Probabilistic stability analysis of an embankment dam considering soil spatial variability

Xiangfeng Guo, Daniel Dias, Qiujing Pan

Probabilistic stability analysis of an embankment dam considering soil spatial variability

Xiangfeng Guo[1*], Daniel Dias[2,1,3], Qiujing PAN[1]

*1 Univ. Grenoble Alpes, CNRS, Grenoble INP**, 3SR, F-38000 Grenoble, France*

2 Distinguished Professor, Hefei University of Technology, School of Automotive and Transportation Engineering, Hefei, China

3Geotechnical Expert, Antea Group, 92160 Antony, France

**Corresponding author (Email: xiangfeng.guo@3sr-grenoble.fr, Address: 1270 Rue de la Piscine, 38610 Gière, France)*

***Institute of Engineering Univ. Grenoble Alpes*

Abstract

A probabilistic stability analysis of an embankment dam taking into account soil spatial variabilities is presented in this paper. The effective cohesion and friction angle of the backfill are modelled as two anisotropic cross-correlated lognormal random fields, by using the Karhunen–Loève Expansion. The probabilistic analyses are performed by a meta-modelling technique SPCE/GSA which combines the Sparse Polynomial Chaos Expansion (SPCE) and the Global Sensitivity Analysis (GSA). This technique is targeted to deal with high dimensional stochastic problems and can provide a variety of interesting results in terms of probability density function (PDF) and statistical moments of the model response, failure probability and sensitivity index of each input variable. Concerning the deterministic calculations, two models (limit equilibrium and finite difference) are created and compared in a probabilistic framework. The influences of the soil spatial variability and the cross-correlation between the shear strength parameters on the reliability analysis results are investigated.

Keywords: Embankment dams, soil spatial variability, reliability analysis, sparse polynomial chaos expansion, global sensitivity analysis

1 Introduction

Embankment dams are mainly made of natural materials such as earth or rock-filled materials. Some of their parameters like the shear strength parameters are uncertain. In addition, soil properties of the backfill vary spatially due to the compaction and the layered construction procedure. For this reason, it is highly desirable to employ a probabilistic analysis method for embankment dam designs, which can account for the uncertainties mentioned above and the soil spatial variability.

Several studies concerning probabilistic stability analyses of embankment dams can be found in literature. Liang et al. [1] developed a probabilistic model for assessing the reliability index and the corresponding failure probability of multi-layered embankment dams; Babu and Srivastava [2] carried out a probabilistic stability analysis for four rehabilitated earth dams using the response surface method; Yi et al. [3] proposed a new program 3DSTAB combining slope stability analysis and reliability analysis for a rock-filled dam. Guo et al. [4] studied the reliability of an existing earth dam related to the sliding stability by using in-situ data. The authors employed the SPCE to estimate the failure probability of a dam. A common limitation of the four studies mentioned above is that they all neglected the soil spatial variability and used a random variable approach to account for the soil property uncertainties of embankment dams. In the random variable approach, soil properties are assumed to be homogeneous across the whole site. This assumption is not realistic since the soil properties vary spatially due to natural deposition and post-deposition processes [5]. The proposed study is dedicated to perform a probabilistic stability analysis of an embankment dam under pseudo-static loading conditions, by using a random field approach which permits to account for the soil spatial variability.

In the random field approach, the value of soil properties changes in both the vertical and horizontal directions in situ. As a result of the discretization of random fields, a large number of variables would be involved in order to obtain accurate results. These variables should be treated as input variables in a probabilistic analysis. A high dimensional input space is thus defined. Dealing with high dimensional stochastic problems is a challenge in the field of uncertainty quantification. Recently, several methods were proposed for such problems [6–9]. Among these methods, the Sparse Polynomial Chaos Expansion (SPCE) in combination with the Global Sensitivity Analysis (GSA) presents a high efficiency and can provide a variety of interesting results including failure probabilities (Pf), distributions and statistical moments of the system response and the sensitive index of each input variable. The method (noted as SPCE/GSA) is efficient since a dimension reduction technique is employed. It selects, at first stage, the significant input variables to form a reduced input space. The selection is based on the sensitivity index of each random variable obtained by a GSA [6]. Once the dimension is reduced, a high order SPCE can be constructed accurately with limited calls to the mechanical model. This method was applied to the reliability analyses of strip footings [7] and tunnel face stability [10], showing its good performance.

In summary, this paper aims at presenting a probabilistic stability analysis of embankment dams in spatially varying soils by using the method SPCE/GSA, and to investigate the effect of some key parameters. There are thus mainly two contributions.

1. The first is to show a comprehensive probabilistic stability analysis of embankment dams from soil random fields' simulation and deterministic models to the propagation of uncertainties and reliability results. The adopted SPCE/GSA method is a powerful reliability analysis tool, especially for high dimensional stochastic problems. It is able to create an accurate meta-model with limited calls to the deterministic model when a large number of random variables are considered. In addition, the method can give more interesting

reliability results than other reliability methods. For example, only Pf or reliability index is available for the methods FORM, Subset simulation and Importance sampling, while Pf together with system response distribution and sensibility index can all be found in the SPCE/GSA. Concerning the algorithm of the SPCE/GSA, an improvement regarding the identification of important input variables is proposed in the paper in order to enhance the method performance. It allows reducing the problem dimension and guarantees that the total variance of the input space can be covered with a desired threshold.

2. The second is to provide a view of the effect of soil spatial variability on the embankment dam reliability by conducting some parametric studies. The influences of the autocorrelation length, coefficient of variation and cross-correlation between the random fields on the reliability results are investigated. Such a study can help to better understand the impact of these parameters on the dam reliability, and to choose appropriate values in dam designs.

2 Presentation of the method SPCE/GSA

This section presents briefly the SPCE/GSA method. It is started by the SPCE presentation. Then, the principles of the GSA are described. In the end, the SPCE is combined with the GSA.

2.1 Sparse Polynomial Chaos Expansions (SPCE)

Basic principles of the PCE

The SPCE is an extension of the meta-modelling tool Polynomial Chaos Expansions (PCE) which approximates an original model by expanding the model response on a suitable basis. The basis is constituted of a series of multivariate polynomials that are orthogonal with respect to the joint probability density function of the random variables [7]. A PCE meta-model can be expressed as follows:

$$Y = \boldsymbol{D}(\boldsymbol{\xi}) \cong \sum_{\boldsymbol{\alpha} \in \mathbb{N}^M} k_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) \tag{1}$$

where $\boldsymbol{\xi}$ is a vector of independent random variables $\boldsymbol{\xi} = \{\xi_1, \xi_2, \dots, \xi_M\}$ with M representing the number of input variables, $\Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$ are multivariate polynomials, $k_{\boldsymbol{\alpha}}$ are unknown coefficients to be computed and $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_M\}$ is a multidimensional

index. The multivariate polynomial $\Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$ is the tensor product of univariate orthonormal polynomials. In this paper, standard normal random variables in conjunction with the Hermite polynomials are used.

For practical application, the Eq. (1) needs to be truncated to a finite number of terms. The truncation scheme, called the hyperbolic truncation scheme, proposed in [11] is adopted in this paper. It introduces a so-called q-quasi-norm. The formula is as follows:

$$\mathrm{A}^{M,p,q} = \left\{ \boldsymbol{\alpha} \in \mathbb{N}^M : |\boldsymbol{\alpha}|_q = \left(\sum_{i=1}^{M} a_i^q \right)^{\frac{1}{q}} < p \right\} \ (0 < q < 1) \qquad (2)$$

Once the truncated basis is determined, the coefficients $\{k_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in A^{M,p,q}}$ shall be computed by using the least-square regression method [12] with a design of experiment (DoE). A DoE includes a set of input variables and the corresponding system responses calculated by a deterministic model. In general, higher order of PCE requires bigger size of DoE, namely higher number of calls to a deterministic model. The accuracy of the obtained PCE can be estimated by the empirical mean-square residual error estimation R^2 and the leave-one-out error estimation Q^2. Concerning the details of the calculation of these two error estimates, readers are referred to [12].

The procedure of the SPCE

The SPCE is based on the principles of the PCE. It was proposed by Blatman and Sudret [11–13] for the purpose of further reducing the number of the involved multivariate polynomials. The idea came from the fact that the non-zero coefficients in the PCE form a sparse subset of the truncation set obtained by the standard truncation scheme or the hyperbolic truncation scheme [14]. Building an SPCE can be achieved by a stepwise regression technique [11] which can select the significant coefficients of the PCE. This technique is briefly presented as follows:

1. Generate randomly an initial experimental design χ and compute the corresponding model responses $\widehat{\boldsymbol{Y}}$ by using the deterministic model
2. Determine the user-defined parameters: the target accuracy Q_{tg}^2, the maximal PCE degree P_{max} and the cut-off values $\varepsilon_1, \varepsilon_2$.
3. For any PCE order $p \in \{1, \dots, p_{max}\}$:
   - Forward step: compute the increase in the determination coefficient R^2 by adding each candidate multivariate polynomial from a PCE basis set ($A^{M,p,q}$). Retain eventually those candidate terms that lead to a significant increase in R^2, i.e. greater than ε_1. Let $A^{p,+}$ be the final truncation set at this stage.
   - Backward step: compute the decrease in R^2 by removing each candidate term in $A^{p,+}$ of then degree not greater than p. Discard eventually from $A^{p,+}$ those terms that lead to an insignificant decrease in R^2, i.e. less than ε_2. Let A^p be the final truncation set.
   - If $Q_{A^p}^2 \geq Q_{tg}^2$, stop

In this paper, the following values are selected for the user-defined parameters in the SPCE according to the works in [4,10]: $Q_{tg}^2 = 0.99$, $P_{max} = 5$, $\varepsilon_1 = \varepsilon_2 = 5 \times 10^{-5}$ and the q-quasi-norm $q = 0.7$.

2.2 Global Sensitivity Analysis (GSA)

The GSA allows quantifying the effects of input random variables on the model response variance [6]. The first order Sobol index of a given variable ξ_i $(i = 1, \dots, M)$ can give the contribution of this variable to the system response variability. It is given as follows [15]:

$$S(\xi_i) = \frac{Var\left[E(\Gamma_{\xi_i})\right]}{Var(\Gamma)} \tag{3}$$

where Γ is system response, $E(\Gamma_{\xi_i})$ is expectation of Γ conditional on a fixed value of ξ_i, and Var represents the variance. Sudret [6] introduced an analytical way for computing the Sobol index as a post-processing of the PCE/SPCE coefficients. The first order Sobol index can be calculated by an available SPCE model as follows:

$$S(\xi_i) = \frac{\sum_{\alpha \in A_{\xi_i}} (k_j)^2 E[(\Psi_\alpha)^2]}{\sum_{\alpha \in A} (k_j)^2 E[(\Psi_\alpha)^2]} \quad (4)$$

where k_j are the SPCE coefficients, A is the truncation set ($A^{M,p,q}$), A_{ζ_i} is a subset of A in which the multivariate polynomials Ψ_α are only functions of the random variable ζ_i (i.e., they only contain the variable ζ_i) and $E[(\Psi_\alpha)^2]$ is the expectation of $(\Psi_\alpha)^2$.

