SPATIAL VARIABILITY OF SHEAR WAVE VELOCITY AND ITS EFFECTS ON SEISMIC GROUND RESPONSE

by

Laura Luna

Submitted in partial fulfillment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
August 2021

Table of Contents

List of Tables

List of Figures

Abstract

Since it is unfeasible to sample every point at a site, there will always be a level of uncertainty in the mechanical properties that are used in geotechnical design. As the geotechnical community transitions towards reliability and risk-based designs, the formal treatment of these uncertainties throughout the analysis process becomes increasingly important. This study has two objectives. The first is to quantify the spatial variability of shear wave velocity based on 206 seismic cone penetration tests available for sites in British Columbia, Canada. The second is to assess how this spatial variability affects the response of a soil mass subject to earthquake ground motions.

The statistical properties of shear wave velocity were estimated using a multiplicative form, where the random shear wave velocity ($V(d)$) was expressed as the product of a deterministic trend ($\hat{v}(d)$) and a lognormal random variable (Y). The distribution of Y was estimated using the method of moments, and the correlation length was estimated using two approaches, a commonly used *direct-fitting* method, as well as a *bias-matched* method. Additionally, a comparison is presented to the first-order, auto-regressive method to randomize shear wave velocity proposed by Toro (1995). Finally, a discrete-time, two-state Markov chain is used to generate realizations of soil layering, modeling the transitions between *clay-like* and *sand-like* materials.

Several probabilistic, equivalent-linear ground response analyses (GRA) were completed to assess how the spatial variability of shear wave velocity affects the peak ground acceleration, cyclic stress ratio, and spectral acceleration response spectra. In general, randomizing shear wave velocity was found to result in lower mean stresses and accelerations when compared against the results obtained from a deterministic approach using the mean shear wave velocity. The sensitivity of the results to the random field parameters was also explored. The distribution of peak ground acceleration (PGA) and cyclic stress ratio (CSR) were found to be most sensitive to the coefficient of variation of shear wave velocity, with correlation length and correlation anisotropy having a smaller influence on the results.

Acknowledgements

The work presented here would not have been possible without the invaluable support of many people.

First and foremost, I would like to extend my deepest gratitude to my supervisor, Dr. Gordon Fenton, for his support and encouragement throughout the completion of my master's degree. I am also extremely grateful for the mentorship and guidance provided by Dr. Roberto Olivera over the past five years. I feel incredibly lucky to have such wonderful mentors, and will be forever thankful for their support.

I would also like to thank my committee members, Dr. Farzaneh Naghibi and Dr. Craig Lake, for their feedback and interest in my work. Finally, thank you to Dr. Serguei Iakovlev, Dr. Wendy Gentleman, and Claire Chisholm for their guidance during my time at Dalhousie.

To Anthony, Gis, Juan, Tow, and the Haliente family: thank you for keeping me sane through a pandemic. A special thanks to Saif for being an infinite source of happiness. Pero mas que todo, gracias a mis papas y a mi hermana, a quien les debo todos los logros que he alcanzado y los que quedan por alcanzar.

Chapter 1

Introduction

1.1 Background

Geotechnical engineers must contend with the fact that natural soils, unlike most other engineering materials, are not man-made and therefore do not have mechanical properties that can be manufactured to meet an established set of standards. Whereas the properties of most construction materials, such as steel and concrete, can be controlled throughout the manufacturing process, there is little to no control over the soil or rock properties at any given site. Even more challenging is the fact that the design of geotechnical systems must proceed with incomplete information, since it is not feasible to sample every single point at a site, and yet soil properties are known to vary spatially. For this reason, there will always be a significant level of uncertainty in the mechanical properties that are used throughout geotechnical analyses.

Historically, the process of site characterization involves the selection of *characteristic* parameters for individual soil layers, which are assumed to represent the overall behavior of the system (Fenton et al., 2008). This is a *deterministic* approach towards site characterization, since the uncertainty in the estimate is not considered directly. What a practicing engineer considers to be the *characteristic* value may change from person to person, but it is generally taken as a cautious estimate of the mean in North America. Of course, what one person considers risky may seem overly-conservative to another, since each engineer has a different tolerance to risk. In addition, some geotechnical problems are preferably represented through some quantile of the distribution instead of the mean. For example, liquefaction triggering analyses have, by convention, adopted the 33^{rd} percentile as the *characteristic* value.

As the geotechnical community transitions towards reliability and risk-based designs, the formal treatment of uncertainties throughout the analysis process becomes of critical importance. In recent years, random field theory has been combined with numerical geotechnical analysis (such as finite element method or limit equilibrium

method) to incorporate the uncertain nature of ground properties into the design process. This has been completed for a variety of problems such as slope stability, bearing capacity, and settlement problems (for example: Fenton et al. (2003), Fenton and Griffiths (2003), and Christodoulou et al. (2020)). However, there has been limited adoption in earthquake geotechnical engineering applications.

1.2 Motivation

This work is inspired by the challenges associated with the seismic design of flood-protection dikes in the Lowe Mainland of British Columbia, a region that is subject to a combination of high seismic and flood hazards. As the region densifies and increasingly relies on earthen dikes as a major flood-defense mechanism, there are growing concerns about the ability of these geotechnical systems to survive major earthquake events.

A regional assessment of flood vulnerability estimates that a major flood could result in losses of up to $22.9 billion dollars (Northwest Hydraulic Consultants, 2016). If this happens, it would become the costliest natural disaster in Canadian history, and its risk of happening is expected to worsen due to climate change (Fraser Basin Council, 2016). Furthermore, a study of current dike infrastructure concluded that few dikes meet provincial standards. In detail, 53% of dikes were judged to be seismically unstable and another 29% were not assessed due to insufficient information (Northwest Hydraulic Consultants and Thurber Engineering Ltd., 2015)

The design of flood-protection infrastructure in the Lower Mainland is best tackled through a risk-based approach due to the uncertain nature of flood and seismic hazards, the identified shortcomings of the dike system, and the high consequences of failure. A risk-based design provides advantages over a deterministic one as it allows practitioners to directly incorporate uncertainty into the design process, prioritize spending on higher risk elements, and find optimal solutions that balance cost and reliability.

An important step in any risk-based design is to understand how the uncertainties in the design affect the probability of failure of the system. This is particularly challenging for dike systems, as the dikes can extend for hundreds of kilometers over spatially varying soils.

1.3 Objectives and Structure

There are two main goals in this study. The first is to characterize the spatial variability of shear wave velocity based on seismic cone penetration tests available at sites in British Columbia, Canada. The second, is to explore the effects of spatial variability of shear wave velocity on the seismic response of a soil mass.

To achieve these objectives, the following work is completed in this study:

- In Chapter 2, a literature review is presented which provides a brief overview of the techniques traditionally used to quantify the spatial variability of ground properties, random field theory, and ground response analyses.
- In Chapter 3, the spatial variability of shear wave velocity is characterized at select sites in the province of British Columbia, Canada. The marginal distribution and correlation structure are established such that the spatial variability can be modelled using random field theory. The proposed approach is also compared against an existing method to randomize shear wave velocity, which consists of a first-order auto-regressive model.
- In Chapter 4, a series of equivalent-linear ground response analysis are completed in which the spatial variability of shear wave velocity is modeled using random fields. First, the properties of the random fields are based on the recommended values provided in Chapter 3. Results are presented in terms of the distribution of peak ground acceleration at the surface, the cyclic stress ratio, and spectral acceleration response spectra. Then, the sensitivity of the results to random field parameters is assessed. Comparisons between the deterministic and probabilistic approaches are presented, as well as between 1- and 2-dimensional models.
- Finally, Chapter 5 summarizes the findings of the study, outlines the limitations of the work, and proposes opportunities for future extension of this work.

Chapter 2

Literature Review

2.1 Quantifying Spatial Variability of Soil Properties

The spatial variability of ground properties is one of the largest sources of uncertainty in geotechnical analyses. Treating soil properties as random fields with an estimated mean, variance, and correlation structures can allow for the estimation of probabilities of failure or reliability indexes for geotechnical systems.

The following sections provide a brief overview of the statistical concepts that have been applied to complete risk or reliability based geotechnical engineering designs. Mainly, the focus is on characterizing how a soil property varies at a single point (marginal distribution), and how that variability changes when considering multiple points in space (correlation structure).

2.1.1 Marginal Distribution

Instead of considering a soil property deterministically through selection of a characteristic value, a probabilistic approach involves modelling the property in question as a random variable; that is, a variable that can take one of many possible outcomes. Because most soil properties are continuous, probability densities are used instead of probability values directly. The probability density function $f_X(x)$ describes the relative likelihood that the property X lies between x and $x + dx$. That is:

$$P[x < X \leq x + dx] = f_x(x)dx \tag{2.1}$$

Because it is impossible to know the *true* distribution of a soil property, measurements collected can be used to fit a theoretical distribution to the data. The selected distribution should match the histogram of the data reasonably well, and should be physically reasonable. For example, properties that are strictly non-negative are generally modelled using the lognormal distribution so that it is impossible to obtain

negative values.

The most important descriptors of a random variable are its *central tendency* and *variability*, which are most commonly quantified through the *mean* (μ) and *variance* (σ^2) , respectively. Unbiased estimates of the true mean and variance can be estimated from a set of observations as follows:

$$\hat{\mu}_X = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2.2}$$

$$\hat{\sigma}_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\mu}_X)^2 \tag{2.3}$$

where the hat ($\wedge$) indicates that the parameter is an estimate of the true value. These sample moments can then be equated to the moments of common distribution functions to fit a theoretical distribution such as the normal, lognormal, exponential distributions. This then is the marginal distribution of the random variable, representing the variability at a single point.

A popular measure of variability used in engineering applications is the coefficient of variation (ν), which provides a dimensionless measure of variability relative to the mean of the property, as follows:

$$\hat{\nu}_X = \frac{\hat{\sigma}_X}{\hat{\mu}_X} \tag{2.4}$$

where $\hat{\sigma}_X$ is the sample standard deviation, corresponding to the square root of the sample variance.

The above estimators assume that the random variable is stationary, that is, that the mean and variance of the process do not depend on position. However, mechanical properties of natural soils generally increase in magnitude with depth, as soils under increased confining stresses are typically stronger and stiffer. For this reason, it is common to represent soil properties through the sum of a deterministic trend component and a random residual, as follows:

$$X(d) = \hat{x}(d) + \varepsilon \tag{2.5}$$

where $\hat{x}(d)$ is the mean soil property at a depth d and ε is a normally distributed random variable with zero mean. The approach then consists of providing an estimate

of the trend, and of the statistical properties of the residuals.

2.1.2 Spatial Correlation

The probability distribution function of a random variable as described above (Section 2.1.1) allows for the description of variability at a point. However, in geotechnical applications it is typically necessary to estimate the property in 1D, 2D, or 3D space instead of a single point. For this reason, it is necessary to account for the spatial dependence of the soil property; that is, to estimate how the random process changes in space.

Consider, for example, the three random processes shown in Figure 2.1. Even though the three fields have the same marginal distribution (same point mean and point variance), the processes look quite different. This is because each field has vastly different correlation structures, with the field to the left being the least correlated and the field on the right being the most strongly correlated. It is thus necessary to establish the mathematical models required to account for the spatial correlation structure in order to fully model soil properties probabilistically.

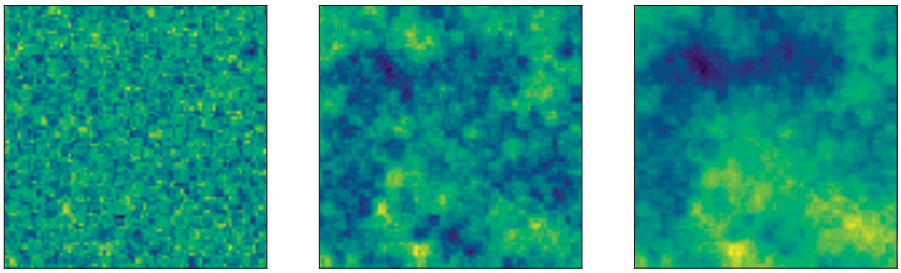

Figure 2.1: Effect of correlation structure on random processes.

The correlation structure of a random variable describes how strongly correlated points in the process are as a function of separation distance. It is logical that two measurements closely together will be more correlated to each other than measurements that are separated by large distances. There are a variety of tools available to quantify the correlation structure of a random process, which are commonly referred to as second-order structural analysis. These tools include, but are not limited to (Fenton, 1999a):

1. The *sample covariance function*, $\hat{C}(\tau)$, which describes the covariance between

two points in the random process separated by a distance τ. For 1-dimensional processes, it can be estimated by the following expression:

$$\hat{C}(\tau = j\Delta) = \frac{1}{n-j-1}\sum_{i=1}^{n-j}(x_i - \hat{\mu}_X)(x_{i+j} - \hat{\mu}_X) \tag{2.6}$$

where $(j\Delta)$ is the separation distance, for $j = 0, 1, ...n-2$. The sample covariance function is commonly transformed to the *sample correlation function* as follows:

$$\hat{\rho}(\tau) = \frac{\hat{C}(\tau)}{\hat{\sigma}_X^2} \tag{2.7}$$

where $\hat{\sigma}_X^2 = \hat{C}(0)$ is the estimate of the point variance of the random process. The sample correlation function is arguably the most commonly used tool to describe the spatial correlation of geotechnical properties. However, one of the major difficulties is that it is heavily dependent on the estimate of the mean $(\hat{\mu}_X)$. In the presence of strong correlation, the estimate of the mean is generally a poor estimate of the true mean. This fact then leads to a biased estimate of the covariance (and correlation) function whenever there is significant correlation throughout the entire sampling domain.

2. The *sample semi-variogram*, $\hat{V}(\tau)$, which is more commonly used in mining geostatistics and hydrological applications. The sample semi-variogram is closely related to the sample correlation function according to:

$$\begin{aligned}\hat{V}(\tau_j) &= \hat{\sigma}_X^2 - \hat{C}(\tau_j) \\ &= \hat{\sigma}_X^2\left(1 - \hat{\rho}(\tau_j)\right)\end{aligned} \tag{2.8}$$

The semi-variogram is advantageous in that it does not depend on the estimated mean of the process; however, there is a high variability in its estimate and therefore requires more data to accurately estimate.

3. The *sample variance function*, $\hat{\gamma}(D)$, measures the reduction in variance that results from averaging an increasing number of sequential random variables. In the one-dimensional case, the variance reduction function is related to the correlation function as follows:

$$\hat{\gamma}(D) = \frac{2}{D^2}\int_0^D (D-\tau)\hat{\rho}(\tau)d\tau \tag{2.9}$$

The sample variance is also biased in the presence of correlation between observations. For the sample variance function to become unbiased, the correlation function must decrease sufficiently rapidly within the averaging region.

4. The *sample spectral density function*, $\hat{G}(\omega)$, which contains the equivalent information to the covariance function but expressed in a different manner. $\hat{G}(w_j)$ can be obtained by first computing the Fourier transform of the data for each Fourier frequency (w_j) and then squaring the magnitude of the complex Fourier coefficients. However, it can also be obtained by taking the inverse Fourier transformation of the covariance function, as follows:

$$\hat{G}(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} \hat{C}(\tau)cos(\omega\tau)d\tau \tag{2.10}$$

Fenton (1999a) explores the usefulness of the above tools to distinguish whether a soil property is best modelled as a *finite-scale* or as a *fractal* process. A finite-scale process is one in which the area under the correlation function is finite. Vanmarcke (1998) defined the correlation length (θ) as the area under the correlation function from $-\infty$ to $+\infty$, as follows:

$$\theta = \int_{-\infty}^{\infty} \rho(\tau)d\tau = 2\int_0^{\infty} \rho(\tau)d\tau \tag{2.11}$$

Finite-scale models are simply those with a finite correlation length. In contrast, fractal models (also referred to as long-memory processes), have an infinite correlation length. Fenton (1999a) shows that a significant advantage of the sample spectral density function is that it allows to detect whether a process exhibits a fractal nature, in which case the spectral density function is linear with a negative slope in log-log space. However, a large data set is required to distinguish whether finite-scale or fractal models are more applicable.

Despite the large variety of estimation tools, the vast majority of geotechnical research has focused on using the sample correlation function to describe the correlation structure of soil parameters, with particular emphasis on the estimation of

correlation length. For this reason, the focus now turns to describing the available methods to estimate the correlation length.

Estimating Correlation Length

Correlation length (θ) can be roughly described as the distance beyond which soil properties are no longer significantly correlated. Arguably the most common approach to estimate correlation length is to fit a theoretical correlation function to a sample correlation function estimated from measurements of the soil property (using Equations 2.6 and 2.7). There exists many different theoretical functions that can be fitted to the sample function, a few of which are summarized in Table 2.1.

Model	**Correlation Function**
Markov	$\rho(\tau) = \exp\left[\frac{-2\lvert\tau\rvert}{\theta}\right]$
Gaussian	$\rho(\tau) = \exp\left[-\pi\left(\frac{\tau}{\theta}\right)^2\right]$
Binary Noise	$\rho(\tau) = \begin{cases} 1 - \lvert\tau\rvert/\theta \text{ if } \tau \ \leq \theta \\ 0 \text{ otherwise} \end{cases}$
Whittle-Matern	$\frac{2}{\Gamma(\nu)}\left(\frac{\sqrt{\pi}\Gamma(\nu+0.5)\lvert\tau\rvert}{\Gamma(\nu)\theta}\right)^{\nu} K_\nu\left(\frac{\sqrt{\pi}\Gamma(\nu+0.5)\lvert\tau\rvert}{\Gamma(\nu)\theta}\right)$

Table 2.1: Common theoretical correlation functions.

The Markov correlation function is the most commonly used to estimate correlation length in geotechnical engineering applications; however, the Gaussian model has the advantage of being mean square differentiable. For this reason, the Gaussian model is commonly used in analytical methods that do not involve some degree of averaging, such as level-crossing statistics, since any amount of averaging makes the process mean square differentiable (gradients having finite variance). The Whittle-Matern model requires an additional parameter besides correlation length, which is called the smoothness parameter (ν), and therefore allows for more flexibility when fitting the sample autocorrelation function (Cami et al., 2020). The model uses the gamma function ($\Gamma(\nu)$) as defined in Abramowitz and Stegun (1970). Note that the Whittle-Matern model corresponds to the Markov model when $\nu = 0.5$ and the Gaussian model when $\nu = \infty$.

Although a variety of theoretical correlation functions are available, much of the focus has been in estimating the correlation length, with less emphasis on which type of assumed correlation function is used. Phoon and Kulhawy (1999) and more recently Cami et al. (2020) present a literature review of common correlation lengths reported in geotechnical studies for a variety of soil types and soil properties. Particular focus has been on estimating correlation lengths from cone penetration test (CPT) data. This is likely due to the large number of data points that are measured during a CPT sounding, as well as the relatively small sampling length when compared to other types of tests.

The correlation length may be found through the *Method of Moments*, in which the error between sample moments and theoretical moments is minimized usually through least squares regression. Another approach is *Maximum-Likelihood Estimation*, in which model parameters are found by maximizing the likelihood of observing the known data under the assumed distribution. More recently, *Bayesian analysis* has also been used and may provide useful insights when limited data is available at a given site (Ching et al., 2015).

2.2 Random Field Theory

Random field theory has been applied to geotechnical engineering problems with the goal of formally accounting for the spatial variability in soil properties throughout the design process. The usefulness of using random fields arises because it is impossible to know the value of a given property at every single point at a site; therefore, the use of probabilistic models allows the uncertain nature of the ground to be properly modeled (Vanmarcke, 1998).

A continuous-state random processes $X(d)$ is one in which the random variable X can take on an infinite number of possible values at each location d. Since values of X at different locations (x_1, x_2 ...) are generally correlated, the complete probabilistic description of $X(d)$ requires the joint probability density function:

$$f_{X_1,X_2...}(x_1, x_2...) \tag{2.12}$$

Dealing with such a multivariate probability density function is impractical, therefore the following assumptions are generally adopted in geotechnical engineering applications (Fenton and Griffiths, 2008):

1. The random field is a *Gaussian process*; that is, the joint probability density function is normally distributed.
2. The process is *stationary* (or *statistically homogeneous*), such that the mean is independent of position and the covariance between two points only depends on the separation distance, or lag, τ.
3. The process is *isotropic*, such that the covariance between two points only depends on the absolute distance between two points, and not on their orientation.

Under these assumptions, all that is needed to fully characterize the random field are the first two moments of the random process, i.e., its mean and covariance structure, which can be estimated through the tools described in Section 2.1. Once the properties of the field are established, the task then becomes generating random realizations of said process that can then be used in geotechnical analyses.

There exists various algorithms to generate random fields, such as moving average methods, discrete fourier transform, turning band methods, among others (Fenton and Griffiths, 2007). However, the two methods that have been more commonly used in geotechnical engineering applications are the Covariance Matrix Decomposition method and the Local Average Subdivision Method (LAS), which are briefly described in the following sections.

2.2.1 Covariance Matrix Decomposition

Covariance matrix decomposition is a direct method of producing a homogeneous random field with a prescribed covariance structure. A discrete process X with zero mean can be produced using the expression:

$$X = LG \tag{2.13}$$

where G is a vector of independent, zero mean, unit variance, Gaussian random variables and L is a lower triangular matrix that is obtained through Cholesky decomposition, so that it satisfies the expression:

$$LL^T = C \tag{2.14}$$

where C is the positive-definite covariance matrix with elements $C_{ij} = C(\tau_{ij})$. This method is advantageous in that it is exact, conceptually easy to understand, and straightforward to code. However, the method becomes time-consuming and prone to round-off error for large fields (roughly larger than 128 random variables in size).

2.2.2 Local Average Subdivision (LAS)

Local average subdivision, LAS, (Fenton, 1990) is an approximate method to produce realizations of Gaussian random fields. The development of LAS was motivated by the fact that engineering properties are generally measured and represented as *local averages* over a finite domain, and by the need to generate values which are naturally used in averaging continuum models such as the finite element method.

The simulation of a random field using LAS is completed through a top-down approach. A global average is generated, which is then subdivided into two regions whose *local* averages must average to the global value. In subsequent stages, whenever each *parent* cell is subdivided, two normally distributed values are selected so that their mean and variance are consistent with local averaging theory, are properly correlated to one another, and average to the parent value. The method is *approximate* for the following reasons, which are discussed in detail in Fenton and Vanmarcke (1990):

1. The correlation between adjacent cells across parent boundaries is accomplished through the parent values and not directly.
2. The range of parent cells on which to condition the distributions is limited to some neighborhood (usually 3 or 5).

A particular advantage of LAS is that it is well suited to be combined with the finite element method (FEM), as each value in the random field realization can be mapped to an element in the finite element mesh and represent the average property within that element. The preservation of local averages allows for different resolutions of the finite element mesh while still preserving the proper statistics of the

random field. This fact led to the development of the Random Finite Element Method (RFEM), which combines finite element method with random field models to study the reliability of geotechnical systems (Fenton and Griffiths, 2008).

RFEM has been applied to a wide variety of problems to assess the reliability of geotechnical systems, such as slope stability (Fenton et al., 2003), bearing capacity (Fenton and Griffiths, 2003), settlement problems (Griffiths and Fenton, 2009), lateral earth pressures (Griffiths et al., 2008), deep foundations (Christodoulou et al., 2020), among others. More recently, RFEM as well as the random limit equilibrium method (RLEM) were incorporated into the popular commercial software Slide2D by RocScience. The adoption of probabilistic methods in commercial software allows practicing engineers to more feasibly carry out full reliability-based designs for ongoing projects.

2.3 Ground Response Analysis

The main objective of ground response analysis (GRA) is to evaluate the effects of local ground conditions on the response of a soil mass subject to earthquake ground motions directly beneath it. GRA can be applied to a variety of problems, including the development of design response spectra for infrastructure, evaluation of liquefaction hazards, and stability of slopes and earth-retaining structures during seismic events (Kramer, 1996).

Detailed guidelines on carrying out ground response analysis are presented in NCHRP (2012), which includes identifying when GRA are necessary, an overview of the available methods, and how to develop input parameters. GRA can be classified in a variety of ways, including but not limited to: the domain in which the calculations are performed (frequency or time domain), the complexity of the models that describe soil behavior (linear, equivalent-linear, or nonlinear), the consideration of pore water pressure generation (total-stress versus effective-stress analysis), and the dimensionality of the analysis (1-, 2-, or 3-dimensional models).

The dimensionality selected for the GRA depends on the nature of the system that is being modeled. One-dimensional GRA assume that all boundaries are horizontal, and that the response is governed by shear waves propagating vertically through the

soil mass. Since 1D models imply that the soil mass extends infinitely in the horizontal direction, it may not be well-suited for sites with sloping ground, stratigraphy that changes horizontally, or sites with embedded structures. In these cases, two- or three-dimensional analysis may be required. When one dimension is considerably larger than the others, a two-dimensional plain strain formulation may be sufficient; otherwise, a three-dimensional model would be required.

Regardless of the type of GRA selected, the development site-specific acceleration time histories that match the seismic hazard at the site is required. This process involves first making a selection of records from historical databases that match, to some degree, the tectonic environment, magnitude and distance, and response spectra characteristics. Then, modification of the time histories is necessary to represent the specific hazard at the site. This can be accomplished either through a *simple scaling* approach, in which the entire time history is linearly scaled so that its spectrum matches a given target, or through *spectrum matching*, an iterative approach where wavelets are added until a satisfactory response spectra is achieved.

Then, a definition of subsurface stratigraphy and corresponding soil properties is required to characterize the behavior of the soil mass. However, the level of information required depends on the complexity of the soil model employed. Generally, frequency-domain, equivalent-linear analyses are the least complex and require the least amount of information to complete.

In an equivalent-linear approach, the true nonlinear properties of soil are approximated through the secant shear modulus G and an equivalent damping ratio ζ. The values of G and ζ must be consistent with the level of shear strain at each layer; however, the computed strain levels depend on this properties. The interdependency of strain levels and soil properties results in the need of an iterative approach which converges towards strain-compatible properties. The general procedure is as follows (Kramer, 1996):

1. Select the initial estimates of shear modulus and damping ($G^{(i=0)}$ and $\zeta^{(i=0)}$), which are generally taken as the low-strain values.
2. Use the values of $G^{(i)}$ and $\zeta^{(i)}$ to compute the ground response for iteration i, and obtain the shear strain time history for each layer.
3. The effective shear strain (γ_{eff}) is determined as a function of the maximum

shear strain in the time history (γ_{max}) for each layer j and the earthquake magnitude M, as follows:

$$\gamma_{\text{eff},\,j}^{(i)} = \frac{M-1}{10}\,\gamma_{\text{max},\,j}^{(i)} \tag{2.15}$$

4. Determine the stiffness and damping values for the next iteration ($G^{(i+1)}$ and $\zeta^{(i+1)}$) using the new effective shear strain.
5. Repeat steps 2 to 4 until the differences in successive iterations fall below a predetermined value in all layers.

