



Probabilistic analysis of a strip footing resting on a spatially varying soil using Kriging and enhanced active learning

A.-K. El Haj and A.-H. Soubra

University of Nantes, GeM, UMR CNRS 6183, Bd. de l'université, CS 70152, 44603 Saint-Nazaire cedex, France

T. Al-Bittar

Lebanese University, Faculty of Engineering Tripoli, Al Kobeh, Liban

ABSTRACT

As is well known, the classical Kriging-based probabilistic approaches such as the Active learning method combining Kriging and Monte Carlo Simulations MCS (named AK-MCS method) or the method combining Kriging and Importance Sampling IS (named AK-IS method) involve the construction of a surrogate Kriging metamodel based on the responses of a small design of experiments computed using the mechanical model. This approximate Kriging meta-model is then successively updated via an enrichment process by selecting new training points that are close to the limit state surface using a learning function. The essential issues in these approaches are that both the choice of a 'best new point' and the stopping criterion are defined from the perspective of individual responses, which may lead to some extra evaluations of unnecessary added training points. To overcome this shortcoming, a reliable and efficient probabilistic methodology based on an enhanced Kriging model (called Global Sensitivity Analysis-enhanced Surrogate GSAS modeling) was proposed by [Hu and Mahadevan \(2016\)](#). In this method, both the convergence criterion and the strategy of selecting new training points are defined from the perspective of reliability estimate instead of individual responses of MCS or IS points. A global sensitivity analysis is performed to select the optimal new training point and the convergence criterion is reached based on the desired accuracy of the reliability estimate. This method is applied in this paper in combination with the classical MCS or IS approach in order to reduce the number of calls of the mechanical model with respect to the corresponding classical AK-MCS and AK-IS approaches. It is used for the probabilistic analysis of the ultimate limit state of a strip footing resting on a spatially varying soil. The aim is the computation of the failure probability against soil punching.

The mechanical model was based on numerical simulations using the finite difference code *FLAC^{3D}*. The soil behavior was modeled using a conventional elastic-perfectly plastic model based on Mohr-Coulomb failure criterion. The soil cohesion c and angle of internal friction ϕ were modeled as two anisotropic non-Gaussian random fields. An anisotropic square exponential autocorrelation function was used for the two random fields. EOLE methodology was used to discretize these fields. All the other soil parameters are assumed to be deterministic.

The numerical results obtained from the combination of GSAS with either MCS or IS are compared to those obtained from the AK-MCS and AK-IS approaches. A significant reduction in the number of calls to the mechanical model was observed in both cases where GSAS was introduced in the probabilistic model. It should be noted that the probabilistic models allow one to obtain not only the failure probability but also the reliability index and the corresponding design point. The critical realization obtained at the design point was shown to be symmetrical with respect to the central vertical axis of the foundation with the weaker soil properties near the footing, the stronger soil being far from the footing.

1 INTRODUCTION

As is well known, the probabilistic analysis of geotechnical structures involving spatially varying soil properties is generally time-consuming since the corresponding deterministic models are based on finite element/finite difference methods. Furthermore, the probabilistic analysis of very heterogeneous soils significantly increases the computation time as compared to the time required for a homogeneous soil due to the increase in the stochastic dimension of the treated problem.

Because of the two mentioned issues related to the mechanical modeling and the soil spatial variability, conventional probabilistic methods such as Monte Carlo Simulation (MCS) are very expensive for the computation of the failure probability. It becomes unaffordable if an accurate value of the failure probability (i.e. with a small value of the coefficient of variation of the estimation) is desired. Consequently, a more advanced probabilistic approach is needed.

Recently, several Kriging-based metamodeling approaches (e.g. AK-MCS, AK-IS, etc.) have been developed by [Echard et al. \(2011, 2013\)](#) and were shown to be efficient with respect to the classical MCS methodology as one can obtain an accurate probability of failure needing a smaller number of calls to the mechanical model. Notice however that the essential issues in these approaches are (i) the choice of a best new training point for the construction of the metamodel and (ii) the stopping criterion related to the addition of new training points. Indeed, these issues are defined from the perspective of individual responses. This may lead to some extra evaluations of unnecessary added training points.

