Subset Simulation and its Application to a Spatially Random Soil

Ashraf Ahmed[1] and Abdul-Hamid Soubra, M.ASCE[2]

[1] PhD student, University of Nantes, GeM, UMR CNRS 6183, Bd. de l'université, BP 152, 44603 Saint-Nazaire cedex, France, Tel: 0033240905106,
Fax: 0033240905109, E-mail: Ashraf.Ahmed@univ nantes.fr
[2] Corresponding author, Professor, University of Nantes, GeM, UMR CNRS 6183, Bd. de l'université, BP 152, 44603 Saint-Nazaire cedex, France,
Tel: 0033240905108, Fax: 0033240905109, E-mail: Abed.Soubra@univ-nantes.fr

ABSTRACT

The failure probability of a footing resting on a spatially random soil is currently computed using Monte Carlo simulation (MCS) methodology. This approach is well known to be very time-consuming especially for computing small failure probabilities. One alternative to MCS is the subset simulation approach. This approach is used in this paper to perform a probabilistic analysis at the serviceability limit state of a strip footing resting on a soil with a spatially varying Young modulus. The random field was discretized using Karhunen-Loeve expansion. The probabilistic numerical results have shown that the failure probability calculated by subset simulation is very close to that computed by MCS methodology but with a significant reduction in the number of simulations.

INTRODUCTION

The failure probability of geotechnical problems modeled by random fields is generally calculated using Monte Carlo simulation (MCS) methodology. This method is very expensive for the computation of a small failure probability. Au and Beck (2001) proposed an efficient approach (called subset simulation) to calculate the small failure probabilities. Except Au et al. (2010) and Santoso et al. (2010) who applied this approach to the random field problems, the subset simulation method was mainly applied in literature to problems where the uncertainties of the different parameters were modeled by random variables. These authors have considered only one-dimensional random field problems. It should be noted that Au et al. (2010) have discretized the random field into a finite number of random variables equal to the number of elements of the deterministic model. Thus the random dimension depends on the number of elements of the deterministic model. In this paper, the subset simulation method was proposed to calculate the failure probability in case of a two-dimensional random field discretized by K-L expansion. As will be explained hereafter, the random dimension in the present work does not depend on the number of elements of the deterministic model but on the size M of the K-L expansion. The proposed procedure was applied to perform a probabilistic analysis at the serviceability limit state (SLS) of a shallow strip footing subjected to a central vertical load (P_V). The uncertain random field considered in the analysis is the soil Young modulus (E). The footing vertical displacement was used to represent the system response. The deterministic model used to compute the system response is based on numerical simulations using the commercial software $FLAC^{3D}$.

KARHNUEN-LOEVE EXPANSION

Consider a random process H(X, θ) where X denotes the spatial coordinates and θ indicates the random nature of the corresponding quantity. If μ is the mean of the process, then the process can be expanded as follows Spanos and Ghanem (1989):

$$H(X,\theta) = \mu + \sum_{i=1}^{\infty} \sqrt{\lambda_i}\phi_i(X)\xi_i(\theta) \quad (1)$$

where λ_i and ϕ_i are the eigenvalues and eigenfunctions of the covariance function $C(X_1, X_2)$ and ξ(θ) is a vector of standard uncorrelated random variables. The series expansion in Equation (1) is referred to as the K-L expansion. For practical implementation, the approximate random process is defined by truncating the series in Equation (1) to a finite number of terms M. The choice of the number of terms M depends on the desired accuracy of the problem being treated.

SUBSET SIMULATION

Subset simulation was proposed by Au and Beck (2001) to compute the small failure probabilities. The basic idea of subset simulation can be described as follows: Consider a failure region F defined by the condition G<0 where G is the performance function and let (s_1, …, s_k, ..., s_{Nt}) be N_t samples located in the space of the uncertain variables. It is possible to divide the failure region F into a number of nested failure regions F_1, …, F_j, ..., F_m of increasing size where $F_1 \supset .. \supset F_j \supset .. \supset F_m = F$ (Fig. 1).

