Proposition de stage de fin d'étude école d'ingénieur (Centre National de Recherches Météorologiques)

Titre du stage

Utilisation des observations radar et foudre dans le fine-tuning de Météo-France IA pour la prévision d'évènements extrêmes précipitants

Description du sujet

- **Contexte**

Au cours des dernières années, des progrès significatifs ont été réalisés dans la prévision numérique du temps (PNT) par IA, à la fois à l'échelle mondiale et régionale (Ben Bouallègue et al., 2024 ; Alexe et al., 2024 ; Nipen et al., 2026). Plusieurs systèmes de prévision fondés sur l'IA ont démontré des compétences prédictives comparables à celles des systèmes de PNT les plus avancés et sont déjà employés de manière opérationnelle dans certains centres météorologiques (CEPMMT, DWD par exemple). Malgré ces progrès rapides, les systèmes de prévision reposant sur l'IA peinent encore à représenter la structure fine des phénomènes météorologiques, en particulier pour les phénomènes de méso-échelle à fort impact, comme la prévision des systèmes orageux et des évènements fortement précipitants.

Le modèle de prévision par IA de Météo-France, appelé « Météo-France IA », est un réseau de neurones sur graphe, basé sur des données d'analyses produites par le modèle opérationnel de prévision météorologique de Météo-France, AROME. Le réseau est entraîné dans une configuration dite de « grille étirée », avec un premier apprentissage global, puis un entraînement secondaire régional, nommé *fine-tuning*. Une solution envisagée pour améliorer la prévision des évènements fortement précipitants est d'intégrer à cette étape de *fine-tuning* des observations sensibles à la pluie et aux évènements extrêmes tels que les orages. Cette approche permet d'exploiter des données multi-sources, telles que des observations de natures différentes et des champs modèles, plutôt que de se baser exclusivement sur une seule source de données. Des travaux récents menés à l'Institut météorologique norvégien, par exemple, ont montré le potentiel de combiner des observations avec des informations météorologiques dans des systèmes de prévision basés sur les données (Nipen et al., 2025).

- **Objectifs du stage**

L'objectif de ce stage est d'étudier l'intérêt d'utiliser des observations sensibles aux variables précipitantes afin d'enrichir la version IA du modèle de PNT à l'échelle kilométrique AROME de Météo-France (Météo-France IA). Il s'agit ici des observations d'éclairs mesurées par le *lightning imager* du satellite MTG et les observations issues du réseau radar de Météo-France. Ce travail sera effectué dans le cadre du logiciel open-source ANEMOI, développé en collaboration avec plusieurs centres météorologiques, dont Météo-France. Ce logiciel permet le développement, l'entraînement et l'inférence de modèles de prévision par IA.

Le stage proposé pourra se diviser en plusieurs étapes :

1. Génération d'un jeu d'observations sur la France métropolitaine et intégration de ce jeu de données dans le logiciel ANEMOI. Plusieurs types d'observations complémentaires seront utilisées. Les données d'éclairs de l'imageur embarquées à bord du satellite géostationnaire européen MTG-LI seront utilisées pour essayer d'améliorer la prévision des éclairs. Pour les précipitations, les données du réseau radar de Météo-France seront exploitées (ex : réflectivités radars, vitesses Doppler).
2. Entraînement d'un nouveau modèle de prévision par IA intégrant à la fois les observations et les analyses AROME. Les prévisions de ce nouveau modèle seront comparées au modèle actuellement employé, qui inclus uniquement les analyses AROME. Ces évaluations pourront être effectuées sur des cas météorologiques à enjeux.
3. Afin d'identifier quelles sont les variables les plus informatives sur les champs prévus, des entraînements seront effectués en modifiant les combinaisons d'observations utilisées.
4. Enfin, cette étude permettra d'évaluer si les modèles d'IA permettent de mieux exploiter le contenu en informations apporté par ces observations, par rapport aux systèmes d'assimilation traditionnellement utilisés en prévision numérique du temps (méthode variationnelle 4DEnVar par exemple).

Un rapport écrit pour mettre en forme les résultats ainsi qu'une présentation orale sont attendus en fin de stage.

Responsables du stage

Responsable principal : Pauline Combarnous (équipe GMAP/OBS)
pauline.combarnous@meteo.fr 05.61.07.84.50

Autre responsable : Clément Brochet (équipe GMAP/PREV)
clement.brochet@meteo.fr 05.61.07.85.63

Informations complémentaires

Lieu du stage : Le stage aura lieu au Centre National de Recherches Météorologiques de Toulouse, unité mixte de recherche de Météo-France et du CNRS. Le candidat ou la candidate intégrera l'équipe « Observations » du Groupe de Modélisation et d'Assimilation pour la Prévision. Ce groupe a en charge le développement des futurs modèles opérationnels de prévision numérique du temps de Météo-France. L'équipe Observations développe l'intégration des observations météorologiques dans les modèles pour définir l'état initial de la prévision. Le stage bénéficiera aussi de l'expertise de l'équipe « Prévisibilité » qui est en charge de du développement du modèle de PNT par IA «Météo-France IA ». Le stage ne peut pas être effectué à distance et ne présente pas de caractère de confidentialité.

Durée : 6 au premier semestre 2027

Profil et compétences recherchées : élève d'école d'ingénieur avec une forte composante informatique. Bon niveau de programmation en Python. Attrait pour la météorologie/les sciences atmosphériques et la prévision numérique du temps. Capacités de communication, esprit de synthèse, autonomie et curiosité intellectuelle.

Un autre stage est proposé sur cette thématique, mais en explorant l'intérêt des observations des stations Netatmo pour Météo-France IA. Les deux stagiaires seront amenés à travailler en

collaboration car ils utiliseront les mêmes outils.

Bibliographie

Alexe, M., E. Boucher, P. Lean, E. Pinnington, P. Laloyaux, A. McNally et al. (2024). GraphDOP: Towards skilful data-driven medium-range weather forecasts learnt and initialised directly from observations. *arXiv* 2412.15687. https://doi.org/10.48550/arXiv.2412.15687

Ben Bouallègue, Z. et al. (2024). The Rise of Data-Driven Weather Forecasting: A First Statistical Assessment of Machine Learning–Based Weather Forecasts in an Operational-Like Context. *Bull. Amer. Meteor. Soc.*, **105**, E864–E883, https://doi.org/10.1175/BAMS-D-23-0162.1

Nipen, T. N. et al. (2026). Regional Data-Driven Weather Modeling with a Global Stretched Grid. *Artificial Intelligence for the Earth Systems*, 5(2). https://doi.org/10.1175/AIES-D-25-0001.1