Kriging-based Approaches for the Probabilistic Analysis of Strip Footings Resting on Spatially Varying Soils

Jawad Thajeel

UNIVERSITE BRETAGNE LOIRE

UNIVERSITÉ DE NANTES

Thèse de Doctorat

Jawad Kadhim THAJEEL

Mémoire présenté en vue de l'obtention du
grade de Docteur de l'Université de Nantes
sous le sceau de l'Université Bretagne Loire

École doctorale : Sciences pour L'ingénieur, Géosciences, Architecture

Discipline : Génie Civil
Spécialité : Géotechnique
Unité de recherche : Génie Civil et Mécanique UMR CNRS 6183

Date du soutenue le 08/12/2017
Thèse N° :

Kriging-based Approaches for the Probabilistic Analysis of Strip Footings Resting on Spatially Varying Soils

JURY

Président du jury:	**Fabrice EMERIAULT**	**Professeur, Université Grenoble Alpes**
Rapporteurs:	**Olivier DECK**	Professeur, Université de Lorraine
	Michael HICKS	Professor, Delft University of Technology
Examinateurs:		
	Ashraf AHMED	Assistant Professor, Aswan University
	Tamara AL-BITTAR	Assistant Professor, Lebanese University
	Franck SCHOEFS	Professeur, Université de Nantes
Directeur de Thèse:	**Abdul-Hamid SOUBRA**	Professeur, Université de Nantes

ACKNOWLEDGEMENTS

First and foremost, praises and thanks to the God, the Almighty, for His granting of blessings throughout my research work to complete the research successfully.

I would like to express my deep and sincere gratitude to my supervisor: Prof. **Abdul-Hamid SOUBRA** for the continuous support of my PhD study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better supervisor and mentor for my PhD study.

Besides my supervisor, I wish to thank the members of my dissertation committee, Prof. Fabrice EMERIAULT, Professeur, Université Grenoble Alpes, for having accepted to be its president. The rapporteurs: Prof. Michael HICKS, Delft University of Technology and Prof. Olivier DECK, Professor, University of Lorraine. I wish to express my gratefulness to them for their careful reading and rating of my thesis report and for their insightful comments and encouragement. I also thank all the jury members of my thesis: Prof. Franck SCHOEFS, Professor, University of Nantes, DR. Ashraf AHMED, Assistant Professor, Aswan University and Dr.Tamara AL-BITTAR, Assistant Professor, Lebanese University for having accepted to be part of the jury and for their relevant questions after my presentation.

I would also like to thank the members of my CST committee Prof. Luc THOREL, Professor, IFSTTAR-Nantes and Prof. Daniel Dias, Professor, Grenoble Alpes University for their comments and good advice for this research.

I would like also to acknowledge the financial supporter of my PhD studies from the Ministry of Higher Education and Scientific Research (MOHERS) of my country, Republic of Iraq and special thanks for the University of Thi Qar, Faculty of Engineering. Also I would to express my special thanks to the Campus France to facilitate all the obstacles in France.

Most especially, I am grateful to my big family, my mother, my father and brothers, and to my small family, my wonderful wife and my sons Almustafa and Almurtada for their understanding and support.

At last, if I were to individually mention all those who aided in my achievement, the list would be long. However, my sincere apologies and my ultimate gratitude to all those who helped from behind the curtains, and who I could not manage to mention in person.

ABSTRACT

The probabilistic analysis of geotechnical structures involving spatially varying soil properties is generally performed using Monte Carlo Simulation methodology. This method is not suitable for the computation of the small failure probabilities encountered in practice because it becomes very time-expensive in such cases due to the large number of simulations required to calculate accurate values of the failure probability. Three probabilistic approaches (named AK-MCS, AK-IS and AK-SS) based on Active learning and combining Kriging and one of the three simulation techniques (i.e. Monte Carlo Simulation MCS, Importance Sampling IS or Subset Simulation SS) were developed. Within AK-MCS, a Monte Carlo simulation without evaluating the whole population is performed. Indeed, the population is predicted using a kriging meta-model which is defined using only a few points of the population, thus significantly reducing the computation time with respect to the crude MCS. In AK-IS, a more efficient sampling technique 'IS' is used instead of 'MCS'. In the framework of this approach, the small failure probability is estimated with a similar accuracy as AK-MCS but using a much smaller size of the initial population, thus significantly reducing the computation time. Finally, in AK-SS, a more efficient sampling technique 'SS' is proposed. This technique overcomes the search of the design points and thus it can deal with arbitrary shapes of the limit state surfaces. All the three methods were applied to the case of a vertically loaded strip footing resting on a spatially varying soil. The obtained results are presented and discussed.

KEY WORDS: spatial variability, Kriging metamodeling, probability of failure, Monte Carlo Simulation, Importance Sampling, Subset Simulation.

RÉSUMÉ

L’analyse probabiliste des ouvrages géotechniques est généralement réalisée en utilisant la méthode de simulation de Monte Carlo. Cette méthode n’est pas adaptée pour le calcul des faibles probabilités de rupture rencontrées dans la pratique car elle devient très coûteuse dans ces cas en raison du grand nombre de simulations requises pour obtenir la probabilité de rupture. Dans cette thèse, nous avons développé trois méthodes probabilistes (appelées AK-MCS, AK-IS et AK-SS) basées sur une méthode d’apprentissage (Active learning) et combinant la technique de Krigeage et l’une des trois méthodes de simulation (i.e. Monte Carlo Simulation MCS, Importance Sampling IS ou Subset Simulation SS). Dans AK-MCS, la population est prédite en utilisant un méta-modèle de krigeage qui est défini en utilisant seulement quelques points de la population, ce qui réduit considérablement le temps de calcul par rapport à la méthode MCS. Dans AK-IS, une technique d'échantillonnage plus efficace 'IS' est utilisée. Dans le cadre de cette approche, la faible probabilité de rupture est estimée avec une précision similaire à celle de AK-MCS, mais en utilisant une taille beaucoup plus petite de la population initiale, ce qui réduit considérablement le temps de calcul. Enfin, dans AK-SS, une technique d'échantillonnage plus efficace 'SS' est proposée. Cette technique ne nécessite pas la recherche de points de conception et par conséquent, elle peut traiter des surfaces d’état limite de forme arbitraire. Toutes les trois méthodes ont été appliquées au cas d'une fondation filante chargée verticalement et reposant sur un sol spatialement variable. Les résultats obtenus sont présentés et discutés.

MOTS-CLÉS : Varabilité spatiale, krigeage, probabilité de ruine, Simulation de Monte Carlo, importance sampling, Subset Simulation.

RÉSUMÉ ÉTENDU

Traditionnellement, l'analyse et le dimensionnement des ouvrages géotechniques reposent sur des approches déterministes. Dans ces approches, les paramètres d'entrée du sol sont représentés par des valeurs conservatrices sans prendre en compte rigoureusement les incertitudes de ces paramètres et sans tenir compte de leur variabilité spatiale. Dans ces approches, un facteur de sécurité global basé sur l'expérience de l'ingénieur est utilisé pour tenir compte des différentes incertitudes.

Ces dernières années, de nombreux efforts ont été consacrés à l'analyse probabiliste des ouvrages géotechniques. Les méthodes probabilistes simplifiées décrivent les différents paramètres incertains du sol par des variables aléatoires où le sol est considéré comme un matériau uniforme. Cependant, dans la nature, les paramètres du sol varient spatialement dans les directions horizontale et verticale en raison des processus de dépôt et post-dépôt. Cela conduit à la nécessité de représenter les paramètres du sol par des champs aléatoires. À cet égard, des approches probabilistes plus avancées ont été proposées dans la littérature. Ces approches sont généralement basées sur la méthode des éléments finis ou la méthode des différences finies.

L'analyse probabiliste des ouvrages géotechniques présentant une variabilité spatiale des propriétés du sol est généralement réalisée à l'aide de la méthode de simulation de Monte Carlo (Monte Carlo Simulation MCS). Cette méthode n'est pas adaptée au calcul des faibles probabilités de rupture rencontrées dans la pratique en raison du grand nombre de simulations nécessaires pour calculer la probabilité de rupture avec une faible valeur du coefficient de variation sur cette probabilité de rupture.

Afin de surmonter les difficultés liées au grand nombre d'appels au modèle mécanique par la méthodologie de Monte Carlo, des techniques d'échantillonnage plus efficaces connues sous le nom de techniques de réduction de variance ont été proposées dans la littérature. On peut citer, entre autres, l'Importance Sampling (IS) et le Subset Simulation (SS). La caractéristique la plus attrayante de ces techniques est que l'on peut obtenir le même niveau de précision que celui de MCS, mais en utilisant un nombre plus réduit d'évaluations de la fonction de performance.

Bien que les techniques de réduction de variance soient remarquablement plus efficaces que MCS, elles ne sont toujours pas pertinentes pour des problèmes où des modèles d'éléments

finis/différences finies sont utilisés pour évaluer la fonction de performance. Par conséquent, un effort considérable a été consacré au développement des techniques dites de metamodeling.

Les techniques de metamodeling permettent de remplacer le modèle mécanique coûteux par un méta-modèle (c'est-à-dire une équation analytique simple). Plusieurs types de techniques de metamodeling peuvent être trouvées dans la littérature, telles que les surfaces de réponse quadratique, les expansions de chaos polynomiales, les réseaux de neurones, le support vector machines et le krigeage. L'idée commune de base de ces méthodes est de construire un méta-modèle en utilisant un certain nombre d'évaluations de la fonction de performance (basé sur le modèle mécanique coûteux) pour un certain nombre d'échantillons désignés comme plan d'expérience (i.e. Design of Experiments DoE). Une fois que le méta-modèle est construit, on peut facilement évaluer la fonction de performance pour n'importe quel échantillon en employant ce méta-modèle au lieu d'utiliser le modèle mécanique coûteux.

Cette thèse vise à développer des approches probabilistes pour l'analyse probabiliste des ouvrages géotechniques en considérant la variabilité spatiale des propriétés de sol. Trois approches probabilistes (AK-MCS, AK-IS et AK-SS) combinant la technique de krigeage et l'une des trois techniques de simulation (comme MCS, IS ou SS) ont été présentées et discutées. L'objectif est de tirer profit des avantages des deux techniques (i.e. metamodeling et réduction de variance) afin d'aboutir à des approches probabilistes plus efficaces nécessitant un temps de calcul réduit. Notons que le but de cette thèse est le calcul de la probabilité de rupture vis-à-vis du poinçonnement d'une fondation superficielle filante reposant sur un sol spatialement variable et soumise à une charge verticale centrée où les paramètres de résistance au cisaillement du sol sont considérés comme deux champs aléatoires anisotropes et non gaussiens. Le modèle mécanique est basé sur des simulations numériques utilisant le code de différences finies FLAC^{3D}.
Nous présentons ci-après un bref aperçu des différents chapitres présentés dans cette thèse sans se préoccuper des résultats obtenus. Puis, nous présenterons dans un dernier paragraphe une synthèse des résultats numériques obtenus par les différentes méthodes probabilistes présentées dans cette thèse.

Chapitre 2:

Avant la présentation des différentes approches probabilistes développées dans cette thèse (voir chapitres 3, 4 et 5), une revue de la littérature est présentée au chapitre 2. Ce chapitre fournit les

types d'incertitudes rencontrées en géotechnique en mettant l'accent sur la variabilité spatiale (caractérisation et discrétisation des champs aléatoires). Il décrit également les principales méthodes probabilistes utilisées dans la littérature pour la propagation de l'incertitude. Les trois méthodes de simulation (Monte Carlo Simulation MCS, Importance Sampling IS et Subset Simulation SS) ainsi que la technique de krigeage sont présentées en détail car elles constituent la base des approches probabilistes développées dans les chapitres suivants. Les principales conclusions du chapitre I sont :

- La combinaison d'une technique de metamodeling et d'une technique de réduction de variance peut conduire à une alternative plus efficace que chacune des méthodes utilisée individuellement. Le krigeage a été choisi dans cette thèse comme technique de metamodeling en raison de l'utilité de ses résultats (moyenne de la prédiction et variance de la prédiction) pour le processus d'apprentissage. En ce qui concerne la technique de réduction de variance, trois types de techniques d'échantillonnage (MCS, IS et SS) ont été utilisés et vérifiés pour leur performance au sein de cette thèse. Ceci fait l'objet des chapitres 3, 4 et 5 respectivement.
- Les méthodes d'expansion en série sont les plus performantes pour la discrétisation des champs aléatoires. La raison est liée au fait que ces méthodes fournissent le nombre optimal de variables aléatoires nécessaires pour discrétiser le champ aléatoire, les autres méthodes étant dépendantes du maillage. Dans cette thèse, la méthode EOLE a été adoptée pour la discrétisation des champs aléatoires. Outre le fait qu'il s'agit d'une méthode d'expansion en série, cette méthode permet de déterminer le nombre minimal de variables aléatoires pour une valeur prescrite de la variance de l'erreur.

Chapitre 3:

Le chapitre 3 a été consacré à la méthode AK-MCS qui est une méthode d'apprentissage alliant krigeage et simulation de Monte Carlo pour évaluer efficacement la probabilité de rupture P_f [Echard et al. (2011)]. La méthode consiste d'abord à générer une population de Monte Carlo de grande dimension suivant les distributions des variables aléatoires. Ensuite, on sélectionne arbitrairement un nombre initial de points parmi les points générés pour être évalués par le modèle mécanique. Le plan, d'expérience et les réponses correspondantes obtenus par le modèle mécanique permettent de construire un modèle de krigeage approximatif moyennant l'outil DACE du logiciel Matlab. Ce modèle de krigeage servira à estimer la prédiction de la fonction de

performance $\check{G}(x)$ ainsi que sa variance $\sigma_{\check{G}}^2(x)$ en tous les points de la population de Monte Carlo. Grâce au calcul des $\check{G}(x)$ et des $\sigma_{\check{G}}^2(x)$ associées, il est possible de construire un critère permettant de classer les points de la population de Monte Carlo par ordre de pertinence pour l'enrichissement du plan d'expérience. Ce critère est défini pour un point donné comme le rapport entre la valeur absolue de la prédiction et son écart-type [i.e. $U(x) = |\check{G}(x)| / \sigma_{\check{G}}(x)$]. Notons que U(x) représente le nombre d'écart-types qui séparent l'estimateur au point x de la surface d'état limite. Les points dont la valeur de U est petite (typiquement inférieures à 2), se trouvent dans une zone très incertaine où le signe du point x est très incertain (i.e. les points très proches de la surface d'état limite). Le point qui minore la fonction U(x) est donc ajouté au plan d'expérience pour être évalué par le modèle mécanique et utilisé pour la construction d'un nouveau modèle de krigeage. L'enrichissement du plan d'expérience se poursuit jusqu'à ce que la valeur de U(x) soit suffisamment grande, c'est à dire que les points restants soient en dehors de la zone incertaine. Dans ce chapitre, le critère d'arrêt du processus d'enrichissement est [min (U)>2]. A la fin du processus d'enrichissement, la probabilité de rupture P_f et le coefficient de variation $COV\left(P_f\right)$ peuvent être estimées. Si le coefficient de variation obtenu n'est pas satisfaisant, le nombre de points de la population de Monte Carlo est augmenté, et la procédure est reprise en partant du dernier plan d'expérience utilisé.

La méthode AK-MCS est très efficace car la probabilité de rupture obtenue est très précise (puisqu'elle correspond à un faible coefficient de variation sur la valeur calculée de la probabilité de rupture) nécessitant un nombre d'appels au modèle mécanique beaucoup plus petit (et donc un temps de calcul réduit) par rapport à la méthodologie de Monte Carlo. Notons cependant que le temps de calcul de l'approche AK-MCS reste important dans le cas des faibles valeurs de la probabilité de rupture (bien que dans cette méthode on utilise les prédictions calculées à l'aide du méta-modèle de krigeage) car une grande population est requise par MCS pour conduire à une petite valeur du coefficient de variation de la probabilité de rupture.

Chapitre 4:

Le chapitre 4 a été consacré à la méthode AK-IS qui est une technique d'apprentissage alliant krigeage et tirage d'importance (Importance Sampling IS). Dans le cadre de cette méthode, l'inconvénient majeur de la procédure AK-MCS impliquant la grande population nécessaire pour évaluer les faibles probabilités de rupture a été surmonté en utilisant une technique

d'échantillonnage plus efficace. En effet, le tirage d'importance a été employé à la place de la technique d'échantillonnage de Monte Carlo. Notons que la procédure AK-IS a été proposée par Echard et al. (2013) dans le cas simple d'une fonction de performance analytique. Dans cette thèse, une extension de la méthode AK-IS au cas d'un sol spatialement variable est présentée. Cette méthode se compose de deux étapes principales. Tout d'abord, le point de rupture le plus probable (point de conception) est déterminé suivant une procédure itérative à l'aide d'un méta-modèle de krigeage approximatif basé sur un petit nombre d'échantillons. Ensuite, le méta-modèle approximatif de krigeage obtenu précédemment est successivement amélioré par un processus d'enrichissement en ajoutant chaque fois un nouvel échantillon sélectionné à partir d'une fonction de densité de probabilité centrée au point de conception. Lorsque l'apprentissage s'arrête, les valeurs estimées de la probabilité de ruine P_f et du coefficient de variation sont déterminées. Notons que la méthode AK-IS utilise l'avantage du krigeage (moyenne et variance de la prédiction) pour le choix du meilleur nouveau point à calculer par le modèle mécanique. Elle utilise aussi l'avantage de l'importance sampling pour générer des points dans la zone de défaillance. Il convient de noter ici que le critère d'arrêt d'apprentissage est similaire à celui présenté dans la procédure AK-MCS [i.e. min (U)>2]. Il convient aussi de souligner que le temps de calcul des prédictions est significativement plus petit que ce qui serait nécessaire si l'on utilisait l'approche AK-MCS où un échantillonnage MCS aurait été utilisé pour déterminer les échantillons candidats. En effet, le nombre d'échantillons candidats utilisés dans AK-IS est beaucoup plus faible pour la même valeur du coefficient de variation sur P_f conduisant ainsi à une réduction considérable du temps de calcul.

Chapitre 5:

La méthode AK-SS est une méthode d'apprentissage alliant krigeage et subset simulation (SS) pour évaluer efficacement la probabilité de ruine P_f . Ce travail a été motivé par le fait que certaines surfaces d'état limite peuvent présenter plusieurs points de conception et par conséquent, l'application de la procédure AK-IS n'est pas simple pour ces cas. Dans l'approche AK-SS, une technique d'échantillonnage plus efficace (Subset Simulation SS) a été utilisée à la place de la technique d'échantillonnage par Importance Sampling. Cette technique permet de surmonter la recherche des points de conception (comme c'est le cas dans AK-IS). Par conséquent, elle peut traiter des surfaces d'état limite de forme arbitraire. La procédure AK-SS se compose de deux étapes principales. Dans la première étape, un méta-modèle approximatif est

construit par krigeage en utilisant un faible nombre d'échantillons appelés plan d'expérience initial (Initial Design of Experimnts (DoE)). Ces échantillons sont choisis arbitrairement parmi une grande population générée par MCS. Dans la deuxième étape, l'approche de Subset Simulation est utilisée pour générer des échantillons qui sont dirigés vers la surface d'état limite en employant le méta-modèle de krigeage approximatif obtenu précédemment. Le méta-modèle approximatif est successivement amélioré par un processus d'enrichissement en ajoutant à chaque fois un échantillon supplémentaire au DoE initial. Le nouvel échantillon est sélectionné (en utilisant la fonction d'apprentissage U) parmi les échantillons obtenus au dernier niveau d'un échantillonnage par Subset Simulation appliqué sur le méta-modèle mis à jour. Le processus d'ajout d'un nouvel échantillon (processus d'enrichissement) est répété jusqu'à ce qu'un critère prescrit sur la valeur de la probabilité de rupture soit obtenu. Dans ce chapitre, le critère d'arrêt récemment proposé par Schöbi et al. (2015) est utilisé. Ce critère se base non pas sur la précision du méta-modèle mais plutôt sur la précision du paramètre de sortie (la probabilité de rupture dans le cas présent). À la fin du processus d'enrichissement, le nombre d'échantillons ajoutés est considéré comme suffisant pour pouvoir calculer la valeur finale de la probabilité de rupture et la valeur correspondante du coefficient de variation en utilisant le méta-modèle final.

La méthode AK-SS proposée dans ce chapitre utilise les avantages de la méthode SS pour l'évaluation de petites probabilités de défaillance (moyennant un temps de calcul plus faible que la méthode MCS) et du modèle de krigeage (qui fournit la moyenne et la variance de la prédiction pour l'approximation de la fonction de performance). Le principal avantage de AK-SS sur AK-MCS devient visible en cas de faibles probabilités de défaillance. Comme il a été mentionné précédemment, l'effort de calcul de la procédure AK-MCS augmente considérablement dans le cas où des grandes populations sont nécessaires pour évaluer des faibles probabilités de ruine. AK-SS résout ce problème en exprimant la faible probabilité de ruine en tant que produit de probabilités de ruine conditionnelles plus importantes de plusieurs événements de défaillance intermédiaires.

Résultats numériques:

Cette section vise à présenter l'effet de la variabilité spatiale du sol sur la probabilité de ruine vis-à-vis du poinçonnement et ce, pour une fondation superficielle filante soumise à une charge verticale centrée. La fonction de performance utilisée dans l'analyse est donnée par l'équation suivante:

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

où q_u est la capacité portante ultime calculée en utilisant le logiciel FLAC^{3D} et q_s est la charge appliquée à la fondation. La cohésion c et l'angle de frottement interne φ sont considérés comme deux champs aléatoires anisotropes non-Gaussiens. L'angle de frottement interne φ est supposé suivre une loi bêta, et la cohésion c est supposée lognormale. Les valeurs moyennes et les coefficients de variation des deux champs aléatoires sont donnés comme suit: $\mu_c = 20kPa, Cov_c = 25\%; \mu_\varphi = 30^o, Cov_\varphi = 10\%$. La fondation superficielle filante de largeur B=1m et de 0.25m de hauteur est supposée non pesante et élastique. Le domaine du sol adopté dans l'analyse est de 13B de large par 5B de profondeur. Nous présentons ci-dessous une comparaison entre les résultats obtenus par les trois méthodes AK-MCS, AK-IS et AK-SS.

a. Effet des distances d'autocorrélation des deux champs aléatoires sur P_f et $COV\left(P_f\right)$

La figure 1 présente l'effet de la distance d'autocorrélation isotrope (a_x = a_y) sur la probabilité de défaillance obtenue à partir des trois méthodes AK-MCS, AK-IS et AK-SS. D'autre part, la figure 2 et la figure 3 présentent respectivement l'effet de la distance d'autocorrélation verticale et horizontale sur la probabilité de défaillance calculée par les trois méthodes.

Les figures 1, 2 et 3 montrent qu'il existe un bon accord entre les résultats obtenus à partir des trois méthodes. Pour la quasi majorité des configurations traitées, l'écart maximum par rapport à la méthode AK-MCS (considérée comme une référence) reste inférieur à 7%. Comme on peut le voir dans ces figures, la probabilité de ruine est maximale pour un sol homogène. Ceci montre l'intérêt de calculer la probabilité de ruine en prenant en compte la variabilité spatiale des propriétés du sol.

Les figures 4, 5 et 6 montrent les valeurs du coefficient de variation obtenues à partir des trois méthodes AK-MCS, AK-IS et AK-SS pour les différentes variabilités du sol. A partir de ces figures, on peut observer que les valeurs obtenues à partir de la méthode AK-SS sont inférieures à 4% pour tous les cas traités. Notons cependant que, pour les deux autres méthodes (i.e. AK-MCS et AK-IS), les valeurs du coefficient de variation restent inférieures à 7% pour la majorité des cas considérés. En conclusion, les petites valeurs du coefficient de variation obtenues (<7%) indiquent que les valeurs de P_f calculées dans cette thèse sont fiables et peuvent être utilisées dans la pratique.

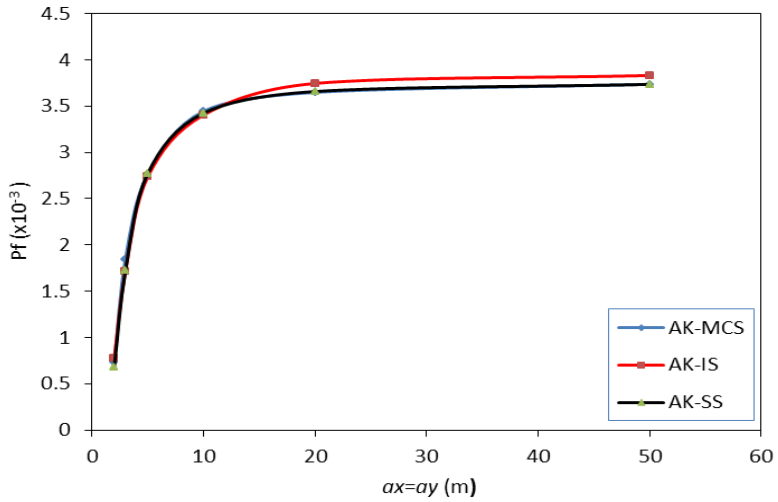


Figure 1. Effet de la distance d'autocorrélation isotrope a_x=a_y sur P_f

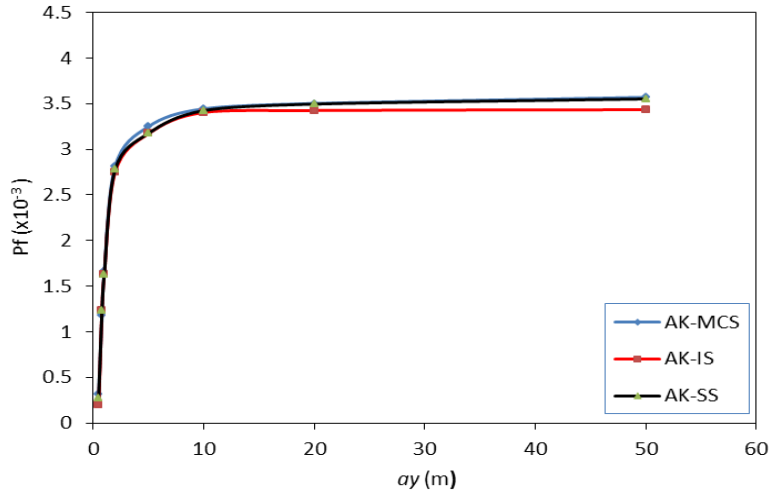


Figure 2. Effet de la distance d'autocorrélation verticale a_y sur P_f pour a_x =10 m

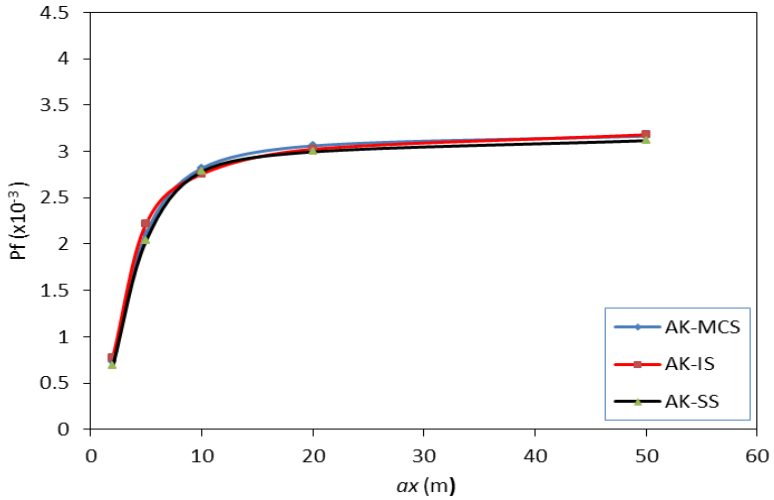


Figure 3. Effet de la distance d'autocorrélation horizontale a_x sur P_f pour a_y=2 m

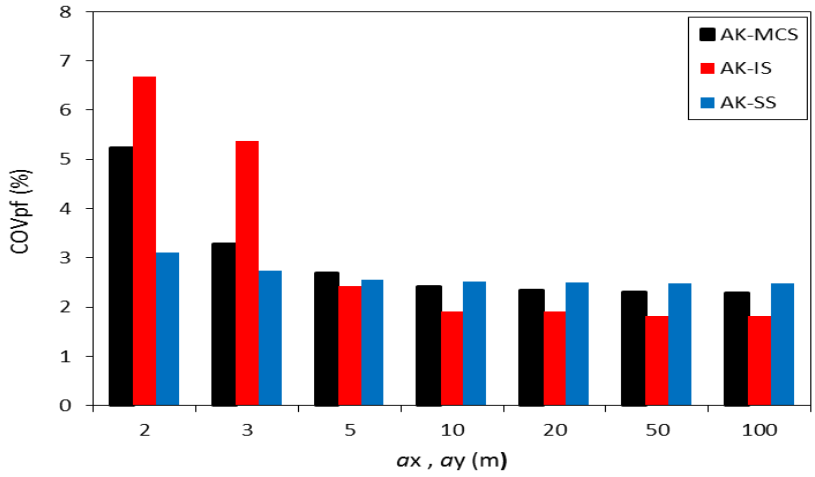


Figure 4. Valeurs de $COV\left(P_f\right)$ pour différentes valeurs de la distance d'autocorrélation isotrope a_x=a_y

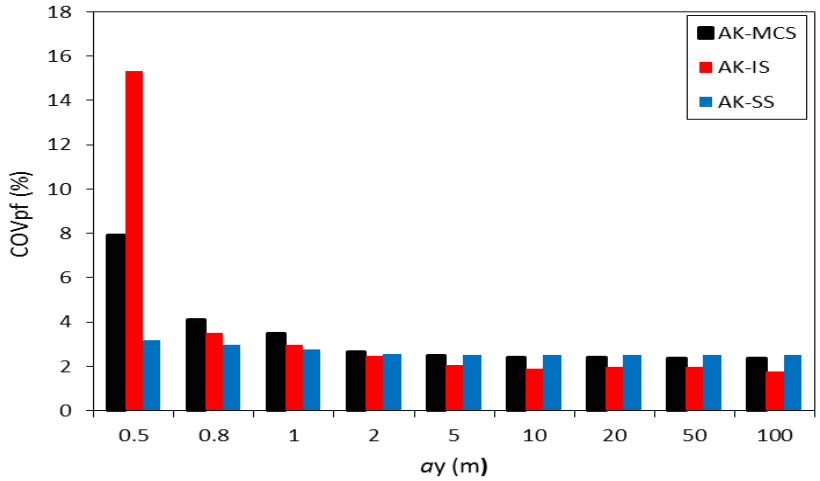


Figure 5. Valeurs de $COV\left(P_f\right)$ pour différentes valeurs de la distance d'autocorrélation verticale a_y pour a_x =10 m

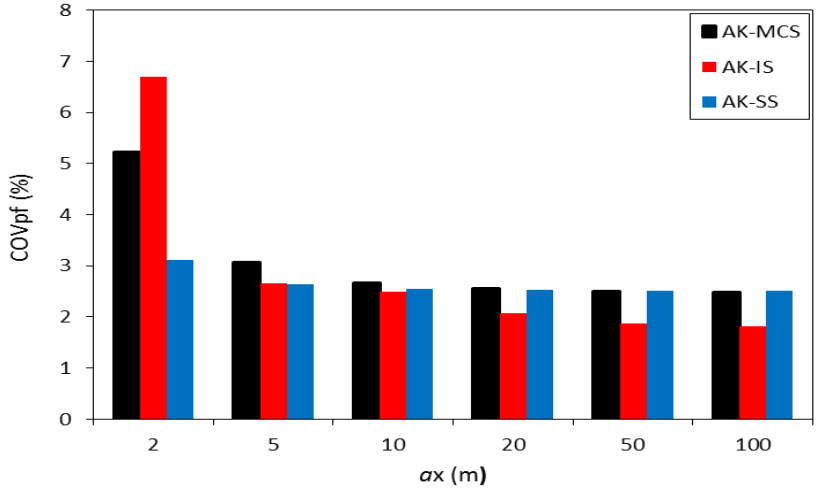
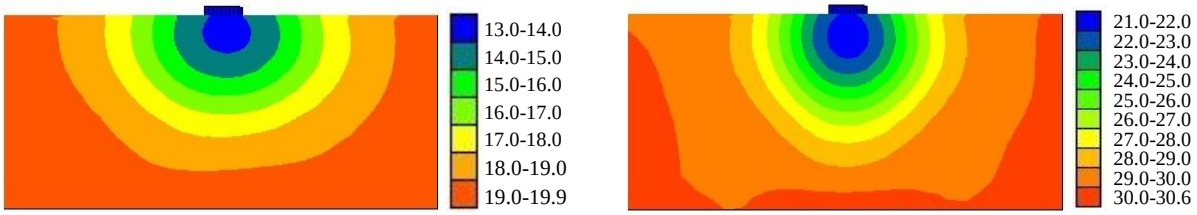


Figure 6. Valeurs de $COV\left(P_f\right)$ Pour différentes valeurs de la distance d'autocorrélation horizontale a_x pour a_y=2 m

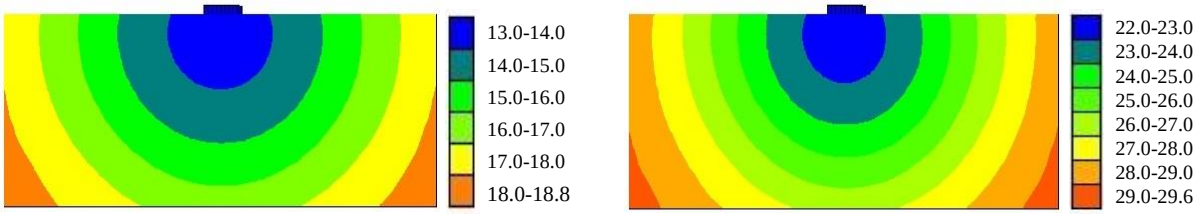
b. Réalisation critique au point de conception

La figure 7 montre les réalisations critiques des champs aléatoires au point de conception pour deux valeurs de la distance d'autocorrélation isotrope. Cette figure montre une répartition symétrique des paramètres de résistance au cisaillement du sol et ce par rapport à l'axe vertical central de la fondation. La zone de sol de faibles caractéristiques mécaniques est concentrée autour de la fondation, le sol de plus fortes valeurs de caractéristiques mécaniques étant loin de la fondation. Comme on peut le constater à partir de la figure 7, la taille de la zone de sol de faibles caractéristiques mécaniques augmente avec l'augmentation de la longueur d'autocorrélation conduisant à une plus grande probabilité de défaillance.



Cohésion (kPa) **Angle de friction (degrés)**

(a) $a_x=a_y$**=2m**



Cohésion (kPa) **Angle de friction (degrés)**

(b) $a_x=a_y$**=5m**

Figure 7. Réalisations critiques pour deux valeurs de la distance d'autocorrélation isotrope

c. Synthèse critique des différentes méthodes probabilistes

Trois méthodes probabilistes (AK-MCS, AK-IS et AK-SS) basées sur le krigeage ont été présentées dans cette thèse pour l'analyse probabiliste d'une fondation superficielle filante reposant sur un sol spatialement variable. Ces méthodes sont très efficaces car la probabilité de défaillance obtenue est très précise, nécessitant un nombre d'appels réduit au modèle mécanique par rapport à une technique de simulation utilisée individuellement telles que Monte Carlo Simulations (MCS), Importance Sampling (IS) et Subset Simulation (SS).

Les trois méthodes AK-MCS, AK-IS et AK-SS ont été comparées en termes de temps de calcul et du nombre total d'appels au modèle mécanique et ce, dans le cas pratique où la distance d'autocorrélation horizontale est égale à 10m et la distance d'autocorrélation verticale est égale à 2m. Les trois méthodes ont fourni des valeurs quasi similaires pour les sorties probabilistes. Notez cependant que la méthode AK-IS était la plus avantageuse en termes de temps de calcul (environ un jour) suivie d'AK-SS (environ deux jours) puis de AK-MCS (environ 4 jours).

En conclusion, la méthode AK-IS devrait être utilisée en cas de présence d'un seul et unique point de conception. Pour les problèmes présentant des surfaces d'état limite ayant plus d'un point de conception, la méthode AK-IS perd son intérêt et, par conséquent, on peut surmonter cet inconvénient en utilisant l'approche AK-SS qui est légèrement plus chère en temps de calcul.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

NOTATION

The following acronyms and symbols are used in this thesis:

ACRONYMS

ACF	Autocorrelation Function
AK-IS	Active learning reliability method combining Kriging and Importance Sampling
AK-MCS	Active learning reliability method combining Kriging and Monte Carlo Simulation
AK-SS	Active learning reliability method combining Kriging and Subset Simulation
ANN	Artificial Neural Networks
BLUP	Best Linear Unbiased Predictor
CDF	Cumulative Density Function
CIUC	Consolidated Isotropic Undrained triaxial Compression test
COV	Coefficient Of Variation
CPT	Cone Penetration Test
DMT	Dilatometer Test
DoE	Design of Experiments
DST	Direct Shear Test
EOLE	Expansion Optimal Linear Estimation method
FORM	First Order Reliability Method
IP	Integration Point method
IS	Importance Sampling
ISD	Importance Sampling Density
KL	Karhunen-Loeve expansion method
LAS	Local Average Subdivision
LH	Latin Hypercube
LSS	Limit State Surface
MCS	Monte Carlo Simulation
MP	Midpoint method
OLE	Optimal Linear Estimation method
OSE	Orthogonal Series Expansion method
PCE	Polynomial Chaos Expansion
PDF	Probability Density Function

P_f	Probability of failure
PMT	Pressuremeter Test
RSM	Response Surface Methodology
SA	Spatial Average method
SF	Shape Function method
SORM	Second Order Reliability Method
SPCE	Sparse Polynomial Chaos Expansion
SPT	Standard Penetration Test
SS	Subset Simulation
SVM	Support Vector Machine
UC	Unconfined Compression test
UCS	Unconfined Compression Strength
ULS	Ultimate Limit State
UU, TUU	Unconsolidated–Undrained triaxial compression test
VST	Vane Shear Test

Latin letters

a	Autocorrelation distance (m)
a_x	Horizontal autocorrelation distance (m)
a_y	vertical autocorrelation distance (m)
a_β	unknown PCE coefficients
a_i,b_i	coefficients obtained by the least squares method used in the RSM method
B	width of footing (m)
c	Soil cohesion (kPa)
C_c	Compression index
C_v	Coefficient of consolidation
C_j	intermediate failure threshold, chapter 2
D_r	relative density, Table (2.2), chapter 2
e	initial void ratio
E	Young's modulus (kPa)
E_D	Dilatometer modulus
F	deterministic part defined by a regression model, Eq. (2.27), chapter 2
f_s	friction of sleeve measured by CPT (kN/m^2)

F_R	normalized friction ratio $= \frac{f_s}{q_c - \sigma_{v0}}.100$
F_s	safety factor
G	Performance function
G	non-Gaussian marginal cumulative density function, Eq. (2.6), chapter 2
$\hat{G}$	realization of a random function, Eq. (2.27), chapter 2
h_x	optimal probability density function Eq. (2.16), chapter 2
Δh	separation distance between the data pairs, Eq. (2.1), chapter 2
I_D	Dilatometer material index
I(X)	indicator function of the failure domain, Eq. (2.14), chapter 2
k	constant sets the confidence level, Eqs. (5.3) and (5.4), chapter 5
K	maximum number of lags, Eq. (2.1); number of samples, Eq. (2.13), chapter 2
K_D	Dilatometer horizontal stress index
K_s	Shear stiffness of the interface (GPa)
K_n	Normal stiffness of the interface (GPa)
m	Number of levels in subset simulation method
M	Number of terms (expansion order) retained in the EOLE method
n	Porosity of soil
N	Number of SPT blows
N	Total number of data samples, chapter 2
N_1	initial Design of Experiment
N_2	number of iterations, chapter 4
$N\gamma$, Nq , Nc	Bearing capacity factors
N_{MC}	Number of Monte Carlo samples
N_{IS}	Number of samples used in chapter 4
N_{SS}	Number of samples per level used in chapter 5
P_0	intermediate failure probability, chapter 5
P_f	Probability of failure
P_f^0	original failure probability based on the kriging predictor values, chapter 5
P_f^+	upper boundary of the failure probability, chapter 5
P_f^-	lower boundary of the failure probability, chapter 5

$\bar{p}_m$	maximum past pressure (kN/m^2)
q_c	cone tip resistance measured by CPT (kN/m^2)
q_s	applied footing load (kN/m^2)
q_T	corrected cone tip resistance measured by CPT (kN/m^2)
q_u	ultimate bearing capacity (kN/m^2)
q_{c1N}	normalized cone tip resistance
R_D	Relative Density, Table (2.3), chapter 2
R	square matrix of dimension $N \times N$, Eqs. (2.31) and (2.32), chapter 2
s	total number of grid points in the EOLE method.
S_u	undrained shear strength (kPa)
U	learning function
u_1, u_2	two standard normal random variables
u_c	Standard Gaussian random variable for random field c
u_φ	Standard Gaussian random variable for random field φ
w	water content
w_l	liquid limit
w_P	plastic limit
X	vector that represents the location in chapter 2
X	vector that represents random variables, chapter 2
Z	The property of interest, chapter 2
$Z(X_i)$	value of the property Z at location X_i , chapter 2
$Z(X_{i+k})$	value of the property Z at the location X_{i+k} , chapter 2
$Z(X)$	fluctuation around the mean value, Eq. (2.27), chapter 2

Greek letters

β_{HL}	Reliability index
β	scalar deterministic part, Eq. (2.31), chapter 2
γ	unit weight of soil (kN/m^3)
δ	Scale of fluctuation (m)
ε_{P_f}	tolerance for two consecutive iteration steps, Eq. (5.2), chapter 5
θ	correlation parameter vector, Eq. (2.30), chapter 2

λ	eigenvalues of the autocorrelation matrix
μ	Mean value of the random field
μ_Z	The mean of the property Z , chapter 2
$\mu_{\hat{G}(X)}$	Kriging mean prediction
ν	Poisson's ratio
ξ	Standard normal random variable
ρ	Coefficient of correlation
σ	Standard deviation value of the random field
σ_{v0}	total vertical stress (kN/m^2)
$\sigma^2_{\hat{G}(X)}$	Kriging prediction variances
σ_Z^2	Random field variance, Eq. (2.32), chapter 2
φ	Angle of internal friction of soil (°)
φ'	Effective angle of internal friction of soil (°)
ϕ_j	eigenvector of the autocorrelation matrix
$\Phi(.)$	standard normal cumulative density function
χ	vector data samples represents the values of the property Z in chapter 2
ψ	soil dilation angle (°)
$\Sigma^G_{\chi;\chi}$	Gaussian autocorrelation matrix
$\Sigma^{NG}_{\chi;\chi}$	non-Gaussian autocorrelation matrix

CHAPTER 1. GENERAL INTRODUCTION

Traditionally, the analysis and design of geotechnical structures is based on deterministic approaches. In these approaches, the soil input parameters are represented by conservative values without rigorously taking into account the uncertainties of these parameters and without considering their spatial variability. Indeed, within these approaches, a global safety factor based on engineering judgment is used to take into account the different uncertainties.

In recent years, much effort has been paid to the probabilistic analysis of geotechnical structures. Simplified probabilistic methods describe the different soil uncertain parameters by random variables where the soil is considered as a uniform material. However, in nature, the soil parameters vary spatially in both the horizontal and vertical directions as a result of depositional and post-depositional processes. This leads to the necessity of representing the soil parameters by random fields. In this regard, more advanced probabilistic approaches were proposed in the literature. These approaches are generally based on the finite element or the finite difference method.

The probabilistic analysis of geotechnical structures presenting spatial variability in the soil properties is generally performed using Monte Carlo Simulation (MCS) methodology. This method is not suitable for the computation of the small failure probabilities encountered in practice. This is because it is very time-expensive in such cases due to the large number of simulations required to calculate the failure probability to within a small value of the coefficient of variation of this failure probability.

In order to overcome the shortcoming related to the excessive number of calls of the mechanical model by MCS methodology, more efficient sampling techniques known as variance reduction techniques were proposed in literature. One may cite among others the Importance Sampling (IS)

and Subset Simulation (SS) techniques. The most attractive feature of these techniques is that one can obtain the same accuracy level as that of MCS but using a significantly smaller number of evaluations of the performance function. Although the variance reduction techniques are remarkably more efficient than MCS, they are still impractical to be used in problems where computationally expensive finite element/finite difference models are employed to evaluate the performance function. Therefore, a considerable amount of effort has been devoted to developing so-called meta-modeling techniques. These techniques allow one to substitute the computationally expensive mechanical model by a meta-model (i.e. a simple analytical equation). Several kinds of meta-modeling techniques can be found in literature such as quadratic response surfaces, polynomial chaos expansions, neural networks, support vector machines and kriging. The basic common idea of these methods is to build a meta-model by using a certain number of evaluations of the performance function (based on the computationally expensive mechanical model) for a number of samples designated as Design of Experiments DoE. Once the meta-model is constructed, one may easily evaluate the performance function for any sample by employing this meta-model instead of using the computationally expensive mechanical model.

In this thesis, the kriging metamodeling technique is used in combination with a simulation method (e.g. MCS, IS or SS) to perform the probabilistic analysis. The aim is to take benefit from the advantages of both techniques in order to lead to more efficient probabilistic approaches needing a reduced computation time as compared to each technique employed individually. Notice that the aim of this thesis is the computation of the probability of failure against soil punching of a strip footing resting on a spatially varying soil and subjected to a vertical load where the soil shear strength parameters were considered as two anisotropic non-Gaussian random fields.

Before the presentation of the different probabilistic approaches developed in this thesis (cf. chapters 3, 4 and 5), a literature review is presented in chapter 2. This chapter provides the types of uncertainties encountered in geotechnical engineering with an emphasis on the soil spatial variability (characterization and discretization of random fields). It also describes the principal probabilistic methods used in literature for the uncertainty propagation. The three simulation methods (Monte Carlo Simulation MCS, Importance Sampling IS and Subset Simulation SS) together with the kriging metamodeling technique are presented in some detail because they constitute the basis of the probabilistic approaches developed in the following chapters.

Chapter 3 presents an Active learning reliability method combining Kriging and Monte Carlo Simulation (called AK-MCS) for the probabilistic analysis of geotechnical structures involving spatially varying soil properties. This method is an extension to the case of random fields, of the AK-MCS approach proposed by Echard et al. (2011) in the case where the uncertain parameters are modeled by random variables. Within this method, one performs a Monte Carlo simulation without evaluating the whole population using the original computationally expensive mechanical model. Indeed, the population is predicted using a kriging meta-model which is defined using only a few points of the population that are evaluated employing the mechanical model. As a result, AK-MCS significantly reduces the number of calls of the mechanical model (and thus the computation time) with respect to MCS. The approach proposed in this chapter and the two other approaches developed in the two following chapters are applied to a strip footing resting on a spatially varying soil and subjected to a prescribed vertical load. The objective is the computation of the probability of failure of the footing against soil punching. The ultimate aim of the thesis is to test the performance of the different probabilistic approaches in terms of computation time.

Chapter 4 presents an Active learning reliability method combining Kriging and Importance Sampling (called AK-IS) for the probabilistic analysis of geotechnical structures involving

spatially varying soil properties. This method is an extension to the case of random fields, of the AK-IS approach proposed by Echard et al. (2013). In this approach, a more efficient sampling technique (Importance Sampling) is used instead of the Monte Carlo sampling technique employed in Chapter 3. Indeed, the computation time of AK-MCS remains important when computing the small failure probabilities encountered in practice (although in this method, one makes use of the predictions computed using the kriging meta-model) since a large population is required by MCS to lead to a small value of the coefficient of variation on the failure probability. In the framework of the present AK-IS approach, the small failure probability can be estimated with a similar accuracy as AK-MCS but using a much smaller size of the initial population, thus significantly reducing the computation time with respect to AK-MCS.

Chapter 5 presents an Active learning reliability method combining Kriging and Subset Simulation (called AK-SS). This work was motivated by the fact that certain limit state surfaces may present several design points and thus, the application of AK-IS is not straightforward for those cases. Within this approach, a more efficient sampling technique (Subset Simulation SS) is used instead of Importance Sampling. This technique allows one to overcome the search of the design points within AK-IS and thus it can deal with arbitrary shapes of the limit state surface.

The thesis ends by a general conclusion of the principal results obtained from the analyses.

CHAPTER 2. LITERATURE REVIEW

2.1 INTRODUCTION

The analysis and design of geotechnical structures are generally based on deterministic approaches. In these approaches, the mean values of the input soil properties are used without considering the uncertainties of these parameters and their spatial variability. Also, the output parameter (ultimate load, safety factor, displacement, etc.) is given by a single mean value without any information on the uncertainty level associated with this parameter. Engineers try to solve these problems using the concept of the global safety factor, but this factor cannot explicitly deal with uncertainty.

In recent years, reliability analyses and probabilistic methods have been applied in order to provide a rational mathematical framework to incorporate the different types of uncertainties into a geotechnical design. They aim at rigorously evaluating the system response in the form of a PDF taking into account the uncertainties in the soil parameters together with their spatial variability. Addressing uncertainty in the system response does not in itself increase the level of safety, but it allows the engineer to rationally evaluate the performance level of the geotechnical system. Being able to rigorously determine the performance level and reduce the undesired conservatism is generally beneficial in the economic sense.

In traditional probabilistic analyses, the uncertain parameter is interpreted as a random variable defined only by its probability density function (PDF). In other words, the soil is considered as a homogeneous material having the same random value of the uncertain parameter in the entire soil domain. However, in nature, the soil parameters (shear strength parameters, elastic properties, etc.) vary spatially in both the horizontal and vertical directions as a result of depositional and post-depositional processes. This leads to the necessity of modeling soil parameters by random fields characterized not only by their marginal PDFs, but also by their autocorrelation functions.

After the input uncertainties have been appropriately modeled by random variables or random fields, the task remains to quantify the influence of these uncertainties on the output of the model. This task is referred to as the uncertainty propagation. In other words, the uncertainty propagation aims to study the impact of the input uncertainty on the probabilistic outputs. The probabilistic outputs may be the statistical moments (mean and standard deviation) of the system response or the failure probability (or the reliability index) against a given threshold of this response. In this thesis, our focus is to compute the failure probability against a given threshold.

During recent years, different approaches (especially the meta-modeling techniques) were developed for the uncertainty propagation. These approaches allow one to substitute the computationally expensive mechanical model by a meta-model (i.e. a simple analytical equation). The meta-modeling approaches are detailed later in this chapter. Of particular interest is kriging metamodeling. The kriging metamodeling technique was used in this thesis in combination with different simulation techniques (i.e. Monte Carlo Simulation MCS, Importance Sampling IS or Subset Simulation SS). The objective was to perform a probabilistic analysis with a reduced computation time as compared to each technique employed individually.

The aim of this thesis is to investigate the effect of the soil spatial variability on the failure probability against soil punching of a strip footing resting on a spatially varying soil and subjected to a vertical load.

This chapter aims at first presenting the different sources of uncertainties related to the geotechnical parameters with an emphasis on the soil spatial variability. Then, some probabilistic methods that are used in the upcoming chapters for the computation of the failure probability are presented and discussed. The chapter ends with a conclusion.

2.2 UNCERTAINTY IN GEOTECHNICAL ENGINEERING

Uncertainty is common in almost every field of engineering, and geotechnical engineering is no exception. Natural soils are heterogeneous and anisotropic in physical properties due to their composition and complex depositional processes.

It is generally agreed that the uncertainty in geotechnical engineering can be divided into two main categories. These are aleatory uncertainty and epistemic uncertainty [Vanmarcke (1977); Baecher et al. (1983); Tang (1984); Phoon and Kulhawy (1999); Der Kiureghian and Ditlevsen (2009)].

In geotechnical engineering, two types of epistemic uncertainties can be faced: The measurements and the transformation uncertainties. The first one is due to the sampling error that results from limited amount of information. The second one is introduced when field or laboratory measurements are transformed into design soil properties using empirical or other correlation models. As for the aleatory (inherent) uncertainty, it primarily results from the natural geologic processes. In this thesis, only aleatory uncertainty which is the spatial variability of the soil properties is considered.

2.3 SPATIAL VARIABILITY OF THE SOIL PROPERTIES

Soils are geological materials formed by weathering, erosion and sedimentation processes. When transported by physical means to their present locations [Baecher and Christian (2003)], soils have been subjected to various stresses, and physical and chemical changes. Thus, it is hardly surprising that the physical properties of soils vary from place to place within resulting deposits. It should be mentioned that the spatial variability of the soil properties occurs at different scales of variability depending on the type of problem. Several authors [Vanmarcke (1978); Burrough (1983); Jaksa (1995); Fenton (1999); Borja (2011); Christakos (2012); Chen et al. (2012) among others] have recognized the multiple scales of soil variability ranging from the microstructure

scale to the regional scale (i.e. from the micro level at the grain size scale to the geological scale of several tens and hundreds of meters as shown in Figure (2.1)). In this context, heterogeneity can be defined as the opposite of homogeneity and is further used as a synonym of spatial variability. The geotechnical level is between the specimen scale and the geological scale; therefore, it is important to keep in mind that there is not a single spatial scale, but multiple spatial scales contributing to soil variability. Of course, this plays a role in the evaluation of spatial variability of soil properties as well in the evaluation of its effects on the system response.

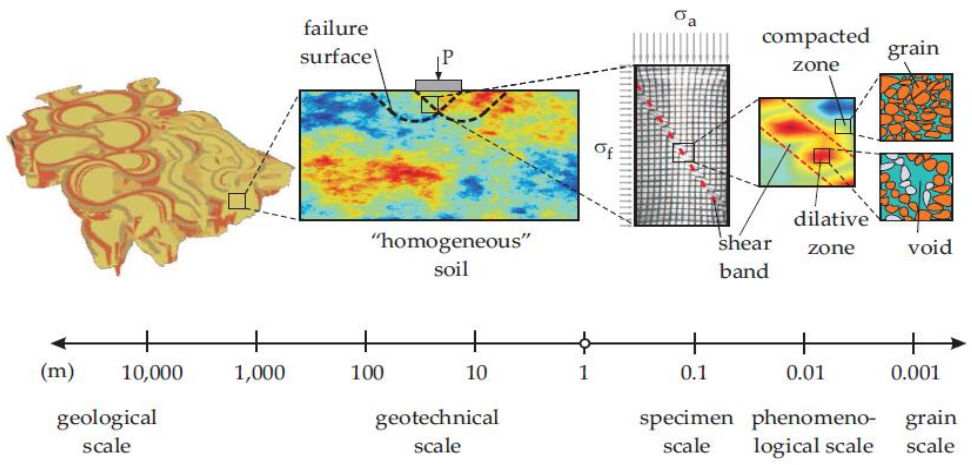


Figure 2.1. Illustration of the multi-scale nature of soil spatial variability [after Borja (2011), Chen et al. (2012), and Christakos (2012)]

When dealing with geotechnical structures, the inherent spatial variability has been considered as a major source of uncertainties in soil properties [Phoon and Kulhawy (1999); Baecher and Christian (2003)]. It significantly affects the performance of geotechnical structures that, under a probabilistic framework, is commonly measured by the probability of failure P_f (or reliability index β_{HL}). Note that the determination of the mean and standard deviation of the soil property is performed using conventional statistical analysis. This analysis provides the variability of the soil property; however, it does not provide the spatial variability of this property. Thus; to characterize the spatial variation of the soil property, one needs to determine the autocorrelation

function and the corresponding value of the autocorrelation distance (*a*). For this purpose, two mathematical techniques can be found in literature to identify the autocorrelation structure of a soil property. These are the random field theory and the geostatistical tools. In this chapter, only the random field theory is presented. For the geostatistical tools, the reader may refer to [Journel and Huijbregts (1978)].

2.3.1 Random field theory

Soil properties at each location within the soil mass are considered to be random variables and typically exhibit considerable variation from point to point. Therefore, it is essential to consider the spatial variability of the soil domain. Vanmarcke (1977,1983) provided a major contribution to the study of spatial variability of geotechnical materials using random field theory. In order to describe a soil property stochastically, Vanmarcke (1983) stated that three statistical parameters need to be described: (i) the mean; (ii) the variance (or standard deviation or coefficient of variation); and (iii) the autocorrelation distance (*a*) (or the scale of fluctuation). The scale of fluctuation accounts for the distance within which the soil property shows relatively strong correlation from one point to another. When the soil property is plotted as a function of the distance, the scale of fluctuation is related to the distance between the intersections of the trend and the fluctuating soil property [i.e. the distance δ_v in Figure (2.2.a) where a typical spatially variable soil profile showing the trend, the fluctuating component and the vertical scale of fluctuation is presented]. Small values of scale of fluctuation imply rapid fluctuations about the mean (i.e. a highly heterogeneous soil mass), whereas large values suggest a slowly varying property, with respect to the average. Vanmarcke (1977) demonstrated a simple procedure to evaluate an approximate value of the vertical scale of fluctuation (which is approximately equal to 0.8 times the average distance between the intersections of the trend and the fluctuating soil property), as shown in Figure (2.2.b).

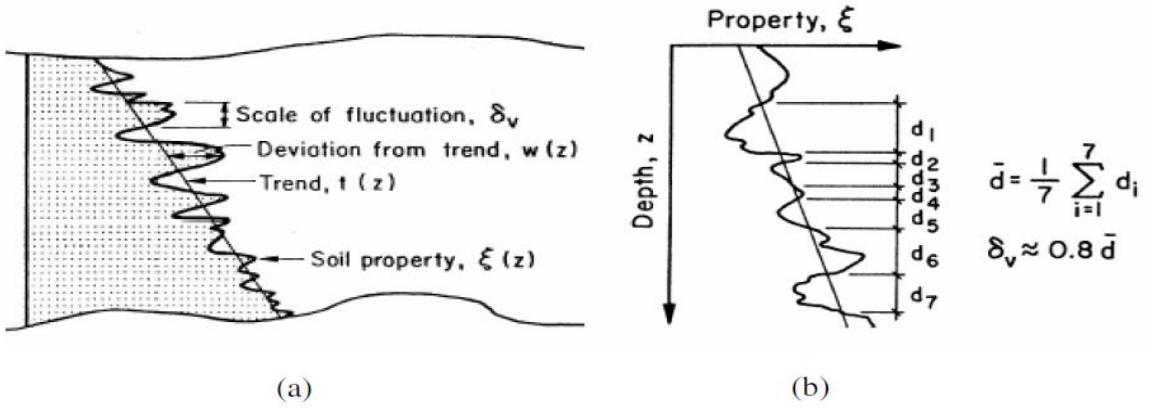


Figure 2.2. (a) Definition of various statistical parameters of a soil property [Phoon and Kulhawy (1996)]; (b) approximate definition of the scale of fluctuation [Vanmarcke (1977)]

As a conclusion, a random field is a conceivable model to characterize the continuous spatial fluctuations of a soil property within a soil unit. It is fully described by the autocorrelation function (ACF) together with the PDF of the uncertain parameter. The ACF is introduced in addition to the classical statistical parameters (i.e. the mean and standard deviation or coefficient of variation). It is used to identify the autocorrelation distance (a) or the scale of fluctuation (δ).

2.3.1.1 Identification of the autocorrelation function and the corresponding values of the autocorrelation distances

The autocorrelation function (ACF) is often used to determine the distance over which a property exhibits strong correlation. Two measures of this distance which are the autocorrelation distance (a) or the scale of fluctuation (δ) may be evaluated. In practice, the ACF must be estimated from some available data samples gathered in a vector $\chi = \{Z(X_1),...,Z(X_N)\}$ where N is the number of these data samples and Z is the property of interest. The sample autocorrelation function is given as follows:

$$\rho_k = \rho(k\,\Delta h) = \frac{\sum_{i=1}^{N-k}\left[Z(X_i)-\mu_Z\right]\left[Z(X_{i+k})-\mu_Z\right]}{\sum_{i=1}^{N}\left[Z(X_i)-\mu_Z\right]^2} \qquad k=0,\ 1,\ \ldots,\ K \tag{2.1}$$

where X is a vector that represents the location; $Z(X_i)$ is the value of the property Z at location X_i; $Z(X_{i+k})$ is the value of the property Z at the location, X_{i+k}; Δh is the separation distance between the data pairs and μ_Z is the mean of the property Z. The sample ACF is thus the graph of ρ_k for lags $k = 0, 1, 2, \ldots K$, where K is the maximum number of lags (data intervals) that ρ_k should not be calculated beyond. While the sample autocorrelation function can be evaluated for all lags up to $N-1$, it is not advisable since, as k tends towards N, the number of pairs reduces and, as a consequence, the reliability of the estimate of the true autocorrelation function ρ_k also decreases. Generally, $K = N/4$ [Box and Jenkins (1970)], where N is the total number of data samples.

The accuracy of the autocorrelation function ACF depends on the number N of data samples. Referring to Jaksa (1995), the minimum number of data samples (observations) has received a little effort in literature. Box and Jenkins (1970), Anderson (1976), and Davis (1986) recommended at least 50 observations. Lumb (1975) suggested that, for a full three-dimensional analysis, the minimum number of observations is of order 10^4. On the other hand, this author recommended that the best that can be achieved in practice is to study the one-dimensional variability, either vertically or horizontally, using a number of observations that ranges from 20 to 100 observations.

The autocorrelation distance (a) is defined as the distance required for the autocorrelation function to decay from 1 to e^{-1} (0.3679) [Diaz Padilla and Vanmarcke (1974)]. On the other hand, the scale of fluctuation is defined as the area under the ACF [Fenton (1999)].

The determination of the autocorrelation distance (a) is done (i) by first fitting the sample ACF to several theoretical ACFs (e.g. the ACFs given in Table 2.1) where $k\Delta h$ is the lag distance and (a) is the autocorrelation distance and (ii) by choosing the theoretical ACF that best fits the data samples. The value of the scale of fluctuation may be determined by using the corresponding formula given in the third column of Table 2.1.

Table 2.1. Theoretical ACF used to determine the autocorrelation distance (*a*) [Vanmarcke (1983)]

Model	Autocorrelation function	Scale of fluctuation (δ)
Single exponential	$\rho_k = \exp\left(\frac{-\lvert k\,\Delta h\rvert}{a}\right)$	$\delta = 2a$
Square exponential	$\rho_k = \exp\left(-\left[\frac{\lvert k\,\Delta h\rvert}{a}\right]^2\right)$	$\delta = \sqrt{\pi}a$
Cosine exponential	$\rho_k = \exp\left(-a\lvert k\,\Delta h\rvert\right)\cos\left(ak\,\Delta h\right)$	$\delta = \frac{1}{a}$
Second-order Markov	$\rho_k = \left(1+a\lvert k\,\Delta h\rvert\right)\exp\left(-a\lvert k\,\Delta h\rvert\right)$	$\delta = \frac{4}{a}$

Finally, it should be mentioned that the modeling of the spatial variability is greatly facilitated by the data being stationary [Uzielli et al. (2005)]. Stationarity is insured if (i) the mean is constant with distance (i.e. no trend exists in the data); (ii) the variance is constant with distance. In random field theory, it is common practice to transform a non-stationary data set to a stationary one by removing a low-order polynomial trend (i.e. a first or a second order polynomial) using the ordinary least square method.

