

[English below]

Le potentiel des méthodes de Monte Carlo multi-niveaux pour l'assimilation de données : exploiter les différentes résolutions de nos modèles

Orateur:

Mayeul Destouches (Cerfacs, Toulouse, France)

En collaboration avec:

Paul Mycek, Selime Gürol, Anthony Weaver, Serge Gratton, Ehouarn Simon, Jérémy Briant.

Résumé:

En assimilation de données variationnelle (EnVar), la matrice \mathbf{B} de covariance des erreurs d'ébauche est estimée à chaque cycle d'assimilation à partir d'un ensemble de prévisions. Une manière possible de générer cet ensemble est d'utiliser un Ensemble d'Assimilations (*Ensemble of Data Assimilation*, EDA), qui échantillonne toutes les entrées incertaines du système : erreur des observations, des conditions aux bords, du modèle de prévision. Plus cet ensemble comporte de membres, plus l'estimation de la matrice \mathbf{B} sera précise. En pratique cependant, la génération de chaque membre est numériquement très coûteuse. Une solution courante consiste à utiliser dans l'EDA des membres à résolution dégradée par rapport au modèle déterministe, ce qui réduit le coût de génération de chaque membre, et permet donc une plus grande taille d'ensemble.

Les méthodes Monte-Carlo multi-niveaux (MLMC) proposent d'aller au-delà de ce compromis entre taille de l'ensemble et résolution des membres. Les estimateurs multi-niveaux s'expriment comme une combinaison astucieuses des estimateurs obtenus à partir de plusieurs EDA de résolutions différentes. À coût de calcul constant, ces méthodes permettent de réduire la variance de l'estimateur de la matrice de covariance, c'est-à-dire de réduire le bruit d'échantillonnage qui affecte la matrice \mathbf{B} .

Nous présenterons dans un premier temps les bases théoriques du MLMC, en particulier dans le cas de l'estimation d'une matrice de variance-covariance. Nous présenterons ensuite une application à un modèle quasi-géostrophique. Le potentiel des méthodes MLMC sera d'abord illustré sur un cycle d'assimilation, dans des conditions où le gain attendu est optimal. Nous aborderons ensuite le cas d'expériences cyclées assimilation-prévision, et discuterons les limitations de l'approche.

Multilevel Data Assimilation: Leveraging Models of Varying Resolutions in Ensemble Variational Data Assimilation

Speaker:

Mayeul Destouches (Cerfacs, Toulouse, France)

In collaboration with:

Paul Mycek, Selime Gürol, Anthony Weaver, Serge Gratton, Ehouarn Simon, Jérémy Briant.

Abstract:

In ensemble variational (EnVar) data assimilation systems, the background error covariance matrix \mathbf{B} is sampled from an ensemble of forecasts evolving with time. One possible way of generating this ensemble is by running an Ensemble of Data Assimilations (EDA) that samples all possible error sources (observation error, boundary condition errors, model errors). Large ensemble sizes are desirable to minimize sampling errors, but generating a single ensemble member is usually expensive due to the cost of integrating the physical model. In practice, ensembles with coarser spatial resolutions are sometimes used, allowing for cheaper generation of individual members, and thus larger ensemble sizes.

Multilevel Monte Carlo (MLMC) methods propose to go beyond this usual trade-off between grid resolution and ensemble size, by expressing a fine-grid estimator as an astute combination of estimators computed on a hierarchy of spatial grids. Without increasing the computational cost, MLMC effectively reduces the variance of the covariance estimator, *i.e.* reduces the sampling noise on covariances.

We first present the theoretical basis of MLMC and how it can apply to the estimation of covariance matrices. An illustration with a quasi-geostrophic model is then presented. The potential of using an MLMC \mathbf{B} matrix is first assessed on one data assimilation cycle, in an optimal setting where MLMC is expected to perform best. We then evaluate the theoretical gain we can expect from MLMC in cycled experiments, and discuss the limitations of the approach. The role of covariance localization is also briefly discussed.