

WITHIN-SITE VARIABILITY AND RELIABILITY OF
FOUNDATION DESIGNS BASED ON LOAD TESTS

A Dissertation presented to the Faculty of the Graduate School
University of Missouri-Columbia

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy
Civil & Environmental Engineering

by

ANDREW Z BOECKMANN

Dr. J. Erik Loehr, Dissertation Supervisor

DECEMBER 2019

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

WITHIN-SITE VARIABILITY AND RELIABILITY OF
FOUNDATION DESIGNS BASED ON LOAD TESTS

presented by Andrew Boeckmann,

a candidate for the degree of Doctor of Philosophy in Civil and Environmental Engineering,

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor J. Erik Loehr, Dissertation Supervisor
Department of Civil and Environmental Engineering

Professor John Bowders, Committee Member
Department of Civil and Environmental Engineering

Professor Brent Rosenblad, Committee Member
Department of Civil and Environmental Engineering

Professor Paul Speckman, Committee Member
Department of Statistics

Acknowledgments

In my master's thesis, I acknowledged Erik Loehr for being "more than a good adviser...for motivating me to be a geotechnical engineer and teaching me to think critically. I am sincerely thankful for the unique honor of working with Dr. Loehr." Some 13 years later, I remain ever-thankful for the opportunity to learn from and work with Dr. Loehr. He is everything one could want in a good mentor – enthusiastic, always giving credit, and genuinely invested in my growth – and in a good boss – supportive, thoughtful, and understanding of life outside work. He has gotten me out of several ruts: mostly figurative, one literal.

I am also grateful to have worked with John Bowders, Brent Rosenblad, and Paul Speckman over the last eight years. Each of them has helped considerably. Dr. Bowders' enthusiasm and perseverance are truly inspirational. Dr. Rosenblad's commitment to excellence is similarly inspiring and reassuring to this Mizzou alumnus. All of the reliability work I have completed in the last eight years would be lost without Paul Speckman, whose ability to understand any engineering problem we threw at him is remarkable.

A debt of thanks is also owed to the many excellent students (mostly former students) I have had the pleasure to work with: Elgin Burton, Dan Ding, Tayler Day, Kyle Friedman, Sarah Grant, Wyatt Jenkins, Eric Koenig, Tasneem Khan, Eric Lindsey, Ronnie Roustio, Sam Runge, Aaron Schoen, Zakaria El Tayash, Minh Uong, Isaiah Vaught, and Martin Wallace. It is humbling to have worked with so many students who are already leaders in the profession.

Funding for this work was provided by the National Cooperative Highway Research Project and the California Department of Transportation. I sincerely appreciate the support. I also appreciate the corresponding opportunity to collaborate with Tom Shantz, whose perspective and experience has been invaluable. A huge debt of thanks is also owed to Walter Vanderpool and Shawna Munn, whose generous provision of micropile load test data facilitated much of this work.

Finally, thank you to the family and friends who have supported me throughout this work. I have nothing but love and appreciation for Todd, Charley, Thom and Pete, Chris, Jessica, and Laura, and mom and dad.

Table of Contents

Acknowledgments	ii
Table of Contents	iii
Tables	vi
Figures	vii
Abstract	xii
1. Introduction	1
1.1 Motivation & Background	1
1.2 Objective	7
1.3 Hypothesis	7
1.4 Research Methodology	8
1.5 Organization of Dissertation	10
2. Literature Review	12
2.1 Foundation Reliability	12
2.1 Bayes' Theorem and Bayesian Inference	15
2.2 Bayesian Updating Framework for Foundation Design with Load Test Information	18
2.3 Bayesian Updating of Foundation Resistance for Load Tests Achieving Failure	22
2.3.1 Kay (1976)	23
2.3.2 Zhang and Tang (2002)	24
2.3.3 Baecher and Rackwitz (1982)	28
2.3.4 Summary of Approaches for Bayesian Updating with Load Tests that Achieve Failure	32
2.4 Bayesian Updating of Foundation Resistance for Proof Tests	34
2.4.1 Zhang (2004)	35
2.4.2 Najjar and Gilbert (2009a, 2009b)	40
2.4.3 Najjar et al. (2017)	44
2.4.4 Summary of Approaches for Bayesian Updating with Proof Tests	46
2.5 Published Information Regarding the Value of Within-site Variability	48
2.6 Summary	52
3. Characterization of Within-site Variability for Micropiles	54
3.1 Description of Datasets	55
3.2 Interpretation of Within-site Variability	55
3.2.1 Sites with Multiple Tests to Failure	56

3.2.2	Site with Many Proof Tests, Some to Failure.....	59
3.2.3	Site with Many Tests with Load-displacement Curves	60
3.3	Distribution of Within-site Variability Among Sites	64
3.4	Comparison of Findings with Published Values of Within-site Variability	66
3.5	Summary	68
4.	Numerical Tool for Bayesian Updating and Reliability Evaluation	70
4.1	Overview of the Numerical Tool	71
4.2	Numerical Method for Bayesian Updating of Foundation Resistance based on Load Test Information.....	73
4.2.1	General Framework and Notation.....	73
4.2.2	Bayesian Framework	76
4.2.3	Computational Methods	81
4.3	Numerical Method for Evaluating Foundation Reliability from Results of Bayesian Updating	85
4.4	Summary	89
5.	Effect of Deterministic vs. Probabilistic Within-site Variability for Cases with “Perfect” Sampling of Load Test Results	91
5.1	Evaluation Methodology	92
5.1.1	Parameters Defining Prior Distributions of μ_{within} and σ_{within}	93
5.1.2	Parameters Defining “True” Site Conditions and Perfect Sampling	98
5.1.3	Parameters Used for Evaluation of Results	99
5.2	Reliability at the Limits: Zero Load Tests and Many Load Tests	101
5.2.1	Reliability without Load Tests	101
5.2.2	Reliability with a Large Number Load Tests	109
5.3	Trends in Posterior Distributions	115
5.3.1	Trends versus Number of Tests for Case 0.5-B	116
5.3.2	Trends in μ_{μ} for all Cases	120
5.3.3	Trends in σ_{μ} for all Cases	122
5.3.4	Trends in μ_{σ} for all Cases	125
5.3.5	Trends in σ_{σ} for all Cases	127
5.3.6	Summary of Posterior Distribution Trends and Support for Hypothesis	129
5.4	Summary	131
6.	Effect of Deterministic vs. Probabilistic Within-site Variability for Cases with Random Sampling of Load Test Results	134
6.1	Evaluation Methodology	135

6.1.1	Parameters Defining Prior Distributions.....	137
6.1.2	Generation of Sites by Random Sampling.....	138
6.1.3	Randomly Sampled Test Results.....	139
6.1.4	Parameters Used for Evaluation of Results.....	139
6.2	Results for Case 0.5-B	141
6.2.1	Convergence of Posterior Distribution Parameters	141
6.2.2	Ratio of Estimated to true Value of $R_{1/1500}$	145
6.3	Errors and Convergence Trends in $R_{1/1500}$	149
6.3.1	Sample Mean and Sample Deviation of the Ratio of Estimated to true $R_{1/1500}$	150
6.3.2	Satisfaction of Target Reliability.....	154
6.3.3	Correlation of Errors with True Site Parameters	160
6.4	Increase in $R_{1/1500}$ with Load Tests	166
6.5	Summary	172
7.	Summary, Conclusions, and Recommendations	178
7.1	Summary of Research and Significant Findings	178
7.2	Conclusions	188
7.3	Recommendations for Practice	191
7.4	Recommendations for Future Research	192
	References.....	194
	Vita	200

Tables

Table 1: Comparison of methods for evaluating reliability of designs based on load tests to failure.	34
Table 2: Summary of results from Najjar and Gilbert (2009b) for an example test site under various assumptions regarding the lower bound resistance.	44
Table 3: Comparison of methods for evaluating reliability of designs based on proof load tests. .	48
Table 4: Within-site variability data reported by Evangelista et al. (1977).	49
Table 5: Within-site variability data reported by Kay (1976).	51
Table 6: Summary of cases with within-site variability calculated from multiple load tests to failure at the same site.	58
Table 7: Summary of within-site variability information for micropiles at five sites.	66
Table 8: Summary of input parameters used to generate perfectly sampled load test results.	93
Table 9: Summary of 12 sets of prior distribution parameters used for perfect sampling analysis.	98
Table 10: Summary of parameters used to evaluate results.	101
Table 11: Summary of posterior distribution trends.	131
Table 12: Summary of input parameters used to generate randomly sampled load test results. .	136
Table 13: Summary of 5 sets of prior distribution parameters used for random sampling analysis.	138

Figures

Figure 1: Failure of a pier supporting an elevated portion of the Lee Roy Selmon Crosstown Expressway in Tampa, Florida (from Dapp et al., 2013).	2
Figure 2: Conceptual representation of among and within-site variability for a presumptive model of foundation resistance.	5
Figure 3: Overview of reliability: (a) distribution of load effect, Q , and resistance, R , and (b) distribution of the safety margin, g	13
Figure 4: Probability of failure versus reliability index for various safety margin distributions. From Baecher and Christian (2003).	14
Figure 5: Procedure for application of Bayesian updating to foundation design based on load test information. Distributions shown in black are for a presumptive design without load tests. Data and distributions in red are for updating based on three load tests.	20
Figure 6: Required factor of safety recommended by Kay (1976) based on Bayesian updating of mean pile resistance from load test information.	24
Figure 7: Example distributions of RMP prior and posterior to conducting one load test from Zhang and Tang (2002).	27
Figure 8: Required factor of safety to achieve target reliability index (beta) as a function of observed RMP and number of tests from Zhang and Tang (2002).	28
Figure 9: Prior distributions of site precision, h , from Baecher and Rackwitz (1982). Precision is the reciprocal of within-site variance.	29
Figure 10: Required factor of safety after updating design from results of one load test: (a) assuming broad distribution of site variability and (b) assuming narrow distribution of site variability. Different lines correspond to different values of reliability index. From Baecher and Rackwitz (1982).	31
Figure 11: Required factor of safety after updating design from results of load tests assuming narrow distribution of site variability: (a) two tests and (b) three tests. The target reliability index (β) is 2 for all plots. Different lines correspond to different values of variability among load test results. From Baecher and Rackwitz (1982).	32
Figure 12: Predictive distributions of foundation resistance: prior (solid line) and after conducting various numbers of successful proof tests (dashed lines) with $x_t = 1$. From Zhang (2004).	38
Figure 13: Predictive distributions of foundation resistance: prior (solid line) and after conducting three proof tests (dashed lines) with $x_t = 1$ and varying numbers of successes. From Zhang (2004).	39
Figure 14: Required factor of safety to achieve a reliability index of 2.5 versus number of successful tests for various target proof load levels. All lines assume the result of all tests in the program is positive. From Zhang (2004).	40
Figure 15: Mixed lognormal distribution used by Najjar and Gilbert (2009a) to characterize the distribution of foundation resistance in the presence of a lower bound.	41

Figure 16: Required design factor of safety as a function of proof load magnitude relative to median resistance for different reliability indices. From Najjar and Gilbert (2009b).	42
Figure 17: Probabilistic distributions used in Najjar et al. (2017) approach: (a) conditional distribution of individual foundation resistance for a given value of mean resistance and (b) mean foundation resistance and lower-bound resistance.	45
Figure 18: Comparison of reliability impacts from approaches by (a) Najjar et al. (2017) and (b) Zhang (2004). Both approaches are based on similar inputs.	46
Figure 19: Histogram and fitted normal distribution of the within-site variability data presented by Baecher and Rackwitz (1982).	52
Figure 20: Histogram of resistance data for proof test program with fitted normal distribution.	60
Figure 21: Load-displacement curves for the micropiles documented by Vanderpool et al. (2002). ..	62
Figure 22: Histogram and fitted distribution of resistance interpreted at different values of displacement at top of micropile: (a) 0.5 in., (b) 1 in., (c) 1.5 in., (d) and 2 in.	62
Figure 23: Within-site variability versus top of pile displacement for project documented by Vanderpool et al. (2002).	64
Figure 24: Probability density functions of COV_R , with individual sites in blue and distributions for the collection of sites in red.	66
Figure 25: Histograms and fitted normal distributions for data from Baecher and Rackwitz (1982) and for micropile data evaluated in Chapter 3.	68
Figure 26: Overview of the numerical tool for Bayesian analysis and reliability analysis of designs based on load test information.	72
Figure 27: Depiction of numerical calculation method for Bayesian updating of foundation resistance in light of load test information.	83
Figure 28: Random sampling procedure used to define posterior predictive distribution.	89
Figure 29: Overview of methodology used to evaluate hypothesis via the numerical tool.	93
Figure 30: Among-site variability versus within-site variability to achieve total variability in resistance of $COV = 0.1$ for different values of uncertainty in within-site variability.	96
Figure 31: Among-site variability versus within-site variability to achieve total variability in resistance of $COV = 0.3$ for different values of uncertainty in within-site variability.	97
Figure 32: Among-site variability versus within-site variability to achieve total variability in resistance of $COV = 0.5$ for different values of uncertainty in within-site variability.	97
Figure 33: Difference between deterministic and probabilistic values of $R_{1/1500}$ for cases without load tests versus standard deviation in within-site variability.	105
Figure 34: Difference between deterministic and probabilistic values of $R_{1/1500}$ for cases without load tests versus COV of within-site variability.	105
Figure 35: Difference between deterministic and probabilistic normalized values of $R_{1/1500}$ for cases without load tests versus standard deviation in within-site variability.	106

Figure 36: Difference between deterministic and probabilistic normalized values of $R_{1/1500}$ for cases without load tests versus COV of within-site variability.	106
Figure 37: Computed total <i>COV</i> of predictive distributions versus the ratio of within-site to among-site variability.....	108
Figure 38: Ratio of posterior to true $R_{1/1500}$ versus true mean for deterministic within-site variability. Case numbers refer to Table 9.	112
Figure 39: Ratio of posterior to true $R_{1/1500}$ versus true standard deviation for deterministic within-site variability.....	112
Figure 40: Ratio of posterior to true $R_{1/1500}$ versus true mean for probabilistic within-site variability.	113
Figure 41: Ratio of posterior to true $R_{1/1500}$ versus true standard deviation for probabilistic within-site variability.....	114
Figure 42: Ratio of posterior to true $R_{1/1500}$ versus within-site variability.....	115
Figure 43: Trend in μ_μ versus number of tests for perfect sampling of Case 0.5-B with true site mean at ninetieth percentile of μ_{within} and true within-site variability at tenth percentile of σ_{within} . . .	118
Figure 44: Trend in σ_μ versus number of tests for perfect sampling of Case 0.5-B with true site mean at ninetieth percentile of μ_{within} and true within-site variability at tenth percentile of σ_{within} . . .	118
Figure 45: Trend in μ_σ versus number of tests for perfect sampling of Case 0.5-B with true site mean at ninetieth percentile of μ_{within} and true within-site variability at tenth percentile of σ_{within} . . .	119
Figure 46: Trend in σ_σ versus number of tests for perfect sampling of Case 0.5-B with true site mean at ninetieth percentile of μ_{within} and true within-site variability at tenth percentile of σ_{within} . . .	120
Figure 47: Error reduction in μ_μ for large number of tests versus ratio of within-site to among-site variability for prior distributions of μ_{within} and σ_{within}	122
Figure 48: Number of tests to reduce error in the posterior value of μ_μ by 75 percent versus ratio of within-site to among-site variability for prior distributions of μ_{within} and σ_{within}	122
Figure 49: Error reduction in σ_μ for large number of tests versus ratio of within-site to among-site variability for prior distributions of μ_{within} and σ_{within}	124
Figure 50: Number of tests to reduce the posterior value of σ_μ by 75 percent of its prior value versus ratio of within-site to among-site variability for prior distributions of μ_{within} and σ_{within} . . .	125
Figure 51: Error reduction in μ_σ for large number of tests versus ratio of within-site to among-site variability for prior distributions of μ_{within} and σ_{within} . Results are for probabilistic within-site variability only.....	126
Figure 52: Number of tests to reduce error in the posterior value of μ_σ by 75 percent versus uncertainty in within-site variability for prior distribution of σ_{within} . Results are for probabilistic within-site variability only.....	127

Figure 53: Error reduction in σ_σ for large number of tests versus ratio of within-site to among-site variability for prior distributions of μ_{within} and σ_{within} . Results are for probabilistic within-site variability only.....	129
Figure 54: Number of tests to reduce the posterior value of σ_σ by 75 percent of its prior value versus uncertainty in within-site variability for prior distribution of σ_{within} . Results are for probabilistic within-site variability only.	129
Figure 55: Overview of methodology used to evaluate trials with randomly sampled load tests via the numerical tool.....	136
Figure 56: Among-site variability versus within-site variability, showing five cases analyzed for random sampling analysis.	138
Figure 57: Trend in μ_μ versus number of tests for perfect and random sampling of Case 0.5-B with true site mean at ninetieth percentile of μ_{within} and true within-site variability at tenth percentile of σ_{within}	142
Figure 58: Trend in σ_μ versus number of tests for perfect and random sampling of Case 0.5-B with true site mean at ninetieth percentile of μ_{within} and true within-site variability at tenth percentile of σ_{within}	143
Figure 59: Trend in μ_σ versus number of tests for perfect and random sampling of Case 0.5-B with true site mean at ninetieth percentile of μ_{within} and true within-site variability at tenth percentile of σ_{within}	144
Figure 60: Trend in σ_σ versus number of tests for perfect and random sampling of Case 0.5-B with true site mean at ninetieth percentile of μ_{within} and true within-site variability at tenth percentile of σ_{within}	144
Figure 61: Histograms with fitted lognormal distributions for ratio of estimated to true $R_{1/1500}$, with one dataset for deterministic within-site variability and another for probabilistic within-site variability. Results for prior are shown in (a). Posteriors are defined after (b) one load test, (c) two load tests, (d) three load tests, (e) five load tests, (f) ten load tests, (g) fifty load tests, and (h) 200 load tests.....	148
Figure 62: Sample mean of values of ratio of estimated to true $R_{1/1500}$ versus number of load tests for Case 0.5-B with random sampling.	149
Figure 63: Sample standard deviation of values of ratio of estimated to true $R_{1/1500}$ versus number of load tests for Case 0.5-B with random sampling.....	149
Figure 64: Sample mean of values of ratio of estimated to true $R_{1/1500}$ versus number of load tests for all cases with random sampling.	152
Figure 65: Sample standard deviation of values of ratio of estimated to true $R_{1/1500}$ versus number of load tests for all cases with random sampling.....	153
Figure 66: Depiction of overreliable, satisfactory, and underreliable outcomes for (a) relatively uniform and (b) relatively variable distributions of the ratio of estimated to true $R_{1/1500}$	155
Figure 67: Percent of underreliable, overreliable, and satisfactory outcomes versus number of load tests for Case 0.5-A.....	158

Figure 68: Percent of underreliable, overreliable, and satisfactory outcomes versus number of load tests for Case 0.5-B.....	159
Figure 69: Percent of underreliable, overreliable, and satisfactory outcomes versus number of load tests for Case 0.5-C.	159
Figure 70: Percent of underreliable, overreliable, and satisfactory outcomes versus number of load tests for Case 0.3-B.....	159
Figure 71: Percent of underreliable, overreliable, and satisfactory outcomes versus number of load tests for Case 0.1-B.....	160
Figure 72: Ratio of estimated to true $R_{1/1500}$ versus within-site mean for Case 0.5-A.....	162
Figure 73: Ratio of estimated to true $R_{1/1500}$ versus within-site variability for Case 0.5-A.	162
Figure 74: (a) Slope and (b) intercept parameters versus number of load tests for all cases with $COV_{total} = 0.5$. Parameters are from linear regression of ratio of estimated to true $R_{1/1500}$ versus true site mean (μ_{within}).	164
Figure 75: (a) Slope and (b) intercept parameters versus number of load tests for all cases with $COV_{total} = 0.5$. Parameters are from linear regression of ratio of estimated to true $R_{1/1500}$ versus true within-site variability (σ_{within}).	166
Figure 76: Increase in $R_{1/1500}$ versus number of tests for different values of the ratio of within-site to among-site variability. Nominal predictive COV is 0.5 for all cases.	169
Figure 77: Increase in $R_{1/1500}$ versus ratio of within-site to among-site variability for different numbers of tests. Nominal predictive COV is 0.5 for all cases.....	170
Figure 78: Increase in $R_{1/1500}$ versus number of tests for different values nominal prior predictive COV . Ratio of within-site to among-site variability is 1.0 for all cases.	171
Figure 79: Increase in $R_{1/1500}$ versus ratio of nominal prior predictive COV for different numbers of tests. Ratio of within-site to among-site variability is 1.0 for all cases.	172

Abstract

Probabilistic evaluations of the reliability of foundation designs based on site-specific load test information are somewhat limited in number. Published evaluations have generally relied on Bayesian techniques. A primary input for Bayesian analysis is *within-site variability*, which describes the variability of foundation resistance across a site. Within-site variability is attributed to geologic variation across a site and to differences in construction outcomes among foundation elements. Published evaluations have generally used a deterministic value of within-site variability wherein within-site variability is treated as a known parameter and is not subject to updating based on load test results. In contrast, probabilistic within-site variability treats within-site variability as an uncertain parameter with its own probability distribution that is updated based on load test results. Probabilistic within-site variability has not been applied commonly. This research examines differences in reliability outcomes between deterministic and probabilistic within-site variability.

Analysis of micropile load test results from five different sites was used to develop a distribution of within-site variability. The resulting distribution is relatively variable (i.e. the value of within-site variability is, itself, variable), which demonstrates that there is, in fact, considerable variability and uncertainty in the value of within-site variability.

A numerical tool was developed to perform reliability evaluations using both deterministic and probabilistic within-site variability. The tool uses site-specific load test information to update prior distributions of the mean and standard deviation of the foundation resistance. The prior distribution of the mean characterizes *among-site variability*, the variation in average resistance across sites that is inherent to empirical methods of foundation design. The standard deviation of foundation resistance is within-site variability. For deterministic within-site variability, a constant value is used. For probabilistic within-site variability, a prior distribution of the standard deviation represents the magnitude of within-site variability and the uncertainty in the magnitude. The numerical tool was used to perform Bayesian updating of various prior distributions used to

represent a wide range of site conditions from relatively low capacity to relatively high capacity and relatively uniform to relatively variable. Updating was performed for a range of load test outcomes. The posterior distributions were used to numerically estimate $R_{1/1500}$, the 1/1500th value of foundation resistance. $R_{1/1500}$ is the value at which 1/1500 values are smaller and 1499/1500 are larger. $R_{1/1500}$ is comparable to factored resistance from load and resistance factor design approach, although $R_{1/1500}$ does not account for load variability.

The findings support the hypothesis: for evaluations with many load tests, the posterior estimate of $R_{1/1500}$ with probabilistic within-site variability converges to the true value of $R_{1/1500}$ while the posterior estimate of $R_{1/1500}$ for deterministic within-site variability does not, unless the prior estimate of within-site variability happens to be correct. Simulations comparing the posterior estimates of $R_{1/1500}$ to the true value were characterized as underreliable when the estimated value of $R_{1/1500}$ was more than ten percent greater than the true value. For small numbers of tests, the proportion of underreliable outcomes for deterministic within-site variability ranged from 10 percent with low within-site variability to 30 percent with high within-site variability. In all cases, the proportion of underreliable cases with deterministic within-site variability was about twice the proportion with probabilistic within-site variability. The prevalence of underreliable outcomes for deterministic within-site variability is attributed to its neglect of uncertainty in within-site variability.

Results from this research suggest that the practically achievable reliability benefit of load testing is to reduce the effect of among-site variability. The results indicate significant reductions in among-site variability can be achieved for one load test. In fact, when the ratio of within-site variability to among-site variability is 0.2 and the nominal *COV* of the prior predictive distribution of resistance is 0.5, the prior value of $R_{1/1500}$ was increased by a factor of three, on average, after Bayesian updating from one randomly sampled load test. Additional simulations suggest the reliability benefit of load testing is significantly greater when total variability is dominated by among-site variability rather than within-site variability. The increase in $R_{1/1500}$ due to load testing is most significant for the first test, with diminishing increases for subsequent tests.

1. Introduction

Foundation load tests are commonly used as a basis for engineering design of foundations. Most designs based on load tests are not reliability-based, which limits evaluation of the benefits of load testing. Previous work has established probabilistic methods to evaluate the reliability of designs based on load tests. Most of the previous work to evaluate reliability of designs based on load tests has treated the variability of foundation resistance at a site, or within-site variability, as a constant, deterministic value. In fact, within-site variability is known to vary, which likely alters the results of reliability evaluations.

This research uses a numerical tool to perform reliability analysis of designs based on load tests both with and without the assumption of constant, deterministic within-site variability. The research also uses a large collection of micropile load test results to develop a probabilistic distribution of within-site variability. Results from the analyses are used to evaluate the effect of constant, deterministic within-site variability versus variable, probabilistic within-site variability on interpretations of foundation reliability. Additional explanation of the motivation is provided in this chapter, along with a statement of objectives and hypothesis, an overview of the research methodology, and a description of the contents of this dissertation.

1.1 Motivation & Background

Foundation design is an inherently uncertain business. Experience has shown foundation resistance and performance to vary widely, even for foundations designed and built to the same standards and even for foundations installed at the same site. Consider, for instance, the Lee Roy Selmon Crosstown Expressway in Tampa, Florida. During construction of an elevated portion of the expressway, one of the bridge piers suddenly sank more than 10 ft upon failure of the drilled shaft supporting the pier. A photograph of the bridge following the shaft failure is shown in Figure 1. The failed shaft was designed based on a boring located approximately 8 ft from the shaft. The boring indicated the shaft was socketed 2 ft into hard material, but coring through the failed shaft indicated more than 10 ft of soft material below the shaft.



Figure 1: Failure of a pier supporting an elevated portion of the Lee Roy Selmon Crosstown Expressway in Tampa, Florida (from Dapp et al., 2013).

Foundation engineers typically address uncertainties in foundation performance by collecting information to reduce the uncertainty, by adopting conservative design postures to reduce the chance of poor performance, or by some combination of both information gathering and conservatism. The most obvious example of information collection to reduce uncertainty is geotechnical site investigation to characterize the subsurface and develop models to predict foundation performance. An increasingly common form of information collection is foundation load testing, which involves installing a foundation element at the project site and observing its performance under controlled loading. The information from a foundation load test provides site-specific performance information, which generally is associated with a considerable reduction in uncertainty regarding foundation performance.

Another trend in foundation engineering is toward reliability-based foundation design. The objective of reliability-based design is to balance uncertainties and conservatism to achieve an appropriate level of risk. Accordingly, designs based on foundation load tests are commonly associated with more aggressive design postures, i.e. lower safety factors (or higher resistance factors). However, while the lower safety factors are consistent with the spirit of reliability-based

design (i.e. reduced uncertainty is associated with less conservatism), the specific values of the lower safety factors (or higher resistance factors) are almost always selected or assigned based on judgment rather than explicit probabilistic analysis. Explicit probabilistic analysis of designs based on load test information is desirable to facilitate “apples-to-apples” comparisons with probabilistic analysis of designs based on other methods (e.g. presumptive designs, designs based on predictive models with laboratory measurements, etc.). Comparison of different design methods that achieve the same reliability is the truest and fairest means for evaluating the relative merits of the different approaches to foundation design.

Previous studies have successfully performed probabilistic reliability analyses of designs based on load test information. The previous work has evaluated reliability for different types of load tests, different types of foundations, and different assumptions regarding available foundation performance information. The primary limitations of the previous work result from assumptions that have been invoked in the reliability analyses, typically to facilitate exact theoretical solutions or to simplify consideration of the problem. One particular assumption that has been applied commonly may have particular consequence to the results of the reliability analyses. The assumption is that within-site variability is a known, constant parameter rather than an uncertain, probabilistic parameter.

Within-site variability refers to the dispersion of foundation resistance values for foundation elements installed at the same site. Statistically, it can be defined as the variability of resistance for similar foundation elements installed at a site where the mean is known perfectly. Within-site variability is generally attributed to differences in geology across a site, and to differences in construction outcomes among foundations at a site. Within-site variability is commonly discussed in conjunction with *among-site variability*, which refers to the dispersion of the average foundation resistance across a collection of sites.

To understand the concepts of within-site variability and among-site variability, it is instructive to consider the relatively simple case of a presumptive design method. In a presumptive design, a single value of foundation resistance is assumed to apply across sites,

typically for a particular type of foundation installed in a particular type of ground (e.g. spread footings on clay, driven piles in sand, drilled shafts in rock, etc.). Figure 2 is a conceptual representation of data used to develop a presumptive model. The hypothetical presumptive model is based on data from a large number of sites. For each site, observations of foundation resistance are averaged and a single value is reported. The complete set of results – one average value from each site – is then averaged to produce a single value, μ_R , for presumptive designs. Application of the presumptive design value to a new site introduces two sources of variability. The first source of variability is introduced by using results from the “average” site (i.e. μ_R) to represent resistance at the new site. The uncertainty associated with using the average results from other sites at a new site depends largely on how variable foundation resistance is among different sites. This among-site variability is represented by the width (or spread, or dispersion) of the distribution of μ_R shown at the bottom of Figure 2. Among-site variability is primarily a result of geologic differences between locations; differences in geologic conditions produce different strengths, different construction results, and various other factors that impact foundation resistance.

The second source of variability is introduced by assuming one value of resistance for the entire design site. The validity of this assumption depends largely on how variable foundation resistance is within the site. The within-site variability is represented by the width of each of the site distributions shown in the top half of Figure 2; within-site variability is the variance of the site-specific distribution of resistance. As the distributions in Figure 2 indicate, some sites are more variable than others. Within-site variability results from geologic variation across a site and from variation in construction outcomes from one foundation element to another. The value of within-site variability for a given site and a given set of construction practices is generally fixed but strictly unknown. Within-site variability also depends on how a site is defined. If a site is defined by the project boundaries, greater within-site variability might be observed compared to the approach of splitting the project site into smaller “sites” for design purposes. By this reasoning, the designer has some subjective control of within-site variability, although some measure of within-site variability will always be present. As in other aspects of foundation design, the

engineer must weigh the benefits of increased precision with practical considerations related to the degree of complication of the design and its construction.

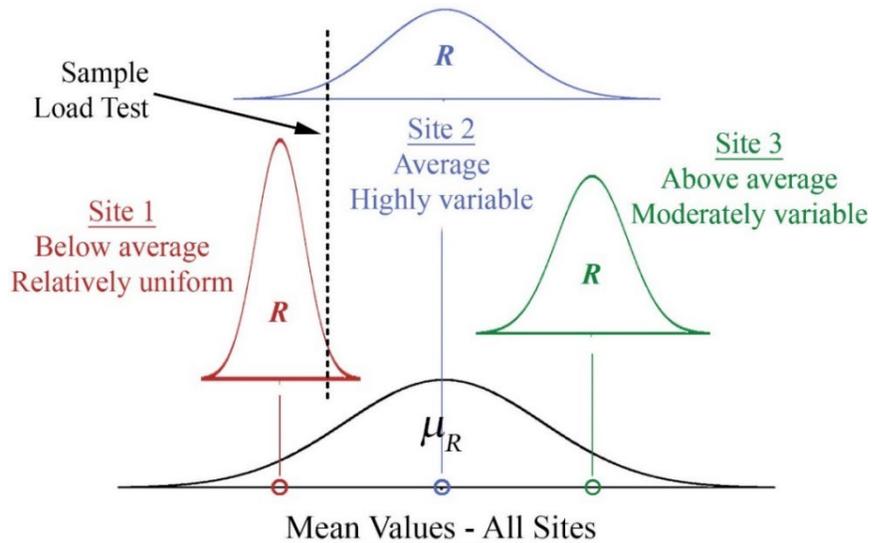


Figure 2: Conceptual representation of among and within-site variability for a presumptive model of foundation resistance.

By performing load tests, a designer is effectively sampling directly from the site distribution of foundation resistance; each test is a measurement sampled at random from a distribution like the ones for Sites 1, 2, and 3 in Figure 2. As for all sampling applications, the information garnered from a load test program generally improves with the number of measurements because the average load test result converges to the mean of the site distribution of foundation resistance. The rate of convergence is generally high; for most sites, a large number of tests is not required to establish a reliable estimate of the mean. The rate of convergence is greater for more uniform sites. Although the average result converges to the mean relatively quickly, a single load test result could be misleading. Consider the result represented by the dashed line in Figure 2. The result could be a relatively high result from Site 1, a relatively low result from Site 2, or, more likely, an average result from some site with mean resistance between that of Site 1 and Site 2. If multiple tests are performed at a site, it is possible to estimate within-site variability, but the number of tests required to produce a reliable estimate of within-site variability is greater than the number of tests required to produce a reliable estimate of the mean.

Estimates of within- and among-site variability provide useful context for consideration of foundation load test information, allowing a foundation engineer to rationally weigh empirical information from past performance at other sites with site-specific information at the project site. Such consideration can be performed informally, i.e. by weighing the evidence via engineering judgment, or formally through probability analysis using *Bayesian updating*. Bayesian updating is the basis for nearly all published reliability studies of foundation designs based on load tests, including this research.

Within-site variability is therefore an important input for reliability analysis of designs based on load tests. The assumption that within-site variability is known and constant simplifies analysis of foundation designs based on load tests. That the assumption has been employed often also likely reflects the fact that information regarding within-site variability is relatively limited. However, the assumption of constant within-site variability is demonstrably false: even previous efforts that have invoked the assumption have used a dataset of non-constant values of within-site variability to select a constant value for analysis.

The consequences of the constant within-site variability assumption have not been examined. The consequences could be significant. Presume the constant value of within-site variability that is assumed is, in fact, representative of the “average” site. Application of the average constant value to a highly variable site may inappropriately (1) give the foundation engineer too much confidence in the observed load test results and (2) result in underestimation of the variability of foundation resistance in reliability analysis of the design based on load test information. The second consequence would produce a design with a probability of failure greater than intended. Conversely, application of the average constant value of within-site variability to a relatively uniform site may inappropriately (1) lead the engineer to under-weigh the foundation load test information and (2) result in overestimation the variability of foundation resistance for reliability analysis. The second consequence would produce overly reliable designs that cost more than necessary to achieve the target probability of failure. Finally, even if application of the average constant value to an “average” site is, in fact, appropriate, the uncertainty in within-site

variability makes it impossible for an engineer to know that the assumed, constant value of within-site variability is appropriate. Failure to consider the variability of within-site variability, and in particular failure to consider the possibility that a site is considerably more variable than assumed, likely produces probability of failure values that are below the true value.

Evaluation of the consequences of assuming constant, known within-site variability will provide useful context for current methods of reliability analysis for designs based on load tests. Importantly, evaluating the effect of the common assumption may lead to reconsideration of the conclusions that have been stated regarding the reliability of designs based on load tests. If the consequences are significant – for example, if assumption of constant, known within-site variability leads to overestimating the reliability of designs based on load tests – the results could have a meaningful impact on the practice of load testing in foundation engineering.

1.2 Objective

The objective of this research is to evaluate the effect of assuming a deterministic value of within-site variability that is known and constant versus using a probabilistic value of within-site variability that is uncertain and subject to revision based on the results of load testing. The evaluation of the assumption focuses on the effect on the left tail of the resistance distribution, but effects on the mean and variability of foundation resistance are also considered in this research. The hypothesized effects associated with the objective are presented in the next section, followed by an overview of the methodology for achieving the objective and evaluating the hypothesis.

1.3 Hypothesis

The hypothesized consequences of the effect of deterministic consideration of within-site variability follow from the discussion of motivation in Section 1.1. The hypothesis is stated in terms of $R_{1/1500}$, the 1/1500th value of resistance, which is the resistance value for which 1/1500 of resistance values are lower and 1499/1500 of resistance values are greater. $R_{1/1500}$ is of interest because it characterizes the left tail of the foundation resistance distribution, and the left tail governs foundation reliability. $R_{1/1500}$ is similar to the concept of factored resistance in load

and resistance factor design (LRFD) methodology, although LRFD analyses would also consider load variability, which is not considered here. $R_{1/1500}$ is also comparable to allowable resistance from working stress design methods, which are typically not reliability-based.

For cases without load test information, assuming a deterministic value of within-site variability is hypothesized to yield estimates of $R_{1/1500}$ that are greater than estimates from probabilistic treatment of within-site variability, when all other factors (average resistance, among-site variability, and average value of within-site variability) are the same. As greater numbers of load tests are considered, probabilistic consideration of within-site variability is hypothesized to produce estimates of $R_{1/1500}$ that converge to the true value. In contrast, deterministic consideration of within-site variability is hypothesized to produce estimates of $R_{1/1500}$ that generally do not converge to the true value.

1.4 Research Methodology

The assumption of deterministic within-site variability is commonly invoked primarily for the sake of mathematical convenience. To avoid the need to invoke the assumption, a numerical tool was developed to perform Bayesian updating and reliability analysis of foundation designs based on load test information. The numerical tool performs Bayesian updating via numerical integration and reliability calculations via Monte Carlo simulations in order to eliminate strict statistical distribution requirements associated with previous efforts. The computational and distributional flexibility of the tool facilitates reliability analysis of foundation designs based on load tests with either deterministic or probabilistic consideration of within-site variability and without any strict distributional requirements.

Another reason for the common assumption of a single value of within-site variability is likely that observations of within-site variability are relatively rare. A single observed value of within-site variability requires multiple tests to failure of “identical” foundation elements at the same site, and such information is limited, at least in literature. To develop a realistic approach for probabilistic consideration of within-site variability, this research examines a large body of

micropile load test data scoured from literature and from submissions by foundation contractors, consultants, and transportation agencies. Micropiles were selected because load testing is performed relatively often for micropiles compared to other types of foundations. In particular, load tests that achieve failure are more common for micropiles; the size of micropiles makes application of failure loads more practical than for larger foundation types. The collection of micropile load test data is analyzed to produce a probabilistic distribution of within-site variability values for micropile foundations.

To evaluate the effect of deterministic-vs.-probabilistic within-site variability on Bayesian updating and reliability analysis of foundation designs based on load tests, the numerical tool is used to analyze simulated cases with both deterministic and probabilistic within-site variability. Both considerations of within-site variability are applied to the same set of cases, which cover a range of site conditions and load test scenarios. The site conditions vary in mean resistance (i.e. relatively low capacity to relatively high capacity) and within-site variability (i.e. relatively uniform to highly variable). The simulated site conditions are created to cover a range of among-site and within-site variability values that is informed by empirical data. Two methods are used to simulate load test results from each simulated site. In the first, test results perfectly satisfy the assumed site conditions (i.e. mean and standard deviation). In the second, test results are randomly sampled from the assumed site conditions. The first method is useful for evaluating why certain trends are observed; the second method is useful for evaluating the practical implications of the trends.

The effect of within-site variability assumption is evaluated by comparisons of the updated distributions of foundation resistance – one for deterministic within-site variability, one for probabilistic within-site variability – with the “true” distribution of resistance. For each simulated site, the true distribution of foundation resistance is defined by the mean parameter and the standard deviation parameter. Thus the true distribution is known perfectly for the purposes of the research. Importantly, in real applications of Bayesian updating for foundation resistance, the parameters are never known perfectly. As explained in the hypothesis (Section 1.3), the

evaluations focus on the left tail of the updated distributions of foundation resistance, which governs foundation reliability. The left tail is characterized using $R_{1/1500}$, which was defined in the statement of hypothesis.

Comparison of estimates of $R_{1/1500}$ computed for each case with the true value of $R_{1/1500}$ facilitates evaluations of the effect of treatment of within-site variability on reliability analyses for various site conditions and numbers of tests. To evaluate the hypothesis, the estimates of $R_{1/1500}$ are compared for deterministic within-site variability and probabilistic within-site variability under different assumptions of site conditions and testing results. Comparisons focus on results without testing and results for large numbers of tests for the purpose of evaluating the hypothesis, but other meaningful conclusions are evident from comparisons at small numbers of tests, for example in examining the magnitude of errors in estimates of $R_{1/1500}$ for small numbers of tests.

1.5 Organization of Dissertation

This chapter has provided an introduction to the motivation for this research and an overview of the research objectives, hypothesis, and methodology. In Chapter 2, additional background information and a review of literature are presented. The chapter focuses primarily on Bayesian updating, beginning with an introduction to Bayes' Theorem and an overview of the application of Bayes' Theorem to foundation design with load tests before proceeding to a detailed review of literature on the topic. Other foundation reliability topics, including a summary of available literature regarding within-site variability, are also covered in Chapter 2. Chapter 3 presents the results of gathering and characterizing within-site variability information for micropile foundation systems, including a comparison of the findings with previous published distributions of within-site variability.

Chapter 4 presents details of the numerical tool for Bayesian updating and reliability analyses of foundation designs based on load test information. In Chapters 5 and 6, the results of Bayesian updating and reliability analyses are compared for cases evaluated with deterministic and probabilistic within-site variability, consistent with the discussion of methodology in Section

1.4. The cases are considered for different numbers of load tests and for various site conditions (i.e. different prior methods of predicting resistance, different “true” values of mean foundation resistance, different “true” values of within-site variability). While Chapter 5 addresses cases with load test results perfectly satisfying assumed site conditions, Chapter 6 addresses cases with load test results randomly sampled from assumed site conditions. The dissertation closes with Chapter 7, which summarizes significant findings and conclusions and makes recommendations for foundation engineering practice and for future research.

2. Literature Review

Bayesian updating for foundation designs based on load test information is typically performed in conjunction with analysis of foundation reliability, the first topic of this chapter. After introducing the concept of foundation reliability, an introduction to Bayes' theorem and Bayesian inference is presented. After the general presentation of Bayesian updating, a conceptual overview of Bayesian updating of foundation designs based on load test information is presented prior to a more detailed review of literature regarding the topic. The final topic in this chapter is within-site variability, with Section 2.5 summarizing published information regarding within-site variability.

2.1 Foundation Reliability

For engineering applications, reliability is typically defined as the probability of success. Reliability, r , is therefore the mathematical complement of probability of failure, p_f :

$$r = 1 - p_f \quad \text{Eq. 2-1}$$

For design of structures, including foundations, failure is defined as an event that achieves a limit state. As shown in Figure 3 both load effects, Q , and resistance, R , are uncertain quantities that can be considered as random variables that contribute to the reliability of a particular foundation. A performance function, g , is often established to facilitate calculation of the probability of failure, as shown in Figure 3(b). For strength limit states, the performance function is often defined as the difference between the resistance and load:

$$g = R - Q \quad \text{Eq. 2-2}$$

With this definition, g is sometimes referred to as the safety margin (e.g. Allen, 2005), with negative values of g indicating failure as shown by the shaded region in Figure 3(b). Thus, the probability of failure can be computed as $p_f = p[g < 0]$.

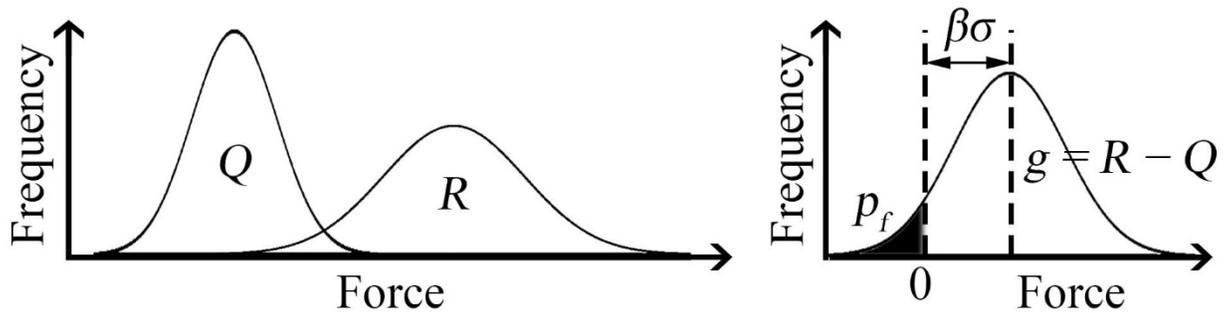


Figure 3: Overview of reliability: (a) distribution of load effect, Q , and resistance, R , and (b) distribution of the safety margin, g .

Because reliability for a particular design is dependent on the nominal magnitude of both load effects and resistance, as well as the uncertainty in load effects and resistance, design reliability can be improved in several ways. The most obvious and direct means for improving reliability as part of foundation design is to increase the magnitude of foundation resistance, effectively shifting the distribution for R in Figure 2.1(a) to the right. Reliability can also be improved by reducing the magnitude of load effects, which shifts the distribution for Q to the left. Reducing the variability and uncertainty associated with the distributions will also improve reliability. For instance, designs based on presumptive values are typically more uncertain than designs based on values predicted from empirical relationships (e.g. driven pile side resistance in clay as a function of undrained shear strength). Predictive values of foundation resistance are, in turn, generally more uncertain than designs based on site-specific load tests. Accordingly, one would expect designs based on presumptive values to be the least reliable and designs based on load tests to be the most reliable, with the reliability of designs based on predictive values between the others.

Reliability is often represented using the reliability index, β , instead of the probability of failure. As shown in Figure 3(b), the reliability index represents the number of standard deviations between the mean value of the safety margin and zero. Greater values of the reliability index therefore correspond to lesser values of the probability of failure. If both load and resistance follow normal distributions, the safety margin is also normally distributed and the probability of failure and the reliability index are related analytically through the standard normal cumulative distribution function, designated by Φ (Baecher and Christian, 2003):

$$p_f = \Phi(-\beta) \quad \text{Eq. 2-3}$$

As shown in Figure 4, if other distributional assumptions are made for the safety margin, the relationship between probability of failure and reliability index changes. For reliability index values less than 2.0, the effect of distribution is limited. For reliability index values greater than 2.0, the assumption of having normally distributed load effects and resistances results in greater probabilities of failure than would be true if load effects and resistances were log-normally distributed, and the difference increases with increasing variability of the lognormal distribution. In Figure 4 and commonly throughout this report and elsewhere, variability is represented using the coefficient of variation, or *COV*. The *COV* is the ratio of the standard deviation to the mean and is therefore a normalized measure of variability:

$$COV = \frac{\sigma}{\mu} \quad \text{Eq. 2-4}$$

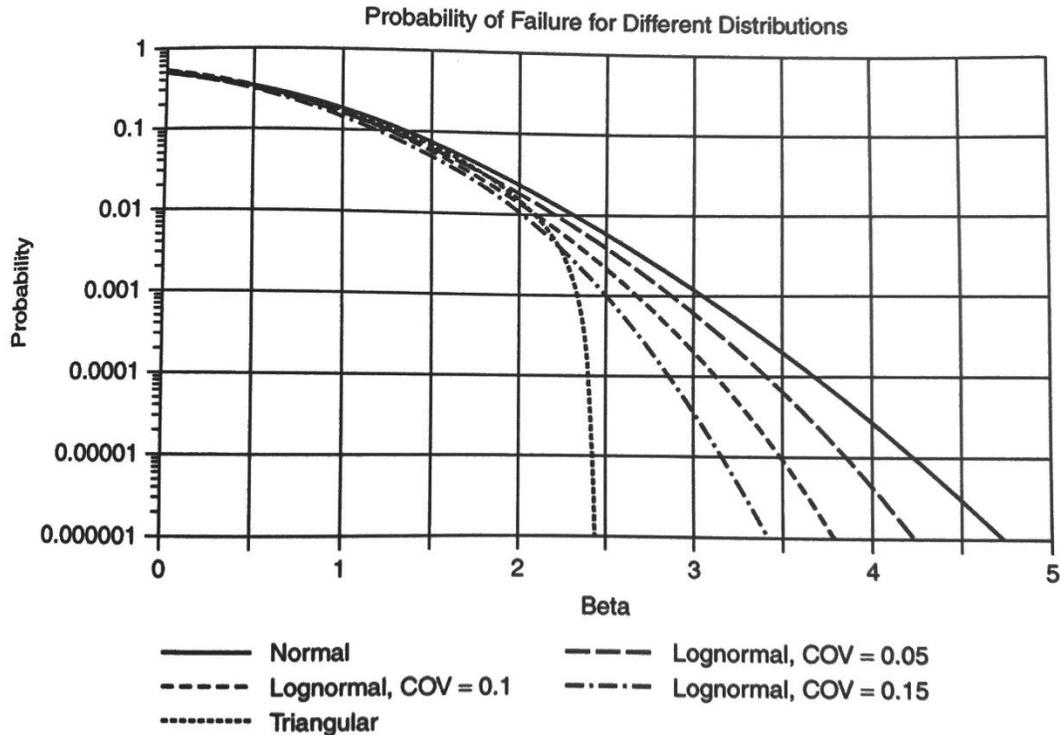


Figure 4: Probability of failure versus reliability index for various safety margin distributions. From Baecher and Christian (2003).

All civil engineering design philosophies intend to achieve reliable structures, but the means for ensuring reliability varies. Allowable stress design (ASD) methods achieve reliability by applying a factor of safety, defined as the ratio of available resistance to applied load, with greater factors of safety resulting in greater reliability. Traditionally, ASD methods have not explicitly considered quantitative reliability, but engineering judgment commonly and appropriately dictates greater factors of safety for more critical structures and for designs with greater uncertainty in load and resistance values. Load and resistance factor design (LRFD) methods were conceived as a more explicit and formal means for considering design uncertainties. Whereas ASD methods apply one factor, the factor of safety, to designs in order to achieve reliability, LRFD methods apply separate factors to expected values of load and resistance. The separation of load factors and resistance factors is intended to reflect the separate uncertainties that surround each quantity. In LRFD for strength limit states, a design is considered satisfactory if the sum of the factored load effects is less than the factored resistance: $\sum \gamma Q \leq \phi R$, where γ is a load factor and ϕ is a resistance factor. Load factors are typically greater than one, with greater load factors being applied to more variable or uncertain load effects. Resistance factors are typically less than one, with smaller resistance factors applied to resistances having greater uncertainty. Greater load factors and/or smaller resistance factors produce greater reliability (or lower probability of failure) for a given design. For foundations, resistance factors can be calibrated by fitting to required factor of safety values from historical ASD practice, or they can be calibrated probabilistically to achieve some selected target reliability.

2.1 Bayes' Theorem and Bayesian Inference

Bayes' theorem is a powerful tool for updating probabilities in consideration of new information. Bayes' theorem underpins a major branch of the field of probability and statistics with applications throughout the sciences. Probably the most familiar form of Bayes' theorem is presented in Eq. 2-5:

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)}, \text{ where} \quad \text{Eq. 2-5}$$

A and B are events

$p(A|B)$ is the conditional probability of observing A given B

$p(B|A)$ is the conditional probability of observing B given A ,

$p(A)$ is the probability of observing A (independent of B), and

$p(B)$ is the probability of observing B (independent of A).

With this form, the conditional probability of A given B reflects the updated probability of observing A in light of the information provided by the occurrence of event B . The updated probability of observing A is calculated from the likelihood of the new information (event B) given the prior information regarding event A , and from the independent probabilities of events A and B . Consistent with that explanation, $p(B|A)$ is commonly referred to as the *likelihood* and $p(A)$ is commonly referred to as the *prior*.

Bayes' theorem is commonly applied to medical tests, e.g. a test for strep throat. In this example, A is the event that a patient is infected with strep throat and B is the event that the patient tests positive for strep throat. The conditional probability $p(A|B)$ then represents the probability of the patient being infected given a positive test outcome. The conditional probability depends on $p(B|A)$, which represents the chance of a true positive test; $p(A)$, which represents the percentage of the population infected with strep throat; and $p(B)$, which represents the total probability of observing a positive strep tests among all patients, including those who are not infected. Better medical tests generally have greater values of $p(B|A)$, which implies a small chance of false positives, and values of $p(B)$ close to $p(B|A)$, which implies a small chance of false negatives.

The field of Bayesian inference is a specific application of Bayes' theorem. In Bayesian inference, the parameters of probability distributions (e.g. the mean of the normal distribution, the standard deviation of the normal distribution, the probability of success of the binomial distribution, etc.) are treated as random variables. This is a departure from the more familiar form

of statistical inference, the frequentist approach. In the frequentist approach to statistical inference, probability distribution parameters are treated as deterministic constants. An insightful discussion comparing Bayesian inference and the frequentist approach is provided by Bain and Engelhardt (1992). Bayes' theorem is re-stated in Eq. 2-6 for application to Bayesian inference:

$$p(\theta|D) \propto p(D|\theta) \cdot p(\theta) \quad , \text{ where} \quad \text{Eq. 2-6}$$

$p(\theta|D)$ is the probability of the parameter(s) θ given the data D , i.e. the posterior distribution of the parameter(s),

$p(D|\theta)$ is the probability of observing the data given prior estimates of the parameter(s), also referred to as the likelihood function, and

$p(\theta)$ is the probability of the parameter(s), i.e. the prior distribution of the parameters(s)

The proportional operator is used in Eq. 2-6 because the right side of the equation must be normalized by a constant (equal to $p(D)$) in order to produce a probability distribution function (i.e. to satisfy $\int_{-\infty}^{\infty} f(x)dx = 1$).

As presented in the rest of this chapter, Bayesian updating is the theoretical basis for nearly all of the literature regarding foundation reliability analyses with load test information. To apply Bayesian updating, a "prior distribution" representing knowledge of the parameter prior to testing is estimated based on previous data, expert opinions, or other means. New data regarding the value of the parameter are collected. Using Bayes' theorem, the prior distribution is updated to reflect the new information, yielding a "posterior distribution" of the parameter that represents knowledge following collection of test data. In most circumstances, the posterior distribution is associated with less variability than the prior distribution, which is evidence of the value of the new information. Tang (1971) was among the first to recommend applying Bayesian inference for geotechnical engineering design, citing the reduction in variability as motivation for adopting Bayesian inference. The application of Bayes' theorem to foundation load test information is the primary topic of the rest of this chapter.

2.2 Bayesian Updating Framework for Foundation Design with Load Test Information

Before presenting details of the literature regarding Bayesian updating based on load test information, it is helpful to introduce the conceptual framework for the topic. The conceptual framework applies the Bayesian inference topic of the last section to the concepts of within-site variability and among-site variability introduced in Chapter 1. This section also functions as an introduction to the two immediately following it, which present details from published literature regarding Bayesian updating based on foundation load test information. The methods presented in literature all generally conform to the framework presented in this section.

Foundation load testing is generally considered an effective method for improving the reliability of foundation designs. Primarily, the reliability is improved by reducing the among-site variability associated with predictions of foundation resistance. If no load tests are performed and a presumptive design method is applied, the design is uncertain because of the among-site variability as shown for the distribution of μ_R in Figure 2, and because of within-site variability, which is not wholly captured by the distribution of μ_R . If load tests are performed, among-site variability is reduced. Designs based on load tests have site-specific information as their primary basis; there is no among-site variability in load test information. However, it is important to note that some amount of among-site variability will be present if the designer considers presumptive or predictive design information in addition to the load test results, as is usually the case.

When considering within-site variability, it is generally important to consider that the definition of “site” should not necessarily refer to the area defined by project boundaries. For large bridge projects, for instance, it is often advantageous to define several different “sites” within the overall project footprint in order to establish more uniform conditions within each site. Lumping all project conditions into one site could produce a distribution of resistance that resembles that for Site 2 in Figure 2 whereas creating several sites for the project might produce distributions of resistance with more uniformity, similar to conditions for Site 1. AASHTO LRFD *Specifications* (2014) encourage such an approach.

Most of the published approaches for considering the reliability of designs based on load tests apply Bayesian updating to the mean of the foundation resistance, following a general methodology for updating geotechnical design parameters first presented by Tang (1971). The procedure is outlined in Figure 5. In the first step of the procedure, the prior distribution of the mean foundation resistance is established. The mean foundation resistance, μ_R , is the same parameter considered in the previous section and shown in Figure 2. In terms of the Bayesian inference (Section 2.1), the prior distribution of the mean represents $p(\theta)$ from Eq. 2-6, where the parameter to be updated is the mean foundation resistance: $\theta = \mu_R$.

The prior distribution of μ_R is typically assumed to follow a lognormal distribution, with a mean value from an appropriate rational prediction method (e.g. beta method for drilled shafts, Meyerhoff's SPT method for driven piles, etc.) and the standard deviation from published data regarding the variability of the prediction method. Alternatively, the prior distribution of mean foundation resistance could be established from presumptive values. The mean of the prior distribution is, in fact, a mean of the mean: it is the resistance we would expect to observe, on average, at the "average" site. The variability of the prior distribution of the mean primarily reflects variability among-sites; the effect of within-site variability on the prior distribution of the mean is diminished by the averaging performed in developing the prior method.

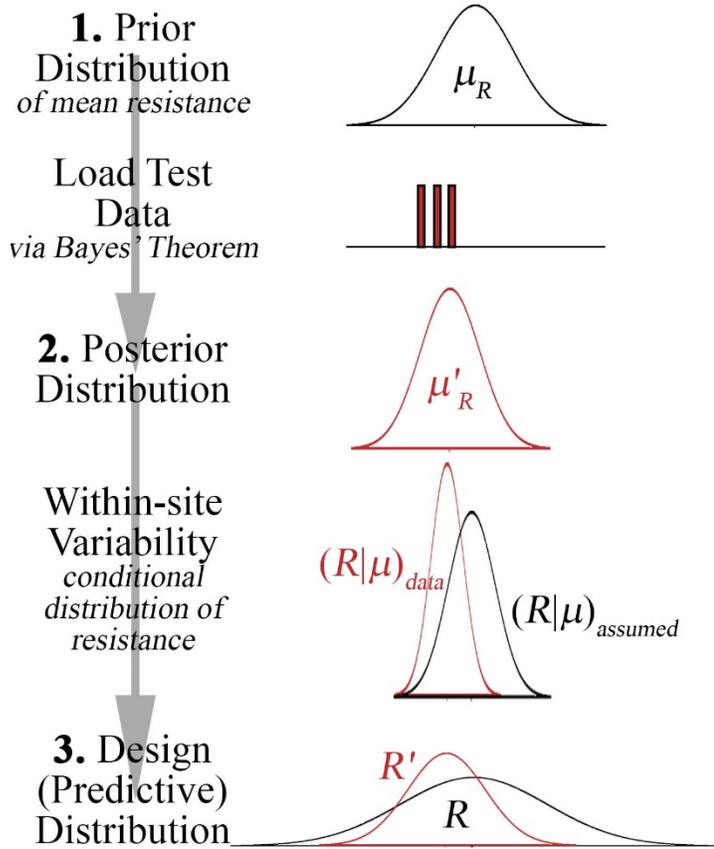


Figure 5: Procedure for application of Bayesian updating to foundation design based on load test information. Distributions shown in black are for a presumptive design without load tests. Data and distributions in red are for updating based on three load tests.

