



Sujet de stage – année scolaire 2017/2018

Analyse de sensibilité en grande dimension impliquant des variables catégorielles

CONTEXTE :

Au sein d'EDF R&D, le département PRISME a pour mission de proposer des solutions innovantes pour une exploitation plus performante des différents moyens de production du groupe EDF : maîtrise des risques, amélioration de la sûreté, optimisation des performances technico-économique, maîtrise de la durée de vie. Il contribue notamment au projet MOVIE dédié aux développements et à la diffusion de méthodes de traitement des incertitudes dans les outils de calcul scientifique, et notamment des méthodes d'analyse de sensibilité (AS) qui visent à classer des variables incertaines en fonction de leur impact sur une quantité d'intérêt évaluée numériquement. Les méthodes classiques d'AS ne sont pas adaptées à la prise en compte de variables catégorielles ou de quantité d'intérêt fonctionnelle (méthode de Morris¹) ou à la grande dimension (plusieurs dizaines de variables incertaines au total) lorsque le budget computationnel est limité (méthodes d'estimation usuelles des indices de Sobol²). Il faut donc proposer ou étudier de nouvelles méthodes qui répondent à ces défis.

OBJECTIF DU STAGE :

Le stage portera essentiellement sur l'utilisation d'indices de sensibilité appelés HSIC (*Hilbert-Schmidt independence criterion*)³ afin d'analyser des modèles de simulation numérique d'accident grave. Il s'articulera selon trois phases (qui pourront se chevaucher) :

- 1) un travail bibliographique qui fera la synthèse des éléments théoriques disponibles au sujet des HSIC (propriétés, avantages, limites, liens avec d'autres indices),
- 2) l'implémentation en langage Python de l'estimation statistique de ces indices à partir d'un échantillon de calculs, mais également d'un test du caractère significatif de l'impact de chaque variable incertaine⁴ sur la quantité d'intérêt (mise en œuvre sur des exemples jouets simples et comparaison avec d'autres indices de sensibilité)
- 3) et un travail d'application à la simulation d'un accident nucléaire grave qui nécessitera d'exploiter une ferme de calcul (*cluster*) d'EDF pour obtenir les échantillons nécessaires aux analyses.

D'autres méthodes d'AS pourront par ailleurs être identifiées et étudiées⁵. Concernant 2), une réflexion pourra être menée quant à la manière d'introduire les HSIC dans la plate-forme logicielle OpenTURNS, co-développée par EDF⁶. Notons enfin qu'une réflexion inévitable, théorique et pratique, portera sur le choix des noyaux qui interviennent dans la définition des HSIC. Ils devront être adaptés aux spécificités des modèles d'accident grave : entrées catégorielles et sorties fonction du temps.

¹ https://en.wikipedia.org/wiki/Morris_method

² https://en.wikipedia.org/wiki/Variance-based_sensitivity_analysis

³ <http://www.tandfonline.com/doi/pdf/10.1080/00949655.2014.945932>

⁴ <https://hal-cea.archives-ouvertes.fr/hal-01090475/document>

⁵ Par exemple www.theses.fr/2015PA066045.pdf (section 2.6.2).

⁶ www.openturns.org

PROFIL :

Etudiants de M2 probabilités / statistiques ou grandes écoles.

COMPETENCES SOUHAITEES :

Bases solides en statistiques et en mathématiques appliquées.

Aisance en informatique.

ENVIRONNEMENT INFORMATIQUE :

Linux, Latex, Python, logiciel R.

CONTACTS :

mathieu.couplet@edf.fr

bertrand.iooss@edf.fr

DUREE ENVISAGEE :

5 - 6 mois à partir de mars 2018

LIEU :

EDF R&D

Département Performance, Risque Industriel et Surveillance pour la Maintenance et l'Exploitation

Groupe Modélisation sous Incertitudes, Simulation et Physique du Soudage et des Procédés

6, Quai Watier - 78401 Chatou