Kriging-Based Reliability Analysis of Strip Footings Resting on Spatially Varying Soils

Tamara Al-Bittar[1]; Abdul-Hamid Soubra, M.ASCE[2]; and Jawad Thajeel[3]

Abstract: The probabilistic analysis of geotechnical structures presenting spatial variability in the soil properties is generally performed using Monte Carlo simulation (MCS) methodology. Despite being robust and accurate, MCS has low efficiency when considering the small failure probabilities encountered in practice. This is because it is very time-expensive in such cases due to the large number of simulations required to calculate a small failure probability with a small value of the coefficient of variation of this failure probability. In order to reduce the number of calls of the mechanical model when performing a probabilistic analysis, this paper uses the active learning reliability method combining kriging and Monte Carlo simulation (AK-MCS). This method is shown to be very efficient because the obtained probability of failure is very accurate, needing only a small number of calls to the computationally expensive mechanical model compared with MCS methodology. This study involves a probabilistic analysis at the ultimate limit state of a strip footing resting on a spatially varying soil using the AK-MCS approach. The soil cohesion and angle of internal friction are considered as random fields. The mechanical model is based on numerical simulations using the finite-difference code FLAC^{3D}. The obtained probabilistic numerical results are presented and discussed.
DOI: 10.1061/(ASCE)GT.1943-5606.0001958.

Author keywords: Probabilistic analysis; Strip footing; Bearing capacity; Spatial variability; Kriging; Monte Carlo simulation; Probability of failure.

Introduction

As is well known, the spatial variability of the soil properties has a significant impact on the probabilistic outputs of geotechnical structures. Several authors have considered the effect of the spatial variability of the soil properties on the statistical moments (mean and standard deviation) of their system response or on the failure probability against a prescribed threshold of this response. The probabilistic analysis of shallow foundations, which is the subject of the present paper, has been studied often (e.g., Fenton and Griffiths 2001; Griffiths et al. 2002; Fenton and Griffiths 2003; Popescu et al. 2005; Cho and Park 2010; Ching et al. 2011; Al-Bittar and Soubra 2013; Al-Bittar and Soubra 2014a, b; Li et al. 2015; Al-Bittar and Soubra 2016; Ching et al. 2016; Li et al. 2016).

When dealing with the computation of the failure probability of geotechnical structures involving spatially varying soils, the classical Monte Carlo simulation (MCS) methodology is generally used. This method is known to be very time-consuming. This is because (1) it usually makes use of finite-element or finite-difference models, which are generally time-expensive, and, more importantly, (2) it requires a large number of calls of the mechanical model for the computation of the small failure probabilities encountered in practice. The computation time becomes excessive when considering a small target value of the coefficient of variation of the failure probability. Thus, a method is needed that keeps to a minimum the number of calls to the mechanical model when performing a probabilistic analysis.

In order to overcome the shortcoming related to the excessive number of calls of the mechanical model when performing a probabilistic analysis, Echard et al. (2011) proposed an active learning reliability method combining kriging and Monte Carlo simulation (AK-MCS). This method consists of constructing a metamodel, i.e., an analytical equation which substitutes the original mechanical model. The computation of the failure probability can thus be easily performed using this metamodel.

The AK-MCS approach is based on the kriging theory and it makes use of a powerful learning function. In this method, an initial approximate kriging metamodel is constructed based on a small number of samples [called design of experiments (DoE)] computed using the mechanical model. This metamodel is then successively updated by adding each time a new sample chosen according to a powerful learning function until satisfying a prescribed stopping condition. Because the computation of the failure probability requires only the sign of the performance function values, the objective of the learning function is to choose samples that have a high uncertainty on the sign of their performance function values, i.e., those that are close to the limit state surface (LSS).

Echard et al. (2011) validated the AK-MCS approach by considering several academic examples involving nonlinear limit state surfaces and high-dimensional stochastic problems, in which the performance function was given by an analytical equation. They showed this method to be very efficient because the obtained probability of failure was very accurate, needing a smaller number of calls to the mechanical model than the crude MCS methodology.

This paper extends the AK-MCS approach by Echard et al. (2011) to the study of geotechnical structures involving spatially varying soil properties. More specifically, this paper presents a probabilistic analysis at the ultimate limit state of a strip footing

[1]Associate Professor, Faculty of Engineering, Lebanese Univ., Tripoli, Al Kobeh, Lebanon (corresponding author). ORCID: https://orcid.org/0000-0002-3382-7069. Email: Tamara.albittar@gmail.com

[2]Professor, Dept. of Civil Engineering, Univ. of Nantes, Blvd. de l'Université, 44603 Saint-Nazaire, France. Email: Abed.Soubra@univ-nantes.fr

[3]Ph.D. Student, Dept. of Civil Engineering, Univ. of Nantes, Blvd. de l'Université, 44603 Saint-Nazaire, France. Email: Jawad.thajeel@etu.univ-nantes.fr

Note. This manuscript was submitted on May 27, 2017; approved on May 1, 2018; published online on July 24, 2018. Discussion period open until December 24, 2018; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Geotechnical and Geoenvironmental Engineering***, © ASCE, ISSN 1090-0241.

resting on a spatially varying soil. The objective is the computation of the probability P_f of exceeding the ultimate bearing capacity of the footing under a prescribed vertical load. The soil cohesion and angle of internal friction (c and φ, respectively) are considered as two anisotropic non-Gaussian random fields. They are characterized by two specified marginal distribution functions and a common autocorrelation function. The expansion optimal linear estimation method (EOLE) methodology proposed by Li and Der Kiureghian (1993) was used to generate these two random fields. The mechanical model used to calculate the system response (i.e., the ultimate bearing capacity) was based on numerical simulations using the finite-difference code FLAC^{3D}.

This paper is organized as follows. The next section presents the probabilistic models. This is followed by an illustration of the AK-MCS approach via a simple bearing capacity problem (without spatial variability). Then a probabilistic analysis of a strip footing resting on a spatially varying soil and the corresponding numerical results are presented and discussed. The paper ends with a conclusion of the main findings.

Probabilistic Models

This section presents the probabilistic models used in this paper. The method of discretization of the random fields is presented first. It is followed by a brief presentation of the kriging metamodeling technique and the combined use of the kriging and the Monte Carlo simulation (i.e., the AK-MCS methodology) as it can be used for the probabilistic analyses of geotechnical structures involving spatially varying soil properties.

Random Fields Discretization Method

Assume that the soil cohesion c and friction angle φ are two non-Gaussian random fields that share the same autocorrelation function. These two random fields are denoted $Z_i^{\mathrm{NG}}(x,y)$ $(i=c,\varphi)$ and they are described by two non-Gaussian marginal cumulative density functions G_i $(i=c,\varphi)$ and a common autocorrelation function ρ_Z^{NG} $[(x,y),(x',y')]$ which gives the values of the correlation function between two arbitrary points (x,y) and (x',y').

This paper uses the following anisotropic square exponential autocorrelation function:

$$\rho_Z^{\mathrm{NG}}[(x,y),(x',y')]=\exp\left[-\left(\frac{|x-x'|}{a_x}\right)^2-\left(\frac{|y-y'|}{a_y}\right)^2\right] \quad (1)$$

where a_x and a_y = autocorrelation distances along x and y, respectively.

