UCLouvain

AGRO LOUVAIN

Faculté des bioingénieurs

Cartographie détaillée des types de formations forestières dans le bassin du Congo par télédétection optique (Sentinel-2)

Concessions forestières de la Congolaise Industrielle des Bois (CIB-OLAM), Ouesso, République du Congo

Auteur :	Alexis del Marmol
Promoteurs :	Pierre Defourny (UCLouvain/ELI/ELIE)
	Sophie Bontemps (UCLouvain/ELI/ELIE)
Lecteurs :	Quentin Ponette (UCLouvain/ELI/ELIE)
	Vincent Istace (CIB-OLAM)
Année Académique :	2018 - 2019
Bioingénieur :	gestion des forêts et des espaces naturels

REMERCIEMENTS

Mes remerciements vont tout d'abord au Prof. Pierre Defourny, mon co-promoteur, qui m'a toujours poussé à donner le meilleur de moi-même, m'a transmis sa motivation pour le progrès des technologies géomatiques et toutes les incroyables applications qu'elles permettent. C'est également lui qui m'a donné l'opportunité de réaliser un voyage au Congo qui fut une source de découvertes, de rencontres et d'apprentissages professionnels. Ils vont également à ma copromotrice Sophie Bontemps qui a pu m'apporter son expertise sur les technologies de traitement d'images Sentinel-2 et qui m'a conseillé dans les démarches d'analyse numérique à suivre.

Ce mémoire put en grande partie être réalisé grâce à mon séjour à Pokola. Pour cela, je souhaite remercier tout particulièrement Vincent Istace pour son encadrement, ses conseils et ses indications, mais surtout pour sa disponibilité et pour les longues conversations que nous avons eues et qui m'ont beaucoup inspiré dans mes questionnements professionnels et personnels. Merci aussi à Romain et Kakao qui m'ont chaque fois accompagné en forêt pour me donner toutes les informations dont j'avais besoin et veilles à ce que je ne me perde pas ! Merci aussi à eux de m'avoir fait découvrir la liane à eau ou « mozambi » grâce à laquelle il est impossible d'être assoiffé en forêt ! Mes remerciements vont aussi aux autres membres de la CIB qui m'ont fait découvrir leur métier et partager le quotidien de leur travail : Patrick pour les bons cafés (double !) du matin à 5 h 30, Cody, Stéphane, Jean-Marie, Ruci et Sarah, les membres de la cellule aménagement et tous les autres qui composent mes souvenirs de ce voyage..

Merci à Juliette, Élie, Quentin et Julia pour leur aide technique et leurs bons conseils !

Mes parents ont toujours été présents pour me soutenir, me conseiller et m'encourager malgré les nombreux découragements, stress et énervements dont je suis un expert, je tiens encore et toujours à les remercier du fond du cœur !

Pour leurs bons conseils, leur soutien au quotidien, mes pensées vont à mes amis, à mes cokotteurs et en particulier à Maxime et Claire pour la relecture attentive et leurs conseils rédactionnels précieux ! Merci les potes ;)

Je remercie vivement également l'Académie Royale des Sciences d'Outre-Mer (ARSOM) pour avoir mis leur confiance dans mon travail au travers d'une généreuse bourse qui a permis le financement d'une bonne partie du séjour de terrain.

Table des matières

Introduction **1**

1 Synthèse bibliographique **2**
1.1 Les forêts tropicales humides . . . 2
1.1.1 Généralités . . . 2
1.1.2 Les types de forêts . . . 3
1.1.3 La structure des forêts . . . 3
1.1.4 Typologie forestière . . . 5
1.2 L'observation des forêts par la télédétection . . . 6
1.2.1 Les technologies de télédétection des forêts tropicales . . . 6
1.2.1.1 Les systèmes passifs . . . 7
1.2.1.2 Les systèmes actifs . . . 7
1.2.2 Les différentes résolutions spatiales et leurs applications . . . 8
1.2.2.1 Très haute résolution spatiale . . . 9
1.2.2.2 Haute résolution spatiale . . . 10
1.2.2.3 Moyenne résolution spatiale . . . 10
1.2.2.4 Basse résolution spatiale . . . 11
1.2.2.5 Le potentiel de Sentinel-2 . . . 11
1.3 La classification des types forestiers . . . 12
1.3.1 Réflectance des surfaces . . . 16
1.3.2 Les indices de végétation . . . 16
1.3.3 Les textures en télédétection . . . 17
1.3.4 *Object-based* et *pixel-based* . . . 20
1.3.5 Les méthodes de classification . . . 21
1.3.5.1 Random Forest . . . 22
1.3.6 Études de trois cas en forêt tropicale . . . 22

2 Objectifs **25**

3 Site d'étude **27**

4 Matériel et méthode **30**
4.1 Données satellitaires . . . 30
4.2 Mission de terrain . . . 33
4.3 Traitement des images . . . 35
4.3.1 Correction en L2A et application du masque de nuage . . . 35
4.3.2 Calcul des indices de végétation . . . 37
4.3.3 Segmentation et calcul des features object-based . . . 37
4.3.4 Construction du jeu de données d'entrainement . . . 38
4.3.5 Classification et validation . . . 38

5 Résultats et discussion **42**
5.1 Types forestiers et clé d'interprétation visuelle 42
5.1.1 Les forêts adultes 42
5.1.2 Les forêts denses humides semi-décidues secondarisées jeunes . . 44
5.1.3 Les forêts à marantacées 44
5.1.4 Les forêts denses humides sempervirentes à limbalis 46
5.1.5 Les forêts ripicoles 47
5.1.6 Les forêts marécageuses 48
5.1.7 Les forêts à raphia 49
5.1.8 Les zones herbacées 50
5.2 Analyse de la qualité du prétraitement 52
5.3 Variables utilisées pour la classification 58
5.3.1 Mise en évidence des bandes les plus discriminantes 58
5.3.2 Calcul des indices de végétation 58
5.3.3 Résultats de la segmentation 60
5.4 Comparaisons de classifications finales 61
5.5 Recommandations 76

Conclusion et perspectives **79**

Annexe A Images Pléiades utilisées pour la mission de terrain **85**

Annexe B Images S2 utilisées durant la mission de terrain **90**

Annexe C Matrices de confusion des différentes combinaisons **102**

Table des figures

1 Exemple de distribution des classes de diamètre pour des populations d'arbres en forêts tropicales humides 4
2 Comparaison des bandes spectrales de Landsat 7 et 8 avec Sentinel-2 . . . 13
3 Vue d'ensemble du Land Cover Classification System (LCCS), de ses deux phases et de ses classificateurs (DI GREGORIO *et al.*, 1997) 14
4 Comparaison d'une classification d'occupation du sol du continent africain à partir de 8 images composites entre 1982 et 1983 et d'une classification à 20m de résolution réalisée par l'Université catholique de Louvain (UCLouvain) en 2015 - 2016 à partir d'image Sentinel-2 15
5 Différents types de formation forestière sont reconnaissables par leurs textures différentes 18
6 Situation géographique des cinq concessions forestières de la Congolaise Industrielle des Bois (CIB) 27
7 Couverture de la zone d'étude par les tuiles des images de Sentinel-2 (S2) . 28
8 Positionnement des prises de vue des satellites S2 en fonction de leur passage sur leurs orbites relatives au-dessus de la zone d'étude. 31
9 Évolution du normalized difference vegetation index (NDVI) et de la probabilité d'avoir une image sans nuages dans les forêts autour d'Ouesso de façon hebdomadaire au cours de l'année (http://maps.elie.ucl.ac.be/CCI/viewer/, générée le 25 mai 2019). 32
10 Répartition des placettes échantillonnées durant la mission de terrain . . . 34
11 Résumé des étapes suivies dans ce mémoire pour le traitement d'images L1C vers l'obtention de classifications 35
12 Comparaison des histogrammes des valeurs de réflectances des bandes à 20 m de résolution d'image traitée en L2A par l'algorithme *Multi-sensor Atmospheric Correction and Cloud Screening* (MACCS) de Sen2-Agri et par l'algorithme Sen2Cor. 36
13 Angles de calcul des textures selon la méthode d'Haralick 38
14 Sous-bois relativement ouvert d'une FA 43
15 Schéma de profil type de FA 43
16 FA vue par S2 selon la composition colorée RGB = B06 B12 B08. 43
17 Sous-bois dense d'une FJ 44
18 Schéma de profil type de FJ 45
19 FJ vue par S2 selon la composition colorée RGB = B06 B12 B08. 45
20 Enchevêtrement extrêmement dense de plantes marantacées 45
21 Schéma de profil type de FM 46
22 FM vue par S2 selon la composition colorée RGB = B06 B12 B08. 46
23 Vue d'un sous-bois en peuplement monospécifique de limbalis 47
24 Schéma de profil type de FL 47
25 FL vue par S2 selon la composition colorée RGB = B06 B12 B08. 47

26 Couverture de lianes euphorbiacées en FR 48
27 Schéma de profil type de FR 48
28 FR vue par S2 selon la composition colorée RGB = B06 B12 B08. 48
29 Sous-bois relativement dense de M 49
30 Schéma de profil type de M 49
31 M vue par S2 selon la composition colorée RGB = B06 B12 B08. 49
32 Sol inondé d'une R 50
33 Schéma de profil type de R 50
34 R vue par S2 selon la composition colorée RGB = B06 B12 B08. 50
35 Clairière inondée couverte de plantes herbacées au cœur de la forêt 51
36 Schéma de profil type de ZH 51
37 ZH vue par S2 selon la composition colorée RGB = B06 B12 B08. 51
38 Mosaïque des images S2 prises le 22 décembre 2018 sur l'orbite relative 50 et le 24 décembre 2018 sur l'orbite relative 7 52
39 Mosaïques des images S2 prises le 22 décembre 2018 sur l'orbite relative 50 et le 24 décembre 2018 sur l'orbite relative 7 pour différentes bandes S2 53
40 Comparaison des fonctions de réponse spectrale normalisées à un des satellites S2A (traits continus) et S2B (traits pointillés) (ESA, 2017) 54
41 Classification sur la base des bandes 6 à 12, du normalized difference infrared Index (NDII) et du *new vegetation index* (NVI)1. On observe une présence élevée de forêts à limbali dans le Nord et une présence très élevée de forêt marécageuse autour de Pokola et d'Ouesso. 55
42 Comparaison des bandes 1 (L1C) et 4 (L2A par Sen2cor) prises le 22 décembre 2018 sur l'orbite relative 50 56
43 Comparaison des mosaïques des bandes 4 S2 prises le 22 décembre 2018 et le 7 mars 2018 sur l'orbite relative 50 traitées en L2A par Sen2cor. 57
44 Bande 11 de la prise de vue du 7 mars 2018 57
45 Signatures spectrales des différents types de forêts selon (a) la représentation classique et (b) une représentation en bâtonnets 59
46 Zoom sur certains polygones obtenus par segmentation des features pixel-based retenues pour les classifications. 61
47 Comparaison des classifications de la C2 par objets (a) et de la C5 par classe majoritaire (b). 66
48 Comparaison des classifications de la C3 par pixels (a) et par classe majoritaire (b) 67
49 Présentation des classifications de la C4 par classe majoritaire (a) et par objets (b) qui ressortent comme les deux meilleures classifications de cette étude 68
50 Comparaison des classifications de la C1 par pixels (a) et de la C2 par pixels (b) 69

51 Présentation des classifications de la C6 par pixels (a) et de la C7 par pixels (b) 71
52 Comparaison des classifications de la C2 par objets (a) et de la C12 par objets (b) 73
53 Résultat du vote majoritaire réalisé sur les combinaisons C1, C3, C4, C5 et C8 par classe majoritaire et C2 et C12 par polygone 75
A.1 Image pléiade en fausse couleur de l'assiette annuelle de coupe de Pokola 2017 86
A.2 Image pléiade en fausse couleur de l'assiette annuelle de coupe de Pokola 2018 87
A.3 Image pléiade en fausse couleur de l'assiette annuelle de coupe de Pokola 2017 88
A.4 Image pléiade en fausse couleur d'une zone intéressante au nord de la zone de développement rurale de Pokola 89
B.1 Image S2 au dessus de l'assiette annuelle de coupe Pokola 2017. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge. Le quadrillage noir représente les parcelles d'exploitation. 91
B.2 Image S2 au dessus de l'assiette annuelle de coupe Pokola 2018. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge. Le quadrillage noir représente les parcelles d'exploitation. 92
B.3 Image S2 au dessus de l'assiette annuelle de coupe Pokola 1980. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge. 93
B.4 Image S2 au dessus de la route nationale entre Pokola et Loundoungou. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge. 94
B.5 Image S2 au dessus de l'assiette annuelle de coupe Loundoungou 2017. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge. Le quadrillage noir représente les parcelles d'exploitation. 95
B.6 Image S2 au dessus de l'assiette annuelle de coupe Loundoungou 2018. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge. Le quadrillage noir représente les parcelles d'exploitation. 96
B.7 Image S2 proche du village de Kabo (à gauche). Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge. 97

B.8 Image S2 au dessus de l'unité forestière d'aménagement de Kabo. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en bleu et la limite des UFA est en rouge. 98
B.9 Image S2 au dessus de l'unité forestière d'aménagement de Loundoungou. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en bleu et la limite des UFA est en rouge. 99
B.10 Image S2 au dessus de la route transversale qui relie Kabo à Mokobo. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge. 100
B.11 Image S2 au dessus du camp forestier de Berandjokou. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge. 101

Liste des tableaux

1 Matrice de corrélation des bandes spectrales de la prise de vue du 7 mars 2018 de l'orbite relative 50 sur la tuile 33NXB 60
2 Différentes combinaisons de *features* essayées pour la réalisation d'une classification 62
3 Matrice de confusion de la C2 par objets. Les classes non végétales sont grisées. 63
4 *overall accuracy* (OA) et κ calculés à partir des matrices de confusions de trois types de produits pour les 13 combinaisons 63
5 Matrice de confusion de la C5 par classe majoritaire 67
6 Matrice de confusion du résultat de classification consolidé par vote majoritaire 76
C.1 Matrice de confusion de la C1 par pixels 103
C.2 Matrice de confusion de la C1 par classe majoritaire 103
C.3 Matrice de confusion de la C1 par objets 104
C.4 Matrice de confusion de la C2 par pixels 104
C.5 Matrice de confusion de la C2 par classe majoritaire 105
C.6 Matrice de confusion de la C2 par objets 105
C.7 Matrice de confusion de la C3 par pixels 106
C.8 Matrice de confusion de la C3 par classe majoritaire 106
C.9 Matrice de confusion de la C3 par objets 107
C.10 Matrice de confusion de la C4 par pixels 107
C.11 Matrice de confusion de la C4 par classe majoritaire 108
C.12 Matrice de confusion de la C4 par objets 108
C.13 Matrice de confusion de la C5 par pixels 109
C.14 Matrice de confusion de la C5 par classe majoritaire 109
C.15 Matrice de confusion de la C5 par objets 110
C.16 Matrice de confusion de la C6 par pixels 110
C.17 Matrice de confusion de la C6 par classe majoritaire 111
C.18 Matrice de confusion de la C6 par objets 111
C.19 Matrice de confusion de la C7 par pixels 112
C.20 Matrice de confusion de la C7 par classe majoritaire 112
C.21 Matrice de confusion de la C7 par objets 113
C.22 Matrice de confusion de la C8 par pixels 113
C.23 Matrice de confusion de la C8 par classe majoritaire 114
C.24 Matrice de confusion de la C8 par objets 114
C.25 Matrice de confusion de la C9 par pixels 115
C.26 Matrice de confusion de la C9 par classe majoritaire 115
C.27 Matrice de confusion de la C9 par objets 116
C.28 Matrice de confusion de la C10 116
C.29 Matrice de confusion de la C11 117

C.30 Matrice de confusion de la C12 117
C.31 Matrice de confusion de la C13 118

Liste des Acronymes

ARSOM Académie Royale des Sciences d'Outre-Mer

AVHRR *Advanced Very High Resolution Radiometer*

CCI *Climate Change Initiative*

CDB Convention sur la Diversité Biologique

CIB Congolaise Industrielle des Bois

COMIFAC Commission des forêts d'Afrique centrale

CCNUCC Convention-cadre des Nations Unies sur les changements climatiques

DT *decision tree*

EFIR exploitation forestière à impact réduit

Envisat *Environment Satellite*

ESA *European Space Agency*

EVI *enhanced vegetation index*

FAO *Food and Agriculture Organisation*

FOTO *fourier transform textural ordination*

FSC *Forest Stewardship Council*

GLCM *gray level co-occurence matrix*

GLDV *gray level difference vector*

LCCS Land Cover Classification System

LCML Land Cover Meta-Language

lidar *light detection and ranging*

MACCS *Multi-sensor Atmospheric Correction and Cloud Screening*

MERIS *Medium Resolution Imagery Spectrometer*

MODIS Moderate-resolution Imaging Spectroradiometer

MSI *Multispectral Instrument*

MV maximum de vraisemblance

NASA National Aeronautics and Space Administration

NDII normalized difference infrared Index

NDVI normalized difference vegetation index

NVI *new vegetation index*

OA *overall accuracy*

OFAC Observatoire des forêts d'Afrique centrale

OTB Orfeo Toolbox

PA *producer's Aacuracy*

PROBA-V *Project for On-Board Autonomy - Vegetation*

REDD+ Réduction des Émissions dues à la Déforestation et à la Dégradation forestière

RF *random forest*

S1 Sentinel-1

S2 Sentinel-2

SVM *support vector machine*

THR très haute résolution

UA *user's accuracy*

UCLouvain Université catholique de Louvain

UE Union européenne

UFA unité forestière d'aménagement

INTRODUCTION

Dans un contexte environnemental international où les forêts tropicales sont un enjeu primordial, ce mémoire de fin d'études cherche à apporter un outil de classification cartographique par télédétection multispectrale des types de forêts du bassin du Congo dans la région d'Ouesso (République du Congo). L'outil développé se base sur les nouvelles technologies européennes des plateformes satellitaires S2.

Dans un premier temps, l'étude réalise un bref état des connaissances de quelques concepts de physionomie forestière et leurs applications en forêts tropicales, ainsi que des enjeux les concernant. Celui-ci est suivi d'une introduction aux différentes technologies de télédétection existantes pour l'observation de ces forêts avec un point d'attention sur les satellites S2. Enfin, un état de l'art de la classification par télédétection appliquée aux types forestiers est réalisé. Cet état de l'art met en exergue les caractéristiques des images multispectrales ainsi que les méthodes utilisées dans la littérature pour réaliser ce genre de classification.

Dans un second temps, le contexte et les deux questions de recherche du mémoire sont posés. La première question est de savoir s'il est possible de réaliser une interprétation visuelle satisfaisante de types de forêts à partir de compositions colorées des images S2. La seconde question est de comprendre quelles sont les variables ou *features* utiles à la classification numérique des types de forêts. Ils sont suivis d'une description de la zone d'étude et des enjeux techniques qui lui sont propres.

Ensuite, le support multispectral est décrit avec une justification des choix réalisés. Ces images sont le support d'une mission de terrain dont le protocole de récolte de données est explicité. Elles sont aussi le point de départ du traitement numérique qui mène à la classification.

Enfin, il est répondu aux questions de recherches de ce travail. D'une part, l'analyse des données de terrain permet de répondre à la première question et d'autre part, l'analyse des résultats obtenus tout au long du traitement numérique permet de répondre à la seconde question. Sur base de ces résultats, une série de recommandations sont données comme bases pour d'autres études similaires.

CHAPITRE 1

SYNTHÈSE BIBLIOGRAPHIQUE

Cette revue de la littérature permet de mettre en contexte l'étude réalisée et se divise en trois parties. La première fait un état des lieux des généralités concernant les forêts tropicales, leur physionomie et les typologies existantes. La seconde aborde les questions de télédétection et les technologies utilisées selon leurs domaines d'applications. Enfin, la troisième s'intéresse davantage aux méthodes d'analyses d'images obtenues par ces technologies et à la façon dont ces méthodes sont mobilisées pour réaliser des classifications d'écosystèmes forestiers tropicaux.

1.1. Les forêts tropicales humides

1.1.1. Généralités

Les forêts tropicales humides se développent dans les régions chaudes où les précipitations sont élevées avec des saisons sèches peu, voire pas du tout marquées. En effet, les températures mensuelles moyennes sont toutes supérieures à 18°C et les précipitations annuelles ne passent jamais sous les 1600 mm et peuvent atteindre plus de 5000 mm dans certaines zones. Les forêts tropicales abritent la plus grande richesse en espèces animales et végétales au monde avec un niveau d'endémisme très élevé (ERINJERY *et al.*, 2018 ; OLSON *et al.*, 2001).

Les écosystèmes forestiers présentent des fonctions d'une importance considérable pour la biosphère et le climat à l'échelle mondiale ainsi qu'à l'échelle locale. Les forêts tropicales constituent un intérêt tout particulier par la grande biodiversité qu'elles hébergent et par les surfaces importantes qu'elles couvrent sur la planète (ROSENQVIST *et al.*, 2000). Ces forêts sont au cœur des enjeux du climat actuels et de conservation de la biodiversité. Le massif forestier d'Afrique centrale qui couvre en partie le bassin du fleuve Congo est le deuxième plus grand écosystème forestier au monde après l'Amazonie. Son importance n'est pas seulement écologique, mais également sociale : près de 60 millions de personnes vivent aux dépens directs de cette forêt pour la nourriture, les substances thérapeutiques d'origine naturelle, l'énergie (bois de chauffage), les fibres et tous les autres produits non ligneux. Par l'exportation de ces produits, 40 autres millions de personnes dépendent de cette forêt. Au-delà des matières premières, la forêt est à

l'origine de modes de vies, de liens sociaux et possède une fonction culturelle très importante (de WASSEIGE *et al.*, 2014).

1.1.2. Les types de forêts

La Convention sur la Diversité Biologique (CDB) définit un «type de forêt» comme un groupe d'écosystèmes forestiers de compositions généralement similaires qui peut être différencié des autres groupes semblables par sa composition spécifique, sa productivité, sa structure ou son taux de fermeture de canopée (UNITED NATIONS ENVIRONMENT, 2018). Leur cartographie est particulièrement importante du fait de leur sensibilité et potentiel d'adaptation très variables par rapport aux changements climatiques et aux pressions anthropiques. En milieu forestier tropical, les informations spatiales détaillées et les cartes des types de forêts sont particulièrement rares. La complexité d'accès à la jungle profonde limite fortement les missions de terrain, réduisant ainsi fortement la quantité d'information spatiale de ces milieux. Ces informations sont pourtant spécifiquement intéressantes pour les zones d'intérêt des *reporting* nationaux du programme UN-REDD Réduction des Émissions dues à la Déforestation et à la Dégradation forestière (REDD+) des Nations Unies. Comme son nom l'indique, ce programme vise à réduire les émissions liées à la déforestation et à la dégradation des forêts (ERINJERY *et al.*, 2018 ; LAURIN *et al.*, 2016 ; PHELPS *et al.*, 2010).

1.1.3. La structure des forêts

La structure de la forêt est une des caractéristiques majeures qui définit un type de formation forestière. Elle représente l'agencement des composantes végétales de la forêt dans l'espace et traduit l'hétérogénéité, le tempérament des espèces et la dynamique des espèces végétales. PASCAL (2003) différencie l'étude des paramètres structuraux en forêt tempérée et en forêt tropicale. Les forestiers s'intéressent principalement aux variables de structure totale : la distribution des tiges par classes de diamètres, la densité et la surface terrière sont les trois principales variables. En forêt tropicale, la distribution des classes de diamètres suit l'allure d'une exponentielle décroissante. La différence de diversité végétale entre les écosystèmes tempérés et tropicaux rend possible l'étude séparée de la distribution des espèces dans ces deux milieux. Plusieurs types de distributions sont alors observés. La figure 1 montre une distribution suivant une exponentielle décroissante pour les espèces d'ombre restant sous couvert ou atteignant la canopée à l'âge adulte, une courbe de Gauss d'écart-type élevé pour les essences de lumière et des courbes irrégulières avec des classes en déficit ou en sureffectifs. Elle montre également, au graphe B, des distributions particulières que peuvent suivre certaines espèces. Ces courbes hétérogènes sont liées à la dynamique particulière de ces forêts non plantées et très naturelles : fructification massive et germination importante des plantules aidées

par une période climatique favorable, accumulation d'une espèce dans une classe de diamètre qui reste bloquée sous la canopée... Il est important de prendre en compte le fait que l'espèce figurée par l'histogramme 1 de la figure 1.A n'est représentée que par très peu d'arbres répartis parmi quelques classes de grosseur.

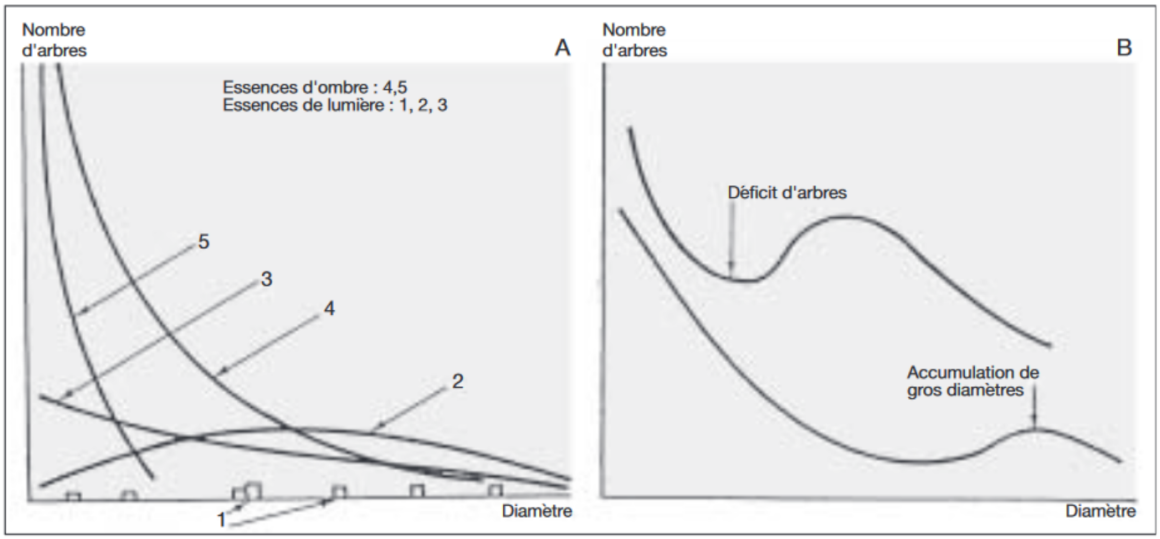


FIGURE 1 – *Exemple de distribution des classes de diamètre pour des populations d'arbres en forêts tropicales humides (*PASCAL*, 2003)*
À : espèces d'ombre et de lumière
B : exemple de distribution particulière

En forêt tropicale, la densité est très variable selon les types de forêts, et dépend des conditions topographiques, hygrométriques et d'expositions locales. On atteint des valeurs allant de 450 à 750 tiges de plus de 10 cm de diamètre pour une moyenne de 600 tiges/ha dans les forêts sans contraintes particulières au niveau des conditions citées précédemment. Cette donnée est un bon indicateur de la dynamique forestière. Enfin, la surface terrière est une des mesures les plus complexes à réaliser à cause de la difficulté à obtenir des sections précises d'arbres possédants des contreforts ou des troncs non circulaires (PASCAL, 2003).

La structure d'un peuplement forestier peut également s'analyser à partir de la spatialisation des éléments de la forêt. Le profil vertical renseigne sur la hauteur du peuplement, son degré de fermeture, la densité et l'étalement des couronnes ainsi que l'étagement de la végétation. Ce dernier peut être un continuum foliaire très différent des profils tempérés où l'étagement est assez net. La structuration horizontale d'un peuplement donne une idée de la distribution des arbres dans l'espace qui semble varier selon la topographie. Les grands arbres, par exemple, sont répartis de façon plus homogène en zone peu accidentée et de façon agrégée en pente accusée (PÉLISSIER, 1997). La structuration horizontale au niveau des espèces traduit les préférences écologiques des espèces au niveau édaphique (information à prendre avec précaution, car d'autres facteurs que les propriétés du sol entrent en jeu pour cela). Il est aussi possible d'étudier

la structuration floristique, richesse et diversité dont les valeurs en milieu tropical sont impressionnantes : en moyenne, 120 à 200 espèces différentes à l'hectare, mais, alors que certaines sont très présentes, d'autres peuvent se trouver qu'une fois en 50 hectares (PASCAL, 2003).

La dynamique forestière est un moteur de la structuration des forêts. Lorsque ce n'est pas la chute d'un arbre qui laisse un trou béant dans la canopée, un chablis, ce sont les dommages physiques sur les feuilles et les branches liés aux maladies, aux espèces herbivores, aux variations météorologiques et saisonnières, à l'altération phénologique naturelle des plantes et à l'intervention de l'homme. L'ensemble de ces effets variant d'une composition forestière à une autre crée la structure du couvert végétal (WIRTH *et al.*, 2001). La structure de la canopée n'est pas seulement une conséquence de la dynamique de la forêt, mais également une cause. En effet, la pénétration du rayonnement solaire vers les strates inférieures de la forêt est directement influencée par l'agencement des arbres, de leurs branches et de leurs feuilles. Une canopée fortement fermée empêche le développement d'un sous-étage dense et ne permet que très peu de régénération là où la fréquence de chablis ou la chute des feuilles laisse se développer une myriade de plantules et de lianes. L'interaction entre la végétation et la lumière est un point développé plus en profondeur dans la suite de ce travail.

Au-delà de la structure de la canopée, la physiographie est une méthode qui permet l'explication de l'organisation spatiale d'une forêt. Celle-ci est liée à certaines conditions physiques du milieu et notamment aux conditions édaphiques et topographiques (GROGAN *et al.*, 2006). Cette information est utile à la caractérisation de ces types de forêts.

1.1.4. Typologie forestière

Malgré l'importance qu'elles représentent pour la cartographie, les typologies des forêts ont rarement fait l'unanimité du fait de la multitude de sémantiques utilisées dans la terminologie forestière. Cette diversité de langages empêche bien souvent les différents acteurs de se comprendre et de pouvoir collaborer efficacement sur les projets de suivi forestier. Elle affaiblit aussi la compréhension du fonctionnement des différents types des forêts et limite les possibilités de comparaisons entre pays (JUNGERS, 2016).

En Afrique Centrale, la cartographie des forêts est un thème abordé depuis des décennies. Il existe aujourd'hui différentes propositions de légende unique dont les plus répandues sont celles dites de « Yangambi » (AUBRÉVILLE, 1957) ou de WHITE (1983). Sur bases de celles-ci, d'autres propositions ont été formulées suivant les évolutions des technologies d'observation, du choix de l'échelle et de la localisation. Il est important de tenir compte de l'augmentation de la capacité d'observation de la diversité végétale suivant l'échelle : une plus petite échelle limite la complexité de la diversité observée,

car elle prend en compte plus d'écosystèmes différents. Ces écosystèmes varient fortement d'un endroit du monde à un autre, ce qui nécessite une adaptation de la légende (JUNGERS, 2016).

1.2. L'observation des forêts par la télédétection

La cartographie, la surveillance et le suivi de l'évolution de ces forêts peuvent appuyer des méthodes de gestion durable des forêts, de conservation des écosystèmes et d'exploitation forestière à impact réduit (EFIR). Ces méthodes peuvent avoir un impact non négligeable sur les conditions de vie de millions de personnes, sur le cycle du carbone et de l'eau à l'échelle mondiale et donc sur les changements climatiques ainsi que sur la préservation d'une grande partie de la biodiversité. Mieux connaitre et comprendre les forêts tropicales est donc un enjeu majeur pour émettre des politiques de gestion alliant l'environnement, le social et l'économie. Face à cet enjeu, la télédétection est un outil puissant.

Aujourd'hui, plusieurs accords internationaux se concentrent sur l'état des forêts tropicales : la Convention-cadre des Nations Unies sur les changements climatiques (CCNUCC) dans laquelle se retrouvent le programme REDD+ et la CDB. Ces accords considèrent les rôles importants joués par les forêts dans le cycle du carbone et dans la richesse des habitats et, par conséquent, de la biodiversité (de WASSEIGE *et al.*, 2014). La caractérisation et le suivi des forêts demandent un accès aux données satellites de haute résolution spatiale et temporelle qui sont adaptées à un suivi régulier de grandes zones (LAURIN *et al.*, 2016).

Cette partie aborde les différentes technologies de télédétection utilisées pour les forêts tropicales ainsi que leurs propriétés, avantages et inconvénients. Une attention particulière est portée à l'utilisation des plateformes Sentinel-2 (S2) dont les données sont utilisées dans cette étude.

1.2.1. Les technologies de télédétection des forêts tropicales

La télédétection de la végétation se fait actuellement principalement sur base des propriétés spectrales de la matière en tenant compte des propriétés biophysiques des plantes. Il est bien connu aujourd'hui que la réflectance du spectre électromagnétique dépend du type de plante et du contenu en eau des tissus ou plus précisément des caractéristiques chimiques et structurelles de la surface des organes des feuilles. Par conséquent, le signal obtenu dans ces bandes permet de faire un lien direct avec les caractéristiques biotiques et donc, avant tout, avec le type de végétation et les espèces présentes. Pour certaines applications plus spécifiques, un lien avec la santé, la vigueur des

plantes, la présence d'un stress hydrique ou encore d'une défoliation est possible et, dans une certaine mesure, cela peut même mener à la détermination du contenu en eau, en sucres, en protéines et en pigments.

1.2.1.1 *Les systèmes passifs*

Les capteurs dits «passifs» utilisent la lumière solaire. Celle-ci atteint les surfaces observées et interfère avec elles pour être renvoyée, en partie, vers le capteur qui enregistre l'intensité de cette lumière réfléchie. La réflectance est le rapport entre la lumière solaire reçue par une surface et la lumière renvoyée et peut se mesurer pour différentes longueurs d'onde du spectre électromagnétique.

Un système de télédétection passif multispectral possède plusieurs matrices de capteurs photosensibles à certaines bandes spectrales de largeurs variables qui permettent de spatialiser l'information obtenue dans ces différentes bandes. Un système hyperspectral est, quant à lui, un système capable d'enregistrer une scène dans un nombre élevé de bandes (souvent plus d'une centaine) étroites (moins de 10 nm) et contigües. Le résultat est l'acquisition d'un spectre continu dans une gamme de longueurs d'onde définie, souvent dans le visible ou le proche infrarouge, avec un pas de 10 nm. Cette technologie est surtout utilisée pour connaitre la signature spectrale d'un matériau, d'une surface dans une gamme de fréquences données afin de pouvoir le caractériser (SOCIÉTÉ FRANÇAISE DE PHOTOGRAMMÉTRIE ET TÉLÉDÉTECTION, 2013). Contrairement aux capteurs multispectraux, qui sont capables d'obtenir de l'information sur des bandes larges et relativement éloignées dans le spectre électromagnétique (visible, infrarouge proche, moyen et lointain...), les capteurs hyperspectraux n'utilisent qu'une gamme de longueurs d'onde très réduite. À cause de cette grande spécificité, la technologie hyperspectrale n'est pas applicable à la classification d'occupations du sol. Les capteurs multispectraux sont donc ceux utilisés dans ce travail et seront détaillés dans la section 1.2.2.

1.2.1.2 *Les systèmes actifs*

Les capteurs actifs sont des appareils qui ont la particularité d'émettre le signal qui interagit avec les surfaces pour enregistrer leur réponse par la suite.

Le radar est un de ces types d'appareils qui utilise les ondes électromagnétiques de type polarisées selon les plans verticaux et horizontaux. Selon les propriétés de l'objet observé (pente, rugosité, propriétés diélectriques) et la distance à cet objet, le temps de parcours de l'onde et son atténuation par rapport au signal émis peuvent être mesuré. Ces mesures donnent l'information utilisée ensuite pour caractériser l'occupation du sol. Parmi ces technologies, il existe le radar latéral (Side Looking Aperture Radar - SLAR)

dont la résolution et l'intensité du signal retourné sont directement liées à la distance de l'objet observé. Il existe aussi le radar à ouverture synthétique (Synthetic Aperture Radar - SAR) capable de viser la direction d'émission du faisceau selon différentes positions, et de focaliser le signal reçu par effet Doppler afin d'augmenter la résolution.

Les images récupérées par les systèmes radars sont soumises à des déformations liées à la géométrie du terrain et à l'angle d'incidence du signal par rapport à l'objet. Il est important de prendre en compte ces effets dans le traitement et l'interprétation des données.

Un second type de technologie active est le *light detection and ranging* (lidar). Le lidar consiste en une émission d'un laser vers le sol à partir d'une plateforme aérienne et une réception de ce même signal après réflexion par les objets au sol. Le temps de retour du signal indique la distance de l'objet et donc sa hauteur. Un laser plus étroit donnera une information plus fine (car moins d'obstacles sont rencontrés), mais demandera une fréquence d'émission plus élevée pour obtenir une couverture représentative du terrain. Cette technologie permet l'obtention de modèles numériques de terrains et de surface qui sont de plus en plus utilisés dans des domaines d'étude variés tels que l'archéologie ou la mesure de biomasse forestière. C'est ce qu'ont réalisé LAURIN *et al.* (2014), par exemple, ajouté les données lidar à laser étroit à des mesures hyperspectrales pour améliorer leur modèle de prédiction de la quantité de biomasse aérienne des forêts tropicales par régression. Concrètement, cette étude a montré qu'en combinant ces deux mesures, le R^2 de leur régression de la biomasse prédite par la biomasse mesurée passe de 0,64 à 0,70. Une autre étude a montré l'utilité du lidar pour le calcul de biomasse forestière aérienne en comparant les résultats obtenus avec les méthodes d'estimation par inventaire de terrain (ASNER *et al.*, 2014). Par cette étude, ils ont montré l'importance de la calibration des mesures lidar sur les mesures de terrain afin d'obtenir des résultats pertinents à petite échelle. Les résultats donnent une correspondance de 90% entre les deux méthodes avec une résolution spatiale d'un hectare sur une grande gamme de types de végétation. En plus d'une mesure d'état, il est possible de mettre en évidence la dynamique de cette biomasse et son évolution dans le temps comme l'ont fait DUBAYAH *et al.* (2010) pour des forêts du Costa Rica. Leur étude a mis en évidence la différence de croissance entre les forêts secondaires et les forêts adultes. Les premières ont eu un gain net de 2,08 m en hauteur et 25 Mg/ha de biomasse contre une perte de 0,33 m en hauteur et un gain de 2 Mg/ha pour les forêts adultes.

1.2.2. Les différentes résolutions spatiales et leurs applications

Les capteurs spectraux utilisés pour la télédétection sont définis par quatre grands types de résolutions : la résolution spatiale, la résolution spectrale, la résolution radiométrique et la résolution temporelle.

— La résolution spatiale est souvent assimilée à la taille d'un pixel en coordonnées terrain. Cette résolution est directement liée au nombre de senseurs et à la qualité de l'appareillage optique qui permet de diriger la lumière vers ces senseurs. Plus la résolution d'une image est élevée, plus un même objet visible sur cette image est précis et composé de nombreux pixels. Cette résolution est directement liée au nombre de détecteurs et à la qualité de l'appareillage optique qui permet de diriger la lumière vers ces détecteurs.
— La résolution spectrale est la capacité d'un capteur à différencier des raies du spectre électromagnétique de faibles intervalles de longueur d'onde. Plus les raies sont nombreuses et fines, plus la résolution est élevée.
— La résolution radiométrique est la précision du signal en sortie du capteur. Elle correspond au nombre de bits associés à une valeur de pixel et donc au nombre de valeurs différentes qu'il est possible d'obtenir avec le capteur. Une image codée sur 8 bits, permettra d'obtenir 2^8 soit 256 valeurs différentes et permettra donc de voir plus de nuances qu'une image codée sur 4 bits (2^4 soit 32 valeurs possibles).
— La résolution temporelle est la fréquence à laquelle le capteur reprend une image d'une même zone.

La résolution spatiale est l'une des contraintes principales des capteurs satellitaires à cause de leur éloignement. Cette résolution est ce qui différencie le plus les images disponibles sur le marché de la télédétection aéroportée et satellitaire. Elle est également inverse à la résolution temporelle : les images avec la plus faible résolution spatiale sont mises à jour le plus rapidement. La littérature définit cette résolution en trois niveaux.

1.2.2.1 *Très haute résolution spatiale*

Les normes de grandeurs couramment admises dans la littérature actuelle considèrent la très haute résolution d'une précision inférieure ou égale à cinq mètres.

Ce type d'image est obtenu à l'aide de capteurs numériques embarqués à bord d'avions ou de drones qui peuvent extraire des informations très précises sur les forêts dont, entre autres, les changements subtils dans la canopée. Ces méthodes permettent également d'obtenir des orthophotoplans par photogrammétrie. Chaque mission demande, cependant, une préparation et une exécution coûteuses en temps, en compétence et en matériel qui justifient le coût financier élevé pour des surfaces imagées relativement faibles (LAURIN *et al.*, 2016). La complexité de mise en place de ces missions rend la résolution temporelle particulièrement faible, de l'ordre de plusieurs années à plusieurs dizaines d'années selon la zone et le pays.

Depuis 1999, des missions spatiales commerciales fournissent également des images

de résolution métrique. La constellation des deux satellites Pléiades est un système d'imagerie spatiale très haute résolution (THR) qui permet d'obtenir des images de n'importe quel point du globe en moins de 24 h. Placées sur une orbite héliosynchrone à 694 km d'altitude, ces plateformes embarquent des capteurs capables de saisir des clichés allant jusqu'à 70 cm de résolution. La sensibilité du capteur permet des temps d'exposition suffisamment faibles que pour prendre plus de 500 images par jour et par satellite (GLEYZES *et al.*, 2012). D'autres missions satellitaires permettent d'observer la Terre avec une très haute résolution spatiale : les plateformes IKONOS, GeoEye, QuickBird, WorldView, SPOT 6 et 7... Ces satellites commerciaux fournissent des images «à la demande» à des prix relativement élevés. Cela limite l'utilisation à des zones relativement petites, ce qui n'est pas le cas de la région étudiée dans ce travail. Les satellites de ce type ne réalisent pas de couverture systématique réduisant la résolution temporelle aux besoins des clients.

1.2.2.2 *Haute résolution spatiale*

Une image de haute résolution est une image dont la taille des pixels varie de 10 à 20 m en coordonnées terrain.

Depuis 1986, les principales données utilisées en télédétection spatiale des forêts à cette résolution étaient les données SPOT. L'arrivée des plateformes S2 du programme Copernicus a changé la donne et ouvert de nouvelles possibilités. Les données obtenues par cette mission S2, spécialisée dans l'observation des terres, sont celles utilisées dans la suite de cette étude et les détails sont expliqués ci-après.

1.2.2.3 *Moyenne résolution spatiale*

À partir de 30 m jusque 100 m à 150 m, on parle aujourd'hui de moyenne résolution.

La mission Landsat offre le plus long enregistrement continu du globe. Ce programme, mené par l'agence gouvernementale américaine National Aeronautics and Space Administration (NASA), a commencé dès le début des années 1970. Depuis lors, il a nourri en continu les archives d'images de la Terre permettant ainsi aux scientifiques d'étudier les changements dans les paysages terrestres. Il s'agit d'une source de données temporelles historique, inégalée en qualité, en détail et en couverture. De plus, ces images prises de façon systématique sont gratuites et en libre accès. Cependant, seul le capteur panchromatique donne une haute résolution à 15 m. Les 4 bandes spectrales visibles par le capteur embarqué (bleu, vert, rouge et proche infrarouge) sont disponibles à 30 m de résolution, ce qui place les satellites Landsat parmi les moyennes résolutions. Cette richesse en information a permis aux scientifiques de développer les techniques d'observation des

forêts et le suivi de la déforestation comme l'ont fait HANSEN *et al.* (2013). Cette dernière étude a donné lieu à la cartographie de l'étendue de la couverture forestière mondiale sur la période de 2000 à 2012. Les zones de perte et de gain en surface ont été localisées et quantifiées pour chaque année permettant ainsi d'établir des tendances quantitatives du taux de déforestation à l'échelle mondiale et locale à une résolution spatiale de 30 m. La perte de forêt a été définie comme la perte de canopée forestière à l'échelle du pixel de Landsat et le gain de forêt a été défini comme l'inverse de la perte, c'est-à-dire l'établissement d'une canopée forestière à partir d'un état non forestier. Les travaux de leur observatoire ont amené une connaissance tout à fait nouvelle sur l'évolution des écosystèmes forestiers tropicaux. Ils ont aussi permis de lancer une dynamique de réflexion par rapport à la protection et la gestion de ces espaces forestiers à des échelles difficilement accessibles jusque là. En effet, ils ont pu démontrer une perte de 2,3 millions de kilomètres^2 de forêt pour cause de perturbation sur la période étudiée et un gain de 0,8 million de kilomètres^2.