2.3.1.2 Typical values of the soil statistical parameters

Various researchers have worked on the description and quantification of the spatial variability of soil properties [i.e. quantification of the statistical parameters of the soil properties which are the mean, the variance (or standard deviation or coefficient of variation) and the autocorrelation distance (or the scale of fluctuation)].

This section aims at presenting the commonly encountered values of the autocorrelation distances and the coefficients of variation COV of different soil types. It also provides some information regarding the type of the probability density function (PDF) encountered in geotechnical engineering and the coefficient of correlation that may exist between some soil parameters.

2.3.1.2.1 Autocorrelation distances

Several authors have provided horizontal and vertical autocorrelation distances for some soil properties from laboratory and in-situ test measurements. The most commonly used test data are Cone Penetration Test (CPT) measurements [Jaksa et al. (1999); Liu and Chen (2010)]. Other tests were also used such as the Standard Penetration Test (SPT) [Zhang and Chen (2012)], the Vane Shear Test (VST) [Asaoka and Grivas (1982)] and other laboratory tests [Diaz Padilla and Vanmarcke (1974)]. A literature review showing the autocorrelation functions and the corresponding values of the autocorrelation distances of different soil types is presented in Table 2.2. The database contains the type of the test carried out, the considered soil property, the type of soil (sand, clay, silt ...), the type of the autocorrelation function and the corresponding values of the vertical and horizontal autocorrelation distances (or the corresponding values of the scale of fluctuation if these values are given between parentheses). From these tables, one may conclude that the exponential and the square exponential autocorrelation functions are the most suitable forms of the autocorrelation function that fit most of the soil properties. The autocorrelation distances in the vertical and horizontal directions are never the same, but in general, differ by an order of magnitude, with the horizontal scale of fluctuation being higher than that in the vertical direction as was stated by Uzielli et al. (2005).

Table 2.2. Autocorrelation function and the corresponding values of the autocorrelation distance (or scale of fluctuation) of certain soil properties obtained from different tests

Test type	Soil property	Soil type	Autocorrelation function		Autocorrelation distance (or scale of fluctuation when the value is given between parentheses) [m]		Author
			Vertical	Horizontal	Vertical	Horizontal	
-	e	soft organic silty clay	Exponential	-	1.2 - 1.8	-	Diaz Padilla and Vanmarcke (1974)
-	w	soft organic silty clay	Exponential	-	1.2 - 1.8	-	Diaz Padilla and Vanmarcke (1974)
-	$\overline{p}_m$	soft organic silty clay	Exponential	Exponential	0.3 - 0.6	122	Diaz Padilla and Vanmarcke (1974)
-	e	soft silty loam	Exponential sine decaying	Exponential sine decaying	0 – 4.5	16.8 – 22	Alonso and Krizek (1975)
-	γ	soft silty loam	Exponential sine decaying	Exponential sine decaying	0 – 3.0	17 – 22	Alonso and Krizek (1975)
-	n	gravelly sand	Exponential sine decaying	Exponential sine decaying	0 – 2.2	8.5 – 12	Alonso and Krizek (1975)
-	w	soft silty loam	Exponential sine decaying	-	0 – 5	-	Alonso and Krizek (1975)
CPT	F_R	clean sand	Exponential	-	0 - 2	-	Alonso and Krizek (1975)
CPT	q_c	Clay	Exponential sine decaying	-	0 – 0.6	-	Alonso and Krizek (1975)
SPT	N	clean sand	Exponential	-	0 – 4	-	Alonso and Krizek (1975)
UCS	S_u	clay of Mexico	-	Exponential	-	0 – 0.9	Alonso and Krizek (1975)
VST	S_u	soft clay	-	Exponential sine decaying	-	0 - 0.5	Alonso and Krizek (1975)
TUU	S_u	marine clay	Exponential	-	1.5	-	Matsuo and Asaoka (1977)
CPT	q_c	Clay	Exponential	-	(1.3)	-	Vanmarcke (1977)
-	w	Clay	Exponential	-	(2.7)	-	Vanmarcke (1977)
-	e	Clay	Exponential	-	(3.0)	-	Vanmarcke (1977)
CPT	q_c	North sea clay	-	Square exponential	-	30	Tang (1979)
SPT	ln D_r	sand deposit	Exponential	Exponential	1.8	33.5	Fardis and Veneziano (1981)
VST	S_u	organic soft clay	Exponential	-	1.2 – 3.1	-	Asaoka and Grivas (1982)
UU	S_u	offshore soil	Exponential	-	3.6	-	Keaveny et al. (1989)
CPT	q_c	North sea clay	-	Exponential	-	13.9 – 37.5	Keaveny et al. (1989)
UCS	S_u	soft clay	Exponential	Exponential	2	40	Honjo and Kuroda (1991)
VST	S_u	very soft clay	Square exponential	Square exponential	1.1	22.1	Bergado et al. (1994)
CPT	q_c	sand, clay	Exponential	Exponential	(0.1 – 2.2)	(3.0 – 80.0)	Phoon and Kulhawy (1996)
CPT	q_T	clay	Exponential	Exponential	(0.2 – 0.5)	(23.0 – 66.0)	Phoon and Kulhawy (1996)
VST	S_u	Clay	Exponential	Exponential	(2.0 – 6.2)	(46.0 – 60.0)	Phoon and Kulhawy (1996)

Test type	Soil property	Soil type	Autocorrelation function		Autocorrelation distance (or scale of fluctuation when the value is given between parentheses) [m]		Author
			Vertical	Horizontal	Vertical	Horizontal	
CPT	q_c	Clay	Exponential	Exponential	-	9.6 – 37.5	Lacasse and Nadim (1997)
CPT	q_c	Clay	Exponential	-	(0.06 – 0.24)	-	Jaksa et al. (1999)
CPT	S_u	Clay	-	Exponential	-	(1.0 – 2.0)	Jaksa et al. (1999)
CPT	q_c	Clay	Exponential	-	(1.24 – 3.3)	-	Fenton 1999
CPT	q_t	Clay	Exponential	-	(0.2 – 0.5)	-	Phoon and Kulhawy (1999)
VST	S_u	clay	Exponential	-	(2.0 – 6.0)	-	Phoon and Kulhawy (1999)
SPT	N	Sand	Exponential	-	(2.4)	-	Phoon and Kulhawy (1999)
Lab.Test	S_u	clay	Exponential	-	(1.0 – 2.0)	-	Phoon and Kulhawy (1999)
CPT	q_c	Clay	Exponential	-	(0.2 – 0.4)	-	Cafaro and Cherubini (2002)
CPT	q_c	Sand	Two Peak	Two Peak	0.95	12.1	Assimaki et al. (2003)
CPT	q_{net}	clay and silt	Two Peak	Exponential	2.62	16.0	Assimaki et al. (2003)
CPT	F_R	clay, silty clay	Exponential	-	(0.1 – 0.5)	-	Uzielli et al. (2005)
CPT	F_R	clean sand, silty sand	Exponential	-	(0.2 – 0.6)	-	Uzielli et al. (2005)
CPT	q_{c1N}	clay, silty clay	Exponential	-	(0.1 – 0.8)	-	Uzielli et al. (2005)
CPT	q_{c1N}	clean sand, silty sand	Exponential	-	(0.3 – 1)	-	Uzielli et al. (2005)
CPT	f_s	organic soft clay	Exponential	-	(0.127–0.415)	-	Dyminski et al. (2006)
CPT	q_t	organic soft clay	Exponential	-	(0.077– 0.38)	-	Dyminski et al. (2006)
SPT	N	organic soft clay	Exponential	-	(0.8857)	-	Dyminski et al. (2006)
CPT	q_c	granular soil	Exponential	Exponential	(0.1– 2.2)	(3 – 80)	Babu and Dasaka (2008)
-	w	Gravel	Exponential	-	(0.2 – 0.4)	-	Tillmann et al. (2008)
CPT	q_c	Sand	Exponential	Exponential	(1.72 – 2.53)	(286 – 597)	Liu and Chen (2010)
CPT	f_s	Sand	Exponential	Exponential	(0.78 – 0.81)	(287 – 555)	Liu and Chen (2010)
Lab.Test	φ	Sand	Exponential	Exponential	5.1	242	Suchomel and Mašín (2010)
SPT	N	sand	Exponential	-	(1.36 – 3.01)	-	Zhang and Chen (2012)
SPT	N	sand	Square exponential	-	(0.92 – 1.63)	-	Zhang and Chen (2012)
CPT	q_{c1N}	Sand	Exponential	Exponential	(0.45 – 0.5)	(1.2 – 2)	Firouzianbandpey et al. (2014)
CPT	F_R	Sand	Exponential	Exponential	(0.2)	(1.2 – 1.4)	Firouzianbandpey et al. (2014)

1.3.1.2.2. Coefficients of variation COV

The coefficient of variation COV, which is obtained by dividing the standard deviation by the mean value of the random variable, is commonly used in quantifying the geotechnical uncertainty because of its advantage of being dimensionless as well as providing a significant measure of the relative dispersion of data around the mean. Table 2.3 presents the COVs of several soil properties for different soil types.

Studies reported in the literature have shown that each soil property can follow a different probability distribution function PDF and a different coefficient of variation depending on the site (type of test, soil type ...). For example, Phoon and Kulhawy (1999) carried out an extensive investigation to determine the range of variation of the coefficient of variation (COV) for soil parameters from in-situ tests (CPT, SPT, Vane Shear Test ...) and laboratory tests (triaxial tests CD, UU, CU) and found that for clay the COV of s_u varies between 10% and 60%. The COV of the friction angle φ for sand and clay soil was found to be between 5% and 20%. Also, the COV of the cone tip resistance (q_c) of sand varies between 10% and 81% and that of clay between 5% and 40%. Notice however that the corrected cone tip resistance (q_T) of the CPT varies between 5% and 15%.

Regarding the Young modulus, Srivastava and Babu (2009) proposed a COV value of 34% from CPT tests, while Duncan (2000) proposed values that ranges between 15% and 70%. Other authors such as Nour et al. (2002) and Baecher and Christian (2003) proposed values that ranges between 2% and 50%. Finally, notice that Kulhawy et al. (1991) have estimated the COVs of some physical properties of the soil (water content w, liquid limit w_l, plastic limit w_p, density γ, void ratio e) and some mechanical properties of the soil (effective friction angle, shear strength, and compression index). These authors have concluded that the average COVs of the mechanical properties of the soil are larger than those of the physical properties of the soil. Kulhawy et al. (1992), Phoon and Kulhawy (1996) and Dyminski et al. (2006) provided a COV value for *N* (the

number of blows) obtained from the in-situ SPT measurements. The COV values obtained vary between 15% and 150% depending on the studied sites.

Table 2.3. Coefficients of variation of certain soil properties obtained from in-situ and laboratory tests

Test type	Soil property	Soil type	COV (%)	Authors
UC	*c*	-	39 – 49	Fredlund and Dahlman (1972)
UC	*c*	-	30 – 50	Lumb (1972)
Laboratory test	RD	sand	11 – 36	Haldar and Tang (1979)
VST	Su	clay	18 – 30	Asaoka and Grivas (1982)
Laboratory test	Cc	different types of soils	26 – 48	Kulhawy et al. (1991)
Laboratory test	E	different types of soils	13 – 26	Kulhawy et al. (1991)
Laboratory test	γ	different types of soils	2 – 12	Kulhawy et al. (1991)
Laboratory test	Su	different types of soils	16 – 61	Kulhawy et al. (1991)
Laboratory test	W	different types of soils	9 – 32	Kulhawy et al. (1991)
Laboratory test	Wl	different types of soils	3 – 19	Kulhawy et al. (1991)
Laboratory test	Wp	different types of soils	7 – 17	Kulhawy et al. (1991)
Laboratory test	φ’	different types of soils	6 – 21	Kulhawy et al. (1991)
CPT	q_c	-	15 – 37	Kulhawy et al. (1992)
VST	Su	-	10 – 20	Kulhawy et al. (1992)
DMT	q	sand	5 – 15	Kulhawy et al. (1992)
SPT	N	-	15 – 45	Kulhawy et al. (1992)
CPT	q_c	clay	20 – 40	Phoon and Kulhawy (1996)
CPT	q_c	sand	20 – 60	Phoon and Kulhawy (1996)
SPT	*N*	clay and sand	25 – 50	Phoon and Kulhawy (1996)
VST	Su	clay	10 – 40	Phoon and Kulhawy (1996)
DMT	I_D	sand	20 – 60	Phoon and Kulhawy (1996)
DMT	K_D	sand	20 – 60	Phoon and Kulhawy (1996)
DMT	E_D	sand	15 – 65	Phoon and Kulhawy (1996)
PMT	P_L	clay	10 – 35	Phoon and Kulhawy (1996)
PMT	P_L	sand	20 – 50	Phoon and Kulhawy (1996)
CPT	q_c	sand	10 – 81	Phoon and Kulhawy (1999)
CPT	q_c	clay,silt	5 – 40	Phoon and Kulhawy (1999)
CPT	q_T	clay	5 -15	Phoon and Kulhawy (1999)
VST	Su	clay	14 – 47	Phoon and Kulhawy (1999)
CIUC	Su	clay	21 – 43	Phoon and Kulhawy (1999)
UC	Su	fine grains	21 – 57	Phoon and Kulhawy (1999)
UU	Su	clay,silt	11 – 34	Phoon and Kulhawy (1999)
Laboratory test	φ	sand and clay	5 – 20	Phoon and Kulhawy (1999)
-	E	-	15 – 70	Duncan (2000)
-	Su	-	13 – 40	Duncan (2000)
-	C_c	clay	10 – 37	Duncan (2000)
-	C_v	clay	33 – 68	Duncan (2000)
DST	φ	clay	28	El-Ramly et al. (2003)
Piezometer	φ	sand	5.6	El-Ramly et al. (2003)
VST	Su	-	42.9 - 63.4	Dyminski et al. (2006)
UU	Su	-	48.5 - 94.6	Dyminski et al. (2006)
SPT	*N*	-	47 – 150	Dyminski et al. (2006)
CPT	Su	-	39	Srivastava and Babu (2009)
CPT	E	-	34	Srivastava and Babu (2009)
CPT	φ'	-	11	Srivastava and Babu (2009)

2.3.1.3 Types of the probability distribution function PDF

Several authors have investigated the type of probability density function PDF of some geotechnical parameters. One may cite among others El-Ramly et al. (2003) who found that the internal friction angle of the sand obtained from the DST (Dynamic shear test) follows a Lognormal distribution with a COV of 28%. Assimaki et al. (2003) found that q_c and q_{net} of the CPT follows the Beta distribution.

2.3.1.4 Coefficient of correlation (ρ)

The coefficient of correlation between two soil parameters represents the degree of dependence between these parameters. For the shear strength parameters *c* and φ, Lumb (1970) noted that the correlation coefficient $\rho_{c,\varphi}$ ranges from -0.7 to -0.37. Yucemen et al. (1973) proposed values in a range between -0.49 and -0.24, while Wolff (1985) reported that $\rho_{c,\varphi}$=-0.47. Finally, Cherubini (2000) proposed that $\rho_{c,\varphi}$=-0.61.

2.3.2 Methods of discretization of random fields

In order to introduce the soil spatial variability in the analysis of geotechnical structures, the random field *Z* which may be represented by an infinite number of random variables has to be discretized in order to yield a finite number of random variables $\{\chi_j, j = 1,...,s\}$. If the finite element/finite difference method is the method used in the mechanical analysis, it is convenient to evaluate the random field values in the same way as the finite element/finite difference model (i.e. at the nodes of the deterministic mesh or at the elements' mid points of this deterministic mesh). The discretization methods can be divided into three main groups which are [Sudret and Der Kiureghian (2000)]:

1- the point discretization methods

2- the average discretization methods

3- the series expansions methods

Each group involves a number of discretization methods. Among the point discretization methods, one may cite: the Midpoint (MP) method, the Shape Function (SF) method, the Integration Point (IP) method, and the Optimal Linear Estimation (OLE) method. Concerning the average discretization methods, one may cite: the Spatial Average (SA) method and the Local Average Subdivision (LAS) method. In the group of series expansion methods, one may cite the Karhunen-Loeve (KL) expansion method, the Orthogonal Series Expansion (OSE) method and the Expansion Optimal Linear Estimation (EOLE) method. The EOLE method is used in this thesis and it will be presented in more details in the following section. For a detailed description on all the other discretization methods cited above, the reader may refer to Appendix A.

As was stated by Sudret and Der Kiureghian (2000); in the MP, SF, IP, and SA methods, the discretized random field can be expressed as a finite summation as follows:

$$\tilde{Z}(X) = \sum_{j=1}^{N} \chi_j \phi_j(X) \tag{2.2}$$

where N is the number of terms retained in the discretization procedure, $\phi_j(X)$ are deterministic functions and χ_j are random variables obtained from the discretization procedure. They can be expressed as weighed integrals of the real random field Z over the volume Ω of the system as follows:

$$\chi_j = \int_{\Omega} Z(X)\omega(X)d\Omega \tag{2.3}$$

where $\omega(X)$ is the weight function. The values of the weight functions and the deterministic functions for all the above mentioned methods are given in Sudret and Der Kiureghian (2000) and they are reported in Appendix A of this thesis.

Sudret and Der Kiureghian (2000) have stated that the deterministic functions ϕ_j given in Equation (2.2) are not optimal in the case of midpoint (MP), Spatial Average (SA), Shape Function (SF) and Integration Point (IP) methods. This is also true for the Local Average Subdivision LAS method by Fenton and Vanmarcke (1990). This means that the number of random variables involved in the discretization scheme is not minimal. Thus, of particular interest are the series expansion methods. In all these methods, the number of the deterministic functions ϕ_j is optimal and thus, the number of random variables involved is minimal.

As a conclusion, all the discretization methods presented in the first two groups provide non optimal solution which makes them unattractive tools for random field discretization. This is because the number of random variables needed to discretize the random fields using these methods is mesh depending. Thus, one obtains a large number of random variables for large finite element/finite difference models. The series expansion methods solve this problem. They provide the optimal number of random variables needed to accurately discretize the random field which makes them powerful tools for random field discretization. From this group, the eigenvalue problem of the KL method given in Equation (A. 2) can be solved analytically only for few types of autocorrelation functions and geometries. As for the OSE method, it avoids solving the eigenvalue problem of the KL method given in Equation (A. 2). On the other hand, this method is less attractive in terms of accuracy when compared to the KL and the EOLE method [cf. Sudret and Der Kiureghian (2000)]. For this reason, the EOLE method which uses the concept of OLE method is selected herein to perform the random field discretization. This method is described in some details in the following section.

2.3.2.1 Expansion Optimal Linear Estimation (EOLE) method

This method was originally proposed by Li and Der Kiureghian (1993) for the case of uncorrelated Gaussian random fields. It was then extended by Vořechovský (2008) to cover the case of cross-correlated non-Gaussian random fields. This method is presented herein in the case of non-Gaussian and uncorrelated random fields.

A non-Gaussian random field $Z^{NG}(X)$ is described by: (i) constant mean and standard deviation (μ_Z, σ_Z), (ii) non-Gaussian marginal cumulative density function G, and (iii) an autocorrelation function $\rho_Z^{NG}[(X), (X')]$ which gives the values of the correlation between two arbitrary points (x, y) and (x', y'). It should be mentioned here that the most suitable form of autocorrelation function that fits most of the soil properties is the square exponential autocorrelation function (see Table 2.2). This function is given as follows:

$$\rho_z^{NG}[(x, y), (x', y')] = \exp\left(-\left(\frac{|x - x'|}{a_x} \right)^2 - \left(\frac{|y - y'|}{a_y} \right)^2 \right) \tag{2.4}$$

where a_x and a_y are respectively the horizontal and vertical autocorrelation distances. In order to discretize the random field using the EOLE method, one should first define a stochastic grid composed of s grid points (or nodes) and determine the non-Gaussian autocorrelation matrix Σ^{NG} which gives the correlation between each grid point of the stochastic mesh and the other grid points of this mesh using Equation (2.4). The non-Gaussian autocorrelation matrix Σ^{NG} should then be transformed into the Gaussian space using the Nataf transformation (Nataf 1962).

As a result, one obtains a Gaussian autocorrelation matrix Σ^{G} that can be used to discretize the Gaussian random field Z as follows:

$$\tilde{Z}^{G}(X) = \mu_Z + \sigma_Z \sum_{j=1}^{s} \frac{\xi_j}{\sqrt{\lambda_j}} \cdot \left(\phi_j \right)^T \cdot \Sigma_{Z(X);\chi} \tag{2.5}$$

where μ_Z and σ_Z are the mean and standard deviation values of the random field *Z*; (λ_j, ϕ_j) are the eigenvalues and eigenvectors of the Gaussian autocorrelation matrix Σ^G; $\Sigma_{Z(X);\chi}$ is the correlation vector between the value of the field at an arbitrary point (*x*, *y*) and its values at the different grid points and ξ_j is a standard normal random variable, and *s* is the total number of grid points.

Once the Gaussian random field is obtained, it should be transformed to the non-Gaussian space by applying the following formula:

$$\tilde{Z}^{NG}(X) = G^{-1}\left\{\Phi\left[\tilde{Z}^{G}(X)\right]\right\} \tag{2.6}$$

where $\Phi(.)$ is the standard normal cumulative density function. It should be mentioned here that the series given by Equation (2.5) is truncated for a number of terms *M* (expansion order) smaller than the number of grid points *s*, after sorting the eigenvalues (λ_j ; *j*=1, …, *s*) in a descending order. This number should assure that the variance of the error is smaller than a prescribed tolerance. Notice that the variance of the error for EOLE is given by Sudret and Der Kiureghian (2000) as follows:

$$Var\left[Z(X) - \tilde{Z}(X)\right] = \sigma_Z^2\left\{1 - \sum_{j=1}^{M}\frac{1}{\lambda_j}\left(\left(\phi_j\right)^T \Sigma_{Z(X);\chi}\right)^2\right\} \tag{2.7}$$

where $Z(X)$ and $\tilde{Z}(X)$ are respectively the exact and the approximate values of the random field at a given point *X* and $\left(\phi_j\right)^T$ is the transpose of the eigenvector ϕ_j.

2.4 PROBABILISTIC METHODS FOR UNCERTAINTY PROPAGATION

Development of efficient methods for uncertainty propagation in order to perform probabilistic analyses has gained much attention in recent years due to the importance of introducing uncertainties in the model parameters. The uncertainty propagation aims to study the impact of

input uncertainty on the probabilistic outputs. Three steps are involved in the probabilistic analysis. The first step consists in identifying the uncertain input parameters and modeling them by random variables or random fields (Step A in Figure 2.3). The second step (Step B in Figure 2.3) consists in the computation of the system response for different realizations using the mechanical model. The final step consists in the post-treatment of the obtained results to provide the probabilistic outputs (Step C in Figure 2.3).

In the probabilistic framework, all of the relevant information regarding the model output is contained in its PDF. Thus, determining the PDF of the system response is the main goal in all uncertainty propagation methods. However, the fact that we are considering numerical models implies that the relation between the model uncertain inputs and the system response cannot be represented by an analytical expression. Consequently, it is impossible to obtain a simple analytical expression of the PDF of the system response. However, for practical purposes, not all the information contained in the PDF is necessary. Thus, depending on the type of study that is carried out, only a set of probabilistic outputs can be used. These probabilistic outputs may be the statistical moments (mean and standard deviation) of the system response or the probability of failure (i.e. the probability of exceeding a given threshold value of this system response). In this thesis, the probabilistic output of interest is the probability of failure P_f. Several methods exist for its computation. These methods can be divided into three main categories which are the approximate methods, the simulation methods and the metamodeling techniques.

In the following, one will first present some basic reliability concepts. This is followed by a presentation of the probabilistic methods used for the estimation of P_f (i.e. the approximate methods, the simulations methods and the metamodeling techniques).

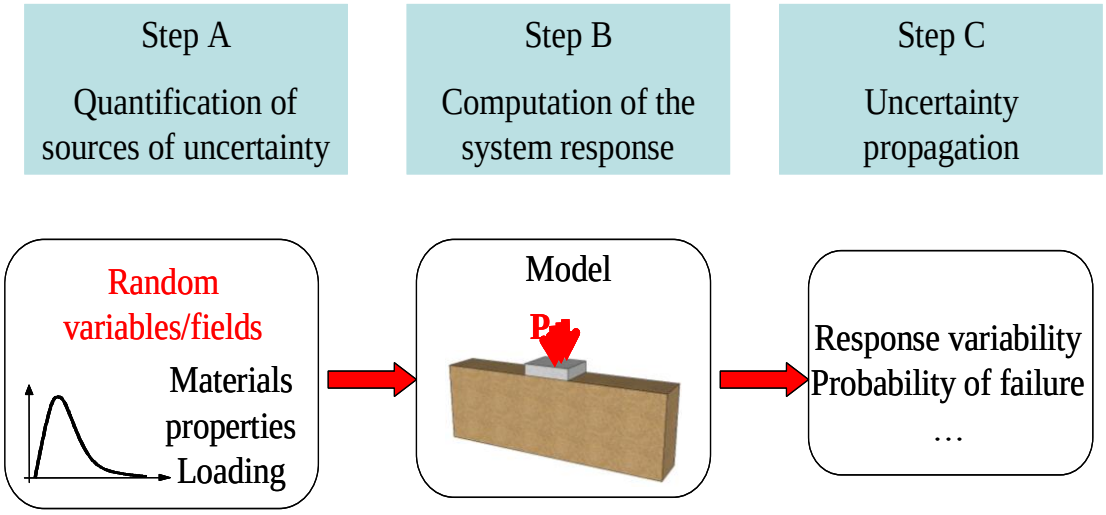


Figure 2.3. General sketch for the probabilistic analyses [Al-Bittar (2012)]

2.4.1 Some basic reliability concepts

2.4.1.1 Performance function and limit state surface

The performance function G is a function by which one can distinguish if a given set of values of the random variables leads to system failure or to system safety. The performance function can be expressed, for a given problem, in different ways. For example, for vertically loaded footings, there are two different forms of the performance function with respect to soil punching: (i) $G=(P_u/P_s)-1$ or (ii) $G=P_u-P_s$ where P_u and P_s are respectively the ultimate vertical load and the footing applied vertical load. In this example, if the random variables are the soil shear strength parameters (c and φ) and the applied load P_s is assumed to be deterministic; then, all pairs (c, φ) that make $G<0$ (i.e. $P_u<P_s$) lead to failure. However, all pairs (c,φ) that make $G>0$ (i.e. $P_u>P_s$) lead to system safety (cf. Figure 2.4a).

The limit state surface of a given mechanical system is defined as the surface that joins the set of values of the random variables (c and φ in the present case) for which failure just occurs (i.e. for which $G=0$). As shown in Figure (2.4a), the limit state surface divides the space of random variables into two zones: (i) a safe zone (characterized by $G>0$) for which all combinations of

random variables (c, φ) lead to safety and (ii) failure zone (characterized by $G<0$) for which all combinations of random variables (c, φ) lead to failure.

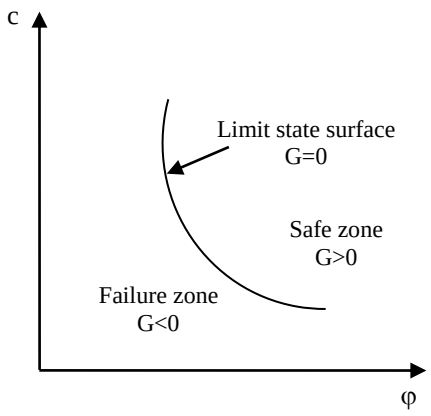


Figure 2.4a. Limit state surface in the space of random variables

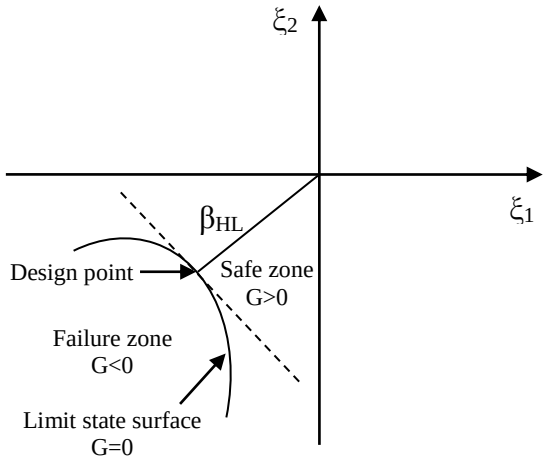


Figure 2.4b. Hasofer-Lind reliability index in the standard space of random variables

2.4.1.2 The reliability index β_{HL}

The reliability index of a given structure is a measure of the safety that takes into account the inherent uncertainties of the input variables. A widely used reliability index is the Hasofer-Lind reliability index β_{HL} [Hasofer and Lind (1974)]. This index is defined as the shortest distance that separates the limit state surface expressed in the space of standard normal uncorrelated random variables and the origin of this space (Figure 2.4b).

In the case where the limit state surface is known analytically, β_{HL} can be easily calculated by minimization of the following formula:

$$\beta_{HL} = \min_{G=0} \left(\sqrt{\sum_{i=1}^{M} \xi_i^2} \right) \tag{2.8}$$

where ξ_i (i=1, 2, …, M) are M standard normal uncorrelated random variables corresponding to the M physical uncertain parameters. The computation of β_{HL} may be described by the two following steps:

- In the first step, the physical (original) random variables should be transformed to the standard normal random variables. In this step, isoprobabilistic transformation is used to transform the physical random variables to standard normal random variables as follows:

$$\xi_i = \Phi^{-1}\left[F_{X_i}(X_i) \right] \tag{2.9}$$

in which, X_i is a physical random variable, F_{X_i} is the cumulative density function (CDF) of the physical random variable X_i and $\Phi^{-1}(\cdot)$ is the inverse of the CDF of the standard normal random variable. Notice that, if the original random variables are correlated, they should be transformed into uncorrelated random variables.

- In the second step, it is required to search for the minimal distance between the limit state surface and the origin in the standard space of uncorrelated random variables. The point corresponding to the minimal value of the reliability index is commonly referred to as the design point. This point corresponds to the most probable point leading to failure. In our case of a vertically loaded footing, this point corresponds to the pair of values of c and φ that is the most critical in a probabilistic framework (i.e. the one corresponding to the maximal value of the failure probability).

2.4.2 Probabilistic methods for the estimation of the failure probability P_f

If the random vector X that represents the uncertain parameters of a given mechanical system is described by a joint probability density function $f_X(X)$, then one can define the failure probability P_f of this mechanical system as:

$$P_f = P(G(X) \leq 0) \tag{2.10}$$

where G(X) is the performance function. The failure probability P_f given by Equation (2.10) may be calculated as the integral of the joint probability density function in the failure domain as follows:

$$P_f = \int_{G(X)\leq 0} f_X(X)dX \tag{2.11}$$

The integral in Equation (2.11) represents the volume between the joint probability density function $f_X(X)$ and the limit state surface G(X) = 0 (see Figure 2.5 in the case of two random variables). Note that the calculation of the integration in Equation (2.11) is practically impossible in the general case. For this reason, several numerical methods were proposed in literature to calculate the failure probability P_f. The following subsections are devoted for the presentation of the three main categories of methods for the computation of P_f which are the approximate methods, the simulation methods and the metamodeling techniques.

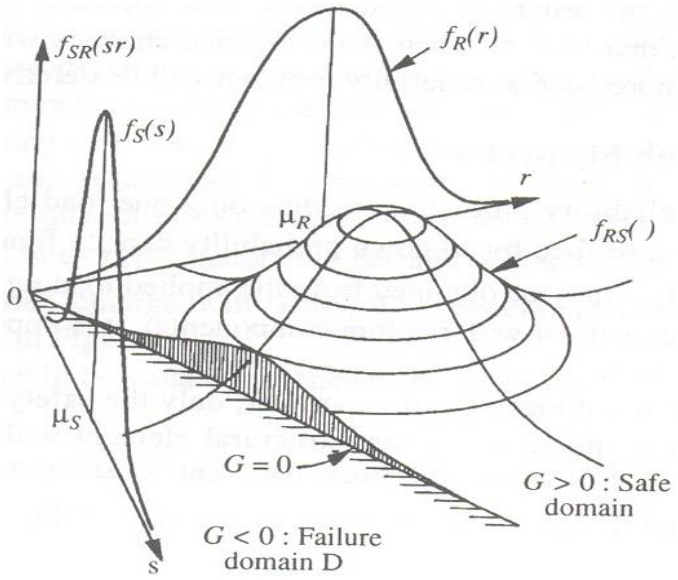


Figure 2.5. Joint probability density function and limit state surface in case of two random variables R and S [after Melchers (1999)]

2.4.2.1 The approximate methods

These methods are based on approximating the limit state surface locally at a reference point. This class of methods can be very efficient (in that only a relatively small number of model evaluations is needed to calculate P_f), but it tends to become unreliable in the presence of complex, non-linear limit state surfaces. These methods can be broadly grouped into First-Order Reliability Method (FORM) and Second-Order Reliability Method (SORM). A detailed discussion of these methods for application to geotechnical problems is presented by many authors [Ang and Tang (1975); Harr (1987); Haldar and Mahadevan (2000); Baecher and Christian (2003)].

2.4.2.1.1 First Order Reliability Method (FORM)

The Hasofer-Lind reliability index presented previously in this chapter can be used to calculate an approximate value of the failure probability using the First Order Reliability Method (FORM) as follows:

$$P_f \approx \Phi(-\beta_{HL}) \tag{2.12}$$

where $\Phi(\cdot)$ is the standard normal cumulative density function (CDF) and β_{HL} is the Hasofer-Lind reliability index. Within this method, the estimation of P_f is based on the approximation of the limit state surface by an hyperplane tangent to this limit state surface at the most probable failure point called "design point" (see Figure 2.4b).

2.4.2.1.2 Second Order Reliability Method (SORM)

The second-order reliability method (SORM) is a second-order refinement of the FORM P_f estimate. The computational costs associated to this refinement increase rapidly with the number of input random variables. After the design point is identified by FORM, the failure probability is

approximated by a tangent hyper parabolic defined by the second order Taylor expansion (see Figure 2.6).

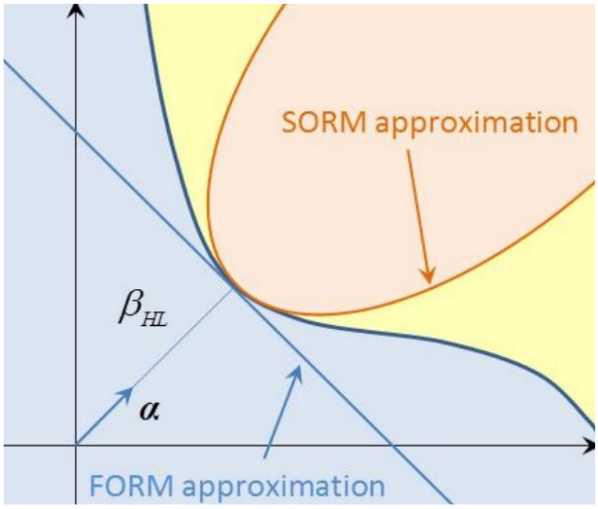


Figure 2.6. Comparison between FORM and SORM approximations of the failure domain for a simple 2-dimensional case [after Marelli et al. (2016)]

2.4.2.2 The simulation methods

This section is devoted to the presentation of the simulation methods used for the computation of the failure probability. This category regroups the universal Monte Carlo simulation (MCS) methodology and other more advanced simulation techniques (i.e. the Importance sampling (IS) and the Subset simulation (SS)). In spite of being rigorous and robust, the crude MCS is well-known to be very time-expensive especially when dealing with finite element or finite difference models which do not offer an analytical solution of the involved problem. Also, this method shows a low efficiency when considering the small failure probabilities encountered in practice. Indeed, MCS is very time-expensive due to (i) the long time required per simulation to calculate the system response when finite element or finite difference codes are involved and (ii) because one needs a large number of simulations to calculate a small failure probability with a small value of its coefficient of variation. The advanced simulation techniques (i.e. the IS and the SS) are all based on a modification of the MCS method in order to simulate more points in a particular zone

of interest and thus they are very attractive for the computation of the failure probability. Consequently, the MCS methodology remains the origin of all the advanced simulation techniques and deserves to be firstly presented. This is followed by a brief presentation of the more advanced simulation techniques (i.e. IS and SS) which are extensively employed in this thesis in combination with other probabilistic techniques (kriging meta-modeling) for the computation of the probability of failure.

2.4.2.2.1 Monte Carlo Simulation (MCS) method

The Monte Carlo simulation is a universal method to evaluate complex integrals. It consists of generating K samples which respect the joint probability density function $f_X(X)$ of the M random variables ($X_1, \ldots, X_M$) gathered in a vector X. For each sample, the performance function is calculated. Thus; for the K samples, one obtains K values of the performance function gathered in a vector $G = \left\{G\left(X^{(1)}\right),...,G\left(X^{(K)}\right)\right\}$ which may be used to determine the estimator of the probability of failure as follows:

$$\tilde{P}_f = \frac{1}{K}\sum_{j=1}^{K} I_j(X) \tag{2.13}$$

where I(X) is an indicator function of the failure domain given by:

$$I(X) = \begin{cases} 1 \ if \quad G(X) \le 0 \\ 0 \ if \quad G(X) > 0 \end{cases} \tag{2.14}$$

In this equation, I(X) = 1 when the random input vector X causes the system to fail and I (X) = 0 otherwise. As may be seen from Equation (2.13), the estimation of the failure probability by MCS methodology consists of computing the ratio between the number of samples that belong to the failure domain and the total number of samples. An important measure for assessing the accuracy of a MCS estimator is given by the coefficient of variation COV defined as:

$$COV\,(P_f) = \sqrt{\frac{1-P_f}{K \cdot P_f}} \tag{2.15}$$

As may be seen from Equation (2.15), the coefficient of variation of the failure probability decreases with the number of samples and it increases with decreasing P_f . To give an example, to estimate a P_f = 10^{-3} with 10% accuracy, 10^5 samples are needed. The COV is often used as a convergence criterion to adaptively increase the MC sample size until some desired COV is reached.

2.4.2.2.2 Importance sampling (IS)

In MCS methodology, a very large number of samples is required to reach the failure region and hence to accurately calculate the failure probability especially when computing small values of this failure probability. To overcome this shortcoming, the importance sampling (IS) technique was proposed in literature [Schuëller and Stix (1987); Ibrahim (1991)]. In the IS technique, the sampling is shifted towards the region of interest (failure region) and the random variables should be sampled from the 'optimal' probability density function $h_x(X)$. Notice that the optimal importance sampling density (ISD) is unknown since it requires knowledge of the failure probability as a priori. One popular choice for ISD is to move the sampling center to the design point (which should thus be determined beforehand) and to use a standard Gaussian PDF $h_x(X)$ centered at this design point. By making use of this ISD, the computation of the failure probability is given by:

$$P_f = \int_{\Omega} I_F(X)\,\frac{f_x(X)}{h_x(X)}\,h_X(X)\,dX = \frac{1}{K}\sum_{i=1}^{K} I_F(X_i)\,\frac{f_X(X_i)}{h_X(X_i)} \tag{2.16}$$

in which, P_f is the failure probability estimate and K is the number of samples. It is to be emphasized here that a much smaller number of samples (as compared to MCS methodology) is required by the IS technique to reach the failure region and thus to accurately calculate the failure

probability estimate P_f . The accuracy of P_f is measured by its coefficient of variation $COV\left(P_f\right)$. This coefficient of variation is calculated as follows:

$$COV\left(P_f\right)=\frac{\sqrt{Var(P_f)}}{P_f} \tag{2.17}$$

where P_f is the value of the failure probability estimate and $Var(P_f)$ is the variance of this estimate. The variance $Var(P_f)$ is calculated by the following equation:

$$Var(P_f)=\frac{1}{K-1}\left[\frac{1}{K}\sum_{i=1}^{K}\left(I_F(X_i)\left(\frac{f(X_i)}{h(X_i)}\right)^2\right)-P_f^2\right] \tag{2.18}$$

2.4.2.2.3 Subset Simulation (SS) approach

The subset simulation (SS) approach was proposed by Au and Beck (2001) as an alternative to MCS methodology to compute the small failure probabilities. Its aim is to reduce the number of calls of the mechanical model as compared to MCS methodology. This method was used by several authors [Au and Beck (2003); Schuëller et al. (2004); Au et al. (2007); Au et al. (2010); Ahmed and Soubra (2011, 2012a, 2012b); Thajeel et al. (2015) among others]. The basic idea of subset simulation is that the small failure probability in the original probability space can be expressed as a product of larger conditional failure probabilities.

Consider a failure region F defined by the condition $G<0$ where G is the performance function and let $\left(X^{(1)},...,X^{(K)}\right)$ be a sample of K realizations of the vector X composed of M random variables ($X_1, \ldots, X_M$). It is possible to define a sequence of nested failure regions $F_1, \ldots, F_j, \ldots, F_m$ of decreasing size where $F_1 \supset \ldots \supset F_j \supset \ldots \supset F_m = F$ (Figure 1.7). An intermediate failure region F_j can be defined by $G<C_j$ where C_j is an intermediate failure threshold whose value is larger than zero. Thus, there is a decreasing sequence of positive failure thresholds $C_1, \ldots, C_j, \ldots, C_m$ corresponding respectively to $F_1, \ldots, F_j, \ldots, F_m$ where $C_1>\ldots>C_j>\ldots> C_m=0$.

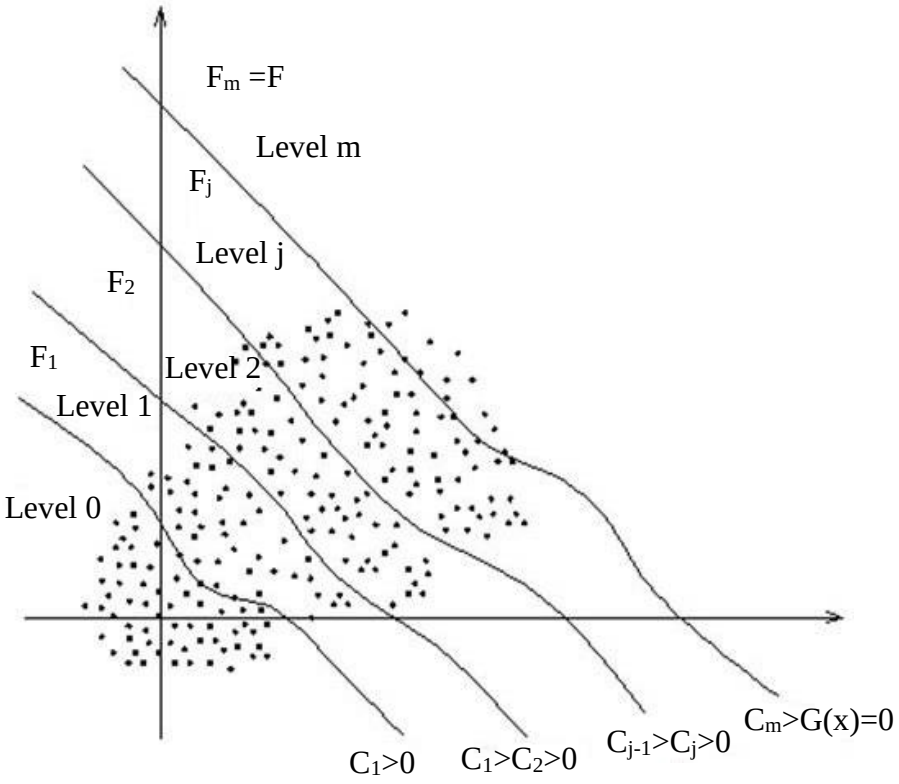


Figure 2.7. Nested Failure domain

In the SS approach, the space of uncertain parameters is divided into a number *m* of levels with equal number *K* of realizations $\left(X^{(1)},...,X^{(K)}\right)$. An intermediate level *j* contains a safe region and a failure region defined with respect to a given failure threshold C_j. The failure probability corresponding to this intermediate level *j* is calculated as follows:

$$P(F_j \left| F_{j-1}\right.) = \frac{1}{K}\sum_{k=1}^{K} I_{F_j}\left(X^{(k)}\right) \tag{2.19}$$

where $I_{F_j}\left(X^{(k)}\right)=1$ if $X^{(k)} \in F_j$ and $I_{F_j}\left(X^{(k)}\right)=0$ otherwise. Notice that in the SS approach, the first *K* realizations are generated using MCS methodology according to a target joint probability density function PDF (Pt). The next realizations of each subsequent level are obtained using a Markov chain method based on the Metropolis-Hastings (M-H) algorithm according to a proposal PDF (Pp).

The failure probability $P(F)=P(F_m)$ of the failure region F can be calculated from the sequence of conditional failure probabilities as follows [Au and Beck (2001)]:

$$P(F)=P(F_m)=P(F_1)\cdot\prod_{j=2}^{m}P(F_j\mid F_{j-1}) \tag{2.20}$$

It should be mentioned that the intermediate failure probability $P(F_j\mid F_{j-1})$ of a given level j (j=1, 2,. . ., m) was chosen equal to 0.1 [Au and Beck (2001)].

The coefficient of variation $COV\ (P_f)$ of the failure probability estimator is given by:

$$COV\ (P_f)=\sqrt{\sum_{j=1}^{m}COV_j^{\,2}} \tag{2.21}$$

where COV_j is the coefficient of variation for a given level j (j=1, 2,. . ., m). It should be mentioned that the coefficient of variation COV_1 (the first level using MCS methodology) is given according to Equation (2.15) as follows:

$$COV_1=\sqrt{\frac{1-P(F_1)}{K\cdot P(F_1)}} \tag{2.22}$$

As for the coefficient of variation, for next level j (j=2, 3,. . ., m), it is given by Au and Beck (2001) as follows:

$$COV_j=\sqrt{\frac{1-P(F_j\mid F_{j-1})}{K\cdot P(F_j\mid F_{j-1})}\cdot(1+\gamma_j)}\qquad j=2,3,...,m \tag{2.23}$$

where γ_j is a factor accounting for the correlation between the K realizations. Its expression is given by Au and Beck (2001).

2.4.2.3 The metamodeling techniques

Although IS and SS approaches allow one to significantly reduce the number of calls to the mechanical model with respect to MCS methodology, these techniques remain very time-consuming. In order to overcome the inconvenience of the simulation methods related to the computation time, the metamodeling techniques were proposed in this regard. The aim of these techniques is to replace the original computationally-expensive mechanical model by a meta-model (i.e. an analytical equation) using a smaller number of calls of the mechanical model (as compared to IS or SS techniques).

Various types of metamodeling techniques can be found in literature such as Response Surface Methodology (RSM) [Box et al. (1978); Bucher and Bourgund (1990); Myers and Montgomery (1995)], Polynomial Chaos Expansion (PCE) and its extension the Sparse Polynomial Chaos Expansion (SPCE) [Spanos and Ghanem (1989); Isukapalli et al. (1998); Xiu and Karniadakis (2002); Berveiller et al. (2006); Sudret et al. (2006); Sudret and Berveiller (2008); Huang et al. (2009); Blatman and Sudret (2010); Mao et al. (2012); Al-Bittar and Soubra (2013, 2014a, 2014b, 2016)], Artificial Neural Networks (ANN) [Papadrakakis and Lagaros (2002)], the Support Vector Machine (SVM) [Hurtado (2004) ; Zhou et al. (2013)] and the Kriging method [Sacks et al. (1989); Booker et al. (1999); Santner et al. (2003); Kaymaz (2005); Bichon et al. (2008)].

The basic common idea of the metamodeling techniques is to use a certain number of evaluations of the performance function to build a meta-model. Once this meta-model is built, it is used to evaluate the performance function instead of using the computationally-expensive mechanical model.

In this thesis, the kriging metamodeling technique is used in combination with a simulation method (e.g. MCS, IS or SS) to perform the probabilistic analysis. The combination between a metamodeling technique and a simulation approach takes the advantages of each technique used

individually and has been employed by several authors. Bourinet et al. (2011) proposed a method combining Subset Simulation and Support Vector Machines. Papadopoulos et al. (2012) proposed a methodology incorporating Neural Network and Subset Simulation. Finally, Balesdent et al. (2013) developed a kriging adaptive importance sampling approach.

The next subsections aim to briefly present commonly used metamodeling techniques (e.g. RSM and PCE/SPCE). This is followed by a more detailed presentation of the kriging methodology which is the metamodeling technique employed in this thesis. As will be shown in the following sections, kriging metamodeling provides not only the predicted value at an unknown sample but also an uncertainty indication of the meta-model at this sample. This advantage becomes visible in the upcoming chapters.

2.4.2.3.1 The Response Surface Methodology (RSM)

The Response Surface Methodology (RSM) aims at approximating the performance function $G(X)$ by an explicit function of the random variables. The most popular form of this function is a second order polynomial model, which can be expressed as:

$$G_{RSM}\left(X\right) = a_0 + \sum_{i=1}^{M} a_i X_i + \sum_{i=1}^{M} b_i X_i^2 \tag{2.24}$$

where X_i ($i=1, \ldots, M$) are the random variables, M is the number of random variables and $\left(a_i, b_i\right)$ are coefficients obtained by the least squares method, which minimizes the sum of the squares of the differences between the predicted values and the model values. It should be emphasized here that the second order polynomial used in the RSM method has limited capability to accurately model highly nonlinear response surfaces. Higher-order polynomial models can be used to model highly nonlinear response surfaces; however, instabilities may arise [cf. Barton (1992)]. Furthermore, this requires a large number of sample points. This enormously increases the computation time and make the RSM solution inadequate in this case.

2.4.2.3.2 The Polynomial Chaos Expansion (PCE) and its extension the (SPCE)

The polynomial chaos expansion (PCE) aims at replacing a complex mechanical model (i.e. finite element/finite difference numerical model) by a meta-model. This allows one to calculate the system response (when performing MCS) using an approximate simple analytical equation.

Consider a mechanical model with M input uncertain parameters gathered in a vector $X = \{X_1, \ ..., \ X_M\}$. The system response can be expanded onto an orthogonal polynomial basis as follows:

$$G_{PCE}(\xi) = \sum_{\beta=0}^{\infty} a_\beta \Psi_\beta(\xi) \cong \sum_{\beta=0}^{P-1} a_\beta \Psi_\beta(\xi) \tag{2.25}$$

where ξ is the vector resulting from the transformation of the random vector X into an independent standard normal space, P is the number of terms retained in the truncation scheme, a_β are the unknown PCE coefficients to be computed and Ψ_β are multivariate (or multidimensional) Hermite polynomials which are orthogonal with respect to the joint probability density function of the standard normal random vector ξ.

In practice, one should truncate the PCE representation by retaining only the multivariate polynomials of degree less than or equal to the PCE order p. Using this method of truncation, the number P of unknown PCE coefficients is given by:

$$P = \frac{(M + p)!}{M\,!\,p\,!} \tag{2.26}$$

As may be seen from Equation (2.26), the number P of the PCE coefficients which is the number of terms retained in Equation (2.25) dramatically increases with the number M of random variables and the order p of the PCE. This number becomes very high in the case of random fields where the number of random variables is significant.

Note that unknown coefficients a_β are computed using the regression approach. In this approach, one needs to generate a DoE using the Monte Carlo simulation methodology or any other sampling scheme (e.g. Latin Hypercube (LH)). Note that the size K of the DoE must insure the stability of the regression problem. This implies that the size of DoE should be larger than the number of unknown coefficients.

Once the coefficients a_β of the PCE given by Equation (2.25) have been computed, the statistical moments (mean, standard deviation, skewness, and kurtosis) can be calculated with no additional cost. This can be done by performing Monte Carlo simulations on the meta-model and not on the original computationally-expensive finite element/finite difference numerical model.

The sparse polynomial chaos expansion (SPCE) which is an extension of the PCE methodology was proposed by Blatman and Sudret (2009, 2010) to deal with high dimensional stochastic problems (i.e. when a large number of random variables is involved). The idea behind the SPCE came from the fact that the number of significant terms in a PCE is relatively small [see Blatman (2009)] since the multidimensional polynomials Ψ_β corresponding to high-order interaction are associated with very small values of coefficients a_β. Based on these observations, a new truncation strategy was proposed by Blatman and Sudret (2009, 2010) in which the multidimensional polynomials Ψ_β corresponding to high-order interaction were penalized. The proposed SPCE methodology leads to a sparse polynomial chaos expansion that contains a small number of unknown coefficients. These coefficients can be calculated from a reduced number of calls of the mechanical model with respect to the classical PCE methodology.

It should be mentioned here that using the PCE/SPCE will lead to an accurate estimation of the first two statistical moments of the system response. However, the probability of failure which is the output of interest in this thesis is generally not well estimated with this approach because the generated samples are usually located around the mean values of the different random variables.

2.4.2.3.3 Kriging metamodeling technique

The kriging metamodeling technique is employed in this thesis to predict the response of a given model at any sample, provided that the responses at a few other samples (called Design of Experiments DoE) are known. With this kind of metamodeling, one can approximate arbitrary functions with high accuracy. The kriging meta-model does not assume an underlying global functional form. However, it is based on the assumption that there is a spatial correlation between the values of the function to be approximated.

In this thesis, the aim is to determine a meta-model for the performance function based on some performance function values $G(X)$ obtained from the FLAC^{3D} software. Within the kriging theory, the value of the performance function $G(X)$ at an unknown sample X is considered as a realization $\hat{G}(X)$ of a random function, which includes a regression part and a centered stochastic process as follows [cf. Sacks et al. (1989)]:

$$\hat{G}(X) = F(X, \beta) + Z(X) \tag{2.27}$$

where $F(X, \beta)$ is a deterministic part defined by a regression model that gives an approximation to $G(X)$ in the mean and $Z(X)$ represents the fluctuation around the mean value. It is given by a stationary Gaussian random process with zero mean and covariance *COVAR* that interpolates the errors between the regression model prediction and the true performance function values at the different N samples of the DoE. This covariance is given by:

$$COVAR\left[Z(X_i), Z(X_j)\right] = \sigma_Z^2 \; R(X_i, X_j) \quad (i, j = 1, 2, 3, ..., N) \tag{2.28}$$

where σ_z^2 is the random field variance; X_i, X_j are two arbitrary samples from the whole design of experiments and $R(X_i, X_j)$ is the spatial correlation function between these two samples with a correlation parameter vector θ. Notice that in this thesis, ordinary kriging is used which means

that $F(X,\beta)$ is a scalar [i.e. $F(X,\beta)=\beta$] to be determined. So, the estimated performance function $\widehat{G}(X)$ can be simplified as:

$$\widehat{G}(X)=F(X,\beta)+Z(X)=\beta+Z(X) \tag{2.29}$$

Notice also that the most widely used correlation function for reliability analysis is the anisotropic square exponential function. This function is employed in this thesis. It is given by:

$$R(X_i,X_j)=\prod_{i=1}^{M}e^{\left(-\theta_i\left(X_i-X_j\right)^2\right)} \tag{2.30}$$

where M is the number of random variables and θ_i is a scalar which is equal to the inverse of the correlation length in the i^{th} direction.

As may be seen from Equation (2.27) or Equation (2.29), the kriging meta-model consists of two parts. The first part is a regression model which approximates the performance function over the whole design space. The second part is a stochastic process which creates localized deviations from the regression model. This meta-model is completely defined by the scalar β, the correlation parameter vector θ and the process variance σ_z^2. These parameters may be estimated by fitting the kriging meta-model to the design of experiments and the corresponding performance function values. Notice that the correlation parameter vector θ can be obtained through the maximum likelihood estimation using DACE toolbox in Matlab [Lophaven et al. (2002)]. Once θ is determined, β and σ_Z^2 can also be estimated using the DACE toolbox. Their expressions are given by:

$$\widehat{\beta}=\left(F^TR^{-1}F\right)^{-1}F^TR^{-1}G \tag{2.31}$$

$$\widehat{\sigma}_Z^2=\frac{1}{N}\left(G-\beta F\right)^T R^{-1}\left(G-\beta F\right) \tag{2.32}$$

where R is a square matrix of dimension $N \times N$, i.e. $R = \left[R\left(X_i, X_j \right) \right]_{NxN}$ and F is a unit vector of dimension *N*. At this stage, all the three parameters of Equation (2.29) are completely determined.

The Best Linear Unbiased Predictor (BLUP) of the performance function $\hat{G}(X)$ at an unknown sample X is shown to be a Gaussian random variate $\hat{G}(X) \sim N\left(\mu_{\hat{G}(X)}, \sigma_{\hat{G}(X)} \right)$ [see Santner et al. (2003)] where:

$$\mu_{\hat{G}(X)} = \beta + r(X) R^{-1} (G - \beta F) \tag{2.33}$$

$$\sigma^2_{\hat{G}(X)} = \sigma_Z^2 (X) \left(1 + u^T (X) \left(F^T R^{-1} F \right)^{-1} u^T (X) - r^T (X) R^{-1} r (X) \right) \tag{2.34}$$

in which $r(X) = \left[R(X, X_1), R(X, X_2), ..., R(X, X_N) \right]$ is the correlation vector between the arbitrary sample *X* and all the samples X_i $(i = 1, 2, 3, ..., N)$ in the DoE, and $u(X)$ is given by:

$$u(X) = F^T R^{-1} r(X) - 1 \tag{2.35}$$

The computation of the mean prediction $\mu_{\hat{G}(X)}$ and the prediction variance $\sigma^2_{\hat{G}(X)}$ as given by Equations (2.33) and (2.34) can be obtained by the DACE toolbox in Matlab making use of the already obtained values of $\hat{\beta}$, $\hat{\sigma}_Z^2$ and θ.

It should be emphasized here that contrary to other types of meta-models, the kriging meta-model provides not only a predicted value at an unknown sample but also an estimate of the prediction variance which gives an uncertainty indication on the kriging meta-model at this sample.

Notice finally that the prediction variance $\sigma^2_{\hat{G}(X)}$ is defined as the minimum of the mean squared error between $\hat{G}(X)$ and $G(X)$. The variances of samples in the initial DoE are zero, but the variances of the other samples are always different from zero. A large value of $\sigma^2_{\hat{G}(X)}$ means that

the prediction is not exact. Therefore, the prediction variance $\sigma^2_{\tilde{G}(x)}$ is an important indicator in the unexplored areas and presents a good index to improve the initial DoE as will be shown later in this thesis.

2.5 CONCLUSION

In this chapter, a literature review on the different types of soil uncertainties encountered in geotechnical engineering (with an emphasis on the soil spatial variability) was firstly presented. This was followed by the methods of characterization and modeling of the soil spatial variability. In this thesis, the soil spatial variability was modeled by random fields described by their probability density functions and their autocorrelation functions. The review on the existing literature allows one to determine the frequently encountered values of the different soil statistical parameters (coefficient of variation of the different soil parameters, correlation between parameters, autocorrelation functions, etc.).

The review on the methods of discretization of random fields has shown that the series expansion methods are the best ones to be employed for random field discretization. This is because these methods provide the optimal number of random variables needed to discretize the random field, the other methods being mesh-dependent. In this thesis, EOLE method was adopted for the discretization of the random fields. Besides the fact that it is a series expansion method, EOLE allows one to determine the minimal number of random variables for a prescribed value of the variance of the error.

This chapter was also devoted to the principal existing methods used for uncertainty propagation. The different classical methods of computation of the failure probability were briefly presented. These methods were divided into three main categories which are the approximation methods, the simulation methods and the metamodeling techniques. The approximation methods include the First Order Reliability Method (FORM) and the Second Order Reliability Method (SORM). The simulation methods involve the Monte Carlo Simulation (MCS) methodology, Importance

Sampling (IS) and Subset Simulation (SS). Finally, the metamodeling techniques include, among others, Response Surface Method (RSM), Polynomial Chaos Expansion (PCE) and its extension Sparse Polynomial Chaos Expansion (SPCE), and kriging.

The literature review on different methods of uncertainty propagation has shown that the combination between a metamodeling technique and a simulation approach takes advantages of each technique used individually. The kriging metamodeling technique was suggested as being a relevant approach because of its powerful outputs (prediction mean and prediction variance) and it is retained as the meta-modelling approach within this thesis. This technique will be employed in combination with one of the three simulation techniques (MCS, IS or SS) to lead to more efficient probabilistic approaches with a reduced computation time as compared to each method employed individually.

CHAPTER 3. PROBABILISTIC ANALYSIS OF STRIP FOOTINGS RESTING ON SPATIALLY VARYING SOILS USING KRIGING AND MONTE CARLO SIMULATION

3.1 INTRODUCTION

As is well known, the spatial variability of the soil properties has a significant impact on the probabilistic outputs of geotechnical structures [Hicks and Samy (2002a, 2002b, 2004)]. Several authors have considered the effect of the spatial variability of the soil properties on the statistical moments (mean and standard deviation) of their system response or on the failure probability against a prescribed threshold of this response. For the probabilistic analysis of shallow foundations, which is the subject of the present thesis, one may cite among others [Fenton and Griffiths (2001); Griffiths et al. (2002); Fenton and Griffiths (2003); Popescu et al. (2005); Cho and Park (2010); Ching et al. (2011); Al-Bittar and Soubra (2013, 2014a, 2014b, 2016); Li et al. (2015)].

When dealing with the computation of the failure probability of geotechnical structures involving spatially varying soils, the classical Monte Carlo Simulation (MCS) methodology is generally used. This method is known to be very time-consuming. This is because (i) it usually makes use of finite element or finite difference models which are generally time-expensive and more importantly (ii) it requires a large number of calls of the mechanical model for the computation of the small failure probabilities encountered in practice. The computation time becomes excessive when considering a small target value of the coefficient of variation on the failure probability. Thus, one needs a method that keeps to a minimum the number of calls to the mechanical model when performing a probabilistic analysis.

In order to overcome the shortcoming related to the excessive number of calls of the mechanical model when performing a probabilistic analysis, Echard et al. (2011) proposed an Active learning reliability method combining Kriging and Monte Carlo Simulation (called AK-MCS). This

method consists in constructing a meta-model (i.e. an analytical equation which substitutes the original mechanical model). The computation of the failure probability may thus be easily performed using this meta-model.

