Bayesian Perspective on Ground Property Variability for Geotechnical Practice

Yu Wang[1], Tengyuan Zhao[2], and Zijun Cao[3]

[1]Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong.
E-mail: yuwang@cityu.edu.hk
[2]Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong.
E-mail: tengyzhao2@cityu.edu.hk
[3]State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, China.
E-mail: zijuncao@whu.edu.cn

Abstract: Soils and rocks are natural heterogeneous geo-materials, and their properties exhibit site-specific spatial variability as an outcome of the previous geological processes that the soils and rocks in the site have undergone. Spatial variability of ground properties and other geotechnical uncertainties may be modelled probabilistically using random variables or random field. Some questions are frequently raised by practicing geotechnical engineers when they consider using probabilistic methods. For example, what is the physical meaning of failure probability and random variable or random field modeling? Is a large amount of data necessary for using probabilistic methods? This paper aims at providing answers to these questions from a Bayesian perspective. Bayesian methods and tools are also presented that were recently developed for characterization of ground property variability from sparse site investigation data.

Keywords: Spatial Variability; uncertainty; random variable; random field; physical meaning; Bayesian statistics.

1 Introduction

One indispensable task in geotechnical practice is to characterize engineering properties of subsurface soils and rocks for geotechnical analysis and design. Soils and rocks are natural geo-materials, whose properties are affected by many spatially varying factors during the geological process, such as properties and textures of their parent materials, weathering and erosion processes, transportation agents, and sedimentation conditions (e.g., Lacasse and Nadim 1996; Phoon and Kulhawy 1999a,b; Baecher and Christian 2003). Soils and rocks are, therefore, heterogeneous, and ground properties exhibit site-specific spatial variability (e.g., Webster 2000; Wang et al. 2016a) as an outcome of the previous geological processes that the soils and rocks in the site have undergone. Spatial variability of ground properties and other geotechnical uncertainties may be modelled probabilistically using either a single random variable (SRV) or a series of spatially distributed and correlated random variables, i.e., a random field (RF). Both SRV and RF have been widely used in literature on geotechnical risk and reliability, and they were a key element in the calibration and development of new generation of geotechnical design codes around the world, such as the Load and Resistance Factor Design codes (Phoon et al. 2003a&b; Paikowsky et al. 2004&2010) in the United States, the Geo-Code 21 (JGS 2006; Honjo et al. 2010) in Japan, and the 2014 Canadian Highway Bridge Design Code in Canada (Fenton et al. 2016). However, direct application of probabilistic methods to routine projects by practicing geotechnical engineers seems rather limited, even though it is required in some national standards, such as the new safety standards for flood defenses in the Netherlands (e.g., Schweckendiek et al. 2015; Wang et al. 2016b). Some questions are frequently raised by practicing geotechnical engineers when they consider using probabilistic methods. For example, what is the physical meaning of failure probability and SRV or RF modeling? Is a large amount of data a must for using probabilistic methods? For facilitating geotechnical practitioners to adopt probabilistic methods in engineering practices, this paper aims at providing answers to these questions from a Bayesian perspective. Bayesian methods and tools are also presented that were recently developed for characterization of ground property variability from sparse site investigation data.

After this introduction, two different schools of probability (i.e., frequentist and Bayesian) are introduced, followed by a review of the general process of geotechnical site characterization, the occurrence of uncertainties and variabilities during geotechnical characterization of a site, and propagation of uncertainties during ground property characterization. Then, the physical meaning of SRV or RF modeling of spatial variability is provided from a Bayesian perspective. In addition, to deal with the issue of sparse data that are often encountered in geotechnical site investigation, Bayesian equivalent sample methods (Wang and Cao 2013a; Wang and Aladejare 2015) and Bayesian compressive sampling – Karhunen -Loève expansion method (Wang et al. 2018) are presented, respectively, for SRV and RF modeling of ground property spatial variability.

2 Two Schools of Probability: Frequentist vs. Bayesian

Proceedings of the 7th International Symposium on Geotechnical Safety and Risk (ISGSR)
Editors: Jianye Ching, Dian-Qing Li and Jie Zhang

Published by Research Publishing, Singapore.
ISBN: 978-981-11-2725-0; doi:10.3850/978-981-11-2725-0_key5-cd

Although probability and statistics are well-developed subjects, there are philosophically different schools of probability, the most prominent being the frequentist and Bayesian schools. Probability in frequentist theory is defined as the long-run or limiting expected frequency of occurrence. This is the one usually taught in most undergraduate courses and naturally linked with the classical examples of rolling a dice or flipping a coin. A large amount of data that satisfies the independent and identical distribution (IID) assumption is often needed for frequentist school. The frequentist definition of probability works well for some geotechnical applications, such as defining target failure probabilities, p_T, in geotechnical design codes. For example, a $p_T = 6.9 \times 10^{-4}$ (or a target reliability index of 3.2) was adopted for transmission line (and similar) structure foundations in North America (Phoon et al. 2003a,b). For a large number, saying 10,000, of foundation designs in 10,000 different projects (or sites) following this design code, it is expected to observe about 7 failure cases among these 10,000 projects. However, this frequentist definition may encounter difficulty in some other applications, such as interpretation of site-specific geotechnical investigation data. The site-specific geotechnical data are spatially varying and correlated, and they are often only sparsely measured. A large amount of IID data that is often needed for frequentist school simply does not exist in many problems that geotechnical engineers routinely face. The Bayesian school of probability may be useful in these cases.

In Bayesian school, probability represents a degree-of-belief of occurrence, with full certainty of occurrence or no-occurrence at the probability of 1 or 0, respectively. The probability is a quantitative measure of uncertainties associated with incomplete information or limited knowledge. As further information is gathered and the knowledge improves, the uncertainties might reduce, and the probability is updated. Therefore, Bayesian inference is often known as Bayesian updating. The degree-of-belief (i.e., the Bayesian school of probability) may be interpreted as a formalization of “engineering judgment” (e.g., Vick 2002; Cao et al. 2016a), which is a key element in geotechnical practice. From this perspective, most geotechnical engineers are intuitive Bayesians whether they know it or not (Baecher 2017). For example, Terzaghi’s observational method is a qualitative form of Bayesian thinking (Lacasse 2016). Bayesian methods have been used across a wide spectrum of geotechnical applications since 1970’s, such as Tang (1971), Einstein et al. (1978), Gilbert and Tang (1995), Juang et al. (1999), Zhang et al. (2004), Najjar and Gilbert (2009), Zhang et al. (2009), Wang et al. (2010), Papaioannou and Straub (2012), Ching and Phoon (2014), Wang et al. (2016a), and Ching et al. (2016). In addition, Baecher (2017) showed a number of successful Bayesian applications in geotechnical engineering. Early applications of Bayesian methods often relied on conjugate pairs of prior distributions and likelihood functions to avoid computational problems, but this nevertheless introduces artificial limitations to the choices of prior distributions and likelihood functions (e.g., Wang et al. 2010). In recent applications, numerical simulations, such as Markov Chain Monte Carlo (MCMC), are often integrated with Bayesian formulation to by-pass the computational complexity, while maintaining the user's flexibility to use non-conjugate prior distributions and likelihood functions for realistic modeling of the concerned applications (e.g., Wang and Cao 2013a). The simulation-based Bayesian methods may be further developed as computer software (e.g., Wang et al. 2016c) for removing mathematical hurdles and facilitating applications of Bayesian methods by geotechnical practitioners in engineering practice. This paper focuses on simulation-based Bayesian methods and tools that were recently developed for characterization of ground properties from sparse site investigation data.