1.2.2.4 *Basse résolution spatiale*

Au-delà de 150 m, la résolution spatiale est considérée comme basse.

L'instrument Moderate-resolution Imaging Spectroradiometer (MODIS), embarqué à bord des satellites Terra de la NASA puis du satellite Aqua, a aussi permis de cartographier des zones forestières et notamment en milieu tropical en Afrique centrale (GOND *et al.*, 2013) à partir d'une résolution spatiale de 250 m à 1 km. Grâce à la précision de cet outil il a été possible de réaliser des premières cartes de couverture du sol, mais pas une observation précise des écosystèmes forestiers tropicaux. MODIS n'est qu'un exemple. Il est possible de citer également les instruments SPOT-Végétation ou encore *Medium Resolution Imagery Spectrometer* (MERIS) embarqué à bord d'*Environment Satellite* (Envisat).

1.2.2.5 *Le potentiel de Sentinel-2*

La mission S2 du programme européen Copernicus, mis en œuvre par l'Agence Spatiale Européenne, ou ci-après dénommée *European Space Agency* (ESA), est composée de deux plateformes jumelles sur une orbite héliosynchrone à 786 km d'altitude, phasées entre elles de 180° et dont la première a été lancée le 23 juin 2015 et la seconde le 7 mars 2017. Grâce à ses capteurs à haute résolution spatiale, cette mission est une opportunité pour la caractérisation des types de forêts à petite échelle. Les images sont disponibles pour l'utilisateur sous forme d'ortho-image à deux niveaux de traitement : le niveau 1C (L1C) qui reprend les réflectances mesurées au-dessus de l'atmosphère et le niveau 2A où l'application de la correction atmosphérique laisse apparaitre les réflec-

tances au niveau du sol (sous atmosphère). Les produits finaux sont une compilation de « granules élémentaires », aussi appelées « tuiles », de taille fixe de 100 km^2 avec l'information d'une seule orbite. Cette grille de tuiles suit la projection UTM/WGS84 qui divise le globe en 60 zones de 6° de longitude pour 8° de latitude. (DRUSCH *et al.*, 2012 ; LAURIN *et al.*, 2016).

Chacune des plateformes possède un Instrument MultiSpectral ou *Multispectral Instrument* (MSI) à 13 bandes, de 443 à 2190 nm, et possède un cycle de retour au même point nadir (point situé au sol, à l'exacte verticale de la localisation du satellite) de 10 jours. Cette résolution temporelle est réduite à 5 jours à l'équateur lorsqu'on prend en compte les deux plateformes (DRUSCH *et al.*, 2012). Les 13 bandes disponibles sont les suivantes : les bandes du bleu (B02), du vert (B03), du rouge (B04) et du proche infrarouge (B08) sont disponibles à 10 m de résolution ; les bandes *Red Edge* (B05, B06 et B07), du très proche infrarouge (B8A) et les bandes de moyens infrarouges (B11 et B12)sont disponibles à 20 m de résolution ; la bande sensible aux aérosols (B01), celle sensible à la vapeur d'eau (B09) et celle sensible aux cirrus [1] (B10) sont disponibles à 60 m de résolution et sont utilisées pour le traitement de correction atmosphérique qui permet de passer du niveau 1C au niveau 2A. La figure 2 compare les bandes spectrales mesurées par S2 avec celles des plateformes Landsat 7 et 8. Les trois niveaux de bandes pour S2 correspondent aux résolutions spatiales 10, 20 et 60 m. Cette figure révèle clairement la capacité du capteur MSI à obtenir une mesure dans un large panel de bandes relativement étroites. Les caractéristiques très intéressantes de la mission S2 sont la gratuité de ses données ainsi que la couverture systématique mondiale à haute résolution temporelle qui permet un suivi en temps quasi réel des évolutions n'importe où sur le globe. Les trois bandes *Red Edge* du capteur sont particulièrement utiles pour détecter les fines différences dans les pigments chlorophylliens. En effet, la région *Red Edge* du spectre électromagnétique est la région qui est la plus sensible aux variations du contenu en chlorophylle, lui-même directement lié au contenu en azote et donc à la fertilité du sol. DAUGHTRY (2000) ont montré cela en comparant les signatures spectrales de feuilles de plantes à plusieurs niveaux de fertilisation en azote (180, 90, et 0kgN/ha). La réflectance augmente significativement dans la région du *Red Edge* lorsque la teneur en chlorophylle diminue.

1.3. La classification des types forestiers

La classification d'une prise de vue consiste à simplifier la réalité en discrétisant les types d'occupation du sol dans un certain nombre de classes suivant une typologie préalablement établie. Ce traitement de l'image utilise l'information spectrale contenue dans

1. Il s'agit de fins nuages composés de cristaux de glaces présents dans les couches supérieures de l'atmosphère (5000 à 14 000 m d'altitude)

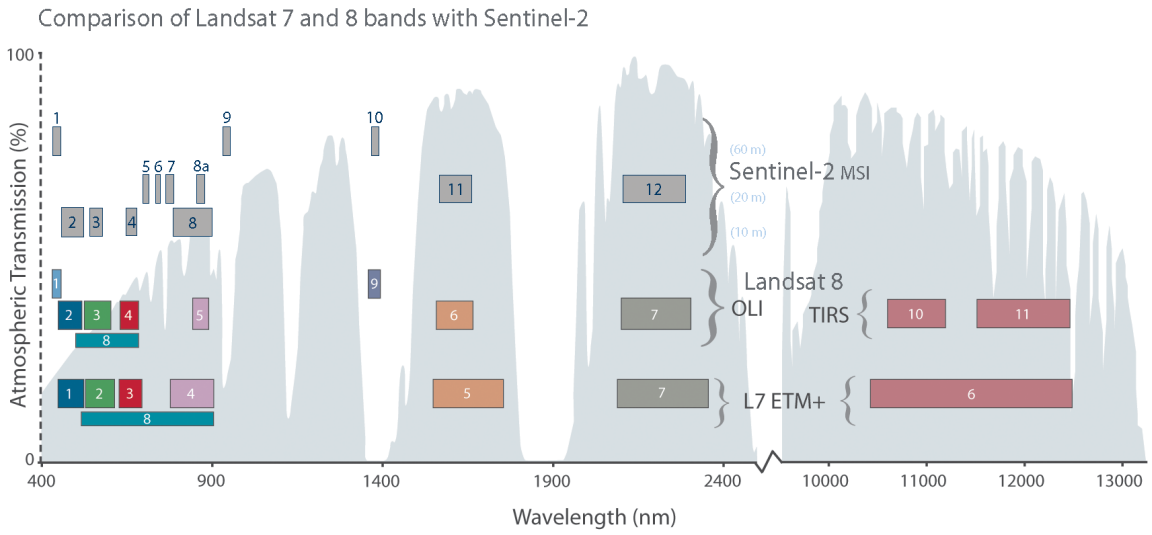


FIGURE 2 – *Comparaison des bandes spectrales de Landsat 7 et 8 avec Sentinel-2 (https://landsat.gsfc.nasa.gov/sentinel-2a-launches-our-compliments-our-complements/)*

les différentes bandes spectrales pour caractériser et discriminer les différents éléments à cartographier. Cette information spectrale peut être synthétisée dans une série de *features* qui peuvent être de différents ordres : le niveau de réflectance directe obtenu dans chacune des bandes, les indices de végétation calculés à partir de ces réflectances ainsi que la texture des éléments. Ces *features* donnent une information spatiale pour la classification. Il est également fréquent de classer non pas une seule image, mais une série temporelle de ces images. L'information temporelle contenue dans ce jeu de donnée est alors utilisée en plus par l'algorithme de classification.

L'occupation du sol est définie comme étant la couverture (bio) physique observée à la surface de la Terre. Cela reflète l'occupation du sol par la végétation présente et les couvertures artificielles. Ces données sont souvent utilisées pour la réalisation d'outils cartographiques. Pour simplifier la complexité du monde réel, des systèmes de classification de ces occupations ont été développés. En 1996, la *Food and Agriculture Organisation* (FAO) apporte sa contribution en proposant la première version de son Land Cover Classification System (LCCS) qui est publié dès 2000 après une période d'essai sur le projet FAO Africover. Une seconde version est développée sur base des retours de la communauté scientifique internationale pour être publiée en 2005. Le principe clé du LCCS est de qualifier une classe d'occupation du sol par une série d'attributs-diagnostiques indépendants. DI GREGORIO *et al.* (2005) explique que cette qualification par attributs se fait en deux étapes : la phase dichotomique et la phase modulaire hiérarchique (Figure 3). La première étape est une clé dichotomique qui permet la classification en 8 grands types d'occupation du sol à partir de 3 classificateurs : la présence de végétation, les conditions édaphiques et l'artificialité de la couverture. À ce stade, la hiérarchisation de ces classificateurs n'a pas d'importance et conduira dans tous les cas aux 8 mêmes classes générales. La deuxième phase de création de classes est donnée par la combi-

naison d'un ensemble de classificateurs prédéfinis «pures» différent pour chacune des 8 classes principales. Ces classificateurs «purs» doivent être utilisés en premier par l'utilisateur, mais celui-ci peut en utiliser plus ou moins selon le niveau de détail voulu. À cet ensemble peut être ajoutée une série d'attributs pour approfondir la définition de la classe. Ces attributs se distinguent en deux types non hiérarchisés : les attributs environnementaux qui influencent la couverture du sol sans être inhérents à celle-ci et les attributs techniques spécifiques, liés à la discipline technique (p. ex. le type de culture en zone cultivée). La troisième version du LCCS utilise le Land Cover Meta-Language (LCML). Cette méthode est une façon de caractériser les éléments du paysage, considérés comme objets (arbre, buisson, bâtiment ...) par une série d'attributs particuliers appelés «méta-éléments» pour construire des bases de données cartographiques. L'utilisation de ce même langage par différents cartographes permet une cohérence et une compréhension entre les multiples classifications (nationales ou autres) qui sont créées grâce à une légende normalisée (DI GREGORIO *et al.*, 2016).

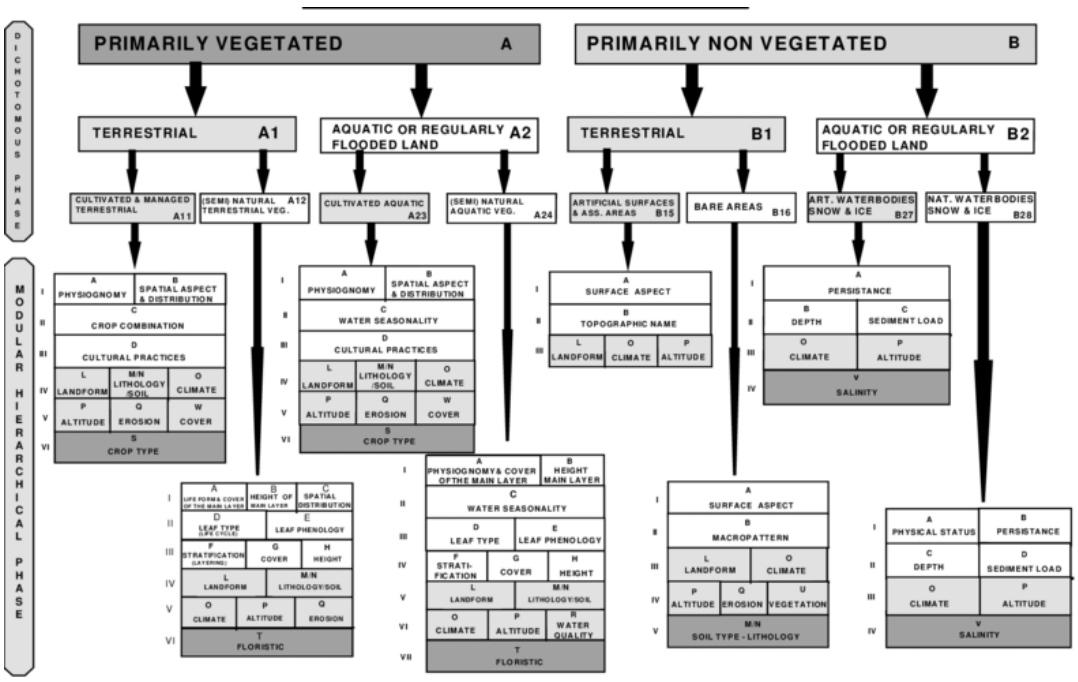
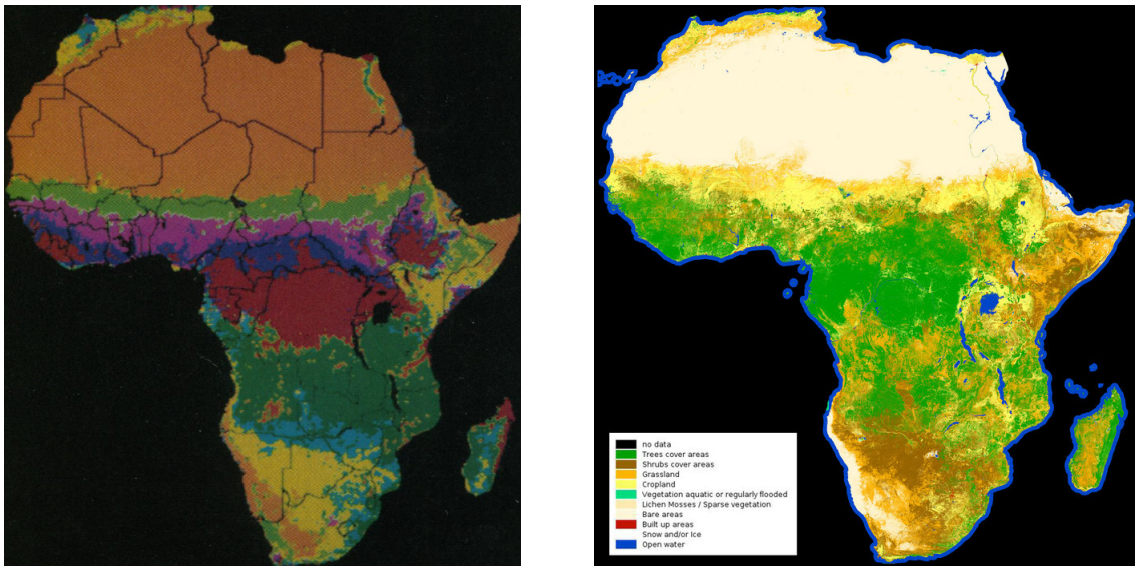


FIGURE 3 – *Vue d'ensemble du LCCS, de ses deux phases et de ses classificateurs (*DI GREGORIO *et al., 1997)*

Les méthodes de classification de l'occupation du sol par télédétection à des échelles continentales ont été théoriquement possibles dès le début des missions Landsat en 1972 (ANDERSON, 1976 ; MAYAUX *et al.*, 2004, 1997). Les premières missions Landsat avaient des périodes de retour au même point nadir de 16 jours, limitant ainsi les chances de contourner la couverture nuageuse souvent importante au-dessus des zones tropicales. Les performances de traitement des images récoltées à 30 m de résolution spatiale étaient relativement limitées et il a longtemps été impossible de réaliser des cartes de la couverture du sol d'un continent entier avec des données récoltées sur une seule année spéci-

fique. Le travail de TUCKER *et al.* (1985) a permis de montrer la faisabilité d'une première carte de l'entièreté de l'Afrique sur une année spécifique en utilisant des données à 4 km de résolution et récoltée journalièrement par la *Advanced Very High Resolution Radiometer* (AVHRR) à bord des satellites en orbite polaire de l'Administration Océanique et Atmosphérique Nationale des États-Unis (Figure 4.a).



(a) *(b)*

FIGURE 4 – *Principaux composants de la classification de l'occupation du sol d'Afrique à partir de huit images composites sur 21 jours (du 12 avril au 2 mai 1982, du 24 mai au 13 juin 1982, du 5 au 25 juillet 1982, du 6 au 26 août 1982, du 27 septembre au 16 octobre 1982, du 8 au 28 novembre 1982, du 20 décembre 1982 au 10 janvier 1983, et du 31 janvier au 19 février 1983) (*TUCKER *et al., 1985) (a). Occupation du sol du continent africain à 20 m de résolution. Cette classification a été réalisée par l'ESA sur base des données récoltées par la plateforme européenne Sentinel-2A durant l'année 2015 à 2016 (b). La légende est accessible sur le visualiseur de l'ESA : http://2016africalandcover20m.esrin.esa.int/viewer.php*

Depuis lors, les technologies ont fortement évolué et le nouveau potentiel des missions européennes permet de réaliser des observations d'une précision dans l'espace et le temps encore inégalée. La couverture pour le monde entier a été publiée par le laboratoire de géomatique de l'Université catholique de Louvain (UCLouvain), dans le cadre du projet *Climate Change Initiative* (CCI) de l'ESA, puis mise en ligne sur une plateforme de visualisation, avec le soutient financier de l'ESA, à d'adresse : http://maps.elie.ucl.ac.be/CCI/viewer/. Cette cartographie mondiale à 300 m de résolution a été créée à partir de séries temporelles composites de 7 jours couvrant une période de 1992 à 2015 et issues des instruments : AVHRR (5 bandes, 1km de résolution spatiale et couverture mondiale de 1992 à 1999), MERIS (15 bandes, 300 ou 1000 m de résolution spatiale et couverture mondiale, de 2003 à 2012) et *Project for On-Board Autonomy - Vegetation* (PROBA-V) (4 bandes à 300 m de résolution spatiale et couverture mondiale, à partir de 2014). Le produit continue d'être mis à jour avec les nouvelles séries temporelles de PROBA-V et avec

des données provenant des instruments de Sentinel-3 (SANTORO *et al.*, 2017). L'évolution du produit permet de mettre en évidence l'évolution des grandes occupations du sol année après année à l'échelle mondiale.

En 2016, l'UCLouvain a publié une nouvelle carte de l'occupation du sol du continent africain à 20 m de résolution sur la base des images composites sans nuages prises par la plateforme Sentinel-2A pendant un an entre 2015 et 2016 (EUROPEAN SPACE AGENCY, 2016; SANTORO *et al.*, 2017). Ce prototype est la donnée à l'échelle du continent la plus précise à ce jour (Figure 4.b). Il ouvre des possibilités de développement de cartes d'occupation du sol mondiales à 20 m de résolution et par conséquent d'un suivi plus précis avec une légende plus détaillée.

1.3.1. Réflectance des surfaces

Les capteurs multispectraux recueillent l'information de la réflectance, la lumière réfléchie par les objets observés. Selon la sensibilité du capteur, la réflectance est mesurée dans l'un ou l'autre longueur d'onde et le niveau de réponse pour chaque longueurs d'onde est indicateur de l'objet observé. Ainsi, la majorité des algorithmes actuels se prêtent aisément au classement d'une étendue d'eau, une surface goudronnée, un champ ou encore une forêt (YUAN *et al.*, 2005). Cependant, la différentiation des espèces végétales ou des compositions d'espèces entre elles demande l'utilisation de méthodes plus complexes. Ainsi, la comparaison des réflectances à différentes longueurs d'onde peut augmenter la puissance de cette discrimination de classes, pour autant que la résolution spatiale et parfois temporelle soit suffisante. En plus de la réflectance des surfaces, il est possible d'utiliser d'autres caractéristiques de l'image décrites ci-après.

1.3.2. Les indices de végétation

En extrayant séparément les informations de réflectances dans les bandes, les résultats obtenus sont souvent limités et trop sensibles à l'hétérogénéité d'une canopée. Ils peuvent ne pas suffire pour la discrimination des types de couverts végétaux. Afin de résoudre ce problème, les bandes sont souvent combinées entre elles pour mettre en évidence des propriétés de la végétation selon les objectifs souhaités. Il existe des dizaines d'indices spectraux utilisés dans la littérature créés selon l'intérêt des chercheurs. Dans le cadre de ce mémoire, les indices de végétation utilisés sont ceux présentés ci-après. Souvent, ils mettent en évidence une caractéristique particulière de la végétation.

L'indice de végétation le plus connu et le plus couramment utilisé est le normalized difference vegetation index (NDVI). Le fait que cet indice différentiel soit normalisé permet de limiter l'impact du type de capteur ainsi que des perturbations résiduelles liées à

l'atmosphère. Cet indice d'une valeur comprise entre -1 et 1 donne une idée de l'état de santé de la végétation en utilisant le rouge (R) et le proche infrarouge (PIR) dont les réflectances sont sensibles à la teneur en chlorophylle dans les feuilles. Il se calcule comme suit :

$$\text{NDVI} = \frac{\text{PIR} - \text{R}}{\text{PIR} + \text{R}}$$

Parmi tous les autres indices existants, le portail www.sentinel-hub.com reprend les principaux utilisables à partir des données récoltées par S2.

On retrouve notamment le normalized difference infrared Index (NDII) qui met en évidence la teneur en eau d'un corps. Il utilise le PIR et le moyen infrarouge (SWIR) à 1600 nm :

$$\text{NDII} = \frac{\text{PIR} - \text{SWIR1}}{\text{PIR} + \text{SWIR1}}$$

En milieu à forte densité de végétation où le Leaf Area Index (LAI) est élevé, la prise en compte du bleu permet de corriger les effets d'humidité atmosphérique. Sur cette base, le *enhanced vegetation index* (EVI) a été développé suivant l'équation :

$$\text{EVI} = \text{G} \cdot \frac{\text{PIR} - \text{R}}{(\text{PIR} + \text{C}_1 \cdot \text{R} - \text{C}_2 \cdot \text{B}) + \text{L}}$$

où G = 2.4, $\text{C}_1 = 6$, $\text{C}_2 = 7.5$ et L = 1000 pour S2

La bande bleue peut parfois poser problème à cause de la pollution électromagnétique souvent présente. Ces longueurs d'onde sont très sensibles aux effets atmosphériques qui ont un effet de diffusion du rayonnement dans l'atmosphère. Sur cette base, un nouvel indice, évolution de l'EVI, n'utilisant que deux bandes a été développé, le EVI 2 :

$$\text{EVI2} = \text{G} \cdot \frac{\text{PIR} - \text{R}}{\text{PIR} + \text{C} \cdot \text{R} + \text{L}}$$

où G = 2.4, C = 1 et L = 1000 pour S2

Ces quelques indices sont les principaux utilisés dans la littérature en milieu forestier tropical.

1.3.3. Les textures en télédétection

Sur une image, la texture est la métrique qui correspond à la granulométrie de la surface, c'est-à-dire la variabilité et la répartition des points de différents niveaux d'intensité et de couleur (Figure 5). Le concept est fondamental dans le processus de classification d'objets et la mise en évidence de régions distinctes. Cependant, la texture est directement liée au type d'occupation du sol et joue un rôle central dans la cartographie numérique de types forestiers.



FIGURE 5 – *Différents types de formation forestière sont reconnaissables par leurs textures différentes (*DEFOURNY*, 2017)*

L'utilisation de cette propriété de l'image lors de la classification numérique nécessite l'usage d'algorithmes d'analyse d'images spécifiques. Plusieurs méthodes ont été développées pour extraire les textures d'une image. Les deux méthodes les plus utilisées en télédétection sont celles des *gray level co-occurence matrix* (GLCM) et des *gray level difference vector* (GLDV) (KHAZENIE *et al.*, 1991, 1992 ; WELCH *et al.*, 1990). Ces méthodes permettent de mettre en évidence une série de variables de texture de l'image dont la moyenne, la corrélation, l'homogénéité, l'entropie, la variance, le contraste, la dissimilarité et le second moment angulaire qui sont alors utilisées comme caractéristiques d'entrées dans les algorithmes de classification. Les noms de ces variables sont liés aux opérations effectuées sur les valeurs des matrices de co-occurrence de niveaux de gris et certaines sont détaillées ci-après. (CHAN *et al.*, 2003 ; KHAZENIE *et al.*, 1992). Une troisième méthode est également de plus en plus utilisée dans la caractérisation des peuplements forestiers par leur texture. Il s'agit de la méthode *fourier transform textural ordination* (FOTO). Enfin, d'autres méthodes telles que la transformation statistique simple, le filtrage de Gabor ou le spectre de texture ont aussi été présentées dans la littérature, mais sont restées marginales dans les études de classification forestière les plus récentes. Elles ne seront pas détaillées dans ce travail.

Ci-dessous le fonctionnement des méthodes GLCM et GLDV est expliqué et illustré par quelques exemples d'études qui ont prouvé leur utilité.

Méthode GLCM Cette technique utilise une image en niveaux de gris et associe à chacun d'entre eux une valeur créant ainsi une matrice. Elle compare ensuite les valeurs selon leurs positions et distances relatives et chacune de ces comparaisons est rentrée dans une nouvelle matrice dite spatio dépendante des niveaux de gris. Les paramètres obtenus sont les résultats de différentes opérations appliquées sur la matrice (HARALICK *et al.*, 1973). Selon COSMAN (2010), pour une image à trois niveaux de gris exprimés sous

forme de matrice (ci-dessous à gauche). Chaque emplacement correspond à un pixel et chaque valeur à un niveau de gris. En appliquant l'opérateur «un pixel à droite et un pixel en bas», on obtient la GLCM C (à droite). La taille de cette matrice vaut le nombre de niveaux de gris existant dans l'image et chacune des valeurs correspond au nombre de fois qu'une co-occurrence de niveau de gris se produit dans la direction donnée par l'opérateur.

$$\begin{matrix} 0 & 0 & 0 & 1 & 2 \\ 1 & 1 & 0 & 1 & 1 \\ 2 & 2 & 1 & 0 & 0 \\ 1 & 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{matrix} \qquad \mathrm{C} = \frac{1}{16} \begin{bmatrix} 4 & 2 & 1 \\ 2 & 3 & 2 \\ 0 & 2 & 0 \end{bmatrix}$$

Cela revient à dire que chaque valeur C_{ij} est le nombre de fois que $\mathrm{F}(x,y) = i$ et $\mathrm{F}(x+1, y+1) = j$. Précisions que, en informatique, le point de référence (0;0) se trouve en haut à gauche et l'axe y est dirigé vers le bas. Le 4 en C_{11} signifie donc qu'on retrouve un 0 en bas à droite d'un autre 0 quatre fois dans l'image. Le facteur $\frac{1}{16}$ permet ici de normaliser les valeurs de la matrice par rapport à la probabilité de co-occurrence sachant qu'il y a 16 paires de pixels. Il est important de tenir compte du fait qu'une image avec plus de contraste (plus de niveau de gris) construit une GLCM plus grande nécessitant plus de temps de calcul.

En pratique, une image complète est traitée par l'algorithme en déplaçant une fenêtre de pixels de dimension choisie par l'utilisateur. Cette fenêtre correspond alors à la matrice de gauche et la valeur de texture calculée est attribuée au pixel central de la fenêtre. Cette dernière se déplace ensuite sur l'image pour la traiter entièrement. À ce stade, il est important de considérer qu'une fenêtre plus grande permet plus de données de calcul, mais risque de perdre en précision si la texture change au sein de la fenêtre. Au lieu d'utiliser une fenêtre de pixel qui se déplace, il est possible de calculer les métriques de texture pour un polygone fixe défini. Le fonctionnement est le même, mais la valeur obtenue est attribuée à l'ensemble des pixels composant le polygone.

Les variables de texture sont obtenues à partir de la matrice C en y calculant différentes métriques. Il s'agit donc bien d'une opération réalisée sur les co-occurrences de niveaux de gris et non pas sur les valeurs directes de niveaux de gris dans la fenêtre. Voici quelques exemples de métriques utilisées dans ce travail et de leur expression (COSMAN, 2010 ; HARALICK *et al.*, 1973) :

— La moyenne (μ)

$$\mu = \frac{\sum_i \sum_j \mathrm{C}ij}{i^2}$$

— Le second moment angulaire (SMA)

$$\mathrm{SMA} = \sum_i \sum_j \mathrm{C}_{ij}^2$$

— La variance (σ^2)

$$\sigma^2 = \sum_i \sum_j (i - \mu)^2 C_{ij}$$

— La corrélation (ρ)

$$\rho = \frac{\sum_i \sum_j (ij) C_{ij} - \mu_x \mu_y}{\sigma_x \sigma_y}$$

Où μ_x,μ_y, σ_x et σ_y sont les moyennes et les écarts types de C_x et C_y.

Méthode GLDV Le vecteur obtenu par cette méthode est la somme des diagonales de la matrice de co-occurrence. Il estime une fonction de densité de probabilité et est basé sur la différence absolue entre des paires de niveaux de gris I et J séparés d'une distance d et selon un angle ϕ dans une direction fixée. La fonction de densité de probabilité du vecteur de différence $P(m)_{d,\phi}$ est, par exemple, définie pour $m = I - J$, où I et J sont les niveaux de gris correspondants, et est obtenue en divisant la fréquence de co-occurrence des niveaux de gris par les fréquences totales (TRIMBLE, 2018 ; WELCH *et al.*, 1990).

À partir de la fonction de densité de probabilité $P(m)_{d,\phi}$, voici quelques exemples d'opération pour extraire les textures (WELCH *et al.*, 1990) :

— μ

$$\mu = \sum_m m P(m)_{d,\phi}$$

— Le SMA

$$SMA = \sum_m \left[P(m)_{d,\phi}\right]^2$$

— L'entropie (S)

$$S = -\sum_m P(m)_{d,\phi} \log P(m)_{d,\phi}$$

1.3.4. *Object-based* et *pixel-based*

L'analyse spatiale d'image peut être abordée de deux façons. La première étudie les caractéristiques spectrales et les textures à l'échelle des pixels (pixel-based) et la seconde se concentre sur ces mêmes caractéristiques, mais à l'échelle de l'objet (object-based). La méthode par objet regroupe des pixels en une seule et même entité afin de simplifier le contenu de l'image. Dans le cas de la classification de forêts, cela permet de rassembler des structures végétales qui paraissent spatialement contigües sur l'image et ainsi de les étudier séparément. Cette simplification peut améliorer le résultat de la classification en calculant les textures à l'intérieur d'un objet, textures qui seront donc bien représentatives d'une formation végétale. Au contraire, cette méthode fait perdre de l'information par la simplification du signal en une valeur appliquée à l'ensemble de l'objet. Ce mémoire met en évidence ces problématiques.

Afin de séparer l'image en objets relativement homogènes du point de vue du signal spectral, il est possible de recourir à un découpage visuel réalisé par l'opérateur ou à une segmentation automatique par un logiciel. Les logiciels utilisés peuvent parfois calculer directement les propriétés de chaque objet, dont les signatures spectrales ou la texture, suivant les principes de HARALICK *et al.* (1973).

Les différentes caractéristiques de réflectance, d'indices de végétation ou de texture obtenues par pixel ou par objet constituent le support pour la classification.

1.3.5. Les méthodes de classification

Une grande variété de méthodes de classifications a été utilisée pour réaliser des cartes d'occupation du sol. Deux grandes approches se distinguent : l'approche supervisée et l'approche non supervisée. La première demande à l'opérateur de définir les classes a priori sur base de connaissances du terrain et d'utiliser ces données en entrée de l'algorithme afin qu'il s'y réfère pour réaliser sa classification. La seconde ne demande pas de connaissance préalable du terrain : l'opérateur fixe un nombre de classes désiré et l'algorithme se charge de déterminer les classes en suivant des règles algorithmiques qui lui sont propres. Parmi les algorithmes non supervisés, il existe l'ISODATA ou le K-means par exemple et le maximum de vraisemblance (MV) est un exemple d'algorithme supervisé. Il existe également des algorithmes de *machine learning* tels que le *support vector machine* (SVM), les réseaux de neurones artificiels, les arbres de décision ou les ensembles de plusieurs méthodes. Les algorithmes de *machine learning* sont apparus comme les plus efficaces et les plus précis en comparaison avec les méthodes paramétriques conventionnelles, notamment pour de larges zones et pour des espaces complexes. Cette efficacité est en partie due au fait que ces méthodes ne se basent pas sur des hypothèses de distribution statistiques des données (p. ex. la normalité). Cependant, les réseaux de neurones ou le SVM demandent un apprentissage avec un important jeu de données. Cela provoque également une demande en temps de calcul très important (RODRIGUEZ-GALIANO *et al.*, 2012).

Récemment, les méthodes qui procèdent par ensemble de classifications ont été mises en avant. Cela consiste à réaliser une série de classifications répétées sur le même jeu de données et de les combiner par des règles mathématiques telles que la majorité, le produit, la somme, le système bayésien ou par une minimisation de l'erreur par itération en réduisant le poids relatif des échantillons correctement classifiés (RODRIGUEZ-GALIANO *et al.*, 2012). La méthode la plus utilisée est le *random forest* (RF).

1.3.5.1 *Random Forest*

Dans le domaine de la classification par télédétection, la méthode du RF est la plus couramment utilisée pour l'occupation du sol. Elle a été développée par BREIMAN (2001) et consiste en l'utilisation d'un ensemble d'arbres de décision qui sont combinés entre eux pour identifier la meilleure classe à attribuer sur la base de la méthode du vote majoritaire.

Une combinaison d'arbres de décision est créée en se basant sur des échantillons d'entrainements aléatoires basés sur les données d'entrainement définies par l'utilisateur. Ils sont tous créés de façon indépendante et les nœuds qui les composent sont séparés suivant un nombre de variables prédéfinies par l'utilisateur. L'ensemble forme ce qu'on appelle la forêt d'arbres de décision. Le modèle construit sur la base des zones d'entrainement peut alors être appliqué à toute l'image. Chaque arbre de décision calcule une probabilité d'attribution d'une classe et chacun d'eux évalue, ensuite, la moyenne des probabilités calculées par l'ensemble des arbres. Se fondant sur ces calculs, ils attribuent leur vote pour le choix de la classe finale qui sera celle ayant obtenu le plus de votes (BREIMAN, 2001 ; GISLASON *et al.*, 2006).

Cette méthode est particulièrement intéressante, car elle possède un faible biais lié au grand nombre d'arbres de décision utilisés. Le faible nombre de variables utilisées pour établir les divisions de chaque nœud des arbres de décision réduit fortement le temps de calcul nécessaire ainsi que la corrélation entre chacun de ces nœuds. En effet, la minimisation de cette corrélation est un argument pour le choix des variables par l'algorithme. Tout cela explique la pertinence de son utilisation dans les études de classification (GISLASON *et al.*, 2006).

La méthode RF présente néanmoins quelques désavantages. Les règles de construction des arbres de décision sont inconnues de l'utilisateur, mais il est possible d'indiquer un ordre général d'importance dans les variables qui discriminent le plus les classes. Ce n'est, cependant, pas possible de réaliser cet ordre d'importance pour chacune des classes et donc d'orienter précisément les choix des arbres de décisions. Enfin, une tendance à effectuer un *overfitting*, c'est-à-dire un ajustement des résultats par rapport aux aires d'entrainement peut limiter l'exactitude.

1.3.6. Études de trois cas en forêt tropicale

CHAN *et al.* (2003) ont tenté de montrer l'utilité des textures pour la classification d'une zone d'étude dans le Nord de la République du Congo à partir d'images Landsat en utilisant trois algorithmes de classification différents : un réseau de neurones artificiels, le SVM et un arbre de décision. Ils expliquent que les deux premiers algorithmes ont

le défaut d'être bien plus chronophages pour l'entrainement que le *decision tree* (DT). Par leur étude, ils ont montré que l'utilisation de 12 textures GLCM et GLDV permet d'augmenter de façon significative (de 10% à 46,7%) la précision de la plus mauvaise classification, celle des forêts exploitées. Les autres classes ont toutes un score supérieur à 80 et l'ajout de texture n'a presque pas d'impact. Cela mène à une faible augmentation de l'indice kappa des matrices de confusion (de 84% à 85%).

L'étude menée par LAURIN *et al.* (2016) visait à comparer l'efficacité de classifications par image hyperspectrale d'un capteur aéroporté avec celle effectuée par image multispectrale S2 simulée à partir d'un découpage en 8 bandes de l'image hyperspectrale. Le but était de voir si les résolutions spatiale et spectrale de S2 peuvent suffire à obtenir une classification de qualité similaire à celle obtenue avec des données hyperspectrales. Cette étude réalisée sur deux forêts au Ghana (forêt humide sempervirente et humide semi-décidue) montre que l'utilisation unique des réponses spectrales dans les différentes bandes en combinaison avec des indices de végétation ne suffit pas à donner un résultat de classification satisfaisant. Au contraire, par leur travail ils ont mis en évidence que l'ajout des variables de texture augmente la séparabilité de façon suffisante pour que le test de validation se révèle concluant et aussi performant que la classification par image hyperspectrale. En effet, la faible résolution spatiale de l'image multispectrale simulée (10 m) ne permettait pas d'obtenir des indicateurs de formes de houppiers et de composition spécifique précise contrairement à l'image hyperspectrale dont la résolution spatiale est à 1 m. Les métriques de texture utilisées dans cette étude sont celles de la moyenne, de la variance, du second moment et de la corrélation calculées à partir des bandes de S2 2, 3 et 5 pour la première zone d'étude (Ankasa) et 2, 3 et 4 pour la seconde zone (Bia). LAURIN *et al.* (2016) précisent que les images S2 réelles possèdent des données supplémentaires à leur simulation : la plupart des bandes infrarouges ainsi que la dimension temporelle de la prise de données ne se trouvent pas dans les données utilisées et peuvent apporter des facteurs discriminants supplémentaires qui pourraient s'avérer très utiles pour la classification. La revue de la littérature publiée par TRANSON *et al.* (2018) conduit à nuancer et critiquer l'étude de LAURIN *et al.* (2016) en démontrant que la technologie hyperspectrale ne répond pas aux attentes des scientifiques qui l'utilisent depuis des décennies. De plus, cette revue met en évidence les qualités du système S2 qui surpasseront probablement les limitations des capteurs hyperspectraux.

Enfin, une troisième étude, récemment publiée par ERINJERY *et al.* (2018) reprend, entre autres, l'étude de LAURIN *et al.* (2016) pour démontrer la difficulté à distinguer les types de végétation par télédétection à cause de l'hétérogénéité, de la grande diversité de types de forêt et de la couverture nuageuse très présente dans certaines régions du monde. Elle tente une classification des types de végétation des forêts tropicales de la chaîne de montagnes des Ghats occidentaux sur la côte ouest de l'Inde en combinant des données multispectrales S2 avec des données radar Sentinel-1 (S1).

Avant de s'intéresser à l'apport des données radar, ERINJERY *et al.* (2018) mettent en évidence l'importance de l'utilisation du NDVI et des textures pour la discrimination des types de forêts. L'étude distingue sept occupations différentes du sol : la forêt humide, la forêt sèche semi-décidue, l'eau, 3 types de plantations (ombragées, en motif, ouvertes) et les zones urbaines. ERINJERY *et al.* (2018) utilisent des images S2 prises en saison sèche, seule période permettant un niveau de couverture nuageuse inférieure à 1 % de l'image, niveau tel que souhaité ici. Ces images sont corrigées de leurs perturbations atmosphériques. Les auteurs calculent également une couche NDVI de ces images ainsi que les couches de textures GLCM de la moyenne et de l'entropie obtenues à partir des deux bandes rouges et proche infrarouge (quatre couches supplémentaires).

L'utilisation des bandes 8, 4, 3 et 2, de la couche de NDVI et des textures de moyenne et d'entropie calculées sur les bandes 4 et 8 leur permet d'obtenir une classification avec un niveau d'exactitude globale supérieur à 75%. Ce résultat est considéré comme élevé et la conclusion de l'étude insiste sur l'importance de l'utilisation des indices de végétation et des textures dérivées dans la classification des types de végétation dans les forêts tropicales de la région étudiée. L'ajout de bandes infrarouges à 20 m de résolution en supplément des quatre bandes utilisées n'améliore que très légèrement l'exactitude et n'est pas utile pour classer les forêts sèches semi-décidues. Dans cette étude, la classification a été réalisée avec les méthodes de MV et des RF à 10 et 100 arbres de décision. Les deux méthodes donnent de bons résultats, mais le MV est légèrement plus précis pour cette étude. L'étude envisage également l'hypothèse que l'exactitude élevée de leur méthode en comparaison à d'autres classifications réalisées avec d'autres données satellitaires est probablement due à la haute résolution spatiale du capteur MSI de S2. Finalement, cette étude montre que l'utilisation des données de S2 est une option viable pour la classification des types forestiers, mais leur méthode peut encore être améliorée en ajoutant la dimension temporelle des données notamment pour réduire les effets de couverture nuageuse.

CHAPITRE 2

OBJECTIFS

Dans une optique d'améliorer les connaissances sur les forêts du bassin du Congo qui couvrent une grande partie de leur territoire, les états membres de la Commission des forêts d'Afrique centrale (COMIFAC) ont créé l'Observatoire des forêts d'Afrique centrale (OFAC) en 2007. Les activités de cet observatoire consistent en la mise à disposition de données actualisées et pertinentes sur les forêts et écosystèmes de la région en vue d'éclairer la prise de décisions politiques et de promouvoir une meilleure gouvernance et une plus grande durabilité de la gestion des ressources naturelles. Elles sont mises en œuvre par un consortium scientifique dont l'UCLouvain fait partie et dont le bailleur principal est l'Union européenne (UE). Ce mémoire s'inscrit dans le contexte de ce projet de collaboration internationalle entre l'Europe et les pays d'Afrique centrale.

L'objectif du travail est d'utiliser les données des plateformes spatiales du programme Copernicus, Sentinel-2 (S2), pour développer une méthode de reconnaissance visuelle et numérique des types de formations forestières dans la région d'Ouesso, dans le nord de la République du Congo, mis en évidence durant une mission de terrain. Le but est la classification de ces types de forêts.

L'étude a été réalisée en collaboration avec la Congolaise Industrielle des Bois (CIB) qui a accueilli la mission de terrain dans ses concessions forestières. Celle-ci a eu pour objectif d'identifier différents types de forêts dans la région d'Ouesso, de les décrire d'un point de vue physionomique et écologique et enfin de les localiser par positionnement GPS. Cet échantillonnage est ensuite utilisé pour la reconnaissance visuelle d'une part et numérique, avec la classification, d'autre part.

Deux questions de recherche sont posées dans ce travail :

— Est-t-il possible d'identifier visuellement les types forestiers sur certaines compositions colorées des images S2 ?

— Quelles sont les variables qui permettent d'obtenir une bonne classification numérique de ces types forestiers ?

La réponse à la première question est donnée par la création d'une clé d'interprétation visuelle des types de forêts. La seconde question trouve sa réponse dans les la comparaison des classifications obtenues par différentes combinaisons de *features* : les différentes bandes spectrales, différents indices de végétation et l'utilisation de valeurs de textures. La mise en évidence de la différence d'une approche par pixel ou par objet

est également un élément de réponse.

En conséquence, ce mémoire permettra de mettre en évidence la capacité de S2 pour l'observation détaillée des forêts tropicales et leur cartographie à grande échelle.

CHAPITRE 3

SITE D'ÉTUDE

La région étudiée dans ce mémoire se situe au nord de la République du Congo, dans le département administratif de la Sangha et plus précisément dans les concessions forestières de la CIB (Figure 6). Situées principalement entre le premier et le troisième parallèle, ces concessions couvrent plus de 2 millions d'hectares et sont couvertes de forêts denses humides. À Ouesso, la moyenne de pluviométrie annuelle est de 1686 mm avec des pics en mai et en octobre et une nette diminution autour du mois de janvier. Les températures moyennes mensuelles varient faiblement autour de 25°C et l'hygrométrie moyenne annuelle est de 85%. Ces informations positionnent la zone dans la région bioclimatique de la cuvette congolaise et dans le domaine de la mousson atlantique permanente (CIB, 2007). Il est important de tenir compte de ces conditions climatiques pour le traitement des images satellites et leur interprétation.