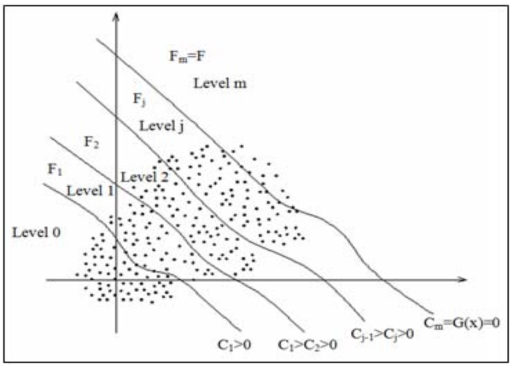

Figure 1. Nested Failure domain

An intermediate failure region F_j can be defined by $G_j<C_j$ where $C_j>0$. Thus, there is a decreasing sequence of positive numbers C_1, …, C_j, ..., C_m corresponding respectively to F_1, …, F_j,…, F_m where $C_1>…>C_j>...> C_m=0$. The N_t samples (s_1, …, s_k, ..., s_{Nt}) will be divided into groups where each group contains a given number N_s

of samples (s_1, …, s_k, ..., s_{Ns}). The failure probability corresponding to an intermediate failure region F_j is calculated as follows:

$$P(F_j) = \sum_{k=1}^{Ns} I_{F_j}(s_k) \qquad (2)$$

I_{F_j} =1 if $s \in F_j$ and I_{F_j} =0 otherwise. The conditional failure probability $P(F)=P(F_m)$ can be calculated from the previous nested sequence of conditional failure regions as follows:

$$P(F)=P(F_m)=P(F_m|F_{m-1})xP(F_{m-1}|F_{m-2})xP(F_{m-2}|F_{m-3})x \ldots xP(F_2|F_1)xP(F_1) \qquad (3)$$

IMPLEMENTATION OF SUBSET SIMULATION ALGORITHM IN THE CASE OF A SPATIALLY VARYING SOIL PROPERTY

As mentioned previously, this paper aims at employing the subset simulation methodology for the computation of the failure probability in the case of a spatially varying soil property. In order to achieve this objective, K-L expansion described before is used to discretize the random field. Therefore, the random field is transformed into a finite number of random variables $\{\xi_i\}_{i=1,\ldots,M}$. Due to this transformation, the use of subset simulation becomes an easy task. The algorithm of subset simulation in case of a spatially varying soil property can thus be described by the following steps:

1- Choose the number M of terms of K-L expansion. This number must be sufficient to rigorously represent the target random field.
2- Generate a vector {ξ_1, …, ξ_k, ..., ξ_M} by direct Monte Carlo simulation with the target probability density function P_t. Notice that P_t is normal in our case.
3- Use the K-L expansion to obtain the first realisation of the random field. Then, use the deterministic model to calculate the corresponding response.
4- Repeat steps 2 and 3 to obtain N_s realisations and their corresponding responses. Then, evaluate the corresponding values of the performance function to obtain the vector $\mathbf{G}_0$={g_1, …, g_k, ..., g_{Ns}}.
5- Propose an intermediate failure probability $P(F_j)$ and evaluate the first failure threshold C_1 which corresponds to the failure region F_1 where C_1 is equal to the $[(N_s xP(F_j))+1]^{th}$ value in the increasing list of elements of the vector $\mathbf{G}_0$. This ensures that the value of $P(F_1)$ wi equal to the proposed value of $P(F_j)$.
6- Among the N_s realisations, there are $[N_s xP(F_j)]$ ones whose values of the performance function is less than C_1 (i.e. they are located in the failure region F_1). The corresponding vectors {ξ_1, …, ξ_k, ..., ξ_M} of these realisations are used as ‘seeds’ to generate an additional $[(1-P(F_j))N_s]$ vectors of {ξ_1, …, ξ_k, ..., ξ_M} using Markov chain method based on Metropolis-Hasting algorithm. These new vectors are used to obtain the random field realisations of level 1 using K-L expansion. The corresponding responses are evaluated using the deterministic model and the corresponding vector of performance function $\mathbf{G}_1$={g_1, …, g_k, ..., g_{Ns}} is computed.

7- Evaluate the second failure threshold C_2 as the $[(N_s x P(F_j))+1]^{th}$ value in the increasing list of the vector $\mathbf{G}_1$.

8- Repeat steps 6-7 until reaching the value of the last failure threshold (i.e. C_m=0). For the last failure threshold, thowing equation may be used:

$$P(\mathrm{F_m} | F_{m-1}) = \sum_{k=1}^{Ns} I_{F_m}(s_k) \tag{4}$$

Where I_{F_m} =1 if the performance function g_k corresponding to the sample s_k is negative and I_{F_m} =0 otherwise.

