



Université de Toulouse

MASTER 2 GEOMATIQUE

« Sciences Géomatiques en environneMent et AménageMent » (SIGMA)

<http://sigma.univ-toulouse.fr>

MEMOIRE DE RECHERCHE

**Suivi d'un couvert ligneux tropical par
cartographie d'indicateurs biophysiques et
écologiques à partir de données satellites Sentinel-2**

DEFOSSEZ Alexandre

**Centre de coopération Internationale en Recherche Agronomique
pour le Développement (CIRAD)**

**Maître de stage : Jean-Baptiste Féret
Enseignant-référent : David Sheeren**



Septembre 2021

Table des Matières

Abstract.....	1
Introduction.....	2
Présentation de la structure d'accueil.....	5
Etat de l'Art.....	6
I. Caractérisation du couvert ligneux par les Indices de Végétation.....	6
II. Caractérisation biophysique du couvert ligneux.....	7
III. Cartographier la biodiversité du couvert ligneux.....	16
IV. Problématique et hypothèses de recherche.....	20
Matériel & Méthodes.....	21
I. Diagramme de Gantt.....	21
II. Présentation de la zone d'étude.....	22
III. Sélection et acquisition des images.....	22
IV. Production des variables biophysiques.....	24
V. Production des indices de biodiversité α et β	27
VI. Photo-interprétation des résultats.....	30
VII. Exploration des liens entre les quatre variables biophysiques.....	30
VIII. Confrontation des résultats avec les données de structure.....	31
Résultats.....	32
I. Cartographie des variables biophysiques.....	32
II. Cartographie des indices de biodiversité.....	36
III. Motifs de variations biophysique et écologique de la végétation.....	38
IV. Relations entre les variables biophysiques.....	42
V. Comparaison avec les données de structure de la canopée.....	43
Discussion.....	47
I. Evaluation de la cohérence des motifs et limites de la photo-interprétation.....	47
II. Limites de l'inversion de PROSAIL et identification des artéfacts.....	48
III. Interprétation des corrélations entre les quatre variables biophysiques.....	50
IV. Evaluation de la cohérence des variables par confrontation avec les données de texture.....	50
Conclusion & Perspectives.....	52
Annexes.....	53
Bibliographie.....	59

Table des Figures

Figure 1. Cartes du couvert ligneux en Amérique du Sud et Afrique centrale	3
Figure 2. Comparaison de deux cartes d'estimation de l'AGB en Amérique du Sud et leurs différences.....	8
Figure 3. MNC d'une forêt tropicale humide.....	10
Figure 4. Etapes d'estimation de l'AGB à l'aide des analyses de texture.....	11
Figure 5. Schéma récapitulatif du principe de fonctionnement de PROSAIL	14
Figure 6. Principe de détermination des trois composantes de la diversité fonctionnelle.....	17
Figure 7. Résumé des quatre approches principales pour cartographier la diversité de la végétation	18
Figure 8. Diagramme de Gantt détaillant la répartition temporelle des tâches prévues au cours du stage.....	21
Figure 9. Zone d'étude au Nord-Ouest de la Guyane Française	22
Figure 10. Diagramme de flux des étapes de la chaîne de traitement de production des variables biophysiques.....	25
Figure 11. Diagramme de flux des étapes de la chaîne de traitement de production des indices de diversité α et β	27
Figure 12. Exemple de Continuum Removal.....	28
Figure 13. Présentation des données complémentaires de structure dans la zone d'étude.	31
Figure 14. Carte des réflectances de la zone d'étude.....	33
Figure 15. Cartographie des quatre variables biophysiques couvrant la zone d'étude.....	34
Figure 16. Distributions des valeurs de pixels des quatre variables biophysiques.....	35
Figure 17. Composantes principales issues de la SPCA	36
Figure 18. Cartographie des indices de biodiversité couvrant la zone d'étude	37
Figure 19. Distributions des pixels de diversité α	37
Figure 20. Exemple d'utilisation du MNT pour appuyer la photo-interprétation	38
Figure 21. Exemple de motifs des quatre variables biophysiques associés aux variations du relief.....	39
Figure 22. Exemple de motifs de diversité α et β associés aux variations du relief.....	40
Figure 23. Exemple de motifs des quatre variables biophysiques associés à une zone humide.....	41
Figure 24. Exemple de motifs de diversité α et β associés à une zone humide.....	42
Figure 25. Tendances observées sur les corrélations entre les moyennes à l'hectare des variables biophysiques.....	43
Figure 26. Distribution des écart-types en fonction des moyennes du MNC à l'hectare	44
Figure 27. Distribution des moyennes et des écart-types à l'hectare des variables biophysiques sur la zone d'étude.....	44
Figure 28. Cercle et coefficients des corrélations de l'ACP	45
Figure 29. Comparaison des cartes de diversité β et des indices de texture dérivés de la méthode FOTO.....	46
Figure 30. Corrélations entre les axes de la diversité β et des indices de texture de la méthode FOTO	46
Figure 31. Zone humide suspectée en composition colorée « Vraies couleurs » et en diversité β	48
Tableau 1. Les seize variables (ou paramètres) intégrées dans le modèle PROSAIL.....	15
Tableau 2 Dates sélectionnées d'après le critère d'enneuagement.....	23
Tableau 3. Bandes spectrales mesurées par Sentinel-2A et Sentinel-2B	24
Tableau 4. Listes des principaux paramètres d'entrée utilisés pour simuler la réflectance avec PROSAIL.....	26
Tableau 5. Tableau récapitulatif des tendances observées pour les différents motifs.....	47
Équation 1. Formule de calcul du NDVI	6
Équation 2. Formules de calcul de l'EWT, LMA et LAI.....	25
Équation 3. Formule de calcul de l'indice de Shannon utilisée pour déterminer la diversité α	29
Équation 4. Formule de calcul de la distance de Bray-Curtis utilisée pour déterminer la diversité β	30

Abstract

Cartographier les forêts tropicales par télédétection est d'importance mondiale, afin d'assurer leur suivi et lutter contre la déforestation. Les satellites Sentinel-2, par l'acquisition de données multispectrales à haute résolution et une période de revisite de cinq jours, ouvrent de nouvelles perspectives pour la cartographie régionale des forêts tropicales. Ce stage a pour objectif d'explorer le potentiel de variables biophysiques et écologiques produites à partir de données Sentinel-2 pour aider à l'analyse écologique des milieux de forestiers tropicaux. Dans ce but, nous avons estimé quatre variables biophysiques (CHL, EWT, LMA, LAI) par inversion hybride du modèle de transfert radiatif PROSAIL, ainsi que les indices de diversité α et β , sur un milieu de forêt tropicale humide situé en Guyane Française, à l'aide des packages R *prosail* et *biodivMapR*. Notre travail a consisté à évaluer la cohérence des produits cartographiques obtenus, en nous appuyant sur un Modèle Numérique de Canopée dérivé de données LiDAR, et un produit issu de l'analyse de texture de données d'imagerie très haute résolution spatiale obtenu avec la méthode FOTO. Nous avons montré l'existence de motifs cohérents et de corrélations entre les différentes variables, suggérant la possibilité de repousser les limites des sources de données LiDAR et THRS habituellement utilisées pour l'analyse des forêts tropicales. Pour confirmer nos résultats, la validation par comparaison avec des données terrain est envisagée dans une prochaine phase du projet. Nos résultats constituent une première étape vers une cartographie fiable et reproductible du couvert ligneux tropical, sur la base de variables biophysiques et écologiques.

Mapping tropical forests by remote sensing is of global importance to ensure their monitoring and fight deforestation. The Sentinel-2 satellites, providing high resolution multispectral data with a five-day revisit period, open new perspectives for regional mapping of tropical forests. The objective of this internship is to explore the potential of biophysical and ecological variables produced from Sentinel-2 data to support the ecological analysis of tropical forest environments. For this purpose, we estimated four biophysical variables (CHL, EWT, LMA, LAI) by hybrid inversion of the PROSAIL radiative transfer model, as well as the α and β diversity indices, on a tropical rainforest environment located in French Guiana, using the *prosail* and *biodivMapR* R packages. Our work mainly consisted in assessing the consistency of the resulting maps, using a Canopy Height Model derived from LiDAR data, and a product derived from textural analysis of very high spatial resolution imagery data obtained with the FOTO method. We found consistent patterns and correlations between the different variables, suggesting the possibility of pushing the limits of LiDAR and THRS data sources usually used for tropical forest analysis. To confirm our results, validation by comparison with field data is envisaged in a next stage of the project. Our results are a first step towards a reliable and reproducible mapping of tropical woody cover, based on biophysical and ecological variables.

Mots-clés : couvert ligneux, variables biophysiques, forêts tropicales, canopée, PROSAIL, diversité α et β , Sentinel-2, photo-interprétation, *biodivMapR*, *prosail*.

Introduction

Les forêts tropicales constituent l'un des plus importants réservoirs de biodiversité à l'échelle globale (Gentry, 1992). Parmi les vingt-cinq *hotspots* de biodiversité définis par Myers *et al.*, (2000), quinze sont des aires occupées par des forêts tropicales (Mittermeier *et al.*, 2011). On estime qu'elles abritent au moins les deux tiers de la biodiversité mondiale (Gardner *et al.*, 2009) et représentent environ 1.8 milliards d'hectares, soit 45 % de la surface forestière globale (*'Global Forest Resources Assessment 2020'*, FAO). Les forêts tropicales sont à l'origine de multiples services écosystémiques, notamment le stockage du carbone atmosphérique (Pan *et al.*, 2011; Phillips and Lewis, 2014; Mackey *et al.*, 2020). Entre 1990 et 2007, les forêts tropicales « intactes » (non directement affectées par les activités humaines) sont responsables d'environ la moitié du carbone total absorbé par l'ensemble des forêts sur la même période (Pan *et al.*, 2011). Leur rôle dans la limitation du changement climatique s'avère primordial (Townsend *et al.*, 2011; Malhi, 2012; Mitchard, 2018). Les écosystèmes forestiers tropicaux représentent également une source de molécules au potentiel thérapeutique encore largement inexploré (Balunas and Kinghorn, 2005) et abritent d'importantes populations humaines, en particulier indigènes (Bennett, 1992; Pert *et al.*, 2015). Malgré l'importance des forêts tropicales pour l'humanité, une déforestation massive se poursuit (entre 1990 et 2020, 7 millions d'hectares sont perdus en moyenne par an), ainsi qu'une « dégradation », plus discrète, liées aux activités humaines et aux événements climatiques (*e.g. El Niño*), dont l'ampleur commence à être perçue (Vancutsem *et al.*, 2021).

L'importance de protéger les forêts tropicales dans la lutte contre le changement climatique et l'érosion de la biodiversité est aujourd'hui communément admise. À titre d'exemple, le programme REDD+¹ cherche à évaluer les flux de carbone à l'échelle mondiale et limiter leur émission à travers la lutte contre la déforestation et le commerce illégal des bois tropicaux. Néanmoins, suivre l'évolution des forêts tropicales représente un défi à divers égards pour les écologues : diversité et étendue des milieux, difficulté d'accès, changements rapides, oppositions locales liées à des intérêts économiques... Le développement des techniques de télédétection a permis de s'affranchir d'une partie de ces contraintes. La télédétection se définit comme étant « *le processus de détection et de suivi des caractéristiques physiques d'une surface par la mesure des radiations réfléchies et émises à distance*² ». Le plus souvent, les mesures sont effectuées depuis un avion ou un satellite. La télédétection rassemble une diversité d'approches et d'outils qui ne cessent de se développer (imagerie optique multispectrale, hyperspectrale, LiDAR...). L'objectif étant de pouvoir caractériser divers phénomènes à la surface de la Terre qui évoluent au cours du temps et dans l'espace (*e.g.* urbanisation, éruptions volcaniques, blooms de phytoplancton). Dans le domaine de l'étude de la végétation, la télédétection a constitué une avancée décisive, bien que demeurent certaines limites, principalement liées à la disponibilité des données et à la précision des capteurs.

Les « couverts ligneux », comprenant les milieux forestiers dans leur diversité, plus ou moins denses, naturels, semi-naturels voire d'origine anthropique (plantations) ont été principalement cartographiés à des échelles globales, à partir d'images satellitaires à moyenne résolution spatiale (MODIS, Landsat). Bien que des cartographies mondiales du couvert ligneux soient indispensables à l'estimation des stocks de carbone et au suivi de la déforestation (Pan *et al.*, 2011), leur utilisation à des échelles locales se heurte aux incertitudes liées à la résolution modérée des données d'imagerie satellite généralement utilisées dans ces applications (Mitchard *et al.*, 2014 ; Réjou-Méchain *et al.*, 2019) (**Figure 1**). Pourtant, ce type de cartes est couramment utilisé pour répondre à des problématiques locales malgré leurs limites à des échelles trop fines, leurs utilisateurs (chercheurs ou gestionnaires de l'environnement) leur accordant souvent une confiance trop importante. D'autre part, l'approche commune en cartographie consistant à découper un territoire en classes d'occupation du sol (*Land cover*) s'accorde peu avec la diversité et la complexité des couverts ligneux, notamment tropicaux, qui demandent des approches permettant de décrire des phénomènes continus plutôt que des classes discrètes.

¹ <http://www.fao.org/redd/fr/>

² <https://www.usgs.gov>

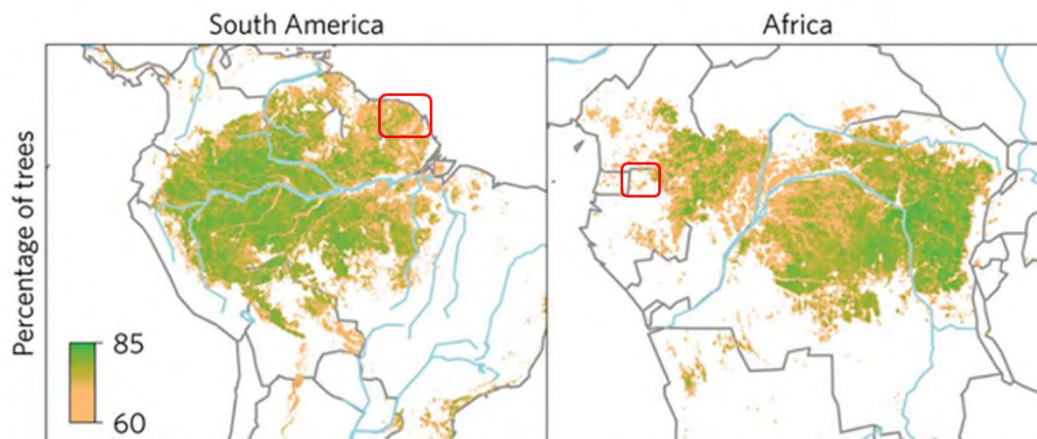


Figure 1. Cartes du couvert ligneux en Amérique du Sud et Afrique centrale, déterminées par les séries temporelles de NDVI MODIS. Figure extraite de Verbesselt *et al.*, (2016) (modifiée). Les encadrés rouges indiquent des zones où le pourcentage d'arbres est sous-estimé, limitant leur utilisation à des échelles régionales et dans le cadre d'applications locales.

L'utilisation de classes d'occupation recèle toujours d'une part d'arbitraire, car elle implique des choix parmi les critères possibles de définition des classes, parfois difficiles à justifier concernant les couverts ligneux. Leur interprétation peut devenir difficile, comme dans le cas des classes de « transitions » entre milieux (*e.g.* « savane arborée », « transition forêt-savane »). Or, les enjeux scientifiques et opérationnels liés au suivi des couverts ligneux nécessitent des indicateurs stables et validés (Mitchard *et al.*, 2014 ; Prestele *et al.*, 2017). La détermination des Indices de Végétations (VIs) et des variables biophysiques associées à la canopée constituent une avancée en ce sens (Huete *et al.*, 2012 ; Baret and Buis, 2008 ; Berger *et al.*, 2018). D'autres approches, notamment celles basées sur les données issues de capteurs LiDAR, se sont révélées également très prometteuses (Wang and Fang, 2020). Les données LiDAR sont notamment moins affectées par la saturation du signal, un problème fréquemment rencontré dans le cas des forêts tropicales (Couteron *et al.*, 2012). À l'avenir, ces approches devraient faire émerger de nouvelles cartographies du couvert ligneux, reflétant sa structure et son état physiologique, basées sur une caractérisation biophysique de la canopée, y compris à des échelles régionales.

Une autre dimension ne peut être ignorée si l'on souhaite caractériser le plus précisément possible la diversité des couverts ligneux : il est également nécessaire de prendre en compte des variables qui fassent explicitement référence à des niveaux de diversité biologique (Wang and Gamon, 2019). Cartographier la biodiversité est à la fois une nécessité scientifique, car ce type de suivi régulier permettrait de donner accès aux dynamiques spatiales et temporelles de multiples composantes de la biodiversité de la végétation (Shurr *et al.*, 2012 ; Rocchini *et al.*, 2019), et un besoin sociétal de plus en plus pressant, les changements dans la répartition géographique et la composition des communautés animales et végétales étant potentiellement plus rapides et dramatiques que les changements climatiques (Rockstrom *et al.*, 2009). La notion de biodiversité recouvre une réalité complexe, comprenant les multiples facettes de la diversité biologique : spécifique, fonctionnelle, taxonomique, génétique, phylogénétique... Les images issues de la télédétection peuvent donner accès à certaines dimensions de la biodiversité du couvert ligneux, notamment à travers la diversité spectrale (Palmer *et al.*, 2000, 2002). Les synergies entre télédétection et écologie ouvrent des perspectives non seulement de suivi global de la biodiversité, mais trouvent également des applications à des échelles locales et opérationnelles (*e.g.* Gómez-Rodríguez *et al.*, 2008 ; Mazor *et al.*, 2013).

Parmi les produits satellitaires disponibles, les données Sentinel-2 du programme européen Copernicus³ ont la particularité de posséder à la fois une haute résolution spatiale, jusqu'à 10 m pour certaines bandes spectrales, et une haute résolution temporelle, avec une nouvelle prise d'image tous les cinq jours. Ces données constituent une opportunité pour caractériser à des échelles régionales, voire globales, les dynamiques de la végétation potentiellement marquées par des changements rapides et des points de basculement vers de nouveaux états (Gonzalez, 2001; Ten Hove *et al.*, 2012). C'est notamment le cas des forêts tropicales, exposées à la déforestation, la dégradation et aux feux. Une autre particularité des images Sentinel-2 incite à leur utilisation : la présence de plusieurs bandes dans

³ <https://www.copernicus.eu/fr>

l'Infrarouge (IR) et le *Red Edge* (domaine spectral situé entre le domaine visible et le Proche Infrarouge), contribuant notamment à mieux caractériser les phénomènes d'absorption induits par la chlorophylle (*a* et *b*), et à mieux séparer la contribution de ce pigment de celle de propriétés de structure, comme l'Indice de surface foliaire (*Leaf Area index*, LAI). Dans le contexte de recherches visant à améliorer la précision des méthodes de cartographie du couvert ligneux, reflétant des caractéristiques biophysiques et écologiques, l'exploitation des images Sentinel-2 est encore loin du plein potentiel offert par ces données.

Le projet de recherche « sé2coul », financé par le Centre National d'Études Spatiales (CNES) dans le cadre d'un financement TOSCA (Terre, Océans, Surfaces Continentales, Atmosphère), impliquant des équipes des UMRs TETIS⁴ et AMAP⁵, a pour objectif de mettre en place les bases d'un système de suivi des forêts tropicales en se fondant sur l'estimation de variables biophysiques et écologiques associées à la canopée, en explorant plus particulièrement le potentiel des images Sentinel-2. Le stage faisant l'objet de ce mémoire s'inscrit dans ce projet et propose d'étudier les couverts ligneux tropicaux, à l'aide de méthodes de cartographie de variables biophysiques et écologiques associées à la canopée, issues de l'analyse de données Sentinel-2.

⁴ UMR TETIS, 500 Rue Jean François Breton, 34000 Montpellier, <https://www.umr-tetis.fr/index.php/fr/>

⁵ UMR AMAP, Parc Scientifique Agropolis, 2196 Bd de la Lironde, 34980 Montferrier-sur-Lez, <https://amap.cirad.fr/fr/index.php>

Présentation de la structure d'accueil

La Maison de la Télédétection en Languedoc-Roussillon⁶ (MTD), fondée en 1994 à Montpellier, est une structure de recherche dédiée à la télédétection et à l'analyse de différents types de données spatialisées. Elle est située à proximité du campus Montpellier-Agropolis, au 500 rue Jean-François Breton et est actuellement hébergée par l'INRAE.

La MTD héberge deux Unités Mixtes de Recherche : l'UMR TETIS et l'UMR « Espace-dev⁷ ». En novembre 2020, elle accueillait 108 agents dont 73 permanents (incluant 47 chercheurs), 23 doctorants, 12 CDD et post-docs, ainsi que de nombreux stagiaires. Les activités de la MTD se caractérisent par leur haut niveau d'interdisciplinarité, aussi bien du point de vue des compétences techniques que thématiques, mais aussi par leurs diverses finalités : recherche, formation, expertise et valorisation. En 2014, la MTD s'agrandit avec la construction d'un nouveau bâtiment, accroissant sa capacité d'accueil, ainsi que d'une antenne de réception satellitaire. En raison de la présence d'instituts comme le CIRAD ou l'IRD à la MTD, une partie importante des travaux sont tournés vers les pays du Sud, portant sur des problématiques locales nécessitant le recours aux outils de télédétection ou à divers traitements des données spatialisées.

La recherche à la MTD s'organise principalement autour de quatre équipes de recherche, affiliées à différents instituts (CIRAD, IRD, INRAE et AgroParisTech) : l'équipe ATTOS (« Acquisition et Traitement de données de Télédétection et d'Observations Spatialisées »), MISCA (« Modélisation de l'Information Spatiale extraction de Connaissance et Analyse »), USIG (« USages de l'Information spatiale et Gouvernance ») et Ingénierie (Métrologie, données, code, intelligence artificielle, géomatique). Elle est également dynamisée par des groupes de travail « *ad hoc* » constitués dans le but de favoriser l'interaction entre chercheurs appartenant à différentes équipes sur des thématiques interdisciplinaires. Par exemple, le groupe « MDL4EO⁸ » développe des méthodologies d'analyse des données spatialisées en s'appuyant sur l'apprentissage automatique (*machine learning*) et l'apprentissage profond (*deep learning*).

Les UMRs TETIS et Espace-dev sont également très impliquées dans les projets « Equipement d'Excellence GEOSUD⁹ » et « Pôle thématique des surfaces continentales Théia¹⁰ ». La MTD se distingue, non seulement comme un pôle de recherche important dans le domaine de la télédétection, mais aussi en tant que plateforme technologique. Enfin, la MTD accueille d'autres structures comme l'association OpenIG et la start-up TeleScop.

⁶ <http://www.teledetection.fr/index.php/fr/>

⁷ <https://www.espace-dev.fr/>

⁸ <https://mdl4eo.irstea.fr/>

⁹ <https://ids.equipex-geosud.fr/web/guest>

¹⁰ <https://www.theia-land.fr/>

Etat de l'Art

I. Caractérisation du couvert ligneux par les Indices de Végétation

a. Principes et applications

L'approche la plus classique de caractérisation du couvert ligneux par télédétection repose sur la détermination des Indices de Végétation (*Vegetation Indices* ou VI). Il en existe une grande diversité et de nouveaux sont régulièrement proposés, à mesure que les capteurs se développent (Xue and Su, 2017). L'hypothèse sous-jacente à la production des VIs (comme dans d'autres approches basées sur l'imagerie optique) est que la signature spectrale de la végétation renseigne sur son état, sa composition biogéochimique, ainsi que sur la structure de la feuille et de la canopée. Le principe de base des VIs consiste à mesurer le rapport entre deux bandes spectrales, dont l'une est absorbée par l'objet ciblé (*i.e.* la végétation) et l'autre non (Huete, 2012). C'est par exemple le cas du très classique NDVI :

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)}$$

Équation 1. Formule de calcul du NDVI. Avec NIR : Réflectance d'une bande spectrale dans le domaine du Proche Infrarouge (absorption faible) ; Red : Réflectance dans le domaine spectral du Rouge (absorption forte).

Le but principal des VIs consiste à estimer les propriétés biophysiques de la végétation et fournir un appui pour le suivi de la végétation, dans le cadre d'applications visant par exemple au suivi de surfaces agricoles, ou à la gestion des ressources forestières (Huete, 2012). Leur potentiel d'utilisation est fortement conditionné par la résolution spatiale permise par le capteur : des résolutions modérées (d'ordre kilométrique) permettront de mettre en évidence des tendances fonctionnelles ou physiologiques à des échelles régionales, tandis que des résolutions plus fines (d'ordre métrique/décamétrique) aideront à caractériser des dynamiques à l'échelle de communautés forestières voire d'individus (Huete, 2012). Les VIs ont des applications en particulier dans les domaines de l'étude du climat et de la gestion forestière, notamment en lien avec la problématique du stockage du carbone. Ils sont aussi régulièrement utilisés pour l'agriculture de précision (Tanriverdi, 2006; Candiago *et al.*, 2015).

Certains indices, tel que l'*Enhanced Vegetation Index* (EVI) sont produits de manière à réduire certains biais associés à la canopée ou à la variation de aérosols (Huete, 2012), en combinant un nombre plus important de bandes. Dans le cas des forêts tropicales, l'EVI permet de minimiser les biais que connaît le NDVI, *i.e.* saturation liée à la fermeture de la végétation, effets de l'« arrière-plan » de la canopée, présence de bruit malgré la procédure Maximum Value Composite (MVC) et impossibilité de sélectionner avec certitude les pixels les plus clairs et les angles de vue les plus petits malgré la MVC (Zhengxing, Chuang and Alfredo, 2003). Il permet également de révéler des variations saisonnières et annuelles (Brando *et al.*, 2010). Outre le travail d'optimisation, l'utilisation conjointe de multiples indices permet également de mieux caractériser la canopée en raison de leur potentielle complémentarité, le comportement spectral de la canopée pouvant varier selon les bandes. Par exemple, l'EVI est particulièrement sensible aux variations du Proche Infrarouge et vont résumer des informations provenant de diverses couches de la canopée, contrairement au NDVI, qui est sensible aux longueurs d'onde dans le Rouge et sera donc sensible uniquement à la couche supérieure de la canopée (Huete, 2012).

Les VIs peuvent également être utilisés dans le but d'approcher des variables biophysiques associées à la végétation, comme le *Leaf Area Index* ou LAI (Cohen *et al.*, 2003) ou le *fractions of Photosynthetically Active Radiation absorbed by vegetation* ou fAPAR (Zhang *et al.*, 2005). En dépit du manque de base théorique pour décrire la relation entre VIs et variables biophysiques, ils demeurent informatifs sur l'état fonctionnel de la forêt. En particulier concernant la Productivité Primaire (PP) et certaines variables phénologiques comme le pic de croissance saisonnière (Huete, 2012).

b. Biais et limites

Les bandes spectrales utilisées pour déterminer les VIs sont généralement trop larges pour identifier de potentielles variations fines de la réflectance, associée à des changements d'ordre biochimiques ou liés à l'état de stress de la végétation. Néanmoins, le développement d'indices dérivés de l'imagerie hyperspectrale apporte des perspectives pour dépasser cette limite (Carter and Knapp, 2001; Huete, 2012).

D'autre part, les modèles de prédiction des variables biophysiques basées sur les VIs présentent des relations généralement assez faibles (Cohen *et al.*, 2003) ou liées de manière partielle à la composante biophysique considérée, *e.g.* le NDVI approxime mieux le fAPAR absorbé par la chlorophylle, que le fAPAR absorbé par la canopée dans son ensemble, une partie du fAPAR de la canopée étant influencée par des éléments non-photosynthétiques, comme, au niveau des feuilles, les autres pigments, la paroi cellulaire, etc. (Zhang *et al.*, 2005). Ces relations sont souvent contexte-dépendante (Huete, 2012) et varient selon le type de végétation (*e.g.* décidus vs résineux) ou la phénologie (Huemmrich *et al.*, 2005).

Enfin, les VIs sont impactés par le problème bien documenté de la saturation du signal (Lu, 2006; Lu *et al.*, 2012; Zhao, Zhong and Zhang, 2016), associé à la fermeture et à l'épaisseur importante de la canopée, principalement dans les milieux tropicaux. Cette difficulté restreint l'utilisation des VIs sur les forêts tropicales denses, et biaise les estimations des variables biophysique comme la biomasse du couvert ligneux (Gibbs *et al.*, 2007).

II. Caractérisation biophysique du couvert ligneux

D'autres approches de télédétection permettent d'estimer certaines variables biophysiques à partir de la réflectance de la canopée. Ces propriétés renseignent sur la structure et le fonctionnement écologique ou écophysologique du couvert ligneux. Les approches qui permettent d'accéder à ces propriétés ont été testées dans une diversité de biomes, notamment dans les forêts tropicales et les savanes, où les enjeux en termes de conservation et de gestion forestière peuvent être très importants. Dans la littérature, nous avons distingué trois principaux types d'approches dans la caractérisation biophysique du couvert ligneux : 1- les méthodes d'estimation de la biomasse forestière épigée (désignée par le terme assez générique d'*Aboveground biomass* ou AGB) visant à établir des cartographies à des échelles régionales voire globales, 2- la caractérisation de la structure du couvert ligneux (AGB, LAI) à l'aide de données LiDAR ou d'analyses texturales, et 3- l'extraction de variables biophysiques (ou biochimiques) foliaires à travers l'inversion de modèles de canopée. Pour chacune de ces grands types d'approches, nous tenterons de donner un aperçu des méthodes principales ainsi que leurs limites, avec une attention particulière concernant l'estimation des variables foliaires de la canopée.

a. Estimation de la biomasse forestière épigée par imageries optique et radar

Les méthodes d'estimation de l'AGB par télédétection se sont révélées plus fiables et économiquement plus avantageuses que les méthodes conventionnelles basées sur le déploiement de mesures sur le terrain de grande ampleur, pouvant comprendre des inventaires forestiers et la détermination de relations allométriques entre certaines variables de structure comme la hauteur des arbres et la biomasse (Lu, 2005, 2006). Elles permettent de cartographier les forêts à des échelles régionales et globales, notamment les forêts tropicales, qui possèdent la biomasse forestière la plus importante à l'échelle mondiale (Pan *et al.*, 2013). L'estimation de l'AGB par télédétection est essentielle pour évaluer les stocks de carbone et mettre en place des plans de gestion forestières qui limitent leur impact sur le climat. Pour estimer l'AGB, il existe une diversité de méthodes s'appuyant sur les images de télédétection. À des niveaux différents, tous les types d'images ont été utilisés : multispectrales, hyperspectrales, radar et LiDAR, de manière indépendante ou combinées dans des approches multi-capteurs (Dube *et al.*, 2016).

Les images multispectrales à moyenne résolution spatiale (MODIS, Landsat) ont été largement utilisées pour estimer l'AGB et cartographier les forêts. Les méthodes d'estimation s'appuient sur des modèles prédisant la relation entre certaines caractéristiques spectrales de l'image et l'AGB, mesuré sur

le terrain par des méthodes conventionnelles (équations allométriques) ou estimé à l'aide de données LiDAR. Jusqu'à présent, l'estimation de la biomasse forestière à l'aide d'images avec une résolution plus fine, comme Sentinel-2, a été assez peu explorée (Pandit, Tsuyuki and Dube, 2018; Navarro *et al.*, 2019; Wang *et al.*, 2019). Le plus souvent, Sentinel-2 (optique) et Sentinel-1 (radar) ont été combinés dans des approches multi-capteurs pour estimer l'AGB. Ces approches basées sur les images multispectrales possèdent des faiblesses, en raison d'incertitudes liées à la résolution spatiale modérée (*e.g.* 250 m à 1 km concernant MODIS), à un potentiel manque d'images (*e.g.* ennuagement fort sous les tropiques), à la disponibilité limitée des données de validation ainsi qu'au problème d'instabilité et de saturation du signal. Le problème de saturation du signal, qui se manifeste dans le cas des forêts à couvert fermé, a été bien documenté (Lu, 2006; Lu *et al.*, 2012; Zhao, Zhong and Zhang, 2016) et constitue une limite majeure à l'utilisation de l'imagerie optique satellitaire pour estimer l'AGB ou d'autres variables de structure. Ces incertitudes peuvent aboutir à des divergences importantes dans l'estimation de l'AGB et la cartographie de forêts, notamment en milieu tropical (voir **Figure 2**, issue de Mitchard *et al.*, 2014, comparant des cartes de Saatchi *et al.*, 2011, et Baccini *et al.*, 2012). Néanmoins, il existe des méthodes permettant d'améliorer les résultats. Par exemple, concernant le problème des pixels mixtes (lié à une résolution spatiale insuffisante), l'utilisation d'analyses de démixage du signal (*spectral mixture analysis*) pour décomposer la végétation, le sol et l'ombre à l'intérieur des pixels, avant d'intégrer ces composantes à des données radar (*Discret Wavelet Transform* ou DWT), permettent d'améliorer les estimations de l'AGB (Basuki *et al.*, 2013). Des approches orientées-objet permettent également de contourner ce problème (Kajisa *et al.*, 2009).