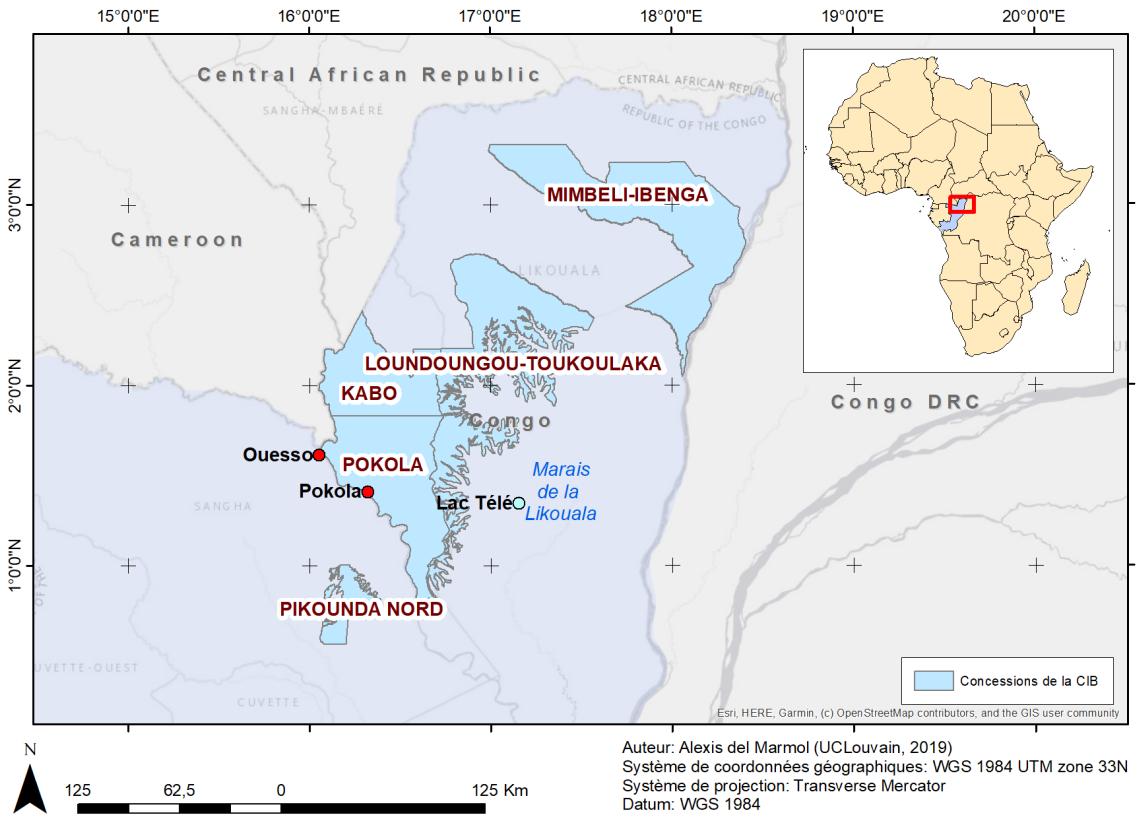


FIGURE 6 – *Situation géographique des cinq concessions forestières de la CIB*

La figure 7 met en évidence que la majeure partie des concessions de la CIB sont couvertes par 7 tuiles de la zone UTM 33 N.

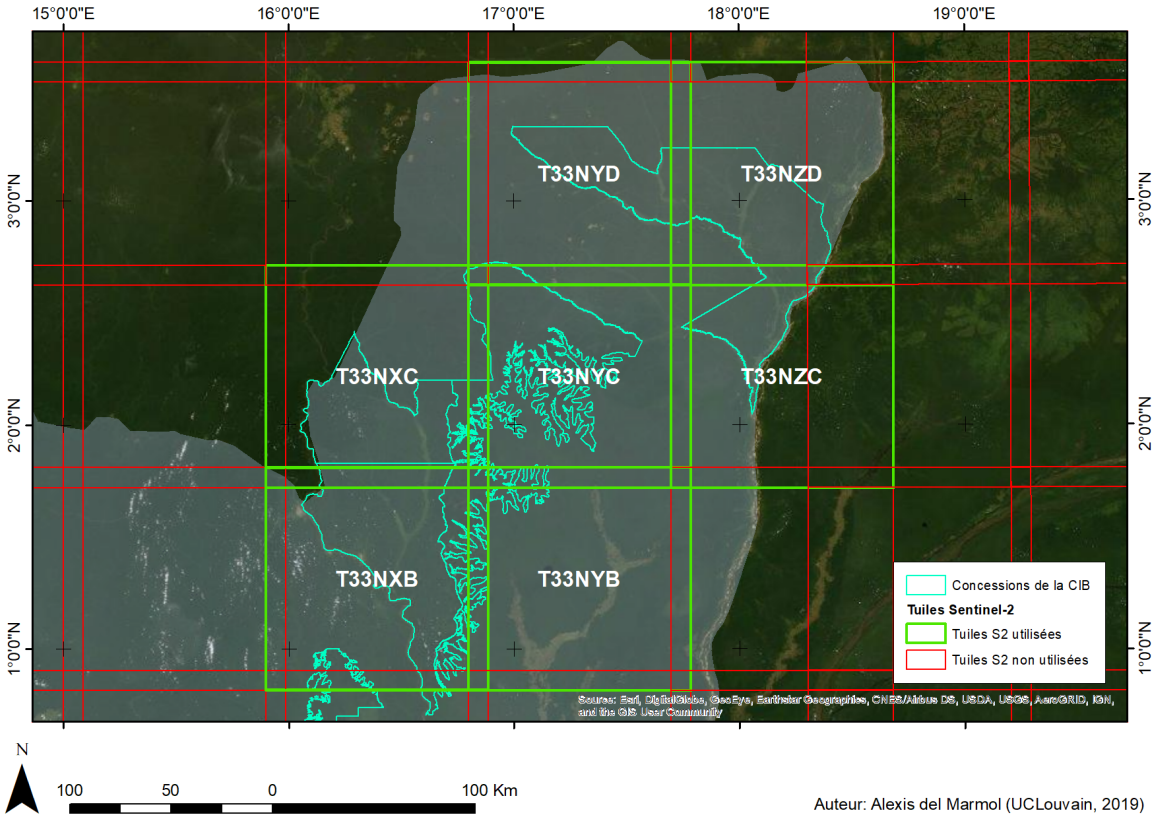


FIGURE 7 – *Couverture de la zone d'étude par les tuiles des images de S2*

La majeure partie du territoire se trouve sur le bassin versant de la Sangha. On y retrouve donc une série de petits affluents qui rejoignent ce fleuve ou la rivière Ndoki qui se jette elle-même dans la Sangha. Malgré un réseau hydrographique bien présent, le relief est relativement faible(CIB, 2007). Cela crée une prédisposition à l'apparition de grands marécages. La présence de la grande zone marécageuse de la Likouala, autour du lac Télé à l'Est (Figure 6), en est la preuve. Les limites sinueuses de ce marécage sont la raison du découpage complexe de l'est de la concession de Loundoungou-Toukoulaka. La région est aussi intéressante pour les larges formations monodominantes de Limbali (*Gilbertiodendron dewevrei*), espèce emblématique des forêts du bassin du Congo. Les forêts de terre ferme sont actuellement exploitées pour leur richesse en espèces commerciales d'intérêt dont le sapelli (*Entandrophragma cylindricum*) est le représentant principal. D'autres espèces telles que le sipo (E*ntandrophragma utile*), l'ayous (*Triplochiton scleroxylon*), l'iroko (*Milicia excelsa*), le bossé (G*uarea cedrata*), le doussié (*Afzelia bipindensis*) et le wengué (*Millettia laurentii*) sont également utilisées pour leurs usages commerciaux nobles.

La CIB est une entreprise d'exploitation forestière appartenant au groupe Olam international. Cette société se démarque par son label *Forest Stewardship Council* (FSC) attribué à trois des quatre concessions exploitées, aussi appelées unité forestière d'aménagement (UFA). À ce jour, il s'agit de la plus grande entreprise d'exploitation forestière en milieu tropicale certifiée FSC au monde avec quelques 395 000 m^3 de bois récoltés en

2017. Près de 90% de ce volume est transformé à Pokola même, où se trouve le siège de l'entreprise et les installations de transformation. Le reste est exporté sous forme brute.

L'entreprise congolaise a l'habitude de travailler en collaboration avec la recherche scientifique (DAÏNOU *et al.*, 2013 ; FORNI *et al.*, 2015 ; KLEINSCHROTH *et al.*, 2018). Le réseau routier, l'infrastructure et la logistique développés pour l'exploitation sont des outils particulièrement précieux pour réaliser des études dans les environnements particulièrement difficiles d'accès que sont les jungles équatoriales. En plus d'être au cœur des forêts concernées par le projet de l'OFAC, le site d'étude est donc particulièrement propice à la réalisation d'une mission de terrain dont l'objectif est de décrire un maximum de types de forêts différents sur une grande zone de façon à constituer une base de données la plus riche possible.

CHAPITRE 4

MATÉRIEL ET MÉTHODE

Afin de réaliser la clé d'interprétation visuelle et les classifications finales pour répondre aux questions de recherches, ce mémoire utilise une série d'images satellites. Ce chapitre décrit et justifie les images choisies pour mener à bien la mission de terrain d'une part et pour la réalisation des classifications d'autre part. Après cette description du matériel utilisé se trouve l'explication détaillée de la méthode suivie pour réaliser au mieux la mission de terrain et pour construire la clé d'interprétation visuelle à partir des données satellitaires prétraitées. Enfin, le chapitre décrit la méthodologie suivie pour obtenir les classifications finales et répondre à la seconde question de recherche.

4.1. Données satellitaires

Cette étude utilise des images satellites des S2. L'ambition du projet est de réaliser une classification sur la région la plus grande possible. Néanmoins, la zone de ce travail est limitée aux tuiles qui comprennent les placettes échantillonnées sur le terrain : 33NXB, 33NXC, 33NYC, 33NYD (Figure 7). En effet, l'objectif de ce travail est de développer une méthode performante sur une zone relativement restreinte, méthode qui pourra être utilisée, adaptée et combinée à d'autres méthodes par la suite pour une généralisation sur de plus grandes zones. Travailler sur une zone restreinte permet principalement de réduire le temps de calcul des différentes opérations. Les fauchées de deux orbites couvrent les tuiles 33NXB et 33NXC : l'orbite 7 et l'orbite 50 (Figure 8). Cela impose l'utilisation d'images composites de prises de vues à deux dates différentes. Ces images ont été le support des deux activités de ce mémoire : la collecte d'information de terrain et la classification.

Pour la mission de terrain, l'utilisation de dates de prises de vues proches les unes des autres pour les différentes tuiles ainsi que la précision des valeurs de réflectance n'ont pas été un enjeu majeur. Les images utilisées sont des images non corrigées au niveau atmosphérique 1C (L1C) prises aux dates suivantes : le 7 juin 2018 pour les zones couvertes par l'orbite relative 7, le 7 mars 2018 pour la zone couverte par l'orbite relative 50 de la tuile 33NXC et le 18 janvier 2018 pour la zone couverte par l'orbite relative 50 de la tuile 33NXB. Ces images ont été choisies parmi les plus proches de la période de la mission qui présentaient une couverture nuageuse la plus faible possible.

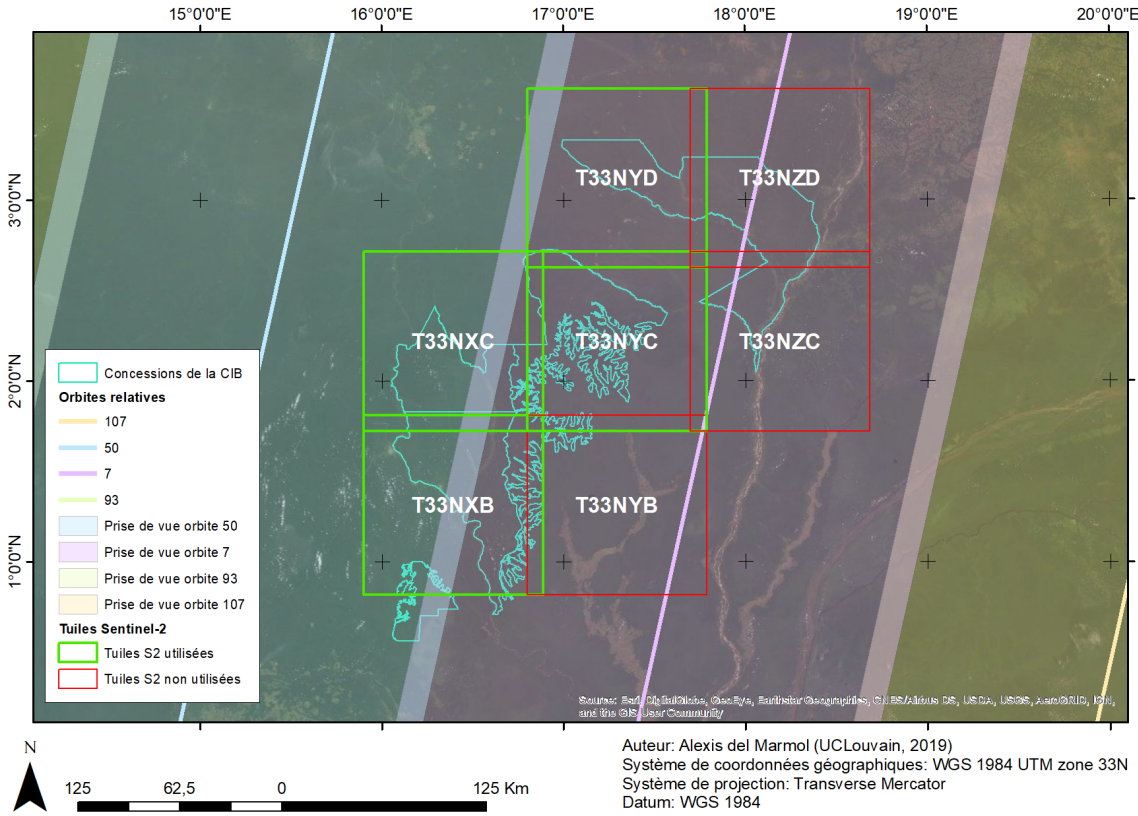


FIGURE 8 – *Positionnement des prises de vue des satellites S2 en fonction de leur passage sur leurs orbites relatives au-dessus de la zone d'étude. Il est possible de distinguer clairement la découpe des tuiles 33NXB et 33NXC suivant les deux prises de vues des orbites relatives 7 et 50.*

Des images optiques des satellites Pléiade ont aussi été commandées pour l'identification des types forestiers en vue d'un déplacement sur le terrain voire d'une classification. La commande effectuée auprès de la compagnie Airbus pour des images d'archives sur des zones indiquées par la CIB comme étant potentiellement intéressantes pour leur richesse floristique et leur accessibilité concerne des images de 2017 à différentes périodes de l'année.

Malgré leur très haute résolution spatiale (bande panchromatique à 0,5 m et bandes multispectrales à 2 m de résolution) permettant de voir jusqu'au houppier de certains arbres, leur prix par unité de surface élevé et leur faible résolution spectrale (4 bandes disponibles : le bleu, le vert, le rouge et le proche infrarouge) ont rapidement mené à la décision de ne pas utiliser ces données pour la seconde partie du travail. En effet, les caractéristiques techniques, le libre accès et la couverture mondiale systématique des images S2 répondent mieux aux attentes. Cependant, les images Pléiades ont été utilisées en complément des images S2 pour le déplacement sur le terrain.

La mission de terrain de ce mémoire se base principalement sur ces images S2 L1C avec l'aide, localement, des images Pléiades prétraitées par une procédure de *pan sharpening* afin d'obtenir les bandes multispectrales à 0,5 m de résolution.

Il a été opté de développer la méthode de classification à partir d'une seule date de prise de vue et ce pour les raisons suivantes. La figure 9, générée par l'outil de visualisation du CCI le 25 mai 2019 (http://maps.elie.ucl.ac.be/CCI/viewer/), montre une probabilité élevée d'obtenir une image composite hebdomadaire sans nuages localement autour de Pokola tout au long de l'année avec une légère baisse durant l'été boréal (pic de pluviométrie annuel). Cependant, les images qui couvrent l'ensemble de la zone d'étude et qui ne comprennent pas de cumulus au-dessus des régions les plus importantes (surfaces échantillonnées) sont particulièrement rares. En effet, des images sans nuages n'ont été prises, depuis le début de la mission S2 (mise en orbite de la plateforme S2A en 2015), que durant les périodes de janvier, mars et juin. En plus de la couverture nuageuse, l'hygrométrie de forêts équatoriales est très élevée et induit des effets atmosphériques qui peuvent perturber fortement la mesure de réflectance et le traitement numérique des données. Enfin, la zone d'étude est exploitée et les trouées réalisées par l'abattage font passer les assiettes annuelles de coupe d'une classe de forêt (forêt adulte) à une autre (forêt secondarisée jeune) sur de très courtes échelle temporelles. Comme certaines placettes ont été échantillonnées dans des zones exploitées tout récemment, travailler avec des images prises plus de quelques semaines avant la mission de terrain risque d'amener des erreurs dans la classification. Ces deux dernières contraintes diminuent encore le nombre de prises de vue utilisables.

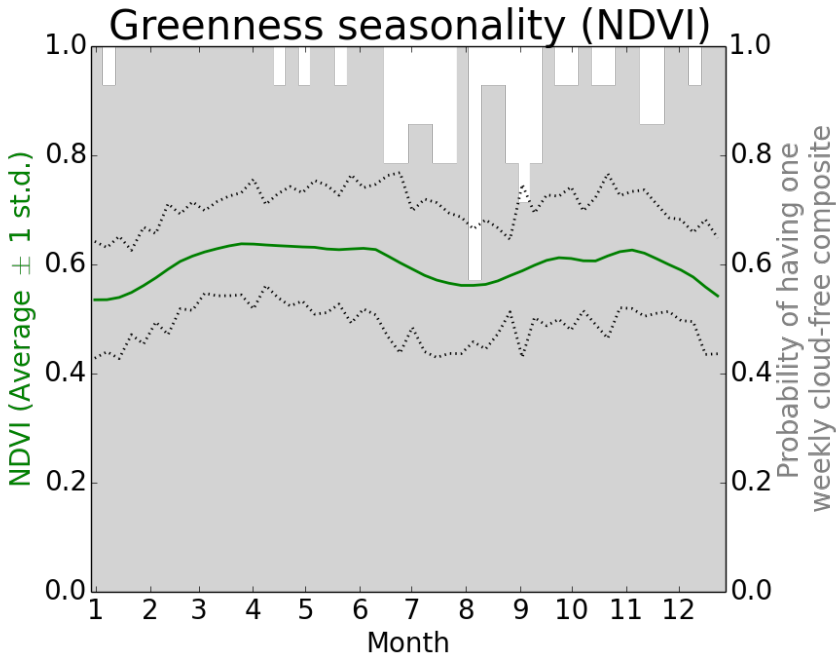


FIGURE 9 – *Évolution du NDVI et de la probabilité d'avoir une image sans nuages dans les forêts autour d'Ouesso de façon hebdomadaire au cours de l'année (http://maps.elie.ucl.ac.be/CCI/viewer/, générée le 25 mai 2019).*

Finalement, les meilleures images ont été trouvées aux dates suivantes : le 24 décembre 2018 pour l'orbite 7 et les 22 décembre, 1er juin et 7 mars 2018 pour l'orbite 50. Ces images ont été choisies pour leur qualité (absence quasi totale de nuages et faible

impact de l'humidité atmosphérique après traitement de niveau 2A - L2A) ainsi que pour leur proximité avec la période durant laquelle la mission de terrain a été réalisée.

4.2. Mission de terrain

Avec l'objectif d'identifier, de positionner et de décrire différents types de formations forestières pour la création d'une clé d'interprétation visuelle des images satellite et pour leur classification, la mission de terrain a été réalisée en quatre temps.

La première étape a consisté en la construction d'un support visuel permettant d'aller sur le terrain par le téléchargement des images satellites L1C de la région (Annexe B). La composition colorée RGB = B06 B12 B08 [RedEdge SWIR2 NIR] a été déterminée comme étant celle qui met en évidence le plus de contraste visuel dans les formations forestières. Les images pléiades obtenues auprès d'Airbus ont été composées en fausse couleur infrarouge classique (Annexe A). Les cartes disponibles aux annexes A et B, sont les outils bruts utilisés sur le terrain. Elles contiennent quelques erreurs de mise en page et, il est possible de voir la superposition des différentes prises de vues est visible sur certaines.

La deuxième étape avait pour objectif la reconnaissance préalable des zones d'intérêt. Par une analyse visuelle des images satellites S2 et avec l'appui éventuel des images Pléiades, les zones qui apparaissent différentes visuellement ont été retenues pour être visitées en forêt. Une optimisation préalable des déplacements était nécessaire afin de limiter les kilomètres parcourus tant en voiture qu'à pied. Les visites de trois à quatre placettes ont été planifiées par jour de terrain pour un total de 64 placettes durant l'ensemble de la mission. Une identification, dans la littérature, des types de formations existantes dans les forêts du nord de la République du Congo a aussi permis de se référer à des connaissances existantes pour les premières observations de terrain.

L'étape suivante a été la récolte de données sur le terrain. Le déplacement en forêt s'est fait soit en accompagnant les ouvriers de la CIB, soit de façon indépendante. Pour la mission, un macheteur autochtone ouvrait la voie et un botaniste permettait l'identification des végétaux et des caractéristiques abiotiques de la zone. Les zones identifiées à la deuxième étape ont été retrouvées au moyen des cartes de l'annexe B emportées sur le terrain et à l'aide d'un GPS Garmin 62s. Dans chaque zone visitée, une placette de 50 x 50 m a été tracée par un relevé de points GPS. Les conditions de déplacement en forêt ont rendu la forme des placettes relativement aléatoire. Les données étaient ensuite récoltées, à l'intérieur des placettes, suivant un formulaire préétabli :

— La zone est-elle proche d'un marqueur spatial particulier ?
— Quelles sont les espèces indicatrices de la formation ?
— Quelles sont les espèces dominantes en hauteur, qui sont les plus visibles depuis

le ciel ?
- Quelles sont les strates végétales en présence ?
- Quel est l'ordre de grandeur de la hauteur dominante ?
- Y a-t-il une saisonnalité des conditions hydromorphiques du sol ?
- Y a-t-il une saisonnalité de la couverture foliaire ?
- Quel est le pourcentage d'ouverture de la canopée ?
- Quelle est l'origine des trouées dans la canopée ?
- De quelle année date la dernière exploitation forestière si la zone est exploitée ?
- A quel type de forêt appartient la zone échantillonnée ?

La première question du formulaire a permis de vérifier si le positionnement GPS relevé était relativement correct par rapport à la disposition des éléments sur la carte. Les réponses à la dernière question ont été premièrement basées sur les types de forêts identifiés dans la littérature et ont ensuite évolué sur base de l'expérience de terrain, par regroupement de certains types et par subdivision d'autres. En plus de ces questions, plusieurs photos de la végétation ont été prises sous la canopée. Les placettes échantillonnées durant la mission de terrain sont réparties dans l'ensemble des concessions de la CIB. La majorité d'entre elles se trouve sur les concessions de Pokola et de Kabo, mais deux missions à Loundgoungou et à Bérandjokou (un camp forestier au nord de la concession de Mimbeli) ont été l'occasion d'enrichir les données (Figure 10).

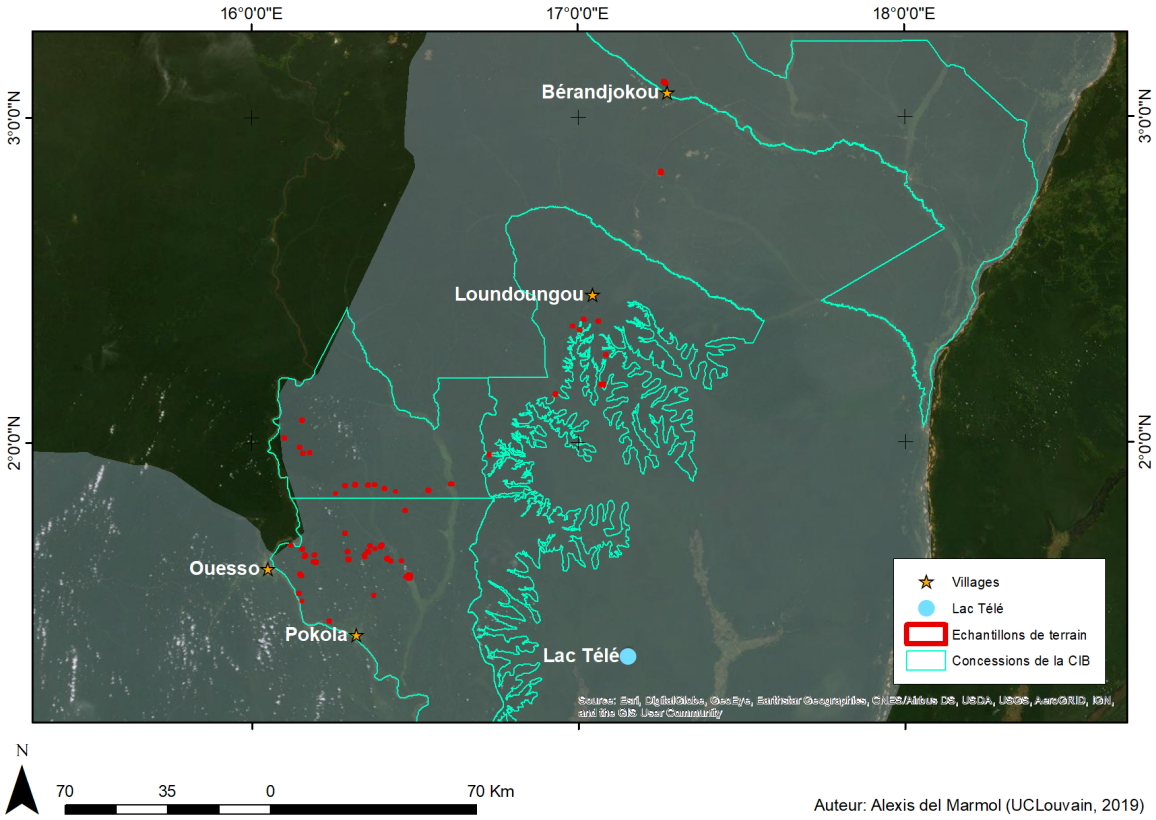


FIGURE 10 – *Répartition des placettes échantillonnées durant la mission de terrain*

Enfin, un classement des placettes échantillonnées dans les types de forêts finale-

ment retenus a été fait. Ce classement a permis de mettre en évidence les caractéristiques biotiques et abiotiques propres à chacun de ces types de forêts. Une mise en parallèle des observations de terrain avec les observations sur la composition colorée des images S2 a également permis de créer une clé d'identification visuelle des types de forêts et de répondre ainsi à la première question de recherche (section 5.1).

4.3. Traitement des images

L'obtention d'une classification à partir d'images satellites L1C nécessite une série d'opérations. Premièrement, des corrections atmosphériques doivent être appliquées afin d'obtenir un niveau de traitement 2A. Les nuages doivent ensuite être supprimés pour ne pas interférer dans les opérations suivantes. Différents indices de végétations sont calculés à partir des bandes spectrales. Une segmentation est réalisée à partir des *features pixel-based* et produit des objets utilisés dans le calcul des *features object-based* dont les textures. La combinaison de toutes ces *features* permet finalement la réalisation d'une classification au moyen d'un algorithme choisi. L'organigramme de la figure 11 résume ces différentes étapes.

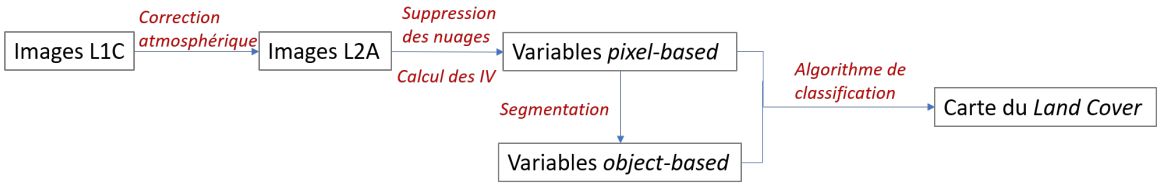
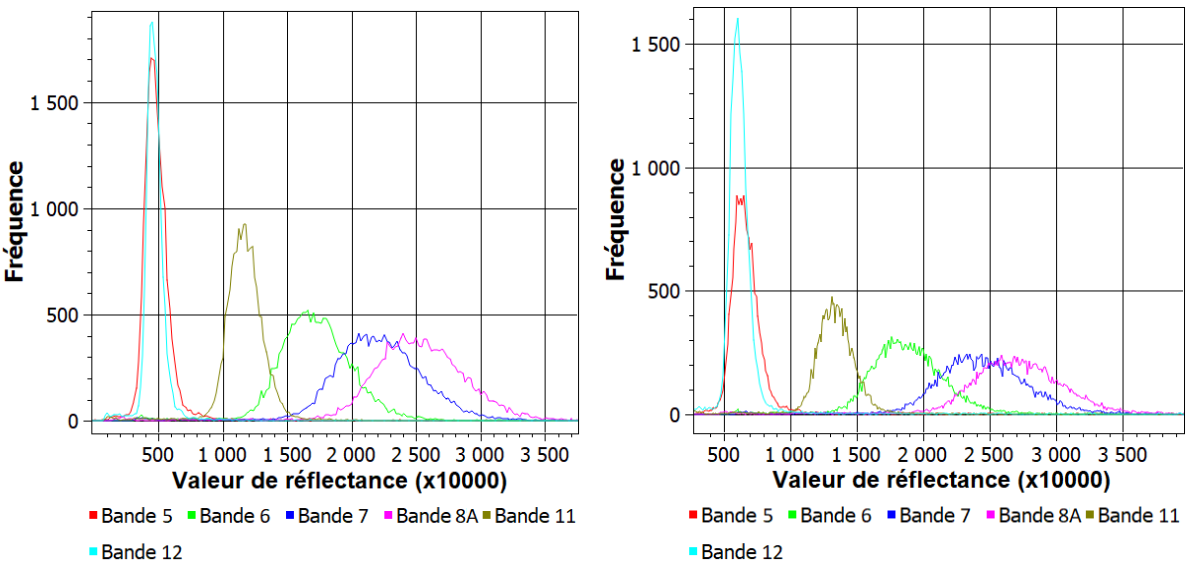


FIGURE 11 – *Résumé des étapes suivies dans ce mémoire pour le traitement d'images L1C vers l'obtention de classifications*

4.3.1. Correction en L2A et application du masque de nuage

Deux algorithmes de traitement d'images L1C ont pu être utilisés et comparés : Sen2Cor distribué par l'ESA et *Multi-sensor Atmospheric Correction and Cloud Screening* (MACCS) intégré au logiciel Sen2-Agri développé par un consortium scientifique dont l'UCLouvain fait partie et dont l'ESA est le bailleur de fonds (SEN2-AGRI_CONSORTIUM, 2018). La figure 12 compare les performances d'un traitement par MACCS et par Sen2Cor. Le bruit présent sur l'histogramme traité par Sen2Cor montre que cet algorithme est moins performant dans la correction des effets atmosphérique.

Malheureusement, à cause d'un problème de délai de mise à jour, le système Sen2-Agri n'a pas pu traiter les images de la tuile 33NXB où se trouvent la plupart des échantillons de terrain. Face à ce problème, la solution a été de travailler avec des images traitées par Sen2Cor.



(a) *(b)*

FIGURE 12 – *Histogramme des valeurs de réflectances pour les bandes à 20 m de résolution de S2 de la prise de vue du 7 mars 2018 sur l'orbite 50 de la tuile 33NXC après un traitement L2A par l'algorithme MACCS de Sen2-Agri (a) et par l'algorithme Sen2Cor (b). Cette comparaison permet de mettre en évidence la capacité de MACCS à réduire fortement les effets de bruit par rapport à Sen2Cor et donc son potentiel de correction supérieur.*

Le service de distribution des données S2 en ligne ne fournit des images L2A qu'à partir de janvier 2019. Or, à partir de cette date, aucune image suffisamment claire n'a été acquise. Les images L1C téléchargées aux dates décrites ci-dessus ont donc été traitées dans le cadre de ce mémoire par le logiciel Sen2cor disponible en téléchargement gratuit sur le site de l'ESA (http://step.esa.int/main/third-party-plugins-2/sen2cor/). Le processus de classification de ce mémoire se base sur ces images L2A.

Le traitement vers le niveau 2C réalisé par l'algorithme Sen2cor utilise les bandes à 60 m de résolution 1, 9 et 10, respectivement sensibles aux aérosols, à la vapeur d'eau et aux cirrus, pour corriger la luminance des autres bandes affectées par la présence de ces éléments dans l'atmosphère. Cette opération permet donc de réaliser la correction atmosphérique des bandes à 10 et 20 m de résolution utilisées pour la suite du travail.

La correction requise pour obtenir la réflectance de surface (L2A) par Sen2cor procède par une première classification, assez sommaire, des principaux éléments de la scène, dont les ombres des nuages ainsi que les pixels ayant une probabilité faible, moyenne et élevée d'être un nuage (LOUIS *et al.*, 2016). L'analyse visuelle de cette classification permet d'évaluer sa qualité et, sur cette base, de choisir les éléments qu'il est pertinent de masquer pour la classification à venir. Les pixels sont alors retirés sur l'ensemble des bandes spectrales.

4.3.2. Calcul des indices de végétation

Les différents indices présentés au début de ce mémoire ont été calculés sur la base de leur description dans la littérature : le NDVI, le NDII, le EVI et le EVI2. Le EVI n'a pas été gardé pour la suite du travail parce qu'il utilise la bande 2 du satellite dont la correction atmosphérique par l'algorithme Sen2cor est peu convaincante. En plus du NDVI, du NDII et du EVI2, de nouveaux indices normalisés ont été calculés sur la base de la formule du NDVI, mais en remplaçant les bandes utilisées par les bandes les moins corrélées entre elles. En effet, le niveau de réflectance dans certaines bandes est lié à des caractéristiques de l'objet observé. Par conséquent, moins les bandes utilisées dans cet indice sont corrélées entre elles, plus l'indice permettra de faire ressortir les caractéristiques propres à chacune des bandes. Ainsi l'infrarouge proche et le rouge, utilisés dans le NDVI, sont respectivement fortement réfléchis et fortement absorbés par les feuilles en bonne santé ce qui permet à l'indice de mettre en évidence le bon état sanitaire d'une couverture végétale. Plusieurs matrices de corrélations ont alors été calculées à partir des images L2A afin d'obtenir de nouveaux couples de bandes complémentaires et créer de nouveaux indices de végétation appelés *new vegetation index* (NVI) : le NVI1, le NVI2, le NVI3 et le NVI4

La finalisation de cette étape permet d'obtenir les variables *pixel-based* utilisées pour la classification : les réflectances à 20 m de résolution dans les bandes 5, 6, 7, 8, 8A, 11 et 12 traitées en L2A ainsi que les indices de végétation NDVI, NDII, EVI2, NVI1, NVI2, NVI3 et NVI4. Le choix de ces variables a fait l'objet d'un travail itératif qui a mené à tester le prétraitement d'images prises à différentes dates, la classification de plusieurs combinaisons de ces variables et différents indices de végétation construits au moyen de plusieurs matrices de corrélations différentes

4.3.3. Segmentation et calcul des features object-based

La segmentation d'une image permet le découpage de celle-ci en polygones utilisés comme objets vectoriels. Les différentes variables d'une image sont utilisées par l'algorithme qui y reconnait des ensemble de pixels cohérents. La reconnaissance nécessite de fixer trois paramètres qui définissent les polygones (la taille, la forme et la compacité). Cette opération est réalisée au moyen des variables *pixel-based* retenues pour les classifications et utilisées comme données d'entrée de l'algorithme [2]. La bande 5 est également utilisée car la rupture entre les tuiles est peu marquée et n'affecte pas la classification.

Pour chacun des objets issus de la segmentation, les textures sont calculées selon la méthode de HARALICK *et al.* (1973). Celui-ci propose de les calculer en suivant un des quatre angles du plus proche pixel voisin comme illustré à la figure 13 ou en considérant

2. La segmentation est réalisée par le logiciel propriétaire eCognition dans sa version 9.4.

les quatre directions à la fois. Le choix des textures et de l'angle utilisé s'est fait par analyse visuelle des résultats. Ainsi les textures GLCM de la moyenne, de l'écart-type et du second moment angulaire et la texture GLDV de l'entropie calculées suivant l'ensemble des directions ont été retenues.

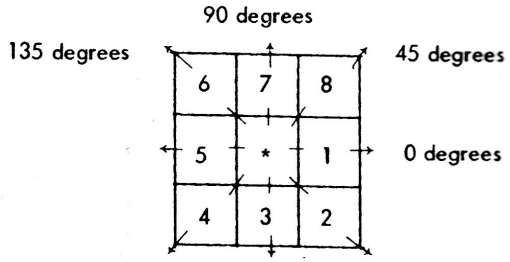


FIGURE 13 – *Angles de calcul des textures selon la méthode d'Haralick. Les pixels 1 et 5 sont les plus proches voisins du pixel central selon l'angle 0° (horizontal), les pixels 2 et 6 le sont à 135°, les pixels 3 et 7 le sont à 90° et les pixels 4 et 8 le sont à 45°. Cette représentation est uniquement spatiale et n'a pas de lien avec les valeurs de nuances de gris. (*HARALICK *et al., 1973).*

Les autres variables *object-based* sont obtenues à partir des polygones de la segmentation : chaque polygone se voit attribuer la valeur moyenne des réflectances et des indices de végétation présentes sur sa surface. En plus de cette moyenne, l'écart-type est aussi calculé afin d'obtenir une série de variables supplémentaires.

4.3.4. Construction du jeu de données d'entrainement

Le modèle créé en vue de la classification de l'image entière est construit à partir d'aires d'entrainement préablement définies et classifiées manuellement. Ces aires d'entrainement sont choisies et leur classe est attribuée par interprétation visuelle des images sur la base de la clé (Section 5.1). Ce travail a subi un processus itératif d'affinage des aires d'entrainement choisies pour être représentatives de la zone d'étude, pour leur ressemblance aux placettes échantillonnées et pour donner la meilleure classification possible. Cette itération a, entre autres, mené à retirer les placettes échantillonnées présentes dans le premier jeu de données d'entrainement afin d'être indépendant de la validation.

4.3.5. Classification et validation

Plusieurs combinaisons des *features* calculées sont testées comme données entrée [3]. La comparaison des résultats de ces différentes classifications permet de mettre en évi-

3. La classification est réalisée au moyen de la librairie *open source* Orfeo Toolbox (OTB) capable de classer et de valider les classifications.

dence les *features* qui ont un impact important sur la reconnaissance des types de forêts de façon générale ou pour certains types en particulier.

Concrètement, pour chaque combinaison de *features* trois approches sont utilisées :

— Une classification par pixel en entrant les *features pixel-based* dans l'algorithme de RF

— Un filtre par objet en appliquant un masque majoritaire sur la première classification par pixel. Ce masque est basé sur les polygones obtenus par segmentation et le filtre attribue à chaque polygone la classe majoritaire présente dans la surface qu'il couvre. Cela permet de lisser et d'enlever l'effet «poivre et sel» d'une classification par pixel.

— Une classification par objets en entrant les *features object-based* dans le l'algorithme de RF.

Le choix de ces approches se justifie par la volonté de mettre en évidence l'utilité de travailler par objets dans le cas d'un traitement d'images de grands massifs forestiers équatoriaux. En effet, les types forestiers sont souvent difficiles à délimiter spatialement et leurs signatures spectrales peuvent être relativement hétérogènes à l'intérieur d'une zone d'un même type. L'utilisation du filtre par classe majoritaire permet d'obtenir un produit qui combine les performances de classification par pixel avec l'aspect lissé d'une approche par objet qui rend la classification plus réaliste et plus pertinente pour l'utilisateur. Les combinaisons de *features* qui utilisent les textures ou les écarts-types ne peuvent être utilisées que dans la troisième approche. En effet, ces features n'existent que par objets.

Chaque classification a été validée en calculant les matrices de confusion sur la base de points générés dans les placettes échantillonnées sur le terrain. L'utilisation de ces zones apporte un biais en limitant la validation au centre de l'image, les extrémités n'ayant pas été visitées. Cependant, cela permet d'assurer l'exactitude par rapport au terrain en s'affranchissant de l'erreur d'interprétation visuelle dans le cas où la validation se ferait uniquement sur la base de cette interprétation. Seuls les polygones des classes de l'eau et des zones urbaines ont été tracés uniquement sur la base de leur interprétation visuelle qui est aisée. Concrètement, 100 points par classe sont générés aléatoirement à l'intérieur des polygones tracés sur la base des points GPS récoltés sur le terrain. Les polygones ont été tracés autour des relevés GPS mais sur des surfaces légèrement plus grandes que les placettes 50 x 50 m.En effet, la surface de celles-ci ne prends en compte pas plus d'une dizaine de pixels alors qu'il importe d'avoir une bonne représentativité du type de forêt. Pour réaliser la génération aléatoire des points, l'ensemble des polygones d'une même classe sont considérés comme un seul objet permettant ainsi une répartition des points indépendante du nombre de polygones par classe mais dépendante de la taille des polygones : une taille plus grande augmente la probabilité de se voir attribuer des points. L'utilisation d'un nombre de points identiques pour chaque classe permet de ne pas influencer le résultat de la validation par les classes dont les aires de validations

seraient plus grandes (polygones plus grands ou plus nombreux). L'eau, par exemple, qui est une classe qui est parfaitement classée, améliorerait considérablement le résultat de la validation si un nombre de points supérieur lui était attribué. Le nombre et la taille des polygones par classe sont liés d'une part à la souplesse du déplacement sur le terrain et d'autre part à la façon dont les polygones ont été tracés à partir des positionnements GPS relevés. En effet, durant la mission de terrain l'objectif était de relever au minimum 6 à 7 placettes par type de forêt et le fait que certaines types possèdent une ou deux placettes supplémentaires est lié à la facilité d'accès à ces forêts plutôt qu'à leur fréquence d'apparition.

Plusieurs métriques statistiques permettent d'évaluer la qualité d'une classification : l'exactitude de l'utilisateur (*user's accuracy* (UA)), l'exactitude du producteur (*producer's Aacuracy* (PA)), l'exactitude globale (*overall accuracy* (OA)), l'indice kappa (κ) et le F-score. Ces métriques sont comprises entre 0 et 1 et peuvent être interprétées de la façon suivante :

$X < 0,5$	Insatisfaisant
$0,5 < X < 0,7$	Satisfaisant
$0,7 < X < 0,8$	Bon
$0,8 < X$	Très bon

Où X est l'indicateur.

L'OA d'une matrice de confusion est la proportion totale des points correctement classés. Elle se calcule par la formule :

$$OA = \frac{\sum_{i=1}^{n} m_{i,i}}{M}$$

Où

- n est le nombre de classes
- i est le numéro de la classe
- M est le nombre de points total
- $m_{i,i}$ est le nombre de points de la classe i qui ont été correctement classifiés en i (les valeurs sur la diagonale de la matrice de confusion)
- C_i est le nombre total de points prédits dans la classe i
- G_i est le nombre total de points réels dans la classe i

L'indice (κ) est relativement similaire à l'OA et mesure l'accord entre les valeurs prédites et la réalité de terrain. Cependant, il tient également compte de l'accord lié à la chance qui est indépendant de la qualité de la classification et qui fausse donc la validation.

Il est calculé par la formule :

$$\kappa = \frac{OA - P(e)}{1 - P(e)}$$

P(e) est la probabilité d'accord lié à la chance :

$$P(e) = \sum_{i=1}^{n} \left(\frac{C_i \cdot G_i}{M} \right)$$

Où

— C_i est le nombre total de points prédits dans la classe i

— G_i est le nombre total de points réels dans la classe i

Le F-score est dérivé des UA et PA. La UA, aussi appelée «précision», est la proportion de points correctement classés parmi l'ensemble des points prédits dans cette classe. La PA, aussi appelée «sensibilité», est la proportion de points correctement classés parmi l'ensemble des points de référence de la classe. Le F-score est alors calculé pour chaque classe par :

$$F_{score} = \frac{2 \cdot précision \cdot sensibilité}{précision + sensibilité}$$

La comparaison des exactitudes globales et des indices κ des matrices de confusion résultantes permet de trouver les combinaisons qui génèrent majoritairement les meilleures classifications. L'utilisation du F-score permet de détailler en mettant en évidence l'influence de chaque classe sur la classification globale et donc de comprendre quelles *features* améliorent la classification d'une classe ou d'une autre en particulier. Une analyse visuelle des produits de classification obtenus permet aussi d'en relever les meilleurs.

