Research Paper

Efficient estimation of the failure probability of a monopile foundation using a Kriging-based approach with multi-point enrichment

Abdul-Kader El Haj, Abdul-Hamid Soubra

University of Nantes, Bd. de l'Université, 44603 Saint-Nazaire Cedex, France

ABSTRACT

In this paper, a probabilistic analysis of an offshore monopile foundation embedded in a spatially varying clayey soil was performed. An efficient Kriging-based probabilistic approach using a multipoint enrichment was adopted for the analysis. The aim is to compute the failure probability P_f against exceeding a threshold value on the monopile head rotation. The soil undrained cohesion was considered as a random field with a mean value that linearly varies with depth. The proposed probabilistic approach was shown to significantly reduce the computation time with respect to the so-called AK-MCS classical Active learning method combining Kriging and Monte Carlo Simulation (MCS) methodology. Some probabilistic numerical results are presented and discussed.

1. Introduction

The probabilistic analysis of geotechnical structures taking into account the spatial variability of the soil properties has been performed for several years using Monte Carlo Simulation (MCS) methodology (e.g. [1–5]). More recently, various reduction techniques such as Subset Simulation (SS) or Asymptotic Sampling (AS) were used for the probabilistic analysis (e.g. [6–9]). These methods aim to reduce the number of calls to the mechanical model especially when dealing with the small practical values of the failure probability.

In order to reduce once again the computation time induced by the use of a time-consuming mechanical model, several authors have resorted to the combination of a simulation method [e.g. Subset Simulation (SS), Importance Sampling (IS) or Monte Carlo Simulation (MCS)] with a metamodeling technique. For instance, [10] proposed a method combining the subset simulation and the support vector machines, [11] proposed a combination between Kriging and Monte Carlo Simulation, [12] proposed a methodology incorporating neural network and subset simulation and [13] developed a Kriging-based adaptive importance sampling approach. The Active learning method combining Kriging with Monte Carlo Simulation methodology (called AK-MCS) (see [11]) has gained a significant popularity in the litterature. This approach consists in replacing the time-consuming mechanical model by a simple Kriging meta-model calibrated by a limited number of mechanical model evaluations. The aim is to apply MCS methodology on the calibrated metamodel (called also surrogate model) with a quasi negligible computational time.

AK-MCS approach was used by [14] for the probabilistic analysis

against soil punching of a strip footing resting on a spatially varying soil. Within AK-MCS approach, a preliminary surrogate model is constructed by Kriging metamodeling using a small Design of Experiments DoE (i.e. a small set of samples). The obtained approximate meta-model is then successively improved through an enrichment process in which a powerful learning function is used for the selection of the 'best' samples to be evaluated by the computationally expensive mechanical model. Notice that in AK-MCS method, a single sample is selected per iteration of the enrichment process. Indeed, AK-MCS is unable to parallelize different mechanical computations. This is a drawback in the case where distributed (or parallel) computing facilities are to be used in the aim to reduce the computation time.

In this paper, a multipoint enrichment technique called AK-MCSm (where m stands for multi-enrichment) is presented. The aim is to allow several evaluations of the performance function to be carried out simultaneously. A relevant clustering technique (see [15]) was adopted in this paper. This technique allows one to consider a set of learning samples that ensure a suitable coverage of the limit state surface. Concerning the stopping condition of the enrichment process, this paper employs the stopping criterion that was recently proposed by [16]. This criterion is more relevant than the one used in AK-MCS because it is based on the convergence of the quantity of interest (i.e. the failure probability).

This paper focuses on the probabilistic analysis of an offshore monopile foundation embedded in a spatially varying clayey soil and subjected to vertical and lateral loadings. The aim is to compute the failure probability P_f against exceeding a threshold value on the monopile head rotation. Previous work by the authors of this paper has

considered this problem by making use of a Global Sensitivity-enhanced Surrogate (GSAS) modelling (see [17]). This method is a kind of variant of the AK-MCS method in which the global sensitivity analysis was used to help a better selection of the training sample to be computed by the original mechanical model. As is the case of AK-MCS approach, GSAS technique suffers from the fact that a unique sample is used in each iteration of the enrichment process.

In this paper, one performs a probabilistic analysis at the ultimate limit state (ULS) of a monopile foundation embedded in a spatially varying soil making use of the relevant multi-enrichment approach AK-MCSm. A clayey soil for which the soil undrained cohesion varies with depth is considered in the analysis. Compared to AK-MCS approach, AK-MCSm allows not only to consider several computations of the mechanical model during each step of the enrichment process, but also a better choice of the training samples (or points) by making use of the information provided by the learning function. This results in a significant reduction in the computation time with respect to the classical AK-MCS approach.

The paper is organized as follows: The next section presents the mechanical model. This is followed by an overview on the soil spatial variability and the discretization of random fields. Then, the probabilistic model is described. Finally, some probabilistic numerical results are presented and discussed. The paper ends by a conclusion of the main findings.

2. Mechanical model

Generally, offshore monopile foundations are subjected to high horizontal loadings (in addition to the vertical loading due to the structure weight) which will result in significant horizontal and moment loadings applied at the mudline. The horizontal and moment loadings may lead to large horizontal displacement and rotation of the monopile at this level. DNV code [18] recommended to prescribe a practical criterion on the lateral deflection or on the rotation of the pile at mudline. This is because excessive lateral displacement and rotation may occur at the pile head before attaining the ultimate lateral soil resistance. Indeed, the lateral soil resistance cannot be locally mobilized in the zones near the pile rotation point, regardless of how much the pile head deflects laterally. In this regards, focus is put in this paper on the monopile rotation at the mudline.

Both the ultimate rotation (i.e. the rotation corresponding to the ultimate load) and the rotation induced by the applied loadings are sought in this section. The mechanical model has been carried out based on numerical simulations using the commercial finite element software Abaqus/Standard [19].

An open-ended steel monopile of diameter $D = 4m$ was studied in this paper. The monopile of 0.05 m thickness and an embedment depth L of 24 m was extended of 1.0 m above the seabed to prevent the soil from going over the monopile which would generate unrealistic results [20]. The steel monopile has a density of 7840 kg/m^3. It was assumed to behave as a linear elastic material with Young's modulus E_p of 210 GPa and Poisson's ratio ν_p of 0.3.

The soil consists of an undrained normally consolidated clay. It was supposed to follow an elastic-perfectly plastic model based on Tresca criterion. This model is defined by the undrained cohesion c_u, the undrained Young's modulus (E_u) and the Poisson's ratio (ν_u). In this paper, the soil has a submerged unit weight of 7 kN/m^3 and a Poisson's ratio of 0.495. The undrained cohesion was supposed to linearly vary with depth as follows:

$$c_u = c_{u,m} + k_{cu} \cdot \sigma'_{v0} \tag{1}$$

where $c_{u,m}$ is the value of the undrained cohesion at mudline, k_{cu} is a material constant for the clay [21] and σ'_{v0} is the effective vertical overburden stress. [22] proposed a relationship between k_{cu} and the soil plasticity index PI as follows:

$$k_{cu} = 0.11 + 0.0037PI \tag{2}$$

where PI is the value of the plasticity index given in percentage. Using the relationship given by Eq. (2) and assuming a plasticity index of 35%, a value of $k_{cu} = 0.24$ is obtained. This represents an increase in c_u from its initial value at mudline (taken here equal to 2 kPa) at a rate of about 1.68 kPa/m. These adopted values are typical for a normally consolidated offshore clay [23]. Similar values were observed for the Gulf of Mexico normally consolidated clays [24,25]. Finally, notice that the soil undrained Young modulus was assumed to be linearly related to the soil undrained cohesion via the relationship $E_u = K_c \times c_u$ where K_c is a correlation factor that is dependent on the clay plasticity index PI and the over-consolidation ratio OCR [26]. Thus, the soil Young modulus was also supposed to linearly vary with depth. For a plasticity index $PI = 35\%$ and an over-consolidation ratio of 1, a value of $K_c = 500$ is adopted in this work [26].

A wished-in-place assumption was adopted in the analysis. Such an assumption consists in considering the initial geostatic stresses in the soil around the monopile (i.e. K_0 stress state) without taking into account the change in the stress state in the disturbed soil induced by the monopile installation process. As may be seen from Fig. 1, the numerical model has a length of 12D, a width of 6D and a height of 1.6 L. It was verified that with these model dimensions, the behavior of the monopile was not influenced by the artificial boundary effects. Details on the numerical modeling of the soil-monopile system are not provided herein. The reader may refer to [17].