In order to overcome this shortcoming, a Global Sensitivity Analysis enhanced Surrogate (GSAS) modeling was developed by [Hu and Mahadevan \(2016\)](#). Within GSAS, a powerful new stopping criterion was suggested and a new way of selecting a new training point was proposed. In this respect, both the convergence criterion and the strategy for selecting new training points are defined from the perspective of reliability estimate instead of individual responses of MCS points. Indeed, the new training points are identified according to their contribution to the uncertainty in the reliability estimate and the selection of new training points stops when the accuracy of the reliability estimate reaches a specific target. It should be mentioned that [Hu and Mahadevan \(2016\)](#) have validated the proposed method based on several academic examples where the performance function was given by an analytical equation. The aim of this paper is to extend the GSAS approach proposed by [Hu and Mahadevan \(2016\)](#) to the case of random field problems in order to study geotechnical structures involving spatial variability of the soil properties. More specifically, this paper presents a probabilistic analysis at the Ultimate Limit State (ULS) of a strip footing resting on a spatially varying soil and subjected to a vertical load. The objective is the computation of the probability P_f of exceeding the ultimate bearing capacity of the footing under a prescribed vertical load.

2 PROBABILISTIC MODEL

The mechanical model is based on numerical simulations using the finite difference code *FLAC*^{3D}. A strip footing of breadth $B = 1\text{m}$ that rests on a soil domain of width $13B$ and depth $5B$ was considered in the analysis ([Al-Bittar and Soubra, 2014](#)). It was verified that with these model dimensions, the system response was not influenced by the artificial boundary effects. As suggested by [Der Kiureghian and Ke \(1988\)](#), the length of the largest element of the deterministic mesh in a given direction (horizontal or vertical) was chosen such that it does not exceed 0.5 times the autocorrelation distance in that direction. Concerning the displacement boundary conditions, the bottom boundary was assumed to be fixed and the vertical boundaries were constrained in motion in the horizontal direction.

The uncertain soil parameters considered in this paper may be described as follows: the soil cohesion c and angle of internal friction φ were modeled as two anisotropic non-Gaussian random fields. Notice also that the soil dilation angle ψ was considered to be related to the soil angle of internal friction φ by $\psi = 2\varphi/3$. This means that the soil dilation angle was implicitly assumed as a random field that is perfectly correlated to the soil angle of internal friction random field. The statistical parameters of the two random fields as used in the present paper are presented in Table 1.

Table 1: Statistical characteristics of the random fields

Random field	Mean μ	Coefficient of variation COV (%)	Type of the probability density function (PDF)
c	20 kPa	25	Log-normal
φ	30^0	10	Beta (bounds = 0^0 ; 45^0)

The EOLE methodology by [Li and Der Kiureghian \(1993\)](#) was used to discretize the two random fields (i.e. to obtain realizations of the soil cohesion c and angle of internal friction φ that respect the correlation structure of those fields). It should be noted that the same square exponential autocorrelation function was used for both c and φ . This autocorrelation function is given by the following equation

$$\rho = \exp \left(- \left(\frac{|\Delta x|}{a_x} \right)^2 - \left(\frac{|\Delta y|}{a_y} \right)^2 \right) \quad (1)$$

where a_x and a_y are the autocorrelation distances along x and y , respectively. In the present work, the horizontal and vertical autocorrelation distances are taken equal to 10 m and 5 m respectively.

3 GSAS GENERAL PROCEDURE FOR SPATIAL VARIABILITY PROBLEMS

This section aims at presenting the general procedure of the GSAS method as adapted to the case of random fields. Firstly, the GSAS-MCS model (which may be considered as an improvement of AK-MCS) is introduced. Then, the GSAS-IS model based on importance sampling will be briefly described.

3.1 GSAS-MCS procedure

The general procedure of the GSAS-MCS method as adapted to the case of random fields problems can be summarized as follows:

1. Generation by Monte Carlo simulation of $x^{(i)}$ ($i = 1, 2, \dots, N_{MCS}$) points. In this work, N_{MCS} was taken equal to 500,000. Each point $x^{(i)}$ consists of M standard Gaussian random variables where M is the number of random variables needed by EOLE methodology ([Li and Der Kiureghian, 1993](#)) to accurately discretize the cohesion and friction angle random fields. In this work, 6 random variables were adopted for each random field, thus leading to a small value of the variance of the error (=1.6%). This means that sufficiently accurate random fields discretization was adopted in the analysis.
2. Random selection of a small design of experiments (DoE) from the generated population (a DoE of 20 points was used in this work). Then, use of EOLE methodology to transform each point into realizations of the soil cohesion c and angle of internal friction φ that respect the correlation structure of those fields. For each selected point, the performance function G is evaluated using the following equation:

$$G = \frac{q_u}{q_s} - 1 \quad (2)$$

where q_u is the ultimate bearing capacity computed based on *FLAC*^{3D} software and q_s is the vertical load applied to the footing.

