HAL open science

Apport des données Sentinel-1 pour le suivi continu de la forêt tropicale : cas de la Guyane

Marie Ballere

To cite this version:

Marie Ballere. Apport des données Sentinel-1 pour le suivi continu de la forêt tropicale : cas de la Guyane. Ingénierie de l'environnement. Université Gustave Eiffel, 2021. Français. NNT : 2021UEFL2019 . tel-03629552

HAL Id: tel-03629552

https://theses.hal.science/tel-03629552v1

Submitted on 4 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Apport des données Sentinel-1 pour le suivi continu de la forêt tropicale : Cas de la Guyane

Thèse de doctorat de l'Université Gustave Eiffel

École doctorale n°532 MSTIC - Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité de doctorat : Sciences et Technologies de l'Information Géographique
Unité de recherche : LASTIG

Thèse présentée et soutenue à l'Université Gustave Eiffel, le 22/11/2021, par

Marie BALLERE

Composition du Jury

Laetitia THIRION Professeure, CentraleSupélec, laboratoire SONDRA	Rapporteure et Présidente
Jérôme CHAVE Directeur de recherche, CNRS, laboratoire EDB	Rapporteur
Inge JONCKHEERE Docteure, FAO (Food and Agriculture Organization)	Examinatrice
Pierre COUTERON Directeur de recherche, IRD, UMR AMAP	Examinateur
Pierre-Louis FRISON Maitre de conférences, HDR, Université Gustave Eiffel, laboratoire LASTIG	Directeur de thèse

Encadrement de la thèse

Thierry KOLECK Docteur, CNES	Examinateur et Co-Encadrant
Stéphane Mermoz Docteur, GlobEO	Co-Encadrant
Alexandre Bouvet Ingénieur de recherche, IRD, UMR CESBIO	Co-Encadrant

Remerciements

L'arrivée dans le monde professionnel n'est pas chose facile, du moins c'est ce que je pensais. C'est une étape que j'appréhendais, n'étant initialement que peu familière avec le monde des ingénieurs et des scientifiques.

Mais quelle chance j'ai eu : durant presque cinq ans, entre stage, CDD et thèse, j'ai pu travailler sur des sujets captivants qui me tiennent à cœur, avec des personnes compétentes et que j'apprécie. Ce fut très agréable et je mesure ma chance d'avoir eu un cadre de travail idéal, dans lequel j'ai pu m'épanouir facilement. J'ai pris énormément de plaisir à réaliser cette thèse, et je souhaite remercier tous ceux qui ont participé à cette aventure et qui ont permis de la façonner.

Commençons par l'équipe rapprochée. J'aimerais avant tout exprimer ma gratitude envers Thierry Koleck ingénieur CNES/CESBIO, sans qui je ne serais sûrement pas en train d'écrire ces lignes. Tu m'as permis d'évoluer dans les domaines de la télédétection et de la forêt, des sujets passionnants. Ta polyvalence et ta façon d'être avec les gens ont été une inspiration. Je tiens également à remercier sincèrement Stéphane Mermoz, fondateur de GlobEO, et Alexandre Bouvet, ingénieur de recherche au CESBIO, qui m'ont intégré à leurs travaux sur le déboisement, et auprès de qui j'ai énormément appris, tant sur la thématique et sur la théorie que sur les aspects « pratiques » de la recherche. Je souhaite également remercier Pierre-Louis Frison, maître de conférences (HDR) à l'Université Gustave Eiffel, qui a accepté de diriger ma thèse. Merci pour tes conseils pertinents, ton enthousiasme pour mes travaux, ton soutien et ton franc-parler. Je n'oublie pas de remercier Thuy Le Toan du CESBIO, pour ses retours sur mes travaux qui les ont sans aucun doute rendus meilleurs, et que j'admire pour sa lucidité et son ouverture scientifique. J'ai vraiment passé de très bons moments à travailler avec vous tous, dans la joie et la bonne humeur, autant qu'avec rigueur et sérieux. Je vous remercie de m'avoir encadré, mais aussi pour votre confiance qui m'a permis de faire mes propres choix pour que cette thèse me corresponde.

Je remercie également les rapporteurs de ma thèse, Laetitia Thirion, professeure à CentraleSupElec et Jérôme Chave, directeur de recherche à EDB : d'abord pour le travail que cela représente, et particulièrement pour leurs retours pertinents et complémentaires, que j'ai su apprécier tant ils attestaient d'une lecture attentive de mon manuscrit. Je remercie également les autres examinateurs du jury, Inge Jonckheere, de la FAO, et Pierre Couteron, directeur de recherche à l'AMAP, pour l'intérêt qu'ils ont porté à mes travaux.

Grâce à Céline Tison qui m'a accueillie au sein du service DSO/SI/TR du CNES, j'ai bénéficié d'un environnement de travail idéal. J'ai apprécié les moments passés avec ce service, ainsi qu'avec nos fidèles acolytes de DSO/SI/IP. Damien, Alejandro, Nadine, Roger, Jérôme, Christophe, Bérangère, Raquel, Claire, Laura, Louise, Denis, Sophie, et tous les autres : merci à tous pour vos conseils, votre bonne humeur et votre soutien (et pour les gaufres, Alain !). J'ai vraiment pris plaisir à évoluer à vos côtés.

Pour m'avoir permis de passer du temps (et ma soutenance !) au CESBIO, je remercie ses deux directeurs successifs Laurent Polidori à mon arrivée, et Mehrez Zribi à mon départ. L'équipe BIOMASS, c'est quelque chose… et j'ai énormément apprécié votre compagnie : Hoa, Milena, Selma, David et tous ceux qui en ont fait partie un moment. Merci Ludo de t'être lancé à corps perdu dans la forêt guyanaise avec Thierry pour essayer de vérifier des détections de chablis. Ceux qui connaissent le terrain savent qu'en pleine forêt, faire 500 mètres représente un effort considérable : bravo à tous les deux d'en être revenus vivants !

Je remercie le CNES et le WWF France pour avoir financé mes recherches. Je n'oublie pas que la mise en relation de ces deux organismes est l'initiative d'Amandine Guillot, que je remercie, et sans qui cette thèse n'aurait pas bénéficié de ce savoureux mélange entre un acteur du spatial et une ONG utilisatrice de produit sur le terrain. Côté WWF, je remercie Daniel Vallauri et Emma Do Khac pour avoir suivi ma thèse et porté de l'intérêt à mes travaux.

Évidemment, je remercie profondément toutes les personnes avec qui j'ai dialogué en Guyane, Mathilde André et Caroline Bedeau de l'ONF, Sébastien Linarès de l'EMOPI, Clément Villien du WWF, Benoit de Thoisy de l'institut Pasteur, pour tous nos échanges et le travail réalisé ensemble aux cours de ces 3 années. Les connaissances des acteurs locaux sont une mine d'or (sans jeu de mots) pour qui travaille à développer des méthodes et des algorithmes sur leur territoire. J'ai beaucoup appris grâce à vous, et ça a rendu le travail d'autant plus passionnant. Merci aussi à Emmanuelle Desaintmartin, de la DAAF, qui a réalisé la campagne terrain et qui m'a appris beaucoup sur la forêt et sur les guyanais. J'en profite pour remercier Vincent Polveche et Hervé Legrand, du Ministère de l'Agriculture, qui ont permis l'organisation de cette campagne.

Je tiens également à souligner le plaisir que j'ai eu de rencontrer Cédric Lardeux de l'ONFI et Juan Doblas de l'INPE, qui travaillent sur les mêmes sujets que moi et qui sont des personnes curieuses, motivées et inspirantes.

Merci à Théo Chombart pour avoir relu mon manuscrit à une vitesse incroyable, et également à l'équipe du centre de production de données du CNES, Baptiste, Virginie, Gwen, Raquel, Clément, Sylvie, à qui j'ai donné du fil à retordre en étant l'utilisatrice numéro 1 pendant des mois, lors de grosses phases de téléchargement des images Sentinel-1.

Enfin, je tiens à remercier Brice, pour son soutien et son aide précieuse, mes amis : les Guillaumes, Alae, Jorge, Adeline, et compagnie ; et en particuliers Zinedine, Michele, Yanis, Emma et Sandra qui m'ont donné beaucoup de force en ayant foi en moi et en m'encourageant. Pour finir, une pensée pour mes parents, qui ne m'ont pas vraiment poussé à faire cette thèse, mais qui m'ont permis de découvrir le monde et de faire des études, et qui sont maintenant fiers de leur fille docteure.

Résumé

Les forêts tropicales abritent la biodiversité la plus riche de notre planète et jouent un rôle particulièrement important dans le stockage du carbone et le cycle de l'eau. Le suivi de la dégradation de ce milieu est un défi d'actualité car il représente un danger immédiat pour l'environnement et la biodiversité.

Par ailleurs, depuis 2014, la mise en orbite des satellites Sentinel dans le cadre du programme européen Copernicus constitue une réelle révolution dans le panorama des capteurs d'observation de la Terre existants jusqu'alors. En effet, des données optiques (Sentinel-2) et radar (Sentinel-1) sont, depuis lors, accessibles librement à tous et permettent des acquisitions à une résolution décamétrique et une répétitivité allant de 5 à 12 jours selon la région analysée.

L'objectif de cette thèse consiste donc à développer des méthodes pour le suivi de la forêt tropicale basées sur des données radar Sentinel-1 et des produits de cartographie exogènes.

Plus spécifiquement, l'étude a pour but de fournir des outils pour observer le déboisement, en produisant des cartes en quasi-temps réel, puis en caractérisant les zones détectées avec la cause de leur déboisement. Ces informations sont nécessaires à l'estimation et au suivi du déboisement sur les zones tropicales, et à la lutte contre les défriches illicites.

La région test est la Guyane. Couverte à plus de 95% par la forêt, elle est caractérisée par de nombreux types de perte forestière permettant de valider la méthode sur de nombreuses caractéristiques. De plus, ce territoire étant relativement bien surveillé, beaucoup de données cartographiques d'occupation du sol sont disponibles pour situer les déboisements ou pour servir de référence au niveau spatial.

La première phase du travail est l'amélioration d'une méthode de détection du déboisement à partir des données Sentinel-1 et son évaluation sur la Guyane. La validation de la carte produite s'appuie sur 1 867 données (de différents types de déboisement) produites de manière indépendante, et représentant un total de 2 124.5 ha à travers toute la Guyane sur une période de 2 ans.

Les résultats obtenus sur la Guyane au niveau spatial sont très satisfaisants : 96% de précision sur les surfaces déboisées et 81.5% de rappel. La mise en place d'une campagne terrain a rendu possible l'estimation du côté quasi-temps réel de la méthode sur 26 parcelles et avance un délai médian de détection de 3.5 jours.

Une comparaison avec un produit optique souvent pris comme référence, a démontré l'avantage des données Sentinel-1 pour ces régions caractérisées par de fortes couvertures

nuageuses. Cette meilleure performance s'observe tant sur le plan spatial (rappel du radar supérieur de plus de 35%), que sur le plan temporel (un tiers des zones déboisées nécessitant une surveillance est détecté avec 3 mois de retard par les images optiques).

La deuxième partie du travail consiste à caractériser les zones déboisées détectées afin de déterminer leur cause. Pour cela, des méthodes d'apprentissage automatique ont été utilisées pour choisir des indicateurs cartographiques pertinents, permettant de produire un modèle prédictif simple, fiable, automatique et utilisable à l'échelle de la Guyane en temps-réel.

Les indicateurs F-scores de chaque type de déboisement sont de 97% pour l'orpaillage et l'agriculture, 95% pour l'exploitation forestière, 87% pour l'urbanisation et 76% pour la classe « autre », pouvant se rapporter à des perturbations naturelles. Ce travail a montré la pertinence d'utiliser des indicateurs cartographiques pour déterminer la cause des déboisements en Guyane, permettant cette identification en temps-réel.

Ces travaux qui démontrent le potentiel des données Sentinel-1 pour le suivi continu des forêts tropicales ont été menés en interaction avec les acteurs locaux. Ils montrent également la complémentarité de systèmes provenant de capteurs différents et pourront être poursuivis en ce sens.

Mots clés : Déboisement, forêt tropicale, suivi continu, Sentinel-1, cause du déboisement, apprentissage automatique.

Abstract

Tropical forests are home to the richest biodiversity on our planet and play a particularly important role in carbon storage and the water cycle. Monitoring the degradation of this environment is currently a challenge because it represents an immediate danger for the environment and biodiversity.

Moreover, since 2014, the launch of the Sentinel satellites within the framework of the European Copernicus program constitutes a real revolution in the panorama of Earth observation sensors existing until then. Indeed, optical (Sentinel-2) and radar (Sentinel-1) data are, since then, freely accessible to all and allow acquisitions with a decametric resolution and a repeatability ranging from 5 to 12 days depending on the region analyzed.

The objective of this thesis is to develop methods to monitor the tropical forest based on Sentinel-1 radar data and exogenous mapping products.

More specifically, the study aims to provide tools to observe deforestation, by producing maps in near-real time, and then characterizing the detected areas with the driver of their deforestation. This information is needed to estimate and monitor deforestation in tropical areas, and to fight against illegal deforestation.

The study site is French Guiana. With 95% forest coverage, it is characterized by many types of forest loss allowing the method to be validated on multiple characteristics. Moreover, as this area is already relatively well monitored, a lot of cartographic data is also available in order to locate forest loss or to be a reference at the spatial level.

The first phase of the work is the improvement of a deforestation detection method from Sentinel-1 data and its evaluation over French Guiana. The validation of the map produced is based on 1 867 data (of different types of deforestation) produced from distinct organizations, and representing a total of 2 124.5 ha throughout French Guiana over a period of 2 years.

The results obtained in French Guiana at the spatial level are very satisfactory: 96% accuracy in detecting deforested areas and 81.5% producer’s accuracy. The implementation of a field campaign has made it possible to estimate the quasi-real time side of the method on 26 plots and gives a median detection time of 3.5 days.

A comparison with an optical product often taken as a reference, demonstrated the advantage of Sentinel-1 data in those regions characterized by high cloud cover. This best performance is observed both spatially (radar recall higher than 35%), and temporally (one third of the areas deforested requiring monitoring are detected within a 3-month delay by optical images).

The second part of the work consists in characterizing the detected deforested areas in order to determine their cause. For this, machine learning methods were used to select relevant cartographic indicators, allowing to produce a simple, reliable, automatic and usable predictive model at French Guiana's scale in real time.

The F-scores for each type of deforestation are 97% for gold panning and agriculture, 95% for logging, 87% for urbanization, and 76% for the "other" class, which may refer to natural disturbances. This work has shown the relevance of using cartographic indicators to determine the drivers of deforestation in French Guiana, allowing this identification in real time.

This work, which demonstrates the potential of Sentinel-1 data for continuous monitoring of tropical forests, was conducted in interaction with local stakeholders. They also show the complementariness of systems from different sensors and efforts can be continued in that regard.

Keywords: Deforestation, tropical forest, continuous monitoring, Sentinel-1, drivers of deforestation, machine-learning.

THESE PREPAREE AU CENTRE NATIONAL D'ETUDES SPATIALES, DANS LE SERVICE DSO/SI/TR, 18 AVENUE EDOUARD BELIN, 31400 TOULOUSE.

Table des matières

Résumé 3

Abstract 8

Chapitre 1 - Introduction 26

1. Importance de la forêt 26

2. Objectifs de recherche et structure du manuscrit 29

Chapitre 2 - Les pertes forestières tropicales 30

1. Causes des pertes forestières tropicales 30

2. Pratiques diverses de déboisement 33

3. État de l'art de l'utilisation de la télédétection pour le suivi des forêts 34

Chapitre 3 - Site d'Etude et Données 40

1. Site d'étude – La Guyane 41

1.1. La Guyane 41

1.2. Les pertes forestières en Guyane 42

1.2.1. L'agriculture 42

1.2.2. L'exploitation forestière 44

1.2.3. L'orpaillage 46

1.2.4. L'urbanisation 48

2. Données cartographiques 49

2.1. Données de référence 49
2.1.1. Données agriculture 49
2.1.2. Données exploitation forestière 49
2.1.3. Données orpaillage 50
2.2. Données auxiliaires 50
2.2.1. BD TOPO de l'IGN 50
2.2.2. CAMINO 51
2.2.3. BRGM 52
2.2.4. ECOSEO 52
2.2.5. GLAD-FA 52
3. Données de terrain acquises dans le cadre de cette thèse 53
3.1. Motivations 53
3.2. Mode opératoire – données collectées 54
3.3. Analyse a posteriori des données 57
4. Données RSO 59
4.1. Télédétection RSO 59
4.1.1. Radar imageur 59
4.1.2. Radar a Synthèse d'Ouverture (RSO) 62
4.2. Les images RSO 63
4.2.1. Résolution spatiale 63
4.2.2. Taille de pixel 64
4.2.3. Caractéristiques du coefficient de rétrodiffusion 65
4.2.4. Propriétés géométriques 67
4.2.5. Propriétés radiométriques 68
4.3. Sentinel-1 70

Chapitre 4 - Analyse et interprétation 77
1. Analyse des données de référence 77

1.1. Hauteur des forêts 78
1.2. Topographie 79
1.3. Taille des surfaces déboisées 80
1.4. Caractéristiques géographiques 82
2. Interprétation des signatures temporelles Sentinel-1 sur les données de référence 87

Chapitre 5 - Méthodes et validation 95
1. Détection du déboisement 96
1.1. Méthode initiale et choix des paramètres pour la Guyane 96
1.1.1. Choix des images 99
1.1.1.1. Polarisation 99
1.1.1.2. Capteur 102
1.1.1.3. Direction de Visée 103
1.1.1.4. Angle d'incidence 104
1.1.2. Prétraitements 105
1.1.3. Calcul du Radar Change Ratio 106
1.1.4. Détection des ombres 106
1.1.5. Détection des zones potentiellement déboisées 107
1.1.6. Discussion sur la résolution spatiale des produits recherchés 107
1.2. Améliorations apportées à la méthode initiale 109
1.2.1. Prise en compte de la dimension temporelle 109
1.2.2. Topographie 111
1.2.3. Indice de confiance 115
2. Méthode de validation et de comparaison de carte de déboisement 120
2.1. Méthode de validation spatiale 120
2.2. Méthode de comparaison temporelle 122
2.3. Données utilisées pour la validation 123
3. Caractérisation des déboisements 126

3.1. Idée générale de la méthode et concept généraux 126
3.1.1. Arbre de décision 126
3.1.1.1. Construction d'un arbre de décision 127
3.1.1.2. Avantages et inconvénients 128
3.1.2. Algorithme du Random Forest (RF) 129
3.2. Préparation : Définition d'indicateurs potentiels et création de la base d'apprentissage .. 131
3.3. Création du modèle 133
3.3.1. Importance des indicateurs 133
3.3.2. Nombre d'indicateurs et taille de l'arbre choisi 134
3.3.3. Arbre de décision choisi comme modèle pour la caractérisation des pertes forestières 134
3.4. Validation du modèle de caractérisation du déboisement 134

Chapitre 6 - Résultats 137
1. Résultats – Détections des pertes forestières 138
1.1. Validation spatiale 139
1.2. Comparaison spatiale avec un produit basé sur des données optiques 141
1.3. Evaluation temporelle 147
1.4. Comparaison temporelle 149
1.4.1. Agriculture vivrière 149
1.4.2. Exploitation forestière 151
1.4.3. Orpaillage 152
1.5. Réduction du délai de détection 153
1.6. Observations 156
1.6.1. Taux annuel de changements forestiers 156
1.6.2. Surface des parcelles détéctées 157
2. Résultats – Caractérisation des déboisements 158

2.1. Création du modèle 158
2.1.1. Importance des indicateurs 158
2.1.2. Nombre d'indicateurs et taille de l'arbre choisi 161
2.2. Validation du modèle 164
2.2.1. Résultats globaux 165
2.2.2. Analyse de la confusion des classes 165
2.2.3. Carte de confiance 167