In this paper, the monopile was considered to be subjected to a horizontal load $H = 1.6MN$ acting at a height $h = 38.6m$ above the seabed level. This results in an additional moment at mudline of $M = 61.7\ MN.m$. A vertical load V of 2 MN that represents the structure weight was also considered in the analysis. The applied loadings adopted in this study induced a head rotation at mudline (as computed by the mechanical model) of nearly 0.55°.

Concerning the determination of the ultimate rotation at the monopile head, a load-controlled procedure was used. The corresponding moment-rotation curve is shown in Fig. 2. As it may be seen from this figure, the ultimate rotation was determined by the tangent intersection method (point A) leading to a limit rotation of $\theta_{ULS} = 1.5°$ and a corresponding ultimate moment capacity at mudline of
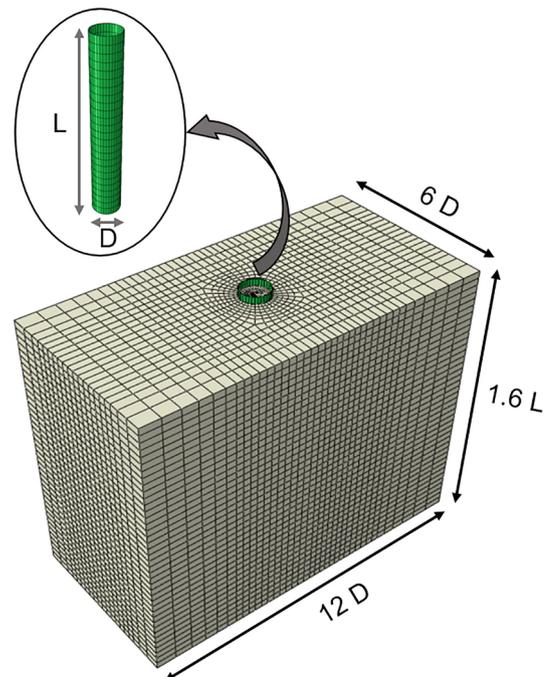
Fig. 1. Three-dimensional numerical model.
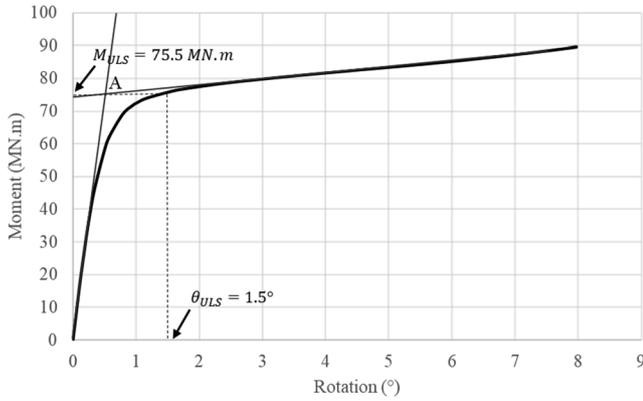
Fig. 2. Moment-rotation curve of the monopile at mud-line level.

$M_{ULS} = 75.5\ MN \cdot m$. This corresponds to an ultimate horizontal force of $1.95\ MN$. The obtained ultimate value of the monopile head rotation will be used afterwards in this paper as a threshold value when handling the probabilistic analysis (see Eq. (6)).

Fig. 3 shows the displacement field corresponding to the ultimate rotation where $M_{ULS} = 75.5\ MN \cdot m$ and $\theta_{ULS} = 1.5°$. From Fig. 3, one may observe that the monopile exhibits a rigid body rotation around a rotation point. Above the rotation point, the pile was shown to kick forward and move against the soil on the front side of the loading direction. Below the rotation point, the pile was shown to kick backward and move against the soil on the rear side of the loading direction. The soil displacement field may be divided into two zones: In the upper part of the model, the soil was shown to fail within a conical wedge mechanism that extends to the soil surface. A passive wedge is mobilised on the front side of the pile (in the direction of loading) with a gap opening up on the rear side of the pile. At a deeper depth level, soil was shown to fail according to a rotational mechanism.

3. Soil spatial variability

The soil undrained cohesion c_u was considered as a log-normal random field with a constant coefficient of variation of 25%. The mean values of the soil undrained cohesion are those of the deterministic analysis provided in the preceding section. Concerning the auto-correlation function, a square exponential function $\rho_Z^{LN}(X, X')$ was used

in this paper. This function provides the values of the correlation between two arbitrary points $X(x, y, z)$ and $X'(x', y', z')$ as follows:

$$\rho_Z^{LN}(X, X') = \exp\left[-\left(\frac{|x - x'|}{a_x}\right)^2 - \left(\frac{|y - y'|}{a_y}\right)^2 - \left(\frac{|z - z'|}{a_z}\right)^2\right] \tag{3}$$

where a_x and a_y are the horizontal autocorrelation distances in the x and y directions (y being the loading direction) and a_z is the vertical autocorrelation distance. Notice that the soil undrained Young modulus was implicitly considered as a random field having the same distribution as the soil undrained cohesion since it was assumed to be linearly related to the soil cohesion.

The discretization of the cohesion random field was performed in this paper using EOLE method proposed by [27]. The step-by-step procedure of this method in the present case of a lognormal random field can be briefly described as follows:

- Define a stochastic grid composed of s grid points. A uniform grid is adopted with five grid elements within each autocorrelation distance as recommended by [27].
- Determine the log-normal autocorrelation matrix Σ^{LN} making use of the autocorrelation function given by Eq. (3). This matrix provides the correlation between each grid point of the stochastic mesh and all the grid points of this mesh.
- Transform this log-normal autocorrelation matrix to the Gaussian space using the Nataf transformation (cf. [28]).
- Obtain the eigenvalues and eigenvectors λ_j and ϕ_j ($j = 1, 2, ..,s$) of the Gaussian autocorrelation matrix. The dicretisation of the Gaussian random field Z is given as follows:

$$\widetilde{Z}(x, y, z) \cong \mu_{\ln Z} + \sigma_{\ln Z} \sum_{j=1}^{s} \frac{\xi_j}{\sqrt{\lambda_j}} (\phi_j)^T \Omega \tag{4}$$

where $\mu_{\ln Z}$ and $\sigma_{\ln Z}$ are respectively the mean and standard deviation values of the underlying normal distribution [i.e. $\ln(Z)$], ξ_j ($j = 1, 2, ..,s$) is a set of independent standard normal random variables, Ω is the correlation vector between the value of the random field at an arbitrary point (x, y, z) and its values at the different grid points of the stochastic grid and $(\phi_j)^T$ is the transpose of the eigenvector ϕ_j. Notice finally that ϕ_j and Ω are two vectors of dimension s.

It should be mentioned here that the series given by Eq. (4) should be truncated for a number of terms M (expansion order) smaller than the number of grid points s. This is done after sorting the s eigenvalues
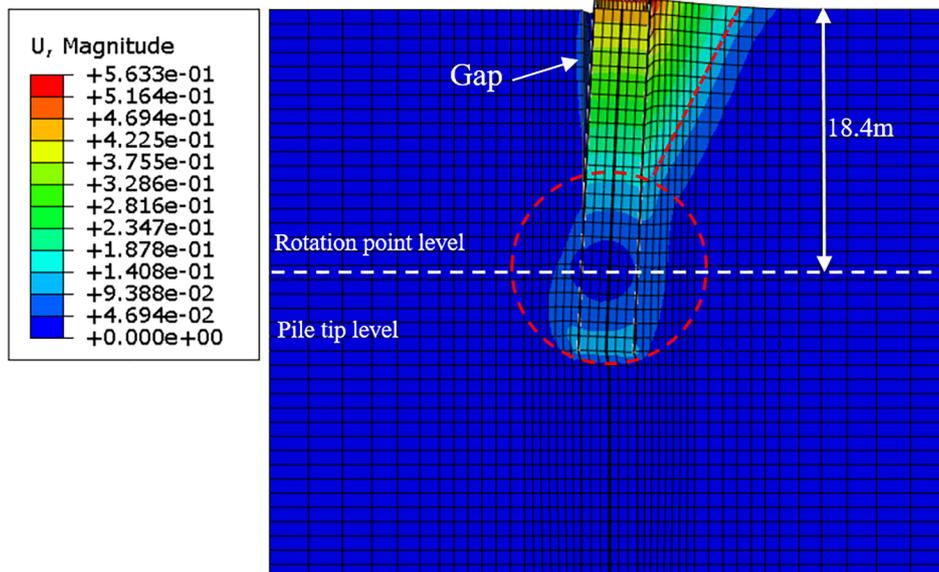
Fig. 3. Displacement field at the ultimate rotation.