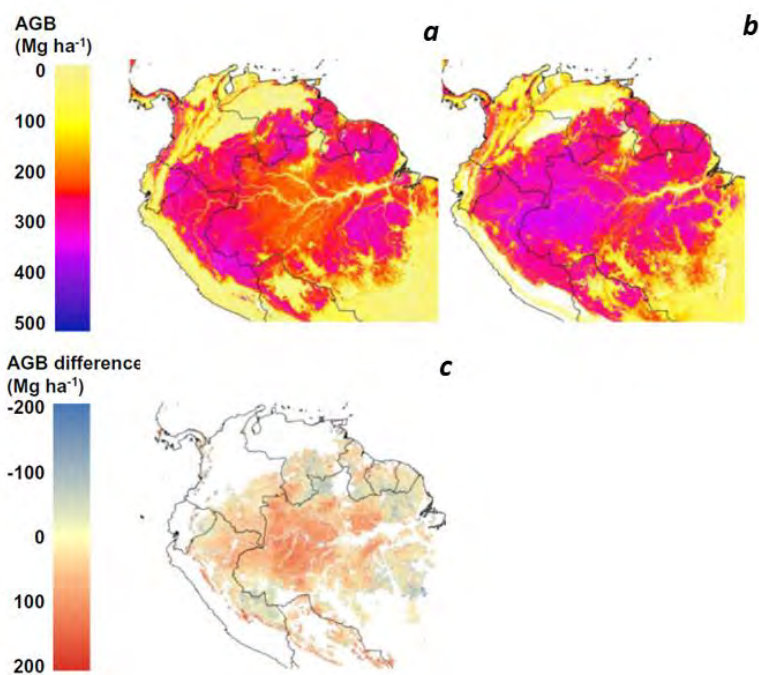


Figure 2. Comparaison de deux cartes d'estimation de l'AGB en Amérique du Sud et leurs différences. Figure extraite de Mitchard *et al.*, (2014) (modifiée). **a.** Carte de l'AGB à partir dans Saatchi *et al.*, 2011 ; **b.** Carte de l'AGB d'après Baccini *et al.*, 2012. **c.** Différence entre les deux cartes. L'AGB a été estimé de manière similaire dans les deux études à l'aide des hauteurs de canopée mesurées par le capteur LiDAR satellitaire GLAS. Elles s'appuient sur des données MODIS pour réaliser l'extrapolation de l'AGB à l'échelle de la région nord de l'Amérique du Sud. Les différences les plus importantes sont concentrées au centre la carte **c.** au niveau du bassin amazonien.

Les approches basées sur les images hyperspectrales ont donné des estimations de l'AGB relativement satisfaisantes bien qu'avec une tendance à la sous-estimation, si l'on compare à l'AGB estimé à l'aide de l'imagerie LiDAR (Clark *et al.*, 2011). La combinaison des deux sources (hyperspectrale et LiDAR) pourrait conduire des estimations de l'AGB plus précises qu'à travers deux approches séparées (Anderson *et al.*, 2008). Néanmoins, la complexité de traitement des données hyperspectrales et la redondance de l'information entre les bandes spectrales « voisines » demeurent un frein important à leur utilisation pour estimer l'AGB (Dube *et al.*, 2016). Comme les images multispectrales, les images hyperspectrales sont sensibles à la saturation dans les milieux forestiers denses.

Pour quantifier l'AGB, il existe également des approches basées sur des mesures radar en particulier concernant le *Synthetic Aperture Radar* (SAR). Le recours aux images radar présente divers avantages : mesures non affectées par les nuages, indépendantes des conditions d'éclairage (elles peuvent être effectuées la nuit), sensibles à l'humidité et à la texture, mesures volumétriques... (Sinha *et al.*, 2015). Toutefois, comme les images optiques, elles sont impactées par le problème de saturation du signal dans le cas des forêts à couvert fermé et dense (Mermoz *et al.*, 2015; Joshi *et al.*, 2017). Le traitement des données radar implique également de faire face à des difficultés spécifiques, notamment les limites intrinsèques liées au type de polarisation et l'influence des caractéristiques du peuplement forestier dont des effets liés au sol. Les données radar ne peuvent être appliquées facilement à tous les types de végétation (Dube *et al.*, 2016). De manière plus générale, l'utilisation des images radar comporte des niveaux d'incertitude qui lui sont propres et demande des efforts de compréhension et de traitement plus importants par rapport aux images issues de l'optique. Pour estimer l'AGB, de meilleurs résultats peuvent toutefois être obtenus en combinant différents types de SAR (polarisé et interférométrique) ou en combinant des capteurs SAR à des capteurs optiques. Les approches combinant des images radar SAR et optiques présentent également l'avantage d'utiliser des sources d'images plus facilement accessibles par rapport aux acquisitions LiDAR (Dube *et al.*, 2016).

b. Utilisation des mesures LiDAR et des analyses texturales pour caractériser la structure de couvert ligneux : focus sur le l'AGB et le LAI

L'utilisation de la technologie LiDAR permet de s'affranchir de certaines contraintes liées aux images radar et optiques. En particulier, les mesures LiDAR ne sont pas affectées par le problème de saturation du signal. Zolkos, Goetz et Dubayah (2013) ont mis en lumière l'intérêt des méthodes d'estimation de l'AGB dérivées du LiDAR : elles se révèlent plus précises que les estimations issues de l'imagerie optique ou radar seules, les estimations dérivées des approches combinant capteurs optique et radar présentaient plus de variabilité et n'étaient pas meilleures que celles dérivées du LiDAR. Il a également été montré que la précision des estimations dérivées du LiDAR variait selon le type de forêt, et que l'augmentation de la taille de la parcelle d'acquisition permettait de réduire les erreurs des modèles prédisant l'AGB.

Il existe différentes métriques dérivées du LiDAR, selon le type de plateforme de mesure (spatiale, aéroportée ou terrestre), la méthode de scanning, le type de signal-retour, et la taille d'emprise de la mesure. Les acquisitions LiDAR permettent par exemple d'établir des Modèles de Hauteur de Canopée ou MNC (**Figure 3**). Il existe aussi de nombreuses méthodes permettant de traiter les données issues des acquisitions LiDAR afin d'estimer l'AGB. Dans le cas de l'étude des forêts tropicales humides, des modèles de régression par apprentissage automatique de type *Random Forest* ont montré de bonnes performances (Mohd Zaki and Abd Latif, 2017). Il a été montré que la hauteur de la canopée et la hauteur moyenne de la canopée (basées sur le MNC) sont des prédicteurs importants pour estimer l'AGB dans les milieux tropicaux et subtropicaux (Mohd Zaki and Abd Latif, 2017).

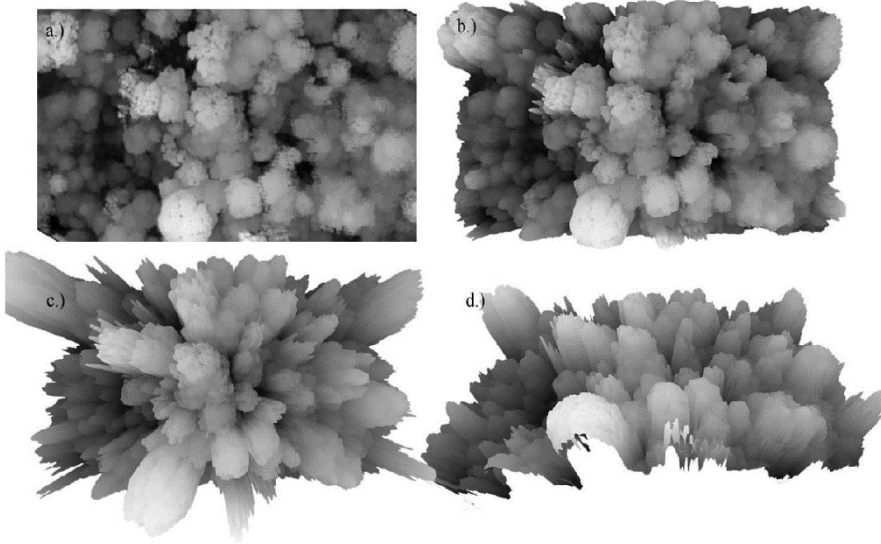


Figure 3. MNC d'une forêt tropicale humide. Figure extraite de Mohd Zaki and Abd Latif (2017). **a.** MNC vu du dessus. **b.** Visualisation 3D du MNC, du dessus. **c.** Interpolation de la hauteur de la canopée, du dessus. **d.** Vue en perspective du MNC.

Malgré le potentiel des mesures LiDAR pour estimer l'AGB, notamment dans le cas des forêts tropicales avec une valeur relativement faible (~ 100 Mg/ha), l'acquisition demeure limitée à des surfaces relativement restreintes (Ploton *et al.*, 2017). La raison principale est liée au coût de l'opération, en particulier concernant le LiDAR aéroporté. Enfin, la précision de l'estimation de l'AGB demeure dépendante de la fiabilité aux modèles allométriques établissant la relation entre biomasse et hauteur de la canopée (Ploton *et al.*, 2017). Les méthodes s'appuyant sur les mesures LiDAR demeurent néanmoins les plus précises à ce jour pour estimer l'AGB et sont particulièrement utiles en l'absence de données terrain pour calibrer les modèles prédisant l'AGB à partir d'images optiques ou radar.

Le *Leaf Area Index* (LAI) est une autre variable biophysique essentielle pour caractériser le couvert forestier donnant accès à une meilleure compréhension des échanges gazeux de la végétation avec l'atmosphère, à différentes échelles, de la feuille au paysage (Zheng and Moskal, 2009). Cet indice se définit classiquement comme le degré de couverture foliaire par unité de surface au sol. Il existe aujourd'hui plusieurs versions du LAI (Zheng and Moskal, 2009). Cartographier le LAI donne accès à la variation dans le temps et dans l'espace de multiples processus et fonctions associés à la végétation : variations de la photosynthèse, évapotranspiration, interception de la pluie, flux de carbone ou encore un indicateur de stress face à la pression d'herbivorie (Zheng and Moskal, 2009).

Différentes méthodes permettent d'estimer le LAI à partir des mesures LiDAR, qui peuvent varier selon le type de LiDAR utilisé (terrestre, aéroporté ou satellitaire) et être confrontées à des limites qui leur sont propres (Wang and Fang, 2020). La méthode d'estimation du LAI la plus fréquemment utilisée repose sur le calcul des moyennes des corrélations avec la fraction d'écart (*gap fraction*) (Wang and Fang, 2020). Le LiDAR terrestre (*Terrestrial Laser Scanning* ou TLS) permet de quantifier le LAI à l'échelle de parcelles d'échantillonnage et est affecté par la complexité de la structure de la canopée (*i.e.* effets d'occlusion, d'agglutination, le type de bois et la taille du voxel). Le LiDAR aéroporté (*Airborne Laser Scanning* ou ALS) permet de quantifier le LAI sur de vastes étendues mais se trouve limité par les effets de seuillage de la hauteur de la canopée, la taille de l'échantillonnage et le type de signal-retour. Le LiDAR satellitaire (*Spaceborne Laser Scanning* ou SLS) donne accès à des estimations du LAI à des échelles globales et est affecté par la topographie et la taille de l'emprise du faisceau du LiDAR (Wang and Fang, 2020). Le LAI peut également être estimé de manière indirecte, par l'intermédiaire de relations allométriques impliquant d'autres métriques dérivées du LiDAR (comme la hauteur de la canopée ou la densité du feuillage) (Wang and Fang, 2020). À ce jour, il existe des lacunes sur la démarche de validation des estimations du LAI, notamment à grande échelle. À l'avenir, de nouvelles méthodes d'estimation basées sur le processus d'inversion à partir des mesures LiDAR appellent à être explorées (Wang and Fang, 2020).

Un dernier type d'approche suscite un intérêt croissant pour estimer l'AGB à partir des données de télédétection. Les analyses de textures basées sur les images panchromatiques à Très Haute Résolution Spatiale (THRS) émergent comme une alternative prometteuse aux approches dépendantes des

acquisitions LiDAR (Ploton *et al.*, 2017). Elles bénéficient à la fois de l'avantage des données optiques satellitaires, à savoir de conditions d'accès relativement simples et peu coûteuses, sur de vastes étendues, et comme dans le cas des mesures LiDAR, elles ne sont pas affectées par le problème de saturation du signal associés aux couverts fermés, comme les forêts tropicales (Bastin *et al.*, 2014). Le principe des analyses de texture (Haralick, 1979) consiste à quantifier l'arrangement des niveaux de gris au sein de l'image de canopée par contraste entre les zones éclairées et les zones ombragées : cette caractérisation de la texture donne accès à la distribution spatiale et à la taille des couronnes des arbres et des trouées entre celles-ci.

La méthode *Fourier Transform Textural Ordination* ou méthode FOTO (Couteron, 2002; Proisy, Couteron and Fromard, 2007) est l'une des principales analyses de texture. Elle vise à extraire les variations pseudo-périodiques de la taille des couronnes des arbres et des trouées à l'intérieur de fenêtres unitaires prédéfinies (*e.g.* 1 ha de surface ou plus). Le principe de la méthode consiste à appliquer, à l'intérieur de chaque fenêtre, une transformation de Fourier sur la fonction d'autocovariance des niveaux de gris, puis à extraire d'un grand nombre de fenêtres les spectres de Fourier obtenus avant de les comparer à l'aide d'une Analyse en Composantes Principales (ACP). Les indices de texture, correspondant aux valeurs projetées sur les axes principaux de l'ACP, peuvent alors être mis en relation via des régressions statistiques à des variables de structure comme l'AGB. Après une étape de calibration, cette relation peut servir à prédire et cartographier les variables de structure en dehors des plots d'échantillonnage (voir **Figure 4** issue de Ploton *et al.*, 2017). De manière similaire aux approches basées sur les mesures LiDAR, le modèle statistique permettant de lier la texture à la structure de la canopée demeure influencé par le site d'acquisition et le type de la végétation locale. En vue d'une généralisation de l'approche, il serait nécessaire de tester la relation sur une diversité de sites comprenant des types de végétation différents à travers différents biomes (Ploton *et al.*, 2017).

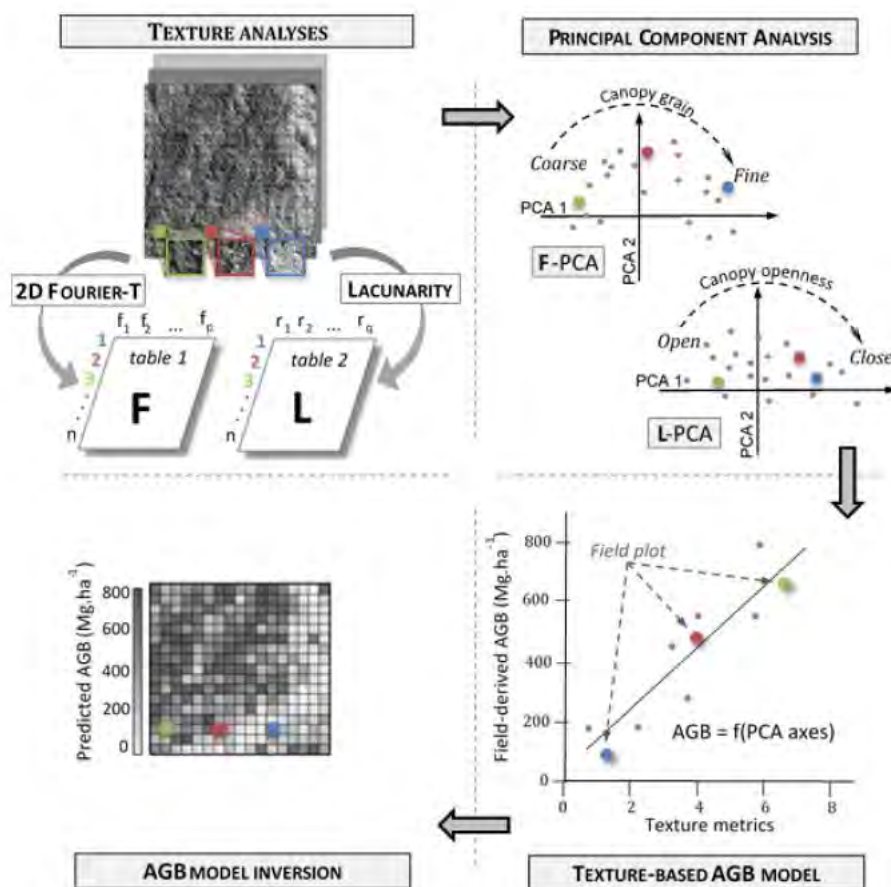


Figure 4. Etapes d'estimation de l'AGB à l'aide des analyses de texture (méthode FOTO et analyse de lacunarité). Figure extraite de Ploton *et al.*, (2017).

La méthode FOTO présente néanmoins deux limites importantes. Les indices de texture ne peuvent distinguer les trouées au sein de la canopée et les couronnes des arbres, qui apparaissent dans les deux cas comme des variations de niveaux de gris. Cette limite appelle à utiliser des analyses de texture complémentaires telles que l'analyse de lacunarité (Frazer, Wulder and Niemann, 2005; Ploton *et al.*, 2017). Une autre limite concerne l'instrumentation : il peut être ardu de différencier les variations de la canopée d'effets liés à la configuration de l'instrumentation (Ploton *et al.*, 2017). Ces effets peuvent provenir de la géométrie du capteur (Barbier *et al.*, 2011; Barbier and Coutron, 2015), de ses propriétés optiques ou de la résolution spatiale (Proisy, Coutron and Fromard, 2007; Ploton *et al.*, 2012). Des approches de simulations 3D de la canopée permettent de faire varier les différents paramètres du capteur (simulé) et aident à mieux évaluer les effets de configuration de l'instrumentation (Frazer, Wulder and Niemann, 2005; Schneider, 2014).

c. Modélisation des propriétés foliaires de la canopée

En dehors de l'estimation de la biomasse épigée et la caractérisation texturale de la canopée, il existe d'autres variables biophysiques qui peuvent être déterminées à l'aide des outils de télédétection sans avoir recours aux modèles dérivés des VIs. Ces variables doivent permettre de caractériser la canopée à travers l'estimation de certaines propriétés foliaires, reliées à des processus d'ordre phénologique (*e.g.* pic de croissance, chute des feuilles), écophysiological (*e.g.* niveau de stress hydrique) ou écologique (*e.g.* stratégies d'adaptation définies dans le *Leaf Economic Spectrum*, voir Wright *et al.*, 2004).

Le *Leaf Mass per Area* (LMA) est un indicateur, par certains aspects, complémentaire au LAI. Il se définit comme le rapport entre la masse sèche d'une feuille et sa surface. Comme le LAI, le LMA peut être relié à une diversité de processus physiologiques (*e.g.* taux de croissance des arbres) et constitue un indicateur de la stratégie adaptative de la végétation (Niinemets, 1999; Westoby *et al.*, 2002). Il renseigne également sur la concentration dans les feuilles de certaines composantes biochimiques comme la concentration en azote (Wright *et al.*, 2004). Le LMA semble dépendre de divers facteurs comme la température et la disponibilité en eau (Poorter *et al.*, 2009), mais avant tout de la quantité de rayonnement incidente (Cunningham, Summerhayes and Westoby, 1999; Niinemets, 1999). Ainsi, des feuilles plongées l'ombre auront tendance à présenter un LMA plus faible par rapport à des feuilles plus exposées à la lumière, de façon à maximiser la surface foliaire en contact avec la lumière par unité d'azote (Evans, 1998). Sur de grandes échelles, le LMA est également un indicateur central pour estimer la variation de la diversité fonctionnelle (Schimel *et al.*, 2015).

Le contenu en eau des feuilles constitue une dimension importante à prendre en compte pour caractériser la structure biophysique de la végétation. Il peut être estimé par l'*Equivalent Water Thickness* (EWT) qui se définit comme la masse volumique d'eau des feuilles par unité de surface. Cet indicateur présente l'intérêt de pouvoir renseigner sur le niveau de tolérance à la déshydratation et sur les stratégies d'utilisation de l'eau de la végétation. Associé au LMA, l'EWT permet d'obtenir une estimation de la quantité de matière combustible (*Fuel Moisture Content*, FMC) en faisant le rapport entre la différence entre masse fraîche et masse sèche d'une part, et la masse sèche d'autre part. L'estimation du FMC permet potentiellement de contribuer à l'évaluation des risques d'incendie (Féret *et al.*, 2019).

Connaître le contenu en chlorophylle de la végétation (*Chlorophyll content* ou CHL) est d'une importance majeure pour comprendre une diversité de processus écophysiologicals à l'échelle de la feuille et de la canopée (Danks, 1983). Les variations du CHL informent notamment sur la productivité primaire, le stade de croissance, l'état de stress de la canopée et le contenu en azote (Ustin *et al.*, 1998; Gitelson *et al.*, 2012; Peng and Gitelson, 2012). A des échelles régionales et globales, estimer la variabilité spatiale du CHL peut donner accès à une meilleure compréhension de la réponse des cultures aux fertilisants et de leur bilan carbone (Gitelson *et al.*, 2005). La chlorophylle n'est néanmoins pas le seul pigment renseignant sur l'écophysiological de la canopée. Par exemple, les anthocyanes sont fortement associés à la sénescence et à la chute des feuilles en milieu tempéré et peuvent être quantifiés à partir des propriétés optiques des feuilles.

Différentes méthodes permettent d'estimer les variables biophysiques décrivant des propriétés foliaires de la canopée à partir de données de télédétection (optiques, radar ou LiDAR). On peut

distinguer deux types de modèle auxquels ces méthodes font appel : des modèles empiriques (*data-driven*) et des modèles physiques de transfert radiative (RTMs) (Féret *et al.*, 2017). Dans les deux cas, l'objectif consiste à prédire la variation des variables biophysiques à des échelles régionales voire globales.

Pour prédire les variables biophysiques de la canopée, les modèles empiriques visent à établir des relations entre un jeu de données de référence et des mesures spectrales (le jeu de données de référence peut être de différente nature : par exemple, dans le cas du LAI, il peut s'agir de mesures destructives sur le terrain ou de données LiDAR). Ces modèles comprennent aussi bien des approches statistiques univariées, multivariées (Gitelson, Keydan and Merzlyak, 2006; Asner and Martin, 2009) que des algorithmes d'apprentissage automatique (*machine learning*) (Verrelst *et al.*, 2015). La robustesse des résultats obtenus dépend fortement de la qualité du jeu de données utilisé pour paramétrer les modèles (Féret *et al.*, 2017).

Le second type de modèle cherche à simuler le processus du transfert radiatif à l'échelle de la feuille ou de la canopée. A l'échelle de la feuille, le modèle PROSPECT (Jacquemoud and Baret, 1990) est le plus connu et le plus utilisé (Jacquemoud *et al.*, 2009). Il consiste à modéliser la réflectance et la transmittance directionnelle-hémisphérique d'une feuille sur le domaine optique allant de 400 nm à 2500 nm. Pour simuler les propriétés optiques d'une feuille, PROSPECT s'appuie sur la modélisation d'une structure de feuille sous la forme d'une ou plusieurs plaques compactes homogènes empilées, dans laquelle diffuse la lumière, d'après Allen *et al.*, (1969). Les propriétés d'absorption de chacune de ces couches sont définies à partir du contenu en différents constituants par unité de surface foliaire, le coefficient d'absorption spécifique de ces constituants étant connu. Les constituants pris en compte par PROSPECT sont la chlorophylle totale (*a* et *b*), les caroténoïdes, les anthocyanes, le LMA et l'EWT (Jacquemoud and Baret, 1990; Féret *et al.*, 2008, 2017). Plusieurs versions de PROSPECT se sont succédées, chaque nouvelle version intégrant de nouveaux pigments, affinant son potentiel d'utilisation pour des applications dans le domaine de l'étude de la végétation. Les versions PROSPECT-4 et PROSPECT-5 développées par Féret *et al.*, (2008) ont permis de séparer les effets des chlorophylles et des caroténoïdes (au lieu de simuler un effet global des pigments combinés). La version suivante, PROSPECT-D (Féret *et al.*, 2017) intègre les anthocyanes qui jouent un rôle important en lien avec le stress et le processus de sénescence, notamment en milieu tempéré, permettant ainsi de simuler le comportement spectral de la végétation à travers les différents moments de son cycle de vie. Enfin, la dernière version, PROSPECT-PRO (Féret *et al.*, 2021), permet de séparer les effets des protéines et des composés carbonés (*e.g.* lignine, cellulose) contenus dans la matière sèche (LMA). PROSPECT-PRO rend possible l'estimation des contenus en azote et en carbone foliaires, ouvrant la voie à des nouvelles applications en écologie et en agriculture de précision (Féret *et al.*, 2021), mais nécessite une information spectralement plus riche que celle fournie actuellement par les satellites multispectraux comme Sentinel-2 et Landsat.

A l'échelle de la canopée, le modèle SAIL (*Scattering by Arbitrary Inclined Leaves*), développé il y a une trentaine d'années (Verhoef, 1985), a permis de modéliser la structure de la canopée sous la forme d'une couche homogène turbide horizontale contenant des particules absorbantes et diffusantes. Il s'agit du modèle de canopée le plus populaire (Jacquemoud *et al.*, 2009). A l'instar de PROSPECT, SAIL a été développé sous différentes versions, afin de modéliser des structures de canopées plus complexes et potentiellement plus proches de la réalité, tout en conservant un nombre limité de paramètres, ce qui facilite son inversion. C'est notamment le cas de la version robuste et numériquement optimisée 4SAIL (Verhoef *et al.*, 2007). Au cours de son évolution, SAIL a pu être combiné avec des modèles géométriques (GeoSail), intégrer les effets d'« agglutination » des couronnes des arbres (4SAIL2) ainsi que des effets d'émissions de radiations, comme la fluorescence de la chlorophylle *a* (FSAIL, FluorSAIL) ou les émissions thermiques (4SAIL) (Jacquemoud *et al.*, 2009).

L'inversion de SAIL, qui permet d'estimer les variables biophysiques de la canopée à partir de la réflectance, nécessite une acquisition d'images dans de nombreuses directions. La plupart des capteurs satellitaires n'effectuent des acquisitions que sous un unique angle. Très tôt, le modèle SAIL a été combiné avec PROSPECT de manière à résoudre ce problème tout en intégrant la dimension biochimique de la canopée (Jacquemoud *et al.*, 2009). Le modèle PROSAIL combine la réflectance et la transmittance des feuilles simulées par PROSPECT à une structure de canopée telle qu'elle est modélisée par SAIL, simulant finalement la réflectance globale de la canopée (Baret *et al.*, 1992). PROSAIL prend

également en compte la réflectance du sol. Il existe une diversité de versions de PROSAIL, selon les versions de PROSPECT et de SAIL utilisées. Le **Figure 5**, issue de Berger *et al.*, (2018), résume le principe de fonctionnement de PROSAIL (version PROSPECT-D + 4SAIL). Le **Tableau 1**, tiré du même article, expose les principales variables du modèle avec les gammes de valeurs typiques pour différentes cultures déterminées à partir de la littérature.

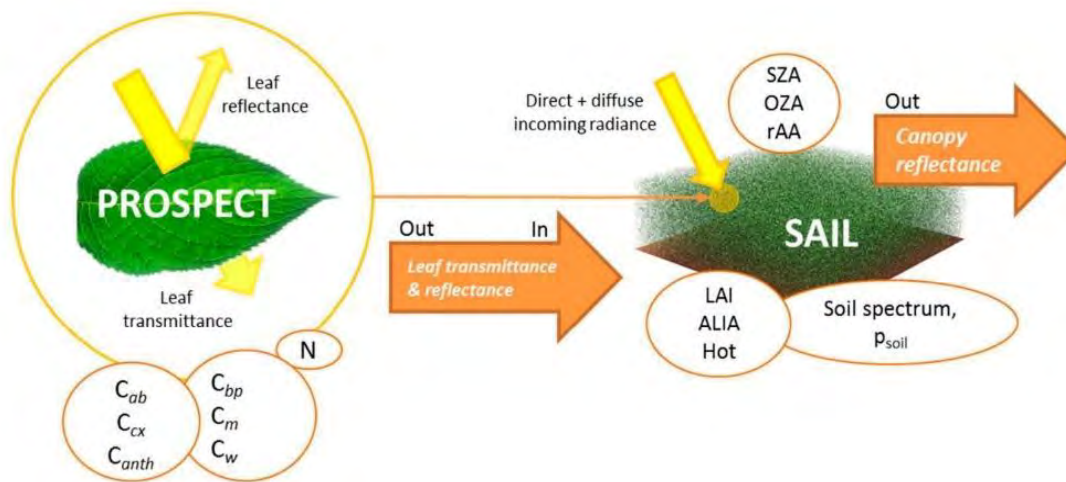


Figure 5. Schéma récapitulatif du principe de fonctionnement de PROSAIL. Figure extraite de Berger *et al.* (2018).

PROSAIL est utilisé pour différentes applications (*e.g.* déterminer de nouveaux indices de végétation, estimer la sensibilité de la réflectance de la canopée aux propriétés optiques). La plus courante étant l'estimation des variables biophysiques de la canopée par inversion du modèle, à partir des réflectances issues d'images multispectrales (*e.g.* MODIS, Landsat) ou, dans une moindre mesure, en raison de leur disponibilité plus récente, hyperspectrales (principalement capteurs au sol de type ASD, ou aéroportés tel le CASI – *Compact Airborne Spectromgraphic Imager*) (Berger *et al.*, 2018). Les techniques d'inversion sont de trois types : statistique, physique ou hybride (Berger *et al.*, 2018). Les méthodes statistiques s'appuient sur des modèles empiriques qui peuvent être paramétriques ou non-paramétriques. Les modèles paramétriques cherchent à ajuster les paramètres du modèle à l'aide de régressions linéaires basées sur certaines variables, comme des rapports d'indices de végétation, des indices de végétation orthogonaux ou encore sur des indices dérivés de mesures du continuum du spectre lumineux. Les modèles non-paramétriques s'appuient sur d'autres méthodes linéaires (régressions multiples de type *stepwise* ou des régressions de composantes principales) ou sur des régressions non-linéaires basées sur des méthodes d'apprentissage automatique, notamment les réseaux de neurones artificiels. Les méthodes physiques s'appuient sur la recherche de la meilleure combinaison possible entre des valeurs spectrales mesurées et simulées. Elles font appel à des méthodes d'optimisation itératives et des méthodes Monte Carlo de chaînes de Markov (complexes et demandant une importante puissance de calcul) ou à des Tables des Correspondances (*Look-Up Tables* ou LUT). Cette dernière méthode est l'une des plus utilisées jusqu'à aujourd'hui pour obtenir l'inversion de PROSAIL (Berger *et al.*, 2018). L'utilisation des LUT conservent certains désavantages, notamment un temps de calcul prolongé dans le cas de traitements de données hyperspectrales. Certains algorithmes permettent d'améliorer significativement la recherche de la meilleure combinaison dans les LUT, notamment la méthode de régression par processus gaussiens (*Gaussian Processes Regression* ou GPR). Les méthodes hybrides, utilisant conjointement des LUT et des régressions non-linéaires et non-paramétriques basées sur des algorithmes d'apprentissage automatique ont permis de faciliter le processus d'inversion de PROSAIL. Quel que soit la méthode utilisée, il n'existe pas de solution unique au problème de l'inversion (puisque différentes combinaisons de variables peuvent conduire à des réflectances similaires) (Berger *et al.*, 2018).

Tableau 1. Les seize variables (ou paramètres) intégrées dans le modèle PROSAIL (PROSPECT-D + 4SAIL) avec leurs gammes de valeurs typiques pour cinq cultures d'après la littérature. La figure est extraite de Berger *et al.*, (2018).

Parameter	Symbol	Units	Typical Ranges for Crops				
			Maize [35–37]	Wheat [37,38]	Rice [39,40]	Soybean [41,42]	Sugar Beet [43–45]
<i>Leaf Model: (PROSPECT-D)</i>							
Leaf structure index	N	Unit less	1.2–1.8	1.0–2.5	1.0–2.0	1.2–2.6	1.0–1.5
Chlorophyll a + b content	C _{ab}	($\mu\text{g}/\text{cm}^2$)	0–80	0–80	0–80	0–80	20–45
Total carotenoid content	C _{cx}	($\mu\text{g}/\text{cm}^2$)	1–24	1–24	4–17	-	-
Total anthocyanin content	C _{anth}	($\mu\text{g}/\text{cm}^2$)	-	-	-	-	-
Brown pigments	C _{bp}	Unit less	0–1	0–1	0–1	0–1	0–1
Dry matter content, or leaf mass per area	C _m /LMA	(g/cm^2)	0.004–0.0075	0.001–0.02	0.001–0.02	0.001–0.02	0.004–0.007
Equivalent water thickness, or water depth	EWT/C _w	(cm)	0.01–0.03	0.001–0.05	0.001–0.002	0.001–0.05	0.03–0.08
<i>Canopy Model: (4SAIL)</i>							
Leaf area index	LAI	(m^2/m^2)	0–7	0–8	0–10	0–7	0–4
Average leaf inclination angle * or: Leaf inclination distribution function **	ALIA LIDF _{a/b}	($^\circ$) ($^\circ$)	20–70 [30]	20–90	20–80	10–75	20–40
Hot spot parameter	Hot	(m/m)	0.01–0.2	0.01–0.5	0.01–0.1	0.2	0.2–0.4
Soil reflectance	ρ_{soil}	(%)					
Soil brightness factor	α_{soil}	Unit less				0.5–1.5 *** or 0–1 ****	
Fraction of diffuse illumination	sky1	Unit less				23% for a standard clear sky [33]	
Sun zenith angle	SZA/ θ_s	($^\circ$)					
Viewing (observer) zenith angle	OZA/ θ_v	($^\circ$)					
Relative azimuth angle between sun and sensor	rAA/ θ_{sv}	($^\circ$)					
* characterizes an ellipsoidal leaf inclination model; ** spherical, planophile, erectophile, uniform, extremophile or plagiophile types. LIDF is characterized by LIDF _a , which controls the average leaf slope and LIDF _b , which controls the distribution's bimodality; *** to be multiplied with single ρ_{soil} spectrum; **** scaling factor between the two model-implemented ρ_{soil} spectra (wet versus dry).							