The AK-MCS approach is based on kriging theory and makes use of a powerful learning function. In this method, an initial approximate kriging meta-model is constructed based on a small number of samples (called Design of Experiments DoE) computed using the mechanical model. This meta-model is then successively updated by adding each time a new sample chosen according to a powerful learning function. It should be emphasized here that since the computation of the failure probability requires only the sign of the performance function values, the objective of the learning function is to choose samples that have a high uncertainty on the sign of their performance function values (i.e. those that are close to the limit state surface).

The AK-MCS approach has been validated by Echard et al. (2011) by considering several academic examples involving non-linear limit state surfaces and high-dimensional stochastic problems, where the performance function was given by an analytical equation. This method was shown to be very efficient as the obtained probability of failure is very accurate needing a smaller number of calls to the mechanical model as compared to the crude MCS methodology.

The aim of this chapter is to extend the AK-MCS approach by Echard et al. (2011) to the study of geotechnical structures involving spatially varying soil properties. More specifically, this chapter presents a probabilistic analysis at the ultimate limit state, of a strip footing resting on a spatially varying soil and subjected to a vertical load. The objective is the computation of the probability of failure of the footing against soil punching. The soil cohesion and angle of internal friction (c and φ) were considered as two anisotropic non-Gaussian random fields. They are characterized by two specified marginal distribution functions and a common autocorrelation function. The Expansion Optimal Linear Estimation (EOLE) methodology proposed by Li and Der Kiureghian (1993) was used to generate these two random fields. The mechanical model used to calculate the

system response (i.e. the ultimate bearing capacity) was based on numerical simulations using the finite difference code FLAC^{3D}.

This chapter is organized as follows: The next two sections aim at presenting the EOLE method used for the discretization of the two anisotropic non-Gaussian random fields and the AK-MCS methodology as it may be employed for the probabilistic analysis of geotechnical structures involving a spatially varying soil. This is followed by an illustration of the AK-MCS approach *via* a simple bearing capacity problem (without spatial variability). Then, a probabilistic analysis of a strip footing resting on a spatially varying soil and the corresponding numerical results are presented and discussed. The chapter ends by a conclusion of the main findings.

3.2 THE EXPANSION OPTIMAL LINEAR ESTIMATION (EOLE) METHOD

Assume that the soil cohesion and friction angle are two non-Gaussian random fields that share the same autocorrelation function. These two random fields will be denoted by $Z_i^{NG}(x, y)\,(i = c, \varphi)$ and they will be described by two non-Gaussian marginal cumulative density functions G_i $(i = c, \varphi)$ and a common autocorrelation function ρ_Z^{NG} [(*x*, *y*), (*x*', *y*')] which gives the value of the correlation between two arbitrary points (*x*, *y*) and (*x*', *y*').

In this thesis, an anisotropic square exponential autocorrelation function was used. It is given as follows:

$$\rho_z^{NG}[(x,\ y),\ (x\ ',\ y\ ')] = \exp\left(-\left(\frac{|x - x\ '|}{a_x} \right)^2 - \left(\frac{|y - y\ '|}{a_y} \right)^2 \right) \tag{3.1}$$

where a_x and a_y are the autocorrelation distances along *x* and *y* respectively.

The discretization of the two random fields *via* EOLE may be described as follows (Li and Der Kiureghian 1993): one should first define a stochastic grid composed of s grid points (or nodes) and determine the common non-Gaussian autocorrelation matrix $\Sigma_{\chi;\chi}^{NG}$. Then, this matrix should be transformed into the Gaussian space using the Nataf correction functions proposed by Nataf

(1962) since the discretization of the random fields using EOLE is performed in Gaussian space. As a result, one obtains two Gaussian autocorrelation matrices $\Sigma^{c}_{\chi;\chi}$ and $\Sigma^{\varphi}_{\chi;\chi}$ that can be used to discretize the two Gaussian random fields at any point using the following equations:

$$\tilde{Z}_i(x, y) = \mu_i + \sigma_i \sum_{j=1}^{M} \frac{\xi_j^i}{\sqrt{\lambda_j^i}} \cdot \left(\phi_j^i\right)^T \cdot \Sigma^i_{z(x,y);\chi} \qquad i = c, \varphi \tag{3.2}$$

where μ_i and σ_i $(i = c, \varphi)$ are respectively the mean and standard deviation values of the two random fields, ξ_j^i $(i = c, \varphi)$ are two blocks of independent standard normal random variables, λ_j^i, ϕ_j^i $(i = c, \varphi)$ are the eigenvalues and eigenvectors of the two Gaussian autocorrelation matrices $\Sigma^{c}_{\chi;\chi}$ and $\Sigma^{\varphi}_{\chi;\chi}$ respectively, $\Sigma^{i}_{z(x,y);\chi}$ is the correlation vector between the values of the random field at the different nodes of the stochastic grid and its value at the arbitrary point (*x*, *y*) as obtained using Equation (3.1), and finally M is the number of terms (expansion order) retained in the EOLE method. This number will be determined later in this section.

Once the two Gaussian random fields [Equation (3.2)] are obtained, they should be transformed to non-Gaussian space by applying the following formula:

$$\tilde{Z}_i^{NG}(x, y) = G_i^{-1}\left\{\Phi\left[\tilde{Z}_i(x, y)\right]\right\} \qquad i = c, \varphi \tag{3.3}$$

where $\Phi(.)$ is the standard normal cumulative density function.

It should be mentioned that the series given by Equation (3.2) is truncated for a number of terms M (expansion order) smaller than the number of grid points s, after sorting the eigenvalues λ_j^c and λ_j^{φ} $(j = 1, \ldots, M)$ in a descending order. This number should assure that the variance of the error is smaller than a prescribed tolerance. Remember that the variance of the error for EOLE is given by Li and Der Kiureghian (1993) as follows:

$$Var\left[\, Z_i(x,\, y) - \tilde{Z}_i(x,\, y)\,\right] = \sigma_Z^2 - \sum_{j=1}^{M} \frac{1}{\lambda_j^i} \left(\left(\phi_j^i \right)^T \Sigma^i_{Z(x,y);\chi} \right)^2 \qquad (i = c,\, \varphi) \qquad (3.4)$$

where $Z_i(x,\, y)$ and $\tilde{Z}_i(x\,,\, y)$ are respectively the exact and the approximate values of the random fields at a given point (*x*, *y*).

3.3 AK-MCS METHODOLOGY FOR GEOTECHNICAL STRUCTURES INVOLVING SPATIALLY VARYING SOIL PROPERTIES

The basic idea of AK-MCS approach may be described as follows: In the AK-MCS method, one randomly selects a small Design of Experiments (DoE) from a large Monte Carlo population. Then, the kriging metamodeling technique is used to construct an approximate kriging meta-model based on the responses of this DoE (as computed using the mechanical model). This approximate kriging meta-model is successively improved (i.e. updated) by considering each time a new 'best' sample that is computed using the mechanical model. This process (called enrichment process) is repeated until one obtains a sufficiently accurate kriging meta-model that can be used for the computation of the failure probability.

Notice that the 'best' new sample is chosen using a powerful learning function that makes use of the prediction mean $\mu_{\hat{G}(X)}$ and the prediction variance $\sigma^2_{\hat{G}(X)}$ of the already-obtained kriging meta-model. As will be shown later, the chosen 'best' new samples are those that are close to the limit state surface since our interest is not to determine the accurate values of the performance function for the different samples but rather to accurately determine the signs of the performance function values for these samples. This is because the computation of the failure probability only requires the knowledge of the signs of the performance function values for the different samples. For more details on the AK-MCS approach as presented by Echard et al. (2011) in the case where the uncertain parameters are modeled by random variables, the reader may refer to Appendix B.

As was stated in the introduction, the aim in this chapter is to show the application of AK-MCS approach for the probabilistic analysis of geotechnical structures involving spatially varying soil

properties. This is achieved in this section by presenting a comprehensive step-by-step procedure describing the implementation of the AK-MCS methodology for the specific case of a strip footing resting on a spatially varying soil (where c and φ are two random fields) and the numerical software used for the mechanical analysis is FLAC^{3D}.

This procedure was implemented in Matlab software. It includes the random field discretization by EOLE and the construction of a kriging meta-model (for the computation of the failure probability) using the DACE toolbox. The implemented Matlab procedure makes several calls to the FLAC^{3D} code for the computation of the system response (i.e. ultimate bearing capacity on a spatially varying soil) or the corresponding performance function value for the different soil realizations. It may be described as follows (see also the flowchart presented in Figure 2.1):

1. Generate a population S of N_{Mc} (say N_{Mc} = 500,000) samples of M standard Gaussian random variables $\left\{\left(\xi_1^1,...,\xi_M^1\right),\left(\xi_1^2,...,\xi_M^2\right),...,\left(\xi_1^{N_{MC}},...,\xi_M^{N_{MC}}\right)\right\}$ where M is the number of random variables needed by EOLE methodology to discretize the two random fields c and φ. It should be emphasized here that each sample of standard Gaussian random variables provides (when substituted into Equations 3.2 and 3.3) typical spatial variations of c and φ that respect the correlation structure of these fields, i.e. the so-called 'realizations' of c and φ. This is performed by computing, for this sample, the values of c and φ at the centroids of the different elements of the FLAC^{3D} mesh. Notice that the difference between the different realizations lies in the position of the weak and strong soil zones although all realizations respect the correlation structure of the corresponding random fields. It should be emphasized here that the computation of the performance function values for the generated samples (based on FLAC^{3D} software) is not required at this stage. These samples will be denoted as candidate samples in this chapter.

2. Randomly select from the population S a small number of samples, i.e. a small DoE of size N_1 (say N_1 = 20). Then, use EOLE methodology to transform each sample into realizations of c and

φ that provide the spatial distribution of the soil cohesion and angle of internal friction respectively. For the N_1 samples, evaluate the performance function values using the following equation:

$$G = \frac{q_u}{q_s} - 1 \tag{3.5}$$

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

3. Based on the DoE and the corresponding performance function values, construct an approximate kriging meta-model in the standard space of random variables using the DACE toolbox.

4. Use again the DACE toolbox in order to compute (for the whole population *S* containing the N_{MC} samples) both the kriging predictor values $\mu_{\hat{G}}$ and their corresponding kriging variance values $\sigma^2_{\hat{G}}$ using the meta-model. From the obtained values of the kriging predictors $\mu_{\hat{G}}$, obtain an estimation of the probability of failure P_f by counting the number of negative predictors $N_{\hat{G}\leq 0}$ and dividing it by the total number of samples in S as follows:

$$P_f = \frac{N_{\hat{G}\leq 0}}{N_{MC}} \tag{3.6}$$

Also, compute the coefficient of variation of P_f as follows:

$$COV\left(P_f\right) = \sqrt{\frac{1-P_f}{P_f \cdot N_{MC}}} \tag{3.7}$$

5. Identify the 'best' next candidate sample in *S* for which one will compute the performance function value using FLAC^{3D}. This is performed by evaluating a learning function U for each sample in *S*. The learning function U is given by:

$$U\left(X_i\right)=\frac{\left|\mu_{\hat{G}(X_i)}\right|}{\sigma_{\hat{G}(X_i)}} \qquad i=1,\ \ldots,\ N_{MC} \tag{3.8}$$

The 'best' next sample is the one with minimum value of U (more details on this criterion for the choice of the 'best' sample will be given later on in this section).

6. If this minimum value of U is smaller than 2, the performance function value based on FLAC^{3D} is evaluated for this 'best' candidate and the initial DoE is updated. Then, one should go back to step 3 and evaluate a new kriging model based on the updated DoE. Steps (3), (4) and (5), which constitute the enrichment process, are repeated until the minimum value of U becomes larger than 2. More details on this stopping criterion will be given later on in this section.

At this stage, the learning stops and the meta-model is considered accurate enough for the computation of the estimated values of both the probability of failure P_f and the coefficient of variation $COV\left(P_f\right)$.

It should be emphasized herein that a small initial DoE is chosen within the present approach (see step 2) in order to keep to a minimum the number of calls to the computationally expensive mechanical model. This initial DoE is successively increased by a single sample at each time (see step 5). The chosen sample is the one that mostly improves the meta-model because Equation (3.8) searches for the sample that has a small kriging predictor (i.e. a sample that is close to the limit state surface) and/or a high kriging variance (i.e. a high uncertainty in the sign of its performance function value). Notice that the samples with high uncertainties in the sign of their performance function values (positive or negative) are those that are close to the limit state surface. Finally, notice that the stopping criterion [min(U)]>2 corresponds to a maximal probability of making a mistake on the sign of the performance function value of $\Phi(-2)=0.023$ [see Echard et al. (2011)]. This means that the stopping criterion is relevant. It makes use of the

samples with a small probability of making a mistake in the signs of their performance function values.

At the end, it should be emphasized that the number of predictions by kriging can be important as the whole Monte Carlo population is estimated. Notice however that the computation time of the predictions is much smaller than that required to evaluate the performance function values using the computationally expensive mechanical model.

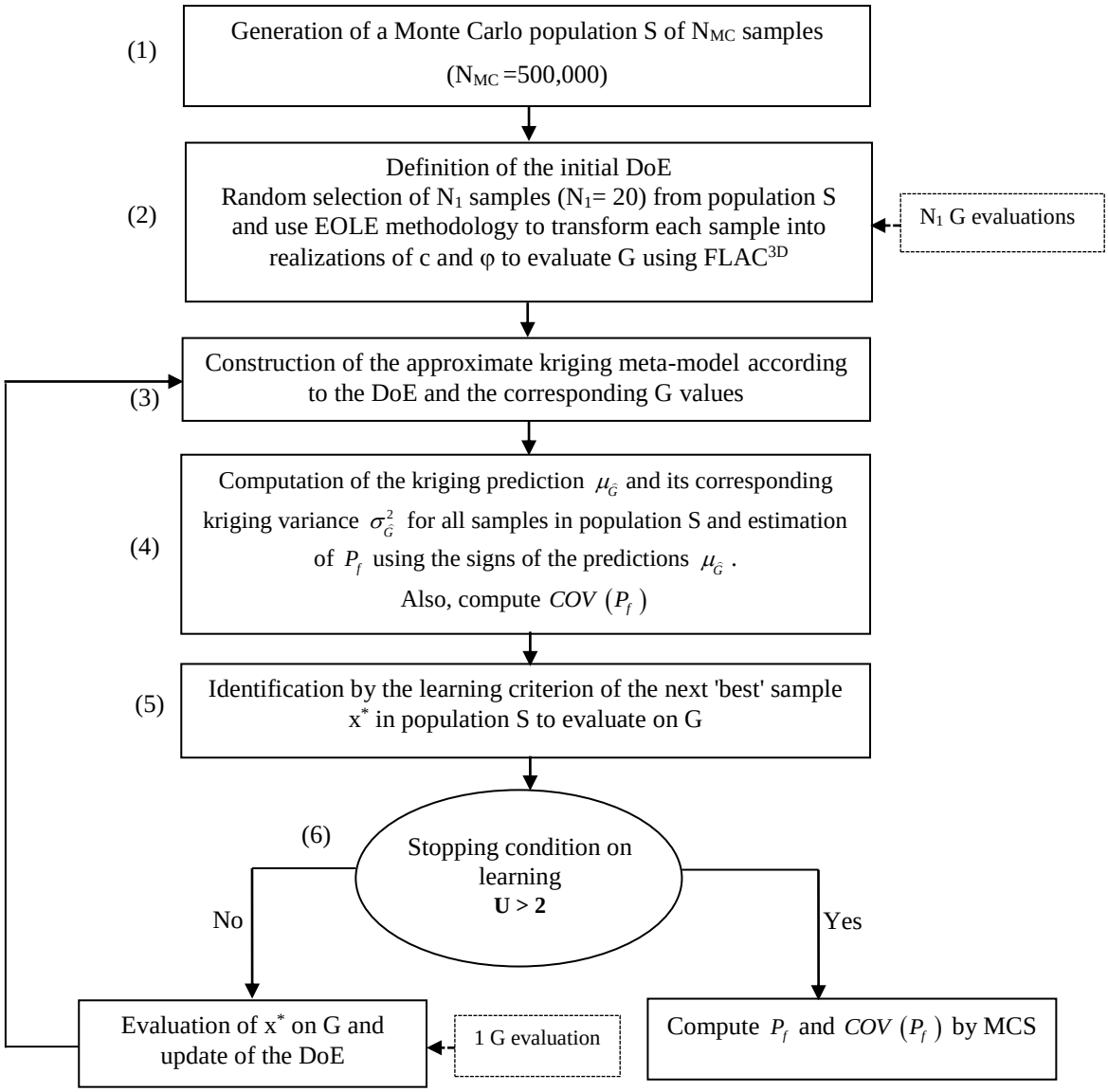


Figure 3.1. AK-MCS flowchart

3.4 ILLUSTRATION OF AK-MCS PROCEDURE VIA AN ANALYTICAL EXAMPLE

This section aims at illustrating the performance of the AK-MCS procedure *via* a simple bearing capacity problem (i.e. without considering the soil spatial variability) where the system response is given by a simple analytical equation with a quasi-negligible computation time. The performance of AK-MCS approach was illustrated by comparison of its results to those given by the crude MCS methodology.

The probabilistic analysis presented in this section involves the computation of the failure probability against soil punching of a shallow strip footing of breadth B=2m resting on a homogeneous (c, φ) soil and subjected to a surcharge loading q=10 kN/m^2. The soil unit weight was taken equal to γ=18 kN/m^3. The applied footing load was equal to q_s=400 kN/m. The uncertain parameters considered in the analysis are the soil shear strength parameters c and φ. The illustrative statistical parameters of these two random variables, as used in the present analysis, are those commonly encountered in practice [cf. Phoon and Kulhawy (1999) among others] and they are presented in Table 3.1. No cross-correlation between c and φ is considered in this analysis.

Table 3.1. Statistical characteristics of the random variables

Random variable	Mean μ	Coefficient of variation COV (%)	Type of the probability density function (PDF)
C	20 kPa	25	Log-normal
Φ	30^o	10	Beta

The performance function used to calculate the failure probability is given by the following equation:

$$G = \frac{q_u}{q_s} - 1 \tag{3.9}$$

where q_u is the ultimate bearing capacity and it is given by:

$$q_u = \frac{1}{2}\gamma B N_\gamma + cN_c + qN_q \tag{3.10}$$

In this equation, *N*γ, *Nq* and *Nc* are the bearing capacity factors due to the soil weight, surcharge loading and cohesion, respectively. The adopted factors employed in this section are those suggested by Vesic (1973). These factors are widely used in routine foundation design. They are functions of only the soil angle of internal friction φ.

Notice that the reference values adopted for both the failure probability P_f and the coefficient of variation $COV\left(P_f\right)$ are those obtained by the crude MCS runs with $N_{Mc} = 10^6$ samples. As may be seen from Table 3.2, both methods (MCS and AK-MCS) provide the same values for the probabilistic outputs (i.e. P_f and $COV\left(P_f\right)$) although AK-MCS method needs only 30 calls of the mechanical model (which correspond to 20 samples from the initial DoE plus 10 added samples during the enrichment process). This number of calls to the mechanical model is to be compared to that required by MCS that needs 10^6 samples to lead to the same value of $COV\left(P_f\right)$. This clearly illustrates the benefit of using the AK-MCS approach instead of the crude MCS methodology. It should be mentioned here that the deterministic safety factor for the studied configuration was equal to $F_s = q_u / q_s = 1190/400 \approx 3.0$.

In order to better understand the impact of the number of added samples (during the enrichment process) on the probabilistic outputs, Figure 3.2(a) and Figure 3.2(b) present the plots of P_f and $COV\left(P_f\right)$ versus the number of added samples. On the other hand, Figure 3.2(c) presents the value of the learning function U that was obtained each time a new added sample with index *i* (*i*=1, …, 10 in this example) was computed using the mechanical model.

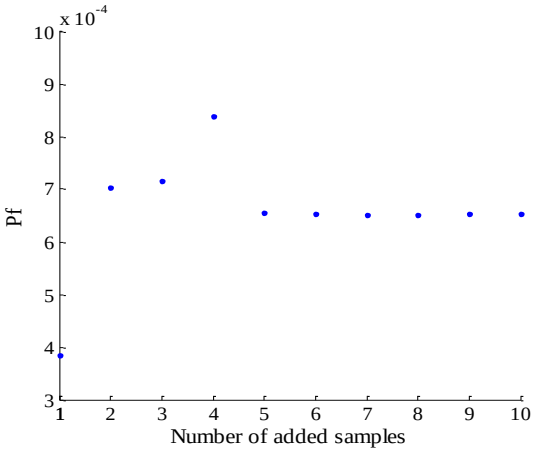
As may be seen from Figure 3.2(c), the new samples chosen during the enrichment process for the computation by the mechanical model have increasing values of U. The process of adding new samples has stopped (as required by the algorithm of the AK-MCS approach) when the value

of U became greater than 2. In the present analytical problem, 10 added samples were necessary to achieve this goal. On the other hand, one may observe from Figure 3.2(a) and Figure 3.2(b) that P_f and $COV\left(P_f\right)$ attain an asymptote when the number of added samples becomes greater than 5. This means that this number of added samples could be used to obtain the values of P_f and $COV\left(P_f\right)$ corresponding to convergence. For this number of added samples, the meta-model becomes sufficiently accurate (i.e. with no bias) to compute P_f and $COV\left(P_f\right)$.

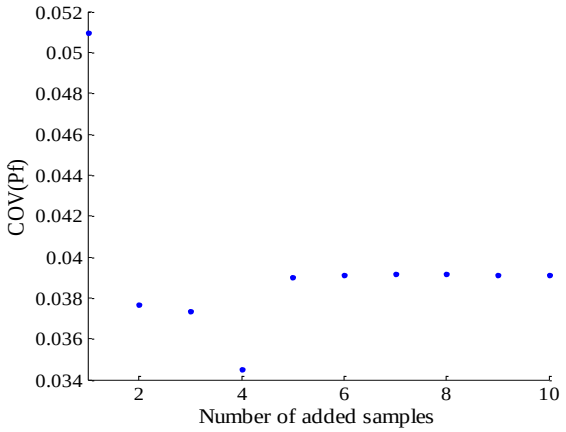
It should be emphasized here that in the subsequent computations within this chapter, the process of adding new samples has stopped only when one obtains a value of U greater than 2. This criterion was shown to provide convergence of P_f and $COV\left(P_f\right)$ for all the soil configurations studied in this chapter as may be seen from Appendix D.

Table 3.2. Probability of failure P_f , coefficient of variation $COV\left(P_f\right)$, and number of calls of the mechanical model N_{calls} as obtained by MCS and AK-MCS

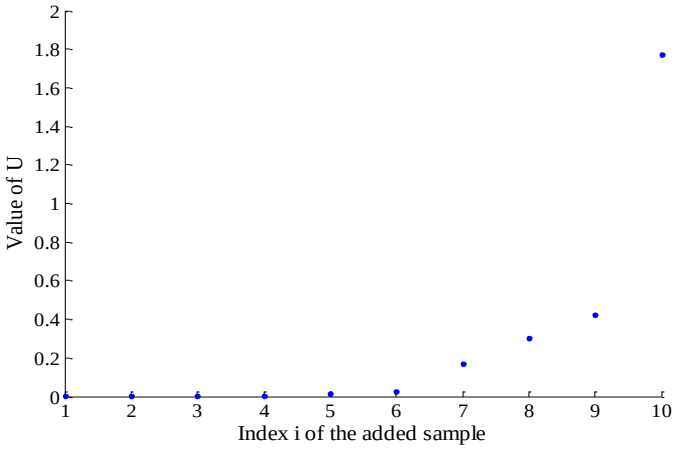
Method	P_f × 10^{-4}	$COV\left(P_f\right)$ (%)	N_{calls}
MCS	6.530	3.912	10^6
AK-MCS	6.530	3.910	20 samples for the initial DoE + 10 added samples = 30 samples



(a) P_f versus the number of added samples



(b) $COV\left(P_f\right)$ versus the number of added samples



(c) Value of U for the different added samples (i=1,…,10)

Figure 3.2. AK-MCS results of the analytical example

3.5 PROBABILISTIC ANALYSIS OF A STRIP FOOTING RESTING ON A SPATIALLY VARYING SOIL MASS

This section aims at presenting the impact of the soil spatial variability on the probability of failure against soil punching of a strip footing subjected to a vertical load. Thus, the system response involves the ultimate bearing capacity (q_u) of a vertically loaded strip footing resting on a spatially varying soil. The performance function employed in the analysis was the one defined by Equation (3.9) where q_u was based on numerical simulations using FLAC^{3D} model.

The soil cohesion *c* and angle of internal friction φ were modeled as two anisotropic non-Gaussian random fields. The EOLE methodology 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). The illustrative statistical parameters of these two random fields as used in the present analysis are those presented in Table 3.1. Recall here that the same autocorrelation function (square exponential) was used for both *c* and φ. 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. All the other parameters of the soil, footing and interface were assumed to be deterministic.

Notice finally that the number of Monte Carlo samples N_{Mc} used in all subsequent computations was equal to 500,000.

3.5.1 The mechanical model

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 Young modulus E and Poisson ratio υ were assumed as follows: $E = 60MPa$ and $\nu = 0.3$. Concerning the footing, a weightless strip foundation of 1m width and 0.25m height was used. It was assumed to follow an elastic linear model ($E = 25GPa$, $\nu = 0.4$). The connection between the footing and the soil mass was modeled by interface elements having the same mean values of the soil shear strength parameters in order to simulate a perfectly rough soil-footing interface. Concerning the elastic properties of the interface, their values were as follows: $K_s = 1GPa$, $K_n = 1GPa$ where K_s and K_n are respectively the shear and normal stiffness of the interface.

As shown in Figure 3.3, a strip footing of width B that rests on a soil domain of width 13B and depth 5B was considered in the analysis. For the displacement boundary conditions, the bottom boundary was assumed to be fixed and the vertical boundaries were constrained in motion in the horizontal direction. For the computation of the ultimate bearing capacity of the rigid rough strip footing subjected to a central vertical load using FLAC^{3D}, the following method was adopted:

An optimal controlled downward vertical velocity of 5 $\times 10^{-6}$ m/time step (i.e. displacement per time step) was applied to the bottom central nodes of the footing. Damping of the system is introduced by running several cycles until a steady state of plastic flow is developed in the soil underneath the footing. At each cycle, the vertical footing load is obtained by using a FISH function that calculates the integral of the normal stress components for all elements in contact with the footing. The value of the vertical footing load at the plastic steady state is the ultimate footing load. For more details on FLAC^{3D} software and its application to the computation of the ultimate bearing capacity of a strip footing, the reader may refer to Appendix C.

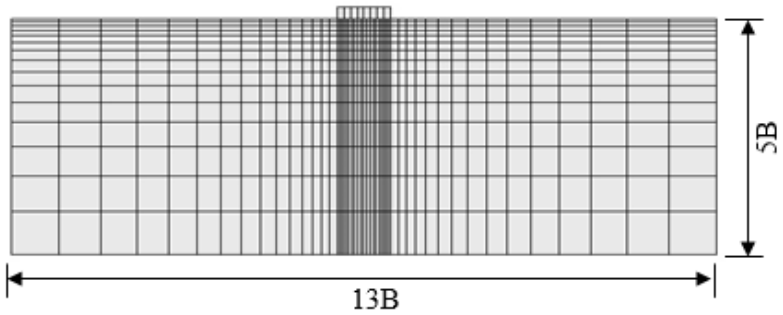


Figure 3.3. Soil domain and the corresponding mesh in the FLAC^{3D} mechanical model

3.5.2 Probabilistic numerical results

For the considered soil domain, and for the different values of the autocorrelation distances (a_x, a_y) used in the analysis, the adopted number N of eigenmodes (or random variables) that is used to discretize the two random fields c and φ is chosen in such a manner that the variance of the error is smaller than 5% for most of the studied configurations (see columns 2 and 3 of Table

3.5). Only for some configurations with small values of the autocorrelation distances, a greater value of the variance of the error was adopted in the analysis. Indeed, these configurations with big fluctuations require a larger number of eigenmodes or random variables (greater than 50) to lead to a smaller variance of the error. For this high number of random variables, the meta-modeling technique by AK-MCS becomes very time-expensive. More advanced probabilistic approaches are needed for these configurations.

It should be mentioned here that a much greater value of the variance of the error (of 10%) was found sufficient to accurately approximate the first two statistical moments (mean and standard deviation) of the same bearing capacity problem treated herein [cf. Al-Bittar and Soubra (2013, 2014b, 2016)]. This is to be expected since the small fluctuations of the random fields have no significant effect on the first two statistical moments of the system response.

3.5.2.1 Probability of failure P_f and $COV\left(P_f\right)$ versus the number of added realizations during the enrichment process

Figure 3.4 presents the probability of failure P_f and the coefficient of variation $COV\left(P_f\right)$ versus the number of added realizations during the enrichment process in the case where a_x=10m and a_y=1m. The number of random variables adopted for this configuration was equal to 32 as may be seen from Table 3.5(b). Figure 3.4 also provides the learning function values for the different added realizations. It should be mentioned here that 752 realizations were needed during the enrichment process in addition to the initial DoE before the algorithm stops [min(U)]>2. The final obtained values of P_f and $COV\left(P_f\right)$ are respectively 1.656×10^{-3} and 3.47%. Figure 3.4 shows that the probability of failure starts to converge at about 727 calls to the mechanical model. This means that there is no bias in the meta-model beyond this number of calls of the mechanical model.

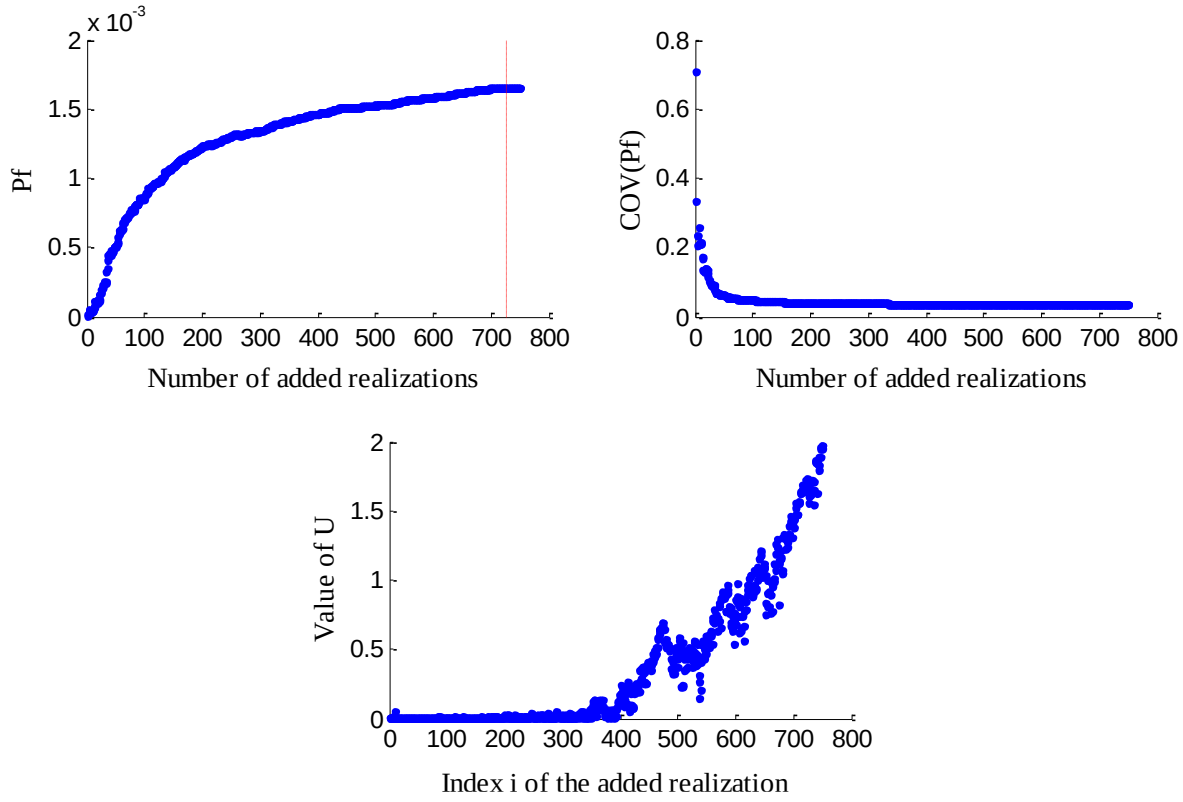


Figure 3.4. AK-MCS results for a spatially varying soil (a_x=10 m, a_y=1 m)

3.5.2.2 Effect of number of added realizations on the limit state surface

Figure 3.5 shows the evolution of the limit state surface (LSS) corresponding to G=0, with the number of added realizations (from 5 to 35 realizations where the number of needed realizations was equal to 37) in the case where a_x= a_y=10,000m. The number of random variables adopted for this configuration was equal to 2 as may be seen from Table 3.5(a). This small number of random variables allows one to visualize the evolution of the LSS since only a two-dimensional space is needed.

As may be easily seen from Figure 3.5, the increase in the number of added realizations improves the limit state surface in the zone that has an impact on the value of the failure probability (i.e. in the zone corresponding to the high values of the probability density which is close to the origin of the standard space); the other zones (far from the origin) being with a non-significant effect on

the value of the failure probability. It should be mentioned that the LSS starts to converge after 30 added samples.

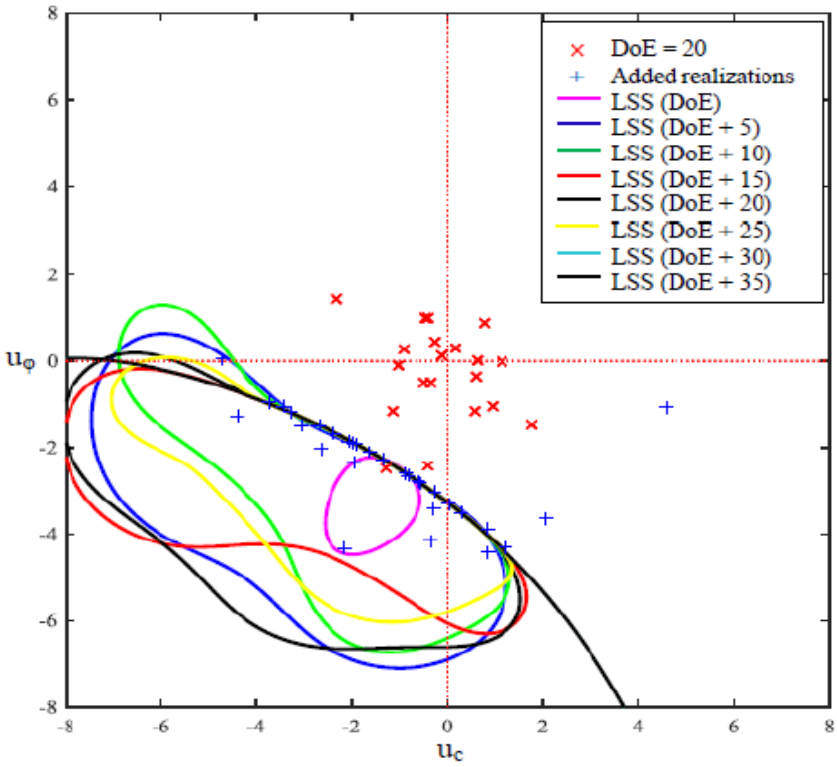


Figure 3.5. Effect of the number of added realizations on the limit state surface when a_x = a_y = 10,000 m

3.5.2.3 Effect of the variance of the error of EOLE methodology on the probabilistic outputs

In order to reduce the effect of the error of discretization of random fields on the estimated value of the failure probability, the variance of the error of EOLE methodology must be sufficiently small. Table 3.3 provides the effect of the number of random variables used in the discretization scheme (and the corresponding value of the variance of the error of EOLE methodology) on the value of P_f in the case where a_x=10m and a_y=5m.

As may be seen from this table, the percent difference on P_f (with respect to the 'accurate' solution corresponding to 24 variables) is equal to 19.51% when adopting 6 random variables and it decreases to 7.3% if one adopts 12 random variables. Notice however that a significant number

of added realizations (406 realizations) is needed to achieve such accuracy. The small value of error increases the stochastic dimension of the problem and thus a larger number of added realizations will be needed to achieve such accuracy. In this thesis, a threshold of 5% may be adopted as an acceptable value for the variance of the error within EOLE. This value may be expected to lead to an acceptable value of the failure probability.

Table 3.3. Effect of the number of random variables on P_f and $COV\left(P_f\right)$ when a_x =10 m and a_y =5 m

Number of random variables for the two random fields (c, φ)	Variance of the error %	$P_f \times 10^{-3}$	$COV\left(P_f\right)$ %	δ_{Pf} %	Number of added realizations
6	9.876	2.822	2.66	19.51	115
12	1.682	3.25	2.48	7.3	406
24	0.0811	3.506	2.38	-	695

3.5.2.4 Effect of the safety factor on the probabilistic outputs

Table 3.4 provides the values of P_f and $COV\left(P_f\right)$ for three values of the safety factor F_s when a_x=a_y=10,000m. This table also provides the number of added realizations which was needed for each case. This number is smaller for the configurations corresponding to the higher values of the safety factor (i.e. those corresponding to the smaller values of the failure probability or the greater values of the reliability index). This may be explained by the fact that for a higher value of the safety factor, the limit state surface becomes farther away from the origin of the standard space (as may be easily seen from Figure 3.6). Thus, the probability density becomes much lower along the limit state surface in the case F_s=3.5 leading to a smaller number of available samples for the prescribed number of Monte Carlo simulations. Notice that the smaller number of available samples leads to a larger value of $COV\left(P_f\right)$. Consequently, a larger number of Monte Carlo simulations is needed in the case of larger values of the safety factor to lead to smaller values of $COV\left(P_f\right)$.

Table 3.4. Effect of the safety factor F_s on P_f , $COV\left(P_f\right)$ and β_{HL} for the case a_x=a_y=10,000 m

F_s	$P_f \times 10^{-3}$	$COV\left(P_f\right)\%$	β_{HL}	Number of added realizations
2.5	3.806	2.2879	2.638	37
3	0.748	5.1689	3.126	20
3.5	0.138	12.037	3.587	12

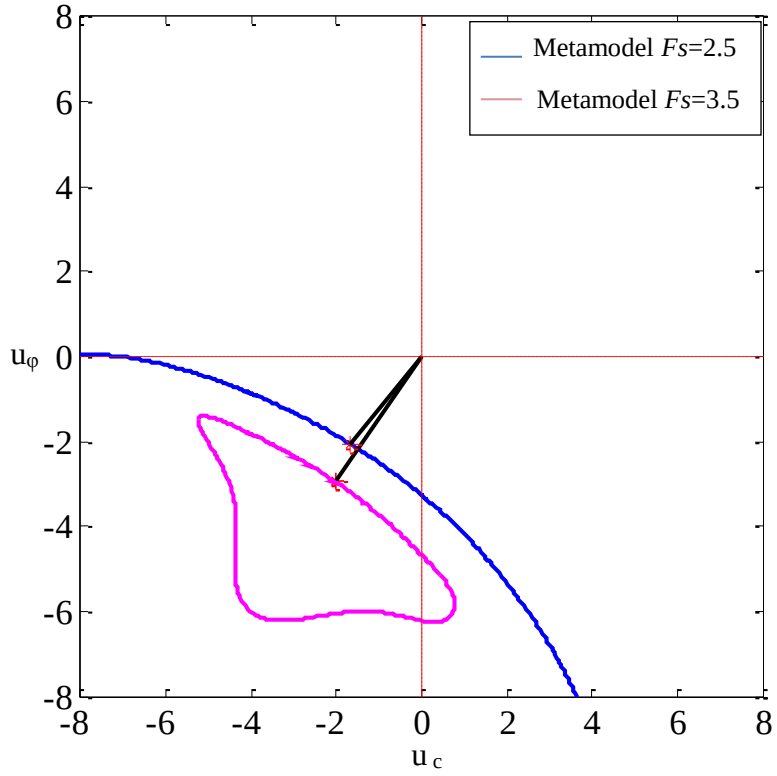


Figure 3.6. Effect of the safety factor on the limit state surface when a_x=a_y= 10,000 m for two values of Fs

3.5.3 Probabilistic parametric study

Columns 2 and 3 of Table 3.5 provide the number of eigenmodes and the corresponding variance of the error of EOLE methodology for different values of the autocorrelation distances. Columns 4, 5 and 6 of the same table provide the failure probabilities, the corresponding values of the coefficient of variation together with the number of added realizations. As may be seen from this table, the number of random variables is small for the very large values of the autocorrelation distances and significantly increases for the small values of the autocorrelation distances.

In order to show that no bias in the meta-model exists at the end of the enrichment process for all the configurations presented in Table 3.5, Appendix D presents the plots of P_f for the different values of the autocorrelation distances considered in this table. For completeness, this appendix also provides the plots of $COV\left(P_f\right)$ and U for these configurations.

3.5.3.1 Effect of the autocorrelation distances on P_f and β_{HL}

For the case of isotropic random fields, Table 3.5(a) presents the effect of the isotropic autocorrelation distance ($a_x = a_y$) on the failure probability P_f and the corresponding value of the coefficient of variation $COV\left(P_f\right)$. As may be seen from Figure 3.7, P_f increases with the increase in the isotropic autocorrelation distance. However, the rate of increase gets smaller for the large values of the autocorrelation lengths (when $a_x = a_y > 10$ m) to attain an asymptote corresponding to the case of a homogeneous soil [see also Thajeel et al. (2016)]. Indeed, for the small values of the isotropic autocorrelation distance, the soil heterogeneity leads (for most realizations) to relatively high values of the ultimate bearing capacity due to the averaging phenomenon. This means that the number of realizations leading to failure is very small in this case. On the contrary, the number of realizations leading to failure is higher in the case of a homogeneous soil due to the fact that the realizations are homogeneous in this case with either small or high values of the soil resistance.

For the case of anisotropic random fields, Table 3.5(b) presents the effect of the vertical autocorrelation distance (a_y) on the failure probability P_f when a_x=10m. Similarly, Table 3.5(c) presents the effect of the horizontal autocorrelation distance (a_x) on the failure probability when a_y=2m. For the very large values of the horizontal or vertical autocorrelation distance, the failure probability tends to a constant maximal value corresponding to the case of a 1D random field as may be seen from Figure 3.8 and Figure 3.9. The reason is similar to that of the isotropic case.

Indeed, the increase in the soil heterogeneity leads (for most realizations) to relatively higher values of the ultimate bearing capacity due to the averaging phenomenon. This means that the number of realizations leading to failure is very small in this case and becomes higher for a lesser degree of heterogeneity in the soil mass.

For both cases of isotropic and anisotropic random fields, Figures 3.7, 3.8 and 3.9 show that the Hasofer Lind reliability index (as obtained from optimization and making use of the final kriging meta-model) decreases as expected with the increase in the autocorrelation distances. Also, one may observe from Table 3.5 that a small value of the coefficient of variation of the failure probability (smaller than about 5% for most cases) was obtained for the adopted value of N_{MC} which indicates that the obtained results are sufficiently accurate for practical use. The number of added realizations required to lead to a good approximation of the kriging model seems to be larger for the smaller values of the autocorrelation distance (because of the increasing fluctuations of the highly heterogeneous soils), although there is no regular increase in the number of added realizations with the decrease in the autocorrelation distance. Indeed, this number depends on the evolution of the kriging meta-model during the enrichment process.

Table 3.5. Adopted number of random variables and the corresponding value of the variance of error of EOLE together with the values of P_f , $COV\left(P_f\right)$ and number of added realizations for various soil variabilities

a. Case of an isotropic case (a_x=a_y)

a_x=a_y (m)	Adopted number of random variables	Variance of the error %	P_f $\times 10^{-3}$	$COV\left(P_f\right)$ %	Number of added realizations
2	48	9.447	0.730	5.23	742
3	32	4.647	1.846	3.29	995
5	24	0.953	2.762	2.69	870
10	10	0.815	3.444	2.41	286
20	8	0.170	3.648	2.34	210
50	6	0.016	3.736	2.31	105
100	6	0.001	3.772	2.29	100
10000	2	0.000048	3.806	2.28	37

b. Case of an anisotropic case (a_x=10 m with varying a_y)

a_y (m)	Adopted number of random variables	Variance of the error %	P_f $\times 10^{-3}$	$COV\left(P_f\right)$ %	Number of added realizations
0.5	44	9.119	0.318	7.93	427
0.8	38	4.798	1.178	4.12	790
1	32	4.212	1.656	3.47	752
2	24	1.437	2.818	2.66	672
5	12	1.682	3.250	2.48	406
10	10	0.815	3.444	2.41	286
20	8	0.855	3.502	2.39	190
50	8	0.297	3.570	2.36	239
100	8	0.099	3.616	2.35	232
10000	8	0.033	3.690	2.32	227

c. Case of an anisotropic case (a_y=2 m with varying a_x)

a_x (m)	Adopted number of random variables	Variance of the error %	P_f $\times 10^{-3}$	$COV\left(P_f\right)$ %	Number of added realizations
2	48	9.447	0.730	5.23	742
5	30	4.101	2.116	3.07	824
10	24	1.437	2.818	2.66	672
20	16	1.415	3.060	2.55	494
50	12	1.272	3.166	2.51	357
100	10	0.842	3.202	2.49	256
10000	10	0.432	3.208	2.49	257

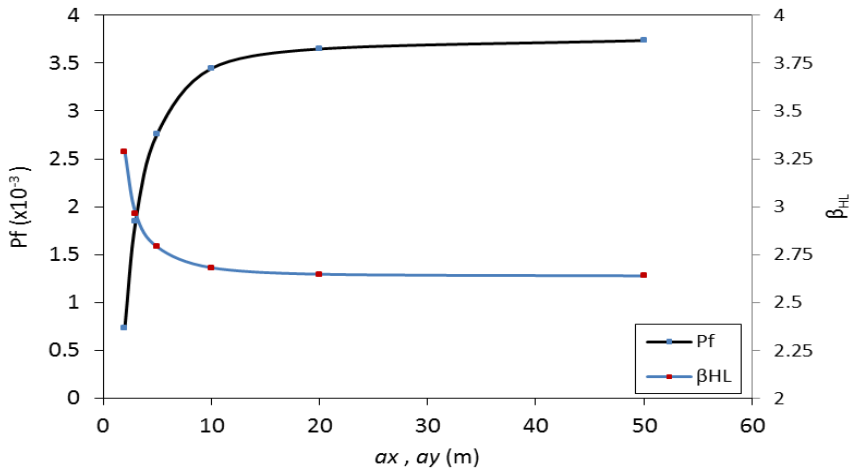


Figure 3.7. Effect of the isotropic autocorrelation distance a_x=a_y on P_f and β_{HL}

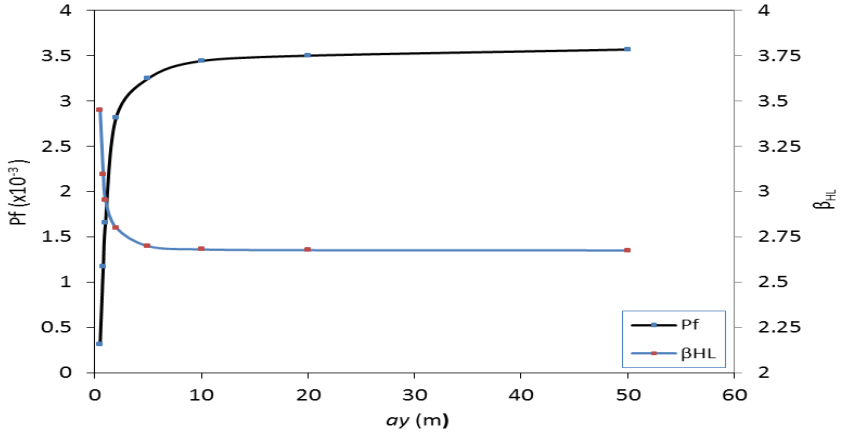


Figure 3.8. Effect of the vertical autocorrelation distance a_y on P_f and β_{HL} when a_x =10 m

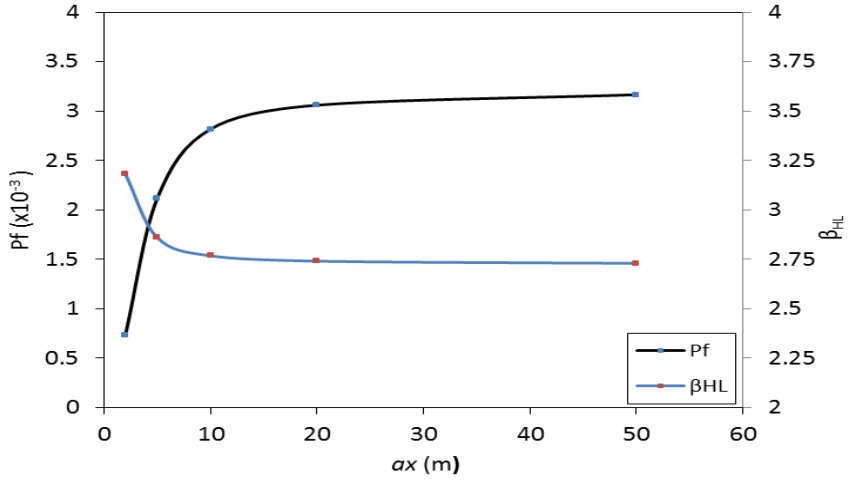


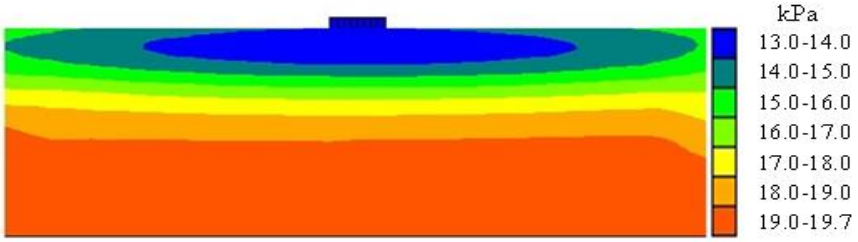
Figure 3.9. Effect of the horizontal autocorrelation distance a_x on P_f and β_{HL} when a_y =2 m

3.5.3.2 Effect of the autocorrelation distances on the critical realizations

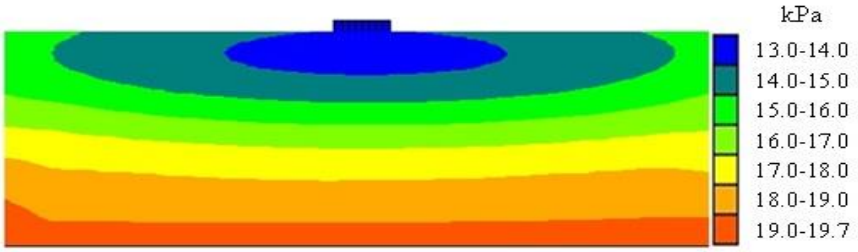
Figure 3.10 provides the realizations corresponding to the design point (i.e. those corresponding to the minimal value of the Hasofer-Lind reliability index) in the case where a_x=10m and for various values of a_y. Only the realizations of the soil cohesion are presented; those of the soil angle of internal friction being of similar trend. These realizations show a symmetrical pattern with respect to the vertical symmetrical axis of the foundation, whatever are the values of the autocorrelation distances for both cases of isotropic and anisotropic random fields. Furthermore, the small values of the soil cohesion and angle of internal friction are located in the neighborhood of the foundation, the higher values of the soil resistance being far from the footing.

The symmetrical pattern with respect to the central vertical axis of the foundation naturally leads to a failure mechanism that develops symmetrically within the weak soil zone near the foundation as may be seen from Figure 3.11 in the case where a_x=10 m and a_y=1 m.

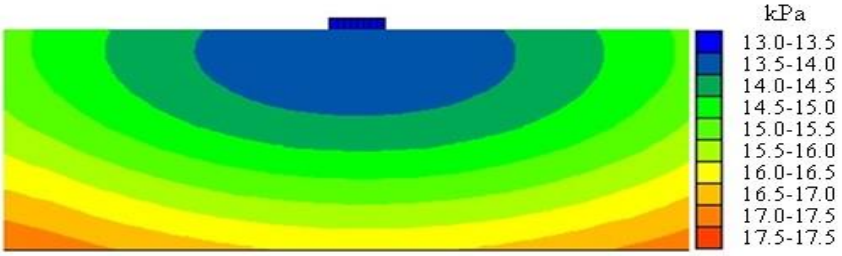
Finally, it should be noted that the critical realizations presented in this section are those that are the most probable ones (i.e. those corresponding to the maximal value of the probability density) for a soil punching under a vertical load. This explains the symmetrical pattern of the soil parameters with the smallest values near the foundation.



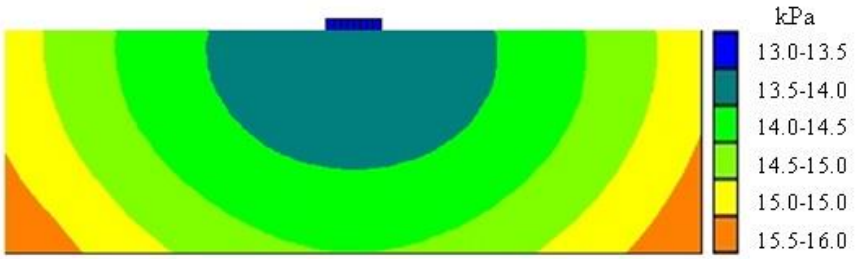
(a) a_x=10m and a_y=1m



(b) a_x=10m and a_y=2m

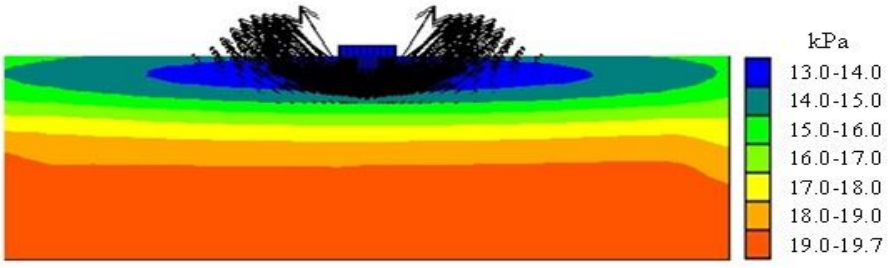


(c) a_x=10m and a_y=5m



(d) a_x=10m and a_y=10m

Figure 3.10. Realizations of the cohesion random field as obtained at the design point for different values of the autocorrelation distances



(a) a_x=10 m and a_y=1 m

Figure 3.11. Realization of the cohesion random field as obtained at the design point and the superimposed failure mechanisms for the case where a_x=10 m and a_y=1 m

3.6 CONCLUSION

The probabilistic analysis of shallow foundations resting on spatially varying soils was generally performed in literature using Monte Carlo Simulation (MCS) methodology. The mean value and the standard deviation of the system response were extensively investigated. This was not the case for the failure probability because MCS methodology requires a large number of calls of the mechanical model to accurately calculate the small failure probabilities encountered in practice. This chapter presented a probabilistic analysis at the ultimate limit state of a strip footing resting on a spatially varying soil using an Active learning reliability method combining Kriging and Monte Carlo Simulation (called AK-MCS). Within this method, one performs a Monte Carlo simulation without evaluating the whole population using the original computationally expensive mechanical model. Indeed, the population is predicted using a kriging meta-model which is defined using only a few points of the population that are evaluated employing the mechanical model. The main findings of this study can be summarized as follows:

1. The present AK-MCS method was shown to be very efficient as the obtained probability of failure is very accurate (i.e. with a small value of the coefficient of variation on this failure probability of about 5% maximum for most configurations) needing a smaller number of calls to the mechanical model (and thus, a smaller computation time) as compared to the crude MCS methodology. Indeed, for a target coefficient of variation on P_f of 5% and a value of the failure probability of 10^{-3}, the crude MCS requires about 800,000 simulations; however, the present AK-MCS approach requires (for the same accuracy) a maximal number of simulations of about 1000 simulations.
2. The computation time of AK-MCS remains important for the small values of the failure probabilities encountered in practice (although in this method one makes use of the predictions computed using the kriging meta-model) since a large population is required

by MCS in these cases to lead to a small value of the coefficient of variation on the failure probability.

3. The failure probability P_f increases with the increase in the autocorrelation distances. Notice however that the rate of increase gets smaller for the large values of the autocorrelation lengths (when a_x=a_y >10 m, a_x>10m, a_y>5m). Indeed, these cases closely resemble to the limit cases corresponding to a homogeneous soil or a 1D random field.
4. The numerical results have shown that the number of added realizations required to lead to a good approximation of the kriging model seems to be larger for the smaller values of the autocorrelation distance (because of the increasing fluctuations of the highly heterogeneous soils), although there is no regular increase in the number of added realizations with the decrease in the autocorrelation distance. Indeed, this number depends on the evolution of the kriging meta-model during the enrichment process.
5. The realizations corresponding to the design point have shown a symmetrical pattern with respect to the vertical symmetrical axis of the foundation, whatever are the values of the autocorrelation distances for both cases of isotropic and anisotropic random fields. Furthermore, the small values of the soil cohesion and angle of internal friction are located in the neighborhood of the foundation, the higher values of the soil resistance being far from the footing. This naturally leads to a failure mechanism that develops symmetrically within the weak soil zone near the foundation.
6. The prescribed value of the variance of the error of EOLE methodology was shown to have a significant influence on the value of the failure probability. A threshold of 5% may be adopted to obtain accurate values of the failure probability. Consequently, the results of the few configurations corresponding to a greater value of the variance of the error may be used with caution. Indeed, these configurations need a significant number of random variables (about 60 random variables) to lead to a variance of the error that is smaller than

5%. This number of random variables cannot be used within the AK-MCS approach because of the huge time necessary for those cases. A more advanced probabilistic method is needed to handle such configurations.

7. The numerical results have shown that the meta-model does not exhibit any bias beyond a number of added realizations that is smaller than the one required by the stopping criterion. This means that the present stopping criterion is not optimal and another more advanced stopping criterion may lead to a smaller number of calls of the mechanical model and thus to a smaller computation time.

CHAPTER 4. PROBABILISTIC ANALYSIS OF STRIP FOOTINGS RESTING ON SPATIALLY VARYING SOILS USING IMPORTANCE SAMPLING AND KRIGING METAMODELING

4.1 INTRODUCTION

In the previous chapter, an efficient approach was proposed for the computation of the failure probability against soil punching of a strip footing resting on a spatially varying soil. An active learning reliability method combining kriging and Monte Carlo simulation (called AK-MCS) was presented. This method consists in constructing a meta-model (i.e. an analytical equation which substitutes the original mechanical model). The computation of the failure probability may thus be easily performed using this meta-model. Notice that AK-MCS is based on the kriging theory and it makes use of a powerful learning function for the selection of the 'best' samples to be computed by the computationally expensive mechanical model. Thus, it overcomes the shortcoming of the crude MCS related to the excessive number of calls of the mechanical model when performing a probabilistic analysis.

When dealing with the small failure probabilities encountered in practice, the computation time of AK–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 high number of samples is required by MCS to lead to a small value of the coefficient of variation on the failure probability. For instance, 1,000,000 samples are required for the computation of P_f values in the order of 10^{-4} for a coefficient of variation on P_f of 10%. Consequently, a method that can reduce the computation time of the probabilistic analysis (as compared to AK-MCS) by reducing the size of the sampling population is needed.

In order to overcome the shortcoming of AK-MCS involving the large population needed to assess the small failure probabilities, Echard et al. (2013) suggested the use of importance sampling instead of Monte Carlo sampling. The method by Echard et al. (2013) is called AK-IS.

It is an Active learning method combining Kriging and Importance Sampling. In the framework of this approach, the small failure probability can be estimated with a similar accuracy as AK-MCS but using a much smaller size of the population.

The AK-IS approach has been validated by Echard et al. (2013) by considering several academic examples where the performance function was given by an analytical equation. This method was shown to be very efficient as the obtained probability of failure is very accurate needing a smaller number of calls to the kriging meta-model as compared to AK-MCS methodology.

The aim of the present chapter is to extend the AK-IS approach by Echard et al. (2013) to the probabilistic analysis at the ultimate limit state of a strip footing resting on a spatially varying soil and subjected to a vertical load. The mechanical model considered in the analysis is the one used in the previous chapter. The same deterministic and uncertain parameters considered in the previous chapter are also conserved in this chapter for comparison purposes. As will be shown later on, the proposed AK-IS approach allows one to accurately compute the small failure probability (i.e. with a small value of the coefficient of variation on P_f) with a reduced computation time as compared to AK-MCS methodology.

The chapter is organized as follows: The next section aims at presenting the proposed combination between the kriging meta-modeling technique and the importance sampling (i.e. AK-IS procedure) in the case of geotechnical structures involving spatially varying soil properties. This is followed by a validation of the present AK-IS approach *via* a simple academic example. Then, some probabilistic results involving a strip footing resting on a spatially varying soil are presented and discussed. The chapter ends with a conclusion.

4.2 THE PROPOSED AK-IS PROCEDURE FOR GEOTECHNICAL STRUCTURES INVOLVING SPATIALLY VARYING SOIL

A brief description of the AK-IS approach as presented by Echard et al. (2013) in the case where the uncertain parameters are modeled by random variables is provided in Appendix E. As

previously mentioned in the introduction, this chapter aims at extending the AK-IS approach by Echard et al. (2013) to the case of a spatially varying soil where the computationally expensive mechanical model based on FLAC^{3D} software is used in the analysis.

The present AK-IS procedure consists of two main stages. First, the most probable failure point (design point) is determined using an approximate kriging meta-model based on a small number of samples called Design of Experiments DoE. Second, the obtained approximate kriging meta-model is successively improved *via* an enrichment process (by adding each time a new sample selected from a probability density function $h_x(X)$ centered at the design point) until reaching a sufficiently accurate meta-model for the computation of the failure probability. These two stages are described in more detail in the next subsections.

4.2.1 Determination of the design point

When dealing with problems that are characterized by an explicit performance function, the design point may be easily determined by minimizing the Hasofer-Lind reliability index subjected to the constraint that the performance function equal to zero [see Echard et al. (2013)]. Notice however that when dealing with analytically-unknown performance functions as is the case in the present work, the determination of the design point is less straightforward. The problem is even more difficult when a high-dimensional stochastic problem is involved [cf. Tandjiria et al. (2000)] as is the case in this chapter where a spatially varying soil mass is considered in the analysis. Indeed, the discretization of the two random fields of c and φ leads to a significant number of random variables (between 2 and 50 random variables) as was shown in the previous chapter. The large number of random variables requires a significant number of calls of the mechanical model for the determination of the design point.