In the second step of procedure, information from load tests is used to update the prior distribution of the mean foundation resistance by way of Bayes' theorem and a likelihood function. The likelihood function, $p(D|\theta)$ from Eq. 2-6, describes the probability of observing the data, given the prior distribution. Methods for establishing the likelihood function depend on the nature of the load test information (e.g. if failure is achieved during the test), and potentially on assumptions made to simplify calculations. Techniques for defining the likelihood functions vary among the literature; the techniques are described in the following sections. Importantly, the updated distribution, like its prior, describes the distribution of the *mean* foundation resistance. In terms of Eq. 2-6, $p(\theta|D) = p(\mu_R|\text{load tests})$. Typically, the variability of the updated distribution of μ_R is less than the variability of the prior distribution, which primarily reflects among-site variability. The reduction in variability reflects the benefit of site-specific load test information.

In order to evaluate the reliability of designs based on load tests, it is necessary to consider the distribution of the individual values of foundation resistance across the project site, rather than just the distribution of the mean. In the third step outlined in Figure 5, the site-specific distribution of foundation resistance is calculated. The site-specific distribution of foundation resistance defines the collection of resistance values one would expect to observe at a site considering the prior information, updated for load test information, and considering the anticipated within-site variability. This collection of resistance values is often termed the *predictive distribution*. The predictive distribution describes the resistance values that one would expect to observe if many load tests were performed at the site; the predictive distribution is therefore the basis for reliability calculations. Constructing the predictive distribution from the updated distribution of the mean is the third step of the Bayesian updating procedure. To construct this distribution, the *posterior predictive distribution*, it is assumed that the predictive distribution has the same mean as the distribution of the updated mean resistance. The variability of the predictive distribution includes the variability of the mean foundation resistance from the updated distribution (i.e. the reduced among-site variability) as well as within-site variability.

In the procedure outlined in Figure 5 and described above, Bayes' theorem is applied only to the distribution of mean foundation resistance. The other parameter defining foundation resistance, within-site variability, is treated as a deterministic parameter and is therefore not updated. As explained in the Introduction, the approach of updating only the mean and treating within-site variability deterministically has been adopted in nearly every published study of foundation reliability. A prominent exception is the work by Baecher and Rackwitz (1982). Baecher and Rackwitz used load test information to perform Bayesian updating of both the mean parameter and the standard deviation parameter, where the standard deviation parameter represents within-site variability. Details of the approach by Baecher and Rackwitz are presented in Section 2.3.3. The effect of deterministic versus probabilistic treatment of within-site variability on reliability is the primary focus of this research.

The literature described in the two sections immediately following this one documents various methodologies for Bayesian updating of foundation designs based on load tests. Two types of load tests are considered: tests that achieve failure, and tests that do not achieve failure. In statistics parlance, tests that achieve failure are described as “uncensored,” and tests that do not achieve failure are “censored” because test result represents a lower bound on resistance rather than a specific value of resistance. Methods for Bayesian updating based on load tests generally fall into two categories. The first category deals strictly with uncensored load tests. The second category addresses “proof” load tests, which may be either censored or uncensored. Literature regarding Bayesian updating for uncensored tests and proof tests are presented in the two sections below, respectively.

2.3 Bayesian Updating of Foundation Resistance for Load Tests Achieving Failure

The three publications described in this section introduced the two primary approaches for performing Bayesian updating of foundation resistance for load tests that achieve failure. Kay (1976) introduced the first such method, which he developed specifically for driven piles in sand. Zhang and Tang (2002) presented a method that is essentially the same approach presented by Kay, but mathematically more general. Baecher and Rackwitz’s (1982) approach differs from Zhang and Tang’s approach primarily in its treatment of within-site variability. Baecher and Rackwitz treated within-site variability as a probabilistic parameter to be updated, whereas Zhang and Tang treated within-site variability as a deterministic constant. Details of each method are presented in this section. At the end of this section, the three methods are summarized and compared.

The presentations include equations and symbols that follow the notation presented in each of the respective publications. The symbols and notation from the literature are different from the notation used throughout the rest of this dissertation. The symbols and notation used for this research are described in Chapter 4.

2.3.1 Kay (1976)

Kay (1976) published the first probabilistic evaluation of geotechnical foundation designs based on load test information. Kay generally followed the approach outlined in the Section 2.2, applying it to driven piles in sand. Kay worked in terms of foundation resistance bias, defined as the ratio of measured to predicted foundation resistance:

$$bias = \frac{R_{measured}}{R_{predicted}}, \text{ where} \quad \text{Eq. 2-7}$$

$R_{measured}$ is the measured foundation resistance and

$R_{predicted}$ is the predicted foundation resistance.

Kay presented bias statistics for several static and dynamic methods for predicting pile resistance and used the statistics to establish prior distributions of the mean pile resistance. The methods were divided into three groups based on the variability of the bias, with the least variable (most reliable) methods in Group 1 and the most variable (least reliable) methods in Group 3. Group 1 methods had bias standard deviations around 0.3, Group 3 methods had bias standard deviations around 1, and Group 2 methods were between Groups 1 and 3. Kay also estimated within-site variability by considering published data from three sites with multiple load tests of similar piles. He concluded the distribution of the logarithm of individual pile resistance should be assumed to be normal with a standard deviation of 0.12. Mathematically, this is equivalent to the arithmetic (i.e. non-log) individual pile resistance following a lognormal distribution and having a *COV* equal to 0.12. Kay acknowledged the data supporting the within-site variability value recommendation were limited. The data regarding within-site variability is discussed in more detail in Section 2.5. Kay calculated the likelihood function for Bayesian updating by assuming the logarithm of pile resistance and its mean follow a normal distribution.

The crux of Kay's paper is a Bayesian procedure for calculating allowable pile resistance from a prior estimate of the mean pile resistance, load test results, and a factor of safety. The factor of safety values recommended by Kay depend on the target reliability index, the number of load tests, and the prediction method for the prior estimate. As shown in Figure 6, the required

factor of safety decreases as the number of load tests increases, and more reliable prediction methods (i.e. those from Group 1) are associated with lower required factors of safety. The sharp decrease in required factor of safety resulting from performing one load test indicates load test information is quite valuable, particularly if prior information is from an unreliable estimate. Kay's procedure assigns little marginal value to subsequent load tests, as evidenced by the flattening of curves in Figure 6 beyond one test. The marginal value is unsurprising considering the relatively small value of within-site variability assumed by Kay; if the variability of pile resistance is known to be small within a site, one test is greatly informative. The values in Figure 6 do not consider any variability of loading.

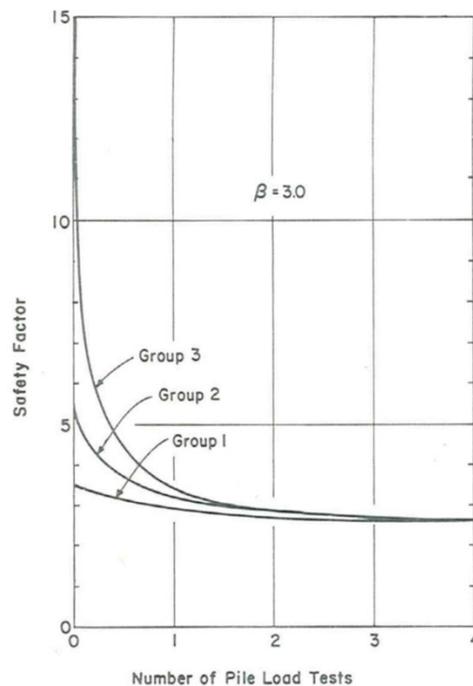


Figure 6: Required factor of safety recommended by Kay (1976) based on Bayesian updating of mean pile resistance from load test information.

2.3.2 Zhang and Tang (2002)

Zhang and Tang (2002) presented a method for evaluating the reliability of designs based on load tests that is analytically similar to Kay's (1976) procedure, but more general in its presentation, primarily in its applicability to all types of foundations. Zhang and Tang assume the ratio of measured to predicted foundation resistance, which they denote x , follows a lognormal distribution. Note that x is mathematically equivalent to the foundation resistance bias defined in

Eq. 2-7. Instead of working with the foundation resistance directly, Zhang and Tang worked in terms of the ratio of measured to predicted resistance for mathematical convenience. Also for mathematical convenience, Zhang and Tang mostly worked in terms of the natural logarithm of x ; since x is assumed to follow a lognormal distribution, $\ln x$ follows a normal distribution by definition of the lognormal distribution. The normal distribution is associated with theoretical solutions for Bayesian calculations; the lognormal distribution is not.

The mean of x , μ , is the parameter to be updated with Bayes' theorem. The standard deviation of x , σ , represents within-site variability. Zhang and Tang assumed σ to be known and constant, the same assumption employed by Kay. Unlike Kay's procedure, Zhang and Tang's method is presented generally without assigning a specific value for within-site variability. Throughout their analyses, Zhang and Tang assume a value of σ equal to 0.2, which they support with within-site variability data from Kay (1976) and Evangelista (1977). The data regarding within-site variability is discussed in more detail in Section 2.5.

The parameter to be updated, μ , is assumed to follow a normal distribution with its own mean μ' and standard deviation σ' . Prior estimates for the parameters μ' and σ' are from bias statistics for a given static or dynamic pile resistance prediction method. For unbiased methods, μ' is 1. The prior variability of μ , σ' , represents among-site variability. This framework for Bayesian updating and assignment of site variability sources is consistent with the discussion in Section 2.2.

Zhang and Tang assume n load tests are completed, each representing one observation of x (the foundation resistance normalized by the predicted value). Based on the assumption of constant, known within-site standard deviation σ , the exact theoretical solution for the updated parameters of μ can be calculated using the Bayesian conjugate for the normal distribution (e.g. Ang and Tang, 2007). Zhang and Tang denote the updated (posterior) parameters of μ as μ'' , the updated mean of μ , and σ'' , the updated standard deviation of μ . As shown in Eq. 2-8 and Eq. 2-9, the updated parameters are a function of the average load test result, \bar{x} , the number of tests (n), the prior parameters (μ' and σ'), and the assumed value of within-site variance (σ^2).

$$\mu'' = \frac{\bar{x}(\sigma')^2 + \mu' \left(\frac{\sigma^2}{n} \right)}{(\sigma')^2 + \left(\frac{\sigma^2}{n} \right)} \quad \text{Eq. 2-8}$$

$$\sigma'' = \frac{\sqrt{\left(\frac{\sigma^2}{n} \right) (\sigma')^2}}{\sqrt{\left(\frac{\sigma^2}{n} \right) + (\sigma')^2}} \quad \text{Eq. 2-9}$$

The updated mean value of μ , μ'' , is based on information from both the prior estimate and the load test information, with each source weighted by the variability of the other so that highly variable prior estimates will result in greater consideration of load test results and highly variable sites will result in greater consideration of prior estimates. Greater numbers of load tests reduce the effect of within-site variability (because $\frac{\sigma^2}{n}$ decreases with the number of tests), which increases consideration of load test information and reduces the updated standard deviation of μ , σ'' . Regardless of the number of tests, the value of within-site standard deviation, σ , is constant.

To evaluate foundation reliability, Zhang and Tang used the updated distribution of μ to update the predictive distribution of individual foundation resistance at the site. Zhang and Tang assume the individual foundation resistance follows a lognormal distribution with a mean value corresponding to μ'' (but transformed to lognormal) and a standard deviation corresponding to $\sqrt{\sigma''^2 + \sigma^2}$ (also transformed to lognormal), so that the variability of the predicted foundation resistance includes the updated variability of the mean as well as the assumed within-site variability. Zhang and Tang applied mean value, first-order second-moment method to evaluate reliability from the individual foundation resistance distributions.

In their presentation of results, Zhang and Tang use *RMP* (from Ratio of Measured to Predicted resistance), which is equivalent to the variable x they used to present the methodology and equations. Example results from Zhang and Tang are shown in Figure 7, which includes distributions of individual, predictive *RMP* (rather than distributions of mean *RMP*, or μ). The prior distribution *RMP* is based on Meyerhoff's SPT method, a conservatively biased prediction method

with a mean *RMP* value of 1.3 (i.e. on average, measured values are 30 percent greater than predicted values) and a *COV* of 0.5. Three posterior distributions of *RMP* are shown. All three are based on performing one load test, with one posterior based on an observed *RMP* of 0.5 (i.e. $\bar{x} = 0.5$), another based on an observed *RMP* of 1.0, and the other based on an observed *RMP* of 1.3. The effect of load test information on the mean of the posterior distribution is intuitive. Greater observed values of *RMP* yield greater posterior mean values, with the posterior mean between the prior mean and the observed *RMP* (except when the observation is equal to the prior, in which case the posterior mean is virtually unchanged). The effect of load test information on the variability of the posterior distribution is perhaps less intuitive. The posterior distribution for the case of observed *RMP* of 0.5 – the most unexpected result – has the smallest standard deviation, considerably less than that for *RMP* of 1.3, the expected result. This follows from all posterior distributions having the same *COV*, which Zhang and Tang attribute to all three posterior distributions having the same values of prior variability, within-site variability, and number of load tests. Zhang and Tang note the equation for σ'' does not depend on the load test results.

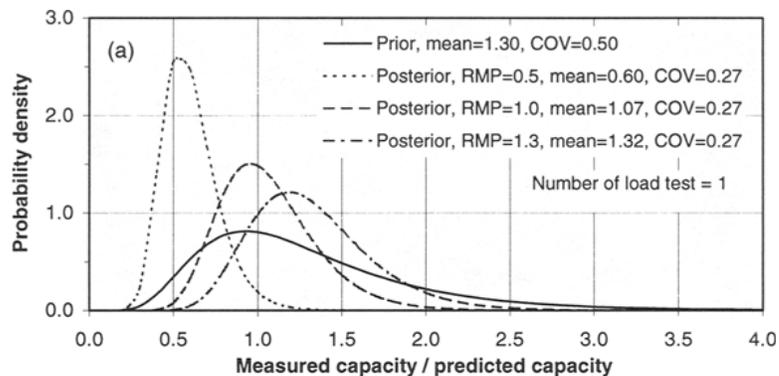


Figure 7: Example distributions of *RMP* prior and posterior to conducting one load test from Zhang and Tang (2002).

Zhang and Tang applied mean value, first-order second moment reliability methods and their load test updating methodology to evaluate the effect of load test results on design reliability. Results of their analysis are shown in Figure 8, which includes plots of the required factor of safety to achieve a specified value of reliability index (2.0 or 2.5) as a function of the observed load test result (as *RMP*) and the number of load tests. The required factor of safety decreases as

the load test results increase. This trend results from the mean value of foundation resistance increasing with the observed load test values. Zhang and Tang's results indicate little marginal value in load testing beyond the first test. For a given value of the reliability index, the required factor of safety lines for different numbers of tests are essentially atop one another. If numerous load tests were more valuable, the additional tests would have the effect of further decreasing the required factor of safety, resulting in more space between lines. The observation of small marginal value of load testing beyond the first test is consistent with the results presented by Kay (1976), as discussed above. A possible explanation for this observation is that both Kay and Zhang and Tang assumed relatively small values of within-site variability and relatively variable prior methods that correspond to high among-site variability.

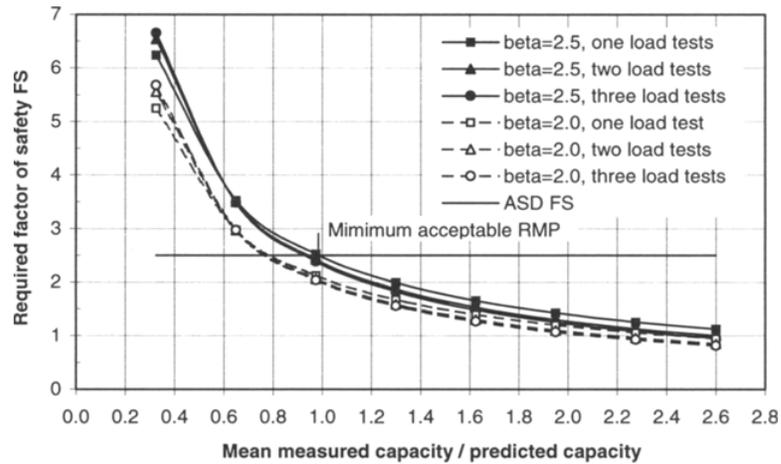


Figure 8: Required factor of safety to achieve target reliability index (beta) as a function of observed RMP and number of tests from Zhang and Tang (2002).

2.3.3 Baecher and Rackwitz (1982)

A method by Baecher and Rackwitz (1982) uses load test information to update not only average foundation resistance but also within-site variability. Baecher and Rackwitz assumed the logarithm of the ratio of measured to predicted foundation resistance follows a normal distribution with mean μ and standard deviation σ . Baecher and Rackwitz's distribution of foundation resistance is therefore equivalent to that assumed by Kay (1976) and Zhang and Tang (2002), but Baecher and Rackwitz applied Bayesian updating to both μ and σ , whereas the other methods updated only μ .

Updating σ in addition to μ adds considerable complexity. Instead of working in terms of the parameter σ , Baecher and Rackwitz use precision, h , which is simply the reciprocal of variance: $h = \frac{1}{\sigma^2}$. Precision was used rather than standard deviation in order to facilitate theoretical solutions for Bayes' theorem. Baecher and Rackwitz presented prior information on precision as shown in Figure 9. The precision data come from the same within-site variability datasets presented by Kay (1976) and Zhang and Tang (2002), but with additional information from a German publication (Petermann et al., 1958). The distributions fitted to the data shown in Figure 9 are gamma. The broader distribution is for all data, while the narrower distribution excludes three sites with the greatest precision (i.e. the least variable sites). Data regarding within-site variability is discussed in more detail in Section 2.5.

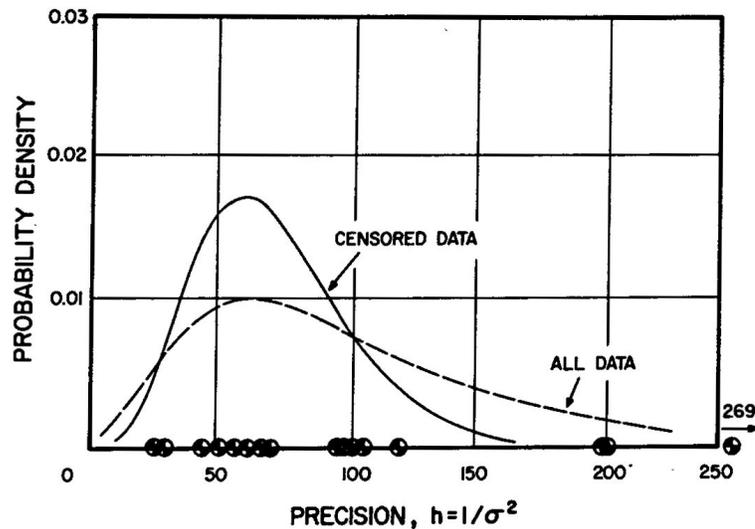


Figure 9: Prior distributions of site precision, h , from Baecher and Rackwitz (1982). Precision is the reciprocal of within-site variance.

Baecher and Rackwitz assumed the prior joint distribution of μ and h to be normal-inverse gamma, a four-parameter probability distribution function that primarily (if not exclusively) is used because it is the Bayesian conjugate prior for the normal distribution when both normal parameters are to be updated. (In contrast, the normal distribution is its own conjugate prior when the standard deviation is assumed to be known, as in the methods by Kay (1976) and Zhang and Tang (2002).) Two of the four normal-inverse gamma parameters are related to the mean of the foundation resistance, and the other two are related to the standard deviation of the foundation

resistance. Prior values of the mean parameters are estimated from databases of bias values for the foundation resistance prediction method (e.g. Meyerhoff's SPT method for driven piles), similar to the approaches used by Kay (1976) and Zhang and Tang (2002). Prior values of the standard deviation parameters are estimated from precision data like that presented in Figure 9.

Application of Bayes theorem results in updated values of the four normal-inverse gamma parameters to reflect the likelihood of observing the results of n load tests. The parameters are updated using both the average load test result as well as the sample variance of the load test results; this is in contrast to the approaches by Kay (1976) and Zhang and Tang (2002), which consider only the average load test result, not the variability of the results. After updating the normal-inverse gamma parameters describing the distribution of μ and h , the predictive distribution of individual foundation resistance is computed. They show the updated predictive distribution follows a Student's t -distribution by integrating over the joint distribution of the foundation resistance ratio (equivalent to RMP used by Zhang and Tang) and the updated distribution of μ and h . In general, the Student's t -distribution is similar to the normal distribution, but has heavier tails. That the updated predictive distribution of individual foundation resistance follows a Student's t -distribution is appropriate: the Student's t -distribution is typically used in frequentist statistics to make inferences on the mean when the variance of the population is unknown.

Baecher and Rackwitz used the predictive t -distribution to evaluate the effect of load test results on design reliability. The results from the t -distribution were adjusted to account for the tail area differences between the t and normal distributions so that reliability index values would be comparable to those from other methods. Results of their analysis are shown in Figure 10 and Figure 11. All four plots shown in the two figures are graphs of required factor of safety versus the ratio of observed to predicted pile resistance (equivalent to RMP). The factor of safety is defined with respect to the prior estimate of resistance (i.e. the value from an empirical distribution without consideration of load test results). For a single load test (Figure 10), the impact of the test result on the reliability was calculated for the two precision distributions of Figure 9. If a broad

distribution of precision is assumed (Figure 10(a)), corresponding to greater uncertainty in within-site variability, then the required factor of safety is most reduced when the observed result is equivalent to that predicted by the prior method. Observations less than the prior estimate reduce the mean and increase the variability of the predictive distribution, resulting in required factor of safety values that are quite high. Load test observations greater than the prior estimate increase both the mean and the variability of the predictive distribution. The effect of the variability increasing overwhelms increases in the mean, resulting in the increase in required factor of safety for greater than expected load test results. When a narrower distribution of precision is assumed, corresponding to less uncertainty in within-site variability, the results of a load test are more informative. For the values assumed in Figure 10(b), the effect of increasing mean and increasing variability more or less offset one another, resulting in mostly flat required factors of safety beyond the anticipated load test result.

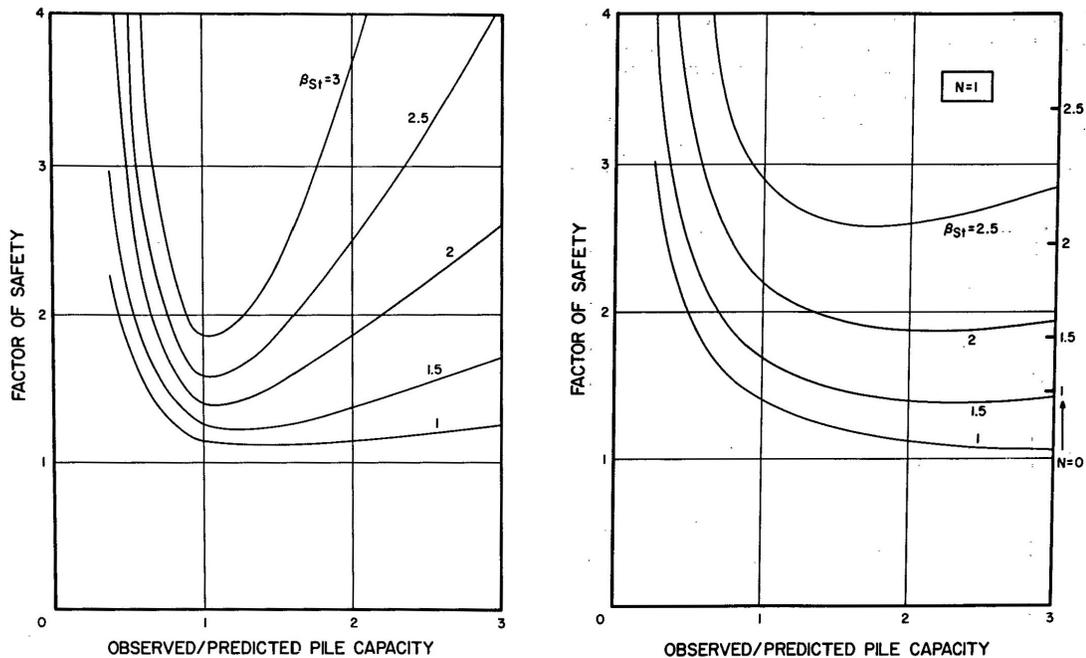


Figure 10: Required factor of safety after updating design from results of one load test: (a) assuming broad distribution of site variability and (b) assuming narrow distribution of site variability. Different lines correspond to different values of reliability index. From Baecher and Rackwitz (1982).

Results from Baecher and Rackwitz considering more than one load test are shown in Figure 11. Both plots of Figure 11 are based on the narrow distribution of precision assumed for

Figure 9(b) and a reliability index, β , of 2. The trends in are similar to those for a single load test, but the required factors of safety are lower compared with the $\beta = 2$ line from Figure 10(b). Baecher and Rackwitz concluded the design factors of safety could be reduced 5 to 10 percent based on the result of one load test, and an addition 5 to 10 percent for a second load test. Unlike the findings for the broad distribution of precision and one load test (Figure 10(a)), the required factors of safety from Figure 11 show no tendency to increase for load test results that are considerably greater than the prior prediction. The results in Figure 11 also demonstrate the effect of load test result variability. For both two tests and three tests, as the variability of the observed load tests increases, the required factor of safety also increases. This is a logical conclusion.

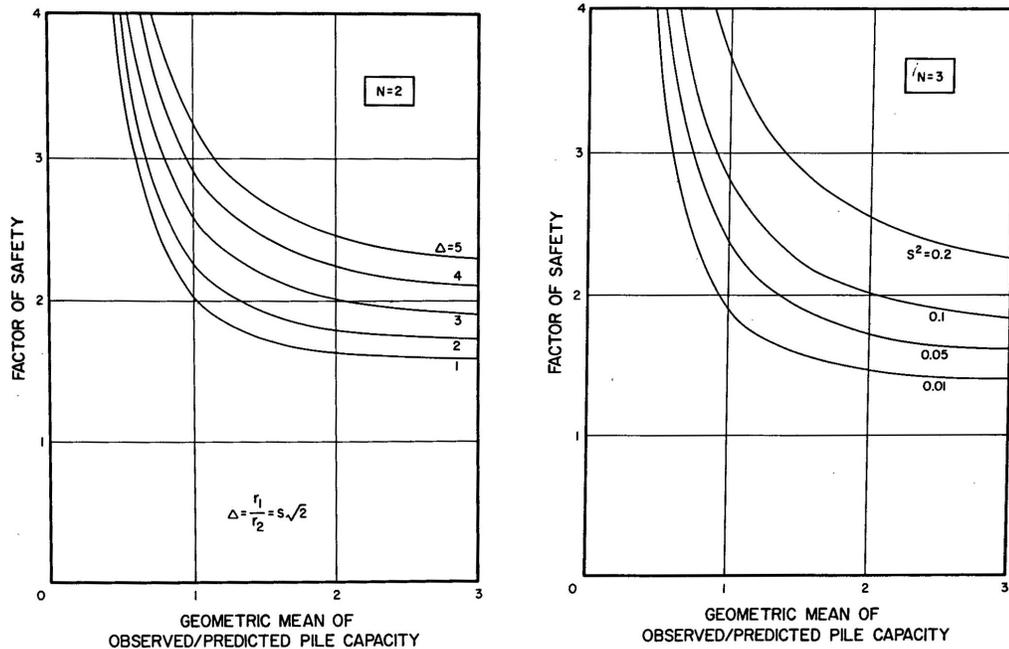


Figure 11: Required factor of safety after updating design from results of load tests assuming narrow distribution of site variability: (a) two tests and (b) three tests. The target reliability index (β) is 2 for all plots. Different lines correspond to different values of variability among load test results. From Baecher and Rackwitz (1982).

2.3.4 Summary of Approaches for Bayesian Updating with Load Tests that Achieve Failure

Two main approaches for considering the results of uncensored load tests were identified from literature. Both approaches follow the general procedure for Bayesian updating outlined in Section 2.2, but they make different assumptions regarding the distributions of foundation

resistance. The simpler of the two approaches was presented by Zhang and Tang (2002), who generalized a specific application to driven piles originally developed by Kay (1976). A more complex approach with fewer assumptions was presented by Baecher and Rackwitz (1982). Both approaches are summarized in Table 1 and in the discussion below.

The relative benefits and limitations of the methods considered for evaluating reliability of designs based on load tests to failure are summarized in Table 1. The method by Zhang and Tang provides a generalized technique for evaluating the reliability of designs based on load tests. The method has been relatively widely applied and cited by other researchers who have investigated the reliability of designs based on load test information. The method by Zhang and Tang is relatively straightforward because it assumes the within-site variability is known and constant, which simplifies the Bayesian computations. The assumption also produces a posterior predictive distribution with variability that is independent of the load test results; the variability of the posterior predictive is a function of the prior variability, the constant value of within-site variability, and the number of tests, but not the test results. Thus a load test program that indicates a highly variable site will produce the same reliability as a load test program with the same average but indicative of a relatively uniform site. The approach by Baecher and Rackwitz treats within-site variability probabilistically, using load test results to update the within-site variability along with the mean resistance. The approach produces posterior distributions that depend not only on the magnitude but also on the variability of load test results, unlike the methods in Kay and Zhang and Tang. The approach by Baecher and Rackwitz is computationally intensive and relies on distributions that may be difficult to implement. In addition, the Baecher and Rackwitz approach requires not just an estimated single value of within-site variability (as in the methods by Kay and Zhang and Tang), but a probabilistic distribution of all possible values of within-site variability. Baecher and Rackwitz's results indicate their approach is sensitive to the assumed distribution of within-site variability, and information regarding within-site variability is limited.

Table 1: Comparison of methods for evaluating reliability of designs based on load tests to failure.

Reference	Overview	Benefits	Limitations
Zhang and Tang, 2002	Use results of load tests to update distribution of the mean parameter of resistance.	<ul style="list-style-type: none"> • Computational simplicity • Relatively widely cited and applied • Approach can be adapted to also consider proof tests. 	<ul style="list-style-type: none"> • Assumes within-site variability is known and constant. • Posterior variability not affected by load test results, except for considering the number of tests.
Baecher and Rackwitz, 1982	Use results of load tests to update distributions of the mean and standard deviation parameters of resistance.	<ul style="list-style-type: none"> • Updates the value of within-site variability. • Considers variability of test results. 	<ul style="list-style-type: none"> • Probabilistic distribution of within-site variability is an important input, and published prior information regarding within-site variability is limited. • Approach has not been used to consider proof test results. • Computational challenges, mostly associated with use of less conventional distributions

2.4 Bayesian Updating of Foundation Resistance for Proof Tests

The previous section focused exclusively on uncensored load tests: tests performed to loads that achieve failure of the test foundation, where failure can be defined as the ultimate or plunging resistance or, more commonly, using some displacement criterion (e.g. Davisson's Offset Limit, 10 percent of the diameter, etc.). Load test programs commonly are based on proof load tests, which are not intended to achieve failure. Instead, proof test foundations are loaded to some target load, often equal to the design load or some factor thereof. If the proof test passes, the test result is censored. If the proof test fails, the test result is uncensored.

Bayesian updating with proof tests is complicated by many of the same topics discussed for Bayesian updating with uncensored tests above: combining resistance information from multiple sources (i.e. a prior prediction and load tests), the influence of within-site variability, and the influence of number of tests. In addition, Bayesian updating with proof tests is complicated by another significant factor: the nature of the test information is different. Instead of measuring a unique value of resistance, a successful proof test measures a lower bound on resistance for one foundation element. The censored test result is less informative than the uncensored test result. Importantly, not all proof tests are successful. Some proof tests may fail, in which case they *do*

achieve failure resistance. Combining the results of failed proof tests with successful proof tests introduces another challenge for Bayesian updating based on proof load test programs.

Three approaches for performing Bayesian updating with proof tests are presented in this section. The first, by Zhang (2004), is similar to the framework by Zhang and Tang (2002) for tests to failure as described in Section 2.3.2, but is applicable for censored tests. Najjar and Gilbert (2009a, 2009b) presented the second approach, which uses truncated distributions to represent the distribution of foundation resistance. The truncated distributions applied by Najjar and Gilbert have the left tail removed to represent a lower bound in foundation resistance; proof test information is used to update the lower bound parameter. The third method, by Najjar et al. (2017), combines aspects of the first two. Details of each of the three methods are presented in separate subsections prior to a summary and comparison of the methods at the end of this section.

2.4.1 Zhang (2004)

Zhang (2004) introduced a method for analyzing the reliability of designs based on proof tests that closely follows the statistical framework he and Tang proposed for tests to failure in 2002. Similar to the previous approach, Zhang (2004) applies Bayesian updating to the mean of the foundation resistance, again invoking the assumption of known and constant within-site variability. Just as for all of the Bayesian methods for load tests achieving failure (Kay, 1976; Zhang and Tang, 2002; Baecher and Rackwitz, 1982), a prior estimate of mean foundation resistance can be derived from any prediction method, with the variability of the mean reflecting among-site variability. The primary difference between Zhang's approach for proof tests and the Zhang and Tang approach for tests to failure is in the definition of the likelihood function. Whereas for tests to failure the likelihood function defined the probability of achieving a specific value of *RMP*, for proof tests the likelihood function is defined with respect to a target maximum test load, x_t . The value of x_t is normalized by the predicted resistance, similar to *RMP*, so that $x_t = 1$ indicates the target load is equal to the predicted resistance. In most conventional proof load test programs, $x_t < 1$; for example, if a foundation is designed with a factor of safety of 2.0,

and the target proof load is equal to the design load, $x_t = 0.5$. Zhang defines three load test program scenarios: (1) all proof tests pass, (2) some proof tests pass, while others fail, and (3) all load tests are conducted to failure. The third scenario is identical to the Zhang and Tang (2002) approach described in Section 2.3.2, so only the first two scenarios are considered here.

For both the first and the second proof test scenarios, the definition of the likelihood function is based on the probability of one test pile surviving the proof test, i.e. the probability of “success.” Zhang assumes that the logarithm of RMP is normal; the same assumption was involved in all methods for load tests to failure. Based on that assumption, the probability of success is

$$L(\mu) = P(x \geq x_t) = \Phi\left(-\frac{\ln(x_t) - \eta}{\xi}\right), \text{ where} \quad \text{Eq. 2-10}$$

x is equivalent to the quantity RMP considered by Zhang and Tang (2002),

η is the mean of the $\ln x$,

ξ is the standard deviation of $\ln x$, and

Φ is used to designate the cumulative distribution function of the standard normal distribution.

For the first scenario, in which all proof tests are successful, the likelihood function for the mean foundation resistance is the joint probability of a series of n foundation elements all having successful tests. Assuming the tests are independent, the joint probability is simply the product of n individual probabilities of success:

$$L(\mu) = \Phi^n\left(-\frac{\ln(x_t) - \eta}{\xi}\right) \quad \text{Eq. 2-11}$$

For the second scenario, in which some of the proof tests are successful and others fail, the likelihood function is the joint probability of some successes and some failures. This is an application of the binomial distribution. The likelihood function for the mean foundation resistance

is binomial, with probability of success $P(x \geq x_t)$ as in Eq. 2-10 and with r successes out of n total tests:

$$L(\mu) = \binom{n}{r} (P(x \geq x_t)|\mu)^r (1 - P(x \geq x_t)|\mu)^{n-r} \quad \text{Eq. 2-12}$$

It is worth noting that this definition of the likelihood function does not include information regarding the specific failure load of the tests that fail; rather, it simply counts how many tests fail and how many pass.

For both scenarios, the likelihood function and prior distribution of mean foundation resistance are combined through Bayes theorem to update the distribution of the mean foundation resistance. The application of Bayes' theorem to update the distribution of the mean resistance is fundamentally the same as the approach in Zhang and Tang (2002), but with proof tests, Bayes theorem does not result in convenient formulae for the updated parameters (i.e. there are no simple formula for μ'' and σ'' as in Eq. 2-8 and Eq. 2-9 from Zhang and Tang (2002)). Instead, Bayes theorem involves the product of lognormal distributions; Zhang states the result is another lognormal distribution, and that the distribution can be solved using a spreadsheet. To evaluate foundation reliability, the updated distribution of mean foundation resistance is used to update the predictive distribution of individual foundation resistance at the site through integration over all possible values of x and μ , again similar to the method for tests achieving failure by Zhang and Tang. The predictive distribution is also lognormal.

Zhang presented example predictive distribution results for both scenarios, as shown in Figure 12 (all proof tests pass) and Figure 13 (some pass, some fail). For both figures, $x_t = 1$, indicating the target proof load is equal to the predicted resistance. The predicted resistance values used to normalize x are not corrected for any bias in the prediction method; hence the prior predictive distribution of x from a conservatively biased method would have a mean value greater than 1, and, on average, an individual test foundation would be expected to survive at $x_t = 1$. In Figure 12, the effect of successful proof tests is shown to shift the predictive distribution of foundation resistance to the right toward greater resistance, reflecting an increase in the mean

resistance with no change in the standard deviation. The shift is significant for the first test, with subsequent successes resulting in additional gains of decreasing significance.

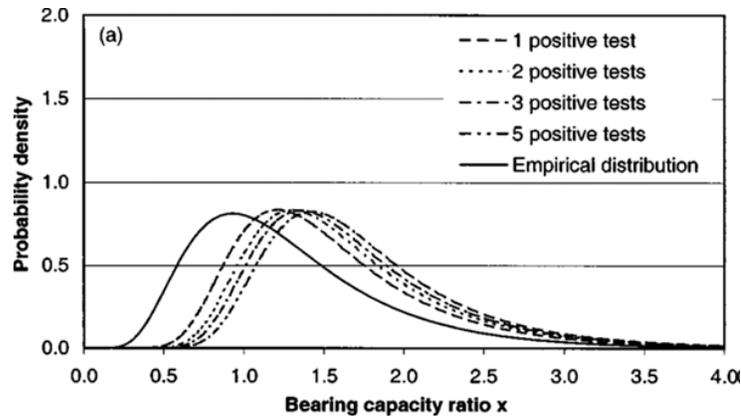


Figure 12: Predictive distributions of foundation resistance: prior (solid line) and after conducting various numbers of successful proof tests (dashed lines) with $x_t = 1$. From Zhang (2004).

The effects of proof tests are more mixed when the proof test program yields both successes and failures, as shown in Figure 13. When all three of three tests fail, the mean *RMP* decreases from 1.3 (the prior prediction is conservatively biased) to 0.72. The updated mean value does not reflect the actual failure loads of the proof tests, merely the likelihood that three tests would fail before reaching a target load equal to the biased predicted failure load. The failed proof tests also significantly decrease the variability of the predictive distribution of individual foundation resistance. Reduced variability generally increases reliability, but comparison of the left tail locations from the distribution updated based on three failed tests with the prior (empirical) distribution indicates the effect of reduced variability is not sufficient to overcome the effect of the reduction in the mean. When the proof test program includes both successes and failures, the effect on mean *RMP* is less significant (1.14 if two tests pass; 0.96 if two tests fail), but the variability is still reduced, resulting in generally favorable predictive distributions. The left tail location for the updated distribution located to the right of that for the prior distribution (i.e. the updated distribution would be associated with greater factored resistance values).

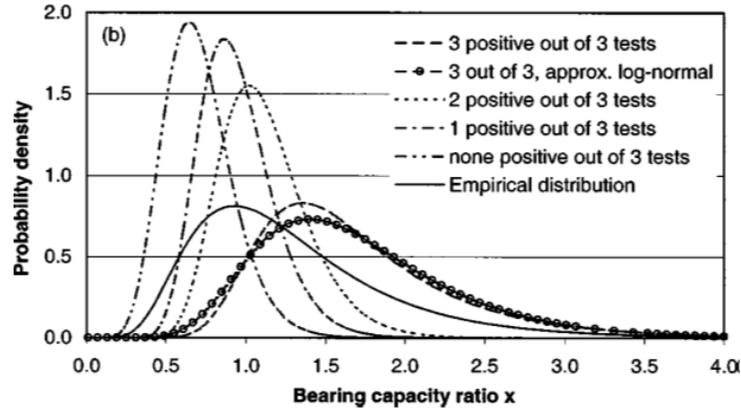


Figure 13: Predictive distributions of foundation resistance: prior (solid line) and after conducting three proof tests (dashed lines) with $x_t = 1$ and varying numbers of successes. From Zhang (2004).

Zhang evaluated the impact of proof load test programs on reliability under his analysis framework by determining the factor of safety required to achieve a reliability index of 2.5 given various target load levels, as shown in Figure 14. A primary conclusion from the graph is that the reliability benefits are particularly sensitive to the proof load level. Without proof test information, a factor of safety just less than 3.5 is required to achieve a reliability index of 2.5. A factor of safety of 3.5 is within the range of values that might typically be applied in practice for foundations designed using empirical prediction methods. Successful proof tests conducted to 1.5 times the predicted resistance reduce the required factor of safety considerably, to about 1.7 after just one test and about 1.25 after five tests. The reduction indicates proof tests to 1.5 times the predicted resistance greatly increase the mean of the predictive distribution of foundation resistance. A less significant effect is noted for proof tests conducted to one half the predicted resistance. For this target load level, the required factor of safety decreases from 3.5 to just 3.0 after five successful tests. The relatively small benefit of tests to $x_t = 0.5$ compared to tests to $x_t = 1.5$ results from smaller increases in the posterior mean *RMP* for the smaller proof load level.

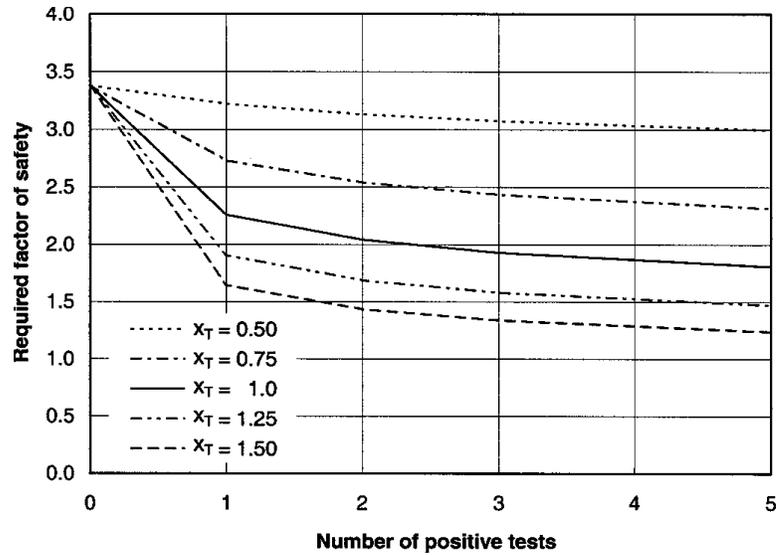


Figure 14: Required factor of safety to achieve a reliability index of 2.5 versus number of successful tests for various target proof load levels. All lines assume the result of all tests in the program is positive. From Zhang (2004).

2.4.2 Najjar and Gilbert (2009a, 2009b)

Najjar and Gilbert (2009a, 2009b) introduced a different approach for considering the effect of proof tests on foundation reliability. Rather than using proof test results to update the mean resistance, Najjar and Gilbert used proof test information to establish a lower bound that removes the left tail of the probability distribution of foundation resistance. The initial motivation for their lower bound concept was more general than the specific application of proof tests: they contended (2009a) that a lower bound was rational (since soil strength cannot be zero) and could help explain why observed rates of failure for foundations are much lower than would be predicted based on target reliability values typically employed in reliability-based design codes. In his dissertation, Najjar (2005) analyzed large databases of load tested driven piles to demonstrate the existence of the lower bound and develop models for predicting its value for various pile types and geologies.

Najjar (2005) also evaluated various probability distribution forms for implementing a lower bound. Truncated distributions modify an existing distribution (e.g. truncated normal, truncated lognormal) by cutting off one tail (or both) and redistributing the tail area among the remaining portion of the probability distribution function. (The tail area must be redistributed to keep the total area under the probability density function equal to one.) Truncated distributions

are commonly applied in other scientific fields, but the redistribution can produce significant changes in the distribution's overall mean and standard deviation. Najjar and Gilbert point out such changes present complications to applying truncation to commonly used distributions of foundation resistance. Instead, Najjar and Gilbert recommend the "mixed lognormal" distribution, which is shown in Figure 15. The concept is similar to the truncated lognormal, but instead of redistributing the tail area, the tail area is simply considered as a discrete probability at the lower bound value. Thus, the mixed lognormal consists of one discrete "probability mass function" value at the lower bound, and then a continuous lognormal distribution above the lower bound. The continuous portion of the mixed lognormal distribution is equal to that for the conventional lognormal distribution. The mixed lognormal distribution is a mix of discrete and continuous distributions, which can present computational challenges. Unlike truncated distributions, the mixed lognormal does not alter the mean or standard deviation values compared to the non-truncated distributions. The mixed lognormal is more conservative than the truncated lognormal by virtue of placing the left tail area right at the point of truncation rather than redistributing the left tail area to greater values of resistance.

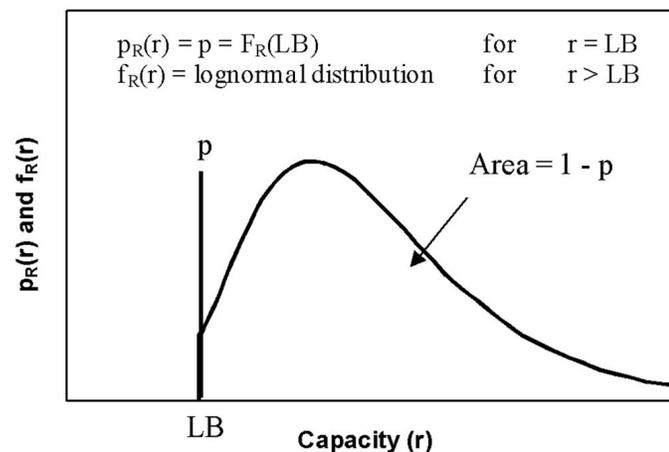


Figure 15: Mixed lognormal distribution used by Najjar and Gilbert (2009a) to characterize the distribution of foundation resistance in the presence of a lower bound.

Najjar and Gilbert showed that implementation of the lower bound via the mixed lognormal distribution can have a dramatic effect on reliability. Figure 16 is a plot of required factor of safety versus the ratio of proof load (lower bound) to median predicted resistance (i.e. the median of the lognormal predictive distribution of foundation resistance, assumed to be

unbiased) for three different target reliability indices. For all three, there is essentially no reduction in required factor of safety below a “threshold” ratio between 0.2 and 0.4. Beyond the threshold, the required factor of safety decreases sharply. Also, the influence of reliability index diminishes as the lower bound increases, likely reflecting major reductions in the probability of failure as the proof load approaches the median resistance. For a reliability index of 2.5, the required factor of safety decreases from 2.8 without proof testing to about 2.2 with a proof test at half the predicted median resistance. This compares with the results from Zhang (2004) in Figure 14, in which the same proof load ratio resulted in a decrease in required factor of safety from 3.4 to 3.0. (The scale difference in factor of safety is at least partially a result of Zhang assuming the foundation resistance $COV = 0.5$, whereas Najjar and Gilbert assumed $COV = 0.4$.)

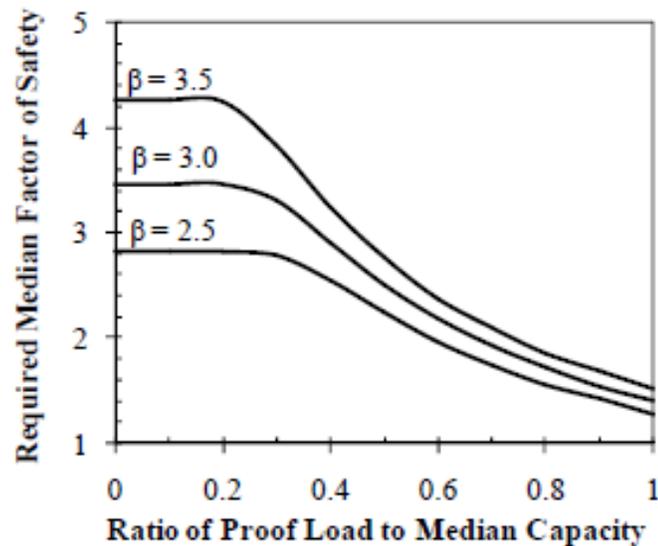


Figure 16: Required design factor of safety as a function of proof load magnitude relative to median resistance for different reliability indices. From Najjar and Gilbert (2009b).

The results in Figure 16 assume a fixed value of the lower bound at the proof load. This neglects any uncertainty in the lower bound, which implies the reliability effects from Figure 16 are only applicable to a specific, individual foundation that has been successfully proof tested. Fixing the lower bound value at the proof load is inappropriate when only some foundation elements at a site are proof tested. To account for cases when only some foundations are tested, Najjar and Gilbert (2009b) introduced a Bayesian framework for updating the lower bound parameter. The Bayesian framework was applied to a truncated lognormal distribution of the pile

resistance, which has three parameters: median, *COV*, and lower bound. (Alternatively, the distribution can be defined using the mean instead of the median and the standard deviation instead of the *COV*.) The median and coefficient of variation were assumed to be deterministic parameters. The lower bound parameter was updated by assuming it follows a prior distribution based on the database analysis by Najjar (2005) described previously. Najjar and Gilbert applied Bayes theorem to update the distribution of the lower bound parameter using the results of proof tests. Numerical integration was used to perform the Bayesian computations, and also to update the predictive distribution of pile resistance. Notably, the approach assumes all proof tests are successful; the methodology cannot consider failing proof tests. In light of this, Najjar and Gilbert recommend a relatively large number of piles be proof tested (to reduce uncertainty in the lower bound) at relatively small proof loads (to reduce the probability of failed proof tests).

Najjar and Gilbert (2009b) present one example of Bayesian updating the lower bound of resistance based on the results of a proof load test program. Results of the example are summarized in Table 2. The baseline case, which does not apply a lower bound, requires a factor of safety of 3.5 be applied to the predicted resistance in order to achieve the target reliability index of 3. If an uncertain lower bound is applied using information from the database analysis of Najjar (2005), the required factor of safety is reduced to 3.2. The distribution of the lower bound based on the database analysis is uniform, with the ratio of lower bound to predicted resistance ranging from 0.2 to 0.6. This distribution is used as the prior for Bayesian updating with proof load test results. The proof load test program consists of testing 3 percent of piles to 0.6 times the predicted resistance. All tests are successful, a required condition to use Najjar and Gilbert's method. After updating, the required factor of safety is reduced to 2.5. If there is no uncertainty in the lower bound, i.e. if all piles are proof tested to 0.6 times the predicted resistance, the required factor of safety is reduced to 2.2 based on the results shown in Figure 16. In summary, these results indicate incorporation of a lower bound is significantly beneficial for reliability, but uncertainties in the lower bound reduce the benefits.

Table 2: Summary of results from Najjar and Gilbert (2009b) for an example test site under various assumptions regarding the lower bound resistance.

Lower Bound Assumption	Required Factor of Safety for $\beta = 3$
No lower bound (conventional case)	3.5
Prior from Najjar (2005): uniform distribution of the ratio of lower bound to predicted resistance from 0.2 to 0.6	3.2
Bayesian updating of prior with 3% of piles tested to 0.6 times the predicted resistance	2.5
Lower bound fixed (zero uncertainty) at 0.6 times the predicted resistance (Figure 16)	2.2

Although the application of Bayesian updating to the lower bound addresses potential uncertainties in the value of the lower bound, Najjar and Gilbert's application of Bayes theorem is not functionally equivalent to the Bayesian approaches adopted by Zhang (2004), Zhang and Tang (2002), or Baecher and Rackwitz (1982). The previous Bayesian approaches all used the variance of the mean foundation resistance parameter to quantify among-site variability and the variance of the distribution of individual foundation resistance at a site, conditional on knowledge of the mean resistance, to quantify within-site variability. (The predictive distribution of individual foundation resistance, not conditional on the mean, includes both within-site and among-site variability.) The approach by Najjar and Gilbert does not distinguish between within-site and among-site variability. As a result, the relationship among the number of proof tests, within-site variability, and uncertainty in the lower bound is unclear. Application of the approach requires treating all sites as equally variable for a given prediction method, or all prediction methods as equally reliable for a given site variability.

2.4.3 Najjar et al. (2017)

Najjar et al. (2017) developed an approach that applies the lower bound concept from Najjar and Gilbert (2009a, 2009b) to the Bayesian framework employed by Zhang (2004) for proof tests. The Bayesian framework by Najjar et al. can be considered with the histograms shown in Figure 17. The histogram of Figure 17(a) is the conditional distribution of individual foundation resistance for certain knowledge of the mean foundation resistance; this is the distribution that would be expected if all foundation elements at a site were loaded to failure (and the average resistance were known). The variability of the histogram in Figure 17(a) represents within-site

variability. Two quantities are characterized by the histograms in Figure 17(b): the mean resistance and lower-bound resistance. The mean resistance histogram reflects the probabilistic distribution of the mean parameter used to define the conditional distribution of individual foundation resistance from Figure 17(a). Variability of the mean resistance distribution reflects among-site variability. The other distribution shown in Figure 17(b) represents the probabilistic distribution of the lower-bound on resistance, which Najjar et al. apply to the mean of foundation resistance, rather than applying the lower bound directly to individual resistance, as per Najjar and Gilbert (2009a, 2009b). To apply the lower bound to the mean resistance, Najjar et al. employ a truncated lognormal distribution for mean resistance. Lognormal distributions are used for the individual foundation resistance and the lower bound resistance. Bayesian updating is used to update the distribution of mean resistance and the distribution of lower bound resistance in consideration of the proof test data.

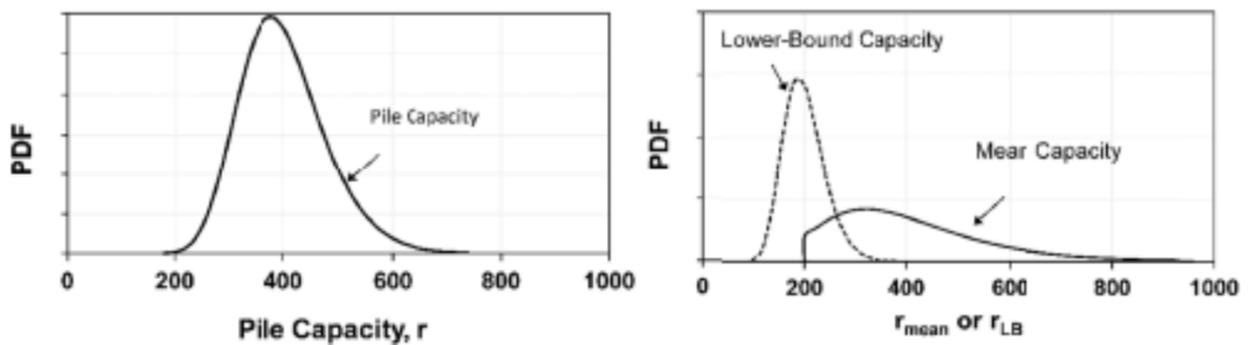


Figure 17: Probabilistic distributions used in Najjar et al. (2017) approach: (a) conditional distribution of individual foundation resistance for a given value of mean resistance and (b) mean foundation resistance and lower-bound resistance.

In order to accommodate the possibility for load test failures as well as successes, Najjar et al. defined the likelihood function using a binomial distribution, similar to the approach described for Zhang's (2004) approach with Eq. 2-12. Implementing the Bayesian framework based on this likelihood function and the general framework described in the discussion of Figure 17 is computationally intensive because of the number of probabilistic inputs. Najjar et al. discretized each of the continuous probability distribution functions and applied numerical integration techniques to apply Bayes theorem, to update the individual foundation resistance distribution, and to calculate the total probability of failure.

The approach by Najjar et al. appears to produce reliability impacts that are more significant than those reported by Zhang (2004). A comparison of results is shown in Figure 18. Both graphs in the figure plot reliability index versus the number of positive proof tests. For both methods, without proof testing, the reliability index is about 1.5 (red points in Figure 18). Both methods predict a similar baseline reliability index because they are based on similar inputs regarding the variability of loads, design factor of safety, and within-site variability; only the prior COV differs, with Najjar et al. assuming a value of 0.4 whereas Zhang assumed a value of 0.5. After 5 successful proof tests to the predicted resistance, the approach by Najjar et al. indicates the reliability index has improved to about 3.6 (blue point in Figure 18(a)). For the same number of proof tests at the same proof load, Zhang indicates the reliability index is less, about 2.7 (blue point in Figure 18(b)). There are several possible explanations for why the approach by Najjar et al. produces greater improvements in reliability. One is that the approach by Najjar et al. applies a lower bound to the distribution of the mean while also updating the mean parameter of the distribution of the mean. Zhang's approach only updates the mean parameter, without consideration of a lower bound. Another possible explanation is the lower prior value of COV assumed in the analysis by Najjar et al.

