

SUJET DE STAGE

Champs gaussiens non-stationnaires et ondelettes pour la planification d'expériences en quantification des incertitudes

J. Baccou, F. Perales
IRSN,
CE Cadarache,
13115 Saint Paul-Lez-Durance,
France

D. Ginsbourger, C. Chevalier
IMSV, Dép. de Mathématiques et de Statistiques,
Alpeneggstrasse 22,
3012 Berne,
Suisse

J. Liandrat
Ecole Centrale Marseille,
38 rue F. Joliot Curie,
13451 Marseille Cedex 20,
France

Lieu : IRSN, Centre de cadarache, 13115 St Paul-Lez-Durance

Contact : jean.baccou@irsn.fr

Mots-clés : analyse de risque, champs gaussiens, plan d'expériences, ondelettes, micromécanique

Contexte et motivation :

La **quantification des incertitudes** est un enjeu de société majeur concernant aussi bien le monde économique, la modélisation des processus naturels, que l'industrie et en particulier la production d'énergie. Dans le domaine de la sûreté nucléaire où le recours à l'expérience physique est souvent inenvisageable, les systèmes complexes sont typiquement modélisés par simulation numérique. Or, en pratique, le temps nécessaire pour l'exécution de chaque simulation haute-fidélité limite le nombre de scénarios sur les entrées pour lesquels la réponse pourra être calculée, pouvant mener à des incertitudes mal quantifiées.

Le sujet de stage s'inscrit dans la thématique du **machine learning**, un domaine de recherche très vivant à la croisée entre mathématiques, statistiques, et informatique appliquées. Une idée centrale de l'apprentissage statistique pour les expériences numériques est de construire un modèle mathématique (ou « métamodèle ») du simulateur en se basant sur les résultats de simulations déjà disponibles et de se servir de ce **métamodèle** pour explorer au mieux l'espace des variables d'entrée lors des prochaines simulations. L'IRSN utilise et participe déjà au développement de méthodes de planification adaptative d'expériences numériques assistée par « métamodèles » probabilistes (de type « Krigeage ») empruntant à la théorie des **champs aléatoires gaussiens**. Plusieurs stratégies d'évaluation de simulateurs numériques reposant sur les champs gaussiens ont fait leurs preuves dans des contextes tels que l'optimisation globale et/ou l'exploration de zones d'intérêt. Toutefois, l'hypothèse de stationnarité souvent faite par défaut est peu adaptée pour approcher des fonctions dont les variations sont fortement hétérogènes d'une région à l'autre de l'espace des entrées.

Le but de ce stage est de développer de **nouvelles méthodes pour la planification adaptative en contexte non-stationnaire**, en tirant à la fois partie des approches séquentielles sur base de champs gaussiens, mais aussi des approches multi-échelles de type **ondelettes**, permettant de modéliser finement les non-stationnarités. On s'appuiera pour cela sur un couplage entre segmentation à base d'ondelettes pour construire un noyau de covariance tenant compte du caractère multi-échelles d'un signal et krigeage. Des critères d'échantillonnage ad hoc pour la quantification d'incertitudes pourront ensuite être définis et étudiés. Enfin, les méthodes développées seront d'abord validées sur des cas synthétiques puis appliquées à des cas réels issus des activités en mécanique du laboratoire commun IRSN-CNRS-UM2 MIST.

Profil du candidat:

En 3^{ième} année de grande école d'ingénieur et/ou étudiant en master2 en mathématiques, mathématiques appliquées, statistiques ou disciplines connexes, la/le stagiaire devra avoir une grande motivation pour la modélisation mathématique, le goût du travail en équipe et des problématiques multidisciplinaires. Une certaine aisance dans l'implémentation informatique (expérience de projets d'étude avec des langages tels que R, Matlab, Java ou C/C++) et dans la familiarisation avec de nouveaux environnements sera un plus.

Le travail pourra être poursuivi par une thèse en cotutelle entre l'Ecole Centrale Marseille (Prof. Jacques Liandrat) et l'Université de Berne (Dr. David Ginsbourger) dans le cadre de la participation de l'IRSN au [consortium ReDICE](#) qui réunit partenaires industriels (EDF, Renault, IFP, CEA, IRSN) et académiques (Mines de St Etienne, Centrale Lyon, Université de Berne) autour de méthodes innovantes pour l'étude des expériences numériques.

Bibliographie

- [1] C.E. Rasmussen and C.K.I. Williams, [Gaussian Processes for Machine Learning](#), The MIT Press, 2006.
- [2] E. Chojnacki, J. Baccou et S. Destercke, *Numerical accuracy and efficiency in the treatment of epistemic and aleatory uncertainty*, International J. of General Systems, 39(7), 683-704, 2010.
- [3] V. Picheny, D. Ginsbourger, O. Roustant, R.T. Haftka, *Adaptive designs of experiments for accurate approximation of a target region*, J. Mech. Des., 132(7), 2010.
- [4] I. Daubechies, *Ten lectures on wavelets*, SIAM, 1992
- [5] J. Baccou, J. Liandrat, *Kriging-based interpolatory subdivision schemes*, Appl. and Comp. Harm. Anal., in Press, 2012.
- [6] Y. Xiong, W. Chen, D. Apley et X. Ding, *A non-stationary covariance-based Kriging method for metamodeling in engineering design*, Int. J. for Num. Meth. in Engineering, 71, 733-756, 2007.
- [7] D.J.J. Toal and A.J. Keane, *Non-Stationary Kriging for Design Optimization*, Engineering Optimization, 44(6), 741-765.
- [8] S. Ba and V. Roshan Joseph, *Composite Gaussian Process Models for Emulating Expensive Functions*, Annals of Applied Statistics 6(4), 1838-1860, 2012.
- [9] *tgpp: An R Package for Bayesian Nonstationary, Semiparametric Nonlinear Regression and Design by Treed Gaussian Process Models*, J. of Statistical Software, 19(9), 2007.
- [10] D. Castano et A. Kunothe, *Robust regression of scattered data with adaptive Spline-Wavelets*, IEEE Trans. Imag. Proc., 15 (6), 1621-1632, 2006.
- [11] F. Perales, F. Dubois, Y. Monerie, B. Piar, L. Stainier, *A NonSmooth Contact Dynamics-based multi-domain solver. Code coupling (Xper) and application to fracture*, Eur. J. Comp. Mech. 19, 389-417, 2010.