Chapitre 7 - Valorisations et perspectives 170
1. Besoins utilisateurs 170
2. Exemple d'applications en Guyane 173
2.1. Détection des dommages causés par les barges d'orpaillage aux rives du Maroni 173
2.2. Détections de sites d'orpaillage clandestins 176
3. Utilisation des images Sentinel-1 à l'échelle globale 177
3.1. Géneralisation de la méthode de détection du déboisement 177
3.2. Opérationnalisation 180
3.2.1. Confiance dans la détection 180
3.2.2. Délai d'alerte et retards de détection 180
3.2.3. Délais de production 181
4. Perspectives 181
4.1. Perspectives d'amélioration 181
4.1.1. Combinaison de données multi-capteurs 181
4.1.2. Evolution dynamique du système de caractérisation 182
4.2. Perspectives d'utilisation 183

Chapitre 8 - Conclusion 186

Bibliographie 190

Annexes.....201

1. Mesures communes d'évaluation.....201

2. Description des indicateurs créés pour la caractérisation du déboisement en Guyane.....204

3. Visualisations graphiques des indicateurs de caractérisation par type de déboisement 212

Table des figures

Figure 1 - Couverture spatiale des différentes forêts dans le monde. Source (The State of the World's Forests 2020. In brief, 2020) 27
Figure 2 - Estimations à l'échelle continentale de (A) la proportion de chaque facteur de déforestation, (B) la variation de la superficie forestière (km2/an-1 ; (The State of the World's Forests 2020. In brief, 2020) pour la période 2000-2010 par facteurs de déforestation, et (C) la proportion de la superficie forestière perturbée par facteur de dégradation. L'étude a été faite sur la base de données provenant de 46 pays tropicaux et subtropicaux. Source : (Hosonuma et al., 2012). 28
Figure 3 - Photo d'une parcelle agricole obtenue par : abattage des arbres (à gauche). Ils sont encore au sol et la végétation se dessèche ; après brûlis (à droite), il ne reste que des troncs au sol. 42
Figure 4 - Vue satellitaire optique et radar d'abattis (3°7 N, 54°0 O). L'image est une mosaïque biannuelle d'images Planet Juin 2019 (www.planet.com/nicfi). L'image radar est une image Sentinel-1 en polarisation VH du 11/10/2019 et sa dynamique est réglée entre -30 (noir) et -7 dB (blanc). 43
Figure 5 - Photo de routes (à gauche) et de parcs à grumes (à droite) construits pour l'exploitation forestière. Dans les deux cas, la forêt a été remplacée par un sol nu et il n'y a rien au sol à part des grumes. Le sol est lisse dans la plupart des cas mais peut être rendu rugueux avec l'activité des machines. © ONF 44
Figure 6 – Vue satellitaire optique et radar de pistes et parcs à grumes construits pour l'exploitation forestière (4°0 N, 51°8 O). L'image est une mosaïque biannuelle d'images Planet Juin 2020 (www.planet.com/nicfi). L'image radar est une image Sentinel-1 en polarisation VH du 13/08/2020 et sa dynamique est réglée entre -30 (noir) et -7 dB (blanc). 45
Figure 7 - Photographie aérienne des opérations légales d'extraction d'or : la zone est vaste et composée de bassins d'eau et de sol nu (à gauche). Photographie aérienne des opérations d'exploitation aurifère illégale : seule une petite surface de forêt a été défrichée (à droite). © ONF 46
Figure 8 - Vue satellitaire optique et radar de mines aurifères légales. L'image est une mosaïque biannuelle d'images Juin 2020 (www.planet.com/nicfi). L'image radar est une image Sentinel-1 en polarisation VH du 13/08/2020 et sa dynamique est réglée entre -30 (noir) et -7 dB (blanc). 47
Figure 9 - Vue satellitaire optique et radar d'une zone urbaine et de constructions illicites récentes (4°9 N, 52°3 O). L'image est une mosaïque biannuelle d'images Planet Décembre 2019 (www.planet.com/nicfi). L'image radar est une image Sentinel-1 en polarisation VH du 22/01/2020 et sa dynamique est réglée entre -30 (noir) et -7 dB (blanc). 48

Figure 10 - Exemple de fond de carte fourni à l'enquêteur à mettre dans l'application OruxMaps. Les parcelles à visiter sont encerclées en rouge et les parcelles déjà visitées ont des hachures bleues. Les routes sont représentées en orange et les déboisements détectés depuis 2016 sont les parcelles en pleine couleur. Le fond de carte est une image Sentinel-2 55
Figure 11 - Photos d'enquêteurs pendant la campagne terrain 56
Figure 12 - Exemple d'informations récupérées à la suite de la campagne terrain : parcelles spatialisées et labélisées (à gauche), fichier Excel regroupant toutes les informations associées aux parcelles (en haut à droite), images prises sur le terrain (en bas à droite) 58
Figure 13 - Principe de fonctionnement du radar imageur : l'antenne émet un signal (émission d'ondes), le sol reçoit l'énergie transmise par le radar et renvoi vers l'antenne des ondes en retour. Auteur : Lucas Ternynck - icône "satellite" par Jose Luis Algara - icône "arbre" par James Keuning 60
Figure 14 - Géométrie d'acquisition d'une radar. Source (Lardeux, 2008) 61
Figure 15 - Visée en distance d'un radar. Source (Lardeux, 2008). 64
Figure 16 - Illustration d'une réflexion diffuse sur de la végétation, spéculaire sur une surface lisse et en coin (mécanisme de double-rebond). Crédit image : Gédéon (Planète Sciences Midi-Pyrénées) 66
Figure 17 - Principales distorsions géométriques sur les images RSO avec leur dépendance à la géométrie d'acquisition : (a) le foreshortening de l'image, (b) le layover et (c) l'ombre. Copyright : (Flores-Anderson et al., 2019). 67
Figure 18 – (a) Origine physique du speckle : effet de la nature cohérente des ondes électromagnétiques utilisées pour le radar et combinaison de la contribution de chaque diffuseurs dans chaque pixel. (b) Distribution radiométrique de l'image RSO en résultant. 68
Figure 19 - Image Sentinel-1 avec différents types de filtrage. Les flèches rouges sont des repères pointant des zones d'intérêt de l'image. 69
Figure 20 - Exemple de plan d'acquisition de Sentinel-1 qui contient les orbites imagées, le temps de revisite et les polarisations disponibles. 71
Figure 21 - Orbites Sentinel-1 couvrant la Guyane 72
Figure 22 - Découpage de la Guyane en tuiles MGRS 73
Figure 23 – Gauche : Boîtes à moustaches des hauteurs des arbres en mètres avant déboisement sur les parcelles déboisées présentes dans les jeux de données. Droite : Carte de hauteur de la forêt en mètres sur la Guyane (Simard et al., 2011), avec localisation des parcelles de référence. 78
Figure 24 - Distribution des pentes sur les différents types de déboisements observés en Guyane (à gauche). Carte des pentes en degrés, avec localisation des parcelles de références (à droite). Les pentes sont dérivées du MNT SRTM à 30 mètres (http://dwtkns.com/srtm30m/). 79

Figure 25 - Illustration de polygones de référence provenant de la campagne terrain à Maripasoula (en orange), et de polygones de référence provenant des données du PAG (en bleu). Le fond de carte provient d'une mosaïque d'images Planet (www.planet.com/nicfi). 80
Figure 26 - Distribution des surfaces (en haut) et de la compacité (périmètre ² / aire) (en bas) des parcelles déboisées par jeu de données de référence. 81
Figure 27 - Distance entre les polygones des jeux de données et la couche ZONE_D_HABITATION de la BD TOPO. 82
Figure 28 - Distance entre les polygones des jeux de données et la couche SURFACE_HYDROGRAPHIQUE de la BD TOPO. 83
Figure 29 - Distance entre les polygones des jeux de données et les données de l'OAM antérieures à 2016. 84
Figure 30 - Distance entre les polygones des jeux de données et la couche TRONCON_DE_ROUTE de la BD TOPO. 85
Figure 31 - Rétrodiffusion du signal radar sur la forêt dense en Guyane (en dB) pour 3 différentes orbites (numéros relatifs : 47, 120, 141). La légende, de la forme s1a-047-vv, donne le capteur, l'orbite et la polarisation des signaux. 87
Figure 32 - Facteurs affectant le signal Sentinel-1, et donc la détection d'un déboisement. (Mermoz et al., 2018). 88
Figure 33 - Haut : signal Sentinel-1 rétrodiffusé en ascendant (gauche), et en descendant (droite). La ligne verticale verte montre le jour estimé de la coupe. La ligne verticale rouge montre la date du brûlis. Bas : image des pentes de la zone et direction des orbites ascendantes et descendantes. La flèche jaune représentant la pente pointe vers le point le plus bas. 89
Figure 34 - Photos prises le 10/08/2019 lors de la campagne terrain du haut de la parcelle et du bas de la parcelle (image du bas). Le signal radar sur la parcelle est visible Figure 33. Cette parcelle apparaît juste après sa coupe dans le reportage Notre Part d'Amazonie. 90
Figure 35 - Haut : signal Sentinel-1 rétrodiffusé en ascendant (gauche), et en descendant (droite). La ligne verticale verte représente un jour de visite sur la parcelle où le sous-bois avait commencé à être nettoyé (23/10/2019). La ligne verticale rouge représente un jour de visite sur la parcelle où les gros arbres étaient coupés (08/10/2019). Bas : photos prises de la parcelle le 10/08/2019 lors de la campagne terrain. 91
Figure 36 - Haut : signal Sentinel-1 rétrodiffusé en ascendant (gauche), et en descendant (droite). La ligne verticale verte représente un jour de visite sur la parcelle où la végétation avait commencé à sécher (16/09/2019). Milieu : image des pentes de la zone et direction des orbites ascendantes et descendantes. La flèche jaune représentant la pente pointe vers le point le plus bas. Bas : Photos prises du haut (gauche) et du bas (droite) de la parcelle le 16/09/2019 lors de la campagne terrain. 92

Figure 37 - Illustration de l'effet d'ombre du radar à synthèse d'ouverture (RSO) à la frontière entre les forêts et les zones déboisées, pour la configuration d'orbite descendante. (Image de fond arrière-plan : ConservationDrones.org/Orangutan Conservancy). (Bouvet et al., 2018). 93
Figure 38 - Schéma décrivant l'algorithme de détection du déboisement produite à la date d_{k-Xa+1} à l'arrivée de l'image de la date $d_{k.}$ 98
Figure 39 - Méthode de détection du déboisement basée sur Sentinel-1. A gauche, série temporelle du RCR d'un pixel d'une zone déboisée à laquelle les deux seuils sont appliqués. Dans cet exemple, ce pixel sera un déboisement potentiel. A droite, pixels détectés en utilisant deux seuils différents (rouge et bleu) et zone déboisée en orange (polygone bleu contenant un polygone rouge). 99
Figure 40 - Fonction de densité de probabilité (PDF) de l'indicateur min(RCR) en polarisations VV et VH pour des données de référence sur une surface totale de 64 771 ha (tuiles 22NCJ et 22NCK). Le terme "ombre" désigne les 20% de pixels de chaque parcelle ayant les valeurs min(RCR) les plus basses. 100
Figure 41 - Courbe ROC-AUC décrivant la séparabilité du min(RCR) entre les forêts intactes et les zones de pertes forestières en polarisation VH (courbe rouge) et VV (courbe bleue). 101
Figure 42 - Signal radar en VV et VH au cours de l'année 2017 pour les satellites S1A et S1B. ... 102
Figure 43 - Ombres qui apparaissent ou disparaissent lorsqu'une parcelle de forêt est déboisée, en fonction de l'orientation de l'orbite (ascendante ou descendante), de l'orientation de la parcelle, et de la configuration (parcelle à l'intérieur d'une forêt, à la lisière d'une forêt, ou isolée). La figure est simulée pour la configuration de Sentinel-1 (orientation vers la droite avec une inclinaison de l'orbite de 98 et 18) qui est résumée dans le cadre en haut à droite. (Bouvet et al., 2018) 103
Figure 44 – Schéma de la relation entre l'angle d'incidence θ, la hauteur d'un arbre Ht et la taille d'une ombre projetée sur sol plat. 104
Figure 45 - Taille de l'ombre en fonction de la hauteur de la forêt et de l'angle d'incidence de Sentinel-1. Le nombre de pixels correspondant est basé sur une taille de pixel de 10 mètres. (Bouvet et al., 2018) 105
Figure 46 – Détection d'une fausse alarme due à un agrégat de pixels dont le min(RCR) passe en dessous du seuil α à différentes dates. La figure à gauche représente les pixels formant l'ombre détectées (min(RCR)<α). La série temporelle du RCR de chaque pixel formant l'ombre est disponible dans la Figure 47. La couleur représente la date et le fond gris représente l'emprise de la détection finale. La figure du milieu représente les pixels formant la zone potentiellement déboisée (min(RCR)<β). La détection finale (la fausse alarme) est représentée sur la figure de droite. 109
Figure 47 – RCR calculé sur chaque pixel du cas d'étude de la Figure 48 au cours du temps. Les lignes verticales oranges représentent le min(RCR) de la série et la ligne verticale rouge représente le seuil α=-4.5 dB. 110

Figure 49 - Schéma de la relation entre l'angle d'incidence θ, la hauteur d'un arbre Ht et la taille d'une ombre projetée sur sol avec une pente ρ. 111
Figure 50 - Largeur de l'ombre radar (en mètres) en fonction de la pente au sol pour des angles d'incidence de 29° (bleu) et 46° (orange), avec une hauteur de forêt fixée à 35 mètres. 112
Figure 51 – Taille de la cellule de résolution (en mètres) en fonction de la pente au sol pour des angles d'incidence de 29° (bleu) et 46° (orange). 113
Figure 52 - Pourcentage de fausses détections en fonction de la topographie pour des plages de pentes données 113
Figure 53 - Détection d'un déboisement dans une zone de forte topographie. Les images du haut montrent l'évolution du déboisement mis en évidence par les contours noirs. Les fonds de carte sont des mosaïques biannuelles sans nuages d'images Planet (www.planet.com/nicfi). Les ombres (image bas milieu) ont toutes été détectées dans une pente < 15° (image bas gauche). Il est possible d'observer la temporalité de ce déboisement qui commence par la partie inférieure puis remonte (image bas droite), comme le montrent également les images optiques. Ce déboisement est celui du mont Fortuné et est due à la création de logements illégaux (Guyane : démolition d'un squat en forêt classée). 114
Figure 54 - Les polygones verts représentent la zone de forêt intacte sur laquelle les fausses détections sont comptabilisées. 116
Figure 55 - Indice de confiance $P\alpha$ en fonction du seuil de détection des ombres α. La courbe bleue est la courbe trouvée de façon expérimentale. La courbe orange est la fonction d'approximation. 117
Figure 56 - À gauche : Guyane française avec les zones de validation pour chaque type de déboisement (orange : agriculture, vert : exploitation forestière, jaune : orpaillage) ; à droite : couverture nuageuse moyenne annuelle de la Guyane française (Crédits : Imagery produced by the NASA Earth Observations team based on data provided by the MODIS Atmosphere Science Team, NASA Goddard Space Flight Center.). 125
Figure 57 - Illustration de la division d'un nœud t et de son ensemble Q_t à l'aide d'une règle binaire δ. 127
Figure 58 – Pertes forestières sur l'ensemble de la Guyane à partir de la méthode développée entre 2016 et 2019 à 10 mètres de résolution. Les images sur les bords montrent des zones de 8×12 km, mettant en évidence les différentes tailles et distributions des zones perturbées. 138
Figure 59 - Cartes des détections de pertes forestières résultant de la création d'abattis (à gauche) et d'orpaillage (à droite). Les couleurs indiquent la date de coupe détectée par l'algorithme. Les lignes vertes sont les limites des polygones de référence. La carte de gauche est centrée sur 3°40 N et 54°3O. Arrière-plan : Image RSO Sentinel-1 139

Figure 60 – Images Landsat du 22/10/2017 (gauche) et du 23/11/2017 (milieu) sur lesquelles on voit un abattis récemment coupé, et détections GLAD-FA associées en rouge (droite). Cartes centrées sur 3°2 N et 52°4 O. 143
Figure 61 - Comparaison spatiale des cartes de perte de forêt basée sur Sentinel-1 (S1) (RSO) (à gauche - Fond : image Sentinel-1) et de GLAD-FA basé sur Landsat (L7/8) (optique) (à droite - Fond : image optique aérienne Bing). Les données de référence sont dessinées en vert. Les deux cartes supérieures sont centrées sur 3°12 N et 52°25 O. 144
Figure 62 - Histogramme du pourcentage de polygones de référence couverts par les deux méthodes. 145
Figure 63 – Dates de coupe observées par la méthode basée sur Sentinel-1 des 26 parcelles visitées sur le terrain, comparées à la date de coupe déduite des observations faites sur place et des images Sentinel-2. La ligne en pointillés noire représente une date observée par Sentinel-1 égale à une date déduite du terrain. La ligne en pointillés grise représente une date observée par Sentinel-1 6 jours après la date de terrain déduite : ce délai peut être attribué aux acquisitions de Sentinel-1. 148
Figure 64 – Agriculture vivrière : nombre d'échantillons de référence détectés par mois pour les systèmes d'alertes forestières Sentinel-1 (S1) (vert) et Landsat (L7/8) (jaune) (gauche). Histogramme de la différence entre les dates de détection des polygones de référence, où les valeurs négatives indiquent une détection plus précoce par le système radar (droite) : en orange les délais pouvant provenir de l'écart entre les acquisitions de chaque capteur ; en bleu, les délais liés aux différences inhérentes aux capteurs ou aux méthodes utilisées. 149
Figure 65 - Exploitation forestière : nombre d'échantillons de référence détectés par mois pour les systèmes d'alertes forestières Sentinel-1 (S1) (vert) et Landsat (L7/8) (jaune) (gauche). Histogramme de la différence entre les dates de détection des polygones de référence, où les valeurs négatives indiquent une détection plus précoce dans le système radar (droite) : en orange les délais pouvant provenir de l'écart entre les acquisitions de chaque capteur ; en bleu, les délais liés aux différences inhérentes aux capteurs ou aux méthodes utilisées. 151
Figure 66 - Orpaillage : nombre d'échantillons de référence détectés par mois pour les systèmes d'alertes forestières Sentinel-1 (S1) (vert) et Landsat (L7/8) (jaune) (gauche). Histogramme de la différence entre les dates de détection des polygones de référence, où les valeurs négatives indiquent une détection plus précoce dans le système radar (droite) : en orange les délais pouvant provenir de l'écart entre les acquisitions de chaque capteur ; en bleu, les délais liés aux différences inhérentes aux capteurs ou aux méthodes utilisées. 152
Figure 67 – RCR en fonction de X_a sur des pixels de forêt intacte tirés aléatoirement. 154

