
Contribution à la méthode de la surface de réponse stochastique - Application à l'analyse de stabilité d'un tunnel

Guilhem Mollon* — Daniel Dias* — Abdul-Hamid Soubra **

* *Laboratoire de Génie Civil et d'Ingénierie Environnementale, INSA Lyon, Université de Lyon, Domaine Scientifique de la Doua, 20 Avenue Einstein, 69621 Villeurbanne*

** *GeM, Université de Nantes, Boulevard de l'Université, BP152, 44603 Saint-Nazaire cedex*

RÉSUMÉ. La méthode de la surface de réponse stochastique par collocation (CSRSM) est décrite et détaillée dans le cas particulier d'un modèle à deux variables aléatoires. Cette méthode est appliquée à un modèle déterministe d'analyse de la stabilité d'un front de taille pressurisé. Une validation est effectuée par comparaison entre les distributions de sortie du modèle déterministe et de son approximation par CSRSM. Dans une dernière partie, une extension de la CSRSM est proposée, permettant de réaliser une étude paramétrique sur les caractéristiques probabilistes des variables d'entrée pour un faible nombre d'appels au modèle déterministe.

ABSTRACT. The Collocation-based Stochastic Response Surface Methodology (CSRSM) is described and detailed in the case of a model with two random variables. This method is applied to a deterministic model for the stability analysis of a pressurized tunnel face. A validation is carried out by comparison between the output distributions of the deterministic model and of its approximation by CSRSM. Finally, an extension of the CSRSM is proposed, allowing the realization of a parametric study on the probabilistic characteristics of the input variables, for a limited number of calls of the deterministic model.

MOTS-CLÉS : Méthode non intrusive, CSRSM, étude paramétrique, Simulation de Monte-Carlo, Propagation d'incertitude

KEYWORDS: Non-intrusive methodology, CSRSM, parametric study, Monte-Carlo Simulation, Uncertainty propagation

1. Introduction

La méthode de surface de réponse stochastique par collocation (abrégée par CSRSM pour *Collocation-Based Stochastic Response Surface Method*) est une méthode non intrusive qui permet d'étudier un modèle numérique complexe de manière probabiliste et fiable en le remplaçant par un méta-modèle analytique approché. On s'intéresse dans cette communication à la CSRSM par régression. Cette méthode est très performante pour caractériser de manière fiable un modèle complexe dont les paramètres d'entrée sont bien connus, mais présente dans son formalisme classique l'inconvénient d'imposer un nouveau jeu de calculs déterministes au moindre changement dans l'un des paramètres d'entrées (comme par exemple la corrélation entre deux variables aléatoires d'entrée). Cet état de fait empêche la réalisation d'une étude paramétrique dans un temps raisonnable, alors même que les propriétés statistiques des variables d'entrée sont souvent assez mal connues, et donc qu'une étude paramétrique est souvent nécessaire.

Cette communication propose une extension de la CSRSM dite « classique », permettant l'utilisation d'un seul et même jeu de résultats déterministes pour plusieurs jeux de paramètres statistiques des variables d'entrée. Cette approche repose sur le fait que la détermination des coefficients du chaos polynomial par régression autorise l'utilisation de points de collocation variables dans l'espace standard. Ceci permet une étude paramétrique complète à partir d'un nombre très réduit d'appels au modèle déterministe. Le formalisme classique de la CSRSM (Isukapalli 1999, Sudret et al. 2006, Sudret 2007, Phoon et Huang 2007, Mollon et al. 2010) et le modèle déterministe utilisé dans cette étude sont d'abord présentés, puis la CSRSM est validée par une approche de type Monte-Carlo. Finalement, une extension de la CSRSM est proposée pour l'étude paramétrique des caractéristiques probabilistes des variables d'entrée.