2.3 The combination of the SPCE and the GSA

Using the SPCE for high dimensional stochastic problems is found to be very time-consuming [7,10] in the context of random fields since a number of variables are required for the discretization. In light of this, a combination of the SPCE with the GSA was proposed by [6,7]. The main idea is to firstly reduce the dimension of the input space by selecting significant variables. This can be achieved by a GSA based on a small order SPCE saying $p = 2$. Note that the PCE order has almost no influence on the Sobol indices according to [6,7]. Then, an accurate meta-model with a higher order SPCE is created w.r.t. the selected variables. As a conclusion, the SPCE/GSA procedure leads to a reduction in the dimension of the input space and consequently a limited number of calls to the deterministic model. The procedure of performing a reliability analysis by using the SPCE/GSA is given in Table 1.

Table 1 Procedure of an SPCE/GSA – based reliability analysis

Step	Description
1.	Construct a 2-order SPCE with the whole dimension of the input variables
2.	Compute the Sobol index for each input variable by Eq. (4); Select the important input variables according to the corresponding Sobol index
3.	Construct a high-order SPCE w.r.t. the selected variables
4.	Perform a Monte Carlo Simulation (MCS) on top of the high-order SPCE
5.	Post-processing on the MCS results to obtain failure probability and PDF of system response etc.

It can be seen from Table 1 that the method SPEC/GSA mainly consists of 5 steps. The two first steps aim at reducing the input dimension by selecting the significant

variables. Once the dimension is reduced, a higher-order SPCE can be created and an MCS can be performed then. Concerning the construction of SPCE meta-models, one can use *Uncertainty Quantification* toolboxes such as the UQlab [16] based on Matlab or the OpenTURNS [17] based on Python etc. There are a variety of reliability methods in the mentioned toolbox and one can easily obtain a meta-model (SPCE, kriging or Support vector machine etc.) with their own data or by linking their mechanical model with the toolbox via a wrapper [16,17].

Concerning the selection of the important input variables in step 2, it is proposed in this paper to use a threshold for the sum of the obtained Sobol indices. Firstly, the input variables are sorted with a descending order according to their Sobol index. Then, the first N_{GSA} variables would be selected so that the sum of the selected Sobol indices is greater than a target value, saying 0.98 in this paper. This guarantees that the reduced dimension covers at least 98% of the total input variance. The previous studies [7,10] about the SPCE/GSA, employ a threshold for individual input variables rather than for the sum of Sobol indices. For example, [7] used the 2% of the Sobol index (Max_S) of the most weighed random variable as an acceptance threshold to identify significant input variables. This approach may encounter problems in two cases: 1) the Max_S is relatively high. So the acceptance threshold is high. Few variables will be selected and the sum of the selected Sobol indices cannot be guaranteed; 2) a low acceptance threshold is induced by a relatively small Max_S. Few variables shall be rejected and the purpose of dimension reduction is not realized. By using the proposed sum threshold to identify the important input variables, the two problems mentioned above can be addressed.

The last step of the SPCE/GSA is to perform an MCS on top of the SPCE meta-model. MCS is the most straightforward and crude way of estimating failure probabilities. However, it could be very time-consuming if a complex deterministic

model is involved or the target failure probability is small. Coupling the MCS with meta-modelling techniques can bypass this issue since the computational effort of one realization by using the meta-model is quasi-negligible. For an MCS with N_{MCS} calls to the deterministic model, the failure probability Pf can be expressed as follows:

$$Pf = \frac{1}{N_{MCS}} * \sum_{i=1}^{N_{MCS}} I_{MCS} \tag{5}$$

where I_{MCS} is a failure indicator; I_{MCS} is set to 1 if the system fails and $I_{MCS} = 0$ otherwise. The number of N_{MCS} should be large enough in order to obtain an accurate estimate. The coefficient of variation of Pf for an MCS can be calculated by [18]:

$$CoV_{Pf} = \sqrt{(1 - Pf)/(N_{MCS} * Pf)} * 100\% \tag{6}$$

The MCS is dimension-independent, and often regarded as a standard reference to assess other reliability methods [8]. In this paper, the direct MCS is performed for the reference case to compare with the results obtained by the SPCE/GSA.

3 Generation of random fields by the K-L expansions

There are mainly three groups of methods for discretization of random fields: point discretization, average discretization and series expansions [19]. The K-L expansions adopted in the paper is part of the third group. The methods in this group allow representing exactly a Gaussian field by using a series of random variables and deterministic spatial functions [20]. The K-L expansions has been widely used in geotechnical engineering such as slopes [20], tunnels [10] and foundations [21] etc. since it is the most efficient in terms of the number of random variables required for a given accuracy and is independent to the division of meshes according to [19,20,22]. On the contrary, the number of random variables involved in the point discretization methods increases rapidly with the size of the problem [19].

Representation of a Gaussian random field

A random field can be defined as a collection of random variables indexed by a continuous parameter. A Gaussian random field can be completely described by its mean μ, variance σ^2 and an autocorrelation function $\rho(\boldsymbol{x}, \boldsymbol{x'})$ [23]. In this paper, the exponential autocorrelation function is used and reads as follows:

$$\rho(x, y) = \exp(-\frac{|x - x'|}{l_x} - \frac{|y - y'|}{l_y}) \tag{7}$$

where (x, y) and (x', y') are the coordinates of two arbitrary points in the random field space, l_x and l_y are respectively the horizontal and vertical autocorrelation length which is defined as the distance to which the autocorrelation function decays to $1/e$.

Basic principles of the K-L expansions

The K-L expansion is based on the spectral decomposition of the autocovariance function $C(\boldsymbol{x}, \boldsymbol{x'})$ which is the autocorrelation function $\rho(\boldsymbol{x}, \boldsymbol{x'})$ multiplied by the standard deviation $\sigma(x)$ and $\sigma(x')$. This function being symmetrical and positive definite, by definition, has all its eigenfunctions mutually orthogonal, and they form a complete orthogonal basis of Ω. Any realization of $H(\boldsymbol{x}, \theta)$ can thus be expanded over this basis as follows [19]:

$$H(\boldsymbol{x}, \theta) = \mu + \sigma \sum_{i=1}^{\infty} \sqrt{\lambda_i} \phi_i(\boldsymbol{x}) \xi_i(\theta) \cong \mu + \sigma \sum_{i=1}^{S} \sqrt{\lambda_i} \phi_i(\boldsymbol{x}) \xi_i(\theta) \tag{8}$$

where μ and σ are respectively mean value and standard deviation of the random field, $\{\xi_i(\theta), i = 1, \dots\}$ denotes the coordinates of the realization θ of the random field in the expanded space, λ_i and ϕ_i are the eigenvalues and eigenfunctions of the autocovariance function respectively and S is the size of the series expansion. The term $\xi_i(\theta)$ is, in fact, a set of uncorrelated standard normal variables as long as the realization of the random field is fixed. The value of S strongly depends on the

desired accuracy, the autocorrelation length and dimension of the random field. It can be determined by evaluating the error estimation of the truncated series expansion. The error estimate based on the variance of the truncated error for a KLE with S terms is given by [18]:

$$\varepsilon = \frac{1}{\Omega}\int_{\Omega}\left[1 - \sum_{i=1}^{S}\lambda_i\phi_i^2(\boldsymbol{x})\right]d\Omega \tag{9}$$

where Ω is the domain of the random field. In order to obtain a sufficient accuracy in terms of the variance error for random fields, Li and Der Kiureghian (1993) [24] recommended that the stochastic grid size can be set as 0.2 times the autocorrelation length.

Log-normal random fields and cross-correlation

In the case of log-normal random fields, the K-L expansion given in Eq. (8) becomes [25]

$$H(\boldsymbol{x},\theta) = exp\left[\mu_{ln} + \sigma_{ln}\sum_{i=1}^{S}\sqrt{\lambda_i}\phi_i(\boldsymbol{x})\xi_i(\theta)\right] = exp[\mu_{ln} + \sigma_{ln}G(\boldsymbol{x},\theta)] \tag{10}$$

Where μ_{ln} and σ_{ln} are respectively mean value and standard deviation of the log-normal random field, and $G(\boldsymbol{x},\theta)$ is a standard normally distributed random field with S terms. Note that only a single soil property is explicitly modelled in Eq. (10). In practice, two or even more soil properties might be considered as uncertain parameters in geotechnical probabilistic analysis. In this case, the correlation between the soil properties, namely cross-correlation, should be taken into account. Take a cohesive frictional $(c-\phi)$ soil as an example, the two cross-correlated log-normal random fields (c and ϕ) with a cross-correlation $\rho_{c,\phi}^{ln}$ can be expressed:

$$H_c(\boldsymbol{x},\theta) = exp\left[\mu_{ln_c} + \sigma_{ln_c} G_c(\boldsymbol{x},\theta)\right] \quad (11)$$

$$H_\phi(\boldsymbol{x},\theta) = exp\left\{\mu_{ln_\phi} + \sigma_{ln_\phi}\left[G_c(\boldsymbol{x},\theta).\rho_{c,\phi}^{ln} + G_\phi(\boldsymbol{x},\theta).\sqrt{1-(\rho_{c,\phi}^{ln})^2}\right]\right\} \quad (12)$$

291

292 where $\rho_{c,\phi}^{ln}$ is the cross-correlation coefficient between $\ln(c)$ and $\ln(\phi)$. It can be
293 calculated from the cross-correlation coefficient between c and ϕ [26]. For a sake of
294 simplicity, the $\rho_{c,\phi}^{ln}$ is noted as $\beta_{c,\phi}$ for the following parts of the paper.

295 4 The deterministic models

296 Two deterministic models are established in this study for computing the dam factor
297 of safety (FoS) under steady state flow conditions. The first one, based on the strength
298 reduction method, is created within the finite difference program Flac2D [27] while
299 the second one is constructed in Matlab using the limit equilibrium theory.

300

301 This section firstly presents the studied academic embankment dam in terms of the
302 geometry and soil properties. It is followed by a description of the numerical model.
303 Then, the analytical model is presented.

304 4.1 Presentation of the studied dam

305 Figure 1 shows the layout of the studied embankment dam where a free surface
306 groundwater flow occurs. Flow takes place from the left side to the right side. The
307 crest of the dam has a width of 10 m. The full reservoir level is equal to 11.88 m. A
308 horizontal drain is installed at the toe of the downstream slope for the purpose of
309 lowering the phreatic surface [28]. The ratio between the drain length and the dam
310 base (AB in Figure 1) is equal to 0.3, corresponding to a horizontal drain length of
311 28m.

312

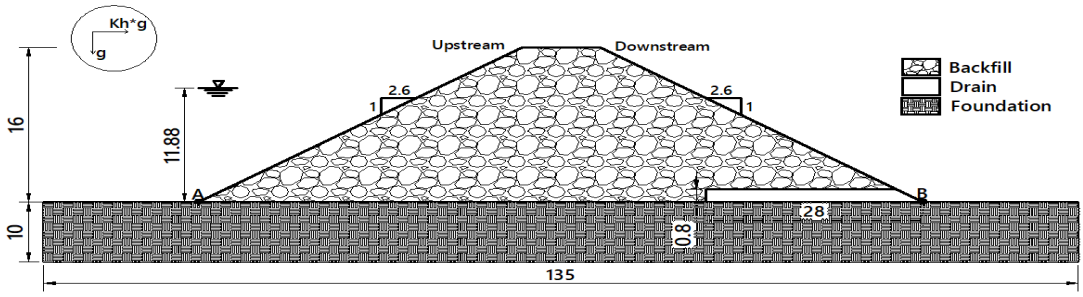


Figure 1. The geometry of the embankment dam case (g represents the gravitational acceleration)

The soil is assumed to follow a linear elastic perfectly plastic behaviour characterized by the Mohr Coulomb shear failure criterion. The deterministic values of the soil parameters adopted in this study are based on the data of an existing dam [29] and are presented in Table 2. In case where the drain is taken into account, just one parameter is modified in the backfill zone parameters. Its hydraulic conductivity is increased to 10^{-2} m/s.