Because the equivalent-linear method is an approximation of nonlinear behavior, there are limitations to using this approach to model ground response. Challenges associated with the equivalent-linear method include (Kramer (1996), NCHRP (2012), and Stewart et al. (2014)):

- Large strain response is not well-captured by equivalent-linear methods.
- Inability to capture the effects of excess pore water pressures (soil liquefaction), since the method is a total-stress approach, as well as the effects of cyclic degradation in clay sites exposed to long-duration motions.
- High levels of amplification an when a strong component of the input motion matches the natural frequency of the soil mass that is obtained from the equivalent-linear properties. Since the actual soil properties change throughout seismic shaking, such high resonances should not develop.
- The use of the effective shear strain (γ_{eff}) may result in an over-softened and over-damped system (when the maximum shear strain is much larger than the overall time history), or to an under-softened and under-damped system (when the shear strain time history is somewhat uniform).

Some of the limitation of equivalent-linear approach can be overcome through the use of more complex nonlinear methods. Effective stress nonlinear methods can also model the generation, redistribution, and dissipation of excess pore pressure during and after shaking. However, calibrating the constitutive models that describe nonlinear behavior is significantly more complex and may require substantial field and laboratory testing programs. Both equivalent and nonlinear techniques can be used

to successfully carry out GRA, but their use and interpretation requires a thorough knowledge of the underlying assumptions and inherent limitations.

Chapter 3

Spatial Variability of Shear Wave Velocity

3.1 Motivation

As the Canadian and world-wide geotechnical community increasingly transition towards reliability-based and risk-based design methods, the need for models that can account for the spatial variability of ground properties is becoming increasingly important.

In recent years, the use of random fields to describe the spatial variability of soils has gained popularity in reliability-based designs. Random fields can be readily combined with finite elements or limit-equilibrium analyses to assess the reliability of geotechnical systems. These types of analyses are becoming increasingly feasible in engineering practice, as popular commercial software are starting to offer probabilistic features like incorporating the generation of random fields into the finite element and limit equilibrium formulations.

While several studies have characterized the randomness of soil properties, such as cone penetration resistance, undrained shear strength, and friction angle, there are only limited studies on the statistical properties of shear wave velocity in soils. Shear wave velocity is a parameter that is widely used in site response analyses, as it is directly related to the small shear-strain stiffness modulus of soils.

A challenge in estimating the random properties of soils is that a large number of samples is required to obtain reasonably accurate estimates, particularly for the covariance structure. The task then becomes unfeasible in many geotechnical projects where limited budgets constrain the amount of testing at any given site.

The purpose of this chapter is to use a database of seismic cone penetration test (SCPT) completed in the province of British Columbia, Canada, to estimate the statistical parameters of shear wave velocity. The results of this process can then be used to inform the selection of random properties in reliability-based analyses for similar sites where less geotechnical data is available.

3.2 SCPT Database

A collection of 206 seismic cone penetration tests (SCPT) were compiled from projects in the province of British Columbia (BC). The majority of the tests were collected as part of four large site investigations shown in plan view in Figure 3.1 and labelled A through D. An additional 11 tests were completed in smaller projects located in the Lower Mainland of BC, for which detailed coordinate information is not available.

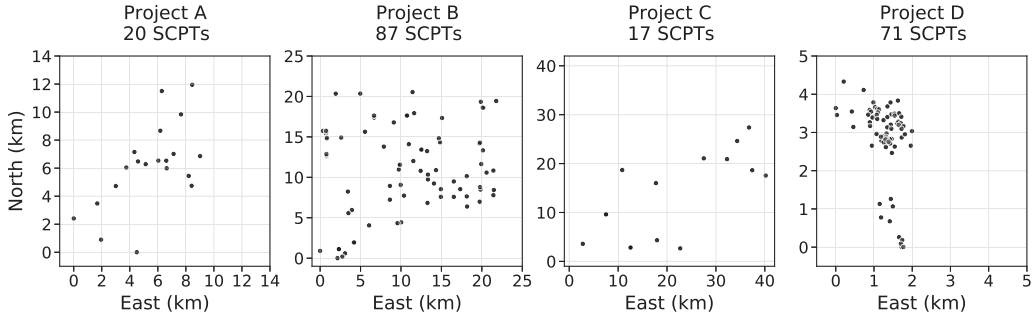

Figure 3.1: Plan view of SCPT collected in site investigations.

Projects A and C consist of 20 and 17 SCPTs, respectively, and correspond to investigations in south-western BC for sites underlain by Fraser River sediments. Project B includes 87 SCPTs that are part of a compilation of shear wave velocities freely available in the Geological Survey of Canada Open File 3622 (Hunter et al. 1998). The SCPTs in site B correspond to unconsolidated sediments of the Fraser River Delta. Although this Open File includes measurements from a variety of testing methods such as downhole boreholes and surface shear wave refraction, only the SCPT results are considered here for consistency with the other data collected. Finally, project D consists of 71 SCPT and is located on the northern coast of British Columbia in an area predominantly underlain by sand, gravel, and minor silt.

The average shear wave velocity in the top 30 meters of soil (v_{s30}) is an important parameter that has been traditionally used in earthquake geotechnical engineering applications to separate sites into different classes that define a different seismic response. The value of v_{s30} is calculated as follows:

$$v_{s30} = \frac{\sum_{i=1}^{n} t_i}{\sum_{i=1}^{n} (t_i / v_i)} \tag{3.1}$$

where t_i and v_i are the thickness and shear wave velocity of the i^{th} layer, respectively,

and only layers in the top 30 meters of the profile are considered.

Table 3.1 and Figure 3.2 summarize the ranges in v_{s30} estimated for each SCPT across all data sources. Of the profiles considered, 99 SCPTs (48%) are classified as Site Class D (stiff soil, between 180m/s and 360m/s) and 107 SCPTs (52%) as Site Class E (soft soil, less than 180 m/s), according to the 2015 National Building Code of Canada.

Project	**Number of SCPTs**	**Range** v_{s30} (m/s)	**Average** v_{s30} (m/s)
A	20	132 - 185	155
B	87	97 - 239	172
C	17	151 - 204	173
D	71	134 - 283	203
Others	11	100 - 286	183

Table 3.1: Summary of v_{s30} for all data sources.

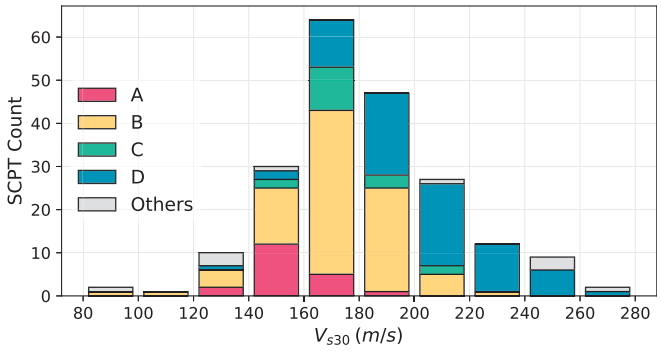

Figure 3.2: Histogram of v_{s30} across all data sources.

3.3 Marginal Distribution

3.3.1 De-trending Approach

It is common for geotechnical properties to show a dependency with depth, as soils under increased effective stresses are generally stronger and stiffer. However, most random field models assume stationary; that is, that the mean and covariance of the random process are not dependent on position. In an attempt to satisfy the stationarity requirement, it is common practice in geotechnical analyses to represent spatial variability as follows:

$$V(d) = \hat{v}(d) + \varepsilon \tag{3.2}$$

where the depth-dependent randomized shear wave velocity ($V(d)$) is composed of two components: (1) a trend with respect to depth ($\hat{v}(d)$) which is assumed to be deterministic, and (2) a random residual component (ε) with zero mean that represents the variability about the trend (Baecher and Christian, 2003). Then, the analysis consists of providing an estimate of the trend and the statistical properties of the residuals.

Since shear wave velocity is a strictly non-negative parameter, it is desirable to avoid the possibility of negative values. For this reason, Equation 3.2 is replaced by a multiplicative form:

$$V(d) = Y\hat{v}(d) \tag{3.3}$$

in which Y is a lognormally distributed random variable with unit mean. Taking the logarithm of Equation 3.3 gives:

$$ln\left(V(d)\right) = ln(Y) + ln(\hat{v}(d)) \tag{3.4}$$

However, it is important to note that separating the trend, as in Equations 3.2 or 3.3, can be problematic in inferential statistic analysis and should be approached with caution. As recommended by Fenton (1999a), de-trending should only be completed when there is: 1) a physical basis for the trend, and 2) when similar trends can be expected to occur at other target sites. Otherwise, using residual statistics at different target sites may lead to estimates that are grossly in error, because residual statistics usually have a reduced variance and reduced spatial dependency when compared to the original parameter.

There are two alternatives to avoid prescribing a deterministic trend. One option is to randomize the trend component in Equation 3.2, so that the randomness in the trend is quantified through a normal or lognormal random variable. Jiang and Huang (2018) summarized different expressions that have been used in literature to randomize the trend, with a focus on slope stability applications. Another option is to estimate spatial variability through a Bayesian instead of frequentist approach, as proposed in Ching et al. (2015). In this case, the functional form of the trend,

its parameters, and variability around the trend are randomized within a Bayesian framework (Ching and Phoon, 2017). The scope of this study is limited to using a deterministic trend; however, either of these alternatives is worth exploring in subsequent studies.

3.3.2 Trend Estimation

Baecher and Christian (2003) suggest that the functional form of the trend should be selected to be "as simple as possible without doing injustice to the set of data or ignoring the geological setting". Generally, statistical studies of CPT data opt for a linear trend estimation between depth and cone penetration resistance (see for example Fenton (1999b), Lloret-Cabot et al. (2014)).

However, shear wave velocity is generally seen to have a power relationship with depth. A mapping of shear wave velocity beneath the Fraser River Delta completed by Hunter et al. (1999) reported a relationship between depth and shear wave velocity as follows:

$$\hat{v}(d) = 35.26d^{0.4362} + 71.22 \tag{3.5}$$

For consistency with the Hunter et al. (1999) study completed for similar soils, the relationship between depth (d) and estimate of the shear wave velocity mean ($\hat{v}(d)$) was determined using the same functional form, that is:

$$\hat{v}(d) = c_1 d^{c_2} + c_3 \tag{3.6}$$

where c_1, c_2 and c_3 are fitting constants found through least-squares regression.

Resonant column tests and theoretical analyses completed by Cascante and Santamarina (1996) provide some physical justification for the above functional form. Cascante and Santamarina (1996) found that the small-strain shear modulus is strongly controlled by the behavior of particle contacts, and that a power relationship between small-strain shear modulus and confining stress can be expected. Since depth and confining stress are proportional, a similar can be expected with respect to depth. Since the functional form shown in Equation 3.6 can be expected to occur at similar sites and has a physical basis, it is considered reasonable to detrend the available shear wave velocity measurements.

The regression can be performed individually for each SCPT or globally over an ensemble of tests (Fenton, 1999b). Here, it is considered that a global regression at each project is preferable, so that a single estimate of the trend at each site can be produced. An additional regression is completed where all the data from different projects is pooled together, resulting in a total of 5 different regressions.

Figure 3.3 shows the estimated trends at each site separately, as well as an additional case where all the data is pooled together. The results for the constants c_1, c_2, and c_3 are also summarized in Table 3.2.

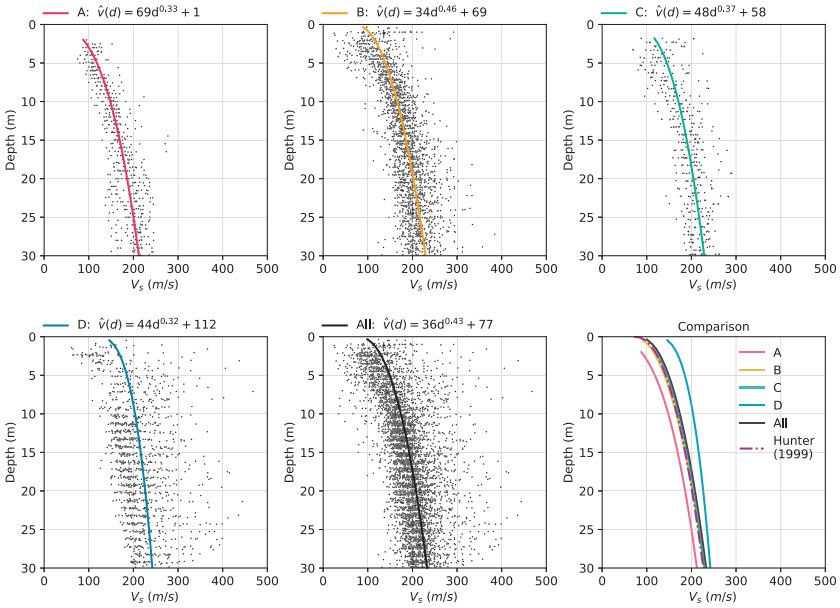

Figure 3.3: Estimation of trend for each site considered.

Site	Num. SCPT	c_1	c_2	c_2
A	20	68.72	0.3298	1.41
B	87	33.56	0.4601	6.91
C	17	47.81	0.3746	5.77
D	71	43.65	0.3217	1.12
All	206	35.85	0.4319	7.73
Hunter (1995)	-	35.26	0.4362	71.22

Table 3.2: Results from trend analysis.

Comparing the trends at each site shows a very consistent relationship with respect to depth, even for sites that are in entirely different regions. Therefore, it can be reasonably expected that similar trends could be expected at other target sites.

3.3.3 Mean and Variance of Y

Now, focus turns to characterizing the statistical properties of the random variable Y. The purpose of de-trending in the preceding section was primarily to achieve a stationary process: one that has a mean and variance independent of depth. The stationarity of Y can be checked by estimating the mean and variance in binned intervals with respect to depth, and confirming that these parameters remain reasonably constant regardless of position (Fenton and Griffiths, 2008). Figure 3.4 shows the results of this procedure for all sites using 2 m depth intervals.

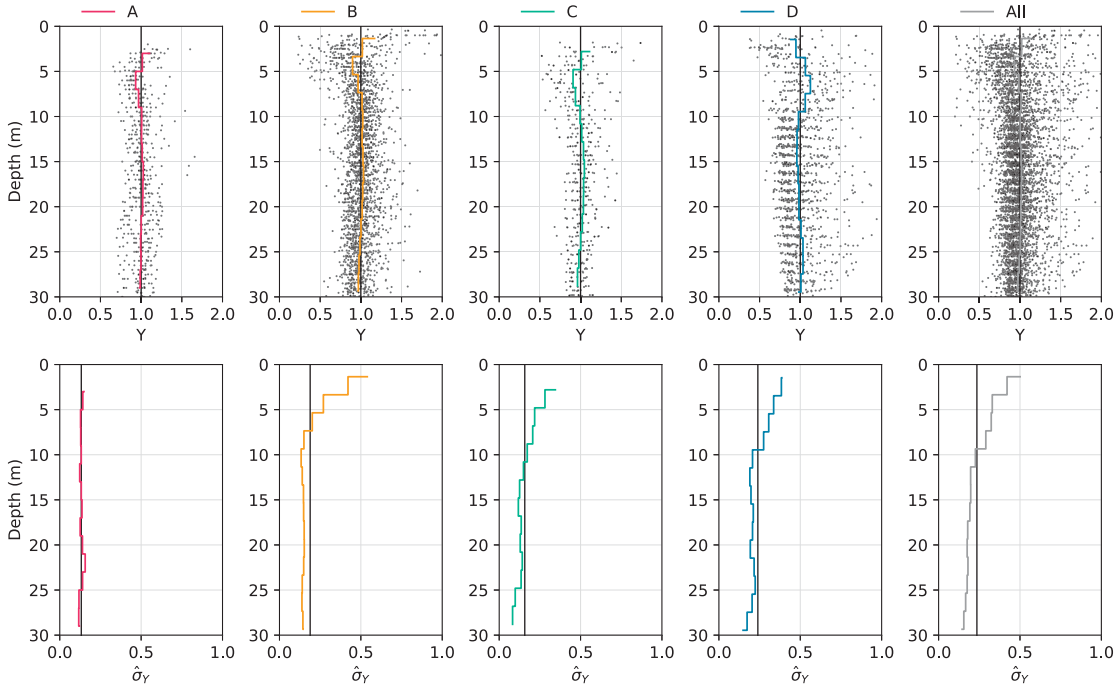

Figure 3.4: Profiles of binned mean and standard deviation of Y.

While the binned estimate of the mean of Y remains relatively constant with depth (at a value of $\hat{\mu}_Y = 1$), the estimate of the standard deviation of Y is largest at the surface but decreases in the top 5 to 10 meters at sites B, C, and D. It is difficult to judge whether this characteristic is likely to hold at other sites, particularly when considering that site A did not exhibit such a trend in the variance.

It is possible to account for a non-stationary variance when randomizing soil parameters (Fenton and Griffiths, 2008), so that one could feasibly prescribe a decreasing variance in the top 10 meters of the soil profile. If the purpose here was to solely

describe the soils at these sites, then this might be a preferable option. However, doing so would yield results that are likely only applicable when a similar trend in variance is observed at other target sites. Since it is yet to be determined whether the decrease in variance observed at sites B, C, and D is widely applicable, it is considered best not to prescribe this in the statistical analyses. Instead, it is assumed that the mean and variance of Y can be represented reasonably well when assumed to be independent of depth.

A lognormal distribution can then be easily fit to the data through the method of moments. Table 3.3 summarizes the estimated mean and standard deviation of Y for all data sets, as well as the lognormal transformation of the parameters. The resulting fitted lognormal distributions are compared to histograms of the measured data in Figure 3.5, showing a reasonably good agreement between the theoretical and measured distributions.

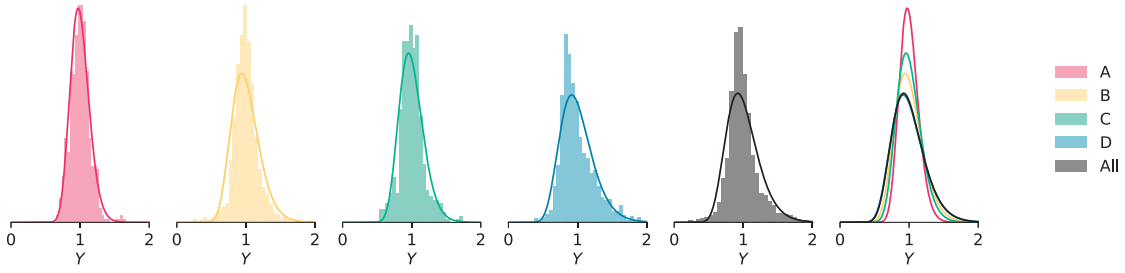

Figure 3.5: Fitted lognormal distributions of Y for each site.

The mean and standard deviation of $V(d)$ can then be expressed as:

$$\mu_V = \mu_Y \hat{v}(d) \tag{3.7}$$

$$\sigma_V = \sigma_Y \hat{v}(d) \tag{3.8}$$

Therefore, the coefficient of variation of shear wave velocity, ν_V, is independent of depth and equal to the coefficient of variation of Y, (ν_Y) since the trend component cancels out when dividing σ_V over μ_V. Since μ_Y is equal to 1, it follows that σ_Y is equal to ν_V, which is a more intuitive parameter to compare the variability at different sites.

Site	$\hat{\mu}_Y$	$\hat{\sigma}_Y$ or $\hat{\nu}_V$	$\hat{\mu}_{lnY}$	$\hat{\sigma}_{lnY}$
A	1.00	0.138	-0.0093	0.1376
B	1.00	0.204	-0.0190	0.2018
C	1.00	0.178	-0.0152	0.1761
D	1.00	0.244	-0.0299	0.2408
All	1.00	0.240	-0.0277	0.2368

Table 3.3: Estimated statistical parameters of Y.

3.4 Correlation Structure

3.4.1 Assumed Correlation Structure

Shear wave velocity data is collected in 3-dimensional space, since each data point has a corresponding northing, easting, and depth coordinate. To characterize the corresponding 2-dimensional random field, it is necessary to first establish the type of correlation structure that will be assumed in higher dimensions.

One of the simplest forms that can be assumed is a separable correlation structure (Fenton and Griffiths, 2008). The covariance function can then be expressed as the product of the correlation function in each dimension times the variance of the process, as follows:

$$C(\tau_{h_1}, \tau_{h_2}, \tau_v) = \sigma^2 \rho_{h_1}(\tau_{h_1}) \rho_{h_2}(\tau_{h_2}) \rho_v(\tau_v) \tag{3.9}$$

where h_1 and h_2 refer to two orthogonal horizontal directions and v refers to the vertical direction (depth). This form is advantageous in geotechnical applications, since depositional history generally results in soils that have much longer horizontal than vertical correlation lengths. Some depositional environments may also result in soils that have a stronger correlation along one horizontal correlation. For example, river deposits may display stronger correlation along the course of the river rather than perpendicularly. A compilation of correlation lengths by Phoon and Kulhawy (1999) shows that horizontal correlation lengths are generally at least an order of magnitude larger than the vertical ones.

An alternative to Equation 3.9 is to assume that the horizontal directions (h_1 and h_2) are similar enough to prescribe an equivalent correlation structure; that is, the correlation structure can be assumed to be isotropic in the horizontal plane. Under

this assumption, the correlation structure can be determined as follows:

$$C(\tau_{h_1}, \tau_{h_2}, \tau_v) = \sigma^2 \rho_h \left(\sqrt{\tau_{h_1}^2 + \tau_{h_2}^2} \right) \rho_v (\tau_v) \tag{3.10}$$

The advantage of using Equation 3.10 is that SCPT pairs used to estimate the horizontal correlation function do not need to be perfectly aligned in the h_1 or h_2 directions. Instead, any SCPT pair can be used to estimate the horizontal correlation structure regardless of the test orientation. This is particularly important since a much larger number of tests is needed to estimate horizontal rather than vertical correlation structure. Unfortunately, the available SCPT data in this study is spaced too far apart to reliably estimate the correlation structure in the horizontal direction. For this reason, only the vertical correlation structure is discussed further.

3.4.2 Sample Correlation Function

The sample correlation function for each SCPT can be obtained by first estimating the covariance between points separated by a distance $(j\Delta v)$, for $j = 0, 1, ..., n-2$ using the following estimator:

$$\hat{C}(\tau = j\Delta v) = \frac{1}{n-j-1} \sum_{i=1}^{n-j} (\ln Y_i - \hat{\mu}_{\ln Y})(\ln Y_{i+j} - \hat{\mu}_{\ln Y}) \tag{3.11}$$

where n is the number of data points in the SCPT. The sample correlation function $(\rho_{\ln Y}(\tau))$ is then easily found by dividing $\hat{C}(\tau)$ by the estimate of the variance of the random process $(\hat{\sigma}^2_{\ln Y} = \hat{C}(0))$.

A disadvantage of using Equation 3.11 is that it requires equispaced measurements separated by a distance Δv. Although SCPTs typically measure shear wave velocity at 1 meter intervals, some of the collected profiles had missing data points (likely due to drill-out), and some had intervals with different spacing that ranged between 0.7 m to 1.4 m. For these reasons, the sample correlation functions were not calculated directly for each individual SCPT. Instead, the correlation coefficients between all possible pairs of data points in one SCPT are calculated and plotted against separation distance. Then, the average of these correlation coefficients across all SCPTs is taken in binned intervals of separation distance and is used as an estimate of the

sample correlation function. The estimation process is illustrated in Figure 3.6, with the sample correlation function being shown in blue.

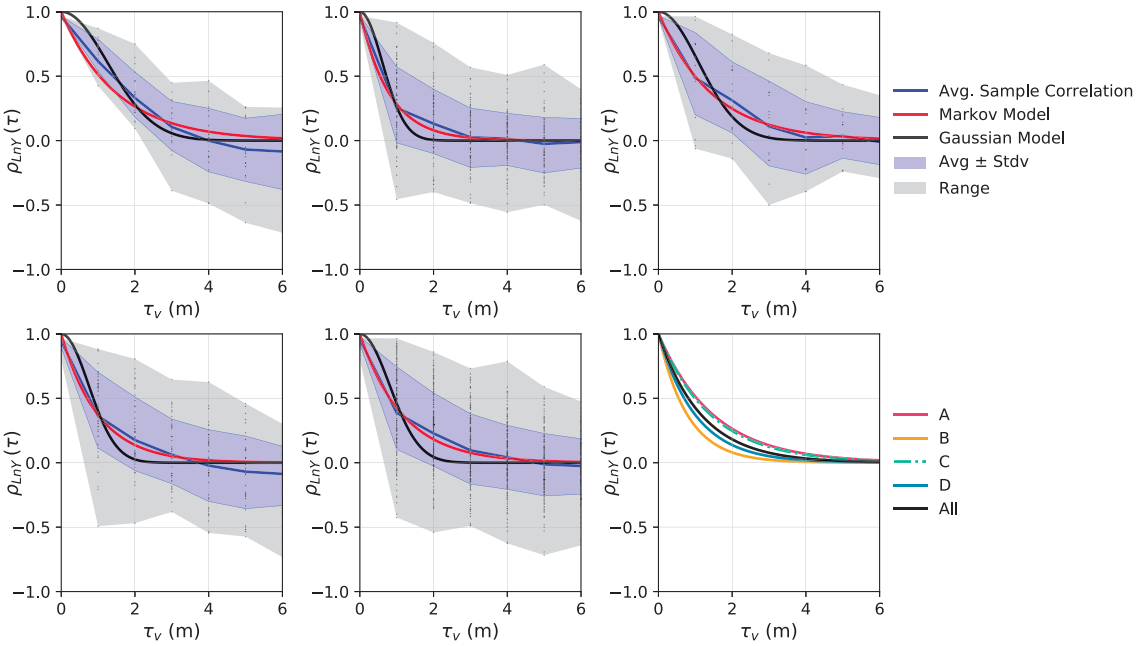

Figure 3.6: Estimation of vertical correlation function.

3.4.3 Estimation of Correlation Length by Direct Fitting

An estimate of correlation length can be produced by fitting a theoretical correlation model to the sample correlation functions through least-squares regression. One of the most common assumed models is the Markov correlation function (also referred to as the single exponential function), which is a function of the separation distance (τ) and the correlation length (θ), as follows:

$$\rho(\tau) = \exp\left[\frac{-2|\tau|}{\theta}\right] \tag{3.12}$$

A disadvantage of the Markov form is that it is not mean squared differentiable at $\tau = 0$, which leads to unstable calculations when using an analytical approach that does not involve some degree of averaging. For this reason, the Gaussian correlation function is also commonly used, which is expressed as:

$$\rho(\tau) = \exp\left[-\pi\left(\frac{|\tau|}{\theta}\right)^2\right] \tag{3.13}$$

Here, both the Markov and the Gaussian models are fitted to the estimated correlation functions, so that either model could be used in further studies. The resulting correlation lengths from the above procedure are summarized in Table 3.4 for both the Markov and the Gaussian models. While the correlation lengths obtained from the Markov and Gaussian models are fairly similar, Figure 3.6 shows that the Markov model more closely follows the sample correlation function, and so it might be a better choice to represent the correlation structure of $\ln(Y)$ at these sites.