2. Présentation du modèle déterministe

Le modèle déterministe utilisé comme exemple dans cette étude est analytique et repose sur le théorème cinématique de l'analyse limite. Celui-ci permet la détermination de la pression nécessaire à appliquer au front de taille d'un tunnel percé en milieu frottant et/ou cohérent pour en garantir la stabilité (Mollon et al. 2009). Il s'agit d'un mécanisme tridimensionnel de rupture du front par plusieurs blocs de sol en translation. Les différents blocs sont de forme conique et subissent une translation selon leur axe. Les conditions de normalité imposées par l'analyse limite sont donc respectées du fait que les surfaces de discontinuité de vitesse (i.e. les surfaces externes des blocs en mouvement) forment toujours un angle φ avec le vecteur vitesse, φ étant l'angle de frottement interne du sol. Le modèle proposé donne des valeurs de la pression de rupture cohérentes avec les simulations numériques (Mollon et al. 2009).

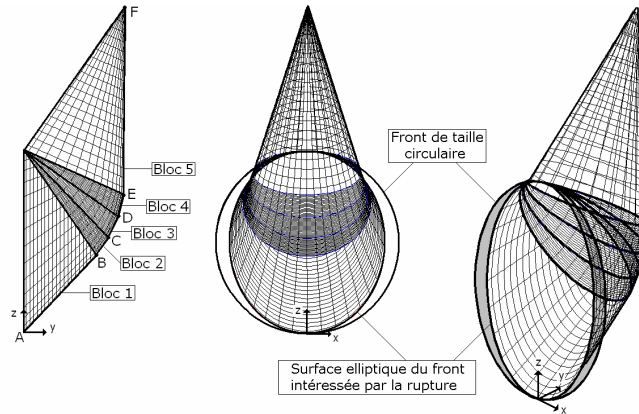


Figure 1. Allure du mécanisme de ruine proposé par le modèle déterministe

Il peut être prouvé que l'augmentation du nombre de blocs n'améliore que très peu la solution au-delà de 5, un mécanisme à 5 blocs est donc utilisé dans cette étude. Ce modèle relativement simple, implémenté sous Excel en Visual Basic, a été choisi par commodité pour son temps de calcul quasi-instantané (0.1s). La variable de sortie du modèle est donc la pression limite d'effondrement du front de taille (nommée σ_c et exprimée en kPa), tandis que les variables aléatoires d'entrée sont au nombre de deux : l'angle de frottement interne du sol (ϕ), et la cohésion (c). Dans une approche probabiliste, ces deux variables peuvent être munies de différentes lois de distribution, et être corrélées ou non.

3. Présentation de la CSRSM

De nombreux problèmes d'ingénierie nécessitent le recours à des modèles numériques « lourds », reposant par exemple sur la méthode des éléments finis, et dont les temps de calculs se prêtent mal à une étude probabiliste. L'objectif de la CSRSM est de définir un méta-modèle approché qui vient se substituer au modèle déterministe dans les études probabilistes, et dont le temps de calcul est quasi-instantané. Ce méta-modèle (ou PCE pour *Polynomial Chaos Expansion*) s'exprime dans une base nommée chaos polynomial. Pour un chaos polynomial d'ordre n , les vecteurs de cette base sont des polynômes orthogonaux (Hermite, Legendre, etc.) multidimensionnels de degrés $\leq n$. Pour un jeu donné de paramètres d'entrée (i.e. corrélations et distributions statistiques des différentes variables aléatoires d'entrée), on peut déterminer les coefficients du chaos par un nombre réduit d'appels au modèle déterministe. Dans l'étude proposée, le chaos est formé par des polynômes d'Hermite multidimensionnels, et les coefficients du chaos sont déterminés par régression à partir de la réponse du modèle déterministe en un certain nombre de points de collocation. Cette partie décrit brièvement l'utilisation de la CSRSM dans le cas de deux variables. Pour une présentation plus générale, le lecteur peut se pencher sur Isukapalli (1999), ou Sudret (2007).