$\lambda_j(j = 1, 2, .., s)$ in a descending order. The adopted number of terms should assure that the variance of the error is smaller than a prescribed tolerance. The variance of the error corresponding to a truncation with M terms is given by the following equation [29]:

$$Var\left[Z(x, y, z) - \widetilde{Z}(x, y, z)\right] = \sigma_{\ln Z}^2 \left\{ 1 - \sum_{j=1}^{M} \frac{1}{\lambda_j} [(\phi_j)^T \Omega]^2 \right\} \tag{5}$$

where $Z(x, y, z)$ and $\widetilde{Z}(x, y, z)$ are respectively the exact and the approximate values of the random field at a given point (x, y, z).

It should be emphasized here that the number of terms M represents the number of eigenmodes (or the number of random variables in the vector ξ_j). For a prescribed value of the variance of the error, the required number of eigenmodes is significant for the small values of the autocorrelation distances (i.e. for the case of a very heterogeneous soil) and it decreases with the increase of these distances. This may be explained by the fact that a greater number of eigenmodes is necessary to simulate the tight fluctuations that may occur when dealing with a very heterogeneous soil medium. The required number M of eigenmodes will be determined later in this paper.

- Once the Gaussian random field is obtained, transform it to the lognormal space by exponentiating the approximated Gaussian random field $\widetilde{Z}(x, y, z)$ given by Eq. (4).

Fig. 4 (a, b and c) present the distribution of the soil undrained cohesion for three typical random field realizations as generated by EOLE method for a one-dimensional vertical random field where $a_z = 2$ m. Notice that the realizations of the soil Young modulus random field can be easily obtained from the realizations of the soil cohesion random field by multiplying the values of the soil cohesion by 500.

4. Probabilistic analysis at the ultimate limit state (ULS)

The probabilistic analysis at ULS aims at computing the failure probability P_f against exceeding a threshold value on the ultimate rotation of the monopile head. The performance function is given by:

$$G = \frac{\theta_{ULS}}{\theta} - 1 \tag{6}$$

where θ_{ULS} is the ultimate rotation of the monopile head ($\theta_{ULS} = 1.5°$) as was determined before using the mean values of E_u and c_u and θ is the monopile head rotation under the applied loading for typical realizations of c_u and E_u. The computation time of θ for each simulation *via* Abaqus software was equal to about one hour.

4.1. Probabilistic method

The probabilistic method adopted in this paper makes use of the Kriging metamodeling theory (see [30]). The Kriging metamodeling consists in constructing a meta-model (i.e. a substitute of the original mechanical model) based on a few number of samples computed using the original mechanical model. The predicted response at an unknown sample (based on the constructed Kriging meta-model) is a Gaussian random variate characterized by a mean prediction value and a corresponding prediction variance. Thus, a major advantage of Kriging is that it provides not only a predicted value at an unknown sample but also an estimate of the prediction variance for this sample (which gives an uncertainty indication on the predicted sample response). Notice that contrary to many types of meta-models, the Kriging meta-modeling technique does not present the potential deficit of not capturing local extrema since the interpolation of the sampled data is carried out using a Maximum Likelihood Estimation (MLE) procedure. Notice also that a considerable number of samples must be considered within the exploited area. It should be emphasized that only a particular zone of the variable space, namely the zone of the Limit State Surface LSS with a high probability density (i.e. the zone of LSS that is the most close to the origin of the standard space of random variables) is required to be well exploited in this work in order to accurately estimate the failure probability.

The classical AK-MCS Kriging-based approach consists of two main stages. First, a preliminary approximate Kriging meta-model is constructed based on a small number of samples. Second, the obtained approximate Kriging meta-model is successively improved *via* an enrichment process (by adding each time a new training sample) until reaching a sufficiently accurate meta-model for the computation of the failure probability. The AK-MCSm probabilistic method used in this paper is based on the traditional AK-MCS approach but employing a multipoint enrichment strategy and a relevant clustering technique. Thus, the major advantage of the present AK-MCSm approach with
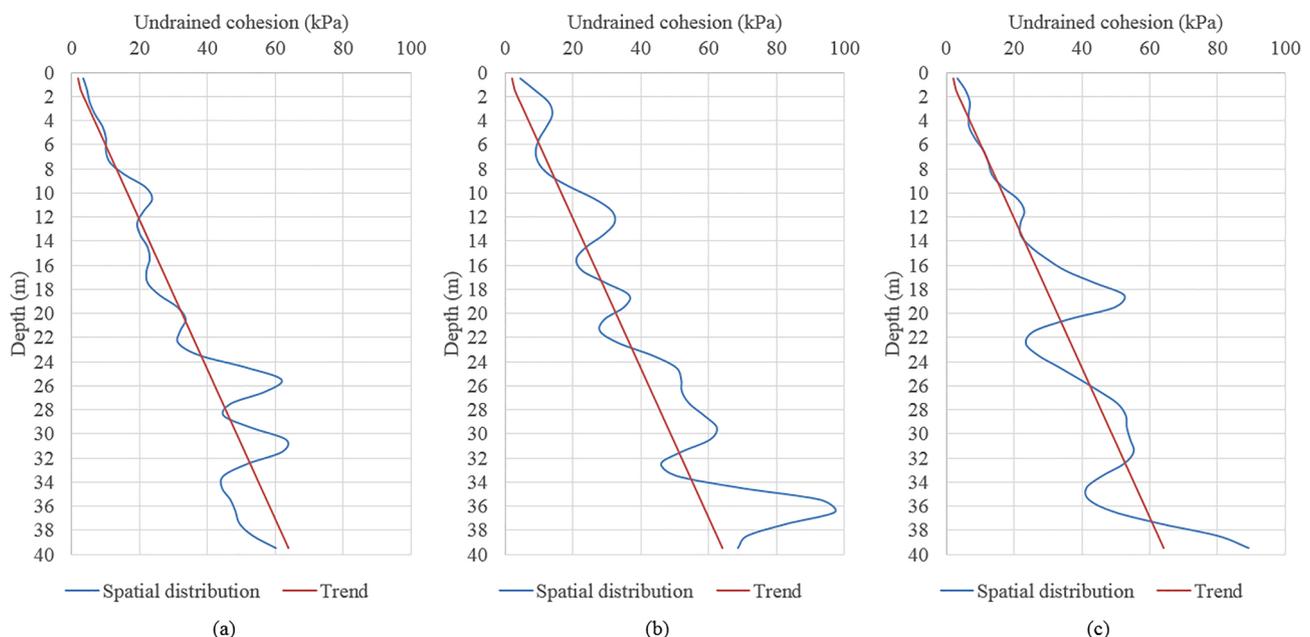
Fig. 4. Typical realizations of the cohesion random field generated by EOLE method.

respect to the classical AK-MCS is that it can consider simultaneous mechanical computations of different samples at each iteration of the enrichment process. Furthermore, the selected samples are better chosen due to the use of the information provided by the learning function as it will be shown later in this paper.

AK-MCSm approach consists of two stages. These stages are described in some details in the two following subsections in the present case of a spatially varying soil.

4.1.1. Construction of a preliminary Kriging metamodel

In this stage, one first generates a MCS population of 5×10^5 samples $\boldsymbol{x}^{(i)}$ ($i = 1, 2, \cdots, 5 \times 10^5$). Each sample $\boldsymbol{x}^{(i)}$ is a vector of M standard Gaussian random variables where M is the number of random variables required by EOLE methodology to accurately discretize the random field. Secondly, a small design of experiment DoE (taken equal to 15 samples) is randomly selected from the generated population. Each sample of the DoE is transformed (using EOLE) into a realization of c_u and a corresponding realization of E_u using $E_u = 500 \times c_u$. These realizations are used as inputs for the mechanical model while computing the sample system response (i.e. monopile head rotation θ) and the corresponding performance function value.

By using the DACE toolbox (cf. [31]), an approximate Kriging metamodel may be constructed in the standard space of random variables on the basis of the DoE and the corresponding performance function values. This meta-model may be used to compute the MCS failure probability P_f given by:

$$P_f = \sum_{i=1}^{N_{MCS}} I(G_p(x^{(i)})) \bigg/ N_{MCS} \tag{7}$$

The meta-model random responses $G_p(x^{(i)})$ in this equation are replaced by the mean prediction values $\hat{g}(x^{(i)})$ of the Kriging meta-model $\hat{g}(x^{(i)})$. Notice also that N_{MCS} in Eq. (7) is the number of MCS samples (i.e. 5×10^5 samples) and $I(G_p(x^{(i)})) = 1$ if $G_p(x^{(i)}) \leq 0$; otherwise, $I(G_p(x^{(i)})) = 0$. The coefficient of variation of the failure probability $COV(P_f)$ is given by the following equation:

$$COV(P_f) = \sqrt{\frac{1 - P_f}{P_f . N_{MCS}}} \tag{8}$$

It should be noted that the value of P_f and the corresponding value of $COV(P_f)$ computed so far are not sufficiently accurate. This is because of the very small number of samples (DoE) used to construct the present preliminary Kriging metamodel. Thus, an enrichment process is needed.

