***Version française***

Proposition de Sujet de thèse 2021

Laboratoire (et n° de l’unité) dans lequel se déroulera la thèse :

**CNRM UMR 3589 (Météo-France, CNRS, Univ. de Toulouse)**

Titre du sujet proposé :

Synergie entre observations issues de capteurs passifs, actifs et in-situ pour la caractérisation des aérosols à l'échelle régionale : Méthodes d'assimilation et d’intelligence artificielle

**Nom et statut** (PR, DR, MCf, CR, …) du *(des)* responsable*(s)* de thèse **(préciser si HDR)** :

- EL AMRAOUI Laaziz (CR – HDR)

**Coordonnées** (téléphone et e-mail) du (des) responsable(s) de thèse :   
[laaziz.elamraoui@meteo.fr](mailto:laaziz.elamraoui@meteo.fr) (Tél. 05-61-07-97-67)

**Cadre de la thèse :**

Ce sujet de thèse est proposé dans le cadre de l’amélioration de la prévision de la qualité de l’air particulaire à l’échelle régionale et locale. 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. 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. Ces techniques ont l‘avantage d’être précis mais représentent l’inconvénient de faire appel à beaucoup de ressources. D’autres approches plus innovantes basées sur des méthodes d’intelligence artificielle commencent à voir le jour. Elles ont l’avantage de fournir des prévisions assez rapidement avec un moindre coût, mais en contrepartie les prévisions peuvent généralement présenter quelques biais selon le degré d’apprentissage et le choix des prédicteurs.

Ainsi, les objectifs de la thèse sont comme suit :

1. Évaluer la valeur ajoutée de l’utilisation, au sein d’un système d’assimilation, des observations issues de capteurs passifs en synergie avec les mesures de capteurs actifs et in-situ pour profiter de la complémentarité de l’ensemble des observations.
2. Mettre en place un système de prévision d’indicateurs de qualité de l’air issu des techniques d’intelligence artificielle et basé sur l’exploitation de l’ensemble des observations satellite et in-situ ainsi que les paramètres météorologiques clés comme prédicteurs.
3. Évaluer et comparer les performances des deux approches de prévision de la qualité de l’air dans un contexte pré-opérationnel.

**Nature du travail attendu :**

Les méthodologies actuelles concernant la détection des aérosols s’appuient sur des observations issues de capteurs passifs (imageurs) ou actifs (type lidars). Les imageurs se caractérisent généralement par une bonne résolution horizontale qui permettent de générer des cartes d'aérosols avec une résolution verticale limitée. De l’autre côté les lidars ont une très bonne résolution verticale permettant la mesure de profils verticaux des propriétés des aérosols, mais avec une résolution spatiale très faible. Quant à la modélisation, elle a l’avantage de fournir à chaque point de grille du modèle une représentation spatio-temporelle réaliste et détaillée des aérosols. Cependant, les modèles sont généralement confrontés à des problèmes liés essentiellement à la condition initiale, la résolution spatiale et à l’inventaire des émissions, ce qui peut induire un biais quant à la distribution des aérosols. Pour remédier à ces difficultés, on fait souvent appel à l’assimilation de données qui consiste à corriger l’état global du modèle en y injectant des observations de façon continue au cours de son intégration.

Dans un premier temps, les performances de la complémentarité entre les mesures issues de capteurs passifs et celles issues de lidars au sein du système d’assimilation seront évaluées. Les deux types d’observations seront assimilées séparément et conjointement. Une attention particulière sera apportée à l’étude des conséquences du confinement liées au Covid-19 sur la qualité de l’air en région d’Occitanie afin de contribuer aux efforts qui visent à améliorer les connaissances quant au rôle de chaque type d’aérosols dans la dégradation de la qualité de l’air régionale et locale.

Dans un deuxième temps une autre approche de prévision de la qualité de l’air basée sur des méthodes d’intelligence artificielle sera utilisée afin de la mettre en concurrence avec la méthode présentée précédemment. L’objectif est d’explorer la possibilité d’utiliser ce genre de techniques pour mieux prévoir des indicateurs de qualité de l’air basés essentiellement sur les concentrations des PM2.5 et PM10 en utilisant les différentes quantités issues des différentes observations satellite et in-situ comme prédicteurs. Plusieurs méthodes de classification supervisées pour prévoir le régime de qualité de l’air en fonction des données issues des différentes observations seront testées et évaluées, la méthode la plus performante sera choisie et implémentée.

