



DOCTORANT

« Propagation des incertitudes et calibration des simulations numériques pour l'évaluation des coûts liés aux inondations marines !

Contenu et problématique

Les modèles d'estimation des coûts liés aux risques naturels, utilisés notamment par la caisse centrale de réassurance CCR (<http://www.ccr.fr/>) dans le cadre de la modélisation des catastrophes naturelles en France, sont sujets à des incertitudes importantes. Une part de celles-ci est imputable à la modélisation de l'aléa qui nécessite une description fine des phénomènes. Une autre part de l'incertitude provient des données de vulnérabilité et de leur croisement avec l'aléa modélisé.

Dans le domaine de la submersion marine, le traitement des incertitudes sur l'évaluation des coûts associés aux dommages doit considérer le modèle dans son ensemble incluant tous les paramètres et composantes :

- pour l'aléa : forçages marins au large (vagues, niveaux d'eaux), topographie, propagation de l'inondation, etc. ;
- pour la vulnérabilité : géolocalisation des biens, caractéristiques du bâti, etc.

Une difficulté supplémentaire est associée au faible nombre de données de validation (il y a peu d'événements de submersion marine ayant fait l'objet de retours d'expérience ou disposant d'un nombre important de sinistres d'assurance), ce qui impose d'aborder le problème des incertitudes sous l'angle de l'analyse de sensibilité.

Sur un plan opérationnel, un verrou important est lié à l'utilisation de modèles numériques pour reproduire/prédire la submersion. Leur force est de permettre de reproduire une majeure partie de la complexité des phénomènes (e.g. submersion côtière en milieu urbain à haute résolution spatio-temporelle) mais au détriment d'un coût en temps de calcul très important (plus long que la durée réelle de l'événement) : cela rend difficile une utilisation directe pour un traitement des incertitudes (e.g. par méthode Monte-Carlo) ou pour la calibration sous incertitude (i.e. l'estimation des paramètres incertains grâce à des observations de dommages par approche bayésienne par exemple). En effet, ces deux types d'analyse exigent souvent un nombre très important de simulations (plusieurs milliers) difficilement gérables malgré l'utilisation de grille de calculs.

Objectifs

La propagation des incertitudes le long de la chaîne de modélisation des coûts associés aux dommages estimés pour un événement de submersion marine constitue l'objectif principal de cette thèse. Les points suivants seront abordés :

- quantification de l'incertitude globale du modèle d'évaluation des coûts et ainsi que celle des différentes composantes (données d'entrée, modèle d'aléa, modèle de vulnérabilité, calcul des dommages);
- clarification du rôle joué par les différentes sources d'incertitudes dans une démarche d'analyse de sensibilité ;
- développement de méthodes statistiques basées sur la méta-modélisation (voir paragraphe suivant) pour combiner différents résultats de simulations numériques issus de différentes chaînes de modélisation de l'aléa (une chaîne au BRGM développée au sein de l'unité Risques Côtiers et Changement Climatique¹ de la direction Risques et Prévention et une autre développée au sein de l'équipe Modélisation et R&D à la CCR) et ainsi surmonter les problèmes de temps de calcul;
- développement de méthodes statistiques pour la gestion de données de différentes natures (carte, séries temporelles, scénarios de modélisation, etc.) ;
- application à des cas concrets (par exemple événement Xynthia sur le site des Boucholeurs).

