



HAL
open science

Efficient sparse polynomial chaos expansion methodology for the probabilistic analysis of computationally-expensive deterministic models

Tamara Al-Bittar, Abdul-Hamid Soubra

► To cite this version:

Tamara Al-Bittar, Abdul-Hamid Soubra. Efficient sparse polynomial chaos expansion methodology for the probabilistic analysis of computationally-expensive deterministic models. *International Journal for Numerical and Analytical Methods in Geomechanics*, 2014, 38, pp.1211-1230. 10.1002/nag.2251 . hal-01006785

HAL Id: hal-01006785

<https://hal.science/hal-01006785>

Submitted on 1 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Efficient sparse polynomial chaos expansion methodology for the probabilistic analysis of computationally-expensive deterministic models

T. Al-Bittar^{1,*,\dagger} and A.-H. Soubra²

¹*Univ. of Balamand, Koura, North Lebanon, Lebanon*

²*University of Nantes, Bd. de l'université, BP 152, 44603 Saint-Nazaire cedex, France*

The sparse polynomial chaos expansion (SPCE) methodology is an efficient approach that deals with uncertainties propagation in case of high-dimensional problems (i.e., when a large number of random variables is involved). This methodology significantly reduces the computational cost with respect to the classical full PCE methodology. Notice however that when dealing with computationally-expensive deterministic models, the time cost remains important even with the use of the SPCE. In this paper, an efficient combined use of the SPCE methodology and the Global Sensitivity Analysis is proposed to solve such problem. The proposed methodology is firstly validated using a relatively non-expensive deterministic model that involves the computation of the PDF of the ultimate bearing capacity of a strip footing resting on a weightless spatially varying soil where the soil cohesion and angle of internal friction are modeled by two anisotropic non-Gaussian cross-correlated random fields. This methodology is then applied to an expensive model that considers the case of a ponderable soil. A brief parametric study is presented in this case to show the efficiency of the proposed methodology.

KEY WORDS: spatial variability; sparse polynomial chaos expansion; global sensitivity analysis; Sobol indices

1. INTRODUCTION

An efficient approach to deal with uncertainties propagation in case of high-dimensional problems (i.e., when a large number of random variables is involved) was recently presented by [1]. This approach is based on sparse polynomial chaos expansion (SPCE) for the system response and leads to a reduced computational cost as compared with the classical PCE methodology. Notice that both, the PCE and SPCE methodologies, aim at replacing the original complex deterministic model which may be an analytical model or a finite element/finite difference model by a meta-model. This allows one to calculate the system response [when performing the probabilistic analysis using Monte Carlo simulations (MCSs)] using a simple analytical equation [1–11]. It should be emphasized that when dealing with spatially varying soils together with computationally-expensive deterministic models, the time cost remains important even with the use of the SPCE methodology. Consequently, a method that can reduce once again the cost of the probabilistic analysis is needed.

In this paper, an efficient approach is proposed. The basic idea of this method is that, for a given discretized random field, the obtained random variables do not have the same weight in the variability of the system response. The variables with a very small contribution in the variability of

*Correspondence to: Tamara Al-Bittar, Univ. of Balamand, Koura, North Lebanon, Lebanon.

^{\dagger}E-mail: tamara.albittar@balamand.edu.lb

the system response can be discarded, which significantly reduces the dimensionality of the treated problem. This allows one to perform a probabilistic analysis using a reduced experiment design and thus a smaller number of calls of the computationally-expensive deterministic model. The main challenge remains in detecting the most influential random variables in order to reduce the dimensionality of the problem. An efficient combination between the SPCE methodology and the Global Sensitivity Analysis (GSA) is proposed in this regard.

The proposed methodology is firstly validated using a relatively non-expensive deterministic model. This model was extensively investigated by [9, 10] using the classical SPCE methodology. It involves the computation of the ultimate bearing capacity of a strip footing resting on a weightless spatially varying soil, where the soil cohesion and angle of internal friction (c and φ) are modeled by two anisotropic non-Gaussian (NG) cross-correlated random fields. In a second stage, a computationally-expensive deterministic model that involves the computation of the PDF of the ultimate bearing capacity of a strip footing resting on a ponderable spatially varying (c and φ) soil was briefly investigated to show the efficiency of the proposed methodology.

The paper is organized as follows: The next two sections aim at briefly presenting both the SPCE methodology and the GSA. Then, the proposed efficient combination between the SPCE methodology and the GSA is presented. It is followed by the numerical results. The paper ends with a conclusion.

2. SPARSE POLYNOMIAL CHAOS EXPANSION METHODOLOGY

In this section, one first presents the PCE and then its extension, the SPCE. The PCE methodology allows one to replace a complex deterministic model which may be an analytical model or a finite element/finite difference numerical model by a meta-model. Thus, the system response may be calculated (when performing the probabilistic analysis using MCS methodology) using a simple analytical equation. This equation is obtained by expanding the system response on a suitable basis which is a series of multivariate polynomials that are orthogonal with respect to the joint probability density function of the random variables. The PCE theory was originally formulated with standard Gaussian random variables and Hermite polynomials [12]. It was later extended to other types of random variables that use other types of polynomials [13]. In this paper, standard normal random variables in conjunction with Hermite polynomials are used. The coefficients of the PCE may be efficiently computed using a non-intrusive technique where the deterministic calculations are carried out using for example an analytical model or a finite element/finite difference software treated as a black box. The most used non-intrusive method is the regression approach [1–11]. This method is used in the present work. The PCE methodology can be briefly described as follows:

For a deterministic model with M random variables, the system response Γ can be expressed by a PCE as follows:

$$\Gamma_{PCE}(\xi) = \sum_{\beta=0}^{\infty} a_{\beta} \Psi_{\beta}(\xi) \cong \sum_{\beta=0}^{P-1} a_{\beta} \Psi_{\beta}(\xi), \quad (1)$$

where P is the number of terms retained in the truncation scheme, $\xi = \{\xi_i\}_{i=1, \dots, M}$ is a vector of M independent standard random variables that represent the M random variables, a_{β} are unknown coefficients to be computed, and Ψ_{β} are multivariate Hermite polynomials which are orthogonal with respect to the joint PDF of the standard normal random vector ξ . These multivariate Hermite polynomials can be obtained from the product of one-dimensional Hermite polynomials as follows:

$$\Psi_{\beta} = \prod_{i=1}^M H_{\alpha_i}(\xi_i), \quad (2)$$

where α_i ($i = 1, \dots, M$) are a sequence of M non-negative integers, and $H_{\alpha_i}(\cdot)$ is the α_i^{th} one-dimensional

Hermite polynomial. The expressions of the one-dimensional Hermite polynomials are given in [14] among others.

In fact, for a PCE of order p , the infinite series in Eq. 1 should be truncated by retaining only the multivariate polynomials Ψ_β of degree less than or equal to p . For this purpose, the classical truncation scheme based on the first order norm is usually used. The first order norm (which is defined as $\|\alpha\|_1 = \sum_{i=1}^M \alpha_i$) should be less than or equal to the order p of the PCE. This leads to a number P of the unknown PCE coefficients equal to $\frac{(M+p)!}{M!p!}$. This number is significant in the present case of random fields (especially when considering small values of the autocorrelation distances) and thus, one needs a great number of calls of the deterministic model. The SPCE methodology presented by [1] is an efficient alternative that can significantly reduce the number of calls of the deterministic model. In this methodology, [1] have shown that the number of significant terms in a PCE is relatively small because the multivariate polynomials Ψ_β corresponding to high-order interaction (i.e., those resulting from the multiplication of the H_{α_i} , with increasing α_i values) are associated with very small values for the coefficients a_β . Thus, a truncation strategy (called the hyperbolic truncation scheme) based on this observation was suggested by these authors. Within this strategy, the multivariate polynomials Ψ_β corresponding to high-order interaction were penalized. This was performed by considering the hyperbolic truncation scheme which suggests that the q -norm¹ should be less than or equal to the order p of the PCE. The q -norm is given by:

$$\|\alpha\|_q = \left(\sum_{i=1}^M (\alpha_i)^q \right)^{1/q} \quad (3)$$

where q is a coefficient ($0 < q < 1$). In this formula, q can be chosen arbitrarily. [1] have shown that sufficient accuracy is obtained for $q \geq 0.5$. The hyperbolic truncation scheme will be employed in this paper to build up a SPCE of the system response using a value of q equal to 0.7. Notice, finally, that the iterative procedure suggested by [1] for building up the SPCE is also employed herein. For more details on this procedure, the reader may refer to [1] (see also the flowchart presented in [10] and [11]).

