Application of improved Kriging-based approaches to the analysis of monopile foundations

Abdul-Kader El Haj
Ph.D. Student, University of Nantes, Saint-Nazaire, France

Abdul-Hamid Soubra
Professor, University of Nantes, Saint-Nazaire, France

ABSTRACT: In this paper, a probabilistic analysis of an offshore monopile foundation embedded in a spatially varying clayey soil was performed. The aim is to compute the failure probability P_f against exceeding a threshold value on the monopile head rotation. A multipoint enrichment technique within a Kriging-based approach was used for the probabilistic analysis. An improved K-means clustering technique was employed. The number of training points used in the enrichment process was determined based on the closeness of the estimated failure probability to its upper and lower confidence values. Some probabilistic numerical results are presented and discussed.

The probabilistic analysis of geotechnical structures involving spatially varying soil properties has been performed for several years using Monte Carlo Simulation (MCS) methodology [e.g. Griffiths and Fenton (2004)]. This method is known to be time-consuming especially when dealing with the small practical values of the failure probability.

In order to reduce the computation time with respect to MCS, the Active learning method by Echard et al. (2011) combining Kriging and Monte Carlo simulation (called AK-MCS) was recently employed by Al-bittar et al. (2018) for the probabilistic analysis of strip footings resting on spatially varying soils. The AK-MCS approach consists in replacing the time-consuming mechanical model by a simple Kriging meta-model calibrated by a limited number of mechanical model evaluations making use of an adaptive learning technique. The aim is to apply MCS methodology on the calibrated metamodel (called also surrogate model) with a quasi negligible computational time.

Within AK-MCS approach, a preliminary surrogate model is constructed by Kriging metamodeling using a small design of experiments. The obtained approximate meta-model is then successively improved through an enrichment process in which a powerful learning function is employed for the selection of the 'best' samples to be evaluated by the computationally expensive mechanical model. The best sample is the one with the highest probability of misclassification [see Echard et al. (2011)]. Notice that in AK-MCS method, a single sample is selected per iteration of the enrichment process. This is a drawback in the case where distributed (or parallel) computing facilities are to be used to reduce the computation time.

In this paper, a multipoint enrichment technique proposed by Lelièvre et al. (2018) is used. The aim is to allow several evaluations of the performance function to be carried out simultaneously. This approach is based on a relevant clustering technique that makes use of the learning function U employed by Echard et al. (2011). It has the advantage of considering the information provided by the U function in order to obtain the optimal samples for the enrichment process. The resulting technique is named K-weighted-means clustering algorithm (K-w-means).

Concerning the stopping condition on learning used in this paper, the present probabilistic method makes use of a criterion that was recently proposed by Schöbi et al. (2017).

This criterion is more relevant than the one used in AK-MCS approach because it is based on the convergence of the quantity of interest (i.e. the failure probability).

This paper aims at applying the above mentioned probabilistic techniques to the case of a spatially varying soil. The objective is to compute the failure probability P_f at the ultimate limit state of an offshore monopile foundation embedded in a spatially varying clayey soil. A prescribed threshold value on the monopile head rotation was considered in the analysis.

1. MONOPILE MECHANICAL MODEL

The mechanical model of the 3D soil-monopile system has been carried out using the commercial finite element software Abaqus/Standard.

An open-ended steel monopile of diameter *D*=4m was considered in this study. 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. The steel monopile material with a density of $7840\ \mathrm{kg/m^3}$ was assumed to be linear elastic 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 assumed to follow the elastic-perfectly plastic Tresca constitutive model. In this paper, the soil was assumed to have a submerged unit weight of $7\ \mathrm{kN/m^3}$ and a Poisson's ratio of 0.495. The undrained cohesion was supposed to vary linearly with depth as given by the following equation:

$$c_u = c_{u,m} + k_{cu}.\sigma'_{v0} \qquad (1)$$

where $c_{u,m}$ is the value of the undrained cohesion at mudline (taken here equal to 2 kPa), k_{cu} is a material constant for the clay (taken here equal to 0.23) and σ'_{v0} is the effective vertical overburden stress. Note here that the soil undrained Young modulus was assumed to be linearly related to the soil undrained cohesion such that $E_u = K_c \times c_u$ where K_c is a correlation factor taken equal to 500 in this paper.

Figure 1 shows the soil domain and the mesh used in the analysis. The soil mesh was constructed using C3D8 and C3D6 linear brick elements. Incompatible mode linear brick elements (C3D8I) were used for the monopile.