Finalement, un résultat consolidé est calculé par vote majoritaire. Le vote se fait sur différentes classifications choisies parmi les meilleures, rendu visuel et matrice de corrélation confondus. Le choix de la combinaison est réalisé au moyen d'un processus itératif et le meilleur résultat est retenu. Concrètement, cette opération consiste à attribuer à chaque pixel la classe qui lui est attribuée le plus de fois dans les classifications choisies. Ce résultat est une confirmation des classifications produites.

CHAPITRE 5

RÉSULTATS ET DISCUSSION

Ce chapitre commence par la présentation de la typologie et de la clé d'interprétation visuelle issues du travail de terrain. Ensuite viennent l'analyse et l'explication des résultats du traitement des données satellitaires et des classifications, à savoir les produits du prétraitement des images, les *features* et les opérations intermédiaires pour les calculer et, pour finir, les rendus des classifications et leur validation.

5.1. Types forestiers et clé d'interprétation visuelle

Le premier résultat, issus de la mission de terrain, est une description des types de forêts avec leurs caractéristiques biotiques et abiotiques ainsi que la création d'une clé d'interprétation visuelle des images S2 affichées selon le code RGB = B06 B12 B08 [*RedEdge* SWIR2 NIR]. La légende choisie résulte d'un compromis entre les types de forêts réellement observés sur le terrain et les classes qu'il a été possible de reconnaître effectivement sur les images satellites en tenant compte des limites techniques et des délais de ce mémoire. En réalité, d'une zone à l'autre les nuances dans la végétation sont très nombreuses et il est presque possible d'identifier autant de types de forêts qu'il y a de placettes.

5.1.1. Les forêts adultes

Les forêts adultes (FA) peuvent être des forêts intactes ou des forêts secondarisées adultes. Sur le terrain il est certes possible de distinguer une forêt secondarisée adulte d'une forêt intacte, mais l'imagerie satellite, elle, ne permet pas encore, de les discriminer si ce n'est par la présence du réseau de pistes d'exploitation encore parfois visible dans les forêts secondarisées. VIVIEN *et al.* (1985) et WHITE (1983) classent ces forêts parmi les forêts tropicales semi-sempervirentes, une transition entre les forêts sempervirentes et les forêts semi-décidues. En effet, la saisonnalité de la couverture foliaire est présente, mais l'absence de réelle saison sèche induit une défoliation relativement réduite durant la période moins pluvieuse de l'année. Cependant, la plupart des autres références, bibliographiques et orales les positionnent en semi-décidues. Les FA possèdent une canopée fermée jusqu'à 70 % et dont la hauteur dominante atteint les 50 mètres. Elle est composée d'espèces très variables, dont les essences commerciales re-

cherchées dans la région (sapelli, sipo, ayous ...). La fermeture importante de la canopée induit un sous-étage relativement ouvert où il est facile de se déplacer. Les forêts secondarisées adultes sont des forêts ayant déjà été perturbées, d'où le terme «secondarisées» qui témoigne de leur évolution par rapport au stade de forêts secondarisées jeune. La végétation a eu le temps de se refermer pour atteindre une structure proche d'une forêt n'ayant jamais été exploitée. C'est ce dernier point qui définit les forêts intactes (Figures 14 et 15). Par conséquent, la composition en espèces commerciales est bien plus importante dans une forêt intacte et cette caractéristique permet de différencier les deux types. Il est important de noter, toutefois, que l'adjectif «intacte» utilisé ici n'est défini que par l'absence d'intervention industrielle passée.

Il est possible de distinguer les FA sur l'image satellite : il s'agit des larges étendues violettes ponctuées de vert selon une granulométrie fine (Figure 16).



FIGURE 14 – *Sous-bois relativement ouvert d'une FA*

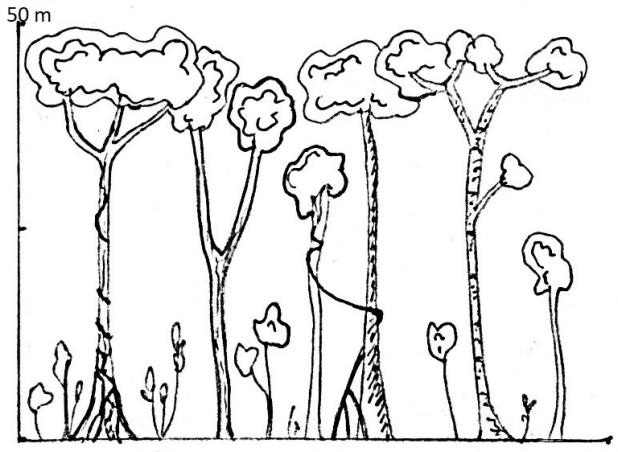


FIGURE 15 – *Schéma de profil type de FA*

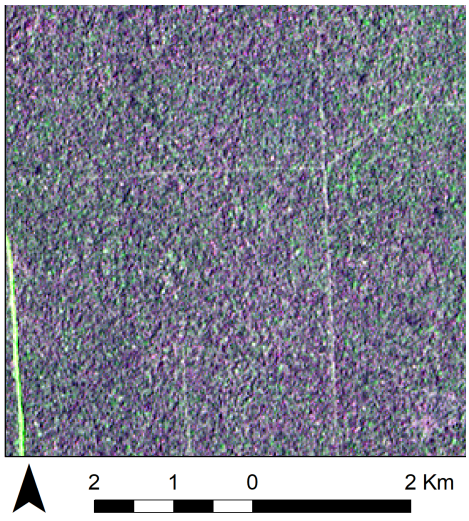


FIGURE 16 – *FA vue par S2 selon la composition colorée RGB = B06 B12 B08.*

5.1.2. Les forêts denses humides semi-décidues secondarisées jeunes

Les forêts secondarisées jeunes (FJ) sont des forêts mélangées de terre ferme de même diversité végétale que les forêts secondarisées adultes. Elles se classent également dans les forêts tropicales de terre ferme semi-décidues. La fermeture de la canopée varie entre 10 et 50 %, la rendant ainsi discontinue. En conséquence, les couronnes d'arbres sont séparées des trouées liées à l'exploitation récente. Cette dernière est l'origine des adjectifs «secondarisées jeunes» employés pour qualifier ces forêts. En effet, l'exploitation est jeune et a mené à une physionomie hétérogène pouvant faire penser à une forêt secondaire issue d'un recru forestier. Cependant, les plus grands arbres atteignant les 50 mètres de hauteur montrent que la forêt est ancienne. Les trouées permettent l'arrivée de lumière dans le sous-étage qui se densifie fortement avec des herbacées lianescentes (*Haumania...*) et de grandes herbes des familles des Marantacées (*Megaphrynium...*), Zingiberacées (*Aframomum, Costus,...*) et Commelinacées (*Palisota...*) comme on peut le voir aux figures 17 et 18. Il est donc assez difficile d'y circuler. L'exploitation provoque, dans ces forêts, une diminution de la proportion d'essences commerciales.

Ce sont principalement les trouées qui permettent l'identification de ce type forestier sur l'image satellite en apparaissant d'un vert plus clair. Les larges étendues vertes ponctuées de violet selon une granulométrie fine sont des FJ (Figure 19). La transition entre les FA et les FJ se fait par dégradé qui donne la mesure de la continuité du niveau d'ouverture de la canopée : plus il y a de vert, plus la canopée est ouverte.



FIGURE 17 – *Sous-bois dense d'une FJ*

5.1.3. Les forêts à marantacées

Les forêts à marantacées (FM) sont des formations végétales de terre ferme dont la canopée ne dépasse pas les 10 %. Ces types de forêts sont composés en grande majorité de plantes de la famille des marantacées (*Megaphrynium* étant un des représentants

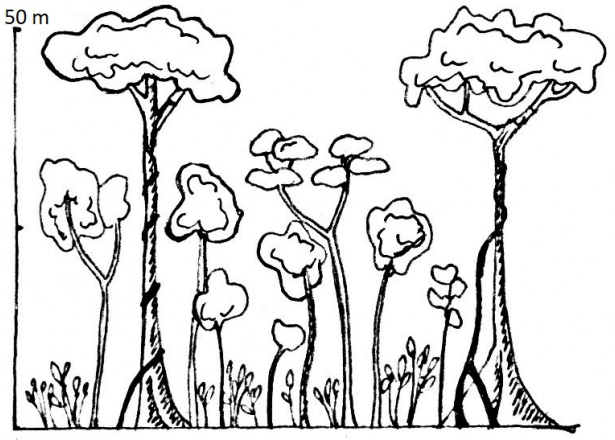


FIGURE 18 – *Schéma de profil type de FJ*

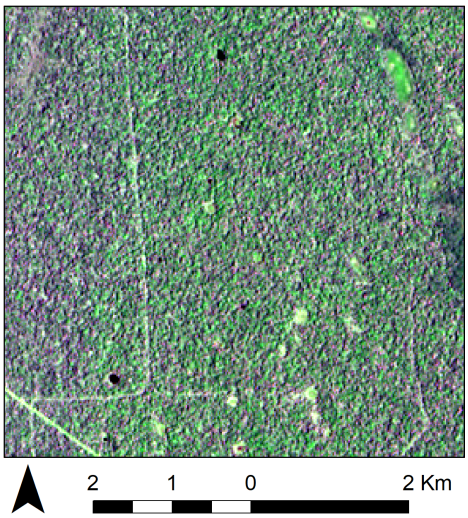


FIGURE 19 – *FJ vue par S2 selon la composition colorée RGB = B06 B12 B08.*

principaux) pouvant atteindre près de trois mètres de haut et quelques arbres émergents ou avec un couvert de parasoliers (*Musanga cecropioides*), arbres héliophiles colonisateurs à feuillage décidu. L'origine de ces ouvertures est variable : l'action d'une violente tempête, la déstabilisation du couvert à la lisière des eyangas (clairières inondées ou inondables) ou une exploitation trop intense.

Leur composition floristique confère à ces formations végétales une signature spectrale particulière qui les rend visibles sur les images satellites. On peut les observer sous la forme de taches blanches nettes autour de eyangas, ou encore de ponctuations blanches plus ou moins denses dans une matrice de forêt jeune ou adulte. Les FJ, les FA et les FM forment un continuum relativement difficile à délimiter tant sur le terrain que par télédétection (Figure 22).



FIGURE 20 – *Enchevêtrement extrêmement dense de plantes marantacées*

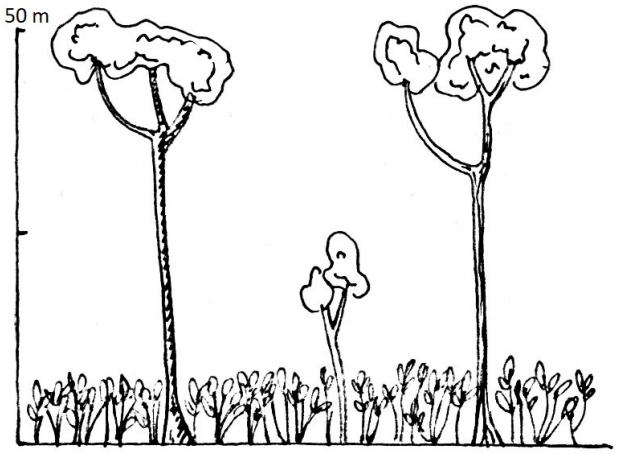


FIGURE 21 – *Schéma de profil type de FM*

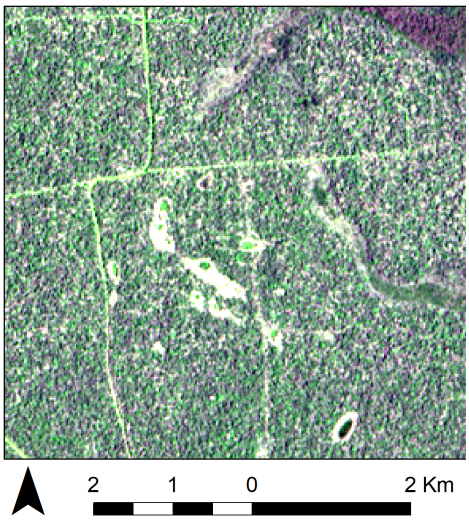


FIGURE 22 – *FM vue par S2 selon la composition colorée RGB = B06 B12 B08.*

5.1.4. Les forêts denses humides sempervirentes à limbalis

Les forêts à limbalis (FL) sont des formations végétales sempervirentes monodominantes composées de limbalis (*Gilbertiodendron dewevrei*) en peuplements presque purs. Ces formations de limbalis sont typiques de cette région de la grande forêt équatoriale africaine et peuvent occuper des surfaces relativement grandes, jusqu'à plus de 2500 hectares pour certaines. Souvent, les peuplements de limbalis se retrouvent en bordure des forêts ripicoles ou des zones marécageuses. Ces arbres sont tolérants à une submersion temporaire du sol en période plus humide, ce qui permet de les retrouver en zones inondables. La canopée peut monter jusqu'à 50 mètres de haut et est très homogène (Figure 24). Une autre caractéristique de cette canopée est sa fermeture qui va de pair avec un sous-bois relativement ouvert, à l'instar d'une FA. Le renouvellement très rapide des feuilles rend cette forêt sempervirente et empêche la variation de la couverture foliaire au cours de l'année (Figure 23).

Enfin, ce type de forêts est facilement identifiable sur les images satellites par les tâches vert foncé caractéristiques et homogènes qu'il forme soit en bordure du réseau hydrographique, soit au milieu de la terre ferme (Figure 25).



FIGURE 23 – *Vue d'un sous-bois en peuplement monospécifique de limbalis*

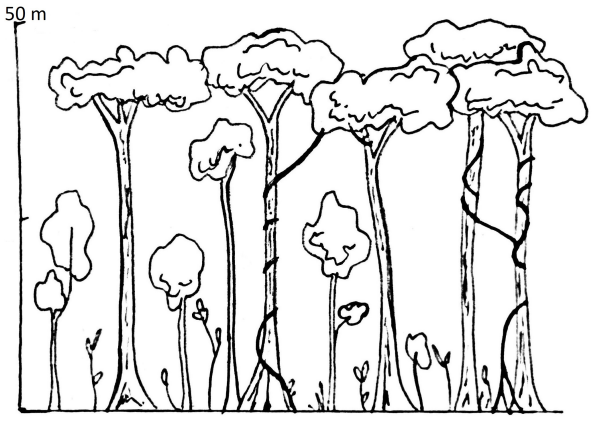


FIGURE 24 – *Schéma de profil type de FL*

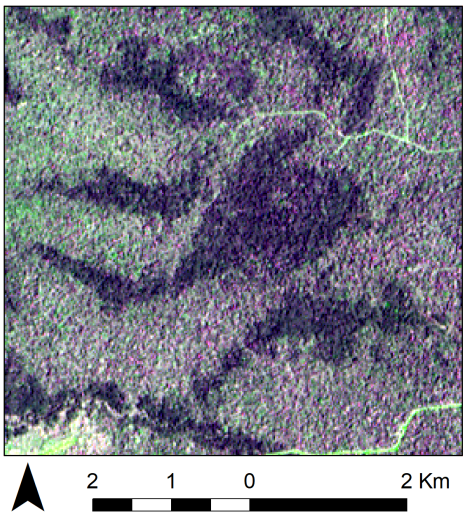


FIGURE 25 – *FL vue par S2 selon la composition colorée RGB = B06 B12 B08.*

5.1.5. Les forêts ripicoles

Les forêts ripicoles (FR) sont des formations végétales qui se trouvent en bordure de petits cours d'eau permanents ou non, aux embranchements fins du réseau hydrographique jusqu'aux extrémités de celui-ci. On les retrouve soit directement dans le lit de la rivière soit sur ses flancs périodiquement inondables ou inondés en permanence suivant le débordement du cours d'eau (Figure 27). La végétation qui s'y trouve est assez basse : la hauteur maximale se situe généralement entre 10 et 20 mètres. Elle est un mélange d'espèces ligneuses basses, de lianes denses et souvent couvertes par un enchevêtrement de plantes grimpantes de la famille des euphorbiacées qui déploient leurs feuilles au-dessus de la canopée (Figure 26). La saisonnalité de la couverture foliaire est impactée par l'alternance entre inondations et périodes sèches. Cette spécificité permet

de reconnaitre ces formations qui apparaissent blanc-vert à rosé sur la composition colorée (Figure 28).



FIGURE 26 – *Couverture de lianes euphorbiacées en FR*

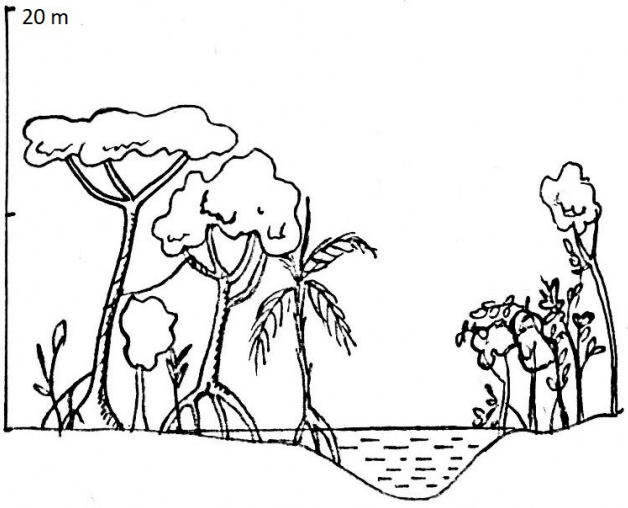


FIGURE 27 – *Schéma de profil type de FR*

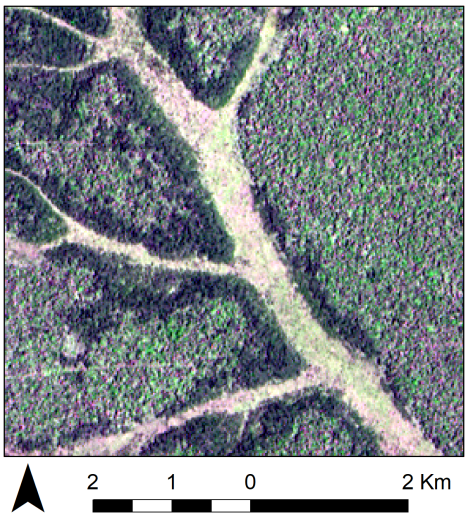


FIGURE 28 – *FR vue par S2 selon la composition colorée RGB = B06 B12 B08.*

5.1.6. Les forêts marécageuses

Les forêts marécageuses (M) sont des forêts inondées en permanence dans les plaines alluviales des artères principales du réseau hydrographique. Ces forêts sont dominées par diverses essences mélangées telles que le bahia (*Hallea ciliata*), le padouk d'eau (*Pterocarpus osun*), le bubinga d'eau (*Guibourtia demeusei*), le limbali (*Gilbertiodendron dewevrei*) ou encore le rikio (Uapaca spp) reconnaissable par ses racines échasses pouvant mesurer plusieurs mètres. Dans ces écosystèmes, le niveau de l'eau est toujours au-dessus de la surface du sol et varie peu au cours de l'année (Figure 29). La couverture

foliaire n'est donc pas affectée par les changements de pluviométrie intra-annuels. La hauteur dominante de ce type forestier ne dépasse pas 30 mètres (Figure 30).

Les M sont visibles sur l'image satellite sous la forme de larges bandes roses sur les cours d'eau les plus importants du réseau hydrographique (Figure 31).



FIGURE 29 – *Sous-bois relativement dense de M*



FIGURE 30 – *Schéma de profil type de M*

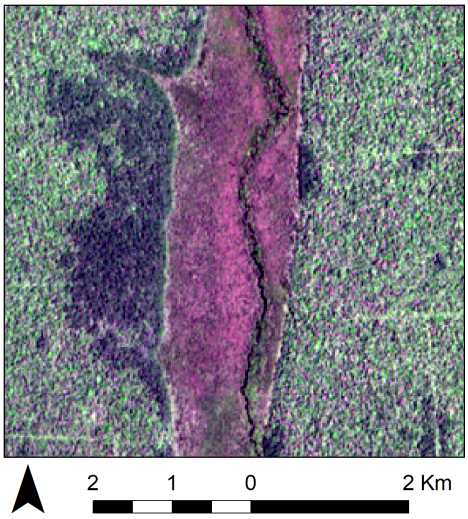


FIGURE 31 – *M vue par S2 selon la composition colorée RGB = B06 B12 B08.*

5.1.7. Les forêts à raphia

Les forêts à raphia ou raphiales (R) sont des forêts marécageuses de formations monospécifiques de raphias (*Raphia spp.*). Ces palmiers d'une hauteur maximale de 20 mètres se développent dans des milieux inondés en permanence et leurs longues feuilles forment une canopée dense qui laisse peu de place à l'apparition d'une végétation en sous-étage. Comme les M d'essences mélangées, les R se retrouvent dans les plaines

alluviales inondées par des cours d'eau principaux ou regroupées en peuplements de quelques hectares dans des dépressions inondées des forêts de terre ferme (Figure 32 et 33).

Les raphias ont une réflectance particulière qui rend leurs peuplements visibles sur les images satellites sous forme de taches vert foncé homogènes (Figure 34).



FIGURE 32 – *Sol inondé d'une R*

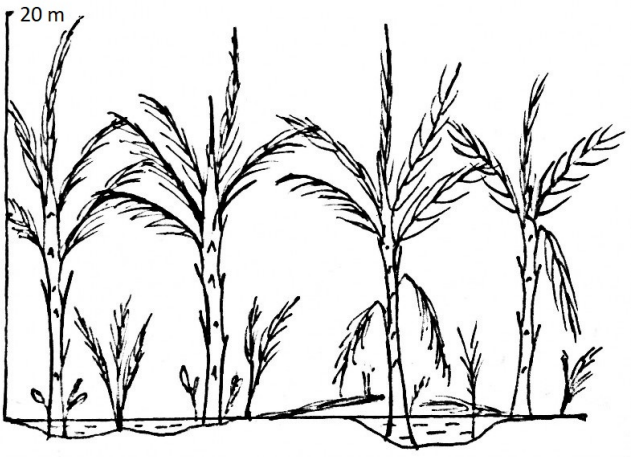


FIGURE 33 – *Schéma de profil type de R*

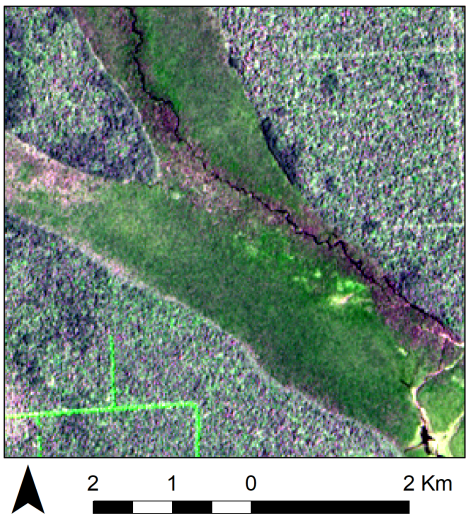


FIGURE 34 – *R vue par S2 selon la composition colorée RGB = B06 B12 B08.*

5.1.8. Les zones herbacées

Les zones herbeuses (ZH) regroupent toute une série d'écosystèmes ouverts occupés par une végétation herbacée. Parmi ces écosystèmes, on retrouve les clairières inondées ou inondables (baï et eyanga) souvent couvertes des plantes de la famille des Cypéracées, des Poacées ou des Mélastomacées, les regroupements de plantes aquatiques dans

les méandres les plus lents des rivières, et les plaines ouvertes à l'approche des zones habitées (Figure 35). Cette végétation est souvent d'une hauteur inférieure au mètre (Figure 36).

La signature spectrale de ces plantes herbacées est caractérisée par une réponse forte dans le moyen infrarouge qui les fait apparaitre vert intense sur les images satellites (Figure 37).



FIGURE 35 – *Clairière inondée couverte de plantes herbacées au cœur de la forêt*

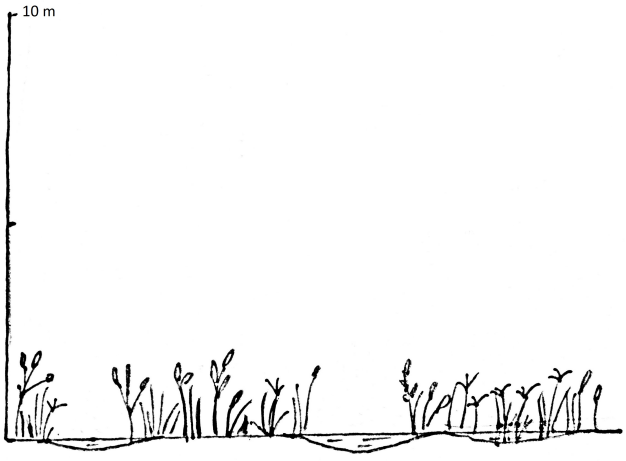


FIGURE 36 – *Schéma de profil type de ZH*

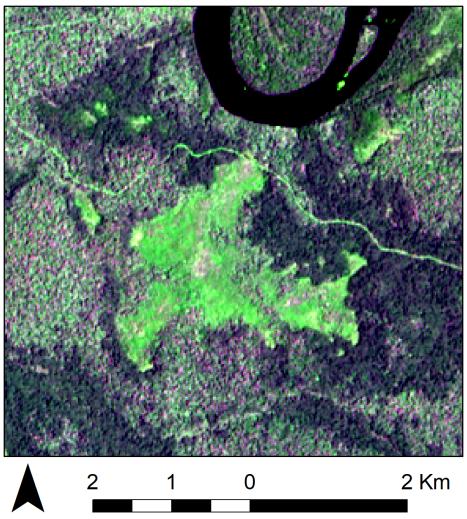


FIGURE 37 – *ZH vue par S2 selon la composition colorée RGB = B06 B12 B08.*

5.2. Analyse de la qualité du prétraitement

Les quatre tuiles 33NXB, 33NXC, 33NYC et 33NYD ont été traitées pour les dates du 22 décembre 2018 sur l'orbite relative 50, et du 24 décembre 2018 sur l'orbite relative 7 qui offraient la combinaison du ciel le plus clair sur l'ensemble de la zone et de la plus petite différence temporelle entre les prises de vue. Malgré cet écart réduit de deux jours, les mosaïques résultantes montrent des artéfacts à la jointure des deux orbites. La figure 38 utilise la mosaïque de la composition colorée RGB = B06 B12 B08 comme fond de carte et la différence entre les deux orbites est bien visible. La figure 39 montre que, pour les premières bandes, le problème se pose aussi entre les images d'une même orbite relative. Ces phénomènes peuvent s'expliquer de plusieurs manières.

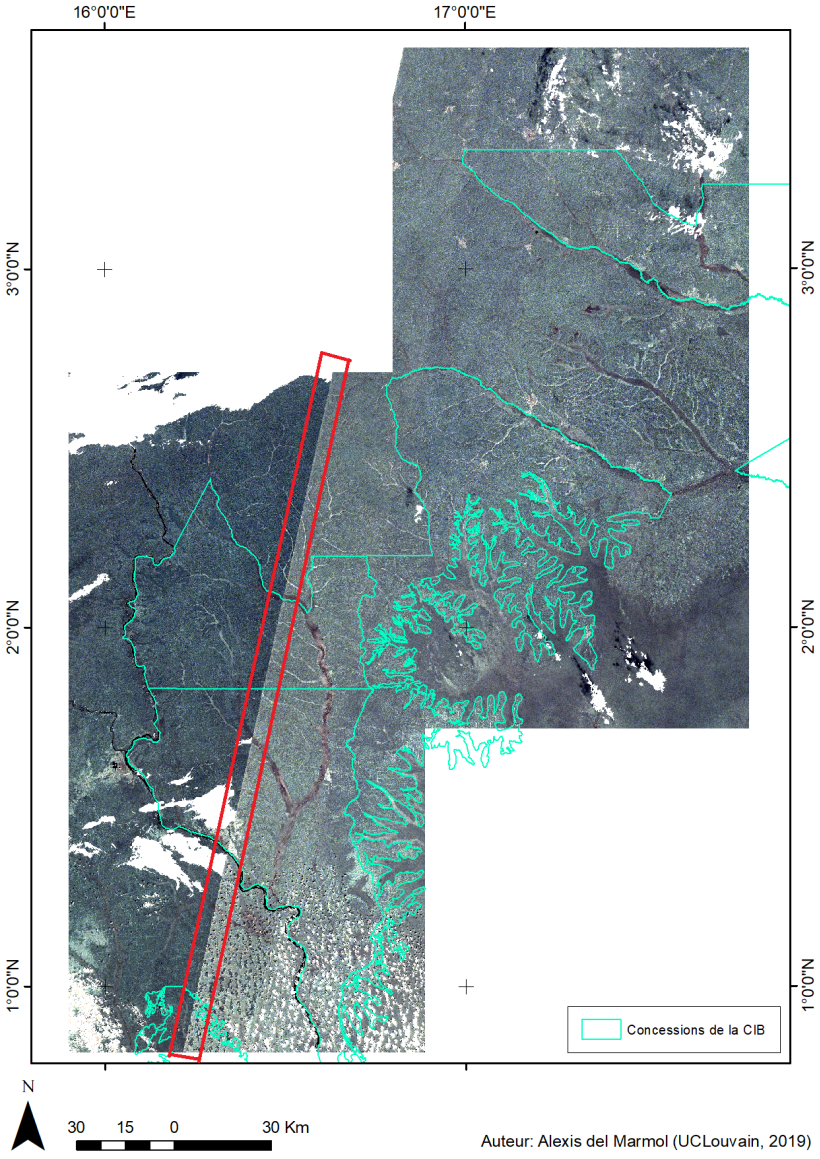
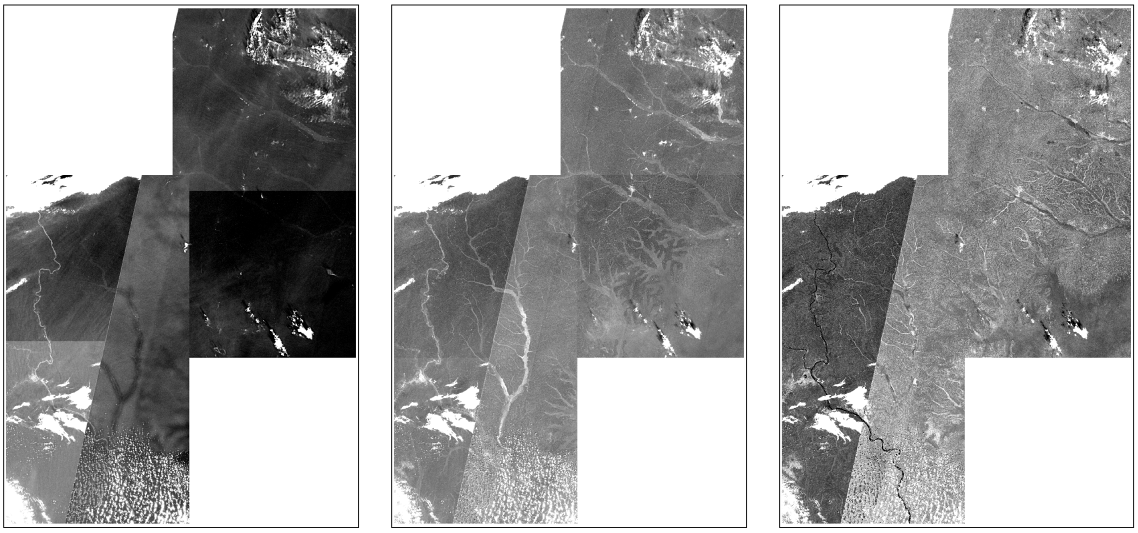


FIGURE 38 – *Mosaïque des images S2 prises le 22 décembre 2018 sur l'orbite relative 50 et le 24 décembre 2018 sur l'orbite relative 7 suivant la composition colorée RGB = B06 B12 B08. Le cadre rouge met en évidence les artéfacts à la limite des deux orbites relatives.*

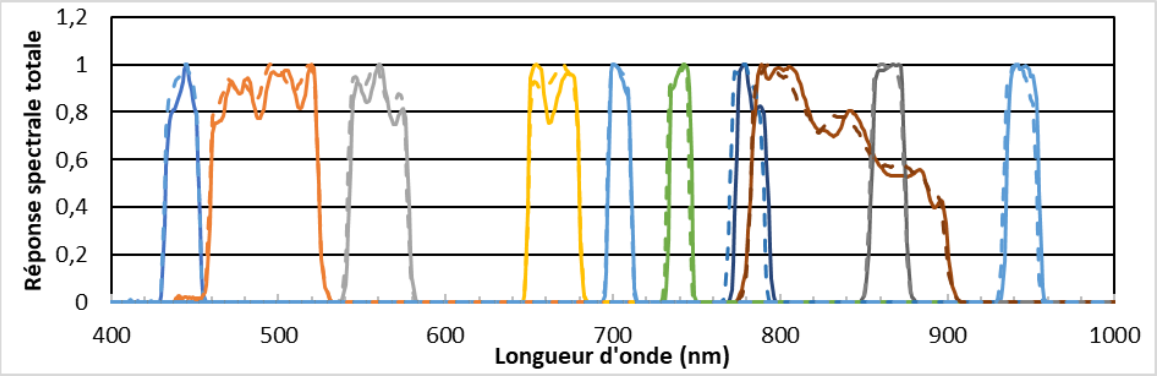


Bande 2 *Bande 5* *Bande 6*

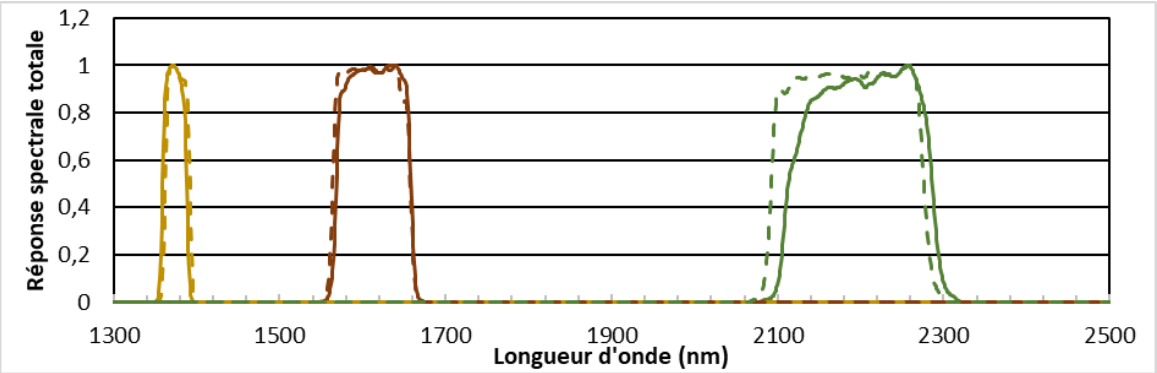
FIGURE 39 – *Mosaïques des images S2 prises le 22 décembre 2018 sur l'orbite relative 50 et le 24 décembre 2018 sur l'orbite relative 7 pour différentes bandes S2. La mauvaise uniformisation du signal lors de la création de la mosaïque est clairement visible entre chaque tuile pour les bandes 2 à 5 et n'est plus que visible entre les deux orbites relatives à partir de la bande 6.*

Premièrement, les images du 22 décembre sur l'orbite relative 50 ont été prises par la plateforme S2B et celles du 24 décembre, sur l'orbite relative 7, l'ont été par la plateforme S2A. Les deux satellites ont des spécifications techniques identiques et sont sensés être parfaitement calibrés l'un à l'autre afin que leurs produits soient tout à fait comparables et utilisables. L'algorithme de correction atmosphérique de Sen2cor de l'ESA utilise des coefficients de calibrations différents pour S2A et S2B afin de compenser les différences qu'il existe entre les fonctions de réponse spectrale des deux capteurs MSI dans les bandes 1, 2 et 12. CHEN *et al.* (2018) ont mesuré ces différences qui s'avèrent être minimes. La figure 40 compare les signaux obtenus par les plateformes et ces très faibles différences. Cette hypothèse n'est donc pas réellement convaincante.

Ensuite, les forêts tropicales sont des écosystèmes qui influencent fortement le microclimat local. Les conditions atmosphériques proches de la surface peuvent fortement varier d'un jour à l'autre et localement d'une tuile à l'autre selon l'ensoleillement et la pluviométrie qui influencent la dynamique des courants aériens et donc des aérosols et de l'eau émise par les arbres, selon l'hydrographie ou encore selon la végétation. De manière générale, la quantité d'aérosols dans l'atmosphère au-dessus de ces forêts est particulièrement élevée. L'ensemble de ces facteurs influence les signaux obtenus dans les bandes de longueurs d'onde courtes (bandes 2, 3, 4 et 5). Afin d'estimer l'effet des aérosols et de procéder aux corrections, Sen2cor ne se base pas sur les tuiles, mais sur des tables (lookup table - LUT) prédéfinies qui reprennent des valeurs pour le profil atmosphérique (hiver ou été), le type d'aérosol (rural ou maritime) et le niveau d'ozone. Ces tables sont construites de façon à couvrir la plupart des conditions atmosphériques pré-



Bandes du très proche infrarouge. De gauche à droite : les bandes 1 à 9. La bande est en marron.



Bandes du moyen infrarouge. De gauche à droite : les bandes 10 à 12.

FIGURE 40 – *Comparaison des fonctions de réponse spectrale normalisées à un des satellites S2A (traits continus) et S2B (traits pointillés) (*ESA*, 2017)*

sentes sur Terre pour la mission S2 (MUELER-WILM *et al.*, 2019). Les LUT pourraient ne pas être suffisantes pour les conditions atmosphériques extrêmes en termes d'aérosols, et empêcheraient ainsi Sen2cor de réaliser une correction suffisante. Les différences de conditions atmosphériques d'un jour à l'autre étant plus élevées que localement au sein d'une même journée, la difficulté à homogénéiser les corrections entre deux dates explique donc l'apparition d'artéfacts sur la limite des deux orbites sur toutes les bandes tandis que les artéfacts entre tuiles d'une même date disparaissent au-delà de la bande 5.

Afin de s'affranchir des effets de mosaïque entre deux dates, la zone d'étude a été réduite aux tuiles de l'orbite relative 50, où la plupart des placettes ont été échantillonnées. Cette réduction de la zone de travail permet également la réduction du temps de calcul pour les différentes opérations de traitement d'image. La zone échantillonnée totale est relativement étendue, l'ensemble des UFA couvrant une surface de près de 2 millions de kilomètres carrés, et possède donc une variabilité naturelle de la flore. En effet, celle-ci augmente avec la distance : des espèces présentes dans l'UFA de Pokola peuvent ne plus se retrouver du côté de Bérandjokou et inversement. La perte d'une partie des données de terrain résultant du choix de réduire la zone est donc compensée par la diminution

de la variabilité floristique qui augmente la représentativité des classes choisies et présentées à la section 5.1.

À partir des deux tuiles retenues, plusieurs classifications ont été réalisées. Les résultats peu convaincants des premières classifications ont été améliorés en modifiant les aires utilisées pour l'entrainement (voir section 4.3) de l'algorithme pour finalement aboutir au résultat présenté à la figure 41. Malgré ces modifications, un autre phénomène contraignant dans la classification de l'image n'a pas pu être compensé : la présence d'un léger gradient de réflectances diminuant dans la direction du nord-est. Difficilement perceptible visuellement sur les bandes au delà de la bande 6, l'effet devient très clair dans les longueurs d'onde inférieures et particulièrement dans la bande 4 (Figure 43). Malgré le fait que ces dernières ne soient pas utilisées pour la classification, le phénomène affecte le résultat (Figure 41) : les forêts à limbalis ont une réponse spectrale généralement inférieure aux autres types forestiers et en plaçant quelques aires d'entrainement de ces forêts autour de Pokola, toute la zone nord se retrouve polluée par des pixels mal positionnés dans cette classe.

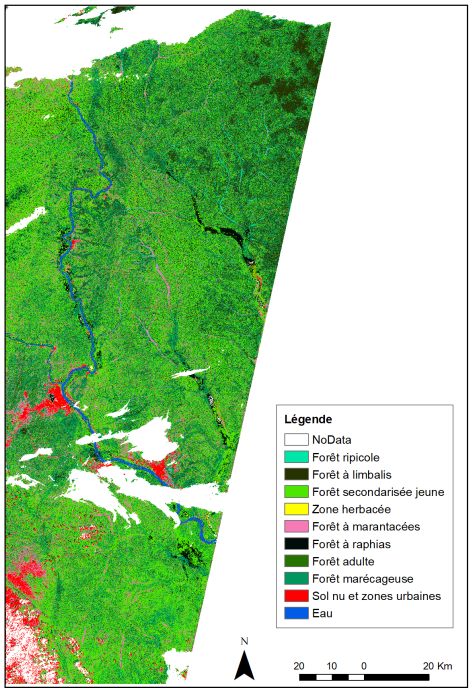
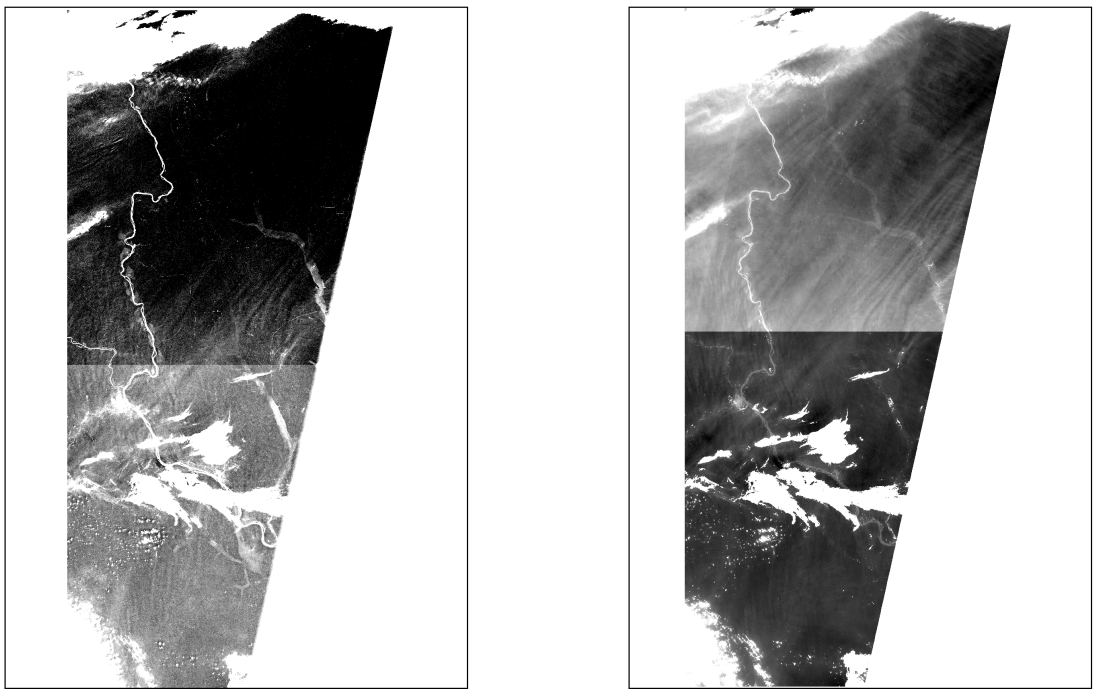


FIGURE 41 – *Classification sur la base des bandes 6 à 12, du NDII et du NVI1. On observe une présence élevée de forêts à limbali dans le Nord et une présence très élevée de forêt marécageuse autour de Pokola et d'Ouesso.*

L'explication de ce phénomène est la même que celle qui permet d'expliquer les artéfacts de jointure des mosaïques, à savoir la présence d'effets atmosphériques résiduels non corrigés par Sen2cor. La figure 42 montre que les aérosols en sont la cause : les effets de gradient nord-est visibles sur les bandes L2A sont les mêmes que ceux visibles sur la bande 1 qui comprend les longueurs d'onde les plus sensibles aux aérosols.

