
<th>Model 2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP</td>
<td>9.00</td>
<td>0.00</td>
<td>5.00</td>
<td>4.00</td>
<td></td>
</tr>
<tr>
<td>TN</td>
<td>6.00</td>
<td>6.00</td>
<td>6.00</td>
<td>6.00</td>
<td></td>
</tr>
<tr>
<td>FP</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>FN</td>
<td>0.00</td>
<td>9.00</td>
<td>4.00</td>
<td>5.00</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>1.00</td>
<td>0.40</td>
<td>0.73</td>
<td>0.67</td>
<td></td>
</tr>
</tbody>
</table>

4.5 Conclusions

In this study, a prediction model for the factor of safety (FS) of homogeneous unreinforced soil slopes is developed by combining statistical design of experiment (DOE), artificial neural networks (ANN), and limit equilibrium (LE) analysis to generate suitable combinations of input factors and for data analysis. Six factors are selected for developing the proposed model (TG Model). The performance results indicated that the predicted values of FS have a high correlation with the measured values. Hence, the prediction tool modeled the response well.

Verification and comparison of the TG model with the existing prediction equations indicated that the TG model estimates the FS with much higher accuracy than the models 2004-2014 and slightly higher accuracy and broaden constraints that the 2016 model. Along with this, the proposed model has the highest accuracy to predict FS when compared to the field condition for a 93% accuracy for a 1.20 FS threshold and 100% for 1.30 FS for the studied cases. The results indicated that the developed model compares to the use of performing analysis in LE software without the need for special packages. The applicability of the TG model can be broadened by adding levels to the input factors (i.e., lower and higher heights, lower and higher friction and cohesion), or input factors for layered slopes. In that case, the TG model could be
used for layered problems and high volumes landslides. It should be noted that the proposed approach should be used as a preliminary prediction tool and for simplified approximations within the range of parameters for which it is developed. For ultimate decision-making, each slope case must be carefully analyzed, and a combination of methods and engineering judgment must be exercised.

References


Rocscience Inc. (2016). “Slide 7.0 2D limit equilibrium slope stability analysis program.”


CHAPTER 5. STATISTICAL INSIGHTS REGARDING FULLY SOFTENED SHEAR STRENGTH

Yuderkat Trinidad-Gonzalez¹, Vernon R. Schaefer², and Derrick K Rollins³

Modified from a paper under review by the Geotechnical and Geological Journal

Abstract

The fully softened shear strength (FSS) concept is a practical approximation of the mobilized drained shear strength of first-time slides in stiff-fissured clays. There has been a recent increase in interest in measurement and estimation of FSS to develop correlations for preliminary design and cost approximation. However, such correlations do not help in understanding the cause and effect relationship between soil properties and FSS. In this study, a laboratory database containing FSS values (output) and soil properties (inputs) of several overconsolidated clays is used to develop a predictive model for the FSS secant friction angle (response). The goal is to detect which inputs from the whole parameter space dominated the response while creating an accurate prediction tool to provide statistical insights regarding the FSS. The proposed methodology is used to assess and quantify the relationships among variables, estimate testing device effects on FSS, and analyze the danger of extrapolation due to model constraints. The applicability of predicted FSS is also evaluated, comparisons, and recommendations regarding the studied prediction tools for slope stability design in stiff-fissured clays are given.

Keywords: Fully softened shear strength. Slope stability. Stiff-fissured clays. Predictive model. Correlation analysis. Statistical insights

¹Graduate Research Assistant, Department of Civil, Construction and Environmental Engineering, Iowa State University, Ames, IA 50011 yuderkat@iastate.edu, Corresponding author
²James M. Hoover Professor of Geotechnical Engineering, Department of Civil, Construction and Environmental Engineering, Iowa State University, Ames, IA 50011, vern@iastate.edu
³Professor, Department of Chemical and Biological Engineering, Department of Statistics Iowa State University, Ames, IA 50011, drollins@iastate.edu
5.1 Introduction

Determination of soil parameters for conducting stability analysis of stiff-fissured clays has long been a significant challenge in the geotechnical field. Since Skempton's elucidation of the issues (1964, 1967, 1970, 1977), it has become well-established that the mobilized strength for first-time slides in stiff-fissured clays subjected to softening corresponds to strengths below the peak and above the residual. Skempton (1970) introduced the concept of fully softened strength (FSS), relating mobilized shear strength to shear strength measured in the laboratory by application of the critical state concept presented by Schofield & Wroth (1968). Skempton (1970) established that FSS parameters $c'_s$ (effective cohesion), and $\Phi'_s$ (effective friction angle) are numerically equal to peak-strength parameters of remolded normally consolidated clay (based on the Brown London clay test). Currently, FSS parameters are recognized as controlling shear strength for cuts and embankments in stiff-fissured clays, shales, and mudstones (James 1970; Stark and Eid 1997; Mesri and Shahien 2002; Castellanos 2013; Eid and Rabie 2016).

In recent years, there has been increased interest in measurement and estimation of FSS to develop correlations, useful as tools for preliminary design and cost estimation. Several authors have developed correlations of FSS with soil index properties such as liquid limit, clay fraction, and plasticity index, among others. The majority of this correlation are tools for indirect estimation of a secant FSS. Among these are the correlations presented by Castellanos et al. (2016) in which strength coefficients are related to clay fraction and plasticity index, using the power function suggested by Lade (2010) to characterize the failure envelope nonlinearity. Previous to that, Stark and Hussain (2012) and previous studies (Stark and Eid (1997), Eid (1996)) presented a correlation in which a secant FSS friction angle is correlated with liquid limit, clay fraction, and effective normal stress. Specimens were tested using a ring-shear device, with a correction factor of $+2.5^\circ$ introduced to convert ring shear FSS secant friction angles to
consolidated drained triaxial compression (CD-TX) FSS secant friction angles. Such a correction factor was introduced because, according to Stark and Eid (1997), the most relevant shear mode of failure in first-time slides in natural-cut slopes and compacted embankments is nearer to drained triaxial compression. Tiwari and Ajmera (2011) had presented a correlation of FSS with plasticity index, liquid limit, or clay-sized fraction. Kaya (2009) presented a prediction model developed using a back-propagation artificial neural network (ANN) based on laboratory data reported by Stark et al. (2005). The coefficient of determination was reported as $R^2 = 0.81$. Hence, performance measurements of the model related to a test dataset were not given since $R^2$ measures training and not testing data performance Wright (2005), using a dataset from Stark et al. (2005) for soils with a 50 percent or higher clay fraction, developed a correlation for FSS and the logarithm of liquid limit using the method of least squares. Saleh and Wright (1997) proposed a correlation based on liquid limit. From all these presented correlations, there is not a clear consensus as to which soil properties should be used to estimate FSS.

This study introduces a rigorous statistical approach to estimate the FSS secant friction angle as a mathematical mean of indirect estimation of FSS accounting for non-linearity of the failure envelop. For this purpose, first, Pearson correlation coefficients among the studied variables (inputs) are determined to identify multicollinearity issues and to reduce input space. Next, an artificial neural network (ANN) is used to provide statistical insights revealing the hidden relationships between soil index properties, effective stress, and FSS. A reduced prediction model based on variable importance assessment and the results from the input correlation analysis is then generated. This prediction model accounted for FSS differences resulting from different testing conditions. The constraints of both present and previous prediction tools are studied to avoid extrapolation. This new approach is presented using in the
following sequence: The next section first defines the proposed approach and discusses artificial neural-network concepts essential to its understanding in this application, with the prediction model, results, and discussion presented in separate sections. Finally, two case histories of slope stability analysis are presented to demonstrate the application of predicted FSS in slope stability analyses and to compare present and previous approaches. Concluding remarks are provided in the last section.

5.2 Materials and Methods

5.2.1 Data collection

Two hundred one laboratory strength test results from 61 different soils, i.e., soils from different locations, are collected from the literature, primarily from Stark and Hussain (2012); and Castellanos et al. (2013). Those FSS results had been determined using ring shear (RS) and direct shear (DS) devices. The procedure for determining FSS followed Skempton’s (1970) definition of the FSS, viz., “peak strength parameters of the remolded normally consolidated clay,” with each specimen prepared by soaking, mixing in a blender, and/or ball-milling to ensure disaggregation of the clay particles. These preparation methods were intended to simulate the natural processes that affect soil and lead to a reduction in its strength. Summary statistics for index properties and testing confining stresses are presented in Table 5.1. LL, PL, PI refer to liquid limit, plastic limit, and plasticity index as defined by ASTM D4318 (ASTM International 2017). CL and A are clay fraction and activity, respectively. The clay fraction is defined as the fraction of soil clay-sized <0.002mm. Activity is the ratio of PI to the CF. The testing device input is D (where 1 is DS test and 2 is RS), and the effective normal stress is $\sigma_n'$. All the inputs are used to determine the best set of inputs for predicting the secant FSS.
Table 5.1. Summary statistics of the range in input properties of the dataset used in this study.

<table>
<thead>
<tr>
<th>Property</th>
<th>Clay fraction (CF) (%)</th>
<th>Liquid Limit (LL) (%)</th>
<th>Plastic Limit (PL) (%)</th>
<th>Plasticity Index (PI) (%)</th>
<th>Activity (A)</th>
<th>Effective stress ($\sigma'$) (kPa)</th>
<th>FSS ($^\circ$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min-Max</td>
<td>10-88</td>
<td>20-288</td>
<td>13-55</td>
<td>4-244</td>
<td>0.22-3.15</td>
<td>12-400</td>
<td>8.5-45</td>
</tr>
<tr>
<td>Mean</td>
<td>51.58</td>
<td>76.8</td>
<td>26.45</td>
<td>50.39</td>
<td>0.93</td>
<td>112.22</td>
<td>25.03</td>
</tr>
<tr>
<td>Std Dev</td>
<td>18.64</td>
<td>45.08</td>
<td>8.75</td>
<td>39.82</td>
<td>0.52</td>
<td>121.53</td>
<td>6.09</td>
</tr>
<tr>
<td>Std Error Mean</td>
<td>1.07</td>
<td>2.64</td>
<td>0.51</td>
<td>2.33</td>
<td>0.03</td>
<td>7.12</td>
<td>0.35</td>
</tr>
<tr>
<td>Upper 95% Mean</td>
<td>53.70</td>
<td>82.05</td>
<td>27.45</td>
<td>54.99</td>
<td>0.99</td>
<td>126.24</td>
<td>25.74</td>
</tr>
<tr>
<td>Lower 95% Mean</td>
<td>49.46</td>
<td>71.64</td>
<td>25.44</td>
<td>45.80</td>
<td>0.87</td>
<td>98.20</td>
<td>24.33</td>
</tr>
</tbody>
</table>

Table 1 shows that the properties of the soils tested are typical of properties for most clay soils encountered in nature. In addition, testing stresses included a wide range of stresses ranging from relatively low to relatively high confining pressures. The range for liquid limit indicated high variability among clays that translated into high plasticity index variability; given that, the range of values for plastic limit is relatively narrow. The measured FSS values covered a wide variety of friction angles from values as low as residual friction angles to values as high as peak friction angles.

By analyzing the FSS data, it is found that an FSS distribution resembles a normal probability curve (approximation, with all positive values), as shown in Fig. 5.1, with an estimated $\mu$ (mean) of 25.03 degrees and an estimated $\sigma$ (standard deviation) of 6.02 degrees. The box plot above the histogram indicates a symmetrical distribution with low variability and a couple of outliers. The outliers are handled by robust fitting and fitting a Johnson Su distribution; more details are given in the ANN generalities section.
5.2.2 Pre-processing: data classification and variable selection

The dataset, classified according to the testing device used during the tests, is subdivided into training and validation sets. The selection of input variables is conducted to ensure that the inputs contained adequate amounts of information for effectively representing the output space. This optimization task addressed both highly-correlated and lowly-correlated variables, among others (Fodor 2002). Geisser (1993) described predictive modeling as “the process by which a model is created or chosen to predict an outcome accurately.” In the practice of predictive modeling, the procedure of developing a model is outlined in an understandable way so that the model’s prediction accuracy with respect to future, yet-to-be-seen data can be quantifiable (Kuhn and Johnson 2013).

