8-2018

A Geostatistical Approach toward Shear Wave Velocity Modeling and Uncertainty Quantification in Seismic Hazard

Wenxin Liu
Clemson University, wenxinl@clemson.edu

Follow this and additional works at: https://tigerprints.clemson.edu/all_dissertations

Recommended Citation

Liu, Wenxin, "A Geostatistical Approach toward Shear Wave Velocity Modeling and Uncertainty Quantification in Seismic Hazard" (2018). *All Dissertations*. 2206.
https://tigerprints.clemson.edu/all_dissertations/2206

A GEOSTATISTICAL APPROACH TOWARD SHEAR WAVE VELOCITY
MODELING AND UNCERTAINTY QUANTIFICATION IN
SEISMIC HAZARD ANALYSIS

A Dissertation
Presented to
the Graduate School of
Clemson University

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

---

by
Wenxin Liu
August 2018

Accepted by:
Dr. Qiushi Chen, Committee Chair
Dr. C. Hsein Juang, Co-Chair
Dr. Ronald Andrus
Dr. Jie Zhang

ABSTRACT

The ground motion parameters such as amplitude, frequency content and the duration can be affected by the local site condition and may result in amplification or de-amplification to the original bedrock motion. Shear wave velocity is an important site parameter to describe the site condition and is widely used in estimating site response, classifying sites in recent building codes and loss estimation. This dissertation is aimed at: modeling the spatial variation of shear wave velocity using geostatistical tools; improve the random field framework for estimating soil properties to account for multiple sources of data; develop finite element model to qualify the uncertainty propagation in dynamic site response; introduce the response surface concept into seismic hazard analysis and quantifying the uncertainty propagation in dynamic site response caused by the variation of shear wave velocity and design parameters.

To model the spatial variation of shear wave velocity, a multiscale random field-based framework is presented and applied to mapping V_{s30} - the time-averaged shear wave velocity in the top 30 meters of subsurface material - over extended areas. In this framework, the random field concept is employed to model the horizontal variation of shear wave velocity. Suzhou Site is selected as research area and its measured shear wave velocity data is combined with U.S. Geological Survey (USGS) slope-based V_{s30} map for mapping the V_{s30} around whole research area. Moreover, a different method of integrating multiple sources of data is used and tested based on a synthetic digital field.

To quantify the uncertainty propagation in dynamic site response caused by the variation of shear wave velocity, the finite element method (FEM) is developed and combined with random field realizations of shear wave velocity profiles. A viscoelastic constitutive model is implemented in the FEM model to account for the non-linear hysteresis response of subsurface materials under cyclic loadings. The analyzed site responses as well as the input parameters generated with Monte Carlo simulations (MCS) are then used to study the peak acceleration at site surface subjected to a given input seismic wave. Finally, the response surface method and the first order second moment method (FOSM) are integrated into dynamic site response analysis to characterize variation of site performance caused by spatial variation of shear wave velocity. Through illustrative examples, the effectiveness, advantage, practicability and significance of improved random field framework and developed uncertainty propagation evaluation methodology are demonstrated.

DEDICATION

I dedicate this dissertation to my parents for their love and support all these years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

1. INTRODUCTION

1.1 Motivation and Background

The 1933 Long Beach, 1957 San Francisco, 1967 Caracas, 1985 Mexico City, 1989 Loma Prieta, and 1994 Northridge earthquake events left evidence of how the local site condition can affect the characteristics of propagating earthquake wave from the bedrock through the topsoil (Bhuiyan 2015). The ground motion amplitude, frequency content and the duration can be affected by the local site condition and thus can cause significant amplification or de-amplification to the original bedrock motion which can seriously affect the structures.

For describing local site conditions, shear wave velocity (V_s) is a simple, effective and representative parameter. Also, it is an important input for any seismic site response study. Normally, there are two expression forms of shear wave velocity: the V_s profile along the depth, and the time-averaged shear wave velocity in the top 30 meters of the subsurface material, denoted as V_{s30}. As a simplified expression of shear wave velocity, V_{s30} integrates a whole V_s profile into one specific value. Owing to its convenience in engineering usage, National Earthquake Hazards Reduction Program (NEHRP) uses V_{s30} value to classify the soil and describe soil stiffness. Lots of ground-motion prediction equations also use it as a key indicator such as Abrahamson and Silva (2008),

Abrahamson et al. (2014), Boore and Aktinson (2008), Boore et al. (2014), Campbell and Bozorgnia (2008, 2014), Choi and Youngs (2008, 2014), and Idriss (2008, 2014). Because of its importance and effectiveness as a site parameter for site response prediction, the NGA-West2 project (Ancheta et al., 2014; Seyhan et al., 2014) made a project-level decision to compile a site database in terms of V_{s30}. The U.S. Geological Survey (USGS) earthquake hazard program also provides and maintains a global V_{s30} map server. While the V_{s30} can be computed directly given a shear-wave velocity measurement, such geophysical measurements are typically very sparse. Therefore, various descriptors or quantitative metrics of site condition have been proposed for the purpose of estimating V_{s30} in the absence of geophysical measurements.

In the past, proxy-based methods have been developed to estimate V_{s30}, including the geology-based (Wills et al. 2015), topography-based (Wald and Allen 2007) or hybrid method (Thompson et al. 2014). One limitation is that, while initially derived from observed V_{s30} values, these approaches fail to directly incorporate the V_{s30} measurements used back into the map that has been created.

In this dissertation, a geostatistical approach that accounts for the spatial variability of V_{s30} across different length scales and incorporates the compiled database of direct geophysical measurements and proxy-based V_{s30} values is presented.

Regional estimation of soil properties necessitates not only the geostatistical model considering spatial variance of soil properties, but also the means to account for heterogeneous sources of information. It is necessary to account for spatial variability of soil properties across scales consistently, while simultaneously preserving constraints

2

imposed by geological boundaries (Wang and Chen 2017). Thus, a hybrid geotechnical and geological data-based framework is presented and tested in this dissertation. Recommendations on how to act in each specific mapping situation and step-by-step instructions for accomplishing mapping are provided including the methodology to calibrate Markov Bayes coefficient B and integrate secondary data.

Coupling the random field model with Monte Carlo simulations, the expected V_{s30} values across the specific site and the associated uncertainties can be obtained. It is found that the uncertainty of shear wave velocity is ubiquitous, in both horizontal and vertical direction. With the understanding of the distribution of V_{s30} and the importance of its uncertainty, the uncertainty propagation in dynamic site response is evaluated.

Dynamic analysis of site effect on ground motions is a challenging task, coupled with the uncertainty of soil parameter, making it even more difficult to quantify. In the past few decades, lots of research has been done to estimate site response. Some of them use empirical models which are based on statistical analysis and fitting of field data like Abrahamson and Silva (2008), Boore and Atkinson (2008), Campbell and Bozorgnia (2008), Chiou and Youngs (2008) and Idriss (2008). Others use numerical methods which are based on dynamic site response analysis like Seed and Idriss (1969), Park and Hashash (2008), Kamalian et. al (2006), Martin et al. (1982) and Elgamal et al. (2002). Few researchers consider the uncertainty of input parameters. Typical works include Wang and Hao (2002), Bahrampouri et al. (2018), Stewart and Kwok (2008) and Tombari and Stefanini (2017).

In this dissertation, a comprehensive study is conducted to evaluate the impacts of soil parameters uncertainty in nonlinear dynamic site response. Numerical analysis with finite element method is applied to describe the uncertainty propagation in dynamic site response and its performance under different site condition. Modified Davidenkov model with simplified loading-reloading rules is used to describe the stress-strain relationship under irregular cyclic loading. Furthermore, in the following chapters, the response surface and first order second moment (FOSM) concept are integrated to quantify the uncertainty in dynamic site response and a response surface model is established for the further reliability analysis.
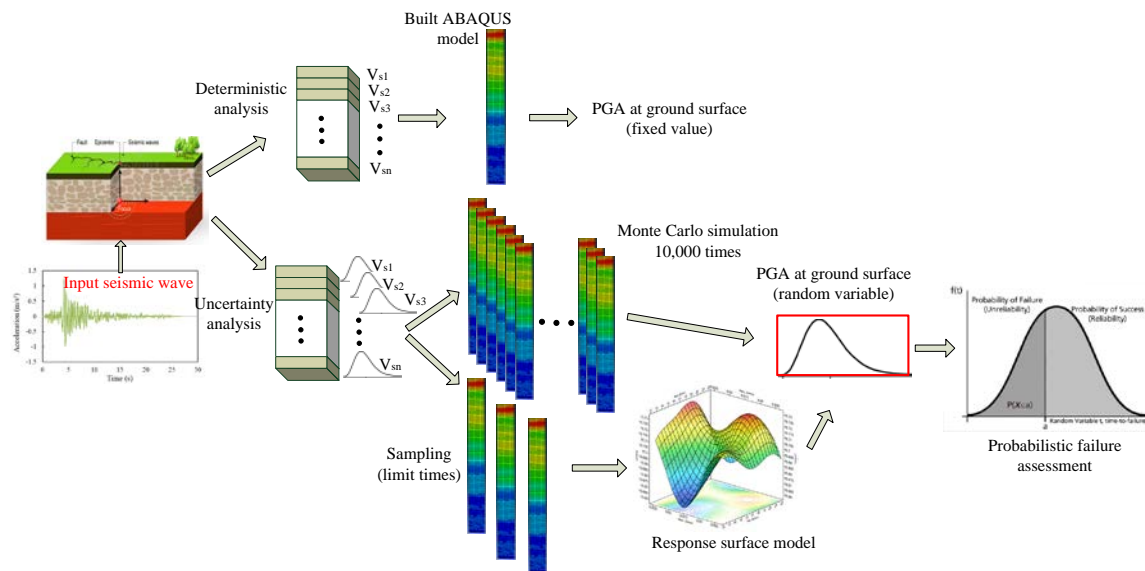
Figure 1.1 Flow chart detailing the uncertainty analysis in the dynamic site response

1.2 Objective and Dissertation Organization

The objectives of this research are to (1) investigative the distribution and uncertainty of regional soil properties (2) improve the random field framework for estimating soil properties to account for multiple sources of data, (3) develop finite element model to qualify the uncertainty propagation in dynamic site response, (4) introduce the response surface concept into seismic hazard analysis, (5) formulate the uncertainty propagation in site response that can explicitly consider longitudinal variation of input soil parameters and design parameters.

This dissertation consists of five chapters. In Chapter II, a multiscale random field-based framework is presented and applied to map V_{s30} over extended areas. Here, the framework accounts for spatial variations of V_{s30} values across different length scales and is able to adaptively refine around areas of high interest while maintaining a consistent description of spatial dependence. In Chapter III, a hybrid geotechnical and geological data-based random field framework is presented and tested, which is an updated model based on the framework introduced in Chapter II. Here, the effects of site investigation plans, the Markov-Bayes coefficient and the element size for the predefined grid of secondary data on the random field-based mapping of soil properties have been tested. Based on the sensitivity analysis, an overall workflow for integrating multiple sources of data in the random field model for regional soil properties mapping is established. In Chapter IV, a finite element model is developed for the dynamic site response analysis. In this model, the modified Davidenkov model with simplified

loading-reloading rules is compiled and applied to simulate the stress-strain relationship of soil. Monte Carlo simulations are conducted to qualify the uncertainty propagation in dynamic site response. In Chapter V, the response surface and first order second moment (FOSM) concepts are applied based on the deterministic numerical solutions (developed in Chapter IV) for building computationally efficient models for complex geotechnical reliability problems. Finally, in Chapter VI, the main conclusions and recommendations of this dissertation are presented.

CHAPTER II

2. SPATIALLY CORRELATED MULTISCALE V_{S30} MAPPING AND A

CASE STUDY OF THE SUZHOU SITE*

2.1 Introduction

The time average shear-wave velocity in the first 30 m of subsoil, denoted as V_{s30},

is an important site parameter used in estimating site response, classifying sites in recent

building codes and loss estimation (Boore, 2004). Because of its importance and

effectiveness as a site parameter for site response prediction, the NGA-West2 project

(Ancheta et al., 2014; Seyhan et al., 2014) made a project-level decision to compile a site

database in terms of V_{s30}. The U.S. Geological Survey (USGS) earthquake hazard

program also provides and maintains a global V_{s30} map server. Site databases in terms of

V_{s30} give useful site information that allows engineers to choose appropriate site

conditions for various design and analysis purposes.

While the V_{s30} can be computed directly given a shear-wave velocity

measurement, such geophysical measurements are typically very sparse. Therefore,

various descriptors or quantitative metrics of site condition have been proposed for the

purpose of estimating V_{s30} in the absence of geophysical measurements. For instance, Wald and Allen (2007) proposed a technique to derive first-order site-condition maps directly from topographic data, where the V_{s30} values are correlated with the topographic slope. Wills and Clahan (2006) and Wills and Gutierrez (2008) grouped shear-wave velocity data by corresponding geologic units to determine the shear-wave velocity characteristics of each geologic unit. Then, the geologic unit designation and shear-wave velocity characteristics are applied to sites without shear-wave velocity data. This revised geologic designation improves the previous geology-based V_{s30} method by Wills et al. (2000) and Wills and Silva (1998). In addition, geology-topography hybrid (Scasserra et al., 2009) and geomorphometry-based proxy relationships (Yong et al., 2012) have been proposed for estimating V_{s30}.

A major limitation of proxy-based methods is that, while initially derived from or constrained by observed V_{s30} values, these methods do not directly incorporate the V_{s30} measurements into the generated site condition map. This, along with the increasing amount of available direct geophysical measurement data, motivates the application of geostatistical methods to V_{s30} and site condition mapping. Examples of recent work along this line include the work of Thompson and his coworkers (Thompson et al., 2014, 2011, 2010), where a new map of V_{s30} for California is developed accounting for geology, topography and most importantly, site-specific V_{s30} measurements. The geostatistical approach of regression kriging (RK) is applied to combine these constraints to predict V_{s30}. This approach allows the resulting V_{s30} map to be locally refined to reflect the rapidly expanding database of V_{s30} measurements. Yong et al. (2013) and Wald et al.

(2011) applied the kriging-with-a-trend method to mapping V_{s30}, where the baseline model was derived from topographic slope. Also, Lee and Tsai (2008) established the spatial relationship between the shear-wave velocity (V_s) and the N value of the standard penetration test (SPT-N) and adopted the kriging with varying local means to update the V_{s30} maps in Taiwan. Thompson et al. (2007) modeled the horizontal variability of near-surface soil shear wave velocity in the San Francisco Bay Area using geostatistical methods.

In this chapter, a multiscale random field-based approach is presented and applied to mapping V_{s30} over an extended region. Unlike existing geostatistical methods for V_{s30} mapping, the presented approach explicitly accounts for the spatial variability of V_{s30} across different length scales and incorporates the compiled database of direct geophysical measurements and proxy-based V_{s30} values. High resolution predictions of V_{s30} can be obtained by adaptively refining coarse-scale values into finer scales in areas where deemed necessary while retaining appropriate spatial correlation, which is a particular useful feature for analyzing fine scale quantities of interest, such as estimation of uncertainties. Coupled with Monte Carlo simulations, the multiscale random field models also allow the quantification of uncertainties in the V_{s30} maps. The resulting V_{s30} maps preserve known V_{s30} data, uphold appropriate spatial correlation and have multiscale resolutions with information on associated uncertainties.

The order of presentation of this chapter goes as follows: Section 2.2 summarizes the engineering geology, field data and secondary V_{s30} data of the Suzhou site; In Section 2.3, key components of the developed geostatistical tools for mapping V_{s30} are presented;

Statistical and spatial characterizations of the known V_{s30} data will be discussed in detail in Section 2.4; In Section 2.5, new V_{s30} maps will be represented and applications of those new V_{s30} maps will be discussed in Section 2.6.

2.2 The Suzhou site: engineering geology and field data

Suzhou is a populous city on the alluvial plain of the Yangtze River Delta in the southeast of Jiangsu Province, China. In this section, the engineering geology and field data of the Suzhou site are briefly summarized. The dominating alluvial deposits beneath the studied site are soft and sensitive. In addition to geotechnical engineering challenges associated with construction on soft soil, long-period ground motions of far earthquakes may also cause serious damage to engineering projects in this area (Zhan et al., 2009).

2.2.1 Engineering geology

The studied area of Suzhou City is covered by Quaternary deposits of fluvial, lake, lagoon and marine origins. Most of the area is a combination of a lacustrine plain and delta plain. Some layers of the lake and river deposits are rich in over-consolidated clay. Most of the lagoonal and marine deposits, however, consist of soft clays, which are dark in color and rich in organic matters.
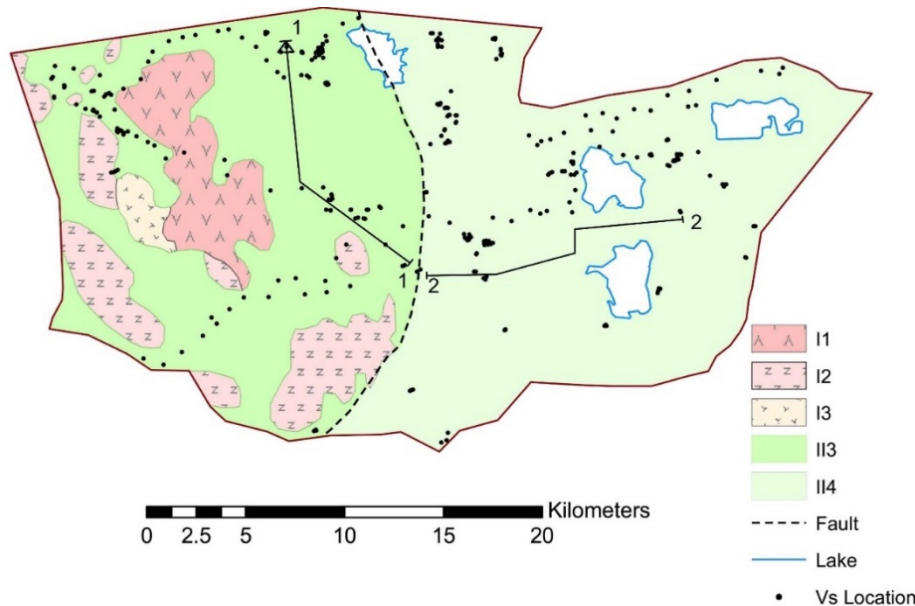
Figure 2.1 Surficial geology map of the Suzhou site and locations of shear-wave velocity measurements (black dots in the figure). II_3 is the Taihu alluvial plain; II_4 is the lake-swamp plain; I_1, I_2 and I_3 are outcrops with different rock types. Cross sections 1-1 and 2-2 are used to plot example soil profiles for the top 50 m. The little triangle shows the location of the sample V_s profile in Figure 2.3.

Figure 2.1 shows the boundaries of the studied area, the major surficial geology units and locations of shear-wave velocity measurements. As shown in Figure 2.1, the western portion of the studied area belongs to the Taihu alluvial plain (II_3) with interspersed outcrops (I_1, I_2 and I_3). The eastern portion belongs to the lake-swamp plain (II_4). Almost all of the shear- wave velocity measurements were taken in the geological units II_3 and II_4. Example profiles of the top 50 m soil are plotted in Figure 2.2(a) for the Taihu alluvial plain (II_3) (cross-section 1-1 in Figure 2.1) and in Figure 2.2(b) for the lake-swamp plain (II_4) (cross-section 2-2 in Figure 2.1), respectively. Explanations of the soil type number are summarized in Table 2.1.

11
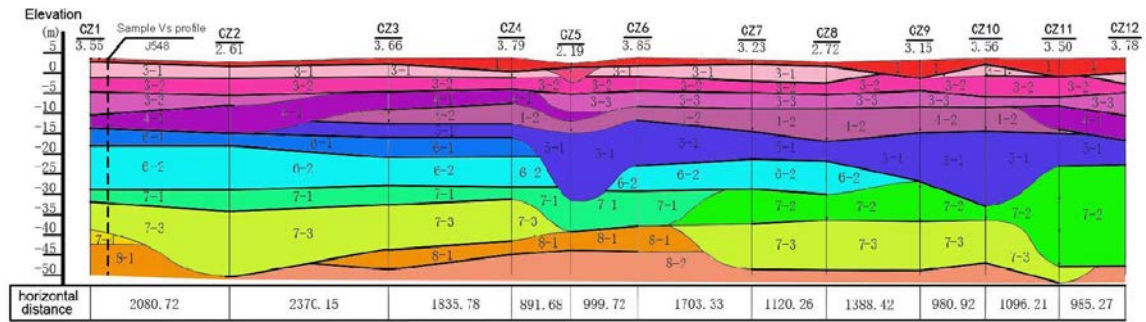
(a) Cross-section 1-1
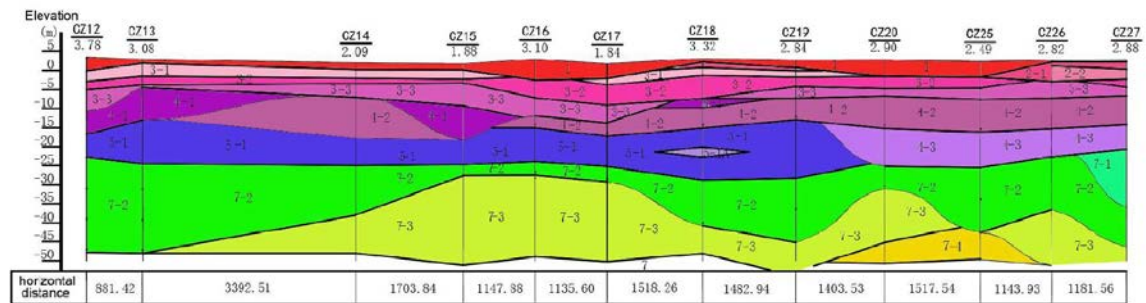
(b) Cross-section 2-2

Figure 2.2 Example soil profiles in the top 50 m for the cross-sections 1-1 and 2-2 shown in Figure 2.1.

Table 2.1 Explanation of soil type numbers used in Figure 2.2.

Number	Soil type	Property	Number	Soil type	Property
1	ground fill	loose	5–1A	clay	medium dense
2–1	silt clay	plastic	5–2	fine	dense
2–2	silt clay	soft plastic	6–1	clay	hard plastic
3–1	clay	hard plastic	6–2	silt clay	plastic
3–2	silt clay	plastic	7–1	silt clay	soft plastic
3–3	silt	medium dense	7–2	silt with fine	dense
4–1	silt clay	flow plastic	7–3	silt clay	soft plastic
4–2	sandy clay	medium dense	7–4	silt	dense
4–3	silt	dense	8–1	silt clay	hard plastic
5–1	silt with clay	flow plastic	8–2	silt clay	plastic

12

2.2.2 Field data

The field data compiled for this study consists of shear-wave velocity measurement data and soil parameters from lab tests from Institute of Earthquake Engineering for Jiangsu Province, China, performed 309 shear wave velocity tests in the Suzhou site using the suspension P-S velocity logging method. The suspension P-S logging system uses a probe that contains a source and two receivers spaced 1 m apart. The probe is lowered into the borehole to a specified depth, where the source generates a pressure wave in the bore- hole fluid to be received by the receivers. The elapsed time between arrivals of the waves at the receivers is used to determine the average velocity of a 1-meter-high column of soil around the borehole. An example sequential waveform arrival along depth profile is shown in Figure 2.3(a) and the corresponding shear-wave velocity profile is shown in Figure 2.3 (b).

(a)   (b)
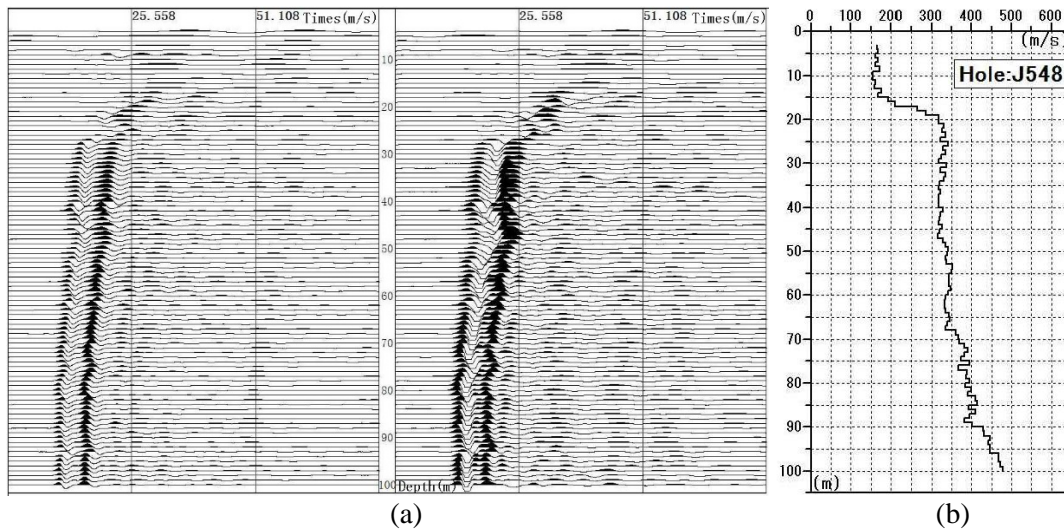
Figure 2.3 Sample shear-wave velocity data obtained from the suspension P-S velocity logging method: (a) depth sequential waveform arrivals; (b) shear wave velocity (Vs) versus depth.

The location of this profile is marked in Figure 2.1 as an triangle. In general, the shear-wave velocity profile corresponds well with the expected soil conditions. For the top 20 m, the shear wave velocity is relatively small (around 150 m/s), which corresponds to the soft soil layers (types 3-1 to 4-1 in Table 2.1). When the depth reaches below 20 m, the shear wave velocity increases significantly (to around 350 m/s) and remains constant from 20 to 50 m, which corresponds well to the relatively hard soil layers (type 6-1 to 8-1 in Table 2.1).

Table 2.2 Summary of soil parameters obtained from borehole samples.

	ρ_{sat} (g/cm^3)	ρ_d (g/cm^3)	LL	PL
Min	1.73	1.14	22.9	11.5
Max	2.96	2.59	70.1	34.4
Mean	2.81	1.51	35.6	20.1

Soil samples were also collected at selected boreholes and analyzed to obtain various soil parameters of interest including the saturated density (ρ_{sat}), the dry density (ρ_d), the liquid limit (LL) and the plastic limit (PL). Table 2.2 summarizes ranges of soil parameters obtained from borehole samples. The water table is found to be at 1.35 to 1.97 m below ground surface.

2.2.3 Calculation of V_{s30} at measurement locations

Given the shear-wave velocity measurement data, a time averaged shear-wave velocity to a profile depth z, denoted as V_{sz}, can be calculated at each measurement location as

$$V_{sz} = \frac{z}{\Delta t_z} \qquad (2.1)$$

$$\Delta t_z = \int_0^z \frac{dz}{V_s(z)} \qquad (2.2)$$

where Δt_z is the travel time for shear waves from depth z to the ground surface; V_{sz} is the shear-wave velocity at depth z; the integral is usually evaluated in practice through summation across velocities taken as constant within depth intervals. When the shear wave velocity profile extends to depths of 30 m or greater, z is taken as 30 m, and the resulting velocity is V_{s30}. When $z < 30$ m, V_{s30} cannot be calculated directly and various correlations between V_{sz} and V_{s30} have been developed to estimate V_{s30} (Boore, 2004; Boore et al., 2011). For this study, all shear wave velocity measurements reach over 30 m. Figure 2.4 plots the V_{s30} values at 309 measurement locations as well as their histogram (the inset).
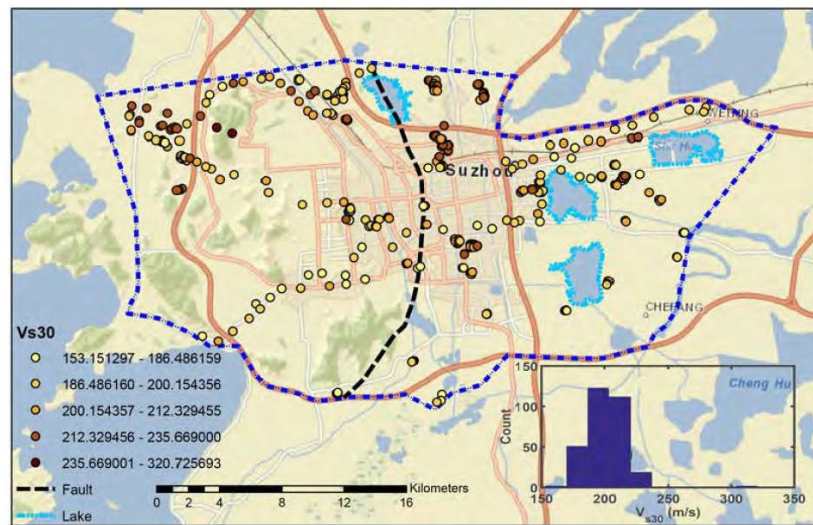
Figure 2.4 Map of V_{s30} measurements in Suzhou City, with histogram inset.

Those V_{s30} values shown in Figure 2.4 are only available at locations with measured shear-wave velocity profiles. To estimate and map V_{s30} values across the region of interest, geostatistical tools and multiscale random field models will be developed and presented in Section 3. Statistical and spatial characterization of the known V_{s30} will be discussed in Section 4.

2.2.4 Secondary V_{s30} data

In addition to the calculated V_{s30} values at measurement locations, proxy-based V_{s30} values are also collected in this study from the U.S. Geological Survey (USGS) global V_{s30} map server (http://earthquake.usgs.gov/hazards/apps/vs30/). Those V_{s30} values are based on a simplified approach that correlates V_{s30} value with the topographic slope (Wald et al., 2004; Allen and Wald, 2009). Such secondary V_{s30} data are necessary because almost all V_{s30} measurements (307 out of 309) are within the Taihu alluvial plain (II$_3$) and the lake-swamp plain (II$_4$), i.e., within relatively soft soils. There is little information on V_{s30} values in hilly areas (I$_1$, I$_2$ and I$_3$). The USGS V_{s30} data will be used to improve V_{s30} predictions in hilly areas, which will be discussed in more detail in Section 2.5. Figure 2.5 plots the USGS V_{s30} data along with its histogram. It is clear from the map that the hilly areas in the western part of the city have much higher V_{s30} values. Moreover, in the alluvial plain, the mean of the USGS V_{s30} is 219 m/s and the minimum is 180 m/s. The mean of the measured V_{s30} values is 200 m/s and the minimum is 153 m/s. Distributions of the USGS and measurement V_{s30} values have also been compared.