Figure 68 – Fausses alarmes dues à la présence d'un évènement pluvieux important sur l'image Sentinel-1 du 18/10/2020. Cet évènement pluvieux important est caractérisé par la masse noire visible au milieu de l'image. 155
Figure 69 - Taux de variation du couvert forestier en Guyane, donné par différentes méthodes. La méthode basée sur Sentinel-1 est en vert, GLAD-FA en jaune et l'année de perte de couvert forestier de (Hansen et al., 2013) en orange. 156
Figure 70 - Surface déboisée (ha) d'après les tailles de parcelles (gauche). Nombre de parcelles déboisées d'après leur surface (droite). Ces données proviennent des détections effectuées avec les données Sentinel-1 en Guyane sur 2016, 2017 et 2018. 157
Figure 71 – Boîtes à moustache (gauche) et graphiques en violon (droite) de la distance aux abattis en km (haut) et de la finesse (bas), par type de déboisement (O : Orpaillage ; Ag : Agriculture ; U : Urbanisation ; EF : Exploitation forestière ; Au : Autre) 158
Figure 72 - Importance des indicateurs (les 8 plus importants) calculée sur des RF pour chaque classe. 159
Figure 73 - Score kappa en fonction du nombre d'indicateurs fournis à l'arbre de décision pour son apprentissage, pour 4 profondeurs d'arbres différentes. Les lignes verticales montrent le nombre d'indicateurs avec les scores les plus élevés pour un arbre de profondeur de 10 (grise : 3 indicateurs – kappa = 0.931, noire : 13 indicateurs – kappa = 0.9328). La ligne noire verticale montre le score le plus élevé pour un arbre de profondeur 10 162
Figure 74 - F-score par classe en fonction pour un nombre d'indicateurs de 3 et de 13. 162
Figure 75 - Classes des zones déboisées prédites par le modèle à l'échelle de la Guyane sur l'année 2018 (pour les zones ayant des références) 164
Figure 76 - Confusion en fonction de chaque classe (en pourcentage). Le titre de chaque sous-graphe donne la classe réelle. La classe prédite est donnée par les étiquettes de couleur. 165
Figure 77 – Exemples d'erreur de classification de zones déboisées. A : route construite pour l'exploitation aurifère classée comme de l'agriculture (entre Cacao et la mine de Devez). B : confusions entre agriculture et urbanisation aux alentours de Saint Laurent du Maroni. 166
Figure 78 - Photo des barges sur le Maroni (gauche), et les dégâts qu'elles entrainent sur les berges (droite). Source : Clément Villien, WWF Guyane. 174
Figure 79 - Images Sentinel-2 de berges intactes le 23/08/2018 (à gauche) et défrichées par les barges le 01/12/2019 (à droite) sur le Maroni (3°71 N, 54° O). Les images proviennent de EO Browser (https://apps.sentinel-hub.com/eo-browser/). 174
Figure 80 - Détection des berges défrichées à l'aide de Sentinel-1. La couleur correspond à la date du déboisement. 175

Figure 81 - Détection des pertes forestières via Sentinel-1 sur l'exploitation de la crique Nelson et en aval. Image centrée en 5°N 52°9O 176
Figure 82 - Boîtes à moustache (gauche) et graphiques en violon (droite) des 51 indicateurs exploités par classe (O : Orpaillage ; Ag : Agriculture ; U : Urbanisation ; EF : Exploitation forestière ; Au : Autre). Les distances et les périmètres sont exprimés en km, l'élévation est donnée en mètres et les surfaces sont exprimés en ha. 224

Table des tableaux

Tableau 1 - Tableau récapitulatif du nombre de parcelles visitées au cours de la campagne terrain. A = ancien abattis, N = nouvel abattis, D = parcelles diverses, EL = erreur de localisation, NV = parcelle non visitée.....59
Tableau 2 - Tableau récapitulatifs des ensembles de données utilisés dans cette thèse. 74
Tableau 3 - Nombre et taille des polygones de référence pour les différents types de déboisement en Guyane. 124
Tableau 4 - Nombre d'observations (pixels) disponibles par classe 133
Tableau 5 - Validation de la méthode développée avec les données Sentinel-1 (S1) (RSO) et de GLAD-FA (optique). 140
Tableau 6 - Moyenne de la moyenne et de l'écart-type du RCR.....153
Tableau 7 - Rang d'importance des indicateurs pour la caractérisation des déboisements en Guyane 160
Tableau 8 – Score de validation du modèle prédictif pour la caractérisation du déboisement.....165
Tableau 9 – Moyenne et écart-type des indices de confiance des pixels de chaque classe. 167
Tableau 10 - Matrice de confusion associée à la détection de la perte de forêt.....201
Tableau 11 - Indicateurs utilisés pour le développement d'un arbre de décision visant à caractériser les déforestations en temps-réel en Guyane, avec leurs sources. La source est décrite de la manière suivante : NOM_DE_LA_COUCHE : liste des conditions sur les classes sélectionnées (séparées par une virgule). DI signifie : date à laquelle on veut l'état de la variable. 204

Chapitre 1 - Introduction

Contenu

1. Importance de la forêt 26

2. Objectifs de recherche et structure du manuscrit 29

1. Importance de la forêt

Les forêts du monde fournissent de nombreux services écosystémiques à l'humanité. Elles abritent environ 80% de la biodiversité terrestre mondiale, procurant un habitat à 80% des espèces d'amphibiens, 75% des espèces d'oiseaux et 68% des espèces de mammifères (*Évaluation des ressources forestières mondiales 2020*, 2021). Plus d'un milliard de personnes en dépendent directement pour leur alimentation, leur logement, leur énergie ou leurs revenus (Pacheco et al., 2021). De plus, les forêts interviennent significativement dans l'atténuation du changement climatique par le stockage de grandes quantités de carbone (662 gigatonnes en 2020 (*Évaluation des ressources forestières mondiales 2020*, 2021)). Elles jouent également un rôle important dans le cycle de l'eau et l'approvisionnement en eau douce (Mbow et al., 2017).

Les forêts comptent pour 31% de la superficie totale des terres émergées (*The State of the World's Forests 2020. In brief*, 2020) et 45% de leur superficie se situe en milieu tropical (Figure 1). Environ 50% de la forêt tropicale se trouve sur le continent américain (Amazonie), 30% en Afrique (Bassin du Congo) et 20% en Asie (Asie du Sud-Est). Les forêts tropicales, en particulier, sont essentielles pour participer à l'atténuation du changement climatique (Canadell and Raupach, 2008; Le Quéré et al., 2009; Stern and Stern, 2007). Elles abritent la biodiversité la plus riche de notre planète (Laurance et al., 2012; Lewis et al., 2015). Près de la moitié des espèces animales vivent dans ces forêts et dépendent d'elles pour survivre : tigres, orangs-outans, toucans, etc. C'est également le milieu naturel d'environ 60% de l'ensemble des plantes vasculaires.

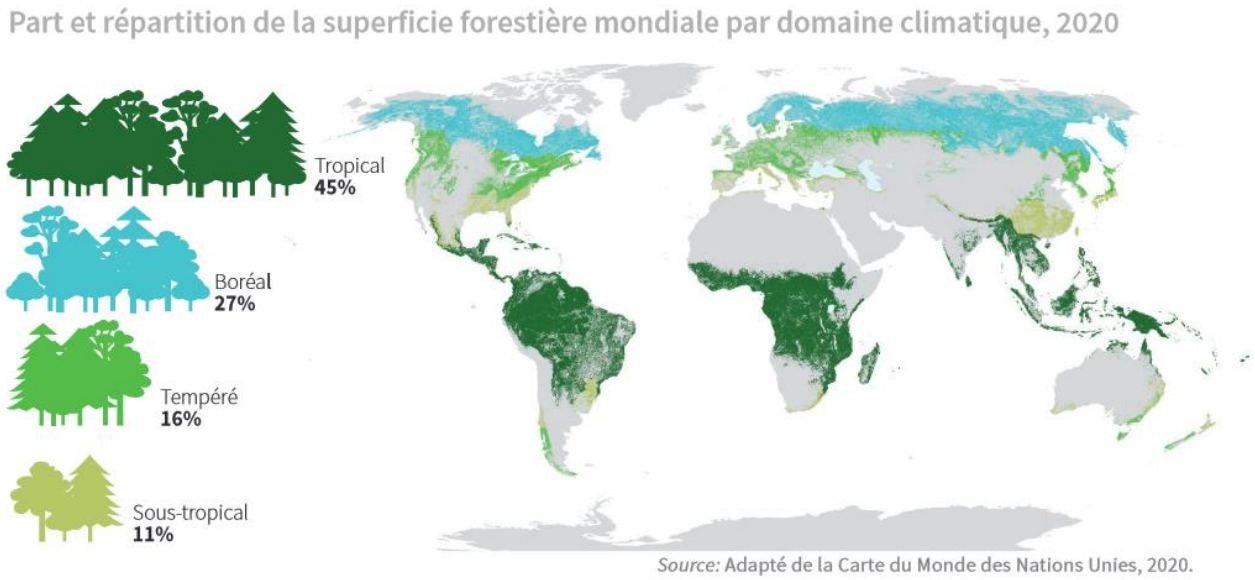


*Figure 1 - Couverture spatiale des différentes forêts dans le monde. Source (*The State of the World's Forests 2020. In brief, *2020)*

Depuis l'an 2000, plus de 3.5 millions de kilomètres carrés de couverture forestière mondiale ont été perdus ou dégradés (Hansen et al., 2013). Les forêts tropicales sont gravement touchées et disparaissent à un rythme alarmant : en 2019, l'équivalent d'un stade de football y est détruit toutes les deux secondes (*The State of the World's Forests 2020. In brief*, 2020)

La Forest Ressource Assessment 2020 (*The State of the World's Forests 2020. In brief*, 2020) définit la déforestation comme étant la conversion de la forêt à d'autres affectations, comme l'agriculture et les infrastructures. Elle implique une régression durable voire définitive des surfaces couvertes par les forêts. Quand la perte forestière est temporaire, on parle de déboisement, défrichement ou déforestage. La dégradation forestière correspond à une perturbation forte, qui engendre une diminution de la capacité de la forêt à fournir des produits et services. Dans ce cas, même si la perte de couvert et/ou de surface terrière est assez forte, la forêt est toujours considérée comme telle. Quand un déboisement ne met pas en cause la capacité de la forêt à se renouveler naturellement, on parle de perturbations ou d'ouvertures de la canopée limitées.

Le facteur principal de ces pertes forestières est l'expansion de l'agriculture industrielle (pâturages et cultures) et de subsistance (Figure 2). Ces déboisements augmentent, entraînés par l'accroissement démographique, l'augmentation de la demande en nourriture, et la création d'infrastructures permettant l'accès à des zones jusqu'ici préservées (Pacheco et al., 2021).

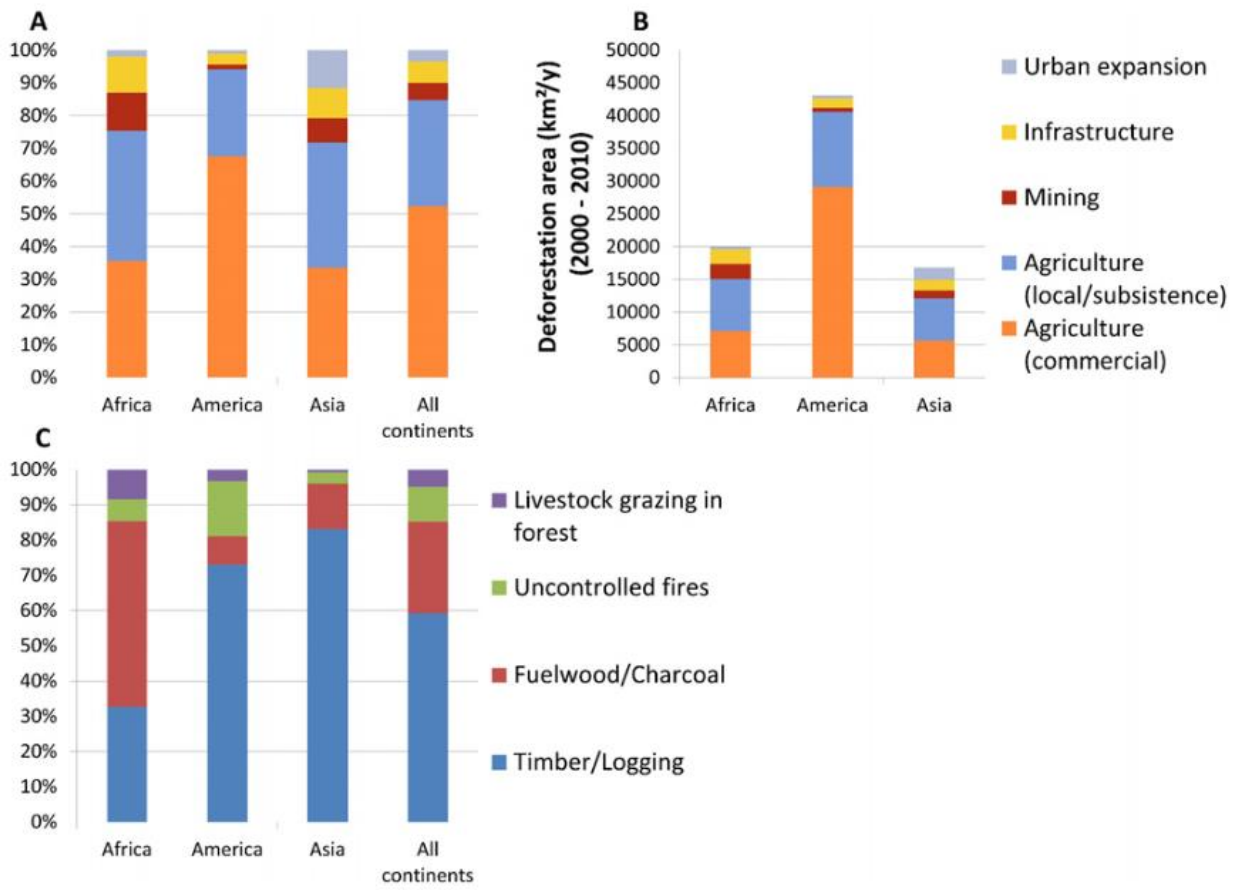


*Figure 2 - Estimations à l'échelle continentale de (A) la proportion de chaque facteur de déforestation, (B) la variation de la superficie forestière (km2/an-1 ; (*The State of the World's Forests 2020. In brief, *2020) pour la période 2000-2010 par facteurs de déforestation, et (C) la proportion de la superficie forestière perturbée par facteur de dégradation. L'étude a été faite sur la base de données provenant de 46 pays tropicaux et subtropicaux. Source : (Hosonuma et al., 2012).*

La conversion des forêts tropicales pour un usage du sol différent constitue une préoccupation majeure de la communauté internationale, car ces déforestations et dégradations ont des conséquences importantes.

Le déboisement génère des émissions de CO_2 importantes et diminue la capacité de stockage du carbone par les forêts. La déforestation et la dégradation de la forêt tropicale ont entraîné respectivement 77% et 13% du flux net total des émissions anthropiques de CO2 attribuable aux changements d'utilisation et d'occupation du sol entre 1850 et 2015 (estimé à 145 ± 16 PgC à l'échelle mondiale, dont 102 ± 5.8 PgC dans les tropiques, selon (Houghton and Nassikas, 2017)).

Celles-ci contribuent aussi grandement à la perte de biodiversité actuelle (Mittermeier et al., 1998; Smith et al., 2014). De plus, la santé et le bien-être des personnes sont étroitement liés aux forêts : la disparition de l'habitats de certaines espèces suite à une modification des massifs forestiers et à l'expansion démographique peut mener à l'émergence de zoonoses (maladies transmises entre l'humain et les animaux vertébrés) (Guégan et al., 2020).

Le suivi des perturbations subies par ce milieu est donc un défi d'actualité car il représente un danger immédiat pour l'environnement et la biodiversité (Finer et al., 2018). Les études menées ces dernières décennies sur l'évolution des forêts et les conséquences associées ont permis une prise de conscience de la situation actuelle par les sociétés. En réaction, plusieurs initiatives internationales comme les Objectifs de Développement Durable des Nations Unies (objectifs 13 et 15) et les accords d'Aichi pour la biodiversité (objectif 5) ont été instaurées. Aujourd'hui, les scientifiques ont un rôle à jouer pour atteindre ces objectifs en fournissant aux acteurs des forêts, des méthodes et outils pour les aider à gérer leur territoire.

2. Objectifs de recherche et structure du manuscrit

Dans cette thèse, en partenariat CNES-WWF-UGE, on propose d'étudier des méthodes pour suivre l'évolution de l'écosystème forestier et pour cartographier ses changements, avec la Guyane comme site d'étude.

L'effort de recherche est mis sur la détection des déforestations par imagerie radar et leur caractérisation dans un contexte quasi-temps réel. La comparaison des performances de détection des données radar avec des détections basées sur l'imagerie optique, très utilisées, est également un des objectifs.

Les méthodes développées visent à être valides à l'échelle de territoires tropicaux, sur des parcelles de surface minimale de 0.2 ha. L'effort est également mis sur la possibilité de les rendre opérationnelles pour qu'elles puissent être utilisées par les acteurs locaux.

Ce manuscrit est constitué de 8 chapitres incluant cette introduction (**chapitre 1**) et la conclusion. Le **chapitre 2** apporte une vue globale de la déforestation tropicale, ainsi que de ses causes et de ses pratiques. L'état de l'art de l'utilisation de la télédétection pour le suivi de la forêt est également abordé. Le site d'étude et les données utilisées dans cette thèse sont présentés dans le **chapitre 3**. Celui-ci inclut la présentation des données cartographiques disponibles sur la zone, les données de terrain collectées, ainsi que les données radar Sentinel-1. L'analyse de ces données cartographiques et l'interprétation des images radar sur les données de référence sont abordées **chapitre 4**. De ces observations découlent les méthodes développées, ainsi que les méthodes de validation associées aux produits. Celles-ci sont exposées dans le **chapitre 5**. Dans le **chapitre 6** les résultats sont présentés, ainsi que leurs évaluations. Ensuite, la valorisation des données produites et les perspectives de travaux futurs en rapport avec cette thèse sont abordées **chapitre 7.** Pour finir, le **chapitre 8** conclut ce manuscrit avec une synthèse des différents travaux effectués au cours de cette thèse.

Chapitre 2 - Les pertes forestières tropicales

Contenu

1. Causes des pertes forestières tropicales 30

2. Pratiques diverses de déboisement 33

3. État de l'art de l'utilisation de la télédétection pour le suivi des forêts 34

Ce chapitre donne une vue d'ensemble des causes et des pratiques associées au déboisement. Les comprendre est important pour déterminer comment l'imagerie satellitaire peut aider à cartographier et à suivre les forêts tropicales. L'état de l'art de l'utilisation de la télédétection pour le suivi des forêts est exposé à la fin de ce chapitre.

1. Causes des pertes forestières tropicales

Récemment, la fondation WWF (World Wide Fund for Nature) a synthétisé différentes études sur les causes de la déforestation dans son rapport (Pacheco et al., 2021). La déforestation tropicale est principalement due à :

- L'expansion agricole (comprenant par exemple l'agriculture à large échelle, l'agriculture vivrière, les plantations d'arbres à moyenne et large échelle, et les pâturages) ;
- L'extraction minière (or, bauxite, etc.) ;
- L'urbanisation (comprenant la construction d'infrastructures de transport, hydroélectriques et l'expansion urbaine) ;
- L'exploitation forestière (comprenant l'exploitation forestière à large et petite échelle, l'extraction pour le bois de chauffage et le charbon de bois) ;
- Les feux naturels ou d'origine anthropique (pouvant être lancés dans le but de déboiser).

Parmi ces facteurs directs de déforestation, c'est l'expansion de l'agriculture industrielle qui a la plus grande influence sur la déforestation en Amazonie et en Asie du Sud-Est. Cependant, la contribution des petits exploitants est croissante et constitue même la cause principale de déforestation dans le Bassin du Congo.

L'exploitation forestière est aussi un facteur important de la déforestation dans certains pays, même si elle a tendance à diminuer globalement ces dernières années (Pacheco et al., 2021). La pression sur les forêts s'accroît particulièrement dans les zones où le réseau routier se développe, comme en Amazonie. Ces causes de déforestation sont influencées par des facteurs indirects comme

la croissance démographique et l'augmentation de la demande de bois, de nourriture et de matériaux (Pacheco et al., 2021).

