

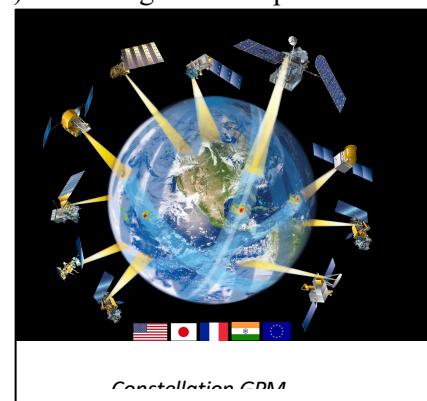
## Offre de Stage IPSL 2022

(Soutenu par le programme EUR IPSL-Climate Graduate School)

Titre du sujet de stage :

Observation spatiale des précipitations : Estimation par apprentissage profond de la qualité du produit IMERG

L'estimation des champs précipitants depuis l'espace à partir de satellites de télédétection est un enjeu primordial tant sur la caractérisation globale des précipitations que sur son évolution dans un contexte de changement climatique. La mission GPM (Global Precipitation Measurement) a pour objectif de fournir des produits de précipitations à l'échelle du globe à partir d'une constellation de satellites de télédétection défilants à basse altitude (LEO) munis de capteurs hyperfréquences [1]. Parmi ces produits, nous nous intéressons au produit IMERG (Integrated Multi-satellitE Retrievals for GPM) estimé principalement à partir des observations de la constellation GPM ainsi qu'à partir d'observations en infrarouge issues de satellites d'observation non défilants en orbite géostationnaire (GEO). L'avantage d'un tel produit est de fournir des observations des précipitations à l'échelle globale avec une résolution temporelle de 30 minutes et une résolution spatiale d'une dizaine de kilomètre sous nos latitudes, le rendant ainsi particulièrement adapté pour des études à l'échelle régionale. La qualité de restitution de ce produit dépend de plusieurs facteurs inhérents à la mesure (relation complexe entre températures de brillance issues des radiomètres spatiaux et intensité de pluie) mais aussi à la configuration géométrique de la constellation de satellites. En effet le nombre de passages journaliers des satellites sur une zone géographique donnée est relativement limité (temps de revisite de 2 à 3 heures en moyenne), ce qui est peu compatible avec la dynamique des systèmes précipitants. Il en résulte un sous échantillonnage spatio temporel important. Afin de remédier à ce problème, l'algorithme de restitution IMERG utilise des observations additionnelles en infrarouge (IR) issues des satellites GEO. Cependant, ces derniers sont plutôt dédiés à l'observation des nuages qu'aux précipitations. Au final, des erreurs d'estimation importantes peuvent exister tant sur la localisation des cellules de pluie que sur leur intensité. Afin d'évaluer la qualité de ce produit, celui-ci a fait l'objet de plusieurs études dans lesquelles les cartes de précipitation IMERG sont comparées à des observations de pluviomètres ou de radars météorologiques [2][3]. Il apparaît toutefois que ces études n'utilisent pas les informations complémentaires fournies par ce produit. Parmi celles-ci, certaines sont fondamentales pour l'estimation de la qualité d'une carte comme par exemple l'heure de passage des satellites ou l'introduction ou non d'information IR.



L'objectif du stage proposé est d'évaluer la qualité du produit IMERG au-dessus de la France métropolitaine en se basant sur un produit pluie de référence fourni par Météo France issu du réseau de radars météorologiques ARAMIS. Il s'agira d'estimer pour chaque pixel de pluie d'une carte donnée un indicateur de qualité. Le modèle à développer s'appuiera sur une base de quatre années d'observation d'une zone de 800x800 km<sup>2</sup> centrée sur la France. Une approche à base de réseaux de neurones profonds (U-Net par exemple [4]) devra être mise en œuvre et permettra de fusionner l'ensemble des informations (hétérogènes) disponibles dans le produit IMERG. Nous envisagerons trois approches possibles, une approche du type régression quantile [5] et permettant d'estimer un intervalle de confiance, l'estimation de la variance de l'erreur, et l'approche Expanded Interval Minimization (EIM) [6]. Le déroulement envisagé est le suivant :

