# PROBABILISTIC ANALYSES OF SLOPES AND FOOTINGS WITH SPATIALLY VARIABLE SOILS CONSIDERING CROSS-CORRELATION AND CONDITIONED RANDOM FIELD

M.K. Lo[1], and Y.F. Leung[2]

## ABSTRACT

This paper presents probabilistic analyses of slopes and strip footings, with spatially variable soil modeled by the random field theory. Random fields are simulated using Latin hypercube sampling with dependence (LHSD), which is a stratified sampling technique that preserves the spatial autocorrelation characteristics. LHSD is coupled with polynomial chaos expansion (PCE) to approximate the probability density function of model response. The LHSD-PCE approach is applied to probabilistic slope analyses for soils with cross-correlated shear strength parameters, and is shown to be more robust than raw Monte Carlo simulations, even with much smaller numbers of model simulations. The approach is then applied to strip footing analyses with conditioned random fields of Young's modulus and shear strength parameters, to quantify the reductions in settlement uncertainty when soil samples are available at different depths underneath the footing. The most influential sampling depth is found to vary between 0.25 to 1 times the footing width, depending on the strength mobilization and spatial correlation features. Design charts are established with practical guidelines for quick estimations of uncertainty in footing settlements.

**Keywords:** probabilistic analyses, conditioned random field, sampling location, shallow foundation, slope stability

[1]Ph.D. Candidate, Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong.

[2]Assistant Professor, Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong. E-mail: yfleung@polyu.edu.hk

## INTRODUCTION

Variability in soil and rock properties and its potential impacts on geotechnical performance have been well recognized in the profession. For example, Phoon and Kulhawy (1999a), Phoon and Kulhawy (1999b) and Baecher and Christian (2003) presented detailed discussions on the characterization and evaluation of geotechnical variability, while their spatial correlation features have also been studied earlier by Vanmarcke (1977), Soulié et al. (1990), DeGroot (1996), and Zhang and Dasaka (2010), among others. Meanwhile, consideration of such variability in geotechnical applications have been investigated by a number of researchers, including Li and Lumb (1987), Christian et al. (1994), Griffiths et al. (2009), Cho (2010) and Jiang et al. (2015), who focused on slope stability analysis; while Duncan (2000), Griffiths and Fenton (2009), Kasama and Whittle (2011), Al-Bittar and Soubra (2014) and Li et al. (2015) discussed its influence on the performance of shallow foundations. Some typical approaches include first-order second-moment (FOSM) methods, the Hasofer-Lind approach (also known as the first-order reliability method) (Hasofer and Lind 1974) and the random finite element method (RFEM).

The total uncertainty associated with geotechnical processes is often separated between aleatory (natural variation) and epistemic (limited knowledge) uncertainty. To account for the former source, random field modelling can be adopted where variations of properties are represented by the spatial correlation structure. Meanwhile, since the actual pattern of variations is unknown from limited soil samples, normally a large number of random field realizations are required to account for such epistemic source of uncertainty.

The Monte Carlo simulation technique, implemented through RFEM, has become a popular approach to study the influence of soil variability on geotechnical performance. To improve the accuracy of the Monte Carlo estimator, a stratified sampling scheme, known as Latin hypercube sampling (LHS), is often applied in the implementation. LHS was originally proposed by McKay et al. (1979), with the fundamental assumption that components of the vector of random variables are independent of each other. However, this does not correspond to most scenarios of spatially-correlated soil properties. To resolve this issue, Cho (2010) represented the random process by the Karhunen-Loève expansion in terms of uncorrelated (orthonormal) random variables, while Jiang et al. (2015) applied LHS in the polynomial chaos expansion (Wiener 1938; Ghanem and Spanos 1991) of the system response.

This paper presents the application of a stratified sampling technique called the Latin hypercube sampling with dependence (LHSD), recently developed by Packham and Schmidt (2010). LHSD preserves the covariance structure of random variables, and is extended to simulate cross-correlated and conditioned random fields in the current study. Probabilistic assessments of slopes and strip footings are presented to demonstrate the capabilities of the approach, and in the latter case, to quantify the importance of various sampling locations through conditioned random field simulations. Comparisons between conditioned and unconditioned models further advance the understanding on the influence of the epistemic source of uncertainty. These lead to the development of design charts for quick assessments of uncertainties in footing settlements according to the locations of sampled points.

## LATIN HYPERCUBE SAMPLING WITH DEPENDENCE (LHSD)

There have been a number of previous attempts (Vanmarcke 1984) to model soil properties as spatially-correlated random variables. In general, the spatial variability of geotechnical properties can be represented by the following general linear model:

$$\boldsymbol{z} = \boldsymbol{\mu} + \boldsymbol{e} \tag{1}$$

where $\boldsymbol{z}$ is the vector representing soil properties at various locations, and $\boldsymbol{\mu}$ is the deterministic trend. When there is no prevalent trend in the data, $\boldsymbol{\mu}$ will become a constant mean vector. $\boldsymbol{e}$ is a random (residual) vector with mean $\mathbf{0}$ and spatial covariance matrix $\mathbf{V}$. $\boldsymbol{z}$ is said to be a Gaussian random field if the distribution of $\boldsymbol{e}$ is multivariate Gaussian. The random field is usually assumed to be second-order stationary, meaning that the variance $\sigma^2$ of $\boldsymbol{e}$ is constant across the domain, and hence the covariance matrix can be factored as

$\mathbf{V} = \sigma^2\mathbf{R}$, with $\mathbf{R}$ defined as the spatial correlation matrix. Methods of inferring the $\boldsymbol{\mu}$, variance ($\sigma^2$), and spatial correlation ($\mathbf{R}$) from site investigation data have been discussed by DeGroot (1996) and Liu et al. (2017).

Engineers are often interested in assessing the system reliability, or the probability of failure due to geotechnical variability. In this case, a limit state function $D(\boldsymbol{z})$ can be defined where $D(\boldsymbol{z}) > 0$ represents failure of the system, and the failure zone refers to the corresponding set of $\boldsymbol{z}$. The term 'failure' here is used in a broad sense, which does not necessarily mean a collapse, but can also refer to inadequate factors of safety or occurrence of excessive displacements. The failure probability, $P_f$, is the area of the joint probability distribution of $\boldsymbol{z}$ (denoted as $f(\boldsymbol{z})$) in the failure zone. $P_f$, and its Monte Carlo estimator, $\hat{P}_f$, can be represented by:

$$P_f = \int_{\boldsymbol{z}} \mathrm{I}\left[D(\boldsymbol{z}) > 0\right] f(\boldsymbol{z}) d\boldsymbol{z} \tag{2a}$$

$$\hat{P}_f = \frac{1}{n}\sum_{i=1}^{n} \mathrm{I}\left[D(\boldsymbol{z}_i) > 0\right] = \frac{n_f}{n} \tag{2b}$$

where $n_f$ is the number of failure cases within $n$ simulations; I is an indicator function which equals 1 when $D(\boldsymbol{z}) > 0$ and 0 otherwise. In the current study, $D(\boldsymbol{z}) > 0$ or factor of safety (FS) $< 1$ refers to stability failure. In practice, when $P_f$ is small, a large number of realizations are required to achieve a satisfactory accuracy. For example, in order to reduce the maximum estimation error to 0.01, Griffiths et al. (2009) used 2,000 simulations to evaluate the slope failure probability. To improve the robustness of $\hat{P}_f$ estimation and reduce number of simulations, LHS was proposed as a multi-dimensional stratified sampling scheme, which ensures a uniform placement of random realizations in the sample domain. However, LHS requires uncorrelated random variables. Although there are transformation techniques to introduce spatial dependence into these samples, Packham (2015) noted that such operation would damage the original stratification and effectiveness of LHS.

Packham and Schmidt (2010) proposed the Latin Hypercube Sampling with Dependence (LHSD), which is an extension of LHS and aims to introduce stratification while maintaining the covariance structure. In addition, LHSD offers two critical advantages. It can be applied to any covariance structure, and its implementation does not depend on the specific geotechnical problem. These enable LHSD to be a robust and flexible algorithm for a wide range of geotechnical applications, as will be illustrated in this paper through analyses on slopes and shallow foundations.

LHSD ensures a uniform placement of random realizations in a $d$-dimensional unit cube. In each dimension, a permutation has to be performed to decide which sample is placed into which stratum. The core concept of LHSD is that the permutation for a particular dimension is calculated using the rank statistic of the simulated samples. For example, if the simulated samples in dimension $j$ is $\boldsymbol{u}^{(j)} = u_1^{(j)}, u_2^{(j)}, \ldots, u_n^{(j)}$, the permutation ($\boldsymbol{r}^{(j)}$) can be obtained by the following rank statistic:

$$r_i^{(j)} = \sum_{k=1}^{n} \mathrm{I}\left[u_k^{(j)} \leq u_i^{(j)}\right] \tag{3}$$

with I being an indicator function, which returns 1 if $u_k^{(j)} \leq u_i^{(j)}$ and 0 otherwise. Through the permutation $\boldsymbol{r}^{(j)}$, $\boldsymbol{u}$ can be converted into the LHSD sample, $\boldsymbol{v}$. Fig. 1 shows an example of the conversion from $\boldsymbol{u}$ to $\boldsymbol{v}$, and details of the implementation will be illustrated with an application on Gaussian random field in the next section.

In statistical terms, LHSD can be applied to a random vector, with distribution of each component being standard uniform, i.e. U(0,1). The property of the LHSD samples has been proven, and a stated key property (Packham and Schmidt 2010) is that the empirical cumulative distribution of the LHSD sample converges to its theoretical cumulative distribution, given the sample size $n$ is sufficiently large.