In general, it is found that the USGS V_{s30} values tend to predict a higher estimate in the alluvial plain.
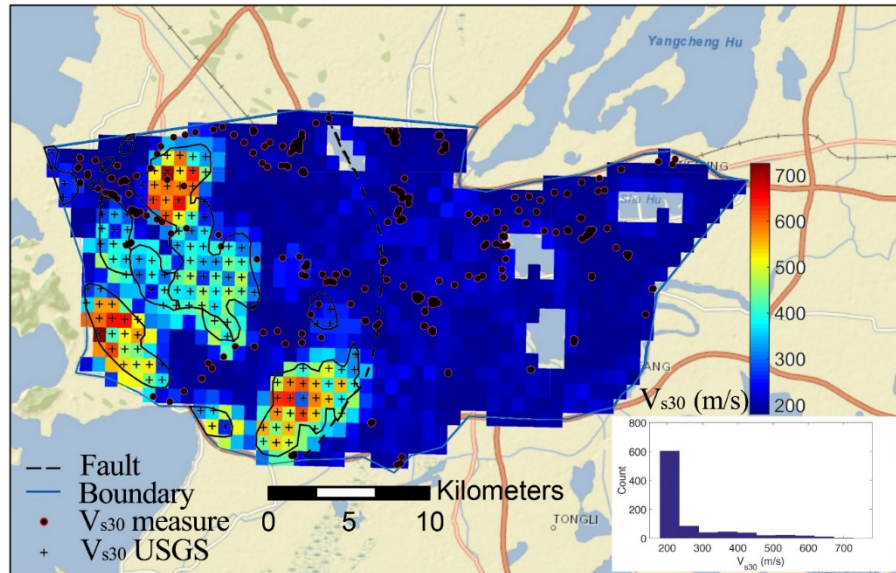
Figure 2.5 USGS global slope-based V_{s30} data: map of the V_{s30} values in Suzhou City and the corresponding histogram (inset).

2.3 Geostatistical approach to characterize spatial variability across scales

In this section, key components of the developed geostatistical tools and random field-based models to map V_{s30} are presented. The rationality behind a geostatistical approach is the fact that the measured soil parameters at one location are more similar to those at neighboring locations than those further away, i.e., soil parameters are spatially correlated. It is desirable to characterize the spatial structure of soil parameters of interest to improve the accuracy of predictions at unsampled locations.

In this study, a form of covariance called the semivariogram is used to describe the spatial structure, which is equal to one half of the variance of two random variables separated by a distance h as

$$\gamma(\mathbf{h}) = \frac{1}{2} Var[Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})] \qquad (2.3)$$

where $Z(\mathbf{u})$ is the variable under consideration at location u and $Z(\mathbf{u} + \mathbf{h})$ is the lagged version of the variable.

Under the condition of second-order stationarity, the semivariogram is related to the spatial correlation $\rho(\mathbf{h})$ by

$$\rho(\mathbf{h}) = 1 - \frac{\gamma(\mathbf{h})}{COV(\mathbf{0})} \qquad (2.4)$$

where COV(0) is the covariance at $\mathbf{h} = \mathbf{0}$. The semivariogram $\gamma(\mathbf{h})$ is typically preferred by the geostatistics community because it only requires the increment $Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})$ to be second-order stationary, which is a weaker requirement than the second-order stationarity of the variable itself. In the following examples, the spatial structure of the soil parameter under consideration (i.e., the V_{s30} value) is characterized by the semivariogram model, which can be converted to ρ and implemented within a random field model.

To account for the multiscale nature of soil variability Chen et al. (2012) and Baker et al. (2011) extended the definition of spatial correlation to multiple scales based on the notion that material properties at the coarser scale are the arithmetically averaged values of the properties over corresponding areas at the finer scale. Such notion is

formally similar to the block kriging (Goovaerts, 1997) but with a different intention to consistently and adaptive refine a coarse scale random field. The multiscale random field allows a higher resolution field to be adaptively generated around areas of high interest.

In this work, two scales of interest are considered and all the subsequent developments apply to variables following the standard Gaussian distribution, i.e., variables after the normal score transformation. The variable of interest Z_I^c at the coarse scale is defined as the arithmetically averaged fine scale values over corresponding areas as (Chen et al., 2012)

$$Z_I^c = \frac{1}{N} \sum_{i=1}^{N} Z_{i(I)}^f \tag{2.5}$$

where the superscripts "c" and "f" correspond to coarse and fine scales, respectively; N is the number of fine scale elements within a corresponding coarse scale area (element) I.

Defining the variable of interest at the fine scale and using the relation of Equation (2.5), the expression for the mean, the variance and the spatial correlation of coarse scale variables of interest can be explicitly derived. The mean of a coarse scale element Z_I^c can be derived by taking the expectation of Equation (2.5) as

$$\mu_{Z^c} = E\left[Z_I^c\right] = \frac{1}{N} \sum_{i=1}^{N} \mu_{Z_{i(I)}^f} = 0 \tag{2.6}$$

where $\mu_{Z_{i(I)}^f}$ is the mean at the fine scale, which equals to zero for variables following the standard Gaussian distribution. Accordingly, if the variance of the fine scale variable is unity, the coarse scale variance, denoted as $\sigma_{Z^c}^2$, can be computed as

$$\sigma_{Z^c}^2 = E\left[\left(Z_I^c\right)^2\right] - 0 = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \rho_{Z_i^f, Z_j^f} \sigma_{Z_i^f} \sigma_{Z_j^f} \tag{2.7}$$

where $\rho_{Z_i^f, Z_j^f}$ is the correlation between two fine scale element i and j with variance $\sigma_{Z_i^f}$

and $\sigma_{Z_j^f}$, respectively. The covariance between any two elements Z_i and Z_j within the

random field is defined as

$$COV\left[Z_i, Z_j\right] = \rho_{Z_i, Z_j} \sigma_{Z_i} \sigma_{Z_j} \tag{2.8}$$

The correlations between all considered scales can be calculated by rearranging

the definition of covariance such that

$$\rho_{Z_i, Z_j} = \frac{COV\left[Z_i, Z_j\right]}{\sigma_{Z_i} \sigma_{Z_j}} \tag{2.9}$$

where Z_i and Z_j are two elements within the random field at any scale with variance $\sigma_{Z_i}^2$

and $\sigma_{Z_i}^2$. By making appropriate substitutions at each scale using Equations. (2.8) and

(2.9), the correlation between elements at different scales can be obtained as (Chen et al.,

2015, 2016)

$$\rho_{Z_I^c, Z_{II}^c} = \frac{\sum_{i=1}^{N} \sum_{k=1}^{N} \rho_{Z_{i(I)}^f, Z_{k(II)}^f}}{\sqrt{\sum_{i=1}^{N} \sum_{j=1}^{N} \rho_{Z_{i(I)}^f, Z_{j(I)}^f}} \sqrt{\sum_{i=1}^{N} \sum_{j=1}^{N} \rho_{Z_{i(II)}^f, Z_{j(II)}^f}}} \tag{2.10}$$

$$\rho_{Z^f, Z_I^c} = \frac{\sum_{i=1}^{N} \rho_{Z^f, Z_{i(I)}^f}}{\sqrt{\sum_{i=1}^{N} \sum_{j=1}^{N} \rho_{Z_{i(I)}^f, Z_{j(I)}^f}}} \tag{2.11}$$

where the Roman numerals I, II... are used for the coarse scale element number; $\rho_{z_I^c, z_{II}^c}$ is the correlation between two coarse-scale elements I and II; ρ_{z^f, z_I^c} is the correlation between a fine-scale element and a coarse scale element I; $\rho_{z_{i(I)}^f, z_{k(II)}^f}$ is the correlation between a fine element i and a fine element k, which belong to two different coarse scale elements I and II, respectively. Given the correlation ρ between elements at different scales, the corresponding covariances COV can be easily obtained via Equation. (2.8).

Once the covariance COV between any two elements at any scale in the random field is determined, a conditional sequential simulation approach is taken for the simulation procedure. The process simulates each value individually, conditional upon all known data and previously simulated values. Using such a process, the conditional distribution of the next value to be simulated in the random field, denoted as Z$_n$, is given by a univariate normal distribution with the updated mean and the variance as

$$\left(Z_n \middle| \mathbf{Z}_p \right) \sim N\left(\mathbf{\Sigma}_{np} \cdot \mathbf{\Sigma}_{pp}^{-1} \cdot \mathbf{Z}_p, \sigma_n^2 - \mathbf{\Sigma}_{np} \cdot \mathbf{\Sigma}_{pp}^{-1} \cdot \mathbf{\Sigma}_{pn} \right) \qquad (2.12)$$

where $\mathbf{Z}_p$ is a vector of all known or previously simulated points; $\mathbf{\Sigma}_{np}$, $\mathbf{\Sigma}_{pp}$, $\mathbf{\Sigma}_{pn}$ are covariance matrices; σ_n^2 is the covariance of the next simulated point; the subscription "p'' and "n'' refer to the "previous" simulated point(s) and the "next" point to be simulated, respectively. Eq. (2.12) means that the unknown value Z_n at an unmeasured location can be drawn from the conditional normal distribution with the mean $\mathbf{\Sigma}_{np} \cdot \mathbf{\Sigma}_{pp}^{-1} \cdot \mathbf{Z}_p$ and the variance $\sigma_n^2 - \mathbf{\Sigma}_{np} \cdot \mathbf{\Sigma}_{pp}^{-1} \cdot \mathbf{\Sigma}_{pn}$. Once Z_n is generated, it is inserted into the "previous" vector, i.e., $\mathbf{Z}_p$, upon which the "next" unknown value at another

unsampled location will be generated. Such process is repeated until all locations within a random field are simulated. A key advantage of such conditional simulation is that it preserves the field data in the random field. Moreover, as pointed out by Baker et al. (2011), such a simulation approach is particular suitable for an adaptive refinement process, where additional fine-scale simulations can be progressively added in the random field in locations deemed necessary.

2.4 Data inference - statistical and spatial characterizations of the known V_{s30} data

The multiscale random field models require as inputs the statistical distributions and the spatial structures of the variable under consideration. In the Suzhou site, a total of 309 V_{s30} values are obtained from direct shear-wave velocity measurements. Figure 2.6 plots the histogram of the 309 V_{s30} measurements. Among those 309 V_{s30} measurements, 307 measurements are located in the two dominating surficial geological units: the Taihu alluvial plain (II_3) and the lake-swamp plain (II_4) as shown in Figure 2.1. Those V_{s30} measurements are grouped by geological units II_3 and II_4 to see whether significant differences exist. Table 2.3 summarizes the statistical characteristics (e.g., mean, variance, maximum, upper quantile, median, lower quantile, minimum) of the two groups. As can be seen from Table 2.3, the statistical characteristics do not differ significantly between the two dominant surficial geological units. In subsequent characterizations and examples, geologic units II_3 and II_4 are grouped together in random field models. In the outcrop areas (I_1, I_2 and I_3), no direct shear-wave velocity

22

measurement is available. The USGS proxy-based V_{s30} data are collected (refer to Figure 2.5) and incorporated as known data in those outcrop areas in subsequent random field simulations.
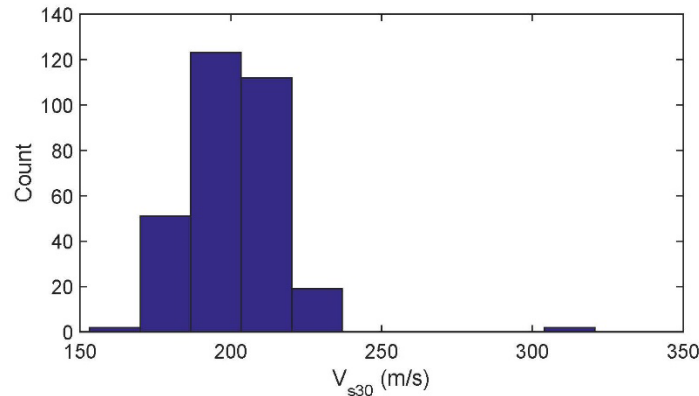
Figure 2.6 Histogram of all 309 V_{s30} values calculated from shear-wave velocity measurements.

Figure 2.7 plots all measurement data projected in the east–west and north–south directions along with the trend lines. The trend line along the west–east direction is almost a horizontal line, indicating little trend in this direction. On the other hand, Figure 2.7 (b) shows slightly increased V_{s30} values from north to south. However, the change is still relatively mild to make any significant impact. It should be pointed out that 307 of the 309 V_{s30} measurements are in the Taihu alluvial plain (II$_3$) and lake-swamp plain (II$_4$). So, the trend analysis reveals the trend (or no trend) of V_{s30} in those geological units only.

Table 2.3 Statistical characteristics of the known V_{s30}.

Statistical parameter	II$_3$	II$_4$	Combined II$_3$ and II$_4$
Data count	143	164	307
Mean	198	202	200
Variance	205	216	192
Maximum	236	233	236
Upper quantile	208	212	201
Median	196	203	200
Lower quantile	188	193	191
Minimum	172	153	153

(a) West-east   (b) North-south
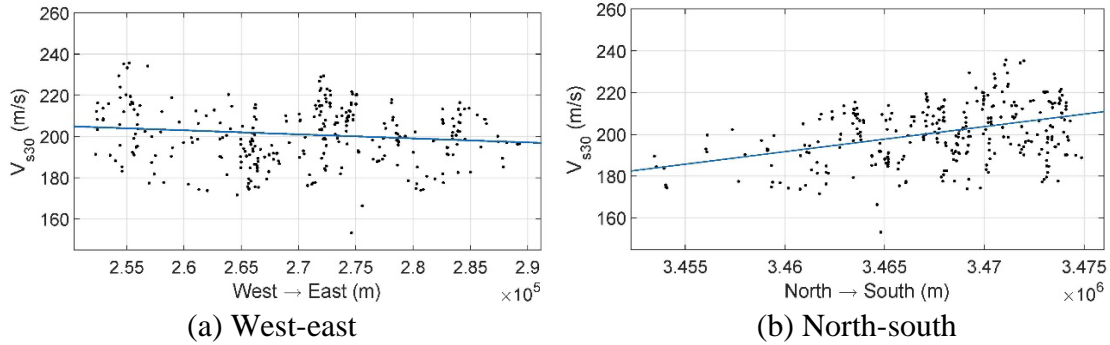
Figure 2.7 Trend of the known V_{s30} values at measurement locations along (a) the west–east direction and (b) the north–south direction.

      The empirical or sample semivariogram of V_{s30} measurements are also computed to infer their spatial structure in the studied region. The empirical semivariogram, denoted as $\hat{\gamma}(\mathbf{h})$, is calculated as (Goovaerts, 1997)

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} \left[Z(\mathbf{u}_\alpha) + Z(\mathbf{u}_\alpha + \mathbf{h}) \right]^2 \tag{2.13}$$

where $N(\mathbf{h})$ is the number of pairs of data ($Z(\mathbf{u}_\alpha)$ and $Z(\mathbf{u}_\alpha + \mathbf{h})$) separated by a vector distance $\mathbf{h}$. To facilitate the incorporation of the semivariogram into random field models, the empirical semivariogram is typically fitted by a basic semivariogram model or a linear combination of several basic semivariogram models that are permissible (Goovaerts, 1997). Fig 2.8 plots the empirical semivariogram model as well as the fitted exponential model of the form

$$\gamma(h) = \omega \left[1 - \exp\left(-\frac{3h}{a} \right) \right] + \tau \tag{2.14}$$

where h is a scalar measure of the separation distance between a pair of points; a is the range, i.e., the distance at which the semivariogram levels off and beyond which

24

the semivariance is constant; $\omega + \tau$ is the sill, which is the constant semivariance beyond

the range. The fitted range for this study site is 2973 m and the sill is 0.9833.
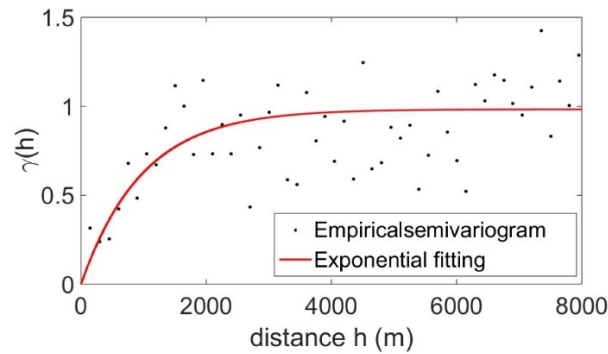
Figure 2.8 Empirical and fitted semivariogram based on known V_{s30} at measurement locations.

2.5 V_{s30} mapping of the Suzhou site

With the inferred model parameters, the known V_{s30} at measurement locations and

the secondary V_{s30} information in the outcrop areas (I_1, I_2, I_3), the multiscale random field

models are used to generate V_{s30} maps of the Suzhou site. An initial coarse grid with an

element size of 500×500 m is used. Lakes are excluded from the V_{s30} maps. The new

maps account for and preserve the site-specific shear-wave velocity measurements and

the inherent multiscale soil spatial structure. When coupled with Monte Carlo

simulations, uncertainties associated with the generated V_{s30} maps can also be estimated.

The generated V_{s30} maps will be compared with the available topography-based V_{s30} map

obtained from the U.S. Geological Survey global V_{s30} database.

A typical set of V_{s30} realizations (single and multiscale) is shown in Figure 2.9. In the multiscale realization, each coarse grid neighboring a measurement location is refined into 36 fine scale elements, where high resolution V_{s30} are generated through the multiscale model described in Section 2.3. Such fine scale field enables predictions across different scales and can facilitate estimation of uncertainties at much finer scales without sacrificing computation efficiency. The secondary V_{s30} data from USGS, placed on a grid with a spacing of 800 m, are incorporated as known point data values in the random fields in the outcrop areas. It should be noted that for the current study, the amount of the secondary data is fixed. A preliminary work (Liu et al., 2017) is undergoing to investigate the effect of secondary data on the spatial structure and the distribution of the resulting V_{s30} realizations.
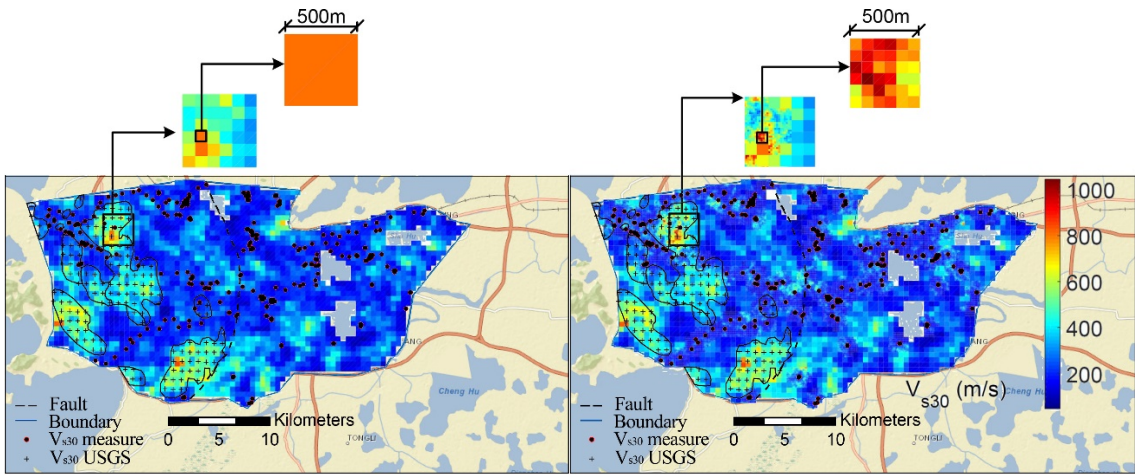
(a) Single scale   (b) Multi scale

Figure 2.9 Sample random field realizations of V_{s30} in Suzhou site.

The corresponding histograms and empirical semivariograms of the simulated V_{s30} are shown in Fig 2.10. Both single and multi-scale simulations preserve the statistical characterizations and the spatial structures of V_{s30} inferred from the known measurement data.

(a) Single scale   (b) Multiscale
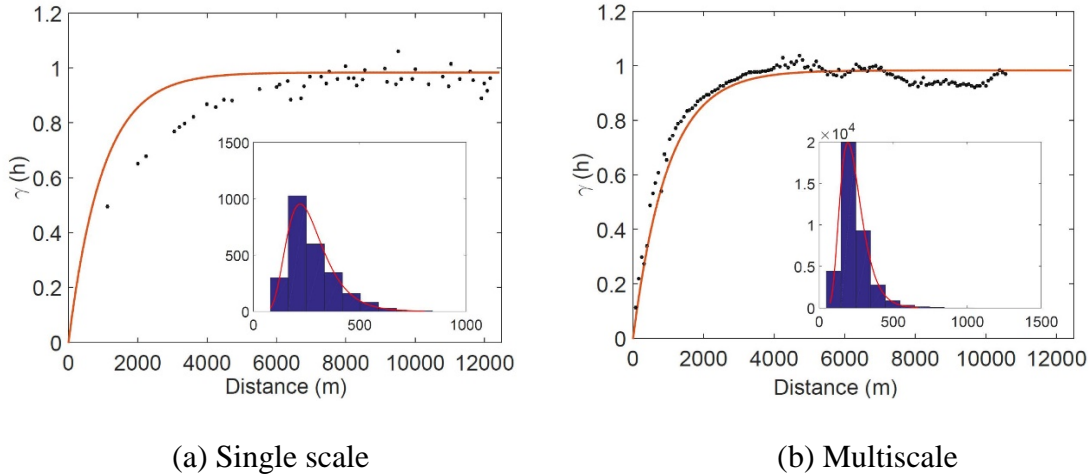
Figure 2.10 Semivariograms and histograms (the insets) of simulated V_{s30} from one set of random field realizations in Suzhou site. Black dots are the empirical semivariogram and the red solid line is the specified exponential model. The red solid line in the histogram inset is the fitted probability density function.

Coupling the random field model with Monte Carlo simulations, the expected V_{s30} values across the Suzhou site as well as the associated uncertainties can be obtained. Maps of the expected V_{s30} values, averaged from 1000 independent Monte Carlo simulations, are shown in Figure 2.11 (a) and (b). An obvious trend manifested in the map is that high V_{s30} values occur in the southern and western part of the city, especially the hilly areas. Low values are common in the northern and eastern part, which are consistent with the trends observed in the measurement data and the knowledge about the geology of this studied area. It should be noted that, in the current study, geological

27

boundaries are not explicitly incorporated in the data reference or in random field simulations.
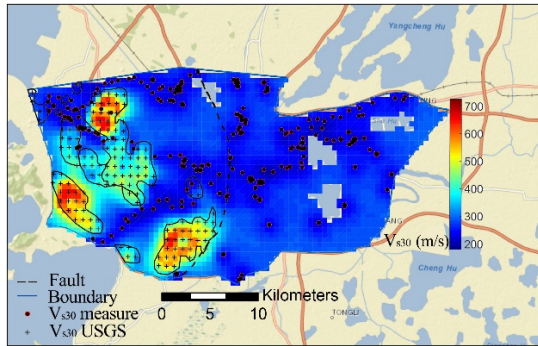
(a) Single scale
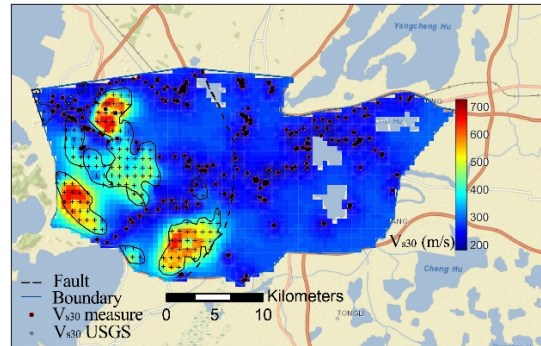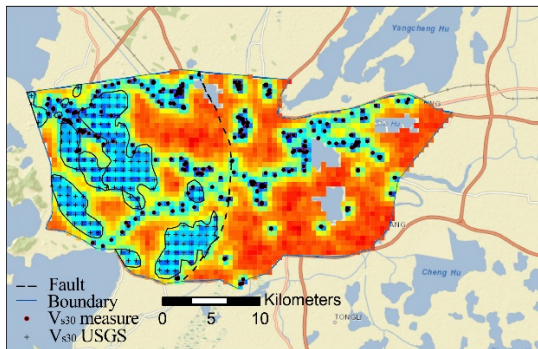
(b) Multiscale

(c) Single scale
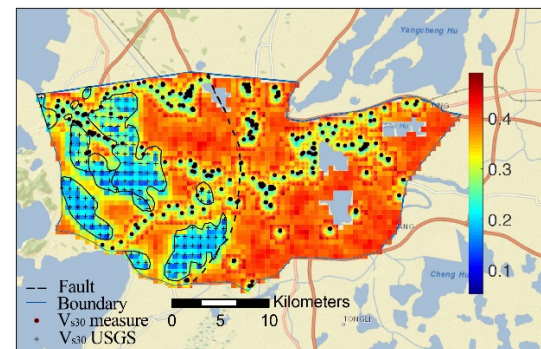
(d) Multiscale

Figure 2.11 Expected Vs30 values and associated uncertainties (coefficient of variations) at the Suzhou site.

One of the strengths of the proposed method is its ability to estimate uncertainties associated with generated V_{s30} maps. To quantify uncertainties, coefficients of variation (COV) from 1000 independent Monte Carlo simulations are calculated at each location and plotted in Figure 2.11 (c) and (d). As shown in the figure, the COVs are generally

very small and approach zero around locations with measurement data. It is interesting to note that the uncertainties associated with single scale map are smaller compared to the multiscale counterpart, especially around locations with known data. Recall that the coarse (single) scale field can be seen as the average of the corresponding fine (multi) scale realizations and such averaging process results the reduced uncertainties observed in Figure 2.11(c) and (d).

The empirical semivariograms of the predicted V_{s30} values are calculated and shown in Figure 2.12 along with the error bars indicating ±one standard deviation. It can be seen from Figure 2.12 that the specified exponential spatial structure, which is inferred from measurement data, is preserved well in the simulations. It is noted that the spatial structures, quantified here by the semivariogram, are different between single and multiscale. This is because the coarse (single) scale spatial correlation is derived based on the notion that a coarse scale element is the average of the corresponding fine scale element. This averaging of the fine scale points will effectively increase the correlation of a given distance relative to the fine scale. This effect has been previously reported and studied in Chen et al. (2012).

| (a) Single scale | (b) Multiscale |
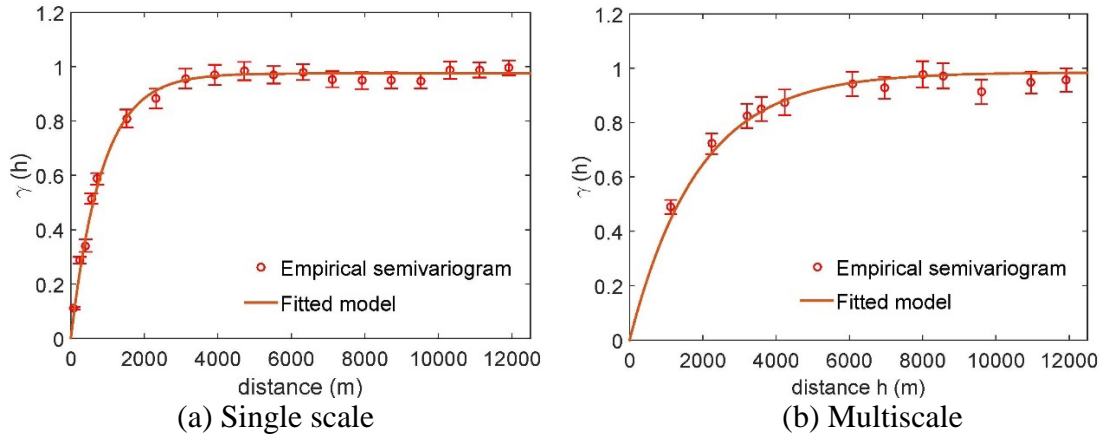
Figure 2.12 Empirical semivariograms of predicted V_{s30}. Error bars indicate ±one standard deviation.

2.5.2 Comparison with USGS V_{s30} maps

The newly generated multiscale random field-based V_{s30} maps incorporate and preserve the site-specific shear wave velocity measurement data and their spatial dependency. To understand the effect of local measurement data and spatial dependency on V_{s30} mapping, Figure 2.13 plots side-by-side the V_{s30} map from the current study and the one from the USGS global V_{s30} map server. Note that the upper limit of the color map is set to $V_{s30} = 360$ m/s, which corresponds to the upper bound of the NEHRP site class D (refer to Table 2.4). Since most of the Suzhou site has soft soil with relatively low V_{s30} values, such scale makes the difference among two maps more distinguishable. As can be seen from Figure 2.13, while both maps capture the general trend of high V_{s30} values in the western hilly area and low V_{s30} values in the eastern region, the current map has significantly higher resolution and has captured the transition from hilly to plain region fairly well. The current V_{s30} map captures a northeast–southwest band with low V_{s30}, as

reflected from the V_{s30} measurement data, which is missed in the proxy-based USGS map. Moreover, the current map precisely preserves the known V_{s30} values at measurement locations and provides multiscale resolution, which contains small-scale V_{s30} information. Such information can be used to estimate uncertainties at a much higher resolution without sacrificing the overall computational efficiency.

(a) USGS   (b) Multiscale map
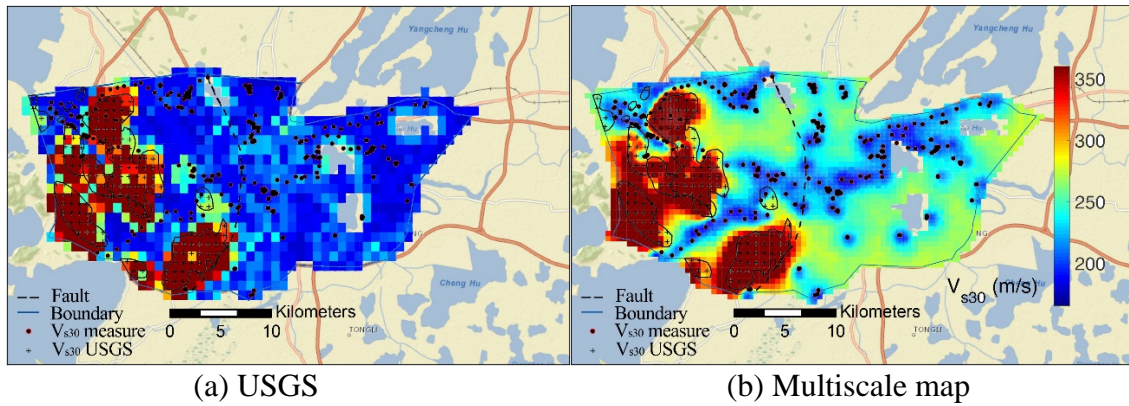
Figure 2.13 Comparison of Vs30 maps: (a) USGS topography-based proxy; (b) current study.