Table 2. Deterministic soil parameters for the studied embankment dam

	Backfill	Foundation
Geotechnical parameters		
Effective cohesion (kPa)	8.9	100
Effective friction angle (^{o})	34.8	34.1
Unit weight (kN/m^3)	20	18
Elastic parameters		
Young Modulus (MPa)	100	600
Coefficient of Poisson (/)	0.3	0.25
Hydraulic parameters		
Hydraulic conductivity (m/s)	5×10^{-5}	1×10^{-6}
Porosity (/)	0.2	0.15
Horizontal seismic acceleration (g)	0.15	

4.2 The numerical model

A numerical model is created in the software FLAC2D. The boundary conditions used in this model are the following ones: the displacements are blocked following the horizontal and vertical axis on the base of the model; the horizontal displacements are blocked on the lateral edges of the model. Figure 2 presents the mesh used for the following calculations. The mesh includes around 1800 4-node quadrilateral plane

elements for the whole dam body. The selected number of elements used in this model is determined by a mesh refinement study. A further reduction in the element size compared to the one shown in Figure 2 has a negligible effect on the computed FoS value.

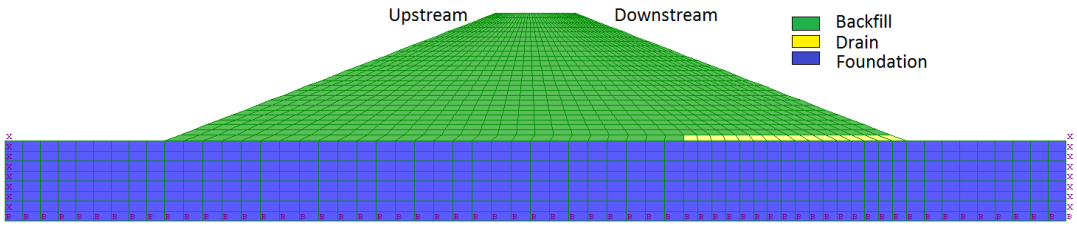


Figure 2 The numerical model mesh of the studied embankment dam

The stability analysis of the embankment dam considering soil spatial variability done by the numerical model is carried out using the following steps [27]:

1. **Initial equilibrium calculation**. Once the mesh is realized and the boundary conditions fixed, a mechanical calculation of the initial stress state is performed under the loading of the soil's own weight.
2. **Flow only calculation**. A hydrostatic head is applied in the upstream and the pore pressure distribution is obtained under steady state flow condition. The obtained pore pressure distribution will also be used for the analytical model.
3. **Assignment of the random fields to the mesh.** Two random fields (c *and* ϕ) are generated within the Matlab code using the Karhunen-Loève expansion method. They are then assigned to each element of the mesh according to the coordinates. Figure 3 gives an example of two negatively correlated random fields (c *and* ϕ) mapped to the mesh. Der Kiureghian and Ke (1988) suggested that the largest element of the deterministic mesh in a given direction (horizontal or vertical) should not exceed 0.5 times the autocorrelation length in that direction [30]. This condition is respected in the present study.
4. **Effective stress calculation with a horizontal acceleration**. A mechanical calculation is performed with the pore pressure distribution determined by step 2 under a pseudo-static loading condition with the new assignment of the soil properties. This step permits to calculate the effective stress state of the soil.
5. **Stability analysis**. The FoS is determined by using the strength reduction

method in this step.

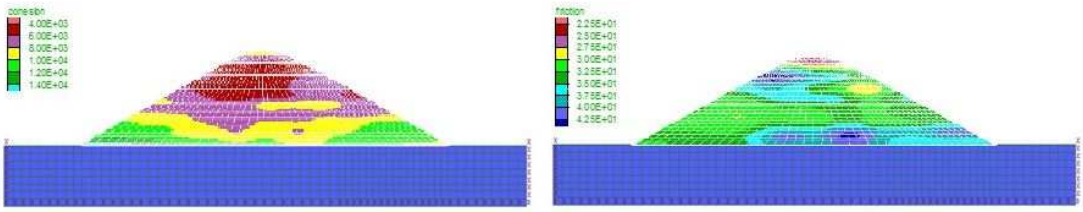


Figure 3. Example of two negatively correlated random fields ($\beta_{c,\phi}$=-0.5) mapped to the numerical model

4.3 The analytical model

An analytical model based on the limit equilibrium theory is created using the Matlab software. This model combines the Morgenstern Price method [31] and the pore water pressure distribution determined by the numerical model for estimating the FoS of a given slip surface. Particularly, the algorithm proposed by [32] is adopted in this paper to perform the Morgenstern Price method. The slip shape is assumed to be circular and the critical slip surface is located by mapping: the entry point and the exit point of the slip surface (A and B in Figure 4) and the deepest passage of the slip arc (D in Figure 4).

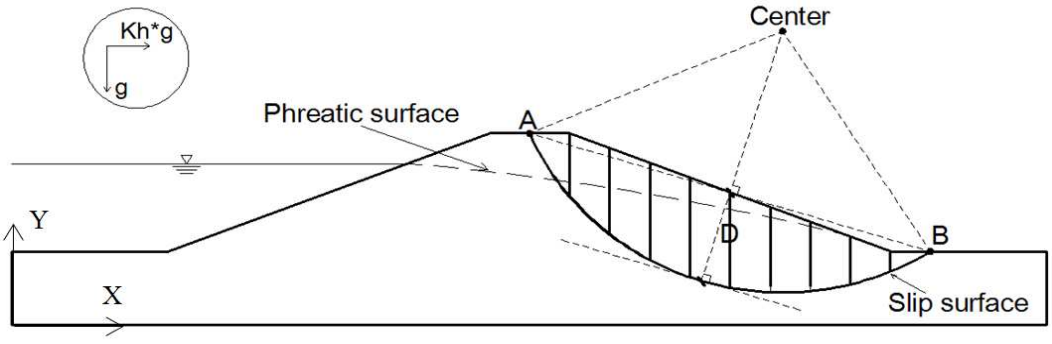


Figure 4. Slip surface and slices of sliding mass

For a first illustration of the analytical model, Figure 5 gives a comparison of the two deterministic models in terms of FoS values and the corresponding slip surface for three sets of random fields. The three cases of random fields represent three levels of

soil spatial variability with different autocorrelation lengths: corresponding to highly, intermediately and slightly non-homogeneous random fields. It can be observed that the FoS values obtained with the two models are always similar, and the analytical model leads to a slightly smaller FoS. The maximum difference is 1.9%. Concerning the critical slip surface, the analytical model can give a good approximation for all the three cases. The slip surfaces determined by the analytical model agree well with those of the numerical model, even for the case of high non-homogeneity.

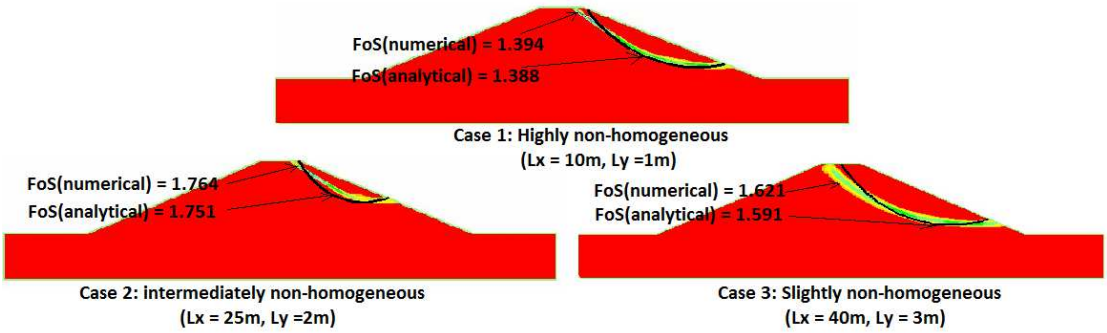


Figure 5 Comparison of the two deterministic models for three sets of random fields

This comparison justifies the slip shape assumption in the analytical model and shows a good accuracy for the computed FoS. However, in order to conduct a probabilistic stability analysis of embankment dams, the analytical model needs still a validation in a probabilistic framework, since the probabilistic analysis needs to call many times of the deterministic model. This is one of the aims of the next section.

Besides, it is important to note that the analytical model can dramatically reduce computational time. The consuming time for the numerical model to calculate the FoS of one realization on an Intel Xeon CPU E5-2609 v4 1.7GHz (2 processors) PC is about 5 minutes whereas it is reduced to 4 seconds with the analytical model. This shows the high efficiency of the analytical model compared to the numerical analysis. Reduction of computational time for deterministic models is very significant for a reliability analysis since it usually needs a large number of calls to deterministic calculations.

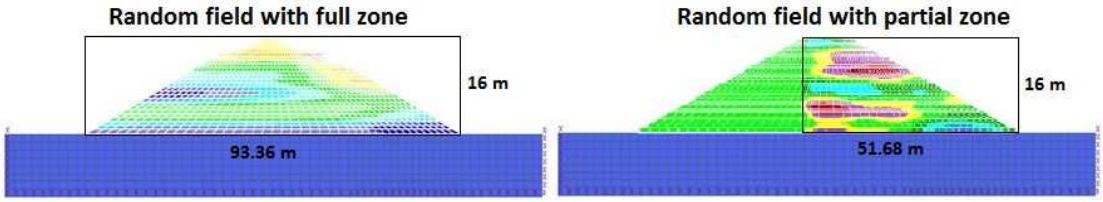
5 Numerical simulations and results

This section aims at presenting a preliminary reliability analysis of the considered dam by using the SPCE/GSA. The analytical model and the numerical model are both coupled with the reliability analysis and the obtained results are compared. Such a comparison allows validating the analytical model in a probabilistic framework. In addition, a direct MCS w.r.t. the analytical model is performed. This could assess the accuracy and efficiency of the method SPCE/GSA.

