

Faculté des bioingénieurs

Cartographie de types de forêt et de l'occupation du sol en République Centrafricaine à l'aide de télédétection optique (Sentinel-2)

Aires Protégées de Dzanga-Sangha, RCA

Auteur : Dorian Lessire
Promoteurs : Prof. Pierre Defourny
Dr. Sophie Bontemps
Lecteurs : Prof. Quentin Ponette
Dr. Philippe Mayaux
Année Académique : 2018 - 2019
Bioingénieur : Sciences et technologies de l'environnement

Remerciements

Je tiens tout d'abord à remercier mon promoteur, le professeur Pierre Defourny, de m'avoir aidé et guidé tout au long de ce mémoire, et même plus largement durant ces deux années de master. C'est aussi grâce à lui que j'ai eu la chance de profiter d'une formidable expérience de terrain. Pour tout cela, je lui en suis très reconnaissant.

Un grand merci aussi à ma co-promotrice, Dr. Sophie Bontemps, pour ses relectures et conseils avisés.

Je remercie aussi le Professeur Quentin Ponette et le Dr. Philippe Mayaux d'avoir accepté d'être les lecteurs de ce travail.

Je souhaite ensuite remercier l'ensemble des chercheurs du labo pour leur aide précieuse.

Ce mémoire n'aurait pas pu se faire sans la campagne de terrain menée en collaboration avec le WWF-RCA. Pour cela, je tiens à remercier l'ensemble des personnes qui ont permis ce séjour possible et pour leur formidable accueil. À la fois tout le personnel du WWF et de l'administration locale. Particulièrement, je voudrais remercier Luis Arranz pour son implication dans cette aventure. Un tout grand merci aussi la population de Bayanga et à Liwa et Félix, deux pisteurs qui m'ont toujours guidé avec courage dans la forêt.

Un grand merci aussi à mes compagnons de mémoire et à tous mes camarades durant ces années universitaires.

Je tiens à remercier particulièrement Baptiste, pour ces bons moments passés en RCA.

Enfin, je remercie aussi mon père et mes grands-parents pour leurs soutiens indéfectibles tout au long de mon parcours scolaire.

Table des matières

Table des figures	V
Liste des tableaux	VI
Liste des acronymes	VII
Introduction	1
1 Revue Bibliographique	2
1.1 Généralités sur la forêt tropicale	2
1.1.1 Enjeux	2
1.1.2 Types Forestiers	4
1.1.3 Dynamique forestière	6
1.2 Télédétection	6
1.2.1 Télédétection appliquée aux forêts tropicales	6
1.2.2 Types d'images	7
1.2.3 Le potentiel de Sentinel-2	8
1.2.4 Indices de végétation & Texture	9
1.3 Classification	11
1.3.1 Classification non supervisée	11
1.3.2 Classification supervisée	12
1.3.3 Évaluation des classifications	14
1.4 Conclusion	16
2 Objectifs	17
3 Zone d'étude	18
3.1 Localisation	18
3.2 Description physique	18
3.2.1 Climat	18
3.2.2 Topographie	18
3.2.3 Hydrographie	20
3.2.4 Géologie & Pédologie	20
3.2.5 Typologie de végétation	20
3.3 Planification agricole	21
4 Matériel & Méthode	23
4.1 Images Sentinel-2	23
4.1.1 Traitement des images Sentinel-2	23
4.2 Campagne de terrain	23
4.2.1 Protocole de terrain	25
4.2.2 Zones d'entraînement	28
4.3 Comparaison des méthodes de classification	30
4.3.1 Choix de la zone de développement	30
4.3.2 Description des méthodes de classification	30
4.3.3 Protocole de validation	33
4.4 Production cartographique	35
4.4.1 Protocole de validation	35

4.5	Estimation de la dynamique agricole	36
4.5.1	Images satellites	36
4.5.2	Évaluation des surfaces non forestières	36
5	Résultats & Discussion	38
5.1	Typologie de végétation	38
5.1.1	Forêt sempervirente - <i>Gilbertiodendron dewevrei</i>	38
5.1.2	Forêt dense humide semi-décidue	40
5.1.3	Forêt secondaire vieille	41
5.1.4	Forêt secondaire jeune	42
5.1.5	Forêt marécageuse	43
5.1.6	Forêt ripicole haute	44
5.1.7	Forêt inondable	45
5.1.8	Clairière saline	46
5.1.9	Savane	47
5.1.10	Zone agricole	48
5.1.11	Autres occupations du sol	48
5.2	Comparaison des méthodes de classification	51
5.2.1	Performance de la classification non supervisée - Approche par pixel	51
5.2.2	Performance de la classification supervisée - Approche par pixel	52
5.2.3	Performance de la classification supervisée - Approche hiérarchique par pixel	54
5.2.4	Performance de la classification supervisée - méthode par objet	57
5.2.5	Comparaison des classifications	59
5.3	Production cartographique	61
5.3.1	Développer la méthode à plus large échelle	64
5.4	Dynamique Agricole	65
	Conclusion et perspectives	68
	A Matrices de confusion	74
	B Masque nuageux	77
	C Dynamique agricole	81

Table des figures

1	Illustration de la forêt tropicale en RCA	4
2	Source d'énergie interceptée par un capteur multispectral (Source : Centre canadien de télédétection (n.d.))	8
3	Comparaison d'images multispectrales et hyperspectrales (Source : Aboras et al. (2015))	8
4	Sentinel-2 : Résolutions spatiales des différentes bandes spectrales . .	9
5	Carte des Aires Protégées de Dzanga Sangha	19
6	Altitude au sein des APDS	20
7	Défrichement de zones forestières au profit de zones agricoles	21
8	Bandes culturales au sein des APDS	22
9	Workflow complet de la méthode générale	24
10	Tuiles de Sentinel-2 recouvrant les APDS	25
11	Composition colorée d'images Sentinel-2 sur la zone d'étude, 27/12/2018, Bandes : B07-B11-B05	27
12	Répartition des zones d'entraînements	29
13	Zone de développement	30
14	Méthode de classification hiérarchique	31
15	Définition des objets par segmentation - Image Sentinel-2, composi- tion colorée : B07-B11-B05	32
16	Répartition des zones d'entraînement et de validation dans la zone de développement	33
17	Répartition des aires de validation pour la comparaison de classifica- tion dans la zone de développement	34
18	Répartition des aires de validation pour la zone d'étude	35
19	Différenciation des zones non forestières - Secteur Lidjombo	37
20	Forêt sempervirente Composition colorée : B07-B11-B05 (Sentinel-2) .	38
21	Forêt sempervirente - <i>Gilbertiodendron dewevrei</i>	39
22	Forêt dense humide semi-décidue Composition colorée : B07-B11-B05 (Sentinel-2)	40
23	Forêt dense humide semi-décidue	40
24	Forêt secondaire vieille Composition colorée : B07-B11-B05 (Sentinel-2)	41
25	Forêt secondaire vieille	41
26	Ancien chemin d'exploitation au sein des forêts secondaires vieilles . .	41
27	Forêt secondaire jeune Composition colorée : B07-B11-B05 (Sentinel-2)	42
28	Forêt secondaire jeune	42
29	Forêt marécageuse Composition colorée : B07-B11-B05 (Sentinel-2) .	43
30	Forêt marécageuse - ici monospécifique de raphia	43
31	Forêt ripicole haute Composition colorée : B07-B11-B05 (Sentinel-2) .	44
32	Forêt ripicole haute	44
33	Forêt inondable - Composition colorée : B07-B11-B05 (Sentinel-2) . .	45
34	Forêt inondable	45
35	Clairière saline - <i>Dzanga Bai</i>	46
36	Saline Composition colorée : B07-B11-B05 (Sentinel-2)	46
37	Clairière saline	46
38	Savane Composition colorée : B07-B11-B05 (Sentinel-2)	47
39	Savane arbustive de type savane incluse	47

40	Zone agricole autour de Bayanga	48
41	Séquence de culture classique de l'agriculture sur brûlis pratiquée au sein des APDS	48
42	Zone d'habitation autour de Bayanga	49
43	Eau : rivière Sangha aux abords de Bayanga	49
44	Photo-interprétation de l'occupation du sol de la zone d'étude	50
45	Classification non supervisée par ISODATA - Zoom à proximité de Bayanga	52
46	Classification supervisée non hiérarchique basée sur les pixels - Zoom à proximité de Bayanga	54
47	Classification supervisée hiérarchique basée sur les pixels (12 classes) - Zoom à proximité de Bayanga	56
48	Comparaison des résultats avec l'utilisation d'un filtre majoritaire	57
49	Classification supervisée via une approche par objet - Zoom à proximité de Bayanga	58
50	Variabilité de réflectance au sein de différentes classes d'occupation du sol	60
51	Cartographie des types forestiers et de l'occupation du sol au sein des APDS (RCA)	62
52	Cartographie des types forestiers et de l'occupation du sol au sein des APDS (RCA) - Zoom	63
53	Évolution des surfaces agricoles dans le secteur de Bayanga centre	67
B.1	Couvertures nuageuses pour la tuile 33 NXD (Sentinel-2) pour la dernière année (partie 1)	77
B.2	Couvertures nuageuses pour la tuile 33 NXD (Sentinel-2) pour la dernière année (partie 2) En vert, les images utilisées lors de ce travail	78
B.3	Couvertures nuageuses pour la tuile 33 NXD (Sentinel-2) pour la dernière année (partie 3) En vert, les images utilisées lors de ce travail	79
B.4	Couvertures nuageuses pour la tuile 33 NXD (Sentinel-2) pour la dernière année (partie 4)	80
C.1	Évolution des surfaces agricoles dans le secteur de Babongo Nord	82
C.2	Évolution des surfaces agricoles dans le secteur de Babongo Sud	83
C.3	Évolution des surfaces agricoles dans le secteur de Moussapoula Nord	84
C.4	Évolution des surfaces agricoles dans le secteur de Moussapoula Sud	85
C.5	Évolution des surfaces agricoles dans le secteur de Bayanga centre	86
C.6	Évolution des surfaces agricoles dans le secteur de Yadoumbe	87
C.7	Évolution des surfaces agricoles dans le secteur de Lindjombo	88
C.8	Évolution des surfaces agricoles dans le secteur de Bomandjokou	89

Liste des tableaux

1	Exemple d'une matrice de confusion	14
2	Évaluation de la qualité d'une classification selon l'indice kappa	15
3	Détail des aires d'entraînement	29
4	Paramètres de segmentation	32
5	Détail des aires de validation pour la zone d'entraînement	34
6	Détail des zones de validation pour la cartographie complète	36
7	Dates des images utiles pour la comparaison des surfaces agricoles dans chacun des secteurs agricoles	36
8	Tableau Récapitulatif de la matrice de confusion pour la classification par ISODATA.	52
9	Analyse de la matrice de confusion de la classification non hiérarchique basée sur les pixels - 12 classes	53
10	Analyse de la matrice de confusion de la classification hiérarchique (4 classes) - Approche par pixel	55
11	Analyse de la matrice de confusion de la classification hiérarchique via une approche par pixel - 12 classes	55
12	Comparaison des qualités de classification avec l'apport de filtre majoritaire	57
13	Analyse de la matrice de confusion de la classification supervisée via une approche par objet sur la zone de développement	58
14	Comparaison des méthodes de classification sur la zone de développement	59
15	Analyse de la classification sur la zone d'étude	61
16	Évolution des surfaces des zones agricoles au sein des APDS	66
A.1	Matrice de Confusion de la classification par Isodata sur la zone de développement	74
A.2	Matrice de Confusion de la classification supervisée non hiérarchique via une approche par pixel - 12 classes	74
A.3	Matrice de Confusion de la classification supervisée hiérarchique via une approche par pixel - 4 classes	74
A.4	Matrice de Confusion de la classification supervisée hiérarchique via une approche par pixel - 12 classes	75
A.5	Matrice de Confusion de la classification supervisée via une approche par objet - Zone de développement	75
A.6	Matrice de Confusion de la classification supervisée via une approche par objet - Zone d'étude	76
C.1	Détails des bandes culturales au sein des APDS	81

Liste d'acronymes

APDS	Aires Protégées de Dzanga-Sangha
BdC	Bassin du Congo
COMIFAC	Commission des Forêts d'Afrique Centrale
ESA	European Space Agency
EVI	Enhanced Vegetation Index
GLCM	Gray-Level Co-Occurrence Matrix
ISODATA	Iterative Self-Organizing Data Analysis Technique
LAI	Leaf Area Index
MLC	Maximum Likelihood Classification
MODIS	Moderate-Resolution Imaging Spectroradiometer
NDVI	Normalized Difference Vegetation Index
NIR	Near-Infrared
OA	Overall Accuracy
OB	Object-Based
OTB	Orfeo Toolbox
PA	Producer Accuracy
PB	Pixel-Based
RCA	République Centrafricaine
RDC	République Démocratique du Congo
RF	Random Forest
ROI	Region Of Interest
SAR	Synthetic Aperture Radar
SIG	Système d'Information Géographique
SIWSI	Shortwave Infrared Water Stress Index
SVM	Support Vector Machine
SWIR	Short-Wave Infrared
TNS	Tri-National de la Sangha
UA	User Accuracy
WWF	World Wildlife Fund

Introduction

Les forêts tropicales sont un écosystème majeur de notre planète de par la diversité animale et végétale qu'elles abritent, mais aussi de par leur rôle (économique, culturel...) auprès des populations locales. De plus, leur impact sur le reste de la planète à travers leur rôle dans le bilan carbone et donc le climat devient un sujet de plus en plus traité dans le monde scientifique. Or cet habitat particulier est soumis à diverses pressions : anthropiques avec notamment la déforestation, climatique avec les changements en cours, etc. Pouvoir quantifier ces phénomènes et évaluer les résistances et résiliences de ces milieux en demande une connaissance accrue. Cependant, une grande partie de ces surfaces restent peu ou mal connues, et ce pour diverses raisons dont, notamment, la difficulté d'accessibilité.

C'est pourquoi l'utilisation de la télédétection est un atout majeur pour la surveillance de ces milieux, car elle permet de collecter des informations et de détecter des phénomènes à distance. Sur une initiative de la Commission européenne, en partenariat avec l'Agence Spatiale Européenne (ESA), le programme Copernicus a été mis en place. Son but est de fournir des données, d'observer et d'analyser l'état de la planète. Dans cet objectif, le lancement des satellites Sentinel-2 est une opportunité majeure puisqu'ils permettent d'obtenir des images précises (10 mètres) et régulières (5 jours) de l'entièreté du globe.

Suite aux avancées dans le domaine, les pays membres de la COMIFAC (Commission des forêts d'Afrique Centrale) ont sollicité l'Observatoire des Forêts du Bassin du Congo pour développer une cartographie détaillée des peuplements forestiers.

Dans ce cadre, ce mémoire a pour but de tenter d'améliorer les méthodes de classification et de cartographie des forêts d'Afrique centrale. Plus spécifiquement, il se concentre sur l'établissement d'un protocole de terrain applicable à large échelle ainsi qu'une méthode de cartographie.

Cette étude se concentre sur une partie restreinte des forêts d'Afrique centrale et plus particulièrement sur des zones protégées dans le sud-ouest de la République Centrafricaine (RCA) : les Aires Protégées de Dzanga-Sangha (APDS). En plus d'une cartographie des peuplements forestiers, une carte d'occupation du sol sera produite sur l'ensemble du territoire de la réserve en réponse à la demande des gestionnaires des aires protégées.

Ce travail sera structuré de la manière suivante. Tout d'abord, une revue de la littérature scientifique existante sur le sujet permettra de faire le bilan sur l'avancement des recherches actuelles. Ensuite, les objectifs du travail seront posés. Dans la section suivante sera présentée la zone d'étude. Dans la quatrième section seront expliquées les sources des données utilisées ainsi que les différents traitements opérés. Enfin, la partie finale se concentrera sur les résultats obtenus ainsi que leur discussion.

1 Revue Bibliographique

La revue bibliographique présentée ci-dessous a pour but de remettre dans le contexte scientifique le travail qui est présenté. Cela passe par un état des lieux des connaissances actuelles dans les domaines qui seront traités.

La première partie est consacrée à la description des forêts tropicales à travers ses enjeux, ses caractéristiques et sa dynamique ainsi qu'à une définition des types forestiers. Dans un second temps, une description des outils de télédétection utilisés dans ce travail est nécessaire avec notamment une attention spécifique apportée à la plateforme Sentinel-2. Enfin, la dernière partie s'intéressera à la méthode de classification employée. De plus, une considération particulière sera portée sur l'évaluation de ces classifications.

La littérature étant vaste et fournie, la traiter dans son intégralité n'est pas possible, un choix quant aux articles lus et utilisés est nécessaire. Dans un maximum de cas, les informations seront présentées de manière générale dans un premier temps avant d'être plus détaillées par la suite.

1.1 Généralités sur la forêt tropicale

Les forêts tropicales sont logiquement caractéristiques des régions chaudes et humides. Elles se situent entre les tropiques du Cancer et du Capricorne et comme tout type de végétation, elles sont étroitement liées au climat de ces régions, qualifié de tropical humide. En RCA, deux saisons sont en alternance, la saison des pluies de mars à novembre et la saison sèche, de décembre à février. Les précipitations annuelles varient entre 1400 mm et 1700 mm (Réjou-Méchain et al., 2008).

Ce type de végétation se retrouve en Amérique du Sud, en Asie et en Afrique. L'Afrique centrale possède la deuxième plus grande étendue de forêt tropicale au monde, derrière la forêt amazonienne, avec près de 1,8 million de km² de forêts denses humides (Mayaux and Achard, 1998). Le bassin forestier centrafricain s'étend sur 6 pays : la République Démocratique du Congo (RDC), le Cameroun, le Gabon, la Guinée équatoriale, la République du Congo et enfin la RCA.

Les forêts tropicales contiennent la plus importante réserve de biodiversité, animale et végétale, de notre planète et de nombreuses espèces y sont par ailleurs endémiques (Laurin et al., 2016). De plus, la forêt est une ressource majeure de la région pour la population, car près de 60 millions de personnes en dépendent directement : via la chasse et la cueillette pour 20% d'entre eux et via l'agriculture sur brûlis pour 80% (Mayaux and Achard, 1998, de Wasseige et al., 2014).

1.1.1 Enjeux

Les forêts tropicales sont au centre de beaucoup d'attention, car de nombreux enjeux pèsent sur elles. Tout d'abord, les forêts tropicales sont un stock majeur de carbone sur la planète et leurs contributions dans la balance globale sont encore fort méconnues. Une meilleure connaissance de ce type de biotope ainsi qu'une évaluation des changements d'occupation des sols (*land cover*) permettraient de quantifier la masse de CO₂ libérée et l'impact sur le changement climatique global (Baccini et al., 2017).

Mieux connaître et comprendre les forêts tropicales permettrait d'évaluer leurs résistances face au changement climatique (Laurin et al., 2016). Il a été en effet démontré que ces changements climatiques risquent d'augmenter les températures moyennes, de modifier les schémas actuels de précipitations et d'augmenter les périodes de saison sèche (Zelazowski et al., 2011). Mieux comprendre la façon dont la végétation actuelle, du point de vue de la structure et de la productivité, est influencée par les conditions climatiques peut permettre de mieux anticiper ses réactions face à des changements brutaux (Gond et al., 2013, 2015).

Par ailleurs, une menace importante pour les forêts tropicales est la déforestation d'origine anthropique. Il a par exemple été montré par Yoneda et al. (2009) qu'en Indonésie, la composition de certaines forêts a été modifiée en faveur d'espèces pionnières suite à leur exploitation. La région d'Afrique centrale est relativement épargnée par l'exploitation forestière, et ce principalement grâce à l'absence de marchés locaux de grande importance et d'infrastructures de transport. Néanmoins, des différences spatiales existent entre les zones plus accessibles et donc plus facilement exploitables, comme le long des routes, des cours d'eau ou en bordure de massif par rapport aux zones plus reculées et difficiles d'accès (Ernst et al., 2012). La plus grande pression sur les zones forestières provient de l'exploitation agricole qui est directement liée à l'augmentation démographique. Les villageois étendent les zones destinées à l'agriculture autour des villages et des routes au détriment de la forêt. Ernst et al. (2012) montrent qu'entre 1990 et 2000 la forêt du Bassin du Congo (BdC) a diminué de 0,13% et que durant les 5 années suivantes (période 2000-2005) ce taux a doublé. Ces tendances ne sont pourtant pas uniformes au travers du bassin et varient selon les pays. Ainsi, ces déforestations sont plus marquées et en hausse en RDC avec une déforestation nette de 0,22% (+0,11% par rapport à la période précédente) alors qu'elles restent stables en RCA (déforestation nette de 0,06%) (de Wasseige et al., 2014).

Au vu de ces éléments, la cartographie des forêts tropicales du BdC peut s'avérer être un outil puissant et indispensable, et ce pour diverses raisons.

Tout d'abord et de manière générale, pour pouvoir mettre en place des politiques et pour les suivre, à la fois à l'échelle régionale et à l'échelle locale, des informations fiables et récentes sont nécessaires. Cela peut permettre de mesurer l'impact et les effets des politiques de conservation menées par les organismes nationaux et internationaux (Ernst et al., 2012). Pour cela, disposer de cartes homogènes, du point de vue des échelles, des terminologies et des niveaux de détail, entre les différents pays d'Afrique Centrale est essentiel, or ce n'est actuellement pas encore le cas (Gond et al., 2015).

Ensuite, un autre intérêt pourrait être de parvenir à cartographier les différents types de forêts. Cela aurait pour avantage de mettre en place des pratiques d'extraction forestière plus sélectives et plus durables ainsi qu'une surveillance des espèces commerciales. Néanmoins pour avoir de tels résultats et bien que certaines couronnes des arbres tropicaux soient fort larges, une cartographie de ce type sur de larges surfaces demanderait des moyens considérables (Laurin et al., 2016).



FIGURE 1 – Illustration de la forêt tropicale en RCA

1.1.2 Types Forestiers

La question de la caractérisation des forêts est essentielle : sur quels critères se baser pour définir l'appartenance d'un peuplement ou d'une surface forestière à une classe définie ? Quels sont les traits spécifiques et comment les quantifier ?

Selon la "*Convention sur la diversité biologique*", un type forestier se définit comme "*un groupe d'écosystème forestier de composition généralement similaire qui peut être différencié d'autres groupes par sa composition d'espèces, sa productivité et/ou le degré de fermeture de la canopée*" (Convention on Biological Diversity, 2006, Laurin et al., 2016).

de Wasseige and Defourny (2002) ont identifié 3 éléments-clés sur lesquels se baser pour définir le type de végétation : la composition floristique, la structure de la forêt et la position physiogéographique.

Composition floristique Un premier élément de définition vient des espèces présentes afin de déterminer s'il en existe une dominante. Ce critère seul n'est pas suffisant vu le nombre d'espèces présentes en forêt tropicale. Elle peut néanmoins être utile dans le cas de massifs homogènes. La variation de types de végétation relève de plusieurs éléments et Réjou-Méchain et al. (2008) relèvent 3 théories en discussion au sein de la communauté scientifique concernant l'origine et le maintien d'espèces dans une région donnée. La première estime que la distribution des espèces est originellement uniforme dans l'espace et que seule l'apparition d'espèces plus rares varie (Pitman et al., 2001). La deuxième promeut une distribution spatiale qui suit un processus lié à la capacité de chaque espèce de se disperser. Cette théorie est connue sous le nom de "*The neutral Theory of Biodiversity*" (Hubbel, 2001). La distribution relève alors d'un processus démographique stochastique. La troisième théorie considère que la distribution des espèces est surtout liée aux conditions environnementales. De nombreuses études s'accordent par ailleurs pour mettre en évidence le rôle important du type de sol sur la végétation (Réjou-Méchain et al., 2008). Les distributions de type de forêt peuvent aussi être liées à des variations locales. Parmentier et al. (2005) montrent qu'en Afrique centrale, au niveau local la compo-

sition floristique est principalement reliée au type de sol, mais qu'à une échelle plus large, régionale par exemple, ce sont les processus historiques et les limitations à la dispersion qui sont les facteurs les plus importants.

Structure de la forêt Le deuxième élément clé, la structure, est défini par Pascal (2003) comme "*l'agencement des divers composants de la forêt*". Il distingue 2 types de structures : celles liées et celles non liées à la dimension spatiale. Les paramètres structuraux non spatialisés se basent sur ceux développés par les forestiers pour les forêts tempérées et sont repris de la même manière pour les forêts tropicales. Nous pouvons citer par exemple :

- **La distribution des diamètres.** Dans les forêts tropicales, la distribution est généralement une exponentielle décroissante, c'est-à-dire que l'on retrouve un grand nombre d'arbres de faible diamètre et un petit nombre d'arbres de diamètre plus imposant. Cela peut cependant être sujet aux exceptions selon le positionnement du peuplement (pente, exposition au vent, etc.) ou suite à des perturbations.
- **La densité ou la surface terrière.** La densité est le nombre d'arbres par unité de surface. Il faut bien entendu fixer une section minimum en dessous de laquelle les arbres ne sont pas pris en compte. Pour une forêt tropicale dite "classique", sans contrainte particulière, la moyenne se situe aux alentours des 600 tiges/ha. La surface terrière, quant à elle, est la somme des sections de troncs à 1,30 m reportée par unité de surface au sol. Cela nécessite donc des mesures de diamètre qui peuvent s'avérer difficiles à cause des contreforts présents sur de nombreuses espèces. La surface terrière en forêt tropicale varie entre 25 et 50 m²/ha avec des moyennes généralement plus hautes pour les forêts africaines en comparaison avec celles d'Amérique du Sud (Pascal, 2003).

Les paramètres structuraux spatialisés comprennent à la fois les dimensions verticale et horizontale. Ils sont sujets à des variations au niveau local dues aux dynamiques forestières, mais restent stables à partir d'une certaine échelle. La dimension verticale permet d'obtenir des informations liées à la hauteur du couvert forestier, aux strates présentes, aux degrés de fermeture et à l'étalement des couronnes. En forêt tropicale, les strates semblent moins marquées que dans les forêts tempérées et donc plus difficiles à utiliser comme caractéristique (Pascal, 2003). La dimension horizontale concerne la distribution des individus et des espèces au sein de l'espace. Cela fait intervenir toute une série de facteurs externes comme le type de sol, le drainage, la quantité de lumière disponible, la topographie ou encore la compétition entre espèces, mais aussi les activités humaines (de Wasseige and Defourny, 2002). Décrire la structure est donc quelque chose de complexe et qui demande un coût financier élevé puisqu'il faut caractériser les positions relatives des individus en fonction des autres individus (Coops and Culvenor, 2000).

Un bon indicateur de la structure est la rugosité de la canopée, qui indique la quantité de lumière capable de traverser la surface supérieure de la forêt. La rugosité est fonction de différentes caractéristiques : la taille des couronnes, la forme, la densité et enfin la structure verticale (de Wasseige and Defourny, 2002). Elle peut aussi se définir autour de la notion de trouées dans la canopée et est alors fonction de la profondeur, de la taille et de la distribution de celles-ci (Runkle, 1992).

Position physiogéographique. La localisation d'un peuplement forestier influe aussi sur sa composition, sa structure et sa dynamique. La pente, l'exposition, le type de sol sont des facteurs limitant ou, au contraire, favorisant certaines espèces. Par exemple, un peuplement situé en fond de vallée peut être soumis à un régime hydrique particulier, comme des inondations durant une partie de l'année.

Dans le cas de l'Afrique Centrale, Gond et al. (2013) ont identifié 22 types de végétation différents. Parmi ces 22 types, 11 sont des forêts divisées en 3 groupes : forêts semi-décidues denses, forêts sempervirentes denses et forêts marécageuses.

1.1.3 Dynamique forestière

La dynamique des forêts dépend de deux facteurs principaux : la dynamique naturelle et les facteurs anthropiques. Un équilibre se fait entre dégradation et régénération. La dégradation naturelle a plusieurs causes possibles, mais la majeure partie des ouvertures dans la canopée est due à des chutes d'arbres. Ces dernières peuvent avoir des origines différentes, mais cela influencera la taille des trouées résultantes. Par exemple, dans le cas d'un feu de forêt ou d'un cyclone important, les ouvertures seront plus importantes. Ces surfaces auront un impact sur le type de régénération qui en découlera : dans le cas de larges surfaces, des espèces héliophiles auront plus tendance à coloniser les surfaces. À noter que les dynamiques forestières dépendent des conditions physiques de leur environnement, du drainage du sol ou de l'exposition au vent impliquant des risques différents (Pascal, 2003).

Les activités humaines peuvent aussi influencer cette dynamique en jouant à la fois sur la dégradation, via des coupes, ou sur la régénération, via des plantations.

1.2 Télédétection

"La télédétection est la technique qui, par l'acquisition d'images, permet d'obtenir de l'information sur la surface de la Terre sans contact direct avec celle-ci. La télédétection englobe tout le processus qui constitue à capter et à enregistrer l'énergie d'un rayonnement électromagnétique émis ou réfléchi, à traiter et à analyser l'information, pour ensuite mettre en application cette information" (Centre canadien de télédétection, n.d.)

La télédétection s'est développée à grande vitesse durant les dernières décennies. Elle a pour avantage, pour les grandes surfaces où les informations sont difficiles à récolter sur le terrain, de permettre d'obtenir des données locales pour des surfaces importantes. L'utilisation de la télédétection pour étudier les forêts tropicales se fait depuis quelques dizaines d'années avec comme objectif le suivi de la déforestation et une caractérisation grossière (Laurin et al., 2016, Erinjery et al., 2018).

1.2.1 Télédétection appliquée aux forêts tropicales

L'utilisation de la télédétection dans les forêts tropicales n'est pas toujours aisée, car dans ces écosystèmes, de nombreuses espèces ont des réflectances similaires, la structure de la canopée est complexe et le comportement phénologique est irrégulier (Laurin et al., 2016). De plus, il existe une grande variabilité d'écosystèmes ce qui entraîne un nombre important de types de "land cover" différents.

Ensuite, le manque d'inventaire de terrain limite les données disponibles. (Erinjery et al., 2018, Ernst et al., 2012). Ceux-ci sont rendus compliqués par le manque d'infrastructures routières, qui rendent l'accès difficile, ainsi que le manque de stabilité politique dans certains pays (Mayaux and Achard, 1998).

Enfin, les régions tropicales sont soumises à une forte couverture nuageuse presque constante. Il est dès lors nécessaire de fixer un seuil au-delà duquel une image ne sera pas utilisable (Ernst et al., 2012). Erinjery et al. (2018), pour une zone située dans le sud de l'Inde ne sélectionne que les images avec moins de 1% de couverture nuageuse. De plus, il attire l'attention sur le choix judicieux de prendre des images en saison sèche. Cela permet de faire ressortir les arbres à feuilles caduques dans des forêts humides. Pour parer aux problèmes liés aux nuages récurrents, Gond et al. (2013), qui ont travaillé sur base d'images obtenues via le capteur *Moderate-Resolution Imaging Spectroradiometer* (MODIS), ont calculé sur base de 8 images et pour chacune des bandes, la valeur minimum du pixel qui est celle ayant la probabilité la plus faible d'être affectée par des effets atmosphériques. Un filtre est ensuite appliqué afin d'éliminer les éventuelles contaminations.

1.2.2 Types d'images

Différents types d'images potentiellement utiles pour la cartographie existent. Celles-ci peuvent provenir de deux grands types de capteurs : les actifs et les passifs. Les capteurs actifs fournissent des images radars tandis que les capteurs passifs donnent des images multispectrales.

Les images radars Les images radars proviennent de capteurs actifs, c'est-à-dire des capteurs qui vont émettre un signal dans une longueur d'onde. Ce signal va interagir avec les surfaces et retourner au capteur qui obtiendra alors de l'information sur les propriétés des objets. Ce type de capteur équipe le satellite Sentinel-1 et les images radar, "*Synthetic Aperture Radar*" (SAR) ont montré un potentiel intéressant dans le cadre de discrimination de forêts tropicales. Un de leurs avantages majeurs, et contrairement aux capteurs passifs, est qu'elles ne sont pas limitées par la couverture nuageuse. Elles permettent donc de réaliser des cartographies d'occupation du sol plus facilement en zone tropicale qu'une méthode qui se base uniquement sur des images multispectrales (Erinjery et al., 2018, Balzter et al., 2015).

Les images multispectrales Les images multispectrales proviennent de capteurs passifs. Ce type de capteur va détecter l'énergie émise et réfléchié naturellement dans son champ de vision. Cette énergie peut être émise par différentes sources : le soleil directement, l'atmosphère (1), réfléchié du sol (2), émise par le sol (3) ou transmise du sous-sol (4) (Figure 2) (Centre canadien de télédétection, n.d.). Les capteurs multispectraux vont pouvoir stocker les informations obtenues dans différentes bandes de longueurs d'onde.

Les capteurs hyperspectraux sont une sous-catégorie de multispectrale. Contrairement à ces dernières, les informations sont stockées dans un nombre de bandes plus important, mais d'intervalles de longueurs d'onde moindres. Cela donne donc un spectre d'aspect plus continu (Figure 3) (Société Française de Photogrammétrie et Télédétection, n.d.). Elles ont aussi prouvé leur efficacité dans la discrimination de

forêts tropicales et l'amélioration des classifications notamment par l'apport d'informations sur les structures forestières (Laurin et al., 2016, Asner et al., 2015). Cependant, les hyperspectrales ne sont pas applicables à de larges zones.

Les capteurs passifs sont cependant sensibles aux conditions atmosphériques et contrairement aux images radar, ils ne peuvent pas capter des signaux au travers d'une couche nuageuse.

Enfin, et pour une résolution spatiale égale, une classification basée sur des données multispectrales représente sans doute mieux la réalité qu'une classification basée uniquement sur des données radar. Mais c'est en combinant les 2 types de données que la meilleure exactitude est obtenue (Erinjery et al., 2018).

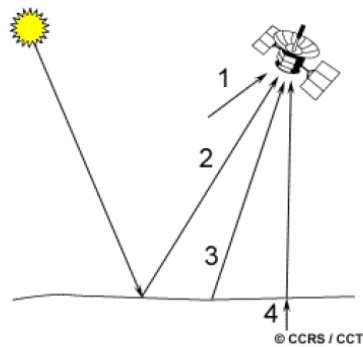


FIGURE 2 – Source d'énergie interceptée par un capteur multispectral (Source : Centre canadien de télédétection (n.d.))

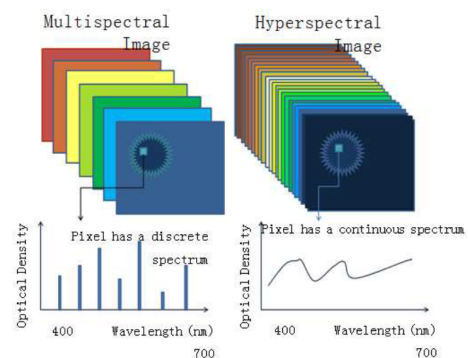


FIGURE 3 – Comparaison d'images multispectrales et hyperspectrales (Source : Aboras et al. (2015))

1.2.3 Le potentiel de Sentinel-2

La mission Sentinel-2 fait partie du programme Copernicus de la Commission européenne. Elle est composée de deux satellites (Sentinel-2A et Sentinel-2B) placés sur la même orbite en phase de 180° entre eux (ESA, 2019). Sentinel-2A a été mis en orbite en juin 2015 suivi par Sentinel-2B en mars 2017 (Immitzer et al., 2016).

Le choix du satellite et donc des images varie selon les auteurs et les études. Cependant, Sentinel-2 possède de nombreux avantages. Son temps de visite est de 10 jours par satellite, donc 5 jours au total et il possède un capteur multispectral composé de 13 bandes spectrales qui varient entre 443 nm et 2190 nm. Différentes résolutions spatiales existent : 4 bandes à 10m dans le visible et dans le proche infrarouge, 6 bandes à 20m dans le proche infrarouge et l'infrarouge court ou *Short-Wave Infrared* (SWIR) et enfin 3 bandes de corrections atmosphériques à 60m de résolution (Figure 4).