Dans le **Bassin du Congo**, l'agriculture vivrière est la première cause de déforestation et continue à croître (Tegegne et al., 2016). L'agriculture à large échelle augmente dans ces régions, et impacte déjà significativement le sud-ouest du Cameroun, notamment avec les plantations d'hévéas (l'arbre à caoutchouc), et de palmiers à huile (Ordway et al., 2017). La dégradation forestière dans le Bassin du Congo est causée notamment par l'exploitation forestière artisanale et industrielle, ainsi que par la demande de bois-énergie. Au Cameroun, le marché intérieur (principalement approvisionné par l'exploitation forestière à petite échelle) est plus important que le marché d'exportation approvisionné par les grands opérateurs commerciaux. En République du Congo, la demande en bois-énergie est significative, et impacte particulièrement les alentours des grandes villes, dans lesquelles on estime que plus de 90% de la population utilise le charbon de bois pour cuisiner, en raison du manque d'alternatives énergétiques abordables (Ickowitz et al., 2015). L'exploitation minière artisanale, notamment aurifère, reste une cause moindre de la déforestation, mais augmente sensiblement depuis une quinzaine d'années. Au Gabon, elle est déjà bien installée et s'étend aux paysages forestiers, y compris à la frontière avec le Cameroun (Bank, 2012).

La perte de la forêt tropicale en **Asie du Sud-Est** est principalement due à l'agriculture et à la plantation d'arbres industrielle. Au Laos, les pressions liées à ces facteurs augmentent. De vastes zones y ont été converties en cultures annuelles (maïs, manioc, canne à sucre, etc.) et en plantations commerciales pérennes (caoutchouc, café, cacao, poivre, etc.) pour les marchés régionaux et mondiaux (Costenbader et al., 2015). De plus, les petites plantations d'arbres à croissance rapide comme l'eucalyptus et l'acacia ont été encouragées par le gouvernement (Chanthalath, others, 2017). En Indonésie et en Malaisie, ces causes restent principales. Cependant, la déforestation liée à l'agriculture à large échelle (principalement due au palmier à huile) diminue, ainsi que les plantations d'arbres pour la pâte à papier. En revanche, l'agriculture à petite échelle augmente, ce qui est lié à une augmentation significative des petits exploitants indépendants fournissant de l'huile de palme (Bosc, others, 2018; Jelsma et al., 2019). Les feux, à la base utilisés pour « nettoyer » les terres après des coupes forestières, sont devenus un facteur important et grandissant de la déforestation, notamment par leur propagation dans les tourbières. Ces dernières ont été drainées par l'humain car les palmiers n'y poussaient pas bien naturellement, seulement, cette tourbe sèche prend feu très rapidement. Pour finir, la construction de routes exerce également des pressions importantes et grandissantes sur la forêt tropicale en Asie du Sud-Est (Sloan et al., 2019). La

distance par rapport aux routes principales est même l'un des facteurs prédictifs les plus importants de la déforestation entre 2006 et 2012 au Laos (Phompila et al., 2017).

En **Amazonie**, la déforestation provient principalement de l'agriculture : une activité qui augmente constamment. L'élevage bovin s'est largement développé via des systèmes extensifs et peu productifs au Brésil, en Colombie, en Bolivie et au Pérou (Finer et al., 2018). Ces pâturages sont parfois liés à la spéculation foncière : leur établissement se faisant pour justifier la propriété foncière (Armenteras et al., 2013; Seymour and Harris, 2019). Bien que l'agriculture à grande échelle soit une cause importante de la déforestation en Amazonie, c'est l'agriculture à petite échelle qui a le plus d'impact. L'agriculture itinérante est persistante et importante au Vénézuela, au Guyana, en Guyane et au Suriname (*WRI - Global Forest Watch Dashboard*, 2020). Au Pérou et en Colombie, la culture de la coca croît depuis une dizaine d'année sur des petites parcelles non contiguës (González et al., 2018). Les cultures de rente (comme le café ou le cacao), au Brésil et au Pérou se développent également (Piotrowski, 2019 ; Dávalos, 2016).

Les opérations minières sont présentes partout en Amazonie. Peu importantes en Bolivie et en Colombie, elles prennent de l'importance au Brésil, au Vénézuela et au Pérou, et deviennent aujourd'hui des causes primaires de la déforestation au Suriname et au Guyana où elles représentent 84% de la déforestation (Laing, 2019). L'orpaillage illégal à petite échelle exerce une forte pression sur la forêt, sur l'environnement et sur la biodiversité, notamment à cause de l'utilisation de mercure disséminé dans la nature (Alvarez-Berríos and Aide, 2015 ; Piotrowski, 2019 ; Caballero Espejo et al., 2018 ; Diele-Viegas and Rocha, 2020 ; Kalamandeen et al., 2018 ; Asner and Tupayachi, 2017).

L'expansion des routes est un facteur croissant de déforestation au Brésil, en Colombie, en Bolivie et au Pérou. Ces routes, en plus d'être elles-mêmes des changements forestiers, facilitent l'accès à des zones de forêt difficilement accessibles, et provoquent l'accroissement d'opérations d'exploitation forestière (Perz et al., others, 2008), l'expansion de petits exploitants (Rojas Briceño et al., others, 2019), et encouragent l'occupation des terres à long terme (Baraloto et al., others, 2015). Par exemple, la transamazonienne construite dans les années 1970 au Brésil a favorisé le développement agricole (Godar et al., 2012 ; Li et al., 2019).

En Guyane, les déboisements sont principalement liés à l'orpaillage (Rahm et al., others, 2015) et à l'agriculture vivrière, même si le gouvernement actuel encourage un développement agricole à des fins lucratives (Demaze and Manusset, 2008). Dans une moindre mesure, l'urbanisation liée à la croissance démographique entraîne également des défrichements, ainsi que l'exploitation forestière.

L'augmentation de la population humaine, la demande globale en nourriture (notamment d'origine animale), en énergie et en produit de base continuera à entrainer le déboisement des forêts tropicales, à moins que les décideurs politiques régissant les forêts et les entreprises privées opérant dans ces régions ne fassent de la conservation des forêts une priorité (Zeppetello et al., 2020).

2. Pratiques diverses de déboisement

Observer efficacement les pertes forestières passe par la connaissance des mécanismes qui ont lieu lorsqu'elle se produit. Cette section expliquera les diverses pratiques qui mènent à la déforestation et à la dégradation forestière.

En Asie du Sud Est, notamment au Vietnam, les déforestations liées aux plantations entraînent des **coupes rases** (Mermoz and Le Toan, 2016) : les parcelles coupées sont complètement nettoyées en enlevant la végétation restante au sol et laissent place (juste après la coupe) à du sol nu. Ce type de coupe se retrouve également sur d'autres territoires, dans les grands projets d'exploitation minière et lors de la création d'infrastructures urbaines ou de transport.

Le **feu** est un moyen très utilisé en Amazonie pour l'établissement de pâturage ou pour l'agriculture itinérante. Il en est de même en Indonésie et en Malaisie pour les plantations de palmiers à huile. Les parcelles sont d'abord grossièrement coupées, puis, quand elles sont sèches, brûlées afin de les « nettoyer ». Cette opération peut être répétée plusieurs années d'affilée avant que la terre ne soit utilisée (comme pour les pâturages au Brésil).

La **coupe sélective** entraîne des dégradations : c'est une pratique qui se retrouve principalement dans l'exploitation forestière où seuls les arbres qui ont une valeur commerciale sont coupés.

En Guyane, ces trois pratiques sont présentes, la première étant principalement utilisée pour l'orpaillage légal, la deuxième pour la construction d'abattis et l'exploitation forestière pouvant s'apparenter à la troisième (bien que l'exploitation forestière en Guyane est telle qu'elle n'engendre pas de dégradations).

La plupart des déboisements tropicaux ont lieu pendant les saisons sèches, plus favorables au travail, à la conduite de machines éventuelles et aux feux. Cependant, les pratiques ont évolué : au Brésil, un programme de surveillance opérationnel de l'Agence Spatiale Brésilienne (INPE : Instituto Nacional de Pesquisas Espaciais) basé sur des images satellitaires Landsat a permis de diviser par trois la déforestation entre 2004 et 2015. Ce programme permet de détecter les coupes illégales survenant en saison sèche comme c'était le cas auparavant. Aujourd'hui, les coupes se

déroulent davantage en saison humide afin d'être moins détectables par les autorités, et pour compliquer leurs interventions potentielles. En revanche, certaines activités illégales comme l'orpaillage clandestin se déroulent tout au long de l'année.

Aussi, Kalamandeen et al., 2018 a montré qu'en Amazonie, la taille moyenne des parcelles déboisées avait diminué entre 2001 et 2014, provenant d'une augmentation des parcelles inférieures à 6.25 ha (seuil correspondant à la surface minimale détectée par PRODES (Valeriano et al., 2004)), et en particulier de celles inférieures à 1 ha (+ 34%) entre 2001-2007 et 2008-2014, et d'une diminution du nombre de parcelles supérieures à 6.25 ha traduisant les tentatives des grands propriétaires fonciers d'échapper aux systèmes de surveillance de la déforestation. Entre 2008 et 2014, les parelles inférieures à 1 ha représentaient environ 20% de la surface déboisée en Amazonie. Le déboisement de petites parcelles n'a pas seulement augmenté en nombre, il a aussi augmenté en terme d'étendue géographique : entre 2001-2007 et 2008-2014, la densité des pertes forestières (en km^2 pour 100 km^2) a augmenté d'un ordre de grandeur dans de grandes zones de l'Amazonie généralement associées à de très faibles taux de déboisement. Ce phénomène se produit dans tout le bassin amazonien, soulignant la nature éparse des petits événements de perte de forêt (Kalamandeen et al., 2018).

En Guyane, les activités liées à l'agriculture se déroulent principalement en saison sèche, alors que l'orpaillage a lieu tout au long de l'année et les parcelles déboisées dépassent rarement 10 ha.

Afin de pouvoir répondre à ces perturbations, des outils de surveillance à grande échelle sont nécessaires. La télédétection apparait comme un outil pertinent pour suivre les forêts qui peuvent se révéler immenses et inaccessibles, et dans lesquelles un suivi au sol serait impossible.

3. État de l'art de l'utilisation de la télédétection pour le suivi des forêts

L'utilisation de l'imagerie satellitaire pour le suivi des changements forestiers a commencé dès 1988, année depuis laquelle l'INPE surveille avec succès les coupes rases opérées sur son territoire avec son système de détection de la déforestation PRODES (Valeriano et al., 2004). Historiquement basé sur les images optiques Landsat 5, ce système permet de produire des cartes annuelles avec une surface de détection minimale de 6.25 ha. Ce système a cependant l'inconvénient de nécessiter une étape de photo-interprétation impliquant une action humaine.

GFW[1] produit et distribue des cartes annuelles à l'échelle mondiale de manière automatique, basées sur les images Landsat à une résolution de 30 mètres (Hansen et al., 2013). Ces cartes annuelles sont utiles pour de nombreuses applications, comme l'estimation des émissions de CO2 (Friedlingstein et al., others, 2019). Elles sont également utilisées par les programmes internationaux de conservation des forêts, comme REDD+.

Les systèmes d'alerte de perte forestière complètent ces observations annuelles par une surveillance plus fréquente. Opérationnels, ils devraient permettre aux États ou aux gestionnaires de forêts de lutter, principalement, contre les activités illégales. Ils peuvent également être utilisés pour la gestion des aires protégées, la gestion des concessions agricoles et autres concessions productives, le suivi des forêts habitées, et pour la sensibilisation du grand public. Ils jouent un rôle précieux dans l'application de la loi et la conservation des forêts, mais la plupart d'entre eux ne sont pas conçus pour quantifier l'ampleur des changements.

Plusieurs organismes de recherche et organisations gouvernementales ont mis au point de tels systèmes d'alerte, principalement basés sur des données de télédétection optique. Avec une faible résolution spatiale (données MODIS, à 250 mètres), les systèmes FORMA (Wheeler et al., 2014) Terra-I (Reymondin et al., 2019) fournissent des alertes à l'échelle pantropicale, et IDEAM fait de même pour la Colombie. Ils sont respectivement mis à jour toutes les deux semaines, tous les mois et tous les trimestres. DETER-B (Diniz et al., others, 2015), un système opérationnel brésilien, fournit des résultats avec une résolution spatiale de 60 mètres et une fréquence de 5 jours. Il est développé par l'INPE, basé sur les données AWiFS et inclut une étape de photo-interprétation. Enfin, avec la résolution des données Landsat à 30 mètres, le MINAM (Pérou) et l'Université du Maryland (UMD) produisent chaque semaine des cartes d'alerte : PNCB Early Warning Alerts et Global Land Analysis and Discovery (GLAD) FA (Hansen et al., 2016).

Cependant, ces systèmes sont basés sur l'imagerie satellitaire optique sensible à la couverture nuageuse. Dans les tropiques, cela entraîne des délais de détection pendant les saisons humides. Ces délais ne sont cependant pas quantifiés en absolu à cause du manque de données adéquates.

Les images radar à ouverture synthétique (RSO), insensibles à la couverture nuageuse, ont un grand potentiel dans les zones tropicales pour construire des systèmes opérationnels d'alertes forestières. Cependant, elles sont encore rarement utilisées pour le suivi des déboisements par rapport à l'imagerie optique (malgré plusieurs travaux de recherche, voir (Lardeux et al., 2019 ; Mermoz and Le Toan, 2016 ; Lohberger et al., 2018 ; Reiche et al., 2018b) Le système JJ FAST

[1] Global Forest Watch (GFW) est une plateforme qui fournissant des données libres et des outils sur le monde entier afin de mieux protéger les forêts (https://www.globalforestwatch.org/).

développé par la JAXA/JICA, fait exception : basé sur les données radar ALOS-2, il produit des alertes sur 77 pays tropicaux tous les 1.5 mois et avec une résolution spatiale de 5 ha (Hayashi et al., others, 2017). Cette faible utilisation des données RSO provient en partie de leur faible disponibilité (Reiche et al., 2016) avant le programme Sentinel-1.

Avant le lancement de ce programme, les méthodes développées étaient basées sur les calculs de changement du signal entre 2 dates, comme Joshi et al., 2015 qui utilise des données RSO en bande L, mais l'absence de revisite fréquente ne permettait pas de mettre en place un suivi régulier. Avec l'arrivée de Sentinel-1, les images RSO sont désormais accessibles à l'échelle mondiale, librement, avec des acquisitions systématiques tous les 6 à 12 jours (selon le lieu), quelles que soient les conditions météorologiques et à une résolution spatiale de 20 mètres. De plus, la continuité du programme est assurée jusqu'en 2030 au moins. Toutefois, Sentinel-1 opère en bande C et par rapport aux longueurs d'ondes plus grandes (telles que la bande L ou P), cette bande de fréquence n'est pas la plus adaptée pour détecter la déforestation sur une image, car il peut y avoir ambiguïté entre la rétrodiffusion sur la forêt et les zones déboisées. En effet, la rétrodiffusion de ces dernières est variable en fonction des conditions de surface au sol (rugosité de la surface, teneur en humidité du sol, restes de végétation et débris). Plusieurs méthodes ont tout de même vu le jour, en se basant sur les séries temporelles des images Sentinel-1, que ce soit pour détecter la perte forestière (Bouvet et al., 2018 ; Lardeux et al., 2019), ou pour détecter certaines causes de déboisement (Forkuor et al., 2020 ; Reiche et al. 2018b). La plupart des méthodes de détection des déboisements se basent sur l'hypothèse que celui-ci est caractérisé par une baisse du signal radar.

Certaines études ont également exploré la fusion des données radar et optiques, notamment (Reiche et al., 2018a) et (Hirschmugl et al., 2020).

Plus un système d'alerte sera précis et rapide, plus il sera utile (Finer et al., 2018). Aujourd'hui, la détection des déboisements supérieurs à 3 ha est maîtrisée, mais les sites éphémères (d'orpaillage par exemple) sont sous-estimés car la végétation reprend avant qu'une image optique puisse les détecter (Kalamandeen et al., 2018). Aussi, plusieurs études se rejoignent pour dire que la plupart des systèmes opérationnels en place sous-estiment les déboisements inférieurs à 2 ha (Milodowski et al., 2017), alors qu'en parallèle, comme vu précédemment, les parcelles déboisées sont de plus en plus petites.

L'effort doit donc être mis sur la détection des déboisements de faible dimension, « isolés », et sur la réduction du délai de détection.

Par ailleurs, pour que la surveillance de la déforestation en quasi-temps réel soit efficace, ces alertes ne sont pas suffisantes. Finer et al., 2018 a mis en place un protocole dans lequel la détection des pertes forestières n'est que la première étape. La deuxième étape est la hiérarchisation des alertes pour déterminer les zones sur lesquelles il est le plus urgent d'agir. Ceci peut être fait par exemple en intégrant des données spatiales telles que les zones protégées, les territoires autochtones ou toute autre zone d'intérêt spécifique, ou en identifiant les zones avec une densité d'alertes forte (Weisse et al., 2017). Vient ensuite la caractérisation des déforestations : l'identification des causes de la déforestation en temps réel est nécessaire aux systèmes d'alertes précoces afin de pouvoir traiter ces alertes (Mora, 2018). La réponse des acteurs forestiers ne sera pas la même qu'il s'agisse d'un déboisement dû à l'orpaillage ou à la création d'une petite parcelle agricole. Ces déforestations n'auront d'ailleurs pas le même impact sur l'environnement.

Aujourd'hui, les données disponibles pour identifier les facteurs de déforestation en temps réel sont limitées (Kissinger et al., 2012), et cette tâche se fait principalement par inspection visuelle (Finer et al., 2018). Les études liées aux facteurs de déforestation sont principalement des évaluations régionales à mondiales (Curtis et al., 2018 ; DeFries et al., 2010 ; Boucher et al., 2011 ; Leblois et al., 2017), et leur but principal est d'expliquer le « pourquoi » de la déforestation. Cependant, ces méthodes ne sont pas adaptées au temps réel.

Des méthodes ont été développées pour classifier rapidement les causes de perturbations détectées : seuillage d'un indice spectral, modèle de *random forest* (RF) avec un ou plusieurs indices spectraux (Murillo-Sandoval et al., 2018) et modèle RF avec des variables spatio-temporelles. Cependant, ces méthodes sont basées sur des images optiques dépendantes de la couverture nuageuse persistante sur les tropiques, ce qui implique des délais de classification. Cela a été identifié comme une limitation actuelle des systèmes d'alerte (Mora, 2018).

L'utilisation des images RSO pourrait répondre au besoin d'identification des facteurs de déforestation en quasi-temps réel, mais peu d'études se sont penchées sur le sujet. Reiche et al., 2018b a combiné la détection en temps réel de la déforestation avec Sentinel-1, au système de détection d'alerte de feux basé sur VIIRS (Giglio et al., 2016) afin de déduire les déforestations dont la cause directe est le feu. La méthode développée par (Nicolau et al., 2021) est basée sur l'utilisation, dans un arbre de décisions, des séries temporelles Sentinel-1 en amplitude, ainsi que des métriques statistiques pour chaque classe. Cette méthode nécessite 3 mois d'images pour le calcul des métriques, ce qui entraîne un délai de classification fixe, mais important. De plus, sa précision globale de 52% tend à montrer que l'utilisation de Sentinel-1 n'est pas adaptée à la caractérisation exhaustive des déforestations sur un territoire.

L'effort doit donc être mis sur le développement de méthodes de classification des causes de la déforestation ne nécessitant pas l'attente de nouvelles données.

