



## PROPOSITION DE CDD

**Intitulé :** Estimation des traits fonctionnels de forêts méditerranéennes : évaluation pluri-capteurs selon le taux de couverture boisée

**Début du contrat :** dès que possible

**Date limite de candidature :** quand poste pourvu

**Durée :** 12 mois (possibilité de renouvellement de 12 mois dans le cadre du projet)

### Mots clés

Imagerie spectroscopique, Sentinel-2, données LiDAR, modélisation 3D, transfert radiatif, inversion, forêts méditerranéennes

### Profil et compétences recherchées

Formation : traitement du signal/image ou de données de télédétection, méthodes de machine learning, écologie/environnement, programmation en python

Notions en modélisation du transfert radiatif, traitement de l'information 3D LiDAR

Une première expérience sur l'utilisation du code DART (mais pas un prérequis)

### Présentation du projet post-doctoral, contexte et objectif

Les forêts méditerranéennes couvrent une surface relativement restreinte à l'échelle mondiale pour une forte concentration en diversité végétale, avec notamment des espèces endémiques qui ont su s'adapter aux aléas du climat méditerranéen (été long, chauds et sec, hiver doux et humide). Cependant, il s'agit de zones critiques, car d'ici 2100 leur biodiversité changera beaucoup plus que celle d'autres écosystèmes terrestres [1], en particulier du fait de la recrudescence des incendies, sécheresses, et exploitations anthropiques (urbanisation, bois, terres agricoles) [2]. Il est donc important de suivre l'évolution temporelle de ces forêts et l'état de santé de certaines de leurs espèces d'arbres, pour pouvoir guider les politiques environnementales et préserver la biodiversité de ces milieux fragiles. Les observations de télédétection permettent à la fois de répondre à cette problématique et une cartographie à grande échelle des variables essentielles de biodiversité (EBV) définies par des groupes d'experts scientifiques (GEO-BON)[3] et notamment l'étude de la classe "EBV-species traits" qui vise à caractériser les traits fonctionnels des espèces à travers leurs attributs phénologiques, morphologiques et physiologiques. Le suivi de ces traits est en partie réalisable à travers l'estimation de certaines propriétés biophysico-chimiques à l'échelle de la feuille et de la canopée accessibles par télédétection. Parmi ces propriétés, les plus couramment étudiées sont les pigments foliaires (CAB - chlorophylles et CAR - caroténoïdes), le contenu en eau et en matière sèche foliaires (EWT, LMA), et l'indice foliaire ou LAI (Leaf Area Index). Sur un cycle phénologique de la plante, ces derniers représentent des indicateurs permettant d'évaluer un stress hydrique, d'observer des évolutions de biomasse et une défoliation précoce.

Face à la multiplication des données de télédétection des satellites actuels et futurs, un enjeu majeur est d'évaluer le potentiel de la combinaison de tous ces jeux de données pour mieux décrire et étudier la végétation, au vu des caractéristiques spectrales, spatiales et temporelles des différents capteurs [4]. Pour notre cas d'étude, l'estimation simultanée des propriétés biophysico-chimiques précitées nécessite une grande richesse spectrale pour tirer parti de leur sensibilité sur l'ensemble du spectre mesuré. Avec son plus grand nombre de bandes spectrales, l'imagerie hyperspectrale sur [0,4-2,5  $\mu\text{m}$ ] est plus adaptée que l'imagerie multispectrale qui a cependant souvent un temps de revisite plus court. Beaucoup d'études ont déjà évalué les performances obtenues lors de l'estimation de ces propriétés pour des forêts denses en climat tropical et tempéré [5-6], mais peu ont porté sur des forêts peu denses ou des forêts avec une large variabilité du taux de couverture boisée (cas des forêts en climat Méditerranéen par exemple). En effet pour ce dernier cas, une résolution spatiale qui n'est pas assez fine (de l'ordre décimétrique) peut faire apparaître des pixels mixtes dont la réflectance est due à la contribution de plusieurs éléments (i.e., arbre, sol, sous-bois). Une solution pour estimer les propriétés biophysico-chimiques de la végétation arborée est l'emploi de méthodes hybrides d'inversion qui combinent des modèles de transfert radiatif (RTM), des méthodes d'apprentissage (machine learning) et des données de télédétection. Le RTM simule l'image de télédétection