4.1.2. Enrichment process

The enrichment process is done *via* an active learning technique. The learning phase stops once the metamodel becomes sufficiently accurate for the computation of the failure probability, which is indicated by a stopping criterion. The aim of the next two subsections is to present the way of selection of the new training samples during the enrichment process and the adopted stopping criterion.

4.1.2.1. Selection of new training samples.
The enrichment process of the AK-MCS method is performed using a learning function U defined by the following equation (see [11]):

$$U(x^{(i)}) = \frac{|\hat{g}(x^{(i)})|}{\sigma_{G_p}(x^{(i)})} \tag{9}$$

where σ_{G_p} is the square root of the Kriging prediction variance. The sample that has the minimum value of U is selected for the enrichment since it is considered to have the highest probability of being misclassified (i.e. the highest probability to have a wrong performance function sign) as was stated by [11]. It should be remembered that AK-MCS method involves a single sample per iteration of the enrichment process. In order to overcome this

shortcoming, a multipoint enrichment procedure is adopted in this paper. A simple approach would be to randomly select the set of samples having high probabilities of misclassification (i.e. the samples having the minimal values of U) to be evaluated simultaneously. However, a better solution could be obtained by using a clustering technique that ensures a better coverage of the samples along the limit state surface.

The conventional K-means clustering technique aims at finding the geometric centroid of each cluster using its arithmetic mean [32]. However, this technique does not consider the information provided by the learning function and thus, the obtained centroids are not the optimal ones for the enrichment. In order to account for the relative importance of the samples in a cluster, a weighted K-means clustering algorithm may be used [33]. In this algorithm, larger weights are dedicated to the samples with high information values according to the learning function.

Lelièvre et al. [15] proposed a clustering technique, named K-weighted-means clustering algorithm (K-w-means), that takes benefit of the information provided by the AK-MCS learning function. The K-w-means technique consists in replacing the mean of each cluster by a weighted one making use of the learning function U. In this way, each sample will be weighted by the corresponding uncertainty of being misclassified (i.e. the uncertainty to have a wrong sign of its performance function) and thus, the obtained centroids of the different clusters will be the optimal ones for the enrichment. In other words, the selected samples will be situated in the highly uncertain zone (i.e. the zone corresponding to a high uncertainty on the sign of the performance function) all along the limit state surface leading to an efficient multipoint enrichment of the Kriging meta-model. This approach is used in this paper.

The step-by-step procedure of the K-w-means clustering algorithm can be described as follows:

1. Let K be the number of clusters used in the analysis. Select among the whole MCS population a number of $n_c \times K$ samples (n_c is taken equal to 5 in this paper) that have the minimal values of U. The selected samples are those that will be used in the clustering procedure.
2. Among the selected samples in the previous step, randomly select K samples and consider these samples as initial centroids for the K clusters. These centroids are denoted $[\boldsymbol{c}_1^{(1)}, \boldsymbol{c}_2^{(1)}, \cdots, \boldsymbol{c}_K^{(1)}]$.
3. Split the $n_c \times K$ samples into K sets according to Voronoi diagram [34] depending on the nearest centroid (i.e. each sample is attributed to the centroid with which it forms the smallest Euclidian distance).
4. Determine the new centroid $\boldsymbol{c}_k$ of each cluster k ($k = 1, 2, \cdots, K$) by computing the corresponding weighted mean as follows:

$$\boldsymbol{c}_k^{(i+1)} = \frac{\sum_{j=1}^{n_k} \left(\frac{1}{U_j}\right)^2 \boldsymbol{x}_j}{\sum_{j=1}^{n_k} \left(\frac{1}{U_j}\right)^2} \quad k = 1, 2, \cdots, K \tag{10}$$

where the index i stands for the considered iteration, $\boldsymbol{c}_k$ is a vector composed of M components (where M is the number of random variables), n_k is the number of samples in the kth cluster, $[\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_{n_k}]$ is the set of samples corresponding to this cluster (where each sample consists of a vector of M components) and $[U_1, U_2, \cdots, U_{nk}]$ is a set containing the learning function values of the different samples of the kth cluster. It should be noted that the number of samples n_k within a given iteration i may be different from one cluster to another one.

From Eq. (10), one may observe that the sample $\boldsymbol{x}_j$ having a high U value (i.e. the sample $\boldsymbol{x}_j$ having a small uncertainty in the sign of its performance function value) is affected with a small weight and vice versa.

5. Calculate the error expressing the sum of the squared distances between each couple of successive centroids (i.e. those corresponding to iterations and $i + 1$), as follows:

$$\sum_{k=1}^{K} (\boldsymbol{c}_k^{(i)} - \boldsymbol{c}_k^{(i+1)})^2 \tag{11}$$

If this error is below a prescribed threshold ε (taken here as 5%), the algorithm stops. Otherwise, the algorithm goes to step 3 to split the samples according to the new centroids.

It should be noted that the obtained centroids do not generally belong to the initially selected samples. Hence, the nearest sample to each centroid is chosen for the enrichment.

4.1.2.2. Stopping condition.
In AK-MCS method, the enrichment process stops when the learning function U is sufficiently large for all the MCS samples. Echard et al. [11] suggested a minimum value of $U = 2$ on these samples. This corresponds to a probability of a wrong sign of the performance function that is lower than 0.0228 (see [11]). One main issue about this criterion is that it is defined from the perspective of individual responses (not the quantity of interest P_f), which may lead to some unnecessary extra evaluations of the mechanical model. A more relevant stopping condition that is based on the convergence of the quantity of interest (i.e. P_f) was proposed by [16]. This criterion was used in this paper in the aim to reduce the computation time. Indeed, the adopted criterion relies on the convergence of the failure probability, which could be attained before reaching the stopping condition indicated by AK-MCS.

Schöbi et al. [16] define a limit state margin characterized by upper and lower boundaries of the limit state surface that takes into account the prediction uncertainty in the Kriging metamodel. These authors stated that when these boundaries become close to each other, a thin limit state margin is obtained and thus, the estimated failure probability can be considered as accurate. The proposed stopping criterion is given as follows:

$$\varepsilon_{P_f} = \frac{P_f^+ - P_f^-}{P_f^0} \leq \varepsilon_t \tag{12}$$

where P_f^0 is the original failure probability based on the Kriging prediction values $P(\hat{g}(x) \leq 0)$ and, P_f^+ and P_f^- are respectively the upper and lower boundaries of the failure probability defined as follows:

$$P_f^+ = P[(\hat{g}(x) + t. \sigma_{G_p}(x)) \leq 0] \tag{13}$$

$$P_f^- = P[(\hat{g}(x) - t. \sigma_{G_p}(x)) \leq 0] \tag{14}$$

where $\hat{g}(x) + t. \sigma_{G_p}(x) = 0$ and $\hat{g}(x) - t. \sigma_{G_p}(x) = 0$ are respectively the upper and lower boundaries of the limit state surface defined by $\hat{g}(x) = 0$, t is a constant ($t = 2$ in this paper) that sets the confidence level equal to $2 = \Phi^{-1}(97.7\%)$ and ε_t is a given tolerance taken as $\varepsilon_t = 10\%$ in this paper.

5. Probabilistic numerical results

The probabilistic numerical results begin with a simple example involving a non-linear analytical equation of the performance function. Only two random variables are considered in the analysis. The aim of this example is to visualize the evolution of the limit state surface during the enrichment process and to demonstrate the efficient computation of the failure probability when using the present AK-MCSm approach. Afterwards, AK-MCSm is applied to the case of the monopile embedded in a spatially varying soil.

5.1. Case of a simple example involving two random variables

The example considered in this section is based on the following non-linear analytical equation of the performance function:

$$G = 0.4 \times (u_1 - u_2)^2 - 0.4 \times (u_2 - 5)^3 - 10 \tag{15}$$

where u_1 and u_2 are two standard normal variables.

5.1.1. Evolution of the limit state surface during the enrichment process
In this section, only the enrichment process (i.e. training samples selection) was investigated. The same stopping condition used in AK-MCS method (i.e. $U > 2$) was adopted herein. The effect of considering the stopping criterion by [16] will be investigated in a subsequent section.