🔔 **Ce qu'il faut retenir**

- Les déboisements en forêt tropicale sont principalement liés à l'agriculture (industrielle et de subsistance), mais aussi à l'urbanisation, l'extraction minière, l'exploitation forestière, et aux incendies ;
- Les coupes peuvent être rases, sélectives ou basées sur l'utilisation du feu ;
- Les parcelles déboisées de manière illégale l'étaient surtout en saison sèche, mais le sont de plus en plus en saison humide, et leurs surfaces diminuent ;
- La surveillance des zones supérieures à 3 ha est considérée comme maitrisée, ce qui n'est pas le cas des surfaces plus petites ;
- Les systèmes d'alerte et de caractérisation des déboisements sont principalement basés sur l'utilisation des images optiques, entraînant des délais de détection, même si les méthodes basées sur le radar sont de plus en plus étudiées depuis l'arrivée de Sentinel-1.

📎 **Ce que nous proposons**

Au regard de l'état de l'art de l'utilisation de la télédétection pour le suivi des forêts, nous proposons donc 2 axes d'étude : la détection des déboisements à partir de 0.2 ha en quasi temps-réel en utilisant les données RSO Sentinel-1, et la labélisation de ces déboisements avec leur cause, en temps-réel et de manière automatique, en utilisant des données exogènes. Ces méthodes auront pour objectifs d'être appliquées et validées à large échelle.

Le site d'étude choisi pour le développement et la validation de ces méthodes est l'ensemble du territoire de Guyane, région française d'outre-mer. Celle-ci présente de multiples intérêts en tant que site d'étude pour le suivi de la déforestation. D'abord, c'est un territoire avec de nombreuses données de télédétection et de terrain disponibles. De plus, il bénéficie d'un effort de partage de ces données entres acteurs locaux, notamment via le portail Géo Guyane (https://www.geoguyane.fr/). Ces acteurs sont sensibilisés à l'utilisation de la donnée spatiale, particulièrement pertinente pour suivre une vaste étendue peu accessible telle que la forêt Guyanaise, ce qui facilite le dialogue.

De plus, de nombreuses déforestations en Guyane sont de petite taille (inférieures à 1 ha), en dessous des surfaces minimales détectées par la plupart des systèmes de détection existants. Leur détection représente donc un défi permettant d'évaluer les méthodes développées. Elles sont également fréquemment de faible densité et peuvent être isolées. Enfin, les déforestations en

Guyane sont dues à plusieurs facteurs, offrant un large éventail de coupes avec des techniques, des temporalités, et des tailles différentes.

✋ Ma contribution

Le premier axe de cette thèse (la détection des pertes forestières en quasi temps-réel) était défini dès le début de celle-ci. Suite à mes discussions avec les acteurs locaux, j'ai compris que connaître les causes des déboisements était important pour que les cartes de pertes forestières soient plus largement utilisées, notamment sur le terrain. Ainsi, pour le deuxième axe de cette thèse, j'ai moi-même proposé d'étudier la caractérisation des déboisements car cela ressortait comme une information nécessaire à la lutte contre la déforestation dans la littérature et dans les conférences.

Chapitre 3 - Site d'Etude et Données

Contenu

1. Site d'étude – La Guyane 41
1.1. La Guyane 41
1.2. Les pertes forestières en Guyane 42
1.2.1. L'agriculture 42
1.2.2. L'exploitation forestière 44
1.2.3. L'orpaillage 46
1.2.4. L'urbanisation 48
2. Données cartographiques 49
2.1. Données de référence 49
2.1.1. Données agriculture 49
2.1.2. Données exploitation forestière 49
2.1.3. Données orpaillage 50
2.2. Données auxiliaires 50
2.2.1. BD TOPO de l'IGN 50
2.2.2. CAMINO 51
2.2.3. BRGM 52
2.2.4. ECOSEO 52
2.2.5. GLAD-FA 52
3. Données de terrain acquises dans le cadre de cette thèse 53
3.1. Motivations 53
3.2. Mode opératoire – données collectées 54
3.3. Analyse a posteriori des données 57
4. Données RSO 59

4.1. Télédétection RSO 59
4.1.1. Radar imageur 59
4.1.1. Radar a Synthèse d'Ouverture (RSO) 62
4.2. Les images RSO 63
4.2.1. Résolution spatiale 63
4.2.2. Taille de pixel 64
4.2.3. Caractéristiques du coefficient de rétrodiffusion 65
4.2.4. Propriétés géométriques 67
4.2.5. Propriétés radiométriques 68
4.3. Sentinel-1 70

Ce chapitre a pour but de présenter le site d'étude ainsi que les données utilisées dans cette thèse pour répondre aux problématiques définies dans le chapitre 2. Dans la première section, les caractéristiques du site d'étude sont présentées. Sont ensuite exposées les données disponibles sur cette zone et utilisées durant cette thèse. Les données collectées sur le terrain sont présentées dans une troisième section. Pour finir, ce chapitre se termine par l'introduction des données de télédétection utilisées pour la méthode de détection des pertes forestières.

1. Site d'étude – La Guyane

Dans le chapitre 2, nous avons vu que la Guyane est un site d'étude présentant de nombreux avantages pour étudier les méthodes de suivi des déboisements. Cette partie la présente et détaille les causes et les pratiques de déboisements sur ce territoire.

1.1. LA GUYANE

La Guyane est un territoire français situé au nord de l'Amérique du Sud, près de l'équateur (entre 2°2 et 5°6 N, et entre 51°6 et 54°4 O). Elle s'étend sur environ 83 530 km², et est couverte à environ 95% de forêts primaires (Keenan et al., 2015). L'altitude en Guyane varie entre 0 et 850 mètres et 12% des pentes sont supérieures à 15°.

Le climat de forêt tropicale humide de la Guyane est caractérisé par une saison sèche, allant de juillet à novembre (100 à 180 mm de précipitations par mois), et une saison humide plus longue, allant de décembre à juin (précipitations de 250 à 550 mm par mois). Cette saison humide est

nuancée par le petit été de mars : une courte saison sèche qui peut s'étendre sur un mois et demi, avec des précipitations moins importantes que pendant la saison des pluies (170 à 370 mm par mois). La couverture nuageuse annuelle moyenne est de 67% (https://earthobservatory.nasa.gov/global-maps/MODAL2_M_CLD_FR).

1.2. LES PERTES FORESTIERES EN GUYANE

Les déboisements en Guyane se font à petite échelle (Alvarez-Berríos and Aide, 2015) et sont de faible densité. Les principaux facteurs de déboisement en Guyane sont l'extraction de l'or (légale et illégale), l'agriculture de subsistance, l'urbanisation et l'exploitation forestière.

1.2.1. L'AGRICULTURE

L'agriculture en Guyane est principalement vivrière. Elle est pratiquée par des petits exploitants qui recourent à l'abattis-brulis (Figure 3, Figure 4), caractérisé par la coupe et le brûlage de petites surfaces (de 0.2 à 1 ha) entre juillet et novembre, pendant la saison sèche. Chaque parcelle est exploitée (et donc coupée et brûlée avant la plantation) pendant 2 à 3 ans avant une période de jachère. Les abattis se trouvent près des habitations, des villages et des fleuves. Les exploitations se situent près des pistes et des voies de communication pour pouvoir vendre la production plus facilement. De nos jours, la politique française encourage le développement agricole à des fins lucratives (Demaze and Manusset, 2008). Ceci, combiné à la croissance démographique et à la sédentarisation de la population, entraine une augmentation de la demande de terres agricoles, et donc une perte plus importante de la surface forestière. Néanmoins, les coupes sont encore principalement dues à une agriculture vivrière et n'entraînent pas d'effets environnementaux significatifs.

Figure 3 - Photo d'une parcelle agricole obtenue par : abattage des arbres (à gauche). Ils sont encore au sol et la végétation se dessèche ; après brûlis (à droite), il ne reste que des troncs au sol.

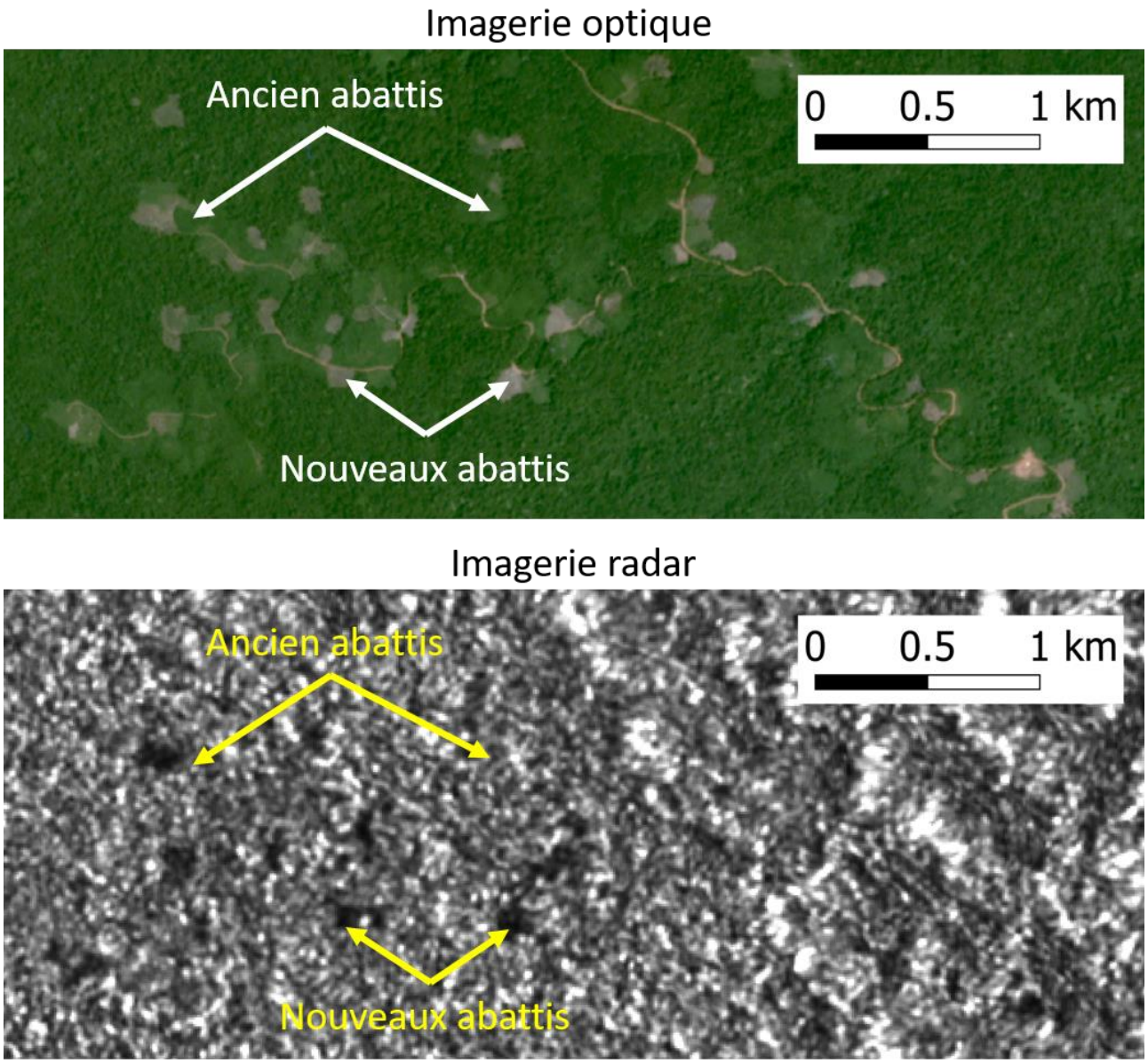


Figure 4 - Vue satellitaire optique et radar d'abattis (3°7 N, 54°0 O). L'image optique est une mosaïque biannuelle d'images Planet Juin 2019 (www.planet.com/nicfi). L'image radar est une image Sentinel-1 en polarisation VH du 11/10/2019 et sa dynamique est réglée entre -30 (noir) et -7 dB (blanc).

1.2.2. L'EXPLOITATION FORESTIÈRE

La gestion des forêts pour l'exploitation du bois est réglementée par l'Office National des Forêts (ONF) (Figure 5, Figure 6). Des pratiques d'exploitation à impact réduit sont mises en œuvre : le prélèvement de seulement 25 m^3 de bois (4 à 5 arbres) par hectare est autorisé, et les parcelles ne peuvent être exploitées qu'une fois tous les 65 ans. L'abattage des arbres et la création de pistes de débardage ont lieu toute l'année et représentent des modifications très faibles du couvert forestier. La construction de routes a plus d'impact, suivant le type de route (axes routiers principaux ou routes secondaires). Pour stocker le bois, des débarcadères (de taille allant de 0.1 à 0.5 ha) et des parcs à grumes (jusqu'à 1 ha) sont créés au bord de ces routes. Les routes, les parcs à grumes et les débarcadères sont généralement créés pendant la saison sèche (de juillet à décembre), bien que certaines étapes de la création des débarcadères aient lieu pendant la saison des pluies (comme le défrichement du sous-bois et parfois la coupe d'arbres).

Figure 5 - Photo de routes (à gauche) et de parcs à grumes (à droite) construits pour l'exploitation forestière. Dans les deux cas, la forêt a été remplacée par un sol nu et il n'y a rien au sol à part des grumes. Le sol est lisse dans la plupart des cas mais peut être rendu rugueux avec l'activité des machines. © ONF

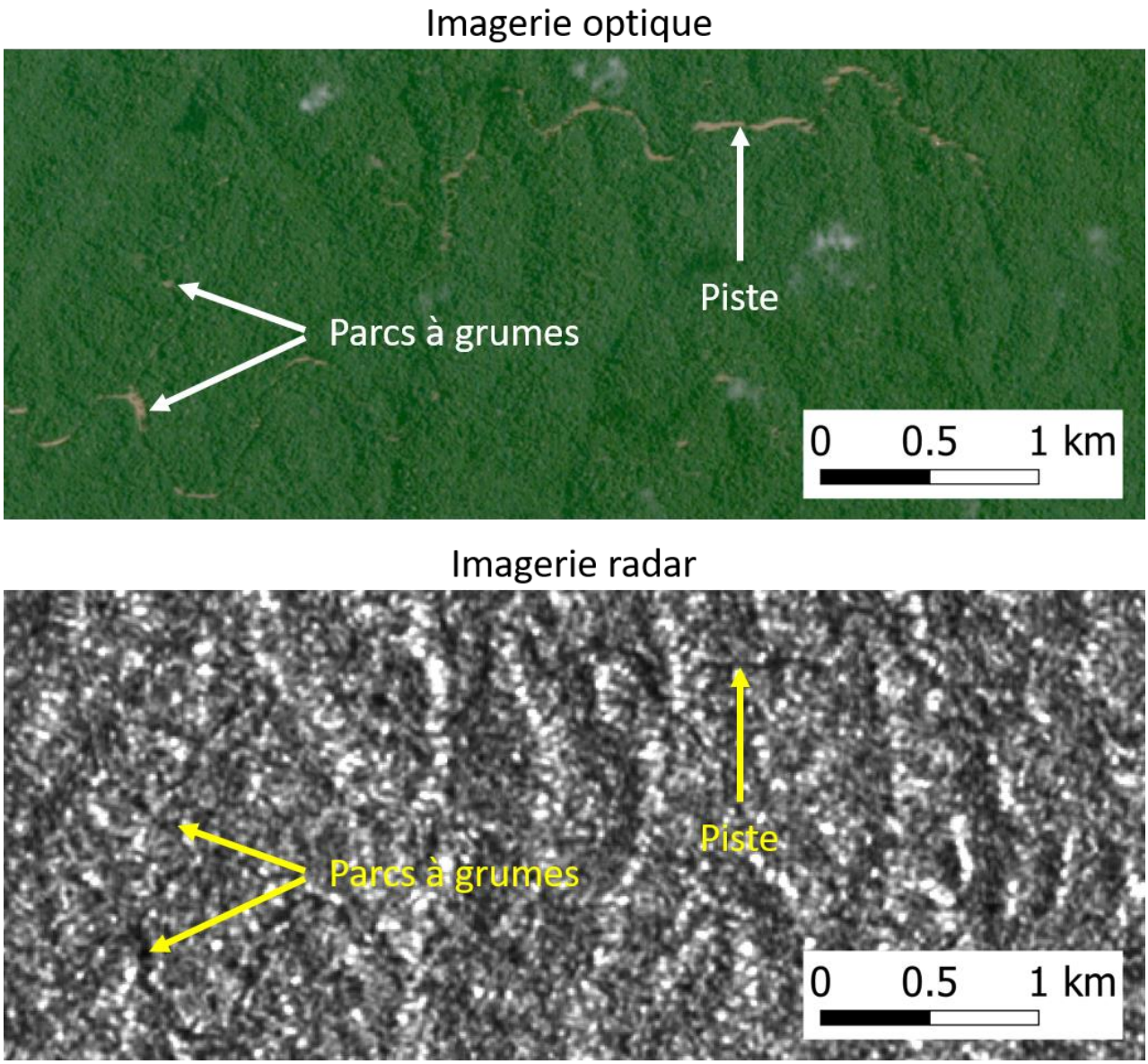


Figure 6 – Vue satellitaire optique et radar de pistes et parcs à grumes construits pour l'exploitation forestière (4°0 N, 51°8 O). L'image optique est une mosaïque biannuelle d'images Planet Juin 2020 (www.planet.com/nicfi). L'image radar est une image Sentinel-1 en polarisation VH du 13/08/2020 et sa dynamique est réglée entre -30 (noir) et -7 dB (blanc).

1.2.3. L'ORPAILLAGE

L'exploitation de l'or est la principale cause de déforestation en Guyane. Une étude (Rahm et al., others, 2015) a rapporté que plus de 25 000 ha de forêt avaient été touchés jusqu'en 2015.
La majorité des mines d'or (Figure 7, Figure 8) en Guyane sont exploitées tout au long de l'année. Elles sont principalement alluvionnaires (récoltes des paillettes d'or arrivées par exemple dans les cours d'eau) mais peuvent également être primaires (recherche de l'or dans sa forme primaire à même la roche). Les exploitations minières légales s'étendent souvent sur de vastes surfaces en eau. La loi française oblige les entreprises à restaurer la zone exploitée et interdit l'utilisation du mercure. La frontière entre l'exploitation légale et illégale est ténue, et des pratiques illégales sont parfois observées au sein de projets légaux.

Les opérations illégales sont généralement de plus petite envergure et peuvent donc être situées dans des vallées plus escarpées. Elles ont un impact environnemental fort et entraînent des problèmes sanitaires (Miller et al., 2003 ; De Kom et al., 1998 ; Eisler, 2004). Dans les opérations illégales, le mercure et le cyanure sont utilisés pour extraire l'or, puis sont rejetés dans les rivières, empoisonnant la faune et les populations locales. De plus, la présence d'orpailleurs à proximité de ces populations amène de l'insécurité, la perte de culture indigène, et des conflits sociaux (WWF, 2018 ; Forte, 1999; Colchester, 1997). L'augmentation des sédiments rejetés dans les rivières entraîne une perte de biodiversité. Sur le plan sanitaire, les mauvaises conditions de vie dans les camps miniers clandestins contribuent à la propagation de maladies comme l'infection au VIH et la malaria (Palmer et al., 2002). La détection de l'exploitation illégale de l'or revêt donc une importance environnementale et sociale pour la Guyane.