To quantify the performance of the proxy-based USGS map, the difference between USGS V_{s30} values and the measured V_{s30} normalized by the measured V_{s30} value is calculated and the histogram of all 309 data is plotted in Figure 2.14. As shown in Figure 2.14, many of the normalized differences are within 0 to 40% range with a few points indicating over 100% difference.

Table 2.4 NEHRP site class and corresponding V_{s30} range.

Site Class	Description	V_{s30}
A	Hard rock	>1500 m/s
B	Firm to hard rock	760 to 1500 m/s
C	Dense soil, soft rock	360 to 760 m/s
D	Stiff soil	180 to 360 m/s
E	Soft clay	<180 m/s
F	Soil requiring site specific evaluation	-
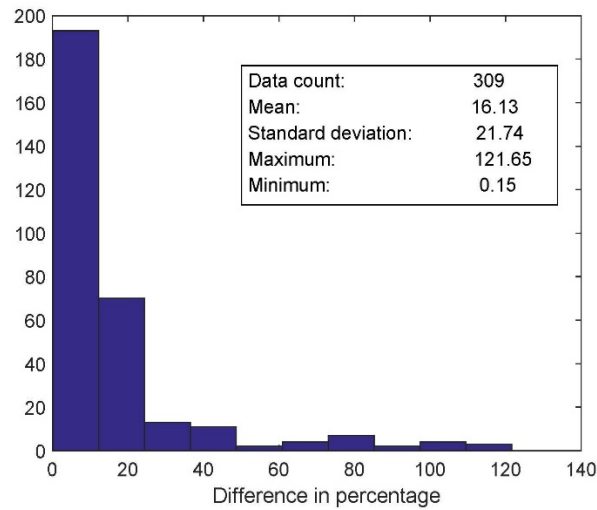
Figure 2.14 Statistic characteristic of the difference (in percentage) between USGS V_{s30} prediction and known V_{s30} at 309 measurement locations.

2.6 Applications of the new V_{s30} maps

V_{s30} is a key indicator of site response in many earthquake engineering applications, such as ground-motion prediction equations, site classification, and earthquake hazard maps. In this section, two of the applications of the newly generated V_{s30} maps will be presented: V_{s30}-based site classification in Section 2.6.1 and the estimation of site amplification factors in Section 2.6.2.

2.6.1 V_{s30}-based site classification

The National Earthquake Hazards Reduction Program (NEHRP) classifies a site into 5 groups and provides the range of V_{s30} values for each class as shown in Table 2.4.

Given a V_{s30} map, the site of interest can be classified based on V_{s30} values following the NEHRP criteria.
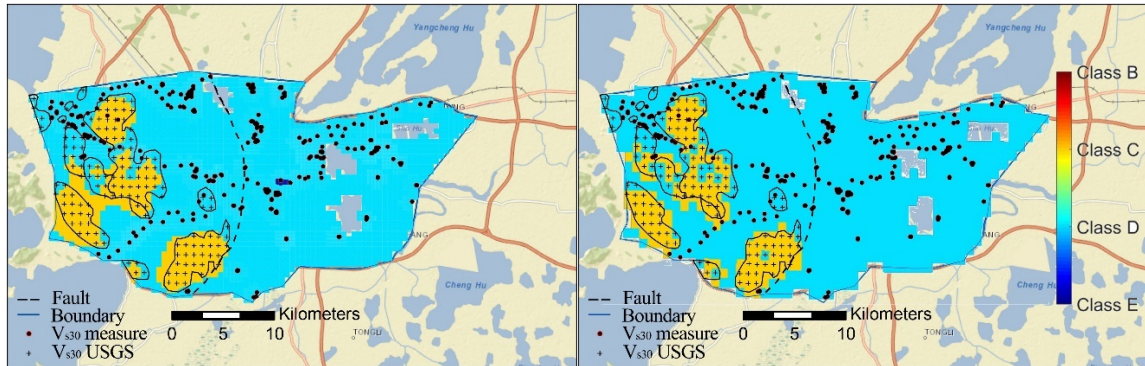
(a) Multiscale   (b) USGS

Figure 2.15 Vs30-based NEHRP site classification (Table 2.4): (a) based on the new multiscale Vs30 map; (b) based on the USGS V_{s30} map.

Figure 2.15 shows the site classification maps for the Suzhou site based on the new multiscale random field-based V_{s30} and the USGS proxy- based V_{s30} maps. The classification map of Figure 2.15 (a) shows that most of the studied region can be classified as NEHRP soil type D, where V_{s30} ranges from 180 to 360 m/s. In the hilly area in the western part, the site is classified as soil type C with V_{s30} values ranging from 360 to 760 m/s. This is consistent with the known engineering geology of this region previously described in Section 2.2.
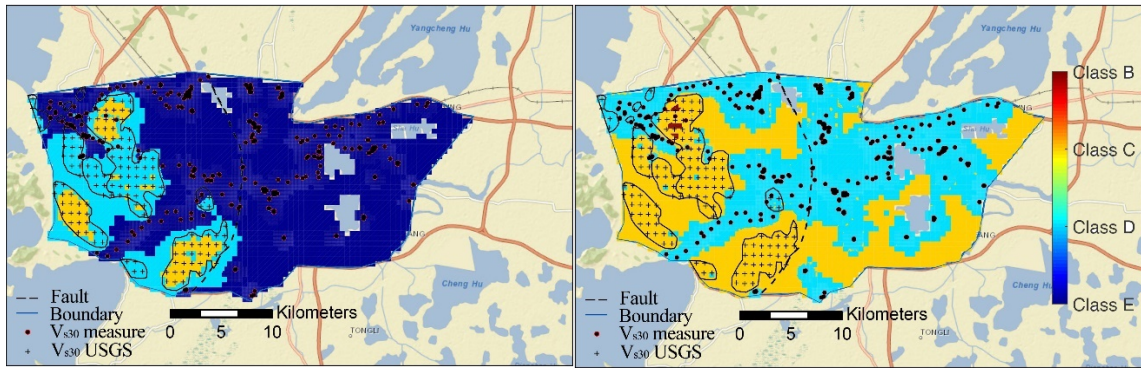
(a) Lower bound   (b) Upper bound

Figure 2.16 Uncertainties associated with the site classification maps based on expected Vs30 values ±one standard deviation: (a) lower bound (mean −one standard deviation); (b) upper bound (mean +one standard deviation).

The site classification shown in Figure 2.15 (a) is based on the expected V_{s30} values averaged from 1000 Monte Carlo simulations as previously shown Figure 2.11 (b). To quantify the associated uncertainties in the site classification, upper and lower bound site classification maps are also generated by using ± one standard deviation of the expected V_{s30} values. The results are shown in Figure 2.16. Compared to the mean V_{s30}-based site classification shown in Figure 2.15 (a), most of the hilly areas in the western part of the city remain in the site class C, but the eastern plain changes to site E when the lower bound (mean minus one standard deviation) V_{s30} map is used, which is considered to be a more conservative estimation.

2.6.2 Amplification factor mapping

The second application of the new V_{s30} map is the estimation and mapping of site amplification factors. Among various commonly used models for estimating site amplification factor, the model by Choi and Stewart (2005) is used in this work to

illustrate the application. In the Choi and Stewart (2005) model, the model for estimating the amplification factor F_{ij} is expressed as

$$\ln\left(F_{ij}\right) = c\ln\left(\frac{V_{s30_{ij}}}{V_{ref}}\right) + b\ln\left(\frac{PHA_{r_{ij}}}{0.1}\right) + \eta_i + \varepsilon_{ij} \tag{2.15}$$

where PHA_r is the peak horizontal acceleration for the reference site condition and is expressed in the unit of the gravitational acceleration g; b is a function of the regression parameters as given in Equation.(2.16); c and V_{ref} are the regression parameters; η_i is a random effect term for the i-th earthquake event with zero median and a standard deviation denoted as τ; ε_{ij} represents the intra-event model residual for the j-th motion in i-th earthquake event, which has a median near zero for well-recorded events with a standard deviation denoted as σ.

The variation of model parameter b is described in the following model (Choi and Stewart, 2005):

$$\begin{aligned} b &= b_1 & V_{s30} < 180 \ (m/s) \\ b &= b_2 + \left(V_{s30} - b_v\right)^2 \frac{b_1 - b_2}{\left(180 - b_v\right)^2} & 180 < V_{s30} < b_v \ (m/s) \\ b &= b_2 & b_v < V_{s30} < 520 \ (m/s) \\ b &= b_2 - \left(V_{s30} - 520\right)\frac{b_2}{240} & 520 < V_{s30} < 760 \ (m/s) \\ b &= 0 & V_{s30} > 760 \ (m/s) \end{aligned} \tag{2.16}$$

where the units of V_{s30} are in m/s; b_1, b_2 and b_v are model parameters. For this reference model, Abrahamson and Silva (1997) provided values of site factor model parameters from regression analysis, which are summarized in Table 2.5.

Table 2.5 Regression parameters for site amplification factors after Abrahamson and Silva (1997).

Parameter	b_1	b_2	b_v	c	V_{ref} (m/s)	τ	σ
$F_a(0.3)$	-0.41	-0.11	300	-0.46	532	0.35	0.54
$F_v(1.0)$	-0.39	0.02	300	-0.69	519	0.41	0.55
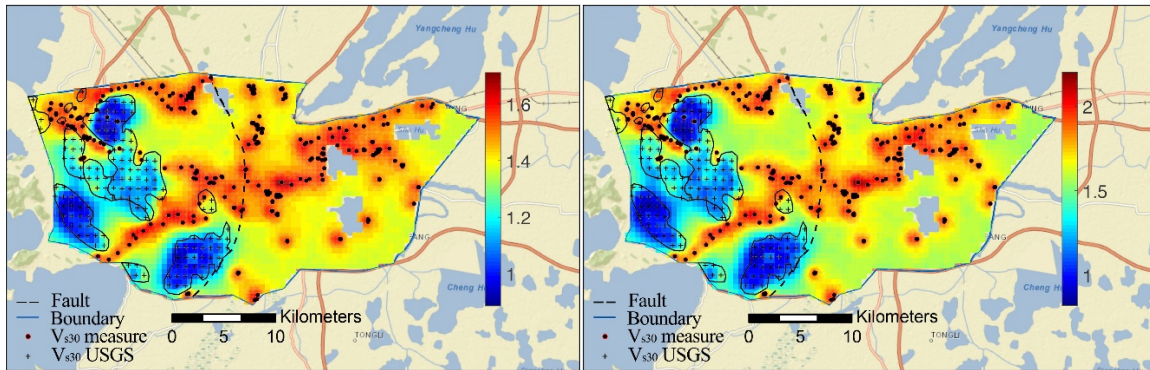
(a) T=0.3s   (b) T=1.0s

Figure 2.17 Maps of amplification factors in Suzhou City based on the Choi and Stewart (2005) model: (a) Fa (T = 0.3 s) and (b) Fv (T = 1.0 s).

With the amplification model Equation. (2.15), Equation. (2.16) and the fitting parameters in Table 2.5, site factors F_a (corresponding to a low-period range with T = 0.1 – 0.5 s) and F_v (corresponding to a mid-period range with T = 0.4 – 2.0 s) are calculated based on an assumed PHA_r of 0.1 g. Results of the site factors are plotted in Figure 2.17 for F_a (T = 0.3 s) and F_v (T = 1.0 s). Figure 2.17 shows that most of the eastern and central areas have relatively high amplification factors with a maximum of 1.7 for T = 0.3 s and 2.2 for T = 1.0 s, which correlates well with the softer soils (NEHERP site classes D and E, refer to Figure 2.15 (a)).

2.7 Summary

In this work, a multiscale random field-based framework is presented to map V_{s30} values over extended areas. The random field model explicitly accounts for the spatial variability of V_{s30} across different scales while incorporates and preserves measured V_{s30} data. The framework is applied to map V_{s30} over the Suzhou site, where 309 shear-wave velocity measurements and topography-based V_{s30} values are compiled. Monte Carlo simulations are coupled with the random field model to quantify uncertainties of the generated multiscale V_{s30} map. The new map is then applied to site classification and amplification factor characterization in the studied region. In summary, it is found that:

1. Quantitatively consistent V_{s30} estimates over different length scales over the entire studied region can be obtained using the multiscale random field model. The resulting map has multi- scale resolutions and is particular convenient to incorporate and preserve local measurement data into a regional V_{s30} map.

2. Comparison of the new V_{s30} map with existing USGS topography-based Vs30 map shows that the new V_{s30} map provides more accurate and more detailed V_{s30} values, especially in the eastern plain region of the studied site because of the incorporated local Vs30 measurements and their spatial dependency.

3. Uncertainties associated with the new V_{s30} map are quantified in terms of the coefficient of variation (COV) calculated from Monte Carlo simulations. In general, the COVs approach zero around locations with measurement data and

gradually increase in areas without any known V_{s30} values. COVs in single scale random field map are found to be slightly smaller when compared to the multiscale counterpart.

4. The site application map based on the newly generated V_{s30} map shows that relatively stiff soil (NEHRP site class C) is found in the northwestern part of the city and the soil tends to be softer in the southeastern region (NEHRP site class D and E). This trend in the soil type correlates well with the calculated amplification factor map, where high amplification factors are predicted in the southeastern part of the city, indicating potential seismic amplification effect in this region.

CHAPTER III

3. PARAMETER STUDY OF HYBRID GEOTECHNICAL AND GEOLOGIC

DATA-BASED RANDOM FIELD FRAMEWORK*

3.1 Introduction

The time-averaged shear-wave velocity in the top 30 meters of the subsurface material, denoted as V_{s30}, is an important site parameter used in estimating site response, classifying sites in recent building codes and loss estimation (Boore, 2004). While the V_{s30} can be computed directly given a shear-wave velocity measurement, such geophysical measurements are typically very sparse. Therefore, various descriptors or quantitative metrics of site condition have been proposed for the purpose of estimating V_{s30} in the absence of geophysical measurements and mapping V_{s30} in the area of interest. Correlation between geologic units and shear wave velocity form the basis of a series of V_{s30} maps developed over the past 15 years. Wills et al. (2006) developed a map of geologic units that can be distinguished by their shear wave velocity and used that map to supply velocity estimates. After that, Wills et al. (2015) updated the previous map and generated a modified V_{s30} map for California based on geology and topography. All these works focus on large regions as they provide the general trend of V_{s30} in at least state scale. The geological data, to some extent, provide information on large-scale (regional scale) material heterogeneity.

---

*A similar form of this chapter has been submitted at the time of writing: Liu, W, Chen, Q, Wang, C, Juang, CH. Parameter study of hybrid geotechnical and geologic data-based random field framework.

The previously mentioned work ignores the inherent spatial variability of soil properties and limits the accuracy. To improve the accuracy and to take advantage of the ever-expanding high-quality geotechnical database, geostatistical methods have received increased attention in recent years. Geostatistical methods allow the incorporating of geotechnical data and slow the explicit consideration of the soil property spatial variability. For instance, Thompson and his coworkers (Thompson et al., 2014, 2011, 2010), used regression kriging (RK) to develop a new map of V_{s30} for California which accounting for geology, topography and site-specific V_{s30} measurements. Yong et al. (2013) and Wald et al. (2011) applied the kriging-with-a-trend method to mapping V_{s30}, where the baseline model was derived from the topographic slope. Lee and Tsai (2008) established the correlation between the shear-wave velocity (V_s) and the N value of the standard penetration test (SPT-N) and adopted the kriging with varying local means to update the V_{s30} maps in Taiwan. Thompson et al. (2007) modeled the horizontal variability of near-surface soil shear wave velocity in the San Francisco Bay Area using geostatistical methods. Lee et al. (2017) developed a 3D model of the geologic structure and associated seismic velocities in the Canterbury, New Zealand. More recently, Chen and co-workers (Chen et al., 2016a, 2016b; Liu et al., 2017a; Liu et al., 2017b), developed a multiscale random field model to account consistently for soil spatial variability across multiple length scales and was applied to mapping soil properties and liquefaction potentials across a region.

The geostatistical methods consider the spatial variability of the mapped properties derived from geotechnical data. However, the spatial structure of the mapped

properties could vary within and across different geological units. The insufficient and unevenly located geotechnical data cannot represent the large-scale (regional scale) material heterogeneity.

To overcome the limitations in the aforementioned mapping methods, Wang and Chen (2017) proposed a novel hybrid geotechnical and geological data-based framework. The framework offers an idea to consistently accounts for the spatial variability of soil properties across scales, while simultaneously preserving constraints imposed by geological boundaries. Both geotechnical and geological data are integrated into a random field model through a conditional sequential simulation technique.

But this framework has not been well-rounded tested and several significant parameters have not been calibrated. In this chapter, sensitivity analysis, and parameter learning are operated based on a synthetic digital field and an overall workflow for integrating multiple sources of data in random field model for regional soil properties mapping is established based on Wang and Chen (2017) hybrid geotechnical and geological data-based framework. Recommendations on how to act in each specific mapping situation and step-by-step instructions for accomplishing mapping are provided including the methodology to calibrate Markov Bayes coefficient B and integrate secondary data.

The order of presentation of this chapter goes as follows: Section 3.2 summarizes the key components of the developed geostatistical tools for mapping V_{s30}; in Section 3.3, spatially correlated synthetic digital soil field is presented; random field realizations of V_{s30} and Monte Carlo analysis is discussed in detail in Section 3.4; in Section 3.5,

parameter study of Hybrid Geostatistical and geological data-based model is operated; Section 3.6 describes the overall workflow for integrating multiple sources of data in the random field model is established; and all the findings will be discussed in Section 3.7.

3.2 Hybrid Geostatistical and geological data-based model

In this section, key components of the hybrid geotechnical and geological data-based framework to map V_{s30} are presented. The rationality behind this hybrid approach is the fact that other than spatially correlated, soil parameters are related to the geologic units to a certain extent. The geological data provide information on large-scale (regional scale) material heterogeneity. It is desirable to combine the spatial structure of soil parameters of interest and its characteristic affected by its geologic properties to improve the accuracy of predictions at unsampled locations.

In this study, a form of covariance called the semi-variogram is used to describe the spatial structure, which is equal to one half of the variance of two random variables separated by a distance $\mathbf{h}$ as

$$\gamma(\mathbf{h}) = \frac{1}{2} Var[Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})] \tag{3.1}$$

where $Z(\mathbf{u})$ is the variable under consideration at location u and $Z(\mathbf{u}+\mathbf{h})$ is the lagged version of the variable.

Under the condition of second-order stationarity, the semi-variogram is related to the spatial correlation $\rho(\mathbf{h})$ by

$$\rho(\mathbf{h}) \;=\; 1 - \frac{\gamma(\mathbf{h})}{COV(\mathbf{0})}$$

(3.2)

where COV(0) is the covariance at $\mathbf{h} = \mathbf{0}$. The semi-variogram $\gamma(\mathbf{h})$ is typically preferred by the geo-statistics community because it only requires the increment $Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})$ to be second-order stationary, which is a weaker requirement than the second-order stationarity of the variable itself. In the following examples, the spatial structure of the soil parameter under consideration (i.e., the V_{s30} value) is characterized by the semi-variogram model, which can be converted to ρ and implemented within a random field model.

To facilitate the incorporation of the semi-variogram into random field models, the empirical semi-variogram is typically fitted by a basic semi-variogram model or a linear combination of several basic semi-variogram models that are permissible (Goovaerts, 1997). The form of the exponential model used in this study is

$$\gamma(h) = \omega\left[1 - \exp\left(-\frac{3h}{a}\right)\right] + \tau$$

(3.3)

where h is a scalar measure of the separation distance between a pair of points; a is the range, i.e., the distance at which the semi-variogram levels off and beyond which the semi-variance is constant; $\omega + \tau$ is the sill, which is the constant semi-variance beyond the range.

A conditional sequential simulation algorithm is implemented in this work to generate random field realizations of V_{s30}. This algorithm integrates and preserves

43

multiple sources of known data (e.g., primary and secondary V_{s30} data). In this algorithm, the realization of a random variable Z_n is represented by a joint distribution as follows

$$\begin{bmatrix} Z_n \\ \mathbf{Z}_p \\ \mathbf{Z}_s \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_n \\ \boldsymbol{\mu}_p \\ \boldsymbol{\mu}_s \end{bmatrix}, \begin{bmatrix} \sigma_n^2 & \Sigma_{np} & \Sigma_{ns} \\ \Sigma_{pn} & \Sigma_{pp} & \Sigma_{ps} \\ \Sigma_{sn} & \Sigma_{sp} & \Sigma_{ss} \end{bmatrix} \right) \tag{3.4}$$

where $N(\mu, \Sigma)$ denotes the vector of random variables following a joint normal distribution with the mean vector $\boldsymbol{\mu}$ and the covariance matrix Σ; Z_n is the random variable to be generated with the expected value μ_n; $\mathbf{Z}_p$ is the vector of previously generated or known primary random variables with the vector of expected values $\boldsymbol{\mu}_p$; $\mathbf{Z}_s$ is a vector of secondary random variables with the vector of expected values $\boldsymbol{\mu}_s$; σ_n is the standard deviation of Z_n; Σ is the covariance matrix with subscripts 'n', 'p' and 's' denoting 'next', 'previous primary' and 'secondary', respectively. The individual terms in the covariance matrix are defined as

$$COV\left[Z_i, Z_j \right] = \rho_{Z_i, Z_j} \sigma_{Z_i, Z_j} \tag{3.5}$$

where ρ_{Z_i, Z_j} is the correlation between two elements Z_i and Z_j within the random field at any scale with a standard deviation of σ_{Z_i} and σ_{Z_j}, respectively.

Given the joint distribution Equation (3.4), the distribution of the random variable Z_n, conditional upon all previously simulated and known primary and secondary data, is given by a univariate normal distribution with the updated mean and variance as

$$\left(Z_n \middle| \left(\mathbf{Z}_p, \mathbf{Z}_s \right) \right) \sim N\left(\widetilde{\mu}_n, \widetilde{\sigma}_n \right) \tag{3.6}$$

with

$$\widetilde{\mu}_n = \begin{bmatrix} \Sigma_{np} & \Sigma_{ns} \end{bmatrix} \begin{bmatrix} \Sigma_{pp} & \Sigma_{ps} \\ \Sigma_{sp} & \Sigma_{ss} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{Z}_p \\ \mathbf{Z}_s \end{bmatrix} \tag{3.7}$$

$$\widetilde{\sigma}_n = \sigma_n^2 - \begin{bmatrix} \Sigma_{np} & \Sigma_{ns} \end{bmatrix} \begin{bmatrix} \Sigma_{pp} & \Sigma_{ps} \\ \Sigma_{sp} & \Sigma_{ss} \end{bmatrix}^{-1} \begin{bmatrix} \Sigma_{pn} \\ \Sigma_{sn} \end{bmatrix} \tag{3.8}$$

where symbols in equations (3.7) and (3.8) have all been defined after equation (3.4). The value of a random variable Z_n at an unsampled location is drawn from the above joint distribution. Once generated, Z_n becomes a data point in the vector $\mathbf{Z}_p$ to be conditioned upon by all subsequent simulations. This process is repeated by following a random path to each unknown location until all the values in the field have been simulated – that is, a map of the primary variable for the region of interest is generated.

To perform the conditional sequential simulation in Equation (3.6), three covariances must be determined: one for the primary variable, one for the secondary variable and a cross-covariance describing the relationship between these variables. With a relatively sufficient amount of V_{s30} data calculated from shear wave velocity test, the covariance of the primary variable can be easily obtained from the inferred spatial correlation between primary data. However, the direct calculation of the secondary and cross-covariances can be challenging. In this work, one simplified approach is adopted based on the Markov–Bayes hypothesis described by Goovaerts (1997) to derive the secondary and cross-covariances by calibrating them to the primary covariance as (Goovaerts, 1997; Moysey et al., 2003)

$$COV_s(\mathbf{h}) = \begin{cases} |B| \cdot COV_p(\mathbf{h}) & \text{for } \mathbf{h} = 0 \\ B^2 \cdot COV_p(\mathbf{h}) & \text{for } \mathbf{h} > 0 \end{cases} \tag{3.9}$$

45

$$COV_{ps}(\mathbf{h}) = B \cdot COV_p(\mathbf{h}) \tag{3.10}$$

where B is the Markov–Bayes coefficient; COV_p is the covariance for the primary variable, COV_s is the covariance for the secondary variable and COV_{ps} is the covariance between the primary and the secondary variables; $\mathbf{h}$ is the distance vector separating two random variables. The Markov–Bayes coefficient B generally varies between 0 and 1 when primary and secondary variables are positively correlated. Its value affects the relative importance of primary data and secondary data and this effect will be illustrated in the results section.

The coefficient B can be chosen based on a calibration procedure recommended by Deutsch and Journel (1998) such that the value B is determined as the difference between two conditional expectations as follows

$$B = E_1 - E_0 \tag{3.11}$$

where the two conditional expectations are defined as

$$E_1 = E\left(\text{Prob}\{Z_s \le z\} \middle| Z_p \le z\right) \tag{3.12}$$

$$E_0 = E\left(\text{Prob}\{Z_s \le z\} \middle| Z_p > z\right) \tag{3.13}$$

where E is the expectation operator; E is the expectation; Z_s is the secondary variable (e.g. the geological data-based LPI value) and $\text{Prob}\{Z_s \le z\}$ is the probability of Z_s less than or equal to a threshold value z (e.g. a given V_{s30} threshold value); Z_p is the primary variable (e.g. the geotechnical data-based V_{s30} value). The conditional expectation E_1 will be close to 1 if the primary and secondary data support each other – that is, the two data

predict similar liquefaction hazard levels. The conditional expectation E_0 will be close to 1 if the primary and secondary data contradict each other – that is, the two data predictions of liquefaction hazard contradict each other.

<u>3.3 Spatially correlated synthetic digital soil field</u>

To test the hybrid approach presented in Section 3.2, a spatially correlated synthetic digital soil field is created and its V_{s30} fields are assumed as the benchmark. The dimension of the synthetic digital soil field is set as 2000×4000 m and a soil element size is correspondingly set as 4×4 m. There are a total of 500,000 soil elements in the field. In order to test the framework's performance in mapping in the field with different site conditions, this synthetic field is bisected into two sub-field which belongs to two geologic units. The left part, denoted as Unit 1, is assumed as crystalline rock (Xtaline) and the right part (denoted as Unit 2) is assumed as quaternary alluvium (Qal2).

Within this two sub-field, two two-dimensional and spatially correlated V_{s30} sub-fields are generated separately and their values are assigned to each soil element as shown in Figure 3.1. The known data points and distribution used to generate the synthetic field are based on the experience gained from the V_{s30} database in Alameda County of California (https://earthquake.usgs.gov/data/vs30/us/). The spatial correlation length is assumed as 3.7 km based on Thompson, E.M. et. al (2007). This assumption made on spatial correlation is for the convenience of generating the digital field. In this

47

synthetic field, two sub-field follow different distribution but share the same spatial correlation (Figure 3.2).
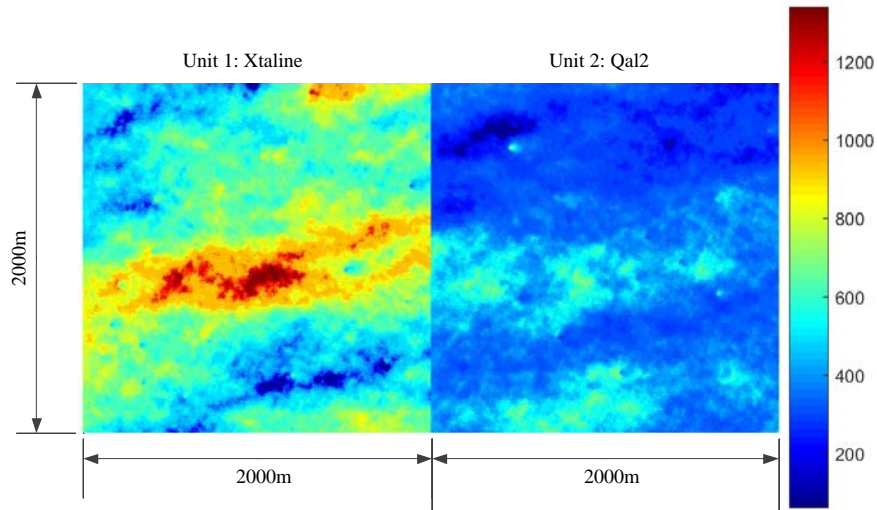
Figure 3.1 The 2D view of synthetic digital V_{s30} field
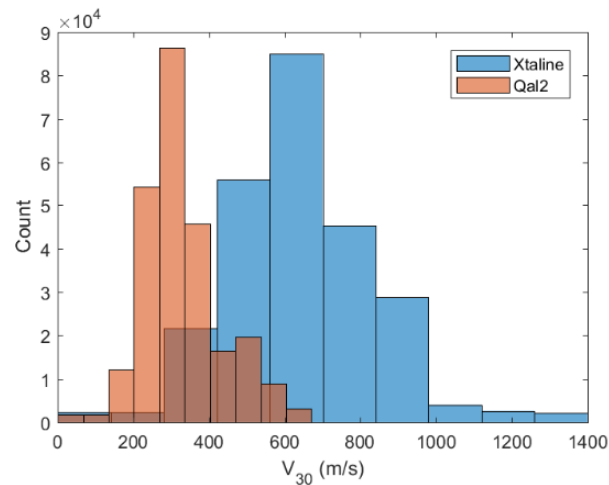
Figure 3.2 Histogram of the synthetic digital V_{s30} field

For simplicity, the synthetic digital V_{s30} field is noted as the "true" field for use in this research. It should be noted that the true distribution and spatial structure of this digital soil field are unknown to random field-based modeling and mapping, the same as in the case of a real soil field.

