

Titre

Évolution des observables lidar nuage en fonction de leur environnement quand le climat se réchauffe : utilisation du *machine learning*

Nom de l'encadrante

Marjolaine Chiriaco (LATMOS, marjolaine.chiriaco@latmos.ipsl.fr)

Lieu de travail

LATMOS, Guyancourt

Niveau requis

Stage de Master 2

Savoirs/connaissances nécessaires ou souhaités

Formation initiale en physique, Master en sciences du climat. Connaissances en analyse de données, des notions de statistiques seraient un plus.

Qualités humaines attendues

Curiosité scientifique, esprit d'initiative, travail en équipe

Interactions/mobilité éventuelle

Collaboration avec le Laboratoire de Météorologie Dynamique (École Polytechnique, Palaiseau) et avec le Laboratoire d'Aérodynamique (Toulouse)

Description du Sujet :

En s'appuyant sur de nombreuses publications scientifiques, les rapports du GIEC réaffirment les uns après les autres que les rétroactions nuageuses sont responsables de la plus grande part d'incertitudes dans la prévision du climat futur. Récemment, l'article de synthèse de Sherwood et al. (2020)¹ conclut que pour progresser dans notre connaissance des rétroactions climatiques, il faut s'appuyer sur des observations satellite de haute qualité via des nouveaux diagnostics nuageux communs aux observations et aux modèles. Les travaux menés dans le cadre de la mission spatiale CALIPSO lancée en 2006 (et toujours active) et ayant à son bord un lidar ont permis de définir plusieurs de ces nouveaux diagnostics². L'objectif de ce stage est de **caractériser comment les nuages évoluent quand leur environnement change en se basant sur ces nouveaux diagnostics lidar**. Nous pensons en effet qu'en caractérisant au mieux comment ces diagnostics évoluent en fonction de la variabilité naturelle environnante, nous serons à même de mieux comprendre comment ils évoluent si l'environnement change sous l'effet du réchauffement climatique.

Les diagnostics lidar utilisés seront la quantité de nuages optiquement fins C_{thin} , la quantité de nuages opaques C_{opaque} , et l'altitude à laquelle le signal lidar devient complètement atténué Z_{opaque} . Si les méthodes d'analyse utilisées le permettent, nous utiliserons également le signal lidar en lui-même. Ces grandeurs présentent l'avantage d'être déduites directement de la mesure, sans inversion du signal et donc avec un minimum d'hypothèses. Concernant la variabilité de l'environnement, nous utiliserons un grand nombre de variables issues des réanalyses ERA5. Toutes ces données sont déjà

¹ <https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1029/2019RG000678>

² <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2016JD025946>

disponibles au bon format sur le serveur de l'IPSL : le ou la stagiaire pourra donc se concentrer sur l'analyse des données.

Les données seront analysées par la méthode de *machine learning* des Forêts Aléatoires. Cette méthode permet de déterminer les relations non linéaires entre des variables, ce qui est le cas ici, et ceci sans *a priori*. Nous chercherons comment les précurseurs (ici les N variables issues des réanalyses ERA5) agissent sur le diagnostic lidar. La méthode des forêts aléatoires permet également d'avoir accès au poids de chaque précurseur, un résultat important pour la question posée dans ce stage.

Les étapes du travail seront les suivantes :

- 1) Bibliographie et prise en main des jeux de données ;
- 2) Application de la méthode des forêts aléatoires au diagnostic lidar Z_{opaque} pour la zone Euro-Méditerranée ;
- 3) Reprise des analyses du 2) en fonction des régimes de temps qui influencent le climat en Europe de l'ouest.
- 4) Extension aux autres diagnostics lidar et à d'autres zones du globe clé pour le changement climatique : Pôles, bande tropicale, *Warm Pool*, *Southern Ocean*.

Outils utilisés

Jeux de données satellites CALIPSO-GOCCP, outils statistiques type forêts aléatoires, réanalyses ERA5. La programmation pourra se faire en Matlab, ou préférentiellement en Python si le ou la stagiaire est autonome avec ce langage.

Acquisitions envisagées

Variabilité du climat, télédétection, analyse de données, observations, programmation Matlab ou Python

Possibilité de poursuivre en thèse : Oui

Informations pratiques

- Indemnités de stage : Environ 550 euros par mois ;
- Durée du stage : 4 à 6 mois ;
- Possibilité d'apporter une aide pour trouver un logement à proximité via le CROUS.