After the generation of 500,000 samples (where each sample consists of two standard Gaussian random variables), an initial Design of Experiment DoE of seven samples was randomly selected from the generated samples. A preliminary Kriging meta-model was then constructed on the basis of this DoE. The multi-enrichment process was performed using 4 clusters. It was shown to require three enrichment iterations resulting in 12 added samples before attaining the U-criterion (i.e. $U > 2$). A computation time of 3.28×10^2 sec was needed in this case. Fig. 5a, b and c present the evolution of the limit state surface during the three successive iterations of the enrichment process (i.e. from the initial DoE to DoE + 4 added samples, from DoE + 4 added samples to DoE + 8 added samples and from DoE + 8 added samples to DoE + 12 added samples, respectively). For each enrichment iteration, the candidate samples are presented in blue stars. These samples are classified in groups separated by blue lines plotted using the Voronoi Matlab function. For each group of samples, a corresponding centroid is presented in a red circle. The centroids of the four different groups are then added to the DoE of the metamodel. One can see that in each enrichment iteration, the metamodel is successively improved in the neighbourhood of the limit state surface by means of the four added samples distributed along the limit state surface.

Fig. 6 presents the evolution of the limit state surface LSS at the different stages of the Kriging meta-model construction. As may be seen from this figure, the LSS of the preliminary Kriging meta-model based on the DoE was very different from that of the true performance function presented in continuous red on the same figure. Notice however that the LSS corresponding to the obtained meta-model (in continuous black) is perfectly matched with that of the true performance function (in continuous red), especially in the zone where the probability density is not negligible (i.e. the zone that is close to the origin of the standard normal coordinate system).

In order to investigate the efficiency of the weighted K-w-means clustering technique with respect to the classical K-means clustering technique, the same computation presented above was performed using the classical technique. Notice that the conventional K-means clustering technique computes the geometric centroid of each cluster using its arithmetic mean determined as follows:

$$\boldsymbol{c}_k^{(i+1)} = \frac{1}{n_k} \sum_{j=1}^{n_k} \boldsymbol{x}_j \tag{16}$$

The calculation has resulted in 16 added samples instead of 12 added samples for the case of K-w-means clustering technique with an increase in the calculation time by 17%. This may be explained by the fact that the weighted K-w-means clustering technique is considered as more relevant than the classical K-means technique since it makes use of the information provided by the learning function U. Thus, it leads to a better choice of the centroids used in the meta-model enrichment process.

5.1.2. Effect of the stopping condition on the computation time
This section aims at presenting the benefits of the adopted stopping

(a)

(b)

(c)
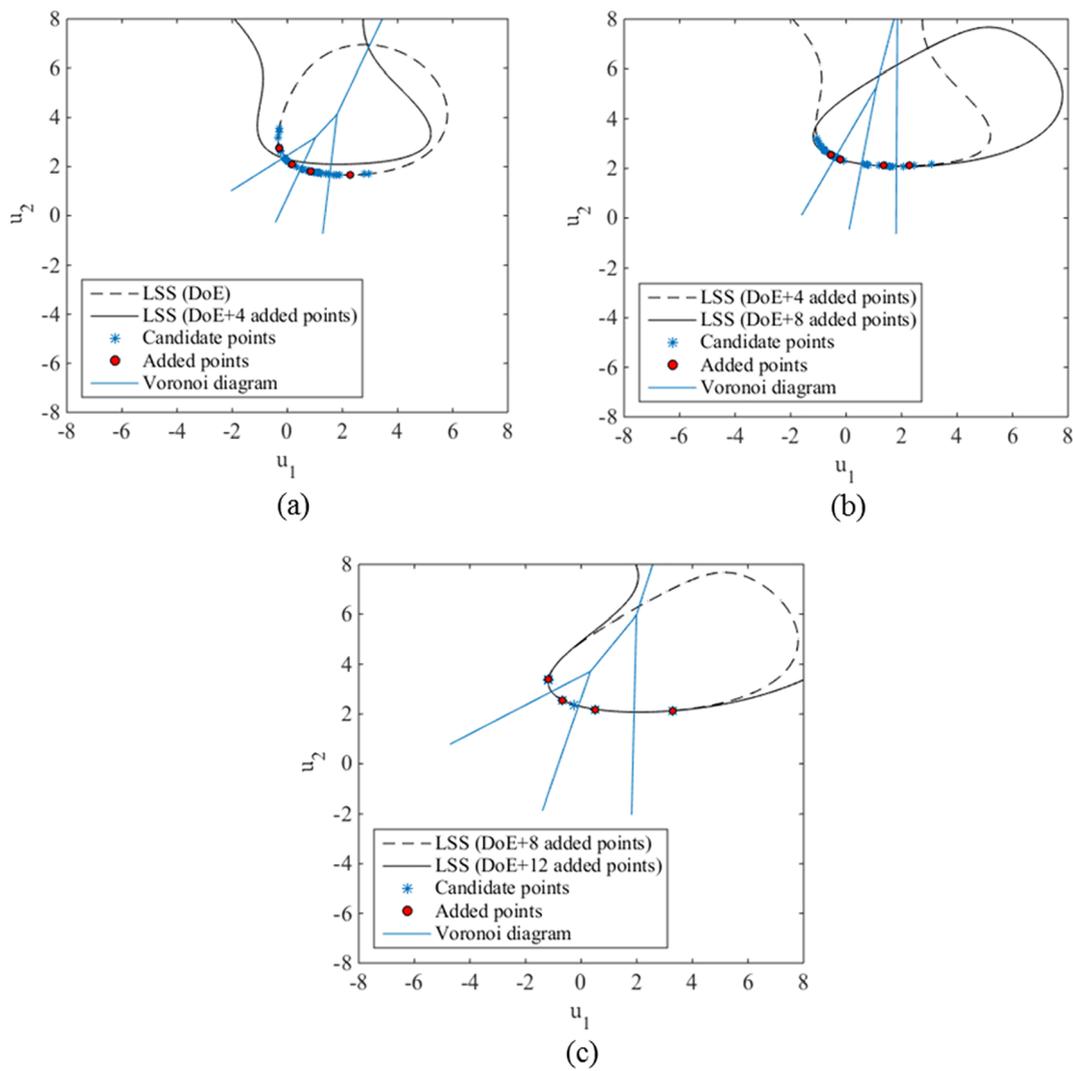
Fig. 5. Evolution of the limit state surface during the three successive iterations of the enrichment process (a) from the initial DoE to DoE + 4 added samples (b) from DoE + 4 added samples to DoE + 8 added samples (c) from DoE + 8 added samples to DoE + 12 added samples.
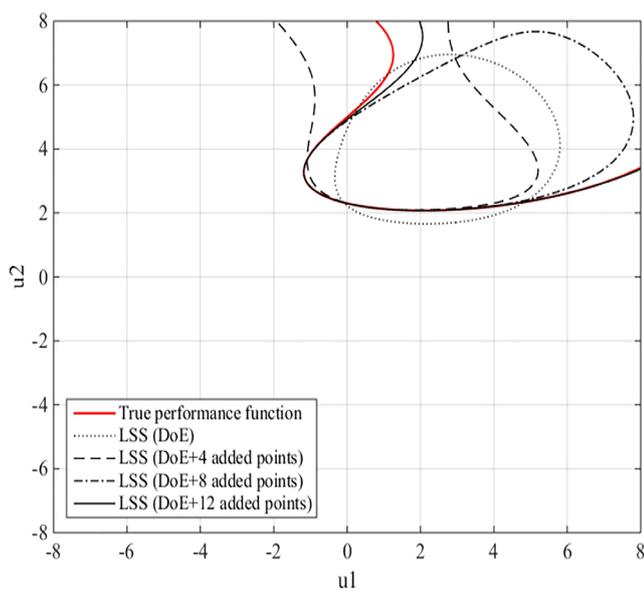
Fig. 6. Evolution of the limit state surface during the three iterations of the enrichment process.
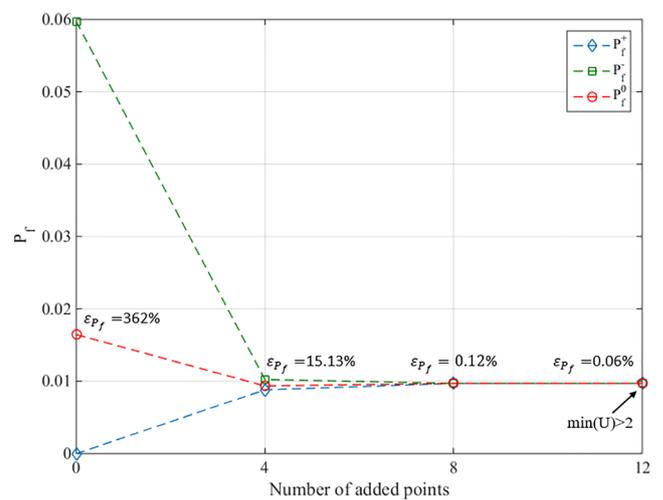
Fig. 7. Evolution of the failure probability with the number of added points.

criterion proposed by [16] concerning the number of added samples and the computation time. Fig. 7 presents the evolution of the failure probability $P_f = P_f^0$ with the number of added samples. The upper and lower boundaries of the failure probability (P_f^+ and P_f^- respectively) are

Table 1
Failure probability P_f and the corresponding value of the coefficient of variation $COV(P_f)$, together with the error ε_r with respect to the crude MCS methodology, the required number of evaluations and the computation time.