3.4 Random field realizations of V_{s30} and Monte Carlo analysis

Based on the synthetic field built in Section 3.3, the site investigation is operated for getting 'measured data'. As suggested by Webster and Oliver (1992), a sample size of 100 should give acceptable confidence to estimate distribution and semi-variograms of soil properties. Hence, the investigation plan is designed with a total of 128 evenly spaced V_{s30} sample locations. Among them, 64 samples locate in Unit 1 and 64 samples locate in Unit 2. The average distance between two sample locations is 250m. The element size for random field realizations is designed with 40 m × 40 m, which is one hundred times larger than that of the synthetic digital V_{s30} field. The V_{s30} value is extracted from the digital soil field at each sample locations (denoted as 'measured data'), the evenly spaced sampling plan is considered in the current study.
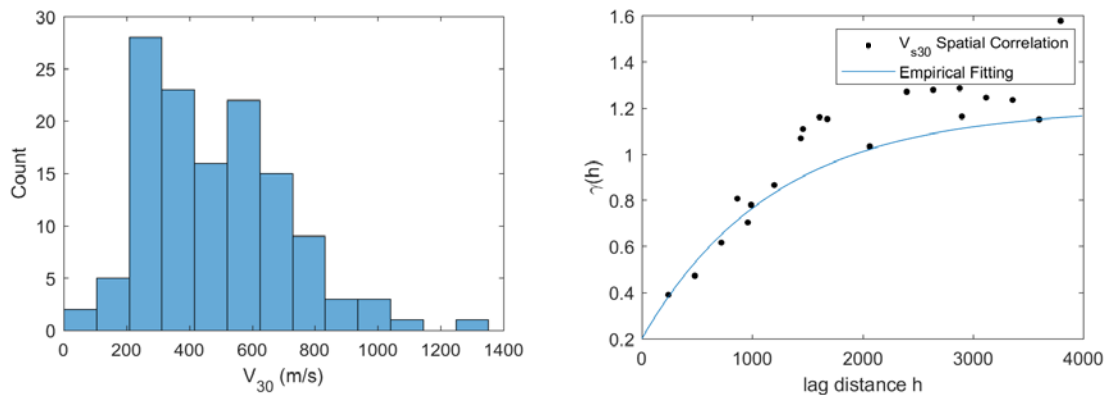
Figure 3.3 Histogram and semi-variograms of 128 measured data

The histogram and semi-variogram of the measured data are shown in Figure 3.3, and the correlation length is 3606 m. The measured V_{s30} data follows a lognormal

distribution, and the mean μ and standard deviation σ are 482.56 m/s and 224.42 m/s, respectively.

As described in Section 3.2, the geological map is considered in this random field framework as secondary information. Wills et al. (2006) and Wills et al. (2015) provide a site condition map to show simplified geologic units and corresponding V_{s30} values. In this map, the mean and standard deviation of each geologic unit are defined (Table 3.1). This map is selected as secondary information in this study. With the given characteristic of Vs30 in each geologic unit, random variable realizations of secondary can be generated and assigned to the predefined grid. As described before, Unit 1 is assumed as crystalline rock (Xtaline) and Unit 2 is assumed as quaternary alluvium (Qal2). Herein, the corresponding mean and standard deviation for two geologic units can be found in Table 3.1 as $\mu_1 = \mu_{xtaline} = 710.1\,\text{m/s}$, $\sigma_1 = \sigma_{xtaline} = 710.1\,\text{m/s}$, $\mu_1 = \mu_{Qal2} = 293.5\,\text{m/s}$ and $\sigma_2 = \sigma_{Qal2} = 73.5\,\text{m/s}$.

With the parameters for statistics and spatial correlation of the test samples, the hybrid random field-based V_{s30} models can be generated using procedures discussed before. Coupling the random field model with Monte Carlo simulations for 1000 times, the expected V_{s30} values across the field and the associated uncertainties can be obtained.

Table 3.1 Geologic units and V_{s30} characteristics. Modified according to Wills et al. 2015

Geologic units	Number of profiles	Mean (m/s)	Standard Deviation (m/s)
Qi	19	176.1	47.6
Af/qi	95	225.6	113.3
Qal1 (flat)	117	228.2	48
Qal2 (mod.)	161	293.5	73.5
Qal3 (steep)	114	351.9	112.2
Qoa	183	386.6	145.1
Qs	13	307.6	33.7
QT	17	444.0	159.7
Tsh	32	385.1	129.4
Tss	62	468.4	212.6
Tv	11	518.9	172.0
Serpentine	3	571.6	87.0
Kss	19	502.5	227.9
KJf	41	733.4	340.1
Xtaline	35	710.1	393.8

3.5 Parameter study of hybrid geostatistical and geological data-based model

Knowledge of surficial geology is an important piece of information for regional soil properties mapping as it typically provides broader area coverage and information on large-scale material heterogeneity. Following steps are used to integrate multiple sources of data into the mapping process:

1. Identify the boundaries of each geological unit within the study region.

2. Get the statistical distributions of the target parameter of each geologic units.

3. Generate random variable realizations of secondary V_{s30} values within each geological unit according to the characterized or assumed statistical distributions. Once generated, the secondary data will be kept constant for the following random field realizations.

4. Assign the generated secondary V_{s30} values to a predefined grid with a given element size, which will be integrated into the conditional sequential simulation algorithm as secondary data. For simplicity, the element size of the predefined grid is denoted as the "esize" for use in subsequent analysis.

In this section, sensitivity analyses are operated to calibrate B, integrate secondary data and test this framework.

3.5.1 Relationship between Z and B

Implementation of the Markov-Bayes hypothesis requires the selection of a calibration parameter, B. In general, when the primary and secondary variables are positively correlated, B can vary between 0 and 1; if B=0, the secondary data are ignored and if B=1, the secondary data are not updated by neighboring hard or soft data.

As described in section 3.2, a calibration procedure recommended by Deutsch and Journel (1998) is selected in this work. In this procedure, the value of B is determined as the difference between two conditional expectations which detailed in Equation (3.11), (3.12) and (3.13). Within this procedure, the selection of threshold dominates the outcome B value. A threshold is used to distinguish ranges of values where the behavior predicted by the model varies in some important way. Some indices have a specific value to distinguish soil behavior, like for liquefaction potential index (LPI), a threshold value of 5 can be chosen as it appropriately separates liquefaction-prone and non-liquefied units. Other than its physical significance, the distribution of primary and secondary data should also be taken into consideration.

Here, 120 sample locations are selected with given measured data (primary) and generated secondary data. The threshold value is selected on an interval from 0 to 1800 m/s with steps of 15. The result is shown in Figure 3.4. Grey dash lines link each pair of primary and secondary data. When threshold smaller than 200 m/s, the corresponding B value is 0, and when threshold larger than 1400m/s, the corresponding B value is 1. Therefore, the threshold value should be selected within the intersection value range of the primary and secondary data.
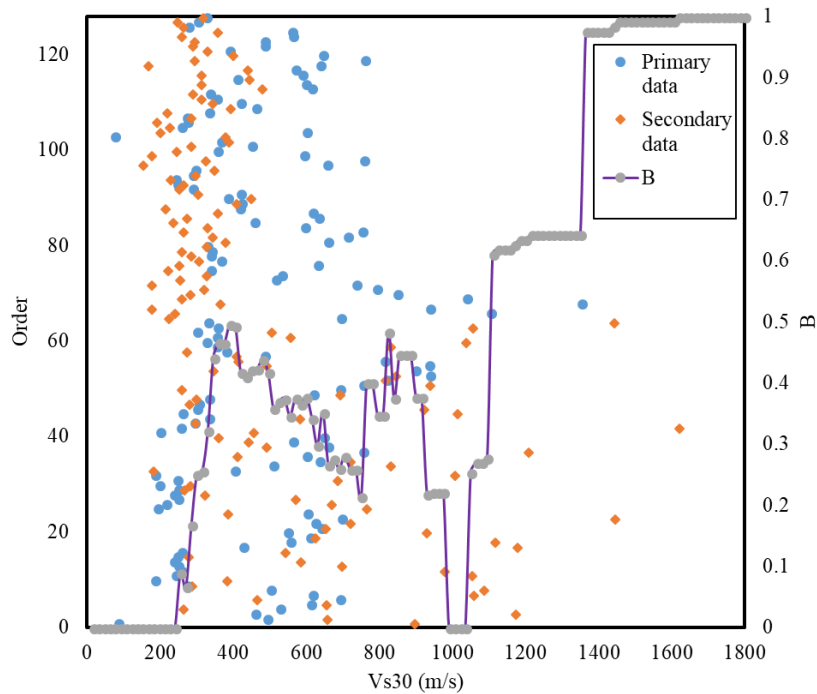
Figure 3.4. Different thresholds with its corresponding B value

3.5.2 Sufficient data & insufficient data

In the real-world field investigations, especially for shear wave velocity test, most sampling locations are clustered in some specific geologic units and few of them locate in the hilly area. Herein, one more investigation plan with all samples locate in one geologic unit is designed in this chapter to compare the model performances under the condition of sufficient and insufficient sample (Figure 3.5).
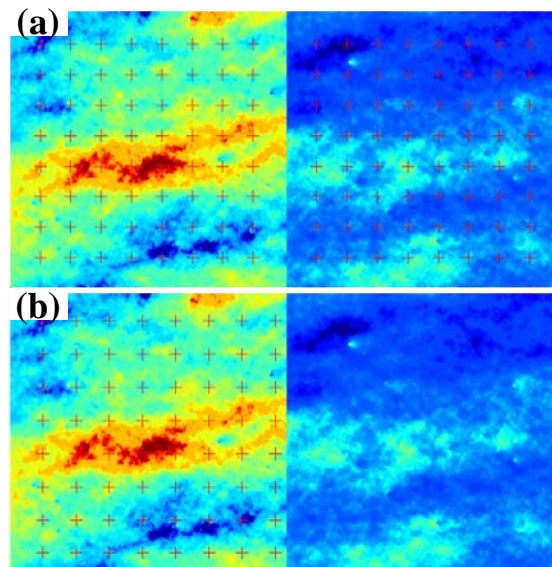
Figure 3.5. Two investigation plans with different number of sample points. (a) Plan A: 128 sample points located in the whole field evenly. (b) Plan B: 64 sample points located only in Unit 1.

Three typical scenarios with different esize and B value ((1)esize=500, B=0.2; (2) esize=400, B=0.4; (3) esize=300, B=0.6) are performed based on both two invagination plans, results are shown in Figure 3.6. It can be seen that by changing esize and B value, maps based on plan A (Figure 3.6(a), (c) and (e)) have a little change while maps based on plan B (Figure 3.6(b), (d) and (f)) have a significant change especially in Unit 2. This

shows the performance of the framework is influenced by the number of measured data significantly. Hence, in the rest section, all the analysis are based on both two investigation plans: Plan A, with 128 evenly placed sample points and Plan B, with 64 points located in Unit 1.
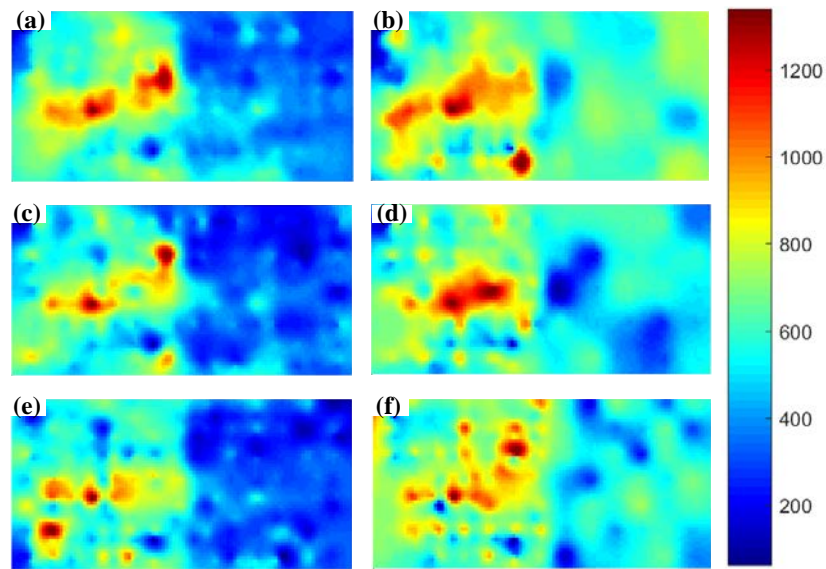
Figure 3.6. Maps of expected V_{s30} values based on two investigate plans. Left three maps are based on the investigate plan with 128 points, right three maps are based on the investigate plan with 64 points. Each map is obtained by averaging results from 1000 Monte Carlo simulations: (a) and (b) esize=500, B=0.2;(c) and (d) esize=400, B=0.4; (e) and (f) esize=300, B=0.6

3.5.3 Relationship between B and output

To investigate the influence of the Markov–Bayes coefficient B introduced in equation (3.9) and (3.10), six B values are used – namely, 0, 0.2, 0.4, 0.6, 0.8 and 1.0. For each B value, 1000 Monte Carlo simulations are performed. Figure 3.7 shows maps of

expected V_{s30} values for all six cases of Markov–Bayes coefficient B for sufficient investigate plan: Plan A. Each of the six maps is obtained by averaging results from 1000 Monte Carlo simulations.
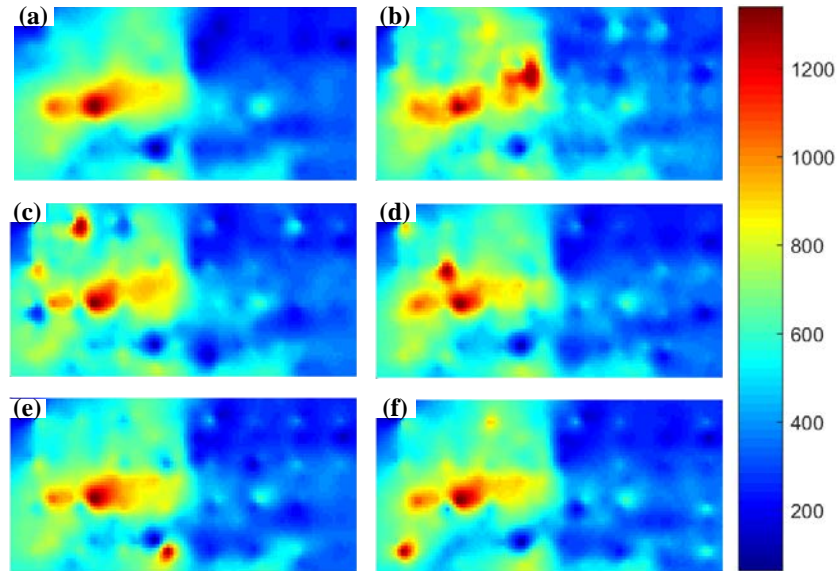
Figure 3.7. Maps of expected V_{s30} values for six cases of the Markov–Bayes coefficient B for investigate Plan A. Each map is obtained by averaging results from 1000 Monte Carlo simulations: (a) B=0; (b) B=0.2; (c) B=0.4; (d) B=0.6; (e) B=0.8; (f) B=1.0

As shown in Equations (3.9) and (3.10), the Markov–Bayes coefficient B is essentially a 'scaling' factor between the primary covariance and the secondary covariance matrices. The larger the coefficient B, the stronger influence the secondary data has on the generated V_{s30} maps. In this work, the secondary Vs30 data come from geological information. Therefore, as the value of B increases, the geological boundaries should become more distinguishable in the resulting V_{s30} maps. However, from Figure 3.7 (a) to (f), there is little change in the boundary. Sufficient amount of measured data is the main reason that causes this observation. With the sufficient measured data, random

field model can make an accurate prediction of V_{s30} at unsampled locations. With the increasing number of measured data, the accuracy improved. So in this case, the boundary can be distinguished, even there is no secondary information.

A more detailed comparison of V_{s30} values for six cases of Markov–Bayes coefficient B for investigation plan A can be seen in Figure 3.8. Each bar in Figure 3.8(a) represents the V_{s30} of soil in the whole field, and segments in the bar represent different site classes of that field. In this field, nearly 50% of the soil belongs to site class C, nearly 30% of the soil belongs to site class D. With the increasing of B value, barely change can be seen in Figure 3.8(a). But the change of mean value for each geologic units can be seen in Figure 3.8(b). For B changing from 0 to 1, the mean V_{s30} value of Unit 1 gradually approaches the mean V_{s30} value of Unit 1 in the synthetic field which equals to 636 m/s. A similar trend can be observed in Unit 2, the mean V_{s30} value decreases slightly and reaches to 328 m/s which is the mean V_{s30} value of Unit 2 in the synthetic field. Summing up the above, adding 'weight' of secondary information can make a slight distribution in improving V_{s30} mapping when there is sufficient measurement data.
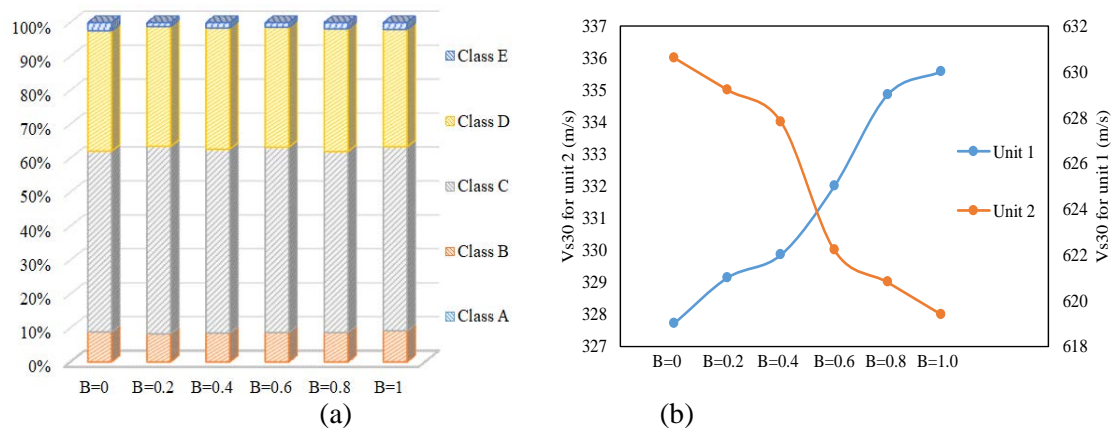
Figure 3.8. Components of V_{s30} values for six cases of Markov–Bayes coefficient B for investigate Plan A: (a) stacked column (b) mean value for two units

Figure 3.9 shows maps of expected V_{s30} values for all six cases of Markov–Bayes coefficient B for insufficient investigate plan: Plan B. Same as previous, each of the six maps is obtained by averaging results from 1000 Monte Carlo simulations. In Figure 3.9(a), no geological constraint is applied to the V_{s30} map, only 64 measured data is considered here. The realization of V_{s30} in Unit 2 holds relativity large value, which is incompatible with the synthetic field. In Figure 3.9(b) to (f), the inclusion of secondary data significant lower the V_{s30} value appears in Unit 2 and the geological boundaries become more distinguishable in the resulting V_{s30} maps.
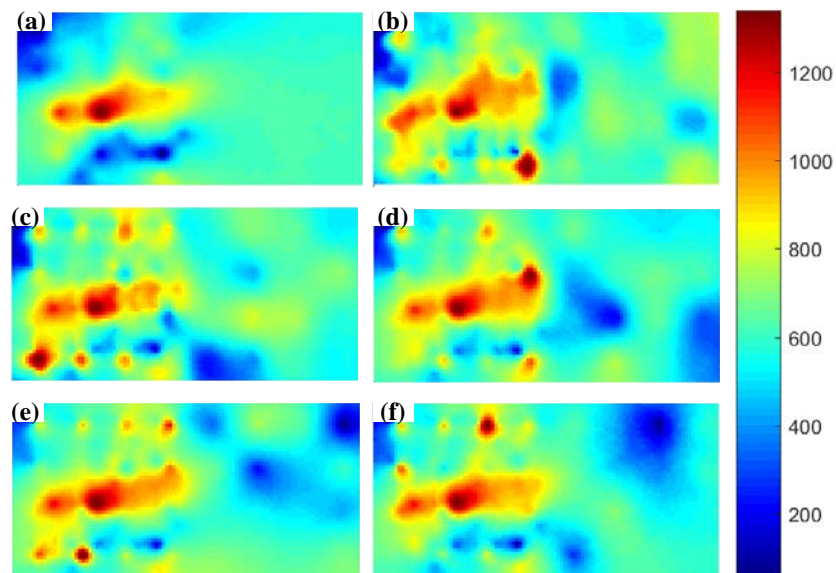
Figure 3.9. Maps of expected V_{s30} values for six cases of Markov–Bayes coefficient B for investigate Plan B. Each map is obtained by averaging results from 1000 Monte Carlo simulations: (a) B=0; (b) B=0.2; (c) B=0.4; (d) B=0.6; (e) B=0.8; (f) B=1.0

Components of V_{s30} values for six cases of Markov–Bayes coefficient B for investigation Plan B are shown in Figure 3.10. In Figure 3.10(a), nearly 80% of the soil

belongs to site class C, only 2% of the soil belongs to site class D when B value equals to zero. For B changing from 0 to 1, the percentage of soil belongs to site class D reaches 8%. In Figure 3.10(b), a similar trend in the mean V_{s30} value of two units can be found: for B changing from 0 to 1, the mean V_{s30} value decreases towards the mean V_{s30} value of Unit 2 in the synthetic field. But for the mean V_{s30} value of Unit 1, other than approaches to mean V_{s30} value of Unit 1 in synthetic field, the mean V_{s30} value excesses it and reaches to 710m/s, which is the given mean V_{s30} value for this geologic unit (Table 3.1) This result indicates, for the scenario with insufficient measured data, the secondary information may control the distribution of estimated value other than primary data.

(a)   (b)
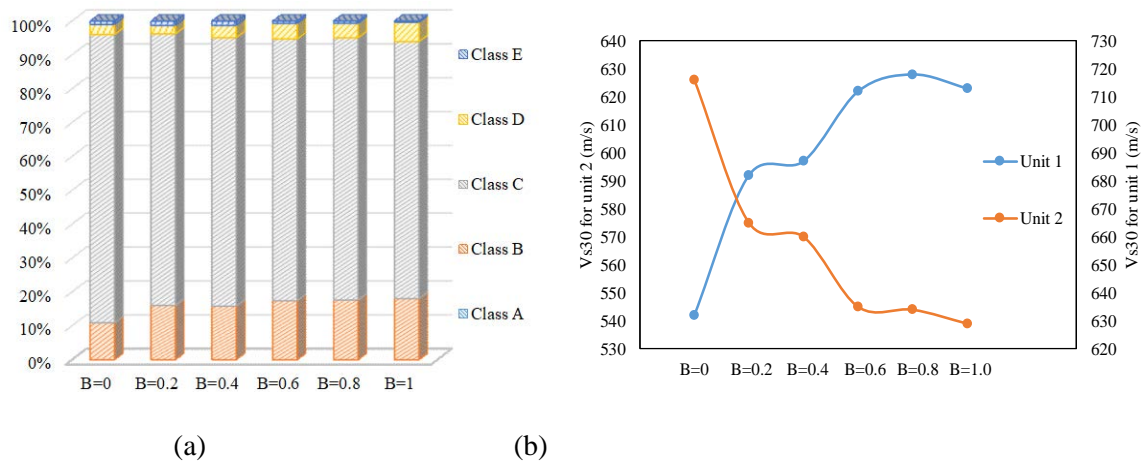
Figure 3.10. Components of V_{s30} values for six cases of Markov–Bayes coefficient B for investigate Plan B: (a) stacked column (b) mean value for two units

3.5.4 Relationship between dense of secondary data and output

To investigate the influence of the size of predefined grid for integrating secondary data, six esize values are used – namely, 500m, 400m, 300m, 250m, 200m and

100m. For each esize value, 1000 Monte Carlo simulations are performed. Figure 3.11 shows maps of expected V_{s30} values for all six cases of esize for sufficient investigate plan: Plan A. Each of the six maps is obtained by averaging results from 1000 Monte Carlo simulations. Since the size for the whole filed is 2000m×4000m, for esize changing from 500 to 100m, the number of secondary data changes from 50 to 800.

As described in the procedure of integrates the secondary data with primary data, esize controls the number of control seeds that sprinkled on the field. The smaller the esize, the more secondary data distributes in the generated V_{s30} maps. Since secondary V_{s30} data come from geological information, as the value of esize decreases, the geological boundaries should become more distinguishable in the resulting V_{s30} maps.

From Figure 3.11 (a) to (f), as the esize decreases, the geological boundaries become more distinguishable in the resulting V_{s30} maps. In Figure 3.11(a) and (d), Unit 2 turns more blue which means the estimated V_{s30} value in Unit 2 become lower. In Figure 3.11(e) and (f), the boundary between two units can be clearly seen as a straight line. But for these two cases, The V_{s30} values gradually turn randomly.
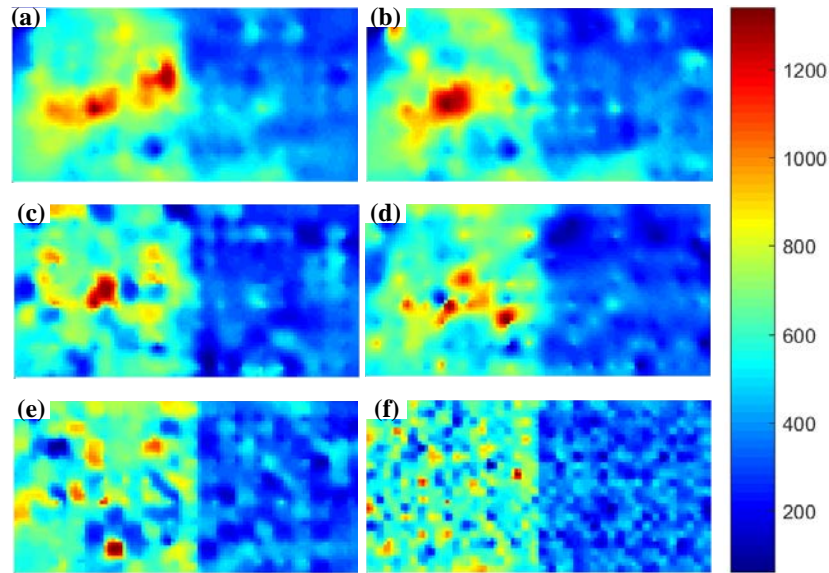
Figure 3.11. Maps of expected V_{s30} values for six cases of esize for investigate Plan A. Each map is obtained by averaging results from 1000 Monte Carlo simulations: (a) esize =500m; (b) esize =400m; (c) esize =300m; (d) esize =250m; (e) esize =200m; (f) esize =100m

In Figure 3.12(a), with the decreasing of esize, the class D soil keeps increasing till esize reaches 250 m. For esize changes from 250 to 100 m, the percentage of class D soil decreases. In Figure 3.12(b), for esize changing from 500 to 250 m, the mean V_{s30} value of Unit 1 gradual approaches to the mean V_{s30} value of Unit 1 in the synthetic field. After that, with the changing of esize from 250 to 100 m, the averaging V_{s30} decreases. A similar trend can be observed in Unit 2, for esize changing from 500 to 250 m, the mean V_{s30} value decreases slightly and reaches the mean V_{s30} value of Unit 2 in the synthetic field. For esize changing from 250 to 100 m, averaging V_{s30} increases. All these results show a inflection point at esize equal to 250 m. Meanwhile, for primary data, the average distance between any two closest points is 250m. When the average distance of secondary data larger than the average distance of primary data, primary data dominates

the simulation. If the average distance of secondary data is smaller than the average

distance of primary data, which means secondary data are more close to the unstimulated

points, the secondary data will dominate the estimation which makes the estimated V_{s30}

value random.
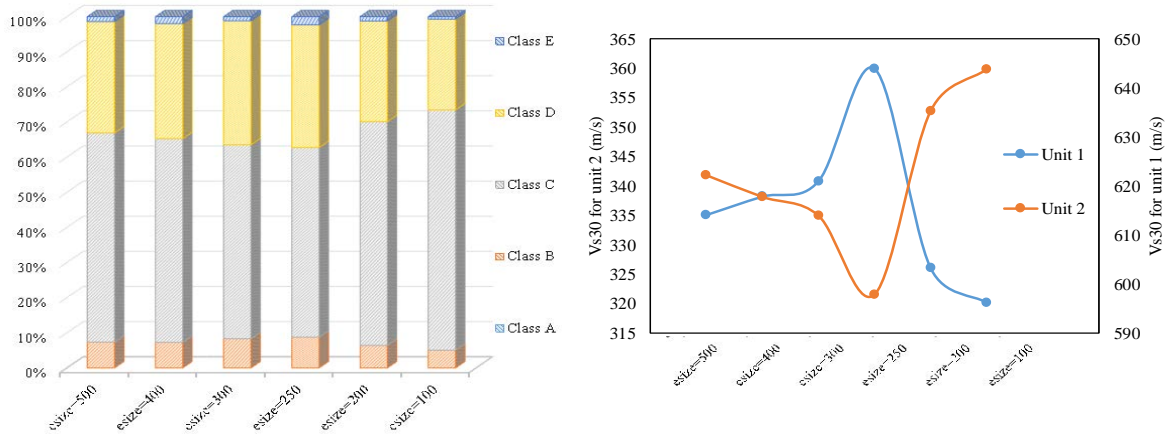
Figure 3.12. Components of V_{s30} values for six cases of esize for investigate Plan A: (a) stacked column (b) mean value for two units

For insufficient investigate plan B, the similar trend can be obtained in Figure

3.13. For changing esize from 500 to 100 m, the boundary shows more distinguishable.

More lower value appears in Unit 2. Still, for Figure 3.13(e) and (f), V_{s30} value become

randomly.
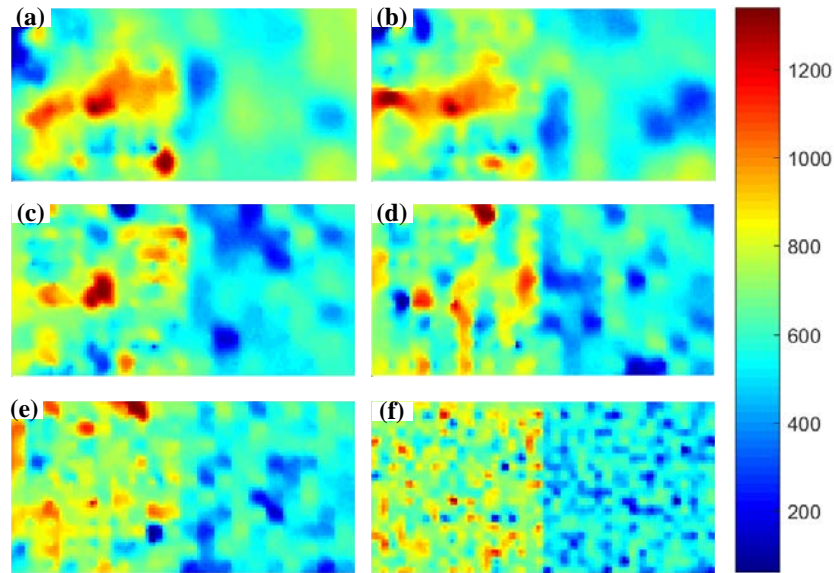
Figure 3.13. Maps of expected V_{s30} values for six cases of esize for investigate Plan B. Each map is obtained by averaging results from 1000 Monte Carlo simulations: (a) esize =500m; (b) esize =400m; (c) esize =300m; (d) esize =250m; (e) esize =200m; (f) esize =100m