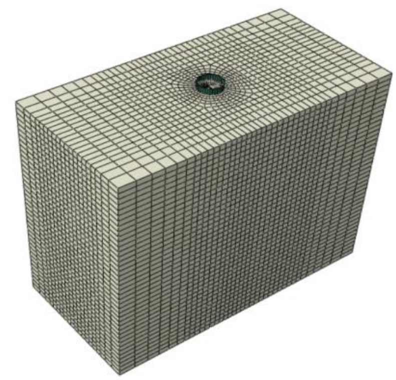

Figure 1: Soil-monopile numerical model

Surface-to-surface master/slave contact formulation was used to model the interaction between the monopile and the soil. The monopile was selected as the master surface while the soil in contact with the monopile was taken as the slave surface. The frictional behavior was modelled using Coulomb friction law where the friction coefficient μ was taken equal to 0.24.

The numerical simulation was executed step-wised. A geostatic step was first performed for the generation of the initial stress state of the soil in the whole model consisting of soil elements only. In a second step, the monopile was simulated by (i) removing the soil elements located at the monopile position and generating the steel elements representing the monopile, (ii) activating the contact conditions between the monopile and the soil and (iii) applying the weight of the generated monopile. Finally, in a third step, the horizontal and vertical forces (H and V) and the corresponding moment M are applied in increments at a reference point (taken here at the top of the monopile) where the applied moment was equal to $\mathrm{M} = \mathrm{H} \times (\mathrm{h} - 1)$, h being the vertical distance between the applied horizontal force and the mudline.

The monopile head rotation value corresponding to the ultimate limit state was determined by applying the tangent intersection method on the moment-rotation curve (see Figure 2) as obtained from the numerical simulation. From this figure, one may observe that the limit rotation corresponding to the ultimate moment is equal to 1.5°, the corresponding horizontal force being equal to about 2 MN. The obtained value of the monopile head rotation is used later in this paper as a threshold value for the probabilistic analysis.

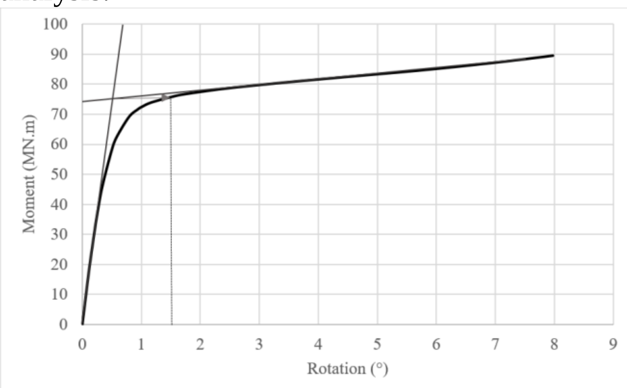

Figure 2: Moment-rotation curve of the monopile at mud-line level

2. NUMERICAL MODELING OF THE SOIL SPATIAL VARIABILITY

The soil undrained cohesion c_u was considered as a random field. It was asuumed to follow a lognormal distribution with a constant coefficient of variation of 25%, which is in accordance with the investigations by Lacasse and Nadim (1996) on the variability of seabed soils. The mean values of the soil undrained cohesion are those of the deterministic analysis provided in the preceding section. Concerning the autocorrelation 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:

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

where a_x and a_y are the horizontal autocorrelation distances and a_z is the vertical autocorrelation distance.

Remember here that the soil undrained Young modulus was assumed in this paper to be linearly related to the soil undrained cohesion such that $E_u = 500 \times c_u$. Thus, the soil undrained Young modulus was implicitly considered as a random field having the same distribution as the soil undrained cohesion.

Notice that the discretization of the cohesion random field was performed using EOLE method proposed by Li and Der Kiureghian (1993). Notice also that the discretization of a random field by EOLE leads to an expression that provides the value of this random field at each point of the soil mass as a function of M standard Gaussian random variables (this number M is equal to the number of eigenmodes). For more details on the discretisation of a log-normal random field by EOLE, the reader may refer to Al-bittar and Soubra (2014). Notice finally that the realizations of the Young modulus random field can be easily obtained from the realizations of the cohesion random field by multiplying the values of the soil cohesion by 500.

3. PROBABILISTIC MODEL

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

$$G = \frac{\theta_{ULS}}{\theta} - 1 \quad (3)$$

where $\theta_{ULS} = 1.5°$ is the monopile head rotation at the Ultimate Limit State (ULS) as was determined before and θ is the rotation corresponding to typical realizations of c_u and E_u.