Method	$P_f \times 10^{-3}$	% COV (P_f)	$\varepsilon_r(\%)$	Nb. evaluations	Time (sec)
Crude MCS	9.694	1.42	–	5×10^5	–
AK-MCS	9.694	1.42	0	7 + 13	9.37×10^2
K-means + U	9.694	1.42	0	7 + 16	3.97×10^2
K-w-means + U	9.694	1.42	0	7 + 12	3.28×10^2
K-w-means + Schöbi	9.694	1.42	0	7 + 8	2.38×10^2

presented as well, together with the computed value of the error for each enrichment stage (i.e. each four added samples). This figure shows that the failure probability converges with an error of 0.12% (smaller than 10%) for 8 added samples when using the Schöbi criterion, whereas the *U*-criterion required 12 added samples. Thus, the use of Schöbi et al. [16] stopping criterion was shown to reduce the number of added samples from 12 to 8 and the computation time from 3.28×10^2 sec to 2. 38×10^2 sec.

5.1.3. Comparison with AK-MCS and the Crude MCS

This section aims at comparing the efficiency of the proposed AK-MCSm method with that of the classical Kriging-based method AK-MCS. Both methods are compared to the conventional crude Monte Carlo simulation in order to verify the accuracy of the results.

Table 1 presents for the different methods, the obtained value of the failure probability P_f together with the corresponding value of the coefficient of variation $COV(P_f)$, the error ε_r with respect to the crude MCS, the number of evaluations of the performance function and the computation time. The first two rows of this table present respectively the results obtained using the crude Monte Carlo Simulation methodology and the AK-MCS method by [11]. The AK-MCS was shown to provide the same result (i.e. the same obtained failure probability) as the crude MCS with an important reduction in the number of evaluations of the performance function. Indeed, only 7 evaluations corresponding to the initial DoE and 13 evaluations corresponding to the enrichment process were required by this method. The third row of the table corresponds to the application of the classical K-means clustering technique and the *U*-stopping criterion of the AK-MCS method. This method has led to the same failure probability value with a higher number of evaluations and a reduced computation time with respect to AK-MCS method (3.97×10^2 sec instead of 9.37×10^2 sec when using

AK-MCS). The fact of obtaining a reduced computation time may be explained by the fact that a multipoint enrichment is adopted herein where 4 training points are added simultaneously at each step of the enrichment process. The fourth row of the table corresponds to the application of the weighted K-w-means clustering technique by [15] and the same *U*-stopping criterion. By applying these techniques, the same value of the failure probability was also obtained with a reduction in the number of added points (12 added points were required) and consequently in the computation time. Finally, the fifth row presents the combined use of the K-w-means clustering technique and the stopping criterion proposed by [16], which corresponds to the adopted AK-MCSm method. The proposed method has led to an accurate value of the failure probability with a further reduction in the number of added points and the computation time with respect to the techniques mentioned previously. Thus, the present AK-MCSm method can be considered as an efficient probabilistic approach to be used for complex geotechnical problems involving a spatially varying soil.

5.2. Case of a monopile foundation embedded in a spatially varying soil

This section considers the effect of the soil spatial variability on the failure probability against exceeding a threshold value on the ultimate head rotation of a monopile embedded in a spatially varying clayey soil. Firstly, a reference case involving a vertical one-dimensional random field with a vertical autocorrelation distance $a_z = 2m$ is considered in the analysis. Secondly, the effect of the individual autocorrelation distance in the x, y or z direction on the failure probability is investigated. For both cases, an undrained soil cohesion that linearly increases with depth (as described before) is considered in the analysis.

5.2.1. Reference case

In this section, a reference case where the vertical autocorrelation distance was taken equal to 2 m was considered, the soil being assumed to be homogeneous in the horizontal direction. The number of random variables required by EOLE to accurately discretize the random field with a small variance of error ($< 5\%$) was found equal to 20. This number of random variables was adopted in the present paper.

Fig. 8(a) and (b) present respectively the evolution of the failure probability P_f and the corresponding value of the coefficient of variation $COV(P_f)$ with the number of added samples (with a step of 4 samples since the computation of the failure probability was carried out using 4 clusters). A failure probability of 1.274×10^{-3} was obtained with a corresponding coefficient of variation of 3.95% indicating a rigorous estimation of the failure probability. It should be noted that a computation time of 7.28 days was required for this calculation.
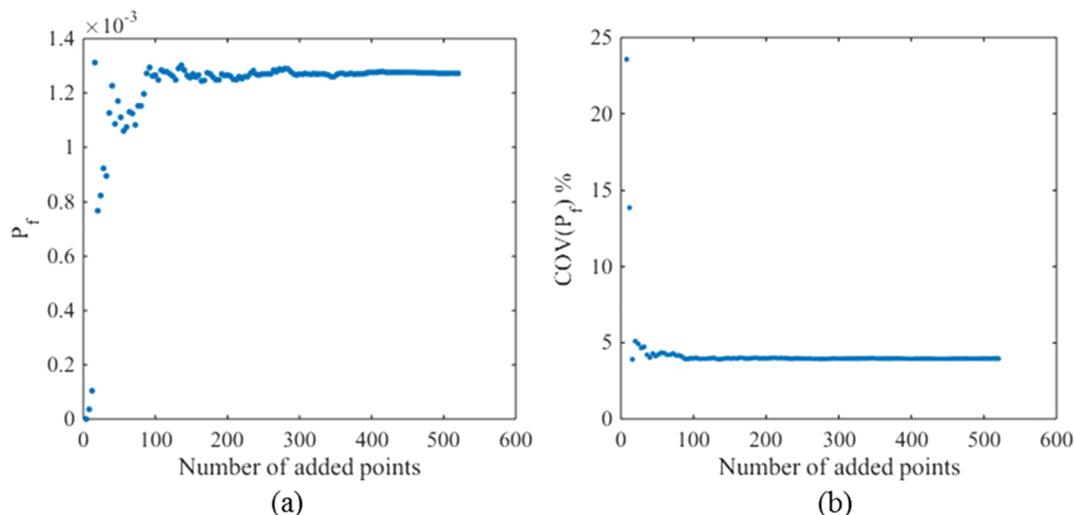
Fig. 8. Evolution of (a) the failure probability P_f and (b) the corresponding value of the coefficient of variation $COV(P_f)$, with the number of added samples.
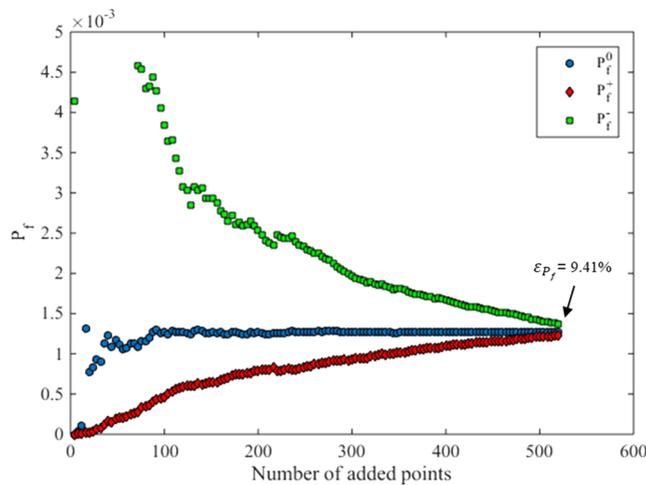
Fig. 9. Evolution of the original failure probability P_f^0 and its upper and lower boundaries P_f^+ and P_f^- with the number of added samples.