As a conclusion, the proposed SPCE methodology leads to a SPCE that contains a small number of unknown coefficients which can be calculated from a reduced number of calls of the deterministic model. The next subsection is devoted to the method used for the computation of the coefficients a_β of the SPCE using the regression approach.

2.1. Computation of the sparse polynomial chaos expansion coefficients by the regression approach

Consider a set of K realizations $\{\zeta^{(1)} = (\zeta_1, \dots, \zeta_M), \dots, \zeta^{(K)} = (\zeta_1, \dots, \zeta_M)\}$ of the standard normal random vector ζ . These realizations are called experimental design (ED) and can be obtained from MC simulations or any other sampling scheme (e.g., Latin Hypercube sampling or Sobol set). In this paper, the MC simulations were used. We note $\Gamma = \{\Gamma(\zeta^{(1)}), \dots, \Gamma(\zeta^{(K)})\}$ the corresponding values of the response determined by deterministic calculations. The computation of the SPCE coefficients using the regression approach is performed using the following equation:

$$\widehat{a} = (\eta^T \eta)^{-1} \eta^T \Gamma \quad (4)$$

where the matrix η is defined by:

$$\eta_{i\beta} = \Psi_\beta(\zeta^{(i)}), \quad i = 1, \dots, K, \quad \beta = 0, \dots, J - 1 \quad (5)$$

¹Equation 3 actually defines a quasi-norm rather than a norm. However, this has no incidence on the derivations of the SPCE algorithm.

In equation 5, J is the number of the retained SPCE coefficients. In order to ensure the numerical stability of the treated problem in Eq. 4, the size K of the ED must be selected in such a way that the matrix $(\eta^T \eta)^{-1}$ is well conditioned. This implies that the rank of this matrix should be larger than or equal to the number of unknown coefficients. This test was systematically performed while solving the linear system of equations of the regression approach. It should be noted that the quality of the output approximation *via* a SPCE closely depends on the SPCE order p . To ensure a good fit between the meta-model and the true deterministic model (i.e., to obtain the optimal SPCE order), one successively increases the SPCE order until a target accuracy was obtained. The coefficients of determination R^2 and Q^2 used in this paper are presented in Appendix A.

Once the coefficients a_β have been computed, the PDF of the system response and the corresponding statistical moments (mean, standard deviation, skewness, and kurtosis) can be calculated with no additional cost. This is performed by generating a large number of realizations of the standard normal random vector ξ_i ($i=1, \dots, M$) and by computing the corresponding system responses using the meta-model.

3. GLOBAL SENSITIVITY ANALYSIS

Once the SPCE coefficients are determined, a GSA based on Sobol indices can be easily performed. Notice that the first-order Sobol index of a given random variable ξ_i ($i=1, \dots, M$) gives the contribution of this variable in the variability of the system response. The first-order Sobol index is given by [15, 16] as follows:

$$S(\xi_i) = \frac{\text{Var}[E(\Gamma|\xi_i)]}{\text{Var}(\Gamma)} \quad (6)$$

where Γ is the system response, $E(\Gamma|\xi_i)$ is the expectation of Γ conditional on a fixed value of ξ_i , and Var denotes the variance. In the present paper, the system response is represented by a SPCE. Thus, by replacing Γ in Eq. 6 with the SPCE expression, one obtains the Sobol index as a function of the different terms of the SPCE (cf. [17]) as follows:

$$S(\xi_i) = \frac{\sum_{\beta \in I_i} (a_\beta)^2 E[(\Psi_\beta)^2]}{\sum_{\beta=0}^{p-1} (a_\beta)^2 E[(\Psi_\beta)^2]} \quad (7)$$

where a_β are the obtained SPCE coefficients, Ψ_β are the multivariate Hermite polynomials, $E[.]$ is the expectation operator, and $E[(\Psi_\beta)^2]$ is given by [17] as follows:

$$E[(\Psi_\beta)^2] = \prod_{i=1}^M a_i! \quad (8)$$

where the a_i are the same sequence of M non-negative integers $\{a_1, \dots, a_M\}$ used in Eq.(2). Notice, finally, that I_i that appears in the numerator of Eq. (7) denotes the set of indices β for which the corresponding Ψ_β terms are only functions of the random variable ξ_i (i.e., they only contain the variable ξ_i).

In order to illustrate the construction of a PCE and the derivation of the equations providing Sobol indices, an illustrative example of a PCE of order $p=3$ using only $M=2$ random variables (ξ_1 and ξ_2) is presented in Appendix B.

4. EFFICIENT COMBINATION BETWEEN THE SPARSE POLYNOMIAL CHAOS EXPANSION METHODOLOGY AND THE GLOBAL SENSITIVITY ANALYSIS

As mentioned previously, the time cost of the probabilistic analysis remains important even with the use of the SPCE when dealing with spatially varying soils and computationally-expensive deterministic models. Consequently, a procedure that can reduce once again this time cost is needed. An efficient combination between the SPCE methodology and the GSA is proposed in this section. This method is called hereafter SPCE/GSA procedure. In this method, a small SPCE order is firstly selected to approximate the system response by a meta-model. It should be noted that the random variables involved in the system response are those that result from the discretization of the random fields into a finite number of random variables. A GSA based on Sobol indices is then performed on this small SPCE order to determine the weight of each random variable in the variability of the system response. The variables with very small values of their Sobol indices (i.e., those that have a small weight in the variability of the system response) can be discarded. Consequently, a response which only depends on a smaller number of random variables is obtained. In other words, one obtains a response with an 'effective dimension'. This dimension is smaller than the initial dimension where the total number of random variables was considered. As it will be shown later, the use of a small SPCE order to perform the GSA is not a concern because higher SPCE orders lead to the same influential random variables. Once the 'effective dimension' was determined, a higher SPCE order that makes use of only the most influential random variables can be used. This significantly reduces the computation time. The use of a higher SPCE order is necessary in order to lead to an improved fit of the SPCE. The SPCE/GSA procedure can be described in more details by the following steps:

- a. Discretize the random field(s): This step was made in this paper using the expansion optimal linear estimation method (EOLE) by [18] and its extensions by [19] to cover the case of cross-correlated NG random fields. Let us consider N_{RF} anisotropic NG cross-correlated random fields $Z_i^{NG}(x, y)$, where $i = 1, \dots, N_{RF}$. These random fields are described by (i) constant means and standard deviations μ_i and σ_i ; (ii) NG marginal CDFs named G_i ; (iii) a target cross-correlation matrix C^{NG} ; and (iv) a common square exponential autocorrelation function $\rho_z^{NG}[(x, y), (x', y')]$ that gives the values of the correlation between the two arbitrary points (x, y) and (x', y') . This autocorrelation function is given as follows:

$$\rho_z^{NG}[(x, y), (x', y')] = \exp\left(-\left(\frac{x-x'}{a_x}\right)^2 - \left(\frac{y-y'}{a_y}\right)^2\right) \quad (9)$$

where a_x and a_y are the autocorrelation distances along x and y , respectively. In order to discretize the abovementioned cross-correlated NG random fields, one should first define a stochastic grid composed of s grid points (or nodes). Then, one should determine the NG autocorrelation matrix, Σ^{NG} , that gives the correlation between each grid point of the stochastic mesh and the other grid points of this mesh by using Eq. (9). The common NG autocorrelation matrix Σ^{NG} and the target NG cross-correlation matrix C^{NG} should be transformed into the Gaussian space by using Nataf model [20], because the discretization of the random fields by using EOLE is carried out in the Gaussian space. As a result, one obtains N_{RF} Gaussian autocorrelation matrices Σ^i (where $i = 1, \dots, N_{RF}$) and a Gaussian cross-correlation matrix C that can be used to discretize the two random fields. The value of the random field i (where $i = 1, \dots, N_{RF}$) at an arbitrary point (x, y) as obtained using this method is given by the following equation [cf. 19]:

$$\tilde{Z}_i(x, y) \approx \mu_i + \sigma_i \sum_{j=1}^N \frac{\kappa_{i,j}^D}{\sqrt{\lambda_j^i}} \cdot (\phi_j^i)^T \cdot \Omega, \quad (10)$$

where $\kappa_{i,j}^D$ are N_{RF} cross-correlated blocks of independent standard normal random variables obtained using the Gaussian cross-correlation matrix C between the N_{RF} fields (cf. [10] for more details),