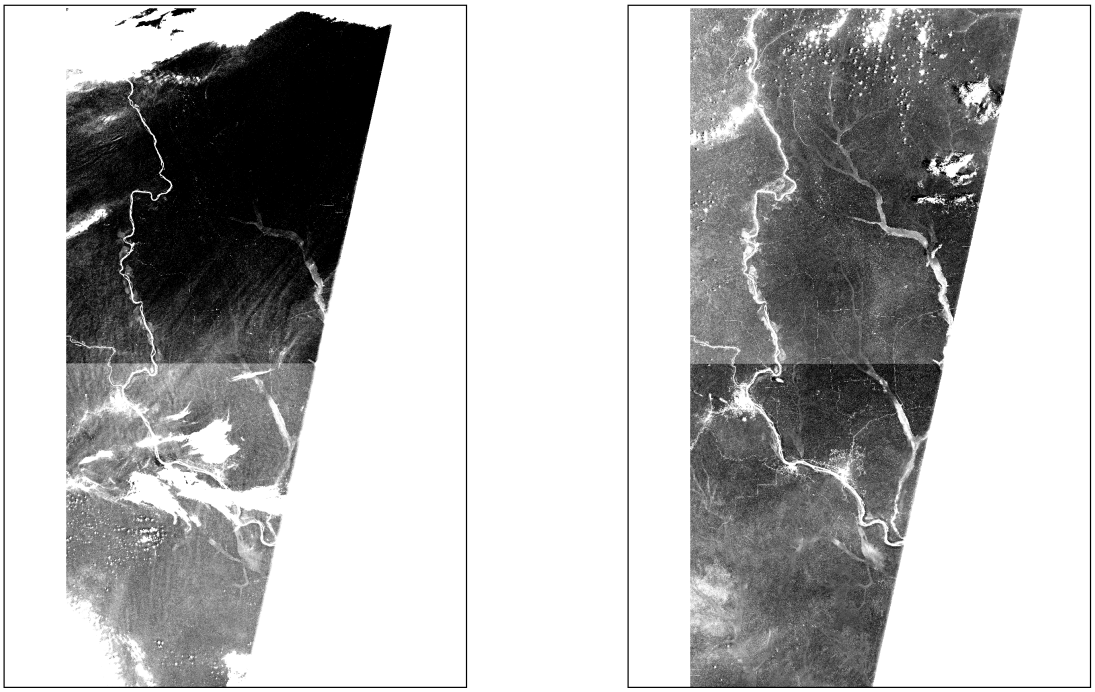


Bande 4 *Bande 1*

FIGURE 42 – *Comparaison des bandes 1 (L1C) et 4 (L2A par Sen2cor) prises le 22 décembre 2018 sur l'orbite relative 50. Les effets atmosphériques visibles sur la bande 4 sont les mêmes que ceux formés par les aérosols sur la bande 1.*

L'utilisation d'autres images avec une atmosphère plus claire est la seule solution pour répondre à ce problème d'aérosols. Des images du 7 mars 2018 ont été choisies, car la réduction de la zone de travail à l'orbite relative 50 a levé la contrainte liée au rapprochement des dates de prise de vue, avec une atmosphère claire, sur deux orbites différentes. Sur ces image son observe l'absence presque totale de nuage dans le sud de la zone ainsi que l'absence du dégradé de réflectance vers le nord-est. La différence est visible sur la comparaison des bandes 4 reprise à la figure 43.

Un dernier effet remarquable lié à la mosaïque est la présence de lignes parallèles à la trajectoire du satellite et formant des bandes de près de 30 km de large visibles dans le moyen infrarouge de la bande 11 (Figure 44) et dans les indices qui utilisent cette bande spectrale. La direction de ces lignes, la largeur des bandes formées et leur présence au sein d'une image prise indépendamment des autres rejette l'hypothèse d'effets atmosphériques ou d'effets propres à la scène observée. Il s'agirait donc d'un phénomène lié au MSI et plus particulièrement à l'intercalibration des 12 détecteurs différentes du MSI. Ce phénomène a une forte influence sur les classifications réalisées à la date du 22 décembre 2018 (Figure 41) et une légère influence sur certaines classifications. L'utilisation d'autres *features* permet de la faire disparaitre (Section 5.4).



22 décembre *7 mars*

FIGURE 43 – *Comparaison des mosaïques des bandes 4 S2 prises le 22 décembre 2018 et le 7 mars 2018 sur l'orbite relative 50 traitées en L2A par Sen2cor. On observe d'une part le gradient de réflectance très présent dans l'infrarouge proche à la date du 22 décembre et d'autre part la différence de propreté du signal et couverture nuageuse entre les deux prises de vues.*

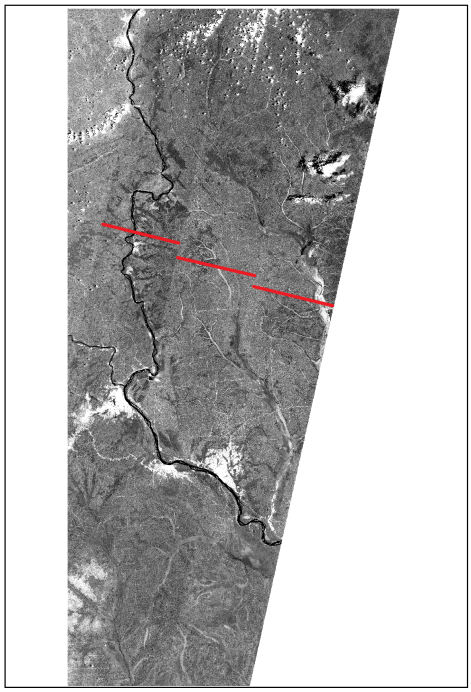


FIGURE 44 – *Bande 11 de la prise de vue du 7 mars 2018. On distingue des bandes plus foncées et plus claires de 30 km de large mises en évidence par les traits rouges qui marquent leur largeur.*

5.3. Variables utilisées pour la classification

Les classifications ont été réalisées sur la base de trois types de variables : les réflectances traitées en L2A, les indices de végétation et les textures. Les deux premiers types ont été calculés par pixels et par objets tandis que les textures n'ont été traitées que par objets.

5.3.1. Mise en évidence des bandes les plus discriminantes

Les bandes spectrales L2A n'ont pas toutes le même pouvoir discriminant pour les différents types forestiers. En comparant les moyennes de réflectance des aires d'entrainement par bandes et par types, on fait ressortir les bandes les plus utiles. La figure 45(a) est une graphique des signatures spectrales de chaque type et l'analyse visuelle de ces signatures montre que les bandes 6 à 12 sont les plus discriminantes. Cette analyse est un argument supplémentaire pour ne pas utiliser les longueurs d'onde les plus courtes. Afin de mieux visualiser les écarts-types, la figure 45(b) présente les mêmes valeurs sous forme d'un graphique en bâtonnets.

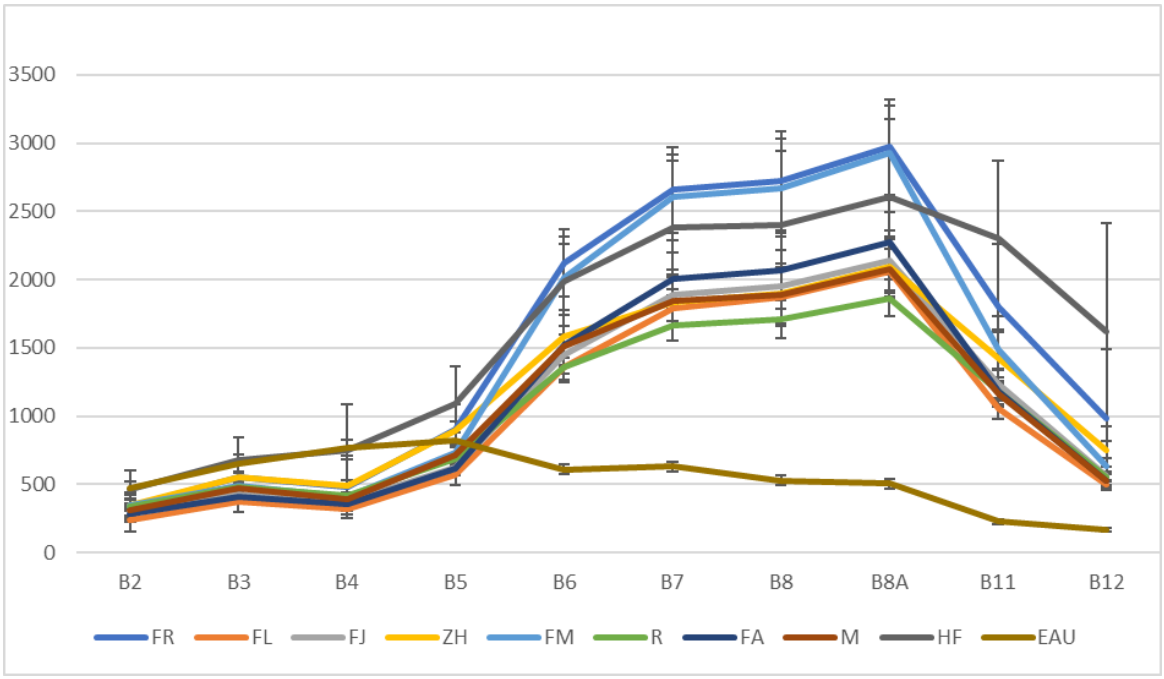
5.3.2. Calcul des indices de végétation

Les formules pour calculer les NVI ont été déduites des valeurs de différentes matrices de corrélations (voir section 4.3). Les premières matrices ont été calculées sur les mosaïques d'images prises à différentes dates sans tenir compte des effets de jointure. Les effets atmosphériques encore visibles dans les plus courtes longueurs d'onde (bandes 2, 3, 4 et 5) et empêchant les jointures de se faire correctement entre les tuiles d'une même prise de vue n'ont pas non plus été pris en compte. Les classifications réalisées avec les indices construits sur la base des plus faibles corrélations de ces matrices ont donc donné de mauvais résultats. Malgré cela, quatre indices sont retenus en raison des résultats visuellement convaincants qu'ils ont permis d'obtenir sur certaines classifications. Il s'agit des indices NVI basés sur les couples de bandes 11 et 4 (NVI1), 12 et 8A (NVI2), 11 et 6 (NVI3) et 12 et 8 (NVI4) obtenus par les formules suivantes :

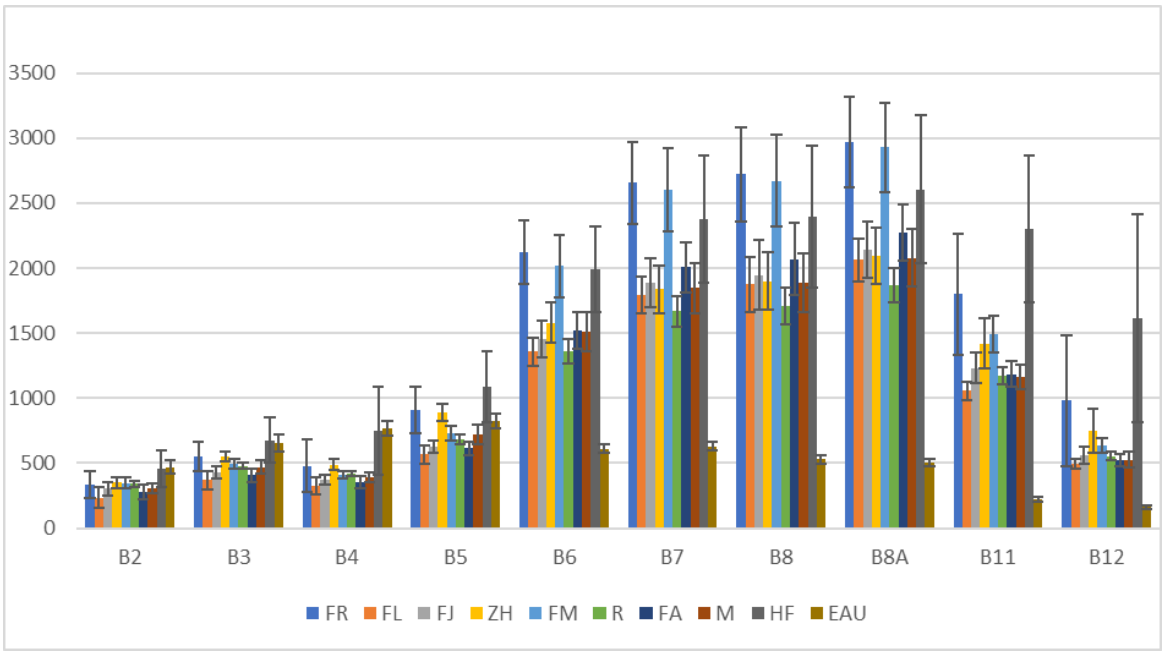
$$\text{NVI1} = \frac{\text{B11} - \text{B04}}{\text{B11} + \text{B04}} \qquad \text{NVI2} = \frac{\text{B8A} - \text{B12}}{\text{B8A} + \text{B12}}$$

$$\text{NVI3} = \frac{\text{B06} - \text{B11}}{\text{B06} + \text{B11}} \qquad \text{NVI4} = \frac{\text{B08} - \text{B12}}{\text{B08} + \text{B12}}$$

La matrice de corrélation retenue est celle calculée à partir de la prise de vue du 7 mars 2018 de l'orbite relative 50 sur la tuile 33NXB où l'absence de nuages est presque complète et l'atmosphère particulièrement claire (Tableau 1). Les bandes 2 à 5 n'étant



(a)



(b)

FIGURE 45 – *Signatures spectrales des différents types de forêts selon (a) la représentation classique et (b) une représentation en bâtonnets*

pas considérées, les deux valeurs les plus intéressantes pour un calcul d'indice de végétation sont celles des deux couples de bandes 12 (SWIR2) et 8 (NIR) ou 8A (veryNIR), affichées en bleu dans le tableau 1. Ces couples sont à la base des indices NVI2 et NVI4 et confirment la pertinence de ceux-ci. Dans cette même matrice, le couple à la base du NVI1 a aussi une corrélation relativement bonne (parmi les plus faibles) tandis que celui du NVI3 possède une valeur peu intéressante avec une valeur supérieure à 0,5 (en jaune dans le tableau 1).

TABLEAU 1 – *Matrice de corrélation des bandes spectrales de la prise de vue du 7 mars 2018 de l'orbite relative 50 sur la tuile 33NXB. Les cellules bleues correspondent aux valeurs les plus pertinentes pour la création d'indices de végétation sur cette prise de vue tandis que celles en jaune les couples de bandes pour les deux autres indices essayés dans l'étude.*

Bandes	B02	B03	B04	B05	B06	B07	B08	B8A	B11	B12
B02	1,00	-	-	-	-	-	-	-	-	-
B03	0,92	1,00	-	-	-	-	-	-	-	-
B04	0,89	0,89	1,00	-	-	-	-	-	-	-
B05	0,75	0,90	0,78	1,00	-	-	-	-	-	-
B06	-0,05	0,14	-0,14	0,37	1,00	-	-	-	-	-
B07	-0,14	0,01	-0,25	0,23	0,97	1,00	-	-	-	-
B08	-0,13	0,02	-0,22	0,20	0,85	0,87	1,00	-	-	-
B8A	-0,19	-0,05	-0,30	0,18	0,94	0,97	0,87	1,00	-	-
B11	0,12	0,26	0,17	0,54	0,72	0,66	0,59	0,67	1,00	-
B12	0,36	0,46	0,50	0,66	0,41	0,33	0,29	0,32	0,87	1,00

5.3.3. Résultats de la segmentation

La segmentation réalisée dans le but d'obtenir les polygones utilisés pour le calcul des *features* par objets a permis un découpage qui respecte bien les limites des éléments visibles dans la scène de composition RGB = B06 B12 B08. La figure 46 montre une petite partie de la zone d'étude où se trouvent plusieurs types de forêts bien distingués par le logiciel.

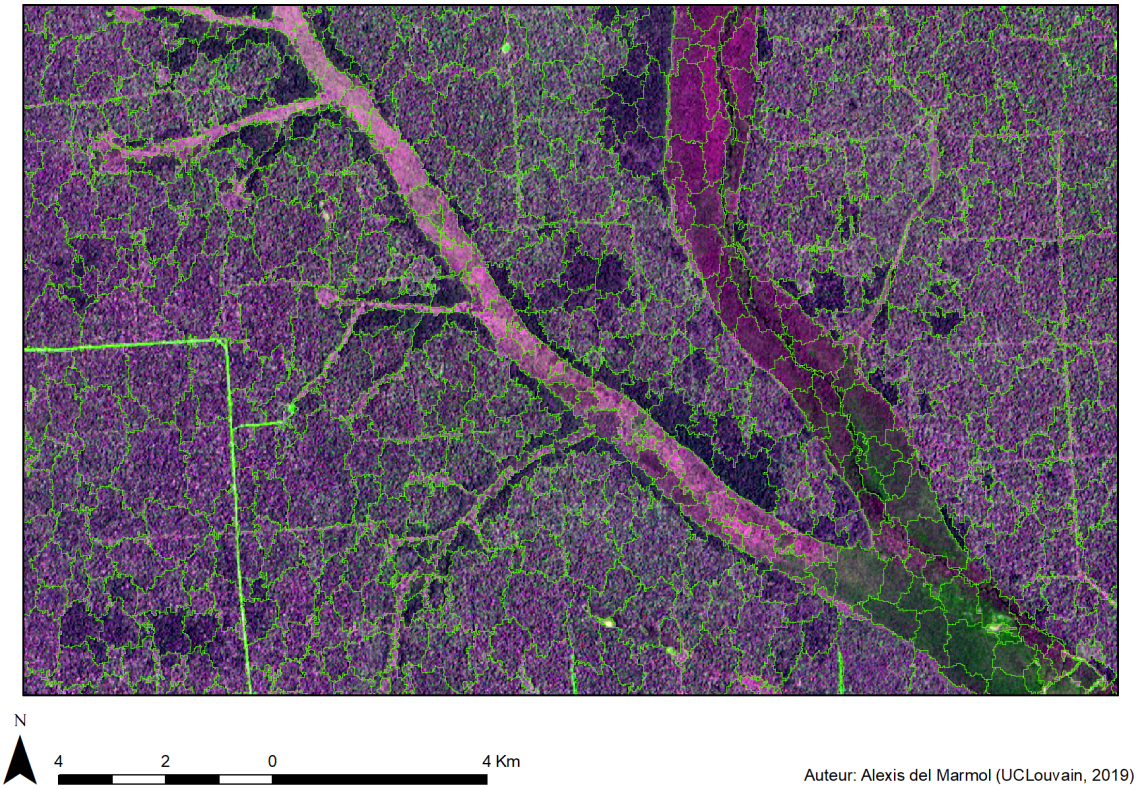


FIGURE 46 – *Zoom sur certains polygones obtenus par segmentation des features pixel-based retenues pour les classifications. Ce zoom permet de voir que le découpage suit bien les différents objets que l'on devine dans la scène du fond de carte de composition RGB = B06 B12 B08.*

5.4. Comparaisons de classifications finales

Treize combinaisons de *features* présentées dans le tableau 2 ont été utilisées comme données d'entrées des différentes classifications. Parmi ces combinaisons, neuf ont subi les trois approches : la classification par pixel, l'application du masque de classe majoritaire par polygones sur la première classification et la classification par objets. Il est important de préciser que la classification par objet des neuf premières combinaisons a été réalisée en utilisant les moyennes par objet des réflectances et des indices de végétation. Les réflectances des bandes 2 à 4 ne sont utilisées que dans une classification pour montrer les artéfacts qu'elles provoquent. Les 4 dernières combinaisons utilisent des *features* qui n'existent que par objet et n'ont donc subi que la troisième approche. Les combinaisons de *features* ont été posées au fur et à mesure de la génération des résultats des premières combinaisons sur base de leur analyse visuelle, en testant l'ajout ou le retrait d'une ou de plusieurs *features*. Au total, 31 produits ont été générés. ce résultat n'exclue nullement que l'utilité d'autres *features* puisse être mis en évidence par une multitude d'autres combinaisons.

La liste exhaustive des matrices de confusions générées à partir des classifications

TABLEAU 2 – *Différentes combinaisons de features considérées pour la discrimination forestière. Les 9 premières combinaisons ont subi les trois types de classification et les 4 dernières n'ont été classifiées que par polygones. Pour celles-ci, les moyennes et écarts-types s'appliquent sur l'ensemble des features de la combinaison indiquée.*

Combinaison	Bandes spectrales	Indices de végétation	
C1	[2 :12]	[NDVI, NDII, EVI2, NVI1 :NIV4]	
C2	[5 :12]	[NDVI, NDII, EVI2, NVI1 :NIV4]	
C3	[5 :12]	-	
C4	[5 :12]	[NDII, NVI2 :NIV4]	
C5	[5 :12]	[NDII, EVI2, NVI2 :NIV4]	
C6	[6 :12]	[NDII, NVI2,NIV4]	
C7	[6 :12]	[NDVI, NDII, NVI2,NIV4]	
C8	[5 :12]	[NDVI, NDII, NVI2,NIV4]	
C9	[5 :12]	[NDVI, NDII, EVI2, NVI2,NIV4]	
	Moyenne	Écart-type	Textures
C10	C2		X
C11	C6		X
C12	C2	C2	
C13		C2	

se trouve à l'annexe C. Dans ces matrices, les classes HF et EAU sont respectivement les zones hors forêts (bâtiments et sol nu) et l'eau. Elles sont considérées comme classes non végétales et sont donc grisées comme le montre la matrice de confusion de la C2 par objet dont l'OA est la plus élevée (Tableau 3).

Les OA ainsi que les κ calculés à partir de ces matrices sont synthétisés dans le tableau 4. Les valeurs des cellules colorées témoignent des meilleures classifications sur la base de la méthode de validation choisie dans cette étude. Ce tableau révèle à première vue que les classifications par objet donnent en moyenne de meilleurs résultats que les classifications par classe majoritaire qui sont, elles-mêmes, plus satisfaisantes que les classifications par pixel qui n'atteignent pas les 0,6 d'OA et 0,5 de κ. Les OA sont toujours supérieures aux κ qui leur correspondent du fait de la formule mathématique. On peut remarquer que les deux approches qui utilisent un découpage des classes par objets ont des résultats remarquablement différents. Cela montre que la classification peut varier de façon importante selon que l'on utilise les valeurs de réflectances moyennées par objets ou les valeurs brutes par pixels.

L'article publié par PONTIUS *et al.* (2011) met en lumière l'utilité de l'indice κ, ses propriétés mathématiques et ses limites graphiques dans les applications de la télédétection. Les données de ce mémoire seront, par conséquent, principalement validées sur base des OA en plus de l'analyse visuelle.

TABLEAU 3 – *Matrice de confusion de la C2 par objets. Les classes non végétales sont grisées.*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	0	0	45	0	5	0	0	0	87	0,43
FL	0	23	35	0	0	0	31	0	0	0	89	0,26
FJ	0	0	16	0	13	0	63	0	0	0	92	0,17
FA	0	0	2	83	5	0	3	0	5	0	98	0,85
FM	0	0	3	0	78	0	0	0	0	0	81	0,96
M	0	0	0	0	0	76	1	0	0	0	77	0,99
R	0	0	5	0	0	0	89	0	0	0	94	0,95
ZH	0	0	8	0	10	52	2	21	0	0	93	0,23
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	37	23	69	83	151	128	194	21	105	100		
UA	1,00	1,00	0,23	1,00	0,52	0,59	0,46	1,00	0,95	1,00	OA :	0,68
F-score	0,60	0,41	0,20	0,92	0,67	0,74	0,62	0,37	0,98	1,00	κ :	0,65

TABLEAU 4 – *OA et κ calculés à partir des matrices de confusions de trois types de produits pour les 13 combinaisons. Les cellules colorées en vert contiennent les meilleures valeurs (supérieures à 0,6*

	Overall Accuracy			κ		
	Pixel	Classe maj.	Objet	Pixel	Classe maj.	Objet
C1	0,5423	0,6081	0,5346	0,4911	0,5640	0,4828
C2	0,5104	0,5971	0,6839	0,4560	0,5525	0,6490
C3	0,5439	0,6367	0,6268	0,4932	0,5964	0,5853
C4	0,5401	0,6630	0,6586	0,4890	0,6256	0,6209
C5	0,5368	0,5653	0,6729	0,4853	0,5169	0,6368
C6	0,4885	0,5357	0,6784	0,4319	0,4842	0,6429
C7	0,4984	0,5236	0,6334	0,4426	0,4708	0,5929
C8	0,5214	0,5829	0,6235	0,4682	0,5364	0,5817
C9	0,5357	0,5664	0,5851	0,4841	0,5181	0,5390
C10			0,6345			0,5940
C11			0,5928			0,5478
C12			0,6026			0,5585
C13			0,4415			0,3778

L'étude plus approfondie des matrices de confusion permet de comprendre l'origine des erreurs dans la classification.

En premier lieu, les deux classes non végétales sont presque toujours parfaitement classées. Cela s'explique par leurs signatures spectrales particulièrement différentes des classes qui contiennent de la végétation. En second lieu, les types de forêts confondus coïncident dans les différents types de classifications. Enfin, les classifications par pixel présentent les erreurs les plus hétérogènes tandis que les classifications par classe majoritaire et par objet conjuguent les contaminations dans une ou deux autres classes en moyenne.

On remarque, par exemple, que les FR sont souvent confondues avec les FM. Intuitivement, cela pourrait s'expliquer par une ressemblance forte entre les propriétés microscopiques internes des feuilles de marantacées et de celles des euphorbiacées qui ont la même allure macroscopique. Cela engendre une similitude importante dans la signature spectrale des deux types.

Les FJ sont, quant à elles, souvent confondues avec les FA, surtout par classe majoritaire. Cela s'explique par le fait que ces deux types de forêts sont une continuité l'un de l'autre. Ces deux classes sont également parfois confondues avec les limbalis et les M. Ces dernières sont également composées d'un mélange de feuillus qui peut porter à confusion. Les limbalis sont également des arbres feuillus et sont presque toujours en contact direct avec une forêt jeune ou adulte, ce qui peut provoquer des erreurs lorsque les points de validation sont trop proches d'une limite entre les deux types. Cette proximité des points de validation avec des frontières, parfois mal marquées entre types forestiers, peut aussi expliquer que, pour certaines compositions de *features*, les R sont classés en FL (p. ex. les combinaisons 3 et 4). Cette explication n'est toutefois valide que lorsque la classification se fonde sur les polygones issus de la segmentation.

Enfin, les zones herbacées sont confondues le plus souvent avec les milieux urbains où l'on trouve également des pelouses et prairies. Les quelques confusions avec les FM peuvent s'expliquer par le fait que des marantacées se trouvent souvent en lisières des clairières de types éyangas et baïs.

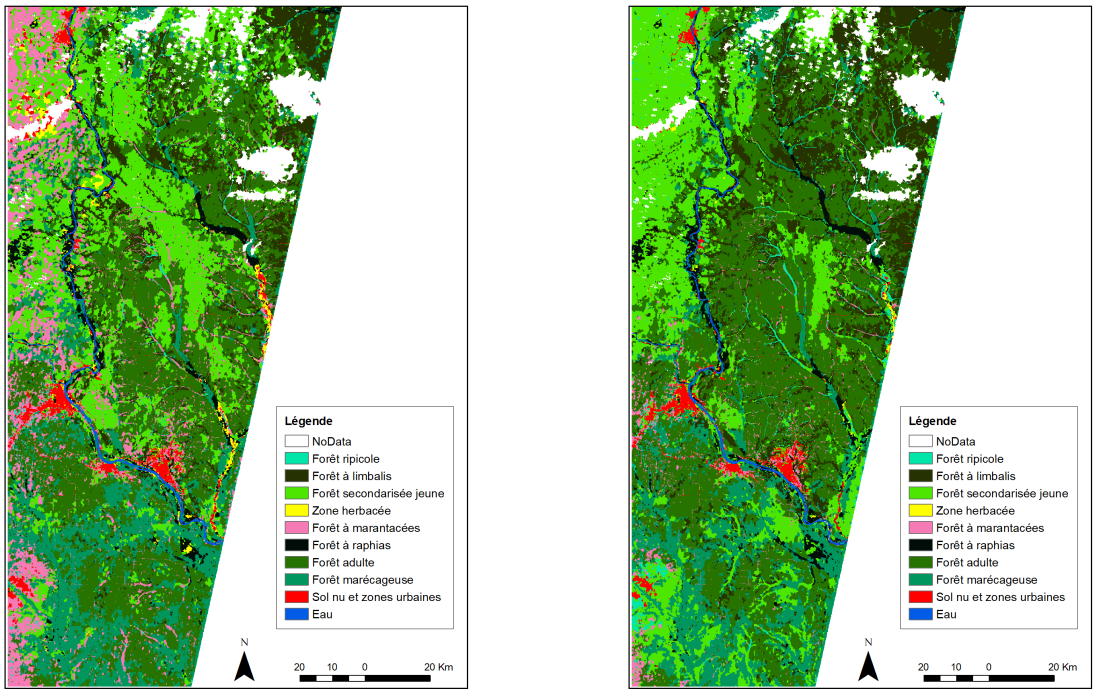
La nature même des types de forêts provoque ces confusions, mais celles-ci peuvent être réduites par une amélioration des aires d'entrainement et des points de validation. Les aires d'entrainement sont dessinées sur la base de la clé d'interprétation visuelle. Celle-ci a ses limites : elle n'explique pas toutes les nuances visibles dans les images et sur le terrain pour chacune des classes. L'interprétation de cette clé peut être, elle-même, source d'erreurs. Enfin, les aires d'entrainement sont choisies pour être les plus représentatives des classes et les nuances et dégradés qui existent à l'intérieur des pixels d'une même classe peuvent poser problème à l'algorithme. Malgré la certitude de la classe qui leur est associée, les points de validation sont eux aussi soumis à la diversité à l'intérieur

de chacune des classes. En effet, les nuances visibles sur le terrain sont telles qu'elles pourraient mener à la distinction de plus de 40 classes différentes, mais les limites de ce travail ont demandé de réduire ce nombre à 10 pour simplifier le développement de la méthodologie. Les polygones qui représentent les placettes visitées et qui ont servi à la génération des points de validation sont soumis à une erreur propre à la complexité du déplacement en forêt. En effet, chaque sortie en forêt avait pour objectif d'être la plus rentable possible en termes de nombre de placettes visitées. Certaines placettes sont donc proches les unes des autres et se situent parfois dans des formations végétales peu étendues. Si les effets de bordures ont été réduit au maximum en s'éloignant d'une centaine de mètres ou plus des limites visibles sur les images satellites et sur le terrain, cela ne les empêche pas d'être encore présents. La segmentation peut aussi amplifier ces effets de bordure par un découpage décalé de quelques pixels aux endroits échantillonnés.

Pour rencontrer ces problématiques il conviendra de travailler la pertinence et la représentativité des aires d'entrainement et des points de validation. Cela peut se faire en testant plusieurs classifications sur la base de différentes séries d'aires d'entrainement et en améliorant le protocole de terrain et la pertinence des zones visitées avec les classes prédéfinies à l'avance. Cela peut aussi se faire en travaillant sur la pondération des points de validation par la surface de l'image occupée par chacune des classes afin d'améliorer la représentativité de cette validation. une telle pondération est illustrée par OLOFSSON *et al.* (2013) dans leur étude qui présente différentes façons de réaliser des validations en tenant compte des surfaces. Concrètement, cela revient à donner plus de poids aux points qui représentent les classes les plus présentes. De cette façon, le résultat n'est pas tiré vers le bas par les classes moins représentées telles que les FM souvent difficiles à discriminer au cœur des massifs de FJ et de FA.

Les limitations et biais provoqués par les points de validation impliquent que les matrices de confusion, l'OA et les indices κ ne sont pas les seuls indicateurs de la qualité des classifications. Leur étude visuelle joue un rôle tout aussi important et peut révéler une classification bonne dans son ensemble, mais mauvaise autour de lieux visités sur le terrain (mauvaise matrice de confusion) ou, au contraire, une classification bonne autour de points de validation (bonne matrice de confusion), mais montrant des classifications aberrantes hors de la zone de terrain. Ceci peut être illustré en comparant la meilleure classification du point de vue statistique (C2 par objets d'OA = 68,39%) (Tableau 3) avec la classification de la combinaison 5 par classe majoritaire (OA = 56,53%) comme le montre la figure 47. La première donne un meilleur résultat statistique que la seconde, mais montre une contamination de M hors du réseau hydrographique à l'ouest et au sud de la rivière Sangha ainsi que dans le nord-est. Les FM sont également trop représentées à l'ouest de la Sangha là où devraient se trouver plus de FJ. La seconde montre, quant à elle, une présence trop importante de FL dans le nord-est et une représentation de FA plus importante en comparaison aux FJ légèrement plus présentes en réalité. Visuellement, la classification de la combinaison 5 par classe majoritaire donne

un résultat plus satisfaisant, mais la matrice de confusion donne un résultat moins bon (Tableau 5).



(a) *(b)*

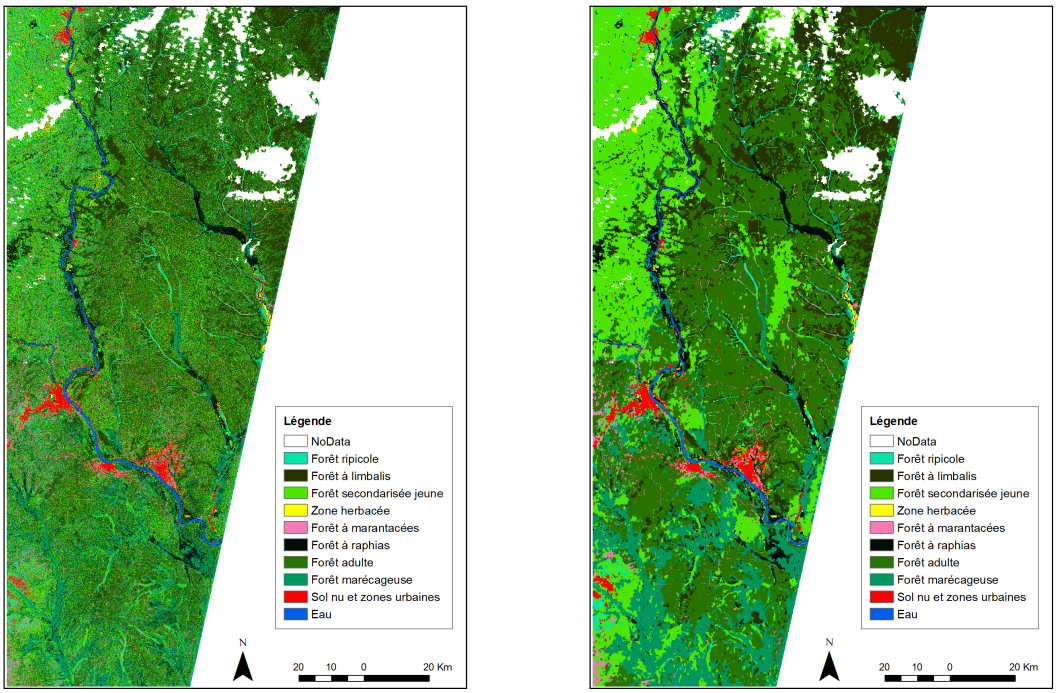
FIGURE 47 – *Comparaison des classifications de la C2 par objets (a) et de la C5 par classe majoritaire (b).*

Les mauvais résultats obtenus par les classifications par pixel s'expliquent par l'effet «poivre et sel» bien visible. En considérant les pixels indépendamment et non pas par ensemble, l'algorithme n'est pas capable de s'affranchir de la variabilité existante au sein d'un même type de forêt. Les pixels correspondent à une surface de 20 x 20 m sur le terrain et certains arbres émergents peuvent remplir facilement cette surface. De cette façon, le pixel ne représente pas la structure et les propriétés d'une formation forestière, mais uniquement les caractéristiques d'un seul arbre qui, pris indépendamment du reste, peut se rapprocher fortement d'un limbali ou d'un arbre de marécage. Ainsi, même si les formes générales prisent par les différentes classes et perçues en prenant du recul sur l'image permettent de distinguer correctement les contours des grandes zones tels que le réseau hydrographique ou les zones urbaines, le résultat n'est pas pertinent lorsqu'on regarde de près par rapport à l'objectif de représenter des ensembles homogènes. En attribuant à chaque polygone la valeur de la classe majoritaire par pixel, on corrige cet effet en passant à l'échelle d'ensemble forestier. Cependant, cette technique peut également amplifier les erreurs d'une classification par pixel (Figure 48).

Il est donc important de prendre en considération à la fois les mesures statistiques (OA) et l'analyse visuelle pour juger de la qualité d'une classification de la couverture du sol. Sur cette base qui combine une part d'objectivité biaisée par la validation et une

TABLEAU 5 – *Matrice de confusion de la C5 par classe majoritaire*

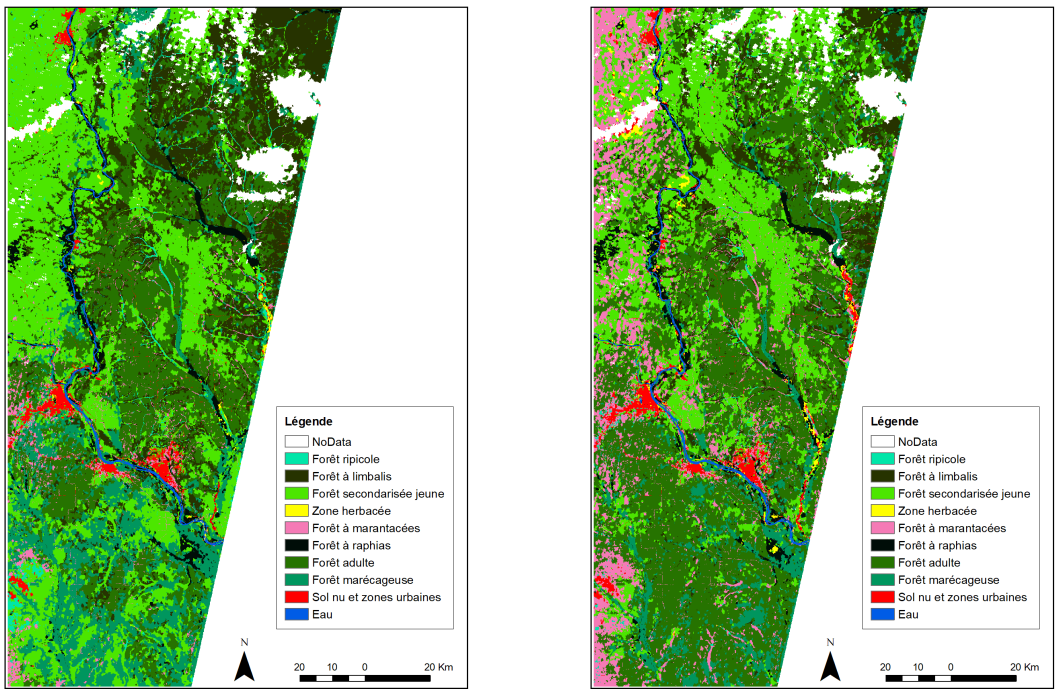
Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	Σ	PA
FR	37	0	50	0	0	0	0	0	0	0	87	0,43
FL	0	41	0	0	0	0	13	35	0	0	89	0,46
FJ	0	0	8	0	0	0	84	0	0	0	92	0,09
ZH	0	0	44	27	1	15	0	0	11	0	98	0,28
FM	72	0	3	0	6	0	0	0	0	0	81	0,07
R	0	0	0	0	0	76	1	0	0	0	77	0,99
FA	0	0	0	0	0	0	94	0	0	0	94	1,00
M	0	5	10	0	0	52	0	26	0	0	93	0,28
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
Σ	109	46	115	27	7	143	192	61	111	100		
UA	0,34	0,89	0,07	1,00	0,86	0,53	0,49	0,43	0,90	1,00	OA :	0,57
F-score	0,38	0,61	0,08	0,43	0,14	0,69	0,66	0,34	0,95	1,00	κ :	0,52



(a) *(b)*

FIGURE 48 – *Comparaison des classifications de la C3 par pixels (a) et par classe majoritaire (b). La seconde classification semble avoir fortement augmenté la présence de forêts marécageuses dans le Sud de la zone.*

part de subjectivité, la meilleure composition est la C4 qui donne un assez bon résultat au niveau du score de l'OA par classe majoritaire et par objet (> 62% pour les deux approches). L'interprétation visuelle est cependant à nuancer, car le résultat de l'approche par classe majoritaire respecte bien mieux la réalité de terrain. Ces deux classifications, par classe majoritaire et par objet de la combinaison 4, sont présentées à la figure 49.



(a) *(b)*

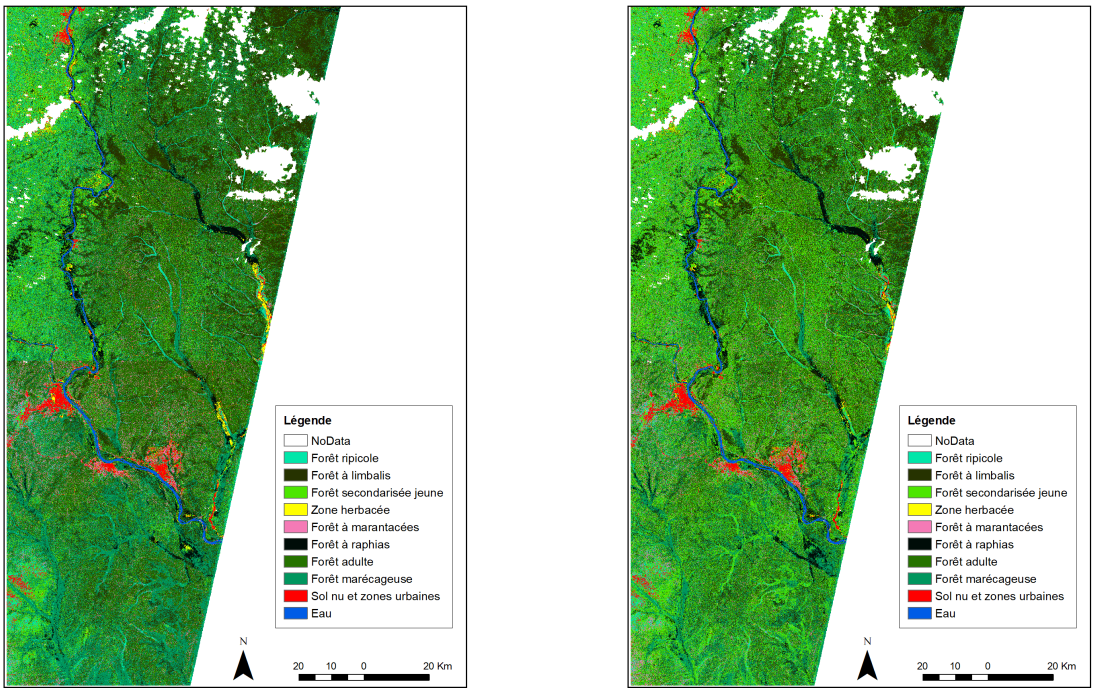
FIGURE 49 – *Présentation des classifications de la C4 par classe majoritaire (a) et par objets (b) qui ressortent comme les deux meilleures classifications de cette étude*

Outre la détermination de la combinaison qui offre la meilleure classification, la comparaison des résultats issus des différentes combinaisons permet de mettre en évidence l'utilité de certaines *features* dans la classification. Mise à part la première comparaison, les résultat d'une approche par pixel ne sont pas étudiées à cause de leur κ médiocre, de leur OA peu satisfaisant et de leur visuel peu représentatif. Chaque comparaison est accompagnée d'un tableau reprenant uniquement les valeurs d'OA des combinaisons concernées extraites du tableau 4 .

— La comparaison de la C1 avec la C2 montre l'impact sur la classification des bandes les plus sensibles à l'effet de jointure de la mosaïque. Les résultats issus de la C1, ne sont intéressant que pour l'approche par classe majoritaire. Cela s'explique par l'analyse visuelle où la jointure de la mosaïque est très nettement visible alors qu'elle ne l'est pas du tout sur les classifications issues de la combinaison 2. Ce phénomène est particulièrement visible sur l'approche par pixel présentée à la figure 50.