Table 2
Influence of t parameter on the number of added samples and the computation time.

t	$P_f \times 10^{-3}$	$\varepsilon_r(\%)$	Number of added samples	Time (days)
2	1.274	–	520	7.28
1.5	1.276	0.15	440	6.04
1	1.272	0.15	360	5.01
0.5	1.290	1.26	280	4.05
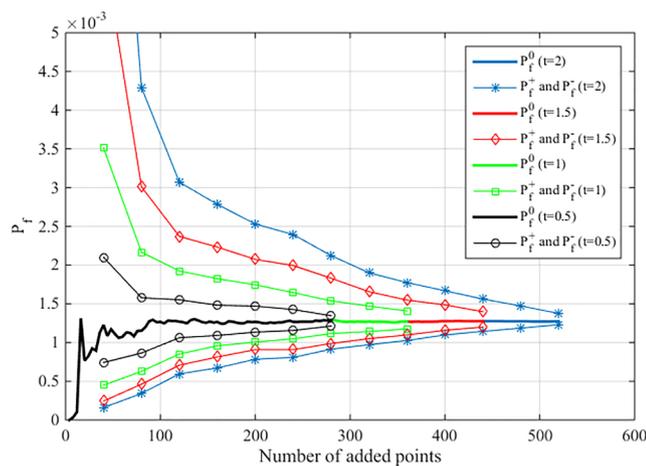
Fig. 10. Evolution of the original failure probability P_f^0 and its upper and lower boundaries P_f^+ and P_f^- with the number of added points, for various values of t.

In order to illustrate the stopping condition used in this paper, Fig. 9 shows the evolution of the upper and lower boundaries of the failure probability (P_f^+ and P_f^-, respectively) and the original failure probability P_f^0 with the number of added samples. From this figure, one may see that the enrichment process has stopped when P_f^+ and P_f^- converge towards the original failure probability P_f^0 within an error of 9.41% ($< 10\%$). Notice that at this stage, the minimum value of the U function is equal to 1.20 (i.e. smaller than the value of 2 imposed in AK-MCS approach), thus showing the efficiency of the adopted stopping criterion as compared to the U-criterion used in AK-MCS. Notice also that even when using the proposed stopping criterion, one can see from Fig. 9 that the probability of failure P_f^0 has already stabilized for a much smaller number of added samples. This can be explained by the fact that severe conditions have been adopted in this paper for the parameters

Table 3
Results of AK-MCS and AK-MCSm approaches.

Method	$P_f \times 10^{-3}$	Number of added samples	Time (days)
AK-MCS	1.274	447	27.89
AK-MCSm (4 clusters)	1.274	360	5.01

employed in the stopping criterion. Indeed, a smaller value of the constant t (i.e. a smaller confidence level) may lead to a faster convergence of the upper and lower boundaries of the failure probability and thus, to an earlier stopping of the enrichment process. For this purpose, the effect of the parameter t on the number of added samples was investigated (see Table 2 and Fig. 10).

From Table 2, one can observe that when t decreases, the relative error on the P_f value (with respect to the case of $t = 2$) increases. This relative error ramains lower than 1% when reaching $t = 1$. Thus, a value of $t = 1$ may be a reasonable choice in the present case. Such a choice reduces the number of added samples to 360 (instead of 520 in the case of $t = 2$) and the computation time to 5.01 days (instead of 7.28 days in the case of $t = 2$).

5.2.1.1. Comparison with AK-MCS approach. The efficiency of the proposed approach (as applied to the problem of a monopile foundation embedded in a spatially varying soil) was compared to that of AK-MCS approach. For this purpose, AK-MCS approach was applied on the same problem in order to allow the comparison. Four clusters were considered when dealing with AK-MCSm approach.

Table 3 presents the results of AK-MCS approach and those of the proposed AK-MCSm method. As may be seen from this table, the different approaches result in the same value of the failure probability. Concerning the computation cost, the proposed method leads to a significant reduction in the computation time as compared to AK-MCS (5.01 days for AK-MCSm instead of 27.89 days for AK-MCS).

5.2.2. Effect of the autocorrelation distances a_x, a_y and a_z on P_f

Fig. 11 presents the evolution of the failure probability with a_x, a_y and a_z where a_x and a_y are the horizontal autocorrelation distances in the directions orthogonal and parallel to the loading plane respectively and a_z is the vertical autocorrelation distance. The aim of this figure is to identify (for the present embedded length and diameter of the monopile) the values of the autocorrelation distances for which the soil spatial variability has a significant effect on the computed failure probability. For all the treated configurations, the number of random variables was chosen in such a manner that the variance of the error of EOLE methodology is smaller than 5%. For this threshold, a maximal number of random variables of 24 was found necessary for the configuration where $a_y = 2m$.

For all the three curves of Fig. 11, the failure probability increases with the different autocorrelation distances a_x, a_y or a_z. The rates of increase of the failure probability largely decrease beyond given values of the autocorrelation distances [i.e. beyond an autocorrelation distance of 24 m ($\approx$ L) for a_z, 10 m ($\approx$ 2.5D) for a_y and 6 m ($\approx$ 1.5D) for a_x] to attain a constant value of the failure probability corresponding to the case of a homogeneous soil (i.e. with no spatial variability). The failure probability corresponding to the homogeneous case was found equal to 0.12. It should be emphasized here that all the three identified distances (i.e., L, 2.5D and 1.5D) are directly related to the soil displacement field obtained due to the ultimate loading (i.e. they are related to the mechanics of the problem and not connected to the input probabilistic data). Indeed, the deformed soil at the ultimate limit state (i.e. at the ultimate rotation) is concentrated along the embedded length L of the monopile in the vertical direction, and it is limited to 2.5D and 1.5D in the y- and x- horizontal directions as it may be seen from Fig. 12a and 12b, respectively. These observations explain the reason why the increase in the autocorrelation distances a_x, a_y and a_z beyond the
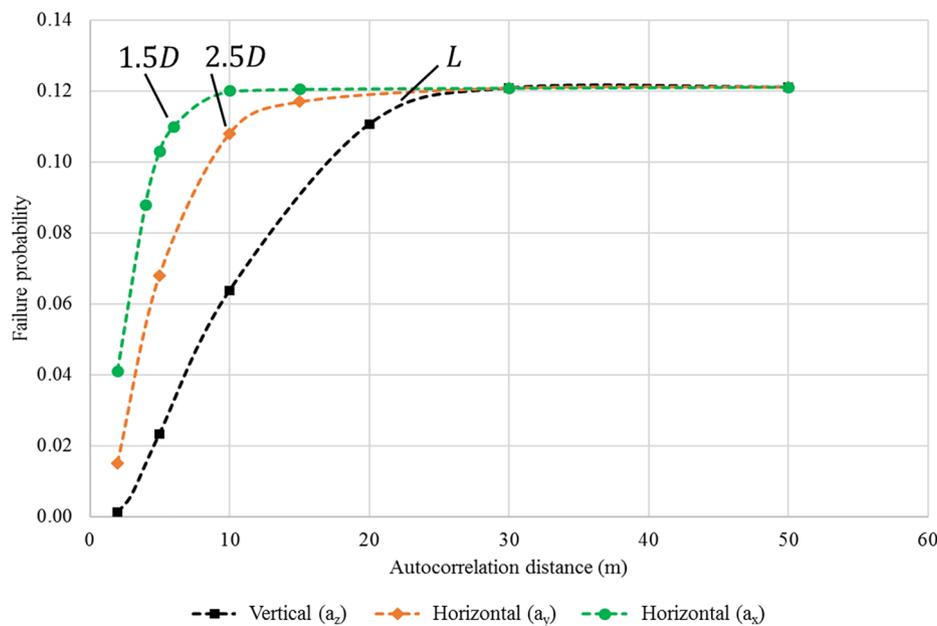
Fig. 11. Effect of the autocorrelation distances a_x, a_y and a_z on P_f.

distances 1.5D, 2.5D and L respectively does not significantly change the value of the failure probability. A reduction in the failure probability with respect to the case of a homogeneous soil may be obtained if at least one of the autocorrelation distances a_x, a_y or a_z is smaller than its corresponding limit value (i.e. 1.5D, 2.5D, L) since the monopile head rotation will be affected by the soil spatial variability in this case, thus inducing a change in the failure probability.

Fig. 11 shows that the rate of increase in the failure probability (for the small and moderate values of the autocorrelation distances) is the highest when considering the horizontal autocorrelation distance a_x. This rate of increase in the failure probability is smaller for the horizontal autocorrelation length a_y and it becomes the smallest one for the vertical autocorrelation distance a_z. These observations indicate that a reduction in the failure probability with respect to the case of a homogeneous soil (i.e. the case of no spatial variability) may be foreseen by mostly considering the soil spatial variability in the vertical direction. This is because (i) the vertical limit value L is bigger than the horizontal limit values of 1.5D and 2.5D and (ii) the vertical autocorrelation distance is generally one order of magnitude smaller than the horizontal one. This finding is interesting since it allows one to simplify a complex 3D stochastic problem into a 1D vertical stochastic problem with a much lower computational burden for the cases where the horizontal autocorrelation distances are greater than 2.5D. Another advantage from this simplification is that the identification of the soil

spatial variability in the vertical direction is easier and cheaper than that in the horizontal direction (e.g. by using a CPT test).