Figure 7 - Photographie aérienne des opérations légales d'extraction d'or : la zone est vaste et composée de bassins d'eau et de sol nu (à gauche). Photographie aérienne des opérations d'exploitation aurifère illégale : seule une petite surface de forêt a été défrichée (à droite). © ONF

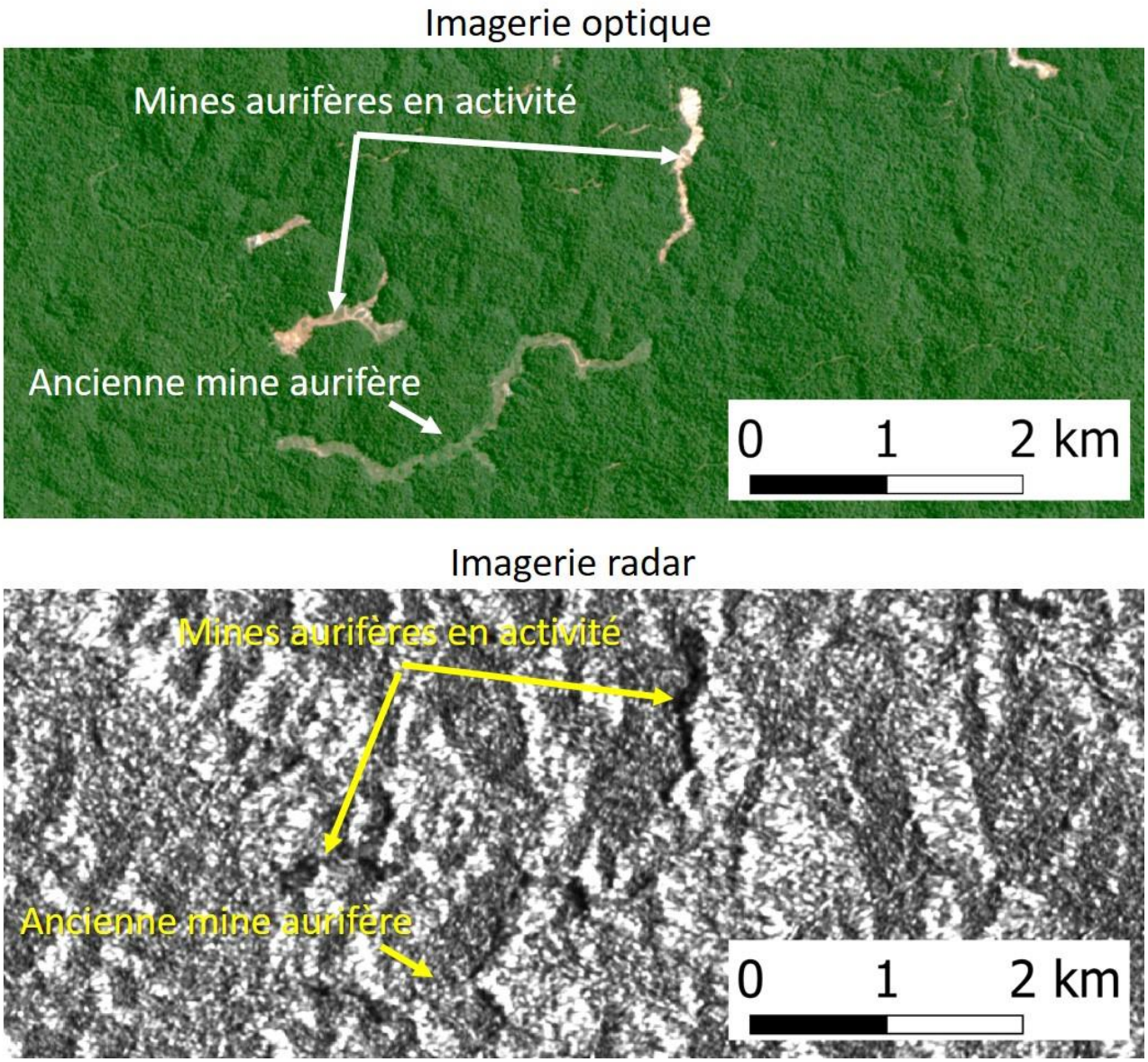


Figure 8 - Vue satellitaire optique et radar de mines aurifères légales. L'image optique est une mosaïque biannuelle d'images Juin 2020 (www.planet.com/nicfi). L'image radar est une image Sentinel-1 en polarisation VH du 13/08/2020 et sa dynamique est réglée entre -30 (noir) et -7 dB (blanc).

1.2.4. L'URBANISATION

Les villes et villages de Guyane se situent principalement sur le littoral ou aux frontières. La forte croissance démographique entraîne une demande en logement de plus en plus importante et le développement de l'urbanisation, notamment illégale (Figure 9). La densité des habitations est très variable, selon leur localisation en milieu urbain ou en milieu rural. Entre 2001 et 2015, les constructions spontanées (édifiées sans permis de construire) recensées ont augmenté en moyenne de 5.4% par an, et représentaient 41% des constructions en 2015 (OBS Habitat, 2015). 11% du bâti spontané se situe en zone à risques et 31% en zone agricole.

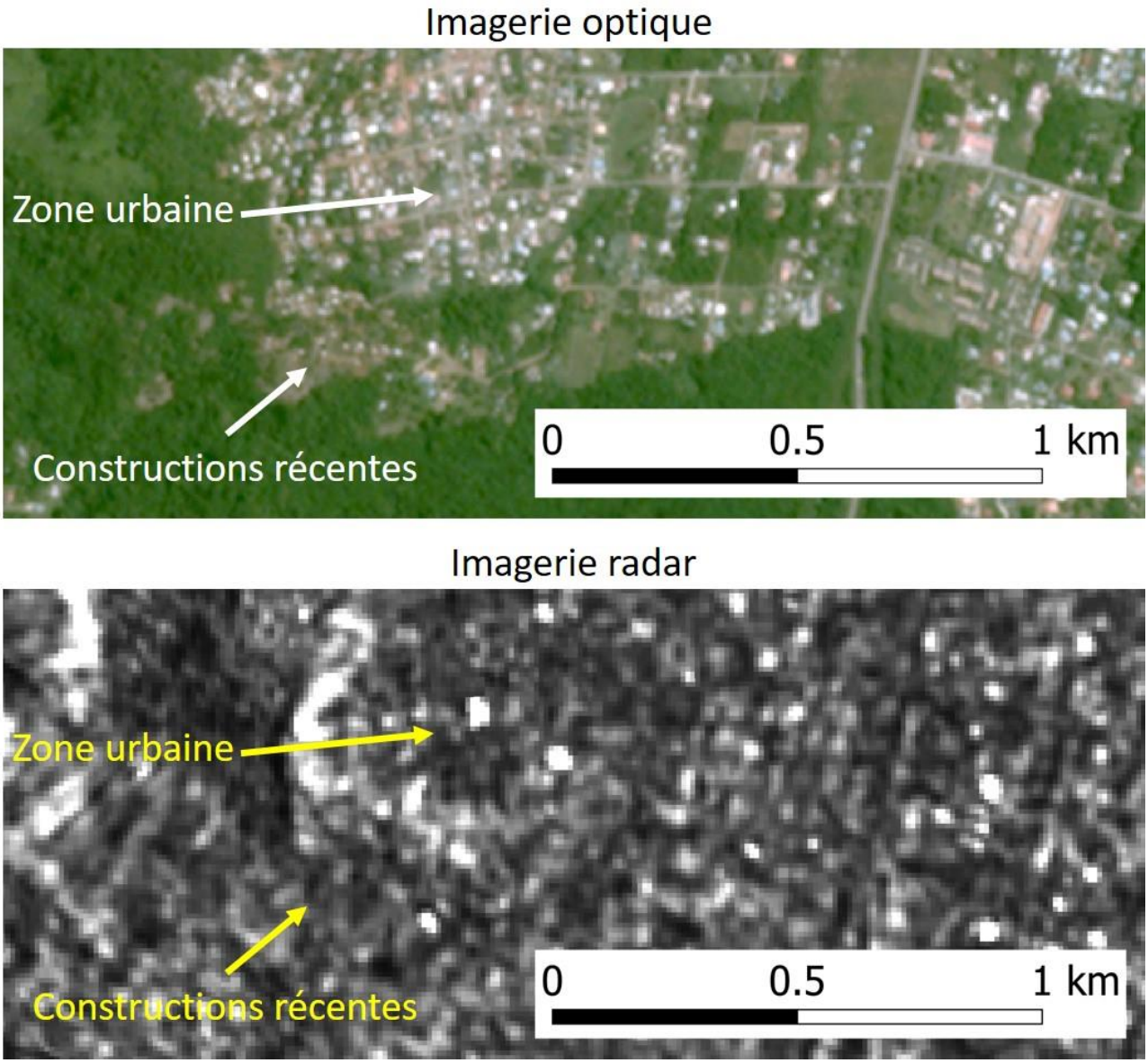


Figure 9 - Vue satellitaire optique et radar d'une zone urbaine et de constructions illicites récentes (4°9 N, 52°3 O). L'image optique est une mosaïque biannuelle d'images Planet Décembre 2019 (www.planet.com/nicfi). L'image radar est une image Sentinel-1 en polarisation VH du 22/01/2020 et sa dynamique est réglée entre -30 (noir) et -7 dB (blanc).

2. Données cartographiques

En Guyane, les données cartographiques d'occupation du sol disponibles sont nombreuses. Cette partie présente de manière globale les ensembles de données cartographiques ayant servi dans cette thèse. Les données utilisées parmi ces ensembles seront définies précisément par la suite en fonction de leur utilisation.

Cette partie se décompose en trois sections : les données de référence, les données auxiliaires et les données terrain acquises dans le cadre de cette thèse. Les données de référence ont permis d'évaluer les cartes de déboisement et de les labéliser avec leurs causes pour obtenir un jeu de données d'apprentissage. Les indicateurs géographiques utilisés pour la caractérisation des déboisements ont été dérivés des données auxiliaires. Certaines d'entre elles ont été utilisées comme données de comparaison. Pour finir, les données terrain ont servi d'exemple pour interpréter le signal radar et pour évaluer le délai de détection de la méthode développée.

2.1. DONNEES DE REFERENCE

Les données présentées dans cette partie sont des données d'usage du sol. Elles indiquent des déboisements dues à plusieurs pratiques. Elles sont toutes issues d'organismes français et utilisées comme référence pour le suivi des forêts de Guyane.

2.1.1. DONNEES AGRICULTURE

Le Parc Amazonien de Guyane (PAG) fournit des observations annuelles (et leur historique) de l'occupation du sol sur son territoire sous forme de fichiers vecteurs disponibles via le portail GéoGuyane (www.geoguyane.fr). Ces données peuvent servir de référence car les nouveaux déboisements dus à l'agriculture peuvent être retrouvés en prenant une parcelle labellisée « abattis frais » à l'année n et « forêt » à l'année n-1. Les polygones sont détectés par photo-interprétation de données SPOT, Pléiades, Landsat ou Sentinel-2 et la plupart d'entre eux ont été vérifiés sur le terrain.

2.1.2. DONNEES EXPLOITATION FORESTIERE

Dans le cadre de cette thèse, l'ONF a mis à disposition des données de référence sur les routes nouvellement construites, les parcs à grumes et les débarcadères pour l'exploitation forestière. Les tracés des routes ont été acquis in situ puis numérisées. Elles sont représentées par des lignes dans un fichier vecteur, sur lesquelles une zone tampon de 20 mètres (taille réelle) a été

appliquée pour représenter leur surface. Les données de référence des parcs à grumes ont été localisées sur le terrain puis numérisées en s'appuyant sur des images optiques. Elles sont datées par la date de la dernière image satellitaire exploitable avant la perturbation, et par la date de l'image satellitaire où l'on observe pour la première fois la perturbation. Trente pour cent de ces données sont datées au mois près. Sur les routes, la date de référence est l'année de la perturbation. Ces données ne sont pas exhaustives.

2.1.3. DONNEES ORPAILLAGE

Les données sur les mines d'or sont collectées en continu par l'Observatoire de l'Activité Minière (OAM) et gérées par l'État-Major de l'Orpaillage et de la Pêche Illicite (EMOPI). L'ONF utilise la photo-interprétation d'images SPOT, Pléiades, Landsat ou Sentinel-2 pour produire une donnée spatialisée permettant de suivre l'orpaillage en Guyane (Linarès et al., 2008). Ces données, systématiquement vérifiées par des campagnes héliportées, servent à orienter les missions de lutte contre l'orpaillage clandestin, et à vérifier le respect du cadre donné aux mines légales. La date de référence associée à chaque parcelle d'orpaillage est la date de la première observation de la perturbation sur les images optiques. Ces données étant sensibles, cette thèse a pu en bénéficier à la suite d'un accord de confidentialité. Le jeu de données de référence présenté ci-dessus est suffisamment précis pour une utilisation opérationnelle car les experts analysent diverses sources d'images par photo-interprétation, y compris des images Pléiades à très haute résolution (jusqu'à 2.8 mètres), et parce que la plupart des polygones numérisés sont vérifiés sur le terrain.

Il est à noter que les jeux de données de référence ne contiennent pas de dates précises de coupe en raison de l'utilisation d'images optiques ou car celles-ci n'ont pas été rapportées.

2.2. DONNEES AUXILIAIRES

Les données décrites dans cette partie sont des données cartographiques renseignant sur la géographie ou l'occupation du sol de la Guyane et sont toutes disponibles librement.

2.2.1. BD TOPO DE L'IGN

Depuis 2021, les données publiques de l'IGN sont devenues libres et accessibles gratuitement via le site https://geoservices.ign.fr/. Parmi ces données, l'on retrouve la BD TOPO : une base de données au format vectoriel rassemblant de nombreuses données géographiques sur la France.

Au sein de cette base de données, les couches ayant été identifiées comme pouvant entrer en jeu dans la caractérisation des causes de déboisement en Guyane sont les suivantes :

- « BATIMENT » : bâtiments agricoles, commerciaux, indifférenciés, industriels, résidentiels, religieux et sportifs. Cette couche contient beaucoup de bâtiments indifférenciés. Il n'est pas possible d'en sortir une information fiable et exhaustive pour les bâtiments résidentiels, industriels et commerciaux. Seuls les bâtiments agricoles et les serres semblent être exploitables ;
- « COURS_D_EAU » : squelettes (lignes) des cours d'eau, classés par importance ;
- « PLAN_D_EAU » : emprise du barrage de Petit-Saut ;
- « SURFACE_HYDROGRAPHIQUE » : emprise des surfaces hydrographiques, naturelles ou artificielles, intermittentes ou persistantes ;
- « ZONE_D_HABITATION » : emprise spatiale des carbets, lieux-dits, quartiers ;
- « ZONE_DE_VEGETATION » : décrit la végétation au sol, souvent forêt ouverte ou fermée mais aussi rizière, verger ;
- « TRONCON_DE_ROUTE » : squelettes (lignes) des routes, chemins, sentiers.

Dans la BP TOPO chaque objet comporte une série de dates :

- la date d'apparition représente la date de création, de construction ou d'apparition de l'objet, ou la date la plus ancienne à laquelle on peut attester de sa présence sur le terrain ;
- la date de confirmation est la date la plus récente à laquelle on peut attester de la présence de l'objet sur le terrain. La date de confirmation peut être antérieure à la date de création de l'objet ;
- la date de création correspond à la date et l'heure à laquelle l'objet a été saisi pour la première fois dans la base de données de production de l'IGN. Elle n'a pas de rapport avec la date d'apparition réelle de l'objet sur le terrain.

La date d'apparition sera donc antérieure à la date de confirmation qui sera elle-même antérieure de la date de création.

2.2.2. CAMINO

Le gouvernement français met à disposition le cadastre minier sur toute la France, accessible librement depuis le site https://camino.beta.gouv.fr/. Ce site répertorie et fournit une information spatialisée sur les titres miniers en cours de demande et en cours d'exploitation.

2.2.3. BRGM

Des cartes géologiques sont mises à disposition par le BRGM (Bureau de Recherches Géologiques et Minières) librement via le site : https://infoterre.brgm.fr/. Ces cartes contiennent une information spatialisée des types de roches présentes en France. L'or présent en Guyane se trouve principalement sur la ceinture de roches vertes du Paramaca.

2.2.4. ECOSEO

Le projet ECOSEO (Observatoire des Services Ecosystémiques du plateau des Guyanes) a pour but de promouvoir une approche transnationale en vue d'une meilleure gestion des écosystèmes sur le Plateau des Guyanes (Guyane, Guyana, Suriname et Amapa). Il est mis en œuvre par l'ONFI sur financement du WWF-France et co-financement du Programme Interreg de Coopération Amazonie (PCIA) et de l'Office de l'Eau de Guyane. Les partenaires impliqués sont la Collectivité Territoriale et la Préfecture de Guyane, l'ONF Guyane, le Parc amazonien de Guyane, la Commission des Forêts du Guyana, le Secrétariat d'état à l'environnement de l'Amapa au Brésil (SEMA-AP), les Universités de Hanovre et de Guyane, le SBB (Suriname) et le WWF Guianas.

Dans le cadre de ce projet, une carte d'occupation des sols en 2015 a été produite à 30 mètres de résolution, en utilisant les données d'occupation du sol disponibles sur chaque territoire et en traitement des images Landsat (voir détail dans Rahm et al., 2020). Les classes contenues dans cette couche sont les suivantes : infrastructures, établissements humains, extractions minières, cultures herbacées, cultures ligneuses, agriculture de subsistance, prairies, forêts, mangroves, zones arbustives, sol nu, zones humides ouvertes, eaux intérieures.

2.2.5. GLAD-FA

GLAD-FA (Hansen et al., 2016) est un produit d'alerte de perte forestière généré chaque semaine. Ce produit est fourni par le laboratoire GLAD de l'université du Maryland (UMD). Il est gratuit, disponible à large échelle (de 30 degrés nord à 30 degrés sud, en plus de l'Extrême-Orient russe), et basé sur des images optiques Landsat d'une résolution de 30 mètres. Une alerte est déclenchée lorsqu'un pixel perd 50% de sa couverture végétale. Cette alerte est confirmée si au moins 2 des 4 produits Landsat exploitables et consécutifs montrent également une perte de canopée sur une période de 180 jours : ce qui revient à confirmer l'alerte entre 8 et 180 jours. L'alerte confirmée est annotée de la date de la première observation du déboisement. La couverture nuageuse persistante sur les zones tropicales constitue une limitation importante de ce système de détection rapide, car peu d'images sont exploitables.

GLAD-FA n'a pas pour but de fournir des estimations de surface précises, mais est destiné à compléter les produits de perte de couverture forestière mondiale annuelle.

3. Données de terrain acquises dans le cadre de cette thèse

Une campagne terrain a été organisée afin de compléter les données précédemment exposées car elles ne répondaient pas à tous les besoins de cette thèse. La campagne terrain et les données collectées sont présentées dans la section ci-dessous.

3.1. MOTIVATIONS

La validation d'une méthode de détection rapide doit être spatiale pour évaluer la performance de détection, mais aussi temporelle afin d'évaluer la capacité du système à agir vite. Si les données de référence des perturbations forestières (orpaillage, agriculture, exploitation forestière) sont nombreuses en Guyane, leur datation est peu précise. Souvent, seule l'année de coupe, la première observation faite par imagerie satellitaire, ou un intervalle de temps (dernière image sans coupe - première image avec coupe) est disponible. Ceci permet une bonne validation spatiale, mais pas temporelle. Cette validation temporelle, rarement abordée dans les études, est pourtant essentielle pour évaluer les algorithmes visant une détection rapide des changements, notamment dans des situations critiques comme la détection précoce de l'orpaillage. Pour répondre à ce manque, nous avons mis en place une campagne terrain permettant de suivre régulièrement et précisément les perturbations forestières, avec l'aide du Service de la Statistique et de la Prospective (SSP) du Ministère de l'Agriculture et de l'Alimentation et de la DAAF de Guyane. Celui-ci a réquisitionné un enquêteur pour mener à bien cette campagne durant la saison sèche de 2019 (août à novembre inclus).