Sentinel-2 a pour le moment été de manière générale encore peu utilisé pour des suivis écologiques en milieu tropical bien que certaines études aient été quand même menées notamment sur la texture des forêts tropicales (Laurin et al., 2016).

Le choix des bandes de Sentinel-2 utilisées pour la classification forestière varie d'un auteur à l'autre. Erinjery et al. (2018) ont travaillé sur les bandes 2 (490 nm),

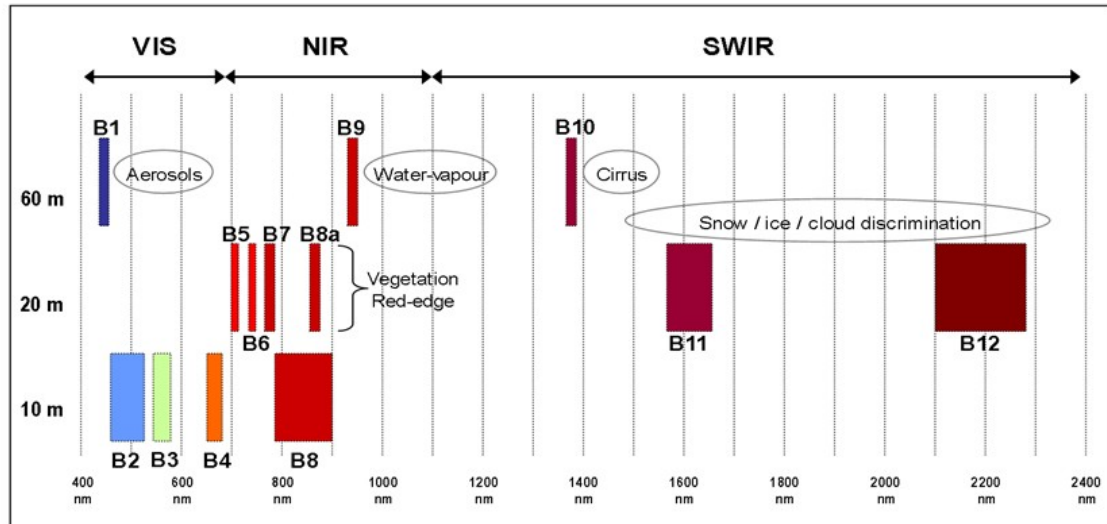


FIGURE 4 – Sentinel-2 : Résolutions spatiales des différentes bandes spectrales

3 (560 nm), 4 (665 nm) et 8 (842 nm). Immitzer et al. (2016) montrent quant à eux que dans le cas d'une classification par pixel, la bande Red Edge 1 (bande 5) est celle qui apporte le plus d'information. Suivent par ordre d'importance, les bandes SWIR (bandes 11 et 12) et la Red Edge 2 (bande 6). Ils obtiennent de manière plus surprenante un moindre apport d'information pour les bandes proche-infrarouge (NIR 1 et NIR 2). Ils supposent ces derniers résultats comme liés principalement à leur zone d'étude et à la période d'acquisition des images.

Les 3 bandes Red-Edge (bandes 5, 6 et 7) sont particulièrement utiles pour de la discrimination basée sur des taux de chlorophylle différents entraînant des différences de densité au niveau de la canopée. Ce sont d'ailleurs ces bandes, ainsi que la bande 3 (verte), qui fournissent les meilleures informations pour la différenciation des types forestiers (Laurin et al., 2016, Alvarez-Añorve et al., 2008)

1.2.4 Indices de végétation & Texture

Différents indices de végétation permettent de faire ressortir des caractéristiques de la végétation et donc d'aider à la classification de celle-ci. L'indice le plus courant et facile à obtenir de façon empirique est le *Normalized Difference Vegetation Index* (NDVI). Verhegghen and Defourny (2011) se basent d'ailleurs sur cet indice pour distinguer les différentes classes de végétation. Le NDVI est basé sur l'idée que seule une petite partie de la réflectance obtenue dans la partie rouge du spectre (R) provient de la végétation foliaire. En effet la majorité de ces longueurs d'onde sont absorbées par les pigments pour la photosynthèse. Au contraire, la majeure part du signal obtenu au niveau du proche-infrarouge (NIR) provient de la réflectance de ces surfaces foliaires. Le NDVI est alors calculé de cette manière, avec ρ = la réflectance (Équation 1) :

$$NDVI = \frac{\rho_{NIR} - \rho_R}{\rho_{NIR} + \rho_R} \quad (1)$$

Un autre indice très répandu est le *Leaf Area Index* (LAI). Ce dernier est un indice qui quantifie la surface foliaire par unité de surface de sol et est influencé par les quantités d'eau, de carbone et d'échange d'énergie entre la végétation et l'atmo-

sphère (Wulder et al., 1996). De plus, la distribution du LAI est utile pour pouvoir déterminer la structure de la forêt. Il faut cependant être attentif à son utilisation, car en forêt tropicale les valeurs de LAI sont généralement hautes, comprises entre 4 et 6 ce qui peut entraîner un phénomène de saturation (de Wasseige and Defourny, 2002).

La relation entre le NDVI et le LAI n'est pas toujours simple et diffère par exemple s'il s'agit de forêt feuillue ou de forêt résineuse (Wulder et al., 1996).

Un autre indice de végétation qui permet de quantifier la biomasse verte est le *Enhanced Vegetation Index* (EVI). Gond et al. (2013) l'utilisent pour traiter des images qui proviennent du capteur MODIS. Le EVI semble plus adapté que le NDVI à l'étude de forêt tropicale composée de végétation dense, car il a une tendance moindre à saturer lorsque les quantités de biomasses vertes sont élevées.

Pour discriminer les zones forestières des zones non forestières, Gond et al. (2013) ont calculé le *Shortwave Infrared Water Stress Index* (SIWSI). Ce dernier caractérise le stress hydrique et permet de différencier une zone forestière, où la canopée garde un niveau hydrique assez stable durant toute l'année, par rapport à des zones non forestières soumises à de grandes variations hydriques entre les saisons sèches et les saisons humides. Dans le cas de Gond et al. (2013), à la fois EVI et SIWSI sont basés sur la bande de réflectance minimum du bleu, soit celle qui est la plus sensible aux effets atmosphériques.

La texture représente la variation de la structure, la distribution spatiale de la végétation (horizontale comme verticale) et se traduit sur une image comme une variation de tons, de nuances entre différents pixels (Franklin et al., 2000, Wulder et al., 1996). Le ton est la variation d'ombre qui existe entre les pixels (Haralick et al., 1973).

L'utilisation d'information concernant la texture lors de classification permet généralement d'obtenir de meilleurs résultats, mais cela est source de discussion dans la communauté scientifique (Laurin et al., 2016, Stumpf and Kerle, 2011). Franklin et al. (2000), dans le cas d'une étude sur les forêts canadiennes à partir d'images aéroportées, obtiennent par exemple une amélioration de leur exactitude de 17% rien que par l'ajout de la texture dans la classification. A contrario, Immitzer et al. (2016), pour une étude sur les forêts d'Europe centrale à l'aide de Sentinel-2, n'obtiennent quant à eux pas d'amélioration de leur classification avec l'ajout de texture.

La taille de la fenêtre sur laquelle la texture est mesurée influence les résultats et par conséquent doit être choisie avec attention. Une large fenêtre de pixel (9x9) a l'avantage de diminuer les erreurs aléatoires et de fournir une meilleure estimation de la distribution par rapport à une fenêtre plus petite (3x3). Néanmoins, le risque d'englober différents types de végétation avec une fenêtre plus grande est supérieur (Franklin et al., 2000, Wulder et al., 1996).

De plus, une liste des sources d'erreurs lors d'ajout de texture dans une classification est donnée par Franklin et al. (2000) :

1. Le choix de la texture en elle-même
2. La surface des fenêtres de texture
3. La manière dont la mesure de texture est utilisée dans l'analyse

4. Le nombre et le type de classes utilisées
5. Les caractéristiques propres de l'image liées à l'acquisition
6. Le but de la classification et son niveau de précision

1.3 Classification

La complexité de la classification sur base de l'imagerie satellite est que des réponses spectrales similaires ne correspondent pas forcément à des réalités similaires et inversement. Les conditions d'exposition, d'échelle ou encore d'agrégation peuvent interférer (Belgiu and Drăguț, 2016). Cela implique que des méthodes simples de classification ne sont pas suffisantes.

Deux grandes méthodes de classification existent : les classifications non supervisées et celles supervisées. Dans une classification supervisée, le nombre de classes est fixé au départ alors que dans le cas d'une classification non supervisée, ce nombre va être ajusté automatiquement en fonction des paramètres d'entrée.

1.3.1 Classification non supervisée

Les classifications non supervisées sont des techniques de classification dites d'apprentissage non supervisé. À partir d'objet ou de pixel non identifiés, c'est à dire dont la classe n'est pas connue, l'algorithme de classification va former des groupes, ou *cluster* en fonction de leurs propriétés.

Cette technique se distingue des classifications supervisées par le fait qu'aucune zone, d'occupation du sol connue, ne lui est fournie au préalable. Elle va donc déterminer elle-même les critères qui servent à l'établissement de catégories.

Un des algorithmes les plus utilisés pour ce type de classification est *l'Iterative Self-Organizing Data Analysis Technique* (ISODATA). Lors de l'utilisation de ISODATA, le nombre de classes est fixé dès le départ, mais aucune information sur celles-ci n'est connue. Chaque pixel va être associé à une classe sur base de sa distance euclidienne minimum avec le centroïde, le centre, de chacun des *clusters*. C'est une méthode itérative donc répétée un certain nombre de fois (Irvin et al., 1997).

Cet algorithme est utilisé régulièrement pour des classifications basées sur des images multispectrales notamment dans le cadre de cartographie d'occupation du sol sur base d'images Landsat (Rozenstein and Karnieli, 2011, Arababah and Alhamad, 2006). Ces deux études fixent, pour des classifications avec entre 6 et 8 classes finales, le nombre de classes initiales à 80.

Les chercheurs comparent ensuite la qualité de classifications obtenue via ISODATA et une classification supervisée basée sur le *Maximum Likelihood Classification* (MLC). Les résultats des deux études divergent alors. Rozenstein and Karnieli (2011) obtiennent des résultats meilleurs pour la classification non supervisée alors qu'à l'inverse, une classification supervisée se montre plus efficace chez Arababah and Alhamad (2006).

Ces différences s'expliquent principalement par la zone d'étude et par la qualité des aires d'entraînement. La méthode ISODATA va par exemple avoir du mal à différencier deux types d'éléments proches. L'ajout de données de terrain, via les

aires d'entraînement, augmente alors la différenciation des deux classes. Cependant, dans le cas où ces aires contiennent des contaminations ou sont intrinsèquement hétérogènes en termes de composition, par exemple des zones de taches d'arbustes, l'effet est inverse. Une solution est de mixer les deux types de classification et d'utiliser un algorithme hybride reprenant à la fois de la classification supervisée et non supervisée (Rozenstein and Karnieli, 2011, Alrababah and Alhamad, 2006).

1.3.2 Classification supervisée

Aires d'entraînement et de validation Dans le cas de classification supervisée et afin de développer un modèle qui permette de définir une classe pour chaque point de la zone d'étude, des zones dont l'occupation du sol est connue sont nécessaires. L'algorithme va pouvoir se baser sur ces dernières pour définir les critères qui vont permettre d'attribuer un point à une classe ou à une autre.

Ces zones doivent cependant répondre à 3 critères principaux (Belgiu and Drăguț, 2016).

- Les aires d'entraînement et de validation doivent être indépendantes entre elles
- Le nombre et la taille des aires d'entraînements doivent être proportionnels dans les différentes classes
- Les aires d'entraînement doivent être représentatives des classes qu'elles sont censées représenter

Généralement, 2/3 des échantillons servent au développement du modèle tandis que le dernier tiers est utilisé pour donner une estimation de l'exactitude du produit final via un procédé de validation (Belgiu and Drăguț, 2016). Il est généralement plus intéressant de séparer les 2 jeux de données lorsqu'ils sont à l'état de polygone. Cela permet d'éviter une redondance de l'utilisation de mêmes pixels pour l'entraînement et la validation (Pelletier et al., 2016).

Colditz (2015) montre en outre l'importance d'avoir des zones d'entraînement proportionnelles à la réalité du terrain. En effet, les résultats de classification sont meilleurs lorsque les classes qui occupent le plus de surface sont bien représentées au sein des aires d'entraînements. Cela doit cependant être nuancé en fonction du but de la classification. Si l'objectif est de montrer au mieux les zones minoritaires en termes de surface, l'accent est à mettre sur ces zones au détriment de la classe majoritaire.

Le nombre de zones d'entraînement et de validation n'est pas un critère majeur. Ceci dit, à large échelle, une augmentation du nombre d'aires d'entraînements accroît les performances de classification lors de l'utilisation de l'algorithme *Random Forest* (RF) (Pelletier et al., 2016).

Algorithme de classification Le choix de l'algorithme de classification est évidemment important. Celui-ci doit pouvoir répondre à plusieurs critères : être capable de gérer des observations comportant une certaine quantité de bruit, des variables spatiales complexes et un nombre d'échantillons relativement faible comparé à la surface en question (Rodriguez-Galiano et al., 2012). Les classifications supervisées sont généralement plus employées, car elles sont capables, sur base de zones d'entraî-

nement prédéfinies, de ressortir les caractéristiques d'intérêt et de les utiliser pour prédire la classe de toute la zone.

L'algorithme RF a été développé par Breiman (2001) et se base sur la création d'arbres de décision multiples. La classification par cet algorithme fonctionne en deux temps. Durant la phase d'entraînement, chaque arbre de décision sera créé à partir d'un jeu de variables d'entrée choisi aléatoirement. À chaque nœud, c'est ensuite l'une de ces variables qui va intervenir. Durant la phase de classification, chaque arbre de décision va classer la même donnée et donner un vote quant à la classification. La classe qui reçoit la majorité des votes va être la classe associée à la donnée (Pelletier et al., 2016).

Deux paramètres doivent être pris avec une considération importante : le nombre d'arbres de décision à générer (*Ntree*) et le nombre de variables à sélectionner et tester lors des séparations dans l'arbre de décision (*Mtry*) (Belgiu and Drăguț, 2016). *Ntree* est dans la plupart des études fixé à 500, car l'erreur se stabilise avant ce nombre, mais d'autres études ont utilisé des valeurs allant de 100 à 5000. Un résultat corroboré par Ghosh and Joshi (2014) qui affirme que les résultats obtenus via RF sont plus sensibles à *Mtry* qu'à *Ntree*. *Mtry* est d'ailleurs la plupart du temps fixé comme la racine carrée du nombre de variables d'entrée (Gislason et al., 2006, Belgiu and Drăguț, 2016). Le nombre et la qualité des variables sélectionnées influent directement sur le résultat de la classification, mais aussi sur le temps de calcul (Pelletier et al., 2016).

Par rapport à d'autres méthodes de classification, comme le MLC ou *Support Vector Machine* (SVM), RF ne nécessite pas comme condition de départ une distribution normale des données. De nombreuses études ont été menées sur la comparaison des méthodes de classification via RF ou via MLC ou SVM. Les résultats dépendent d'une étude à l'autre, mais de manière générale les 3 méthodes donnent des qualités de classification similaires. RF est cependant bien plus performant lors d'un travail sur des surfaces plus importantes grâce à son faible temps de calcul comparé aux autres algorithmes (Pelletier et al., 2016).

Méthode via une approche par pixel La méthode basée sur les pixels (*Pixel-based*) va attribuer à chacun des pixels une classe. La précision peut sembler donc accrue puisque le travail se fait à l'échelle de chaque pixel. Néanmoins, Blaschke (2010) estime que le manque de considération par cette méthode des informations géométriques et contextuelles supprime cet avantage et au contraire dessert la précision. Cela se remarque davantage au niveau des images à haute résolution spatiale pour lesquelles la variance spectrale est plus grande, ce qui augmente la variabilité au sein d'une même classe et donc les probabilités d'erreur de classification (Stumpf and Kerle, 2011). De plus, puisque chaque pixel est traité individuellement, sans tenir compte des autres pixels qui l'entourent, la classification ne se base que sur sa valeur spectrale propre. Cela donne parfois des résultats avec un effet "poivre et sel", car deux pixels spatialement proches et relativement similaires peuvent être mis dans des classes différentes (Xiong et al., 2017).

Méthode via une approche par objet Les objets correspondent à des entités renvoyant une réflectance similaire. Ce sont donc des regroupements de pixels avec des caractéristiques similaires pour former une seule unité. Ces unités peuvent

correspondre à des peuplements forestiers, à des trouées ou même à un individu selon le choix du niveau d'agrégation. Ces regroupements de pixels peuvent sur faire sur base de critères préétablis, d'une homogénéité spectrale, de caractéristiques de texture ou encore de valeurs d'indice similaires (de Wasseige and Defourny, 2002, Immitzer et al., 2016).

L'avantage majeur de cette méthode est qu'elle permet d'apporter des informations supplémentaires aux objets comme leur taille, leur géométrie, leur forme et leur agencement. De plus elle peut aussi bien fonctionner avec des résolutions plus grossières (pixel de 10 à 30m) qu'avec des hautes résolutions (pixel de 2 à 10m).

Le regroupement de pixels nécessite un découpage de l'image satellite en objets. Ce procédé est appelé la segmentation. La difficulté réside souvent dans le choix du bon niveau d'agrégation. Cela peut demander dans certains cas une connaissance du terrain, contrairement à la méthode basée sur les pixels, car aucune méthode générale pour la segmentation en objets n'existe, celle-ci étant spécifique à chaque cas et nécessitant des ajustements selon les situations. L'évaluation de la qualité de la segmentation et du bon niveau d'agrégation passe par une interprétation visuelle des images. Cela peut nécessiter beaucoup de temps pour des grandes surfaces et peut apporter une part de subjectivité (Xiong et al., 2017).

De plus, ces segmentations peuvent, selon les conditions de l'étude, utiliser de grandes quantités de données et augmenter de ce fait considérablement le temps de calcul.

Les résultats obtenus par cette méthode fournissent en général des informations cartographiques plus cohérentes que par la méthode basée sur les pixels (de Wasseige and Defourny, 2002). Cette méthode permet aussi de diminuer l'effet poivre et sel.

1.3.3 Évaluation des classifications

Pour évaluer une classification, les études se basent généralement sur des matrices de confusions (Tableau 1) et de métriques qui en dérivent. Les métriques les plus couramment utilisées sont "*Overall Accuracy*" (OA), le "*Producer Accuracy*" (PA), le "*User Accuracy*" (UA), le F-score et le coefficient kappa (Immitzer et al., 2016). Tous ne sont cependant pas utilisés dans toutes les études, Pelletier et al. (2016) se basent quasi exclusivement sur le OA pour déterminer la qualité de la classification alors que Rodriguez-Galiano et al. (2012) se basent sur l'indice kappa.

		Prédiction			
		classe 1	classe 2	...	classe m
Référence	classe 1	n_{11}	n_{12}	...	n_{1m}
	classe 2	n_{21}	n_{22}	...	n_{2m}

	classe m	n_{m1}	n_{m2}	...	n_{mm}

TABLE 1 – Exemple d'une matrice de confusion

L'OA donne la proportion d'observations classées correctement. Il est le résultat de la division du nombre de pixels correctement classés (n_{ii}) par le nombre total de pixels (N_{tot}) (Congalton, 1991) (Équation 2).

$$OA = \sum n_{ii} / N_{total} \quad (2)$$

L'indice kappa compare la classification obtenue avec une classification aléatoire, uniquement basée sur la chance. L'indice est compris entre 0 et 1. 0 représente une classification aléatoire et 1 une classification en parfaite adéquation avec les données de terrain. Cependant, le coefficient kappa est fort critiqué par Pontius and Millones (2011) qui considèrent qu'il n'apporte pas d'informations supplémentaires en comparaison à l'OA. Au contraire, il est même difficile à comprendre et à interpréter. Dans ce travail la qualité de la classification selon l'indice kappa sera évaluée comme ceci (Tableau 2) :

Indice kappa	Qualité de la classification
$k \leq 0,4$	Faible
$0,4 < k \leq 0,6$	Satisfaisante
$0,6 < k \leq 0,8$	Bonne
$0,8 < k$	Excellente

TABLE 2 – Évaluation de la qualité d'une classification selon l'indice kappa

Le PA, ou l'exactitude du producteur et le UA, l'exactitude de l'utilisateur, sont tous deux des ratios. Le PA est le nombre de pixels correctement classés pour une catégorie sur le nombre de pixels total pour cette catégorie. C'est une mesure de la précision de chacune des catégories, car elle montre combien de fois un élément présent sur le terrain est correctement classé.

Le UA est plutôt une mesure de l'erreur, car il consiste en la division du nombre de pixels correctement classés pour une catégorie par le nombre de pixels qui ont été classés dans cette catégorie. Il va donc montrer une probabilité qu'un élément classé dans cette catégorie se retrouve en effet bien sur le terrain.

Le F-score ou F_1 -score (Équation 5) prend en compte à la fois la *precision* (Équation 3) et le *recall* (Équation 4). Il varie entre 0 et 1 et au plus le F-score est élevé, meilleur est la classification de cette catégorie.

$$Precision = \frac{n_{ii}}{\sum_{j=1}^n n_{ij}} \quad (3)$$

$$Recall = \frac{n_{jj}}{\sum_{i=1}^n n_{ij}} \quad (4)$$

$$F - Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (5)$$

1.4 Conclusion

Les forêts tropicales constituent un écosystème majeur dans l'équilibre de la planète, de part la biodiversité qu'elles renferment, les activités économiques, culturelles et sociales qu'elles développent et leur contribution à la balance carbone mondiale. Peu de cartographies précises des forêts d'Afrique Centrale cohérentes à large échelle existent. Or mieux connaître la forêt est essentiel pour comprendre et prévenir les changements futurs .

De nombreuses études concernent la réalisation de cartographies d'occupation du sol. Pourtant, peu d'études se focalisent sur la différenciation des types forestiers en milieu tropical. De ce fait, peu d'informations précises existent sur les classes forestières en Afrique Centrale, à la fois précises au niveau typologique et large au niveau spatial.

La disponibilité d'images de haute résolution avec la plateforme Sentinel-2 offre de nouvelles opportunités pour l'étude de végétation et la surveillance de forêts. Les 13 bandes multispectrales, de résolutions spatiales de 10 ou 20 m couplées avec une résolution temporelle de 5 jours, offrent des précisions potentielles bien supérieures à ce qui existait.

A ce jour, peu d'études ont utilisé Sentinel-2 pour de la cartographie de type forestier avec un niveau de détails élevé. La plupart des classifications utilisent un nombre restreint de catégories forestières. Sentinel-2 et ses images à hautes résolutions permettent d'apporter plus de précision et de nuances dans ces catégories forestières.

Les classifications qui semblent se montrer les plus performantes sont des classifications supervisées. Des zones d'entraînement, d'occupation du sol connu sont donc nécessaires. L'approche via des pixels reste majoritairement utilisée dans la littérature bien que l'approche par objet montre un potentiel intéressant dans plusieurs cas.

2 Objectifs

L'objectif premier de ce travail est de continuer à participer au développement et à l'amélioration des méthodes de classification des forêts d'Afrique Centrale à l'aide de Sentinel-2. Dans ce but, plusieurs objectifs secondaires sont nécessaires :

1. Premièrement, établir une typologie forestière qui permette de rendre compte de la variété des milieux présents sur place. Cette typologie doit être transposable à une échelle spatiale plus large. Elle ne doit donc être ni trop grossière, ni trop détaillée.

2. Deuxièmement, la mise en place d'une clé de photo-interprétation réalisée grâce à une campagne de terrain. L'idée est de pouvoir, sur base de compositions colorées de Sentinel-2 discerner et identifier les types forestiers. La campagne sert aussi à l'obtention des données nécessaires à la classification et à la cartographie des APDS.

3. Sur base de cette campagne et de l'expérience acquise, un protocole généralisable à toute cartographie de peuplements forestiers en forêt tropicale sera établi.

4. Plusieurs méthodes de classification seront testées et leurs résultats comparés. Le but est de déterminer le type de classification le plus adapté au contexte de ce travail.

5. Cette méthode sera ensuite utilisée pour cartographier l'ensemble des APDS. Le produit final attendu est une cartographie générale des types forestiers ainsi que de l'occupation du sol des APDS. La qualité de la classification sera aussi évaluée.

6. Enfin, un suivi des surfaces agricoles dans les APDS fera l'objet d'une section.

3 Zone d'étude

3.1 Localisation

La zone d'étude concerne les Aires Protégées de Dzanga Sangha (APDS), situées au sud-ouest de la RCA entre 2°13' et 3°24' de latitude Nord, et 15°30' et 16°35' de longitude Est. Sa superficie totale est d'environ 4000 km².

Créées en 1990 sous l'impulsion du World Wildlife Fund (WWF) et du gouvernement de RCA, les APDS comprennent 2 parcs nationaux : le parc national de Dzanga (495 km²) et celui de Ndoki (725km²) ainsi que la Réserve Spéciale de Forêt Dense de Dzanga-Sangha. Une zone de chasse communautaire relie les deux parcs nationaux (Figure 5). Les APDS font partie d'un projet transfrontalier avec les parcs nationaux contigus de Nouabalé-Ndoki en République du Congo et celui de Lobéké au Cameroun. Ensemble, ils forment le Tri-National de la Sangha (TNS) et couvrent une superficie totale de 7682 km² classés en parc national.

Le but des APDS est la protection de l'écosystème local qui comprend aussi bien les espèces animales que végétales, tout en satisfaisant les besoins des populations locales. Elles mettent en avant différents thèmes allant de la conservation aux études scientifiques, au tourisme, à l'économie ou encore aux valeurs culturelles (APDS, 2016, N'Goran et al., 2016).

Depuis 2012, les APDS sont inscrites au patrimoine mondial pour l'humanité de l'UNESCO pour leurs richesses écologiques et culturelles (UNESCO, 2019). Le village principal de Bayanga est situé le long de la rivière Sangha. Il est le chef-lieu de la préfecture de Sangha-Mbaéré.

3.2 Description physique

3.2.1 Climat

Le climat de la zone d'étude est de type guinéen forestier caractérisé par une double saisonnalité : une saison sèche de décembre à février et une saison des pluies durant les 9 autres mois. Cette dernière est néanmoins entrecoupée d'une autre saison sèche, plus courte, en juin et juillet. L'humidité relative reste élevée toute l'année grâce aux réserves en eau dans le sol. Les précipitations annuelles moyennes sont comprises entre 1400 et 1600 mm. La température moyenne est de 26°C et est très stable tout au long de l'année puisque l'amplitude thermique annuelle est de seulement 1,8°C. Cette faible variabilité ainsi que la longue saison sèche peut avoir des répercussions importantes aux niveaux des végétaux (APDS, 2016, Climate-Data.org, n.d., Gond et al., 2013).

3.2.2 Topographie

L'altitude au sein de la zone d'étude varie entre 340 m et 690 m (Figure 6) (APDS, 2016). Deux grandes zones se distinguent :

- le sud de la réserve est une grande plaine d'altitude moyenne de 350 m.
- le nord est quant à lui plus vallonné et composé de 2 plateaux situés à l'est (Carnot et Gadzi) et à l'ouest (Bilolo), tous deux d'une altitude moyenne de 650 m.

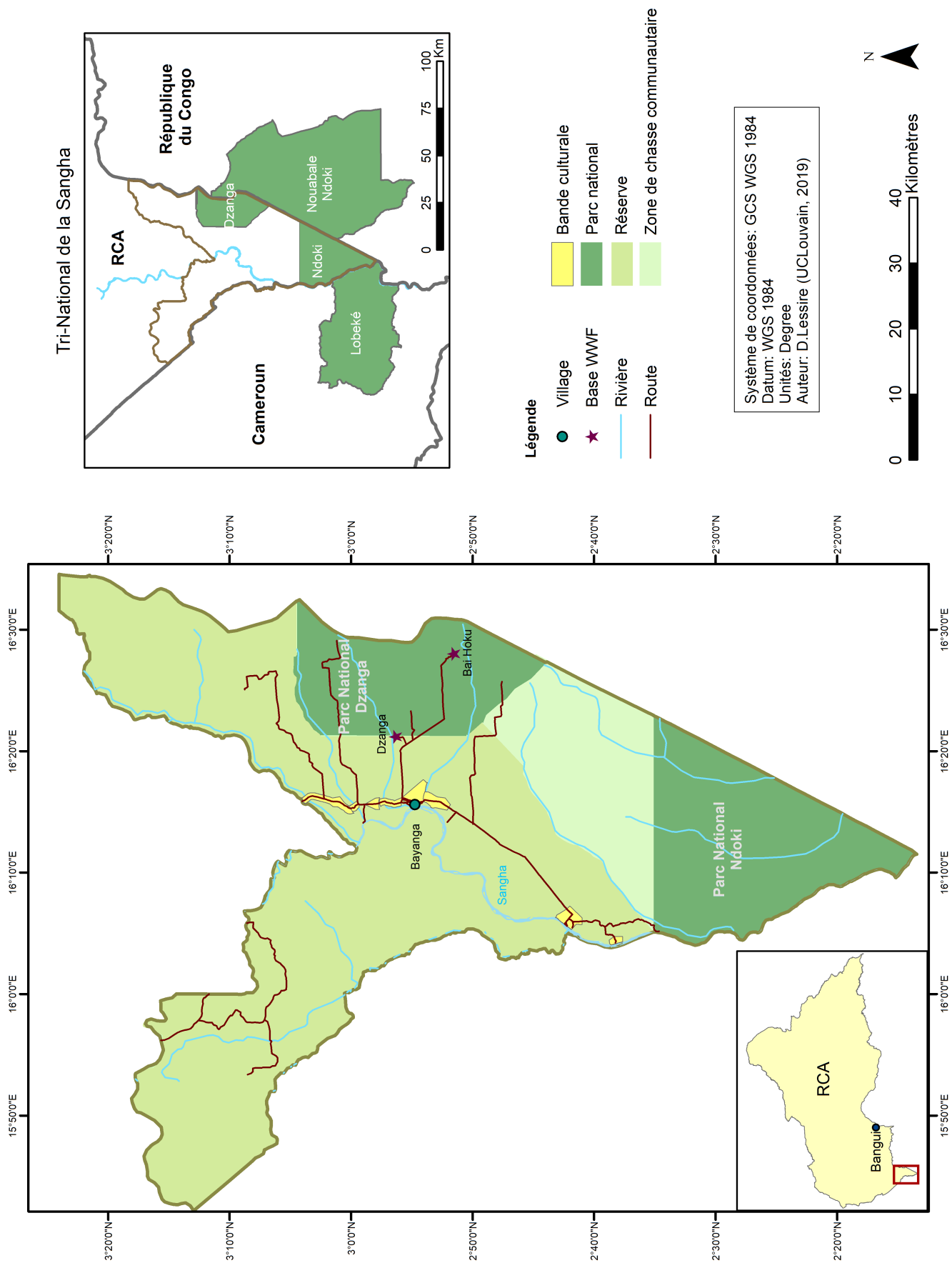


FIGURE 5 – Carte des Aires Protégées de Dzanga Sangha

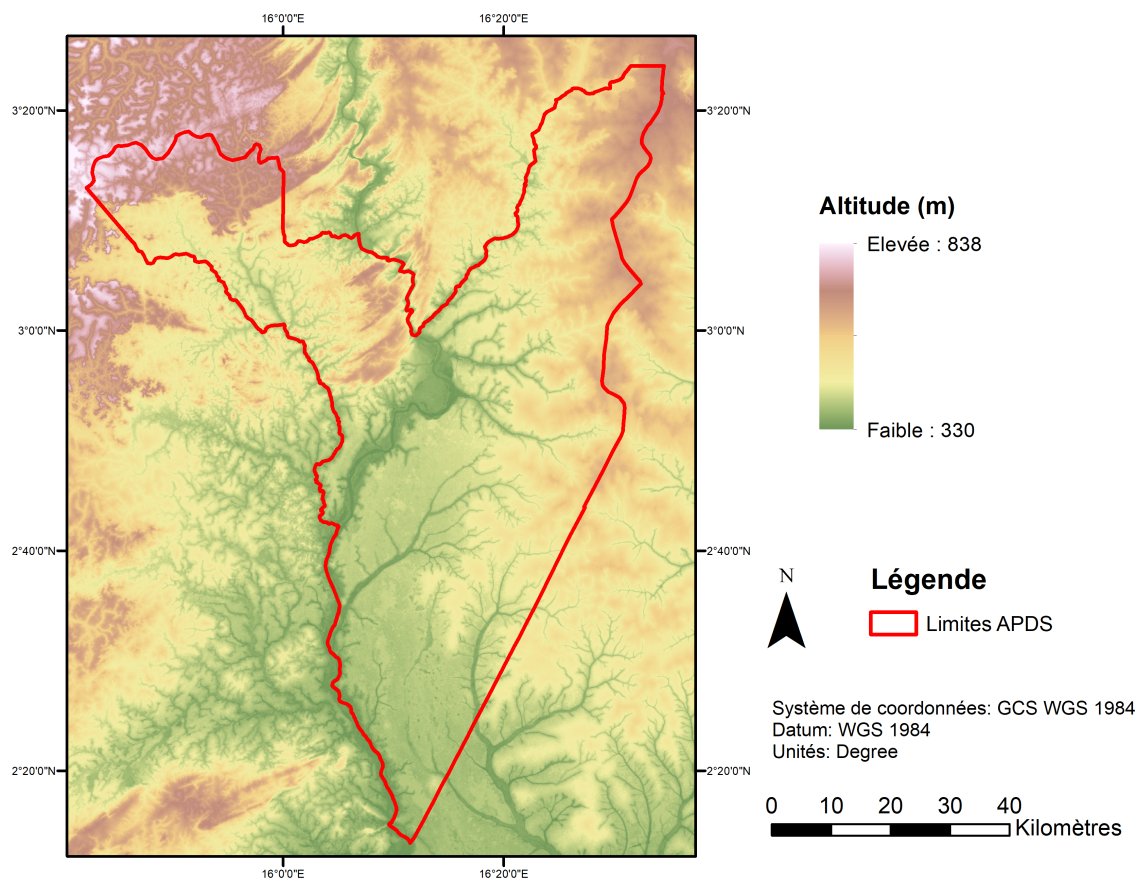


FIGURE 6 – Altitude au sein des APDS

3.2.3 Hydrographie

Le parc est traversé du nord au sud par la rivière Sangha, un affluent du fleuve Congo. Il est le deuxième fleuve le plus important du pays et était utilisé historiquement pour le transport fluvial, notamment des grumes. En période sèche, le niveau descend de 3,5 m par rapport à son niveau de crue occasionnant l'apparition de nombreux bancs de sable. Au sein des APDS, diverses rivières alimentent la Sangha dont Yobé, Massapoula, Babongo, Kényé, Goboumo et Lobéké (APDS, 2016).

3.2.4 Géologie & Pédologie

Les sols des plaines de la Sangha sont des sols ferrallitiques fortement désaturés et mal drainés. Les sols sont à prédominance sableuse et pauvre en argiles. Au nord-ouest de la réserve, se trouvent des sols minéraux bruts d'érosion (Boulvert, 1986, APDS, 2016). Les forêts du sud des APDS se trouvent sur un plateau sableux ce qui leur confère des propriétés particulières par rapport aux forêts des pays limitrophes présentant plus de forêts sempervirentes (Gond et al., 2013).