Le succès que connaît PROSAIL tient en grande partie à ce processus d'inversion, qui impose une contrainte spectrale forte pour estimer les variables biophysiques de la végétation, limitant le nombre de variables inconnues. De plus, l'inversion ne nécessite pas de calibration préliminaire avec mesures sur le terrain, simultanément au passage du capteur (Berger *et al.*, 2018). Un consensus émerge actuellement sur la pertinence de PROSAIL, qui permet des estimations relativement précises, compte tenu du niveau de complexité du modèle, avec un temps de calcul acceptable (Berger *et al.*, 2018).

PROSAIL, dans sa version initiale, a été développé avec un intérêt singulier pour les applications qu'il permet en agriculture : la structure de canopée modélisée par SAIL sous forme d'une couche homogène turbide s'applique particulièrement bien aux cultures. Environ la moitié des études (49%) utilisant l'inversion de PROSAIL (jusqu'à l'année 2018, date de la dernière *review*) ont été appliquées à des cultures, en particulier le blé et le maïs, dans des contextes géographiques et climatiques différents (Berger *et al.*, 2018). Bien qu'initialement non adapté aux végétations possédant une canopée complexe et hétérogène, Berger *et al.*, (2018) recensent 15 % d'utilisation de l'inversion de PROSAIL (dans sa version de couche homogène turbide) sur des milieux forestiers et 14 % dans d'autres milieux « naturels » comme les prairies (GéoSail et les versions multicouches de SAIL prennent en compte des canopées plus complexes). Le LAI, suivi par le CAB (*Chlorophyll a + b content*), constitue de loin la variable biophysique la plus ciblée par les études, en raison de son importance centrale pour approcher les processus écophysiologiques de la végétation (Berger *et al.*, 2018). Le CAB est également fréquemment estimé, en raison de son lien avec le contenu en azote des feuilles (Berger *et al.*, 2018). L'estimation du LAI et du CAB s'est révélée satisfaisante pour de nombreux types de cultures, y compris pour des cultures avec des structures de canopée plus complexes (*e.g.* coton, pommes de terre, tournesol). Néanmoins, les conditions d'application de l'inversion de PROSAIL, en fonction de la structure de la végétation, doit être davantage étudiée, notamment l'influence du sol dans la détermination du LAI qui peut être plus importante dans des cultures discontinues ou pour des stades de croissance précoce (un faible LAI laisse entrevoir la réflectance du sol) (Berger *et al.*, 2018).

À ce jour, la caractérisation biophysique de certains types de milieux par inversion de PROSAIL a été très peu investiguée. C'est notamment le cas des forêts tropicales. Les raisons principales de cette lacune sont liées à la complexité de la structure des canopées tropicales (qui ne rencontre pas *a priori* les conditions d'application de SAIL, surtout dans sa forme initiale) et la difficulté d'acquisition des images, notamment en raison du fort taux d'enneuagement. Par ailleurs, la difficulté d'acquisition de données de validation (terrain souvent difficile d'accès) a également pu constituer un frein à leur investigation (qui n'est pas exclusif à ce type d'approche). Néanmoins, l'efficacité et la précision des estimations des variables biophysiques comme le LAI et le CAB dont PROSAIL a été capable dans de nombreuses études, y compris concernant des végétations plus complexes (*e.g.* champs de coton) suggère que le potentiel du modèle mériterait d'être exploré sur les forêts tropicales, au regard des enjeux majeurs qu'elles concentrent en termes de conservation (Gardner *et al.*, 2009; Mittermeier *et al.*, 2011) et de régulation du climat (Townsend *et al.*, 2011; Malhi, 2012; Mitchard, 2018).

Dans ce contexte, le potentiel des images à haute résolution spatiale, et avec une diversité de bandes permettant une caractérisation fine de la signature spectrale de la végétation, constituerait une opportunité pour tester l'inversion de PROSAIL sur des végétations complexes et spectralement plus riches. Les images Sentinel-2 en particulier, ont l'avantage de posséder plusieurs bandes du domaine *Red Edge* (associé à une forte réflectance de la chlorophylle) une haute résolution spatiale (jusqu'à 10 m) et temporelle (les satellites repassant tous les 5 jours) permettant de pallier le nombre limité d'images faiblement enneuagées. Enfin, les images Sentinel sont plus simples à traiter par rapport aux images hyperspectrales. L'utilisation de Sentinel-2 avec PROSAIL est relativement récente, comme source unique (Masemola, Cho and Ramoelo, 2020; Sinha *et al.*, 2020) ou dans des approches combinant d'autres images issues de l'imagerie optique satellitaire (Punalekar *et al.*, 2018; Darvishzadeh *et al.*, 2019).

III. Cartographier la biodiversité du couvert ligneux

L'estimation de la diversité du couvert ligneux est une dimension importante pour comprendre son fonctionnement. Un pan entier de recherche en écologie (*Biodiversity-Ecosystem Functions* ou BEF ; *e.g.* Naeem and Wright, 2003; Tilman, Isbell and Cowles, 2014) s'est consacré à établir des relations entre la biodiversité et certaines fonctions des écosystèmes (*e.g.* la relation entre la Richesse Spécifique et la production de biomasse ; Cardinale *et al.*, 2007). Connaître la diversité associée au couvert ligneux permet de donner accès à une information complémentaire à la caractérisation biophysique et potentiellement d'améliorer la compréhension de leur distribution spatiale. Avec le déploiement de cartographies à larges échelles basées sur des techniques de télédétection, cette caractérisation constitue également un appui indispensable pour établir des stratégies de conservation à des échelles régionales et globales.

a. Principales approches pour cartographier la biodiversité par télédétection

La biodiversité peut être estimée de multiples façons, selon la dimension à laquelle on s'intéresse (taxonomique, fonctionnelle, spécifique, phylogénique...) et la méthode employée. Une approche qui facilite la comparaison entre les écosystèmes ou les communautés d'espèces passe par le calcul d'indices de diversité. Il en existe de nombreux : les indices basés sur le nombre d'espèces (*e.g.* Richesse Spécifique, Indice de Margaleff, Chao 1 et Chao2), sur l'équitabilité (*e.g.* Equitabilité de Simpson, indice de Berger Parker), les indices hétérogènes qui prennent en compte à la fois le nombre d'espèces et l'équitabilité (*e.g.* Indice de Shannon, Indice de Simpson), les indices basés sur les dimensions taxonomique, fonctionnelle et phylogénétique (*e.g.* Entropie quadratique) et enfin les indices décomposant la diversité spatialement ou temporellement (diversité α , β et γ). À titre d'exemple, la **Figure 6**, extraite de Villéger, Mason and Mouillot, (2008) (modifiée) illustre des trois composantes de la Diversité Fonctionnelle, qui fait partie des concepts majeurs en écologie : Richesse, Equitabilité et Divergence Fonctionnelles. Elle se base sur la mesure des traits fonctionnels, définis comme : « *toute caractéristique morphologique, physiologique ou phénologique mesurable au niveau individuel, de la cellule à l'organisme entier, sans référence à l'environnement ou à tout autre niveau d'organisation* » (Violle *et al.*, 2007) impactant le succès reproductif de l'individu.

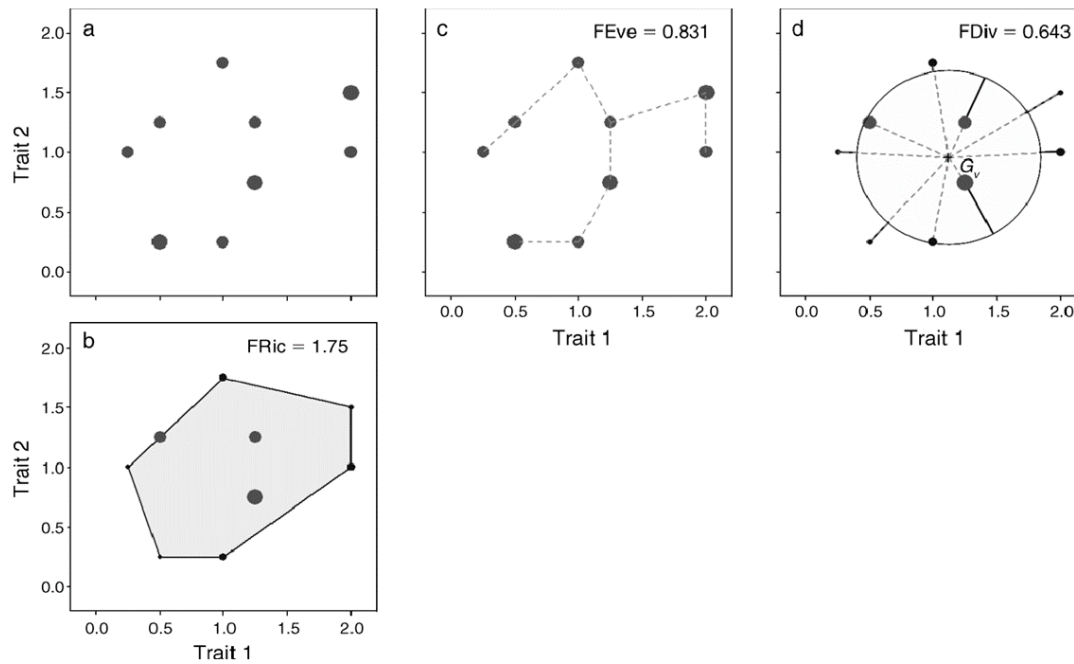


Figure 6. Principe de détermination des trois composantes de la diversité fonctionnelle. Figure extraite de Villéger *et al.*, (2008) (modifiée). Dans cet exemple, deux traits fonctionnels et neuf espèces sont considérées. **a.** Les espèces sont positionnées dans l'espace fonctionnel multidimensionnel en fonction de leurs valeurs de traits (chaque point représentant une espèce et la taille du point correspondant à son abondance). **b.** La Richesse Fonctionnelle correspond au volume défini par la position des espèces au sein de l'espace fonctionnel. **c.** L'Équitabilité Fonctionnelle mesure la régularité des points le long de l'Arbre Couvrant de poids Minimal (ACM). **d.** Les lignes pointillées représentent la distance des points à au centre de gravité du nuage, le cercle noir la distance moyenne à ce centre et les lignes pleines l'écart des points à la distance moyenne. Pour les espèces abondantes (gros points), plus l'écart à la distance moyenne est important, plus la Divergence Fonctionnelle est élevée.

Ces différents indices permettent d'accéder à des dimensions complémentaires de la biodiversité. Une caractérisation relativement complète de la diversité d'un écosystème ou d'une communauté requiert l'utilisation d'indices complémentaires. Les approches statistiques multivariées sont aujourd'hui largement utilisées pour explorer ces différentes dimensions de manière synthétique en évaluant leur niveau de complémentarité.

L'un des défis majeurs pour les écologues consiste aujourd'hui à suivre la variation de la biodiversité à de grandes échelles. Pouvoir cartographier la biodiversité est également une nécessité au déploiement de stratégies globales de conservation. La télédétection s'est imposée comme l'un des outils majeurs pour cartographier la biodiversité (Cohen *et al.*, 2003; Gillespie *et al.*, 2008). Les techniques de télédétection ont été utilisées en conservation avec divers objectifs. Par exemple, dans le but de cartographier la pollution lumineuse affectant certaines espèces (*e.g.* Mazor *et al.*, 2013), localiser des sites de reproduction (*e.g.* Gómez-Rodríguez *et al.*, 2008) ou estimer des indices de diversité (*e.g.* Walters and Scholes, 2017). Wang and Gamon, (2019) distinguent quatre approches principales permettant de caractériser la diversité de la végétation à partir des techniques de télédétection : la cartographie des habitats (une approche indirecte de la diversité), l'étude de la distribution spatiale des individus (en vue de la caractérisation de la composition et de la diversité des communautés), l'estimation de la diversité fonctionnelle (par les traits fonctionnels) et la cartographie des motifs de diversité basés sur la variabilité spectrale (l'hypothèse sous-jacente étant que les individus d'une même espèce auront des signatures spectrales proches). Ces quatre approches sont résumées ci-dessous dans le **Figure 7**, extraite de Wang et Gamon (2019).

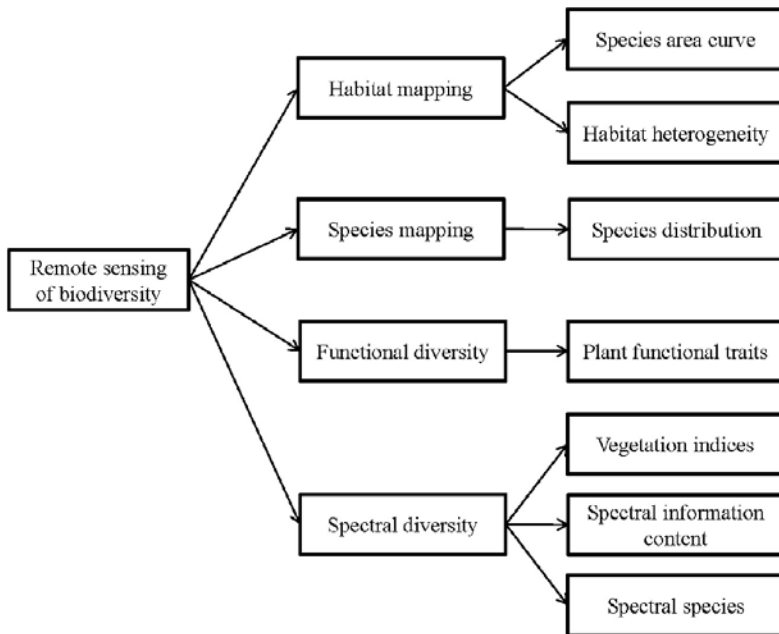


Figure 7. Résumé des quatre approches principales pour cartographier la diversité de la végétation. La Figure est extraite de Wang and Gamon (2019). NB. L'approche indirecte par cartographie d'habitat peut également s'appliquer à d'autres groupes taxonomiques (e.g. insectes, oiseaux).

La diversité des habitats constitue l'un des facteurs les plus importants déterminant la Richesse Spécifique. Il a été montré que la biodiversité augmente avec l'hétérogénéité du milieu, augmentant le nombre d'habitats potentiels rencontrant les exigences écologiques d'un plus grand nombre d'espèces, conformément à la théorie de la niche écologique (Hutchinson, 1957; Tews *et al.*, 2004; Stein, Gerstner and Kreft, 2014; Tuanmu and Jetz, 2015). La cartographie des types d'occupation du sol (*Land cover mapping*) s'appuyant sur les techniques de télédétection est une approche classique pour identifier les habitats et leur distribution spatiale (Stoms and Estes, 1993; Tomaselli *et al.*, 2013). Cette approche a notamment pour application la cartographie des habitats d'espèces faisant l'objet de programmes de conservation. La précision de la cartographie des habitats, et de l'estimation de la biodiversité qui en découle, dépend largement de l'échelle considérée et finalement de la résolution spatiale du capteur. Par exemple, l'utilisation d'images à 500 m de résolution comme MODIS ne peut généralement pas rendre compte de la variabilité intra-habitats et tend à sous-estimer la diversité. Les photographies aériennes peuvent s'avérer utiles pour cartographier les habitats à travers leur résolution spatiale plus fine, mais sont limitées par l'information spectrale qui ne permet pas de discriminer les habitats de manière suffisamment pertinente. Enfin, cette approche indirecte par cartographie des habitats rend davantage compte la biodiversité potentielle qu'effective. Elle ne permet pas par exemple de prédire la distribution d'espèces rares ou des espèces dont les exigences écologiques sont peu spécifiques. Un autre type d'approche permettant d'estimer indirectement la biodiversité se base sur les courbes aires-espèces (Stoms and Estes, 1993; Turner *et al.*, 2003).

Les mesures LiDAR et les images hyperspectrales aéroportées ont grandement contribué à la cartographie des espèces et des types fonctionnels de la végétation, notamment dans le cas des espèces à larges couronnes (Wang and Gamon, 2019). La richesse de l'information spectrale des images hyperspectrales à des résolutions spatiales fines peut être reliée à la diversité des traits fonctionnels, dans la mesure où la variation de nombreux traits fonctionnels sont associés à des variations des propriétés optiques de la végétation. À travers l'estimation des traits fonctionnels, il devient possible de cartographier la diversité des niches écologiques et des grandes fonctions de la végétation sans nécessité d'identifier les espèces individuellement (Wang and Gamon, 2019). Des méthodes statistiques comme les *Partial Least Square Regressions* (PLSR) permettent de déterminer des coefficients spectraux (associé à des combinaisons de valeurs de traits biophysiques ou biochimiques) pouvant être utilisés comme des proxys de la Diversité Fonctionnelle. La cartographie des traits fonctionnels de la végétation peut être réalisée à l'échelle de la feuille ou de la canopée. Bien qu'il ne soit généralement pas possible d'identifier directement les espèces individuelles, il est parfois possible, à partir de certaines caractéristiques spectrales, d'identifier les espèces spatialement dominantes dans un milieu (e.g.

espèces invasives) ou des types de communauté ou des groupes fonctionnels distincts (*e.g.* fixateurs vs non-fixateurs d'azote) (Wang and Gamon, 2019).

Enfin, il existe diverses approches permettant de cartographier la diversité de la végétation en utilisant la diversité spectrale comme un proxy. L'Hypothèse de Variabilité Spectrale (*Spectral Variation Hypothesis* ou SVH) proposée par Palmer *et al.*, (2000, 2002) suggère que la variabilité spectrale peut être considérée comme un proxy de la variabilité spatiale de la biodiversité. En effet, la SVH, en s'appuyant sur la théorie de la niche écologique, propose que plus la variabilité spectrale est importante (pour une zone donnée), plus l'hétérogénéité de l'environnement l'est également et donc que le nombre de niches augmente et avec lui le nombre d'espèces, quel que soit le groupe taxonomique considéré (Schlemmer *et al.*, 2013). En pratique, la SVH s'applique principalement à la diversité végétale, qui est la plus directement détectée à travers la canopée. L'hypothèse a été testée dans diverses études, pour des images multispectrales et hyperspectrales, montrant que les résultats dépendent de divers facteurs, comme les caractéristiques du capteurs (dont sa performance), le type de végétation et les métriques utilisées pour estimer l'hétérogénéité spatiale (Rocchini *et al.*, 2010, 2018). Les approches s'appuyant sur la SVH peuvent être classées à travers la détermination de trois types de métriques (Wang and Gamon, 2019) : les métriques reposant sur les Indices de Végétation, celles basées sur la théorie de l'information (*e.g.* entropie spectrale), et enfin celles faisant appel à la méthode des « espèces spectrales ».

b. Cartographier la diversité α et β à partir des espèces spectrales

Une manière aujourd'hui largement éprouvée de quantifier la biodiversité, spatialement explicite, repose sur l'estimation de la diversité α et β (Whittaker, 1972). L'objectif consiste à comparer la diversité des communautés d'espèces entre différents sites (la méthode peut s'appliquer à des communautés séparées spatialement ou temporellement). La diversité α ou diversité « locale » correspond au nombre de taxons et leurs abondances dans un site donné. La diversité β représente la variation de la composition entre deux sites, autrement dit, entre deux communautés (Whittaker, 1972; Anderson *et al.*, 2011). Elle peut être vue comme la diversité exprimée à une échelle « régionale » (inter-sites) ou exprimant un turn-over d'espèces entre deux sites. α et β peuvent être estimés par différentes métriques, comme l'indice de Shannon pour la diversité α , et la dissimilarité de Bray-Curtis pour la diversité β (Féret and Asner, 2014). L'étude de la diversité α et β a donné lieu à une littérature très importante, appliquée à une diversité de communautés, du microbiome aux ensembles forestiers (*e.g.* Valencia, Balslev and Paz Y Miño C, 1994; Prober *et al.*, 2015; Schmidt *et al.*, 2017).

Le plus souvent, les images satellitaires (optiques) ont été utilisées afin de localiser des *hot spots* de biodiversité (Rocchini *et al.*, 2010) qui peuvent être révélés à travers les variations de la diversité α . La diversité β , plus complexe à estimer, a également suscité un intérêt important (Féret and Asner, 2014; Rocchini *et al.*, 2018). L'estimation de la diversité α et β par télédétection repose sur la SVH. Différents types d'approches, supervisées et non-supervisées, permettent de quantifier la diversité spectrale (Rocchini *et al.*, 2018). Dans le premier cas, le potentiel d'utilisation demeure limité, notamment à cause de la faible disponibilité des jeux de données de validation. Les approches non-supervisées, telle que celle proposée par Féret et Asner (2014), permettent de caractériser l'hétérogénéité environnementale à l'échelle du paysage de manière plus étendue. Ces méthodes ont recours à plusieurs niveaux d'abstractions avant d'estimer des indicateurs comme la diversité α et β (Rocchini *et al.*, 2018). L'approche proposée par Féret et Asner (2014) consiste à affecter à chaque pixel de l'image une « espèce spectrale » (c'est-à-dire un groupe contenant des pixels de valeurs proches) à partir d'une méthode de *clustering* (*e.g.* méthode des *k-means*). Elle fait l'hypothèse que les individus d'une même espèce auront tendance à se retrouver dans le même groupe, autrement dit, à être affecté à une même espèce spectrale (partageant des signatures spectrales proches). Les cartes d'espèces spectrales obtenues peuvent alors être utilisées pour calculer des indices de diversité, notamment des métriques pour estimer la diversité α et β . Cette approche constitue l'une des principales pour cartographier la diversité α et β (Rocchini *et al.*, 2016, 2018). La méthode présente l'avantage d'être moins affectée par la présence de bruit, provoqué par la présence de pixels indésirables non-filtrés (eau, sol nu, zones urbaines, etc.) mais est aussi limitée par la faible disponibilité de données de validation issue de la télédétection, peu de capteurs possédant à la fois une haute résolution spectrale et spatiale (Rocchini *et al.*, 2016).

L'application de cette méthode peut nécessiter des ajustements selon le type d'environnement qu'elle échantillonne. Les milieux tropicaux se révèlent plus complexes à traiter en raison du nombre extrêmement élevé d'espèces qu'ils contiennent. Il est permis d'espérer que les missions spatiales futures (e.g. EnMAP, HYPXIM) permettront de fournir une information spectrale adaptée pour quantifier des environnements avec de hauts niveaux de diversité. Certaines difficultés demeurent néanmoins persistantes, notamment le problème de l'ennuagement qui peut être particulièrement limitant (Rocchini *et al.*, 2016). Dans les milieux tempérés et boréaux, les difficultés sont moins importantes, principalement en raison des variations phénologiques associées aux espèces qui permettent de les différencier plus facilement et un ennuegement plus faible (Rocchini *et al.*, 2016).

Par ailleurs, certaines difficultés mériteraient d'être davantage explorées. Choisir la taille de la fenêtre d'échantillonnage des pixels (avant de calculer les métriques de diversité α et β) influence de manière importante les résultats : une taille trop « petite » (contenant peu de pixels) limite par exemple la diversité potentiellement calculable au sein de chaque fenêtre. La taille « idéale » de la fenêtre d'échantillonnage peut également varier selon les milieux. Par exemple, dans le cas d'une forêt tropicale dense, Féret et de Boissieu (2019) suggèrent d'utiliser une taille de fenêtre entre 4 et 0.25 ha pour des données à haute résolution spatiale (10-20 m). La résolution spatiale peut également influencer l'estimation de la diversité. En effet, des pixels trop « petits » (par exemple, avec une résolution spatiale de 5 m) risquent de conduire à une augmentation du bruit, en particulier à cause des effets d'ombre qui augmentent artificiellement l'hétérogénéité (Nagendra and Rocchini, 2008; Stickler and Southworth, 2008), tandis que des pixels trop « grands » (résolution spatiale plus faible) risquent de lisser des motifs de diversité plus subtiles. La méthode mériterait d'être testée à travers différentes tailles de fenêtres et résolutions, de manière à identifier les combinaisons les plus adaptées au taxon ou au milieu étudié. Dans une optique de conservation, cette approche basée uniquement sur la télédétection ne peut entièrement suffire. En effet, la détermination d'indices de diversité, comme les indices α et β ne permet pas de renseigner sur l'importance écologique relative des milieux analysés et risque de masquer certaines spécificités qui participent aux choix des zones prioritaires de conservation. Les éléments de configuration du paysage, qui impactent directement la capacité des espèces à se disperser et permet d'expliquer les niveaux de diversité observé ne sont pas observables par cette approche (Rocchini *et al.*, 2016). Dans de futures études, il serait intéressant d'explorer la piste d'une combinaison de la cartographie de la diversité dérivée des espèces spectrales avec des modélisations explicites des corridors et îlots caractérisant les paysages, qui sont rattachées à des concepts centraux de l'écologie du paysage (Turner, 1989). D'une manière plus générale, des efforts doivent être faits pour mieux éclairer les mécanismes qui lient la variabilité spectrale et l'hétérogénéité environnementale sur laquelle repose l'estimation de la biodiversité.

IV. Problématique et hypothèses de recherche

Nous cherchons à mettre en évidence quels pourraient être les variables biophysiques et les indices de biodiversité les plus pertinents, parmi un ensemble de méthodes proposées dans les packages *prosail* et *biodivMapR*, pour caractériser un couvert ligneux tropical humide à une échelle régionale. Pour répondre à cette question, nous cherchons à tester les hypothèses suivantes :

1. *Les variables biophysiques et indices de diversité produites à partir de l'information spectrale des données Sentinel-2 ont une pertinence écologique contribuant potentiellement à l'analyse des motifs taxonomiques et fonctionnels des forêts tropicales.*
2. *Ces informations issues de Sentinel-2 sont complémentaires aux informations issues d'autres sources de télédétection et peuvent contribuer au passage d'une analyse de site à une analyse régionale.*

Notre travail est essentiellement de nature exploratoire, l'enjeu central étant d'estimer l'apport spécifique du modèle PROSAIL et des indices de biodiversité dérivés de *Sentinel-2*, par rapport à d'autres méthodes de caractérisation du couvert ligneux qui peuvent être employées à partir de données de télédétection. A l'heure actuelle, différents produits issus de données de télédétection sont utilisés par les écologues, comme les modèles numériques de hauteur de canopée (MNC) issus d'acquisition LiDAR aéroportées (*Airborne Laser Scanning*, ALS) et les indices de texture calculés à partir de données de très

haute résolution spatiale (THRS), en s'appuyant par exemple sur la méthode FOTO, qui permet d'étudier la texture de la canopée par une analyse en transformée de Fourier et de relier les motifs fréquentiels à des stades de développement ou des niveaux de biomasse forestière. Bien que leur utilité ait été démontrée, la disponibilité de telles données (ALS et THRS) est limitée à des zones de quelques centaines à des milliers de kilomètres carrés, les analyses à échelle régionale sont donc impossibles. Dans le cadre de ce travail, nous cherchons à évaluer la cohérence de nos différentes variables dérivées de l'optique et le niveau de précision qu'elles apportent. Ne disposant pas d'un véritable jeu de données de validation, notre stratégie méthodologique consistera à confronter nos résultats à la connaissance de la zone d'étude dont nous disposons par diverses sources et à les comparer aux données texturales précédemment acquises sur cette zone. Nous espérons pouvoir ouvrir la voie à de nouvelles perspectives d'utilisation de *Sentinel-2* et plus généralement de l'imagerie optique à haute résolution dans le domaine de la cartographie du couvert ligneux tropical.

Matériel & Méthodes

I. Diagramme de Gantt

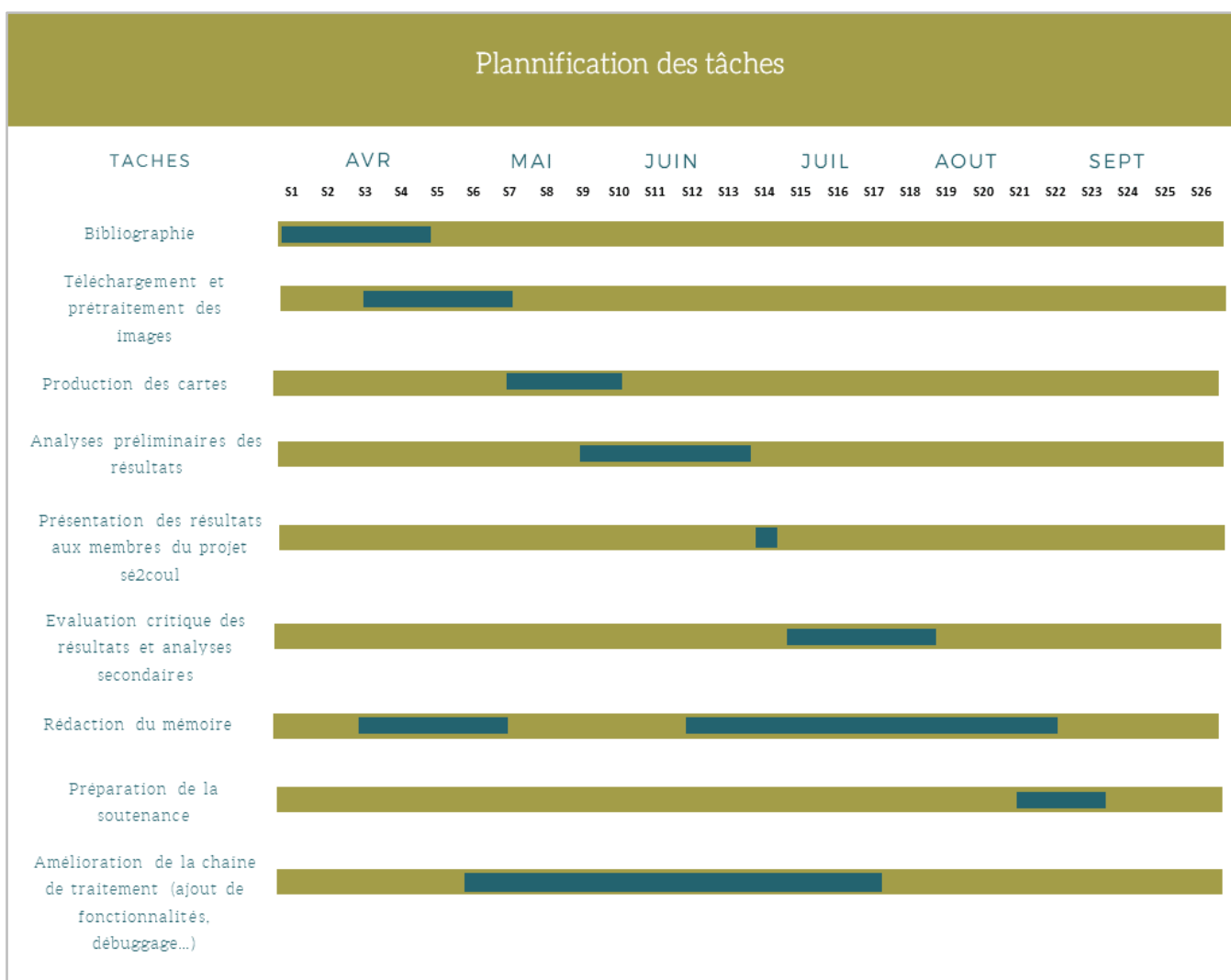


Figure 8. Diagramme de Gantt détaillant la répartition temporelle des tâches prévues au cours du stage.

II. Présentation de la zone d'étude

Notre travail s'est concentré sur une zone d'étude située au nord-ouest de la Guyane Française (**Figure 9**), centrée autour de la localisation 5°05' N, 54°W, dans le contexte de la région géologique sud-américaine du Bouclier Guyanais. La zone s'étend sur environ 340 000 ha, elle contient plusieurs sites suivis par l'UMR AMAP et l'ONF dans le cadre de différents projets de recherche, notamment le projet « Habitats » piloté par l'ONF. Nous bénéficions sur cette zone d'étude, à travers de précédents travaux (Couteron *et al.*, 2005; Guitet *et al.*, 2015) et des études en cours, de données de texture et d'une connaissance du terrain qui constitueront une source d'information complémentaire à notre approche.

