

PROJET DE FIN D'ETUDES
INGENIEURS DE L'ECOLE NATIONALE DE LA METEOROLOGIE
FICHE DE PROPOSITION DE SUJET

Titre du sujet proposé Apprentissage automatique d'un préconditionneur spectral pour le 4DVar ARPEGE

Organisme ou service proposant le sujet : DESR/CNRM/GMAP/ASSIM

Responsable principal du stage :

Responsable principal (le responsable principal est l'interlocuteur direct de l'Ecole. C'est à lui, en particulier, que seront adressés les courriers ultérieurs) :

NOM : DESTOUCHES

Prénom : Mayeul

Téléphone : +33 5 61 07 90 49

Mél : mayeul.destouches@meteo.fr

Autres responsables :

Le stage présente-t-il un caractère de confidentialité ? : Non

Le stage peut-il être effectué à distance ? : Non

1) Description du sujet – livrables attendus

Une part importante des erreurs de prévisions météorologiques à l'échelle globale peut être liée aux erreurs dans l'état initial fourni au modèle de prévision. Ces erreurs initiales sont contrôlées grâce au processus d'assimilation de données, qui fournit la meilleure estimation de l'état actuel de l'atmosphère. Ce processus combine de manière optimale la dernière prévision disponible avec les mesures, même indirectes, de l'état de l'atmosphère, en prenant en compte les incertitudes associées à chacune de ces sources d'information.

A Météo-France, le processus d'assimilation pour le modèle global ARPEGE repose sur un schéma d'assimilation variationnel, le 4DVar incrémental, qui consiste à résoudre toutes les six heures une série de problèmes d'optimisation quadratique en très grande dimension. Le calcul, contraint par le temps, est accéléré en préconditionnant chaque problème à l'aide de l'information spectrale récupérée lors de la résolution du problème précédent (Tshimanga et al. 2008; Fisher et al. 2009).

Récemment, plusieurs approches ont été proposées pour préconditionner également le premier problème quadratique, pour lequel il n'existe pas de « problème précédent ». Un

préconditionneur spectral peut en effet être construit soit par randomisation de la matrice Hessienne (Daužickaitė et al. 2021), soit par apprentissage automatique (Trappler and Vidard 2025). Cette dernière étude repose aussi, bien qu'implicitement, sur la randomisation de la matrice Hessienne. Nous nous proposons de tester cette approche avec une version simplifiée du modèle ARPEGE, en intégrant une dépendance au réseau d'observation courant, et en adaptant la cible de l'entraînement pour se défaire de la dépendance implicite à la randomisation.

La tâche principale du stagiaire consistera à construire et entraîner un réseau de neurone pour apprendre un preconditionneur spectral optimal qui puisse être utilisé dès le début de résolution du premier problème quadratique. Selon l'avancement, l'efficacité du preconditionneur ainsi obtenu pourra être comparé à celle d'un preconditionneur obtenu par randomisation. Un objectif secondaire du stage est de parvenir à une meilleure compréhension de ce qui distingue fondamentalement l'information spectrale obtenue lors de la résolution d'un problème d'optimisation quadratique de l'information spectrale obtenue par randomisation.

Un plan de travail pourrait être :

- Etude bibliographique sur le 4DVar incrémental et les preconditionneurs spectraux de second niveau.
- Conceptualisation, construction et entraînement du réseau de neurones.
- Eventuellement en parallèle, comparaisons empiriques des preconditionneurs construits par randomisation ou par résolution d'un problème précédent.
- Evaluation de l'impact du preconditionneur obtenu par *machine learning* sur la convergence du problème d'assimilation de données.
- Mise en forme des résultats et synthèse sous forme d'un rapport écrit et de présentation orale.

Prérequis :

- Une formation en intelligence artificielle en requise.
- Un solide bagage mathématique serait très utile, en particulier en algèbre linéaire.
- Une formation aux sciences de l'atmosphère ou à l'assimilation de données serait profitable, sans être indispensable.
- Le stage demande une appétence pour le travail sur Linux et le codage Python, voire en C++ si affinités.

Références :

1. Tshimanga, J., Gratton, S., Weaver, A. T. & Sarternaer, A. Limited-memory preconditioners, with application to incremental four-dimensional variational data assimilation. *Q. J. R. Meteorol. Soc.* **134**, 751–769 (2008).
2. Fisher, M., Nocedal, J., Trémolet, Y. & Wright, S. J. Data assimilation in weather forecasting: a case study in PDE-constrained optimization. *Optim. Eng.* **10**, 409–426 (2009).
3. Daužickaitė, I., Lawless, A. S., Scott, J. A. & Van Leeuwen, P. J. Randomised preconditioning for the forcing formulation of weak-constraint 4D-Var. *Q. J. R. Meteorol. Soc.* **147**, 3719–3734 (2021).
4. Trappler, V. & Vidard, A. State-dependent preconditioning for the inner-loop in Variational Data Assimilation using Machine Learning. Preprint at <https://doi.org/10.48550/arXiv.2501.04369> (2025).

2) lieu du stage, durée ou période

Dans les locaux du CNRM, sur la météopôle.

5 à 6 mois, à débuter entre février et avril 2026.