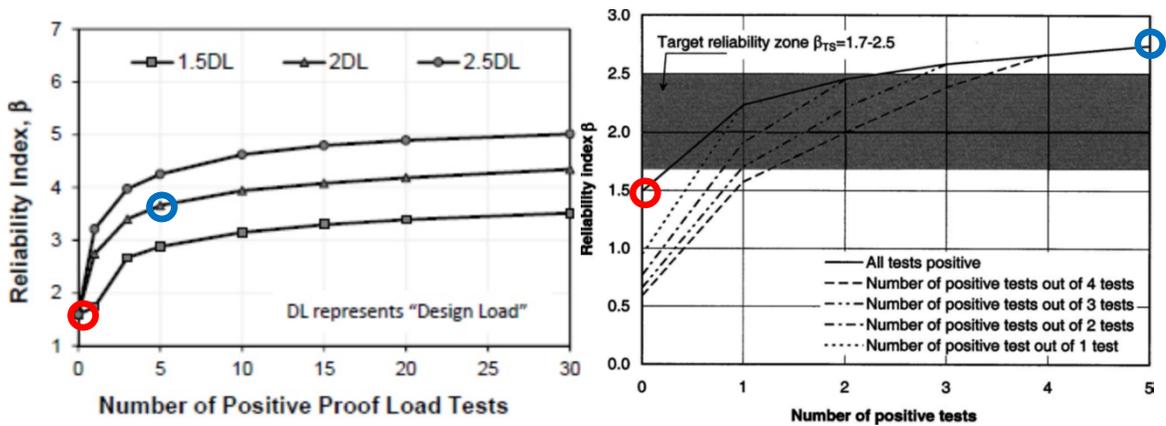


Figure 18: Comparison of reliability impacts from approaches by (a) Najjar et al. (2017) and (b) Zhang (2004). Both approaches are based on similar inputs.

2.4.4 Summary of Approaches for Bayesian Updating with Proof Tests

Three approaches for evaluating reliability of designs verified by load tests not intended to achieve failure have been considered. The approaches are summarized in Table 3. The first

approach, by Zhang (2004), applies a Bayesian framework to update the prior distribution of mean resistance using proof test information. The Bayesian framework of Zhang's approach is consistent with previous methods for updating the mean (e.g. Kay, 1976; Zhang and Tang, 2002).

The second approach, by Najjar and Gilbert (2009a, 2009b), applies truncated probability distributions for the distribution of individual foundation resistance at a site. The truncated distribution removes the left tail of the distribution at some value specified using an additional probability distribution parameter (e.g. the truncated normal distribution has a mean, standard deviation, and a third parameter representing the lower bound). Najjar and Gilbert contended truncated distributions are appropriate for foundation resistance, generally, showing with a large database of driven pile load tests from many sites that there appears to be a lower bound for resistance. Najjar and Gilbert treat proof load tests as a special case of the lower bound application, using Bayesian updating to update the lower bound parameter. Najjar and Gilbert's reliability analyses produced findings of reliability benefits greater than those produced by Zhang (2004). However, Najjar and Gilbert's approach is limited in its treatment of uncertainty, all sources of which are lumped together in one parameter. This makes it difficult to distinguish among effects due to the number of proof tests, within- and among-site variability, and uncertainty in the lower bound parameter. Establishing such effects is likely important since Najjar and Gilbert found that uncertainty in the lower bound parameter has a considerable impact on reliability. Najjar and Gilbert's approach offers no method for evaluating reliability if any of the proof tests fail.

The third approach, by Najjar et al. (2017), applies the lower bound concept from Najjar and Gilbert within the Bayesian framework from Zhang (2004). Comparison of results from the method by Najjar and Gilbert with results from the method by Zhang indicates including a lower bound parameter on the distribution of the mean may produce greater values of calculated reliability, but it is difficult to establish the comparison with certainty since the prior predictive *COV* values in the results were not equal.

Table 3: Comparison of methods for evaluating reliability of designs based on proof load tests.

Reference	Overview	Advantages	Limitations
Zhang, 2004	Use results of proof load tests to update mean value of mean resistance.	<ul style="list-style-type: none"> • Statistical development is rational and consistent with Zhang and Tang's (2002) approach for tests to failure. • Considers successful and unsuccessful proof tests. • Separate inputs for within- and among-site variability 	<ul style="list-style-type: none"> • Does not consider specific failure load information from proof tests that fail. • Within-site variability is assumed to be known and is not updated.
Najjar and Gilbert, 2009	Remove left tail of the predictive distribution of foundation resistance at a lower bound value. Use results of proof load tests to update the lower bound parameter.	<ul style="list-style-type: none"> • Intuitive and relatively simple approach to interpreting proof test results. 	<ul style="list-style-type: none"> • Assumes all proof tests pass; no method for considering failures. • One value of variability is used to account for all sources of uncertainty. • Unclear relationship among site variability, number of tests, and uncertainty in lower bound • Requires prior estimate of lower bound resistance.
Najjar et al., 2017	Lower bound is applied to the distribution of the mean resistance. Bayesian updating is applied to both the mean and lower bound of mean resistance.	<ul style="list-style-type: none"> • Considers successful and unsuccessful proof tests. • Separate inputs for within- and among-site variability • Concept of lower bound intuitively reflects nature of proof testing. 	<ul style="list-style-type: none"> • Does not consider specific failure load information from proof tests that fail. • Application of lower bound to mean, rather than individual foundation resistance may limit effect on reliability. • Requires prior estimate of lower bound resistance.

2.5 Published Information Regarding the Value of Within-site Variability

The quantity of published information regarding specific values of within-site variability is limited. The information that has been published mostly traces back to two publications from the 1970s, one by Evangelista et al. (1977) and the other by Kay (1976), the same work discussed in Section 2.3.1.

The only work focusing exclusively on the topic of within-site variability is by Evangelista et al. (1977). Evangelista et al. examined within-site variability at five different sites in South Italy and Sicily. The subsurface at all five sites was described as sandy, with varying proportions of gravel and/or silt. A total of 94 deep foundations were installed and proof tested among the five sites. Franki piles were installed at three of the sites; bored piles were installed at the other two.

At one of the driven pile sites, piles of three different lengths were installed. At another, piles of two different diameters were installed.

In all, then, eight different sets of deep foundations were proof tested, with the reported results summarized in Table 4. The table includes foundation characteristics, the average of displacement at the service load and at 1.5 times the service load, and the *COV* of loads among the piles. The *COV* of loads among the piles was calculated using the applied load on each pile that produced the average displacement values among all piles at the service load and 1.5 times the service load. As an example, consider the 25-ft driven piles at Site A (the first row of Table 4). Among the 12 piles, the average displacement observed at the service load is 0.07 in. The *COV* of the displacement values at the service load is 0.41. To quantify the *COV* of applied loads, Evangelista et al. determined the load on each pile that produced 0.07 in. of displacement (the average displacement at the service load). Presumably the average of these loads is near the service load value. The *COV* of the loads is 0.23.

Table 4: Within-site variability data reported by Evangelista et al. (1977).

Site	Foundation Information				Average Displacement, \bar{w} , in.		<i>COV</i> of Load	
	Number of Piles	Type	Length, ft	Diameter, ft	At Service Load	At 1.5x Service Load	For \bar{w} at Service Load	For \bar{w} at 1.5x Service Load
A	12	Franki Pile	25	1.2	0.07	0.17	0.23	0.25
	4		37	1.2	0.09	0.22	0.19	0.12
	17		50	1.6	0.13	0.46	0.16	0.19
B	3	Franki Pile	22	1.1	0.08	0.14	0.11	0.12
	16			1.4	0.07	0.11	0.15	0.14
C	4	Franki Pile	47	1.6	0.06	0.11	0.22	0.20
D	16	Bored Pile	45	1.5	0.04	0.09	0.11	0.12
E	22	Bored Pile	62	2.5	0.04	0.08	0.13	0.13

The magnitude of average displacement values in Table 4 is noteworthy. The average displacement values range from 0.04 to 0.13 in. at the service load and 0.08 to 0.46 in. at 1.5 times the service load. Load displacement curves presented in Evangelista et al. (1977) confirm these displacements do not represent ultimate or plunging conditions, and likely do not satisfy most failure criteria (e.g. Davisson Offset Limit), except for perhaps three of the piles at Site A. Accordingly, the load *COV* values in Table 4 are representative of the variability of resistance at

working loads – not at displacement-based failure criterion loads or at ultimate or plunging loads. That the load *COV* values are for working loads, not failure conditions is noteworthy not as a critique of the work by Evangelista et al. but for context, since much of the load test reliability research has applied the load *COV* values in Table 4 to reliability evaluations of the strength limit state (e.g. Kay, 1976; Zhang and Tang, 2002; Baecher and Rackwitz, 1982; Zhang, 2004). None of published literature regarding within-site variability has established how the variability is affected by the relative magnitude of loads (i.e. working loads versus service limit state versus strength limit state).

For the work summarized in Section 2.3.1, Kay (1976) quantified within-site variability using data from load tests of driven piles at three different sites. Results are presented in Table 5. Kay noted “it is now considered appropriate to use for design a variability equivalent to that of the Bremerhaven [Germany] tests,” applying the corresponding standard deviation of the base-10 logarithm of resistance (0.12) to represent within-site variability in his reliability analyses. (Standard deviation of the base-10 logarithm of resistance is approximately equivalent to the *COV* value of 0.26 reported for the German site in Table 5.) References for the works cited by Kay are listed in Table 5. The Komornik, et al. (1971) reference was not located. The work by Chellis (1961) was consulted to identify further details of the test programs for which Kay reported *COV* values. The data are reported from a large table of load test results in an appendix of the textbook by Chellis. The table does not list displacement values for the load tests, so there is no indication of whether the tests reported in Table 5 were near failure at the reported test loads. For the Bremerhaven, Germany site, embedded pile length varied from 23 to 43 ft. The wide range of pile lengths could have resulted in additional variation beyond within-site variability; in other words, the $COV = 0.26$ value in could be biased toward high *COV*. For the San Francisco site, pile length was more consistent, with values between 41 and 45 ft.

Table 5: Within-site variability data reported by Kay (1976).

Location	Number of Piles	Soil Type	Load Test Result			Reference
			Mean, tons	St. Dev., tons	COV	
Ashdod, Israel	12	Sand	154	34	0.22	Komornik, et al. (1971) ¹
Bremerhaven, Germany	9	Sand	61	16	0.26	Chellis (1961)
San Francisco	15	Sand with clay	120	32	0.27	Chellis (1961)

¹Reference source unavailable.

Much of the work presented in Sections 2.3 and 2.4 has applied or built upon the findings reported by Evangelista et al. (1977) and Kay (1976). Zhang and Tang (2002) listed the load *COV* values from the Franki piles in Table 4 and all sites in Table 5 in one table, adopting *COV* = 0.20 to represent within-site variability in their analyses of reliability of designs based on load tests. Zhang (2004) and Najjar et al. (2017) also used *COV* = 0.20 for within-site variability.

Baecher and Rackwitz (1982) also used the results from Kay (1976) and Evangelista et al. (1977) in their consideration of within-site variability. In addition, Baecher and Rackwitz considered within-site variability results from five sites reported by Petermann, et al. (1958). The Petermann et al. reference is in German and was not located. Unlike the other works (e.g. Zhang and Tang, 2002; Najjar et al., 2017, etc.), Baecher and Rackwitz treated within-site variability probabilistically rather than deterministically. Therefore, Baecher and Rackwitz established a prior distribution of within-site variability rather than a single representative value. As discussed in Section 2.3.3 and presented in Figure 9, Baecher and Rackwitz worked in terms of precision h , the reciprocal of variance, in order to facilitate the theoretical solutions of their Bayesian analysis. To facilitate comparisons with the values in Table 4 and Table 5, data from the prior distribution by Baecher and Rackwitz shown in Figure 9 was transformed from precision of the base-10 logarithm of resistance to *COV* of resistance, COV_R :

$$COV_R = \ln(10) h^{-0.5} \quad \text{Eq. 2-13}$$

A histogram of the transformed data is shown in Figure 19. Also shown in the figure is a normal distribution fitted to the data, which has a mean parameter of 0.27 and a standard

deviation parameter of 0.09. These are the mean and standard deviation of the value of within-site COV ; in other words, the average value of COV among the sites considered by Baecher and Rackwitz is 0.27. This is greater than the value of $COV = 0.20$ assumed by Zhang and Tang (2002) and used in subsequent works that cited Zhang and Tang (e.g. Najjar et al., 2017). In Chapter 3, the distribution of within-site variability for a new collection of data is presented and compared to the results from Baecher and Rackwitz.

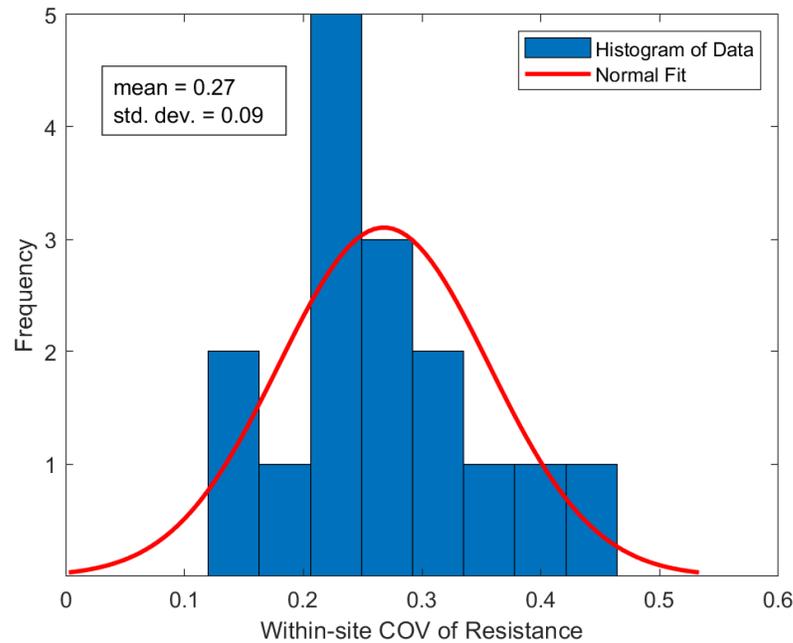


Figure 19: Histogram and fitted normal distribution of the within-site variability data presented by Baecher and Rackwitz (1982).

2.6 Summary

In this chapter, previous work was presented for topics pertinent to the investigation of within-site variability and its effect on Bayesian updating for foundation reliability. The first topic presented was foundation reliability, which provides context for the motivation of Bayesian updating for foundation designs based on load tests. A general introduction to Bayesian updating and Bayesian inference followed the discussion of foundation reliability. The emphasis of the Bayesian inference section was the treatment of probability distribution parameters as probabilistic rather than deterministic inputs. The former approach is unique to Bayesian statistics; the latter is associated with frequentist statistics.

Most of the content in this chapter was presented in Sections 2.3 and 2.4, which address Bayesian updating for designs based on tests to failure and proof test programs, respectively. The methods presented in these sections treat the mean parameter of the distribution of foundation resistance as a probabilistic input that is updated based on the results of load testing. Except for the method by Baecher and Rackwitz (1982), all of the methods presented in Sections 2.3 and 2.4 treat the standard deviation parameter of the distribution of foundation resistance as a deterministic input, an approach that assumes within-site variability is known and will not be updated, irrespective of load test results. Baecher and Rackwitz treat within-site variability probabilistically and report significant effects on reliability depending on both the prior distribution of within-site variability and the within-site variability information observed in test data.

The final topic addressed in this chapter is the value of within-site variability. Previous efforts to quantify within-site variability were summarized, with most work having settled on a value corresponding to $COV = 0.20$. Data supporting the estimate of $COV = 0.20$ originate from about 10 sites. Many of the load tests used to conclude $COV = 0.20$ were terminated at working loads without significant displacement. Information for about one third of the sites originates from dated sources that are no longer available, and supporting information, specifically displacement information, is missing for several other sites. In the next chapter, a new dataset regarding within-site variability is evaluated.

3. Characterization of Within-site Variability for Micropiles

Three of the primary findings from the literature review of Chapter 2 are that

- (1) prior assumptions of within-site variability can have a considerable effect on the reliability interpreted for designs based on load tests (Baecher and Rackwitz, 1982);
- (2) most of the published work regarding reliability of designs based on load tests has assumed a single deterministic value of $COV = 0.2$ to represent within-site variability; and
- (3) the data supporting the value of $COV = 0.2$ is limited in quantity and, arguably, quality, considering many of the load tests used to develop the estimate were loaded only to working loads without significant displacement and displacement information is unavailable for other sites.

Taken together, these findings are strong motivation to collect and evaluate additional data regarding within-site variability. Micropile foundations are an appealing candidate for evaluating within-site variability because micropiles are commonly load tested, and load tests of micropiles achieve failure more often than those for other foundation types. Micropile resistance is relatively small compared to other deep foundation types, making application of failure loads more practical.

This chapter presents an evaluation of within-site variability from five sites with micropile load tests. For most of the sites, a small number of tests to failure were reported. For one site, a large number of proof tests were performed, with a small proportion achieving failure. For another site, a large number of load displacement curves are available. Values of the COV of resistance, COV_R , are fitted to each site, and a distribution of COV_R is then fitted to the collection of sites. The distribution is compared to the prior distribution reported by Baecher and Rackwitz (Figure 9, Figure 19) and to the value assumed in most of the other previous work, $COV_R = 0.2$.

3.1 Description of Datasets

In Chapter 1, within-site variability was defined as the variability of resistance for similar piles at a site where the mean value of resistance is known perfectly. The variability can be characterized using the COV of resistance values measured among similar piles installed at the same site, COV_R . COV_R is the parameter evaluated in the review of previous within-site variability information in Section 2.5. One limitation noted among the COV_R data in Section 2.5 is the displacement values observed among test piles used to calculate COV_R ; the displacement was typically either small or unreported.

To satisfy the definition of COV_R , micropile data used in this analysis are from datasets where all piles for a site have similar diameter, similar bond zone length, are installed in similar geologic conditions, and are loaded in the same direction. These qualities – similar diameter, similar bond zone length, similar geologic conditions, same direction of loading – define “similar” micropiles for this work. To remedy the displacement issue, micropile data used in this analysis are from datasets where significant displacements were observed during load tests, at least for some of the micropiles.

Based on a review of micropile literature, three sites with similar micropiles all tested to failure were identified. In addition, records were obtained for one site with a large proof load test program wherein 17 of 177 test micropiles failed. The fifth dataset was for a site with 536 load tests of similar micropiles. Load-displacement curves were available for each of the 536 micropiles.

3.2 Interpretation of Within-site Variability

As described in Section 3.1, datasets were identified for three types of sites with recorded within-site variability information: four sites with multiple tests performed to failure, one site with a large proof test program with some tests to failure, and one site with many load displacement curves. Interpreted within-site variability results for each type of site are presented in the corresponding sections below.

3.2.1 Sites with Multiple Tests to Failure

Review of micropile literature revealed three sites with similar micropiles, all load tested to failure. For each site, COV_R was estimated using the definition of coefficient of variation:

$$COV_R = \frac{s_r}{\bar{r}} \quad , \text{ where} \quad \text{Eq. 3-1}$$

s_R is the sample standard deviation of resistance, corrected for small sample size, and

\bar{r} is the sample mean of resistance.

It is necessary to correct the sample standard deviation for small sample size to overcome bias in the estimated standard deviation that, on average, leads to underestimating the true standard deviation.

$$s_r = \frac{\sqrt{s_r^2}}{c(n)} \quad , \text{ where} \quad \text{Eq. 3-2}$$

s_r^2 is the unbiased sample variance and

$c(n)$ is the correction factor for small sample sizes, a function of sample size, n .

The unbiased sample variance is calculated from the individual observations of resistance, r_i :

$$s_r^2 = \frac{\sum_{i=1}^n (r_i - \bar{r})^2}{n - 1} \quad \text{Eq. 3-3}$$

The use of $n - 1$ rather than n in the denominator of the unbiased sample variance is known as Bessel's correction, which overcomes the bias between population variance and sample variance (i.e. use of Bessel's correction is what makes the unbiased sample variance unbiased). However, Bessel's correction does not fully eliminate bias in the sample standard deviation (because square root is a concave function). To eliminate the bias, another correction factor is applied (Cureton, 1968):

$$c(n) = \sqrt{\frac{2}{n-1}} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}, \text{ where} \quad \text{Eq. 3-4}$$

Γ refers to the gamma function.

The correction factor defined in Eq. 3-4 is approximately 0.8 for a sample size of $n = 2$ and increases to 1 for large sample sizes.

The methodology captured in Eq. 3-1 through Eq. 3-4 was applied to each of the three sites identified in literature. Results of the within-site variability interpretations are presented in Table 6. The first case was documented by Wolosick (2009). Two micropiles were installed in Yazoo clay, a high plasticity clay with measured undrained shear strength at the site of 3300 psf. The micropiles were post-grouted and load tested in compression, with one micropile failing at 188 kips and the other at 208 kips. Both micropiles were displaced at least 1.5 in. at failure, defined at plunging. COV_R was computed using the bond zone stress at failure as resistance in Eq. 3-1 through Eq. 3-4. The bond zone stress at failure was calculated using the nominal bond zone diameter and length. The failure loads reported in Wolosick are based on the Davisson Offset Limit (1972).

Drbe and El Naggar (2015) documented compression load testing of four hollow bar micropiles in lean clay in London, Ontario. Two of the micropiles had a nominal diameter of 7 in.; the nominal diameter was 9 in. for the other two micropiles. Bond zone stress at failure was calculated from the nominal bond zone diameter and length and used to calculate COV_R . Failure was defined according to the criteria by Fuller and Hoy (1970). At least 1 in. displacement was achieved for all micropiles.

The last case with multiple tests to failure at the same site is Ahsanuzzaman et al. (2019). Ahsanuzzaman et al. installed eight hollow bar micropiles in medium dense sand at a site near the coast of North Carolina in order to study the effect of construction techniques on capacity. The drill bit insertion rate, grout flow rate, and grout water-cement ratio were varied among the eight micropiles; it is likely these variations contributed to the relatively high value of

COV_R , 0.20, listed in Table 6. Indeed, the value of COV_R calculated from Ahsanuzzaman et al. was greater than the values for the other two cases. However, it is not uncommon for installation techniques to vary within production micropile projects. In light of this, and acknowledging the limited availability of within-site variability data, the results from Ahsanuzzaman et al. were maintained for the development of a distribution of within-site variability described in Section 3.3. The micropiles by Ahsanuzzaman et al. were tested in tension. Pullout capacities were reported as the ultimate load for each test, which Ahsanuzzaman et al. report was close to the definition of failure by Fuller and Hoy (1970).

The last column in Table 6 lists the standard error in the estimated values of COV_R . The COV_R standard error was calculated from the standard error in the standard deviation estimate, which was computed by MATLAB. The standard error in COV_R is 0.05 or 0.06 for all three cases. That the value is similar for all cases reflects two offsetting trends: greater errors for greater variability (i.e. larger errors in the standard deviation estimate are expected when the estimate, itself, is relatively high), and lesser errors for greater sample sizes. The standard error in COV_R for the Ahsanuzzaman et al. case is therefore similar to the value from the Wolosick case; the former was based on four times as many points as the latter, but the standard deviation estimate for the former case was significantly higher.

Table 6: Summary of cases with within-site variability calculated from multiple load tests to failure at the same site.

Reference	Location	Number of Micropiles	Micropile Information			Bond Zone Material	Bond Resistance			Std. Error in COV_R
			Type	Bond Length, ft	Diameter, in.		\bar{r} , ksf	s_r , ksf	COV_R	
Wolosick (2009)	Jackson, MS	2	D	25 to 30	8	Yazoo clay	3.46	0.26	0.07	0.059
Drbe and El Naggari (2015)	London, Ontario (Canada)	4	E	19	7 to 9	Lean clay	3.43	0.46	0.13	0.060
Ahsanuzzaman et al. (2019)	Coastal NC	8	E	25	7 to 9	Medium dense sand	2.19	0.43	0.20	0.053

3.2.2 Site with Many Proof Tests, Some to Failure

Micropile load test records were obtained for a project outside Toronto. Type D micropiles with a 6.5-in. diameter by 36-ft long nominal bond zone were installed in glacial till characterized as silty sand and silty clay. A proof test program was performed on 177 micropiles at the site. Seventeen of the micropiles failed at or before reaching the target proof load of 236 kips. A small number of micropiles were proof tested and survived greater loads.

To characterize within-site variability for the project micropiles, a normal distribution was fit to the proof test results. The resulting distribution is shown with the proof test results in Figure 20. The blue bars in the histogram represent results from all micropiles, and the shorter orange bars represent the failed micropiles. Tests that did not fail are censored observations; rather than representing an observation of micropile resistance, the survived proof load represents a lower bound of resistance for the micropile. Censored observations were used with the uncensored observations to compute maximum likelihood estimators of the normal distribution parameters using MATLAB. As shown in Figure 20, the resulting distribution corresponds to estimated $COV_R = 0.21$ with a standard error in the COV_R estimate of 0.04. Note that the distribution fitted to the proof test results defines resistance in terms of load rather than unit bond stress, which was used in Section 3.2.1. Because all of the project micropiles in Figure 20 had the same bond zone geometry, the approaches are equivalent in terms of COV_R .

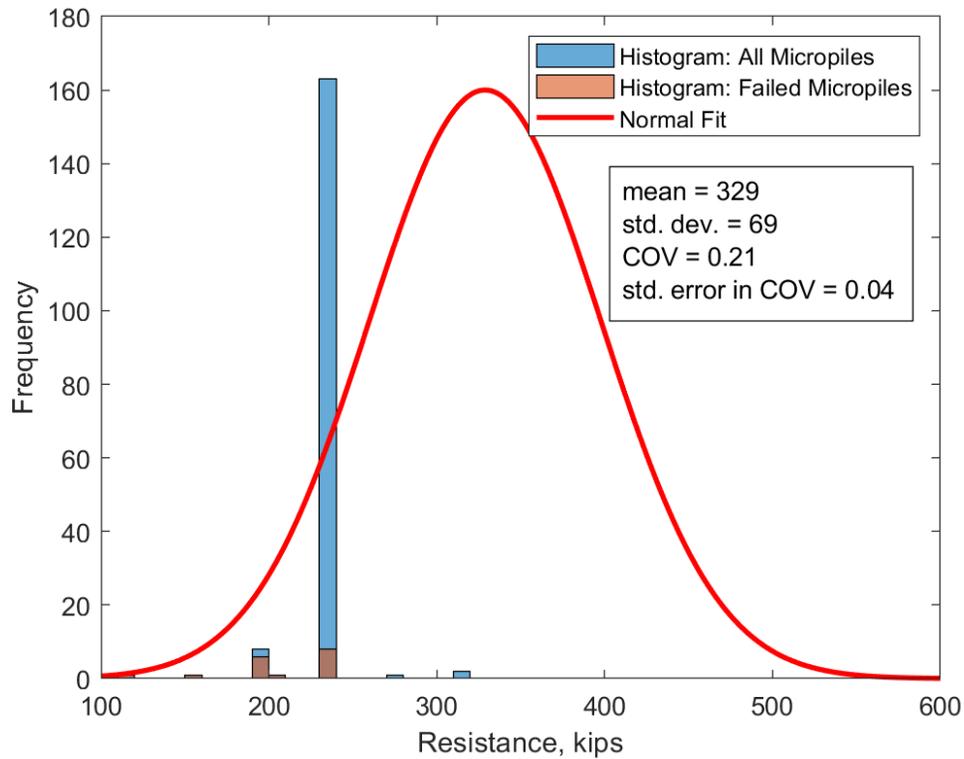


Figure 20: Histogram of resistance data for proof test program with fitted normal distribution.

3.2.3 Site with Many Tests with Load-displacement Curves

Vanderpool et al. (2002) documented a load test program for a site in Las Vegas where 536 micropiles were load tested. The Type B micropiles were installed to a depth of 200 ft with an 8-in. nominal diameter bond zone in caliche, a clayey sand with zones of clay and variable cementation. The 536 micropiles were tested in axial compression to a maximum load of about 600 kips. Load and displacement at the top of the micropile were measured for each test, with the resulting set of load-displacement curves for all 536 micropiles shown in Figure 21. The shape of the load-displacement curves varies considerably: some are nearly linear with final displacement values less than 0.5 in., others show significant curvature with final displacement values exceeding 3.5 in., and most range between these extremes.

Although the observed displacement values shown in Figure 21 are significant, the magnitude of the displacement is qualified by the micropile length, 200 ft. Vanderpool et al. interpreted the load-displacement curves by applying criteria for micropiles from Jeon and

Kulhawy (2001) and Hirany and Kulhawy (2002). Results are shown in Figure 21 as the red line, which marks the beginning of the transition region of the mean load-displacement curve (i.e. the curve interpreted by Vanderpool et al. using the average of all results). The transition region is between the initial linear region and the final linear region of the interpretations by Hirany and Kulhawy. At the maximum load of 600 kips, load-displacement curves for 69 percent of the micropiles indicate more displacement than the red line. The finding that a majority of the micropiles were at least transitioning toward ultimate resistance offers support for the qualitative observation above that many of the load-displacement curves of Figure 21 show considerable curvature.

To evaluate within-site variability from the load-displacement data, load values were interpolated for each of the 536 load-displacement curves at displacement values of 0.5 in., 1 in., 1.5 in., and 2 in. The displacement values are shown with dashed blue lines in Figure 21. For each of the four displacement values, a normal distribution was fitted to the interpolated load values. The fitted distributions are shown with histograms of the interpolated loads in Figure 22. Censored observations at the maximum applied load of 600 kips were included in the development of each distribution to represent micropile load tests that terminated at 600 kips prior to reaching the displacement value for distribution fitting. The number of censored observations increases with increasing displacement. For example, displacement of 0.5 in. was achieved for most micropiles, with only 88 tests terminated at 600 kips prior to achieving 0.5 in. displacement. In contrast, displacement of 2 in. was not achieved for most micropiles; 503 of the 536 tests were terminated prior to reaching 2 in. displacement. As for the site with many proof tests documented in Section 3.2.2, the censored observations were used with the uncensored observations to compute maximum likelihood estimators of the normal distribution parameters using MATLAB.

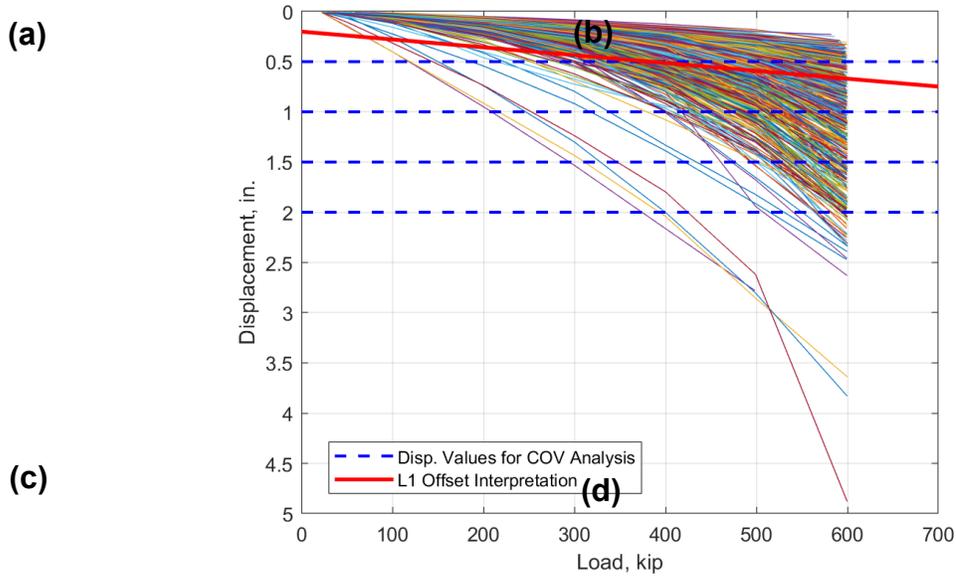


Figure 21: Load-displacement curves for the micropiles documented by Vanderpool et al. (2002).

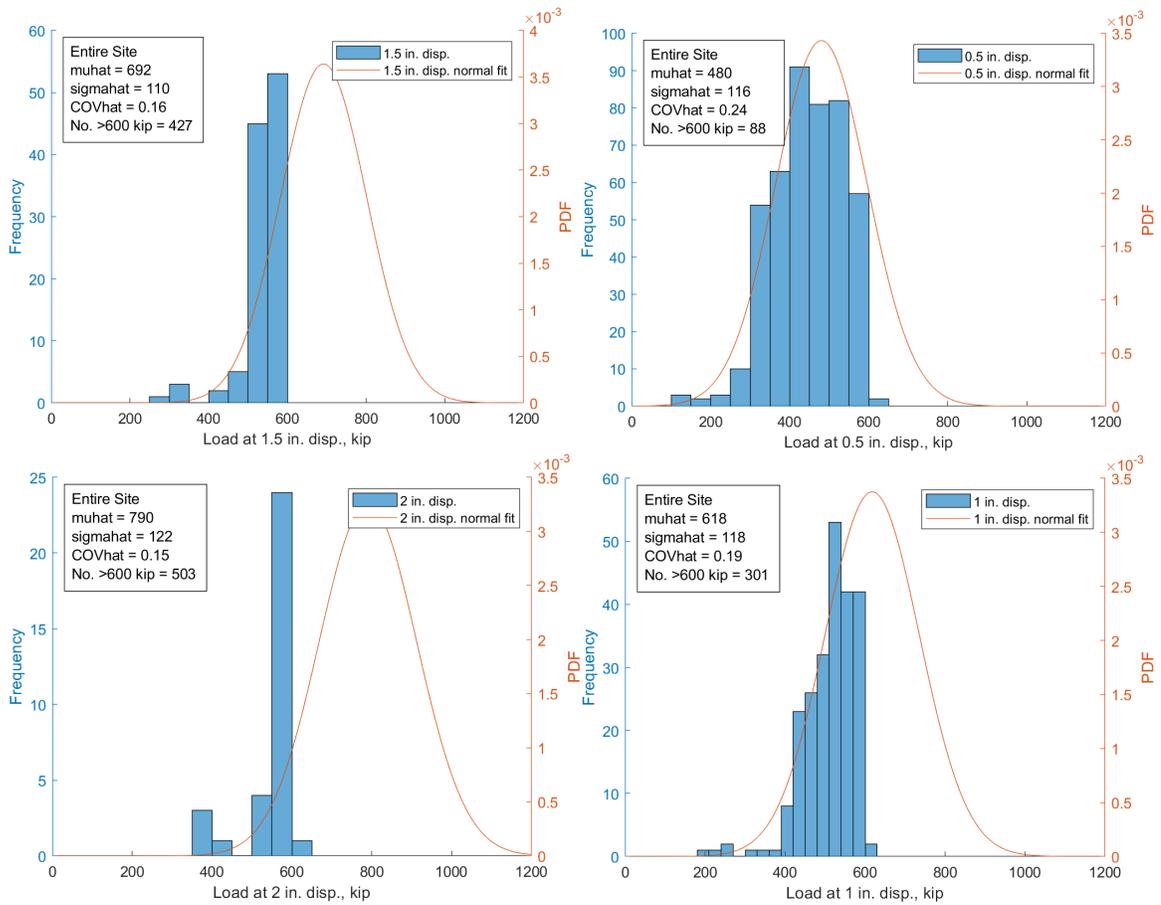


Figure 22: Histogram and fitted distribution of resistance interpreted at different values of displacement at top of micropile: (a) 0.5 in., (b) 1 in., (c) 1.5 in., (d) and 2 in.

Values of COV_R were computed for each of the distributions shown in Figure 22; the values are listed in textboxes (as “COVhat”) in Figure 22. The values are also plotted versus the displacement value in Figure 23. Figure 23 also includes values of COV_R for the north half of the project site, which includes 253 micropiles, and for the south half of the project site, which includes 283 micropiles. The north and south halves of the project site were evaluated separately because Vanderpool et al. noted greater displacements were observed for micropiles on the south side. Vanderpool et al. cited three potential explanations for the greater displacement: (1) “relaxation of the construction technique” as the project proceeded, with north micropiles installed first; (2) “failure to maintain the quality control feedback”; and (3) spatial variability. The potential explanations are consistent with the understanding that within-site variability can result from construction variability or geologic variability across a site. Although the displacements were greater for the south micropiles, there is no clear difference in variability between north and south micropiles. For two of the displacement values, the north side micropiles are less variable; for the other two displacement values, the south side micropiles are less variable.

In fact, the one pronounced effect evident in Figure 23 is a decrease in variability with increasing displacement. The outlying point for the north side at 2 in. displacement is best explained by the small number of micropiles that actually achieved 2 in. displacement on the north side: six of 253, or 2 percent. The resulting fitted distribution is therefore based predominately on censored observations. Neglecting the outlying point, the variability data in show a nearly linear decrease in COV_R with displacement. The observation that variability in micropile resistance is greatest at small displacements suggests that micropile stiffness is more variable than micropile ultimate resistance.

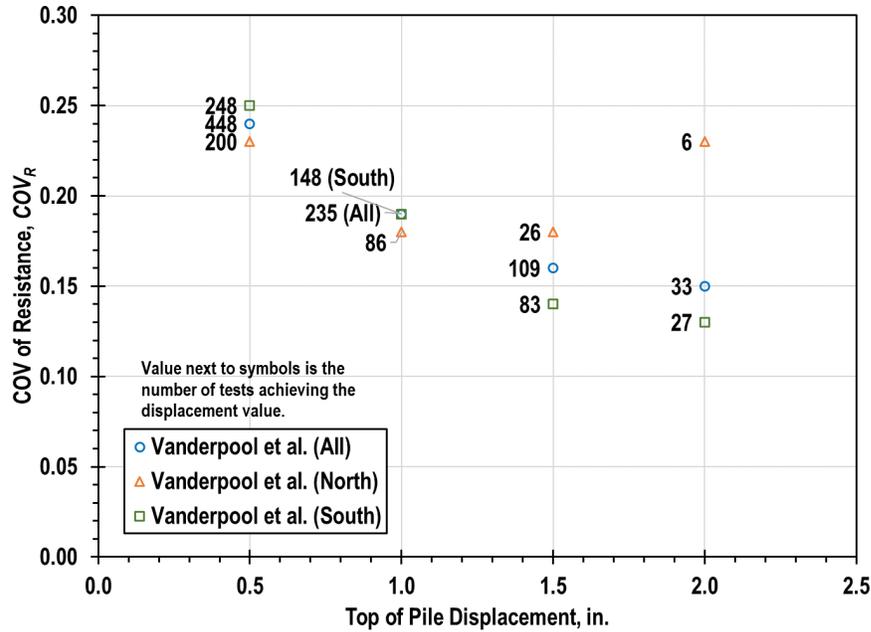


Figure 23: Within-site variability versus top of pile displacement for project documented by Vanderpool et al. (2002).

The number of data points in Figure 23 begs the question: What single value of COV_R is appropriate to characterize within-site variability for the site documented by Vanderpool et al.? For the analysis presented in the next section, a value of $COV_R = 0.16$ was estimated based on the entire site and a displacement value of 1.5 in. The entire site was used because the interpretation of north versus south did not produce any conclusive differences in variability between the two halves. Displacement of 1.5 in. was used because the load-displacement curves in Figure 21 show considerable curvature at 1.5 in., and because 20 percent of the micropiles actually achieved 1.5 in. displacement. Only 6 percent of micropiles achieved 2 in. displacement, so the 2 in. distribution is more dependent on censored observations. The estimated value of $COV_R = 0.16$ is associated with a standard error in COV_R of 0.01.

3.3 Distribution of Within-site Variability Among Sites

Results from the five sites evaluated in Section 3.2 are summarized in Table 7. The table lists project information, micropile characteristics, bond zone material, the estimated coefficient of resistance at the site (COV_R), and the standard error in the COV_R estimate. The COV_R results are shown graphically as the blue normal distribution probability density functions (PDFs) in Figure

24. One normal PDF was assigned to each of the five sites. For each, the mean parameter is equal to the estimated value of COV_R listed in Table 7, and the standard deviation value is equal to the standard error of COV_R listed in Table 7.

Among the individual site distributions of COV_R in Figure 24, the distribution based on analysis of the load-displacement curves from Vanderpool et al. (2002) (Section 3.2.3) stands out as considerably less variable than the other estimates. The relatively small variability associated with the estimate of COV_R associated with the Vanderpool et al. site indicates the estimate of COV_R is less uncertain than the estimates for other sites, which is appropriate considering the quantity of data interpreted for the Vanderpool et al. site. The estimated value of COV_R for the Vanderpool et al. site is also the median of the five sites; that the median of the dataset is also the least uncertain estimate is perhaps unsurprising.

Also shown in Figure 24 are PDFs of two distributions fit to the COV_R estimates of Table 7, one normal distribution and one lognormal distribution. The normal distribution has a mean parameter of 0.15 and a standard deviation of 0.06. The lognormal distribution is characterized by parameters of $\mu = -1.9$ and $\sigma = 0.45$; these are equivalent to the mean and standard deviation of the logarithm of COV_R . The 99 percent confidence interval of COV_R values included in the normal distribution is 0.01 to 0.30, compared to an interval of 0.05 to 0.45 for the lognormal distribution. At both ends of the interval, the normal distribution seems more realistic.

Table 7: Summary of within-site variability information for micropiles at five sites.

Reference or Description	Location	Number of Micropiles (Number Failed)	Micropile Information			Bond Zone Material	COV_R Estimate	Std. Error in COV_R
			Type	Bond Length, ft	Diameter, in.			
Wolosick (2009)	Jackson, MS	2 (2)	D	25 to 30	8	Yazoo clay	0.07	0.059
Drbe and El Naggar (2015)	London, Ontario (Canada)	4 (4)	E	19	7 to 9	Lean clay	0.13	0.060
Ahsanuzzaman et al. (2019)	Coastal NC	8 (8)	E	25	7 to 9	Medium dense sand	0.20	0.053
Proof test program from Section 3.2.2	Toronto, ON	177 (17)	D	36	6.5	Glacial till	0.21	0.042
Vanderpool et al. (2002)	Las Vegas, NV	536 (109)	B	200	8	Caliche	0.16	0.013

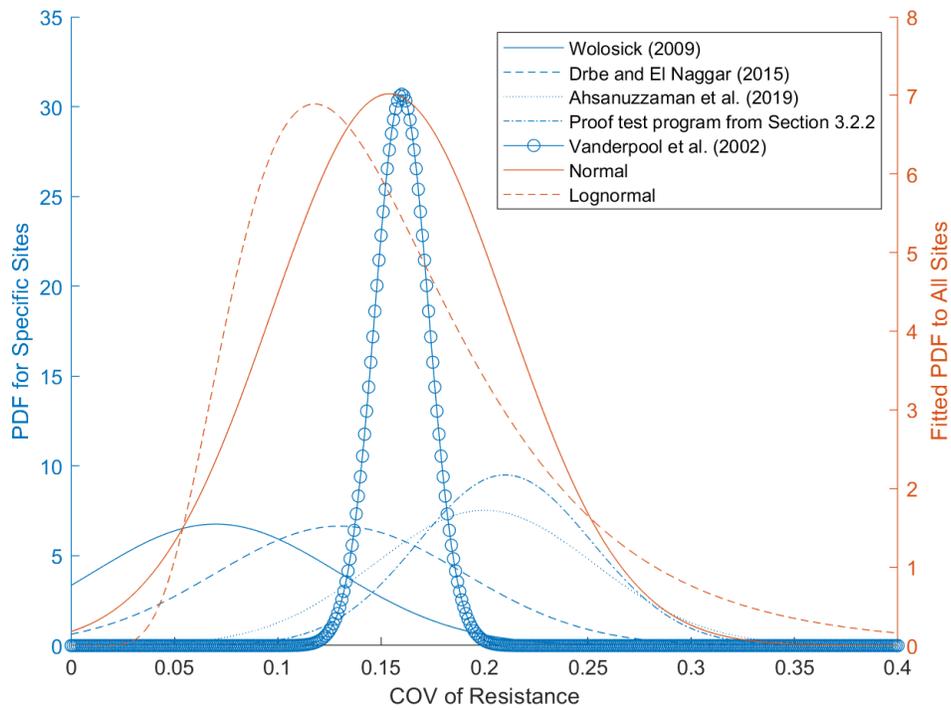


Figure 24: Probability density functions of COV_R , with individual sites in blue and distributions for the collection of sites in red.

3.4 Comparison of Findings with Published Values of Within-site Variability

The normal distribution fitted to COV_R estimates for micropiles in Section 3.4 is shown alongside the distribution of COV_R from Baecher and Rackwitz’s (1982) results (Section 2.5) in

Figure 25. Also shown in the figure are histograms of the data to which the distributions are fitted. The distribution of micropile COV_R estimates has a notably smaller mean value than the distribution from Baecher and Rackwitz: 0.15 versus 0.27. The mean value of 0.15 is also less than the value of $COV_R = 0.2$ assumed by Zhang (2004) and cited subsequently by others (e.g. Najjar et al., 2017). There is also more variability associated with the Baecher and Rackwitz distribution, although both distributions have about the same COV at approximately 0.3 (this is the COV of COV_R).

One possible explanation for the smaller COV_R values associated with the micropile data from this chapter is that the micropile COV_R data were interpreted from load tests that achieved significantly greater displacement than the values interpreted by Baecher and Rackwitz (using data from Evangelista et al., 1977). This explanation is consistent with the finding of decreasing COV_R with increasing displacement discussed in Section 3.2.3. Another possible explanation is that micropile resistance is inherently less variable than resistance of the larger-diameter Franki piles and driven piles considered by Baecher and Rackwitz and Evangelista et al. Micropile resistance is predominately from side resistance; the total resistance of a micropile element can therefore be considered an average of the unit side resistance along the entire length of the micropile. For larger-diameter deep foundations, base resistance contributes more significantly. Total base resistance is generally not averaged over as a large volume of soil as total side resistance. Finally, these two possible explanations are not mutually exclusive; the effect of base resistance variability could be more pronounced at the small displacement values used to define COV_R in the Baecher and Rackwitz dataset.

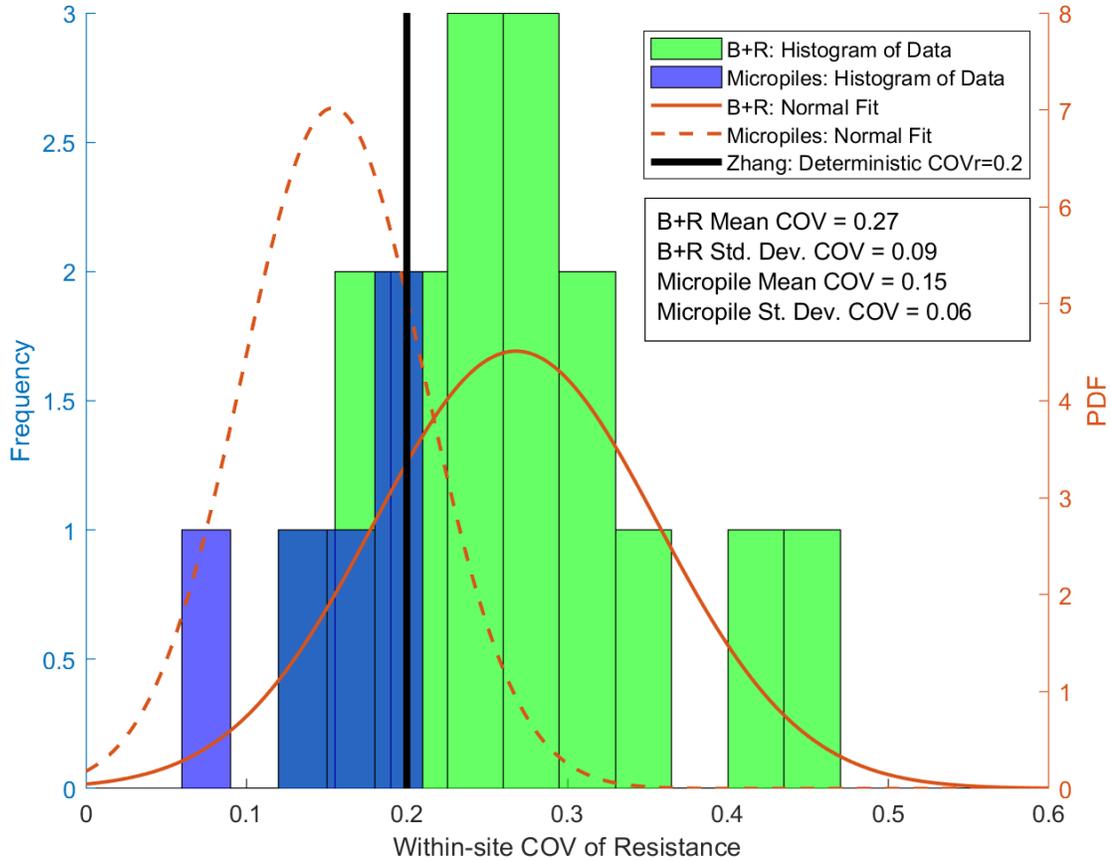


Figure 25: Histograms and fitted normal distributions for data from Baecher and Rackwitz (1982) and for micropile data evaluated in Chapter 3.

3.5 Summary

Data from micropile load tests at five sites were used to estimate the coefficient of variation of resistance (COV_R) at each site. For three of the sites, the estimates were based strictly on load tests that achieved failure, and COV_R was estimated as the ratio of unbiased sample standard deviation to sample mean. For another site, the estimate of COV_R was based on fitting a normal distribution to resistance data based on a proof load test program of 177 micropiles, 17 of which achieved failure. The distribution fitting was based on maximum likelihood estimation with censored observations for the tests that did not achieve failure. For the remaining site, the estimate of COV_R was based on analysis of interpolated values of applied loading at 1.5 in. displacement for 536 load tests documented by Vanderpool et al. (2002). Additional analysis for the site showed a clear trend of smaller values of COV_R for greater micropile displacement, suggesting ultimate resistance is less variable than stiffness.

Results from each site were compiled, and a normal distribution was fitted to the collection of COV_R estimates. The normal distribution has a mean parameter of 0.15 and a standard deviation of 0.06. The estimated COV_R for micropiles represents considerably less within-site variability than previous estimates. For instance, the mean value of COV_R for a distribution presented by Baecher and Rackwitz is 0.27; the deterministic value used by Zhang (2004) and cited by others (e.g. Najjar et al., 2017) is 0.20. A likely explanation for the reduced variability interpreted in this chapter is the use of load test results that achieved failure; the previous estimates were based on load tests that were generally performed to maximum loads representing working or service loads. The explanation that reduced within-site variability is observed for greater displacement is consistent with the trend observed for interpretations of data from Vanderpool et al. (2002). Lastly, and importantly, the investigation of within-site variability documented in this chapter indicates there is significant uncertainty in within-site variability, as evidenced by the standard deviation of 0.06. The standard deviation of 0.06 represents a COV of the value of within-site variability of 0.40. The uncertainty in within-site variability is a significant factor in the differences between deterministic and probabilistic within-site variability, as examined in the remaining chapters.

4. Numerical Tool for Bayesian Updating and Reliability Evaluation

In Chapter 1, the concepts of among-site and within-site variability were introduced, and a general discussion of the importance of the concepts for the reliability of foundation designs based on load tests was presented. In Chapter 2, the discussion of among-site variability, within-site variability, and load tests proceeded in more detail, first with a general overview of why and how Bayesian updating is applied to designs based on load tests (Section 2.2), and then with detailed presentations of previous applications of Bayesian updating for designs based on load tests (Section 2.3). The presentation of previous applications focused on approaches developed by Zhang and Tang (2002), which in turn traced back to a method by Kay (1976), and by Baecher and Rackwitz (1982). The method by Zhang and Tang adopted a deterministic approach to within-site variability, treating it as a constant, known value. The method by Baecher and Rackwitz adopted a probabilistic approach to within-site variability, treating it as unknown value to be updated based on load test information. Both methods relied on theoretical solutions with strict distributional requirements for prior information.

This chapter presents the methodology developed for this research to (1) perform Bayesian updating of foundation designs based on load tests and (2) use the results of Bayesian updating to evaluate foundation reliability. The methodology was implemented through a numerical tool developed for the research to perform both tasks (Bayesian updating and reliability evaluation). The tool can be used to perform Bayesian updating with either a deterministic or a probabilistic treatment of within-site variability. Importantly, the tool uses numerical methods to perform Bayesian updating and reliability evaluation; use of numerical methods eliminates the need to satisfy distributional requirements for the prior information (to perform Bayesian updating) or for the posterior results (to evaluate reliability).

The first section of the chapter provides a relatively concise summary of the methodology used in the tool and its important input and output. In the second section, the numerical method

used in the tool to apply Bayesian updating to designs based on load tests is presented. The presentation follows from the discussion of Chapter 2: it adds detail and equations to the overview from Section 2.2, and it also contains several statements of comparison to distinguish the numerical approach of the tool from the previous approaches described in Section 2.3. The third section of the chapter presents the numerical method for evaluating reliability based on the posterior results from Bayesian updating. The chapter closes with a summary.

4.1 Overview of the Numerical Tool

The numerical tool has two computational objectives: (1) perform Bayesian updating of foundation resistance to reflect load test information and (2) characterize the Bayesian output to provide useful measures for interpretation of the posterior results, particularly for reliability evaluation. A summary of the tool, including important tool input and output, is presented in Figure 26.

The most important inputs for the numerical tool follow from the framework for Bayesian updating described in Section 2.2. The inputs include prior distributions for both of the parameters used to define site-specific resistance, μ_{within} and σ_{within} , and load test information. μ_{within} is the mean parameter for site-specific resistance; σ_{within} is the standard deviation parameter of site-specific resistance. The prior distributions of μ_{within} and σ_{within} are defined numerically, without distributional requirements. Although no specific distribution is required for either parameter, normal distributions were assumed for the sake of example in parts of Section 4.2, which presents additional information about how the two parameters are defined. The other important tool input is load test results. The tool can consider any number of tests. The magnitude of the failure load from each load test is input for the tool.

The first numerical analysis of the tool is used to perform Bayesian updating. The numerical analysis generally conforms to the framework described in Section 2.2. Both deterministic and probabilistic within-site variability can be used in the tool. In the former, used in most of the previous work related to Bayesian updating for foundation load tests (e.g. Kay, 1976; Zhang and Tang, 2002; Najjar et al., 2017), the within-site variability parameter is treated as a

known, constant value and is not updated to reflect load test results. In the latter, previously used only in Baecher and Rackwitz (1982), the within-site variability is treated as an unknown value to be updated based on load test information. Regardless of the within-site variability approach (i.e. deterministic or probabilistic), the previous work presented in Sections 2.3 and 2.4 all used theoretical solutions to perform Bayesian updating. In contrast, the numerical tool uses numerical solutions that facilitate use of prior distributions without strict distributional requirements.

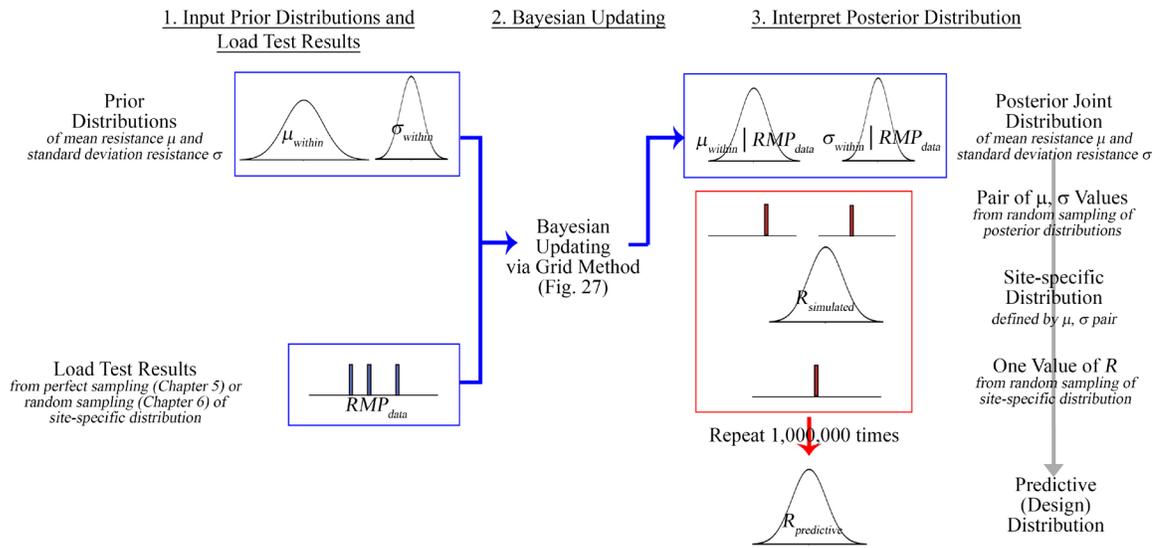


Figure 26: Overview of the numerical tool for Bayesian analysis and reliability analysis of designs based on load test information.

The primary output from the Bayesian analysis is the posterior joint distributions of μ_{within} and σ_{within} , the parameters used to define site-specific resistance. The output from the tool includes a numerical posterior joint probability density functions of μ_{within} and σ_{within} , as well as posterior marginal distributions of μ_{within} and σ_{within} . In addition, the tool fits normal distribution parameters to the posterior marginal distributions of both μ_{within} and σ_{within} , yielding four parameters: μ_{μ} and σ_{μ} for the mean and standard deviation of μ_{within} , respectively, and μ_{σ} and σ_{σ} for the mean and standard deviation of σ_{within} , respectively. In Chapters 5 and 6, values of the four parameters are used to evaluate the results of Bayesian updating.