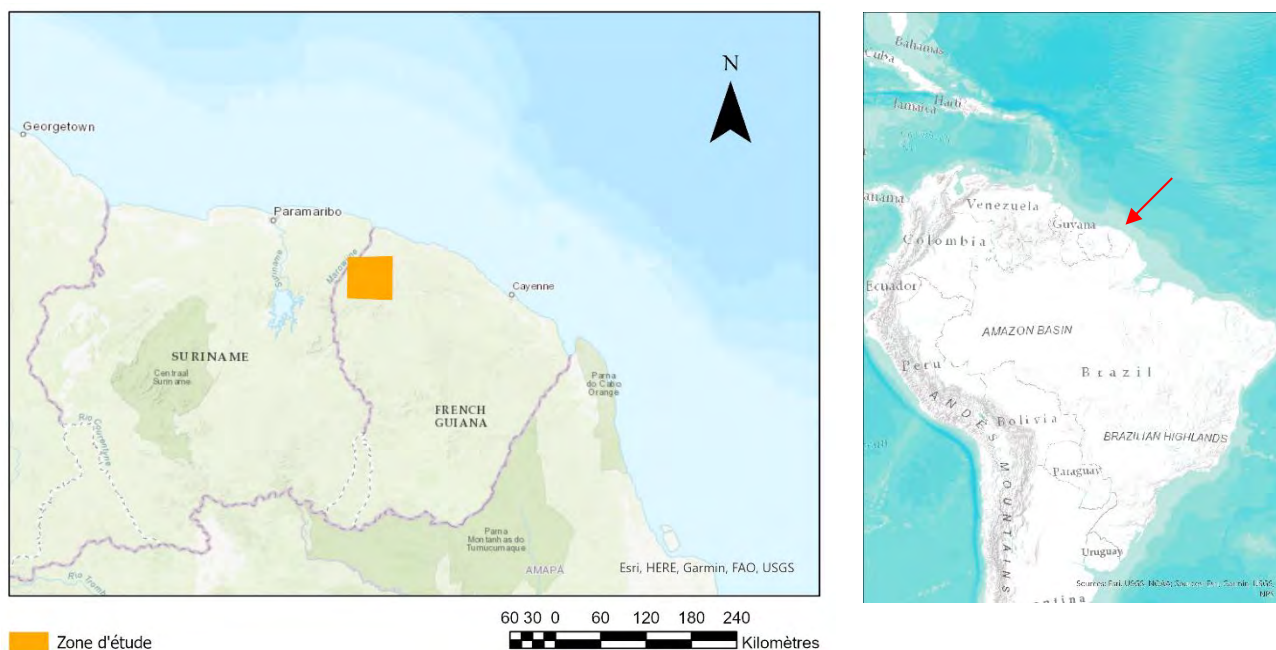


Figure 9. Zone d'étude au Nord-Ouest de la Guyane Française. Sources : IRD, ESRI, HERE, Garmin, FAO, NOAA, NPS.

Le climat est de type équatorial avec une saison plus sèche entre juillet et novembre. Les températures moyennes demeurent élevées (25,5°C) et s'accompagnent de précipitations importantes (entre 2500 et 3000 mm par an). La zone possède un taux d'enneuagement important, quoique plus faible en comparaison au Nord-Est de la Guyane. Elle est dominée par la forêt tropicale humide. La végétation est dense, avec un couvert « fermé » au feuillage persistant ou semi-persistant, et une hauteur de canopée comprise entre 25 et 50 m. Pour cette zone, l'ensemble des cartes produites ont été projetées en UTM 22 Nord dans le système de coordonnées géographiques de référence WGS84 (EPSG 32622).

III. Sélection et acquisition des images

a. Sélection des images via Sentinel Hub

Dans le but d'identifier les images Sentinel-2 disponibles avec un taux d'enneuagement acceptable, nous avons eu recours à la plateforme en ligne *Sentinel Hub*¹¹. La plateforme permet de visualiser facilement, sans téléchargement ni traitement de données, des images Sentinel, MODIS, Landsat ou ENVISAT en appliquant des critères de filtrage comme la période d'intérêt (*e.g.* 1er juillet - 31 octobre 2020) et le taux d'enneuagement maximum autorisé (entre 0 et 100%). Via son application *EO Browser*, *Sentinel Hub* permet de choisir entre différentes compositions colorées (soit prédéfinies, *e.g.* « Fausses couleurs », « Vraies couleurs », ou d'après un ordre de bandes défini par l'utilisateur), d'appliquer des Indices de végétation et de calculer des éléments de statistiques descriptives bande par bande ou sur l'indice sélectionné. Le calcul du taux d'enneuagement proposé par *Sentinel Hub* est effectué

¹¹ <https://www.sentinel-hub.com/>

à l'aide du package *s2cloudless*¹² développé par l'entreprise, et correspond à un taux d'ennuagement à l'échelle de la tuile.

Sentinel Hub s'est avéré être un outil puissant pour visualiser les images *Sentinel-2* et identifier celles qui étaient *a priori* exploitables dans le cadre de la production de nos indices et variables. Sur la tuile 21NZF, englobant notre zone d'étude, nous avons filtré les images *Sentinel-2* de niveau 1C avec un taux d'ennuagement maximal de 30% depuis les premières acquisitions de *Sentinel-2A* en juin 2015, jusqu'à la fin de l'année 2020. Parmi les résultats de notre recherche, nous avons ensuite sélectionné l'image à la date présentant le moins de nuages sur notre zone d'étude.

Après un examen visuel approfondi, sur 111 résultats obtenus avec notre critère de recherche, nous n'avons identifié que six images *Sentinel-2* potentiellement exploitables du fait de leur ennuagement modéré (**Tableau 2**). Le taux d'ennuagement utilisé pour présélectionner les acquisitions étant basé sur des valeurs à l'échelle de la tuile, son efficacité pour identifier les zones sans nuages couvrant une fraction de la surface totale d'une tuile de surface est parfois limitée. La sélection de données *Sentinel-2* via *Sentinel Hub* dans le contexte de zones fortement ennuagées nécessite donc un examen visuel complémentaire, mais permet de gagner un temps non négligeable.

Tableau 2. Dates sélectionnées d'après le critère d'ennuagement. Pour ces six dates, l'image *Sentinel-2* satisfait le critère d'ennuagement ($\leq 30\%$) et une répartition acceptable des nuages (examen visuel) pour la tuile 21NZF.

ANNEE	JOURS / MOIS	ENNUAGEMENT (%)
2017	23/08	0.3
2018	23/08	0.1
2018	07/09	0.7
2020	13/07	1.9
2020	11/09	0.1
2020	16/10	0.1

Parmi les six dates sélectionnées, nous avons choisi de travailler sur l'image acquise le 16 octobre 2020, qui présente la couverture nuageuse la plus faible, avec la présence d'un nombre réduit de nuages isolés sur notre zone d'étude. Un examen approfondi sur l'ensemble des bandes spectrales a cependant permis d'identifier un artefact dû au capteur sur la bande 11 de cette image. Afin de limiter l'effet de ces artefacts sur certains produits calculés à partir des images *Sentinel-2*, une seconde image acquise le 23 août 2018 (**Annexe 1**) fut aussi sélectionnée pour être utilisée spécifiquement pour les produits impactés par cet artefact.

b. Téléchargement des images et correction atmosphérique

Nous avons initialement prévu d'utiliser le package R *sen2r*¹³ qui permet de télécharger les images *Sentinel-2* (**Tableau 3**) depuis la plateforme en ligne *Copernicus Open Access Hub*¹⁴. Cette plateforme en ligne mise en place par l'ESA propose notamment les images *Sentinel-2* au niveau 1C ou *Top of Atmosphere* (TOA), c'est-à-dire non corrigées des effets atmosphériques. Pour l'Europe continentale et certaines régions régulièrement téléchargées, des images au niveau 2A ou *Bottom of Atmosphere* (BOA), corrigées des effets atmosphériques à l'aide du logiciel de correction atmosphérique *Sen2Cor*¹⁵, sont aussi disponibles au téléchargement. Concernant les images acquises en Guyane Française, seules les données au niveau 1C sont disponibles. Dans ce cas, le package *sen2r* permet aussi d'appliquer les corrections atmosphériques à l'aide de *Sen2Cor*.

¹² <https://github.com/sentinel-hub/sentinel2-cloud-detector>

¹³ <https://github.com/ranghetti/sen2r>

¹⁴ <https://scihub.copernicus.eu/>

¹⁵ <https://github.com/alekfal/Auto-Sen2Cor>

Pour notre étude, nous avons décidé d'utiliser *LaSRC* (*Landsat Surface Reflectance Code*), une méthode de correction alternative à *Sen2Cor* développée par la NASA. La méthode, initialement développée pour Landsat, a fait l'objet d'une adaptation récente pour les images Sentinel-2 (Vermote *et al.*, 2018). Ce choix a été fait en raison de résultats préliminaires obtenus à l'UMR TETIS montrant une stabilité des données de réflectance 2A supérieure. Dans la chaîne de traitement finale, il est prévu de laisser à l'utilisateur la possibilité de traiter des images provenant de différentes sources : *Sen2Cor/sen2r*, Théia¹⁶ ou encore *LaSRC*. Pour appliquer la méthode *LaSRC*, nous avons bénéficié d'un programme développé par Florian de Boissieu (UMR TETIS, communication personnelle). Après téléchargement, nous obtenons 13 images à 10 m de résolution correspondant aux bandes individuelles de l'image Sentinel-2 d'origine (visualisée dans *Sentinel Hub*) avec correction atmosphérique (niveau 2A) et un masque de « non-végétation » (nuages, ombres, eau, éléments urbains) au format TIF. Les bandes initialement à 20 m de résolution ont été rééchantillonnées à 10 m, les bandes à 60 m n'ont pas été utilisées.

Tableau 3. Bandes spectrales mesurées par Sentinel-2A et Sentinel-2B. Tableau extrait de <https://sentinel2.cnes.fr/fr/instruments-2> (modifié).

Résolution spatiale	Domaine spectral	Bande	Satellite Sentinel-2A		Satellite Sentinel-2B	
			Longueur d'onde (nm)	Largeur de bande (nm)	Longueur d'onde (nm)	Largeur de bande (nm)
10m	Bleu	2	496.6	98	492.1	98
	Vert	3	560.0	45	559	46
	Rouge	4	664.5	38	665	39
	PIR	8	835.1	145	833	133
20m	Végétation Red Edge	5	703.9	19	703.8	20
	Végétation Red Edge	6	740.2	18	739.1	18
	Végétation Red Edge	7	782.5	28	779.7	28
	PIR « étroit »	8a	864.8	33	864	32
	SWIR	11	1613.7	143	1610.4	141
	SWIR	12	2202.4	242	2185.7	238
60m	Aérosol côtier	1	443.9	27	442.3	45
	Vapeur d'eau	9	945.0	26	943.2	27
	SWIR – Cirrus	10	1373.5	75	1376.9	76

IV. Production des variables biophysiques

Les variables biophysiques ont été obtenues après application d'une chaîne de traitement développée en langage R permettant d'utiliser le package *prosail*, actuellement en développement (auteurs : Jean-Baptiste Féret et Florian de Boissieu). Ce package mobilise le modèle de transfert radiatif de canopée PROSAIL, dans une version qui fait appel au modèle foliaire PROSPECT-D (Féret *et al.*, 2017). La méthode d'inversion hybride de PROSAIL intégrée dans *prosail* permet l'estimation de différentes variables biophysiques à partir des données Sentinel-2. Dans le cadre de notre étude, nous avons calculé puis analysé les produits correspondant aux variables biophysiques suivantes : le Contenu en chlorophylle par unité de surface foliaire (leaf *Chlorophyll content* ou CHL), l'Épaisseur équivalente en eau (*Equivalent Water Thickness* ou EWT), l'Indice de Masse Foliaire (*Leaf Mass per Area* ou LMA) et l'Indice Surfaique Foliaire (*Leaf Area Index* ou LAI).

¹⁶ <https://www.theia-land.fr/>

- $CHL = \text{Leaf Chlorophyll Content} \quad (\mu\text{g} \cdot \text{cm}^{-2})$
- $EWT = \frac{\text{Leaf Dry Mass} \times \text{Leaf Fresh Mass}}{\text{Area}} \quad (\text{g} \cdot \text{cm}^{-2} = \text{cm}^3 \cdot \text{cm}^{-2} = \text{cm})$
- $LMA = \text{Leaf Dry Mass} / \text{Surface} \quad (\text{g} \cdot \text{cm}^{-2})$
- $LAI = \frac{\text{Leaf Area}}{\text{Ground Area}} \quad (\text{m}^2 / \text{m}^2)$

Équation 2. Formules de calcul de l'EWT, LMA et LAI. Le CHL est obtenu par spectrométrie.

La chaîne de traitement comporte trois étapes principales (**Figure 10**) : a) *Extraction et empilement des bandes de l'image Sentinel-2*, b) *Application du filtre radiométrique avec mise à jour du masque* et c) *Cartographie des variables biophysiques par inversion de PROSAIL*.

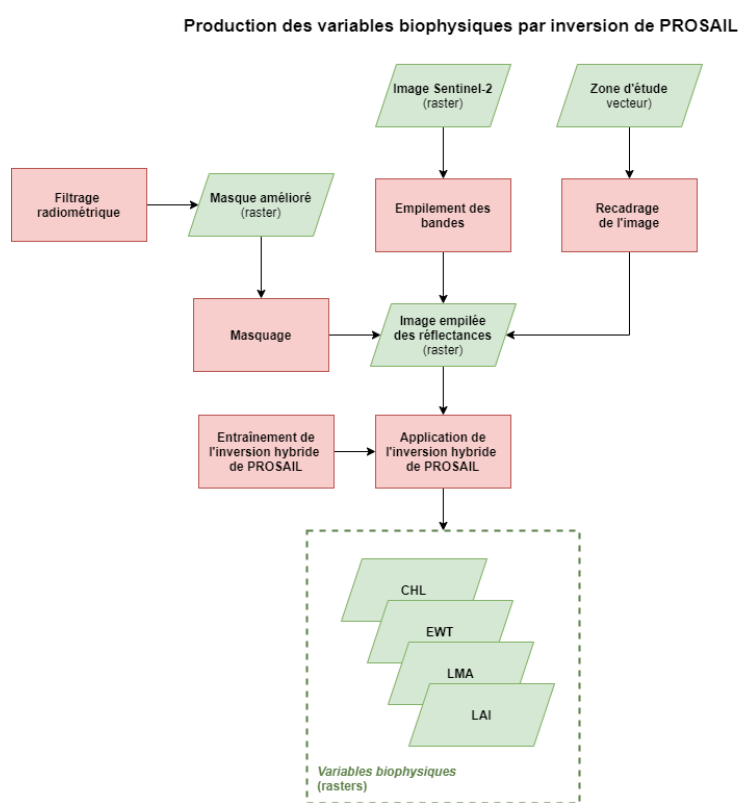


Figure 10. Diagramme de flux résumant les grandes étapes de la chaîne de traitement de production des variables biophysiques. Les phases d'entraînement et application s'appuient sur le package prosail.

a. Extraction et concaténation des bandes.

Les images Sentinel-2 corrigées et téléchargées par LaSRC (à l'exception des bandes 1, 9 et 10 qui ne sont pas utilisées ici, voir **Tableau 3**) ainsi que le masque « non-végétation », ont été recadrés pour correspondre à notre zone d'étude, préalablement définie par une couche vecteur, puis concaténés sous la forme d'une liste de bandes. Dans R, cette liste correspond à un objet « stars » (défini dans le package *stars*¹⁷) soit un type particulier de tableau de données multidimensionnel, spécialement conçu pour les données à caractère spatio-temporel. Certaines fonctions, initialement développées pour traiter des données téléchargées au format SAFE à l'aide du package *sen2r* et corrigées avec *Sen2Cor*, ont dû être adaptées aux données issues de LaSRC.

¹⁷ <https://r-spatial.github.io/stars/articles/stars1.html>

b. Application du filtre radiométrique avec mise à jour du masque.

Malgré le masque produit par LaSRC après une étape de détection de nuages, certaines zones affectées par des perturbations atmosphériques (bordures de nuages, voile nuageux très faible) ou correspondant à des ombres de nuages peuvent subsister. La présence de ces pixels pouvant affecter le calcul des variables, nous avons appliqué un filtre radiométrique complémentaire dans le but de masquer les pixels résiduels d'ombre et de nuages avant d'ajouter ces pixels filtrés au masque de LaSRC. Concrètement, la bande de masque présente dans l'image « empilée » a été convertie en valeurs binaires (« 1 » correspond à un pixel ensoleillé de végétation, « 0 » à un pixel de nuage ou de non-végétation) puis mise à jour en affectant une valeur de 0 aux pixels dont la réflectance dans le Bleu (B2) était supérieure à 5% (nuage suspecté) et inférieure à 20 % dans le proche Infrarouge (ombre suspectée). À l'issue de cette étape, une image des réflectances contenant les 10 bandes spectrales et un masque « amélioré » ont été produits sous la forme de rasters au format ENVI.

c. Cartographie des variables biophysiques par inversion de PROSAIL.

La production des variables biophysiques par inversion hybride de PROSAIL repose sur deux étapes : l'entraînement du modèle sur un jeu d'apprentissage, suivi de son application à l'image entière des réflectances. Le principe de l'inversion hybride consiste ici à déterminer, pour chaque variable d'intérêt, la meilleure combinaison parmi une multiplicité de scénarios reliant des valeurs de réflectances à des propriétés foliaires, déterminées par un modèle de transfert radiatif (PROSAIL), à l'aide de modèles de régression en apprentissage automatique. Pour l'apprentissage, l'algorithme *Support Vector Regression (SVR)* est utilisé. Ces deux étapes ont été réalisées au sein de la chaîne de traitements à l'aide de fonctions en développement du package *prosail*.

L'entraînement du modèle se déroule en trois grandes étapes. Pour commencer, la sélection d'un ensemble de paramètres d'entrée de PROSAIL est effectuée, en fixant les bornes inférieures et supérieures et en suivant une distribution d'échantillonnage aléatoire uniforme. Ces paramètres d'entrée incluent les quatre variables biophysiques d'intérêt {CHL, EWT, LMA, LAI}, ainsi qu'un certain nombre de paramètres additionnels influençant les mesures de réflectance et pour lesquelles la valeur peut être connue (géométrie d'acquisition) ou bien inconnue (distribution angulaire foliaire) (**Tableau 4**). La distribution des valeurs de chaque paramètre est ainsi définie de manière plus ou moins précise en fonction d'une connaissance *a priori*. Chacune de ces combinaisons de paramètres d'entrée est alors utilisée pour produire la réflectance correspondant à l'aide du modèle PROSAIL. Dans notre cas, nous avons simulé 1000 échantillons de réflectance, chaque spectre de réflectance répondant aux caractéristiques du capteur de Sentinel-2 et étant associé à une combinaison {CHL, EWT, LMA, LAI}. Un bruit gaussien est alors appliqué de façon indépendante à chacune de bandes afin de prendre en compte le bruit du capteur et les incertitudes associées aux simulations avec le modèle PROSAIL. Le jeu de variables biophysiques d'intérêt et les données de réflectance correspondantes sont alors utilisées pour réaliser l'apprentissage du modèle de régression par SVR. Ce modèle est ensuite appliqué aux données de réflectance extraites des données Sentinel-2 afin de produire une cartographie des variables estimées au format raster.

Tableau 4. Listes des principaux paramètres d'entrée utilisés pour simuler la réflectance avec PROSAIL.

Paramètres d'entrée	Explication
<i>Parms2Estimate</i>	Variables biophysiques : CHL, EWT, LMA et LAI
<i>Bands2Select</i>	Liste des bandes spectrales à utiliser pour le modèle de régression, pour chacun des paramètres d'entrée
<i>SpecPROSPECT</i>	Paramètres incluant les constantes optiques nécessaires à PROSPECT
<i>SpecSOIL</i>	Paramètres définissant aux propriétés optiques du sol (sec et humide)
<i>SpecATM</i>	Paramètres définissant les effets de l'atmosphère dans des conditions « claires » sur les radiations directes et diffuses
<i>Minval</i>	Liste des valeurs minimales des paramètres échantillonnés
<i>Maxval</i>	Liste des valeurs maximales des paramètres échantillonnés

V. Production des indices de biodiversité α et β

Pour produire les indices de diversité α et β nous avons eu recours au package *biodivMapR* (Féret and de Boissieu, 2019). Le principe de la méthode consiste à déterminer les indices de diversité en s'appuyant sur l'hypothèse de variabilité spectrale (*Spectral Variation Hypothesis*, SVH). Cette hypothèse propose que les variations de l'information spectrale mesurées par satellite puissent être utilisées comme indicateurs de la diversité biologique, autrement dit, que la variabilité spectrale peut être associée à une hétérogénéité d'habitats, une hétérogénéité de conditions écologiques, environnementales, ou une diversité d'espèces, selon l'échelle et la résolution des données utilisées. Le traitement des données d'imagerie optique suivi par *biodivMapR* comporte trois étapes (**Figure 11**) : 1) réduction de dimensionnalité par Analyse en Composantes Principales (ACP) des données de réflectance, après masquage des zones ombragées, ennuagées et non végétalisées et application d'une méthode de normalisation (*continuum removal*) puis sélection des composantes pertinentes par examen visuel, 2) définition des espèces spectrales à l'échelle du pixel (*k-means clustering* à partir des composantes sélectionnées) et 3) calcul des indices de biodiversité à partir de l'inventaire des espèces spectrales.

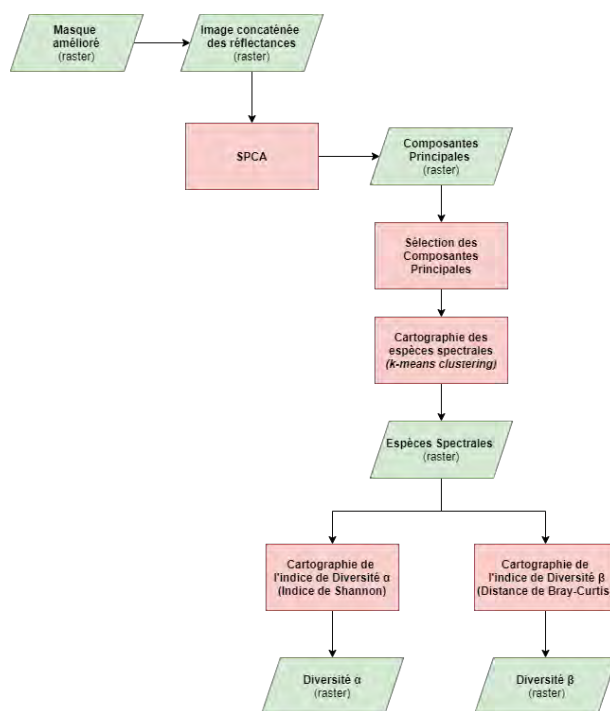


Figure 11. Diagramme de flux des principales étapes de la chaîne de traitement de production des indices de diversité α et β . Les traitements s'appuient sur le package *biodivMapR*.

Préalablement aux traitements, il est nécessaire de déterminer la taille de la fenêtre de pixels qui sera utilisée pour calculer les indices (soit le nombre de pixels de l'image initiale à partir duquel l'inventaire d'espèces spectrales sera réalisé pour calculer les indices de diversité). Les fenêtres comprises entre 0,25 et 4 ha conviennent généralement pour l'analyse de zones correspondant à des forêts, et ces dimensions sont généralement compatibles avec celles des inventaires forestiers. Il convient également que la taille de la fenêtre choisie ne soit pas trop réduite car un nombre de pixels limité diminue la variabilité spectrale et donc la pertinence des indices de diversité. Pour nos analyses, nous avons choisi une fenêtre de 10 x 10 pixels. La résolution initiale étant à 10 m, nous obtenons une résolution finale pour nos indices de 10 000 m² soit 1 ha (Féret and de Boissieu, 2019). En entrée de la chaîne, nous avons utilisé l'image des réflectances multi-bandes précédemment obtenue avant l'application de *prosail*, c'est-à-dire l'image Sentinel-2 empilée contenant 10 bandes de 10 m de résolution spatiale avec correction atmosphérique via LaSRC, ainsi que le masque complété par application d'un filtre radiométrique précédemment mentionné.

a. Analyse en Composantes Principales et sélection des axes

Parmi les méthodes de réduction dimensionnelle de l'information spectrale intégrées dans *biodivMapR*, nous avons eu recours à l'Analyse en Composantes Principales sur les données standardisées (*standardized Principal Component Analysis* ou SPCA). Cette étape a pour but de réduire la dimensionnalité des données initiales, ainsi que le bruit présent dans les données et de sélectionner l'information spectrale susceptible de correspondre à des motifs de biodiversité. À cette étape, la variabilité spectrale des différentes bandes de l'image est résumée sous la forme d'axes synthétiques, chaque bande correspondant à une dimension de l'analyse multivariée. Les axes obtenus résument ainsi un nombre limité de bandes et sont plus facilement interprétables.

Une étape de *Continuum Removal* est ensuite réalisée, dans le but de réduire la variabilité due aux variations de luminosité au sein des houppiers (particulièrement pertinent pour des données de résolution spatiale métrique, moins critique pour les données Sentinel-2), ainsi que les effets de changement de luminosité dus à l'orientation des capteur (Féret and de Boissieu, 2019). Elle consiste à définir une enveloppe convexe à la limite supérieure du spectre de réflectance en tant que continuum de référence (**Figure 12**, voir Huang *et al.*, 2004). Pour chaque pixel de l'image (non-masqué), le spectre de réflectance est ensuite divisée par ce continuum. À l'aide de cette technique (couramment utilisée en spectroscopie) les pics de réflectances sont redéfinis en diminuant la variabilité due aux effets évoqués précédemment.

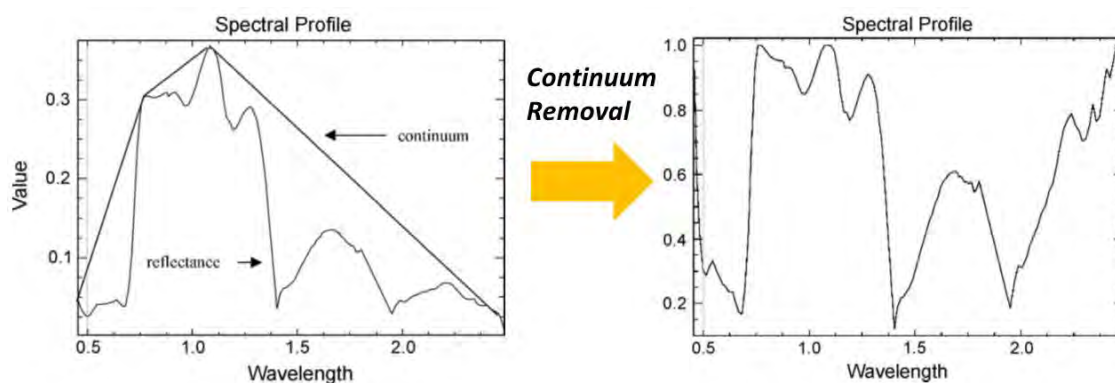


Figure 12. Exemple de Continuum Removal. La figure est extraite de Huang *et al.*, (2004) (modifiée).

Lors de réalisation de la SPCA sur l'image en entrée, *biodivMapR* permet d'appliquer un second filtre basé sur un seuillage des valeurs extrêmes des composantes principales, de manière à éliminer d'éventuels pixels résiduels de non-végétation (par exemple des pixels mixtes en bordure de nuages) qui pourraient « attirer » les axes synthétiques de l'analyse et ainsi brouiller la détermination des motifs de diversité. Après réalisation de la SPCA, les 8 composantes (axes synthétiques) ont été visualisées dans QGis de manière à pouvoir sélectionner « manuellement » celles qui présentent le moins de bruit et présentent des motifs s'apparentant à des motifs associés aux propriétés de la végétation potentiellement pertinents pour estimer les variations spatiales de la biodiversité (l'hypothèse sous-jacente étant que la variabilité spectrale reflète en partie la variabilité écologique, Palmer *et al.*, 2002). Cette étape de sélection permet également de repérer d'éventuels artéfacts présents dans l'image initiale, comme le *Spectral Response Non-Uniformity* affectant potentiellement les données Sentinel-2 (Féret and de Boissieu, 2019) et finalement d'écarter les composantes les plus affectées.

b. Cartographie des espèces spectrales

Les espèces spectrales sont calculées à partir des composantes sélectionnées à l'étape précédente. Elle se base sur la méthode du *k-means clustering*, qui permet de partitionner les données en *k* groupes de valeurs de manière optimisée. Pour commencer, un pool de pixels sélectionnés aléatoirement est séparé en différentes « jeux », chaque jeu correspondant à une répétition. Un *k-means clustering* est ensuite appliqué à chaque jeu de pixels avec un nombre défini de 50 clusters (un nombre raisonnable dans le cas d'une forêt tropicale fortement diversifiée, Féret and de Boissieu, 2019). Chaque

pixel de l'image entière est ensuite affecté à un cluster, sur la base du centroïde du cluster le plus proche, pour chaque jeu de pixels ou répétition. Un ensemble de cartes des espèces spectrales (rasters) est produit à l'issue de cette étape (une par répétition) qui seront utilisées comme base de calcul des indices biodiversité.

c. Cartographie des indices de biodiversité

Les indices de diversité α et β sont déterminés à partir des cartes des espèces spectrales, à l'aide de fonctions de *biodivMapR*. Le calcul est réalisé à partir de l'inventaire des pixels contenus dans chaque fenêtre de 1 ha de l'image (soit au maximum 100 pixels, et moins si des pixels sont masqués). Les valeurs obtenues à partir de chaque répétition d'espèce spectrale (correspondant à des partitionnements différents) sont ensuite moyennées pour obtenir la valeur finale de l'indice spectrale. Pour la calculer la diversité α nous avons choisi de calculer l'indice de Shannon, très utilisé en écologie pour déterminer notamment la richesse spécifique d'un habitat.

$$H' = - \sum_{i=1}^S p_i \log_2 p_i$$

avec

$$p(i) = n_i / N$$

H' : L'indice de diversité de Shannon.

S : La Richesse Spécifique, soit le nombre total d'espèces dans le milieu.

$p(i)$: La proportion d'une espèce i dans le milieu.

n_i : Le nombre d'individus de l'espèce i .

N : L'effectif total comprenant toutes les espèces.

Équation 3. Formule de calcul de l'indice de Shannon utilisée pour déterminer la diversité α .

Appliqué à la carte des espèces spectrales : $S = 50$ (nombre de clusters ou d'espèces spectrales), N le nombre total de pixels, n_i le nombre de pixels de l'espèce spectrale i . Le calcul est répété sur chaque fenêtre spatiale (1 ha) qui correspond au « milieu ».

La diversité β a été déterminée à partir du calcul de la Dissimilarité de Bray-Curtis sur les cartes des espèces spectrales. La carte de diversité β a été obtenue à l'issue de deux étapes successives : *d'abord*, un espace à trois dimensions est obtenu par ordination d'une matrice de dissimilarité issue d'un échantillon de fenêtres spatiales (1 ha), la dissimilarité entre chaque fenêtre de l'image entière et les fenêtres de l'échantillon est alors calculée, *puis* les coordonnées des fenêtres de l'espace à trois dimensions sont déterminées à partir des trois fenêtres les plus similaires utilisées à l'étape précédente, à l'aide de la distance pondérée inverse. Ce processus a été défini de façon à limiter le temps de calcul (Féret and de Boissieu, 2019). La méthode d'ordination utilisée ici est l'Analyse en Coordonnées Principales (*Principal Coordinates Analysis*, ou PCoA).

$$BC_{jk} = 1 - \frac{2 \sum_{i=1}^p \min(N_{ij}, N_{ik})}{\sum_{i=1}^p (N_{ij} + N_{ik})}$$

BC_{jk} : L'indice de dissimilarité de Bray-Curtis.

N_{ik} : L'abondance d'une espèce i dans un échantillon k .

N_{ij} : L'abondance d'une espèce i dans un échantillon j .

$\min(\dots, \dots)$: pour une espèce i présente dans j et k , le compte minimum d'individus entre j et k .

p : Le nombre total d'espèces.

Équation 4. Formule de calcul de la distance de Bray-Curtis utilisée pour déterminer la diversité β .

VI. Photo-interprétation des résultats

Les variables biophysiques et les indices de biodiversité correspondant à l'image la moins ennuagée ont été visualisés à l'aide du logiciel QGIS. Pour chaque variable, nous avons cherché à repérer des contrastes dans la distribution spatiale qui pourraient correspondre à des motifs de végétation (on s'attend à ce que des communautés végétales différentes soient associées à des variables biophysiques et des indices de biodiversité contrastés). L'image de réflectance a également été visualisée en composition colorée et comparée aux résultats, de manière à aider l'interprétation des motifs. Cette phase de photo-interprétation des résultats a également servi à identifier d'éventuels artefacts, soit liés à la méthode, ou provenant d'un biais dans l'acquisition de Sentinel-2.

