

Métriques de similarité inspirées par la physique pour la prévision météo par IA

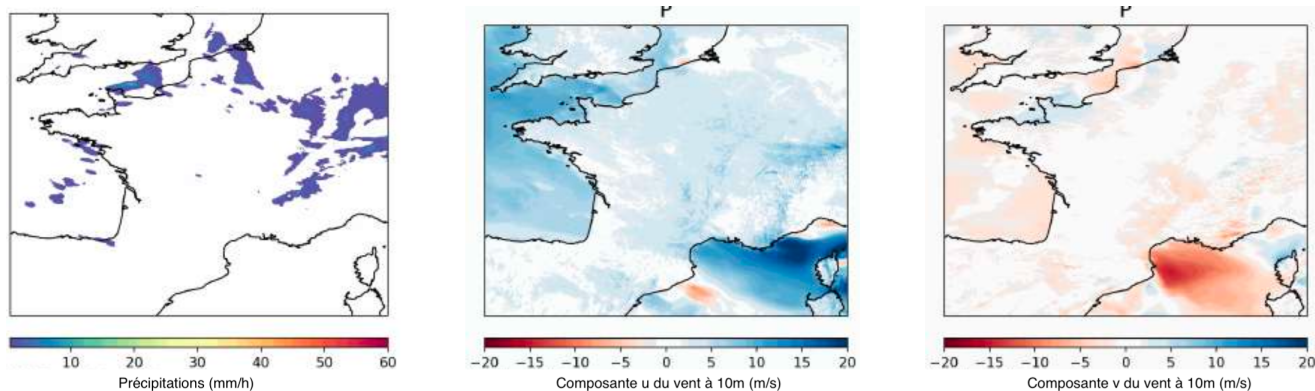
Résumé

L'utilisation de réseaux neuronaux a récemment renouvelé différents champs de l'ingénierie mathématique. En particulier, le domaine de la prévision météorologique a vu l'émergence de nouveaux acteurs issus du domaine de l'apprentissage automatique [Lam, 2022], [Kaifeng, 2023], dont les modèles de prévision numérique du temps (NWP) ont une qualité qui s'approche de celle des modèles conventionnels. Plus généralement, les réseaux neuronaux profonds (DNN) ont ouvert de nouvelles perspectives dans l'analyse ou la simulation des phénomènes physiques régis par les équations différentielles partielles (PDE) [Ruthotto & Haber, 2020], [Eliasof, 2021], [Raissi et al., 2019].

Une caractéristique clé des DNNs est que les mécanismes sous-jacents à leurs prédictions sont appris à partir de données d'apprentissage. Dans un contexte de prévision météorologique, un DNN prédit typiquement les futures conditions météorologiques à partir de l'état courant de l'atmosphère. L'apprentissage de ses règles de prédiction s'effectue en comparant des prédictions à ce qui a été observé dans des données réelles préalablement acquises. Une question au coeur du stage se pose alors : Quelles métriques de similarité (losses) utiliser pour comparer des prédictions météo à des observations réelles ?

Le ou la stagiaire travaillera sur l'utilisation de métriques plus adaptées aux données météo que des métriques purement locales comme les pertes d'erreur moyenne-carrée standard (MSE). De premières stratégies seront testées en appliquant les fonctions de pertes développées dans [Doury, 2022], [Swarzman, 2024] et [Subich, 2025] au système de prévision météorologique régional par DNN développé chez Météo-France. Une approche plus ambitieuse consistera ensuite à partiellement utiliser des équations fondamentales des modèles de prévision météorologique pour favoriser des prédictions physiquement cohérentes. A l'aide de l'équipe encadrante, la ou le stagiaire sera en charge de comprendre la pertinence de ces équations et de les interpréter sous forme de métriques de similarité pour l'apprentissage.

Les apprentissages et tests s'effectueront sur le jeu Titan de Météo-France <https://meteofrance.github.io/ai-lab-blog/weather%20forecasting/machine%20learning/python/dataset/2024/10/24/titan.html>. Les développements seront de même intégrés au package py4cast (<https://github.com/meteofrance/py4cast>) qui permet d'entraîner et d'inférer des prédictions



météo par DNN.

Utilisation du jeu Titan de Météo-France avec le package py4cast

Le stage se déroulera dans le cadre des travaux de la chaire ExplEarth du cluster IA ANITI. Il sera co-encadré par *Laurent Risser* de l'Institut de Mathématiques de Toulouse (IMT CNRS), *Laure Raynaud* du Centre National de Recherches Météorologiques (CNRM, CNRS et Météo-France) et *Luciano Drozda* du Centre Européen de Recherche et de Formation

Avancée en Calcul Scientifique (CERFACS). Le stage sera hébergé à cheval entre le CERFACS et l'IMT. Sous réserve de validation par le comité scientifique d'ANITI, le stage pourrait être poursuivi par une thèse à l'IMT en co-encadrement avec le CNRM et le CERFACS.

Profil recherché

La ou le stagiaire aura une formation solide en mathématiques appliquées aux sciences numériques (EDPs, signal, statistique, ...) et des connaissances générales en apprentissage automatique (machine learning). Une bonne maîtrise de Python sera nécessaire pour intégrer les stratégies développées durant le stage aux packages développés par Météo-France. Un intérêt pour les sciences atmosphériques et climatiques sera un plus.

Vous pouvez envoyer votre CV, un relevé de notes récent et une lettre de motivation courte à Laurent Risser (lrissier@math.univ-toulouse.fr) pour candidater.

Bibliographie

- [Lam, 2022] Lam R et al: GraphCast: Learning skillful medium-range global weather forecasting, Science 382, pp. 1461-1421. 2022
- [Kaifeng, 2023] Kaifeng B et al.: Accurate medium-range global weather forecasting with 3D neural networks. Nature 619, pp. 533-538. 2023
- [Ruthotto & Haber, 2020] : Deep neural networks motivated by partial differential equations. Journal of Mathematical Imaging and Vision, 62(3), 352–364. 2020
- [Eliasof, 2021] Eliasof M. et al.: PDE-GCN: Novel Architectures for Graph Neural Networks Motivated by Partial Differential Equations. Proc NeurIPS. 2021
- [Raissi et al., 2019] Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics 378. Pp. 686-707. 2019
- [Doury, 2022] Doury A. : PhD thesis: Estimation robuste du changement climatique régional : construction d'une approche hybride entre réseaux de neurones profonds et modèles de climat.
- [Swarcman, 2024] Swarcman D. et al: Quantizing reconstruction losses for improving weather data synthesis. Natrue Scientific reports. 2024
- [Subich, 2025] Subich et al.: Fixing the Double Penalty in Data-Driven Weather Forecasting Through a Modified Spherical Harmonic Loss Function. arXiv:2501.19374, 2025