The posterior joint distribution of μ_{within} and σ_{within} serves as input for the second numerical analysis of the tool, which is a numerical analysis of reliability. The posterior joint

distribution of μ_{within} and σ_{within} is used to develop a posterior predictive distribution of resistance. The posterior predictive distribution of resistance represents the expected magnitude and variability of resistance in light of both empirical methods and load test information. The tool computes the 1/1500th value of resistance from the posterior predictive distribution, $R_{1/1500}$. $R_{1/1500}$ describes the left tail of the posterior predictive distribution; it is the design value of resistance with a probability of failure of 1/1500 if the loads are perfectly known and perfectly balanced with the design value. In Chapters 5 and 6, the value of $R_{1/1500}$ is used to evaluate reliability implications of the Bayesian analyses with deterministic and probabilistic treatment of within-site variability.

4.2 Numerical Method for Bayesian Updating of Foundation Resistance based on Load Test Information

As outlined in Section 2.2, resistance information from foundation load tests is typically considered in conjunction with results from empirical design methods that are based on results from a collection of other sites. Bayesian updating is a rational method for considering both the results of foundation load tests and the predictions from empirical design methods. The methodology for Bayesian updating used in the numerical tool is presented in this section, which details the notation, Bayesian framework, and computational methods applied in the tool.

4.2.1 General Framework and Notation

The random variable used to characterize foundation resistance is ***RMP***, the ratio of measured foundation resistance, R_m , to predicted foundation resistance, R_p :

$$\mathbf{RMP} = \frac{R_m}{R_p} \quad \text{Eq. 4-1}$$

Bold text is used throughout the equations in this chapter to denote random variables; normal, non-bold text is used for deterministic constants or parameters. The measured foundation resistance, R_m , is a random variable, while the predicted resistance, R_p , is treated as deterministic. (In other applications, e.g. evaluation of reliability of empirical designs, R_p should be

treated as a random variable, but for the purposes of evaluating designs based on load tests, the value of R_p is fixed.) Because R_p is deterministic, the mean and standard deviation of **RMP** (μ_{RMP} and σ_{RMP} , respectively) are linear transformations of the mean and standard deviation of R_m (μ_{R_m} and σ_{R_m} , respectively):

$$\mu_{RMP} = \frac{1}{R_p} \mu_{R_m} \quad \text{Eq. 4-2}$$

$$\sigma_{RMP} = \frac{1}{R_p} \sigma_{R_m} \quad \text{Eq. 4-3}$$

It follows from the definition of *COV* that the *COV* of **RMP**, COV_{RMP} , is equal to the *COV* of R_m , COV_{R_m} :

$$COV_{RMP} = \frac{\frac{1}{R_p} \sigma_{R_m}}{\frac{1}{R_p} \mu_{R_m}} = COV_{R_m} \quad \text{Eq. 4-4}$$

COV_{R_m} is equivalent to the quantity used to evaluate within-site variability in Chapter 3, COV_R . By Eq. 4-4, then, the results of Chapter 3 can be used to estimate COV_{RMP} .

RMP is defined equivalently to the bias factor commonly used in evaluation of design methods. **RMP** can be considered a normalized version of foundation resistance. Use of **RMP** is mathematically convenient compared to working in terms of un-normalized foundation resistance because distributions of **RMP** are readily compared regardless of the magnitude of resistance. The distribution of **RMP** is an indication of both the accuracy and precision of predictions: more accurate prediction sets have a mean ratio closer to 1, and more precise prediction sets have smaller variability. Use of **RMP** is also advantageous since bias statistics are available for many foundation design methods (e.g. Paikowsky et al., 2004).

The natural logarithm of the ratio, $\ln(\mathbf{RMP})$, is also used for convenience. The distribution of $\ln(\mathbf{RMP})$ is assumed to be normal, which is equivalent to assuming the distribution of RMP is lognormal. A lognormal distribution of foundation resistance has been used commonly (e.g.

Barker et al., 1991; Paikowsky et al., 2004). One advantage of the lognormal distribution is that its support is strictly positive, an appealing characteristic for modeling strictly positive quantities such as foundation resistance.

Another characteristic of the lognormal is convenient for modeling purposes. As stated previously, if a random variable is lognormally distributed, then the natural logarithm of the variable is normally distributed. Further, the standard deviation of the natural logarithm of the variable is approximately equivalent to the *COV* of the variable, itself (Ang and Tang, 2007):

$$\ln(\mathbf{RMP}) \sim N(\mu_{\ln(\mathbf{RMP})}, \sigma_{\ln(\mathbf{RMP})}) \quad , \quad \text{where} \quad \text{Eq. 4-5}$$

\mathbf{RMP} is defined as above and is lognormally distributed

$\mu_{\ln(\mathbf{RMP})}$ is the mean of $\ln(\mathbf{RMP})$, and

$\sigma_{\ln(\mathbf{RMP})}$ is the standard deviation of $\ln(\mathbf{RMP})$

$$\sigma_{\ln(\mathbf{RMP})} = \sqrt{\ln(1 + COV_{RMP}^2)} \approx COV_{RMP} \quad \text{Eq. 4-6}$$

The approximation of Eq. 4-6 holds for $COV \leq 0.3$ (Ang and Tang, 2007). As established in Eq. 4-4, $COV_{RMP} = COV_{R_m}$. Accordingly, the normal distribution of Eq. 4-5 can be restated:

$$\ln(\mathbf{RMP}) \sim N(\mu_{\ln(\mathbf{RMP})}, COV_{R_m}), \quad COV_{R_m} \leq 0.3 \quad \text{Eq. 4-7}$$

Combining Eq. 4-4 and Eq. 4-8 produces the useful result captured in Eq. 4-8. The result means that distributions of the *COV* of resistance, including the one developed in Chapter 3, can be used as the standard deviation of $\ln(\mathbf{RMP})$, provided the *COV* of resistance is less than 0.3. This result is useful because the numerical tool is defined in terms of $\ln(\mathbf{RMP})$.

$$COV_R = COV_{RMP} \approx \sigma_{\ln(\mathbf{RMP})} \quad \text{Eq. 4-8}$$

In the Bayesian framework presented in Section 4.2.2, various forms of Eq. 4-5 and Eq. 4-7 are used to represent different distributions of $\ln(\mathbf{RMP})$ for the Bayesian prior, likelihood, and

posterior. Although the “ $\ln(RMP)$ ” subscript is dropped in the equations of Section 4.2.2 to simplify notation, the Bayesian framework is developed in terms of $\ln(RMP)$.

Finally, note that the random variable $\ln(RMP)$ is equivalent to the variable x considered by Zhang and Tang (2002) (Section 2.3.2) and by Zhang (2004) (Section 2.4.1). It is also similar to the variables considered by Kay (1976) (Section 2.3.1) and Baecher and Rackwitz (1982) (Section 2.3.3). Kay and Baecher and Rackwitz worked in terms of the base-10 logarithm of RMP , which is linearly related to the natural logarithm:

$$\ln RMP = \ln 10 \cdot \log_{10} RMP \quad \text{Eq. 4-9}$$

4.2.2 Bayesian Framework

In this section, the notation developed in Section 4.2.1 is applied to the Bayesian methodology outlined in Sections 2.1 and 2.2. To simplify notation, the “ $\ln(RMP)$ ” subscript is dropped in the equations of this section, but the framework is intended for application to $\ln(RMP)$. For example, μ_{within} is used to represent the mean value of $\ln(RMP)$ at a specific site.

For a specific site, the distribution of $\ln(RMP)$ can be defined according to Eq. 4-10 for the case in which the mean $\ln(RMP)$ for the site, μ_{within} , and the within-site variability, σ_{within} , are known. The case of known μ_{within} and σ_{within} is represented by the conditional statement on the left side of Eq. 4-10. If μ_{within} and σ_{within} are known, the only uncertainty in design predictions is due to the within-site variability. Such a case is not realistic, but the conditions are useful for developing the Bayesian framework in this section.

$$(\ln(\mathbf{RMP})|\mu_{within}, \sigma_{within}) \sim N(\mu_{within}, \sigma_{within}) \text{ , where} \quad \text{Eq. 4-10}$$

μ_{within} is the mean value of $\ln(\mathbf{RMP})$ within the site,

$(\ln(\mathbf{RMP})|\mu_{within}, \sigma_{within})$ is the random variable $\ln(\mathbf{RMP})$ conditioned on knowledge of μ_{within} and σ_{within} , and

σ_{within} is the standard deviation of $\ln(\mathbf{RMP})$ within the site, i.e. within-site variability

Consider the case of a design based on an empirical method developed from experience at other sites and without any site-specific load tests. This is the typically the prior case for Bayesian updating to reflect site-specific load tests. For the prior case, the mean value assigned to $\ln(\mathbf{RMP})$ is μ_{among} , the value from an empirical method. As explained in Section 2.1, distributional parameters (e.g. mean and standard deviation for the normal distribution) in the Bayesian framework are treated probabilistically, i.e. as random variables. In contrast, the frequentist approach treats distributional parameters deterministically, i.e. as constant values. Accordingly, the assumption $\mu_{within} = \mu_{among}$ associated with the prior method can be considered in Bayesian terms by assuming a normal distribution for μ_{within} , and using μ_{among} as the mean for the distribution, as shown in Eq. 4-11:

$$\mu_{within} \sim N(\mu_{among}, \sigma_{among}) \text{ , where} \quad \text{Eq. 4-11}$$

μ_{among} is the mean value of $\ln(\mathbf{RMP})$ among the sites used to develop an empirical design method

σ_{among} is the standard deviation of the prior estimate of μ_{among} for an empirical design method, i.e. among-site variability

Both μ_{among} and σ_{among} are treated as known, deterministic parameters; within Eq. 4-11, only μ_{within} is to be updated. The among-site variability represented by σ_{among} reflects variation in the mean results for each site used to develop the empirical design method. Importantly, the conditional distribution of Eq. 4-10 includes only the uncertainty in $\ln(\mathbf{RMP})$ due to within-site variability, and the distribution of μ_{within} in Eq. 4-11 includes only the uncertainty due to among-

site variability. For new predictions of $\ln(\mathbf{RMP})$ based on empirical design methods, a distribution of $\ln(\mathbf{RMP})$ that includes the effects of both within- and among-site variability is necessary to evaluate the total variability and uncertainty that applies to new predictions of resistance, σ_{total} . For deterministic within-site variability, in which within-site variability is assumed known and constant (e.g. Zhang and Tang, 2002), the distribution of $\ln(\mathbf{RMP})$ for new predictions is shown in Eq. 4-12 and Eq. 4-13. In Bayesian parlance, this distribution is known as the prior predictive distribution. Generalized derivations of the prior predictive shown in Eq. 4-12 and Eq. 4-13 are relatively widely available (e.g. Taboga, 2017).

$$(\ln(\mathbf{RMP})|\sigma_{within}) \sim N(\mu_{among}, \sigma_{total}) \quad \text{Eq. 4-12}$$

$$\sigma_{total} = \sqrt{\sigma_{among}^2 + \sigma_{within}^2}, \sigma_{within} \text{ known} \quad \text{Eq. 4-13}$$

The prior predictive of Eq. 4-12 is conditioned on knowledge of σ_{within} . The normal distribution of Eq. 4-12 is only appropriate when the within-site variability is assumed to be known and constant. When σ_{within} is probabilistic, the prior predictive of $\ln(\mathbf{RMP})$ (without knowledge of σ_{within}) follows a Student's T distribution, as presented by Baecher and Rackwitz (1982). The Student's T distribution is wider than the normal distribution, which reflects the additional uncertainty introduced by unknown σ_{within} .

Bayes' theorem is used to update the distributions of μ_{within} and σ_{within} in light of site-specific load test information. Updated distributions of μ_{within} and σ_{within} can be used to evaluate reliability for a specific foundation site. Bayes' theorem states that in light of new information, prior distributions of parameters can be updated to determine posterior distributions of the parameters. A form of Bayes' theorem is stated mathematically in Eq. 4-14:

$$p(\theta|D) \propto p(D|\theta) \cdot p(\theta) \quad , \text{ where} \quad \text{Eq. 4-14}$$

$p(\theta|D)$ is the probability of the parameter(s) θ given the data D , i.e. the posterior distribution of the parameter(s),

$p(D|\theta)$ is the probability of observing the data given prior estimates of the parameter(s), also referred to as the likelihood function, and

$p(\theta)$ is the probability of the parameter(s), i.e. the prior distribution of the parameters(s)

The proportional operator is used in Eq. 4-14 because the right side of the equation must be normalized by a constant (equal to $p(D)$) in order to produce a probability distribution function (i.e. to satisfy $\int_{-\infty}^{\infty} f(x)dx = 1$).

Bayes' theorem is restated in Eq. 4-15, with "load test" replacing data (D in Eq. 4-14) to represent the results of the load test. The parameters to be updated are the site-specific mean and standard deviation of the foundation resistance, μ_{within} and σ_{within} , respectively (i.e. the parameters in Eq. 4-10). Note that most published approaches for Bayesian updating based on load test information use test results to update only the distribution of μ_{within} , while σ_{within} is treated as a deterministic parameter. For deterministic within-site variability, the σ_{within} terms drop out of Eq. 4-15. The approach by Baecher and Rackwitz (1982) is a notable exception.

$$p(\mu_{within}, \sigma_{within} | \text{load test}) \propto p(\text{load test} | \mu_{within}, \sigma_{within}) \cdot p(\mu_{within}, \sigma_{within}) \quad \text{Eq. 4-15}$$

The assumption presented previously that $\mu_{within} = \mu_{among}$ is used to define the prior distribution of μ_{within} . Based on Eq. 4-11, the prior of μ_{within} can be calculated from the probability density function of the normal distribution, denoted by f_{normal} in Eq. 4-16. The term $\ln(RMP)$ in Eq. 4-16 is used to reflect the potential values that can be taken by μ_{within} (in terms of $\ln(RMP)$).

$$p(\boldsymbol{\mu}_{within}) = f_{normal}(\ln(RMP), \mu_{among}, \sigma_{among}) \quad \text{Eq. 4-16}$$

For cases where σ_{within} is to be updated, a prior distribution of σ_{within} is also required. The distribution of σ_{within} is discussed in Section 2.5 and throughout Chapter 3, in which a new distribution of σ_{within} for micropiles was presented. For the numerical tool, no distributional constraints were placed on σ_{within} . However, the distribution of σ_{within} in Eq. 4-17 is shown as normal for the sake of example and clarity, and because the distribution of σ_{within} developed in Chapter 3 was normal.

$$p(\sigma_{within}) = f_{normal}(\sigma, \mu_{COV_R}, \sigma_{COV_R}) \quad , \text{ where} \quad \text{Eq. 4-17}$$

σ are the potential values of σ_{within} to be evaluated in the probability density function,
 μ_{COV_R} is the mean parameter of the distribution of σ_{within} , and
 σ_{COV_R} is the standard deviation parameter of the distribution of σ_{within} .

For the distribution of within-site variability developed in Chapter 3, $\mu_{COV_R} = 0.15$ and $\sigma_{COV_R} = 0.06$. These parameters were developed with respect to the COV of resistance, COV_R , but by Eq. 4-8, they also apply to COV_{RMP} and $\sigma_{\ln(RMP)}$.

The prior distributions of μ_{within} and σ_{within} are assumed to be independent, so their joint probability is the product of their marginal probabilities:

$$p(\boldsymbol{\theta}) = p(\boldsymbol{\mu}_{within}, \sigma_{within}) = p(\boldsymbol{\mu}_{within}) \cdot p(\sigma_{within}) \quad \text{Eq. 4-18}$$

The likelihood function to be used in the Bayesian updating depends on the nature of the load test information. Two cases are considered: (1) load tests that achieve failure and (2) load tests that do not achieve failure.

If failure is achieved during the load test, the likelihood is defined by applying the normal distribution, as presented in Eq. 4-19:

$$p(D|\theta) = p(\ln(\mathbf{RMP}) = \ln(\mathbf{RMP}_{fail}) | \boldsymbol{\mu}_{within}, \boldsymbol{\sigma}_{within}) = \dots$$

Eq. 4-19

$$= f_{normal}(\ln(\mathbf{RMP}_{fail}), \boldsymbol{\mu}_{within}, \boldsymbol{\sigma}_{within}) , \text{ where}$$

$$\mathbf{RMP}_{fail} = \frac{\text{failure load observed during load test}}{R_p}, \text{ i.e. the normalized failure load, and}$$

f_{normal} denotes the probability density function of the normal distribution

Because the load test is an observation at a particular site, it is appropriate for the variability used to evaluate the likelihood to be the within-site variability, $\boldsymbol{\sigma}_{within}$. Use of $\boldsymbol{\sigma}_{within}$ is also consistent with Eq. 4-10 since the likelihood is a function of $(\ln(\mathbf{RMP}) | \boldsymbol{\mu}_{within}, \boldsymbol{\sigma}_{within})$.

The definition of the likelihood function in Eq. 4-19 is equivalent to that adopted by Baecher and Rackwitz (1982) (Section 2.3.3). However, the approach used in this work to update foundation resistance based on the results of load tests to failure was based on a numerical approach, as described below. The approach by Baecher and Rackwitz was theoretical, with the joint prior of $\boldsymbol{\mu}_{within}$ and $(\boldsymbol{\sigma}_{within})^{-2}$ (i.e. within-site precision, the inverse of within-site variance) following a normal-gamma distribution. (Note that Eq. 4-18 produces a normal-gamma distribution if a gamma distribution is used for $(\boldsymbol{\sigma}_{within})^{-2}$.)

4.2.3 Computational Methods

A numerical approach is adopted for the computations required to implement the Bayesian framework presented in Section 4.2.2. The numerical approach has two primary advantages: (1) it prevents the need for distributional assumptions for the parameters and results, and (2) it produces results that are readily interpreted via the Monte Carlo method for reliability analyses. The numerical method for reliability analyses is discussed in 4.3.

To understand the computational complexity of Bayesian calculations, consider the result of substituting the priors from Eq. 4-16 and Eq. 4-17 and the likelihood from Eq. 4-19 into the statement of Bayes' theorem of Eq. 4-15. The result is shown in Eq. 4-20. On the right side of the proportionality, the first line is the likelihood, the second line is the prior for the site-specific mean $\boldsymbol{\mu}_{within}$, and the third line is the prior for the site-specific within-site variability $\boldsymbol{\sigma}_{within}$. Computing

the right side of the proportionality requires integration because the likelihood is a function of μ_{within} and σ_{within} , both of which are random variables with their own probability distributions (in the second and third lines, respectively).

$$\begin{aligned}
 p(\mu_{within}, \sigma_{within} | \text{load test}) \propto & f_{normal}(\ln(RMP_{fail}), \mu_{within}, \sigma_{within}) \cdot \dots \\
 & \cdot f_{normal}(\ln(RMP), \mu_{among}, \sigma_{among}) \cdot \dots \quad \text{Eq. 4-20} \\
 & \cdot f_{normal}(\sigma, \mu_{COV_R}, \sigma_{COV_R})
 \end{aligned}$$

The integration required by Eq. 4-20 is simplified if certain distributional assumptions are satisfied. For example, if σ_{within} is assumed to be known and constant and μ_{within} is normally distributed, the posterior of μ_{within} will also be normally distributed, and simple solutions are available for the updated parameters of μ_{within} . This is the approach taken by Zhang and Tang (2002) (Section 2.3.2). Alternatively, if μ_{within} is normally distributed and $(\sigma_{within})^{-2}$ follows a gamma distribution, both the joint prior and the joint posterior follow the four-parameter normal-gamma distribution, and relatively simple formulas can be used to update the four parameters. This is the approach taken by Baecher and Rackwitz (1982) (Section 2.3.3). These solutions are referred to as conjugate distributions or conjugate pairs. Ang and Tang (2007) present a summary table with solutions for various conjugate distributions.

The numerical tool performs integration numerically. The numerical integration requires more computational effort than conjugate pairs, but any distributional form of μ_{within} and σ_{within} can be assumed with numerical integration. The method used for numerical integration in the tool is based on the grid method presented by Zhang (2009) and summarized by Juang and Zhang (2017).

Application of the grid method to Eq. 4-20 is depicted in Figure 27. The grid consists of two axes. The horizontal axis represents μ_{within} . The distribution shown below the horizontal axis is the prior distribution of μ_{within} . Values of μ_{within} near the center of the axis, i.e. near μ_{among} , are more likely than values of μ_{within} further from μ_{among} . Similarly, the vertical axis represents σ_{within} , and the distribution to the left of the vertical axis is the prior distribution of σ_{within} . Values

of σ_{within} near the center of the axis, i.e. near μ_{COV_R} , are more likely than values further from μ_{COV_R} . Each grid point is defined by one value of μ_{within} , $\mu_{within-i}$, and one value of σ_{within} , $\sigma_{within-j}$. Each grid point is associated with a joint prior, defined in Eq. 4-18 as the product of the independent prior distributions.

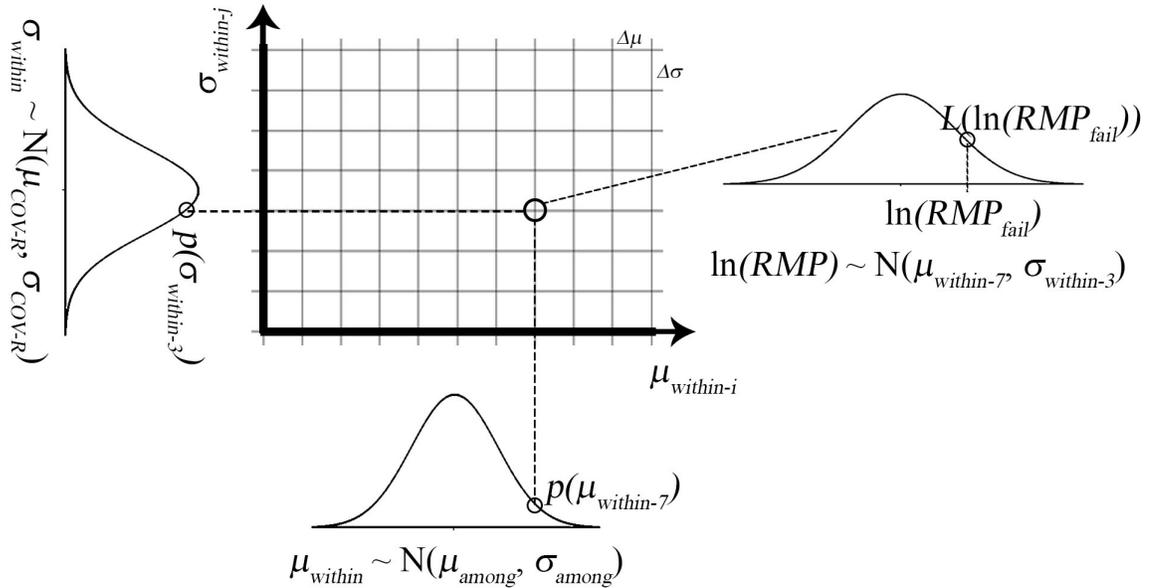


Figure 27: Depiction of numerical calculation method for Bayesian updating of foundation resistance in light of load test information.

Each grid point also defines a unique normal distribution, i.e.

$\ln(RMP) \sim N(\mu_{within-i}, \sigma_{within-j})$, that is used to evaluate the likelihood function for the grid point.

An example of one of the unique distributions for the likelihood is shown on the right side of

Figure 27. Because the example is for the grid point at the bottom right of the grid, the resulting

distribution has a relatively high mean ($\mu_{within-i} > \mu_{among}$) and relatively narrow dispersion

($\sigma_{within-j} < \mu_{COV_R}$). In the example shown in Figure 27, the observed load test resistance in terms

of $\ln RMP$, $\ln(RMP_{fail})$, is greater than μ_{among} but significantly less than $\mu_{within-i}$, so the likelihood

for the example grid point is relatively small.

The right side of the proportionality of Eq. 4-20 is solved with the grid method by computing the joint prior (p) and likelihood (L) for each point on the grid (similar to the example outlined in the preceding paragraph):

$$p(\mu_{within-i}, \sigma_{within-j} | \ln(RMP_{fail})) \propto L(\ln(RMP_{fail}) | \mu_{within-i}, \sigma_{within-j}) \cdot p(\mu_{within-i}, \sigma_{within-j}), \text{ where} \quad \text{Eq. 4-21}$$

$$L(\text{load test} | \mu_{within-i}, \sigma_{within-j}) = f_{normal}(\ln(RMP_{fail}), \mu_{within-i}, \sigma_{within-j}) \quad \text{Eq. 4-22}$$

$$p(\mu_{within-i}, \sigma_{within-j}) = f(\mu_{within-i}, \mu_{among}, \sigma_{among}) \cdot f(\sigma_{within-j}, \mu_{COV_R}, \sigma_{COV_R}) \quad \text{Eq. 4-23}$$

The likelihood function of Eq. 4-22 is based on the normal distribution PDF; defining the likelihood of observing load test data with respect to the normal distribution is the only distributional assumption of the numerical tool. In the examples throughout this section (e.g. the parameterization of Eq. 4-22 and Eq. 4-23), a normal distribution has been used for both μ_{within} and σ_{within} , but the tool accommodates any distribution for both μ_{within} and σ_{within} .

The left side of Eq. 4-21 is the unnormalized joint posterior distribution of μ_{within} and σ_{within} . It is proportional to the joint posterior, but integrating the unnormalized joint posterior over μ_{within} and σ_{within} would not yield a value of 1, so it is not a true probability distribution. To normalize the result of Eq. 4-21, it is necessary to divide all values of the unnormalized joint posterior by a constant, k . The constant is related to the sum of all points in the joint posterior:

$$k = \Delta_{\mu} \Delta_{\sigma} \sum_j \sum_i L(\text{load test} | \mu_{within-i}, \sigma_{within-j}) \cdot p(\mu_{within-i}, \sigma_{within-j}), \text{ where} \quad \text{Eq. 4-24}$$

Δ_{μ} is the width of the grid increments on the μ_{within} axis, and

Δ_{σ} is the width of the grid increments on the σ_{within} axis.

And therefore

$$p(\mu_{within-i}, \sigma_{within-j} | \ln(RMP_{fail})) = \frac{L(\ln(RMP_{fail}) | \mu_{within-i}, \sigma_{within-j}) \cdot p(\mu_{within-i}, \sigma_{within-j})}{k} \quad \text{Eq. 4-25}$$

Finally, the joint posterior distribution is used to define marginal posteriors for μ_{within} and σ_{within} , $p(\mu_{within-i} | \ln(RMP_{fail}))$ and $p(\sigma_{within-j} | \ln(RMP_{fail}))$, respectively:

$$p(\mu_{within-i} | \ln(RMP_{fail})) = \frac{\sum_j p(\mu_{within-i}, \sigma_{within-j} | \ln(RMP_{fail}))}{\frac{k}{\Delta_\sigma}} \quad \text{Eq. 4-26}$$

$$p(\sigma_{within-j} | \ln(RMP_{fail})) = \frac{\sum_i p(\mu_{within-i}, \sigma_{within-j} | \ln(RMP_{fail}))}{\frac{k}{\Delta_\mu}} \quad \text{Eq. 4-27}$$

The posterior marginal distributions define the expected values and variability of μ_{within} and σ_{within} , based on the prior distributions but updated to reflect the results of load tests, $\ln(RMP_{fail})$. The marginal distributions are defined numerically: one value of the posterior probability density function for each point on the corresponding axis (i.e. one value for each horizontal grid point i in Figure 27 for μ_{within} ; one value for each vertical grid line j in Figure 27 for σ_{within}). To interpret the posterior marginal distributions, normal distribution parameters are fitted to each: μ_μ and σ_μ for the mean and standard deviation of μ_{within} , respectively; μ_σ and σ_σ and for the mean and standard deviation of σ_{within} , respectively.

The posterior joint and marginal distributions are the primary output of the Bayesian updating component of the numerical tool. The posterior joint distribution is used to interpret the posterior predictive distribution, as described in Section 4.3. Also, the normal parameters fitted to the posterior marginal distributions are used in the interpretation of results in Chapters 5 and 6.

4.3 Numerical Method for Evaluating Foundation Reliability from Results of Bayesian Updating

The numerical tool uses the results of Bayesian updating to evaluate reliability of designs based on load tests. The reliability evaluations are based on the posterior predictive distribution of foundation resistance. As explained in Section 2.2, the posterior predictive distribution is used to characterize the expected value and expected variability of new estimates of foundation resistance based on both empirical methods and site-specific load tests. The posterior predictive distribution can more simply be termed the design distribution, as it describes the distribution of estimated resistance used in design. In this section, the numerical methodology for developing the posterior predictive distribution from the Bayesian updating results (Section 4.2) is described.

The section also includes explanation for how reliability is interpreted from the posterior predictive distribution. The interpretation is relatively simple, using a left-tail descriptor of the posterior predictive distribution as a measure of reliability.

The numerical method used in the tool to develop the posterior predictive distribution follows directly from previous discussions of among-site variability, within-site variability, and the Bayesian updating framework (Sections 1.1, 2.2, and 4.2). Based on those discussions, the uncertainty in resistance estimates can be characterized as originating from three distinct sources:

1. Within-site variability
2. Uncertainty in the mean value of foundation resistance (i.e. among-site variability)
3. Uncertainty in the within-site variability of foundation resistance (i.e. to account for the fact that true within-site variability is unknown)

The first source of uncertainty in resistance predictions is within-site variability. As discussed in Sections 1.1 and 2.2, within-site variability characterizes the distribution of foundation resistance for a site, given that the average resistance is known. Within-site variability results from geologic variation across a site and from variation in construction outcomes from one foundation element to another. The value of within-site variability for a given site and a given set of construction practices is generally fixed but strictly unknown.

The second source, uncertainty in the mean value of foundation resistance, is equivalent to among-site variability for cases in which there are no site-specific load tests and all foundation resistance information is derived from empirical methods. As explained in Sections 1.1 and 2.2, among-site variability represents the variation in the mean resistance from site to site, or among sites. Based on the findings from previous work regarding the reliability of designs based on load tests (e.g. Kay, 1976; Baecher and Rackwitz, 1982; Zhang and Tang, 2002), the uncertainty in the mean diminishes as site-specific load test information is considered. Stated differently, the effect of among-site variability is reduced as load test information becomes available.

The third source of uncertainty in resistance estimates based on load test information is uncertainty in the variability of the foundation resistance. This source refers not to within-site variability, itself, but rather to the fact that the within-site variability is not known precisely. The uncertainty in the within-site variability is characterized using a probabilistic distribution of within-site variability values, i.e. by assigning a standard deviation to the standard deviation of foundation resistance at the site. For example, consider the distribution of COV_R developed in Chapter 3 and shown in Figure 25. The mean value of COV_R , 0.15, represents the value of within-site variability; this is used to characterize the first source of uncertainty above. The standard deviation value of COV_R , 0.06, represents the uncertainty in within-site variability. The standard deviation of COV_R is used to characterize the third source of uncertainty. (As explained in Section 4.2, Eq. 4-8 establishes that the distribution from Chapter 3, defined in terms of COV_R , can be applied directly in the numerical tool, which performs Bayesian updating in terms of $\ln(RMP)$.) Note that if deterministic within-site variability is used, i.e. the within-site variability is treated as known and constant, the third source of uncertainty is neglected (e.g. Kay, 1976; Zhang and Tang, 2002; Zhang, 2004; Najjar et al., 2017). As noted in Sections 2.3.3, 2.5, and 3.4, the approach by Baecher and Rackwitz (1982) is the only previous work to consider the uncertainty in within-site variability. The evaluations of Chapters 5 and 6 compare results from both the deterministic and probabilistic within-site variability.

Each of the three sources of uncertainty is represented in the posterior results of Bayesian updating of Section 4.2. The posterior results include a posterior joint distribution of the site-specific mean parameter, μ_{within} , and the site-specific within-site variability parameter, σ_{within} . The first source of uncertainty described above, within-site variability, is represented by the magnitude of the values of σ_{within} represented in the posterior joint distribution. The second source of uncertainty, among site variability, is represented by the variability of the posterior joint distribution with respect to μ_{within} . The third source of uncertainty, uncertainty in the value of within-site variability, is represented by the variability of the posterior joint distribution with respect to σ_{within} .

The computational procedure for building the posterior predictive distribution from the posterior joint distribution of μ_{within} and σ_{within} is depicted in Figure 28. (Although the marginal distributions of μ_{within} and σ_{within} are shown atop Figure 28 for simplicity, in reality, the joint distribution is three-dimensional.) The computational procedure involves three random samplings:

1. Randomly sample a pair of μ_{within} and σ_{within} values from the posterior joint distribution of μ_{within} and σ_{within} . Call these values $\mu_{within-k}$ and $\sigma_{within-k}$. The pair represents a site-specific distribution of resistance: $\ln(RMP_k) \sim N(\mu_{within-k}, \sigma_{within-k})$.
2. Randomly sample one value of resistance from the site-specific distribution, $\ln(RMP_k)$. The value, $\ln(RMP_k)$, represents a randomly sampled value from the posterior predictive distribution of $\ln(RMP)$.
3. Compute the exponential of the result from (4) to find the result in arithmetic rather than exponential terms: $\exp(\ln(RMP_k)) = RMP_k$
4. Repeat steps (1) through (3) one million times.

Performing the four steps produces a collection of one million values of resistance in terms of RMP . The values are used to numerically define the posterior predictive distribution of RMP . Consistent with the discussion presented earlier in this section, each of the random samplings from the procedure accounts for one of the three sources of uncertainty:

- By sampling from μ_{within} , Step (1) above accounts for among-site variability
- By sampling from σ_{within} , Step (1) above also accounts for uncertainty in within-site variability
- Step (2) above accounts for within-site variability, itself.

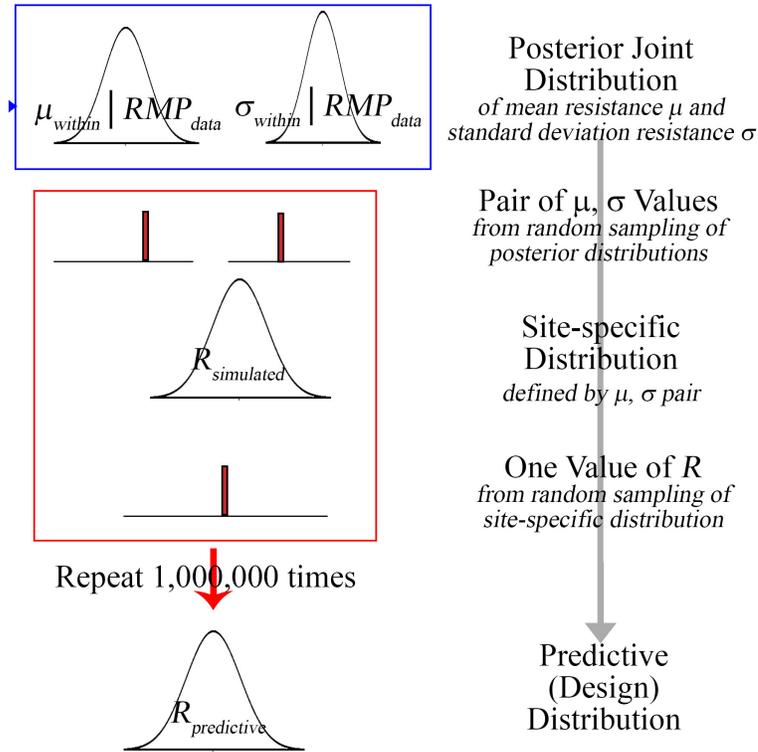


Figure 28: Random sampling procedure used to define posterior predictive distribution.

To characterize the reliability from the posterior predictive distribution (defined numerically by the one million values of RMP), a left-tail statistic is used. The statistic is the 1/1500th value of the posterior predictive distribution, $R_{1/1500}$. The 1/1500th value is the resistance value for which 1/1500 of resistance values are less and 1499/1500 of resistance values are greater. The 1/1500th value is also the design (or nominal) value of resistance associated with a probability of failure of 1/1500 if the loads are known perfectly and are perfectly balanced with the design resistance value. To calculate the 1/1500th value, the one million resistance values defining the posterior predictive distribution are sorted in ascending order, and the 667th (1,000,000 divided by 1500) value is designated as the 1/1500th value.

4.4 Summary

This chapter documents the methodology, notation, and output of the numerical tool used for Bayesian updating and reliability evaluation in this research. Prior distributions of the site-specific mean and site-specific within-site variability are input for the tool, along with site-specific

load test information. Estimates of $R_{1/1500}$ are the primary output of the tool. Posterior distributions of the mean and standard deviation of site-specific resistance are also output.

The first numerical analysis of the tool is Bayesian updating, which yields posterior distributions of the site-specific mean and site-specific within-site variability parameters. The tool fits normal distributions to posterior distributions of the mean and the within-site variability; the fitted normal distribution parameters are used in Chapters 5 and 6 to interpret the results of Bayesian updating. In the tool, Bayesian updating can be performed with either deterministic or probabilistic within-site variability. In the former, within-site variability is assumed to be a known and constant value and is not updated based on load test results. In the latter, within-site variability is treated as an unknown parameter to be updated.

The posterior joint distribution of the site-specific mean and site-specific within-site variability parameters is used as input for the second numerical analysis of the tool, which is a reliability analysis. The posterior joint distribution of the mean and standard deviation is used to develop a posterior predictive distribution that represents the magnitude and variability of new predictions of foundation resistance at a site with load test information. The tool outputs the 1/1500th value of resistance, which characterizes the left tail of the posterior predictive resistance distribution. The 1/1500th value is used in Chapters 5 and 6 to evaluate the reliability implications of Bayesian updating with deterministic and probabilistic within-site variability for different site conditions and load test results.

5. Effect of Deterministic vs. Probabilistic Within-site Variability for Cases with “Perfect” Sampling of Load Test Results

The primary objective of this chapter is to use the numerical tool described in Chapter 4 to evaluate the hypothesis. Specifically, the tool is used to evaluate the effect of deterministic versus probabilistic consideration of within-site variability on the interpretation of the 1/1500th value of resistance, $R_{1/1500}$. As presented in Chapter 1, the hypothesis includes two main statements about $R_{1/1500}$. First, for cases without load tests, $R_{1/1500}$ is hypothesized to be greater for deterministic within-site variability because uncertainty in within-site variability is neglected for deterministic within-site variability, whereas uncertainty in within-site variability is considered for probabilistic within-site variability. Second, for cases with very many load tests (i.e., in the limit), $R_{1/1500}$ is hypothesized to converge to the true value for probabilistic within-site variability, while $R_{1/1500}$ is hypothesized to not converge to the true value for deterministic within-site variability in most cases.

A secondary objective of this chapter is to explain any trends observed in the evaluation of the hypothesis with corresponding trends in the posterior distributions of the mean and standard deviation of foundation resistance. Evaluation of trends in the posterior distributions of the mean and standard deviation of resistance is facilitated by use of “perfect” sampling of load test results for computational tool. The descriptor “perfect” is used here to indicate that the load test statistics, namely the sample mean and sample standard deviation, are equal to the values intended for analysis and evaluation. Perfect sampling is in contrast to random sampling, in which load test values are sampled at random from a distribution, e.g. from a normal distribution with a mean and standard deviation intended for analysis and evaluation. Use of perfect sampling with the numerical tool is analogous to sensitivity studies with a theoretical model, i.e. varying the prior and load test inputs for the posterior parameter equations from Zhang and Tang (2002) or Baecher and Rackwitz (1982). Chapter 6 includes similar evaluations to those presented in this chapter, but with random rather than perfect sampling. Random sampling is more realistic than

perfect sampling; the evaluation of Chapter 6 therefore focuses on rates of convergence and the practical ramifications of the trends observed in this chapter.

In the first section of this chapter, details of the methodology are presented, including perfect sampling and how the numerical tool from Chapter 4 was used. The second section focuses on results of cases without load tests and cases with many load tests to evaluate the hypothesis as described in the first paragraph above. The third section of the chapter focuses on trends in the posterior distributions of site-specific mean and standard deviation, and how the trends relate to the observations in the limit. The chapter closes with a summary of significant findings.

5.1 Evaluation Methodology

The methodology employed to evaluate the hypothesis is depicted in Figure 29. As listed atop the figure, the methodology includes three major steps: (1) simulate load test data from the prior distributions of mean and standard deviation of resistance, (2) use load test data to update the prior distributions, and (3) interpret reliability from the posterior distributions. Details of the second and third steps were provided in Chapter 4 (in Sections 4.2 and 4.3, respectively); this section focuses on the first step.

The procedure used to simulate load test results was designed to evaluate the hypothesis under a range of conditions: site conditions, foundation design conditions, and load test result conditions. The parameters used to define the various site conditions are defined in Table 8. The parameters are divided into two categories: parameters used to define the prior distributions of μ_{within} and σ_{within} , and parameters used to define the “true” site conditions. The last row in the table, number of load tests, is the final parameter that was varied among the simulations. One subsection below is dedicated to each of the two categories of parameters. The final subsection discusses how the results of analyses are interpreted.

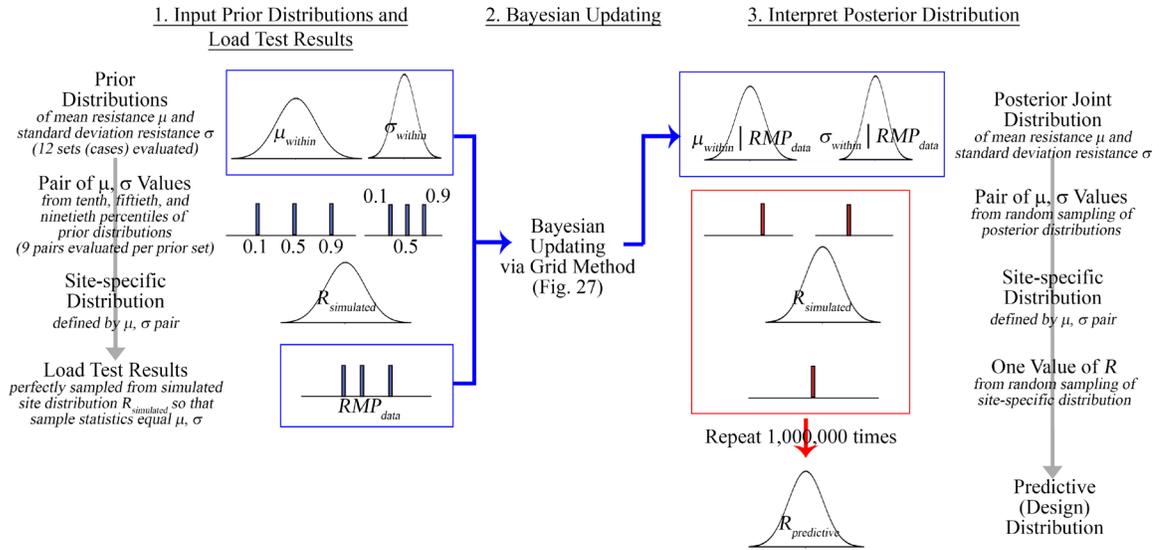


Figure 29: Overview of methodology used to evaluate hypothesis via the numerical tool.

Table 8: Summary of input parameters used to generate perfectly sampled load test results.

Parameter		Input Levels	Values
Prior Distributions of μ_{within} and σ_{within}	Total variability of resistance, σ_{total}	3	0.1 (Figure 30) 0.3 (Figure 31) 0.5 (Figure 32)
	Ratio of within-site to among-site variability, $\mu_{COV_R}/\sigma_{among}$	3	Approximately 0.2 (Points A in figures referenced above) 1 (Points B and D) 2.0 (Points C);
	Uncertainty in within-site variability, $\sigma_{COV_R}/\mu_{COV_R}$	2	Approximately 0.25 (Points A, B, C in figures referenced above) 0.5 to 0.8 (Points D)
"True" Site Condition	μ_{within}	3	10th, 50th, and 90th percentiles
	σ_{within}	3	10th, 50th, and 90th percentiles
Number of Load Tests		15	0, 1, 2, 3, 4, 5, 7, 10, 20, 30, 40, 50, 200, 400, and 800

5.1.1 Parameters Defining Prior Distributions of μ_{within} and σ_{within}

As explained in Chapter 4, the prior distributions of μ_{within} and σ_{within} are critical inputs for Bayesian updating and for the evaluation of foundation resistance. The prior distribution of μ_{within} , the site-specific mean, represents the information from an empirical design method, with the variability of μ_{within} representing the uncertainty associated with the design method. In Chapter 4, the variability of μ_{within} was represented by σ_{among} , since the uncertainty of empirical design methods is dominated by among-site variability, at least in the convention of Chapter 4.

The prior distribution of σ_{within} represents prior information regarding within-site variability. The mean parameter of σ_{within} , as presented in Chapter 4, is μ_{COV_R} , which represents the expected magnitude of within-site variability. The literature regarding Bayesian updating for load test information summarized in Section 2.3 has shown the prior distributions of μ_{within} and σ_{within} govern the weight given to load test results (e.g. Zhang and Tang, 2002; Baecher and Rackwitz, 1982). In addition, the prior distributions of μ_{within} and σ_{within} define the prior predictive distribution of resistance.

As presented in Table 8, the three parameters used to define the prior distributions of μ_{within} and σ_{within} are the total variability of RMP , σ_{total} , the ratio of within-site variability to among site variability, $\mu_{COV_R}/\sigma_{among}$, and the uncertainty in within-site variability, $\sigma_{COV_R}/\mu_{COV_R}$. By the convention presented in Chapter 4, σ_{total} is the standard deviation of $\ln(RMP)$, which is approximately equal to the COV of resistance by Eq. 4-8 (Section 4.2.1). Thus, the three nominal values of σ_{total} considered, 0.1, 0.3, and 0.5, are approximately equivalent to predictive distributions of resistance with COV values of 0.1, 0.3, and 0.5. These values are used to represent relatively low variability (0.1), medium variability (0.3), and relatively high variability (0.5). The values are described as nominal because the total predictive COV values cannot be perfectly satisfied by both deterministic and probabilistic within-site variability with the same input values since deterministic within-site variability neglects uncertainty in within-site variability. This concept is addressed in Section 5.2.1.

The ratio of within-site variability to among-site variability, $\mu_{COV_R}/\sigma_{among}$, is used to evaluate conditions in which the total variation of resistance is driven largely by within-site variability, conditions in which within-site variability and among-site variability contribute approximately equally to the total variation of resistance, and conditions in which total variation of resistance is governed by among-site variability. As indicated in Table 8, these cases are represented by values of $\mu_{COV_R}/\sigma_{among}$ approximately equal to 0.2, 1, and 2, respectively. Each of the three values of $\mu_{COV_R}/\sigma_{among}$ was considered for each level of σ_{total} . The resulting combinations of σ_{total} and $\mu_{COV_R}/\sigma_{among}$ are shown as points in Figure 30, Figure 31, and Figure

32. Each of the figures is a plot of among-site variability, σ_{among} , versus within-site variability, μ_{COV_R} , for a constant value of total variability, σ_{total} . Cases where the total variability is dominated by among-site variability are represented in the figures by Points A, which are near the vertical, among-site variability axis. Cases where among-site and within-site variability contribute approximately equally to total variability are represented in the figures by Points B, which fall near a line inclined 45-deg., halfway between the vertical and horizontal axes. Cases where the total variability is dominated by within-site variability are represented in the figures by Points C, which are near the horizontal, within-site variability axis.

In addition, for each value of σ_{total} , two cases with $\mu_{COV_R}/\sigma_{among}$ equal to 1 are considered, with one representing the same level uncertainty in the value of within-site variability used for the other values of $\mu_{COV_R}/\sigma_{among}$ (Points A in Figure 30, Figure 31, and Figure 32), and the other representing significantly greater uncertainty in the value of within-site variability (Points D in Figure 30, Figure 31, and Figure 32). Specifically, the uncertainty in within-site variability used for most cases corresponds to $\sigma_{COV_R}/\mu_{COV_R}$ approximately equal to 0.25; for the special cases with greater uncertainty in within-site variability, $\sigma_{COV_R}/\mu_{COV_R}$ ranged from 0.5 to 0.8.

As explained in Section 4.2.2, σ_{total} can be calculated directly from the among-site variability and within-site variability parameters for the case of deterministic within-site variability (Eq. 4-13). For the case of probabilistic within-site variability considered here, Eq. 4-13 does not apply because it does not account for uncertainty in the within-site variability parameter. To satisfy the target values of σ_{total} and $\mu_{COV_R}/\sigma_{among}$ listed in Table 8, a numerical procedure was used. In the procedure, the prior distribution of within-site variability was specified, and the value of among-site variability was computed by iteration to satisfy the target value of σ_{total} .

The results are shown in Figure 30, Figure 31, and Figure 32 for σ_{total} equal to 0.1, 0.3, and 0.5, respectively. In the figures, one line representing among-site variability σ_{among} , versus within-site variability, μ_{COV_R} , is shown for each of several different values of uncertainty in within-site variability, σ_{COV_R} . For small uncertainty in the within-site variability, the shape of the lines is

nearly circular, which is consistent with the equation for deterministic total variability of resistance, Eq. 4-13. The deterministic case, $\sigma_{COV_R} = 0$, is represented by the solid black line in each figure. For small prior predictive variability (Figure 30), the deterministic case intercepts both the among-site (vertical) and within-site (horizontal) axes at the value corresponding to the prior predictive COV of 0.1. For greater prior predictive variability, especially $\sigma_{total} = 0.5$ (Figure 32), the deterministic case intercepts both axes at values slightly less than the prior predictive COV value. The difference between the axes values, which are parameters of the logarithm of RMP , $\ln RMP$, and the prior predictive COV , which describes the arithmetic resistance, is equal to the error associated with using the approximation of Eq. 4-6. As noted in Section 4.2.1, the difference is insignificant below $COV = 0.3$. For $COV = 0.5$, the difference is about 6 percent (0.47 versus 0.5).

As uncertainty in within-site variability increases, the lines are less circular, indicating the influence of uncertainty in within-site variability is more significant. Also shown in each figure is a dashed line inclined 45 deg. from horizontal. The dashed line represents equal contributions from within-site and among-site variability. Points above the line have greater among-site variability than within-site variability; points below the line have greater within-site variability than among-site variability.

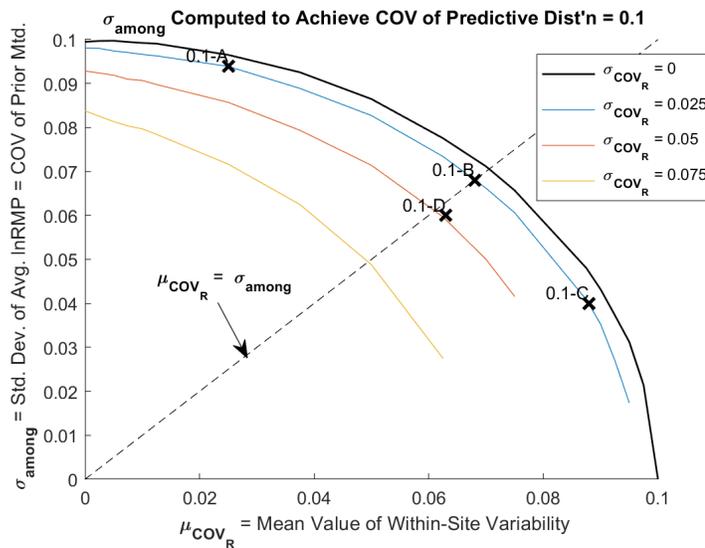


Figure 30: Among-site variability versus within-site variability to achieve total variability in resistance of $COV = 0.1$ for different values of uncertainty in within-site variability.

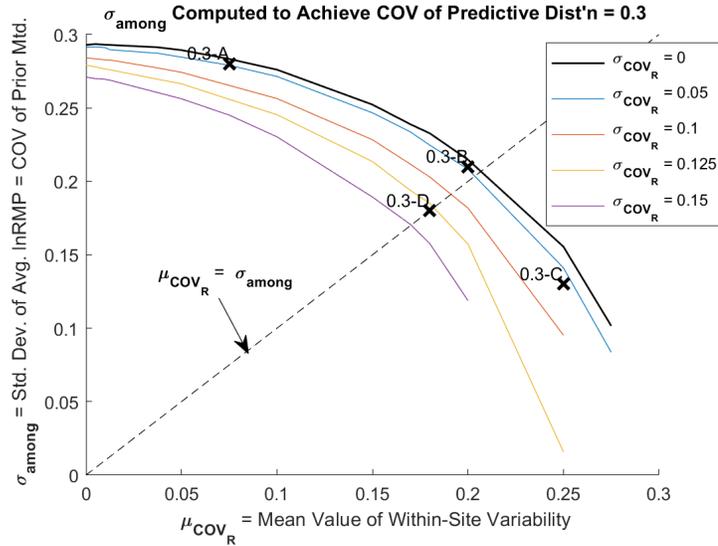


Figure 31: Among-site variability versus within-site variability to achieve total variability in resistance of $COV = 0.3$ for different values of uncertainty in within-site variability.

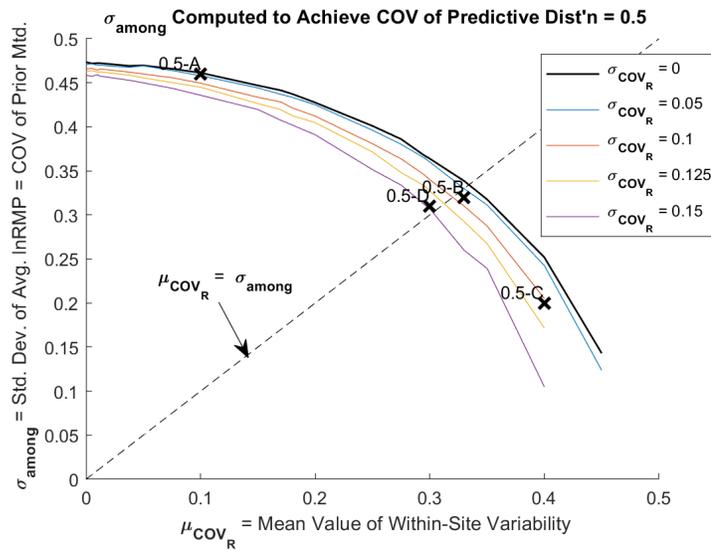


Figure 32: Among-site variability versus within-site variability to achieve total variability in resistance of $COV = 0.5$ for different values of uncertainty in within-site variability.

A summary of the 12 sets of prior distributions represented by Points A, B, C, and D in Figure 30, Figure 31, and Figure 32 is presented in Table 9. The values listed in the first column of the table are used to reference the cases in presentation of results later in this chapter and in Chapter 6. The next three columns present calculated parameters – total variability, ratio of within-site variability to among-site variability, and uncertainty in within-site variability – while the

final four columns present the normal distribution parameters used to define the prior distributions of μ_{within} and σ_{within} in terms of $\ln(RMP)$.

Table 9: Summary of 12 sets of prior distribution parameters used for perfect sampling analysis.

Case	σ_{total}	$\mu_{COV_R}/\sigma_{among}$	$\sigma_{COV_R}/\mu_{COV_R}$	μ_{within}		σ_{within}	
				μ_{among}	σ_{among}	μ_{COV_R}	σ_{COV_R} ¹
0.1-A	0.1	0.26	0.30	-0.005	0.094	0.025	0.0075
0.1-B	0.1	1.01	0.25	-0.002	0.07	0.07	0.018
0.1-C	0.1	2.14	0.29	-0.001	0.04	0.088	0.025
0.1-D	0.1	1.04	0.80	-0.002	0.06	0.063	0.05
0.3-A	0.3	0.26	0.33	-0.040	0.28	0.075	0.025
0.3-B	0.3	0.97	0.25	-0.021	0.21	0.20	0.05
0.3-C	0.3	1.97	0.30	-0.008	0.13	0.25	0.08
0.3-D	0.3	0.98	0.69	-0.017	0.18	0.18	0.13
0.5-A	0.5	0.22	0.25	-0.11	0.46	0.10	0.025
0.5-B	0.5	1.02	0.23	-0.053	0.32	0.33	0.075
0.5-C	0.5	1.97	0.25	-0.021	0.20	0.40	0.10
0.5-D	0.5	0.97	0.50	-0.048	0.31	0.3	0.15

¹Values of σ_{COV_R} listed are for probabilistic within-site variability. For deterministic within-site variability, $\sigma_{COV_R} = 0$.

5.1.2 Parameters Defining “True” Site Conditions and Perfect Sampling

The second category of parameters described in Table 8 are the “true” site conditions: the real values of site-specific mean, μ_{within} , and site-specific within-site variability, σ_{within} . The true values are known because they are specified in the simulation process. In reality, i.e. outside the simulation process, the true values are not known exactly unless all foundations at a site (i.e. the population) are load tested to failure. The true values are defined with respect to the prior distributions of μ_{within} and σ_{within} discussed in Section 5.1.1. Nine combinations are considered from three levels of μ_{within} and three levels of σ_{within} . For both μ_{within} and σ_{within} , the levels are defined at the tenth, fiftieth, and ninetieth percentiles, as shown in Figure 29 and listed in Table 8. For μ_{within} , the tenth percentile corresponds to a relatively site with relatively weak foundation resistance, the fiftieth percentile corresponds to a site with average resistance, and the ninetieth percentile corresponds to a site with relatively strong resistance. For σ_{within} , the tenth percentile corresponds to a site with relatively uniform foundation resistance, the fiftieth percentile corresponds to a site with average variability of resistance, and the ninetieth percentile

corresponds to a site with relatively variable resistance. The nine combinations of μ_{within} and σ_{within} were evaluated for each of the posterior distribution sets evaluated. In other words, the nine combinations were evaluated for each of the Points A, B, C, and D in Figure 30, Figure 31, and Figure 32.