Project	**Markov** $\hat{\theta}_{\ln Y}$	**Gaussian** $\hat{\theta}_{\ln Y}$	**Difference**
A	3.00	3.14	4.6 %
B	1.58	1.53	3.2 %
C	2.86	2.73	4.7 %
D	2.02	1.85	8.4 %
All	2.34	2.01	15.3 %

Table 3.4: Estimated vertical correlation length.

The estimation of correlation length through this approach can be highly biased. The first issue is that the large sampling intervals of SCPTs (1 meter) along with the relatively small sampling length (30 meters), means that a single SCPT provides very limited data to accurately estimate the sample correlation function. At a lag of zero, 30 data pairs are averaged to estimate $\rho(\tau = 0)$, but this number decreases by one for each meter of separation distance. At a lag of 6 meters, only 24 data points are averaged. This issue is somewhat alleviated by averaging over the ensemble of collected SCPTs instead of each SCPT individually.

Nonetheless, the limited data is an important disadvantage when compared against estimating correlation lengths for cone penetration resistance from CPT which are generally taken at intervals between 10 and 50 mm (Lunne et al., 2002). At the larger sampling length of 50 mm, a 30 m CPT results in 600 data points being averaged at a lag of 0 meters, and 480 at a lag of 6 meters. This is illustrated in Figure 3.7 by comparing the data points available to estimate $\rho(\tau)$ as a function of separation distance for 30 meter CPT and SCPT tests.

Another issue in the estimation process is that if the sampling depth of 30 meters is not significantly larger than the *true* correlation length, then the sample correlation function is highly biased and the estimated correlation lengths may be grossly

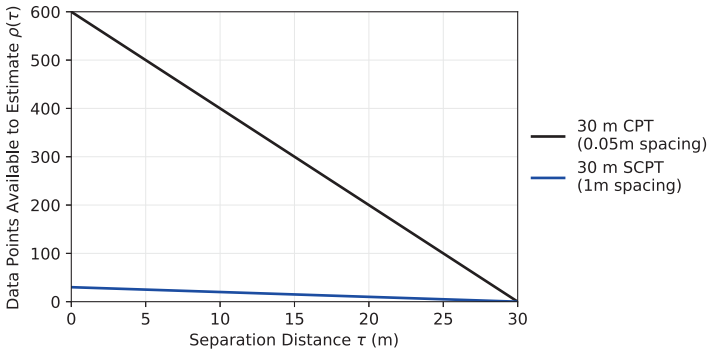

Figure 3.7: Simulated sample corelation functions.

in error. This problem is illustrated in Figure 3.8 through a simulation approach. It compares the *true* correlation length of a simulated random process against the estimated correlation length.

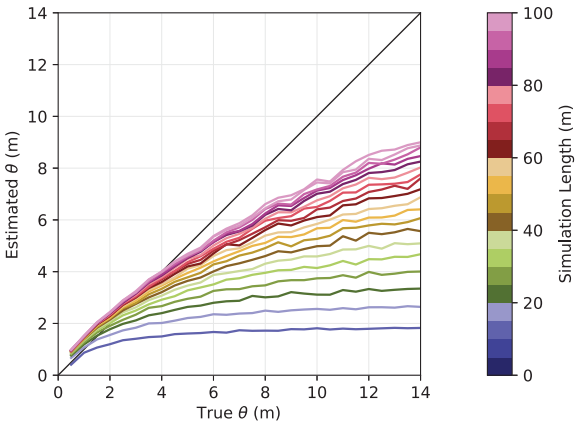

Figure 3.8: Simulated versus estimated correlation lengths.

The steps used to generate Figure 3.8 are as follows:

1. Generate a collection of 1-dimensional realizations of a random process using a sampling interval of 1 meter, a prescribed sampling depth, and a prescribed correlation length (θ_{true}) using Local Average Subdivision (LAS).
2. Determine the average *sample* correlation function using the generated realizations. Fit a Markov function through least-squares regression to determine the estimated correlation length (θ_{est}).
3. Repeat the above process for increasing values of true correlation length and sampling depth.

The results of this process, shown in Figure 3.8, illustrate how the *true* correlation length must be much smaller than the simulation length in order to obtain an accurate estimated correlation length. If the simulation length is short (lowermost blue curves), then the estimated correlation length becomes much smaller than the actual, *true* correlation length, with the error increasing for longer *true* correlation lengths.

3.4.4 Bias-Matched Estimation of Correlation Length

An alternative, bias-matched approach to the estimation of correlation length is presented here. The goal of this process is to find the correlation length that, upon generating an ensemble of random realizations, results in the closest match to the observed sample correlation functions from the collected SCPT data. Through this approach, the sample correlation functions taken from the simulations will provide the closest match to the bias of the sample correlation functions taken from the actual SCPT measurements. The steps of this approach are as follows:

1. Guess at a value of the correlation length (θ_{guess}).
2. Simulate $n = 1,000$ realizations of the random process through Local Average Subdivision. Use a Markov correlation function with correlation length θ_{guess}, and a sampling length and sampling interval equal to that of the SCPT measurements.
3. Estimate the *sample* correlation function from the ensemble of random realizations, in the same way that the sample correlation function is estimated from the SCPT data. This is referred to as $\rho_{sim}(\tau)$.
4. Repeat steps 1 to 3 for a variety of θ_{guess}, making sure to cover a range wide enough to cover what is expected to be the *true* correlation length.
5. For each simulated sample correlation function, determine the sum of squared errors between the sample correlation functions obtained from the simulation and that obtained from the original SCPT data. That is:

$$\varepsilon^2 = \sum_{all\ \tau} \left(\rho_{\ln Y}(\tau_i) - \rho_{sim}(\tau_i)\right)^2 \tag{3.14}$$

6. Determine the value of θ_{guess} that results in the lowest sum of squared errors calculated in Equation 3.14.

Although the above process can be improved to use an updating rule for θ_{guess} instead of the *brute-force* approach shown above, the computational requirements are small enough that using the above approach is not too time consuming.

Figure 3.9 displays the sample correlation functions that are obtained from the simulation with sampling depth 30 meters, a sampling interval of 1 meter, and various choices of correlation length (corresponding to step 4). The larger the correlation length, the more negative the sample correlation function becomes at larger separation distances due to the bias in the correlation function estimate.

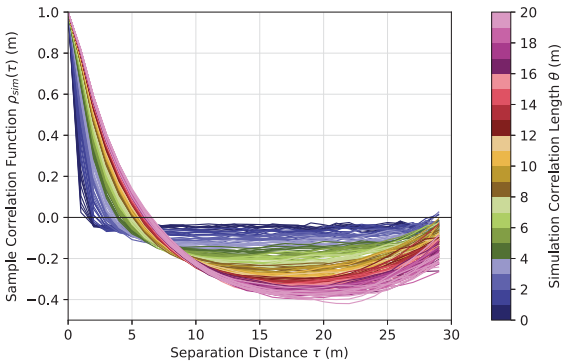

Figure 3.9: Simulated sample corelation functions.

Figure 3.10 shows the sum of squared errors between the simulated and estimated correlation functions (Equation 3.14) as a function of simulated correlation length. Because the calculated sum of squared errors display some noise, a 15-point rolling average was applied before minimizing the error, shown as a blue line in Figure 3.10. The simulation correlation length that minimizes the errors is shown by a red circle.

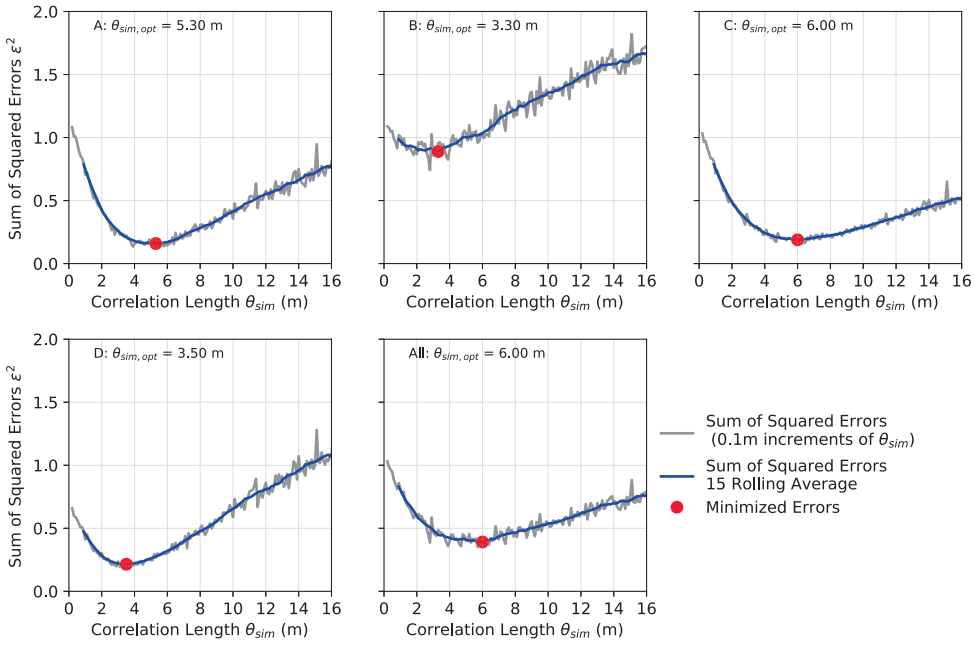

Figure 3.10: Selection of correlation length by minimizing squared errors.

Finally, a comparison of the sample correlation function of $\ln Y$ against the closest simulated correlation function is presented in Figure 3.11 for all sites. Table 3.5 summaries the estimated correlation lengths for each project site, and compares them against the direct fitting approach used previously. The estimated correlation lengths from this process are significantly larger than those obtained from fitting the theoretical correlation function directly, and can be expected to better represent the *true* correlation structure at the sites considered.

Project	**Direct Fitting** $\hat{\boldsymbol{\theta}}_{v,\,\ln Y}$ **(m)**	**Bias-Mached** $\hat{\boldsymbol{\theta}}_{v,\,\ln Y}$ **(m)**
A	3.00	5.30
B	1.58	3.30
C	2.86	6.00
D	2.02	3.50
All	2.34	6.00

Table 3.5: Comparison of estimated correlation lengths.

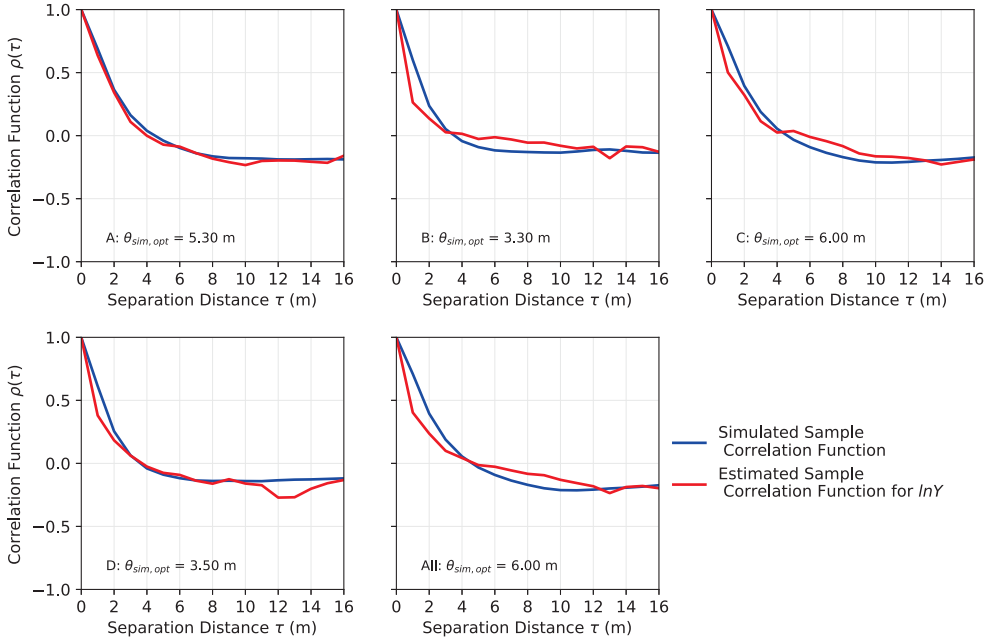

Figure 3.11: Simulation-based estimation of correlation length.

3.4.5 Averaging over SCPT Ensembles

A challenging aspect of estimating the correlation structure of natural soil is that abundant data is needed in order to make reliable estimates. Most geotechnical engineering projects do not carry out enough tests to estimate the correlation structure, due to time and/or budget constraints. Additionally, SCPT have the particular disadvantage of having a large sampling interval of 1 meter when compared against CPT, so that tests that are in the order of 30 meters provide limited data to estimate correlation lengths.

The above challenges can be overcome by averaging over an *ensemble* of SCPT available for similar soils, instead of estimating the correlation length directly for each test or single site. Two questions may arise from this process:

1. Is there a significant difference between fitting to the average sample correlation function for all SCPTs, versus fitting each individual sample correlation function and then averaging the resulting correlation lengths?
2. How many SCPT need to be averaged in order to reliably estimate the correlation length?

The first question can be explored by estimating the correlation length individually for each SCPT, and then comparing the results to the correlation length estimated for the ensemble of SCPT. This is completed here for all the SCPT data available, and compared against the estimate of $\hat{\theta}_{v,\ln Y} = 6$ m that was previously obtained for the full ensemble (Section 3.4.4).

First, SCPT that do not contain enough information to return an individual sample correlation function are discarded from the dataset. The minimum criteria was set as having enough data points to estimate the correlation function up to a lag of 5 meters, with a minimum of 10 data point pairs for each lag distance. Using this criteria, a total of 32 SCPT are discarded (mostly from project D). This is a clear disadvantage of trying to estimate the correlation function individually for each SCPT: the shorter SCPT must be discarded whereas by averaging over the ensemble, data from shorter SCPT tests can still contribute to the estimate at shorter lags. This becomes more important when the database includes relatively short SCPT, but might not be as relevant for databases with deeper SCPT.

Individual sample correlation functions were determined for the remaining 174 SCPT, and the correlation length of each was estimated through the bias-matched approach described in Section 3.4.4. Figure 3.12 shows a histogram of estimated correlation lengths for each individual SCPT. The average of the estimated correlation lengths results in a correlation length of $\hat{\theta}_{v,\,\ln Y} = 7.0$ m, which is reasonably close to the previous estimate of $\hat{\theta}_{v,\,\ln Y} = 6.0$ m. The wide range in estimates of correlation length obtained from single SCPTs reinforces the fact that a large number of tests may be needed to reach a reliable estimate of correlation structure.

The longer correlation lengths ($\hat{\theta}_{v,\,\ln Y} > 15$ m) observed in a few SCPT is a consequence of using a single, global function to de-trend the data instead of fitting a different trend for each SCPT. Any SCPT that does not follow the global trend closely will have larger sample correlation length, increasing as the agreement with the global trend decreases. Conversely, the better an SCPT matches the global trend, the smaller the correlation length of $\ln Y$ becomes. Using a single trend for the ensemble still seems preferable, as the random realizations of shear wave velocity assume a constant trend within the domain. In other words, if different trends at each location were allowed, then that would lead to the problem of which single trend to assume when generating realizations of the random parameter for use in design. A possible solution to this problem is to treat the trend component as a random variable, as discussed by Jiang and Huang (2018) for slope stability problems.

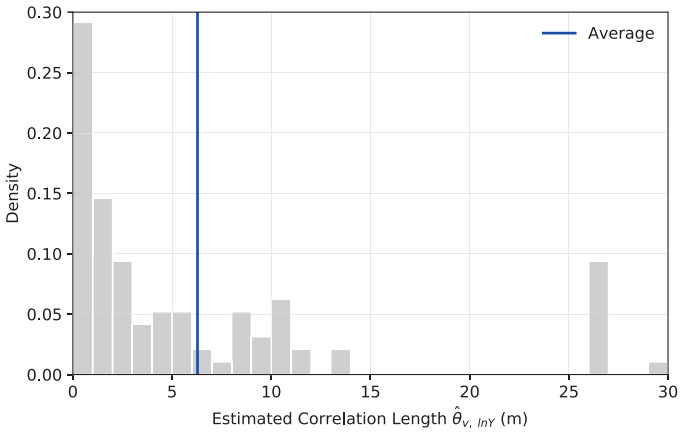

Figure 3.12: Estimation of correlation length for individual SCPT.

The next question then becomes how many SCPT are required to estimate correlation length. This problem is explored through a simulation approach, using the

following procedure:

1. Select a random permutation order of the available SCPT database.
2. Estimate the sample correlation function for an ensemble of the first $i = 1$ SCPT in the permutation, and then estimate the optimal correlation length using the simulation-based approach described in Section 3.4.4.
3. Repeat Step 2 with an increasing number of SCPT in the ensemble; that is, for values of $i = \{2, 3, \ \ldots \ n\}$ where n is the total number of available SCPT.
4. Calculate the percent difference in estimated correlation length (δ) as a function of the number of SCPT in the ensemble (i), using the formula:

$$\delta(i) = 100\% \left| \frac{\hat{\theta}_i - \hat{\theta}_{i-1}}{\hat{\theta}_i} \right| \tag{3.15}$$

5. Repeat Steps 1 to 4 for the desired number of permutations of the SCPT database to consider.

Figure 3.13 illustrates the results of the above procedure for 5,000 trials only, since the total number of permutations if far too large to complete ($n = 206!$). For a small number of SCPT (i.e., less than 50), the distribution of estimated correlation lengths is highly skewed so that the median estimate is much less than the final estimate, yet the mean is larger. This is because the limited number of SCPT that predict very long correlation lengths result in a highly skewed distribution. However, the median estimate reaches the final value of $\hat{\theta}_{v,\,\ln Y} = 6$m once around 36 SCPT are included.

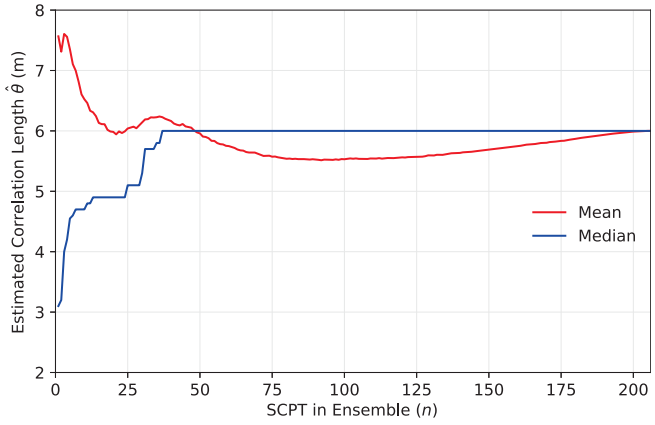

Figure 3.13: Percent difference in estimate of $\hat{\theta}_{v,\,\ln Y}$ for $n \leq 50$.

Figure 3.14 displays the percent difference in the estimate of $\hat{\theta}_{v,\,\ln Y}$ as a function of the number of SCPT in the ensemble, up to 50 SCPT, as well as the 7-point rolling average of the results. It is evident that using less than 20 SCPT is not likely to result in a reliable estimate. On average, a 9.5% difference is observed when using 20 SCPT in the ensemble, whereas the mean plus three standard deviations difference is roughly 53%.

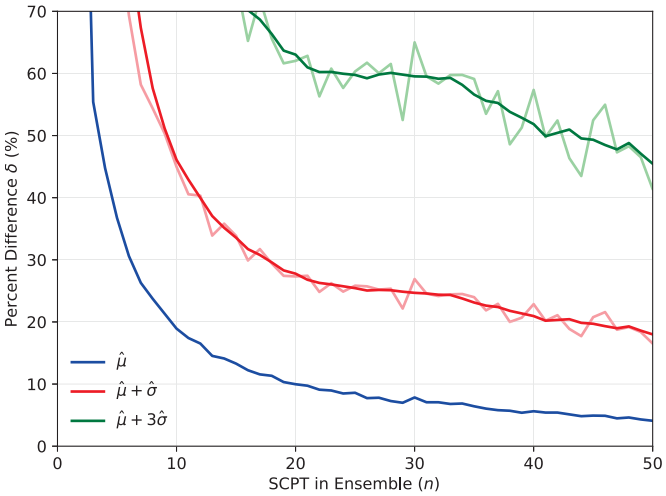

Figure 3.14: Percent difference in estimate of $\hat{\theta}_{v,\,\ln Y}$ for $n \leq 50$.

The mean percent difference greatly improves when using 50 SCPT, with a percent difference of only 4%, although the mean plus 3 standard deviations remains relatively high at 48%. Figure 3.15 shows how the percent difference continues to decrease for larger number of SCPT included in the ensemble. By using 100 SCPT, the mean plus three standard deviations drops below 20%, whereas by using 200 SCPT this becomes only 3.5% with a mean of 0.1%.

These results indicate that the database in this study, consisting of 206 SCPT, should be sufficient information to yield reliable estimates of correlation length when all the data is pooled together. Since projects A and C have the least number of SCPTs at 20 and 17, respectively, their estimates of correlation length may not be as reliable as those estimated for projects B and D.

The minimum number of SCPT is a somewhat arbitrary decision based on the level of precision that is desired before providing an estimate. Using anywhere between 20 and 50 SCPT should yield, at the very least, a good *ball-park* estimate of the correlation structure at a site, and using more that 100 SCPT seems like enough data

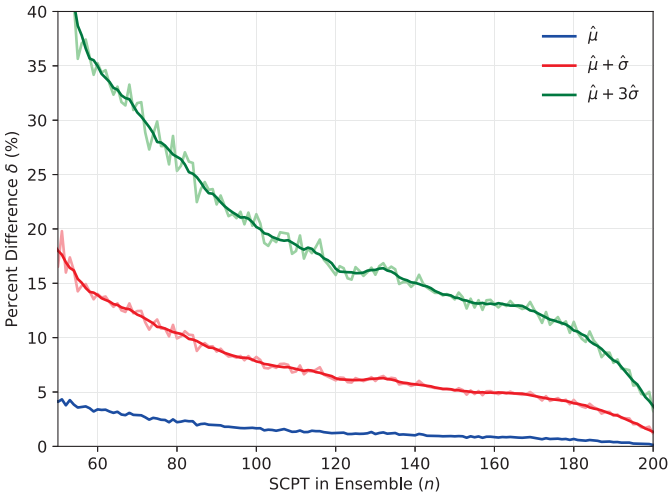

Figure 3.15: Percent difference in estimate of $\hat{\theta}_{v,\ln Y}$ for $n \geq 50$.

to provide a reliable estimate. However, the large amount of tests cannot correct for the fact that the correlation function is biased whenever there remains significant correlation throughout the sampling domain. For this reason, it is important to examine the sensitivity of the reliability of geotechnical systems to the selection of correlation length.

3.5 Randomized Soil Layering

In all the analyses so far, soil layering has not been considered so that all the collected SCPTs are assumed to consist of a single layer. However, it may be of interest to divide the collected data over depth, in terms of soil classification. Of particular interest in earthquake geotechnical engineering is whether the soils will behave as *clay-like* (cohesive) or *sand-like* (cohesionless) upon seismic shaking. This soil classification is only available at site A, and so the following sections only discuss that project. The layering available at 19 SCPTs are shown in Figure 3.16.

If the observed soil layering were fairly constant, then it might not be necessary to randomize soil stratigraphy, or one could simply model the thickness of each layer as a continuos random variable. However, the layering data available for project A shows a high variability in soil layering, both in terms of layer thickness and the depositional order. The goal here is then to randomize this layering such that this variability can be captured in probabilistic analyses.

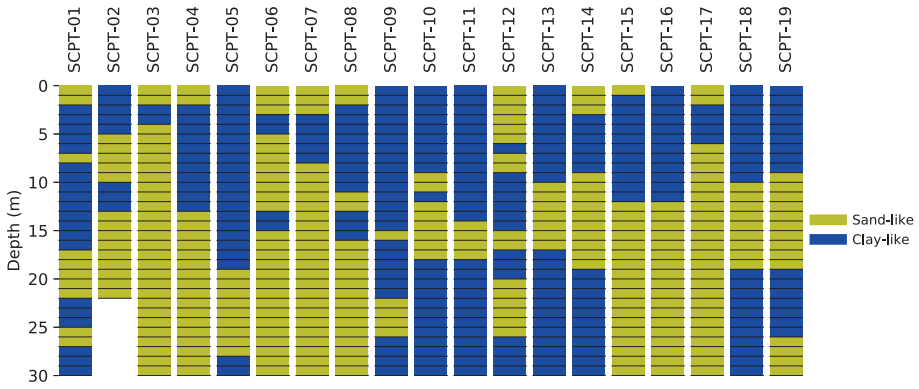

Figure 3.16: Available layering data at Project A.

3.5.1 Modeling Layer Transitions

To model the layering at project A, the following simplifying assumptions about the random process are adopted:

1. The layering is assumed to be discrete, such that changes in layering only occur in multiples of 1 meter intervals.
2. The thickness of the layers and depositional order is independent of depth. This assumption seems reasonable based on the data from project A.

Under the above assumptions, the layer variability at project A can be modelled through a discrete random process (X) with two possible states: *sand-like* (represented as $x = 0$), or *clay-like* (represented as $x = 1$). The transition between layers can then be modelled as a discrete-time, discrete-state Markov chain shown in Figure 3.17, where P_{ij} represents the probability of transitioning from state i to state j. Discrete-time (where *time* is actually *depth*) because the soil layering is randomized in constant depth intervals of 1 meter, and discrete-state because there are only two possible values of the random process.

The transition probabilities (P_{00}, P_{01}, P_{10}, P_{11}) are estimated directly from the layering data by counting the number of transitions that occur for each path, and dividing by the total number of trials. The resulting transition matrix is shown in Equation 3.16.

$$P = \begin{bmatrix} P_{00} & P_{10} \\ P_{01} & P_{11} \end{bmatrix} = \begin{bmatrix} 0.896 & 0.109 \\ 0.104 & 0.891 \end{bmatrix} \tag{3.16}$$

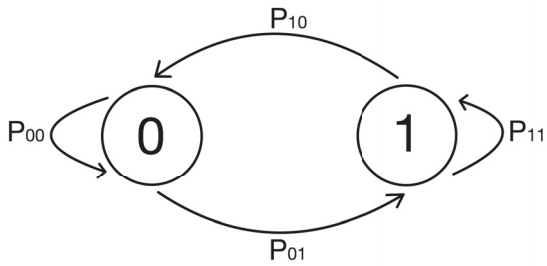

Figure 3.17: Markov chain where $x = 0$ is *sand-like* and $x = 1$ is *clay-like*.