The model in this study is fitted to create an approximating function simpler than running a laboratory shear tests. If the observed response from laboratory tests is a function $y(x_1, x_2, ..., x_i)$, the modeled response is $\hat{y}(x_1, x_2, ..., x_i)$, and the goal, rather than increasing the number of input factors (using more soil properties), is to detect those inputs from the whole parameter space that have a dominant effect on the response. Insignificant input variables are
therefore, eliminated to obtain a parsimoniously accurate prediction model. In this study, the soil properties suggested by VandenBerge et al. (2013), obtained from the findings of a 2011 workshop on the use of fully softened shear strength (FSS) for the stability of slopes in highly-plastic clays, are used as the starting point for model generation. All the inputs are preliminary used to evaluate whether FSS needs to be correlated to more soil properties, as suggested by the workshop. A preliminary correlation analysis is conducted, and the results are presented to illustrate why simple linear regression approaches are not recommended because of the highly correlated soil properties involved in FSS estimation.

5.2.3 Artificial neural network (ANN) generalities

ANNs encompass a large class of model structures and learning methods, and by definition, represent a non-linear (in model parameters) empirical modeling approach. More generally, an ANN uses a two-stage non-linear regression or classification modeling approach and generally consists of an input layer, hidden layers, and an output layer (Hastie et al. 2009). The input layer represents the properties or variables (soil properties and effective stress). This layer takes in the inputs and performed calculations via nodes/neurons to be transmitted to subsequent layers. The hidden layers are the units in the middle that transform input effects into the output space of the response or responses. The output layer is responsible for producing the final results. It performs the final calculations that represent the response or responses of the problem that is being modeled (in this study, the response is secant FSS).

The ANN used in this study is a fully-connected multiple-layer perceptron (decision-making node). The data set reflected $n$ observations, $m$ input variables denoted by the vector $x_i$, and a continuous response $y_i$. The first step of model generation is mapping the input variables (soil properties and effective stress) into a design matrix. The constituents of the lowest hidden layers are comprised of an activation function applied to linear combinations of the design row
vector. The activation function is the function in the artificial neuron that delivers an output based on the inputs. In other words, the activation function introduced non-linearity into the output and decide whether a neuron should be activated or not updating the calculated weighted sum and adding bias with it. Hence, the activation function transforms the inputs into non-linear making them capable of learning and performing complex tasks. The activation functions used in the hidden layer are tanh and the identity transformation (Gaussian). The covariates are transformed by preprocessing the continuous variables to fit a Johnson Su distribution, using maximum likelihood to transform the inputs closer to normality, thereby mitigating the effects of outliers and skewed distributions. The general fitting approach is to minimize the negative log-likelihood of the data plus a penalty function applied to a scaled and centered subset of the parameters. The goal of a penalty function is to get the sets of model parameters that lead to a better fit and improve the optimization of the model. Penalty to the parameters is a way to combat the overfitting problem that can happen with ANN models. The penalty function used is

\[ e_i = y_i - \hat{y}_i \]

Cross-validation, also of vital importance to avoid overfitting and for testing the effectiveness of the model, refers to assessing the performance of the predictive model with an independent data set (Ghasemi et al. 2019; Hastie et al. 2009; Cheng and Titterington 1994). Because of the amount of data limitations, we used a K-fold cross-validation procedure. K-fold cross-validation refers to randomly partitioning the dataset into K equal subsamples. The model is then fit to the K - 1 portion of the data, and the remaining subsample is used for validation. The prediction error of the fitting model when using the validation set is estimated. This procedure is followed for all folds, and the K estimates of prediction errors are combined. The number of folds, specified for each model, is determined through an iteration process. A quasi-
Newton method, BFGS (Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm), iterates to optimize the penalty parameter during the training stage, while the BFGS algorithm simultaneously monitors the likelihood function of the validation set when the crossvalidation no longer reflects improvement, the early-stopping rule terminates the iteration. The performance results for the final model are based on the following statistics:

The “average difference (AD)” as an estimate of the model bias

\[
AD = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i) \quad (5.1)
\]

where \( n \) is the number of input vectors, \( y_i \) is the \( i^{th} \) measured response value, and \( \hat{y}_i \) is the \( i^{th} \) fitted response value. The “average absolute difference (AAD)” is the averaged absolute distance between fitted and measured values, defined as

\[
AAD = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \quad (5.2)
\]

The correlation between the measured and fitted values, \( r_{fit} \), is defined as

\[
r_{fit} = \frac{n \sum_{i=1}^{n} y_i \hat{y} - (\sum_{i=1}^{n} y_i)(\sum_{i=1}^{n} \hat{y})}{\sqrt{n \sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2} \sqrt{n \sum_{i=1}^{n} \hat{y}^2 - (\sum_{i=1}^{n} \hat{y})^2}} \quad (5.3)
\]

The closer \( r_{fit} \) is to 1, the better the fit.

A generalized coefficient of determination, \( r_{generalized}^2 \), is given as

\[
r_{generalized}^2 = 1 - \frac{2}{e_n} (L_{\hat{\beta}} - L_{Null}) \quad (5.4)
\]

where \( L_{Null} \) is the Laplacian log-likelihood using the sample median and mean absolute deviation as scaling parameters, \( L_{\hat{\beta}} \) is the negative loglikelihood of the set using the model parameters on the training data (\( \hat{\beta} \) is a generic name for model parameters, likelihood functions are a product of probability density evaluated at the observed value. Hence, given an observed
data point, maximum likelihood seeks to find values that maximize the probability of occurrence. Using the negative loglikelihood, the problem is reformulated seeking minimized the occurrence of $L\hat{\beta}$. $L\beta$ tell us how bad the model is performing (a loss function, the lower, the better). The importance of these statistics is that they allow comparison across different nonlinear modeling problems. Initially, the “full” (i.e., all inputs) model is fit to establish a baseline of performance, after which a reduced model uses a subset of the inputs. A schematic representation of a six neurons (in the hidden layer) ANN, as applied in this study, is presented in Fig. 5.2.

Fig. 5.2. Schematic illustration of ANN structure using JMP-PRO 14 (JMP pro 14, 2019), where the first layer represents input variables such as LL, PI, and so on, the second layer shows the activation neurons of the hidden layer, and *FSS represents the response, secant FSS friction angle.

The mathematical representation of the model is given by matrices representing the inputs ($X$), an additional node representing the bias (not shown in Fig. 5.2), and the hidden layer nodes or neurons ($A$). The arrows represent the weights ($W$), the contributions of the inputs to the activation nodes. $b_0$ is the intercept representing the bias at the output layer ($O$). Each input element $x_i$ is multiplied by a connection weight $w_{ij}$ and $w_{ir}$. The weighted input signals are summed, and $b_{aj}$ and $b_{ar}$, the bias values at the nodes, are added. The combined input is passed
through nonlinear activation functions $f(r)$ and $f(j)$ to produce the output $y_a$ of the hidden layer that serves as an input to the next layer. The weights are adjusted using $w_{or}$ and $w_{oj}$ through learning using the training data set. The sum of the adjusted weights is then added to the bias of the output layer $b_0$. A general form of the model equation can be written as:

$$\hat{y} = b_0 + \sum_{j=1}^{n} w_{oj} \cdot f(j) \left( b_{aj} + \sum_{i=1}^{n} w_{ij} x_i \right)$$

$$+ \sum_{r=1}^{n} w_{or} \cdot f(r) \left( b_{ar} + \sum_{i=1}^{n} w_{ir} x_i \right)$$

(5.5)

where $b_0$ is the bias of the output layer, $w_{oj}$ is the connection weight between neuron $j$ of the hidden layer ($j = 1$ to 3, number of neurons with this type of activation function) and the output layer; $b_{aj}$ is the bias at neuron $j$ of the hidden layer; $w_{ij}$ is the connection weight between the input variable $i$ (for $i = 1$ to 7, number inputs) and neuron $j$ of the hidden layer; $x_i$ is input $i$; and $f(j)$ is the Gaussian activation function at the hidden layer.

Similarly, $w_{or}$ is the connection weight between neuron $r$ of the hidden layer ($r = 1$ to 3, same as $j$) and the output layer; $b_{ar}$ is the bias at neuron $r$ of the hidden layer; $w_{ir}$ is the connection weight between input variable $i$ and neuron $r$ of the hidden layer; and $f(r)$ is the tanh activation function at the hidden layer. The activation functions used in this study are defined as:

$$f(r) = \tanh \left( \frac{b_{ar} + \sum_{i=1}^{n} w_{ir} x_i}{2} \right)$$

(5.6)

$$f(j) = e^{(b_{aj} + \sum_{i=1}^{n} w_{ij} x_i)}$$

(5.7)
5.3 Results and discussions

5.3.1 Correlation analysis (Pearson correlation coefficient)

The value of the Pearson correlation coefficient between two variables can range from -1 to +1, and the closer to +1 or -1, the more linearly associated are the two variables. The sign indicates the direction in which the variables are linearly related. Results with absolute values of 0.5 or higher are shown in bold and red text in Table 5.2 to indicate variables that are highly correlated.

Table 5.2. Correlation matrix (Pearson correlation coefficients) for inputs and output.

<table>
<thead>
<tr>
<th></th>
<th>CF</th>
<th>LL</th>
<th>PL</th>
<th>PI</th>
<th>A</th>
<th>$\sigma'$</th>
<th>FSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF</td>
<td>1.0000</td>
<td>0.6284</td>
<td>0.4134</td>
<td>0.6249</td>
<td>0.2451</td>
<td>-0.0582</td>
<td>-0.5423</td>
</tr>
<tr>
<td>LL</td>
<td>0.6284</td>
<td>1.0000</td>
<td>0.6879</td>
<td>0.9880</td>
<td>0.8530</td>
<td>0.0822</td>
<td>-0.6781</td>
</tr>
<tr>
<td>PL</td>
<td>0.4134</td>
<td>0.6879</td>
<td>1.0000</td>
<td>0.5674</td>
<td>0.5033</td>
<td>0.1090</td>
<td>-0.5173</td>
</tr>
<tr>
<td>PI</td>
<td>0.6249</td>
<td>0.9880</td>
<td>0.5674</td>
<td>1.0000</td>
<td>0.8606</td>
<td>0.0700</td>
<td>-0.6591</td>
</tr>
<tr>
<td>A</td>
<td>0.2451</td>
<td>0.8530</td>
<td>0.5033</td>
<td>0.8606</td>
<td>1.0000</td>
<td>0.1103</td>
<td>-0.5201</td>
</tr>
<tr>
<td>$\sigma'$</td>
<td>-0.0582</td>
<td>0.0822</td>
<td>0.1090</td>
<td>0.0700</td>
<td>0.1103</td>
<td>1.0000</td>
<td>-0.4854</td>
</tr>
<tr>
<td>FSS</td>
<td>-0.5423</td>
<td>-0.6781</td>
<td>-0.5173</td>
<td>-0.6591</td>
<td>-0.5201</td>
<td>-0.4854</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

As shown in Table 5.2, the soil properties under study are highly correlated with one another, with correlation coefficients >0.5. The lowest correlation among the inputs is between activity and clay fraction. The input factors with the highest correlation with the response are the liquid limit and the plasticity index. This explains why the general trend from the literature is the use of either plasticity index or liquid limit to estimate the FSS empirically.

The normal stress reflected low correlation with all the soil properties and the lowest correlation among all inputs with the output variable. However, as shown in the next section, because correlation does not imply causation, normal stress is still one of the most critical variables needed in the analytical model for FSS prediction. If multiple linear regression modeling is applied, multicollinearity can be an issue leading to inflation of prediction errors. Multicollinearity refers to the linear relationship between multiple predictors (i.e., inputs).
specifically, highly-correlated predictors in linear regression can result in magnified errors in fitted model coefficients and consequent degraded predictive performance (Kuhn and Johnson 2013). Collinearity can cause loss of information when reducing correlated variables.