5.1 The conducted analyses

As the slip surfaces are generated mostly in the downstream part of the dam under steady state flow conditions, it seems that the soil variability in the upstream part of the dam has very limited effects on the FoS. The dimension of the random field can, therefore, be reduced in order to decrease the number of random variables involved in random field generations. Based on this assumption, a random field in a partial zone of the backfill is proposed (see in Figure 6). It is relatively smaller than the random filed in the full backfill. It allows reducing the number of random variables. The number of random variables involved in the K-L expansion is equal to the size of the series expansion S (see Eq. (9)), which depends on the desired accuracy, the autocorrelation length and dimension of the random field. It can be determined by evaluating the error estimation of the truncated series expansion using Eq. (9). For a variance error between 9% and 10%, $l_x = 40m$ and $l_y = 3m$, it needs at least 120 variables for a random field with a full zone of the backfill (left of Figure 6) while the number is reduced to 70 for a random field with a partial zone (right of Figure 6). This means that the random fields considering the full zone can be accurately expressed by a vector of 120 standard normal variables which will then be treated as input variables in the reliability analysis. The dimension of the input space is thus 241 (241 = 2×120 for the two random fields + 1 variable of the pseudo-static acceleration coefficient)

426 which represents a high dimensional stochastic problem. Using the reduced-size
427 random fields, the input dimension decreases from 241 to 141.
428



429
430 **Figure 6 Presentation of the random field with full zone and partial zone of the backfill.**

431 In order to validate the analytical model, verify the assumption about the size
432 reduction and to assess the method SPCE/GSA, four different probabilistic analyses
433 are performed (see Table 3). For the MCS, 2.1×10^5 samples are generated and
434 evaluated by the analytical model. Such a number of model evaluations (N_m) is
435 determined by an adaptive process with a target CoV_{Pf} smaller than 5%. Table 4 gives
436 the statistical models of the input parameters for the probabilistic analyses in the paper.
437 The effective cohesion and friction angle (c and ϕ) of the backfill are modelled by
438 log-normal random fields while the coefficient of pseudo-static acceleration (K_h) is
439 simulated by a log-normal random variable. In the context of the random field
440 approach, the value of c or ϕ is represented by a vector of random variables. The size
441 of the vector is respectively equal to 120 and 70 for a full and a partial zone. This
442 leads to a high dimensional stochastic problem. The parameters of the foundation
443 zone are considered as homogeneous and deterministic. It is noted that the
444 variabilities of the hydraulic parameters are neglected in the present study since they
445 are expected to have very limited effects on the reliability results. According to [33],
446 the spatial variability of soil permeability does not significantly add variability to the
447 free surface location for the dams with successfully designed filter drains (drain
448 permeability should at least 120 times the mean permeability of the dam itself and the
449 drain length should be larger than the 0.25 of the base dimension). The considered
450 drain in the paper satisfies the conditions of [33]. The phreatic surface inside the dam

is thus considered to be representative, even computed by deterministic values of hydraulic parameters. For the sake of simplicity, the spatial variabilities of the soil hydraulic parameters are neglected in the study as recommend in [33].

Table 3 The four probabilistic analyses

Probabilistic analysis	Random field dimension	Deterministic model	Reliability method
N. 1	Full zone	Numerical	SPCE/GSA
N. 2	Partial zone	Numerical	SPCE/GSA
N. 3	Partial zone	Analytical	SPCE/GSA
N. 4	Partial zone	Analytical	MCS

Table 4 Statistical model of the input parameters (backfill) for the probabilistic analyses

Variable	Mean	COV (%)	Modelling approach
Effective cohesion (kPa)	8.9	30	Log-normal random field
Effective friction angle (^{o})	34.8	15	Log-normal random field
Horizontal acceleration (g)	0.15	20	Log-normal random variable

COV: coefficient of variation

5.2 The obtained results

Figure 7 presents the PDFs of the FoS obtained by the four analyses with $l_x = 40m$, $l_y = 3m$ and $\beta_{c,\phi} = 0$. The PDF curve of MCS is directly plotted by estimating the probability density with a normal kernel function on the FoS values (2.1×10^5 samples). The other three curves are based on the FoS values of a MCS population of 10^6 samples, estimated by the SPCE meta-model. It can be seen from Figure 7 that the curves of the first two analyses are almost overlapping with each other. This observation indicates that neglecting the soil variability in the upstream part of the backfill has no influence on the PDF of FoS for the studied dam. The random fields with partial zone can then be used for the following probabilistic analyses in the paper, which can significantly reduce the number of random variable. Another remark is that the curves obtained with the analytical model are pretty close to those of the

numerical model and they have almost the same shape. This shows a good accuracy of the analytical model in the probabilistic analysis. It is also found that the curve N.3 is similar to the one obtained by a direct MCS (N.4). This means that the meta-model constructed by the method SPCE/GSA is able to capture the essential information of the whole input space. The meta-model can then accurately replace the original model. It is noted that only 1500 calls to the deterministic model are required in the analysis 3 (see Table 5). This represents a good efficiency of the method SPCE/GSA compared to the direct MCS which needs 2.1×10^5 model evaluations.

In addition, some other results of the four analyses are provided in Table 5. The failure probability is estimated between 0.0015 and 0.0021, and the mean value of the FoS is around 1.47. The method SPCE/GSA (N.3) gives similar results to the direct MCS (N.4). This finding is consistent with the observation made in Figure 7 and shows a good accuracy of the method SPCE/GSA. The results obtained with the partial zone (numerical or analytical) are all in good agreement with that of full zone using the numerical model. This validates once again the accuracy of the analytical model and the assumption about the dimension reduction of the random field. Concerning the analytical model, it is found to be slightly more conservative than the numerical model because the model estimates a smaller mean value of FoS and a bigger failure probability compared to those of the numerical model.

As a GSA is coupled with the probabilistic analysis, the sensitivity indices of each input variables are also available. Firstly, the first three analyses give similar Sobol index of the three input parameters according to Table 5, indicating validation of the analytical model with the partial zone in terms of the sensitivity study. Secondly, it can be observed that the friction angle takes the first place among the three parameters which has the most important influence on the FoS variability with a Sobol index around 0.58, the second most important factor is the acceleration coefficient with a

Sobol index around 0.35, and the cohesion is the least important one with a Sobol index around 0.06.

The following concluding remarks can be made based on the four analyses: 1) neglecting the soil variability in the upstream part of the backfill has almost no effect on the probabilistic analysis; 2) the method SPCE/GSA is efficient and has the capacity to provide accurate estimates of reliability results; 3) the analytical model is validated in a probabilistic framework considering soil spatial variability and random pseudo-static accelerations. According to the three remarks, it is recommended to couple the analytical model, more efficient than the numerical one as aforementioned, with the SPCE/GSA procedure in the partial zone domain for the following parametric studies.

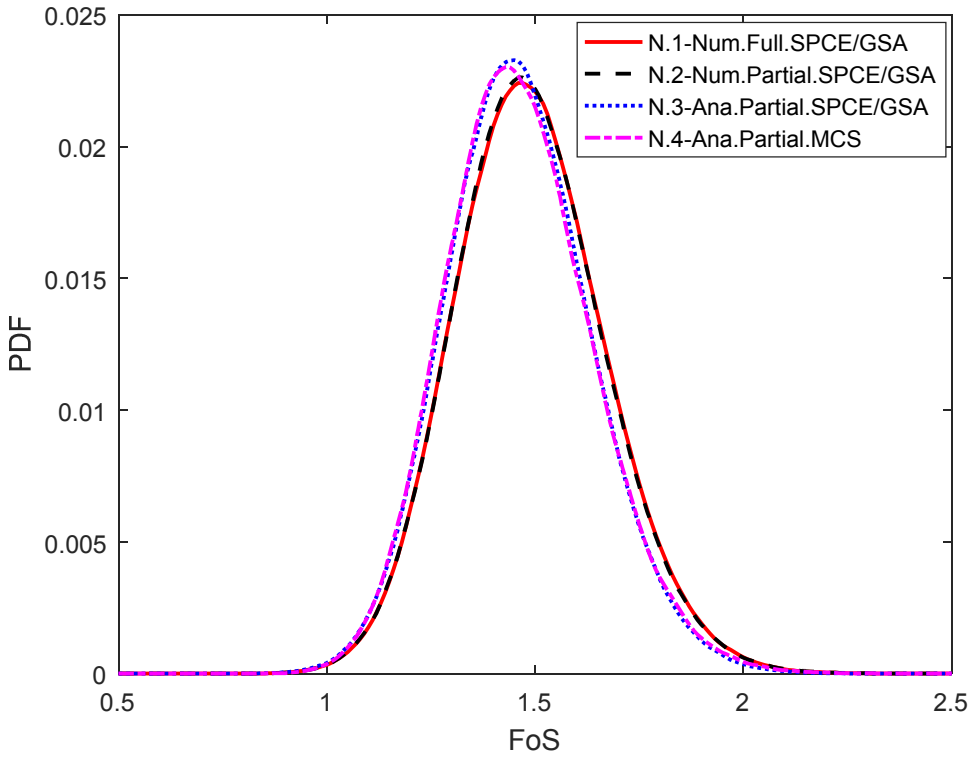


Figure 7 PDFs of the FoS obtained by the four probabilistic analyses ($l_x = 40, l_y = 3, \beta_{c,\phi} = 0$)

Table 5 Probabilistic results obtained by the four analyses ($l_x = 40, l_y = 3, \beta_{c,\phi} = 0$)

Number of analysis	P_f	$Mean_{FoS}$	Std_{Fos}	S_c	S_ϕ	S_{K_h}	N_m
N. 1	0.0015	1.488	0.180	0.063	0.561	0.369	3000
N. 2	0.0016	1.487	0.180	0.066	0.558	0.374	1500
N. 3	0.0021	1.462	0.174	0.045	0.607	0.347	1500
N. 4	0.0019	1.463	0.175	-	-	-	2.1×10^5

P_f: Probability of failure; $Mean_{FoS}$: Mean value of FoS; Std_{Fos}: Standard deviation of FoS; S_C: Sobol index of the cohesion; S_ϕ: Sobol index of the friction angle; S_{K_h}: Sobol index of the acceleration coefficient.

6 Parametric studies

This section takes advantage of the previous conclusion, using the analytical model for deterministic calculations and the partial zone for random field generation, to investigate the influences of the spatial variability ($l_x, l_y, COV_c \ and \ COV_\phi$) and the cross-correlation coefficient ($\beta_{c,\phi}$) on the PDF of the studied dam safety factor, on the sensitivity index of the shear strength parameters and on the failure probability. The cohesion and friction angle (c and ϕ) of the backfill are modelled by random fields based on the statistical moments presented in Table 4, while the coefficient of the pseudo-static acceleration is regarded as deterministic with a value of 0.15 in order to focus on the effects of soil variability.

