

## Modélisation de la covariance pour l'assimilation variationnelle de données océaniques

Ce stage de 6 mois se déroulera dans l'équipe **Algorithmes Parallèles (ALGO)** du **CERFACS**, située à la Météopole de Toulouse.

Les superviseurs principaux seront Anthony Weaver (CERFACS-CECI/CNRS; [weaver@cerfacs.fr](mailto:weaver@cerfacs.fr)) et Olivier Goux (CERFACS-CECI/CNRS ; [goux@cerfacs.fr](mailto:goux@cerfacs.fr)).

### 1. Contexte général

L'assimilation de données (AD) est une procédure qui consiste à combiner les observations (provenant des satellites et des réseaux *in situ*, par exemple) de l'état d'un système avec une estimation préalable (appelée ébauche) afin de produire la meilleure estimation possible. Dans l'assimilation variationnelle des données, l'estimation optimale est obtenue itérativement en minimisant une fonction de coût non linéaire pondérée des moindres carrés qui est la somme de deux termes : l'un mesurant l'adéquation d'une estimation à l'ébauche ; l'autre mesurant l'adéquation de l'estimation aux observations, sous réserve de contraintes physiques sur l'évolution dans le temps de l'estimation. Les matrices de pondération de ces termes sont définies en termes de matrices de covariance des erreurs d'ébauche et d'observation. En pratique, cette procédure extrait et fusionne les informations les plus fiables de chaque type d'observations disponible, sur la base leurs incertitudes respectives. L'AD variationnelle est couramment utilisée pour l'estimation de l'état de systèmes météorologiques, océanographiques ou climatologiques.

Le travail effectué dans le cadre de ce stage contribuera au développement d'un modèle de covariance de l'erreur d'ébauche au sein du code d'assimilation de données océanique NEMOVAR. NEMOVAR est utilisé de manière opérationnelle pour les applications océaniques, météorologiques et climatiques au Centre Européen pour les Prévisions Météorologiques à Moyen Terme (CEPMMT) et au UK Met Office. Il est développé conjointement par le CERFACS, le CEPMMT, le UK Met Office et l'INRIA. Des interactions collaboratives avec des scientifiques de ces instituts auront lieu pendant le stage. Ce travail contribuera également à un projet de l'AD océanique financé par le Copernicus Climate Change Service.

### 2. Sujet du stage

Il y aura deux axes de cette étude.

1. **Corrélations anisotropes.** Dans NEMOVAR, les corrélations d'erreur d'ébauche sont modélisées à l'aide d'un opérateur de diffusion, formulé comme la solution d'une équation de diffusion discrétisée à l'aide d'un schéma d'Euler implicite (Mirouze et Weaver 2010 ; Weaver et Mirouze 2013 ; Weaver et al. 2016). Le tenseur de diffusion contrôle les propriétés de lissage spatial du modèle de corrélation. Dans la mise en œuvre de NEMOVAR par le CEPMMT, le tenseur de diffusion pour chaque variable de contrôle est estimé à partir d'un ensemble d'assimilations de données (Weaver et al. 2021 ; Chrust et al. 2021). Le processus d'estimation fournit des estimations des éléments diagonaux et non diagonaux du tenseur. Actuellement, les éléments hors diagonale sont négligés.

Une manière plus efficace d'exploiter l'information d'ensemble consiste à généraliser l'opérateur de diffusion pour tenir compte des termes de dérivation croisée qui résultent de l'utilisation d'un tenseur non diagonal. Le défi consiste à introduire une discrétisation numérique qui génère une réponse anisotrope précise, tout en préservant les attributs de symétrie et de positivité de l'opérateur de diffusion discrétisé. La symétrie et de positivité sont des propriétés cruciales qui doivent être satisfaites pour que l'opérateur de diffusion produise une matrice de corrélation valide (symétrique et finie positive). Dans le cadre de ce stage, l'étudiant étudiera le schéma symétrique de différences finies de van Es et al. (2014) pour tenir compte des corrélations anisotropes horizontales dans NEMOVAR.

2. **Corrélations adaptées à la bathymétrie.** Le modèle océanique NEMO (Madec et al. 2023), sur lequel est basé NEMOVAR, utilise un système de coordonnées verticales "z", avec un "pas partiel" dans le niveau du fond de l'océan pour une meilleure représentation de la bathymétrie. Avec les cellules de fond partielles, les quantités dans les cellules horizontalement adjacentes sont définies à des profondeurs différentes. Afin de calculer les dérivées horizontales requises dans les équations du modèle (par exemple, dans le terme de diffusion), la quantité dans le niveau le plus profond est d'abord interpolée au niveau le moins profond pour une meilleure précision. Un système de coordonnées z avec des pas partiels est également utilisé dans NEMOVAR. Cependant, l'opérateur de diffusion utilisé pour modéliser les corrélations d'erreur d'ébauche ne comprend pas de traitement spécial des dérivées horizontales dans les cellules du fond, ce qui entraîne des corrélations parasites dans les régions à topographie abrupte (par exemple, près des plateaux continentaux). Dans ce stage, l'étudiant adaptera le modèle de corrélation pour inclure la procédure d'interpolation utilisée dans NEMO. Comme pour le schéma de diffusion anisotrope, le défi consistera à mettre en œuvre un schéma numérique (d'interpolation) qui préserve l'attribut SPD de la matrice de corrélation.

Le stagiaire devra mettre en œuvre les marches partielles et l'anisotropie horizontale 2D dans le modèle de diffusion de NEMOVAR, et valider la mise en œuvre numérique à l'aide de tests unitaires standard. Des expériences d'observation unique et de corrélation ponctuelle seront utilisées pour illustrer l'impact de ces nouvelles fonctionnalités. Si le temps le permet, des expériences d'assimilation de données pluriannuelles seront menées en collaboration avec nos partenaires afin d'évaluer l'impact dans un cadre réaliste de modèle océanique global. Les principaux développements et tests seront effectués en utilisant l'infrastructure informatique du CERFACS. Le développement du logiciel impliquera de la programmation en Fortran et en python.

## References

1. Chrust, M., Balmaseda, M. A., Browne, P., Martin, M., Storto, A., Vidard, A., Weaver, A. T., Zuo, H. Ensemble of Data Assimilations in the ocean for better exploitation of surface observations. ECMWF Newsletter 168, Summer 2021.
2. Madec, G. and the NEMO System Team, 202. NEMO Ocean Engine Reference Manual. Zenodo, <https://doi.org/10.5281/zenodo.8167700>.
3. Mirouze I, Weaver AT. 2010. Representation of correlation functions in variational assimilation using an implicit diffusion operator. *Q. J. R. Meteorol. Soc.* **136**: 1421-1443.
4. van Ens, B., Koren, B. and H. J. de Blank, 2014. Finite-difference schemes for anisotropic diffusion. *J. Comput. Phys.*, **272**, 526—549.
5. Weaver, A. T., Chrust, M., Ménétrier, B., Piacentini, A., 2021: An evaluation of methods for normalizing diffusion-based covariance operators in variational data assimilation. *Q. J. R. Meteorol. Soc.*, **147**, 289—320, DOI:10.1002/qj.3918.
6. Weaver AT, Tshimanga, J, Piacentini P. 2016. Correlation operators based on an implicitly formulated diffusion equation solved with the Chebyshev iteration. *Q. J. Roy. Meteorol. Soc.* **142**: 455-471.
7. Weaver AT, Mirouze I. 2013. On the diffusion equation and its application to isotropic and anisotropic correlation modelling in variational assimilation. *Q. J. R. Meteorol. Soc.* **139**: 242-260.