- Travail bibliographique sur le produit IMERG ainsi que sur les études existantes.
- Prise en main des données et familiarisation avec l'environnement de travail : Des travaux antérieurs basés sur les données IMERG et Météo France ayant été menés au laboratoire, le stagiaire pourra s'appuyer sur les codes Python développés et se familiariser avec l'environnement de travail (Serveur GPU de l'IPSL, framework Pytorch).
- Développement d'un premier modèle à l'aide de réseaux de neurones convolutionnels (régression quantile), validation, analyse des performances
- Développement du second modèle et comparaison des performances avec le premier modèle
- En fonction du temps disponible : développement du troisième modèle et comparaison avec les autres modèles

**Pré-requis** : Machine learning, programmation (python), environnement de travail du type Pytorch, Keras

**Environnement de recherche** : Le stagiaire sera intégré au sein du département SPACE dans une équipe de chercheurs composée de spécialistes dans l'observation et la modélisation des précipitations ainsi qu'en modélisation statistique (machine learning) dans le domaine du cycle de l'eau atmosphérique

**Poursuite en thèse** : une bourse de thèse sera demandée sur un sujet connexe

## Bibliographie

- [1] NASA precipitation Measurement Mission <https://pmm.nasa.gov/gpm>
- [2] Sungmin, O, S., Foelsche, U., Kirchengast, G., Fuchsberger, J., Tan, J., and Petersen, W. A.: Evaluation of GPM IMERG Early, Late, and Final rainfall estimates using WegenerNet gauge data in southeastern Austria, Hydrol. Earth Syst. Sci., 21, 6559-6572, <https://doi.org/10.5194/hess-21-6559-2017>, 2017.
- [3] LINDA BOGERD, AART OVEREEM, HIDDE LEIJNSE, AND REMKO UIJLENHOET : A Comprehensive Five-Year Evaluation of IMERG Late Run Precipitation Estimates over the Netherlands, Journal of Hydrometeorology, 2021, 22, 1855–1868
- [4] Ronneberger, Olaf; Fischer, Philipp; Brox, Thomas, U-Net: Convolutional Networks for Biomedical Image Segmentation, <https://arxiv.org/pdf/1505.04597.pdf> , 2015
- [5] Koenker, Roger, and Kevin F. Hallock. 2001. "Quantile Regression." Journal of Economic Perspectives, 15 (4): 143-156. DOI: 10.1257/jep.15.4.143
- [6] Dongqi Su, Ying Yin Ting and Jason Ansel : Tight Prediction Intervals Using Expanded Interval Minimization, 2018, arXiv:1806.11222.

Résumé en anglais (5 lignes) :

Responsable du stage (Nom/prénom/statut) : Laurent Barthes

Collaborations spécialistes ML de l'équipe SPACE du LATMOS

Contact Laurent Barthes ([laurent.barthes@latmos.ipsl.fr](mailto:laurent.barthes@latmos.ipsl.fr)),

Laboratoire concerné :LATMOS

Equipe de recherche concernée (si pertinent) : SPACE (Statistique Processus Atmosphere Cycle de l'EAU)

Niveau du stage (Licence, M1, M2, internship) : M2

Licence ou Master(s) où sera proposé le sujet : formations **de master et écoles d'ingénieur**

Thème scientifique de l'IPSL concerné : SAMA et Cycle de l'eau

Gratification de stage : 580€/mois environ+ indemnité pass navigo

Période et durée : 5-6 mois à partir de mars 2022

Période : 01/02/2022 → 30/09/2022

Est-il prévu une thèse dans le prolongement du stage ? un financement de thèse sur un sujet également lié à l'utilisation des algorithmes d'apprentissage profond pour l'observation spatiale des précipitations sera demandé au programme UdopIA de Paris-saclay