9- Finally, the failure probability P(F) is evaluated according to Equation (3).

EXAMPLE PROBLEM

In this section, the proposed procedure for the computation of the failure probability is illustrated through an example problem. In this example, the failure probability at SLS (i.e. the probability of exceeding a tolerable vertical displacement) of a strip footing resting on a soil having spatially varying Young modulus and subjected to a central vertical load (P_v) is calculated. The Young modulus was considered as a log-normal random field with mean value μ=60MPa and coefficient of variation COV=15%. An exponential covariance function was used in this paper. The random field was discretized using K-L expansion. Although an isotropic random field is often assumed in theature; in this paper, different values of the horizontal and vertical autocorrelation distances were studied and analyzed. The performance function used to calculate the failure probability is defined as follows:

$$G=v_{max}-v \tag{5}$$

where v_{max} is a prescribed tolerable vertical displacement and v is the footing vertical displacement due to the applied load. The deterministic model used to calculate the footing vertical displacement (v) is based on the commercial numerical code $FLAC^{3D}$. Due to the non-symmetrical footing movement created by the soil heterogeneity, the entire soil domain shown in Fig. 2 was considered to calculate the footing displacement. The values of the parameters of the soil, footing and interface are given in Table (1).

It should be mentioned here that a normal PDF was used as a target probability distribution P_t. However, a uniform PDF was used as a proposal probability distribution P_p. Notice also that the intermediate failure probability $P(F_i)$ was chosen equal to 0.1. Notice finally that v_{max} was assumed equal to 4cm and P_v was taken equal to 1000kN/m throughout the paper. As mentioned before, the accuracy of the approximated random field depends on the number M of terms of the K-L expansion. For the most critical configurations of the autocorrelation distances used in this paper [(L_x=10m and L_y=0.5m), (L_x=5m and L_y=1m)], the eigenvalue vanishes when the number of terms is $M \approx 100$ terms. As a conclusion, the number of terms will be set to M=100 terms for all the probabilistic calculations presented in this paper.

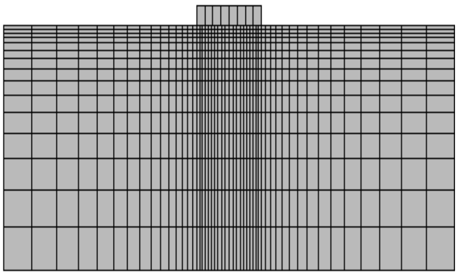

Figure 2. Soil domain and mesh used in the numerical simulation

Table 1. Shear Strength and Elastic Properties of Soil, Footing, and Interface

Variable	Soil	Footing	Interface
c	20kPa	N/A	20kPa
φ	30^o	N/A	30^o
$\psi_s=2/3\ \varphi$	20^o	N/A	20^o
E	60MPa	25GPa	N/A
ν	0.3	0.4	N/A
K_n	N/A	N/A	1GPa
K_s	N/A	N/A	1GPa

Comparison between subset simulation and MCS methodologies

In this section, the failure probabilities calculated at each level of the subset simulation were compared to those obtained by MCS. To perform this comparison, a random field with L_x=10m and L_y=1m (called hereafter the reference case) was considered. For each level j of the subset simulation, the corresponding failure threshold C_j was calculated and presented in Table (2) for different values of the number of realisations N_s.

Table 2. Evolution of the performance function with the different levels of the subset simulation and with the number of realisations (N_s) at each level

Performance function C_j for each level j	Number of realisations at each level (N_s)				
	50	100	150	200	250
C_1	0.0086	0.0077	0.0080	0.0076	0.0076
C_2	0.0058	0.0048	0.0050	0.0041	0.0040
C_3	0.0044	0.0015	0.0019	0.0011	0.0011
C_4	0.0017	0	0	0	0
C_5	0	-	-	-	-