In order to determine the design point in the present work using a relatively small number of calls to the mechanical model, an iterative procedure based on kriging metamodeling was proposed. This procedure may be described as follows (see also the flowchart presented in Figure 4.1):

1. Generate a large MCS population of N_{Mc} samples (say N_{Mc}=500,000 samples) of M standard Gaussian random variables $\left\{\left(\xi_1^1,...,\xi_M^1\right),\left(\xi_1^2,...,\xi_M^2\right),...,\left(\xi_1^{N_{Mc}},...,\xi_M^{N_{Mc}}\right)\right\}$ where M is the number of random variables adopted in EOLE methodology for the discretization of both c and φ. It should be emphasized here that each sample of M standard Gaussian random variables provides (when substituted into Equations 3.2 and 3.3) typical spatial variations of c and φ that respect the correlation structure of these fields, i.e. the so-called 'realizations' of c and φ. The difference between the different realizations lies in the position of the weak and strong soil zones although all realizations respect the correlation structure of the corresponding random fields.
2. From this population, randomly select a small number of samples (say N_1=20 samples) of M standard Gaussian random variables. Then, use EOLE methodology to transform each sample into realizations of c and φ that provide the spatial distribution of the soil cohesion and angle of internal friction respectively. These realizations are obtained through the computation of the values of c and φ at the centroids of the different elements of the FLAC^{3D} mesh using Equations (3.2 and 3.3**)**.
3. Use the software FLAC^{3D} to calculate the performance function value corresponding to each sample. Based on DACE toolbox, construct an initial approximate kriging meta-model in the standard space using the N_1 samples and the corresponding performance function values.
4. Find the minimum value of the Hasofer-Lind reliability index and the corresponding value of the design point by making use of the already-obtained kriging meta-model and by employing

the optimization toolbox available in Matlab. This procedure gives an approximate value of the reliability index and its corresponding design point.

5. Generate a small number of samples of *M* standard Gaussian random variables (5 samples are used in this work) so that they are centered at the design point obtained in the previous step. Then, transform each sample into realizations of c and φ that provide the spatial distribution of the soil cohesion and angle of internal friction respectively and finally compute for the five samples, the corresponding values of the performance function using $FLAC^{3D}$.
6. Construct a new kriging meta-model in the standard space using all samples of standard Gaussian random variables generated so far (i.e from steps (2) to step (5)).
7. Compute an updated Hasofer-Lind reliability index and its corresponding design point using the obtained kriging meta-model.
8. Steps 5 to 7 are repeated several times until the absolute difference between two successive values of the Hasofer-Lind reliability index becomes smaller than a given tolerance. The number of iterations is denoted hereafter as N_2. Consequently, the DoE which is considered to represent the number of samples needed to obtain the design point is given by $DoE = N_1 + 5 \times N_2$.

The aforementioned procedure does not intend to accurately determine the performance function over the entire design space but it focuses on the computation of the design point using a relatively small number of evaluations of the computationally expensive mechanical model.

4.2.2 Enrichment process

Further improvement of the already-obtained kriging meta-model is achieved in this stage *via* an enrichment process according to the following steps (see also the flowchart presented in Figure 4.2):

1. Generate a population of N_{IS} samples (say N_{IS} =10,000 samples) of *M* standard Gaussian random variables according to a probability density function $h_x(X)$ centered at the obtained

design point. These samples are called candidate samples hereafter. Remember that M is the number of random variables needed by EOLE methodology to discretize the two random fields c and φ. Notice also that the population size N_{IS} is relatively small herein as compared to the one generated in the AK-MCS procedure where N_{Mc} =500,000 samples.

2. Use the DACE toolbox in order to compute (for the whole population containing the N_{IS} samples) both the kriging predictor values $\mu_{\hat{G}}$ and their corresponding kriging variance values $\sigma^2_{\hat{G}}$ using the obtained meta-model. From the obtained values of the kriging predictors $\mu_{\hat{G}}$, obtain an estimation of the probability of failure and the corresponding value of the coefficient of variation using Equations (2.16 to 2.18).
3. Identify the 'best' next candidate sample among the whole population for which one will compute the performance function value using FLAC^{3D}. This is performed by evaluating a learning function U for each sample in the population. The learning function U is given by:

$$U\left(X_i\right)=\frac{\left|\mu_{\hat{G}(X_i)}\right|}{\sigma_{\hat{G}(X_i)}} \qquad i = 1, \ldots, N_{IS} \tag{4.1}$$

The 'best' next sample is the one that has the smallest U value (the details on this criterion for the choice of the 'best' sample were given in the second chapter of this thesis).

4. If the obtained minimum value of U is smaller than 2, evaluate the performance function value based on FLAC^{3D} for this 'best' candidate and update the DoE by adding the new 'best' sample. Also, re-construct the kriging meta-model again with the updated DoE.
5. Repeat the steps 2 to 4 several times until the smallest U value becomes larger than 2 (the details on this stopping criterion were given in the second chapter of this thesis).

At this stage, the learning stops and the meta-model is considered sufficiently accurate for the computation of the failure probability. When the learning stops, one must compute the estimated

values of both the probability of failure P_f and its corresponding coefficient of variation $COV\left(P_f\right)$.

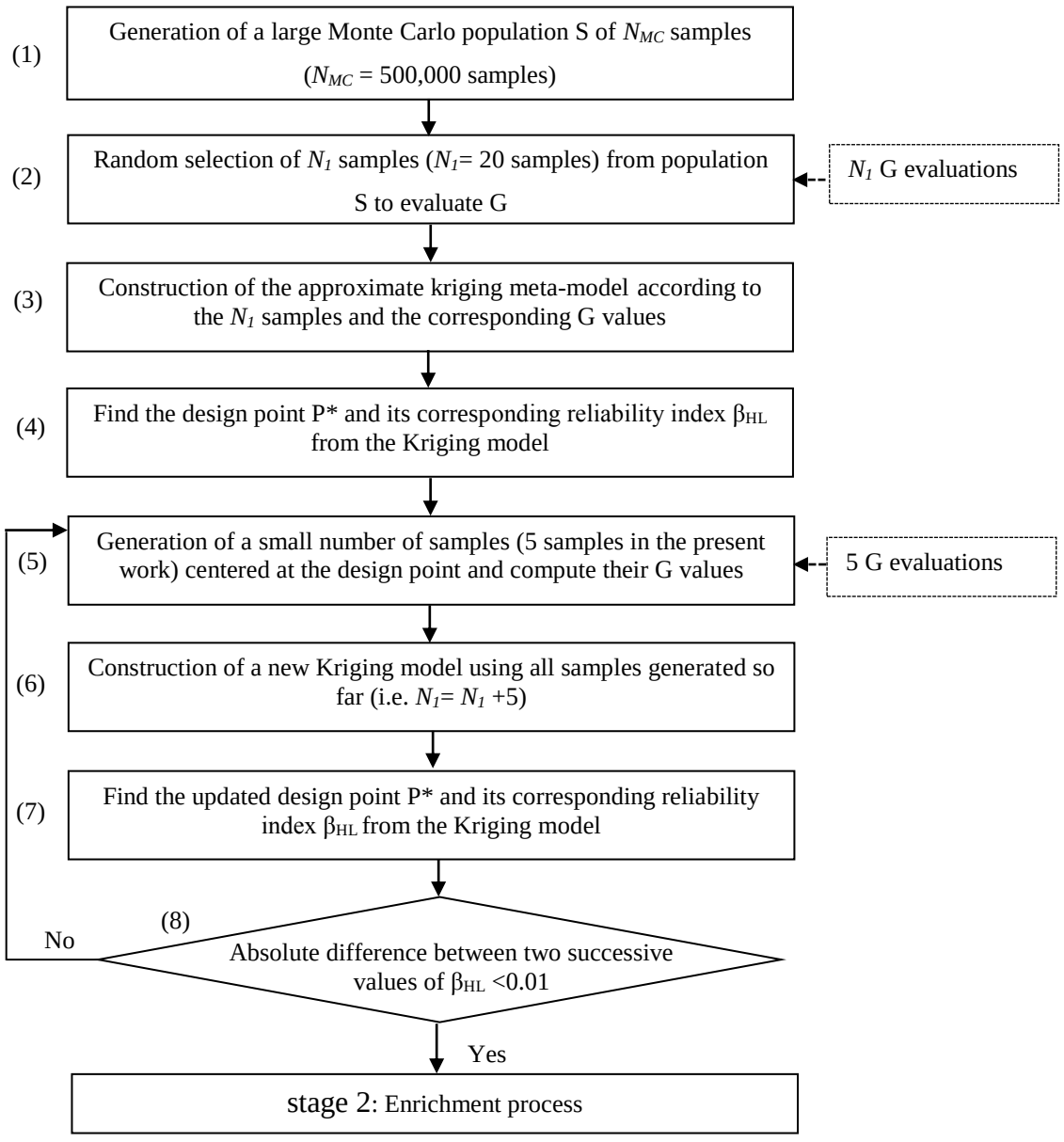


Figure 4.1. Flowchart of the proposed AK-IS procedure (Stage 1: Determination of the design point)

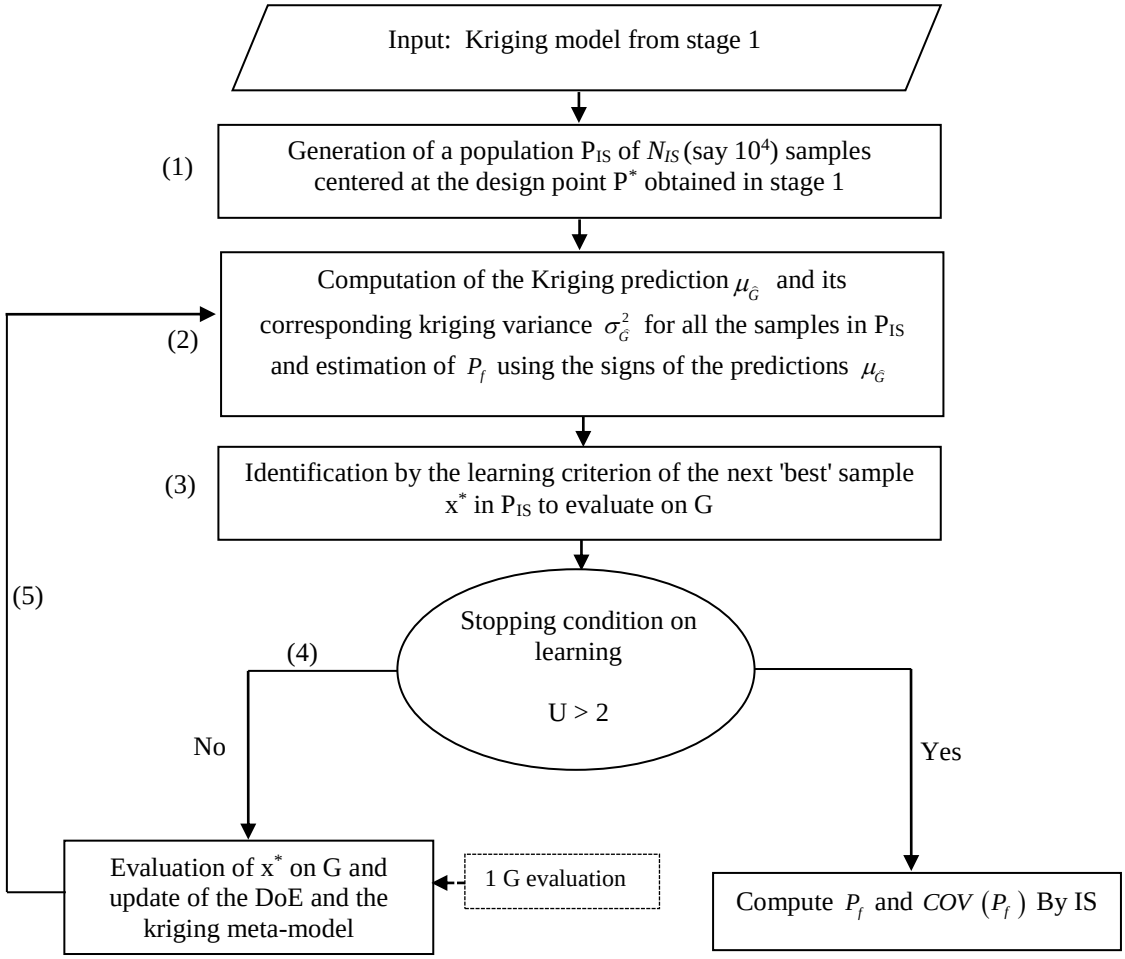


Figure 4.2. Flowchart of the proposed AK-IS procedure (Stage 2: Enrichment process)

4.2.3 Numerical implementation and computational issue

The comprehensive step-by-step procedure described above involves the probabilistic analysis at the ultimate limit state of a strip footing resting on a spatially varying soil (where c and φ are two random fields). However, it may be easily applied to other mechanical problems by substituting the present mechanical model with that required for the analysis.

This procedure was implemented in Matlab software. It includes the random field discretization by EOLE, the determination of the design point by an iterative procedure and the construction of a kriging meta-model for the computation of the failure probability. The implemented Matlab procedure makes several calls to the FLAC^{3D} code for the computation of the system response

(i.e. ultimate bearing capacity on a spatially varying soil) or the corresponding performance function value for the different soil realizations.

It should be emphasized herein that the number of predictions by kriging can be important as the whole importance sampling population (i.e. N_{IS} samples) is estimated. Notice however that the computation time of the predictions is significantly smaller than the one required by the AK-MCS approach presented in the previous chapter where a MCS sampling was employed to determine the candidate samples. Indeed, the number of candidate samples used in AK-IS is much smaller for the same value of the coefficient of variation on P_f thus leading to a considerable reduction in the computation time.

4.3 PROBABILISTIC NUMERICAL RESULTS

Before the presentation of the probabilistic results of a spatially varying soil, it seems necessary to validate the present AK-IS procedure by comparison of its results with those obtained by Echard et al. (2013) when considering a simple analytical equation. This is the aim of the next subsection.

4.3.1 Validation of the present AK-IS procedure via a simple analytical equation

This section focuses on the validation of the present AK-IS procedure through an analytical example. The corresponding performance function is given as follows:

$$G(u_1,u_2) = 0.5(u_1 - u_2)^2 - 1.5(u_2 - 5)^3 - 3 \tag{4.2}$$

where u_1 and u_2 are two standard normal random variables. A comparison between the results obtained by the present AK-IS procedure and those provided by Echard et al. (2013) was presented in Table 4.1.

Notice that in Echard et al. (2013), the design point was simply determined by minimizing the reliability index using the optimization toolbox in Matlab and making use of the analytical performance function. However; in the present AK-IS procedure, this design point is determined

by employing the iterative procedure proposed in the previous section. The aim is to check and validate the proposed iterative procedure which will be employed hereafter in the complex case of the spatially varying soil properties.

As may be seen from Table 4.1, the approximate kriging meta-model (which was needed for the determination of the design point) was constructed using an initial DoE of 15 samples and five iterations with 2 samples per iteration. The enrichment process required 4 additional samples. Thus, the total number of samples needed in our procedure is equal to 29 samples. This number is close to that needed by the classical FORM analysis by Echard et al. (2013) (i.e. 26 samples) with the advantage that the present approach may be applied to analytically-unknown performance functions.

As a conclusion, the iterative procedure proposed in this chapter for the computation of the design point can be considered as a powerful tool and may be used for more complex cases involving spatially varying soil properties.

Table 4.1. Probabilistic outputs and the corresponding number of used samples N_{call} as obtained from the two AK-IS methods

Method	N_{call}	$P_f \times 10^{-5}$	$COV\left(P_f\right)$ (%)	β_{HL}	Design point (u_1, u_2)
AK-IS by Echard et al. (2013)	19 (DoE) + 7 (enrichment)= 26 samples	2.86	2.39	3.93	(0.788, 3.853)
Present AK-IS approach	15 + (2 × 5) + 4 =29 samples	2.83	2.40	3.93	(0.786, 3.853)

4.3.2 Probabilistic results in the case of a spatially varying soil

This section aims at presenting the impact of the soil spatial variability on the failure probability against soil punching of a strip footing subjected to a vertical loading. The mechanical model considered in the analysis is the one used in the previous chapter and it was presented in section 3.5. The same deterministic and uncertain parameters described in the previous chapter are

considered in this section. Also, the input data employed in chapter 3 for both the deterministic and the uncertain parameters remain the same herein. The objective is to compare the results from the two approaches in terms of computational effort. The performance function used in the analysis is given by Equation (3.5) of the previous chapter. Notice finally that the number of samples N_{IS} used in all subsequent computations was equal to 10,000 samples. The small size of the sampling population may be explained by the fact that the sampling is performed using a probability density that is centered at the design point leading to a much larger number of samples lying in the failure domain as compared to MCS methodology.

4.3.2.1 Evolution of the limit state surface during the computational process

As was previously mentioned in this chapter, the AK-IS procedure consists of two main stages: The first stage (called stage 1) consists in computing the design point from an approximate kriging meta-model constructed using a small number of samples. In the second stage (called stage 2), the approximate meta-model is successively improved through an enrichment process. In this section, the evolution of the limit state surface with the addition of new samples (or realizations) during the two stages (i.e. stage 1 and stage 2) was investigated (see Figures 4.3 and 4.4). A typical case where a_x=10,000m and a_y=10,000m was considered in these figures. This configuration was chosen because it requires only two random variables and thus, the limit state surface can be easily visualized since only a two-dimensional space is needed in this case.

Figure 4.3 presents the evolution of the limit state surface with the addition of new samples during the different iterations of stage 1. Also, Table 4.2 presents the evolution of the reliability index β_{HL} for the different iterations. This table shows that the accurate value of the reliability index was obtained since the first iteration in the present case of a quasi-homogeneous soil where a_x=a_y=10,000m. Notice however that a greater number of iterations (between 2 and 13) were

found necessary for moderately to highly heterogeneous soil mediums as may be seen from the sixth column of Table 4.4.

It should be emphasized here that the addition of new samples during Stage 1 does not intend to accurately determine the limit state surface over the entire range of random variables, but rather to accurately determine the reliability index and the corresponding design point. Thus, only the point of the limit state surface which is the closest one to the origin of the limit state surface is expected to be correct at the end of the first stage of AK-IS; the other points of this limit state being in general not correctly estimated within this stage.

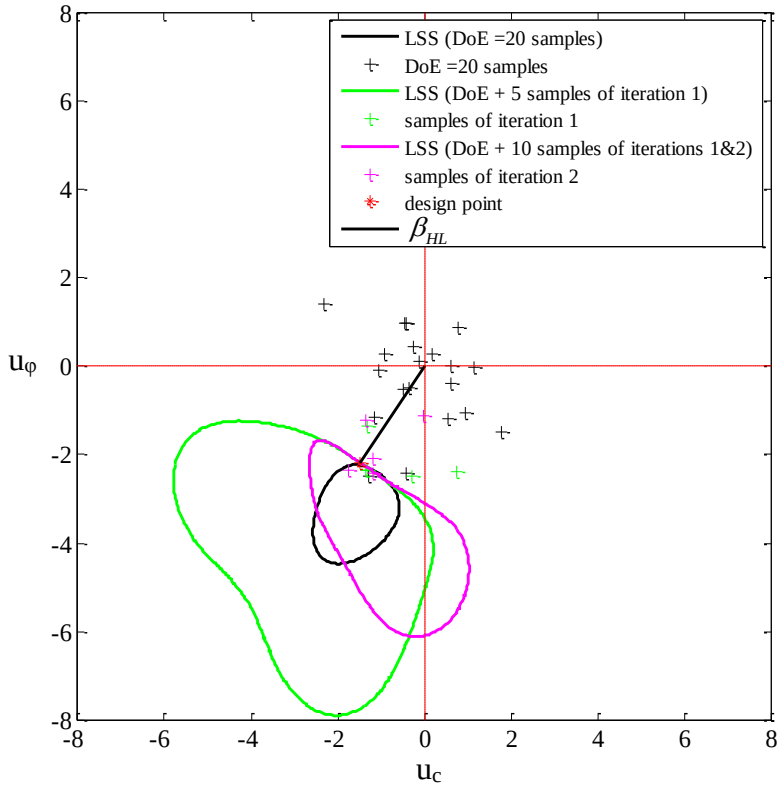


Figure 4.3. Effect of the number of iterations of stage 1 on the limit state surface when $a_x = a_y$ = 10,000 m

Table 4.2. The evolution of the reliability index for the different iterations

Case	β_{HL}	δ
Initial DoE = 20 samples	2.6205	-
Initial DoE + 5 samples of iteration 1 = 25 samples	2.6345	0.014
Initial DoE + 5 samples of iteration 1 + 5 samples of iteration 2 = 30 samples	2.6340	0.0005

Figure 4.4 presents the evolution of the limit state surface with the addition of new samples (from zero to 28 samples) during stage 2; the number 28 being the needed number of added realizations. From this figure, one may notice that the limit state surface is successively improved with the addition of new samples. Notice however that for the last two iterations, the two curves representing the limit state surface are superimposed. Thus, the limit state surface cannot be further improved beyond 24 samples. This means that there is no bias in the meta-model beyond 24 samples.

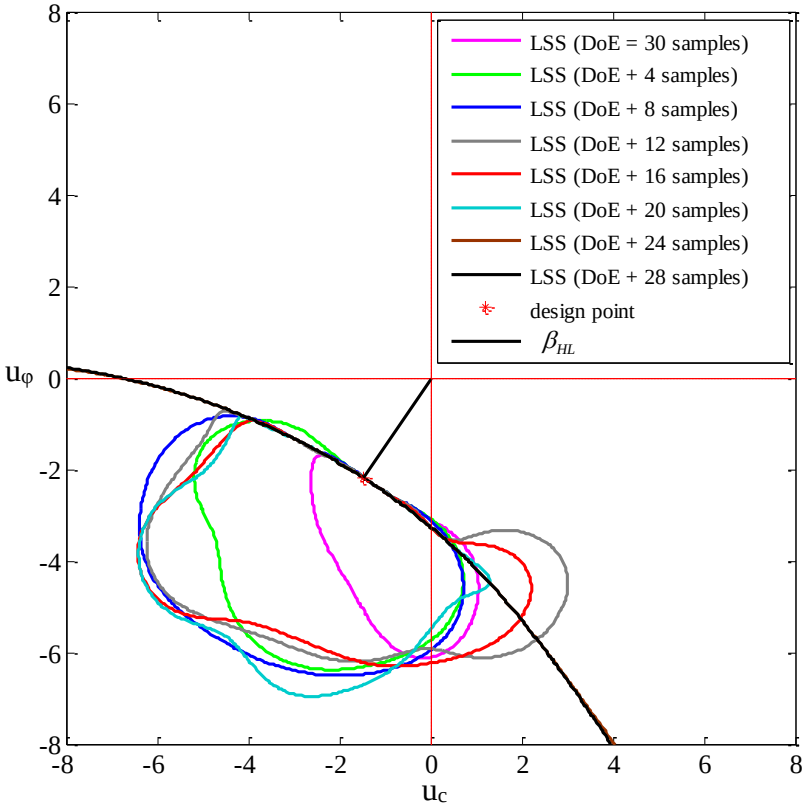


Figure 4.4. Effect of the number of added samples during the enrichment process on the limit state surface when a_x=a_y=10,000 m

Table 4.3 presents the evolution of the probability of failure P_f and its corresponding coefficient of variation $COV\left(P_f\right)$ as function of the added samples. This Table shows that the values of P_f and $COV\left(P_f\right)$ converge after the addition of 24 samples. This is in conformity with Figure 4.4 in which no further improvement in the limit state function was obtained in the last two iterations.

Table 4.3. The evolution of the probability of failure and its corresponding coefficient of variation as a function of the added samples during the enrichment process

Number of added samples	$P_f \times 10^{-3}$	$COV\left(P_f\right)\%$
0	3.606	1.957
4	4.589	1.736
8	4.171	1.763
12	4.111	1.878
16	3.919	1.772
20	3.828	1.772
24	3.830	1.771
28	3.830	1.771

4.3.2.2 Evolution of the probabilistic outputs during the enrichment process

First of all, recall here that the failure probability is computed each time a new sample is added during the enrichment process. Figure 4.5 shows the effect of the number of added samples in the enrichment process on P_f and $COV\left(P_f\right)$ values for a typical case where a_x=10m and a_y=1m [see also Al-Bittar et al. (2017)]. This figure also provides the learning function values for the different added samples. The configuration (a_x=10m and a_y=1m) was studied because it represents a practical case requiring a significant number of random variables (32 random variables in the present case).

Figure 4.5 shows that for the small number of added samples, both P_f and $COV\left(P_f\right)$ are not constant. This is due to the inaccuracy of the kriging meta-model when only a small number of realizations were considered. Notice however that both P_f and $COV\left(P_f\right)$ tend to be constant as the number of added samples increases. It should be mentioned here that 921 samples were

needed in the enrichment process in addition to the DoE before the algorithm stops [min(U)]>2. The final obtained values of P_f and $COV\left(P_f\right)$ are respectively 1.628×10^{-3} and 2.99%. As may be shown from Figure 4.5 the values of P_f and $COV\left(P_f\right)$ reach an asymptote when the number of added samples is equal to 823. An additional increase in the number of added samples does not lead to a significant change in the values of P_f and $COV\left(P_f\right)$. This means that when the number of added samples becomes equal to 823, the kriging meta-model is accurate enough (i.e. with no bias) and it can be used to calculate a rigorous value of the failure probability.

In order to show that no bias in the meta-model exists at the end of the enrichment process for all the configurations presented in Table 4.4, Appendix F presents the plots of P_f for the different values of the autocorrelation distances considered in this table. For completeness, this appendix also provides the plots of $COV\left(P_f\right)$ and U for these configurations.

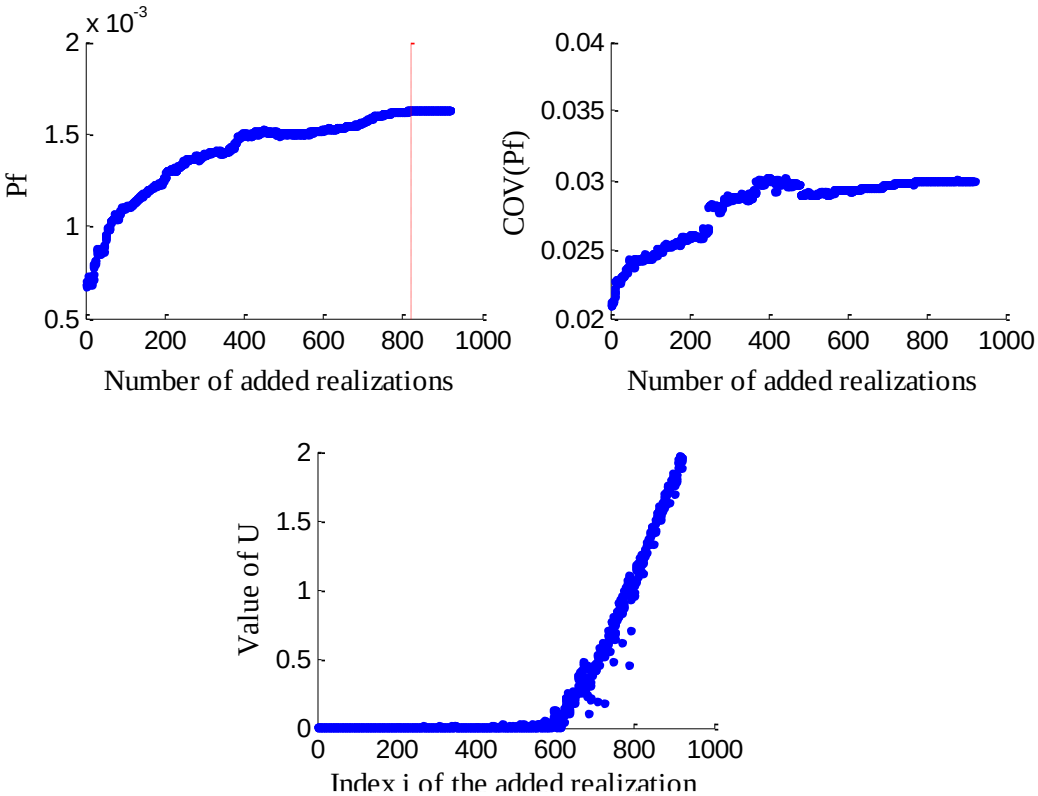


Figure 4.5. AK-IS results for a spatially varying soil (a_x=10 m, a_y=1 m)

4.3.2.3 Parametric study

This section aims at presenting the effect of (i) the autocorrelation distances of the random fields, (ii) the coefficients of variation of these fields, and (iii) the type of the probability distribution function of these fields, on the probabilistic outputs.

4.3.2.3.1. Effect of the autocorrelation distances of the random fields

Figure 4.6 presents the effect of the isotropic autocorrelation distance on P_f and β_{HL} as obtained from AK-MCS and AK-IS methodologies. Also, Figure 4.7 and Figure 4.8 present the effect of the autocorrelation distance (a_y or a_x) on P_f and β_{HL} as obtained from the same two methodologies. The values of the probabilistic outputs corresponding to the different soil variabilities were given in Table 4.4.

As may be easily seen from Table 4.4, the N_1 value was taken equal to 20 (as mentioned in the flowchart of Figure 4.1) for all configurations except for two configurations corresponding to a high number of random variables (or eigenmodes). Indeed, these two configurations correspond to very heterogeneous soils and require a greater number of samples for the construction of the meta-model. In this chapter, the N_1 value adopted for these configurations was arbitrarily taken equal to the number of eigenmodes (i.e. 48 and 44 samples respectively).

From Figures 4.6, 4.7 and 4.8, one may observe that the two methods lead to similar results. Except for the very heterogeneous case where a_x=10m and a_y=0.5m, the maximal percent difference between the two approaches for all the other configurations is smaller than 7 %.

It should be emphasized here that although the sampling population adopted in AK-IS (i.e. 10,000 samples) is much smaller than that adopted in AK-MCS (i.e. 500 000 samples), the coefficient of variation on the failure probability obtained from AK-IS is small (smaller than 7% for most configurations) and it is quite close to that obtained by AK-MCS (see Figures 4.9, 4.10 and 4.11) which means that both methods lead to accurate results.

The AK-IS procedure used herein is more efficient than AK-MCS. It provides an accurate value of the failure probability (i.e. with small values of the coefficient of variation on the failure probability) using a much smaller number of calls to the meta-model as compared to AK-MCS (10,000 calls instead of 500,000 calls). This significantly reduces the computation time. For instance; when considering the typical case where a_x=10m and a_y=2m, 12 days (in average) were necessary to complete the AK-MCS computation, whereas only 3 days were needed in average to perform a complete calculation using the AK-IS method.

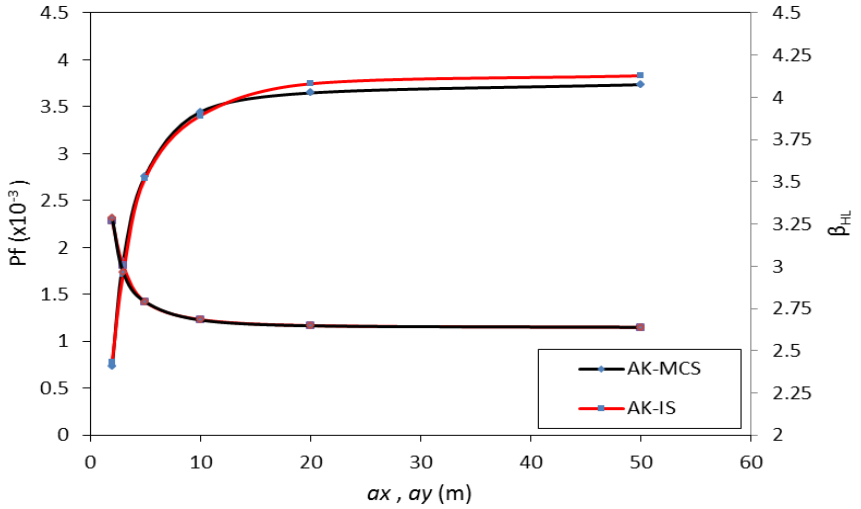


Figure 4.6. Effect of the isotropic autocorrelation distance a_x=a_y on P_f and β_{HL}

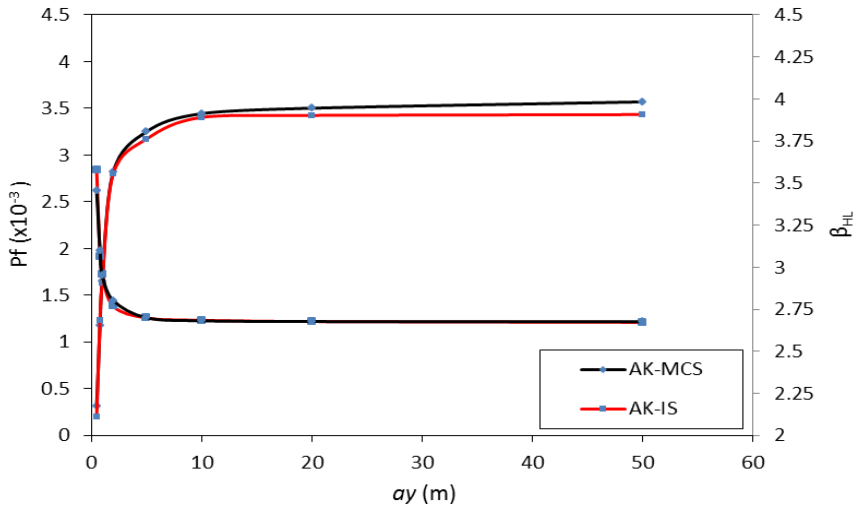


Figure 4.7. Effect of the vertical autocorrelation distance a_y on P_f and β_{HL} when a_x =10 m

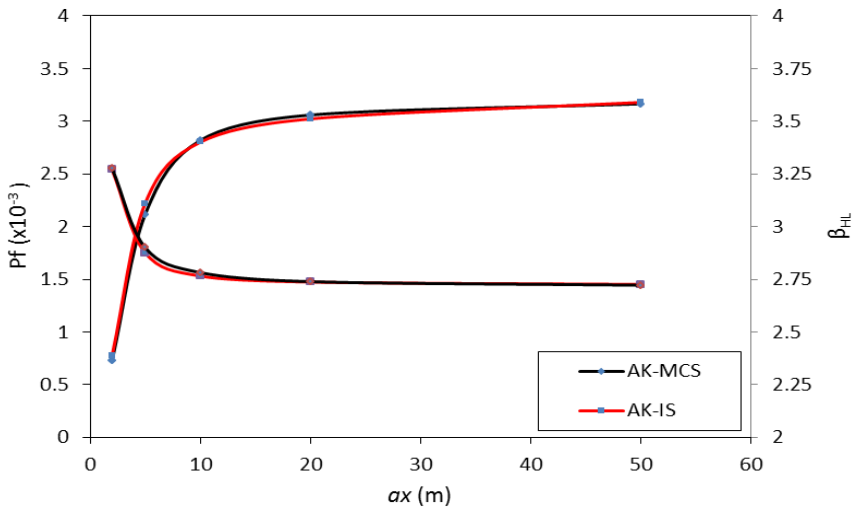


Figure 4.8. Effect of the horizontal autocorrelation distance a_x on P_f and β_{HL} when a_y=2 m

Table 4.4. Adopted number of random variables and the corresponding value of the variance of error of EOLE together with the values of P_f , $COV\left(P_f\right)$, size of DoE and number of added realizations for various soil variabilities

a. Case of an isotropic case (a_x=a_y)

a_x=a_y (m)	No. of random variables	Variance of the error %	P_f × 10^{-3}	$COV\left(P_f\right)$ %	Size of DoE = N_1 +5× N_2	No. of added realizations
2	48	9.447	0.777	6.69	48+5×3	1363
3	32	4.647	1.718	5.38	20+5×6	1076
5	24	0.953	2.738	2.42	20+5×10	812
10	10	0.815	3.404	1.91	20+5×10	243
20	8	0.170	3.745	1.91	20+5×3	200
50	6	0.016	3.831	1.82	20+5×6	74
100	6	0.001	3.933	1.82	20+5×6	90

b. Case of an anisotropic case (a_x=10 m with varying a_y)

a_y (m)	No. of random variables	Variance of the error %	P_f × 10^{-3}	$COV\left(P_f\right)$ %	Size of DoE = N_1 +5× N_2	No. of added realizations
0.5	44	9.119	0.199	15.34	44+5×2	1237
0.8	38	4.798	1.234	3.51	20+5×8	1192
1	32	4.212	1.628	2.99	20+5×8	921
2	24	1.437	2.755	2.68	20+5×5	644
5	12	1.682	3.172	2.06	20+5×8	354
10	10	0.815	3.404	1.91	20+5×10	243
20	8	0.855	3.425	1.98	20+5×5	228
50	8	0.297	3.434	1.99	20+5×4	210
100	8	0.099	3.595	1.78	20+5×10	194

c. Case of an anisotropic case (a_y=2 m with varying a_x)

a_x (m)	No. of random variables	Variance of the error (%)	P_f × 10^{-3}	$COV\left(P_f\right)$ %	Size of DoE = N_1 +5× N_2	No. of added realizations
2	48	9.447	0.777	6.69	48+5×3	1363
5	30	4.101	2.221	2.65	20+5×13	988
10	24	1.437	2.755	2.68	20+5×5	644
20	16	1.415	3.023	2.07	20+5×9	437
50	12	1.272	3.180	1.87	20+5×12	313
100	10	0.842	3.191	1.82	20+5×12	244

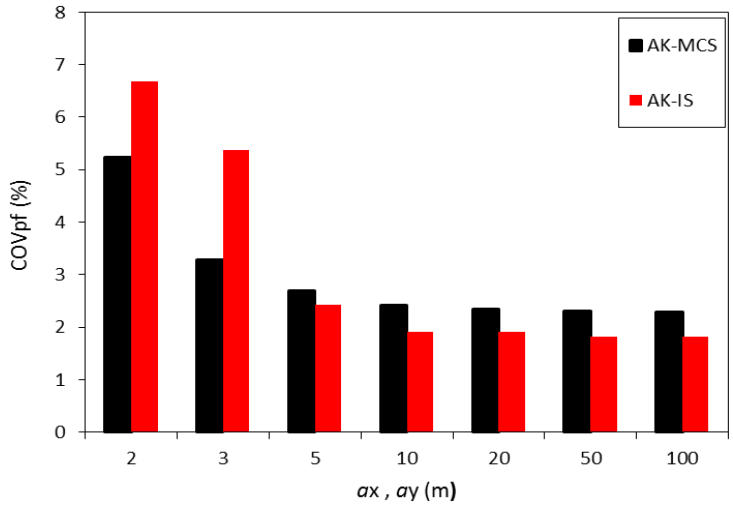


Figure 4.9. Values of $COV\left(P_f\right)$ for different values of the isotropic autocorrelation distance a_x=a_y

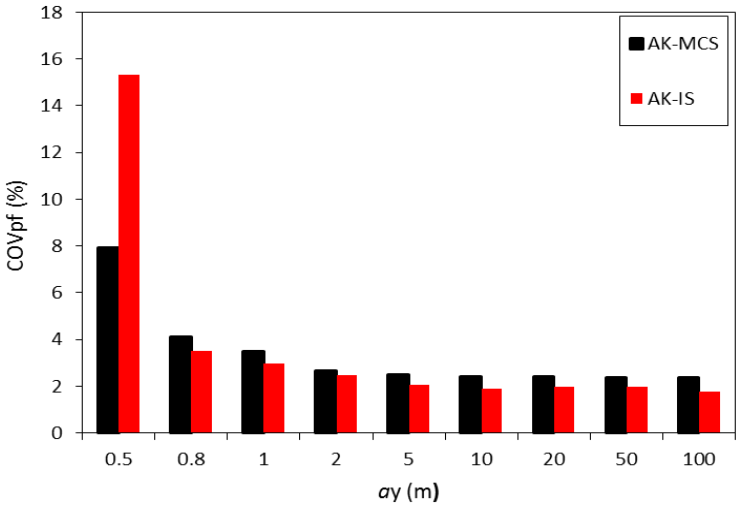


Figure 4.10. Values of $COV\left(P_f\right)$ for different values of the vertical autocorrelation distance a_y when a_x=10 m

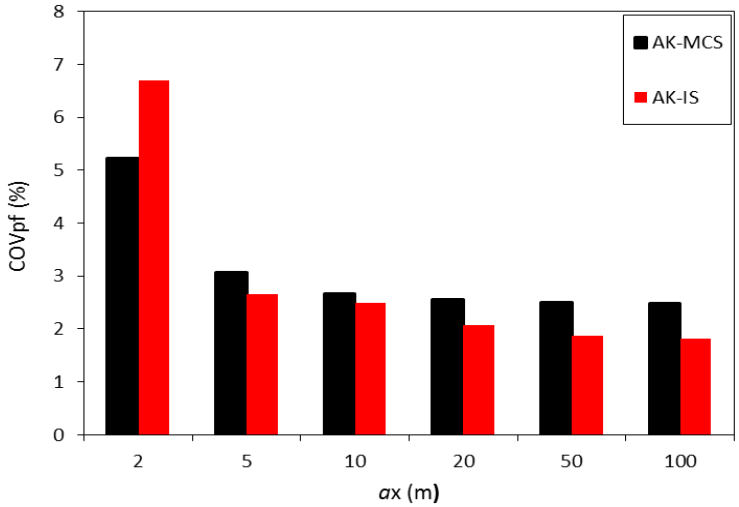
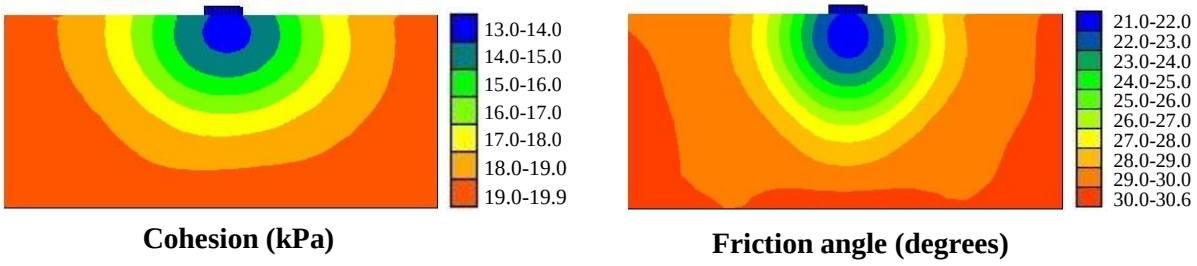


Figure 4.11. Values of $COV\left(P_f\right)$ for different values of the horizontal autocorrelation distance a_x when a_y=2 m

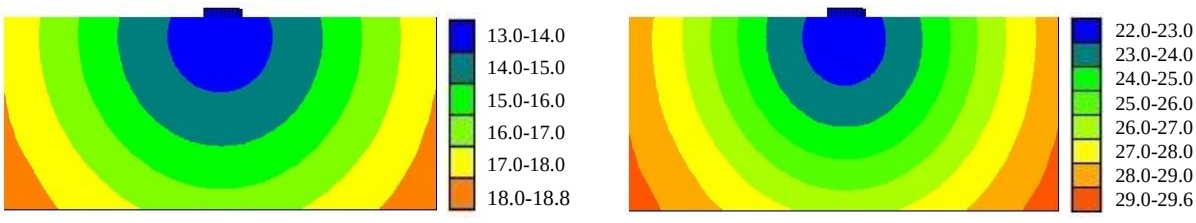
Figure 4.12 shows the critical realizations of the random fields at the design point (i.e. the most probable failure point) for two values of the isotropic autocorrelation distance. This figure shows, as in chapter 3, that a symmetrical distribution of the soil shear strength parameters was obtained. The weaker soil zone is concentrated around the foundation, the stronger soil being far from the foundation. As may be observed from Figure 4.12, the size of the weaker soil zone under the footing increases with an increase in the autocorrelation length leading to a greater failure probability. The weak soil zone under the foundation allows the failure mechanism to develop through this zone reflecting the most prone soil to punching.

Figure 4.13 shows the critical realizations of the random fields at the design point for non-isotropic soils. Two values of a_y (with a_x=10m) were considered in the analysis. As in the isotropic case, the critical realizations at the design point exhibit a symmetrical pattern. From Figure 4.13, one may observe an increase in the depth of the weaker soil zone with the increase in the vertical autocorrelation distance leading naturally to a greater failure probability.



Cohesion (kPa) **Friction angle (degrees)**

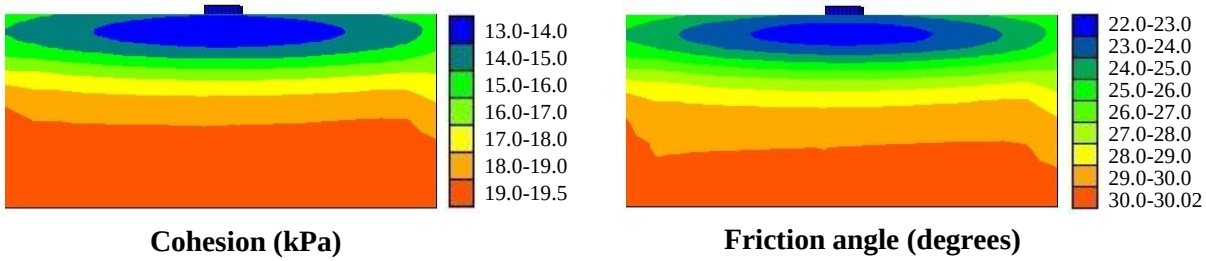
(a) $a_x=a_y$**=2m**



Cohesion (kPa) **Friction angle (degrees)**

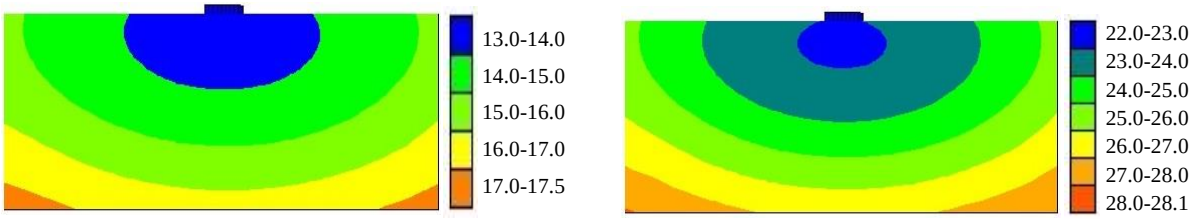
(b) $a_x=a_y$**=5m**

Figure 4.12. Critical realizations for two values of the isotropic autocorrelation distance



Cohesion (kPa) **Friction angle (degrees)**

(a) **a_x=10m and a_y=1m**



Cohesion (kPa) **Friction angle (degrees)**

(b) **a_x=10m and a_y=5m**

Figure 4.13. Critical realizations for two values of a_y when a_x=10m

4.3.2.3.2 Effect of the coefficients of variation of the random fields

Table 4.5 presents the impact of the coefficients of variation of the random fields (COV_c *and* COV_φ) on the values of P_f and $COV\left(P_f\right)$. As expected, this table shows that P_f increases with the increase in the coefficient of variation of a random field. From this table, it can be observed that the effect of COV_φ is significant compared to that of COV_c. For instance, increasing COV_c by 50% of its reference value (COV_c = 25%) increases P_f by 228.17 %. However, an increase in COV_φ by 50 % of its reference value (COV_φ = 10%) increases P_f by 741.34 %.

Table 4.5. Effect of the coefficients of variation of the random fields *c* and φ on P_f ,$COV\ (P_f)$, the size of the DoE and the number of added realizations in case of anisotropic random fields where a_x=10m and a_y=2m

COVc %	*COV* φ %	$P_f \times 10^{-3}$	$COV(P_f)$ %	Size of DoE = $N_1 + 5 \times N_2$	Number of added realizations
12.5	10	0.708	6.61	20+5×3	678
25	10	2.755	2.68	20+5×5	644
37.5	10	9.041	1.94	20+5×8	637
25	5	0.0129	2.83	20+5×10	547
25	15	23.179	2.09	20+5×5	885

4.3.2.3.3 Effect of the type of the probability density functions (PDFs) of the random fields

The effect of the type of PDF of the random fields is presented in Table 4.6. Two cases were considered (normal and non-normal random fields). For the case of normal random fields, both the soil cohesion and friction angle were considered to follow a normal distribution. However, for the case of non-normal random fields, the soil cohesion was considered to follow a lognormal distribution while the internal friction angle was considered to follow a beta distribution. Table 4.6 shows that the P_f value corresponding to the normal random fields is larger than that corresponding to the non-normal random fields. This may be explained by the fact that, in the case of non-normal random fields, the limited range of variation of the soil shear strength parameters may led to quasi similar responses for the different realization and thus, to a smaller variability of the system response. The opposite occurs in the case where the random fields follow a normal distribution. As a result, a higher value of the failure probability was obtained in case of normal random fields.

Table 4.6. Effect of the type of the probability density function of the random fields *c* and φ on P_f , $COV\ (P_f)$, the size of the DoE and the number of added realizations in case of anisotropic random fields where a_x=10 m and a_y=2 m

Type of the probability density function	$P_f \times 10^{-3}$	$COV(P_f)$%	Size of DoE = $N_1 + 5 \times N_2$	Number of added realizations
Normal fields	5.042	2.53	20+5×7	756
Non-normal fields	2.755	2.68	20+5×5	643

4.4 CONCLUSION

This chapter was devoted to the AK-IS method which is an active learning reliability procedure combining kriging and importance sampling. Within this approach, the shortcoming of AK-MCS involving the large population needed to assess the small failure probabilities was addressed by providing a more efficient sampling technique. Indeed, importance sampling was used instead of Monte Carlo sampling technique. Furthermore, the AK-IS procedure makes use of the advantages of both kriging (by using the prediction mean and prediction variance for the determination of the 'best' new candidate point) and importance sampling (for the generation of samples about the most probable point).

AK-IS procedure consists of two main stages. First, 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 samples. Second, the obtained approximate kriging meta-model is successively improved *via* an enrichment process by adding each time a new sample selected from a probability density function centered at the design point. This dramatically reduces the size of the population needed to assess the small failure probability *via* the meta-model and thus, this significantly reduces the computation time with respect to AK-MCS approach.

After the presentation of the probabilistic AK-IS approach, this chapter presented (as in chapter 3) a probabilistic analysis at the ultimate limit state of a strip footing resting on a spatially varying soil using the same deterministic and uncertain parameters. The main findings of this study can be summarized as follows:

1. The iterative procedure proposed in this chapter for the computation of the design point was shown to be a powerful tool and it has been used successfully in the complex cases involving very heterogeneous soil media.

2. Although the sampling population adopted in AK-IS (i.e. 10,000 samples) is much smaller than that adopted in AK-MCS (i.e. 500,000 samples), the coefficient of variation on the failure probability obtained from AK-IS is small (smaller than 7% for most configurations) and it is quite close to that obtained by AK-MCS which means that both methods lead to accurate results.
3. The comparison between the results of the failure probability obtained using the AK-MCS and the AK-IS methodologies has shown good agreement. Except for the very heterogeneous case where a_x=10m and a_y=0.5m, the maximal percent difference between the two approaches for all the other configurations is smaller than 7 %.
4. The present AK-IS procedure was shown to be more efficient than AK-MCS. It provides an accurate value of the failure probability (i.e. with a small value of the coefficient of variation on this failure probability) using a much smaller number of calls to the meta-model as compared to AK-MCS (10,000 calls instead of 500,000 calls). This significantly reduces the computation time. For instance; when considering the typical case where a_x=10m and a_y=2m, 12 days (in average) were necessary to complete the AK-MCS computation, whereas only 3 days were needed in average to perform a complete calculation using the AK-IS method.
5. 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. The weak soil zone increases with an increase the autocorrelation length leading to a greater failure probability.
6. The study of the effect of the coefficient of variation of a random field on the failure probability has shown as expected that the increase in the coefficient of variation of a random field increases the P_f value, the coefficient of variation of the internal friction angle being of more impact on the failure probability than that of the soil cohesion.

7. The P_f value corresponding to the normal random fields was found to be larger than that corresponding to non-normal random fields. This was explained by the limited range of variation of the shear strength parameters within a given realization in case of non-normal random fields.

CHAPTER 5. PROBABILISTIC ANALYSIS OF STRIP FOOTINGS RESTING ON SPATIALLY VARYING SOILS USING KRIGING META-MODELLING AND SUBSET SIMULATION

5.1 INTRODUCTION

In the previous chapter, the shortcoming of AK-MCS procedure involving the large population needed to assess the small failure probabilities was addressed by providing a more efficient sampling technique. Indeed, importance sampling was used instead of the Monte Carlo sampling technique. As a result, a combined use of kriging metamodeling and importance sampling (called AK-IS) was proposed. In the framework of this approach, the small failure probability can be estimated with a similar accuracy as AK-MCS but using a much smaller size of the initial population. Notice that the AK-IS procedure consists of two main stages. First, the most probable failure point (design point) is determined using an approximate kriging meta-model 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 sample selected from a probability density function centered at the design point. This dramatically reduces the size of the population needed to assess the small failure probability.

Since the AK-IS method requires the computation of the design point, it may not be suitable when dealing with limit state surfaces that possess more than one design point. Consequently, a method that overcomes such shortcoming is needed. In this chapter, a more efficient method was proposed. Within this approach, a combined use of kriging metamodeling and subset simulation (called AK-SS) was employed. This method makes use of a more efficient sampling technique than AK-IS and thus, it is more suitable for problems involving several design points.

The aim of this chapter is to describe the present AK-SS procedure and to perform a probabilistic analysis at the ultimate limit state of a strip footing resting on a spatially varying soil and subjected to a vertical load. As in chapter 4, the mechanical model and the corresponding input

data considered in the analysis are the ones used in chapter 3. Indeed, the ultimate aim of this thesis is to compare the results from the three approaches AK-MCS, AK-IS and AK-SS in terms of computation time.

The chapter is organized as follows: The next section aims at presenting the proposed combination between the kriging meta-modeling technique and the subset simulation (i.e. AK-SS procedure). This is followed by a validation of the AK-SS approach *via* an academic example involving several design points. Then, some probabilistic results involving a strip footing resting on a spatially varying soil (although a single design point exists in this case) are presented and discussed. The chapter ends with a conclusion of the main findings.

5.2 PROPOSED AK-SS PROCEDURE

In this section, we propose a technique that combines the kriging meta-modeling and the subset simulation approach (i.e. the AK-SS procedure). The present AK-SS procedure consists of two main steps.

In the first step, an approximate kriging meta-model is constructed using a small number of samples called initial Design of Experiments (DoE). These samples are randomly selected from a large population generated by MCS. In the second step, the subset simulation approach is used to generate samples which are directed towards the limit state surface by employing the obtained approximate kriging meta-model. Then, the approximate kriging meta-model is improved by adding a ‘best’ new sample to the initial DoE. The new sample is selected among the samples obtained in the final level of the subset simulation and making use of a powerful learning function. This learning function allows one to select the sample that is the closest one to the limit state surface G=0. The process of adding a new sample (process of enrichment) is repeated until a prescribed criterion on the value of the failure probability is obtained. At the end of the enrichment process, the number of added samples is considered sufficient to compute the final

value of the failure probability P_f and the corresponding value of the coefficient of variation $COV\left(P_f\right)$ making use of the classical subset simulation approach on the final meta-model.

It should be emphasized here that the initial approximate kriging meta-model constructed based on the initial DoE should be sufficiently accurate within AK-SS, since it is used by subset simulation to locate the points that are close to the limit state surface for the selection of the best one to be computed by the mechanical model. For this reason, the DoE was selected from a large Mote Carlo population of 1,000,000 simulations. Furthermore the size of this DoE was arbitrarily taken equal to 20 samples for all the configurations corresponding to a number of eigenmodes smaller than or equal to 20. However, this DoE was arbitrarily taken equal to the number of eigenmodes for the other configurations corresponding to a very heterogeneous soil. These suggestions were found suitable to build an initial kriging meta-model that was easily improved by the added realizations.

A step by step description of the AK-SS procedure is given as follows (see also the flowchart presented in Figure 5.1):

1. Generate by Monte Carlo Simulation (MCS) a large population composed of N_{Mc} samples (say 1000,000 samples) where each sample is composed of M standard normal random variables, M being the number of eigenmodes required by EOLE methodology for the discretization of both *c* and φ.

2. Randomly select from the N_{Mc} samples generated in step 1, a small number of N_1 samples (N_1=20 samples at least as was explained before). The N_1 samples are called hereafter the initial Design of Experiment (DoE).

3. For each selected sample of the initial DoE, obtain realizations of the two random fields using the EOLE method. Then, calculate the system response value corresponding to each sample based on FLAC^{3D} software.

4. Construct an approximate kriging meta-model using the samples of the initial DoE and the corresponding values of the system response. It should be mentioned here that the kriging meta-model is constructed using the DACE toolbox in Matlab software.

5. Apply the subset simulation approach to the approximate kriging meta-model obtained in the previous step. This provides a number of samples (vectors of standard normal random variables) which are directed towards the limit state surface. Notice that in the SS approach, the space of uncertain parameters is divided into a number *m* of levels with equal number N_{SS} of samples. The first N_{SS} samples are generated using MCS methodology and the next samples of each subsequent level are obtained using a Markov chain method based on Metropolis-Hastings (M-H) algorithm. It should be emphasized here that the number of samples N_{SS} to be used per level of the SS approach should be sufficient to accurately calculate the P_f value. Therefore, N_{SS} =100,000 samples will be considered in the subsequent probabilistic calculations. Notice finally that the samples of the final level of subset simulation are expected to be close to the limit state surface G=0. Therefore, they represent candidate samples to improve the kriging meta-model.

6. For the samples of the final level of subset simulation, identify the ‘best’ sample to be used in the enrichment processes to improve the approximate kriging meta-model. Notice that the 'best' sample is the one that has the smallest value of the learning function U. This learning function is given by the following equation:

$$U(X_i) = \frac{\left|\mu_{\hat{G}(X_i)}\right|}{\sigma_{\hat{G}(X_i)}} \qquad i = 1, \ldots, N_{SS} \tag{5.1}$$

7. If the stopping condition by Schöbi et al. (2015) is not satisfied, evaluate the performance function value based on FLAC^{3D} software for this best candidate. Then, add this sample to the initial DoE and reconstruct the kriging meta-model using the updated DoE to obtain an updated kriging meta-model. Steps (5), (6) and (7) are repeated until the process of adding a new sample stops (i.e. until the stopping criterion by Schöbi et al. (2015) is satisfied).

8. If the stopping condition is satisfied, the learning stops and the final kriging meta-model becomes sufficiently accurate for the computation of the failure probability. Finally, the subset simulation approach is applied on the obtained final meta-model to calculate the final value of the failure probability P_f (SS) and the coefficient of variation $COV\left(P_f\right)$(SS). It should be mentioned that the intermediate failure probability P_0 of a given level j (j=1, 2, …, m) was chosen equal to 0.1 in this chapter.

5.2.1 Sample selection and stopping condition

In the two previous chapters, the stopping condition for both AK-MCS and AK-IS procedures was based on the accuracy of the meta-model around the limit state surface rather than on the estimation of the statistics of interest. Thus, the 'best' chosen sample was the one that mostly improves the meta-model.

In this chapter, use is made of a recent stopping criterion by Schöbi et al. (2015) for the estimation of failure probabilities. The aim of this criterion is to maximize the accuracy of the statistics of interest while minimizing the computational costs rather than improving the accuracy of the meta-model. This stopping condition is very interesting in the present AK-SS procedure since the majority of the samples are very close to the limit state function, the use of the stopping

criterion min (U)>2 would add a large number of samples that will not necessarily lead to a significant improvement in the estimated values of P_f .

According to Schöbi et al. (2015), the stability of the estimate of the statistics of interest can be measured by the size of the limit state margin and consequently by the associated values of the upper and lower boundaries of the limit state surface. When the boundaries are close to each other (i.e. a small limit state margin), the estimate of the statistics of interest is accurate. Therefore, the stopping criterion for estimating the failure probability P_f is given as follows:

$$\frac{P_f^+ - P_f^-}{P_f^0} \leq \varepsilon_{P_f} \tag{5.2}$$

where P_f^0 is the original failure probability based on the kriging predictor values $P(\mu_{\hat{G}}(x) \leq 0)$ and P_f^+ and P_f^- are respectively the upper and lower boundaries of the failure probability. The upper and lower bound failure probabilities are defined as:

$$P_f^+ = P(\mu_{\hat{G}}(x) + k.\sigma_{\hat{G}}(x) \leq 0) \tag{5.3}$$

$$P_f^- = P(\mu_{\hat{G}}(x) - k.\sigma_{\hat{G}}(x) \leq 0) \tag{5.4}$$

where $\mu_{\hat{G}}(x) + k.\sigma_{\hat{G}}(x) = 0$ and $\mu_{\hat{G}}(x) - k.\sigma_{\hat{G}}(x) = 0$ are respectively the upper and lower boundary of the limit state surface defined by $\mu_{\hat{G}}(x) = 0$ and k is a constant that sets the confidence level typically equal to 1.96 = Φ^{-1} (97.5%). In this chapter, a value of k=2 was adopted. Finally, ε_{P_f} in Equation (5.2) is a given tolerance for two consecutive iteration steps.