3. Based on the DoE and the corresponding performance function evaluations, an approximate Kriging meta-model is constructed in the standard space of random variables using the DACE toolbox ([Lophaven et al., 2002](#)).

4. For each Monte Carlo point $x^{(i)}$, the random response predicted by the approximate Kriging surrogate model is a Gaussian variate as follows:

$$G_p(x^{(i)}) \sim N(\hat{g}(x^{(i)}), \sigma_{G_p}^2(x^{(i)}))$$

where $\hat{g}(x^{(i)})$ and $\sigma_{G_p}^2(x^{(i)})$ are the mean prediction and the corresponding Kriging variance, respectively.

The Kriging mean predictions values $\hat{g}(x^{(i)})$ and their corresponding Kriging variance $\sigma_{G_p}^2(x^{(i)})$ values are determined for the whole MCS points using the DACE toolbox. Then, the failure probability \hat{P}_f is estimated using the following equation (after replacing the meta-model random responses $G_p(x^{(i)})$ by the mean prediction values $\hat{g}(x^{(i)})$):

$$\hat{P}_f = \sum_{i=1}^{N_{MCS}} I(G_p(x^{(i)})) / N_{MCS} \quad (3)$$

In this equation, $I(G_p(x^{(i)})) = 1$ if $G_p(x^{(i)}) \leq 0$, otherwise, $I(G_p(x^{(i)})) = 0$. Thus, \hat{P}_f is estimated by counting the number of negative mean predictors and dividing it by the total number of MCS points. The corresponding coefficient of variation $COV(\hat{P}_f)$ is given by the following equation:

$$COV(\hat{P}_f) = \sqrt{\frac{1 - \hat{P}_f}{\hat{P}_f \cdot N_{MCS}}} \quad (4)$$

5. It should be emphasized here that the value of the failure probability computed at this stage is far from being accurate because of the small DoE used so far. An enrichment process is thus needed. Within AK-MCS approach, the best next candidate point adopted during the enrichment process is selected as the one that is the most close to the limit state surface. This point could be considered as the one that mostly reduces the uncertainty in \hat{P}_f if the sample responses $G_p(x^{(i)})$ predicted from the surrogate model were completely independent. Notice however that these sample responses are correlated normal variables according to the property of the Kriging model. The GSAS approach allows one to overcome this shortcoming. The basic idea of this approach is to treat the probability of failure estimate \hat{P}_f as a random variate representing the output of the system presented in Figure 1 where the system inputs are the random responses $G_p(x^{(i)})$ predicted by the Kriging meta-model. In other words, the uncertainty in the input random variates $G_p(x^{(i)})$ is propagated through the system given by Eq. (3) and thus, the uncertainty in the failure probability estimate can be quantified.

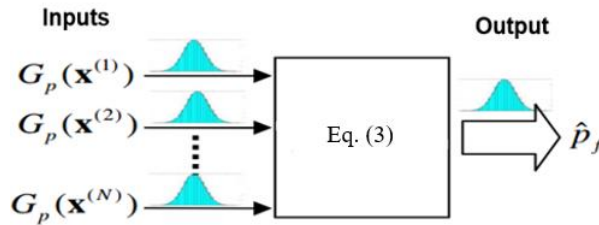


Figure 1: Probability of failure estimate as a system response

For an efficient enrichment of the Kriging meta-model within GSAS, the new training point is selected based on its contribution to the uncertainty of the quantity of interest (i.e. \hat{p}_f). It should reduce the uncertainty in \hat{p}_f in the most significant way. This is done *via* a global sensitivity analysis method extended to the case of models with dependent inputs. The extended FAST method developed by [Xu and Gertner \(2007\)](#) was used in this paper. The enrichment process within GSAS-MCS approach can be briefly described as follows: The MCS points are firstly classified into two