du couvert forestier à partir des propriétés structurelles et optiques de ses éléments (ex : hauteur et distribution des arbres, réflectance et transmittance foliaires, etc.). Via la variation de ces propriétés, le RTM peut créer une base de données de réflectances simulées qui permet d'entraîner et de valider une méthode d'apprentissage (type PLSR, RFR, SVMR, ANN). Cette méthode est ensuite appliquée sur les images spectrales pour produire des cartes de traits de végétation qui sont comparés avec des mesures terrain à des fins de validation.

Un challenge majeur actuel est de connaître le niveau de complexité nécessaire pour modéliser une scène arborée selon le taux de couverture boisée (paramétrisation du RTM), la méthode d'apprentissage la plus appropriée pour l'ensemble des traits de végétation étudiés, et les adaptations spectrales et spatiales à mener entre données spatiales multi-spectrales Sentinel-2 à 10/20m et missions hyperspectrales à 10m (BIODIVERSITY) et 30m (PRISMA, EnMAP, CHIME, SBG).

Le travail proposé vise à répondre à ces problématiques en utilisant le code RTM à l'échelle canopée 3D – DART [7], incluant le code RTM à l'échelle feuille PROSPECT [8] (lien entre les propriétés optiques de la feuille et les traits foliaires). Il s'inscrit dans la deuxième année du projet APR TOSCA CNES SentHyMED [9] visant à caractériser l'état de santé et le stress hydrique de deux forêts méditerranéennes situées au Nord de Montpellier (Puéchabon et Pic Saint Loup) avec deux espèces majoritaires de chênes (décidu : chêne blanc et sempervirent : chêne vert). Le travail considèrera une date donnée correspondant à la campagne de mesures MEDOAK (juin 2021), réalisée en parallèle d'acquisitions aéroportées hyperspectrales NASA-ESA pour les missions CHIME et SBG.

Dans un premier temps, les travaux porteront sur les données hyperspectrales:

- 1) Simulation des images satellitaires à 10m (BIODIVERSITY) et 30m (CHIME, SBG) à partir de celles aéroportées AVIRIS-Next Generation à 1 et/ou 3 m grâce aux outils existants à l'ONERA permettant de considérer le bruit instrumental,
- 2) Création des maquettes 3D DART des parcelles d'étude sur les 2 sites, à partir de données LiDAR drone et d'inventaires forestiers. Les nuages de points LiDAR seront convertis en matrice 3D de voxels "turbides" avec l'outil Amapvox. Puis, ces matrices 3D seront utilisées pour créer les maquettes DART,
- 3) Génération de bases spectrales simulées avec DART à partir des maquettes créées et d'un plan d'expériences pour la variation des propriétés de scène,
- 4) Entraînement de méthodes d'apprentissage et leur application sur l'ensemble des images (aéroportées et simulées satellitaires) pour l'estimation et la cartographie des propriétés biophysico-chimiques. Des méthodes sont déjà disponibles via des travaux sur des sites forestiers méditerranéens en Californie (thèse de Thomas Miraglio [10-12] et APR CNES HyperMED [13]) et des sites forestiers tropicaux en Guyane (thèse Dav Ebengo [14] et APR CNES HyperTropik [15]),
- 5) Comparaison avec des données terrain pour l'évaluation des performances.

Dans un deuxième temps, les travaux porteront sur les données Sentinel-2 :

- 1) Sélection d'images sur la plateforme THEIA acquises au plus proche de la campagne MEDOAK,
- 2) Co-registation avec les images aéroportées hyperspectrales à partir de l'outil Gefolki,
- 3) Ré-échantillonnage spectral et spatial des bases spectrales simulées par DART pour correspondre aux caractéristiques des données Sentinel-2,
- 4) Inversion avec les méthodes ci-dessus mais adaptées aux indices spectraux du fait de la réduction du nombre de bandes spectrales,
- 5) Comparaison des résultats avec les données terrain.