## Application to Gaussian and transformed random fields

The distribution of a Gaussian random field is obviously not standard uniform. However, any random vector $\boldsymbol{z}$ (which may represent Gaussian random field of geotechnical properties) with arbitrary distribution can be transformed into a random vector $\boldsymbol{u}$ with uniform distribution, by the use of cumulative distribution functions (CDF). In theory, to apply LHSD to an arbitrary distribution, it is possible to first transform the simulated $\boldsymbol{z}$ into the standard uniform sample $\boldsymbol{u}$, obtain $\boldsymbol{v}$ and then back-transform $\boldsymbol{v}$ into the LHSD sample of $\boldsymbol{z}$. In fact, Packham and Schmidt (2010) stated that if all dimensions have the same distribution, then the rank statistic $\boldsymbol{r}$ computed using the original sample ($\boldsymbol{z}$) will be equal to $\boldsymbol{r}$ computed using the corresponding standard uniform sample ($\boldsymbol{u}$). In other words, it is actually not necessary to perform the CDF transformation. The implementation of LHSD on a Gaussian random field can then be summarized by the following steps:

1. Simulate $n$ Gaussian vectors $\boldsymbol{e}_i = \left\{e_i^{(1)}, e_i^{(2)}, \ldots, e_i^{(d)}\right\}^T$ with mean $\mathbf{0}$ and spatial correlation $\mathbf{R}$, using Cholesky decomposition: $\boldsymbol{e}_i = \mathbf{L}\boldsymbol{s}_i$, where $\mathbf{L}$ is the Cholesky factor of $\mathbf{R}$, $\boldsymbol{s}_i$ is a $d \times 1$ vector of independent standard Gaussian random variables.
2. Stack the $n$ vectors by rows to form a matrix, with the $i^{th}$ row and $j^{th}$ column denoted by $\boldsymbol{e}_i$ and $\boldsymbol{e}^{(j)}$. For each column, compute the rank statistic according to Eq. (3).
3. Calculate $v_i^{(j)} = (r_i^{(j)} - 0.5)/n$, where $\boldsymbol{v}_i = \left\{v_i^{(1)}, v_i^{(2)}, \ldots, v_i^{(d)}\right\}^T$ is the $i^{th}$ LHSD standard uniform sample.
4. Back-transform the standard uniform sample into a multivariate Gaussian sample using the inverse cumulative distribution function with consideration of the deterministic trend:

$$\begin{aligned} \boldsymbol{z}_i &= \left\{\mathrm{F}_1^{-1}(v_i^{(1)}), \mathrm{F}_2^{-1}(v_i^{(2)}), \ldots, \mathrm{F}_d^{-1}(v_i^{(d)})\right\}^T \\ \text{where} \qquad \mathrm{F}_j^{-1}(v_i^{(j)}) &= \boldsymbol{\mu}_j + \sqrt{\sigma^2}\ \Phi^{-1}(v_i^{(j)}) \end{aligned} \tag{4}$$

   with $\sqrt{\sigma^2}$ being the standard deviation of Gaussian random field, and $\Phi^{-1} : [0,1] \to \mathbb{R}$ being the inverse cumulative standard Gaussian distribution function.

In some cases, the original random field is non-Gaussian, and a transformation on the original field is necessary. A non-Gaussian original random field is denoted as $\boldsymbol{z}^*$ herein. For example, log-transform is common in modeling positive-valued soil properties such as Young's modulus and undrained shear strength ($\boldsymbol{z}_i = \log \boldsymbol{z}_i^*$), while Box-Cox transform may be applied to ensure stationarity assumptions are satisfied, i.e. normality and constant variance across the spatial domain (Liu et al. 2017). It should be noted that both log and Box-Cox transform are monotonic. If a transform T is monotonic, then its back-transform (or inverse) function $\mathrm{T}^{-1}$ will also be monotonic. Meanwhile, the percentile for a distribution is unchanged under a monotonic transformation, as noted by Lark and Lapworth (2012). In other words, the $k^{th}$ percentile in a normal distribution is still the $k^{th}$ percentile in the monotonic back-transformed distribution. Therefore, stratification is preserved when the LHSD sample, $\boldsymbol{z}_i$ from Eq. (4), is back-transformed to the original space (i.e., $\boldsymbol{z}_i^* = \mathrm{T}^{-1}(\boldsymbol{z}_i)$). This back transformation will be an additional step to the implementation of LHSD in cases of non-Gaussian random field.

## Application to cross-correlated random fields

Geotechnical properties are often found to correlate with each other. For example, the shear strength and stiffness of many soils are observed, or assumed, to be positively-correlated. Fenton and Griffiths (2003) outlined the simulation of cross-correlated random fields, making use of the lower triangular matrix, $\mathbf{L}_\rho$, from Cholesky decomposition of the correlation matrix between the two properties. If $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ are two cross-correlated standard Gaussian random fields with cross correlation coefficient $\rho$, at each spatial location $\boldsymbol{x}_i$:

$$\begin{bmatrix} \boldsymbol{z}_1(\boldsymbol{x}_i) \\ \boldsymbol{z}_2(\boldsymbol{x}_i) \end{bmatrix} = \mathbf{L}_\rho \begin{bmatrix} \boldsymbol{s}_1(\boldsymbol{x}_i) \\ \boldsymbol{s}_2(\boldsymbol{x}_i) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} \boldsymbol{s}_1(\boldsymbol{x}_i) \\ \boldsymbol{s}_2(\boldsymbol{x}_i) \end{bmatrix} \tag{5}$$

where $\boldsymbol{s}_1$, and $\boldsymbol{s}_2$ are the two independent standard Gaussian random fields. To adopt this approach in the LHSD framework, the spatial correlation matrix ($\mathbf{R}$) in Step (1) needs to incorporate cross-correlation between the considered properties, and the updated matrix is denoted as $\mathbf{R}_\rho$ herein. $\mathbf{R}_\rho$ can be obtained by considering a block correlation matrix $\mathbf{R}_{12}^{(ij)}$, whose $ij^{th}$ block represents the correlation of the two properties (1 and 2) between locations $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$:

$$\mathbf{R}_\rho^{(ij)} = \mathbf{L}_\rho \mathbf{R}_{12}^{(ij)} \mathbf{L}_\rho^T$$

$$\text{where} \qquad \mathbf{R}_{12}^{(ij)} = \begin{bmatrix} \mathbf{R}_1(\boldsymbol{x}_i, \boldsymbol{x}_j) & 0 \\ 0 & \mathbf{R}_2(\boldsymbol{x}_i, \boldsymbol{x}_j) \end{bmatrix} \tag{6}$$

The rows and columns of $\mathbf{R}_\rho$ need to be re-arranged such that Rows 1 to $d$ correspond to Property 1 and Rows $(d+1)$ to $2d$ correspond to Property 2.

## Application to conditioned random field

In most engineering applications, a number of sampled points (e.g., boreholes, CPT soundings) are available, providing limited amount of information upon which engineering assumptions of geotechnical properties are based. The characterization of spatial variability in geological profiles and geotechnical properties have been discussed by Lloret-Cabot et al. (2012), Dasaka and Zhang (2012), Liu et al. (2017), Li et al. (2016), etc., with consideration on the observed data at sampled locations. These recent efforts illustrate how site information can be utilized in the construction of conditioned random fields.

In the current study, the importance of sample locations will be demonstrated through probabilistic analyses involving conditioned random fields generated using the LHSD approach. Consider the case with $k$ observed sample points, denoted as $\boldsymbol{z}_0 = z_0^{(1)}, z_0^{(2)}, \ldots, z_0^{(k)}$. The generated random field, $\boldsymbol{z}$, can reflect the sampling data by becoming conditional on $\boldsymbol{z}_0$, (i.e. $\boldsymbol{z}|\boldsymbol{z}_0$). If $\boldsymbol{z}$ and $\boldsymbol{z}_0$ are both multivariate Gaussian, then $\boldsymbol{z}|\boldsymbol{z}_0$ is also multivariate Gaussian, and the conditional mean and covariance can be derived as:

$$\boldsymbol{\mu}_{cond} = \mathrm{E}\left[\boldsymbol{z}|\boldsymbol{z}_{\mathbf{0}}\right] = \boldsymbol{\mu} + \mathbf{V}_c^T \mathbf{V}_0^{-1} (\boldsymbol{z}_0 - \boldsymbol{\mu}_0) \tag{7a}$$

$$\mathbf{V}_{cond} = \mathrm{cov}[\boldsymbol{z}|\boldsymbol{z}_0] = \mathbf{V} - \mathbf{V}_c^T \mathbf{V}_0^{-1} \mathbf{V}_c \tag{7b}$$

In Eq. (7b), $\mathbf{V}$ represents the covariance of the application domain, and $\mathbf{V}_0$ is a $k \times k$ spatial covariance matrix between the $k$ sampled locations, which is a subset of $\mathbf{V}$. $\mathbf{V}_c$ is the $k \times d$ covariance matrix between the sampled and unsampled (simulated) locations,

$\boldsymbol{\mu}_0$ represents the expected values at the sampled locations. Under the ordinary kriging formulation (Cressie 1993), $\boldsymbol{\mu}_{cond}$ represents the kriging predictor, and the diagonal terms of $\mathbf{V}_{cond}$ are equivalent to the prediction variance, $\boldsymbol{\sigma}_z^2$. The conditioned random field is no longer second-order stationary, since the diagonal of $\mathbf{V}_{cond}$ is not constant. The following transforms $\mathbf{V}_{cond}$ into a conditional correlation matrix $\mathbf{R}_{cond}$:

$$\mathbf{R}_{cond} = \mathbf{D}^{-\frac{1}{2}}\mathbf{V}_{cond}\mathbf{D}^{-\frac{1}{2}} \tag{8}$$

where $\mathbf{D}$ is a $d \times d$ diagonal matrix formed by the $d$ terms in $\boldsymbol{\sigma}_z^2$. To apply the LHSD approach to a conditioned random field, $\mathbf{R}$ will be replaced by $\mathbf{R}_{cond}$ in Step (1) described earlier, and $\boldsymbol{\mu}_j$ and $\sqrt{\sigma^2}$ will be replaced by $\boldsymbol{\mu}_{cond,j}$ and $\sqrt{\boldsymbol{\sigma}_{z,j}^2}$, respectively, in Step (4).