A simple single-layer ANN model is fit to evaluate improvement concerning the prediction capabilities of the model. ANN is a nonlinear regression method and consequently is not subject to the collinearity limitation of structures with linear parameters. More specifically, ANN uses projection to linearly combine the input variables before being passed through nonlinear transformations (De Veaux and Ungar 1994). Performance results from comparing general linear regression with the ANN model are presented in Table 5.3. The number of folds for the complete models is five. The models are compared for both the raw and the corrected database (corrected means accounting for +2.5° in the friction angle from Stark and Eid (1997)).

Table 5.3. Summary of Statistics for performance evaluation (averaged five folds).

<table>
<thead>
<tr>
<th>Statistics</th>
<th>ANN</th>
<th>ANN (+2.5°)</th>
<th>Regression</th>
<th>Regression (+2.5°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>0.026</td>
<td>0.018</td>
<td>-0.074</td>
<td>-0.017</td>
</tr>
<tr>
<td>AAD</td>
<td>0.770</td>
<td>0.883</td>
<td>1.996</td>
<td>1.995</td>
</tr>
<tr>
<td>$r_{fit}$</td>
<td>0.97</td>
<td>0.97</td>
<td>0.90</td>
<td>0.90</td>
</tr>
<tr>
<td>$R^2$</td>
<td>*0.97</td>
<td>*0.97</td>
<td>0.82</td>
<td>0.81</td>
</tr>
</tbody>
</table>

* $R^2$ generalized ($R^2$ only applicable for training sets)

The results from the best models for both ANN and regression indicated that the predicted values of FSS given by the ANN models have a higher coefficient of determination, $R^2$, than the predicted values from the regression models. The closer the coefficient of determination to 1, the better. In addition, the $r_{fit}$ (correlation between the measured and fitted values), and the average absolute deviations performed better for the ANN models. Both ANN models have higher statistical performance (small differences, higher $r_{fit}$, and coefficient of determination) than the regression models.
5.3.2 Variable importance assessment

Variable importance is assessed by analyzing the results of the independent uniform inputs from a Predictor Profiler. The results are presented in Table 5.4. The Prediction Profiler displays profile traces for each input variable. A profile trace is the predicted response with changing a variable while the others are held constant. The Profiler re-compute the profiles and predicted response (in real-time) as the value of the input variables vary (SAS Institute Inc. 2018).

Table 5.4. Variable importance: Independent uniform inputs complete model raw data ANN.

<table>
<thead>
<tr>
<th>Input</th>
<th>Main Effect</th>
<th>Total Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>PI</td>
<td>0.568</td>
<td>0.63</td>
</tr>
<tr>
<td>$\sigma_n$</td>
<td>0.11</td>
<td>0.159</td>
</tr>
<tr>
<td>CF</td>
<td>0.076</td>
<td>0.131</td>
</tr>
<tr>
<td>LL</td>
<td>0.067</td>
<td>0.127</td>
</tr>
<tr>
<td>A</td>
<td>0.012</td>
<td>0.045</td>
</tr>
<tr>
<td>D</td>
<td>0.007</td>
<td>0.024</td>
</tr>
<tr>
<td>PL</td>
<td>0.005</td>
<td>0.023</td>
</tr>
</tbody>
</table>

It can be seen that from the soil properties, the PI and the CF are the coefficients with greater influence on the response. The effects of interactions on the response are seemed by differentiating between main effects and total effects. The main effects are due to changes in the inputs independently, and interactions refer to changes in the response by pairing the inputs. The normal stress with the lowest correlation with the response is the second independent input affecting it. The variable importance results in conjunction with the correlation analysis are used to optimize the inputs using a process of elimination until the minimum number of inputs associated with lower variation of model performance (statistics) is attained. The results are presented in the next section.
The prediction profiler of the complete model is used to visually determine how changing the settings of a given input influences the response. Fig. 5.3 shows that the findings are in agreement with the results presented by Castellanos et al. (2013), given that the model dataset contains the dataset from this study.

Fig. 5.3. Prediction profiler, the influence of input factors in FSS.
As CF, $\sigma'_n$, PI, and LL increase, the FSS decreases. The profiler shows that for PI values around 100%, the steep slope flattens, indicating a cutoff point at which the PI effects on FSS decrease. The CF profiler, on the other hand, shows a negative influence of clay fraction on FSS at a constant rate. The curvature of the failure envelope can be easily explained by interpreting the $\sigma'$ profiler results. The curvature of the failure envelope increases with decreasing normal effective stresses because of the higher effect of lower normal stresses on the FSS. The profiler shows that the steep influence from normal stress decreases after a cutoff value of around 150kPa.

The curvature of the failure envelope is a currently accepted concept; this profiler helps defines the cutoff values between curvature and linearity within the envelope. The activity profiler shows an increase of FSS as activity increases, for clays with <0.75 activity, (inactive clays). After a peak is reached around 0.75, the effects of activity on the response are minor, characterized by a flat line. The results indicate that in active clays (A>1.25) activity does not considerably affect the FSS. The individual effects of LL and PL on the response are negligible after 75% and 25% values, respectively.

A lower secant FSS is found for condition 2 (RS), of the nominal input “testing device,” indicating that the measured FSS from DS tests is slightly higher than the FSS measured from a ring shear device under a non-corrected condition (no 2.5$^\circ$ correction factor from Stark and Eid (1997)). In other words, the measured secant FSS using a DS device is higher than the measured non-corrected FSS using RS device but close to the corrected secant FSS from RS device. Other authors, such as Hvorslev (1939), Thompson and VandenBerge (2017), and Stark and Eid (1997), have found small differences between RS and DS tests measured FSS values.
The greatest differences among RS FSS and DS FSS are those reported by Castellanos et al. (2013). The general trend of the model indicated that, while testing device influences the secant FSS, soil properties variability and confining pressure are more relevant on affecting or changing the response when comparing corrected secant FSS using RS.

5.3.3 Reduced optimized model

After the reduction process based on variable importance is performed, a simplified model with four inputs is developed (σ′ n, PI, CF, and testing device). Using the variable “testing device” and the case histories, the reduced model, is set to the optimal model to attain a small difference between a back-analysis factor of safety and a factor of safety of one, resulting in the elimination of the variable testing device (D) from the model equation.

The iteration process consisted of selecting a condition closer to the field one. It is shown in the case histories section that since the raw dataset (using RS without correction) yielded very conservative results, a RS device would not be recommended for measuring FSS unless a correction is applied. The optimized reduce model contained three input variables and fifteen neurons. The general components and performance of the optimized model are given as follows:

\[ \hat{y} = 32.23 + 2.22f_1 - 9.06f_2 - 5.85f_3 - 3.66f_4 - 0.45f_5 + 0.26f_6 \\
+ 3.34f_7 - 3.80f_8 + 2.11f_9 - 3.78f_{10} - 0.76f_{11} - 1.54f_{12} \\
- 7.83f_{13} - 0.41f_{14} - 0.18f_{15} \]  

(5.8)

Each of the \( f(r) \) (r from 1-3) and \( f(j) \) (j from 4-15) are the corresponding activating functions presented in equations (6) and (7). The performance results for the reduced model are presented in Table 5.5 for the validation sets for all folds (for the optimized model \( K = 5 \)). \( R^2 \) for the training set is found to be 0.92.
Table 5.5. Performance statistics of the reduced final model.

<table>
<thead>
<tr>
<th>Statistics</th>
<th>K=1</th>
<th>K=2</th>
<th>K=3</th>
<th>K=4</th>
<th>K=5</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>0.57</td>
<td>0.003</td>
<td>-0.12</td>
<td>0.84</td>
<td>-0.08</td>
</tr>
<tr>
<td>AAD</td>
<td>1.24</td>
<td>1.14</td>
<td>1.52</td>
<td>1.62</td>
<td>1.07</td>
</tr>
<tr>
<td>( r_{fit} )</td>
<td>0.96</td>
<td>0.96</td>
<td>0.94</td>
<td>0.92</td>
<td>0.97</td>
</tr>
</tbody>
</table>

These results indicated that the prediction equation modeled the response well for all sets of data. Due to the extent of the model enclosed solution (considerably bias and weights), a spreadsheet is facilitated for the application of this method.

To avoid extrapolation, a series of constraints are placed on the developed model. The constraints on the dataset from the laboratory tests used to develop the correlations and the presented model are given in Table 5.6.

Table 5.6. Maximum and minimum constraints for the use of prediction tools based on the referenced tests.

<table>
<thead>
<tr>
<th>Reference</th>
<th>CF</th>
<th>LL</th>
<th>PI</th>
<th>( \sigma_{n}' )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Castellanos et al. (2016)</td>
<td>10-70</td>
<td>22-88</td>
<td>6-63</td>
<td>12-690 (varies)</td>
</tr>
<tr>
<td>Stark and Eid (1997)</td>
<td>18-88</td>
<td>20-288</td>
<td>4-244</td>
<td>12-400</td>
</tr>
<tr>
<td>Eid and Rabie (2017)</td>
<td>Same database of RS test results that Stark and Eid (1997)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wright (2005)</td>
<td>Only applicable for high plasticity and high CF&gt;50 fine-grained soils</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Predicting for inputs beyond the range of the data used to develop the model should be avoided due to the “danger of extrapolation” (see Devore, 2011) that can lead to invalidity of the fitted relationship for values outside the input space of the data set. The linear constraints for the final model are set, as shown in Tables 5.1 and 5.7, in the facilitated spreadsheet. The use of this model outside the presented constraints is not recommended and could result in gross prediction error, and an example of poor prediction due to such extrapolation is presented in the next section.
5.4 Application: Case Histories

While FSS secant friction angle correlations are mainly intended to provide an estimate of the laboratory-determined FSS, the real application of the strength values from correlations would be to estimate the mobilized shear of first-time slides in stiff-fissured clays. Nevertheless, Abdel-Ghaffar 1990; Stark and Eid 1997; Mesri and Shahien 2002; presented that some overconsolidated clays (primary heterogeneous ones) can mobilize a strength lower than FSS in a first-time slide.

As mentioned by Castellanos et al. (2016), the use of the FSS to analyze the entire cut of a highly plastic slope is an empirical simplification to obtain a factor of safety. However, the location of the failure surface does not always agree with the location on the field. The applications of FSS in this study are limited to soils mobilizing values larger than the residual in first-time slope failures.

Two case histories are selected from the literature and evaluated, using two approaches for validation:

- Back-analysis methodology considering a known location of the failure surface with the soil functions from predictive tools (correlations and model),
- Designing (assuming the failure surface is unknown) the slope with the soil function from predictive tools (correlations and model), and comparing the resulting failure surface with the one from the field.

Four correlations from the main published correlations for estimating secant FSS are selected for case-history verification Castellanos et al. (2016) (CI x PI), Stark and Eid (1997) (with and without correction factor), Eid and Rabie (2017)(with correction factor), and Wright (2005).
As presented in Table 6, most of the published correlations are based on the RS test results presented by Eid (1996), the exception being the correlations by Castellanos et al. (2013), Castellanos et al. (2016) developed using a DS device. The non-linearity of the FSS envelope is currently widely recognized. Hence, to perform the slope stability analysis in the limit equilibrium software SLOPE/W 2019 from GEO-SLOPE International Ltd (2019), a shear/normal function is generated with the results from the prediction tools. The Spencer method is used. The Spencer method satisfies all the requirements for static equilibrium (Duncan et al. 2014). Non-circular mode of failure with surface optimization is used.

5.4.1 Cut at Kingsbury Station, London

The failure at Kingsbury occurred in a cut located ~122m (400 ft) south of Kingsbury Station in Brown London Clay about 16 years after the cut was made (Skempton 1977). The rail line was opened in 1932, in a cutting ~6.0 m (20 ft) deep with a slope of 2 or 2 ¼:1. The first slips at the site occurred in 1947, and remedial measures were undertaken that were not fully successful, and instability was noted again in 1968. A section of 1947 corresponding to the first-time failure is analyzed. The slip plane assumed from the failure profile, and the slope geometry presented by James (1970) is used. The properties of the soil are LL of 82%, PL of 30%, PI of 52%, and CF of 55%. The pore water pressure coefficient is 0.3, and the unit weight is 18.8kN/m³. The soil functions generated by each prediction tool are presented in Fig. 5.4.

