THE APPLICATION OF GEOSTATISTICAL METHODS FOR THE QUANTIFICATION OF MULTIPLE-SCALE UNCERTAINTY DUE TO ALEATORY GEOLOGIC VARIABILITY

by

David Lane Boyd
Copyright by David Lane Boyd 2019

All Rights Reserved
A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in Partial fulfillment of the requirements for the degree of Doctor of Philosophy (Geological Engineering).

Golden, Colorado

Date ________________________

Signed: ________________________
David Lane Boyd

Signed: ________________________
Dr. Gabriel Walton
Thesis Advisor

Signed: ________________________
Dr. Whitney Trainor-Guitton
Thesis Advisor

Golden, Colorado

Date ________________________

Signed: ________________________
Dr. Wendy Bohrson
Professor and Department Head
Geology and Geological Engineering
ABSTRACT

Tunneling projects in rock are characterized by a high degree of spatial uncertainty, which is due in part to the natural, random (aleatory) variability the rock possesses. Some degree of variability is intrinsic to all rock, and is present due to the complex nature of its deposition or emplacement and subsequent tectonics. This variability is present at multiple spatial scales, from heterogeneous grains to the project scale, where tectonics cause variability in discontinuity properties. As this variability contributes to overall uncertainty in tunneling projects, it is critical to understand and characterize this variability at multiple relevant scales. This research isolated the component of spatial uncertainty associated with aleatory geologic variability and evaluated statistical and geostatistical methods for quantification and characterization of this variability. Geostatistics has been commonly used in natural resource extraction and other data-sparse environments, and has been used extensively in this research as a means by which to better predict, characterize or quantify spatial uncertainty associated with aleatory geologic variability. As the first contribution of this thesis, 2-D covariance maps were generated for rock core specimen photos and were analyzed to identify the number of specimens required in order to adequately represent rock strength. This contribution identified a method by which to quantify this without testing large numbers of specimens at great cost. Next, sequential indicator cosimulation was used to integrate sparse borehole data with a geologist’s interpretation of subsurface lithology, identifying the value added by having a geologist’s interpretation over borehole data alone in uncertainty quantification. This identifies uncertainty in a geologist’s interpretation for use in tunneling projects, whereas geologist interpretations do not typically reflect spatial uncertainty besides boundary uncertainty (besides qualitative indications of confidence in specific parts of geologic boundaries). Finally, indicator kriging was used to quantify uncertainty in ground conditions both prior to and during excavation of the Caldecott Fourth Bore Tunnel in California, USA, demonstrating an approach
by which engineers and geologists could quantify uncertainty to inform high-level decision making. The completion of these works provides valuable insight into aleatory variability at multiple spatial scales and demonstrates novel approaches to integrate different types of geotechnical data, including subjective and interpreted, into geostatistical algorithms to better understand spatial uncertainty in the context of tunneling.
# TABLE OF CONTENTS

ABSTRACT ................................................................................................................................. iii

LIST OF FIGURES ..................................................................................................................... viii

LIST OF TABLES ....................................................................................................................... xiii

ACKNOWLEDGEMENTS ......................................................................................................... xiv

CHAPTER 1 INTRODUCTION .................................................................................................... 1

CHAPTER 2 RELEVANT LITERATURE ON THE APPLICATION OF GEOSTATISTICAL APPROACHES TO GEOLOGICAL ENGINEERING PROBLEMS ................................................................................................................................. 7

2.1 Variogram-Based Geostatistical Tools for Quantifying Spatial Uncertainty ............... 8

2.1.1 Kriging ........................................................................................................... 10

2.1.2 Simulation ..................................................................................................... 13

2.2 Other Statistical Tools for Quantifying Spatial Uncertainty ....................................... 16

2.2.1 Random Fields ............................................................................................... 16

2.2.2 Markov Chains .............................................................................................. 18

2.2.3 T-PROGS ...................................................................................................... 20

2.2.4 Multiple Point Statistics ................................................................................ 21

2.2.5 Decision Aids for Tunneling ......................................................................... 22

CHAPTER 3 ASSESSMENT OF ROCK UNIT VARIABILITY THROUGH USE OF SPATIAL VARIOGRAMS ................................................................................................................................. 26

3.1 Introduction ................................................................................................................. 27

3.1.1 Variability and Uncertainty in Engineering and Rock .................................. 28

3.1.2 Factors Influencing UCS Variability in Rock Cores ..................................... 29

3.2 Methodology ............................................................................................................... 31

3.2.1 Spatial Covariance and the Variogram .......................................................... 32

3.2.2 Generation of a 2-D Covariance Map ............................................................ 33
3.2.3 Sampling and Parameterization of 1-D Variograms ........................................ 36
3.2.4 Compilation of Data for a Single Rock Unit ..................................................... 38

3.3 Removing Sources of Error from Image Analysis .................................................. 38
3.3.1 Lighting Irregularities ...................................................................................... 38
3.3.2 Cropping out unusable portions of images ...................................................... 41
3.3.3 Presence of Drill Marks ................................................................................. 42
3.3.4 Wet vs Dry Imaging ...................................................................................... 44

3.4 Analysis and Results ............................................................................................ 45
3.4.1 Analysis of a₀ and c ....................................................................................... 46
3.4.2 Results from Analysis ................................................................................... 49
3.4.3 Discussion of a₀ – UCS variability relationship (Γ vs. κ) .................................. 50
3.4.4 Discussion of c – UCS variability relationship (θ vs. κ) .................................. 55

3.5 Relationship to the number of specimens required ................................................. 57

3.6 Conclusions ......................................................................................................... 60

CHAPTER 4 QUANTIFYING SPATIAL UNCERTAINTY IN ROCK THROUGH
GEOSTATISTICAL INTEGRATION OF BOREHOLE DATA AND A
GEOLOGIST’S CROSS-SECTION ........................................................................ 61

4.1 Introduction ......................................................................................................... 62

4.2 Project Methodology .......................................................................................... 64
4.2.1 Cosimulation and Cokriging Algorithms ...................................................... 65
4.2.2 Sequential Indicator Cosimulation Algorithms ............................................. 67
4.2.3 Creation of Cross-Sections for COSISIM ..................................................... 70
4.2.4 Calculations of Entropy Metrics .................................................................. 71
4.2.5 Modeling Decision in COSISIM ................................................................. 75

4.3 Input Parameter Sensitivity Analysis .................................................................. 85
4.3.1 Variogram Range ......................................................................................... 85
| Figure 2-1: | An example of a variogram and associated covariance. .................................................. 10 |
| Figure 2-2: | Kriged map with categorical variables using nitrate monitoring wells in New Zealand (left) and kriging variance map (right) (Baalousha, 2010). .............. 12 |
| Figure 2-3: | Three equally-probable realizations of soil water content over an area (Delbari et al., 2009). .................................................................................................................. 15 |
| Figure 2-4: | Shear strength random fields and resulting equilibrium deformation for a soil slope modeled utilizing two different correlation lengths – (a) 0.2 units and (b) 2.0 units (Griffiths et al., 2009). ................................................................. 17 |
| Figure 2-5: | Transition probability matrix in a carbonate-siliciclastic sedimentary environment (Miall, 1973). .................................................................................................................. 18 |
| Figure 2-6: | Rock type estimated using Markov Chains and borehole data (Qi et al., 2016).............. 19 |
| Figure 2-7: | Six realizations from T-PROGS (Fleckenstein et al., 2006).............................................. 20 |
| Figure 2-8: | (a) ‘True’ geology on a grid; (b) Measured data locations; (c) Training image used for simulating the true geology; (d) Training image of the ellipse structure patterns in the training image; (e-f) Two realizations of an algorithm using the training image and measured data to simulate the true geology (Strebelle, 2002). 22 |
| Figure 2-9: | Ground classes resulting from Markov chain simulation of lithology and water inflow within ‘metamorphic rocks’ (Haas and Einstein, 2002).............................................. 23 |
| Figure 2-10: | Example Time-Cost scattergram from DAT using prior information (Phase I) or updated information (Phase II) (Min et al., 2008).......................................................... 25 |
| Figure 3-1: | Comparison of UCS variability for five different rock units ............................................. 30 |
| Figure 3-2: | Flowchart showing the general method of obtaining geologic variability for a rock type; *Coefficient of Variation................................................................. 33 |
| Figure 3-3: | Process of obtaining a 2-D covariance map and N-pairs map. ........................................ 35 |
| Figure 3-4: | Example 1-D variogram showing the location of c and 3*\(a_0\) for a Stanstead Granite specimen. .................................................................................................................. 37 |
| Figure 3-5: | Effect of homomorphic filtering on a metabasalt sample with a distinct vein. ............. 40 |
| Figure 3-6: | The effect of cropping on a Stanstead Granite BTS specimen. ..................................... 43 |
| Figure 3-7: | The effect of horizontal laminations from drilling on the rock core and associated covariance map. ............................................................................................................. 44 |
Figure 3-8: Dry vs. damp Smaland Granite and associated covariance maps ............... 45

Figure 3-9: Covariance map of a metagranodiorite sample with noticeable orientation in fabric ................................................................. 48

Figure 3-10: (a) Relationship between $\Gamma$ and $\kappa$ for all cases ($y = 1600 \times / 100$);
(b) Relationship between $\theta$ and $\kappa$ for cases with no structural defects
($y = 0.55x$); * - no data were collected regarding failure type .................. 51

Figure 3-11: Comparison of three sedimentary units ........................................... 52

Figure 3-12: Comparison of three metamorphic units ........................................ 54

Figure 3-13: Comparison of a rock that commonly fails along structure versus one that has no structural failures .................................................. 56

Figure 3-14: (a) 2-D contour map of the number of specimens required to obtain a mean UCS value within a certain percentage $E$ of the true mean at 95% confidence;
(b) the same figure at 99% confidence ....................................................... 59

Figure 4-1: Flowchart of the development of the proposed methodology ............... 66

Figure 4-2: (a) A cross-section with borehole locations. (b) An entropy map created using this data and given input parameters .................................................. 73

Figure 4-3: (a) 2-D entropy map with 1-D Sample marked. (b) 1-D sample from the 2-D map in (a) ................................................................. 74

Figure 4-4: (a) Locations on the cross-section that possess the value ‘Rock 1’ in dark gray, with boreholes used to create the variogram presented as black lines. (b) Computation of the experimental variogram in two directions for ‘Rock 1’ using the borehole data only ........................................... 78

Figure 4-5: (a) A simple cross-section with two lithologies separated by an undulating boundary. Comparison of entropy maps created using (b) an isotropic and (c) an anisotropic variogram .................................................. 80

Figure 4-6: Results obtained based on varying the degree of decimation – (a) Entropy map with cross-section data decimated every 5 units (b) Corresponding perpendicular 1-D sample across a zone of high entropy. (c) Entropy map with the cross-section not decimated. (d) Corresponding perpendicular 1-D sample across a zone of high entropy .................................................. 82

Figure 4-7: Variogram of output realization using (a) no decimation, (b) decimation value of 5 units, and (c) decimation value of 10 units ........................................... 83

Figure 4-8: Results of (a) maximum entropy and (b) $W_{0.5}$ obtained by stochastically selecting a given number of realizations ........................................... 84
Figure 4-9: Results obtained based on varying the variogram range –
(a) Entropy map with a variogram range of 50m. (b) Corresponding
perpendicular 1-D sample across a zone of high entropy. (c) Entropy map
with a variogram range of 100m. (d) Corresponding perpendicular 1-D
sample across a zone of high entropy. ................................................................. 86

Figure 4-10: Results obtained based on varying the number of conditioning data –
(a) Entropy map with a maximum of 24 conditioning data. (b)
Corresponding perpendicular 1-D sample across a zone of high entropy.
(c) Entropy map with a maximum of 48 conditioning data. (d) Corresponding
perpendicular 1-D sample across a zone of high entropy. ........................................ 88

Figure 4-11: Results obtained based on varying the Markov-Bayes Parameter –
(a) Entropy map with a Markov-Bayes value of 0.10. (b)
Corresponding perpendicular 1-D sample across a zone of high entropy.
(c) Entropy map with a Markov-Bayes value of 1.00.
(d) Corresponding perpendicular 1-D sample across a zone of high entropy........ 89

Figure 4-12: Average W0.3 values with a given combination of input parameters.
(b) Average W0.5 values with a given combination of input parameters. .......... 91

Figure 4-13: (a) Complex ‘true’ geology with the increasing number of boreholes iteratively
given to a geologist to produce cross-sections with higher geological
understanding; (b) cross-sections made by a geologist based on the
varying numbers of boreholes................................................................. 93

Figure 4-14: Deterministic solutions from the most accurate parameter input
combinations, along with a resulting entropy map and accuracy plot,
using a cross-section with (a) four boreholes, (b) six boreholes, and (c) seven
boreholes. ................................................................................................................. 95

Figure 4-15: Deterministic solutions from the most accurate parameter input combinations,
along with a resulting entropy map and accuracy plot, using (a)
six boreholes and geologist’s interpretation and (b) simulation
with no secondary data............................................................................................. 99

Figure 5-1: An example of a variogram and associated covariance. .............................. 106

Figure 5-2: Diagram of the Caldecott Tunnel Fourth Bore site, with boreholes, alignment,
and the geologic units used in this study (blue) (after Geomatrix, 2008). ............. 110

Figure 5-3: Histogram of RQD to Ground Class for the Second Sandstone and
Orinda Formation geologic units (after Jacobs Associates, 2008). ....................... 112

Figure 5-4: (a) Modal outcome and actual Ground Class encountered during excavation;
(a) Example point-scale certainty map of modal outcome along with
correctness from the face maps............................................................................. 115
Figure 5-5: RQD values recorded in the borehole that transects the Second Sandstone geologic unit along with three realizations of the probability of Ground Class 1 occurring based on the RQD values................................. 117

Figure 5-6: RQD values recorded in the borehole that transects the Second Sandstone geologic unit and the corresponding probabilities of each Ground Class occurring based on the RQD values................................................................. 118

Figure 5-7: Verification plots for the Hard Data approach in the Second Sandstone using indicator kriging with different sub-vertical (SV) ranges based on multiples of the sub-horizontal (SH) range estimated from borehole data........ 124

Figure 5-8: Verification plots for the Soft Data approach in the Second Sandstone using indicator kriging with different sub-vertical (SV) ranges based on multiples of the sub-horizontal (SH) range estimated from borehole data....... 125

Figure 5-9: Verification plots for the Hard Data approach in the Orinda Formation using indicator kriging with different sub-vertical (SV) ranges based on multiples of the sub-horizontal (SH) range estimated from borehole data........ 126

Figure 5-10: Verification plots for the Soft Data approach in the Orinda Formation using indicator kriging with different sub-vertical (SV) ranges based on multiples of the sub-horizontal (SH) range estimated from borehole data........ 127

Figure 5-11: Verification plots for the Hard Data approach in the Orinda Formation West using indicator kriging with different sub-vertical (SV) ranges based on multiples of the sub-horizontal (SH) range estimated from borehole data........ 128

Figure 5-12: Verification plots for the Hard Data approach in the Orinda Formation East using indicator kriging with different sub-vertical (SV) ranges based on multiples of the sub-horizontal (SH) range estimated from borehole data........ 129

Figure 5-13: Verification plots for the Soft Data approach in the Orinda Formation West using indicator kriging with different sub-vertical (SV) ranges based on multiples of the sub-horizontal (SH) range estimated from borehole data........ 130

Figure 5-14: Verification plots for the Soft Data approach in the Orinda Formation East using indicator kriging with different sub-vertical (SV) ranges based on multiples of the sub-horizontal (SH) range estimated from borehole data........ 131

Figure 5-15: Change in modal certainty in the Second Sandstone geologic unit by adding face maps overlaid with the modal certainty prior to excavation; gray indicates that the specified distance ahead of face map is outside the problem domain............................................................................................................. 133

Figure 5-16: Change in modal certainty in the Orinda Formation geologic unit by adding face maps overlaid with the modal certainty prior to excavation; gray
indicates that the specified distance ahead of face map is outside the problem domain........................................................................................................................................................................ 134

Figure 5-17: Pythagorean total probability change (\(\Delta P\)) in the Second Sandstone geologic unit at distances 1, 2, 5, and 10 m ahead of the excavation, along with the prior modal certainty at each location; locations where the prior modal Ground Class outcome changed when face maps were added are identified. .................. 135

Figure 5-18: Pythagorean total probability change (\(\Delta P\)) in the Orinda Formation geologic unit at distances 1, 2, 5, and 10 m ahead of the excavation, along with the prior modal certainty at each location; locations where the prior modal Ground Class outcome changed when face maps were added are identified. ........ 136

Figure 5-19: Pythagorean total probability change (\(\Delta P\)) in the Second Sandstone at various distances ahead of the face for individual face maps and with mean and median values shown; results are segmented based on whether the a priori results correctly or incorrectly identified the Ground Class at each location of interest.................................................................................................................. 137

Figure 5-20: Pythagorean total probability change (\(\Delta P\)) in the Orinda Formation at various distances ahead of the face for individual face maps and with mean and median values shown; results are segmented based on whether the a priori results correctly or incorrectly identified the Ground Class at each location of interest.................................................................................................................. 138
### LIST OF TABLES

| Table 3-1: | Minimum number of specimens required* to know the true mean of UCS within 10% of the true value at 95% confidence | 58 |
| Table 4-1: | Calculated entropy values with given probabilities of lithology occurring. | 73 |
| Table 4-2: | Color code for 1-D samples. | 75 |
When I first started at Mines in August of 2012, one of the speakers at orientation told all of us “you can’t complete Mines on an island.” That stuck with me as a terrified 18 year old, and my experience here has verified that. The last seven years have been the best I’ve had. These years have introduced me to people who will be my friends, colleagues, and adopted family for the rest of my life. I would like to first acknowledge my two advisors, Drs. Gabriel Walton and Whitney Trainor-Guitton. They took a senior in Geological Engineering with no research experience and taught me a ton about this field, what it means to be a Ph.D., and how to learn. Further, my committee, Drs. Kathleen Smits, Paul Santi, and Mike Mooney were instrumental in developing my skills to be the best engineer and geologist I can be. Their insight and experience have been critical in my personal and professional development. I cannot thank them enough, especially when things got hard, but it was their expertise that pushed me to work harder and do better. They empowered me to develop skills that will benefit me for the rest of my career. For that, I am eternally grateful.

I’ve met an astonishing number of people at Mines whom I believe will be a part of my life for the next half century or more. My boyfriend, Daniel, has been a massive source of support. My close colleagues have been extremely helpful: Juliet, Heather, Caroline, Meriel, Luke, Sankhaneel, Deepanshu, Rami, Carlos, Brian, Kendall, and Mehmet. I also appreciated working with my fellow GSG executive members: Allie, Caitlin, Gauen, Nora, and Joe.

Finally, I want to give thanks to all of my family (included those who I consider family), my brothers in FIJI, and other friends. You have all made this easier through your love and support.
I dedicate this work to all of the strong women in my life.

My advisor, my committee member, the academics and non-academics;

From my mother, who's been gone 10 years next month,

to my great aunt, who turns 90 tomorrow;

My friends, my family, and those whom I've adopted as family; Both cis and trans -

I count your support as paramount to my success
CHAPTER 1
INTRODUCTION

The process of planning, designing, and executing rock engineering projects, especially in the subsurface, is complicated by uncertainty associated with variable geological conditions. Throughout the life of a project, from conception to excavation and construction, geologic uncertainty at differing spatial scales impacts performance and cost (Haas and Einstein, 2002; Langford, 2013). Unanticipated ground conditions during excavation lead to cost and time overruns, poor ground support design, and can ultimately be hazardous to equipment and workers (Haas and Einstein, 2002; Bernardos and Kaliampakos, 2004). For these reasons, an understanding of the geological uncertainty that exists at different spatial scales for a project is critical to ensuring a successful project for all stakeholders.

Overall geologic uncertainty is epistemic, and can be theoretically minimized with repeated sampling (Bedi, 2013). However, this is impractical for time and cost effectiveness. Accordingly, a large amount of research has been performed in order to better understand geologic uncertainty at multiple scales, including uncertainty in index testing due to variability in rock core specimens at the centimeter scale (e.g. Ruffalo and Shakoor, 2009; Pepe et al., 2017), spatial uncertainty in rock type or physical properties at the project scale (e.g. Tacher et al., 2006, Xiong et al., 2018), and uncertainty over time during the tunneling process (e.g. Haas and Einstein, 2002; Min et al., 2008).

As uncertainty in geology has become an increasingly popular topic of study (e.g. Frodeman, 1995; Miranda et al., 2009; Langford, 2013; Bond, 2015), researchers have attempted to isolate components of uncertainty into different categories; this has led to varying definitions of the components that make up uncertainty. Bedi et al. (2013) argues that overall uncertainty is solely
a combination of epistemic uncertainty and aleatory variability, whereas Mann (1993) divides overall uncertainty into aleatory variability, measurement uncertainty, sampling uncertainty, and modeling uncertainty. Regardless of the definition of uncertainty applied, aleatory variability is a critical component that, unlike other components of uncertainty, is impossible to reduce through improved data collection or analysis.

Aleatory variability is present due to the complex nature of the depositional environment or nature of emplacement of a given rock. Diagenetic, tectonic, and/or metamorphic effects also contribute to variability. Variability in rock is present at all spatial scales; at the rock specimen (centimeter) scale, it presents as differences in grain size, shape, or mineralogy. At the rock unit (meter) scale, tectonic or metamorphic effects can cause zones of weakness or zones with higher abrasion and strength. At the project (meter to kilometer) scale, variable, irregular boundaries and complex lithological relationships can limit structural understanding of the local and regional geology, in addition to local variability within geologic or geotechnical units. This compounding geologic variability makes understanding geologic uncertainty challenging without a keen understanding of the geologic variability.