λ_j^i, ϕ_j^i (where $i = 1, \dots, N_{RF}$) are the eigenvalues and eigenvectors of the N_{RF} Gaussian autocorrelation matrices Σ^i , Ω is the correlation vector between the value of the field at an arbitrary point (x, y) and its values at the different grid points, and N is the number of terms (expansion order) retained in the expansion. This number N is obtained (i) by sorting the eigenvalues λ_j ($j = 1, \dots, s$) in a descending order and (ii) by choosing the number N of eigenmodes that leads to a variance of the error which is smaller than a prescribed tolerance (say $\varepsilon \approx 10\%$). Notice that the variance of the error for EOLE is given by [18] as follows:

$$\text{Var}[Z_i(x, y) - \tilde{Z}_i(x, y)] = \sigma_i^2 \left\{ 1 - \sum_{j=1}^N \frac{1}{\lambda_j^i} \left((\phi_j^i)^T \Omega \right)^2 \right\}, \quad (11)$$

where $Z_i(x, y)$ and $\tilde{Z}_i(x, y)$ are respectively the exact and the approximate values of the random field i at a given point (x, y) and $(\phi_j^i)^T$ is the transpose of the eigenvector ϕ_j^i (where $i = 1, \dots, N_{RF}$). Notice, finally, that $(\phi_j^i)^T$ and Ω in Eq. 10 are two vectors of dimensions $(1 \times s)$ and $(s \times 1)$, respectively. Once the Gaussian random fields are obtained, they should be transformed into the NG space (in case of NG random fields) by applying the following formula:

$$\tilde{Z}_i^{NG}(x, y) = G_i^{-1} \{ \Phi[\tilde{Z}_i(x, y)] \} \quad (12)$$

where $\Phi(\cdot)$ is the standard normal CDF. For more details about the EOLE method and its extensions to cover the case of cross-correlated NG random fields, the reader may refer to [19]. After the discretization procedure, a random field is represented by N independent standard normal random variables. For the N_{RF} random fields that have the same autocorrelation function, the total number of random variables is $N_T = N_{RF} \times N$, which can be relatively large especially for the small values of the autocorrelation distances. It should be noted that for each realization, the values of the N_{RF} random fields were determined at the centroid of each element of the deterministic mesh. Once the different elements of the mesh are filled with the values of the different fields, the system response for this specific realization can be determined.

b. Select a preliminary small order of the sparse polynomial chaos expansion (e.g., $p = 2$) to approximate the system response by a meta-model. The main reason for selecting a small order is the exploration of the most influential random variables (i.e., those that have a significant weight in the variability of the system response) using a small ED. It should be emphasized here that the small value of the SPCE order leads to a significant decrease in the size of the ED, that is, in the number of calls of the deterministic model.

c. Perform a GSA based on Sobol indices (using the obtained second-order SPCE) to determine the weight of each random variable (of the different random fields) in the variability of the system response. The variables with very small values of their Sobol indices have no significant weight in the variability of the system response and can thus be discarded. Consequently, a response that only depends on a smaller number of random variables is obtained. In other words, one obtains a response with an 'effective dimension' N_e that is smaller than the initial dimension where the total number N_T of random variables was considered. It should be mentioned here that the small SPCE order (i.e., $p = 2$) used to perform the GSA is sufficient to provide the weight of each random variable in the variability of the system response because higher SPCE orders lead to the same influential random variables as will be seen later in the numerical results.

d. Use the same ED that was employed in step (b) but this time by only keeping the most influential random variables. By reducing the number of random variables from N_T to N_e (where $N_e < N_T$), one has the possibility to use a higher SPCE order (i.e., $p > 2$). The use of a higher SPCE order is necessary to lead to an improved fit of the SPCE because the coefficient of determination Q^2 given in Eq. (A.3) increases when the SPCE order increases as it will be shown in the numerical results.

As a conclusion, the use of the SPCE/GSA procedure is expected to provide a good fit of the deterministic model with a reduced number of model evaluations as compared with the classical SPCE approach.

5. NUMERICAL RESULTS

The aim of this section is to show the efficiency of the SPCE/GSA approach for the probabilistic analysis of geotechnical problems considering spatially varying soils and computationally-expensive deterministic models. The deterministic model considered in this paper involves the computation of the ultimate bearing capacity of strip footings resting on a ponderable soil. Because the SPCE approach was unable to consider the case of a ponderable soil (because of the significant computational cost of a single-deterministic computation), only the case of a weightless soil was considered by [9, 10]. In this section, the SPCE/GSA procedure was applied to compute the PDF of the ultimate bearing capacity of a strip footing resting on a ponderable spatially varying (c and φ) soil. The soil cohesion and friction angle are modeled by two anisotropic cross-correlated NG random fields. The deterministic model is based on numerical simulations by using FLAC^{3D}. The inputs of the deterministic and probabilistic models, which are common for both cases of weightless and ponderable soils, are briefly presented in Table I. It should be noted that the size of a given element in the deterministic mesh depends on the autocorrelation distances of the soil properties. [21] have 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. Notice also that [18] have shown that the number of grid points in the stochastic mesh strongly depends on the autocorrelation distances. These authors have shown that a ratio of about $l_{RF}/a=1/5$ provides a sufficient accuracy in terms of the variance of the error where l_{RF} is the typical element size in the stochastic grid and a is the autocorrelation distance. The two conditions mentioned previously have been considered when constructing the deterministic and stochastic meshes.

Before the presentation of the probabilistic results in the case of a ponderable soil, it seems necessary to validate the present SPCE/GSA procedure by comparison of its results with those obtained by the use of the SPCE (in the case of a weightless soil). This is the aim of the next subsection.

5.1. Validation of the SPCE/GSA

The aim of this section is the validation of the present SPCE/GSA approach. For this purpose, a comparison between the results obtained in the case of a weightless soil using the classical SPCE method (cf. [10]) and those obtained with the proposed SPCE/GSA procedure is presented herein. The adopted soil domain considered in the analysis is 15-m wide by 6-m deep. The footing breadth is equal to 2 m. It should be mentioned here that when neglecting the soil weight γ , the computation time decreases from 10 to 5 min per simulation. Although this difference may not seem to be significant for a single simulation, it becomes dramatically important during the probabilistic analyses where a large number of simulations are needed for each probabilistic analysis.

The validation of the SPCE/GSA procedure is carried out for the illustrative case [$a_x=10$ m, $a_y=1$ m, $r(c, \varphi)=0$]. For this configuration, the discretization of the two random fields c and φ has led to a total number of random variables N_T equal to 24 (12 random variables for each random field). By using the classical SPCE approach with the total number of random variables N_T , a fourth-order SPCE was sufficient to reach the target value of 0.999 for the coefficients of determination R^2 and Q^2 . An ED involving 800 points was needed to solve the regression problem given in Eq. 4 (i.e., to obtain a well-conditioned regression problem for which the rank of the matrix $(\eta^T \eta)^{-1}$ is larger than or equal to the number of unknown coefficients). On the other hand, by using the present SPCE/GSA procedure, a GSA was performed to detect the most influential random variables. Different SPCE orders (i.e., orders 2, 3, and 4) were considered in order to check if the SPCE order has an impact on the determination of the most influential random variables.

Table I. Deterministic and probabilistic input parameters.