In Figure 3.14(b), the mean V_{s30} value decreases towards the mean V_{s30} value of Unit 2 in the synthetic field when esize changing from 500 to 100 m. But for the mean V_{s30} value of Unit 1, the mean V_{s30} value excesses mean V_{s30} value of Unit 1 in the synthetic field and leading to the given mean V_{s30} value for geologic unit 1 (Table 3.1). In accordance with the finding in Section 3.5.3, for the scenario with insufficient measured data, the secondary information controls the distribution of estimated value other than primary data.
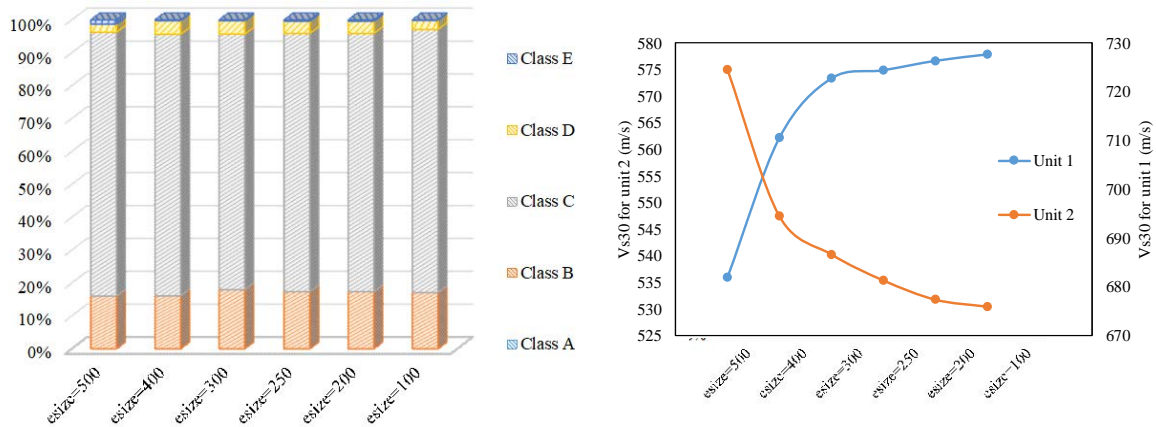
Figure 3.14. Components of V_{s30} values for six cases of esize for investigate Plan B: (a) stacked column (b) mean value for two units

<u>3.6 Suggestions in using the hybrid random field framework</u>

With the operated parameter study, the overall workflow for integrating multiple sources of data in random field model for regional soil properties mapping can be established based on a hybrid geotechnical and geological data-based framework (Figure 3.15). This overall workflow combines guideline and a detailed procedure of operating the proposed random field framework. It provides all advice on how to act in each specific mapping situation. Following the given procedure listed in Figure 3.15, the mapping can be accomplished with step-by-step instructions.
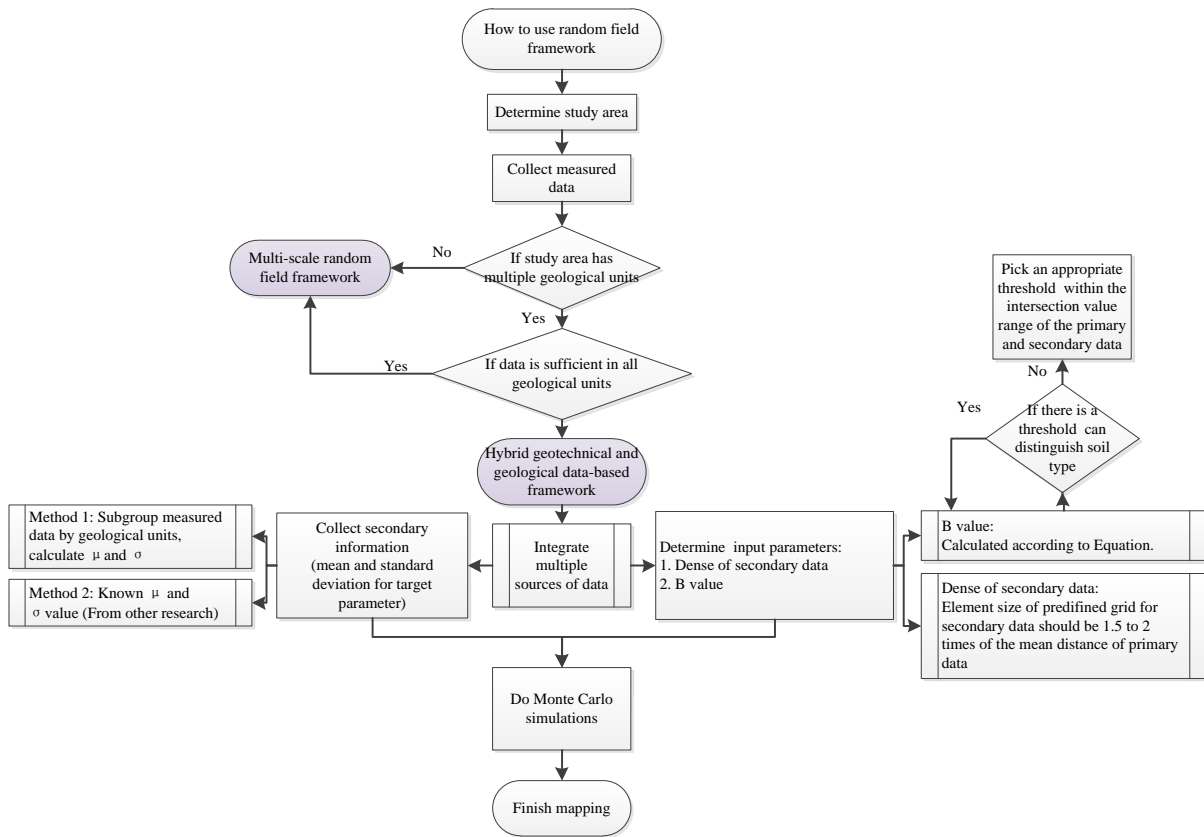
Figure 3.15. Overall workflow to integrate multiple sources of data in random field model

Summary

In this chapter, a guideline to integrate multiple sources of data in the random field model for regional soil properties mapping is established based on a hybrid geotechnical and geological data-based framework. The geotechnical data such as the V_{s30} measured data are used as the primary information, showing clear spatial correlation. Geological information is considered as secondary information that essentially enforces

geological constraints to the generated soil properties maps. Both primary and secondary data are integrated into random field models through a conditional sequential simulation technique. The integration effect caused by Markov–Bayes coefficient B and the size of the predefined grid of secondary information are emphatically discussed within two sampling conditions: sufficient measured data and insufficient measured data. A well-designed 2-D synthetic digital field is applied here to test the proposed framework. In summary, it is found that:

1. Since the secondary data represents the geologic information, as the weight of the secondary data increases, the geological boundaries become more distinguishable in the generated map.

2. It is necessary to select an appropriate threshold value for calibrating the Markov–Bayes coefficient B. Other than its physical meaning, the threshold value should be selected within the intersection value range of the primary and secondary data.

3. The element size of the predefined grid for secondary data cannot be smaller than the mean distance of any closest two primary data points. The overdense secondary data may result in the randomly estimated value.

4. Integration of secondary data can modify mapping scenario with insufficient, unevenly distributed measured data. For the mapping scenario with sufficient measured data in all geologic units, it can modify the resultant mapping slightly.

CHAPTER IV

4. UNCERTAINTY PROPAGATION IN DYNAMIC SITE EFFECT CAUSED

BY UNCERTAINTY OF SOIL PARAMETERS*

4.1 Introduction

Predicting the influence of local soil conditions on expected earthquake ground motions is a critical aspect of the seismic design process. In the past few decades, lots of research has been done to estimate site response. All these works can be divided into two main groups based on research methods, one is the empirical method which is based on statistical analysis and fitting of field data, the other one is the numerical method which is based on dynamic site response analysis.

"Next Generation of Ground-Motion Attenuation Models" for the western United States (NGA-West) program, a comprehensive multidisciplinary research project coordinated by the Pacific Earthquake Engineering Research Center (PEER), has made some major advances in seismic hazard estimation for the western US. As a typical work using the empirical model method, one main contribution of NGA-West project is that a new ground-motion prediction relationship is developed through a comprehensive and highly interactive research program. In phase 1 (known as NGA-West 1), five sets of Ground Motion Prediction Equations (GMPEs) were developed by Abrahamson and

Silva (2008), Boore and Atkinson (2008), Campbell and Bozorgnia (2008), Chiou and Youngs (2008) and Idriss (2008). In phase 2 (known as NGA-West 2), based on updated NGA-West database, Abrahamson et al. (2014), Boore et al. (2014), Campbell and Bozorgnia (2014), Chiou and Youngs (2014), and Idriss (2013, 2014) updated and superseded GMPEs which is developed in NGA-West 1 Project for both small-to-moderate and moderate-to-large magnitude database. Apart from the NGA-West program, other efforts on improving GMPEs have been made by Steward et al. (2015), Choi and Stewart (2015), Atkinson (2015).

With these GMPEs, the PGA at site surface or amplification factor can be easily obtained. But in many situations, researchers still use dynamic analysis to obtain a detailed, accurate, thorough and site-specific ground response. The non-linear dynamic behavior of soils during a seismic event has a predominant role in current site response analysis.

Dynamic seismic site response analysis represents the effects of soils conditions on ground shaking that assessed through dynamic simulations of wave propagation. This analysis propagates rock acceleration-time histories through the local soil profile to compute acceleration-time histories at the ground surface. The site response analysis provides an assessment of surface acceleration time histories, surface acceleration response spectra, amplification factors, as well as an evaluation of the induced shear stresses and shear strains within the soil profile. Lots of work has been done in dynamic seismic site response analysis like Seed and Idriss (1969), Park and Hashash (2008), Kamalian et al. (2006), Martin et al. (1982) and Elgamal et al. (2002). Constitutive

models are at the core of dynamic site response analysis. A broad range of simplified (e.g., Ramberg and Osgood, 1943; Duncan and Chang 1970; Hardin and Drnevich, 1972a&b; Pyke, 1979; Martin and Seed, 1982) and advanced soil constitutive models (Prevost, 1977, Li et al., 1997, Borja and Amies 1994) have been proposed in non-linear site response analysis (Hashash et al., 2010). Some advanced constitutive models are able to capture important features of soil behavior such as anisotropy, generation of excess pore pressure, and dilation (Prevost, 1977, Li et al., 1997, Borja and Amies 1994). However, advanced constitutive models typically require more detailed information on soil behavior for parameter calibration. For many engineering applications, the only soil behavior information available is usually the modulus reduction and damping curves. To this end, simplified models, especially those that belong to the family of hyperbolic models, are often used and will be adopted in this work.

Within dynamic site response, it has been recognized that the severity of ground motion at a site significantly depends on the soil characteristics of the layers below the surface (Campbell, 1979; Toro, 1993). At most sites, however, the soil profiles and the parameters that are responsible for the site dynamic response are not known with certainty. The lack of in-situ geotechnical investigation data, the effect of sample disturbance and scaling associated with laboratory tests and the natural heterogeneity of soil profiles are notable sources of uncertainty in nonlinear site response analysis. Herein, the effect of soil parameter uncertainty on the prediction of site response needs to be examined. Several studies have been published on this topic: Li and Assimaki (2010) investigated the effect of 1D spatial variability of shear wave velocity, material damping,

modulus reduction and input motion using randomized realizations of a based soil profile. Bazzurro and Cornell (2004a) used SUMDES, a nonlinear procedure for response analysis to study the ground response variability due to uncertainty in the ground motion site condition and show that spectral acceleration at the bedrock of the input record is the single most helpful parameter for the prediction of amplification factor at the same oscillator frequency. Bazzurro and Cornell (2004b) presented effective probabilistic procedures for evaluating ground-motion hazard at the free-field surface of soil deposit. Wang and Hao (2003) considered the effect of groundwater level during site response analysis. Bahrampouri et al. (2018), Stewart and Kwok (2008) studied the effect of uncertainty in modulus reduction and damping curves on the uncertainty in the amplification factors. Tombari and Stefanini (2017) used a fuzzy set method to simulate uncertainty of input parameters include shear elastic modulus, unit density, damping ratio and thickness of the soil deposit.

In this chapter, a comprehensive study is conducted to evaluate the impacts of soil parameters uncertainty in nonlinear dynamic site response. Extensive geotechnical data on variability statistics of soil at three downhole sites with different site classes are used in this study. Also, Davidenkov model with simplified loading-reloading rules is used to describe the stress-strain relationship under irregular cyclic loading. For each of these three sites, the peak ground acceleration (PGA) at surface subjected to a historical seismic wave motion is estimated while considering the uncertainty of soil parameters. The values of PGA, are obtained by driving a real rock ground motion through different representations of the soil model. Each representation is characterized by a different but

plausible combination of soil parameter values generated by Monte Carlo Simulation. To obtain a reliable estimation of soil response, a finite element computer program capable of adding subroutine on user-defined materials is used to perform the analysis in the time domain.

The structure of this paper is as follows: Firstly, numerical analysis with finite element method is introduced, including constitutive model and numerical implementation in ABAQUS VUMAT. Then, uncertainties of soil parameters are quantified and modeled. Thereafter, the model for numerical analysis is described and three study sites with different site condition are simulated. Finally, the results from the site response analyses are then used to discuss the uncertainty propagation in PGA caused by shear wave velocity.

4.2 The modified Davidenkov constitutive model

In this section, the modified Davidenkov constitutive model is introduced to model the non-linear soil behavior under irregular cyclic loadings. The model is then implemented as a user-defined material subroutine VUMAT in ABAQUS/Explicit, where details of the numerical implementation algorithm are presented.

4.2.1 Model formulation

The stress-strain curves of the modified Davidenkov model under irregular cyclic loadings are illustrated in Figure 4.1. This hyperbolic-type model can be described using two sets of equations. The first set of equations defines the stress-strain relationship for loading and is shown as the backbone curve (also called the skeleton curve) in Figure 4.1. The second set of equations defines the stress-strain relationship for unloading and reloading (e.g., loading path 1-2-1, 5-6-7 in Figure 4.1)
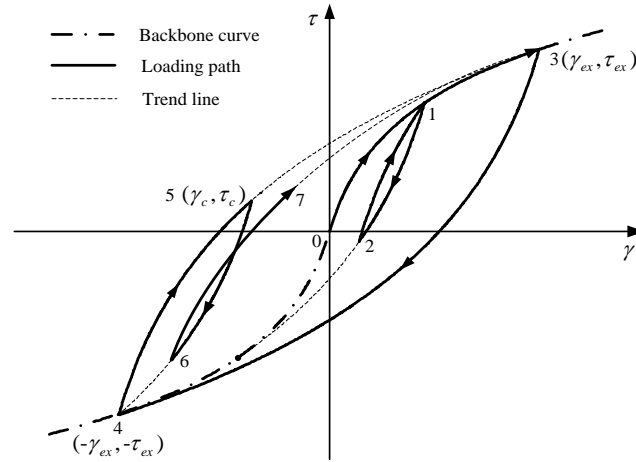
Figure 4.1 Stress-strain curves of the modified Davidenkov model under irregular cyclic loadings

The backbone curve of the modified Davidenkov model is constructed using the following set of equations (Martin and Seed, 1982)

$$\tau = G \cdot \gamma = G_{max} \cdot \gamma \cdot \left[1 - H(\gamma) \right] \tag{4.1}$$

$$H(\gamma) = \left\{ \frac{\left(\gamma/\gamma_0 \right)^{2B}}{1 + \left(\gamma/\gamma_0 \right)^{2B}} \right\}^{2A} \tag{4.2}$$

where τ and γ are the shear stress and shear strain, respectively; G_{max} is the initial shear modulus; A, B and γ_0 are fitting parameters.

For unloading and reloading (before reaching the previous loading reversal point), the following variation of the original Davidenkov model is adopted (Pyke 1979)

$$\tau = \tau_c + G_{max} \cdot (\gamma - \gamma_c) \cdot \left[1 - H\left(\frac{|\gamma - \gamma_c|}{2n} \right) \right] \tag{4.3}$$

where γ_c is the shear strain at the point of loading reversal; n is a scaling factor of the initial loading curve. In this work, the "n-times method" by Pyke (1979) is adopted instead of the original factor of two (Martin 1974). Once reloading passes the previous reversal point (e.g., loading path 1-3 in Figure 4.1), the stress-strain curve will follow the original backbone curve.

With both sets of equations defined, the stress-strain curve of a soil under irregular cyclic loadings follows the extended Masing rule (Martin and Seed, 1982; Pyke, 1979):

(1) For initial loading (e.g., path 0-1 in Figure 4.1), the loading path follows the backbone curve defined by Equations (4.1) and (4.2).

(2) If a loading reversal occurs at a point defined by (γ_c, τ_c), subsequent stress-strain curve (e.g., path 1-2) should move down along the path defined by Equation (4.3). The path extends from the loading reversal point to the historical maximum (minimum) point (e.g., the trend line 2-4).

(3) For reloading (e.g., path 2-1-3), if the reloading curve passes the maximum past strain (e.g., point 1) before unloading, it will follow the backbone curve until the next stress

reversal.

With these three rules, the stress-strain curve under irregular cyclic loading can be fully defined with stress and strain values at the loading reversal (τ_c, γ_c), stress and strain values at the loading reversal at the historical maximum point (τ_{ex}, γ_{ex}), and some constant model parameters (G_{max}, A, B, γ_0), minimizing the information the model needs to "memorize". This is a great advantage when it comes to numerical implementation, which will be detailed in the next section.

4.2.2 Numerical implementation in ABAQUS VUMAT

In this section, the numerical integration and implementation of the modified Davidenkov model as a user-defined material subroutine VUMAT in ABAQUS/Explicit are presented. ABAQUS/Explicit is a robust and well-tested commercial finite element code capable of three-dimensional linear and nonlinear dynamic analysis. The unique feature of ABAQUS/Explicit is that it allows the integration of user-defined material models through the VUMAT interface, which makes it a particularly attractive tool for this research.

Given the values of stress and internal state variables at time t_n, and the strain increment $\Delta\gamma$, the objective of the numerical integration is to find the stress and state variable values at time t_{n+1}. In the modified Davidenkov model, the shear modulus G has a non-linear dependency on strain and accounts for the hysteresis behavior of the soil. The key in the integration is therefore to obtain shear modulus update. Once the modulus is updated, the value of stress can be easily calculated using Equation (4.1).

When the loading path follows the backbone curve, the shear modulus at time t_{n+1} can be obtained by taking the derivative of stress τ with respect to strain γ per Equation (4.1) as

$$G_{n+1} = \frac{d\tau}{d\gamma} = G_{max}\left[1-\left(1+\frac{2AB\gamma_0^{2B}}{\gamma_0^{2B}+\gamma^{2B}}\right)\cdot H(\gamma)\right]$$ (4.4)

For unloading and reloading (before reaching the previous loading reversal point), the shear modulus can be obtained by taking the derivative of stress in Equation (4.3) with respect to γ-γ_c as

$$G_{n+1} = \frac{d(\tau-\tau_c)}{d(\gamma-\gamma_c)} = G_{max}\left\{1-\left[1+\frac{2AB(2n\gamma_0)^{2B}}{(2n\gamma_0)^{2B}+|\gamma-\gamma_c|^{2B}}\right]H\left(\frac{|\gamma-\gamma_c|}{2n}\right)\right\}$$ (4.5)

where $(2n\gamma_0)^{2B}$ is determined by both current point of loading reversal and historical maximum point as

$$(2n\gamma_0)^{2B} = (\gamma_{ex}\pm\gamma_c)^{2B}\cdot\left(\frac{1-R}{R}\right)$$ (4.6)

By substituting the historical maximum point (τ_{ex}, γ_{ex}) into Equation (4.3), R can be expressed as

$$R = \left(\frac{\tau_{ex}\pm\tau_c}{G_{max}\cdot(\gamma_{ex}\pm\gamma_c)}\right)^{\frac{1}{A}}$$ (4.7)

The "$\pm$" in this equation takes "-" during the loading process and takes "+" during the unloading process.

This scheme is advantageous in the sense that it only needs five state variables, i.e., the strain of the previous time step (γ_n), the strain and stress values at the historical maximum point ($\tau_{ex(n)}$, $\gamma_{ex(n)}$), and the strain and stress values at the current point of

75

loading reversal ($\tau_{c(n)}$, $\gamma_{c(n)}$). The numerical integration algorithm is summarized in Algorithm 1.

Algorithm 1: Explicit integration algorithm

Input: state parameters γ_n, $\gamma_{ex(n)}$, $\tau_{ex(n)}$, $\gamma_{c(n)}$, $\tau_{c(n)}$

Incremental strain and constants: $\Delta\gamma_{(n+1)}$, G_{max}, A, B, γ_0 , β

Output: $\gamma_{(n+1)}$, $\tau_{(n+1)}$

Step 1: Update current strain

$\gamma_{n+1} = \gamma_0 + \Delta\gamma_{n+1}$

Step 2: Update historical maximum point

If $\gamma_n \geq \gamma_{ex(n)}$ && $\gamma_{n+1} < \gamma_n$ then

$\gamma_{ex(n+1)} = \gamma_n$

$\tau_{ex(n+1)} = \tau\left(\gamma_{ex(n+1)}\right)$ Equation (4.1) and (4.2)

End if

Step 3: Update inflexion point

If $\gamma_{n+1} < \gamma_{ex(n)}$ && $\gamma_{n+1} > \gamma_n$ then

$\gamma_{c(n+1)} = \gamma_n$

$\tau_{c(n+1)} = \tau\left(\gamma_{c(n+1)}\right)$ Equation (4.2) and (4.3)

End if

Step 4: Update shear modulus G

If $\gamma_{n+1} \geq \gamma_{ex(n)}$

$G = G(\gamma_{n+1})$ Equation (4.4)

Else

$G = G\left(\gamma_{n+1}, \gamma_{ex(n)}, \tau_{ex(n)}, \gamma_{c(n)}, \tau_{c(n)}\right)$ Equation (4.5), (4.6) and (4.7)

End if

Step 5: Update stress

$\tau_{n+1} = G \cdot \gamma_{n+1}$

An additional viscous mechanism is usually available and exploited in most finite element (FE) codes (Pisanò and Jeremić, 2014). The viscous component implies smoother cycles and avoids the sharp transitions at stress reversal. Ronaldo I. B. et al. (2000) point out that incorporating the viscous effect provides for a smoother transition

of the stress-strain curve on a reverse loading. In contrast, the inviscous response shows a

sharp corner at the unloading on reverse loading.

The form for viscous component σ^v is:

$$\sigma^v = \beta D^{el} \dot{\varepsilon} \tag{4.8}$$

$$\beta = {2\xi_0}\Big/{\omega_1} \tag{4.9}$$

where $\dot{\varepsilon}$ is the strain rate; D^{el} is the elastic matrix; β is a scaling coefficient; ω_1 is the first

order frequency of structure; and ξ_0 is the value of damping ratio in the limit zero shear

strain. It is assumed that the energy in soil gets dissipated even in the elastic domain.

Finally, the total stress can be obtained by summing two components as

$$\sigma = \sigma^f + \sigma^v \tag{4.10}$$

where σ^f is the stress calculated from Algorithm 1 and σ^v is the viscous force defined in

Equation (4.8).

Algorithm 2:
Extra viscous damping algorithm viscous component
Input: Gmax, ν, εn+1
Output: σ n+1
Step 1: Calculate σv according to elastic theory
$\sigma^v_{n+1} = \sigma\left(G_{max},\ \nu,\ \varepsilon_{n+1}\right)$ Equation (4.8)
Step 2: Calculate total stress
$\sigma_{n+1} = \sigma^f_{n+1} + \sigma^v_{n+1}$ Equation (4.10)

4.3 Sensitivity analysis and shear-wave velocity modeling

The modified Davidenkov model has six parameters, i.e., G_{max}, ν, A, B, γ_0 and β. In this section, sensitivity analysis is first performed to understand the relative importance of these parameters in terms of their impact on amplifying seismic motions. Then, the most significant parameter is selected for subsequent uncertainty quantifications and case studies. A shear-wave velocity model is presented which accounts for the spatial variability shear-wave velocity along the depth.

4.3.1 Sensitivity analysis of model parameters

As previously mentioned in Section 4.2, there are six parameters in the modified Davidenkov model, i.e., G_{max}, ν, A, B, γ_0 and β. Among them, the initial shear modulus G_{max} is related to the shear-wave velocity (V_s) and the unit weight of a soil. In this study, the unit weight of soil is treated as a constant while the shear-wave velocity is considered as an independent model parameter, from which the initial shear modulus G_{max} can be derived. The Poisson's ratio ν is assumed to be a constant. The fitting parameters A, B, γ_0 can be obtained from modulus reduction and damping curves and the range of their values are set based on recommendations from Chen et al. (2005). The scaling coefficient β for the additional viscous component is used to control high-frequency oscillations (detailed in Section 2.2) and its value is set as zero here. The range and interval of the model parameters are listed in Table 4.1.

For sensitivity analysis, a homogeneous soil model with 4 meters high and 4 meters wide is used here. The displacements of the lateral and bottom sides are fixed along the vertical direction, the ground motion is added at the bottom in the horizontal direction. Four-node bilinear plane strain element type (CPE4R) is assigned to each element. With given input seismic motion, the maximum acceleration at top of the model can be obtained. This simplified homogeneous quadrate model is computationally efficient in single factor sensitivity analysis.

Table 4.1 Range of model parameters in sensitivity analysis

	V_s (m/s)	A	B	γ_0	
Maximum	760	1.2	0.45	9×10^{-3}	9×10^{-4}
Minimum	180	0.9	0.35	1×10^{-3}	1×10^{-4}
Interval	20	0.05	0.02	1×10^{-3}	1×10^{-4}

A motion record from Coyote Lake earthquake, August 6th, 1979 is selected as input motion in this study (obtained from PEER Strong Motion Database: https://ngawest2.berkeley.edu/). This seismic motion is recorded 17.2 km far away from fault rupture, with a magnitude of 5.7 and a maximum acceleration of 1.23 m/s². The acceleration time history of the Coyote Lake earthquake is shown in Figure 4.2.
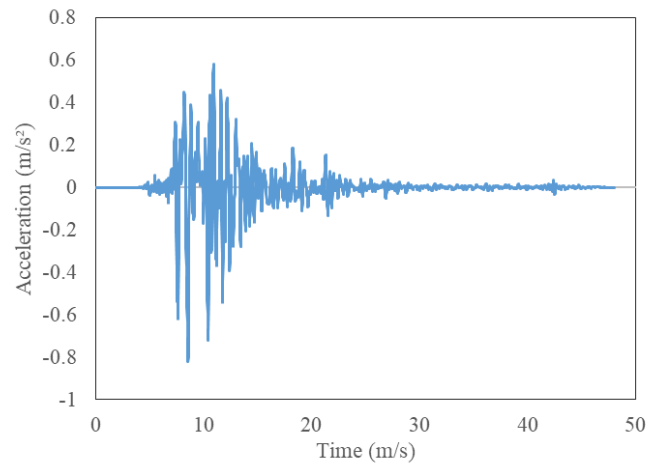
Figure 4.2. Acceleration time history of the Coyote Lake earthquake record

Each parameter changes with a given range and step (e.g. V_s changes from 180 m/s to 760 m/s with an increment of 20 m/s). For each scenario, only one parameter changes its value, other parameters are kept constant. Scenarios with same changed parameter are combined as one case. Since four parameters are tested here, all scenarios are divided into 4 cases, Case A, Case B, Case γ_0, and Case V_s. The calculated PGA values with 4 cases are shown in Figure 4.3. It indicates that V_s caused a huge influence especially for a relatively soft soil, other parameters' influence can be ignored compared with the effect caused by V_s. Hence, V_s is selected as the random variable in the further analysis. It can be much more effective, pertinent and time-saving in further analysis with one significant variable comparing to six variables.
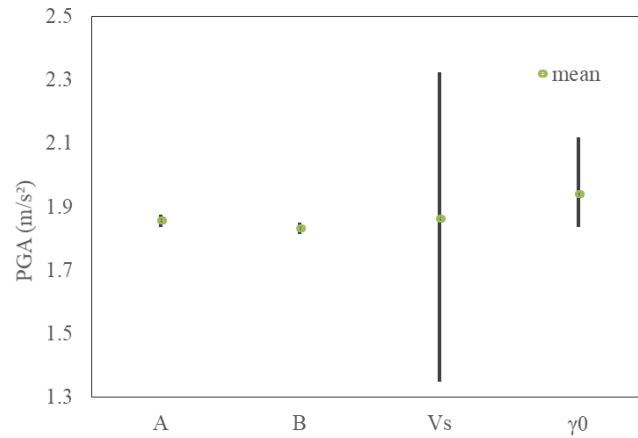
Figure 4.3. Range of PGA caused by variation of each parameter

4.3.2 The expression form of shear wave velocity

In accordance with the sensitivity analysis results in the previous section, the importance of shear wave velocity in the dynamic analysis has been acknowledged for a long time. Boaga et al. (2011) point out in the case of seismic ground motion scenarios, shear wave velocity models are the primary input for the computation of site ground response amplification. Rathje et al. (2010) suggest engineers must put thought into the selection of the interlayer correlation coefficient and the standard deviation of the shear-wave velocity.

Also, in the statistical method, shear wave velocity is considered as an important parameter to describe soil properties in estimating amplification factor. But other than V_s profile, they used a simplified expression V_{s30} to describe the shear wave velocity. Choi and Stewart (2005) develop empirical relationships to predict nonlinear amplification factors for 5% damped response spectral acceleration as a continuous function of average

shear wave velocity in the upper 30m. Also, Boore and Atkinson (2008), Campbell and Bozorgnia (2008) and Abrahamson and Silva (2008) use V_{s30} as an important soil properties parameter to establish empirical ground-motion models.

However, as the time-averaged shear wave velocity in the upper 30 m, V_{s30} is a simplified expression of shear wave velocity that only considers the soil properties in the upper 30 m. We expand the 4 m soil column model in previous sensitivity analysis into 30 m and 100 m to test how soil layers below 30 meters affect the site response. For each model, a homogeneous soil is assumed with shear wave velocity equals to 155m/s, 266m/s, 489m/s and 913m/s which are typical values for NEHRP site class E, D, C and B. Model with 30 m depth simulates the case only considering top 30 m. In this case, as shown in Figure 4.4, the PGA value increases first then decreases when the soil becomes harder. This trend is similar to the trend shown in Stewart and Seyhan (2013). But for the model with 4 m and 100m depth, the trends are different. These three different trends indicate the depth of soil column has a significant influence on the site response. Using V_{s30} is not accurate enough for estimating PGA at the ground surface. The soil layers below 30 meters have a contribution in seismic wave prorogation from rock to the surface.
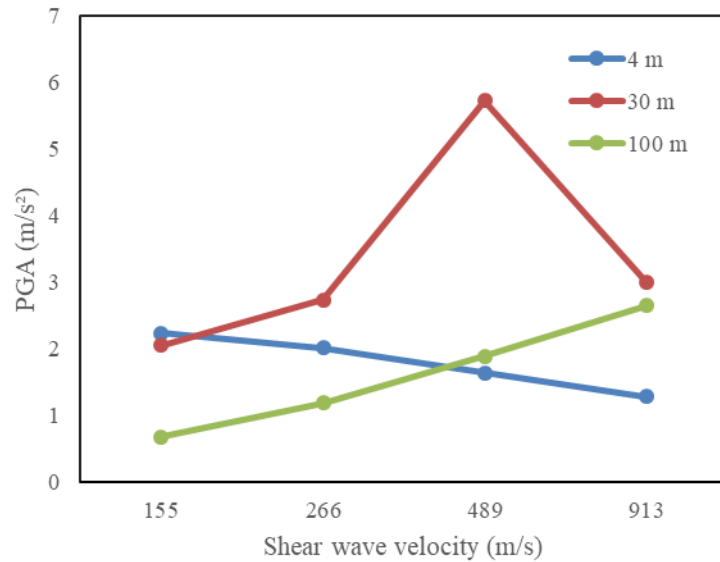
Figure 4.4. Resultant PGA trend of 4 types of Site Classes with 3 different model sizes

Above all, the shear wave velocity (V_s) profile is the most critical parameter influencing the site response and the effect of the variability in V_s can be taken into account by performing 1D analyses for multiple realizations of the V_s profile.