3 Uncertainty and Variability in Geotechnical Characterization of a Specific Site

Geotechnical site characterization is a multi-step process that can be divided into six stages: desk-study, site reconnaissance, in-situ investigation, laboratory testing, interpretation of site observation data, and inferring soil and rock properties and underground stratigraphy (e.g., Clayton et al. 1995; Mayne et al. 2002). A general procedure of geotechnical site characterization is shown schematically in Figure 1. Geotechnical characterization of a project site often starts with desk-study and site reconnaissance, which provide prior knowledge about the site such as geological maps, regional guides, soil survey maps and records, published reports and studies, engineering judgment (or expertise) and/or engineering experience (e.g., Trautmann and Kulhawy 1983; Cao et al. 2016a). After desk-study and site reconnaissance, in-situ investigation (e.g., in-situ boring and testing) and laboratory testing are performed to obtain project-specific test results (i.e., site observation data) of a soil or rock property X_M measured by the tests. Note that the measured soil or rock property X_M is not necessarily the design property X_D that is used directly in geotechnical analysis. The design property X_D (e.g., undrained Young’s modulus E_u) might be estimated from the test data of X_M (e.g., the “*N*” values) obtained from in-situ/laboratory testing (e.g., standard penetration test, SPT) through a transformation model M_T (e.g., empirical regressions and theoretical relationships) between X_D and X_M. Engineers then utilize both the results interpreted from site-specific observation data and prior knowledge to estimate soil and rock properties for engineering applications.

As shown in Figure 2, both site-specific testing data and prior knowledge may be spatially variable and involve measurement errors, statistical uncertainty, and transformation uncertainty (e.g., Phoon and Kulhawy 1999b; Baecher and Christian 2003; Wang et al. 2016a). The spatial variability exists long before the engineering project is planned, and it is usually categorized as “natural variability”, which is independent of the state of knowledge about the soil and rock properties and cannot be reduced as the knowledge improves. On the other hand, measurement errors, statistical uncertainty, and transformation uncertainty are resulted from imperfect test equipment and/or procedural-operator errors, lack of test data, and insufficient knowledge about the relationship between X_D and X_M, respectively. They are usually categorized as "knowledge uncertainty" and can be reduced gradually as the knowledge improves.

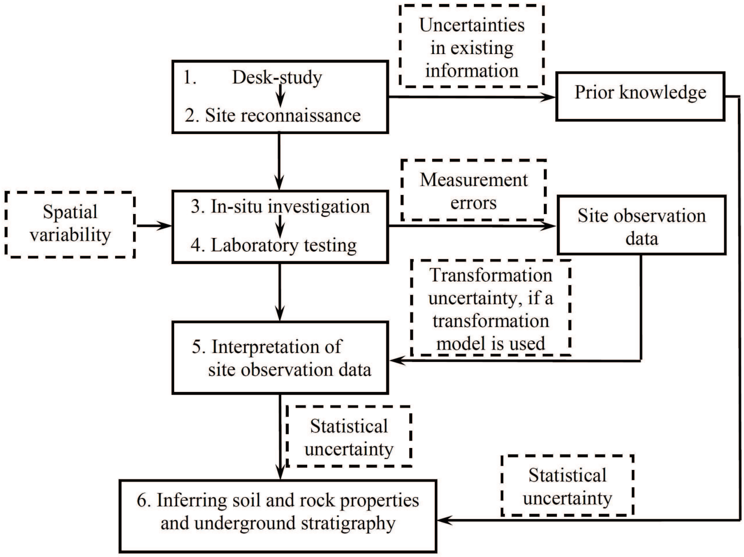

Figure 1. Geotechnical site characterization and uncertainty propagation (after Wang et al. 2016a).

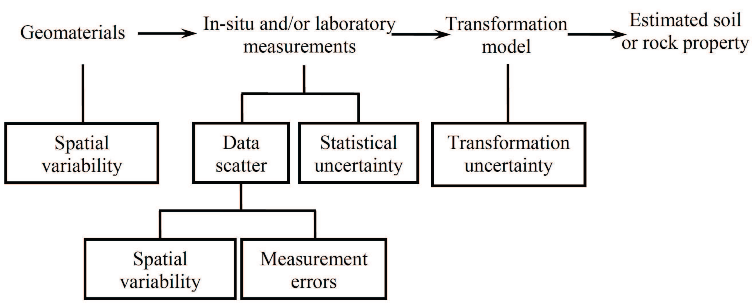

Figure 2. Variability and uncertainties in soil and rock properties (after Phoon and Kulhawy 1999a; Wang et al. 2016a).

4 Uncertainty Propagation during Ground Property Characterization

Figure 1 also shows explicitly propagation of the uncertainties, including spatial variability and various knowledge uncertainties, in ground properties during different stages of geotechnical characterization of a specific site. The prior knowledge is obtained during desk-study and site reconnaissance, and it is associated with spatial variability and measurement errors incorporated in existing data, subjective uncertainties of engineers' expertise, and so forth (e.g., Baecher 1983; Vick 2002). The site observation data fluctuate because of spatial variability and measurement errors.

When the design property X_D is not measured directly, a transformation model M_T is needed to relate the observation data (i.e., measured property X_M) to X_D. The uncertainty (i.e., transformation uncertainty) associated with the transformation model is then combined with the uncertainties included in site observation data (i.e., spatial variability and measurement errors), propagating together into the interpretation outcomes of the site observation data. Finally, engineers commonly utilize both the interpretation outcomes of site-specific observation data and prior knowledge to infer ground properties. Because the inferred ground properties are based on information at some locations or depths (i.e., some samples of the population), rather than information at every location and depth (i.e., the whole population), statistical uncertainty occurs and is added to the inferred ground properties.

The knowledge uncertainties (i.e., measurement errors, statistical uncertainty, and transformation uncertainty) arise from insufficient knowledge or lack of data during geotechnical site characterization. Although such knowledge uncertainties inevitably affect the prediction or results of reliability studies, they have no influence on the existing condition of the project site or the observed performance (i.e., actual response) of geotechnical structures. For example, the actual settlement of a foundation does not change no matter how accurately the soil and rock properties under the foundation are measured, estimated, and subsequently used in the analysis. Accurate estimates of soil and rock properties only contribute to the accuracy of predicted response of geotechnical structures, e.g., the predicted settlement of foundations. On the other hand, the spatial variability of ground properties exists long before the project and affects significantly the actual response of geotechnical structures (e.g., Fenton and Griffiths 2002; Wang et al. 2011; Liu et al. 2019). The knowledge uncertainties and spatial variability, therefore, should be differentiated in probabilistic analyses. As described in the next section, a Bayesian inverse analysis framework directly and explicitly characterizes spatial variability of design property and account simultaneously for various knowledge uncertainties in a rational manner.