3.2.5 Typologie de végétation

Actuellement, les cartes existantes de la végétation de la RCA n'offrent pas un niveau de détail très élevé. Elles classent généralement les forêts présentes sur

le territoire centrafricain comme des forêts décidues tout en distinguant celles sur terres fermes (*Terra-Firm*) et les forêts marécageuses (Gond et al., 2013). Pour Ernst et al. (2012), les forêts sèches denses sont typiques de la RCA et sont des reliques des forêts denses humides. Elles sont caractérisées par une strate herbacée éparsée sous un couvert forestier haut de 15 à 20m. Les savanes boisées et les savanes arborées couvrent aussi de larges étendues en RCA.

3.3 Planification agricole

Les APDS sont des aires protégées et par conséquent les activités au sein de la réserve sont limitées. L'agriculture ne fait pas exception et dans un objectif de protection et de gestion du territoire foncier, le gouvernement a mis en place des zones affectées à l'agriculture. Ces zones sont appelées les *bandes culturelles* et sont au nombre de 8 dans l'ensemble des APDS (Figure 8). Elles se retrouvent logiquement aux abords des villages et de la route principale. Ces zones ont été fixées en 1992 et modifiées pour la dernière fois en 2005 (APDS, 2016).

L'agriculture pratiquée est une agriculture extensive sans rotation et sans intrants externes. La production agricole est en très grande majorité destinée à l'autoconsommation (APDS, 2016). Le système pratiqué est celui de l'abattis-brûlis. C'est-à-dire que la forêt est abattue et brûlée pour créer de nouvelles terres agricoles (Figure 7). Sur les terres agricoles situées en bordure de village, proches des zones d'habitation et cultivées depuis longtemps, les cultures poussent moins vite et moins bien. Les rendements sont donc bien inférieurs. Les paysans locaux sont alors en recherche de nouvelles terres en périphérie des zones agricoles existantes, au détriment de la forêt.

Les surfaces agricoles suivent l'augmentation démographique et sont donc en hausse dans la majorité des bandes culturelles. Dans certaines zones, la pression foncière est trop forte et l'extension des surfaces agricoles se fait en dehors des surfaces affectées à l'agriculture, au détriment des surfaces forestières de la réserve.



FIGURE 7 – Défrichement de zones forestières au profit de zones agricoles

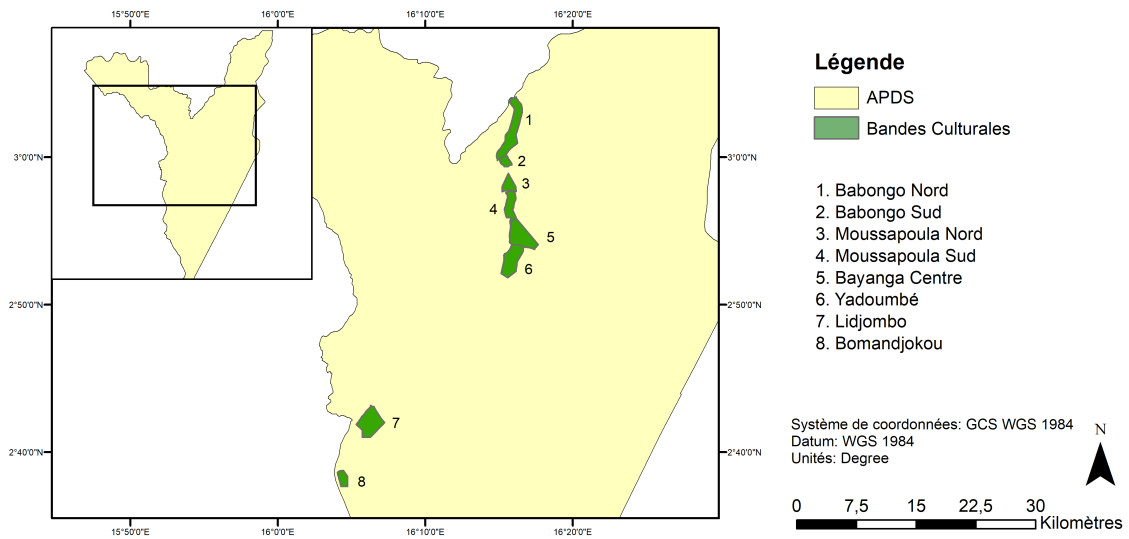


FIGURE 8 – Bandes culturelles au sein des APDS

4 Matériel & Méthode

Cette section présente les données et la méthode utilisées pour la réalisation de ce travail. Le schéma général suivi est représenté à la figure 9.

Les données proviennent de deux sources : celles issues d’images satellites et celles obtenues lors d’une campagne de terrain. La combinaison des deux permet de mettre au point une clé de photo-interprétation. Les zones d’entraînements pour le développement du modèle de classification sont définies sur base de parcelles d’échantillonnages issues de la campagne de terrain et de photo-interprétation de compositions colorées. Quatre classifications sont faites sur une partie restreinte de la zone d’étude, la zone de développement. La qualité de ces classifications est ensuite évaluée et une seule méthode est retenue pour être développée à l’échelle de la zone d’étude complète.

4.1 Images Sentinel-2

Les différentes classifications sont faites sur base des images Sentinel-2. La zone d’étude recouvre 3 tuiles : 33NWD, 33NXD et 33NXC (Figure 10).

Les images sont directement téléchargées à partir site de l’ESA (<https://scihub.copernicus.eu/>). Ces images correspondent à des réflectances de surfaces (niveau L2A). Une correction atmosphérique a donc déjà été apportée et les masques de nuages identifiés.

Le nombre d’images avec une couverture nuageuse faible disponibles pour la zone d’étude et pour l’année 2019 est très limité. En annexe B est présenté la couverture nuageuse pour la dernière année sur une partie de la zone d’étude. Dans cette étude, les images sélectionnées, durant la saison sèche et sans nuages, pour développer et réaliser la classification sont celles acquises aux dates suivantes :

- 27/12/2018
- 12/03/2019

4.1.1 Traitement des images Sentinel-2

Dans le cadre de ce travail, toutes les bandes de Sentinel-2 sont utilisées. L’enjeu est d’identifier les combinaisons qui permettent d’identifier au mieux les différents massifs et types forestiers. Des combinaisons telles que celle présentée à la figure 11 sont réalisées.

Afin d’obtenir une résolution commune à toutes les bandes spectrales, un ré-échantillonnage est appliqué à chacune des bandes. La résolution finale est de 10m. Les masques nuageux sont aussi identifiés par l’algorithme *Sens2Cor*.

4.2 Campagne de terrain

Dans le but de comprendre et décrire les différents types forestiers existant dans le sud-ouest de RCA, une collecte de terrain a été réalisée durant 7 semaines, en février et mars 2019. Le choix de la localisation est basé sur la diversité des types de forêts, mais aussi les infrastructures existantes. Une collaboration a été mise en place avec le WWF RCA pour permettre la collecte des données. Les équipes du

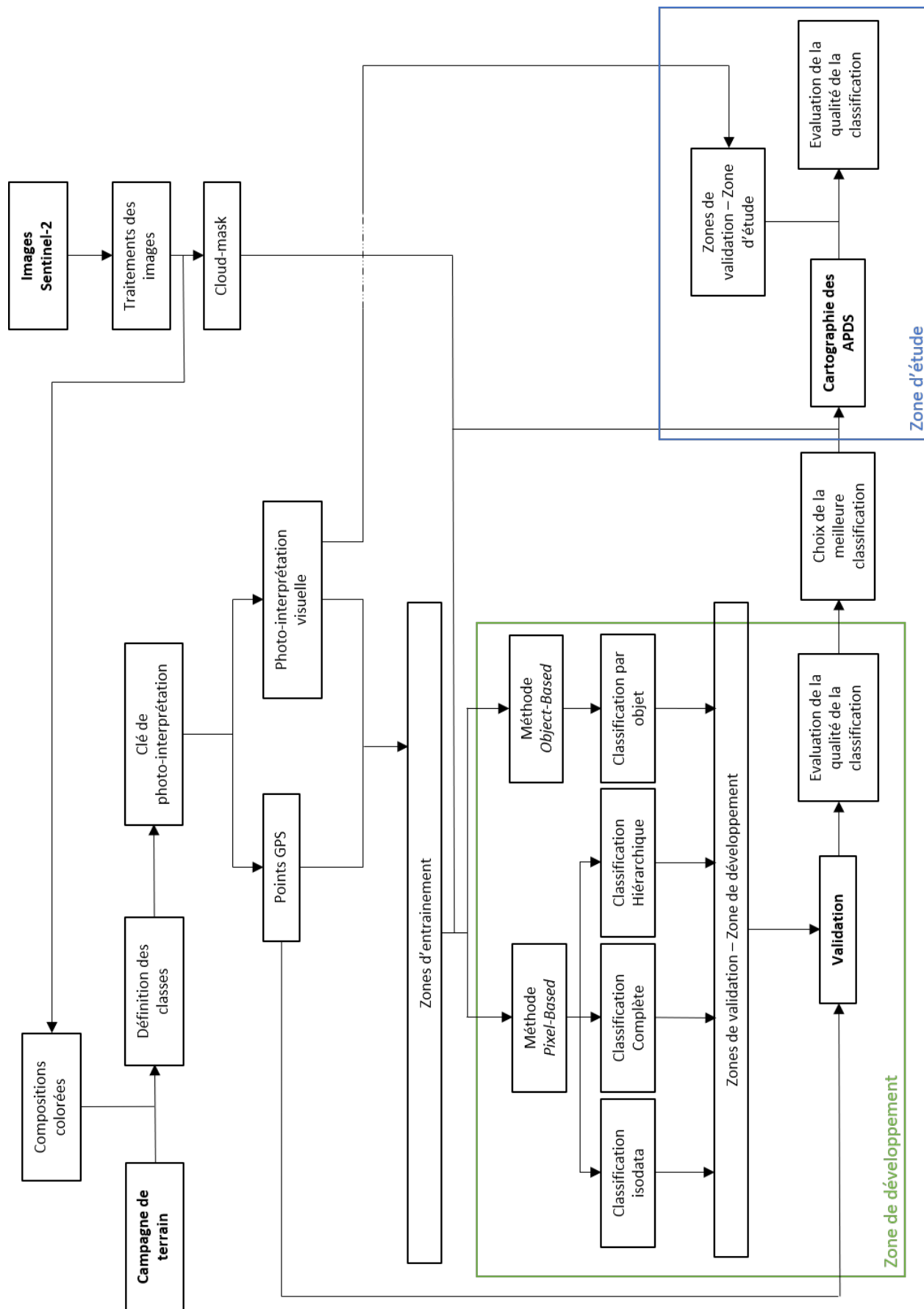


FIGURE 9 – Workflow complet de la méthode générale

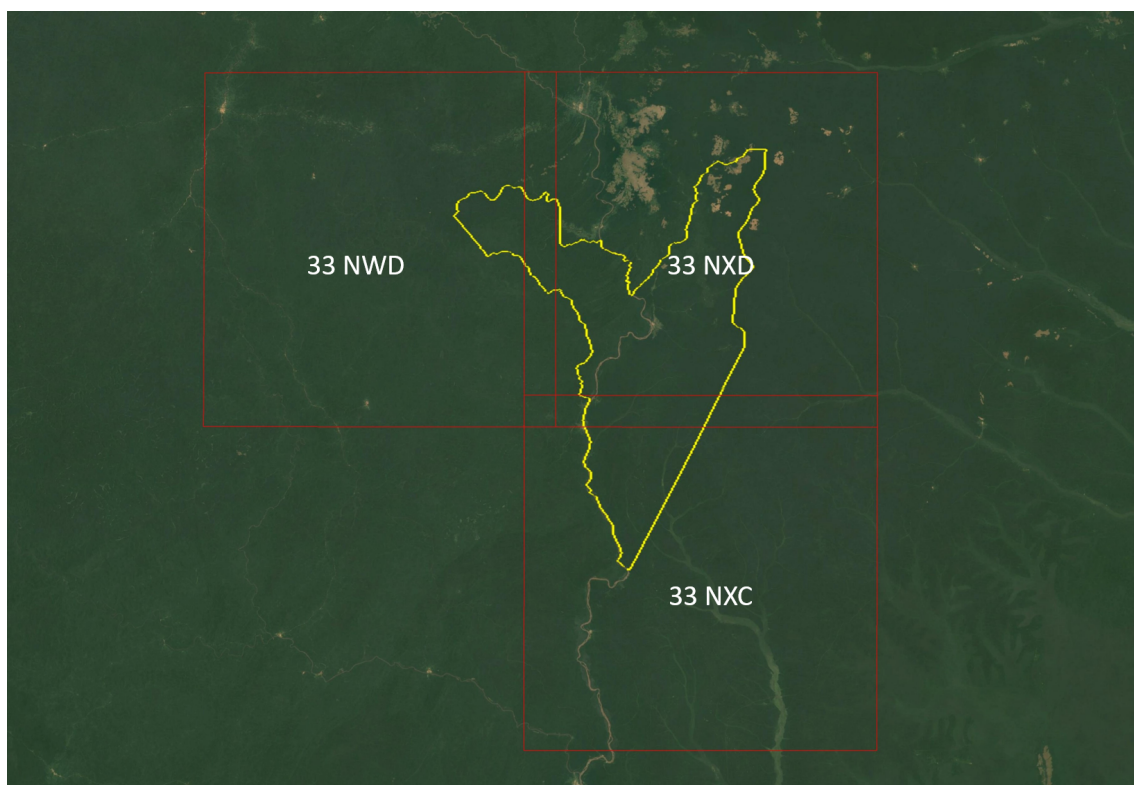


FIGURE 10 – Tuiles de Sentinel-2 recouvrant les APDS

WWF qui ont assuré l'accueil ont de plus fourni les accès, les autorisations locales et diverses aides ainsi qu'un logement.

4.2.1 Protocole de terrain

Afin d'estimer la diversité de la végétation sur place, une première approche a été menée à partir de diverses compositions colorées Sentinel-2, comme celle présentée à la figure 11. Cela permet de mettre en évidence différents massifs forestiers de par leur réflectance, leurs textures et leur morphologie. Les informations disponibles dans la littérature sont aussi utiles afin d'identifier les critères déterminants de la végétation et les types forestiers typiques de la zone d'étude. Une première liste de la diversité a ainsi été créée a priori. Elle ne met pas obligatoirement en relation les types de forêts avec leur image spectrale, mais permet d'avoir une base de recherche et d'établir un plan d'action une fois sur place.

Sur le terrain, et sur base des compositions colorées, les coordonnées GPS de points dans ces différents types sont pris afin de retrouver la zone d'intérêt et de mettre en place des placettes d'échantillonnages.

La répartition de ces placettes au sein de la zone d'étude est limitée par les déplacements possibles. La majorité de ceux-ci ont été faits à pied ou via des opportunités (voiture, moto, etc.). Cela limite donc l'échantillonnage à des zones proches de Bayanga. Deux séjours ont cependant été réalisés dans des camps du WWF en forêt (*Baï Hoku* et *Dzanga*). L'avantage est de permettre de couvrir de nouvelles zones avec des faciès de végétations différentes.

Une fois sur place, des placettes d'échantillonnage sont mises en place. Chaque placette doit avoir une taille minimum de 50 x 50 m pour tenir compte de la ré-

solution spatiale de Sentinel-2 qui est de 20m pour les bandes spectrales SWIR et Red-Edge. Ainsi l'entièreté du signal du pixel est comprise dans la placette. Chaque parcelle doit aussi être homogène à un maximum d'égards (végétation, hydromorphie, pente, etc).

Dans chacune des parcelles, une série de caractéristiques est relevée :

- Sa position GPS est prise à l'aide du Garmin GPSMAP 62. Ces points sont ensuite reportés sur les données satellites via un logiciel d'information géographique (SIG).
- Ses caractéristiques physiques : la pente, l'exposition, le type de sol et enfin l'hydromorphie. Cette dernière n'est pas toujours observable (par exemple dans le cas des forêts inondées seulement une partie de l'année), les informations proviennent alors des discussions avec le guide local.
- Les espèces végétales présentes. Il ne s'agit pas de faire un relevé phytosociologique exhaustif, mais une manière de détecter la présence d'une essence potentiellement indicatrice. Les espèces ont été identifiées par le guide local puis traduites depuis le dialecte autochtone vers le nom scientifique.
- À travers une description de chaque strate, la structure du peuplement est décrite. La hauteur et l'ouverture de chaque strate est estimée, la présence de dominants, de trouées ou encore de chablis est aussi relevée.

Sept classes de forêts en sont ressorties, leurs descriptions et caractéristiques principales sont reprises à la section 5.1. Au total, 180 placettes ont été relevées dans le milieu forestier réparties équitablement dans chacune des classes.

Dans un deuxième temps, une campagne similaire a été réalisée sur les zones non-forestières. Le protocole est semblable à celui pour les milieux forestiers avec 97 points répartis en 5 catégories (clairière saline, zone d'habitation, eau, savane et zone agricole).

Une autre campagne de terrain, similaire dans le protocole, est réalisée simultanément par Baptiste Delhez dans le cadre d'un mémoire portant aussi sur la cartographie de forêt au sein des APDS (Delhez, 2019). Les données relevées dans ce cadre servent d'échantillons de validation.

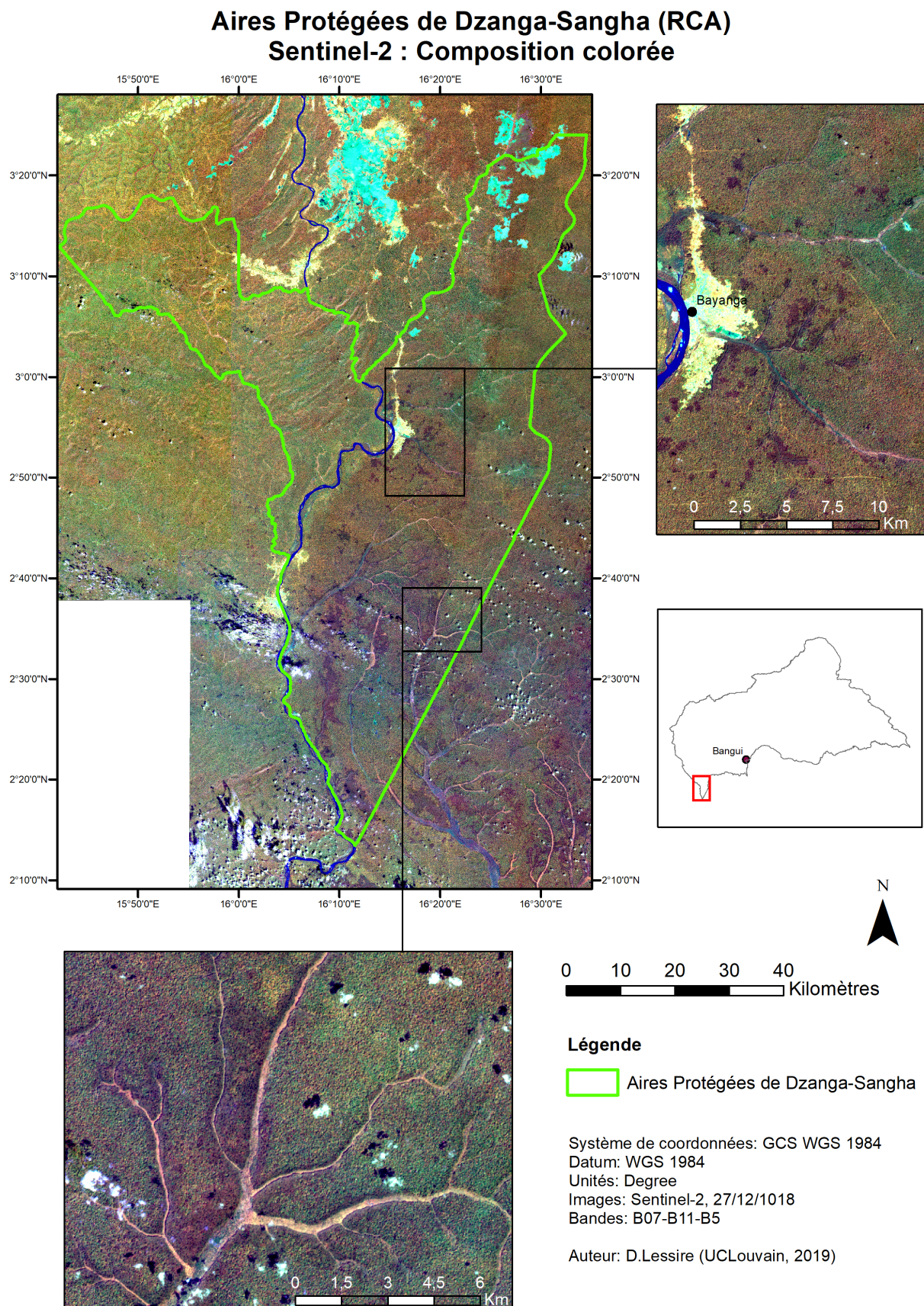


FIGURE 11 – Composition colorée d’images Sentinel-2 sur la zone d’étude, 27/12/2018, Bandes : B07-B11-B05

4.2.2 Zones d'entraînement

Pour chacun des points GPS correspondant à une parcelle d'échantillonnage, la classe de végétation observée lui est associée. Autour de chacun de ces points est dessiné un polygone homogène du point de vue de la végétation. La taille de celui-ci varie, il sera moins grand dans une forêt ripicole que dans une forêt secondaire avec de plus grandes surfaces homogènes. Ces polygones sont tracés sur base des images Sentinel-2 et des données relevées sur le terrain.

Les zones dont l'occupation du sol est connue sont divisées en 2 catégories : la première sert à développer le modèle de classification. Ces zones sont appelées les aires d'entraînement. La seconde va évaluer l'exactitude de la classification et comprend les aires de validation. La répartition des zones dans ces deux catégories est faite de manière aléatoire.

Afin d'augmenter le nombre de zones d'entraînement, 127 nouvelles zones sont délimitées à partir de photo-interprétation réparties dans les différentes classes. À partir des différentes compositions colorées et sur bases de données préexistantes, des zones dont l'occupation du sol peut être identifiée sont aussi désignées comme zones d'entraînement. A noter cependant que lors de la campagne de terrain aucun point n'a été pris dans la classe "eau". Dès lors les zones d'entraînement et de validation pour cette classe, se font uniquement sur base de photo-interprétation.

Une vue d'ensemble des zones d'entraînement ou *Region of Interest* (ROI) est disponible au tableau 3 et leur répartition spatiale est présentée à la figure 12. En accord avec la littérature, les zones d'entraînement ont été réalisées de manière à respecter autant que possible les proportions des classes de végétation sur le terrain. Ces proportions ne sont pas connues avec exactitude a priori, mais se basent sur des estimations visuelles sur base des compositions colorées. La définition des zones d'entraînement est aussi liée à la campagne de terrain et aux zones visitées.

Près de 40% des pixels utilisés comme entraînement sont des pixels représentatifs des forêts secondaires. Cela s'explique par la présence majoritaire en termes de surface, de cette classe dans la zone échantillonnée. De plus, ces forêts sont homogènes et couvrent de larges surfaces. Les zones sont donc généralement de grandes dimensions. Le phénomène est similaire pour les forêts denses humides semi-décidues avec une représentativité de 18% dans les pixels d'entraînement. A contrario, les forêts ripicoles hautes et les zones d'habitations ne représentent chacune que 0,6% du jeu d'entraînement. Cela s'explique par leur faible représentativité en termes de surface dans la zone d'étude. Néanmoins, un nombre similaire d'échantillons ont été réalisés pour ces deux classes en comparaison aux autres catégories. Seulement, les zones de ces deux classes ne contiennent qu'un faible nombre pixels à chaque fois. Et cela du fait que ces types de végétation ou d'utilisation du sol ne couvrent que de faibles surfaces.

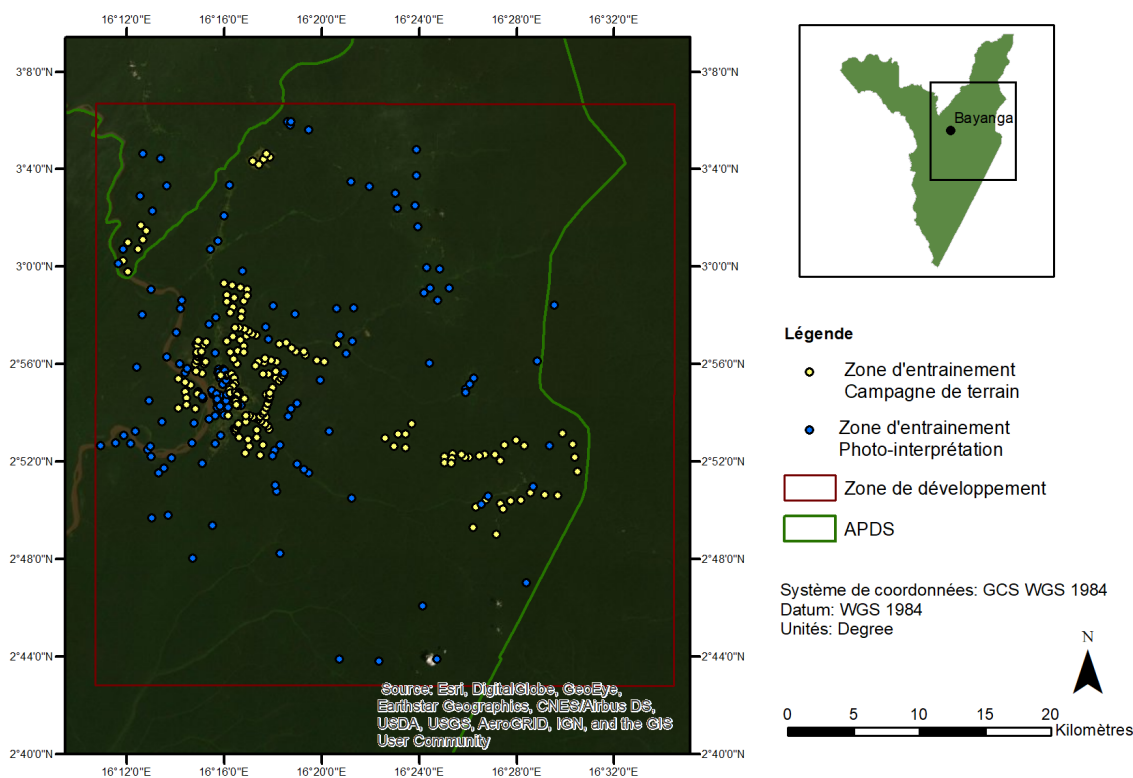


FIGURE 12 – Répartition des zones d’entraînements

	Nombre ROI	ROI Terrain	ROI Photo-interprétation	Nombre Pixel	% en pixel
F sempervirente	31	15	16	3925	4,7
F dense humide	44	32	12	15025	18,2
F marécageuse	34	28	6	4062	4,9
F secondaire vieille	37	27	10	36181	43,8
F secondaire jeune	30	26	4	6748	8,2
F ripicole haute	21	16	5	474	0,6
F inondable	21	11	10	4898	5,9
Clairière saline	14	3	11	1658	2,0
Zone habitation	20	0	20	462	0,6
Eau	10	0	10	4540	5,5
Savane	8	5	3	2401	2,9
Zone agricole	45	25	20	2297	2,8

TABLE 3 – Détail des aires d’entraînement

4.3 Comparaison des méthodes de classification

L'objectif poursuivi est de comparer 4 méthodes de classification¹ : une non supervisée par un algorithme ISODATA et 3 supervisées par l'algorithme RF. Ces trois dernières se font via une approche soit par des pixels, soit pas des objets. La comparaison des classifications se fait sur une zone plus petite que la zone d'étude, appelée zone de développement.

4.3.1 Choix de la zone de développement

Pour réduire le temps de calcul et permettre plus de tests, la classification est d'abord réalisée non pas sur l'ensemble de la zone d'études mais sur une partie seulement. Cette zone est appelée la zone de développement et correspond à une surface de 45 km x 45 km (Figure 13). Le choix de la délimitation de cette zone vient directement de la campagne de terrain puisqu'elle correspond à la surface parcourue et échantillonnée. Elle comprend donc l'ensemble des placettes.

De plus, cette zone est relativement bien connue, car parcourue durant la campagne de terrain. Cela permet donc une évaluation des résultats plus facile, notamment grâce à l'interprétation visuelle.

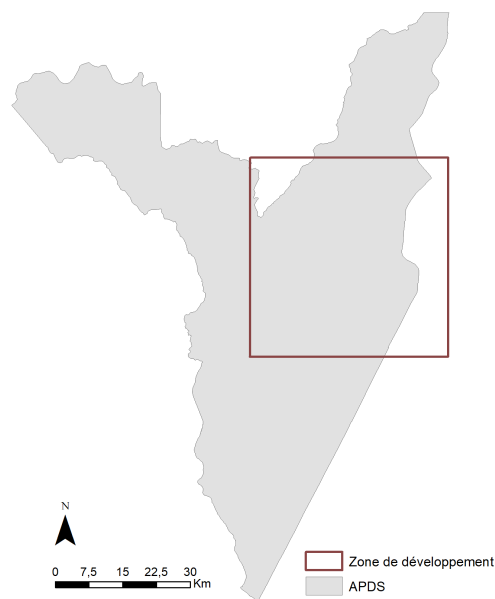


FIGURE 13 – Zone de développement

4.3.2 Description des méthodes de classification

Classification non supervisée - Isodata La première approche consiste à appliquer un algorithme de classification non supervisé sur les images Sentinel-2. L'avantage est que cela ne nécessite pas de connaissance à priori du terrain puisqu'aucune zone d'entraînement n'est nécessaire. L'algorithme fonctionne sur base d'un *clustering*. À partir de *clusters* locaux sélectionnés aléatoirement, chaque pixel va être associé à un de ces groupes. Les différents *clusters* peuvent être regroupés ou séparés en fonction de leur ressemblance/différence. La moyenne et la déviation standard sont recalculées à chaque itération. Les différentes classes créées sont ensuite identifiées et rassemblées par photo-interprétation. Le nombre de classes fixé est de 50 en accord avec la littérature sur le sujet.

Classification supervisée - Approche par pixel La deuxième approche est une classification supervisée réalisée à l'échelle des pixels. C'est une des techniques les plus utilisées dans la littérature. La classification est faite aussi ici via une approche

1. Toutes les classifications sont réalisées à l'aide de *Orfeo Toolbox* (OTB) développé par le CNES

par pixel, mais avec l'algorithme RF. Les avantages de cette technique comparée aux autres classifications supervisées sont sa simplicité et son gain de temps.

La classification est dite supervisée, car des zones dont l'occupation du sol est connue servent à définir les critères qui permettent de distinguer une classe d'une autre. Ces zones sont appelées les "zones d'entraînements". Une description de l'établissement de ces zones est disponible à la section 4.2.2.

Classification supervisée - Approche hiérarchique par pixel La troisième approche consiste aussi en une classification supervisée basée sur les pixels. Comme la précédente, l'algorithme utilisé est ici aussi RF. Toutefois, contrairement à la classification précédente, l'ensemble des classes ne sont pas introduites dès le départ. La classification se fait donc en plusieurs temps.

Tout d'abord, 4 classes principales sont identifiées : les forêts denses humides semi-décidues, les forêts secondaires, les forêts liées aux conditions hydriques et enfin les zones non forestières. Sur base de ces 4 groupes, une première classification est réalisée. Ensuite, chaque classe est affinée séparément et des classifications secondaires sont réalisées jusqu'à obtenir le nombre de classes final, soit 12 dans ce cas-ci. La répartition des classes dans chaque catégorie principale est représentée à la figure 14.

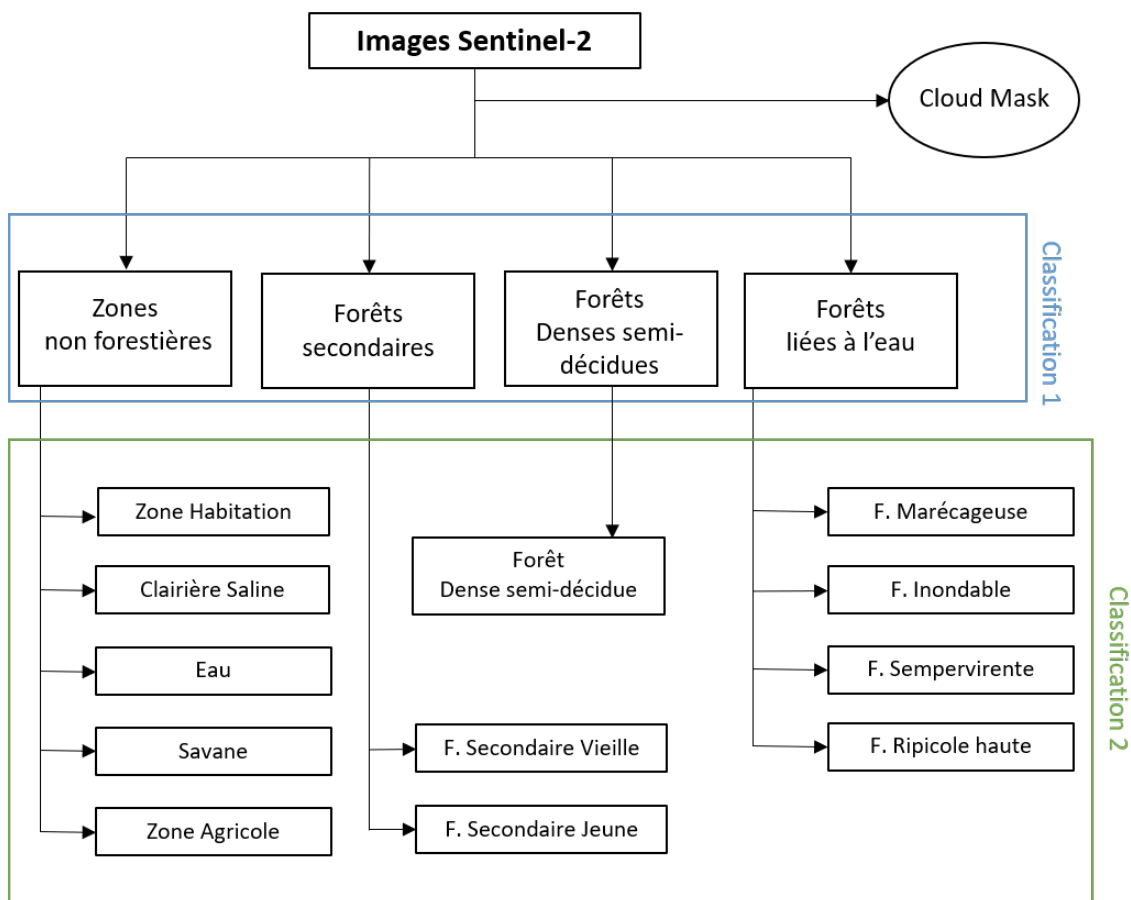


FIGURE 14 – Méthode de classification hiérarchique

Classification supervisée - Approche par objet La quatrième méthode de classification testée est une classification supervisée basée sur les objets. Pour cela, il convient tout d'abord de définir ceux-ci via une segmentation². Pour cela plusieurs paramètres sont fixés préalablement. Tout d'abord les bandes utilisées et les poids associés à chacune d'elles. Ici la segmentation est réalisée sur les 10 bandes de Sentinel-2 avec la même importance accordée à chacune d'elles. Le paramètre d'échelle (*scale*), le paramètre de forme (*shape*) et enfin le paramètre de compacité (*compactness*) sont fixés comme défini dans le tableau 4. Ceux-ci ont été identifiés par essai-erreur et sont ceux qui donnent les meilleurs résultats (Figure 15).