6.1 Effect of the autocorrelation length

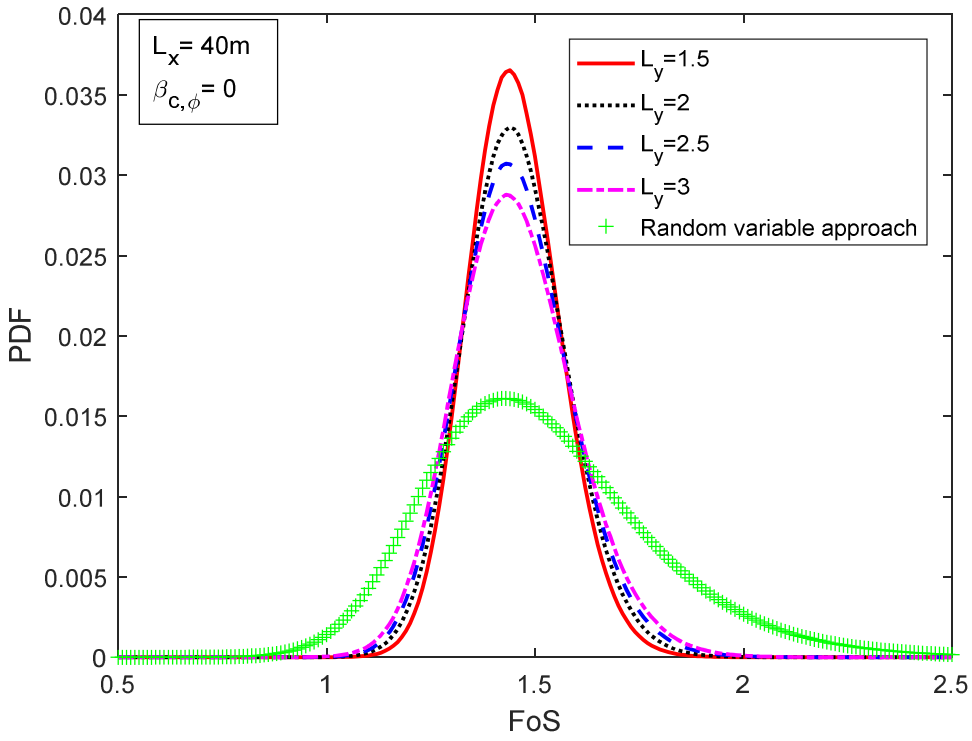
The autocorrelation length is defined as the distance to which the autocorrelation function decays to 1/e. It, together with the COV of the parameter, describes the soil spatial variability. A smaller autocorrelation length means that the soil is more severe non-homogeneous. According to a literature review given by El-Ramly et al (2003) [34], the autocorrelation length is within a range of 10-40m in the horizontal direction, while it ranges from 1 to 3m in the vertical direction. Several values of the autocorrelation lengths selected from its physical range are tested here to investigate its influence on the probabilistic analysis, with $\beta_{c,\phi} = 0$. An extreme case is that the

534 horizonal and vertical autocorrelation lengths are both equal to infinite, which
535 corresponds to the case of random variables.
536
537 Figure 8 and Figure 9 present the PDFs of the studied dam FoS corresponding to
538 different autocorrelation lengths together with the case of random variables. The
539 results are obtained for l_y varying between 1.5 to 3m with l_x being 40m in Figure 8,
540 and l_x varying between 25 to 40m with l_y being 3m in Figure 9. Table 6 gives an
541 overview of the probabilistic analysis results.
542
543 It can be observed from Figure 8 that taller and narrower PDF curves are produced
544 when decreasing the vertical autocorrelation length. This implies that the variability of
545 the studied dam safety factor decreases with the increase in the soil spatial
546 heterogeneity. In addition, the case of random variables gives the most spread-out
547 distribution of the FoS. From a physical point of view, a higher vertical
548 autocorrelation length indicates that the random variables (each discretized point) are
549 more strongly correlated. It therefore results in relatively low variation of the
550 simulated values of soil strength parameters along the depth and presents a greater
551 chance to form more uniform zones during one random field generation. However, the
552 global average of the strength parameter random fields changes a lot from one
553 realization to another. This subsequently leads to a higher variability of the system
554 response. On the contrary, a smaller vertical autocorrelation length can result in a
555 smaller uncertainty in the global average value from one realization to another and
556 thus a smaller variability of the system response. As the case of random variables
557 corresponds to the infinite value of the autocorrelation length which means a perfectly
558 correlated random field, it presents the biggest variability of the system response and
559 thus the most spread-out distribution of the FoS.
560

The same observations, concerning the effect of the autocorrelation length on the PDFs of the FoS, can also be remarked in Figure 9. However, it is worth to note, by comparing the two figures, that the variability of the studied dam safety factor is more sensitive to the change of the vertical autocorrelation length.

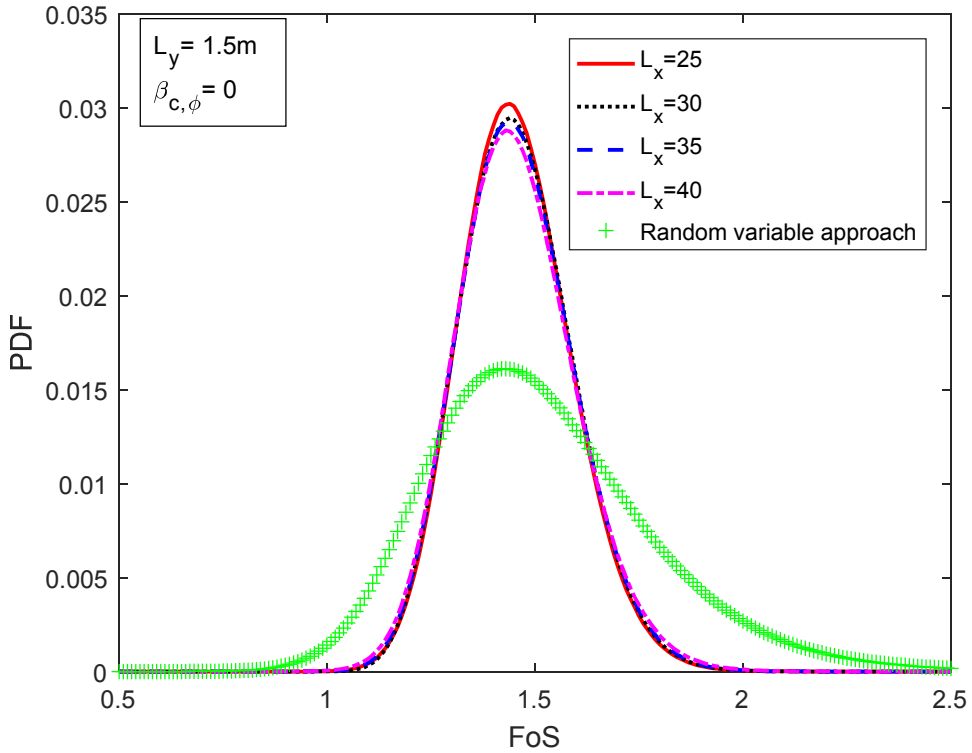
From Table 6, it is clearly seen that the failure probability decreases with the autocorrelation length, which would be expected. A small autocorrelation length induces significant fluctuation of the soil strength parameters along the slip surface. Therefore, the probability of calculations with a safety factor lower than 1.0 decreases since the simulated values are averaged to the mean value along the slip surface. The failure probability decreases from 20.5×10^{-5} to 2.36×10^{-5} when l_y varies from 3 to 1.5m with l_x being 40m. Particularly, the case of random variables leads to a much greater failure probability than all the cases of random fields. This implies that using the random variables approach in dam design, which neglects the soil spatial variability, is conservative. Concerning the statistical moments of the FoS, one can observe that the mean value remains almost constant while the standard deviation decreases with the autocorrelation lengths deceases. This characteristic results in a taller and narrower PDF curve when a smaller autocorrelation length is involved.

Table 6 presents the Sobol indices of $c\ and\ \phi$ for different autocorrelation lengths. It is observed that the autocorrelation length has almost no influence on the Sobol indices. The Sobol index of the friction angle (equal to 0.92) is bigger than that of the cohesion (equal to 0.08). This indicates that the variability of the FoS is more sensitive to the friction angle than to the cohesion.



585

586 **Figure 8 Influence of the vertical autocorrelation length on the PDF of the FoS**



587

588 **Figure 9 Influence of the horizontal autocorrelation length on the PDF of the FoS**

589 **Table 6 Comparison of the probabilistic results obtained with different autocorrelation lengths ($\beta_{c,\phi} = 0$)**

Autocorrelation length (m)	P_f	$Mean_{FoS}$	Std_{Fos}	S_C	S_φ
$l_x = 40, l_y = 1.5$	2.36×10^{-5}	1.451	0.111	0.07	0.93
$l_x = 40, l_y = 2$	4.07×10^{-5}	1.451	0.123	0.08	0.92
$l_x = 40, l_y = 2.5$	8.99×10^{-5}	1.454	0.133	0.07	0.93
$l_x = 40, l_y = 3$	20.5×10^{-5}	1.455	0.142	0.09	0.90
Random variables ($l_x = \infty, l_y = \infty$)	400×10^{-5}	1.517	0.270	0.07	0.93
$l_x = 40, l_y = 3$	20.5×10^{-5}	1.455	0.142	0.09	0.90
$l_x = 35, l_y = 3$	6.45×10^{-5}	1.456	0.137	0.08	0.92
$l_x = 30, l_y = 3$	3.75×10^{-5}	1.457	0.135	0.08	0.92
$l_x = 25, l_y = 3$	1.23×10^{-5}	1.457	0.133	0.08	0.92

590

591 **6.2 Effect of the cross-correlation**

592 It is commonly recognized that there is a negative correlation between c and ϕ, and
593 the value of correlation coefficient is between -0.7 and -0.24 according to [35,36].
594 Besides, Wolff (1985) [37] reported a positive value of 0.25 for the correlation
595 coefficient based on the consolidated drained soil test. In order to explore the effects
596 of cross-correlation between the two random fields (c and ϕ) on the dam reliability,
597 7 cases of correlation coefficients ranging from -0.6 to 0.6 are considered in this
598 section.

599

600 Figure 10 presents the PDFs of the studied dam safety factor corresponding to 4
601 correlation coefficients for a sake of clarity, together with the cases of random
602 variables. It shows that taller and narrower PDF curves are induced when decreasing
603 the cross-correlation coefficients. This implies that the variability of the studied dam
604 safety factor decreases with the cross-correlation between the two random fields
605 (c and ϕ). For the cases of negative correlation coefficients, the increase of one
606 parameter value decreases the other value. The variation of the total shear strength is

reduced and consequently the safety factor variation is also reduced. Similar results, concerning the effect of the correlation between c and ϕ on the PDF curves, can also be observed for the cases of random variables and their PDF curves are always spread compared to those of random fields approach.

Figure 11 shows the failure probability as a function of the cross-correlation coefficients. It is seen that the failure probability is greatly influenced by the cross-correlation coefficients. It changes several orders of magnitude (i.e., increasing from 8.00×10^{-6} to 1.09×10^{-3}) when the correlation coefficient varies from −0.6 to 0.6 for the cases of random fields. The failure probability decreases when the negative cross-correlation becomes stronger and increases with positive cross-correlation coefficient. This implies that the assumption of independence between c and ϕ may be severely biased if the actual cross-correlation is positive or negative. In fact, a negative correlation implies that low values of cohesion are associated with high values of friction angle and vice versa. In other words, a negative correlation between $c\ and\ \phi$ means that the uncertainty in the calculated shear strength is smaller compared to the independent and positively correlated cases. Besides, the overestimation of the failure probability by the random variables approach compared to the random fields approach is observed. This indicates once again that using the random variables approach in dam design is conservative.

Table 7 gives a summary of all the probabilistic analyses in terms of failure probability, FoS statistical moments and Sobol indices of c and ϕ for the cases of random fields. Particularly, a slight decrease of the mean value of FoS when increasing $\beta_{c,\phi}$ can be found in Table 7. On the contrary, the standard deviation presents an opposite trend to the mean value.

Concerning the Sobol indices, it can be seen that the change in $\beta_{c,\phi}$ has almost no effect on the sensitivity index of c and ϕ. The S_c remains around 0.08 and the S_ϕ around 0.92.