Cette enquête de terrain a pour objectifs :

1) D'évaluer les résultats sur le plan temporel, c'est-à-dire d'estimer au mieux la date effective de la coupe. L'accès à une information terrain précise au niveau temporel est rare et cruciale dans la validation d'algorithmes de détection en quasi temps-réel ;

Mais aussi de :

2) Valider les résultats obtenus par la méthode d'un point de vue spatial, c'est-à-dire la détection effective des parcelles déboisées, et en particulier des parcelles de surface inférieure 1 ha ;

3) Mieux comprendre l'interprétation des images radar. La prise de photos sur place permet de connaître les types de déboisements effectifs et d'interpréter l'imagerie radar disponible ;
4) Comprendre comment la parcelle déboisée évolue dans le temps. Les revisites des parcelles déboisées permettent de caractériser l'évolution de la parcelle.

Les parcelles suivies dans cette enquête sont des abattis fraichement déboisés. La finesse de l'échelle temporelle pour la détection de ces parcelles n'est pas critique, mais ces zones sont relativement accessibles et leur accès est peu dangereux en comparaison, par exemple, des zones d'orpaillage. De plus, les abattis se trouvant près des villages et le long de pistes, il est plus facile de connaître leur historique (par le visuel). Pour finir, nous imaginons que la population est plus encline que des orpailleurs clandestins (par exemple) à répondre aux questions d'un enquêteur pour révéler le plus précisément possible la date de coupe.

3.2. MODE OPERATOIRE – DONNEES COLLECTEES

Le **principe de l'enquête** est le suivant : chaque parcelle détectée par télédétection est directement transmise à un enquêteur pour que celui-ci aille sur le terrain dans un délai raisonnable, afin de définir à quel moment la coupe a eu lieu (le plus précisément possible). L'enquêteur réalise de 5 à 10 observations/jour (selon l'éloignement et la praticabilité des pistes) et est disponible 2 jours/mois, ce qui permet d'observer de 10 à 20 zones mensuellement.

Il a également été souhaité qu'un déboisement récent détecté par l'enquêteur soit identifié et intégré au jeu de données. Cela permet de voir si ce déboisement est détecté ou non (quel qu'en soit l'usage in fine). Ce fonctionnement en double sens est important car il permet de s'assurer de l'intégralité des possibilités de détection.

Le **protocole** des visites sur le terrain a évolué tout au long de la campagne, et a abouti à un protocole final efficace pour la personne sur le terrain et pour celle qui analyse les résultats. Tous les mois, entre mi-août et mi-décembre 2019, l'enquêteur a reçu un fichier *xml* contenant des parcelles à visiter et un fond de carte à mettre dans l'application *Oruxmaps* gratuite au moment de la campagne terrain (https://www.oruxmaps.com/). Le fond de carte était composé des routes, d'un zonage défini par l'enquêteur, de parcelles d'abattis détectées les années antérieures et a été produit à partir de *Qgis* (Figure 10).

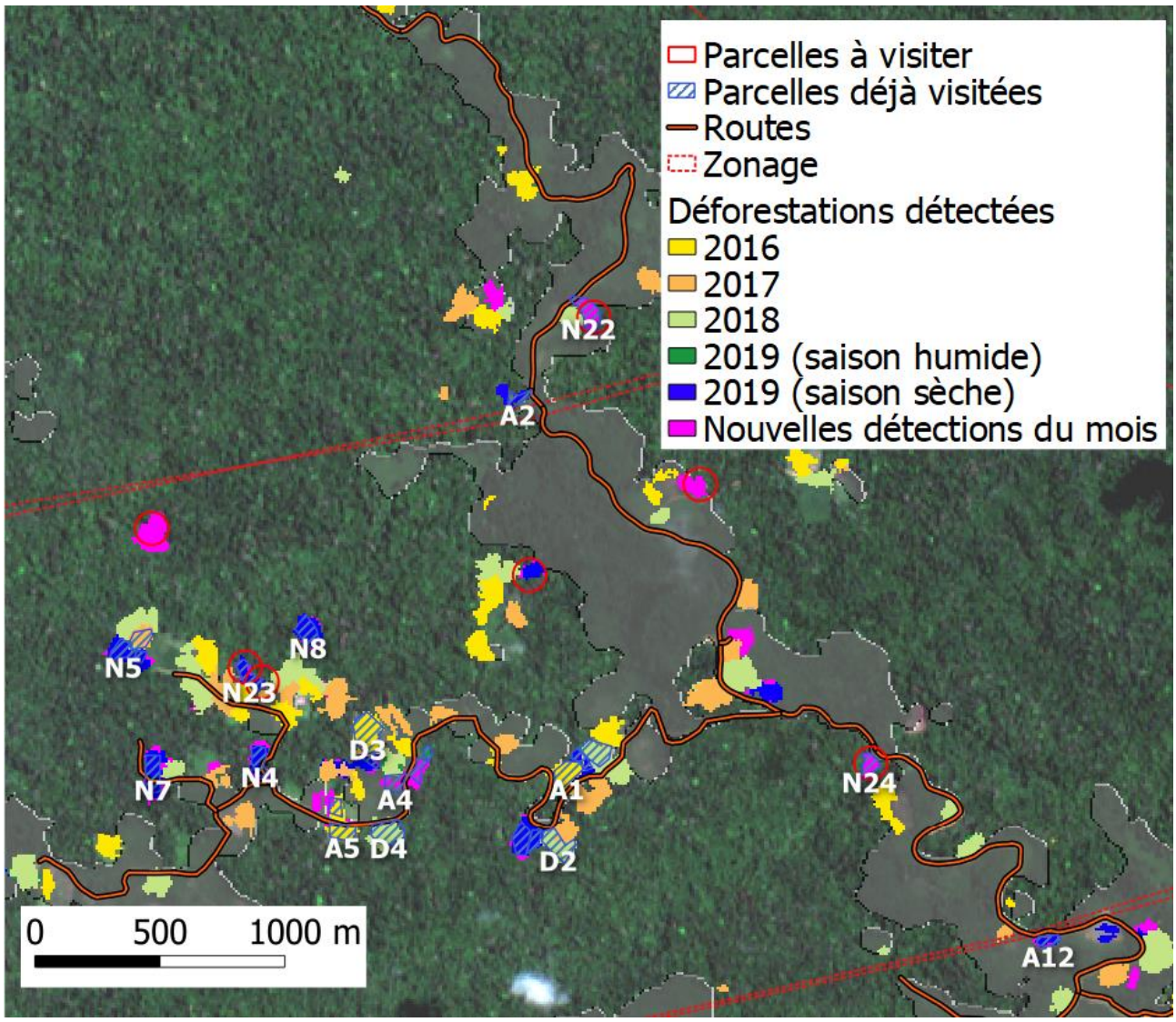


Figure 10 - Exemple de fond de carte fourni à l'enquêteur à mettre dans l'application OruxMaps. Les parcelles à visiter sont encerclées en rouge et les parcelles déjà visitées ont des hachures bleues. Les routes sont représentées en orange et les déboisements détectés depuis 2016 sont les parcelles en pleine couleur. Le fond de carte est une image Sentinel-2.

Ces informations ont largement contribué à faciliter les visites, permettant à l'enquêteur de mieux se repérer qu'avec le seul fond de carte Open Street Maps (https://www.openstreetmap.fr/). Avant de partir, l'enquêteur suit une procédure pour mettre le fond de carte dans *OruxMaps*, ainsi que les points à visiter. Il devait aussi s'assurer que l'option de géolocalisation des photos était activée dans son appareil. Pendant la campagne, pour chaque parcelle, l'enquêteur enregistrait un jalon dans *OruxMaps*, avec un identifiant qu'il définissait. Ces jalons sont récupérables directement dans *Qgis* sous forme de points géolocalisés.

Figure 11 - Photos d'enquêteurs pendant la campagne terrain

Lors de la visite de ces parcelles (Figure 11), les points à noter sont les suivants :

- Identifiant de la parcelle (associé au jalon) ;
- Date de la visite de la parcelle ;
- État de la végétation au sol : Sol nu/brûlé/végétation restante au sol + commentaire/description/remarque ;
- Estimation de la taille de la parcelle à l'œil ;
- Estimation (si possible) de la date de coupe de la parcelle ;
- Nature de la parcelle avant la coupe (forêt, recrue, etc.) ;
- Photos géolocalisées de la parcelle ;
- Zone (définie par le zonage : PK10, New Wacapou, etc.).

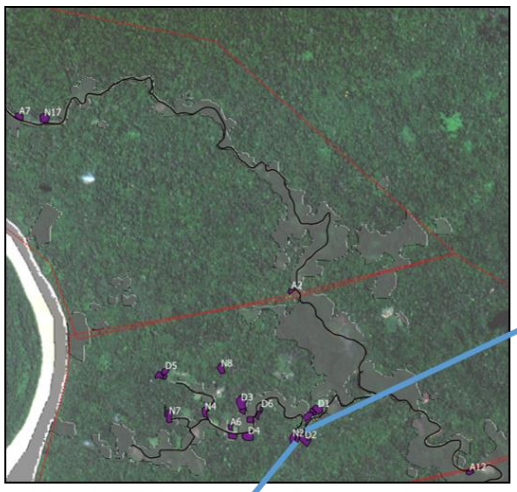
Ces informations sont ensuite transmises et analysées dans le cadre de cette thèse. Chaque parcelle a été nommée, spatialisée et associée à un fichier Excel regroupant toutes les informations recueillies (Figure 12). La liste des parcelles visitées a été complétée par les coupes détectées sur le terrain.

3.3. ANALYSE A POSTERIORI DES DONNEES

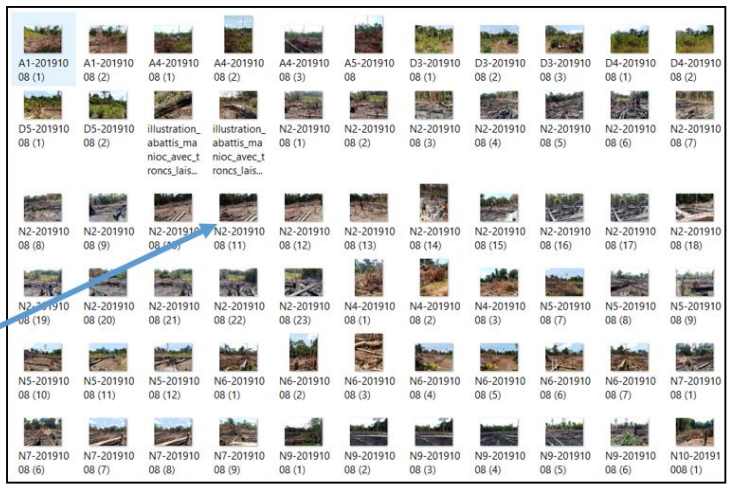
Plusieurs difficultés ont été rencontrées pendant cette campagne terrain. D'abord, il y avait une ambiguïté entre les premières parcelles à visiter et les parcelles effectivement visitées. Au début des enquêtes, seules les coordonnées GPS des parcelles et un fond de carte Open Street Map étaient fournis. Cela ne permettait pas à l'enquêteur de se repérer correctement et d'être sûr qu'il était sur la bonne parcelle. Il a aussi été difficile d'estimer la taille de la parcelle à l'œil nu, ce qui rend cette information peu fiable. L'estimation de la date de coupe s'est également révélée compliquée : dans la plupart des cas, les agriculteurs n'étaient pas présents sur les parcelles, ou ne savait pas dire avec précision quand ils avaient coupé leur parcelle. Pour contourner cette difficulté et estimer au mieux la date de coupe de chaque parcelle, un intervalle de coupe a été défini avec des images Sentinel-2. Dans cet intervalle, une date de coupe a été déduite en fonction de l'état de séchage ou la présence de brûlis sur la végétation de la parcelle à la date de la visite, en se basant sur les règles suivantes :

- Feuilles vertes et masse végétale haute : < 7 jours ;
- Changement de couleurs des feuilles au sol + baisse de la masse végétale : 20 jours ;
- Feuilles sèches : 35 jours ;
- Brûlis en cours ou parcelle fumante : 40 – 60 jours ;
- Re-brûlis : 70 jours ;
- Mise en culture : 80 jours.

Ces règles ont été définies à partir de ces observations et des revisites des parcelles, et permettent ainsi d'estimer au plus près les dates effectives de coupe des parcelles.



Carte des parcelles visitées géolocalisées



Images associées aux parcelles par date de visite

Point	Visite de la parcelle	État de la végétation au sol	Estimation de la taille de la parcelle	Estimation de la date de déforestation	Nature de la parcelle avant	Vérification	Zone
N1	08/10/2019	Présence de gros troncs coupés (en plus grande quantité que A1) au sol. Moins de cecropias que sur la parcelle A1 collée. Parcelle brûlée mais mal brûlée. Ça a mal brûlé il y a toujours plein de bois au sol. Un peu de masse végétal au sol.		Difficile d'estimer mais suffisamment longtemps pour avoir laissé le temps à la végétation de sécher pour faire le brûlis	Forêt ancienne qui était entre deux abattis et qui a été coupé cette année	Confirmé par le PAG : forêt ancienne jsuqu'à 2019	PK10
N2	19/09/2019	Parcelle brûlée, encore fumante, reste bois carbonisé au sol, le reste de la matière est en cendre.	1ha	Emmanuelle connait l'agricultrice, abattis autour du 10/08/2019	Forêt ancienne	Confirmé par l'agricultrice	PK10
	08/10/2019	Bien brûlée. Recoupée à droite					
N3	08/10/2019	Brûlis en cours. Présence de gros troncs comme de palmiers et de petits troncs. Plus trop de masse végétale.		Extension de la parcelle N2	Forêt ancienne	Confirmé par l'agricultrice	PK10
N4	19/09/2019	Végétation restante au sol, pas encore de brûlis. Beaucoup de masse végétale.	1ha	Récente, la plupart des feuilles sont vertes.	Forêt ancienne, on voit que la plupart des troncs sont gros.	Confirmé par le PAG : forêt ancienne jsuqu'à 2019	PK10
	08/10/2019	Gros tronc. Toujours en train de sécher. Cependant il y a toujours des feuilles en train de changer de couleur. Le volume de la masse végétale a beaucoup diminué par rapport à la dernière visite, ça veut dire que ça a sécher. Toujours pas de brûlis.					
N5	19/09/2019	Végétation restante au sol, pas encore de brûlis	2ha	Difficile à estimer on voit que les **feuilles sont déjà sèches**	Forêt ancienne	Confirmé par le PAG et par Emmanuelle	PK10
	23/09/2019	Brûlis en cours					
	08/10/2019	Gros troncs encore fumants					
N6	08/10/2019	Végétation restante au sol, pas encore de brûlis.		Extension de la parcelle N5. Les feuilles sont en train de sécher. Il y a des feuilles qui sont en train de changer de couleur et celles des cecropias sont déjà sèches. Vient juste d'être coupé. L'agricultrice avait commencé à nettoyer le sous-bois le 23 septembre et pas encore coupé les gros arbres. Aujourd'hui les gros arbres sont coupés.	Forêt ancienne	Confirmé par le PAG et par Emmanuelle	PK10

Informations associées aux parcelles par date de visite

Figure 12 - Exemple d'informations récupérées à la suite de la campagne terrain : parcelles spatialisées et labélisées (à gauche), fichier Excel regroupant toutes les informations associées aux parcelles (en haut à droite), images prises sur le terrain (en bas à droite)

Au final, 57 parcelles ont été observées au cours de 70 visites (incluant les revisites) (Tableau 1). Ces parcelles ont été classées en 3 catégories : anciens abattis (16), nouveaux abattis (26) (anciennement forêt dense), parcelles diverses (15). Les anciens abattis sont des abattis fraîchement coupés qui étaient anciennement des plantations ou des jachères. Ils proviennent majoritairement des détections faites par l'enquêteur qui ne connaissait pas la nature de la parcelle avant la coupe. Cette nature a été déterminée par la suite à l'aide de couches d'occupation du sol. Les plantations diverses proviennent d'observations faites sur le terrain et auraient été potentiellement intéressantes dans le cas où la campagne aurait été reconduite l'année suivante.

Tableau 1 - Tableau récapitulatif du nombre de parcelles visitées au cours de la campagne terrain. A = ancien abattis, N = nouvel abattis, D = parcelles diverses, EL = erreur de localisation, NV = parcelle non visitée

Mois de visite	Parcelles détectées par imagerie satellitaire	Parcelles visitées	Parcelles non visitées ou erreur de localisation	Parcelles revisitées	Parcelles ajoutées par l'enquêteur
Août	6	1 A, 1 D	4 EL		
Septembre	9	8 N	1 NV	1N	5 N, 5 A
Octobre	7	7 N		6 A, 6 N	10 A, 14 D
Novembre	20	6 N	1 EL, 13 NV		
Total	42	21 N, 1A, 1 D	5 EL, 14 NV	7 N, 6 A	5 N, 15A, 14 D

4. Données RSO

Comme présenté dans le chapitre 2, la télédétection est un outil pertinent pour le suivi de vastes zones inaccessibles. Les données Radar à Synthèse d'Ouverture (RSO, SAR pour Synthetic Aperture Radar en anglais) présentent l'avantage d'être relativement insensibles à la couverture nuageuse, permettant d'avoir une information exploitable à chaque acquisition, ce qui est bien adapté au suivi quasi-temps réel des zones tropicales. Dans cette partie, nous abordons, dans un premier temps, le principe d'acquisition des RSO ainsi que les propriétés et les caractéristiques de leurs images. Puis, sont détaillées les caractéristiques et le traitement appliqués aux images Sentinel-1 utilisées dans cette thèse.

4.1. TELEDETECTION RSO

4.1.1. RADAR IMAGEUR

Le RADAR (RAdio Detection And Ranging) est un instrument de télédétection actif, basé sur l'émission et la réception d'une onde électromagnétique (Figure 13).

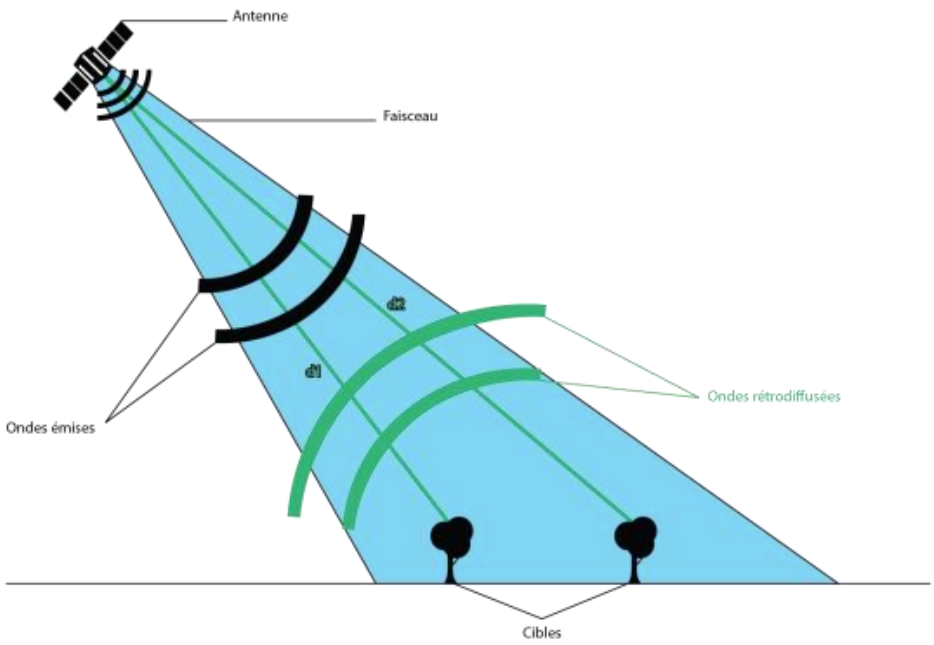


Figure 13 - Principe de fonctionnement du radar imageur : l'antenne émet un signal (émission d'ondes), le sol reçoit l'énergie transmise par le radar et renvoi vers l'antenne des ondes en retour. Auteur : Lucas Ternynck - icône "satellite" par Jose Luis Algara - icône "arbre" par James Keuning

Le système radar va générer une onde électromagnétique (Figure 14) à une fréquence f (correspondant à une longueur d'onde $\lambda =$ c/f, c étant la vitesse de propagation l'onde) sous la forme d'une impulsion de durée τ_p, transmise latéralement par une antenne de gain G avec une fréquence de répétition des impulsions PRF. Cette onde va se propager dans l'espace jusqu'à la cible observée, sur une distance notée R. Les cibles associées sont localisées à l'aide du temps de retour de l'onde, duquel on déduit la distance R $= (c \times t)/2$ car la vitesse de propagation c de l'onde est connue (dans l'air ou le vide).