4.3.3 Shear wave velocity uncertainty

As a selected random variable, shear wave velocity needs to be characterized with its variability and distribution.

Toro (1993) velocity model is adopted here to describe the variability of V_s within each layer and its correlation with adjacent layers. Toro (1993) studied the probability distribution of $\ln(V_s)$ using the cumulative distribution of standardized variables shown in Equation (4.11) for generic soil profiles

$$Z_i = \frac{\ln(v_i) - \ln(v_{m,i})}{\sigma_{\ln v}} \tag{4.11}$$

where V_i is the velocity at the midpoint of layer i; $V_{m,i}$ is the median velocity of the same layer; $\sigma_{\ln v}$ is the standard deviation of $\ln(V_s)$ for all layers; Z_i represents the number of standard deviations from the mean value of $\ln(v_{m,i})$ in log space.

The lognormal distribution of V_s and the V_s layer to layer correlation based on an analysis of over 500 shear-wave velocity profiles at the study site are characterized using a first-order autoregressive model as Equation (4.12)

$$Z_i = \begin{cases} \varepsilon_i & (i=1) \\ \rho Z_{i-1} + \sqrt{1-\rho^2}\varepsilon_i & (i>1) \end{cases} \tag{4.12}$$

where Z_{i-1} is the standard normal variable of the previous layer; ρ is the interlayer correlation coefficient; ε_i are independent normal random variables with zero mean and unit standard deviation. The standard normal variable for the surface layer (Z_1, i=1) is independent of all other layers and Z_i is correlated with the layer above it. The parameters ρ and $\sigma_{\ln v}$, and the median V_s profile define completely the probabilistic velocity model. Toro (1993) estimated the parameters ρ and $\sigma_{\ln v}$ using data from generic soil profiles via linear regression as $\rho=0.577$ and $\sigma_{\ln v}=0.39$ (corresponding to a velocity coefficient of variation COV = 41%), and those values are adopted in this study. Note that if more detailed geotechnical data are available at the site, the parameters ρ and $\sigma_{\ln v}$ of the lognormal distribution will be calibrated based on site-specific information, and the V_s stochastic model will be characterized by a lower COV.

4.4 Study site and analyses performed

In this case study, multiple 1D profile realizations are generated by statistically varying site V_s profiles using Monte Carlo simulation. Then, all these realizations are used as input for dynamic site response analysis. A fully nonlinear soil model (modified Davidenkov model) in the site response simulations with realistic statistical descriptions of the soil properties is implemented. Besides, 3 base cases with different site classes are selected in order to obtained the site response uncertainty under different site conditions.

4.4.1 Deterministic model for site response

Using ABAQUS/Explicit, a soil column of 152m was modeled by 100 layers with 1.52 m thickness each. Four-node bilinear plane strain element type (CPE4R) is selected here. The displacements of the lateral and bottom sides are fixed along the vertical direction, ground motion is added at the bottom in the horizontal direction. The FE model is shown in Figure 4.5. Coyote Lake earthquake is selected as the input motion (Figure 4.2). Detailed information about the model is listed in Table 4.2.

Table 4.2. Model information and values for soil parameters used in this work

Model & Element		Soil parameters	
Soil column size (m)	152×10	Density (kg/m^3)	1980
Element size (m)	1.52×1	Poisson ratio	0.49
Number of layers	100	A	1.1
Input motion	Coyote Lake Earthquake	B	0.78
Element type	CPE4R	γ_0	change with depth
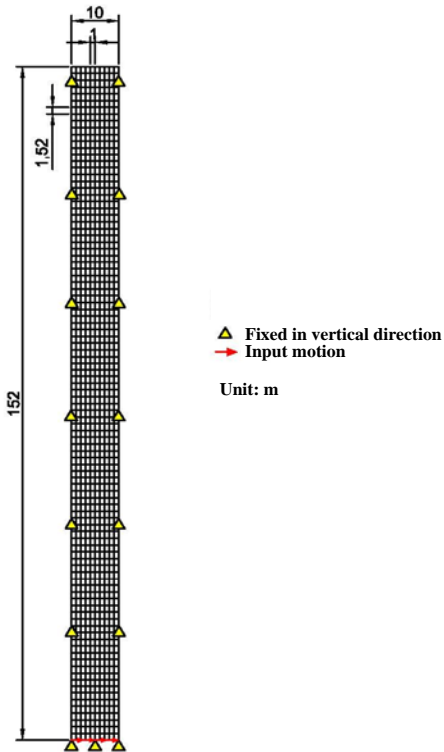
Figure 4.5. Model size and boundary condition

4.4.2 Example: Stanford area (Site class D)

With the input uncertainty model and deterministic finite element model introduced before, a whole procedure of uncertainty analysis of a class D site is detailed described in this section. Using Toro's (1993) probabilistic velocity model, random realizations of the V_s fields are generated using one standard truncation of the lognormal distribution to eliminate potential outliners. Realizations of V_s random profile at Stanford site are shown in Figure 4.6, where the thick black line corresponds to the median profile

at the site, and the thin chromatic lines correspond to 1000 random realizations of V_s profile.
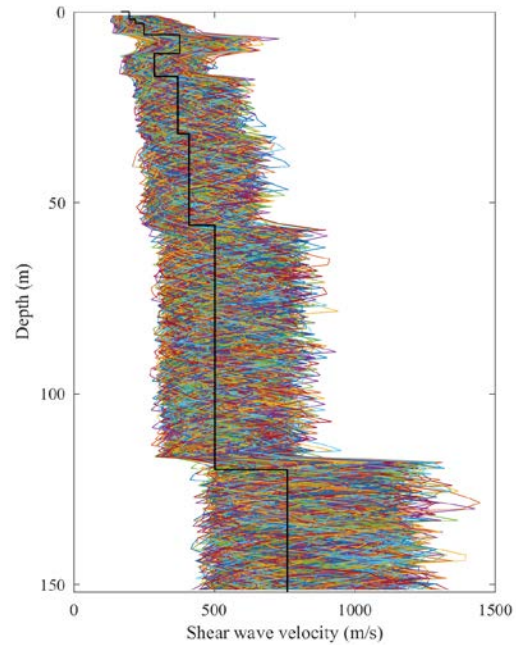
Figure 4.6. Sample realizations of V_s random profile at the Stanford site

Monte Carlo Simulation is implemented to evaluate the effect of soil parameter uncertainty on the site response variability. For the ground response variability, PGA at site surface is set as the target parameter.
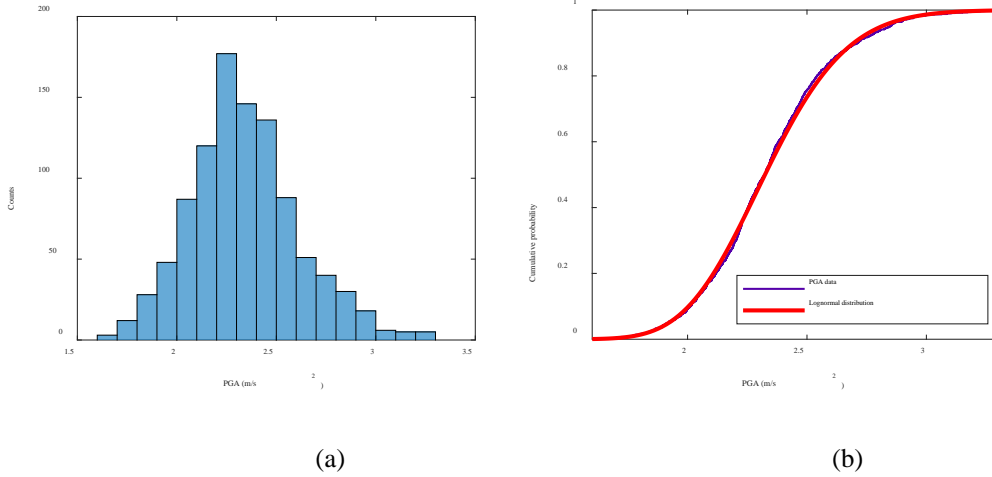
Figure 4.7. Variability in PGA caused by uncertainties in soil parameters for a strong seismic excitation. (a) Histogram for PGA with 1000 times realizations; (b) Cumulative distribution function (CDF) of PGA.
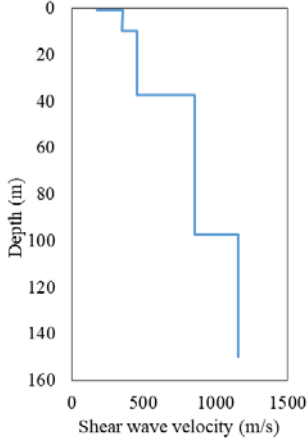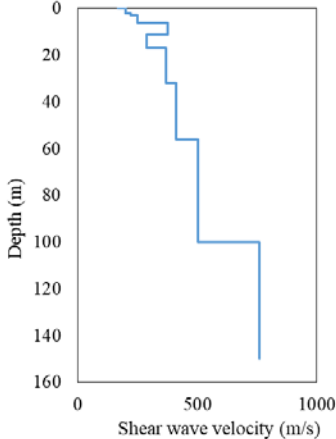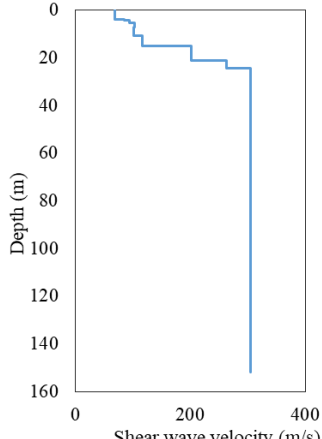
Figure 4.7 shows the statistical distribution of PGA values from 1000 time realizations. As shown in Figure 4.7 (b), the PGA distribution is approximately lognormal, with the exception of a small deviation at the tail of the distribution. The mean value of resultant PGA is 2.43 m/s^2 and the standard deviation is 0.27 m/s^2.

4.4.3 Uncertainty propagation of PGA under different site classes

Besides Stanford case, two more calculations with different site classes are done for seeking the PGA distribution under different site condition. Basic information for three sites is listed in Table 4.3. Three typical sites for site class C, D and E are selected, which corresponding to the stiff soil, relative stiff soil and soft soil. During the

calculation, the same model size as 152m×10 m is applied to reduce the difference caused

by the depth of the soil.

Table 4.3. Information of three sites

ID	1	2	3
NEHRP site class	C	D	E
V_{s30} (m/s)	414	315	134
City	Salt Lake Valley	Stanford	Hilo
State	UT	CA	HI
Vs profile			

Based on these three sites, MCSs for 3000 times realizations are conducted to

evaluate the uncertainty propagation in ground motion caused by uncertain soil. The PGA

distributions for three sites are shown in Figure 4.8. The base case is shown as red solid

line.

As expected, all three distributions of resultant PGA (Figure 4.8(a) to (c)) are

approximately lognormal which corresponding to the results in Section 4.2. Figure 4.8(d)

shows the scatter plot of the resultant PGA and corresponding V_{s30} value. An obvious trend manifested in Figure 4.8(d) is that relatively higher PGA value occurs in the Class E site, relatively lower PGA value is more common in the Class C site. This indicates that an increasing shear wave velocity brings about an increasing resultant PGA value. This trend is different from Stewart (2013) and Seyhan and Stewart (2014) which suggest the PGA will increases first and then decreases with the increasing of V_{s30}. Several reasons cause this discrepancy. First of all, unlike the empirical equation proposed by Seyhan and Stewart (2014) which use V_{s30} as a parameter to describe the soil property, V_s profile is used to describe the real soil property of a specific site in this finite element analysis which considers variabilities along the depth. Secondly, instead of upper 30-meters, the depth of the soil model down to the bedrock is considered here. Thus, deep-seated soil (beneath upper 30m) can make a contribution to the site response analysis, which is consistent with the reality. Thirdly, Soil is made up of distinct horizontal layers which have different properties and characteristics from the adjacent layers above and below. The combination of distinct horizontal layers makes a significant contribution in site response, which is not included in the empirical equation. These three reasons have been verified in Section 3.1.

One of the advantages of the Monte Carlo simulation is its ability to characterize uncertainties. To quantify uncertainties, coefficient of variation (COV) and standard deviation from 3000 independent Monte Carlo simulations are calculated for each site. The mean value of PGA keeps increasing from 0.7 to 3.31. The standard deviation is also increasing from 0.13 to 0.44 for site class E to C. This indicates the stiff soil may cause a

relatively higher standard deviation than soft soil. The corresponding coefficient of variation (COV) for site class E, D and C are 0.19, 0.12 and 0.13 respectively. It is interesting to note that the uncertainties associated with soft soil (Class E) are higher compared to the stiff soil (Class D and C), no obvious difference in between Site class D and C. Variability in soil properties significantly increases the standard deviation of the PGA but has a lesser effect on the COV. A similar trend is also found in Rathje et al. (2010).
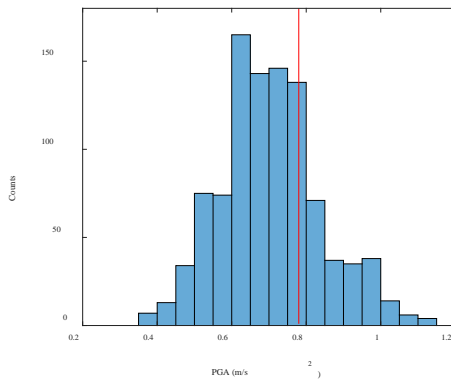
(a) Class E
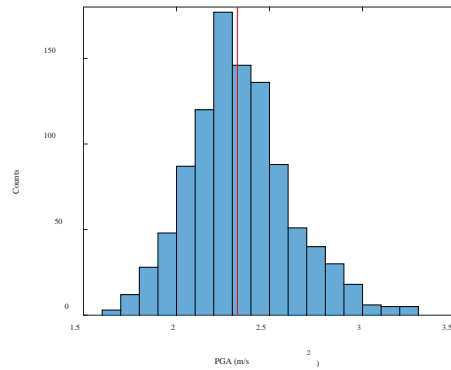
(b) Class D
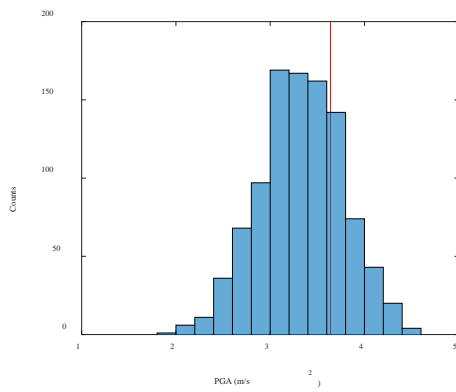
(c) Class C
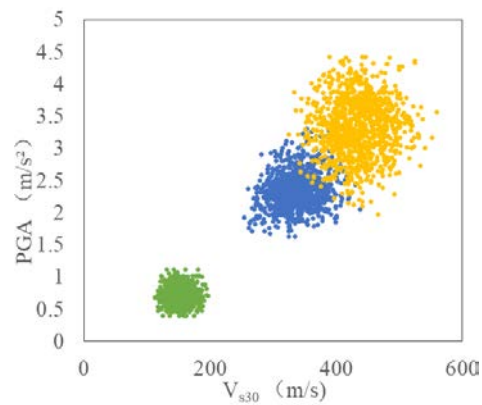
(d) Scatter plot for 3000 times realizations

91

Figure 4.8. Histogram and scatter plot of PGA for three sites, base case result shown as red line. (a) , (b) and (c) are histogram of site class E, D and C, (d) is the scatter plot for 3000 times realizations.

4.4.4 Effect of input distribution in uncertainty of PGA

To conform the conclusion made in Section 4.4.3, and testing the influence on output distribution caused by input distribution, another set of shear wave velocity distribution model is used. In Section 4.3, a set of lognormal distributed shear wave velocity with COV=0.39, ρ=0.577 is used. In this section, shear wave velocity follows the lognormal distribution with COV=0.2 (which is commonly used in engineering), no layer correlation is operated. With 3000 Monte Carlo realizations and Finite Element analyses, the results are obtained and shown in Figure 4.9.

Figure 4.9(a) shows the boxplot of PGA corresponding to the lognormal distribution shear wave velocity with COV=0.39 while considering the inter-layer correlations. For Figure 4.9 (b), the input shear wave velocity follows the lognormal distribution with COV=0.2, no layer correlation considered. Comparing Figure 4.9(a) and Figure 4.9(b), resultant PGA follows the same trend in that the mean value of PGA keeps increasing when soil become harder. Figure 4.9(c), (d) and (e) represent the mean, standard deviation and COV of PGAs that are calculated by two sets of shear wave

velocity profiles following different distributions. It is obvious that under two different sets of input shear wave velocity profiles, the trend of mean, standard deviation and COV of PGA still keep the same with little fluctuation. These results confirm the trend observed in Section 4.4.3 and also indicate both the COV and layer correlation of shear wave velocity have no effect on the trend of output uncertainty, including mean, standard deviation and COV for different site condition.

(a)　　　　　　　　(b)

(c)　　　　　　(d)　　　　　　(e)
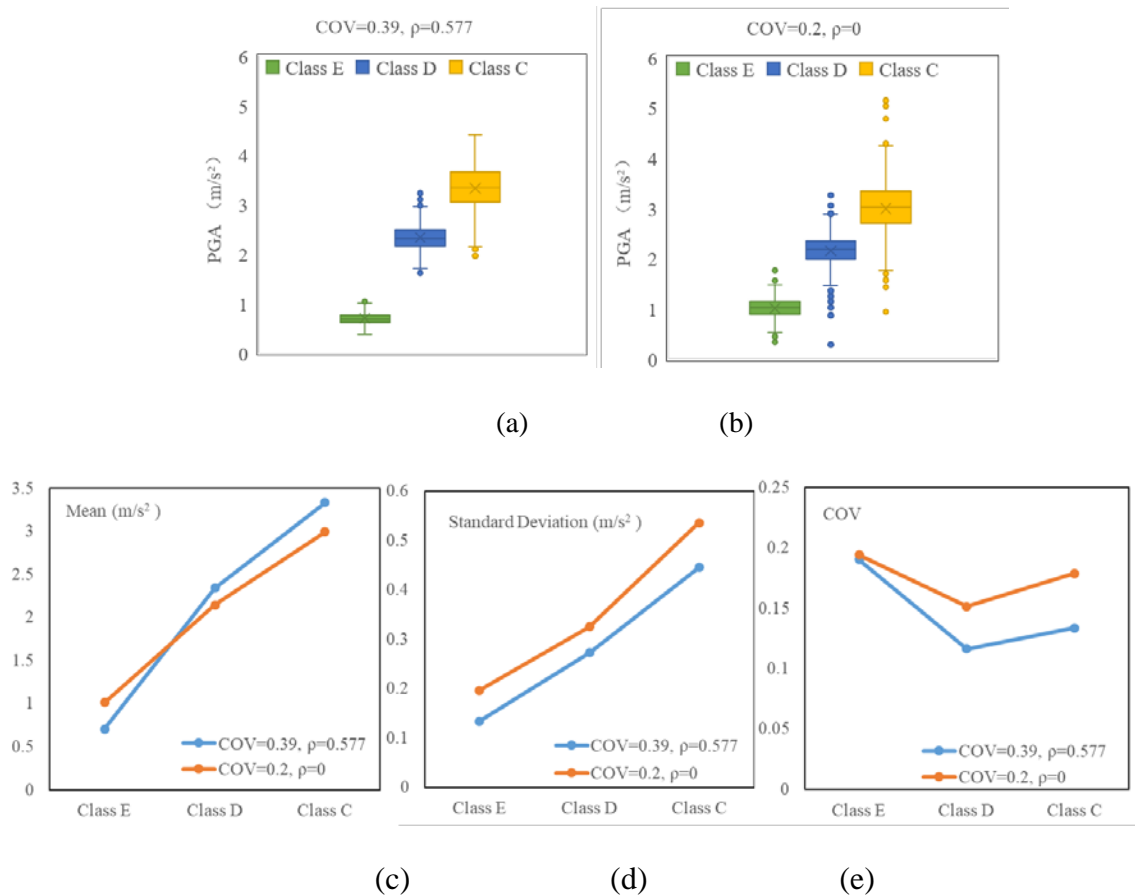
Figure 4.9. Comparison of two different input distributions. (a) and (b) are box plots of PGA with two different V_s distributions, (c) is the mean value of PGA, (d) is standard deviation of PGA, and (e) is the COV of PGA.

4.5 Summary

In this chapter, the effect of soil layers with uncertainty properties on the amplification of surface ground motion with respect to incident rock motion is investigated from a statistical perspective. More specifically, three downhole sites in Stanford, Salt Lake and Hilo which represent three types of site conditions are considered and applied for developing realistic probability models for the nonlinear soil properties based on site-specific geotechnical data. The nonlinear response of the soil was considered here. Modified Davidenkov constitutive model was implemented in ABAQUS/Explicit. The uncertainty of the soil properties and the imperfect correlation of the parameter values in different layers were considered via a Monte Carlo simulation procedure. The site amplification of three site classes was studied here. Each soil column was subjected to a real rock recording applied to its base, and the peak ground acceleration at the ground surface was computed for each run.

In summary, it is found that:

1. Sensitivity analyses performed on the calculated PGA and six input soil parameters revealed that PGA strongly depends on the shear wave velocity. Other parameters in the subroutine are found to be insignificant comparing to shear wave velocity. Uncertainties in the velocity structure of soil profiles (V_s) are shown to be the most intensity dependent variable that governing the amplification potential of the site.

2. Using V_{s30} is not accurate enough for estimating PGA at the ground surface. The

soil layers below 30 meters have contributions in seismic wave prorogation from rock to the surface. Also, the variation of shear wave velocity along depth should be taken into consideration. Thus, shear wave velocity should be expressed as V_s profile instead of V_{s30} value.

3. The results of Monte Carlo Simulation show that with the given lognormally distributed shear wave velocity, PGA values follow lognormal distribution with the exception of a small deviation at the tail of the distribution.

4. With Coyote Lake Earthquake input motion, the calculated PGA at soil surface keeps increasing when soil become harder.

5. For the uncertainty of PGA at site surface, variability in soil properties significantly increases the standard deviation of the ground peak acceleration at site surface but has a lesser effect on the COV. Uncertainties associated with soft soil (Class E) are higher compared to the stiff soil (Class D and C), no obvious difference between Site class D and C.

6. Distribution of input parameter has little influence of uncertainty propagation in dynamic site response analysis. The trend of output uncertainty, including mean, standard deviation and COV, will not change under different distributions of input parameter.

CHAPTER V

5.   UNCERTAINTY QUANTIFICATION IN SEISMIC HAZARD USING

RESPONSE SURFACE METHOD*

5.1 Introduction

Predicting the influence of local soil conditions on expected earthquake ground motions is a critical aspect of the seismic design process. Regarding dynamic site response, it is well known that the soil characteristics of the layers below the surface are indicators of the severity of ground motions at any given site (Campbell, 1979; Toro, 1993). At most sites, however, the soil profiles and the parameters that control the site dynamic response are not known with certainty. The lack of in-situ geotechnical investigation data, the effect of sample disturbance and scaling associated with laboratory tests and the natural heterogeneity of soil profiles are notable sources of uncertainty in nonlinear site response analysis. Therefore, the effect of soil parameter uncertainty on the prediction of site response has been the subject of study. Specifically, randomized realizations of a based soil profile have been used to determine the effect of shear wave velocity, material damping, modulus reduction and input motion (Li and Assimaki 2010). Further, the spectral acceleration at the bedrock of the input record was deemed the single most helpful parameter for the prediction of amplification factor at the same oscillator fre

_____

*A similar form of this chapter has been submitted at the time of writing: Liu, W, Juang, CH, Chen, Q. Uncertainty quantification in seismic hazard using response surface method.

-quency (Bazzurro and Cornell, 2004a). Other studies have entailed evaluating the i) ground-motion hazard at the free-field surface of soil deposits (Bazzurro and Cornell, 2004b), ii) the effect of groundwater level during site response analysis (Wang and Hao, 2002) and iii) the effect of uncertainty in modulus reduction and damping curves as associated with the uncertainty of the amplification factors (Bahrampouri et al. 2018; Stewart and Kwok 2008). Finally, the uncertainty of such input parameters as shear elastic modulus, unit density, damping ratio and the thickness of the soil deposit have also been the subject of simulations (Tombari and Stefanini, 2017).

Although advanced modeling techniques such as the finite element method have been used successfully in site response analysis, their use in uncertainty analysis is hindered by prohibitive computational costs, a problem that has become accentuated with the evolution of time consuming and sophisticated computer codes (Wong, 1985). Also, developing such numerical programs requires expertise in both geotechnical numerical analysis and geotechnical uncertainty analysis.

To bridge the gap, research is being pursued to develop statistical analysis methodologies that are compatible with long-running, numerical simulation codes. Herein, the authors introduce the response surface method to provide computationally efficient statistical approximations to a time consuming model, and to derive the uncertainty propagation through computationally efficient models that approximate the associated deterministic numerical solutions.

Although the response surface analytical method is widely used in slope reliability analysis (Li et al, 2015, Wong, 1985, Zhang et al, 2010, Cho, 2009), in the

reliability analysis of ground–support interaction in circular tunnels (Lü et al., 2011) and in the design of supported excavation (Khoshnevisan et al., 2017), it has never been used in seismic site response analysis.

Herein, the authors describe a method for using the response surface method for seismic site response analysis to assess both the response uncertainty and its dependence on the randomness of the design variables (Figure 5.1). The results of this effort suggest its efficacy in extending the capability of the profession for the reliability analysis of complex geotechnical problems.
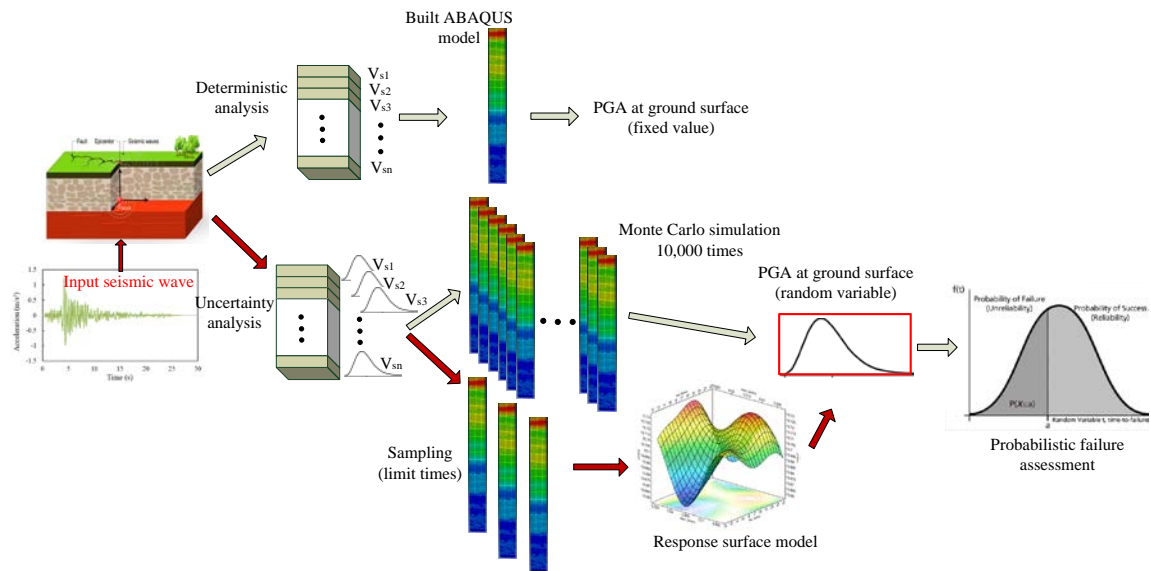
Figure 5.1 Flow chart detailing the uncertainty analysis in the dynamic site response

The structure of this paper is as follows: First, a numerical analysis derived by the finite element method is introduced, specifically as regards to the Constitutive model and the 1D Finite element model. The soil and design parameters, considered as input variables in the response surface in site response analysis, are next characterized. Next, the response surface, which is the process of identifying and fitting an approximate

response surface model from input and output data obtained from the numerical analysis in which each run is regarded as an experiment, is undertaken. Finally, FOSM is used to characterize the uncertainty propagation in site response analysis for extension into such reliability analyses as the probability of liquefaction (Juang, 2006).

5.2 Constitutive model

Constitutive models that describe the non-linear stress-strain behavior of soils under cyclic loadings play a central role in site response analysis. As such, a broad range of simplified (e.g., Ramberg and Osgood, 1943; Duncan and Chang 1970; Hardin and Drnevich, 1972a&b; Pyke, 1979; Martin and Seed, 1982) and advanced soil constitutive models (Prevost, 1977, Li et al., 1997, Borja and Amies 1994) have been proposed in non-linear site response analysis (Hashash et al., 2010). Some of these advanced constitutive models have been used to capture important features of soil behavior such as anisotropy, the generation of excess pore pressure, and dilation (Prevost, 1977, Li et al., 1997, Borja and Amies 1994). However, advanced constitutive models typically require more detailed information on soil behavior for parameter calibration, which for many engineering applications is unsuitable in that the only soil behavior information available is the modulus reduction and damping curves. To this end, simplified models, especially those within the family of hyperbolic models, are often used and will be adopted in this work.

In the next section, the modified Davidenkov constitutive model is introduced to derive the non-linear soil behavior under irregular cyclic loadings. The model is then implemented as a user-defined material subroutine VUMAT in ABAQUS/Explicit, where the details of the numerical implementation algorithm are presented.