OA des C1 et C2 suivant les trois approches

	Pixel	Classe Maj.	Objet
C1	0,5423	0,6081	0,5346
C2	0,5104	0,5971	0,6839



(a) *(b)*

FIGURE 50 – *Comparaison des classifications de la C1 par pixels (a) et de la C2 par pixels (b). On distingue clairement la rupture entre les deux tuiles lorsque les bandes 2 à 4 sont utilisées.*

— La comparaison de la C2 avec la C3 pose la question de l'utilité de l'ensemble des indices de végétation. L'étude des matrices de confusions montre que les indices de végétation augmentent l'OA par objet, mais leur absence l'augmente pour l'approche par classe majoritaire. Visuellement, les deux approchent donnent des résultats très différents pour une même combinaison. Ce n'est donc pas possible de conclure à l'importance des indices de végétation avec les résultats obtenus dans ce travail.

OA des C2 et C3 suivant les trois approches

	Pixel	Classe Maj.	Objet
C2	0,5104	0,5971	0,6839
C3	0,5439	0,6367	0,6268

— La comparaison de la C2 avec la C4 montre l'impact des indices de végétation lorsqu'ils sont marqués par l'effet de jointure de la mosaïque dans l'utilisation de

la bande 4. Enlever le NDVI et le NVI1 augmente la proportion de FA et de M dans la classification et améliore les matrices de confusion. Les différences sont faibles mais les interprétations visuelles confirment que l'absence de ces deux indices donne un résultat de meilleure qualité.

OA des C2 et C4 suivant les trois approches

	Pixel	Classe Maj.	Objet
C2	0,5104	0,5971	0,6839
C4	0,5401	0,6630	0,6586

— La comparaison de la C4 avec la C5 montre l'utilité du EVI2. L'approche par polygone ne montre que très peu de nuances, mais l'approche par classe majoritaire met en évidence l'utilité de l'indice pour augmenter l'OA. L'indice permet également d'augmenter la proportion de FJ dans la zone proche de Pokola, ce qui correspond mieux à la réalité.

OA des C4 et C5 suivant les trois approches

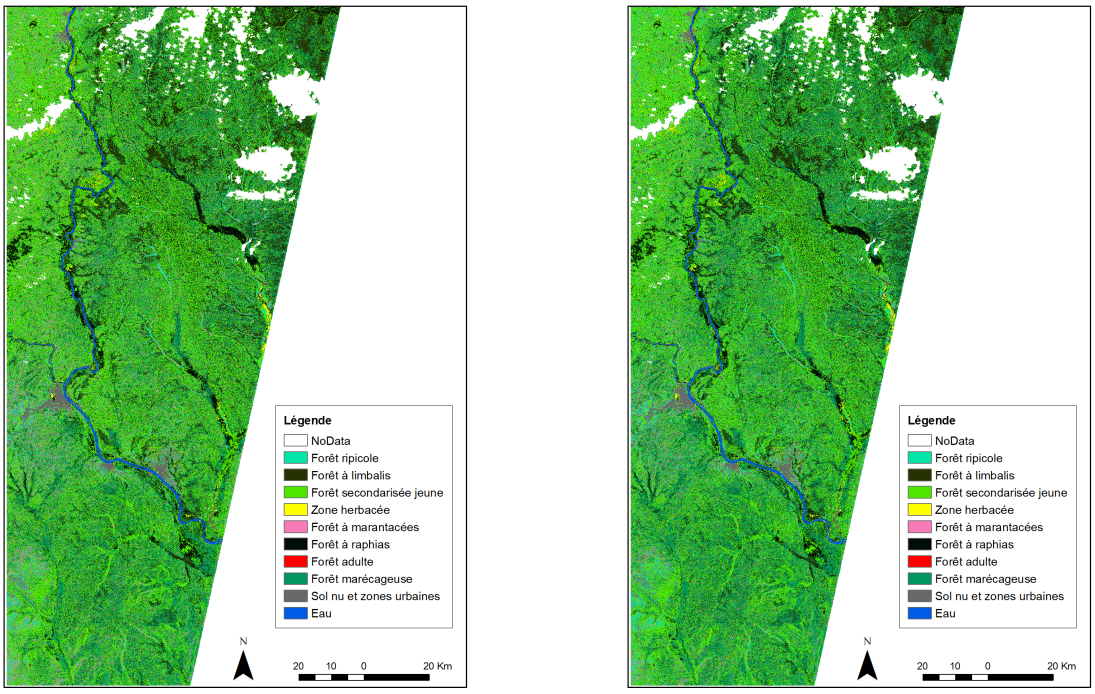
	Pixel	Classe Maj.	Objet
C4	0,5401	0,6630	0,6586
C5	0,5368	0,5653	0,6729

— La comparaison de la C8 avec la C7 et C6 montre d'une part l'importance de l'utilisation de la bande 5 qui reste un peu touchée par l'effet de jointure de la mosaïque et d'autre part, l'utilité du NDVI. Un phénomène étonnant se produit lors de la classification par pixel (et donc par classe majoritaire également) en l'absence de la bande 5 : la classe de FA est presque complètement omise (Figure 51). Afin de mettre en évidence le phénomène, la légende a été modifié et le rouge a été attribué à la classe FA. Cela ne se produit pas lors d'une classification par objet. Une explication à ce phénomène réside dans la répartition des valeurs de réflectances des deux classes dans la bande 5. Malgré une superposition importante, les courbes des valeurs par pixels des deux classes se distinguent l'une de l'autre. En effet, la moyenne de réflectance (en valeur unitaire x 10 000) pour les FA se situe autour de 600 avec des maxima limités à 700. La moyenne pour les FJ se positionne autour de 640 avec des maxima dépassant les 800 et les écarts-types de leur répartition ne dépasse pas 70 alors qu'il va de 100 à plus de 350 pour ces deux classes dans les bandes 6 à 11. La bande 12 donne un écart-type du même ordre que la bande 5 mais les maxima et minima des deux classes se superposent quasiment contrairement courbes dans cette dernière. La bande 5 est donc essentielle pour distinguer ces deux classes dans une approche par pixel. Lors d'une classification par polygones, le phénomène ne se produit pas, car les écarts-types sont bien plus faibles (inférieur à 20 dans les deux classes pour la

plupart des bandes). Les courbes sont alors différenciables par l'algorithme dans toutes les bandes.
Une autre constatation est que l'utilisation du NDVI apporte une contamination plus importante de M en l'absence de FA. Enfin, la différence de résultat avec ou sans cet indice lors de la classification par objet n'est pas importante, mais elle peut se remarquer légèrement dans la matrice de confusion où l'on observe que les ZH sont mieux classées et plus présentes lorsque cette *features* est absente.

OA des C6, C7 et C8 suivant les trois approches

	Pixel	Classe Maj.	Objet
C6	0,4885	0,5357	0,6784
C7	0,4984	0,5236	0,6334
C8	0,5214	0,5829	0,6235



(a) *(b)*

FIGURE 51 – *Présentation des classifications de la C6 par pixels (a) et de la C7 par pixels (b). Ces deux classifications sont marquées par l'absence presque totale de la classe FA légendée en rouge afin de mettre en évidence cette absence.*

— La comparaison de la C9 avec la C8 montre l'impact de NVI3, indice construit à partir d'une corrélation élevée entre deux bandes. Au niveau des matrices de confusion ainsi qu'au niveau du visuel, l'absence de cet indice améliore la classification. L'ajout de *features* non pertinentes influence donc négativement l'algorithme. Cette conclusion est étonnante lorsqu'on utilise l'algorithme RF mais le

nombre relativement réduit de *features* utilisées au total qui donne un poids plus important à chacune d'elles peut être la cause de ces erreurs.

OA des C8 et C9 suivant les trois approches

	Pixel	Classe Maj.	Objet
C8	0,5214	0,5829	0,6235
C9	0,5357	0,5664	0,5851

— La comparaison de la C10 avec la C2 montre l'apport des textures en plus des bandes et des indices de végétation. La C2 classée par polygone permet d'obtenir l'OA la plus élevée. Cependant, la différence est trop faible pour être significative et pour pouvoir conclure de l'utilité des textures. Cependant, elles ont tendance à contaminer les FJ et FA par des M et des FM.

OA des C2 et C10 suivant les trois approches

	Pixel	Classe Maj.	Objet
C2	0,5104	0,5971	0,6839
C10			0,6345

— La comparaison de la C11 avec la C10 et la C2 questionne l'utilité relative des indices de végétations par rapport aux textures. Prises sans les indices de végétation, les textures ne parviennent pas à rendre compte correctement de la présence des M. Ces *features* contaminent l'ensemble de la zone d'étude par cette classe sans suivre le réseau hydrographique. Le résultat de la matrice de confusion de la C11 est également inférieur aux résultats de celles des C10 et C2.

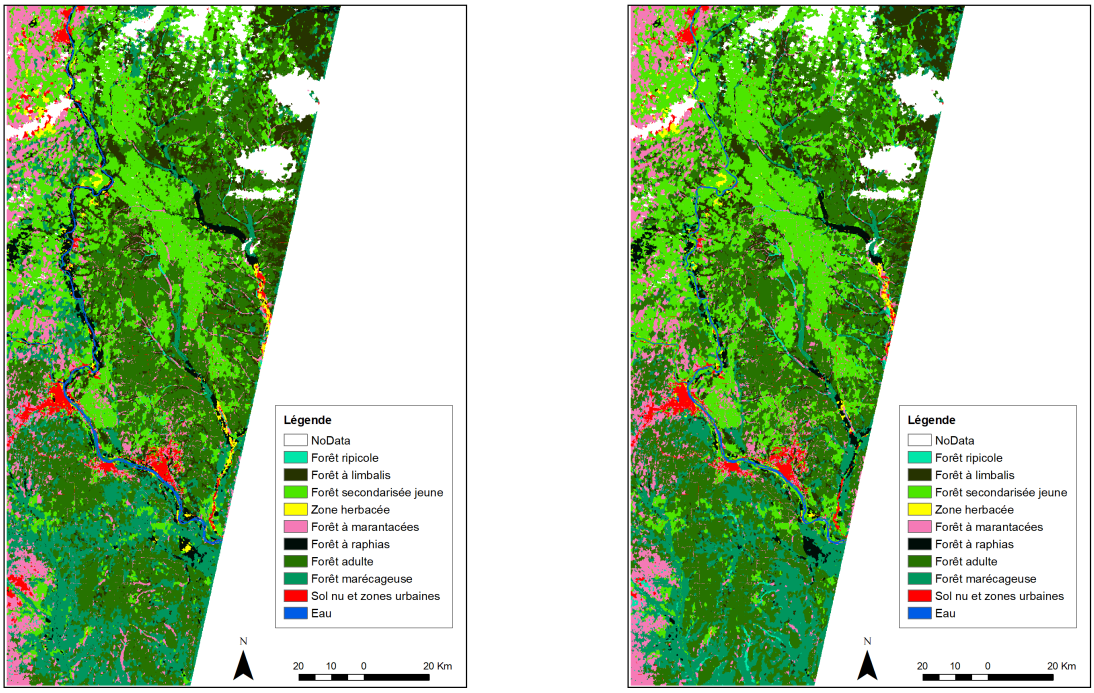
OA des C2, C10 et C11 suivant les trois approches

	Pixel	Classe Maj.	Objet
C2	0,5104	0,5971	0,6839
C10			0,6345
C11			0,5928

— La comparaison de la C12 avec la C2 montre l'impact de l'ajout des écarts-types aux moyennes par polygones. L'ajout des écarts-types donne un résultat qui ne donne pas les OA les plus élevées mais montre une amélioration visuelle nette du découpage des M au sud et à l'ouest de la Sangha, contrairement à la C2. Dans le nord, les contaminations de cette même classe sont également réduites. Cependant, on observe une diminution visible des R sur les berges de la Sangha et des erreurs plus importantes dans la zone de validation qui explique une valeur d'OA supérieure pour la C2 (Figure 52).

OA des C2 et C12 suivant les trois approches

	Pixel	Classe Maj.	Objet
C2	0,5104	0,5971	0,6839
C12			0,6026



(a) *(b)*

FIGURE 52 – *Comparaison des classifications de la C2 par objets (a) et de la C12 par objets (b). L'ajout des écarts-types a un effet positif sur la classification des M dans le sud et l'ouest de la Sangha, mais provoque des erreurs dans la zone de validation ainsi que la disparition de R le long de la Sangha.*

— La comparaison de la C13 avec la C12 et la C2 met en lumière le résultat d'une utilisation des seuls écarts-types sans les moyennes. cette approche produit la plus mauvaise des classifications. Visuellement, une grosse partie des M et des R est classée de façon incorrecte en FL.

OA des C2, C12 et C13 suivant les trois approches

	Pixel	Classe Maj.	Objet
C2	0,5104	0,5971	0,6839
C12			0,6026
C13			0,4415

Cette analyse montre la plus-value d'une classification qui se base sur un découpage d'objets. En effet, l'étude s'intéresse à la composition des forêts à l'échelle du peuple-

ment et au-delà, ce que l'approche par pixel ne permet pas de faire de façon pertinente. À cela s'ajoute l'importance du choix des *features* qui peuvent diminuer la qualité de la classification si elles sont inutiles ou lorsqu'elles sont de mauvaise qualité. C'est le cas des bandes spectrales influencées par l'effet de jointure de mosaïque ainsi que des indices de végétation qui utilisent ces bandes. Enfin, les indices de végétation, les textures et les écarts-types qui vont de pair avec une diminution de la qualité ne semblent pas apporter d'information pertinente à l'algorithme.

La consolidation par vote majoritaire est le résultat analysé ci-après (Figure 53). Il est construit à partir des classifications suivantes :

— C1 par classe majoritaire
— C3 par classe majoritaire
— C4 par classe majoritaire
— C5 par classe majoritaire
— C8 par classe majoritaire
— C2 par polygone
— C12 par polygone

L'analyse visuelle de ce produit donne un résultat qui pourrait être considéré comme le meilleur parmi tous. Cependant, quelques contaminations continuent à être présentes par rapport aux autres classifications : une contamination de la classe HF (en rouge) au bord des cours d'eau principaux du réseau hydrographique à la place de la classe ZH, la présence de FM dans les zones ripicoles aux embranchements les plus fins du réseau et de la contamination de FJ dans les zones marécageuses.

La matrice de confusion est présentée au tableau 6. L'OA de cette matrice est du même ordre de grandeur que les valeurs des classifications qui sont à l'origine de ce produit consolidé (65%). L'analyse plus approfondie des F-score des différentes classes montre que les classes les plus mal classées sont les FJ et les M. Les premières sont très souvent classées en FA tandis que les secondes sont prédites en R et en FJ dans la plupart des cas. Cette observation confirme les analyses réalisées sur les autres classifications et les difficultés liées aux similitudes des caractéristiques géographiques et physionomiques entre ces différents types forestiers.

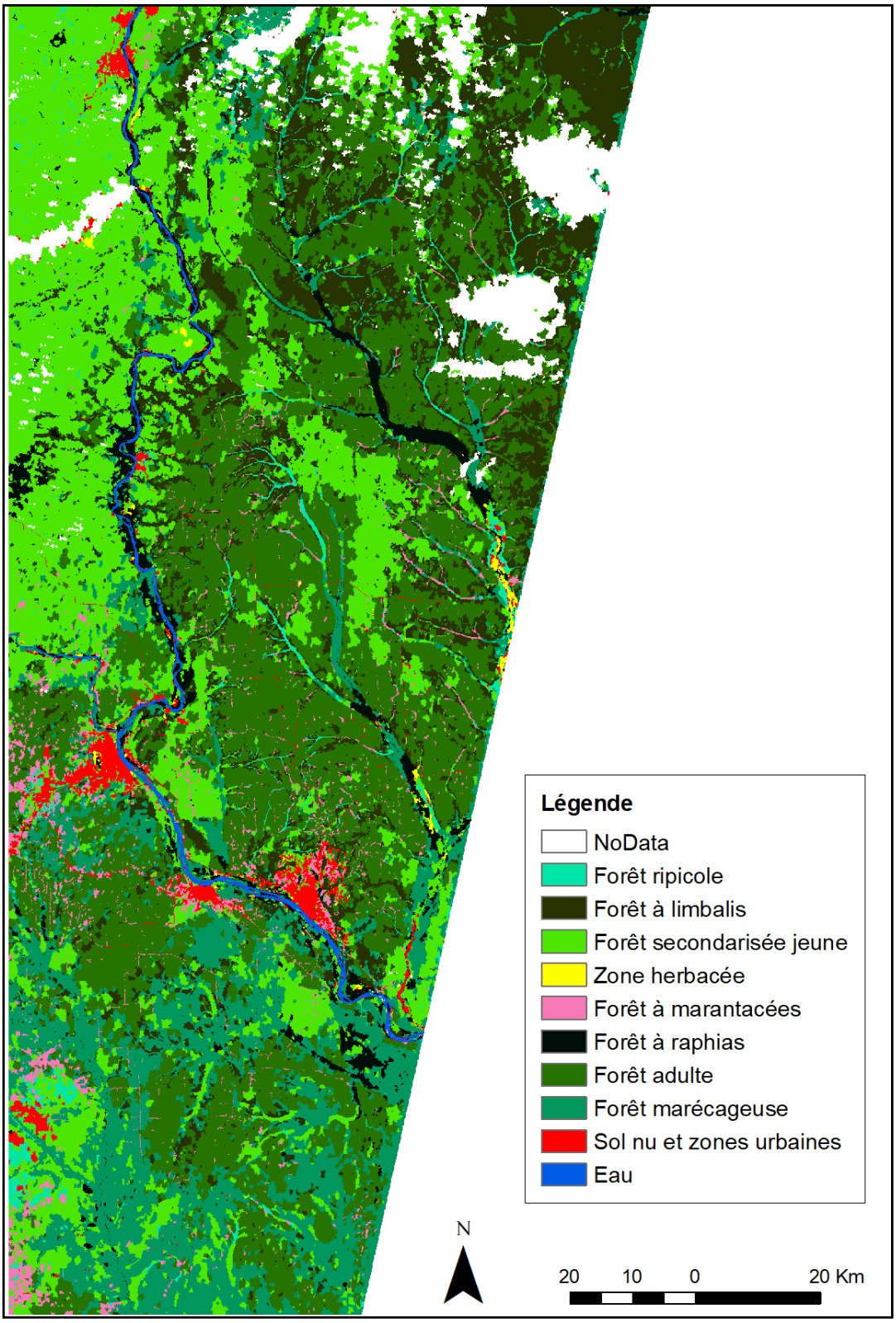


FIGURE 53 – *Résultat du vote majoritaire réalisé sur les combinaisons C1, C3, C4, C5 et C8 par classe majoritaire et C2 et C12 par polygone. Il s'agit de l'un des meilleurs rendus visuels obtenus dans ce mémoire.*

TABLEAU 6 – *Matrice de confusion du résultat de classification consolidé par vote majoritaire*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	50	0	0	0	0	0	0	0	87	0,43
FL	0	32	0	0	0	0	22	35	0	0	89	0,36
FJ	0	0	15	0	0	0	77	0	0	0	92	0,16
ZH	0	0	30	49	1	15	3	0	0	0	98	0,50
FM	14	0	3	0	64	0	0	0	0	0	81	0,79
R	0	0	0	0	0	76	1	0	0	0	77	0,99
FA	0	0	4	0	0	0	90	0	0	0	94	0,96
M	0	5	10	0	0	52	0	26	0	0	93	0,28
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	51	37	112	49	65	143	193	61	100	100		
UA	0,73	0,86	0,13	1,00	0,98	0,53	0,47	0,43	1,00	1,00	OA :	0,65
F-score	0,54	0,51	0,15	0,67	0,88	0,69	0,63	0,34	1,00	1,00	κ :	0,61

5.5. Recommandations

Les réserves qui doivent être placées sur les résultats présentés dans ce travail invitent à la réflexion sur les moyens à envisager et les précautions à adopter pour les prochaine études sur le sujet.

Premièrement, le prétraitement des images en L2A au moyen d'un processeur plus performant que Sen2Cor, tel que MACCS, pourrait probablement éviter les problèmes survenus lors de la création des mosaïques. Par conséquent, l'utilisation des bandes 2, 3 et 4 comme *features* et, par la même occasion, des indices de végétations qui en découlent, n'aurait plus d'impact négatif sur la continuité de la classification au niveau de la jointure des tuiles. Les couples ayant les valeurs les plus faibles de la matrice de corrélation, concernant souvent ces bandes, pourront alors être sélectionnés pour la construction de nouveaux indices performants. Malheureusement, l'ESA fournit par défaut des données de réflectance de surface prétraitées par Sen2Cor ou des données brutes (L1C) qu'il convient de prétraiter soi-même avec MACCS.

La mission de terrain est une étape cruciale dans la réalisation d'un produit cartographique de ce genre. Pourtant, les contraintes qui existent dans ce genre de milieu rendent le déplacement en forêt difficile. Les infrastructures et la logistique mises à disposition de la CIB ont grandement aidé à la réalisation de cette étape du mémoire. Cependant, quelques aménagements faciliteraient la construction d'un jeu de données plus pertinent. Premièrement, la détermination des classes a priori aurait permis de guider l'échantillonnage de façon à choisir un nombre de placettes par type de forêt propor-

tionnel à la surface occupée par chacun d'eux. Deuxièmement, le fait que les ensembles d'un type de forêt sont relativement étendus et que les pixels des images traitées ne sont que de 20 x 20 m rend la dimension de placettes de 50 x 50 m peu pertinente, car elles ne couvrent que de toutes petites fractions de la surface totale de l'ensemble. Il serait donc intéressant de travailler par transects de 500 à 1000 m en positionnant des points GPS sur l'ensemble du parcours et en prenant note, le cas échéant, des changements de type de forêt afin d'avoir la possibilité de réaliser plusieurs polygones de validation pour un même déplacement en forêt. Ce type d'échantillonnage est facilité par la présence des layons d'exploitations passées et en cours de traçage ainsi que par la présence des sentiers tracés par les autochtones pour la chasse et la cueillette. Enfin, le nombre de sorties sur le terrain pourrait être augmenté en rallongeant la durée du séjour et en prévoyant l'agenda quotidien dans une meilleures coordination avec l'entreprise d'accueil de façon à se synchroniser avec sa logistique.

La région étudiée est soumise à des variations saisonnières faibles, entre janvier et juin, qui pourraient être intéressantes à introduire comme *features* supplémentaires. En effet, les variations de signature spectrales peuvent différer au cours de l'année selon les types de forêts, notamment par la saisonnalité des conditions hydromorphiques en milieu inondable. Cette analyse temporelle demande d'avoir accès à des images à faible couverture nuageuse durant ces deux périodes de l'année. Au fur et à mesure des années, la quantité d'images et le recul dans le temps augmentera, rehaussant la probabilité de bonnes prises de vues à différentes saisons. Cette hypothèse est soutenue par la figure 9 qui démontre la probabilité élevée d'obtenir de bonnes images sur une grande période de l'année.

Dans le même ordre d'idée, une observation temporelle avec la résolution spatiale de S2 conduirait à étudier la re-fermeture du couvert après exploitation forestière et de mettre ainsi en évidence l'impact des pratiques en comparant l'évolution des FJ en FA dans le temps et l'espace. Dès lors que cette résolution spatiale permet de voir à l'échelle de la couronne de certains arbres dominants, il pourrait être intéressant de se baser sur cette propriété des images pour développer une méthode qui met en évidence la diminution d'essences commerciales entre une forêt secondarisée adulte et une forêt primaire.

La classification en elle-même doit être peaufinée. Les produits réalisés par ce travail sont tous issus d'une classification unique, mais une façon d'améliorer la classification serait de procéder de façon hiérarchique : une première classification permet de séparer les milieux forestiers des milieux non forestiers. Les premiers sont ensuite soumis à une seconde classification pour isoler le réseau hydrographique des zones de terre ferme. Les classes M, R et FR sont alors uniquement utilisées dans la classification du réseau hydrographique et les classes FJ, FA et FM ne sont classées que dans les zones de terre ferme. Les classes présentes sur les deux types de sols (FL, ZH) peuvent être utilisées pour l'en-

trainement des deux classifications. Cela permettrait d'une part d'assurer l'absence de contaminations des classes spécifiques à un milieu dans un autre et de discuter de l'utilité des *features* hiérarchiquement (p. ex. l'utilité d'un indice de végétation pour mettre en évidence les milieux inondés).

Enfin, un des objectifs de ce mémoire est d'apporter une méthode qui aidera à la réalisation d'une cartographie nationale avec une grande diversité floristique. La diversité qui existe sur une zone telle que l'ensemble du bassin du Congo nécessite de prendre des mesures nécessaires pour ne pas multiplier les marges d'erreurs avec la surface étudiée. Il conviendra d'avoir une bonne connaissance du terrain dans les différentes régions afin d'avoir un jeu de données de validation solide et afin de pouvoir construire une clé d'interprétation visuelle variable selon la région pour la construction des données d'entrainement. Enfin, le choix des classes est particulièrement important : il convient de procéder à une balance des intérêts entre un choix entre un petit nombre de classes générales et un grand nombre de classes spécifiques pour trouver un équilibre entre la complexité de la réalisation de la classification et la richesse du produit cartographique. Une méthode pourrait, par exemple, consister en la réalisation de plusieurs classifications localisées qui seraient combinées manuellement pour obtenir un produit cartographique final de haute qualité.

CONCLUSION ET PERSPECTIVES

En se basant sur les connaissances actuelles en télédétection des forêts tropicales, cette étude de la faisabilité d'une classification des types de forêts par imagerie S2 a pu être réalisée dans les forêts peu connues de la région d'Ouesso, dans le nord de la République du Congo.

Au cours d'une mission de terrain, les positions GPS de placettes représentatives de huit types de forêts ont été récoltées et leurs caractéristiques physionomiques ont été décrites. Ces connaissances ont mené à la création d'une clé d'interprétation visuelle des images S2 colorées selon la composition RGB = B06 B12 B08. La photo-interprétation qu'il est possible de réaliser avec la clé visuelle réponds à la première question en aboutissant à l'identification des types de forêts de manière exacte.

Le prétraitement de ces images a mis en évidence des problèmes liés à la création de mosaïques d'images de dates différentes, d'orbites relatives différentes et de tuiles différentes. L'analyse des bandes spectrales a permis la création de nouveaux indices de végétation qui mettent en évidence certaines propriétés des types de forêts. Une segmentation a mené à la création des variables par objet nécessaires à la comparaison des différentes approches de classification, à savoir par pixel, par classe majoritaire et par objet. Les classifications ont été réalisées par un algorithme de *random forest* en utilisant la clé d'interprétation visuelle pour tracer les aires d'entrainement. Le bon fonctionnement de cette clé a pu être mis en évidence. La validation des classifications s'est basée sur les données de terrain. Cette validation couplée à l'interprétation visuelle a permis d'identifier la méthode par objet de la quatrième composition de variables comme la plus performante. En revanche, l'utilisation des textures et des écarts-types ainsi que celle des bandes et indices de végétation soumis à des erreurs de correction atmosphérique n'améliore pas les classifications et peuvent même les détériorer. Enfin, la création d'un résultat consolidé par vote majoritaire de différentes combinaisons confirme les résultats issus des classifications. Néanmoins, la détermination d'une méthodologie de classification nécessite des développements supplémentaires, soutenus par une série de recommandations, afin d'en obtenir des similaires à ceux de l'interprétation de la clé visuelle. Ceci répond à la seconde question de recherche.

Par la richesse du spectre électromagnétique perçu par le capteur MSI, les plateformes S2 ont ouvert la voie à de nouvelles possibilités d'étude des forêts tropicales par télédétection. Les applications cartographiques sont nombreuses et les connaissances qui en découleront pourront inspirer des politiques de gestion et d'exploitation toujours plus durables.

RÉFÉRENCES

ANDERSON, J. R. (1976). *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*. U.S. Government Printing Office.

ASNER, G. P. & MASCARO, J. (2014). Mapping Tropical Forest Carbon: Calibrating Plot Estimates to a Simple LiDAR Metric. *Remote Sensing of Environment, 140*, 614-624. doi :10.1016/j.rse.2013.09.023

AUBRÉVILLE, A. (1957). Accord à Yangambi sur la nomenclature des types africains de végétation. *Bois et Forêts des Tropiques, 51*.

BREIMAN, L. (2001). Random forests. *Machine learning, 45*(1), 5-32.

CHAN, J. C.-W., LAPORTE, N. & DEFRIES, R. S. (2003). Texture Classification of Logged Forests in Tropical Africa Using Machine-Learning Algorithms. *International Journal of Remote Sensing, 24*(6), 1401-1407. doi :10.1080/0143116021000050538

CHEN, F., MING, C., LI, J., WANG, C. & CLAVERIE, M. (2018). A Comparison of Sentinel-2A and Sentinel-2B with Preliminary Results. *University of Waterloo*, 4.

CIB. (2007). Plan d'aménagement de l'unité forestière d'aménagement de Pokola (2007-2036).

COSMAN, P. (2010). Gray-Level Co-occurrence Matrices (GLCMs). Récupérée 2 janvier 2019, à partir de http://code.ucsd.edu/pcosman/glcm.pdf

DAÏNOU, K. & DOUCET, J.-L. (2013). *Évaluation de l'aménagement et l'état des populations des essences exploitées par la société Congolaise Industrielle des Bois (CIB–OLAM Congo). Unités Forestières d'Aménagement de Pokola, Kabo et Loundoungou-Toukoulaka.* Laboratoire de Foresterie des Régions tropicales et subtropicales-ULg Gx ABT.

DAUGHTRY, C. (2000). Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sensing of Environment, 74*(2), 229-239. doi :10.1016/S0034-4257(00)00113-9

de WASSEIGE, C., FLYNN, J., LOUPPE, D., HIOL HIOL, F. & MAYAUX, P. (Éd.). (2014). *The forests of the Congo basin - State of the forest 2013*. Belgium : WEYRICH Edition. Récupérée 17 octobre 2018, à partir de https://theredddesk.org/sites/default/files/resources/pdf/the_forests_of_the_congo_basin_state_of_the_forest_2013.pdf

DEFOURNY, P. (2017). Géomatique appliquée. Cours de master bioingénieur à l'UCLouvain (LBIRE 2102).

DI GREGORIO, A., HENRY, M., DONEGAN, E., FINEGOLD, Y., LATHAM, J., JONCKHEERE, I. & CUMANI, R. (2016). Land Cover Classification System - Classification Concepts Software Version 3. *Food and agriculture organisation of the United Nations*, 40.

DI GREGORIO, A. & JANSEN, L. J. (1997). A new concept for a land cover classification system. In *Proceedings of the Earth Observation and Environmental Information 1997 Conference. Alexandria, Egypt* (p. 13-16).

DI GREGORIO, A. & JANSEN, L. J. (2005). *Land cover classification system: Classification concepts and user manual: LCCS*. Food & Agriculture Org.

DRUSCH, M., DEL BELLO, U., CARLIER, S., COLIN, O., FERNANDEZ, V., GASCON, F., HOERSCH, B., ISOLA, C., LABERINTI, P., MARTIMORT, P., MEYGRET, A., SPOTO, F., SY, O., MARCHESE, F. & BARGELLINI, P. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment, 120*, 25-36. doi :10.1016/j.rse.2011.11.026

DUBAYAH, R. O., SHELDON, S. L., CLARK, D. B., HOFTON, M. A., BLAIR, J. B., HURTT, G. C. & CHAZDON, R. L. (2010). Estimation of Tropical Forest Height and Biomass Dynamics Using Lidar Remote Sensing at La Selva, Costa Rica: FOREST DYNAMICS USING LIDAR. *J. Geophys. Res. 115*(G2), n/a-n/a. doi :10.1029/2009JG000933

ERINJERY, J. J., SINGH, M. & KENT, R. (2018). Mapping and Assessment of Vegetation Types in the Tropical Rainforests of the Western Ghats Using Multispectral Sentinel-2 and SAR Sentinel-1 Satellite Imagery. *Remote Sensing of Environment, 216*, 345-354. doi :10.1016/j.rse.2018.07.006

ESA. (2017). Sentinel-2 Spectral Response Functions (S2-SRF). European Space Agency. Récupérée à partir de https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/document-library/-/asset_publisher/Wk0TKajiISaR/content/sentinel-2a-spectral-responses

EUROPEAN SPACE AGENCY. (2016). ESA CCI LAND COVER – S2 prototype Land Cover 20m map of Africa 2016. Récupérée 19 mai 2019, à partir de http://2016africalandcover20m.esrin.esa.int/

FORNI, E., DAINOU, K., GOURLET-FLEURY, S., BAUWENS, S., LOUBOTA, G. J., FAYOLLE, A., DOUH, C., BIWOLÉ, A., FEYCON, V. & BRACKE, C. (2015). *Rapport annuel 2014 du projet dynaffor*. Centre de coopération International pour la Recherche Agronomique et le Développement (CIRAD).

GISLASON, P. O., BENEDIKTSSON, J. A. & SVEINSSON, J. R. (2006). Random Forests for Land Cover Classification. *Pattern Recognition Letters, 27*(4), 294-300. doi :10.1016/j.patrec.2005.08.011

GLEYZES, M. A., PERRET, L. & KUBIK, P. (2012). Pleiades: System Architecture and Main Performances. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXIX-B1*, 537-542. doi :10.5194/isprsarchives-XXXIX-B1-537-2012

GOND, V., FAYOLLE, A., PENNEC, A., CORNU, G., MAYAUX, P., CAMBERLIN, P., DOUMENGE, C., FAUVET, N. & GOURLET-FLEURY, S. (2013). Vegetation Structure and Greenness in Central Africa from Modis Multi-Temporal Data. *Philosophical Transactions of the Royal Society B: Biological Sciences, 368*(1625), 20120309-20120309. doi :10.1098/rstb.2012.0309

GROGAN, J. & GALVÃO, J. (2006). Physiographic and Floristic Gradients across Topography in Transitional Seasonally Dry Evergreen Forests of Southeast Pará, Brazil. *Acta Amaz. 36*(4), 483-496. doi :10.1590/S0044-59672006000400009

HANSEN, M. C., POTAPOV, P. V., MOORE, R., HANCHER, M., TURUBANOVA, S. A., TYUKAVINA, A., THAU, D., STEHMAN, S. V., GOETZ, S. J., LOVELAND, T. R., KOMMAREDDY, A., EGOROV, A., CHINI, L., JUSTICE, C. O. & TOWNSHEND, J. R. G. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science, 342*(6160), 850-853. doi :10.1126/science.1244693

HARALICK, R. M., SHANMUGAM, K. & DINSTEIN, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-3*(6), 610-621. doi :10.1109/TSMC.1973.4309314

JUNGERS, Q. (2016). Construction d'un typologie forestière harmonisée en Afrique Centrale. CEOFAC3.

KHAZENIE, N. & RICHARDSON, K. (1991). Classification of cloud types based on spatial textural measures using NOAA-AVHRR data. In *Geoscience and Remote Sensing Symposium, 1991. IGARSS'91. Remote Sensing: Global Monitoring for Earth Management., International* (T. 3, p. 1701-1705). IEEE.

KHAZENIE, N. & RICHARDSON, K. (1992). *Comparison of Texture Analysis Techniques in Both Frequency and Spatial Domains for Cloud Feature Extraction.* Naval research lab Monterey ca.

KLEINSCHROTH, F., GARCIA, C. & GHAZOUL, J. (2018). Reconciling Certification and Intact Forest Landscape Conservation. *Ambio.* doi :10.1007/s13280-018-1063-6

LAURIN, G. V., CHEN, Q., LINDSELL, J. A., COOMES, D. A., FRATE, F. D., GUERRIERO, L., PIROTTI, F. & VALENTINI, R. (2014). Above Ground Biomass Estimation in an African Tropical Forest with Lidar and Hyperspectral Data. *ISPRS Journal of Photogrammetry and Remote Sensing, 89*, 49-58. doi :10.1016/j.isprsjprs.2014.01.001

LAURIN, G. V., PULETTI, N., HAWTHORNE, W., LIESENBERG, V., CORONA, P., PAPALE, D., CHEN, Q. & VALENTINI, R. (2016). Discrimination of Tropical Forest Types, Dominant Species, and Mapping of Functional Guilds by Hyperspectral and Simulated Multispectral Sentinel-2 Data. *Remote Sensing of Environment, 176*, 163-176. doi :10.1016/j.rse.2016.01.017

LOUIS, J., DEBAECKER, V., PFLUG, B., MAIN-KNORN, M., BIENIARZ, J., MUELLER-WILM, U., CADAU, E. & GASCON, F. (2016). Sentinel-2 Sen2cor:L2A Processor for Users, 8.

MAYAUX, P., BARTHOLOMÉ, E., FRITZ, S. & BELWARD, A. (2004). A New Land-Cover Map of Africa for the Year 2000: New Land-Cover Map of Africa. *Journal of Biogeography, 31*(6), 861-877. doi :10.1111/j.1365-2699.2004.01073.x

MAYAUX, P., JANODET, E., BLAIR-MYERS, C. M. & LEGEAY-JANVIER, P. (1997). Vegetation map of Central Africa at 1: 5,000,000. *Publication Series D, EUR, 17322.*

MUELER-WILM, U., DEVIGNOT, O. & PESSIOT, L. (2019). Sen2Cor configuration and user manual. ESA. Récupérée 10 août 2019, à partir de http://step.esa.int/thirdparties/sen2cor/2.8.0/docs/S2-PDGS-MPC-L2A-SUM-V2.8.pdf

OLOFSSON, P., FOODY, G. M., STEHMAN, S. V. & WOODCOCK, C. E. (2013). Making Better Use of Accuracy Data in Land Change Studies: Estimating Accuracy and Area and Quantifying Uncertainty Using Stratified Estimation. *Remote Sensing of Environment, 129*, 122-131. doi :10.1016/j.rse.2012.10.031

OLSON, D. M., DINERSTEIN, E., WIKRAMANAYAKE, E. D., BURGESS, N. D., POWELL, G. V. N., UNDERWOOD, E. C., D'AMICO, J. A., ITOUA, I., STRAND, H. E., MORRISON, J. C., LOUCKS, C. J., ALLNUTT, T. F., RICKETTS, T. H., KURA, Y., LAMOREUX, J. F., WETTENGEL, W. W., HEDAO, P. & KASSEM, K. R. (2001). Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience, 51*(11), 933. doi :10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2

PASCAL, J.-P. (2003). Notions sur les structure et dynamique des forêts tropicales humides. *Rev. For. Fr.* 118. doi :10.4267/2042/5765

PÉLISSIER, R. (1997). Heterogeneite spatiale et dynamique d'une foret dense humide dans les Ghats occidentaux de l'Indie. *nstitut français de Pondichéry - Publications du Département d'Écologie, 37*.

PHELPS, J., WEBB, E. L. & AGRAWAL, A. (2010). Does REDD+ Threaten to Recentralize Forest Governance? *Science, 328*(5976), 312-313. doi :10.1126/science.1187774

PONTIUS, R. G. & MILLONES, M. (2011). Death to Kappa: Birth of Quantity Disagreement and Allocation Disagreement for Accuracy Assessment. *International Journal of Remote Sensing, 32*(15), 4407-4429. doi :10.1080/01431161.2011.552923

RODRIGUEZ-GALIANO, V., GHIMIRE, B., ROGAN, J., CHICA-OLMO, M. & RIGOL-SANCHEZ, J. (2012). An Assessment of the Effectiveness of a Random Forest Classifier for Land-Cover Classification. *ISPRS Journal of Photogrammetry and Remote Sensing, 67*, 93-104. doi :10.1016/j.isprsjprs.2011.11.002

ROSENQVIST, A., SHIMADA, M., CHAPMAN, B., FREEMAN, A., DE GRANDI, G., SAATCHI, S. & RAUSTE, Y. (2000). The Global Rain Forest Mapping Project - A Review. *International Journal of Remote Sensing, 21*(6-7), 1375-1387. doi :10.1080/014311600210227

SANTORO, M., KIRCHES, G., WEVERS, J., BOETTCHER, M., BROCKMANN, C., LAMARCHE, C., BONTEMPS, S., MOREAU, I. & DEFOURNY, P. (2017). Land Cover CCI - Product User Guide Version 2.0. European Space Agency. Récupérée 7 août 2019, à partir de http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf

SEN2-AGRI_CONSORTIUM. (2018). Sentinel-2 for Agriculture. Récupérée 25 mai 2019, à partir de http://www.esa-sen2agri.org/

SOCIÉTÉ FRANÇAISE DE PHOTOGRAMMÉTRIE ET TÉLÉDÉTECTION. (2013). Principe de l'imagerie hyperspectrale – imagerie hyperspectrale. Récupérée 1 mai 2019, à partir de https://www.sfpt.fr/hyperspectral/?page_id=168

TRANSON, J., d'ANDRIMONT, R., MAUGNARD, A. & DEFOURNY, P. (2018). Survey of Hyperspectral Earth Observation Applications from Space in the Sentinel-2 Context. *Remote Sensing, 10*(3), 157. doi :10.3390/rs10020157

TRIMBLE, T. (2018). ECognition Developer 9.4 Reference Book. *Trimble Germany GmbH: Munich, Germany*, 319-328.

TUCKER, C. J., TOWNSHEND, J. R. & GOFF, T. E. (1985). African Land-Cover Classification Using Satellite Data. *Science*, *227*(4685), 369-375. doi :10.1126/science.227.4685.369

UNITED NATIONS ENVIRONMENT. (2018). Definitions. Récupérée 15 octobre 2018, à partir de https://www.cbd.int/forest/definitions.shtml

VIVIEN, J. J. & FAURE, J.-J. (1985). Arbres des forêts denses d'Afrique centrale. *Agence de Coopération Culturelle et Technique*, 565.

WELCH, R. M., KUO, K.-S. & SENGUPTA, S. K. (1990). Cloud and surface textural features in polar regions. *IEEE*, *28*(4), 520-528.

WHITE, F. (1983). The vegetation of Africa: A descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa by F White. *Natural Resources Research Report XX, UNESCO, Paris, France*.