The purpose of this research is to quantify and characterize the aleatory variability component of overall uncertainty at different spatial scales. While numerous academic papers (e.g. Frodeman, 1995; Miranda et al., 2009; Zhang et al., 2009; Miranda et al., 2013; Wellman et al., 2014; Wellmann, 2017) and entire doctoral theses (e.g. Bedi, 2013; Langford, 2013) have focused on spatial uncertainty in the geosciences and geotechnical engineering, works have largely ignored the geological variability component of overall uncertainty. Aleatory geological variability in the context of rock engineering projects is unique in that it presents both a scientific (geology) and engineering (geotechnical) problem, and this research combines elements from both fields. Using
various statistical and geostatistical approaches, this research identifies which specific algorithms
and approaches are appropriate for use at different spatial scales and in different contexts.
Geostatistical approaches used in this research have rarely been used in the field of tunnel
engineering, especially when considering geological data that have not been transformed into an
engineering framework. This is largely due to the fact that geostatistical methods have historically
been developed for natural resource extraction (e.g. Krige, 1951; Strebelle and Journel, 2001;
Soltani et al., 2014) or for hydrological applications (e.g. Lin et al., 2001; Liu et al., 2004).
Therefore, when attempting to use geostatistics for the rock engineering applications in this thesis,
there are challenges involved in integrating relevant geological data and generating results that are
useful and relevant.

In this thesis, the first analysis focused on the grain (millimeter) scale, and was to quantify
the variability between specimens of rock core from a single rock unit that are sampled during a
project. These rock core specimens are subject to index testing such as the Uniaxial Compressive
Strength (UCS) test to evaluate a rock core’s (and in turn, a geologic unit’s) geomechanical
properties. However, UCS tests performed on different rock core specimens from the same rock
unit for a project return different values based on variable geologic features such as heterogeneous
grain size and shape (Wong, 1982) or due to the presence of discrete features such as argillaceous
wisps and fossils in carbonate rock (Day et al., 2017). This effect of spatial geologic variability
between rock core specimens within rock units at the project scale leads to uncertainty in
evaluating the most appropriate UCS value for the unit or recognizing the variability in the UCS
values; the International Society of Rock Mechanics recognizes that increasing geologic variability
in a rock unit for a project increases the variability in the geomechanical properties of the rock
unit, and advises engineers to test enough rock core specimens of the same rock unit to “adequately
represent the rock sample” and that the appropriate number should be “a function of the intrinsic variability of the rock” (Ulusay, 2014). These guidelines are inherently vague, and do not provide engineers with a quantitative measure of variability nor do they state a range of acceptable values, leaving these decisions up to the engineers. Based on this information, there was a need identified to relate geological and geomechanical variability for a rock unit at the project scale using rock core specimens. This analysis used the geostatistical concept of the variogram and associated covariance to generate 2-D covariance maps from images of rock core specimens, and parameterization of these maps led to a relationship between geologic (parameterization of 2-D covariance maps) and geomechanical (summary statistics of UCS values) for 18 different rock units. These results were used to develop recommendations for the number of rock core specimens required to test in order to capture the overall variability within the rockmass.

The second analysis completed in this thesis addresses a need to better understand the uncertainty in geologist’s cross-sections generated prior to a tunneling project. During the site investigation phase of a project, geologists create cross-sections of subsurface geology based on borehole data and a general geologic understanding of the region. These cross-sections are used for preliminary large-scale decision making regarding a project (Fookes et al., 2000) and to make geotechnical cross-sections to support decision making (de Vallejo and Ferrer, 2011). Variable subsurface geologic conditions lead to some degree of inherent uncertainty in these cross-sections. Further, as these cross-sections are important components that contribute to the success of a project, a lack of understanding in spatial uncertainty in these cross-sections can lead negative project outcomes. Using a synthetic 2-D geologic section of rock type, sparse boreholes were given to a geologist for him to generate his estimate of the rock type on the section. The geostatistical approach for estimating the spatial uncertainty was done through a variogram-based simulation.
algorithm, as these algorithms are ideal for generating numerous realizations that inherently are not locally but globally accurate, and are a good metric for uncertainty due to random variability (Maironi, 2003). The algorithm Sequential Indicator Cosimulation (COSISM) was used as there are two types of indicator-coded (rock type) data: the borehole data and the geologist’s cross-section interpretation. The information theory concept of entropy was used to generate 2-D maps of uncertainty in the cross-section based on variability between realizations. Following an intensive parameter sensitivity analysis using a simple base case, an increasing number of boreholes from the section was iteratively given to the geologist, who created new cross-sections. Increasing the number of boreholes given to the geologist improved the accuracy of the created model, and as the number of boreholes increased, a different set of input parameters more faithfully recreated the spatial distribution of uncertainty in the geologist’s cross-section. The results of this project were also used to provide guidelines for geologists and engineers to better understand the uncertainty in their cross-sections, creating a powerful tool for integrating geological interpretation into a structured geostatistical framework for uncertainty quantification.

The final analysis completed in this thesis was performed in the context of a finished high-profile tunnel in the San Francisco Bay region in California, USA. Variable geologic conditions in a highly tectonized rockmass led to a high degree of uncertainty in ground conditions during excavation. Due to the geological environment, ground conditions changed over scales on the order of several meters, leading to unexpected changes in support requirements and potentially hazardous conditions for workers and equipment. Both data acquired prior to excavation (boreholes) and data acquired during excavation (face maps) were used to quantify uncertainty in ground conditions both prior to and during excavation. Rock Quality Designation (RQD) in boreholes correlated reasonably well to the engineering Ground Class, while true Ground Class
conditions were recorded in face maps. The correlation between RQD and ground conditions is not only inherently uncertain, but it is also challenging to use such a correlation to appropriately estimate the Ground Class distribution along a tunnel alignment and the associated uncertainty. In this case, indicator kriging was used to estimate the Ground Class and uncertainty. Two histogram sampling methods between RQD and Ground Class were evaluated, and face maps were used to compare these methods. Once this was completed, face maps were iteratively added as input data to the geostatistical simulations in order to evaluate the reduction in uncertainty ahead of the maximum extent of excavation (the farthest point excavated) at each excavation interval.

Collectively, the contributions in this thesis advance the capabilities of scientists and engineers to account for spatial uncertainty in rock across a range of scales. This research integrates many different types of data that are outside the bounds of the conventional geostatistics, including high resolution image data and uncertain human interpretations. Additionally, this research carefully combined multiple data types, including data of the same variable with large differences in extent and coverage, and data that were conflicting (e.g. the prior information did not agree with information obtained through subsequent sampling). The next Chapter summarizes relevant research on uncertainty quantification in tunneling projects and geostatistics. The following three chapters present the three projects introduced above. Chapters 3 and 4 are papers that have been accepted for publication in the journal “Engineering Geology” as written, while Chapter 5 is intended for journal submission immediately following the submission of this dissertation document. Chapter 6 provides some an overview of the conclusions from this body of research as a whole.
CHAPTER 2

RELEVANT LITERATURE ON THE APPLICATION OF GEOSTATISTICAL APPROACHES TO GEOLOGICAL ENGINEERING PROBLEMS

In rock mechanics, geologic uncertainty has been studied as early as the first ISRM Congress in 1966 (Hadjigergiou and Harrison, 2012), as scientists and engineers have long considered an understanding of uncertainty and error essential to successful rock engineering project planning, design, construction, and excavation. As rock is a heterogeneous and variable material for engineering purposes, many different approaches have been developed in order to help understand the rock and aid engineers take on difficult and challenging projects. Much of this work considered the grain scale, with an attempt to understand how the individual grains of the rock at the millimeter and centimeter scale behave and how this affects how rocks break and fail (Wong, 1982; Martin, 1994; Martin and Chandler, 1994). This research took a much different approach, using geostatistics to map the rocks and quantify heterogeneity and anisotropy within rock specimens.

Numerous empirical systems including the Q-System (Barton et al., 1974), Rock Mass Rating (Bieniawski, 1976), and the Hoek-Brown Failure Criterion (Hoek and Brown, 1980) were developed to help engineers predict rock mass behavior and evaluate support needs for underground excavations. Many of these systems rely on index testing of core specimens sampled from the rock, and it is critical that a sufficient number of specimens are tested to fully capture the variation present. Much of this work has been based on statistical analyses that used large data sets to constrain the most appropriate number of specimens (e.g. Gill et. al., 2005; Ruffolo and Shakoor, 2009; Pepe et. al., 2017), but this is inherently costly and labor-intensive. This research uses the understanding of rock specimen behavior outlined by numerous authors (Wong, 1982; Martin,
1994; Martin and Chandler, 1994) and used the concept of geostatistics to constrain the number of core specimens by quantifying variability within each specimen.

These systems have been further refined as engineers have found that variable rockmass conditions are difficult to characterize in the framework of an empirical system (Palmstrong and Broch, 2006; Cai, 2011). Others have found that uncertainty is not always properly recorded or reported when establishing rock and discontinuity property inputs for empirical systems, and that under-sampling with respect to the geologic variability of the material leads to unreliable results when attempting to use an empirical system for rock mass behavior, design, and support (Priest and Hudson, 1981; Gill et al., 2005; Ruffolo and Shakoor, 2009).

In natural resource exploration, statistical and geostatistical methods for spatial uncertainty quantification have been widely used (Goovaerts, 1997; Deutsch and Journel, 1998). Many of these methods are based on the variogram approach for spatial correlation quantification (Cressie, 1985), and range from deterministic approaches such as kriging (Delhomme, 1978; Cressie, 1990) to probabilistic approaches such as stochastic simulation (Deutsch and Cockerham, 1994). Multiple point statistics including training images (Strebelle, 2002) are also used for spatial uncertainty quantification, as well as other approaches including Markov Chains (Miall, 1973; Elfeko and Dekking, 2001) and Transition Probabilities (Carle, 1999).

The remainder of this section will discuss some of the many different algorithms and approaches introduced above, as well as the specific applications of each approach.

2.1 Variogram-Based Geostatistical Tools for Quantifying Spatial Uncertainty

Variogram-based algorithms have been used extensively in the geosciences, including applications in natural resource evaluation (e.g. Xu et al., 2002; Horata and Soares, 2010) and
environmental remediation (e.g. Cattle, 2002; Liu et al., 2004). All of these algorithms are based around the variogram, which is a function that represents the spatial correlation of a data set (Bohling, 2005).

The variogram is a 1-D representation of the spatial continuity of data in a given direction (Bohling, 2005). The variogram is a function of lag distance, or the distance between data points, and is given by Matheron (1963) as Equation 2-1.

\[ \gamma(h) = \frac{1}{2N(h)} \sum_{N(h)} (z(u + h) - z(u))^2 \]  

(2-1)

In Equation 2-1, \( \gamma(h) \) is the value of the variogram at a distance of \( h \) meters, \( N(h) \) is the number of data pairs at a lag distance of \( h \) meters, and \( (z(u+h) - z(u))^2 \) is the squared difference of a pair of data separated by a distance of \( h \) meters.

An example of a variogram with associated covariance can be found in Figure 2-1. As lag distance increases, the variogram value increases and the covariance value decreases. The variogram value at which the no more increase is observed is called the sill value, and the lag distance at which that occurs is called the range.

Each point on the variogram is representative of all of the data pairs that are separated by a certain lag distance within a certain lag tolerance. The covariance values are a transformation of the variogram that are used to represent the variogram in the majority of variogram-based algorithms, and, assuming no small-scale error is present, these values are commonly calculated as the sill of the variogram minus the variogram value at each lag distance (Christakos, 1984).
2.1.1 Kriging

A simple algorithm that utilizes the variogram is kriging. Kriging was developed for the purposes of estimating spatial ore concentrations in gold mines in South Africa, in an environment where measured data are sparse (Krige, 1951). Kriging is a linear unbiased estimator that estimates values at locations in space where data have not been measured by calculating the weights of nearby (measured) data points, honoring the spatial correlation established by the variogram and reducing the effect of data clustering by decreasing the weights of clustered data (Matheron, 1963). With the weights solved using the covariance (from the variogram) with the distance between measured points and the unknown point, as well as the distance between measured points to limit the effect of data clustering, the value at the unknown point is calculated using Equation 2-2.
$$X^*(\vec{u}_0) = \lambda_0 + \sum_{i=1}^{N} \lambda_i X_i(\vec{u}_0)$$  \hspace{1cm} (2-2)

In Equation 2-2, $X^*(\vec{u}_0)$ is the kriged value at unknown location $\vec{u}_0$, $\lambda_0$ is the global mean value over the region, $\lambda_i$ is the kriging weight for the $i^{th}$ conditioning point, and $X_i$ is the value at the $i^{th}$ conditioning point (Yamamoto, 2000).

Additionally, kriging returns a value for uncertainty in the result, which is called the kriging variance (Yamamoto, 2000).

$$\sigma^2(\vec{u}_0) = \text{Var}[X(\vec{u}_0) - X^*(\vec{u}_0)] = C(0) - \lambda \cdot c_{i\vec{u}_0}$$  \hspace{1cm} (2-3)

In Equation 2-3, $\sigma^2(\vec{u}_0)$ is the kriging variance at a location $\vec{u}_0$, $C(0)$ is the value of the sill of the variogram, and $\lambda \cdot c_{i\vec{u}_0}$ is the dot product of the vector of kriging weights, $\lambda$, and the vector of data-to-unknown covariances. Figure 2-2 shows an example of a kriged map of nitrate vulnerability transformed into ten categories in New Zealand; the nitrate monitoring wells used in the analysis are shown on both maps (Baalousha, 2010).

As this figure shows, kriging is a powerful tool for deterministically estimating properties away from known data points, and is useful for a baseline analysis of the potential spatial distribution of a property. Additionally, the kriging variance gives an associated estimate of the spatial error in the deterministic solution (Haas, 1990), and as the data in Figure 2-2 show, away from data points where the kriging variance is highest, the estimated value trends towards the mean.
Figure 2-2: Kriged map with categorical variables using nitrate monitoring wells in New Zealand (left) and kriging variance map (right) (Baalousha, 2010).
While kriging does output a metric for uncertainty (the kriging variance), it is not an ideal metric for comparing uncertainty between simulation grids, as this value is homoscedastic, meaning it is independent of the data values for the conditioning data used to obtain the kriging estimate (Olea, 1991; Yamamoto, 2000). This makes the kriging variance useful for only comparing different sections of the same grid or for testing sensitivity in the input parameters for the kriging algorithm.

Since its introduction, kriging has been modified significantly for a broad array of applications, including the ability to take on indicator-coded values and categorical values (e.g. Bierkens and Burrough, 1993) and the ability to incorporate secondary data, also known as cokriging (e.g. Xu et al., 1992). In a general cokriging approach, the sparse measured data are used as primary data, and spatially extensive, low-resolution data are used as secondary data to inform the spatial correlation of the primary data, which is the data property being evaluated throughout the grid (Zhu and Journel, 1993; Goovaerts and Journel, 1995).

Indicator kriging, used in this research, is commonly used when continuous data possess one or more critical data thresholds, such as a maximum allowable value of contaminant concentration in groundwater (e.g. Liu et al., 2004) or minimum value of permeability in an aquifer (e.g. Ritzi et al., 1994). We use indicator kriging differently in this research by not defining cutoffs, which returns probabilities of each categorical variable occurring.

2.1.2 Simulation

Variogram-based stochastic simulation algorithms draw upon kriging for interpolating random or categorical variables between locations of known information. However, whereas kriging is a deterministic approach, stochastic simulation algorithms are a probabilistic approach
that use the kriging estimate and the kriging variance to produce multiple equally-probable simulations of the region (Deutsch and Journel, 1998). In practice, these algorithms are primarily used in fluid transport or to identify spatial patterns such as geological features or regions of uncertainty (Lin et al., 2001; Lee et al., 2007; Soltani et al., 2014). In the field of tunneling engineering, simulation algorithms have been used to estimate rockhead elevations or for SPT (standard penetration test) values (Grasmick, 2019).

In order to generate a realization using a stochastic simulation algorithm, a random path to every location on the simulation grid is created. At each location along the path, the kriging estimate and kriging variance are calculated. The kriging estimate and kriging variance become the mean and variance of a conditional cumulative distribution function (ccdf) from which a Monte Carlo draw is used to assign a value at the location (Soares, 2001). This process is completed for the remainder of the grid; the incorporation of previously-simulated locations as conditioning data allows for a ‘smoother’ result than if all of the locations were simulated independently of each other, and new realizations are created using a different path around the grid (Soares, 2001). Due to this approach, realizations from simulation algorithms are inherently not locally accurate, but are a good metric for uncertainty due to random variability (Maironi, 2003). An example of this is presented in Figure 2-3.

Figure 2-3 shows three realizations of spatial soil water content in a region using the same set of input parameters. As can be seen in the realizations, there are a number of global features or trends that do not change considerably, but there is a high degree of local spatial variation. This is due to the changing path that populates the simulation grid; the algorithm takes known data and the variogram as inputs and honors the known data and the spatial correlation, including the presence of global-scale features. However, away from known data, there is a higher degree of
uncertainty, and the algorithm will use the known data and the spatial correlation as defined by the variogram in conjunction with previously-simulated values in order to generate a realization.

Figure 2-3: Three equally-probable realizations of soil water content over an area (Delbari et al., 2009).

Like kriging, variogram-based stochastic simulation algorithms have been modified in a number of ways to be more flexible to meet the varying needs of geoscientists. While sequential Gaussian simulation, which transforms variables into Gaussian space prior to simulation (e.g. Lin et al., 2001; Delbari et al., 2009) and direct sequential simulation are common (Caers, 2000; Soares, 2001), simulation has been modified similar to kriging to successfully allow for the use of
categorical variables (indicator simulation) (e.g. Journel and Isaaks, 1984; Deustch, 2006) or to incorporate secondary data (cosimulation) (Zhu and Journel, 1993; Le Ravalec-Dupin and De Viega, 2011; Azevedo et al., 2015; Nunez et al. 2017). In this research, however, cosimulation is extended to include cosimulation with both primary and secondary variables of the same data type, which has not been used in practice. This was done using heavy modification of the variogram of the secondary data and through decimation of these data, which has only been previously extensively used in the work of Koch et al. (2014), who showed that decimation is a viable as long as the spatial correlation of the data is maintained.

2.2 Other Statistical Tools for Quantifying Spatial Uncertainty

Numerous non-variogram based geostatistical and statistical methods have been applied to characterize uncertainty due to variability in spatial uncertainty for geotechnical problems. This section will introduce several of these, as principles from these approaches were incorporated into this research.

2.2.1 Random Fields

As discussed earlier, spatial uncertainty in geotechnical and geological engineering problems is considered to result largely from aleatory variability, or effectively random chance. As this is considered to be a random process from an engineering perspective, much work for probabilistic approaches in geotechnical engineering has utilized the concept of random fields to estimate spatial attributes when data are unknown (Fenton and Vanmarcke, 1990; Vanmarcke, 2010). A spatial random field can be constructed using a number of different stochastic approaches, with the end result being numerous realizations of spatial attributes that can be used
for probabilistic analyses, such as for a factor of safety calculation for slope stability (Griffiths and Fenton, 2004; Griffiths et al., 2009) or for estimating fluid transport (Mantoglou and Wilson, 1982; Sorbie et al., 1994; Oliver and Chen, 2011).

In their work, Griffiths and Fenton (2004) and Griffiths et al. (2009) utilized random fields for slope stability in soils to estimate the probability of slope failure and the factor of safety of a slope. The random fields in their work were built using the random finite element method, or RFEM. In RFEM, spatial soil properties, including the shear strength (Griffiths and Fenton, 2004) or the cohesion and friction angle (Griffiths et al., 2009) are assumed to have a normal or lognormal distribution with a mean and a standard deviation, as well as a correlation length. The correlation length is a metric that defines the anticipated spatial distance for which a random process is correlated (Griffiths et al., 2009). Figure 2-4 shows two random fields of soil shear strength for a slope, utilizing two different correlation lengths.

Much like variogram-based simulation algorithms, random fields are not locally accurate, but are better suited for a probabilistic analysis of problems with unknown spatial parameters.

Figure 2-4: Shear strength random fields and resulting equilibrium deformation for a soil slope modeled utilizing two different correlation lengths – (a) 0.2 units and (b) 2.0 units (Griffiths et al., 2009).


2.2.2 Markov Chains

Markov Chains have been used extensively in geology, as processes such as transition probability in sequence stratigraphy are considered to be Markov processes (Krumbein and Dacy, 1969). In sedimentary layers, Markov Chains are used to estimate the probability of changing state (lithology) or remaining in the same state based on the current state; these probabilities are viewed as matrices (Krumbein and Dacy, 1969). In many cases, the transition probability matrix in the vertical direction is informed by well logs, whereas the horizontal transition probability matrix is typically informed by geological maps or other indirect methods such as geophysical data (Elfeko and Dekking, 2001). A simple transition probability matrix using a Markov Chain in a carbonate-siliciclastic depositional environment is found in Figure 2-5.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red sandstone</td>
<td>1</td>
<td>0.00</td>
<td>0.11</td>
<td>0.06</td>
<td>0.39</td>
<td>0.03</td>
<td>0.05</td>
<td>0.21</td>
</tr>
<tr>
<td>Red siltstone</td>
<td>2</td>
<td>0.59</td>
<td>0.00</td>
<td>0.06</td>
<td>0.28</td>
<td>0.03</td>
<td>0.04</td>
<td>0.22</td>
</tr>
<tr>
<td>Red shale</td>
<td>3</td>
<td>0.08</td>
<td>0.11</td>
<td>0.00</td>
<td>0.26</td>
<td>0.03</td>
<td>0.00</td>
<td>0.21</td>
</tr>
<tr>
<td>Grey coarse to medium sandstone</td>
<td>4</td>
<td>0.11</td>
<td>0.14</td>
<td>0.07</td>
<td>0.00</td>
<td>0.04</td>
<td>0.05</td>
<td>0.27</td>
</tr>
<tr>
<td>Grey fine sandstone</td>
<td>5</td>
<td>0.08</td>
<td>0.11</td>
<td>0.05</td>
<td>0.26</td>
<td>0.00</td>
<td>0.04</td>
<td>0.20</td>
</tr>
<tr>
<td>Grey siltstone</td>
<td>6</td>
<td>0.08</td>
<td>0.11</td>
<td>0.05</td>
<td>0.26</td>
<td>0.03</td>
<td>0.00</td>
<td>0.21</td>
</tr>
<tr>
<td>Grey shale</td>
<td>7</td>
<td>0.10</td>
<td>0.13</td>
<td>0.07</td>
<td>0.31</td>
<td>0.04</td>
<td>0.05</td>
<td>0.00</td>
</tr>
<tr>
<td>Grey dolomite</td>
<td>8</td>
<td>0.10</td>
<td>0.15</td>
<td>0.07</td>
<td>0.33</td>
<td>0.04</td>
<td>0.05</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Figure 2-5: Transition probability matrix in a carbonate-siliciclastic sedimentary environment (Miall, 1973).