5 Bayesian Perspective on Spatial Variability Modeling

Both prior knowledge and site-specific observation data are utilized during geotechnical site characterization to estimate ground properties. From a Bayesian perspective, such a process can be considered as a process of updating the prior knowledge using observation data from a project site, as shown in Figure 3. The prior knowledge obtained during desk-study and site reconnaissance is represented quantitatively by a prior distribution. The site observation data obtained from in-situ/laboratory testing is reflected through a likelihood function. After the prior distribution and likelihood function are specified, using the Bayes' Theorem (e.g., Sivia and Skilling 2006) leads to a posterior distribution that quantitatively reflects the updated knowledge (i.e., posterior knowledge) and combines the prior knowledge and site observation data. This process corresponds to the last two stages (i.e., Stages 5 and 6 shown in Figure 1) of geotechnical site characterization, where soil and rock properties are estimated using prior knowledge and site observation data. The six stages of geotechnical site characterization shown in Figure 1 are mapped to the three key elements (i.e., prior distribution, likelihood function, and posterior distribution) of the Bayesian framework, as shown in Figure 3. Because of the page limit, details of prior distribution quantification, likelihood function formulation, and MCMC simulation of the posterior distribution are not included in this paper, but referred to Cao et al. (2016a), Wang et al. (2016a), and Wang and Cao (2013a), respectively.

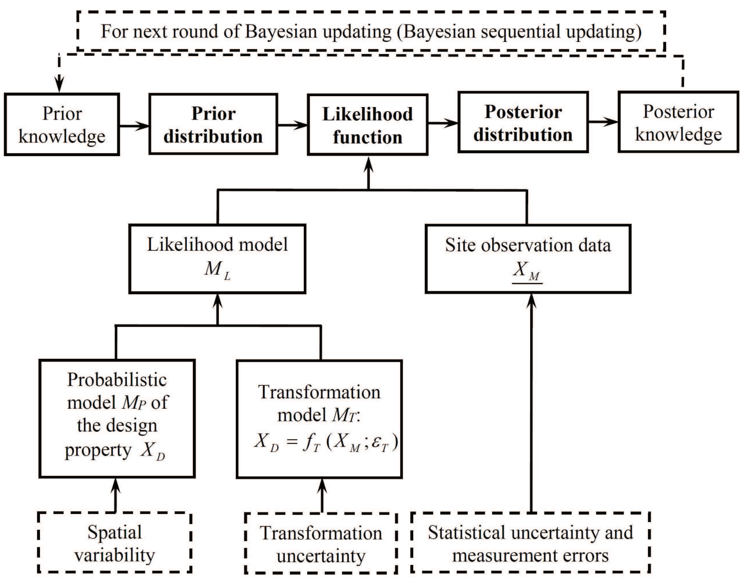

Figure 3. Bayesian framework for geotechnical characterization of a project site (after Wang et al. 2016a).

The Bayesian updating may be performed in a sequential manner when additional information on the ground properties is available (Cao et al. 2016b). For example, after the Bayesian updating is finished with existing site observation data, if additional or other types of in-situ or laboratory tests are subsequently performed to obtain additional site observation data, a new round of Bayesian updating can be performed to incorporate these additional data. The posterior knowledge obtained in the previous round of Bayesian updating can be taken as the prior knowledge and used together with the additional site observation data to further update the knowledge on ground properties, as shown by the dashed line in Figure 3. The Bayesian framework offers a systematic way of gradually improving knowledge of ground properties for a project site as observation data are accumulated.

As shown in Figure 3, a probabilistic model M_P (e.g., an SRV or RF) is needed to represent the spatial variability of X_D in the Bayesian framework. Consider, for example, a statistically homogenous soil layer. The spatial variability of X_D in the soil layer (i.e., the actual X_D value at any location within the soil layer) can be modelled by an SRV (e.g., Wang and Cao 2013a). An SRV can be uniquely defined by a probability distribution function, PDF, with its associated parameters. Note that PDF for a specific site may not necessarily follow any commonly used PDF, such as normal or lognormal PDF (Wang et al. 2015). When explicit consideration of spatial coordinates is needed, a RF $\underline{X_D}$ can be used to represent the spatial variability of X_D in the soil layer (e.g., Wang et al. 2018). A RF may be fully specified by its marginal PDF and spatial auto-correlation structures (Vanmarcke 1983; Cao and Wang 2014a). Bayesian quantification of spatial variability of ground properties is therefore reduced to characterization of SRV's PDF or RF's marginal PDF and spatial auto-correlation structures from prior knowledge and site-specific observation data, with simultaneous consideration of various knowledge uncertainties in a rational manner.

5.1 Physical meaning of single random variable modeling

In SRV modeling of spatial variability, spatial coordinates of the ground property are not considered explicitly. Therefore, even when the ground property is measured at every location, it is still a population or random variable following a PDF. From a Bayesian perspective, the missing information that leads to uncertainties in the SRV modeling is the spatial coordinates. Because the spatial coordinates have never been used in the Bayesian updating of the ground properties, SRV modeling of spatial variability is often categorized as "natural variability" from a frequentist perspective, which is independent of the state of knowledge about the ground properties and cannot be reduced as the knowledge improves (Baecher and Christian 2003). The SRV modeling of spatial variability is suitable for geotechnical applications where variation of spatial coordinates of the ground properties is not important, such as foundation design (e.g., Wang 2011; Wang and Cao 2013b). For example, side resistance of a pile foundation is a summation of unit pile side resistance along a pre-specified surface (i.e., the interface between the pile shaft and soils). However, the SRV model may not perform well when variation of spatial coordinates of the ground properties is important, such as slope stability analysis (e.g., Griffiths and Fenton 2004; Li et al. 2019; Liu et al. 2019) where a search for the minimum resistance path is needed from a large number of potential slip surfaces. In these cases, RF modeling of spatial variability, with explicit consideration of spatial coordinates, is preferred.

5.2 Physical meaning of random field modeling

In RF modeling of spatial variability, the spatial coordinates of ground properties are modelled explicitly. However, site-specific tests are only performed at a number, often limited, of locations. From a Bayesian perspective, the lack of information that leads to uncertainties in the RF modeling is the missing measurements at unsampled locations. Therefore, RF modeling of spatial variability reflects uncertainties in the prediction of ground property at unsampled locations from prior knowledge and the often sparse site-specific measurement data. If the ground property is measured at every location in a specific site, the spatial variability is fully specified and deterministic. No uncertainty in this case, and it is a fully certain case. The Bayesian RF modeling of spatial variability may be considered as an interpolation from sparse measurements with quantified interpretation uncertainty. This is similar to conditional RF modeling using Kriging. When measurements are taken at every location in a specific site, Kriging and conditional RF modeling converge to the measurements themselves with negligible uncertainty. It is like a deck of shuffled cards which has definite order although the shuffling process involves randomness (e.g., Baecher and Christian 2003). Drawing a card from a specific deck of shuffled cards produces a deterministic result even though it is unknown before the drawing. Note that, however, this is very different from the frequentist perspective of RF modeling, where the RF samples are still random even when measurements are taken at every location in a specific site. The Bayesian perspective of RF modeling is more suitable for modeling site-specific (or intra-site) spatial variability, while the frequentist perspective of RF modeling is more suitable for modeling inter-site spatial variability, or natural variability across different sites. Applications of RF model of spatial variability to development and calibration of geotechnical design codes may be considered as modeling of inter-site spatial variability.

In the following two sections, two recently developed Bayesian methods are presented for characterization of ground properties from sparse site-specific data: (1) MCMC-based Bayesian equivalent sample method for SRV modeling in Section 6; and (2) Bayesian Compressive Sampling and Karhunen–Loève Expansion method for RF modeling in Section 7.