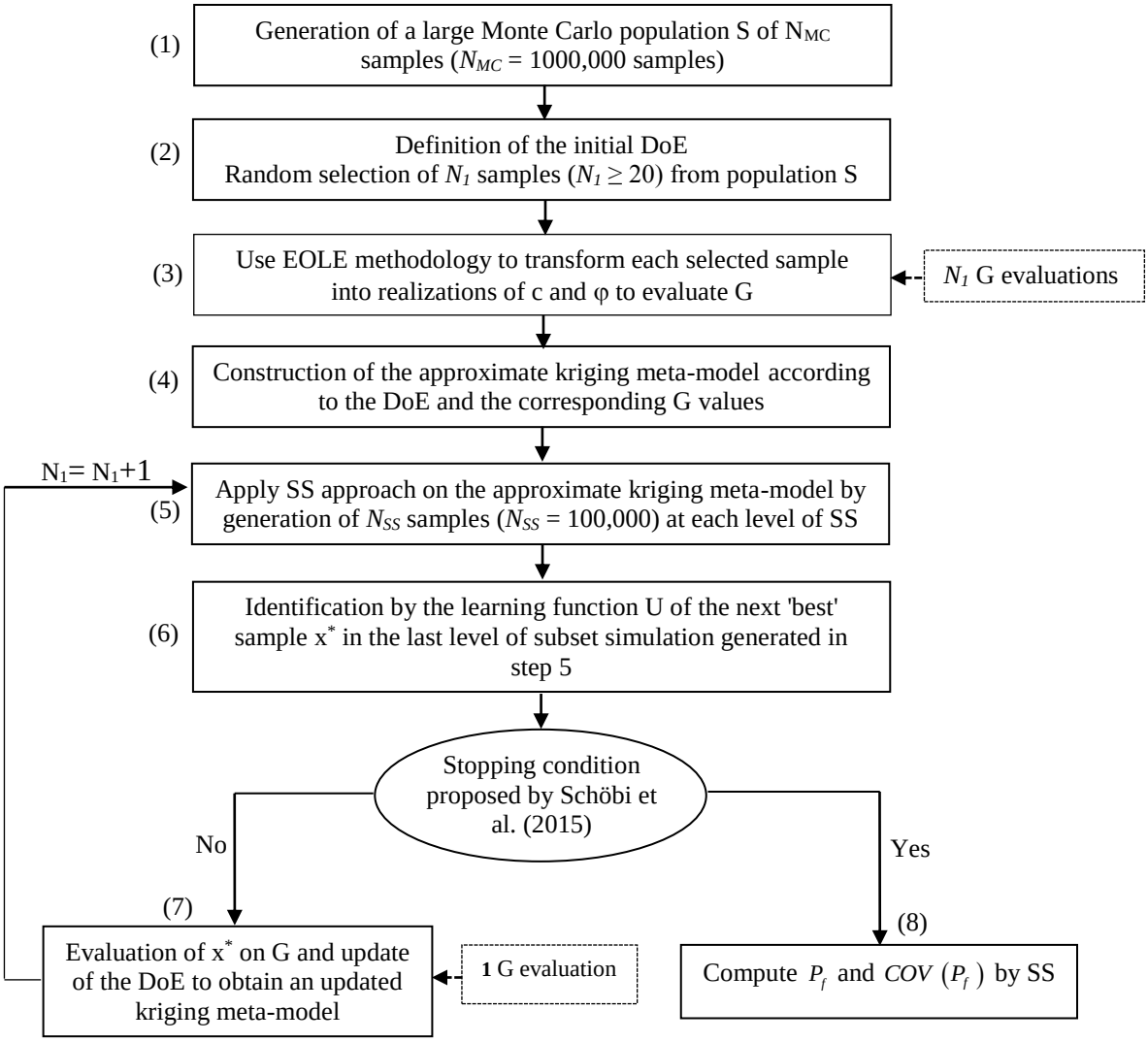


Figure 5.1. The flowchart of the proposed AK-SS method

5.3 VALIDATION OF THE PRESENT AK-SS PROCEDURE VIA AN ANALYTICAL EXAMPLE

In order to check the performance of the present AK-SS procedure, an academic example involving an explicit non-linear limit state function is considered. It consists of a series system with four branches. This example exhibits four design points. It has been studied by Schueremans and Van Gemert (2005), Echard et al. (2011), Bourinet et al. (2011), Balesdent et al. (2013) and Huang et al. (2016) among others. The validation of the present AK-SS procedure was performed by comparison of the results of the present example with those of AK-MCS and the crude MCS and SS methods.

The performance function G of the series system with four branches is defined as follows:

$$G\left(x_1,x_2\right)=\min\begin{Bmatrix}3+\left(x_1-x_2\right)^2/10-\left(x_1+x_2\right)/\sqrt{2}\\3+\left(x_1-x_2\right)^2/10+\left(x_1+x_2\right)/\sqrt{2}\\\left(x_1-x_2\right)+7/\sqrt{2}\\\left(x_2-x_1\right)+7/\sqrt{2}\end{Bmatrix} \quad (5.5)$$

where x_1 and x_2 are two standard normal random variables.

The approximate kriging meta-model was constructed herein using an initial DoE of 80 samples. It should be emphasized that the large number of samples required in stage 1 (i.e. 80 samples) may be explained by the high non-linearity of the limit state surface; a smaller number of samples led to difficulties in constructing the initial kriging meta-model. Notice that the process of adding new samples has stopped according to the stopping condition proposed by Schöbi et al. (2015).

Table 5.1 provides the value of the failure probability P_f and the corresponding value of the coefficient of variation $COV\left(P_f\right)$ together with the total number of calls of the performance function for different values of the tolerance ε_{P_f}. As may be seen from this table, one obtains convergence of the failure probability for ε_{P_f} =5% which corresponds to 130 added samples. This means that the meta-model does not present any bias beyond this number of samples in the zone of interest for the computation of the failure probability (i.e. around the four design points, which are the closest ones to the origin of the standard normal space) as may be seen from Figures 5.2 to 5.9.

Figure 5.10 presents the plots of P_f and $COV\left(P_f\right)$ versus the number of added samples for the adopted value of 5% for ε_{P_f}. Although the enrichment process required 130 samples, Figure 5.10 shows that the convergence of the failure probability was achieved at about 100 samples and that at this number of samples, the meta-model does not present a significant bias. The use of additional samples beyond 100 samples ensures the accuracy of the statistics as required by the

adopted criterion. Finally, notice that the small fluctuations observed beyond 100 samples may be explained by the fact that the different values of the failure probability are computed from different subset simulations with different random samples for each subset.

Table 5.1. Results of present AK-SS approach for the four- branches series system when using different values of ε_{P_f}

ε_{P_f} %	$P_f \times 10^{-3}$	$COV\left(P_f\right)$ (%)	N_{call} = DoE + number of added samples
50	9.973	2.33	80 +7= 87
10	4.719	2.38	80 +27= 107
7	3.236	2.47	80 +82= 162
6	2.725	2.48	80 +97= 177
5	2.236	2.58	80 +130= 210
1	2.259	2.53	80 +193= 273
0.5	2.237	2.54	80 +236= 316

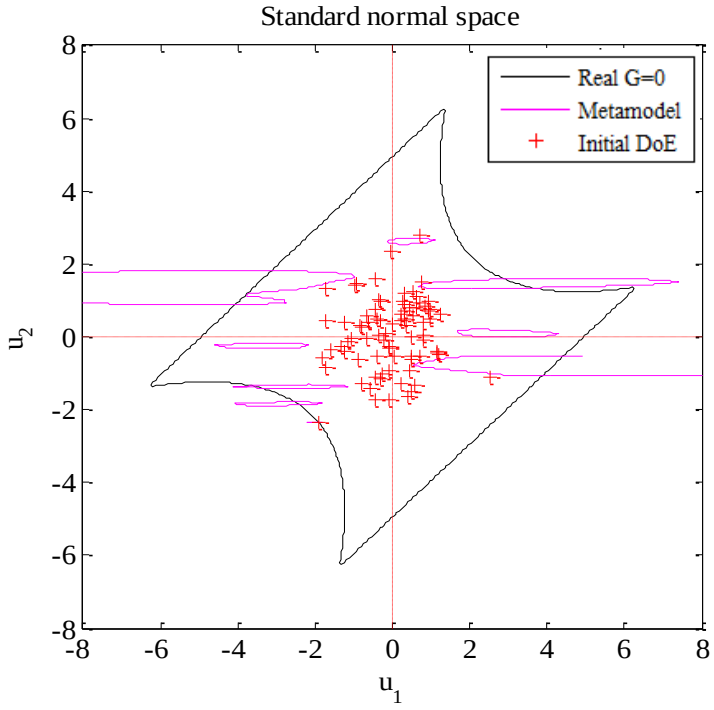


Figure 5.2. Results of present AK-SS approach for the four- branches series system (using DoE only)

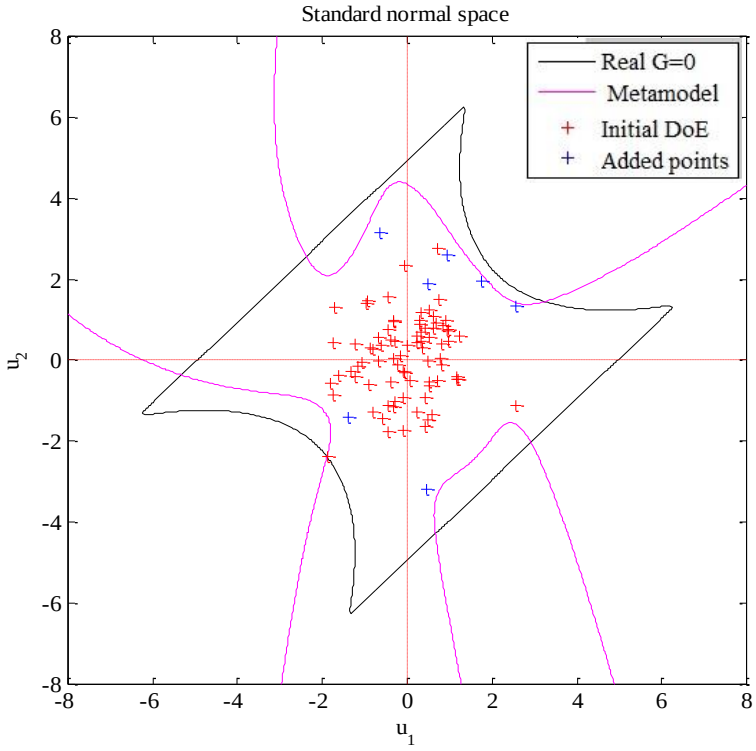


Figure 5.3. Results of present AK-SS approach for the four- branches series system (with ε_{P_f} = 50%)

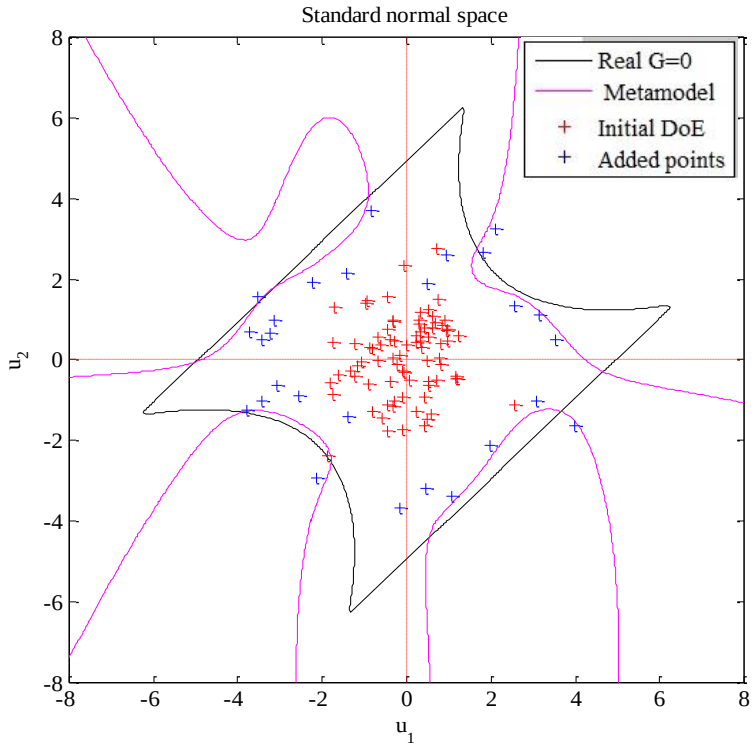


Figure 5.4. Results of present AK-SS approach for the four- branches series system (with ε_{P_f} = 10%)

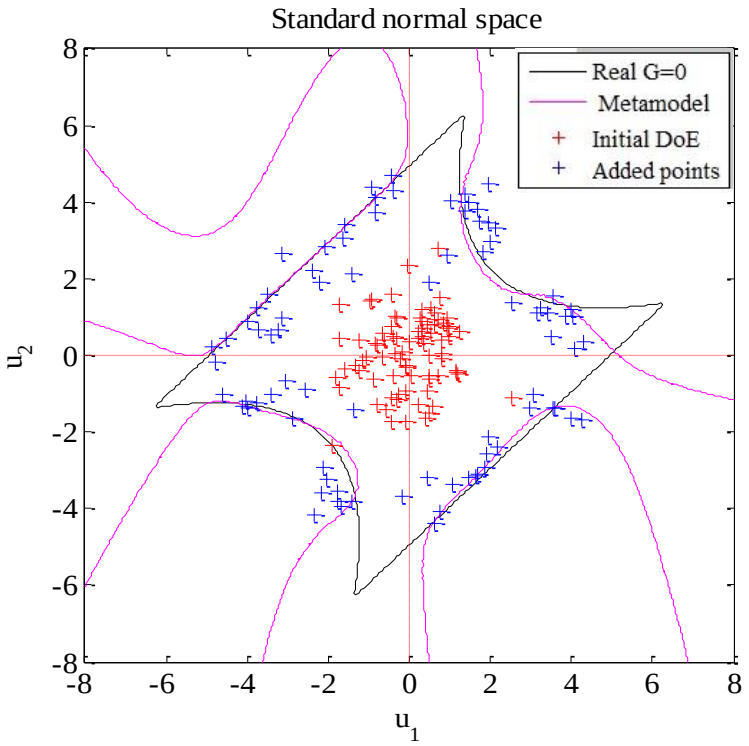


Figure 5.5. Results of present AK-SS approach for the four- branches series system (with ε_{P_f} =7%)

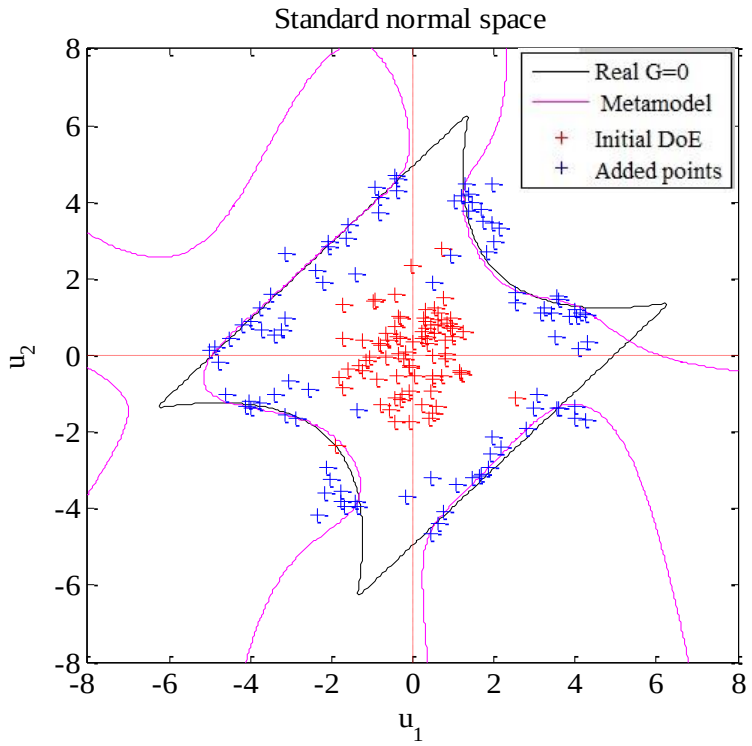


Figure 5.6. Results of present AK-SS approach for the four- branches series system (with ε_{P_f} = 6%)

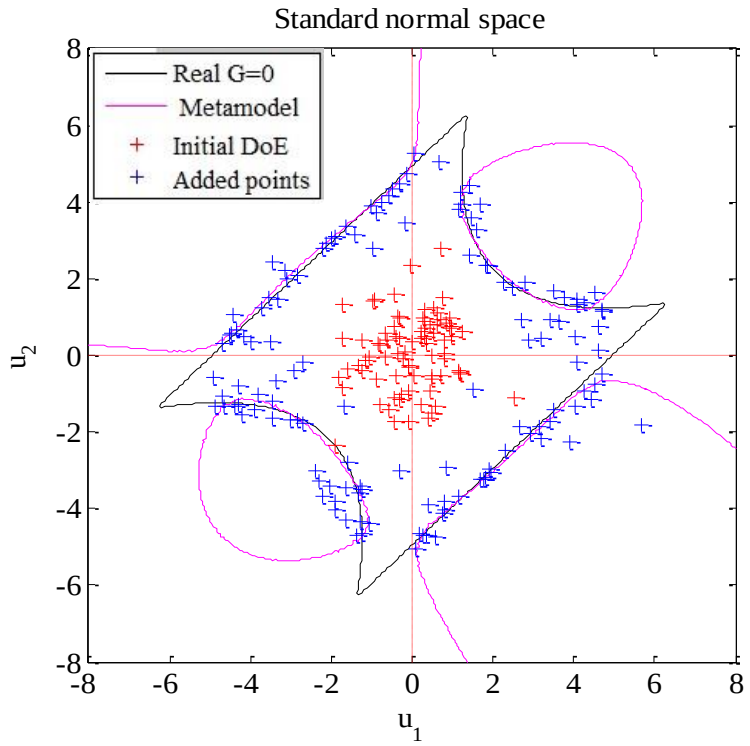


Figure 5.7. Results of present AK-SS approach for the four- branches series system (with ε_{P_f} = 5%)

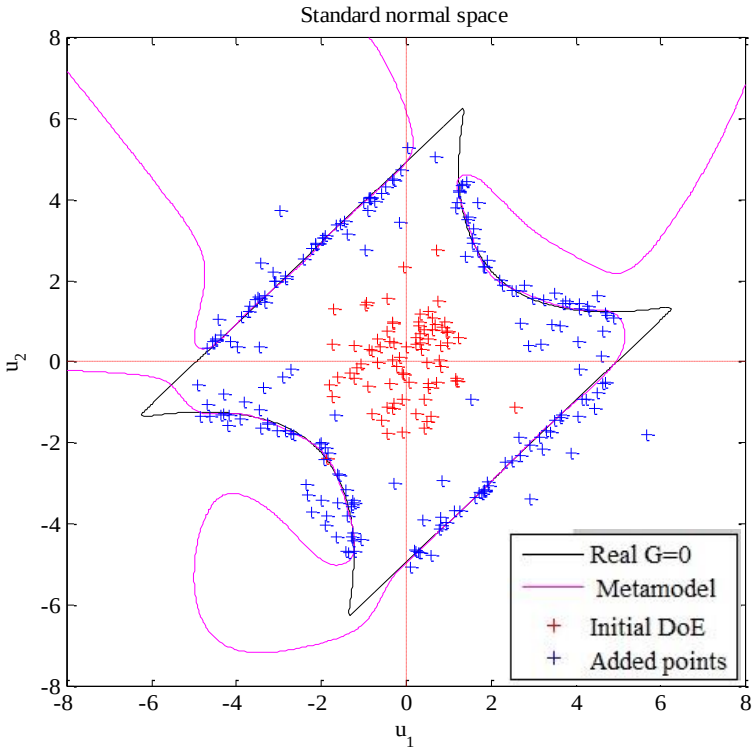


Figure 5.8. Results of present AK-SS approach for the four- branches series system (with ε_{P_f} = 1%)

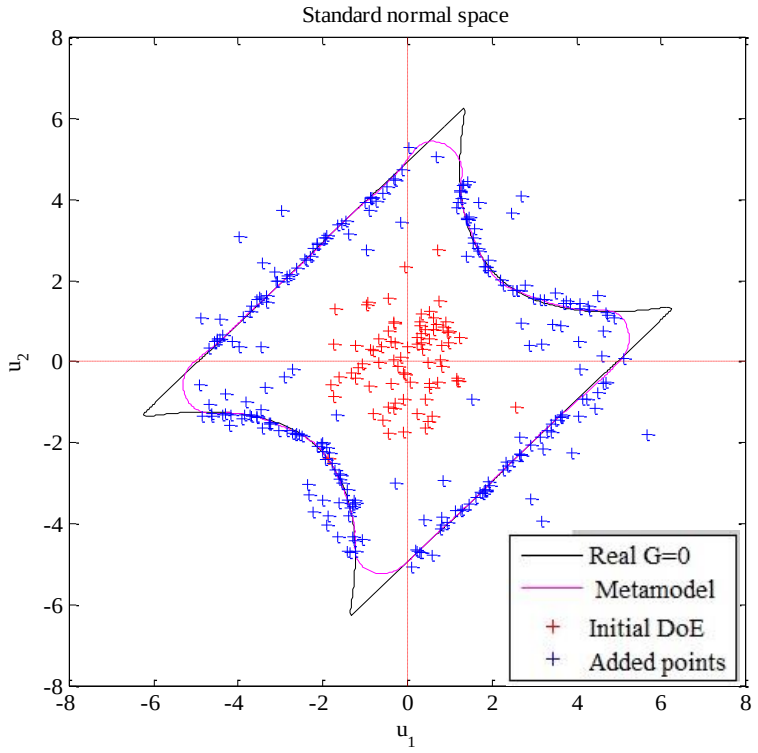


Figure 5.9. Results of present AK-SS approach for the four- branches series system (with ε_{P_f} = 0.5%)

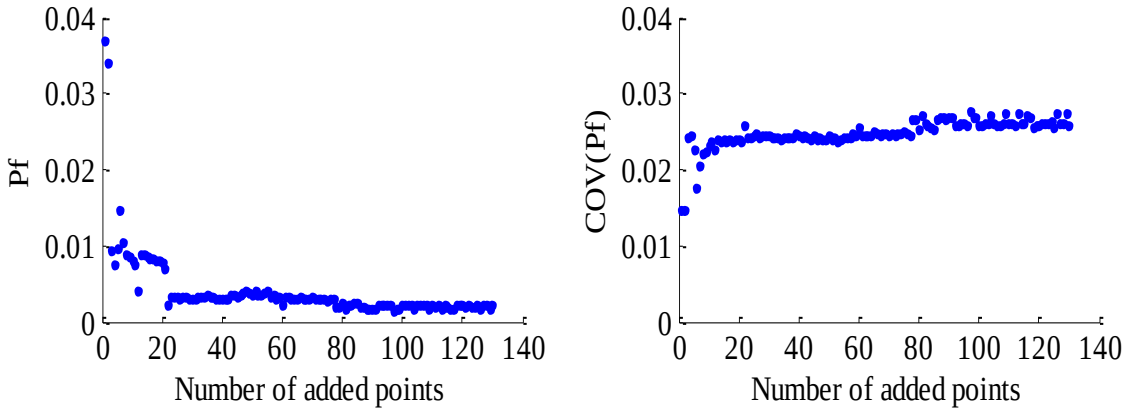


Figure 5.10. P_f and $COV\ (P_f)$ values as function of the added points for the four- branches series system

Table 5.2 provides a comparison between the results obtained by the present AK-SS approach and those given by AK-MCS and the crude MCS and SS methods in terms of the failure probability and the number of calls of the performance function. It should be noted here that the results of AK-MCS are presented for both criteria (or stopping conditions) [i.e. (min U>2) and

the one proposed by Schöbi et al. (2015)] and using the same initial DoE as AK-SS (i.e. 80 samples) as may be seen from Figures 5.11 and 5.12 and Table 5.2. Notice also that the reference values adopted for both the failure probability P_f and the coefficient of variation $COV\left(P_f\right)$ are those obtained by the crude MCS runs with $N_{Mc} = 10^6$ samples. This table shows that there is quite good agreement between the results of the present AK-SS approach and those of the other approaches, the AK-SS being less time-consuming than AK-MCS in terms of the number of calls of the performance function. Indeed, the present AK-SS approach requires 50 less samples than the AK-MCS approach using the stopping condition proposed by Schöbi et al. (2015)). Finally, it should be noted here that the adopted value of ε_{P_f} used in Table 5.2 when using Schöbi's criterion was equal to 1%.

Table 5.2. Results of the failure probability P_f , the corresponding coefficient of variation $COV\left(P_f\right)$ and the number of calls N_{call} for different methods

Method	$P_f \times 10^{-3}$	$COV\left(P_f\right)$ (%)	N_{call} = DoE + number of added samples
MCS	2.233	2.11	10^6
SS	2.233	8.87	3×10^4
Present AK-MCS (U criterion)	2.262	2.10	80 + 409 = 489
Present AK-MCS (criterion of Schöbi et al. (2015))	2.268	2.10	80 + 243 = 323
Present AK-SS	2.259	2.53	80 + 193 = 273

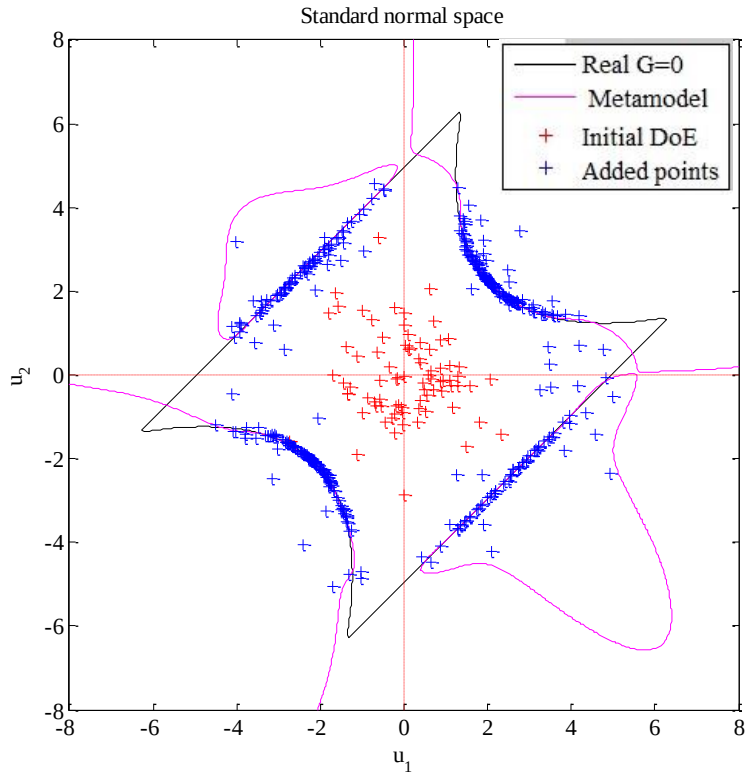


Figure 5.11. Results of present AK-MCS approach (U criterion) for the four- branches series system

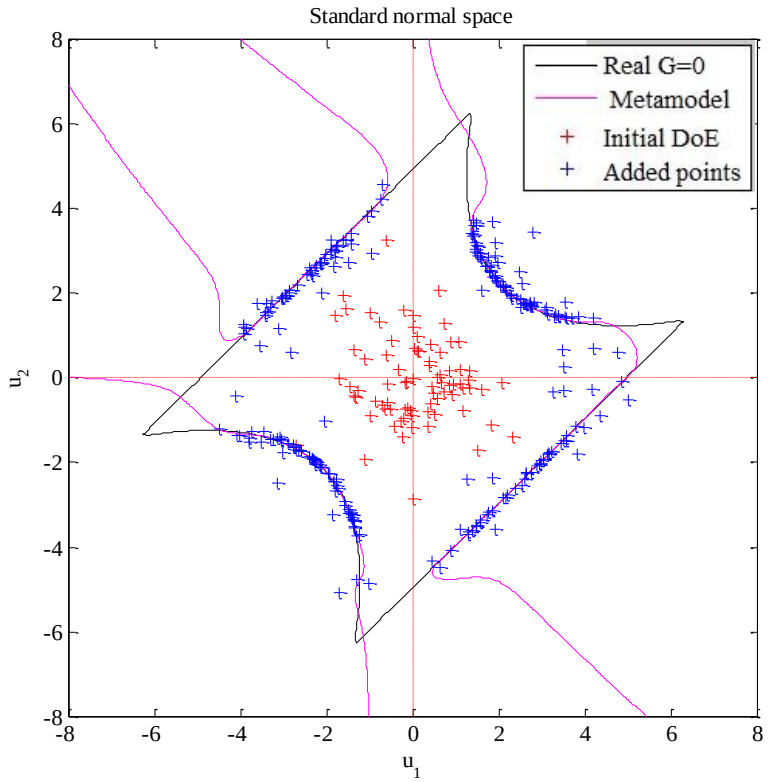


Figure 5.12. Results of present AK-MCS approach (criterion of Schöbi et al. (2015)) for the four- branches series system

5.4 PROBABILISTIC RESULTS IN THE CASE OF A SPATIALLY VARYING SOIL

This section aims at presenting the impact of the soil spatial variability on the failure probability against soil punching of the same strip footing presented in the two previous chapters. The same deterministic and uncertain parameters are considered in this section. Also, the input data employed in chapters 3 and 4 for both the deterministic and the uncertain parameters remain the same herein.

In this section, a value of 1% was adopted for ε_{P_f} except for the configurations corresponding to smaller values of the autocorrelation distances (i.e. where the number of eigenmodes was larger than 20) as may be seen from Table 5.4. Indeed; for these configurations, the ε_{P_f} value was taken equal to 5% because these configurations correspond to a very heterogeneous soil and thus, they require a much greater computation time to achieve a better accuracy.

5.4.1 Evolution of the limit state surface during the enrichment process

Figure 5.13 presents the evolution of the limit state surface (LSS) with the number of added realizations for a typical case where a_x=10,000m and a_y=10,000m. This LSS was obtained using the present AK-SS approach. The adopted DoE was equal to 20 samples and the needed number of added samples during the enrichment process was equal to 27 samples. As may be seen from this figure, the limit state surface is successively improved with the addition of new samples during the enrichment process. The final obtained LSS was similar to that provided in the two previous chapters.

It should be emphasized here that the adopted value of ε_{P_f} was taken equal to 0.1% in the present section (not 1% as mentioned above). This small value was necessary to lead to an accurate LSS over the entire range of variation of the random variables. A greater value of ε_{P_f} (ε_{P_f} =1%) would also be acceptable if only the computation of P_f value was required. Using Schöbi criterion with ε_{P_f} =1%, the needed number of added samples for the computation of the failure probability would be equal to only 8 samples instead of 27 samples.

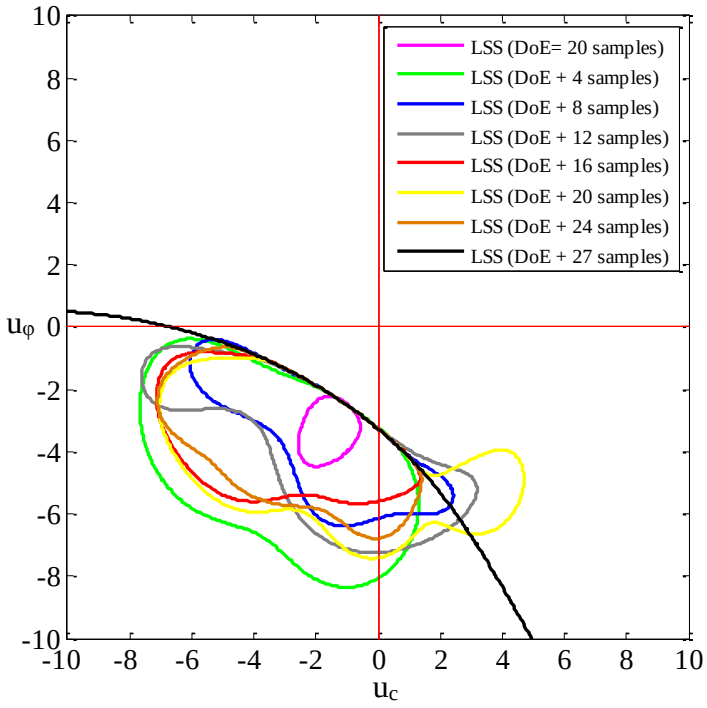


Figure 5.13. Effect of number of added realizations on the limit state surface when a_x = a_y = 10,000 m

5.4.2 Evolution of the probabilistic outputs during the enrichment process

Figure 5.14 presents the evolution of P_f^0 , P_f^+ and P_f^- as function of the number of added samples in the case where a_x=10m and a_y=1m. As mentioned above (see also Table 5.4), the adopted value of ε_{P_f} was equal to 5% in the present case. From Figure 5.14, one may observe that 511 calls to the mechanical model were necessary to achieve such accuracy. The final obtained values of P_f and $COV\left(P_f\right)$ were respectively 1.658×10^{-3} and 2.75%.

It should be emphasized here that although the stopping condition by Schöbi is less severe than the U criterion (and thus, it significantly reduces the needed number of added samples), the Schöbi criterion remains quite severe. The reason is due to the fact that one continues to add new samples until the marge between the upper and lower boundaries of the failure probability attains a prescribed tolerance, although the value of P_f could vary only very slightly with the addition of new samples (as may be shown from Figure 5.14 and Table 5.3).

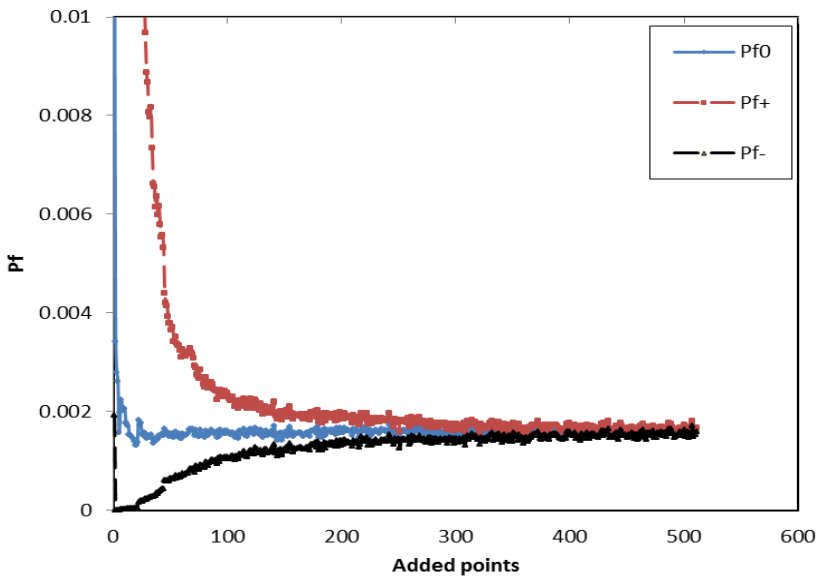


Figure 5.14. AK-SS results for a spatially varying soil (a_x=10 m, a_y=1 m) with ε_{P_f} = 5%

Table 5.3. Results of present AK-SS approach for a spatially varying soil (a_x=10 m, a_y=1 m) when using different values of ε_{P_f}

ε_{P_f} %	$P_f \times 10^{-3}$	$COV(P_f)$ (%)	number of added samples
90	1.5735	2.77	89
80	1.608	2.77	100
70	1.461	2.79	110
60	1.535	2.77	123
50	1.561	2.77	145
40	1.624	2.77	173
30	1.554	2.79	216
20	1.624	2.77	276
10	1.666	2.76	398
5	1.658	2.75	511

5.5 PROBABILISTIC PARAMETRIC STUDY

Similarly to the two previous chapters, a parametric study is performed in this section using the present AK-SS approach to investigate the effect of the autocorrelation distances of the two random fields on the P_f value. A comparison between the results obtained by the three methods AK-MCS, AK-IS and AK-SS is presented and discussed.

5.5.1 Effect of the autocorrelation distances of the random fields on P_f and $COV\left(P_f\right)$

The values of the probabilistic outputs given by AK-SS and corresponding to the different soil variabilities were given in Table 5.4. In order to show that no bias in the meta-model exists at the end of the enrichment process for all configurations presented in Table 5.4, Appendix G presents the plots of P_f for different values of the autocorrelation distances considered in this table. For completeness, also the values of $COV\left(P_f\right)$ are also given in this appendix.

Figure 5.15 presents the effect of the isotropic autocorrelation distance (a_x=a_y) on the failure probability P_f as obtained from AK-MCS, AK-IS and AK-SS methodologies. Also, Figure 5.16 and Figure 5.17 present the effect of the vertical or horizontal autocorrelation distance on the failure probability P_f as computed by the three methods. Figure 5.15, 5.16 and 5.17 show that there is a good agreement between the results of P_f obtained from the three methods. Except for the case where a very heterogeneous soil case was considered (i.e. a_x =10m and a_y=0.5m), the maximum percentage difference with respect to the AK-MCS method (considered as a reference) does not exceed 7%.

Figures 5.18, 5.19 and 5.20 show the values of $COV\left(P_f\right)$ as obtained from the three methods AK-MCS, AK-IS and AK-SS for the different soil variabilities. From these figures, one may observe that the values of $COV\left(P_f\right)$ obtained from the present AK-SS are smaller than 4% for all the treated cases. Notice however that for the two other methods (i.e. AK-MCS and AK-IS), the values of $COV\left(P_f\right)$ are smaller than 7% for all the treated cases expect for the case of very heterogeneous soil (i.e. a_x=10m, a_y=0.5m) where $COV\left(P_f\right)$ is equal to 15% in AK-IS method. In this case, a larger number of IS samples is required in order to reduce the obtained value of $COV\left(P_f\right)$. As a conclusion, the small values of $COV\left(P_f\right)$ (<7%) obtained from the different

methods indicate that the values of P_f computed in this thesis can be used with confidence in practice.

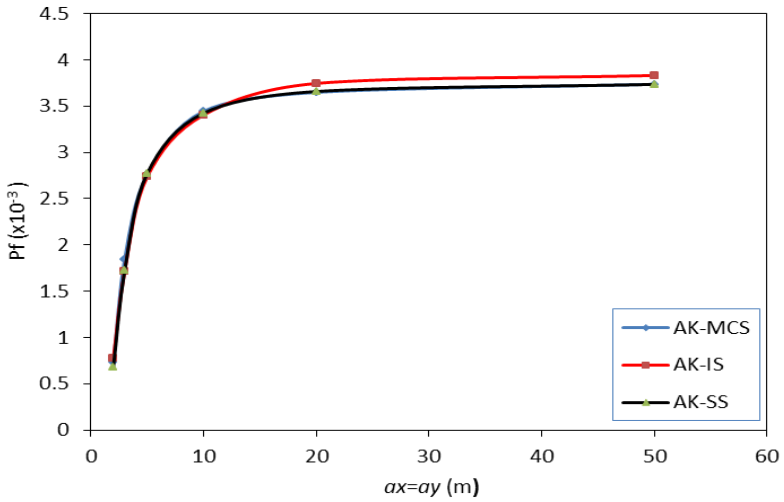


Figure 5.15. Effect of the isotropic autocorrelation distance a_x=a_y on P_f

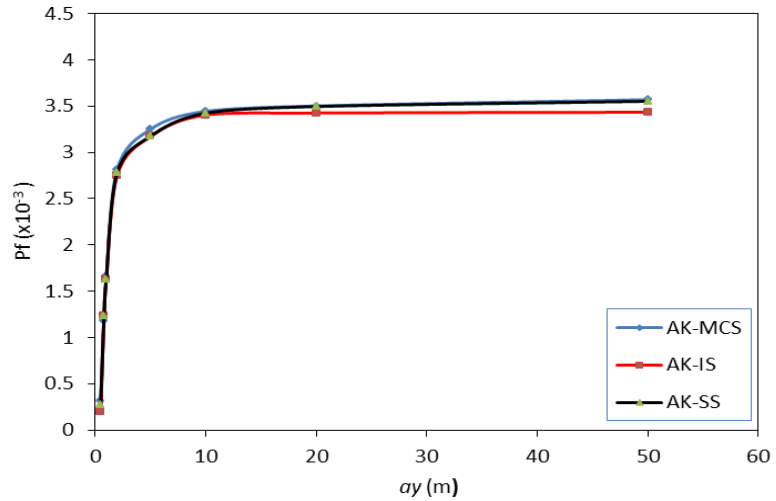


Figure 5.16. Effect of the vertical autocorrelation distance a_y on P_f when a_x =10 m

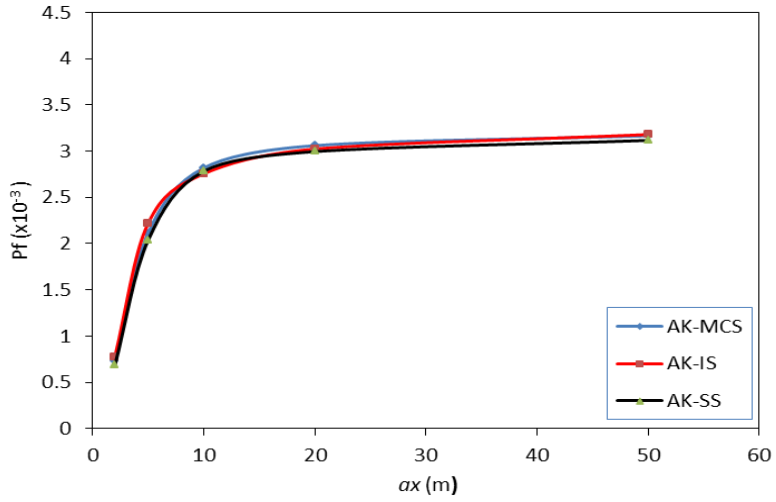


Figure 5.17. Effect of the horizontal autocorrelation distance a_x on P_f when a_y=2 m

Table 5.4. Adopted number of random variables and the corresponding value of the variance of error of EOLE together with the values of ε_{P_f} , P_f , $COV\left(P_f\right)$ and number of added realizations for various soil variabilities

a. Case of an isotropic case (a_x=a_y)

a_x=a_y (m)	Adopted number of random variables	Variance of the error %	ε_{P_f} %	$P_f \times 10^{-3}$	$COV\ (P_f)$ %	Number of added realizations
2	48	9.447	5	0.684	3.11	>584
3	32	4.647	5	1.722	2.74	383
5	24	0.953	5	2.772	2.56	247
10	10	0.815	1	3.424	2.50	84
20	8	0.170	1	3.658	2.48	57
50	6	0.016	1	3.735	2.48	44
100	6	0.001	1	3.761	2.49	46

b. Case of an anisotropic case (a_x=10 m with varying a_y)

a_y (m)	Adopted number of random variables	Variance of the error %	ε_{P_f} %	$P_f \times 10^{-3}$	$COV\ (P_f)$ %	Number of added realizations
0.5	44	9.119	5	0.278	3.18	>601
0.8	38	4.798	5	1.163	3.01	546
1	32	4.212	5	1.658	2.75	511
2	24	1.437	5	2.783	2.53	256
5	12	1.682	1	3.175	2.51	136
10	10	0.815	1	3.424	2.50	84
20	8	0.855	1	3.495	2.49	83
50	8	0.297	1	3.552	2.49	72
100	8	0.099	1	3.582	2.52	74

c. Case of an anisotropic case (a_y=2 m with varying a_x)

a_x (m)	Adopted number of random variables	Variance of the error %	ε_{P_f} %	$P_f \times 10^{-3}$	$COV\ (P_f)$ %	Number of added realizations
2	48	9.447	5	0.684	3.11	>584
5	30	4.101	5	2.039	2.63	471
10	24	1.437	5	2.783	2.53	256
20	16	1.415	1	2.995	2.54	169
50	12	1.272	1	3.116	2.53	106
100	10	0.842	1	3.164	2.53	95

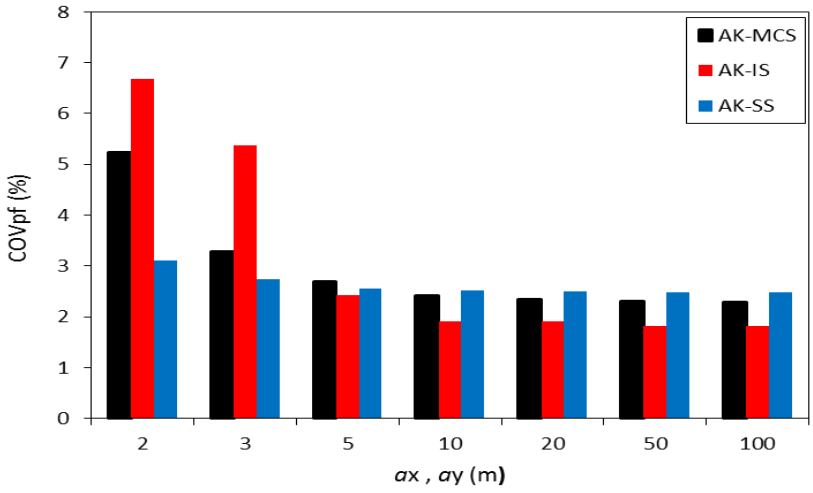


Figure 5.18. Values of $COV\left(P_f\right)$ for different values of the isotropic autocorrelation distance a_x=a_y

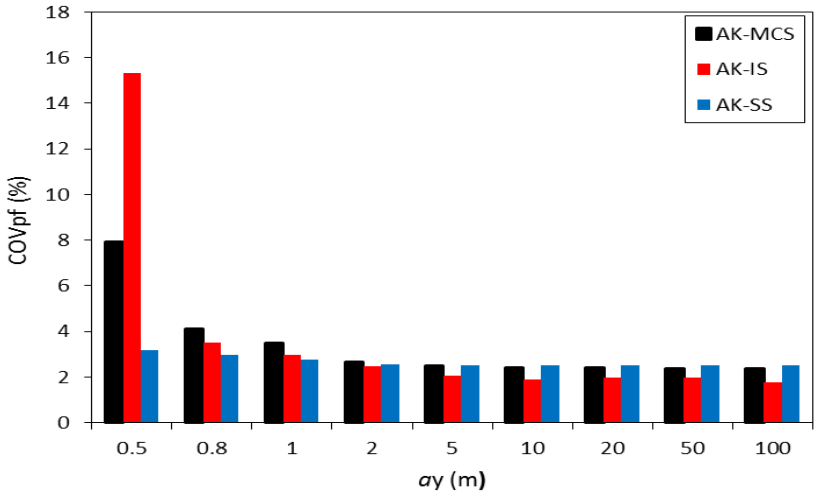


Figure 5.19. Values of $COV\left(P_f\right)$ for different values of the vertical autocorrelation distance a_y when a_x=10m

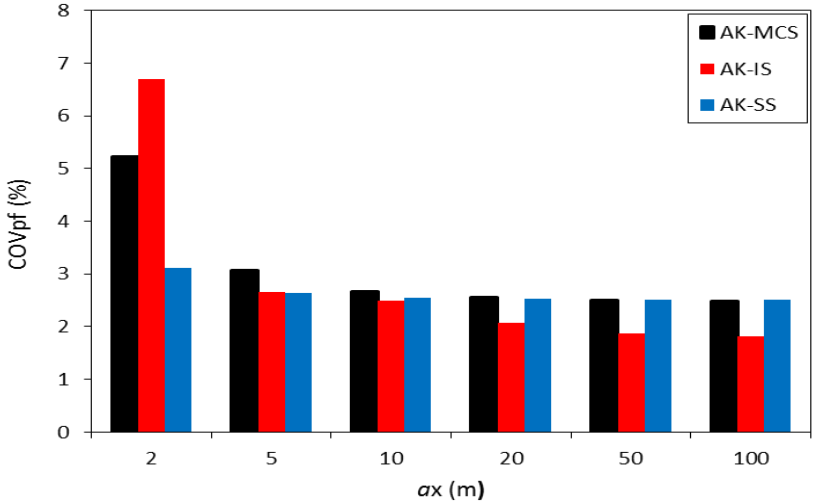


Figure 5.20. Values of $COV\left(P_f\right)$ for different values of the horizontal autocorrelation distance a_x when a_y=2m

5.6 CONCLUDING REMARKS REGARDING THE DIFFERENT PROBABILISTIC METHODS

The kriging-based reliability methods AK-MCS, AK-IS and AK-SS have been presented in this thesis for the probabilistic analysis of a strip footing resting on a spatially varying soil. These methods are very efficient as the obtained probability of failure is very accurate needing a smaller number of calls to the mechanical model as compared to the corresponding simulation technique used individually such as Monte Carlo Simulations (MCS), Importance Sampling (IS) and Subset Simulation (SS).

Table 5.5 presents a comparison between these methods in terms of computational time and total number of calls to the computationally expensive mechanical model in the case where a_x=10m and a_y=2m. It should be mentioned here that for the present case, the AK-MCS and AK-IS methodologies were repeated in this section using the stopping condition proposed by Schöbi et al. (2015) with ε_{P_f} =5% as may be seen from Figures 5.21 and 5.22. Thus; for those two methods (i.e. AK-MCS and AK-IS), the results are presented for both stopping conditions [i.e. (min U>2) and the one proposed by Schöbi et al. (2015)].

As may be seen from Table 5.5, the three methods AK-MCS (with U criterion), AK-IS (with U criterion) and AK-SS have provided quasi-similar values for the probabilistic outputs (i.e. P_f and $COV\left(P_f\right)$). Notice however that there is some discrepancy with the solutions given by both AK-MCS and AK-IS when the Schöbi criterion is used with ε_{P_f} =5%. For a value of ε_{P_f} =1%, the P_f values given by AK-MCS (Schöbi) and AK-IS (Schöbi) become equal to 2.66 and 2.73 respectively.

From Table 5.5, one may observe that the AK-IS method with Schöbi criterion was the most advantageous in terms of computation time (about one day) followed by AK-SS (about two days) and then by AK-MCS with Schöbi criterion (about 4 days).

As a conclusion, the AK-IS with Schöbi criterion should be used in case of a single design point. When dealing with limit state surfaces that possess more than one design point, the AK-IS method loses its interest and thus, one can overcome this shortcoming by using the slightly more expensive AK-SS approach.

Table 5.5. Comparison of the results of the different methods when a_x =10 m and a_y =2 m

Method	Candidate population	$P_f \times 10^{-3}$	$COV\left(P_f\right)\%$	Total number. of calls to the mechanical model	Computation time (days)
AK-MCS_U	5×10^5	2.818	2.66	692	12
AK-MCS based on Schöbi criterion	5×10^5	2.348	2.92	221	4
AK-IS based on U criterion	1×10^4	2.755	2.68	689	3
AK-IS based on Schöbi criterion	1×10^4	2.666	2.32	200	1
AK-SS	1×10^5	2.783	2.53	280	2

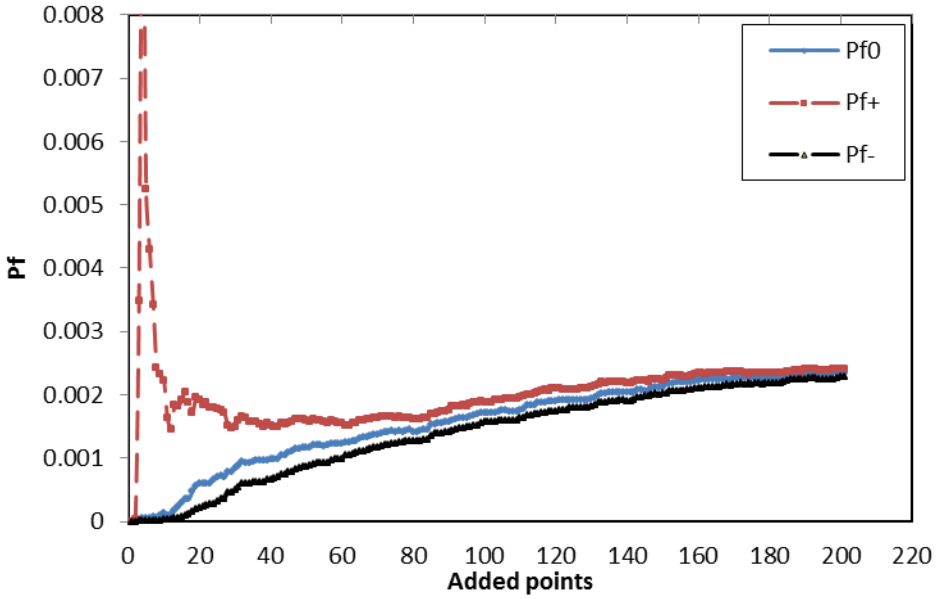


Figure 5.21. AK-MCS results based on Schöbi criterion for a spatially varying soil (a_x=10 m, a_y=2 m) with ε_{P_f} = 5%

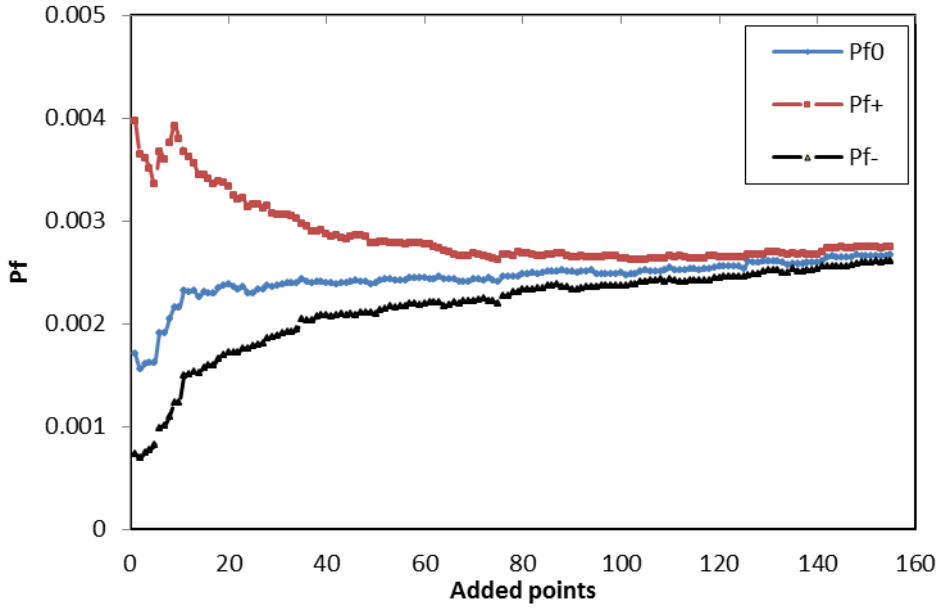


Figure 5.22. AK-IS results based on Schöbi criterion for a spatially varying soil (a_x=10 m, a_y=2 m) with ε_{P_f} = 5%

5.7 CONCLUSION

This chapter mainly presented a probabilistic analysis at the ultimate limit state of a strip footing resting on a spatially varying soil using an active learning reliability method combining kriging and Subset Simulation (called AK-SS). The new method AK–SS takes advantages of subset simulation for evaluating the small failure probabilities and the Kriging model with active learning and updating characteristic for approximating performance function.

Within AK-SS, an efficient sampling technique (Subset Simulation SS) is used instead of Monte Carlo or Importance Sampling employed in Chapters 3 and 4 respectively. The present AK-SS procedure consists of two main steps. In the first step, an approximate kriging meta-model is constructed using a small DoE. These samples are randomly selected from a large population generated by MCS. In the second step, the subset simulation approach is first used to generate samples which are directed towards the limit state surface by employing the obtained approximate kriging meta-model. Then, the approximate kriging meta-model is improved by adding a 'best' new sample to the initial DoE. The new best sample is selected among the

samples obtained in the final level of the subset simulation making use of the learning function U. The process of adding a new sample (process of enrichment) is repeated until a prescribed criterion on the value of the failure probability is obtained. The AK-SS technique allows one to overcome the search of the design points (as is the case in AK-IS) and thus, it can deal with arbitrary limit state surfaces.

The objective of this chapter was the computation of the probability of failure against soil punching of a strip footing resting on a spatially varying soil and subjected to a prescribed vertical load. The soil cohesion and angle of internal friction were considered as two non-isotropic non-Gaussian random fields. The deterministic model was based on numerical simulations using the finite difference code FLAC^{3D}. The main findings of this study can be summarized as follows:

1. The main advantage of AK-SS over AK-MCS is visible when dealing with small failure probabilities. In fact, the computational effort of AK-MCS increases dramatically due to the large populations required for the assessment of small probabilities while AK–SS solves this problem by expressing the small failure probability as a product of larger conditional failure probabilities of several intermediate failure events.
2. AK-SS approach is not based on any assumption about the shape of the limit state surface like AK-IS. Thus, this method can handle complex problems with several design points as was shown in the analytical academic example.
3. The new stopping criterion adopted in the AK-SS method (i.e. Schöbi criterion) which monitors the convergence of the statistics of interest is better than the stopping criterion which was based on the accuracy of the meta-model (i.e. the U criterion). This result in a significant reduction in the computation time needed because this stopping criterion leads to the smallest number of calls to the mechanical model.

4. Although the stopping condition by Schöbi is less severe than the U criterion, this stopping condition remains quite severe. The reason is due to the fact that one continues to add new samples until the marge between the upper and lower boundaries of the failure probability attains a prescribed tolerance.
5. The values of $COV\left(P_f\right)$ obtained from the present AK-SS are smaller than 4% for all the treated cases. Notice however that for the two other methods (i.e. AK-MCS and AK-IS), the values of $COV\left(P_f\right)$ are smaller than 7% for all the treated cases expect for the case of very heterogeneous soil (i.e. a_x=10m, a_y=0.5m) where $COV\left(P_f\right)$ is equal to 15% in AK-IS method. In this case, a larger number of IS samples is required in order to reduce the obtained value of $COV\left(P_f\right)$. As a conclusion, the small values of $COV\left(P_f\right)$ (<7%) obtained for most configurations indicate that the values of P_f obtained in this thesis can be used with confidence in practice.
6. A comparison between the results of P_f obtained using the different methods AK-MCS, AK-IS and AK-SS has shown good agreement. The AK-IS is the most advantageous method concerning the computation time followed by the AK-SS method and then by the AK-MCS method. Thus, AK-IS is the most suitable method in case of a single design point. AK-SS may be considered as an excellent alternative for problems involving several design points since it overcomes the shortcoming of AK-IS involving the necessity of computation of the design points.

CHAPTER 6. GENERAL CONCLUSIONS

6.1 INTRODUCTION

The aim of this thesis is to investigate the effect of the soil spatial variability on the failure probability against soil punching of a strip footing resting on a spatially varying soil and subjected to a vertical load. When dealing with the computation of the failure probability of geotechnical structures involving spatially varying soils, the classical Monte Carlo Simulation (MCS) methodology is generally used. This method is known to be very time-consuming. This is because it usually makes use of finite element or finite difference models which are generally time-expensive and more importantly it requires a large number of calls of the mechanical model for the computation of the small failure probabilities encountered in practice. For this reason, only the two first statistical moments (i.e. the mean and the standard deviation) of the system response were extensively investigated in literature using this approach.

In order to overcome the shortcoming related to the excessive number of calls of the mechanical model when performing a probabilistic analysis, this thesis investigates the efficiency of three new metamodeling approaches (AK-MCS, AK-IS and AK-SS). These approaches consist in combining the kriging metamodeling and one of the three simulation techniques (i.e. MCS, IS or SS). The computation of the failure probability may thus be easily performed using a meta-model.

The three methods were applied to the computation of the probability of failure against soil punching of a strip footing resting on a spatially varying soil and subjected to a vertical load. The system response involves the ultimate bearing capacity (q_u) of the vertically loaded strip footing. The soil cohesion and angle of internal friction were considered as anisotropic non-Gaussian random fields. The mechanical model employed for the computation of the system response was based on numerical simulations using the finite difference code FLAC^{3D}.

6.2 REASERCH CONCLUSIONS

The main findings of this study can be summarized as follows:

1. Kriging-based approaches (AK-MCS, AK-IS and AK-SS) were found to significantly reduce the number of calls to the mechanical model with respect to the corresponding classical simulation methods (i.e. MCS, IS and SS).

2. The values of the probability of failure P_f obtained from the three methods AK-MCS, AK-IS and AK-SS are very accurate because the maximum obtained value of the coefficient of variation $COV\left(P_f\right)$ (for most configurations) is about 7%. Thus, these values of P_f may be used with confidence in practice.

3. The prescribed value of the variance of the error of EOLE methodology has a significant influence on the value of the failure probability. A threshold of 5% may be adopted to obtain accurate values of the failure probability. Consequently, the results of the few configurations corresponding to a greater value of the variance of the error may be used with caution. Indeed, these configurations need a significant number of random variables. This large number of random variables cannot be used within the present approaches because of the significant computation time needed for those cases. A more advanced probabilistic method is needed to handle such configurations.

4. The iterative procedure proposed in AK-IS approach for the determination of the design point is a powerful tool that has been shown to perform successfully even in complex cases involving very heterogeneous soil media.

5. The three approaches (AK-MCS, AK-IS and AK-SS) have shown similar patterns. The probability of failure P_f increases with the increase in the isotropic autocorrelation distance and then, it attains an asymptote in the case of a homogeneous soil. Also, the failure

probability P_f increases with the increase in the vertical or horizontal autocorrelation distance and then, it attains an asymptote corresponding to the case of a 1D random field.

6. There is a good agreement between the results of P_f obtained from the three approaches AK-MCS, AK-IS and AK-SS. In case of a single design point, AK-IS is the most advantageous method concerning the computation time followed by AK-SS and then by AK-MCS. AK-SS may be an excellent alternative for problems involving several design points.
7. For the two cases of isotropic and anisotropic random fields, the critical realizations at the design point exhibit a symmetrical distribution of the soil shear strength parameters with respect to the central vertical axis of the foundation. In addition, the small values of the soil cohesion and angle of internal friction are located in the vicinity of the foundation, the higher values of the soil resistance being far from the footing. This naturally leads to a failure mechanism that develops symmetrically within the weak soil zone near the foundation. The soft soil zone increases with the increase of the autocorrelation distances leading to a larger failure probability.
8. The new stopping condition by Schöbi et al. (2015) which controls the convergence of the statistics of interest is better than the stopping criterion based on the accuracy of the meta-model (i.e. the U criterion). It leads to a significant reduction in the calculation time.
9. Although the stopping condition by Schöbi is less severe than the U criterion, this stopping condition remains quite severe. The reason is due to the fact that one continues to add new samples until the marge between the upper and lower boundaries of the failure probability attains a prescribed tolerance.

6.3 RECOMMENDATIONS FOR FUTURE WORK

Based on the work that was performed in this thesis, the following directions for future work can be proposed as follows:

1. Apply the probabilistic approaches used in this thesis in the case of a multilayer soil medium (i.e. soft over stiff soil or stiff over soft soil) that exhibits spatial variability because the foundations are mostly supported by multi-layered soil profiles.
2. Consider the case of a rectangular or a circular footing resting on a soil with three-dimensional spatial variability of the soil properties. This problem is challenging since (i) the three-dimensional deterministic problems are known to be very time-consuming and (ii) the three-dimensional variation of the soil properties significantly increases the number of random variables needed in the discretization of random fields and this leads to a significant number of calls to the mechanical model.
3. Try to combine the Kriging meta-modeling technique with other simulation methods (i.e. directional simulation or Markov Chain Monte Carlo MCMC) that may lead to additional reduction in the computation time.
4. Extend the presented methods (AK-MCS, AK-IS and AK-IS) to the cases of multiple performance functions. This may be of practical use for some geotechnical problems involving more than a single mode of failure (e.g. retaining wall problem).
5. Try to improve the efficiency of adaptive Kriging, by considering a new stopping criterion and a new manner of selecting new training points. This may be performed by introducing the Global Sensitivity Analysis enhanced Surrogate (GSAS) modeling method.

REFERENCES

1. Ahmed, A., Soubra, A. H. (2011). “Subset simulation and its application to a spatially random soil.” Geo-Risk 2011: Risk Assessment and Management, 209–216.