**LHSD coupled with polynomial chaos expansion**

If the input to a geotechnical model consists of independent standard Gaussian random variables, the probability density function of the response can be approximated by the polynomial chaos expansion (PCE), which was described in detail by Ghanem and Spanos (1991) and recently applied by Al-Bittar and Soubra (2014) and Jiang et al. (2015) in reliability analyses of slopes and footings. In the current study, PCE is coupled with the LHSD to further enhance the robustness of the estimator. Following earlier description, the LHSD sample, with spatial correlation $\mathbf{R}$, can be transformed into a set of independent standard Gaussians through the principal component analysis, which is a standard multivariate statistical technique as outlined below. A spectral decomposition is first performed on the spatial correlation matrix $\mathbf{R}$:

$$\mathbf{R} = \mathbf{H}\boldsymbol{\Lambda}\mathbf{H}^T \tag{9}$$

where $\mathbf{H} = [\boldsymbol{h}_1\ \boldsymbol{h}_2\ \ldots\ \boldsymbol{h}_d]$ is a matrix containing $d$ orthonormal eigenvectors; and $\boldsymbol{\Lambda}$ is a diagonal matrix with $d$ positive descending eigenvalues $(\lambda_1, \lambda_2, \ldots, \lambda_d)$. The number of principal components could be less than $d$. For example, if one needs to preserve 95% of the total variance, the number of components, $M$, and the corresponding principal components,

$\boldsymbol{\xi}$, can be obtained by:

$$\min_{M} \sum_{i=1}^{M} \lambda_i > 0.95d \tag{10a}$$

$$\xi_i = \frac{\boldsymbol{h}_i^T \Phi^{-1}(\boldsymbol{v})}{\sqrt{\lambda_i}} \qquad i = 1, 2, \ldots, M \tag{10b}$$

with $\boldsymbol{v}$ being the standard uniform sample discussed earlier. $\xi_i$ are independent standard Gaussians, which are used directly to construct the PCE. The system response, $g$, can be expressed as PCE of order $p$:

$$g(\boldsymbol{\xi}) = \sum_{\beta=0}^{P-1} a_\beta \Psi_\beta$$
$$\text{where} \qquad P = \frac{(M+p)!}{M!p!} \tag{11}$$

where $\Psi_\beta$ are polynomials constructed by $\boldsymbol{\xi}$, with details shown in the Appendix. The coefficients $a_\beta$ can be computed by the regression approach (Blatman and Sudret 2010; Al-Bittar and Soubra 2014). This involves geotechnical analyses of $n$ realizations of $\boldsymbol{\xi}$ which, in the current study, are performed using the finite difference software, *FLAC*. Results from the $n$ *FLAC* analyses are compiled into a $\boldsymbol{\Gamma}$ vector for the regression analyses to obtain $a_\beta$:

$$\hat{\boldsymbol{a}} = \left(\boldsymbol{\eta}^T \boldsymbol{\eta}\right)^{-1} \boldsymbol{\eta}^T \boldsymbol{\Gamma}$$
$$\text{where} \qquad \boldsymbol{\eta}_{ij} = \Psi_{j-1}(\boldsymbol{\xi}^{(i)}) \qquad i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, P$$
$$\boldsymbol{\Gamma} = \left\{ g(\boldsymbol{\xi}^{(1)}), g(\boldsymbol{\xi}^{(2)}), \ldots, g(\boldsymbol{\xi}^{(n)}) \right\}^T \tag{12}$$

With $a_\beta$ coefficients determined by Eq. (12), the mean and variance of $g(\boldsymbol{\xi})$ are given by:

$$\mathrm{E}\left[g(\boldsymbol{\xi})\right] = a_0 \tag{13a}$$

$$\mathrm{Var}\left[g(\boldsymbol{\xi})\right] = \sum_{\beta=1}^{P-1} (a_\beta)^2 \ \mathrm{E}\left[(\Psi_\beta)^2\right] \tag{13b}$$

Once the coefficients are determined using the $n$ *FLAC* analyses, the probability density function of $g(\boldsymbol{\xi})$ can be constructed by computing Eq. (11) with many sets ($n_{PC}$ sets) of $\boldsymbol{\xi}$. Al-Bittar and Soubra (2014) described this as the 'metamodel' evaluation, and the computation time for this step is short because it does not involve any *FLAC* analyses. To evaluate the accuracy of the PCE, the coefficient $Q^2$ is used (Blatman and Sudret 2010), which is based on leave-one-out cross validation, and reflects the prediction capability of PCE better than the traditional $R^2$ in linear regression.

The proposed approach involves $n$ *FLAC* analyses used to construct the PCE. While the mesh density, i.e. number of elements in the *FLAC* model, controls the size of $\mathbf{R}$ matrix, and affects the efficiency of each model simulation, it does not necessarily affect the 'efficiency of the LHSD-PCE approach' per se. The efficiency of such would, instead, depend on the number of model simulations ($n$) required to achieve a stable PCE. This will be determined by correlation parameters such as the autocorrelation distances ($\theta$), which control the number of principal components ($M$). In fact, as will be shown later, incorporating LHSD would lead to more robust constructions of PCE, which means fewer model simulations are required.

In the following analyses, the spatial correlation is assumed to follow a squared exponential function characterized by $\theta$. In theory, $M$ would vary with the choice of correlation function, if the same value of $\theta$ is adopted. However, from a practical standpoint, the more fundamental issue is the actual autocorrelation of the concerned properties at different separation distances. For example, for a particular project site, the spatial correlation may be represented (fitted) by various functions (single/squared exponential or spherical function), but each of them will correspond to a different value of $\theta$, and the subsequent values of $M$ should still be similar. In other words, the choice of correlation function itself is not the determining factor of $M$ or the efficiency of the approach.

## APPLICATION TO SLOPE STABILITY ANALYSIS

Two example applications are presented in the current study. The first involves analyses of a slope with $c - \phi$ soils, where the slope geometry, soil properties and spatial correlation features are identical to those studied by Cho (2010) and Jiang et al. (2015). Comparisons between the results serve as a validation for the formulation of the current approach, meanwhile illustrating the capabilities and features of LHSD and PCE. In this section, the failure probability in slope analyses is defined as $P_f$ for simplicity, despite the slightly different notations between Eqs. (2)(a) and (b).

Table 1 shows the input parameters of the slope example. The slope has a height of 10 m, slope angle of 45°, with the model boundary at 15 m below the top of slope, and water table is not considered in the analyses. The mean values of shear strength parameters, $c$ and $\phi$, are 10 kPa and 30°, respectively. A deterministic analysis is first performed with uniform soil properties, and the corresponding FS is found to be 1.201 using the strength reduction method implemented in *FLAC*. This is comparable to the value of 1.204 reported by Cho (2010) and 1.206 by Jiang et al. (2015).

In the probabilistic analyses, $c$ and $\phi$ are assumed to be lognormally distributed with coefficients of variation (COV) of 0.3 and 0.2, respectively. For both parameters, the spatial correlation structure is represented by a squared exponential function:

$$\mathbf{R}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \mathbf{R}((x_i, y_i), (x_j, y_j)) = \exp\left[-\left(\frac{|x_i - x_j|}{\theta_{\ln,x}}\right)^2 - \left(\frac{|y_i - y_j|}{\theta_{\ln,y}}\right)^2\right] \tag{14}$$

where $\theta_{\ln,x}$ and $\theta_{\ln,y}$ are the autocorrelation distances in $x$ and $y$ directions, which are taken as 20 m and 2 m, respectively. A typical realization of the two random fields is shown in Fig. 2. Before applying the proposed LHSD approach, 'benchmark' failure probabilities ($P_{f,MC}$) are developed through Monte Carlo simulation with 10,000 *FLAC* analyses, with the cross-correlation coefficients between $c$ and $\phi$ ($\rho_{c-\phi}$) ranging from -0.7 to 0. The LHSD approach, coupled with PCE, is then applied to the same settings to obtain $P_f$, which are then compared with $P_{f,MC}$ and also the findings from Jiang et al. (2015). During construction of the PCE, the principal components are extracted according to Eq. (9), and 96% of the total variance is preserved with 26 to 30 principal components (developing on $\rho_{c-\phi}$) in the current study, resulting in second-order PCE with 378 to 496 terms. The $\hat{\boldsymbol{a}}$ coefficients are obtained based on 1,000 or 1,500 realizations and *FLAC* analyses ($n$=1,000 or 1,500 depending on $\rho_{c-\phi}$) (Eq. (12)), and the probability density function of FS is then reconstructed using the PCE, through 50,000 sets of $\boldsymbol{\xi}$ ($n_{PC}$=50,000). Failure probability is calculated as the proportion of cases with FS$<1$, out of the 50,000 cases.

Fig. 3 shows that the proposed approach is able to reproduce the failure probabilities estimated by the 'raw' Monte Carlo simulation, despite the much smaller number of *FLAC* analyses required in LHSD coupled with PCE. These $P_f$ are, however, higher than those estimated by Jiang et al. (2015) under the same material parameters and correlation features. This may be attributed to the fact that in their evaluation of FS, Jiang et al. (2015) adopted the limit equilibrium method (LEM) with circular slip surfaces; while the current study utilizes the strength reduction method by finite difference analyses (FDM), without any assumptions on slip surfaces or interslice forces. These effects are reflected both in the slightly lower estimates of deterministic FS and higher estimates of $P_f$ by the current approach. In view of the differences between LEM and FDM, the estimates in $P_f$ are comparable. Also, it is deemed that non-circular slip surface may better represent the actual failure mechanism for soils with significant spatial variation.

The features of the proposed approach is further illustrated using the case with $\rho_{c-\phi} = -0.5$ as an example, through six series of analyses tabulated in Table 2: (1a) LHSD with 500 realizations of soil profiles; (1b) PCE with 500 realizations; (1c) LHSD coupled with PCE with 500 realizations; (2a) LHSD with 1,000 realizations and (2b) PCE with 1,000 realizations; (2c) LHSD coupled with PCE with 1,000 realizations. Each of the six series of analyses are repeated 30 times, and the PCE coefficients (for 1b, 1c, 2b and 2c) are re-estimated for each of the 30 repetitions. An empirical standard deviation (SD$_e$) of the $P_f$ estimator is then obtained from each series of analyses. This can be compared with the analytical SD of failure probability (tail probability) by the raw Monte Carlo simulation:

$$\mathrm{SD}_a = \sqrt{\frac{P_f(1-P_f)}{n}} \tag{15}$$

and reductions in $\mathrm{SD}_e$ (compared with $\mathrm{SD}_a$) demonstrate the capabilities of the proposed approach in obtaining a robust estimate of $P_f$.

The LHSD approach leads to a reduction of standard deviation compared with raw Monte Carlo Simulation. For both sample sizes, coupling LHSD with PCE (1c and 2c) will give a greater reduction in standard deviation compared with LHSD alone (1a and 2a). With the sample size $n$ of 1,000, the reduction of standard deviation ($\mathrm{SD}_e$ compared with $\mathrm{SD}_a$) using LHSD with PCE reaches 67%. The standard deviation of the estimator is 0.00172, and is similar to the standard deviation of estimator using Monte Carlo simulation with sample size of 10,000, which is 0.00169 as calculated by Eq. (15). Therefore, a 90% reduction of simulation size is possible with LHSD and PCE, while the estimation accuracy is preserved. This is because when calculating $P_{f,MC}$ using raw Monte Carlo simulation, the profiles are simply separated into two categories, namely those with FS$<1$ and those with FS$>1$. In other words, the calculated FS is not fully utilized. On the other hand, the exact FS values ($g(\boldsymbol{\xi})$) of all cases are used to construct the PCE. In addition, LHSD ensures that the samples are well-spread across the sample domain, which are also reflected in PCE through $\boldsymbol{\xi}$. As a result, the PCE utilizes more information from each model simulation, and has extrapolating power for the tail region of the response. This enables a stable reconstruction of the tail probability, leading to more robust $P_f$ estimates.