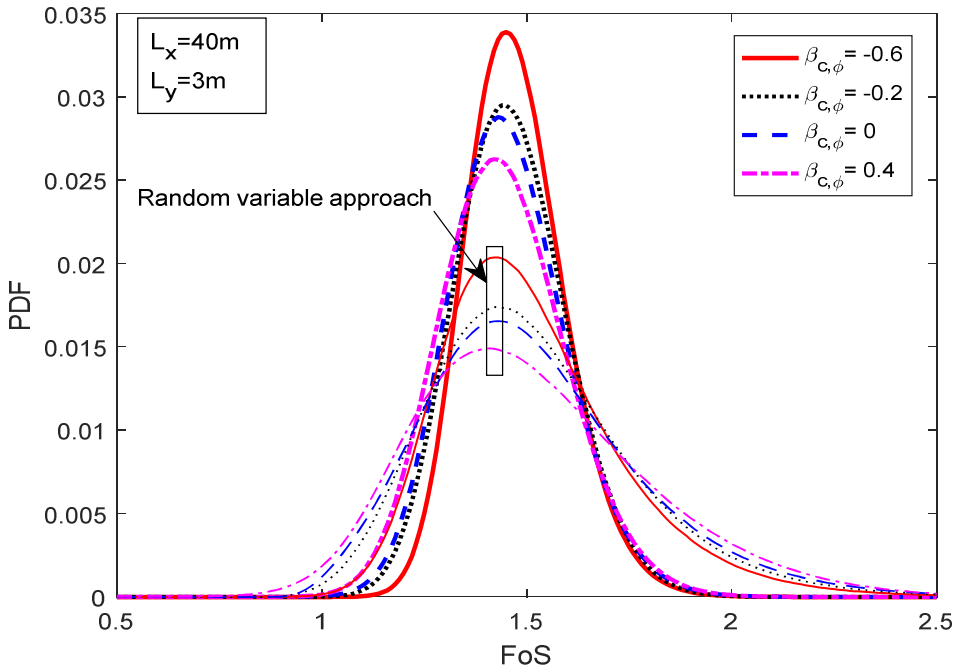


Figure 10 Effect of the cross-correlation coefficients between c and ϕ on the PDF of FoS

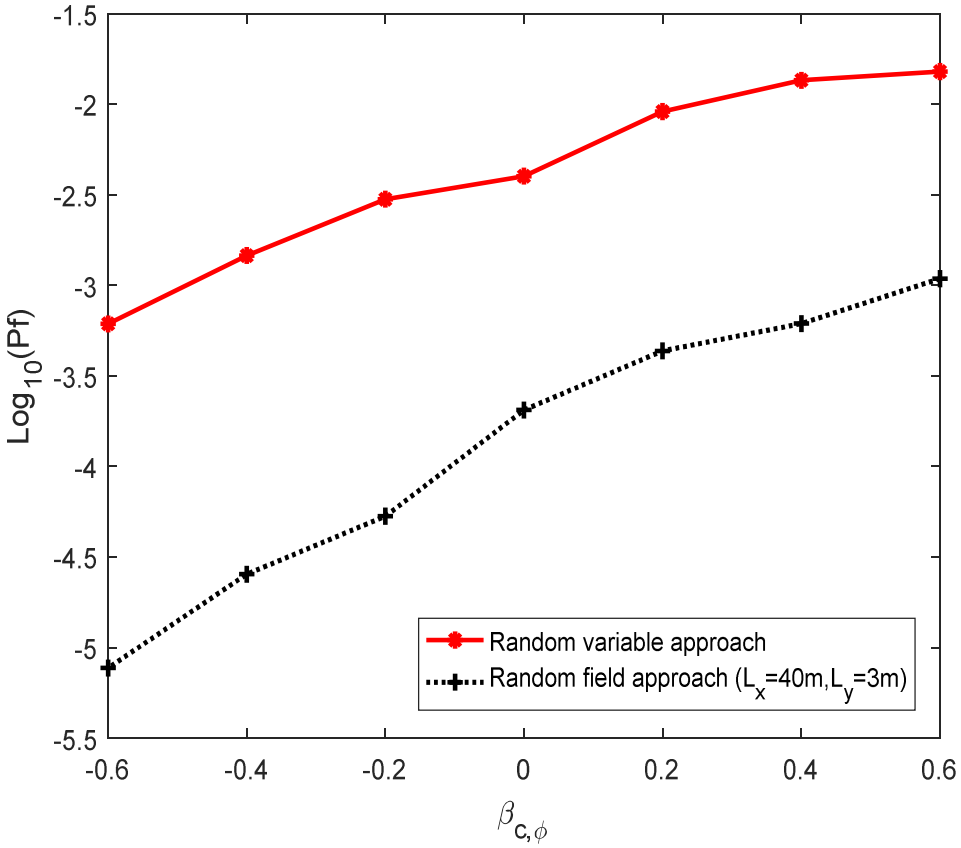


Figure 11 Effect of cross-correlation coefficients between c and ϕ on the failure probability

Table 7 Comparison of the probabilistic results obtained with different cross-correlation coefficients for the cases of random fields ($L_x = 40m$, $L_y = 3m$)

$\beta_{c,\phi}$	P_f	Mean_{FoS}	Std_{Fos}	S_C	S_ϕ
-0.6	0.08×10^{-4}	1.475	0.121	0.07	0.92
-0.4	0.25×10^{-4}	1.469	0.128	0.08	0.91
-0.2	0.53×10^{-4}	1.463	0.136	0.08	0.91
0	2.05×10^{-4}	1.455	0.142	0.09	0.90
0.2	4.35×10^{-4}	1.450	0.148	0.07	0.92
0.4	6.15×10^{-4}	1.443	0.153	0.08	0.91
0.6	10.9×10^{-4}	1.439	0.159	0.08	0.91

6.3 Effect of the COV

For the previous analyses, the COVs of c and ϕ are fixed as shown in Table 4. In order to discuss the effect of the COV on the probabilistic results, three values were used for the COV of c and ϕ to perform the probabilistic analysis. According to [38], the COV of cohesion ranges from 10% to 55%, while for friction angle, it ranges from 5% to 15%.

Figure 12 presents the PDFs of the studied dam safety factor obtained with three COVs of the cohesion (10%, 30% and 50%), with COV_ϕ being 15%. It can be seen that a small COV_c produces a taller and narrower PDF curve, which indicates that the variability decreases with the COV_c decreases. For a same value of COV_c, it is always the random variables approach which produces the most spread out PDF curve. The similar observations can be remarked in Figure 13 which gives the PDFs corresponding to three COVs of the friction angle (5%, 10% and 15%), with COV_c being 30%. By comparing these two figures, it can be found that COV_ϕ has a more important influence on the shape of the PDF curve which means that the variability of FoS is more sensitive to the change of COV_ϕ. This finding is consistent to the computed Sobol indices of c and ϕ as show in Table 6. In addition, it can also be clearly observed in Figure 13 that COV_ϕ has a non-negligible effect on the mean value and standard deviation of the FoS. It seems that the mean value increases while the standard deviation decreases when decreasing the COV_ϕ. The PDF curve will become a vertical line as show in Figure 13, which in fact corresponds to the deterministic calculation using the parameters of Table 2 ($COV_c = COV_\phi = 0$).

Figure 14 presents the failure probability as a function of the COV_c and COV_ϕ. It shows that the COV has a significant effect on the failure probability. It decreases with the COV decreases. In addition, the random variables approach always results in a bigger failure probability than the random fields approach.

Figure 15 presents the Sobol indices of c and ϕ for different COVs. It is not surprising to find that the Sobol indices are greatly affected by the COV. Another remark is that the random variables and random fields approaches give almost the same results for the Sobol Indices.

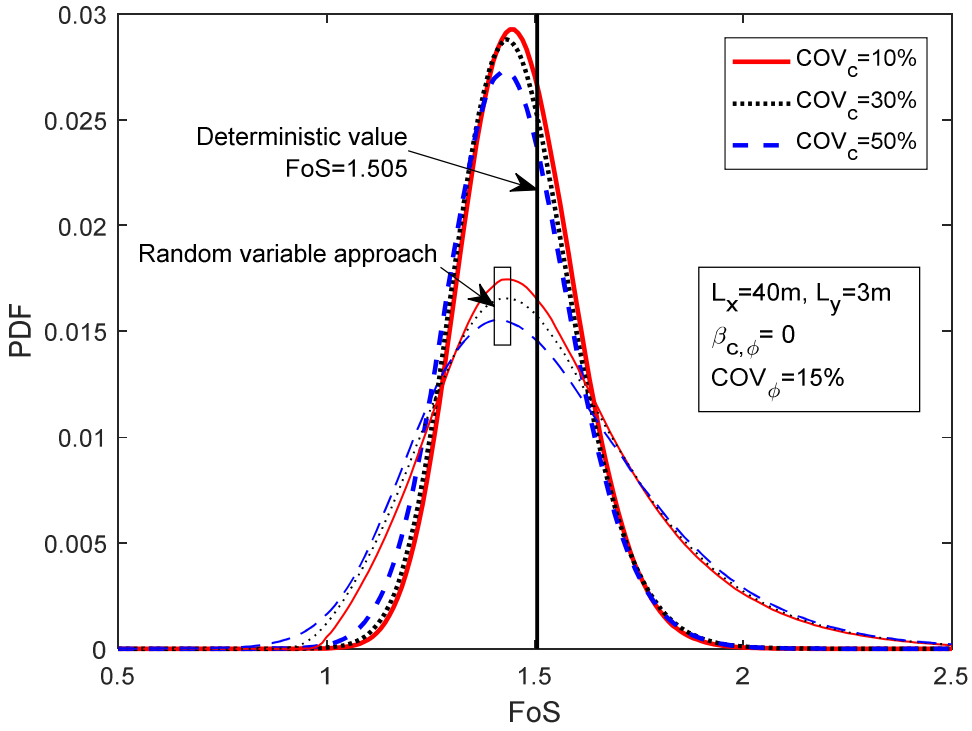
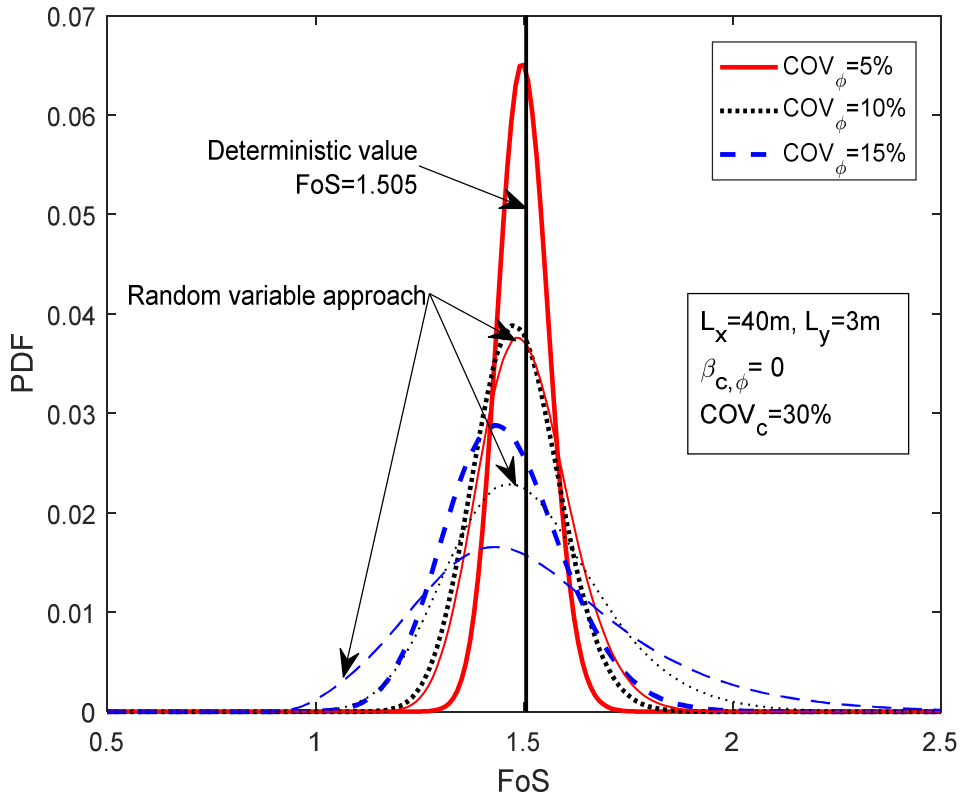