2. Ahmed, A., and Soubra, A. H. (2012a). “Extension of subset simulation approach for uncertainty propagation and global sensitivity analysis.” Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards 6(3), 162–176.

3. Ahmed, A., and Soubra, A. H. (2012b). “Probabilistic analysis of strip footings resting on a spatially random soil using subset simulation approach.” Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards 6(3), 188–201.

4. Al-Bittar, T. (2012). “Probabilistic Analysis of Shallow Foundations Resting on Spatially Varying Soils.” PhD. Thesis, University of Nantes, France.

5. Al-Bittar, T., and Soubra, A. H. (2016). “Bearing capacity of spatially random rock masses obeying Hoek – Brown failure criterion.” Georisk : Assessment and Management of Risk for Engineered Systems and Geohazards.

6. Al-Bittar, T., and Soubra, A. H. (2013). “Bearing capacity of strip footings on spatially random soils using sparse polynomial chaos expansion.” International Journal for Numerical and Analytical Methods in Geomechanics, 37(13), 2039–2060.

7. Al-Bittar, T., and Soubra, A. H. (2014a). “Efficient sparse polynomial chaos expansion methodology for the probabilistic analysis of computationally-expensive deterministic models.” International Journal for Numerical and Analytical Methods in Geomechanics, 38(12), 1211–1230.

8. Al-Bittar, T., and Soubra, A. H. (2014b). “Probabilistic analysis of strip footings resting on spatially varying soils and subjected to vertical or inclined loads.” Journal of Geotechnical and Geoenvironmental Engineering, ASCE 140(4).

9. Al-Bittar, T., Ahmed, A., Soubra, A. H., and Thajeel, J. (2017). “Probabilistic Analysis of Strip Footings Resting on Spatially Varying Soils Using Importance Sampling and Kriging Metamodeling.” Proceedings of the 6th International Symposium on Geotechnical Safety and Risk (ISGSR). Denver, Colorado, 5–10 June 2017.

10. Alonso, E. E., and Krizek, R. J. (1975). “Stochastic formulation of soil properties.” Proceedings of the 2nd International Conference on Applications of Statistics and Probability in Soil and Structural Engineering, 9–32.

11. Anderson, O. D. (1976). “Time series analysis and forecasting: The box-Jenkins approach.” Butterworths London, p.182.

12. Ang H-S, A., and Tang, H. W. (1975). “Probability concepts in engineering planning and design.” Vol. 1: Basic principles. John Wiley & Sons, New York.

13. Asaoka, A., and Grivas. (1982). “Spatial variability of the undrained strength of clays.” International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts, vol. 108, Pp.743–756.

14. Assimaki, D., Pecker, A., Popescu, R., and Prevost, J. (2003). “Effects of Spatial Variability of Soil Properties on Surface Ground Motion.” Journal of Earthquake Engineering, 7(sup001), 1–44.

15. Au, S. K., and Beck, J. L. (2001). “Estimation of small failure probabilities in high dimensions by subset simulation.” Probabilistic Engineering Mechanics, 16(4), 263–277.

16. Au, S. K., and Beck, J. L. (2003). “Subset simulation and its application to seismic risk based on dynamic analysis.” Journal of Engineering Mechanics, American Society of Civil Engineers, 129(8), 901–917.

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

18. Au, S. K., Ching, J., and Beck, J. L. (2007). “Application of subset simulation methods to reliability benchmark problems.” Structural Safety, 29(3), 183–193.

19. Babu, G. L. S., and Dasaka, S. M. (2008). “The effect of spatial correlation of cone tip resistance on the bearing capacity of shallow foundations.” Geotechnical and Geological Engineering, Springer, 26(1), 37–46.

20. Baecher, G. B., Marr, W. A., Lin, J. S., and Consla, J. (1983). “Critical parameters for mine tailings embankments.” Denver, CO, US Bureau of Mines.

21. Baecher, G. B., and Christian, J. T. (2003). “Reliability and statistics in geotechnical engineering.” John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichewter, West Sussex PO19 8SQ, England, ASBN 0-471-49833-5. 205-242.

22. Balesdent, M., Morio, J., and Marzat, J. (2013). “Kriging-based adaptive Importance Sampling algorithms for rare event estimation.” Structural Safety, 44, 1–10.

23. Barton, R. R. (1992). “Metamodels for simulation input-output relations.” Proceedings of the 24th conference on Winter simulation, 289–299. ACM.

24. Bergado, D. T., and Anderson, L. R. (1985). “Stochastic analysis of pore pressure uncertainty for the probabilistic assessment of the safety of earth slopes.” Soils and foundations, The Japanese Geotechnical Society, 25(2), 87–105.

25. Bergado, D. T., Chai, J. C., Alfaro, M. C., and Balasubramaniam, A. S. (1994). “Improvement techniques of soft ground in subsiding and lowland environment.” CRC Press

26. Berveiller, M., Sudret, B., and Lemaire, M. (2006). “Stochastic finite element: a non intrusive approach by regression.” European Journal of Computational Mechanics,15(1–3), 81–92.

27. Bichon, B. J., Eldred, M. S., Swiler, L. P., Mahadevan, S., and McFarland, J. M. (2008). “Efficient Global Reliability Analysis for Nonlinear Implicit Performance Functions.” AIAA Journal, 46(10), 2459–2468.

28. Blatman, G., (2009). “Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis.” Ph.D thesis, Université Blaise Pascal, Clermont II Ferrand, France.

29. Blatman, G., and Sudret, B. (2010). “An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis.” Probabilistic Engineering Mechanics, Elsevier Ltd, 25(2), 183–197.

30. Booker, A. J., Dennis Jr, J. E., Frank, P. D., Serafini, D. B., Torczon, V., and Trosset, M. W. (1999). “A rigorous framework for optimization of expensive functions by surrogates.” Structural optimization, Springer, 17(1), 1–13.

31. Borja, R. I. (2011). “Multiscale and Multiphysics Processes in Geomechanics.” Springer.

32. Bourinet, J.-M., Deheeger, F., and Lemaire, M. (2011). “Assessing small failure probabilities by combined subset simulation and support vector machines.” *Structural Safety*, Elsevier, 33(6), 343–353.

33. Box, G. E., and Jenkins, G. M. (1970). “Time series analysis: forecasting and control.” Holden-Day, San Fransisco.

34. Box, G. E. P., Hunter, W. G., and Hunter, J. S. (1978). “Statistics for experimenters: an introduction to design, data analysis, and model building.” JSTOR.

35. Brooker, P. I., Winchester, J. P., and Adams, A. C. (1995). “A geostatistical study of soil data from an irrigated vineyard near Waikerie, South Australia.” Environment international, Elsevier, 21(5), 699–704.

36. Bucher, C. G., and Bourgund, U. (1990). “A fast and efficient response surface approach for structural reliability problems.” Structural safety, Elsevier, 7(1), 57–66.

37. Burrough, P. A. (1983). “Multiscale sources of spatial variation in soil. I. The application of fractal concepts to nested levels of soil variation.” European Journal of Soil Science, Wiley Online Library, 34(3), 577–597.

38. Cafaro, F., and Cherubini, C. (2002). “Large sample spacing in evaluation of vertical strength variability of clayey soil.” Journal of geotechnical and geoenvironmental engineering, American Society of Civil Engineers, 128(7), 558–568.

39. Chen, Q., Seifried, A., Andrade, J. E., and Baker, J. W. (2012). “Characterization of random fields and their impact on the mechanics of geosystems at multiple scales.” International Journal for Numerical and Analytical Methods in Geomechanics, Wiley Online Library, 36(2), 140–165.

40. Cherubini, C. (2000). “Reliability evaluation of shallow foundation bearing capacity on c’, φ’ soils.” Canadian Geotechnical Journal, NRC Research Press, 37(1), 264–269.

41. Ching, J., Hu, Y. G., Yang, Z. Y., Shiau, J. Q., Chen, J. C., and Li, Y. S. (2011). “Reliability-based design for allowable bearing capacity of footings on rock masses by considering angle of distortion.” International Journal of Rock Mechanics and Mining Sciences, 48(5), 728–740.

42. Cho, S. E., and Park, H. C. (2010). “Effect of spatial variability of cross-correlated soil properties on bearing capacity of strip footing.” International Journal for Numerical and Analytical Methods in Geomechanics, 34(1), 1–26.

43. Christakos, G. (2012). “Random field models in earth sciences.” Courier Corporation.

44. Davis, J.C. (1986). “Statistics and data analysis in geology.” 2nd ed., John Wiley and Sons, New York, p.646.

45. Der Kiureghian, A., and Ke, J.-B. (1988). “The stochastic finite element method in structural reliability.” Probabilistic Engineering Mechanics, 3(2): 83-91.

46. Der Kiureghian, A., and Dakessian, T. (1998). “Multiple design points in first and second-order reliability.” Structural Safety, Elsevier, 20(1), 37–49.

47. Der Kiureghian, A., and Ditlevsen, O. (2009). “Aleatory or epistemic? Does it matter?” Structural Safety, Elsevier, 31(2), 105–112.

48. Diaz Padilla, J., and Vanmarcke, E. (1974). “Settlement of structures on shallow foundations.” Dept. of Civil Engrg. Cambridge: MIT.

49. Ditlevsen, O. (1996). “Dimension reduction and discretization in stochastic problems by regression method.” in F. Casciati and B. Roberts, Editors, Mathematical models for structural reliability analysis, CRC Mathematical Modelling Series, 2, 51-138.

50. Duncan, J. M. (2000). “Factors of safety and reliability in geotechnical engineering.” Journal of geotechnical and geoenvironmental engineering, American Society of Civil Engineers, 126(4), 307–316.

51. Dyminski, Andrea S and Howie, J and Shuttle, D and Kormann, A. (2006). “Variability Analysis of SPT and CPT data for a Reliability-based Embankment Design of a Southern Brazilian Port Site.” Proceedings of XIII Brazilian Conference on Soil Mechanics and Geotechnical Engineering.

52. Dyminski, A. S., Howie, J., Shuttle, D., and Kormann, A. (2006). “Soil variability study for embankment design of port of navegantes, brazil.” GeoCongress 2006: Geotechnical Engineering in the Information Technology Age, 1–6.

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

54. Echard, B., Gayton, N., Lemaire, M., and Relun, N. (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.

55. El-Ramly, H., Morgenstern, N. R., and Cruden, D. M. (2003). “Probabilistic stability analysis of a tailings dyke on presheared clay shale.” Canadian Geotechnical Journal, 40(2002), 192–208.

56. El Gonnouni, M., Riou, Y., and Hicher, P. Y. (2005). “Geostatistical method for analysing soil displacement from underground urban construction.” Géotechnique, 55(2), 171–182.

57. Fardis, M. N., and Veneziano, D. (1981). “Estimation of SPT-N and relative density.” Journal of Geotechnical and Geoenvironmental Engineering, 107(ASCE 16590 Proceeding).

58. Fenton, G. a. (1999). “Estimation for Stochastic Soil Models.” Journal of Geotechnical and Geoenvironmental Engineering, 125(6), 470–485.

59. Fenton, G. a., and Griffiths, D. V. (2001). “Bearing capacity of spatially random soil: the undrained clay Prandtl problem revisited.” Géotechnique, 51(4), 351–359.

60. Fenton, G.A., and Griffiths, D.V. (2002), “Probabilistic foundation settlement on spatially random soil.” Journal of Geotechnical and Geoenvironmental Engineering, ASCE, 128(5): 381-390.

61. Fenton, G. A., and Griffiths, D. V. (2003). “Bearing-capacity prediction of spatially random c - φ soils.” Canadian geotechnical journal, 40(1), 54–65.

62. Fenton, G.A., and Griffiths, D.V. (2005). “Three-dimensional probabilistic foundation settlement.” Journal of Geotechnical and Geoenvironmental Engineering, ASCE, 131(2): 232-239.

63. Fenton, G., and Vanmarcke, E. (1990). “Simulation of random fields via local average subdivision.” Journal of Engineering Mechanics, 116(8), pp.1733-1749.

64. Fenton, G.A., Zhou, H., Jaksa, M.B., and Griffiths, D.V. (2003). “Reliability analysis of a strip footing designed against settlement.” Applications of statistics and probability in civil engineering, Der Kiureghian, Madanat & Pestana (eds) © 2003 Millpress, Rotterdam, ISBN 90 5966 004 8.

65. Firouzianbandpey, S., Griffiths, D. V, Ibsen, L. B., and Andersen, L. V. (2014). “Spatial correlation length of normalized cone data in sand: case study in the north of Denmark.” Canadian Geotechnical Journal, NRC Research Press, 51(8), 844–857.

66. Fredlund, D. G., and Dahlman, A. E. (1972). “Statistical geotechnical properties of glacial lake Edmonton sediments.” *Statistics and Probability in Civil Engineering*, Oxford University Press: London.

67. Griffiths, D. V, Fenton, G. a., and Manoharan, N. (2002). “Bearing Capacity of Rough Rigid Strip Footing on Cohesive Soil: Probabilistic Study.” Journal of Geotechnical and Geoenvironmental Engineering, 128(9), 743–755.

68. Haldar, A., and Mahadevan, S. (2000). “Probability, reliability, and statistical methods in engineering design.” John Wiley and Sons, Inc., New York

69. Haldar, A., and Tang, W. H. (1979). “Probabilistic evaluation of liquefaction potential.” Journal of Geotechnical Engineering, ASCE , 105 (2), 145-163.

70. Harr, M. E. (1987). “Reliability-based design in civil engineering.” *McGraw-Hill Book Company*, New York, p. 290.

71. Hasofer, A. M., and Lind, N. C. (1974). “Exact and invariant second-moment code format.” Journal of the Engineering Mechanics division, ASCE, 100(1), 111–121.

72. Hicks M. A., Samy K. (2002a). “Influnce of heterogeneity on undrained clay slope stability.” Quarterly Journal of Engineering Geology and Hydrogeology, vol. XXXV, pp. 41–49.

73. Hicks M. A., Samy K. (2002b). “Influnce of anisotropic spatail variability on slope reliability.” Proceedings of the 8th International symposium on Numerical Models in Geomechanics, Rome, Italy, pp. 535-539.

74. Hicks M. A., Samy K. (2004). “Stochastic evaluation of hetergeneous slpe stability.” Italian geotechnical Journal, vol. 38(2), pp. 54–66.

75. Honjo, Y., and Kuroda, K. (1991). “A new look at fluctuating geotechnical data for reliability design.” Soils and Foundations, The Japanese Geotechnical Society, 31(1), 110–120.

76. Huang, S. P., Liang, B., and Phoon, K. K. (2009). “Geotechnical probabilistic analysis by collocation-based stochastic response surface method: An Excel add-in implementation.” Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 3(2), 75–86.

77. Huang, X., Chen, J., and Zhu, H. (2016). “Assessing small failure probabilities by AK-SS: An active learning method combining Kriging and Subset Simulation.” Structural Safety, 59, pp.86-95.

78. Hurtado, J. E. (2004). “An examination of methods for approximating implicit limit state functions from the viewpoint of statistical learning theory.” Structural Safety, 26(3), 271–293.

79. Ibrahim, Y. (1991). “Observations on applications of importance sampling in structural reliability analysis.” Structural Safety, Elsevier, 9(4), 269–281.

80. Isukapalli, S. S., Roy, A., and Georgopoulos, P. G. (1998). “Stochastic response surface methods (SRSMs) for uncertainty propagation: Application to environmental and biological systems.” Risk Analysis, 18(3), 357–363.

81. Jaksa, M. B. (1995). “The influence of spatial variability on the geotechnical design properties of a stiff, overconsolidated clay.” Ph.D. Thesis, Faculty of Engineering, University of Adelaide.

82. Jaksa, M. B., Brooker, P. I., and Kaggwa, W. S. (1997). “Modelling the spatial variability of the undrained shear strength of clay soils using geostatistics.” Proc. of 5th Int. Geostatistics Congress, 1284–1295.

83. Jaksa, M. B., Kaggwa, W. S., and Brooker, P. I. (1999). “Experimental evaluation of the scale of fluctuation of a stiff clay.” Proc. 8th Int. Conf. on the Application of Statistics and Probability, 415–422.

84. Journel, A. G. and Ch J. Huijbregts. (1978). “Mining Geostatistics.” Academic Press.

85. Kaymaz, I. (2005). “Application of kriging method to structural reliability problems.” Structural Safety, 27(2), 133–151.

86. Keaveny, J. M., F. Nadim, and S. Lacasse. (1989). “Autocorrelation Functions for Offshore Geotechnical Data.” Structural Safety and Reliability, pp.263–270.

87. Kulhawy, F. H., M. J. S. Roth, and M. D. Grigoriu. (1991). “Some Statistical Evaluations of Geotechnical Properties.” in Proc. ICASP6, 6th Int. Conf. Appl. Stats. Prob. Civ. Eng.

88. Kulhawy, Fred H., B. Birgisson, and M. D. Grigoriu. (1992). “Reliability-Based Foundation Design for Transmission Line Structures.” (No. EPRI-EL-5507-Vol. 4). Electric Power Research Inst., Palo Alto, CA (United States); Cornell Univ., Ithaca, NY (United States). Geotechnical Engineering Group

89. Lacasse, S., and Nadim, F. (1997). “Uncertainties in characterising soil properties.” Publikasjon-Norges Geotekniske Institutt, Norwegian Geotechnical Institute, 201, 49–75.

90. Li, C. C., and Der Kiureghian, A. (1993). “Optimal discretization of random fields.” J. Eng. Mech.-ASCE, 119(6), 1136–1154.

91. Li, J. L., Cassidy, M. J., and Tian, Y. (2015). “Failure Mechanism and Bearing Capacity of Footings Buried at Various Depths in Spatially Random Soil.” Journal of Geotechnical and Geoenvironmental Engineering, 141(2), 4014099.

92. Li, K. S. and I. K. Lee. (1991). “The Assessment of Geotechnical Safety.” Selected Topics in Geotechnical Engineering, Lumb Volume.

93. Liu, C., Chen, C. (2010). “Estimating spatial correlation structures based on CPT data.” Georisk, Taylor & Francis, 4(2), 99–108.

94. Liu, W.-K., Belytschko, T., and Mani, A. (1986a). “Probabilistic finite elements for non linear structural dynamics.” Computer Methods in Applied Mechanics and Engineering, 56: 61-86.

95. Liu, W.-K., Belytschko, T., and Mani, A. (1986b), “Random field finite elements.” International Journal for Numerical Methods in Engineering, 23(10): 1831-1845.

96. Lophaven, S. N., Nielsen, H. B., and Sondergaard, J. (2002). “DACE: A Matlab Kriging Toolbox. 2002.” Technical University of Denmark.

97. Lumb, P. (1970). “Safety factors and the probability distribution of soil strength.” Canadian Geotechnical Journal, NRC Research Press, 7(3), 225–242.

98. Lumb, P. (1972). “Precision and accuracy of soil tests.” Proceedings, 1st international conference on applications of statistics and probability in soil and structural engineering, Hong Kong, 329–345.

99. Lumb, P. (1975). “Spatial variability of soil properties.” Proc. 2nd Int. Conf. on Appl. of Statistics and Probability in Soil and Struct. Eng.(ICASP) Aachen, 397–422.

100. Mao, N., Al-bittar, T., and Soubra, A. (2012). “Probabilistic Analysis and Design of Strip Foundations Resting on Rocks Obeying Hoek – Brown Failure Criterion.” International Journal of Rock Mechanics and Mining Sciences 49:45–58.

101. Marelli, S., Schöbi, R., and Sudret, B. (2016). "UQLAB User Manual – Structural Reliability (Rare Events Estimation)." Structural Reliability, Report UQLab-V0.9-107, Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich, 2015.

102. Matsuo, M., and Asaoka, A. (1977). "Probability models of undrained strength of marine clay layer." Soils and Foundations, The Japanese Geotechnical Society, 17(3), 53–68.

103. Matthies, G., Brenner, C., Bucher, C., and Guedes Soares, C., (1997). "Uncertainties in probabilistic numerical analysis of structures and solids – stochastic finite elements." Struct. Safety, 19, 3, 283-336.

104. Melchers, R. E. (1999). "Structural Reliability: Analysis and Prediction." *Ellis Horwood Ltd.*, Chichester, U. K., 437p.

105. Myers, R. H., and Montgomery, D. C. (1995). "Response Surface Methodology: Process and Product Optimization Using Designed Experiments", John Wiley & Sons, New York.

106. Nataf, A. (1962). "Détermination Des Distributions de Probabilités Dont Les Marges Sont Données." Comptes Rendus de l'Académie des Sciences 225:42–43.

107. Nour, A., Slimani, A., and Laouami, N. (2002). "Foundation settlement statistics via finite element analysis." Computers and Geotechnics, Elsevier, 29(8), 641–672.

108. Papadopoulos, V., Giovanis, D. G., Lagaros, N. D., and Papadrakakis, M. (2012). "Accelerated subset simulation with neural networks for reliability analysis." Computer Methods in Applied Mechanics and Engineering, Elsevier, 223, 70–80.

109. Papadrakakis, M., and Lagaros, N. D. (2002). "Reliability-based structural optimization using neural networks and Monte Carlo simulation." Computer Methods in Applied Mechanics and Engineering, 191(32), 3491–3507.

110. Phoon, K. K., and Kulhawy, F. H. (1996). "On quantifying inherent soil variability." Uncertainty in the geologic environment: From theory to practice, 326–340.

111. Phoon, K.-K., and Kulhawy, F. H. (1999). "Evaluation of geotechnical property variability." Canadian Geotechnical Journal, 36(4), 625–639.

112. Popescu, R., Deodatis, G., and Nobahar, A. (2005). "Effects of random heterogeneity of soil properties on bearing capacity." Probabilistic Engineering Mechanics, 20(4), 324–341.

113. Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989). "Design and Analysis of Computer Experiments." Statistical Science, 409–423.

114. Santner, T. J., Williams, B. J., and Notz, W. I. (2003). "The Design and Analysis of Computer Experiments." New York: Springer-Verlag.

115. Schöbi, R., Sudret, B., and Marelli, S. (2015). "Rare Event Estimation Using Polynomial-Chaos Kriging." ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, American Society of Civil Engineers, D4016002.

116. Schuëller, G. I., Pradlwarter, H. J., and Koutsourelakis, P. S. (2004). "A critical appraisal of reliability estimation procedures for high dimensions." Probabilistic Engineering Mechanics, 19(4), 463–474.

117. Schuëller, G. I., and Stix, R. (1987). "A critical appraisal of methods to determine failure probabilities." Structural Safety, 4(4), 293–309.

118. Schueremans, L., and Van Gemert, D. (2005). "Benefit of splines and neural networks in simulation based structural reliability analysis." Structural safety, Elsevier, 27(3), 246–261.

119. Spanos, P. D., and Ghanem, R. (1989). "Stochastic finite element expansion for random media." Journal of Engineering Mechanics, ASCE, 115(5), 1035–1053.

120. Srivastava, A., and Babu, G. L. S. (2009). "Effect of soil variability on the bearing capacity of clay and in slope stability problems." Engineering Geology, Elsevier, 108(1), 142–152.

121. Suchomel, R., and Mašín, D. (2010). "Spatial variability of soil parameters in an analysis of a strip footing using hypoplastic model." Proceedings of 7th European conference on numerical methods in geomechanic, 383–388.

122. Sudret, B., and Berveiller, M. (2008). "Stochastic finite element methods in geotechnical engineering." Reliability-based design in geotechnical engineering: computations and applications. Taylor & Francis, 260–296.

123. Sudret, B., Berveiller, M., and Lemaire, M. (2006). "A stochastic finite element procedure for moment and reliability analysis." European Journal of Computational Mechanics, 15(7–8), 825–866.

124. Sudret, B., and Der Kiureghian, A. (2000). "Stochastic finite element methods and reliability: A State-of-the-Art Report." Report No. UCB/SEMM-2000/08 Structural Engineering, Mechanics and Materials Department of Civil & Environmental Engineering, University of California, Berkeley

125. Tandjiria, V., Teh, C. I., and Low, B. K. (2000). "Reliability analysis of laterally loaded piles using response surface methods." Structural Safety, 22(4), 335–355.

126. Tang, W. H. (1979). "Probabilistic evaluation of penetration resistances." Journal of Geotechnical and Geoenvironmental Engineering, 105(GT10), 1173-1191.

127. Tang, W. H. (1984). "Principles of probabilistic characterizations of soil properties." Probabilistic characterization of soil properties: Bridge between theory and practice, 74–89.

128. Thajeel, J., AL-Bittar, T., Ahmed, A., and Soubra, A. H., (2015). "Probabilistic Analysis at Ultimate Limit State of Strip Footings Resting on a Spatially Varying Soil Using Subset Simulation Approach." Proceedings of the 5th International Symposium on Geotechnical Safety and Risk (ISGSR 2015). Netherlands, Rotterdam, 13-16 October. *V*, IOS Press 326-331

129. Thajeel, J., AL-Bittar, T., Issa, N., and Soubra, A. H., (2016). "Bearing Capacity of Strip Footing on Spatially Random Soils Using Kriging and Monte Carlo Simulation."

Proceedings of the VII European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS Congress 2016). Crete, Greece, 5-10 June 2016. vol. 3, 6260-6271

130. Tillmann, A., Englert, A., Nyári, Z., Fejes, I., Vanderborght, J., and Vereecken, H. (2008). "Characterization of subsoil heterogeneity, estimation of grain size distribution and hydraulic conductivity at the Krauthausen test site using cone penetration test." Journal of Contaminant Hydrology, Elsevier, 95(1), 57–75.

131. Uzielli, M., Vannucchi, G., and Phoon, K. K. (2005). "Random field characterisation of stress-normalised cone penetration testing parameters." Geotechnique, London: Geotechnical Society, 1948, 55(1), 3–20.

132. Vanmarcke, E. H. (1977). "Probabilistic modeling of soil profiles." Journal of the Geotechnical Engineering Division, ASCE, 103(11), 1227–1246.

133. Vanmarcke, E. H. (1978). "Probabilistic characterization of soil profiles." Site Characterization & Exploration, 199–219.

134. Vanmarcke, E. (1983). "Random fields: Analysis and synthesis." M.I.T, Press, Cambridge, Mass., 382 p.

135. Vanmarcke, E., and Grigoriu, M., (1983). "Stochastic finite analysis of simple beams." Journal of Engineering Mechanics, 109: 1203-1214.

136. Vesic, A. S. (1973). "Analysis of Ultimate Loads of Shallow Foundations." Journal of Soil Mechanics & Foundations Division, 99(1), 45–73.

137. Vořechovský, M. (2008). "Simulation of simply cross correlated random fields by series expansion methods." Structural Safety, 30(4), 337–363.

138. Wolff, T. F. (1985). "Analysis and design of embankment dam slopes: A probabilistic approach." Purdue University.

139. Xiu, D., and Karniadakis, G. E. (2002). "The Wiener--Askey polynomial chaos for stochastic differential equations." SIAM journal on scientific computing, SIAM, 24(2), 619–644.

140. Yucemen, M. S., Tang, W. H., and Ang, A.-S. (1973). "A probabilistic study of safety and design of earth slopes." University of Illinois Engineering Experiment Station. College of Engineering. University of Illinois at Urbana-Champaign.

141. Zhang, L., and Chen, J.-J. (2012). "Effect of spatial correlation of standard penetration test (SPT) data on bearing capacity of driven piles in sand." Canadian Geotechnical Journal, NRC Research Press, 49(4), 394–402.

142. Zhang, J., and Ellingwood, B. (1994). "Orthogonal series expansion of random fields in reliability analysis." Journal of Engineering Mechanics, ASCE, 120(12): 2660-2677.

143. Zhou, C. ., Lu, Z. ., and Yuan, X. . (2013). "Use of relevance vector machine in structural reliability analysis." Journal of Aircraft, 50(6), 1726–1733.

APPENDIX A

Random field discretization methods

The most commonly used methods of random field discretization in geotechnical engineering can be divided into three main categories [Sudret and Kiureghian (2000)]. Each category involves a number of discretization methods as may be seen below.

A.1. Point discretization methods

In these methods, the random variables used in the analysis are selected values of Z at some given points X_j. This group involves the following methods:

a) Midpoint (MP) method

This method was introduced by Der Kiureghian and Ke (1988). In this method, the random field is discretized by associating to each element of the finite element/finite difference mesh, a single random variable defined as the value of the field at the centroid of that element.

b) Shape Function (SF) method

This method was presented by Liu et al. (1986a, b). It is similar to the MP method, with the difference being that the random field is discretized by associating a single random variable to each node of the finite element/finite difference mesh. Thus, the value of the random field within an element is described in terms of these nodal values and the corresponding shape functions.

c) Integration Point (IP) method

This method was proposed by Matthies et al. (1997). In this method, the random field is discretized by associating a single random variable to each of the integration points appearing in the finite element resolution scheme.

d) Optimal Linear Estimation (OLE) method

This method was presented by Li and Der Kiureghian (1993). It is sometimes referred to as the kriging method. It is a special case of the regression method on a linear function [Ditlevsen

(1996)]. The fact of knowing the values of the soil property Z at some given points may allow one to approximate the value of Z at an arbitrary point X using the optimal linear estimation method OLE. Indeed, OLE makes use of the experimental data samples to estimate the values of a soil property at unsampled locations. It should be noted that the concepts used in OLE method were employed for the discretization of a random field by the expansion optimal linear estimation EOLE method which was used in this thesis.

A.2. Average discretization methods

a) Spatial Average (SA) method

This method was proposed by Vanmarcke and Grigoriu (1983). It consists in approximating the random field in each element of the finite element/finite difference mesh by a constant computed as the average of the original field over that element.

b) Local Average Subdivision (LAS) method

This method was developed by Fenton and Vanmarcke (1990) in order to produce data with specified statistical parameters, which are simultaneously spatially correlated. LAS theory follows a recursive fashion, where a global average is created in stage zero. In stage one, the field is divided into two equal parts for 1D (and four equal parts for 2D) whose local average equals to the parent global value. In stage two, two absolute normally distributed values are generated whose means and variances are selected so as to satisfy three criteria: (a) they show the correct variance according to local averaging theory, (b) they are properly correlated with one another, and (c) they average to the parent value. It should be mentioned here that this method is widely used by the geotechnical committee as in [Fenton and Griffiths (2002, 2005) and Fenton et al. (2003) among others].

A.3. Series expansion methods

In the series expansion discretization methods, the random field is approximated by an expansion that involves deterministic and stochastic functions. The deterministic functions depend on the

coordinates of the point at which the value of the random field is to be calculated. This group involves the following methods:

a) Karhunen-Loeve (KL) expansion method

This method was presented by Spanos and Ghanem (1989). In this method, the random field is expressed as follows:

$$\tilde{Z}(X) = \mu_Z + \sigma_Z \sum_{j=1}^{M} \sqrt{\lambda_j}\,\xi_j \phi_j(X) \tag{A.1}$$

where μ_Z and σ_Z are the mean and standard deviation of the random field *Z*, (λ_j, ϕ_j) are the eigenvalues and eigenfunctions of the autocorrelation function ρ_Z of the random field *Z*, ξ_j is a vector of uncorrelated standard normal random variables and *M* is the number of terms retained in the KL expansion. It should be noticed here that ξ_j are stochastic variables that represent the random nature of the uncertain soil parameter. However, the eigenfunctions $\phi_j(X)$ are the deterministic functions of the KL expansion. They can be evaluated as the solution of the following integral equation:

$$\int_{\Omega} \rho_Z(X, X')\, f_j(X')\, dX' = \lambda_j f_j(X) \tag{A.2}$$

This integral can be solved analytically only for few types of the autocorrelation functions (triangular and first order exponential functions) and for simple geometries. Otherwise, it has to be solved numerically.

b) Orthogonal series expansion (OSE) method

This method was proposed by Zhang and Ellingwood (1994). It was introduced to avoid solving the eigenvalue integral of Equation (A.2) using a complete set of orthogonal functions $h_j(X)$

(i.e. Legendre or Hermite polynomials). Thus, in this method, the random field is expressed as follows:

$$\tilde{Z}(X) = \mu_Z + \sigma_Z \sum_{j=1}^{M} \chi_j h_j(X) \tag{A.3}$$

where χ_j are zero mean random variables with unit variance and M is the number of terms retained in the expansion.

c) Expansion optimal linear estimation (EOLE) method

This method was introduced by Li and Der Kiureghian (1993). It makes use of the (OLE) or the kriging method concept in the special case of a Gaussian random field. This method uses a spectral representation of the autocorrelation matrix of the Gaussian random field and it is used in this thesis. Thus, it was presented in more detail.

Weight functions and deterministic basis of the MP, SF, SA and OLE methods

Table A.1 represent the weight functions and deterministic basis of the MP, SF, SA and OLE methods:

Table A.1. Weight functions and deterministic basis of the MP, SF, SA and OLE methods

Method	Weight function $\omega(X)$	Deterministic basis $\varphi_j(X)$
MP	$\delta(X - X_c)$	$1_{\Omega_e}(X)$
SF	$\delta(X - X_j)$	$1_{\Omega_e}(X)$
SA	$\frac{1_{\Omega_e}(X)}{\lvert\Omega_e\rvert}$	Polynomial shape function $N_j(X)$
OLE	$\delta(X - X_j)$	$\left(\Sigma^{-1}_{\chi;\chi}.\Sigma_{Z(X);\chi}\right)_j$

In Table A.1, X is the vector of the coordinates of an arbitrary point, X_c is the vector of the coordinates at the centroid element of the finite element/finite difference mesh, X_j is the vector of

the coordinates at a node *j* in the SF method and at a the sample point *j* in the OLE method, $\delta(.)$ denotes the Dirac function, $1_{\Omega_e} = \begin{cases} 1 & X \in \Omega_e \\ 0 & otherwise \end{cases}$ and Ω_e is the mesh element.

APPENDIX B

AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation

The AK-MCS approach as described by Echard et al. (2011) is given in Figure B.1. It consists of 10 stages:

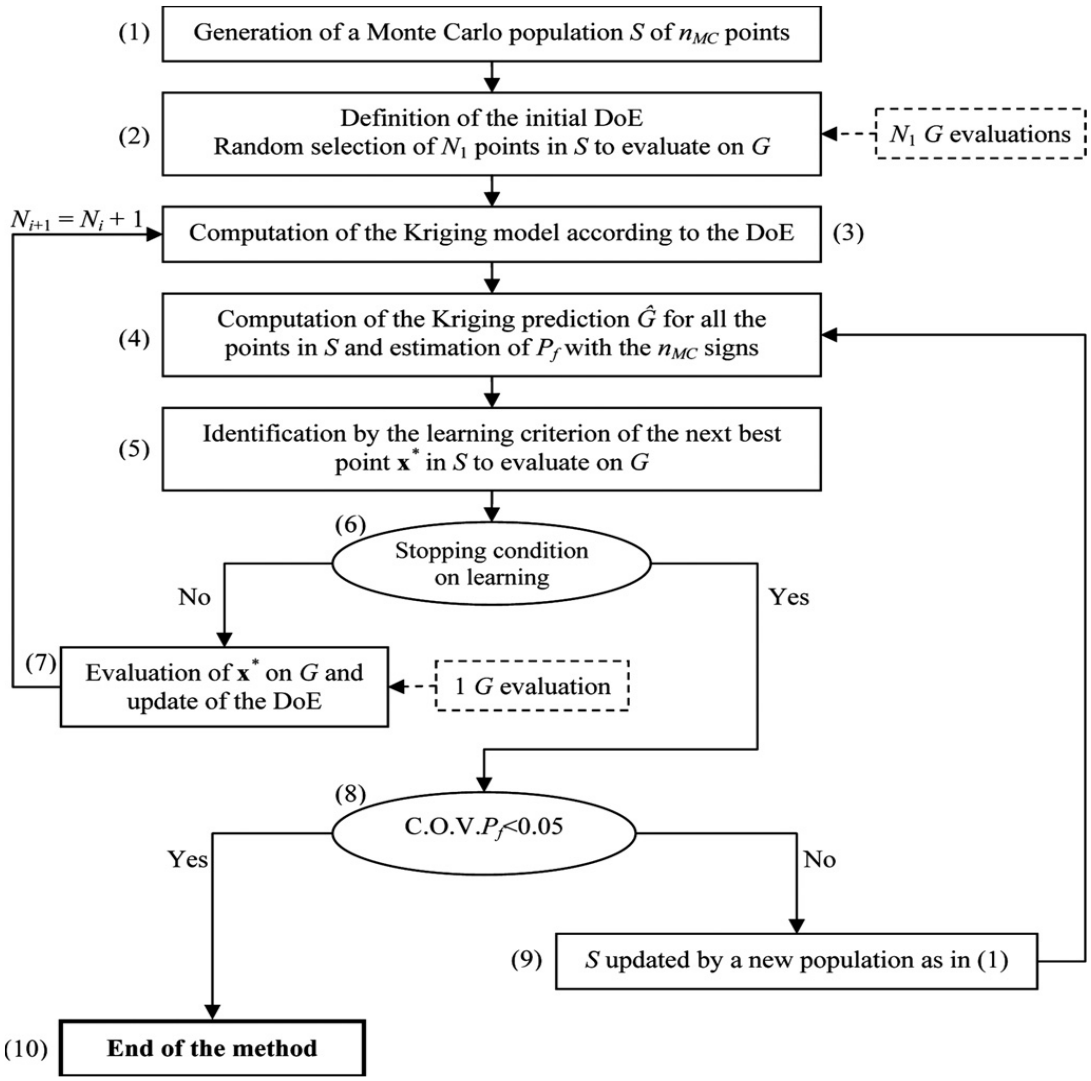


Figure B.1. AK-MCS flowchart from Echard et al. (2011)

1. Generation of a large Monte Carlo population S in the design space. This population is composed of n_{MC} samples (the terms samples and points are used interchangeably within this thesis). At this stage, none of these samples is evaluated on the real performance

function (i.e. using the mechanical model). They represent candidate samples to be evaluated by the mechanical model if the active learning requires it. This population remains the same during the whole process of learning in AK-MCS unless Stage 9 is reached.

2. Definition of the initial design of experiments (DoE): To perform kriging, a design of experiments is required. Here, it consists of a random selection of N_1 samples among the population S. These samples are evaluated on the real performance function and are used to construct the kriging meta-model. The initial design of experiments is preferred to be small at the beginning and to be successively enriched by adding each time the next 'best' sample that best improves the meta-model. The aim of this procedure is to reduce to a minimum the number of calls of the mechanical model. Notice that although a small design of experiments is preferred, a greater DoE is needed for problems involving higher stochastic dimension (i.e. where each sample is composed of more than about 10 random variables).

3. Computation of the kriging meta-model according to the design of experiments: This stage is performed using the DACE toolbox in MATLAB. It should be mentioned that ordinary kriging is used where the regression function is considered constant. As for the correlation model, a Gaussian correlation function is selected.

4. Prediction by kriging and estimation of the probability of failure: First, use DACE to compute the Kriging predictions $\mu_{\hat{G}}$ and their corresponding kriging variance values $\sigma^2_{\hat{G}}$ for the whole population S (i.e. for i = 1,. . . ,n_{MC}) making use of Equations (1.33) and (1.34). The probability of failure is then estimated with the signs of the predictions. It is obtained as the ratio of the points (or samples) in the population S with a negative or null Kriging prediction and the total number of points in S using the following equation:

$$P_f = \frac{n_{\hat{G}\leq 0}}{n_{MC}} \tag{B.1}$$

5. Identification of the best next point in S to evaluate on the real performance function: This is performed by evaluating a learning function U for each point in S. The learning function value U is obtained as the ratio of the absolute value of the prediction and its corresponding kriging variance value using the following equation:

$$U\left(X_i\right) = \frac{\left|\mu_{\hat{G}(X_i)}\right|}{\sigma_{\hat{G}(X_i)}} \qquad i = 1, \ldots, n_{MC} \tag{B.2}$$

The criterion to identify the best next point consists in choosing the point with minimum value of U.

6. Stopping condition on learning: The learning stops if the minimum value of U is greater than 2.
7. Update of the previous design of experiments with the best point: If the stopping condition given in Stage 6 is not satisfied, the learning carries on and the best point x^* in S is evaluated on the real performance function. This best point x^* is then added to the design of experiments: $N_{i+1} = N_i+1$. The method, then, goes back to Stage 3 to compute the new kriging model with the updated design of experiments composed of N_{i+1} points. This process of learning is repeated until the stopping condition is satisfied.
8. Computation of the coefficient of variation of the probability of failure: If the stopping condition given in Stage 6 is satisfied, the learning stops and the metamodel is said to be accurate enough for the computation of the estimated values of both the probability of failure P_f and the coefficient of variation $COV\left(P_f\right)$. It should be mentioned here that a value of the coefficient of variation below 5% is seen as acceptable. It is calculated based on the following equation using the final kriging metamodel:

$$COV\left(P_f\right) = \sqrt{\frac{1-P_f}{P_f \cdot n_{MC}}} \tag{B.3}$$

9. Update of the population S: If the coefficient of variation is found to be larger than 5%, the population S is increased by adding new points generated using Monte Carlo simulation (similar to Stage 1). The AK-MCS procedure goes back to Stage 4 to compute the predictions using the updated population and the active learning method carries on until the stopping condition is met again.

10. End of AK-MCS procedure: If the coefficient of variation of $COV\left(P_f\right)$ is low enough (i.e. smaller than 5%), AK-MCS stops and the last estimation of the probability of failure is considered as the result of the procedure.

APPENDIX C

Description of the software FLAC^{3D}

FLAC^{3D} (Fast Lagrangian Analysis of Continua) is a computer code which allows one to perform three dimensional (3D) numerical simulations. It should be mentioned that FLAC^{3D} allows the application of stresses (stress control method) or velocities (displacement control method) on the geotechnical system. The application of stresses or velocities creates unbalanced forces in this system. The solution of a given problem in FLAC^{3D} is obtained by damping these forces to reduce them to very small values compared to the initial ones. The stresses and strains are calculated at several time intervals (called cycles) until a steady state of static equilibrium or a steady state of plastic flow is achieved in the soil mass.

It should be mentioned here that the programming language FISH in FLAC^{3D} allows one to create functions that calculates the stresses, displacements, etc. at any point in the soil mass.

Computation of the ultimate load of a vertically loaded footing

For the computation of the ultimate footing load using FLAC^{3D}, the displacement control method was used. In this method, a small vertical velocity (5×10^{-6} m/time step in this thesis) is applied to the lower nodes of the footing and then, several cycles are run until reaching a steady state of plastic flow. The steady state of plastic flow is assumed to be reached when the two following conditions are satisfied:

- The load becomes constant with the increase in the number of cycles. In other words, increasing the number of cycles no longer changes the footing load (see Figure C.1).
- The unbalanced forces tend to a very small value as shown in Figure (C.2).

At each cycle, the footing load is obtained by using a FISH function that computes the summation of stresses of all elements of the soil-footing interface. The value of the footing load when reaching the steady state of plastic flow is the ultimate failure load of the footing.

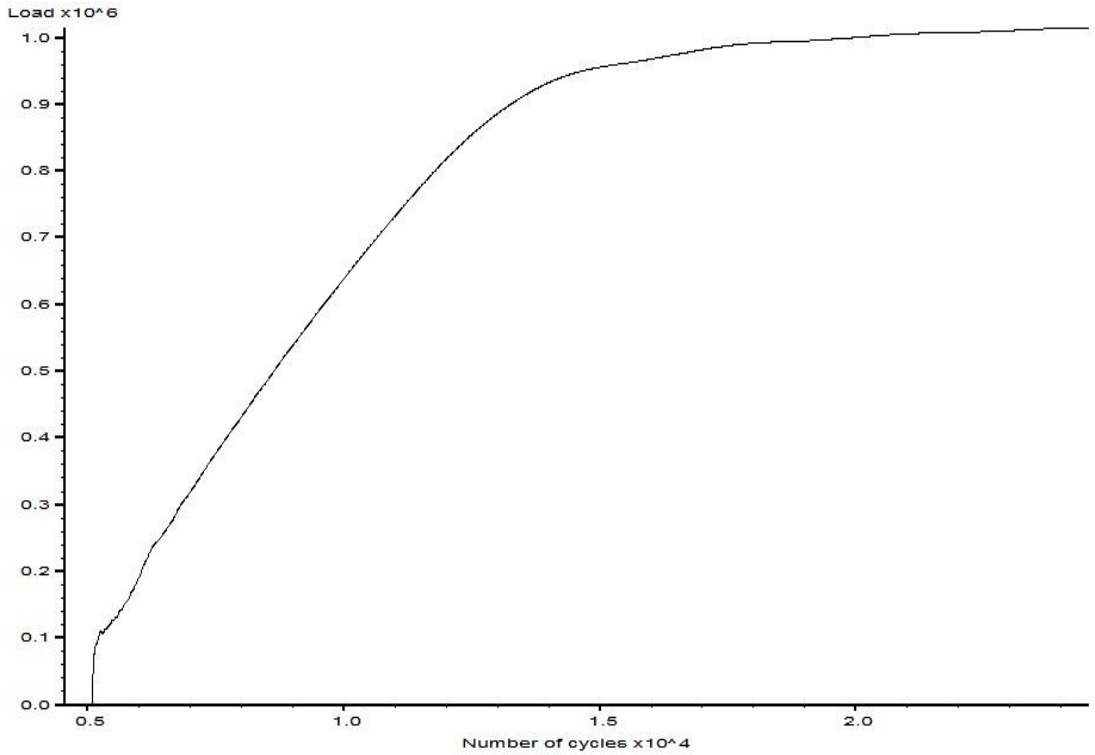


Figure C.1. Load versus the number of cycles

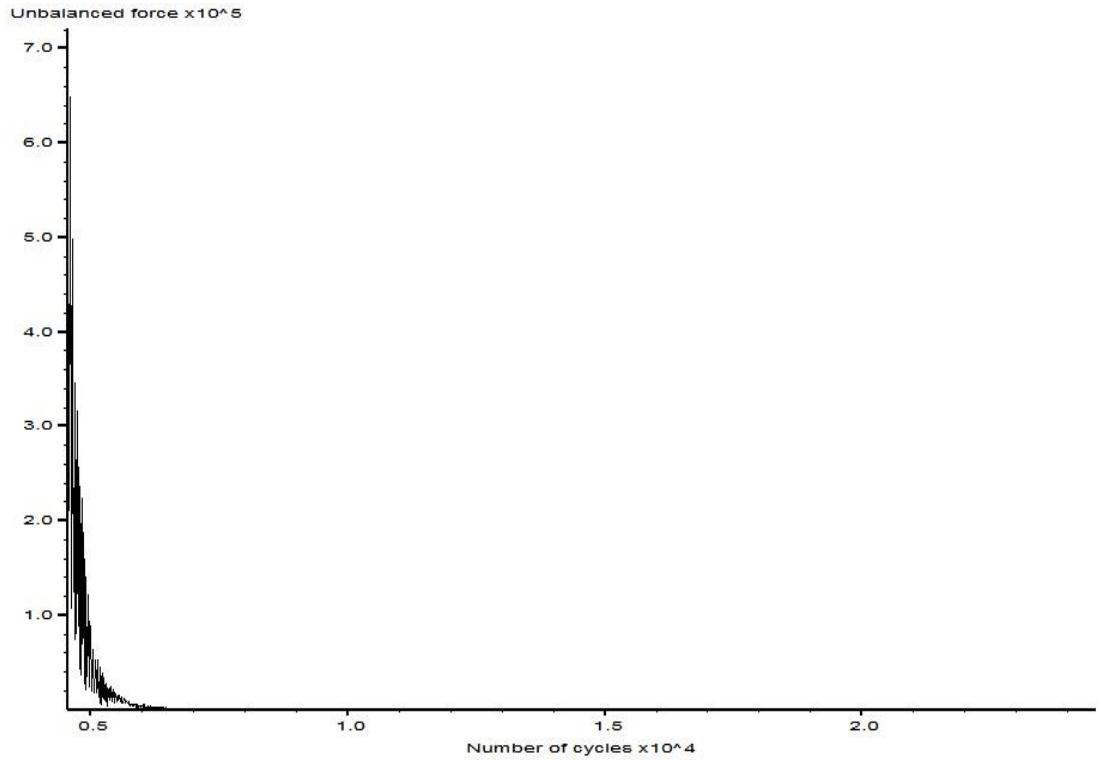


Figure C.2. Unbalanced forces versus the number of cycles

APPENDIX D

Failure probability P_f and $COV\left(P_f\right)$ versus the number of added realizations obtained using AK-MCS method

The figures below present the failure probability P_f and the coefficient of variation $COV\left(P_f\right)$ versus the number of added realizations for the different values of the autocorrelation distances (a_x, a_y) used in the analysis (see Table 3.5). These figures also provide the learning function values for the different added realizations.

Case of an isotropic soil (a_x=a_y)

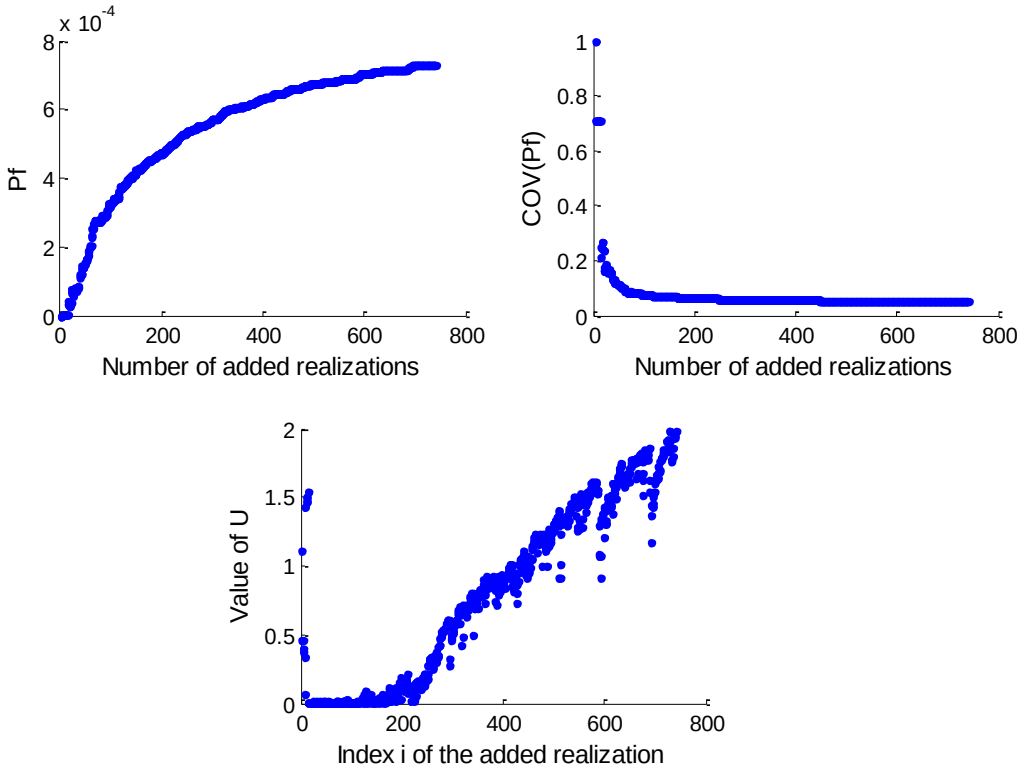


Figure D.1. AK-MCS results for a spatially varying soil for the case (a_x=a_y=2m)

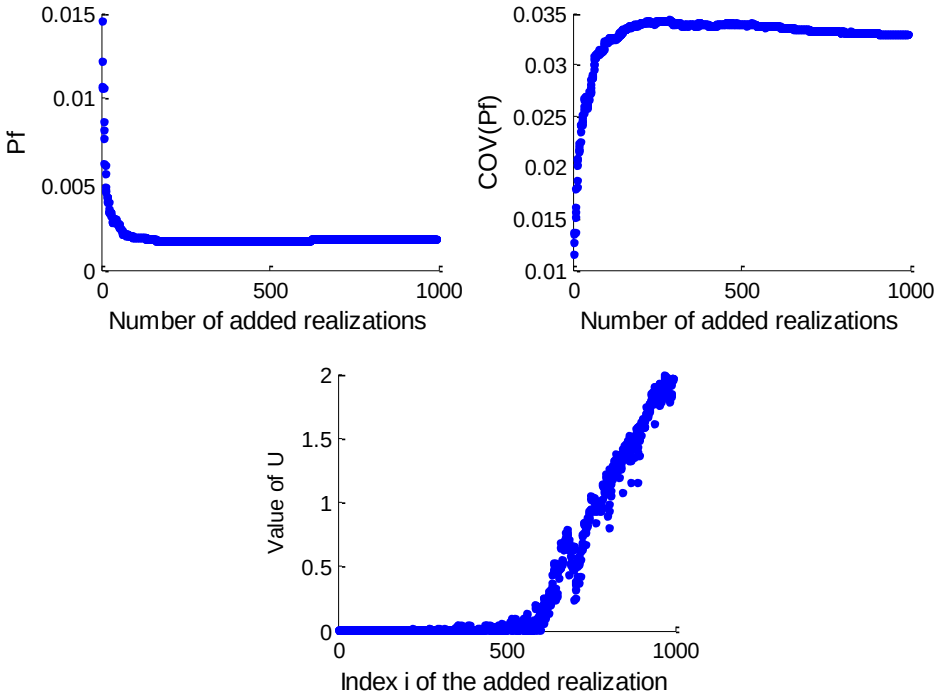


Figure D.2. AK-MCS results for a spatially varying soil for the case (a_x=a_y=3m)

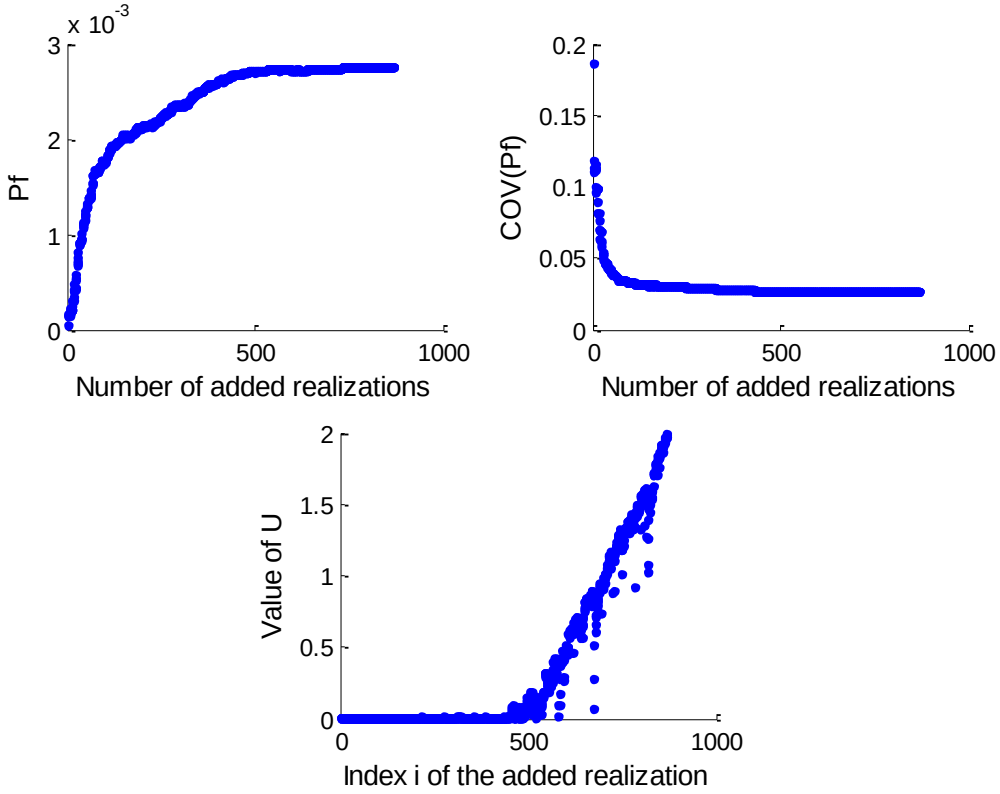


Figure D.3. AK-MCS results for a spatially varying soil for the case (a_x=a_y=5m)

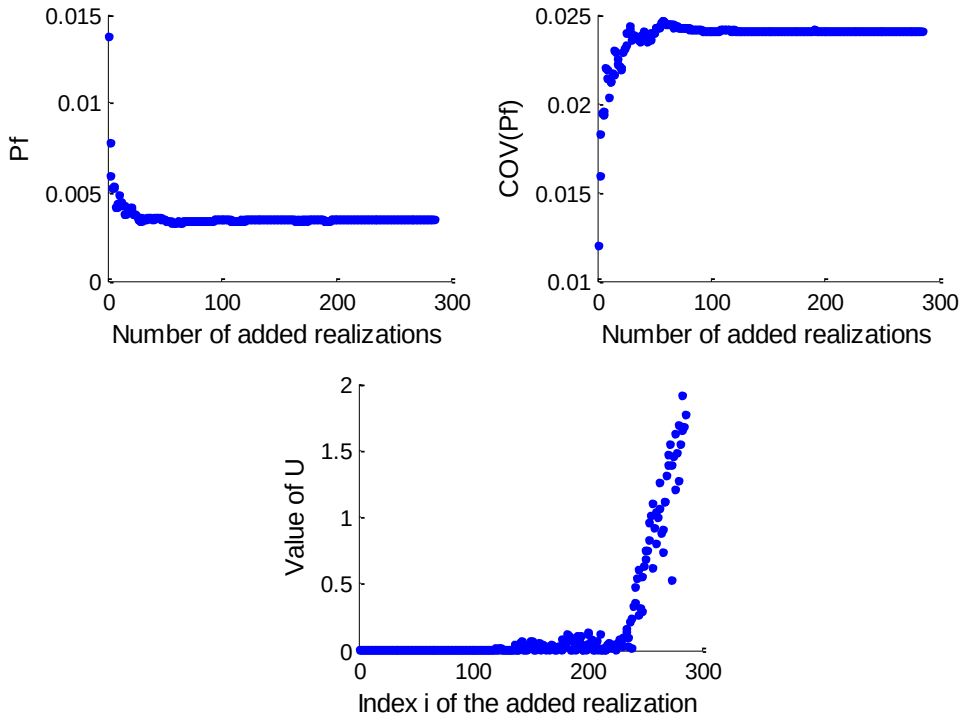


Figure D.4. AK-MCS results for a spatially varying soil for the case (a_x=a_y=10m)

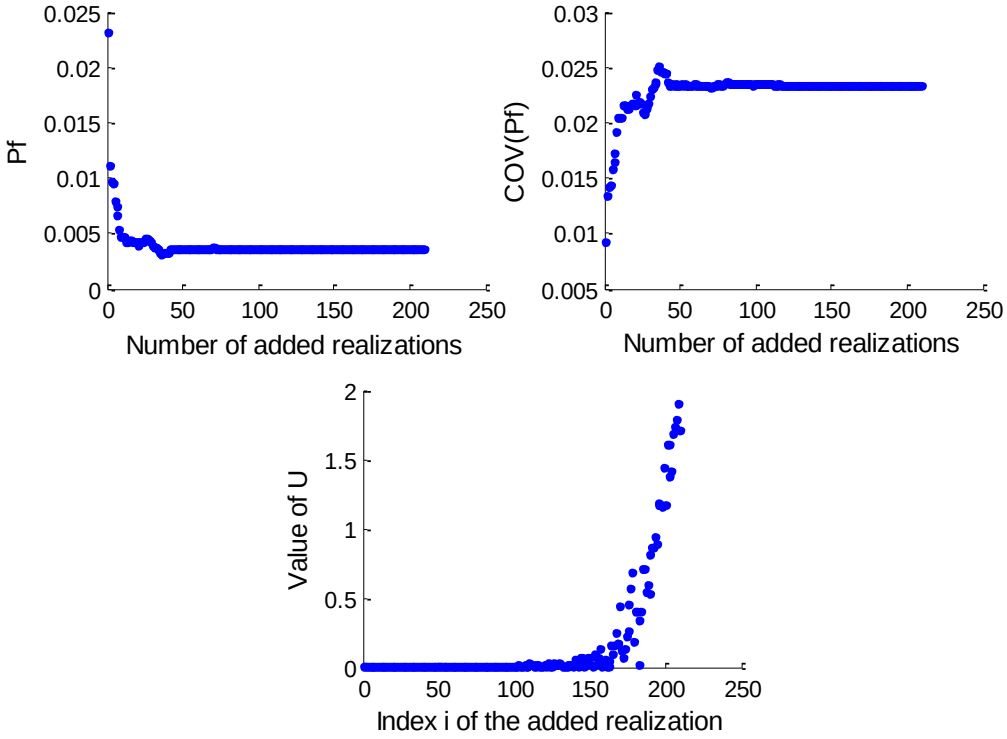


Figure D.5. AK-MCS results for a spatially varying soil for the case (a_x=a_y=20m)

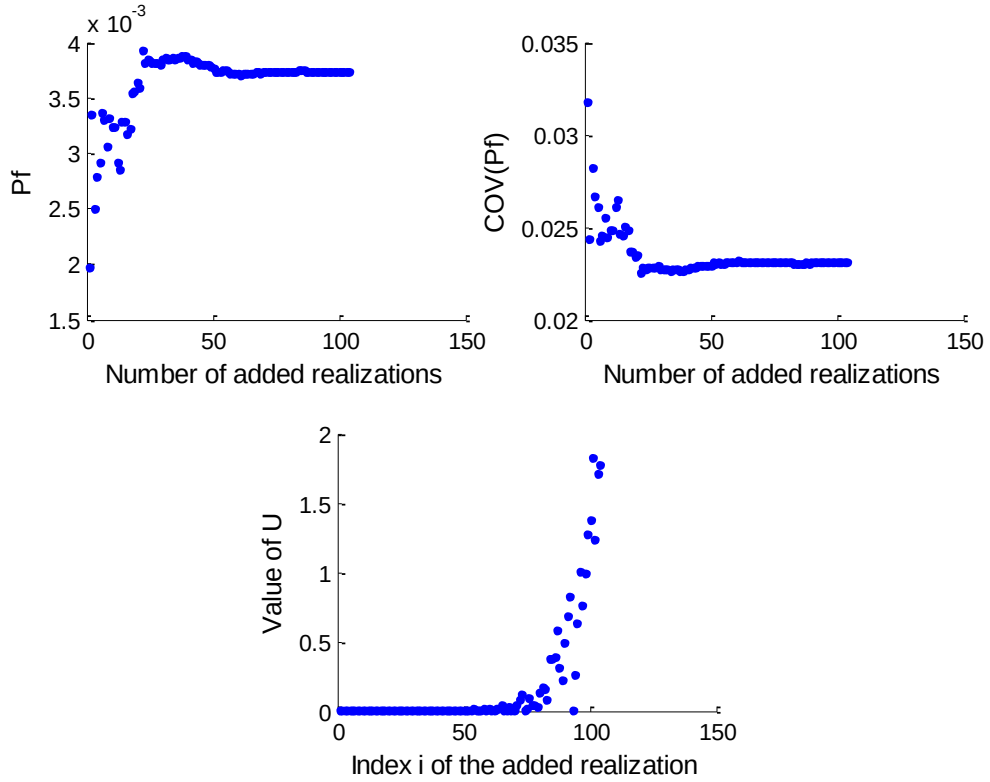


Figure D.6. AK-MCS results for a spatially varying soil for the case (a_x=a_y=50m)