6 Bayesian Equivalent Sample Method for Single Random Variable Modeling

Starting from the Bayes Theorem, Eq. (1) can be derived to update statistical parameters Θ_P (e.g., mean μ and standard deviation σ) of a design ground property X_D (e.g., soil effective friction angle ϕ'), which is treated as an SRV, given a set of site-specific test data as:

$$P(\Theta_P|Data, Prior) = K \cdot P(Data|\Theta_P)P(\Theta_P) \quad (1)$$

where K is a normalizing constant independent of the statistical parameters Θ_P of X_D; Data = $\underline{X_M}$ is the site-specific measurement data (e.g., a set of standard penetration test SPT-N values); $P(\Theta_P)$ is the prior distribution of the statistical parameters in the absence of site-specific measurement data; and $P(Data|\ \Theta_P)$=$P(\underline{X_M}|\Theta_P)$ is the likelihood function.

6.1 Bayesian formulation and MCMC simulation

The likelihood function $P(\underline{X_M}|\Theta_P)$ is a PDF of site-specific measurement data $\underline{X_M}$ for a given set of statistical parameters Θ_P. It quantifies probabilistically the Θ_P information provided by $\underline{X_M}$. Formulation of the likelihood function (i.e., $P(\underline{X_M}|\Theta_P)$) requires a likelihood model that probabilistically describes the relationship between the statistical parameters Θ_P of a design property X_D and project-specific test data $\underline{X_M}$. Generally speaking, the likelihood model should reflect sound physical insights into the relationship between the design property X_D and the measurement data $\underline{X_M}$ and the propagation of various uncertainties that occurred during site characterization. As much as possible insights from soil or rock mechanics should be incorporated in the likelihood model. For example, insights from soil mechanics suggest that undrained shear strength, S_u, of clay is not a fundamental soil property, but depends on the vertical effective stress, σ_v'. It is therefore a better likelihood model to consider S_u/σ_v' than S_u as an SRV (Cao and Wang 2014b). In addition, the design property X_D might not be measured directly, and a transformation or regression model is needed to relate $\underline{X_M}$ to X_D. The uncertainty (i.e., transformation uncertainty) associated with the transformation model should also be incorporated in the likelihood model. Based on the likelihood model, X_M (e.g., SPT-N value) can be derived as a random variable that has a (e.g., normal or lognormal) PDF (Wang et al. 2016a). Statistical parameters for the random variable X_M are a function of the statistical parameters Θ_P for the random variable X_D and the transformation uncertainty. This establishes a link between the site-specific measurement data $\underline{X_M}$ and the statistical parameters Θ_P for the design property X_D and allows the likelihood function to be formulated mathematically. Therefore, the statistical parameters Θ_P for X_D (e.g., μ and σ for the soil effective friction angle ϕ') can be updated from $\underline{X_M}$ (e.g., SPT-N values), as shown in Eq. (1).

Using the theorem of total probability, the posterior PDF of the design property X_D can be further expressed as (Wang and Cao 2013a; Wang et al. 2016a):

$$P(X_D\ |Data,\text{Prior}) = \int P(X_D|\Theta_P)P(\Theta_P|Data,Prior)d\Theta_P \quad (2)$$

where $P(X_D|\Theta_P)$ is the conditional (e.g., normal or lognormal) PDF of X_D for a given set of statistical parameters Θ_P (e.g., μ and σ); and $P(\Theta_P\ |Data,\text{Prior})$ is obtained from Eq. (1). When the prior knowledge and likelihood function in geotechnical practice are sophisticated, the X_D PDF might be complicated or difficult to express analytically or explicitly. To remove this mathematical hurdle in engineering practice, MCMC simulation (e.g., Robert and Casella 2004) is used to depict the X_D PDF numerically. The generated MCMC samples collectively reflect the posterior PDF of X_D (i.e., $P(X_D\ |Data,\text{Prior})$ in Eq. (2)), and they are referred to as Bayesian equivalent samples of the design property X_D (Wang and Cao 2013a). Note that although a normal or lognormal PDF is often used for $P(X_D|\Theta_P)$ because of the Gaussian assumption often used in the development of transformation or regression model, $P(X_D\ |Data,\text{Prior})$ in Eq. (2) is a weighted summation of many $P(X_D|\Theta_P)$, and it may turn out to be another distribution (Wang et al. 2015).

6.2 Computer software

Although the Bayesian framework described above is general and applicable for various soil or rock properties, its formulations vary for different properties when they are estimated from various in-situ and laboratory tests. For example, the formulation for Bayesian characterization of uniaxial compressive strength, UCS, of a rock using point load test data (e.g., Wang and Aladejare 2015) is different from the formulation for characterizing effective friction angle of soil using SPT-N values (e.g., Wang et al. 2015). Therefore, extensive backgrounds in probability, statistics, and simulation algorithms are needed to formulate the method for various properties. To remove this mathematical hurdle for geotechnical practitioners, a user-friendly Microsoft Excel-based toolkit, called Bayesian Equivalent Sample Toolkit (BEST), has been developed for implementing the Bayesian equivalent sample method (Wang et al. 2016c). The BEST was developed using the Visual Basic for

Applications (VBA) in a commonly available Microsoft Excel spreadsheet platform. The Excel-based BEST Add-in can be obtained without charge from https://sites.google.com/site/yuwangcityu/best/1.

6.3 Application examples

Bayesian equivalent sample method has been applied to various soil or rock properties (e.g., undrained shear strength (Cao and Wang 2014b) and Young's modulus (Wang and Cao 2013a) of clay, sand friction angle (Wang and Zhao 2017a), and UCS (Wang and Aladejare 2015), Geological Strength Index (Wang and Aladejare 2016a), and Young's modulus (Wang and Aladejare 2016b) of rock) and validated by real site data. Consider, for example, characterization of the undrained Young's modulus E_u of clay using SPT-N value data obtained from the clay site of the United States National Geotechnical Experimentation Sites (NGES) at Texas A&M University (Briaud 2000). A limited number of SPT-N values (i.e., 5 SPT-N values) were obtained within top stiff clay layer of the clay site, as illustrated in Figure 4(a). Figure 4(b) shows the results of 42 pressuremeter tests performed in the same clay layer (Briaud 2000) which are used for validating the BEST results.

The design property X_D in this example is the E_u of clay, and its corresponding measured data are the SPT-N values. A relationship between SPT-N and E_u of clay developed by Kulhawy and Mayne (1990) is used, since it is a built-in model in BEST Add-in. A set of non-informative prior knowledge is used, and it is taken as a joint uniform distribution with a mean of E_u varying between 5 MPa and 15 MPa and a standard deviation of E_u ranging from 0.5 MPa to 13.5 MPa. This set of prior knowledge is consistent with the typical ranges of E_u of clay reported in literature (e.g., Phoon and Kulhawy 1999a,b; Cao et al. 2016a). Using this set of prior knowledge and the 5 SPT-N values shown in Figure 4(a), BEST is executed to generate 30,000 equivalent samples of E_u. Conventional statistical analysis, such as calculation of mean and standard deviation and plotting histogram for PDF or cumulative distribution function, CDF, can be easily performed on the 30,000 equivalent samples using built-in functions in Excel. Table 1 shows statistics of the E_u samples obtained from BEST and their comparison with those obtained directly from the pressuremeter tests. Figures 5 and 6 show, respectively, the E_u PDF and CDF estimated from the BEST equivalent samples, together with the pressuremeter test results. The results from the Bayesian equivalent samples agree well with those obtained from 42 pressuremeter tests.