As described in the introduction to this chapter, perfect sampling was employed for evaluation of the hypothesis. Perfect sampling means that for all simulation sets, i.e. all combinations of prior distributions, site parameters, and numbers of tests, the sample mean and sample standard deviation are equal to the true values of μ_{within} and σ_{within} , respectively. Perfect sampling removes one source of variability, random sampling, from the Bayesian analysis and interpretation since the load test result statistics match the true site conditions for all numbers of tests, as opposed to only for large sample sizes, as in the case of random sampling. For evaluating posterior trends, especially in Section 5.3, perfect sampling is preferable to random sampling because the additional variability associated with random sampling is avoided.

To achieve perfect sampling, values of $\ln(RMP)$ were sampled at random from the target site parameters, $\mu_{within-true}$ and $\sigma_{within-true}$. All randomly sampled values, $\ln(RMP)_i$, were then adjusted according to Eq. 5-1 to produce the “perfect” sample, with values denoted $\ln(RMP)_{i-perfect}$.

$$\ln(RMP)_{i-perfect} = \frac{\ln(RMP)_i - \hat{\mu}_{\ln(RMP)_i}}{\hat{\sigma}_{\ln(RMP)_i}} \cdot \sigma_{within-target} + \mu_{within-target} \quad , \text{ where} \quad \text{Eq. 5-1}$$

$\hat{\mu}_{\ln(RMP)_i}$ is the sample mean of the randomly sampled values $\ln(RMP)_i$, and

$\hat{\sigma}_{\ln(RMP)_i}$ is the sample standard deviation of the randomly sampled values $\ln(RMP)_i$.

5.1.3 Parameters Used for Evaluation of Results

Five parameters were used in the interpretation of results: $R_{1/1500}$, μ_μ , σ_μ , μ_σ , and σ_σ . The parameters are defined in Table 10, which also lists how each parameter is interpreted. All five parameters are output of the numerical tool, and all were described in Chapter 4. $R_{1/1500}$ is the 1/1500th value of resistance from the predictive distribution: the value for which 1/1500 of

resistance values are smaller and 1499/1500 are larger. $R_{1/1500}$ is therefore a left-tail statistic; it is the value of resistance with a probability of failure of 1/1500 if loads are known perfectly and are perfectly balanced with $R_{1/1500}$. While $R_{1/1500}$ characterizes reliability from the posterior predictive distribution, the other four parameters are fitted directly to the marginal posterior distributions of $\boldsymbol{\mu}_{within}$ and $\boldsymbol{\sigma}_{within}$: μ_{μ} and σ_{μ} for the mean and standard deviation of $\boldsymbol{\mu}_{within}$, respectively; μ_{σ} and σ_{σ} and for the mean and standard deviation of $\boldsymbol{\sigma}_{within}$, respectively.

The evaluations in the rest of this chapter (and in Chapter 6) compare the five parameters resulting from the Bayesian analyses and interpretation to their “true” values. The true values are derived from the known simulation input for the tool, as summarized in Table 10. For $\boldsymbol{\mu}_{within}$, the true value of μ_{μ} is $\mu_{within-true}$, which was introduced in Section 0 and Eq. 5-1. The value of $\mu_{within-true}$ depends on the prior distribution of $\boldsymbol{\mu}_{within}$, which, as described in Section 5.1.1, depends on the total variability σ_{total} and the ratio of within-site variability to among-site variability, $\mu_{COV_R}/\sigma_{among}$. The value of $\mu_{within-true}$ also depends on the percentile of $\boldsymbol{\mu}_{within}$, which was specified as either the tenth, fiftieth, or ninetieth percentile as described in Section 0. Similarly, for $\boldsymbol{\sigma}_{within}$, the true value of μ_{σ} is $\sigma_{within-true}$, which was also introduced in Section 0 and Eq. 5-1. The value of $\sigma_{within-true}$ depends on σ_{total} , the ratio $\mu_{COV_R}/\sigma_{among}$, and the percentile of $\boldsymbol{\sigma}_{within}$.

For both $\boldsymbol{\mu}_{within}$ and $\boldsymbol{\sigma}_{within}$, the true standard deviation, σ_{μ} and σ_{σ} , respectively, is zero. Resistance at a site is defined only by the true mean, i.e. the average resistance at the site, and the true standard deviation, i.e. the true within-site variability. The parameters σ_{μ} and σ_{σ} account for uncertainty in the mean resistance and uncertainty in within-site variability, respectively, but for a simulated site where $\mu_{within-true}$ and $\sigma_{within-true}$ are known, there is no uncertainty in $\boldsymbol{\mu}_{within}$ or $\boldsymbol{\sigma}_{within}$, so σ_{μ} and σ_{σ} are both zero. As stated previously, the case of known μ_{within} and σ_{within} (and therefore zero σ_{μ} and σ_{σ}) is only true for simulation and generally not for reality.

The true value of $R_{1/1500}$ is derived from the site-specific distribution of resistance, which is defined by $\mu_{within-true}$ and $\sigma_{within-true}$. To compute the true value of $R_{1/1500}$, a simulation

procedure was used. The procedure mirrors the procedure used to compute $R_{1/1500}$ from the prior and posterior predictive distributions of resistance, which was detailed in Section 4.3. First, one million values are randomly sampled from the distribution of $\ln(RMP)$ defined by $\mu_{within-true}$ and $\sigma_{within-true}$. The 1/1500th value of the distribution is found by sorting the randomly sampled values and selecting the 667th smallest value ($\frac{1,000,000}{1500} = 667$). The exponential of the result is the true value of $R_{1/1500}$.

Table 10: Summary of parameters used to evaluate results.

Parameter	Definition	Interpretation from Numerical Tool	“True” Value
$R_{1/1500}$	1/1500th value of predictive distribution	Sampling procedure described in Section 4.3	1/1500th value of distribution defined by $\mu_{within-true}$ and $\sigma_{within-true}$
μ_{μ}	Mean value of posterior distribution of μ_{within}	Fit to posterior distribution of μ_{within} from numerical tool (Section 4.2.3)	$\mu_{within-true}$
σ_{μ}	Standard deviation value of posterior distribution of μ_{within}		0
μ_{σ}	Mean value of posterior distribution of σ_{within}	Fit to posterior distribution of σ_{within} from numerical tool (Section 4.2.3)	$\sigma_{within-true}$
σ_{σ}	Standard deviation value of posterior distribution of σ_{within}		0

5.2 Reliability at the Limits: Zero Load Tests and Many Load Tests

In this section, the hypothesis is evaluated directly by comparing $R_{1/1500}$ for deterministic and probabilistic consideration of within-site variability. The comparisons focus on two cases: cases without load testing and cases with many load tests. All results were obtained using the numerical tool detailed in Chapter 4 and following the methodology presented in Section 5.1. Results for cases without load testing are presented first; results for cases with many load tests are presented second.

5.2.1 Reliability without Load Tests

For cases without load tests, it was hypothesized in Chapter 1 that deterministic consideration of within-site variability would produce greater estimates of $R_{1/1500}$ than probabilistic consideration of within-site variability. Probabilistic within-site variability includes the effect of

uncertainty in within-site variability; the uncertainty should produce additional variability in the predictive distribution of resistance. The additional variability should produce a wider distribution and therefore a smaller value of $R_{1/1500}$ than deterministic within-site variability, which does not consider uncertainty in within-site variability.

To evaluate the hypothesis, estimates of $R_{1/1500}$ are compared for deterministic and probabilistic within-site variability for 12 cases without load tests. The cases represent the 12 sets of prior distributions of μ_{within} and σ_{within} described in Section 5.1.1, shown in Figure 30, Figure 31, and Figure 32, and summarized in Table 9. As explained in Section 5.1.1, the cases represent three nominal levels of total variability in the predictive distribution (σ_{total}), three levels of the ratio of within-site variability to among-site variability ($\mu_{COV_R}/\sigma_{among}$), and two levels of uncertainty in the within-site variability parameter ($\sigma_{COV_R}/\mu_{COV_R}$). For each of the cases, the numerical tool described in Chapter 4 was used to compute $R_{1/1500}$ with deterministic within-site variability and with probabilistic within-site variability. “True” values of $R_{1/1500}$ were also calculated by assuming the site-specific distribution for each case is defined by the mean parameters of μ_{within} and σ_{within} :

$$\mu_{within-true} = \mu_{among} \quad \text{Eq. 5-2}$$

$$\sigma_{within-true} = \mu_{COV_R} \quad \text{Eq. 5-3}$$

The assumptions stated in Eq. 5-2 and Eq. 5-3 are unique to this analysis of cases without load tests. For cases with load tests (Sections 5.2.2 and 5.3), the true values of site-specific parameters were evaluated at the tenth, fiftieth, and ninetieth percentiles, as described in Section 0. For the case without load tests considered in this section, there is no need to sample load tests, so the comparison focuses on just the fiftieth percentile (i.e. mean for the normal distribution, i.e. Eq. 5-2 and Eq. 5-3) values.

Results of the analyses are presented in Figure 33, Figure 34, Figure 35, and Figure 36. Each figure plots the difference between $R_{1/1500}$ from deterministic within-site variability,

$R_{1/1500-deterministic}$, and $R_{1/1500}$ from probabilistic within-site variability, $R_{1/1500-probabilistic}$, versus uncertainty in the within-site variability. Note the value of uncertainty in within-site variability is considered only in probabilistic within-site variability; for deterministic within-site variability, uncertainty in within-site variability is neglected. In each of the plots, points are labeled to indicate the prior distribution parameter cases as summarized in Table 9. The difference among the figures is in how the axis quantities are normalized. In Figure 33, neither quantity is normalized: $(R_{1/1500-deterministic} - R_{1/1500-probabilistic})$ versus σ_{COV_R} . In Figure 34, the uncertainty in within-site variability is normalized by the average value of within-site variability (i.e. the horizontal axis is COV of within-site variability): $(R_{1/1500-deterministic} - R_{1/1500-probabilistic})$ versus $\sigma_{COV_R}/\mu_{COV_R}$. In Figure 35, the difference in $R_{1/1500}$ is normalized by the true value of $R_{1/1500}$, but the uncertainty in within-site variability is not normalized: $\frac{(R_{1/1500-deterministic} - R_{1/1500-probabilistic})}{R_{1/1500-true}}$ versus σ_{COV_R} . As explained in Section 5.1.3, the true value of $R_{1/1500}$ is found from the predictive distribution of resistance defined by the known site-specific distribution parameters used for the simulations. In Figure 36, both quantities are normalized: $\frac{(R_{1/1500-deterministic} - R_{1/1500-probabilistic})}{R_{1/1500-true}}$ versus $\sigma_{COV_R}/\mu_{COV_R}$.

Results for all 12 cases support the hypothesis: the value of $R_{1/1500}$ from deterministic within-site variability is always greater than the value from probabilistic within-site variability for cases without load tests, as indicated by positive y-values for all points. That deterministic within-site variability yields greater values of $R_{1/1500}$ without load tests compared to deterministic within-site variability is evident from the fact that all of the $R_{1/1500}$ difference values plotted in Figure 33, Figure 34, Figure 35, and Figure 36 are positive.

Results from the 12 cases also support the explanation of the hypothesis provided at the beginning of this section, i.e. that $R_{1/1500}$ from deterministic within-site variability is greater because it does not consider uncertainty in within-site variability. In each of the plots of Figure 33, Figure 34, Figure 35, and Figure 36, the difference in the estimate of $R_{1/1500}$ between deterministic and probabilistic within-site variability is greater for greater uncertainty in within-site

variability. A linear trend is observed in each of the plots, i.e. for all forms of normalization. However, normalization of the uncertainty in within-site variability in Figure 34 and Figure 36 confuses the trend somewhat, most likely because the 12 cases are largely divided into two values of $\sigma_{COV_R}/\mu_{COV_R}$. Indeed, the linear trend is stronger in Figure 33 and Figure 35 (without normalization of within-site variability uncertainty), and especially in the normalization of the $R_{1/1500}$ difference in Figure 35.

The linear trend in Figure 35 indicates that the difference between the value of $R_{1/1500}$ with deterministic within-site variability and with probabilistic within-site variability is (1) greater for greater uncertainty in within-site variability, as quantified by the standard deviation in COV of within-site resistance, and (2) that the difference in $R_{1/1500}$ is also greater for greater values of $R_{1/1500}$. The first observation is supported by the linear trend in the data. As stated previously, the first observation is also consistent with the explanation of the hypothesis presented at the beginning of this section. The second observation is supported by the normalization of the $R_{1/1500}$ difference by the true value, and by how the normalization improved linearity (i.e. from Figure 33 to Figure 35).

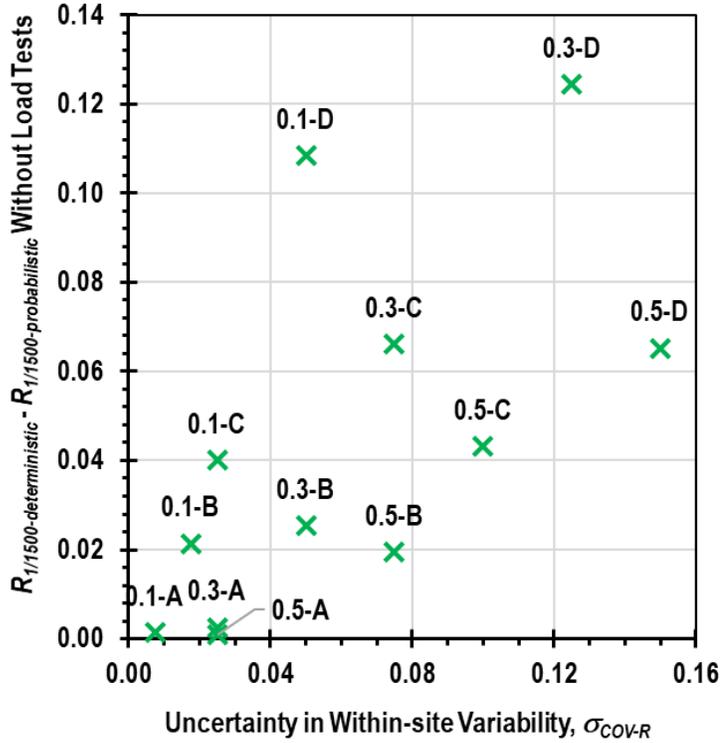


Figure 33: Difference between deterministic and probabilistic values of $R_{1/1500}$ for cases without load tests versus standard deviation in within-site variability.

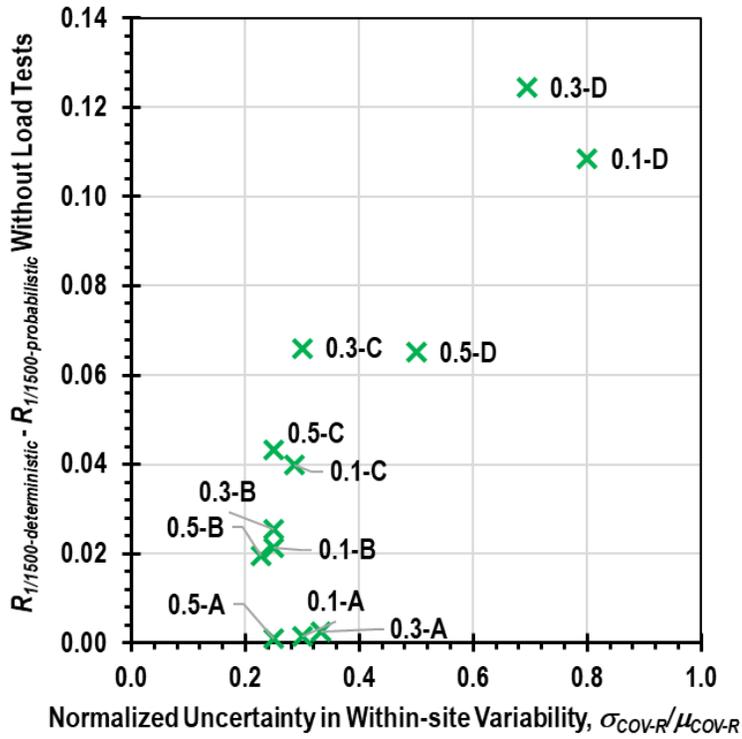


Figure 34: Difference between deterministic and probabilistic values of $R_{1/1500}$ for cases without load tests versus COV of within-site variability.

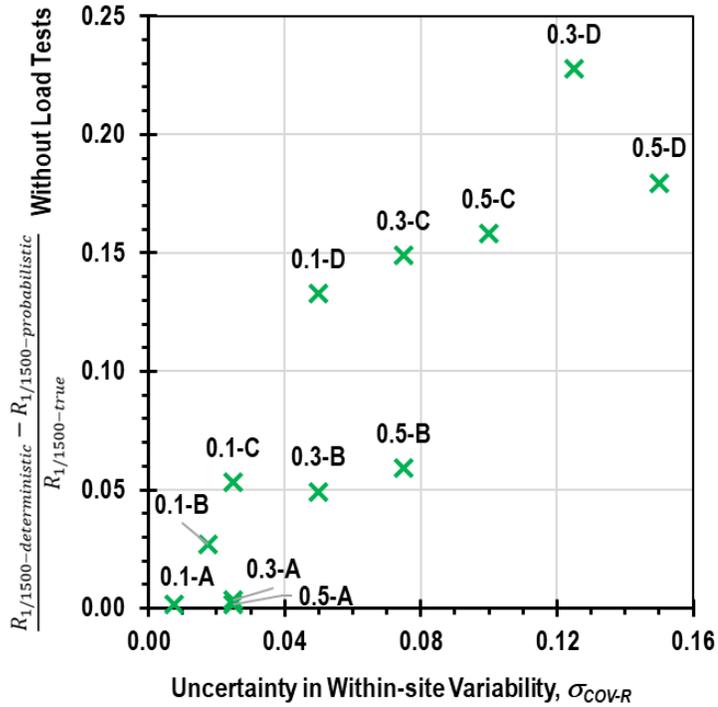


Figure 35: Difference between deterministic and probabilistic normalized values of $R_{1/1500}$ for cases without load tests versus standard deviation in within-site variability.

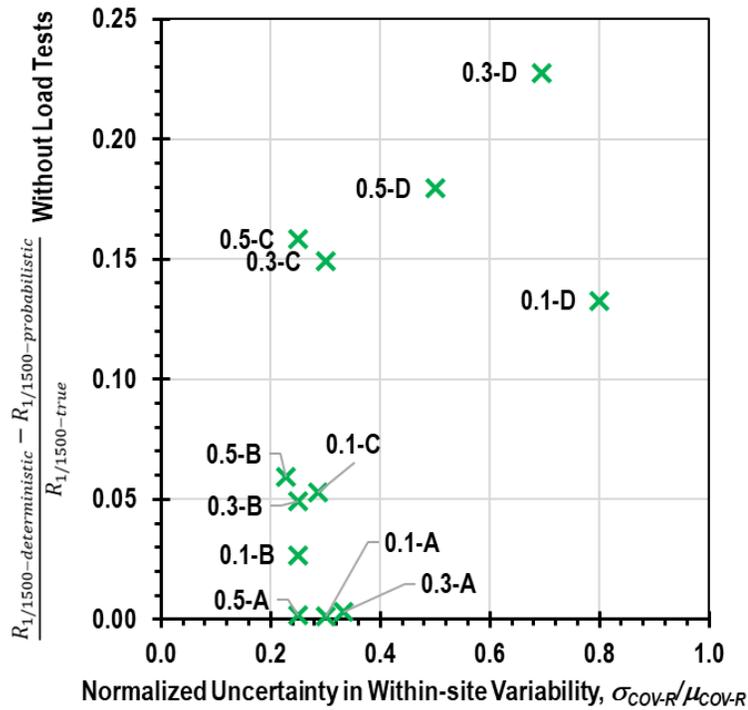


Figure 36: Difference between deterministic and probabilistic normalized values of $R_{1/1500}$ for cases without load tests versus COV of within-site variability.

In Section 5.1.1, the total predictive COV values were identified as nominal values because the total predictive COV values cannot be perfectly satisfied with both deterministic and probabilistic within-site variability with the same input values. For deterministic within-site variability, uncertainty in within-site variability is neglected (i.e. $\sigma_{COV_R} = 0$), which reduces the total predictive COV . That deterministic within-site variability produces a lower total predictive COV compared with probabilistic within-site variability is supported by Figure 37, which is a plot of the prior (i.e. no load tests) predictive COV versus the ratio of within-site to among-site variability for the cases with nominal COV of 0.5. Values of the total prior predictive COV were computed from the sample statistics of the random sample used to define the prior predictive distribution numerically (Section 4.3). For probabilistic within-site variability, the true value of total COV is always within 1 percent of 0.5. For deterministic within-site variability, the true value varies from 0.45 (Case 0.5-D) to 0.5 (Case 0.5-A), the best agreement is for the case with least within-site variability (Case 0.5-A), and the difference between the nominal and true value is greatest for the case with relatively great within-site variability (Case 0.5-C) and the case with great uncertainty in the value of within-site variability (0.5-D). The differences between nominal and true COV for deterministic within-site variability are consistent with the findings of greater estimates of $R_{1/1500}$ with deterministic within-site variability compared with estimates of $R_{1/1500}$ with probabilistic within-site variability.

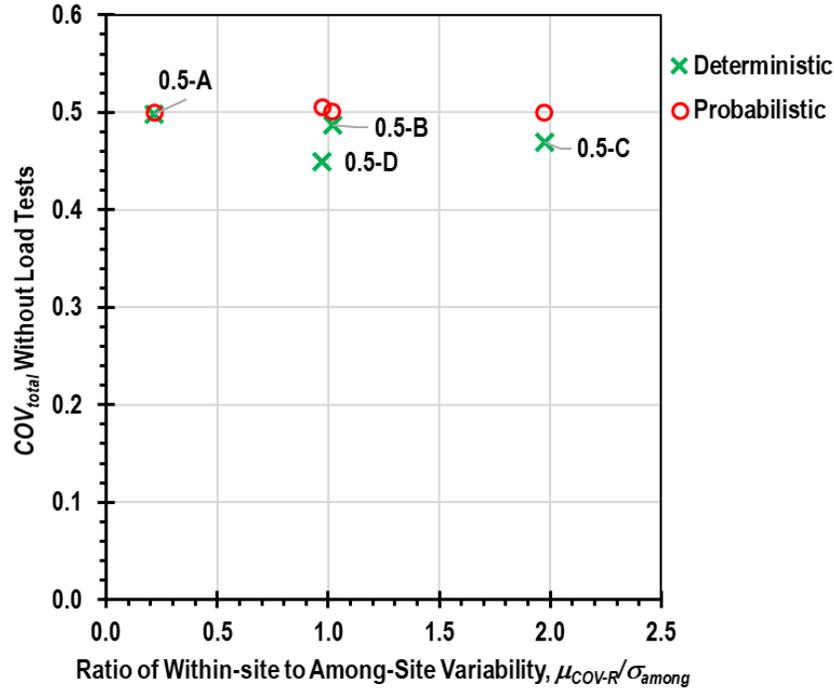


Figure 37: Computed total COV of predictive distributions versus the ratio of within-site to among-site variability.

For perspective on the results, consider case 0.3-D, which has a nominal total COV of 0.3 with approximately equal contributions from within-site and among-site variability and relatively great uncertainty in within-site variability (Table 9). Suppose the predicted capacity at the site (R_p) is 1000 kips. Deterministic within-site variability yields a value of $R_{1/1500} = 0.43$, corresponding 430 kips. Probabilistic within-site variability yields a value of $R_{1/1500} = 0.31$, corresponding to 310 kips. The true values, calculated from site-specific parameters known only in the simulation, are $R_{1/1500} = 0.55$, corresponding to 550 kips. The difference in $R_{1/1500}$ between deterministic and probabilistic within-site variability is 120 kips, which is about 22 percent of the true $R_{1/1500}$, as shown in Figure 35.

The observed relative magnitude of $R_{1/1500}$ without load tests, i.e.

$$R_{1/1500\text{-probabilistic}} < R_{1/1500\text{-deterministic}} < R_{1/1500\text{-true}} \quad \text{Eq. 5-4}$$

is noteworthy. It is important to note the inequality of Eq. 5-4 only holds for the assumption that the true values of average resistance and within-site variability ($\mu_{\text{within-true}}$ and $\sigma_{\text{within-true}}$,

respectively) are equal to the mean values used to define the prior distributions of the twelve cases evaluated (as stated in Eq. 5-2 and Eq. 5-3). The assumption is inherent to all of the analyses reported in this section. The relative magnitude of the three values of $R_{1/1500}$ is explained by consideration of uncertainty in site-specific parameters μ_{within} and σ_{within} . For the true value of $R_{1/1500}$, there is no uncertainty in either quantity ($\sigma_{\mu} = \sigma_{\sigma} = 0$), as explained in Section 5.1.3. For deterministic within-site variability, uncertainty in within-site variability is neglected ($\sigma_{\sigma} = 0$), but uncertainty in the mean is included ($\sigma_{\mu} \neq 0$), and the uncertainty in the mean reduces $R_{1/1500}$ compared to the true value. Probabilistic within-site variability considers uncertainty in both the mean and within-site variability, producing the smallest value of $R_{1/1500}$. It is perhaps tempting to conclude deterministic within-site variability is better than probabilistic within-site variability for cases without load tests by virtue of its estimate of $R_{1/1500}$ being closer to the true value. However, such reasoning neglects the possibility that deterministic within-site variability produces unreliable estimates of $R_{1/1500}$ in some cases because of neglecting uncertainty in within-site variability. Errors in estimates of $R_{1/1500}$ are examined in Chapter 6.

5.2.2 Reliability with a Large Number Load Tests

For cases with a large number of site-specific load tests, it was hypothesized in Chapter 1 that the posterior value of $R_{1/1500}$ would converge to the true value of $R_{1/1500}$ with probabilistic within-site variability but not with deterministic within-site variability. Probabilistic within-site variability treats both the mean and within-site variability of foundation resistance as unknown parameters to be updated via Bayes' theorem. Assuming both parameters converge for large numbers of tests, the posterior predictive distribution of resistance should converge to the true distribution for probabilistic within-site variability. Deterministic within-site variability treats only the mean of foundation resistance as unknown and to be updated; the within-site variability parameter is treated as a constant and is not updated. Accordingly, the posterior predictive distribution of resistance should not converge to the true distribution for deterministic within-site variability, unless the prior estimate of within-site variability happens to be equal to the true value.

The hypothesis is evaluated in this section by comparing the results from the numerical tool described in Chapter 4 for cases with many load tests evaluated with deterministic within-site variability and with probabilistic within-site variability. The cases were outlined in Section 5.1. The cases are derived from the twelve sets of prior distributions summarized in Table 9 and evaluated in Section 5.2.1 for the cases without load tests. However, whereas the evaluation without load tests considered only the fiftieth percentile values of mean foundation resistance and within-site variability (μ_{within} and σ_{within} , respectively), the evaluation in this section for cases with many load tests considers the results from load tests sampled at the tenth, fiftieth, and ninetieth percentiles of both μ_{within} and σ_{within} . Thus, for each of the twelve sets of prior distributions of μ_{within} and σ_{within} , nine combinations of true site parameters were considered, resulting in a total of 108 cases.

As indicated in Table 8 (Section 5.1), cases with as many as 800 load tests were evaluated with the numerical tool. Indeed, for most of the 108 cases evaluated in this section, “a large number of tests” is 800 load tests. However, for some of the cases with 800 tests, overflow occurred in the numerical tool. The consequence of the overflow is negligible: for all cases with overflow, the results of the numerical tool had converged, typically at a significantly smaller number of tests. For the overflow cases, “a large number of tests” refers to 200 or 400 load tests.

Results of the analyses are presented in terms of the ratio of posterior to true values of $R_{1/1500}$. A ratio value of 1 indicates the value of $R_{1/1500}$ from Bayesian updating for many load tests is equal to the true value of $R_{1/1500}$, presumably because the Bayesian distributions of μ_{within} and σ_{within} have converged to the true distributions. Values of the ratio of posterior to true $R_{1/1500}$ less than one indicate the Bayesian results are conservative, producing estimates of $R_{1/1500}$ that are lower than the true value. Values of the ratio of posterior to true $R_{1/1500}$ greater than one indicate the Bayesian results are unconservative, producing estimates of $R_{1/1500}$ that are greater than the true value.

The ratio of posterior to true $R_{1/1500}$ is presented in Figure 38 and Figure 39 for deterministic within-site variability. Figure 38 is a plot of the ratio versus the true mean simulated for each site; Figure 39 is a plot of the ratio versus the true within-site variability (i.e. site standard deviation) simulated for each site. The results support the hypothesis: for most of the cases, the value of the ratio is not equal to one. The plot of Figure 38 indicates a wide distribution of ratio values for each of the three site mean values considered. In fact, the distribution of ratios appears similar, if not identical, for each of the three mean values considered. In addition, for each of the 12 prior distribution cases considered, the ratio of posterior to true $R_{1/1500}$ appears to be independent of site mean for the deterministic case.

The presentation of results in Figure 39, a plot of ratio of posterior to true $R_{1/1500}$ versus within-site variability, suggests the value of true within-site variability relative to the assumed, deterministic value of within-site variability controls the accuracy of $R_{1/1500}$ estimates for deterministic within-site variability. For cases when the deterministic value of within-site variability overestimates the true value, represented by the tenth percentile cases in Figure 39, the values of the ratio of posterior to true $R_{1/1500}$ are all less than one, suggesting conservative results. This is logical: when within-site variability is over-estimated, factored resistance values are likely to be conservative. The converse is also supported by the results: for cases when the deterministic value of within-site variability underestimates the true value (the ninetieth percentile cases in Figure 39), the ratio values are all greater than one, suggesting unconservative results. The results of Figure 39 suggest deterministic within-site variability only produces accurate estimates of $R_{1/1500}$ when the estimated value of within-site variability happens to be equal to the true value.

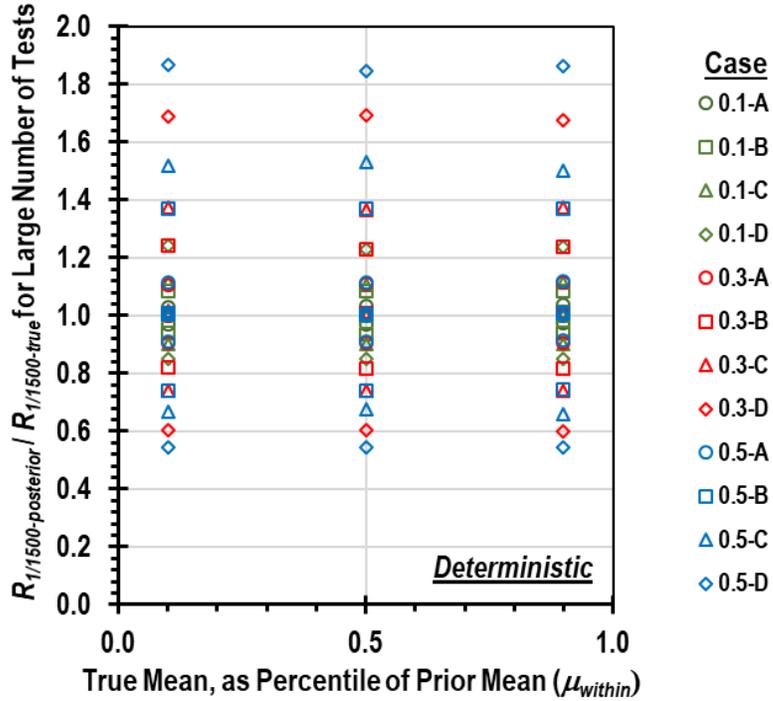


Figure 38: Ratio of posterior to true $R_{1/1500}$ versus true mean for deterministic within-site variability. Case numbers refer to Table 9.

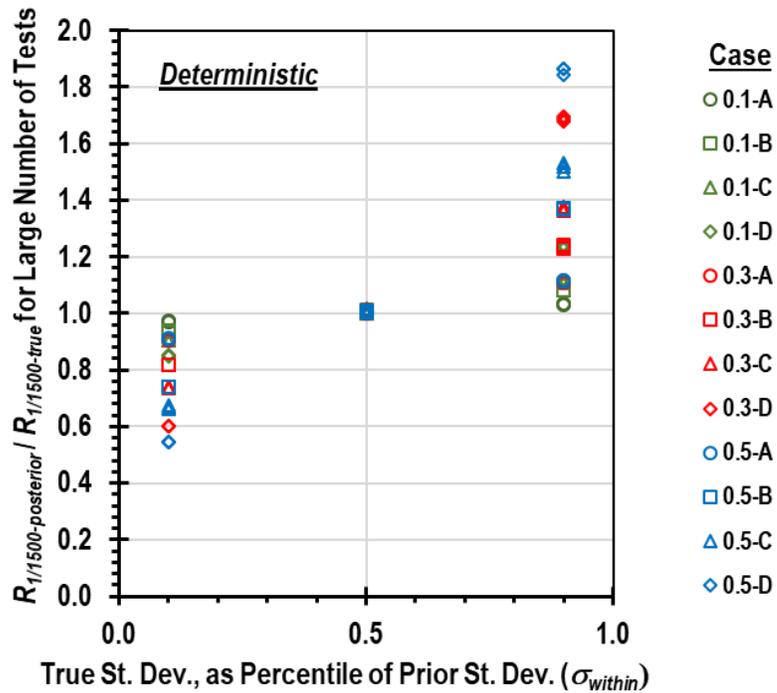


Figure 39: Ratio of posterior to true $R_{1/1500}$ versus true standard deviation for deterministic within-site variability.

The ratio of posterior to true $R_{1/1500}$ is presented in Figure 40 and Figure 41 for probabilistic within-site variability. Figure 40 is a plot of the ratio versus the true mean simulated for each site; Figure 41 is a plot of the ratio versus the true within-site variability simulated for each site. The results shown in both figures also support the hypothesis: for probabilistic within-site variability, all cases had values of the ratio of posterior to true $R_{1/1500}$ of 1.0. The value of the ratio ranged from 0.98 to 1.01. The observation of ratio values of 1.0 is true of all prior distributions of site mean and within-site variability, and of all true values of site parameters considered within the prior distributions. The resulting conclusion is straightforward, but powerful: in the limit, Bayesian updating with probabilistic within-site variability produces accurate estimates of $R_{1/1500}$, regardless of prior distributions of site mean and within-site variability, and regardless of where true site parameters fall with respect to the prior distributions.

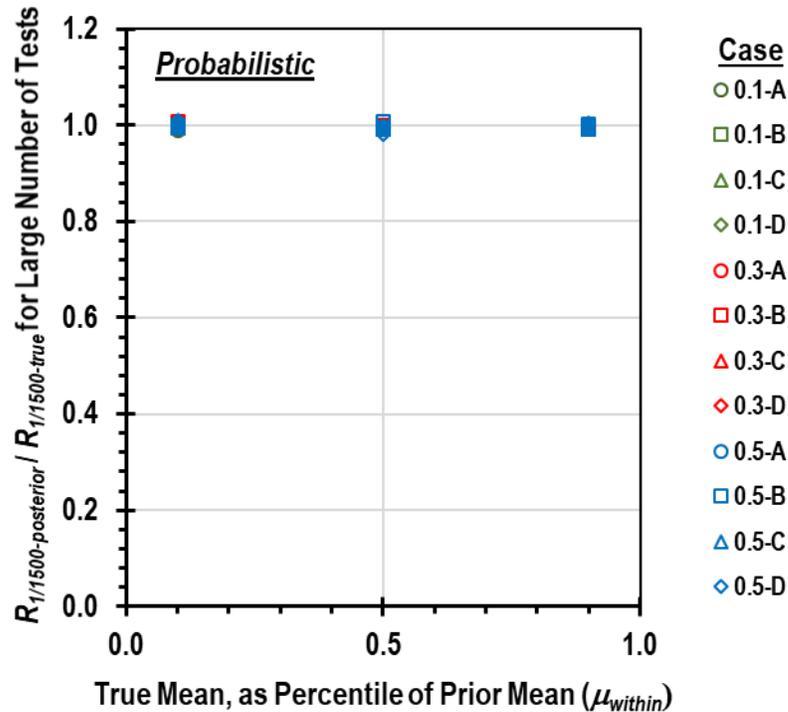


Figure 40: Ratio of posterior to true $R_{1/1500}$ versus true mean for probabilistic within-site variability.

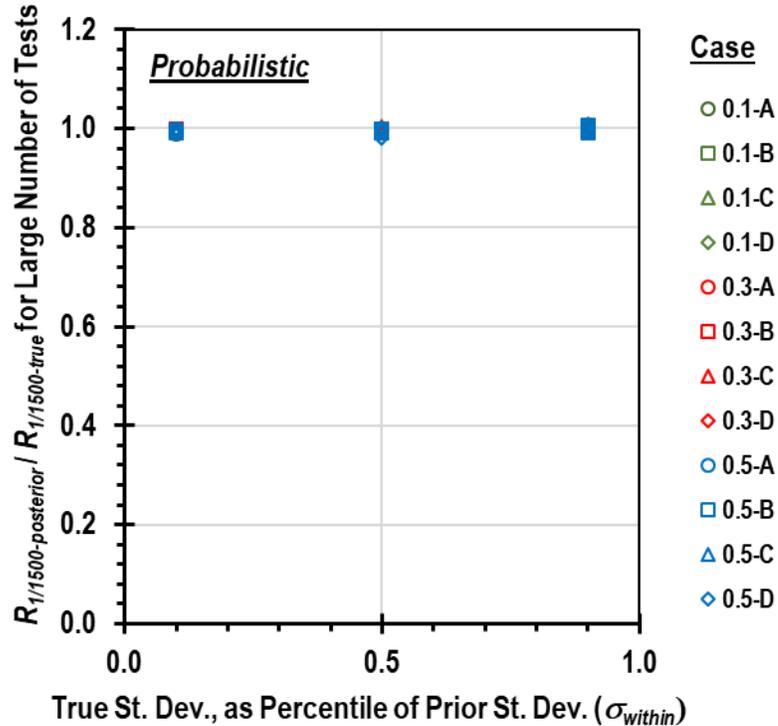


Figure 41: Ratio of posterior to true $R_{1/1500}$ versus true standard deviation for probabilistic within-site variability.

For deterministic within-site variability, the results shown in Figure 39 are strong evidence that the accuracy of reliability estimates with deterministic within-site variability depends on the accuracy of the deterministic estimate of within-site variability. It is also clear from the plot of Figure 39 that the degree of inaccuracy depends on other factors. For instance, for the cases where $R_{1/1500}$ is overestimated (i.e. the cases for which the ratio of posterior to true $R_{1/1500}$ is greater than one, i.e. the cases in Figure 39 for which the true within-site variability is in the ninetieth percentile of the prior), the value of the ratio varies from just greater than 1.0 to about 1.9. To examine this variation, the ratio of posterior to true $R_{1/1500}$ is plotted versus the prior value of mean within-site variability in Figure 42. Results from deterministic and probabilistic within-site variability are presented, with the probabilistic results all falling along a horizontal line corresponding to the ratio of 1.0. For deterministic within-site variability, the deviation from ratio values of 1.0 is greater for greater prior mean values of within-site variability, suggesting both the overestimates and underestimates of $R_{1/1500}$ are worse when within-site variability is greater. Errors in reliability estimating are examined more closely in Chapter 6.

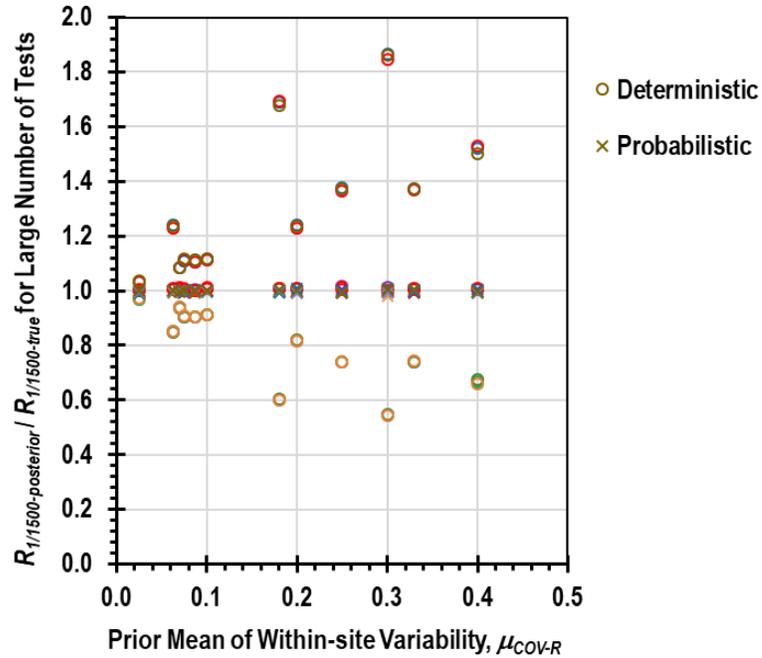


Figure 42: Ratio of posterior to true $R_{1/1500}$ versus within-site variability.

5.3 Trends in Posterior Distributions

Results presented in Section 5.2 supported the hypothesis: (1) without load tests, deterministic within-site variability produces greater estimates of $R_{1/1500}$ than probabilistic within-site variability and (2) with many load tests, the estimate of $R_{1/1500}$ from probabilistic within-site variability converges to the true value while the estimate of $R_{1/1500}$ from deterministic within-site variability generally does not converge to the true value. In this section, the support for the second statement of the hypothesis is examined in greater detail by evaluating trends in the four parameters used to define the posterior distributions of μ_{within} and σ_{within} : μ_μ and σ_μ for the mean and standard deviation of μ_{within} , respectively; μ_σ and σ_σ and for the mean and standard deviation of σ_{within} , respectively. Details regarding how the parameters are interpreted from the Bayesian analyses are presented in Sections 4.2.3 and 5.1.3.

The evaluation methodology used to evaluate trends in the posterior distributions was presented in Section 5.1. In summary, the methodology involves using the numerical tool (Chapter 4) to perform Bayesian updating for 12 sets of prior distributions of μ_{within} and σ_{within} . The 12 sets of distributions were discussed in 5.1.1 and summarized in Table 9. For each of the

12 sets of distributions, load test results are sampled at nine types of sites. The nine types of sites are defined at all combinations of the tenth, fiftieth, and ninetieth percentiles of the distributions of μ_{within} and σ_{within} . As explained in Section 0, “perfect” sampling was used to obtain load test results with sample statistics that match the true values. The analyses were performed with different numbers of load tests (1, 2, 3, 4, 5, 7, 10, 20, 30, 40, 50, 200, 400, 800) to evaluate the trends in the parameters with increasing numbers of tests.

The first subsection below presents results for one of the cases, Case 0.5-B, for the sake of example and to introduce the criteria for interpreting results for other cases. The results for all cases are presented in the following four subsections, with each focusing on a specific posterior parameter: μ_μ , σ_μ , μ_σ and σ_σ , respectively.

5.3.1 Trends versus Number of Tests for Case 0.5-B

For the sake of example, results from the numerical tool for Case 0.5-B are presented in this section. As presented in Table 9, Case 0.5-B has relatively high total variability ($\sigma_{total} = 0.5$) with approximately equal contributions from within-site and among-site variability ($\mu_{COV_R}/\sigma_{among} = 1.0$). The specific site presented here is for true site conditions corresponding to the ninetieth percentile of site mean, i.e. a site with relatively strong foundation resistance, and the tenth percentile of site standard deviation, i.e. a site with relatively uniform foundation resistance. Importantly, the specific values used to describe the trends reported in this section apply only to the specific example case. The sections following this example evaluate general trends for all cases considered.

Plots of posterior μ_μ , σ_μ , μ_σ and σ_σ versus the number of load tests are presented in Figure 43, Figure 44, Figure 45, and Figure 46, respectively. A logarithmic scale is used for the horizontal axis (number of load tests) of each of the plots. In each plot, the prior value of the parameter is shown with a green line, the true value of the parameter is shown with an orange line, and posterior results for deterministic and probabilistic within-site variability are shown with red and blue lines, respectively. Also shown in each figure is the error, defined as the difference between the true value and the prior value of each parameter. As explained in Section 5.1.3, the

true value of σ_μ and σ_σ is zero because for a specific site, there is no variability or uncertainty in the mean or within-site variability parameters. Finally, each figure includes lines to indicate a 75 percent reduction in the error. The solid horizontal black line indicates the parameter value that corresponds to a 75 reduction in the difference between the true value and prior value. The dashed vertical black line indicates the number of tests at which the posterior results have achieved a 75 percent reduction in error.

The plot of μ_μ versus number of tests in Figure 43 indicates results for both deterministic and probabilistic within-site variability converge to the true value of μ_{within} . Both converge at nearly identical rates with respect to the number of tests. Relatively large reductions in error are achieved for small numbers of tests: about half the error is reduced by the first test, and 75 percent is reduced by the third test. The rate of error reduction slows considerably for greater numbers of tests.

The plot of σ_μ versus number of tests in Figure 44 shows decreasing uncertainty in the value of μ_{within} for both deterministic and probabilistic within-site variability. The results mirror the trends in μ_μ shown in Figure 43: as μ_μ converges to the true value of μ_{within} , the corresponding uncertainty in μ_{within} (i.e. σ_μ) decreases. Both deterministic and probabilistic within-site variability produce reductions in σ_μ to approximately zero for large numbers of tests. The rate of reduction is the same for both deterministic and probabilistic within-site variability, with a significant reduction, about one-third of the prior σ_μ , for the first load test and diminishing reductions for subsequent tests.

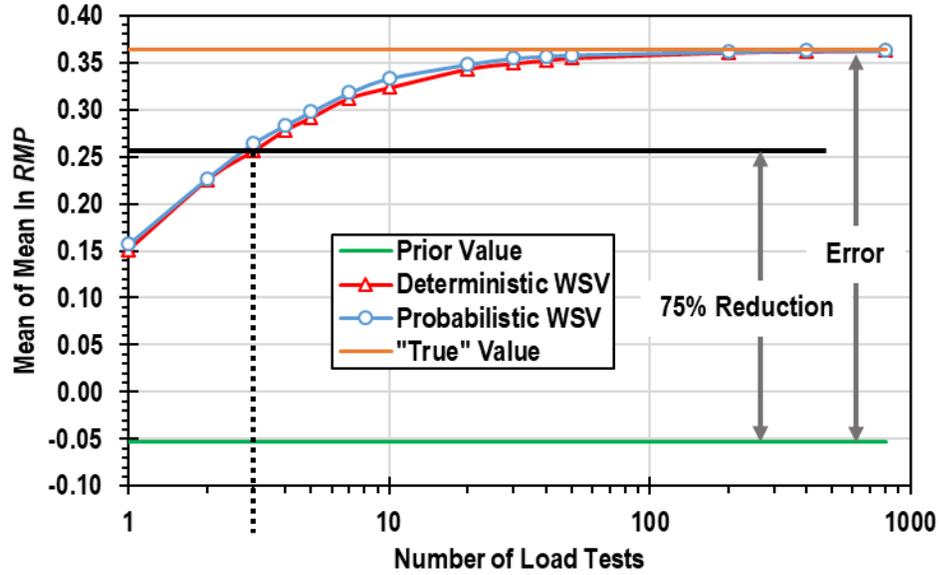


Figure 43: Trend in μ_μ versus number of tests for perfect sampling of Case 0.5-B with true site mean at ninetieth percentile of μ_{within} and true within-site variability at tenth percentile of σ_{within} .

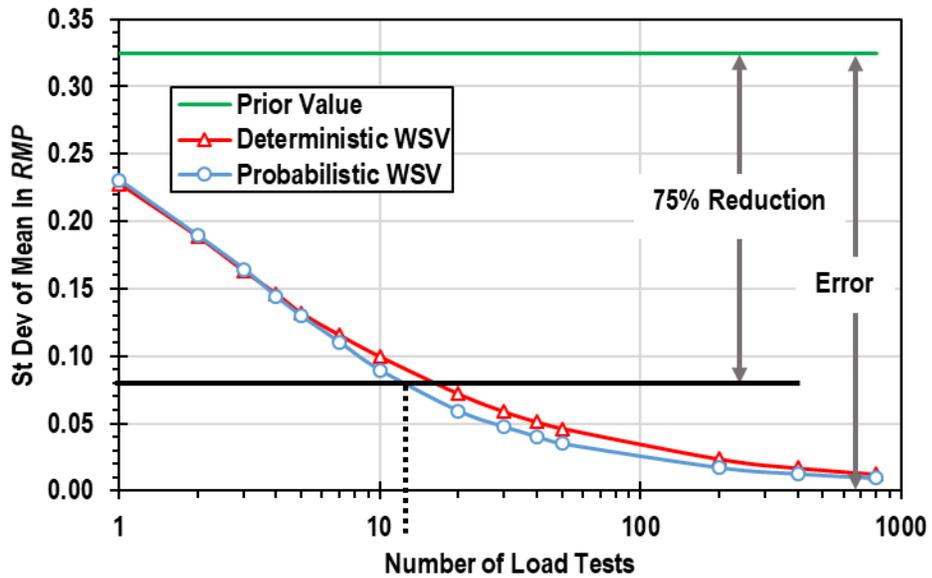


Figure 44: Trend in σ_μ versus number of tests for perfect sampling of Case 0.5-B with true site mean at ninetieth percentile of μ_{within} and true within-site variability at tenth percentile of σ_{within} .

The plot of μ_σ versus the number of load tests presented in Figure 45 shows convergence to the true value of σ_{within} for probabilistic within-site variability. For deterministic within-site variability, σ_{within} is treated as a constant value and is not updated; that deterministic within-site variability does not converge to the true value of σ_{within} is therefore to be expected.

The rate of convergence for probabilistic within-site variability is slow compared to convergence in μ_μ and σ_μ discussed above. The first test produces essentially no reduction in the error, which is logical considering one cannot calculate sample standard deviation with only one test. About 30 load tests are required to achieve a 75 percent reduction in error.

The plot of σ_σ versus number of load tests in Figure 46 shows decreasing uncertainty in the value of σ_{within} with increasing tests for probabilistic within-site variability. Similar to the rate of convergence in μ_σ observed in Figure 45, the rate of decrease in σ_σ in Figure 46 is quite slow compared to the rate of decrease in σ_μ discussed above. More than 100 load tests are required to achieve a 75 percent reduction in the uncertainty in σ_{within} with probabilistic within-site variability. For deterministic within-site variability, σ_{within} is treated as a constant value with zero uncertainty.

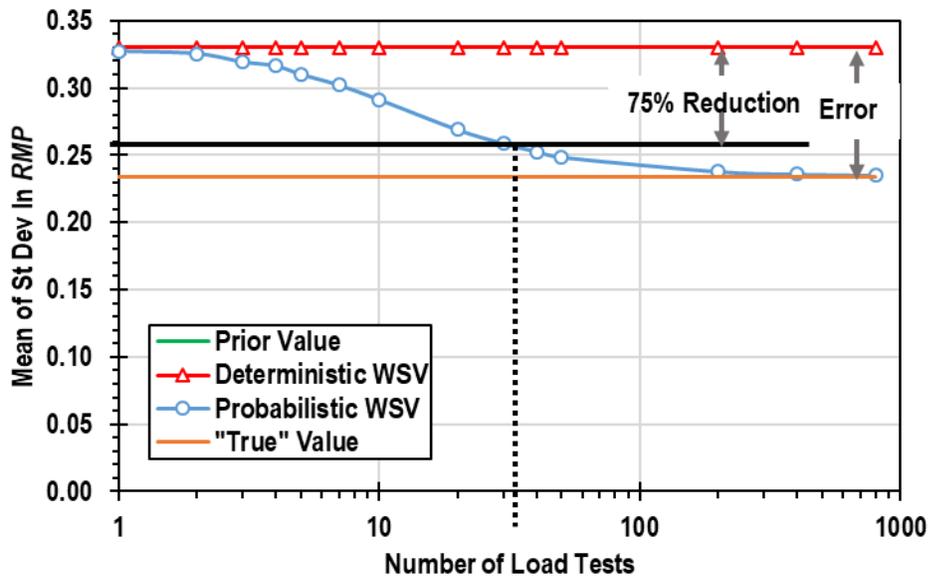


Figure 45: Trend in μ_σ versus number of tests for perfect sampling of Case 0.5-B with true site mean at ninetieth percentile of μ_{within} and true within-site variability at tenth percentile of σ_{within} .

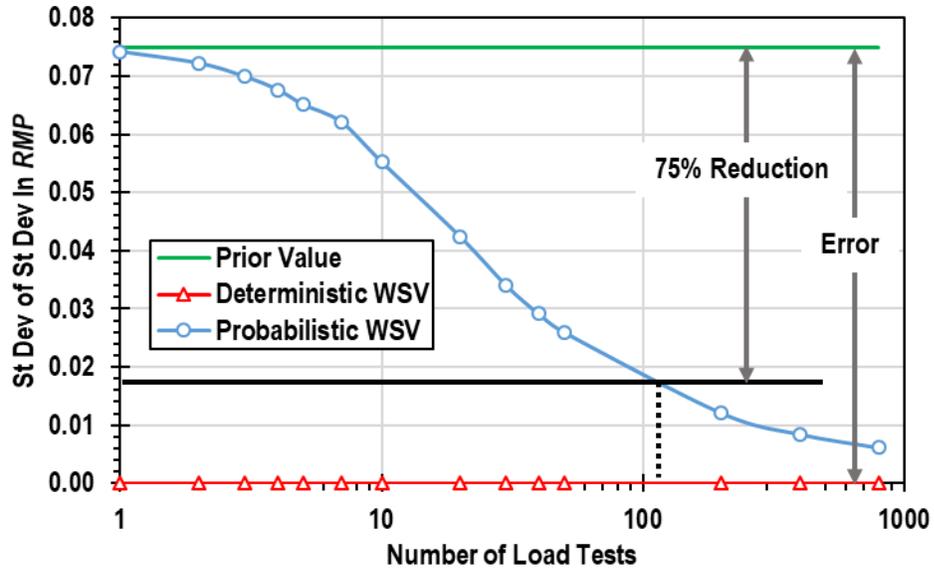


Figure 46: Trend in σ_σ versus number of tests for perfect sampling of Case 0.5-B with true site mean at ninetieth percentile of μ_{within} and true within-site variability at tenth percentile of σ_{within} .

5.3.2 Trends in μ_μ for all Cases

For the example of Case 0.5-B, the mean value of μ_{within} , μ_μ , converged to the true value for large numbers of tests, with most of the error in the prior estimate of μ_μ eliminated for small numbers of tests. To examine convergence and the rate of error reduction for other cases, Figure 47 and Figure 48 are presented. Both figures present results for all cases, with “all cases” indicating all 12 sets of prior distributions of μ_{within} and σ_{within} as well as all nine combinations of true site parameters (tenth, fiftieth, and ninetieth percentiles of μ_{within} and σ_{within}).

Figure 47 is a plot of the error reduction in μ_μ at large numbers of tests for all cases and for both deterministic and probabilistic within-site variability. The error reduction is plotted against the ratio of within-site to among-site variability for the prior distributions of μ_{within} and σ_{within} ($\mu_{COV_R}/\sigma_{among}$). For all cases, the error reduction in μ_μ is at least 94 percent, and for most it is 100 percent. The results indicate both deterministic and probabilistic within-site variability achieve convergence of μ_{within} to the true value for large numbers of tests, regardless of prior conditions or true site conditions.

Figure 48 is a plot of the number of tests required to reduce the error in the posterior value of μ_μ by 75 percent versus the ratio of within-site to among-site variability. Reducing the error by 75 percent is equivalent to achieving convergence to within 25 percent of the true value. Results are shown for all cases, with circles used for deterministic within-site variability results and “x” shapes for probabilistic within-site variability results. The results indicate that for both deterministic and probabilistic within-site variability, the posterior value of μ_μ converges to within 25 percent of the true value for significantly fewer tests when the ratio of within-site to among-site variability is small. For the cases with $\mu_{COV_R}/\sigma_{among} = 0.1$, convergence to within 25 percent of the true value was achieved with the first test; for cases with $\mu_{COV_R}/\sigma_{among} = 2$, convergence to within 25 percent of the true value required 10 to 30 tests.

Results in Figure 48 also indicate differences in the rates of convergence between results for deterministic and probabilistic within-site variability. The rate appears to be the same, on average, but for a given value of $\mu_{COV_R}/\sigma_{among}$, the number of tests required to achieve convergence to within 25 percent of the true value is constant for deterministic within-site variability but variable for probabilistic within-site variability. For probabilistic within-site variability, greater numbers of tests are required when the true within-site variability is high (i.e. for the sites at the ninetieth percentile of σ_{within}) than when the true within-site variability is low (i.e. for sites at the tenth percentile of σ_{within}). When the true within-site variability is average, the number of tests for convergence to within 25 percent of the true value is the same for both deterministic and probabilistic within-site variabilities. Taken together, the results of Figure 48 suggest convergence for deterministic within-site variability is independent of the true within-site variability, depending primarily on the assumed deterministic value of within-site variability. Convergence for probabilistic within-site variability requires greater numbers of tests when the true within-site variability is relatively high compared to when it is relatively low.

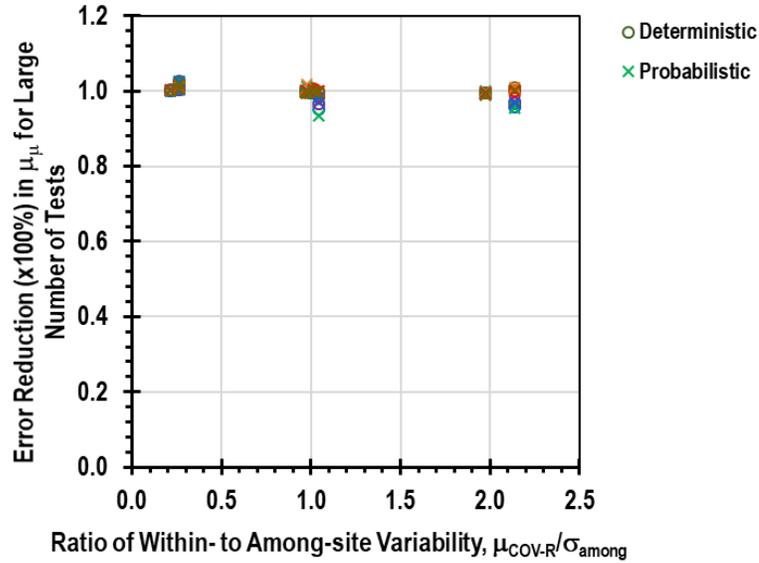


Figure 47: Error reduction in μ_μ for large number of tests versus ratio of within-site to among-site variability for prior distributions of μ_{within} and σ_{within} .

