

PROJET DE FIN D'ETUDES

INGENIEURS DE L'ECOLE NATIONALE DE LA METEOROLOGIE

FICHE DE PROPOSITION DE SUJET

Titre du sujet proposé : Utilisation de réseaux de neurones informés par la physique (PINNs) pour le développement de paramétrisations 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 61079622

Mél : blanka.balogh@meteo.fr

Autres responsables :

- David Saint-Martin (david.saint-martin@meteo.fr)

- Olivier Geoffroy (olivier.geoffroy@meteo.fr)

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.

Il existe une famille de techniques permettant d'implémenter des contraintes physiques aux NN pendant l'apprentissage, appelée « NNs informés par la physique » (ou « *physics informed neural networks* », PINNs). Dans ce stage, nous proposons d'explorer ces approches dans le cadre du développement d'une paramétrisation NN pour la convection profonde. 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 de contraintes physiques, à travers l'utilisation d'une des deux méthodes suivantes : (i) via un terme de pénalité ajouté à la *loss* pendant l'apprentissage ou (ii) en ajoutant une couche spécifique à la fin du NN dont les paramètres sont imposés manuellement.

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 en implémentant la ou les contraintes physiques ;
- (4) comparaison des résultats obtenus avec chacune des méthodes (et avec la version 'sans contraintes physiques');
- (5) test de la paramétrisation obtenue en ligne

Les tests en ligne se feront dans le modèle ARP-GEM1, une version d'ARPEGE/IFS, optimisée et contenant une physique à l'état de l'art (Saint-Martin et Geoffroy, 2024 ; Geoffroy et Saint-Martin, 2024, en préparation) 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 optimale des algorithmes de *machine learning* pour la modélisation climatique pourront être explorées.

Références sur le sujet

- Sur l'utilisation des PINNs pour les paramétrisations : Beucler, T., Pritchard, M., Rasp, S., Ott, J., Baldi, P., & Gentine, P. (2021). Enforcing Analytic Constraints in Neural Networks Emulating Physical Systems. *Physical Review Letters*, 126(9). <https://doi.org/10.1103/physrevlett.126.098302>
- 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 les PINNs : Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2017, November 28). Physics Informed Deep Learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations. *arXiv.org*. <https://arxiv.org/abs/1711.10561>

Balogh, B., Saint-Martin, D., & Ribes, A. (2021). A toy model to investigate stability of AI-based dynamical systems. *Geophysical Research Letters*, 48, e2020GL092133.

<https://doi.org/10.1029/2020GL092133>

Balogh, B., Saint-Martin, D., & Ribes, A. (2022). How to calibrate a dynamical system with neural network based physics? *Geophysical Research Letters*, 49, e2022GL097872.

<https://doi.org/10.1029/2022GL097872>

Balogh, B., Saint-Martin, D., & Geoffroy, O. (2024). Developing a Fortran-Python Interface in ARP-GEM: Application to a Deep Convection Neural Network. *Submitted to AIES*.

Saint-Martin, D., & Geoffroy, O. (2024). The ARP-GEM Global Atmospheric Model Version 1, part I: Model Description and Speed-Up Analysis. <https://arxiv.org/abs/2409.19083v1>

Geoffroy, O., & Saint-Martin, D. (2024). The ARP-GEM global atmospheric model version 1, part II : multiscale evaluation from 50 to 6 km. <https://arxiv.org/abs/2409.19089>

Prérequis

Un bon niveau en python est souhaité, avec des bases en machine learning. 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

Durée : 5 à 6 mois entre février 2025 et août 2025