Figure 5.4 shows the differences between the slopes of the soil functions from the different correlations and the TG model (spreadsheet from this study). The lowest fitted line corresponds to the correlation by Stark and Eid (1997), with no correction value, while the highest line on the plot corresponds to the Eid and Rabie (2017) correlation. Table 5.7 summarizes the factors of safety resulting from fixing the critical surface to the field condition and varying the strength function, as well as the results from the design approach.
Fig. 5.4. Shear/Normal function for BLC = Brown London Clay at Kingsbury Station for slope stability analysis Case 1.

By combining the results from Fig. 5.4 and Table 5.7, it can be seen that, although most of the soil functions from the predictions tools have similar slopes, some resulted in conservative factors of safety.

Table 5.7. Comparison of factors of safety from back analysis and pre-failure analysis (design).

<table>
<thead>
<tr>
<th>Prediction Tools</th>
<th>FS (Back Analysis)</th>
<th>FS (Design)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Castellanos et al. (2016)</td>
<td>1.11</td>
<td>1.07</td>
</tr>
<tr>
<td>Stark and Eid (1997)+2.5°</td>
<td>0.96</td>
<td>0.92</td>
</tr>
<tr>
<td>Stark and Eid (1997)</td>
<td>0.86</td>
<td>0.83</td>
</tr>
<tr>
<td>Eid and Rabie (2017)</td>
<td>0.98</td>
<td>0.94</td>
</tr>
<tr>
<td>Wright (2005)</td>
<td>0.93</td>
<td>0.89</td>
</tr>
<tr>
<td>TG model</td>
<td>1.03</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Results presented before by Eid (1996), and Stark and Eid (1997) indicated that a correction factor is needed when measuring FSS with a ring-shear device, because non-corrected correlation may yield unduly conservative results. This statement is in agreement with the results in Table 7. The larger difference between a factor of safety of one and the back-calculated factor
of safety came from using the non-corrected Stark and Eid (1997) correlation, followed by Castellanos et al. (2016). The lowest difference with the unity factor of safety occurred when using the Eid and Rabie (2017). Nonetheless, if ±% difference is introduced to account for variability, most prediction tools are considered accurate for preliminary analysis with different degrees of uncertainties. As suggested by Castellanos et al. (2016), users should adjust their failure envelope parameters to account for uncertainty in the correlation.

For the design approach, when considering the surface unknown to simulate a pre-failure of predictive analysis, the results show more conservativism for most of the correlations. The TG model yielded closer to one factor of safety and matched better the shape of the in-situ scar. Results from considering the critical surface unknown are presented in Fig. 5.5. BLC indicates Brown London Clay. A superposition of all the results indicated that the location of the critical surface slightly varies among predictive tools. However, the shape and location of the critical surfaces from the analysis resemble the field slip. The critical surface from correlations with lower factors of safety, Stark and Eid (1997), with and without correction, Wright (2005), and Eid and Rabie (2017) agreed well with the slip plane from the field profile along the top and bottom portion of the failure. The slip found using Castellanos et al. (2016) agreed, correlating with most of the in-situ slip except for the top. The critical surface from the TG model (closer to a factor of safety of one) closely resembles the shape of the slip. A comparison of the back-analysis and design factors of safety presented in Table 5.7, indicating that for the conditions of the selected case, pre-failure analysis effectively matches the post-failure findings. Stark and Eid (1997), without correction, should not be used due to its conservative results. Castellanos et al. (2016) and Wright (2005) should be used at the designer’s discretion, and all other prediction tools should carefully be applied as preliminary estimation tools only.
Fig. 5. Comparison of slip plane presented by James (1970), (dashed-red line), and critical surface from slope stability analyses using various FSS correlations.

5.4.2 The Lesueur Landslide, North Saskatchewan River, Canada

A summary from the literature (James 1970; Mesri and Shahien 2002; Castellanos et al. (2016)) indicated that most first-time slides related to softening occurred in over-consolidated clays with $PI \leq 70$, (soils with PI beyond 70 are much less likely to be encountered), and while under such conditions all the studied prediction tools are equally applicable to the upper bound of PI, in certain cases such as “The Lesueur Landslide,” layers of highly overconsolidated clays (bentonitic clay shale) are found. Observing the constraints presented in Table 5.6, Castellanos et al. (2016) should not be used for cases such as Lesueur, and Wright (2005) should not be used for soil layers with $PI \leq 50$. The dangers of extrapolation are exemplified by using Castellanos et al. (2016) in this case.
The Lesueur Landslide, located on the outside of a bend of the North Saskatchewan River, was originally analyzed by Thomson (1971) and re-evaluated by Abdel-Ghaffar (1990) and Mesri and Shahien (2002). The results from these analyses indicated that for the lower part of the failure (horizontal-like), the strength mobilized was less than the peak but greater than the residual. The lower portion of the slope was composed of a thin layer of bentonitic clay shale with a PI of about 170, a clay fraction of 55, and averaged unit weight of 18 kN/m³, and the original water table conditions presented by Thomson (1971) are used. The soil functions generated by each prediction tool are presented in Fig. 5.6. The larger difference between soil functions generated by Castellanos et al. (2016) and the other prediction tools can be attributed to the danger of extrapolation.

![Graph](image)

**Fig. 5.6.** Shear/Normal function for Bentonitic Clay at Lesueur Landslide for slope stability analysis Case 2.

Table 5.8 shows the results from the back-analyses and the design cases. It can be seen that for highly overconsolidated clays, for the back-analysis, the correlation by Wright (2005) yields a factor of safety of one. Yet, for the design case, the same correlation resulted in a
conservative 0.90 factor of safety. Extrapolation from a correlation in a range of values beyond the inputs from which it is developed can result in an unsafe factor of safety, as shown in Table 8 for the Castellanos et al. (2016) results.

Table 5.8. Comparison of factors of safety from back analysis and pre-failure analysis (design).

<table>
<thead>
<tr>
<th>Prediction Tools</th>
<th>FS (Back Analysis)</th>
<th>FS (Design)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Castellanos et al. (2016)</td>
<td>*1.50</td>
<td>*1.27</td>
</tr>
<tr>
<td>Stark and Eid (1997) + 2.5°</td>
<td>0.98</td>
<td>0.88</td>
</tr>
<tr>
<td>Stark and Eid (1997)</td>
<td>0.93</td>
<td>0.81</td>
</tr>
<tr>
<td>Eid and Rabie (2017)</td>
<td>1.01</td>
<td>0.92</td>
</tr>
<tr>
<td>Wright (2005)</td>
<td>1.00</td>
<td>0.90</td>
</tr>
<tr>
<td>TG model</td>
<td>1.04</td>
<td>0.95</td>
</tr>
</tbody>
</table>

*Extrapolation

The results for both back-analysis and design using Castellanos et al. (2016) correlation are enormously high factors of safety related to the instability of the prediction tool at the range of inputs of this case. In other words, this exemplifies how a good correlation can be misused and how the end-user of prediction tools should be aware of the constraints or input space to which a prediction tool should be applied. As stated before, the use of any model or correlation outside its constraints is not recommended and could result in a gross prediction error.

The results from the TG model are the closest to the factor of safety of one for the design case. The superimposed failure surfaces from the design analyses and the field condition are presented in Fig. 5.7. Figure 5.7 shows that most of the critical slip surfaces from the analyses are similar and slightly differ from the observed slip surface. The factors of safety values from the analyses suggested that, although accurate results are obtained from back-analyses, for this case, the design factor of safety is somewhat conservative for some of the correlations that are correctly applied at this PI range. The verification from case histories indicates that, in general, most of the available prediction tools are accurate in estimating the FSS secant friction angle,
especially for back-analysis. Moreover, the proposed TG model shows the lowest difference to a factor of safety of one for the design approach, given to the model higher reliability for pre-failure analyses (assuming unknown field scar location) and prediction.

The approach followed herein is an accurate and effective way of studying the secant FSS, not only for analysis of variable relationships but also for enhancing prediction capability for design. The results from this approach could be further enhanced by increasing the size of the dataset used for the generation of the ANN model. In general, the use of ANN is left to developing models with ANN basics that guarantee repeatability of the results.
However, to simplify the use of the TG model as a preliminary tool for designing stiff-fissured clay slopes subjected to softening, the optimum model is facilitated as a spreadsheet.

5.5 Conclusions and Recommendations

The FSS is a practical approximation to estimate the mobilized shear strength acting on first-time slides of stiff clays prone to softening. Findings from the literature suggested that caution and engineering judgment are needed to apply FSS to design since not all overconsolidated soils seem to mobilize strength values as high as the FSS in a first-time slide. An analysis is conducted using the basics of regression to create a predictive model that simulates the fully softened shear strength (FSS) secant friction angle and allows for the study of the relationships among the variables involved. The main results and recommendations subjected to the assumptions and dataset analyzed herein are summarized as follows:

- Regarding the studied factors; liquid limit, plastic limit, plasticity index, clay fraction, activity, testing device, and effective normal stress, the Pearson correlation pairwise analysis shows a strong association among the inputs, and between the inputs and the output variable (FSS). The FSS is highly correlated with all the studied soil properties, (correlation coefficients >0.5), and moderately correlated with the normal effective stress (0.3<correlation coefficients <0.49). The liquid limit and the plasticity index are the properties with the highest correlation with the FSS.

- The results from the variable importance assessment indicated that the main factors affecting FSS are the plasticity index, the clay fraction, and the normal effective stress (facilitated in the spreadsheet). Hence, to generate an efficient predictive tool, there is no need to correlate FSS to other soil properties such as
activity, liquid limit, or plastic limit. In the absence of the plasticity index, a model containing the liquid limit, the clay fraction, and the normal effective stress also yield accurate approximations of FSS. Testing devices affect the FSS, but that effect is smaller than the one caused by changes in soil properties or confining stresses.

- The predictor profiler visually shows how changing the settings (values) of a given input influences the response. The main findings are summarized as follows:
  - For plasticity index values around 100%, the steep slope flattens, showing a cutoff point at which the plasticity index effects on FSS decrease.
  - The clay fraction profiler shows a negative influence of clay fraction on FSS at a constant slope, indicating that for all cases in the studied input space, an increase in the clay fraction decreases the FSS at a constant rate.
  - The curvature of the failure envelope increases with decreasing normal effective stresses because of the larger effect of lower normal stresses on the FSS. The strong influence from normal stress decreases after a cutoff value of around 150kPa. The curvature of the failure envelope is a currently accepted concept; this profiler helps defines the cutoff values between curvature and linearity within the envelope.
  - The activity profiler shows an increase of FSS as activity increases, for clays with <0.75 activity, (inactive clays).
  - The individual effects of LL and PL on the response are negligible after 100% and 25%, respectively.
• The performance results from comparing the best models between artificial neural networks (ANN) and multiple linear regression indicate that the ANN models performed better than the regression models. Numerically, the results indicate that ANN is a powerful tool for predicting the secant (FSS) friction angle of soils using the mentioned index properties. The ANN model identified hidden patterns between input and output variables in a linear and a non-linear form. Consequently, it is not subject to the collinearity limitations of structures with only linear parameters. As is factual for all models, the accuracy of the ANN model presented herein is limited to the constraints of the dataset used to develop such a model.

• The practical application of the developed model is studied by a comparative analysis based on two case histories. Four correlations from the main published correlations for estimating the FSS envelope are selected. Overall, for back-calculation analyses, no practical difference is found among the compared correlations. The predictive tools studied yield good agreement with the field condition, but are overly conservative when designing a slope. The proposed TG model is the prediction tool with the lowest difference to a factor of safety of one for the design approach, given the model higher reliability for pre-failure analysis (assuming unknown field scar location) and prediction. The correlations based on Ring shear test results presented by Eid (1996) with the +2.5° correction factor introduced by Stark and Eid (1997) seem to yield suitable estimates. The non-corrected Stark and Eid (1997) correlation yielded significantly conservative estimates. On the other hand, Castellanos et al. (2016), for values within the
correlation constraint, yield slightly unconservative estimates. Wright (2005) correlation seems to be more suitable for soils with a very high plasticity index $\geq 50\%$. Otherwise, the results are conservative. In general, correlations are useful tools for preliminary estimation of laboratory-measured FSS if used properly, placing constraints that guarantee prediction stability to avoid extrapolation is a must.