Based on the transition probability matrix in Figure 2-5, the probability of being in Red Sandstone and transitioning into a Red Siltstone is 0.11, while the probability of the inverse occurring is 0.59. As there are zeros along the diagonal, this was created using an embedded Markov Chain, whereas typical Markov Chains allow for non-zero probability values to be associated with the outcome of remaining in the same state (Elfeko and Dekking, 2001).
Markov Chains are regarded as conceptually simple, and have been used in a wide array of spatial uncertainty applications. While many of these include coal seams and other applications involving sequence stratigraphy (e.g. Jones and Dixon, 1976; Jones et al., 2005; Dindarloo et al., 2015; Ju et al., 2019), Markov Chains have been used to estimate spatial uncertainty in soils (e.g. Ching and Wang, 2016; Qi et al., 2016) and for estimating discrete fracture characteristics in rock (Snyder and Waldron, 2018).

Figure 2-6 shows an example of a simulated 2-D geologic section based on boreholes in a setting with three rock types. The vertical transition probability is calculated from the actual transitions in the boreholes, and the horizontal transition probability is calculated based on a regional understanding of the geology. A Monte Carlo draw is used to estimate the rock type based on the transition probabilities (Qi et al., 2016).

---

**Figure 2-6:** Rock type estimated using Markov Chains and borehole data (Qi et al., 2016).
2.2.3 T-PROGS

An extension of the Markov Chain approach for estimating spatial uncertainty is the development of Transition Probability Geostatistical Software, known as T-PROGS (Carle 1999). This software has been used extensively in geological and hydrological applications for spatial uncertainty estimation. This software utilizes the approach in Carle and Fogg (1996) and Carle and Fogg (1997) regarding Markov Chains with categorical variables, and has been utilized in a large number of applications, including hydrogeology (e.g. Fleckenstein et al., 2006; dell’Arciprete et al., 2012) and geotechnical engineering (Felletti and Beretta, 2009; Zetterlund et al., 2011; Koch et al., 2013). An example of three realizations from T-PROGS can be found in Figure 2-7.

Figure 2-7: Three realizations from T-PROGS (Fleckenstein et al., 2006).
2.2.4 Multiple Point Statistics

Much of the work in early geostatistics utilized variogram-based approaches for spatial interpolation or to stochastically generate realizations of subsurface geology. Variogram-based approaches, which were introduced previously, are considered to be two-point statistics, as variograms are created using pairs of data and the lag distances between them. However, one of the drawbacks of these approaches is the challenge of incorporating geologic information or known geometries, as variogram-based approaches solely consider spatial correlation.

This limitation of two-point geostatistics led to the development of multiple point statistics, where a variogram is no longer required, and spatial patterns are given in the form of a training image (Strebelle, 2002). Training images generated for a location identify patterns anticipated in the geologic environment of interest and use these patterns to stochastically interpolate between measured data (Caers and Zhang, 2004). An example of this approach is found in Figure 2-8.

Training images have been widely utilized in reservoir characterization (e.g. Strebelle and Journel, 2001; Caers, 2002; Mariethoz and Caers, 2014) when the geologic environment can be inferred by a geologist. For example, a training image for a deltaic environment or a fluvial environment will contain different geometries based on the anticipated geometrical relationships for that specific environment. This creates a powerful tool for replicating anticipated spatial correlation and patterns without an overly-smooth or unrealistic result that might be obtained using a variogram. Figure 2-8 shows the creation of a geologically-reasonable output using this approach with training images. Using sample data and geometries that are reasonable, training images allow for the creation of multiple realizations that honor both the obtained data and the expected geometries.
2.2.5 Decision Aids for Tunneling

Quantification of uncertainty at the project scale during the tunneling process can be performed using Decision Aids for Tunneling, commonly known as DAT. DAT is a computer-
based tool used to estimate cost and time information for a tunnel before and during construction (Einstein et al., 1999). This approach was developed to incorporate spatial geological and construction information for the project (Einstein et al., 1999), and was later modified to include resource allocation (Min and Einstein, 2016). Essentially, this tool was developed to help engineers simulate the construction of a tunnel a priori, and to evaluate and incorporate spatial uncertainty in geology and the nuances of construction to estimate the total cost of constructing the tunnel and the length of time the tunnel it will take to be constructed (Min et al., 2003).

One of the primary elements of DAT is the simulation of geology and ground class using geological information obtained from site investigation. First, the alignment of the tunnel is subdivided into different geological or geotechnical units, if applicable (Haas and Einstein, 2002). Within these geological units (and occasionally for the geological units themselves), the length of each lithology or geotechnical unit is simulated through the use of Markov Chains (Haas and Einstein, 2002; Min et al., 2008). The result is lengths of ground classes anticipated during tunnel construction. An example of this process is found in Figure 2-9.

![Figure 2-9: Ground classes resulting from Markov chain simulation of lithology and water inflow within ‘metamorphic rocks’ (Haas and Einstein, 2002).](image-url)
The construction simulation component relates the construction process to each anticipated ground class from the simulated geology. Information such as support installation and excavation method over the course of the alignment, which contain cost and time components, are used for estimating the overall anticipated length of construction and overall cost (Haas and Einstein, 2002). Another component of this includes simulation of resources such as muck and support materials (Einstein et al., 1999; Min et al., 2008; Min et al., 2016). These components rely on a reasonable estimation of spatial geological uncertainty, as uncertainty in these values contributes to error in estimation of construction time or cost, as well as the cost of resources.

DAT was originally intended to be used only prior to construction (Einstein et al., 1992), but more recent modifications have allowed for updating during construction with additional information (Min et al., 2016). This is primarily performed using face mapping data to better inform the geological transition probabilities in the Markov chains (Min et al., 2008). Example results from DAT can be found in Figure 2-10.

As shown in Figure 2-10, updating information (Phase II) narrowed the distribution of results in the case of both analysis methods that were employed. This is due to a better understanding of the uncertainty in the geological conditions anticipated in the tunnel during construction.

As opposed to DAT, this research used variogram-based geostatistics as the basis of estimation of uncertainty of ground conditions within a tunnel. This inherently does not give a measure of time and cost uncertainty for a tunnel, but can be used to aid engineers and geologists to better understand what factors will influence the cost and timeline for the project.
Figure 2-10: Example Time-Cost scattergram from DAT using prior information (Phase I) or updated information (Phase II) (Min et al., 2008).
CHAPTER 3
ASSESSMENT OF ROCK UNIT VARIABILITY THROUGH USE OF SPATIAL VARIOGRAMS

This article was published by Engineering Geology, 233, D. Lane Boyd¹, Whitney Trainor-Guitton², and Gabriel Walton³, Assessment of rock unit variability throughput use of spatial variograms, 200-212, Copyright Elsevier (2018).

Abstract

All rock units contain a certain degree of variability, which is an intrinsic property of the material. This variability can present itself in differences in mineralogy, grain size, grain shape, porosity, or a number of other ways. This presents a challenge when attempting to identify the number of specimens required in order to capture the geomechanical variability of a rock unit. For instance, while a homogeneous granite may only require a few specimens to characterize the spectrum of geomechanical behavior anticipated within the unit, a moderately to highly metamorphosed rock unit such as a gneiss, a schist, or a meta-igneous or meta-sedimentary rock may require a significantly larger number of specimens. This discrepancy can lead to over-testing, which induces an unnecessary excess cost, or under-testing, which could lead to an under-representation of the geomechanical variability possible within a rock unit. While previous works have considered robust statistical approaches such as Monte Carlo simulations and confidence interval analysis with large data sets, this work presents a practical empirical methodology of

¹ Graduate Student and author for correspondence
² Co-Advisor and Assistant Professor, Department of Geophysics and Geophysical Engineering, Colorado School of Mines, Golden, CO
³ Co-Advisor and Assistant Professor, Department of Geology and Geological Engineering, Colorado School of Mines, Golden, CO
assessing geologic and geomechanical variability by analyzing images and the respective uniaxial compressive strength (UCS) of core specimens. Once corrected for lighting irregularities and other deleterious influences, two-dimensional covariance maps and one-dimensional variogram samples are calculated for each rock core and used to extract several metrics for rock unit geologic variability. These metrics are then correlated to geomechanical variability based on UCS testing results. Ultimately, these correlations can be used to find the number of specimens required to estimate the rock unit’s mean UCS within a specified margin of error. This methodology allows one to quickly analyze core images of a specific rock type and evaluate how many specimens are required for testing.

3.1 Introduction

Rock, unlike most engineered materials, possesses a large degree of variability in physical characteristics, even within the same geologic or geotechnical unit. This degree of variability can be relatively small, such as within a homogeneous granite, or be large, such as within a highly foliated gneiss with mineral segregation and folding. This variability is often apparent in the mechanical properties of rocks, including their uniaxial compressive strength (UCS) values. Consider the five rock units outlined in Figure 3-1. These rocks range from a granite with low variability in both geologic characteristics and UCS values to a gneiss with a large degree of geologic variability and a large range of UCS values. It would be reasonable to assume that only a few samples of the granite would be needed in order to estimate its geomechanical properties, while a greater number of specimens would be needed for the gneiss. This is problematic when attempting to determine how many specimens to test to adequately characterize the geomechanical properties of the rock. ISRM guidelines suggest testing enough specimens to “adequately represent
the rock sample”, and indicate that the appropriate number “should be a function of the intrinsic variability of the rock”, but these guidelines are inherently vague (Ulusay, 2014). Currently, work done to attempt to quantify the number of specimens needed to characterize the rock unit has employed the use of robust statistical approaches on a large number of UCS specimens (Gill et. al., 2005; Ruffolo and Shakoor, 2009; Pepe et. al., 2017). While useful, these methods are not practical for an engineer to replicate for each rock unit encountered during a project. The aim of this work is to develop a reliable system that can be used in the field and in testing facilities to easily and objectively characterize the degree of a rock unit’s variability and to establish how many specimens are required to characterize its mechanical properties. This is achieved through the use of a geostatistical approach to characterize lab-scale specimen geologic variability, which can then be correlated with geomechanical variability.

3.1.1 Variability and Uncertainty in Engineering and Rock

In the geosciences, there exists an inherent problem of incomplete knowledge in the understanding of geologic and geomechanical characteristics of rock units, especially with respect to the subsurface. This incomplete knowledge is commonly ascribed to two different categories: epistemic uncertainty and aleatory variability (Bedi, 2013; Langford, 2013). Epistemic uncertainty is a result of unpredictability due to a lack of knowledge (Bedi, 2013). Eliminating this factor for subsurface properties is practically impossible, as it would require access to complete and fully accurate information. Aleatory variability, conversely, is unpredictability due to inherent randomness (Bedi, 2013). It is also a function of scale, with variability ranging from grain to field scale influencing unpredictability in the geologic and geomechanical aspects of an engineering project (Langford, 2013). This, unlike epistemic uncertainty, can be reasonably quantified given
sufficient data. Previous work done in epistemic uncertainty has typically consisted of updating probability values each time new data became available.

Much of the recent work has been performed using Bayes’ Theorem (Bardossy and Fodor, 2013; Miranda et. al., 2009) in addition to numerous types of numerical modeling and probabilistic methods (Day et. al., 2012; Day, 2017; Einstein, 1996; Jing, 2003). While quantifying epistemic uncertainty has been extensively studied and remains is a target for future research, this work will focus on aleatory variability only. In order to understand aleatory variability, large amounts of high quality data are required. This is in contrast to characterizing epistemic uncertainty, where the quantity or quality of data, or both, are lower (Bedi and Harrison, 2013).

3.1.2 Factors Influencing UCS Variability in Rock Cores

Rock strength has been extensively studied, often with the intent of ultimately applying laboratory results to evaluate the behavior of an in-situ rock mass as a whole (Ghazvinian, Diederichs, and Martin, 2012; Hoek and Brown, 1980; Hoek and Brown, 1998; Martin and Chandler, 1994). The most commonly used index tests for rock strength is the Uniaxial Compressive Strength (UCS), where an unconfined rock core of known dimension is subjected to uniaxial compression until failure. The failure mechanism of brittle rock specimens under such loading conditions is the development and eventual coalescence of microcracks within the core. Initial crack development occurs at the grain boundary, with inter-granular microcracks developing around the time of coalescence. As a result, grain scale inhomogeneity leads to inherent variability in the location, length, and density of microcracks (Wong, 1982).
<table>
<thead>
<tr>
<th>Granite</th>
<th>Limestone</th>
<th>Metagranite</th>
<th>Granodiorite</th>
<th>Gneiss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Variability</td>
<td>Low Variability</td>
<td>Moderate Variability</td>
<td>High Variability</td>
<td>Very High Variability</td>
</tr>
<tr>
<td>UCS 129-143 MPa</td>
<td>UCS 38-57 MPa</td>
<td>UCS 147-259 MPa</td>
<td>UCS 83-197 MPa</td>
<td>UCS 29-170 MPa</td>
</tr>
</tbody>
</table>

Figure 3-1: Comparison of UCS variability for five different rock units. Photos and UCS information were provided by Natural Resources Canada.
For example, in the Cobourg Limestone, microcracks develop around the larger mineral grains, and preferentially traveled through the argillaceous wisps and fossils (Day et al., 2017). In Westerly Granite, once intergranular cracks form, its behavior has been observed to be dependent on the mineralogy, with cleavage cracks forming in the feldspar grains and sub-vertical cracks forming in the quartz grains (Wong, 1982). This implies that heterogeneity in grain size, shape, and mineralogy will affect microcrack properties, which will further have impact on the geomechanical properties of the core.

One of the most important factors that influences uniaxial compressive strength is the presence of discrete macroscopic defects, such as mineralized veins. These large defects commonly lead to lowered strength values and are considered outliers in standard analyses, however these defects impact the overall geomechanical properties of the rock unit. Current ISRM guidelines for specimen preparation recommend that no cores should be tested that contain one or more discrete geologic weaknesses (Ulusay, 2014). However, recent work is beginning to incorporate the presence of veins in rock core in order to characterize a rock mass. These works attempt to understand the influence that discrete veins have on the overall rock mass, and utilize numerical simulations in order to fully characterize both the rock core specimens and the rock mass (Turichshev and Hadjigeorgiou, 2017; Vallejos et. al., 2015). Any attempt to characterize geologic variability in the context of geomechanical behavior must recognize and address the significance of these features.

3.2 Methodology

As previously stated, the aim of this work is to investigate the relationship between geologic and geomechanical variability, and use this relationship to evaluate the number of
specimens required to characterize the UCS of a given rock unit. Thus, two terms must be defined for a given rock unit: geomechanical and geologic variability. In this study, geomechanical variability is represented by the coefficient of variation of all of the UCS values for a given rock unit ($\kappa$). The coefficient of variation was selected (rather than the standard deviation) such that all the rock units studied could be easily compared despite their different mean UCS values. Quantifying geologic variability requires many steps, with the central calculation the 2-D covariance map which maps the spatial correlation of the rock cores.

Consider the flowchart in Figure 3-2 (page 32). A rock unit named ‘X’ has associated rock cores images A-E. These images are first put through an algorithm in MATLAB which generates one 2-D covariance map for each core. Next, each of these 2-D covariance maps is sampled in 64 directions, and the variability from each of the 2-D covariance maps is evaluated by calculating the variability of the 64 1-D samples. Finally, the coefficient of variation of metrics for A-E are calculated, each of which represents a measure of variability for rock unit ‘X’. Each step of this process is explained in greater detail in the following section.

### 3.2.1 Spatial Covariance and the Variogram

Spatial (2-D) covariance and variograms have been used for decades, with their primary use involving interpolating data when there is missing information. For example, 2-D covariances are used in Kriging analysis, when properties such as porosity or permeability are known at discrete boreholes with no information between the borehole locations (Bohling, 2005; Drew et. al., 2004). Consider N pairs of points that are separated by a distance $h$ in a direction $\theta$ degrees from horizontal. The general equation for the ergodic spatial (meaning the average is assumed to be the same over the entire area) covariance for all of these points is defined as Equation 3-1..
\[ \gamma(h) = \frac{1}{2N(h)} \sum_{N(h)} (z(u + h) - z(u))^2 \] 

(3-1)

In Equation 3-1, \( \gamma(h) \) is the value of the variogram at a distance of \( h \) meters, \( N(h) \) is the number of data pairs at a lag distance of \( h \) meters, and \( (z(u+h) - z(u))^2 \) is the squared difference of a pair of data separated by a distance of \( h \) meters.

3.2.2 Generation of a 2-D Covariance Map

In most variogram analyses, the input consists of discrete points that represent sampling locations such as boreholes. This means that there are large portions of the area of interest with missing information. In image analysis, however, there exists large amounts of data with each pixel representing a data point. With a typical resolution of 1024 x 1024, this equates to just over half a trillion data pairs, which leads to computational issues if all pair-wise covariances are computed using a standard spatial approach (Marcotte, 1994); as such, a faster system is required to evaluate the data pairs. The solution lies in computing the covariances in the Fourier domain, which decreases the number of data pairs for a 1024 x 1024 image from 5.5 x 1011 pairs to 46
In order to generate a 2-D spatial covariance map from a rock core image, some modifications are necessary to ensure that the results are representative of the true specimen grain structure. First, the image is converted to grayscale and subjected to filtering to remove lighting irregularities. This process will be described in detail in Section 3. Next, it is cropped to remove any portions of the image that should not be analyzed, such as portions that are not rock or portions of rock covered by writing. Once this is completed, the image is run through the aforementioned code with grayscale intensity values ranging from 0-255 as the variable of interest. The output is a spatial 2-D spatial covariance map and a map showing the number of data pairs at each location in the 2-D spatial covariance map.

Examples of the spatial covariance map and data pairs map for each step obtained through this process is shown in Figure 3-3. The number of pairs does not change due to the applied lighting correction, but decreases when the writing on the core is cropped out. Furthermore, the 2-D spatial covariance map slowly decreases in correlation at larger lag distances with each modification of the original image. As the image becomes less noisy, the faint vertical line on the spatial covariance map disappears. This is from the cropping of the left and right edge of the image, which shows that there are still faint lighting irregularities on the edge of the core, even after the lighting correction was applied. The final output of the 2-D spatial covariance map is therefore the most accurate representation of the rock core as a whole.
Figure 3-3: Process of obtaining a 2-D covariance map and N-pairs map.
3.2.3 Sampling and Parameterization of 1-D Variograms

While the 2-D covariance map provides reliable quantifiable spatial correlation data for the rock specimen that was processed, there is still a need to identify and calculate useful metrics from the 2-D image. This is done by sampling 1-D variograms from the 2-D covariance map. The variogram is chosen because of the number of available empirical models that exist for variograms as opposed to covariograms. As the data are numerous on the image, ergodicity is assumed and the variogram becomes Equation 3-2.

\[
\gamma(h, \theta) = C(0, \theta) - C(h, \theta)
\]

In Equation 3-2, \(C(0, \theta)\) is the maximum value of the covariance map and \(C(h, \theta)\) is the value of the covariance map at the lag distance \(h\) in the direction of \(\theta\).

To obtain 1-D variograms from the 2-D covariance map, a sampling algorithm must be used to ensure that the 2-D maps are sampled in an unbiased manner. Considering the 2-D covariance map as a simple Cartesian plane, the first sample is taken along the positive horizontal axis starting at the center and is transformed using the equation above. The remaining 63 1-D samples are taken radially at increments of 2.8125° clockwise around the image, with the final 1-D sample lying along the negative horizontal axis. It is unnecessary to sample around the entire 2-D covariance map, as the image is symmetrical (only 180° are sampled). Once the 64 variograms are sampled from the 2-D map, they are parameterized according to a nested variogram model. This nested model is expressed as Equation 3-3

\[
\gamma(h) = c \left[ 1 - e^{-\frac{h}{a_0}} \right] + m \cdot h
\]
The first half of this equation represents the empirical exponential model and the second half represents a standard linear model with a slope m. In the exponential model, the parameter c is representative of the behavior of sill (i.e. the maximum covariance value at large lag distances), while the \( a_0 \) (i.e. the practical range) parameter is one-third of the lag distance at which 95% of similarity is lost (Webster and Oliver, 2007). An example of this model can be found in Figure 3-4.

![2D Covariance Map](image1)

![1D Variogram Sample](image2)

Figure 3-4: Example 1-D variogram showing the location of \( c \) and \( 3a_0 \) for a Stanstead Granite specimen.

In Figure 3-4, the blue points represent the variogram values obtained the 2-D covariance values in the positive horizontal direction via Equation 3-1. The red curve is the parameterization of the empirical nested exponential-linear variogram model. The horizontal line represents the variance value of \( c \), which is directly related to the overall mineralogical variability of the sample. The vertical line represents the value of \( a_0/3 \), which is directly related to the extent of correlation.
in the direction along which the 1-D variogram was sampled (Webster and Oliver, 2007). To the left of the vertical line and, the exponential variogram model dominates, while to the right, the linear model dominates.

### 3.2.4 Compilation of Data for a Single Rock Unit

Throughout the process of creating and parameterizing 1-D variograms from one 2-D covariance map, numerous pieces of data are recorded regarding the fit, including all of the parameters from each 1-D variogram and the goodness of fit (as quantified by a least-squares $R^2$). Once this is completed for 64 1-D variograms for a rock core, confidence intervals, mean values, median values, and standard deviations are computed (both absolute and normalized to the mean value).

### 3.3 Removing Sources of Error from Image Analysis

One of the primary concerns when working with images is the sensitivity of data to variability induced by the process of photograph acquisition. During the covariance analysis performed for this study, a large number of factors were found to have potentially deleterious effects on the characterization of geologic variability. As the proposed methodology for quantification of geologic variability relies heavily on accurate digital representation of the visual attributes of core, these issues must be addressed.

#### 3.3.1 Lighting Irregularities

As the covariance generation relies on grayscale pixel values, irregularities in lighting can substantially alter the results. When acquiring an image of a core, the technician taking the
photograph may not consider the lighting of the image, especially if the irregularities in lighting are visually insignificant. However, even minor lighting irregularities can show false-positive correlations in the direction of the irregularity. A simple way to combat this is by homomorphic filtering. This type of filtering has been used to both remove lighting irregularities and homogenize multiple images with different lighting conditions, and in its simplest form, involves fitting a line to a cross-section of average pixel values and adding/subtracting the residuals (Delac, Grgic, and Kos, 2006; Seow and Asari, 2006). For the purposes of this work, code was generated that finds the average values of each column of pixel values on the core, fits a line to the data, and then adds or subtracts the residuals accordingly. The same is done for the rows.