To establish a starting point for the chain, the probability that the first layer is sand or clay is also estimated and shown in Equation 3.17.

$$P = \begin{bmatrix} P_0 \\ P_1 \end{bmatrix} = \begin{bmatrix} 0.526 \\ 0.474 \end{bmatrix} \tag{3.17}$$

3.5.2 Layering Simulation

Because there are only two possible states, the simulation process can be based on the Binomial distribution, in which the probability of *success* is defined as transitioning to clay-like ($x = 1$). To simulate random layering profiles, the following procedure is adopted:

1. Determine whether the first (top) element is sand-like ($x = 0$) or clay-like ($x = 1$) based on the initial state probabilities (Equation 3.17).That is, draw from the Binomial distribution using a probability of success equal to $P_1 = 0.474$.
2. Determine the probability that the next state, at a depth 1 m further down, is clay-like ($x = 1$) based on the current state. If the *current* state is $x = 0$, then the probability of *success* is $P_{01} = 0.104$, but if the *current* state is $x = 1$, then the probability of *success* is $P_{11} = 0.891$. Determine the next state by drawing from the Binomial distribution using the appropriate probability of *success* (transitioning to clay-like).
3. Repeat Step 2 $(n - 1)$ times in order to simulate a random layering process of n meters depth.

Sample simulations generated using the above procedure are shown in Figure 3.18, which can be repeated for any number of realizations that are needed in the analyses.

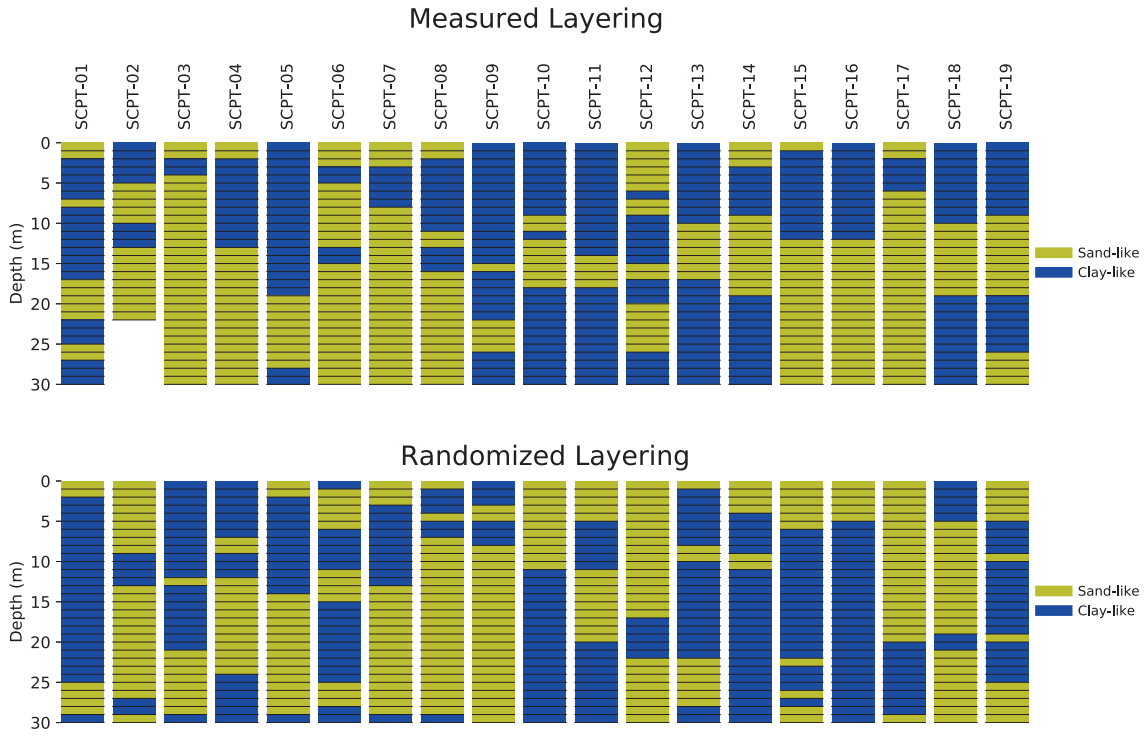

Figure 3.18: Measured versus randomized layering.

3.5.3 Layer-Based Shear Wave Velocity Randomization

If the layering is to be randomized at a site, then it is worth exploring whether each layer should have a different set of statistical properties. This is completed by following the same approach outlined previously (Sections 3.3 and 3.4), except that measurements taken in *sand-like* and *clay-like* soils are considered separately.

Figure 3.19 shows the trend obtained for *sand-like* against *clay-like* materials, and compares it against the trend obtained when both materials are pooled into a single dataset. Although there is a minor difference in the trends, it does reflect the expectation that *clay-like* soils are generally less stiff than sandy soils. Figure 3.20 compares the fitted lognormal distributions of the de-trended parameter Y.

Given that there are only 19 SCPT with layering information, estimating separate correlation structures for the *clay-like* and *sand-like* materials is going to result in estimates that are only a *ball-park* estimate of the true correlation structure. Nonetheless, Figure 3.21 shows the correlation lengths that are obtained from the bias-matched approach, and indicate that the *clay-like* and *sand-like* materials have very similar correlation lengths of $\hat{\theta}_{v,\,lnY} = 4.8$ m and $\hat{\theta}_{v,\,lnY} = 4.7$ m, respectively.

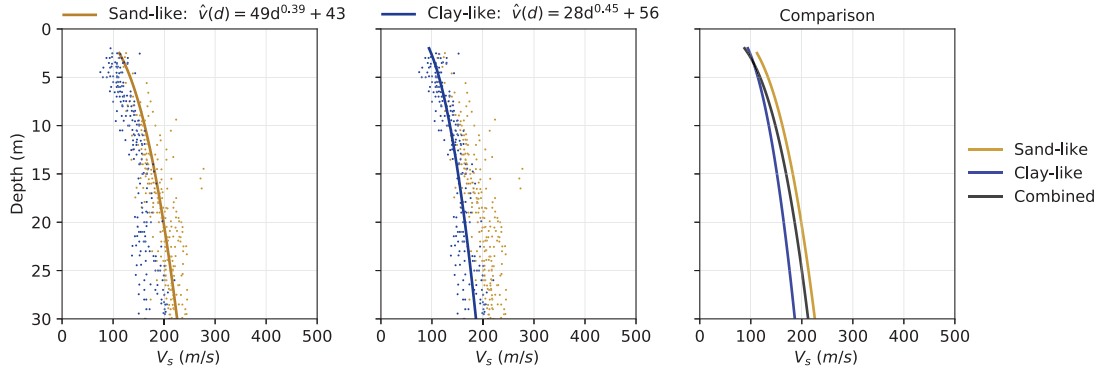

Figure 3.19: Trend estimate for sand-like and clay-like soils.

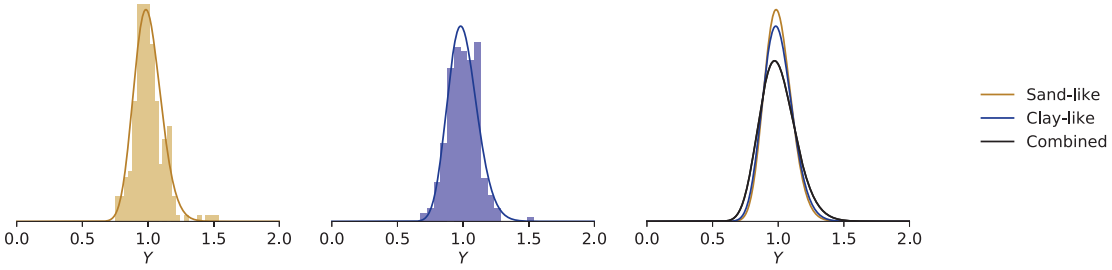

Figure 3.20: Distribution of Y for sand-like and clay-like soils.

Since there is very limited data to determine the estimates of correlation length, it might be preferable to assume that the materials have an equal correlation structure, and use the previous estimate of $\hat{\theta}_{v,\,lnY} = 5.3$ m that is derived from pooling the data together. Another advantage of pooling the data is that the simulated correlation function matches the sample one much better than when dividing the data points into different soil types.

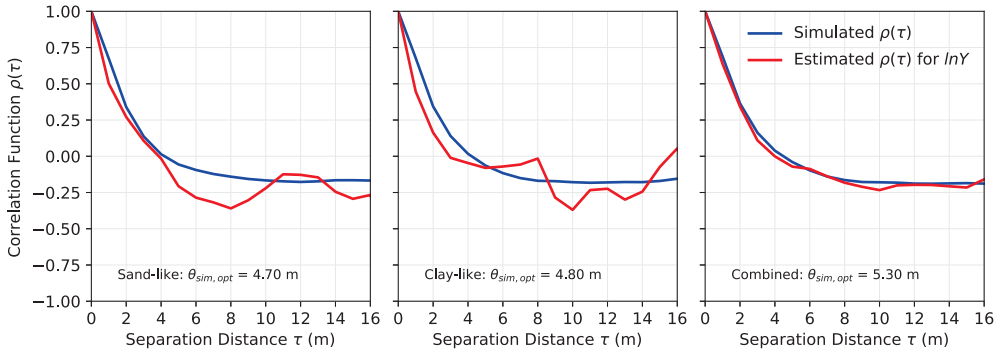

Figure 3.21: Correlation length for sand-like and clay-like materials.

3.6 Comparison to Toro's Model

3.6.1 Description of Toro's Model

Toro (1995) proposed a framework to randomize shear wave velocity and layer thickness, which has been widely adopted in 1-dimensional ground response analyses. Toro's model is a first-order, auto-regressive process in which shear wave velocity is modelled in a series of correlated, lognormally distributed layers of varying thickness. Toro (1995) used a database of 557 shear wave velocity profiles obtained mostly in California, to calibrate the model and provides recommended parameters based on v_{s30} site classification.

The location of layer boundaries is modelled as a Poisson process with a depth-dependent rate, as follows:

$$\lambda(h) = c_3 \left(h + c_1\right)^{-c_2} \tag{3.18}$$

Although the mean layer thickness is depth-dependent, the thickness of layer i is statistically independent of the thickness of layer $i-1$. The randomized shear wave velocity in layer i is calculated using:

$$lnV(i) = \mu_{lnV}(i) + Z_i \sigma_{lnV} \tag{3.19}$$

where $\mu_{lnV}(i)$ and σ_{lnV} are the mean and standard deviation of shear wave velocity in layer i in log-space, respectively. Finally, Z_i is a random variable that accounts for randomness about the mean and for the correlation between layers determined using:

$$Z_i = \begin{cases} \varepsilon_i & \text{for } i = 1 \\ \rho_{\text{IL}} Z_{i-1} + \varepsilon_i \sqrt{1 - \rho_{\text{IL}}^2} & \text{for } i > 1 \end{cases} \tag{3.20}$$

such that the first value of Z is taken as the standard normal ε_i (zero mean and unit variance) and subsequent values of Z are a function of the inter-layer correlation coefficient ρ_{IL}.

The inter-layer correlation model depends on the depth to the layer interface (d) and on the thickness of the layer (h), such that deep and thin layers are more strongly correlated than shallow, thick layers. The expressions for the inter-layer, depth-dependent, and thickness-dependent correlation coefficients are, respectively:

$$\rho_{\mathrm{IL}}(d,h) = [1 - \rho_d(d)]\,\rho_h(h) + \rho_d(d) \tag{3.21}$$

$$\rho_d(d) = \begin{cases} \rho_{200}\left[(d+d_o)/(d+200)\right]^b & \text{for } d \leq 200 \text{ m} \\ \rho_{200} & \text{for } d > 200 \text{ m} \end{cases} \tag{3.22}$$

$$\rho_h(h) = \rho_o exp\,(-h/\Delta) \tag{3.23}$$

where ρ_{200}, d_o, b, and Δ are model parameters that depend on site conditions.

Using the shear wave velocity data base, Toro (1995) provided recommended values for generic soil sites based on site classification according to v_{s30} .Table 3.6 summarizes the recommended parameters for the site classes that most closely match the v_{s30} of the sites considered in the current study.

v_{s30}	σ_{lnV}	ρ_o	Δ	ρ_{200}	h_o	b
$180 - 360$ m/s	0.31	0.99	3.90	0.98	0.0	0.344
≤ 180 m/s	0.37	0.00	5.00	0.50	0.0	0.744

Table 3.6: Recommended parameters in Toro (1995).

3.6.2 Sampling Length and Interval

An important difference between this study and the one completed by Toro (1995), is difference in type of shear wave velocity measurements that are available. Whereas the SCPTs considered in this study have a maximum depth of 30 meters, the dataset used to calibrate Toro's model reached up to 200 m in depth. The observation in Toro's study that deeper layers are more correlated is not important in only the top 30 meters of the soil profile, and so the depth-dependency of inter-layer correlation coefficients may not be justified here. This is explored further in Section 3.6.4

The second difference is that SCPTs measure shear wave velocity at constant intervals of 1 meter, whereas the profiles in Toro's study provide a constant shear wave velocity in layers of varying thickness. Therefore, whereas Toro's study required a thickness randomization model (shown earlier in Equation 3.18), the use of SCPTs results in a deterministic value of layer thickness equal to 1 meter. For this reason, the layering randomization model proposed by Toro is not explored further here.

3.6.3 Mean and Variance

In this study, the measured shear wave velocity data was de-trended by using a prescribed functional form (Equation 3.6). The process in Toro (1995) involves estimating the median of shear wave velocity separately at each layer instead, which corresponds to the mean shear wave velocity in logarithmic space. Figure 3.22 shows a comparison of these two approaches for each site.

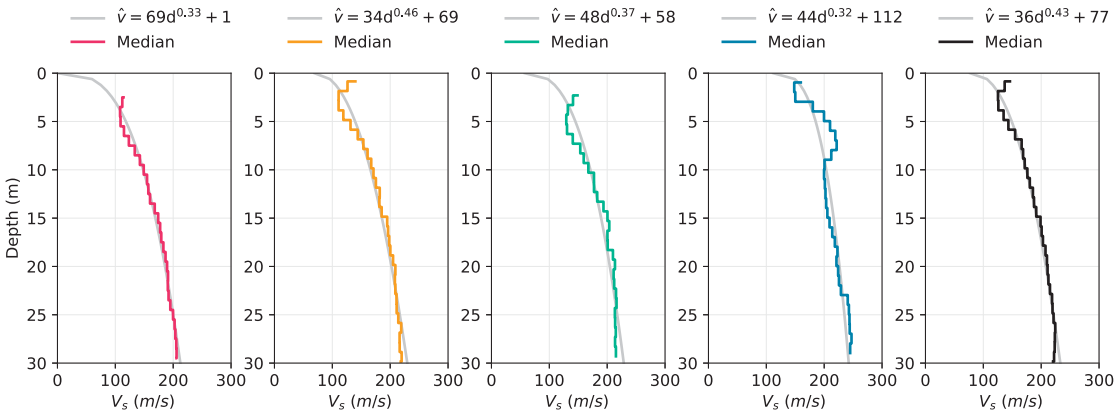

Figure 3.22: Comparison of mean estimation methods.

For the data considered, both approaches (prescribing a functional form or using binned intervals of depth) result in very similar estimations of the median shear wave velocity. The advantage of using binned intervals is that it eliminates the need to assume a somewhat arbitrary relationship between shear wave velocity and depth. On the other hand, using a prescribed functional form allows for an easier comparison to other sites, since a few constants (in this case three) can fully describe the mean at a site. Nonetheless, either approach would provide reasonable realizations of the random process.

Toro (1995) presented recommended values of standard deviation of shear wave velocity in logarithmic space as summarized earlier in Table 3.6. Figure 3.23 compares Toro's recommended values against those estimated in this study. It is unsurprising that the estimated standard deviation is smaller than that in Toro's study, particularly for sites A, B, and D. The data considered in Toro's study originated from a wide variety of projects across California, whereas the investigations at sites A, B, and C tested very similar geological conditions. Project B (corresponding to Hunter et al. (1999) data base), is the most variable data set considered here because it corresponds

to a regional study in the Lower Mainland. Even then, the variability in shear wave velocity is not as high as that observed by Toro (1995).

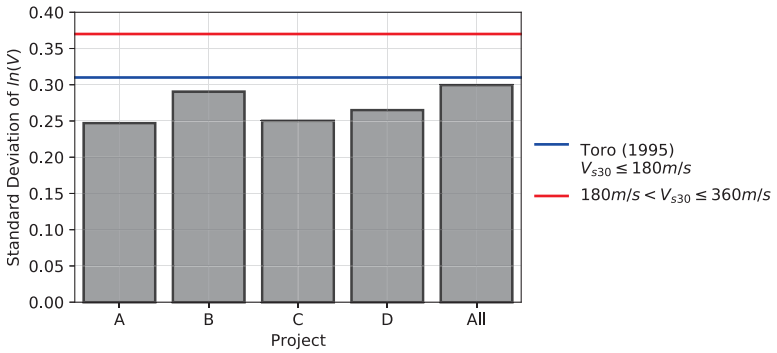

Figure 3.23: Comparison of standard deviation estimation.

3.6.4 Correlation Structure

In the Toro (1995) velocity model, correlation between adjacent layers is quantified through the inter-layer correlation coefficient (ρ_{IL}). Toro (1995) proposed the correlation coefficient be depth-dependent and thickness-dependent, so that deep and thin layers more strongly correlated than shallow, thick layers. Since the layer thickness in this study is set to 1 meter deterministically, the thickness-dependency is not explored here.

To examine the depth-dependency of the inter-layer correlation coefficient, this parameter is estimated as a function of depth for all the project sites, as shown in Figure 3.24. In this study, the inter-layer correlation coefficients do not show a clear trend with respect to depth. This could be because the data considered here is quite shallow in comparison to the data used by Toro (1995). Figure 3.24 compares the inter-layer correlation coefficients for generic sites recommended in Toro (1995). Because a strong trend with respect to depth is not observed, the ρ_{IL} are instead assumed to be constant at each site, and calculated by taking the average of the calculated ρ_{IL} values, as shown by the black dashed line in Figure 3.24.

Table 3.7 displays the estimated inter-layer correlation coefficients. Additionally, analogous correlation lengths are presented, which are calculated by re-arranging the Markov correlation function and using a lag of 1 meter, as follows:

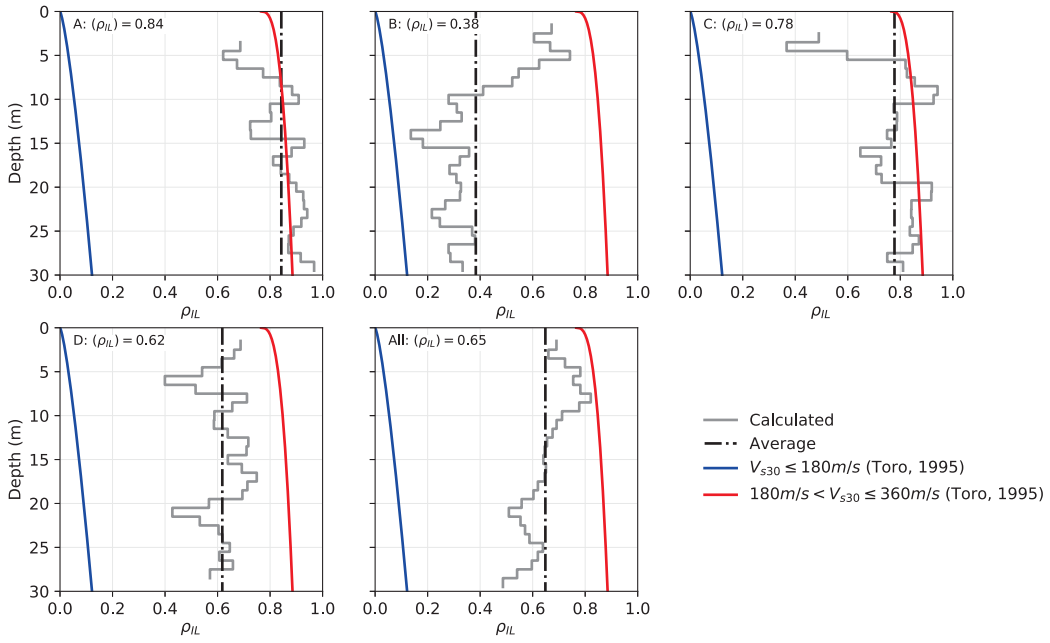

Figure 3.24: Estimation of inter-layer correlation coefficients.

$$\theta = \frac{-2}{ln(\rho_{\text{IL}})} \tag{3.24}$$

Project	$\hat{\rho}_{IL}$	$\hat{\theta}$
A	0.84	11.5
B	0.38	2.1
C	0.78	8.0
D	0.62	4.2
All	0.65	4.6

Table 3.7: Estimated inter-layer correlation coefficients.

3.7 Summary

The purpose of this chapter was to estimate the statistical properties of shear wave velocity based on available seismic cone penetration tests (SCPT) completed at sites in British Columbia, Canada. A total of 206 SCPT were available, which were collected at sites with average shear wave velocity in the top 30 meters between 80 m/s and 280 m/s.

Random shear wave velocity values as a function of depth ($V(d)$) were obtained by multiplying a lognormally distributed variable (Y) times the estimated mean trend

as a function of depth ($\hat{v}(d)$). This results in values that are strictly positive, and the values of Y can be reasonably simulated with stationary random fields.

Estimates of the mean trend, mean of Y and variance of Y are produced at four different sites, as well as for all the data combined. Additionally, estimates of the correlation length are produced in two ways: first, by directly fitting a theoretical correlation function to the sample correlation function obtained from the available data. The second method involved a simulation-based procedure that minimizes the error between the sample correlation function, and the correlation function that is obtained from simulating random fields with the same sampling length and sampling interval as the collected SCPT.

The bias-matched approach results in estimates of correlation length that are roughly twice as large as the direct fitting approach, and are likely a better representation of the *true* correlation structure at a site, since the simulated correlation function is able to more closely follow the sample correlation function. Estimating through the direct fitting approach is likely to yield poor estimates of correlation length for the data considered, since the sample correlation function becomes highly biased whenever the correlation length is not significantly smaller than the sampling domain. In either approach, the estimation of correlation length for soil properties remains very challenging, so that consideration of the sensitivity of the geotechnical analysis to the choice of correlation length may be an important factor to consider.

It was also shown that attempting to estimate correlation length using ensembles of less than 10 SCPT is likely to result in estimates that are grossly in error. However, using roughly 20 to 50 SCPT would likely yield, at the very least, a good *ball-park* estimate of the correlation structure, and using more than 100 SCPT is likely to yield a reasonably good estimate. It is important to note that these estimates are based SCPT with a maximum depth of 30 meters. If the SCPT used in the ensemble are significantly deeper, then it is likely that a smaller number of tests would be required in the ensemble to reach equivalent accuracy.

Finally, a comparison is presented against the auto-regressive model by Toro (1995), which is widely used in probabilistic 1-dimensional ground response analysis. Both the auto-regressive approach and the estimation of correlation length are able to represent the spatial variability of shear wave velocity, and both provide valid

methods to randomize this parameter. Although Toro's method can capture possible changes in correlation structure with depth and layer thickness, it is difficult to obtain enough data at any of the sites considered here to properly calibrate all the parameters in the model. Estimating the correlation length allows for the use of random field theory to randomize shear wave velocity, be it through covariance decomposition or local average subdivision. This has the advantage of being easily extended to multiple dimensions and can accommodate for different mesh resolutions.

The results from this chapter can be used to inform the selection of variance and correlation length that are used in probabilistic ground response analysis performed for sites with similar soil conditions, where not as much data is readily available. The bias-matched approach to estimate correlation length also provides an improvement to the traditional direct-fitting method, while still being a straightforward approach that can be readily completed by practicing engineers.

Chapter 4

Probabilistic Ground Response Analyses

4.1 Motivation

Ground response analysis (GRA) is used in earthquake geotechnical engineering problems to predict ground surface motions for the development of design response spectra. The latter are used to predict the dynamic stresses and strains for evaluation of liquefaction hazards, and to determine the earthquake-induced forces that can lead to instability in geotechnical systems (Kramer, 1996). The task in GRA is to determine the response of a soil mass subject to an earthquake motion at its base.

Probabilistic seismic hazard analysis (PSHA) estimates the likelihood that various levels of shaking will be exceeded at a given location, considering the contribution from potential earthquake sources near the site. PSHA provides a rational framework allowing the significant uncertainties in size, location, and rate of recurrence of earthquakes to be quantified and combined into estimates of seismic risk.

In contrast, GRA are primarily carried out deterministically. In the case of equivalent-linear GRA, the amplification of ground motions at a site depends on the shear wave velocity (V) and the shear modulus reduction and damping (MRD) curves. These quantities carry considerable uncertainties due to the spatial variability of soil conditions and the difficulty in predicting dynamic soil behavior. While this variability may be approximately quantified through parametric studies, a probabilistic approach has the advantage of fully capturing the effects of uncertainties in inputs on the resulting amplification of ground motions.

The goal of this chapter is to explore how the spatial variability of shear wave velocity affects the probability distribution of the amplification of ground motions and dynamic stresses at a site. A series of 1D and 2D equivalent-linear GRA are completed using the program QUAD4M (Hudson et al., 1994), modelling the spatial variability of shear wave velocity through random fields.

4.2 Description of Ground Response Analyses

To explore the effects of spatial variability of shear wave velocity on the amplification of ground motions, a series of GRA are completed, as follows:

1. A series of *deterministic* analyses are completed, in which the shear wave velocity at each element is assumed to be equal to the mean estimate at that depth. The analysis is repeated for a suite of 10 ground motions, which is typical in current geotechnical practice. The purpose of the deterministic analysis is to have a baseline to compare the probabilistic analyses against, both in terms of the mean response and the uncertainty around this estimate.
2. A *base-case* probabilistic GRA is completed. In this scenario, the statistical properties of shear wave velocity estimated in Chapter 3 are used to produce 1,000 random field realizations, which are then each analysed using QUAD4M. The results of this analysis allow for an estimate of the mean and variance of dynamic stresses as well as ground motion amplification. Because of the high computation time required, these analyses are completed using a single ground motion instead of the full suite of motions used in the deterministic analyses.
3. Finally, the sensitivity of the probabilistic GRA results to the choice of random field parameters is explored. The coefficient of variation of shear wave velocity (ν_V), correlation length ($\theta_{\ln Y}$), and correlation anisotropy (r) are systematically varied to explore the effects on their results.

All the analyses described above are completed for 1D and 2D models, in order to also explore how the selection of model geometry affects the results.

A maximum depth of 30 meters is considered to be consistent with the SCPT data in Chapter 3. The following sections describe the finite element program QUAD4M, the model inputs and outputs, as well as the process to automate the probabilistic GRA.

4.2.1 QUAD4M Description

The finite element program QUAD4M, initially developed by Idriss et al. (1973) and later modified by Hudson et al. (1994), was used to complete the ground response analyses presented in this chapter. QUAD4M evaluates the seismic response

of soil structures through a dynamic, time-domain, finite element procedure. It is an equivalent-linear program, such that the nonlinear properties of soils are incorporated through shear modulus reduction and damping (MRD) curves.