This table indicates that the failure threshold decreases with the successive levels until reaching zero at the last level. This means that the realisations generated by the proposed procedure successfully progress towards the limit state surface G=0 which indicates the validity of the proposed procedure. Notice that the value of the failure threshold remains almost constant for $N_s \geq 200$. This means that 200 realisations at each level are sufficient to calculate the failure probability. To confirm the validity of the present approach, a comparison with MCS was carried out. The number of realisations N of MCS should be adequate for a rigorous calculation of the failure probability. The accuracy of MCS can be estimated by calculating the coefficient of variation of the failure probability as follows:

$$COV_{P_f} = \sqrt{(1-P_f)/NP_f} \tag{6}$$

Using this equation, the coefficient of variation of the failure probability was calculated and presented in Fig. 3. It was found to decrease with the increase of the number of realisations. It reaches an asymptote when the number N of realisations is equal to 20,000. Hence, 20,000 realisations were used to calculate the failure probabilities by MCS. The comparison between the subset simulation method (using 4 levels with N_s=200, i.e. 800 realisations) and the MCS method is presented in Fig. 4.

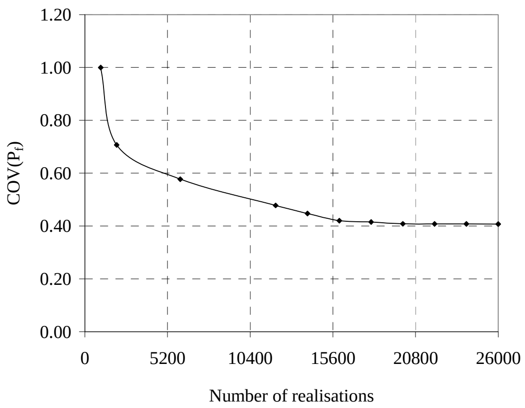

Figure 3. Effect of number of realisations of MCS on COV(P_f)

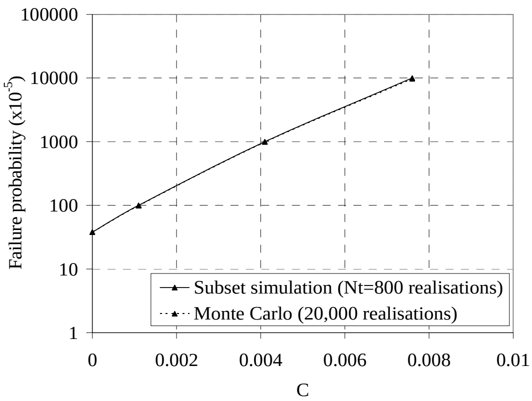

Figure 4. Comparison between P_f obtainedat at each level of subset simulation and that computed by MCS

Notice that to calculate the failure probability at level j using MCS, the performance function should be set equal to C_j. In this case, the failure region is defined as G≤C_j and the safety region is defined as G>C_j. The failure probability can then be calculated as:

$$P(F_i) = \sum_{k=1}^{N} I_F(s_k) \tag{7}$$

where I_F =1 if s_k is located in the failure region F_j and I_F =0 otherwise. Fig. 4 shows that for the case where N_s=200 realisations, the failure probabilities calculated by subset simulation were found to be very close to those computed by MCS. For smaller values of N_s, there is some discrepancy between both approaches (results not shown in this paper). From Fig. 4, one can conclude that the 20,000 realisations required by MCS methodology can be reduced to only 800 realisations by applying the proposed procedure based on subset simulation methodology.

Parametric study

In order to investigate the effect of the anisotropy of the random field, the failure probability was plotted against the vertical and horizontal autocorrelation distance in Figs. 5 and 6 respectively for the prescribed footing breadth. (i.e. B=2m).