Consider the images Figure 3-5. The top row is the original image of a metabasalt sample with a distinct mineralized vein and its associated covariance with the data from a single row of the grayscale grid. Due to lighting irregularities, the left side of the core is darker than the right side of the core, leading to the vein becoming unrecognizable in the covariance. Once the homomorphic filtering is applied (bottom row), the effect of the vein is now visible.

It is not immediately clear from the raw image how substantial the lighting irregularity is on the core. The vein, which is only a few millimeters in width, is not visible due to the dominant trend associated with the lighting irregularity, which shows that the source of light is coming from the right side of the image and creating a shadow on the left side of the core. When homomorphic filtering is applied, the effect of lighting is nearly removed. However, there does exist some latent lighting irregularities on the edges of the image, which creates a faint increase in correlation in the vertical direction. This edge effect is removed by subsequently cropping off the image edges. Typically, 10% off each side is cropped. This was dependent on the specimen, with several requiring 5% more off the sides to remove the effect of the lighting anomaly.
One of the main assumptions made in this section is that the images from each rock unit were taken under similar lighting conditions such that the average grayscale intensity of the image is similar. In the case of all rock units studied here, images of the same rock unit were taken in the same position with consistent lighting. It should be noted for future application of the proposed methodology that if the lighting conditions vary significantly from specimen to specimen, this has the potential to introduce error in the resulting variability metrics.
3.3.2 Cropping out unusable portions of images

Rock core images of UCS samples typically contain more than only rock. Commonly, the core image is taken on a table, in a core box, or on the testing apparatus. Additionally, there commonly exists writing on the core, usually for specimen identification purposes. To accurately capture rock specimen variability, any portion of the image that does not contain images of the rock core under analysis should not be input into the code. For this reason, code was developed to allow manual selection portions of the image that are not desirable for analysis, which are then excluded in the modified code from Kryiakidis (2005a).

Consider the images in

Figure 3-6. Stanstead Granite Brazilian Tensile Strength (BTS) cores were used for analysis of geologic variability because images of the associated UCS specimens were not available. Each BTS specimen was taken from the same location as a corresponding UCS specimen. However, as all the BTS specimens were taken after testing, each of the specimens had a large tension crack through the diameter. Additionally, the edges and corners of the image contained images of the background wall and table. In the original image, there are large spikes of correlation in the top and bottom of the 2-D covariance map. This is due to the presence of the corners, as there is extremely high similarity between large lag distances that connect the corners. There is also a vertical feature in the covariance map, which represents the tension crack induced during the BTS test. This can be seen more clearly in the second row of images, where the corners have been cropped out, but the tension crack is still visible.

The covariance map better represents the rock, with the exception of the tension crack, which leads to higher correlation along the vertical direction than would be expected given an isotropic granite. Removing the tension crack leads to a reasonable covariance map given visual
inspection of the rock. The zone of correlation around the origin is roughly circular, indicating
similar levels of correlation in all directions. There are numerous small patches of higher
correlation at locations away from the origin. These, unlike the large spike at the origin, indicate
marginal correlation at larger lag distances, and are related to pairs that fall on the same grain type,
as expected in a granite (such as pairs that start and stop on a feldspar, a quartz, or bioite grain).
The outside portion of the covariance map, represented by the pale blue color, indicates values
where the covariance map could not be calculated due to lack of data pairs.

In order to avoid irregularities due to insufficient data, the lower limit of data pairs for
which the covariance map is calculated was set at 15,000 pairs which typically represents 3-5% of
the number of pairs at zero lag (the covariance map origin). This number was selected arbitrarily
based on the amount of irregularities near the edge of the image, which contain the lowest number
of data pairs and the most significant amount of irregularity.

3.3.3 Presence of Drill Marks

In some cores, drilling induces superficial horizontal laminations due to the rotational
motion of the cutter bit. These typically create higher horizontal correlation in the covariance map,
as shown in Figure 3-7. The rock, a limestone with calcite veins, has many of these parallel
horizontal drill marks as well as calcite veins that could have an effect on the UCS. When
calculating the covariance map of this image, it can be seen that the horizontal drill marks
dominate, even though they are not part of the rock. Of all of the rock units tested in this analysis,
only three have specimens that are affected by this phenomenon. As these specimens represent a
small sample of typical cores, and as the 1-D samples affect the results, the cores cannot be used
with this approach for quantifying geologic variability.
Figure 3-6: The effect of cropping on a Stanstead Granite BTS specimen.
Figure 3-7: The effect of horizontal laminations from drilling on the rock core and associated covariance map.

3.3.4 Wet vs Dry Imaging

The visual appearance of some rock core is significantly different when it is dampened. In most cases, rocks that are dry appear to be more homogenous than actuality, and can be a result of clay or other materials in the matrix between the grains, or due to residue on the core. Consider the gneiss in Figure 3-8. The top row is a core sample that has been dampened, while the bottom row is the same specimen that has been allowed to air dry.

As can be seen in Figure 3-8, many of the different minerals, especially the darker minerals, are much more vivid when the rock core is damp. This is evident in the covariance map of the image where the damp specimen shows significantly higher correlation due to the increased contrast of the minerals. In the covariance map of the image of the dry specimen, the correlation
is significantly noisier, which presents a less accurate representation of the rock core. While this phenomenon typically affects rocks with clay minerals in the matrix, care should be taken to ensure that the results are not negatively affected by the presence of residue or discoloration. Most importantly, a consistent approach should be applied for all specimens of the same rock unit.

![Dry vs. damp Smaland Granite and associated covariance maps.](image)

**Figure 3-8:** Dry vs. damp Smaland Granite and associated covariance maps.

### 3.4 Analysis and Results

From the procedure outlined in Section 3, two primary outputs from 1-D variograms were modified, which provided meaningful relationships between geologic and geomechanical
variability. With a high number of geologic units tested, these meaningful relationships show the robustness of this approach, as well as providing different metrics with which to compare this variability based on the type of failure and the presence of structural defects. The image and geomechanical datasets presented in this work were collected by Natural Resources Canada (CANMET), SP Swedish National Testing and Research Institute, Colorado School of Mines Earth Materials Institute (EMI), and the University of Vigo Natural Resources and Environmental Engineering Department.

3.4.1 Analysis of $a_0$ and $c$

The parameterization of the 1-D directional variograms identifies two important variables that constrain the curve in the ‘exponential region’: $a_0$ and $c$. The $a_0$ parameter, as mentioned in the methods section, is related to the correlation length (in millimeters) up to which the exponential model of the variogram dominates (Bohling, 2005). The $c$ value, is related to total variability as captured in the pixel grayscale intensity values, which can be related to the total heterogeneity of the specimen.

To establish variability in the $c$ parameter, the coefficient of variation of the mean $c$ of each core for a given rock unit was calculated. Recall that after the 64 1-D variograms were parameterized for a given rock core, the mean $c$ of that rock core’s 64 1-D variograms was obtained. Once this was completed for all of the rock cores for a given rock unit, the mean of all of the mean $c$ values for that rock unit were obtained and the standard deviation of the individual specimen means was normalized with respect for the overall mean rock unit mean to obtain the coefficient of variation. The equation for the coefficient of variation for $N$ specimens of a given rock unit can be found Equations 3-5. The mean $c$ value for a rock core specimen is used as a
metric for the relative mineralogical homogeneity of the specimen compared to other specimens from the same rock unit. These are given as Equation 3-4 through Equation 3-6.

\[
Specimen \ Mean = \bar{x} = \frac{1}{64} \sum_{i=1}^{64} c_i
\]  
(3-4)

\[
Rock \ Unit \ Mean = \mu = \frac{1}{64N} \sum_{i=1}^{64} \sum_{j=1}^{N} c_{(i,j)}
\]  
(3-5)

\[
CV, \ rock \ unit = \theta = \frac{1}{\mu} \sqrt{\frac{1}{N-1} \sum_{j=1}^{N} (c_i - \bar{x}_j)^2}
\]  
(3-6)

In these equations, \(i\) represents a given 1-D variogram, \(j\) represents a given rock core specimen from \(N\) total core specimens for a rock unit. As opposed to simply looking at variability in overall core specimen heterogeneity, the \(a_0\) parameter is ideal for investigating variability that includes the fabric of the rock, including folding, discrete defects, bedding, variation in grain size and shape between cores. As opposed to working with the mean value for each specimen, this parameter is investigated considering variability within each rock core specimen.

Consider the metagranodiorite specimen in Figure 3-9. There is an obvious trend in the fabric from the bottom left to the top right of the image. This is due to metamorphism, at which time the igneous protolith was subjected to a large degree of stresses at a high temperature, which caused a rearrangement of the grains, leading to light foliation within the rock. Many specimens from this same geologic unit possess similar anisotropy that led to failures along this orientation. The orientation of the fabric of the rock is identified by the covariance map, with higher correlation values along the lag vectors that are along the fabric, and much lower correlation at lag distances...
at any other angle, especially perpendicular to the fabric. This means that $a_0$ is significantly higher along the orientation of the fabric, and the standard deviation of $a_0$ from the 64 sub-sampled variograms is relatively high.

![Covariance map of a metagranodiorite sample with noticeable orientation in fabric.](image)

Figure 3-9: Covariance map of a metagranodiorite sample with noticeable orientation in fabric.

In the proposed workflow, in order to find the coefficient of variation of the specimen-level coefficients of variation of $a_0$ for a rock unit with $N$ specimens, first the coefficient of variation of the $a_0$ values for one specimen is calculated (Equation 3-4). The use of normalized values is more robust than using raw standard deviations, as slight variations in cases with very low mean $a_0$ values would be overstated. Completing this task for each specimen gives a metric of variability for each rock core (Equation 3-5). The final step is to use these values to evaluate the overall rock
unit variability. This is done by taking the standard deviation with respect to the mean of the individual specimen values (Equation 3-6).

### 3.4.2 Results from Analysis

As previously stated, the two parameters from the geologic variability methods were the variability in the mean heterogeneity (total variability) in each rock core ($c$), as well as the variability in each rock core’s spatial correlation variability ($a_0$). The relationship of these two parameters to the geomechanical variability in UCS is illustrated in Figure 10 (a and b). The top image represents the relationship between the variability in the coefficient of variation of $a_0$ ($\Gamma$) vs. variability in UCS ($\kappa$) for all rock units investigated, while the bottom image represents the variability between the heterogeneity of a rock ($\theta$) vs. the variability in UCS values ($\kappa$).

The fits shown in Figure 10 are weighted based on the square root of the number of specimens of the UCS values, which is related to the standard weighting for a coefficient of variation correlation found in literature (Solon et al., 2013; Langford, 2013). The confidence intervals were taken from McKay’s approximation of confidence intervals for the coefficient of variation (Payton, 1996), and are given as Equation 3-7.

\[
\left( \frac{X^2 (1 + CV^2)}{nc^2} - 1 \right)^{-0.5} \leq (\Gamma, \theta \rightarrow \kappa) \leq \left( \frac{X^2 (1 + CV^2)}{nc^2} - 1 \right)^{-0.5}
\]

In Equation 3-7, $X^2$ are the upper and lower $\alpha/2$ percentiles in the Chi-squared distribution for $n-1$ degrees of freedom and a sample coefficient of variation $CV$ (Payton, 1996). Lastly, the
curve was weighted orthogonally, as opposed to a simple regression, which only considers error in the y. The orthogonal regression, instead, takes error in the x-direction into account as well.

Figure 3-10 shows the relationships between two variables quantifying geologic variability (which are derived from $a_0$ and $c$) and geomechanical variability (as quantified by $\kappa$.)

### 3.4.3 Discussion of $a_0$ – UCS variability relationship ($\Gamma$ vs. $\kappa$)

The relationship between $\Gamma$ and $\kappa$ was identified to be of the form $y(x) = a^x$, in which $\kappa$ does not appreciably increase until the geologic variability surpasses a certain threshold. In Figure 10a, the data points with the lowest geologic and UCS variability are granitic igneous rocks. This is unsurprising due to the fact that granitic rocks are formed from a slowly-cooling melt, with certain minerals precipitating out of solution at higher temperatures and other minerals precipitating around them until a highly-interlocked granular structure is reached (Bowen, 1922). This structure generates fairly consistent geomechanical properties due to the consistent mineralogy, grain size, orientation, and distribution of grains. The differences in geologic variability between granitic rocks arises from slight variability in the mineralogical makeup of different specimens. However, this variability in comparison to other rock units remains low.

Figure 3-10a shows an exponential relationship between geologic (correlation length) and geomechanical variability for all specimens, regardless of the number of specimens that failed along defects or structure for a given rock type. Figure 3-10b, alternatively, shows a linear relationship between geologic (total heterogeneity) and geomechanical variability for the subset of rock types for which no specimens failed along discrete geological structures. Further discussion will be present in the following sections.
Figure 3-10: (a) Relationship between $\Gamma$ and $\kappa$ for all cases ($y = 1600 \times / 100$); (b) Relationship between $\theta$ and $\kappa$ for cases with no structural defects ($y = 0.55x$); * - no data were collected regarding failure type.
The sedimentary rocks studied generally have both higher geologic and geomechanical variability when investigating the relationship between $\Gamma$ and $\kappa$. Sedimentary rocks, unlike granitic rocks, are deposited over time, with variability in water level, climate, tectonic environment, and location in the basin affecting its geologic properties (Klein and Beukes, 1989; Lubeseder et. al., 2010; Mackenzie and Pigott, 1981).

Figure 3-11: Comparison of three sedimentary units.
Figure 3-11 shows six different sedimentary cores from three different sedimentary rock units that were used in this study. There is a certain degree of variability in rock units that would not likely be present in a rock unit formed from nearly-homogeneous grains. The Cobourg Limestone contains argillaceous whisps from tidal deposition of mud, as well as fossils and burrowing from Ordovician invertebrates (Day, 2017; Ghazvinian et al., 2013). These diagenetic features are variable throughout the length of the core due to the influence of invertebrate lifeforms, amount of mud deposited, and paleoclimate and paleoweather at the time of deposition. The Lyons Sandstone varies spatially between a quartz arenite and an arkosic sandstone and contains local variability in grain size, shape, and mineralogy. This can also vary vertically, as facies changes through time are expected. In these cores, quartz and feldspar grains were deposited in horizontal laminations, with occasional lithic fragments interbedded in the layers. This introduces variability, both geologically and geomechanically, as the different properties of the grains influence the strength of the core. Finally, the Resolution Conglomerate contains obvious variability in grain size, which is related to the velocity of the depositional environment of the transporting fluid. As with the Lyons Sandstone, variability in grain properties affect the geomechanical properties of the overall rock core.

Finally, the rock units that have the highest geologic and geomechanical variability are the highly-metamorphosed rocks. These rocks have been exposed to intense heat and/or a high degree of stress. Such exposures can induce recrystallization of minerals if partial melting is reached, realignment of grains, and occasionally folding if paleostresses are sufficiently high. Examples are shown in Figure 3-12.

Higher variability of $a_0$ indicates that several of the rock cores from a rock unit had a high degree of variability in the correlation length. This may be due in part to the presence of discrete
veins or other defects in the core, or due to variation in the fabric of the rock. In case of the sedimentary rocks, such as the Lyons Sandstone in Figure 3-11, each of the rock cores possesses a high degree of variability in $a_0$ due to the presence of the horizontal laminations. This leads to a higher $a_0$ value in the horizontal direction. However, as each of the cores exhibit similar horizontal laminations, its effect is nullified.

Figure 3-12: Comparison of three metamorphic units.
3.4.4 Discussion of $c$ – UCS variability relationship ($\theta$ vs. $\kappa$)

The relationship between $I$ and $\kappa$ includes all of the rock cores tested for each unit, including those that fail along structural defects, such as discrete veins or foliation. This is because those rock units which have a higher $I$ typically contain these features, as the presence of discrete features leads to large variations of $a_0$ within specimens (and between specimens when only some specimens contain such features). However, while there is a growing trend to incorporate data from rock cores with defects when studying a rock unit’s UCS, this work still aims to provide insights relevant to UCS testing programs that treat test results impacted by such defects as invalid.

The $\theta$ parameter reports the overall variability of the total heterogeneity of the rock core, which is related to mineralogy and grain structures. When plotted against $\kappa$, the variability of the $\theta$ parameter increases linearly for cases that do not contain specimens that possess structural defects. Those specimens that do contain structural defects lie to the left and above the regression line, indicating that for a given level of total mineralogical variability (as characterized by $\theta$), rock units that contain structural defects tend to show higher $\kappa$ values. This is consistent with the meaning of the $\theta$, as cores with structural defects would not appear to have a particularly high degree of total mineralogical heterogeneity. Consider the two metamorphic rocks in Figure 3-13. One of the rock units is a metabasalt, which contains a large amount of structural defects infilled with quartz. The other is a metagranite that has undergone regional metamorphism.

The metabasalt unit contains both specimens that fail along structure, such as the lower specimen pictured (UCS = 64 MPa), and specimens that do not fail along structure, such as the upper specimen pictured (UCS = 399 MPa), but does not appear to show a large degree of variability mineralogically. The only significant difference between the two images is the presence
of a discrete quartz-filled fracture, which increases the geomechanical variability. However, this has little effect on the $\theta$ parameter with the exception of the influence of some higher-valued pixels which represent the presence of quartz existing within the rock core. Conversely, the metagranite pictured on the right shows a notably higher modal percentage of quartz in the top core, despite originating from the same rock unit. As quartz-rich specimens are likely to be stronger than specimens that contain more biotite, pyroxene, and hornblende, the variability in the modal percentages minerals are likely to influence the geomechanical variability of the unit.

Figure 3-13: Comparison of a rock that commonly fails along structure versus one that has no structural failures.
3.5 Relationship to the number of specimens required

Once geologic variability has been quantified through image analysis and $\kappa$ has been estimated through one of the correlations shown in Figure 3-11, it is possible to statistically characterize one’s level of confidence in the mean UCS that would be obtained from testing a given number of specimens. Conversely, with the same information, the number of specimens required to estimate the mean UCS within a given level of accuracy at a stated level of confidence can be estimated.

The number of specimens needed to estimate the population mean of a parameter of interest within a certain range at a specific level of confidence has been well-established (Reiland 2008) as Equation 3-8.

$$n = \left( \frac{z_{\alpha/2} \ast \kappa}{E} \right)^2$$

(3-8)

In Equation 3-8, $z_{\alpha/2}$ is the z-score associated with a confidence level of $1-\alpha$ that the population mean has been estimated within a margin of error $E$ (expressed as a percentage of the population mean) given a coefficient of variation $\kappa$. The z-score is given as 1.96 in all of the remaining calculations as the level of confidence in the standard deviation is selected at a certain confidence, implying that the standard deviation of the ‘true’ population distribution of UCS values is known $a priori$ or the confidence level is based on the quantity and quality of known information.

Once geologic variability for a given rock unit has been captured through image analysis, the $\kappa$ can be estimated from Figure 3-10. The geologic variability parameter used can be either $\Gamma$, $\theta$, or both, depending on whether or not specimens with structural defects are included in the
analysis. Furthermore, while the mean regression relationship can be used in either case, the upper limit of the 67% or 95% prediction intervals can be used, resulting in 83% or 97.5% confidence that $\kappa$ has not been underestimated.

Based on the equation introduced above, 2-D contour maps were developed that show the number of specimens required to achieve a target specimen mean within an error margin $E$ for a given $\kappa$. Contour maps corresponding to 95% and 99% confidence levels is shown in Figure 3-14 (page 59). Note that if the number of specimens falls between values, the number should be rounded up to the next integer. Per ISRM guidelines (Ulusay, 2014), no fewer than four specimens should be tested under any circumstances. Table 3-1 shows five rock units with different $\kappa$ values at 95% confidence and multiple acceptable error percentages from the mean UCS, $E$. The rocks presented are the same rocks that were shown in Figure 3-1 in the beginning of Chapter 3.

Table 3-1: Minimum number of specimens required* to know the true mean of UCS within 10%, 20%, and 30% of the true value at 95% confidence

<table>
<thead>
<tr>
<th>Rock Unit</th>
<th>$\kappa$</th>
<th>$E = 10%$</th>
<th>$E = 20%$</th>
<th>$E = 30%$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Granite</td>
<td>0.10</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Limestone</td>
<td>0.20</td>
<td>16</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Metagranite</td>
<td>0.30</td>
<td>35</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>Granodiorite</td>
<td>0.35</td>
<td>48</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>Gneiss</td>
<td>0.50</td>
<td>97</td>
<td>25</td>
<td>11</td>
</tr>
</tbody>
</table>

* The authors recommend testing no fewer than four specimens in any instance

As can be seen in Table 3-1, with increase in the acceptable deviance from the mean, the number of specimens required to test decreases substantially, especially for rocks with higher coefficients of variation of UCS.
Figure 3-14: (a) 2-D contour map of the number of specimens required to obtain a mean UCS value within a certain percentage $E$ of the true mean at 95% confidence; (b) the same figure at 99% confidence.
3.6 Conclusions

Unlike engineered materials, rock possesses a high degree of inherent variability. This results from variability in mineralogy, diagenetic effects, metamorphism, and other geologic processes. This presents a challenge when attempting to understand the number of specimens required to adequately characterize the geomechanical variability of a rock unit. This work presents an empirical methodology using spatial and directional covariance maps and variograms from rock core images in order to assess the geologic and geomechanical variability within a rock unit. The information obtained from the proposed process can be used to relate the estimated geomechanical variability to the number of specimens required to adequately capture the mean UCS of a given rock unique.

The established method, while imperfect due to a large amount of subjective user input with regard to the geologic variability, is validated by the fact that two metrics of geologic variability correlate well with geomechanical variability. The developed approach can be applied in geotechnical engineering to understand the number of specimens needed to be sent to the lab for testing. This ensures that not too few are being sent, which would require additional cores to be send subsequent to testing, and prevents a large amount of over-testing for a relatively homogeneous geologic unit. This additionally allows for a certain level of confidence to be set that is based on the purpose of the testing and the confidence level required for the project.