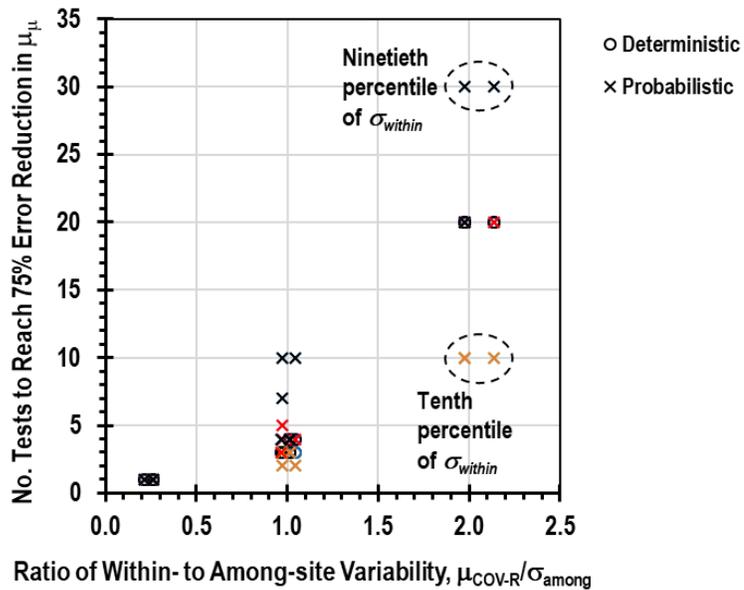


Figure 48: Number of tests to reduce error in the posterior value of μ_μ by 75 percent versus ratio of within-site to among-site variability for prior distributions of μ_{within} and σ_{within} .

5.3.3 Trends in σ_μ for all Cases

For the example of Case 0.5-B, the uncertainty in μ_{within} , σ_μ , converged to zero (the true value, as explained in Section 5.1.3) for large numbers of tests. In addition, the greatest reduction in uncertainty occurred for the first test, with diminishing reductions for greater numbers of tests.

To examine convergence and the rate of error reduction for other cases, Figure 49 and Figure 50 are presented. Both figures present results for all cases, with “all cases” indicating all 12 sets of prior distributions of μ_{within} and σ_{within} as well as all nine combinations of true site parameters (tenth, fiftieth, and ninetieth percentiles of μ_{within} and σ_{within}).

Figure 49 is a plot of the reduction in σ_{μ} at large numbers of tests for all cases and for both deterministic and probabilistic within-site variability. The reduction in uncertainty is plotted against the ratio of within-site to among-site variability for the prior distributions of μ_{within} and σ_{within} ($\mu_{COV_R}/\sigma_{among}$). For all cases, the value of σ_{μ} is reduced by at least 80 percent of its initial value, and for most cases, the reduction in σ_{μ} was at least 95 percent. As shown in Figure 49, the cases with reductions in σ_{μ} of about 80 percent all have high values of the ratio of within-site to among-site variability. For these cases, the uncertainty in μ_{within} is small initially (because among-site variability is small), so in absolute terms, the posterior values of σ_{μ} are quite small. The results indicate both deterministic and probabilistic within-site variability achieve reduction of uncertainty in μ_{within} to nearly zero for large numbers of tests, regardless of prior conditions or true site conditions.

Figure 50 is a plot of the number of tests required to reduce the prior value of σ_{μ} by 75 percent versus the ratio of within-site to among-site variability. Results are shown for all cases, with circles used for deterministic within-site variability results and “x” shapes for probabilistic within-site variability results. Results in Figure 50 regarding the rate of convergence for σ_{μ} are similar to the results presented in Figure 48 and discussed in Section 5.3.2 regarding the rate of convergence of μ_{μ} . The results indicate that for both deterministic and probabilistic within-site variability, the posterior value of σ_{μ} is reduced by 75 percent of the prior value for significantly fewer tests when the ratio of within-site to among-site variability is small. For the cases with $\mu_{COV_R}/\sigma_{among} = 0.1$, convergence to within 25 percent of the true value was achieved with the first or second test; for cases with $\mu_{COV_R}/\sigma_{among} = 2$, convergence to within 25 percent of the true value required 200 tests for most cases.

Differences in the results for σ_μ between deterministic and probabilistic within-site variability are similar to those noted for μ_μ (Section 5.3.2). On average, the number of tests required to achieve 75 percent reduction in σ_μ is the same for both deterministic and probabilistic within-site variability. However, while deterministic within-site variability results are independent of the true within-site variability, probabilistic within-site variability results indicate sites that are less variable (i.e. the tenth percentile sites called out in Figure 50) require fewer tests to achieve 75 percent reduction in σ_μ and sites that are more variable (i.e. the ninetieth percentile sites called out in Figure 50) require more tests to achieve the same reduction.

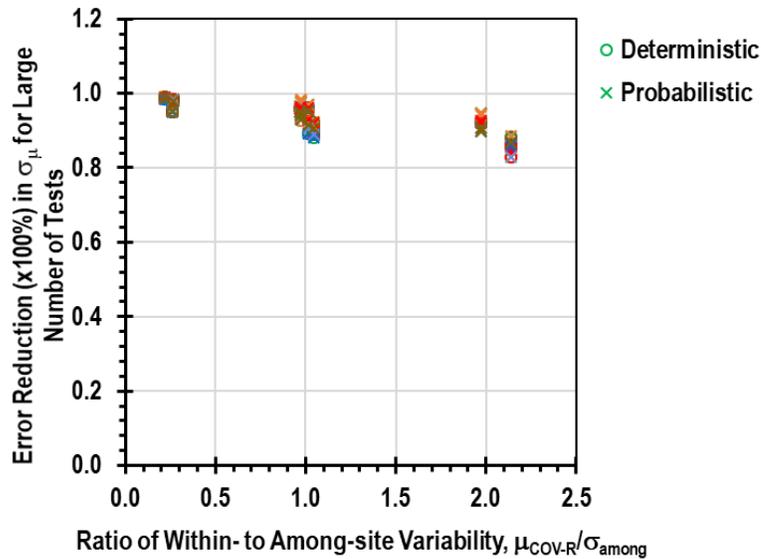


Figure 49: Error reduction in σ_μ for large number of tests versus ratio of within-site to among-site variability for prior distributions of μ_{within} and σ_{within} .

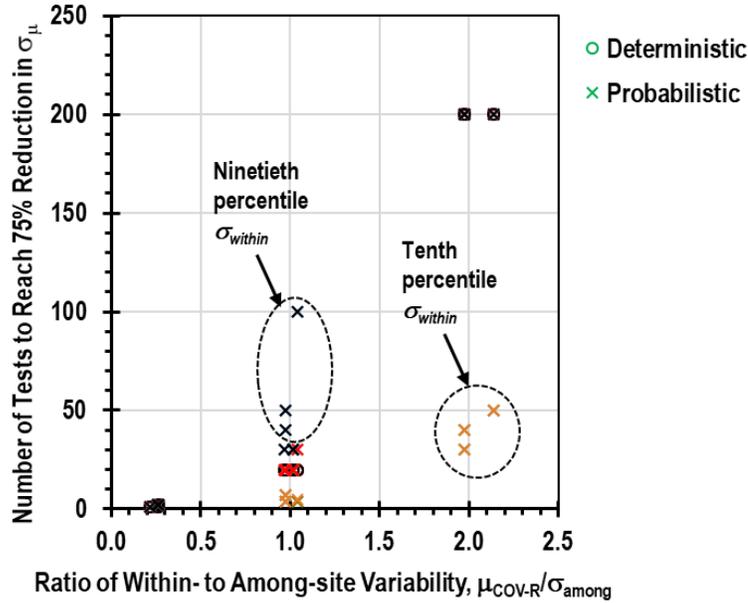


Figure 50: Number of tests to reduce the posterior value of σ_{μ} by 75 percent of its prior value versus ratio of within-site to among-site variability for prior distributions of μ_{within} and σ_{within} .

5.3.4 Trends in μ_{σ} for all Cases

The mean value of σ_{within} , μ_{σ} , is considered a known constant and is not updated for deterministic within-site variability; this is the defining feature of deterministic within-site variability. In contrast, μ_{σ} is treated as an unknown parameter to be updated with probabilistic within-site variability. Because μ_{σ} is not updated for deterministic within-site variability, this section focuses on results for probabilistic within-site variability. For the example of Case 0.5-B, μ_{σ} converged to the true value for large numbers of tests, with convergence requiring more tests than was required for Case 0.5-B for μ_{μ} . To examine convergence and the rate of error reduction for other cases, Figure 51 and Figure 52 are presented. Both figures present results for all cases, with “all cases” indicating all 12 sets of prior distributions of μ_{within} and σ_{within} as well as all nine combinations of true site parameters (tenth, fiftieth, and ninetieth percentiles of μ_{within} and σ_{within}).

Figure 51 is a plot of the error reduction in μ_{σ} at large numbers of tests for all cases with probabilistic within-site variability. The error reduction is plotted against the ratio of within-site to among-site variability for the prior distributions of μ_{within} and σ_{within} ($\mu_{COV-R}/\sigma_{among}$). For all

cases, the error reduction in μ_σ is at least 87 percent, and for most it is 100 percent. The relatively small number of cases with error reduction less than 95 percent are all for prior distributions of σ_{within} with relatively small prior values of μ_σ , so the magnitude of error is relatively small. The results indicate probabilistic within-site variability achieves convergence of σ_{within} to the true value for large numbers of tests, regardless of prior conditions or true site conditions.

The rate of convergence in μ_σ is evaluated with Figure 52, which is a plot of the number of tests required to achieve a 75 percent reduction in error versus the *COV* of the prior distribution of σ_{within} ($\sigma_{COV_R}/\mu_{COV_R}$). The *COV* of σ_{within} represents the uncertainty in σ_{within} , relative to the prior magnitude of within-site variability (μ_{COV_R}). For all cases, the number of tests to achieve convergence to within 25 percent of the true value is 40 or fewer. The plot indicates a strong trend toward convergence at fewer tests as the uncertainty in σ_{within} increases. This is a logical trend: greater uncertainty in a prior estimate should generally be associated with a greater tendency to believe the estimate supported by new evidence.

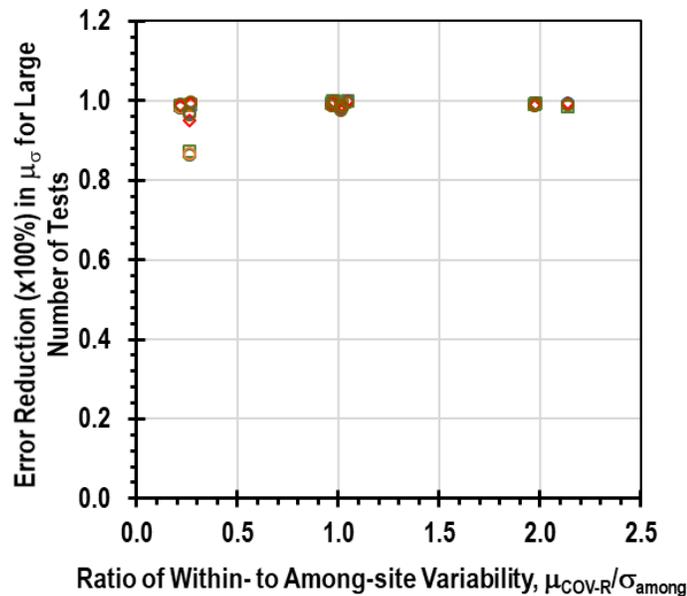


Figure 51: Error reduction in μ_σ for large number of tests versus ratio of within-site to among-site variability for prior distributions of μ_{within} and σ_{within} . Results are for probabilistic within-site variability only.

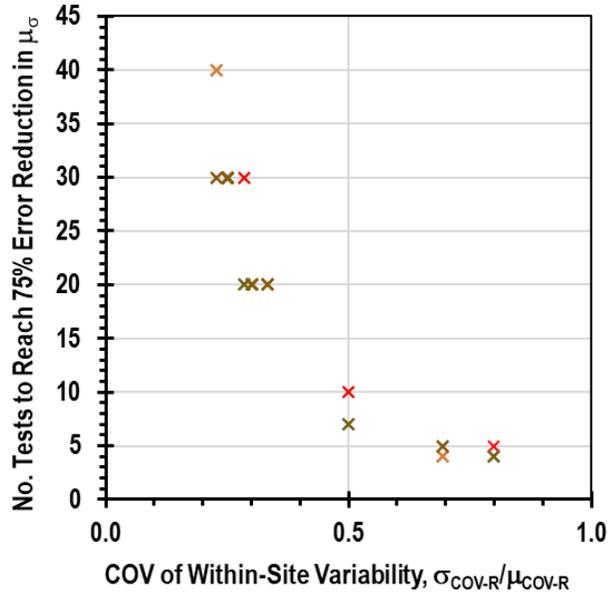


Figure 52: Number of tests to reduce error in the posterior value of μ_{σ} by 75 percent versus uncertainty in within-site variability for prior distribution of σ_{within} . Results are for probabilistic within-site variability only.

5.3.5 Trends in σ_{σ} for all Cases

Results for σ_{σ} , uncertainty in the within-site variability parameter σ_{within} , are similar to those for the mean in σ_{within} , μ_{σ} , presented in Section 5.3.4. As for μ_{σ} , results for σ_{σ} are presented only for probabilistic within-site variability; with deterministic within-site variability, σ_{σ} is zero by definition. For the example of Case 0.5-B, σ_{σ} converged to nearly zero (the true value, as explained in Section 5.1.3) for large numbers of tests. The convergence required many tests: more than 100 to achieve a 75 percent reduction from the prior value. To examine convergence and the rate of error reduction for other cases, Figure 53 and Figure 54 are presented. Both figures present results for all cases, with “all cases” indicating all 12 sets of prior distributions of μ_{within} and σ_{within} as well as all nine combinations of true site parameters (tenth, fiftieth, and ninetieth percentiles of μ_{within} and σ_{within}).

Figure 53 is a plot of the reduction in σ_{σ} at large numbers of tests for all cases with probabilistic within-site variability. The reduction in uncertainty is plotted against the ratio of within-site to among-site variability for the prior distributions of μ_{within} and σ_{within} ($\mu_{\text{COV-R}}/\sigma_{\text{among}}$). For all cases, the value of σ_{σ} is reduced by at least 72 percent of its initial value, and for most

cases, the reduction in σ_μ was at least 85 percent. As shown in Figure 53, the cases with reductions in σ_μ less than 80 percent all have low values of the ratio of within-site to among-site variability. For these cases, the uncertainty in σ_{within} is small initially (because within-site variability is small), so in absolute terms, the posterior values of σ_σ are small. The results indicate that while σ_σ does not always converge to the true value of zero for large numbers of tests, most of the uncertainty in σ_{within} is eliminated, regardless of prior conditions or true site conditions.

Figure 54 is a plot of the number of tests required to reduce the prior value of σ_σ by 75 percent the *COV* of the prior distribution of σ_{within} ($\sigma_{COV_R}/\mu_{COV_R}$). The *COV* of σ_{within} represents the uncertainty in σ_{within} , relative to the prior magnitude of within-site variability (μ_{COV_R}). Results in Figure 54 regarding the rate of convergence for σ_σ are similar to the results presented in Figure 53 and discussed in Section 5.3.4 regarding the rate of convergence of μ_σ . The results indicate that the number of tests required to reduce the posterior value of σ_σ by 75 percent of its prior value is significant, more than 50 for most cases and as many as 400 for some cases. As noted for the results of μ_σ , the plot indicates a strong trend toward convergence at fewer tests as the uncertainty in σ_{within} increases. As explained in Section 5.3.4, this is logical, since greater uncertainty in the prior estimate should result in a greater tendency to believe estimates based on new evidence.

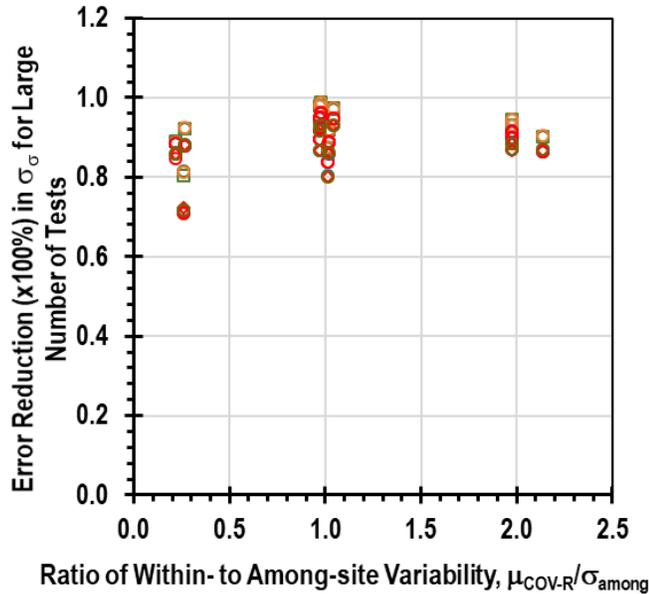


Figure 53: Error reduction in σ_σ for large number of tests versus ratio of within-site to among-site variability for prior distributions of μ_{within} and σ_{within} . Results are for probabilistic within-site variability only.

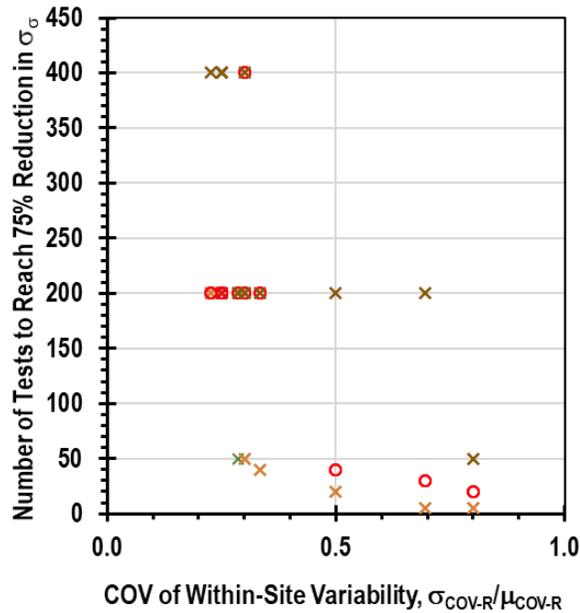


Figure 54: Number of tests to reduce the posterior value of σ_σ by 75 percent of its prior value versus uncertainty in within-site variability for prior distribution of σ_{within} . Results are for probabilistic within-site variability only.

5.3.6 Summary of Posterior Distribution Trends and Support for Hypothesis

A summary of the trends in posterior distribution parameters of μ_{within} and σ_{within} is presented in Table 11. For each of the four parameters considered in this section, the table

presents results in terms of two trends. The first trend is convergence for large numbers of tests: does the parameter converge to the true value for large numbers of tests? The second trend regards the rate of convergence: approximately how many tests are required to achieve a reduction of 75 percent of the error between the prior estimate and the true value of the parameter? For each trend and each parameter, results are presented for both deterministic and probabilistic within-site variability.

For each of the four parameters, results for probabilistic within-site variability converge to the true value at large numbers of tests, as summarized in the first results row of Table 11. For deterministic within-site variability, parameters defining the mean (μ_μ and σ_μ) converge to the true value for large numbers of tests, but parameters defining within-site variability do not. For deterministic within-site variability, within-site variability is treated as a constant parameter and is not updated; by definition, then, the within-site variability will not converge to the true value since it remains unchanged regardless of load test variability. The results summarized in the first results row of Table 11 are consistent with support of the hypothesis: probabilistic within-site variability achieves convergence to the true value of $R_{1/1500}$ because the posterior parameters of μ_{within} and σ_{within} all converge to the true values; deterministic within-site variability does not achieve convergence to the true value of $R_{1/1500}$ because while the posterior parameters of μ_{within} converge to the true values, posterior parameters of σ_{within} do not.

Results regarding the rate of convergence are presented in the second results row of Table 11, which lists the number of tests required to reduce the error in the posterior parameters by 75 percent. A range of values (numbers of tests) is listed for each parameter and for both deterministic and probabilistic within-site variability. The ranges are large, indicating the rate of convergence varies considerably among the different cases. Despite the size of the ranges, it is generally concluded that the mean parameters defining μ_{within} and σ_{within} (μ_μ and μ_σ) converge more quickly than the standard deviation parameters defining μ_{within} and σ_{within} (σ_μ and σ_σ). As noted in the detailed presentation of results earlier in this section, the parameters defining μ_{within} (μ_μ and σ_μ) generally converge more quickly when the ratio of within-site to among-site variability

is small. The parameters defining σ_{within} (μ_σ and σ_σ) generally converge more quickly when the uncertainty in within-site variability is great.

Table 11: Summary of posterior distribution trends.

Trend	μ_{within}				σ_{within}			
	μ_μ		σ_μ		μ_σ		σ_σ	
	D ¹	P ¹	D	P	D	P	D	P
Convergence to True Value for Large Number of Tests	Yes	Yes	Yes	Yes	No	Yes	N/A	Yes
Number of Tests Required to Reduce Error by 75%	1-20 ²	1-30 ²	1-200	1-200	N/A	5-40	N/A	5-400

¹Within-site variability: D = Deterministic; P = Probabilistic

²For μ_μ , the same rate of convergence was found, on average, for deterministic and probabilistic within-site variability.

5.4 Summary

The numerical tool described in Chapter 4 was used to evaluate the hypothesis and trends in posterior distributions of site specific mean, μ_{within} , and within-site variability, σ_{within} . The evaluations were based on a variety of cases. The cases were derived from 12 different sets of prior distributions of μ_{within} and σ_{within} . The sets represent three levels of total variability, three levels of the ratio of within-site to among-site variability, and two levels of uncertainty in within-site variability. For each set of prior distributions of μ_{within} and σ_{within} , nine different “true” site distributions were defined using all combinations of the tenth, fiftieth, and ninetieth percentiles of the distributions of μ_{within} and σ_{within} .

All cases were evaluated with the numerical tool using both deterministic and probabilistic within-site variability to within-site variability. The analyses were performed with different numbers of load tests (0, 1, 2, 3, 4, 5, 7, 10, 20, 30, 40, 50, 200, 400, 800) to evaluate the trends in the parameters with increasing numbers of tests. “Perfect” sampling was used so that the sample mean and sample standard deviation of the load test values were equal to the true values of mean and within-site variability in all cases and for all numbers of tests.

Results of the analyses support the hypothesis statements regarding analyses without load tests and with many load tests. For all analyses without load tests, deterministic within-site variability produces estimates of $R_{1/1500}$ that are greater than estimates with probabilistic within-site variability. Results suggest that $R_{1/1500}$ with deterministic within-site variability is greater because it does not consider uncertainty in within-site variability; neglecting the uncertainty in within-site variability reduces the total variability of the predictive distribution of resistance. The difference between estimates of $R_{1/1500}$ with deterministic within-site variability and $R_{1/1500}$ with probabilistic within-site variability is greater for (1) greater uncertainty in within-site variability and (2) greater values of $R_{1/1500}$.

For analyses with many load tests, results also support the hypothesis: all estimates of $R_{1/1500}$ with probabilistic within-site variability converge to the true value of $R_{1/1500}$, while for deterministic within-site variability, estimates of $R_{1/1500}$ converge to the true value only when the prior estimate of within-site variability happens to be equal to the true value. Errors in estimates of $R_{1/1500}$ for the deterministic case are conservative (i.e. underestimates) when the within-site variability is overestimated and unconservative (i.e. overestimates) when the within-site variability is underestimated. Results also suggest the errors in $R_{1/1500}$ for deterministic within-site variability are greater for greater within-site variability. The reliability implications of the errors in $R_{1/1500}$ for deterministic within-site variability are evaluated in Chapter 6.

Additional support for the hypothesis is provided by the trends in the posterior marginal distributions of μ_{within} and σ_{within} . Probabilistic within-site variability achieves convergence to the true value of $R_{1/1500}$ because the posterior parameters of μ_{within} and σ_{within} all converge to the true values. Deterministic within-site variability does not achieve convergence to the true value of $R_{1/1500}$ because while the posterior parameters of μ_{within} converge to the true values, posterior parameters of σ_{within} do not.

The evaluation of posterior parameter trends also considered the rate of convergence to true values. In general, the mean parameters of both μ_{within} and σ_{within} (i.e. μ_μ and μ_σ) achieve

convergence at fewer numbers of tests than the standard deviation parameters (i.e. σ_μ and σ_σ). Considerable variability was noted in the rate of convergence for each parameter. The parameters defining μ_{within} (μ_μ and σ_μ) generally converge more quickly when the ratio of within-site to among-site variability is small. The parameters defining σ_{within} (μ_σ and σ_σ) generally converge more quickly when the uncertainty in within-site variability is great, which is attributed to greater weight being placed on the new evidence when uncertainty in the prior value is great. Issues pertaining to the rate of convergence, particularly its impact on the reliability of designs based on load tests, are examined in Chapter 6.

6. Effect of Deterministic vs. Probabilistic Within-site Variability for Cases with Random Sampling of Load Test Results

In Chapter 5, the numerical tool was used to evaluate cases with “perfectly” sampled load test results. The results supported the hypothesis: (1) deterministic within-site variability estimates of the 1/1500th value of resistance, $R_{1/1500}$, were greater than estimates based on probabilistic within-site variability for situations without load tests (the prior), and (2) the posterior estimates of $R_{1/1500}$ with probabilistic within-site variability always converged to the true value whereas the estimates of $R_{1/1500}$ for deterministic within-site variability were only correct when the prior estimate of within-site variability happened to be correct.

In this chapter, the numerical tool is used to evaluate cases with randomly sampled load test results. Randomly sampled load tests are more realistic than perfectly sampled load tests. Trials with random sampling are therefore more appropriate for evaluation of the magnitude of errors that can be expected from the Bayesian analysis. Practically, the most meaningful errors are errors in the estimate of $R_{1/1500}$ compared to the true value of $R_{1/1500}$ since $R_{1/1500}$ is akin to the factored or allowable resistance. Errors in $R_{1/1500}$ are therefore a primary focus of the chapter, with a particular focus on how errors resulting from deterministic within-site variability compare with errors from probabilistic within-site variability. Errors in other parameters, specifically within-site mean and within-site variability, are also considered, especially in the context of explaining the errors in $R_{1/1500}$. Random sampling is also used to examine the difference between posterior estimates of $R_{1/1500}$ and prior estimates of $R_{1/1500}$. The difference is an important practical measure of the reliability benefits associated with load test information.

The first section of this chapter presents the evaluation methodology. The methodology is largely similar to that for perfect sampling from Chapter 5, so the discussion primarily focuses on differences between the two methodologies. The second section of the chapter presents detailed results for Case 0.5-B, the case examined in detail in Chapter 5. Comparison of the results for

perfect sampling with those for random sampling is used to evaluate the effect of random sampling. Results for Case 0.5-B are also used to provide a detailed analysis of errors in estimates of $R_{1/1500}$. The analysis of errors is generalized and applied to all analysis cases in the third section of the chapter, which includes a subsection that categorizes outcomes for different cases as either satisfactory, overreliable, or underreliable. The fourth section of the chapter presents an analysis of the difference between posterior and prior estimates of $R_{1/1500}$ to evaluate the reliability benefits associated with load test information. The chapter closes with a summary that includes detailed findings.

6.1 Evaluation Methodology

As noted in the introduction to this chapter, the methodology used for the analyses of randomly sampled load test data follows closely the methodology used the analyses of perfectly sampled load test data from Chapter 5 (presented in detail in Section 5.1). The methodology employed to evaluate trials with random sampling of load tests is depicted in Figure 55. As listed atop the figure, the methodology includes three major steps: (1) simulate load test data from the prior distributions of mean and standard deviation of resistance, (2) use load test data to update the prior distributions, and (3) interpret reliability from the posterior distributions. Details of the second and third steps were provided in Chapter 4 (in Sections 4.2 and 4.3, respectively); this section focuses on the first step.

The procedure used to simulate random load test results was designed to evaluate $R_{1/1500}$ under a range of conditions: site conditions, foundation design conditions, and load test result conditions. The parameters used to define the various site conditions are defined in Table 12. The parameters are divided into two categories: parameters used to define the prior distributions of within-site mean, μ_{within} , and within-site variability, σ_{within} , and parameters used to define the “true” site conditions. The last row in the table, number of load tests, is the final parameter that was varied among the simulations. One subsection below is dedicated to each of the two categories of parameters. Another subsection presents information regarding random sampling of load test results. The final subsection discusses how the results of analyses are

interpreted. Additional details of the interpretations are included with the presentation of results throughout the chapter.

Importantly, the random sampling routines employed throughout the methodology are all unbiased. For both the generation of sites by random sampling (Section 6.1.2) and the random sampling of load test values from the site distributions (Section 6.1.3), randomly sampled values conform to the true distributions; in other words, for large samples, the sample mean and sample standard deviation will match the true values. The implications of the unbiased sampling are discussed in Section 6.1.2.

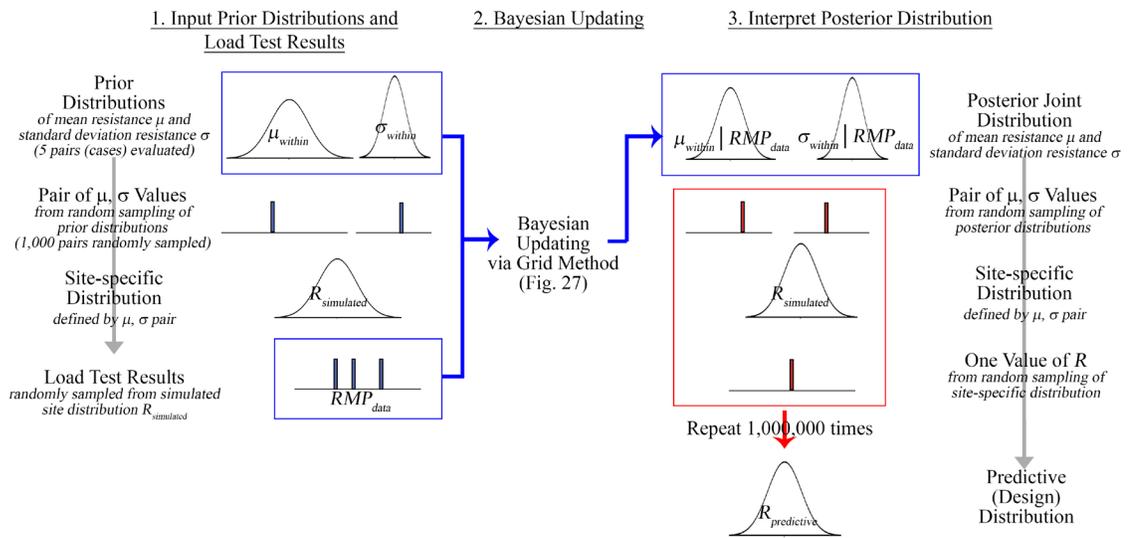


Figure 55: Overview of methodology used to evaluate trials with randomly sampled load tests via the numerical tool.

Table 12: Summary of input parameters used to generate randomly sampled load test results.

Parameter		Input Levels	Values
Prior Distributions of μ_{within} and σ_{within}	Total variability of resistance, σ_{total}	3	0.1 (Figure 30) 0.3 (Figure 31) 0.5 (Figure 32) All three shown together in Figure 56
	Ratio of within-site to among-site variability, $\mu_{COV_R}/\sigma_{among}$	3	Approximately 0.2 (Case 0.5-A) 1 (Case 0.1-B, Case 0.3-B, Case 0.5-B) 2.0 (Case 0.5-C);
"True" Site Condition	μ_{within}	N/A	Randomly sampled from prior distribution of μ_{within}
	σ_{within}	N/A	Randomly sampled from prior distribution of σ_{within}
Number of Load Tests		8	0, 1, 2, 3, 5, 10, 50, 200

6.1.1 Parameters Defining Prior Distributions

In Chapter 5, Section 5.1.1 is a detailed discussion of the prior distributions of μ_{within} and σ_{within} , summarizing the importance of the parameters and explaining how 12 cases from the prior distributions were defined as the basis for evaluating results. The cases were defined from combinations of different levels of three parameters: nominal prior predictive COV , σ_{total} , the ratio of within-site to among-site variability, $\mu_{COV_R}/\sigma_{among}$, and the uncertainty of within-site variability. For the evaluation using randomly sampled load tests in this chapter, five of the 12 cases from Chapter 5 were selected: Case 0.1-B, Case 0.3-B, Case 0.5-A, Case 0.5-B, and Case 0.5-C. The cases are shown in the plot of among-site variability versus within-site variability in Figure 56, and details of the case parameters are presented in Table 13. Note that the plot of Figure 56 includes the same lines of nominal prior predictive COV as presented separately in Figure 30, Figure 31, and Figure 32 in Chapter 5; the combined presentation is useful for perspective on the magnitude of variability. All three levels of nominal prior predictive COV are represented by approximately circular lines centered about the origin, with the smaller COV values considerably closer to the origin than the greater values.

The five cases were selected to focus evaluations on two factors: the nominal prior predictive COV and the ratio of within-site to among-site variability. Recall from Chapter 5 (Section 5.1.1 and Section 5.2.1) that the term nominal is used to indicate that the target value of prior predictive COV is only perfectly satisfied for probabilistic within-site variability; for deterministic within-site variability, prior predictive COV is less than the nominal value because of not considering uncertainty in within-site variability. Evaluation of the effects of nominal prior predictive COV is facilitated by comparison of results from Cases 0.1-B, 0.3-B, and 0.5-B, which have nominal values of prior predictive COV of 0.1, 0.3, and 0.5, respectively, and which all have a value of the ratio of within-site to among-site variability of 1 and uncertainty in the magnitude of within-site variability corresponding to a COV of about 0.25. Evaluation of the effects of the ratio of within-site to among-site variability is facilitated by comparison of results from Cases 0.5-A, 0.5-B,

and 0.5-C, which have values of the ratio of within-site to among-site variability of 0.2, 1, and 2, respectively, and which all have a nominal value of prior predictive COV of 0.5.

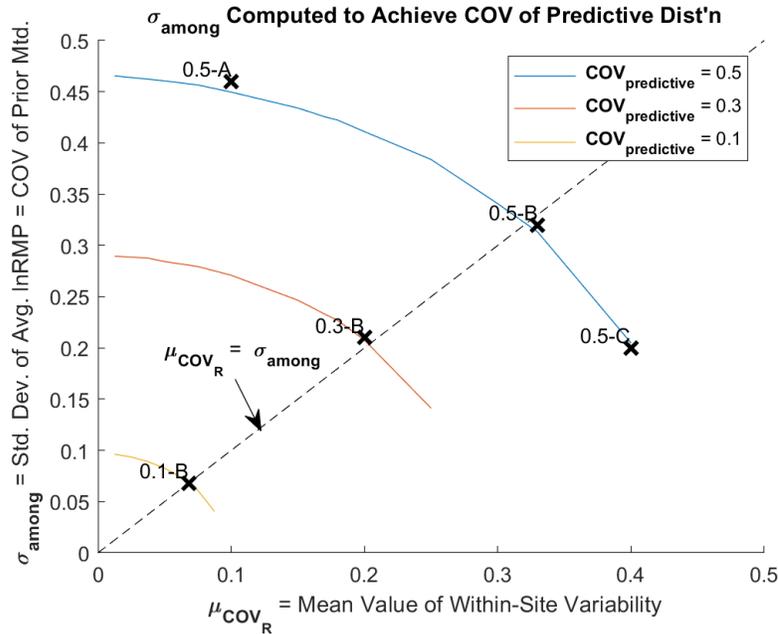


Figure 56: Among-site variability versus within-site variability, showing five cases analyzed for random sampling analysis.

Table 13: Summary of 5 sets of prior distribution parameters used for random sampling analysis.

Case	$\sigma_{total} \approx COV_R$	$\mu_{COV_R} / \sigma_{among}$	$\sigma_{COV_R} / \mu_{COV_R}$	μ_{within}		σ_{within}	
				μ_{among}	σ_{among}	μ_{COV_R}	$\sigma_{COV_R}^1$
0.1-B	0.1	1.01	0.25	-0.002	0.07	0.07	0.018
0.3-B	0.3	0.97	0.25	-0.021	0.21	0.20	0.05
0.5-A	0.5	0.22	0.25	-0.11	0.46	0.10	0.025
0.5-B	0.5	1.02	0.23	-0.053	0.32	0.33	0.075
0.5-C	0.5	1.97	0.25	-0.021	0.20	0.40	0.10

¹Values of σ_{COV_R} listed are for probabilistic within-site variability. For deterministic within-site variability, $\sigma_{COV_R} = 0$.

6.1.2 Generation of Sites by Random Sampling

For the perfect sampling analysis of Chapter 5, true site distributions were defined using all nine combinations of tenth, fiftieth, and ninetieth percentiles of the prior distributions of μ_{within} and σ_{within} (with a set of prior distributions defined for each of the 12 cases). For the random sampling analysis used in this chapter, the true site parameters were sampled randomly from the

prior distributions, as shown in the left pane of Figure 55. For each of the five cases, 1,000 sites were generated by randomly sampling values of μ_{within} and σ_{within} from the prior distributions discussed in Section 6.1.2 and defined in Table 13. Each site is defined by one randomly sampled value of μ_{within} and one randomly sampled value of σ_{within} . The pair of mean and standard deviation values are the “true” site parameters, which are used to define the true distribution of resistance for the site. Importantly, the true value of $R_{1/1500}$ is calculated as the 1/1500th value of the true distribution of resistance for each site. The approach for generating sites by random sampling is similar to the procedure used in the numerical tool to evaluate reliability from posterior distributions of μ_{within} and σ_{within} . The numerical tool procedure is presented in Section 4.3.

As noted in the introduction to this section, the random sampling of site parameters is unbiased; the collection of true site parameters conform to the prior distributions of μ_{within} and σ_{within} . The implication is that errors in the site parameters result only from the variability and uncertainty in the parameters about the mean of the prior distributions and *not* from errors in the prior distribution parameters, themselves. The effect of errors in the prior distribution is not considered here.

6.1.3 Randomly Sampled Test Results

For each of the 1,000 sites generated by random sampling, load test results were randomly sampled from the true distribution of resistance for the site. The procedure for establishing the true distribution of resistance was presented in Section 6.1.2. The random sampling of load test results is depicted in the left pane of Figure 55. Throughout this chapter, evaluations are based on seven levels of the number of load tests: 1, 2, 3, 5, 10, 50, 200. Accordingly, sets of load test data with 1, 2, 3, 5, 10, 50, and 200 resistance values were randomly sampled for each of the 1,000 sites from each of the five cases.

6.1.4 Parameters Used for Evaluation of Results

Interpretation of results from the analysis of randomly sampled load test data is primarily based on two ratio values involving $R_{1/1500}$, the 1/1500th value of the resistance distribution. The

first ratio, the ratio of estimated to true $R_{1/1500}$, is defined mathematically in Eq. 6-1. The ratio is used to evaluate errors in the estimated values of $R_{1/1500}$. Values of the ratio of estimated to true $R_{1/1500}$ less than one indicate the estimate of $R_{1/1500}$ is less than the true value, a conservative outcome. Values of the ratio greater than 1 indicate the estimate of $R_{1/1500}$ is greater than the true value, an unconservative outcome.

$$\frac{(R_{1/1500-estimated})}{R_{1/1500-true}}, \text{ where} \quad \text{Eq. 6-1}$$

$R_{1/1500-estimated}$ are estimates of the 1/1500th value of resistance from either the posterior or prior predictive distributions of resistance, and $R_{1/1500-true}$ are the true values of the 1/1500th value of resistance from the true distribution of resistance as defined in Section 6.1.2.

The second ratio, the ratio of posterior to true $R_{1/1500}$, is defined mathematically in Eq. 6-2. Since the posterior estimates vary within the cases (because of differences among the simulated sites as well as because of sampling load test values within each simulated site), the sample mean is used to represent an average value of posterior resistance. The ratio of Eq. 6-2, then, represents the average increase in resistance one might expect to achieve by load testing, maintaining satisfaction of $p_f = 1/1500$ (without consideration of load variability) and expressed as a factor of the prior estimate.

$$\frac{\text{Mean}(R_{1/1500-posterior})}{R_{1/1500-prior}}, \text{ where} \quad \text{Eq. 6-2}$$

$R_{1/1500-prior}$ are prior estimates of the 1/1500th value of resistance
 $R_{1/1500-posterior}$ are posterior estimates of the 1/1500th value of resistance, and
 $\text{Mean}(R_{1/1500-posterior})$ is the sample mean of posterior values of $R_{1/1500}$.

Details of how the two ratio quantities are interpreted are presented throughout the chapter. Specifically, criteria for categorizing values of the ratio of estimated to true $R_{1/1500}$ are presented in the analysis of errors in Section 6.3.2.

6.2 Results for Case 0.5-B

In the previous chapter, results for perfect sampling of Case 0.5-B were presented for the sake of example (Section 5.3.1), prior to presenting average results for all cases (Sections 5.3.2 through 5.3.6). In this section, a similar approach is adopted for random sampling: results for random sampling of Case 0.5-B are presented for the sake of example and to introduce the methodology used to evaluate the cases collectively in the rest of the chapter. First, the convergence plots for the four posterior parameters (μ_μ , σ_μ , μ_σ and σ_σ) are presented, updating the results from Section 5.3.1; second, histograms of the ratio of estimated to true value of $R_{1/1500}$ are presented.

6.2.1 Convergence of Posterior Distribution Parameters

In Section 5.3.1, plots of μ_μ , σ_μ , μ_σ , and σ_σ versus number of load tests were presented for Case 0.5-B with perfect sampling (Figure 43 through Figure 46, respectively). To evaluate the effect of random sampling on posterior parameter convergence, the same parameters were interpreted from the random sampling analysis of Case 0.5-B. The analysis of Section 5.3.1 considered results for a site with the true mean at the ninetieth percentile of the prior distribution of μ_{within} and the true within-site variability at the tenth percentile of the prior distribution of σ_{within} . To be consistent, the results presented in this section for random sampling are from one of the 1,000 randomly sampled sites (Section 6.1.2) with true site parameters at the same relative values with respect to the prior distributions. Results are presented in Figure 57 (μ_μ versus number of tests), Figure 58 (σ_μ versus number of tests), Figure 59 (μ_σ versus number of tests), and Figure 60 (σ_σ versus number of tests). The figures are updates of Section 5.3.1 figures; all plots include results from perfect sampling in solid lines and random sampling in dashed lines.

Results indicate the effect of random versus perfect sampling on the interpreted parameters is relatively limited, at least for Case 0.5-B. The biggest difference between parameters interpreted from perfect versus random sampling was found for the mean of site-specific mean (μ_μ in Figure 57). For small numbers of tests, five or fewer, the interpreted mean for random sampling is greater than that for perfect sampling. The posterior μ_μ from random sampling is higher than posterior μ_μ from perfect sampling primarily because the randomly sampled resistance for the first load test, corresponding to $\ln RMP$ of 0.62, is significantly greater than the true mean, which corresponds to $\ln RMP$ of 0.36. For all numbers of tests, however, the posterior value of μ_μ never exceeds the true value.

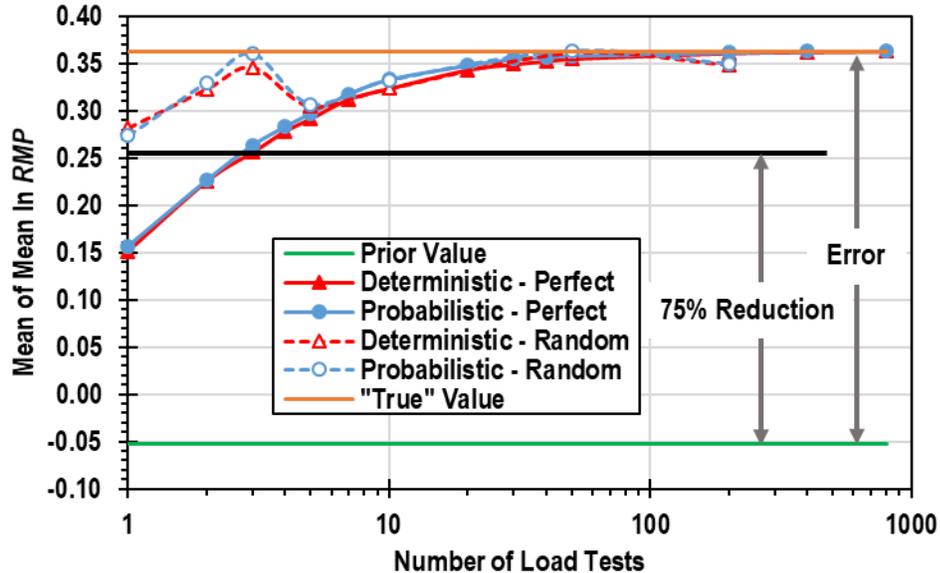


Figure 57: Trend in μ_μ versus number of tests for perfect and random sampling of Case 0.5-B with true site mean at ninetieth percentile of μ_{within} and true within-site variability at tenth percentile of σ_{within} .

For the other parameters, the effect of perfect versus random sampling is limited for Case 0.5-B. For σ_μ , the interpretation based on random sampling is virtually indistinguishable from the interpretation based on perfect sampling. For the deterministic case, this is to be expected: the value of σ_μ for deterministic within-site variability depends only on the value of within-site variability (an assumed constant), the prior value of σ_μ (i.e. σ_{among} , or among-site variability), and the number of tests, based on Eq. 2-9 from Zhang and Tang (2002) (Section 2.3.2). For

probabilistic within-site variability, the posterior value of σ_μ for the randomly sampled interpretation is similar to but slightly higher than the values from the perfect sampling interpretation. It is possible uncertainty in the within-site mean is slower to converge to zero for random sampling because of the additional variability in the sampled load tests, but the effect appears limited. For the within-site variability parameters μ_σ and σ_σ , the effect of random versus perfect sampling is also limited. Parameters interpreted from random sampling show more noise, alternating between being greater than and being less than parameters interpreted from perfect sampling but conforming to the same trends.

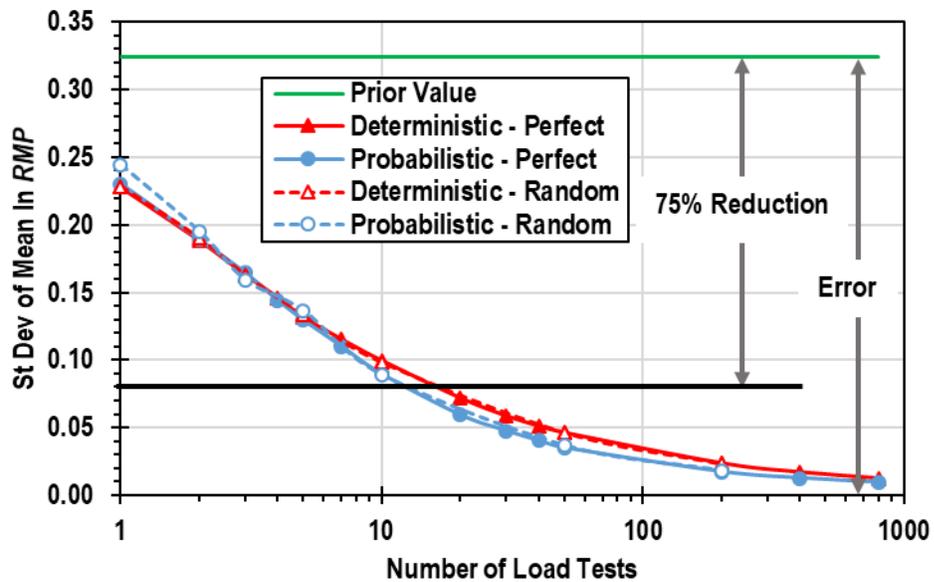


Figure 58: Trend in σ_μ versus number of tests for perfect and random sampling of Case 0.5-B with true site mean at ninetieth percentile of μ_{within} and true within-site variability at tenth percentile of σ_{within} .

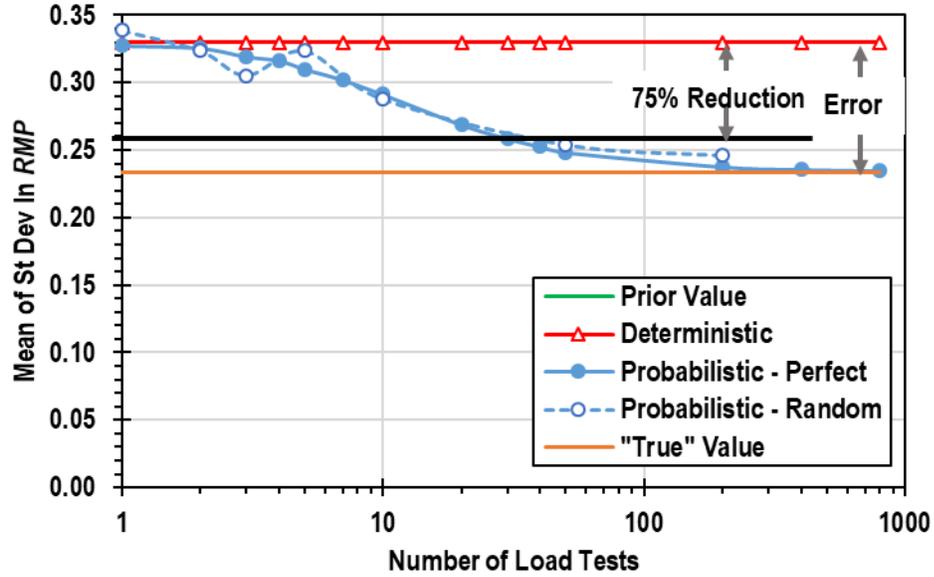


Figure 59: Trend in μ_σ versus number of tests for perfect and random sampling of Case 0.5-B with true site mean at ninetieth percentile of μ_{within} and true within-site variability at tenth percentile of σ_{within} .

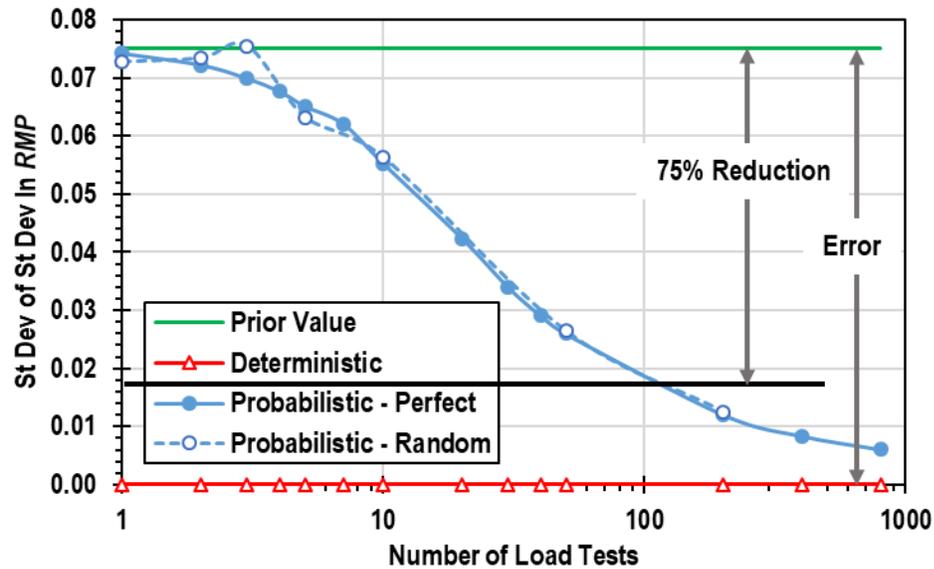


Figure 60: Trend in σ_σ versus number of tests for perfect and random sampling of Case 0.5-B with true site mean at ninetieth percentile of μ_{within} and true within-site variability at tenth percentile of σ_{within} .

It is important to note, again, that the trends documented in this section apply only to Case 0.5-B. Nonetheless, the results provide some evidence that the effect of random versus perfect sampling is limited primarily to μ_μ . The similarity of results between random and perfect

sampling analyses also provides reassurance that the evaluation methodologies employed for the two sets of analyses are consistent with one another.

6.2.2 Ratio of Estimated to true Value of $R_{1/1500}$

As explained in Section 6.1.4, the ratio of estimated to true $R_{1/1500}$ is used to evaluate the accuracy and precision of estimates of $R_{1/1500}$. The ratio was defined in Eq. 6-1. When the ratio of estimated to true $R_{1/1500}$ is exactly 1, the estimate of $R_{1/1500}$ is equal to the true value; the estimate is perfectly accurate. Values of the ratio less than 1 represent conservative estimates of $R_{1/1500}$, and values greater than 1 represent unconservative estimates. The further the value of the ratio is from 1 in either direction, the less accurate the estimate of $R_{1/1500}$ is. For a collection of values of the ratio, it is also possible to evaluate the precision of the estimates by considering the variability of the ratio values. Less variable collections of ratio values indicate the estimates of $R_{1/1500}$ are more precise; more variable collections of ratio values indicate less precision.

For Case 0.5-B, the ratio of estimated to true $R_{1/1500}$ was computed for each of the 1,000 simulated sites, for both deterministic and probabilistic within-site variability, and for all numbers of tests evaluated: 0 (the prior case), 1, 2, 3, 5, 10, 50, and 200. For each of the eight numbers of tests evaluated, two histograms of the ratio of estimated to true $R_{1/1500}$ are presented in Figure 61, one for deterministic within-site variability and one for probabilistic within-site variability. A lognormal distribution was fit to each set of 1,000 values; the lognormal distributions are represented by the curves in Figure 61, and the lognormal parameters associated with each curve are listed in textboxes next to the histograms. Also listed in the textboxes are the sample mean and sample standard deviation for each set of 1,000 values of the ratio. The sample mean and sample standard deviation are the parameters that would be used to define a normal distribution for each dataset. The lognormal distribution fits the ratio data better than the normal distribution, but the normal parameters are clearer for interpretation. Consistent with the previous paragraph, the value of the mean is interpreted to represent accuracy, and the value of the standard deviation is interpreted to represent precision. Values of the mean closer to one indicate greater accuracy, and smaller values of standard deviation indicate greater precision.

For the prior case without load test information (Figure 61(a)), estimates of $R_{1/1500}$ are highly variable and, on average, conservative. For both deterministic and probabilistic within-site variability, the *COV* of the ratio of estimated to true $R_{1/1500}$ is about 0.4, indicating considerable variability and low precision in the estimates. The mean value of the ratio is 0.71 for the deterministic case and 0.65 for the probabilistic case; the deterministic case is slightly less conservative. That prior estimates of $R_{1/1500}$ are less conservative for deterministic within-site variability is consistent with deterministic within-site variability neglecting uncertainty in the within-site variability, as discussed in Chapter 5 (Section 5.2.1).

As load test information is considered, estimates of $R_{1/1500}$ become less conservatively biased and more accurate. This is evident from the histograms in Figure 61, which shift to the right and toward ratio values of 1.0 as the number of load tests increases. The shift toward a ratio of estimated to true $R_{1/1500}$ is noted for both deterministic and probabilistic within-site variability. The trend of increasing value of the ratio of estimated to true $R_{1/1500}$ is also shown in Figure 62, a plot of sample mean versus the number of tests, with the number of tests plotted on a logarithmic scale. For both deterministic and probabilistic within-site variability, the mean value of the ratio increases with the number of tests, and the mean value is approximately 1.0 for the posterior estimates of $R_{1/1500}$ based on 200 load tests. The mean for deterministic within-site variability is greater than the mean for probabilistic within-site variability for all numbers of tests, although the difference diminishes with increasing numbers of tests. The difference is again likely due to deterministic within-site variability neglecting uncertainty in the within-site variability.

As additional tests are considered, there is less similarity between outcomes from deterministic and probabilistic within-site variability. As noted previously, estimates of $R_{1/1500}$ for deterministic and probabilistic within-site variability are similarly variable and imprecise for the prior case, with *COV* of ratio values near 0.4. As load test information is considered, probabilistic within-site variability produces more precise estimates of $R_{1/1500}$, as indicated by the reduced variability evident in histograms of Figure 61. For deterministic within-site variability, the estimates of $R_{1/1500}$ are similarly variable for all numbers of tests. The difference in precision between

estimates from deterministic and probabilistic within-site variability is particularly stark when multiple load tests are considered; for 50 and 200 load tests, the probabilistic histograms are considerably less variable than the deterministic histograms. The trends in variability versus numbers of tests are also evident in Figure 63, a plot of sample standard deviation versus the number of tests, with the number of tests on a logarithmic scale. The standard deviation of the ratio of estimated to true $R_{1/1500}$ decreases significantly with increasing numbers of tests for probabilistic within-site variability, approaching zero for the estimates based on 200 load tests. For deterministic within-site variability, there is only a slight reduction in standard deviation as the number of tests increases, from 0.3 at one test to 0.25 at 200 tests.