Parameters	Parameters type	Deterministic parameters values	Stochastic parameters			
			Type of PDF and statistical moments	Coefficient of correlation r and correlation matrix C^{NG}	Autocorrelation function for c , φ , and ψ	
Soil	Elastic-perfectly plastic (Mohr-Coulomb)	Young modulus $E=60MPa$	—	—	—	
	Poisson ratio ν	$\nu=0.3$	—	—	—	
	Cohesion c	Random field	—	—	—	
	Friction angle φ	Random field	—	—	—	
			$\mu_c = 20kPa$ $Cov_c = 25\%$ $\mu_\varphi = 30^\circ$ $Cov_\varphi = 10\%$	$C^{NG} = \begin{bmatrix} r(c, c) & r(c, \varphi) \\ r(\varphi, c) & r(\varphi, \varphi) \end{bmatrix}$	$\rho_Z^{NG} = \exp\left(-\left(\frac{x-x'}{a_x}\right)^2 - \left(\frac{y-y'}{a_y}\right)^2\right)$ a_x and a_y are the autocorrelation distances along x and y , respectively.	
	Dilation ψ	Random field	—	—	—	
Foundation	Elastic	Young modulus $E=25GPa$	—	—	—	
Interface	Elastic-perfectly plastic (Mohr-Coulomb)	Poisson ratio ν	deterministic	$\nu=0.4$	—	—
		Shear stiffness K_s	deterministic	$K_s=1GPa$	—	—
		Normal stiffness K_n	deterministic	$K_n=1GPa$	—	—
		Cohesion c_{int}	deterministic	$c_{int}=20kPa$	—	—
	Friction angle φ_{int}	deterministic	$\varphi_{int}=30^\circ$	—	—	
	Dilation ψ_{int}	deterministic	$\psi_{int} = \frac{2}{3}\varphi_{int}$	—	—	

Figure 1 depicts the magnitudes of Sobol indices for the 24 random variables, as given by the SPCEs of orders 2, 3, and 4. The first 12 random variables [i.e., ξ_i for $i = 1, \dots, 12$] correspond to the cohesion random field, and the last 12 random variables [i.e., ξ_i for $i = 13, \dots, 24$] are those corresponding to the friction angle random field. Figure 1 shows that whatever the SPCE order is, the two first random variables of both fields (i.e., $\xi_1, \xi_2, \xi_{13}, \xi_{14}$) are the most influential. For the two random fields, a very fast decay in the weight of the random variables is noticed with quasi negligible magnitudes beyond the first two random variables. In fact, the first two random variables of the two random fields, which correspond to the first two eigenmodes of both fields, involve 95% of the response variability as may be seen from Table II. This is logical because the system response (i.e., the ultimate bearing capacity) is a quantity that depends on the average distribution of the soil properties (c and φ) which is therefore quite insensitive to small-scale fluctuations of the spatially varying shear strength parameters c and φ . Notice that the first eigenmodes provide the average distribution of the shear strength parameters over the soil domain; however, the remaining eigenmodes give the small scale fluctuations around this average distribution. Figure 1 clearly shows that the Sobol indices of the different random variables do not significantly change with the SPCE order. A second-order SPCE is thus sufficient to identify the influential random variables (i.e., those that have a significant weight in the variability of the ultimate bearing capacity). The main advantage of a small SPCE order is that a small ED is sufficient to solve the regression problem. As shown in Table III, 150 calls of the deterministic model are needed to solve the regression problem for a second-order SPCE. This number attains 800 for a fourth-order SPCE. This significant increase is because the number of unknown coefficients significantly increases from 29 to 144 when one chooses a fourth-order SPCE instead of a second-order SPCE. It should be emphasized here that the numbers of coefficient that appear in Table III are those retained by the iterative SPCE procedure suggested by [1]. Notice, finally, that the numbers of coefficient of the full PCEs of order 2, 3, and 4 are respectively 325, 2925, and 20475. This clearly shows that the use of the PCE in the case of random fields would not be feasible.

To choose the number of random variables which will be retained hereafter within the SPCE/GSA procedure, the different random variables of the two random fields are firstly sorted in a descending order according to the values of their Sobol indices (cf. the first three columns in Table IV). A threshold of acceptance t_a is then fixed as a percentage of the most influential (weighted) random variable. In the present paper, the most influential random variable is ξ_1 and it has a Sobol index $S_1=0.5$. Different values of the threshold were tested (cf. first line in Table IV). The random variables having a Sobol index smaller than the prescribed threshold t_a are discarded (marked with the symbol $(-)$ in the table). In this paper, a threshold of 2% of the Sobol index of the most weighed random variable is considered as sufficient; the corresponding retained random variables provide 98% of the total variance of the system response as may be seen from the last line of the

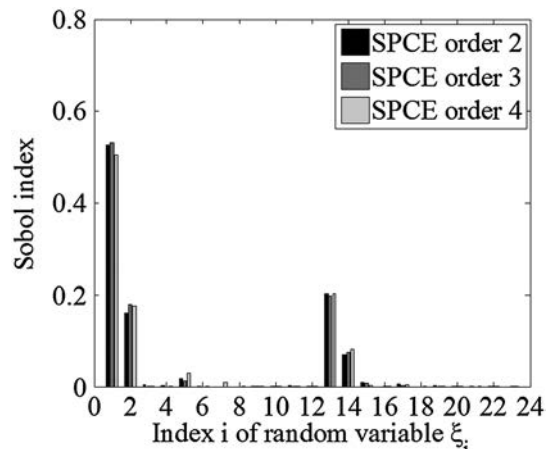


Figure 1. Sobol indices for sparse polynomial chaos expansions of orders 2, 3, and 4 using the classical sparse polynomial chaos expansion method with the total number of random variables ξ_i ($i = 1, \dots, 24$).

Table II. Sobol indices for the case where $a_x = 10$ m, $a_y = 1$ m, and $r(c, \varphi) = 0$.

		ζ_i ($i = 1, \dots, 12$) for the cohesion random field																						
Sobol index	ζ_1	0.500	ζ_2	0.170	ζ_3	0.002	ζ_4	0.002	ζ_5	0.030	ζ_6	0.002	ζ_7	0.009	ζ_8	2×10^{-4}	ζ_9	2×10^{-4}	ζ_{10}	9×10^{-5}	ζ_{11}	2×10^{-4}	ζ_{12}	7×10^{-5}
	ζ_{13}	0.200	ζ_{14}	0.080	ζ_{15}	0.001	ζ_{16}	8×10^{-4}	ζ_{17}	0.002	ζ_{18}	5×10^{-4}	ζ_{19}	6×10^{-4}	ζ_{20}	3×10^{-4}	ζ_{21}	1×10^{-4}	ζ_{22}	4×10^{-5}	ζ_{23}	4×10^{-5}	ζ_{24}	5×10^{-5}
		ζ_i ($i = 13, \dots, 24$) for the friction angle random field																						
Sobol index																								

Table III. Number of unknown coefficients and model evaluations for different sparse polynomial chaos expansion orders.

SPCE order	2	3	4
Number of unknown coefficients	29	35	144
Number of model evaluations	150	350	800

7th column in Table IV. For this threshold, an 'effective dimension' $N_e=5$ is obtained (i.e., five random variables are considered to be the most weighed). The five retained random variables ($\xi_1, \xi_{13}, \xi_2, \xi_{14}, \xi_5$) will now be used with the already existing 150 model evaluations which were firstly employed to approximate the second-order SPCE with the total number of random variables $N_T=24$.

The reduction in the number of random variables from $N_T=24$ to $N_e=5$ provides the possibility to use higher SPCE orders (i.e., $p > 2$) with the same ED (i.e., the 150 model evaluations). The use of a higher SPCE order is necessary to lead to an improved fit of the SPCE because the coefficients R^2 and Q^2 increase when the SPCE order increases as shown in Table V for both the classical SPCE approach (using the total number of random variables $N_T=24$ and the number of model evaluations of Table III) and the present SPCE/GSA procedure (where the effective dimension is equal to 5, i.e., $N_e=5$ and the number of model evaluations is fixed to 150). By using the SPCE/GSA procedure, an SPCE up to $p=8$ was reached using only 150 model evaluations. This order is to be compared with the fourth-order SPCE that was used in the SPCE methodology. From Table V, one can notice that with the use of the SPCE/GSA procedure, the Q^2 and R^2 coefficients increase with the increase of the SPCE order and stabilize beyond the order 5. This means that there is a need to increase the SPCE order to improve the fit; however, there is no improvement in the fit beyond the fifth order. On the other hand, the values of Q^2 and R^2 given by the present approach (i.e., 0.963 and 0.972) are smaller than those of the classical SPCE approach with a fourth order (i.e., 0.994 and 0.999). This is because 19 random variables were discarded which slightly affect the goodness of the fit.