The finite element procedures uses a system of equations represented in matrix form as follows:

$$[M]\ddot{u} + [C]\dot{u} + [K]u = [M]\ddot{u}_g \tag{4.1}$$

where $[M]$ is the mass matrix, [C] is the damping matrix, [K] is the stiffness matrix, and $[M]\ddot{u}_g$ is the load vector given by the product of the mass matrix and the ground acceleration. The selection of QUAD4M to complete the analyses was based on the following considerations:

- Many linear-equivalent software that are commonly used to complete GRA are limited to 1D problems, such as SHAKE2000 and DEEPSOIL. Because this study aims to explore how the 2D soil variability affects the amplification of ground motions, it is not feasible to use these software.
- Although more recent commercial software are available such as RS2 from Rocscience, QUAKE/W from Geoslope, these programs are controlled through a graphical user interface, making the automation of the analyses cumbersome. In contrast, input files for QUAD4M can be readily generated through scripts, such that the execution of the program can be easily automated.
- Advanced numerical models employed in Plaxis or FLAC could be used to model the full non-linear response of the system. However, the large runtime of these model, in the order of hours or days, prohibits the simulation of thousands of realizations within a feasible time frame. Additionally, purchasing of the commercial licenses can be cost-prohibitive, and does not generally allow for multiple models to be run in parallel.
- Another alternative is the software OpenSees (McKenna, 2011), an object-oriented and open-source program that allows for the analysis of geotechnical systems subjected to earthquake ground motions. OpenSees allows users to perform finite element analysis using a scripting language, and is therefore well-suited for automating and parallelizing the analyses. The program also includes

robust constitutive models that can characterize the non-linear behavior of soils. However, the calibration of constitutive models for non-linear analyses generally requires a larger number of parameters than are needed for equivalent-linear methods. Such analyses may be a possible extension of the work presented here.

4.2.2 Model Inputs

Input Ground Motions

The ground motions considered in this study consist of a suite of 10 acceleration time histories which correspond to crustal and inslab earthquakes, matched to a return period of 475 years for a location in the Lower Mainland of BC. The time histories were first spectrally matched to Site Class C conditions and to the seismic hazard at the site. However, because site class C conditions are deeper than 100 meters, deterministic 1-dimensional ground response analyses were completed and time histories were extracted at a depth of 30 meters. The acceleration time histories and the corresponding spectral acceleration response spectra are displayed in Figure 4.1.

Soil Stiffness and Damping

Soil stiffness is described by the shear modulus (G), which is the ratio of shear stress (τ) to shear strain (γ). Material damping is described by the damping ratio (ζ), which is a ratio of the dissipated energy to the maximum elastic strain energy at a given strain amplitude.

Due to the non-linearity of soil, both the shear modulus and the damping ratio depend on the level of shear strain that the soil mass is experiencing. At low strains, the shear modulus is at its maximum but decreases as the strain amplitude increases, whereas the damping ratio is at its minimum at low strains and increases for larger shear strains. For this reason, the characterization of soil stiffness during cyclic loading requires establishing a maximum shear modulus at low strains (G_{max}), as well as a description of how the shear modulus and damping ratio change as a function of shear strain amplitude.

Since most geophysical tests induce shear strains lower than 10^{-6}, the shear wave

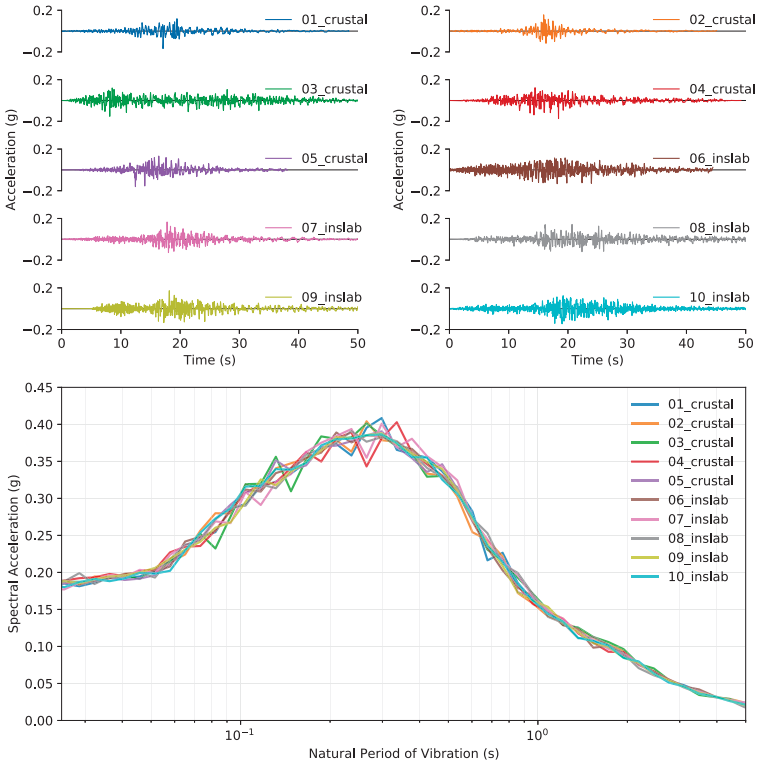

Figure 4.1: Acceleration time histories (above) and response spectra (below).

velocity measured during the field investigations can be used to directly estimate the maximum shear modulus (Kramer, 1996). The relationship between shear wave velocity and maximum shear modulus is as follows:

$$G_{max} = \rho \, V^2 \tag{4.2}$$

where ρ is the material density. Note that the maximum shear modulus is also sometimes referred to as the small strain shear modulus (G_o). Randomizing the shear wave velocity of the soil materials therefore results in randomized values of initial soil stiffness.

The change in stiffness and damping as a function of strain is described through the shear modulus reduction and damping curves (MRD), where the shear modulus reduction is described as the ratio of secant shear modulus to the initial, maximum shear modulus (G/G_{max}). The MRD curves are generally developed empirically based on laboratory tests. In this study, the curves proposed by Darendeli (2001) are used, which were based on numerous soil samples tested at the University of Texas at

Austin. The shear modulus reduction curve is defined as follows:

$$\frac{G}{G_{max}} = \frac{1}{1 + \left(\frac{\gamma}{\gamma_r}\right)^{0.919}} \tag{4.3}$$

$$\gamma_r = \left(0.0352 + 0.00101 \ \cdot \mathrm{PI} \cdot \mathrm{OCR}^{0.325}\right) \left(\sigma_o'\right)^{0.348} \tag{4.4}$$

where γ_r is the reference strain, σ_o' is the mean effective confining stress, OCR is the soil over-consolidation ratio, and PI is the plasticity index. The material damping curve is defined as the addition of the small-strain material damping ratio (ζ_{min}) and the scaled Masing behavior damping, which is determined as follows:

$$\zeta = \left(0.633 - 5.66 \times 10^{-3} \ln(N)\right) \left(\frac{G}{G_{max}}\right)^{0.1} \zeta_{masing} + \zeta_{min} \tag{4.5}$$

where:

$$\zeta_{min} = \left(0.801 + 0.0129 \ \mathrm{PI} \cdot \mathrm{OCR}^{-0.107}\right) \left(\sigma_o'\right)^{-0.289} \left(1 + 0.292 \ \ln \ f\right) \tag{4.6}$$

$$\zeta_{masing,\ a=1} = \frac{1}{\pi} \left(4 \frac{\gamma - \gamma_r \ln\left(\frac{\gamma + \gamma_r}{\gamma_r}\right)}{\frac{\gamma^2}{\gamma + \gamma_r}} - 2 \right) \tag{4.7}$$

$$\zeta_{masing} = 1.02 \ \zeta_{masing,a=1} - 6.76 \times 10^{-3} \ \left(\zeta_{masing,a=1}\right)^2 + 6.15 \times 10^{-5} \ \left(\zeta_{masing,a=1}\right)^3 \tag{4.8}$$

where f is the loading frequency in Hz and N is the number of loading cycles. Because the analysis in this study assume that the underlying soil is sand, the plasticity index is selected as zero and the over-consolidation ratio as one. Darendeli (2001) observed that the loading frequency and number of cycles have minor effects on the MRD curves, and recommends values of $f = 1$ Hz and $N = 10$ cycles, which were adopted in this study. As a result, the MRD curves used in this study vary across the finite element mesh depending solely on the mean confining stress at each element. In order to avoid excessive shear strains near the surface, the minimum confining stressed used to develop the MRD curves is set as 100 kPa. The resulting MRD curves are displayed in Figure 4.2 for the 1D and 2D models.

A unique advantage of using the Darendeli (2001) curves is that curves are provided not only for the mean estimate of the MRD curves, but also for the uncertainty associated with these estimates. Although this study only considers the mean of the curves, the analyses can be easily extended to include this uncertainty in further studies.

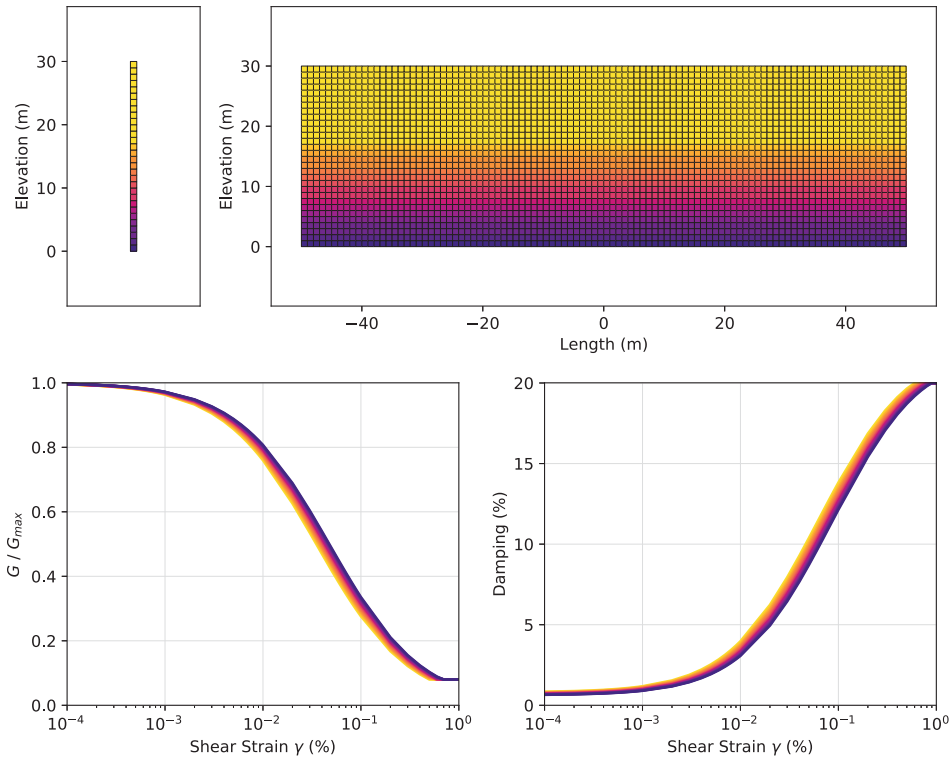

Figure 4.2: 1D (top-left) and 2D (top-right) mesh, and MRD curves (bottom).

Geometry

The 1D models consist of a 30 meter soil column, with elements of 1 meter height and 1 meter width. The use of a 1D model is advantageous in that the runtime is significantly shorter than the 2D models, therefore more configurations of the random field parameters can be explored without requiring excessive computing power. The 2D models are 100 meter in length and 30 meter deep, with level ground, and 1 meter by 1 meter elements. The advantage of 2D models is that more complex geological settings can be modelled, as well as particular geotechnical systems such as retaining walls or sloping ground. The finite element meshes used in this study are displayed

in the top row of Figure 4.2.

In the 1D model, the finite element nodes that are above the base of the column are free to move in the horizontal direction but are constrained in the vertical direction. At the base of the model, the input motions are applied by prescribing the horizontal acceleration time history. Therefore, the soil column represents horizontal acceleration time histories that propagate vertically. In the 2D model, nodes at the left and right boundaries are constrained in the vertical direction, so that the nodes can only move horizontally. The nodes at the base are prescribed the acceleration time history, and all other nodes are free to move in the horizontal and vertical directions.

4.2.3 Model Outputs

There are three main output parameters that are of interest in this study. The first is the cyclic stress ratio (CSR) as a function of depth, which is calculated as follows:

$$\text{CSR} = 0.65 \, \frac{\tau_{max}}{\sigma'_{vc}} \tag{4.9}$$

where τ_{max} is the maximum shear stress experienced at a given soil element and σ'_{vc} is the effective vertical stress. The choice of 0.65 represents a reference stress level that is conventionally used in liquefaction evaluation procedures (Idriss and Boulanger, 2008). This study will examine not only the estimate of the mean of CSR ($\hat{\mu}_{\text{CSR}}$), but also its estimated coefficient of variation ($\hat{\nu}_{\text{CSR}}$). In the 2D models, the CSR is extracted only for the soil column at the middle of the finite element mesh.

The second output parameter of interest is the peak ground acceleration (PGA) at the top of the model, which is assumed to follow a lognormal distribution. The mean ($\hat{\mu}_{\text{PGA}}$) and coefficient of variation ($\hat{\nu}_{\text{PGA}}$) are estimated from the probabilistic analysis results.

The final parameter of interest is the spectral acceleration response spectra (SA), which is calculated from the acceleration time history extracted at the middle of the models (i.e., the top element in the 1D models and the middle top element in the 2D models). The response spectra describes the maximum acceleration of a 5% damped single degree of freedom oscillator when subject to the acceleration time history.

4.2.4 Automation

To complete the probabilistic analysis, it is necessary to automate the generation, running, and post-processing of QUAD4M models. Since each probabilistic analysis generally consists of some 1,000 realizations, it is unfeasible to complete these tasks manually. This section describes the procedure that was used to automate the analysis, with the code attached as an electronic supplement.

QUAD4M is available as an executable program that requires three input files, described in detail in the QUAD4M manual (Hudson et al., 1994). The main purpose of each file is as follows:

- **Input File (*.q4r)**: includes computational switches to control the input and output options, as well as specifying the location of the acceleration file. It includes a table specifying the element properties (node numbers, unit weight, Poisson ratio, stiffness, and soil number), as well as a table with nodal properties (coordinates, boundary conditions, initial conditions, and output options).
- **Soil Properties File (*.dat)**: contains the shear modulus reduction and damping curves to account for soil non-linearity. These are mapped to the finite element mesh using the *soil number* column in the element table of the input file.
- **Acceleration File (*.shk)**: provides the acceleration time history to be used in the analysis.

The generation of QUAD4M input files is completed by using a combination of AutoCAD, Fortran functions, and Python scripts. The steps of the generation process are as follows:

- **Model Geometry**: The geometry of the model is first created in a CAD software and then exported in DXF format. If necessary, different soil layering can be specified by separating the geometry into different CAD layers. A Python module (*geometry.py*) was written to read the DXF file using the *EZDXF* library (Moitzi, 2021) and to create tables with node and element geometry information. An output DXF file is also generated, which contains the node and element numbering.

- **Nodal Properties**: The Python module *props_nodes.py* populates the nodal properties which include boundary conditions, acceleration output options, and initial conditions.
- **Element Properties**: The deterministic soil properties are prescribed for each layer, and then mapped to the finite element mesh. The module *props_elems.py* populates the element properties such as estimated confining stress, stress output options, and initial strain levels.
- **Shear Modulus Reduction and Damping Curves**: The *props_elems.py* module also determines the list of unique element properties, and calls the Python module *darendeli_2011.py* to determine the required number of MRD curves. Each element is then assigned a soil number that maps the appropriate curves to that element.
- **Random Field Generation**: random field realizations are generated for user-specified mean, variance, and correlation structure. The random fields are generated using the Fortran functions LAS1G or LAS2G (Fenton and Griffiths, 2008) available at: *http://random.engmath.dal.ca/rfem*. These functions are wrapped in using the F2PY module (Peterson, 2009). Finally, the module *props_elems.py* maps the random field values to the elements table and converts values of shear wave velocity to maximum shear modulus.
- **Generate Input Files**: Having specified all the necessary input properties, the module *genfiles.py* produces text files to be read by the QUAD4M executable.

The generation process is controlled by a script (*generation.py*), which calls the necessary modules and iterates through different geometries, random field properties, and acceleration time histories as needed. The module *runQ4Ms.py* runs a series of QUAD4M analyses in parallel using the threading and subprocess standard libraries in Python. The post-processing of output files is also completed immediately after running each analysis to avoid excessive storage. Post-processing is completed in the module *post_process.py*, which involves extracting the peak acceleration at each node, peak stress in each element, equivalent soil properties, as well as output acceleration and stress time histories. Finally, the module *extract_results.py* summarizes the key results from each analysis and computes the acceleration response spectra from the output acceleration time histories.

4.3 Deterministic Analysis

4.3.1 Description

Before delving into the probability-based ground response analyses, it is useful to carry out the analysis first through a deterministic approach, so that the results can be used as a point of reference against the probabilistic analysis. In the deterministic approach, the mean shear wave velocity (m/s) profile corresponds to the best-fit of all the shear wave velocity data discussed in Chapter 3 of this study, as follows:

$$\hat{v}(d) = 35.9d^{\,0.432} + 77.3 \tag{4.10}$$

where d is the depth below the ground surface (m). Current practice in GRA involves selecting a suite of 5 to 10 ground motions that have been matched to the seismic hazard at a given location. Therefore, GRA was completed for all the ground motions displayed in Section 4.2.2. It is important to note that only one of those ground motions is used when assessing the effects of spatial variability of shear wave velocity.

4.3.2 Results

When the mean shear wave velocity profile is used, the 1D and the 2D models should yield identical results since the 2D model is just a repetition of the 1D model. The results of the QUAD4M models correctly yielded equivalent results for both geometries; therefore, the results presented in this section apply equally to the 1D and 2D models. Figure 4.3 displays the cyclic stress ratio (CSR) profile. The mean CSR and standard deviation are estimated based on the results of the 10 ground motions. Shown in red is the result from the input motion used in the probabilistic analysis (*06_inslab*), which closely follows the mean CSR profile. Although the coefficient of variation of CSR changes with depth, it can be estimated to be around 7% when averaging over the entire profile. This value can be used to compare the effects of ground motion variability against the effects of shear wave velocity variability as explored in later sections.

Figure 4.4 shows the peak ground acceleration at the top of the models for each ground motion. The mean PGA calculated from all the ground motions is 0.22g, with a standard deviation of 0.03 g. The peak ground acceleration of the selected ground

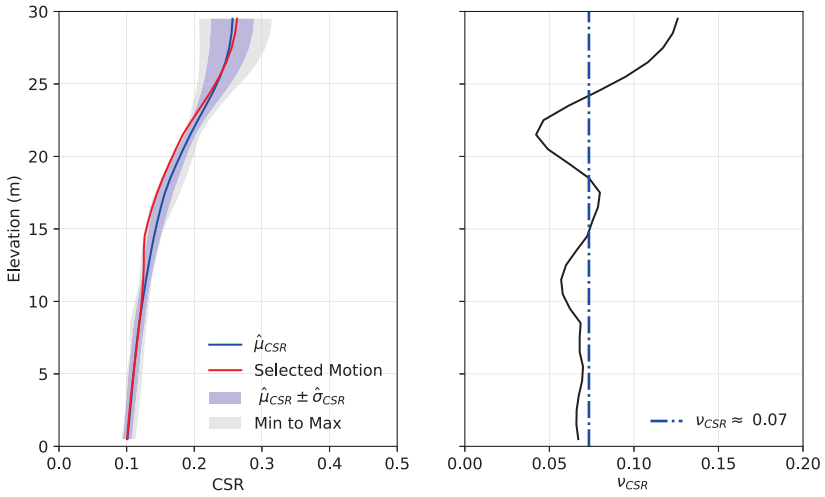

Figure 4.3: Cyclic shear stress ratio results.

motion is 0.224 g (shown in red) which will later be compared to the distribution of PGA that is obtained when randomizing shear wave velocity.

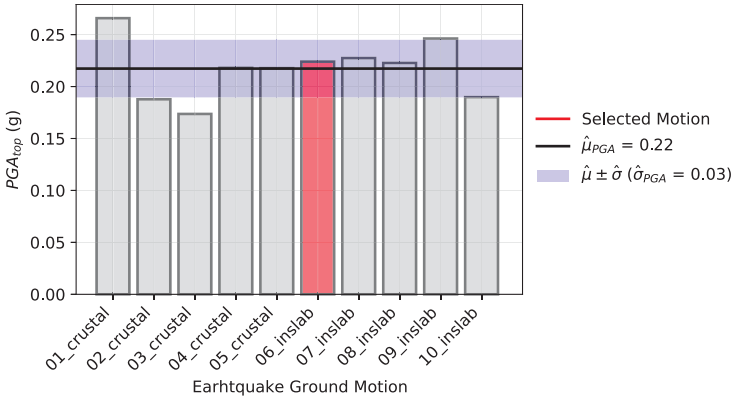

Figure 4.4: Peak ground acceleration results.

Finally, the response spectra is shown in Figure 4.5. The mean response spectra across the 10 ground motions is shown in blue, which is reasonably close to the response spectra of the selected ground motion. An important parameter in ground motion prediction models is the standard deviation of the response spectra in lognormal units (σ_{lnSA}), also shown in Figure 4.5. Although the standard deviation changes for different natural periods of vibration, it is approximated by taking the average across all periods as $\sigma_{lnSA} \approx 0.13$.

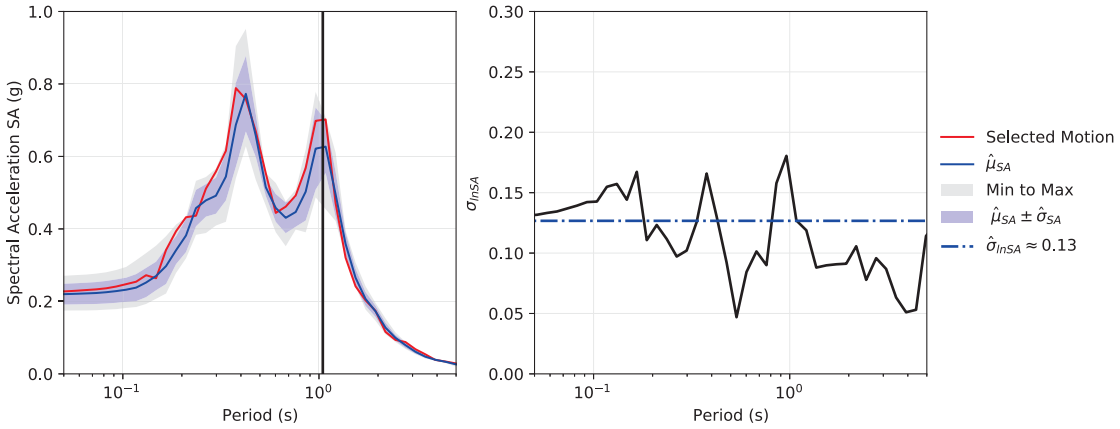

Figure 4.5: Spectral acceleration response spectra.

4.4 Base-Case Probabilistic Analysis

4.4.1 Description

The *base-case* probabilistic analysis consists of randomizing shear wave velocity using the results of the statistical analysis in Chapter 3 of this study. The parameters used correspond to the results obtained when pooling all the available SCPT data together, and represent the *best-guess* at the parameters based on real shear wave velocity data.

The simulation of random shear wave velocity fields starts by establishing the mean field ($\hat{v}(d)$), which is equivalent to the one used in the deterministic analyses (Section 4.3). Then, a lognormally-distributed random field, Y, is generated using local average subdivision (Fenton and Vanmarcke, 1990). The parameters used in the base-case scenario are a mean $\mu_Y = 1$, standard deviation $\sigma_Y = 0.24$ and vertical correlation length $\theta_{v,\ln Y} = 6$ m.

Since there was not sufficient data to estimate horizontal correlation length with a reasonable degree of accuracy, an assumed value needs to be used here for the 2D models. Estimates of vertical and horizontal correlation length of soil properties available in literature, summarized by Phoon and Kulhawy (1999) and Cami et al. (2020), show that the horizontal correlation length is larger than the vertical one by roughly one to two orders of magnitudes, although there is a very high variability in the estimates. Here, it is considered reasonable to assume a horizontal correlation length that is 10 times larger than the vertical one, such that $\theta_{h,\ln Y} = 60$ m. The sensitivity to the choice of horizontal correlation length is explored later (Sections 4.6

and 4.7).

The generation process is illustrated in Figure 4.6 for a sample random field realization of the 2D model. Whereas the mean field $\hat{v}(d)$ remains the same for all realizations, the lognormal field of Y changes for each analysis, resulting in different realizations of random shear wave velocity. An ensemble of 1,000 realizations were generated for each of the 1D and 2D models. Figure 4.6 also displays the locations where CSR profiles and peak ground accelerations are extracted after completing the ground response analyses.

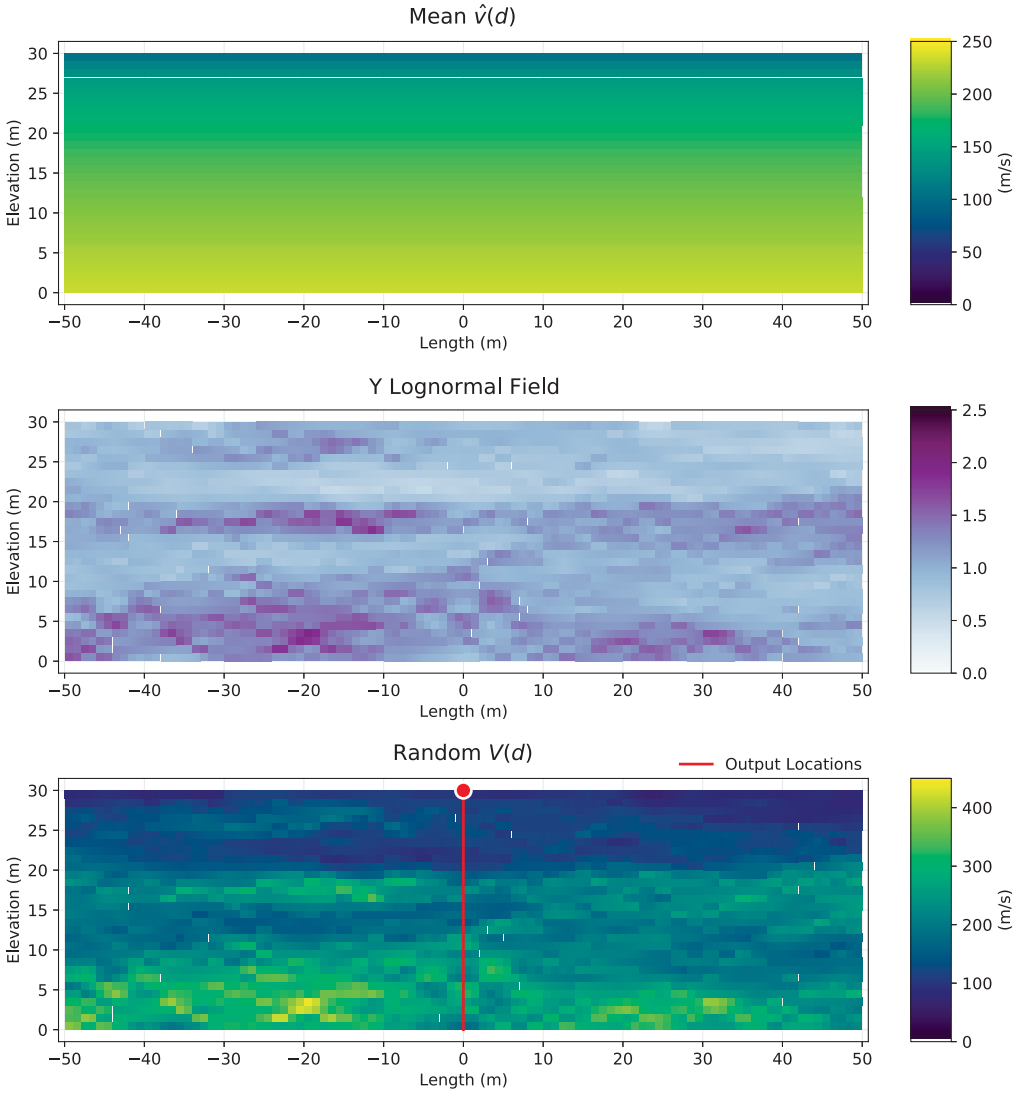

Figure 4.6: Generation of random fields of $V(d)$.