The discretization of the two random fields via EOLE is described as follows (Li and Der Kiureghian 1993): first define a stochastic grid composed of N_q grid points (or nodes), and then compute a common non-Gaussian autocorrelation matrix using Eq. (1). The common non-Gaussian autocorrelation matrix $\Sigma_{\chi;\chi}^{\mathrm{NG}}$ should be transformed into the Gaussian space using the correction functions proposed by Nataf (1962) because the discretization of the random fields using EOLE is performed in the Gaussian space. This obtains two Gaussian autocorrelation matrices $\Sigma_{\chi;\chi}^{c}$ and $\Sigma_{\chi;\chi}^{\varphi}$ that can be used to discretize the two Gaussian random fields at any point using

$$\tilde{Z}_i(x,y)=\mu_i+\sigma_i\sum_{j=1}^{M}\frac{\xi_j^i}{\sqrt{\lambda_j^i}}\cdot(\phi_j^i)^T\cdot\Sigma_{Z(x,y);\chi}^i \qquad i=c,\varphi \quad (2)$$

where μ_i and σ_i $(i=c,\varphi)$ = mean and standard deviation values of the two random fields, respectively; ξ_j^i $(i=c,\varphi;\ j=1,\ldots,M)$ = two blocks of independent standard normal random variables; λ_j^i and ϕ_j^i $(i=c,\varphi;\ j=1,\ldots,M)$ = eigenvalues and eigenvectors of the two Gaussian autocorrelation matrices $\Sigma_{\chi;\chi}^{c}$ and $\Sigma_{\chi;\chi}^{\varphi}$, respectively; $\Sigma_{Z(x,y);\chi}^i$ = correlation vector between the values of the random field at the different nodes of the stochastic grid and its value at an arbitrary point (x,y) obtained using Eq. (1); and M = number of terms (expansion order) retained in the EOLE method. This number is determined later in this section.

Once the two Gaussian random fields [Eq. (2)] are obtained, they should be transformed to the non-Gaussian space by applying the following formula:

$$\tilde{Z}_i^{\mathrm{NG}}(x,y)=G_i^{-1}\{\Phi[\tilde{Z}_i(x,y)]\} \qquad i=c,\varphi \quad (3)$$

where $\Phi(\cdot)$ = standard normal cumulative density function.

The series given by Eq. (2) are truncated for a number of terms M (expansion order) smaller than the number of grid points N_q, after sorting the eigenvalues λ_j^c and λ_j^{φ} $(j=1,\ldots,M)$ in descending order. This number should ensure that the variance of the error is smaller than a prescribed tolerance. The variance of the error for EOLE is given by Li and Der kiureghian (1993) as follows:

$$\mathrm{Var}[Z_i(x,y)-\tilde{Z}_i(x,y)]=\sigma_Z^2-\sum_{j=1}^{M}\frac{1}{\lambda_j^i}\left[(\phi_j^i)^T\Sigma_{Z(x,y);\chi}^i\right]^2 \quad (i=c,\varphi) \quad (4)$$

where $Z_i(x,y)$ and $\tilde{Z}_i(x,y)$ = exact and the approximate values, respectively, of the random fields at a given point (x,y).

Kriging Metamodeling Technique

This paper used a metamodel to predict the outcome from software for any sample, provided that the outcomes for a few other samples (i.e., DoE) are known; the software was used as a black-box. This paper used the kriging metamodeling technique. This kind of metamodeling can approximate arbitrary functions with high accuracy. The kriging metamodel does not assume an underlying global functional form. However, it is based on the assumption that there is a spatial correlation between the values of the function to be approximated.

The aim of this paper was to determine a metamodel of the performance function based on some performance function values $G(X)$ obtained from FLAC^{3D} software. The kriging theory considers the value of the performance function $G(X)$ of an unknown sample X as a realization $\widehat{G}(X)$ of a random function, which includes a regression part and a centered stochastic process as follows (Sacks et al. 1989):

$$\widehat{G}(X)=F(X,\beta)+Z(X) \quad (5)$$

where $F(X,\beta)$ = deterministic part defined by a regression model that gives an approximation of the mean of $G(X)$; and $Z(X)$ = fluctuation around this mean. It is given by a stationary Gaussian random process with zero mean and covariance (COVAR) that interpolates the errors between the regression model prediction and the true performance function values at the different N samples of the DoE. This covariance is given by

$$\mathrm{COVAR}[Z(X),Z(X')]=\sigma_Z^2R(X,X') \quad (6)$$

J. Geotech. Geoenviron. Eng., 2018, 144(10): 04018071

where σ_Z^2 = random field variance; X and X' = two arbitrary samples from the whole design of experiments; and $R(X, X')$ = spatial correlation function between these two samples with a correlation parameter vector θ. This paper selected ordinary kriging, which means that $F(X, \beta)$ is a scalar [i.e., $F(X, \beta) = \beta$] to be determined. Therefore, the estimated performance function $\widehat{G}(X)$ can be simplified as

$$G(X) = F(X, \beta) + Z(X) = \beta + Z(X) \tag{7}$$

Furthermore, the most widely used correlation function for reliability analysis is the anisotropic square exponential function. This paper used this function, which is given by

$$R(X, X') = \prod_{i=1}^{M} e^{[-\theta_i (x_i - x_i')^2]} \tag{8}$$

where M = number of random variables; x_i and $x_i' = i$th coordinates of the points X and X', respectively; and θ_i = scalar which is equal to the inverse of the correlation length in the ith direction.

Eqs. (5) and (7) show that the first part of the kriging metamodel is a regression model, which approximates the performance function over the whole design space. The second part is a stochastic process, which creates localized deviations from the regression model. This metamodel is completely defined by the scalar β, the correlation parameter vector θ, and the process variance σ_z^2. These parameters can be estimated by fitting the kriging metamodel to the design of experiments and the corresponding performance function values. They can be easily computed using the DACE toolbox in MATLAB (Lophaven et al. 2002). At this stage, all the three parameters of Eq. (7) are completely determined.

The best linear unbiased predictor (BLUP) of the performance function $\widehat{G}(X)$ at an unknown sample X is a Gaussian random variate $\widehat{G}(X) \sim N\left(\mu_{\widehat{G}(X)}, \sigma_{\widehat{G}(X)}\right)$ (Santner et al. 2003). The mean prediction $\mu_{\widehat{G}(X)}$ and the prediction variance $\sigma^2_{\widehat{G}(X)}$ of any unknown sample can be easily obtained by DACE toolbox in MATLAB making use of the already obtained values of β, σ_Z^2, and θ.

Contrary to other types of metamodels, the kriging metamodel provides not only a predicted value of an unknown sample but also an estimate of the prediction variance, which gives an uncertainty indication in the kriging metamodel of this sample. The variances of samples in the initial DoE are zero, but the variances of the other samples are always different from zero. A large value of $\sigma^2_{\widehat{G}(X)}$ means that the prediction is not exact. Therefore, the prediction variance $\sigma^2_{\widehat{G}(X)}$ is an important indicator of the unexplored areas and presents a good index to improve the initial DoE. This property is interesting and thus is used in the following paragraph.

AK-MCS Methodology for Geotechnical Structures Involving Spatially Varying Soil Properties

This section presents the AK-MCS methodology for geotechnical structures involving spatially varying soil properties and making use of computationally expensive numerical models (such as the FLAC^{3D} model used in this paper for the mechanical analysis). The basic idea of this approach is as follows.

In the AK-MCS method, a small design of experiments is randomly selected from a large Monte Carlo population. Then the kriging metamodeling technique is used to construct an approximate kriging metamodel based on the responses of this DoE computed using the mechanical model. This approximate kriging metamodel is successively improved (i.e., updated) by considering each time a new best sample that is computed using the mechanical model. This process, called an enrichment process, is repeated until satisfying a prescribed stopping criterion. A best new sample is chosen using a powerful learning function that makes use of the prediction mean $\mu_{\widehat{G}(X)}$ and the prediction variance $\sigma^2_{\widehat{G}(X)}$ of the already-obtained kriging metamodel.