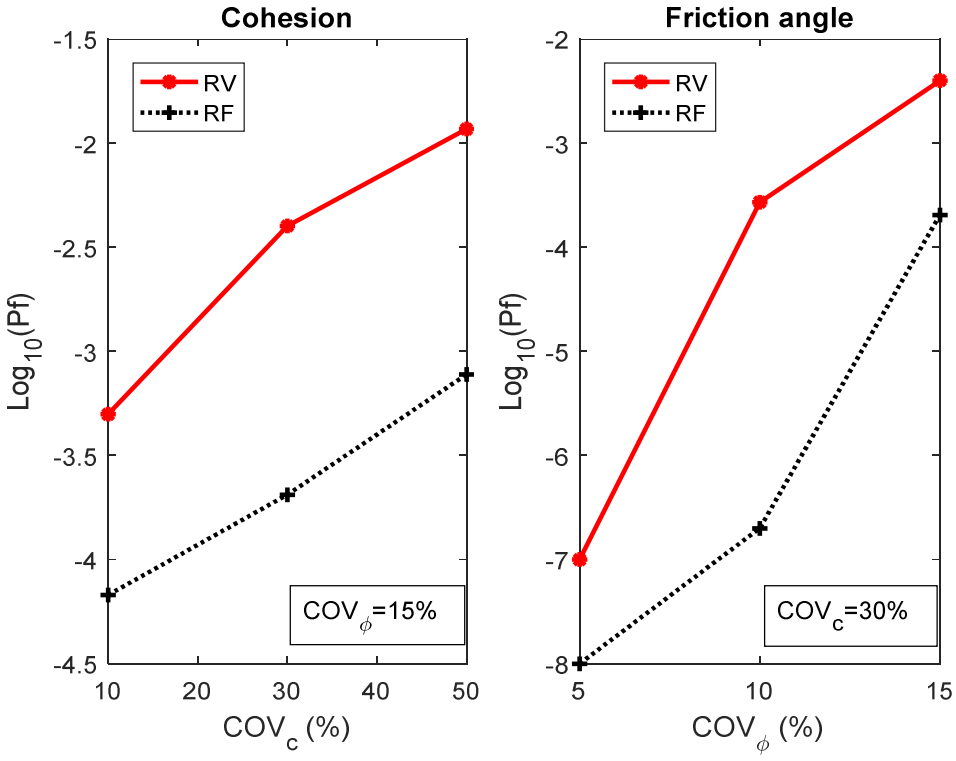
Figure 12 Effect of the COV of cohesion on the PDF of FoS (Random fields and random variables approach)



679

680 **Figure 13: Effect of the COV of friction angle on the PDF of FoS (Random fields and random variables**

681 **approach)**



682

Figure 14: Effect of the COV on the failure probability (RV: Random Variables approach, RF: Random Fields approach)

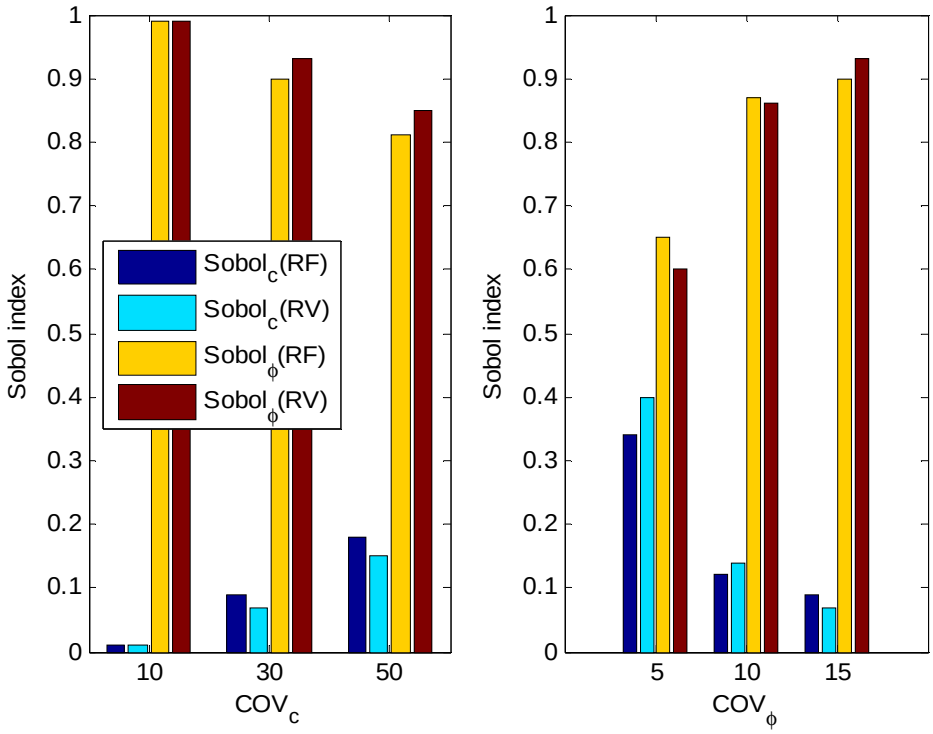


Figure 15: Effect of the COV on the Sobol index (RV: Random Variables approach, RF: Random Fields approach)

7 Discussions

7.1 Simulations of soil variabilities

As presented in the paper, there are two ways to simulate soil variabilities: random variables approach (RVA) and random fields approach (RFA). The first one assumes that a soil property is constant for the whole domain of the considered model, but is varied as a random variable among different realizations of calculations. The latter enables taking into account the soil inherent spatial variabilities by using a random field which is a collection of a series of random variables at different discretized points of the calculation domain. The value at each point is not a constant but varied according to an autocorrelation function. Obviously, the RFA is more realistic than

the RVA. In addition, the results of section 6 show that the RVA always leads to an higher failure probability than the RFA. Above all, it is better to employ the random fields approach in reliability analyses, so that the soil spatial variabilities can be simulated properly and uneconomical designs by the RVA can be avoided.

In practice, the quantity of available measurements is usually not large enough which doesn't allow meaningful estimates of autocorrelation distance for the considered soil properties. Therefore, it is impossible to construct a random field model through measured data. A common solution for this problem is to select the autocorrelation distance from literature as the works in [10,39]. A range of values for this parameter can be found in published papers. For example, [34] indicated that the autocorrelation distance of soils was between 10-40m in the horizontal direction and 1-3m in the vertical direction; [40] found that the offshore soils have a more wide range for this parameter saying 7-9000m and 0.4-7.14m for respectively horizontal and vertical directions. According to the results of section 6.1, the higher the autocorrelation distance is, the bigger the failure probability is obtained. In this case, the maximum values of this parameter can be selected in reliability analyses in order to be safe enough. Otherwise, one can perform the analysis with different values of autocorrelation distance and give finally a range of failure probability.

In some cases, other information during the construction or monitoring phases can help to determine the autocorrelation distance. For example, it is reasonable to assume $l_y = 2m$ for an earth dam if it was constructed by layers with a depth of 2m. Also, one can use physical relation to obtain a representation of random fields for the target soil property by transforming from another soil property random field. Transformation example can be found in [29].

7.2 Comparison between the two deterministic models

Two deterministic models are created in the paper and are coupled with the SPCE/GSA reliability analyses. Results of section 4.3 and section 5.2 indicate that the two models lead to similar results in both deterministic and probabilistic frameworks. It is found that using the analytical model can release dramatically the computational burden compared to the numerical one, especially in reliability analyses. The time reduction could be 5 days to only 2 hours for an analysis with 1500 calls of deterministic models as presented in Table 5. Therefore, it is recommended to use the analytical model for the dam FoS estimation in reliability analyses since the model is efficient and accurate enough. Its efficiency has a huge advantage when parametric studies should be performed or several values of a parameter should be tested in a reliability analysis.

On the contrary, the analytical model used in the paper cannot simulate complex design scenarios such as the settlement of the dam, the dam stability under rapid drawdown conditions or dam response under real seismic loadings. For these cases, a numerical model should be considered to model the complex dam behaviour. In addition, the numerical model based on the strength reduction method (SRM) is considered to be more sophisticated than the limit equilibrium method (the analytical one) since there is no assumption about slip surface shape and inter-forces and the critical slip surface can be determined automatically in SRM [29]. Therefore, the numerical model shall be recommended if there is a high requirement for the modelling of soil behaviour or failure mechanism.

In conclusions, the analytical model can be adopted for the following cases: 1) A simple design scenario, namely the dam (or slope) stability under steady state flow conditions with or without a pseudo static acceleration; 2) Parametric studies where several reliability analyses should be performed; 3) A preliminary design stage which

allows a fast estimate of the order of magnitude for the target system response. For other cases such as complicated design scenarios, the numerical model should be employed.

7.3 Other reliability methods

The present study employs the random field theory to simulate soil spatial variabilities. In total, 141 random variables are considered in the conducted reliability analyses: 70 for the c random field, 70 for the ϕ random field and 1 for the pseudo-static acceleration coefficient.

Traditional methods, like FORM and SORM, cannot handle too many random variables [10]. Particularly, the computational cost increases rapidly with the number of input variable in SORM [41]. These two approximation methods are not efficient in such a problem. Concerning the sampling-based methods, such as Monte Carlo Simulation (MCS) and Subset Simulation (SS), they are independent of the dimension of the input space. However, the efficiency of this class of methods is usually low especially for small failure probabilities. Section 5.2 has shown that the adopted method is much more efficient than the direct MCS which needs 2.1×10^5 model evaluations and takes about 2.5 days. The third class of methods for reliability analyses is the meta-modelling. Meta-Modelling aims at creating a surrogate model (a.k.a. meta-model) of the original mechanical model which is usually expensive to evaluate, by using mathematical tools (e.g. Polynomial Chaos and Kriging). Then, an MCS can be performed using the meta-model. This class of methods can address the difficulties in non-linear problems and rare event estimation encountered by FORM/SORM and Sampling-based methods. However, they suffer from the so-called *Curse of Dimensionality* [8] which means that the number of model evaluations increases rapidly with the number of input variables. The adopted method SPCE/GSA alleviates this issue by two ways: 1) reduce the input dimension by performing a

GSA; 2) construct a meta-model by using sparse PCE which results in fewer terms than a full PCE. The obtained results in the paper also demonstrate that the SPCE/GSA can provide reliable, accurate and efficient solutions for reliability analyses of dams.