Finally, Figure 4.7 compares the shear wave velocity means and standard deviations estimated from the ensemble of 1,000 realizations of the random fields against the random field generation inputs. In the case of the 2D model, the shear wave velocity is extracted at the soil column in the middle of the model. The mean, standard deviation, and correlation length estimated from the ensemble of realization closely match the input parameters.

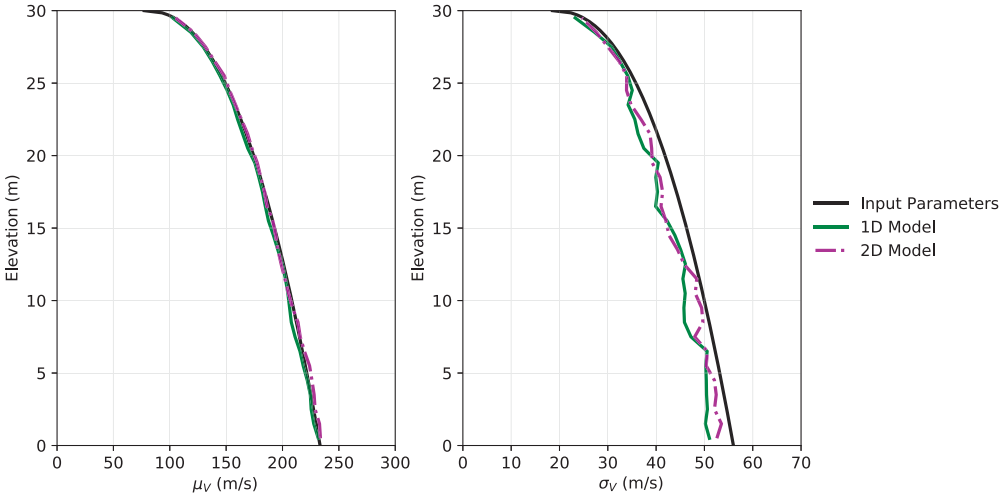

Figure 4.7: Comparison of input and simulated parameters.

4.4.2 Results

Figures 4.8, 4.9, and 4.10 display the cyclic stress ratio, peak ground acceleration, and response spectra, respectively, from the 1D and 2D probabilistic ground response analyses. The results include the mean and standard deviation estimates obtained from the 1,000 realizations of randomized shear wave velocity and compares them against the deterministic results that were obtained in Section 4.3.

The cyclic stress ratio (CSR) profiles, shown in Figure 4.8, show that randomizing shear wave velocity results in a reduction of the mean CSR profile in both the 1D and 2D models when compared against the deterministic case. The difference between the probabilistic and deterministic analyses is smallest at the base of the model, and generally becomes larger towards the ground surface.

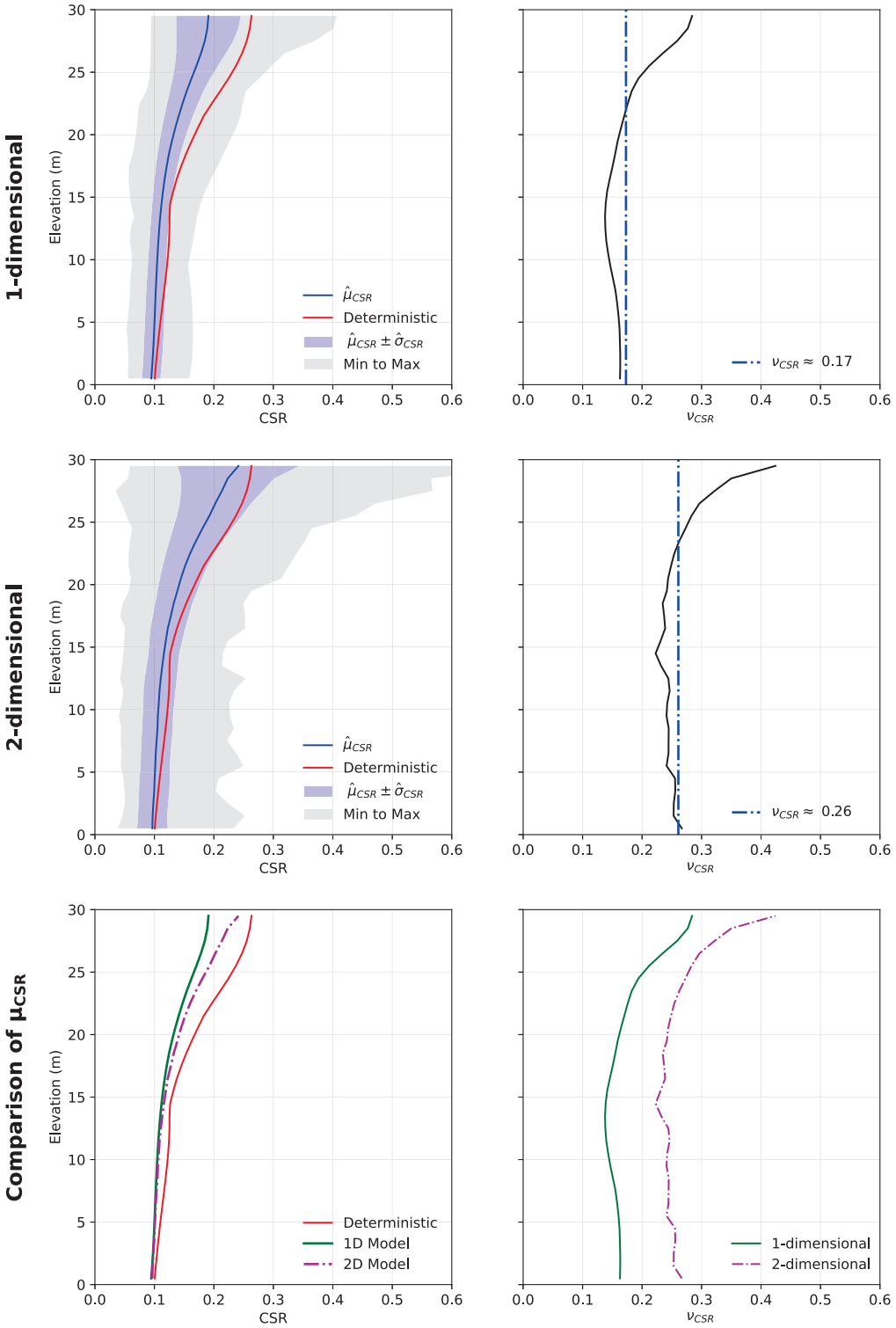

Figure 4.8: CSR profiles for 1D and 2D models.

The reduction in CSR when introducing spatial variability of shear wave velocity is due to the fact that randomizing shear wave velocity allows for the presence of softer layers, which result in larger shear strains and therefore more soil damping.

The reduction in mean CSR is slightly more prominent in the 1D models than in the 2D models. While in the 1D model the mean CSR is between 71% and 93% of the deterministic response, the mean CSR in the 2D model is between 78% and 95%. The coefficient of variation of CSR is generally smaller at the base of the ground model, and increases towards the ground surface. In the 1D case, the coefficient of variation ranges between 14% and 28%, with an average of 17%. In the 2D case, the coefficient of variation ranges between 22% and 42% with an average of 26%. In other words, variability in CSR is much larger in the 2D models tha in the 1D model.

The distribution of peak ground accelerations at the ground surface are displayed in Figure 4.9. The results are consistent with those observed in the CSR profiles: when shear wave velocity is randomized, the mean PGA drops, compared to the deterministic case, in both the 1D and 2D models; however, there is a larger reduction in PGA in the 1D model.

The histogram of the PGA at the surface (shown in grey in Figure 4.9) closely follows a lognormal distribution (shown in black). In the 1D probabilistic model, the mean PGA is 73% of the deterministic PGA. The deterministic PGA roughly correspond to the 90^{th} percentile of the probabilistic distribution. In the 2D models, the probabilistic PGA is 85% of the deterministic PGA, and the deterministic PGA is approximately the 77^{th} percentile of the random PGA.

The results from the probabilistic analysis indicate that using the mean shear wave velocity as the *representative* or *characteristic* value in a deterministic ground response analysis does not lead to the *true* (or, at least, a better approximation of the) mean response for the ground motion considered. Instead, a significantly higher PGA than the mean is obtained when only the deterministic analysis is used. In other words, using only a deterministic analysis may be considered advantageous in that the results are a conservative estimate of the true mean response. However, this also means that a deterministic analysis will almost certainly result in more expensive design or remediation.

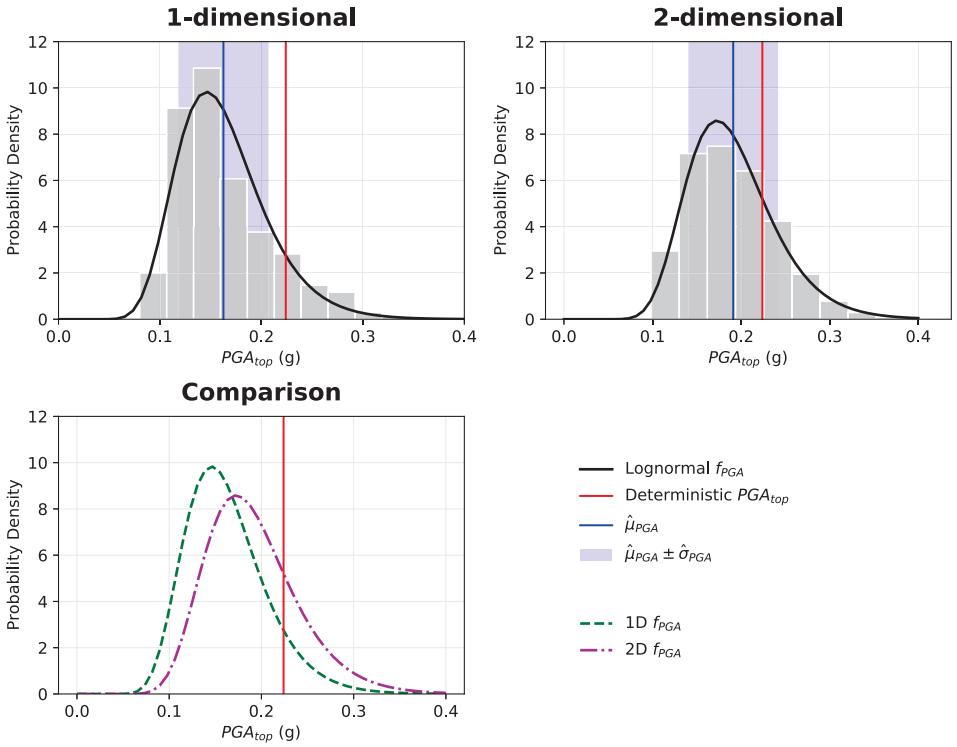

Figure 4.9: Peak ground acceleration for 1D and 2D probabilistic models.

Finally, Figure 4.10 displays the results in terms of the response spectra obtained at the surface. The comparison between the mean probabilistic (blue) and deterministic (red) responses is a function of the natural period of vibration. For periods that are below the natural period of vibration of the system, shown as a black vertical lines for the case considered here, the probabilistic analysis suggest a reduction of the mean response spectra. The reduction is particularly evident at the periods with the highest response in the deterministic analysis ($T = 0.4$s and $T = 1.0$s). As with the CSR and PGA results, the 1D model results in a larger reduction in mean response when compared to either the deterministic or the 2D results for most periods. The largest reduction in the 1D model occurs at a period of 0.38 s, where the probabilistic spectral acceleration is 50% of the deterministic one. In the 2D model, the largest reduction occurs at a period of 0.96s, where its mean spectral acceleration is 47% of the deterministic response. The standard deviation of spectral acceleration, in log-normal units, is also shown in Figure 4.10, with very similar values obtained in the 1D and 2D models.

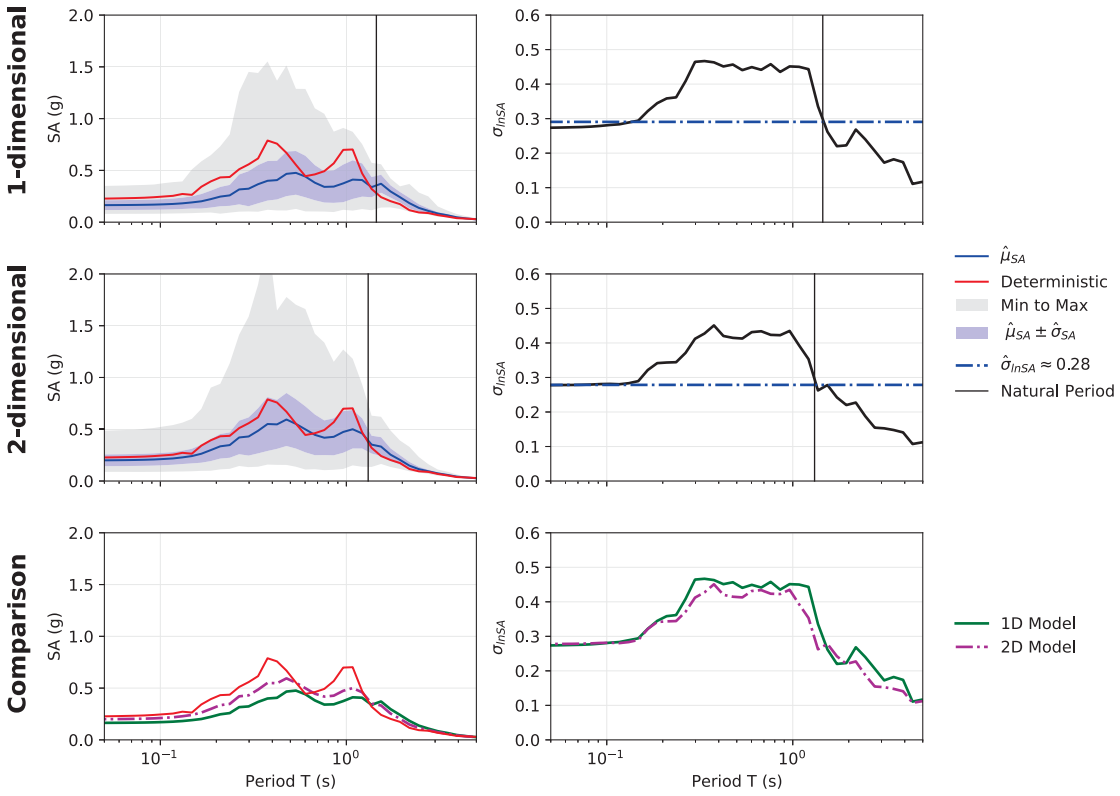

Figure 4.10: Response spectra from 1D and 2D probabilistic models.

4.5 Sensitivity to Coefficient of Variation

4.5.1 Description

In this section, the sensitivity of the probabilistic ground response analysis to variations in the coefficient of variation of shear wave velocity (ν_V) is explored. To do so, a series of probabilistic ground response analyses are completed using different coefficients of variation of shear wave velocity, the latter of which is equal to the standard deviation of Y (σ_Y) as described in Chapter 3.

For the 1D models, a total of 11 different coefficients of variation were considered ($\nu_V = \{0.01, 0.05, 0.10, 0.15, ..., 0.50\}$). Because the 2D models have a substantially longer run-time than the 1D models, it was too time consuming to consider as many sensitivity scenarios. For this reason, the 2D models only consider 6 coefficients of variation, ($\nu_V = \{0.01, 0.10, 0.20, ..., 0.50\}$). It is considered that this range covers most realistic coefficients of variation, and that larger values are unlikely to be found in practice. The correlation length of the random fields are kept constant, and equivalent to the ones used in the *base-case* analysis, at $\theta_{v,\,lnY} = 6$m and $\theta_{h,\,lnY} = 60$m.

Figure 4.12 displays sample realizations for four coefficients of variation considered. An ensemble of 1,000 realizations are generated for each set of random field parameters, such that a total of 17,000 ground response analysis are completed. As done with the *base-case* analysis, the CSR profile and PGA at the surface are extracted at the top middle of each realization of the 2D model. Figure 4.11 displays the mean and standard deviation that result from the ensemble of randomized shear wave velocity.

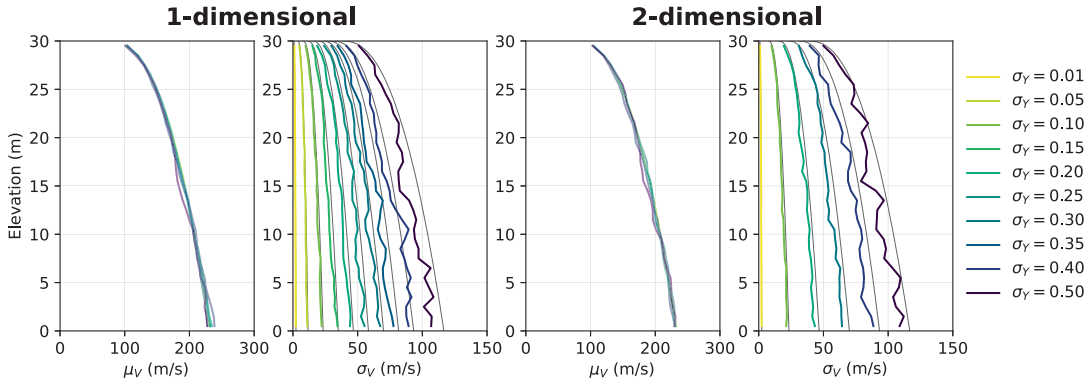

Figure 4.11: Simulated parameters of shear wave velocity for varying ν_V.

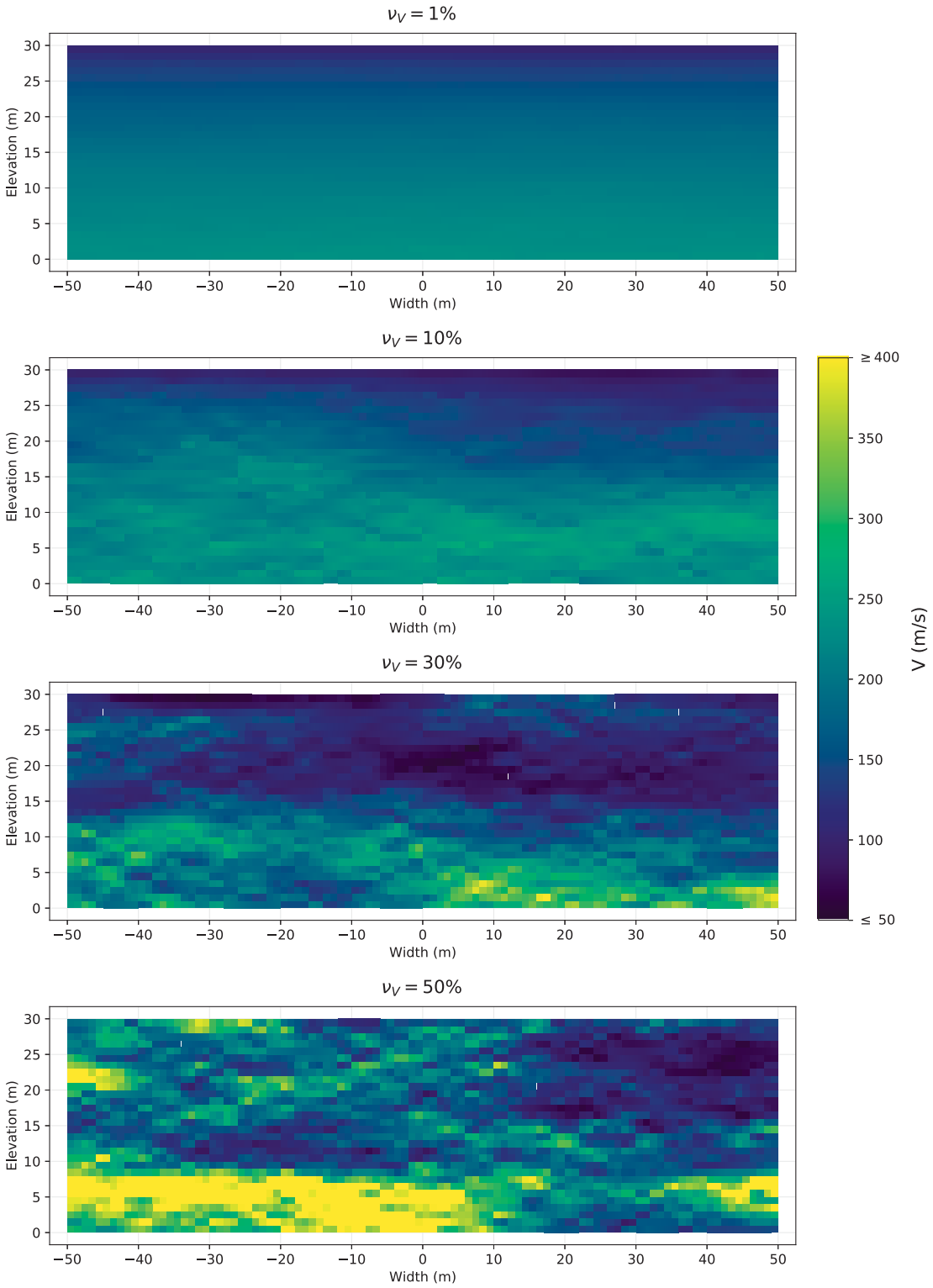

Figure 4.12: Sample realizations with varying ν_V.

4.5.2 Effects on CSR

Figure 4.13 displays the mean CSR profiles that result from the ensemble of 1,000 realizations for each choice of ν_V, for both the 1D and 2D models. As observed in the base case scenario, introducing shear wave spatial variability results in a reduction of the mean cyclic stresses in the system. In general, the reduction in mean CSR is smallest at the base of the model and increases towards the ground surface.

As expected, increasing the coefficient of variation of shear wave velocity results in an increased reduction in the mean CSR. In the 1D model, the overall shape of the mean CSR profile remains constant for all ν_V, with the most increase in CSR happening at elevations of 15 m to 25 m, with less increase in CSR happening on the top 5 meters of the profile. However, in the case of the 2D model, ν_V influences the shape of the CSR profile. For ν_V larger than 20%, the steepest increase in mean CSR occurs in the top 5 meters of the profile.

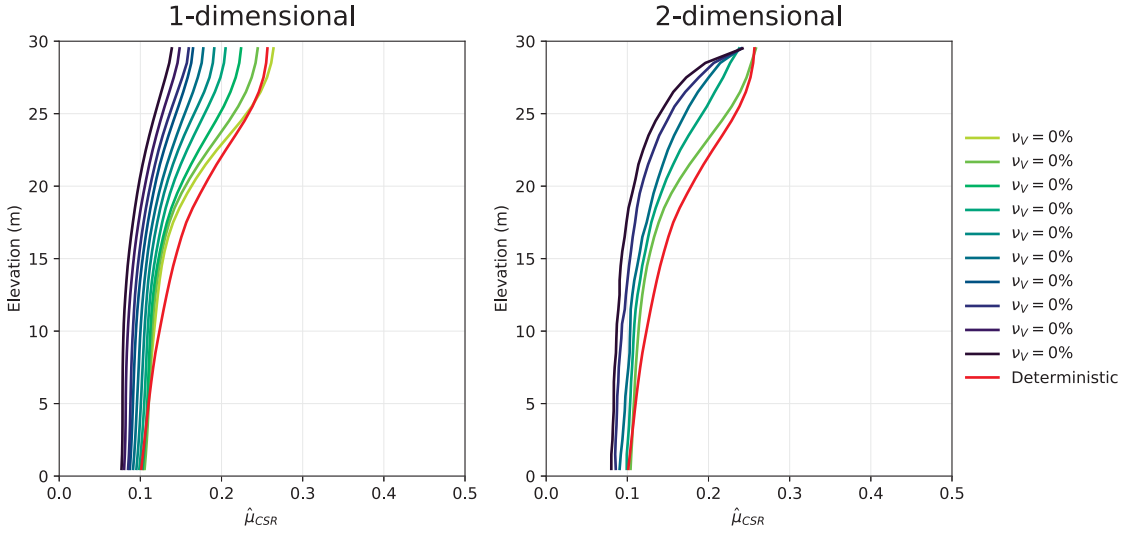

Figure 4.13: Sensitivity of mean CSR profile to ν_Y.

To summarize the sensitivity results, the average reduction in mean CSR profile and the coefficient of variation of CSR are plotted against ν_V, as shown in Figure 4.14. The solid dots in the figure represent estimates from the simulations while the solid lines are the best least square regression fits. The reduction in mean CSR is defined as the average ratio of the mean probabilistic profile to the deterministic one. The reduction seen in the 1D and 2D models can be reasonably represented through a linear fit for the range of ν_V considered. At a coefficient of variation $\nu_V = 0$, the

shear wave velocity profile should revert back to the deterministic one. Therefore, the reduction ratio is assumed to be equal to 1 in both the 1D and 2D models. At $\nu_v = 50\%$, the mean CSR profile is equal to 0.60 and 0.68 of the deterministic value in the 1D and 2D models, respectively.

The average coefficient of variation of CSR (ν_{CSR}) is highly dependent on ν_V, but also on whether the problem is modeled using a 1D or 2D model. In the case of the 2D model, ν_V and ν_{CSR} approximately follow a one-to-one linear relationship, whereas the results of the 1D model are best represented by a quadratic formulation where the increase in ν_{CSR} is not as large as in the 2D model. In both relationships, it is assumed that when the shear wave velocity variance is zero, so also is the variance in CSR.

