

Proposition de sujet de thèse

Approches stochastiques et variationnelles pour l'analyse de sensibilité et la quantification d'incertitudes – Application à un système de prévision des circulations océaniques

Les systèmes numériques de prévision environnementaux (météorologie, océanographie, chimie atmosphérique...) sont des outils complexes. Ils associent en général un modèle numérique (traduisant les équations de la physique du phénomène) et des informations disponibles (mesures directes ou indirectes, statistiques, images...) via des techniques dites « d'assimilation de données ». De plus, la dynamique mise en jeu est le plus souvent assez fortement non-linéaire, et avec un espace des états de grande dimension (typiquement $O(10^6-10^7)$ dans les domaines d'applications citées plus haut).

Toutefois, des incertitudes sont présentes à tous les niveaux dans ces systèmes : les équations sont un modèle imparfait de la réalité, et font souvent intervenir des paramètres (coefficients, conditions initiales ou aux limites) mal connus ; la solution numérique est une approximation de la solution continue du modèle ; les observations sont entachées d'erreurs ; le système d'assimilation de données utilise des métriques (covariances d'erreur) mal connues.

Il est donc important, pour donner du sens à la prévision fournie par le système, de lui associer une estimation de son incertitude. Pour fournir des informations en ce sens, plusieurs approches sont alors possibles, qui ont déjà été, tout au moins partiellement, abordées dans le cadre de tels systèmes.

L'approche variationnelle consiste à calculer le gradient d'un critère vis à vis d'un paramètre incertain. Etant données les dimensions mises en jeu, ce gradient est souvent hors d'atteinte par des approximations de type « taux d'accroissement », et doit être calculé par une méthode adjointe (le modèle adjoint étant un outil souvent disponible dans les systèmes d'assimilation de données). Toutefois, si l'on s'intéresse à la réponse non pas du modèle seul, mais du système modèle+observations, on doit alors calculer un gradient à partir du système d'optimalité (modèle direct, modèle adjoint et équation d'Euler), c'est à dire un Hessien du modèle initial. Le coût de calcul devient alors gigantesque, et des approximations doivent souvent être faites pour avoir accès à (une partie de) cette information.

L'approche stochastique consiste quant à elle à proposer des lois de probabilité sur les paramètres d'entrée et à réaliser des simulations de type Monte-Carlo pour en déduire une estimation de la loi du critère de sortie. On est cependant très limité par la méconnaissance des lois d'erreur en entrée et par le coût de réalisation de chaque simulation (par exemple Météo-France ne réalise quotidiennement que quelques dizaines de simulations, correspondant à la perturbation d'un état initial à plusieurs millions de degrés de liberté). Les perturbations sont alors choisies en tentant de maximiser l'information recueillie, par exemple en utilisant des outils de la théorie des systèmes dynamiques (vecteurs de perturbation optimaux). C'est aussi ce qui explique pourquoi très peu d'études aient pu à ce jour être menées de façon exhaustive (analyse de sensibilité aux entrées, surface de réponse, quantification d'incertitudes) sur de tels systèmes.

Le travail proposé dans le cadre de cette thèse consiste à mettre en œuvre et comparer, sur un même cas-test, à la fois les outils variationnels et les outils stochastiques, et éventuellement à proposer des approches mixtes plus « informatives ». On fera tout d'abord un état des lieux des travaux précédents réalisés dans les deux communautés, afin de déterminer les approches variationnelles et stochastiques les plus pertinentes. Puis on appliquera celles-ci dans le cadre du système de prévision océanique NEMOVAR, utilisé par de nombreux centres opérationnels et équipes de recherche en Europe. On travaillera sur une configuration

idéalisée (de façon d'une part à avoir une stratégie d'expériences jumelles, dans laquelle la « vérité » provient d'une simulation de référence, et d'autre part à limiter les coûts de calcul). Enfin, une synthèse des résultats sera effectuée, dans l'optique de dégager des recommandations, voire des méthodes mixtes variationnelles-stochastiques.

Directeurs de thèse : Clémentine PRIEUR et Eric BLAYO (Laboratoire Jean Kuntzmann, Université Joseph Fourier, Grenoble. Contact : Clementine.Prieur@insa-toulouse.fr, Eric.Blayo@imag.fr)

Profil souhaité pour le ou la candidat(e) : Master en mathématiques appliquées, connaissance des EDP et de la théorie des probabilités, expérience en simulation numérique et calcul scientifique

Les outils (modèle direct, modèle adjoint, configuration idéalisée) sont déjà ou seront disponibles au moment du démarrage de cette thèse. Il y a de plus localement dans l'équipe les compétences pour encadrer la mise en œuvre de ces outils.

Quelques références :

Le Dimet F.-X. H. E. Ngodock, and B. Luong, 1997 : Sensitivity analysis in variational data assimilation. *J. Meteorol. Soc. Japan*, **75-1B**, 245-255.

Le Dimet F.-X., I.M. Navon and D. M. Daescu, 2002 : Second order information in data assimilation. *Monthly Weather Review*, **130**, 629-648.

McKeague I.W., G. Nicholls and K. Speer, 2005: Statistical inversion of South Atlantic circulation in an abyssal neutral density layer. *J. Mar. Res.*, **63**, 683-704.

Oakley J.E. and A. O'Hagan, 2004 : Probabilistic sensitivity analysis of complex models : a Bayesian approach. *J. Roy. Stat. Soc. B*, **66**, 751-769.

Ngepieba P., M. Hussaini and L. Debreu, 2006 : Optimal control and stochastic parameter estimation. *Monte Carlo Methods and Applications*, **12**, 461-476.

Wang Z., I.M Navon, F.-X Le Dimet, and X. Xou, 1992 : The second order adjoint analysis: theory and applications. *Met. Atmos. Phys.*, **50**, 3-20.