This method could be used to develop a tool for geologists selecting core specimens for testing. For example, the worker could take photos of the core, the geological variability could be calculated and correlated to geomechanical variability, and a necessary number of core specimens for testing could be output based on user preferences (e.g. error tolerance).
CHAPTER 4
QUANTIFYING SPATIAL UNCERTAINTY IN ROCK THROUGH GEOSTATISTICAL INTEGRATION OF BOREHOLE DATA AND A GEOLOGIST’S CROSS-SECTION

This article was accepted by Engineering Geology, 2019, D. Lane Boyd4, Gabriel Walton5, and Whitney Trainor-Guitton6, Assessment of rock unit variability through use of spatial variograms, pages unknown7, Copyright Elsevier (2019).

Abstract

In order to increase understanding of spatial uncertainty in subsurface conditions for problems in geological and geotechnical engineering, this study develops a geostatistical approach for incorporating a geologist’s interpreted cross-section to quantify spatial uncertainty of lithology. In this work, the Sequential Indicator Cosimulation (COSISM) algorithm and the information theory concept of entropy are utilized to generate uncertainty models. A sensitivity analysis is used to study how both borehole data and the cross-section are incorporated into the algorithm, which is quantified using entropy as a metric for uncertainty in geologic unit boundary locations. A verification methodology is then developed to identify the most appropriate combination of input parameters for combining these two data types; recommendations from the verification methodology are made by comparing the calculated entropy to accuracy at the global scale on a point-by-point basis. This methodology is then used to investigate how calibrated input parameters vary due to differences in the input cross-section based on different levels of understanding of the

4 Graduate Student and author for correspondence
5 Co-Advisor and Assistant Professor, Department of Geology and Geological Engineering, Colorado School of Mines, Golden, CO
6 Co-Advisor and Assistant Professor, Department of Geophysics and Geophysical Engineering, Colorado School of Mines, Golden, CO
7 This volume of Engineering Geology has not been published
subsurface geology by iteratively providing a geologist with additional borehole data after a cross-section is developed. Results show that two of the parameters, variogram range and the number of conditioning data, can be used to most directly reflect geologist understanding in the simulations. Finally, this study provides guidelines for incorporating a geologist’s understanding of a geological environment into this modeling framework, as results show cross-sections that are locally accurate are better represented by different parameters than cross-sections that are not locally accurate but are globally accurate. These guidelines are intended to aid geologists and engineers in understanding how to incorporate a geologist’s interpretation of the subsurface lithology or categorical geotechnical parameters at a particular worksite into a structured geostatistical framework.

4.1 Introduction

In geotechnical engineering projects in the subsurface, geologists and engineers use geologists’ cross-sections (hereafter referred to as ‘cross-sections’) with some level of uncertainty as a guide to better understand the geological and geotechnical conditions at a particular site. At the preliminary investigation stage, a cross-section is used to make decisions about future site investigation for a particular project and for preliminary decision making regarding a project (Fookes et al., 2000; Martino et al., 2004). These are also used to make geotechnical cross-sections to support engineering decision-making (de Vallejo and Ferrer, 2011). As these cross-sections are important for prediction of factors that contribute to the success of a project, a lack of understanding with respect to spatial uncertainty in these cross-sections can lead to cost overruns, setbacks, and other negative project outcomes.
Spatial uncertainty in geology and geotechnical engineering has become an increasingly popular topic of study in recent years (e.g. Langford, 2013; Bond, 2015; Wang and Cao, 2013; Wang et al., 2015; Wang et al., 2016), as spatial uncertainty in subsurface geology is driven not only by lack of information, but by aleatory (random) variability caused by the natural processes that created the geology (Bedi, 2013). Many researchers have attempted to characterize geological, geotechnical, and structural model uncertainties in a Bayesian framework, updating parameters in various algorithms using observed performance data (e.g. Zhang et al., 2009; Wellmann et al., 2017), or have utilized random fields to populate grids of data properties including lithology or geotechnical characteristics (e.g. Tacher et al., 2006; Li et al., 2016; Wang et al., 2017; Wang et al., 2019). Many of these studies incorporate different software packages and algorithms, and vary input parameters to perform a sensitivity analysis of spatial parameters (e.g. Cho, 2007; Lan et al., 2004; Wellmann et al., 2014; Randle et al., 2018; Zhou et al., 2003). Other researchers have attempted to quantify human error in quantifying spatial uncertainty due to variability; Lark et al. (2014) compared the cross-sections of 28 geologists presented with the same information, and found that there are notable differences in the cross-sections produced by each geologist.

This study focused on the development of a geostatistical approach to combine borehole data and a cross-section for the purposes of quantifying spatial uncertainty using an entropy map. In particular, the approach used the variogram-based algorithm Sequential Indicator Cosimulation (COSISIM) to generate 100 equally probable realizations of the subsurface geology. The resulting entropy map produced is not a deterministic estimation of subsurface geology, but is a quantification of spatial uncertainty in subsurface geologic boundaries. Unlike algorithms like kriging that provide one deterministic solution for a set of input parameters, simulation algorithms like COSISIM provide poor local estimations, but are a good measure of global (larger-scale)
spatial variability (Marioni, 2003). Thus, while a single realization from a simulation algorithm is a poor estimation of the geologic boundaries in the subsurface, the aggregate of numerous realizations with the same input parameters provides a good estimation of spatial uncertainty due to geologic variability. This provides a powerful tool for estimating uncertainty in geologic boundary or lithology location, with targeted applicability to geotechnical engineering and underground construction/excavation problems.

The first part of Section 4.2 provides background information on the well-established COSISIM algorithm and the use of cokriging and cosimulation algorithms as a whole. The second part of Section 4.2 describes parameter and modeling decisions made for generating COSISIM realizations with the unique input combination used in this study in the context of a simple test case. Section 4.3 presents results from a comprehensive sensitivity analysis using the simple test case presented in Section 4.2. Section 4.4 then presents a more complex geology to illustrate the accuracy of the spatial uncertainty quantification approach used. In this case, we iteratively supplied a geologist with additional boreholes from the same geology, with each iteration increasing the geologist’s understanding of the geology, and used the cross-sections developed at each iteration as inputs for separate simulations. This approach illustrates which combinations of simulation inputs are most appropriate to utilize given the different levels of geological understanding used to develop the cross-sections.

4.2 Project Methodology

This work uses COSISIM with borehole data and a cross-section as inputs to create entropy maps for quantification of spatial uncertainty due to natural aleatory variability. This study considers borehole data to be primary data, and a cross-section to be secondary data. As in
traditional cokriging and cosimulation algorithms, borehole data are precise, measured, and sparse, whereas cross-section ‘data’ are inherently imprecise, but spatially extensive. A flowchart of the proposed methodology is provided as Figure 4-1. After introducing the history and basis of simulation algorithms and specifically outlining the COSISM algorithm, Sections 4.2.3 and 4.2.4 introduce how the cross-section for parameter sensitivity was developed and how the entropy metric used to quantify uncertainty is calculated. The remainder of Section 4.2 discusses various modeling decisions, including the selection of variogram parameters, the decimation of the secondary data for appropriate incorporation into the COSISM algorithm, and the establishment of a sufficient number of realizations for each iteration of COSISIM. With these modeling decisions established in Section 4.2, the sensitivity analysis of selected parameters is introduced in Section 4.3 prior to the verification presented in Section 4.4.

**4.2.1 Cosimulation and Cokriging Algorithms**

COSISIM is a simulation algorithm that incorporates a user-defined variogram as well as primary and secondary data in order to populate a grid with specific values of categorical data. While COSISIM simulates categorical variables (e.g. lithology), other cosimulation and cokriging algorithms simulate continuous properties (e.g. porosity, permeability). Cosimulation and cokriging algorithms are used extensively in subsurface estimation problems when two types of data are available in a modeling region, and their joint distribution and relationship is known or assumed (Zhu and Journel, 1993). The primary data are typically measured data and are sparse. The secondary data, which have a lower resolution but are spatially extensive, provide information regarding the spatial continuity of the area. In most reservoir characterization applications, the primary data are sampled well data, whereas the secondary data are seismic or other geophysical
data, though other indirectly measured data types can be used (Horta and Soares, 2010; Le Ravalec-Dupin and De Viega, 2011; Azevedo et al., 2015; Nunez et al. 2017). A general established guideline for choosing secondary data is that the secondary data should be chosen as ‘the data with the most continuous spatial pattern’ (Horta and Soares, 2010).

![Flowchart of the development of the proposed methodology.](image)

Primary and secondary data are assumed to have some level of correlation, but perfect correlation is rarely assumed. The original development of the cokriging and cosimulation
approaches required one or more cross-variograms between the primary and secondary data in addition to the variograms for each the primary and secondary data, which were challenging and time-consuming to generate (Zhu and Journel, 1993). To simplify this, the Markov-Bayes algorithm was developed, which removed the need for tedious cross-variograms by relating the variogram of the secondary data directly to the variogram of the primary data (Zhu and Journel, 1993; Goovaerts and Journel, 1995). As inputs, this approach requires only the variogram of the primary data and a single value between 0 and 1, referred to as the Markov-Bayes parameter. This parameter describes the relationship and correlation of the primary and secondary data, with values near 0 indicating low correlation, and values close to 1 indicating high correlation (Zhu and Journel, 1993).

4.2.2 Sequential Indicator Cosimulation Algorithms

In order to produce a single realization, COSISIM populates a grid of categorical variables by defining a random path around the grid, traveling to each location on the grid using that path, and simulating categorical values using nearby primary and secondary data, as well as previously-simulated variables. The initial step in COSISIM involves using the user-defined variogram to calculate covariance values and the cross-variogram. Finding the covariance involves a simple transformation of the user-defined variogram model. With the variogram model, $\gamma(h)$, defined at each lag distance $h$, the covariance model is found using Equation 4-1.

\[
\gamma(h) = C(0) - \gamma(h)
\]  

(4-1)
In Equation 4-1, $C(h)$ is the covariance at a lag distance $h$, $C(0)$ is the variance of the data (also known as the sill of the variogram), and $\gamma(h)$ is the variogram defined at a lag distance $h$. As full indicator cokriging is used, one variogram for each categorical variable must be calculated (Hill, 1998). From this point onward throughout the algorithm, covariance values are used as opposed to the variogram.

While the cross-variogram, which describes how the primary and secondary data vary in space with respect to each other, can be calculated directly, a simplification of cross-variogram calculation is adopted by utilizing the Markov-Bayes parameter and the primary covariance. Using the Markov-Bayes model, the cross-variogram is a scaled version of the primary covariance. The Markov-Bayes parameter captures the correlation between the primary and secondary data. This is mathematically described by Equation 4-2.

$$C_{12}(h) = \frac{c_{12}(0)}{c_{11}(0)} C_{11}(h) = M \ast C_{11}(h)$$  \hspace{1cm} (4-2)

In Equation 4-2, $C_{12}(h)$ is the data-to-data cross-covariance between the primary and secondary data at a distance of $h$ meters, $C_{12}(0)$ is the maximum cross-covariance between the primary and secondary data, $C_{11}(h)$ is the covariance of the primary data at a distance $h$, and $M$ is the Markov-Bayes parameter. The Markov-Bayes parameter is a user-defined value in the range [0,1], and is a metric for correlation between the primary and secondary data (Zhu and Journel, 1993; Deutsch et al., 1996; Capilla et al., 1999; Boyd et al., 2018). In the proposed methodology, as will be demonstrated, this parameter and other geostatistical parameters are determined through a calibration process.
At a particular location on the grid being simulated, the COSISIM algorithm attempts to simulate the location’s categorical variable value based on nearby data. With a search neighborhood around the location defined, a maximum number of nearby categorical variable values, called conditioning data, are used to interpolate the categorical variable value at the location being simulated. The nearby conditioning data are not weighted equally; the weights for each of the conditioning data are solved using Equation 4-3.

$$\lambda = C^{-1}c$$  \hspace{1cm} (4-3)

In Equation 4-4, $C$ is the data-to-data covariance matrix (the covariances between each of the conditioning data), $c$ is the data-to-unknown covariance array (the covariances between each of the conditioning data and the location being simulated), and $\lambda$ is the array of weights of each of the conditioning data. Equation 4-3 prevents clustered data from being over-weighted, and weights data points that are closer to the location being simulated higher (isotropic variogram) or weights points along the orientation of anisotropy higher (anisotropic variogram). This allows for unbiased weighting that honors the data configuration and the extent of correlation and diminishing correlation with lag distance (Oliver and Webster, 2014).

Stochastic simulation algorithms, such as COSISIM, rely on a Monte Carlo draw from a conditional cumulative distribution function (ccdf) to assign a value at a location. The ccdf is defined by a mean (the solution to the kriging equation) and by a variance (the kriging variance). At a particular location for a particular categorical variable, the indicator kriging equation and kriging variance are defined by Equation 4-4 and Equation 4-5, respectively.

$$I_{k,SK}(u) = p_k + \sum_{\alpha=1}^{N} \sum_{m=1}^{K} \lambda_m(u_\alpha) [I_m(u_\alpha) - p_m]$$  \hspace{1cm} (4-4)
\[ \sigma_{SK}^2(u) = \text{Var}[I(u) - I_{SK}^*(u)] = C(0) - \lambda \cdot c_{12} \] (4-5)

In Equation 4-4, \( I^{*k,SK}(u) \) is the probability of a given categorical variable \( k \) occurring at a location \( u \), \( p_k \) is the marginal probability of categorical variable \( k \) occurring (equivalent to the global mean in traditional simple kriging), \( \alpha \) are the nearby conditioning data, and \( m \) are each of the categorical variables (in this study, ‘Sandstone,’ ‘Limestone,’ or ‘Shale’).

Further, \( \lambda_m(u_\alpha) \) is the kriging weight of categorical variable \( m \) for conditioning datum at location \( u_\alpha \), \( I_m(u_\alpha) \) is the value of the categorical variable \( m \) at location \( u_\alpha \), and \( p_m \) is the marginal probability of categorical variable \( m \) occurring at any location on the grid. In Equation 4-4, \( \sigma_{SK}^2(u) \) is the kriging variance at a location \( u \), \( C(0) \) is the variance of the data (also known as sill of the variogram), and \( \lambda \cdot c_{12} \) is the dot product of the vector of kriging weights, \( \lambda \), and the vector of data-to-unknown cross-covariances. (Carle and Fogg, 1995).

Equation 4-4 and Equation 4-5 are used to create the conditional ccdf with a given mean (Equation 4-4) and variance (Equation 4-5) at a particular location in space, and a Monte Carlo draw from the generated ccdf deterministically applies a categorical variable to the particular location. This process is completed for each of the remaining grid locations that have not been simulated; new realizations are created using a different path around the grid (Soares, 2001).

**4.2.3 Creation of Cross-Sections for COSISIM**

The basic geologic environment used to develop and test the proposed methodology is a hypothetical series of alternating layers of dipping sedimentary rock on a 180 by 250 unit grid. For this particular work, these layers represent the categorical variable values of ‘Sandstone’,
‘Limestone’, or ‘Shale.’ During the sensitivity analysis phase of this study, the cross-sections were created based on a straight-line approximation from three boreholes within the simulation grid.

4.2.4 Calculations of Entropy Metrics

One iteration of COSISIM with a given set of input parameters output a set of 100 equally-probable realizations of the subsurface geology. As mentioned previously, each of these realizations is not locally accurate, but the set of realizations as a whole provides a good measure of spatial variability and uncertainty (Marioni, 2003). While many applications of simulation or cosimulation algorithms use the realizations directly to estimate fluid transport or to identify spatial patterns (e.g. Lee et al., 2007; Soltani et al., 2014), this work uses the set of realizations to generate point-scale entropy, which can then be used to quantify the spatial distribution of uncertainty. Entropy is a well-defined concept in information theory that relates to the consistency of a given event occurring in a stochastic framework, and ranges from 0 (perfect consistency and high predictability) to 1 (perfect inconsistency and low predictability) (Shannon, 1948). A versatile concept, it has been previously applied in geotechnical engineering (e.g. Wang et al., 2013; Zhao and Wang, 2019), and in this case, is the measure of how many realizations simulated different lithologies at a certain point. Point-scale entropy maps were created by calculating the entropy at each location using Equation 4-6.

\[
H(X) = -\sum_{i=1}^{N} P(x_i) \log_N P(x_i)
\]  

(4-6)

Using Equation 4-6, the entropy \( H \) at a given location given \( N \) possible outcomes is taken as the probability \( P(x_i) \) of outcome \( i \) occurring at that location multiplied by the \( N \)-based log of the probability \( P(x_i) \) of outcome \( i \) occurring at that location (Shannon, 1948). In this work, the
probability of occurrence for each of the three lithologies (categorical variables) at a given location \( X \), is the number of occurrences \( (N) \) of that the lithology \( (\theta_i) \) divided by the total number of realizations \( (NR = 100) \).

\[
P(X)_{\theta_i} = \frac{1}{NR} \sum_{i=1}^{NR} N(\theta_i)
\] (4-7)

With three lithologies, the maximum value of entropy at any point would be when the probabilities of each lithology occurring are approximately equal, returning Equation 4-8.

\[
H(X) = -0.33 \log_3 0.33 - 0.33 \log_3 0.33 - 0.34 \log_3 0.34 \approx 1
\] (4-8)

Minimum entropy occurs when the probability of a single lithology occurring is 100\%, returning Equation 4-9.

\[
H(X) = -1 \log_3 1 = 0
\] (4-9)

With three lithologies, the entropy values at a given point with different combinations of probabilities of Rock 1, Rock 2, and Rock 3 existing can be found in Table 4-1.

While the value of entropy can theoretically approach 1, it is unlikely that the entropy values in this case will be significantly higher than 0.63, as near an interpreted geologic boundary of two lithologies, the COSISIM algorithm is unlikely to simulate a third lithology in a large number of realizations with the input parameters provided.
Table 4-1: Calculated entropy values with given probabilities of lithology occurring.

<table>
<thead>
<tr>
<th>Sum of Probability = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rock 1</strong></td>
</tr>
<tr>
<td>0.33</td>
</tr>
<tr>
<td>0.50</td>
</tr>
<tr>
<td>0.66</td>
</tr>
<tr>
<td>0.75</td>
</tr>
<tr>
<td>0.90</td>
</tr>
<tr>
<td>0.95</td>
</tr>
</tbody>
</table>

Figure 4-2: (a) A cross-section with borehole locations. (b) An entropy map created using this data and given input parameters.

An example entropy map using the boreholes and the cross-section in Figure 4-2a can be found in Figure 4-2b. The white vertical lines are the borehole locations, and the zones of higher entropy correspond to geologic boundary locations in the cross-section. As anticipated, areas near the presumed boundaries are associated with more uncertainty than locations the geologist believes...
are closer to the center of a lithological unit. The given set of input parameters affect the width of the zones of maximum entropy, as well as the value of the maximum entropy. For the sensitivity analysis performed, 1-D entropy samples are extracted perpendicularly across the zones of higher entropy. This allowed the 2-D entropy maps to be qualitatively compared while also providing an opportunity to quantitatively identify how altering the input parameters affected presumed boundary locations in 1-D. An example of this is shown in Figure 4-3.

Figure 4-3: (a) 2-D entropy map with 1-D Sample marked. (b) 1-D sample from the 2-D map in (a).

Figure 4-3b is a 1-D cross-sectional representation of the profile line that is marked in Figure 4-3a. The entropy value along the 1-D sample reaches a maximum of 0.62, and this value is close to the theoretical maximum anticipated at an interpreted boundary of two geologic units (Table 4-1: Calculated entropy values with given probabilities of lithology occurring.). Two additional recorded values are \( W_{0.3} \) and \( W_{0.5} \), which are the widths of the zone which correspond to entropy values of 0.3 and 0.5, which correspond to approximately 90\% \( (W_{0.3}) \) and 75\% \( (W_{0.5}) \) probabilities of the majority lithology existing. These were chosen as output metrics for the
sensitivity analysis, as together they can reasonably describe the size of the zone of maximum entropy without being sensitive to the noise outside of this zone. Both the 2-D entropy maps and the entropy data points on the 1-D sample were colored blue, green, yellow, red, and black based on the thresholds defined in Table 4-2. These values correspond to entropy intervals defined in Table 4-1: Calculated entropy values with given probabilities of lithology occurring.

Table 4-2: Color code for 1-D samples.

<table>
<thead>
<tr>
<th>Probability of Most Likely Lithology Occurring (assuming only two major possible lithologies at a given location)</th>
<th>Entropy Range</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>No majority</td>
<td>0.63 - 0.99</td>
<td>Black</td>
</tr>
<tr>
<td>&gt; 0.50</td>
<td>0.58 - 0.63</td>
<td>Red</td>
</tr>
<tr>
<td>&gt; 0.66</td>
<td>0.51 - 0.58</td>
<td>Yellow</td>
</tr>
<tr>
<td>&gt; 0.75</td>
<td>0.30 - 0.51</td>
<td>Green</td>
</tr>
<tr>
<td>&gt; 0.90</td>
<td>0.00 - 0.30</td>
<td>Blue</td>
</tr>
</tbody>
</table>

4.2.5 Modeling Decision in COSISIM

The COSISM algorithm requires a number of input parameters that can be varied. Many of these are standard inputs that are left unchanged, whereas a few are poorly constrained and can significantly affect the uncertainty predictions produced. The following section discusses several of the modeling decisions made in order to simplify the analysis and to identify variables that are varied for the sensitivity analysis in Section 4.3.
Variogram Parameters

The foundation of COSISIM and any variogram-based geostatistical algorithm, a variogram is a function that describes spatial correlation of variables throughout the simulation grid. Typically represented as variogram value versus lag distance, an empirical variogram model is fit to the experimental (measured) variogram values from input data. These variogram models are defined by a nugget effect, a range, and a sill. The nugget effect is a measure of small-scale heterogeneity and error; the range is the lag distance beyond which any pair of points is not anticipated to be correlated, while the sill is the variogram value past the range (Oliver and Webster, 2014).

As shown previously in Section 4.2, the variogram is a critical input for simulation algorithms, as the covariances defined by the variogram (Equation 4-1) define the weights of each of the nearby conditioning data (Equation 4-3) as well as the mean (Equation 4-4) and variance (Equation 4-5) of the ccdf at a particular location; the weights of the conditioning data found in Equation 4-3 are additionally used to identify the mean of the ccdf at a particular location.