Pour guider nos interprétations, nous avons bénéficié de l'expertise thématique de Christophe Proisy et Pierre Couteron de l'UMR AMAP, spécialisés en écologie forestière tropicale et connaisseurs des communautés végétales potentiellement présentes sur le site d'étude. Dans le but d'alimenter les discussions et guider nos interprétations, nous avons réalisé une présentation orale de nos résultats pour les collègues des UMRs AMAP et TETIS impliqués dans le projet sé2coul.

VII. Exploration des liens entre les quatre variables biophysiques

a. Extraction des valeurs moyennes et des écart-types

Pour étudier les relations entre les variables biophysiques, la résolution spatiale des images étant trop fine (10 m) et de manière à extraire des métriques plus facilement comparables, nous avons établi sur l'emprise de la zone d'étude une grille de fenêtres de 1 ha et calculé à l'intérieur de chaque fenêtre la valeur moyenne et l'écart-type des pixels. Cette grille a également été réalisée dans le but de faciliter les comparaisons des variables biophysiques avec les données de texture.

b. Recherche des corrélations entre les variables

Pour explorer les liens entre les quatre variables biophysiques, nous avons testé les corrélations pouvant exister entre elles à l'aide du test non-paramétrique de Spearman (plus robuste dans le cas où les données ne suivraient pas une distribution linéaire) avant de visualiser ces corrélations à l'aide d'un corrélogramme. Après identification des corrélations les plus importantes, celles-ci ont été représentées à l'aide de *scatter plots* de manière à identifier de potentielles tendances et les interpréter.

VIII. Confrontation des résultats avec les données de structure

a. Présentation des données de structure de la canopée

Deux types de données de texture de canopée nous ont été fournies pour réaliser notre approche comparative : des indices de finesse de grain issus de la méthode FOTO (*Fourrier based-textural ordination*) et des hauteurs maximales de canopée provenant d'une acquisition LiDAR aéroportée sous la forme d'un Modèle Numérique de Terrain (MNT) et d'un Modèle Numérique de Canopée (MNC).

Les données issues de la méthode FOTO ont été produites dans le cadre du projet BIOMAP (Programme Guyamazon) à partir de données à THRS Spot7 (acquises le 28/10/2015, programme GEOSUD). Les données ALS (MNC et MNT) ont été mises à disposition par l'ONF et ont bénéficié d'une aide financière de l'Europe (Fonds Européen Agricole pour le Développement Rural). Ces données font l'objet d'un article soumis à publication (Forestier *et al.*, 2021, *submitted*).

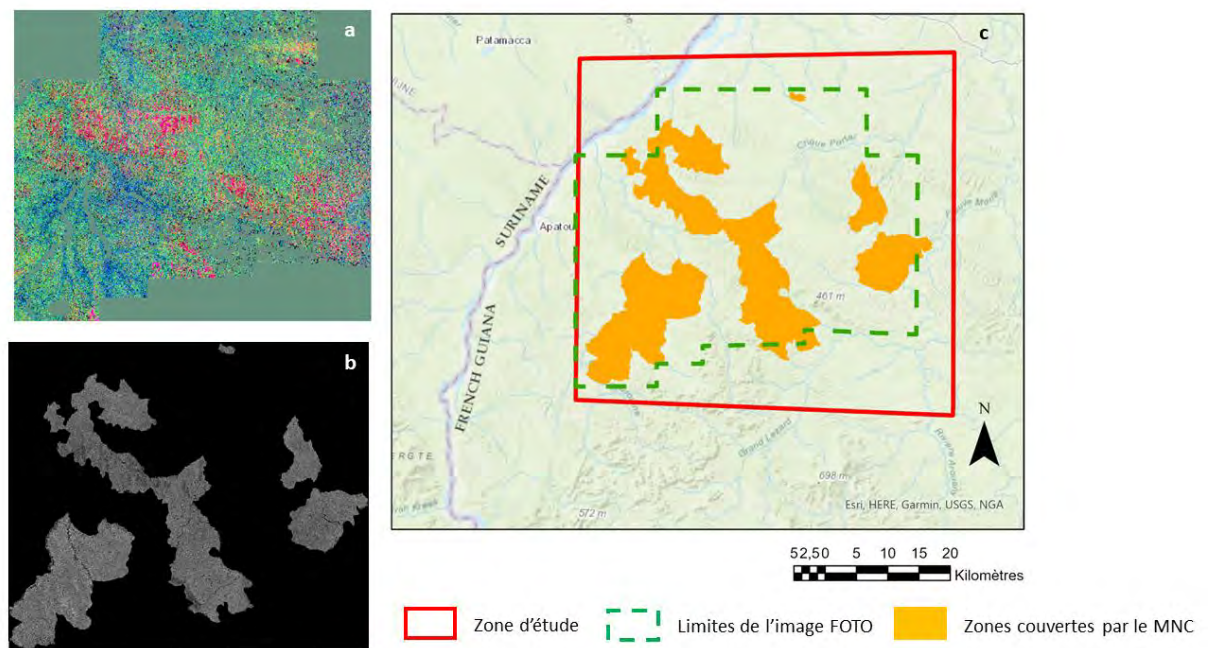


Figure 13. Présentation des données complémentaires de structure dans la zone d'étude. **a.** Image de texture de grain issue de la méthode FOTO. **b.** Modèle Numérique de Canopée. **c.** Emprise des deux images dans la zone d'étude. Sources : IRD, ONF, Esri, HERE, Garmin, USGS, NGA.

b. Étude des relations entre la structure de la canopée estimée par le MNC et les quatre variables biophysiques.

Nous avons cherché à évaluer la cohérence des variables biophysiques à travers une « confrontation » avec le MNC. Cette cohérence est évaluée qualitativement via la caractérisation fonctionnelle de la canopée émergeant de cette confrontation. De manière similaire aux variables biophysiques, nous avons calculé la moyenne et l'écart-type du MNC à l'intérieur d'une grille de fenêtres de 1 ha. Nous avons ensuite comparé visuellement à l'aide de *scatter plots* la distribution de l'écart-type en fonction de la moyenne de chaque variable biophysique avec la distribution équivalente concernant le MNC. L'hypothèse étant que cette distribution (écart-type en fonction de la moyenne du MNC) peut être interprétée en tant que structure de canopée et servir de référence pour évaluer la cohérence des quatre variables biophysiques (Pierre Couteron, communication personnelle).

Nous avons également cherché à identifier les corrélations entre les moyennes et écart-types du MNC et les moyennes et écart-types des variables biophysiques. Dans ce but, nous avons réalisé une Analyse en Composantes Principales (ACP) de manière à pouvoir visualiser ces corrélations et préciser la caractérisation fonctionnelle de la canopée à l'aide des axes synthétiques. La diversité α , également calculée par fenêtre de 1 ha sur une emprise similaire, a été intégrée dans l'ACP de manière à tester si la caractérisation fonctionnelle peut être reliée à des variations de la biodiversité du couvert ligneux (la diversité β a été exclue de l'analyse car elle-même repose sur trois axes synthétiques et ne peut être interprétée comme les autres variables).

c. Recherche des corrélations entre l'estimation de la diversité β et les composantes de l'image FOTO

Afin d'évaluer la cohérence de notre estimation de la diversité β , nous avons cherché à confronter les motifs obtenus avec les variations de l'indice de texture issu de la méthode FOTO. Nous faisons ici l'hypothèse que des communautés d'arbres éloignées (en termes de composition en espèces) auront tendance à présenter une variabilité dans la taille des couronnes d'arbres différente. Nous avons donc cherché à comparer visuellement les motifs de la diversité β avec ceux de l'indice de texture, qui est sensible aux variations de la taille de couronne des arbres (Ploton *et al.*, 2017). Nous avons ensuite testé les corrélations entre les deux ensembles d'axes à l'aide du test de Spearman.

L'image FOTO ayant une résolution de 49,49 m, il ne nous pas été possible de la superposer exactement avec l'image de diversité β , malgré une étape de reprojection dans le même système de coordonnées de référence (WGS84, UTM 22N) et de rééchantillonnage de l'image de diversité β (initialement à 100 m de résolution). De manière à pouvoir extraire des couples de valeurs diversité β - FOTO et tester les corrélations entre les deux images, nous avons procédé à un échantillonnage de points aléatoires desquels nous avons extrait les valeurs pour les deux images. Après suppression des couples de valeurs comprenant une absence de données (« No data ») nous avons extrait et testé la corrélation sur environ 30 000 couples de valeurs diversité β - FOTO.

Résultats

I. Cartographie des variables biophysiques

a. Carte des réflectances en composition colorée « vraies couleurs »

Le taux d'enneuagement sur la zone d'étude est faible (estimé à 0.1% par *Sentinel Hub*) et apparaît sous la forme de nuages isolés, dispersés, de manière relativement homogène sur l'étendue de la zone (**Figure 14**). Le fleuve Maroni est présent au nord-ouest de la zone d'étude, marquant la frontière entre le Suriname à l'ouest, et la Guyane Française à l'est (une route longe le fleuve côté français). Des cours d'eau secondaires sont visibles à l'est et au centre de l'image. La végétation apparaît sombre, dense avec un couvert fermé, sur la quasi-totalité de la zone d'étude.

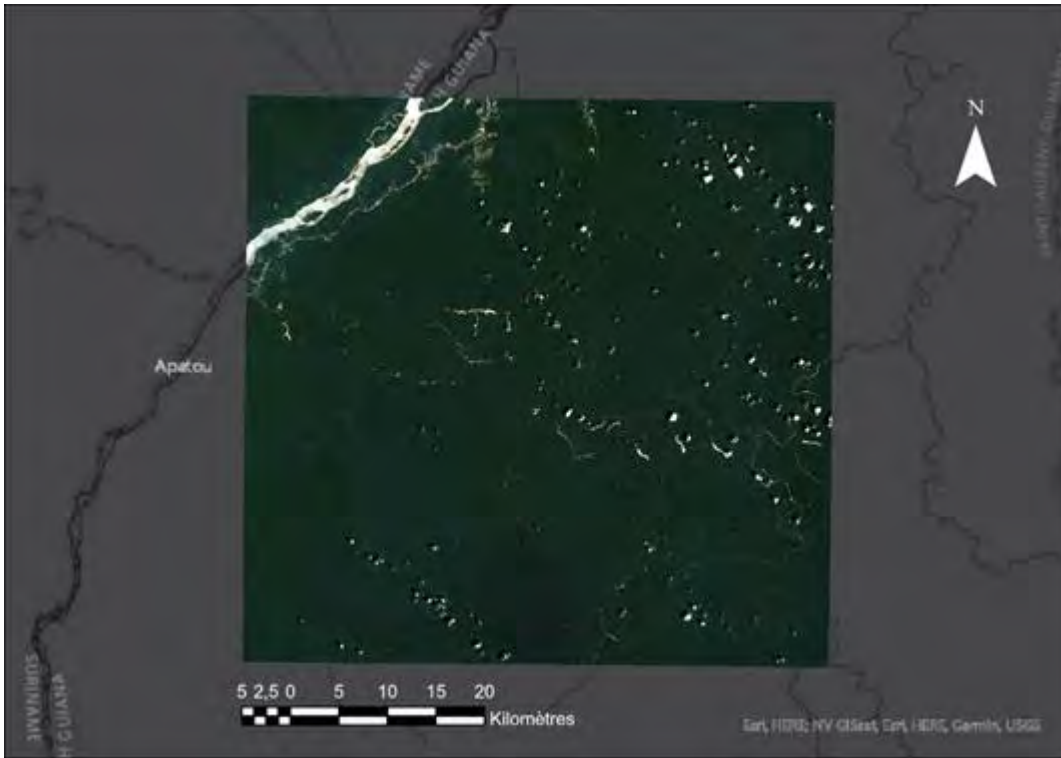


Figure 14. Carte des réflectances de la zone d'étude après correction atmosphérique en composition colorée « Vraies couleurs ». L'image source a été acquise le 16 octobre 2020. Résolution spatiale : 10 m. Sources : ESA, Esri, HERE, Garmin, USGS, NGA.

b. Carte des variables biophysiques : CHL, EWT, LMA et LAI

Les cartes *a* à *d* de la **Figure 15** révèlent des motifs correspondant aux variations de propriétés biophysiques de la végétation (CHL, EWT, LMA, LAI) estimées par inversion du modèle physique PROSAIL. La **Figure 16** présente la distribution des valeurs pour variable biophysique avec les valeurs moyennes sur l'ensemble de la zone d'étude et la variance.

Les valeurs de CHL estimées sont comprises entre 43 et 64 $\mu\text{g}\cdot\text{cm}^{-2}$, avec une moyenne de 53.3 $\mu\text{g}\cdot\text{cm}^{-2}$ les plus élevées ($\geq 60 \mu\text{g}\cdot\text{cm}^{-2}$) se situant à l'ouest de la zone d'étude, notamment autour du fleuve Maroni et au sud du fleuve. La zone située au sud du fleuve Maroni et à l'ouest de la zone d'étude est elle-aussi caractérisée par une teneur en chlorophylle estimée importante, présentant un motif s'apparentant au tracé d'un réseau de cours d'eau, distinguables également sur l'image des réflectances « Vraies couleurs ». Sur le reste de l'image, les valeurs de CHL sont globalement plus faibles ($\leq 60 \mu\text{g}\cdot\text{cm}^{-2}$).

Les valeurs d'EWT sont comprises entre 0.015 cm et 0.022 cm, avec une valeur moyenne de 0.019 cm. Les valeurs les plus fortes d'EWT sont obtenues principalement sur un massif situé au sud de la zone d'étude. Les valeurs apparaissent également plus élevées à l'Est. A l'Ouest, on distingue le même motif de végétation suivant les cours d'eau que pour CHL, mais dans le cas d'EWT, les motifs sont associés à des valeurs plus faibles (autour de 0.016 cm).

Les valeurs estimées de LMA sont distribuées entre 0.015 et 0.025 $\text{g}\cdot\text{cm}^{-2}$ avec une valeur moyenne de 0.021 $\text{g}\cdot\text{cm}^{-2}$. Les faibles valeurs de LMA apparaissent principalement dans les zones interprétées visuellement comme correspondant aux réseaux de cours d'eau, comme ce qui est observé pour l'EWT, mais avec un contraste encore plus important. Les valeurs estimées de LMA apparaissent plus élevées en s'éloignant des réseaux de cours d'eau.

Les valeurs estimées de LAI sont comprises entre 3.4 et 4.5, ce qui apparait relativement modéré et assez peu variable par rapport aux valeurs rapportées dans la littérature pour les forêts tropicales denses, avec des valeurs de LAI pouvant atteindre 8 à 10, et des contrastes importants entre les types forestiers. La valeur moyenne du LAI est de 4.13.

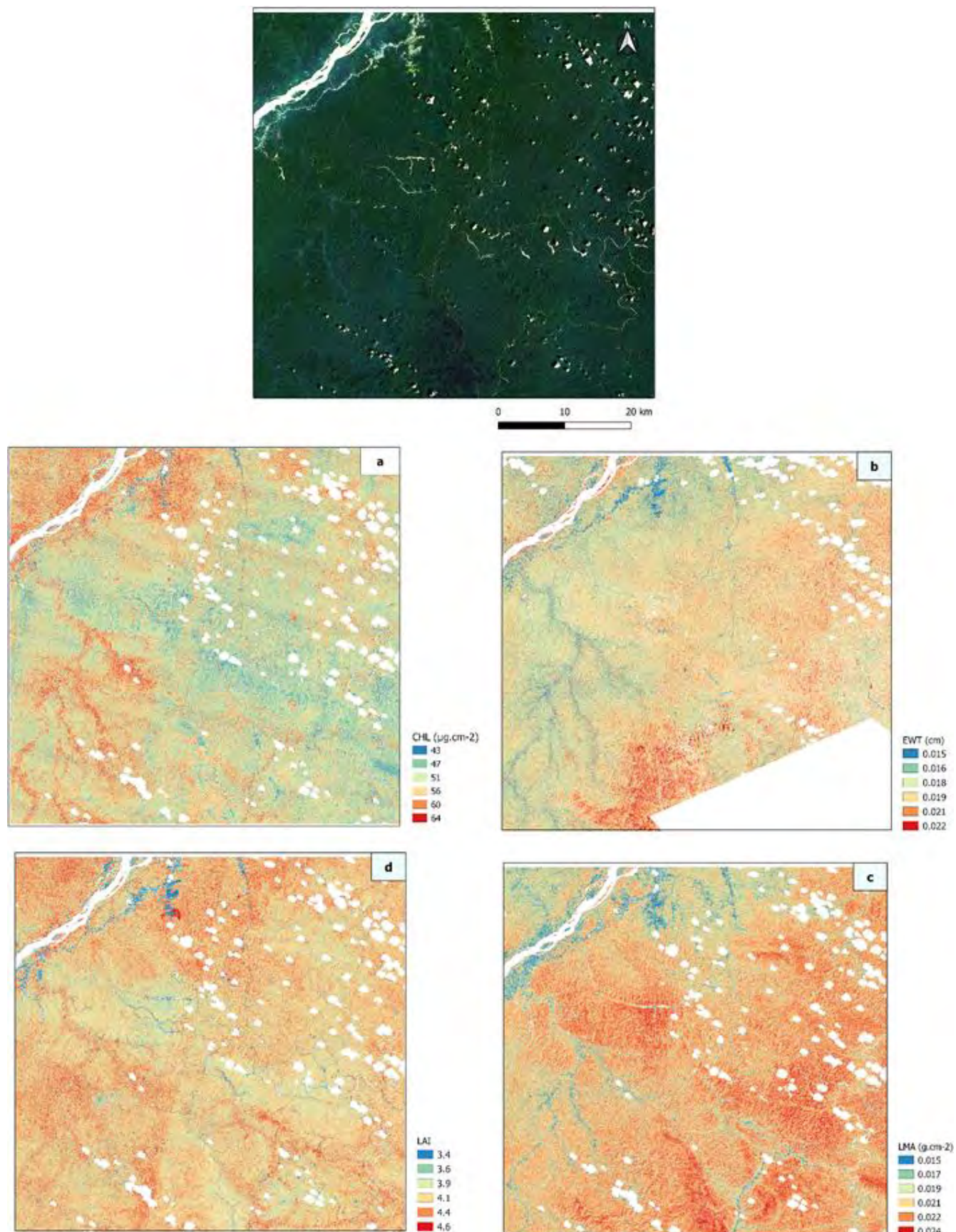


Figure 15. Cartographie des quatre variables biophysiques couvrant la zone d'étude. En haut à gauche l'image des réflectances en composition colorée « Vraies couleurs » puis **a.** Le contenu en chlorophylle foliaire (CHL), **b.** l'Épaisseur équivalente en eau (EWT), **c.** L'Indice de masse foliaire (LMA), **d.** L'Indice surfacique foliaire (LAI). La date d'acquisition de l'image source Sentinel-2 date du 16 octobre 2020 à l'exception de l'EWT, acquise le 23 août 2018. Résolution spatiale : 10 m. Source : ESA.

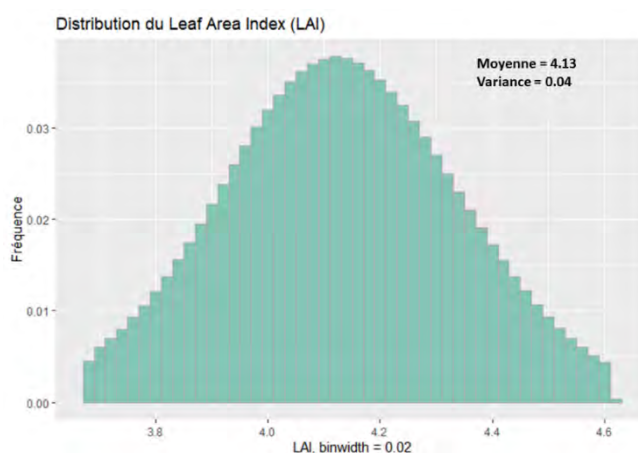
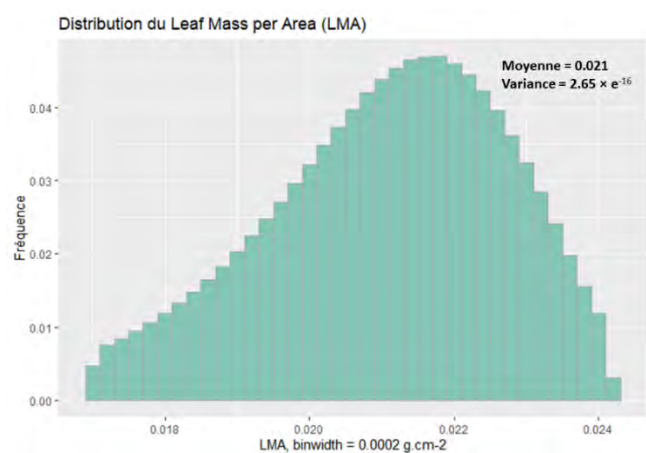
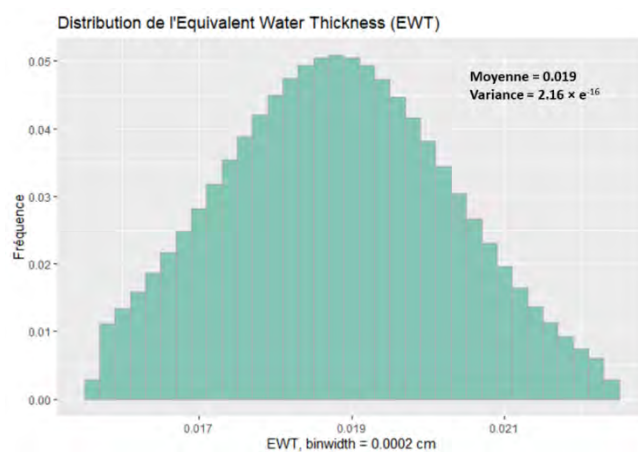
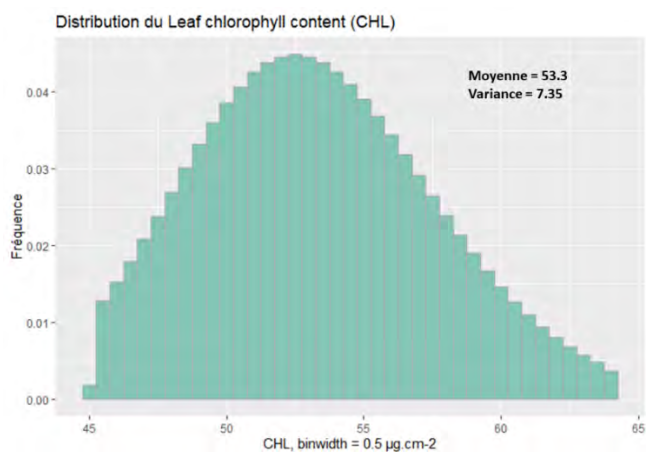


Figure 16. Distributions des valeurs de pixels des quatre variables biophysiques. Le CHL, LMA et LAI à la date du 16 octobre 2020, ainsi que l'EWT à la date du 23 août 2018, sur l'ensemble de la zone d'étude (distributions entre les percentiles 0.05 et 0.98, après exclusion des valeurs extrêmes).

II. Cartographie des indices de biodiversité

a. Sélection des composantes principales

Après application de la SPCA, huit composantes sont obtenues à partir de l'image Sentinel-2 (l'application du *Continuum Removal* réduisant la dimensionnalité de l'image initiale de dix bandes à huit bandes). La **Figure 17** présente ces huit composantes. La sélection des composantes consiste à éliminer les composantes présentant des artefacts potentiellement indésirables pour la cartographie des indices de biodiversité, et à ne sélectionner que les composantes montrant des motifs pouvant être identifiés comme associés à des motifs de types de végétation. Dans notre étude, nous avons sélectionné les composantes principales 1 et 2 pour produire la carte des espèces spectrales (**Figure 17**). La première composante principale (PC1) apparaît peu affectée par d'éventuels artefacts et révèle des motifs qui peuvent être associés à des variations de l'environnement et potentiellement du couvert ligneux. La deuxième composante principale (PC2), apparaît comme la moins affectée par les artefacts parmi les composantes restantes.

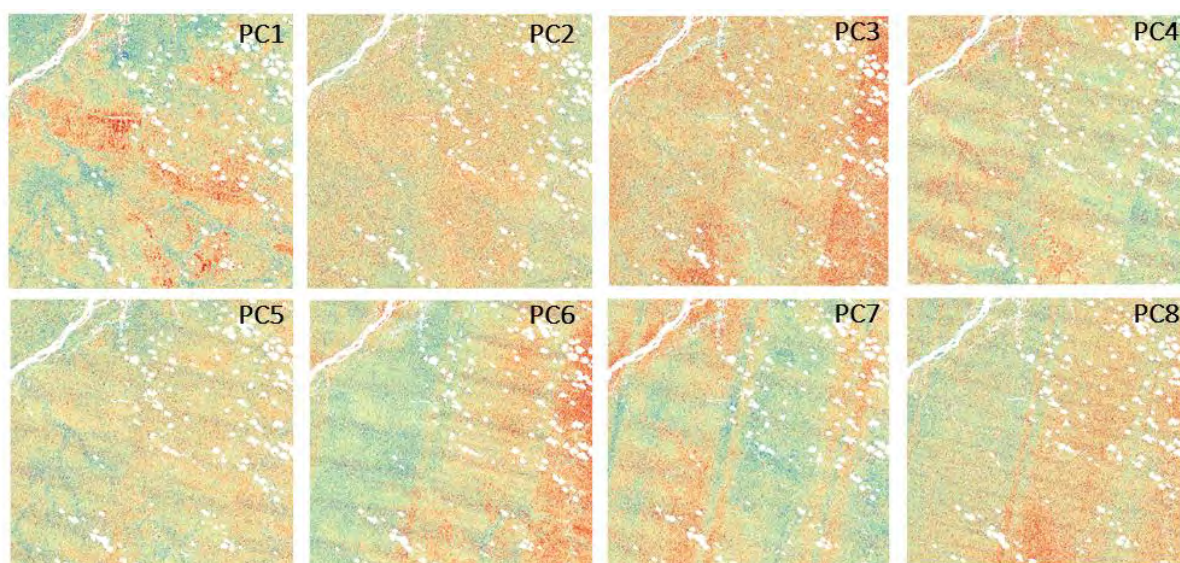


Figure 17. Composantes principales issues de la SPCA. Nous observons la présence d'artefacts liées à l'acquisition de Sentinel-2 de PC2 à PC8, les artefacts les moins importants étant sur PC2. Les artefacts apparaissent sous la forme de stries verticales (PC6, 7 et 8), horizontales (PC5, 6, 4, 7 et 8) ainsi que sous la forme d'une bande large à droite de l'image (PC2, 3, 4 et 6). La première composante semble exempte d'artefact majeur et présente des motifs qui peuvent être associés à des variations de l'environnement.

b. Indices de diversité α et β

La **Figure 18** présente les cartes d'indices de diversité calculés à partir des données Sentinel-2 sur la zone d'étude. La **Figure 18-a** présente la carte de l'indice de Shannon correspondant à l'indice de diversité α , et la **Figure 18-b** présente la carte correspondant au calcul de la PCoA issue de la matrice de dissimilarité de Bray-Curtis, utilisée pour illustrer la diversité β spatialement. L'indice de Shannon calculé sur la zone d'étude varie entre 2.3 et 3.2, avec une valeur moyenne de 2.9 (**Figure 19**). Les valeurs d'indice de diversité calculés à partir du package *biodivMapR* ne peuvent être directement comparées aux indices de diversité obtenus à partir d'inventaires terrain, et la gamme de valeurs obtenue avec *biodivMapR* est directement dépendante du nombre d'espèces spectrales utilisé lors du calcul de ces indices de diversité. Sur la **Figure 18-a**, on observe le contraste entre des zones de faible diversité α (2.3-2.5) et des zones de plus forte diversité α (3.0-3.2). Les motifs de faible diversité α apparaissent aussi sur la carte de diversité β (**Figure 18-b**, zones de couleur verte). La végétation à proximité du réseau hydrographique correspond à des zones de teinte rose pâle apparaissant sur la carte de diversité β . Ces résultats suggèrent l'existence d'une variation de la diversité α et β du couvert ligneux à proximité du réseau hydrographique, avec un type de communauté végétale caractérisé par une plus faible diversité α , possédant une communauté d'espèces très différente de celle dominant le reste de la zone d'étude (diversité β).

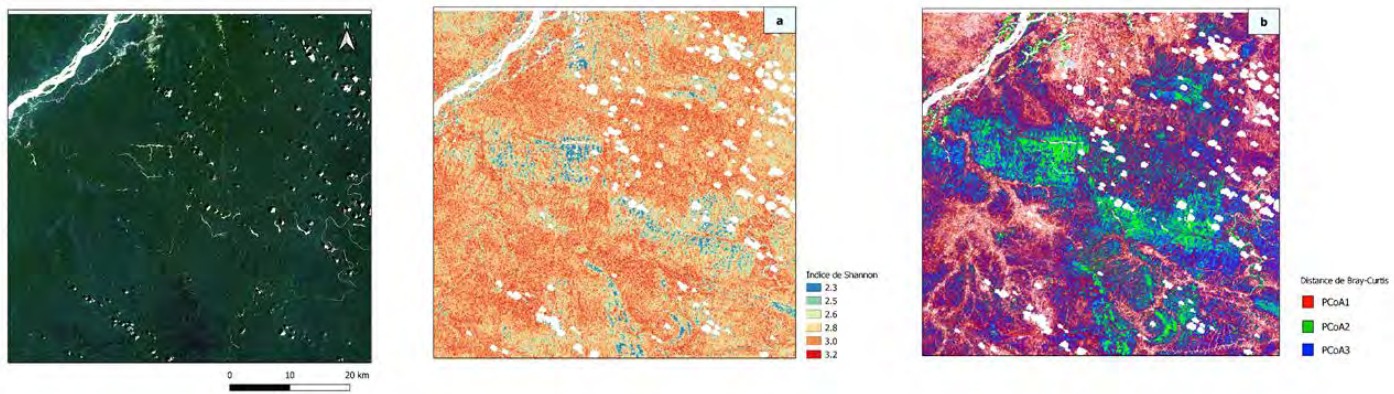


Figure 18. Cartographie des indices de biodiversité couvrant la zone d'étude. En haut à gauche l'image des réflectances en composition colorée « Vraies couleurs » puis **a.** La diversité α estimée par l'indice de Shannon, **b.** La diversité β représentée à travers les trois composantes de la PCoA basée sur la Distance de Bray-Curtis. Résolution spatiale : 100 m. Source : ESA.

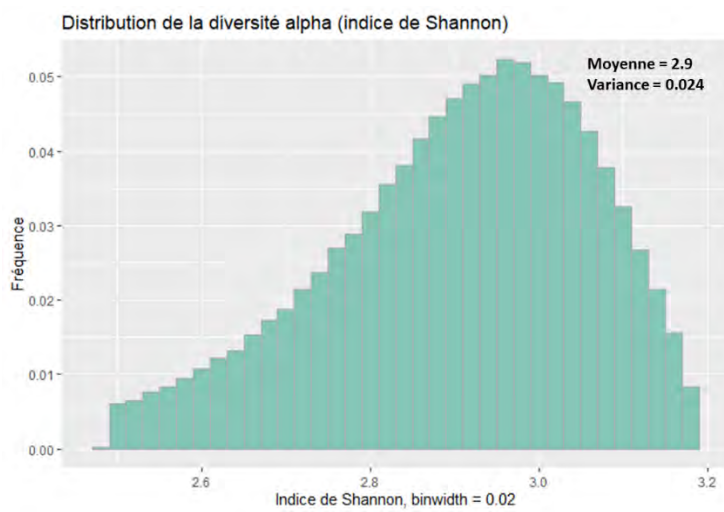


Figure 19. Distributions des pixels de diversité α (indice de Shannon). Résolution des pixels : 100 m.

III. Motifs de variations biophysique et écologique de la végétation

Nous avons effectué un travail de photo-interprétation en étudiant les motifs spatiaux des variables biophysiques et des indices de biodiversité, et en cherchant à les relier à des éléments caractéristiques du paysage, observables notamment sur l'image Sentinel-2 à partir d'une visualisation en « Vraies couleurs ». Nous nous sommes également appuyés sur le MNT (**Figure 19**). Cette photo-interprétation a permis d'identifier trois types de végétation, définis de manière *ad hoc*, sans faire d'hypothèses préalables sur la nature de la végétation qui pouvait être observée dans la zone d'étude, à partir de ses caractéristiques biophysiques ou écologiques. Nous avons identifié des motifs associés aux variations du relief, interprétés comme des motifs permettant de différencier la végétation de plaines inondables et les zones humides de la végétation de terre ferme. Ces motifs se révélaient sous la forme de contrastes visuels relativement marqués (à diverses échelles) considérant les variations de certaines des variables biophysique ou des indices de biodiversité (notamment la diversité β). Ne pouvant ici présenter de manière exhaustive les motifs observés sur la zone d'étude, nous avons choisi de donner ici un exemple pour chaque type de motif identifié.