As is shown subsequently, the chosen best new samples are those that are close to the limit state surface because the interest is not to determine accurate values of the performance function for the different samples but rather to accurately determine the signs of the performance function values for these samples. This is because the computation of the failure probability only requires the knowledge of the signs of the performance function values for the different samples.

A comprehensive step-by-step procedure describing the implementation of the AK-MCS methodology in the specific case of a strip footing resting on a spatially varying soil (where c and φ are two random fields and the numerical software used for the mechanical analysis is FLAC^{3D}) is given as follows:

1. Generate a population S of N_{MC} samples (say, 500,000 samples) of M standard Gaussian random variables $\{(\xi_1^1, \ldots, \xi_M^1), (\xi_1^2, \ldots, \xi_M^2), \ldots, (\xi_1^{N_{MC}}, \ldots, \xi_M^{N_{MC}})\}$, where M is the number of random variables needed by EOLE methodology to discretize the two random fields c and φ. Each sample of standard Gaussian random variables provides [when substituted into Eqs. (2) and (3)] typical spatial variations of c and φ that respect the correlation structure of these fields, i.e., the so-called realizations of c and φ. This is performed by computing, for this sample, the values of c and φ at the centroids of the different elements of the FLAC^{3D} mesh. The difference between the different realizations lies in the positions of the weak and strong soil zones, although all realizations respect the correlation structure of the corresponding random fields. The computation of the performance function values for the generated samples (based on FLAC^{3D} software) is not required at this stage. This paper denotes these as candidate samples.
2. Randomly select from the population S a small number of samples, i.e., a small DoE of size N_1 (say, $N_1 = 20$ samples). Then use EOLE methodology to transform each sample into realizations of c and φ that provide the spatial distribution of the soil cohesion and angle of internal friction respectively. For the N_1 samples, evaluate the performance function values using the following equation:

$$G = \frac{q_u}{q_s} - 1 \tag{9}$$

   where q_u = ultimate bearing capacity computed based on FLAC^{3D} software, and q_s = vertical loading applied to the footing.
3. Based on the DoE and the corresponding performance function values, construct an approximate kriging metamodel in the standard space of random variables using the DACE toolbox.
4. Use the DACE toolbox to compute (for the whole population S containing the N_{MC} samples) both the kriging predictor values $\mu_{\widehat{G}}$ and their corresponding kriging variance values $\sigma^2_{\widehat{G}}$ using the metamodel. From the obtained values of the kriging predictors $\mu_{\widehat{G}}$, obtain an estimation of the probability of failure P_f by counting the number of negative predictors $N_{\widehat{G}\leq 0}$ and dividing it by the total number of samples in S as follows:

$$P_f = \frac{N_{\widehat{G}\leq 0}}{N_{MC}} \tag{10}$$

Then compute the coefficient of variation (COV) of P_f as follows:

$$\mathrm{COV}(P_f) = \sqrt{\frac{1 - P_f}{P_f \cdot N_{\mathrm{MC}}}} \tag{11}$$

5. Identify the best next candidate sample in S for which to compute the performance function value using FLAC^{3D}. This is performed by evaluating a learning function U for each sample in S

$$U(X_i) = \frac{\left|\mu_{\widehat{G}(X_i)}\right|}{\sigma_{\widehat{G}(X_i)}} i = 1, \ldots, N_{\mathrm{MC}} \tag{12}$$

The best next sample is the one with minimum value of U; more details of this criterion for the choice of the best sample are given subsequently in this section. If this minimum value of U is smaller than 2, the performance function value based on FLAC^{3D} is evaluated for this best candidate and the initial DoE is updated. Then return to Step 3 and evaluate a new kriging model based on the updated DoE.

6. Steps 3, 4, and 5, which constitute the enrichment process, are repeated until the minimum value of U becomes larger than 2. More details of this stopping criterion are given subsequently in this section.

At this stage, the learning stops and the metamodel is considered accurate enough for the computation of the estimated values of both the probability of failure P_f and the coefficient of variation $\mathrm{COV}(P_f)$.

A small initial DoE is chosen in the present approach (Step 2) in order to keep to a minimum the number of calls to the computationally expensive mechanical model. This initial DoE is successively increased by a single sample each time (Step 5). The chosen sample is the one that mostly improves the metamodel because Eq. (12) searches for the sample that has a small kriging predictor (i.e., a sample that is close to the limit state surface) and/or a high kriging variance (i.e., a high uncertainty in the sign of its performance function value). The samples with high uncertainties in the sign of their performance function values (positive or negative) are those that are close to the limit state surface. Finally, the stopping criterion $[\min(U)] > 2$ corresponds to a maximal probability of making a mistake in the sign of the performance function value of $\Phi(-2) = 0.023$ (Echard et al. 2011). This means that the stopping criterion is relevant, making use of the samples with a small probability of making a mistake in the signs of their performance function values. The number of predictions by kriging can be important because the whole Monte Carlo population is estimated. However, the computation time of the predictions is much smaller than that required to evaluate the performance function values using the computationally expensive mechanical model.

Illustration of AK-MCS Procedure via Analytical Example

This section illustrates the performance of the AK-MCS procedure via a simple bearing capacity problem (i.e., without considering the soil spatial variability) in which the system response (ultimate bearing capacity) is given by a simple analytical equation with a quasi-negligible computation time. The performance of AK-MCS approach is illustrated by comparing its results with those given by the crude MCS methodology.

The probabilistic analysis presented in this section involves the computation of the failure probability against soil punching of a shallow strip footing of breadth $B = 2$ m resting on a homogeneous (c, φ) soil. A uniform surcharge loading $q = 10$ kN/m^2 was applied at the level of the base of the foundation on both sides of this foundation. The soil unit weight was $\gamma = 18$ kN/m^3. The applied footing load was equal to $q_s = 400$ kN/m. The uncertain parameters considered in the analysis were the soil shear strength parameters c and φ. The illustrative statistical parameters of these two random variables, as used in this analysis, are those commonly encountered in practice (e.g., Phoon and Kulhawy 1999) (Table 1). Eq. (9) is the performance function used to calculate the failure probability; q_u was calculated herein using the following equation:

$$q_u = \frac{1}{2}\gamma B N_\gamma + cN_c + qN_q \tag{13}$$

where N_γ, N_q, and N_c = bearing capacity factors due to the soil weight, surcharge loading, and cohesion, respectively. The adopted factors used in this section are those suggested by Vesic (1973). These factors are widely used in routine foundation design. They are functions only of the soil angle of internal friction φ.

The reference values adopted for both the failure probability P_f and the coefficient of variation $\mathrm{COV}(P_f)$ were those obtained by the crude MCS runs with $N_{\mathrm{MC}} = 10^6$ samples. Both methods (MCS and AK-MCS) provided the same values for the probabilistic outputs, i.e., P_f and $\mathrm{COV}(P_f)$ in Table 2, although the AK-MCS method needed only 30 calls of the mechanical model (which corresponded to 20 samples from the initial DoE plus 10 samples added during the enrichment process). This number of calls to the mechanical model is to be compared to that required by MCS, which needed 10^6 samples to lead to the same value of $\mathrm{COV}(P_f)$. This clearly illustrates the benefit of using the AK-MCS approach instead of the crude MCS methodology. The deterministic safety factor for the studied configuration was $F_s = q_u/q_s = 1190/400 \approx 3.0$.

In order to better understand the impact of the number of samples added during the enrichment process on the probabilistic outputs, Figs. 1(a and b) plot P_f and $\mathrm{COV}(P_f)$ versus the number of added samples. Fig. 1(c) presents the value of the learning function U that was obtained each time the mechanical model computed a new added sample with index i ($i = 1, \ldots, 10$ in this example).