The present probabilistic procedure consists of two main stages. First, a preliminary approximate kriging meta-model based on a small number of samples is generated. Second, the obtained approximate kriging meta-model is successively improved *via* an enrichment process (by adding each time new training samples) until reaching a sufficiently accurate meta-model for the computation of the failure probability. These two stages are described in more details in the next two subsections in the present case of a spatially varying soil.

3.1. Construction of a preliminary Kriging metamodel

The procedure begins with a generation by MCS methodology of 5×10^5 samples $\boldsymbol{x}^{(i)}(i = 1,2,\dots,5 \times 10^5)$. Each sample $\boldsymbol{x}^{(i)}$ consists of M standard Gaussian random variables where M is the number of random variables required by EOLE methodology. Afterwards, a small design of experiment DoE (taken equal to 15 samples) is randomly selected from the generated population. Each sample of the DoE is then transformed (using EOLE) into a realization of c_u and a corresponding realization of E_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. Lophaven et al. (2002)], an approximate Kriging meta-model may be constructed in the standard space of random variables based on the DoE and the corresponding performance function values. This meta-model may be used to compute the MCS failure probability $\hat{P}_f$ given by:

$$\hat{P}_f = \sum_{i=1}^{N_{MCS}} I\left(G_p(x^{(i)})\right) / N_{MCS} \tag{4}$$

The meta-model random responses $G_p(x^{(i)})$ in this equation are replaced by the mean prediction values $\hat{g}(x^{(i)})$. Notice also that N_{MCS} in equation (4) 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 $COV\ (\hat{P}_f)$ is given by the following equation:

$$COV(\hat{P}_f) = \sqrt{\frac{1 - \hat{P}_f}{\hat{P}_f \cdot N_{MCS}}} \tag{5}$$

It should be noted that the value of the failure probability and the corresponding value of the coefficient of variation computed during this stage are not sufficiently accurate because of the very small number of samples (DoE) used so far. Thus, an enrichment process is needed.

3.2. Enrichment process

The enrichment process is done *via* an active learning technique. The learning phase stops once the metamodel is sufficiently improved, 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.

3.2.1. Selection of new training points

The enrichment process of the AK-MCS method is performed using the learning function U defined by the following equation:

$$U(x^{(i)}) = \frac{\left|\hat{g}(x^{(i)})\right|}{\sigma_{G_p}(x^{(i)})} \tag{6}$$

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. It should be emphasized 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 making use of a clustering technique.

The conventional k-means clustering technique aims at finding the geometric centroid of each cluster using its arithmetic mean. 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 (Zaki and Meira, 2014). In this algorithm, larger weights are dedicated to the samples with high information values according to the learning function.

Lelièvre et al. (2018) 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. It 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 and thus, the centroid obtained for each cluster will be the optimal one for the enrichment. In other words, the selected samples will be situated in the uncertain zone all along the limit state surface leading to an efficient multipoint enrichment of the kriging metamodel. This approach is used in this paper for the probabilistic analysis.

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

1. Assume that K is 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 $[\boldsymbol{c}_1^{(1)}, \boldsymbol{c}_2^{(1)}, \dots, \boldsymbol{c}_K^{(1)}]$.
3. Split the samples into K sets according to Voronoi diagram (Aurenhammer, 1991) depending on the nearest centroid.
4. Determine the centroid $\boldsymbol{c}_k$ of each cluster k $(k = 1,2, \dots, K)$ by computing the corresponding weighted mean as follows:
$$\boldsymbol{c}_k^{(i+1)} = \frac{\sum_{j=1}^{n_k} \left(\frac{1}{U_j}\right)^2 \boldsymbol{x}_j}{\sum_{j=1}^{n_k} \left(\frac{1}{U_j}\right)^2} \tag{7}$$
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 points in the k^{th} cluster and $[\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_{n_k}]$ is the set of samples corresponding to this cluster.
5. Calculate the error expressing the sum of the squared distances between each couple of successive centroids, as follows:
$$\sum_{k=1}^{K} \left(\boldsymbol{c}_k^{(i)} - \boldsymbol{c}_k^{(i+1)}\right)^2 \tag{8}$$
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.

3.2.2. Stopping condition

In AK-MCS method, the enrichment process stops when the learning function *U* is sufficiently large for all the MCS samples. A minimum value of *U*=2 is adopted on these samples, which corresponds to a probability of a wrong sign that is lower than 0.0228. 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 Schöbi et al. (2017). 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. (2017) 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. They 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:
$$\frac{P_f^+ - P_f^-}{P_f^0} \leq \varepsilon_{P_f} \tag{9}$$
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:

$$P_f^+ = P(\hat{g}(x) + t.\sigma_{G_p}(x) \leq 0) \quad (10)$$

$$P_f^- = P(\hat{g}(x) - t.\sigma_{G_p}(x) \leq 0) \quad (11)$$

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_{P_f} is a given tolerance taken as $\varepsilon_{P_f} = 10\%$ in this paper.