Les produits attendus sont :

- Cartes d'estimation du LAI, CAB, CAR, EWT et LMA

Les résultats attendus sont :

- Bilan de l'impact de la résolution spatiale et des caractéristiques instrumentales sur les performances d'estimation des propriétés biophysico-chimiques,
- Bilan de la meilleure méthode de machine learning pour l'inversion,
- Bilan des performances d'estimation des propriétés biophysico-chimiques selon le taux de couverture boisée.

Ce CDD de 12 mois sera salarié CESBIO et accueilli sur le centre ONERA à Toulouse. Il pourra être suivi éventuellement d'un CDD de 12 mois porté par le laboratoire DYNAFOR à Toulouse pour prendre en compte les aspects multi-temporels dans la méthode développée et en ajoutant des indicateurs phénologiques pour évaluer le stress hydrique de ces forêts.

**Si vous êtes intéressés par le sujet, veuillez envoyer votre CV et lettre de motivation à J.P. Gastellu-Etchegorry (jean-philippe.gastellu@iut-tlse3.fr) et K. Adeline (karine.adeline@onera.fr)**

#### Références bibliographiques:

- [1] O. E. Sala *et al.*, "Global biodiversity scenarios for the year 2100," *Science*. 2000, doi: 10.1126/science.287.5459.1770.
- [2] E. C. Underwood, J. H. Viers, K. R. Klausmeyer, R. L. Cox, and M. R. Shaw, "Threats and biodiversity in the mediterranean biome," *Divers. Distrib.*, 2009, doi: 10.1111/j.1472-4642.2008.00518.x.
- [3] H. M. Pereira *et al.*, "Essential biodiversity variables," *Science*. 2013, doi: 10.1126/science.1229931.
- [4] J. Transon, R. d'Andrimont, A. Maignard, P. Defourny. Survey of hyperspectral Earth Observation applications from space in the Sentinel-2 context. *Remote Sens.* Vol. 10:2, 1-32, 2018, doi: 10.3390/rs10020157.
- [5] M. P. Ferreira, J. B. Féret, E. Grau, J. P. Gastellu-Etchegorry, Y. E. Shimabukuro, and C. R. de Souza Filho, "Retrieving structural and chemical properties of individual tree crowns in a highly diverse tropical forest with 3D radiative transfer modeling and imaging spectroscopy," *Remote Sens. Environ.*, 2018, doi: 10.1016/j.rse.2018.04.023.
- [6] F. D. Schneider *et al.*, "Mapping functional diversity from remotely sensed morphological and physiological forest traits," *Nat. Commun.*, vol. 8, no. 1, 2017, doi: 10.1038/s41467-017-01530-3.
- [7] <https://dart.omp.eu>
- [8] J.-B. Féret, K. Berger, F. De Boissieu, and Z. Malenovsky, "PROSPECT-PRO for estimating content of nitrogen-containing leaf proteins and other carbon-based constituents," ArXiv200311961 Q-Bio, 2020.
- [9] ONERA, CESBIO, TETIS, DYNAFOR and CEFE, « APR TOSCA CNES SentHyMED « Complémentarité entre de l'imagerie multi-temporelle Sentinel-2 et des imageurs Hyperspectraux pour un meilleur suivi des traits fonctionnels de forêts MEDiterranéennes », 2021-2023
- [10] T. Miraglio, K. Adeline, M. Huesca, S. Ustin, and X. Briottet, "Monitoring LAI, Chlorophylls, and Carotenoids Content of a Woodland Savanna Using Hyperspectral Imagery and 3D Radiative Transfer Modeling," *Remote Sens.*, vol. 12, no. 1, p. 28, Dec. 2019, doi: 10.3390/rs12010028.
- [11] T. Miraglio *et al.*, "Impact of modeling abstractions when estimating leaf mass per area and equivalent water thickness over sparse forests using a hybrid method," *Remote Sens.*, vol. 13, no. 16, 2021, doi: 10.3390/rs13163235.
- [12] T. Miraglio *et al.*, "Assessing Vegetation Traits Estimates Accuracies from the Future SBG and Biodiversity Hyperspectral Missions Over two Mediterranean Forests". *International Journal of Remote Sensing*, 43:10, 3537-3562, doi: 10.1080/01431161.2022.2093143.
- [13] ONERA, CESBIO, CSTARS, BiometLab. APR TOSCA CNES HyperMED: Evaluation des caractéristiques fonctionnelles des essences d'arbres pour le suivi de leur état de santé pour des écosystèmes de forêts MEDiterranéennes pour un imageur Hyperspectral. Préparation de la mission hyperspectrale HYPs. 2019-2022.
- [14] Dav Ebengo Mwampongo, "Apport de la modélisation physique pour la cartographie de la biodiversité végétale en forêts tropicales par télédétection optique," Université de Montpellier.
- [15] TETIS; CESBIO; AMAP; ECO&SOLS; Orsay, E. APR TOSCA CNES HyperTropik "Estimation de la biodiversité des forêts tropicales par imagerie hyperspectrale", préparation de la mission hyperspectrale HYPXIM 2014