- In light of the findings, the use of a ring shear device it is not recommended to measure FSS unless the results are later corrected. While the correction factor (+2.5°) introduced by Eid (1996), and Stark and Eid (1997) seem to yield acceptable results, future research could focus on statistically determining the relationship between FSS values with different testing devices.

In summary, it is possible to affirm that the approach followed herein is an accurate and effective way of studying the FSS, not only for analysis of variable relationships but also for enhancing prediction capability for design. Future efforts should focus on determining whether the mobilized shear of overconsolidated clays subjected to softening could be better characterized by a general framework somewhat beyond the FSS concept.

References


CHAPTER 6. MOBILIZED SHEAR STRENGTH OF STIFF-FISSURED CLAYS: A MULTIPLE RESPONSE APPROACH

Yuderka Trinidad-Gonzalez¹, Vernon R. Schaefer², and Derrick K Rollins³
Modified from a paper to be submitted to the Engineering Geology Journal

Abstract

Assessing the shear strength of the soil is crucial for slope stability design and practice in engineering work. Generally, shear strength parameters are determined with an acceptable degree of accuracy in a laboratory for intact clay soils. However, the laboratory determination of the mobilized shear strength parameters in stiff-fissured clays is difficult due to the effects of softening and its laboratory characterization. The main purpose of this study is to develop a new indirect way of estimation of the mobilized shear strength of stiff-fissured clays for first-time slides, based on the integration of multiple response artificial neural networks and finite element analyzed case histories. The results showed that the proposed equation could be applied effectively for reliable prediction of a continuous and nonlinear failure envelope. The results also indicate that considering the model's good performance, the model can be used for preliminary stages of the rational design of stiff-fissured clay slopes. The prediction studies not only provide practical tools for estimation but also contribute to improving understanding of the main controlling factors of the mobilized shear strength of stiff-fissured clays in first-time slides. A comparison between the current methodology, the proposed method, and the field condition is presented to demonstrate the validity of the modeling approach.


¹Graduate Research Assistant, Department of Civil, Construction and Environmental Engineering, Iowa State University, Ames, IA 50011 yuderkat@iastate.edu, Corresponding author
²James M. Hoover Professor of Geotechnical Engineering, Department of Civil, Construction and Environmental Engineering, Iowa State University, Ames, IA 50011, vern@iastate.edu
³Professor, Department of Chemical and Biological Engineering, Department of Statistics Iowa State University, Ames, IA 50011, drollins@iastate.edu
6.1 Introduction

Slope failures have been responsible for immeasurable economic losses and casualties around the world. In many areas, the instability of slopes is related to strength loss due to time-dependent natural processes primarily affecting stiff-fissured clays. For these types of clays, the determination of the shear strength parameters for conducting stability analysis is still a significant challenge in the geotechnical field. Terzaghi (1936) was the first one to point out the special considerations needed when performing slope stability analysis in stiff-fissured clays. For first time slides, in either soft or stiff, joint-free clays, the mobilized shear falls very close to the peak strength from conventional laboratory test results (Duncan and Dunlop 1968). However, for first-time slides in stiff-fissured clays, the results of studies from several authors such as Henkel and Skempton 1955; Henkel 1957; Skempton 1964; Skempton and LaRochelle 1965; Bjerrum 1967; Duncan and Dunlop 1968; Skempton 1977; Stark and Eid 1993; Mesri and Shahien 2003; Castellanos 2013; Stark and Fernandez 2020) show that conventional methods of analysis do not lead to satisfactory results.

Skempton (1970) introduced the concept of fully softened strength (FSS), relating mobilized shear strength to shear strength measured in the laboratory by application of the critical state concept presented by Schofield & Wroth (1968). FSS is currently accepted as the controlling strength in low plastic homogenous clays (Mesri and Shahien 2003). Indirect estimation methods (i.e., not in the laboratory but from a mathematical expression) of FSS of stiff-fissured clays are conventionally implemented by empirical equations derived mostly from a simple linear regression analysis. In these empirical correlations, the only strength parameter estimated from its relationship with soil index properties and normal stress is the secant friction angle (secant FSS). In this study, the term FSS is used to refer to the secant friction angle in the fully softened state. Moving from the peak to the residual, the cohesion intercept tends to zero.
However, even small cohesion values can greatly influence the location of the failure surface unless the residual strength is the one in control or a softening process had completely destroyed the cohesive bond. There is vast information in the literature regarding measurement and estimation of the FSS to develop empirical correlations for preliminary design. However, several authors such as Potts et al. 1990; Stark and Eid 1997; Mesri and Shahien 2003; Skempton and Vaughan 2009 have found that residual strength can also play a part in the first time slides in clay fills and cuttings. Hence, the mobilized shear of first-time slides in stiff-fissured clays could be less than FSS. In addition, as presented by Castellanos et al. (2016), FSS is a simple concept that may yield a factor of safety close to the unity but may fail to predict the correct location of the failure surface. Generally, computed FSS failure surfaces are shallower than the observed field slips. This could be related to the assumption that cohesion is zero since resulting failure surfaces only considering the friction angle tend to be shallower surfaces. Mesri and Shahien (2003) presented a method in which portions of the failure surface were considered to mobilize the residual strength. This method provided a better agreement of the critical and observed failure surfaces. However, from a prediction point, the method is difficult to apply since the location and shape of the failure surface are unknown before the failure happens.

Building on the work of the aforementioned authors, in this study, multiple response artificial neural networks is paired with finite element (FE) analyzed case histories to indirectly estimate the mobilized shear strength parameters (outputs $a,b,c$ curve fitting coefficients for a non-linear failure envelope) in first-time slides in stiff-fissured clays based on soil index properties, normal stress, time of failure, pore water pressure conditions, and slope geometry properties (the inputs). This study aims to: (1) investigate the potential of multiple response artificial neural networks for the indirect estimation of the mobilized shear strength parameters
of first-time slides in stiff-fissured clays; (2) study the relationship among inputs and outputs using a correlation analysis; (3) improve the understanding of the main controlling factors of the mobilized shear strength of stiff-fissured clays in first-time slides and; (4) compare the performance of the presented approach with the existing methods. Finally, this work developed a python script (Van Rossum and Drake 2009) of the enclosed solution of the model to facilitate its application.

6.1.1 Current methodologies

There are several publications in the literature addressing the indirect estimation of shear strength parameters of the soils using related regression equations (Khanlari et al. (2012), Azimian (2017)). However, such studies were mostly developed from a laboratory database to address the indirect estimation of shear strength for either intact soils or rocks by determining the strength parameters with independent equations without accounting for the non-linearity of the failure envelop. For stiff-fissured clay slopes, studies have been focused on applying the FSS concept for developing correlations to empirically represent the mobilized shear strength. Currently, there are around seven main empirical correlations developed for simplification, as tools for preliminary estimation of FSS. The main published correlations are presented by Stark and Fernandez (2020), summarized in Table 6.1. The soil-related properties correspond to LL, PL, and PI, referring to liquid limit, plastic limit, and plasticity index. CF is clay fraction defined as the fraction of soil and $\sigma'_n$ is the normal effective stress.

Most of the current correlations (except from Castellanos et al. (2016), and Tiwari and Ajmera (2011)) were developed from a laboratory dataset presented by Stark and Eid (1997) in which specimens were tested using a ring-shear device, with a correction factor of + 2.5° introduced to convert ring shear FSS secant friction angles to consolidated drained triaxial compression (CD-TX) FSS secant friction angles.
Table 6.1. Some empirical correlations between FSS (output), index properties, and normal stress from previous studies.

<table>
<thead>
<tr>
<th>FSS correlation</th>
<th>Inputs studied</th>
<th>Output (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stark and Fernandez (2020)</td>
<td>LL, CF, $\sigma'_n$</td>
<td>FSS (secant friction angle)</td>
</tr>
<tr>
<td>Eid and Rabie (2017)</td>
<td>PI, $\sigma'_n$</td>
<td>FSS (secant friction angle)</td>
</tr>
<tr>
<td>Castellanos et al. (2016b)</td>
<td>PI or PI x CF, $\sigma'_n$</td>
<td>a, and b with two separate equations</td>
</tr>
<tr>
<td>Gamez and Stark (2014)</td>
<td>LL, CF, $\sigma'_n$</td>
<td>FSS (secant friction angle)</td>
</tr>
<tr>
<td>Tiwari and Ajmera (2011)</td>
<td>CF, PI, LL</td>
<td>FSS (secant friction angle)</td>
</tr>
<tr>
<td>Wright (2005)</td>
<td>LL, $\sigma'_n$</td>
<td>FSS (secant friction angle)</td>
</tr>
<tr>
<td>Mesri and Shahien (2003)</td>
<td>PI, $\sigma'_n$</td>
<td>FSS (secant friction angle)</td>
</tr>
</tbody>
</table>

Such a correction factor was introduced because, according to Stark and Eid (1997), the most relevant shear mode of failure in first-time slides in natural-cut slopes and compacted embankments is nearer to drained triaxial compression. All current empirical correlations are laboratory-based, and most of them account for non-linearity of the failure envelop.

6.2 Materials and methods

6.2.1 Data collection and correlation analysis

To generate a dataset for developing the prediction model, 68 slope stability finite element analysis from 35 cases histories of 12 different stiff-fissured clays, reported in the literature as potential first-time slides, are gathered. The information regarding the cases is presented in Table 6.2.

The soil-related properties considered in this study correspond to LL, PL, and PI, referring to liquid limit, plastic limit, and plasticity index, as defined by ASTM D4318 (ASTM International 2017). CF is clay fraction defined as the fraction of soil clay-sized <0.002mm. The geometric properties considered are the slope height (H) in meters, the slope angle of inclination ($\beta$) in degrees, and the water conditions are accounted for by determining the pore water pressure coefficient ($r_u$). This set of properties represents the eight inputs of the study.
Table 6.2. Summary of case histories reported as first-time slides in stiff-fissured clays.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Slope location</th>
<th>CF (%)</th>
<th>LL (%)</th>
<th>PL (%)</th>
<th>PI (%)</th>
<th>T (years)</th>
<th>H (m)</th>
<th>β (°)</th>
<th>ru</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skempton (1948)</td>
<td>Watford By-Pass</td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>7.5</td>
<td>4.6</td>
<td>23</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Ardley</td>
<td>59</td>
<td>28</td>
<td>31</td>
<td>33</td>
<td>52</td>
<td>8.47</td>
<td>22</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>Barrowden Rutland</td>
<td>62</td>
<td>31</td>
<td>31</td>
<td>52</td>
<td>83</td>
<td>4.5</td>
<td>28</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>Heyford</td>
<td>57</td>
<td>29</td>
<td>28</td>
<td>36*</td>
<td>126</td>
<td>9.5</td>
<td>23</td>
<td>0.27</td>
</tr>
<tr>
<td>Chandler (1974); Mesri and Shahien (2003); Castellanos et al. (2016)</td>
<td>Hunsbury Hill</td>
<td>60</td>
<td>28</td>
<td>32</td>
<td>43</td>
<td>44</td>
<td>12.8</td>
<td>24</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>Seaton</td>
<td>58</td>
<td>27</td>
<td>31</td>
<td>36</td>
<td>66</td>
<td>10</td>
<td>22</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>Stowehill</td>
<td>58</td>
<td>27</td>
<td>31</td>
<td>36</td>
<td>66</td>
<td>10</td>
<td>22</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>Wothrope (slip B)</td>
<td>68</td>
<td>27</td>
<td>41</td>
<td>48</td>
<td>9</td>
<td>6</td>
<td>18.43</td>
<td>0.2</td>
</tr>
<tr>
<td>Chandler and Skempton (1974); Skempton (1977); Thomson (1971)</td>
<td>Sudbury Hill Sections 1 and 2</td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>46</td>
<td>10.7</td>
<td>18.43</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>Lesueur</td>
<td>98</td>
<td>37</td>
<td>61</td>
<td>77</td>
<td>-</td>
<td>6.6</td>
<td>26</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Northolt S1</td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>35</td>
<td>8.5</td>
<td>21.8</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>Northolt S3</td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>35</td>
<td>7.9</td>
<td>21.8</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>Northolt S6</td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>35</td>
<td>8.5</td>
<td>21.8</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>Northolt S9</td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>35</td>
<td>8.7</td>
<td>21.8</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>Northolt S11</td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>35</td>
<td>8.97</td>
<td>21.8</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>West Acton SA</td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>46</td>
<td>4.9</td>
<td>18.43</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>56</td>
<td>46</td>
<td>4.9</td>
<td>18.43</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>98</td>
<td>9.2</td>
<td>26.5</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Grove Park</td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>98</td>
<td>9.2</td>
<td>26.5</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>98</td>
<td>9.2</td>
<td>26.5</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>16</td>
<td>6.1</td>
<td>23</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>Kingsbury S1</td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>16</td>
<td>6.1</td>
<td>23</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>16</td>
<td>6.1</td>
<td>23</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>Crews Hill A</td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>47</td>
<td>7.6</td>
<td>18.43</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>47</td>
<td>7.6</td>
<td>18.43</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>47</td>
<td>7.6</td>
<td>18.43</td>
<td>0.3</td>
</tr>
</tbody>
</table>
Table 6.2. (Continued).