In addition, the Bayesian equivalent sample method may be extended to model joint probability distribution between two ground properties, such as effective cohesion and friction angle of soil (Wang and Akeju 2016) and uniaxial compressive strength and Young's modulus of rock (Wang and Aladejare 2016b).

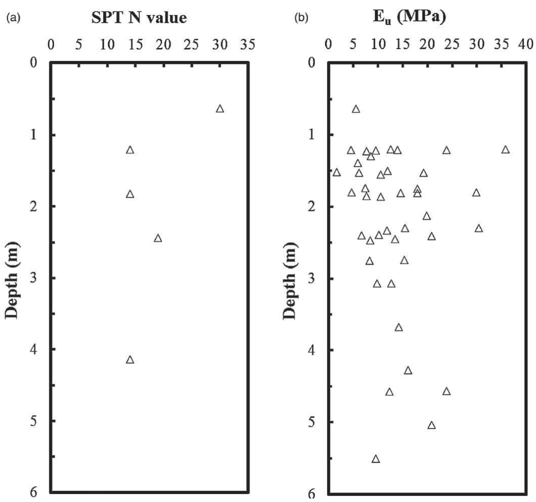

Figure 4. SPT-N values and undrained Young's modulus, E_u, measured by pressuremeter tests at the clay site of the NGES at Texas A&M University (after Briaud 2000).

Table 1. Summary of the E_u statistics.

Statistics	Mean (MPa)	Standard deviation (MPa)
BEST Excel Add-in	11.46	6.00
MATLAB (Wang and Cao 2013a)	11.60	6.00
Pressuremeter Tests	13.50	7.50
Difference between BEST and Pressuremeter Tests	2.04	1.50

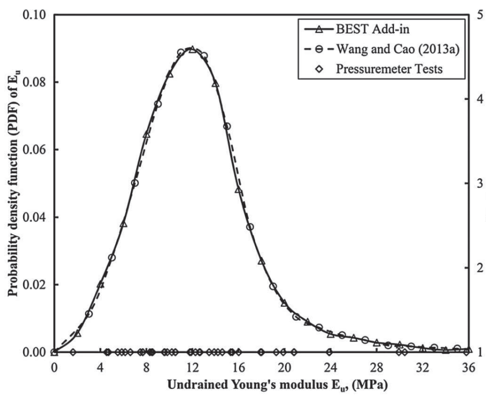

Figure 5. E_u PDF and frequency plots

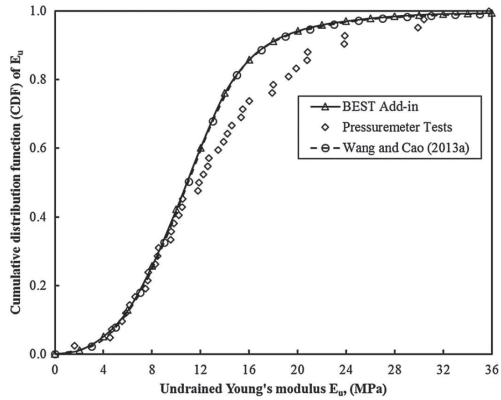

Figure 6. E_u CDF plot.

7 Random Field Modeling by Bayesian Compressive Sampling and Karhunen–Loève Expansion

Random field modeling of spatial variability requires two sets of parameters: marginal PDF for each random variable and spatial auto-correlation structures among random variables at different locations. Estimation of these random field parameters, particularly auto-correlation structures (e.g., auto-correlation function types and correlation length), requires extensive measurements. However, data gathered from geotechnical characterizations of a specific site are usually sparse, particularly for projects with small or medium sizes. Therefore, it is challenging to provide an accurate estimation on random field parameters from sparse measurements, leading to significant uncertainty in the random field samples (RFS) subsequently generated and used in reliability and risk analyses. To simulate RFSs with proper consideration of such uncertainty, a non-parametric and data-driven random field generator was recently developed. The generator is based on Bayesian compressive sampling (BCS) and Karhunen–Loève (KL) expansion, and it is denoted as BCS-KL generator.

Bayesian compressive sampling or sensing is the formulation of compressive sensing (CS) from a Bayesian perspective, and it aims to reconstruct a digital signal (e.g., variations of a quantity with time or space in a discrete form) from sparse measurement data on the digital signal and to quantify the uncertainty associated with the reconstruction. During the reconstruction, the compressibility in many physical signals is exploited (e.g., Wang and Zhao 2017b; Zhang et al. 2017). Compressibility means that a complete digital signal, $\boldsymbol{f}$, with a length of N, can be represented concisely by a weighted summation of a number of proper basis functions. Mathematically, $\boldsymbol{f} = \mathbf{B}\boldsymbol{\omega}$, where $\mathbf{B}$ is an N × N orthonormal matrix composed by columns of pre-specified basis functions and $\boldsymbol{\omega}$ is a vector of length N composed of the corresponding weight coefficients, most of which are virtually zero, except of S non-trivial components with significant magnitude, because of the compressibility of $\boldsymbol{f}$ (e.g., Foucart and Rauhut 2013; Wang and Zhao 2016). Therefore, $\boldsymbol{f}$ can be reconstructed from sparse measurement data $\boldsymbol{y}$ once the non-trivial components of $\boldsymbol{\omega}$ are obtained. $\boldsymbol{y}$ is a real-valued vector with a length of M, where M is much smaller than N. In the context of CS, $\boldsymbol{f}$ is unknown and needs to be estimated from sparse measurement data $\boldsymbol{y}$ using $\boldsymbol{y} = \boldsymbol{\Psi f} = \mathbf{A}\boldsymbol{\omega}$, where $\boldsymbol{\Psi}$ is an M × N matrix representing the locations of components $\boldsymbol{y}$ in $\boldsymbol{f}$. $\mathbf{A}=\boldsymbol{\Psi}\mathbf{B}$ is also an M × N matrix. $\boldsymbol{y} = \boldsymbol{\Psi f} = \mathbf{A}\boldsymbol{\omega}$ is an underdetermined system of linear equations that can be solved by various efficient algorithms, including Bayesian methods (e.g., Ji et al. 2008; Foucart and Rauhut 2013; Wang and Zhao 2017b). Once the non-trivial coefficients in $\boldsymbol{\omega}$ are properly estimated, the $\boldsymbol{\omega}$ can be approximated as $\boldsymbol{\omega}_S$ by setting those trivial elements of $\boldsymbol{\omega}$ as zero. Then, the reconstructed signal $\hat{\boldsymbol{f}}$ can be expressed as $\hat{\boldsymbol{f}} = \mathbf{B}\boldsymbol{\omega}_S$. When Bayesian methods are used to estimate $\boldsymbol{\omega}_S$, both the best estimate and covariance of $\hat{\boldsymbol{f}}$ (i.e., mean vector $\boldsymbol{\mu}_{\hat{f}}$ and covariance matrix $\mathbf{COV}_{\hat{f}}$) are obtained, as illustrated in Figure 7.