Typically, experimental variograms are calculated in multiple orientations to identify potential anisotropy in the data and to determine the most appropriate parameters for the variogram model. In this study, experimental variograms were calculated by using the borehole data to fit the sill for each lithology. By utilizing the Markov Model II and full indicator kriging, one variogram model represented each lithology (Zhu and Journel, 1993; Hill, 1998). Figure 4-4a shows a simple cross-section and the location the boreholes used to create the variograms, while Figure 4-4b shows the variogram data at angles (clockwise from vertical) of 0° and 30°, which is approximately the dip of the beds inferred from the borehole data. Accordingly, the sill is reached at about four times the lag distance using an angle of 30° than at an angle of 0°.
Experimental variograms were computed at intervals of 30° at borehole locations within the simulation grid. This returned the directional variability in the variogram sill and range for each lithology. The nugget effect was set to zero for all variogram models, as no microscale heterogeneity or sampling error is assumed (Oliver and Webster, 2014). Since there was only slight variability (less than 10% deviation from the mean) in the sill values calculated for each lithology (as it is an approximation of the overall variance of the data), an isotropic approximation of the sill was used in this study.

As can be seen in the variograms in Figure 4-4b, there is noticeable uncertainty in fitting a variogram model to the borehole data due to the sparseness of data. The vertical experimental variogram appears to have a much lower range than the 30° variogram, and that is intuitive given the dipping nature of the rock units.

However, the 30° variogram does not show values until a distance of 120 grid units due to borehole spacing, adding uncertainty to any empirical variogram fit. This problem is simple to overcome, as others have previously studied variogram fitting to sparse borehole data like these with promising results (Wang and Zhao, 2017; Wang and Zhao, 2018). However, as the cross-variogram of the primary and secondary data is related to the variogram of the primary data, for this particular study, it was deemed more appropriate to treat the range of the variogram as a variable parameter to account for uncertainty in the secondary data (cross-section) than as an attribute to be calculated from the primary data.

While the uncertainty of fitting a variogram model to experimental variogram values is discussed using the cross-section in Figure 4-4, there also existed a concern with overfitting anisotropy to the end result; this is illustrated by a simpler cross-section in Figure 4-5. This cross-
section contains two lithologies with one boundary. While the boundary between the two lithologies gently dips, there are variations in the dip of the boundary.

Figure 4-4: (a) Locations on the cross-section that possess the value ‘Rock 1’ in dark gray, with boreholes used to create the variogram presented as black lines. (b) Computation of the experimental variogram in two directions for ‘Rock 1’ using the borehole data only.

As can be seen in Figure 4-5 (page 80), the use of an isotropic variogram allows for an entropy band that is less sensitive to the local dip of the boundary in the cross-section. The use of anisotropic variograms creates a narrower entropy band when the dip of the interpreted boundary in the cross-section similar to the angle of the variogram, and creates ‘flares’ in the entropy band.
when the dip of the boundary is in a different orientation than the orientation of the variogram. Near the left borehole in Figure 4-5c, there is a noticeable entropy ‘flare’ using the anisotropic variogram when the cross-section shows a horizontal boundary. This flare is misleading, as it disagrees with the cross-section and is near a source of primary data. Additionally, near the right borehole, the anisotropic variogram produces a narrow entropy band along a 45° angle. This is an issue, as this work intends to capture uncertainty in variable geologic boundaries with all orientations; effectively, an isotropic variogram should be used because the uncertainty in a cross-section tends to be approximately isotropic in many cases.

Due to the concerns presented, the range of the isotropic variogram model was varied to investigate how this value affected the results, and the variogram was set as isotropic. The impacts of the isotropic variogram parameters on the model results are presented in Section 4.3.

Decimation

In any cokriging and cosimulation algorithm, an understanding of the interaction of primary and secondary data is crucial to appropriately integrating secondary data and synthesizing the result. As previously discussed, primary data are typically directly measured, sparse information, and secondary data are typically indirectly measured and spatially extensive and are intended to show spatial continuity of the primary variable being simulated (Horta and Soares, 2010; Le Ravalec-Dupin and De Viega, 2011; Azevedo et al., 2015; Nunez et al. 2017).

As the secondary data are spatially extensive relative to the borehole data, a potential challenge is over-conditioning the secondary data to the simulation grid. This is especially true when using categorical variables (lithologies) as both the primary and secondary data, which is the case in this study.
Figure 4-5: (a) A simple cross-section with two lithologies separated by an undulating boundary. Comparison of entropy maps created using (b) an isotropic and (c) an anisotropic variogram.

The simulation grids in this case are 180 by 250 units; if the entire dataset of secondary data were used, the secondary data would be 45,000 data points compared to the 540 data points of the primary data. This is not only computationally expensive, it highly over-conditions the secondary data to the simulation grid.

Other authors have used an approach called collocated cokriging to limit secondary data to only locations where primary data exist, limiting the amount of secondary data used to avoid computationally-expensive models and over-conditioning (Xu et al., 2002; He et al., 2014). However, as the primary data in this study are vertically continuous borehole data, collocated
cokriging would not be ideal, as this approach would leave large sections of the simulation grid empty. In an approach deviating from the collocated cokriging model, a study of glacially-deposited aquifers in Denmark faced a similar challenge when incorporating secondary data, and overcame this challenge by decimation of the secondary data used (geophysical data) to ensure that these data were not over-conditioning the result (Koch et al., 2014). While the authors conceded that the degree of decimation adopted in this approach was arbitrary, this approach ensured that all of the relevant geological features were captured. Additionally, the authors validated their model by running a simulation with the entire secondary dataset, confirming that the decimation approach did not ignore relevant geologic features that would lead to inaccurate results following decimation. In this study, decimation was used to limit the amount of secondary data and allow for more freedom in COSISIM realization generation, but care was taken to ensure that the decimated data were appropriately representing the geologic structure information and the spatial correlation from the cross-section.

As can be seen in Figure 4-6, increasing the decimation has a significant effect on the resulting entropy maps. When using an un-decimated cross-section, the simulation realizations all reproduce extremely similar results to the cross-section; this results in an entropy map which shows minimal uncertainty with respect to boundary locations, contrary to the expectation that the result should be somewhat uncertain. This is due to the fact that weighting of a nearby point is much higher (non-linearly) than that of a point farther away. If large sections of the secondary data are removed, the weights for points not directly adjacent to the point being simulated become relatively larger.

Different levels of decimation were tested in this study, and the final decision was to decimate the cross-section onto a 5 x 5 unit grid; this value ensured that the secondary data did not
over-condition the results and prevent the generation of entropy maps showing non-negligible uncertainty. Additionally, this level of decimation decreased the run time to create realizations for a single parameter combination from 20 hours using the entire secondary dataset to 3.5 hours, decreasing the run time by over 80%.

Figure 4-6: Results obtained based on varying the degree of decimation – (a) Entropy map with cross-section data decimated every 5 units (b) Corresponding perpendicular 1-D sample across a zone of high entropy. (c) Entropy map with the cross-section not decimated. (d) Corresponding perpendicular 1-D sample across a zone of high entropy.

To ensure that decimating the secondary data does not affect the spatial relationship of the results, the authors created variograms of output realizations for several decimation values. Three such variograms can be found in Figure 4-7. Figure 4-7 shows that decimating the secondary data decreases the sill of the output variogram, but has little, if any, effect on the range. As the range
quantifies the size and shape of the geometries of the resulting output, an unchanging range between each value of decimation shows that the spatial correlation of lithologies in the output realizations are not affected by decimating the secondary data. The sill value, while related to the maximum variance in the output model in a particular direction, does not affect the geometrical configuration of the interpreted features in the cross-section. With this in mind, the selection of a single, consistent 5 x 5 grid for decimation eliminates any meaningful influence of the decimation on the results.

Figure 4-7: Variogram of output realization using (a) no decimation, (b) decimation value of 5 units, and (c) decimation value of 10 units.

**Marginal Probabilities**

The marginal probabilities are the respective probabilities of each lithology occurring globally across the entire simulation grid. While these values can be determined from borehole data or the cross-section, if the geologist has uncertainty in their cross-section and the modal percentages of each lithology, they can vary these values to test the influence of potentially incorrect modal percentages of each lithology. In all cases presented in this study, the three marginal probabilities were all set to 1/3.
Number of Realization

The number of realizations for a given set of input parameters was initially set to 100. This number was selected as it was considered to be a sufficient number of realizations to asymptotically approach a stable solution for maximum entropy and width of entropy zones $W_{0.3}$ and $W_{0.5}$. This was tested by stochastically selecting between 2 to 100 realizations and evaluating the maximum entropy and $W_{0.5}$ values for those realizations. For this analysis, the input parameters were a variogram range of 25 and maximum conditioning data of 12. This combination of parameters has the highest uncertainty out of all combinations as it uses a smaller number of conditioning data and has the smallest search ellipsoid.

The results from this process are found in Figure 4-8. A lower number of realizations resulted in a high variation of maximum entropy and $W_{0.5}$ values; increasing the number of realizations beyond 80 leads to very little change in maximum entropy, while increasing the number of realizations beyond 50 leads to very little change in $W_{0.5}$.

![Figure 4-8: Results of (a) maximum entropy and (b) $W_{0.5}$ obtained by stochastically selecting a given number of realizations.](image)
4.3 Input Parameter Sensitivity Analysis

With the exception of the variogram range, the parameters listed in the previous section were left unchanged during the sensitivity analysis and verification methodology. However, the variogram range, the number of conditioning data used, and the Markov-Bayes parameter appeared to have a large effect on the results.

As there is considerable uncertainty in the most appropriate values to use for this novel application, their effects were investigated using a sensitivity analysis described in this section. A base case, with a variogram range of 50, a conditioning data value of 24, and a Markov-Bayes parameter of 0.10 was selected as a control, and the subsequent analysis varied each of these variables independently to identify their effects on the 2-D entropy maps and the 1-D samples.

4.3.1 Variogram Range

As previously specified, the variogram sill was left unchanged, and the modeling decision was made to generate isotropic variograms. The range of the isotropic variogram was varied to show the effect on the 2-D entropy map and the 1-D sample. Figure 4-9 shows how doubling the range affects both of these results.

Increasing the variogram range results in a wider zone of high entropy. The reason for this phenomenon is that increasing the variogram range puts higher weighting on data points farther away from the location being simulated. Variograms with larger ranges (everything else held constant) result in higher covariance values at larger distances. Thus, the covariance values of data points farther away are higher relative to results using variograms with smaller ranges.
As the variogram range is increased, the simulation will effectively consider more data points from different lithologies. Incorporating data points from different lithologies incorporates these data into the mean of the ccdf (Equation 4-4) as well as the variance of the ccdf (Equation 4-5), leading to a greater likelihood of a Monte Carlo draw from this ccdf to identify a different lithology at a farther away location than if the nearby data points all came from the same lithology. This was found to have the strongest influences on the results, as was expected.
4.3.2 Conditioning Data

The conditioning data are the nearby data points that are included in the kriging equations for estimating the categorical variable value at an unknown location on the grid. The maximum number of conditioning data was modified similarly to range. Consider the images in Figure 4-10: while doubling the number of conditioning data did not have a significant effect on the result, it resulted in a slightly narrower zone of high entropy and a slightly smaller value of maximum entropy. This phenomenon is explained by better conditioning of nearby data; with more nearby data included in producing the ccdf, the simulated point is more likely to follow the global trend.

4.3.3 Markov-Bayes Parameter

Typical application of COSISM using primary and secondary data relies on different types of data for primary and secondary data, and the Markov-Bayes parameter defines the interpreted correlation between the spatial continuity of the data types and is defined as a positive scalar value on the range of 0 to 1 (Zhu and Journel, 1993; Capilla et al., 1999). In this study, the data types for both the primary and secondary data are the lithologies, where the primary data are borehole data from the ‘true’ geology and the secondary data are the cross-section.

While the cross-section is the expert interpretation from the geologist, these data are more subjective than the borehole data, and perfect spatial correlation between the simulated lithologies and cross-section data is not assumed. With a non-zero correlation between borehole data and cross-section data, the Markov-Bayes parameter is given as an input variable for the COSISIM algorithm.
Figure 4-10: Results obtained based on varying the number of conditioning data – (a) Entropy map with a maximum of 24 conditioning data. (b) Corresponding perpendicular 1-D sample across a zone of high entropy. (c) Entropy map with a maximum of 48 conditioning data. (d) Corresponding perpendicular 1-D sample across a zone of high entropy.

Due to subjectivity of the secondary data and the fact that these data are much more spatially numerous than the borehole data, the decision was made to not consider a Markov-Bayes parameter higher than 0.25 for the full sensitivity analysis or the verification methodology, as a value higher than this would incorrectly assume that confidence in the spatial configuration of the spatially-extensive secondary data is high relative to the borehole data. However, in order to show the potential impact of the maximum possible input value, an example with a Markov-Bayes value of 1 is presented in Figure 4-11.
Consider the images in Figure 4-11. Increasing the Markov-Bayes parameter increases the cross-covariance between the borehole and cross-section data. This increases the cokriging variance for the ccdf, which results in decreased overall entropy away from the borehole measurements in the simulation grid. This decreases the potential for deviation from the cross-section using the Monte Carlo draw from the ccdf. Effectively, higher Markov-Bayes parameters increases replication of the cross-section and correspondingly decreases the degree of uncertainty modeled.
4.3.4 Results of Sensitivity Analysis

Based on the sensitivity analysis of parameters outlined earlier in Section 3, the authors tested various combinations of variogram range values, conditioning data, and Markov-Bayes parameters. To visualize the results, the W0.3 and W0.5 values at 20 boundary locations on each entropy map for a given input parameter set were averaged to give a sense of how a combination of input parameters affected the entropy maps as a whole. The results of this analysis are shown in Figure 4-12. Figure 4-12 shows that, in general, as variogram range increases and the number of conditioning data decreases, W0.3 and W0.5 also increase. This is reasonable given that increasing the variogram range increases the weights of points farther away that may be from different lithologies, which is incorporated into the mean (Equation 4-4) and increases the variance (Equation 4-5) of the ccdf at a particular location.

Decreasing the number of conditioning data can selectively increase the weights of a fewer number of points that are from a different lithology, which increases the variance of the ccdf and focuses the mean a different nearby lithology. Additionally, increasing the Markov-Bayes parameter decreases W0.3 and W0.5; this is expected, as increasing the Markov-Bayes parameter decreases the cokriging variance (Equation 4-5) in the ccdf, resulting in lower entropy. However, this trend is not always the case, and some combinations of variogram range and conditioning data show no coherent trend in the results as a function of the Markov-Bayes parameter. This lack of a solid trend is due to the method by which the Markov-Bayes parameter is applied; this value defines the spatial correlation of the secondary data relative to the spatial correlation of the primary data. As the secondary data are much more numerous than the primary data (i.e. every 5 units in both directions as opposed to a continuous vertical borehole every 75 units horizontally), the spatial correlation of the secondary data is less important, as it is easier to interpolate with how
numerous these data are. This shows that while Markov-Bayes parameter does have an effect on the results, the impact of the variogram range and the number of conditioning data has a much larger and more consistent effect on the uncertainty predictions produced.

4.4 Application of the Proposed Method and Demonstration of the Value-Added for Quantification of Geological Uncertainty

The modeling approach proposed in Section 4.2 represents a valuable tool for the simulation of geological uncertainty in the context of an interpreted cross-section. For such an approach to be practically valuable, however, it must be verified; in other words, based on some measure of geological understanding in the cross-section (such as data density used in its creation), it must be possible to select a combination of variogram range, conditioning data, and Markov-
Bayes parameter values that will result in a series of entropy maps that faithfully reflect the actual spatial distribution of uncertainty as compared to reality.

While the previous sections show the process of developing this methodology, in this section, the authors use the developed methodology in a hypothetical complex geologic environment to demonstrate that it is possible obtain proper representations of model uncertainty. Furthermore, the authors show the parameters required to obtain such representations vary in a consistent manner as a function of geologist understanding of the geology (and initial data density as the key control on that understanding). Additionally, we demonstrate that the predictions of spatial distributions of uncertainty using our approach are more accurate than those obtained using a conventional geostatistical approach that does not incorporate geologist expert knowledge.

A complex 2-D geologic environment was generated for the verification methodology to incorporate geologist understanding of the geology into this approach. Figure 4-13a shows the ‘true’ geology generated for this case, with the location of the increasing number of boreholes iteratively given to a geologist. Figure 4-13b shows the cross-section created by the geologist using the various numbers of boreholes. The true geology shown in Figure 4-13a has the same lithology names as in the sensitivity analysis, but this cross-section has units with variable thicknesses, folding, and a fault.

As can be seen in the results of the cross-sections in Figure 4-13, the outcomes slowly begin to look more similar to the true geology, which was expected as more information being provided to the geologist. This was intended to simulate a more geologically reasonable interpretation as more information was provided.
In order to generate cross-sections that were developed using different levels of understanding of the geology, the geologist (a colleague at Colorado School of Mines) was iteratively provided with additional borehole information and asked to create cross-sections at each step, using the borehole (primary) data to create the cross-section (secondary) data with increasing amounts of information. While geophysical investigations are becoming more commonly used to develop geological cross-sections (e.g. Tacher et al., 2006; Bond et al., 2015), geologic cross-sections off of solely borehole data is a common state of practice in subsurface industries (Lark et al., 2014; Randle et al., 2019). This methodology aimed to reflect this common state of practice, which in this project, inherently uses the borehole data twice in the COSISIM simulation, both in the creation of the cross-sections and as the primary data in the algorithm itself.
Initially, the geologist was asked to generate a cross-section based on the knowledge that there are three lithologies with some level of structural complexity. The geologist was given the lithology information associated with Boreholes 1-4 as shown in Figure 4-13. In the next iteration, the geologist was given Boreholes 5 and 6 in addition to Boreholes 1-4 and identified that there is a fault present in the area of interest. On the last iteration, the geologist was given a borehole that transected the fault; Borehole 7 allowed the geologist to better constrain the location and orientation of the fault. Results from this analysis are found in Figure 4-14.

In Section 4.3, the authors established a reasonable range of input parameters that leads to \( W_{0.3} \) and \( W_{0.5} \) values that are appropriate given the thicknesses of the lithologies and the size of the simulation grid used. Within this parametric space, the authors investigated which combinations of parameters provided uncertainty predictions that were most correct at the point scale. This was assessed by comparing two quantities for each point on the simulation grid: (1) the frequency of occurrence of the most commonly simulated lithological unit at a given point (equal to \( X/100 \) realizations, with \( X \) being the number of realizations that the most likely lithology was present); (2) the frequency with which predictions at a given level of confidence as indicated by (1) correctly identified the lithology from the true geology. A perfect prediction of uncertainty would correspond to a 1:1 relationship between these quantities (e.g. a point simulated with a certain level of confidence should correspond to the true geology at a rate equivalent to that value).

The ‘true’ geology used for this case study is compared to the simulated ‘modal’ geological sections in Figure 4-14; note that these modal sections are nearly identical to the cross-sections used as inputs. Also shown in Figure 4-14 are the entropy maps corresponding to the best parameters sets for each of the cases (in terms of the representation of uncertainty), and plots showing the quality of the different uncertainty calibrations. The most appropriate parameter
Figure 4-14: Deterministic solutions from the most accurate parameter input combinations, along with a resulting entropy map and accuracy plot, using a cross-section with (a) four boreholes, (b) six boreholes, and (c) seven boreholes.
combination for each cross-section was selected based on an $R^2$ value comparing the simulation results to the theoretical ‘best case’ line of a 1:1 ratio (see bottom row in Figure 4-14). The calibrated input parameter combinations used were found to produce appropriate representations of uncertainty, with these $R^2$ values being above 0.8 in all three cases.

Comparing the simulated models in Figure 4-14, as the input cross-section improves based on an increased data density (number of boreholes), the zones of high simulated entropy decrease in size, indicating lower uncertainty in the result. As the number of boreholes increased (and the geologist’s understanding of the geology increased), the most appropriate combination of parameters was a lower variogram range and a larger number of conditioning data. This is reasonable, as cross-sections created with more boreholes and an increased understanding of the geology should generally be more locally accurate. Using a small range and large number of conditioning data better conditioned the cross-section to the simulation grid, reproducing the cross-section with less uncertainty.

Conversely, the cross-section that was created using only four boreholes was more appropriately represented by a large variogram range and a smaller number of conditioning data. Using a larger variogram range and fewer conditioning data served the dual purpose of not over-conditioning the cross-section to the simulation grid and returning a large amount of uncertainty. In this case, the cross-section made with only four boreholes is not highly locally accurate but is still reasonably globally accurate; perhaps more importantly, the uncertainty predictions are consistent with the actual local results. Such a data-sparse environment is unlikely to occur in practice in engineering geology, but was used to show how this approach can be used when geologist confidence is very low and their cross-section is relatively ‘incorrect’. Accordingly, this combination of parameters would also be appropriate if more data exist for the geologist to make
a cross-section, but the geologist has lower confidence due to the complexity of the geology or a lack of understanding of the region. The input and result for the four borehole case is considered a ‘poor’ interpretation of the subsurface geology, but reasonable global accuracy is maintained, and high uncertainty bands reflect lack of confidence in local accuracy. In cases where reasonable global accuracy cannot be guaranteed by the geologist, the proposed method is likely not appropriate for uncertainty quantification.

For the cross-sections produced using both six and seven boreholes, the most accurate parameter combination used a low variogram range and a large number of conditioning data. It is envisioned that as experience using this method is further developed, it will eventually become possible to select appropriate input parameter combinations for uncertainty modeling based on geologist understanding and/or data density a priori (e.g. for practical purposes, in the absence of knowledge of the true geology in a given area).

In addition to studying the influence of data density and resulting cross-section quality on appropriate input parameters and associated model uncertainty, the performance of the proposed method was compared to that of a more conventional geostatistical approach. The COSISM algorithm was used without secondary (cross-section) data; this approach was performed with only the borehole data from the first six boreholes given to the geologist.

The deterministic solution produced by the conventional geostatistical approach is largely globally accurate and correctly identifies all of the lithologies present in the six boreholes. Additionally, out of the 45,000 locations on the simulation grid, the conventional geostatistical approach correctly identifies 39,657 locations, whereas the solution using a geologist’s interpretation (cross-section) as secondary data only identifies 588 more correct locations, representing only a slightly more than 1% increase in global accuracy using the cross-section.
However, the solution produced using the cross-section is more locally accurate and provides a more accurate representation of the uncertainty present in the modal simulation solution.