It should be noted that the estimator is biased when only PCE approach is adopted (1b and 2b), or with a sample size $n$ of 500 (1c). In those cases, the mean $P_f$ are higher than the benchmark failure probability, and the variances of FS obtained are larger than the true value (around 0.0122). This means the tails of the FS distributions are flattened, causing overestimation of the tail probability. The overestimation is also reflected by the poor prediction capability of the PCE, with average $Q^2$ of 0.371 (1b), 0.790 (1c) and 0.885 (2b). In the case of 1c, there are 435 unknown PCE constants ($M = 28$) but only 500 data samples, which is insufficient to accurately construct the probability density of FS. The estimation bias can be eliminated with a larger sample size and by combining LHSD with PCE. In case 2c, the PCE can be properly constructed with an average $Q^2$ above 0.95. From this numerical experiment, it is recommended that $Q^2$ of the PCE should be at least 0.95 for estimation of the tail probability.

## APPLICATION TO STRIP FOOTING ANALYSIS

The second application involves strip footings on soils with spatially-correlated properties. Conditioned random fields are generated using the proposed LHSD approach, and the subsequent analyses enable an investigation into the uncertainties of foundation performance considering locations of sampled points with known information. A series of design charts are then developed to provide practical guidelines on foundation reliability according to spatial correlation features of the soil and locations of available samples.

For a rigid strip footing of width $B$ under an applied loading $q$, the settlements will depend largely on the elastic parameters (i.e. Young's modulus, $E$, and Poisson's ratio, $\nu$) and shear strength parameters ($c-\phi$) of the soil. In addition, it is a common perception that that the corresponding soil parameters at depths ($D$) of $0.25B$ to $1B$ will be most influential to the footing response, considering the stress distribution under the load and the failure mechanism as $q$ approaches the ultimate bearing capacity $q_u$. For example, Osman and Bolton (2005) suggested that soil properties at $0.3D$ are most representative in simulating the nonlinear response of circular footings on clay. This proposition will be re-examined from the perspective of spatial variations in geomaterials.

In the following sections, footings on linear-elastic soils will be analyzed, followed by footings on soils of Tresca ($c_u$) model and Coulomb ($\phi$) model, with details of footing geometries, input parameters and their variations summarized in Table 3. In the probabilistic analyses, unconditioned random fields are first generated to establish the coefficients of variation for settlement response ($\mathrm{COV}_\delta$), which represents the situation where no site-specific soil samples are available. These will be compared with analyses of conditioned random fields, where soil samples are available at various depths ($D$), and $\mathrm{COV}_\delta$ are reduced accordingly. The magnitudes of $\mathrm{COV}_\delta$ reductions, which can be interpreted as significance of the information, will be shown to vary with the footing geometries, degrees of strength mobilization, depths of samples and spatial variability of the parameters.

The $\mathrm{COV}_\delta$ reductions here should not be confused with the SD reductions associated with Table 2. $\mathrm{COV}_\delta$ reductions in this case correspond to the reductions of performance uncertainty due to additional soil samples under the footings, whereas the SD reductions in the slope study are used to compare robustness of various approaches, defined as the capability to obtain similar $P_f$ values when multiple (30) probabilistic analyses are repeated.

## Footing on linear-elastic soil

Probabilistic analyses of strip footings on spatially variable linear-elastic soils are presented in this section. The *FLAC* model is 15 m wide, 6 m deep with a strip footing of 2 m width on the ground surface. The footing is subjected to a vertical pressure of 500 kPa, and the soil-footing interface is perfectly rough. The Poisson's ratio ($\nu$) of the soil is taken as a constant of 0.3 throughout the domain, while the Young's modulus ($E$) is modeled as a lognormal random field with a mean ($\mu_E$) of 60 MPa and coefficient of variation ($\mathrm{COV}_E$) of 0.15. The autocorrelation of $E$ is assumed to follow a squared exponential function (Eq. (14)), with the horizontal autocorrelation distance much larger than the domain scale ($\theta_{\ln,x}$=200m). This assumption is made since Al-Bittar and Soubra (2014) observed that for $B = 2$ m, the footing settlements ($\delta$) becomes insensitive to changes in $\theta_{\ln,x}$ once $\theta_{\ln,x} > 10$ m; meanwhile $\theta_{\ln,x}$ for soil properties are often found to be an order of magnitude higher than $\theta_{\ln,y}$, ranging from 10 m to over 80 m (Phoon and Kulhawy 1999a; DeGroot 1996).

In other words, this study focuses on the influence of $\theta_{\ln,y}$ on the footing performance ($\delta$), and the importance of sample depth $D$ on the overall reduction of uncertainties ($\mathrm{COV}_\delta$). To create a basis for comparison, the LHSD approach is coupled with PCE to first simulate 500 unconditioned random fields, which represent scenarios where no samples are available under the footing. In the analyses, 97% of the total variance is preserved by extracting principal components from the unconditioned spatial autocorrelation matrix of $E$. The principal components are then used to construct PCE of the second order, from which $\text{COV}_\delta$ of unconditioned cases are obtained. The $\text{COV}_\delta$ for conditioned cases are evaluated using a similar procedure, except that Eqs. (7) and (8) are applied to simulate conditioned random fields with various sample depths. Also, $Q^2$ is larger than 0.95 in all subsequent analyses.

Fig. 4 shows the reductions in $\text{COV}_\delta$ comparing the unconditioned random fields with conditioned cases, considering different sample depths ($D/B$) and $B/\theta_{\text{ln},y}$ ratios. Apart from the base case described earlier ($B = 2$ m, $\mu_E = 60$ MPa, $\text{COV}_E = 0.15$), three more sets of probabilistic analyses have been performed, with double model scale (i.e. $B = 4$ m and double domain size), reduced mean stiffness ($\mu_E = 30$ MPa) and increased stiffness variation ($\text{COV}_E = 0.4$), respectively. Although they entail different $B/\theta_{\text{ln},y}$ ratios, all the resulting data points are lined up along the corresponding $D/B$ lines, which demonstrates the validity of normalization employed in Fig. 4.

A larger reduction in $\text{COV}_\delta$ represents better value of the sample as the uncertainties in $\delta$ are reduced to a greater extent through knowledge of $E$ at that point. Therefore, according to Fig. 4, the most significant sampling points are at depths of $1B, 0.5B$ and $0.25B$, depending on the ratio between footing width and vertical autocorrelation distance. Two crossover points are observed at $B/\theta_{\text{ln},y} = 1$ and $B/\theta_{\text{ln},y} = 2.75$. With large values of $\theta_{\text{ln},y}$ ($B/\theta_{\text{ln},y} < 1$), the soil properties are relatively uniform. For example, with sample depth $D = B = \theta_{\text{ln},y}$, the sample is representative of the properties from the ground surface to a depth of about 2B, which covers the region where most of the internal work is dissipated in a linear-elastic strip footing analysis. This explains why the sampling depth is most effective at $D = B$ in such cases. On the other hand, with a relatively small value of $\theta_{\text{ln},y}$ ($B > 2.75\theta_{\text{ln},y}$), the properties are highly variable, and the shallow region (e.g. a shallow, highly compressible layer) can become more influential to the overall footing settlement. Therefore, the optimal sampling depth is at $D = 0.25B$. In the transition where $B < \theta_{\text{ln},y} < 2.75B$, the optimal sampling depth is at $D = 0.5B$. Moreover, Fig. 4 shows that the maximum reduction in

$\mathrm{COV}_\delta$ depends heavily on $\theta_{\ln,y}$, ranging from around 20% when $\theta_{\ln,y} = 0.2B$, up to 80% when $\theta_{\ln,y} = 2B$. As $\theta_{\ln,y}$ increases, the conditioning power of the sample point becomes more significant, which means the regions around the sample are less uncertain, and hence more substantial reductions in $\mathrm{COV}_\delta$ can be achieved.

As mentioned earlier, the curves shown in Fig. 4 are insensitive to the footing size, $\mu_E$ and $\mathrm{COV}_E$. It provides general guidelines on the optimal sampling depth and the corresponding percentage reduction in $\mathrm{COV}_\delta$. Although the value of $\theta_{\ln,y}$ cannot be determined with a single sample at the site, it may be reasonably assumed based on understanding of the local geology, or published information from the literature (Phoon and Kulhawy 1999a; DeGroot 1996). Also, Fig. 4 is established through linear-elastic analyses, and is therefore more relevant to footing designs with high factors of safety. In the following sections, plasticity will be introduced in the analyses as the footings are loaded to a factor of safety of 2.0.

## Footing on Tresca ($c_u$) soil

This section investigates the uncertainties in footing performance on spatially variable soils idealized as Tresca material. The footing size, model boundaries and spatial characteristics of soil Young's modulus are identical to the base case in the previous section, but the Poisson's ratio is set as 0.499 for total stress analyses. The undrained shear strength ($c_u$) is perfectly correlated with the Young's Modulus, with a constant $E/c_u$ ratio in the soil domain (i.e. $\mathrm{COV}_{cu} = \mathrm{COV}_E = 0.15$). Two sets of probabilistic analyses are performed with mean $c_u$ values ($\mu_{cu}$) of 120 kPa and 200 kPa, resulting in $E/c_u$ ratios of 500 and 300, respectively. In both cases, the applied loading is assigned such that the 'deterministic' factor of safety, based on ultimate bearing capacity of $q_u = (2+\pi)\mu_{cu}$, equals 2.0.

Similar to the linear-elastic case, the significance of sampling depth $D$ can be assessed by comparing the $\mathrm{COV}_\delta$ obtained from unconditioned and conditioned random field simulations. Fig. 5(a) presents the results for different $B/\theta_{\ln,y}$ ratios, which also shows that under the same deterministic FS, the reductions in $\mathrm{COV}_\delta$ is insensitive to the individual $\mu_{cu}$ value (or $E/c_u$ ratio) adopted. $\mathrm{COV}_\delta$ is, however, sensitive to the variations of $c_u$. Fig. 5(b) shows the analyses with $\text{COV}_{cu} = \text{COV}_E = 0.4$, and the resulting curves of $\text{COV}_\delta$ reductions are substantially different. The significance of sample depth from $0.25B$ to $B$ is greatly enhanced with a large variation in $c_u$, while sample depth at $2B$ becomes even less important.