5.2.1 Model formulation

The stress-strain curves of the modified Davidenkov model under irregular cyclic loadings are illustrated in Figure 5.2. Two sets of equations are used to describe this hyperbolic-type model. The first set of equations, shown as the backbone, or skeleton curve in Figure 5.2, defines the stress-strain relationship for loading. The second set of equations defines the stress-strain relationship for unloading and reloading (e.g., loading path 1-2-1, 5-6-7 in Figure 5.2).
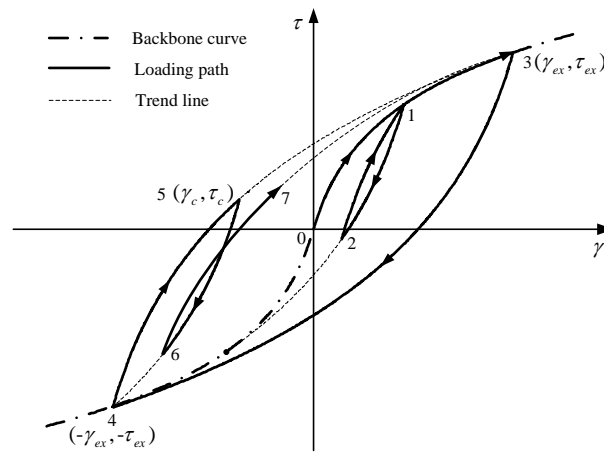
Figure 5.2 Stress-strain curves of the modified Davidenkov model under irregular cyclic loadings

The backbone curve of the modified Davidenkov model is constructed using the following set of equations (Martin and Seed, 1982)

$$\tau = G \cdot \gamma = G_{max} \cdot \gamma \cdot \left[1 - H(\gamma) \right] \tag{5.1}$$

$$H(\gamma) = \left\{ \frac{\left(\gamma / \gamma_0 \right)^{2B}}{1 + \left(\gamma / \gamma_0 \right)^{2B}} \right\}^{2A} \tag{5.2}$$

where τ and γ are the shear stress and shear strain, respectively; G_{max} is the initial shear modulus; A, B and γ_0 are the fitting parameters.

For unloading and reloading (prior to reaching the previous loading reversal point), the following variation of the original Davidenkov model is adopted (Pyke 1979)

$$\tau = \tau_c + G_{max} \cdot (\gamma - \gamma_c) \cdot \left[1 - H\left(\frac{|\gamma - \gamma_c|}{2n} \right) \right] \tag{5.3}$$

where γ_c is the shear strain at the point of loading reversal and n is the scaling factor of the initial loading curve. In this work, the established "n-times method" is adopted instead of the original factor of two (Pyke 1979). The reloading passing the previous reversal point follows the backbone curve.

With both sets of equations defined, the stress-strain curve of a soil under irregular cyclic loadings follows the extended Masing rule (Martin and Seed, 1982; Pyke, 1979):

(4) For initial loading (e.g., path 0-1 in Figure 5.2), the loading path follows the backbone curve defined by Equations (5.1) and (5.2).

(5) If a loading reversal occurs at a point defined by (γ_c, τ_c), subsequent stress-strain curve (e.g., path 1-2) should move down along the path defined by Equation (5.3). The path extends from the loading reversal point to the historical maximum

(minimum) point (e.g., the trend line 2-4).

(6) For reloading (e.g., path 2-1-3), if the reloading curve passes the maximum past strain (e.g., point 1) before unloading, it will follow the backbone curve until the next stress reversal.

With these three rules, the stress-strain curve under irregular cyclic loading can be defined with only stress and strain at loading reversal (τ_c , γ_c), the historical maximum point (τ_{ex} , γ_{ex}) and constant model parameters (G_{max}, A, B, γ_0), minimizing the information the model must "memorize". Such a minimization is a great advantage in terms of numerical implementation, as detailed in the next section.

5.2.2 Numerical implementation in ABAQUS VUMAT

Here, the modified Davidenkov model is implemented as a user-defined material subroutine VUMAT in ABAQUS/Explicit. ABAQUS/Explicit is a robust and well-tested commercial finite element code capable of three-dimensional linear and nonlinear dynamic analysis. The unique feature of ABAQUS/Explicit is that it allows the integration of user-defined material models through the VUMAT interface, which makes it a particularly attractive tool for this research.

Given the values of stress and internal state variables at time t_n, and the strain increment $\Delta\gamma$, the objective of the numerical integration is to find the stress and state variable values at time t_{n+1}. In the modified Davidenkov model, the shear modulus G exhibits a non-linear dependency on strain and also considers the hysteresis behavior of

the soil. The key to its integration entails obtaining a modulus update, which once acquired makes it quite easy to derive the stress value per the elastic theory.

When the loading path follows the backbone curve, the shear modulus at time t_{n+1} can be obtained by taking the derivative of stress τ with respect to strain γ in Equation(5.1).

$$G_{n+1} = \frac{d\tau}{d\gamma} = G_{max}\left[1 - \left(1 + \frac{2AB\gamma_0^{2B}}{\gamma_0^{2B} + \gamma^{2B}}\right) \cdot H(\gamma)\right] \tag{5.4}$$

For unloading and reloading (prior to reaching the previous loading reversal point), the expression of time-varying shear modulus is obtained as Equation (5.5) and (5.6) by the derivative of γ-γ_c in Equation(5.3).

$$G_{n+1} = \frac{d(\tau - \tau_c)}{d(\gamma - \gamma_c)} = G_{max}\left\{1 - \left[1 + \frac{2AB(2n\gamma_0)^{2B}}{(2n\gamma_0)^{2B} + |\gamma - \gamma_c|^{2B}}\right] H\left(\frac{|\gamma - \gamma_c|}{2n}\right)\right\} \tag{5.5}$$

where $(2n\gamma_0)^{2B}$ is determined by both the current point-of-loading reversal and the historical maximum point.

$$(2n\gamma_0)^{2B} = (\gamma_{ex} \pm \gamma_c)^{2B} \cdot \left(\frac{1-R}{R}\right) \tag{5.6}$$

By substituting the historical maximum point (τ_{ex} , γ_{ex}) into Equation (5.3), R is then expressed as follows:

$$R = \left(\frac{\tau_{ex} \pm \tau_c}{G_{max} \cdot (\gamma_{ex} \pm \gamma_c)}\right)^{\frac{1}{A}} \tag{5.7}$$

In this "$\pm$" equation "-" represents the loading procedure and "+" the unloading procedure.

5.3 Elements of response surface in site response analysis

The input variables, specifically the soil and design parameters as related to the response surface, are detailed in this section. Also, the soil parameters are considered as random variables for the balance of this discussion.

5.3.1 Characterization of random soil parameters

The inherent variability, measuring error, and transformation error may result in variations in the soil parameters. For the site response problem shown in Figure 5.1, the main uncertainty parameter is the shear wave velocity for each of the site layers, which is considered as a random variable in this study.

The heterogeneity of soil means that the soil properties may vary in space, thus complicating the effort to place precision soil properties into any analysis. For any proposed engineering project, the soil properties must be simplified through the use of a uniform and deliberate method. Therefore, the shear wave velocity is simplified, the process of which is detailed below.
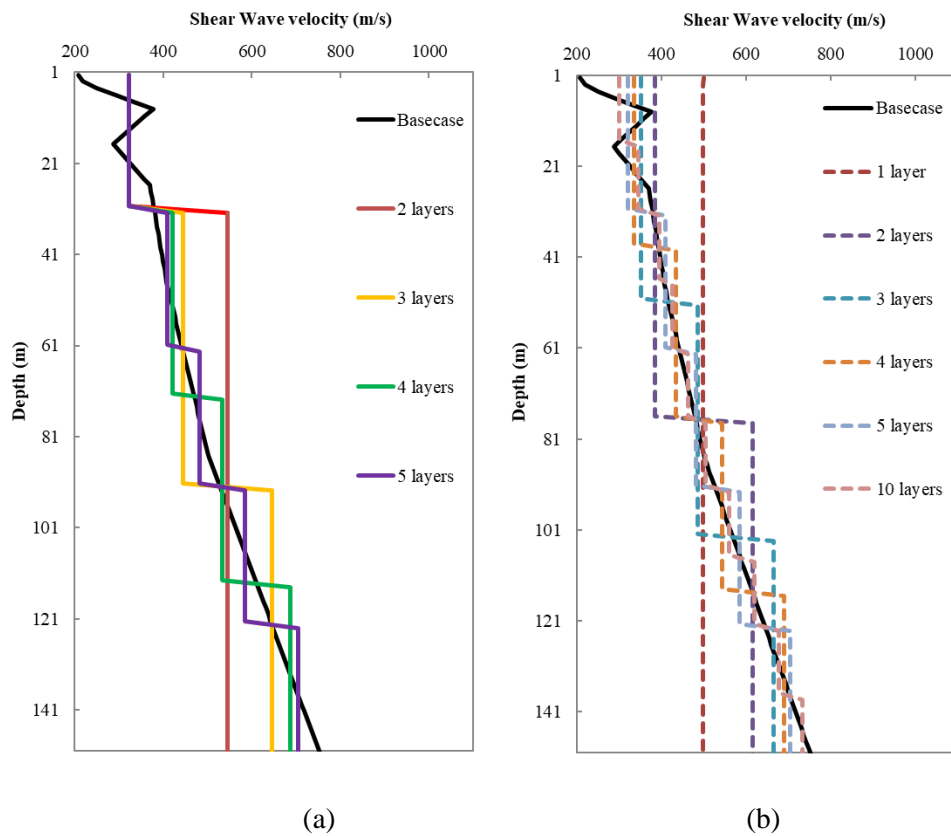
Figure 5.3 The soil profile in different simplification methods: (a) soil profiles divided into 1, 2, 3, 4, 5 and 10 layers equally (b) the top 30m is considered as one layer with the remainder of the 120 m of the soil column divided into 1, 2, 3, 4 extra layers equally (for a total of 2, 3, 4 and 5 layers).

The black solid line in Figure 5.3(b) is assumed as the real soil shear wave velocity profile with 100 layers (denoted as the S7 model). The soil profile is then divided into 1, 2, 3, 4, 5 and 10 layers with the simplified methods shown here. Six dash lines represent the model with the soil profiles equally divided into 1, 2, 3, 4, 5 and 10 layers (model S1, S2, S3, S4 and S5). For that soil column that is less than 4 layers (model S2, S3 and S4), the first layer is larger than 30m. Considering the importance of V_{s30} in describing the soil property, three more models (model S8, S9, and S10) are

added, all of which are represented with chromatic solid lines in Figure 5.3(b). In here, top 30m is considered as one layer, and the rest 120 m of soil column is divided into 1, 2 and 3 extra layers equally (the total of layers 2, 3 and 4).
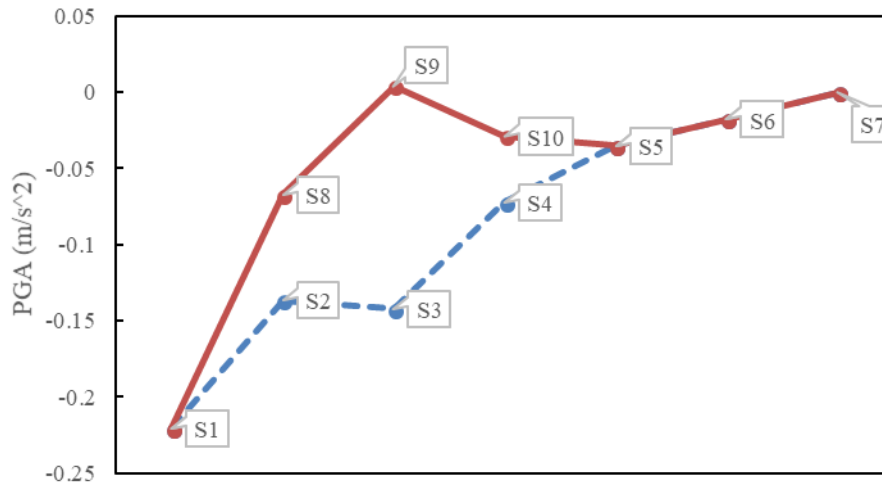
Figure 5.4 Resultant PGA value with different soil profiles

Ten deterministic site response analyses were implemented with ABAQUS/CAE, the results of which are shown in Figure 5.4. The ratio between the difference and the resultant $PGA_{surface}$ in the base case is used to characterize the deviation. In the equipartition cases, represented as a dashed line in Figure 5.4, the deviation continues to approach zero with an increase in the number of layers. When the top 30 m is taken into consideration, however, the accuracy improves greatly. For engineering purposes, those results with a high degree of accuracy and requiring fewer input parameters should be optimal, which means that the three layers exhibited the superior performance in this test. This three-layer profile was again used in a further analysis in this work, in which the top 30 m was considered a single layer with the remainder bisecting the rest of the soil column.

The use of this simplified three-layer profile thus makes it easy to describe the soil properties, with the values given for those uncertain soil parameters $V_{s, \text{layer 1}}$, $V_{s, \text{layer 2}}$, and $V_{s, \text{layer 3}}$ listed in Table 5.2. These listed values are assumed as the mean value. According to Moss (2008), the Coefficient of Variation (COV) changes with the mean value of the shear wave velocity, which is expressed as:

$$COV = 0.000328 \bullet \mu_{V_s} + 0.165967 \qquad (5.8)$$

5.3.2 Characterization of design parameters

Design parameters are those input parameters that easily controlled by the designer, which for purposes of this study are also deemed input parameters. For this site response analysis, the maximum acceleration of the input wave and the class of site are treated as design parameters. The values of each of the design parameters listed in Table 4 were determined via established practice. A discrete space is selected as an example to demonstrate the site response analysis of the nine scenarios considered here. It should be noted that, for site class C, D, and E, different base profiles are used. The base profiles for different site classes (considered as mean value) are listed in Table 5.4 as are the COV values corresponding to the various soil layers.

107

5.4 Deterministic model for site response

An actual site was then analyzed with the Finite element program ABAQUS/CAE with all information provided in Sections 5.2 and 5.3. A soil column of 152 m was meshed into 1000 elements with the element type selected as four-node bilinear plane strain element (CPE4R), which is a general plane strain element with reduced integration and hourglass control.
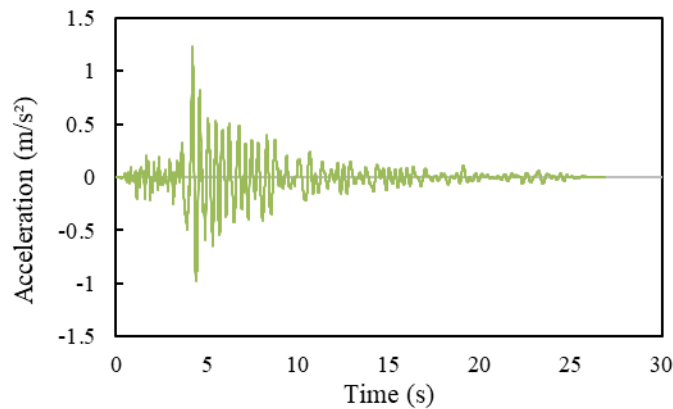
Figure 5.5 Acceleration time-history of the Coyote Lake earthquake record

The displacement of two broadsides and bottom side are fixed in the direction of gravity with ground motion added at the bottom in the horizontal direction. The Coyote Lake earthquake that occurred on the Calaveras Fault near Coyote Lake in Santa Clara County, California on August 6th, 1979, with a magnitude of 5.7 and a maximum acceleration as 1.23m/s² from its SE component is selected as input motion (Figure 5.5). The information about this model is detailed in Table 5.1. This deterministic model was

then used to simulate the site response subjected to a given seismic motion and to obtain the PGA at the site surface.

Table 5.1 Model information and values for the soil parameters used in this study

Model & Element					Soil parameters				γ_0
Model size (m)	Element size (m)	Number of layers	Input motion	Element type	Dense (kg/m³)	Poisson ratio			
152×10	1.52×1	100	Coyote Lake Earthquake	CPE4R	1980	0.49	1.1	0.78	changes with depth

According to the previous section, the simplified three-layer soil model is selected in this study. The detailed soil properties and parameters for the constitutive model in each layer are listed in the Table 5.2. Here, the shear wave velocity for each depth is calculated based on a real site measurement in Stanford University, California, offered by Network for Earthquake Engineering Simulation (NEES) database (https://datacenterhub.org).

Table 5.2 . The basic soil properties adopted in this study

Layer	Depth (m)	Density (kg/m³)	Shear wave velocity (m/s) Class D	A	B	Poisson ratio	γ_0
1	0-30	1980	322	1.1	0.78	0.49	changes with depth
2	30-90	1980	446	1.1	0.78	0.49	changes with depth
3	90-150	1980	646	1.1	0.78	0.49	changes with depth

<u>5.5 Response surface of dynamic site response</u>

The use of a finite element code such as ABAQUS for computing the site response with a given input seismic wave within the uncertainty propagation framework that must consider various random variable scenarios and many designs in a design pool is computationally challenging. Therefore the response surface is selected to ensure the efficiency and practicality of this computation. Here, the response surface was framed as a problem specific-surrogate model rather than a numerical model for computing the site response that is subjected to a given seismic wave.

5.5.1 Response surface procedure

The behavior of the specific site in an earthquake is affected by factors such as site conditions (e.g., the stiffness of soil, dense of soil), site class and input wave. Here, the site class and input wave are set as design parameters. For the deterministic model described in Section 5.4, three soil parameters $\theta = \{\theta_i, i = 3\} = \{V_{s, \, layer \, 1}, V_{s, \, layer \, 2}, V_{s, \, layer \, 3}\}$ are treated as the random variable with the design parameters $d = \{d_i, i = 2\} = \{PGA_{input}, class\}$ also considered in the construction of the response surface. It is noted that, although only three random variables are considered in this study, additional uncertainties may also be considered and included as random variables, the number of random variables are unlimited. For each of these three soil parameters, an

upper limit and lower limit of variation, $\mu + \sigma$ and $\mu - \sigma$ are used to define the region

of interest. Here, μ denotes the mean value and σ denotes the standard deviation. Here,

nine series of parametric studies are conducted to investigate how the dynamic site

response is affected by these factors. For these parametric analyses, the parameter

settings are listed in Table 5.3.

Table 5.3 The nine soil parameter scenarios adopted for developing the response surface

Random variable scenario	$V_{s,\ layer\ 1}$	$V_{s,\ layer\ 2}$	$V_{s,\ layer\ 3}$
1	μ	μ	μ
2	$\mu + \sigma$	μ	μ
3	$\mu - \sigma$	μ	μ
4	μ	$\mu + \sigma$	μ
5	μ	$\mu - \sigma$	μ
6	μ	μ	$\mu + \sigma$
7	μ	μ	$\mu - \sigma$
8	$\mu + \sigma$	$\mu + \sigma$	$\mu + \sigma$
9	$\mu - \sigma$	$\mu - \sigma$	$\mu - \sigma$

Other than the shear wave velocity in each layer, other external factors can also

affect the site response. In this study, these external factors are considered as the 'design

parameters' to extend the response surface model. Here, the peak acceleration of seismic

wave and the site class are taken into consideration. Three values of PGA_{input} are

considered here: 0.1g, 0.3g and 0.5g. Also three site classes, C, D, E, are designed in this

study. For different site class, three real sites are selected as the base case, in which, their

measured shear wave velocity is considered as the mean value. Also, the standard

deviation values are calculated based on mean values. The characteristics of the random

variables for site classes C, D, and E and the assumption of distribution are detailed in

Table 5.4.

With two design parameters and three cases for each parameter, nine design

parameter scenarios are listed in Table 5.5.

Table 5.4 The characteristics of the random variables for site classes C, D, and E in this study

	Site class C		Site class D		Site class E		Distribution
	μ (m/s)	σ	μ (m/s)	σ	μ (m/s)	σ	
Layer 1	424	0.305	322	0.272	170	0.221	Lognormal
Layer 2	805	0.430	446	0.312	287	0.260	Lognormal
Layer 3	1128	0.534	646	0.377	479	0.323	Lognormal

Table 5.5 The selected values for the design parameters in this study

Design parameters scenarios	PGA_{input} (g)	Site class
1	0.1	C
2	0.1	D
3	0.1	E
4	0.3	C
5	0.3	D
6	0.3	E
7	0.5	C
8	0.5	D
9	0.5	E

For each set of design parameters (e.g. design parameter scenario 1), 9 random

variables scenarios are performed (all scenarios are listed in Table 5.3). Thus, a total of

9×9=81 scenarios are obtained.

By repeating the ABAQUS/CAE analysis for each of these scenarios, 81 resultant

PGA values are obtained. For building the response surface model, both the design and

soil parameters are combined as dependent variables for estimating independent variable,

which is the PGA at the site surface.

5.5.2 Model selection

Although the second-order polynomial is a commonly used model for determining the response surface (Bucher and Bourgund 1990; Xu and Low 2006), 18 models were analyzed here, including the first and second order and the intersection to determine the most appropriate response surface model. Different combinations of variables are tested to determine the use of all terms in predicting the PGA values (Figure 5.6 and Table 5.6).
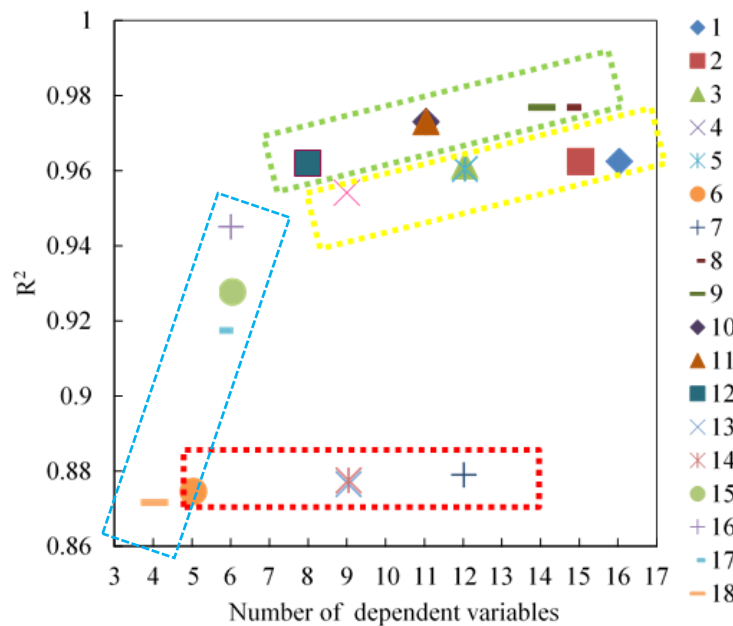
Figure 5.6 R^2 value with different response surface models.

Table 5.6 Terms included in each response surface models and the corresponding R^2 value

Case	No. of fitting parameters	R^2	soil parameters V_{s1}	V_{s2}	V_{s3}	V_{s1}^2	V_{s2}^2	V_{s3}^2	$V_{s1}V_{s2}$	$V_{s1}V_{s3}$	$V_{s2}V_{s3}$	design params pga	ln(pga)	Site class	Class ln(pga)	interaction V_{s1}ln(pga)	V_{s2}ln(pga)	V_{s3}ln(pga)
1	16	0.9629	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			
2	15	0.9628	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓			
3	12	0.9617	✓	✓	✓	✓	✓	✓					✓	✓	✓			
4	9	0.9541	✓	✓	✓								✓	✓	✓			
5	12	0.9609	✓	✓	✓				✓		✓		✓	✓	✓			
6	5	0.8749	✓	✓	✓								✓					
7	12	0.8791	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓					
8	15	0.9772	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓	✓	✓
9	14	0.9772	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓			✓	✓	✓
10	11	0.9734	✓	✓	✓	✓	✓	✓					✓			✓	✓	✓
11	11	0.9734	✓	✓	✓				✓				✓			✓	✓	✓
12	8	0.9619	✓	✓	✓								✓			✓	✓	✓
13	9	0.8769	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓					
14	9	0.8779	✓	✓	✓				✓	✓	✓		✓					
15	6	0.9281	✓	✓									✓			✓	✓	
16	6	0.9452	✓		✓								✓			✓		✓
17	6	0.9178	✓	✓	✓								✓				✓	✓
18	4	0.8720	✓										✓			✓		

114

For engineering purposes, the model should be simplified to a smaller number of parameters with an acceptable R^2 value. Note that 'site class' is considered here as a categorical variable. Given the difficulty of applying this variable to reliability analysis, the possibility of removing the site class was also examined. As shown in Figure 5.6, a downward trend of the R^2 value with a decrease in the number of parameters is clearly evident. The model with the five fitting parameters and four variables received the lowest R^2 value. In Figure 5.6, the points in red box represent the models that did not consider the interaction of input wave and shear wave velocity, and the corresponding R^2 values are lower than other models apparently. The points in both the yellow and green boxes represent the models considering the interaction terms, the relative higher R^2 values indicate the interaction with the input PGA and the shear wave velocity makes a huge contribution in the prediction of the resultant PGA. This interaction term can be expressed as Class×PGA or Vs×PGA. The points in the green box represent the models that include the term Vs×PGA, and the points in the yellow box represent the models that include the term Class×PGA. Although both perform well as categorical variables, but the application of the model with factor variable will be restricted (e.g. the factor variable is nondifferentiable). Therefore, the term Vs×PGA is selected here.

Model No. 12 was selected in that it best synthesized both the R^2 value and number of variables. In this model, only eight fitting parameters are used including the intercept to express the response surface with an R^2 value that is equal to 0.9619. This response surface can be expressed as follows:

$$PGA_{surface} = -2.977 + 0.019 \times V_{s,layer1} + 0.004 \times V_{s,layer2} + 0.001 \times V_{,layers3}$$
$$-1.091 \times \ln(PGA_{input}) + 0.009 \times V_{s,layer1} \times \ln(PGA_{input}) \quad (5.9)$$
$$+0.001 \times V_{s,layer2} \times \ln(PGA_{input}) - 0.001 \times V_{s,layer3} \times \ln(PGA_{input})$$

This response surface model encompasses three random variables and one design variable, all of which are dependent. The PGA (m/s^2) value at site surface can be obtained easily by inserting $V_{s,layer1}$ (m/s), $V_{s,layer2}$ (m/s), $V_{s,layer3}$ (m/s) and PGA_{input} (g) of the input seismic wave into Equation (5.9) rather than the dynamic analysis using the Finite Element method. Note that the application of the developed response surface model requires entering all input parameters with their nominal values. No variation is considered, and the variation in the random variables is detailed in Section 5.4.

5.5.3 Model validation

To determine the accuracy of the obtained response surface model, ten designs were randomly selected from the design pool under the single constraint those designs used for the development of the response surface models are excluded. For each design, the values of the three soil parameters are randomly selected from the assumed probability distributions. For each of these designs, ABAQUS/CAE is used to compute the PGA at the ground surface, with the results then compared with that obtained from the developed response surface model. As shown in Figure 5.7, the accuracy of the response surface model in this example problem was deemed satisfactory. It should be

noted that the response surface model is problem specific, meaning that a new response surface model must be developed for a given set of problems.
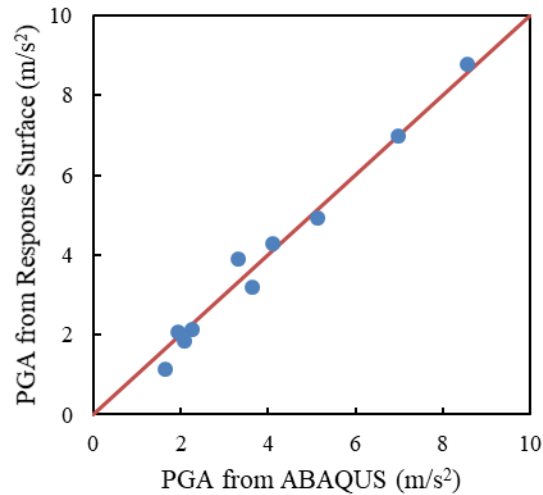
Figure 5.7 Validation of the proposed response surface models for PGA at the site surface

5.5.4 Uncertainty propagation in PGA

The response surface method is also superior in that it can construct a function permitting the ready acceptance of the uncertainty of the dependent variable. Unlike the numerical model, however, only Monte Carlo simulation can be used to study this uncertainty propagation, which while arguably more accurate requires in excess of 10,000 simulations to derive that accuracy. Consequently, a strong computational performance and a long calculation time are required, thus making the convenience of

response surface method, especially in complicated dynamic numerical analysis a better fit.

In this study, the FOSM method is used to calculate the model uncertainty. Formally known as the first-order second-moment method, FOSM derives its name from it basis within a first-order Taylor series approximation of the performance function linearized at the mean values of the random variables, and because of the sole use of the second moment statistic (mean and covariances) of the random variables. Using the response surface model obtained in Section 5.5.2, the performance function can be defined as Equation (5.9). Among these four input parameters, $V_{s,layer1}$, $V_{s,layer2}$, $V_{s,layer3}$ are considered as random variables. In this uncertainty analysis, each of these random variables is assumed to follow a lognormal distribution, which does provide a good fit to the measured geotechnical parameters. Lognormal distribution requires a knowledge of the mean and standard deviation, which for each case in the example studied here, both the mean and the coefficients of variation are available. The mean value for $V_{s,layer1}$, $V_{s,layer2}$, $V_{s,layer3}$ for site class C, D and E are listed in Table 5. According to Equation (5.8) , the standard deviation is expressed as:

$$\sigma_{Vs} = \mu_{Vs} \times COV = 0.000328 \times \mu_{Vs}^2 + 0.165967 \times \mu_{Vs} \tag{5.10}$$

The performance function is generalized for three random variables, denoted by a vector X. Let the performance function be written as:

$$PGA = g(X) = g(X_1, X_2, X_3) \tag{5.11}$$

Followed by a Taylor series expansion of the performance function about the mean value which is expressed as

118

$$PGA = g(\mu_X) + \sum_{i=1}^{3} \frac{\partial g}{\partial X_i}(X_i - \mu_{X_i}) + \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3} \frac{\partial^2 g}{\partial X_i \partial X_j}(X_i - \mu_{X_i})(X_j - \mu_{X_j}) + ... \quad (5.12)$$

where the derivatives are evaluated at the mean values of the random variables (X_1, X_2, X_3) and μ_{Xi} is the mean value of X_i. A linear truncation of terms of this series is then used to derive the first order approximate mean and variance for uncorrelated variables, which is expressed as follows:

$$\mu_{PGA} \approx g\left(\mu_{X1}, \mu_{X2}, \mu_{X3}\right) \quad (5.13)$$

$$\sigma^2_{PGA} \approx \sum_{i=1}^{3}\left(\frac{\partial g}{\partial X_i}\right)^2 Var\left(X_i\right) \quad (5.14)$$

By substituting the Equation (5.9) into Equations (5.13) and (5.14), the approximate mean and variance for the response surface model embedded into Section 5.4 are expressed as:

$$
\begin{aligned}
\mu_{PGA_{surface}} \approx\ & -2.977 + 0.019 \times \mu_{V_{s,layer1}} + 0.004 \times \mu_{V_{s,layer2}} + 0.001 \times \mu_{V_{layers3}} \\
& -1.091 \times \ln(PGA_{input}) + 0.009 \times \mu_{V_{s,layer1}} \times \ln(PGA_{input}) \\
& + 0.001 \times \mu_{V_{s,layer2}} \times \ln(PGA_{input}) - 0.001 \times \mu_{V_{s,layer3}} \times \ln(PGA_{input})
\end{aligned} \quad (5.15)
$$

$$
\begin{aligned}
\sigma^2_{PGA_{surface}} =\ & \left(0.019 + 0.009 \cdot \ln(PGA_{input})\right)^2 \cdot \left(0.000328 \times \mu^2_{Vs,layer1} + 0.165967 \times \mu_{Vs,layer1}\right)^2 \\
& + \left(0.004 + 0.001 \cdot \ln(PGA_{input})\right)^2 \cdot \left(0.000328 \times \mu^2_{Vs,layer2} + 0.165967 \times \mu_{Vs,layer2}\right)^2 \\
& + \left(0.001 - 0.001 \cdot \ln(PGA_{input})\right)^2 \cdot \left(0.000328 \times \mu^2_{Vs,layer3} + 0.165967 \times \mu_{Vs,layer3}\right)^2
\end{aligned} \quad (5.16)
$$

These given equations thus make it possible to easily obtain both the mean value and standard deviation of PGA$_{surface}$ by substituting input variables into Equations (5.15) and (5.16), both of which are effective in a further reliability analysis such as in determining the probability of liquefaction (Juang, 2006). This concept also provides

removes the limitation of current geotechnical reliability analyses in terms of focusing on problems with a relatively simple set of limit-state functions. Most numerical programs lack such an effective reliability analysis. For example, even though the Monte Carlo simulation is a well-developed method for determining uncertainty analysis, it requires strong computational performance and long calculation times. However, this combination of the response surface model and FOSM method is the most effective method for characterizing the uncertainty propagation in site response analysis in that it circumvents those lengthy calculation requirements.