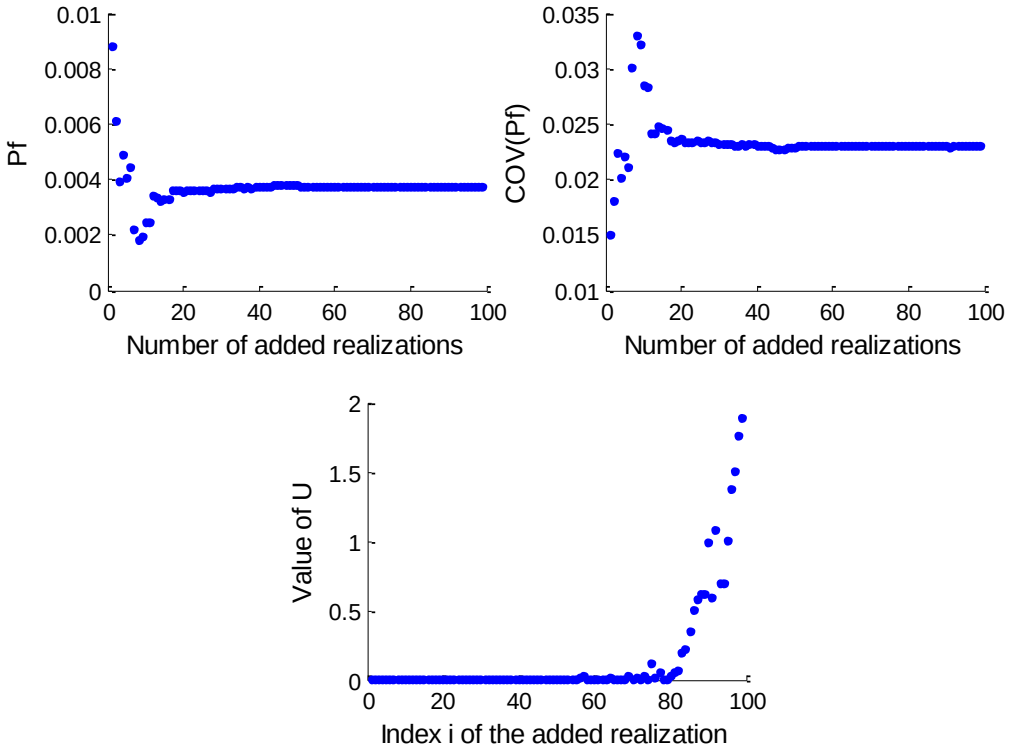


Figure D.7. AK-MCS results for a spatially varying soil for the case (a_x=a_y=100m)

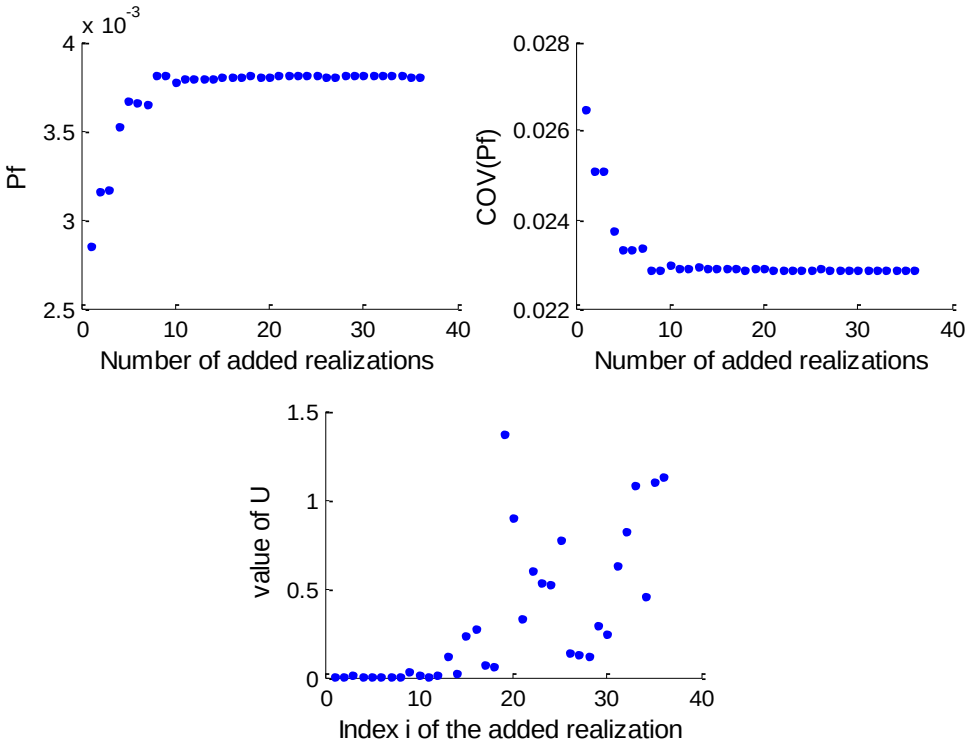


Figure D.8. AK-MCS results for a spatially varying soil for the case (a_x=a_y=10000m)

Case of an anisotropic soil (a_x=10 m with varying a_y)

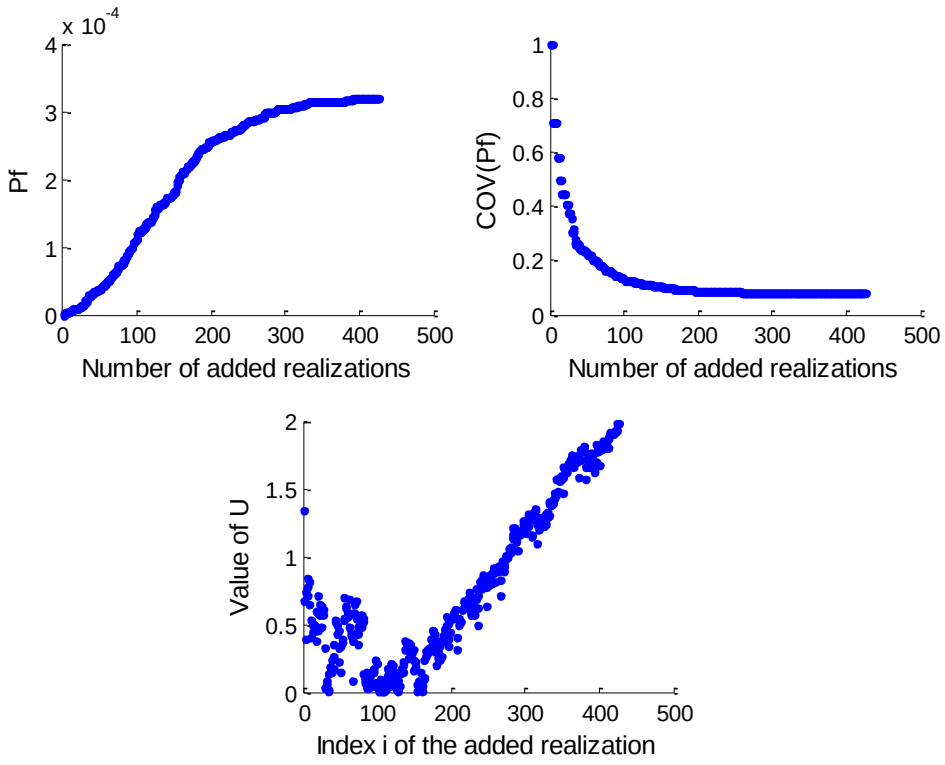


Figure D.9. AK-MCS results for a spatially varying soil for the case (a_x=10m, a_y=0.5m)

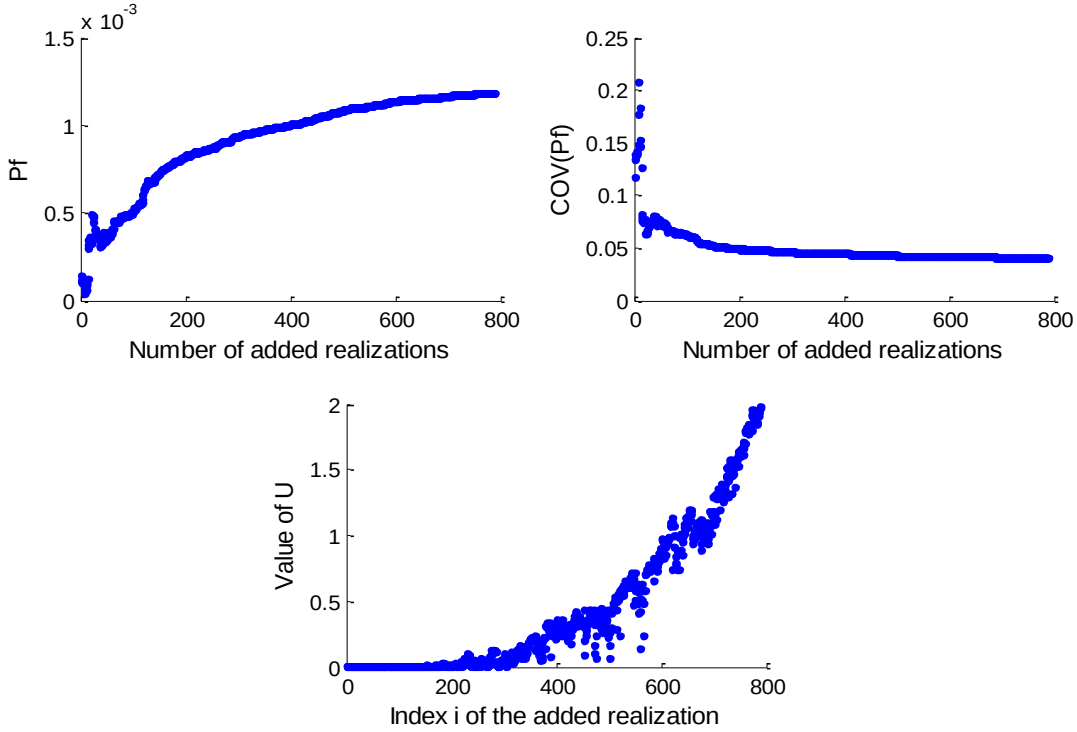


Figure D.10. AK-MCS results for a spatially varying soil for the case (a_x=10m, a_y=0.8m)

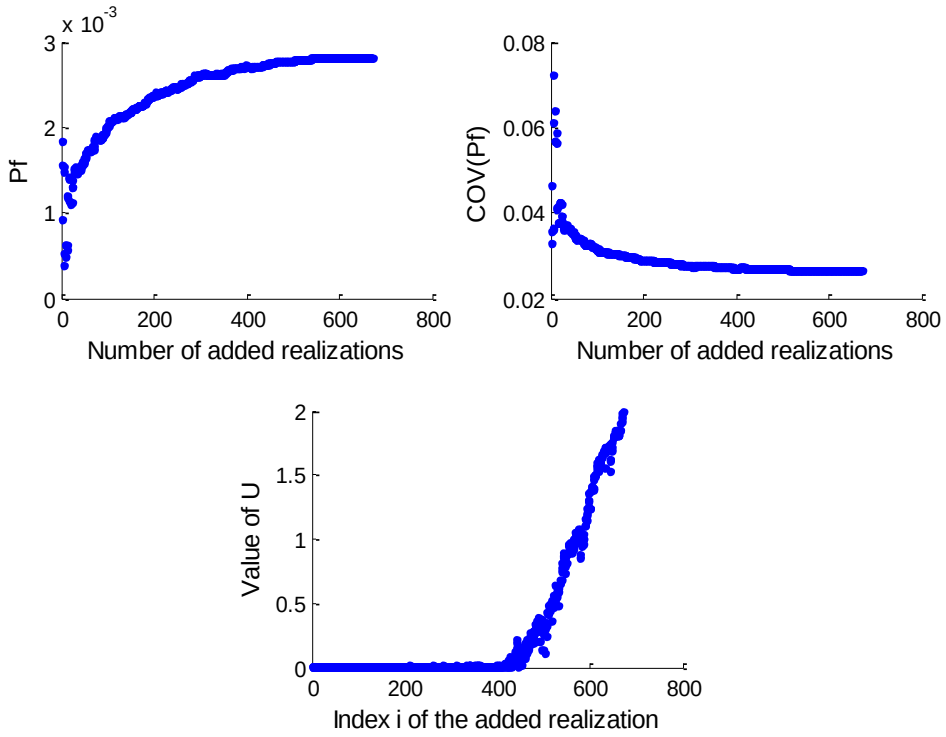


Figure D.11. AK-MCS results for a spatially varying soil for the case (a_x=10m, a_y=2m)

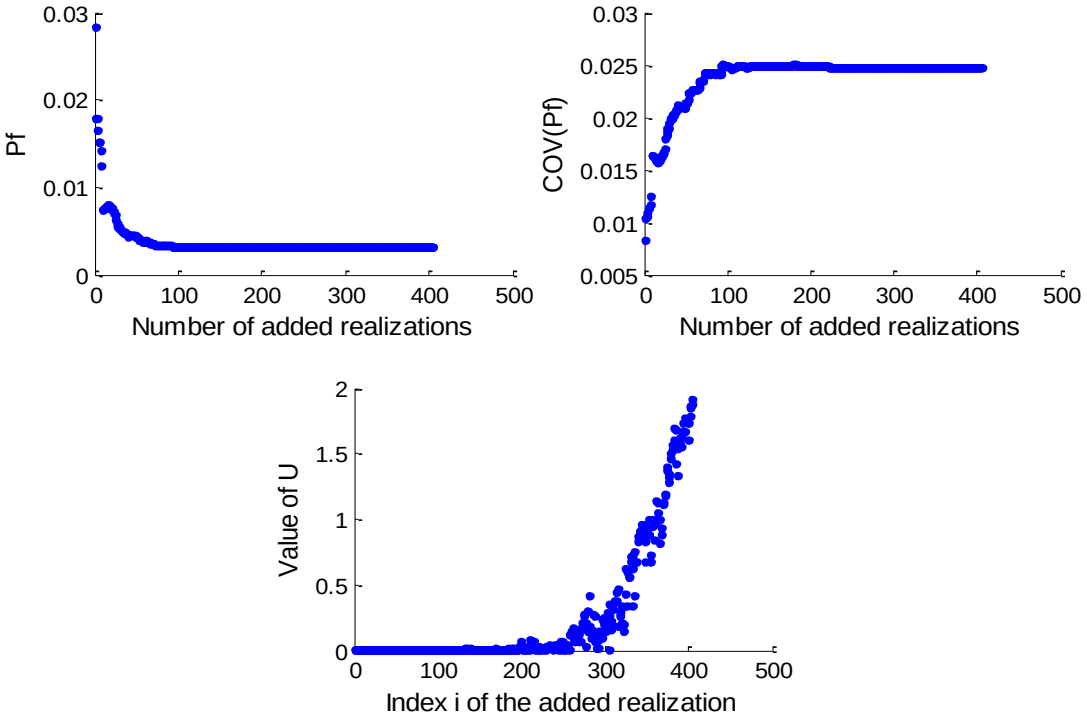


Figure D.12. AK-MCS results for a spatially varying soil for the case (a_x=10m, a_y=5m)

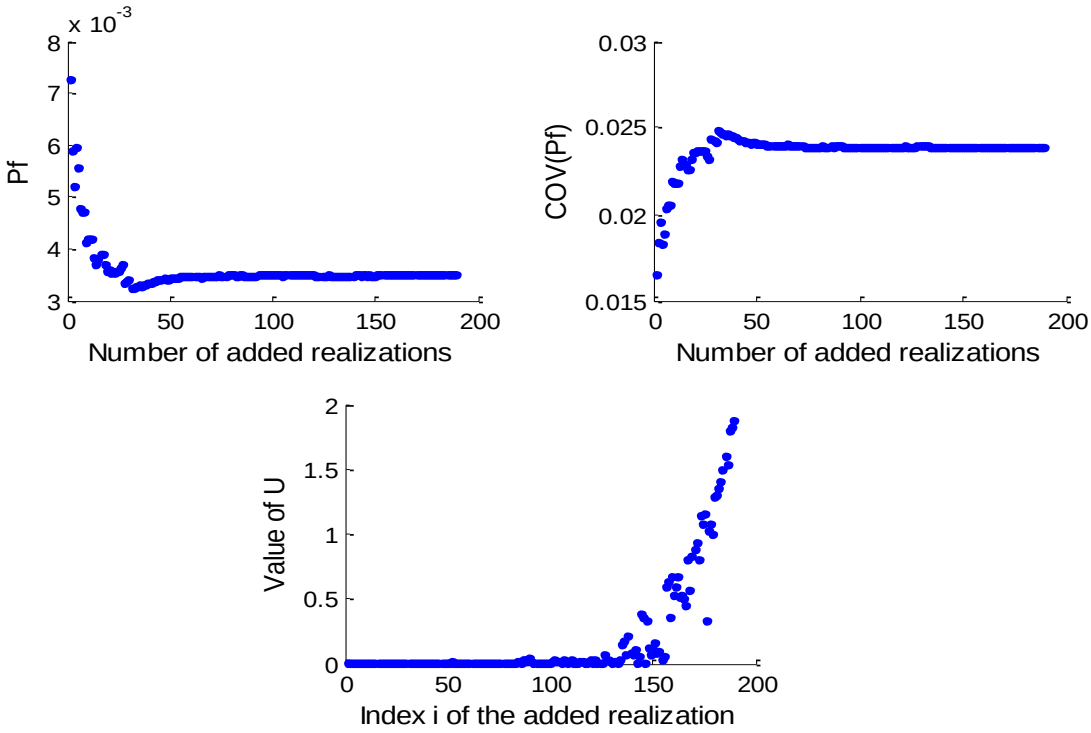


Figure D.13. AK-MCS results for a spatially varying soil for the case (a_x=10m, a_y=20m)

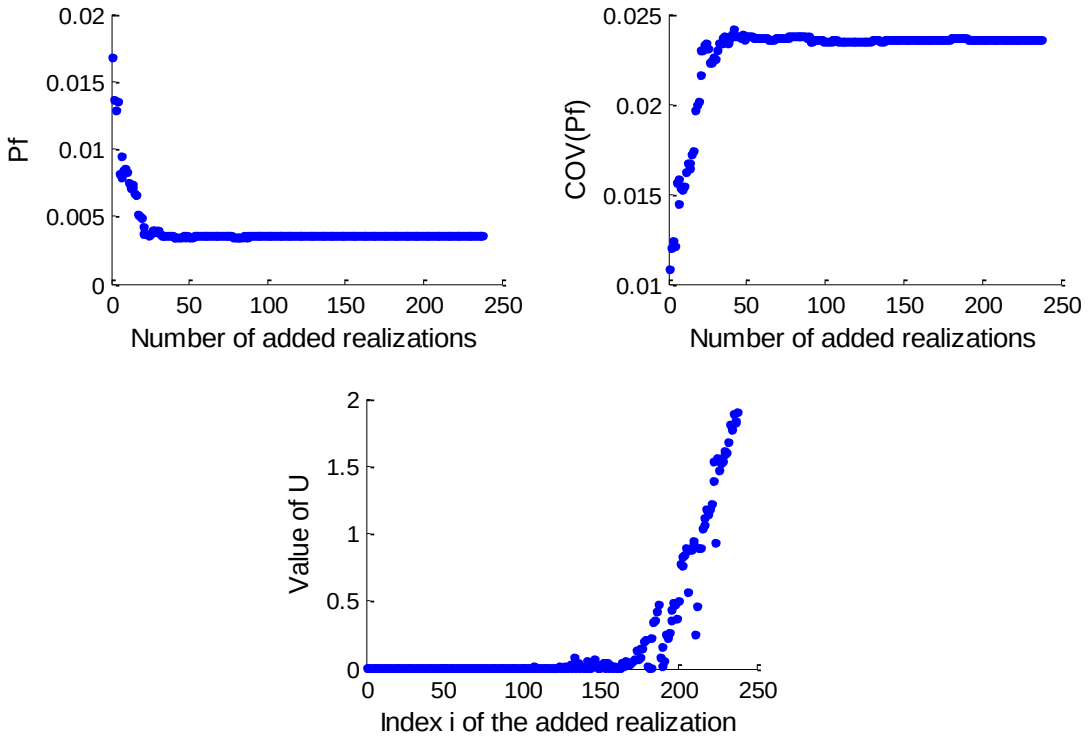


Figure D.14. AK-MCS results for a spatially varying soil for the case (a_x=10m, a_y=50m)

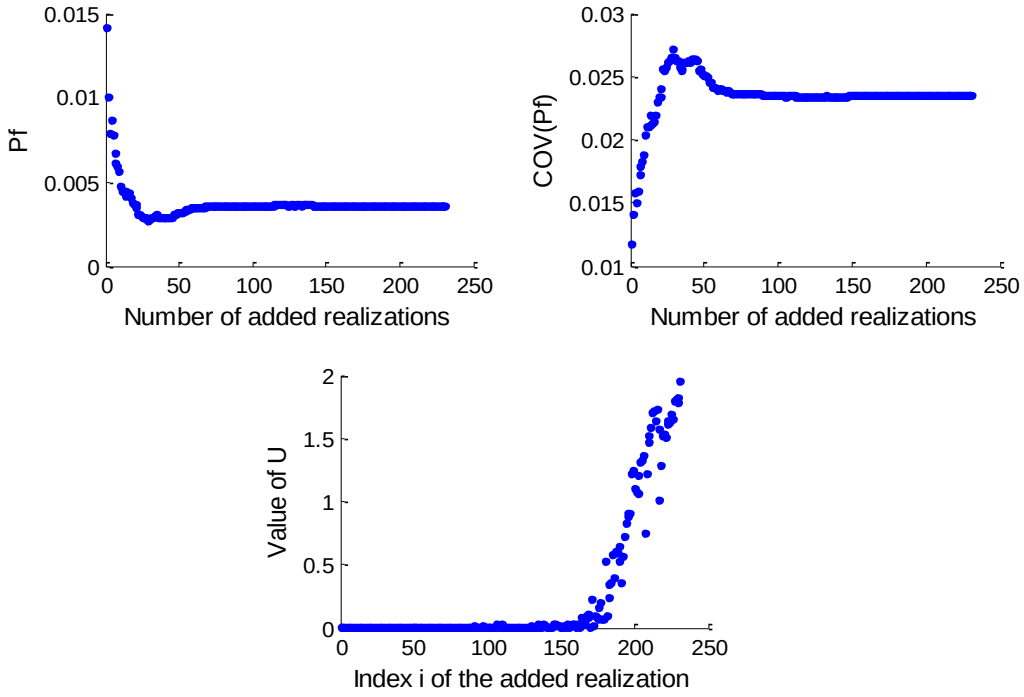


Figure D.15. AK-MCS results for a spatially varying soil for the case (a_x=10m, a_y=100m)

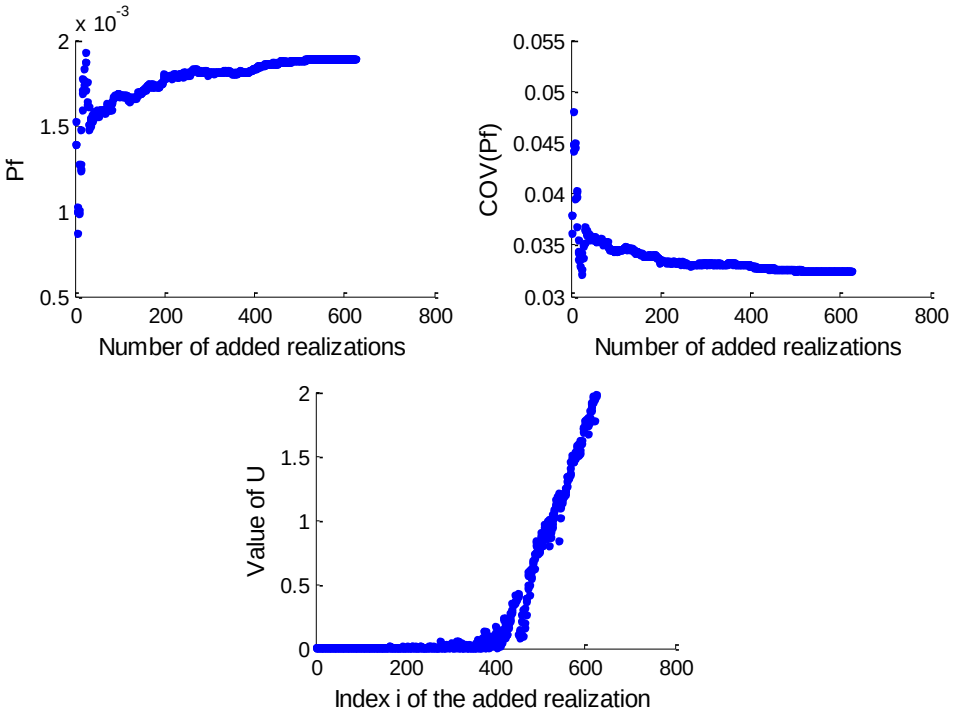


Figure D.16. AK-MCS results for a spatially varying soil for the case (a_x=10m, a_y=10000m)

Case of an anisotropic soil (a_y=2 m with varying a_x)

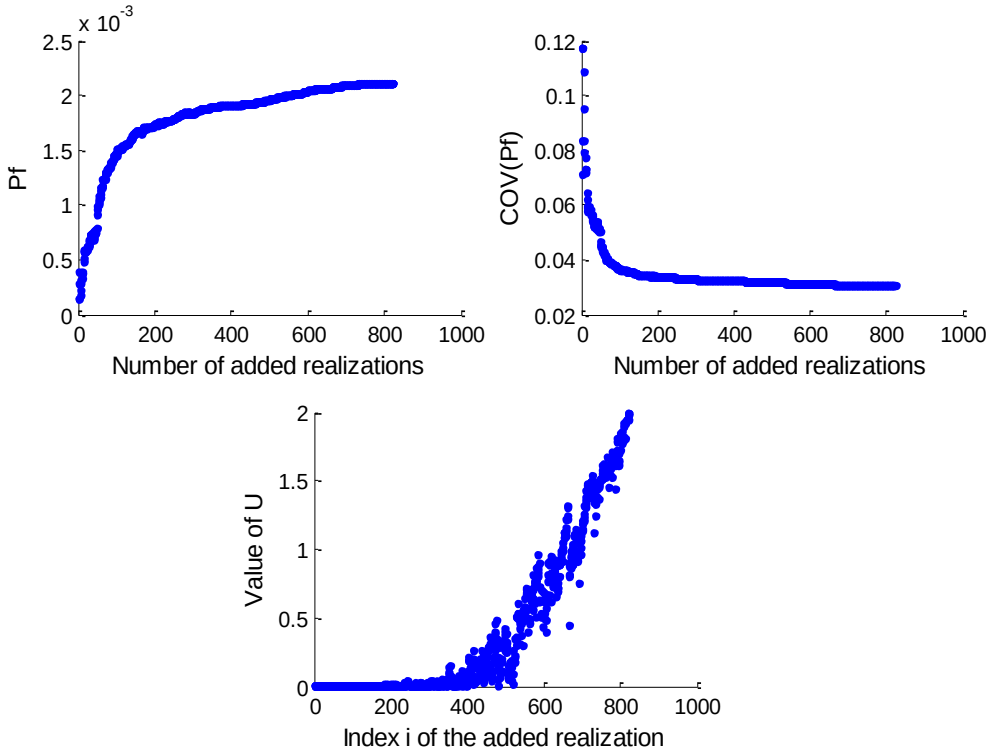


Figure D.17. AK-MCS results for a spatially varying soil for the case (a_x=5m, a_y=2m)

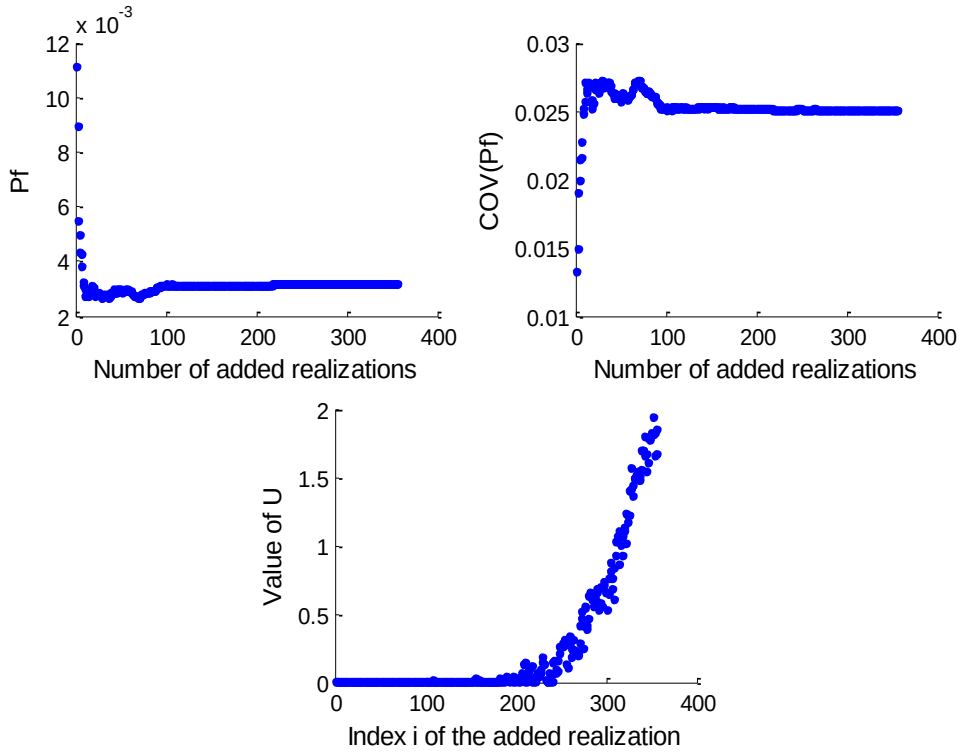


Figure D.18. AK-MCS results for a spatially varying soil for the case (a_x=20m, a_y=2m)

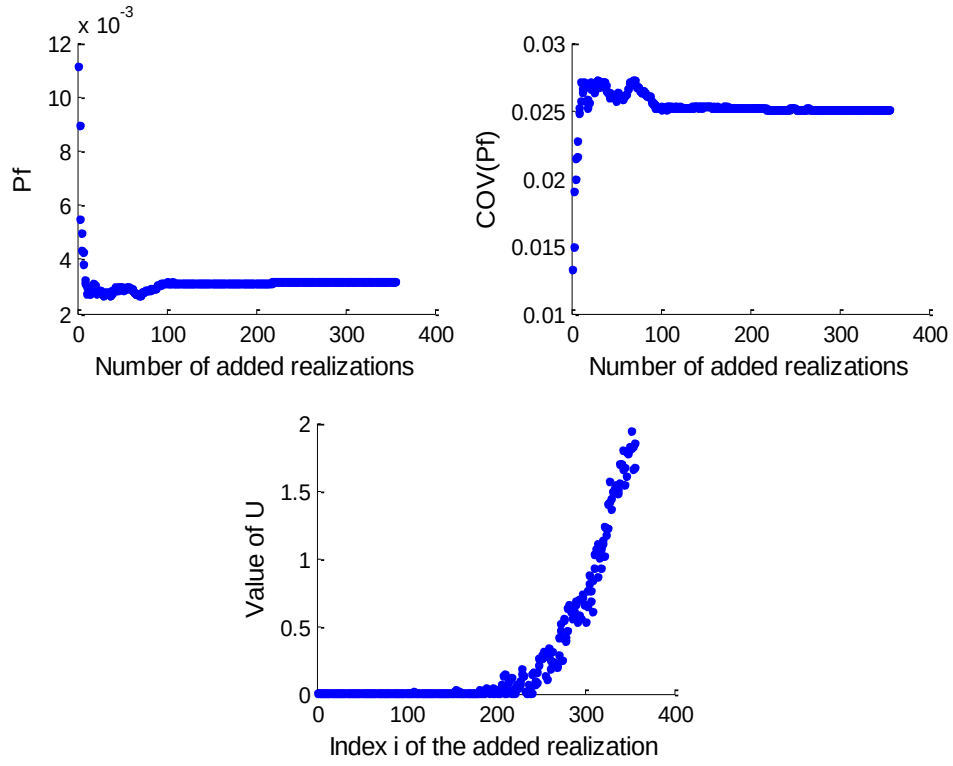


Figure D.19. AK-MCS results for a spatially varying soil for the case (a_x=50m, a_y=2m)

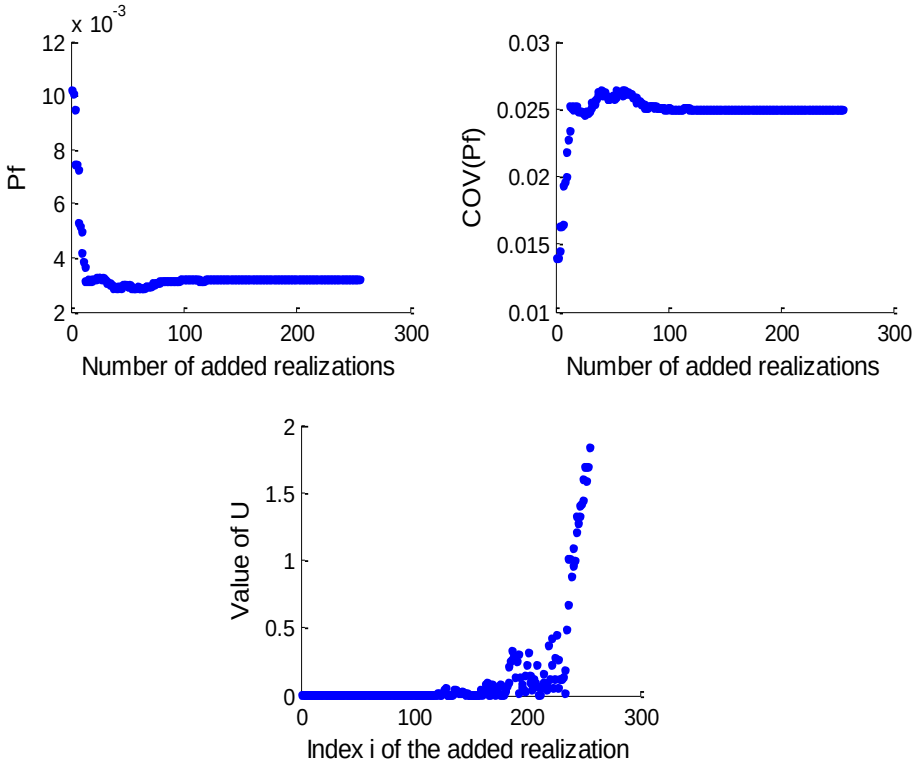


Figure D.20. AK-MCS results for a spatially varying soil for the case (a_x=100m, a_y=2m)

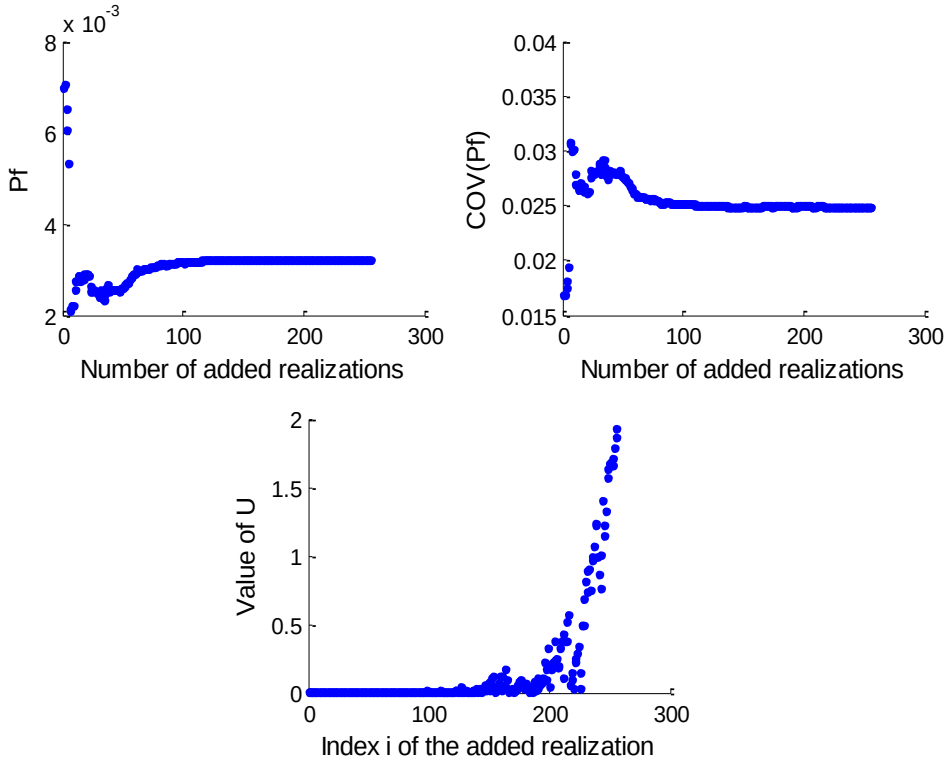


Figure D.21. AK-MCS results for a spatially varying soil for the case (a_x=10000m, a_y=2m)

APPENDIX E

AK-IS: An active learning reliability method combining Kriging and Importance Sampling

The AK-IS approach as described by Echard et al. (2013) is composed of two main steps:

In a first step, the FORM approximation is performed and the most probable failure point P^* is computed using the Hasofer–Lind–Rackwitz–Fiessler algorithm or other. The aim of determining the failure point is to shift the sampling distribution from the center of the standard space of random variables so that it is centered at this failure point. Notice that the determination of the failure point is a rather time-demanding step as it is performed on the real performance function G. Notice however that an accurate most probable failure point is not required and thus a few iterations are enough to shift the center of the sampling population towards the failure domain.

In a second step, the classification by Kriging of a population P_{IS} was performed. The population P_{IS} of N_{IS} samples is simulated using a sampling PDF centered at the approximated most probable failure point found in the first step of the AK-IS procedure. The classification into the safe and failure groups is performed using the technique developed in AK-MCS [Echard et al. (2011)]. The second step of the IK-IS procedure may be described in more details as follows:

1. Generation of a population P_{IS}: P_{IS} is generated using a sampling PDF centered at the approximated most probable failure point determined in the first step of the AK-IS procedure. The generated samples represent candidate samples at which the performance function may be computed if the active learning requires it.
2. Definition of the initial DoE: The initial DoE is composed of the samples at which the performance function has already been computed during FORM approximation at the first stage of AK-IS procedure. These samples are called FORM DoE and are considered to be sufficient to start the procedure. Note here that no other sample is computed at this stage.

3. Computation of the kriging meta-model according to the DoE: This stage is performed using the DACE toolbox of Matlab.
4. Prediction by kriging and estimation of the probability of failure: First, kriging predictions $\mu_{\hat{G}}$ and their corresponding kriging variance values $\sigma^2_{\hat{G}}$ for $i = 1,\ldots,N_{IS}$ are computed using DACE toolbox. Then, the probability of failure is estimated with the signs of predictions $\mu_{\hat{G}}$ using Equation (2.16).
5. Identification of the best next sample x* in P_{IS}: The identification of this sample is performed according to the learning function U proposed by Echard et al. (2011). The learning function U is given by Equation (3.8). The criterion to identify the best next point is the one with minimum value of U.
6. Stopping condition on learning: The learning stops if the minimum value of U is greater than 2.
7. Update of the previous design of experiments with the identified sample: If the stopping condition of Stage 6 is not satisfied, the learning continues and the performance function is computed at the best sample x^*. Following this, it is added to the design of experiments. The method, then, goes back to Stage 3 to compute the new kriging model with the updated design of experiments. This process of learning is repeated until the stopping condition is satisfied.
8. Computation of the coefficient of variation of the probability of failure: If the stopping condition of Stage 6 is satisfied, the learning stops. The population P_{IS} is then checked to be large enough to provide an accurate failure probability prediction. To do so, the coefficient of variation of the probability of failure is calculated by IS on the kriging results using Equation (2.17). Here, a coefficient of variation is compared to a limit value seen as acceptable by the user. It should be mentioned here that a value of the coefficient of variation below 5% is seen as acceptable.

9. Update of the population: If the coefficient of variation is found to be too high, P_{IS} is enlarged with new samples and AK-IS goes back to Stage 4 to predict the new population. The active learning method continues until the stopping condition is met again.
10. End of AK-IS procedure: If the coefficient of variation of P_f is small enough, AK-IS stops and the last estimation of the probability of failure is considered as the result of the method.

APPENDIX F

Failure probability P_f and $COV\left(P_f\right)$ versus the number of added realizations obtained using AK-IS method

The figures below present the failure probability P_f and the coefficient of variation $COV\left(P_f\right)$ versus the number of added realizations for the different values of the autocorrelation distances (a_x, a_y) used in the analysis (see Table 4.4). These figures also provide the learning function values for the different added realizations.

Case of an isotropic soil (a_x=a_y)

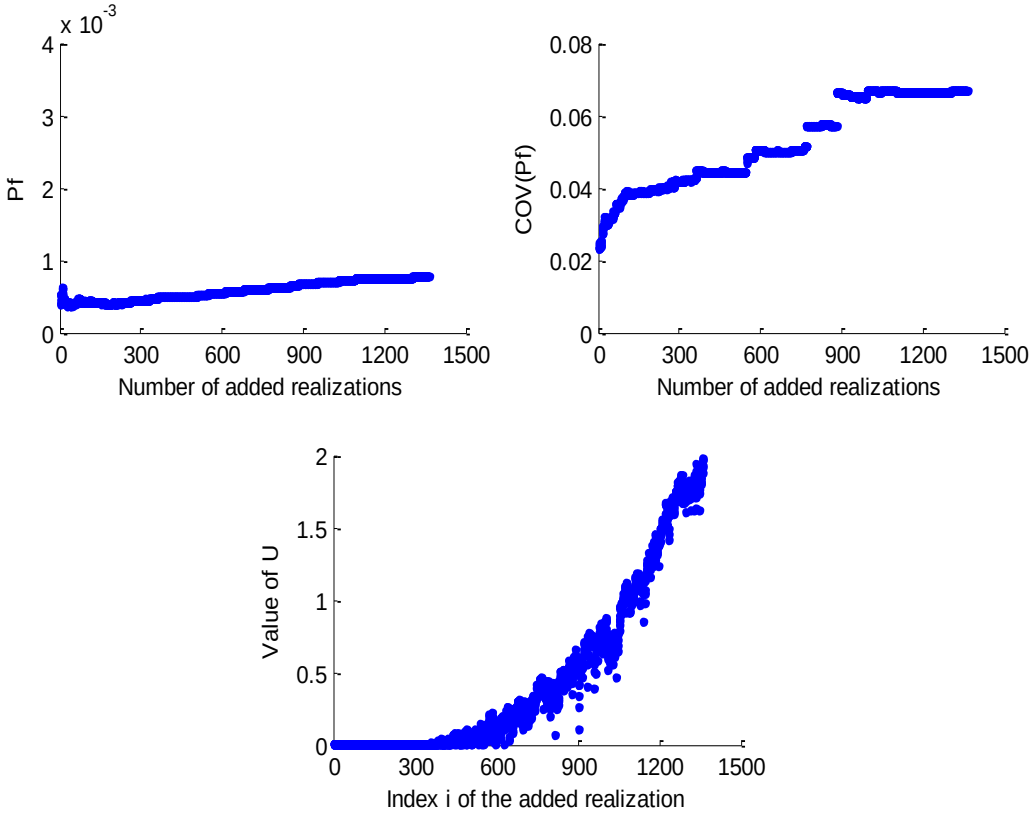


Figure F.1. AK-IS results for a spatially varying soil for the case (a_x=a_y=2m)

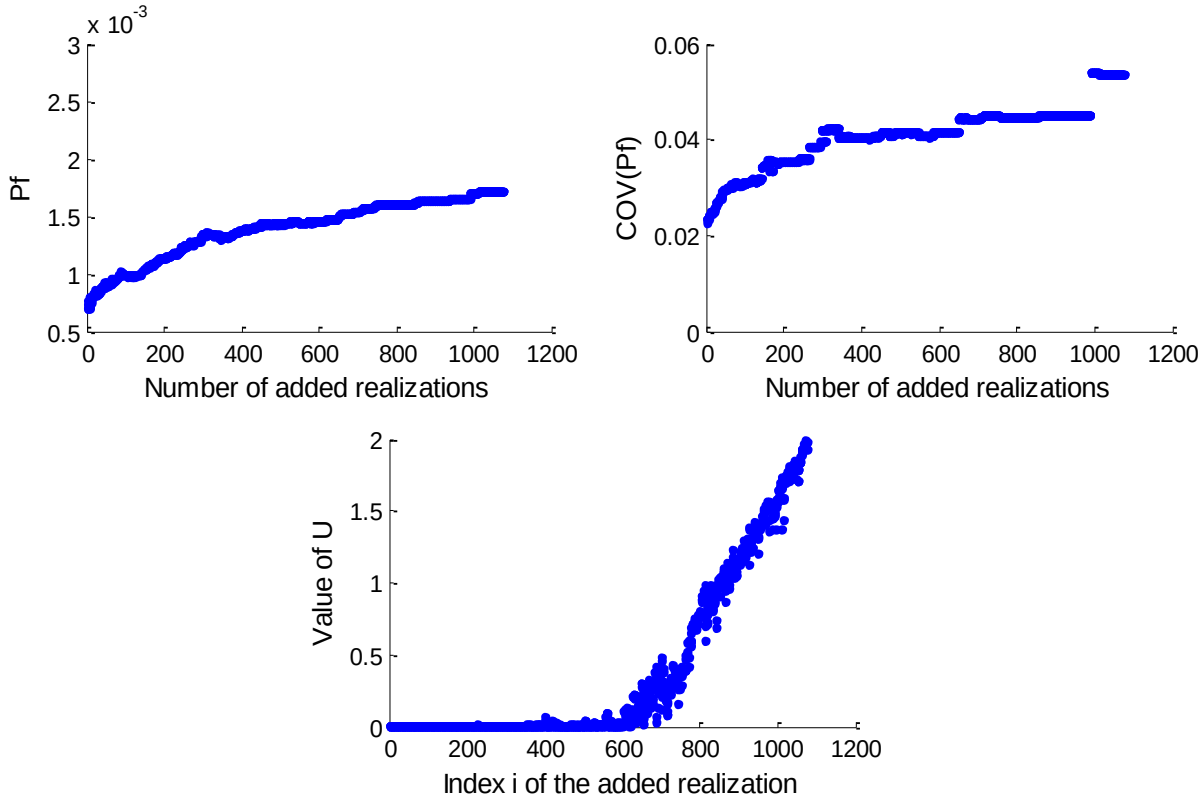


Figure F.2. AK-IS results for a spatially varying soil for the case (a_x=a_y=3m)

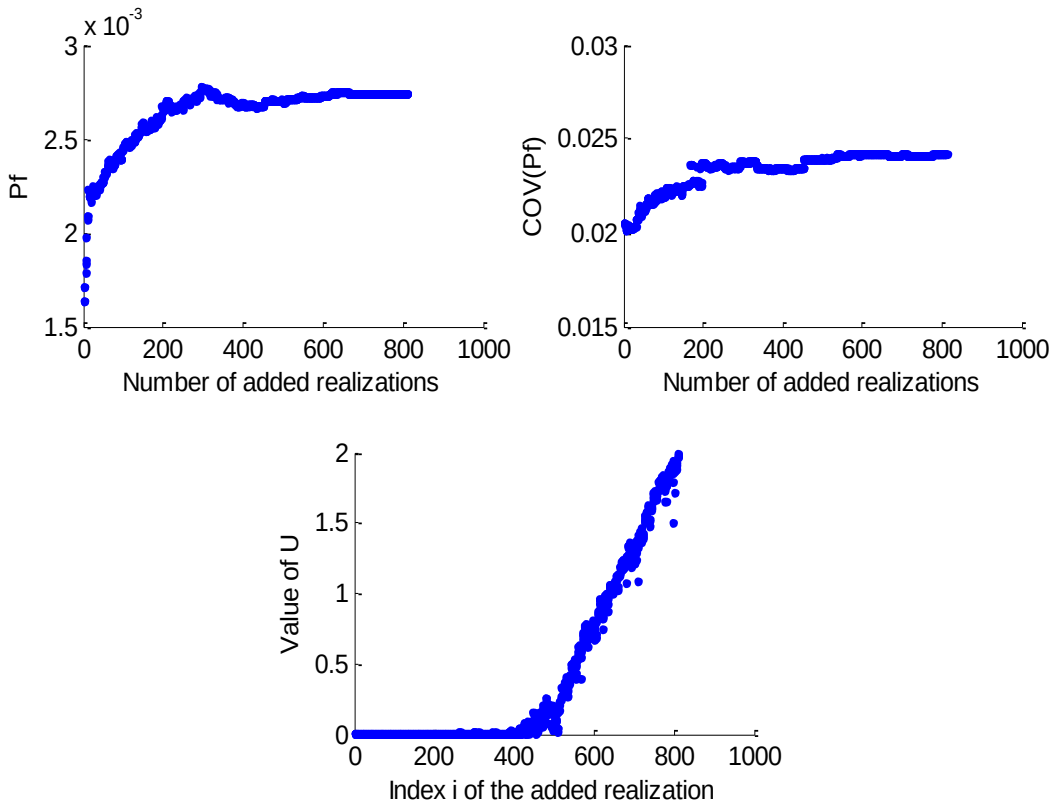


Figure F.3. AK-IS results for a spatially varying soil for the case (a_x=a_y=5m)

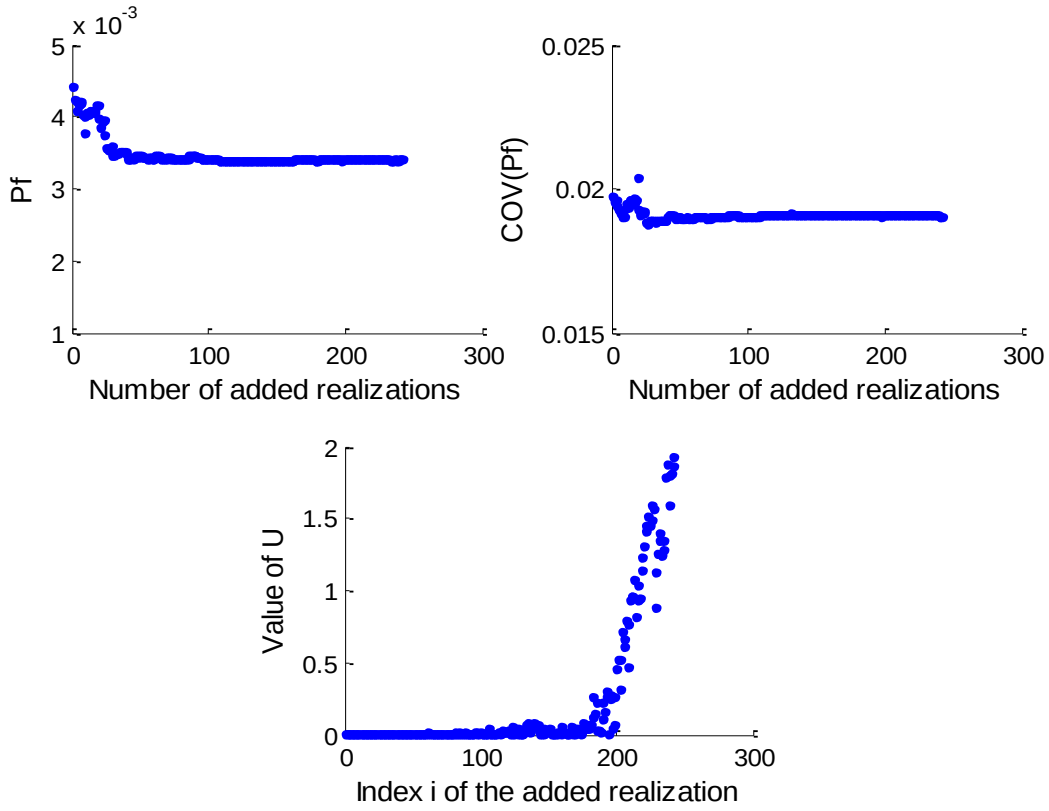


Figure F.4. AK-IS results for a spatially varying soil for the case (a_x=a_y=10m)

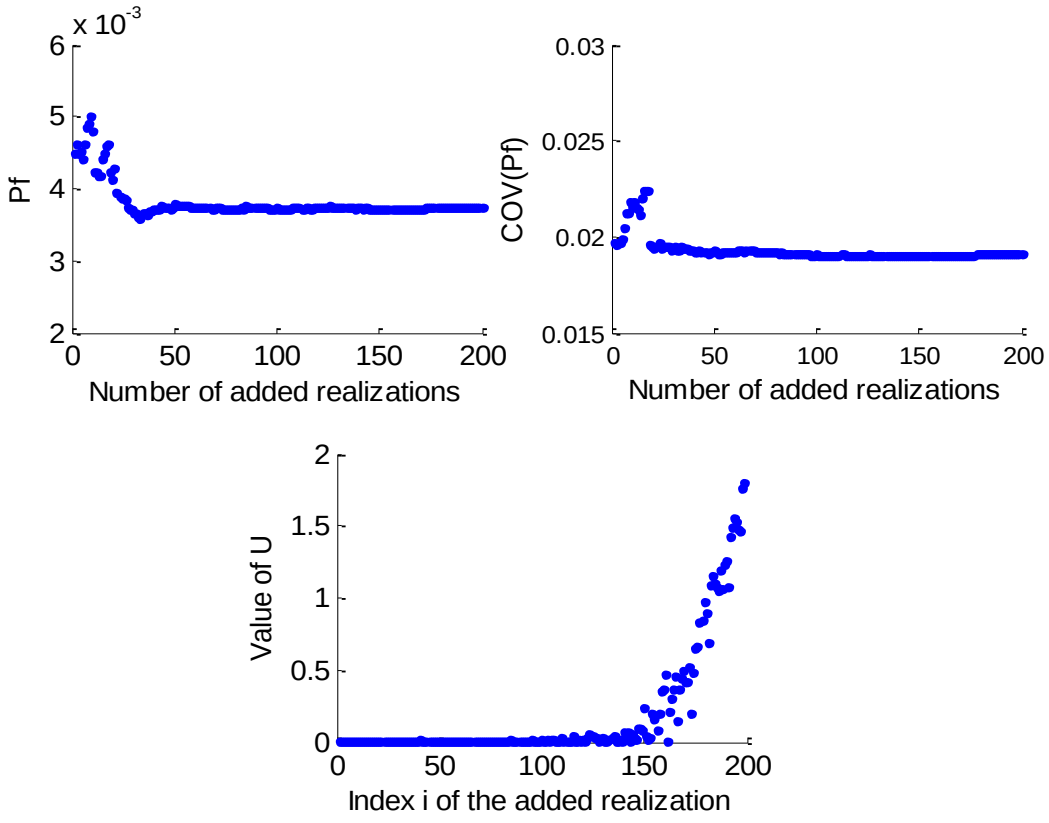


Figure F.5. AK-IS results for a spatially varying soil for the case (a_x=a_y=20m)

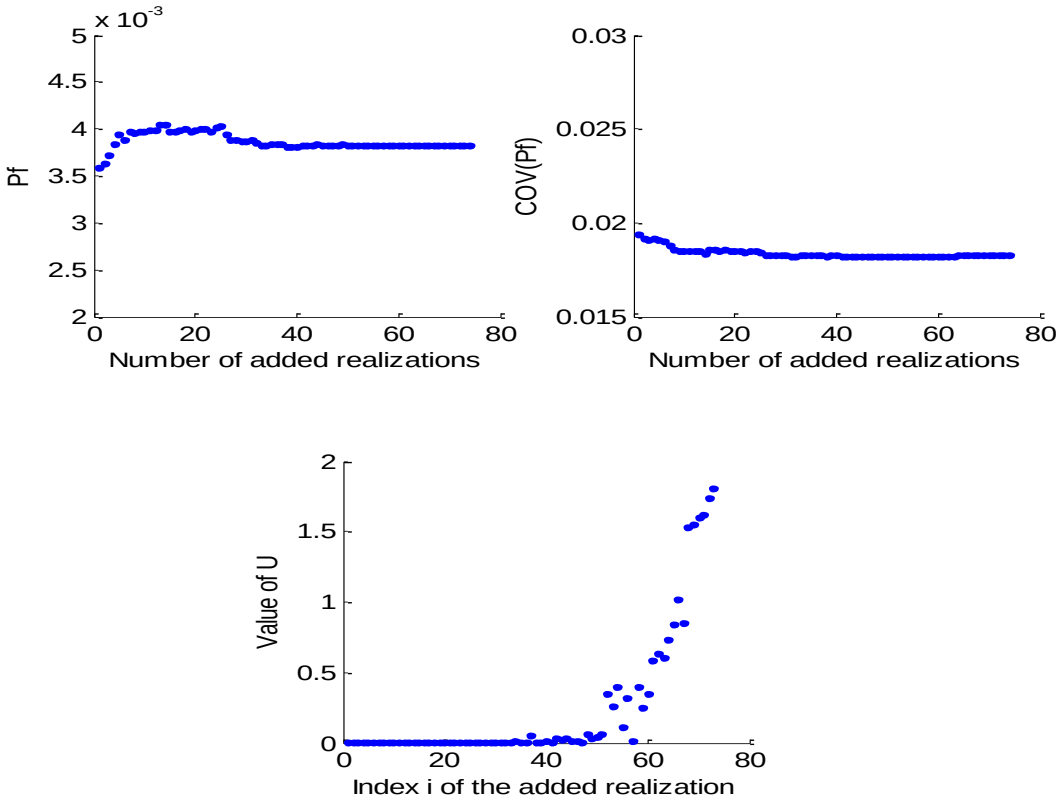


Figure F.6. AK-IS results for a spatially varying soil for the case (a_x=a_y=50m)

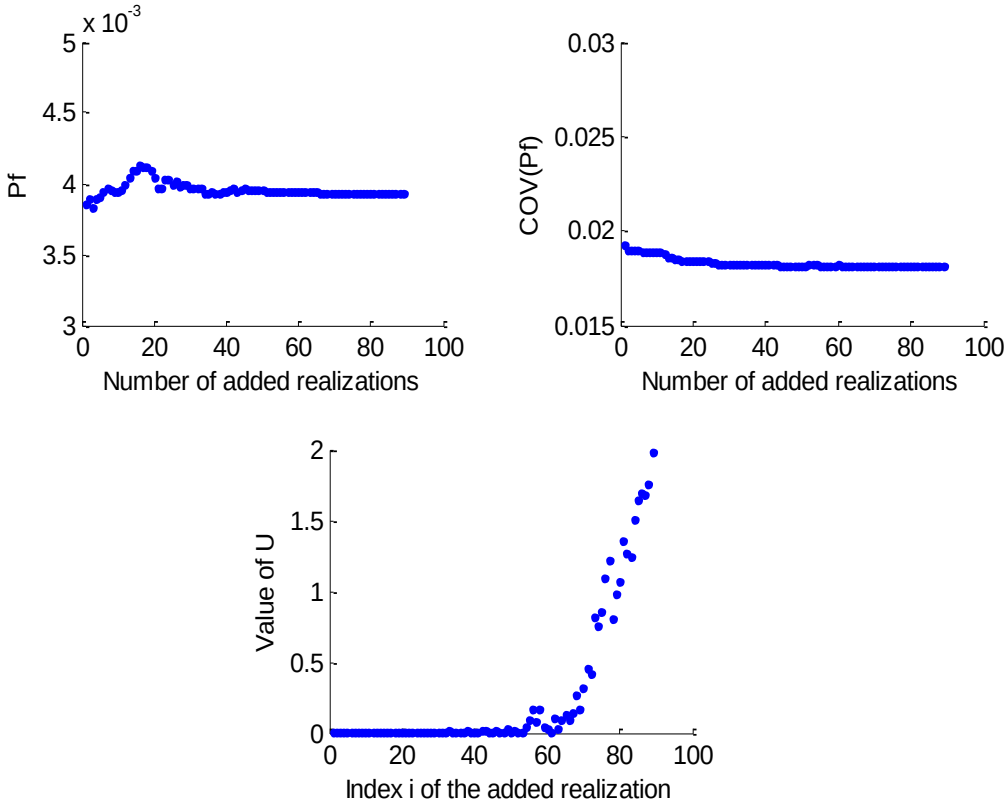


Figure F.7. AK-IS results for a spatially varying soil for the case (a_x=a_y=100m)

Case of an anisotropic soil (a_x=10 m with varying a_y)