Les deux variables aléatoires (φ et c) doivent être représentées dans le chaos polynomial par deux variables standards (variables normales centrées et réduites) nommées respectivement ξ_1 et ξ_2 . Pour un PCE d'ordre n donné, les points de collocation disponibles sont déterminés dans l'espace standard (ξ_1, ξ_2) : chacune des variables standards peut prendre les valeurs des racines du polynôme d'Hermite de degré $n+1$, et les points de collocation disponibles résultent de l'ensemble des combinaisons de ces racines. Il est à noter que cette méthode de choix des points de collocation n'est pas obligatoire, mais correspond à la pratique usuelle de la CSRSM. Pour deux variables standards, la variable de sortie U peut être calculée à partir du chaos polynomial d'ordre n comme suit :

$$U = \sum_{i=1}^p a_i \cdot \Gamma_i(\xi_1, \xi_2) \quad (1)$$

Dans cette expression, les termes Γ_i sont des polynômes d'Hermite multidimensionnels, qui constituent les vecteurs de la base du chaos polynomial. Les coefficients a_i du chaos sont les inconnues du problème. Pour un PCE d'ordre n et un nombre n_v de variables aléatoires d'entrée, le nombre p de coefficients et le nombre M de points de collocation disponibles sont donnés par (Isukapalli 1999) :

$$p = \frac{(n + n_v)!}{n! n_v!} \quad (2)$$

$$M = (n+1)^{n_v} + \begin{cases} 0 & \text{si l'une des racines du polynôme d'Hermite de degré } n+1 \text{ est nulle} \\ 1 & \text{sinon} \end{cases} \quad (3)$$

Ces expressions permettent de remarquer que, dès lors que l'on considère un nombre important de variables aléatoires, le nombre de points disponibles est bien plus élevé que le nombre de coefficients à déterminer. Il est donc nécessaire d'opérer un choix judicieux des points de collocation pour lesquels le calcul déterministe sera effectué, afin de réduire les temps de calcul sans affecter la qualité de l'approximation. Isukapalli (1999) propose d'utiliser un nombre de points de collocation environ deux fois supérieur au nombre de coefficients du PCE. Il est également recommandé par le même auteur de disposer ces points symétriquement, et aussi près de l'origine que possible. Enfin, il est proposé d'ajouter aux points de collocation l'origine de l'espace des variables standard pour le cas où le polynôme d'Hermite de degré $n+1$ n'aurait pas de racine nulle, comme explicité dans l'équation (3). Sudret (2007) a écrit un algorithme de choix des points de collocation, reposant sur l'inversibilité de la matrice $(\underline{N}^t \cdot \underline{N})$, \underline{N} étant définie dans l'équation (7). Dans la présente étude, le faible nombre de variables implique que $2p > M$ quelque soit l'ordre du PCE, il sera donc nécessaire d'utiliser tous les points de collocation disponibles.

Pour être introduits dans le modèle déterministe, les points de collocation (sous forme de couples $(\xi_{1,m}, \xi_{2,m})$ avec m compris entre 1 et M) doivent d'abord être exprimés dans l'espace des variables physiques (couples (φ_m, c_m) correspondants), en utilisant les expressions suivantes :

$$\begin{bmatrix} \xi_{1C,m} \\ \xi_{2C,m} \end{bmatrix} = H \cdot \begin{bmatrix} \xi_{1,m} \\ \xi_{2,m} \end{bmatrix} \quad (4)$$

$$\begin{cases} \varphi_m = F_\varphi^{-1}[\Phi(\xi_{1C,m})] \\ c_m = F_c^{-1}[\Phi(\xi_{2C,m})] \end{cases} \quad (5)$$

avec H =transformée de Cholesky de la matrice de corrélation de c et φ , Φ =fonction de répartition normale centrée réduite, F_φ = fonction de répartition de φ , et F_c = fonction de répartition de c . Après obtention de la réponse du modèle déterministe aux points de collocation, la détermination des coefficients du PCE s'effectue par régression, en résolvant l'équation matricielle suivante :