The BCS results can be used together with KL expansion to generate RFSs directly from sparse measurements (Wang et al. 2018). KL simulation of RFSs generally requires the mean of the random field of interest and deterministic orthogonal eigen-functions and eigenvalues corresponding to the covariance function or covariance matrix (Phoon et al. 2002). As shown in Figure 7, BCS provides both the best estimate (i.e., the mean of the random field) and the covariance matrix for the signal of interest directly from sparse measurements. When BCS and KL expansion is used together as referred to as BCS-KL generator, RFSs may be directly generated from sparse measurement data. The BCS-KL generator is non-parametric and data-driven. No pre-determined function forms are needed for marginal probability density function or covariance function of the random field. Therefore, the BCS-KL generator is readily applicable to non-Gaussian and non-stationary RFSs, including RFSs with non-stationary auto-covariance structure (Montoya-Noguera et al. 2019) and RFSs with unknown trend function without de-trending (Wang et al. 2019). In addition, the BCS-KL generator may be readily extended to simulate cross-correlated bivariate RFSs (Zhao and Wang 2018).

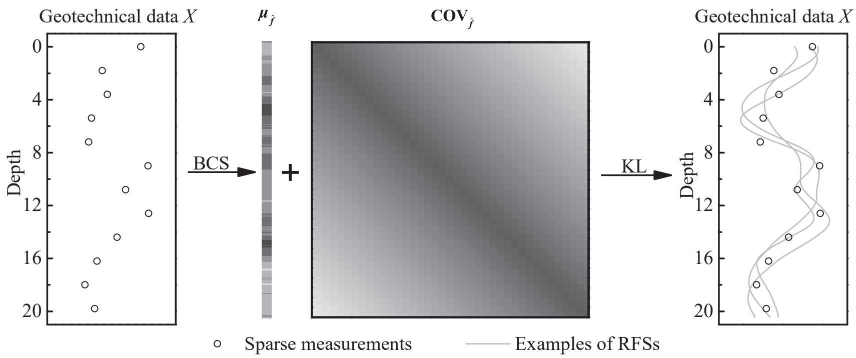

Figure 7. Schematic illustration of BCS-KL random field generator.

To illustrate the BCS-KL generator, a soil property *X* within a soil layer with a thickness of 20.44m is simulated from a random field with a constant mean of 30 and standard deviation of 2, together with an exponential auto-correlation structure, i.e., $\rho = e^{-2\Delta/\lambda c}$. Δ represents the relative difference of *X* data at different depths; λ_c =2m represents the correlation length within which *X* is highly auto-correlated. Using the random field parameters prescribed above, one *X* RFS with a resolution of 0.04m is shown in Figure 8a by a bold solid line. This *X* RFS (i.e., the solid line in Figure 8a) is denoted as the original *X* profile hereafter, which has N = 512 data points in total.

A number (e.g., M =20) of data points are obtained from the original *X* profile, as shown in Figure 8b by open triangles, and used as measurement data ***y*** to the BCS-KL generator. Then, many (e.g., 1000) *X* RFSs are generated accordingly. Figure 8b plots some *X* RFSs generated from the BCS-KL generator by gray lines. Figure 8b shows that the gray lines follow trends that are quite similar to the bold solid line, suggesting that the *X* RFSs generated from the BCS-KL generator is realistic and reasonable. Figure 8b also includes the statistics of the 1000 *X* samples, i.e., the average and 95% confidence interval (CI) of *X* at each depth, by a dashed line and two dotted lines respectively. In Figure 8b, the dashed line is very consistent with the bold solid line, and most local variations of the bold solid line falls within the 95% CI obtained from the 1000 *X* RFSs. These agreements suggest that the BCS-KL generator performs reasonably well and generates realistic RFSs. Figure 9 plots the experimental auto-covariance function (AF) of the original *X* profile (i.e., the bold solid line in Figure 8) and the statistics (i.e., the average and 95% CI) of the AFs from the 1000 *X* RFSs simulated from ***y***. Evidently, the dashed line (i.e., the averaged AF of the *X* samples) is in very good agreement with the bold solid line (i.e., the AF of the original *X* profile). Almost all variations of the bold solid line fall within the 95% CI of AF.

In addition, as the number M of measurements increases, the *X* RFSs generated from the BCS-KL generator gradually converge original *X* profile, as shown in Figures 8c and 8d. The corresponding AF also converges to the AF of the original *X* profile, which is not reported here due to page limit. The convergence of RFSs with increasing M is consistent with the Bayesian RF modeling of site-specific spatial variability. The BCS-KL generator has also been tested with a large number of real CPT data (Wang et al. 2018). Because of page limit, performance of the BCS-KL generator on real CPT data is not reported here, but referred to Wang et al. (2018).

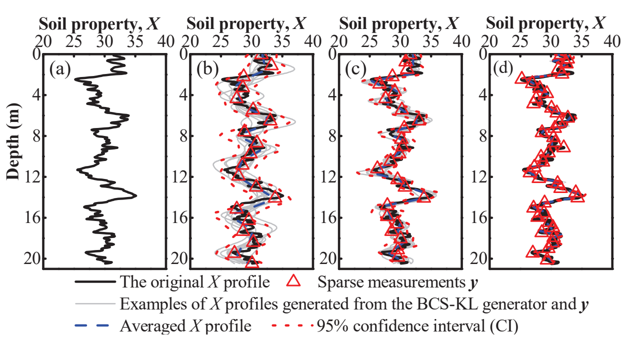

Figure 8. Original *X* profile (a) and *X* RFSs generated from ***y***: (b) M = 20; (c) M = 50, and (d) M = 250

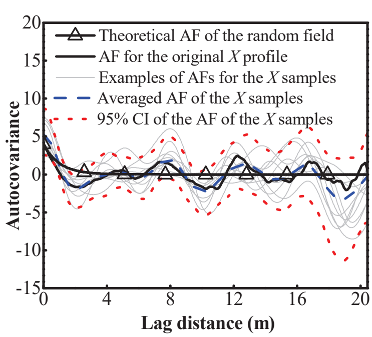

Figure 9. Comparison of AF for *X* RFSs.

8 Concluding Remarks

Bayesian thinking is in line with geotechnical thinking, particularly when dealing with site-specific data or analysis. Most geotechnical practitioners behave as Bayesians, even if they are not aware or deny. A large amount of data is not needed when using Bayesian methods. Strong Bayesian inference can be obtained from prior knowledge (e.g., engineering judgment, data from other sites with similar geological settings) and sparse and incomplete data, which are frequently encountered in geotechnical practices. From a Bayesian perspective, spatial coordinates are the missing information that leads to uncertainties in the SRV modeling of site-specific ground property variability. Because the spatial coordinates have never been used in the Bayesian updating of the ground properties, SRV modeling of spatial variability may also be categorized as "natural variability" from a frequentist perspective. On the other hand, in RF modeling of site-specific spatial variability, the spatial coordinates of ground properties are modelled explicitly. However, site-specific tests are only performed at a number, often limited, of locations. From a Bayesian perspective, the lack of information that leads to uncertainties in the RF modeling is the missing measurements at unsampled locations. Therefore, RF modeling of spatial variability in a specific site reflects uncertainties in the prediction of ground property at unsampled locations from prior knowledge and the often sparse site-specific measurement data. If the ground property is measured at every location in a specific site, the spatial variability is fully specified and deterministic. This is consistent with conditional RF modeling and Kriging in geostatistics. For facilitating geotechnical practitioners to adopt probabilistic methods in engineering practices, Bayesian equivalent sample method and BCS-KL random field generator have been developed for characterization of ground property variability from sparse site investigation data.