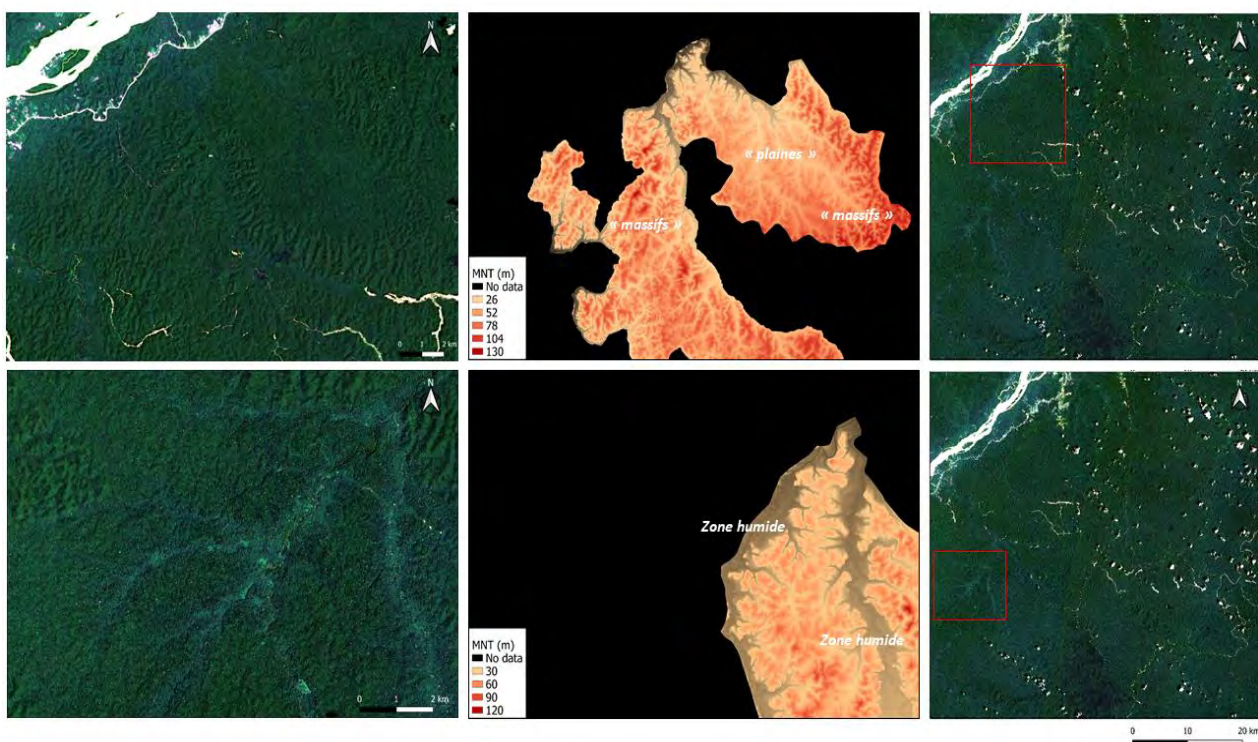


Figure 20. Exemple d'utilisation du MNT pour appuyer la photo-interprétation. En haut : Exemple d'une zone présentant des variations relativement importantes du relief (point culminant à 130 m et point le plus bas à environ 6 m) au nord-ouest de la zone d'étude. En bas : Exemple d'une zone humide au sud-ouest de la zone d'étude. Résolution spatiale : 10 m (composition colorée « vraies couleurs ») et 5 m (MNT). Sources : ESA, ONF.

a. Motifs associés à la variation du relief

Les motifs illustrant la **Figure 20-a**, des motifs de valeurs « plus faibles » du CHL (par rapport à la valeur moyenne sur la zone d'étude, de $53 \mu\text{g}\cdot\text{cm}^{-2}$) apparaissent sur les massifs ($43\text{-}51 \mu\text{g}\cdot\text{cm}^{-2}$) et plus élevées dans les plaines (autour de $60 \mu\text{g}\cdot\text{cm}^{-2}$). Concernant l'EWT (date du 23 août 2018) (**Figure 20-b**) des motifs de valeurs globalement plus élevées apparaissent sur les massifs (autour de $0.019\text{-}0.021 \text{ cm}$) par rapport aux plaines ($0.015\text{-}0.019 \text{ cm}$; moyenne = 0.019 cm). Des motifs de valeurs les plus élevées du LMA (autour de $0.024 \text{ g}\cdot\text{cm}^{-2}$; moyenne = $0.021 \text{ g}\cdot\text{cm}^{-2}$) sont associés de manière étroite aux plis du relief (**Figure 20-c**). Le LMA est également élevé dans les plaines ($0.021\text{-}0.022 \text{ g}\cdot\text{cm}^{-2}$). Des motifs de valeurs plus élevées de LAI apparaissent dans les plaines ($4.1\text{-}4.6$; moyenne = 4.1) (**Figure 20-d**). Sur les massifs, les valeurs sont plus faibles ($3.9\text{-}4.1$).

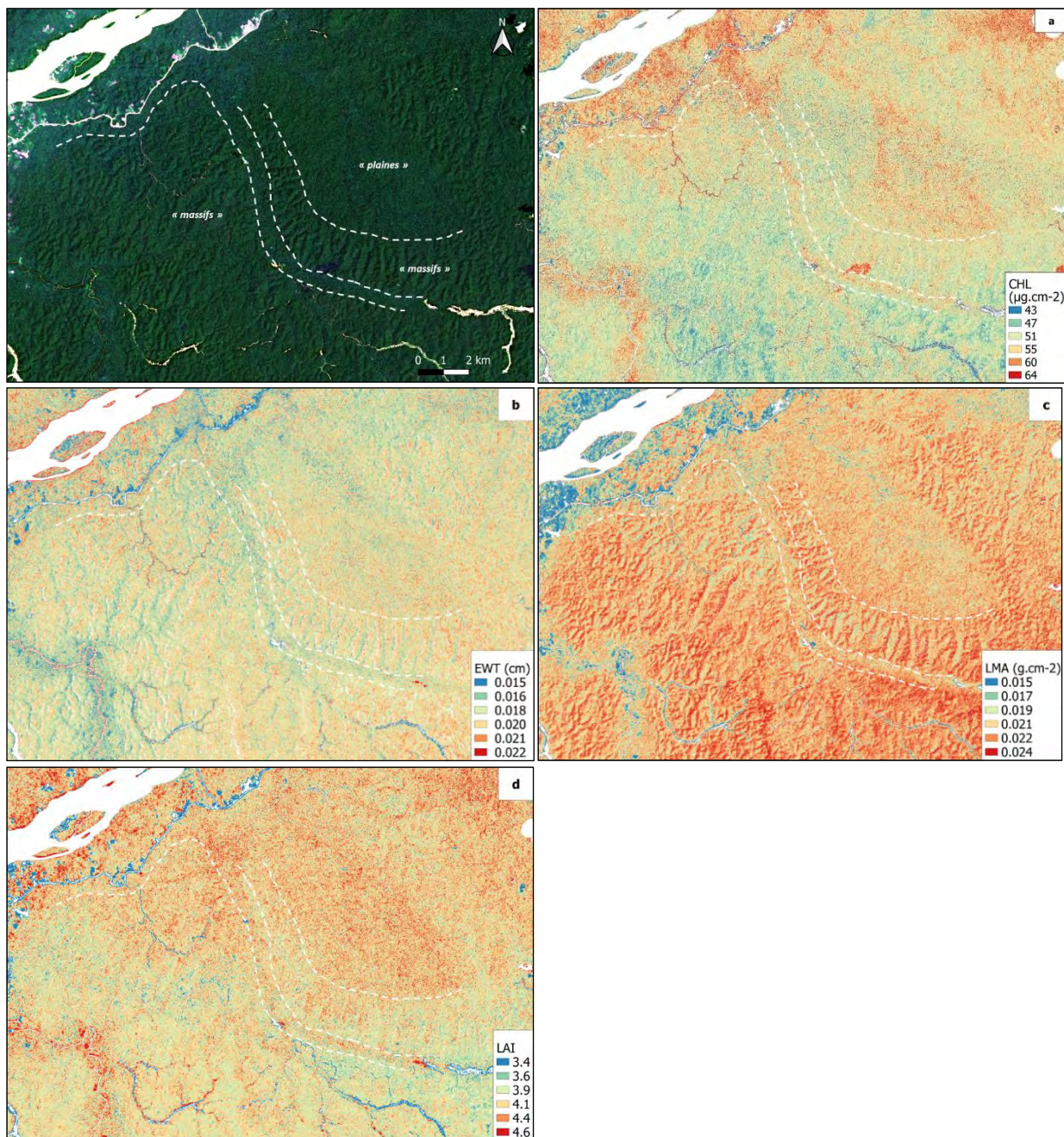


Figure 21. Exemple de motifs des quatre variables biophysiques associés aux variations du relief. La zone est située au Nord-Ouest de la zone d'étude. En haut à gauche l'image des réflectances en composition colorée « Vraies couleurs ». Les traits pointillés indiquent des séparations entre plaines et massifs. **a.** Contenu en chlorophylle de la canopée ou CHL ($\mu\text{g}\cdot\text{cm}^{-2}$). **b.** Epaisseur équivalente en eau ou EWT (cm) (23/08/2018). **c.** Indice de masse foliaire ou LMA ($\text{g}\cdot\text{cm}^{-2}$). **d.** Indice de surface foliaire ou LAI (sans dimensions). Résolution spatiale : 10 m. Source : ESA

Sur la **Figure 22-a**, la diversité α estimée par l'indice de Shannon révèle des motifs de valeurs plus faibles sur une partie des massifs (2.3-2.5) et plus élevées dans les plaines (3.0-3.2 ; moyenne = 2.9). La diversité α apparaît également plus élevée dans les plaines par rapport aux massifs. Concernant la diversité β (**Figure 22-b**), on observe que des contrastes importants apparaissent entre une partie des massifs (vert) et les plaines (rose pâle) suggérant des différences dans la composition de la communauté végétale.

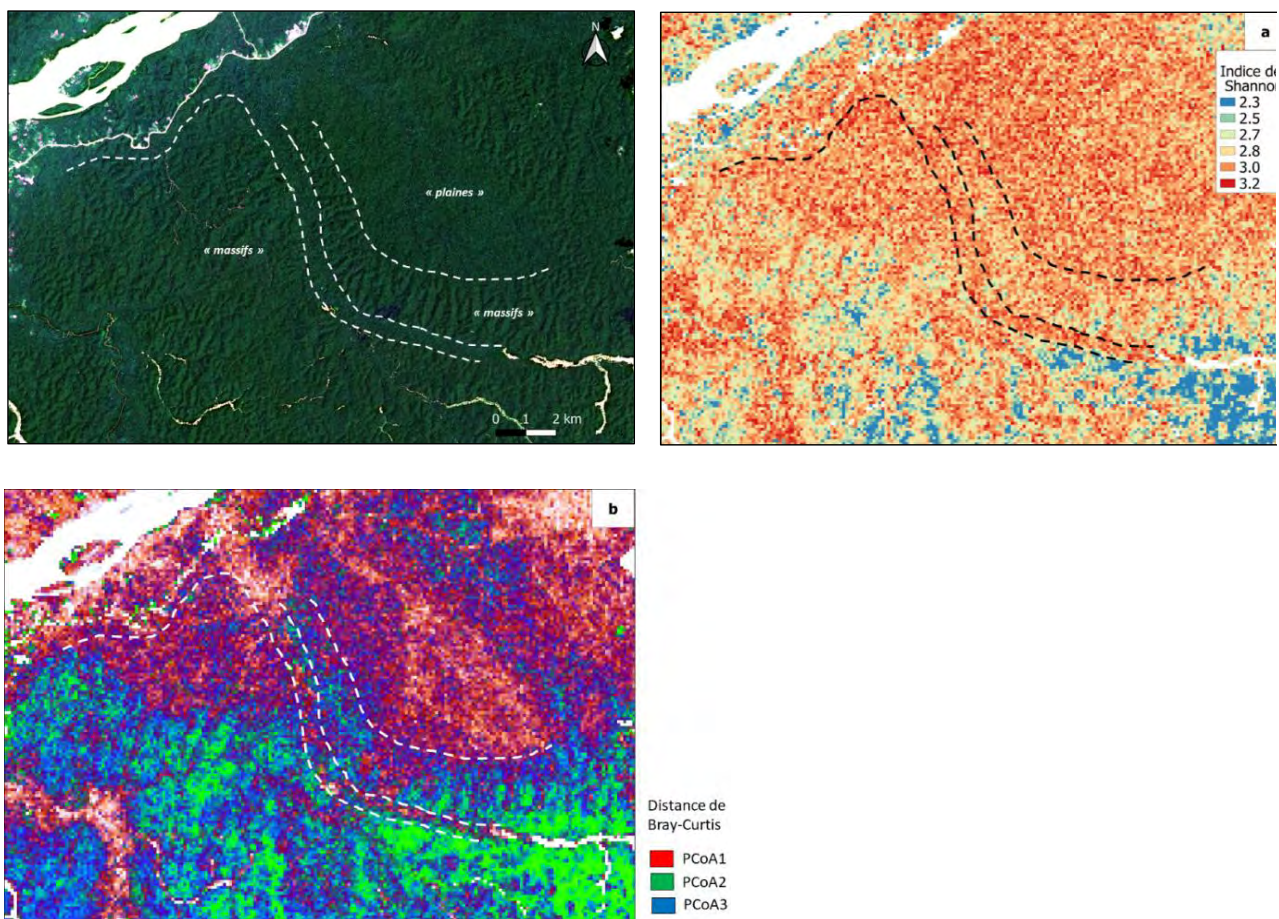


Figure 22. Exemple de motifs de diversité α et β associés aux variations du relief. La zone est située au Nord-Ouest de la zone d'étude. En haut à gauche l'image des réflectances en composition colorée « vraies couleurs ». Les traits pointillés indiquent des séparations entre plaines et massifs. **a.** La diversité α estimée par l'indice de Shannon. **b.** La diversité β représentée par les trois composantes de la PCoA basée sur la dissimilarité de Bray-Curtis. Résolution spatiale : 100 m. Source : ESA.

b. Motifs associés aux zones humides

Sur la **Figure 23-a**, le CHL présente des motifs de valeurs plus élevées dans la zone humide (autour de $60 \mu\text{g}\cdot\text{cm}^{-2}$; moyenne sur la zone d'étude = $53 \mu\text{g}\cdot\text{cm}^{-2}$). Concernant l'EWT (date du 23 août 2018), **Figure 23-b**, des motifs de valeurs plus faibles apparaissent dans la zone humide (0.015-0.018 cm ; moyenne = 0.019 cm). Le LMA (**Figure 23-c**) présente également des motifs de valeurs plus faibles dans la zone humide (autour de $0.015\text{-}0.019 \text{ g}\cdot\text{cm}^{-2}$; moyenne = $0.021 \text{ g}\cdot\text{cm}^{-2}$). Dans le cas du LAI (**Figure 23-d**), les valeurs tendent à augmenter (4.4-4.6 ; moyenne = 3.7) sans former de motif qui puisse être clairement distingué de la végétation des plaines.

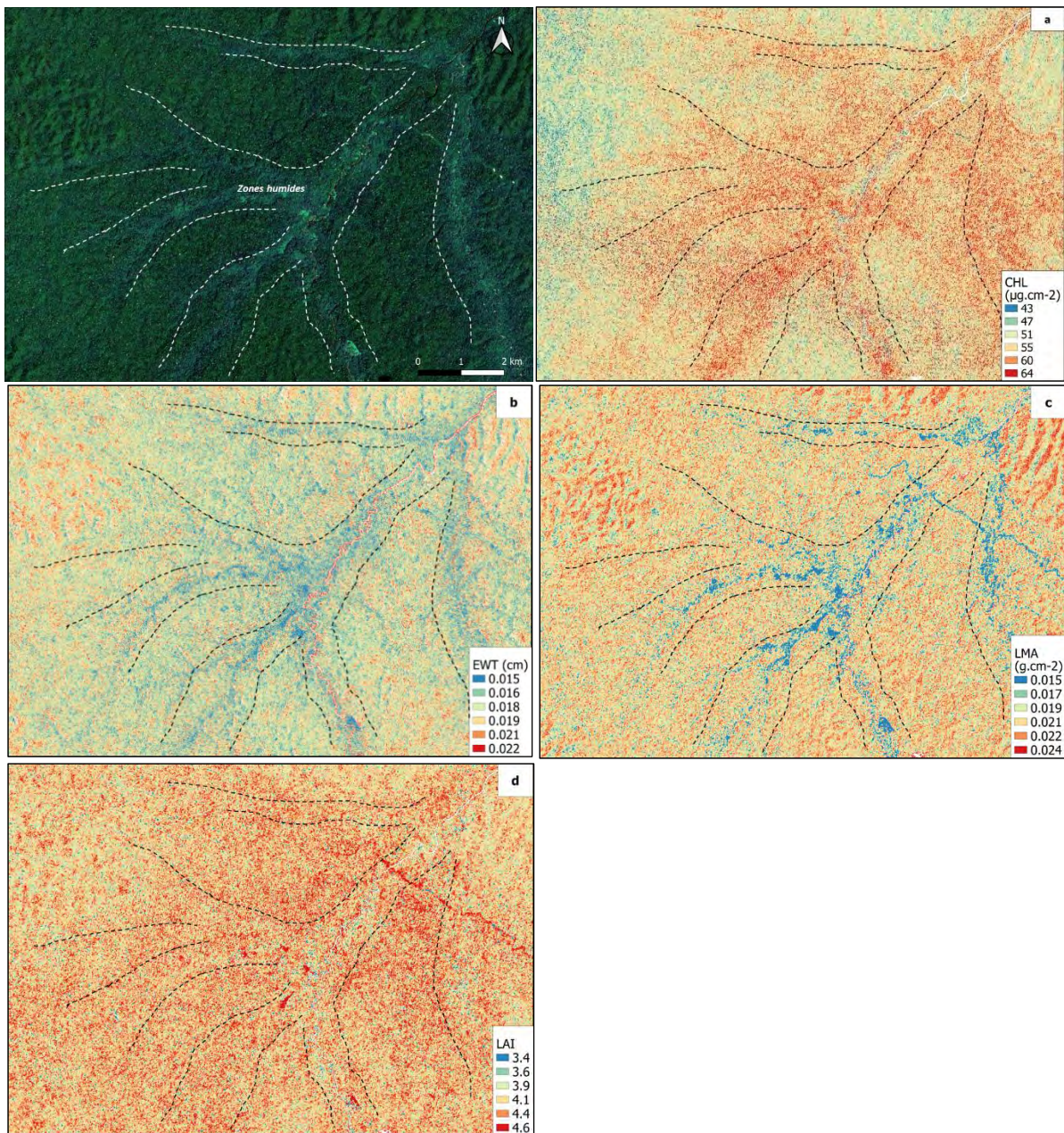


Figure 23. Exemple de motifs des quatre variables biophysiques associés à une zone humide. La zone est située au Sud-Ouest de la zone d'étude. En haut à gauche l'image des réflectances en composition colorée « vraies couleurs ». Les traits pointillés indiquent approximativement les limites de la zone humide et de sa végétation. **a.** Contenu en chlorophylle de la canopée ou CHL ($\mu\text{g}\cdot\text{cm}^{-2}$). **b.** Epaisseur équivalente en eau ou EWT (cm) (23/08/2018). **c.** Indice de masse foliaire ou LMA ($\text{g}\cdot\text{cm}^{-2}$). **d.** Indice de surface foliaire ou LAI (sans dimensions) (moyenne = 3.7). Résolution spatiale : 10 m. Source : ESA.

La diversité α présentée sur la **Figure 24-a** fait apparaître des motifs de valeurs globalement plus élevées au sein de la zone humide (2.8-3.2 ; moyenne = 2.9) avec des valeurs plus faibles correspondant au passage des cours d'eau (2.3-2.5). La diversité β révèle des contrastes entre la zone humide (rose pâle) et les zones situées en dehors (violet) suggérant l'existence de contrastes dans la composition de la communauté végétale (**Figure 24-b**).

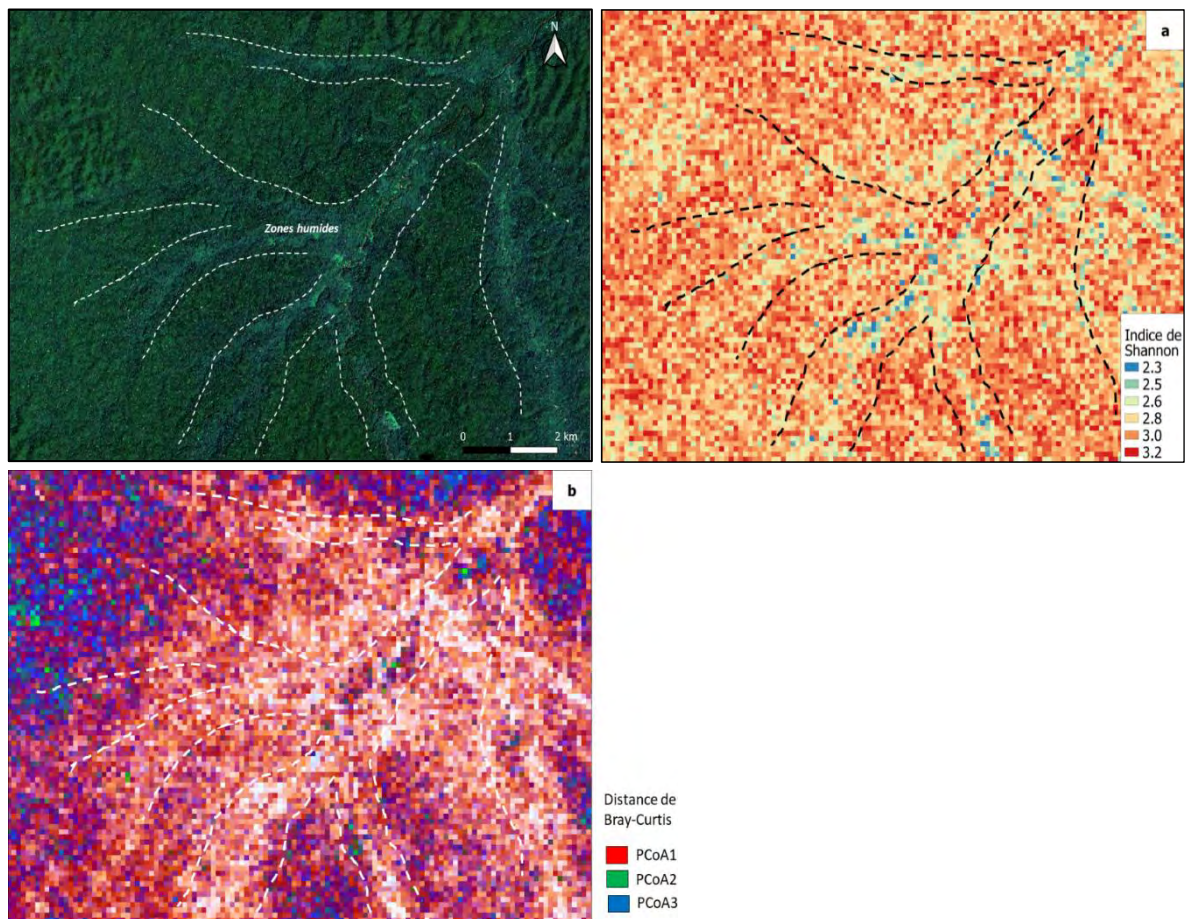


Figure 24. Exemple de motifs de diversité α et β associés à une zone humide. La zone est située au Sud-Ouest de la zone d'étude. En haut à gauche l'image des réflectances en composition colorée « Vraies couleurs ». Les traits pointillés indiquent approximativement les limites de la zone humide et de sa végétation. **a.** La diversité α estimée par l'indice de Shannon. **b.** La diversité β représentée par les trois composantes de la PCoA basée sur la Distance de Bray-Curtis. Résolutions spatiales : 10 m (« vraies couleurs ») et 100 m (diversité diversités α et β). Source : ESA.

IV. Relations entre les variables biophysiques

Une corrélation positive est observée entre les valeurs estimées d'EWT et de LMA ($\rho = 0.54$) (**Figure 25**). Une corrélation négative est observée entre CHL et LMA ($\rho = -0.48$) ainsi qu'entre CHL et EWT ($\rho = -0.38$). Une corrélation positive est obtenue entre CHL et LAI ($\rho = 0.43$). Ces corrélations sont modérées, mais toutes significatives. Elles suggèrent que l'EWT augmente avec le LMA et que le CHL augmente avec le LAI mais diminue avec l'EWT et le LMA.

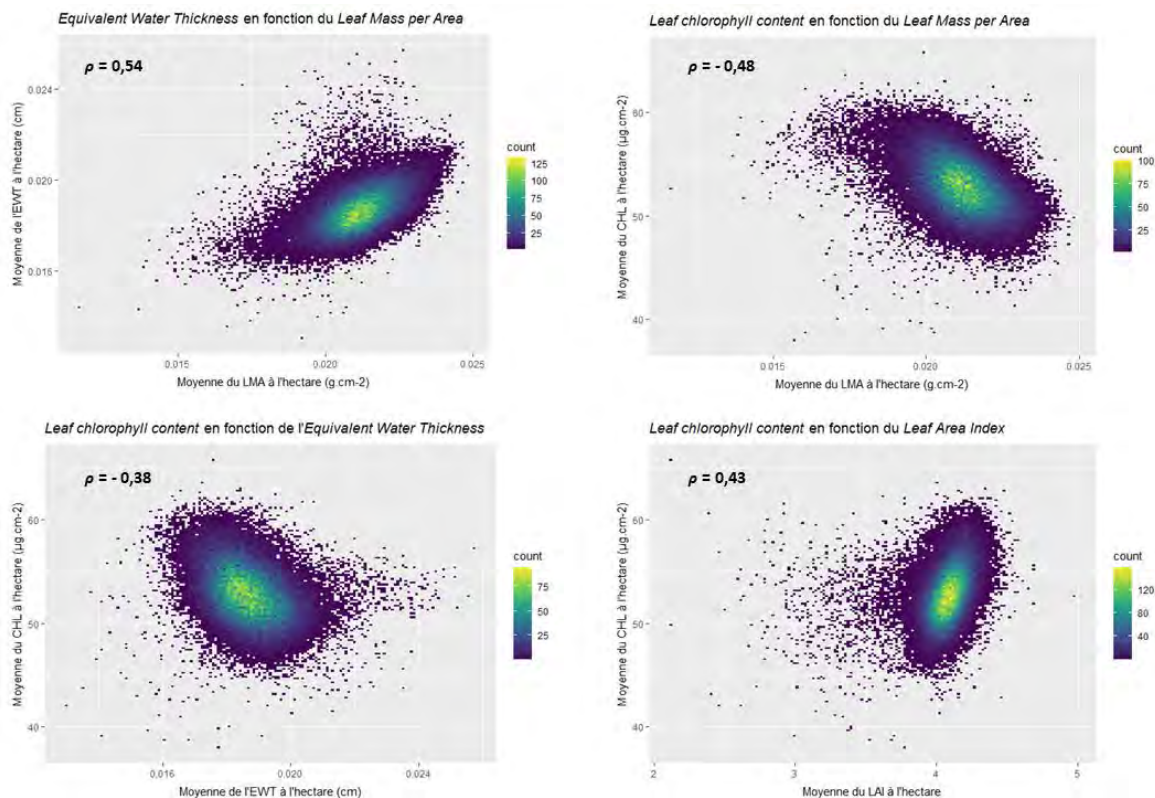


Figure 25. Tendances observées sur les principales corrélations entre les moyennes à l'hectare des variables biophysiques. Les différentes corrélations représentées sont : EWT - LMA, CHL - LMA, CHL - EWT et CHL - LAI. ρ correspond au coefficient de corrélation de Spearman (les quatre corrélations sont significatives avec une p -value < 0,05).

V. Comparaison avec les données de structure de la canopée

a. Relations entre variables biophysiques et le Modèle Numérique de Canopée

Pour évaluer la cohérence des variables biophysiques à l'aide du MNC, nous avons fait l'hypothèse que la structure du couvert ligneux peut être approchée par la distribution des écart-types en fonction de la moyenne de la hauteur maximale de la canopée, permettant ainsi de comparer les motifs de distributions écart-types et moyennes entre chaque variable biophysique et le MNC.

La distribution entre les valeurs moyennes et les écart-types du MNC à l'hectare dans la **Figure 26** apparaît de manière approximativement triangulaire, présentant trois « pôles » : un nuage d'écart-types et de moyennes relativement plus faibles (partie inférieure gauche du graphique, autour de la valeur 16,5 m, ± 3 m), un nuage d'écart-types et moyennes élevées (partie supérieure droite, autour de la valeur 31 m, $\pm 8,5$ m), un nuage d'écart-types très élevés et de moyennes plus faibles (16 m, ± 13 m). La formulation de cette hypothèse s'appuie sur l'expérience des écologues de l'UMR AMAP, mais à notre connaissance, elle n'a pas fait à ce jour l'objet d'une publication scientifique.

La distribution des valeurs entre ces trois « pôles » n'est pas homogène. Une densité plus importante du nuage de points est visible entre 25 et 30 mètres de hauteurs, ± 6 m. Le « pôle » dans la partie supérieure gauche (hauteurs moyennes relativement faibles, écart-types très élevés) possède la densité la plus faible.

Hauteur maximale de la canopée : écart-type en fonction de la moyenne

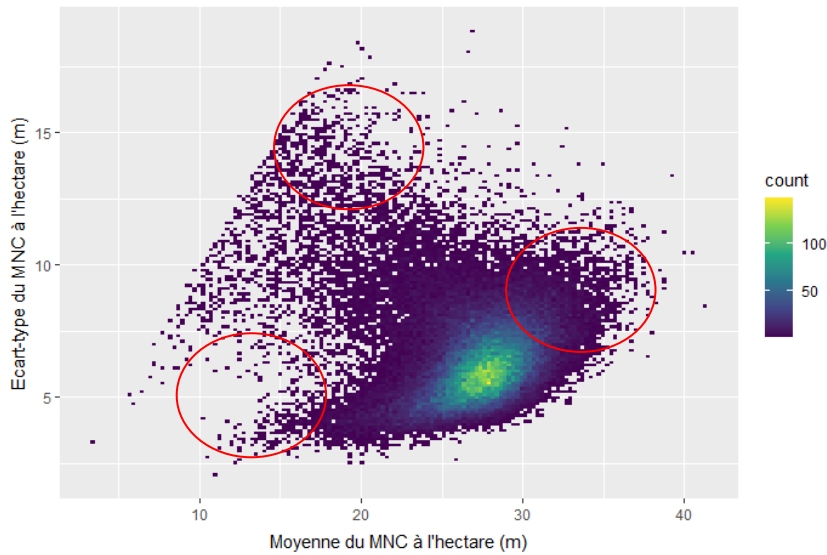


Figure 26. Distribution des écart-types en fonction des moyennes du MNC à l'hectare. Les cercles rouges indiquent les trois « pôles » de la « distribution triangulaire ».

Sur la **Figure 27** des tendances différentes apparaissent entre écart-types et moyennes en fonction du type de variable biophysique. Dans le cas du CHL et du LAI, l'écart-type tend à augmenter avec la moyenne. À l'inverse, les écart-types de l'EWT et du LMA tendent à diminuer lorsque les valeurs moyennes augmentent.