The difference in precision between estimates of $R_{1/1500}$ with deterministic and probabilistic within-site variability, especially for large numbers of tests, is likely due to the observation from Section 5.3 that for probabilistic within-site variability, posterior within-site variability converges to true within-site variability for large numbers of tests, whereas deterministic within-site variability assumes a constant value equal to the prior estimate. The finding was confirmed for random sampling of Case 0.5-B in Section 6.2.1. Because all posterior parameters for probabilistic within-site variability converge to the true values, the posterior predictive distribution of resistance for probabilistic within-site variability converges to the true distribution of resistance at the site. For deterministic within-site variability, the error in the prior estimate of within-site variability is present regardless of the number of tests, resulting in imprecision in the estimates of $R_{1/1500}$ that only diminishes slightly with greater numbers of tests.

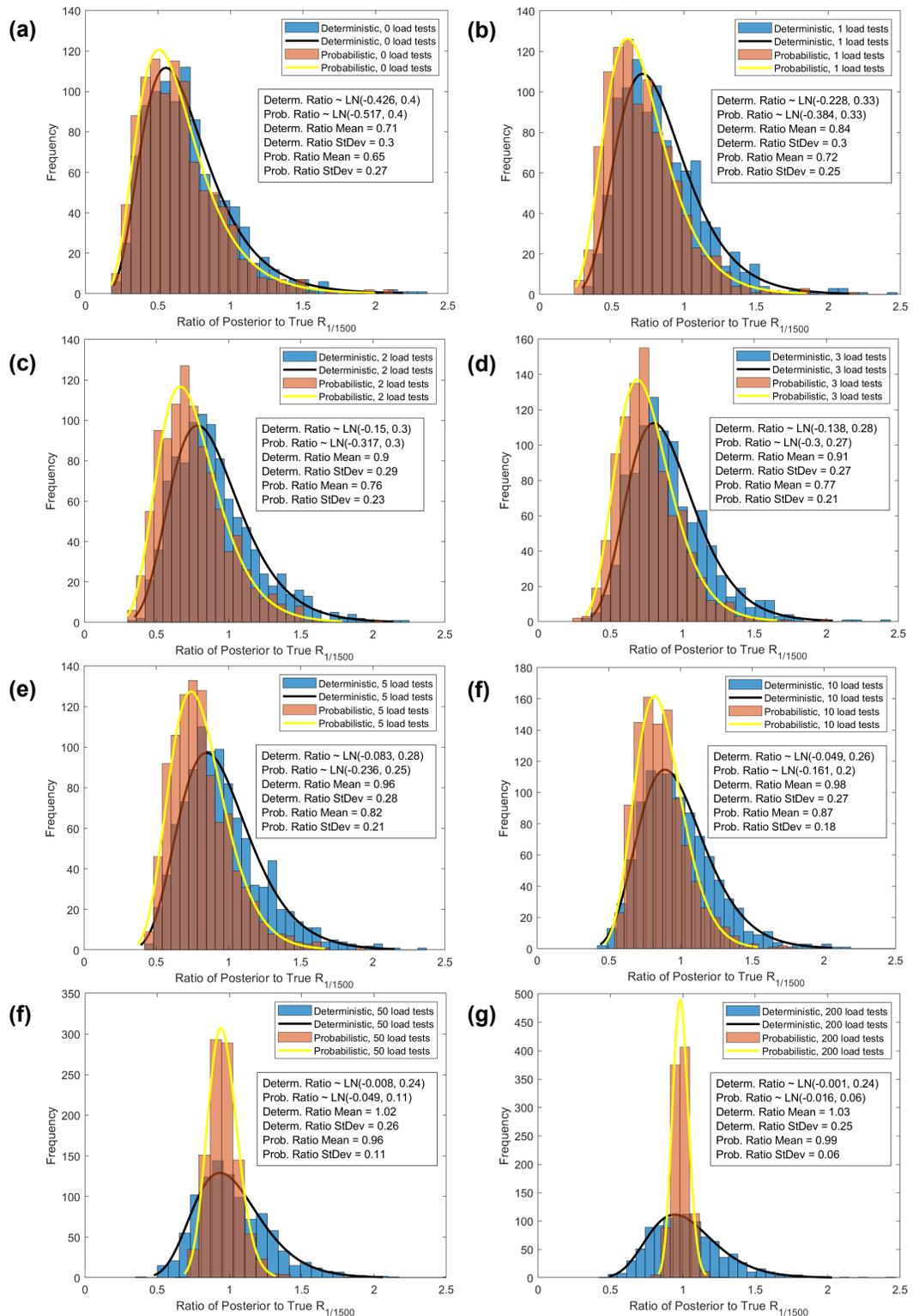


Figure 61: Histograms with fitted lognormal distributions for ratio of estimated to true $R_{1/1500}$, with one dataset for deterministic within-site variability and another for probabilistic within-site variability. Results for prior are shown in (a). Posteriors are defined after (b) one load test, (c) two load tests, (d) three load tests, (e) five load tests, (f) ten load tests, (g) fifty load tests, and (h) 200 load tests..

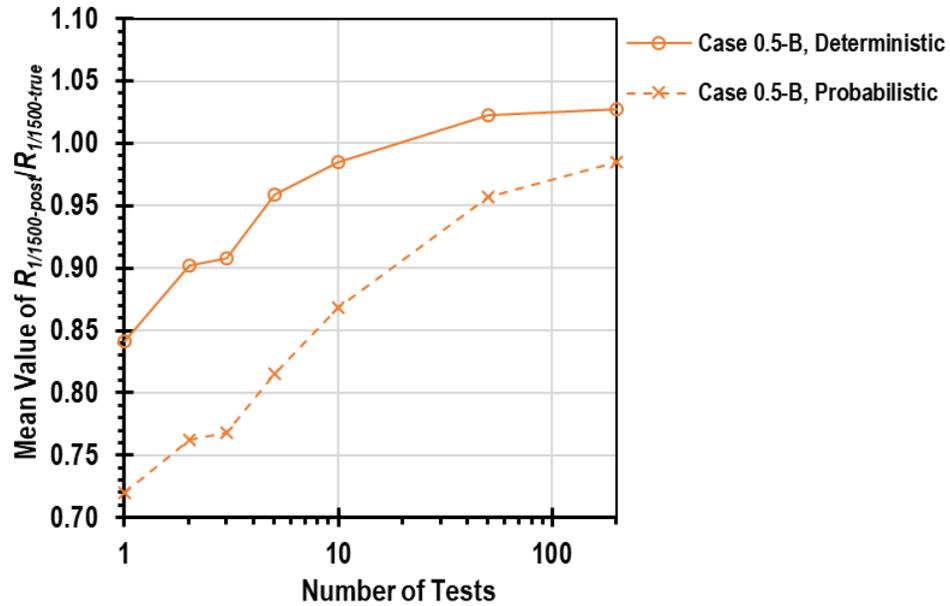


Figure 62: Sample mean of values of ratio of estimated to true $R_{1/1500}$ versus number of load tests for Case 0.5-B with random sampling.

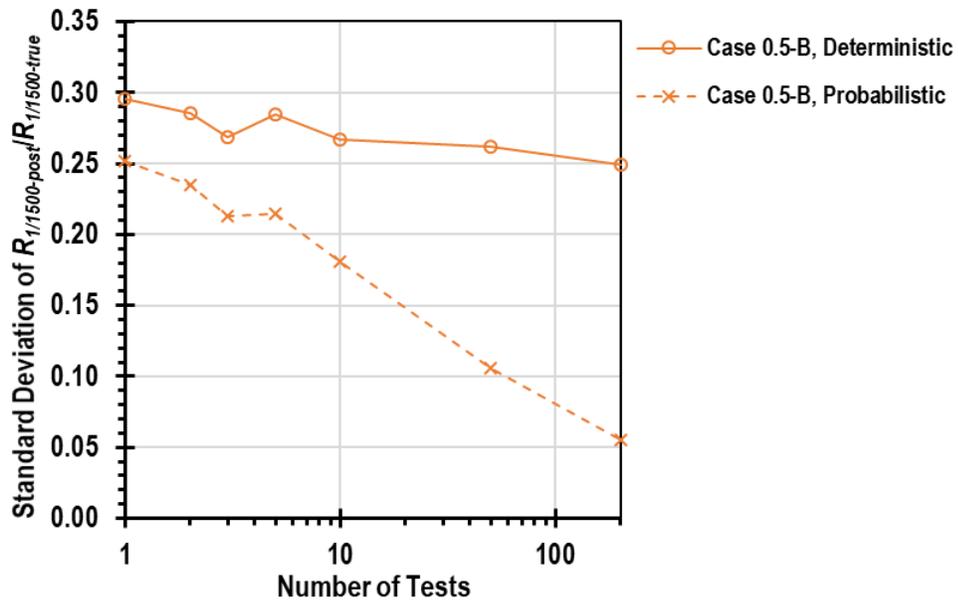


Figure 63: Sample standard deviation of values of ratio of estimated to true $R_{1/1500}$ versus number of load tests for Case 0.5-B with random sampling.

6.3 Errors and Convergence Trends in $R_{1/1500}$

The analysis of results for Case 0.5-B in Section 6.2.2 revealed important trends in the accuracy and precision of estimates of $R_{1/1500}$ for different numbers of tests. The Case 0.5-B results generally indicate prior estimates of $R_{1/1500}$ are on average conservative, and that the

conservatism is diminished with greater numbers of tests, eventually producing estimates of $R_{1/1500}$ that are accurate, on average. The precision of prior estimates of $R_{1/1500}$ was generally low. For deterministic within-site variability, only minor improvements in precision were noted as additional tests were considered; in contrast, considerable improvements were noted for probabilistic within-site variability, with the standard deviation of the ratio of estimated to true ratio of $R_{1/1500}$ diminishing to almost zero for 200 tests.

This section applies the same methodology to other cases (i.e. to collections of sites with different proportions of within-site and among-site variability and to collections of sites with different total variability) to evaluate if similar trends in accuracy and precision are observed. The five cases considered in this section were introduced in Section 6.1.1. One of the five cases is Case 0.5-B from Section 6.2. Three sets of evaluations are presented: (1) sample mean and sample standard deviation of the ratio of estimated to true $R_{1/1500}$ (similar to Section 6.2.2); (2) satisfaction of target reliability, which is a practical interpretation of the accuracy and precision data, and (3) explanation of trends via correlation of errors with true site parameters.

6.3.1 Sample Mean and Sample Deviation of the Ratio of Estimated to true $R_{1/1500}$

In Section 6.2.2, histograms of the ratio of estimated to true $R_{1/1500}$ were presented for Case 0.5-B for various numbers of load tests (Figure 61), and sample mean and sample standard deviation were then plotted versus the number of tests, as presented in Figure 62 and Figure 63, respectively. The same evaluation was completed for analyses of Cases 0.5-A, 0.5-C, 0.3-B, and 0.1-B, with each evaluation based on 1,000 sites generated by random sampling, each evaluation considering results from both deterministic and probabilistic within-site variability, and each evaluation considering the same numbers of load tests (0, 1, 2, 3, 5, 10, 50, and 200). Results of the evaluation are presented in Figure 64, a plot of the sample mean versus the number of tests on a logarithmic scale, and Figure 65, a plot of the sample standard deviation versus the number of tests on a logarithmic scale. In each figure, different colors are used to denote the five cases, solid lines are used to denote deterministic within-site variability results, and dashed lines are used to denote probabilistic within-site variability results.

The trends in the mean value of the ratio of estimated to true $R_{1/1500}$ shown in Figure 64 are generally similar to those discussed in Section 6.2.2 for Case 0.5-B. For all cases, estimates of $R_{1/1500}$ for relatively small numbers of tests, say less than 10 or so, are conservatively biased. The conservatism is a result of uncertainty in the mean value of resistance (represented by σ_μ) and, for probabilistic within-site variability, uncertainty in the within-site variability (represented by σ_σ). As discussed previously (e.g. Section 5.1.3), the true values of uncertainty are zero ($\sigma_\mu = \sigma_\sigma = 0$). Including nonzero values of σ_μ and/or σ_σ produces additional variability in the posterior predictive values of resistance, which, in turn, reduces the value of posterior $R_{1/1500}$. The effect of reduced posterior $R_{1/1500}$ is the conservatively biased ratio values shown in Figure 64. For deterministic within-site variability, the conservatism is considerably reduced with just two or three tests and is essentially eliminated by about ten load tests. This implies the uncertainty in the mean, σ_μ , is diminished with relatively few tests, an observation that is consistent with the results presented in Section 5.3.3. For probabilistic within-site variability, additional tests are required to achieve the same reduction in conservative bias; 50 tests are required to achieve a mean value of the ratio equal to 0.95. One possible explanation is that the number of tests required to effectively eliminate uncertainty in within-site variability, σ_σ , is significantly greater than the number of tests required to eliminate uncertainty in the mean. This explanation is also consistent with the findings of Chapter 5 (Section 5.3.5).

Differences in the mean value of the ratio of estimated to true $R_{1/1500}$ in Figure 64 among the five cases are also noteworthy. As the prior predictive variability decreases (i.e. from 0.5-B to 0.3-B to 0.1-B), the mean value of the ratio increases, most likely because the uncertainty in both the mean and within-site variability parameters are less for cases with smaller prior predictive variability. For the cases with different proportions of within-site to among-site variability, there is no clear trend in the mean value of the ratio of estimated to true $R_{1/1500}$. One might expect the “A” cases with relatively low within-site variability to have greater values of the ratio since for these cases, the uncertainty is predominately in the mean, which diminishes with even one test. However, for these cases, the true value of $R_{1/1500}$ is also greater (since within-site variability is

lower; among-site variability has no effect on the true site distribution). The effect of low within-site variability and greater true $R_{1/1500}$ may offset one another.

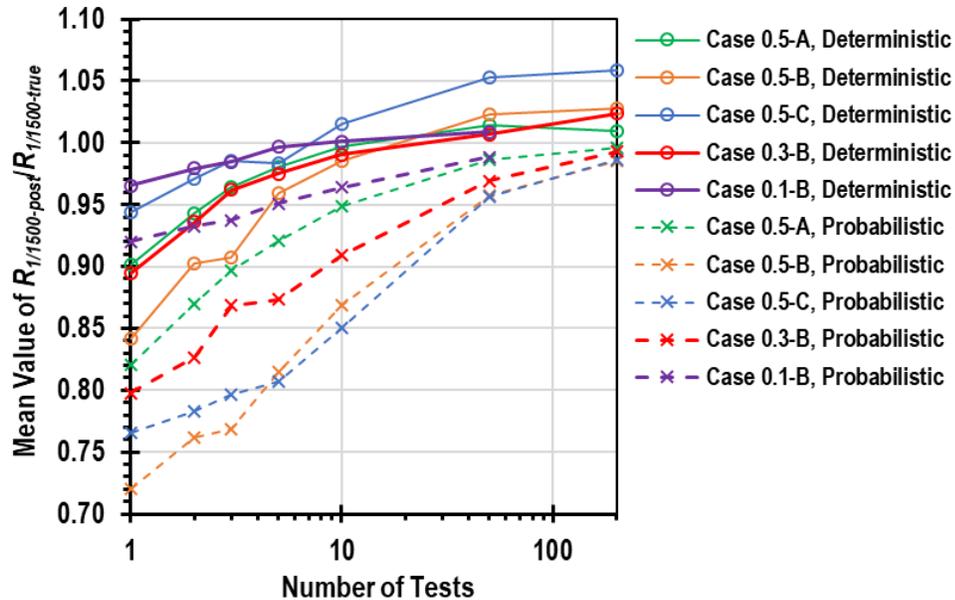


Figure 64: Sample mean of values of ratio of estimated to true $R_{1/1500}$ versus number of load tests for all cases with random sampling.

The trends in the mean value of the ratio of estimated to true $R_{1/1500}$ shown in Figure 65 are also generally similar to those discussed in Section 6.2.2 for Case 0.5-B. For deterministic within-site variability, the standard deviation of the ratio of estimated to true $R_{1/1500}$ is relatively constant, decreasing only slightly as the number of tests increases. This suggests the magnitude of errors in $R_{1/1500}$ with deterministic within-site variability is relatively constant. For probabilistic within-site variability, the standard deviation is initially similar to the value from deterministic within-site variability, but the standard deviation decreases significantly as additional tests are considered, diminishing to near zero for 200 tests. This suggests the magnitude of errors is decreasing, i.e. the precision of the estimates of $R_{1/1500}$ is improved for greater numbers of tests with probabilistic within-site variability. As explained in Section 6.2.2, the finding of relatively constant errors for the deterministic case but diminishing errors for probabilistic within-site variability is consistent with the finding from Chapter 5 that results for probabilistic within-site variability eventually converges to the true value of within-site variability for all cases. In contrast,

any prior error in within-site variability for deterministic within-site variability remains for all posteriors, regardless of the number of load tests.

Unlike the results for the mean of the ratio of estimated to true $R_{1/1500}$, trends in the standard deviation of the ratio among the five cases are relatively straightforward. The standard deviation values are consistent with the magnitude of within-site variability (μ_σ) among the cases, with the least precision noted for the case with greatest within-site variability (Case 0.5-C) and the greatest precision noted for the case with least within-site variability (Case 0.1-B). This finding is consistent with the explanation in the previous paragraph that errors in the estimate of $R_{1/1500}$ are predominately a result of errors in the within-site variability parameter, especially for large numbers of tests. The relative magnitude of precision among the cases is independent of whether deterministic or probabilistic within-site variability is used. However, for probabilistic within-site variability, the difference among cases diminishes for large numbers of tests as the estimated within-site variability improves. For deterministic within-site variability, the difference among cases is independent of the number of tests.

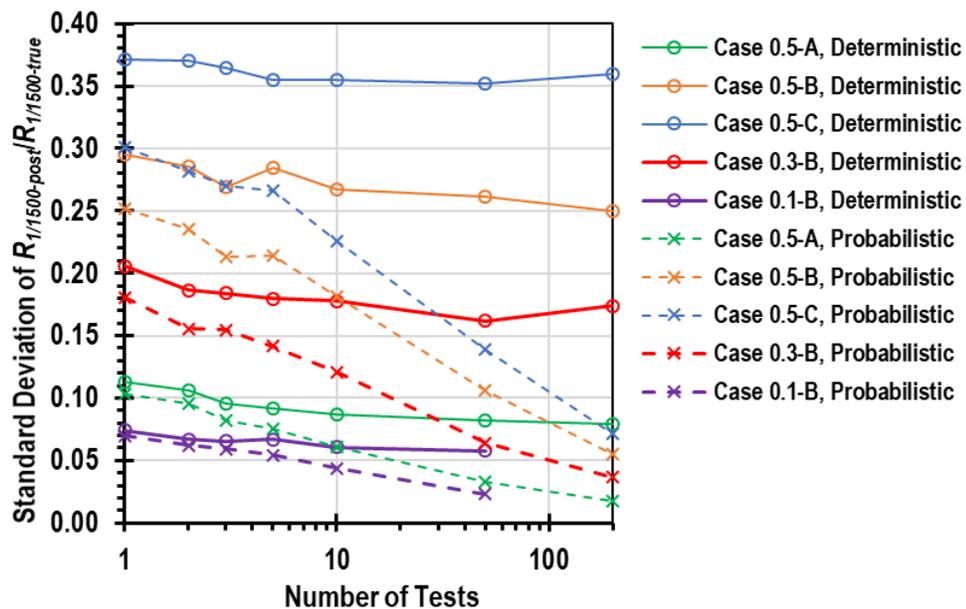


Figure 65: Sample standard deviation of values of ratio of estimated to true $R_{1/1500}$ versus number of load tests for all cases with random sampling.

6.3.2 Satisfaction of Target Reliability

To consider the implications of errors in estimates of $R_{1/1500}$, the ratios of estimated to true $R_{1/1500}$ were assigned to one of three categories describing agreement between the estimate of $R_{1/1500}$ and the true value. The three categories are shown in Figure 66, which presents conceptual distributions of the ratio of estimated to true $R_{1/1500}$. In the categorization, satisfactory outcomes are defined as values of the ratio of estimated to true $R_{1/1500}$ between 0.9 and 1.1. In other words, satisfactory outcomes occur when the estimate of $R_{1/1500}$ is within 10 percent of the true value. When the estimate of $R_{1/1500}$ is less than 90 percent of the true value (ratio value less than 0.9), the outcome is overreliable. When the estimate is greater than 110 percent of the true value (ratio value greater than 1.1), the outcome is underreliable.

The three categories are each shown on the two histograms of the ratio of estimated to true $R_{1/1500}$ in Figure 66. In Figure 66(a), the categories are depicted on a relatively uniform distribution of the ratio of estimated to true $R_{1/1500}$; in Figure 66(b), the categories are depicted on a more variable distribution. The former represents a relatively precise collection of estimates while the latter represents a comparatively imprecise collection of estimates. For both distributions of the ratio, the width of the shaded band representing satisfactory outcomes is the same, but a considerably greater proportion of the precise estimates fall into the satisfactory band compared to the imprecise estimates. It is important to note that both of the histograms shown in Figure 66 are centered at a ratio value of 1.0, indicating the estimates of $R_{1/1500}$ are accurate, on average. In other words, the estimates of $R_{1/1500}$ are unbiased with respect to the true value. As discussed in Section 6.3.1, unbiased estimates of $R_{1/1500}$ are generally only achieved for large numbers of load tests. For small numbers of tests, uncertainty in the mean and standard deviation produces conservatively biased estimates associated with overreliable outcomes.

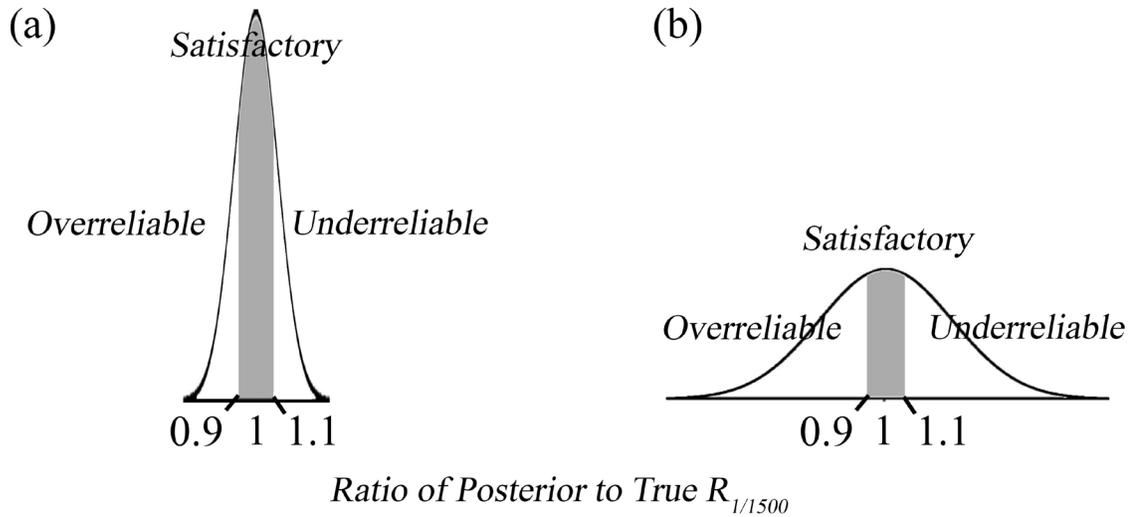


Figure 66: Depiction of overreliable, satisfactory, and underreliable outcomes for (a) relatively uniform and (b) relatively variable distributions of the ratio of estimated to true $R_{1/1500}$.

The categories of Figure 66 are applied to the results of Case 0.5-A in Figure 67, Case 0.5-B in Figure 68, Case 0.5-C in Figure 69, Case 0.3-B in Figure 70, and Case 0.1-B in Figure 71. Each figure is a plot of the percent of outcomes in each category versus the number of tests on a logarithmic scale. Each plot includes six lines, one for each category in blue for deterministic within-site variability and one for each category in red for probabilistic within-site variability. The percent of outcomes in each category for the prior analyses is listed in a textbox at the right of each figure.

The proportion of outcomes for the prior evaluations is instructive. For Case 0.5-A, with relatively small within-site variability compared with among-site variability, the prior outcomes are nearly all overreliable. In contrast, for Case 0.5-C, with relatively high within-site variability compared with among-site variability, only about half of the prior outcomes are overreliable, with a significant proportion of underreliable outcomes (27 percent for deterministic within-site variability, 12 percent for probabilistic within-site variability). The difference in outcomes is consistent with the nature of variability for each case. Both Case 0.5-A and Case 0.5-C have the same prior predictive distribution with $COV = 0.5$; therefore, they have the same prior estimate of $R_{1/1500}$. However, the true value of $R_{1/1500}$ is significantly less for Case 0.5-C because the true sites associated with the case have significantly more variable distributions of resistance. For

Case 0.5-A, the total variability is primarily due to among-site variability, which is uncertainty in the mean. For Case 0.5-C, the total variability is primarily due to within-site variability. In reality, the only variability in the true distribution of resistance is within-site variability; the site has one value for the mean, and it does not vary across the site, by definition. Accordingly, for Case 0.5-A, the estimate of $R_{1/1500}$ is primarily accounting for a source of uncertainty that does not really exist at the site, and the true distribution of resistance at the site is relatively uniform. In contrast, the estimate of $R_{1/1500}$ for Case 0.5-C accounts for a real source of variability at the site, and the true distribution of resistance at the site is relatively variable.

For all cases, common trends are observed in the plots, although the magnitudes of the proportions of outcomes in each category vary widely among the cases and between results for deterministic and probabilistic within-site variability. In general, outcomes for each case are dominated by overreliable outcomes for the prior analysis and for analyses based on relatively small numbers of tests. As additional tests are considered, the proportion of satisfactory outcomes increases as the proportion of overreliable outcomes decreases, indicating load testing is reducing conservatism. For the prior and for all numbers of load tests, the proportion of underreliable outcomes is typically small compared with the proportion of overreliable outcomes, although Case 0.5-C is an exception.

For probabilistic within-site variability, in all cases the proportion of satisfactory outcomes achieves or at least approaches 100 percent for large numbers of tests. For small numbers of tests, the trend toward all satisfactory outcomes is dominated by reducing the proportion of overreliable outcomes; for greater numbers of tests, the proportion of underreliable outcomes is also diminished. For all cases with probabilistic within-site variability, the proportion of underreliable outcomes is relatively small, less than 15 percent for any number of tests, and the proportion of underreliable outcomes never increases as additional tests are considered.

For deterministic within-site variability, a similar trend toward a greater proportion of satisfactory outcomes with increased numbers of load tests is noted. As for probabilistic within-site variability, the trend toward a greater proportion of satisfactory outcomes is driven by a

reduction in the proportion of overreliable outcomes. However, deterministic within-site variability also produces an increase in underreliable outcomes for greater numbers of tests, whereas the proportion of underreliable outcomes generally decreases (and never increases significantly) with additional tests for probabilistic within-site variability. In fact, for all cases considered and for all numbers of tests, deterministic within-site variability produces a greater proportion of underreliable outcomes than probabilistic within-site variability. The proportion of underreliable outcomes for deterministic within-site variability is greater than that for probabilistic within-site variability by about a factor of two for small numbers of tests, and the difference increases as additional tests are considered. The finding of more underreliable outcomes for deterministic within-site variability is consistent with the finding from Chapter 5 that deterministic within-site variability generally produces greater estimates of $R_{1/1500}$ than probabilistic within-site variability because deterministic within-site variability neglects uncertainty in within-site variability.

The trends observed in the preceding paragraphs are consistent with previous discussions of trends in posterior parameters and differences between deterministic and probabilistic within-site variability. For relatively small numbers of tests, the proportion of overreliable outcomes is reduced with additional tests because load test information generally produces significant decreases in uncertainty in the mean (σ_μ) for small numbers of tests (Section 5.3.3). Reducing uncertainty in the mean reduces variability of the posterior predictive distribution, increasing $R_{1/1500}$. This occurs for both deterministic and probabilistic within-site variability. Reducing the other source of uncertainty, uncertainty in within-site variability (σ_σ), is only accomplished with probabilistic within-site variability and generally only for large numbers of load tests (Section 5.3.5), which explains why satisfactory outcomes continue to increase for probabilistic within-site variability for large numbers of tests, approaching 100 percent satisfactory, while the proportion of satisfactory outcomes for deterministic within-site variability stagnates after uncertainty in the mean has been diminished.

The differences among the cases in Figure 67 through Figure 71 are also consistent with previous observations regarding the influence of within-site variability. The differences among the

cases generally indicate more favorable results when within-site variability is small. The more favorable results include a greater proportion of satisfactory outcomes, a smaller proportion of underreliable outcomes, and, for probabilistic within-site variability, convergence to 100 percent satisfactory outcomes at fewer load tests. These findings are consistent with the discussion of the role of within-site variability on posterior parameter convergence in Section 5.3, and with the discussion of the role of within-site variability on precision in Section 6.3.1. Comparison of the “B” cases (Figure 68 versus Figure 70 versus Figure 71) indicates similarly favorable results (greater proportion of satisfactory outcomes, smaller proportion of underreliable outcomes, and convergence to 100 percent satisfactory outcomes at fewer load tests) are produced when the total variability decreases, although the reduction of total variability is associated with decreases in within-site variability, so it is difficult to differentiate between the effect of total variability and within-site variability.

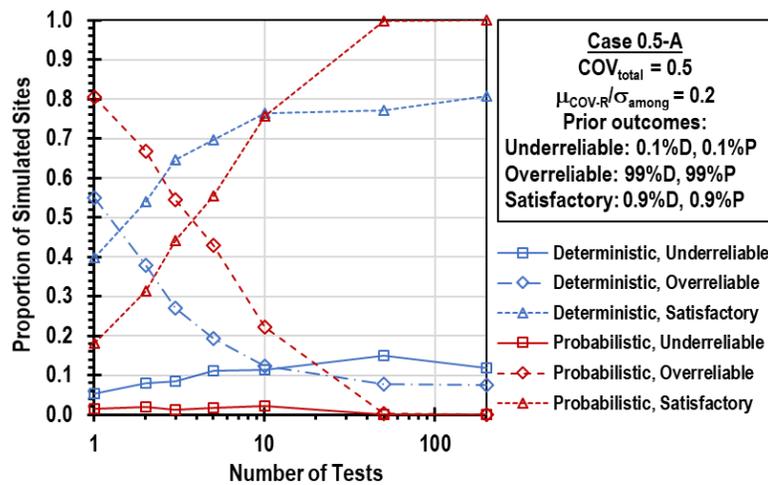


Figure 67: Percent of underreliable, overreliable, and satisfactory outcomes versus number of load tests for Case 0.5-A.

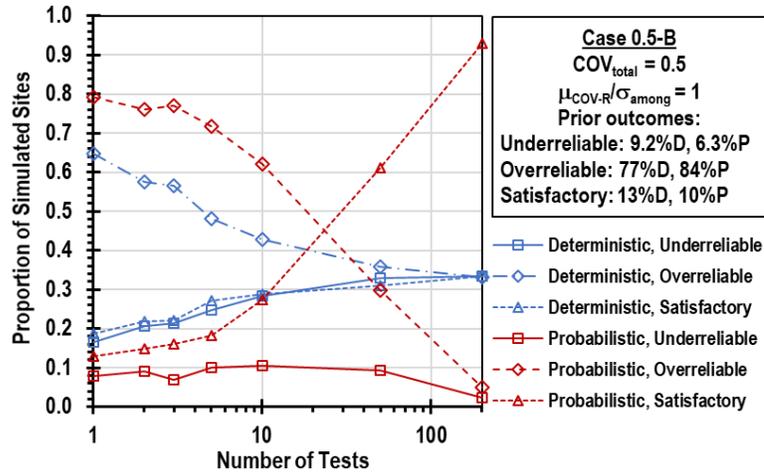


Figure 68: Percent of underreliable, overreliable, and satisfactory outcomes versus number of load tests for Case 0.5-B.

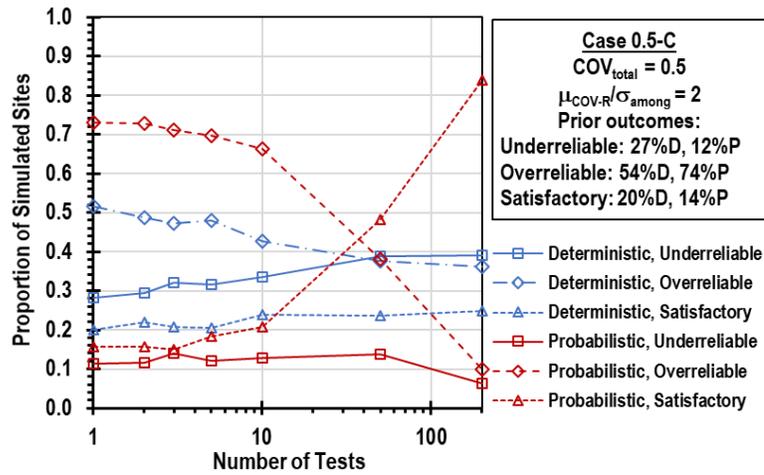


Figure 69: Percent of underreliable, overreliable, and satisfactory outcomes versus number of load tests for Case 0.5-C.

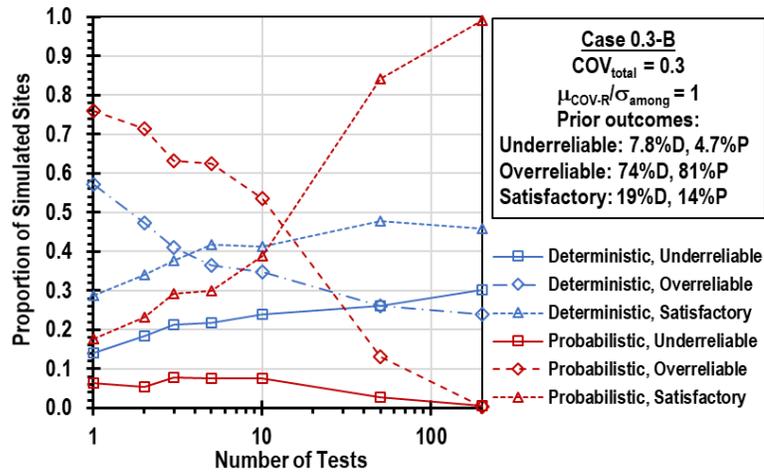


Figure 70: Percent of underreliable, overreliable, and satisfactory outcomes versus number of load tests for Case 0.3-B.

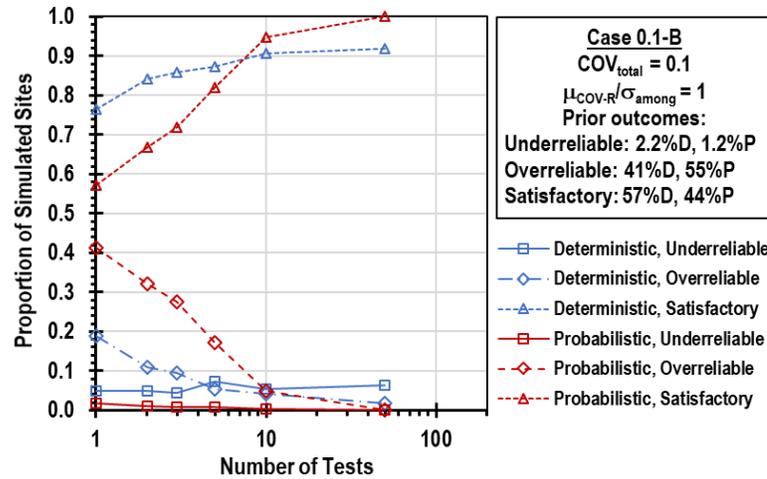


Figure 71: Percent of underreliable, overreliable, and satisfactory outcomes versus number of load tests for Case 0.1-B.

6.3.3 Correlation of Errors with True Site Parameters

The findings in Chapter 5 were based on perfect sampling with the true site mean and true site standard deviation set to the tenth, fiftieth, and ninetieth percentiles of the prior distributions of within-site mean, μ_{within} , and within-site variability, σ_{within} . The findings indicated significant differences between convergence of posterior parameters and $R_{1/1500}$ for different values of the true site mean and true within-site variability. In this section, the effect of true site mean and true within-site variability on the ratio of estimated to true $R_{1/1500}$ are examined with random sampling.

The effect of true site mean and true within-site variability on outcomes for Case 0.5-A are shown in Figure 72 and Figure 73, respectively. The figures are plots of the ratio of estimated to true $R_{1/1500}$ for the 1,000 sites generated by random sampling of Case 0.5-A (Section 6.1.2) versus true values of μ_{within} and σ_{within} , respectively. In both figures, different point symbols are used for different numbers of tests (including zero tests, the prior) and to distinguish between deterministic and probabilistic within-site variability results.

The results shown in Figure 72 indicate a clear trend of decreasing values of the ratio of estimated to true $R_{1/1500}$ with increasing within-site mean, μ_{within} . As explained previously (Sections 6.1.4 and 6.2.2), ratio values less than 1 indicate the estimate of $R_{1/1500}$ is

conservative, with conservatism increasing as the ratio approaches zero. Accordingly, the trend in Figure 72 can be interpreted to indicate that the prior estimate of $R_{1/1500}$ is generally conservative (less than 1 for all but one of the 1,000 sites), and especially conservative for sites with relatively strong foundation resistance (i.e. high μ_{within}). Meanwhile, the results shown for the prior in Figure 73 indicate only random scatter with no discernable effect of within-site variability on the ratio of estimated to true $R_{1/1500}$. This is perhaps confounding, since less conservative results might be anticipated for sites with greater within-site variability. However, Case 0.5-A has a small ratio of within-site variability to among-site variability (0.2). Most of the variability of the prior predictive distribution of resistance is a result of among-site variability, or uncertainty in the mean parameter, which explains the strong effect in Figure 72 and the lack of effect in Figure 73.

Bayesian updating for just one load test essentially eliminates the effect of true μ_{within} on the ratio of estimated to true $R_{1/1500}$. For deterministic and probabilistic within-site variability, there is no discernable trend in the ratio with μ_{within} in Figure 72. In addition, the magnitude of the ratio values has increased compared to the prior, with nearly every value above 0.6 compared to some values near zero for the prior. For one load test, a trend of increasing ratio values with true σ_{within} also emerges, indicating more variable sites do produce less conservative outcomes. The trend is more evident in the results for deterministic within-site variability.

Results for Bayesian updating with 200 load tests indicate probabilistic within-site variability produces estimates of $R_{1/1500}$ that are within about five percent of the true value for all sites, regardless of μ_{within} or σ_{within} . This is consistent with previous findings that all posterior parameters eventually converge for probabilistic within-site variability, producing an accurate and precise posterior predictive distribution. For deterministic within-site variability, values of the ratio of estimated to true $R_{1/1500}$ range from 0.79 to 1.36. There is no trend with μ_{within} , indicating the deterministic posterior mean parameter is equal to the true value. However, there is a clear trend with σ_{within} : sites with greater within-site variability have greater values of the ratio, indicating less conservative (in fact, unconservative) outcomes for these sites. The trend with within-site variability for the case of 200 load tests reflects errors in within-site variability estimates for

deterministic within-site variability, which, unlike probabilistic within-site variability, does not update σ_{within} to reflect results.

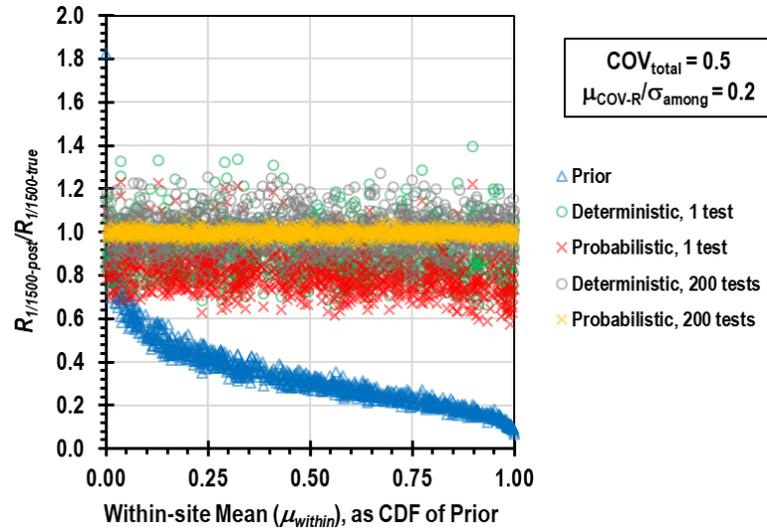


Figure 72: Ratio of estimated to true $R_{1/1500}$ versus within-site mean for Case 0.5-A.

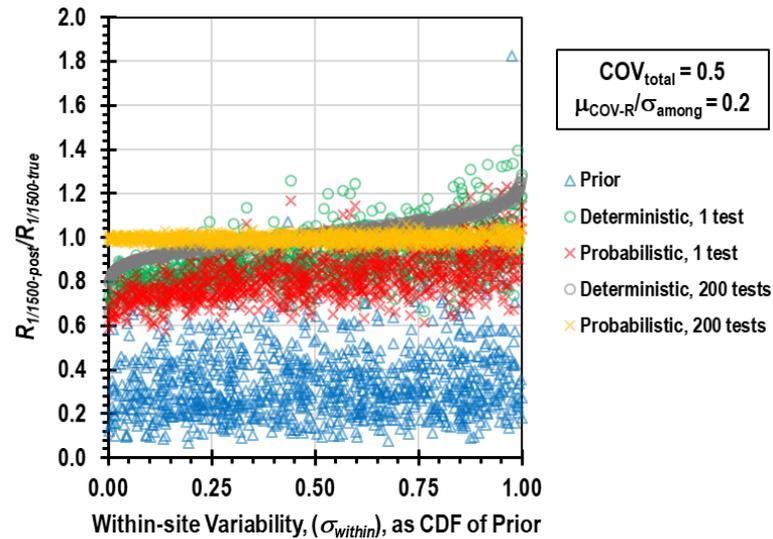


Figure 73: Ratio of estimated to true $R_{1/1500}$ versus within-site variability for Case 0.5-A.

To extend the analysis based on Figure 72 and Figure 73 to other cases, linear regression was applied to similar plots prepared for Case 0.5-A, Case 0.5-B, and Case 0.5-C, with consideration of additional numbers of load tests (1, 2, 3, 5, 10, 50, 200). The resulting linear regression parameters, slope and intercept, are plotted versus the number of load tests in Figure 74 and Figure 75. Figure 74 shows the slope and intercept from regression of the ratio versus

true site mean (μ_{within}), and Figure 75 shows the slope and intercept from regression of the ratio versus true within-site variability (σ_{within}).

Consider the example of Figure 72 and Figure 73. Regression against the site mean (Figure 72) would produce a negative slope for the prior, increasing to zero slope for small numbers of load tests. The intercept would be significantly less than one for the prior, but increase to near 1.0 for one test. Regression against the within-site variability (Figure 73) would produce slope values near zero for the prior. For deterministic within-site variability, the slope would increase as additional load tests are considered, reflecting the dependence on within-site variability discussed above. For probabilistic within-site variability, the slope would remain near zero for all numbers of tests. For both deterministic and probabilistic within-site variability, the intercept value would be relatively low for the prior and small numbers of tests, and increase to about 1.0 for large numbers of tests.

The ideal regression parameters for both site mean and within-site variability are slope parameters equal to zero and intercept parameters equal to one. The ideal slope is zero because the ratio of estimated to true $R_{1/1500}$ – a measure of the accuracy and precision of estimates of $R_{1/1500}$ – should ideally be independent of the site characteristics. Given the ideal slope parameter of zero, the ideal value of the intercept is equal to 1 because values of the ratio equal to 1 indicate the posterior estimate of $R_{1/1500}$ is equal to the true value of $R_{1/1500}$.

For regression of the ratio versus site mean, the slope (Figure 74(a)) converges to zero for all cases and for both deterministic and probabilistic within-site variability. Convergence is achieved for significantly fewer tests for Case 0.5-A, which has the smallest ratio of within-site to among-site variability (0.2), than for the cases with greater values of the ratio. For Case 0.5-A, the slope is relatively small even for one test; for the other two cases, the slope decreases to zero logarithmically, with convergence requiring 50 to 200 tests. Convergence toward a slope of zero indicates success in achieving independence of the average accuracy of posterior $R_{1/1500}$ from the site mean. Independence from the site mean indicates the mean parameter has been correctly updated. This is consistent with findings presented in Chapter 5, which indicated both

deterministic and probabilistic within-site variability produce successful convergence to the mean. That convergence was achieved for fewer numbers of tests when within-site variability is relatively low is also consistent with the findings from Chapter 5 and with other findings in this chapter. The significantly greater number of tests required for convergence for Cases 0.5-B and 0.5-C is most likely a result of high within-site variability resulting in less weight being applied to the evidence from load tests. There could also be an effect from highly variable load test results producing slower convergence.

The intercept from regression of the ratio versus site mean (Figure 74(b)) also converges to the ideal value, 1, for both deterministic and probabilistic within-site variability. (Deterministic within-site variability for Case 0.5-C is a possible exception; after 200 load tests, it is still decreasing but within 10 percent of 1.) The convergence of both the slope and intercept parameters toward the ideal values is consistent with the findings from Chapter 5 that the posterior parameters for the mean converge to true values for both deterministic and probabilistic within-site variability.

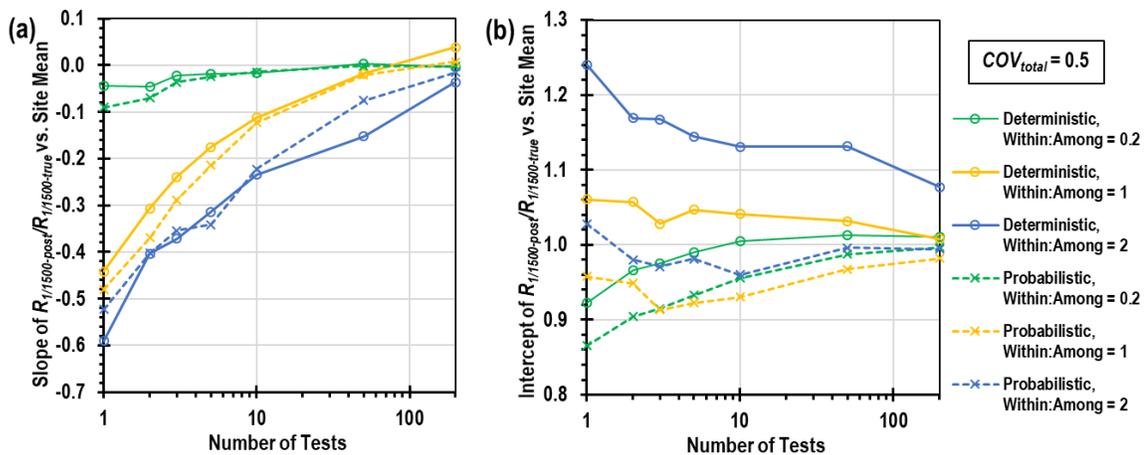


Figure 74: (a) Slope and (b) intercept parameters versus number of load tests for all cases with $COV_{total} = 0.5$. Parameters are from linear regression of ratio of estimated to true $R_{1/1500}$ versus true site mean (μ_{within}).

Unlike for regression versus the site mean, regression of the ratio of estimated to true $R_{1/1500}$ versus true within-site variability (Figure 75) produces starkly different results for the deterministic and probabilistic within-site variability. For deterministic within-site variability, the

slope parameter is positive and essentially constant for all numbers of tests, with the slope values increasing with the ratio of within-site to among-site variability. The slope values reflect the influence of errors in prior estimates of within-site variability on all results with deterministic within-site variability: the errors produce more unconservative results when true within-site variability is greater. For probabilistic within-site variability, the slope values are positive for small numbers of tests, converging to zero for large numbers of tests. Fifty to 200 tests are required to achieve convergence. The large number of tests required for convergence to the true within-site variability (and therefore slope values of zero) is consistent with findings from Chapter 5 (Section 5.3.4).

For the intercept parameter of regression of the ratio versus true within-site variability, a similar trend is noted: the intercept values are relatively constant for all numbers of tests for deterministic within-site variability, whereas probabilistic within-site variability intercept values all converge to the ideal value of 1. As for the other regressions considered in this section, convergence is achieved for fewer tests when the ratio of within-site to among-site variability is small. The convergence to a slope of zero and intercept of one for probabilistic within-site variability indicates accuracy, on average, for all true values of within-site variability. For deterministic within-site variability, the intercept values are less than one, indicating the deterministic target resistance values are conservative for true values of within-site variability less than the assumed value, which is a rational finding. Since the slope values for deterministic within-site variability are all positive, the conservatism is reduced as true within-site variability increases. Based on the findings in Section 6.3.2, the reduction in conservatism can produce a significant proportion of underreliable outcomes.

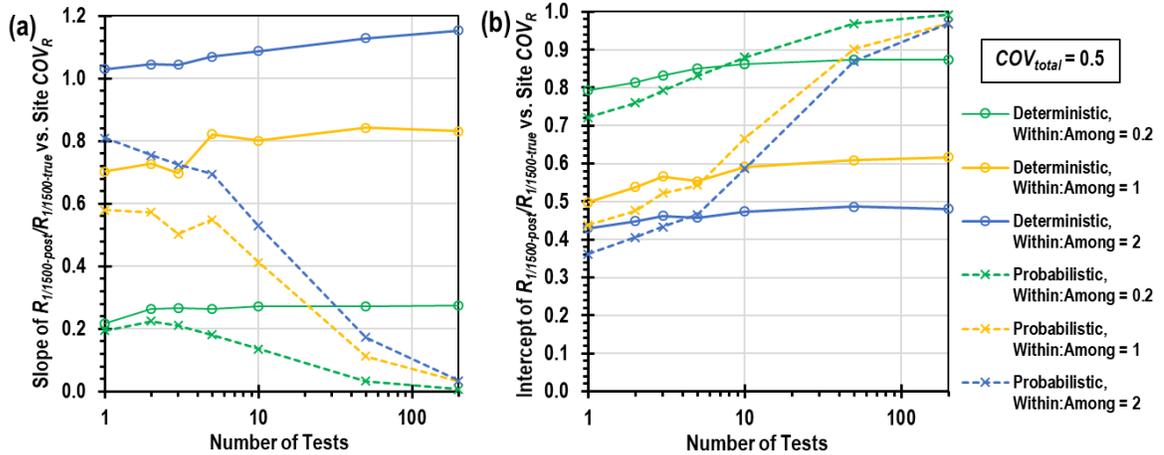


Figure 75: (a) Slope and (b) intercept parameters versus number of load tests for all cases with $COV_{total} = 0.5$. Parameters are from linear regression of ratio of estimated to true $R_{1/1500}$ versus true within-site variability (σ_{within}).

6.4 Increase in $R_{1/1500}$ with Load Tests

The analyses described in this chapter have thus far focused on the accuracy and precision of estimates of $R_{1/1500}$, i.e. how the estimated values compare to true values. It is also useful to consider the magnitude of estimates of $R_{1/1500}$ for different numbers of tests with respect to the prior $R_{1/1500}$ without load tests. Considering the magnitude of estimates of $R_{1/1500}$ is a method for evaluating the practical value of load test information, since the estimate of $R_{1/1500}$ is a comparable to reliability-based estimates of factored resistance in LRFD.

To compare magnitudes of $R_{1/1500}$, the quantity expressed in Eq. 6-2 is used. As explained in Section 6.1.4, the quantity is the ratio of posterior and prior estimates of $R_{1/1500}$. Since the posterior estimates vary within the cases (because of differences among the simulated sites as well as because of sampling load test values within each simulated site), the sample mean is used to represent an average value of posterior resistance. The quantity in Eq. 6-2, then, represents the average increase in resistance one might expect to achieve by load testing, maintaining satisfaction of $p_f = 1/1500$ (without consideration of load variability) and expressed as a factor of the prior estimate.

For example, consider a prior estimate of predicted resistance (R_p) of 1,000 kips. Uncertainty in the prior estimate results in a value of $R_{1/1500} = 0.3$, corresponding 300 kips. If Bayesian updating with load test information yields a value of $R_{1/1500} = 0.6$, corresponding 600 kips, the posterior estimate of $R_{1/1500}$ is twice the prior value, producing a value of 2 for the quantity of Eq. 6-1.

The quantity in Eq. 6-2 was computed for all results of random sampling, including all five cases (Section 6.1.2), all numbers of load tests, and both deterministic and probabilistic within-site variability. As explained in the definition of Eq. 6-1, one value was reported for each combination of case, number of load tests, and deterministic-vs.-probabilistic within-site variability by averaging the 1,000 posterior values associated with each combination. Results are presented in Figure 76 through Figure 79. Figure 76 and Figure 77 are results for Cases 0.5-A, 0.5-B, and 0.5-C, which are considered jointly to evaluate the effect of the ratio of within-site to among-site variability on the increase in estimated $R_{1/1500}$ associated with load testing. Figure 78 and Figure 79 are results for Cases 0.5-B, 0.3-B, and 0.1-B, which are considered jointly to evaluate the effect of nominal prior predictive COV on the increase in estimated $R_{1/1500}$.

The ratio of posterior to prior estimates of $R_{1/1500}$ (i.e. the quantity of Eq. 6-1) is plotted versus the number of tests for Cases 0.5-A, 0.5-B, and 0.5-C in Figure 76. All three cases have a nominal prior predictive COV of 0.5. Each case is shown in a different color, with solid lines representing results from deterministic within-site variability and dashed lines representing results from probabilistic within-site variability. The most striking result from the figure is how much greater the ratio of posterior to prior estimates of $R_{1/1500}$ is for Case 0.5-A than for Cases 0.5-B or 0.5-C: values for Case 0.5-A are between 2.0 and 2.7, while values for the other cases are between 0 and 0.7. The difference suggests the ratio of within-site to among-site variability has a large influence on the increase in estimated $R_{1/1500}$ from load testing. The influence is supported by Figure 77, which is a plot of the same data but with the ratio of posterior to prior estimates of $R_{1/1500}$ plotted against the ratio of within-site to among-site variability rather than number of tests. In Figure 77, the different numbers of tests are represented by different color lines, and again

solid lines are used to represent deterministic within-site variability while dashed lines represent probabilistic within-site variability. The results of Figure 77 indicate a steep decrease in the ratio of posterior to prior estimates of $R_{1/1500}$ between values of the ratio of within-site to among-site variability of 0.2 and 1.0, and a less pronounced decrease between values of 1 and 2.

It is important to note that differences in values of the ratio of posterior to prior estimates of $R_{1/1500}$ represented by the data in Figure 76 and Figure 77 are due only to increases in the estimated value of $R_{1/1500}$ between the prior and the posterior. All three cases have the same prior predictive *COV* and therefore the same prior values of $R_{1/1500}$ (approximately 0.22 for deterministic within-site variability and 0.20 for probabilistic within-site variability). It is also important to emphasize the significance of the difference in results between the case with a low ratio of within-site to among-site variability (Case 0.5-A) and the cases with high ratios of within-site to among-site variability (Cases 0.5-B and 0.5-C). For Case 0.5-A, one load test increases the prior estimate of $R_{1/1500}$ by a factor of 3, on average. For the other cases, one load test increases the prior estimate of $R_{1/1500}$ by a factor of at most 1.25 (again using average posterior results). The difference between the results has practical implications: the expense and complications of load testing are significantly easier to justify when the factored (or allowable) resistance increases by a factor of 3 compared to when the increase corresponds to a factor of 1.25.

The two other factors that can be evaluated from Figure 76 and Figure 77, number of tests and deterministic-vs.-probabilistic within-site variability, have considerably more modest effects on the increase in estimated $R_{1/1500}$ from the prior to the posterior. For all levels of the ratio of within-site to among-site variability and for both deterministic and probabilistic within-site variability, the effect of number of tests is relatively modest beyond the first test. For instance, for Case 0.5-B, the first test provides a considerable improvement compared to the prior case, increasing the prior estimate of $R_{1/1500}$ by a factor of 3. Subsequent tests increase the estimate of $R_{1/1500}$ significantly less; for 200 tests, the estimate of $R_{1/1500}$ is about 3.7 times the prior estimate, with the increase between one test and 200 tests approximately linear on a logarithmic

scale. Indeed, the shape of the increase versus number of test lines in Figure 76 suggests diminishing returns with increasing numbers of tests; on average, additional tests always increase the estimated value of $R_{1/1500}$, but the increases are smaller for each additional test. This finding is consistent with current practices for load testing, which commonly involve only one test per project site.

The effect of deterministic versus probabilistic within-site variability on the estimated value of $R_{1/1500}$ is also modest, with similar increases in estimates of $R_{1/1500}$ noted for both. An exception is for large numbers of tests, say more than 10, for the case where within-site variability is twice among-site variability (Case 0.5-C). For large numbers of tests with Case 0.5-C, probabilistic within-site variability produces greater increases in $R_{1/1500}$ relative to the prior estimate compared with deterministic within-site variability.

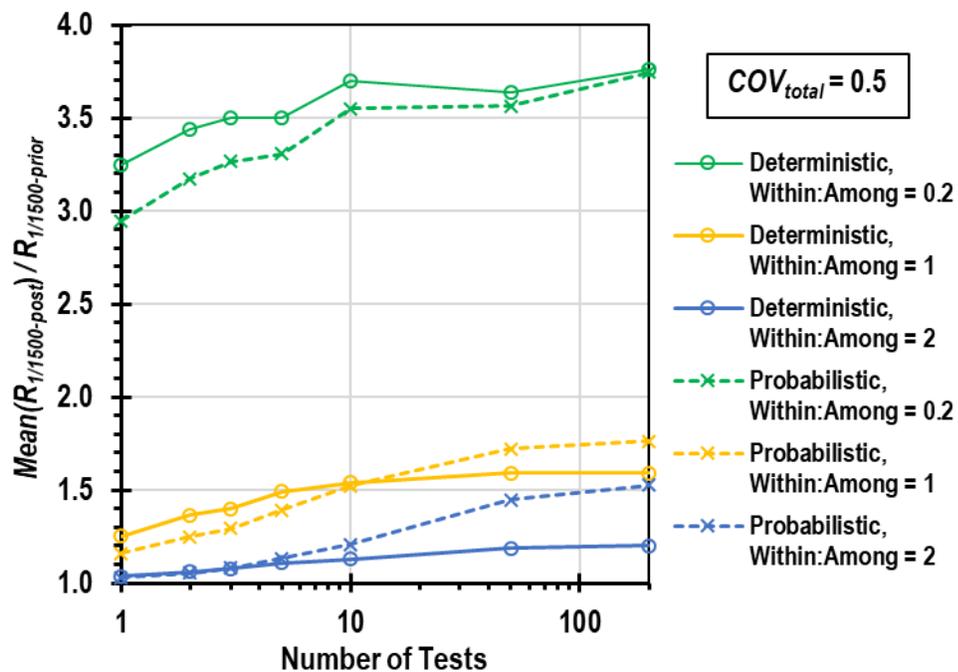


Figure 76: Increase in $R_{1/1500}$ versus number of tests for different values of the ratio of within-site to among-site variability. Nominal predictive COV is 0.5 for all cases.