groups according to their U values where U is a learning function usually employed in the Kriging-based approaches. It is given by the following equation:

$$U(x^{(i)}) = \frac{|\hat{g}(x^{(i)})|}{\sigma_{G_p}(x^{(i)})} \quad (5)$$

Notice that a large value of U indicates a low probability of making an error on the sign of $\hat{g}(x)$. For $U(x^{(i)}) > 2$, the probability of making a mistake on the sign of the performance function value is less than 0.023 (Echard et al., 2011). Based on that, the MCS points $x^{(i)}$ ($i=1,2,\dots,N_{MCS}$) are divided into two groups, namely the group x_{g1}^{MCS} with U values larger than 2 and the group x_{g2}^{MCS} with the remaining points in x^{MCS} . The global sensitivity analysis is then performed on the x_{g2}^{MCS} group points to determine their contributions to the uncertainty of \hat{p}_f since we assume that the uncertainty of \hat{p}_f comes from this group of points. It should be noted that in order to reduce the dimensionality of the problem, only a reduced number n_{can} of candidate points (taken equal to 20 in this work) of the x_{g2}^{MCS} group with the lowest U values are selected to perform the global sensitivity analysis since they have high probability of having wrong performance function signs (i.e. high probability of being the new selected training point).

6. A powerful stopping criterion based on the quantification of the uncertainty in the failure probability was suggested within GSAS approach. Although a prescribed maximal value of the uncertainty on the failure probability would be a quite relevant stopping condition (because it makes sure that the uncertainty in \hat{P}_f is sufficiently small), Hu and Mahadevan (2016) suggest stopping the addition of new points based on the uncertainty of the error on the failure probability ε_r . The error on the failure probability is a measure of the error between the theoretical and the computed values of the failure probability. It is defined by the following equation:

$$\varepsilon_r = \frac{\hat{P}_f - \hat{P}_f'}{\hat{P}_f} \quad (6)$$

where \hat{P}_f is the theoretical failure probability given by Eq. (3) and \hat{P}_f' is the estimate value of the failure probability that can be directly computed based on the Kriging meta-model mean prediction values $\hat{g}(x^{(i)})$. It should be noted here that \hat{P}_f given by Eq. (3) is not a unique scalar value that can be computed (because $G_p(x^{(i)})$ in this equation is a Gaussian variate), but a random variate for which one can quantify the corresponding uncertainty. The uncertainty in ε_r as given by Eq. (6) was thus quantified herein based on the uncertainty quantification of \hat{P}_f . The sampling based method was used: This method consists in generating n_r points ($n_r=600$ in this work) of N_2 correlated normal variables $G_p(x_{g2}^{MCS}(i))$, $i = 1, 2, \dots, N_2$ for each point where N_2 is the number of points in the x_{g2}^{MCS} group. From these points, one can compute n_r samples of the failure probability \hat{P}_f and other n_r corresponding samples of the error ε_r . From the $\varepsilon_r(i)$, $i = 1, 2, \dots, n_r$ samples, the Kernel Smoothing function is employed to fit the distribution of the error ε_r . Based on the fitted distribution, Hu and Mahadevan (2016) have suggested stopping the addition of new points when the quantity ε_r^{max} becomes smaller than a prescribed threshold a (where a is taken equal to 1% in this paper) where ε_r^{max} defined as follows:

$$\varepsilon_r^{max} = \max\{|F_{\varepsilon_r}^{-1}(0.99)|, |F_{\varepsilon_r}^{-1}(0.01)|\} \quad (7)$$

In this equation, $F_{\varepsilon_r}^{-1}$ is the inverse CDF of ε_r . The proposed stopping condition corresponds to a probability of 2% that the actual estimation error on \hat{P}_f is larger than 1%. For more information on this criterion, the reader may refer to Hu and Mahadevan (2016). Notice finally that the value of ε_r^{max} was checked every time the surrogate model was updated except for the case where the number N_2 was too large ($>8,000$ samples). This is related to the fact that the error computation cost is large in this case. Furthermore, this cost would be of no interest since the uncertainty on the failure probability estimate is obviously significant.