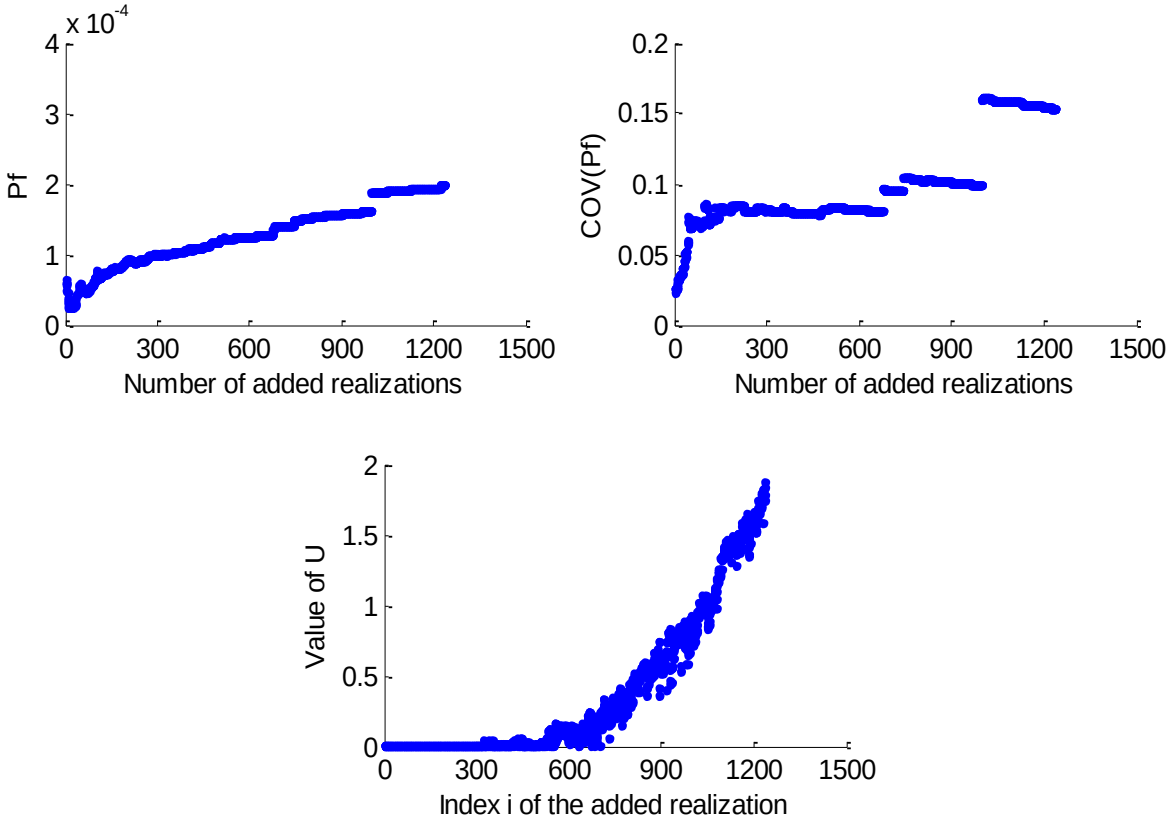


Figure F.8. AK IS results for a spatially varying soil for the case (a_x=10m, a_y=0.5m)

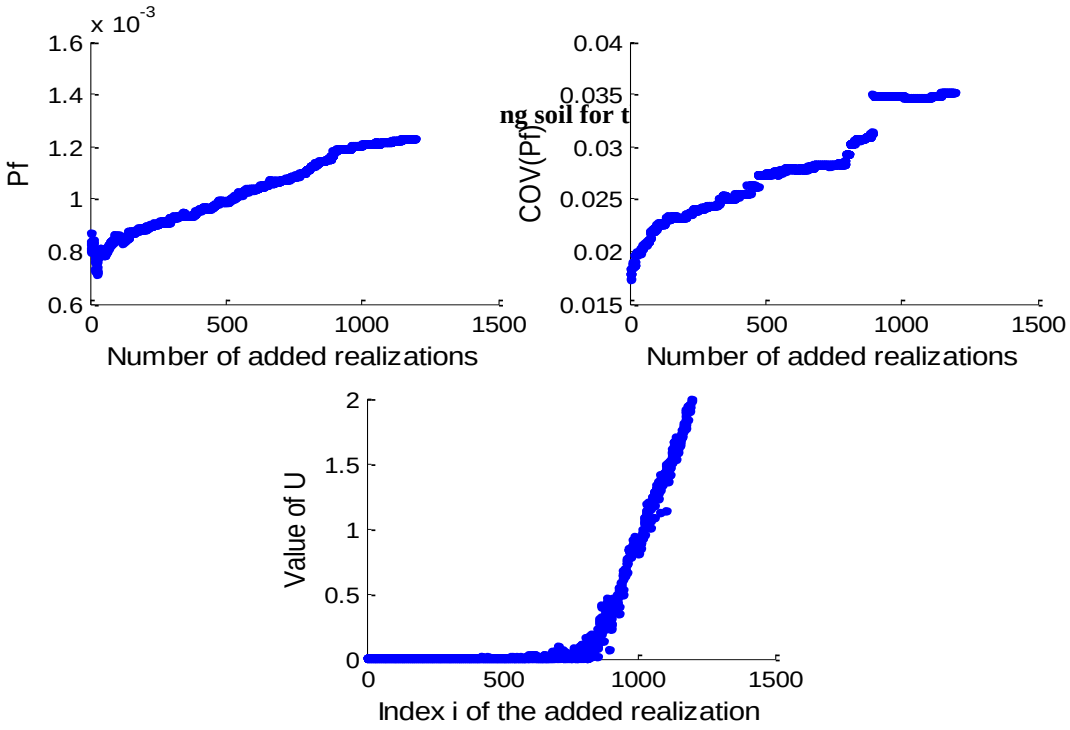


Figure F.9. AK-IS results for a spatially varying soil for the case (a_x=10m, a_y=0.8m)

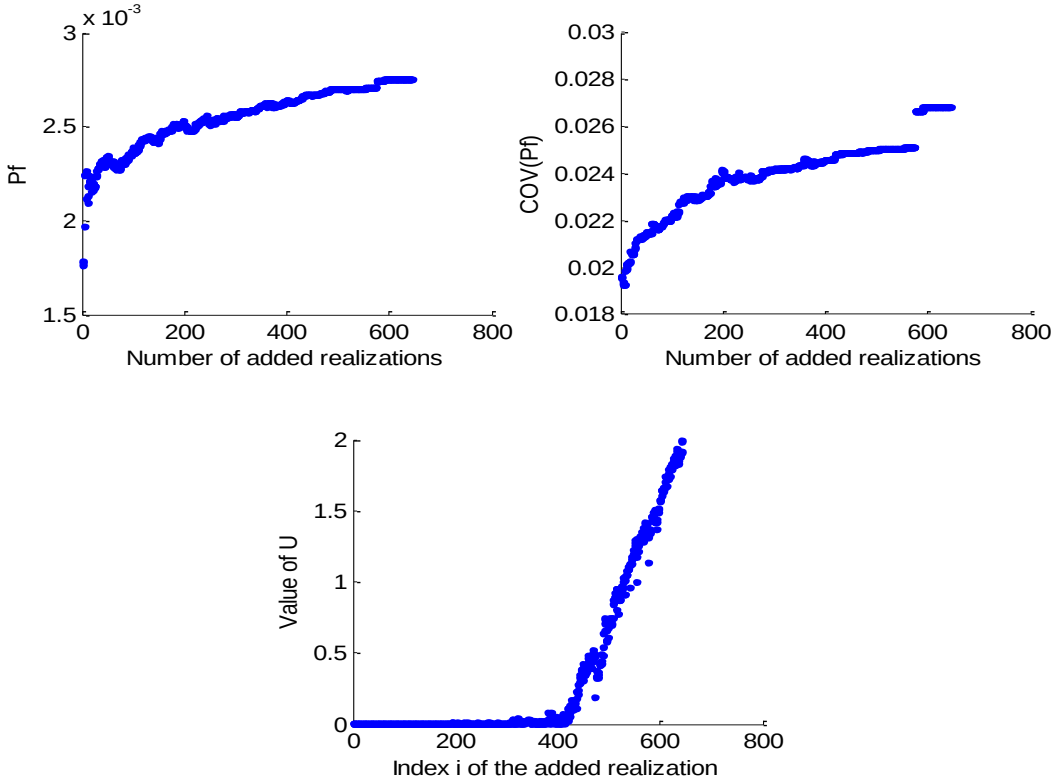


Figure F.10. AK-IS results for a spatially varying soil for the case (a_x=10m, a_y=2m)

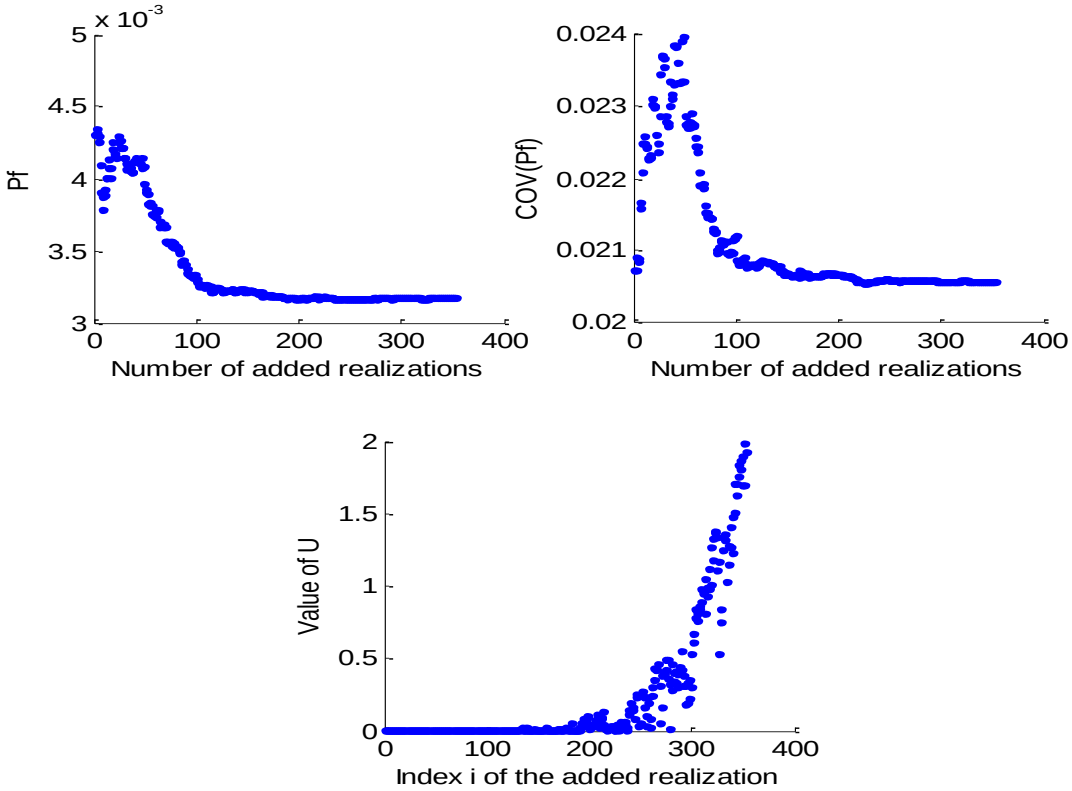


Figure F.11. AK-IS results for a spatially varying soil for the case (a_x=10m, a_y=5m)

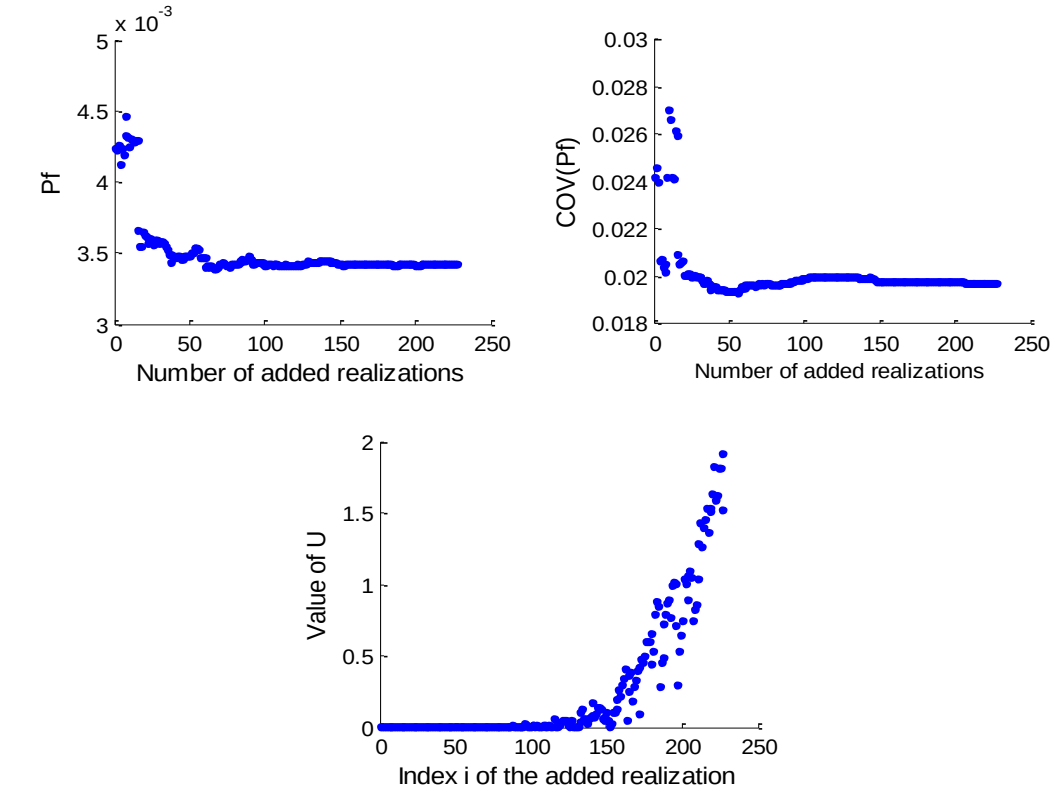


Figure F.12. AK-IS results for a spatially varying soil for the case (a_x=10m, a_y=20m)

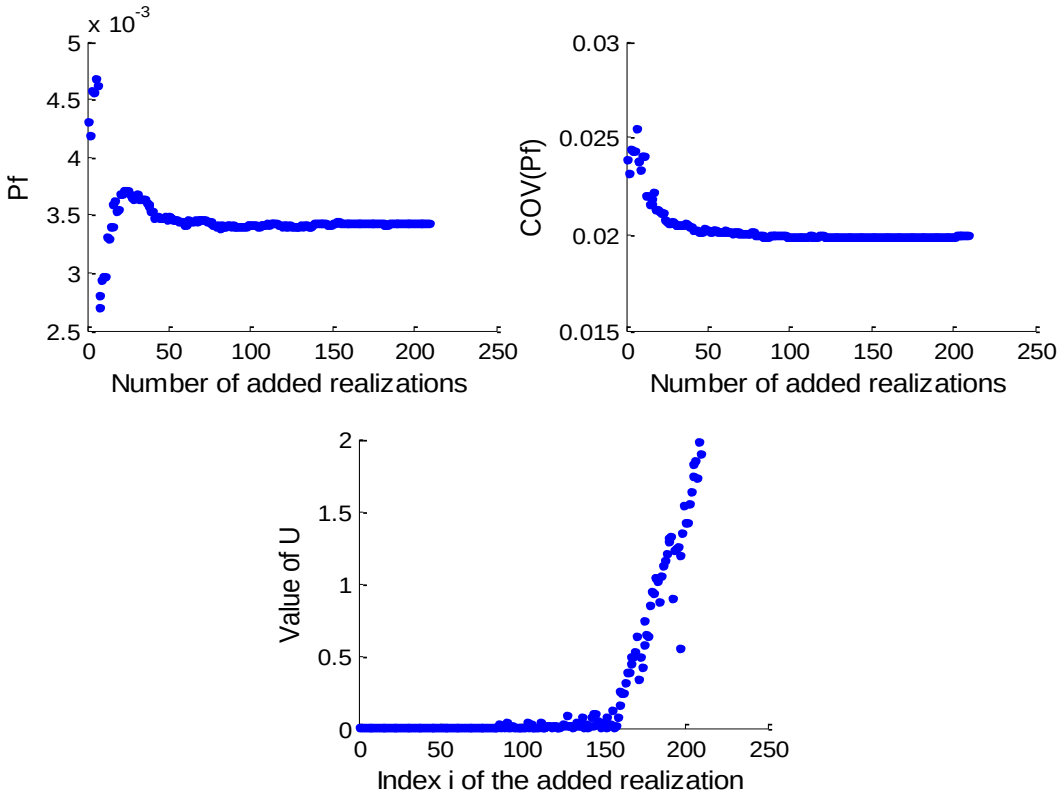


Figure F.13. AK-IS results for a spatially varying soil for the case (a_x=10m, a_y=50m)

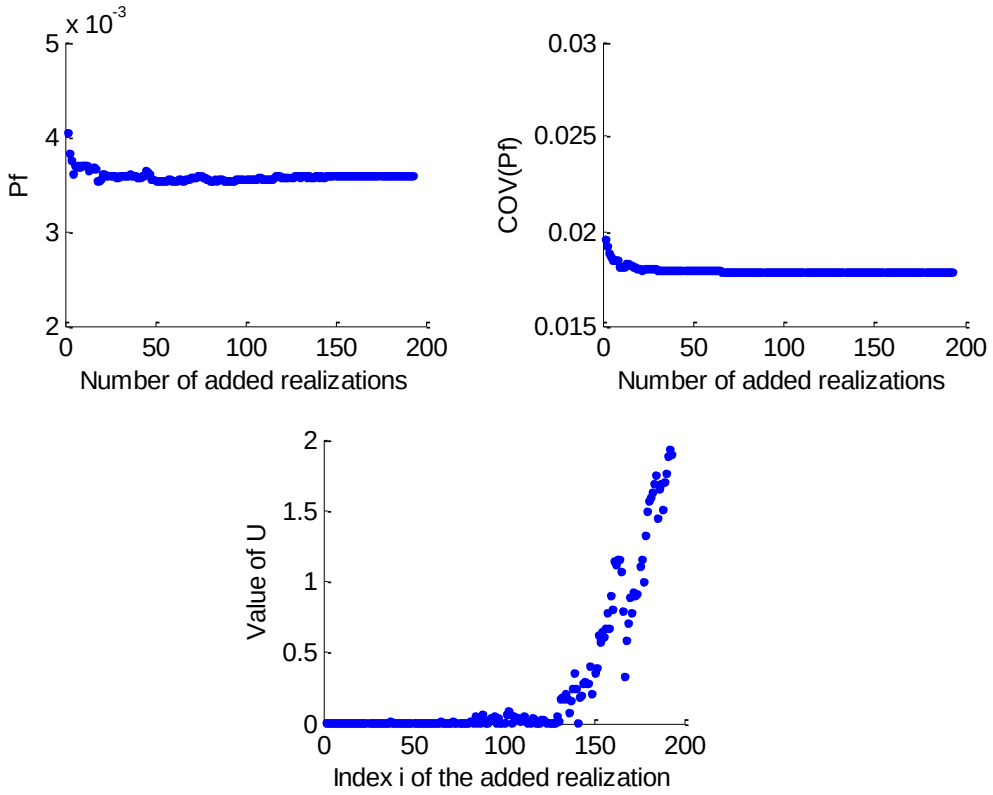


Figure F.14. AK-IS results for a spatially varying soil for the case (a_x=10m, a_y=100m)

Case of an anisotropic soil (a_y=2 m with varying a_x)

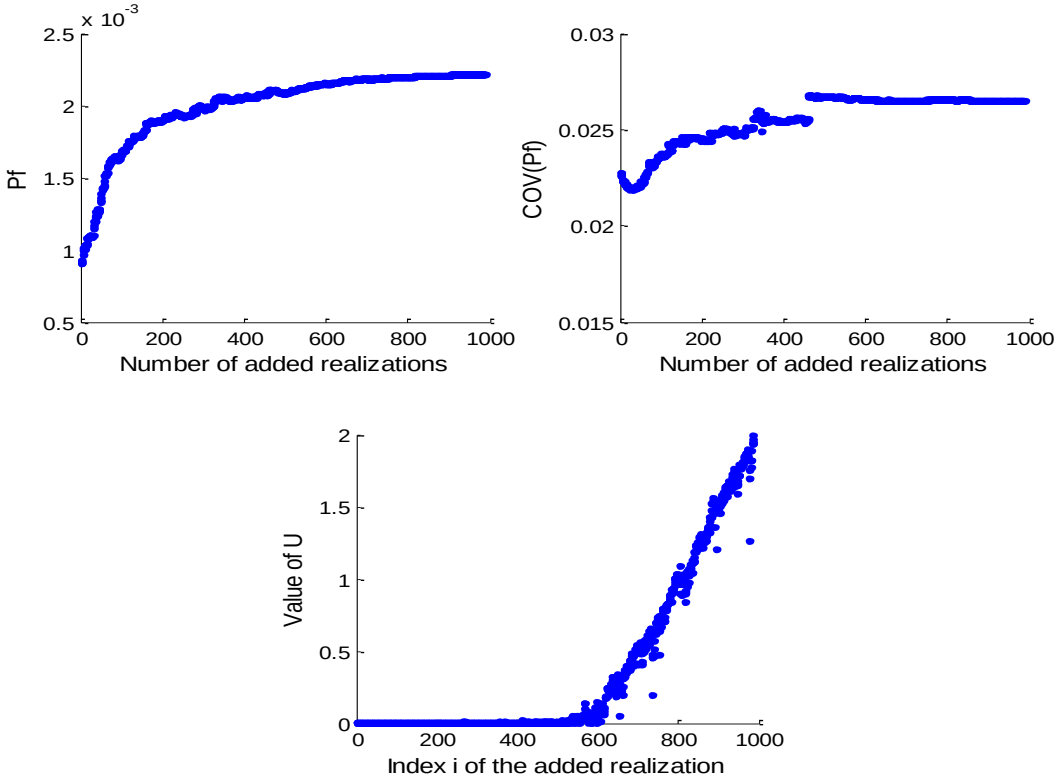


Figure F.15. AK-IS results for a spatially varying soil for the case (a_x=5m, a_y=2m)

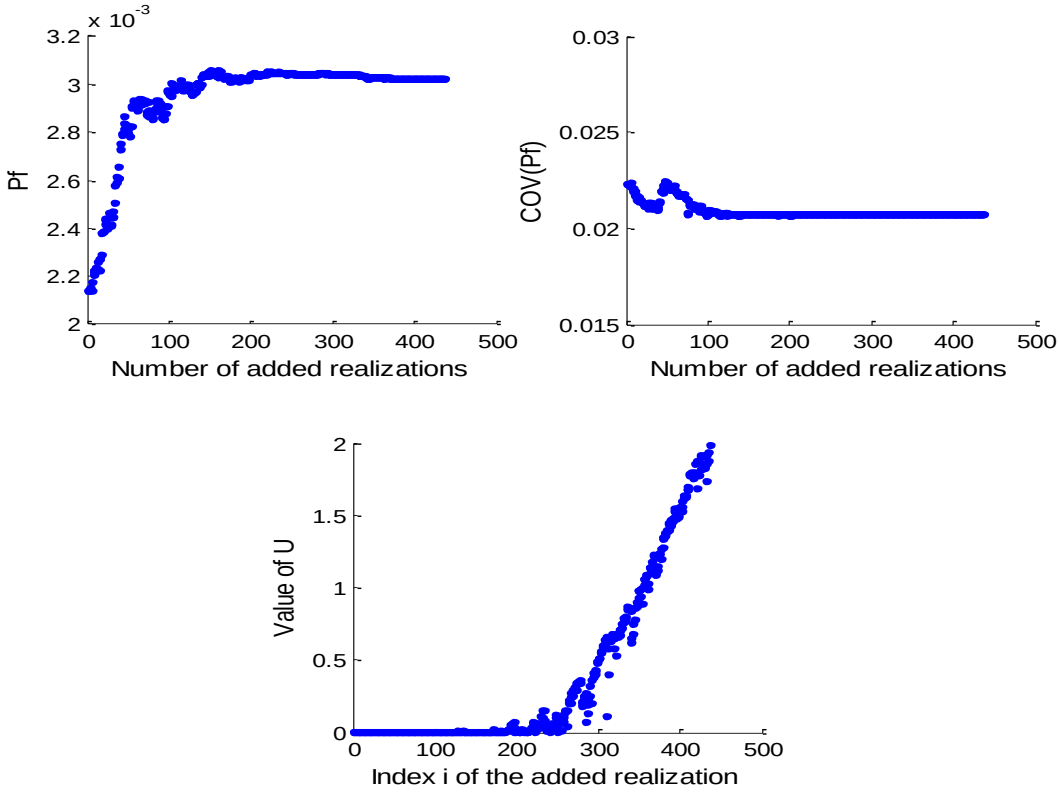


Figure F.16. AK-IS results for a spatially varying soil for the case (a_x=20m, a_y=2m)

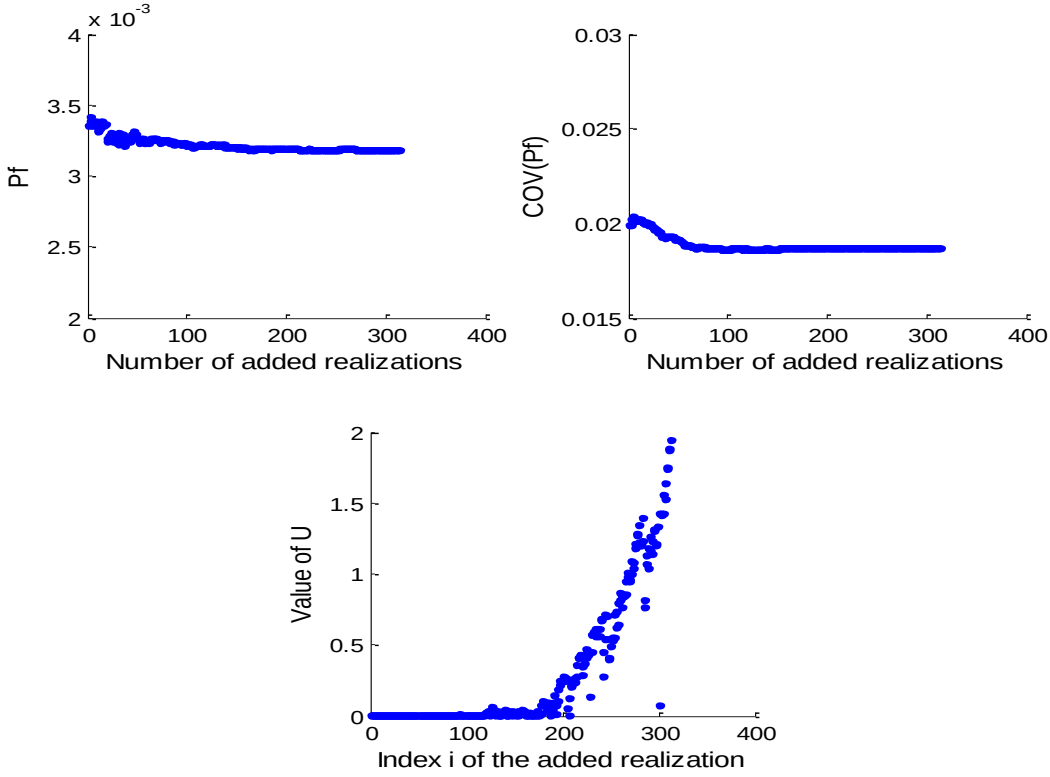


Figure F.17. AK-IS results for a spatially varying soil for the case (a_x=50m, a_y=2m)

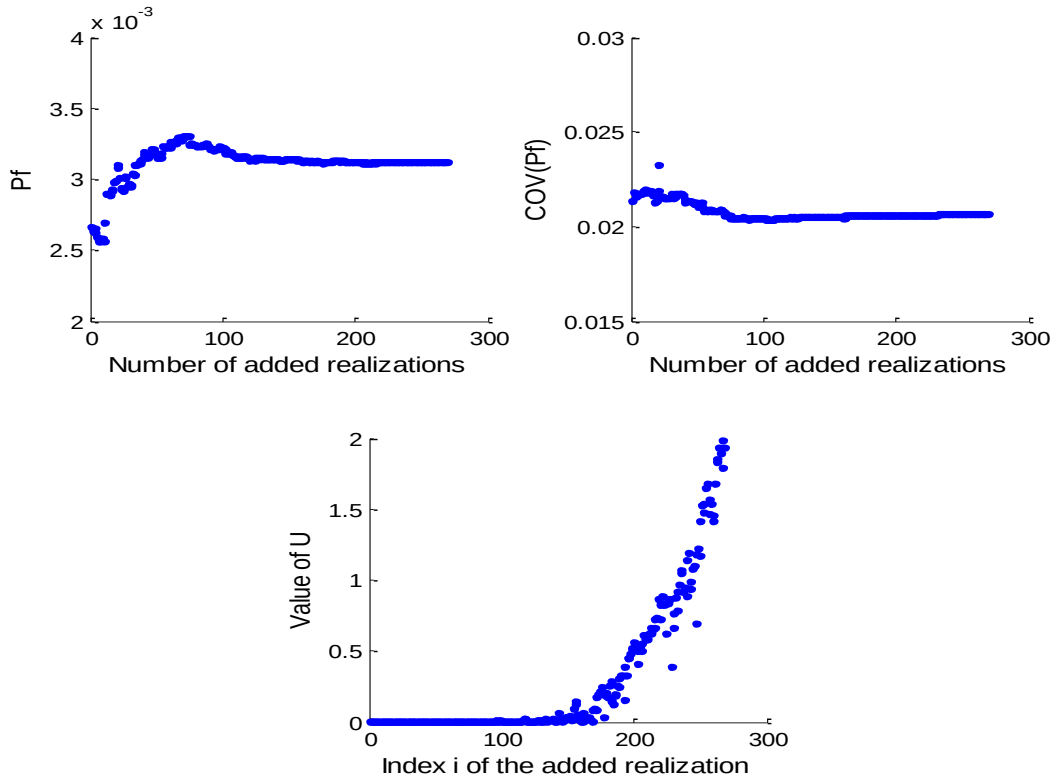


Figure F.18. AK-IS results for a spatially varying soil for the case (a_x=100m, a_y=2m)

APPENDIX G

Failure probability P_f and $COV\left(P_f\right)$ versus the number of added realizations obtained using AK-SS method

The figures below present the failure probability P_f and the coefficient of variation $COV\left(P_f\right)$ versus the number of added realizations for the different values of the autocorrelation distances (a_x, a_y) used in the analysis (see Table 5.4).

Case of an isotropic soil (a_x=a_y)

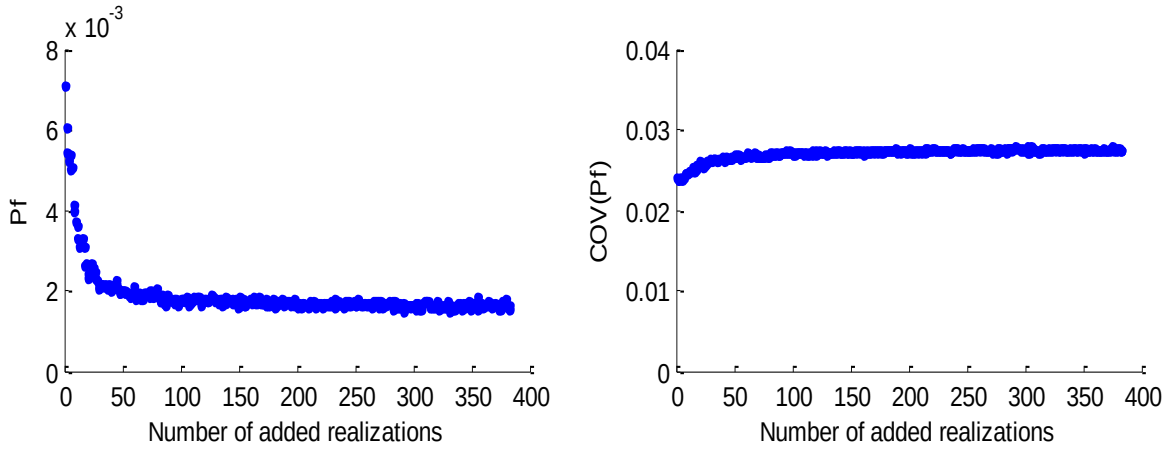


Figure G.1. AK-SS results for a spatially varying soil for the case (a_x=a_y=3m)

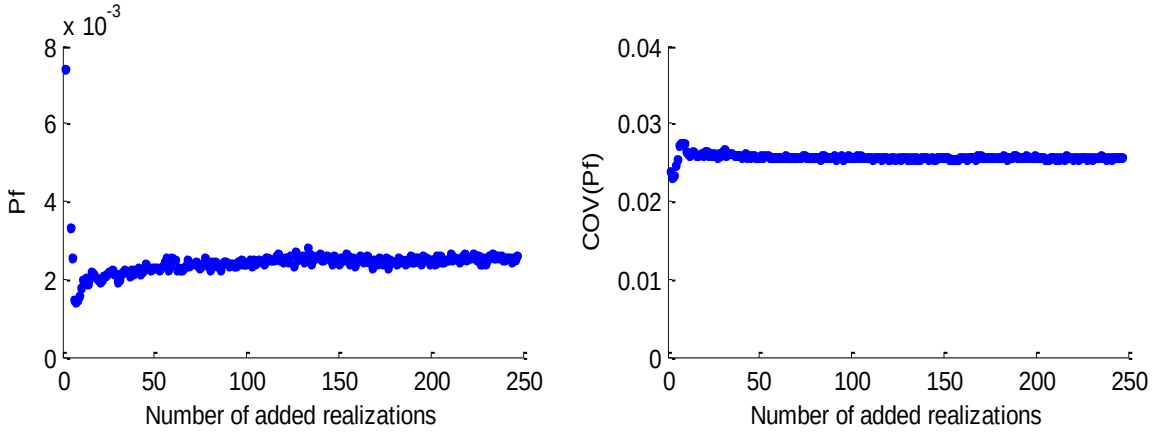


Figure G.2. AK-SS results for a spatially varying soil for the case (a_x=a_y=5m)

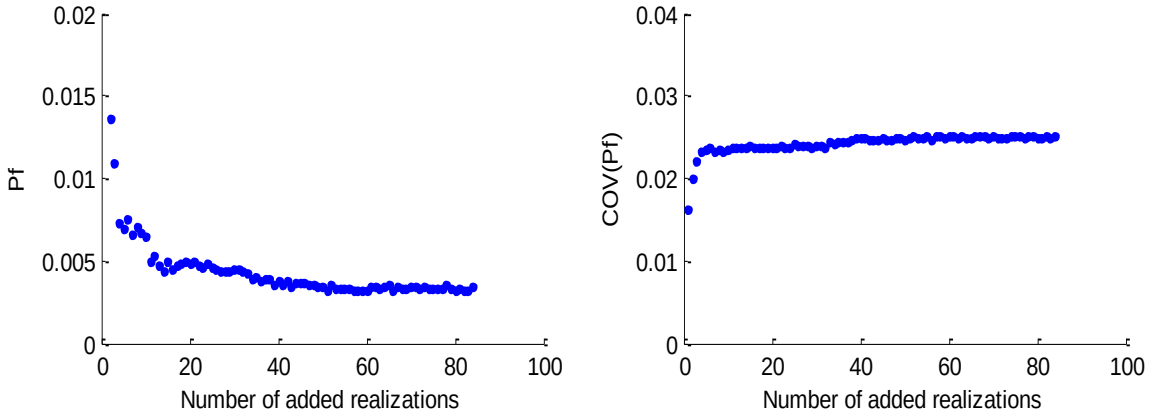


Figure G.3. AK-SS results for a spatially varying soil for the case (a_x=a_y=10m)

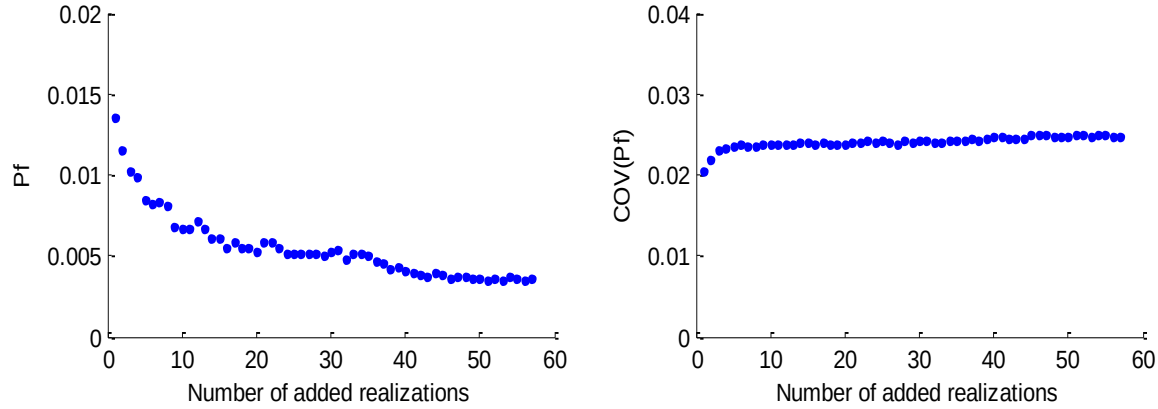


Figure G.4. AK-SS results for a spatially varying soil for the case (a_x=a_y=20m)

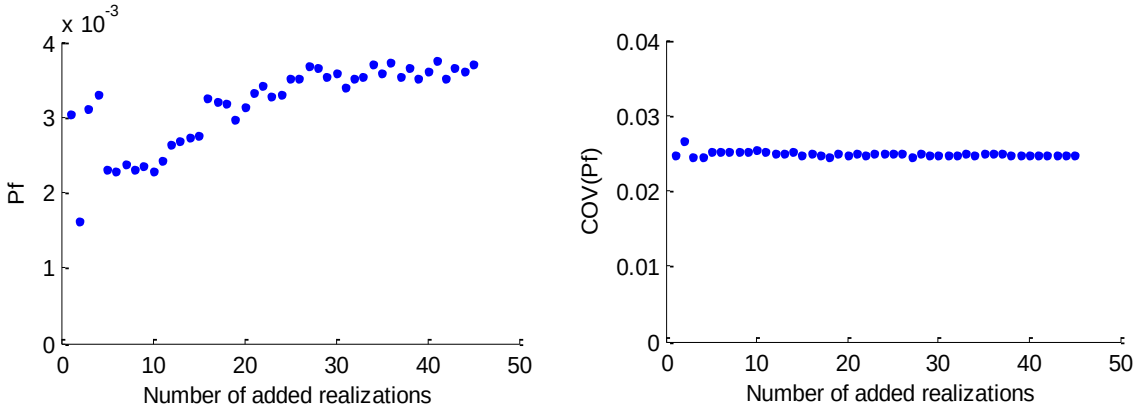


Figure G.5. AK-SS results for a spatially varying soil for the case (a_x=a_y=50m)

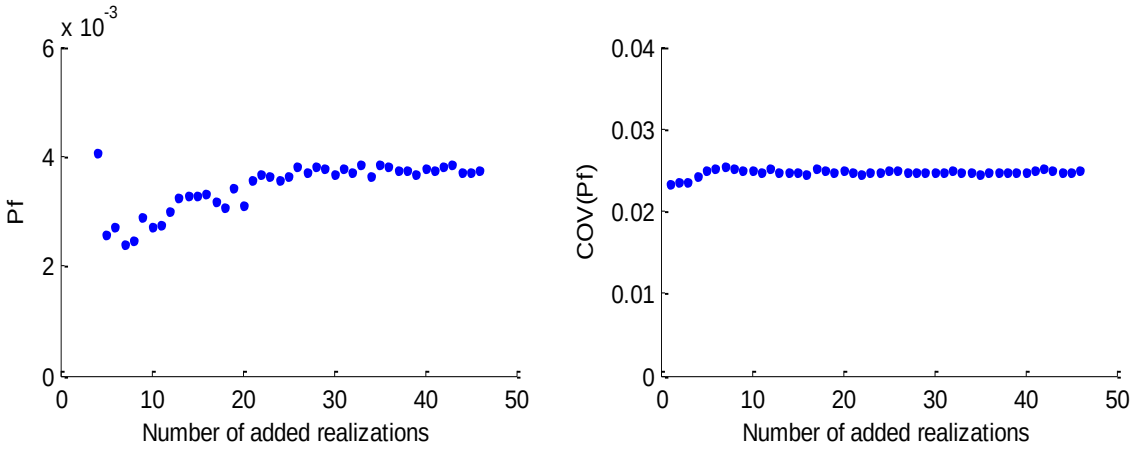


Figure G.6. AK-SS results for a spatially varying soil for the case (a_x=a_y=100m)

Case of an anisotropic soil (a_x=10 m with varying a_y)

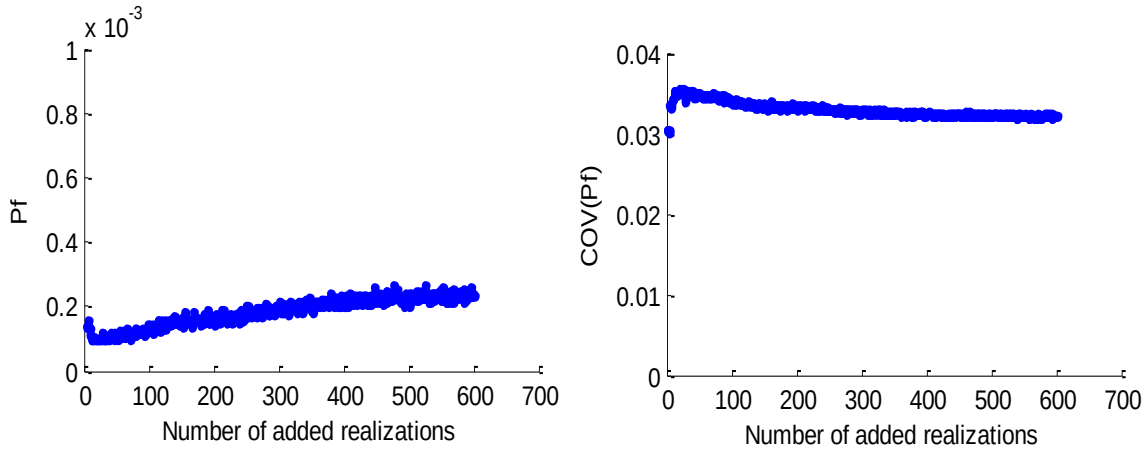


Figure G.7. AK-SS results for a spatially varying soil for the case (a_x=10m, a_y=0.5m)

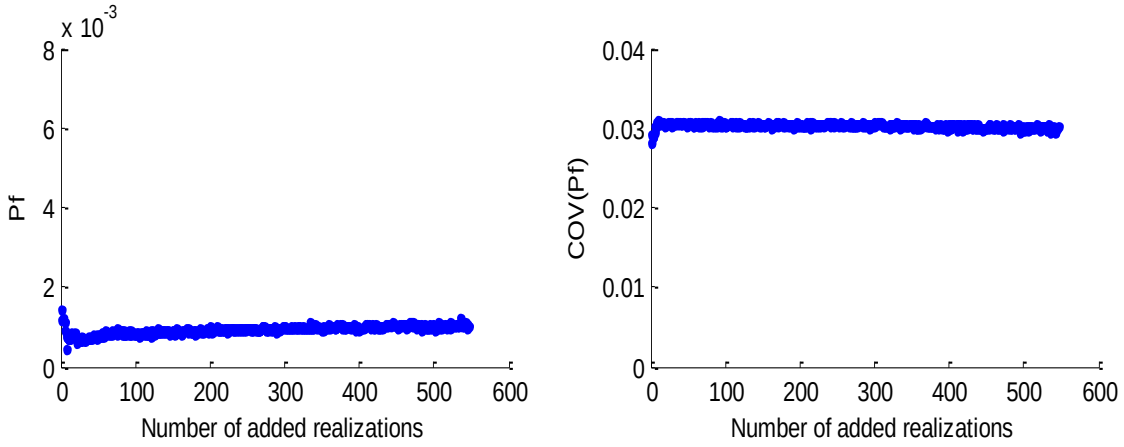


Figure G.8. AK-SS results for a spatially varying soil for the case (a_x=10m, a_y=0.8m)

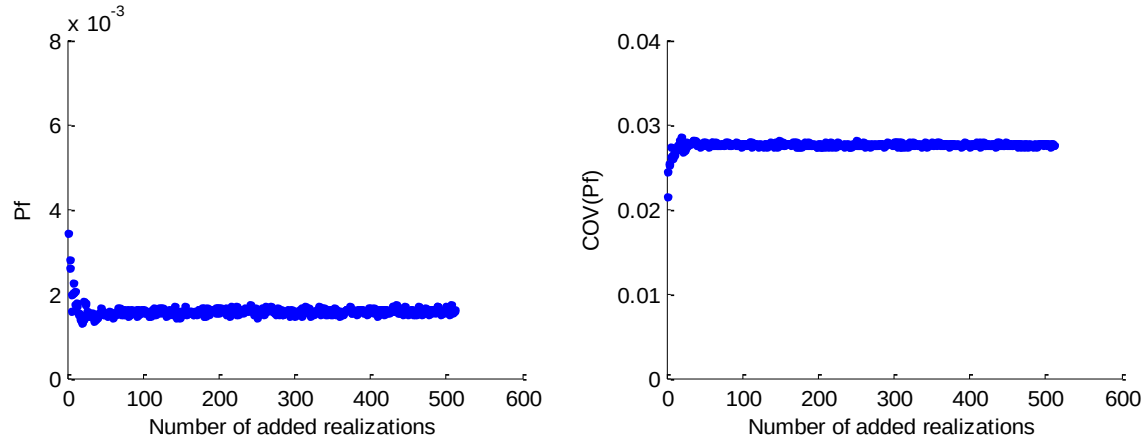


Figure G.9. AK-SS results for a spatially varying soil for the case (a_x=10m, a_y=1m)

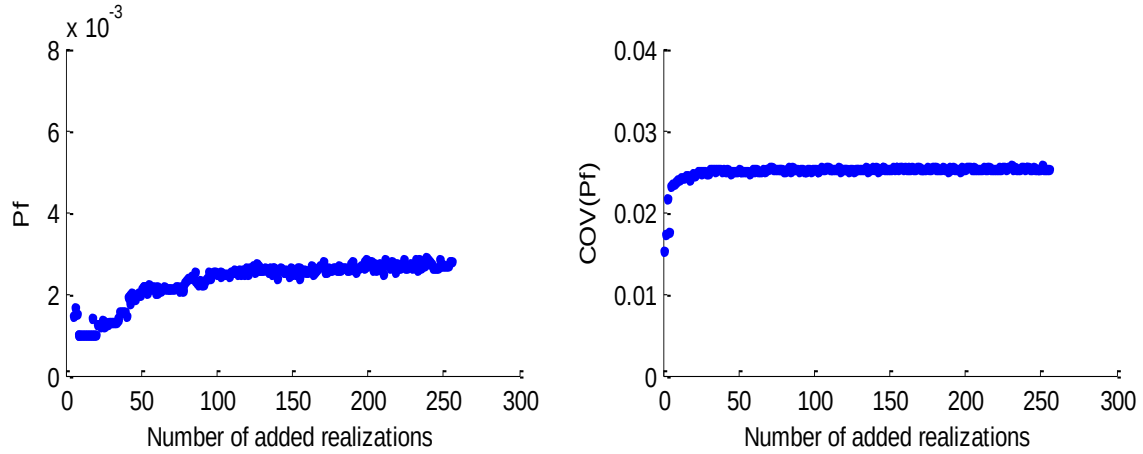


Figure G.10. AK-SS results for a spatially varying soil for the case (a_x=10m, a_y=2m)

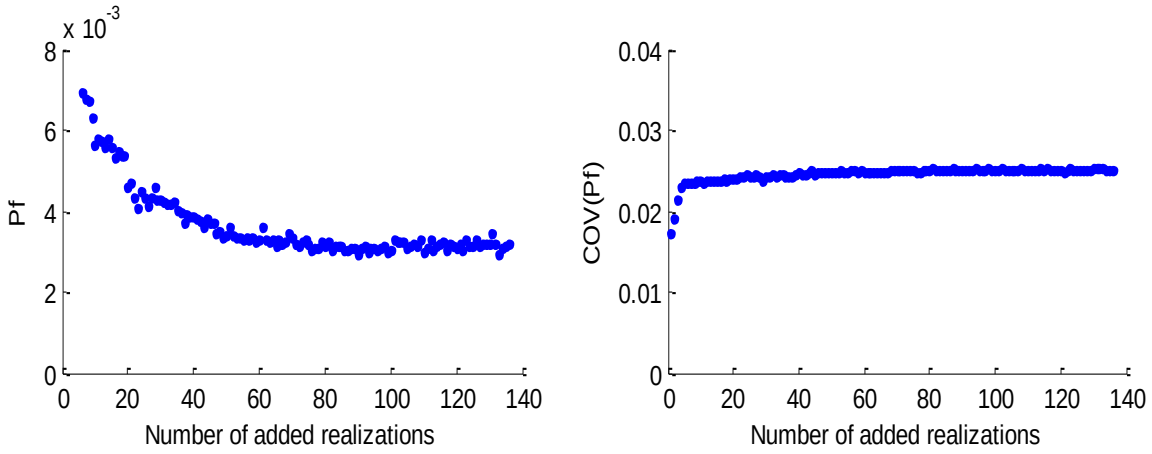


Figure G.11. AK-SS results for a spatially varying soil for the case (a_x=10m, a_y=5m)

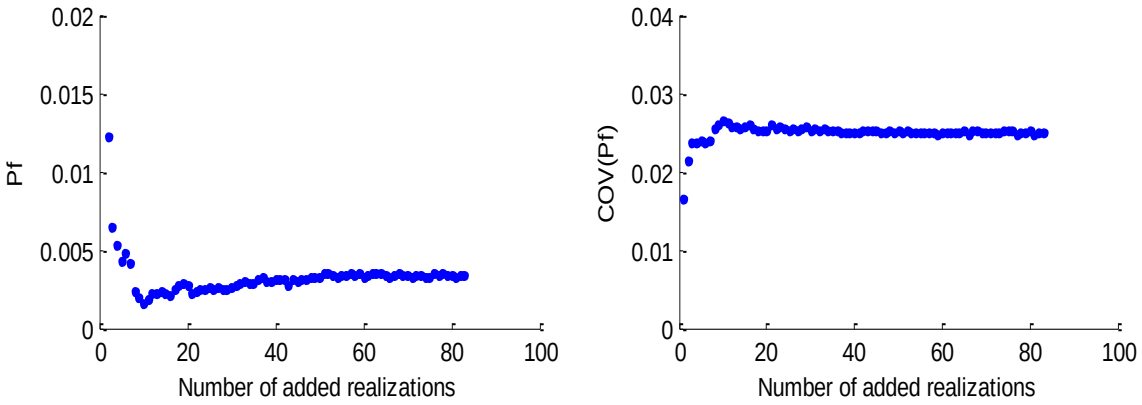


Figure G.12. AK-SS results for a spatially varying soil for the case (a_x=10m, a_y=20m)

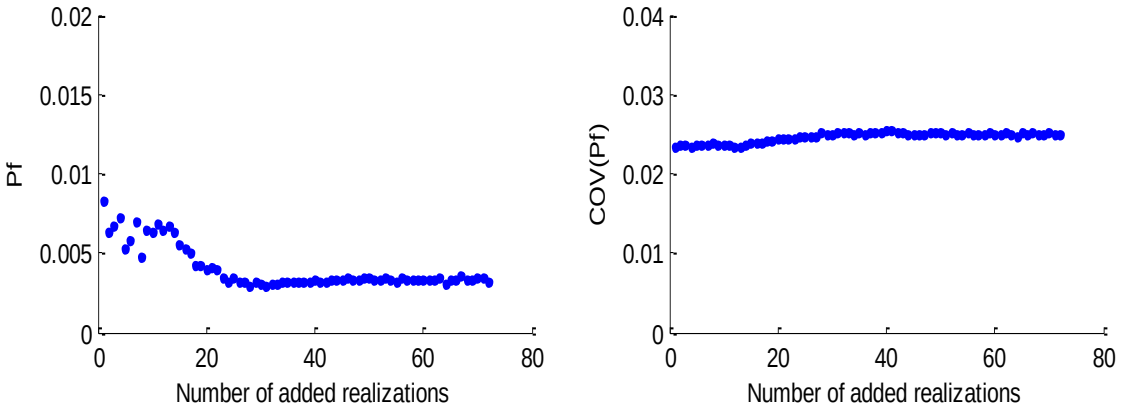


Figure G.13. AK-SS results for a spatially varying soil for the case (a_x=10m, a_y=50m)

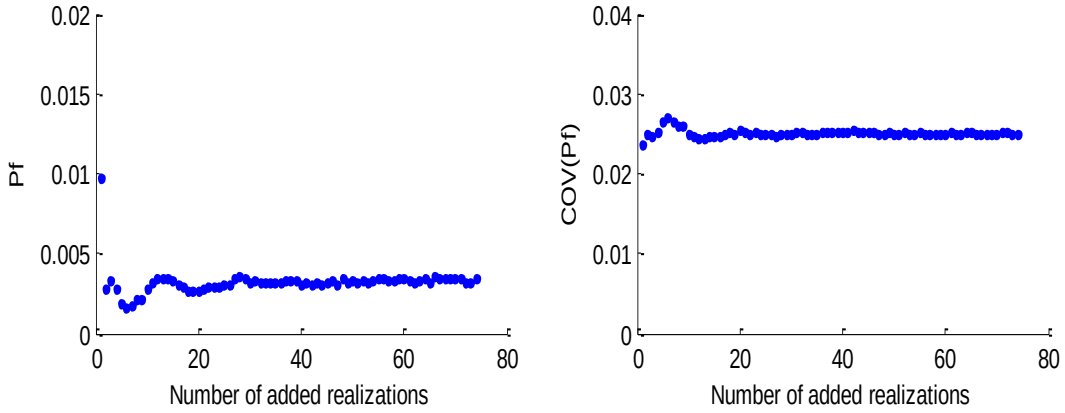


Figure G.14. AK-SS results for a spatially varying soil for the case (a_x=10m, a_y=100m)

Case of an anisotropic case (a_y=2 m with varying a_x)

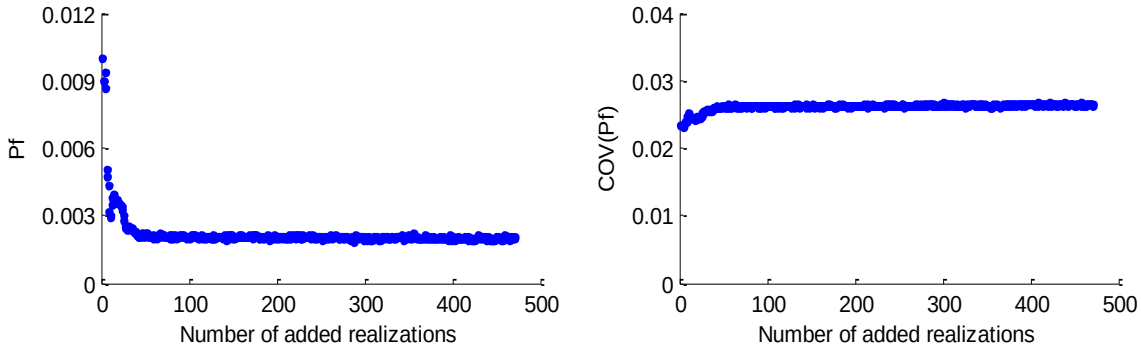


Figure G.15. AK-SS results for a spatially varying soil for the case (a_x=5m, a_y=2m)

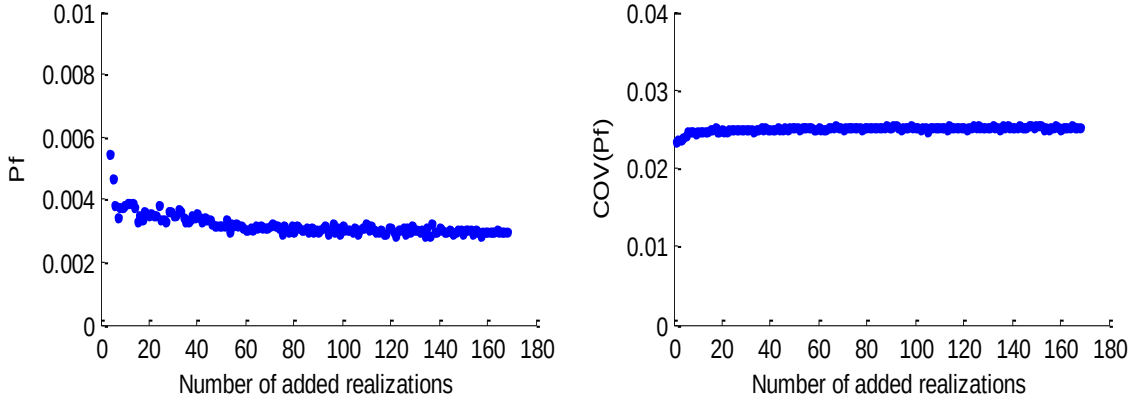


Figure G.16. AK-SS results for a spatially varying soil for the case (a_x=20m, a_y=2m)

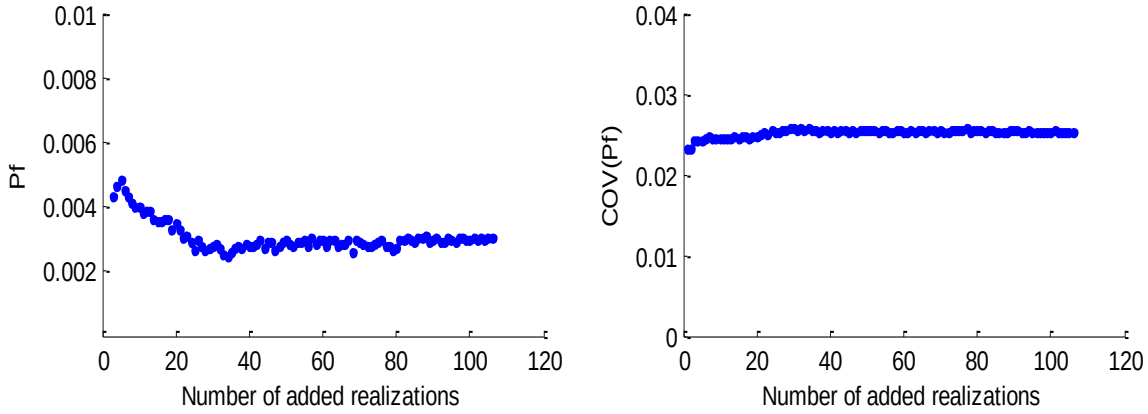


Figure G.17. AK-SS results for a spatially varying soil for the case (a_x=50m, a_y=2m)

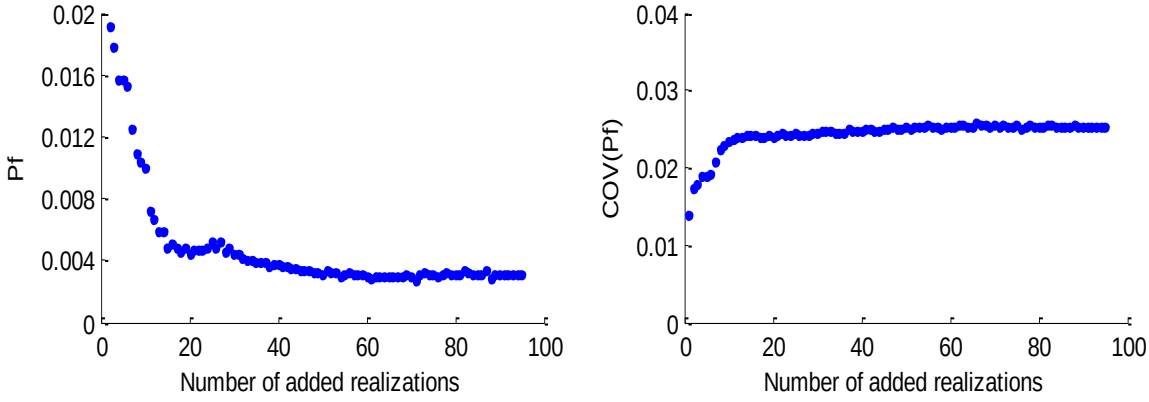


Figure G.18. AK-SS results for a spatially varying soil for the case (a_x=100m, a_y=2m)

UNIVERSITE BRETAGNE LOIRE

UNIVERSITÉ DE NANTES

Thèse de Doctorat

Jawad Kadhim THAJEEL

Kriging-based Approaches for the Probabilistic Analysis of Strip Footings Resting on Spatially Varying Soils

Résumé

L'analyse probabiliste des ouvrages géotechniques est généralement réalisée en utilisant la méthode de simulation de Monte Carlo. Cette méthode n'est pas adaptée pour le calcul des faibles probabilités de rupture rencontrées dans la pratique car elle devient très coûteuse dans ces cas en raison du grand nombre de simulations requises pour obtenir la probabilité de rupture. Dans cette thèse, nous avons développé trois méthodes probabilistes (appelées AK-MCS, AK-IS et AK-SS) basées sur une méthode d'apprentissage (Active learning) et combinant la technique de Krigeage et l'une des trois méthodes de simulation (i.e. Monte Carlo Simulation MCS, Importance Sampling IS ou Subset Simulation SS). Dans AK-MCS, la population est prédite en utilisant un méta-modèle de krigeage qui est défini en utilisant seulement quelques points de la population, ce qui réduit considérablement le temps de calcul par rapport à la méthode MCS. Dans AK-IS, une technique d'échantillonnage plus efficace 'IS' est utilisée. Dans le cadre de cette approche, la faible probabilité de rupture est estimée avec une précision similaire à celle de AK-MCS, mais en utilisant une taille beaucoup plus petite de la population initiale, ce qui réduit considérablement le temps de calcul. Enfin, dans AK-SS, une technique d'échantillonnage plus efficace 'SS' est proposée. Cette technique ne nécessite pas la recherche de points de conception et par conséquent, elle peut traiter des surfaces d'état limite de forme arbitraire. Toutes les trois méthodes ont été appliquées au cas d'une fondation filante chargée verticalement et reposant sur un sol spatialement variable. Les résultats obtenus sont présentés et discutés.

Mots clés
Varabilité spatiale, krigeage, probabilité de ruine, Simulation de Monte Carlo, importance sampling, Subset Simulation.

Abstract

The probabilistic analysis of geotechnical structures involving spatially varying soil properties is generally performed using Monte Carlo Simulation methodology. This method is not suitable for the computation of the small failure probabilities encountered in practice because it becomes very time-expensive in such cases due to the large number of simulations required to calculate accurate values of the failure probability. Three probabilistic approaches (named AK-MCS, AK-IS and AK-SS) based on an Active learning and combining Kriging and one of the three simulation techniques (i.e. Monte Carlo Simulation MCS, Importance Sampling IS or Subset Simulation SS) were developed. Within AK-MCS, a Monte Carlo simulation without evaluating the whole population is performed. Indeed, the population is predicted using a kriging meta-model which is defined using only a few points of the population thus significantly reducing the computation time with respect to the crude MCS. In AK-IS, a more efficient sampling technique 'IS' is used instead of 'MCS'. In the framework of this approach, the small failure probability is estimated with a similar accuracy as AK-MCS but using a much smaller size of the initial population, thus significantly reducing the computation time. Finally, in AK-SS, a more efficient sampling technique 'SS' is proposed. This technique overcomes the search of the design points and thus it can deal with arbitrary shapes of the limit state surfaces. All the three methods were applied to the case of a vertically loaded strip footing resting on a spatially varying soil. The obtained results are presented and discussed.

Key Words
spatial variability, Kriging metamodeling, probability of failure, Monte Carlo Simulation, Importance Sampling, Subset Simulation.