

PROPOSITION DE STAGE 2021-2022

Niveau M2 ou dernière année d'École d'ingénieur

Apprentissage profond pour l'estimation de variables biophysiques du couvert forestier en milieu tropical

Etablissement : CIRAD - UMR Tetis

lieu du stage : Maison de la Télédétection, 500, rue Jean François Breton – Montpellier (34)

Contexte

Le suivi de l'état des forêts tropicales est d'importance mondiale (Herold et al. 2019). Il appelle le suivi dans la durée et sur de grandes surfaces de variables « biophysiques » de structures forestières (couvert, hauteur de canopée, biomasse épigée, ...) et d'indicateurs de diversité spécifique et fonctionnelle. Le suivi du couvert arboré, notion polysémique, a fait l'objet de travaux importants, principalement basés sur les séries d'images satellitaires longues à moyenne résolution ou haute résolution spatiale (MODIS, Landsat, respectivement), Ceci a débouché sur des produits cartographiques ambitieux à des échelles pantropicales et mondiales (e.g. Avitabile et al. 2012; Hansen et al. 2013 ; Kubayashi et al. 2016). Cependant, ces produits recèlent des faiblesses de plusieurs ordres. Si celles-ci n'opèrent pas leur intérêt, tant que l'ampleur des surfaces peut permettre la compensation des imprécisions locales, elles rendent leurs utilisations hasardeuses pour des applications demandant de se focaliser sur des surfaces plus restreintes (*downscaling*), ou requérant un niveau de précision supérieur à celui effectif (qui n'est pas toujours explicite, d'ailleurs).

Le projet **Sé2coul**, a pour objectif de traduire les opportunités, offertes par les données Sentinel-1 et Sentinel-2, pour le suivi et la caractérisation des forêts, en tirant partie :

- i) de la densification temporelle des séries temporelles d'images satellitaires et de leur accroissement en résolution spatiale. Ceci doit permettre d'aller bien au-delà de ce qui peut déjà être fait actuellement avec les séries HR de type « Landsat » (par exemple via le Google Earth Engine) pour corriger les effets instrumentaux dans l'optique de détection de changements francs de couvert (déforestation, feu, plantation, ...), ou de détections de perturbations localisées (exploitation, chute d'arbres, ...).
- ii) de la facilitation de l'usage conjoint des données radar et optiques (Sentinel-1 et -2). En complément de l'optique, l'information radar en bande C (bien que probablement saturante à des niveaux assez bas

de couvert et de biomasse) est sensible à la structure tridimensionnelle, et peut aider à lever des ambiguïtés dans les végétations complexes, ouvertes (par exemple pour le suivi de la dynamique des trouées, naturelles ou non), ou marquées par des variations saisonnières d'humidité (y compris du substrat).

iii) du développement de techniques d'intelligence artificielle (deep-learning) pour tirer parti des atouts des données Sentinel dans la perspective spécifique de la caractérisation biophysique des composantes ligneuses de la végétation dans des contextes tropicaux variés.

Le projet aura pour objectif de traduire ces opportunités en avancées significatives concernant l'estimation de caractéristiques fondamentales de la végétation ligneuse : (i) structure : Hauteur de canopée, LAI, etc. (ii) diversité, (iii) signature fonctionnelle saisonnière.

Bibliography

Herold M, Carter S, Avitabile V et al (2019) The role and need for space-based forest biomass-related measurements in environmental management and policy. *Surv Geophys.*

Avitabile, V.; Baccini, A.; Friedl, M.A.; Schullius, C. Capabilities and limitations of Landsat and land cover data for aboveground woody biomass estimation of Uganda. *Remote Sens. Environ.* 2012, 117, 366–380

Hansen MC, Potapov PV, Moore R et al (2013) High-resolution global maps of 21st-century forest cover change. *Science* 342:850–

Kobayashi T., Tsens-Ayushi J., Tateishi R. (2016). A new global tree-cover percentage map using MODIS data. *International Journal of Remote Sensing* 37(4): 969-992.

Objectifs du Stage

En s'appuyant sur les résultats actuels du projet, qui ont permis d'évaluer le potentiel des images Sentinel pour l'extraction de variables biophysiques sur la structure et la diversité des couverts forestiers, l'objectif de ce stage est de tester des solutions basées sur l'apprentissage profond pour : (i) s'affranchir du problème de la forte nébulosité des images disponibles en optique (Sentinel-2) en s'appuyant sur l'utilisation conjointe d'images optiques et radar; (ii) mettre en place des solutions permettant la mise en opération de la production cartographique de ces variables; (iii) poser les bases pour l'extraction d'information à valeur ajoutée (cartographie de la typologie des couverts forestiers, extraction de traits fonctionnels par analyse de séries temporelles). Le terrain d'étude en Guyane Française sera ciblé en priorité pour tester ces méthodes.

Pour cela, nous allons mobiliser en grande partie des méthodes déjà développées au sein de notre unité, concernant à la fois l'extraction de variables biophysiques et la génération de séries temporelles d'images Sentinel-2 dénuagées, avec l'objectif de les combiner opportunément et de tester la production de cartes sous différentes contraintes en termes de disponibilité d'images (e.g. nébulosité). Selon les acquis de cette phase, un deuxième objectif serait de travailler sur des séries temporelles d'images optiques dénuagées pour évaluer des premières méthodes d'extraction d'information à valeur ajoutée.

Description du contenu du stage

Les principales étapes envisagées pour la réalisation du stage sont :

- Analyse bibliographique sur les méthodes d'apprentissage profond pour la synthèse d'images sans nuages par couplage radar/optique, et sur l'état de l'art concernant l'extraction de variables biophysiques du couvert forestier.
- Mise en place du jeu de données :
 - Prise en main des données existantes
 - Téléchargement et prétraitements de série temporelle d'images Sentinel-1 et Sentinel-2
 - Préparation du jeu de données pour l'évaluation
- Prise en main des méthodes d'apprentissage profond qui sont à mobiliser.
- Mise en place de telle méthode dans le contexte de la cartographie de la typologie du couvert forestier sur la zone d'étude.
- Évaluation de la méthode déployée à travers une analyse quantitative/qualitative.
- Rédaction du rapport et préparation de la soutenance.

Profil souhaité :

- Master II ou 3ème année d'école ingénieur en Traitement du Signal, Informatique ou Télédétection
- Compétences : programmation informatique et géomatique
- Bonne connaissance des méthodes de machine et deep learning (TensorFlow/PyTorch, Keras, ...)
- Goût pour la recherche scientifique et l'interdisciplinarité
- Maîtrise de l'anglais et bonne capacité rédactionnelle

Conditions de réalisation du projet

- Durée du projet : 6 mois à partir du 1er Avril (possibilité de soutenir avant la fin de stage)
- Rémunération : 600 € / mois environ

Contact pour le stage:

- Dr. Raffaele Gaetano: raffaele.gaetano@cirad.fr