The discrepancies between Fig. 5(a) and (b) may be explained by first considering a footing on uniform Tresca material, where the slip surface lies between the ground surface to depths of approximately $0.7B$ at bearing failure. Similarly for spatially variable soils, these depths are also observed to be more dominant as plasticity is developed. Fig. 6 shows the plastic zones developed under the footing in two example *FLAC* analyses with the same $\mu_{cu}$. With a weaker shallow layer, plastic zones are concentrated near the ground surface, whereas the profile with stronger shallow layer is associated with only a small number of (or no) plastic zones. With a larger variation in $c_u$, there is a higher probability of plasticity developing in the shallow layer even with a deterministic FS of 2.0, since this FS is evaluated only based on the mean shear strength. Therefore, the information of $c_u$ at shallow depths ($0.25B$ to $B$) becomes more important, causing the corresponding curves to shift up in Fig. 5(b). On the contrary, the significance of deeper samples ($2B$) appears to diminish further.

### Footing on Coulomb ($\phi$) soil

In this section, the soil is modeled as a Coulomb material, with spatially varying friction angle ($\phi$) and Young's modulus ($E$), while $\nu = 0.3$ and is a constant. Unlike the previous case for $c_u$ material, perfect correlation is rarely adopted between $\phi$ and $E$, meanwhile $\text{COV}_\phi$ is typically smaller than $\text{COV}_E$ at the same site. Therefore, in the current study, cross-correlated, conditioned random fields (Eqs. (6) and (8)) are generated for $E$ and $\phi$, with cross-correlation coefficient $\rho_{\phi-E} = 0.5$, $\text{COV}_\phi = 0.05$ and $\text{COV}_E = 0.15$. $\phi$ is assumed to be lognormally distributed with mean value $\mu_\phi = 35°$. The footing size and model boundaries remain the same as previous cases. Fig. 7 shows an example of the residuals for simulated profiles of $E$ and $\phi$, where both profiles pass through the sample point at the designated depth ($D/B = 0.5$), with positive cross-correlation between the residuals.

Adopting a similar strategy as before, Fig. 8 shows the reductions in $\text{COV}_\delta$ for conditioned random fields, at different sampling depths $D$ and $B/\theta_{\ln,y}$ ratios. In all these analyses, the footing is loaded to a deterministic FS of 2.0, with $q_u = 0.5\gamma B N_\gamma$ ($N_\gamma$ is the bearing capacity factor, taken as a function of $\mu_\phi$). One set of additional analyses is performed with $\mu_\phi = 30°$, and the results show that COV$_\delta$ is insensitive to $\mu_\phi$ under the same FS. While the general pattern of Fig. 5(a) and 8 are similar, the reduction in COV$_\delta$ appear to be slightly lower for $\phi$ soil. Fig. 8 also includes analyses with $\rho_{\phi-E} = 1.0$, and shows that for practical purposes, the influence of different $\rho_{\phi-E}$ values is minimal.

In many practical situations, the mean shear strength parameters ($\mu_{cu}$ or $\mu_\phi$) of soils are not constant with depth. Such effects on footing response are also evaluated in this study, by adopting depth-dependent shear strength profiles in the probabilistic analyses. In general, the associated COV$_\delta$ reductions are largely similar to those of constant $\mu_{cu}$ or $\mu_\phi$ soils. More details are provided in Fig. S1 and Table S1 in the Supplemental Data File.

## PRACTICAL GUIDELINES ON ESTIMATION OF SITE-SPECIFIC COV$_\delta$

Making use of findings from Figs. 4, 5 and 8, a set of design charts and guidelines are established for quick estimates of COV$_\delta$ for strip footings, based on project-specific foundation geometry, depth of soil samples or *in situ* tests and spatial correlation features of the associated properties. The procedures can be described as follows:

1. Deterministic estimates of footing settlement ($\delta_d$) can be obtained by $\mu_{cu}$ or $\mu_\phi$ and $\mu_E$, using common design procedures.
2. The mean settlement from probabilistic analyses ($\mu_\delta$) can be inferred from the ratio of $\mu_\delta/\delta_d$. Based on analyses from this study, the ratio is approximately 1.02 for COV$_E = 0.15$ and 1.12 for COV$_E = 0.4$, with details shown in Fig. S2 of the Supplemental Data File.
3. The 'unconditioned' COV$_\delta$ depends on the COV$_E$, COV$_{cu}$ or COV$_\phi$ for soils at the site, and can be estimated from Fig. 9, which is developed by compiling the results of unconditioned random field analyses described earlier.
4. With soil sample under the footing, the 'conditioned' COV$_\delta$ can be estimated by interpolation from Figs. 5, 8 and S1, since most footings are designed with FS$>$ 2.0. The

standard deviation, or confidence level of $\delta$ can then be assessed with $\text{COV}_\delta$ and $\mu_\delta$.

An interesting feature of the significance of sampling can be revealed by comparing the current study with results by Al-Bittar and Soubra (2014), who conducted probabilistic analyses for unconditioned random fields underneath the footing. One of their cases involved $\theta_{\ln,x} = \infty$, which is comparable to the current study. With $\text{COV}_E = 0.15$ and $B/\theta_{\ln,y}$ ratio of 2, the $\text{COV}_\delta$ obtained by Al-Bittar and Soubra (2014) and the current study are 8.16% and 8.22%, respectively, showing good comparisons between the two approaches. Further, the analyses by Al-Bittar and Soubra (2014) show that $\text{COV}_\delta$ increases monotonically with $\theta_{\ln,y}$, and the same trend is observed in the unconditioned random field analyses herein, as presented in Fig. 9. However, Fig. 9 also shows that under a conditioned random field (with $\text{COV}_{cu} = \text{COV}_E = 0.15$), not only is $\text{COV}_\delta$ reduced, their trend is also substantial altered. With small $B/\theta_{\ln,y}$ ratio (i.e. large $\theta_{\ln,y}$), the soil is more uniform and $\text{COV}_\delta$ is greatly diminished through additional knowledge from the soil sample, while large $B/\theta_{\ln,y}$ ratios indicate more variable soils and the significance of the sample is less pronounced.

Some limitations of the presented design charts and guidelines should be noted. They are developed based on random fields with either constant mean values of modulus and shear strength parameters, or monotonic variations of these parameters (Table S1). As with any design charts, they should be applied with proper engineering judgement, especially when:

- the geological profiles display strong layering effects (e.g. presence of particularly weak seams), where geotechnical properties vary abruptly with depth; or
- the trends of mean stiffness or shear strength parameters differ significantly from previous assumptions. For example, surface dessication often increases the overconsolidation ratio, and hence the shear strength, of clayey soils near the ground surface. In these cases, $c_u$ may reduce with depth near the surface but increase with depth beyond a certain point.

Under these conditions, the $\text{COV}_\delta$ estimates may be different from those presented in earlier sections. The design charts and guidelines may be treated as first-pass assessments of the performance uncertainty, while detailed probabilistic analyses should be conducted by the proposed LHSD-PCE approach for more accurate estimates.

## CONCLUSION

This paper presents the LHSD-PCE approach, which is capable of achieving similar accuracy compared with raw Monte Carlo simulations, but with much smaller numbers of model simulations. The approach is formulated for random field analyses with cross-correlated parameters and conditioning that arises from availability of soil samples. Probabilistic analyses of slopes and strip footings are performed, the latter of which reveals the significance of various sampling depths beneath the footing. In most cases, sampling depths of $0.25B$ to $1B$ are the most influential, depending on the spatial correlation features and adopted FS in deterministic analysis. The investigation also leads to development of a series of design charts and practical guidelines, which allow researchers and practitioners to quickly estimate the uncertainty of foundation performance without performing the probabilistic analyses.


## ACKNOWLEDGEMENTS

The work presented in this paper is financially supported by the Research Grants Council of the Hong Kong Special Administrative Region (Project No. 25201214).


## APPENDIX I. $\Psi$ FUNCTIONS IN POLYNOMIAL CHAOS EXPANSION

$\Psi_\beta$ is a set of zero mean, independent (orthogonal) random polynomials constructed using $\xi_i$. For a second order PCE ($p = 2$), $\Psi_\beta$ are given by:

$$\begin{aligned}
\text{For } \beta = 0: &\qquad \Psi_0 = 1 \\
\text{For } \beta = 1, 2, \ldots, M: &\qquad \Psi_\beta = \xi_i \quad (i = 1, 2, \ldots, M) \\
\text{For } \beta = M+1, \ldots, P-1: &\qquad \Psi_\beta = \xi_{i_1}\xi_{i_2} - \delta_{i_1 i_2} \quad (i_1 = 1, \ldots, M;\ i_2 = i_1, \ldots, M)
\end{aligned} \tag{16}$$

where $\delta_{i_1 i_2}$ represents the Kronecker delta.

As an example, Table 4 is extracted from Ghanem and Spanos (1991), which shows the formulations when $M = 3$. In this study, only PCE of second order have been adopted. The construction of PCE for higher orders have been discussed by Al-Bittar and Soubra (2014).

## SUPPLEMENTAL DATA

Effects of depth-dependent shear strength profiles on $\mathrm{COV}_\delta$, and the ratio between probabilistic and deterministic settlement estimates, including Table S1 and Figs. S1 and S2, are available online in the ASCE Library (www.ascelibrary.org).

## REFERENCES


Al-Bittar, T. and Soubra, A.-H. (2014). "Probabilistic analysis of strip footings resting on spatially varying soils and subjected to vertical or inclined loads." *J. Geotech. Geoenviron. Eng.*, 10.1061/(ASCE)GT.1943-5606.0001046, 04013043.

Baecher, G. and Christian, J. (2003). *Reliability and Statistics in Geotechnical Engineering.* Wiley, New York.

Blatman, G. and Sudret, B. (2010). "An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis." *Probabilistic Engineering Mechanics*, 25(2), 183 – 197.

Cho, S. E. (2010). "Probabilistic assessment of slope stability that considers the spatial variability of soil properties." *J. Geotech. Geoenviron. Eng.*, 10.1061/(ASCE)GT.1943-5606.0000309, 975–984.