Acknowledgments

The work described in this paper was supported by grants from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project Nos. 8779012 (T22–603/15N), 9042172 (CityU 11200115), and 9042516 (CityU 11213117)). The financial support is gratefully acknowledged. The first author is grateful to several former PhD students who worked together on this Bayesian topic in the last ten years, including Drs Zijun Cao, Kai Huang, Adeyemi Emmna Aladejare, Olwatosin Victor Akeju, and Tengyuan Zhao. He also deeply appreciates many inspiring discussions on this topic with Professors Ivan Siu-Kui Au and Kok Kwang Phoon. The authors are very grateful to Professor Gregory B. Baecher for his thorough and enlightening comments and suggestions on this paper.

References

Baecher, G.B. (1983). Professional judgment and prior probabilities in engineering risk assessment. *Proceedings of Fourth International Conference on Application of Statistics and Probability in Soil and Structural Engineering*, Universita di Firenze (Italy), Pitagora Editrice, 635-650.

Baecher, G.B. (2017). Bayesian thinking in geotechnics. *Geo-Risk 2017*, *Geotechnical Special Publication,* GSP 282, 1-18.

Baecher, G.B. and Christian, J.T. (2003). *Reliability and Statistics in Geotechnical Engineering*, Wiley, Chichester, U.K.

Briaud, J.L. (2000). The national geotechnical experimentation sites at Texas A&M University: Clay and sand, a summary. National Geotechnical Experimentation Sites, *Geotechnical Special Publication*, GSP 93, 26–51.

Cao, Z.J. and Wang, Y. (2014a). Bayesian model comparison and selection of spatial correlation functions for soil parameters. *Structural Safety*, 49, 10-17.

Cao, Z.J. and Wang, Y. (2014b). Bayesian model comparison and characterization of undrained shear strength. *Journal of Geotechnical and Geoenvironmental Engineering*, 140(6), 04014018, 1-9.

Cao, Z.J., Wang, Y., and Li, D. (2016a). Quantification of prior knowledge in geotechnical site characterization. *Engineering Geology*, 203, 107-116.

Cao, Z.J., Wang, Y., and Li, D. (2016b). Site-specific characterization of soil properties using multiple measurements from different test procedures at different locations – A Bayesian sequential updating approach. *Engineering Geology*, 211, 150-161.

Ching, J.Y. and Phoon, K.K. (2014) Transformations and correlations among some clay parameters – the global database. *Canadian Geotechnical Journal*, 51(6), 663-685.

Ching, J.Y., Wu, S.S., and Phoon, K.K. (2016). Statistical characterization of random field parameters using frequentist and Bayesian approaches, *Canadian Geotechnical Journal*, 53(2), 285-298.

Clayton, C.R.I., Matthews. M.C., Simons, N.E. (1995). *Site Investigation*, Blackwell Science, Cambridge, Mass., USA.

Einstein, H.H., Labreche, D.A., Markow, J.J., and Baecher, G.B. (1978). Decision analysis applied to rock exploration. *Engineering Geology*, 12(2), 143–161.

Fenton, G.A. and Griffiths, D.V. (2002). Probabilistic foundation settlement on spatially random soil. *Journal of Geotechnical and Geoenvironmental Engineering*, 128(5), 381-390.

Fenton, G.A., Naghibi, F., Dundas, D., Bathurst, R.J., and Griffiths, D.V. (2016). Reliability-based geotechnical design in 2014 Canadian Highway Bridge Design Code. *Canadian Geotechnical Journal*, 53(2), 236-251.

Foucart, S. and Rauhut, H. (2013). *A Mathematical Introduction to Compressive Sensing*, Birkhäuser Basel, Boston.

Gilbert, R.B. and Tang, W.H. (1995). Model uncertainty in offshore geotechnical reliability. *Proceeding of the 27th Offshore Technology Conference, Society of Petroleum Engineers*, Houston, 557-567.

Griffiths, D.V and Fenton, G.A. (2004). Probabilistic slope stability analysis by Finite Elements. *Journal of Geotechnical and Geoenvironmental Engineering*, 130(5), 507 - 518.

Honjo, Y., Kikuchi, Y., and Shirato, M. (2010). Development of the design codes grounded on the performance-based design concept in Japan. *Soils and Foundations*, 50(6), 983-1000.

Japanese Geotechnical Society (2006). *Principles for Foundation Designs Grounded on a Performance-based Design Concept*, JGS 4001-2004.

Ji, S., Xue, Y., and Carin, L. (2008). Bayesian compressive sensing. *IEEE Transactions Signal Processing*, 56, 2346-2356.

Juang, C.H., Rosowsky, D.V., and Tang, W.H. (1999). Reliability-based method for assessing liquefaction potential of soils. *Journal of Geotechnical and Geoenvironmental Engineering*, 125(8), 684–689.

Kulhawy, F.H. and Mayne, P.W. (1990). *Manual on Estimating Soil Properties for Foundation Design*, Rep. EL-6800, Electric Power Research Institution, Palo Alto, CA.

Lacasse, S. (2016). Hazard, reliability and risk assessment - Research and practice for increased safety. *Proceedings of the 17th Nordic Geotechnical Meeting Challenges in Nordic Geotechnics*, Icelandic Geotechnical Society, Reykjavik, 17–42.

Lacasse, S. and Nadim, F. (1996). Uncertainties in characterizing soil properties. *Uncertainty in the Geologic Environment: from Theory to Practice*, Geotechnical Special Publication No. 58 (I), 49–75.

Li, L., Wang, Y., Zhang, L., Choi, C., and Ng, C.W.W. (2019). Evaluation of critical slip surface in limit equilibrium analysis of slope stability by smoothed particle hydrodynamics. *International Journal of Geomechanics*, ASCE, 19(5), 04019032.

Liu, X., Wang, Y., and Li, D. (2019). Investigation of slope failure mode evolution during large deformation in spatially variable soils by random limit equilibrium and material point methods. *Computers and Geotechnics*, 111, 301-312.

Mayne, P.W., Christopher, B.R., and DeJong, J. (2002). *Subsurface Investigations – Geotechnical Site Characterization*, No. FHWA NHI-01-031, Federal Highway Administration, U.S. Department of Transportation, Washington D.C.

Montoya-Noguera, S., Zhao, T., Hu, Y., Wang, Y., and Phoon, K.K. (2019). Simulation of non-stationary non-Gaussian random fields from sparse measurements using Bayesian compressive sampling and Karhunen-Loève expansion. *Structural Safety*, 79, 66-79.

Najjar, S.S. and Gilbert, R.B. (2009). Importance of lower-bound capacities in the design of deep foundations. *Journal of Geotechnical and Geoenvironmental Engineering,* 135 (7), 890–900.

Paikowsky, S.G., Birgisson, B., MvVay, M., Nguyen, T., Kuo, C., Baecher, G., Ayyub, B., Stenersen, K., O'Malley, K., Chernauskas, L., and O'Neill, M. (2004). Load and resistance factor design (LRFD) for deep foundations. *National Cooperative Highway Research Program (NCHRP) Report 507*, Transportation Research Board, National Research Council, Washington, D.C.

Paikowsky, S.G., Canniff, M.C., Lesny, K., Kisse, A., Amatya, S., and Muganga, R. (2010). LRFD design and construction of shallow foundations for highway bridge structures. *National Cooperative Highway Research Program (NCHRP) Report 651*, Transportation Research Board, National Research Council, Washington, D.C.

