

Proposition de post-doctorat pour 2014 :

« Développement d'outils stochastiques pour l'analyse de sensibilité en grande dimension des simulateurs numériques »

Durée du post-doc : 12 mois, à partir du 1^{er} trimestre 2014

Localisation : Laboratoire d'études de Sûreté et de Maîtrise des Risques (LSMR), CEA de Cadarache

Encadrant CEA: Amandine MARREL (DEN/SESI/LSMR)

Encadrant Universitaire : Béatrice LAURENT (INSA de Toulouse / IMT de Toulouse)

Description du sujet

Titre détaillé : « Estimation d'indices de sensibilité basés sur les dérivées à partir du métamodèle de krigeage : application à l'analyse de sensibilité des simulateurs numériques en grande dimension. »

De nombreux phénomènes physiques sont représentés par des équations déterministes qui conduisent, lors de la phase de modélisation numérique, à l'obtention de simulateurs numériques ou de codes de calcul. Ces outils logiciels prennent en entrée un grand nombre de paramètres caractéristiques du phénomène étudié. L'ensemble de ces paramètres sont entachés d'une incertitude plus ou moins importante selon le degré de connaissance et de caractérisation du phénomène modélisé. Il est alors important de prendre en compte ces incertitudes et d'étudier, au travers d'une analyse de sensibilité, comment elles se répercutent sur les sorties du code. Cela permet de valider le modèle mathématique, physique ou numérique, d'orienter des efforts de caractérisation de certains paramètres, de simplifier le modèle et d'améliorer la compréhension du phénomène modélisé, en identifiant des relations entre les variables d'entrée et les variables de sortie (*e.g.* détection d'interactions).

Les méthodes d'analyse de sensibilité globale (Iooss [2009]) permettent de quantifier l'influence de chacune de ces variables d'entrée sur la sortie du code. Elles s'appuient généralement sur une décomposition de la variance de la sortie qui conduit à la définition des indices de Sobol (Sobol [1993], Saltelli et al. [2000]). Une estimation empirique de ces indices nécessite plusieurs milliers d'évaluations du modèle pour chaque paramètre (*i.e.* de simulations du code), ce qui n'est généralement pas réalisable compte-tenu du grand nombre de paramètres et du temps de calcul nécessaire pour chaque simulation.

Récemment, des indices de sensibilité basés sur la dérivée du modèle et appelés DGSM pour « Derivative based Global Sensitivity Measures » ont été proposés (Sobol & Kucherenko [2009], Lamboni et al. [2013]). L'estimation de ces indices nécessite moins de simulations (le nombre de simulations est indépendant du nombre de paramètres incertains) et permettent d'identifier à moindre coût les variables d'entrée non influentes. L'objectif est donc d'utiliser ces indices pour réaliser une première sélection des paramètres potentiellement influents dans le cas des modèles présentant un grand nombre de paramètres d'entrée incertains. Cependant, l'estimation de ces indices nécessite quand même plusieurs centaines de simulations. Les codes de calcul utilisés lors des études de sûreté sont souvent complexes et coûteux en temps de calcul. Par conséquent, leur exploitation directe pour estimer les indices DGSM peut s'avérer difficile. Une solution est alors de remplacer le code de calcul par un modèle simplifié appelé métamodèle, qui requiert un temps de calcul réduit et qui est représentatif du code de calcul dans le domaine d'étude. Parmi les métamodèles classiquement utilisés pour l'approximation des simulateurs numériques, le krigeage (ou processus gaussien) s'avère très performant dans de nombreuses applications (Sacks et al. [1989], Oakley & O'Hagan [2002], Marrel [2008]). De plus, ce métamodèle présente l'avantage de fournir un prédicteur non seulement pour le modèle (sortie du code de calcul) mais aussi pour la dérivée du modèle. Le calcul des DGSM faisant intervenir la dérivée du modèle, un calcul analytique de ces indices à partir du métamodèle de krigeage est à étudier. La variance du prédicteur de krigeage permet aussi de disposer d'un estimateur de l'erreur du métamodèle et, par conséquent, de l'erreur sur les indices DGSM estimés.