The new samples chosen during the enrichment process for the computation by the mechanical model had increasing values of U [Fig. 1(c)]. The process of adding new samples stopped (as required by the algorithm of the AK-MCS approach) when the value of U became greater than 2. This analytical problem required 10 added

Table 1. Statistical characteristics of random variables

Random variable	Mean, μ	Coefficient of variation (COV) (%)	Type of the probability density function (PDF)
c	20 kPa	25	Log-normal
φ	30°	10	Beta

Table 2. Failure probability P_f, coefficient of variation $\mathrm{COV}(P_f)$, and number of calls of mechanical model $\mathrm{N_{calls}}$ obtained by MCS and AK-MCS

Method	P_f ($\times 10^{-4}$)	$\mathrm{COV}(P_f)$ (%)	$\mathrm{N_{calls}}$
MCS	6.530	3.912	10^6
AK-MCS	6.530	3.910	20 samples for the initial DoE + 10 added samples = 30 samples

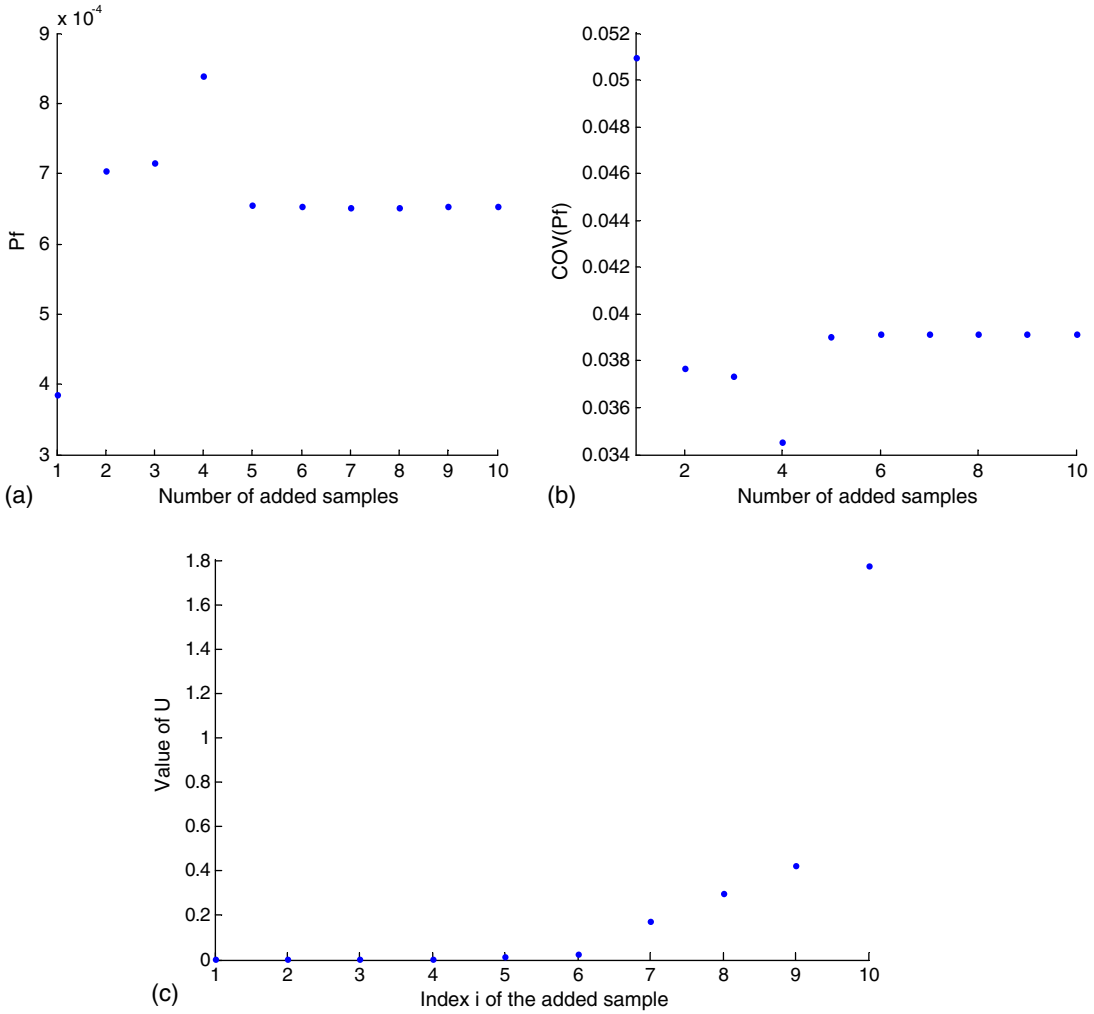

Fig. 1. AK-MCS results for case of homogeneous soil: (a) P_f versus number of added samples; (b) COV(P_f) versus number of added samples; and (c) value of U for different numbers of added samples ($i = 1, \ldots, 10$).

samples to achieve this goal. On the other hand, Figs. 1(a and b) show that P_f and COV(P_f) attained an asymptote when the number of added samples became greater than 5. This means that this number of added samples could be used to obtain the values of P_f and COV(P_f) corresponding to convergence. For this number of added samples, the metamodel became sufficiently accurate to compute P_f and COV(P_f).

As a conclusion, the adopted stopping criterion $[\min(U)] > 2$ was quite severe [more restrictive than the simple visual inspection of convergence observed in Fig. 1(a)] and thus, it led to convergence of P_f because it suggested adding all the samples that were close to the limit state surface ($G = 0$), although the value of P_f likely stabilized. This criterion was adopted for all the soil configurations studied later in this paper.

AK-MCS Results for Spatially Varying Soil Properties

This section presents the impact of the soil spatial variability on the probability of failure against soil punching of a strip footing subjected to a vertical load. Thus, the system response involves the ultimate bearing capacity q_u of a vertically loaded strip footing resting on a spatially varying soil. This analysis used the performance function defined by Eq. (9) in which q_u was based on numerical simulations using the FLAC^{3D} model.

The soil cohesion c and angle of internal friction φ were modeled as two anisotropic non-Gaussian random fields. The EOLE methodology was used to discretize the two random fields, i.e., to obtain realizations of the soil cohesion c and angle of internal friction φ that respect the correlation structure of those fields. Table 1 presents the illustrative statistical parameters of these two random fields as used in the analysis. The same autocorrelation function (square exponential) was used for both c and φ. In addition, the soil dilation angle ψ was considered to be related to the soil angle of internal friction φ by $\psi = 2\varphi/3$. This means that the soil dilation angle was implicitly assumed as a random field that was perfectly correlated with the soil angle of internal friction random field. All the other parameters of the soil, footing, and interface were assumed to be deterministic. Furthermore, the number of Monte Carlo realizations N_{MC} used in all subsequent computations was 500,000.