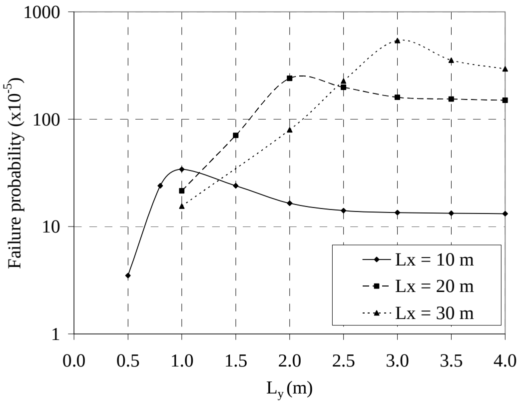

Figure 5. Effect of L_y on the failure probability for different values of L_x

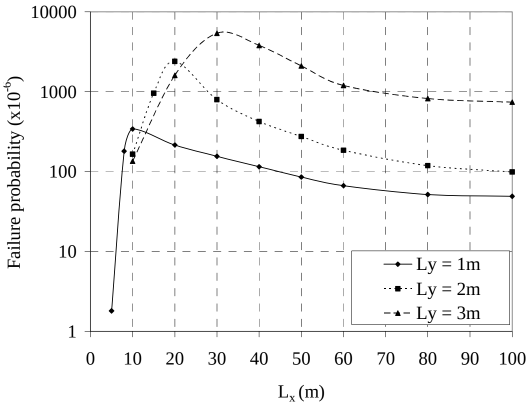

Figure 6. Effect of L_x on the failure probability for different values of L_y

Both figures show that the failure probability presents a maximum when the ratio of L_x to L_y is equal to 10 for B=2m. This observation can be explained as follows: The small value of the autocorrelation distance (L_x or L_y) induces a large soil heterogeneity which results in a large variety (i.e. a great number of high and small values) of the Young modulus beneath the footing. This variety leads to small footing displacement and consequently to small failure probability. The small footing displacement occurs because the rigid footing is prevented to settle due to the presence of a high number of small strong zones beneath this footing; the high number of small weak zones being of little effect in this case; the increase in the autocorrelation distance increases the footing displacement and consequently the failure probability. On the other hand, when the autocorrelation distance is very large in either the vertical or the horizontal direction, the problem becomes similar to that of the one-dimensional random field for which the failure probability is smaller than that of the two-dimensional isotropic case. For medium autocorrelation distances, the soil contains a number of strong zones adjacent to a number of weak zones whose areas are greater than those corresponding to the case of the one-dimensional random field. The soil movement can easily develop throughout the weak zones. As a result, the failure probability increases in such cases with respect to the case of the one-dimensional random field. The values of the autocorrelation distance corresponding to the maximum can be used to define a conservative design case for probabilistic analysis when rigorous information about the autocorrelation distance is not available. Notice finally that for the same ratio of L_x/L_y but greater values of L_x and L_y, the maximum failure probability was found to be higher. This is due to the simultaneous increase of the autocorrelation distances in both the vertical and the horizontal directions which makes the failure probability tend to that of the homogeneous soil. The numerical results of Figs 5 and 6 also indicate that the failure probability is more sensitive to the vertical autocorrelation distance. This is because the rate of change in the failure probability (i.e. rate of increase or decrease) when increasing the vertical autocorrelation distance by a certain percentage is larger than that when increasing the horizontal autocorrelation distance by the same percentage.

CONCLUSION

This paper presents an alternative procedure to Monte Carlo Simulation (MCS) methodology for the computation of the failure probability of a strip footing resting on a soil with spatially varying properties. This procedure is based on the subset simulation approach. The proposed procedure was applied to calculate the failure probability at the serviceability limit state of a strip footing resting on a soil with a spatially varying Young modulus and subjected to a central vertical load. The random field was discretized into a finite number of random variables using K-L expansion. The vertical displacement of the footing center was used to represent the system response. The deterministic model used to compute the system response is based on numerical simulations using the commercial software $FLAC^{3D}$. The failure probability calculated by the proposed procedure was compared to that computed by MCS methodology and was found to be very close with a significant decrease in the number of calls of the deterministic model. For the case of a non-isotropic random field, the failure probability presents a maximum value. The maximum was obtained at a fixed value of L_x/L_y=10 for a footing breadth equal to 2. Notice however that for the same ratio of L_x/L_y but greater values of L_x and L_y, the maximum failure probability was found to be higher. It has also been shown that the failure probability is more sensitive to the vertical autocorrelation distance than the horizontal one.

REFERENCES

Au, S. K., and Beck, J. L. (2001). “Estimation of small failure probabilities in high dimentions by subset simulation.” *J. Prob. Engrg. Mech.*, 16,263-277.

Au, S. K., Cao, Z. J., Wang, Y. (2010). “Implementing advanced Monte Carlo simulation under spreadsheet environment.” *J. Structural Safety*,32(5),281-292.

Santoso, A. M., Phoon, K. K., Quek, S. T. (2010). “Modified Metropolis-Hasting algorithm with reduced chain-correlation for efficient subset simulation.” *J. Prob. Engrg. Mech.*, doi:10.1016/j.probengmech.2010.08.007.

Spanos, P. D., and Ghanem, R. (1989). “Stochastic Finite Element Expansion for Random Media.” *J. Engrg. Mech.*, 115(5),1035-1053.