The results of the sensitivity analysis indicate that for sites with high variability in shear wave velocity (say, more than $\nu_V = 20\%$), completing a probabilistic ground response analysis may be worth the additional computational effort in order to reduce the estimated stresses during seismic shaking, as using the mean profile would result in estimates that are quite conservative. However, the high sensitivity of the results to the selection of ν_V also emphasizes the importance of collecting enough information at a site to be able to reliably estimate not only the mean, but also the variability of shear wave velocity at a site when probabilistic ground response analysis are to be completed.

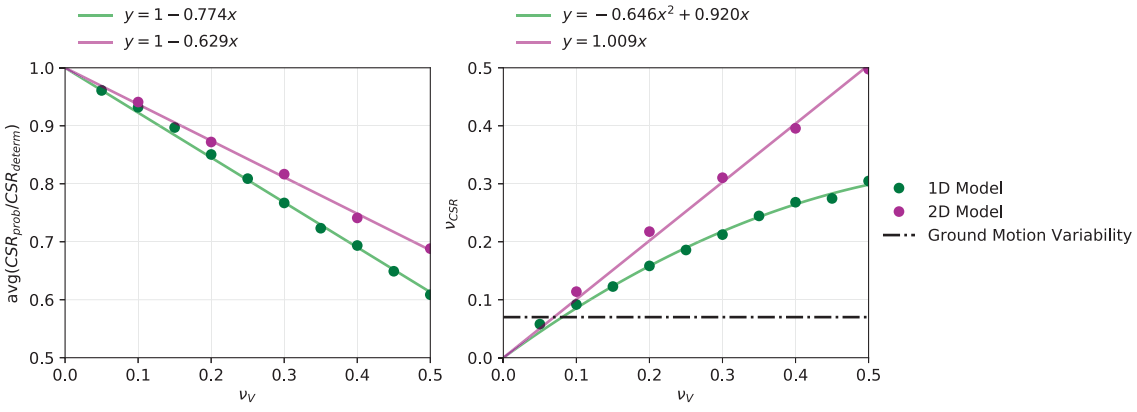

Figure 4.14: Sensitivity of CSR to ν_V.

4.5.3 Effects on PGA

Figure 4.15 plots the ratio of mean probabilistic PGA to the deterministic value, as well as the coefficient of variation of PGA, as a function of the coefficient of variation of shear wave velocity (ν_V). The results are largely consistent with the results of the CSR profile. However, the results here only consider the response at the ground surface, whereas the CSR results involved averaging over the depth profile.

At the values with lowest ν_V, an increase in mean PGA is observed when compared to the deterministic one since the ratio of mean probabilistic to deterministic PGA is larger than one. For the 1D model, this occurs when $\nu_V = 5\%$ and for the 2D model when $\nu_V = 10\%$. This indicates that using the mean shear wave velocity profile results in a sightly under-conservative prediction of peak ground acceleration at surface. As ν_V increases,the mean probabilistic PGA becomes much lower than the deterministic response. At the largest value of $\nu_V = 50\%$, the mean probabilistic PGA is 0.54 and 0.65 of the deterministic one for the 1D and 2D models, respectively. The relationship between the reduction in PGA and ν_V is most closely represented through the quadratic formulas shown in Figure 4.15.

The coefficient of variation of PGA (ν_{PGA}) also increases with larger values of ν_V. Whereas ν_{CSR} was significantly different for the 1D and 2D models, ν_{PGA} does not show such a strong dependency on the geometry. As shown in Figure 4.15, both the 1D and 2D model results can be reasonably approximated using a quadratic relationship. However, the estimated coefficients of variation are very similar when using either geometry.

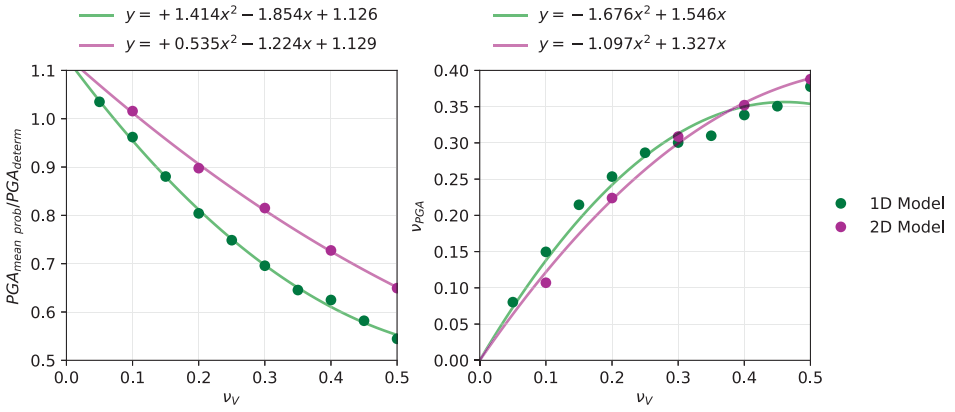

Figure 4.15: Sensitivity of PGA to ν_Y.

4.6 Sensitivity to Correlation Length

4.6.1 Description

Another parameter of the shear wave velocity random fields that may influence the results of the ground response analysis is the correlation length of $\ln Y$. In the base-case analysis, a correlation length of $\theta_{v,\,\ln Y} = 6$m was used in both the 1D and 2D models, with $\theta_{h,\,\ln Y} = 60$m in the 2D model. Here, a series of correlation lengths are considered in order to explore the effects on the CSR profile and surface PGA. A total of 15 and 9 vertical correlation lengths ranging between 1 m and 300 m are considered in the 1D and 2D models, respectively.

The correlation lengths considered in the 1D and 2D models are, respectively:

- $\theta_{v,\,\ln Y} = \{1,\ 1.5,\ 2,\ 3,\ 4,\ 5,\ 10,\ 15,\ 30,\ 40,\ 50,\ 75,\ 100,\ 300\}$
- $\theta_{v,\,\ln Y} = \{1,\ 2,\ 5,\ 10,\ 20,\ 30,\ 100,\ 200,\ 300\}$

In the 2D models, the correlation lengths are chosen such that the ratio between horizontal to vertical correlation is kept constant at 10, as was done in the base-case analyses. Additionally, although vertical correlation lengths less than 1 meter are not expected (since shear wave velocity is measured in 1 meter intervals), additional runs with $\theta_{v,\,\ln Y} = \{0.1,\ 0.25,\ 0.5\}$ were used for the 1D models to illustrate the effects of correlation lengths that are smaller than the finite element mesh.

Figure 4.16 illustrates sample realizations of the 2D shear wave velocity random fields, for four different choices of correlation length. As before, 1,000 realizations are generated for each set of random field properties, resulting in a total of 27,000 ground response analysis being completed.

Additionally, Figure 4.17 compares the mean and standard deviation of shear wave velocity from the ensemble of random field realizations against the input parameters. Whereas the mean of all realizations closely follows the input mean, the set of realizations with smaller correlation lengths (of roughly $\theta_{v,\,\ln Y} \leq 4$ m) result in standard deviations that fall below the input point standard deviation of the random process. The reduction in standard deviation is a result of local averaging that occurs within the finite element that is involved in the LAS procedure. As the correlation length approaches zero, the standard deviation also approaches zero and the value of the element becomes the median of the lognormal field.

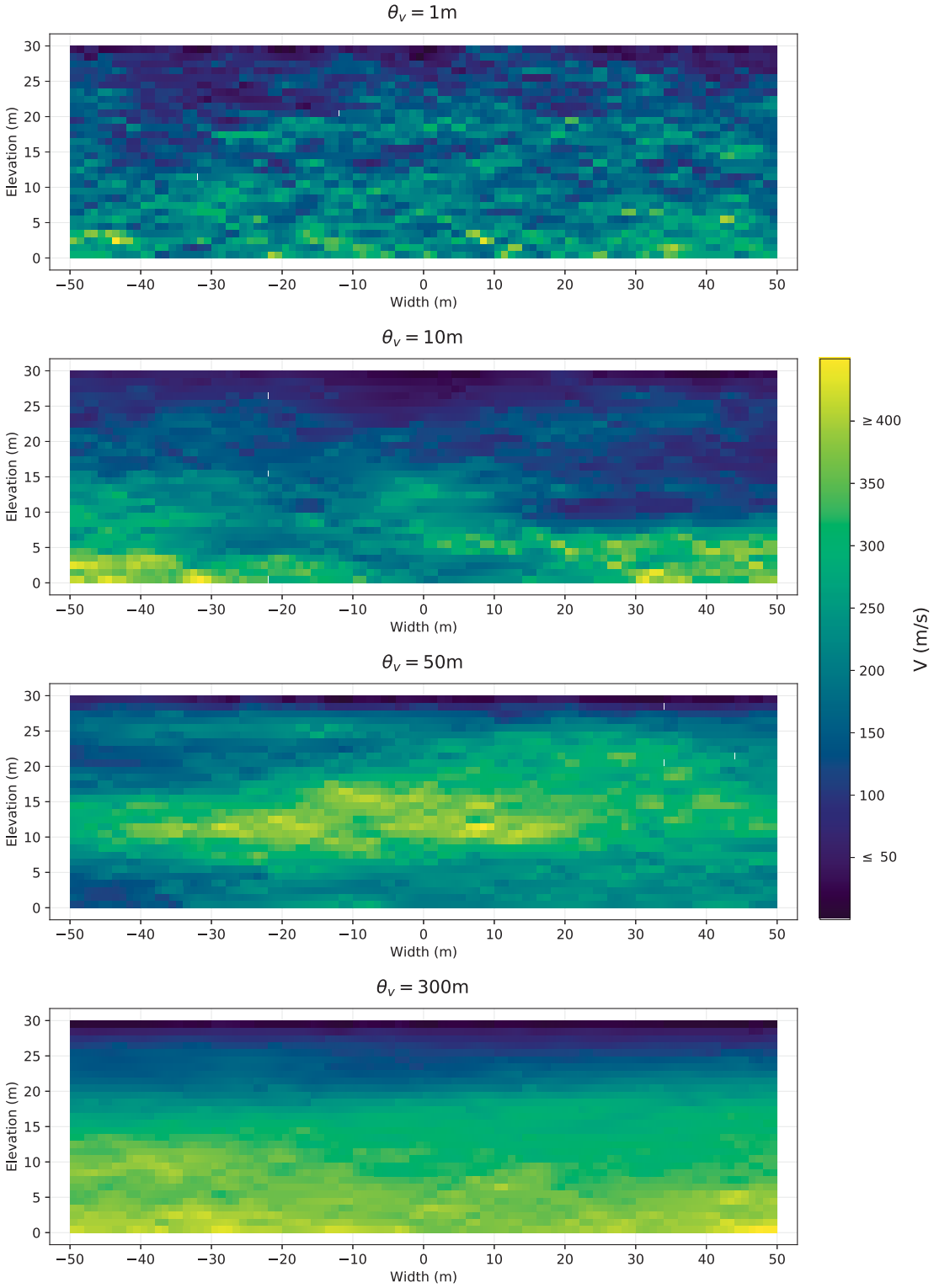

Figure 4.16: Sample realizations with varying correlation lengths.

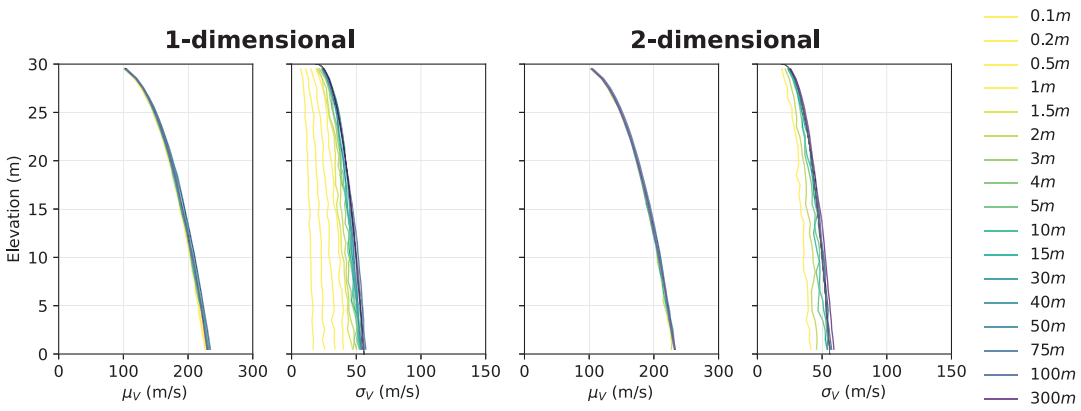

Figure 4.17: Simulated parameters of shear wave velocity for varying $\theta_{v,\ln Y}$.

4.6.2 Effects on CSR

Figure 4.18 illustrates the mean CSR profiles that are obtained from the ensemble of 1,000 realizations for each set of random field properties considered. The probabilistic analysis still result in a mean CSR that is less than the deterministic one, although the magnitude of the reduction is dependent on the choice of correlation length.

Figure 4.19 summarizes these effects of correlation length on CSR mean and variability by plotting the ratio of mean probabilistic CSR over the deterministic CSR (averaged over the elevation profile) as well as the coefficient of variation (ν_{CSR}), as a function of correlation length. Additionally, Figure 4.19 includes duplicate figures with different horizontal scales so that the effects at smaller correlation lengths can be clearly observed.

Generally, for correlation lengths that are roughly $\theta_{v,\ln Y} \geq 5$m (and therefore $\theta_{h,\ln Y} \geq 50$m), increasing correlation length results in a higher mean response which tends closer towards the deterministic one. Therefore, a larger choice of correlation lengths would generally yield more conservative results. The largest change is observed in correlation lengths between 5 m and 100 m, but tapers off for larger correlation lengths. There is little difference between correlation lengths larger than 100 m, likely because the correlation length is already vastly larger than the domain size of 30 m. At such large correlation lengths, the 1D and 2D results also become very similar, with the ratio of probabilistic to deterministic CSR reaching between 0.95 and 0.97, and ν_{CSR} reaching between 33% and 35%.

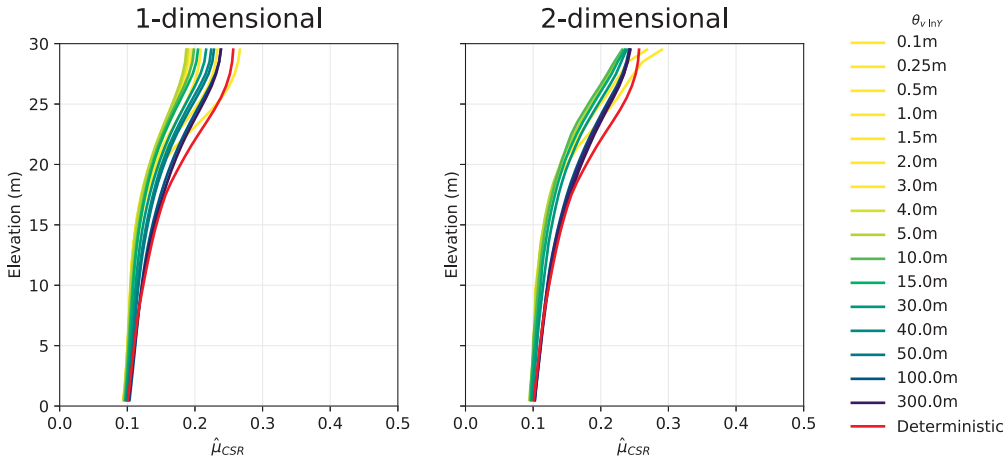

Figure 4.18: Sensitivity of mean CSR profile to correlation length.

For shorter correlation lengths ($\theta_{v,\,\ln Y} \leq 5$m and $\theta_{h,\,\ln Y} \leq 50$m) a reduction in correlation length also results in a higher mean CSR, as shown in the lower left plot of Figure 4.19. This is because as the correlation length approaches and falls below the size of the finite element mesh (1 m), averaging occurs within each finite element according to local average subdivision. This results in a lower variability of shear wave velocity, which in turns increases the mean response. Therefore, there exists a *best-case* correlation length which results in the lowest estimated stresses.

4.6.3 Effects on PGA

Figure 4.20 displays how the probabilistic PGA at the surface changes as a function of the correlation length. The results are plotted against vertical correlation length, with the horizontal length being 10 times larger. The mean PGA is summarized by plotting the ratio of mean probabilistic PGA over the deterministic PGA, as well as by showing the coefficient of variation of PGA.

The effects of correlation length on PGA are consistent with the results observed with CSR. For correlation lengths in the range $\theta_{v,\,\ln Y} \geq 5$m, increases in correlation length result in a higher mean PGA in both the 1D and 2D models. This is because longer correlation lengths result in less variability within each random field realization. However, increasing correlation length beyond 100m, corresponding to 3.3 times the domain size, results in little change in the mean probabilistic response. For small correlation lengths (roughly $\theta_{v,\,\ln Y} \leq 5$m), the mean response also increases, because

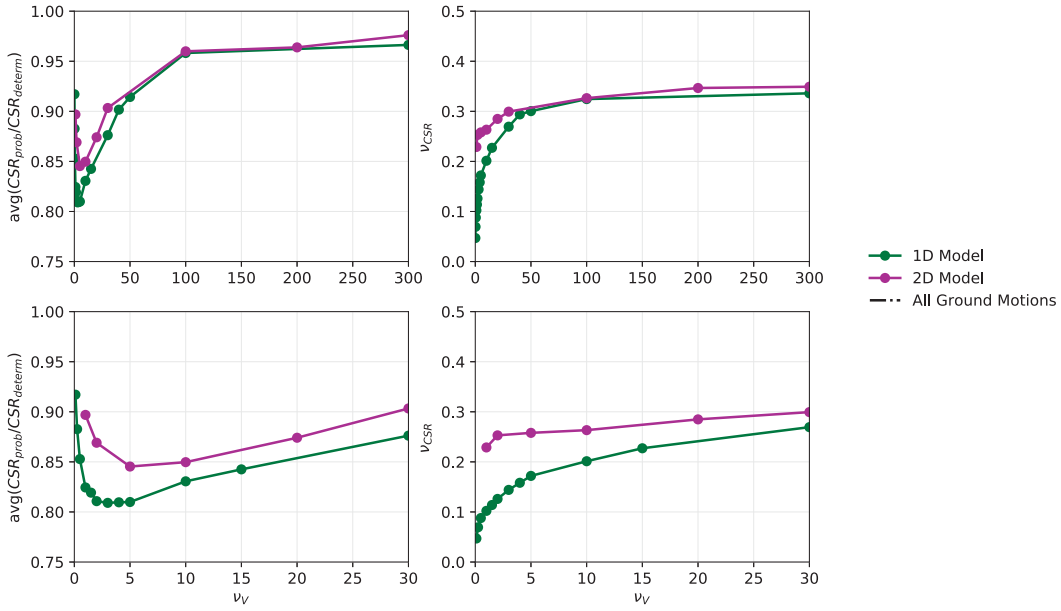

Figure 4.19: Sensitivity of CSR parameters to correlation length.

as the correlation length approaches the element size less variability is obtained across realizations. Therefore, there exists a correlation length that results in the smallest mean PGA in the probabilistic analysis.

As before, the 1D model results in a lower mean response when compared against the 2D model. However, this difference is largest around $\theta_{v,\,\ln Y} \approx 5$m, where the largest reduction in mean response is observed. As the correlation lengths become significantly smaller or larger, the difference between the 1D and 2D models is reduced. The coefficient of variation of PGA generally increases for larger correlation lengths, but plateaus for correlation lengths larger than about 50 m (corresponding to 1.6 times the domain size).

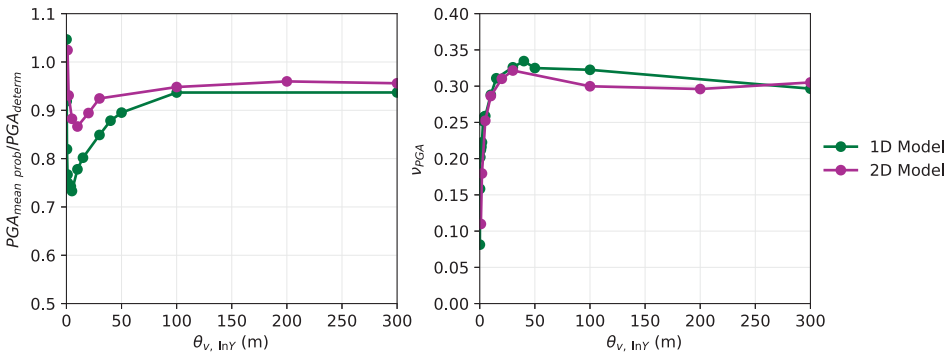

Figure 4.20: Sensitivity of probabilistic PGA to correlation length.

4.7 Sensitivity to Anisotropy

4.7.1 Description

The final parameter that is explored is the anisotropy ratio, defined as:

$$r = \frac{\theta_{h\ \ln Y}}{\theta_{v\ \ln Y}} \tag{4.11}$$

for the 2D model. Because the 1D model is equivalent to a 2-dimensional one where each soil column has the same properties, the 1D model is the same as the 2D model with an infinite horizontal correlation length, i.e., $r = \infty$. The values considered in this study are $r = \{1, 2, 3, 5, 7, 10, 15, 20\}$. Because of the depositional nature of soils, it is generally expected that the horizontal correlation length should be larger or equal to the vertical correlation length. For this reason, values of $r < 1$ are not considered realistic. The other parameters of the random field are kept the same as the base-case scenario, i.e., $\nu_V = 0.24$ and $\theta_{v\ \ln Y} = 6$m, and the value of $\theta_{h\ \ln Y}$ is selected in accordance with the desired anisotropy ratio.

Figure 4.22 displays sample realizations of the random fields for increasing values of r. An ensemble of 1,000 realizations is generated for each value of r considered, resulting in a total of 8,000 ground response analyses. Figure 4.21 compares the point mean and standard deviation used to generate the random field realizations, and compares it against the mean and standard deviation obtained from the ensemble of realizations.

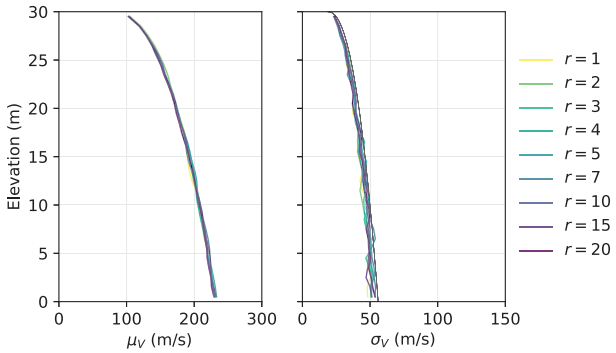

Figure 4.21: Simulated parameters of shear wave velocity for varying r values.

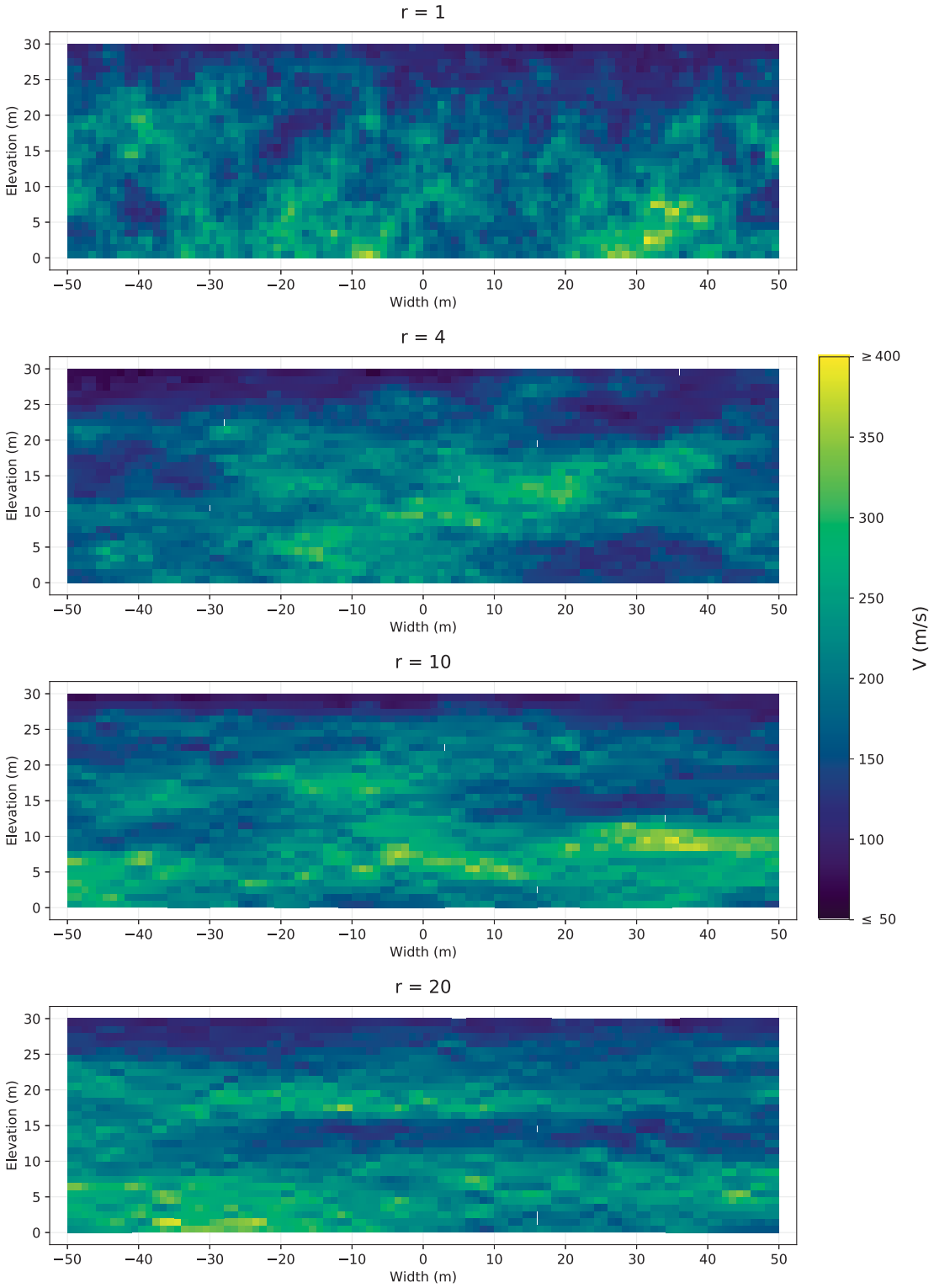

Figure 4.22: Sample realizations with varying values of r.

4.7.2 Effects on CSR

Figure 4.23 displays the mean CSR profiles that are obtained from the probabilistic ground response analysis for all values of r considered. Also included is the CSR profile that is obtained from the deterministic analysis (using the mean shear wave velocity), and the results from the base-case 1D model, which correspond to a value of $r = \infty$ such that the same soil column extends infinitely in the horizontal direction.