The mechanical model considered a strip footing of breadth $B = 1$ m resting on a soil domain of width $13B$ and depth $5B$ (Al-Bittar and Soubra 2014a). Der Kiureghian and Ke (1988) suggested that the length of the largest element of the deterministic mesh in a given direction (horizontal or vertical) should not exceed 0.5 times the autocorrelation distance in that direction. In order to

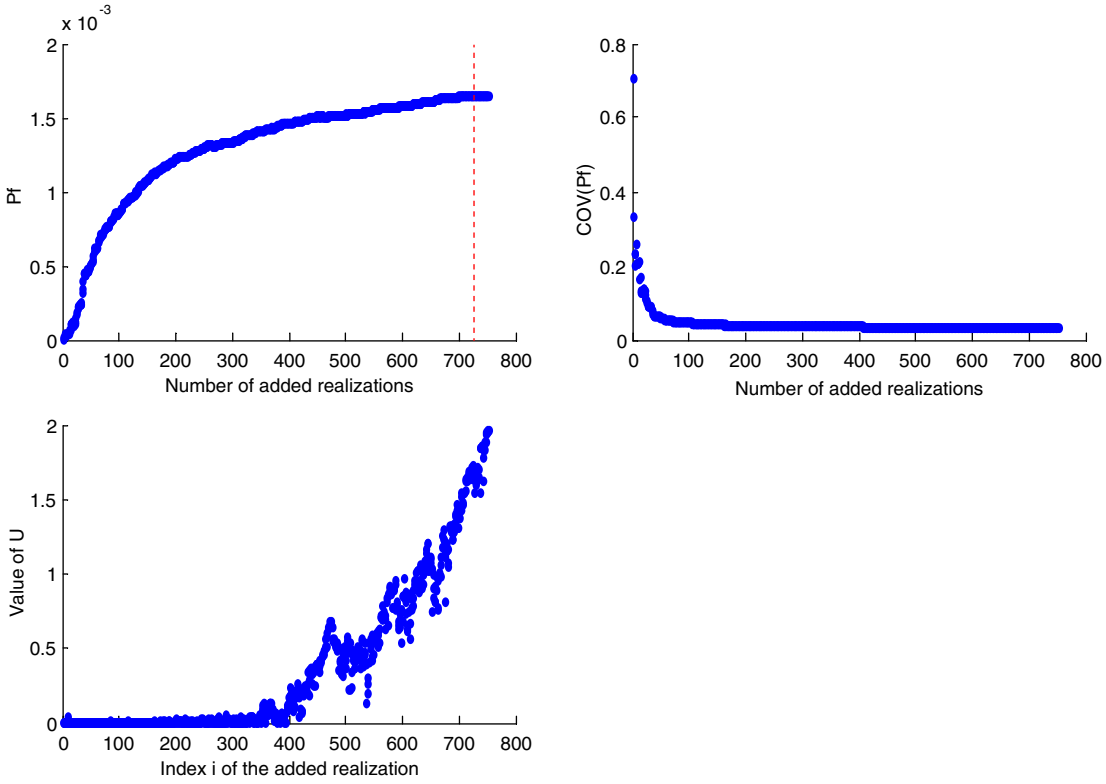

Fig. 2. AK-MCS results for spatially varying soil ($a_x = 10$ m, $a_y = 1$ m).

respect this criterion for the different autocorrelation distances, the FLAC^{3D} model considered two different deterministic meshes (Al-Bittar and Soubra 2014a). The first was devoted to the case of moderate to large values of the autocorrelation distances (i.e., when $a_x \geq 10$ m and $a_y \geq 1$ m) and the second considered small values of the autocorrelation distances (i.e., when $2\text{ m} \leq a_x < 10$ m or $0.5\text{ m} \leq a_y < 1$ m). Al-Bittar and Soubra (2014a) presented other details concerning the stochastic mesh and the deterministic model, and therefore they are not provided here in order to avoid repetition.

Probabilistic Numerical Results

The desired level of accuracy of the random field discretization may be achieved by imposing a threshold value for the variance of error computed using Eq. (4). The adopted number of random variables should lead to a computed variance of error that is smaller than this prescribed value. Several authors proposed a target value of 10% for the variance of error (e.g., Sudret and Der kiureghian 2000; Al-Bittar and Soubra 2013, 2014a).

All configurations considered in the following sections were studied with a target value of the variance of the error of 5% except two configurations [$(a_x = a_y = 2$ m) and ($a_x = 10$ m, $a_y = 0.5$ m)] corresponding to small values of the autocorrelation distance which adopted a greater value of the variance of error of about 10% (Tables 4–6). Consequently, the discretised random field can be considered to be sufficiently accurate for all the configurations considered in the paper.

Failure Probability P_f and COV(P_f) versus Number of Added Realizations

Fig. 2 presents the probability of failure P_f and the coefficient of variation COV(P_f) versus the number of added realizations for the case in which $a_x = 10$ m and $a_y = 1$ m. The number of random variables adopted for this configuration was 32 (Table 5). Fig. 2 also provides the learning function values for the different added realizations; 752 realizations were needed in the enrichment process in addition to the initial DoE before the algorithm stopped $[\min(U) > 2]$. The final obtained values of P_f and COV(P_f) were respectively 1.656×10^{-3} and 3.47%. Fig. 2 shows that the probability of failure started to converge at about 727 calls to the mechanical model. This means that there was no bias in the metamodel (in the zone of interest for the computation of the failure probability) beyond this number of calls of the mechanical model. However, the addition of new samples continued until satisfying the stopping criterion $[\min(U)] > 2$.

Effect of Number of Added Realizations on Limit State Surface

Fig. 3 shows the evolution of the limit state surface corresponding to $G = 0$, with the number of added realizations (from 5 to 35 realizations) for the case in which $a_x = a_y = 10{,}000$ m. Two random variables were adopted for the discretization of the random fields c and φ for this configuration (Table 4). Fig. 3 is represented in the standard Gaussian space in which ξ_c and ξ_φ are the standard Gaussian random variables needed to discretize the random fields c and φ, respectively. This small number of random variables (i.e., 2) allowed visualizing the evolution of the LSS because only a two-dimensional space was needed.

The increase in the number of added realizations improved the limit state surface in the zone that has an impact on the value of the failure probability (i.e., the zone corresponding to the high values of the probability density which is close to the origin of the standard space); the other zones (far from the origin) has a nonsignificant effect on the value of the failure probability (Fig. 3).

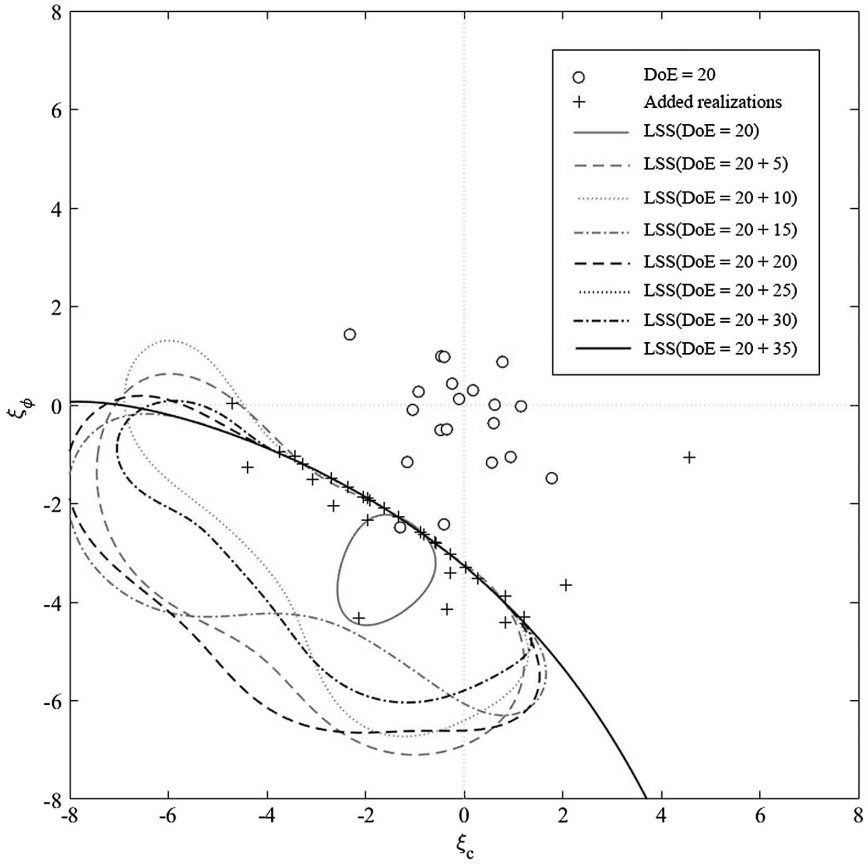

Fig. 3. Effect of number of added realizations on limit state surface when $a_x = a_y = 10{,}000$ m.