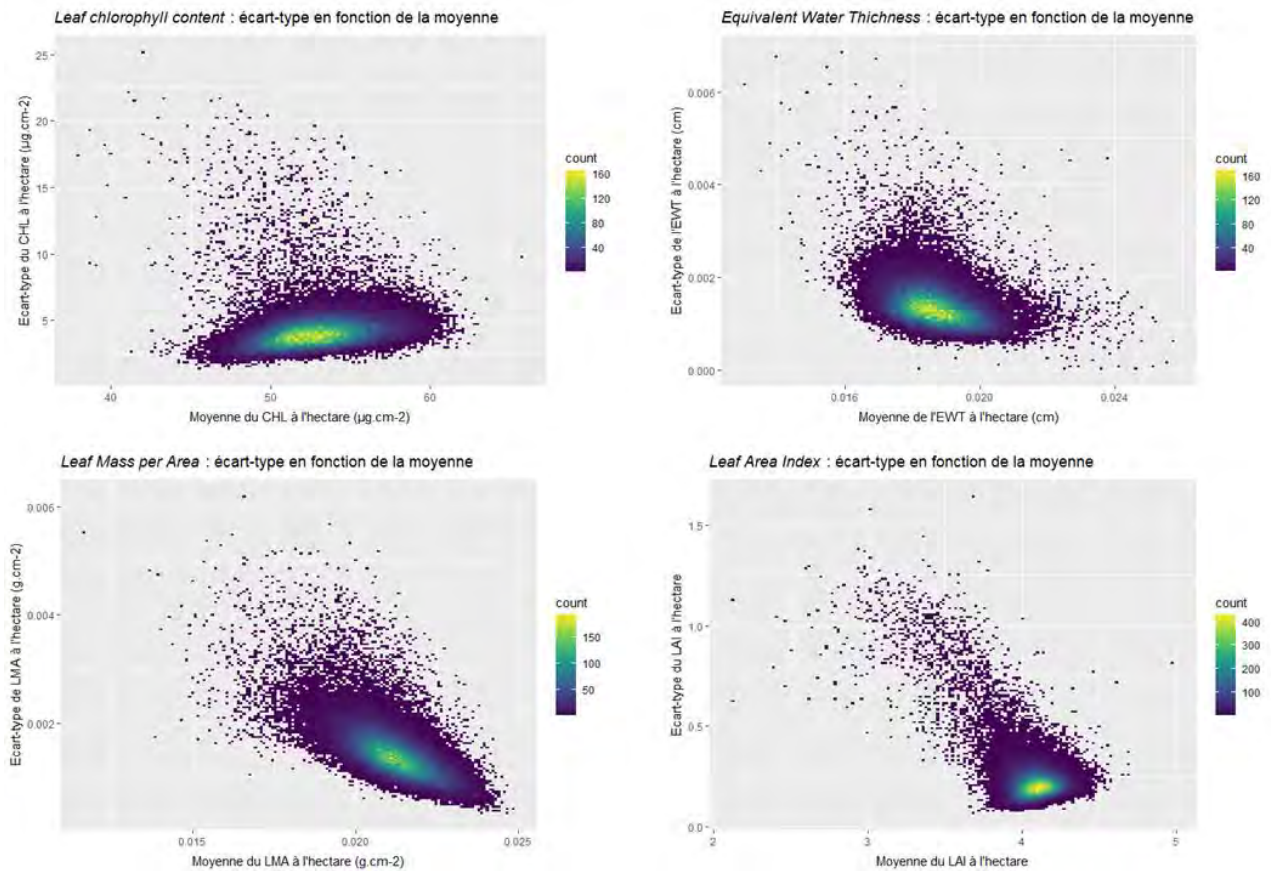
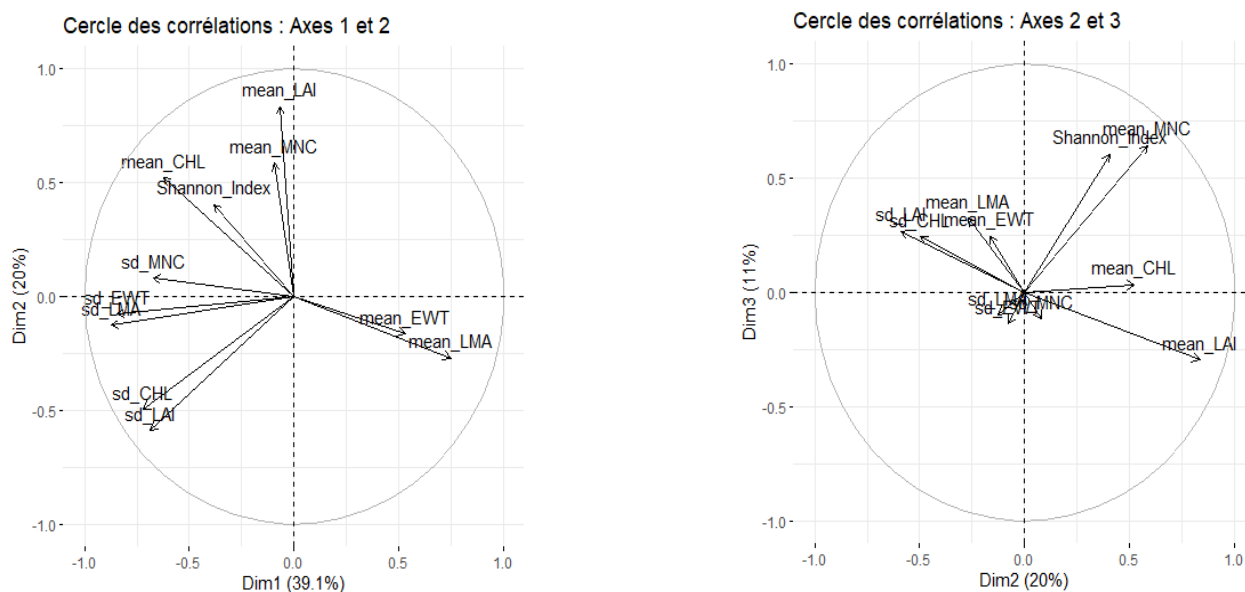


Figure 27. Distribution des moyennes et des écart-types à l'hectare des variables biophysiques sur la zone d'étude.

b. Analyse en Composantes Principales : MNC, variables biophysiques et diversité α

À l'issue de l'examen des valeurs propres des composantes principales de l'ACP (**Annexe 2**), nous avons sélectionné les trois premiers axes, selon le critère de Kaiser, c'est-à-dire les axes possédant des valeurs propres supérieures ou égales à 1. Ces trois axes expliquent environ 70 % de la variabilité, dont 39 % concernant le premier axe. D'après la **Figure 28** (cercle et coefficients des corrélations) l'axe 1 explique principalement l'EWT et le LMA (moyennes et écart-types) et l'écart-type du MNC. L'axe 2 explique principalement les moyennes du LAI et du MNC. L'axe 3 explique mieux la moyenne du CHL et la diversité α (indice de Shannon).



	PC1	PC2	PC3
<u>mean_MNC</u>	-0.092	0.586	0.645
<u>sd_MNC</u>	-0.672	0.0813	-0.117
<u>mean_CHL</u>	-0.622	0.523	0.0347
<u>sd_CHL</u>	-0.720	-0.496	0.245
<u>mean_LMA</u>	0.748	-0.270	0.325
<u>sd_LMA</u>	-0.873	-0.128	-0.099
<u>mean_LAI</u>	-0.066	0.836	-0.296
<u>sd_LAI</u>	-0.685	-0.589	0.268
<u>mean_EWT</u>	0.530	-0.165	0.248
<u>sd_EWT</u>	-0.845	-0.077	-0.136
<u>Shannon_Index</u>	-0.384	0.408	0.605

Figure 28. Cercle et coefficients des corrélations de l'ACP des écart-types (sd) et moyennes (mean) des différentes variables (variables biophysiques, MNC et diversité α par l'indice de Shannon). Plus les flèches sont proches du cercle des corrélations, mieux les variables sont représentées. Plus elles sont proches des axes et plus elles sont corrélées à l'axe considéré. Dans le tableau, les valeurs expriment les coefficients de corrélation entre les variables et les axes, indiquant la contribution relative des différentes variables à la construction des axes (plus la valeur absolue est élevée, plus la corrélation est importante et avec elle la contribution de la variable). Les écart-types et moyennes sont déterminées à l'hectare.

c. Corrélations entre la diversité β et l'analyse de texture de la méthode FOTO

En comparant visuellement les deux systèmes (**Figure 29**), nous pouvons remarquer des similitudes entre les motifs révélés par les deux méthodes. Cette comparaison suggère l'existence d'un lien entre les changements de compositions entre communautés (diversité β) et la variation de la texture de la canopée.

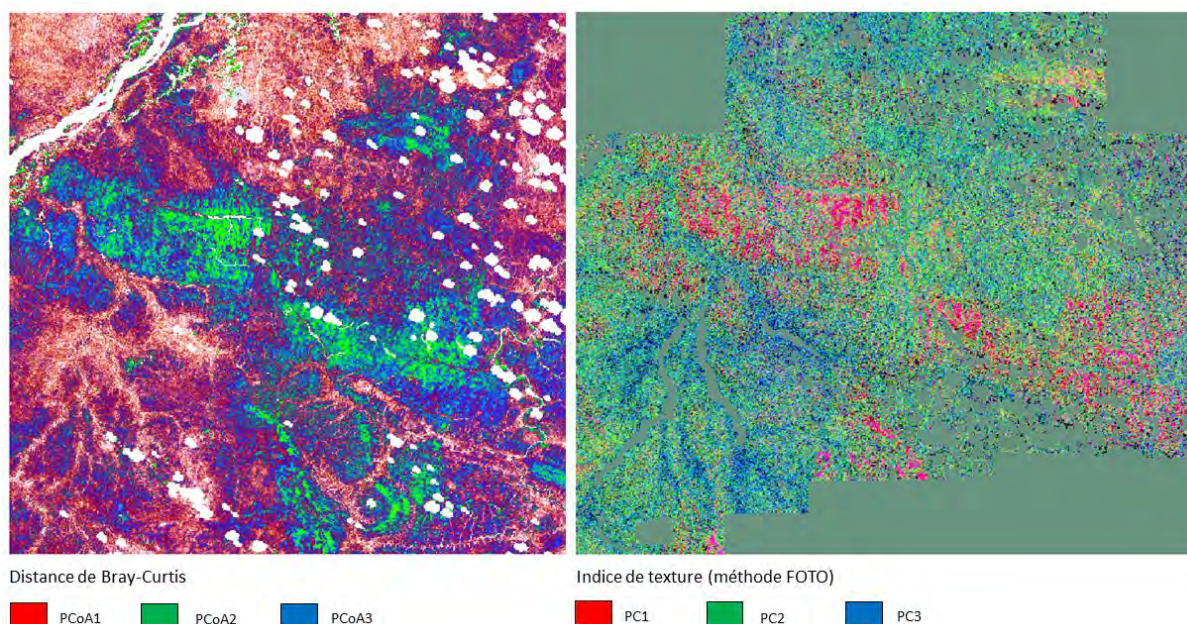


Figure 29. Comparaison des cartes de diversité β (gauche) et des indices de texture dérivés de la méthode FOTO (droite). Résolutions spatiales : 100 m (diversité β) et 50 m (indices de texture). Sources : ESA, UMR AMAP.

La **Figure 30** illustre les corrélations entre les deux systèmes d'axes. Ces corrélations sont relativement faibles, à l'exception de l'axe 1 de la PCoA (diversité β) et de l'axe 1 de l'indice de texture (méthode FOTO) avec $\rho = -0,43$.

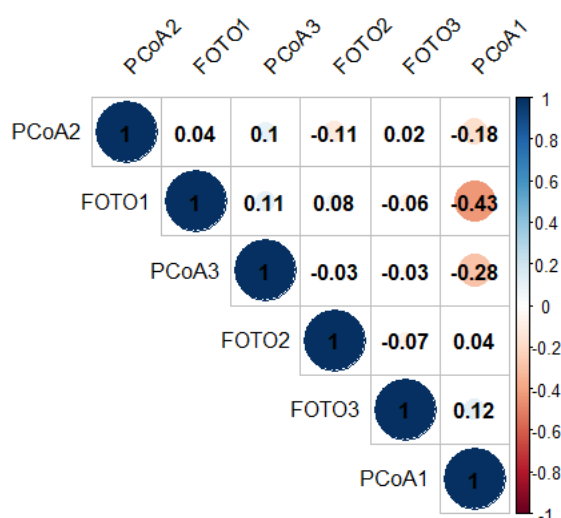


Figure 30. Corrélations entre les axes de la diversité β et les axes de l'indice de texture de la méthode FOTO (coefficients de Spearman). Les corrélations sont toutes significatives (p -value < 0,05).

Discussion

I. Evaluation de la cohérence des motifs et limites de la photo-interprétation

Les cartes des variables biophysiques et des indices de diversité α et β révèlent des motifs qui peuvent être associés avec une confiance relative à des variations du type de végétation : soit du point de vue écophysologique (variables biophysiques) ou en termes de composition de la communauté végétale (diversité α et β). Ces motifs n'apparaissent pas de manière systématique entre les différentes variables (*e.g.* un motif révélé par le LAI ne l'était pas nécessairement par le CHL) et manifestaient des niveaux de singularité qui rendaient possible la démarche de photo-interprétation. Les tendances que nous avons décrit pour nos différentes variables s'expriment par rapport aux valeurs moyennes (**Tableau 5**). Sur l'ensemble de l'image, nous avons pu distinguer deux types de motifs relativement fréquents :

D'abord, des motifs correspondant à des variations du type de végétation en fonction du relief. Les différences se sont révélées plus marquées entre la végétation des « massifs » et celle des « plaines ». Pour approcher les variations topographiques du terrain, nous pouvions nous appuyer sur le MNT et l'image des réflectances en composition « Vraies couleurs ». D'après ce type de motifs, le CHL, l'EWT et le LAI tendaient à augmenter pour la végétation de plaines. Cependant, dans le cas de l'EWT, la tendance était moins nette. L'estimation du LMA s'est révélée impactée par la présence d'un artefact, compte tenu de l'effet d'« orientation » de la distribution des valeurs les plus élevées, correspondant à l'ombre projetée par le relief. Ces valeurs « orientées » du LMA sont très probablement surestimées. La diversité α tendait à diminuer sur les massifs et la diversité β contrastait fortement entre plaines et massifs.

Ensuite, des motifs pouvaient être associés à la végétation des zones humides, repérables via le réseau hydrographique. Le CHL augmentait dans les zones humides, contrairement à l'EWT et au LMA qui tendaient à diminuer. Dans le cas du LMA, le motif présentait des variations spatiales particulières qui différaient des autres variables de manière importante. Le LAI et la diversité α présentaient des valeurs plus élevées dans les zones humides (par rapport à la moyenne) sans toutefois manifester contraste avec la végétation de plaine. La diversité β apparaissait de manière fortement contrastée dans ces zones (dans le rose pâle, selon la composition RVB).

Tableau 5. Tableau récapitulatif des tendances observées pour les différents motifs de variables biophysiques et indices de biodiversité en fonction du type de végétation identifié. Les tendances indiquées (augmentation et diminution) sont relatives aux valeurs moyennes sur l'ensemble de la zone d'étude (après exclusion des valeurs extrêmes).

Motif	Massifs	Plaines	Zones humides	Moyenne	Variance
CHL	Diminue (43-51 $\mu\text{g.cm}^{-2}$)	Augmente (~ 60 $\mu\text{g.cm}^{-2}$)	Augmente (~ 60 $\mu\text{g.cm}^{-2}$)	53.3	7.35
EWT	Augmente (0.019-0.021 cm)	Diminue (0.015-0.019 cm)	Diminue (0.015-0.018 cm)	19	$2.16 \times e^{-16}$
LMA	Augmente (~ 0.024 g.cm^{-2})	Augmente (0.021-0.022 g.cm^{-2})	Diminue (0.015-0.019 g.cm^{-2})	21	$2.65 \times e^{-16}$
LAI	Diminue (3.9-4.1)	Augmente (4.4-4.6)	Augmente (4.4-4.6)	4.13	0.04
Diversité α	Diminue (2.3-3.5)	Augmente (2.8-3.2)	Augmente (2.8-3.2)	2.9	0.024
Diversité β	Vert	Rose pâle	Rose pâle	-	-

La végétation sur les massifs tend à présenter à la fois un LAI et une diversité α plus faibles, avec un contraste très marqué concernant la diversité β . Les échanges avec les écologues de l'UMR AMAP nous ont permis de déterminer que la zone d'étude comprends un certain nombre de forêts se développant sur des sols saprolitiques, issus de l'altération de la partie supérieure de la roche-mère, notamment dans des zones d'affleurement (**Annexe 3**). Sur ce type de sol, la végétation est spécialisée (adaptée aux conditions physico-chimiques particulières) et généralement moins haute et moins diversifiée que la forêt tropicale de plaine. Les motifs associés aux « massifs » sont donc potentiellement caractérisés par ce type de végétation basse et peu diversifiée (faible diversité α), spécialisée (contraste fort dans la diversité β), sur des affleurements à profils saprolitiques.

Néanmoins, notre capacité à photo-interpréter les motifs s'est trouvée rapidement limitée par le manque de données complémentaires sur la zone d'études, l'image en composition colorée et le MNT n'apportant que des informations très partielles sur la composition paysagère. Certains motifs, bien visibles à travers les différentes variables, n'ont pu être associés de manière claire à un type particulier de végétation, sur la base d'un élément spécifique du paysage (**Annexe 4**). L'association entre les plans d'eau et les massifs rendaient également confuse l'interprétation des motifs, ne sachant pas *a priori* quel élément était le plus déterminant sur le type de végétation considéré. C'est également le cas pour la distinction entre la végétation des plaines et celle des zones humides.

Concernant la diversité α et β , un motif ne pouvait être associé de manière unique avec un type de végétation spécifique (*e.g.* végétation de zone humide) puisque des communautés avec des compositions en espèces différentes peuvent présenter des niveaux similaires en termes de richesse spécifique (diversité α) ou de taux de variations de la composition en espèces (diversité β) (**Figure 28**). Dans le cas de la diversité β , il est également possible que des différences initialement présentes dans la matrice de dissimilarité aient été « masquées » lors de la construction des axes synthétiques de la PCoA. L'examen de la matrice de dissimilarité devrait permettre de vérifier l'existence de telles différences (par exemple entre les végétations de massifs et de certaines zones humides apparaissant dans le vert). En l'absence de données de validation, l'interprétation de motifs de diversité α et β sur la base de l'image en composition colorée « Vraies couleurs » doit rester prudente. Ces motifs peuvent néanmoins préciser l'interprétation des variables biophysiques : les valeurs plus faibles de diversité α vont dans le sens de l'interprétation de la végétation des massifs comme de la végétation spécialisée sur sol saprolitique.

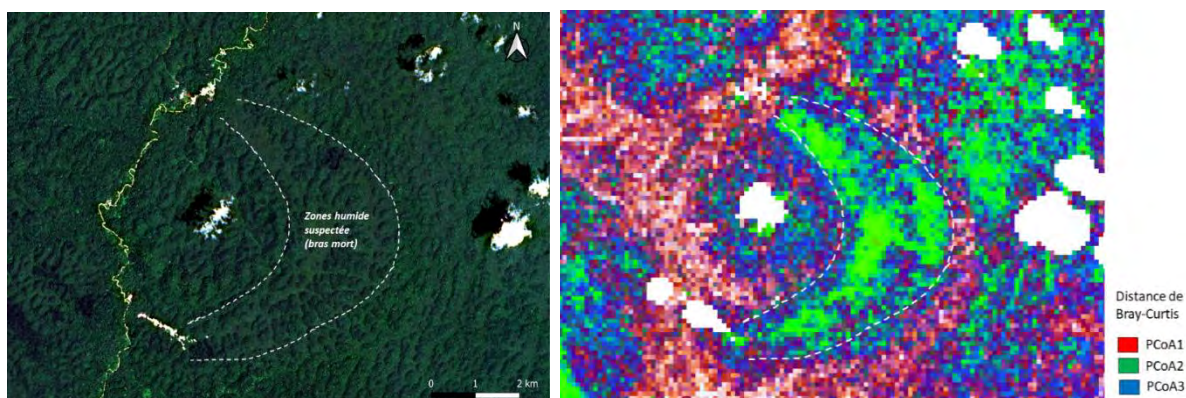


Figure 31. Zone humide suspectée (bras mort de rivière) en composition colorée « Vraies couleurs » (gauche) et en diversité β . La zone se situe au sud-est de la zone d'étude. La couleur verte du motif de diversité β est identique à celle observée avec les « massifs ». Résolutions spatiales : 10 m (« Vraies couleurs ») et 100 m (diversité β). Source : ESA

Les résultats de la photo-interprétation, bien que partiels et limités par l'absence de données complémentaires sur la composition paysagère de la zone d'étude, révèlent néanmoins une certaine cohérence des motifs, en lien avec des types de végétation identifiés de façon *ad hoc* principalement à partir de l'image « Vraies couleurs ». Nous pouvons raisonnablement espérer que le modèle PROSAIL permette de décrire le couvert ligneux de la forêt tropicale humide guyanaise à travers sa dimension fonctionnelle et écophysologique, de manière relativement satisfaisante, sur la base des données Sentinel-2. Les motifs de diversité α et β constituent également un résultat encourageant dans la mesure où des contrastes clairs se manifestent (notamment dans le cas de la diversité β). En l'absence de données de validation, les valeurs de nos différentes variables ne peuvent être interprétées que de manière relative et par leur distribution spatiale, sous forme de motifs.

II. Limites de l'inversion de PROSAIL et identification des artéfacts

Notre démarche de photo-interprétation a également permis de soulever certaines limites dans l'inversion de PROSAIL avec la méthode du package *prosail*. Si l'on s'intéresse au LAI, des motifs de valeurs plus élevées (par rapport à la moyenne) apparaissent sur des zones de végétation basse, le long de certains cours d'eau. Il s'agit probablement de zones d'abattis voire de sites d'orpaillage (d'après la connaissance du terrain des écologues de l'UMR AMAP). Dans ce contexte, il est surprenant d'observer

ces valeurs de LAI plus élevées que celles de la forêt dense, car on s'attendrait au contraire à observer des valeurs faibles appartenant à une végétation basse recolonisant un espace ouvert. Ces valeurs élevées peuvent s'expliquer par l'entraînement de l'inversion de PROSAIL, réalisé sur les caractéristiques biophysiques (simulées) d'une végétation dense. Le modèle ne serait pas en mesure de prédire correctement une végétation basse. Cette limite interroge sur la fiabilité des estimations de PROSAIL concernant l'hétérogénéité potentielle de la végétation présente dans notre zone d'étude : la caractérisation biophysique de couverts ligneux moins dense pourrait se trouver erronée. L'entraînement de l'inversion appelle à être affiné à l'aide de mesures sur le terrain (incluant des types de couverts forestiers plus ou moins denses) mais aussi en fonction des objectifs de cartographie du couvert ligneux (tous les types de végétation ne peuvent être identifiés avec la même précision).

L'estimation du LMA semble également problématique. Sur les massifs, nous pouvons observer que des motifs de valeurs élevées présentent une forme d'orientation, à la manière de l'ombre produite par les plis du relief. Il semblerait que des pixels d'ombre non-masqués soient pris en compte dans la détermination du LMA. En effet, en comparant le masque d'ombre final à l'image « Vraies couleurs », on s'aperçoit qu'une partie de l'ombre n'a pas été masquée (**Annexe 5**). Néanmoins, lorsque l'on s'intéresse à la distribution de valeurs du LMA (**Figure 16**), on ne détecte aucune trace de saturation. Il semble donc que malgré la présence de ce biais, le LMA demeure informatif (bien qu'à interpréter avec précaution). La présence d'ombres résiduelles peut s'expliquer par la définition d'une valeur seuil du filtre radiométrique dans le Proche Infrarouge trop élevée. Nous avons par ailleurs testé différentes valeurs seuils (18 %, 15 %, 10 %) mais nous nous sommes vite aperçus que des valeurs inférieures à 20 % conduisaient à intégrer dans le masque de nuages un grand nombre de pixels de végétation. Nous avons été contraints d'abandonner cette solution. Après examen visuel, les autres variables biophysiques n'apparaissent pas impactées par la présence de pixels d'ombre résiduels.

Malgré la résolution élevée de Sentinel-2, nous avons remarqué la présence de pixels mixtes, c'est-à-dire comprenant à la fois de la végétation et un élément « autre », non-végétalisé. En examinant les motifs à la limite du réseau hydrographique, des valeurs extrêmes apparaissent pour les différentes variables biophysiques. Le problème des pixels mixtes concernant les images à moyenne et haute résolution est bien connu. Des solutions existent, toutefois, pour notre objectif d'exploration de la cohérence de nos variables, cette difficulté impacte peu nos conclusions. Néanmoins, il serait intéressant de quantifier la présence des pixels mixtes au sein de certains types de végétation, notamment dans le cas des zones humides, la présence de mares et de nombreux cours d'eau pouvant conduire à fausser les estimations et limiter l'interprétabilité des résultats. Dans le cas du LMA, il est possible que la présence de pixels mixtes (eau-végétation) dans les zones humides conduise à la formation de motifs à une échelle plus fine que celle que nous avons identifié initialement sur la base de l'image « Vraies couleurs » (**Figure 23-c**) et conduise à biaiser notre interprétation, à travers la formation de motifs correspondant en réalité à des réseaux de pixels eau-végétation.

Enfin, il semble que notre capacité à prédire correctement les variables biophysiques dépende aussi fortement de la qualité des données initiales. Nous nous sommes aperçus que Sentinel-2 présentait un artéfact sur la bande 11 (1613.7 nm, domaine du *Short-wave Infrared* ou SWIR) apparaissant sous la forme d'une longue bande orientée Nord-Est avec une tendance à sous-estimer les valeurs de réflectances (**Annexe 6**). Un artéfact similaire a été relevé, quoique de manière plus légère, sur la bande 5 (740.2 nm, domaine du *Red Edge*). Des différences de configuration de capteurs entre Sentinel-2A et Sentinel-2B ou un effet lié au pré-traitement des images Sentinel-2 pourraient être à l'origine de l'artéfact. La sous-estimation de la réflectance sur la bande 11 conduit à une surestimation de l'EWT sur une grande partie de l'image (**Annexe 6**), nous contraignant à utiliser pour cette variable une autre date ne présentant pas cet artéfact (*i.e.* le 12 août 2018). Sur les six images que nous avons pu télécharger et examiner, nous nous sommes aperçus de l'existence d'artéfacts similaires sur trois d'entre elles (dont la date du 16 octobre que nous avons principalement utilisée) conduisant à des erreurs d'estimation de l'EWT sur une partie importante de la zone d'étude. La haute résolution temporelle de Sentinel-2 constitue un avantage certain, donnant la possibilité de choisir l'image la plus fiable. La présence de ces artéfacts appelle à davantage de contrôle sur la base de données de validation et à mieux comprendre l'impact des pré-traitements sur les résultats.

III. Interprétation des corrélations entre les quatre variables biophysiques

Sur la **Figure 25**, la corrélation entre les moyennes du LMA et de l'EWT peut s'expliquer par le fait que l'augmentation du contenu en eau des feuilles entraîne une augmentation de la masse par unité de surface de la feuille. La diminution du CHL avec le LMA semble aller dans le sens de cette interprétation, le contenu en eau faisant augmenter la masse des feuilles, alors que la proportion du contenu en chlorophylle, participant à la masse des feuilles, diminue. Une explication similaire peut rendre compte de la diminution du CHL avec l'EWT. Enfin, si le LAI augmente, la surface foliaire en contact avec la lumière augmente et le CHL augmente. Ces différentes explications, bien qu'elles doivent être testées formellement, permettent de rendre compte de manière cohérente des principales corrélations observées entre les variables biophysiques.

IV. Evaluation de la cohérence des variables par confrontation avec les données de texture

Pour guider notre interprétation, nous avons fait l'hypothèse que la structure du couvert ligneux pouvait être caractérisée à travers la relation entre la moyenne et l'écart-type de la hauteur maximale de la canopée (MNC). Nous avons pu identifier, sous forme d'une tendance réparable graphiquement, une distribution des valeurs que l'on peut qualifier de « triangulaire », répartie entre trois « pôles » : 1- Moyennes faibles, écart-types faibles, 2- Moyennées élevées, écart-types modérément élevés, 3- Moyennes faibles, écart-types élevés. Nous pouvons proposer une interprétation de ces trois pôles. Si des écart-types élevés semblent indiquer la présence d'une végétation perturbée, et la hauteur de la forêt son degré de fermeture, nous pouvons faire l'hypothèse que les trois pôles correspondent, respectivement : 1- Un couvert ligneux ouvert caractérisé par une végétation relativement basse et faiblement perturbée, 2- Un couvert ligneux fermé caractérisé par une végétation haute faiblement perturbée, 3- Un couvert ligneux ouvert caractérisé par une végétation relativement basse et fortement perturbée. L'existence de cette diversité de structure du couvert ligneux, à l'intérieur d'un continuum entre les trois pôles suggère que nous pourrions distinguer, dans une certaine mesure, une structuration similaire concernant nos variables biophysiques, partant du principe que les différences de maturité et le degré de perturbation entre les forêts influent sur leur traits fonctionnels et leurs caractéristiques écophysiologiques (Enquist *et al.*, 2007; Boukili and Chazdon, 2017).

Parmi les quatre variables biophysiques, le CHL et surtout le LAI présentent une distribution des valeurs qui s'approche de la « distribution triangulaire » observée dans le cas du MNC. Cette ressemblance suggère que le CHL et le LAI tendent à varier avec la structure du couvert ligneux : ces deux variables sont relativement plus élevées et stables dans les couverts ligneux fermés peu perturbés et plus faibles dans les couverts ligneux plus ouverts, plus ou moins stables selon le niveau de perturbation de la végétation. Ces relations apparaissent cohérentes avec ce que l'on pourrait attendre des caractéristiques biophysiques de ces trois types de couverts ligneux. Bien qu'à prendre avec beaucoup de précautions, cette tendance plaide en faveur du potentiel d'estimation du CHL et du LAI de la forêt tropicale humide par inversion de PROSAIL. Dans le cas du LMA, une nette tendance négative est observée. Pour expliquer cette distribution, nous pouvons faire l'hypothèse que cette tendance corresponde en réalité à la partie supérieure droite de la « distribution triangulaire » du MNC, autrement dit, au continuum de valeurs allant de couvert ligneux ouverts très perturbés à des couverts ligneux fermés et faiblement perturbés (continuum entre les pôles 3- et 2- décrits dans le paragraphe précédent). Dans une certaine mesure, cette distribution pourrait également être liée à l'artéfact évoqué précédemment concernant l'ombre du relief. Néanmoins, les valeurs ne présentent pas de saturation autour d'une valeur extrême qui puisse correspondre à un artéfact. Concernant l'EWT, la tendance est beaucoup moins claire et plus difficile à interpréter à l'aide du MNC. Néanmoins, nous pouvons faire l'hypothèse que contrairement aux autres variables, les variations de l'EWT sont davantage impactées par les conditions environnementales (*e.g.* température, humidité) et est de fait moins liée à la structure de la canopée telle que nous y avons accès à travers le MNC. Cette relative indépendance suspectée entre l'EWT et la structure de la canopée semble confirmée par l'orthogonalité entre la moyenne de l'EWT et la moyenne du MNC d'après les axes 1 et 2 de l'ACP (**Figure 28**).

En résumé, au moins concernant la distribution du CHL et du LAI, nos résultats vont dans le sens d'une cohérence générale avec la structure du couvert ligneux, sur la base des ressemblances entre les distributions des écart-types en fonction des moyennes et de la connaissance des types possibles de structure de couverts ligneux. Ces tendances, observées de manière empirique, doivent cependant être complétées par des analyses plus formelles de la distribution des variables biophysiques intégrant la structure du couvert ligneux.

Sur la **Figure 28**, on remarque que l'écart-type du MNC et sa moyenne sont orthogonaux selon les axes 1 et 2 de l'ACP. Ce résultat suggère que le niveau de perturbation du couvert ligneux (accessible via l'écart-type) ne dépend pas de sa hauteur moyenne. D'autre part, les moyennes du LAI et du MNC sont positivement corrélées selon l'axe 2, ce qui suggère que des couverts ligneux plus hauts auront tendance à posséder une surface foliaire plus importante. Cette corrélation renforce également l'interprétation proposée dans le paragraphe précédent, sur la ressemblance entre les « distributions triangulaires » du LAI et du MNC. On remarque que le CHL (moyennes et écart-types) et la diversité α (indice de Shannon) sont corrélés de manière assez équivalente aux trois axes de l'ACP, avec des valeurs relativement faibles (-0.38, 0.41 et 0.60, respectivement). Nous pouvons en conclure que la diversité α varie de façon relativement indépendante de la structure de la canopée, du moins telle que nous l'avons caractérisée (MNC mais aussi LAI). Pour expliquer ce résultat, nous pouvons faire l'hypothèse que des communautés d'arbres structurellement très différentes peuvent présenter des niveaux de diversité α similaires et réciproquement. En effet, les liens entre la diversité fonctionnelle (que l'on peut rapprocher de la structure du couvert ligneux) et la diversité α sont complexes et ne peuvent être interprétés en termes de relations linéaires. Par exemple, il a été démontré que la diversité fonctionnelle atteint un plateau avec l'augmentation du nombre d'espèces (diversité α) en raison de la redondance fonctionnelle à l'intérieur des communautés (Mason *et al.*, 2005; Cadotte, Carscadden and Mirotnick, 2011). Quant au CHL (moyenne et écart-type), il apparaît davantage corrélé à l'axe 1 de l'ACP. Ce résultat semble être contradictoire avec la relative ressemblance entre les « distributions triangulaires » des moyennes du CHL et du MNC (expliquée principalement par l'axe 2 de l'ACP) qui pouvait suggérer un lien entre structure du couvert ligneux et le contenu en chlorophylle foliaire. Toutefois, la ressemblance entre les distributions des écart-types en fonction de la moyenne était moins nette que dans le cas du LAI. Nous pouvons faire l'hypothèse que le CHL n'est lié que de manière partielle à la structure du couvert ligneux telle que décrite par le MNC. L'indépendance entre la moyenne de l'EWT et la structure du couvert ligneux approchée par les moyennes du MNC et du LAI est cohérente avec notre observation précédente concernant l'absence de ressemblance entre leurs « distributions triangulaires » (**Figures 27 et 28**). Cette interprétation vaut également pour le LMA, corrélé avec l'EWT. La nature des relations entre le LMA, l'EWT et la structure du couvert ligneux, appelle à être caractérisée plus en détails.

