

# Offre de Stage IPSL 2024

(soutenu par le programme EUR IPSL-*Climate Graduate School*)

Titre du sujet de stage : **Machine Learning based online bias correction for the LMDZ atmospheric model**

Description du sujet :

Les modèles de climat sont des outils cruciaux pour comprendre le climat actuel et son évolution future. Cependant, les modèles actuels présentent encore des biais significatifs par rapport aux observations en raison de leur faible résolution et de la nécessité de représenter les processus de petite échelle par des paramétrisations. LMDZ (voir <https://lmdz.lmd.jussieu.fr/>), qui constitue la composante atmosphérique du modèle de climat de l'IPSL (voir <https://cmc.ipsl.fr/>), est un modèle de circulation générale de l'atmosphère qui résout les équations de la dynamique des fluides appliquées à l'atmosphère et prend en compte les processus atmosphériques pertinents. Sa résolution est de 144 x 142 x 79 points en longitude, latitude, et sur la verticale. Ce modèle peut être utilisé en mode libre (le modèle génère sa propre trajectoire météorologique) ou en mode guidé par les réanalyses météorologiques (le modèle suit une trajectoire atmosphérique observée). Le terme de relaxation vers les réanalyses du vent horizontal, estimé à une fréquence de 6 heures, peut être vu comme un terme d'erreur qu'on peut souhaiter corriger d'une manière statistique. Des simulations montrent qu'injecter la moyenne du terme d'erreur dans une simulation en mode libre permet d'en réduire les biais par rapport aux observations.

Ce stage de Master 2 offre une opportunité unique de contribuer à améliorer la représentation de ce terme de correction d'erreur pour le rendre plus dynamique de manière à ce qui soit représentatif de l'état de l'atmosphère. On espère ainsi réduire encore plus les biais du modèle atmosphérique et améliorer ses performances par rapport aux observations (en particulier les distributions de la précipitation). Ce travail sera mené en utilisant les dernières avancées dans le domaine de l'apprentissage automatique. Concrètement, le but du stage est de développer un modèle Machine Learning qui permettra d'estimer la fonction qui relie les erreurs du champ de vent avec l'état instantané du modèle, puis d'intégrer ce modèle statistique basé apprentissage automatique dans une simulation LMDZ réalisée en mode libre. Les métriques d'application sont basées sur l'écart entre LMDZ en mode libre et ERA5 une réanalyse météorologique, en fin de simulation et permettront de mesurer la performance de la méthode.

Le modèle ML entraîné devra inférer les corrections à apporter aux variables physiques du modèle à une fréquence de 6 heures. Dans une première étape, la modélisation ML sera 1D, cad sur chacun des points de la grille pris individuellement (comprenant les 79 niveaux de l'atmosphère). La modélisation 1D évitera les communications interprocessus qui surviennent dans le cas d'une modélisation 3D dont le volume serait partagé entre plusieurs domaines. Les modèles ML envisagés seront de type perceptron multicouche ou basés sur des arbres aléatoires parallélisés et distribués sur CPU et/ou GPU.

L'IPSL mettra à disposition du stagiaire ses moyens de calculs (clusters CPU et GPU). Il sera également envisagé un accès au super ordinateur [Jean Zay](#) de l'IDRIS. En plus de l'entraînement du modèle ML, d'autres objectifs sont possibles, en fonction de l'avancement du stagiaire :

**Objectif supp #1** : entraînement en 3D avec des architectures type U-Nets ou basés sur des réseaux de neurones sur graphe.

**Objectif supp #2** : mise en production du modèle Machine Learning en l'intégrant dans le code Fortran de LMDZ. Pour cela un état de l'art sera fait sur les solutions existantes (comme [Fortran-Keras Bridge](#), [SmartSim](#) ou le récent [TorchFort](#) de Nvidia) et une des solutions sera choisie et implémentée.

#### References:

- Watt-Meyer, O., Brenowitz, N. D., Clark, S. K., Henn, B., Kwa, A., McGibbon, J., et al. (2021). Correcting weather and climate models by machine learning nudged historical simulations. *Geophysical Research Letters*, 48, e2021GL092555. <https://doi.org/10.1029/2021GL092555>
- Krinner, G., Beaumet, J., Favier, V., Déqué, M., & Brutel-Vuilmet, C. (2019). Empirical run-time bias correction for Antarctic regional climate projections with a stretched-grid AGCM. *Journal of Advances in Modeling Earth Systems*, 11, 64–82. <https://doi.org/10.1029/2018MS001438>

Résumé en anglais (5 lignes) :

**This internship focuses on enhancing climate models, specifically addressing biases in the LMDZ atmospheric model of IPSL. The LMDZ model, currently experiences biases in representing atmospheric processes. The internship aims to employ machine learning to develop a dynamic correction model for the horizontal wind field errors. The goal is to reduce model biases by integrating the ML model into LMDZ simulations in free mode.**

Responsable du stage (Nom/prénom/statut) : **Sébastien Gardoll, CNRS/IPSL**

Laboratoire concerné : **LMD-IPSL**

Adresse à laquelle a lieu le stage : **4, place de Jussieu, 75005, Paris**

Equipe de recherche concernée (si pertinent) ou autre participant à l'encadrement du stage:

- **Redouane Lguensat, IRD/IPSL**
- **Olivier Boucher, CNRS/IPSL**
- **Frédérique Chérucy, CNRS/LMD-IPSL**

Niveau du stage ( M1, M2, internship) : **M2**

Thème scientifique de l'IPSL concerné : **SAMA**

Durée du stage : **5 ou 6 mois**

Période : **01/03/2024 30/09/2024**

Rémunération de l'ordre de **580 euros par mois**

Compétences souhaitées: **Python, bibliothèques deep learning: Keras ou PyTorch (et ses surcouches), bibliothèques d'analyse de données géoscience comme xarray, familiarité avec les outils Git/Bash/JupyterNotebooks. Des connaissances en modélisation atmosphérique sont souhaitables mais pas nécessaires.**

Est-il prévu une thèse dans le prolongement du stage ? **Suivant les résultats du stage**