Christian, J. T., Ladd, C. C., and Baecher, G. B. (1994). "Reliability applied to slope stability analysis." *J. Geotech. Engrg.*, 10.1061/(ASCE)0733-9410(1994)120:12(2180), 2180–2207.

Cressie, N. A. C. (1993). *Statistics for Spatial Data (Revised Edition).* John Wiley & Sons.

Dasaka, S. M. and Zhang, L. M. (2012). "Spatial variability of in situ weathered soil." *Géotechnique*, 62(5), 375–384.

DeGroot, D. J. (1996). "Analyzing spatial variability of in situ soil properties." *Uncertainty in the Geologic Environment: From Theory to Practice*, Vol. 1, 210–238.

Duncan, J. M. (2000). “Factors of safety and reliability in geotechnical engineering.” *J. Geotech. Geoenviron. Eng.*, 10.1061/(ASCE)1090-0241(2000)126:4(307), 307–316.

Fenton, G. A. and Griffiths, D. V. (2003). “Bearing-capacity prediction of spatially random c - $\phi$ soils.” *Can. Geotech. J.*, 40(1), 54–65.

Ghanem, R. G. and Spanos, P. D. (1991). *Stochastic Finite Element: A Spectral Approach.* Springer, New York.

Griffiths, D. V. and Fenton, G. A. (2009). “Probabilistic settlement analysis by stochastic and random finite-element methods.” *J. Geotech. Geoenviron. Eng.*, 10.1061/(ASCE)GT.1943-5606.0000126, 1629–1637.

Griffiths, D. V., Huang, J., and Fenton, G. A. (2009). “Influence of spatial variability on slope reliability using 2-D random fields.” *J. Geotech. Geoenviron. Eng.*, 10.1061/(ASCE)GT.1943-5606.0000099, 1367–1378.

Hasofer, A. M. and Lind, N. C. (1974). “An exact and invariant first order reliability format.” *J. Eng. Mech. Div.*, 100(EM1), 111121.

Jiang, S.-H., Li, D.-Q., Cao, Z.-J., Zhou, C.-B., and Phoon, K.-K. (2015). “Efficient system reliability analysis of slope stability in spatially variable soils using Monte Carlo simulation.” *J. Geotech. Geoenviron. Eng.*, 10.1061/(ASCE)GT.1943-5606.0001227, 04014096.

Kasama, K. and Whittle, A. J. (2011). “Bearing capacity of spatially random cohesive soil using numerical limit analyses.” *J. Geotech. Geoenviron. Eng.*, 10.1061/(ASCE)GT.1943-5606.0000531, 989–996.

Lark, R. M. and Lapworth, D. J. (2012). “Quality measures for soil surveys by lognormal kriging.” *Geoderma*, 173174, 231 – 240.

Li, J., Tian, Y., and Cassidy, M. J. (2015). “Failure mechanism and bearing capacity of footings buried at various depths in spatially random soil.” *J. Geotech. Geoenviron. Eng.*, 10.1061/(ASCE)GT.1943-5606.0001219, 04014099.

Li, K. S. and Lumb, P. (1987). “Probabilistic design of slopes.” *Can. Geotech. J.*, 24(4), 520–535.

Li, X. Y., Zhang, L. M., and Li, J. H. (2016). "Using conditioned random field to characterize the variability of geologic profiles." *J. Geotech. Geoenviron. Eng.*, 10.1061/(ASCE)GT.1943-5606.0001428, 04015096.

Liu, W. F., Leung, Y. F., and Lo, M. K. (2017). "Integrated framework for characterization of spatial variability of geological profiles." *Can. Geotech. J.*, 54(1), 47–58.

Lloret-Cabot, M., Hicks, M. A., and van den Eijnden, A. P. (2012). "Investigation of the reduction in uncertainty due to soil variability when conditioning a random field using kriging." *Géotechnique Letters*, 2(3), 123–127.

McKay, M. D., Beckman, R. J., and Conover, W. J. (1979). "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code." *Technometrics*, 21(2), 239–245.

Osman, A. S. and Bolton, M. D. (2005). "Simple plasticity-based prediction of the undrained settlement of shallow circular foundations on clay." *Géotechnique*, 55(6), 435–447.

Packham, N. (2015). "Combining Latin hypercube sampling with other variance reduction techniques." *Wilmott*, 2015(76), 60–69.

Packham, N. and Schmidt, W. M. (2010). "Latin hypercube sampling with dependence and applications in finance." *Journal of Computational Finance*, 13(3), 81–111.

Phoon, K. K. and Kulhawy, F. H. (1999a). "Characterization of geotechnical variability." *Can. Geotech. J.*, 36(4), 612–624.

Phoon, K. K. and Kulhawy, F. H. (1999b). "Evaluation of geotechnical property variability." *Can. Geotech. J.*, 36(4), 625–639.

Soulié, M., Montes, P., and Silvestri, V. (1990). "Modeling spatial variability of soil parameters." *Can. Geotech. J.*, 27(5), 617–630.

Vanmarcke, E. H. (1977). "Probabilistic modeling of soil profiles." *J. Geotech. Engrg. Div.*, 103(GT11), 1227–1246.

Vanmarcke, E. H. (1984). *Random Fields: Analysis and Synthesis*. MIT Press, MA.

Wiener, N. (1938). "The homogeneous chaos." *American Journal of Mathematics*, 60(4), 897–936.

Zhang, L. M. and Dasaka, S. M. (2010). "Uncertainties in geologic profiles versus variability in pile founding depth." *J. Geotech. Geoenviron. Eng.*, 10.1061/(ASCE)GT.1943-5606.0000364, 1475–1488.

**TABLE 1. Soil properties and spatial correlations in slope stability analyses**

| Property | | Adopted value |
|---|---|---|
| Shear modulus ($G$) | | 30 MPa |
| Poisson's ratio ($\nu$) | | 0.35 |
| Unit weight ($\gamma$) | | 20 kN/m$^3$ |
| Cohesion | - Mean ($\mu_c$) | 10 kPa |
| | - Coefficient of variation (COV$_c$) | 0.3 |
| Friction angle | - Mean ($\mu_\phi$) | 30° |
| | - Coefficient of variation (COV$_\phi$) | 0.2 |
| Horizontal autocorrelation distance ($\theta_{\mathrm{ln},x}$) | | 20 m |
| Vertical autocorrelation distance ($\theta_{\mathrm{ln},y}$) | | 2 m |
| Cross-correlation coefficient ($\rho_{c-\phi}$) | | -0.7, -0.6, -0.5, -0.4<br>-0.25, -0.1, 0 |

**TABLE 2. Failure probability estimated by six series of LHSD, PCE and LHSD-PCE analyses**

| Comparisons of analyses | LHSD (500 samples) | PCE (500 samples) | LHSD-PCE (500 samples) | LHSD (1000 samples) | PCE (1,000 samples) | LHSD-PCE (1,000 samples) |
|---|---|---|---|---|---|---|
| Mean of LHSD $P_f$ estimator | 0.0285 | 0.0895 | 0.045 | 0.0297 | 0.0331 | 0.0305 |
| Bias exists? | No | Yes | Yes | No | Yes | No |
| $\mathrm{SD}_e$ of LHSD $P_f$ estimator | 0.00582 | 0.01611 | 0.00423 | 0.00513 | 0.00286 | 0.00172 |
| $\mathrm{SD}_a$ of raw Monte Carlo $P_f$ estimator | 0.00758 | 0.00758 | 0.00758 | 0.00536 | 0.00536 | 0.00536 |
| Reduction of SD | 23.2% | -112.5% | 44.2% | 4.3% | 46.6% | 67.9% |
| Average variance in FS estimates | 0.0121 | 0.0221 | 0.0141 | 0.0123 | 0.0130 | 0.0122 |
| Average $Q^2$ for PCE | NA | 0.371 | 0.790 | NA | 0.885 | 0.956 |

**TABLE 3. Strip footing geometries, stiffness, strength and variability parameters adopted in LHSD-PCE analyses of constant $\mu_{cu}$ and $\mu_{\phi}$ cases**

| Parameter | | Linear-elastic soil | Tresca ($c_u$) soil | Coulomb ($\phi$) soil |
|---|---|---|---|---|
| Footing width ($B$) | | 2 m, 4 m | 2 m | 2 m |
| Elasticity parameters | $\mu_E$ | 30 MPa, 60 MPa | 60 MPa | 60 MPa |
| ($E$, $\nu$) | $\mathrm{COV}_E$ | 0.15, 0.4 | 0.15, 0.4 | 0.15, 0.4 |
| | $\nu$ | 0.3 | 0.499 | 0.3 |
| Shear strength | $\mu_{cu}$ | - | 120 kPa, 200 kPa | - |
| parameters | $\mathrm{COV}_{cu}$ | - | 0.15, 0.4 | - |
| ($c_u$, $\phi$) | $\mu_{\phi}$ | - | - | 30°, 35° |
| | $\mathrm{COV}_{\phi}$ | - | - | 0.05 |
| | $\rho_{E-\phi}$ | - | - | 0.5, 1 |

**TABLE 4. Polynomial chaoses ($\Psi_\beta$) and variances with $p = 2$ and $M = 3$ (adapted from Ghanem and Spanos 1991)**

| $\beta$ | Order of the Polynomial Chaos, $p$ | $\Psi_\beta$ | $\mathrm{E}[(\Psi_\beta)^2]$ |
|---|---|---|---|
| 0 | $p = 0$ | 1 | 1 |
| 1 | $p = 1$ | $\xi_1$ | 1 |
| 2 | | $\xi_2$ | 1 |
| 3 | | $\xi_3$ | 1 |
| 4 | $p = 2$ | $\xi_1^2 - 1$ | 2 |
| 5 | | $\xi_1\xi_2$ | 1 |
| 6 | | $\xi_1\xi_3$ | 1 |
| 7 | | $\xi_2^2 - 1$ | 2 |
| 8 | | $\xi_2\xi_3$ | 1 |
| 9 | | $\xi_3^2 - 1$ | 2 |