$$\underline{N}^t \cdot \underline{N} \cdot \underline{a} = \underline{N}^t \cdot \underline{f} \quad (6)$$

avec a =vecteur colonne des coefficients a_i , f =vecteur colonne des réponses du modèle aux points de collocation, et N la matrice suivante :

$$\underline{N} = \begin{bmatrix} \Gamma_{0,0} & \Gamma_{1,0}(\xi_{1,1}) & \Gamma_{0,1}(\xi_{2,1}) & \Gamma_{2,0}(\xi_{1,1}) & \Gamma_{1,1}(\xi_{1,1}, \xi_{2,1}) & \Gamma_{0,2}(\xi_{2,1}) \\ \Gamma_{0,0} & \Gamma_{1,0}(\xi_{1,2}) & \Gamma_{0,1}(\xi_{2,2}) & \Gamma_{2,0}(\xi_{1,2}) & \Gamma_{1,1}(\xi_{1,2}, \xi_{2,2}) & \Gamma_{0,2}(\xi_{2,2}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Gamma_{0,0} & \Gamma_{1,0}(\xi_{1,M}) & \Gamma_{0,1}(\xi_{2,M}) & \Gamma_{2,0}(\xi_{1,M}) & \Gamma_{1,1}(\xi_{1,M}, \xi_{2,M}) & \Gamma_{0,2}(\xi_{2,M}) \end{bmatrix} \quad (7)$$

La connaissance des coefficients du PCE fournit finalement une approximation analytique du modèle déterministe. Le post-traitement de ce PCE consiste dans la présente étude à déterminer la distribution statistique de la réponse par une méthode de Monte-Carlo. On peut aussi envisager d'appliquer à ce meta-modèle les outils de l'analyse fiabiliste (Monte-Carlo, Tirages d'Importance, FORM) afin d'étudier la probabilité de ruine pour une fonction de performance donnée.

4. Validation de la CSRSM

Dans cette section, le méta-modèle fourni par la CSRSM est comparé au modèle déterministe original afin d'évaluer son exactitude et de choisir l'ordre optimal du PCE. Pour cette comparaison, on utilise des données de références pour les deux variables d'entrée, qui sont considérées gaussiennes et non corrélées, avec les

caractéristiques suivantes : $\mu_\phi=17^\circ$ et $COV(\phi)=10\%$; $\mu_c=7$ kPa et $COV(c)=20\%$. Pour ce cas de référence, des tirages de Monte-Carlo sont effectués avec (i) les méta-modèles d'ordres 2, 3, 4, et 5, et (ii) le modèle déterministe original. Cette démarche est rendue possible par le choix d'un modèle analytique au temps de calcul assez faible (0.1s). Le nombre de tirages est de 10^6 .

La figure 2a présente la densité de probabilité de la réponse pour le modèle déterministe et les 4 méta-modèles. Seul le PCE d'ordre 2 peut être discerné, tandis que les PCE d'ordres 3 et supérieurs fournissent une courbe de densité de probabilité confondue avec celle du modèle analytique. La figure 2b propose la probabilité de ruine du tunnel en queue de distribution et apporte plus de précision. On observe que l'ordre 2 donne des résultats peu satisfaisants (même si l'ordre de grandeur est correct), mais que les ordres supérieurs se rapprochent de la probabilité de ruine obtenue pour le modèle analytique. Les ordres 4 et 5 fournissent des probabilités de ruines très similaires aux résultats obtenus par FORM lors d'une précédente étude (Mollon et al. 2009), et assez satisfaisantes.