Table 3. Effect of number of random variables on P_f and $\mathrm{COV}(P_f)$ for case in which $a_x = 10$ m and $a_y = 5$ m

Number of random variables for two random fields (c, φ)	Variance of error (%)	P_f ($\times 10^{-3}$)	$\mathrm{COV}(P_f)$ (%)	δ_{P_f} (%)	Number of added realizations
6	9.9	2.822	2.66	19.5	115
12	1.7	3.25	2.48	7.3	406
24	0.08	3.506	2.38	—	695

Effect of Variance of Error of EOLE Methodology on Probabilistic Outputs

In order to reduce the effect of the error of discretization of random fields on the estimated value of the failure probability, the variance of the error of EOLE methodology must be sufficiently small. Table 3 provides the effect of the number of random variables used in the discretization scheme (and the corresponding value of the variance of the error of EOLE methodology) on the value of P_f for the case in which $a_x = 10$ m and $a_y = 5$ m.

The percentage difference in the P_f value (with respect to the accurate solution corresponding to a very small value of the variance of error of about 0.08%) was 19.5% when adopting a variance of error of about 10%, and it decreased to 7.3% with a variance of error of 1.7%. This means that the percentage difference between the P_f value corresponding to a variance of the error of 5% (the value adopted in this paper) and the accurate P_f value was expected to be about 12%, which may be considered to be sufficiently accurate for engineering applications; the value of the variance of error adopted in the literature (i.e., 10%) led to a greater (although acceptable) error of 19.5% on the value of P_f.

Validation of AK-MCS Approach by Comparison with Crude MCS

In order to show the efficiency of the proposed AK-MCS methodology, it was compared with the crude Monte Carlo simulation method. The computation considered a total of 136,959 simulations.

Fig. 4 presents the evolution of P_f and $\mathrm{COV}(P_f)$ as function of the number of simulations as obtained from MCS. Fig. 4 also presents the values obtained from the AK-MCS approach for comparison.

The values of P_f and $\mathrm{COV}(P_f)$ obtained from the crude MCS were, respectively, 1.701×10^{-3} and 6.54%. Those values were compared with the values obtained from the AK-MCS approach [i.e., $P_f = 1.656 \times 10^{-3}$ and $\mathrm{COV}(P_f) = 3.47\%$]. The results showed good agreement in term of the value of P_f. The AK-MCS approach was clearly more efficient than the MCS because only 772 samples (which corresponds to 20 samples from the initial DoE plus 752 samples added during the enrichment process) were needed to obtain an accurate value of P_f with a value of $\mathrm{COV}(P_f)$ that was smaller than the value obtained using the crude MCS with 136,959 simulations. A better agreement with MCS in terms of P_f and $\mathrm{COV}(P_f)$ could be obtained for a much greater number of MCS simulations (about 500,000), but this was not done because it is very time-comsuming.

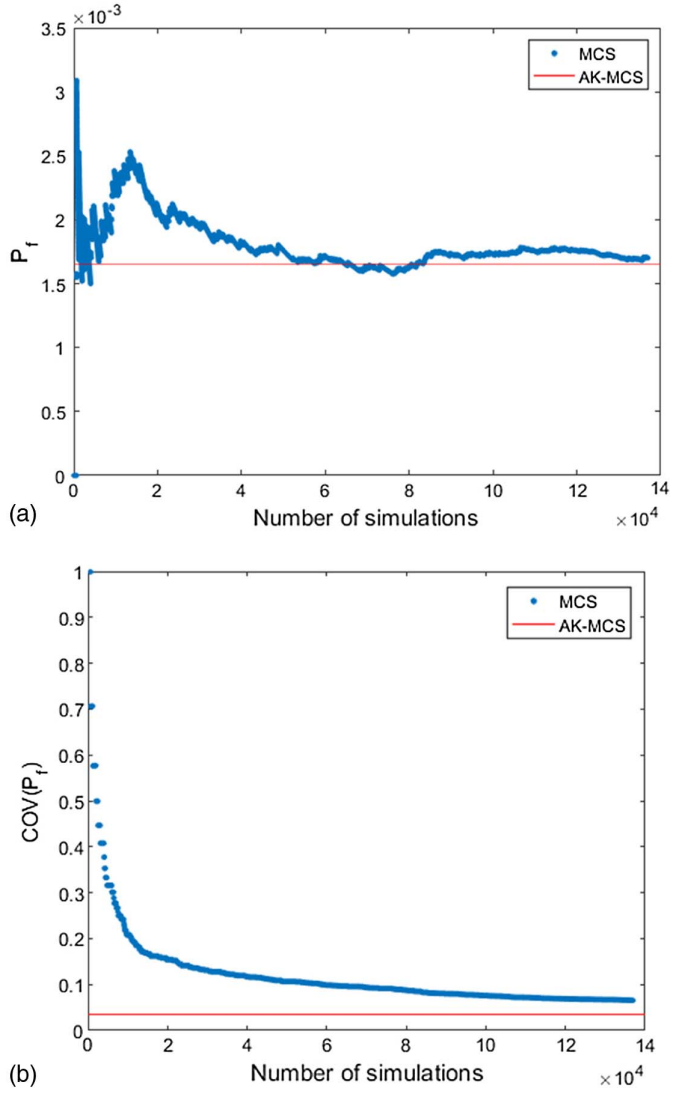

Fig. 4. Comparison of crude Monte Carlo simulation and AK-MCS approach for reference case in which $a_x = 10$ m and $a_y = 1$ m: (a) P_f versus number of simulations; and (b) COV(P_f) versus number of simulations.

Probabilistic Parametric Study

Columns 2 and 3 of Tables 4–6 provide the number of eigenmodes and the corresponding variance of the error of EOLE methodology for different values of the autocorrelation distances. Columns 4, 5, and 6 of the same tables provide the failure probabilities, the corresponding values of the coefficient of variation, and the number of added realizations.

A maximal target value of 5% was adopted for the variance of the error of EOLE methodology except for two configurations in which a variance of error of about 10% was considered. As expected, the number of random variables was small for very large values of the autocorrelation distances and significantly increases for small values of the autocorrelation distances (Tables 4–6). A greater value of the variance of the error was considered for the two configurations because of the huge computation time required for these cases. Small values of the autocorrelation distances required a greater number of random variables, leading to a high-dimensional stochastic problem with greater heterogeneity. These configurations were typical of very small failure probabilities (on the order of 10^{-4}). In order to allow the AK-MCS approach to handle these cases, the population S must be increased from 500,000 to 1,000,000 or more. However, increasing the population S significantly increases the computation time. This is because in the process of enrichment (i.e., the process of adding new samples to the DoE), the AK-MCS approach requires the computation of the U values for all the samples in the population S. The problem of a very large population did not appear in the academic example by Echard et al. (2011) because the value of the probability of failure they considered was relatively large (on the order of 10^{-3}). In addition, the computational model used was an analytical equation (not a computationally expensive numerical model) which made the use of a large population S feasible. In conclusion, the excessive computation time that appeared in the two configurations was not related to the number of random variables but was inherent in the very small values of the failure probability corresponding to these cases.

Table 4 presents the effect of the isotropic autocorrelation distance ($a_x = a_y$) on the failure probability P_f and the corresponding value of the coefficient of variation COV(P_f); P_f increased with the increase in the autocorrelation distance (Fig. 5). However, the rate of increase decreased for large values of the autocorrelation distances (when $a_x = a_y > 10$ m) to attain an asymptote corresponding to the case of a homogeneous soil. Indeed, for small values of the isotropic autocorrelation distance, the soil heterogeneity led (for most realizations) to relatively high values of the ultimate bearing capacity due to the averaging phenomenon. This means that the number of realizations leading to failure was very small in this case. In contrast, the number of realizations leading to failure was higher in the case of a homogeneous soil due to the fact that the realizations were homogeneous in this case, with either small or large values of soil resistance.