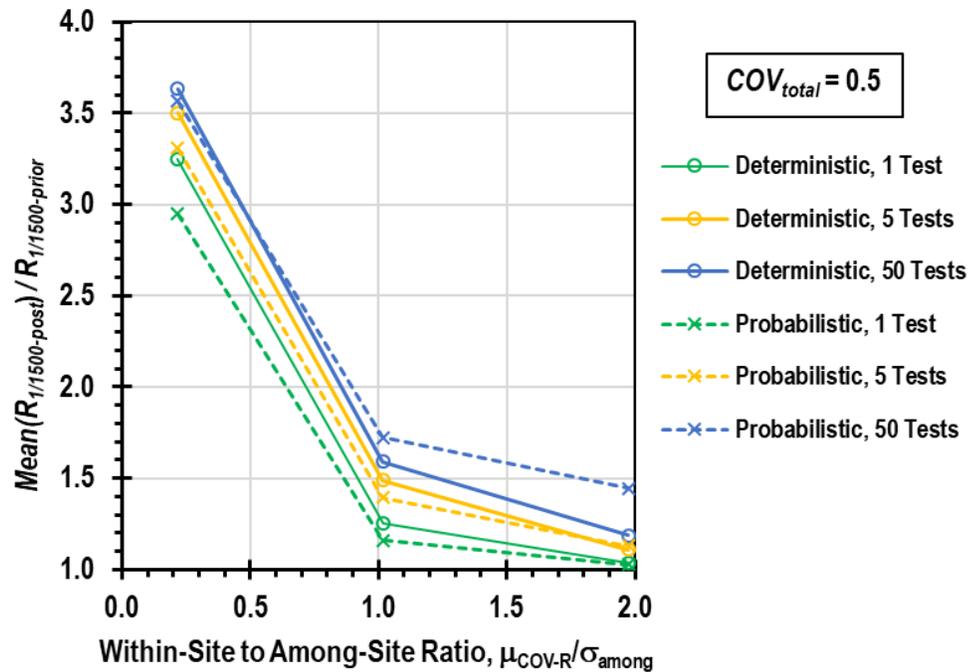


Figure 77: Increase in $R_{1/1500}$ versus ratio of within-site to among-site variability for different numbers of tests. Nominal predictive COV is 0.5 for all cases.

The ratio of posterior to prior estimates of $R_{1/1500}$ (i.e. the quantity of Eq. 6-1) is plotted versus the number of tests for Cases 0.1-B, 0.3-B, and 0.5-B in Figure 78. All three cases have a ratio of within-site variability to among-site variability of 1. Each case is shown in a different color, with solid lines representing results from deterministic within-site variability and dashed lines representing results from probabilistic within-site variability. The results support the findings from Figure 76 and Figure 77 regarding the relatively modest effects of deterministic-vs.-probabilistic within-site variability approach and number of tests. The number of tests factor does look more significant when plotted on the expanded vertical axis scale of Figure 78 versus Figure 76, but the increases are still relatively modest (i.e. the steepest increase is from a ratio of 1.2 to 1.3 from one to two tests when nominal predictive COV is 0.5), and the increase is still linear on a logarithmic scale for the number of tests.

The difference in the ratios of posterior to prior estimates of $R_{1/1500}$ among the three cases in Figure 78 suggest a significant effect of nominal prior predictive COV on the relative increase in $R_{1/1500}$ from prior to posterior. For Case 0.1-B, with the smallest value of nominal prior predictive COV (0.1), the ratio of posterior to prior estimates of $R_{1/1500}$ is small, less than about

1.1 for all numbers of tests. For Case 0.5-B, with the greatest value of nominal prior predictive COV (0.5), the ratio is greater, ranging from 1.2 for one test to 1.8 for 200 tests. The influence of nominal prior predictive COV is supported by Figure 79, which is a plot of the same data as Figure 78 but with the nominal prior predictive COV rather than the number of tests. In Figure 79, the different numbers of tests are represented by different color lines, and again solid lines are used to represent deterministic within-site variability while dashed lines represent probabilistic within-site variability. The results of Figure 79 indicate a linear increase in the ratio of posterior to prior estimates of $R_{1/1500}$ with increasing values of nominal prior predictive COV . This suggests load testing provides greater benefits in terms of factored (or allowable) resistance when total variability is great.

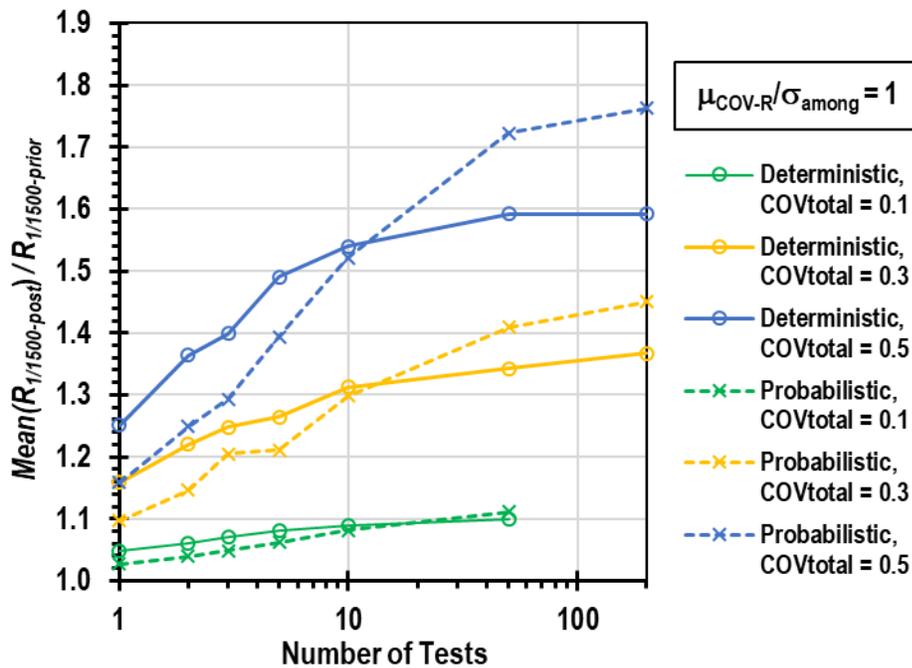


Figure 78: Increase in $R_{1/1500}$ versus number of tests for different values nominal prior predictive COV . Ratio of within-site to among-site variability is 1.0 for all cases.

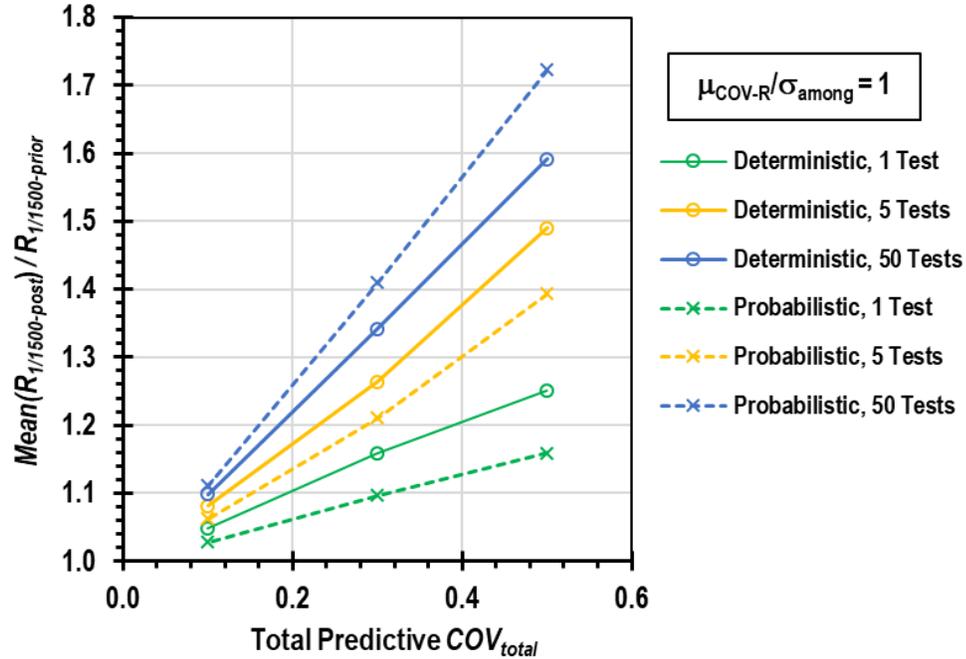


Figure 79: Increase in $R_{1/1500}$ versus ratio of nominal prior predictive COV for different numbers of tests. Ratio of within-site to among-site variability is 1.0 for all cases.

6.5 Summary

The numerical tool from Chapter 4 was used to perform Bayesian updating and interpret the 1/1500th value of resistance, $R_{1/1500}$, for five different cases with randomly sampled load test results. Each case represents a different combination of nominal prior predictive COV and the ratio of within-site variability to among-site variability; the same five cases were analyzed (along with seven additional cases) in Chapter 5 for perfect sampling. Each of the five cases was analyzed for different numbers of load tests: 0 (the prior), 1, 2, 3, 5, 10, 50, and 200. For each combination of case and number of load tests, 1,000 sites were generated by random sampling. The sites were generated by pairing (1) a value of mean foundation resistance randomly sampled from the prior distribution of site specific mean, μ_{within} , with (2) a value of within-site variability randomly sampled from the prior distribution of within-site variability, σ_{within} . The resulting pair of mean and within-site variability defines the “true,” site specific distribution of resistance. Load test results were randomly sampled from the true distribution for each site. For each site generated by random sampling, Bayesian updating was performed using the randomly sampled load test

results, and the posterior results of Bayesian updating were used to interpret $R_{1/1500}$ from posterior predictive distributions of resistance.

A variety of analyses were used to interpret the resulting collections of $R_{1/1500}$ estimates. Primarily, the interpretations focus on the accuracy, precision, and errors associated with posterior estimates of $R_{1/1500}$ compared with the true value of $R_{1/1500}$, which is defined from the site-specific distribution. Interpretations are based on the ratio of estimated to true $R_{1/1500}$. A value of the ratio exactly equal to 1 indicates a perfectly accurate prediction from the posterior; ratio values less than 1 indicate conservative estimates and ratio values greater than 1 indicate unconservative estimates. The variability of a collection of ratio values is used to interpret the precision of a set of $R_{1/1500}$ estimates, with more variable results indicating less precision.

For all cases, results indicate both deterministic and probabilistic within-site variability produce, on average, conservative estimates of $R_{1/1500}$ for the prior evaluation (without load tests). For both deterministic and probabilistic within-site variability, the accuracy of estimates of $R_{1/1500}$ increases (i.e. conservatism is reduced) for greater numbers of load tests, eventually to the position where estimated $R_{1/1500}$ is, on average, nearly equal to the true value. The precision of the estimates of $R_{1/1500}$ differs for deterministic and probabilistic within-site variability. For the prior, both deterministic and probabilistic within-site variability produce relatively imprecise estimates of $R_{1/1500}$, as indicated by high variability of the distributions of the ratio of estimated to true $R_{1/1500}$. For deterministic within-site variability, the variability is relatively constant for all numbers of load tests; in other words, the precision of the deterministic estimates of $R_{1/1500}$ does not improve significantly with additional load tests. For probabilistic within-site variability, the variability decreases with increasing numbers of load tests, approaching zero for large numbers of tests. The increase in precision noted for probabilistic within-site variability is attributed to the Bayesian updating of within-site variability. Results for deterministic and probabilistic within-site variability both indicate convergence to the true value of the site-specific mean for all sites. Because results for probabilistic within-site variability also include convergence to the true value

of within-site variability, the posterior predictive distribution of resistance converges to the true distribution at the site. In contrast, any errors in the prior estimate of within-site variability for deterministic within-site variability are present regardless of the number of load tests.

The implications of the finding that estimates of $R_{1/1500}$ with deterministic within-site variability produce greater errors than the same estimates from probabilistic within-site variability were investigated by dividing outcomes into categories of satisfactory, overreliable, and underreliable. The categories are defined with respect to the ratio of estimated to true $R_{1/1500}$. Outcomes with values of the ratio between 0.9 and 1.1 are deemed satisfactory, outcomes with values of the ratio less than 0.9 are deemed overreliable, and outcomes with values of the ratio greater than 1.1 are deemed underreliable.

For the prior, the proportion of outcomes depends on the ratio of within-site variability to among-site variability. When the prior predictive variability is mainly due to among-site variability, the outcomes are predominately overreliable. When the prior predictive variability is mainly due to within-site variability, significantly more underreliable outcomes are noted. In the evaluation documented in this chapter, both cases have the same prior predictive distribution, so the estimates of $R_{1/1500}$ are the same. The difference in outcomes is therefore due to the true value of $R_{1/1500}$, which is significantly less for the case dominated by within-site variability. The true value of $R_{1/1500}$ is less for the case of significant within-site variability because within-site variability is the true, inherent variability at the site, whereas among-site variability is uncertainty in the mean – a Bayesian construct that is critical in reliability evaluations but which does not manifest itself as true variability at any particular site.

For both deterministic and probabilistic within-site variability, outcomes for small numbers of load tests are dominated by overreliable estimates of $R_{1/1500}$. As additional tests are considered, the proportion of satisfactory outcomes increases as the proportion of overreliable outcomes decreases. For probabilistic within-site variability, the proportion of satisfactory cases converges to 100 percent or near 100 percent for all cases. Deterministic within-site variability

outcomes do not converge to 100 percent satisfactory for any of the cases; for the cases with prior predictive COV of 0.5, deterministic within-site variability achieves 80 percent satisfactory cases when within-site variability is low and less than 40 percent satisfactory cases when within-site variability is high.

For all cases with probabilistic within-site variability, the proportion of underreliable outcomes is relatively small, less than 15 percent for any number of tests, and the proportion of underreliable outcomes decreases to near zero for large numbers of tests. In contrast, the proportion of underreliable cases for deterministic within-site variability with small numbers of load tests was 10 percent for the case with a low ratio of within-site variability to among-site variability and 30 percent for the case with a high ratio of within-site variability to among-site variability. In addition, and importantly, the proportion of underreliable cases generally increased with additional tests for deterministic within-site variability. For all cases considered and for all numbers of tests, deterministic within-site variability produces a greater proportion of underreliable outcomes than probabilistic within-site variability. The proportion of underreliable outcomes for deterministic within-site variability is greater than that for probabilistic within-site variability by about a factor of two for small numbers of tests, and the difference increases as additional tests are considered.

To better understand the sources of errors in the estimates of $R_{1/1500}$, the ratio of estimated to true $R_{1/1500}$ was plotted versus the true (randomly sampled) site parameters, mean and within-site variability, for the different cases. Evaluation of the results indicates that for deterministic and probabilistic within-site variability with small numbers of tests, errors are generally correlated with both true site mean and true within-site variability. For both deterministic and probabilistic within-site variability, the correlation of errors with the mean diminish for relatively small numbers of tests. For greater numbers of tests, correlation of errors with true within-site variability is undiminished for deterministic within-site variability but near zero for probabilistic within-site variability. The difference is consistent with the finding that results for

probabilistic within-site variability converge to true within-site variability compared with persistent errors in within-site variability for deterministic within-site variability.

In addition to the analyses performed to evaluate the accuracy, precision, and errors associated with posterior estimates of $R_{1/1500}$, analyses were also completed to evaluate the increase in estimated $R_{1/1500}$ from the prior evaluations without load tests to the posterior evaluations with load tests. The results indicate the average increase in $R_{1/1500}$ from the prior to posterior depends on many factors but is dominated by the ratio of within-site variability to among-site variability. When the ratio of within-site variability to among-site variability is small (i.e. the total variability results primarily from among-site variability), equal to 0.2, the posterior estimate of $R_{1/1500}$ after one load test is greater than the prior estimate of $R_{1/1500}$ by a factor of three on average. When the ratio of within-site to among-site variability is greater, equal to 1.0, the posterior estimate of $R_{1/1500}$ after one load test is greater than the prior by a factor of 1.25 on average. The finding that load testing increases $R_{1/1500}$ considerably more when total variability is primarily caused by among-site variability rather than within-site variability is consistent with the findings regarding posterior parameters. The analysis of posterior parameters revealed that among-site variability is effectively diminished by load testing, and that among-site variability is generally diminished with relatively few tests. In contrast, the magnitude of within-site variability generally is not reduced by load testing.

The average increase in $R_{1/1500}$ from the prior to posterior is also sensitive to the nominal prior predictive COV , although the effect was less significant than that of the ratio of within-site to among-site variability. The ratio of posterior to prior $R_{1/1500}$ increased linearly between nominal predictive COV values of 0.1 and 0.5. The number of load tests has a relatively modest influence on the average increase in $R_{1/1500}$ from the prior to posterior. When the ratio of within-site to among-site variability is 0.2 (i.e. the smallest value evaluated), the ratio of posterior to prior $R_{1/1500}$ is 3 for one test and increases to 3.7 after 200 load tests: the vast majority of the increase in $R_{1/1500}$ is realized for the first test. The increase in the ratio of posterior to prior $R_{1/1500}$ is

approximately linear when the number of tests is on a logarithmic scale. For cases with greater ratios of within-site to among-site variability, the first test has less effect, but similarly flat slopes are observed when the ratio of posterior to prior $R_{1/1500}$ is plotted versus the number of tests on a logarithmic scale. Finally, the effect of the deterministic-vs.-probabilistic within-site variability on the increase in estimated $R_{1/1500}$ is limited. The difference between ratios of posterior to prior $R_{1/1500}$ for deterministic and probabilistic within-site variability were similar for all cases and numbers of tests.

7. Summary, Conclusions, and Recommendations

The research described in this dissertation focuses on the effect of how within-site variability is considered on the interpreted reliability of foundation designs based on load tests. In the course of examining the effect of within-site variability consideration, findings regarding several related topics were revealed. Among the related topics are the magnitude and uncertainty in within-site variability estimates, the accuracy and precision of reliability-based foundation designs based on load tests, and the magnitude of increases in allowable resistance that can be expected from consideration of site-specific load test information. In this chapter, findings from the research are summarized, conclusions from the findings are discussed, and recommendations for practice and for future research are presented.

7.1 Summary of Research and Significant Findings

Previous work to evaluate reliability of foundation designs based on load test information have predominately adopted a Bayesian approach. In the Bayesian approach, estimates of foundation resistance from an empirical design method are used to define a prior distribution of resistance, and the distribution is updated based on site-specific load test information. The distribution of resistance is updated by treating its mean parameter as a probabilistic input defined by its own probability distribution. The mean of the distribution of the mean characterizes the average resistance; the standard deviation of the distribution of the mean characterizes among-site variability. Among-site variability represents the uncertainty associated with the empirical technique of applying results from other sites to a new site. Application of Bayes theorem to the prior distribution of the mean with site-specific load test information produces an updated estimate of average resistance (i.e. the mean of the mean) and an updated estimate of the uncertainty in the mean. In most cases, the uncertainty in the mean is reduced upon updating with site-specific information; in many cases, the reduction is considerable.

The approach for updating the mean described in the previous paragraph is applied in all of the previous work regarding Bayesian updating with load test information (e.g. Kay, 1976;

Baecher and Rackwitz, 1982; Zhang and Tang, 2002). Where the approaches differ is in treatment of the standard deviation parameter of the resistance distribution. The standard deviation parameter of the resistance distribution characterizes within-site variability. Within-site variability results from geologic variation across a site and from variation in construction outcomes from one foundation element to another. The value of within-site variability for a given site and a given set of construction practices is generally fixed but strictly unknown. While all of the previous work reviewed addresses the concept of within-site variability, consideration of within-site variability is divided between two approaches. In the first approach, within-site variability is treated deterministically: a constant value that is considered known and is not updated based on load test results. Deterministic within-site variability has been adopted in most of the previous efforts (e.g. Kay, 1976; Zhang and Tang, 2002; Najjar et al., 2017). In the second approach, within-site variability is treated as a probabilistic parameter with its own probability distribution that is updated based on load test results. The mean of the distribution is the average value of within-site variability, and the standard deviation of the distribution is the variability and uncertainty in the value of within-site variability. Probabilistic within-site variability was introduced for foundation load tests by Baecher and Rackwitz (1982), but has not been commonly applied. With probabilistic within-site variability, the mean and the standard deviation parameters of foundation resistance are jointly updated via Bayes theorem to reflect site-specific load test information. Updating both parameters is more complicated computationally compared with the case of deterministic within-site variability, which updates only the mean parameter. Updating both parameters requires either application of strict distributional requirements (as in Baecher and Rackwitz's approach) or use of numerical methods to perform Bayesian computations (as in this research). A detailed review of previous approaches for Bayesian updating of foundation resistance is presented in Chapter 2.

The focus of this research is to evaluate the difference in the interpreted reliability of foundation designs between deterministic and probabilistic within-site variability. Specifically, the hypothesis is that

- (1) deterministic within-site variability yields greater estimates of $R_{1/1500}$, the 1/1500th value of resistance, than probabilistic within-site variability for prior evaluations without load tests and
- (2) results for probabilistic within-site variability will converge to the true value of $R_{1/1500}$ for all cases whereas results for deterministic within-site variability will converge to the true value only when the prior estimate of within-site variability happens to be correct.

Deterministic within-site variability is based on a constant value of within-site variability, frequently assumed to be represented by a coefficient of variation (COV) of 0.2 (e.g. Zhang and Tang, 2002; Zhang, 2004; Najjar et al., 2017). Probabilistic within-site variability requires a distribution of within-site variability values. For context on the difference in reliability between deterministic and probabilistic within-site variability, it is instructive to evaluate existing information regarding the magnitude and variability of observed values of within-site variability. Data supporting the estimate of COV of 0.2 are fairly limited, originating from about 10 sites. Many of the load tests used to conclude $COV = 0.2$ were terminated at working loads without significant displacement. Information for about one third of the sites originates from dated sources that are no longer available, and supporting information, specifically displacement information, is missing for other sites. The distribution of within-site variability implemented by Baecher and Rackwitz has a mean value that corresponds to $COV = 0.27$. The data used to develop the distribution mostly overlaps with the data used to support Zhang and Tang's deterministic value of 0.2.

Because the published information regarding within-site variability is relatively limited, new data regarding within-site variability of micropile foundations were compiled. Data from micropile load tests at five sites were used to estimate the COV of resistance at each site. For three of the sites, the estimates of COV were based strictly on load tests that achieved failure, with failure loads for the three sites equal to the values reported in literature. At one of the sites, the Davisson Offset Limit (1972) was used to define failure; for the other two sites, the failure load was defined according to the Fuller and Hoy (1970) criterion. For the three sites where tests

achieved failure, COV was estimated as the ratio of unbiased sample standard deviation to sample mean. For another site, the estimate of COV was based on fitting a normal distribution to resistance data based on a proof load test program of 177 micropiles, 17 of which achieved failure. The distribution fitting was based on maximum likelihood estimation with censored observations for the tests that did not achieve failure. For the remaining site, the estimate of COV was based on analysis of interpolated values of applied loading at 1.5 in. displacement for 536 load tests documented by Vanderpool et al. (2002).

Results from each site were compiled, and a normal distribution was fitted to the collection of COV estimates. The normal distribution has a mean parameter of 0.15 and a standard deviation of 0.06. The estimated value of within-site COV for micropiles is considerably less than the previous estimates, i.e. $COV = 0.20$ from Zhang and Tang (2002) and $COV = 0.27$ from Baecher and Rackwitz (1982). A likely explanation for the reduced variability interpreted in this research for micropiles is the use of load test results that achieved failure; the previous estimates were based on load tests that were generally performed with maximum loads equal to working or service loads. The explanation that reduced within-site variability is observed for greater displacement is consistent with the trend observed for interpretations of data from Vanderpool et al. (2002), which suggested ultimate resistance is less variable than stiffness. Importantly, the investigation of within-site variability documented in this research indicates there is significant uncertainty in within-site variability, as evidenced by the standard deviation of 0.06, which corresponds to a COV of the value of within-site variability equal to 0.4. Note the COV of 0.4 represents uncertainty in the value of within-site variability, rather than the magnitude of within-site variability, itself. Additional information regarding the analysis of within-site variability for micropiles is presented in Chapter 3.

To facilitate comparisons of reliability interpretations between deterministic and probabilistic within-site variability, a numerical tool for Bayesian updating and reliability analysis was developed for this research. The tool can be used to perform Bayesian updating with either deterministic or probabilistic within-site variability. Because the tool uses numerical methods to

perform Bayesian updating, there are no distributional requirements for the prior distributions of within-site mean or within-site variability. The prior distributions of the site-specific mean and site-specific within-site variability are input for the tool, along with site-specific load test information.

The prior distributions and load test information are used to perform two numerical analyses. The first numerical analysis of the tool is Bayesian updating, which yields a posterior joint distribution of the site-specific mean and site-specific within-site variability parameters. (For deterministic within-site variability, the within-site variability is a constant value that is not updated, rather than a distribution.) The posterior joint distribution is used as input for the second numerical analysis of the tool, which is a reliability analysis. The posterior joint distribution of the mean and standard deviation is used to develop a posterior predictive distribution that represents the magnitude and variability of new predictions of foundation resistance at a site with load test information. A primary output of the tool is $R_{1/1500}$, the 1/1500th value of resistance, which characterizes the left tail of the posterior predictive resistance distribution. In this research, $R_{1/1500}$ is used to evaluate the reliability implications of Bayesian updating with deterministic and probabilistic within-site variability under different scenarios. Additional details of the numerical tool, including the equations used to perform Bayesian updating, are presented in Chapter 4.

Two sets of analyses were completed with the numerical tool, one with “perfect” sampling and one with random sampling. In perfect sampling, the sample statistics perfectly match the parameters of the distribution from which the data are sampled, i.e. the sample mean equals the site-specific mean and the sample standard deviation equals the within-site variability parameter. Perfect sampling results, the topic of Chapter 5, were used to evaluate the hypothesis and to evaluate trends in posterior parameters without the effect of random sampling noise. Random sampling results, the topic of Chapter 6, were used to evaluate the magnitude of errors in Bayesian estimates of $R_{1/1500}$, with a particular focus on how errors resulting from deterministic within-site variability compare with errors from probabilistic within-site variability. Random sampling was also used to examine the difference between posterior estimates of $R_{1/1500}$ and

prior estimates of $R_{1/1500}$. The difference is an important practical measure of the reliability benefits associated with load test information.

Twelve cases representing a variety of different prior distributions of the within-site mean, μ_{within} , and within-site variability, σ_{within} , were evaluated with perfect sampling. The distributions varied both in total variability (the combined effect of within-site and among-site variability) and in the ratio of within-site variability to among-site variability. For each set of prior distributions of μ_{within} and σ_{within} , nine different “true” site distributions were defined using all combinations of the tenth, fiftieth, and ninetieth percentiles of the distributions of μ_{within} and σ_{within} . All true site distributions were evaluated with the numerical tool using both deterministic and probabilistic within-site variability. The analyses were performed with different numbers of load tests (0, 1, 2, 3, 4, 5, 7, 10, 20, 30, 40, 50, 200, 400, 800) to evaluate the trends in the parameters with increasing numbers of tests.

Results of the analyses support the hypothesis statements regarding analyses without load tests and with many load tests. For all analyses without load tests, deterministic within-site variability produces estimates of $R_{1/1500}$ that are greater than estimates from probabilistic within-site variability. Results suggest that $R_{1/1500}$ from deterministic within-site variability is greater because it does not consider uncertainty in within-site variability; neglecting the uncertainty in within-site variability reduces the total variability of the predictive distribution of resistance. For analyses with many load tests, results also support the hypothesis: all estimates of $R_{1/1500}$ for probabilistic within-site variability converge to the true value of $R_{1/1500}$, while for deterministic within-site variability, estimates of $R_{1/1500}$ converge to the true value only when the prior estimate of within-site variability happens to be equal to the true value. Errors in estimates of $R_{1/1500}$ for the deterministic case are conservative (i.e. underestimates) when the within-site variability is overestimated and unconservative (i.e. overestimates) when the within-site variability is underestimated.

Additional support for the hypothesis is provided by the trends in the posterior distributions of μ_{within} and σ_{within} . Results for probabilistic within-site variability achieve convergence to the true value of $R_{1/1500}$ because the posterior parameters of μ_{within} and σ_{within} all converge to the true values. Deterministic within-site variability does not achieve convergence to the true value of $R_{1/1500}$ because while the posterior parameters of μ_{within} converge to the true values, posterior parameters of σ_{within} do not. In general, the mean parameters of both μ_{within} and σ_{within} (i.e. μ_μ and μ_σ) achieve convergence at fewer numbers of tests than the standard deviation parameters (i.e. σ_μ and σ_σ).

The numerical tool was used with random sampling to evaluate the magnitude of errors in Bayesian estimates of $R_{1/1500}$. The evaluations are based on five of the 12 sets of prior distributions of μ_{within} and σ_{within} evaluated for the perfect sampling trials. For each case, 1,000 sites were generated by pairing one randomly sampled value of μ_{within} and one randomly sampled value of σ_{within} from the respective prior distributions. The pair of parameters define the true distribution of resistance for the random site. Load test results were randomly sampled from the true distribution for each site. Bayesian updating was performed using the randomly sampled load test results, and the posterior results of Bayesian updating were used to estimate $R_{1/1500}$ from the posterior predictive distributions of resistance.

Evaluation of the accuracy, precision, and errors associated with the posterior estimates of $R_{1/1500}$ compared with the true value of $R_{1/1500}$ are based on the ratio of estimated to true $R_{1/1500}$. A value of the ratio exactly equal to 1 indicates a perfectly accurate prediction from the posterior; ratio values less than 1 indicate conservative estimates and ratio values greater than 1 indicate unconservative estimates. The variability of a collection of ratio values is used to interpret the precision of a set of $R_{1/1500}$ estimates, with more variable results indicating less precision.

For all cases, results indicate both deterministic and probabilistic within-site variability produce, on average, conservative estimates of $R_{1/1500}$ for the prior evaluation (without load tests). For both deterministic and probabilistic within-site variability, the accuracy of estimates of

$R_{1/1500}$ increases (i.e. conservatism is reduced) for greater numbers of load tests, eventually to the position where estimated $R_{1/1500}$ is, on average, nearly equal to the true value. The precision of the estimates of $R_{1/1500}$ differs for deterministic versus probabilistic within-site variability. For the prior, both deterministic and probabilistic within-site variability produce relatively imprecise estimates of $R_{1/1500}$, as indicated by high variability in the distributions of the ratio of estimated to true $R_{1/1500}$. The precision of the deterministic estimates of $R_{1/1500}$ does not improve with additional load tests. In contrast, for probabilistic within-site variability, the variability decreases with increasing numbers of load tests, approaching zero for large numbers of tests. The increase in precision noted for probabilistic within-site variability is attributed to the Bayesian updating of within-site variability.

The implications of the finding that estimates of $R_{1/1500}$ with deterministic within-site variability produce greater errors than estimates with probabilistic within-site variability were investigated by dividing outcomes into categories of satisfactory, overreliable, and underreliable. The categories are defined with respect to the ratio of estimated to true $R_{1/1500}$. Outcomes with values of the ratio between 0.9 and 1.1 are deemed satisfactory, outcomes with values of the ratio less than 0.9 are deemed overreliable, and outcomes with values of the ratio greater than 1.1 are deemed underreliable.

For the prior, the proportion of outcomes depends on the ratio of within-site variability to among-site variability. When the prior predictive variability is mainly due to among-site variability, the outcomes are predominately overreliable. When the prior predictive variability is mainly due to within-site variability, significantly more underreliable outcomes are noted. In the random sampling analysis, both cases have the same prior predictive distribution, so the estimates of $R_{1/1500}$ are the same. The difference in outcomes is therefore due to the true value of $R_{1/1500}$, which is significantly less for the case dominated by within-site variability. The true value of $R_{1/1500}$ is less for the case of significant within-site variability because within-site variability is the true, inherent variability at the site, whereas among-site variability is uncertainty in the mean – a

Bayesian construct that is critical in reliability evaluations but which does not manifest itself as true variability at any particular site.

For both deterministic and probabilistic within-site variability, outcomes for small numbers of load tests are dominated by overreliable estimates of $R_{1/1500}$. As additional tests are considered, the proportion of satisfactory outcomes increases as the proportion of overreliable outcomes decreases. For probabilistic within-site variability, the proportion of satisfactory cases converges to 100 percent or near 100 percent for all cases. Deterministic within-site variability outcomes do not converge to 100 percent satisfactory for any of the cases; for the cases with prior predictive COV of 0.5, deterministic within-site variability achieves 80 percent satisfactory cases when within-site variability is low and less than 40 percent satisfactory cases when within-site variability is high.

For all cases with probabilistic within-site variability, the proportion of underreliable outcomes is relatively small, less than 15 percent for any number of tests, and the proportion of underreliable outcomes eventually decreases to near zero, although the decrease to zero requires as many as 200 load tests. In contrast, the proportion of underreliable cases for deterministic within-site variability with small numbers of load tests was 10 percent for the case with a low ratio of within-site variability to among-site variability and 30 percent for the case with a high ratio of within-site variability to among-site variability. In addition, and importantly, the proportion of underreliable cases generally increased with additional tests for deterministic within-site variability. For all cases considered and for all numbers of tests, deterministic within-site variability produces a greater proportion of underreliable outcomes than probabilistic within-site variability. The proportion of underreliable outcomes for deterministic within-site variability is greater than that for probabilistic within-site variability by about a factor of two for small numbers of tests. The difference in underreliable outcomes between deterministic and probabilistic within-site variability increases as additional tests are considered. As additional tests are considered, results for probabilistic within-site variability converge to the true value of within-site variability

whereas errors associated with deterministic within-site variability persist and result in underreliable outcomes.

In addition to the analyses performed to evaluate the accuracy, precision, and errors associated with posterior estimates of $R_{1/1500}$, analyses were also completed to evaluate the increase in estimated $R_{1/1500}$ from the prior evaluations without load tests to the posterior evaluations with load tests. The results indicate the average increase in $R_{1/1500}$ from the prior to posterior depends on many factors but is dominated by the ratio of within-site variability to among-site variability. When the ratio of within-site variability to among-site variability is small (i.e. the total variability results primarily from among-site variability), equal to 0.2, the posterior estimate of $R_{1/1500}$ after one load test is on average greater than the prior estimate of $R_{1/1500}$ by a factor of three. When the ratio of within-site to among-site variability is greater, equal to 1.0, the posterior estimate of $R_{1/1500}$ after one load test is greater than the prior by a factor of 1.25. The finding that load testing increases $R_{1/1500}$ considerably more when total variability is primarily caused by among-site variability rather than within-site variability is consistent with the findings regarding posterior parameters. The analysis of posterior parameters revealed that among-site variability is effectively diminished by load testing, and that among-site variability is generally diminished with relatively few tests; in contrast, the magnitude of within-site variability generally is not reduced by load testing. The implications of this finding are discussed in the recommendations for practice below.

The average increase in $R_{1/1500}$ from the prior to posterior also increased with increasing nominal prior predictive COV , but the effect was less significant than that of the ratio of within-site to among site variability. The number of load tests has a relatively modest influence on the average increase in $R_{1/1500}$ from the prior to posterior. For the case of relatively low within-site variability, most of the increase in the posterior estimate of $R_{1/1500}$ is observed for the first test, with diminishing returns for additional testing. Finally, the effect of deterministic-vs.-probabilistic within-site variability on the increase in estimated $R_{1/1500}$ is limited. The difference between ratios

of posterior to prior $R_{1/1500}$ for deterministic and probabilistic within-site variability were similar for all cases and numbers of tests.

7.2 Conclusions

Findings from this research highlight significant differences in outcomes between deterministic within-site variability, which treats within-site variability as a known, constant parameter, and probabilistic within-site variability, which treats within-site variability as an uncertain parameter subject to updating based on load test results. The findings support the hypothesis: (1) deterministic within-site variability produces greater estimates of $R_{1/1500}$ than probabilistic within-site variability for evaluations without load tests and all other parameters equal and (2) for evaluations with many load tests, the posterior estimate of $R_{1/1500}$ with probabilistic within-site variability converges to the true value of $R_{1/1500}$ while the posterior estimate of $R_{1/1500}$ with deterministic within-site variability generally does not.

The differences in estimates of $R_{1/1500}$ between deterministic and probabilistic within-site variability produce considerable consequences on satisfaction of the target reliability. Simulations comparing the posterior estimates of $R_{1/1500}$ to the true value were characterized as underreliable when the estimated value of $R_{1/1500}$ was more than ten percent greater than the true value. For all prior distributions evaluated, which included 12 combinations of site parameters ranging from relatively high capacity and uniform to relatively low capacity and variable, deterministic within-site variability produces a significantly greater proportion of underreliable outcomes than probabilistic within-site variability. For small numbers of tests, the proportion of underreliable outcomes for deterministic within-site variability ranged from 10 percent with low within-site variability to 30 percent with high within-site variability. In all cases, the proportion of underreliable cases from deterministic within-site variability was about twice the proportion from probabilistic within-site variability. The prevalence of underreliable outcomes for deterministic within-site variability is attributed to its neglect of uncertainty in within-site variability. As the evaluation of

within-site variability from load test data in Chapter 3 documented, estimates of within-site variability are, in fact, subject to considerable uncertainty.

The analysis of micropile load test results from five different sites to develop a distribution of within-site variability values validates the premise of the hypothesis. The distribution is relatively variable (i.e. the value of within-site variability is, itself, variable), which demonstrates that there is, in fact, considerable variability and uncertainty in the value of within-site variability. The variability and uncertainty of estimates of within-site variability means that within-site variability is truly a probabilistic quantity, and it is therefore worthwhile to examine the consequences of treating within-site variability deterministically.

The distribution of within-site variability values developed for micropile foundation systems is also useful in its own right. The distribution indicates the value of within-site variability commonly cited in literature corresponding to $COV = 0.2$ is an overestimate, at least for micropile foundations, which were observed across five sites to have an average within-site COV of 0.15. In addition, the data analysis used to develop the distribution of within-site variability suggests the stiffness of foundation load-displacement curves is more variable than the ultimate resistance.

To evaluate the value of foundation load testing from a reliability perspective, it is instructive to consider the four parameters used in the Bayesian construction of predictive distributions of resistance: μ_μ , the mean of within-site mean; σ_μ , the standard deviation of within-site mean; μ_σ , the mean of within-site variability; and σ_σ , the standard deviation of within-site variability. Of the four, only μ_μ and μ_σ define the “true” distribution of resistance at a particular site. The other two parameters, σ_μ and σ_σ , are used to represent the uncertainty associated with applying empirical results (values of average resistance and within-site variability, respectively) to a new site. Both of the uncertainty parameters, σ_μ and σ_σ , are zero for the true distribution of resistance at a site. The uncertainty parameters are non-zero for empirical design methods. For highly variable empirical datasets – a description that applies to most empirical foundation design methods – the uncertainty parameters can dominate the total variability of the prior predictive resistance.

From a reliability perspective, the objective of load testing is to reduce the variability of the predictive distribution. Consider the case of unbiased prior information: the prior estimates of μ_μ and μ_σ are correct for an “average” site. In this case, the results of load testing for any given site are just as likely to reduce the value of μ_μ as they are to increase it; likewise, the results of load testing are just as likely to increase the value of μ_σ as they are to decrease it. Accordingly, the more dependable reliability benefit of load testing is to reduce the values of σ_μ and σ_σ . As site-specific information regarding the mean and within-site variability of resistance is collected through load testing, the uncertainty in the parameters is reduced, and the variability of the predictive distribution is reduced in turn.

Results from this research support this explanation of load test reliability benefits. Moreover, the results suggest that the practically achievable load test benefit is to reduce σ_μ , the effect of among-site variability. The results indicate significant reductions in σ_μ can be achieved for one load test. In fact, when the ratio of within-site variability to among-site variability is 0.2 and the nominal prior predictive *COV* is 0.5, the prior value of $R_{1/1500}$ was increased by a factor of three, on average, after Bayesian updating from one randomly sampled load test. While a ratio of within-site to among-site variability of 0.2 and total *COV* of 0.5 was the most favorable case of any evaluated with respect to load test reliability benefits, the conditions are certainly feasible. Three of the five sites considered in the evaluation of micropile within-site variability had within-site variability corresponding to about *COV* = 0.1. Additional simulations suggest the reliability benefit of load testing is significantly greater when total variability is dominated by among-site variability rather than within-site variability. Reducing σ_σ , the uncertainty in within-site variability, can have a significant benefit, but is practically unachievable in many cases since many load tests are required to produce a significant reduction in the uncertainty surrounding within-site variability.

For circumstances when the predictive distribution of resistance is dominated by within-site variability, the value of load testing may be limited. Performing a great number of load tests at a site can effectively characterize the value of within-site variability (and therefore reduce σ_σ), but performing the load tests will not reduce the magnitude of within-site variability, itself, except

inasmuch as the experience of installing test foundations may improve the precision of construction outcomes.

The increase in $R_{1/1500}$ due to load testing is most significant for the first test, with diminishing increases for subsequent tests. This finding is largely consistent with common load testing practices for most types of deep foundations; designs based on load tests commonly include only one test.

7.3 Recommendations for Practice

Each of the major conclusions leads to a corresponding recommendation for foundation engineering practice. Because probabilistic within-site variability yields more satisfactory and considerably fewer underreliable outcomes compared with deterministic within-site variability, it should be used for any reliability-based evaluation of designs based on site-specific load tests. The simplicity of deterministic within-site variability is appealing, but the numerical methods used in this research are not overly-complicated. An alternative to the numerical techniques used here is the approach by Baecher and Rackwitz, which involves equations that can be computed with a calculator.

Until and unless additional studies of within-site variability for other foundation systems are performed, the distribution developed in this research for micropiles (with average $COV_R = 0.15$ and standard deviation of $COV_R = 0.05$) is recommended for use in Bayesian analysis of load test information. Although the distribution is specific to micropiles, the distribution is based on results of tests achieving failure or at least significant displacements. Evidence suggests the commonly used value of within-site variability corresponding to $COV_R = 0.2$ is an overestimate, likely because few of the load tests that serve as its basis achieved failure.

The finding that site-specific load test information is of most value from a reliability perspective when within-site variability is small relative to among-site variability is also instructive. Foundation engineers should consider the sources of variability and uncertainty in their designs and formulate their design methodology appropriately. For a given foundation type, if a precise

empirical method is available (i.e. among-site variability is low) and construction outcomes or the design site are known to be variable, load test information is not likely to be of significant value from a reliability perspective. Precise empirical methods are sometimes available, for instance when an empirical method is based on a considerable amount of data for a specific type of foundation in a specific set of geologic conditions. However, if empirical methods are subject to significant uncertainty and within-site variability is relatively small, load test information can greatly increase reliability-based design resistance. Highly uncertain empirical methods are the norm, not the exception in foundation engineering, and the analysis of micropile within-site variability data from this research suggests within-site variability is relatively low, at least compared to among-site variability. It is therefore reasonable to expect that for many foundation design scenarios, collecting load test information is worthwhile.

7.4 Recommendations for Future Research

This research revealed several topics that would be worthy of future study. First, to build on the micropile within-site variability analysis presented here, it would be useful to gather and evaluate information regarding within-site variability for other types of foundation systems. Ground anchors are a feasible target for such an analysis, since, like micropiles, they are commonly load tested. Driven piles would be a particularly useful type of deep foundation for which to characterize within-site variability, in part because of their prevalence among foundation systems. Moreover, comparison of within-site variability between micropiles and driven piles could facilitate an interesting comparison of variability between drilled and driven foundations.

It would also be worthwhile to perform an evaluation of within-site variability that considers side resistance and end resistance separately. The evaluation of micropile load test data in Chapter 3 is based on load-displacement results at the top of micropiles, without consideration of the force distribution within the foundation. Considering the results of this research, knowledge of the relative variability between side and end resistance would be useful for planning load test programs. For instance, if within-site variability of end resistance were revealed to be significantly greater than that of side resistance, load test programs could perhaps

be more effective by focusing on measuring side resistance. Considering the prevalence of bi-directional testing for drilled shafts and, to a lesser extent, auger-cast piles, such a conclusion would be meaningful since bi-directional tests typically only achieve ultimate conditions in either side resistance or end resistance, not both.

As noted in Chapter 6, the effect of errors in the prior distributions of average resistance and within-site variability were not evaluated in this research; the random sampling was unbiased. Evaluation of posterior outcomes with biased sampling would be useful. Because bias in the prior distributions would be associated with greater errors between the prior and true conditions, it is likely that evaluation of biased priors would reveal even greater differences in outcomes between deterministic and probabilistic within-site variability.

Finally, this research and the previous efforts it builds on provide a framework for combining empirical and site-specific information for reliability-based design, but foundation load testing is hardly the only application of such an approach. It would be worthwhile to apply the methods from this effort to other geotechnical and civil engineering systems: anchored earth retention systems, levees and dams with monitoring systems, solid waste disposal sites, and water distribution networks, to name just a few. Considering the increasing emphasis on data collection and monitoring, an emphasis on thorough and precise consideration of the site-specific data with respect to past performance seems prudent.

References

AASHTO LRFD Bridge Design Specifications (2014), 7th ed. with 2015 and 2016 Interim Revisions, American Association of State Highway and Transportation Officials, Washington, D.C.

Ahsanuzzaman, M., W.A. Smith, M.A. Gabr, R.H. Borden (2019), "Quantifying the Influence of Construction Parameters on Hollow-Bar Micropiles Pullout Capacity in Sandy Soil," *Proceedings of the 2019 Geo-Congress*, Philadelphia, PA, ASCE Geotechnical Special Publication 307, 8 pp.

Allen, T.M. (2005), *Development of Geotechnical Resistance Factors and Downdrag Load Factors for LRFD Foundation Strength Limit State Design*, Federal Highway Administration Publication No. FHWA-NHI-05-052.

Ang, A. H.-S. and W. H. Tang (2007), *Probability Concepts in Engineering Planning and Design*, John Wiley & Sons, New York.

Armour, T., P. Groneck, J. Keeley, and S. Sharma (2000), *Micropile Design and Construction Guidelines: Implementation Manual*, Federal Highway Administration Publication No. FHWA-SA-97-070.

Baecher, G.B., and J.T. Christian (2003), *Reliability and Statistics in Geotechnical Engineering*, John Wiley and Sons, 605 pp.

Baecher, G.B., and R. Rackwitz (1982), "Factors of safety and pile load tests," *International Journal for Numerical and Analytical Methods in Geomechanics*, Vol. 6, No. 4, 409-424.

Bain, L.J. and M. Engelhardt (1992), *Introduction to Probability and Mathematical Statistics*, Brooks/Cole, Second Edition, 644 pp.

Barker, R.M., J.M. Duncan, K.B. Rojiani, P.S.K. Ooi, C.K. Tan, and S.G. Kim (1991), *Manuals for the Design of Bridge Foundations: Shallow Foundations, Driven Piles, Retaining Walls and Abutments, Drilled Shafts, Estimating Tolerable Movements, and Load Factor Design Specifications and Commentary*, NCHRP Report 343, Transportation Research Board, Washington, D.C.

Bea, R.G., Z. Jin, C. Valle, and R. Ramos (1999), "Evaluation of reliability of platform pile foundations," *Journal of Geotechnical and Geoenvironmental Engineering*, Vol. 125, No. 8, 696-704.

Chellis, R.D. (1961), *Pile Foundations*, 2nd Ed., McGraw-Hill Book Co., Inc., New York, NY.

Cureton, E.E. (1968), "Unbiased Estimation of the Standard Deviation," *The American Statistician*, Vol. 22, No. 1, p. 22.

Dapp, S., D. Brown, D.S. Graham, and R.T. McGillivray (2013), "Selmon Expressway in Tampa, Florida: Case History of Drilled Shaft Design for Extreme Variability," Presentation to the 38th Annual Conference on Deep Foundations, Deep Foundations Institute, Phoenix, AZ.

Drbe, O.F.E. and M.H. El Naggar (2014), "Compressive Behaviour of Hollow Core Micropiles in Cohesive Soil," *Proceedings of the DFI/EFFC 11th International Conference on Piling and Deep Foundations*, Stockholm, Sweden, 10 pp.

Evangelista, A., A. Pellegrino, and C. Viggiani (1977), "Variability among piles of the same foundation," *Proceedings of the 9th ICSMFE*, Tokyo, 493-500.

Fuller, F.M. and H.E. Hoy (1970), "Pile Load Tests Including Quick-load Test Method Conventional Methods and Interpretations," Highway Research Board, HRB(333), pp. 78-86.

Gahreman, B. (2009), "Design and load test of micropiles in a power plant in Missouri", Proceedings of the 34th Annual Conference on Deep Foundations, 2009, Kansas City, MO, USA, Deep Foundation Institute.

Gómez, J., A. Cadden and D.A. Bruce (2003), "Micropiles founded in rock. Development and evolution of bond stresses under repeated loading", *presented at 7th International workshop on micropiles*, International Society of Micropiles.

Gómez, J.E., C.J. Rodriguez, H.D. Robinson, J. Mikitka and L. Keough (2007), "Hollow core bar micropile - design parameters interpreted from 260 load tests", *Proceedings of the 32nd Annual Conference on Deep Foundations*, 2007, Colorado Springs, CO, USA, Deep foundation institute.

Gómez, J.E., C.J. Rodriguez, H.D. Robinson, J. Mikitka and L. Keough (2007), "Installation and testing of 260 hollow core bar micropiles", *Proceedings of the 32nd Annual Conference on Deep Foundations*, 2007, Colorado Springs, CO, USA, Deep foundation institute.

Gómez, J.E., C.J. Rodriguez, H.D. Robinson, J. Mikitka and L. Keough (2008), "Hollow core bar micropile - design parameters interpreted from 404 load tests", *Proceedings of the 33rd Annual and 11th International Conference on Deep Foundations*, 2008, New York, NY, USA, Deep foundation institute.

Gómez, J. E., C. J. Rodriguez, H. D. Robinson, J. Mikitka, and L. Keough (2008), "Bond strength of hollow-core bar micropiles." *Proceedings of International conference on case histories*.

Han, J., and S. Ye (2006), "A field study on the behavior of a foundation underpinned by micropiles." *Canadian Geotechnical Journal*, vol. 43, no. 1, pp: 30-42.

Han, J., and S. Ye (2006), "A field study on the behavior of micropiles in clay under compression or tension." *Canadian Geotechnical Journal*, vol. 43, no. 1, pp: 19-29.

Hirany, A. and F.H. Kulhawy (2002), "On the Interpretation of Drilled Foundation Load Test Results," *Deep Foundations 2002*, ASCE, pp. 1018-1028.

Huaco, Daniel R., John J. Bowders, and J. Erik Loehr (2012), "Method to develop target levels of reliability for design using LRFD." *In Compendium of Papers for Transportation Research Board 91st Annual Meeting*.

Jeon, S.S. and F.H. Kulhawy (2001), "Evaluation of Axial Compression Behavior of Micropiles," *Foundations and Ground Improvement*, ASCE GSP 113, Ed. T.L. Brandon, pp. 460-471.

Juang, C.H. and J. Zhang (2017), "Bayesian Methods for Geotechnical Applications – A Practical Guide," *Proceedings of Geo-Risk 2017*, ASCE GSP 286, pp. 215-246.

Kay, J.N. (1976), "Safety factor evaluation for single piles in sand," *Journal of Geotechnical Engineering*, ASCE, Vol. 102, No. 10, 1093-1108.

Komornik, A., Wiseman, G., and Zeitlan, J.G. (1971), "Comparison of Test Performance with Predicted Behavior for Piles Driven in Sand," *Proceedings of the Fourth Asian Regional Conference on Soil Mechanics and Foundation Engineering*, Bangkok, Thailand, pp. 281-286.

Loehr, J.E. A.Z. Boeckmann, P.L. Speckman, and D. Ding (2013), *Procedures for Establishing Geotechnical Design Parameters from Two Data Sources*, Missouri Department of Transportation, Report cmr14-002, 22 pp.

Loehr, J.E., A.J. Lutenecker, B.L. Rosenblad, and A.Z. Boeckmann (2016), *Geotechnical Site Characterization, Geotechnical Engineering Circular No. 5*, FHWA, Report FHWA-NHI-16-072, (in preparation).

Loehr, J.E., D. Ding, and W.J. Likos (2015), "Effect of number of soil strength measurements on reliability of spread footing designs," *Transportation Research Record: Journal of the Transportation Research Board*, TRR No. 2511, Transportation Research Board, pp. 37-44.

Najjar, S.S. and R.B. Gilbert (2009a), "Importance of lower-bound capacities in the design of deep foundations," *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE, Vol. 135, No. 7, 890-900.

Najjar, S.S. and R.B. Gilbert (2009b), "Importance of proof-load tests in foundation reliability," *Proceedings of IFCEE 2009: Contemporary Topics in In-Situ Testing, Analysis, and Reliability of Foundations*, ASCE Geotechnical Special Publication No. 186, 340-347.

Najjar, S., G. Saad, and Y. Abdallah (2017), "Reliability-based framework for designing test programs for piles," *Proceedings of Geo-Risk 2017*, ASCE GSP 285, 101-112.

Paikowsky, S. G., Birgisson, B., McVay, M., Nguyen, T., Kuo, C., Baecher, G., Ayyub, B., Stenersen, K., O.Malley, K., Chernauskas, L., and O'Neill, M. (2004), *Load and Resistance Factor Design (LRFD) for Deep Foundations*, NCHRP Report 507, Transportation Research Board of the National Academies, Washington, D.C., 126 pp.

Petermann, H., E. Lackner, and W. Schenek (1958), "Tragfähigkeit von Pfählen, Großversuche und ihre Auswertung," *Berichte aus der Bauforschung*, H. 49, W. Ernst & Sohn, Berlin.

Rackwitz, R., and Fiessler, B. (1978), "Structural reliability under combined random load sequences," *Computers and Structures*, vol. 9, pp. 489-494.

Reddy, S.C. (2014), *Ultimate and serviceability limit state reliability-based axial capacity of deep foundations*. PhD dissertation submitted to the Oregon State University, 409 pp.

Sidi, I.D. (1986), *Probabilistic prediction of friction pile capacities*, Ph.D. dissertation, University of Illinois.

Sidi, I.D. and W.H. Tang (1987), "Updating friction pile capacity in clay," *Proceedings of the Fifth International Conference on Applications of Statistics and Probability in Soil and Structural Engineering*, University of British Columbia, Vancouver, 938-946.

Taboga, M. (2017), *Lectures on Probability Theory and Mathematical Statistics*, Third Ed., CreateSpace Independent Publishing Platform, 670 p.

Tang, W.H. (1971), "A Bayesian evaluation of information for foundation engineering design," *Proceedings of the First International Conference on Applications of Statistics and Probability to Soil and Structural Engineering*, Hong Kong, 173-185.

Vanderpool, W.E., D.A. Bruce, E.E. Rinne, F.H. Kulhawy (2002), "Micropile performance variability: a fundamental case study", *Proceedings of the 27th Annual Conference on Deep Foundations, 2002, San Diego, CA, The Time Factor in Design and Construction of Deep Foundations*, Deep foundation institute.

Whitman, R.V. (1984), "Evaluating calculated risk in geotechnical engineering," *Journal of Geotechnical Engineering*, ASCE, Vol. 110, No. 2, pp. 143-188.

Wolosick, J.R. (2009), "Ultimate micropile bond stresses observed during load testing in clays and sands," *Proceedings of the 2009 International Foundation Congress and Equipment Expo (IFCEE): Contemporary topics in deep foundations*, pp. 12-22.

Zhang, J. (2009), *Characterizing geotechnical model uncertainty*, PhD Thesis, The Hong Kong University of Science and Technology, Hong Kong SAR, China.

Zhang, L. (2004), "Reliability verification using proof pile load tests," *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE, Vol. 130, No. 11, 1203-1213.

Zhang, L. and W. Tang (2002), "Use of load tests for reducing pile length," *Proceedings of the International Deep Foundations Congress*, Feb. 14-16.

Zhang, L., W. Tang, and C. Ng (2001), "Reliability of axially loaded driven pile groups," *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE, Vol. 127, No. 12, 1051-1060.

Vita

Andrew Boeckmann is a research engineer in the Civil and Environmental Engineering Department at the University of Missouri. His research interests include slope stability, deep foundations, load and resistance factor design and reliability, lateral loading of foundations, soil nails and geotechnical asset management. Prior to joining MU, he worked for five years as a consulting geotechnical engineer for URS Corp. in St. Louis, where he performed design and analysis for large projects, including post-Hurricane Katrina levee design in New Orleans, the new Interstate 64 project in St. Louis, and the 2011 Missouri River high-water event at Eppley Airfield in Omaha. Before working at URS Corp., he obtained his bachelor's and master's degrees in civil engineering. His undergraduate and graduate research with Erik Loehr involved large-scale physical models of earth slopes stabilized with micropiles. He is a registered professional engineer in the state of Missouri.