<table>
<thead>
<tr>
<th>Reference</th>
<th>Slope location</th>
<th>CF (%)</th>
<th>LL (%)</th>
<th>PL (%)</th>
<th>PI (%)</th>
<th>T (years)</th>
<th>H (m)</th>
<th>$\beta^o$</th>
<th>ru</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whitstable</td>
<td></td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>99</td>
<td>9.15</td>
<td>18.43</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>99</td>
<td>9.15</td>
<td>18.43</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>-</td>
<td>15.20</td>
<td>18.43</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>82</td>
<td>30</td>
<td>52</td>
<td>55</td>
<td>-</td>
<td>15.20</td>
<td>18.43</td>
<td>0.3</td>
</tr>
<tr>
<td>Grange Hill</td>
<td></td>
<td>57</td>
<td>24</td>
<td>33</td>
<td>49*</td>
<td>57</td>
<td>15.40</td>
<td>28</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>57</td>
<td>24</td>
<td>33</td>
<td>49*</td>
<td>57</td>
<td>15.40</td>
<td>28</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>64</td>
<td>23</td>
<td>41</td>
<td>59*</td>
<td>103</td>
<td>4.60</td>
<td>18.43</td>
<td>0.35</td>
</tr>
<tr>
<td>Hullavington</td>
<td></td>
<td>57</td>
<td>21</td>
<td>36</td>
<td>51*</td>
<td>100</td>
<td>7.60</td>
<td>24</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>57</td>
<td>21</td>
<td>36</td>
<td>51*</td>
<td>100</td>
<td>7.60</td>
<td>24</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>57</td>
<td>21</td>
<td>36</td>
<td>51*</td>
<td>100</td>
<td>7.60</td>
<td>24</td>
<td>0.35</td>
</tr>
<tr>
<td>Verney Junction SA</td>
<td></td>
<td>60</td>
<td>26</td>
<td>34</td>
<td>30</td>
<td>20</td>
<td>9.80</td>
<td>22</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60</td>
<td>26</td>
<td>34</td>
<td>30</td>
<td>20</td>
<td>9.80</td>
<td>22</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60</td>
<td>26</td>
<td>34</td>
<td>30</td>
<td>20</td>
<td>9.80</td>
<td>22</td>
<td>0.35</td>
</tr>
<tr>
<td>Haddenham</td>
<td></td>
<td>77</td>
<td>27</td>
<td>50</td>
<td>60</td>
<td>55</td>
<td>17.10</td>
<td>19</td>
<td>0.36</td>
</tr>
<tr>
<td>Toddington</td>
<td></td>
<td>55</td>
<td>22</td>
<td>33</td>
<td>43</td>
<td>40</td>
<td>12.70</td>
<td>22</td>
<td>0.28</td>
</tr>
<tr>
<td>Charwelton SA</td>
<td></td>
<td>60</td>
<td>26</td>
<td>34</td>
<td>30</td>
<td>20</td>
<td>9.80</td>
<td>22</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60</td>
<td>26</td>
<td>34</td>
<td>30</td>
<td>20</td>
<td>9.80</td>
<td>22</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60</td>
<td>26</td>
<td>34</td>
<td>30</td>
<td>20</td>
<td>9.80</td>
<td>22</td>
<td>0.35</td>
</tr>
<tr>
<td>Standish Junction</td>
<td></td>
<td>58</td>
<td>25</td>
<td>33</td>
<td>43*</td>
<td>129</td>
<td>7.60</td>
<td>21</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>58</td>
<td>25</td>
<td>33</td>
<td>43*</td>
<td>129</td>
<td>7.60</td>
<td>21</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>58</td>
<td>25</td>
<td>33</td>
<td>43*</td>
<td>129</td>
<td>7.60</td>
<td>21</td>
<td>0.35</td>
</tr>
<tr>
<td>Bath</td>
<td></td>
<td>56</td>
<td>24</td>
<td>32</td>
<td>47</td>
<td>109</td>
<td>8.20</td>
<td>32</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>56</td>
<td>24</td>
<td>32</td>
<td>47</td>
<td>109</td>
<td>8.20</td>
<td>32</td>
<td>0.3</td>
</tr>
<tr>
<td>Haslemere SB</td>
<td></td>
<td>40</td>
<td>19</td>
<td>21</td>
<td>23</td>
<td>104</td>
<td>5.80</td>
<td>25</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40</td>
<td>19</td>
<td>21</td>
<td>23</td>
<td>104</td>
<td>5.80</td>
<td>25</td>
<td>0.3</td>
</tr>
<tr>
<td>Hildenborough</td>
<td></td>
<td>41</td>
<td>19</td>
<td>22</td>
<td>32*</td>
<td>69</td>
<td>11.60</td>
<td>18.43</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>41</td>
<td>19</td>
<td>22</td>
<td>32*</td>
<td>69</td>
<td>11.60</td>
<td>18.43</td>
<td>0.3</td>
</tr>
<tr>
<td>Abdel Ghaffar (1990)</td>
<td></td>
<td>36</td>
<td>18</td>
<td>18</td>
<td>40</td>
<td>5</td>
<td>17.61</td>
<td>26.56</td>
<td>0.35</td>
</tr>
</tbody>
</table>

* Estimated from soil type, not available in the original publication
In all cases, the observed slip surface is used with observed or assumed groundwater conditions from the original references. Additional ru values are analyzed where the original publication indicated an assumed ru value. The statistical description of the examined input properties is given in Table 6.3. Basic descriptive statistics are given. Normal probability plots for each case are analyzed and did not show a significant departure from normality (plots not shown for space considerations).

Table 6.3. Summary statistics of the range in input properties of the dataset used in this study with a sample size of 68 for each case.

<table>
<thead>
<tr>
<th>Statistics</th>
<th>CF (%)</th>
<th>LL (%)</th>
<th>PL (%)</th>
<th>PI (%)</th>
<th>T (years)</th>
<th>H (m)</th>
<th>β (°)</th>
<th>ru</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min-Max</td>
<td>23-77</td>
<td>40-98</td>
<td>19-37</td>
<td>21-61</td>
<td>5-129</td>
<td>4.5-17</td>
<td>18-32</td>
<td>0.38</td>
</tr>
<tr>
<td>Mean</td>
<td>46.62</td>
<td>68.75</td>
<td>27.18</td>
<td>41.56</td>
<td>62.77</td>
<td>8.92</td>
<td>22.00</td>
<td>0.25</td>
</tr>
<tr>
<td>Std Dev</td>
<td>10.14</td>
<td>13.70</td>
<td>3.70</td>
<td>10.64</td>
<td>35.35</td>
<td>3.17</td>
<td>3.52</td>
<td>0.09</td>
</tr>
<tr>
<td>Std Error Mean</td>
<td>1.26</td>
<td>1.71</td>
<td>0.46</td>
<td>1.33</td>
<td>4.41</td>
<td>0.39</td>
<td>0.44</td>
<td>0.01</td>
</tr>
<tr>
<td>Upper 95% Mean</td>
<td>51.19</td>
<td>72.12</td>
<td>28.11</td>
<td>44.22</td>
<td>71.60</td>
<td>9.71</td>
<td>20.90</td>
<td>0.27</td>
</tr>
<tr>
<td>Lower 95% Mean</td>
<td>46.09</td>
<td>65.33</td>
<td>26.26</td>
<td>38.90</td>
<td>53.94</td>
<td>8.12</td>
<td>21.10</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Table 6.3 also shows the statistical variation among the properties for the studied cases. The range T (124 years) corresponds to the largest range among the study inputs. The pore water pressure coefficient, slope height, and the plastic limit are among the properties with smaller ranges. A correlation matrix is presented in Table 6.4 to investigate the strength of the correlation among the variables under study.

The correlation matrix gives a pairwise comparison among the inputs and the outputs (determined by the slope stability analyses). The curve fitting parameters (i.e., the outputs a, b, and c) are related to the Mohr-Coulomb parameters c' and φ' (details are presented in section 3.2). As shown in Table 6.4, the soil properties (LL, PL, PI, and CF) are all significantly correlated (with correlation coefficients ≥ 0.64).
The highest correlations to time of failure with the inputs are the ones with the soil properties, while the higher correlation with the outputs is with the cohesion intercept. \(a^*\) represents the tangent of the secant friction angle for effective normal stress of one atmosphere and has the highest correlations with the slope height and the angle of inclination. \(b\) represents the shape curvature of the failure envelope, and \(c\) represents the cohesion intercept. The output variables are moderate to highly correlated with each other.

**6.2.2 Slope stability analyses**

The FE slope stability analyses are initially calibrated to the field condition for a linear failure envelope with a conceptual model of failure of a deformable, constrained material body with Mohr-Coulomb, elastic-perfectly plastic, and non-associated flow rule assumptions (angle of dilation considered zero). From these results, to account for the non-linearity of the failure envelope, the normal and shear stresses along the critical surfaces are plotted using a mathematical tool generally called “power function” (not a statistical power function).

Curve fitting parameters following the relationship presented by Miller (1988) for rock slopes are then obtained, and an iteration process is performed. The iteration process stops when a factor of safety of one is found, and the critical surface agrees with the field scar. In this way,
the results correspond to a prediction case in which the location of the surface is considered unknown when performing the analysis. The general equation developed by Miller (1988) for modeling shear stress at low normal stresses along rock discontinuities is selected because it is unknown before the analysis if the cohesion intercept is reduced to zero. Overconsolidated clays fall between cemented soils or rocks and soils when the FSS concept is not being applied. The components of Miller (1988) equation are as follows:

\[
\tau = c + a(\sigma'_n + d)^b + \sigma'_n\tan \theta_w
\]  

(6.1)

where:

\(\tau\) = The shear strength of the soil, \(\sigma'_n\) = normal effective stress

\(a, b\) = Shape parameters for power function. The model is linear for b (power coefficient) of one

\(c\) = y-intercept of the power curve, (represents the remaining cohesion on the soil after softening)

\(d\) = Represents the tensile strength of the soil,

\(\theta_w\) = Is the waviness angle that can be set to zero when the modeling does not correspond to a joint.

Although \(a\) is a fitting parameter, it can be related to the Mohr-Coulomb parameter \(\phi'\) by normalization using the relationship presented by Castellanos et al. (2016a):

\[
a^* = \frac{a}{P_a(1-b)}
\]  

(6.2)

where \(a^*\) is the tangent of the secant friction angle for effective normal stress of one atmosphere (from the normalized power function presented by Lade (2010)). For this study, \(d\) and \(\theta_w\) are considered zero. To determine the power coefficient (i.e. \(b\)) and whether or not there is cohesion (i.e. \(c\)), a log-log graph is generated such that its slope is \(b\). In the presence of
an intercept \((c)\), the log-log plot shows curvature, and the plot is adjusted. The most acceptable estimate of the power and intercept coefficients are determined after this adjustment. An example of the results of this procedure is presented in Fig. 6.1 for case 1 (Watford By-Pass) of the analyzed cases.

Fig. 6.1. The power coefficient \(b=0.86\) from Fig. 6.1 (c) with a small intercept (coefficient \(c=1.31\) found by extrapolating the line in Fig. 6.1 (d) to the origin, for case history 1 (Watford By-Pass).