#### **Encadrement**

Jean-Philippe Gastellu-Etchegorry (CESBIO, Toulouse), 05 61 55 61 30

Karine Adeline (ONERA – DOTA, Toulouse), 05 62 25 26 68

#### **Collaborations extérieures**

Laboratoires partenaires du projet SentHyMED (TETIS, DYNAFOR, CEFE)



## SHORT-TERM CONTRACT PROPOSAL

**Title :** Estimation of Mediterranean forest functional traits : a multi-sensor assessment in function of the canopy cover

**Start :** as soon as possible

**Deadline :** when the position is filled

**Duration :** 12 months (possibility to renew for 12 months more in the framework of the project)

### Keywords

Spectroscopy, Sentinel-2, LiDAR data, 3D modelling, radiative transfer, inversion, Mediterranean forests méditerranéennes

### Formation and skills

Engineering schools and/or research masters: signal/image or remote sensing data processing, machine learning methods, ecology/environment, python programming

Notions in radiative transfer modeling, 3D LiDAR information processing

A first experience on the use of the DART code (but not a prerequisite)

### Context, objective and work description

Mediterranean forests cover a relatively small area on a global scale for a high concentration of plant diversity, including endemic species that have been able to adapt to the hazards of the Mediterranean climate (long, hot and dry summers, mild and wet winters). However, they represent critical areas, because by 2100 their biodiversity will change much more than that of other terrestrial ecosystems [1], in particular due to the increase in fires, droughts, and anthropogenic exploitation (urbanization, wood industry, increase in agricultural lands) [2]. It is therefore important to monitor the temporal evolution of these forests and the health of some of their tree species, in order to guide environmental policies and preserve the biodiversity of these fragile environments. Remote sensing observations can be used to address this issue, as well as for large-scale mapping of essential biodiversity variables (EBVs) defined by groups of scientific experts (GEO-BON)[3] and, in particular, the study of the "EBV-species traits" class, which aims to characterize the functional traits of species through their phenological, morphological and physiological attributes. The monitoring of these traits is partly achieved through the estimation of certain biophysical-chemical properties at the leaf and canopy scale accessible by remote sensing. Among these properties, the most commonly studied are leaf pigments (CAB - chlorophylls and CAR - carotenoids), leaf water and dry matter content (EWT, LMA), and leaf area index (LAI). Over a phenological cycle of the plant, the latter are indicators that allow the evaluation of water stress, the observation of biomass changes and early defoliation.

With the multiplication of remote sensing data from current and future satellites, a major challenge is to assess the potential of combining all these datasets to better describe and study vegetation, given the spectral, spatial and temporal characteristics of the different sensors [4]. For our case study, the simultaneous estimation of the above-mentioned biophysical-chemical properties requires a high spectral richness to take advantage of their sensitivity over the whole measured spectrum. With its larger number of spectral bands, hyperspectral imaging over [0.4-2.5  $\mu\text{m}$ ] is more suitable than multispectral imaging, which however often has a shorter revisit time. Many studies have already evaluated the performance of these properties for dense forests in tropical and temperate climates [5-6], but few have focused on sparse forests or forests with a large variability in tree cover (e.g. Mediterranean climate forests). Indeed, for the latter case, a spatial resolution that is not fine enough (decametric order) may reveal mixed pixels whose reflectance is due to the contribution of several elements (i.e. tree, soil, understory). One solution to estimate the biophysical-chemical properties of tree vegetation is the use of hybrid inversion methods that combine radiative transfer models (RTM), machine learning methods and remote sensing data. The RTM simulates the remotely sensed images of the forest canopy based on the structural and optical properties of its elements (e.g. tree height and distribution, leaf reflectance and transmittance, etc.). By varying these properties, the RTM can create a database of simulated reflectances that can be used to train and validate a learning method (PLSR, RFR, SVMR, ANN). This method is then applied to spectral images to produce vegetation trait maps that are compared with field measurements for validation purposes.