Bandes spectrales	10 bandes de valeurs égales
Échelle	65
Forme	0,67
Compacité	1

TABLE 4 – Paramètres de segmentation

Une fois la segmentation réalisée, plusieurs métriques liées à chaque objet sont calculées. Tout d'abord, 14 paramètres de textures dérivées de *Gray-Level Co-Occurrence Matrix* (GLCM) définie par Haralick et al. (1973) sont considérés comme pertinents lors d'analyse d'images. Ensuite, la moyenne et la déviation standard de chacune des bandes ainsi que le moment angulaire, le contraste, la variance, l'homogénéité et la corrélation. La classification est ensuite réalisée par l'algorithme RF sur bases des métriques ci-dessus.

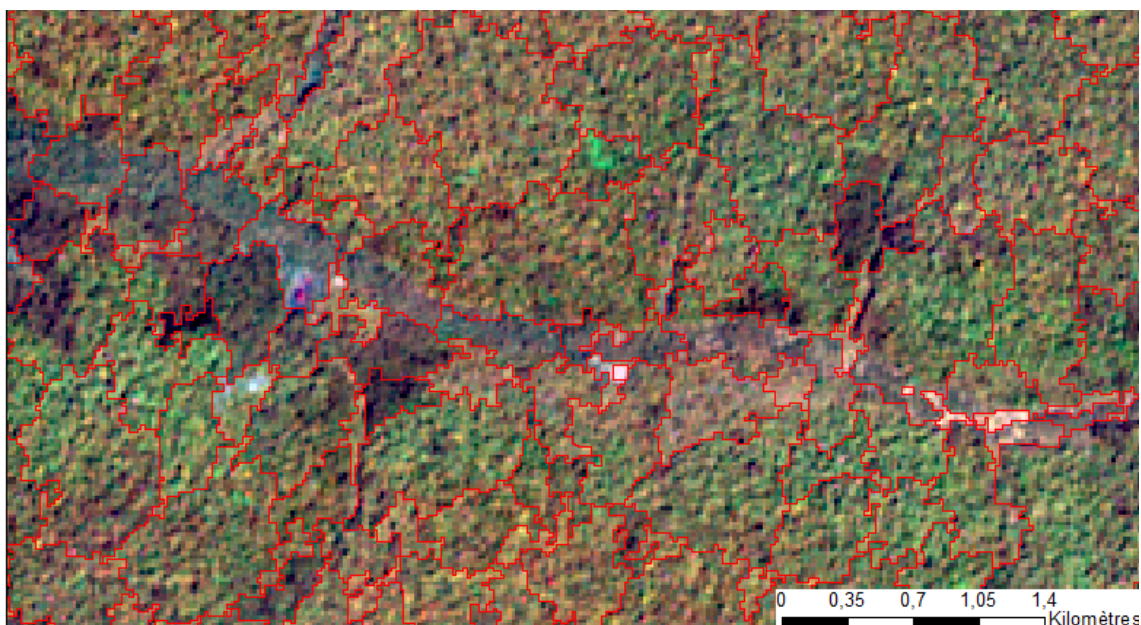


FIGURE 15 – Définition des objets par segmentation - Image Sentinel-2, composition colorée : B07-B11-B05

2. La segmentation est réalisée à l'aide du software de *Trimble : e-Cognition*

4.3.3 Protocole de validation

La validation est faite de deux manières : tout d'abord de manière visuelle et ensuite de manière plus objective avec des zones de validation et des métriques d'évaluation de la qualité de la classification.

La validation visuelle consiste en une évaluation subjective de la classification. Cela permet de donner rapidement une qualité à une classification. Elle se base sur les connaissances de l'utilisateur et nécessite donc une connaissance du terrain. L'évaluation se base sur différents paramètres comme la bonne classification de zones connues, la répartition des classes dans l'espace, l'effet poivre et sel, etc.

Les métriques d'évaluation de la qualité d'une classification sont choisies en accord avec la littérature scientifique. L'OA et l'indice kappa servent d'évaluation globale tandis que le F-score sera une indication sur la qualité de classification de chacune des classes. En plus des F-score, les PA et UA sont aussi calculés. L'obtention de ces métriques découle de la matrice de confusion calculée pour chaque classification. La matrice compare les résultats d'une classification avec la réalité du terrain.

Pour cela des zones de validation dont le type de végétation est connu sont nécessaires. Celles-ci, dans le cas de la zone de développement, proviennent de deux jeux de données : la campagne de terrain menée personnellement (Campagne terrain 1) et une autre campagne menée simultanément par Baptiste Delhez (Campagne terrain 2).

Tous les points obtenus lors de cette dernière sont utilisés uniquement dans des zones de validation. Toutes les classes d'occupation du sol ne sont cependant pas dans ce jeu de données. C'est pourquoi une partie des échantillons obtenus par la campagne principale (Campagne terrain 1) sont répartis entre les zones d'entraînement et les zones de validation. Cette division se fait de manière aléatoire, cependant l'entièreté des échantillons établis sur base de photo-interprétation sert à l'entraînement et non à la validation (Figure 16) .

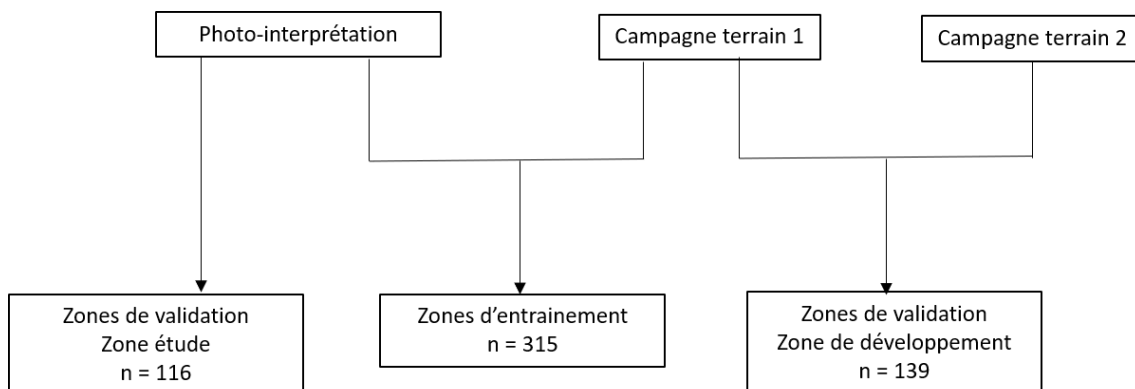


FIGURE 16 – Répartition des zones d'entraînement et de validation dans la zone de développement

Les détails des aires de validation pour la zone de développement sont repris au tableau 5 et visibles à la figure 17.

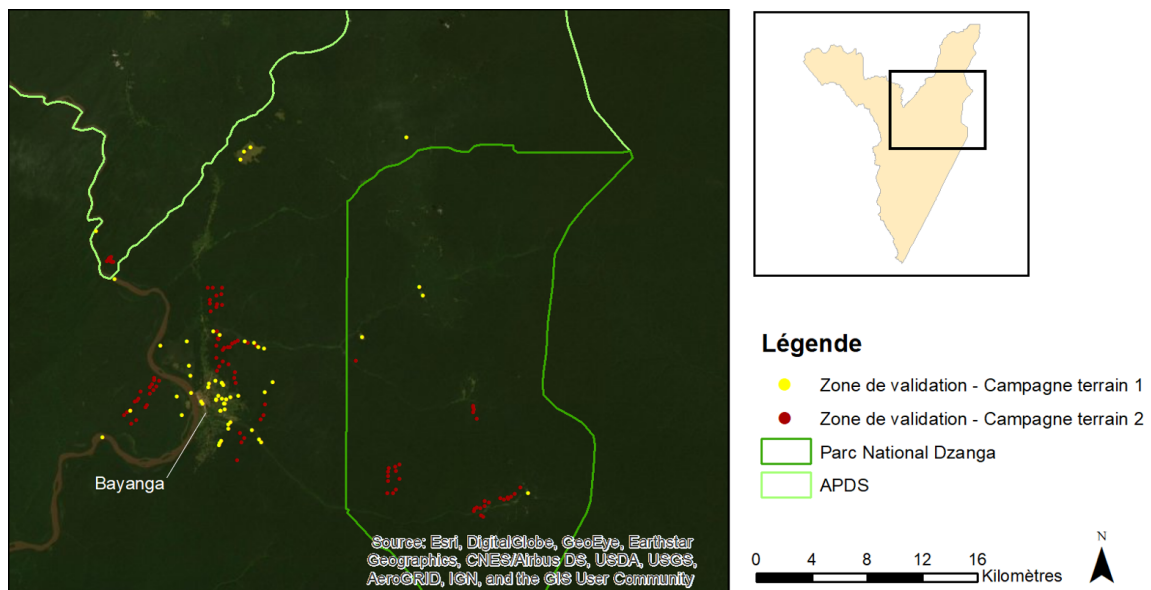


FIGURE 17 – Répartition des aires de validation pour la comparaison de classification dans la zone de développement

	Nombre Zone de validation	Campagne 1	Campagne 2	Nombre Pixel	% en pixel
F sempervirente	11	0	11	494	4,5
F dense humide	34	1	33	2117	19,2
F marécageuse	8	8	0	842	7,6
F secondaire vieille	11	0	11	807	7,3
F secondaire jeune	26	0	26	1344	12,2
F ripicole haute	7	7	0	85	0,8
F inondable	6	6	0	1079	9,8
Clairière saline	5	5	0	339	3,1
Zone habitation	4	4	0	167	1,5
Eau	4	4	0	2435	22,1
Savane	3	3	0	995	9,0
Zone agricole	20	20	0	337	3,1

TABLE 5 – Détail des aires de validation pour la zone d'entraînement

4.4 Production cartographique

Une cartographie des types forestiers et de l'occupation du sol pour l'ensemble de la zone d'étude est réalisée. La méthode de classification choisie est celle qui donne les meilleurs résultats sur la zone de développement.

4.4.1 Protocole de validation

Comme la cartographie des APDS est faite sur l'ensemble de la zone d'étude, la surface couverte est supérieure à celle parcourue lors de la campagne de terrain. Dès lors la validation et l'estimation de la qualité de la cartographie à l'échelle de la zone d'étude nécessitent des zones de validation différentes de celles de la zone de développement.

Un nouveau jeu de données de validation est créé sur base uniquement de photo-interprétation. Le nouveau jeu couvre l'ensemble de la zone d'étude. Dans un but d'objectivité, le choix des zones de validation s'est effectué selon un quadrillage de 41 zones réparties sur les APDS. Comme ce quadrillage n'est pas assez fin pour obtenir suffisamment de diversité au niveau des catégories de végétation, toutes les classes ne sont pas représentées. Dès lors, 75 zones de validation sont ajoutées sur base de photo-interprétation dans les catégories les moins représentées par le quadrillage. La répartition spatiale de ces zones est présentée à la figure 18. Pour les zones non-forestières, la photo-interprétation se base à la fois sur les images Sentinel-2 et sur les images CNES/Airbus disponibles sur *Google Earth*. Le tableau 6 reprend le détail des zones de validation pour chacune des catégories.

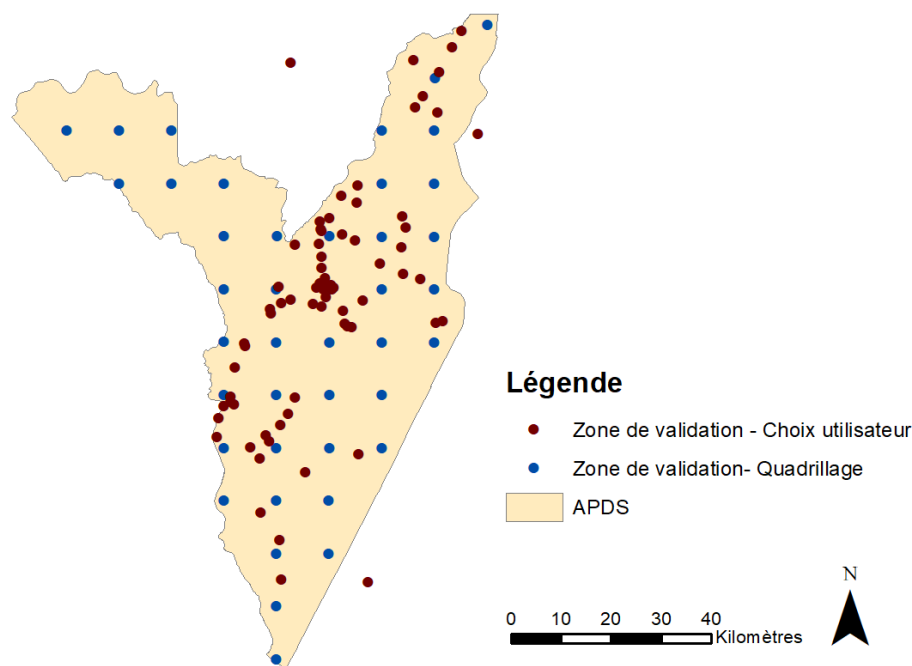


FIGURE 18 – Répartition des aires de validation pour la zone d'étude

	Nombre Zone de validation	Quadrillage	Choix utilisateur	Nombre Pixel
F sempervirente	9	3	6	1285
F dense humide	18	18	0	2622
F marécageuse	7	0	7	1442
F secondaire vieille	12	12	0	1594
F secondaire jeune	8	4	4	1585
F ripicole haute	7	0	7	284
F inondable	9	2	7	2780
Clairière saline	10	0	10	1171
Zone habitation	10	0	10	738
Eau	5	0	5	5358
Savane	9	0	9	1864
Zone agricole	12	2	10	1456

TABLE 6 – Détail des zones de validation pour la cartographie complète

4.5 Estimation de la dynamique agricole

4.5.1 Images satellites

Pour évaluer l'évolution des surfaces agricoles au sein des APDS, des images antérieures à 2015 et donc au lancement de Sentinel-2 sont nécessaires. Les comparaisons se feront donc sur base des images disponibles sur le logiciel *Google Earth*. Celles-ci sont des images CNES/Airbus à très hautes résolutions qui permettent de délimiter avec plus de précision les limites forestières. Cette délimitation est faite sur base de photo-interprétation. De plus, et selon la localisation de la zone de culture, les dates des images disponibles diffèrent. Un tableau récapitulatif reprend les dates de comparaison pour chaque secteur (Tableau 7).

Secteur	Date initiale	Date finale
Babongo Nord	21/11/2013	13/05/2017
Babongo Sud	21/11/2013	13/05/2017
Moussapoula Nord	21/11/2013	13/05/2017
Moussapoula Sud	18/12/2013	13/05/2017
Bayanga Centre	18/12/2013	08/01/2018
Yadoumbé	18/12/2013	08/01/2018
Lidjombo	12/05/2015	27/10/2017
Bomandjokou	12/05/2015	27/10/2017

TABLE 7 – Dates des images utiles pour la comparaison des surfaces agricoles dans chacun des secteurs agricoles

4.5.2 Évaluation des surfaces non forestières

Afin de déterminer l'évolution des surfaces agricoles dans chaque bande culturale, les surfaces non forestières sont mesurées pour deux dates. Pour cela et sur base de

photo-interprétation, les lisières forestières sont mises en évidence comme illustré à la figure 19. Les calculs de surfaces se font sur un logiciel SIG.

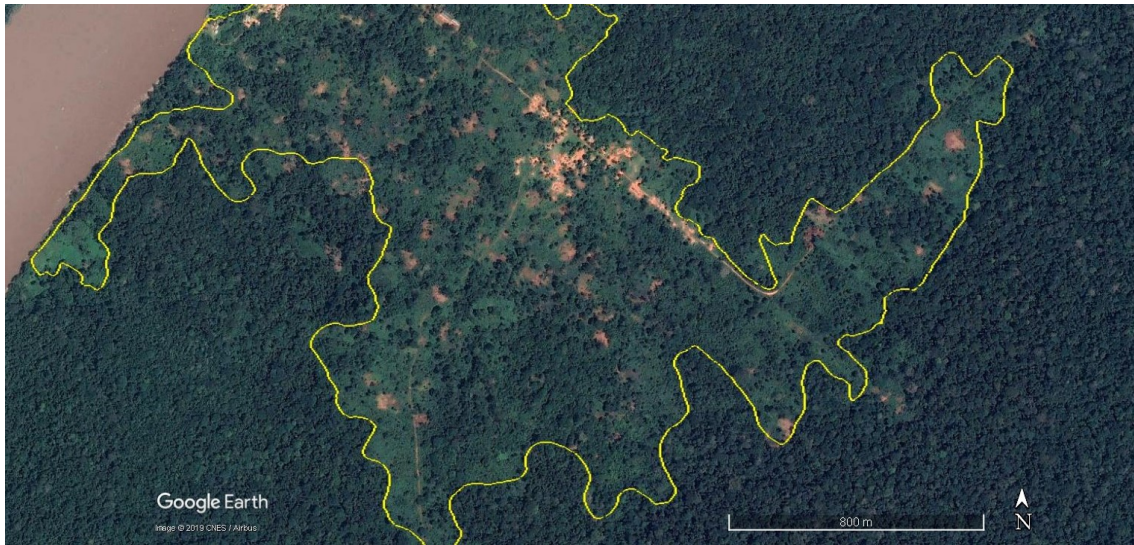


FIGURE 19 – Différenciation des zones non forestières - Secteur Lidjombo

5 Résultats & Discussion

Cette section reprend l'ensemble des résultats obtenus. Premièrement, la typologie de végétation établie lors de la campagne de terrain, ainsi que la photo-interprétation correspondante sont présentées. Dans un deuxième temps, les résultats de chacune des méthodes de classification sont présentés. Le choix de la classification retenue et de ses performances est ensuite discuté. Troisièmement, une cartographie des types de forêts et de l'occupation du sol sur l'ensemble de la zone d'étude est produite. Les résultats comme l'extension de l'échelle de classification sont discutés. Enfin, la dynamique agricole dans les APDS est évaluée.

5.1 Typologie de végétation

Suite à la campagne de terrain, une nouvelle typologie de la végétation des APDS a été créée. Par rapport à la typologie préexistante, elle permet d'apporter plus de détails tout en restant au maximum généralisable à d'autres zones forestières d'Afrique Centrale. Les principaux types de végétation, forestier et non-forestier, sont présentés ci-après. Un récapitulatif de la photo-interprétation est disponible à la figure 44.

5.1.1 Forêt sempervirente - *Gilbertiodendron dewevrei*

Ces massifs forestiers sont caractérisés par la présence en peuplements presque monospécifiques de l'espèce *Gilbertiodendron dewevrei*, aussi souvent appelée *Limballi*. Ils se retrouvent généralement à proximité des cours d'eau, dans les vallées et cuvettes, et se développent dans des endroits humides, parfois sur sol inondable (Harris, 2002, Boulvert, 1986, Meunier et al., 2015, Harris, 2002). Mais certains peuplements sont aussi sur terre ferme. Le sol est plat ou en légère pente et peut être inondé durant certaines périodes de l'année. La canopée est régulière, peu d'arbres émergents ressortent, la hauteur varie entre 30m et 40m. Contrairement à beaucoup d'autres types forestiers présents en forêt tropicale, les peuplements de *Gilbertiodendron dewevrei* ont un sous-bois très ouvert caractérisé par un épais tapis de feuilles mortes (Figure 21). En effet, le renouvellement de la végétation dans une forêt sempervirente n'est pas lié à la succession des saisons et se fait de manière régulière tout au long de l'année (Nziengui et al., 2013). Ce tapis a pour conséquence que la végétation au sol peine à se développer et est donc plus rare que dans d'autres types forestiers.



FIGURE 20 – Forêt sempervirente
Composition colorée : B07-B11-B05
(Sentinel-2)

Le comportement grégaire de cette espèce rend ses massifs très particuliers et reconnaissables aussi bien sur le terrain que par photo-interprétation. Ce sont des tâches foncées et homogènes en termes de texture et de réflectance (Figure 20).



FIGURE 21 – Forêt sempervirente - *Gilbertiodendron dewevrei*

5.1.2 Forêt dense humide semi-décidue

Les forêts denses humides semi-décidues se retrouvent aux sources de cours d'eau, mais aussi dans des zones non anthropisées, généralement dépourvues d'activités humaines. Les conditions hydromorphiques sont généralement sèches bien qu'elles se trouvent à proximité de cours d'eau. Le relief va de plat à légèrement accidenté. Il est caractérisé par une canopée fermée d'une hauteur proche de 35m. La strate arbustive s'élève en moyenne à 15m et est ouverte. Le sous-bois est lui aussi ouvert (Figure 23). On retrouve de plus une densité haute d'arbres dominants de circonférence importante (diamètre > 2m). Ce type de forêt se situe du point de vue de sa structure entre la forêt sempervirente et la forêt secondaire vieille.

Les forêts denses humides se distinguent sur les compositions colorées de par leur couleur particulière et leur position aux sources des cours d'eau (Figure 22).

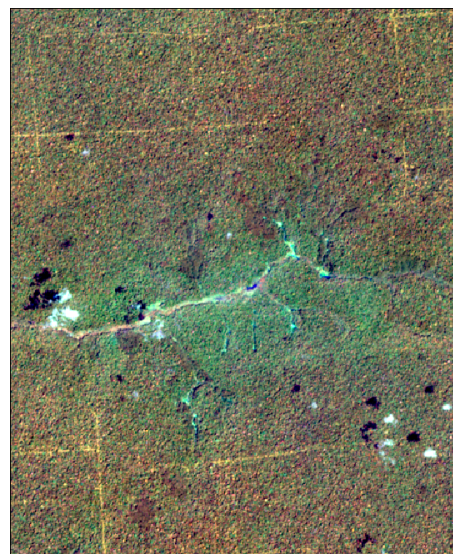


FIGURE 22 – Forêt dense humide semi-décidue
Composition colorée : B07-B11-B05
(Sentinel-2)



FIGURE 23 – Forêt dense humide semi-décidue

5.1.3 Forêt secondaire vieille

Les forêts secondaires vieilles sont comme leur nom l'indique des forêts secondaires, c'est-à-dire des forêts ayant été exploitées par l'homme. Cette exploitation s'est faite il y a déjà un certain temps contrairement aux forêts secondaires jeunes exploitées plus récemment. Dans le parc national de Dzanga, les dernières exploitations sont antérieures au statut de parc naturel et datent des années 80. Cette forêt est une des plus communes au sein des APDS. Elle est caractérisée par la présence d'une densité d'arbres dominants élevée. La strate arborée est d'une hauteur moyenne de 35m, la canopée est généralement fermée et hétérogène de par la présence de dominants. Le sous-bois est quant à lui plus ouvert que dans une forêt semi-décidue (Figure 25). A de nombreux endroits se trouvent aussi des trouées. Ces dernières sont très vite remplies d'une végétation épaisse et très dense composée principalement de *Haumania danckelmaniana*, une plante herbacée vivace de la famille marantaceae. La hauteur de cette dernière peut atteindre plus de 2m.

L'identification par photo-interprétation est simple, car elle couvre une grande surface sur les APDS (Figure 24). Ce type forestier est en réalité composé d'une variabilité importante de réflectance. Les lignes plus claires sur les compositions colorées sont en réalité à anciens chemins d'exploitation (Figure 26).

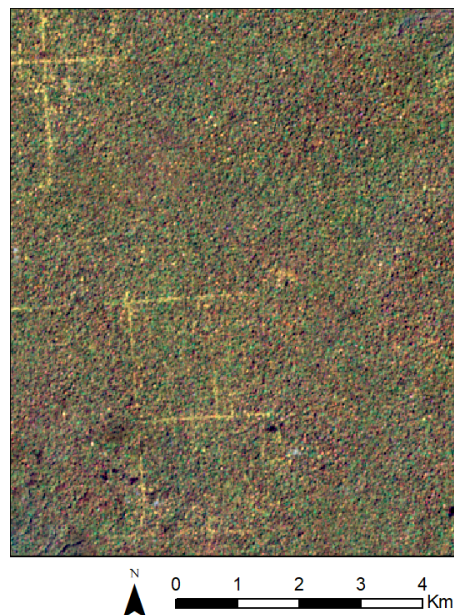


FIGURE 24 – Forêt secondaire vieille
Composition colorée : B07-B11-B05
(Sentinel-2)



FIGURE 25 – Forêt secondaire
vieille

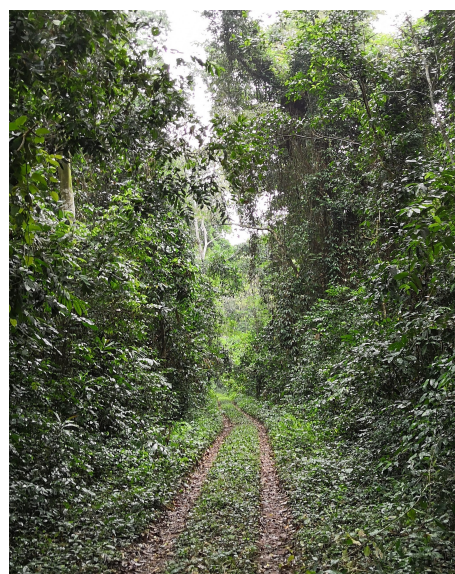


FIGURE 26 – Ancien chemin
d'exploitation au sein des forêts
secondaires vieilles

5.1.4 Forêt secondaire jeune

Les forêts secondaires jeunes correspondent aussi à des forêts exploitées, mais à des périodes plus récentes que les forêts secondaires vieilles. Ces forêts se retrouvent plus proches des zones anthropisées, notamment des villages. Elles sont remarquables par l'absence d'arbres dominants et par le faible diamètre moyen des arbres (diamètre < 70 cm) (Figure 28). La canopée est d'une hauteur moyenne de 20m et est ouverte. La strate arbustive est elle à une hauteur de 5m et est fermée. Les espèces sont variées. Les trouées sont colonisées par *Calamus deeratus* et *Haumania danckelmaniana*. Le sol est généralement plat et sec.

Sur les compositions colorées, ces forêts sont difficilement distinguables des forêts secondaires vieilles. Néanmoins sur la figure 27, les forêts secondaires jeunes ressortent comme un mélange de forêts secondaires vieilles et de forêts denses humides.



FIGURE 27 – Forêt secondaire jeune
Composition colorée : B07-B11-B05
(Sentinel-2)



FIGURE 28 – Forêt secondaire jeune

5.1.5 Forêt marécageuse

Les forêts marécageuses ont la caractéristique d'être inondées en permanence. Elles se retrouvent logiquement de part et d'autre de certains cours d'eau, notamment ceux de taille plus importante, et aux confluences de plusieurs fleuves. Dans cette catégorie se retrouvent aussi les forêts de raphia. Elles se distinguent des forêts ripicoles hautes par une moins grande hétérogénéité au niveau des espèces présentes. Ces forêts peuvent présenter un faciès différent selon les espèces présentes. Par exemple à la confluence de la rivière Massapoula et la rivière Sangha, se retrouvent des raphia (*Raphia hookeri*) quasiment en monospécifique (Figure 30). Cela est notamment dû à leur utilisation par la population locale pour la fabrication d'alcool dérivé. Dans ce cas, la canopée n'est pas continue, les arbres ont une taille moyenne de 15m et un diamètre de 20cm. Le sol est logiquement détrempe et boueux avec un gradient en fonction de la proximité de la rivière. Au plus proche de la rivière, le sol est humide voire sous eau, au contraire en s'éloignant de la rivière, le sol se fait plus sec.

Sur les compositions colorées, les forêts marécageuses apparaissent de couleur foncée le long des cours d'eau (Figure 29).



FIGURE 29 – Forêt marécageuse
Composition colorée : B07-B11-B05
(Sentinel-2)



FIGURE 30 – Forêt marécageuse - ici monospécifique de raphia

5.1.6 Forêt ripicole haute

Les forêts ripicoles hautes se développent le long, et dans des cours d'eau tout comme certaines forêts marécageuses et forêts inondables. Néanmoins, ce type de forêt se développe autour de cours d'eau de moindre importance, généralement de 1 à 2m de large.

En comparaison avec les forêts marécageuses, la végétation est plus dense et moins haute (3-5m). Elle comporte de plus, une grande variété d'espèces (Figure 32).

Sur base d'images satellites, ce type de forêt est une ligne de quelques pixels de large et de couleur foncée, qui suivent le cours d'eau (Figure 31).



FIGURE 31 – Forêt ripicole haute
Composition colorée : B07-B11-B05
(Sentinel-2)



FIGURE 32 – Forêt ripicole haute

5.1.7 Forêt inondable

Ce type de forêt est présent exclusivement en bordure ou à proximité des cours d'eau. Comme son nom l'indique, la végétation est partiellement inondée durant l'année. Les hauteurs d'eau ainsi que la période d'inondation varient selon la localisation du massif. L'allure et les espèces présentes dans ce type de forêt sont très hétérogènes, allant de peuplement quasi monospécifique à de grandes diversités de végétation (Figure 34). Cette diversité est présente aussi sur les compositions colorées avec des faciès hétérogènes en termes de couleur et de granulométrie (Figure 33).

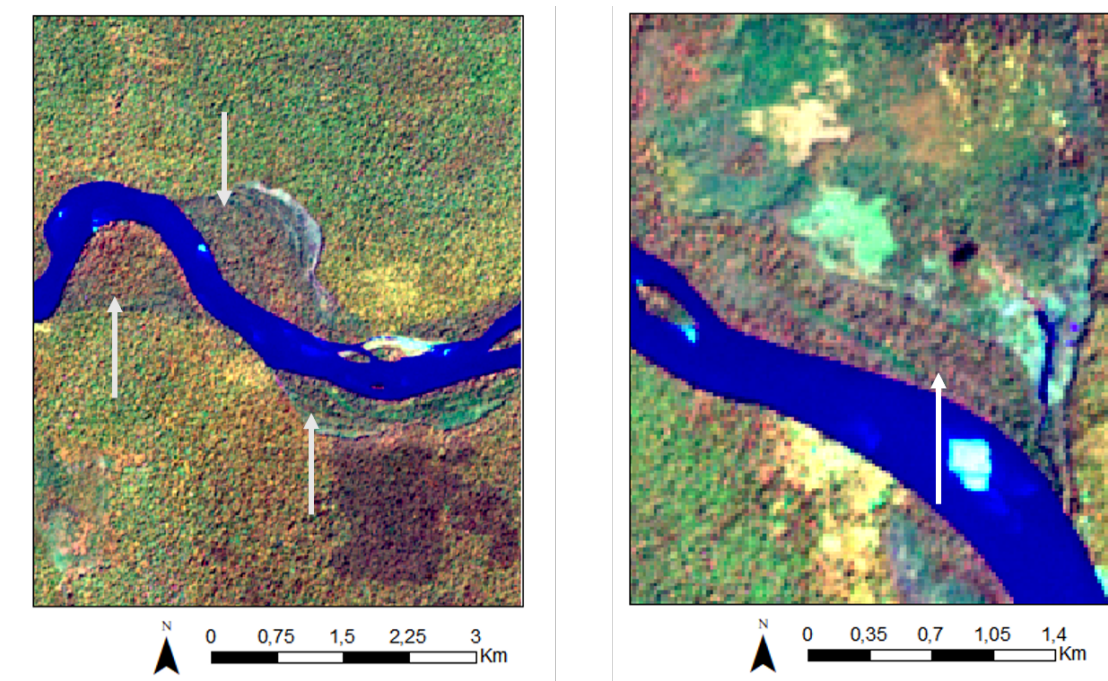


FIGURE 33 – Forêt inondable - Composition colorée : B07-B11-B05 (Sentinel-2)



FIGURE 34 – Forêt inondable

5.1.8 Clairière saline

Les salines sont des marais salés riches en sels minéraux (sulfate de calcium hydraté, chlorure de sodium, chlorure de potassium, etc.) présents au milieu des forêts. Ces milieux sont fort fréquentés par de nombreux animaux qui recherchent ces eaux riches en matière minérale introuvable en forêt tropicale (Figure 37). La présence et les activités des animaux maintiennent le caractère ouvert du milieu (Nziengui et al., 2013). Certaines salines sont de plus aménagées pour l'observation de ces animaux. Elles sont localisées généralement près d'un cours d'eau et peuvent être en partie inondées.

Sur les compositions colorées, les salines apparaissent comme de petites zones claires et de dimension moindre par rapport aux savanes de manière générale (Figure 36).



FIGURE 35 – Clairière saline - *Dzanga Bai*



FIGURE 36 – Saline
Composition colorée :
B07-B11-B05
(Sentinel-2)



FIGURE 37 – Clairière saline

5.1.9 Savane

Les savanes présentes au sein des APDS sont des savanes incluses, c'est à dire entourées de forêts et qui forment des "îles". Dans la région protégée, elles sont toutes situées au nord-est. Ce type de milieu est plus fréquent et de plus grande superficie au nord des APDS sur la route de Nola. Leurs origines relèveraient de facteurs liés au sol (Harris, 2002). Les savanes sont ici des étendues herbeuses parsemées d'arbustes, d'une hauteur maximale de 3m, et de buissons de 1m de haut. Elles sont donc qualifiées de savanes arbustives (Figure 39).

Elles apparaissent comme de grandes zones homogènes en termes de couleur et de textures au milieu de la forêt (Figure 38). Elles se différencient des clairières salines de par leur taille plus importante et du fait que le sol reste sec.



FIGURE 38 – Savane
Composition colorée : B07-B11-B05
(Sentinel-2)



FIGURE 39 – Savane arbustive de type savane incluse

5.1.10 Zone agricole

Les zones agricoles consistent en une mosaïque de parcelles cultivées et en jachères de faible dimension, comprises généralement entre 100 et 2500 m² (Figure 40). La culture principale et largement majoritaire est celle de manioc. L'agriculture pratiquée est une agriculture sur brûlis et suit une séquence cyclique en trois temps (Figure 41). Après une période de jachère, la végétation est coupée et brûlée sur place afin d'apporter plus de fertilité dans le sol. Le manioc est ensuite planté et cultivé pendant 6 mois à 1 an. Ensuite la parcelle est remise en jachère. Les zones de cultures sont cependant limitées dans des zones précises appelées bandes culturelles.



FIGURE 40 – Zone agricole autour de Bayanga



FIGURE 41 – Séquence de culture classique de l'agriculture sur brûlis pratiquée au sein des APDS

5.1.11 Autres occupations du sol

Zone d'habitation Les zones d'habitations sont les surfaces qui comprennent les bâtiments et les jardins autour, les routes, etc. (Figure 42).

