Ecole Nationale de la Météorologie
Direction des Etudes
42, avenue Gaspard Coriolis
BP 45712
31057 TOULOUSE Cedex 1

PROJET DE FIN D'ETUDES
INGENIEURS DE L'ECOLE NATIONALE DE LA METEOROLOGIE
FICHE DE PROPOSITION DE SUJET
/
FINAL YEAR INTERNSHIP PROJECT
ENGINEERING STUDENTS – NATIONAL SCHOOL OF METEOROLOGY
PROPOSAL FORM

1- Titre du sujet proposé / Proposed project title:

Robustesse au changement climatique d'une paramétrisation de la convection profonde par NN

2 - Organisme ou service proposant le sujet / Organization or department proposing the topic:

DESR/CNRM/GMGEC/GLOB-ATM

3 - Responsable principal du stage / Main internship supervisor:

Responsable principal (le responsable principal est l'interlocuteur direct de l'Ecole. C'est à lui, en particulier, que seront adressés les courriers ultérieurs) :
The main supervisor is the primary contact for the School. All official correspondence will be addressed to this person:

NOM / **Last name** : BALOGH Prénom / **First name**: Blanka

téléphone / **phone** : +33 5 61079622 Mél / **Email**: blanka.balogh@meteo.fr

4 - Le stage présente-t-il un caractère de confidentialité ? / Is the internship subject to confidentiality?

Non

5 - Le stage peut-il être effectué à distance ? / Can the internship be carried out remotely?

Non

6 - Description du sujet – livrables attendus / Project description – Expected deliverables

Bien que des simulations climatiques à résolution kilométrique soient désormais réalisables, la plupart des modèles de climat utilisent encore une résolution horizontale de l'ordre de la centaine de kilomètres. À ces résolutions, seuls les processus de grande échelle sont résolus explicitement par la discrétisation des équations primitives. Les processus de plus petite échelle, comme la convection profonde, la turbulence ou encore la microphysique, sont représentés par des paramétrisations physiques : des ensembles d'équations, empiriques ou théoriques, établies à partir d'observations, de modèles conceptuels, de simulations à haute-résolution et de considérations théoriques.

Une approche plus récente consiste à développer ces paramétrisations par apprentissage automatique (machine learning, ML), en particulier à l'aide de réseaux de neurones (NN). Les NN pourraient réduire certains biasi connus des paramétrisations physiques. Leur précision et cohérence physique ne sont cependant pas garanties, ce qui peut se traduire par des biais, voire des instabilités numériques, une fois le réseau couplé au modèle de climat. Ce stage s'inscrit dans la continuité des travaux de l'équipe GLOB-ATM sur le développement d'une paramétrisation de la convection profonde par NN dans le modèle de climat global ARP-GEM (Balogh et al., 2025 ; Germain et al., 2026, pour le déclenchement de la convection par NN ; Moreau et al., en préparation, pour l'approche générative par Flow Matching).

Les paramétrisations par NN sont entraînées sur des simulations en climat présent ou sur des observations, mais doivent rester valides dans des climats plus chauds, en dehors de leur domaine d'apprentissage. Or, les NN sont réputés peu robustes à l'extrapolation (eg., Beucler et al., 2024). L'objectif du stage est de comparer, dans un cadre contrôlé, plusieurs approches existantes visant à améliorer la capacité de généralisation d'une paramétrisation NN de la convection profonde entraînée en climat présent et évalué en climat plus chaud.

Le cadre d'étude est l'émulation du schéma de convection profonde physique d'ARP-GEM à 150 km de résolution (Tiedtke-Bechtold) : les NN sont entraînés en climat présent et testés dans une simulation +8K. Le schéma étant le même dans les deux climats, l'erreur mesurée en climat chaud reflète essentiellement la capacité d'extrapolation du réseau. Deux facteurs seront croisés :
- l'architecture : MLP ou 1D-CNN selon la verticale ;
- la représentation des entrées : normalisation classique par niveau vertical, variables physiques dont la distribution dépend peu du climat (Beucler et al., 2024) ou normalisation par colonne inspirée de l'approche RevIN (Kim et al., 2022) et de Perezhogni et al., 2025.

Les résultats seront comparés à des réseaux de référence entraînés directement en climat +8K, sur plusieurs initialisations. L'évaluation portera d'abord sur des diagnostics offline (calculés sur un échantillon test indépendant), puis sur des simulations ARP-GEM utilisant un NN développé comme paramétrisation de la convection profonde (online).

Références sur le sujet

- Germain, H., B. Balogh, O. Geoffroy, and D. Saint-Martin (2026). Improvement of a neural network convection scheme by including triggering and evaluation in present and future climates. *Artif. Intell. Earth Syst.,* e250108*,* https://doi.org/10.1175/AIES-D-25-0108.1

- Tom Beucler *et al. (*2024). Climate-invariant machine learning. *Science Advances*, **10**, eadj7250. https://doi.org/10.1126/sciadv.adj7250 .
- T. Kim, J. Kim, Y. Tae, C. Park, J. Choi and J. Choo; 2022: Reversible Instance Normalization for Accurate Time Series Forecasting against Distribution Shift. *ICLR, OpenReview.net,* url: https://openreview.net/forum?id=cGDAkQo1C0p.
Perezhogin, P., Zhang, C., Adcroft, A., Fernandez-Granda, C., & Zanna, L. (2024). A stable implementation of a data-driven scale-aware mesoscale parameterization. *Journal of Advances in*

Modeling Earth Systems, **16**, e2023MS004104. https://doi.org/10.1029/2023MS004104

Prérequis
Un bon niveau en Python, en machine learning et un intérêt pour les sciences du climat sont souhaités. Des bases en sciences du climat sont un plus.

Livrables
Rapport de stage (format selon exigences de l'école/université), code produit pendant le stage.

7 - Lieu du stage, durée ou période / Internship location, duration or period

Lieu du stage : CNRM, Toulouse.
Durée : 5 à 6 mois.
Date de début : entre février et mars 2027.