A current major challenge is to determine the level of complexity required to model a tree scene according to the forest canopy cover (RTM parameterisation), the most appropriate learning method for all the vegetation traits studied, and the spectral and spatial adaptations to be carried out between Sentinel-2 multi-spectral data at 10/20m and hyperspectral missions at 10m (BIODIVERSITY) and 30m (PRISMA, EnMAP, CHIME, SBG)

The proposed work aims to address these issues using the 3D canopy-scale RTM code DART [7], including the leaf-scale RTM code PROSPECT [8] (link between the optical properties of the leaf and the leaf traits). It is part of the second year of the TOSCA CNES SentHyMED APR project [9], which aims to characterize the health status and water stress of two Mediterranean forests located to the north of Montpellier (Puéchabon and Pic Saint Loup) with two main oak species (deciduous: white oak and evergreen: green oak). The work will consider a given date corresponding to the MEDOAK measurement campaign (June 2021), carried out in parallel with NASA-ESA airborne hyperspectral acquisitions for the CHIME and SBG missions.

Initially, the work will focus on hyperspectral data:

- 1) Simulation of satellite images at 10m (BIODIVERSITY) and 30m (CHIME, SBG) from AVIRIS-Next Generation airborne images at 1 and/or 3m using existing tools at ONERA that allow to consider instrumental noise,
- 2) Creation of 3D DART mock-ups of the study plots on the two sites, using LiDAR drone data and forest inventories. The LiDAR point clouds will be converted into a 3D matrix of "turbid" voxels with the Amapvox tool. These 3D matrices will then be used to create the DART mock-ups,
- 3) Generation of simulated spectral databases with DART from the created mock-ups and building of a design of experiments for the variation of scene properties,
- 4) Training of machine learning methods and their application on the whole set of images (airborne and satellite simulated) for the estimation and mapping of biophysical-chemical properties. Methods are already available through work on Mediterranean forest sites in California (Thomas Miraglio's thesis [10-12] and APR CNES HyperMED [13]) and tropical forest sites in French Guyana (Dav Ebengo's thesis [14] and APR CNES HyperTropik [15]),
- 5) Comparison with field data for performance evaluation.

In a second phase, the work will focus on Sentinel-2 data:

- 1) Selection of images on the THEIA platform acquired as close as possible to the MEDOAK campaign,
- 2) Co-registration with airborne hyperspectral images using the Gefolki tool,
- 3) Spectral and spatial resampling of the DART-simulated spectral databases to match the Sentinel-2 data characteristics,
- 4) Inversion with the above methods but adapted to the spectral indices due to the reduction of the number of spectral bands,
- 5) Comparison of the results with the field data.

Expected products :

- Estimation maps of LAI, CAB, CAR, EWT and LMA

Expected results :

- Review of the impact of spatial resolution and instrumental characteristics on the performance of biophysical-chemical property estimation,
- Review of the best machine learning method for inversion,
- Assessment of the performance of biophysical-chemical properties estimation according to the forest canopy cover.

This 12-month short-term contract will be on the CESBIO payroll and hosted at the ONERA centre in Toulouse. It may be followed by a 12-month another short-term contract at the DYNAFOR laboratory in Toulouse to take into account multi-temporal aspects in the method developed and by adding phenological indicators to assess the water stress of these forests.