6. Conclusions

A probabilistic analysis at the ultimate limit state (ULS) was performed in this paper for an offshore monopile foundation embedded in a spatially varying clayey soil. The soil undrained cohesion was considered as a random field following a lognormal distribution and the soil undrained Young modulus was assumed to be linearly related to the soil undrained cohesion.

A Kriging-based probabilistic approach using a multipoint enrichment technique was adopted. An improved clustering technique proposed by Lelièvre et al. [15] was used for learning. Compared to the classical AK-MCS Kriging-based approach, the present AK-MCSm method can consider simultaneous mechanical computations of different samples at each iteration of the enrichment process. Furthermore, the selected samples are better chosen due to the use of the information provided by the learning function. Notice also that this paper makes use of a relevant stopping condition recently proposed by Schöbi et al. [16]. The method was shown to be efficient in terms of the computation time with respect to AK-MCS approach.

A reduction in the failure probability with respect to the case of a homogeneous soil may be obtained if at least one of the autocorrelation
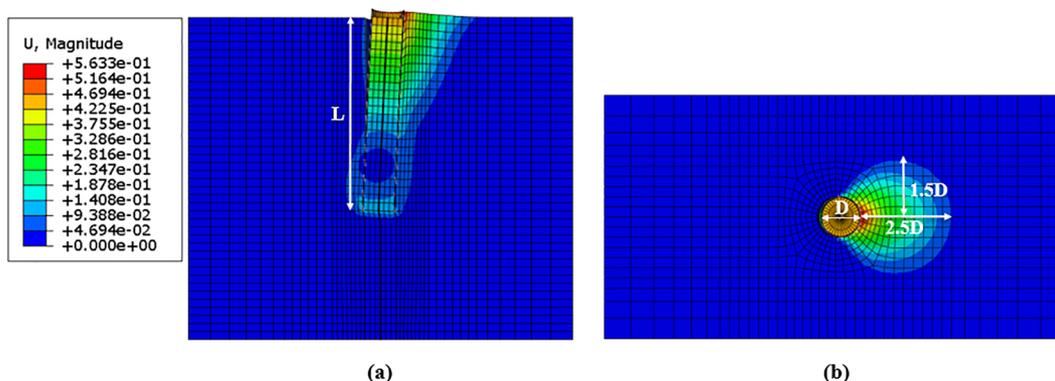
Fig. 12. Deformed soil at the ultimate limit state (a) in the yz plane (b) in the xy plane.

distances a_x, a_y or a_z is smaller than its corresponding limit value (i.e. 1.5D, 2.5D, L). From a practical point of view, a reduction in the failure probability with respect to the case of a homogeneous soil (i.e. the case of no spatial variability) may be foreseen by mostly considering the soil spatial variability in the vertical direction.

Funding

CRediT authorship contribution statement

Abdul-Kader EL AHAJ: Conceptualization, Investigation, Methodology, Project administration, Software, Visualization, Writing - original draft, Writing - review editing. **Abdul-Hamid Soubra:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing - review editing.

Declaration of Competing Interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Acknowledgements

References

[1] Paice GM, Griffiths DV, Fenton GA. Finite element modeling of settlements on spatially random soil. J Geotechnical Eng 1996;122.

[2] Griffiths DV, Fenton GA. Probabilistic slope stability analysis by finite elements. J Geotech Geoenviron Eng 2004;130:507–18.

[3] Hicks MA, Spencer WA. Influence of heterogeneity on the reliability and failure of a long 3D slope. Comput Geotech 2010;37:948–55.

[4] Li YJ, Hicks MA, Nuttall JD. Comparative analyses of slope reliability in 3D. Eng Geol 2015;196:12–23.

[5] Hicks MA, Li Y. Influence of length effect on embankment slope reliability in 3D. Int J Numer Anal Meth Geomech 2018;42:891–915.

[6] Andersen LV, Vahdatirad MJ, Sichani MT, Sørensen JD. Natural frequencies of wind turbines on monopile foundations in clayey soils — A probabilistic approach. Comput Geotech 2012;43:1–11.

[7] Li D, Xiao T, Cao ZJ, Zhou CB, Zhang LM. Enhancement of random finite element method in reliability analysis and risk assessment of soil slopes using Subset Simulation. Landslides 2016;13:293–303.

[8] Jiang SH, Huang JS. Efficient slope reliability analysis at low-probability levels in spatially variable soils. Comput Geotech 2016;75:18–27.

[9] Huang J, Fenton G, Griffiths DV, Li DQ, Zho CB. On the efficient estimation of small failure probability in slopes. Landslides 2017;14:491–8.

[10] Bourinet JM, Deheeger F, Lemaire M. Assessing small failure probabilities by combined subset simulation and Support Vector Machines. Struct Saf 2011;33:343–53.

[11] Echard B, Gayton N, Lemaire M. AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation. Struct Saf 2011;33:145–54.

[12] Papadopoulos V, Giovanis DG, Lagaros ND, Papadrakakis M. Accelerated subset simulation with neural networks for reliability analysis. Comput Methods Appl Mech Eng 2012;223–224:70–80.

[13] Balesdent M, Morio J, Marzat J. Kriging-based adaptive Importance Sampling algorithms for rare event estimation. Struct Saf 2013;44:1–10.

[14] Al-bittar T, Soubra AH, Thajeel J. Kriging-based reliability analysis of strip footings resting on spatially varying soils. J Geotech Geoenviron Eng 2018;144.

[15] Lelièvre N, Beaurepaire P, Mattrand C, Gayton N. AK-MCSi : A Kriging-based method to deal with small failure probabilities and time-consuming models. Struct Saf 2018;73:1–11.

[16] Schöbi R, Sudret B, Marelli S. Rare event estimation using polynomial-Chaos Kriging. ASCE-ASME J Risk Uncertainty Eng Syst, Part A: Civil Eng 2017;3:1–12.

[17] El Haj A-K, Soubra AH, Fajoui J. Probabilistic analysis of an offshore monopile foundation taking into account the soil spatial variability. Comput Geotech 2019;106:205–16.

[18] DNV-OS-J101. Design of offshore wind turbine structures; 2014.

[19] ABAQUS. Dassault Systèmes Simulia Corp, Providence; 2016.

[20] Kellezi L, Hansen PB. Static and dynamic analysis of an offshore mono-pile windmill foundation. BGA international conference on foundations: innovations, observations, design and practice, Dundee; 2003.

[21] Wroth C, Houlsby G. Soil mechanics – property characterization and analysis procedures. Keynote lecture. Proceedings of the eleventh international conference on soil mechanics and foundation engineering, San Francisco, 12-16 August 1985. p. 1–55.

[22] Skempton AW. Discussion: The planning and design of the new Hong Kong airport. Proc Inst Civ Eng 1957;7:305–7.

[23] Dean ETR. Offshore geotechnical engineering: Principles and practice. London: Thomas Telford; 2010.

[24] Semple RM, Gemeinhardt JP. Stress history approach to analysis of soil resistance to pile driving. In: Proceedings of the 13th offshore technology conference, Houston, Texas, vol. 1; 1981. p. 165–72.

[25] Menzies D, Roper R. Comparison of jackup rig spudcan penetration methods in clay. Offshore Technology Conference, Houston, Texas, USA; 2008.

[26] USACE. Engineering and design : settlement analysis. Engineer manual 1110–1–1904, US Army Corps of Engineers; 1990.

[27] Li CC, Der Kiureghian A. Optimal discretization of random fields. J Eng Mech 1993;119:1136–54.

[28] Nataf A. Détermination des distributions de probabilités dont les marges sont données. Comptes Rendus de l'Académie Des Sciences 1962;225:42–3.

[29] Sudret B, Der Kiureghian A. Stochastic finite element methods and reliability: A state-of-the-art report; 2000.

[30] Sacks J, Welch WJ, Mitchell TJ, Wynn HP. Design and analysis of computer experiments. Stat Sci 1989;4:409–23.

[31] Lophaven SN, Nielsen HB, Søndergaard J. DACE: A Matlab Kriging Toolbox Version 2.0. Informatics and mathematical modelling, technical report IMM-TR-2002-12, Technical University of Denmark, DTU; 2002.

[32] MacQueen J. Some methods for classification and analysis of multivariate observations. Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. Berkeley: University of California Press; 1967. p. 281–97.

[33] Zaki MJ, Meira W. Data mining and analysis: fundamental concepts and algorithms. New York: Cambridge University Press; 2014.

[34] Aurenhammer F. Voronoi diagrams – A survey of a fundamental geometric data structure. ACM Comput Surv 1991;23:345–405.