En comparant l'image de texture dérivée de la méthode FOTO avec l'image de diversité β , on s'aperçoit de l'existence d'une ressemblance importante entre les motifs (**Figure 29**). Face à ce constat, nous aurions tendance à confirmer notre hypothèse selon laquelle des différences de texture observées puissent s'expliquer par des communautés différentes. Pourtant, les corrélations entre les deux systèmes d'axes demeurent relativement faibles, contredisant cette interprétation (**Figure 30**). Il est possible que cette correspondance visuelle entre les deux images corresponde en réalité à des transitions entre des zones légèrement plus homogènes en termes de texture, sans qu'elles ne reflètent directement des différences de composition des communautés. Cette ressemblance appelle cependant à examiner plus en profondeur les relations entre la diversité β et la texture du couvert ligneux. Par exemple, l'indice de Jaccard permet de caractériser le degré de ressemblance entre les motifs de deux images (Fletcher and Islam, 2018) et pourrait contribuer à préciser la nature de cette ressemblance entre texture de la couvert ligneux et diversité β . En raison du problème précédemment évoqué de non-superposition entre les deux images, il n'a pas été possible de tester l'indice de Jaccard dans le cadre cette étude.

Conclusion & Perspectives

Notre travail a permis de montrer que les images Sentinel-2 en entrée de nos chaînes de traitement permettent de caractériser certaines propriétés biophysiques et écologiques du couvert ligneux tropical, dans le contexte d'un milieu forestier dense au nord-ouest de la Guyane Française. Les informations produites relatives aux variables biophysiques et aux indices de diversité ont contribué à caractériser différents types de végétation associés aux variations du relief et à la présence d'un réseau hydrographique. Une phase de validation s'appuyant sur des données terrain est maintenant nécessaire. Les motifs biophysiques et écologiques obtenus à l'aide de *prosail* et *biodivMapR* ont révélé une relative cohérence avec le type et la structure du couvert ligneux, décrits à l'aide de données de télédétection acquises avec différents capteurs. Cette cohérence a notamment pu être explorée par photo-interprétation et par confrontation avec un Modèle Numérique de Canopée et une analyse de texture issue de la méthode FOTO. La mise en place d'un cadre de comparaison entre les différents produits issus de données de télédétection a cependant montré des limites en raison de certaines différences entre capteurs (comme la résolution spatiale), empêchant de comparer des informations parfaitement cohérentes spatialement. Nous avons pu néanmoins mettre en évidence l'existence de corrélations : 1- entre certaines variables biophysiques, pouvant donner lieu à une interprétation cohérente des liens entre le CHL et la structure ou la chimie foliaire (LAI, LMA et EWT), 2- entre nos résultats et certaines composantes du MNC (moyennes et écart-types) donnant un aperçu de la structure globale de la canopée. D'autre part, la ressemblance entre les motifs de texture dérivés de la méthode FOTO et les motifs de diversité β suggère l'existence d'une relation entre la variabilité de la taille des couronnes des arbres et la variation de la composition des communautés du couvert ligneux. Cette tendance, qui doit encore être testée de manière plus formelle, plaide en faveur de la pertinence des motifs de diversité β . Un type de couvert ligneux a pu être identifié avec un degré de confiance important, avec l'appui des experts de la végétation locale, révélé en particulier par les motifs de diversité α et le LAI : la végétation de forêt basse peu diversifiée sur sapolites.

Les corrélations que nous avons pu mettre en évidence, ainsi que les motifs correspondant à nos différentes variables, penchent en faveur du potentiel des méthodologies que nous avons employé. Ces résultats peuvent également encourager le financement d'une campagne de terrain visant à acquérir des données de validation et la recherche de jeux de données déjà existants pouvant constituer une forme de validation auprès des instituts travaillant sur la zone (*e.g.* inventaires forestiers de l'ONF). L'évaluation de la précision des estimations des différentes variables biophysiques constituera une étape essentielle vers l'établissement d'une cartographie opérationnelle du couvert ligneux en Guyane Française. Il serait également intéressant d'évaluer comment pleinement exploiter le potentiel de Sentinel-2 face aux difficultés que soulèvent les artéfacts liés au capteur ou à la présence de pixels d'ombre ou de nuages non-filtrés (notamment dans le cas de l'estimation de l'EWT et du LMA). Évaluer le potentiel d'utilisation des images présentant des taux d'ennuage plus importants mériterait d'être exploré, ainsi que le potentiel des synthèses mensuelles fournies par Théia¹. Sentinel-2 devrait également permettre d'accéder à certaines variations phénologiques du couvert ligneux, bien que l'on s'attende à ce que ces variations soient moins importantes en contexte tropical, comparativement aux forêts tempérées. Une fois testés et validés, les outils et méthodes exploités dans le cadre du stage contribuant à la caractérisation du couvert ligneux tropical pourront être utilisés dans d'autres zones géographiques, et à des échelles régionales, comme l'échelle du plateau des Guyanes.

Bibliographie

- Allen, W. A. *et al.* (1969) 'Interaction of isotropic light with a compact plant leaf', *Josa*, 59(10), pp. 1376–1379.
- Anderson, J. *et al.* (2008) 'Integrating waveform lidar with hyperspectral imagery for inventory of a northern temperate forest', *Remote Sensing of Environment*, 112(4), pp. 1856–1870. doi: 10.1016/j.rse.2007.09.009.
- Anderson, M. J. *et al.* (2011) 'Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist: Roadmap for beta diversity', *Ecology Letters*, 14(1), pp. 19–28. doi: 10.1111/j.1461-0248.2010.01552.x.
- Asner, G. P. and Martin, R. E. (2009) 'Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests', *Frontiers in Ecology and the Environment*, 7(5), pp. 269–276. doi: 10.1890/070152.
- Baccini, A. *et al.* (2012) 'Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps', *Nature Climate Change*, 2(3), pp. 182–185. doi: 10.1038/nclimate1354.
- Balunas, M. J. and Kinghorn, A. D. (2005) 'Drug discovery from medicinal plants', *Life sciences*, 78(5), pp. 431–441.
- Barbier, N. *et al.* (2011) 'Bidirectional texture function of high resolution optical images of tropical forest: An approach using LiDAR hillshade simulations', *Remote Sensing of Environment*, 115(1), pp. 167–179. doi: 10.1016/j.rse.2010.08.015.
- Barbier, N. and Coueron, P. (2015) 'Attenuating the bidirectional texture variation of satellite images of tropical forest canopies', *Remote Sensing of Environment*, 171, pp. 245–260. doi: 10.1016/j.rse.2015.10.007.
- Baret, E. *et al.* (1992) 'Modeled Analysis of the Biophysical Nature of Spectral Shifts and Comparison with Information Content of Broad Bands', p. 11.
- Bastin, J.-F. *et al.* (2014) 'Aboveground biomass mapping of African forest mosaics using canopy texture analysis: toward a regional approach', *Ecological Applications*, 24(8), pp. 1984–2001.
- Basuki, T. M. *et al.* (2013) 'Estimating tropical forest biomass more accurately by integrating ALOS PALSAR and Landsat-7 ETM+ data', *International Journal of Remote Sensing*, 34(13), pp. 4871–4888. doi: 10.1080/01431161.2013.777486.
- Bennett, B. C. (1992) 'Hallucinogenic plants of the Shuar and related indigenous groups in Amazonian Ecuador and Peru', *Brittonia*, 44(4), pp. 483–493.
- Berger, K. *et al.* (2018) 'Evaluation of the PROSAIL Model Capabilities for Future Hyperspectral Model Environments: A Review Study', *Remote Sensing*, 10(2), p. 85. doi: 10.3390/rs10010085.
- Boukili, V. K. and Chazdon, R. L. (2017) 'Environmental filtering, local site factors and landscape context drive changes in functional trait composition during tropical forest succession', *Perspectives in Plant Ecology, Evolution and Systematics*, 24, pp. 37–47. doi: 10.1016/j.ppees.2016.11.003.
- Brando, P. M. *et al.* (2010) 'Seasonal and interannual variability of climate and vegetation indices across the Amazon', *Proceedings of the National Academy of Sciences*, 107(33), pp. 14685–14690. doi: 10.1073/pnas.0908741107.
- Cadotte, M. W., Carscadden, K. and Mirotnick, N. (2011) 'Beyond species: functional diversity and the maintenance of ecological processes and services', *Journal of applied ecology*, 48(5), pp. 1079–1087.
- Candiago, S. *et al.* (2015) 'Evaluating multispectral images and vegetation indices for precision farming applications from UAV images', *Remote sensing*, 7(4), pp. 4026–4047.
- Cardinale, B. J. *et al.* (2007) 'Impacts of plant diversity on biomass production increase through time because of species complementarity', *Proceedings of the National Academy of Sciences*, 104(46), pp. 18123–18128. doi: 10.1073/pnas.0709069104.
- Carter, G. A. and Knapp, A. K. (2001) 'Leaf optical properties in higher plants: linking spectral characteristics to stress and chlorophyll concentration', *American journal of botany*, 88(4), pp. 677–684.
- Clark, M. L. *et al.* (2011) 'Estimation of tropical rain forest aboveground biomass with small-footprint lidar and hyperspectral sensors', *Remote Sensing of Environment*, 115(11), pp. 2931–2942. doi: 10.1016/j.rse.2010.08.029.

- Cohen, W. B. *et al.* (2003) 'An improved strategy for regression of biophysical variables and Landsat ETM+ data', *Remote Sensing of Environment*, 84(4), pp. 561–571. doi: 10.1016/S0034-4257(02)00173-6.
- Couteron, P. (2002) 'Quantifying change in patterned semi-arid vegetation by Fourier analysis of digitized aerial photographs', *International Journal of Remote Sensing*, 23(17), pp. 3407–3425. doi: 10.1080/01431160110107699.
- Couteron, P. *et al.* (2005) 'Predicting tropical forest stand structure parameters from Fourier transform of very high-resolution remotely sensed canopy images', *Journal of Applied Ecology*, 42(6), pp. 1121–1128. doi: 10.1111/j.1365-2664.2005.01097.x.
- Cunningham, S. A., Summerhayes, B. and Westoby, M. (1999) 'EVOLUTIONARY DIVERGENCES IN LEAF STRUCTURE AND CHEMISTRY, COMPARING RAINFALL AND SOIL NUTRIENT GRADIENTS', *Ecological Monographs*, 69(4), pp. 569–588. doi: 10.1890/0012-9615(1999)069[0569:EDILSA]2.0.CO;2.
- Danks, S. (1983) 'Photosynthetic systems: Structure, function, and assembly', in.
- Darvishzadeh, R. *et al.* (2019) 'Analysis of Sentinel-2 and RapidEye for Retrieval of Leaf Area Index in a Saltmarsh Using a Radiative Transfer Model', *Remote Sensing*, 11(6), p. 671. doi: 10.3390/rs11060671.
- Dube, T. *et al.* (2016) 'Remote sensing of aboveground forest biomass: A review', p. 8.
- Enquist, B. J. *et al.* (2007) 'A general integrative model for scaling plant growth, carbon flux, and functional trait spectra', *Nature*, 449(7159), pp. 218–222. doi: 10.1038/nature06061.
- EVANS, J. R. (1998) 'Photosynthetic characteristics of fast- and slow-growing species. Inherent variation in plant growth', *Physiological Mechanisms and Ecological Consequences*. Available at: <https://ci.nii.ac.jp/naid/10018062192/en/>.
- Féret, B. and de Boissieu, F. (2019) 'biodivMapR: An r package for α - and β -diversity mapping using remotely sensed images', p. 7.
- Féret, J.-B. *et al.* (2008) 'PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments', *Remote Sensing of Environment*, 112(6), pp. 3030–3043. doi: 10.1016/j.rse.2008.02.012.
- Féret, J.-B. *et al.* (2017) 'PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle', *Remote Sensing of Environment*, 193, pp. 204–215. doi: 10.1016/j.rse.2017.03.004.
- Féret, J.-B. *et al.* (2019) 'Estimating leaf mass per area and equivalent water thickness based on leaf optical properties: Potential and limitations of physical modeling and machine learning', *Remote Sensing of Environment*, 231, p. 110959. doi: 10.1016/j.rse.2018.11.002.
- Féret, J.-B. *et al.* (2021) 'PROSPECT-PRO for estimating content of nitrogen-containing leaf proteins and other carbon-based constituents', *Remote Sensing of Environment*, 252, p. 112173. doi: 10.1016/j.rse.2020.112173.
- Féret, J.-B. and Asner, G. P. (2014) 'Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy', *Ecological Applications*, 24(6), pp. 1289–1296. doi: 10.1890/13-1824.1.
- Fletcher, S. and Islam, M. Z. (2018) 'Comparing sets of patterns with the Jaccard index', *Australasian Journal of Information Systems*, 22.
- Frazer, G. W., Wulder, M. A. and Niemann, K. O. (2005) 'Simulation and quantification of the fine-scale spatial pattern and heterogeneity of forest canopy structure: A lacunarity-based method designed for analysis of continuous canopy heights', *Forest Ecology and Management*, 214(1–3), pp. 65–90. doi: 10.1016/j.foreco.2005.03.056.
- Gardner, T. A. *et al.* (2009) 'Prospects for tropical forest biodiversity in a human-modified world', *Ecology Letters*, 12(6), pp. 561–582. doi: 10.1111/j.1461-0248.2009.01294.x.
- Gentry, A. H. (1992) 'Tropical Forest Biodiversity: Distributional Patterns and Their Conservational Significance', p. 11.
- Gibbs, H. K. *et al.* (2007) 'Monitoring and estimating tropical forest carbon stocks: making REDD a reality', *Environmental Research Letters*, 2(4), p. 045023. doi: 10.1088/1748-9326/2/4/045023.

- Gillespie, T. W. *et al.* (2008) 'Measuring and modelling biodiversity from space', *Progress in Physical Geography: Earth and Environment*, 32(2), pp. 203–221. doi: 10.1177/0309133308093606.
- Gitelson, A. A. *et al.* (2005) 'Remote estimation of canopy chlorophyll content in crops', *Geophysical Research Letters*, 32(8). doi: 10.1029/2005GL022688.
- Gitelson, A. A. *et al.* (2012) 'Remote estimation of crop gross primary production with Landsat data', *Remote Sensing of Environment*, 121, pp. 404–414. doi: 10.1016/j.rse.2012.02.017.
- Gitelson, A. A., Keydan, G. P. and Merzlyak, M. N. (2006) 'Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves', *Geophysical research letters*, 33(11).
- 'Global Forest Resources Assessment 2020' (2020), p. 16.
- Gómez-Rodríguez, C. *et al.* (2008) 'High-resolution remote-sensing data in amphibian studies: identification of breeding sites and contribution to habitat models', *The Herpetological Journal*, 18(2), pp. 103–113.
- Gonzalez, P. (2001) 'Desertification and a shift of forest species in the West African Sahel', *Climate Research*, 17, pp. 217–228. doi: 10.3354/cr017217.
- Guitet, S. *et al.* (2015) 'Catalogue des habitats forestiers de Guyane', p. 66.
- Haralick, R. M. (1979) 'Statistical and structural approaches to texture', *Proceedings of the IEEE*, 67(5), pp. 786–804. doi: 10.1109/PROC.1979.11328.
- Huang, Z. *et al.* (2004) 'Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis', *Remote Sensing of Environment*, 93(1–2), pp. 18–29. doi: 10.1016/j.rse.2004.06.008.
- Huemmerich, K. F. *et al.* (2005) 'Time-series validation of MODIS land biophysical products in a Kalahari woodland, Africa', *International Journal of Remote Sensing*, 26(19), pp. 4381–4398. doi: 10.1080/01431160500113393.
- Huete, A. R. (2012) 'Vegetation indices, remote sensing and forest monitoring', *Geography Compass*, 6(9), pp. 513–532.
- Hutchinson, G. E. (1957) 'A Treatise on Limnology, Volume 1. Geography, physics and chemistry Author: George Evelyn Hutchinson Publication Date: 19'.
- Jacquemoud, S. *et al.* (2009) 'PROSPECT+SAIL models: A review of use for vegetation characterization', *Remote Sensing of Environment*, 113, pp. S56–S66. doi: 10.1016/j.rse.2008.01.026.
- Jacquemoud, S. and Baret, F. (1990) 'PROSPECT: A model of leaf optical properties spectra', *Remote Sensing of Environment*, 34(2), pp. 75–91. doi: 10.1016/0034-4257(90)90100-Z.
- Joshi, N. *et al.* (2017) 'Understanding “saturation” of radar signals over forests', *Scientific Reports*, 7(1), p. 3505. doi: 10.1038/s41598-017-03469-3.
- Kajisa, T. *et al.* (2009) 'Object-based forest biomass estimation using Landsat ETM+ in Kampong Thom Province, Cambodia', *Journal of Forest Research*, 14(4), pp. 203–211. doi: 10.1007/s10310-009-0125-9.
- Lu, D. (2005) 'Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon', *International Journal of Remote Sensing*, 26(12), pp. 2509–2525. doi: 10.1080/01431160500142145.
- Lu, D. (2006) 'The potential and challenge of remote sensing-based biomass estimation', *International Journal of Remote Sensing*, 27(7), pp. 1297–1328. doi: 10.1080/01431160500486732.
- Lu, D. *et al.* (2012) 'Aboveground Forest Biomass Estimation with Landsat and LiDAR Data and Uncertainty Analysis of the Estimates', *International Journal of Forestry Research*, 2012, pp. 1–16. doi: 10.1155/2012/436537.
- Mackey, B. *et al.* (2020) 'Understanding the importance of primary tropical forest protection as a mitigation strategy', *Mitigation and Adaptation Strategies for Global Change*, 25(5), pp. 763–787. doi: 10.1007/s11027-019-09891-4.
- Malhi, Y. (2012) 'The productivity, metabolism and carbon cycle of tropical forest vegetation', *Journal of Ecology*, 100(1), pp. 65–75. doi: 10.1111/j.1365-2745.2011.01916.x.

- Masemola, C., Cho, M. A. and Ramoelo, A. (2020) 'Towards a semi-automated mapping of Australia native invasive alien Acacia trees using Sentinel-2 and radiative transfer models in South Africa', *ISPRS Journal of Photogrammetry and Remote Sensing*, 166, pp. 153–168.
- Mason, N. W. *et al.* (2005) 'Functional richness, functional evenness and functional divergence: the primary components of functional diversity', *Oikos*, 111(1), pp. 112–118.
- Mazor, T. *et al.* (2013) 'Can satellite-based night lights be used for conservation? The case of nesting sea turtles in the Mediterranean', *Biological Conservation*, 159, pp. 63–72. doi: 10.1016/j.biocon.2012.11.004.
- Mermoz, S. *et al.* (2015) 'Decrease of L-band SAR backscatter with biomass of dense forests', *Remote Sensing of Environment*, 159, pp. 307–317. doi: 10.1016/j.rse.2014.12.019.
- Mitchard, E. T. A. (2018) 'The tropical forest carbon cycle and climate change', *Nature*, 559(7715), pp. 527–534. doi: 10.1038/s41586-018-0300-2.
- Mittermeier, R. A. *et al.* (2011) 'Global biodiversity conservation: the critical role of hotspots', in *Biodiversity hotspots*. Springer, pp. 3–22.
- Mohd Zaki, N. A. and Abd Latif, Z. (2017) 'Carbon sinks and tropical forest biomass estimation: a review on role of remote sensing in aboveground-biomass modelling', *Geocarto International*, 32(7), pp. 701–716. doi: 10.1080/10106049.2016.1178814.
- Myers, N. *et al.* (2000) 'Biodiversity hotspots for conservation priorities', *Nature*, 403(6772), pp. 853–858. doi: 10.1038/35002501.
- Naeem, S. and Wright, J. P. (2003) 'Disentangling biodiversity effects on ecosystem functioning: deriving solutions to a seemingly insurmountable problem', *Ecology letters*, 6(6), pp. 567–579.
- Nagendra, H. and Rocchini, D. (2008) 'High resolution satellite imagery for tropical biodiversity studies: the devil is in the detail', *Biodiversity and Conservation*, 17(14), pp. 3431–3442. doi: 10.1007/s10531-008-9479-0.
- Navarro, J. A. *et al.* (2019) 'Integration of UAV, Sentinel-1, and Sentinel-2 data for mangrove plantation aboveground biomass monitoring in Senegal', *Remote Sensing*, 11(1), p. 77.
- Niinemets, Ü. (1999) 'Research review. Components of leaf dry mass per area - thickness and density - alter leaf photosynthetic capacity in reverse directions in woody plants', *New Phytologist*, 144(1), pp. 35–47. doi: 10.1046/j.1469-8137.1999.00466.x.
- Palmer, M. W. *et al.* (2000) 'Opportunities for long-term ecological research at the Tallgrass Prairie Preserve, Oklahoma', in *Proceedings of the ILTER Regional Workshop: Cooperation in Long Term Ecological Research in Central and Eastern Europe, Budapest, Hungary*.
- Palmer, M. W. *et al.* (2002) 'Quantitative tools for perfecting species lists', p. 17.
- Pan, Y. *et al.* (2011) 'A Large and Persistent Carbon Sink in the World's Forests', *Science*, 333(6045), pp. 988–993. doi: 10.1126/science.1201609.
- Pan, Y. *et al.* (2013) 'The Structure, Distribution, and Biomass of the World's Forests', *Annual Review of Ecology, Evolution, and Systematics*, 44(1), pp. 593–622. doi: 10.1146/annurev-ecolsys-110512-135914.
- Pandit, S., Tsuyuki, S. and Dube, T. (2018) 'Estimating above-ground biomass in sub-tropical buffer zone community forests, Nepal, using Sentinel 2 data', *Remote Sensing*, 10(4), p. 601.
- Peng, Y. and Gitelson, A. A. (2012) 'Remote estimation of gross primary productivity in soybean and maize based on total crop chlorophyll content', *Remote Sensing of Environment*, 117, pp. 440–448. doi: 10.1016/j.rse.2011.10.021.
- Pert, P. L. *et al.* (2015) 'Mapping cultural ecosystem services with rainforest aboriginal peoples: Integrating biocultural diversity, governance and social variation', *Ecosystem Services*, 13, pp. 41–56. doi: 10.1016/j.ecoser.2014.10.012.
- Phillips, O. L. and Lewis, S. L. (2014) 'Evaluating the tropical forest carbon sink', *Global Change Biology*, 20(7), pp. 2039–2041. doi: 10.1111/gcb.12423.

- Ploton, P. *et al.* (2012) 'Assessing aboveground tropical forest biomass using Google Earth canopy images', *Ecological Applications*, 22(3), pp. 993–1003. doi: 10.1890/11-1606.1.
- Ploton, P. *et al.* (2017) 'Toward a general tropical forest biomass prediction model from very high resolution optical satellite images', *Remote Sensing of Environment*, 200, pp. 140–153. doi: 10.1016/j.rse.2017.08.001.
- Poorter, H. *et al.* (2009) 'Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis', *New Phytologist*, 182(3), pp. 565–588. doi: 10.1111/j.1469-8137.2009.02830.x.
- Prober, S. M. *et al.* (2015) 'Plant diversity predicts beta but not alpha diversity of soil microbes across grasslands worldwide', *Ecology letters*, 18(1), pp. 85–95.
- Proisy, C., Coutron, P. and Fromard, F. (2007) 'Predicting and mapping mangrove biomass from canopy grain analysis using Fourier-based textural ordination of IKONOS images', *Remote Sensing of Environment*, 109(3), pp. 379–392. doi: 10.1016/j.rse.2007.01.009.
- Punalekar, S. M. *et al.* (2018) 'Application of Sentinel-2A data for pasture biomass monitoring using a physically based radiative transfer model', *Remote Sensing of Environment*, 218, pp. 207–220. doi: 10.1016/j.rse.2018.09.028.
- Rocchini, D. *et al.* (2010) 'Remotely sensed spectral heterogeneity as a proxy of species diversity: Recent advances and open challenges', *Ecological Informatics*, 5(5), pp. 318–329. doi: 10.1016/j.ecoinf.2010.06.001.
- Rocchini, D. *et al.* (2016) 'Satellite remote sensing to monitor species diversity: potential and pitfalls', *Remote Sensing in Ecology and Conservation*, 2(1), pp. 25–36. doi: 10.1002/rse2.9.
- Rocchini, D. *et al.* (2018) 'Measuring β -diversity by remote sensing: A challenge for biodiversity monitoring', *Methods in Ecology and Evolution*. Edited by F. Parrini, 9(8), pp. 1787–1798. doi: 10.1111/2041-210X.12941.
- Saatchi, S. S. *et al.* (2011) 'Benchmark map of forest carbon stocks in tropical regions across three continents', *Proceedings of the National Academy of Sciences*, 108(24), pp. 9899–9904. doi: 10.1073/pnas.1019576108.
- Schimel, D. *et al.* (2015) 'Observing terrestrial ecosystems and the carbon cycle from space', *Global Change Biology*, 21(5), pp. 1762–1776. doi: 10.1111/gcb.12822.
- Schlemmer, M. *et al.* (2013) 'Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels', *International Journal of Applied Earth Observation and Geoinformation*, 25, pp. 47–54. doi: 10.1016/j.jag.2013.04.003.
- Schmidt, F. A. *et al.* (2017) 'Similar alpha and beta diversity changes in tropical ant communities, comparing savannas and rainforests in Brazil and Indonesia', *Oecologia*, 185(3), pp. 487–498. doi: 10.1007/s00442-017-3960-y.
- Schneider, F. D. (2014) 'Mapping functional diversity from remotely sensed morphological and physiological forest traits', *NATURE COMMUNICATIONS*, p. 12.
- Sinha, S. *et al.* (2015) 'A review of radar remote sensing for biomass estimation', *International Journal of Environmental Science and Technology*, 12(5), pp. 1779–1792. doi: 10.1007/s13762-015-0750-0.
- Sinha, S. K. *et al.* (2020) 'Estimation of leaf area index using PROSAIL based LUT inversion, MLRA-GPR and empirical models: Case study of tropical deciduous forest plantation, North India', *International Journal of Applied Earth Observation and Geoinformation*, 86, p. 102027. doi: 10.1016/j.jag.2019.102027.
- Stein, A., Gerstner, K. and Kreft, H. (2014) 'Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales', *Ecology letters*, 17(7), pp. 866–880.
- Stickler, C. M. and Southworth, J. (2008) 'Application of multi-scale spatial and spectral analysis for predicting primate occurrence and habitat associations in Kibale National Park, Uganda', *Remote sensing of environment*, 112(5), pp. 2170–2186.
- Stoms, D. M. and Estes, J. E. (1993) 'A remote sensing research agenda for mapping and monitoring biodiversity', *International journal of remote sensing*, 14(10), pp. 1839–1860.
- Tanriverdi, C. (2006) 'A review of remote sensing and vegetation indices in precision farming', *J. Sci. Eng.*, 9, pp. 69–76.

- Ten Hoope, J. E. *et al.* (2012) 'Recent shift from forest to savanna burning in the Amazon Basin observed by satellite', *Environmental Research Letters*, 7(2), p. 024020. doi: 10.1088/1748-9326/7/2/024020.
- Tews, J. *et al.* (2004) 'Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures', *Journal of biogeography*, 31(1), pp. 79–92.
- Tilman, D., Isbell, F. and Cowles, J. M. (2014) 'Biodiversity and Ecosystem Functioning', *Annual Review of Ecology, Evolution, and Systematics*, 45(1), pp. 471–493. doi: 10.1146/annurev-ecolsys-120213-091917.
- Tomaselli, V. *et al.* (2013) 'Translating land cover/land use classifications to habitat taxonomies for landscape monitoring: a Mediterranean assessment', *Landscape Ecology*, 28(5), pp. 905–930.
- Townsend, A. R. *et al.* (2011) 'Multi-element regulation of the tropical forest carbon cycle', *Frontiers in Ecology and the Environment*, 9(1), pp. 9–17. doi: 10.1890/100047.
- Tuanmu, M.-N. and Jetz, W. (2015) 'A global, remote sensing-based characterization of terrestrial habitat heterogeneity for biodiversity and ecosystem modelling', *Global Ecology and Biogeography*, 24(11), pp. 1329–1339.
- Turner, M. G. (1989) 'Landscape ecology: the effect of pattern on process', *Annual review of ecology and systematics*, 20(1), pp. 171–197.
- Turner, W. *et al.* (2003) 'Remote sensing for biodiversity science and conservation', *Trends in Ecology & Evolution*, 18(6), pp. 306–314. doi: 10.1016/S0169-5347(03)00070-3.
- Ustin, S. L. *et al.* (1998) 'Estimating Canopy Water Content of Chaparral Shrubs Using Optical Methods', *Remote Sensing of Environment*, 65(3), pp. 280–291. doi: 10.1016/S0034-4257(98)00038-8.
- Valencia, R., Balslev, H. and Paz Y Miño C, G. (1994) 'High tree alpha-diversity in Amazonian Ecuador', *Biodiversity and Conservation*, 3(1), pp. 21–28. doi: 10.1007/BF00115330.
- Vancutsem, C. *et al.* (2021) 'Long-term (1990–2019) monitoring of forest cover changes in the humid tropics', *Science Advances*, 7(10), p. eabe1603. doi: 10.1126/sciadv.abe1603.
- Verhoef, W. (1985) 'Earth observation modeling based on layer scattering matrices', *Remote Sensing of Environment*, 17(2), pp. 165–178. doi: 10.1016/0034-4257(85)90072-0.
- Verhoef, W. *et al.* (2007) 'Unified Optical-Thermal Four-Stream Radiative Transfer Theory for Homogeneous Vegetation Canopies', *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, 45(6), p. 15.
- Vermote, E. *et al.* (2018) 'LaSRC (Land Surface Reflectance Code): Overview, application and validation using MODIS, VIIRS, LANDSAT and Sentinel 2 data's', in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Valencia: IEEE, pp. 8173–8176. doi: 10.1109/IGARSS.2018.8517622.
- Verrelst, J. *et al.* (2015) 'Experimental Sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods—A comparison', *ISPRS Journal of Photogrammetry and Remote Sensing*, 108, pp. 260–272.
- Villéger, S., Mason, N. W. H. and Mouillot, D. (2008) 'NEW MULTIDIMENSIONAL FUNCTIONAL DIVERSITY INDICES FOR A MULTIFACETED FRAMEWORK IN FUNCTIONAL ECOLOGY', *Ecology*, 89(8), pp. 2290–2301. doi: 10.1890/07-1206.1.
- Violle, C. *et al.* (2007) 'Let the concept of trait be functional!', *Oikos*, 116(5), pp. 882–892. doi: 10.1111/j.0030-1299.2007.15559.x.
- Walters, M. and Scholes, R. J. (eds) (2017) *The GEO Handbook on Biodiversity Observation Networks*. Cham: Springer International Publishing. doi: 10.1007/978-3-319-27288-7.
- Wang, J. *et al.* (2019) 'Estimating leaf area index and aboveground biomass of grazing pastures using Sentinel-1, Sentinel-2 and Landsat images', *ISPRS Journal of Photogrammetry and Remote Sensing*, 154, pp. 189–201. doi: 10.1016/j.isprsjprs.2019.06.007.
- Wang, R. and Gamon, J. A. (2019) 'Remote sensing of terrestrial plant biodiversity', *Remote Sensing of Environment*, 231, p. 111218. doi: 10.1016/j.rse.2019.111218.

- Wang, Y. and Fang, H. (2020) 'Estimation of LAI with the LiDAR Technology: A Review', *Remote Sensing*, 12(20), p. 3457. doi: 10.3390/rs12203457.
- Westoby, M. *et al.* (2002) 'Plant Ecological Strategies: Some Leading Dimensions of Variation Between Species', *Annual Review of Ecology and Systematics*, 33(1), pp. 125–159. doi: 10.1146/annurev.ecolsys.33.010802.150452.
- Whittaker, R. H. (1972) 'Evolution and measurement of species diversity', *Taxon*, 21(2–3), pp. 213–251.
- Wright, I. J. *et al.* (2004) 'The worldwide leaf economics spectrum', *Nature*, 428(6985), pp. 821–827. doi: 10.1038/nature02403.
- Zhang, Q. *et al.* (2005) 'Estimating light absorption by chlorophyll, leaf and canopy in a deciduous broadleaf forest using MODIS data and a radiative transfer model', *Remote Sensing of Environment*, 99(3), pp. 357–371. doi: 10.1016/j.rse.2005.09.009.
- Zhao, B., Zhong, Y. and Zhang, L. (2016) 'A spectral–structural bag-of-features scene classifier for very high spatial resolution remote sensing imagery', *ISPRS Journal of Photogrammetry and Remote Sensing*, 116, pp. 73–85. doi: 10.1016/j.isprsjprs.2016.03.004.
- Zheng, G. and Moskal, L. M. (2009) 'Retrieving Leaf Area Index (LAI) Using Remote Sensing: Theories, Methods and Sensors', *Sensors*, 9(4), pp. 2719–2745. doi: 10.3390/s90402719.
- Zhengxing, W., Chuang, L. and Alfredo, H. (2003) 'From AVHRR-NDVI to MODIS-EVI: Advances in vegetation index research', *Acta ecologica sinica*, 23(5), pp. 979–987.
- Zolkos, S. G., Goetz, S. J. and Dubayah, R. (2013) 'A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing', *Remote Sensing of Environment*, 128, pp. 289–298. doi: 10.1016/j.rse.2012.10.017.