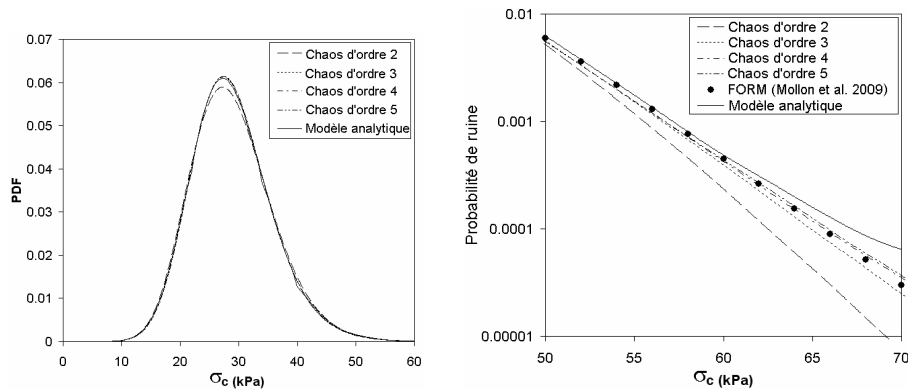


Figure 2. (a) Comparaison entre les PDF obtenus pour le modèle analytique et pour les PCE d'ordres 2 à 5 ; (b) Comparaison entre les probabilités de ruine obtenues pour le modèle analytique et pour les PCE d'ordres 2 à 5, ainsi que les valeurs obtenues par FORM dans une précédente étude

Tableau 1. Estimateurs des moments de la réponse

	Moyenne (kPa)	Variance (kPa ²)	Asymétrie	Aplatissement
PCE d'ordre 2	28.92	47.90	0.496	0.409
PCE d'ordre 3	28.93	46.12	0.555	0.703
PCE d'ordre 4	28.92	46.00	0.552	0.744
PCE d'ordre 5	28.92	46.02	0.552	0.749
Modèle analytique	28.95	46.13	0.537	0.766
COV de l'estimateur	~0.02%	~0.14%	/	/

Le tableau 1 présente les quatre premiers moments statistiques de la distribution de la variable de sortie, pour les PCE d'ordres 2 à 5 et pour le modèle analytique. Il apparaît que l'ordre 3 est suffisant pour obtenir une convergence sur la moyenne, la variance, et l'asymétrie, mais qu'il faut monter à l'ordre 4 pour que la convergence soit satisfaisante également en terme de coefficient d'aplatissement. Il faut noter que les moments présentés ici ont été estimés par Monte-Carlo, mais qu'ils peuvent également être calculés analytiquement à partir des coefficients du PCE (Sudret 2007). Les coefficients de variation des estimateurs de la moyenne et de la variance fournis dans le tableau 1 sont valables pour tous les PCE ainsi que pour le modèle analytique, et attestent d'une très bonne convergence du tirage de Monte-Carlo.

La figure 3 présente les surfaces de réponses fournies dans le plan (φ, c) , en terme de lignes d'égaux valeurs de la pression d'effondrement (en kPa). Ces surfaces sont tracées pour le modèle analytique, ainsi que pour les PCE d'ordre 2 et 4. La correspondance est très satisfaisante pour les deux méta-modèles au centre de l'espace physique (c'est-à-dire dans la zone de forte probabilité d'occurrence des variables φ et c), mais l'ordre 4 est plus performant sur les zones « périphériques », correspondant aux queues de distribution. Ceci explique clairement les résultats fournis par la figure 2b. L'ordre 4 sera donc choisi pour conduire la suite de cette étude.