Figure 2 shows the PDF of the ultimate bearing capacity as obtained by both the classical SPCE approach (with the total number of random variables $N_T=24$) and the proposed SPCE/GSA procedure (using only five random variables). Table VI provides the corresponding statistical moments and coefficients of determination R^2 and Q^2 . Notice that the results of the present SPCE/GSA approach are given in Table VI for different values of the number of model evaluations (from 150 to 800) and for a fifth-order SPCE. From this table, one can see that the coefficients R^2 and Q^2 of the SPCE/GSA procedure are quasi constant with the increase in the number of model evaluations. This means that 150 model evaluations are sufficient and there is no need for more model evaluations to improve the accuracy of the fit. On the other hand, one can observe (Table VI) that the first two statistical moments (μ and σ) are well estimated with the present SPCE/GSA approach using the 150 model evaluations. However, the third and fourth statistical moments (δ_u and κ_u) need more model evaluations (800 model evaluations) in order to converge to their reference values given by the SPCE approach. This demonstrates the efficiency of the present SPCE/GSA procedure to compute only the first two statistical moments with a much reduced number of the model evaluations (150 model evaluations) with respect to the classical SPCE approach (with 800 model evaluations).

As for the Sobol indices of the two random fields c and φ , Table VII shows that the SPCE/GSA procedure with only 150 model evaluations gives the same results obtained by the classical SPCE approach by using 800 model evaluations. This demonstrates once again the efficiency of the present SPCE/GSA procedure. It should be noted that the Sobol index of a random field was computed as the sum of the Sobol indices of the different variables that represent this field.

5.2. Probabilistic numerical results for a ponderable soil mass

The aim of this section is to present some probabilistic numerical results in the case of a ponderable soil mass (with $\gamma = 18kN/m^3$) using the SPCE/GSA procedure. The study involves the computation of the PDF of the ultimate bearing capacity of a shallow strip foundation resting on a ponderable 2D

Table IV. Sobol indices S_i ($i=1, \dots, 24$) of the different random variables ζ_i ($i=1, \dots, 24$) and the retained random variables for the different values of the threshold of acceptance t_a .

<i>Index i of random variable ζ_i ($i=1, \dots, 24$)</i>	<i>Random variable ζ_i ($i=1, \dots, 24$)</i>	<i>Sobol index S_i ($i=1, \dots, 24$)</i>	$t_a = 0.25\% \times S_i$ = 0.0013	$t_a = 1\% \times S_i$ = 0.005	$t_a = 1.5\% \times S_i$ = 0.0075	$t_a = 2\% \times S_i$ = 0.01	$t_a = 2.5\% \times S_i$ = 0.0125
1	ζ_1	$S_1 = 0.500$	0.500	0.500	0.500	0.500	0.500
13	ζ_{13}	$S_{13} = 0.200$	0.200	0.200	0.200	0.200	0.200
2	ζ_2	$S_2 = 0.170$	0.170	0.170	0.170	0.170	0.170
14	ζ_{14}	$S_{14} = 0.080$	0.080	0.080	0.080	0.080	0.080
5	ζ_5	$S_5 = 0.030$	0.030	0.030	0.030	0.030	0.030
7	ζ_7	$S_7 = 0.009$	0.009	0.009	0.009	—	—
6	ζ_6	$S_6 = 0.002$	0.002	—	—	—	—
17	ζ_{17}	$S_{17} = 0.002$	0.002	—	—	—	—
3	ζ_3	$S_3 = 0.002$	0.002	—	—	—	—
4	ζ_4	$S_4 = 0.002$	0.002	—	—	—	—
15	ζ_{15}	$S_{15} = 0.001$	0.001	—	—	—	—
16	ζ_{16}	$S_{16} = 8.0 \times 10^{-4}$	—	—	—	—	—
19	ζ_{19}	$S_{19} = 6.0 \times 10^{-4}$	—	—	—	—	—
18	ζ_{18}	$S_{18} = 5.0 \times 10^{-4}$	—	—	—	—	—
20	ζ_{20}	$S_{20} = 3.0 \times 10^{-4}$	—	—	—	—	—
8	ζ_8	$S_8 = 2.0 \times 10^{-4}$	—	—	—	—	—
9	ζ_9	$S_9 = 2.0 \times 10^{-4}$	—	—	—	—	—
11	ζ_{11}	$S_{11} = 2.0 \times 10^{-4}$	—	—	—	—	—
21	ζ_{21}	$S_{21} = 1.0 \times 10^{-4}$	—	—	—	—	—
10	ζ_{10}	$S_{10} = 9.0 \times 10^{-5}$	—	—	—	—	—
12	ζ_{12}	$S_{12} = 7.0 \times 10^{-5}$	—	—	—	—	—
24	ζ_{24}	$S_{24} = 5.0 \times 10^{-5}$	—	—	—	—	—
22	ζ_{22}	$S_{22} = 4.0 \times 10^{-5}$	—	—	—	—	—
23	ζ_{23}	$S_{23} = 4.0 \times 10^{-5}$	—	—	—	—	—
	<i>Sum of Sobol indices</i>	1.001	0.998	0.989	0.989	0.98	0.98

Table V. Coefficients of determination R^2 and Q^2 for different sparse polynomial chaos expansion orders when using the total and the reduced number of random variables.

SPCE order		2	3	4	5	6	7	8
Total number of random variables N_T	Coefficient of determination R^2	0.998	0.999	0.999	—	—	—	—
	Coefficient of determination Q^2	0.824	0.932	0.994	—	—	—	—
Reduced number of random variables N_e	Coefficient of determination R^2	0.961	0.963	0.968	0.970	0.972	0.972	0.972
	Coefficient of determination Q^2	0.791	0.883	0.957	0.961	0.963	0.963	0.963

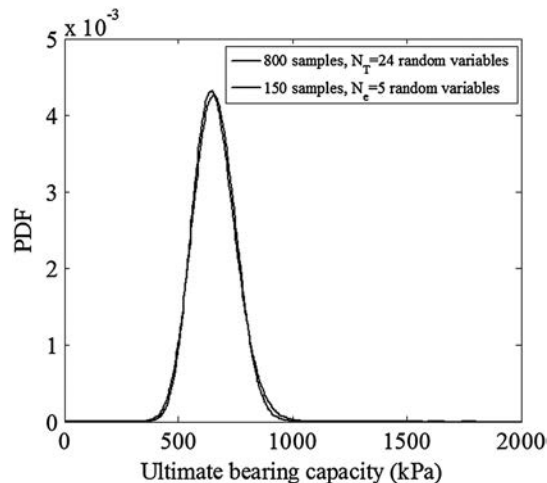


Figure 2. PDFs of the ultimate bearing capacity for both the classical sparse polynomial chaos expansion method with the total number of random variables ($N_T=24$) and the proposed SPCE/GSA procedure with only five random variables ($N_e=5$) when $a_x=10$ m, $a_y=1$ m and $r(c, \varphi)=0$.

spatially varying soil where the soil shear strength parameters (c and φ) are modeled as two anisotropic cross-correlated NG random fields. As shown in Figure 3, the adopted soil domain considered in the analysis is 13-m wide by 5-m deep. The footing breadth is equal to 1 m. It should be noticed here that the random fields of c and φ are discretized into a finite number of random variables. As was shown by [10], this number is small for the very large values of the autocorrelation distance and significantly increases for the small values of the autocorrelation distance. Table VIII provides the total number N_T of random variables needed to discretize the two random fields c and φ within a prescribed variance of the error of 10%. This table also presents the number N_e of the retained random variables as obtained using the SPCE/GSA procedure. One can observe an important reduction in the dimensionality of the treated problem with the use of the proposed SPCE/GSA procedure. For instance, the reduced number N_e of random variables is equal to 21 when the autocorrelation distances $a_x=a_y=0.25$ m. This number is to be compared with the total number $N_T=1760$ that shows once again that the ultimate bearing capacity is not sensitive to the very small fluctuations of the two random fields.