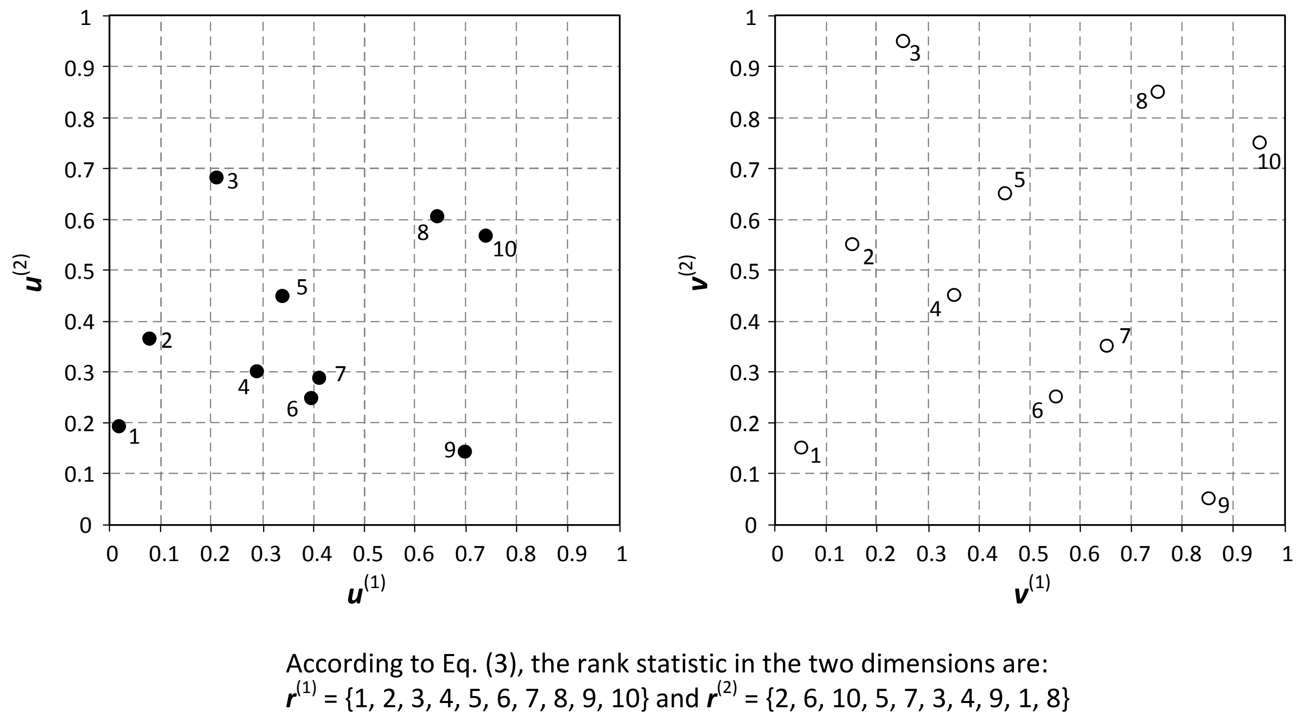


**FIG. 1. Conversion from original sample $u$ to LHSD sample $v$ (based on Packham and Schmidt 2010)**

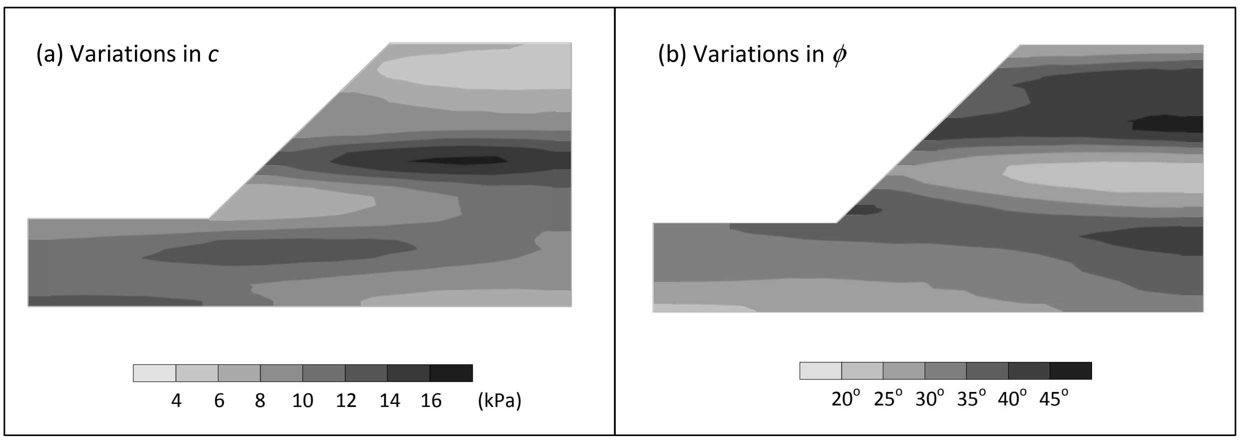


**FIG. 2. Typical realizations of random fields of cohesion and friction angle, with cross-correlation coefficient of** $-0.5$

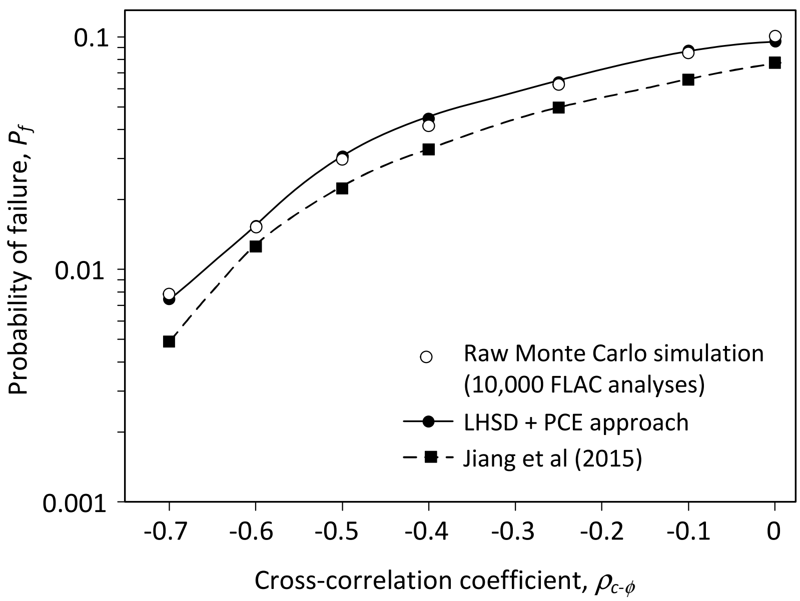


**FIG. 3. Probability of failure in slope stability analyses**

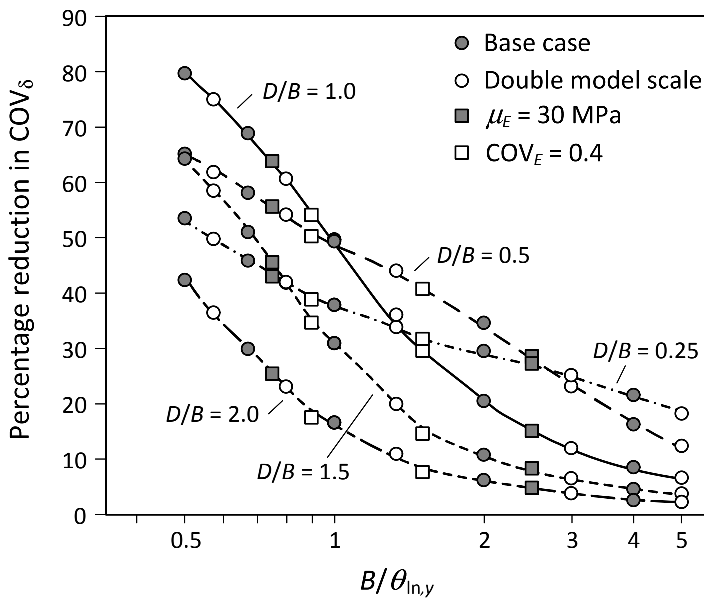


**FIG. 4. Reduction in $COV_\delta$ for strip footings on linear-elastic soil**

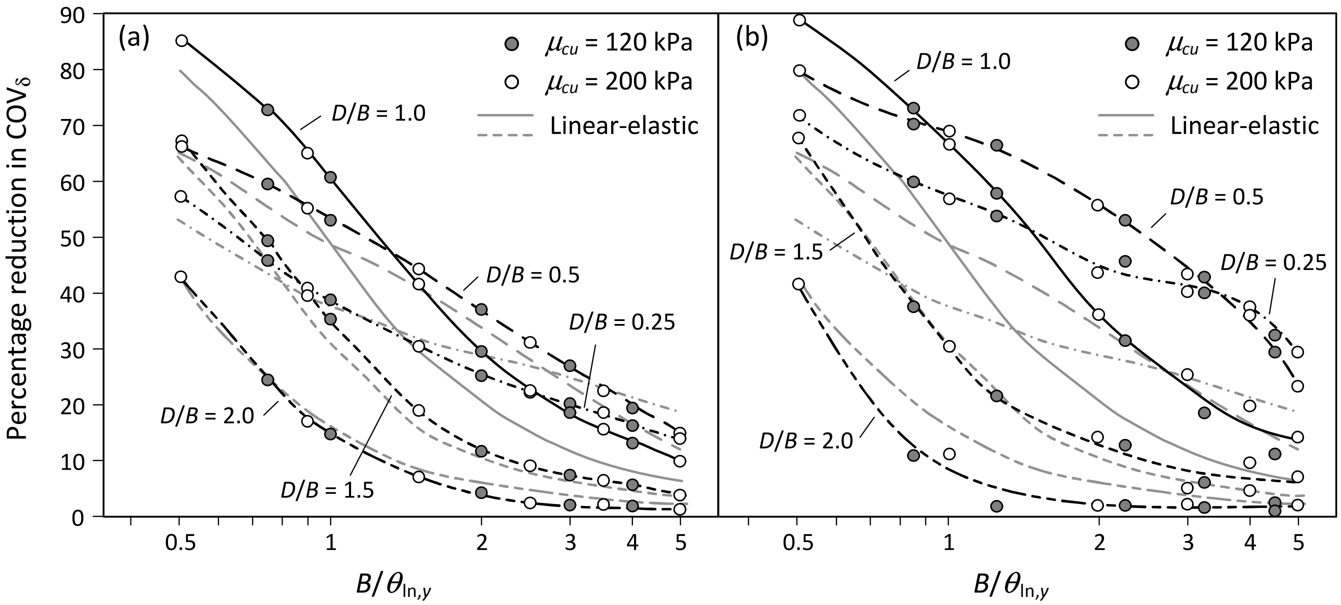


**FIG. 5. Reduction in COV$_\delta$ for strip footings on $c_u$ soil (FS=2.0) with (a) COV$_{cu}$=0.15; and (b) COV$_{cu}$=0.4**