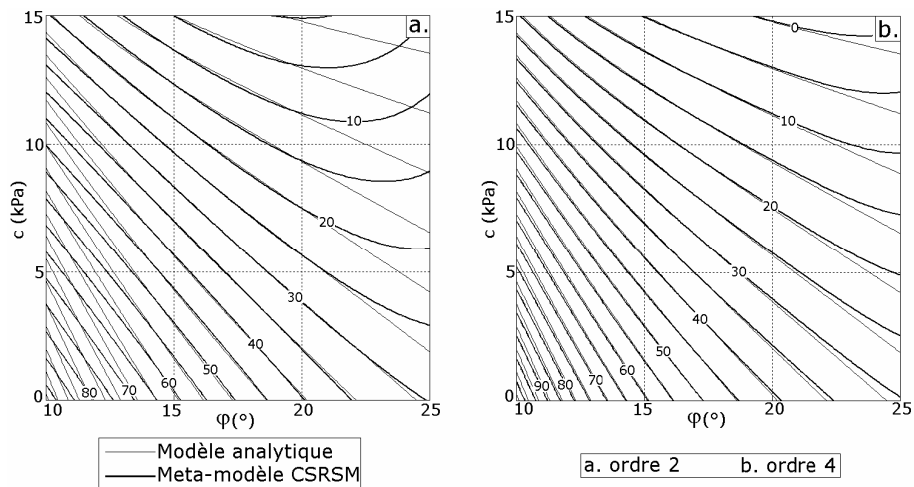


Figure 3. Comparaison entre les surfaces de réponses du modèle analytique et des PCE d'ordres 2 et 4 dans l'espace physique

5. Extension de la CSRSM pour l'étude paramétrique des variables d'entrée

Si la CSRSM apparaît comme un outil à la fois précis et efficace pour l'étude probabiliste d'un modèle déterministe, son formalisme classique présente néanmoins une légère limitation. L'observation des équations (4) et (5) montre ainsi que la

position des points de collocation dans l'espace standard est liée aux distributions des variables aléatoires d'entrée (par l'intermédiaire de H , F_φ et F_c). Dans le cas d'une étude paramétrique des caractéristiques probabilistes des variables d'entrée, il est donc nécessaire d'effectuer un nouveau jeu de calculs déterministes à chaque fois que l'on fait varier l'une de ces caractéristiques (telle que le COV d'une variable ou la corrélation entre deux variables). Cette section propose une extension de la CSRSM conventionnelle, qui repose sur l'unicité de la surface de réponse analytique dans l'espace des variables physiques. Lorsqu'un jeu de calculs déterministes a été effectué et qu'une surface de réponse approchée a été déterminée sous forme de PCE, il semble en effet opportun de conserver cette surface de réponse pour des études ultérieures. La figure 4a présente les surfaces de réponse de deux PCE différents dans l'espace des variables physiques. Il s'agit de deux PCE d'ordre 4 issus de la CSRSM conventionnelle pour deux jeux différents des caractéristiques probabilistes des variables d'entrée : le cas de référence (variables normales et non corrélées, $\text{COV}(\varphi)=10\%$ et $\text{COV}(c)=20\%$), et un cas arbitraire (variables suivant une loi lognormale, avec un coefficient de corrélation $\rho_{\varphi c}=-0.3$, $\text{COV}(\varphi)=8\%$ et $\text{COV}(c)=25\%$). Les moyennes des variables restent inchangées entre les deux cas. On observe que les deux surfaces de réponse sont très similaires.

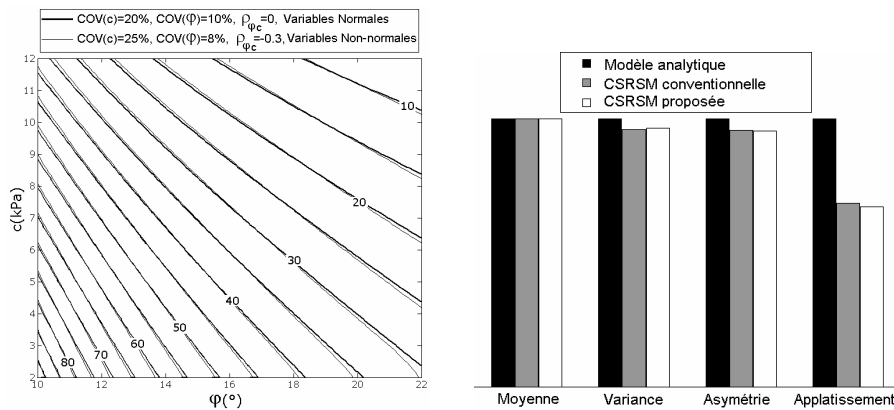


Figure 4. (a) Comparaison entre les surfaces de réponse dans l'espace physique, obtenues par application de la CSRSM pour deux jeux de paramètres probabilistes ; (b) Moments statistiques de la réponse pour le modèle analytique, pour le PCE obtenu par CSRSM conventionnelle, et pour le PCE obtenu par la méthode proposée (en échelle normalisée)