3.2 GSAS-IS procedure

When dealing with the small failure probabilities encountered in practice, the computation time of GSAS-MCS remains important (although in this method, one makes use of the predictions computed using the Kriging meta-model) since a large population with a very large number of points is required by MCS (e.g. 500,000 points) to lead to a small value of the coefficient of variation on the failure probability. As a result, [Hu and Mahadevan \(2016\)](#) suggested a combination between GSAS and importance sampling (IS). This procedure is called herein GSAS-IS. Notice that GSAS-IS can be regarded as an extension of the AK-IS approach proposed by [Echard et al. \(2013\)](#). GSAS-IS approach can reduce the computation time of the probabilistic analysis as compared to GSAS-MCS by reducing the size of the sampling population.

The GSAS-IS procedure consists of two main stages. In the first stage, the most probable failure point (design point) is determined *via* an iterative procedure using an approximate Kriging meta-model based on a small number of points. In the second stage, the obtained approximate Kriging meta-model is successively improved *via* an enrichment process (as in GSAS-MCS). Notice however that in GSAS-IS, the enrichment is performed based on points selected from a probability density function that is centered at the design point.

The main procedure of the GSAS-IS method (as adapted to the case of random fields) is nearly the same as that of GSAS-MCS described in paragraph 3.1, except step (4) which is now modified as follows (the other steps being similar to those of GSAS-MCS and thus they are not re-written herein to avoid repetition).

- a. Find the Hasofer-Lind reliability index and the corresponding value of the design point by making use of the approximate already-obtained Kriging meta-model. This procedure gives an approximate value of the reliability index and its corresponding design point.
- b. Generate a small number of points (5 points are used in this paper) so that they are centered at the design point obtained in the previous step and compute the corresponding values of the performance function using $FLAC^{3D}$.
- c. Construct a new Kriging meta-model in the standard space using all points generated so far. This Kriging meta-model is used to obtain an updated design point and its corresponding Hasofer-Lind reliability index.
- d. Steps (b) and (c) are repeated several times until the absolute difference between two successive values of the Hasofer-Lind reliability index becomes smaller than a given tolerance (taken equal to 0.01).

Once the final design point is obtained, an IS population (with a reduced number of points of say 10,000 points) that is centered at the design point is generated. The failure probability \hat{P}_f and the corresponding coefficient of variation $COV(\hat{P}_f)$ are then calculated after each added point using the corresponding formulas of importance sampling approach. Finally, it should be emphasized herein that the enrichment process is the same as that of GSAS-MCS, except that the selection of enrichment points is done among the reduced IS population (10,000 points). Moreover, the number N_2 is considered herein as being too large if it is bigger than 1,000.

4 NUMERICAL RESULTS

In order to check the efficiency of GSAS-MCS and GSAS-IS with respect to the classical AK-MCS and AK-IS approaches respectively, a probabilistic computation has been performed on the same problem but using AK-MCS and AK-IS.

Table 2: Numerical results of the different methods

Method	$\hat{P}_f \times 10^{-3}$	% COV(\hat{P}_f)	Size of the DoE	Added points
AK-MCS	3.250	2.48	20	406
GSAS-MCS	3.154	2.51	20	71
AK-IS	3.172	2.06	20 + 5 × 8	354
GSAS-IS	3.245	2.08	20 + 5 × 8	102