WIRTH, R., WEBER, B. & RYEL, R. J. (2001). Spatial and Temporal Variability of Canopy Structure in a Tropical Moist Forest. *Acta Oecologica*, *22*(5-6), 235-244. doi :10.1016/S1146-609X(01)01123-7

YUAN, F., SAWAYA, K. E., LOEFFELHOLZ, B. C. & BAUER, M. E. (2005). Land Cover Classification and Change Analysis of the Twin Cities (Minnesota) Metropolitan Area by Multitemporal Landsat Remote Sensing. *Remote Sensing of Environment*, *98*(2-3), 317-328. doi :10.1016/j.rse.2005.08.006

ANNEXE A

IMAGES PLÉIADES UTILISÉES POUR LA MISSION DE TERRAIN

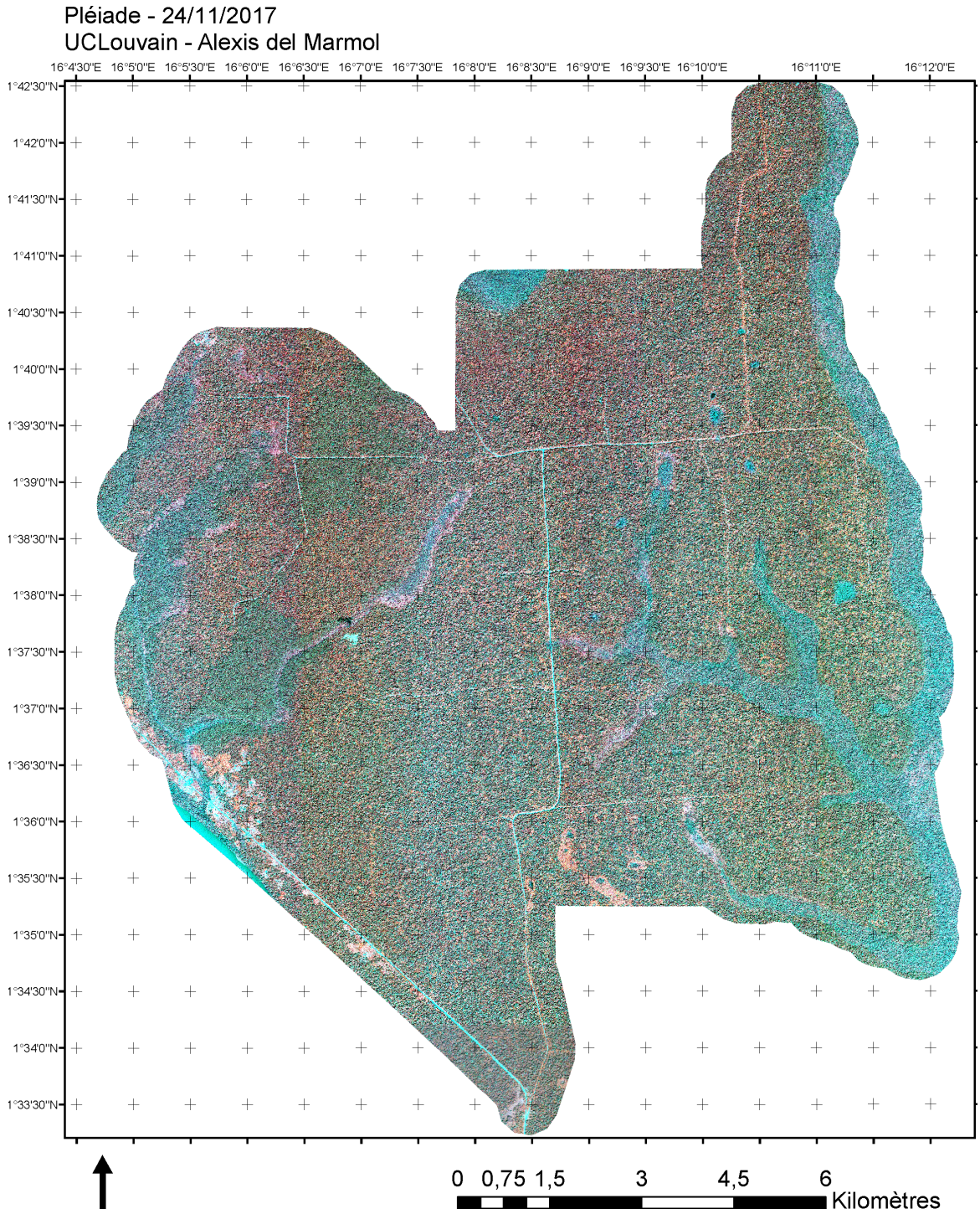


FIGURE A.1 – *Image pléiade en fausse couleur de l'assiette annuelle de coupe de Pokola 2017*

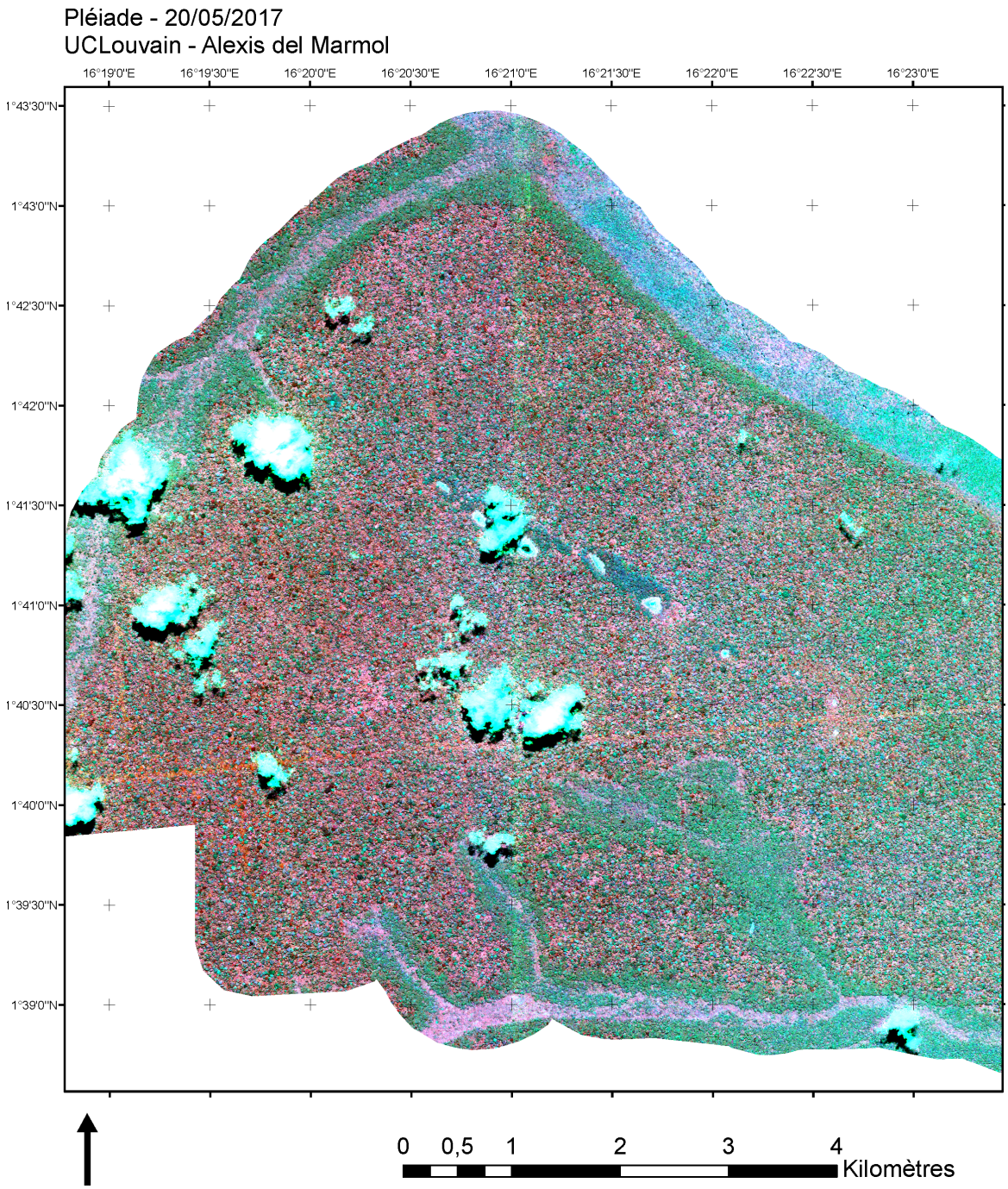


FIGURE A.2 – *Image pléiade en fausse couleur de l'assiette annuelle de coupe de Pokola 2018*

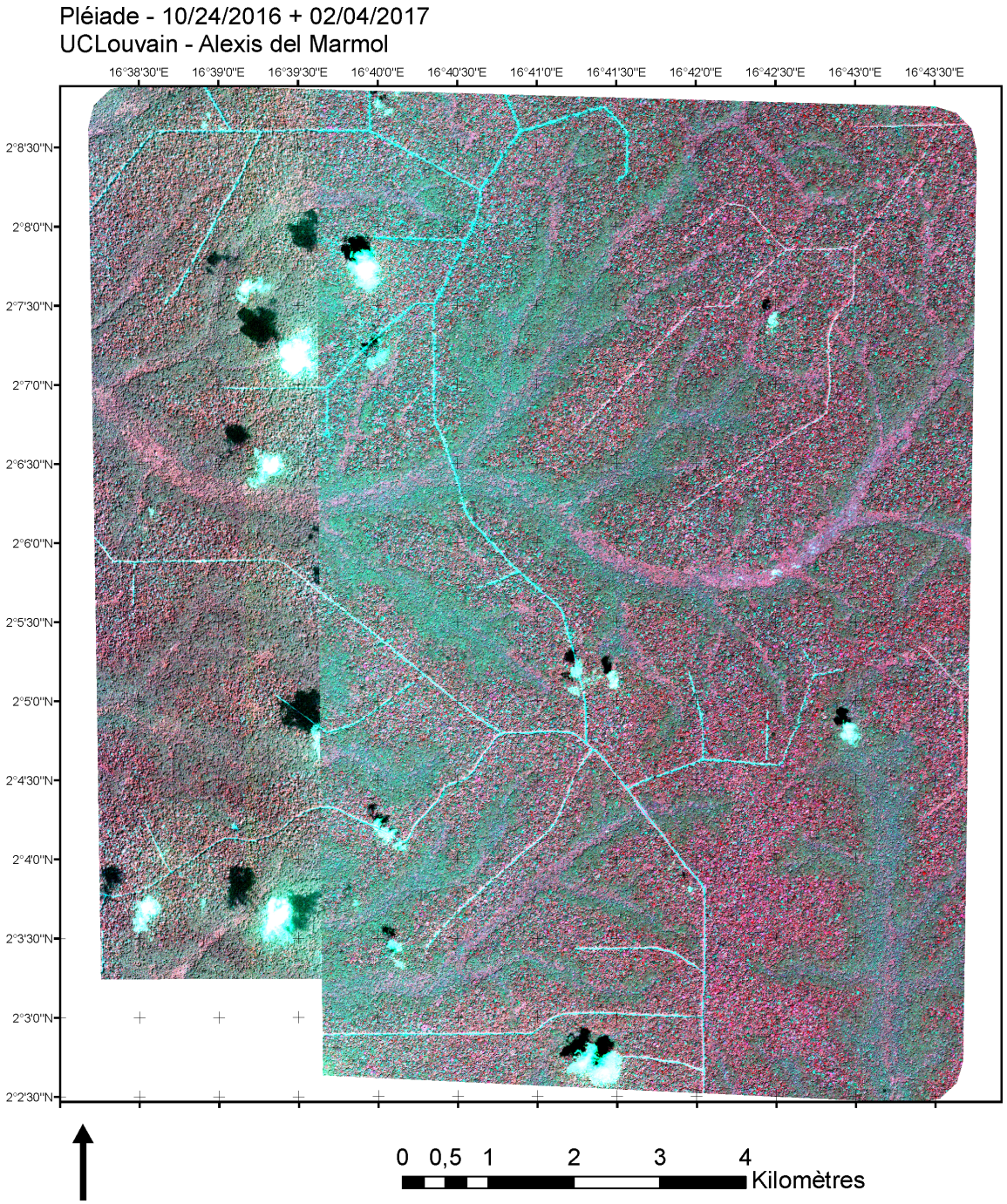


FIGURE A.3 – *Image pléiade en fausse couleur de l'assiette annuelle de coupe de Pokola 2017*

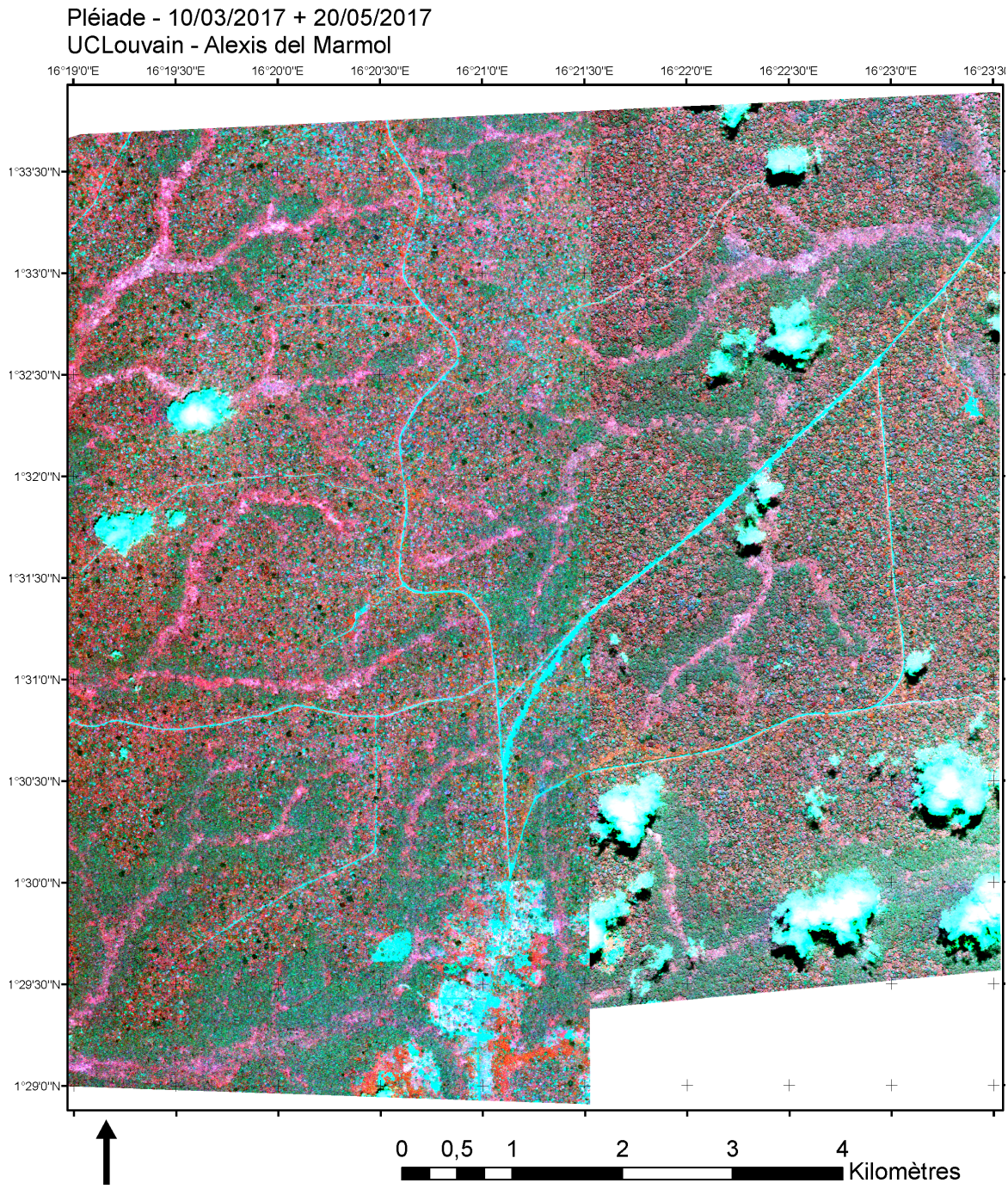


FIGURE A.4 – *Image pléiade en fausse couleur d'une zone intéressante au nord de la zone de développement rurale de Pokola*

ANNEXE B

IMAGES S2 UTILISÉES DURANT LA MISSION DE TERRAIN

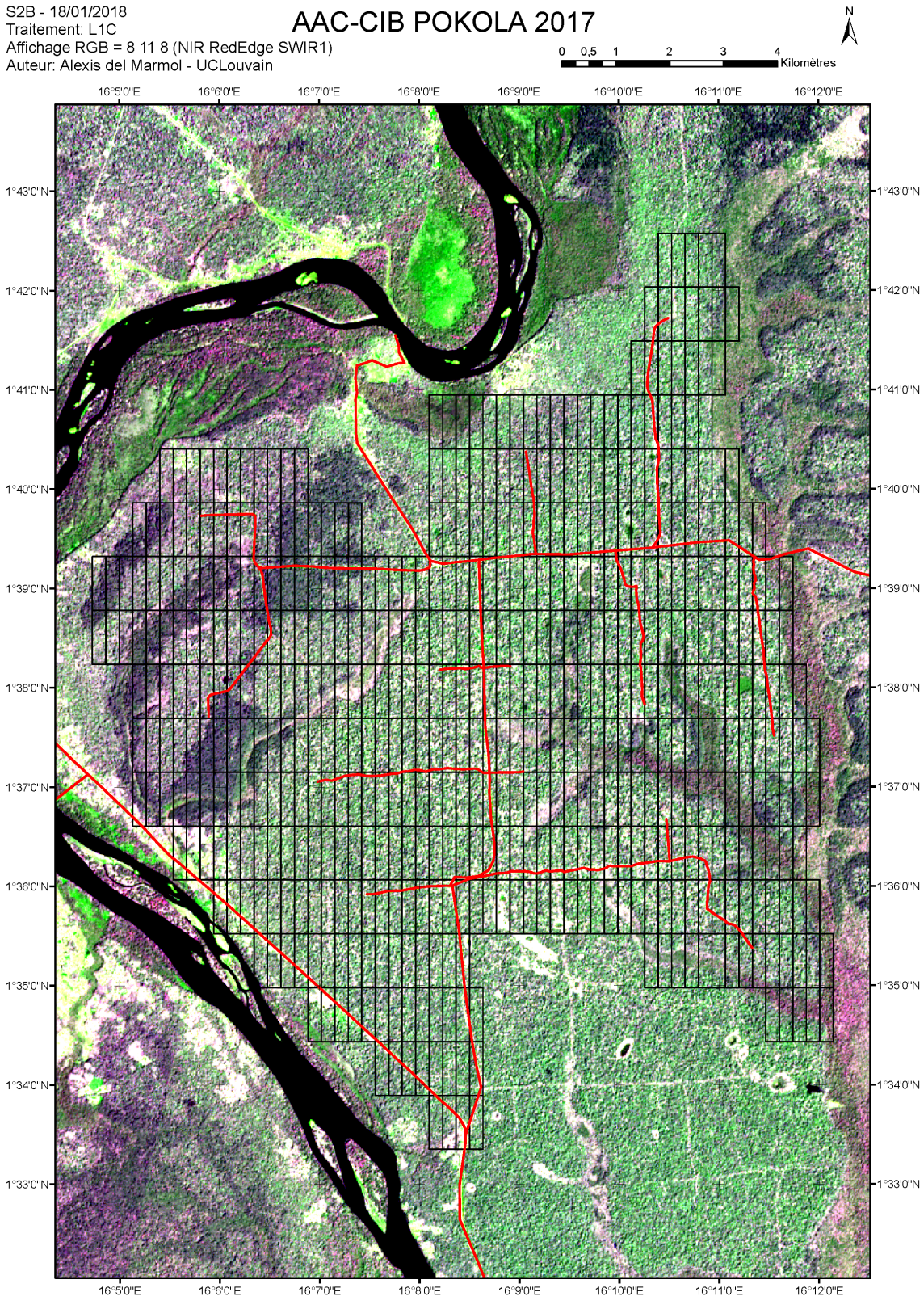


FIGURE B.1 – *Image S2 au dessus de l'assiette annuelle de coupe Pokola 2017. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge. Le quadrillage noir représente les parcelles d'exploitation.*

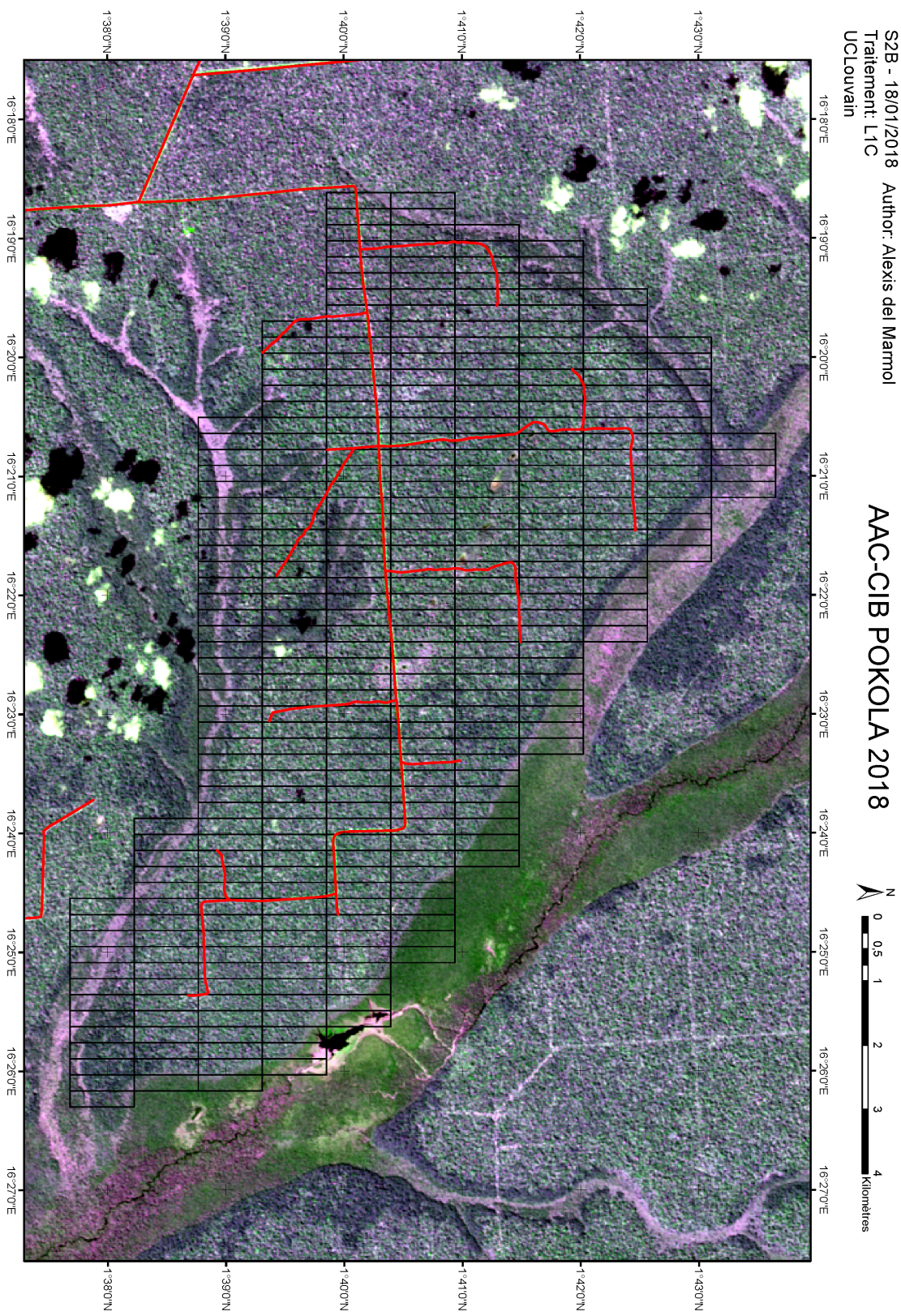


FIGURE B.2 – *Image S2 au dessus de l'assiette annuelle de coupe Pokola 2018. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge. Le quadrillage noir représente les parcelles d'exploitation.*

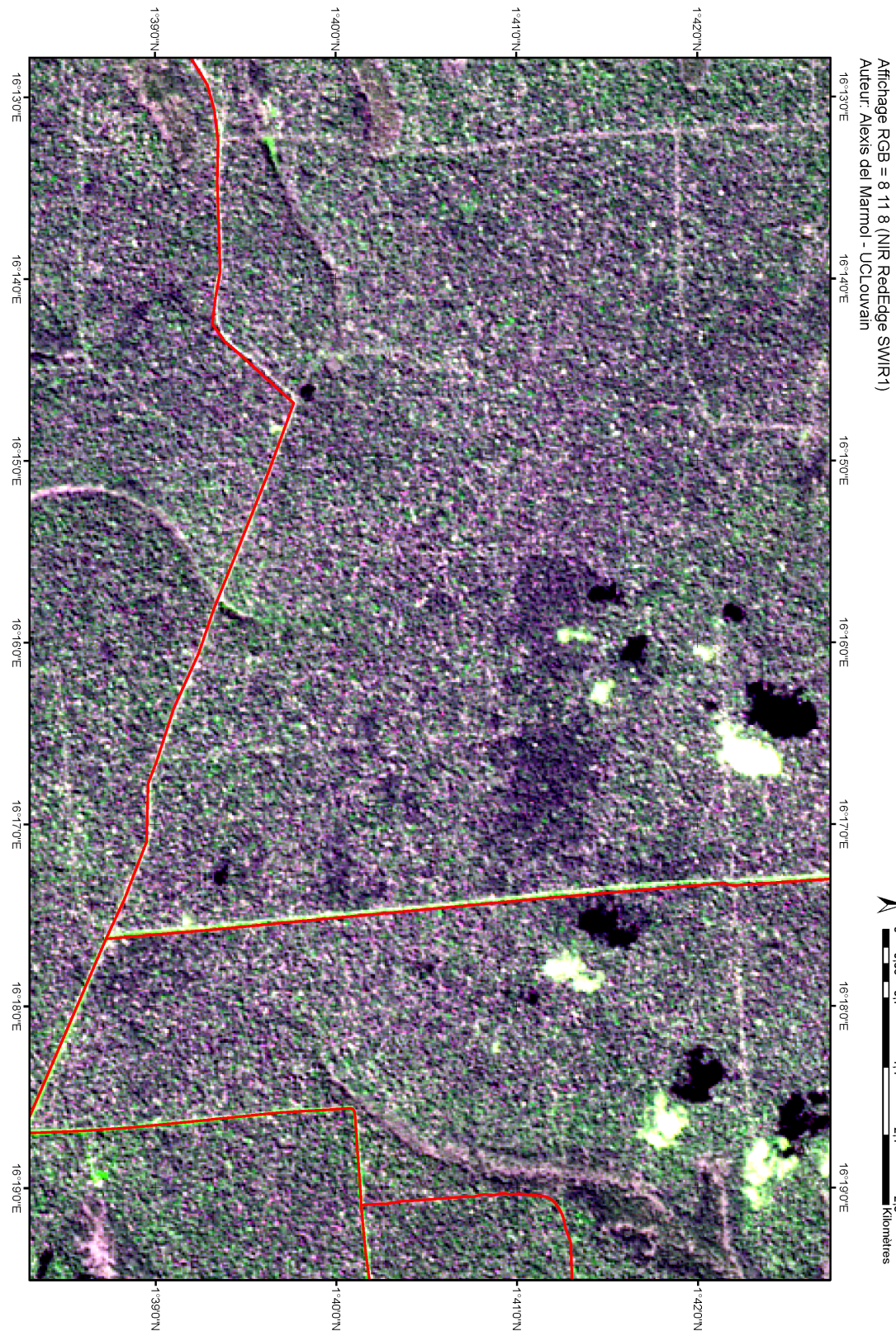


FIGURE B.3 – *Image S2 au dessus de l'assiette annuelle de coupe Pokola 1980. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge.*

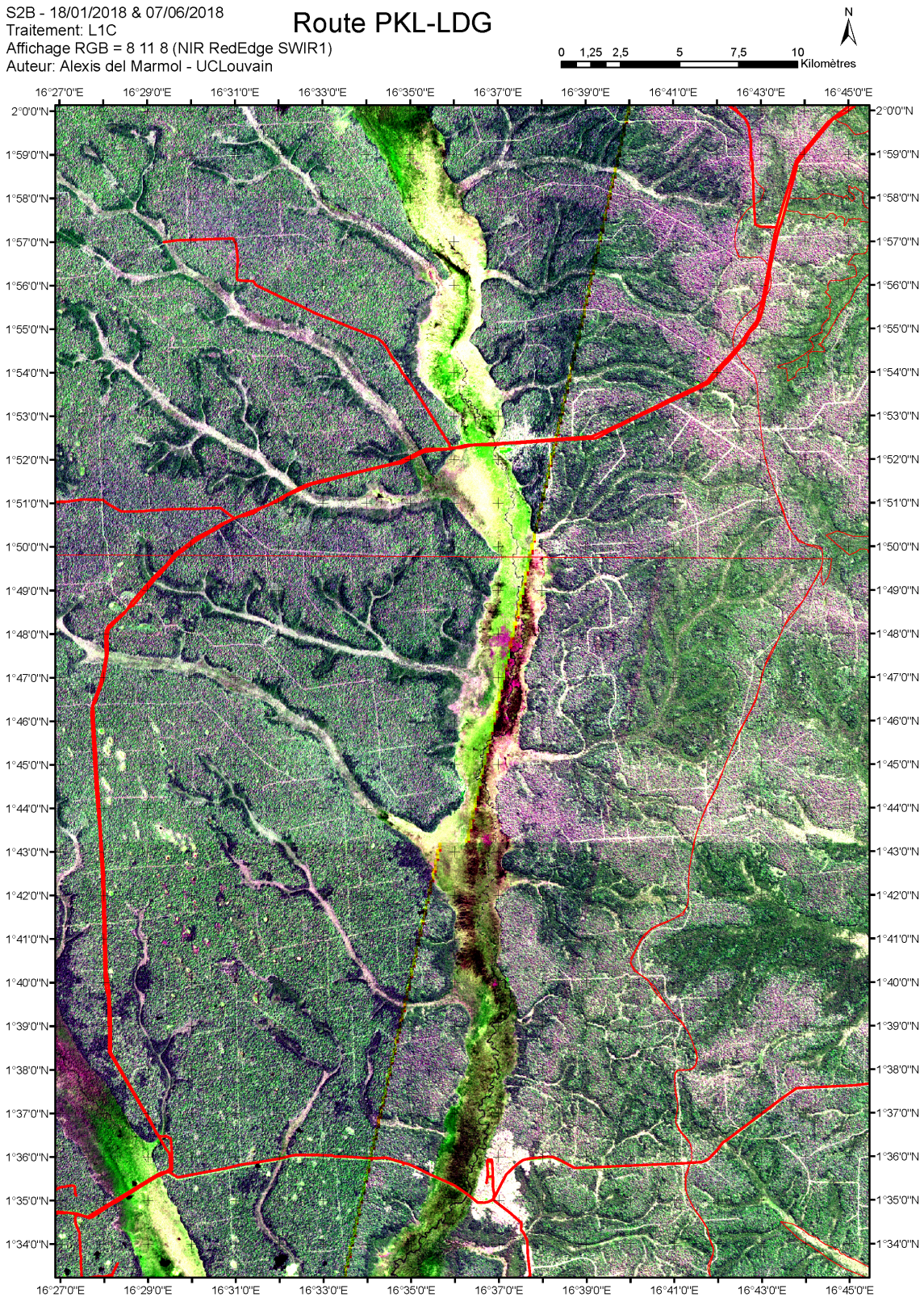


FIGURE B.4 – *Image S2 au dessus de la route nationale entre Pokola et Loundoungou. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge.*

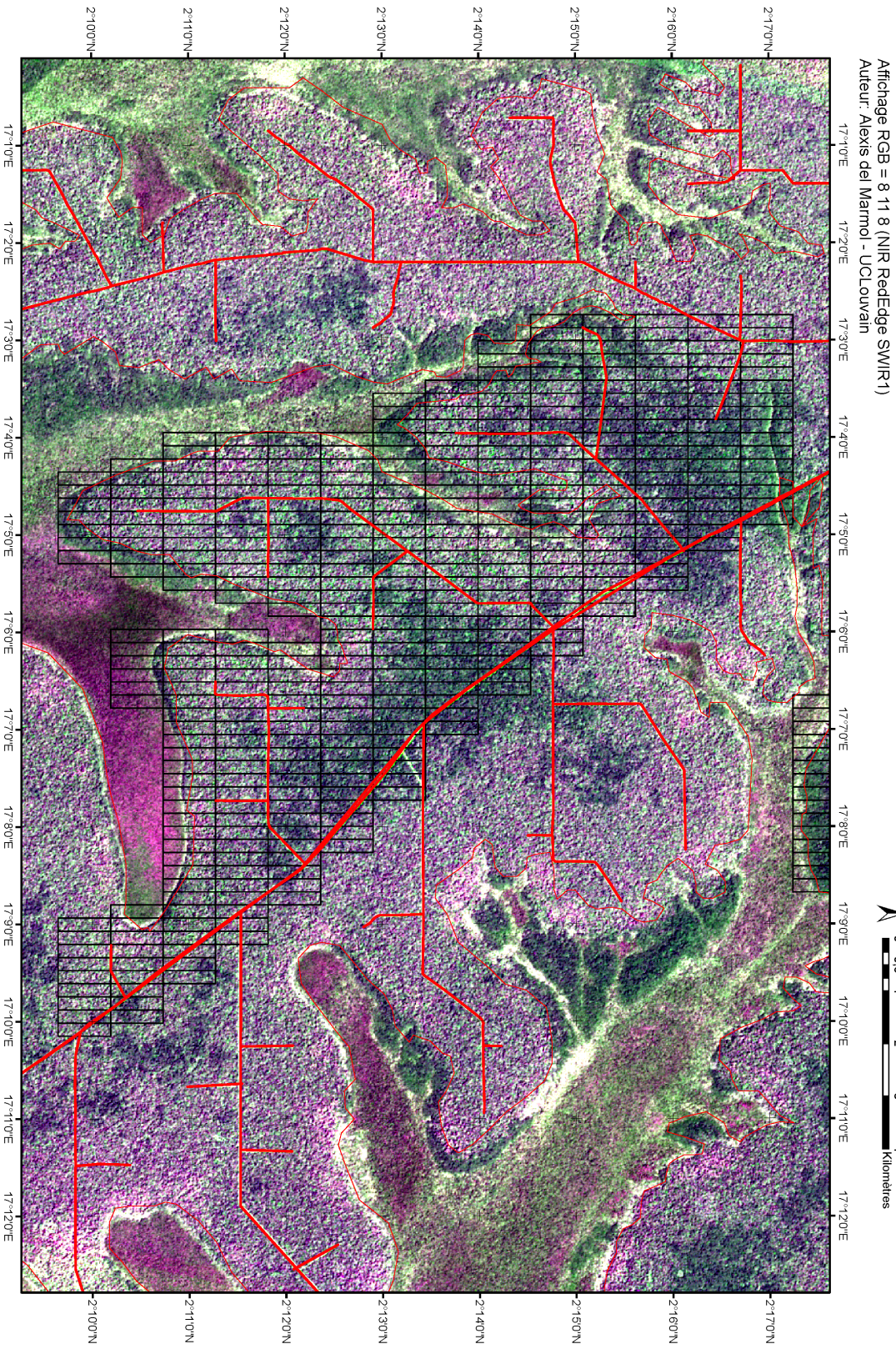


FIGURE B.5 – *Image S2 au dessus de l'assiette annuelle de coupe Loundoungou 2017. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge. Le quadrillage noir représente les parcelles d'exploitation.*

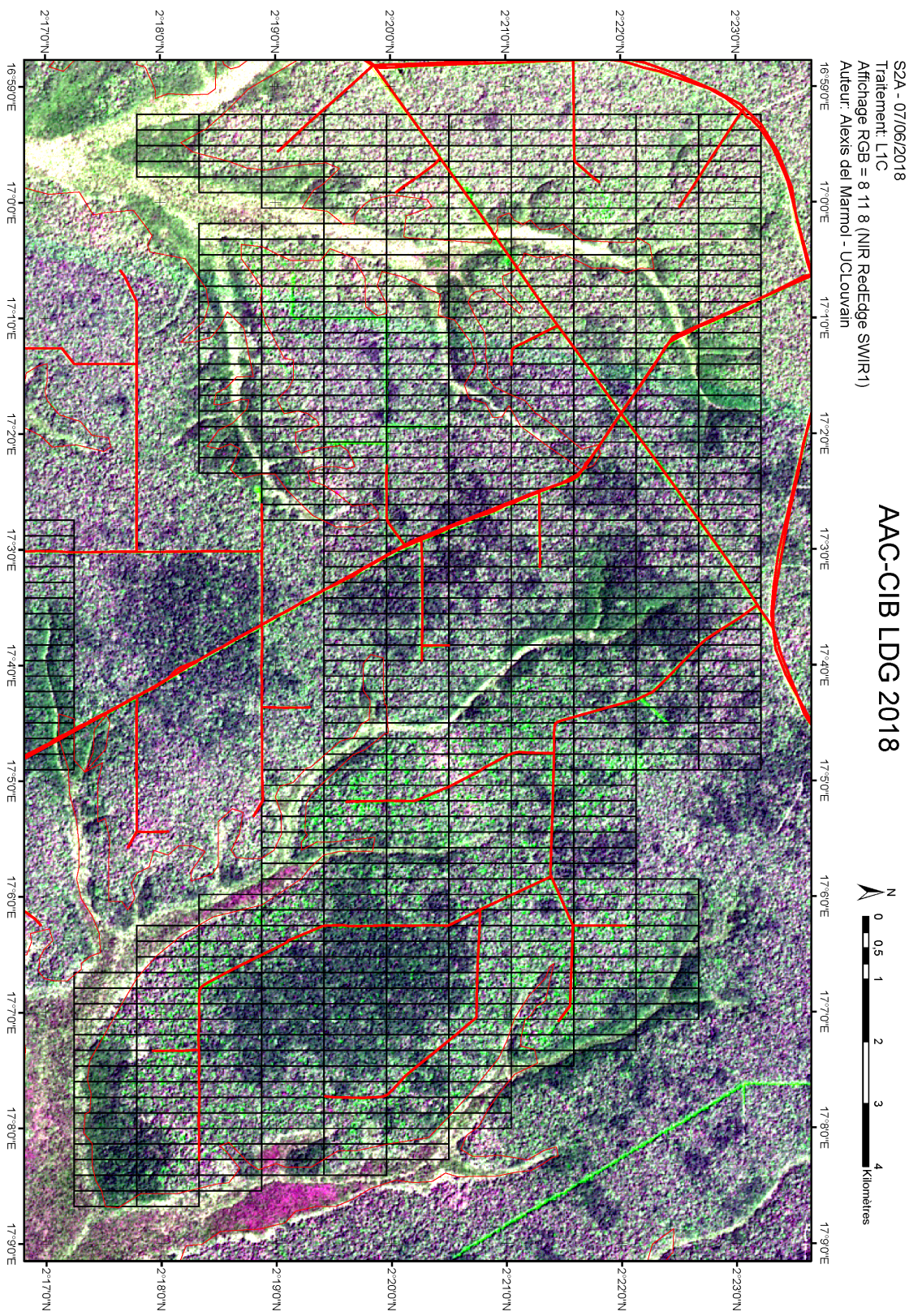


FIGURE B.6 – *Image S2 au dessus de l'assiette annuelle de coupe Loundoungou 2018. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge. Le quadrillage noir représente les parcelles d'exploitation.*

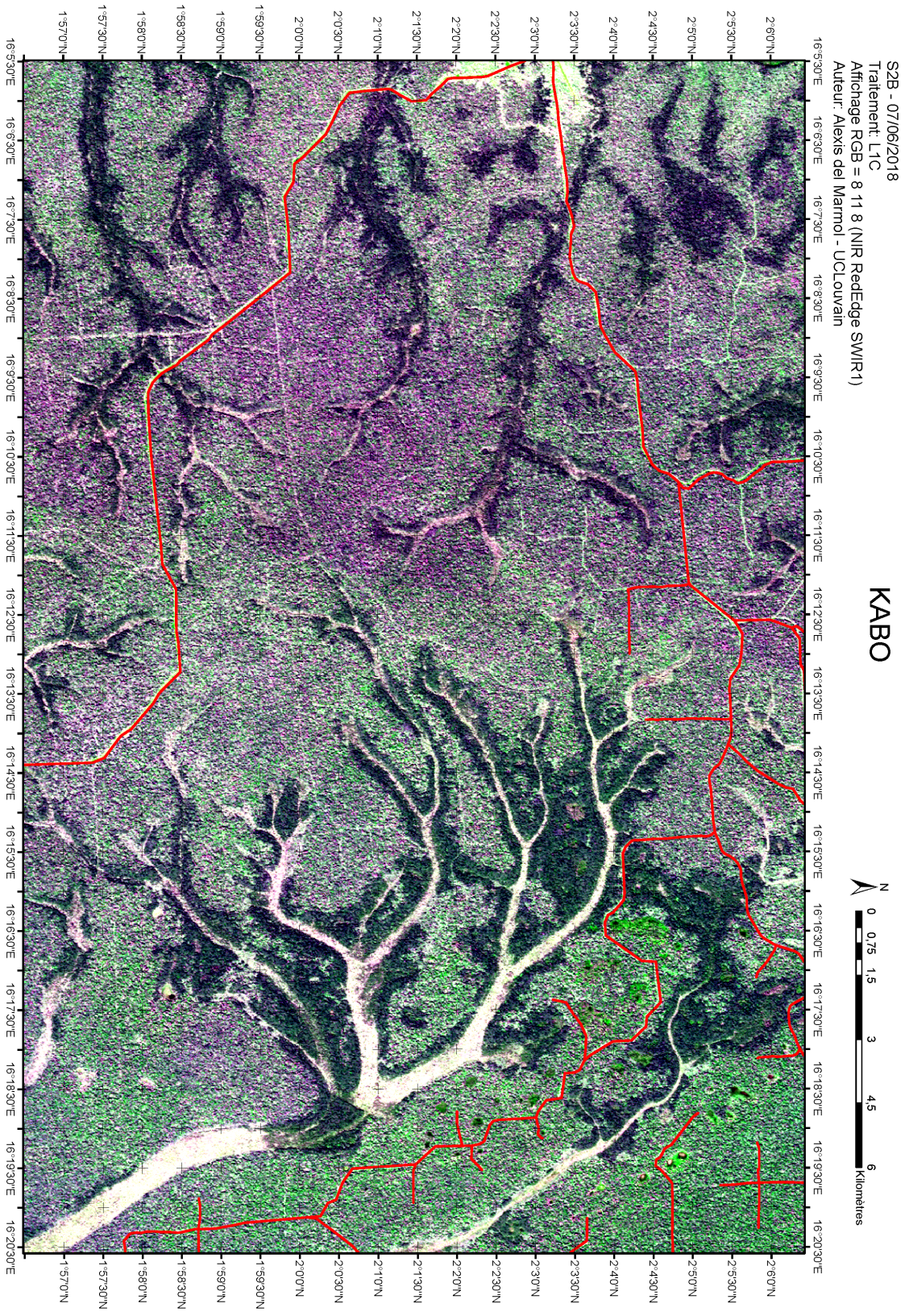


FIGURE B.7 – *Image S2 proche du village de Kabo (à gauche). Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge.*

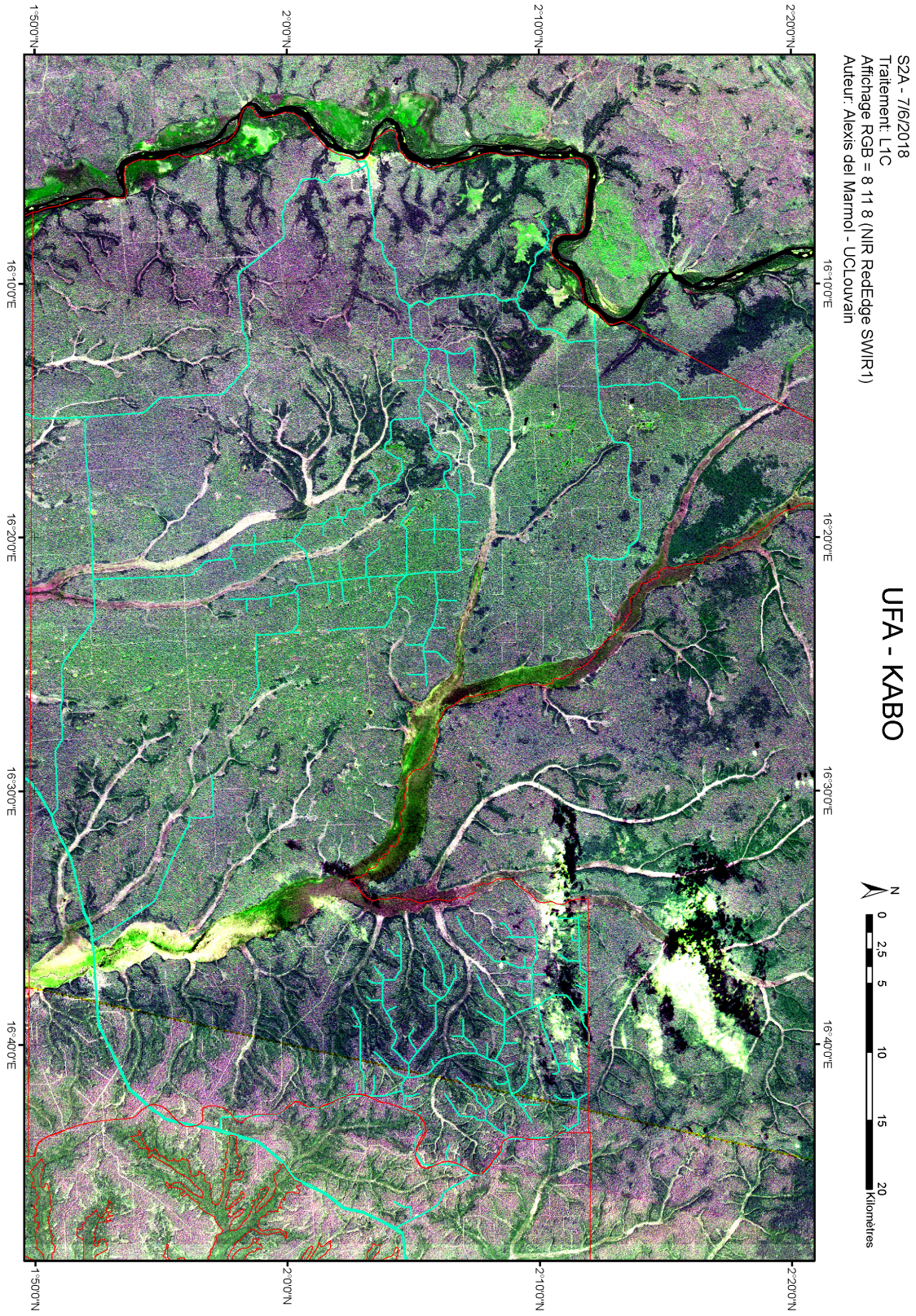


FIGURE B.8 – *Image S2 au dessus de l'unité forestière d'aménagement de Kabo. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en bleu et la limite des UFA est en rouge.*