The solution produced by the conventional geostatistical approach rigidly follows the variogram orientation input into the algorithm, and deviations from that orientation which are captured in the cross-section are not appropriately identified when only the borehole data are used. The conventional geostatistical approach tends to ‘break’ units that deviate from this orientation, whereas the 2-D solution with an isotropic variogram captures deviation as specified by the geologist in their cross-section.

The entropy map created by the conventional geostatistical approach shows much higher spatial uncertainty due to a lack of data. This is especially true in the center of the simulation grid, which is not near any data. In these locations, the algorithm must use the overall trend supplied in the anisotropic variogram along with the borehole data, whereas the algorithm using a cross-section as secondary data can use this information along with the spatial correlation supplied by the variogram and the borehole data in order to generate results. In addition to being visible qualitatively from the 2-D entropy map, it is quantitatively shown by the uncertainty calibration scatterplot, which shows that the result is systematically under-confident in areas of moderate uncertainty.

The results of this analysis show that the incorporation of secondary data in the form of a cross-section provides significant value-added relative to a more conventional approach with respect to prediction of uncertainty associated with geological models. While the conventional geostatistical approach produces a result that is reasonably globally accurate, incorporating a geologist’s interpretation aids in constraining the result to be more locally accurate.
Figure 4-15: Deterministic solutions from the most accurate parameter input combinations, along with a resulting entropy map and accuracy plot, using (a) six boreholes and a geologist’s interpretation and (b) simulation with no secondary data.
This approach reproduces a result that is geologically more accurate, as a geologist supplies the secondary information by using the boreholes to make a geologically reasonable interpretation. This shows that while the borehole data are used twice, these data are significantly sparser than the cross-section, and the cross-section is the primary driver of the geometry of the interpreted geologic features. As shown in Figure 4-14 and Figure 4-15, the geologic cross-section is a strong constraint of the subsurface geology relative to the boreholes; this is especially true when considering 6 boreholes alone or with a geologist’s interpretation (Figure 4-15).

4.5 Conclusions

Uncertainty due to geologic variability in geologists’ interpretations of subsurface geology (cross-sections) is challenging to both characterize and quantify. Additionally, the role of incorporating a geologist’s interpretation and level of understanding of the geology in spatial uncertainty quantification is poorly understood. To address these issues, a geostatistical approach was developed to combine borehole data and a cross-section for the purposes of quantifying spatial uncertainty due to aleatory variability in subsurface conditions, including the investigation of how to appropriately incorporate these two types of information into a cosimulation algorithm. A sensitivity analysis aided in understanding the physical meaning of the input parameters and uncertainty maps generated. Key findings from this study include the following:

- The relationship was identified between the appropriate variogram range and conditioning data required to appropriately represent the geologist’s understanding of the subsurface geology in the simulation. In particular, when the hard data supplied are more numerous and the geologist has a greater understanding of the
geology, the input range should be smaller and a larger number of conditioning data should be used.

- The combination of variogram range and the number of conditioning data showed mathematically that a geologist’s interpretation created with limited amounts of data is still reasonably accurate globally, whereas a geologist’s interpretation created with large amounts of data is both globally accurate and locally accurate.

- A novel approach to combine primary and secondary data of the same variable in cosimulation algorithms was developed by way of decimation, which removes large sections of secondary data while still honoring its data configuration.

- This approach shows that there is a large degree of value added by using a geologist’s cross-section along with conventional borehole data for the simulation of lithology. Incorporating a geologist’s interpretation not only guides the results to be more accurate, but also aids in showing where geologic uncertainty exists. This will allow stakeholders in engineering geology projects to make more informed decisions regarding the characterization and management of uncertainty in project and contract document development.

- The results obtained using this approach can be used by engineers, geologists, and designers to better understand uncertainty associated with a project. Additional work to constrain input parameters could be performed in collaboration with design firms and contractors using data from a variety of projects.
Acknowledgements

The authors would like to thank Colorado School of Mines M.S. student Luke Weidner for the creation of the geological cross-sections in Section 4. The authors would also like to thank Rami Abousleiman, Daniel Franz, and Heather Schovanec for proofreading and edits.
CHAPTER 5
GEOSTATISTICAL ESTIMATION OF GROUND CLASS UNCERTAINTY PRIOR TO AND DURING EXCAVATION FOR THE CALDECOTT TUNNEL FOURTH BORE PROJECT

5.1 Introduction

Tunneling projects in rock are characterized by a high degree of spatial uncertainty, which presents a unique challenge compared to engineering projects at the surface. Much of this uncertainty stems from spatial uncertainty in the subsurface geology, as the geologic and geotechnical conditions along the tunnel alignment are rarely known prior to excavation (Haas and Einstein, 2002). Geologists and engineers typically must rely on sparse borehole data taken during the site investigation phase in order to estimate anticipated tunneling conditions expected during the project. Using this information as well as a knowledge of the regional geology and case histories from nearby projects (if applicable), geologists and engineers develop geological and geotechnical cross-sections of anticipated tunneling conditions that are used for high-level planning and decision making (Fookes et al., 2000; de Vallejo and Ferrer, 2011). As these geological and geotechnical cross-sections are based off sparse data, there is inherently a high degree of uncertainty in their creation. This leads to risk of encountering unanticipated tunneling conditions, which can lead to extreme cost and timeline overruns and a poor understanding of the excavation support required, which can be hazardous to workers and equipment (Haas and Einstein, 2002; Bernardos and Kaliampakos, 2004).

Spatial uncertainty in geology is due in part to the natural, random variability within the rock (Bedi, 2013). This variability is present within rock due to the complex nature of its deposition or emplacement and subsequent tectonics. Spatial uncertainty and variability in geology and
geotechnical engineering has become an increasingly popular topic of research (e.g. Langford, 2013; Wang and Cao, 2013; Bond, 2015; Wang et al., 2015) and is relevant to a wide variety of applications including energy resource extraction (e.g. Wellmann et al., 2014) and environmental engineering (e.g. Baalousha, 2010). Within the field of underground construction and tunneling, spatial uncertainty and variability has been quantified and characterized using variogram-based geostatistics (e.g. Jeon et al., 2009; Stavropoulou et al., 2010), transition probability-based geostatistics (e.g. Felletti and Beretta, 2009), and decision aids for tunneling, also known as DAT (e.g. Einstein at al., 1999; Haas and Einstein, 2002; Min et al., 2003; Min and Einstein, 2016).

In the context of two geologic units within the completed Caldecott Fourth Bore tunnel in northern California, USA, this work used a variogram-based geostatistical algorithm (indicator kriging) with information collected from boreholes to estimate the tunneling conditions anticipated along the alignment, along with their associated uncertainty (Solow, 1986). Rock Quality Designation (RQD) was continuously sampled along the boreholes, and a notable correlation was observed between this value and the Ground Class categories used in practice to categorize anticipated tunneling conditions; however, no guidance was available in the literature how to quantify this relationship, or use this information to predict tunneling conditions prior to excavation. Using histograms of RQD separated by Ground Class, we test two methods from which to apply this relationship in geostatistical calculations. A second phase of this work used the actual Ground Classes encountered during excavation along with the original borehole data to iteratively update the estimated uncertainty ahead of the latest round of excavation throughout excavation of these two geologic units. This provides a valuable tool for updating uncertainty estimates while excavating, increasing understanding of the likelihood of encountering unanticipated tunneling conditions, and reducing uncertainty and risk underground.
Indicator Kriging

The variogram-based geostatistical algorithm indicator kriging was used in this project, as Ground Class is a categorical variable. Kriging is one of the oldest geostatistical interpolators, and is a form of linear regression that takes into account the spatial configuration and dependence of data (Olea, 1991). Originally used to estimate mineral concentrations between known boreholes (Krige, 1951), kriging has been expanded upon for application to many different types of data. Indicator kriging was designed to handle categorical as opposed to continuous variables, and is commonly used when continuous data possess one or more critical data thresholds, such as a maximum allowable value of contaminant concentration in groundwater (e.g. Liu et al., 2004) or minimum value of permeability in an aquifer (e.g. Ritzi et al., 1994).

When performing indicator kriging for any region, the first step is to generate an indicator variogram for the known data within the region. A variogram is a model of spatial correlation of the data and is calculated in multiple directions to capture any anisotropy within the data (Cressie, 1985). Variograms are defined by a range, which is the distance beyond which there is no correlation in the data in a given direction, and a sill, which is the value of the variogram past the range and is a maximum value for the variogram. If median indicator kriging is used, only one indicator variogram is required, as it is assumed the spatial correlation of each categorical variable is roughly the same; commonly, however, full indicator kriging is used, which mandates a single indicator variogram be defined for each categorical variable (Hill, 1998).

With one or more indicator variograms defined, the covariance model is generated from the variograms. This defines the weights of each of the data in the region, and takes into account both anisotropy and data clustering; the covariance model is defined in Equation 5-1.
\[ C(h) = C(0) - \gamma(h) \] (5-1)

In Equation 5-1, \( C(h) \) is the covariance at a lag distance \( h \), \( C(0) \) is the variance of the data (also known as the sill of the variogram), and \( \gamma(h) \) is the variogram defined at a lag distance \( h \) (Hill, 1998). From this point on in the algorithm, covariance values are used as opposed to variogram values. An example variogram and the corresponding covariance model are shown in Figure 5-1:

An example of a variogram and associated covariance.

Figure 5-1: An example of a variogram and associated covariance.
At all locations in a region, the indicator kriging algorithm interpolates categorical variables based on the input critical data thresholds using nearby data, or if no critical data thresholds have been defined (as was the case in this work), the probability of each categorical variable occurring in space. With a search neighborhood defined, a maximum number of allowable data within the search neighborhood, called conditioning data, are used to estimate the values at a location. These conditioning data do not have equal weighting, but rather the weights are computed using Equation 5-2.

$$\lambda = C^{-1}c$$

(5-2)

In Equation 5-2, $C$ is the data-to-data covariance matrix (the covariance values between each of the conditioning data), $c$ is the data-to-unknown covariance array (the covariance values between each of the conditioning data and the location being estimated), and $\lambda$ is the array of weights for each of the conditioning data. This equation prevents clustered data (that are assumed to be similar data) from being overly weighted, as the data-to-data covariance between clustered data will be very high, causing their weights to be lower. This also allows for data along the direction of maximum anisotropy to be weighted higher, as data along this orientation will have higher data-to-unknown covariance values (Oliver and Webster, 2014).

With the weights of the nearby conditioning data defined, the indicator kriging algorithm generates probabilities of each categorical variable being present at all locations within a region. This is done using Equation 5-3.

$$I_k^*(u) = p_k + \sum_{\alpha=1}^{N} \sum_{m=1}^{K} \lambda_m(u_{\alpha})[I_m(u_{\alpha}) - p_m]$$

(5-3)
At a location \( u \), Equation 5-3 calculates the probability of a given categorical variable \( k \) being present based on \( N \) nearby conditioning data with weights \( \lambda \). Additionally, \( p_k \) is the user-defined marginal probability of categorical variable \( k \), which is the anticipated modal outcome of each categorical variable; combined, these marginal probabilities sum to 1 (Carle and Fogg, 1995).

A relative measure of uncertainty in an indicator kriging estimate is given as the kriging variance, which is defined by Equation 5-4.

\[
\sigma^2_S(u) = \text{Var}[I(u) - I_{S_K}(u)] = C(0) - \lambda * c_{12} 
\] (5-4)

In Equation 5-4, \( C(0) \) is the variance of the data (also known as sill of the variogram), and \( \lambda * c_{12} \) is the dot product of the vector of kriging weights, \( \lambda \), and the vector of data-to-unknown cross-covariances (Carle and Fogg, 1995). This value is only reasonable for comparing different areas of the same region, as this value is homoscedastic, meaning it is independent of the data values for the conditioning data to obtain the indicator kriging estimate (Olea, 1991; Yamamoto, 2000).

5.2 Methods

The goal of this project was to apply indicator kriging in a tunneling project using borehole data collected prior to excavation to predict the distribution of Ground Classes along the alignment and associated uncertainty, then integrate data collected during excavation to identify the reduction in uncertainty ahead of the latest point of excavation. Transformation of the data from measured borehole data (RQD) to Ground Class was performed based on histograms developed during site investigation, and two methods for representing this relationship in the indicator kriging algorithm
were evaluated by comparing the resulting outputs to the actual Ground Classes encountered during excavation via face maps.

5.2.1 Case Study and Data

This study was performed in the context of the Caldecott Tunnel Fourth Bore project near San Francisco, California. A diagram of the tunnel, the boreholes drilled during the site investigation, and the two geologic units considered in this study can be found in Figure 5-2; the two units used for this study were selected based on their relatively large size, the availability of site investigation borehole data across the entire lateral extent of each of these units, and the presence of three out of the four Ground Classes that showed notable trends in the histograms of RQD by Ground Class.

The 1 km long tunnel was bored through three sedimentary formations that are divided into seven geologic units. These geologic units are subvertical to vertical due to several transverse faults located within and around the project site, including the large and active Hayward Fault located west of the West Portal (State of California Department of Transportation, 2009).

The Second Sandstone geologic unit is bounded by two faults, while the Orinda Formation is bounded on the west by another fault. This tunnel was constructed using conventional methods (i.e. without a tunnel boring machine) with open-face excavation using a road header and minimal blasting (State of California Department of Transportation, 2009). Work crews began excavating eastbound from the West Portal and westbound from the East Portal and met in the geologic unit that lies between the Second Sandstone and the Orinda Formation.
Due to the heterogeneous and complex geology, and because this project was a high-profile project in a highly populated metropolitan area, the specifications required the contractor to produce face maps regularly during excavation (0.5-1.5 m) to record the tunneling conditions encountered along the alignment. Additionally, the contractor was instructed to perform probe drilling into the face if conditions ahead of the face were highly uncertain (State of California Department of Transportation, 2009).

Prior to excavation, engineers assigned anticipated Ground Classes along the alignment based on laboratory and field tests from boreholes including Rock Quality Designation (RQD), and their strengths were evaluated in terms of Hoek-Brown 2002 Criterion (Hoek et al., 2002; Jacobs Associates, 2008). These Ground Classes are not specific to a geologic unit, but are a global representation of the expected tunneling conditions along the alignment. These assignments are inherently uncertain due to data sparsity and natural variability within the rockmass. Divided into four categories, these were intended to supply engineers and other stakeholders with data to
estimate the timeline and cost of excavation, as well as the support required. Ground Class 1 exhibits blocky to massive structure, with fair to good discontinuity surfaces. Ground Classes 2 and 3 exhibit disintegrated to very blocky structure, but Ground Class 2 exhibits poor to fair discontinuity surfaces while Ground Class 3 exhibits very poor to poor discontinuity surfaces. Ground Class 4 exhibits disintegrated structure with very poor to poor discontinuity surfaces (Jacobs Associates, 2008). This makes Ground Class 1 the “best” and Ground Class 4 the “worst” class from an excavation stability and ground support standpoint. Only Ground Classes 1-3 were present in the two geologic units considered for this study.

During the site investigation, engineers constructed histograms of RQD from the boreholes for each Ground Class based on the estimated Ground Classes along the boreholes. RQD was chosen for Ground Class prediction in this study because it was continuous in the borehole to the meter scale, the differences in Ground Classes are primarily the differences in joint frequency (i.e. disintegrated vs. blocky vs. massive), and RQD is commonly used in practice to estimate the relative size of blocks (Palmstrom, 2005). Data for these histograms were provided by Geomatrix (2008). Histograms of RQD by Ground Class for the Second Sandstone and Orinda Formation geologic units are shown in Figure 5-3.

As seen in Figure 5-3, located on the next page, there is notable relationship between RQD and Ground Class for these geologic units; higher RQD values generally correspond to a greater likelihood of Ground Class 1 being present, whereas lower RQD values generally correspond to a greater likelihood of Ground Classes 2 and 3 being present. Again, these histograms were chosen as they show the most reasonable trends of all of the geologic units present in the project.
5.2.2 Histogram Representation Methods

Two methods of representing information from the boreholes in the indicator kriging algorithm were tested and assessed using face maps collected during excavation of the tunnel that show the true Ground Classes encountered.

![Histogram of RQD to Ground Class for the Second Sandstone and Orinda Formation geologic units (after Jacobs Associates, 2008).](image)

After choosing the Second Sandstone and Orinda Formation geologic unit, the first phase of this work used RQD values from boreholes to estimate the Ground Classes and associated uncertainty prior to excavation using indicator kriging and two RQD histogram representation methods (Hard Data and Soft Data) used to assign Ground Classes within the boreholes. Though Hard Data and Soft Data are defined terms in geostatistics to indicate whether the data were
measured or a proxy for the variable of interest (e.g. Lyon et al., 2006), in this work, these terms are used in a different way to indicate whether Ground Classes are assigned directly to a borehole (Hard Data) or the probabilities of each Ground Class occurring are assigned to the borehole (Soft Data). These two histogram representation methods are described in detail in the following sections. Once Ground Class information was assigned to the borehole data, indicator kriging was used to estimate the spatial distribution of Ground Classes throughout the remainder of the two geological units using these data.

Within the indicator kriging algorithm, the marginal probability of each Ground Class occurring was defined as the proportions of anticipated Ground Classes present in the alignment for a geologic unit that was estimated by engineers in the Geotechnical Baseline Report for the project (Jacobs Associates, 2008). Full indicator kriging was used, which necessitates a variogram for each Ground Class. The bedding and structures that primarily control Ground Class are subvertical, and there is a high degree of anisotropy suggesting that the variogram range is highest in this direction. However, the borehole data given do not allow for complete characterization of this anisotropy, even in the case of the Orinda Formation, where two boreholes have slightly different orientations. For this reason, the variogram range in the subvertical direction was varied to test its effect on the results. The variogram range perpendicular to this direction (subhorizontal) was determined from the borehole data. Note that although the boreholes used in this study are not perfectly aligned with the subhorizontal axis perpendicular to the geologic structure in the area, their orientations are similar enough that the variogram parameters calculated from the boreholes represent reasonable approximations of those in the subhorizontal direction.

For a given set of input Ground Class data in the boreholes, indicator kriging generates three probability maps corresponding to the probability of each of the three Ground Class
occurring in space. The probability values for a given location in space add to 1. Both histogram representation methods were evaluated using the face mapping data, which are the actual Ground Classes encountered during excavation. In each case, a map of the modal outcome at each point was created along with a modal outcome certainty map (Figure 5-4) (Boyd et al., 2019).

The modal outcome map was created by identifying the Ground Class at each location that has the highest probability of occurring, and the modal outcome certainty map was created by identifying the probability at each location that corresponds to the identified modal outcome. These maps also show the location of face maps and the actual Ground Classes encountered during excavation (modal outcome map – Figure 5-4a) and whether the predicted Ground Classes at these locations were correct or not (certainty map – Figure 5-4b).

Prior to Excavation – “Hard Data”

We refer to the first histogram representation method as the Hard Data approach. As can be seen in Figure 5-3, for each bin of RQD within a given geologic unit, a probability between 0 and 1 exists of each for the three Ground Classes occurring at the location along the borehole. In the Hard Data approach, the Ground Class that corresponds to a given RQD value is obtained from a Monte Carlo draw, and the Ground Class from the Monte Carlo draw is assigned to that point. For example, for $N$ realizations, a location with an RQD of 24 would simulate Ground Class 2 ~0.32$N$ times and a location with an RQD of 25 would simulate Ground Class 2 ~0.51$N$ times (refer to Figure 5-3). This differs from the Soft Data approach as it treats each realization as ground truth, and not as a probability of a given Ground Class occurring in space. Thus, in order to capture variability in the histogram, it must be recompleted several times.
Figure 5-4: (a) Modal outcome and actual Ground Class encountered during excavation; (a) Example point-scale certainty map of modal outcome along with correctness from the face maps.

In order to reflect the uncertainty and variability represented by the histogram relationships between RQD and Ground Class, 100 realizations of Ground Classes for each geologic unit in this study were performed. This was deemed sufficient, as a query of any single point along the borehole over 100 realizations returned Ground Class proportions within 5% of the target probability values from the data input, and the input covariance model was retained in every
realization; in the geostatistics literature, these conditions indicate a sufficiently large number of
realizations have been used (Leuangthong et al., 2004). An example of the simulated locations of
Ground Class 1 in three borehole realizations for the Second Sandstone and the corresponding
RQD values recorded in the borehole are shown in Figure 5-5.

This approach treats each realization as hard data, which is known as ground truth in many
geostatistical applications (e.g. Zhu and Journel, 1993; Goovaerts and Journel, 1995), and relies
on creating multiple borehole realizations with different Monte Carlo draw values to represent the
variability present in the histogram. Each borehole realization is then used as an input for a separate
indicator kriging calculation.

The Hard Data approach represents the simulated Ground Class along the borehole as
ground truth; there is no uncertainty at points that are collocated with points along the borehole.
Having multiple realizations does reflect variability within the histogram, but this approach
introduces uncertainty at the inter-realization scale, and not within a single result from indicator
kriging. Additionally, the Hard Data approach identifies the Ground Class at the borehole based
on the Monte Carlo draw alone, ignoring the variogram when simulating the Ground Classes along
the borehole.

*Prior to Excavation – “Soft Data”*

Another histogram representation method was developed that directly provided individual
Ground Class probabilities as inputs to indicator kriging at the borehole locations. We refer to this
histogram sampling method as the “Soft Data” approach. For each RQD value within a borehole,
the bin for which the RQD value is within was identified, and the probability of each Ground Class
occurring at that location is equal to the corresponding probability in the histogram. For example,
in the Second Sandstone, a location with an RQD of 24 would have a 0.32 probability of being Ground Class 2 and a 0.68 probability of being Ground Class 3, whereas a location with an RQD of 25 would have 0.32, 0.61, and 0.07 probabilities of being Ground Classes 1, 2 or 3, respectively.