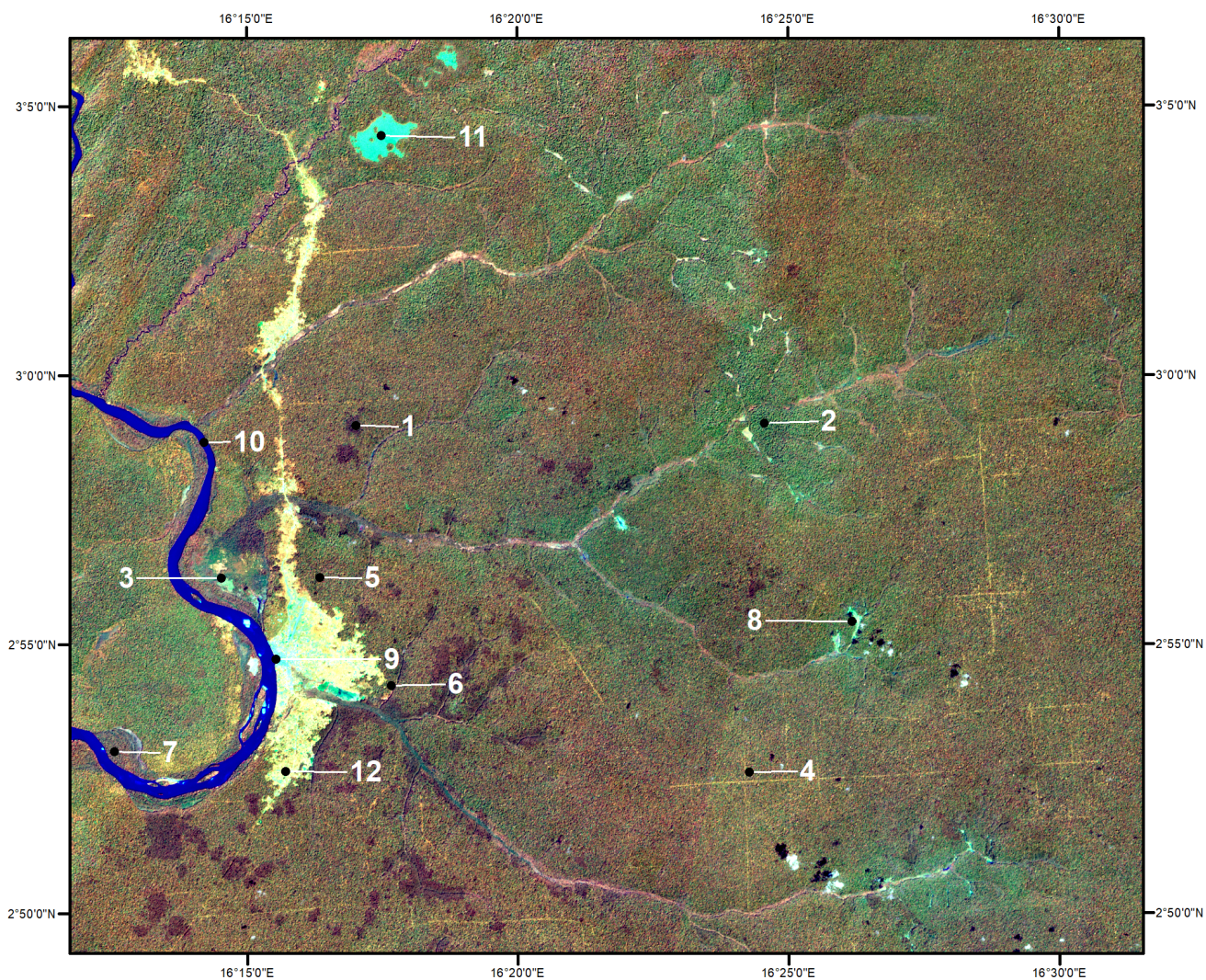
Eau Cette classe d'utilisation du sol comprend les surfaces aquatiques de grandes dimensions. Dans le cas des APDS, seule la rivière Sangha correspond à ces critères (Figure 43).



FIGURE 42 – Zone d’habitation
autour de Bayanga



FIGURE 43 – Eau : rivière Sangha
aux abords de Bayanga



- | | |
|------------------------------------|----------------------|
| 1. Forêt sempervirente | 7. Forêt inondable |
| 2. Forêt dense humide semi-décidue | 8. Clairière saline |
| 3. Forêt marécageuse | 9. Zone d'habitation |
| 4. Forêt secondaire vieille | 10. Eau |
| 5. Forêt secondaire jeune | 11. Savane |
| 6. Forêt ripicole haute | 12. Zone agricole |

Système de coordonnées: WGS 1984 UTM Zone 33N
 Projection: Transverse Mercator
 Datum: WGS 1984
 Unités: Mètre
 Source : Image Sentinel-2, 27/12/2018
 Composition colorée : B 7-11-5
 Auteur : D.Lessire (Uclouvain, 2019)



0 2,5 5 7,5 10
 Kilomètres

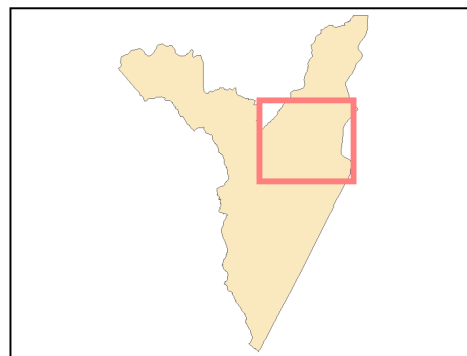


FIGURE 44 – Photo-interprétation de l'occupation du sol de la zone d'étude

5.2 Comparaison des méthodes de classification

Pour l'ensemble des résultats présentés ci-dessous, et pour des raisons de facilité de lecture, des numéros ont été attribués à chacune des classes. Les numéros sont les suivants :

- | | |
|------------------------------------|-----------------------|
| 1. Forêt sempervirente | 9. Forêt inondable |
| 2. Forêt dense humide semi-décidue | 10. Clairière saline |
| 3. Forêt marécageuse | 11. Zone d'habitation |
| 5. Forêt secondaire vieille | 12. Eau |
| 6. Forêt secondaire jeune | 13. Savane |
| 8. Forêt ripicole haute | 14. Zone agricole |

Pour certaines classifications, quatre nouvelles classes ont été créées et correspondent à des regroupements de plusieurs classes. C'est le cas des classes suivantes :

15. Classes non forestières (classes 10 à 14)
16. Forêts secondaires (classes 5 et 6)
17. Forêts liées à l'eau (classes 1, 3, 8 et 9)

5.2.1 Performance de la classification non supervisée - Approche par pixel

La classification non supervisée par ISODATA donne des résultats mitigés. La matrice de confusion (Annexe - Tableau A.1) et son tableau d'analyse associé (Tableau 8) présentent plusieurs éléments :

- Aucun pixel de validation n'a été classé dans la catégorie 9 (forêt inondable) et donc logiquement cette catégorie obtient un F-score de 0. Aucune caractéristique commune à ces zones et propre à la classe ne semble ressortir.
- Les F-scores pour chaque catégorie sont assez faibles ($<0,4$) à l'exception des forêts secondaires vieilles (classe 5, F-score=0,74) et des zones non forestières (classe 15, F-score=0,95). La classification de ces dernières est même très bonne avec un F-score se rapprochant de 1.

La méthode de classification par ISODATA semble fort limitée au vu des indices de qualité généraux. L'OA est seulement de 0,56 et le coefficient kappa de 0,45. Si elle ne permet pas de classifier les forêts avec une bonne exactitude, elle permet néanmoins de différencier les zones forestières des zones non forestières.

Les résultats de la classification sont présentés à la figure 45. La forêt secondaire vieille, bien que possédant un F-score élevé, garde un aspect "poivre et sel" avec notamment des confusions pour certains pixels avec les forêts denses humides semi-caducifoliées et les forêts sempervirentes. Il semble également qu'une partie des savanes au nord soit classée comme forêt marécageuse. Cela peut expliquer les zones non-forestières mal classées. Enfin, une autre constatation est que la plupart des lisières forestières sont classées comme forêt dense humide semi-caducifoliée.

Classes	1	2	3	5	6	8	9	15
Producer accuracy	0,17	0,42	0,19	0,71	0,38	0,35	0	0,98
User accuracy	0,68	0,23	0,42	0,77	0,32	0,48	0	0,92
F-Score	0,27	0,30	0,26	0,74	0,35	0,38	0	0,95
Overall accuracy	0,56							
Indice Kappa	0,45							

TABLE 8 – Tableau Récapitulatif de la matrice de confusion pour la classification par ISODATA.

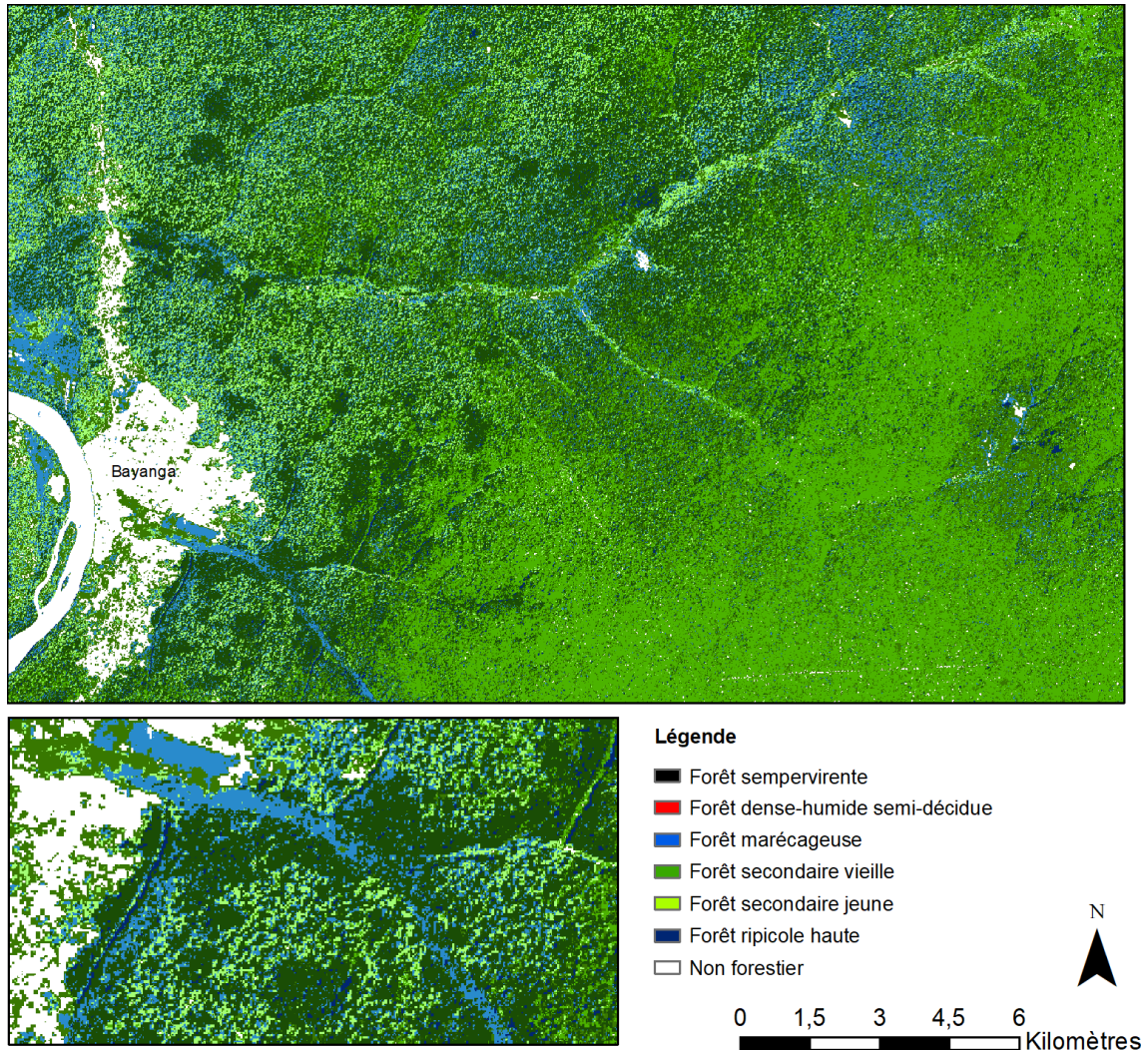


FIGURE 45 – Classification non supervisée par ISODATA - Zoom à proximité de Bayanga

5.2.2 Performance de la classification supervisée - Approche par pixel

Les résultats de la classification supervisée non hiérarchique à l'échelle des pixels sont repris au tableau 9. La matrice de confusion correspondante est disponible en annexe (Tableau A.2). Dans cette classification, toutes les classes de végétation et d'occupation du sol sont reprises (12 classes). La qualité est légèrement supérieure à celle obtenue par la méthode ISODATA avec un indice kappa de 0,53 et une OA

de 0,59.

Au niveau des qualités de classification pour les différentes classes, de grandes différences existent. D'une manière générale, les zones non forestières (classes 10 à 14) sont mieux classées que les zones forestières. Cependant, les clairières salines possèdent un F-score plutôt faible de 0,47.

Les classes forestières présentent une large variabilité de F-score. Les forêts secondaires vieilles (classe 5) ressortent comme bien classées avec un score de 0,76. Au contraire, les forêts denses humides semi-caducifoliées et les forêts inondables sont très mal classées avec des F-scores de 0,18 et 0. Aucune zone de validation n'a été classée dans la catégorie des forêts inondables. Visuellement, cette catégorie est d'ailleurs très peu représentée. Pour les autres classes forestières, les F-scores varient de 0,24 à 0,44.

Plusieurs faibles F-scores sont les résultats des fortes différences entre les PA et les UA. Pour les forêts denses humides (classe 2) et les forêts marécageuses (classe 3), les PA sont bien plus élevées que les UA (0,82 contre 0,10 pour les forêts denses humides et 0,42 contre 0,16 pour les forêts marécageuses). A contrario, les forêts secondaires jeunes (classe 6) ont une UA bien supérieure à leur PA (0,96 contre 0,28).

La classification obtenue sur la zone de développement est présentée à la figure 46. Les zones forestières sont classées majoritairement en forêts secondaires, jeunes et vieilles. Celles-ci sont même surreprésentées et prennent le pas sur d'autres classes, comme les forêts sempervirentes ou les forêts marécageuses. Les forêts inondables sont elles totalement absentes de la classification. Pour les zones non forestières, les zones d'habitation semblent bien représentées. Une confusion subsiste entre les zones agricoles et les clairières salines.

Classes	1	2	3	5	6	8	9	10	11	12	13	14
Producer accuracy	0,47	0,82	0,42	0,69	0,28	0,20	0	0,53	0,87	1,00	0,92	0,63
User accuracy	0,31	0,10	0,16	0,85	0,96	0,51	0	0,42	0,81	1,00	0,99	0,81
F-Score	0,37	0,18	0,24	0,76	0,44	0,29	0	0,47	0,84	1,00	0,95	0,71
Overall accuracy	0,59											
Indice Kappa	0,53											

TABLE 9 – Analyse de la matrice de confusion de la classification non hiérarchique basée sur les pixels - 12 classes

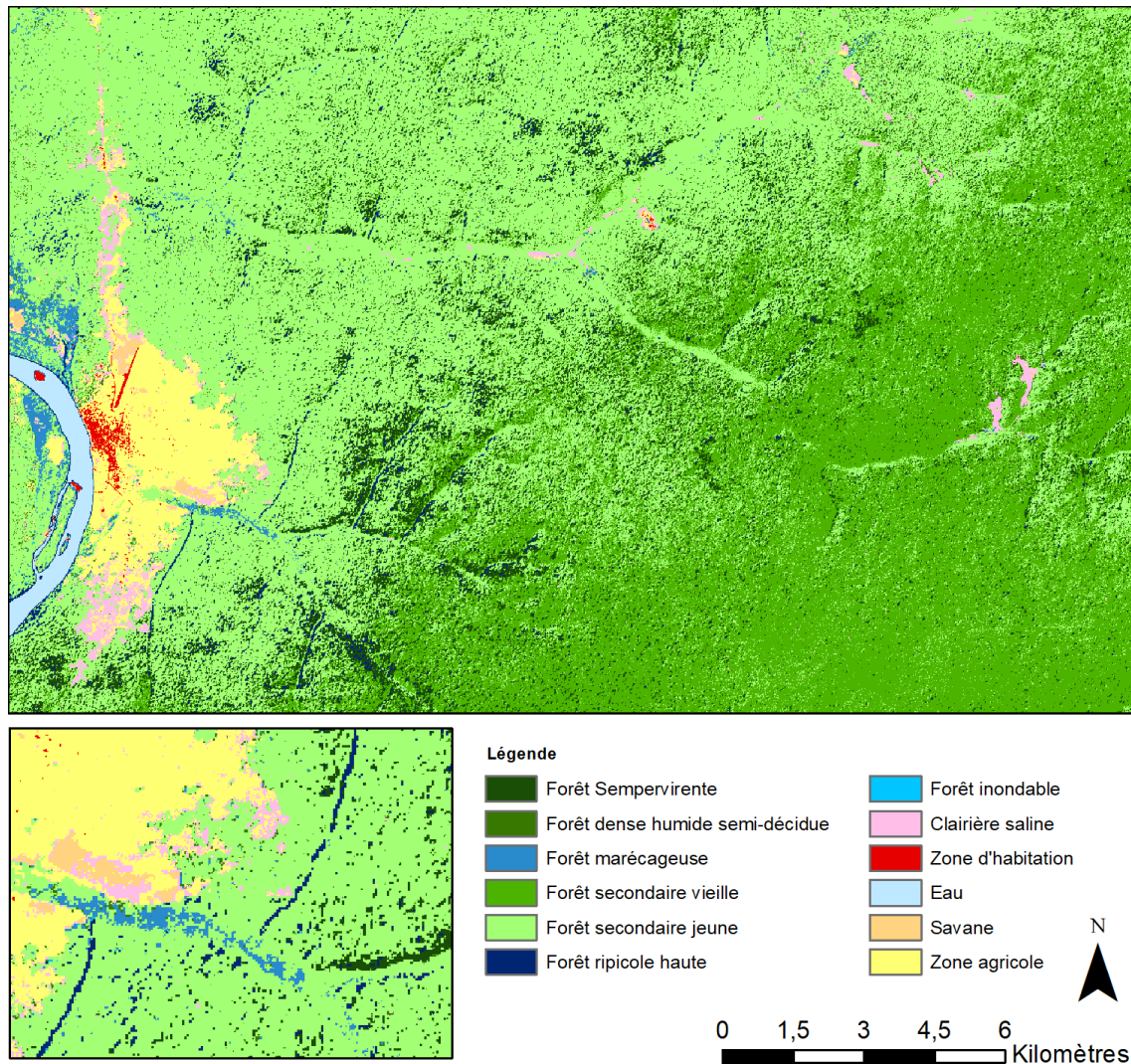


FIGURE 46 – Classification supervisée non hiérarchique basée sur les pixels - Zoom à proximité de Bayanga

5.2.3 Performance de la classification supervisée - Approche hiérarchique par pixel

Classification hiérarchique - 4 classes La matrice de confusion correspondant aux résultats de la classification par une approche hiérarchique avec 4 classes est disponible en annexe (Tableau A.3) et son analyse au tableau 10.

Comme précédemment, les zones forestières et non forestières se distinguent bien comme en témoigne le F-score de 0,97 pour les zones non forestières (classe 15). Les autres classes possèdent des F-score qui varient entre 0,6 et 0,7 ce qui peut être considéré comme bon.

L'OA pour cette classification est de 0,78 et l'indice kappa de 0,70. Sur cette base la classification peut être considérée comme bonne, mais ceci est à nuancer par le fait qu'elle ne comporte que 4 classes.

Classification hiérarchique - 12 classes Les résultats de la classification supervisée de manière hiérarchique complète, c'est-à-dire avec les 12 classes, sont présentés

Classes	2	16	17	15
Producer accuracy	0,64	0,63	0,67	0,98
User accuracy	0,71	0,54	0,70	0,97
F-Score	0,67	0,58	0,68	0,97
Overall accuracy	0,78			
Indice Kappa	0,70			

TABLE 10 – Analyse de la matrice de confusion de la classification hiérarchique (4 classes) - Approche par pixel

au tableau 11. L'OA de la classification est de 0,73 et son indice kappa de 0,68. La qualité peut être donc considérée comme bonne.

L'augmentation du nombre de classes ne diminue que légèrement la qualité de la classification. Lors du passage de 4 à 12 classes, l'OA passe de 0,78 à 0,73 (-0,05) et l'indice kappa de 0,70 à 0,68 (-0,02). Pour les différentes catégories, et comme précédemment, les zones non forestières ont des F-scores élevés. Seules les clairières salines sont moins bien classées (classe 10, F-score = 0,57). Pour les classes forestières, les forêts secondaires vieilles et les forêts denses humides semi-décidues ont des F-scores hauts (0,89 et 0,76). Les autres catégories forestières ont des F-scores compris entre 0,33 et 0,55.

La classification est présentée à la figure 47. Visuellement, la classification semble plutôt bonne. Elle est réalisée sur chaque pixel et un fort effet poivre et sel subsiste. Les forêts marécageuses semblent bien représentées tout comme les forêts denses humides semi-décidues. Les forêts inondables se confondent parfois avec les forêts sempervirentes. Pour les zones non forestières, une partie des zones agricoles est confondue avec les clairières salines. Les zones d'habitations sont elles bien définies. Cependant, dans cette dernière catégorie sont aussi classés des bancs de sable situés au milieu de la rivière Sangha.

Classes	1	2	3	5	6	8	9	10	11	12	13	14
Producer accuracy	0,51	0,66	0,46	0,88	0,54	0,24	0,47	0,64	0,87	1,00	0,91	0,80
User accuracy	0,44	0,88	0,67	0,90	0,24	0,52	0,30	0,52	0,82	1,00	0,97	0,77
F-Score	0,47	0,76	0,55	0,89	0,33	0,33	0,36	0,57	0,85	1,00	0,94	0,78
Overall accuracy	0,73											
Indice Kappa	0,68											

TABLE 11 – Analyse de la matrice de confusion de la classification hiérarchique via une approche par pixel - 12 classes

Pour pallier l'effet "poivre et sel" des classifications via une approche par pixel, des filtres peuvent être appliqués. Ils permettent de réduire les variabilités locales et les bruits éventuels. Dans ce cas, un filtre majoritaire est appliqué. Pour chaque fenêtre de 8x8 pixels, la valeur majoritaire des pixels est attribuée au pixel central. Le double filtre est un deuxième filtre majoritaire appliqué sur les résultats d'un premier filtre.

L'application de filtre majoritaire permet de donner un aspect plus lissé aux résultats (Figure 48). Tout d'abord, le filtre tend à supprimer les pixels isolés. En-

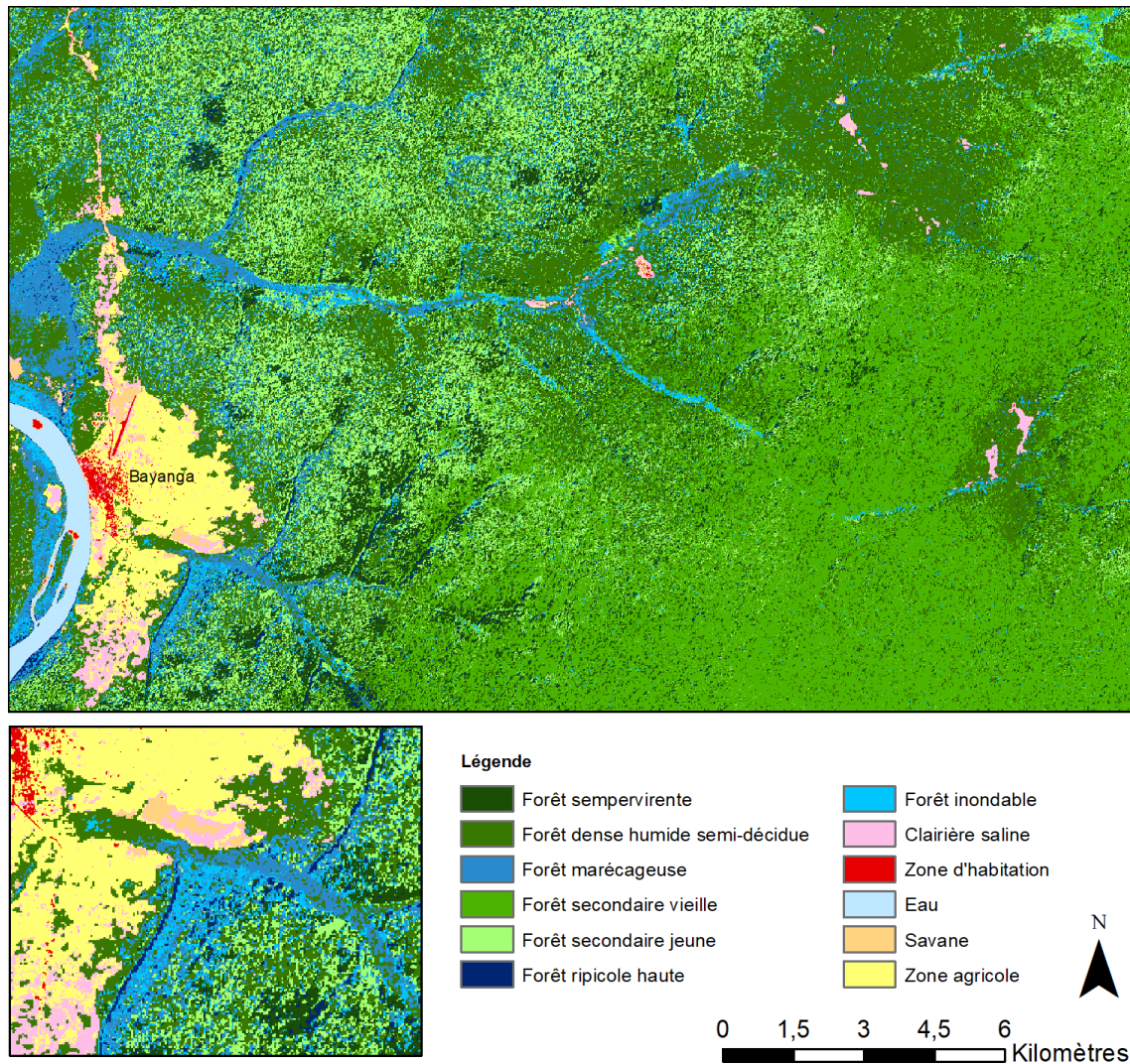


FIGURE 47 – Classification supervisée hiérarchique basée sur les pixels (12 classes)
- Zoom à proximité de Bayanga

suite, il réalise une agrégation des pixels de même classe. L'utilisation d'un double filtre majoritaire accentue ces effets, mais n'apporte pas de différences majeures par rapport au simple filtre. L'ajout de filtre n'enlève pas totalement l'effet poivre et sel au niveau macro mais le diminue au niveau micro.

L'application de filtre augmente très légèrement la qualité de classification comme montrée au tableau 12. Avec l'ajout d'un filtre majoritaire, l'OA passe de 0,73 à 0,74 et l'indice kappa de 0,68 à 0,70. L'ajout d'un deuxième filtre augmente encore légèrement l'OA d'un centième, mais ne modifie pas l'indice kappa. Les F-scores augmentent aussi ou restent stables avec l'ajout d'un filtre. Une exception cependant, le F-score des forêts sempervirentes tend à diminuer lors d'ajout de filtre. Néanmoins, ces augmentations ne sont pas significatives et ne correspondent pas à une meilleure classification. La faible augmentation des différents F-scores, et donc de la qualité globale de classification, provient de la suppression de pixels de contamination. Ces derniers sont des pixels isolés qui ont une réflectance suffisamment différente des pixels qui les entourent et sont donc classés différemment. Ils ne représentent pour autant pas une classe différente et peuvent être juste une variation au sein d'une

classe, comme une trouée au sein d'une forêt secondaire par exemple.

L'ajout de filtre n'augmente donc pas de manière significative la qualité de classification. Dès lors, aucun filtre n'est ajouté.

Classes	F-Score												OA	Indice Kappa
	1	2	3	5	6	8	9	10	11	12	13	14		
Sans Filtre	0,47	0,76	0,55	0,89	0,33	0,33	0,36	0,57	0,85	1,00	0,94	0,78	0,73	0,68
Filtre majoritaire	0,45	0,77	0,58	0,92	0,33	0,33	0,36	0,60	0,89	1,00	0,94	0,81	0,74	0,70
Double filtre majoritaire	0,44	0,78	0,59	0,93	0,33	0,34	0,36	0,60	0,90	1,00	0,94	0,81	0,75	0,70

TABLE 12 – Comparaison des qualités de classification avec l'apport de filtre majoritaire

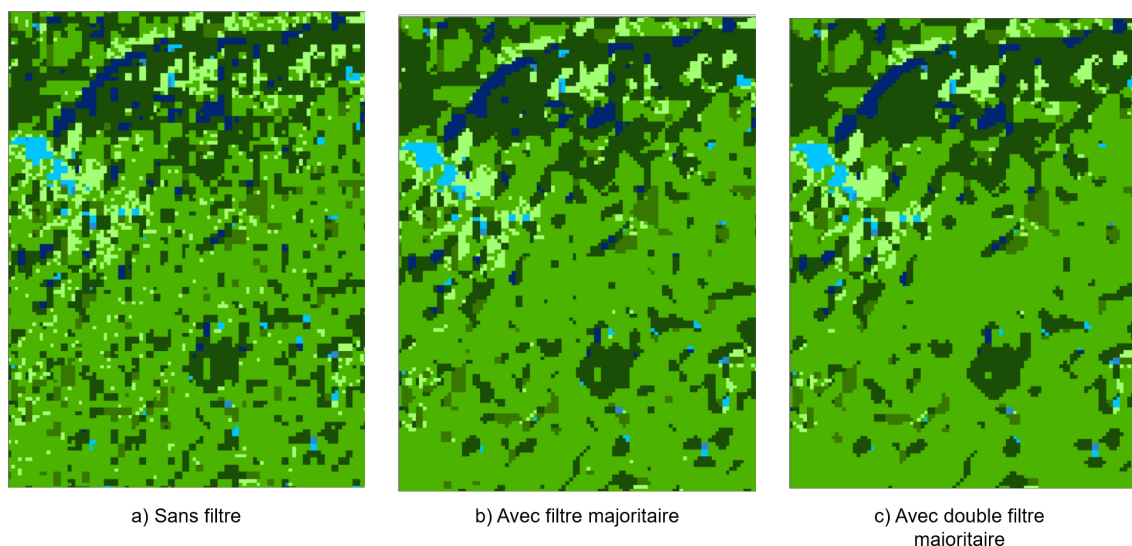


FIGURE 48 – Comparaison des résultats avec l'utilisation d'un filtre majoritaire

5.2.4 Performance de la classification supervisée - méthode par objet

Les résultats de la classification via une approche par objet sur la zone de développement sont présentés à la figure 49. L'analyse de la matrice de confusion correspondante (Annexe - Tableau A.5) est présentée au tableau 13.

La qualité globale de la classification est bonne, avec une OA de 0,78 et un indice kappa de 0,74. Les F-scores des classes non forestières (classes ≥ 10) sont élevés et tous supérieurs à 0,75. Pour les classes forestières, la variabilité des F-scores est plus importante. Deux classes ont des F-scores nuls : les forêts ripicoles hautes (classe 8) et les forêts inondables (classe 9). Les autres classes forestières ont des F-scores supérieurs à 0,5. Les forêts denses humides semi-décidues (classe 2) ont même un score proche de 1 (0,97).

Visuellement, la classification par objet ne contient pas d'effet poivre et sel et est plus lisse que les approches par pixel. Les forêts ripicoles hautes et les forêts marécageuses sont moins bien représentées que par les approches par pixel. Cela peut

provenir d'une segmentation d'objets de taille trop importante et qui par conséquent ne prend pas en compte les objets de petite taille. Ici encore certaines zones agricoles sont classées comme des clairières salines.

Classes	1	2	3	5	6	8	9	10	11	12	13	14
Producer accuracy	0,46	0,75	1,00	0,95	0,46	0	0,09	0,74	0,70	1,00	1,00	0,99
User accuracy	0,55	0,95	0,44	1,00	0,71	0	0,01	0,76	0,98	1,00	0,99	1,00
F-Score	0,50	0,84	0,61	0,97	0,56	0	0,02	0,75	0,82	1,00	0,99	0,99
Overall accuracy	0,78											
Indice Kappa	0,74											

TABLE 13 – Analyse de la matrice de confusion de la classification supervisée via une approche par objet sur la zone de développement

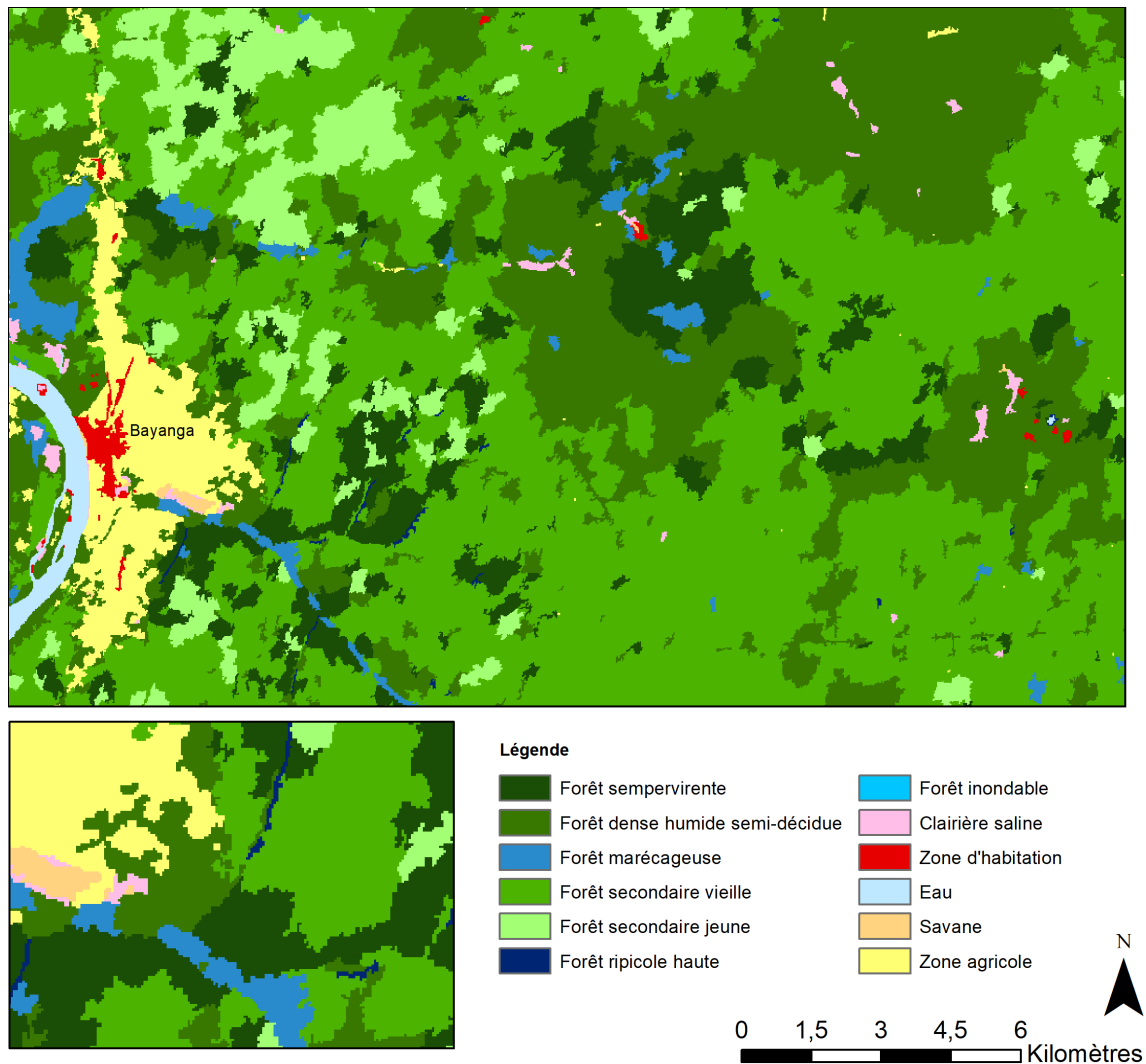


FIGURE 49 – Classification supervisée via une approche par objet - Zoom à proximité de Bayanga

5.2.5 Comparaison des classifications

Le tableau 14 ci-dessous permet de comparer la qualité des différentes classifications.

Méthode de classification	OA	Indice Kappa
Non supervisée - ISODATA	0,56	0,45
Supervisée non hiérarchique, approche par pixel - RF	0,59	0,53
Supervisée hiérarchique, approche par pixel - RF	0,73	0,68
Supervisée approche par objet - RF	0,78	0,74

TABLE 14 – Comparaison des méthodes de classification sur la zone de développement

La classification non supervisée est celle qui donne les moins bons résultats avec une OA de 0,56 et un indice kappa de 0,45.

De plus, la méthode ISODATA nécessite une grande connaissance du terrain puisque la combinaison des différentes classes créées par l'algorithme est faite par l'utilisateur avec son interprétation propre. Elle n'est donc pas généralisable, car elle comporte toujours une grande part de subjectivité. À large échelle non plus il n'est pas possible de la développer, car elle demande une trop grande connaissance de la zone d'étude. De plus la combinaison des classes demande beaucoup de temps.