4. NUMERICAL RESULTS

The monopile was considered to be subjected to a horizontal load $\mathrm{H} = 1{,}6\ MN$ acting at a height h (supposed equal to 38.6 m above the sea bed level) resulting in an additional moment at mudline of $M = H \times h$. A vertical load V of $2\ MN$ representing the structure weight was also considered in the analysis. Notice that the applied loads $(\mathrm{H}, \mathrm{V}, \mathrm{M})$ adopted in this study induces a deterministic value of the rotation at mudline of nearly 0.55°.

In this paper, only the vertical autocorrelation distance was considered in the analysis, the horizontal variability being generally less significant than the one in the vertical direction. The value of the vertical autocorrelation distance used in this paper is equal to 2m. 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.

Figures 3 and 4 present the evolution of the failure probability and the corresponding coefficient of variation with the number of the added samples as obtained from the proposed method (case of 2 clusters). A failure probability of 1.274×10^{-3} is obtained with a corresponding coefficient of variation of 3.95% indicating a rigorous estimation of the failure probability.

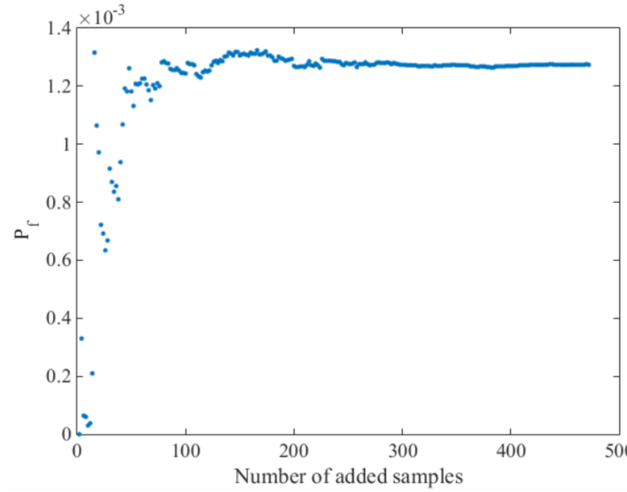

Figure 3: Failure probability vs number of added samples

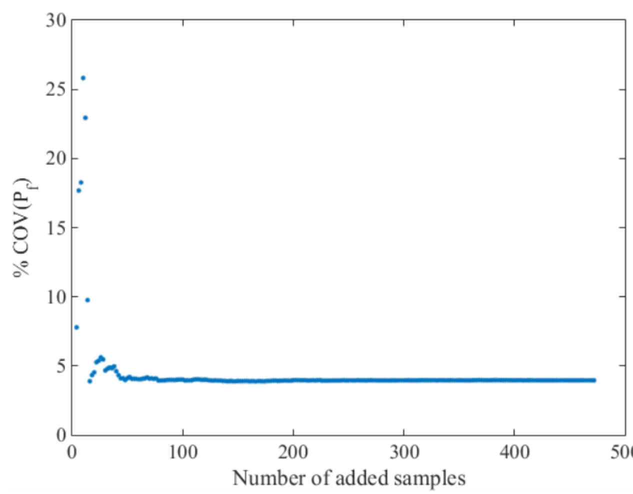

Figure 4: Coefficient of variation of the failure probability vs number of added samples

Figure (5) shows the evolution of the upper and lower boundaries of the failure probability (P_f^+ and P_f^-, respectively) 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 the minimum value of the U function is equal to 1.20 (<2) at this stage thus showing the efficiency of the adopted stopping criterion as compared to the U criterion. Notice also that even when using the proposed stopping criterion, one can see that the probability of failure 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 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.

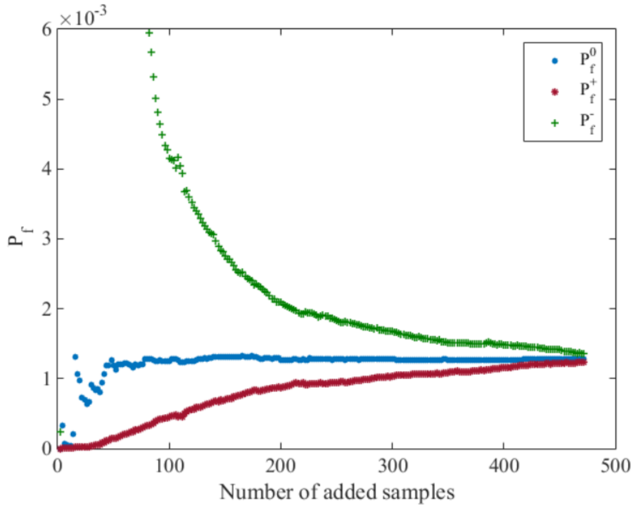

Figure 5: Upper and lower boundaries of the failure probability vs number of added samples