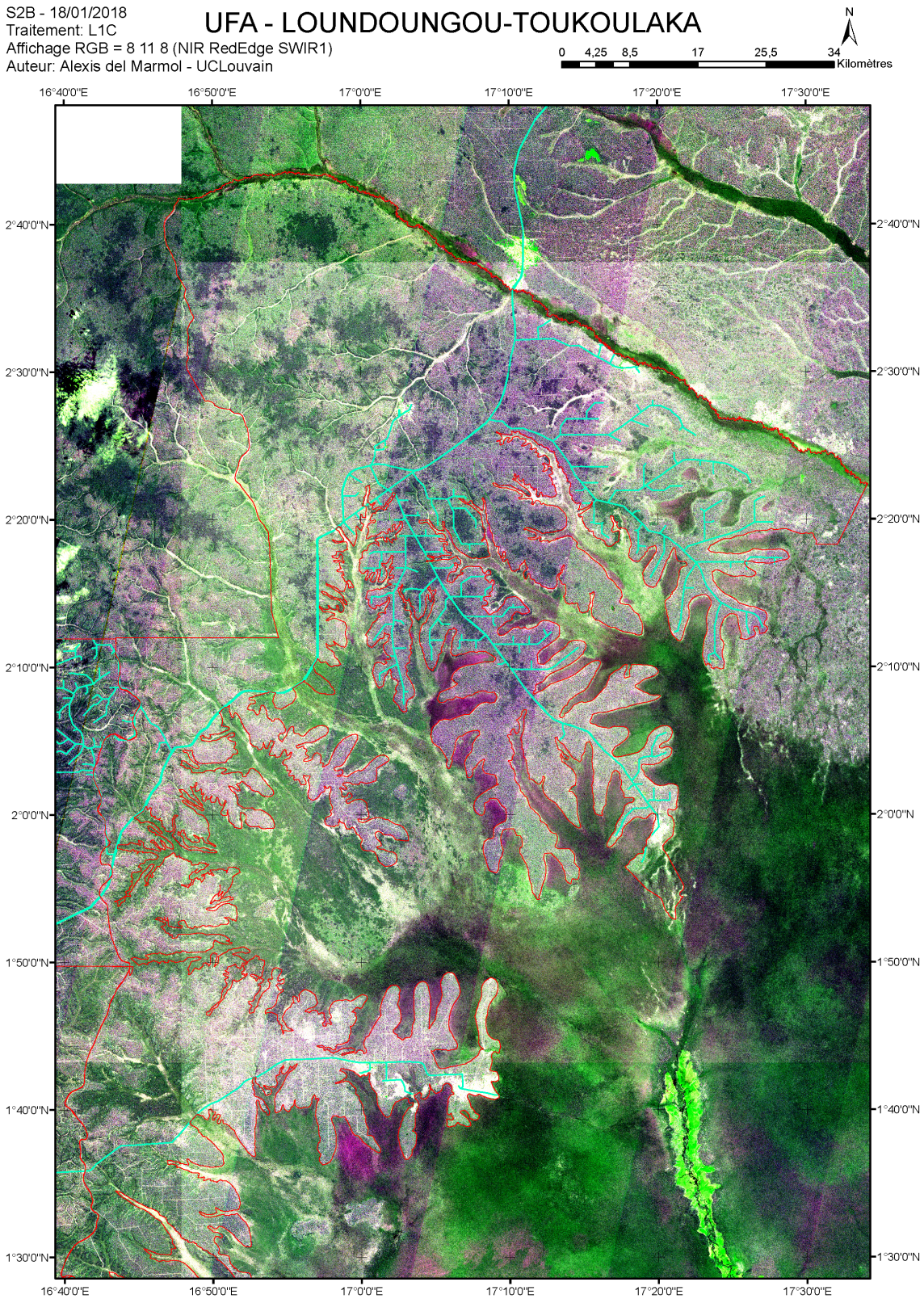


FIGURE B.9 – *Image S2 au dessus de l'unité forestière d'aménagement de Loundoungou. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en bleu et la limite des UFA est en rouge.*

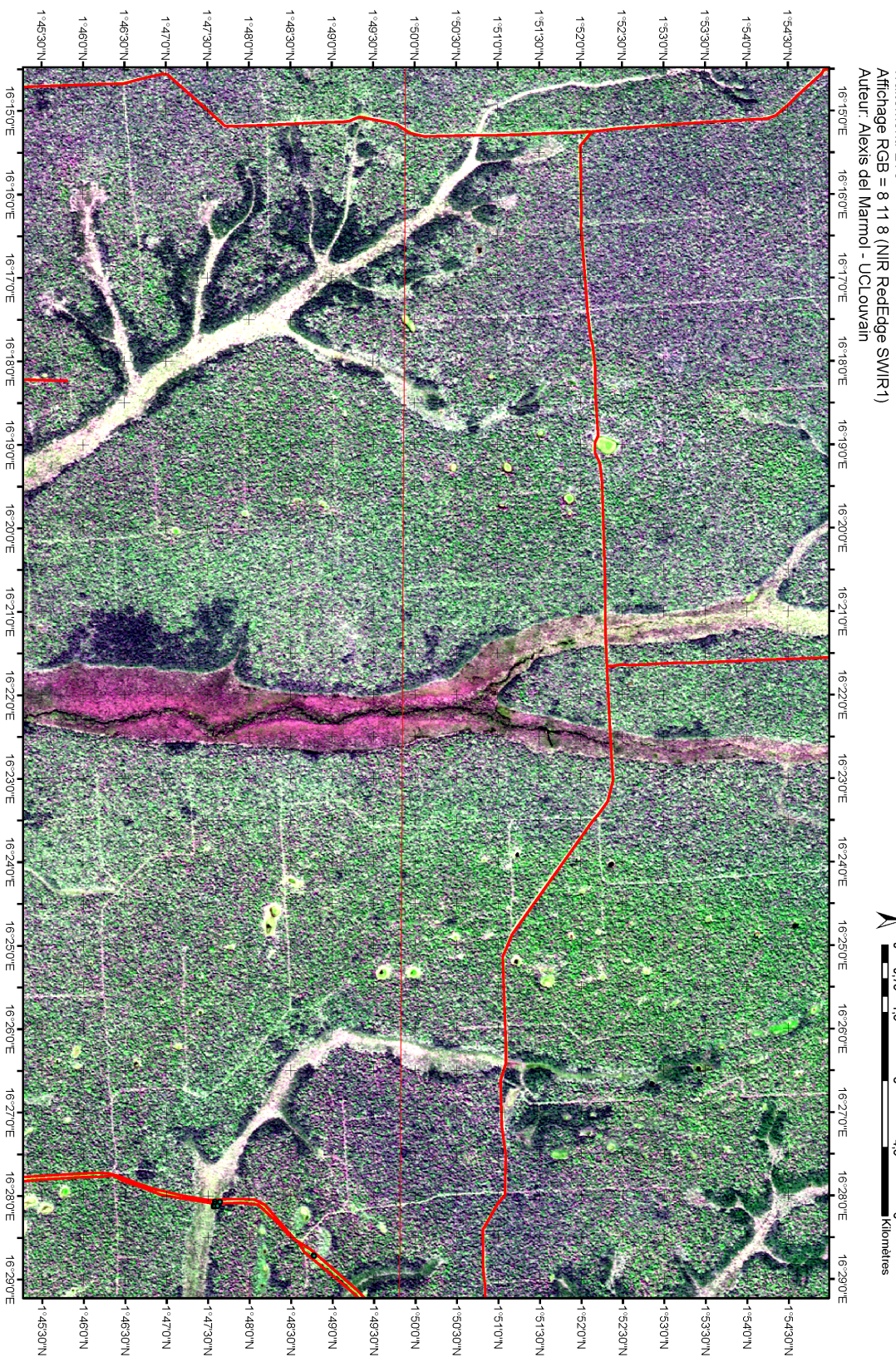


FIGURE B.10 – *Image S2 au dessus de la route transversale qui relie Kabo à Mokobo. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge.*

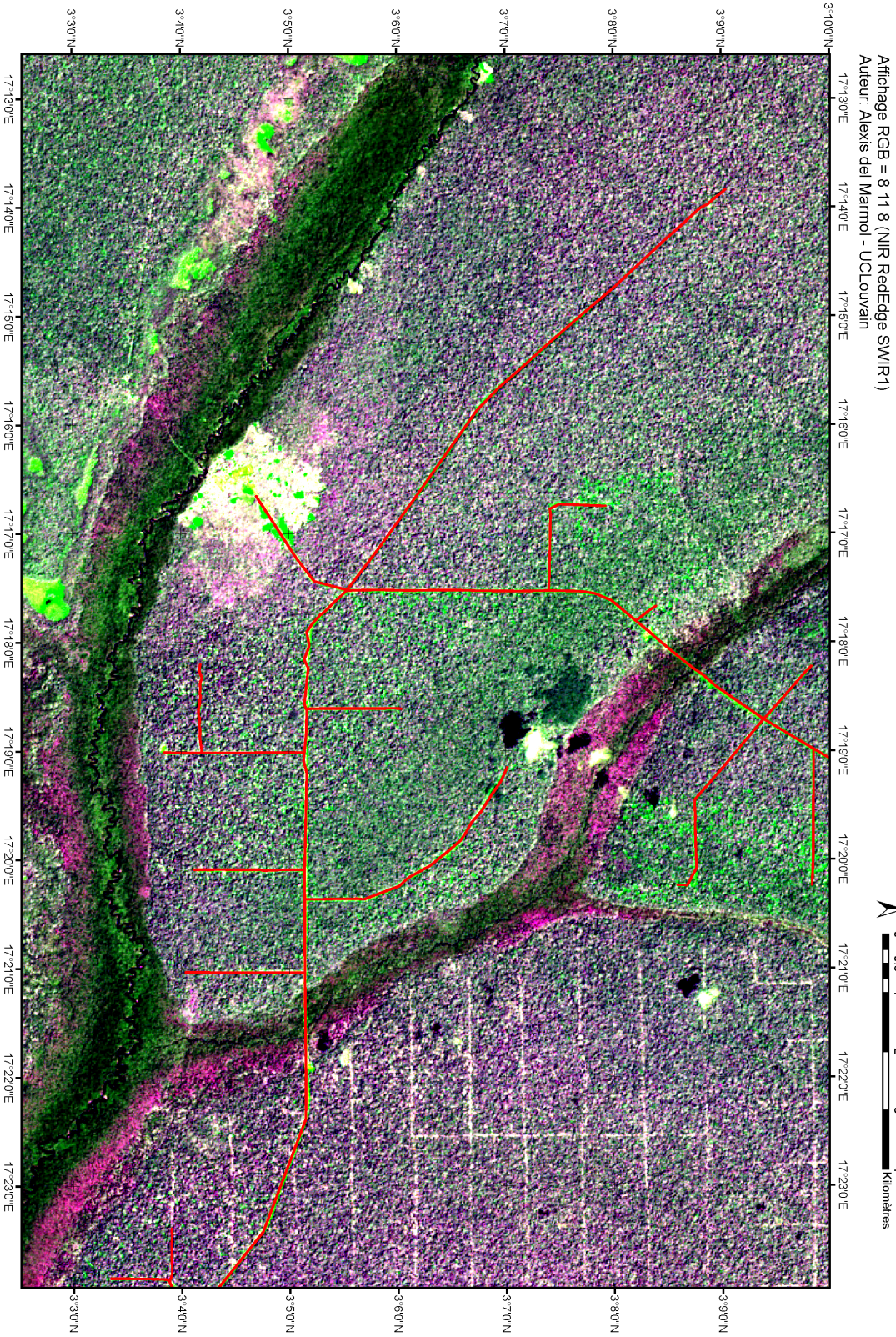


FIGURE B.11 – *Image S2 au dessus du camp forestier de Berandjokou. Attention, l'affichage correct est RGB = B06 B12 B08 [RedEdge SWIR2 NIR]. Les routes principales et secondaires sont indiquées en rouge.*

ANNEXE C

MATRICES DE CONFUSION DES DIFFÉRENTES COMBINAISONS

TABLEAU C.1 – *Matrice de confusion de la C1 par pixels*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	43	0	4	0	9	0	1	19	11	0	87	0,49
FL	0	49	1	0	0	0	27	12	0	0	89	0,55
FJ	1	11	11	0	11	1	48	9	0	0	92	0,12
ZH	5	0	5	52	0	13	1	14	8	0	98	0,53
FM	46	0	2	0	19	0	3	1	10	0	81	0,23
R	0	9	16	3	0	41	1	7	0	0	77	0,53
FA	0	18	2	0	10	0	55	8	1	0	94	0,59
M	3	8	1	2	0	49	0	30	0	0	93	0,32
HF	5	0	0	1	0	0	0	0	94	0	100	0,94
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	103	95	42	58	49	104	136	100	124	100		
UA	0,42	0,52	0,26	0,90	0,39	0,39	0,40	0,30	0,76	1,00	OA :	0,54
F-score	0,45	0,53	0,16	0,67	0,29	0,45	0,48	0,31	0,84	1,00	κ :	0,49

TABLEAU C.2 – *Matrice de confusion de la C1 par classe majoritaire*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	43	0	0	0	0	0	5	39	0	0	87	0,49
FL	0	32	0	0	0	0	22	35	0	0	89	0,36
FJ	0	0	8	0	0	0	84	0	0	0	92	0,09
FA	9	0	1	77	1	6	4	0	0	0	98	0,79
FM	58	0	0	0	20	0	3	0	0	0	81	0,25
M	0	10	0	0	0	44	1	22	0	0	77	0,57
R	0	0	0	0	0	0	94	0	0	0	94	1,00
ZH	0	5	0	0	0	52	0	36	0	0	93	0,39
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	5	95	100	0,95
∑	110	47	9	77	21	102	213	132	105	95		
UA	0,39	0,68	0,89	1,00	0,95	0,43	0,44	0,27	0,95	1,00	OA :	0,60
F-score	0,44	0,47	0,16	0,88	0,39	0,49	0,61	0,32	0,98	0,97	κ :	0,56

TABLEAU C.3 – *Matrice de confusion de la C1 par objets*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	42	0	0	0	8	0	1	26	10	0	87	0,48
FL	0	53	2	0	0	0	23	11	0	0	89	0,60
FJ	1	11	6	0	13	2	47	12	0	0	92	0,07
ZH	6	0	0	34	0	17	0	33	8	0	98	0,35
FM	46	0	0	0	25	0	2	3	5	0	81	0,31
R	0	8	7	1	0	45	1	14	1	0	77	0,58
FA	0	20	0	0	10	0	52	11	1	0	94	0,55
M	2	7	0	2	0	50	0	32	0	0	93	0,34
HF	0	0	0	2	0	0	0	0	98	0	100	0,98
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	97	99	15	39	56	114	126	142	123	100		
UA	0,43	0,54	0,40	0,87	0,45	0,39	0,41	0,23	0,80	1,00	OA :	0,53
F-score	0,46	0,56	0,11	0,50	0,36	0,47	0,47	0,27	0,88	1,00	κ :	0,48

TABLEAU C.4 – *Matrice de confusion de la C2 par pixels*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	45	0	21	0	8	0	0	3	10	0	87	0,52
FL	0	53	6	0	0	1	13	16	0	0	89	0,60
FJ	1	11	32	0	11	1	32	4	0	0	92	0,35
ZH	6	0	26	30	0	17	0	9	10	0	98	0,31
FM	58	0	4	0	7	0	2	0	10	0	81	0,09
R	0	6	21	2	0	48	0	0	0	0	77	0,62
FA	0	22	20	0	8	0	34	9	1	0	94	0,36
M	2	6	9	0	0	54	0	22	0	0	93	0,24
HF	3	0	0	3	0	0	0	0	94	0	100	0,94
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	115	98	139	35	34	121	81	63	125	100		
UA	0,39	0,54	0,23	0,86	0,21	0,40	0,42	0,35	0,75	1,00	OA :	0,51
F-score	0,45	0,57	0,28	0,45	0,12	0,48	0,39	0,28	0,84	1,00	κ :	0,46

TABLEAU C.5 – *Matrice de confusion de la C2 par classe majoritaire*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	50	0	0	0	0	0	0	0	87	0,43
FL	0	48	31	0	0	0	1	9	0	0	89	0,54
FJ	0	33	29	0	0	0	30	0	0	0	92	0,32
ZH	1	0	33	32	0	21	0	0	11	0	98	0,33
FM	72	0	3	0	6	0	0	0	0	0	81	0,07
R	0	0	1	0	0	76	0	0	0	0	77	0,99
FA	0	0	4	0	0	0	90	0	0	0	94	0,96
M	0	5	10	0	0	52	0	26	0	0	93	0,28
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	110	86	161	32	6	149	121	35	111	100		
UA	0,34	0,56	0,18	1,00	1,00	0,51	0,74	0,74	0,90	1,00	OA :	0,60
F-score	0,38	0,55	0,23	0,49	0,14	0,67	0,84	0,41	0,95	1,00	κ :	0,55

TABLEAU C.6 – *Matrice de confusion de la C2 par objets*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	0	0	45	0	5	0	0	0	87	0,43
FL	0	23	35	0	0	0	31	0	0	0	89	0,26
FJ	0	0	16	0	13	0	63	0	0	0	92	0,17
FA	0	0	2	83	5	0	3	0	5	0	98	0,85
FM	0	0	3	0	78	0	0	0	0	0	81	0,96
M	0	0	0	0	0	76	1	0	0	0	77	0,99
R	0	0	5	0	0	0	89	0	0	0	94	0,95
ZH	0	0	8	0	10	52	2	21	0	0	93	0,23
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	37	23	69	83	151	128	194	21	105	100		
UA	1,00	1,00	0,23	1,00	0,52	0,59	0,46	1,00	0,95	1,00	OA :	0,68
F-score	0,60	0,41	0,20	0,92	0,67	0,74	0,62	0,37	0,98	1,00	κ :	0,65

TABLEAU C.7 – *Matrice de confusion de la C3 par pixels*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	40	0	25	0	11	0	0	0	11	0	87	0,46
FL	0	53	5	0	0	1	20	10	0	0	89	0,60
FJ	2	11	20	0	11	2	40	6	0	0	92	0,22
ZH	6	0	13	32	0	18	0	22	7	0	98	0,33
FM	31	0	4	0	33	0	2	0	11	0	81	0,41
R	0	6	17	1	0	47	0	6	0	0	77	0,61
FA	0	19	9	0	11	0	50	4	1	0	94	0,53
M	1	6	7	0	2	53	0	24	0	0	93	0,26
HF	2	0	0	1	0	0	0	0	97	0	100	0,97
EAU	0	0	0	1	0	0	0	0	0	100	101	0,99
∑	82	95	100	35	68	121	112	72	127	100		
UA	0,49	0,56	0,20	0,91	0,49	0,39	0,45	0,33	0,76	1,00	OA :	0,54
F-score	0,47	0,58	0,21	0,48	0,44	0,47	0,49	0,29	0,85	1,00	κ :	0,49

TABLEAU C.8 – *Matrice de confusion de la C3 par classe majoritaire*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	50	0	0	0	0	0	0	0	87	0,43
FL	0	32	0	0	0	0	31	26	0	0	89	0,36
FJ	0	0	8	0	0	0	84	0	0	0	92	0,09
ZH	10	0	20	43	0	21	1	3	0	0	98	0,44
FM	14	0	0	0	64	0	3	0	0	0	81	0,79
R	0	0	1	0	0	76	0	0	0	0	77	0,99
FA	0	0	0	0	0	0	94	0	0	0	94	1,00
M	0	5	10	0	0	52	0	26	0	0	93	0,28
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	61	37	89	43	64	149	213	55	100	100		
UA	0,61	0,86	0,09	1,00	1,00	0,51	0,44	0,47	1,00	1,00	OA :	0,64
F-score	0,50	0,51	0,09	0,61	0,88	0,67	0,61	0,35	1,00	1,00	κ :	0,60

TABLEAU C.9 – *Matrice de confusion de la C3 par objets*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	5	0	45	0	0	0	0	0	87	0,43
FL	0	22	26	0	0	0	41	0	0	0	89	0,25
FJ	0	0	16	0	13	0	63	0	0	0	92	0,17
ZH	10	0	2	83	0	0	2	1	0	0	98	0,85
FM	58	0	3	0	20	0	0	0	0	0	81	0,25
R	0	0	0	0	0	76	1	0	0	0	77	0,99
FA	0	0	5	0	0	0	89	0	0	0	94	0,95
M	10	0	0	0	0	52	0	28	0	3	93	0,30
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	115	22	57	83	78	128	196	29	100	103		
UA	0,32	1,00	0,28	1,00	0,26	0,59	0,45	0,97	1,00	0,97	OA :	0,63
F-score	0,37	0,40	0,21	0,92	0,25	0,74	0,61	0,46	1,00	0,99	κ :	0,59

TABLEAU C.10 – *Matrice de confusion de la C4 par pixels*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	42	0	25	0	9	0	0	0	11	0	87	0,48
FL	0	53	5	0	0	1	13	17	0	0	89	0,60
FJ	2	11	23	0	12	2	36	6	0	0	92	0,25
ZH	4	0	29	36	0	14	0	7	8	0	98	0,37
FM	34	0	3	0	32	0	1	0	11	0	81	0,40
R	0	6	24	1	0	45	0	1	0	0	77	0,58
FA	0	22	11	0	11	0	39	10	1	0	94	0,41
M	2	6	10	0	0	53	0	22	0	0	93	0,24
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	84	98	130	37	64	115	89	63	131	100		
UA	0,50	0,54	0,18	0,97	0,50	0,39	0,44	0,35	0,76	1,00	OA :	0,54
F-score	0,49	0,57	0,21	0,53	0,44	0,47	0,43	0,28	0,87	1,00	κ :	0,49

TABLEAU C.11 – *Matrice de confusion de la C4 par classe majoritaire*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	50	0	0	0	0	0	0	0	87	0,43
FL	0	41	0	0	0	0	13	35	0	0	89	0,46
FJ	0	0	16	0	0	0	76	0	0	0	92	0,17
ZH	0	0	33	49	1	15	0	0	0	0	98	0,50
FM	14	0	3	0	64	0	0	0	0	0	81	0,79
R	0	0	1	0	0	76	0	0	0	0	77	0,99
FA	0	0	4	0	0	0	90	0	0	0	94	0,96
M	0	0	10	0	0	52	0	31	0	0	93	0,33
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	51	41	117	49	65	143	179	66	100	100		
UA	0,73	1,00	0,14	1,00	0,98	0,53	0,50	0,47	1,00	1,00	OA :	0,66
F-score	0,54	0,63	0,15	0,67	0,88	0,69	0,66	0,39	1,00	1,00	κ :	0,63

TABLEAU C.12 – *Matrice de confusion de la C4 par objets*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	0	0	45	0	5	0	0	0	87	0,43
FL	0	22	36	0	0	0	31	0	0	0	89	0,25
FJ	0	0	16	0	13	0	63	0	0	0	92	0,17
ZH	0	0	2	61	10	0	3	0	22	0	98	0,62
FM	0	0	3	0	78	0	0	0	0	0	81	0,96
R	0	0	0	0	0	76	1	0	0	0	77	0,99
FA	0	0	5	0	0	0	89	0	0	0	94	0,95
M	0	0	8	0	10	52	2	21	0	0	93	0,23
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	37	22	70	61	156	128	194	21	122	100		
UA	1,00	1,00	0,23	1,00	0,50	0,59	0,46	1,00	0,82	1,00	OA :	0,66
F-score	0,60	0,40	0,20	0,77	0,66	0,74	0,62	0,37	0,90	1,00	κ :	0,62

TABLEAU C.13 – *Matrice de confusion de la C5 par pixels*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	43	0	23	0	10	0	0	0	11	0	87	0,49
FL	0	54	4	0	0	1	16	14	0	0	89	0,61
FJ	2	11	20	0	11	2	41	5	0	0	92	0,22
ZH	5	0	33	32	0	17	0	1	10	0	98	0,33
FM	41	0	3	0	24	0	2	0	11	0	81	0,30
R	0	5	23	1	0	46	0	2	0	0	77	0,60
FA	0	21	5	0	13	0	47	7	1	0	94	0,50
M	2	7	8	0	1	52	0	23	0	0	93	0,25
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	93	98	119	33	59	118	106	52	133	100		
UA	0,46	0,55	0,17	0,97	0,41	0,39	0,44	0,44	0,75	1,00	OA :	0,54
F-score	0,48	0,58	0,19	0,49	0,34	0,47	0,47	0,32	0,86	1,00	κ :	0,49

TABLEAU C.14 – *Matrice de confusion de la C5 par classe majoritaire*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	50	0	0	0	0	0	0	0	87	0,43
FL	0	41	0	0	0	0	13	35	0	0	89	0,46
FJ	0	0	8	0	0	0	84	0	0	0	92	0,09
ZH	0	0	44	27	1	15	0	0	11	0	98	0,28
FM	72	0	3	0	6	0	0	0	0	0	81	0,07
R	0	0	0	0	0	76	1	0	0	0	77	0,99
FA	0	0	0	0	0	0	94	0	0	0	94	1,00
M	0	5	10	0	0	52	0	26	0	0	93	0,28
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	109	46	115	27	7	143	192	61	111	100		
UA	0,34	0,89	0,07	1,00	0,86	0,53	0,49	0,43	0,90	1,00	OA :	0,57
F-score	0,38	0,61	0,08	0,43	0,14	0,69	0,66	0,34	0,95	1,00	κ :	0,52

TABLEAU C.15 – *Matrice de confusion de la C5 par objets*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	0	0	45	0	5	0	0	0	87	0,43
FL	0	22	36	0	0	0	31	0	0	0	89	0,25
FJ	0	0	16	0	13	0	63	0	0	0	92	0,17
ZH	0	0	2	72	10	0	3	0	11	0	98	0,73
FM	0	0	3	0	78	0	0	0	0	0	81	0,96
R	0	0	0	0	0	76	1	0	0	0	77	0,99
FA	0	0	5	0	0	0	89	0	0	0	94	0,95
M	0	0	8	0	10	52	0	23	0	0	93	0,25
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	37	22	70	72	156	128	192	23	111	100		
UA	1,00	1,00	0,23	1,00	0,50	0,59	0,46	1,00	0,90	1,00	OA :	0,67
F-score	0,60	0,40	0,20	0,85	0,66	0,74	0,62	0,40	0,95	1,00	κ :	0,64

TABLEAU C.16 – *Matrice de confusion de la C6 par pixels*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	38	0	25	0	13	0	0	0	11	0	87	0,44
FL	0	46	7	0	0	2	0	34	0	0	89	0,52
FJ	1	11	50	0	14	2	0	14	0	0	92	0,54
ZH	2	0	42	30	1	17	0	0	6	0	98	0,31
FM	46	0	7	0	17	0	0	0	11	0	81	0,21
R	0	2	21	2	0	52	0	0	0	0	77	0,68
FA	3	12	36	0	11	3	0	28	1	0	94	0,00
M	0	9	11	0	4	51	0	18	0	0	93	0,19
HF	0	0	0	1	0	0	0	0	99	0	100	0,99
EAU	0	0	0	1	0	0	0	0	4	95	100	0,95
∑	90	80	199	34	60	127	0	94	132	95		
UA	0,42	0,58	0,25	0,88	0,28	0,41	-	0,19	0,75	1,00	OA :	0,49
F-score	0,43	0,54	0,34	0,45	0,24	0,51	-	0,19	0,85	0,97	κ :	0,43

TABLEAU C.17 – *Matrice de confusion de la C6 par classe majoritaire*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	50	0	0	0	0	0	0	0	87	0,43
FL	0	31	39	0	0	0	0	19	0	0	89	0,35
FJ	0	0	92	0	0	0	0	0	0	0	92	1,00
ZH	0	0	46	31	0	21	0	0	0	0	98	0,32
FM	78	0	3	0	0	0	0	0	0	0	81	0,00
R	0	0	1	0	0	76	0	0	0	0	77	0,99
FA	0	0	54	0	0	0	0	40	0	0	94	0,00
M	0	5	10	0	0	52	0	26	0	0	93	0,28
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	5	95	100	0,95
∑	115	36	295	31	0	149	0	85	105	95		
UA	0,32	0,86	0,31	1,00	-	0,51	-	0,31	0,95	1,00	OA :	0,54
F-score	0,37	0,50	0,48	0,48	-	0,67	-	0,29	0,98	0,97	κ :	0,48

TABLEAU C.18 – *Matrice de confusion de la C6 par objets*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	0	0	45	0	5	0	0	0	87	0,43
FL	0	22	36	0	0	0	31	0	0	0	89	0,25
FJ	0	0	16	0	13	0	63	0	0	0	92	0,17
ZH	0	0	1	84	10	0	3	0	0	0	98	0,86
FM	0	0	3	0	78	0	0	0	0	0	81	0,96
R	0	0	0	0	0	76	1	0	0	0	77	0,99
FA	0	0	5	0	0	0	89	0	0	0	94	0,95
M	0	5	8	0	10	49	2	16	0	3	93	0,17
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	37	27	69	84	156	125	194	16	100	103		
UA	1,00	0,81	0,23	1,00	0,50	0,61	0,46	1,00	1,00	0,97	OA :	0,68
F-score	0,60	0,38	0,20	0,92	0,66	0,75	0,62	0,29	1,00	0,99	κ :	0,64

TABLEAU C.19 – *Matrice de confusion de la C7 par pixels*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	45	0	27	0	6	0	0	0	9	0	87	0,52
FL	0	40	7	0	0	1	0	41	0	0	89	0,45
FJ	2	10	53	0	9	2	0	16	0	0	92	0,58
ZH	3	0	44	27	0	15	0	0	9	0	98	0,28
FM	46	0	7	0	21	0	0	0	7	0	81	0,26
R	0	2	27	1	0	47	0	0	0	0	77	0,61
FA	5	11	31	0	6	2	1	37	1	0	94	0,01
M	0	8	13	0	2	50	0	20	0	0	93	0,22
HF	1	0	0	1	0	0	0	0	98	0	100	0,98
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	102	71	209	29	44	117	1	114	124	100		
UA	0,44	0,56	0,25	0,93	0,48	0,40	1,00	0,18	0,79	1,00	OA :	0,50
F-score	0,48	0,50	0,35	0,43	0,34	0,48	0,02	0,19	0,88	1,00	κ :	0,44

TABLEAU C.20 – *Matrice de confusion de la C7 par classe majoritaire*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	50	0	0	0	0	0	0	0	87	0,43
FL	0	27	6	0	0	0	0	56	0	0	89	0,30
FJ	0	0	59	0	0	0	0	33	0	0	92	0,64
ZH	1	0	54	31	0	12	0	0	0	0	98	0,32
FM	57	0	3	0	21	0	0	0	0	0	81	0,26
R	0	0	1	0	0	76	0	0	0	0	77	0,99
FA	0	0	53	0	0	0	0	41	0	0	94	0,00
M	0	5	10	0	0	52	0	26	0	0	93	0,28
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	95	32	236	31	21	140	0	156	100	100		
UA	0,39	0,84	0,25	1,00	1,00	0,54	-	0,17	1,00	1,00	OA :	0,52
F-score	0,41	0,45	0,36	0,48	0,41	0,70	-	0,21	1,00	1,00	κ :	0,47

TABLEAU C.21 – *Matrice de confusion de la C7 par objets*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	0	0	45	0	5	0	0	0	87	0,43
FL	0	22	36	0	0	0	31	0	0	0	89	0,25
FJ	0	0	16	0	13	0	63	0	0	0	92	0,17
ZH	0	0	1	43	10	0	3	0	41	0	98	0,44
FM	0	0	3	0	78	0	0	0	0	0	81	0,96
R	0	0	0	0	0	76	1	0	0	0	77	0,99
FA	0	0	5	0	0	0	89	0	0	0	94	0,95
M	0	5	8	0	10	52	2	16	0	0	93	0,17
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	37	27	69	43	156	128	194	16	141	100		
UA	1,00	0,81	0,23	1,00	0,50	0,59	0,46	1,00	0,71	1,00	OA :	0,63
F-score	0,60	0,38	0,20	0,61	0,66	0,74	0,62	0,29	0,83	1,00	κ :	0,59

TABLEAU C.22 – *Matrice de confusion de la C8 par pixels*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	41	0	18	0	13	0	0	4	11	0	87	0,47
FL	0	52	4	0	0	1	19	13	0	0	89	0,58
FJ	2	11	21	0	14	2	37	5	0	0	92	0,23
ZH	4	0	27	32	0	17	0	6	12	0	98	0,33
FM	47	0	3	0	14	0	2	0	15	0	81	0,17
R	0	6	21	1	0	46	0	2	1	0	77	0,60
FA	0	21	9	0	11	0	46	6	1	0	94	0,49
M	2	7	5	0	2	52	0	25	0	0	93	0,27
HF	1	0	0	1	0	0	0	0	98	0	100	0,98
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	97	97	108	34	54	118	104	61	138	100		
UA	0,42	0,54	0,19	0,94	0,26	0,39	0,44	0,41	0,71	1,00	OA :	0,52
F-score	0,45	0,56	0,21	0,48	0,21	0,47	0,46	0,32	0,82	1,00	κ :	0,47

TABLEAU C.23 – *Matrice de confusion de la C8 par classe majoritaire*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	44	0	6	0	0	0	0	0	87	0,43
FL	0	32	0	0	0	0	22	35	0	0	89	0,36
FJ	0	0	15	0	0	0	77	0	0	0	92	0,16
ZH	0	0	30	49	1	15	0	3	0	0	98	0,50
FM	72	0	3	0	6	0	0	0	0	0	81	0,07
R	0	0	0	0	0	76	0	1	0	0	77	0,99
FA	0	0	4	0	0	0	90	0	0	0	94	0,96
M	0	5	10	0	0	52	0	26	0	0	93	0,28
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	109	37	106	49	13	143	189	65	100	100		
UA	0,34	0,86	0,14	1,00	0,46	0,53	0,48	0,40	1,00	1,00	OA :	0,58
F-score	0,38	0,51	0,15	0,67	0,13	0,69	0,64	0,33	1,00	1,00	κ :	0,54

TABLEAU C.24 – *Matrice de confusion de la C8 par objets*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	5	0	45	0	0	0	0	0	87	0,43
FL	0	22	36	0	0	0	31	0	0	0	89	0,25
FJ	0	0	16	0	13	0	63	0	0	0	92	0,17
ZH	1	0	2	92	0	0	3	0	0	0	98	0,94
FM	58	0	3	0	20	0	0	0	0	0	81	0,25
R	0	0	0	0	0	76	1	0	0	0	77	0,99
FA	0	0	5	0	0	0	89	0	0	0	94	0,95
M	0	5	8	0	10	52	2	16	0	0	93	0,17
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	96	27	75	92	88	128	189	16	100	100		
UA	0,39	0,81	0,21	1,00	0,23	0,59	0,47	1,00	1,00	1,00	OA :	0,62
F-score	0,40	0,38	0,19	0,97	0,24	0,74	0,63	0,29	1,00	1,00	κ :	0,58

TABLEAU C.25 – *Matrice de confusion de la C9 par pixels*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	42	0	12	0	10	0	0	12	11	0	87	0,48
FL	0	53	5	0	0	1	19	11	0	0	89	0,60
FJ	2	11	21	0	11	2	38	7	0	0	92	0,23
ZH	5	0	1	30	0	17	0	34	11	0	98	0,31
FM	54	0	3	0	14	0	2	1	7	0	81	0,17
R	0	5	8	1	0	51	0	12	0	0	77	0,66
FA	0	21	7	0	10	0	49	6	1	0	94	0,52
M	3	8	2	0	1	51	0	28	0	0	93	0,30
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	106	98	59	31	46	122	108	111	130	100		
UA	0,40	0,54	0,36	0,97	0,30	0,42	0,45	0,25	0,77	1,00	OA :	0,54
F-score	0,44	0,57	0,28	0,47	0,22	0,51	0,49	0,27	0,87	1,00	κ :	0,48

TABLEAU C.26 – *Matrice de confusion de la C9 par classe majoritaire*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	43	0	0	0	0	0	5	39	0	0	87	0,49
FL	0	32	0	0	0	0	22	35	0	0	89	0,36
FJ	0	0	8	0	0	0	84	0	0	0	92	0,09
ZH	1	0	2	21	0	21	0	42	11	0	98	0,21
FM	72	0	3	0	6	0	0	0	0	0	81	0,07
R	0	0	0	0	0	76	0	1	0	0	77	0,99
FA	0	0	0	0	0	0	94	0	0	0	94	1,00
M	0	5	0	0	0	52	0	36	0	0	93	0,39
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	116	37	13	21	6	149	205	153	111	100		
UA	0,37	0,86	0,62	1,00	1,00	0,51	0,46	0,24	0,90	1,00	OA :	0,57
F-score	0,42	0,51	0,15	0,35	0,14	0,67	0,63	0,29	0,95	1,00	κ :	0,52

TABLEAU C.27 – *Matrice de confusion de la C9 par objets*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	5	0	45	0	0	0	0	0	87	0,43
FL	0	27	26	0	0	0	31	5	0	0	89	0,30
FJ	0	0	16	0	13	0	63	0	0	0	92	0,17
ZH	1	0	2	42	0	0	0	3	50	0	98	0,43
FM	58	0	3	0	20	0	0	0	0	0	81	0,25
R	0	0	1	0	0	76	0	0	0	0	77	0,99
FA	0	0	5	0	0	0	89	0	0	0	94	0,95
M	0	5	0	0	10	52	0	26	0	0	93	0,28
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	96	32	58	42	88	128	183	34	150	100		
UA	0,39	0,84	0,28	1,00	0,23	0,59	0,49	0,76	0,67	1,00	OA :	0,59
F-score	0,40	0,45	0,21	0,60	0,24	0,74	0,64	0,41	0,80	1,00	κ :	0,54

TABLEAU C.28 – *Matrice de confusion de la C10*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	37	0	0	0	45	0	5	0	0	0	87	0,43
FL	0	22	35	0	0	0	22	10	0	0	89	0,25
FJ	0	0	16	0	13	0	63	0	0	0	92	0,17
ZH	0	0	2	42	1	0	3	0	50	0	98	0,43
FM	0	0	3	0	78	0	0	0	0	0	81	0,96
R	0	0	0	0	0	76	1	0	0	0	77	0,99
FA	0	0	5	0	0	0	89	0	0	0	94	0,95
M	0	5	8	0	10	52	0	18	0	0	93	0,19
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	37	27	69	42	147	128	183	28	150	100		
UA	1,00	0,81	0,23	1,00	0,53	0,59	0,49	0,64	0,67	1,00	OA :	0,63
F-score	0,60	0,38	0,20	0,60	0,68	0,74	0,64	0,30	0,80	1,00	κ :	0,59

TABLEAU C.29 – *Matrice de confusion de la C11*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	36	0	0	0	46	0	0	5	0	0	87	0,41
FL	0	22	30	0	0	0	28	9	0	0	89	0,25
FJ	0	0	16	0	13	0	63	0	0	0	92	0,17
ZH	0	0	2	41	29	0	3	0	23	0	98	0,42
FM	0	0	3	0	78	0	0	0	0	0	81	0,96
R	0	0	22	0	0	54	1	0	0	0	77	0,70
FA	0	0	5	0	0	0	89	0	0	0	94	0,95
M	0	5	0	0	10	52	17	9	0	0	93	0,10
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	5	95	100	0,95
∑	36	27	78	41	176	106	201	23	128	95		
UA	1,00	0,81	0,21	1,00	0,44	0,51	0,44	0,39	0,78	1,00	OA :	0,59
F-score	0,59	0,38	0,19	0,59	0,61	0,59	0,60	0,16	0,88	0,97	κ :	0,55

TABLEAU C.30 – *Matrice de confusion de la C12*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	14	0	0	0	68	0	5	0	0	0	87	0,16
FL	0	22	41	0	0	0	8	18	0	0	89	0,25
FJ	0	0	16	0	13	0	63	0	0	0	92	0,17
ZH	0	0	2	72	10	0	3	0	11	0	98	0,73
FM	0	0	3	0	78	0	0	0	0	0	81	0,96
R	0	0	44	0	0	32	1	0	0	0	77	0,42
FA	0	0	5	0	0	0	89	0	0	0	94	0,95
M	0	0	5	0	10	52	0	26	0	0	93	0,28
HF	0	0	0	0	0	0	0	0	100	0	100	1,00
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	14	22	116	72	179	84	169	44	111	100		
UA	1,00	1,00	0,14	1,00	0,44	0,38	0,53	0,59	0,90	1,00	OA :	0,60
F-score	0,28	0,40	0,15	0,85	0,60	0,40	0,68	0,38	0,95	1,00	κ :	0,56

TABLEAU C.31 – *Matrice de confusion de la C13*

Obs. \ Préd.	FR	FL	FJ	ZH	FM	R	FA	M	HF	EAU	∑	PA
FR	0	0	5	0	23	0	37	8	0	0	73	0,00
FL	0	16	51	0	0	0	7	15	0	0	89	0,18
FJ	0	0	53	0	0	0	39	0	0	0	92	0,58
ZH	0	20	5	0	1	0	11	50	11	0	98	0,00
FM	0	0	3	0	78	0	0	0	0	0	81	0,96
R	0	66	1	0	0	10	0	0	0	0	77	0,13
FA	0	0	40	0	40	0	13	0	1	0	94	0,14
M	0	50	5	0	0	1	9	28	0	0	93	0,30
HF	0	0	0	0	0	0	0	0	98	2	100	0,98
EAU	0	0	0	0	0	0	0	0	0	100	100	1,00
∑	0	152	163	0	142	11	116	101	110	102		
UA	-	0,11	0,33	-	0,55	0,91	0,11	0,28	0,89	0,98	OA :	0,44
F-score	-	0,13	0,42	-	0,70	0,23	0,12	0,29	0,93	0,99	κ :	0,38

Cartographie détaillée des types de formations forestières dans le bassin du Congo par télédétection optique (Sentinel-2)
Concessions forestières de la CIB, Ouesso, République du Congo

Alexis del Marmol

Ce mémoire s'inscrit dans un projet de cartographie détaillée des forêts du bassin du Congo, dans un contexte international où la gestion durable des forêts équatoriales, qui se base sur de faibles connaissances de ces écosystèmes, est devenue un enjeu majeur. Il utilise les données multispectrales fournies par les satellites Sentinel-2 A et B récemment mis en orbite par l'Agence spatiale européenne. Plus précisément, ce travail étudie différentes méthodes de classification des types de forêts du bassin du Congo dans la région d'Ouesso (nord de la République du Congo) par l'extraction numérique des caractéristiques spectrales de ces formations végétales. Le travail est réalisé en deux temps : une mission de terrain d'une part et un traitement des images pour obtenir les produits cartographiques d'autre part. La mission de terrain, réalisée en collaboration avec la Congolaise industrielle des bois (CIB), a permis de décrire les types de forêt, de créer une clé d'interprétation visuelle des images satellites et d'obtenir des données de validation pour la classification. Le traitement des images a, quant à lui, consisté en l'extraction et le calcul de différentes métriques (réflectances, indices de végétation et textures) à partir des images satellites ayant reçu une correction atmosphérique. Différentes combinaisons de métriques sont utilisées pour la réalisation de différentes classifications dont les résultats sont discutés afin de mettre en lumière la pertinence des différentes métriques. Enfin, un produit consolidé par vote majoritaire est présenté pour appuyer les résultats obtenus par classification. Les résultats de ce mémoire et les recommandations qui sont données ouvrent la porte sur des nouvelles possibilités d'observation des forêts équatoriales par télédétection.

Université catholique de Louvain

Faculté des bioingénieurs

Croix du Sud, 2bte L7.05.01, 1348 Louvain-La-Neuve, Belgique | www.uclouvain.be/agro