Son avantage est par contre qu'elle différencie bien les zones forestières des zones non-forestières. Elle peut donc servir pour la création de masques. Mais encore une fois, cela nécessite de fusionner plusieurs classes, une zone agricole n'étant pas similaire à une savane ou à une zone d'habitation, c'est à l'utilisateur de procéder par interpolation pour les regrouper.

Le dernier avantage potentiel de cette méthode est qu'elle ne nécessite pas de zones d'entraînements, ce qui procure un gain de temps et une campagne de terrain non indispensable. Cependant, ce dernier point est à relativiser au vu de ce qui a été dit précédemment. Pour comprendre la signification des classes créées automatiquement, une connaissance du terrain est quand même indispensable.

Une méthode hiérarchique permet d'augmenter de manière nette la qualité d'une classification via une approche par pixel. En effet, cette méthode va classer des catégories avec des spectres de réflectance relativement proches. Dès lors, les différences entre ces classes seront plus marquées et la distinction plus facile. Cependant, les approches par pixels gardent toujours un fort effet poivre et sel qui diminue les métriques de qualité par les contaminations des zones de validation. L'utilisation de filtre majoritaire permet de diminuer légèrement cet effet, mais n'augmente que très peu la qualité de la classification.

La classification supervisée via une approche par objet est la classification qui donne les meilleures métriques de qualité avec une OA de 0,78 et un indice kappa de 0,74. De plus, elle est la classification qui, à l'exception des forêts marécageuses et des forêts inondables, donne les F-scores les plus élevés en moyenne. La qualité de la classification dépend fortement de la qualité de la segmentation. Si les objets sont bien identifiés et cohérents entre eux, la qualité finale ne sera que plus importante.

D'une manière générale dans les classifications, les zones non forestières sont mieux classées que les zones forestières. Ces différences de performances peuvent s'expliquer par plusieurs éléments :

- La séparation entre zones couvertes d'arbres et les autres occupations de sol est bien réalisée. Cette distinction s'explique principalement, car les faciès et les réflectances des deux milieux sont fort différents (Figure 50).
- La classification des zones non forestières dans les différentes catégories présente généralement des bons F-scores. Cela peut s'expliquer par les différences fortes de réflectances entre ces occupations du sol.
- Enfin concernant les zones forestières, et contrairement aux autres occupations du sol, il existe en leur sein une variabilité plus grande de réflectance. Les forêts au sein d'une même classe peuvent renvoyer à l'échelle du pixel des signaux forts différents. Cela est due à la mixité d'espèce ou de structure en leur sein et que les tailles des pixels sont inférieures à la surface des couronnes. Une même couronne peut donc présenter des réflectances forts différents car une partie sera illuminée tandis que l'autre est ombragée (Figure 50).

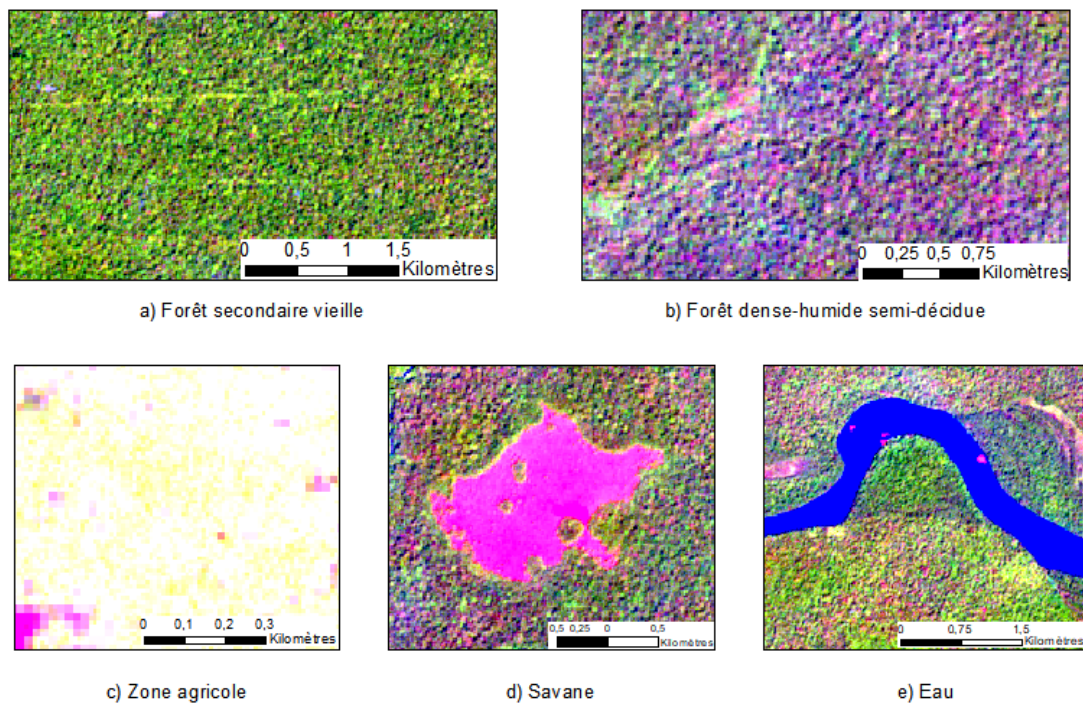


FIGURE 50 – Variabilité de réflectance au sein de différentes classes d'occupation du sol

5.3 Production cartographique

La méthode de classification retenue, pour la cartographie des types forestiers et de l'occupation du sol sur la zone d'étude, est la classification supervisée via une approche par objet. Elle est celle qui permet d'obtenir les meilleurs résultats sur la zone de développement, aussi bien au niveau des métriques que visuellement.

La cartographie obtenue est présentée à la figure 51 et à la figure 52. La matrice de confusion est présentée en annexe (Tableau A.6) et son analyse au tableau 15.

Classes	1	2	3	5	6	8	9	10	11	12	13	14
Producer accuracy	0,90	0,55	0,97	0,34	1,00	1,00	0	0,64	0,80	1,00	0,86	0,93
User accuracy	0,97	0,84	1,00	0,94	0,19	0,26	0	0,61	1,00	1,00	0,75	1,00
F-Score	0,93	0,66	0,99	0,50	0,32	0,41	0	0,62	0,89	1,00	0,80	0,96
Overall accuracy							0,74					
Indice Kappa							0,70					

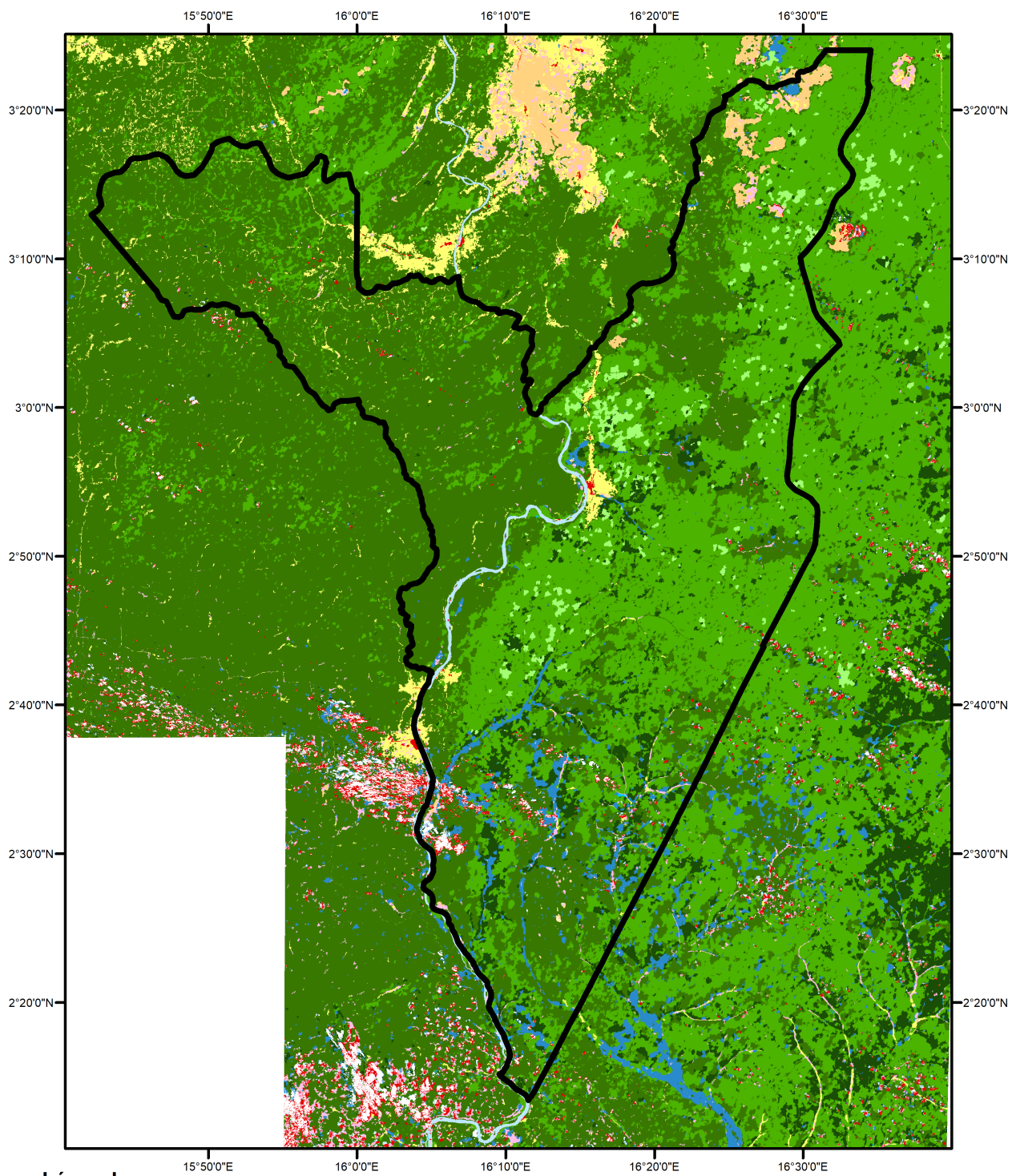
TABLE 15 – Analyse de la classification sur la zone d'étude

La classification finale présente une OA de 0,74 et un indice kappa de 0,70. Elle peut donc être qualifiée de bonne qualité. Cependant les F-scores ne sont pas équivalents pour toutes les catégories.















Pour les zones forestières, Les F-scores sont particulièrement élevés pour 2 catégories : les forêts sempervirentes (classe 1, F-score = 0,93) et les forêts marécageuses (classe 3, F-score = 0,99). Pour les autres catégories forestières, les F-scores varient entre 0,32 et 0,66, à l'exception des forêts inondables (classe 9), pour lesquelles aucune zone de validation n'a été classée dans cette catégorie. Pour les forêts secondaires jeunes (classe 6) et les forêts ripicoles hautes, les PA sont toutes les deux maximums avec un score de 1 mais les UA sont assez faibles (respectivement de 0,19 et 0,26). La classification a donc tendance à surestimer ces catégories en y classant des zones qui ne devraient pas s'y trouver.

Les zones non forestières (zone d'habitation, eau, savane et zone agricole), à l'exception des clairières salines (classe 10, F-score = 0,62) sont très bien classées avec des F-scores supérieurs ou égaux à 0,80. Les savanes se confondent cependant parfois avec les clairières salines.

Les APDS semblent être divisées en 2 grandes parties : à l'ouest de la rivière Sangha se trouve une majorité de forêts denses humides semi-décidues tandis que la partie à l'est est dominée par des forêts secondaires vieilles. Visuellement, et en accord avec la matrice de confusion, les forêts inondables sont absentes de la classification. Les zones non forestières comme la rivière Sangha, les zones agricoles et les savanes ressortent bien et sont facilement identifiables. Enfin, le masque nuageux appliqué ne prend pas en compte l'entièreté des nuages et de leurs ombres. Ainsi une partie de ces nuages est classée comme des zones d'habitation.



Légende

- | | |
|---|---|
|  Forêt sempervirente |  Clairière saline |
|  Forêt dense humide semi-décidue |  Zone d'habitation |
|  Forêt marécageuse |  Eau |
|  Forêt secondaire vieille |  Savane |
|  Forêt secondaire jeune |  Zone agricole |
|  Forêt ripicole haute |  Couverture Nuageuse |
|  Forêt inondable |  APDS |

Système de coordonnées: GCS WGS 1984
 Datum: WGS 1984
 Unités: Degree
 Auteur: D.Lessire (UCLouvain - 2019)
 Images: Sentinel-2, 27/12/2018



FIGURE 51 – Cartographie des types forestiers et de l’occupation du sol au sein des APDS (RCA)

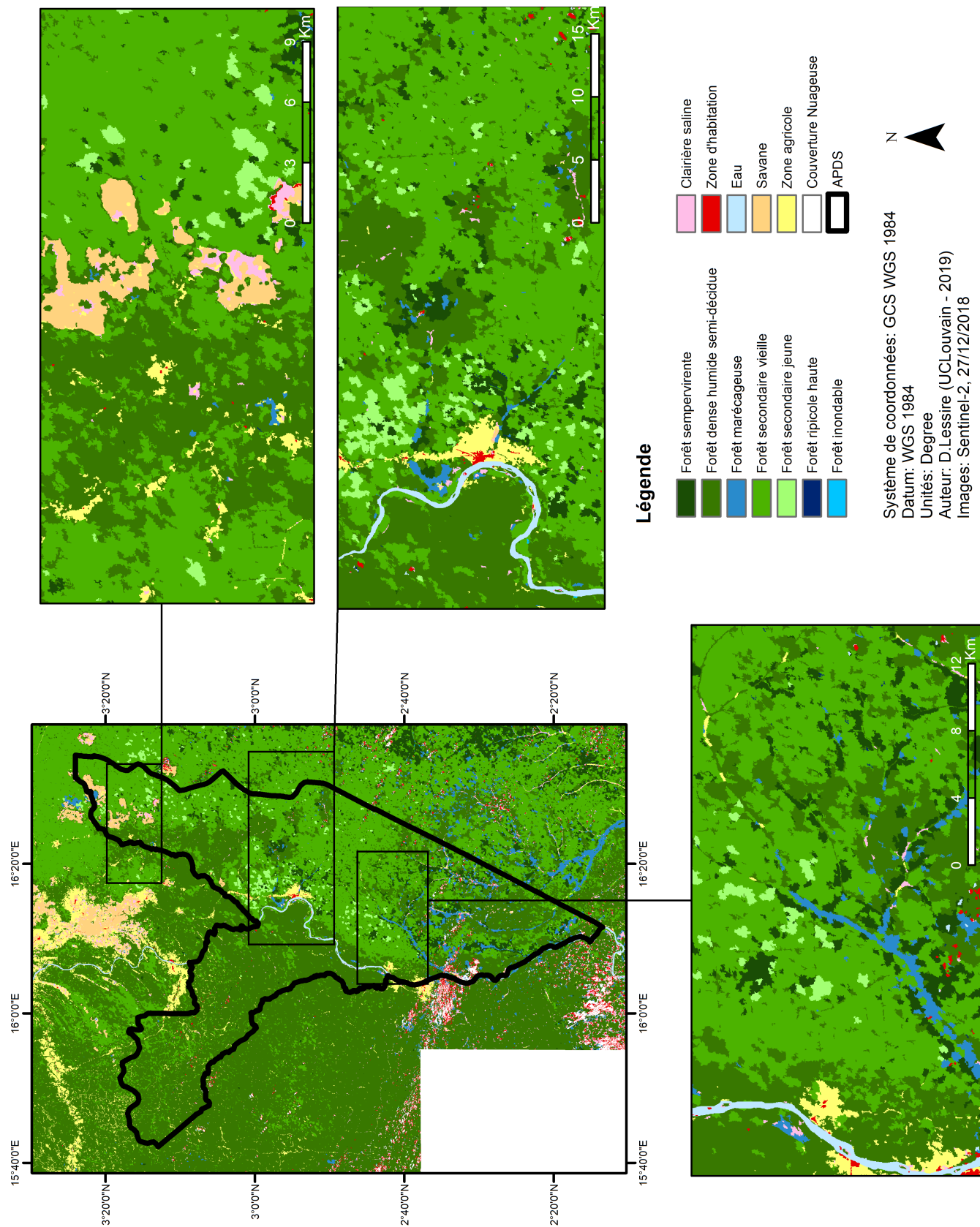


FIGURE 52 – Cartographie des types forestiers et de l'occupation du sol au sein des APDS (RCA) - Zoom

L'absence des forêts inondables dans la classification relève de plusieurs facteurs. Tout d'abord de la catégorie elle-même. Les forêts inondables rassemblent des zones forestières avec des faciès de végétation très différents. Leur point commun est d'être inondées durant une partie de l'année. Or cette caractéristique n'est pas prise en compte dans cette classification. Et ce, car les images utilisées sont des images prises durant la saison sèche, comme recommandé par Erinjery et al. (2018). L'avantage est que cela permet de différencier plus facilement les forêts décidues des forêts sempervirentes. Ce choix d'image est aussi limité de par la couverture nuageuse. La saison sèche est la seule période avec des couvertures assez faibles pour pouvoir rendre utilisables les images.

Les forêts secondaires jeunes sont dispersées dans une grande partie de la zone. Cela ne semble pas logique puisque ce qui définit ces forêts est une exploitation récente. Elles se trouvent donc normalement proches des zones d'habitation. Cette constatation est en adéquation avec les observations faites à partir de la matrice de confusion. Dans de nombreux cas, ces forêts sont confondues avec des forêts secondaires vieilles. Cela peut sembler logique, car les deux types forestiers sont liés. En effet les forêts secondaires jeunes seront amenées à évoluer vers des forêts secondaires vieilles. Il n'y a pas donc de transition brute entre les deux types, mais plutôt un gradient.

5.3.1 Développer la méthode à plus large échelle

Dans le but d'évaluer la capacité à généraliser une méthode de classification basée sur les objets, une comparaison est réalisée entre les qualités de ces classifications à des échelles spatiales différentes.

La première classification est celle réalisée sur la zone de développement. Les métriques de la qualité sont une OA de 0,78 et un indice kappa de 0,74. La deuxième classification est celle réalisée sur l'ensemble de la zone d'étude et donne une OA de 0,74 et un indice kappa de 0,70. Les deux classifications utilisent le même jeu de données comme zones d'entraînement, mais des zones de validation différentes. La classification sur la zone de développement se montre donc légèrement plus performante que celle obtenue pour l'ensemble des APDS.

L'effet de l'extension de la classification à l'ensemble de la zone se répercute différemment sur les classes forestières. Les forêts sempervirentes (+0,43), les forêts marécageuses (+0,38) et les forêts ripicoles hautes (+0,41) ont leurs F-scores qui augmentent. Pour la dernière catégorie, le F-score pour la classification sur la zone de développement est de 0. Lors de la classification sur toute la zone d'étude les forêts ripicoles sont donc mieux classées. Les autres classes forestières ont leurs F-scores qui diminuent. Pour les forêts secondaires vieilles, la diminution du F-score (de 0,97 à 0,50) est due à une forte diminution du PA (de 0,95 à 0,34). Des contaminations apparaissent donc dans les zones de validation. Les forêts denses humides semi-décidues et les forêts secondaires jeunes ont toutes deux aussi leurs F-score qui diminuent, respectivement de 0,18 et 0,24. A noter que les forêts secondaires jeunes ont une PA de 1 mais une UA de 0,19.

Pour les zones non forestières, d'une manière générale, les F-scores diminuent légèrement lors de l'extension de la zone de classification, à l'exception des zones d'habitation (+0,07). Les savanes sont la classe avec la plus forte baisse (-0,19) mais

gardent un F-score haut (F-score = 0,80). Cette diminution du F-score est due à des zones de savane classées comme des clairières salines.

La différence de qualité des classifications peut provenir de plusieurs éléments. Tout d'abord, les zones d'entraînement sont ici les mêmes dans les deux classifications. Or elles se situent toutes dans la zone de développement. Par conséquent, si elles sont bien représentatives de la zone de développement, elles ne le sont pas forcément pour la zone d'étude. De plus, il a été montré par Colditz (2015) l'importance de garder une certaine proportionnalité entre les surfaces des zones d'entraînement et la réalité du terrain, pour chacune des classes. Or si c'est le cas pour la zone de développement, cela n'est plus garanti pour l'ensemble de la zone d'étude.

Ensuite, les classes de végétation ont été développées sur base des observations réalisées dans la zone de développement. Ces classes ne correspondent peut être pas parfaitement à la réalité de l'ensemble de la zone d'étude. Une autre catégorie, non relevée, peut exister et fausser la classification. L'extension à plus large échelle d'une classification est faisable uniquement dans le cas où les catégories de végétations restent similaires. Cela limite donc l'extension spatiale. Ceci est néanmoins à nuancer par le fait que la validation a été réalisée sur base de photo-interprétation. Or durant celle-ci aucune zone n'a semblé fort différente des classes préétablies.

Finalement, la différence de qualité de la classification peut, en partie, provenir des zones de validations. Si dans le cas de la zone de développement, ces zones ont été visitées durant la campagne de terrain, ce n'est pas le cas pour les zones de validation de la zone d'étude. Ces dernières sont uniquement basées sur la photo-interprétation et peuvent comporter une partie d'erreur.

5.4 Dynamique Agricole

Les résultats des calculs de surfaces non forestières au sein des bandes culturelles sont présentés au tableau 16. Pour chaque secteur, les surfaces agricoles sont cartographiées de la même manière que présentée à la Figure 53).

En 2013 les zones agricoles au sein des APDS couvraient 2055 ha contre 2284 ha en 2017, soit une augmentation de près de 3,3% chaque année sur l'ensemble des APDS. Seul un secteur enregistre une diminution des surfaces agricoles entre 2013 et 2017, celui de Moussapoula Nord. En moins de quatre ans, sa surface a diminué de près de 15 ha, soit près de 18,4%. Pour les secteurs en augmentation, les dynamiques ne sont pas similaires partout. L'augmentation la plus importante est dans secteur Yadoumbé, avec une augmentation annuelle de 23,8 ha (+7,6%). Au contraire, dans le secteur de Moussapoula Sud, l'augmentation annuelle n'est que de 1,4 ha soit 0,7%.

Dans les APDS, 4041 ha sont affectés à des zones agricoles et 52,1% de cette surface est utilisée comme tel. Cependant 7,8% des surfaces agricoles effectives s'étendent en dehors de ces zones. Ces extensions ne concernent pas toutes les bandes culturelles. Quatre bandes sont affectées : Babongo Nord, Moussapoula Nord, Bayanga Centre et Bomandjokou. C'est dans ce dernier secteur que la problématique est la plus importante puisque 35% de la surface cultivée est hors la zone d'affectation prévue à cet effet, soit près de 90 ha. Il est aussi la zone d'affectation avec le moins de surface encore non utilisée. Dans un cas où les expansions de surfaces

agricoles sont restreintes à la zone d'affectation, cette dernière sera totalement occupée endéans les 2 ans. Or actuellement, la majeure partie de l'extension des surfaces agricoles dans ce secteur se situe en dehors de ce plan (+ 13,5 ha/an hors de la zone contre +1,3 ha/an dans la zone).

Un autre secteur pour lequel la majorité de l'augmentation de surfaces agricole se fait en dehors de la zone prévue à cet effet est celui de Bayanga Centre. Si sa croissance annuelle est de 15,5 ha, 13,7 ha, soit 88% de l'augmentation, se fait en dehors de la zone prévue.

Dans un scénario avec des évolutions similaires à celles actuelles, le secteur de Yadoumbe risque d'être problématique dans le moyen terme. Les surfaces agricoles respectent actuellement le plan d'affectation, mais au rythme actuel, cette surface sera totalement occupée d'ici 11 ans.

Les limites des bandes culturales datent de 1992. Les dernières modifications datent de 2005 soit 14 ans. La démographie a depuis évolué et la pression sur les zones de la réserve s'est accrue. Repenser les bandes culturales et leurs limites peut être un défi futur pour les gestionnaires des APDS.






Secteur	Surface affectée (ha)	Surface occupée (ha)	Évolution par an (ha)	Surface hors zone d'affectation (ha)	Surface d'affectation restante (ha)
Babongo Nord	886,4	329,0	+2,8 (+0,9%)	5,6	563,0 (63%)
Babongo Sud	140,5	18,8	+1,0 (+6,5%)	0	121,7 (87%)
Moussapoula Nord	233,3	64,6	-4,2 (-5,3%)	1,1	169,8 (73%)
Moussapoula Sud	376,0	195,0	+1,4 (+0,7%)	0	181,0 (48%)
Bayanga Centre	764,5	710,4	+15,5 (+2,4%)	82,6	136,7 (18%)
Yadoumbe	676,3	407,2	+23,8 (+7,6%)	0	269,1 (40%)
Linjombo	771,6	302,1	+12,7 (+4,7%)	0	469,5 (61%)
Bomandjokou	192,8	257,3	+14,8 (+6,7%)	89,5	25,0 (13%)
Total	4041,4	2284,4	67,8 (+3,3%)	178,8	1935,8 (48%)

TABLE 16 – Évolution des surfaces des zones agricoles au sein des APDS

Evolution des surfaces agricoles au sein des APDS (RCA)

Secteur Bayanga Centre, 2013-2018

Légende

-  Surface agricole - 2013
-  Surface agricole - hors zone 2013
-  Surface agricole - extension 2018
-  Surface agricole - hors zone 2018
-  Bande culturale - Secteur Bayanga Centre

	2013	2017
Surface agricole (ha)	620,8	627,8
Surface agricole hors zone (ha)	27,7	82,6
Total (ha)	648,5	710,4

Surface Secteur (ha)	764,5
Evolution par an (ha)	+ 2,8
Surface d'affectation restante (ha)	136,2
Surface d'affectation restante (%)	17,9

Système de coordonnées: GCS WGS 1984

Datum: WGS 1984

Unités: Degree

Auteur: D.Lessire (UCLouvain, 2019)

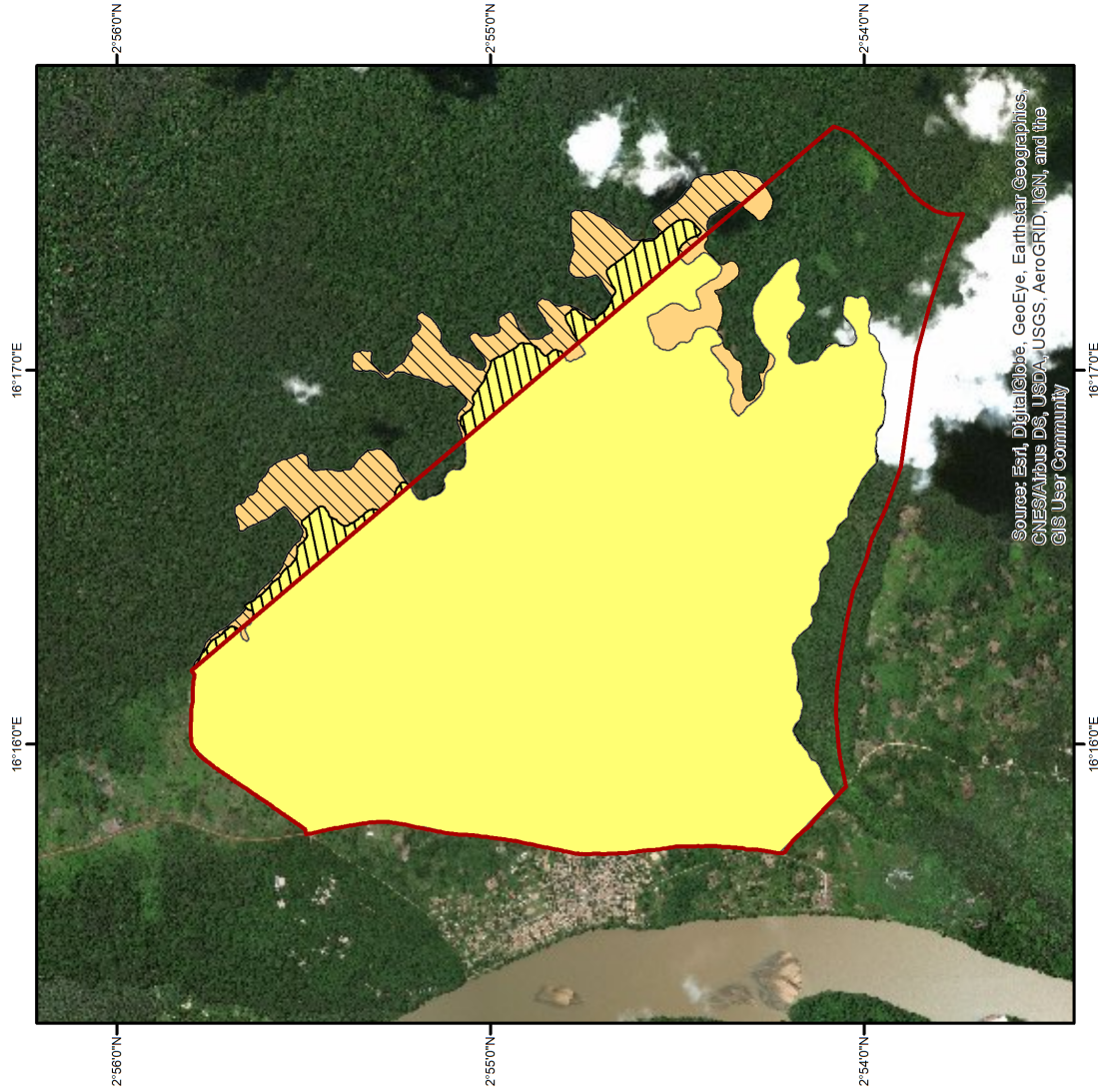


FIGURE 53 – Évolution des surfaces agricoles dans le secteur de Bayanga centre

Conclusion et perspectives

Le but de ce mémoire était de développer une méthode de classification des types forestiers d'Afrique centrale et d'en évaluer sa qualité. Cette étude a été réalisée sur les Aires Protégées de Dzanga-Sangha dans le sud-ouest de la RCA. Une cartographie de l'occupation du sol est aussi réalisée.

La télédétection offre un moyen d'obtenir des informations de manière rapide, régulière et à faible coût. Les plateformes Sentinel-2 lancées dans le cadre du programme Copernicus fournissent des images à hautes résolutions (10-20m) avec 10 bandes spectrales et un temps de revisite de 5 jours. Elles constituent donc un atout majeur pour la surveillance des forêts tropicales.

Une campagne de terrain a permis de récolter des placettes d'échantillonnage qui ont servi à développer un modèle de classification ainsi qu'à l'évaluation de sa qualité. De plus, une typologie des types forestiers présents au sein de la réserve a été mise en place. 7 types forestiers majeurs et 5 occupations du sol en milieu non forestier ont été identifiés et décrits. Une clé de photo-interprétation sur base de composition colorée (RGB = B07-B11-B05) a été mise en place et utilisée pour identifier les types forestiers.

Quatre méthodes de classification ont été réalisées sur une partie seulement de la zone d'étude. Une classification non supervisée via l'algorithme ISODATA et trois classifications supervisées via l'algorithme RF. Dans ces dernières, deux ont été appliquées via une approche par pixel (de manière hiérarchique ou non) et une via une approche par objet. Sur la zone de développement, la classification qui présente les meilleures métriques de qualité est la classification supervisée via une approche par objet. L'OA générale obtenue est de 0,78 et l'indice kappa de 0,74. C'est donc cette méthode qui est retenue pour la cartographie de l'ensemble de la zone d'étude.

Une cartographie de la zone d'étude a ensuite été réalisée sur base de cette méthode. La classification finale obtient une OA de 0,74 et un indice kappa de 0,70. Les zones de validation pour cette classification ont été définies sur base de photo-interprétation.

Toutes les classes de végétation ne présentent pas des qualités de classification similaires. Ainsi si les zones non forestières sont généralement très bien classées, ce n'est pas le cas de toutes les zones forestières. Les forêts inondables ne sont, par exemple, pas reprises dans les classifications et ce principalement due à la période de détection des images utilisée.

La télédétection et notamment l'utilisation de Sentinel-2 révèle un potentiel très intéressant pour la classification forestière. Néanmoins ce capteur est contraint par la couverture nuageuse, ce qui, en région tropicale, limite fortement la période de détection. Or cette période peut s'avérer cruciale pour la différenciation de certaines classes forestières, notamment celles liées aux régimes hydriques comme les forêts inondables. L'utilisation complémentaire d'images radar, avec Sentinel-1, pourrait permettre de se libérer de cette limitation et d'améliorer la classification de certaines classes.

Références

- Aboras, M., Amasha, H. and Ibraheem, I. (2015), ‘Early detection of melanoma using multispectral imaging and artificial intelligence techniques’, p. 6.
- Alrababah, M. A. and Alhamad, M. N. (2006), ‘Land use/cover classification of arid and semi-arid Mediterranean landscapes using Landsat ETM’, *International Journal of Remote Sensing* **27**(13), 2703–2718.
- Alvarez-Añorve, M., Quesada, M. and de la Barrera, E. (2008), Remote Sensing and Plant Functional Groups : Physiology, Ecology, and Spectroscopy in Tropical Systems, in M. Kalacska and Arturo Sanchez-Azofeifa, eds, ‘Hyperspectral Remote Sensing of Tropical and Sub-Tropical Forests’, CRC Press, pp. 27–45.
- APDS (2016), Plan de gestion des Aires Protégées de Dzanga Sangha 2016 -2020, Technical report, Le Ministre des Eaux, Forêts, Chasse et Pêche.
- Asner, G. P., Ustin, S. L., Townsend, P. A., Martin, R. E. and Chadwick, K. D. (2015), ‘Forest Biophysical and Biochemical Properties from Hyperspectral and LiDAR Remote Sensing’, p. 20.
- Baccini, A., Walker, W., Carvalho, L., Farina, M., Sulla-Menashe, D. and Houghton, R. A. (2017), ‘Tropical forests are a net carbon source based on aboveground measurements of gain and loss’, *Science* **358**(6360), 230–234.
- Balzter, H., Cole, B., Thiel, C. and Schullius, C. (2015), ‘Mapping CORINE Land Cover from Sentinel-1A SAR and SRTM Digital Elevation Model Data using Random Forests’, *Remote Sensing* **7**(11), 14876–14898.
- Belgiu, M. and Drăguț, L. (2016), ‘Random forest in remote sensing : A review of applications and future directions’, *ISPRS Journal of Photogrammetry and Remote Sensing* **114**, 24–31.
- Blaschke, T. (2010), ‘Object based image analysis for remote sensing’, *ISPRS Journal of Photogrammetry and Remote Sensing* **65**(1), 2–16.
- Boulvert, Y. (1986), ‘Carte phytogéographique de la République centrafricaine à l’échelle de 1 :1 000 000’.
- Breiman, L. (2001), ‘Random Forests’, *Machine Learning* (45).
- Centre canadien de télédétection (n.d.), ‘Notions fondamentales de télédétection’.
- Climate-Data.org (n.d.), ‘Climat Bayanga’, <https://fr.climate-data.org/afrique/republique-centrafricaine/sangha-mbaere/bayanga-422887/#climate-table>.
- Colditz, R. (2015), ‘An Evaluation of Different Training Sample Allocation Schemes for Discrete and Continuous Land Cover Classification Using Decision Tree-Based Algorithms’, *Remote Sensing* **7**(8), 9655–9681.
- Congalton, R. G. (1991), ‘A review of assessing the accuracy of classifications of remotely sensed data’, *Remote Sensing of Environment* **37**(1), 35–46.