4.1. Comparison with AK-MCS method

This section aims at comparing the efficiency of the proposed approach (as applied to the problem of a monopile foundation embedded in a spatially varying soil) with respect to AK-MCS approach. For this purpose, AK-MCS approach was applied on the same problem in order to allow the comparison.

Table 1: Results of the different approaches

Method	$P_f \times 10^{-3}$	*Nb. Added points*	*Time (days)*
AK-MCS	*1.274*	*447*	*27.89*
AK-MCSm + 2 clusters	*1.274*	*472*	*24.49*
AK-MCSm + 4 clusters	*1.274*	*520*	*7.28*

Table 1 presents the results of AK-MCS approach and those of the proposed approach denoted herein as AK-MCSm (where m stands for multipoint enrichment).

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 reduction in the computation time as compared to AK-MCS, the reduction being more significant when increasing the number of clusters. It should be noted that obtaining a larger number of added points when using the proposed method is not surprising. This may be explained by the fact that the present method is based on a multipoint enrichment process.

5. CONCLUSIONS

This paper presents a probabilistic analysis at the ultimate limit state of an offshore monopile foundation embedded in a spatially varying soil. The finite element model of the monopile foundation being time-consuming, a cost-effective probabilistic approach was carried out.
A Kriging-based approach combined with a multipoint enrichment technique was adopted in this paper. The proposed approach makes use of an improved clustering technique proposed by Lelièvre et al. (2018) for learning. Also, a relevant stopping condition proposed by Schöbi et al. (2017) was employed. The resulting method was applied to the case of random fields and used to perform the reliability analysis.
The applied method was shown to be efficient with respect to the classical Kriging-based approach, namely AK-MCS approach.

ACKNOWLEDGEMENTS

This work was carried out within the framework of the WEAMEC, WEst Atlantic Marine Energy Community, and with funding from the CARENE, Communauté d'Agglomération de la Région Nazairienne et de l'Estuaire.

6. REFERENCES

Al-bittar, T. and Soubra, A.H. (2014). Probabilistic Analysis of Strip Footings Resting on Spatially

Varying Soils and Subjected to Vertical or Inclined Loads. *Journal of Geotechnical and Geoenvironmental Engineering*, 140(4).

Al-bittar, T., Soubra, A.H. and Thajeel, J. (2018). Kriging-based reliability analysis of strip footings resting on spatially varying soils. *Journal of Geotechnical and Geoenvironmental Engineering*.

Aurenhammer, F. (1991). A survey of a fundamental geometric data structure. *ACM Computing Surveys*, 23(3).

Echard, B., Gayton, N. and Lemaire, M. (2011). AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation. *Structural Safety*, 33(2), pp.145–154.

Griffiths, D.V. and Fenton, G.A. (2004). Probabilistic slope stability analysis by finite elements. *Journal of Geotechnical and Geoenvironmental Engineering*, 130(5), pp.507–518.

Lacasse, S. and Nadim, F. (1996). *Proceedings of Uncertainty in the geologic environment: From theory to practice*, Geotechnical Special Publication (GSP 58), ASCE, Edited by CD Shackelford, PP Nelson & MJS Roth, pp. 49–75.

Lelièvre, N. et al. (2018). AK-MCSi : A Kriging-based method to deal with small failure probabilities and time-consuming models. *Structural Safety*, 73, pp.1–11.

Li, C.C. and Der Kiureghian, A. (1993). Optimal discretization of random fields. *Journal of Engineering Mechanics*, 119(6), pp.1136–1154.

Lophaven, S.N., Nielsen, H.B. and Søndergaard, J. (2002). *DACE: A Matlab Kriging Toolbox Version 2.0*, Informatics and Mathematical Modelling, Technical report IMM-TR-2002-12, Technical University of Denmark, DTU.

Schöbi, R., Sudret, B. and Marelli, S. (2017). Rare Event Estimation Using Polynomial-Chaos Kriging. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 3(2), pp.1–12.

Zaki, M. and Meira, W. (2014). *Data mining and analysis: Fundamental concepts and algorithms*, Cambridge University Press, New York.