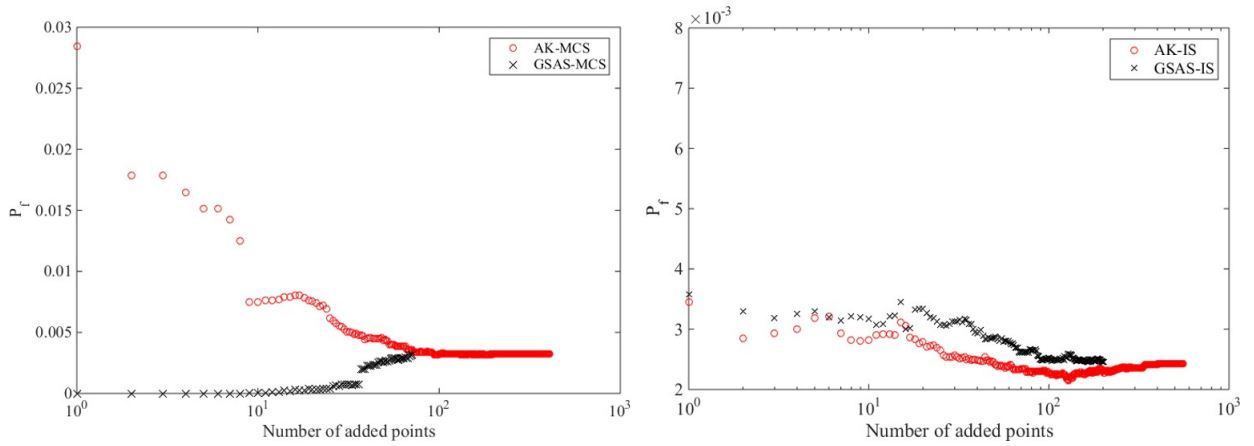

Figure 2: Failure probability vs the number of added points

Table 2 provides the probabilistic results obtained from the four methods. Also, Figure 2 presents the evolution of \hat{P}_f with the number of added points as given by the different methods making use of the corresponding stopping conditions. As may be seen from this figure and from Table 2, GSAS-MCS and GSAS-IS are powerful approaches since they provide quasi similar values of the failure probability as the corresponding AK-MCS and AK-IS classical approaches making use of a much reduced number of calls to the mechanical model. The methods based on IS are much more efficient than those based on MCS in terms of the computation time because a smaller population is used in these methods (10,000 instead of 500,000) while computing the Kriging meta-model estimations. Finally note that when using IS-based methods, the \hat{P}_f values calculated after the first added points were not very far from the final \hat{P}_f value. This may be explained by the fact that the added samples in this case are chosen within the zone of interest for the computation of the failure probability (i.e. around the design point).

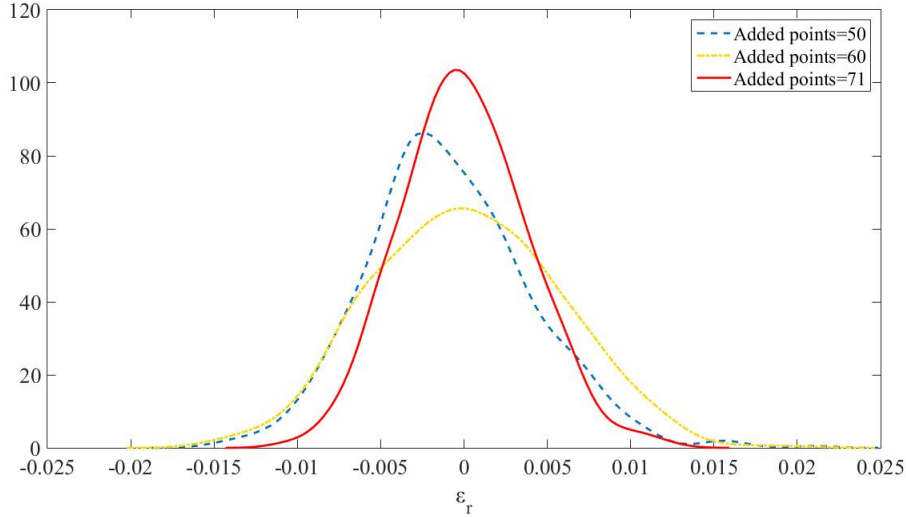


Figure 3: Evolution of the fitted distribution of ε_r with the added points

Figure 3 presents the evolution of the distribution of ε_r with the number of added points for the GSAS-MCS calculation process (similar trend was obtained for GSAS-IS). This figure shows that (i) the mean value of the error converges successively to zero and (ii) the variability of ε_r decreases with the number of added points, the corresponding standard deviation value becomes very small (with a value of 3.78×10^{-3}) when reaching the optimal number of added points (i.e. 71 points). These two observations provide a quite good indication on the convergence of the estimated failure probability to its theoretical value. Table 3 shows the evolution of ε_r^{max} and the corresponding value of the failure probability \hat{P}_f with the added points. ε_r^{max} decreases progressively until reaching the stopping condition ($\varepsilon_r^{max} < 1\%$) which indicates that the mean error on \hat{P}_f is converging to zero and the level of uncertainty on this error has reached the threshold value. The final value of the estimated failure probability \hat{P}_f is thus obtained within the target uncertainty level.

