

M2 SOAC : Fiche de stage de recherche en laboratoire

Laboratoire : CNRM

Titre du stage : Évaluation de l'impact des différents types d'aérosols sur la qualité de l'air par application des méthodes d'intelligence artificielle

Nom et statut du (des) responsable (s) de stage : Laaziz El Amraoui & Thomas Rieutord (CNRM, Toulouse)

Coordonnées (téléphone et email) du (des) responsable (s) de stage :

- Laaziz El Amraoui, 05 61 07 97 67, laaziz.elamraoui@meteo.fr
- Thomas Rieutord, 05 61 07 96 53, thomas.rieutord@meteo.fr

Sujet du stage :

Ce sujet de stage M2 s'inscrit dans le cadre de l'exploration de nouvelles méthodes basées sur des techniques d'intelligence artificielle afin de mieux évaluer l'impact des différents types d'aérosols, primaires et secondaires, sur la qualité de l'air aussi bien à l'échelle locale qu'à l'échelle régionale.

La qualité de l'air est souvent impactée par des événements extrêmes tels que le transport de poussières désertiques, feux de biomasse ou d'éruptions volcaniques qui sont considérés comme d'importantes sources d'aérosols. De ce fait, la surveillance de ces polluants particulaires est primordiale afin d'appréhender leurs effets en particulier sur la santé humaine. La recherche en matière de qualité de l'air est généralement basée sur l'utilisation combinée des observations et des modèles via des techniques d'assimilation de données. Une chaîne pré-opérationnelle d'assimilation des épaisseurs optiques d'aérosols (AOD pour *Aerosol Optical Depth*) a été développée au CNRM. Elle fournit quotidiennement des analyses concernant les différents types d'aérosols en termes d'épaisseurs optiques et de concentrations y compris les PM_{2.5} et les PM₁₀ (matières particulaires dont le diamètre est inférieur respectivement à 2.5 et 10 µm). Les concentrations de ces deux dernières quantités sont souvent utilisées comme des indicateurs pour évaluer à la fois la qualité de l'air ainsi que les dépassements de seuils.

L'objectif de ce stage M2 est d'explorer la possibilité d'utiliser des techniques d'intelligence artificielle pour mieux prévoir des indicateurs de qualité de l'air basés sur les concentrations des PM_{2.5} et PM₁₀ en utilisant les analyses d'AOD issues de la chaîne pré-opérationnelle comme prédicteurs. La première étape du stage consistera à se familiariser avec les données, identifier d'éventuels régimes de qualité de l'air (par exemple : bon, moyen, mauvais) à l'aide de classification non-supervisée sur les données de PM. Ensuite, nous implémenterons plusieurs méthodes de classification supervisées pour prévoir le régime de qualité de l'air en fonction des données issues de la chaîne sur un domaine régional (Europe). Enfin, nous mettrons en concurrence les différentes méthodes implémentées et nous choisirons la plus robuste avant de pouvoir l'évaluer sur une longue période et comparer sa performance par rapport à celle de la chaîne. Pour cela, nous disposons d'archives de la chaîne pré-opérationnelle sur une profondeur de deux ou trois ans. En fonction de l'avancement du stage, nous étudierons l'apport respectif de chaque prédicteur et le bénéfice d'en apporter de nouveaux (par exemple la hauteur de couche limite) ainsi que le bénéfice de regrouper certains points de grilles aux caractéristiques similaires, ou encore de distinguer les différences espèces chimiques dans la classification.

Le (la) candidat(e) retenu(e) sera accueilli(e) dans l'équipe PLASMA, spécialisée en chimie atmosphérique et responsable de la chaîne pré-opérationnelle. Il (elle) devra être à l'aise avec le langage Python et l'environnement Linux. Une connaissance des bibliothèques classiques de machine learning (Scikit-learn, Pandas, Numpy, Matplotlib, Seaborn...) et/ou une expérience préalable dans la manipulation des fichiers NetCDF sera un avantage. Il (elle) utilisera les outils de travail collaboratif (git, Overleaf) qui faciliteront son suivi par ses deux encadrants.