Méthodes envisagées

Le verrou majeur est le temps de calcul important associé à chaque appel du simulateur numérique, ce qui limite le nombre d'appels total pouvant être effectué pour répondre à la problématique du traitement des incertitudes. Une stratégie possible se base sur la méta-modélisation (Roustant et al., 2012) : elle consiste à approximer le modèle coûteux en temps de calcul par un modèle analytique moins coûteux (une surface de réponse) construit à partir d'un nombre limité de couples {données d'entrée ; réponse du simulateur coûteux}. Cette technique a été utilisée dans plusieurs domaines dont celui des risques naturels (e.g. Rohmer et Foerster 2011 ; Rohmer et Idier 2012 ; etc.). Ces méthodes ont montré leur force pour : i. l'optimisation via des méthodes d'échantillonnage adaptatif (e.g. Jones et al., 1998) ; ii. le calcul de probabilité de dépassement via des méthodes d'échantillonnage adaptatif (e.g. Bect et al. 2012). Cependant, la qualité de cette approche dépend fortement du nombre de simulations haute résolution que l'on peut réaliser avec les moyens informatiques (grappe de calcul) disponibles. En pratique, il est souvent possible de se permettre un plus grand nombre de simulations en dégradant la physique (hypothèses simplificatrices). La 1ère stratégie introduit des erreurs liées à l'approximation des résultats de modèles, alors que la 2ème stratégie introduit des erreurs du fait de modèles qui simplifient la physique du phénomène. Des approches récentes (e.g. le cadre des co-krigeages multi-fidélité, Le Gratier et al., 2014) ont été développées pour améliorer la construction des méta-modèles en tirant parti de différentes sources d'informations i.e. en utilisant plusieurs jeux de simulations aux hypothèses de modélisation différentes. Aujourd'hui peu de développements existent pour adapter ce cadre à des problématiques de propagation d'incertitudes ou de calibration sous incertitude, qui sont au cœur de la thèse.

¹ <http://www.brgm.fr/activites/risques/risques-cotiers-littoral>

Conditions et contacts

Profil recherché :

- Mathématique et statistiques ;
- Maîtrise des outils informatiques (expérience en programmation, e.g., R, matlab, python, fortran) ;
- Intérêt pour la théorie des incertitudes, la modélisation statistique et leurs applications à des cas concrets ;
- Bonne autonomie ;
- Bonnes capacités à communiquer (oral et écrit) ;
- Capacité à travailler avec des équipes pluridisciplinaires ;
- Bonne pratique de l'anglais.

Début de la thèse : octobre 2017

Durée : 3 ans

Localisation envisagée :

- site principal : École des Mines de Saint-Étienne (2/3 du temps) ;
- séjours (1 an au total) au BRGM (Orléans) et à la CCR (Paris).

Encadrement :

- Olivier Roustant (directeur de thèse)
- Jérémie Rohmer (BRGM)
- David MONCOULON (CCR)

Votre lettre de motivation, CV actualisé, relevés de notes et références professionnelles sont à adresser dès maintenant à Olivier Roustant par courriel : roustant@emse.fr
Les auditions des candidats auront lieu le 29 juin 2017 dans les locaux de la CCR à Paris.

Bibliographie

- Bect, J., Ginsbourger, D., Li, L., Picheny, V., & Vazquez, E. (2012). Sequential design of computer experiments for the estimation of a probability of failure. *Statistics and Computing*, 22(3), 773-793.
- Jones, D. R., Schonlau, M., & Welch, W. J. (1998). Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(4), 455-492.
- Le Gratiet, L., Cannamela, C., & Iooss, B. (2014). A Bayesian approach for global sensitivity analysis of (multifidelity) computer codes. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1), 336-363.
- Rohmer, J., & Foerster, E. (2011). Global sensitivity analysis of large-scale numerical landslide models based on Gaussian-Process meta-modeling. *Computers & Geosciences*, 37(7), 917-927.
- Rohmer, J., & Idier, D. (2012). A meta-modelling strategy to identify the critical offshore conditions for coastal flooding. *Natural Hazards and Earth System Science*, 12(9), 2943-2955.
- Roustant, O., Ginsbourger, D., & Deville, Y. (2012). DiceKriging, DiceOptim: Two R Packages for the Analysis of Computer Experiments by Kriging-Based Metamodeling and Optimization. *Journal of Statistical Software*, 51(1), 1-55