For the case of anisotropic random fields, Table 5 presents the effect of the vertical autocorrelation distance (a_y) on the failure probability P_f when $a_x = 10$ m; this study introduced some non-practical values in which the vertical autocorrelation distance was

Table 4. Adopted number of random variables and corresponding value of variance of error of EOLE together with values of P_f, COV(P_f), and number of added realizations for case of isotropic soil

$a_x = a_y$ (m)	Adopted number of random variables	Variance of error (%)	P_f ($\times 10^{-3}$)	COV(P_f) (%)	Number of added realizations
2	48	9.45	0.730	5.23	742
3	32	4.65	1.846	3.29	995
5	24	0.95	2.762	2.69	870
10	10	0.81	3.444	2.41	286
20	8	0.17	3.648	2.34	210
50	6	0.02	3.736	2.31	105
100	6	0.001	3.772	2.29	100
10,000	2	0.000048	3.806	2.28	37

Table 5. Adopted number of random variables and corresponding value of variance of error of EOLE together with values of P_f, COV(P_f), and number of added realizations for case of anisotropic soil ($a_x = 10$ m with varying a_y)

a_y (m)	Adopted number of random variables	Variance of error (%)	P_f ($\times 10^{-3}$)	COV(P_f) (%)	Number of added realizations
0.5	44	9.12	0.318	7.93	427
0.8	38	4.80	1.178	4.12	790
1	32	4.21	1.656	3.47	752
2	24	1.44	2.818	2.66	672
5	12	1.68	3.250	2.48	406
10	10	0.81	3.444	2.41	286
20	8	0.85	3.502	2.39	190
50	8	0.30	3.570	2.36	239
100	8	0.10	3.616	2.35	232
10,000	8	0.03	3.690	2.32	227

Table 6. Adopted number of random variables and corresponding value of variance of error of EOLE together with values of P_f, COV(P_f), and number of added realizations for case of anisotropic soil ($a_y = 2$ m with varying a_x)

a_x (m)	Adopted number of random variables	Variance of error (%)	P_f ($\times 10^{-3}$)	COV(P_f) (%)	Number of added realizations
2	48	9.45	0.730	5.23	742
5	30	4.10	2.116	3.07	824
10	24	1.44	2.818	2.66	672
20	16	1.41	3.060	2.55	494
50	12	1.27	3.166	2.51	357
100	10	0.84	3.202	2.49	256
10,000	10	0.43	3.208	2.49	257

larger than the horizontal autocorrelation distance in order to check the behavior of the quantity of interest (i.e., P_f) when the vertical autocorrelation distance varied within a wide range of values. Similarly, Table 6 presents the effect of the horizontal autocorrelation distance (a_x) on the failure probability when $a_y = 2$ m. For very large values of the horizontal or vertical autocorrelation distance, the failure probability tended to a constant maximal value corresponding to the case of a one-dimensional (1D) random field (Figs. 6 and 7). The reason is similar to that of the isotropic case. Indeed, the increase in the soil heterogeneity led (for most realizations) to relatively higher values of the ultimate bearing capacity due to the averaging phenomenon. This means that the number of realizations leading to failure was very small in this case and increased for a lesser degree of heterogeneity in the soil mass.

For both cases of isotropic and anisotropic random fields, Figs. 5–7 show that the Hasofer–Lind reliability index (as obtained from optimization and making use of the final kriging meta-model) decreased as expected with the increase in the autocorrelation distances. Furthermore, a small value of the coefficient of variation of the failure probability (smaller than 6% for most cases) was obtained for the adopted value of N_{MC}, which indicates that the obtained results are sufficiently accurate for practical use (Tables 4–6). The number of added realizations required to lead to a good approximation of the kriging model seemed to be larger

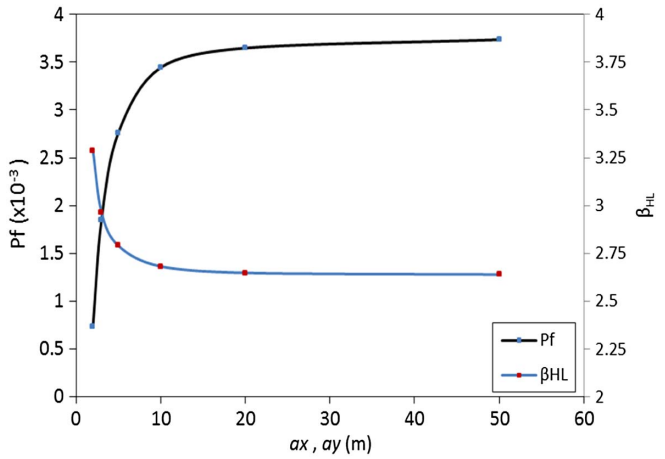

Fig. 5. Effect of isotropic autocorrelation distance $a_x = a_y$ on P_f and β_{HL}.

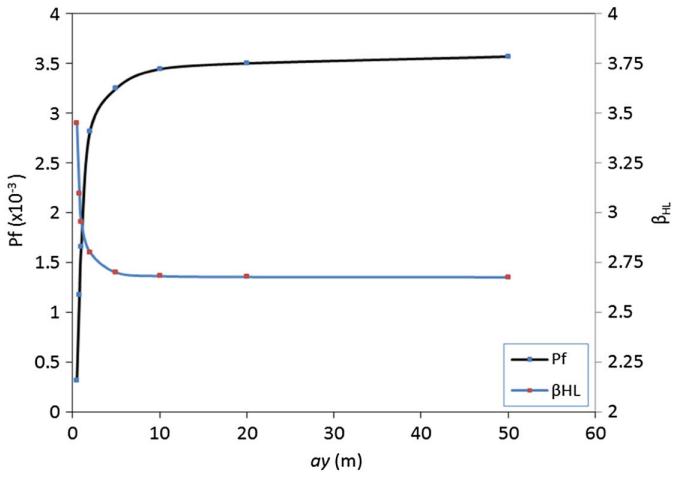

Fig. 6. Effect of vertical autocorrelation distance a_y on P_f and β_{HL} when $a_x = 10$ m.

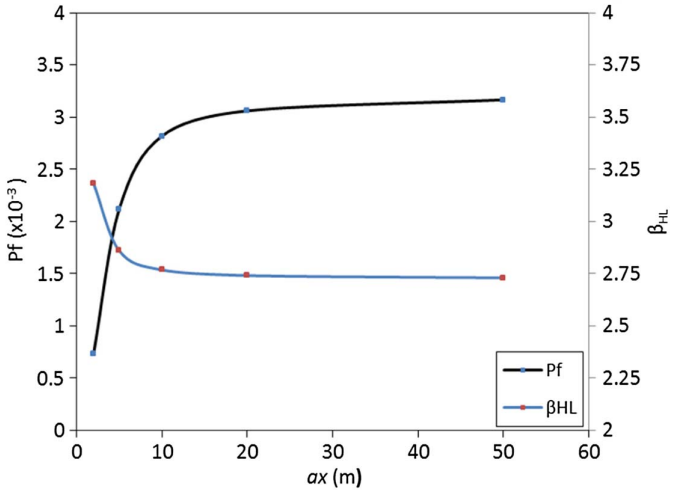

Fig. 7. Effect of horizontal autocorrelation distance a_x on P_f and β_{HL} when $a_y = 2$ m.

for smaller values of the autocorrelation distance (because of the increasing fluctuations of the highly heterogeneous soils), although there was no regular increase in the number of added realizations with the decrease in the autocorrelation distance. Indeed, this number depends on the evolution of the kriging metamodel during the enrichment process.