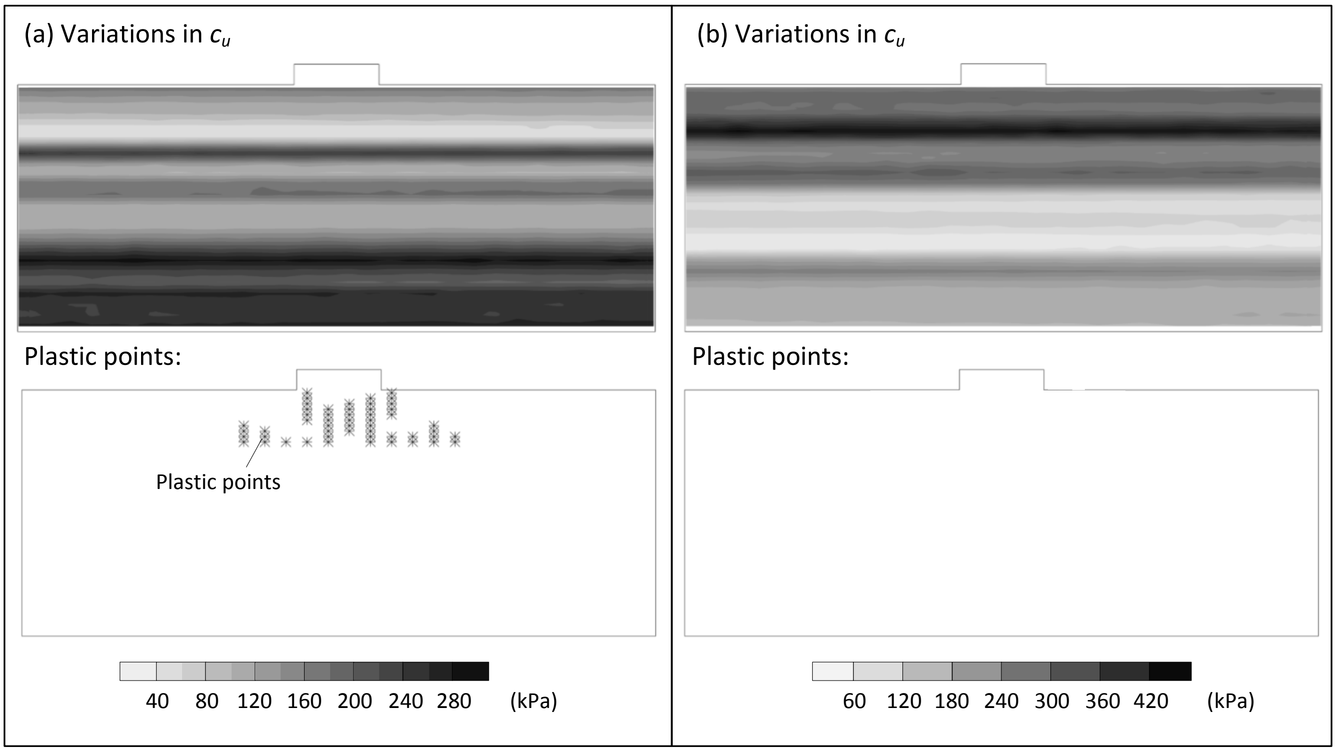


**FIG. 6. Realizations with $\mu_{cu}$ =200 kPa, COV$_{cu}$ =0.4 and $\theta_{\ln,y}$ =0.4 m but different distributions with depth**

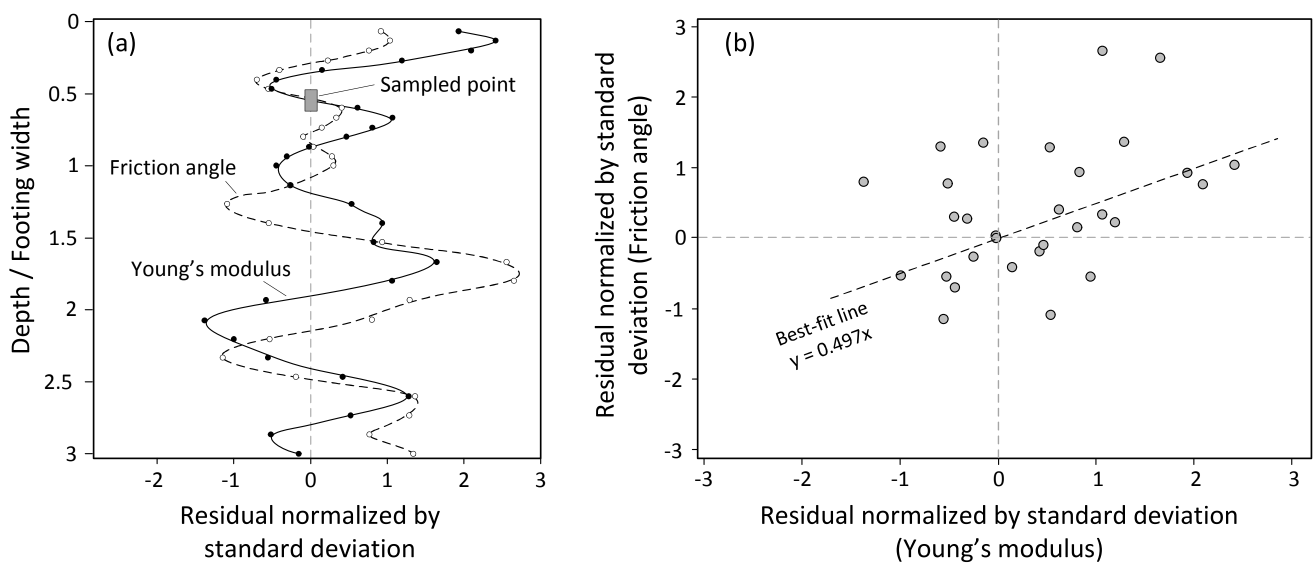


**FIG. 7. (a) Residuals for simulated profiles of Young's modulus and friction angle; (b) Correlations between the normalized residuals**

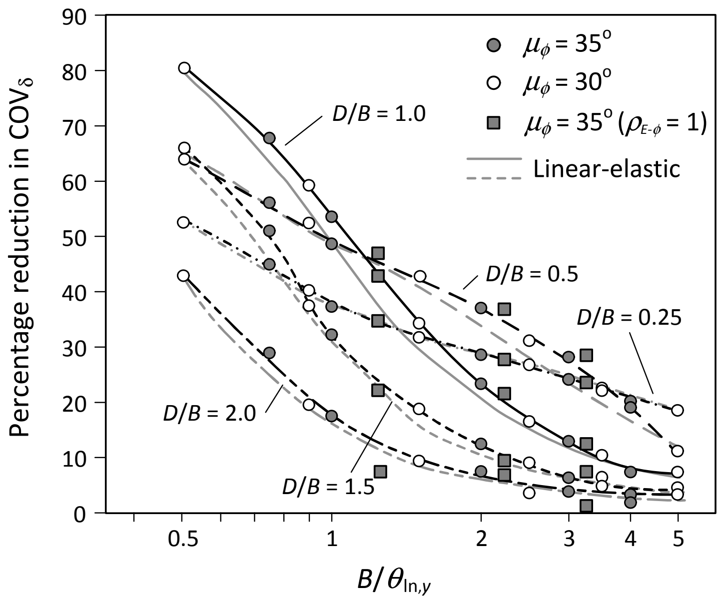


**FIG. 8. Reduction in COV$_\delta$ for strip footings on $\phi$ soil (FS=2.0)**

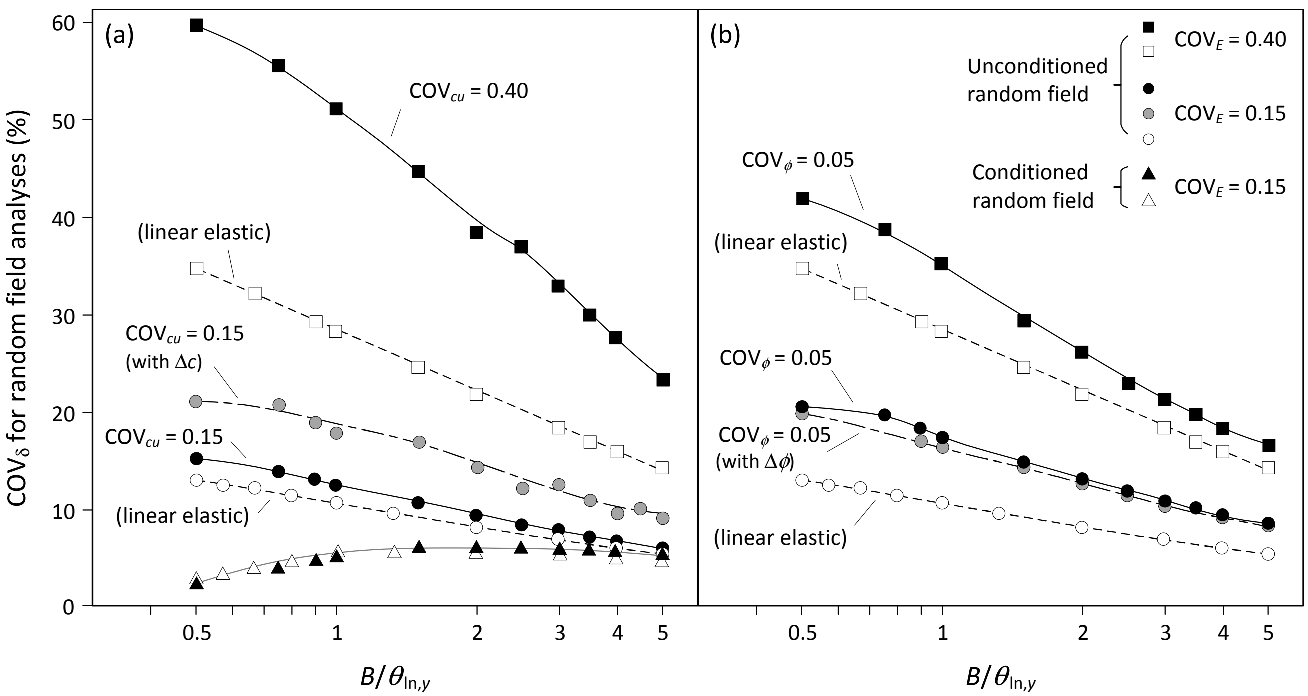


**FIG. 9. COV$_\delta$ for strip footings on (a) $c_u$ soils and (b) $\phi$ soils**