In the following subsections, the effect of the isotropic autocorrelation distance and the cross-correlation coefficient between the two random fields on the ultimate bearing capacity was undertaken using the input values presented in Table I.

5.2.1. Effect of the autocorrelation distances. Although the probabilistic model is able to consider both cases of isotropic and anisotropic random fields, only the isotropic case is considered in this

Table VI. Coefficients of determination R^2 and Q^2 of the sparse polynomial chaos expansion and statistical moments (μ , σ , δ_u , and κ_u) of the ultimate bearing capacity as given by the classical sparse polynomial chaos expansion approach (with $N_T=24$) and by the present SPCE/GSA procedure (with $N_e=5$).

	Number of model evaluations	Mean μ (kPa)	Standard deviation σ (kPa)	Skewness δ_u (-)	Kurtosis κ_u (-)	R^2	Q^2
With the total number of random variables $N_T=24$	800	658.2	93.57	0.287	0.163	0.999	0.995
With the reduced number of random variables $N_e=5$	150	657.84	90.80	0.105	0.013	0.970	0.961
	200	658.98	91.53	0.168	0.056	0.972	0.951
	250	659.90	92.10	0.188	0.063	0.964	0.956
	300	659.73	92.15	0.202	0.060	0.962	0.963
	400	660.05	90.95	0.291	0.050	0.969	0.960
	500	659.50	90.81	0.296	0.043	0.970	0.963
	600	659.75	90.99	0.272	0.116	0.968	0.963
	700	659.50	90.85	0.280	0.164	0.968	0.963
800	659.85	91.20	0.300	0.160	0.970	0.967	

section. This is because this case is the most critical one involving the greatest number of random variables.

Figure 4 shows the PDFs of the ultimate bearing capacity for different values of the isotropic autocorrelation distance $a_x=a_y$, together with the case of random variables when the cross-correlation coefficient is equal to -0.5 . Table IX presents the first two statistical moments of all these PDFs together with those corresponding to other values of the autocorrelation distance (50 m and 100 m). Figure 4 and Table IX show that the variability of the ultimate bearing capacity decreases when the autocorrelation distance $a_x=a_y$ decreases. The variability of the ultimate bearing capacity decreases with the increase in the soil heterogeneity because the zone involved by the possible failure mechanism will have (for the very small values of the autocorrelation distance) somewhat uniform values of the shear strength parameters over this zone because of the large number of high and small values of the shear strength parameters. This leads to close values of the ultimate bearing capacity from one simulation to another one and thus to a smaller variability in this bearing capacity. Figure 5

Figure 5 and Table IX show that the probabilistic mean value of the ultimate bearing capacity presents a minimum when the isotropic autocorrelation distance $a_x=a_y$ is nearly equal to the footing breadth B (i.e., in our case when $a_x=a_y=1$ m). Notice that the minimal probabilistic mean was also observed by [22–24] when isotropic random fields were studied. For a large value of the isotropic autocorrelation distance ($a_x=a_y=100$ m), the probabilistic mean tends to the one obtained in the case of random variables as may be seen from Table IX. On the other hand, for very small values of the autocorrelation distance $a_x=a_y$, the probabilistic mean becomes greater than the minimal value. This is because the weakest path becomes increasingly tortuous and its length is also longer and thus, the failure mechanism will pass through a shorter path and a stronger soil leading to greater bearing capacity values as mentioned by [9–11, 22–24]. Finally, it should be mentioned that the computation time using the SPCE/GSA procedure is relatively small for moderate to great values of the isotropic autocorrelation distance and becomes significant for the small values of the isotropic autocorrelation distance. For instance, in the case of very small values of the autocorrelation distance (i.e., $a_x=a_y=0.25$ m), the reduced number of random variables ($N_e=21$) has led to a computation time of about 3 days; however, for the large values of the autocorrelation distance (i.e., $a_x=a_y=100$ m), the reduced number of random variables ($N_e=2$) has led to a computation time of about 15 min.

5.2.2. *Effect of the cross-correlation coefficient of the two random fields.* Table X presents the two statistical moments (μ and σ) of the ultimate bearing capacity for negatively cross-correlated $r(c, \varphi)=-0.5$ and non-correlated $r(c, \varphi)=0$ random fields when $a_x=10$ m and $a_y=1$ m. Table X shows that the variability of the ultimate bearing capacity decreases when considering a negative correlation between the two random fields. This is because the

Table VII. Sobol indices as computed from the classical sparse polynomial chaos expansion approach and the present SPCE/GSA procedure.

Number of model evaluations		i	S_i ($i=1, \dots, 12$)	i	S_i ($i=13, \dots, 24$)	$S(c) = \sum_{i=1}^{12} S_i$	$S(\phi) = \sum_{i=13}^{24} S_i$
With the total number of random variables $N_T=24$	800	1	0.500	13	0.200	0.715	0.285
		2	0.170	14	0.080		
		3	0.002	15	0.001		
		4	0.002	16	8.0×10^{-4}		
		5	0.030	17	0.002		
		6	0.002	18	5.0×10^{-4}		
		7	0.009	19	6.0×10^{-4}		
		8	2.0×10^{-4}	20	3.0×10^{-4}		
		9	2.0×10^{-4}	21	1.0×10^{-4}		
		10	9.0×10^{-5}	22	4.0×10^{-5}		
		11	2.0×10^{-4}	23	4.0×10^{-5}		
		12	7.0×10^{-5}	24	5.0×10^{-5}		
With the reduced number of random variables $N_e=5$	150	i	S_i ($i=1, 2, 3$)	i	S_i ($i=4, 5$)	$S(c) = \sum_{i=1}^3 S_i$	$S(\phi) = \sum_{i=4}^5 S_i$
		1	0.510	4	0.076		
		2	0.200	5	0.190		
3	0.010					0.721	0.279

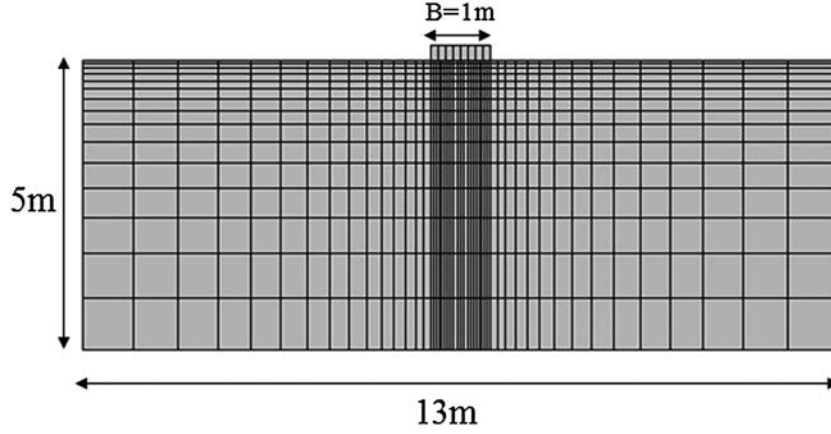


Figure 3. Adopted soil domain and the corresponding deterministic mesh.

Table VIII. Number of random variables used to discretize the two random fields c and ϕ in the case of isotropic random fields.