La dernière étape de la thèse consistera à comparer et évaluer les deux approches. L’objectif étant d’optimiser les ressources et de désigner la méthode la plus performante dans un contexte de prévision pré-opérationnelle de qualité de l’air à l’échelle régionale et locale. Le doctorant travaillera à plein temps sur le sujet proposé au sein du laboratoire d'accueil et sera encadré par les responsables scientifiques, il devra ainsi :

1. Réaliser un travail bibliographique.
2. Prendre en main le modèle MOCAGE et son système d’assimilation des aérosols.
3. Se familiariser avec les différentes méthodes d’intelligence artificielle
4. Mettre en concurrence les deux approches de prévision de qualité de l’air dans un contexte pré-opérationnel
5. Valoriser les résultats par des présentations de conférences et des publications
6. Procéder à la rédaction du manuscrit et à la soutenance.

***English version***

Proposition de Sujet de thèse 2021

Name of the laboratory :

**CNRM UMR 3589 (Météo-France, CNRS, Univ. de Toulouse)**

Title of the thesis :

Synergy between observations from passive, active and in-situ sensors for aerosol characterization at regional scale: Assimilation and artificial intelligence methods

**PhD supervisor / contact** :

- EL AMRAOUI Laaziz ( [laaziz.elamraoui@meteo.fr](mailto:laaziz.elamraoui@meteo.fr) ; Tél. 05-61-07-97-67)

**Background**

This thesis is proposed in the framework of improving the prediction of particulate air quality at regional and local scales. Air quality is often influenced by extreme events such as desert dust transport, biomass fires or volcanic eruptions. Therefore, monitoring of these particulate pollutants is essential to understand their effects, especially on human health. Air quality research is generally based on the combined use of observations and models through data assimilation techniques. These techniques have the advantage of being accurate but have the disadvantage of being resource intensive. Other more innovative approaches based on artificial intelligence methods are beginning to emerge. They have the advantage of providing forecasts fairly quickly and at a lower cost, but on the other hand the forecasts can generally present some bias depending on the degree of learning and the choice of predictors

Therefore, the objectives of the thesis are as follows:

1. Evaluate the added value of using, within an assimilation system, observations from passive sensors in synergy with measurements from active and in-situ sensors in order to take advantage of the complementarity of all observations.
2. Implement an air quality indicator forecasting system based on artificial intelligence techniques and using all satellite and in-situ observations as well as key meteorological parameters as predictors.
3. Evaluate and compare the performance of the two air quality forecasting approaches in a pre-operational context.

**Expected work:**

Current methodologies for aerosol detection are based on observations from passive (imagers) or active sensors (lidar). Imagers are generally characterized by a good horizontal resolution which allows to generate aerosol maps with a limited vertical resolution. On the other hand, lidars have a very good vertical resolution allowing the measurement of vertical profiles of aerosol properties, but with a very low spatial resolution. For modeling, it has the advantage of providing a realistic and detailed representation of aerosols in space and time at each grid point of the model. However, the models are generally confronted with problems related mainly to the initial condition, the spatial resolution and the emission inventory, which may induce a bias in the aerosol distribution. To overcome these difficulties, data assimilation is often used, which consists in correcting the global state of the model by injecting observations continuously during its integration.

The first step of the thesis will consist of evaluating the performance of the complementarity between passive sensor and lidar measurements within the assimilation system. The two types of observations will be assimilated separately and jointly. Particular attention will be paid to the study of the consequences of the Covid-19 lockdown on the air quality in the Occitanie region in order to contribute to the efforts to improve the knowledge of the role of each type of aerosols in the degradation of the regional and local air quality.

In a second step, another air quality forecasting approach based on artificial intelligence methods will be used to compare with the previously presented method. The objective is to explore the possibility of using such techniques to better predict air quality indicators based mainly on PM2.5 and PM10 concentrations using the different quantities from the different satellite and in-situ observations as predictors. Several supervised classification methods for predicting the air quality regime based on the data from the different observations will be tested and evaluated, the best performing method will be selected and implemented.

The last step of the thesis will be to compare and evaluate the two approaches. The objective is to optimize the resources and to designate the most efficient method in a context of pre-operational air quality forecasting at regional and local scales. The PhD student will work full time on the proposed topic within the host laboratory. He will have to :

1. Conduct a bibliographic work on the data assimilation and artificial intelligence techniques applied to the air quality topic.
2. Be trained on the MOCAGE model and its aerosol assimilation system.
3. Become familiar with the different artificial intelligence methods
4. Compare the two air quality forecasting approaches in a pre-operational context
5. Valorize the results through conference presentations and publications
6. Write the manuscript and defend the thesis.