Table 3: Evolution of ε_r^{max} and \hat{P}_f with the added points

Added points	$\varepsilon_r^{max}(\%)$	\hat{P}_f
50	1.74	2.626
60	1.43	2.766
71	0.92	3.154

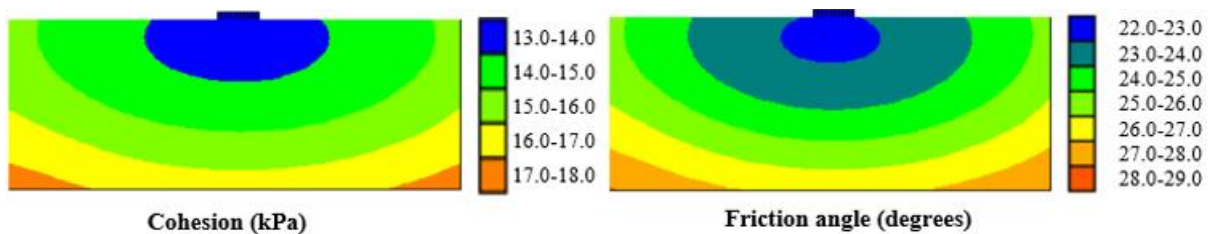


Figure 4: Critical realizations of the soil shear strength parameters

Figure 4 presents the critical realizations of the soil shear strength parameters corresponding to the obtained design point. This figure exhibits a symmetrical distribution of the soil shear strength parameters with respect to the central vertical axis of the foundation. The weaker soil zone is concentrated around the foundation while the stronger soil is far from the foundation. The weak soil zone under the foundation allows the failure mechanism to develop through this zone thus reflecting the most prone soil to

punching.

5 CONCLUSION

A probabilistic analysis was performed at the Ultimate Limit State (ULS) of a strip footing resting on a spatially varying soil and subjected to a vertical load. The soil cohesion c and angle of internal friction φ were modeled as two anisotropic non-Gaussian random fields. EOLE method was used for the generation of realizations of the random fields.

The Global Sensitivity Analysis enhanced Surrogate (GSAS) modeling proposed by [Hu and Mahadevan \(2016\)](#) was extended in this work to the case of random fields and used to perform the reliability analysis. The method was applied in combination with the classical MCS or IS approach. The resulting methods have shown high efficiency as compared to the corresponding AK-MCS and AK-IS approaches since they have led to quasi similar values of the failure probability and coefficient of variation making use of a much reduced number of calls to the mechanical model.

The critical realizations at the design point have shown a symmetrical distribution of the soil shear strength parameters with respect to the central vertical axis of the foundation with a weak soil zone near the footing.

ACKNOWLEDGMENTS

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.

REFERENCES

- Al-Bittar, T. and A. Soubra (2014). Efficient sparse polynomial chaos expansion methodology for the probabilistic analysis of computationally-expensive deterministic models. *International Journal for Numerical and Analytical Methods in Geomechanics* 12(38), 1211–1230.
- Der Kiureghian, A. and J. Ke (1988). The stochastic finite element method in structural reliability. *Probabilistic Engineering Mechanics* 3(2), 83–91.
- Echard, B., N. Gayton, M. Lemaire, and N. Relun (2011). An active learning reliability method combining Kriging and monte carlo simulation. *Structural Safety* 2(33), 145–154.
- Echard, B., N. Gayton, M. Lemaire, and N. Relun (2013). A combined importance sampling and Kriging reliability method for small failure probabilities with time-demanding numerical models. *Reliability Engineering and System Safety* (111), 232–240.
- Hu, Z. and S. Mahadevan (2016). Global sensitivity analysis-enhanced surrogate (GSAS) modeling for reliability analysis. *Structural and Multidisciplinary Optimization* 53, 501–521.
- Li, C. and A. Der Kiureghian (1993). Optimal discretization of random fields. *J. Eng. Mech.-ASCE*, 119(6), 1136–1154.
- Lophaven, S., H. Nielsen, and J. Sondergaard (2002). DACE: A Matlab Kriging toolbox.
- Xu, C. and G. Gertner (2007). Extending a global sensitivity analysis technique to models with correlated parameters. *Computational Statistics and Data Analysis* 51, 5579–5590.