- Convention on Biological Diversity (2006), 'Indicative definitions taken from the Report of the ad hoc technical expert group on forest biological diversity', <https://www.cbd.int/forest/definitions.shtml>.
- Coops, N. and Culvenor, D. (2000), 'Utilizing Local Variance of Simulated High Spatial Resolution Imagery To Predict Spatial Pattern of Forest Stands', *Remote Sensing of Environment* **71**(3), 248–260.
- de Wasseige, C. and Defourny, P. (2002), 'Retrieval of tropical forest structure characteristics from bi-directional reflectance of SPOT images', *Remote Sensing of Environment* **83**(3), 362–375.
- de Wasseige, C., Flynn, J., Louppe, D., Hiol Hiol, F. and Mayaux, P. (2014), *The Forests of the Congo Basin - State of the Forest 2013*, COMIFAC, Weyrich, Belgique.
- Delhez, B. (2019), 'Tropical forest type mapping by remote sensing using RADAR and optical data from Sentinel-1 and -2'. UCLouvain, Mémoire de fin d'étude.
- Erinjery, J. J., Singh, M. and 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.
- Ernst, C., Verhegghen, A., Mayaux, P., Hansen, M. and Defourny, P. (2012), 'Cartographie du couvert forestier et des changements du couvert forestier en Afrique centrale'.
- ESA (2019), 'Sentinel-2 - Missions - Sentinel Online', <https://sentinel.esa.int/web/sentinel/missions/sentinel-2>.
- Franklin, S. E., Hall, R. J., Moskal, L. M., Maudie, A. J. and Lavigne, M. B. (2000), 'Incorporating texture into classification of forest species composition from airborne multispectral images', *International Journal of Remote Sensing* **21**(1), 61–79.
- Ghosh, A. and Joshi, P. (2014), 'A comparison of selected classification algorithms for mapping bamboo patches in lower Gangetic plains using very high resolution WorldView 2 imagery', *International Journal of Applied Earth Observation and Geoinformation* **26**, 298–311.
- Gislason, P. O., Benediktsson, J. A. and Sveinsson, J. R. (2006), 'Random Forests for land cover classification', *Pattern Recognition Letters* **27**(4), 294–300.
- Gond, V., Betbeder, J., Fayolle, A., Viennois, G., Cornu, G., Réjou-Méchain, M., Baghdadi, N., Benedet, F., Doumenge, C. and Gourlet-Fleury, S. (2015), 'New insights in tropical forest diversity mapping in Central Africa using low'.
- Gond, V., Fayolle, A., Pennec, A., Cornu, G., Mayaux, P., Camberlin, P., Doumenge, C., Fauvet, N. and 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.

- Haralick, R. M., Shanmugam, K. and Dinstein, I. (1973), 'Textural Features for Image Classification', *IEEE Transactions on Systems, Man, and Cybernetics SMC-3*(6), 610–621.
- Harris, D. J. (2002), *The Vascular Plants of the Dzanga-Sangha Reserve, Central African Republic*, Vol. 23 of *Scripta Botanica Belgica*, National Botanic Garden of Belgium.
- Hubbel, S. P. (2001), *The Unified Neutral Theory of Biodiversity and Biogeography*, Princeton University Press.
- Immitzer, M., Vuolo, F. and Atzberger, C. (2016), 'First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe', *Remote Sensing* **8**(3), 166–.
- Irvin, B. J., Ventura, S. J. and Slater, B. K. (1997), 'Fuzzy and isodata classification of landform elements from digital terrain data in Pleasant Valley, Wisconsin', *Geoderma* **77**(2-4), 137–154.
- Laurin, G. V., Puletti, N., Hawthorne, W., Liesenberg, V., Corona, P., Papale, D., Chen, Q. and 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.
- Mayaux, P. and Achard, F. (1998), 'Suivi à long terme de la végétation d'Afrique Centrale à partir de l'imagerie satellitaire', *Séminaire FORAFRI de Libreville* p. 14.
- Meunier, Q., Moumbogou, C. and Doucet, J.-L. (2015), *Les arbres utiles du Gabon*, Les Presses agronomiques de Gembloux, Gembloux, Belgium. OCLC : ocn915161076.
- N'Goran, K. P., Ndomba, D. L. and Beukou, G. B. (2016), Rapport de l'inventaire des grands et moyens mammifères dans le segment RCA du paysage national de la Sangha, Technical report, WWF, APDS.
- Nziengui, M., Tchindjang, M., Feizoure, J. and Zanre, N. (2013), 'Apport de la télédétection et des SIG pour la caractérisation des habitats', *Revue Scientifique et Technique Forêt et Environnement du Bssin du Congo* **1**, 25–41.
- Parmentier, I., Stévant, T. and Hardy, O. J. (2005), 'The inselberg flora of Atlantic Central Africa. I. Determinants of species assemblages : Spatial patterns of inselberg flora', *Journal of Biogeography* **32**(4), 685–696.
- Pascal, J. (2003), 'Notions sur les structure et dynamique des forêts tropicales humides.', *Revue Forestière Française* (sp), 118.
- Pelletier, C., Valero, S., Inglada, J., Champion, N. and Dedieu, G. (2016), 'Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas', *Remote Sensing of Environment* **187**, 156–168.

- Pitman, N. C. A., Terborgh, J. W. and Silman, M. R. (2001), 'DOMINANCE AND DISTRIBUTION OF TREE SPECIES IN UPPER AMAZONIAN TERRA FIRME FORESTS', **82**(8), 17.
- Pontius, R. G. and 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.
- Réjou-Méchain, M., Péliissier, R., Gourlet-Fleury, S., Couteron, P., Nasi, R. and Thompson, J. D. (2008), 'Regional variation in tropical forest tree species composition in the Central African Republic : An assessment based on inventories by forest companies', *Journal of Tropical Ecology* **24**(6), 663–674.
- Rodriguez-Galiano, V., Ghimire, B., Rogan, J., Chica-Olmo, M. and 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.
- Rozenstein, O. and Karnieli, A. (2011), 'Comparison of methods for land-use classification incorporating remote sensing and GIS inputs', *Applied Geography* **31**(2), 533–544.
- Runkle, J. R. (1992), *Guidelines and Sample Protocol for Sampling Forest*, General Technical Report, united states department of agriculture edn, Pacific Northwest Research Station.
- Société Française de Photogrammétrie et Télédétection (n.d.), 'Principe de l'imagerie hyperspectrale'.
- Stumpf, A. and Kerle, N. (2011), 'Object-oriented mapping of landslides using Random Forests', *Remote Sensing of Environment* **115**(10), 2564–2577.
- UNESCO (2019), 'Trinational de la Sangha', <https://whc.unesco.org/fr/list/1380/>.
- Verhegghen, A. and Defourny, P. (2011), Phenology of the natural vegetation : A land cover specific approach for a reference dataset in Central Africa, *in* '2011 6th International Workshop on the Analysis of Multi-Temporal Remote Sensing Images (Multi-Temp)', IEEE, Trento, Italy, pp. 257–260.
- Wulder, M. A., Franklin, S. E. and Lavigne, M. B. (1996), 'High Spatial Resolution Optical Image Texture for Improved Estimation of Forest Stand Leaf Area Index', *Canadian Journal of Remote Sensing* **22**(4), 441–449.
- Xiong, J., Thenkabail, P., Tilton, J., Gumma, M., Teluguntla, P., Oliphant, A., Congalton, R., Yadav, K. and Gorelick, N. (2017), 'Nominal 30-m Cropland Extent Map of Continental Africa by Integrating Pixel-Based and Object-Based Algorithms Using Sentinel-2 and Landsat-8 Data on Google Earth Engine', *Remote Sensing* **9**(10), 1065.
- Yoneda, T., Nishimura, S., Fujii, S. and Mukhtar, E. (2009), 'Tree guild composition of a hill dipterocarp forest in West Sumatra, Indonesia', *Tropics* **18**(3), 143–154.

Zelazowski, P., Malhi, Y., Huntingford, C., Sitch, S. and Fisher, J. B. (2011), 'Changes in the potential distribution of humid tropical forests on a warmer planet', *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences* **369**(1934), 137–160.

A Matrices de confusion

		Prédiction								Total
		1	2	3	5	6	8	9	15	
Référence	1	336	42	43	61	0	12	0	0	494
	2	355	483	821	67	361	19	0	11	2117
	3	350	69	350	0	40	13	0	20	842
	5	17	121	1	621	7	0	0	40	807
	6	440	116	307	34	429	13	0	5	1344
	8	42	0	2	0	0	41	0	0	85
	9	414	202	62	87	280	33	0	1	1079
	15	0	118	223	0	0	0	0	3884	4225
	Total	1954	1151	1809	870	1117	118	0	3961	10993

TABLE A.1 – Matrice de Confusion de la classification par Isodata sur la zone de développement

		Prédiction											Total	
		1	2	3	5	6	8	9	10	11	12	13		14
Référence	1	152	0	0	54	262	26	0	0	0	0	0	0	494
	2	63	217	135	253	1370	4	3	59	0	0	0	13	2117
	3	0	13	137	0	539	68	0	49	0	0	0	36	842
	5	46	0	0	689	68	4	0	0	0	0	0	0	807
	6	9	6	12	0	1295	20	0	2	0	0	0	0	1344
	8	13	0	0	0	28	43	1	0	0	0	0	0	85
	9	39	3	3	0	987	47	0	0	0	0	0	0	1079
	10	0	0	32	0	0	4	0	141	21	0	78	63	339
	11	0	0	0	0	0	0	0	0	136	0	5	26	167
	12	0	0	0	0	0	0	0	0	0	2435	0	0	2435
	13	0	2	3	0	0	0	0	2	0	0	988	0	995
	14	0	22	0	0	20	0	0	12	0	0	0	235	289
	Total	322	263	322	996	4569	216	4	265	157	2435	1071	373	10993

TABLE A.2 – Matrice de Confusion de la classification supervisée non hiérarchique via une approche par pixel - 12 classes

		Prédiction				Total
		1	2	3	4	
Référence	1	1493	223	397	4	2117
	2	525	1165	461	0	2151
	3	209	462	1746	79	2496
	4	112	0	7	4032	4151
Total	2339	1850	2611	4115	10915	

TABLE A.3 – Matrice de Confusion de la classification supervisée hiérarchique via une approche par pixel - 4 classes

		Prédiction												Total	
		1	2	3	5	6	8	9	10	11	12	13	14		
Référence	1	215	12	39	66	25	24	113	0	0	0	0	0	494	
	2	32	1869	35	27	66	4	70	10	0	0	0	4	2117	
	3	0	105	566	0	1	38	50	52	0	0	0	30	842	
	5	23	20	0	729	31	4	0	0	0	0	0	0	807	
	6	43	620	228	0	321	23	109	0	0	0	0	0	1344	
	8	14	0	6	0	0	44	21	0	0	0	0	0	85	
	9	96	124	335	4	152	47	321	0	0	0	0	0	1079	
	10	0	2	21	0	0	1	0	177	20	0	96	22	339	
	11	0	0	0	0	0	0	0	25	137	0	5	0	167	
	12	0	0	0	0	0	0	0	0	0	2435	0	0	2435	
	13	0	22	2	0	0	0	0	2	0	0	969	0	995	
	14	0	56	0	0	0	0	0	11	0	0	0	222	289	
	Total		423	2830	1232	826	596	185	684	277	157	2435	1070	278	10993

TABLE A.4 – Matrice de Confusion de la classification supervisée hiérarchique via une approche par pixel - 12 classes

		Prédiction												Total	
		1	2	3	5	6	8	9	10	11	12	13	14		
Référence	1	190	59	0	17	22	0	59	0	0	0	0	0	347	
	2	0	1660	0	21	60	0	0	0	0	0	0	0	1741	
	3	12	0	293	0	281	0	9	71	0	0	0	0	666	
	5	0	0	0	683	0	0	0	0	0	0	0	0	683	
	6	0	319	0	0	774	0	0	0	0	0	0	0	1093	
	8	20	13	0	0	13	0	6	0	0	0	0	0	52	
	9	196	159	0	0	534	0	7	0	0	0	0	0	896	
	10	0	12	0	0	0	0	0	201	52	0	0	0	265	
	11	0	0	0	0	0	0	0	0	119	0	0	2	121	
	12	0	0	0	0	0	0	0	0	0	1977	0	0	1977	
	13	0	7	0	0	0	0	0	0	0	0	793	1	801	
	14	0	0	0	0	0	0	0	0	0	0	0	217	217	
	Total		418	2229	293	721	1684	0	81	272	171	1977	793	220	8859

TABLE A.5 – Matrice de Confusion de la classification supervisée via une approche par objet - Zone de développement

		Prédiction												Total
		1	2	3	5	6	8	9	10	11	12	13	14	
Référence	1	998	3	0	30	0	0	0	0	0	0	0	0	1031
	2	0	1748	30	311	0	0	0	0	0	0	0	2	2091
	3	0	0	1167	0	0	0	0	0	0	0	0	0	1167
	5	0	77	0	1196	0	0	0	0	0	0	0	0	1273
	6	0	147	0	903	249	0	0	0	0	0	0	0	1299
	8	114	26	0	0	0	48	0	0	0	0	0	0	188
	9	0	1170	0	1100	0	0	0	0	0	0	0	0	2270
	10	0	25	0	0	0	0	0	567	78	0	195	69	934
	11	0	0	0	0	0	0	0	0	525	0	0	0	525
	12	0	0	0	0	0	0	0	0	0	4372	0	0	4372
	13	0	0	0	0	0	0	0	322	51	0	1161	15	1549
	14	0	0	0	0	0	0	0	0	0	0	0	1177	1177
	Total	1112	3196	1197	3540	249	48	0	889	654	4372	1356	1263	17876

TABLE A.6 – Matrice de Confusion de la classification supervisée via une approche par objet - Zone d'étude

B Masque nuageux

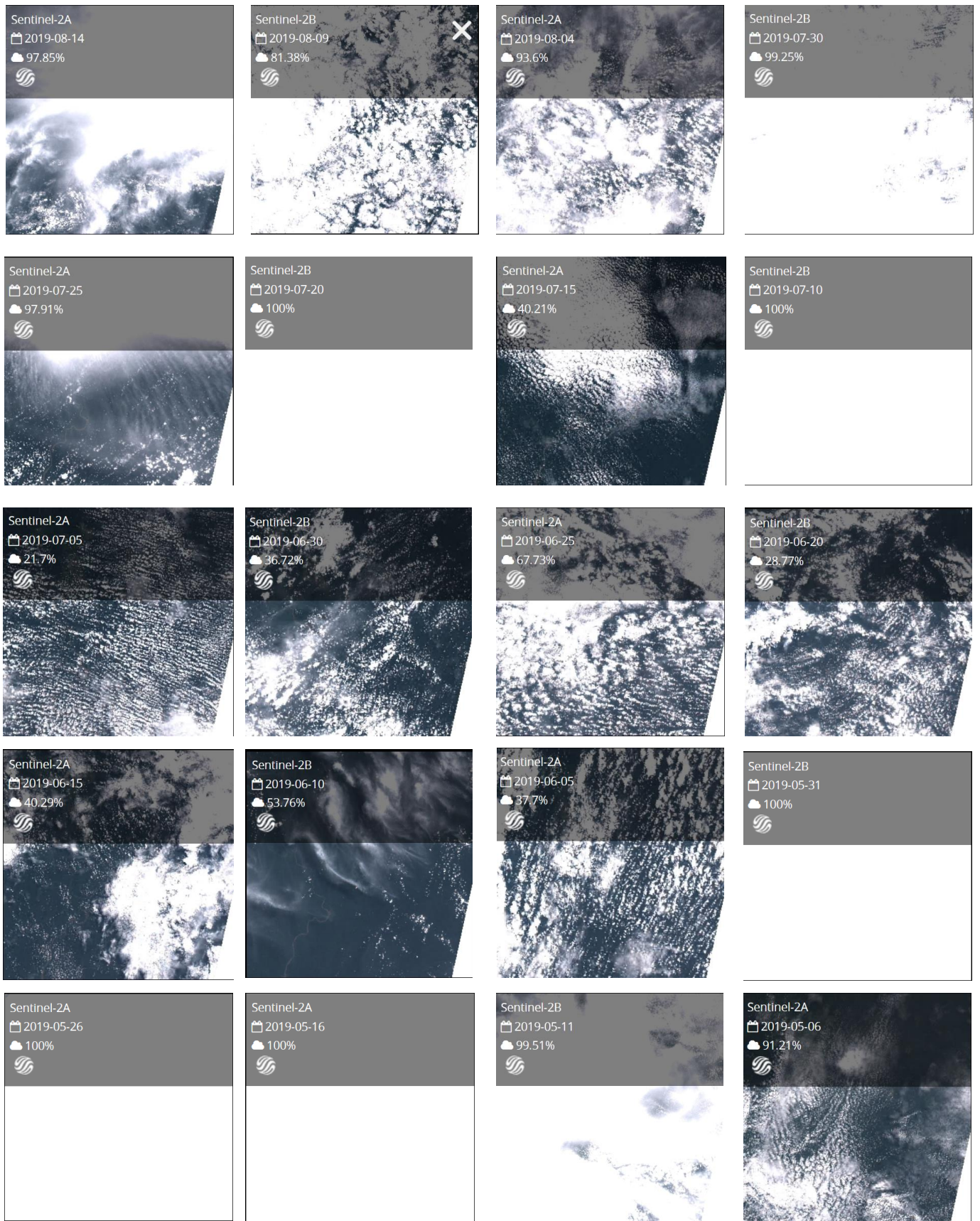


FIGURE B.1 – Couvertures nuageuses pour la tuile 33 NXD (Sentinel-2) pour la dernière année (partie 1)

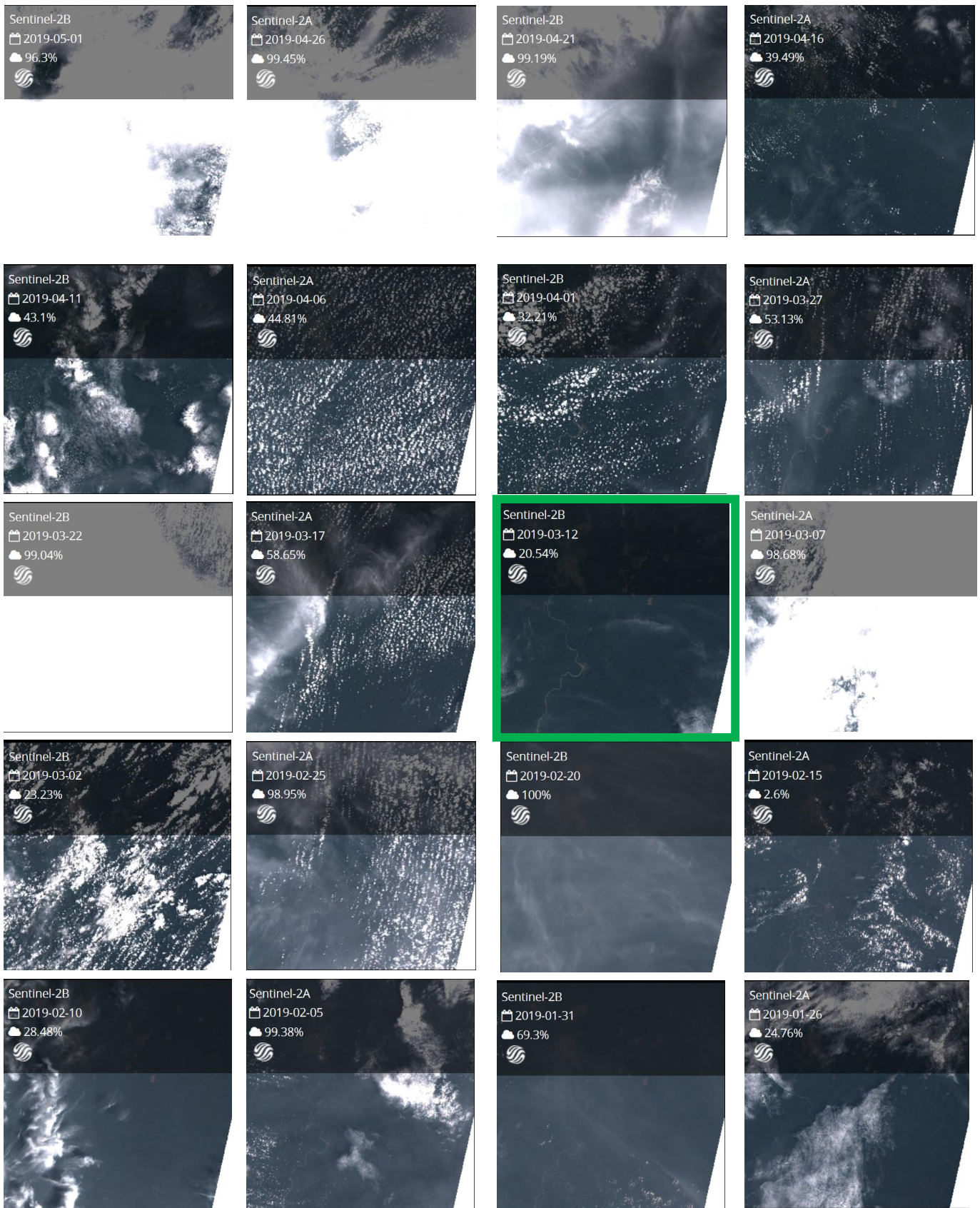


FIGURE B.2 – Couvertures nuageuses pour la tuile 33 NXD (Sentinel-2) pour la dernière année (partie 2)

En vert, les images utilisées lors de ce travail

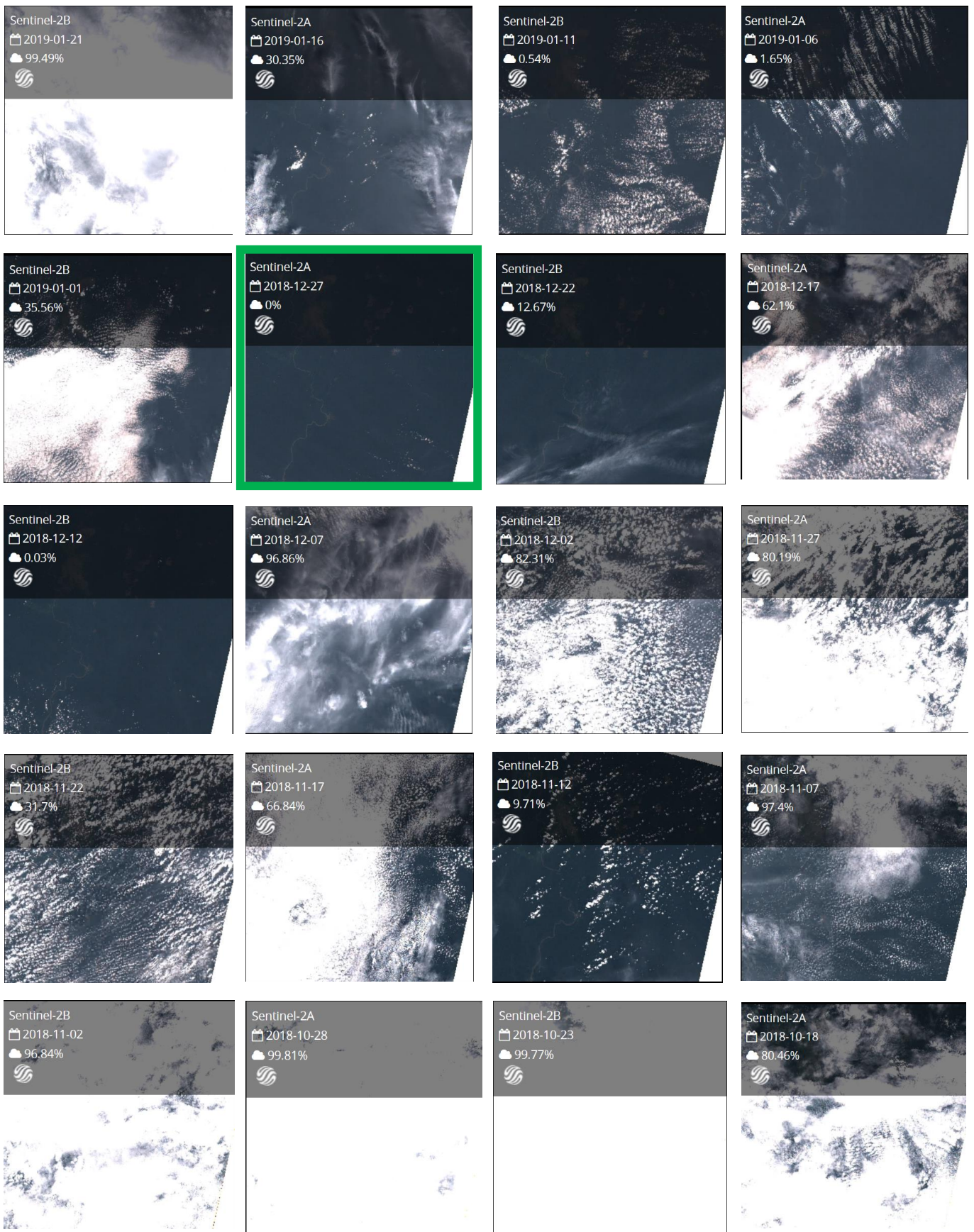


FIGURE B.3 – Couvertures nuageuses pour la tuile 33 NXD (Sentinel-2) pour la dernière année (partie 3)

En vert, les images utilisées lors de ce travail

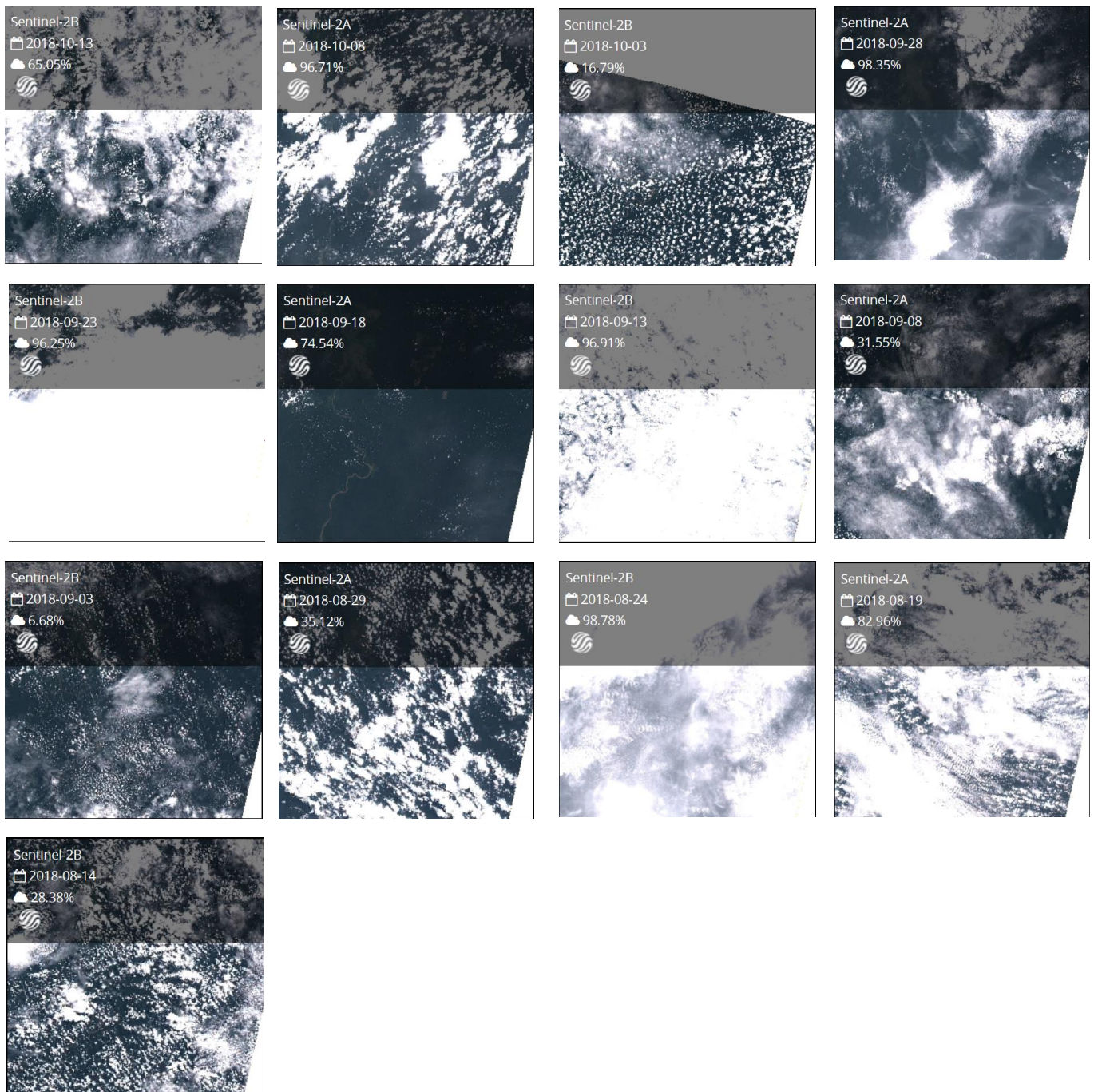


FIGURE B.4 – Couvertures nuageuses pour la tuile 33 NXD (Sentinel-2) pour la dernière année (partie 4)

C Dynamique agricole

Zone	Surface d'affectation agricole (ha)	Surface agricole 2013 (ha)	Surface agricole 2017 (ha)	Évolution de la surface agricole (ha)	Évolution de la surface agricole (%)	Évolution annuelle de la surface agricole (ha)	Évolution annuelle de la surface agricole (ha)
Babongo Nord	886,4	319,1	329,0	+9,9	+3,1	+2,8	+0,9
Babongo Sud	140,5	15,3	18,8	+3,5	+22,9	+1,0	+6,5
Moussapoula Nord	233,3	79,4	64,6	-14,8	-18,6	-4,2	-5,3
Moussapoula Sud	376,0	190,2	195,0	+4,8	+2,5	+1,4	+0,7
Bayanga Centre	764,5	648,5	710,4	+61,9	+9,5	+15,5	+2,4
Yadoumbe	676,3	312,1	407,2	+95,1	+30,5	+23,8	+7,6
Linjombo	771,6	270,3	302,1	+31,8	+11,8	+12,7	+4,7
Bomandjokou	192,8	220,2	257,3	+37,1	+16,8	+14,8	+6,7
Total	4041,4	2055,1	2284,4	+229,3	+11,2	+67,8	+3,3






Zone	Surface agricole hors zone d'affectation 2013 (ha)	Surface agricole hors zone d'affectation 2017 (ha)	Surfaces agricoles hors zone d'affectation 2017 (%)	Évolution des surfaces agricoles hors zone d'affectation (ha)	Surface d'affectation agricole libre (ha)	Surface d'affectation agricole libre (%)
Babongo Nord	0,0	5,6	1,70	+5,6	563,0	63,52
Babongo Sud	0,0	0	0,00	0,0	121,7	86,62
Moussapoula Nord	1,4	1,1	1,70	-0,3	169,8	72,78
Moussapoula Sud	0,0	0	0,00	0,0	181,0	48,14
Bayanga Centre	27,7	82,6	11,63	+54,9	136,7	17,88
Yadoumbe	0,0	0	0,00	0,0	269,1	39,79
Linjombo	0,0	0	0,00	0,0	469,5	60,85
Bomandjokou	55,7	89,5	34,78	+33,8	25,0	12,97
Total	84,8	178,8	7,83	94,0	1935,8	47,90

TABLE C.1 – Détails des bandes culturales au sein des APDS

Evolution des surfaces agricoles au sein des APDS (RCA)

Secteur Babongo Nord, 2013-2017

Légende

-  Surface agricole - 2013
-  Surface agricole - hors zone 2013
-  Surface agricole - extension 2017
-  Surface agricole - hors zone 2017
-  Bande culturale - Secteur Babongo Nord

	2013	2017
Surface agricole (ha)	319,1	323,4
Surface agricole hors zone (ha)	0,0	5,6
Total (ha)	319,1	329,0

Surface Secteur (ha)	886,4
Evolution par an (ha)	+ 2,8
Surface d'affectation restante (ha)	557,4
Surface d'affectation restante (%)	63,9

Système de coordonnées: GCS WGS 1984

Datum: WGS 1984

Unités: Degree

Auteur: D.Lessire (UCLouvain, 2019)

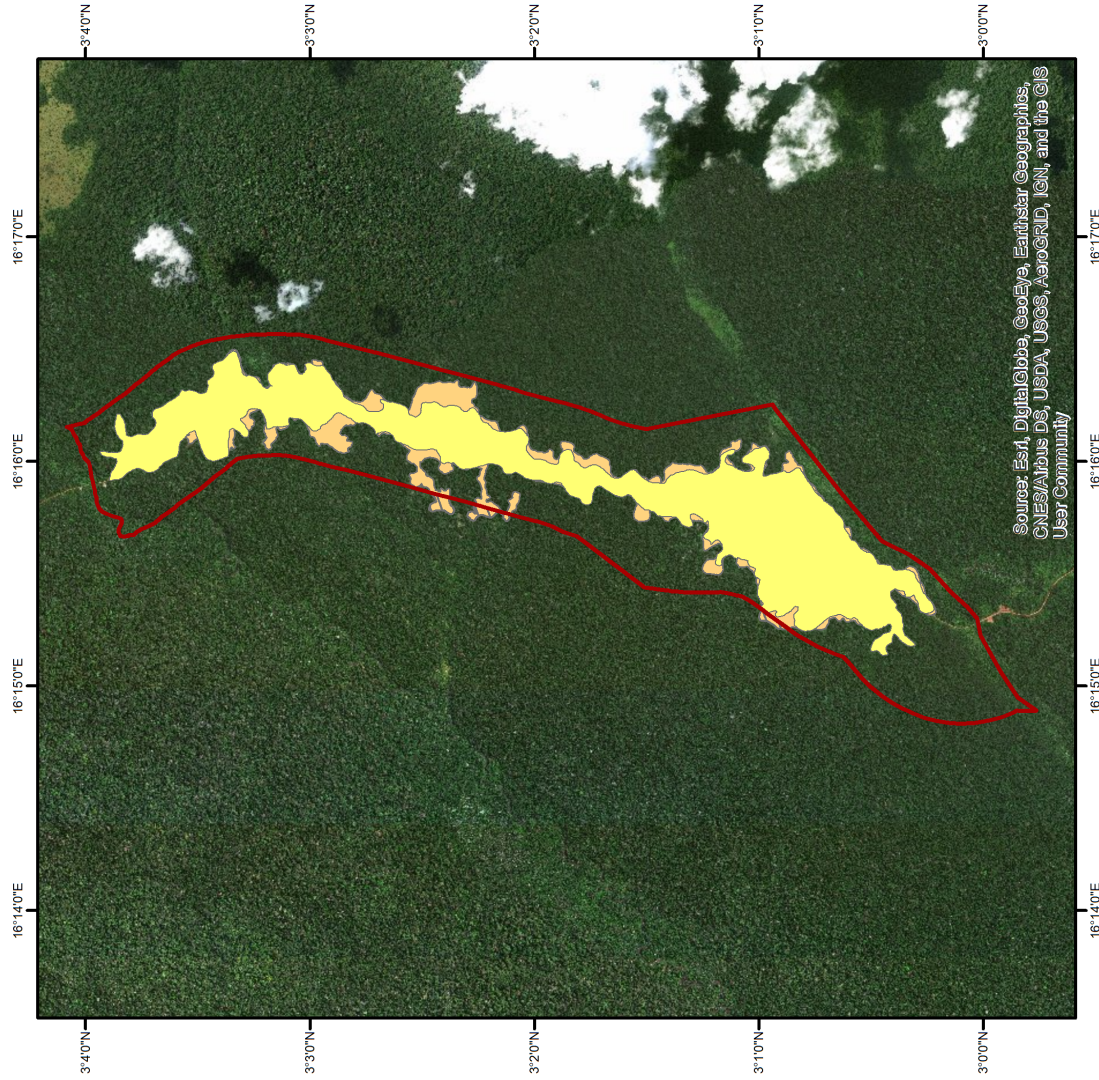
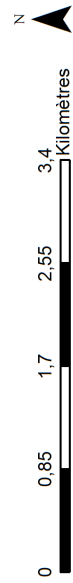