Il est donc proposé de n'effectuer les calculs déterministes que sur un seul jeu de paramètres probabilistes. Il est souhaitable que ce jeu corresponde à des caractéristiques probabilistes « classiques » pour les variables physiques étudiées (ce qui est le cas par exemple pour le jeu de référence utilisé dans la présente étude). Pour étudier un deuxième jeu de caractéristiques probabilistes, il faut calculer les positions des points de collocation dans l'espace standard qui *auraient conduit* aux points de calcul déjà effectués. Ces positions $(\xi'_{1,m}, \xi'_{2,m})$ s'obtiennent par :

$$\begin{cases} \xi'_{1C,m} = \Phi^{-1}(F_\varphi(\varphi_m)) \\ \xi'_{2C,m} = \Phi^{-1}(F_c(c_m)) \end{cases} \quad (8)$$

$$\begin{bmatrix} \xi'_{1,m} \\ \xi'_{2,m} \end{bmatrix} = H^{-1} \begin{bmatrix} \xi'_{1C,m} \\ \xi'_{2C,m} \end{bmatrix} \quad (9)$$

Par cette opération, on dispose d'un ensemble de points de collocation $(\xi'_{1,m}, \xi'_{2,m})$ dans l'espace standard, et des réponses correspondantes du modèle déterministe, sans avoir effectué le moindre calcul supplémentaire. L'équation (6) peut donc être utilisée pour calculer les coefficients d'un nouveau PCE. La seule différence avec la CSRSM conventionnelle réside dans le fait que les points de collocation dans l'espace standard ne correspondent plus à des combinaisons des racines du polynôme d'Hermite de degré $n+1$.

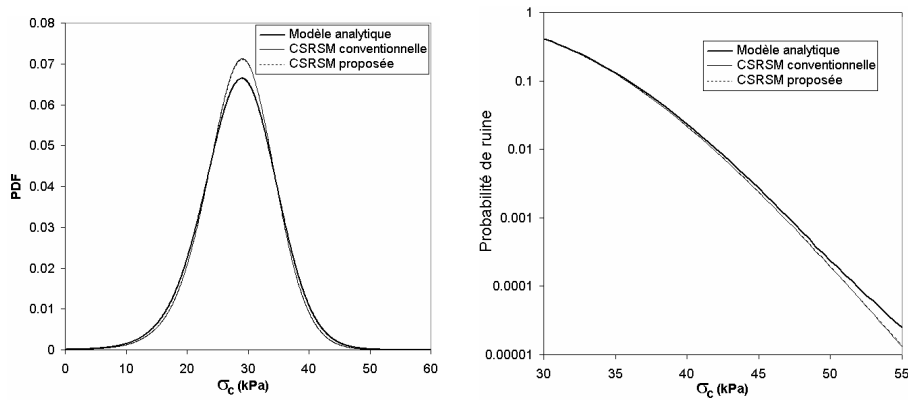


Figure 5. (a) Comparaison entre les PDF de la réponse pour le modèle analytique, pour le PCE obtenu par CSRSM conventionnelle, et pour le PCE obtenu par la méthode proposée ; (b) comparaison entre les probabilités de ruine