Conclusion

The literature generally performs probabilistic analysis of shallow foundations resting on spatially varying soils using the Monte Carlo simulation methodology. The mean value and the standard deviation of the system response have been extensively investigated. This is not the case for the probability of failure, because MCS methodology requires a large number of calls of the mechanical model to accurately calculate the small failure probabilities encountered in practice.

This paper presented a probabilistic analysis at the ultimate limit state of a strip footing resting on a spatially varying soil using an active learning reliability method combining kriging and Monte Carlo simulation. This method involves performing a Monte Carlo simulation without evaluating the whole population using the original computationally expensive mechanical model. Indeed, the population is predicted using a kriging metamodel which is defined based on only a few points of the population that are evaluated using the mechanical model. The main findings of this study are summarized as follows:

- The AK-MCS method was shown to be very efficient because the obtained probability of failure was very accurate [i.e., with a small $\mathrm{COV}(P_f)$] using a smaller number of calls to the computationally expensive mechanical model compared with the crude MCS methodology.
- The failure probability P_f increased with the increase in the autocorrelation distances. However, the rate of increase decreased for large values of autocorrelation distances (when $a_x = a_y > 10$ m, $a_y > 5$ m or $a_x > 10$ m) . Indeed, these cases closely resembled the limit cases corresponding to a homogeneous soil or a 1D random field.
- A more relevant stopping criterion that is based on the convergence of the quantity of interest (i.e., P_f) rather than on the accuracy of the metamodel would lead to a further reduction in the number of calls of the mechanical model. This is the object of future research.

References

Al-Bittar, T., and A. H. Soubra. 2013. "Bearing capacity of strip footings on spatially random soils using sparse polynomial chaos expansion." *Int. J. Numer. Anal. Methods Geomech.* 37 (13): 2039–2060. https://doi.org/10.1002/nag.2120.

Al-Bittar, T., and A. H. Soubra. 2014a. "Efficient sparse polynomial chaos expansion methodology for the probabilistic analysis of computationally-expensive deterministic models." *Int. J. Numer. Anal. Methods Geomech.* 38 (12): 1211–1230. https://doi.org/10.1002/nag.2251.

Al-Bittar, T., and A. H. Soubra. 2014b. "Probabilistic analysis of strip footings resting on spatially varying soils and subjected to vertical or inclined loads." *J. Geotech. Geoenviron. Eng.* 38 (12): 04013043. https://doi.org/10.1061/(ASCE)GT.1943-5606.0001046.

Al-Bittar, T., and A. H. Soubra. 2016. "Bearing capacity of spatially random rock masses obeying Hoek–Brown failure criterion." *Georisk: Assess. Manage. Risk Eng. Syst. Geohazards* 11 (2): 215–229. https://doi.org/10.1080/17499518.2016.1232831.

Ching, J., Y. G. Hu, Z. Y. Yang, J. Q. Shiau, J. C. Chen, and Y. S. Li. 2011. "Reliability-based design for allowable bearing capacity of footings on rock masses by considering angle of distortion." *Int. J. Rock Mech. Min. Sci.* 48 (5): 728–740. https://doi.org/10.1016/j.ijrmms.2011.05.005.

Ching, J. Y., X. W. Tong, and Y. G. Hu. 2016. "Effective Young's modulus for a spatially variable soil mass subjected to a simple stress state." *Georisk* 10 (1): 11–26. https://doi.org/10.1080/17499518.2015.1084426.

Cho, S. E., and H. C. Park. 2010. "Effect of spatial variability of cross-correlated soil properties on bearing capacity of strip footing." *Int. J. Numer. Anal. Methods Geomech.* 34 (1): 1–26. https://doi.org/10.1002/nag.791.

Der Kiureghian, A., and J. B. Ke. 1988. "The stochastic finite element method in structural reliability." *Probab. Eng. Mech.* 3 (2): 83–91. https://doi.org/10.1016/0266-8920(88)90019-7.

Echard, B., N. Gayton, and M. Lemaire. 2011. "AK-MCS: An active learning reliability method combining Kriging and Monte Carlo simulation." *Struct. Saf.* 33 (2): 145–154. https://doi.org/10.1016/j.strusafe.2011.01.002.

Fenton, G. A., and D. V. Griffiths. 2001. "Bearing capacity of spatially random soil: The undrained clay Prandtl problem revisited." *Géotechnique* 51 (4): 351–359. https://doi.org/10.1680/geot.2001.51.4.351.

Fenton, G. A., and D. V. Griffiths. 2003. "Bearing-capacity prediction of spatially random $c - \varphi$ soils." *Can. Geotech. J.* 40 (1): 54–65. https://doi.org/10.1139/t02-086.

Griffiths, D. V., G. A. Fenton, and N. Manoharan. 2002. "Bearing capacity of rough rigid strip footing on cohesive soil: Probabilistic study." *J. Geotech. Geoenviron. Eng.* 128 (9): 743–755. https://doi.org/10.1061/(ASCE)1090-0241(2002)128:9(743).

Li, C. C., and A. Der Kiureghian. 1993. "Optimal discretization of random fields." *J. Eng. Mech.* 119 (6): 1136–1154. https://doi.org/10.1061/(ASCE)0733-9399(1993)119:6(1136).

Li, J. H., M. J. Cassidy, Y. Tian, J. Huang, A. V. Lyamin, and M. Uzielli. 2016. "Buried footings in random soils: Comparison of limit analysis and finite element analysis." *Georisk* 10 (1): 55–65. https://doi.org/10.1080/17499518.2015.1064141.

Li, J. L., M. J. Cassidy, and Y. Tian. 2015. "Failure mechanism and bearing capacity of footings buried at various depths in spatially random soil." *J. Geotech. Geoenviron. Eng.* 141 (2): 04014099. https://doi.org/10.1061/(ASCE)GT.1943-5606.0001219.

Lophaven, S. N., H. B. Nielsen, and J. Søndergaard. 2002. *DACE—A Matlab Kriging toolbox, version 2.0*. Technical Rep. IMM-TR-2002-12. Lyngby, Denmark: Technical Univ. of Denmark.

Nataf, A. 1962. "Détermination des distributions de probabilités dont les marges sont données." *Comptes Rendus de l'Académie des Sciences* 255: 42–43.

Phoon, K.-K., and F. H. Kulhawy. 1999. "Evaluation of geotechnical property variability." *Can. Geotech. J.* 36 (4): 625–639. https://doi.org/10.1139/t99-039.

Popescu, R., G. Deodatis, and A. Nobahar. 2005. "Effects of random heterogeneity of soil properties on bearing capacity." *Probab. Eng. Mech.* 20 (4): 324–341. https://doi.org/10.1016/j.probengmech.2005.06.003.

Sacks, J., W. J. Welch, T. J. Mitchell, and H. P. Wynn. 1989. "Design and analysis of computer experiments." *Stat. Sci.* 4 (4): 409–423. https://doi.org/10.1214/ss/1177012413.

Santner, T. J., B. J. Williams, and W. I. Notz. 2003. *The design and analysis of computer experiments*. New York: Springer.

Sudret, B., and A. Der Kiureghian. 2000. *Stochastic finite element methods and reliability: A state-of-the-art report*. Rep. No. UCB/SEMM-2000/08. Berkeley, CA: Dept. of Civil and Environmental Engineering.

Vesic, A. S. 1973. "Analysis of ultimate loads of shallow foundations." *J. Soil Mech. Found. Div.* 99 (1): 45–73. https://doi.org/10.1016/0148-9062(74)90598-1.