L'objectif du stage post-doctoral est donc d'utiliser le métamodèle de krigeage pour estimer le plus efficacement possible les indices DGSM. La variance du prédicteur du métamodèle pourra être utilisée pour associer un intervalle de confiance aux indices DGSM estimés. A partir des résultats obtenus, un screening (aussi appelé criblage) pourra être réalisé : il s'agira d'identifier les paramètres non influents et les paramètres potentiellement influents.

Plusieurs applications sont envisagées pour mettre en œuvre la méthodologie et les outils développés :

- un modèle de diffusion atmosphérique de radionucléides suite à un rejet dans une installation nucléaire, dans le cadre des études d'impact environnemental. Cette application s'appuie sur un code de dispersion atmosphérique bidimensionnel appelé CERES-MITRA et développé par le CEA pour évaluer les conséquences d'un rejet de produits radioactifs dans l'environnement. Les entrées incertaines de ce simulateur sont nombreuses et de différents types : variables scalaires (vitesse de dépôt des radionucléides, hauteur des points de rejets, ...) et variables fonctionnelles (chroniques de rejet, chroniques météorologiques). Les concentrations de radionucléides dans l'air et au sol à différents instants du scénario constituent les sorties du code de calcul
- un modèle simplifié de thermohydraulique développé dans le cadre des études d'accidents graves sur les réacteurs à neutrons rapides de 4^{ème} génération à caloporteur sodium. Ce modèle permet de simuler un scénario de fusion localisée du cœur initiée par le bouchage total et instantané d'un assemblage combustible. La dégradation de l'assemblage initiateur de l'accident et la progression de ce dernier aux assemblages périphériques, estimée par le simulateur, dépend de nombreux paramètres incertains relatifs à la phénoménologie, à sa modélisation ainsi qu'au comportement des systèmes du réacteur.

Dans ces deux applications, les simulateurs numériques utilisés présentent en entrée un grand nombre de paramètres incertains (plusieurs dizaines) et le temps d'évaluation nécessaire à chaque simulation permet d'envisager au mieux quelques centaines de simulations. Ce post-doctorat est financé par le projet DISN/MCIS/MRISQ.

Compétences requises

Thèse en probabilités/statistiques, bonne connaissance des logiciels Matlab et R.

Contact

Amandine MARREL, CEA Cadarache, CEA/DEN/DER/SESI/LSMR, 13108 Saint-Paul-lez-Durance.

Email : [amandine.marrel \[at\] cea.fr](mailto:amandine.marrel@cea.fr)

Références bibliographiques

Iooss B. (2011). Revue sur l'analyse de sensibilité globale de modèles numériques. *Journal de la Société Française de Statistique*, 152:1–23.

Lamboni M., Iooss B., Popelin A-L. and Gamboa F. (2013). Derivative-based global sensitivity measures: general links with Sobol' indices and numerical tests. *Mathematics and Computers in Simulation*, 87 : 45-54.

Marrel A. (2008). Mise en œuvre et utilisation du métamodèle processus gaussien pour l'analyse de sensibilité de modèles numériques. *Thèse de l'INSA Toulouse*.

Oakley J., O'Hagan A. (2002). Bayesian inference for the uncertainty distribution. *Biometrika* 89: 769–784,

Sacks J., Welch W.J., Mitchell T.J. and Wynn H.P. (1989). Design and analysis of computer experiments. *Statistical Science*, 4:409-435 .

Saltelli A., Ratto M., Andres T., Campolongo F., Cariboni J., Gatelli D., Saisana M., Tarantola S. (2008). *Global Sensitivity Analysis*. Wiley.

Sobol I.M. (1993). Sensitivity estimates for non linear mathematical models. *Mathematical Modelling and Computational Experiments*, 1:407–414.

Sobol I.M., Kucherenko S. (2009). Derivative based global sensitivity measures and their link with global sensitivity indices. *Mathematics and Computers in Simulation*, 79:3009-3017