La figure 4b compare les quatre premiers moments statistiques de la réponse (en échelle normalisée), pour le cas arbitraire déjà évoqué (variables lognormales, avec un coefficient de corrélation $\rho_{\varphi c} = -0.3$, $\text{COV}(\varphi) = 8\%$ et $\text{COV}(c) = 25\%$). Ces moments sont calculés par tirages de Monte-Carlo (10^6 échantillons) pour trois modèles différents : (i) le modèle analytique, (ii) le PCE d'ordre 4 obtenu par CSRSM conventionnelle (c'est-à-dire impliquant 25 nouveaux appels au modèle déterministe), et (iii) le PCE d'ordre 4 obtenu par la CSRSM proposée (c'est-à-dire réutilisant les résultats du cas de référence de cette étude, et n'impliquant donc aucun nouvel appel au modèle déterministe). Il apparaît que les deux méta-modèles donnent des résultats très comparables pour les quatre moments. La figure 5a montre la densité de probabilité obtenue pour ces trois modèles (à partir des mêmes tirages

de Monte-Carlo), et la figure 5b montre la probabilité de ruine (P_f) en queue de distribution. Sur ces deux graphiques, les deux méta-modèles sont presque confondus. Ces résultats prouvent que l'extension de la CSRSM proposée fournit des résultats de qualité comparable à ceux de la CSRSM conventionnelle, pour un coût de calcul supplémentaire quasi-nul. Elle présente en outre l'avantage de conserver le formalisme de la CSRSM, ce qui permet d'utiliser les propriétés naturelles du PCE (possibilité de calculer analytiquement les moments statistiques à partir des coefficients du PCE, indices de Sobol', etc. [Sudret 2007]). Ces résultats confirment que la méthode de choix des points de collocation utilisant les racines du polynôme d'Hermite d'ordre supérieur n'est pas strictement indispensable. Néanmoins, il faut préciser que la méthode proposée ne reste valide que si les régions de l'espace standard couvertes par les points de collocation d'origine et modifiés sont comparables. On s'assurera donc que les variables aléatoires de la première étude sont les plus réalistes possibles, et que les changements de paramètres ou de distribution restent réduits.

6. Conclusion

Cette communication rappelle les principes de la Méthode de la Surface de Réponse Stochastique par Collocation, et valide cette méthode par l'intermédiaire d'un modèle analytique relativement simple d'étude de la stabilité d'un front de taille pressurisé. Cette validation a été effectuée par comparaison entre les résultats probabilistes obtenus par tirages de Monte-Carlo sur le modèle analytique et sur son approximation par chaos polynomial. Un ordre 4 pour ce chaos est jugé optimal en termes de précision et de temps de calcul. Dans une deuxième partie, une extension de cette méthode est proposée, permettant la réutilisation d'un ensemble de calculs déterministes pour un nouveau jeu de caractéristiques probabilistes des variables d'entrée, sans faire appel de nouveau au modèle déterministe. Cette méthode ouvre la voie à des études paramétriques de modèles déterministes « lourds » (de type éléments finis par exemple), pour un coût calculatoire relativement faible.

7. Références Bibliographiques

- Isukapalli, S.S. (1999). An uncertainty analysis of transport-transformation models, *Ph.D. thesis*, The State University of New Jersey, New Brunswick, New Jersey.
- Mollon, G., Dias, D., and Soubra, A.-H. (2009). "Probabilistic analysis and design of circular tunnels against face stability." *Int. J. of Geomech., ASCE*, 9(6), 237-249.
- Mollon, G., Dias, D., and Soubra, A.-H. (2010). "Probabilistic Analysis of Pressurized Tunnels Against Face Stability Using Collocation-based Stochastic Response Surface Method" *J. of Geotech. & Geoenv. Engrg.*, Under review.
- Phoon, K.K., and Huang, S.P. (2007). "Geotechnical probabilistic analysis using collocation-based stochastic response surface method", *Applications of Statistics and Probability in Civil Engineering*, Kanda, Takada and Furada (eds.), Taylor and Francis Group, London.
- Sudret, B., Berveiller, M., and Lemaire, M. (2006). "A stochastic finite element procedure for moment and reliability analysis", *Eur. J. Comput. Mech.*, (15)7-8, 825-866.
- Sudret, B. (2007). "Global sensitivity analysis using polynomial chaos expansion.", *Reliability Engineering and System Safety*, 93, 964-979.