Figure 6.1. (a) Shows the original power function from the results of the slope stability analysis. The log of the original stresses is then plotted to verify the linearity of the function, presented in Fig. 6.1 (b). A slight curvature can be seen. A calibration process is conducted by
subtracting potential intercept values from the dependent variable (shear stress called adjusted shear stress) until the fit is improved to a straight line with the results presented in Fig. 6.1 (c). The slope of the straight line corresponds to $b$. $b$ is used to create the modified plot shown in Fig. 6.1 (d) when the value of the small intercept equals 1.31 can be seen. The value of $c$ is found by plotting shear stresses versus normal stresses to the power of $b$ and extrapolating the line in Fig. 6.1 (d) to the origin.

The coefficient ($a$) the power ($b$) and the intercept ($c$) are used as initial strength parameters for the slope stability analysis of case 1. This procedure is repeated for all analyzed case histories. The values of the cohesion intercept where mainly small values, and for some cases, no intercept ($c=0$) is found. The larger values of $c$ correspond to the cases where the field scarf is considerably deep.

The stability analyses are performed in terms of effective stresses using the readily available computer program from the Rocscience suite (Rocscience Inc. 2019): RS2 2019, corresponding to 2D finite element analysis. The shear strength reduction technique is used for a numerical equivalence of a factor of safety (FS) of the slopes. The shear strength reduction method is used to weaken the soil in steps until the slope “fails” (Dawson et al. 2015, Griffiths and Lane 1999, Pradel et al. 2010, Fu and Liao 2010). The $FS$ is considered to be the ratio of the actual shear strength to the lowest shear strength of a rock or soil material that is required to maintain the slope in equilibrium, also called strength reduction factor ($SRF$). Due to the lack of three-dimensional geometry information, the 3D effects are not considered. The elastic modulus of the soil and the Poison’s ratio are considered constant values of 50 GPa and 0.4, respectively. A generic sketch of a continuum model with the material body meshed is presented in Fig. 6.2.
Fig. 6. 2. Section of a generic slope with inputs.

The section characterizes a slope with a unit width in the z-direction. The nodes in the mesh and the boundaries have two degrees of freedom, (displacements $u$ and $w$, in the $x$- and $y$-directions, respectively). Triangular 6-noded elements are used. The right and left boundaries of the mesh are fixed only in the horizontal direction ($u = 0$ at $x = 0$ and at $x = L$). No displacement is allowed at the base of the slope ($u = w = 0$ at $y = 0$). For each slope geometry, the mesh is adjusted to the minimum number of elements to converge to the approximated solution. The results from a mesh convergence analysis are shown in Fig. 6.3. The number of elements used varies from 3,000-6,000 as the area of the cases varies.

Fig. 6. 3. Mesh convergence study results.
Ten examples of reanalyzed failures are shown in Fig. 6.4 to illustrate the degree of agreement between the observed slip surface and the critical surface from the analyses. Figure 6.5 shows a summary of the procedure followed to develop the dataset for the proposed MRANN. The colored area represents the soil displacement from the FE slope stability analyses, while the black line indicates the observed failure surfaces as presented by the referenced sources.

Fig. 6.4. Example of results from slope stability analyses for SRF=1
For the FE analysis, the solid displacement is represented by a heat map where blue is zero displacement, and red is the largest displacement. In general, the results indicate good agreement in most of the re-analyzed cases when compare to the filed condition. The largest differences between the filed scarf and the critical surface are found at the top of the slope for specific cases (for instance, Westerfield and Verney Junction). This could be related to the heterogeneity of softening conditions or soil properties. However, this last statement is only an assumption due to the lack of information from the case histories to characterize heterogeneity or anisotropicity.

Fig. 6.5. Summary of the procedure followed to develop the dataset for the proposed MRANN
6.2.3 Multiple Response Artificial Neural Network (MRANN)

In general, ANNs are a large class of model structures and learning methods comprised of simple processing elements, called neurons. ANN neurons are inspired by the neurons in the biological nervous system. By definition, ANN represents a non-linear (in model parameters) empirical modeling approach. Generally, an ANN model consists of an input layer, hidden layers, and an output layer.

In this study, the input variables to develop the model are the properties in Table 6.3, while the output variables are the fitting curve parameters \((a, b, c)\) from the results of the slope stability section. The input layer represents the properties or variables (soil and geometric properties). This layer takes in the inputs and performed calculations via nodes/neurons to be transmitted to subsequent layers. The hidden layers are the units in the middle that transform input effects into the output space of the response or responses.

The output layers present the model response or responses (in this study, the responses are the fitting curve parameters for a continuous and nonlinear failure envelope). Hence, the ANN model applied in this study corresponds to a MRANN with a total of three neurons at the output layer. To develop the predictive model, the model is trained with a dataset. During this stage, the network adjusts the interconnected weights between neurons and layers, minimizing the errors. The MRANN used in this study is a fully-connected multiple-layer perceptron with eight neurons at the input layer (for the input variables) and three neurons at its output layer.

The number of neurons at the hidden layer is selected by an optimization process. Initially, the hidden layer has 15 neurons, and then boosting increases the number of layers by an automatic addition process. Figure 6.6 shows the architecture of the MRANN system used in this study.
Each input element $x_i$ is multiplied by a connection weight $w_{ij}$. The weighted input signals are summed, and $b_j$, the bias values at the nodes, are added. The combined input is passed through the nonlinear activation function $f_j$ to produce the outputs $y_j$ of the hidden layer that serves as an input to the next layer. The weights are adjusted $w_{0j}$ through learning using the training data set. The sum of the adjusted weights is then added to the bias of the output layer, $b_0$.

A general form of the model equation can be written as:

$$\text{out}_k = \hat{y}_k = b_{0k} + \sum_{j=1}^{n} \left[ w_{0kj} \cdot f \left( b_j + \sum_{i=1}^{n} w_{hij} x_i \right) \right]$$  \hspace{1cm} (6.3)

where, $\text{out}_k = \hat{y}_k$ is the output $k$ (for $k = 1$ to 3, number responses in this study), $b_0$ is the bias of the output layer, $w_{hij}$ is the connection weight between the input variable and neuron $j$ of the hidden layer; $f_j$ is the Gaussian activation function at the hidden layer. $b_j$ is the bias at neuron $j$ of the hidden layer, $w_{0j}$ is the connection weight between the neuron $j$ of the hidden
layer and the neuron $k$ of the output layer, the vector $x_i$ is correspond to the inputs ($for i = 1 \ to \ 8$, number inputs). The activation functions used in this study is defined as:

$$f(j) = e^{(b_j + \sum_{i=1}^{n} w_{ij}x_i)} \quad (6.4)$$

The activation function in the neuron of the hidden layers delivers an output based on the inputs. In other words, the activation function introduces non-linearity into the output and decides whether a neuron should be activated or not, updating the calculated weighted sum and adding bias with it. Hence, the activation function transforms the inputs into non-linear making them capable of learning and performing complex tasks. The network is model with the neural network platform available from JMP pro 14 (2019). The general fitting approach is to minimize the negative log-likelihood of the data plus a penalty function applied to a scaled and centered subset of the parameters. The goal of the penalty function is to get the sets of model parameters that lead to a better fit and improve the optimization of the model. Penalty to the parameters is a way to combat overfitting. The penalty function used is the sum-of-squares (centered and scaled) residuals (i.e., $e_l = y_l - \hat{y}_l$) penalty. The proposed model computes separate likelihoods for each response and calculates an overall likelihood for the model (as the sum of the loglikelihoods across all the individual responses). A single model equation that minimizes the error among the predicted observed value simultaneously for multiple responses is developed.

The optimum network is reached upon a minimum error by randomly varying weights. To prevent overfitting, the hold-out cross-validation approach is used. The available dataset is randomly partitioned into training, validation, and testing set. The training set is used to create the model, the validation set is used to cross-validate, and the testing sets is used to assess the performance of the model with an independent dataset (Ghasemi et al. 2019; Hastie et al. 2009; Cheng and Titterington 1994). A quasi-Newton method, BFGS (Broyden–Fletcher–Goldfarb–
Shanno (BFGS) algorithm), iterates to optimize the penalty parameter during the training stage, while simultaneously monitoring the likelihood function of the validation set when the cross-validation no longer reflects improvement, the early-stopping rule terminates the iteration (Hastie et al. 2009). The statistical performance of the final model is measured for each output and overall performance based on the following statistics:

The “average difference (AD)” as an estimate of the model bias is

\[
AD = \frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y}_k)
\]  
(6.5)

where \(n\) is the number of input vectors, \(y_i\) is the \(i^{th}\) measured response value, and \(\bar{y}_k\) is the \(k^{th}\) fitted response value. The “average absolute difference (AAD)” is the averaged absolute distance between fitted and measured values, defined as

\[
AAD = \frac{1}{n} \sum_{i=1}^{n} |y_i - \bar{y}_k|
\]  
(6.6)

The correlation between the measured and fitted values, \(r_{fit}\), is defined as

\[
r_{fit} = \frac{n \sum_{i=1}^{n} y_i \bar{y}_k - (\sum_{i=1}^{n} y_i)(\sum_{i=1}^{n} \bar{y}_k)}{\sqrt{n \sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2} \sqrt{n \sum_{i=1}^{n} \bar{y}_k^2 - (\sum_{i=1}^{n} \bar{y}_k)^2}}
\]  
(6.7)

The closer \(r_{fit}\) is to 1, the better the fit.

A generalized coefficient of determination, \(r_{Generalized}^2\), is given as

\[
r_{Generalized}^2 = 1 - e^{\frac{2}{n}(L_{\hat{\beta}} - L_{Null})}
\]  
(6.8)

where \(L_{Null}\) is the Laplacian log-likelihood using the sample median and mean absolute deviation as scaling parameters, \(L_{\hat{\beta}}\) is the negative loglikelihood of the set using the model parameters on the training data (\(\hat{\beta}\) is a generic name for model parameters), likelihood functions are a product of probability density evaluated at the observed value. A summary of the procedure followed when developing the MRANN model is graphically presented in Fig. 6.7.
Fig. 6. 7. Flow diagram of the procedure follow when developing the MRANN model
6.4 Results and discussions

6.4.1 MRANN Performance prediction

The performance of the model measured based on the statistics from Equations 6.5 to 6.8 is presented in Table 6.5. The results indicate that the prediction equation modeled the response well for all sets of data.

Table 6.5. Performance results for the MRANN model

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Training</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>0.08</td>
<td>0.1</td>
</tr>
<tr>
<td>AAD</td>
<td>1.15</td>
<td>1.14</td>
</tr>
<tr>
<td>$r^2_{generalized}$</td>
<td>0.99</td>
<td>-</td>
</tr>
<tr>
<td>$r_{fit}$</td>
<td>-</td>
<td>0.98</td>
</tr>
</tbody>
</table>

The closer the $r_{fit}$ and the $r^2_{generalized}$ to one, the better, while for deviations, the lower they are, the better. Due to the extent of the model enclosed solution (considerably bias and weights), a Python script (Van Rossum and Drake 2009) is facilitated for the application of this method.