**If you are interested, please send your CV and motivation letter to J.P. Gastellu-Etchegorry (jean-philippe.gastellu@iut-tlse3.fr) and K. Adeline (karine.adeline@onera.fr)**

References:

- [1] O. E. Sala *et al.*, "Global biodiversity scenarios for the year 2100," *Science*. 2000, doi: 10.1126/science.287.5459.1770.
- [2] E. C. Underwood, J. H. Viers, K. R. Klausmeyer, R. L. Cox, and M. R. Shaw, "Threats and biodiversity in the mediterranean biome," *Divers. Distrib.*, 2009, doi: 10.1111/j.1472-4642.2008.00518.x.

- [3] H. M. Pereira *et al.*, "Essential biodiversity variables," *Science*. 2013, doi: 10.1126/science.1229931.
- [4] J. Transon, R. d'Andrimont, A. Maignard, P. Defourny. Survey of hyperspectral Earth Observation applications from space in the Sentinel-2 context. *Remote Sens*. Vol. 10:2, 1-32, 2018, doi: 10.3390/rs10020157.
- [5] M. P. Ferreira, J. B. Féret, E. Grau, J. P. Gastellu-Etchegorry, Y. E. Shimabukuro, and C. R. de Souza Filho, "Retrieving structural and chemical properties of individual tree crowns in a highly diverse tropical forest with 3D radiative transfer modeling and imaging spectroscopy," *Remote Sens. Environ.*, 2018, doi: 10.1016/j.rse.2018.04.023.
- [6] F. D. Schneider *et al.*, "Mapping functional diversity from remotely sensed morphological and physiological forest traits," *Nat. Commun.*, vol. 8, no. 1, 2017, doi: 10.1038/s41467-017-01530-3.
- [7] <https://dart.omp.eu>
- [8] J.-B. Féret, K. Berger, F. De Boissieu, and Z. Malenovsky, "PROSPECT-PRO for estimating content of nitrogen-containing leaf proteins and other carbon-based constituents," ArXiv200311961 Q-Bio, 2020.
- [9] ONERA, CESBIO, TETIS, DYNAFOR and CEFE, « APR TOSCA CNES SentHyMED « Complémentarité entre de l'imagerie multi-temporelle Sentinel-2 et des imageurs Hyperspectraux pour un meilleur suivi des traits fonctionnels de forêts MEDiterranéennes », 2021-2023
- [10] T. Miraglio, K. Adeline, M. Huesca, S. Ustin, and X. Briottet, "Monitoring LAI, Chlorophylls, and Carotenoids Content of a Woodland Savanna Using Hyperspectral Imagery and 3D Radiative Transfer Modeling," *Remote Sens.*, vol. 12, no. 1, p. 28, Dec. 2019, doi: 10.3390/rs12010028.
- [11] T. Miraglio *et al.*, "Impact of modeling abstractions when estimating leaf mass per area and equivalent water thickness over sparse forests using a hybrid method," *Remote Sens.*, vol. 13, no. 16, 2021, doi: 10.3390/rs13163235.
- [12] T. Miraglio *et al.*, "Assessing Vegetation Traits Estimates Accuracies from the Future SBG and Biodiversity Hyperspectral Missions Over two Mediterranean Forests". *International Journal of Remote Sensing*, 43:10, 3537-3562, doi: 10.1080/01431161.2022.2093143.
- [13] ONERA, CESBIO, CSTARS, BiometLab. APR TOSCA CNES HyperMED: Evaluation des caractéristiques fonctionnelles des essences d'arbres pour le suivi de leur état de santé pour des écosystèmes de forêts MEDiterranéennes pour un imageur Hyperspectral. Préparation de la mission hyperspectrale HYPH. 2019-2022.
- [14] Dav Ebengo Mwampongo, "Apport de la modélisation physique pour la cartographie de la biodiversité végétale en forêts tropicales par télédétection optique," Université de Montpellier.
- [15] TETIS; CESBIO; AMAP; ECO&SOLS; Orsay, E. APR TOSCA CNES HyperTropik "Estimation de la biodiversité des forêts tropicales par imagerie hyperspectrale", préparation de la mission hyperspectrale HYPXIM 2014

### Supervision

Jean-Philippe Gastellu-Etchegorry (CESBIO, Toulouse), 05 61 55 61 30

Karine Adeline (ONERA – DOTA, Toulouse), 05 62 25 26 68

### Collaborations

Laboratories partners of the project SentHyMED (TETIS, DYNAFOR, CEFE)