8 Conclusions

Probabilistic stability analyses of an embankment dam considering soil spatial variability are presented. The analyses are performed by using the method SPCE/GSA and the soil spatial variability is simulated by the random field theory. Two deterministic models are established for computing the FoS of the dam. By performing several probabilistic analyses with the analytical model, the influences of the soil spatial variability (autocorrelation length and coefficient of variation) and the cross-correlation between the random fields of cohesion and friction angle on the reliability results are investigated.

The main conclusions of this paper can be summarized as follows:

1. The introduced analytical model is validated in a probabilistic framework considering soil spatial variability and random pseudo-static accelerations. The model can reduce dramatically the computation time of a probabilistic analysis. For a first order of estimate, it can be applied at the preliminary design stage for similar problems (earth dam or slope stability) rather than using the complex numerical model.
2. Accounting for soil spatial variabilities in reliability analyses of dams can provide less dispersive FoS values and lower estimates of failure probability compared to the random variables approach. In other words, neglecting soil spatial variabilities leads to conservative designs thus uneconomical constructions in some cases. Additionally, simulation of soil variabilities by the random field theory is realistic since the soil properties are spatially varied due to geological process and engineering construction.
3. According to the obtained Sobol indices, the variability of the studied dam safety factor is much more sensitive to the friction angle than the cohesion. In addition, the autocorrelation length and the cross-correlation coefficients between the random fields of cohesion and friction angle have almost no influence on the Sobol indices. On the contrary, the coefficient of variation has a significant effect.

4. It is found that each of the three considered factors for soil variables (autocorrelation distance, cross-correlation and coefficient of variation) has significant effects on the dam reliability. An order or several orders of magnitude for the dam failure probability can be induced by using different values of the three factors. For example, the failure probability changes 3 orders of magnitude when $\beta_{c,\phi}$ varies from −0.6 to 0.6 for the cases of random fields. More specifically, the dam failure probability is increased with increasing the values of the three factors. Particularly, the vertical autocorrelation distance has a more important effect on the failure probability than the horizontal one.
5. It is found that simulating the soil variability by random variables or random fields can lead to similar results in terms of sensitivity index. This finding could be very useful if a sensitivity analysis would have to be performed by considering the soil spatial variability. The sensitivity index of the soil properties can be approximated by using the random variables approach given that it is always easier to handle several input variables than to deal with a high dimensional stochastic problem.

9 Acknowledgement

The first author thanks gratefully the China Scholarship Council for providing him with a PhD Scholarship for his research work.

10 References

[1] Liang RY, Nusier OK, Malkawi AH. A reliability based approach for evaluating the slope stability of embankment dams. Eng Geol 1999;54:271–85. doi:10.1016/S0013-7952(99)00017-4.

[2] Babu GS, Srivastava A. Reliability Analysis of Earth Dams. J Geotech Geoenvironmental Eng 2010;136:995–8. doi:10.1061/(ASCE)GT.1943-5606.0000313.

[3] Yi P, Liu J, Xu C. Reliability Analysis of High Rockfill Dam Stability. Math Probl Eng 2015;2015. doi:10.1155/2015/512648.

[4] Guo X, Dias D, Carvajal C, Peyras L, Breul P. Reliability analysis of embankment dam sliding stability using the sparse polynomial chaos expansion. Eng Struct 2018;174:295–307. doi:10.1016/j.engstruct.2018.07.053.

[5] Dasaka SM, Zhang LM. Spatial variability of in situ weathered soil. Géotechnique 2012;62:375–84. doi:10.1680/geot.8.P.151.3786.

[6] Sudret B. Global sensitivity analysis using polynomial chaos expansions. Reliab Eng Syst Saf 2008;93:964–79. doi:10.1016/j.ress.2007.04.002.

[7] Al-bittar T, Soubra A-H. Efficient sparse polynomial chaos expansion methodology for the probabilistic analysis of computationally-expensive deterministic models. Int J Numer Anal Methods Geomech 2014;38:1211–30. doi:10.1002/nag.2251.

[8] Pan Q, Dias D. An efficient reliability method combining adaptive Support Vector Machine and Monte Carlo Simulation. Struct Saf 2017;67:85–95. doi:10.1016/j.strusafe.2017.04.006.

[9] Pan Q, Dias D. Sliced inverse regression-based sparse polynomial chaos expansions for reliability analysis in high dimensions. Reliab Eng Syst Saf 2017;167:484–93. doi:10.1016/j.ress.2017.06.026.

[10] Pan Q, Dias D. Probabilistic evaluation of tunnel face stability in spatially random soils using sparse polynomial chaos expansion with global sensitivity analysis. Acta Geotech 2017;12:1415–29. doi:10.1007/s11440-017-0541-5.

[11] Blatman G, Sudret B. Adaptive sparse polynomial chaos expansion based on least angle regression. J Comput Phys 2011;230:2345–67. doi:10.1016/j.jcp.2010.12.021.

[12] Blatman G, Sudret B. An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. Probabilistic Eng Mech 2010;25:183–97. doi:10.1016/j.probengmech.2009.10.003.

[13] Blatman G, Sudret B. Sparse polynomial chaos expansions and adaptive stochastic finite elements using a regression approach. Comptes Rendus Mécanique 2008;336:518–23. doi:10.1016/j.crme.2008.02.013.

[14] Sudret B. Polynomial Chaos Expansions and Stochastic Finite Element Methods. In: Phoon K-K, Ching J, editors. Risk Reliab. Geotech. Eng., CRC

873 Press; 2014, p. 265–300.
874 [15] Sobol' IM. Global sensitivity indices for the investigation of nonlinear
875 mathematical models. Mat Model 2005;17:43–52.
876 doi:10.1002/wilm.42820050114.
877 [16] Marelli S, Sudret B. UQLab: A Framework for Uncertainty Quantification in
878 MATLAB. 2nd Int. Conf. Vulnerability Risk Anal. Manag., Liverpool: 2014, p.
879 2554–2563. doi:10.1061/9780784413609.257.
880 [17] Baudin M, Dutfoy A, Iooss B, Popelin A. OpenTURNS: An Industrial
881 Software for Uncertainty Quantification in Simulation. In: Ghanem R, Higdon
882 D, Owhadi H, editors. Handb. Uncertain. Quantif., Cham: Springer
883 International Publishing; 2017, p. 2001–38. doi:10.1007/978-3-319-12385-
884 1_64.
885 [18] Phoon KK. Numerical recipes for reliability analysis – a primer. In: Phoon K-
886 K, Ching J, editors. Reliab. Des. Geotech. Eng., CRC Press; 2008, p. 545.
887 [19] Sudret B, Der Kiureghian A. Stochastic finite element methods and reliability.
888 A state-of-the-art-report. University of California, 2000.
889 [20] Cho S. Probabilistic Assessment of Slope Stability That Considers the Spatial
890 Variability of Soil Properties. J Geotech Geoenvironmental Eng 2009;136:975–
891 84. doi:10.1061/(ASCE)GT.1943-5606.0000309.
892 [21] Ahmed A, Soubra A. Application of the subset simulation approach to spatially
893 varying soils. In: Phoon K-K, Ching J, editors. Risk Reliab. Geotech. Eng.,
894 CRC Press; 2014, p. 561–77.
895 [22] Ghiocel DM, Ghanem RG. Stochastic Finite-Element Analysis of Seismic Soil-
896 Structure Interaction. J Eng Mech 2002;128:66. doi:10.1061/(ASCE)0733-
897 9399(2002)128:1(66).
898 [23] Baecher GB, Christian JT. Reliability and Statistics in Geotechnical
899 Engineering. John Wiley & Sons; 2005. doi:10.1198/tech.2005.s838.
900 [24] Li C-C, Der Kiureghian A. Optimal discretization of random fields. J Eng

Mech 1993;119:1136–54. doi:10.1061/(ASCE)0733-9399(1993)119:6(1136).

[25] Cho SE, Park HC. Effect of spatial variability of cross-correlated soil properties on bearing capacity of strip footing. Int J Numer Anal Methods Geomech 2010;34:1–26. doi:10.1002/nag.791.

[26] Fenton GA, Griffiths D V. Risk Assessment in Geotechnical Engineering. 2008. doi:10.1002/9780470284704.

[27] Itasca. FLAC 7.0 reference manual. Minneapolis: 2011.

[28] Guo X, Baroth J, Dias D, Simon A. An analytical model for the monitoring of pore water pressure inside embankment dams. Eng Struct 2018;160:356–65. doi:10.1016/j.engstruct.2018.01.054.

[29] Mouyeaux A, Carvajal C, Bressolette P, Peyras L, Breul P, Bacconnet C. Probabilistic stability analysis of an earth dam by Stochastic Finite Element Method based on field data. Comput Geotech 2018;101:34–47. doi:10.1016/j.compgeo.2018.04.017.

[30] Der Kiureghian A, Ke J-B. The stochastic finite element method in structural reliability. Probabilistic Eng Mech 1988;3:83–91. doi:10.1016/0266-8920(88)90019-7.

[31] Morgenstern NR, Price VE. The Analysis of the Stability of general slipe surfaces. Geotechnique 1965;15:725–6. doi:10.1680/geot.1965.15.1.79.

[32] Zhu DY, Lee CF, Qian QH, Chen GR. A concise algorithm for computing the factor of safety using the Morgenstern-Price method. Can Geotech J 2005;42:272–8. doi:10.1139/t04-072.

[33] Fenton GA, Griffiths DV. Extreme hydraulic gradien statistics in stochastic earth dam. J OfGeotechnical Geoenvironmental Eng 1997;123:14775. doi:10.1061/(ASCE)1090-0241(1997)123:11(995).

[34] El-Ramly H, Morgenstern NR, Cruden DM. Probabilistic stability analysis of a tailings dyke on presheared clay-shale. Can Geotech J 2003;40:192–208. doi:10.1139/t02-095.

929 [35] Lumb P. Safety factors and the probability distribution of soil strength. Can
930 Geotech J 1970;7:225–42. doi:10.1139/t70-032.
931 [36] Yucemen MS, Tang WH, Ang A.-S. A probabilistic study of safety and design
932 of earth slopes. University of Illinois at Urbana-Champaign, 1973.
933 [37] Wolff TH. Analysis and design of embankment dam slopes: A probabilistic
934 approach. Purdue University, 1985.
935 [38] Phoon K-K, Kulhawy FH. Characterization of geotechnical variability. Can
936 Geotech J 1999;36:612–24. doi:10.1139/t99-038.
937 [39] Kader A, Haj E, Soubra AH, Fajoui J, Al-bittar T. Probabilistic model of an
938 offshore monopile foundation taking into account the soil spatial variability.
939 Comput Geotech 2019;106:1–19. doi:10.1016/j.compgeo.2018.10.011.
940 [40] Li JH, Zhou Y, Zhang LL, Tian Y, Cassidy MJ, Zhang LM. Random finite
941 element method for spudcan foundations in spatially variable soils. Eng Geol
942 2016;205:146–55. doi:10.1016/j.enggeo.2015.12.019.
943 [41] Marelli S, Schobi R, Sudret B. UQLAB user manual – Structural reliability
944 (Rare event estimation), Report, UQLab-V1.2-107. Zurich: 2019.
945