	N_T : Total number of random variables used to discretize the two random fields (c and ϕ)	N_e : Number of most influent random variables used to construct the SPCE when $t_a = 2\% \times S_I$
$a_x = a_y = 0.25m$	1760	21
$a_x = a_y = 0.5m$	460	21
$a_x = a_y = 1m$	120	20
$a_x = a_y = 1.5m$	70	20
$a_x = a_y = 2m$	50	20
$a_x = a_y = 3m$	24	12
$a_x = a_y = 5m$	20	8
$a_x = a_y = 10m$	10	6
$a_x = a_y = 50m$	10	4
$a_x = a_y = 100m$	10	2

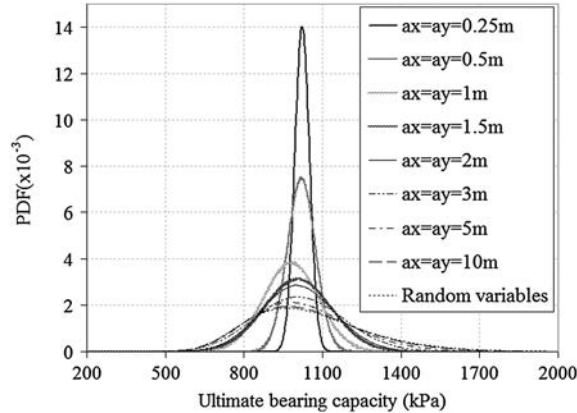


Figure 4. Influence of the isotropic autocorrelation distance $a_x = a_y$ on the PDF of the ultimate bearing capacity in the case where $r(c, \phi) = -0.5$.

increase of one parameter value implies a decrease in the other parameter. Thus, the total shear strength slightly varies. This leads to a reduced variation in the ultimate bearing capacity. It should be mentioned that the probabilistic mean value of the ultimate bearing capacity slightly increases when a negative correlation between the two random fields exists.

Table IX. Effect of the isotropic autocorrelation distance $a_x = a_y$ on the statistical moments (μ , and σ) of the ultimate bearing capacity.

$a_x = a_y (m)$	μ (kPa)	σ (kPa)	COV%
0.25	1022.3	28.5	2.8
0.5	1019.3	53.2	5.2
1	980.2	103.3	10.5
1.5	1001.4	127.0	12.6
2	1005.1	136.9	13.6
3	1012.7	169.2	16.7
5	1021.7	195.1	19.1
10	1040.0	216.9	20.9
50	1051.5	230.1	21.9
100	1052.0	230.9	21.9
Random variables	1052.2	230.9	21.9

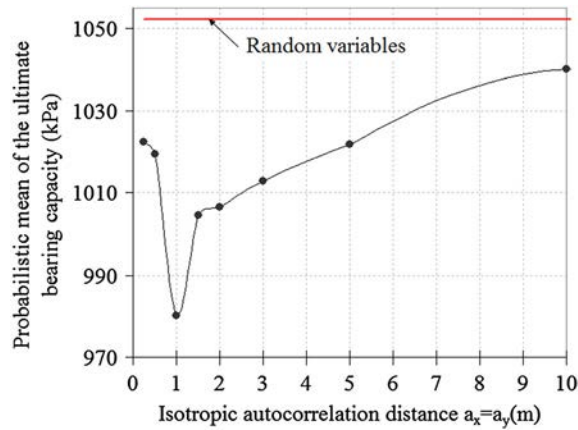


Figure 5. Influence of the isotropic autocorrelation distance $a_x = a_y$ on the probabilistic mean of the ultimate bearing capacity in the case where $r(c, \varphi) = -0.5$.

Table X. Effect of the coefficient of correlation $r(c, \varphi)$ on the statistical moments (μ and σ) of the ultimate bearing capacity when $a_x = 10$ m and $a_y = 1$ m.

$r(c, \varphi)$	μ (kPa)	σ (kPa)	COV%
-0.5	1022.7	172.1	16.8
0	1019.7	275.1	27.0

6. CONCLUSIONS

An efficient combined use of the SPCE methodology and the GSA has been proposed. The aim is to reduce the cost of the probabilistic analysis of computationally-expensive deterministic models. This methodology was firstly validated in this paper using a relatively non-expensive deterministic model. Then, it was applied to a computationally-expensive deterministic model. The validation consists in comparing the results of both the classical SPCE method that uses the total number of random variables and the proposed combination between the SPCE and the GSA that makes use of a reduced number of random variables. Satisfactory results were obtained using the proposed methodology with a smaller number of model evaluations. The first two statistical moments and the Sobol indices show good agreement between the two methods. On the other hand, the third and fourth statistical moments need more model evaluations in order to converge to their reference

values obtained using the classical SPCE approach. The application of the SPCE/GSA procedure to an expensive deterministic model that involves the computation of the PDF of the ultimate bearing capacity of a strip footing in the case of a ponderable spatially varying (c and ϕ) soil has shown the efficiency of the proposed methodology.

APPENDIX A

The coefficient of determination R^2 is given by:

$$R^2 = 1 - \frac{\frac{1}{K} \sum_{i=1}^K \left[\Gamma(\xi^{(i)}) - \Gamma_{SPCE}(\xi^{(i)}) \right]^2}{\frac{1}{K-1} \sum_{i=1}^K \left[\Gamma(\xi^{(i)}) - \bar{\Gamma} \right]^2} \quad (\text{A.1})$$

Where

$$\bar{\Gamma} = \frac{1}{K} \sum_{i=1}^K \Gamma(\xi^{(i)}) \quad (\text{A.2})$$

The value $R^2=1$ indicates a perfect fit of the true model response Γ , whereas $R^2=0$ indicates a nonlinear relationship between the true model response Γ and the SPCE model response Γ_{SPCE} . The coefficient R^2 may be a biased estimate because it does not take into account the robustness of the meta-model (i.e., its capability of correctly predicting the model response at any point which does not belong to the experimental design). As a consequence, one makes use of a more reliable and rigorous coefficient of determination [1]. In this case, one sequentially removes a point from the experiment design composed of K points. Let $\Gamma_{\xi \setminus i}$ be the meta-model that has been built from the experiment design after removing the i^{th} observation and let $\Delta^i = \Gamma(\xi^{(i)}) - \Gamma_{\xi \setminus i}(\xi^{(i)})$ be the predicted residual between the model evaluation at point $\xi^{(i)}$ and its prediction based on $\Gamma_{\xi \setminus i}$. The corresponding coefficient of determination is often denoted by Q^2 :

$$Q^2 = 1 - \frac{\frac{1}{K} \sum_{i=1}^K (\Delta^i)^2}{\frac{1}{K-1} \sum_{i=1}^K \left[\Gamma(\xi^{(i)}) - \bar{\Gamma} \right]^2} \quad (\text{A.3})$$

Both coefficients R^2 and Q^2 will be used in the present paper to check the accuracy of the fit, although only Q^2 is sufficient because it is more restrictive than R^2 .

APPENDIX B

Illustrative Example (for the construction of a PCE and for the computation of Sobol indices)

In order to illustrate the PCE theory in a simple manner, a PCE of order $p=3$ using only $M=2$ random variables (ξ_1 and ξ_2) will be considered in this illustrative example. By using the classical truncation scheme, Table B.1 presents the retained PCE terms that are those having a first-order norm $\|\alpha\|_1$ smaller than or equal to p (i.e., $p=3$). These terms are presented in Table B.1 in bold characters. As may be easily seen from Table B.1, the PCE basis contains $P=10$ terms whose expressions are computed using Eq.2.

Table B.1. Terms retained using the classical truncation scheme for $M=2$ and $p=3$.

α_1	0	1	0	1	2	0	2	1	2	3	0	3	1	3	3
α_2	0	0	1	1	0	2	1	2	2	0	3	1	3	2	3
$\ \alpha\ _1$	0	1	1	2	2	2	3	3	4	3	3	4	4	5	6

Table B.2 presents the expressions of the PCE basis Ψ_β . By using Table B.2, one can write the PCE expression as function of the input random variables (ξ_1 and ξ_2) as follows:

$$\begin{aligned} \Gamma_{PCE}(\xi) &= a_0\Psi_0 + a_1\Psi_1 + \dots + a_9\Psi_9 = \\ &a_0 + a_1\xi_1 + a_2\xi_2 + a_3\xi_1\xi_2 + a_4(\xi_1^2 - 1) + a_5(\xi_2^2 - 1) + a_6(\xi_1^2 - 1)\xi_2 + a_7\xi_1(\xi_2^2 - 1) \\ &\quad + a_8(\xi_1^3 - 3\xi_1) + a_9(\xi_2^3 - 3\xi_2) \end{aligned} \quad (B.1)$$

In this expression, the unknown coefficients can be computed using Eq. (4) by simulating an ED that contains K initial realizations of the two random variables (ξ_1 and ξ_2) and by computing the corresponding responses from deterministic calculations. It should be mentioned here that the size K of the ED should ensure the numerical stability of the regression problem and thus it can be enriched each time the matrix $(\eta^T\eta)^{-1}$ is badly-conditioned.