Increasing the anisotropy ratio r results in a lower mean CSR profile, which approaches the results of the 1D model, as expected. For this reason, the 1D model can represent a lower bound on the 2D model when considering the effects of anisotropy. The mean CSR profiles increase for smaller values of r, with the difference being smallest at the base of the model and increasing towards the ground surface.

Figure 4.24 displays the sensitivity of CSR to the anisotropy ratio. The figure plots the ratio of the mean CSR to the deterministic value as well as the coefficient of variation of CSR. Values in the range $r \leq 10$ show the steepest reduction in mean response, with the reduction tapering off for larger anisotropy where the response approaches that of the 1D models. The coefficient of variation of CSR is slightly sensitive to the choice of r, particularly for values of $r \leq 10$, with larger values of r leading to a lower coefficient of variation.

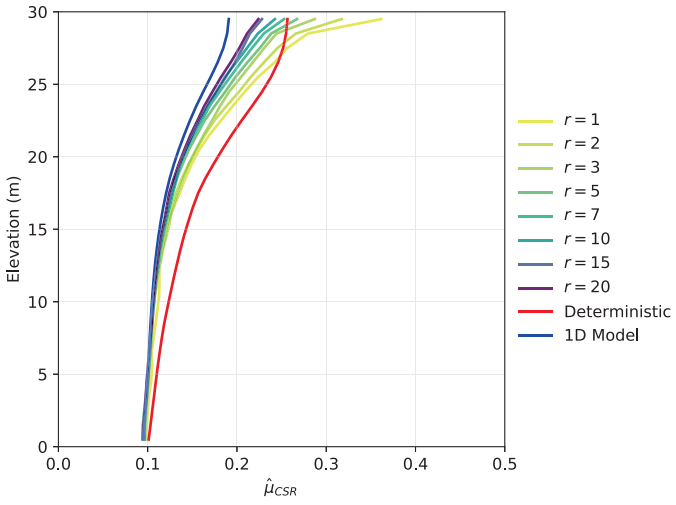

Figure 4.23: Mean CSR profiles for varying values of r.

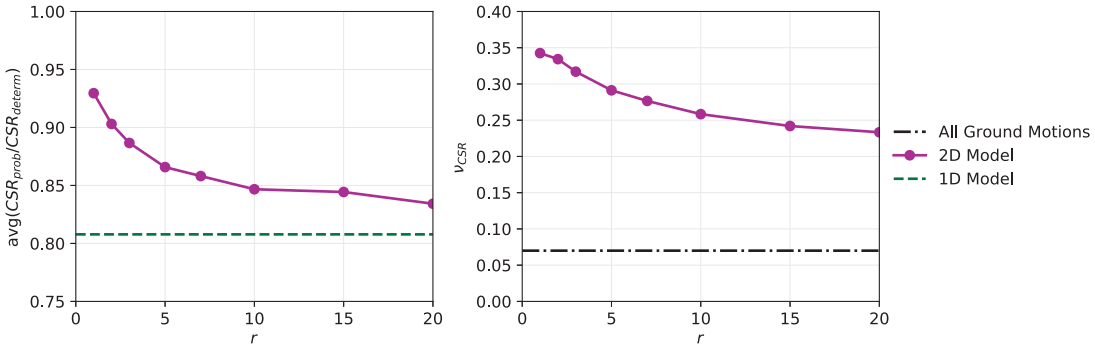

Figure 4.24: Sensitivity of probabilistic CSR to anisotropy ratio.

4.7.3 Effects on PGA

Finally, Figure 4.25 displays the effects of anisotropy on the surface PGA. Similar to the CSR results, increasing values of r result in a lower mean surface PGA, with the largest changes occurring in the range $r \leq 10$. A smaller rate of reduction is seen for larger values of r, with the results approaching the 1D model as r tends towards infinity.

Whereas larger values of r led to a reduction in coefficient of variation of CSR, the opposite is observed here for PGA. For values of $r \leq 10$, a significant increase in coefficient of variation of PGA is observed (right plot in Figure 4.25). However, the effects of anisotropy are relatively small. The other random field parameters considered in the sensitivity analysis, particularly the coefficient of variation of shear wave velocity, were found to have a much larger effect on the distribution of PGA at the surface.

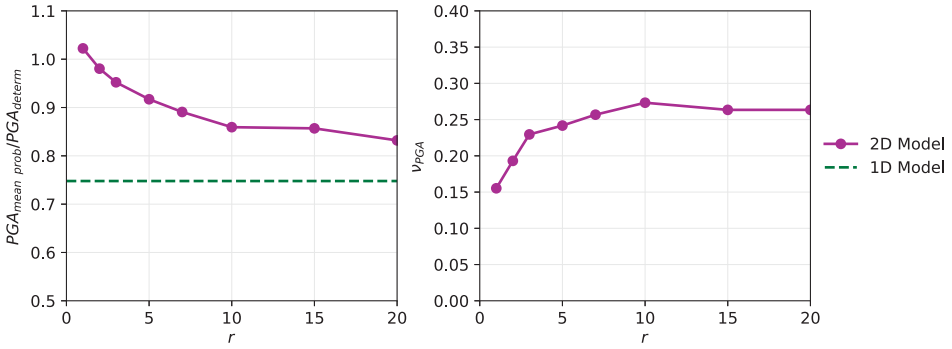

Figure 4.25: Sensitivity of probabilistic PGA to anisotropy ratio.

4.8 Summary

This chapter explored how modelling the spatial variability of shear wave velocity affects the results of 1D and 2D ground response analysis in terms of the cyclic stress ratio, peak ground acceleration at surface, and response spectra. This was accomplished by carrying out a series of equivalent-linear, probabilistic GRA, where spatially-varying shear wave velocity is represented using realizations of random fields with statistical properties based on Chapter 3 of this study.

Incorporating the spatial variability of shear wave velocity into GRA generally results in lower estimates of the mean stresses and accelerations when compared against the deterministic case, in which shear wave velocity is taken as a depth-dependent mean. This is true in both the 1D and 2D models for CSR profiles and PGA. In the case of the response spectra, a reduction in spectral accelerations are obtained below the natural period of vibration of the site; however, the mean probabilistic response is larger than the deterministic one for longer periods.

The variability of PGA, CSR, and response spectra when randomizing shear wave velocity are significantly larger than the uncertainty associated with running the suite of 10 ground motions with the mean shear wave velocity values. Tables 4.1 summarizes the uncertainty in estimates of CSR, PGA, and response spectra associated with the 10 ground motions, as well as when using a single motion with randomized shear wave velocity.

Parameter	Ground Motions	1D Randomized V	2D Randomized V
$\hat{\nu}_{CSR}$	0.07	0.17	0.26
$\hat{\nu}_{PGA}$	0.14	0.27	0.27
$\hat{\sigma}_{\ln SA}$	0.13	0.28	0.28

Table 4.1: Summary of uncertainty estimates.

However, it is important to note that this may be the result of using ground motions that are spectrally matched to the seismic hazard, as supposed to linearly matched. Similar 1D probabilistic GRA studies that use linearly scaled ground motions have found that ground motion variability contributes more to the uncertainty in the response spectra (for example Rathje et al. (2010) and Sun et al. (2020)). This may be due to the fact that this study uses spectrally matched ground motions,

whereas the aforementioned studies use scaled ground motions. Although spectrally matched motions can provide an excellent fit to target spectra, the resulting ground motions do not have a response spectrum that is representative of real, recorded earthquake motions.

When comparing the 1D and 2D probabilistic GRA, the 1D model consistently results in a lower mean CSR, PGA, and response spectra when compared against the 2D model. However, very similar estimates of the variability of CSR, PGA, and response spectra are obtained from the 1D and 2D models.

The sensitivity of the probabilistic GRA results to the statistical properties of the random fields was assessed by repeating the GRA for several values of coefficient of variation of shear wave velocity (ν_V), correlation length ($\theta_{\ln Y}$) and anisotropy ratio ($r = \theta_{h,\,\ln Y}/\theta_{v,\,\ln Y}$). The sensitivity results were presented for both CSR and PGA, in terms of the ratio of the mean probabilistic value to the deterministic one, as well as the coefficient of variation.

Of the random field properties considered, ν_V was found to have the largest effect on the estimated mean CSR profile and mean surface PGA, with the lowest and most variable results of PGA and CSR being obtained when maximizing ν_V. When varying correlation length, it was found that there exist an *optimum* correlation length that results in the smallest mean probabilistic PGA and CSR. As the correlation length approaches and become smaller than the element size, local averaging reduces the variance of shear wave velocity across realizations. On the other extreme, larger correlation lengths reduce the shear wave velocity variance within realizations. Both of these reductions in variance ultimately result in a response that is closer to the deterministic scenario. However, increasing the correlation length beyond ≈ 1.6 times the domain size had little effect on the results. Finally, the anisotropy ratio r was found to have the smallest effect on the probabilistic GRA results in terms of both CSR and PGA. Increasing the anisotropy ratio results in lower values of mean PGA and CSR, but higher coefficients of variation. Since the 1D scenario is equivalent to the 2D model with an infinite anisotropy ratio, the results can be used as a lower bound to the 2D model when examining the effects of anisotropy. The sensitivity of the probabilistic results to the random field properties indicates that careful selection of these parameters is necessary. If probabilistic analyses are to be completed at a site,

it is important to collect enough information to be able to estimate, at a minimum, the coefficient of variation of shear wave velocity at the site. Since estimating the correlation structure at any given site is more challenging, it may be preferable to consider a range of correlation lengths in the analysis instead of a single value.

Chapter 5

Conclusions

5.1 Summary

There were two main objectives in this study. The first, discussed in Chapter 3, was to quantify the spatial variability of shear wave velocity based on 206 seismic cone penetration tests available for sites in British Columbia, Canada. The analyses were completed at four separate sites, as well as for all the available data pooled together. The results of this chapter can be used to inform the selection of random field parameters when randomizing shear wave velocity, as well as to facilitate the estimation of correlation length when more shear wave velocity data is available.

The second objective, addressed in Chapter 4, was to explore how the spatial variability of shear wave velocity affects the response of a soil mass subject to seismic shaking. The results of this chapter can be used to compare how a deterministic approach to GRA compares to probabilistic ones, as well to provide some insight into which random field parameters most affect the results of probabilistic GRA.

The statistical properties of shear wave velocity were estimated by using a multiplicative form, where the random shear wave velocity ($V(d)$) was expressed as the product of a deterministic trend ($\hat{v}(d)$) and a lognormal random variable (Y). The following can be concluded from the statistical analyses:

1. Despite the SCPT being collected from sites that are large distances apart, the estimated trends ($\hat{v}(d)$) at each location were relatively similar, and consistent with previous studies in the area. All the sites showed a strong relationship between shear wave velocity and depth, and therefore de-trending the original measurements was considered appropriate.

2. A lognormal distribution was found to provide a good fit to the values of Y calculated from the SCPT data. Although the mean of Y was shown to be

constant with depth, some sites showed a slight decrease in standard deviation of Y in the top 5 to 10 meters.

3. Vertical correlation lengths of $\ln Y$ were calculated using the traditional *direct-fitting* approach. The correlation lengths ranged from 1.6 m to 3.0 m when using a Markov model and from 1.53 m to 3.14 m when using a Gaussian model. In general, the Markov model was found to provide a better fit to the sample correlation functions calculated from the SCPT data.

4. Using a *bias-matched* estimation approach allowed for a better fit to the sample correlation functions obtained from the SCPT data. The estimated correlation lengths from this process, ranging from 3.3 m to 6.0 m, are significantly larger than those obtained from the direct-fitting approach. The bias-matched approach can be expected to be a closer estimation of the *true* correlation structure at the sites considered.

5. Through a simulation-based approach, it was found that using less than 20 SCPT would have yielded a poor estimate of the vertical correlation length, and that the estimate greatly improves once 50 or more SCPT are available.

6. Finally, the approach was compared to the framework proposed by Toro (1995), which randomizes shear wave velocity using a first-order, auto-regressive process in which shear wave velocity is modeled in a series of correlated, lognormally distributed layers of varying thickness. Because SCPT are measured in consistent 1 meter intervals, only the shear wave velocity randomization model, and not the layer boundary process, was used in this study. Although Toro (1995) uses a depth-dependent inter-layer correlation coefficient (ρ_{IL}), a constant value of ρ_{IL} was assumed at each site, which ranged from 0.62 to 0.84.

7. Both the auto-regressive method described by Toro (1995) and the estimation of correlation length for the de-trended process are valid statistical models to model the spatial variability of shear wave velocity. Although Toro's method is more flexible in that it can capture possible changes in correlation structure with depth and layer thickness, it is difficult to obtain enough data at any given site to properly calibrate all the parameters in the model. Using correlation

lengths allows for a parameter that has been used extensively for other soil properties such as cone penetration resistance, and is arguably a more intuitive parameter.

In Chapter 4, a series of equivalent-linear ground response analysis (GRA) were completed using the finite element program QUAD4M, to explore how the spatial variability of shear wave velocity affects response of soil subject to seismic shaking.

Three sets of analyses were completed. First, a deterministic GRA was carried out, in which the shear wave velocity at each element was assumed to be the mean estimate at that depth. Then, the statistical properties from Chapter 3 were used to carry out a base-case probabilistic GRA, in which random shear wave velocity was modeled using an ensemble of 1,000 random field realizations. Finally, the sensitivity of the probabilistic GRA to the random field parameters was assessed. This was completed for the coefficient of variation of shear wave velocity (ν_V), correlation length ($\theta_{\ln Y}$), and the correlation anisotropy (r).

The following can be concluded from the probabilistic GRA analyses:

1. Randomizing shear wave velocity results in lower estimates of cyclic stress ratio (CSR) and peak ground acceleration (PGA) at the surface. Under the base-case scenario, the 2D models mean CSR was between 78% and 95% of the deterministic values and the mean PGA at the surface was only 85% of the deterministic value. Therefore, using the mean of shear wave velocity as the characteristic value in GRA generally results in estimates of PGA and CSR that are below the mean. This fact highlights the value of information that is collected during site investigations: completing thorough site investigations allows for reliable estimates of the variability of shear wave velocity, which can then be used in probabilistic GRA. These analyses can in turn reduce design costs by avoiding overconservative designs.

2. The difference between the spectral acceleration response spectra that is obtained from the probabilistic and deterministic analysis depend on the natural period of vibration. For periods below the natural period of vibration of the soil mass, the mean accelerations are significantly smaller than the deterministic ones, with the largest reduction of 50% occurring at a period of 0.38 s in

the 2D model. However, for periods larger than the natural period of vibration, the mean accelerations from the probabilistic model were found to be slightly higher than the deterministic response.

3. The coefficient of variation of shear wave velocity (ν_V) used in the probabilistic analysis was shown to have a significant effect on the distribution of CSR and PGA at the surface. Increasing variability resulted in lower mean CSR and PGA. At the highest variability considered ($\nu_V = 0.5$) in the 2D models, the mean CSR was, on average, 70% of the deterministic value and the mean PGA at the surface was 65% of the deterministic value. The reduction in mean PGA and CSR was more significant for the 1D models than the 2D models, resulting in a consistently lower mean response. Increasing variability also resulted in increases in the coefficient of variation of CSR and PGA in both the 1D and 2D models.

4. The choice of correlation length also affected the results of the probabilistic GRA. The largest reduction in mean CSR and PGA was observed for a vertical correlation lengths of roughly 5 m, corresponding to 0.16 times the domain size. Shorter correlation lengths result in local averaging within the finite elements, and longer correlation lengths result in less variability of shear wave velocity within each realization. Therefore, there exists an *optimum* correlation length that minimizes the mean response obtained from probabilistic GRA. In general, increasing the vertical correlation length beyond 1.6 times the domain size had little effect on the distribution of PGA and CSR.

5. Finally, higher choices of correlation anisotropy ratio (r) resulted in lower mean PGA and CSR in the probabilistic GRA. Since the 1D model is equivalent to the 2D one with an infinite anisotropy ratio, the 1D model can be used as a lower bound to the 2D model when examining the effects of anisotropy. However, the anisotropy ratio was found to have a small effect on results of GRA when compared against the other random field parameters, particularly the coefficient of variation of shear wave velocity.

5.2 Limitations and Future Work

Characterizing the spatial variability of soils and assessing ground response during seismic shaking are complex tasks that often require simplifying assumptions to successfully complete. Many such assumptions were adopted in the current study, which then lead to plenty of opportunities for improvement or extensions to this study.

On the spatial variability of shear wave velocity (Chapter 3), the following limitations and opportunities for future work are identified:

1. A particular challenge of modeling the spatial variability of ground properties is that an extensive number of tests is required to estimate the correlation structure. As discussed in Chapter 3, this is particularly true for seismic cone penetration tests due to the large sampling interval of 1 meter when compared to the cone penetration test. A limitation of this study is that the SCPT data considered reached a maximum depth of 30 meters, and that the SCPT were generally too far apart to be able to provide an estimate of the horizontal correlation structure of shear wave velocity.

   Opportunities for future work include repeating the estimation process for shear wave velocity data with larger depths, closer horizontal spacing, and in other regions. A significant effort has been completed in the United States to develop public-access, high-quality shear wave velocity databases for use by the earthquake geotechnical community (Ahdi et al., 2018). Similar efforts in curating public-access databases for Canadian sites would facilitate the transition towards probability-based methods of design.

2. A simplifying assumption in the statistical analysis in Chapter 3 is that the random variable Y, used to randomize shear wave velocity, has a constant mean and standard deviation. However, some of the SCPT sites considered showed a slight dependency between standard deviation of Y (σ_Y) and depth, particularly in the top 10 meters. Future work could involve collecting more SCPT data to determine whether a similar trend is consistently observed at other sites, and whether considering a non-stationary variance is necessary.

3. Finally, the SCPT data was analyzed by represented shear wave velocity as the

product of a deterministic trend ($v(\hat{d})$) and a random variable Y. As discussed in Chapter 3, de-trending should only be completed when there is a physical basis for the trend and when similar trends can be expected to occur at other target sites. Future work could explore randomizing the trend component instead of prescribing it deterministically, as discussed in Jiang and Huang (2018) or Ching and Phoon (2017).

Regarding the probabilistic ground response analysis (Chapter 4), the following limitations and opportunities for future work are identified:

1. There are many sources of uncertainty in equivalent-linear ground response analysis beyond the spatial variability of shear wave velocity, including the selection and scaling of input ground motion time histories and the modulus reduction and damping (MRD) curves. This study explored in detail the uncertainty in shear wave velocity; however, it is important to also explore the contribution from other sources of uncertainty. In particular, the ground motions considered in this study were spectrally matched to the hazard at a single site, resulting in uncertainties that are smaller than expected. It is generally recognized that selection of input ground motions is one of the major sources of uncertainty in ground response analyses (NCHRP, 2012). Future work could explore how uncertainties in a wider variety of ground motions affects the results of probabilistic GRA, as well as incorporate the uncertainties in MRD curves, which have been characterized by (Darendeli, 2001).

2. The general limitations of using an equivalent-linear approach to ground response analysis apply to the work completed in this study. The equivalent-linear approach approximates the true nonlinear behavior of soil by selecting single values of G and ζ that match the expected level of shear strain; however, the method is still linear and therefore carries some error. The linear-equivalent method is unable to capture the effects of excess pore water pressure, it may result in excessive levels of amplification when the input motions match the natural frequency of the soil mass, and the use of an effective shear strain may result in an over- or under-softened and damped response depending on the characteristics of the shear strain time history.

Future work could overcome the limitations of an equivalent-linear approach by adopting towards nonlinear total-stress or effective-stress methods, which use constitutive models to describe soil behavior during seismic shaking. This approach is significantly more complex than the equivalent-linear one, as the calibration of the constitutive model can require extensive laboratory tests and may introduce an additional source of uncertainty. However, nonlinear approaches have the significant advantage of being able to predict displacement associated with seismic shaking.

Bibliography

Abramowitz, M. and Stegun, I. A. (1970). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, volume 55. US Government printing office.

Ahdi, S. K., Sadiq, S., Ilhan, O., Bozorgnia, Y., Hashash, Y. M. A., Kwak, D. Y., Park, D., Yong, A., and Stewart, J. P. (2018). Development of a United States Community Shear Wave Velocity Profile Database. In *Geotechnical Earthquake Engineering and Soil Dynamics V*, pages 330–339, Austin, Texas. American Society of Civil Engineers.

Baecher, G. and Christian, J. (2003). *Reliability and Statistics in Geotechnical Engineering*. John Wiley & Sons, Inc., first edition.

Cami, B., Javankhoshdel, S., Phoon, K.-K., and Ching, J. (2020). Scale of Fluctuation for Spatially Varying Soils: Estimation Methods and Values. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 6(4):03120002.

Cascante, G. and Santamarina, J. C. (1996). Interparticle contact behavior and wave propagation. *Journal of Geotechnical Engineering*.

Ching, J. and Phoon, K.-K. (2017). Characterizing Uncertain Site-Specific Trend Function by Sparse Bayesian Learning. *Journal of Engineering Mechanics*, 143(7):04017028.

Ching, J., Wu, S.-S., and Phoon, K.-K. (2015). Statistical characterization of random field parameters using frequentist and Bayesian approaches. *Canadian Geotechnical Journal*, 53(2):285–298.

Christodoulou, P., Pantelidis, L., and Gravanis, E. (2020). The Effect of Targeted Field Investigation on the Reliability of Axially Loaded Piles: A Random Field Approach. *Geosciences*, 10(5):160.

Darendeli, M. B. (2001). Development of a New Family of Normalized Modulus Reduction and Material Damping Curves. page 393.

Fenton, G. and Griffiths, D. (2007). Random field generation and the local average subdivision method. In *Probabilistic Methods in Geotechnical Engineering*, volume 491, pages 201–223.

Fenton, G. A. (1990). *Simulation and Analysis of Random Fields*. PhD thesis, Princeton University.

Fenton, G. A. (1999a). Estimation for stochastic soil models. *Journal of Geotechnical and Geoenvironmental Engineering*, 125(6):470–485.

Fenton, G. A. (1999b). Random field modeling of CPT data. *Journal of Geotechnical and Geoenvironmental Engineering*, 125(6):486–498.

Fenton, G. A., Griffiths, D., and Urquhart, A. (2003). A slope stability model for spatially random soils. In *Proc. 9th Int. Conf. Applications of Statistics and Probability in Civil Engineering (ICASP9), a. Kiureghian et al. Eds Millpress, San Francisco, CA*, pages 1263–1269.

Fenton, G. A. and Griffiths, D. V. (2003). Bearing-capacity prediction of spatially random c - φ soils. 40:12.

Fenton, G. A. and Griffiths, D. V. (2008). *Risk Assessment in Geotechnical Engineering*. John Wiley & Sons, Inc., Hoboken, NJ, USA.

Fenton, G. A., Griffiths, D. V., and Zhang, X. (2008). Load and resistance factor design of shallow foundations against bearing failure. *Canadian Geotechnical Journal*, 45(11):1556–1571.

Fenton, G. A. and Vanmarcke, E. H. (1990). Simulation of Random Fields via Local Average Subdivision. *Journal of Engineering Mechanics*, 116(8):1733–1749.

Fraser Basin Council (2016). Lower Mainland flood management strategy phase 1 summary. Technical report, Fraser Basin Council.

Griffiths, D., Fenton, G. A., and Ziemann, H. R. (2008). Reliability of passive earth pressure. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 2(2):113–121.

Griffiths, D. V. and Fenton, G. A. (2009). Probabilistic Settlement Analysis by Stochastic and Random Finite-Element Methods. *Journal of Geotechnical and Geoenvironmental Engineering*, 135(11):1629–1637.

Hudson, M., Idriss, I., and Beikae, M. (1994). *User's Manual for QUAD4M*. National Science Foundation, Davis, California.

Hunter, J. A., Harold, C., Harris, J., Britton, J., and Lutternauer, J. (1999). Mapping Shear Wave Velocity Structure beneath the Fraser River Delta Sediments - Preliminary Results. In *8th Canadian Conference on Earthquake Engineering*.

Idriss, I. and Boulanger, R. (2008). *Soil Liquefaction during Earthquakes*. Engineering Monographs on Earthquake Criteria, Structural Design, and Strong Motion Records. Earthquake Engineering Research Institute.

Idriss, I. M., Lysmer, J., Hwang, R., and Bolton Seed, H. (1973). QUAD-4 A Computer Program for Evaluating the Seismic Response of Soil Structures by Variable Damping Finite Element Procedures. Technical Report EERC 73-16, National Science Foundation.

Jiang, S.-H. and Huang, J. (2018). Modeling of non-stationary random field of undrained shear strength of soil for slope reliability analysis. *Soils and Foundations*, 58(1):185–198.

Kramer, S. L. (1996). *Geotechnical Earthquake Engineering*. Prentice-Hall.

Lloret-Cabot, M., Fenton, G. A., and Hicks, M. A. (2014). On the estimation of scale of fluctuation in geostatistics. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 8(2):129–140.

Lunne, T., Powell, J. J., and Robertson, P. (2002). *Cone Penetration Testing in Geotechnical Practice*. CRC Press.

McKenna, F. (2011). OpenSees: A framework for earthquake engineering simulation. *Computing in Science and Engg.*, 13(4):58–66.

Moitzi, M. (2021). Ezdxf Documentation. page 524.

NCHRP (2012). *Practices and Procedures for Site-Specific Evaluations of Earthquake Ground Motions*. Transportation Research Board, Washington, D.C.

Northwest Hydraulic Consultants (2016). Regional assessment of flood vulnerability. Technical Report 3000149, Fraser Basin Council.

Northwest Hydraulic Consultants and Thurber Engineering Ltd. (2015). Lower Mainland dike assessment. Technical Report GS15LMN-054, Fraser Basin Council.

Peterson, P. (2009). F2PY: A tool for connecting Fortran and Python programs. *International Journal of Computational Science and Engineering*, 4(4):296.

Phoon, K.-K. and Kulhawy, F. H. (1999). Characterization of geotechnical variability. *Canadian geotechnical journal*, 36:28.

Rathje, E. M., Kottke, A. R., and Trent, W. L. (2010). Influence of Input Motion and Site Property Variabilities on Seismic Site Response Analysis. *Journal of Geotechnical and Geoenvironmental Engineering*, 136(4):607–619.

Stewart, J. P., Afshari, K., and Hashash, Y. M. A. (2014). Guidelines for Performing Hazard-Consistent One-Dimensional Ground Response Analysis for Ground Motion Prediction. page 152.

Sun, Q., Guo, X., and Dias, D. (2020). Evaluation of the seismic site response in randomized velocity profiles using a statistical model with Monte Carlo simulations. *Computers and Geotechnics*, 120:103442.

Toro, G. (1995). Probabilistic models of site velocity profiles for generic and site-specific ground-motion amplification studies. *Technical Rep*, 779574.

Vanmarcke, E. (1998). *Random Fields: Analysis and Synthesis*. MIT Press.