Table 5.7 The mean estimation and the standard deviation of the PGA at the site surface.

	Class C	Class D	Class E
$\mu_{PGA_{surface}}$	3.57	1.87	1.28
$\sigma_{PGA_{surface}}$	1.90	0.85	0.53
COV	0.53	0.46	0.41

Substituting the V_s listed in Table 5.7 into Equations (5.15) and (5.16) greatly enhances acquisition of the $\mu_{PGA_{surface}}$ and $\sigma_{PGA_{surface}}$ (listed in Table 8). Although both the mean value and standard deviation of PGA$_{surface}$ decreases in site classes C to E, the COV trend is more opaque, possibly in that the output distribution is subjected to the input distribution. Here, the COV of the V_s changes with the mean value. Again in site classes C to E, the mean value of V_s decreases which is the cause of the decreased COV, a decrease that will in turn cause a similar trend in output distribution of the COV value.

Summary

The deterministic, complex and time-consuming procedure that characterized current site response analysis necessitates considering uncertainties in that analysis which is the result of uncertainties in the soil properties. Six findings in this study are of particular note.

(1) A procedure for developing a response surface surrogate model to replace the time consuming finite element code in statistical studies was proposed, through which the power of the finite element method in modeling and the statistical methods in addressing uncertainties are combined into an effective statistical analysis tool.

(2) The finite element method was used to create a response surface model based on the seismic dynamic analysis, and a Modified Davidenkov model was used to describe soil behaviors under irregular loading reloading conditions.

(3) FOSM, based upon the surface model that was described and constructed here, was used to characterize the uncertainty propagation of PGA caused by the uncertainty of the soil parameter. The resultant formulas of the mean estimation and standard deviation of PGA now makes it much easier to undertake geotechnical reliability analysis, such as liquefaction probability. Unlike most geotechnical reliability schemes that emphasize problems with simple limit-state functions, this novel concept was efficient in solving problems in terms of

determining reliability analysis.

(4) The interaction with input ground motion and shear wave velocity were most important in predicting the resultant PGA. However, given that the commonly used second-order polynomial model was not used here, future response surface model selection efforts must be targeted towards test the model with different patterns of manifestation instead of the direct use of the second-order polynomial model.

(5) A uniform and deliberate method is necessary for simplifying the soil properties for a specific engineering purpose. Here, the three layers profile which considers the top 30 m as a single layer and a bisection of the remainder of the soil column, was deemed most suitable for this exercise.

(6) Given the problem-specific of this proposed response surface model, specific response surface models must be developed for a certain set of problem criteria.

CHAPTER VI

6. CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

The ground motion parameters such as amplitude, frequency content or the duration can be affected by the local site condition and may result in amplification or de-amplification to the original bedrock motion. Shear wave velocity is an important site parameter to describe the site condition that is widely used in estimating site response, classifying sites in recent building codes and loss estimation. In this dissertation, the geostatistical approach accounting for the spatial variability of V_{s30} across different length scales and incorporates the compiled database of direct geophysical measurements and proxy-based V_{s30} values is presented. Based on that, a hybrid geotechnical and geological data-based framework is presented and well tested in this dissertation. Recommendations on how to act in each specific mapping situation and step-by-step instructions for accomplishing mapping are provided including the methodology to calibrate B and integrate secondary data.

With the well understood of the variation of shear wave velocity, a comprehensive study is conducted to evaluate the impacts of soil parameters uncertainty especially shear wave velocity in nonlinear dynamic site response. Numerical analysis with finite element method is applied to describe the uncertainty propagation in dynamic

site response and its performance under different site condition. Modified Davidenkov model with simplified loading-reloading rules is used to describe the stress-strain relationship under irregular cyclic loading. Furthermore, the response surface and the first order second moment (FOSM) concept are integrated to quantify the uncertainty in dynamic site response.

In chapter II, a multiscale random field-based framework is presented to map V_{s30} values over extended areas. The random field model explicitly accounts for the spatial variability of V_{s30} across different scales while incorporates and preserves measured V_{s30} data. The framework is applied to map V_{s30} over the Suzhou site, where 309 shear-wave velocity measurements and topography-based V_{s30} values are compiled. Monte Carlo simulations are coupled with the random field model to quantify uncertainties of the generated multiscale V_{s30} map. The new map is then applied to site classification and amplification factor characterization in the studied region. In summary, it is found that: Quantitatively consistent V_{s30} estimates over different length scales over the entire studied region can be obtained using the multiscale random field model; The resulting map has multiscale resolutions and is particularly convenient to incorporate and preserve local measurement data into a regional V_{s30} map; Comparison of the new V_{s30} map with existing USGS topography-based Vs30 map shows that the new V_{s30} map provides more accurate and more detailed V_{s30} values, especially in the eastern plain region of the studied site because of the incorporated local Vs30 measurements and their spatial dependency; Uncertainties associated with the new V_{s30} map are quantified in terms of the coefficient of variation (COV) calculated from Monte Carlo simulations. In general,

the COVs approach zero around locations with measurement data and gradually increase in areas without any known V_{s30} values. COVs in single scale random field map are found to be slightly smaller when compared to the multiscale counterpart; The site application map based on the newly generated V_{s30} map shows that relatively stiff soil (NEHRP site class C) is found in the northwestern part of the city and the soil tends to be softer in the southeastern region (NEHRP site class D and E); This trend in the soil type correlates well with the calculated amplification factor map, where high amplification factors are predicted in the southeastern part of the city, indicating potential seismic amplification effect in this region.

In chapter III, a guideline to integrate multiple sources of data in random field model for regional soil properties mapping is established based on a hybrid geotechnical and geological data-based framework. The geotechnical data such as the V_{s30} measured data are used as primary information, which show clear spatial correlation. Geological information is considered as secondary information which essentially enforces geological constraints to the generated soil properties maps. Both primary and secondary data are integrated into random field models through a conditional sequential simulation technique. The integration effect caused by Markov–Bayes coefficient B and the size of the predefined grid of secondary information are emphatically discussed within two sampling conditions: sufficient measured data and insufficient measured data. A well-designed 2-D synthetic digital field is applied here to test the proposed framework. In summary, it is found that: Since the secondary data represents the geologic information, as the weight of the secondary data increases, the geological boundaries become more

distinguishable in the generated map; It is necessary to selected an appropriate threshold value for calibrating Markov–Bayes coefficient B. Other than it's physical meaning, the threshold value should be selected within the intersection value range of the primary and secondary data; The element size of the predefined grid for secondary data cannot be smaller than the mean distance of any closest two primary data points; The overdense secondary data may result in the randomly estimated value; Integration of secondary data can modified mapping scenario with insufficient, unevenly distributed measured data. For the mapping scenario with sufficient measured data in all geologic units, it can modify the resultant mapping slightly.

In chapter IV, the effect of soil layers with uncertainty properties on the amplification of surface ground motion with respect to incident rock motion is investigated from a statistical perspective. More specifically, three downhole sites in Stanford, Salt Lake and Hilo which represent three types of site conditions are considered and applied for developing realistic probability models for the nonlinear soil properties based on site-specific geotechnical data. The nonlinear response of the soil was considered here. The modified Davidenkov constitutive model was implemented in ABAQUS/Explicit. The uncertainty of the soil properties and the imperfect correlation of the parameter values in different layers were considered via a Monte Carlo simulation procedure. The site amplification of three site classes was studied here. Each soil column was subjected to a real rock recording applied to its base, and the peak ground acceleration at the ground surface was computed for each run. In summary, it is found that: Sensitivity analyses performed on the calculated PGA and six input soil parameters

revealed that PGA strongly depends on the shear wave velocity; Other parameters in the subroutine are found to be insignificant comparing to shear wave velocity. Uncertainties in the velocity structure of soil profiles (V_s) are shown to be the most intensity dependent variable that governing the amplification potential of the site; Using V_{s30} is not accurate enough for estimating PGA at ground surface; The soil layers below 30 meters have contribution in seismic wave prorogation from rock to the surface. Also, the variation of shear wave velocity along depth should be taken into consideration. Thus, shear wave velocity should be expressed as V_s profile instead of V_{s30} value. The results of Monte Carlo Simulation show that with the given lognormally distributed shear wave velocity, PGA values follow the lognormal distribution with the exception of a small deviation at the tail of the distribution. With Coyote Lake Earthquake input motion, the calculated PGA at soil surface keeps increasing when soil become harder. For the uncertainty of PGA at site surface, variability in soil properties significantly increases the standard deviation of the ground peak acceleration at site surface but has a lesser effect on the COV. Uncertainties associated with soft soil (Class E) are higher compared to the stiff soil (Class D and C), no obvious difference between Site class D and C. Distribution of input parameter has little influence of uncertainty propagation in dynamic site response analysis. The trend of output uncertainty, including mean, standard deviation and COV, will not change under different distributions of input parameter.

In chapter VI, response surface method is introduced in dynamic site response analysis to assess both the response uncertainty and its dependence on the randomness of the design variables. The deterministic, complex and time-consuming procedure that

characterized current site response analysis necessitates consider uncertainties in that analysis which is the result of uncertainties in the soil properties. In summary, it is found that: A procedure for developing a response surface surrogate model to replace the time consuming finite element code in statistical studies was proposed, through which the power of the finite element method in modeling and the statistical methods in addressing uncertainties are combined into an effective statistical analysis tool. The finite element method was used to create a response surface model based on the seismic dynamic analysis, and a Modified Davidenkov model was used to describe soil behaviors under irregular loading reloading conditions; FOSM, based upon the surface model that was described and constructed here, was used to characterize the uncertainty propagation of PGA caused by the uncertainty of the soil parameter. The resultant formulas of the mean estimation and standard deviation of PGA now makes it much easier to undertake geotechnical reliability analysis, such as liquefaction probability. Unlike Most geotechnical reliability schemes that emphasize problems with simple limit-state functions, this novel concept was efficient in solving problems in terms of determining reliability analysis. The interaction with input ground motion and shear wave velocity were most important in predicting the resultant PGA. However, given that the commonly used second-order polynomial model was not used here, future response surface model selection efforts must be targeted towards test the model with different patterns of manifestation instead of the direct use of the second-order polynomial model. A uniform and deliberate method is necessary for simplifying the soil properties for a specific engineering purpose. Here, the three-layer profile that considers the top 30 m as a single

layer and a bisection of the remainder of the soil column, was deemed most suitable for this exercise. Given the problem-specific of this proposed response surface model, specific response surface models must be developed for a certain set of problem criteria.

6.2 <u>Recommendations</u>

To further expand the work presented in this dissertation, a number of research topics may be undertaken, which include the following:

(1) Further investigation of a more complex finite element model for site response analysis is suggested, such that the boundary condition can be more accurate included.

(2) Other than the variables that considered in chapter V and chapter VI, depth of soil is also a parameter worth considering. In chapter V, a basic study has been done which related to the depth of soil. Based on that, it can be selected as a variable and considered into the response surface model. It should be noticed that this should be based upon field data.

(3) In the constitutive model described in Chapter IV, the values of fitting parameters are obtained from other researcher's work. This may only apply to the soil in specific area. It is recommended to do more experiments based on the soil form different site conditions to offer a more widely used range of those fitting parameters.

REFERENCES

Abrahamson, N, Silva, WJ, (1997). Empirical response spectral attenuation relations for shallow crustal earthquakes. *Seismol. Res. Lett*. 68 (1), 94–127.

Abrahamson, N. and Silva, W. (2008). Summary of the Abrahamson & Silva NGA ground-motion relations. *Earthquake Spectra*, 24(1), 67-97.

Abrahamson, N.A., Silva, W.J. and Kamai, R., (2014). Summary of the ASK14 ground motion relation for active crustal regions. *Earthquake Spectra*, 30(3), pp.1025-1055.

Atkinson, G.M., (2015). Ground‐motion prediction equation for small‐to‐moderate events at short hypocentral distances, with application to induced‐seismicity hazards. *Bulletin of the Seismological Society of America*, 105(2A), pp.981-992.

Allen, TI, Wald, DJ, (2009). On the use of high-resolution topographic data as a proxy for seismic site conditions V_{s30}. *Bull. Seismol. Soc. Am*. 99 (2A), 935–43.

Ancheta, T, Darragh, R, Stewart, J, Seyhan, E, Silva, W, Chiou, B, Wooddell, K, Graves, R, Kottke, A, Boore, D, Kishida, T, Donahue, J, (2014). NGA-West2 database. *Earthq. Spectra* 30 (3), 989–1005.

Bahrampouri, M., Rodriguez-Marek, A. and Bommer, J.J. (2018). Mapping the uncertainty in modulus reduction and damping curves onto the uncertainty of site amplification functions. *Soil Dynamics and Earthquake Engineering*.

Baker, JW, Seifried, A, Andrade, JE, Chen, Q, (2011). Characterization of random fields at multiple scales: an efficient conditional simulation procedure and applications in geomechanics. *Applications of Statistics and Probability in Civil Engineering* pp. 347–8.

Bazzurro, P. and Cornell, C. A. (2004 a). Ground-motion amplification in nonlinear soil sites with uncertain properties. *Bulletin of the Seismological Society of America*, 94(6), 2090-2109.

Bazzurro, P. and Cornell, C. A. (2004 b). Nonlinear soil-site effects in probabilistic seismic-hazard analysis. *Bulletin of the Seismological Society of America*, 94(6), 2110-2123.

Bhuiyan, M. and A, Haque. 2015. Application of Nonlinear Site Response Analysis in Coastal Plain South Carolina, Ph. D. dissertation, Clemson University, Clemson, SC.

Boaga, J., Vignoli, G., & Cassiani, G. (2011). Shear wave profiles from surface wave inversion: the impact of uncertainty on seismic site response analysis. *Journal of Geophysics and Engineering*, 8(2), 162.

Boore, DM, (2004). Estimating Vs30 (or NEHRP site classes) from shallow velocity models. *Bull. Seismol. Soc. Am*. 94 (2), 591–7.

Boore, D. M. and Atkinson, G. M. (2008). Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s. *Earthquake Spectra*, 24(1), 99-138.

Boore, DM, Thompson, EM, Cadet, H, (2011). Regional correlations of Vs30 and velocities averaged over depths less than and greater than 30 meters. *Bull. Seismol. Soc. Am.* 101 (6), 3046–59.

Boore, D.M., Stewart, J.P., Seyhan, E. and Atkinson, G.M., (2014). NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes. *Earthquake Spectra*, 30(3), pp.1057-1085.

Borja, R. I., and Amies, A. P. (1994). Multiaxial Cyclic Plasticity Model for Clays. *Journal of Geotechnical and Geoenvironmental Engineering*, 120(6), 1051-1070.

Bucher, C.G. and Bourgund, U. (1990). A fast and efficient response surface approach for structural reliability problems. *Structural safety*, 7(1), 57-66.

Campbell, J. B. (1979). SPATIAL VARIABILITY OF SOILS ∗. *Annals of the Association of American Geographers*, 69(4), 544-556.

Campbell, K. W. and Bozorgnia, Y. (2008). NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10 s. *Earthquake Spectra*, 24(1), 139-171.

Campbell, K.W. and Bozorgnia, Y., (2014). NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra. *Earthquake Spectra*, 30(3), pp.1087-1115.

Chen, G., Liu, X., Zhu, D. and Hu, Q. (2005). The experimental study on dynamic shear modulus ratio and damping ratio of recently deposited soil in southern area of Jiangsu province in China. *Geotechnical Engineering for Disaster Mitigation and Rehabilitation*, 269-274.

Chen, Q, Seifried, A, Andrade, JE, Baker, JW, (2012). Characterization of random fields and their impact on the mechanics of geosystems at multiple scales. Int. J. Numer. Anal. *Methods Geomech*. 36 (2), 140–65.

Chen, Q, Wang, C, Juang, CH, (2015). CPT-based evaluation of liquefaction potential accounting for soil spatial variability at multiple scales. J. *Geotech. Geoenviron.* 04015077.

Chen, Q, Wang, C, Juang, CH, (2016). Probabilistic and spatial assessment of liquefaction-induced settlements through multiscale random field models. *Eng. Geol.* 211, 135–49.

Choi, Y, Stewart, JP, (2005). Nonlinear site amplification as function of 30 m shear wave velocity. *Earthq. Spectra* 21 (1), 1–30.

Chiou, B.J. and Youngs, R.R., (2008). An NGA model for the average horizontal component of peak ground motion and response spectra. *Earthquake Spectr*a, 24(1), pp.173-215.

Chiou, B.S.J. and Youngs, R.R., (2014). Update of the Chiou and Youngs NGA model for the average horizontal component of peak ground motion and response spectra. *Earthquake Spectra*, 30(3), pp.1117-1153.

Cho, S.E. (2009). Probabilistic stability analyses of slopes using the ANN-based response surface. *Computers and Geotechnics*, 36(5), 787-797.

Deutsch, CV, Journel, AG, (1998). *Geostatistical software library and user's guide*. New York, NY, USA: Oxford University Press.

Duncan, J.M. and Chang, C.Y. (1970) "Nonlinear Analysis of Stress and Strains in Soils," J. Soil Mechanics and Foundation Division, ASCE, v., 96(5), 1629–1653

Elgamal, A., Yang, Z. and Parra, E. (2002). Computational modeling of cyclic mobility and post-liquefaction site response. *Soil Dynamics and Earthquake Engineering*, 22(4), 259-271.

Goovaerts, P, (1997). Geostatistics for Natural Resources Evaluation. *Oxford University Press, New York*.

Hardin, B.O. and Drnevich, V.P. (1972 a). Shear modulus and damping in soils: design equations and curves. *Journal of Soil Mechanics & Foundations Div*, 98(7), 667-692.

Hardin, B.O. and Drnevich, V.P. (1972 b). Shear modulus and damping in soils: measurement and parameter effects. *Journal of Soil Mechanics & Foundations Div*, 98(sm6).

Hashash, Y., Phillips, C. and Groholski, D.R. (2010). Recent advances in non-linear site response analysis. *International Conferences on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics*. Paper 8.

Idriss, I.M., (2008). An NGA empirical model for estimating the horizontal spectral values generated by shallow crustal earthquakes. *Earthquake Spectra*, 24(1), pp.217-242.

Idriss, I.M., (2014). An NGA-West2 empirical model for estimating the horizontal spectral values generated by shallow crustal earthquakes. *Earthquake Spectra*, 30(3), pp.1155-1177.

Juang, C.H., Fang, S.Y. and Khor, E.H., (2006). First-order reliability method for probabilistic liquefaction triggering analysis using CPT. *Journal of Geotechnical and Geoenvironmental Engineering*, 132(3), pp.337-350.

Kamalian, M., Jafari, M. K., Sohrabi-Bidar, A., Razmkhah, A. and Gatmiri, B. (2006). Time-domain two-dimensional site response analysis of non-homogeneous topographic structures by a hybrid BE/FE method. *Soil Dynamics and Earthquake Engineering*, 26(8), 753-765.

Khoshnevisan, S., Wang, L. and Juang, C.H. (2017). Response surface-based robust geotechnical design of supported excavation–spreadsheet-based solution. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards,* 11(1), 90-102.

Lee, C, Tsai, B, (2008). Mapping vs30 in Taiwan. *Terr. Atmos. Ocean. Sci.* 19 (6), 671–82.

Lee, RL, Bradley, BA, Ghisetti, FC, Thomson, EM, (2017). Development of a 3D Velocity Model of the Canterbury, New Zealand, Region for Broadband Ground‐Motion Simulation. *Bulletin of the Seismological Society of America*, 107(5), 2131-2150.

Li, D.Q., Jiang, S.H., Cao, Z.J., Zhou, W., Zhou, C.B. and Zhang, L.M. (2015). A multiple response-surface method for slope reliability analysis considering spatial variability of soil properties. *Engineering Geology*, 187, 60-72.

Li, W. and Assimaki, D. (2010). Site-and motion-dependent parametric uncertainty of site-response analyses in earthquake simulations. *Bulletin of the Seismological Society of America*, 100(3), 954-968.

Li, X., Wang, Z. L., and Shen, C. K. (1997). SUMDES. A nonlinear procedure for response analysis of horizontally layered sites subjected to multidirectional

earthquake loading. Department of Civil Engineering, *University of California at Davis*.

Liu, W, Wang, C, Chen, Q, Chen, G, Juang, CH, (2017a). Multiscale random field-based shear wave velocity mapping and site classification. *Proceedings of the Geo-Risk 2017 Conference, Denver, Colorado.*

Liu, W, Chen, Q, Wang, C, Juang, CH, (2017b). Spatially correlated multiscale Vs30 mapping and a case study of the suzhou site. *Engng Geol*. 220, 110–122.

Lü, Q., Sun, H.Y. and Low, B.K. (2011). Reliability analysis of ground–support interaction in circular tunnels using the response surface method. *International Journal of Rock Mechanics and Mining Sciences,* 48(8), 1329-1343.

Martin, P. P. and Seed, H. B. (1982). One-dimensional dynamic ground response analyses. *Journal of the Geotechnical Engineering Division*, 108(7), 935-952.

Moss, R.E.S. (2008). Quantifying measurement uncertainty of thirty-meter shear-wave velocity. *Bulletin of the Seismological Society of America*, 98(3), 1399-1411.

Moysey, S, Caers, J, Knight, R, Allen-King, RM, (2003). Stochastic estimation of facies using ground penetrating radar data. *Stochastic Environ. Res. Risk Assessment* 17, No. 5, 306–318.

Prevost, J. H. (1977). Mathematical modeling of monotonic and cyclic undrained clay behavior. *International Journal of Numerical and Analytical Methods in Geomechanics*, 1(2), 195-216.

Park, D., and Hashash, Y. M. A. (2008). Rate-dependent soil behavior in seismic site response analysis. *Canadian Geotechnical Journal*, 45(4), 454-446.

Pisanò, F. and Jeremić, B. (2014). Simulating stiffness degradation and damping in soils via a simple visco-elastic–plastic model. *Soil Dynamics and Earthquake Engineering*, 63, pp.98-109.

Pyke, R. M. (1979). Nonlinear soil models for irregular cyclic loadings. *Journal of the Geotechnical Engineering Division,* 105(6), 715–726.

Ramberg, W. and Osgood, W.R. (1943). Description of stress-strain curves by three parameters. *Technical Note No. 902*. National Advisory Committee for Aeronautics. Washington, DC;

Rathje, E. M., Kottke, A. R., & 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.

Ronaldo I. B., Chao H. L., Kossi M. S., and Gwynn M. M. (2000). Modelling non-linear ground response of non-liquefiable soils. *Earthquake Engng Struct. Dyn.* 2000; 29, 63-83.

Scasserra, G, Stewart, JP, Kayen, RE, Lanzo, G, (2009). Database for earthquake strong motion studies in Italy. *J. Earthq. Eng.* 13 (6), 852–81.

Seed, H. B., & Idriss, I. M. (1969). Influence of soil conditions on ground motions during earthquakes. *Journal of the Soil Mechanics and Foundations Division,* 95(1), 99-138.

Seyhan, E. and Stewart, J. P. (2014). Semi-empirical nonlinear site amplification from NGA-West2 data and simulations. *Earthquake Spectra*, 30(3):1241–1256.

Seyhan, E, Stewart, JP, Ancheta, TD, Darragh, RB, Graves, RW, (2014). NGA-West2 site database. *Earthq. Spectra* 30 (3), 1007–24.

Stewart, J.P., Douglas, J., Javanbarg, M., Bozorgnia, Y., Abrahamson, N.A., Boore, D.M., Campbell, K.W., Delavaud, E., Erdik, M. and Stafford, P.J., (2015). Selection of ground motion prediction equations for the Global Earthquake Model. *Earthquake Spectra,* 31(1), pp.19-45.

Stewart, J. P., and Kwok, A. O. (2008). Nonlinear seismic ground response analysis: Code usage protocols and verification against vertical array data. *Geotechnical engineering and soil dynamics IV,* D. Zeng, M. T. Manzari, and D. R. Hiltunen, eds., ASCE, Reston, VA, 1–24.

Stewart, J. P. and Seyhan, E. (2013). Semi-empirical nonlinear site amplification and its application in NEHRP site factors. *Technical Report PEER 2013/13*, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Tombari, A. and Stefanini, L. (2017). A Fuzzy Logic Approach to Stochastic 1D Site Response Analysis accounting for Soil Uncertainties.

Toro, G. R. (1993). Probabilistic model of soil-profile variability. *Electric Power Research Institute*, EPRI TR-102293, 2.

Thompson, E, Wald, DJ, Worden, C, (2014). A V_{s30} map for California with geologic and topographic constraints. *Bull. Seismol. Soc. Am.* 104 (5), 2313–21.

Thompson, EM, Baise, LG, Kayen, RE, (2007). Spatial correlation of shear-wave velocity in the San Francisco Bay Area sediments. *Soil Dyn. Earthq. Eng.* 27 (2), 144–52.

Thompson, EM, Baise, LG, Kayen, RE, Morgan, EC, Kaklamanos, J, (2011). Multiscale site-response mapping: a case study of Parkfield, California. *Bull. Seismol. Soc. Am.* 101 (3), 1081–100.

Thompson, EM, Baise, LG, Kayen, RE, Tanaka, Y, Tanaka, H, (2010). A geostatistical approach to mapping site response spectral amplifications. *Eng. Geol.* 114 (3), 330–42.

Wald, D, Earle, P, Quitoriano, V, (2004). Topographic slope as a proxy for seismic site correction and amplification. *EOS. Trans. AGU* 85 (47), F1424.

Wald, DJ, Allen, TI, (2007). Topographic slope as a proxy for seismic site conditions and amplification. *Bull. Seismol. Soc. Am.* 97 (5), 1379–95.

Wald, DJ, McWhirter, L, Thompson, E, Hering, AS, (2011). A new strategy for developing V_{s30} maps. *Proc. 4th Int. Effects of Surface Geology on Seismic Motion Symp.* Santa Barbara, California, 23– 26 August 2011, 12 pp.

Wang, C, Chen, Q, (2017). A hybrid geotechnical and geological data-based framework for multiscale regional liquefaction hazard mapping. *Géotechnique*, 68(7), 614-625.

Wang, S. and Hao, H. (2002). Effects of random variations of soil properties on site amplification of seismic ground motions. *Soil Dynamics and Earthquake Engineering*, 22(7), 551-564.

Webster, R, Oliver, MA, (1992). Sample adequately to estimate variograms of soil properties. *European Journal of Soil Science,* 43(1), 177-192.

Wills, C, Clahan, K, (2006). Developing a map of geologically defined site-condition categories for California. *Bull. Seismol. Soc. Am.* 96 (4A), 1483–501.

Wills, C, Gutierrez, C, (2008). Investigation of geographic rules for improving site-conditions mapping. *California Geologic Survey Final Technical Report* (07HQGR0061)

Wills, C, Petersen, M, Bryant, W, Reichle, M, Saucedo, G, Tan, S, Taylor, G, Treiman, J, (2000). A site-conditions map for California based on geology and shear-wave velocity. *Bull. Seismol. Soc. Am.* 90 (6B), S187–S208.

Wills, CJ, Silva, W, (1998). Shear-wave velocity characteristics of geologic units in California. *Earthq. Spectra* 14 (3), 533–56.

Wills, CJ, Gutierrez, CI, Perez, FG, Branum, DM, (2015). A Next Generation V_{s30} Map for California Based on Geology and TopographyA Next Generation V_{s30} Map

for California Based on Geology and Topography. *Bulletin of the Seismological Society of America*, 105(6), 3083-3091.

Wong, F.S. (1985). Slope reliability and response surface method. *Journal of Geotechnical Engineering,* 111(1), 32-53.

Xu, B. and Low, B.K. (2006). Probabilistic stability analyses of embankments based on finite-element method. *Journal of Geotechnical and Geoenvironmental Engineering,* 132(11), 1444-1454.

Yong, A, Hough, SE, Iwahashi, J, Braverman, A, (2012). A terrain-based site-conditions map of California with implications for the contiguous United States. *Bull. Seismol. Soc. Am.* 102 (1), 114–28.

Yong, A, Martin, A, Stokoe, K, Diehl, J, (2013). ARRA-funded vs30 measurements using multi-technique approach at strong-motion stations in California and central-eastern United States. *Reston, Virginia, U.S. Geol. Surv. Open-File Rept.* 2013-1102, 59

Zhan, J, Chen, G, Liu, J, (2009). Empirical relationship between shear wave velocity and soil depth on deep soft sites in urban area of Suzhou City. *World Earthq. Eng.* 25 (2), 11–7.

Zhao, D., Ruan, B., and Chen, G. (2017). Validation of the modified irregular loading-reloading rules based on Davidenkov skeleton curve and its equivalent shear strain algorithm implemented in ABAQUS. *Chinese Journal of Geotechnical Engineering*, 39(5): 888-895.

Zhang, J., Zhang, L.M. and Tang, W.H. (2010). Slope reliability analysis considering site-specific performance information. *Journal of geotechnical and geoenvironmental engineering*, 137(3), 227-238.