6.4.2 Variable importance assessment

The controlling factors among the inputs in the mobilized shear strength coefficients are studied through a variable importance assessment. The results are presented in Table 6.6. The results indicate that the main factors controlling the overall model response are the clay fraction and the slope height. Important interactions between the factors can be seen by differentiating total and main effects. The total effect for each input indicates the sensitivity of the overall model to each studied factor. The lowest total effect corresponds to the pore water coefficient. This is related to the narrow range of the input given by the case histories. Some cases are limited to assumptions of pore water coefficients around the average (0.25-0.30).
Table 6.6. Variable importance ranking by total effects

<table>
<thead>
<tr>
<th>Factor</th>
<th>Main Effect</th>
<th>Total Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF</td>
<td>0.237</td>
<td>0.686</td>
</tr>
<tr>
<td>H</td>
<td>0.037</td>
<td>0.356</td>
</tr>
<tr>
<td>PI</td>
<td>0.081</td>
<td>0.318</td>
</tr>
<tr>
<td>β</td>
<td>0.046</td>
<td>0.219</td>
</tr>
<tr>
<td>T</td>
<td>0.047</td>
<td>0.155</td>
</tr>
<tr>
<td>$r_u$</td>
<td>0.014</td>
<td>0.021</td>
</tr>
</tbody>
</table>

6.4.3 Comparative study

To demonstrate the validity of the modeling approach, a comparison between the current methodology, the proposed method, and the field condition is performed. The current methodology refers to applications of FSS correlation. The correlation selected for the comparison is the one presented by Stark and Fernandez (2020). The case history selected for the comparison is used as a testing case. In other words, this case is not used to develop the predictive model, and it is only used to assess its performance. The case corresponds to Sudbury Hill (section 1) presented by Chandler and Skempton (1974); Skempton (1977). The cut was constructed around 1903, and a slip on the south side occurred around 1949. The soil at the site consisted of Brown London clay with reported LL of 82%, PL of 30, and CF of 55% at this site. An average pore water pressure coefficient of 0.28 is used following the water table information from the original publication. The unit weight used is 18.8kn/m$^3$. The curve fitting parameters from FSS and the presented approach are presented in Table 6.7. The results of the comparison are presented in Fig. 6.8.

Table 6.7. Curve fitting coefficients for slope stability analysis testing case

<table>
<thead>
<tr>
<th>Method</th>
<th>a</th>
<th>b</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSS correlation</td>
<td>0.834</td>
<td>0.851</td>
<td>-</td>
</tr>
<tr>
<td>Presented approach</td>
<td>0.47</td>
<td>0.94</td>
<td>0.2</td>
</tr>
</tbody>
</table>
The results indicate a good agreement between the observed slip surface and the critical surface from the presented method. Results in the literature indicated that the application of FSS to perform slopes stability analysis is an acceptable simplification for shallow failures. If the location of the failure surfaces matter, then other methods of analysis must be used (Castellanos et al. 2016a). These results are in agreement with the findings herein.

The computed SRF for both FSS and MRANN estimation tools are below one for a 0.73 and 0.94, respectively. However, the location of the failure surface from the proposed approach better agrees with the failure surface reported by the case history original source. Indicating the advantage of the MRANN estimation over the FSS approach. Hence, the proposed method could be used as an indirect estimation tool for preliminary analysis and design with a higher degree of accuracy than the existent approach.

6.5. Conclusions and Recommendations

In the present study, by combining finite element re-analyzed case histories with multiple response artificial neural networks (MRANN), a single equation for prediction of a continuous and nonlinear failure envelope is developed. The general form of this equation is given by (6.3).

Due to the extent of the model enclosed solution (considerably bias and weights), a Python script
(Van Rossum and Drake 2009) is facilitated for the application of this method. The presented approach not only provides practical tools for estimation but also contributes to improving understanding of the main controlling factors of the mobilized shear strength of stiff-fissured clays in first-time slides. From the results, several conclusions and recommendations can be drawn:

- The results from developing a suitable power curve for the study cases indicated that although the cohesion intercept is generally zero (the cohesion bond is destroyed by weathering), in some cases, like in deep-seated slip, small amounts of cohesion are still operative. The log-log plot test to determine the true power of the curve is seen to be an appropriate approach.

- The performance results indicate that the proposed MRANN model is suitable for predicting fitting curve parameters that can be used to represent a continuous and nonlinear failure envelope for stability analysis of the first-time slide in stiff-fissured clays. Numerically, the results indicate that the MRANN is a powerful tool for identifying hidden patterns between the inputs and outputs in a non-linear form. Consequently, the model is able to handle the complexity of optimization for multiple responses while not being subject to the collinearity limitations. As is factual for all models, the accuracy of the MRANN model presented herein is limited to the constraints of the dataset used to develop such a model.

- The results from the variable importance assessment indicate that the main factors controlling the overall model response are the clay fraction of the soil and the slope height.
The practical application of the developed model is studied by a comparative analysis. The results indicate the proposed MRANN yield higher agreement between the critical surface and the field observed slip surface than procedures based on the FSS concept for deep-seated slides.

In summary, it is possible to confirm that the approach followed herein is an accurate and effective way of estimating the mobilized shear strength of stiff-fissure clays for the design of first-time slides. However, indirect estimation tools should be used with caution and only as preliminary design tools. Extrapolation must be avoided to guarantee the stability of the estimation tool.

References


CHAPTER 7. GENERAL CONCLUSIONS AND RECOMMENDATIONS

7.1 General conclusions

This study presents the compilation of several statistical assessments conducted with the objective of developing a prediction tool for the study and estimation of the mobilized shear strength of stiff-fissured clays. The analysis is conducted as a progression from studying the slope stability mechanism for general soils, followed by studying changes in the factor of safety of a slope due to the introduction of reinforcing elements, and finally narrowing down the analysis to specific soils prone to softening. The findings from the specific analyses indicate that the proposed statistical approaches contribute to increasing the understanding of the slope stability mechanism in general and specifically focused on stiff-fissured clays. From the results, several conclusions can be drawn:

- For general slope stability analyses:
  - Understanding the factors involved in the slope stability mechanism is crucial to determining an optimum remediation technique. The applied statistical approach, integrations of design of experiments, and Bayes' theorems with slope stability analyses provide understanding regarding the inputs affecting the factor of safety of unreinforced slopes for each type of soil.
  - In general, for all soil types (purely cohesive, purely frictional, and mixed soils), the results indicated that six primary factors ($c'$, $\beta$, $r_u$, $\varphi'$, $\gamma$, $H$) and their interactions have statistically significant effects on $FS$ at $\alpha = 0.05$. The findings statistically demonstrate that the elastic modulus and the Poisson's ratio do not affect $FS$ significantly.
A comparison of mean FS indicated that there is no statistically significant difference between the results of 2DLE, 3DLE, and 2DFE analyses. The difference is found when performing a 3DFE analysis. The results exhibited that introducing 3D space in the slope stability analysis caused more variability in FS than the variabilities due to the method assumptions. These results contribute to the geotechnical profession by quantifying the difference between analytical methods. Hence, allowing the designer to make choices when performing their analyses.

The diagnostic testing results indicate that the analytical methods in this study have relatively low accuracy when a threshold FS value of 1 is used to indicate stability/instability. This suggests that considering a unit value for the threshold between failure and stability may not be analytically correct. To prevent unsafe design, analytically, a threshold FS value greater than 1.20 is recommended based on this study. In light of the findings, the reliability of back-calculated parameters should be considered carefully. These results contribute to increase knowledge regarding reliability studies in slope stability, and they can also be the basics to more rigorous assessment regarding back-calculation techniques.

For pile-reinforced slope stability analyses:

The results from this section provide a general depiction of how FS of slopes is affected by reinforcement specific properties. These results contribute to design optimization for pile-reinforced slopes. From the findings, for designing a pile-reinforced slope, it is recommended to
perform an interactive analysis accounting for the main effects and interactions of the factors. A preliminary design could be achieved following results from the Pareto and the surface plots presented. An optimal combination of pile location, embedment depth, diameter, and spacing, while fixing the slope properties, can be analyzed by applying the basics of statistical design of experiments/regression techniques presented herein.

- The main factors influencing $FS$ of a pile-reinforced slope are soil and geometry properties, as in the unreinforced case. Isolating the pile related factors, to study their effects, the results indicate that the optimum pile location may be a concern only for soils containing both friction and cohesion. Changing the location of a pile in pure soils (either purely cohesive or purely frictional soils) does not significantly affect $FS$.

- The relationship pile location-$FS$ is modified by cohesion, friction angle, slope angle, embedment depth, spacing, diameter, unit weight, pore water pressure, and slope height. The variation of these factors and their interactions modify the optimum pile location ranging from the middle to the top of the slope wherever the middle of the critical surface is located within the slope. For all the soil groups, if the piles are located in a favorable location, increasing the diameter and the embedment depth and decreasing spacing increases $FS$. However, for spacing selection, the decision comes to whether the pile row is intended to work as a group, single piles, or a pile wall (for close piles).
• For slope stability analysis of first-time slides in stiff-fissured clays:
  o The fully softened shear strength (FSS) is a practical approximation to estimate the mobilized shear strength acting on first-time slides of stiff clays prone to softening. Findings from the literature suggested that caution and engineering judgment are needed to apply this concept due to its empirical and simplified nature.
  o Four correlations from the current main published correlations for estimating the FSS envelope are evaluated. Overall, for back-calculation analyses, no practical difference is found among the compared correlations. The predictive tools studied yield good agreement with the field condition, but are overly conservative when designing a slope.
  o An artificial neural network (ANN) model to predict the secant friction angle FSS is proposed. This model shows a slightly better agreement with the factor of safety of one for a design approach (assuming unknown field scar location) and prediction.
  o The findings from the analysis will help to improve the understanding of applicability and limitations of the currently available methods to estimate FSS and also the applicability of the FSS concept in itself. Correlations are useful tools for preliminary estimation of laboratory-measured FSS if used properly by placing constraints that guarantee prediction stability to avoid extrapolation.
The findings from comparing secant FSS values measured using a direct shear test device and a ring shear device indicate that the use of a ring shear device is not recommended unless the results are later corrected.

This study's major contribution to advance the understanding of the mobilized shear strength of stiff-fissured clays resides on providing a general approach that could help overcome the applicability constraints of the FSS concept.

The approach consisted of combining finite element re-analyzed case histories with multiple response artificial neural networks (MRANN) to develop a single equation for prediction of a continuous and nonlinear failure envelope.

The prediction studies not only provide practical tools for estimation but also contribute to improving understanding of the controlling factors of the mobilized shear strength of stiff-fissured clays in first-time slides. The proposed model performed well and showed high agreement with the critical field surfaces where the FSS fail to apply (deep-seated failure surfaces).

In summary, it is possible to confirm that the approaches followed herein are accurate and effective ways of estimating the mobilized shear strength of stiff-fissured clays for the design of first-time slides. However, indirect estimation tools should be used with caution and only as preliminary design tools. Extrapolation must be avoided to guarantee the stability of the prediction.
7.2 Recommendations for future research

Further evaluation of slope stability in first-time slides is recommended because of the limited case histories used for the models developed herein. These limitations are related to the model input space constraints since extrapolation is not recommended for any prediction tool. Some recommendations for future research can be listed as follows:

- While the correction factor (+2.5°) introduced by Eid (1996), and Stark and Eid (1997) seem to yield acceptable results, future research should focus on statistically determining the relationship between FSS values with different testing devices for an accurate correction factor when determining FSS values using a ring shear device. Although extensive research has been conducted in the laboratory determination of FSS, the degree of disturbance that the clay will be prone to change from one state to another remains vague (peak-FSS-residual). Simulating in a laboratory environment, the disturbances that lead to modification of the clay's bounding and particle orientation are complicated. In the same way, the relationship between the strength values obtained from different testing devices such as triaxial, direct shear, ring shear apparatuses is still a controversy concerning the accuracy of the results (mode of shear) and correlation with index properties. Hence, experimental research should be dedicated to understanding FSS of stiff-fissured clays by using rational methodologies, applying statistical concepts to quantify the softening mechanisms that lead to strength reduction, and compared the strength factors from different testing devices. The results from such studies could help to determine how the mode of failure (different testing devices) influence FSS.
To improve the understanding in the estimation of the mobilized shear strength of stiff-fissured clays, it is recommended to evaluate transient analysis that will allow determining a time factor to analytically characterize softening. A general hypothesis for future research considers that studying time as a factor in a multi-factor statistical design experiment combine with transient analysis, the relationship between time of failure, pore water dissipation or buildup, and strength loss can be unhidden to determined or recognize patterns in the time of failure of slopes subjected to softening. The outcomes of such a study will provide the means to define time factors that can be used in FE analysis to account long-term softening assessments of stiff-fissured clays.

To account for the high multicollinearity among the inputs involved in slope stability analysis, further statistical assessments by techniques such as principal component analysis and partial least squares are recommended. In addition, due to the success in the performance of the applied statistical approaches, it is encouraged the introduction of statistical data evaluation for geotechnical processes that are otherwise analyzed applying one factor at a time techniques.

7.3 Data availability statement

Additional data, models, or code generated or used during the study are available from the author by request.