FIGURE C.1 – Évolution des surfaces agricoles dans le secteur de Babongo Nord

Evolution des surfaces agricoles au sein des APDS (RCA)

Secteur Babongo Sud, 2013-2017

Légende

- Surface agricole - 2013
- Surface agricole - extension 2017
- Bande culturale - Secteur Babongo Sud

	2013	2017
Surface agricole (ha)	15,3	18,8
Surface agricole hors zone (ha)	0,0	0,0
Total (ha)	15,3	18,8

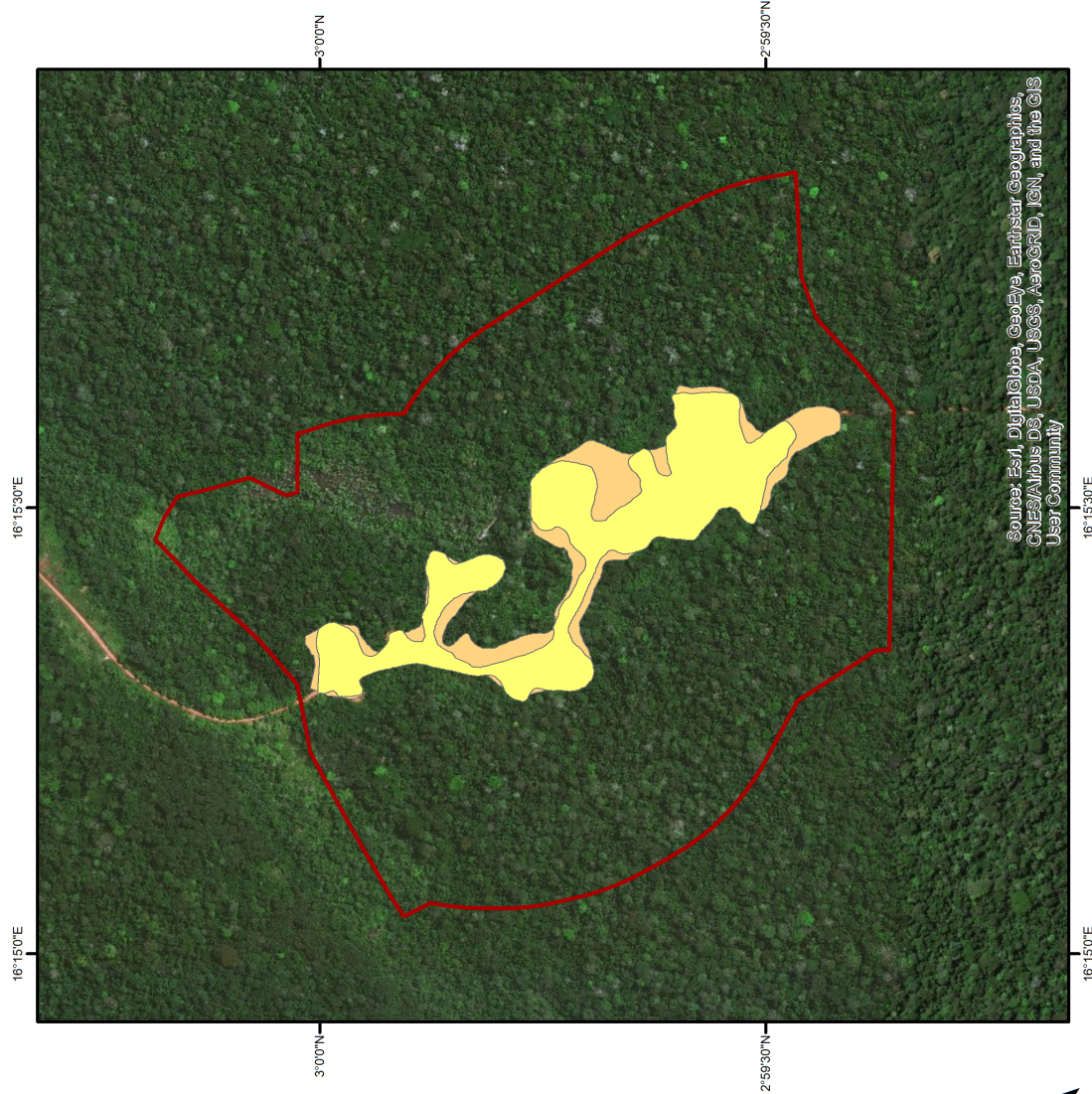
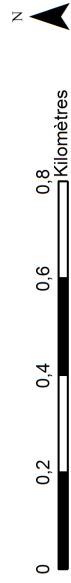
Surface Secteur (ha)	140,5
Evolution par an (ha)	+ 1
Surface d'affectation restante (ha)	121,7
Surface d'affectation restante (%)	86,6

Système de coordonnées: GCS WGS 1984

Datum: WGS 1984

Unités: Degrée

Auteur: D.Lessire (UCLouvain, 2019)








Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community

FIGURE C.2 – Évolution des surfaces agricoles dans le secteur de Babongo Sud

Evolution des surfaces agricoles au sein des APDS (RCA)

Secteur Moussapoula Nord, 2013-2017

Légende

-  Surface agricole - 2013
-  Surface agricole hors zone - 2013
-  Surface agricole - 2017
-  Surface agricole hors zone - 2017
-  Bande culturale - Secteur Moussapoula Nord

	2013	2017
Surface agricole (ha)	78,0	63,5
Surface agricole hors zone (ha)	1,4	1,1
Total (ha)	79,4	64,6

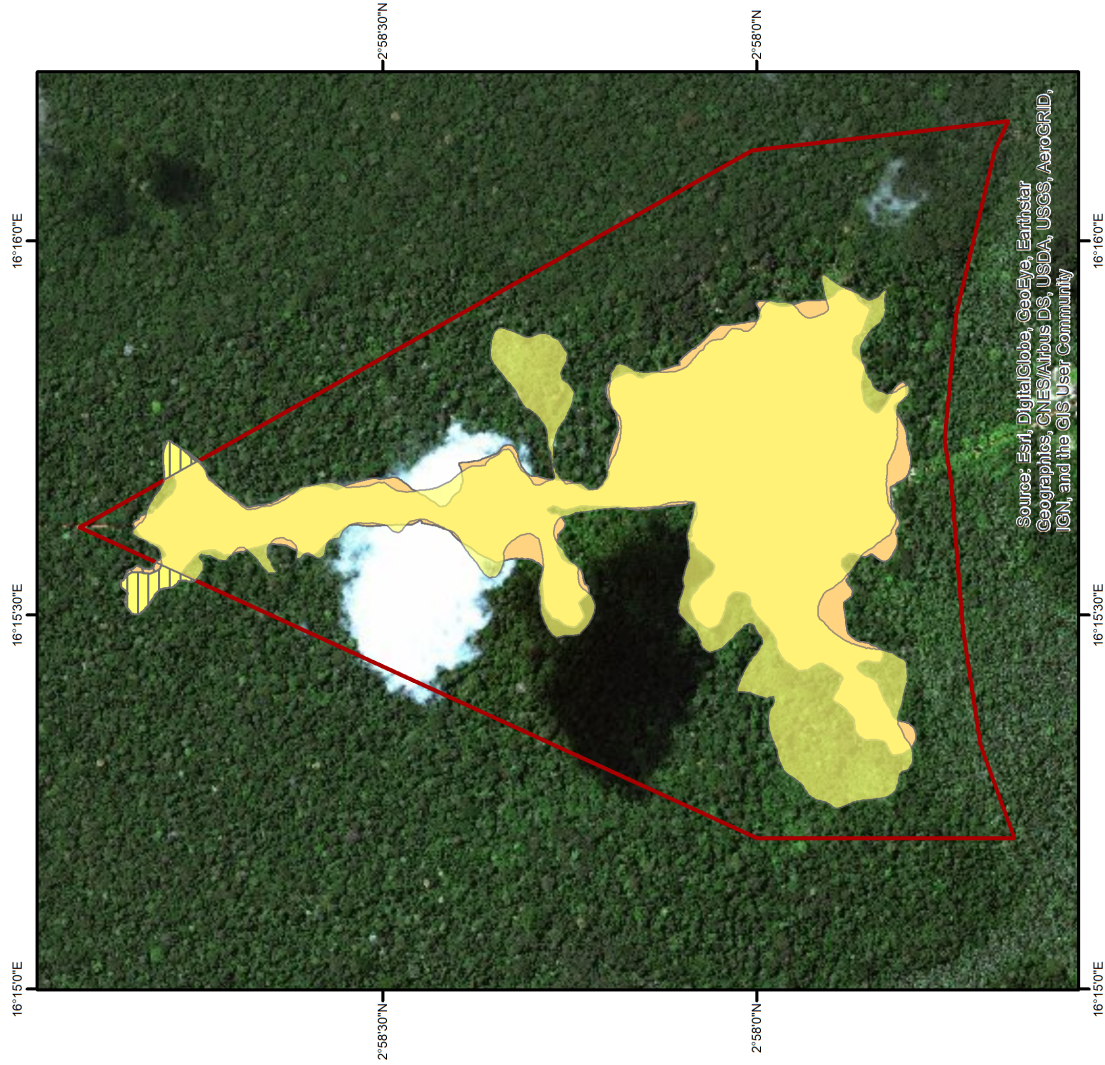
Surface Secteur (ha)	233,3
Evolution par an (ha)	-4,2
Surface d'affectation restante (ha)	168,7
Surface d'affectation restante (%)	72,3

Système de coordonnées: GCS WGS 1984

Datum: WGS 1984

Unités: Degrée

Auteur: D.Lessire (UCLouvain, 2019)



Source: Esri, DigitalGlobe, GeoEye, Earthstar
Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID,
IGN, and the GIS User Community

FIGURE C.3 – Évolution des surfaces agricoles dans le secteur de Moussapoula Nord

Evolution des surfaces agricoles au sein des APDS (RCA) Secteur Moussapoula Sud, 2013-2017

Légende

- Surface agricole - 2013
- Surface agricole - 2017
- Bande culturale - Secteur Massapoula sud

	2013	2017
Surface agricole (ha)	190,2	195,0
Surface agricole hors zone (ha)	0,0	0,0
Total (ha)	190,2	195,0

Surface Secteur (ha)	376,0
Evolution par an (ha)	+1,3
Surface d'affectation restante (ha)	181
Surface d'affectation restante (%)	48

Système de coordonnées: GCS WGS 1984

Datum: WGS 1984

Unités: Degrée

Auteur: D.Lessire (UCLouvain, 2019)

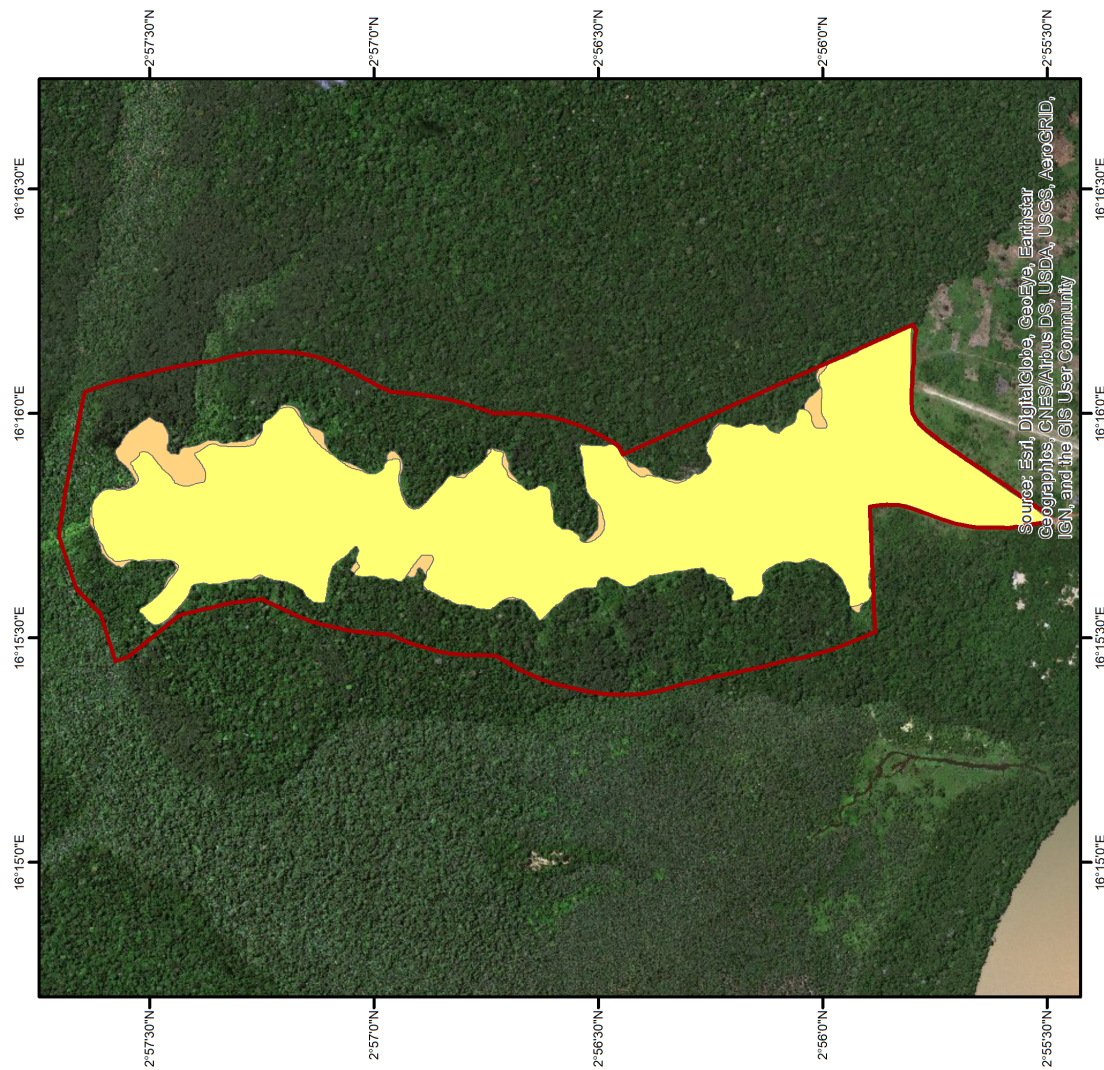
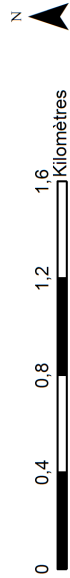







FIGURE C.4 – Évolution des surfaces agricoles dans le secteur de Moussapoula Sud

Evolution des surfaces agricoles au sein des APDS (RCA)

Secteur Bayanga Centre, 2013-2018

Légende

-  Surface agricole - 2013
-  Surface agricole - hors zone 2013
-  Surface agricole - extension 2018
-  Surface agricole - hors zone 2018
-  Bande culturale - Secteur Bayanga Centre

	2013	2017
Surface agricole (ha)	620,8	627,8
Surface agricole hors zone (ha)	27,7	82,6
Total (ha)	648,5	710,4

Surface Secteur (ha)	764,5
Evolution par an (ha)	+ 2,8
Surface d'affectation restante (ha)	136,2
Surface d'affectation restante (%)	17,9

Système de coordonnées: GCS WGS 1984

Datum: WGS 1984

Unités: Degree

Auteur: D.Lessire (UCLouvain, 2019)

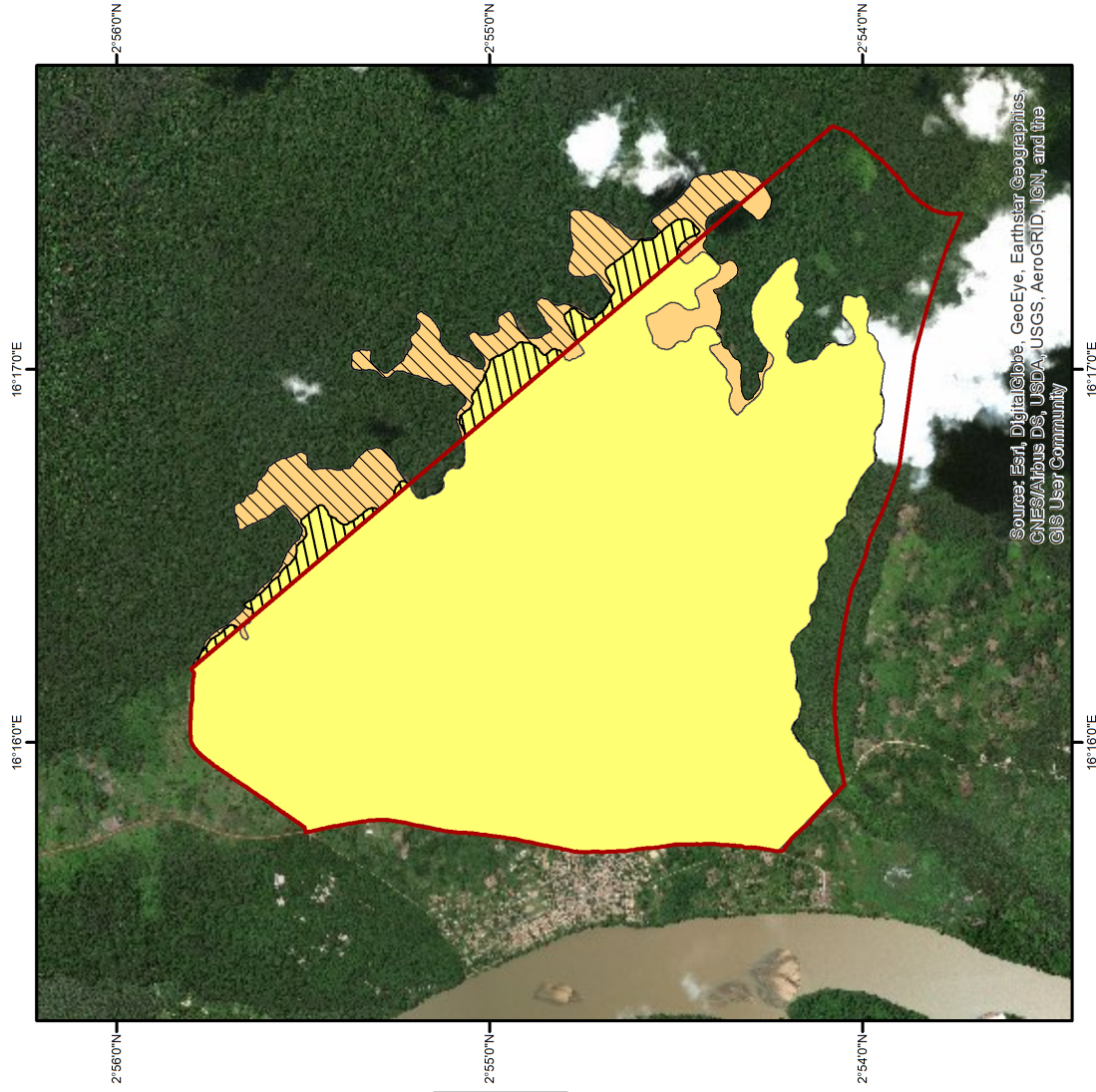


FIGURE C.5 – Évolution des surfaces agricoles dans le secteur de Bayanga centre

Evolution des surfaces agricoles au sein des APDS (RCA) Secteur Yadoumbe, 2013-2018

Légende

- Surface agricole - 2013
- Surface agricole - 2018
- Bande culturale - Secteur Yadoumbe

	2013	2018
Surface agricole (ha)	312,1	407,2
Surface agricole hors zone (ha)	0,0	0,0
Total (ha)	312,1	407,2

Surface Secteur (ha)	676,3
Evolution par an (ha)	+23,8
Surface d'affectation restante (ha)	269
Surface d'affectation restante (%)	39,8

Système de coordonnées: GCS WGS 1984

Datum: WGS 1984

Unités: Degrée

Auteur: D.Lessire (UCLouvain, 2019)

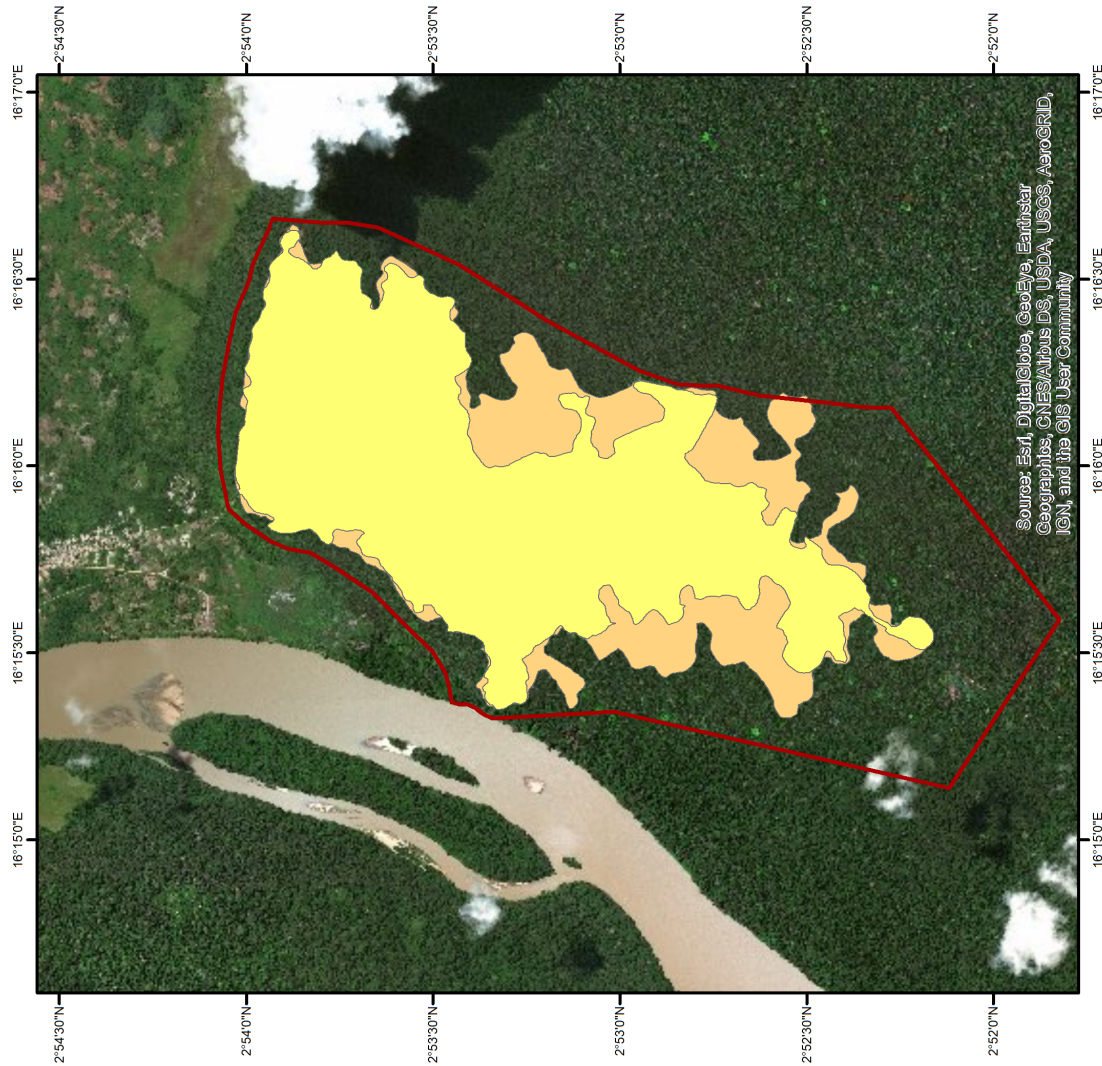
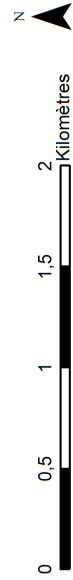


FIGURE C.6 – Évolution des surfaces agricoles dans le secteur de Yadoumbe

Evolution des surfaces agricoles au sein des APDS (RCA) Secteur Lindjombo, 2015-2017

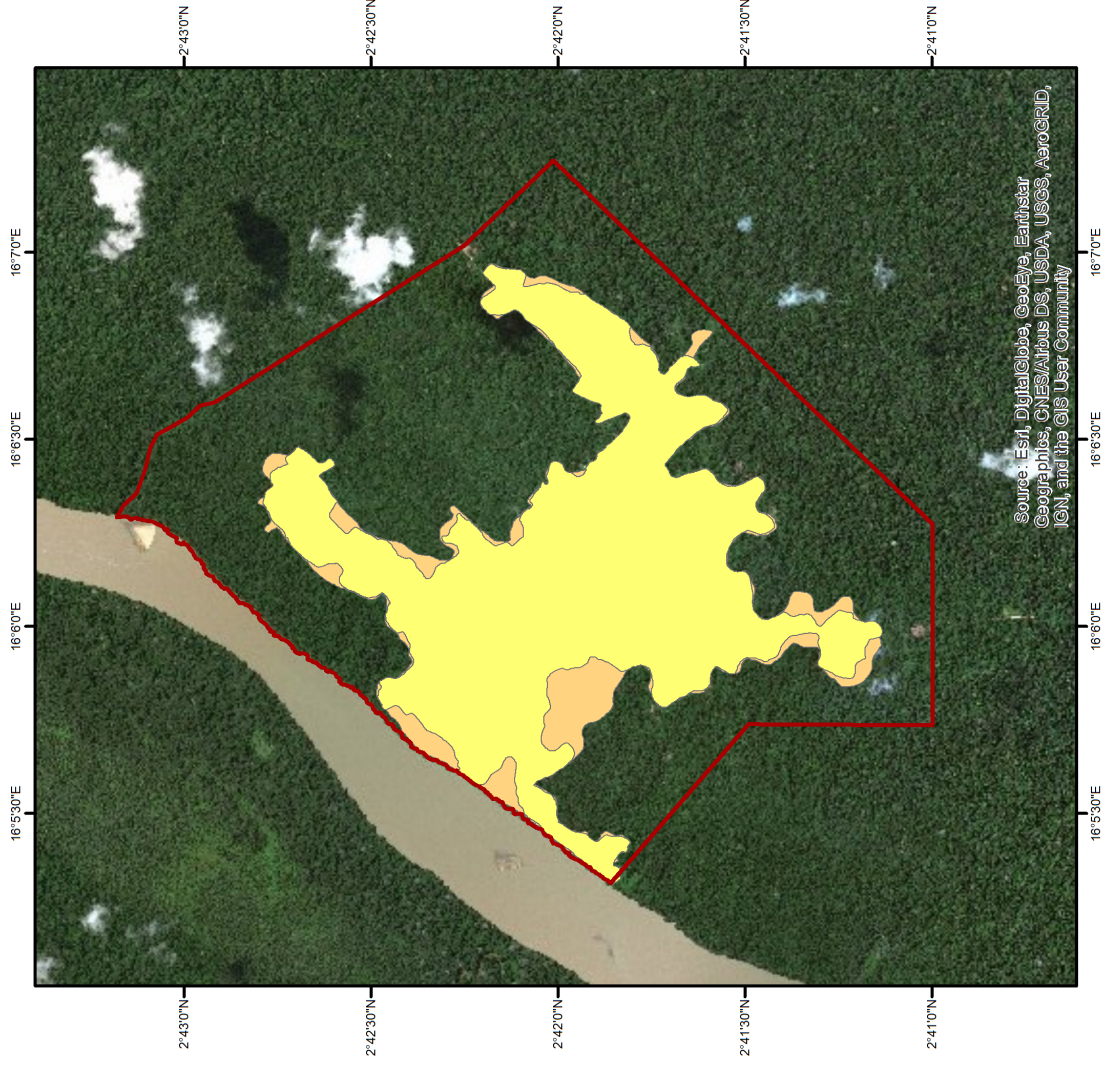
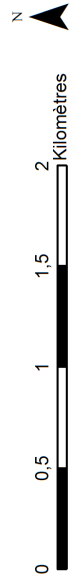
Légende

- Surface agricole - 2015
- Surface agricole - 2018
- Bande culturale - Secteur Lindjombo

	2015	2017
Surface agricole (ha)	270,3	302,1
Surface agricole hors zone (ha)	0,0	0,0
Total (ha)	270,3	302,1

Surface Secteur (ha)	771,6
Evolution par an (ha)	+ 12,7
Surface d'affectation restante (ha)	469,5
Surface d'affectation restante (%)	60,8

Système de coordonnées: GCS WGS 1984
 Datum: WGS 1984
 Unités: Degrée
 Auteur: D.Lessire (UCLouvain, 2019)








Source: Esri, DigitalGlobe, GeoEye, Earthstar
 Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID,
 IGN, and the GIS User Community

FIGURE C.7 – Évolution des surfaces agricoles dans le secteur de Lindjombo

Evolution des surfaces agricoles au sein des APDS (RCA)

Secteur Bomandjokou, 2015-2017

Légende

-  Surface agricole - 2015
-  Surface agricole - hors zone 2015
-  Surface agricole - 2017
-  Surface agricole - hors zone 2017
-  Bande culturale - Secteur Bomandjokou

	2015	2017
Surface agricole (ha)	164,5	167,8
Surface agricole hors zone (ha)	55,7	89,5
Total (ha)	220,2	257,3

Surface Secteur (ha)	192,8
Evolution par an (ha)	+ 14,8
Surface d'affectation restante (ha)	25,1
Surface d'affectation restante (%)	13

Système de coordonnées: GCS WGS 1984

Datum: WGS 1984

Unités: Degrée

Auteur: D.Lessire (UCLouvain, 2019)

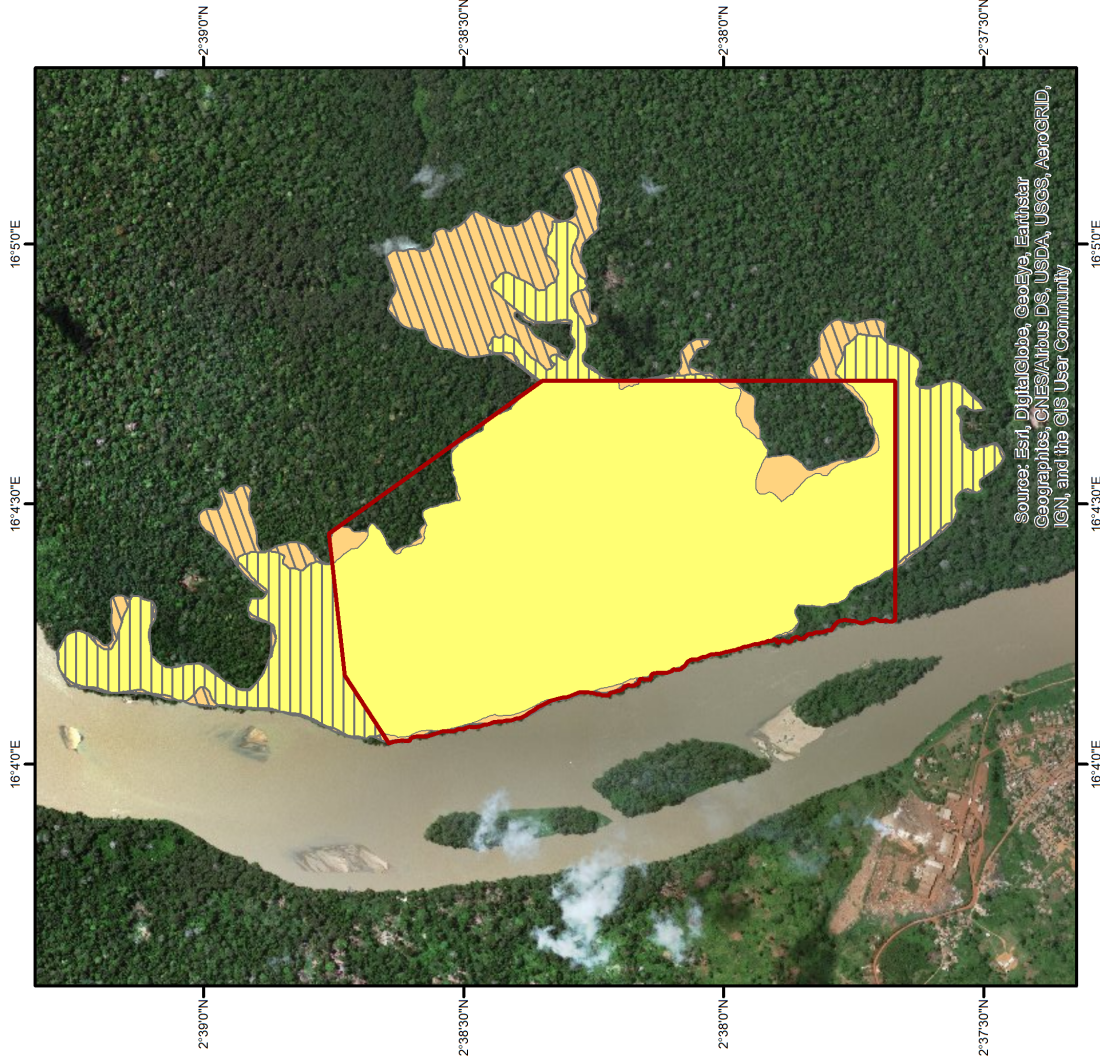
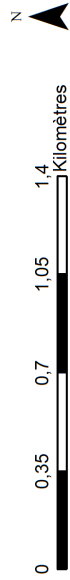


FIGURE C.8 – Évolution des surfaces agricoles dans le secteur de Bomandjokou

Cartographie de types de forêt et de l'occupation du sol en République Centrafricaine à l'aide de télédétection optique (Sentinel-2)

Aires Protégées de Dzanga-Sangha, RCA

Dorian Lessire

Les forêts tropicales sont un écosystème majeur pour la planète aussi bien en termes de biodiversité que de stock de carbone. Elles jouent aussi un rôle capital dans les activités économiques, culturelles et sociales des populations locales. Or diverses pressions pèsent sur ces habitats. Dès lors, mieux les connaître est essentiel pour évaluer et prévoir les changements potentiels.

La télédétection permet d'obtenir des informations très récentes, sur de larges échelles spatiales et à moindre coût. Les lancements récents des plateformes Sentinel-2 de l'ESA, offrent de nouvelles opportunités, de par leurs données à hautes résolutions spatiales (10-20 m) et leur faible temps de revisite (5 jours). Cependant, peu d'études ont encore été menées sur les forêts tropicales à partir de ces capteurs.

L'objectif de ce mémoire est donc d'évaluer le potentiel de discrimination forestière à partir des images multispectrales de Sentinel-2 et de contribuer à l'amélioration des classifications de forêts.

La zone d'étude concerne les Aires Protégées de Dzanga-Sangha dans le sud-est de la République Centrafricaine. Une campagne de terrain, en collaboration avec le WWF, a été réalisée afin de définir une typologie forestière ainsi qu'une clé de photo-interprétation. De plus, un protocole de terrain généralisable à d'autres études en forêts tropicales a été mis au point.

Quatre classifications différentes ont été testées. Il en ressort que la classification supervisée via une approche par objet est celle qui donne les meilleurs résultats en termes de qualité. Cette méthode a donc été utilisée pour cartographier l'entièreté de la zone d'étude à la fois pour les types forestiers et pour l'occupation du sol. La classification finale obtient une exactitude de prédiction générale de 74%.

Une évaluation de la dynamique agricole au sein de la zone d'étude est aussi réalisée.

Mots-clés : Classification forestière - Land cover - Sentinel-2 - République Centrafricaine - Aires Protégées de Dzanga-Sangha - Dynamique agricole

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

Faculté des bioingénieurs

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