Papaioannou, I. and Straub, D. (2012). Reliability updating in geotechnical engineering including spatial variability of soil. *Computers and Geotechnics*, 42, 44-51.

Phoon, K.K., Huang, S.P., and Quek, S.T. (2002). Simulation of second-order processes using Karhunen–Loève expansion. *Computers and Structures*, 80(12), 1049–1060.

Phoon, K.K. and Kulhawy, F.H. (1999a). Characterization of geotechnical variability. *Canadian Geotechnical Journal*, 36(4), 612–624.

Phoon, K.K. and Kulhawy, F.H. (1999b). Evaluation of geotechnical property variability. *Canadian Geotechnical Journal*, 36(4), 625–639.

Phoon, K.K., Kulhawy, F.H., and Grigoriu, M.D. (2003a). Development of a reliability-based design framework for transmission line structure foundations. *Journal of Geotechnical and Geoenvironmental Engineering*, 129(9), 798-806.

Phoon, K.K., Kulhawy, F.H., and Grigoriu, M.D. (2003b). Multiple resistance factor design (MRFD) for shallow transmission line structure foundations. *Journal of Geotechnical and Geoenvironmental Engineering*, 129(9), 807-818.

Robert, C. and Casella, G. (2004). *Monte Carlo Statistical Methods*, Springer, United States.

Schweckendiek, T., Slomp, R., and Knoeff, H. (2015). New safety standards and assessment tools in the Netherlands. *Proceedings of the 5th Siegener Symposium "Sicherung von Dämmen, Deichen und Stauanlagen"*, Siegen, Germany, 19-20 February, 2015.

Sivia, D.S. and Skilling, J. (2006). *Data Analysis: A Bayesian Tutorial*, Oxford University Press, New York.

Tang, W.H. (1971). A Bayesian evaluation of information for foundation engineering design. *Proceedings of the 1st International Conference on Applied Statistics and Probability to Soil and Structure Engineering*, Hong Kong, 174-185.

Trautmann, C.H. and Kulhawy, F.H. (1983). Data sources for engineering geologic studies. *Bulletin of the Association of Engineering Geologists*, 20(4), 439-454.

Vanmarcke, E.H. (1983). *Random Fields: Analysis and Synthesis*, MIT Press, Cambridge, USA.

Vick, S.G. (2002). *Degrees of Belief: Subjective Probability and Engineering Judgment*. Amer Society of Civil Engineers.

Wang, Y. (2011). Reliability-based design of spread foundations by Monte Carlo Simulations. *Geotechnique*, 61(8), 677-685.

Wang, Y. and Akeju, O.V. (2016). Quantifying the cross-correlation between effective cohesion and friction angle of soil from limited site-specific data. *Soils and Foundations*, 56(6), 1057–1072.

Wang, Y., Akeju, O.V., and Cao, Z. (2016c). Bayesian Equivalent Sample Toolkit (BEST): an Excel VBA program for probabilistic characterization of geotechnical properties from limited observation data. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 10(4), 251-268.

Wang, Y. and Aladejare, A.E. (2015). Selection of site-specific regression model for characterization of uniaxial compressive strength of rock. *International Journal of Rock Mechanics and Mining Sciences*, 75, 73-81.

Wang, Y. and Aladejare, A.E. (2016a). Evaluating variability and uncertainty of Geological Strength Index at a specific site. *Rock Mechanics and Rock Engineering*, 49(9), 3559–3573.

Wang, Y. and Aladejare, A.E. (2016b) Bayesian characterization of correlation between uniaxial compressive strength and Young's modulus of rock. *International Journal of Rock Mechanics & Mining Sciences*, 85, 10-19.

Wang, Y., Au, S.K., and Cao, Z.J. (2010). Bayesian approach for probabilistic characterization of sand friction angles. *Engineering Geology*, 114(3-4), 354-363.

Wang, Y. and Cao, Z.J. (2013a). Probabilistic characterization of Young's modulus of soil using equivalent samples. *Engineering Geology*, 159, 106-118.

Wang, Y. and Cao, Z.J. (2013b). Expanded reliability-based design of piles in spatially variable soil using efficient Monte Carlo simulations. *Soils and Foundations*, 53(6), 820–834.

Wang, Y., Cao, Z.J., and Au, S.K. (2011). Practical analysis of slope stability by advanced Monte Carlo Simulation in spreadsheet. *Canadian Geotechnical Journal*, 48(1), 162-172.

Wang, Y., Cao, Z.J., and Li, D. (2016a). Bayesian perspective on geotechnical variability and site characterization. *Engineering Geology*, 203, 117-125.

Wang, Y., Schweckendiek, T., Gong, W., Zhao, T., and Phoon, K. K. (2016b). Direct probability-based design methods. *Reliability of Geotechnical Structures in ISO2394*, 193-228, edited by K.K. Phoon and J.V. Retief, CRC Press.

Wang, Y. and Zhao, T. (2016). Interpretation of soil property profile from limited measurement data: a compressive sampling perspective. *Canadian Geotechnical Journal*, 53(9), 1547-1559.

Wang, Y. and Zhao, T. (2017a). Bayesian assessment of site-specific performance of geotechnical design charts with unknown model uncertainty. *International Journal for Numerical and Analytical Methods in Geomechanics*, 41(5), 781-800.

Wang, Y. and Zhao, T. (2017b). Statistical interpretation of soil property profiles from sparse data using Bayesian compressive sampling. *Géotechnique*, 67, 523-536.

Wang, Y., Zhao, T., and Cao, Z. (2015). Site-specific probability distribution of geotechnical properties. *Computers and Geotechnics*, 70, 159-168.

Wang, Y., Zhao, T., Hu, Y., and Phoon, K.K. (2019). Simulation of random fields with trend from sparse measurements without detrending. *Journal of Engineering Mechanics*, ASCE, 145(2), 04018130.

Wang, Y., Zhao, T., and Phoon, K.K. (2018). Direct simulation of random field samples from sparsely measured geotechnical data with consideration of uncertainty in interpretation. *Canadian Geotechnical Journal*, 55(6), 862-880.

Webster, R. (2000). Is soil variation random? *Geoderma*, 97(3-4), 149-163.

Zhang, J., Zhang, L.M., and Tang, W.H. (2009). Bayesian framework for characterizing geotechnical model uncertainty. *Journal of Geotechnical and Geoenvironmental Engineering*, 135(7), 932-940.

Zhang, L.M., Tang, W.H., Zhang, L.L., and Zheng, J.G. (2004). Reducing uncertainty of prediction from empirical correlations. *Journal of Geotechnical and Geoenvironmental Engineering*, 130(5), 526-534.

Zhang, Y., Comerford, L., Kougioumtzoglou, I.A., Patelli, E., and Beer, M. (2017). Uncertainty quantification of power spectrum and spectral moments estimates subject to missing data. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems*, Part A: Civil Engineering, 3(4), 04017020.

Zhao, T., Montoya-Noguera, S., Phoon, K.K., and Wang, Y. (2018). Interpolating spatially varying soil property values from sparse data for facilitating characteristic value selection. *Canadian Geotechnical Journal*, 55(2), 171-181.

Zhao, T. and Wang, Y. (2018). Simulation of cross-correlated random field samples from sparse measurements using Bayesian compressive sensing. *Mechanical Systems and Signal Processing*, 112, 384-400.