

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

Titre du sujet proposé : Utilisation de techniques d'IA générative pour une paramétrisation (stochastique) de la convection profonde dans ARP-GEM

Organisme ou service proposant le sujet : DESR/CNRM/GMGEC/GLOB-ATM

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 : BALOGH

Prénom : Blanka

téléphone : +33 5 61 07 96 22

Mél : blanka.balogh@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

Même si des études récentes montrent qu'il est possible d'utiliser des résolutions plus fines pour des simulations climatiques, la majorité des modèles de climat utilisent une résolution horizontale de l'ordre de 150 km à 50 km. À ces résolutions, seuls les processus de grande échelle sont décrits par les équations primitives discrétisées. Pour représenter les processus de fine échelle, non explicitement résolus par le modèle, comme la convection profonde, la turbulence ou la microphysique, la démarche classique consiste à utiliser des paramétrisations physiques. Les paramétrisations physiques sont des ensembles d'équations, empiriques ou théoriques, développées à partir d'une combinaison de données issues d'observations, de modèles conceptuels, de résultats de modèles à haute résolution, et/ou d'approches théoriques.

Une nouvelle méthode permettant de développer de nouvelles paramétrisations consiste à utiliser des techniques de *machine learning*, en particulier des réseaux de neurones (NN). Les NN pourraient réduire certains biais connus des paramétrisations physiques. Mais la cohérence physique des paramétrisations NN peut être insuffisante. Cela se traduit alors par l'apparition de biais (et/ou d'explosion numérique) dans les simulations climatiques.

Ce stage propose d'explorer l'utilisation d'approches d'IA génératives pour l'émulation d'une paramétrisation (stochastique) de la convection profonde, telles que les GANs et/ou des modèles de diffusion. Pour ce faire, nous utiliserons la paramétrisation de convection profonde de notre modèle comme cadre idéal de travail. Il s'agira de mettre en place et tester la valeur ajoutée des approches d'IA génératives, y compris en remplacement de la paramétrisation de convection profonde existante dans notre modèle (tests 'en ligne').

Le cadre de travail permettant la construction de l'échantillon d'apprentissage est en place, et la réalisation de tests en ligne de paramétrisations NN est d'ores et déjà possible et a été testé. Un planning prévisionnel du projet pourrait être le suivant :

- (1) lecture de quelques articles clés pour mieux comprendre le problème ;
- (2) construction de l'échantillon d'apprentissage à partir des données brutes ;
- (3) réalisation des apprentissages ;
- (4) comparaison des résultats obtenus par le modèle IA avec le modèle physique ;
- (5) test de la paramétrisation obtenue en ligne.

Les tests en ligne se feront dans le modèle ARP-GEM, une version d'ARPEGE/IFS, optimisée et contenant une physique à l'état de l'art (Geoffroy & Saint-Martin, 2025) pour laquelle un interfaçage souple entre le modèle de climat et la paramétrisation NN a été récemment développée (Balogh et al., 2024, en préparation).

Suivant l'avancement du travail, d'autres pistes pour une utilisation des algorithmes d'IA générative pour la modélisation climatique pourront être explorées.

Références sur le sujet

- Sur les paramétrisations NN : Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G., & Yacalis, G. (2018). Could machine learning break the convection parameterization deadlock? *Geophysical Research Letters*, 45, 5742–5751. <https://doi.org/10.1029/2018GL078202>
- Sur l'utilisation de l'IA générative pour l'apprentissage d'une paramétrisation dans un cadre simplifié : Gagne, D. J., Christensen, H., Subramanian, A., & Monahan, A. H. (2020). Machine learning for stochastic parameterization: Generative adversarial networks in the Lorenz '96 model. *Journal of Advances in Modeling Earth Systems*, 12, e2019MS001896. <https://doi.org/10.1029/2019MS001896>
- Publications de l'équipe sur la paramétrisation IA de la convection profonde : Balogh, B., D. Saint-Martin, and O. Geoffroy, 2025: Online Test of a Neural Network Deep Convection Parameterization in ARP-GEM1. *Artif. Intell. Earth Syst.*, 4, e240100, <https://doi.org/10.1175/AIES-D-24-0100.1>.

Germain, H., Balogh, B., Geoffroy, O., and D., Saint-Martin, 2025: Improvement of a neural network convection scheme by including triggering and tests in present and future climate. *In preparation for AIES*.

- Le modèle ARP-GEM : Geoffroy, O., and D. Saint-Martin, 2025: The ARP-GEM1 Global Atmosphere Model: Description, Speedup Analysis, and Multiscale Evaluation up to 6 km. *J. Climate*, **38**, 4739–4762, <https://doi.org/10.1175/JCLI-D-24-0547.1>.

Prérequis

Un bon niveau en Python et en machine learning est souhaité. Quelques bases en Fortran peuvent être utiles. Des bases en sciences du climat sont un plus.

2) lieu du stage, durée ou période

Lieu du stage : CNRM, Toulouse.

Durée : 5 à 6 mois.

Date de début : libre, entre février et avril 2026.