Figure 5-5: RQD values recorded in the borehole that transects the Second Sandstone geologic unit along with three realizations of the probability of Ground Class 1 occurring based on the RQD values.
The probabilities of each Ground Class occurring in the Second Sandstone geologic unit in the borehole and the corresponding RQD values recorded in the borehole are shown in Figure 5-6. Like the Hard Data approach, the Soft Data approach used the RQD histograms to identify the Ground Class at the borehole, and the overall distribution of data was the same using both approaches. The primary difference is that in the Soft Data approach, the variogram has a larger
role in estimating the Ground Class at the borehole: simulation of the probabilities at the borehole uses nearby probabilities as conditioning data, whereas the Hard Data approach assigned Ground Classes to the borehole as ground truth independent of nearby borehole data. This important distinction governs the spatial dependency of the data, and explains the differences in obtained results.

5.2.3 During Excavation – Evaluation of Face Map Influence on Simulation Results

Once the most appropriate base approach was identified between the Hard Data and Soft Data histogram representation methods, face maps were iteratively added as inputs for indicator kriging. A base certainty map was created using none of the face maps, and each face map was successively added along the direction of excavation (eastbound in the Second Sandstone and westbound in the Orinda Formation). At various locations along the alignment ahead of each added face map, two metrics were calculated: the change in the maximum individual probability value, referred to as the change in modal certainty (see Figure 5-4), and a metric for the overall change in the probabilities of each of the Ground Classes, the Pythagorean total probability change ($\Delta P$) (Equation 5-5):

$$\Delta P = \sqrt{(\Delta_{Gc1})^2 + (\Delta_{Gc2})^2 + (\Delta_{Gc3})^2}$$

(Equation 5-5)

In Equation 5-5, $\Delta P$ is a measure of the total change of all Ground Class probabilities, and $\Delta_{GcX}$ is the change in probability of Ground Class $X$ occurring. This provides a more robust metric for calculating the total deviation of Ground Class probabilities than the change in modal certainty. For example, if the prior probabilities of each Ground Class at a particular location were 0.25,
0.25, and 0.50, respectively, but the addition of face map input data resulted in corresponding probabilities of 0.50, 0.25, and 0.25, this would not be reflected by the change in modal certainty.

Both the change in modal certainty and $\Delta P$ were used to evaluate the impact of the face map data on the geostatistical predictions at distances 1-10 m ahead of each face map for both units. Finally, $\Delta P$ was used to identify the degree of change in the predicted results at locations that were either correctly or incorrectly identified a priori.

5.3 Results

The first part of this section outlines the comparison of the two borehole representation methods in the Second Sandstone and the Orinda Formations. Verification plots were used to identify the modal certainty versus proportion of locations with that modal certainty that were correctly identified a priori.

5.3.1 Prior to Excavation – Comparison of Histogram Representation Methods

Face maps produced periodically during excavation identified the true Ground Class encountered. This allowed for the outputs of the two histogram representation methods (Hard Data and Soft Data) to be directly compared for accuracy in modal Ground Class predictions and for the degree of correctness of the spatial uncertainty predictions. Using the certainty maps produced (e.g. Figure 5-4b), the degree of correctness was calculated by identifying the number of face map locations that fell within a given certainty range, then identifying how many of face maps within that range were actually correct. These were compared to a 1:1 line, which is the ideal verification plot (e.g. of all locations that have a certainty between 60-70%, approximately 65% should be
correct). To obtain a goodness of fit relative to the 1:1 line, a weighted R2 metric was calculated using Equation 5-6 (Boyd et al., 2019).

\[
R^2 = 1 - \sum_{i=1}^{k} \frac{\sqrt{\frac{N_i}{SSR_i}}}{\sqrt{SST}}
\]  
(5-6)

In Equation 5-6, the goodness of fit \(R^2\) for \(N\) data points over \(k\) intervals is calculated by taking the sum of squared error relative to the 1:1 line at the \(y\) value’s respective “ideal” location on the 1:1 line (\(SSR\)), and divide it by the total sum of square difference from the mean of the data points within the interval (\(SST\)) (Boyd et al., 2019). As there is a high variation in the number of face maps in each bin, bins with a large number of face maps have a high degree of heteroscedasticity in their results. For this reason, the weights are taken as the square roots of the total number of face maps in each bin, which reduces heteroscedasticity and allows for more unbiased weighing (Solon et al., 2013).

**Second Sandstone Geologic Unit**

The Second Sandstone was identified as ideal for this work; this unit possessed an RQD by Ground Class histogram that showed a high degree of variability in RQD for each Ground Class, ensuring that there would be a reasonable degree of uncertainty in the resulting outputs. Additionally, the borehole was drilled approximately 50-70 m above the tunnel alignment, allowing for some degree of uncertainty to be present due to the lower covariance values between points on the borehole and points along the alignment (Hill, 1998). Results for both the Hard Data and Soft Data approaches for histogram representation can be found in Figure 5-7 and Figure 5-8, respectively. Although the indicator kriging estimates were averaged in the Hard Data approach, this method produced more results due to the larger number of face maps considered because of
the multiple realizations performed. In Figure 5-8, each point is an average of all of the 100 realizations, whereas each point in Figure 5-7 comes from one set of borehole data inputs (probabilities).

With multiple sub-vertical range values tested, the verification plots with the highest correlation came from vertical range values 10 times larger than the horizontal range in both cases. The lowest anisotropy ratio that was tested (vertical = 4 x horizontal range) also provided reasonable results, but as the range ratio increased past 10, the Ground Class predictions became systematically over-confident. This occurs because as the range value is increased towards infinity, this effectively becomes a ‘1-D’ simulation (i.e. the result at the tunnel alignment will be a direct projection of the result simulated at the borehole).

**Orinda Formation Geologic Unit**

The other geologic unit that was examined in this work was the Orinda Formation. Like in the Second Sandstone, the histogram of RQD by Ground Class shows distinct trends, with higher RQD values corresponding to Ground Class 1 and lower RQD values corresponding to Ground Classes 2 and 3. However, a key difference is that a large section of the tunnel alignment had a single subhorizontal borehole running through its center, making this partially a 1-D problem (Figure 5-2). Figure 5-9 and Figure 5-10 show the verification plots in the Orinda Formation for the Soft Data and Hard Data approaches, respectively.

Results from the Soft Data approach show a reasonable correlation for the two lower anisotropy ratios in the Orinda Formation (again, an anisotropy ratio of 10 produces the best result). However, using the Hard Data approach, the results were systematically over-confident, with the results of the Hard Data approach appearing similar to the Hard Data approach in the
Second Sandstone at high vertical range values. Like in the Second Sandstone results, this highlights an important distinction between the two borehole sampling methods. Using the Hard Data approach, certainty is very high along the alignment due to the presence of the borehole through the tunnel zone. In practice, however, locations along the tunnel zone inherently should not have high certainty, as while borehole data are useful for estimating conditions in the subsurface, there are limits when attempting to scale up results from a narrow borehole to the excavation scale. The Soft Data approach effectively allows for uncertainty to be integrated into indicator kriging, and gives more reliable predictions of Ground Class probabilities at locations at the borehole and away from the borehole. In order to test the effects of each of the two boreholes in the Orinda Formation, indicator kriging was performed on the two boreholes separately for domains within the lateral bounds of each borehole. An inherent limitation of this approach is that the histograms used in the histogram representation methods were intended to be for the Orinda Formation as a whole, not simply a single part of it. Orinda Formation West contains the borehole that is located 20-50 m above the tunnel alignment zone, whereas Orinda Formation East contains the borehole that is coincident with the tunnel alignment zone. Results for the Orinda Formation West and East using the Soft Data approach can be found in Figure 5-11 and Figure 5-12, respectively, while results using the Hard Data approach can be found in Figure 5-13 and Figure 5-16. In the Orinda West, the simulation results were systematically overconfident. In Orinda East, the results highlight the limitations of the Hard Data approach when a borehole is coincident with the tunnel alignment; however, results obtained using the Soft Data approach with Orinda Formation East does show a slight positive correlation in the verification plots, indicating this approach may be useful even when considering boreholes coincident with the tunnel alignment zone.
Figure 5-7: Verification plots for the Hard Data approach in the Second Sandstone using indicator kriging with different sub-vertical (SV) ranges based on multiples of the sub-horizontal (SH) range estimated from borehole data.
Figure 5-8: Verification plots for the Soft Data approach in the Second Sandstone using indicator kriging with different sub-vertical (SV) ranges based on multiples of the sub-horizontal (SH) range estimated from borehole data.
Figure 5-9: Verification plots for the Hard Data approach in the Orinda Formation using indicator kriging with different sub-vertical (SV) ranges based on multiples of the sub-horizontal (SH) range estimated from borehole data.
Figure 5-10: Verification plots for the Soft Data approach in the Orinda Formation using indicator kriging with different sub-vertical (SV) ranges based on multiples of the sub-horizontal (SH) range estimated from borehole data.
Figure 5-11: Verification plots for the Hard Data approach in the Orinda Formation West using indicator kriging with different sub-vertical (SV) ranges based on multiples of the sub-horizontal (SH) range estimated from borehole data.
Figure 5-12: Verification plots for the Hard Data approach in the Orinda Formation East using indicator kriging with different sub-vertical (SV) ranges based on multiples of the sub-horizontal (SH) range estimated from borehole data.
Figure 5-13: Verification plots for the Soft Data approach in the Orinda Formation West using indicator kriging with different sub-vertical (SV) ranges based on multiples of the sub-horizontal (SH) range estimated from borehole data.
Figure 5-14: Verification plots for the Soft Data approach in the Orinda Formation East using indicator kriging with different sub-vertical (SV) ranges based on multiples of the sub-horizontal (SH) range estimated from borehole data.
5.3.2 During Excavation – Soft Data Method

Having been previously verified, the Soft Data method was used to identify the change in modal certainty and $\Delta P$ ahead of the face given new information in the form of face maps. First, the change in modal certainty was calculated; to do this, each face map was iteratively added, and a new modal certainty map was created. Then, the change in modal certainty at distances ahead of the face could be calculated by subtracting the new modal certainty value at a location from the prior modal certainty value. Figure 5-15 and Figure 5-16 show the change in modal certainty ahead of face maps at different distances within the Second Sandstone and Orinda Formation, respectively.

In both geologic units, the change in modal certainty is high in areas with lower a priori modal certainty values at distances 1 and 2 m ahead of the face. There is a notable drop off in the change of modal certainty at distances beyond 5 m. In the Orinda Formation, there is very change in modal certainty in the portion of the alignment coincident with the borehole, but a notable change in modal certainty in the region with no coincident borehole.

However, in the Second Sandstone, the values of $\Delta P$ are notably larger than the corresponding change in modal certainty values, especially near the beginning of the excavation. In these cases, it can be seen that an incorrect Ground Class was often assigned a priori, and the presence of the face map led to a change in the modal Ground Class. For example, if the a priori result predicted Ground Class 1 to be the modal outcome, but a nearby face map identified Ground Class 3 only 10 m away, the Ground Class 3 may become the predicted modal outcome. Although this would not necessarily result in any notable increase in modal certainty, the changes in each of the individual Ground Class probabilities result in a relatively large change in $\Delta P$. 

132
Figure 5-15: Change in modal certainty in the Second Sandstone geologic unit by adding face maps overlaid with the modal certainty prior to excavation; gray indicates that the specified distance ahead of face map is outside the problem domain.
Figure 5-16: Change in modal certainty in the Orinda Formation geologic unit by adding face maps overlaid with the modal certainty prior to excavation; gray indicates that the specified distance ahead of face map is outside the problem domain.
Figure 5-17: Pythagorean total probability change ($\Delta P$) in the Second Sandstone geologic unit at distances 1, 2, 5, and 10 m ahead of the excavation, along with the prior modal certainty at each location; locations where the prior modal Ground Class outcome changed when face maps were added are identified.
Figure 5-18: Pythagorean total probability change ($\Delta P$) in the Orinda Formation geologic unit at distances 1, 2, 5, and 10 m ahead of the excavation, along with the prior modal certainty at each location; locations where the prior modal Ground Class outcome changed when face maps were added are identified.
To understand whether or not the changes in Ground Class predictions caused by the addition of face maps were generally improving Ground Class predictions, the average and median $\Delta P$ values at various distances ahead of all face map locations were calculated. This same procedure was applied considering only cases where the a priori result (1) correctly identified and (2) incorrectly identified the Ground Class at a particular location. These results for the Second Sandstone and Orinda Formation are found in Figure 5-19 and Figure 5-20, respectively.

Figure 5-19: Pythagorean total probability change ($\Delta P$) in the Second Sandstone at various distances ahead of the face for individual face maps and with mean and median values shown; results are segmented based on whether the a priori results correctly or incorrectly identified the Ground Class at each location of interest.
Figure 5-20: Pythagorean total probability change ($\Delta P$) in the Orinda Formation at various distances ahead of the face for individual face maps and with mean and median values shown; results are segmented based on whether the a priori results correctly or incorrectly identified the Ground Class at each location of interest.

In the Second Sandstone, there is a notably higher average $\Delta P$ when the a priori result incorrectly identified the Ground Class at a particular location than when the a prior result correctly identified the Ground Class, even at a distance of 10 m ahead of the face. This shows that the addition of face maps provides a valuable tool for adjusting predictions ahead of the face most
when the prior modal outcome was originally wrong, although in the Second Sandstone, the addition of face maps does adjust predictions even when the prior modal outcome was correct.

The Orinda Formation showed minimal influence of the face maps on the updated results using either metric. This is again due to the fact that a large section of the alignment within the Orinda Formation is collocated with a borehole. However, looking at the individual data points in Figure 5-20, there are still some cases where the face maps did impact the Ground Class predictions within a few meters of the face (only when the original Ground Class predictions at those locations were incorrect).

5.4 Conclusions

Uncertainty in tunneling projects both prior to and during excavation is challenging to quantify and predict with sparse borehole data and limited understanding of how information at the borehole scale relates to the tunneling conditions encountered during excavation. This work compared two methods of applying sparse, small-scale borehole information to estimate Ground Classes and evaluated their accuracy in producing realistic predictions of Ground Class uncertainty. The first method (termed the “Hard Data” method) performed a Monte Carlo draw to convert RQD values to Ground Classes, and assigned the results to the boreholes for a given realization of indicator kriging. The second method (termed the “Soft Data” method) assigned Ground Class probabilities along the boreholes as the input for indicator kriging. This novel comparison identified critical differences in these two methods even when the same input data are used, as the indicator kriging algorithm treats these input data differently. While both of these approaches used the same histogram, the assignment of probabilities (“Soft Data” method) allowed for better utilization of the variogram to honor the spatial relationships inherent in the data.
Additionally, face maps collected during excavation were iteratively incorporated into the calculations to identify their influence on Ground Class prediction. This analysis confirmed that in cases where the pre-excavation data are not coincident with the tunnel alignment, consideration of additional data collected during excavation can significantly influence the results obtained from updated calculations. An examination of the overall influence of added face maps on simulation results ($\Delta P$) showed that in this particular geological environment, face map information can provide information relevant to predictions up to 10 m ahead of the face; this is especially true when information gathered during excavation contradicted initial Ground Class predictions and/or where initial predictions had a high degree of uncertainty.

This work provides project stakeholders with valuable information for the basis of risk assessment and evaluation not both prior to and during excavation; additionally, this work has developed a tool that can aid geologists and engineers to understand where information should be gathered during excavation and when initial predictions are more likely to be incorrect. This work provides engineers and geologists with information about the value of incorporating data gathered during excavation into update estimates of ground conditions and uncertainty. Additionally, the methodology demonstrated in this study could be applied to other projects and geological settings.

Acknowledgements

The authors would like to acknowledge Gall Zeidler Consultants for access to face maps for the Caldecott Tunnel Fourth Bore project.
CHAPTER 6

CONCLUSIONS

The research in this thesis demonstrates how geostatistics can aid geologists and engineers in understanding how different aspects of engineering uncertainty are affected by natural, random, aleatory variability. Spanning from the grain (millimeter) to the project (kilometer) scale, the approaches outlined in this thesis show that geostatistical methods can be successfully adapted using unique approaches at varying spatial scales with different types and configurations of data, leading to better characterization of geologic characteristics for use in subsurface rock engineering projects. The following paragraphs summarize some of the specific findings of this research.

Covariance maps of UCS specimen images were able to successfully capture variability in grain size, shape, and mineralogy based on grayscale values at the millimeter scale. As opposed to traditional covariance maps, which are generated using sparse data and as a basis for other variogram-based geostatistical approaches, the grayscale pixel values in the images were nearly continuous on the grid, creating data-dense 2-D covariance maps and 1-D variograms that were highly certain; this captured minor variability within the rock core specimens with extremely high precision.

Nontraditional use of covariance maps and variograms identified a positive relationship between geologic and geomechanical variability using two output metrics from the parameterization of 1-D variograms from the 2-D covariance maps (geologic variability) within a specimen and the coefficient of variation of the UCS values for each specimen within a geologic unit (geomechanical variability). Within the 18 rock units, a geostatistical parameter was defined for individual UCS specimens that captured not only heterogeneity within the specimen, but fabric orientation, anisotropy, and discrete features that could influence strength. Additionally, a second
parameter was defined that captured heterogeneity between samples, and was found to be more appropriate from UCS variability prediction when discrete features were not present in the specimens. Ultimately, a framework for determination of the number of UCS specimens required to evaluate the mean UCS value for a geotechnical unit was developed based on the parameters and correlations that were identified. This approach is the first to allow for such a determination to be made without the prior availability of a large (and costly) testing database for the purposes of statistical analysis.

At the field-scale, the variogram-based algorithm Sequential Indicator Cosimulation (COSISIM) and the information theory concept of entropy were used to successfully combine a geologist’s interpretation with borehole data to estimate spatial uncertainty in geologists’ cross-sections. This is the first study to develop quantitative estimates of spatial uncertainty in cross-sections created by geologists through integration of interpretive information and hard (borehole) data. As both the primary and secondary data were of the same variable, decimation of the secondary (cross-section) data was required so that the extensive cross-section (interpreted) data did not over-condition the result and ignore the sparse borehole (true) data. While this ultimately removed 96% of the information from these cross-sections, it was found that the variogram of the decimated cross-section reflected the spatial dependency of features in the full cross-sections, and provided the ability to quantify uncertainty.

Additionally, significant reduction of the Markov-Bayes Parameter in the COSISIM algorithm, which limited the cross-variogram between the two data types, signified that there is a weak correlation between the spatial dependency of the borehole and cross-section data. This was found to increase the spatial influence of the borehole data even though they were more sparse than the cross-section data. Ultimately, it was found that as geologist interpretation confidence
increases, the variogram range should be decreased and the number of conditioning data should be
decreased to achieve well-calibrated predictions of spatial uncertainty. The approaches used to
achieve this objective demonstrated that a nontraditional use of cosimulation could successfully
include the primary and secondary variable of the same data type.

Finally, an optimal method for the simulation of engineering Ground Class in the context
of an actual case study was identified, and this method was used to evaluate the impact of
incorporating face mapping data into geostatistical simulation on uncertainty predictions. Using
histograms of RQD (measured in boreholes) and Ground Class (inferred prior to excavation), two
histogram representation methods were compared using face maps for validation, as they identify
the true Ground Class encountered during excavation. The histogram representation method that
was found to provide the most accurate results was to assign probabilities of each Ground Class
being present along the boreholes as inputs for geostatistical simulation based on the measured
RQD values.

Using the aforementioned approach, it was found that the incorporation of face map data
into updated simulations reduced uncertainty up to 10 m ahead of the face throughout the
excavation in the Second Sandstone geologic unit. Results showed that if a face map disagreed
with the a priori Ground Class prediction in a given area, it would have a notable influence on
simulation results up to 10 m ahead of the face, in some cases changing the modal Ground Class
prediction.

In the Orinda Formation, the Soft Data approach prior to excavation was found to have a
reasonably good (weakly correlated) result prior to excavation. While this approach was limited
by the collocated borehole along a large section of this geologic unit, a large change in modal
outcome certainty was observed in the section that was not collocated with the borehole.
In both geologic units, these results indicate that performing face mapping in a tunnel can help prepare engineers for changing conditions and inform decisions with respect to the need for further analysis or probe drilling.

This thesis has shown that nontraditional use of geostatistics can be successfully used to better understand random, aleatory geologic variability at multiple scales and with different types of data. In all of these analysis performed, it was critical to understand the most appropriate tool based on the input data provided and the intended results, as each of these algorithms have advantages and drawbacks for different applications. If used appropriately and with a solid knowledge of all of these considerations, this work has shown that the use of geostatistics provides geologists and engineers with a powerful tool to better capture, characterize, and quantify spatial geologic variability and uncertainty in the field of tunneling engineering, reducing risk and ensuring these projects go more smoothly and with better outcomes.

6.1 Future Work

There are multiple directions for future work that resulted from this thesis. In the 2-D covariance map analysis, only UCS specimens were considered. This can easily be expanded to include triaxial specimens and BTS (Brazilian Tensile Strength) specimens. Additionally, the use of the COSISIM-based approach for integrating a geologist’s interpretation could be applied to a real case study, where the result could be verified based on the rock types encountered in the alignment during excavation; while this was originally intended as a primary methodology for the Caldecott Fourth Tunnel Fourth Bore analysis, the data provided were found to not be applicable for this approach. Further, the COSISIM-based approach could be expanded to include multiple geologists and multiple cross-sections in the verification methodology, allowing to reliably and
quantitatively translate geologist certainty into input parameters in the COSISIM algorithm. With respect to the Caldecott Fourth Bore analysis, the alternative (Hard Data) histogram representation approach could be used to analyze the impact of face maps on updated simulations to see if the results are similar to those already obtained. Additionally, the results obtained from this analysis could be compared to a traditional Markov Chain approach, which uses the transiogram as opposed to the variogram.
REFERENCES


Klein, C. and N. J. Beukes. 1989. Geochemistry and sedimentology of a facies transition from limestone to iron-formation deposition in the early Proterozoic Transvaal Supergroup, South Africa. *Economic Geology*, 84.


Kyriakidis, P. 2005a. raster2structgrid.m. MATLAB code.

Kyriakidis, P. 2005b. geoeas2matlab.m. MATLAB code.

Kyriakidis, P. 2005c. matlab2geoeas.m. MATLAB code.


State of California Department of Transportation. 2009. Notice to bidders and special provisions for construction on state highway in Alameda and Contra Costa Counties from east Temescal Separation to 0.8 km east of Gateway Blvd.


Webster, R., and M. A. Oliver. 2007. *Geostatistics for Environmental Scientists*. 2nd ed. Chichester, West Sussex, England: John Wiley and Sons Ltd.