The first-order Sobol indices for the two random variables (ξ_1 and ξ_2) can be easily obtained once the coefficients a_0, \dots, a_9 are computed using Eq. (7). The only additional step is to compute $E[(\Psi_\beta)^2]$ corresponding to these two random variables. Table B.2 shows the values of $E[(\Psi_\beta)^2]$ computed using Eq. 8 for the different Ψ_β terms. The expression of the first-order Sobol indices of the two random variables ξ_1 and ξ_2 are written as follows:

$$\begin{aligned} S(\xi_1) &= \frac{a_1^2 + 2a_4^2 + 6a_8^2}{a_1^2 + a_2^2 + a_3^2 + 2a_4^2 + 2a_5^2 + 2a_6^2 + 2a_7^2 + 6a_8^2 + 6a_9^2} \\ S(\xi_2) &= \frac{a_2^2 + 2a_5^2 + 6a_9^2}{a_1^2 + a_2^2 + a_3^2 + 2a_4^2 + 2a_5^2 + 2a_6^2 + 2a_7^2 + 6a_8^2 + 6a_9^2} \end{aligned} \quad (B.2)$$

with

$$I_1 = (1, 4, 8); \quad I_2 = (2, 5, 9) \quad (B.3)$$

Table B.2. Basis of the polynomial chaos expansion with the classical truncation scheme for $M=2$ and $p=3$.

β	PCE order p	$\Psi_\beta = \prod_{i=1}^M H_{\alpha_i}(\xi_i)$	$E[(\Psi_\beta)^2] = \prod_{i=1}^M \alpha_i!$
0	$p=0$	$H_0(\xi_1) \times H_0(\xi_2) = 1$	$\alpha_1! \times \alpha_2! = 0! \times 0! = 1$
1	$p=1$	$H_1(\xi_1) \times H_0(\xi_2) = \xi_1$	$\alpha_1! \times \alpha_2! = 1! \times 0! = 1$
2		$H_0(\xi_1) \times H_1(\xi_2) = \xi_2$	$\alpha_1! \times \alpha_2! = 0! \times 1! = 1$
3	$p=2$	$H_1(\xi_1) \times H_1(\xi_2) = \xi_1 \xi_2$	$\alpha_1! \times \alpha_2! = 1! \times 1! = 1$
4		$H_2(\xi_1) \times H_0(\xi_2) = \xi_1^2 - 1$	$\alpha_1! \times \alpha_2! = 2! \times 0! = 2$
5		$H_0(\xi_1) \times H_2(\xi_2) = \xi_2^2 - 1$	$\alpha_1! \times \alpha_2! = 0! \times 2! = 2$
6	$p=3$	$H_2(\xi_1) \times H_1(\xi_2) = (\xi_1^2 - 1)\xi_2$	$\alpha_1! \times \alpha_2! = 2! \times 1! = 2$
7		$H_1(\xi_1) \times H_2(\xi_2) = \xi_1(\xi_2^2 - 1)$	$\alpha_1! \times \alpha_2! = 1! \times 2! = 2$
8		$H_3(\xi_1) \times H_0(\xi_2) = \xi_1^3 - 3\xi_1$	$\alpha_1! \times \alpha_2! = 3! \times 0! = 6$
9		$H_0(\xi_1) \times H_3(\xi_2) = \xi_2^3 - 3\xi_2$	$\alpha_1! \times \alpha_2! = 0! \times 3! = 6$

REFERENCES

- Blatman G, Sudret B. An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. *Probabilistic Engineering Mechanics* 2010; **25**:183–197.
- Isukapalli SS, Roy A, Georgopoulos PG. Stochastic response surface methods (SRSMs) for uncertainty propagation: Application to environmental and biological systems. *Risk Analysis* 1998; **18**:357–363.
- Huang SP, Liang B, Phoon KK. Geotechnical probabilistic analysis by collocation-based stochastic response surface method: An Excel Add-in implementation. *Georisk, Assessment and Management of Risk for Engineered Systems and Geohazards* 2009; **3**:75–86.
- Mollon G, Dias D, Soubra AH. Probabilistic analysis of pressurized tunnels against face stability using collocation-based stochastic response surface method. *Journal of Geotechnical and Geoenvironmental Engineering, ASCE* 2011; **137**:385–397.

5. Houmadi Y, Ahmed A, Soubra AH. Probabilistic analysis of a one-dimensional soil consolidation problem. *Georisk, Assessment and Management of Risk for Engineered Systems and Geohazards* 2011; **6**(1):36–49.
6. Mao N, Al-Bittar T, Soubra AH. Probabilistic analysis and design of strip foundations resting on rocks obeying Hoek-Brown failure criterion. *International Journal of Rock Mechanics and Mining Sciences* 2012; **49**:45–58.
7. Ahmed A, Soubra AH. Extension of subset simulation approach for uncertainty propagation and global sensitivity analysis. *Georisk, Assessment and Management of Risk for Engineered Systems and Geohazards* 2012. doi:10.1080/17499518.2012.656296.
8. Mollon G, Dias D, Soubra AH. Range of the safe retaining pressures of a pressurized tunnel face by a probabilistic approach. *Journal of Geotechnical and Geoenvironmental Engineering, ASCE* 2013; **139**(11):1954–1967.
9. Al-Bittar T, Soubra AH. Bearing capacity of strip footing on spatially random soils using sparse polynomial chaos expansion. *GeoRisk (GSP 224)*, ASCE, Atlanta, USA, 26–28 June 2011.
10. Al-Bittar T, Soubra AH. Bearing capacity of strip footing on spatially random soils using sparse polynomial chaos expansion. *International Journal for Numerical and Analytical Methods in Geomechanics* 2013; **37**(13):2039–2060.
11. Al-Bittar T. Probabilistic analysis of shallow foundations resting on spatially varying soils. *PhD Thesis*, University of Nantes, France, 2012, (cf. <http://www.univ-nantes.fr/soubra-ah>).
12. Spanos PD, Ghanem R. Stochastic finite element expansion for random media. *Journal of Engineering Mechanics* 1989; **115**:1035–1053.
13. Xiu D, Karniadakis GE. The Wiener-Askey polynomial chaos for stochastic differential equations. *Journal of Scientific Computing* 2002; **24**:619–44.
14. Ghanem R, Spanos P. *Stochastic Finite Elements: A Spectral Approach*. Courier Dover Publications, 2003; **224**:63–92.
15. Saltelli A, Chan K, Scott EM. *Sensitivity Analysis*. John Wiley and Sons, 2000, 155–174.
16. Sobol IM. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mechanics and Computation in Simulation* 2001; **55**:271–280.
17. Sudret B. Global sensitivity analysis using polynomial chaos expansion. *Reliability Engineering and System Safety* 2008; **93**:964–979.
18. Li CC, Der Kiureghian A. Optimal discretization of random fields. *Journal of Engineering Mechanics* 1993; **119**:1136–54.
19. Vořechovský M. Simulation of simply cross-correlated random fields by series expansion methods. *Structural Safety* 2008; **30**:337–363.
20. Nataf A. Détermination des distributions de probabilités dont les marges sont données. *Comptes Rendus de l'Académie des Sciences* 1962; **225**:42–3 (in French).
21. Der Kiureghian A, Ke JB. The stochastic finite element method in structural reliability. *Probabilistic Engineering Mechanics* 1988; **3**(2):83–91.
22. Griffiths DV, Fenton GA. Bearing capacity of spatially random soil: The undrained clay Prandtl problem revisited. *Geotechnique* 2001; **51**:351–359.
23. Griffiths DV, Fenton GA, Manoharan N. Bearing capacity of rough rigid strip footing on cohesive soil: Probabilistic study. *Journal of Geotechnical and Geoenvironmental Engineering, ASCE* 2002; **128**:743–755.
24. Fenton GA, Griffiths DV. Bearing capacity prediction of spatially random c-φ soils. *Canadian Geotechnical Journal* 2003; **40**:54–65.