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

Titre du stage

Utilisation d'observations de stations de surface Netatmo à haute densité pour la prévision par IA de phénomènes météorologiques à enjeux de fine échelle

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 en opérationnel dans certains centres météorologiques (CEPMMT, DWD par exemple). Malgré ces progrès rapides, les systèmes de prévision basés 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 ou les îlots de chaleur urbain.

Le modèle de prévision par IA de Météo-France, appelé « Météo-France IA », utilise un réseau de neurones sur graphe dans une configuration « grille étirée », avec un premier apprentissage global, puis un entraînement secondaire régional, nommé fine-tuning, à partir de champs modèle (analyses du modèle de PNT à échelle kilométrique AROME). Une option envisagée pour améliorer les prévisions de Météo-France IA est d'utiliser aussi des observations, en plus des analyses AROME. En effet, des approches émergentes en IA permettent d'exploiter des données multi-sources, comme des observations de natures différentes avec 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.

Les observations issues des stations Netatmo permettent d'avoir des observations in-situ de pression de surface, de température et d'humidité relative avec une très bonne densité spatiale sur la France. Une fois contrôlé, ce réseau d'observations permet de détecter finement les systèmes météorologiques localisés ou avec une forte variabilité spatiale (ex : fronts, îlot de chaleurs urbains, etc... ; Demortier et al. 2024, 2025). Ces observations semblent être prometteuses pour enrichir les bases de données utilisées pour entraîner Météo-France IA, notamment pour la prévision des événements météorologiques à fort enjeux.

- **Objectifs du stage**

L'objectif de ce stage est donc d'étudier l'intérêt d'utiliser des observations denses issues d'objets connectés (stations Netatmo) afin d'enrichir la version IA du modèle de PNT à l'échelle kilométrique AROME de Météo-France (Météo-France IA). 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 (Pression, température, humidité) sur la France métropolitaine issu du réseau Netamo et intégration de ce jeu de données dans le logiciel ANEMOI.
2. Entraînement d'un nouveau modèle de prévision par IA intégrant à la fois des observations NETATMO et les analyses AROME. Les prévisions de ce nouveau modèle seront comparées au modèle actuellement employé, qui inclus uniquement les analyses AROME.
3. Selon l'avancement du stage, une étude sera menée pour évaluer quelles sont les variables les plus impactantes sur les champs prévus. Des entraînements seront alors effectués soit en utilisant uniquement les champs observés par les données Netatmo, ou bien avec la pression de surface uniquement, ou encore les observations d'humidité. L'impact de la densité de stations pourra aussi être évalué en diminuant la densité spatiale des stations.
4. Enfin, cette étude permettra d'évaluer si les modèles d'IA permettent de mieux exploiter le contenu en informations apporté par les réseaux d'observations denses, 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 : Alan Demortier (équipe GMME/PRECIP)
alan.demortier@meteo.fr 05.61.07.93.44

Autres responsables : Mary Borderies, Clément Brochet, Suzanne Angeli

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 ».Ce stage sera aussi co-encadré par l'équipe « Précip» du Groupe de Météorologie de Moyenne Échelle (GMME) du CNRM, et qui détient l'expertise sur les observations de surface Netatmo. 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 en

intelligence artificielle et en traitement de larges volumes de données. 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 données radar et foudre 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

Demortier, A., Mandement, M., Pourret, V., & Caumont, O. (2024). Assimilation of surface pressure observations from personal weather stations in AROME-France. *Natural Hazards and Earth System Sciences*, 24(3), 907-927. https://nhess.copernicus.org/articles/24/907/2024/nhess-24-907-2024.html

Demortier, A., Mandement, M., Pourret, V., and Caumont, O. (2025). Assimilation of temperature and relative humidity observations from personal weather stations in AROME-France, *Nat. Hazards Earth Syst. Sci.*, 25, 429–449, https://doi.org/10.5194/nhess-25-429-2025

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