La cible, en fonction de ses caractéristiques propres (structure, matériaux, humidité, etc.), va rétrodiffuser une partie de l'énergie de l'onde vers le radar. Cette partie de l'onde, va être captée par l'antenne de réception et numérisée par le système radar.

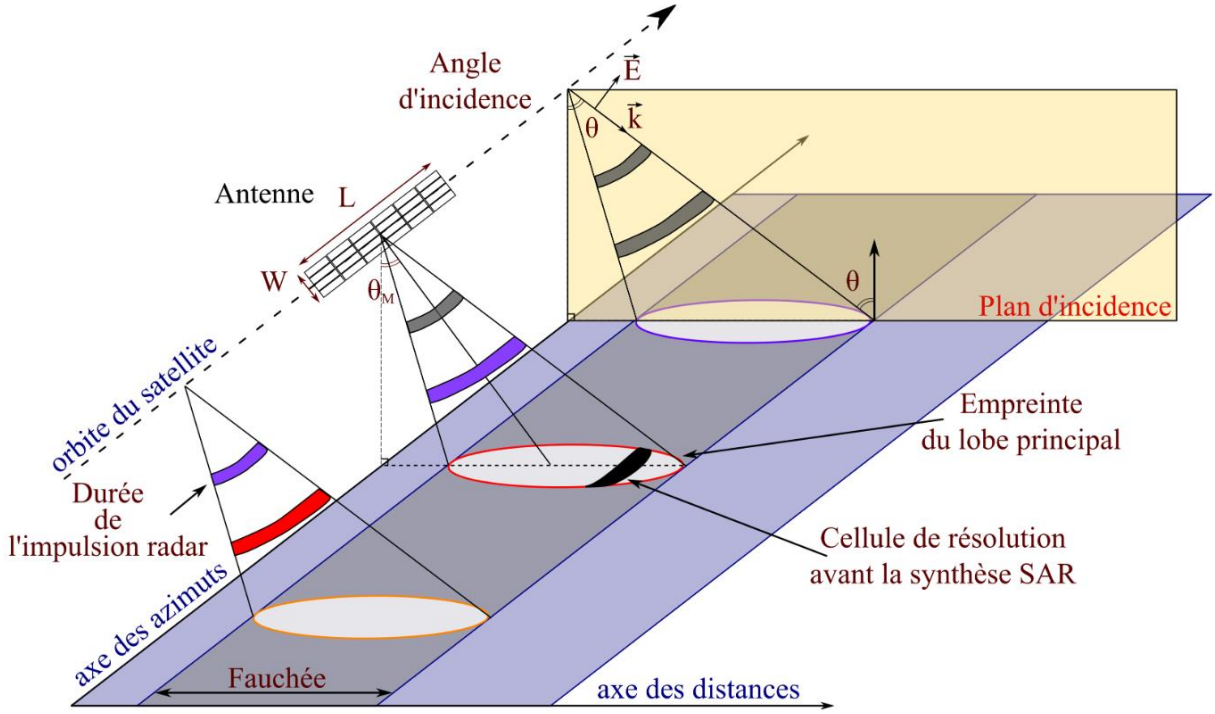


Figure 14 - Géométrie d'acquisition d'une radar. Source (Lardeux, 2008)

Dans une configuration monostatique, caractérisée par une antenne d'émission et de réception identiques, les échos rétrodiffusés par les objets éclairés interagissant avec l'onde incidente sont classés par ordre d'arrivée. L'échantillonnage de ces échos définira les colonnes de l'image (sur la direction dite « en distance »), tandis que les lignes seront formées par les différents échos reçus des impulsions envoyées au cours du déplacement du satellite (axe des azimuts). La Figure 14 montre 3 impulsions radar le long de la direction de déplacement du capteur. Chacune est espacée par la distance $P_a = V/PRF$ (voir équation 3.9). Chaque impulsion est représentée 2 fois correspondant à des temps de propagation différents.

L'onde électromagnétique reçue par le capteur est caractérisée par son amplitude A, sa pulsation $\omega = 2\pi$f et sa phase φ tel que :

$$\psi(r,t) = A\ \cos(\omega t + \ \varphi) \tag{3.1}$$

Avec l'amplitude A de l'onde, ou plutôt son intensité I = A² (assimilable à une puissance) permet de caractériser le coefficient de rétrodiffusion radar σ^0 (m²/m²), donnant des informations sur les caractéristiques du milieu observé.

L'équation radar (simplifiée) relie les signaux, les paramètres du système et les caractéristiques de la cible :

$$P_r = \frac{P_t G^2 \lambda^2}{(4\pi)^3 R^4} \sigma^0 S \Leftrightarrow \sigma^0 = \frac{P_r}{P_t} \frac{(4\pi)^3 R^4}{G^2 \lambda^2 S} \quad (3.2)$$

Avec P_r (watts) la puissance reçue par le système, P_t (watts) la puissance transmise, S la surface effective vue par l'antenne, λ la longueur d'onde utilisée (mètres) et G le gain de l'antenne. Le coefficient de rétrodiffusion σ^0 traduit les propriétés de rétrodiffusion de la surface. Elles sont déterminées à la fois par ses caractéristiques physiques (couverture, topographie, rugosité, humidité), et par les paramètres du radar (angle d'incidence, longueur d'onde, polarisation). Il est l'analogue à la réflectance bidirectionnelle en optique (dans une configuration de visée de hotspot). En raison de sa grande dynamique, σ^0 est habituellement exprimé en dB, donné par :

$$\sigma^0(dB) = 10log_{10}\left(\sigma^0(m^2m^{-2})\right) \quad 3.3$$

Le principe de fonctionnement du radar, décrit très simplement ici, est plus détaillé dans la vaste littérature disponible sur le sujet (Ulaby et al., 1982) (Massonnet and Souyris, 2008).

4.1.2. RADAR A SYNTHESE D'OUVERTURE (RSO)

Pour être exploitables, des données radar doivent avoir une résolution suffisante et compatible avec les applications visées. Sans traitement complémentaire, un système radar spatial ne peut pas descendre à des résolutions inférieures à quelques kilomètres, ce qui le rend totalement inutile. Cependant, la technique de la synthèse d'ouverture (RSO) permet d'améliorer considérablement sa résolution, pouvant ainsi atteindre des valeurs métriques.

Initialement, la résolution du radar dépend de la taille de son antenne. Ainsi, plus l'antenne sera grande, meilleure sera la résolution (voir paragraphe 4.2.1). Par un calcul numérique du premier ordre, on montre que pour atteindre une résolution d'un mètre avec un radar spatial, on a besoin d'une antenne de plusieurs kilomètres de long. Or, un satellite a des contraintes d'encombrement et de masse et ne peut physiquement pas supporter une telle antenne.

Le concept du radar à synthèse d'ouverture (RSO) permet de simuler une très grande antenne, afin d'atteindre les résolutions souhaitées, en utilisant des techniques de traitement du signal ad-hoc. Pour se faire, on utilise le déplacement du radar le long de sa trajectoire, ce qui permet de mesurer la rétrodiffusion de la cible depuis plusieurs points de vue (typiquement quelques milliers de mesures). Puis, en combinant toutes ces mesures (à la fois l'information d'amplitude et de phase de la rétrodiffusion) dans le processus du traitement RSO, on obtient la réponse du radar pour cette antenne « synthétique » de très grande taille (durée de l'illumination de la cible le long de la trajectoire).

4.2. LES IMAGES RSO

Les mesures de la rétrodiffusion des cibles (ou plus généralement de la surface observée par le radar) peuvent être représentées sous la forme d'une image, selon les deux axes d'azimut et de distance. Chaque mesure synthétique (c'est-à-dire après le traitement RSO) est alors représentée sur la forme d'un pixel dont la valeur correspond à la valeur du coefficient de rétrodiffusion σ^0. Les images RSO ont plusieurs caractéristiques qui les rendent uniques :

- Elles fournissent des images hautes résolutions disponibles de jour comme de nuit (instrument actif) et quelle que soit la couverture nuageuse (micro-ondes).
- La valeur du pixel informe sur les caractéristiques physiques de la cible, comme par exemple la structure, la rugosité et l'humidité.

Dans cette partie, nous verrons la résolution spatiale des images RSO, la taille de leurs pixels, ainsi que leurs propriétés géométriques et radiométriques.

4.2.1. RESOLUTION SPATIALE

Les systèmes RSO collectent des échos dans des fauchées latérales pour former une dimension de l'image (en distance) et font un échantillonnage le long de la trajectoire (en azimut) pour former la deuxième dimension de l'image. Par définition, la résolution de l'image radar est la plus petite distance séparant deux objets pouvant être distingués. Les RSO effectuant des traitements distincts sur les 2 axes, on définit une résolution pour chaque axe de l'image.

La résolution en distance ($\mathrm{R_d}$) correspond basiquement à la largeur de l'impulsion radar τ_p, inversement proportionnelle à sa bande passante B (Maître et al., 2001) :

$$\mathrm{R_d} = \frac{\mathrm{c}}{2\mathrm{B}} \qquad 3.4$$

Avec c la vitesse de propagation de l'onde. Le facteur 2 provient de l'aller-retour du signal au capteur. Ce résultat provient de l'utilisation de la technique de compression d'impulsion, qui permet de générer des impulsions de grande bande passante tout en maintenant une durée τ_p suffisamment longue pour garder une puissance d'impulsion adaptée. $\mathrm{R_d}$ est la résolution du système en géométrie radar (Figure 15).

La résolution effective de l'image au sol est donnée par la projection de cette résolution $\mathrm{R_d}$ sur le sol :

$$\mathrm{R_{d,s}} = \frac{\mathrm{R_d}}{\sin\theta} = \frac{\mathrm{c}}{2\mathrm{B}\sin\theta} \qquad 3.5$$

Avec θ l'angle de visée du radar (ou angle d'incidence), définit comme l'angle formé par la direction de visée du radar et la normale à la surface observée (Figure 13).

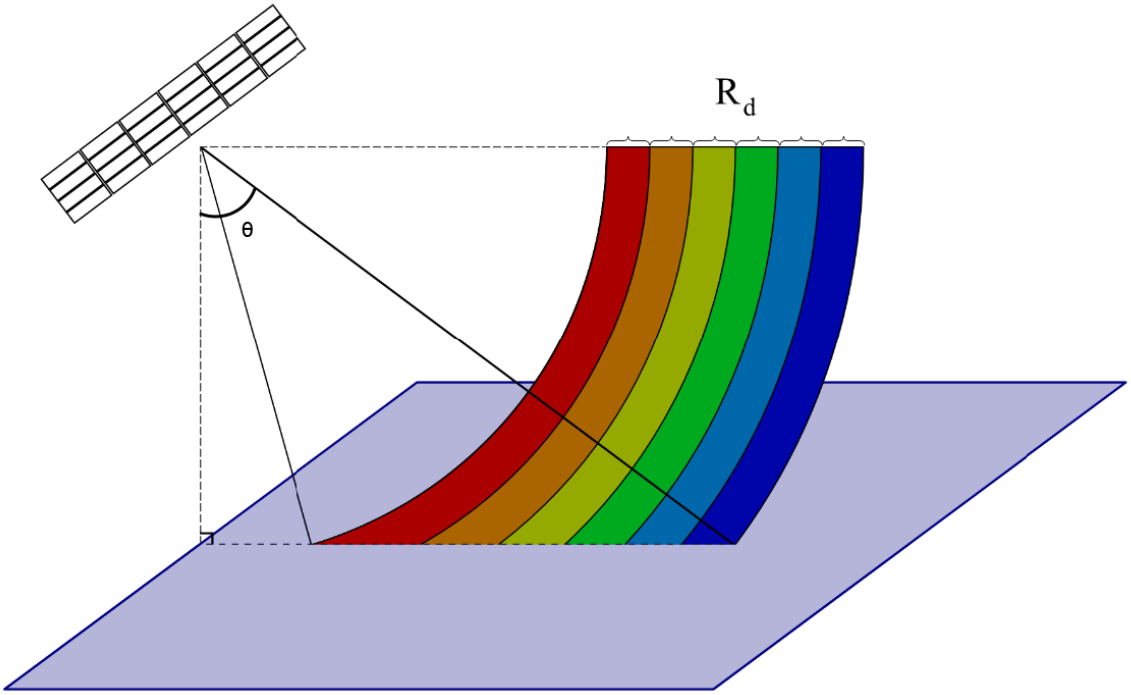


Figure 15 - Visée en distance d'un radar. Source (Lardeux, 2008).

Dans la direction azimutale, on peut montrer qu'au premier ordre, après le traitement RSO, la résolution peut être exprimée directement à partir de la largeur L de l'antenne (mètres) :

$$\mathrm{R_a} = \frac{L}{2} \qquad 3.6$$

Il s'agit bien d'une expression simplificatrice, car en réalité, les résolutions dépendent également de tous les paramètres du traitement RSO.

4.2.2. TAILLE DE PIXEL

En distance, la taille du pixel radar « brut » (à ne pas confondre avec la résolution spatiale du radar), est donnée par :

$$P_d = \frac{c}{2\,P_f} \qquad 3.7$$

Avec P_f la fréquence d'échantillonnage de l'écho rétrodiffusé. Cet échantillonnage régulier étant fait sur l'axe distance (géométrie radar), la surface de sol réelle observée dans le pixel varie avec l'angle d'incidence. Les pixels radar représentent donc des zones au sol plus grandes quand ils ont un angle d'incidence plus faible. En géométrie sol, on a :

$$P_{d,s} = \frac{P_d}{\sin\theta} = \frac{c}{2\,P_f \sin\theta} \qquad 3.8$$

En azimut, chaque impulsion émise constitue une ligne de l'image radar séparée de la suivante par la distance parcourue par le radar entre 2 impulsions (échantillonnage en azimut) :

$$P_a = \frac{V}{PRF} \qquad 3.9$$

Avec V la vitesse de déplacement du radar.

4.2.3. CARACTERISTIQUES DU COEFFICIENT DE RETRODIFFUSION

L'amplitude du signal renvoyé permet de calculer le coefficient de rétrodiffusion de la surface observée, noté σ^0. Celui-ci varie suivant plusieurs paramètres, comme par exemple la rugosité de la surface, son humidité ou encore la structure du couvert.

La rugosité d'une surface correspond la variation moyenne de ses aspérités par rapport à une surface plane. L'influence de la rugosité sur l'interaction onde/surface dépend de la longueur d'onde λ utilisée. Sur des surfaces "lisses" ($\lambda \gg$ rugosité), comme de l'eau plate, l'onde reçue est réfléchie de façon spéculaire (Figure 16). Le coefficient de rétrodiffusion sera donc faible. En revanche, sur une surface "rugueuse" ($\lambda \ll$ rugosité), comme un champ labouré, l'onde est réfléchie dans plusieurs directions et le signal reçu par l'antenne est plus important.

Pour certaines structures de la cible, un mécanisme de double-rebond peut avoir lieu. Il apparaît quand deux surfaces forment un dièdre (telles que des constructions en zones urbaines, ou des troncs d'arbres en lisière de forêt), renvoyant ainsi la majorité du signal dans la direction du satellite. On retrouve alors un fort coefficient de rétrodiffusion.

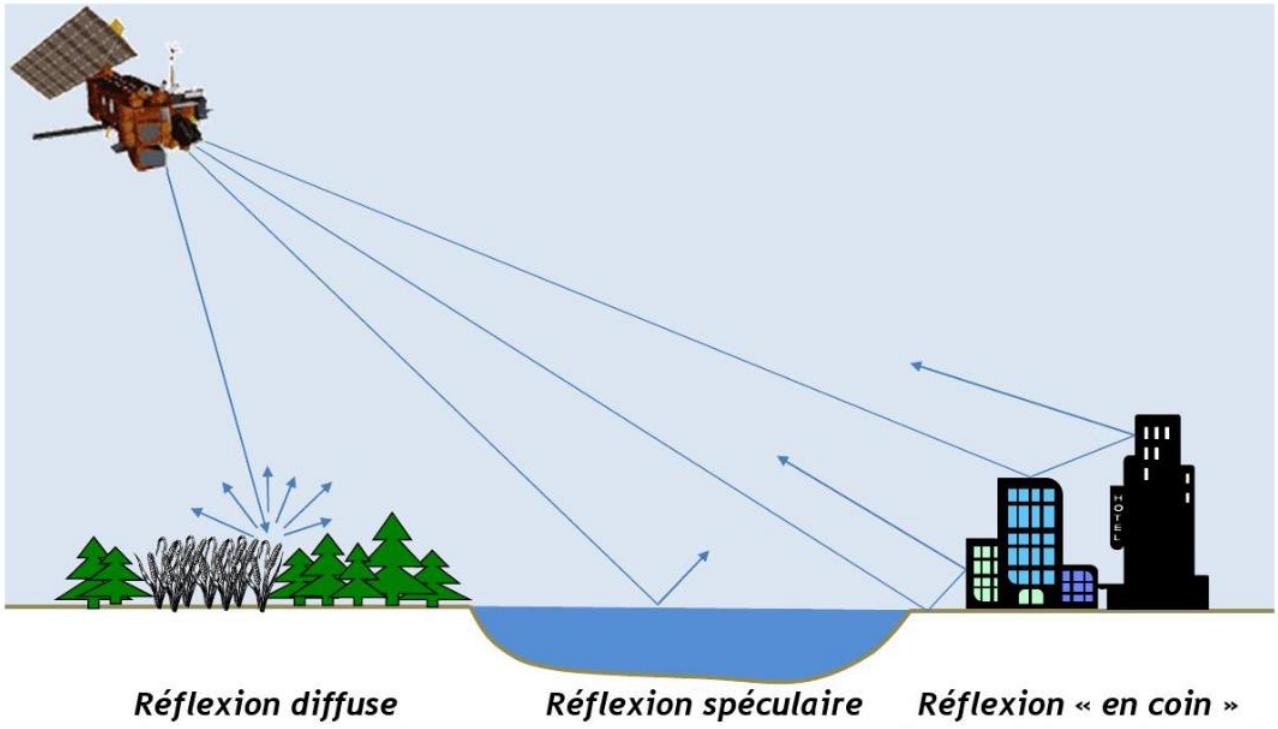


Figure 16 - Illustration d'une réflexion diffuse sur de la végétation, spéculaire sur une surface lisse et en coin (mécanisme de double-rebond). Crédit image : Gédéon (Planète Sciences Midi-Pyrénées)

De plus, l'interaction d'une onde électromagnétique avec un milieu dépend de sa polarisation. La polarisation représente l'orientation du champ électromagnétique dans le plan d'incidence (Figure 14). Il est possible de la contrôler en émission et en réception afin de mesurer le coefficient de rétrodiffusion en fonction de chaque polarisation. Par exemple, si une onde est émise en polarisation verticale et que le radar reçoit uniquement la polarisation verticale, on parlera de rétrodiffusion de co-polarisation (même polarité du transmetteur et du récepteur), dans ce cas VV (Vertical/Vertical). Si la mesure de l'énergie est retournée avec un décalage 90° par rapport à l'onde transmise, on parlera de polarisation croisée.

Connaître les coefficients de rétrodiffusion suivant plusieurs configurations de polarisation permet de déduire certaines caractéristiques géométriques des cibles observées. Citons le cas d'un bâtiment en zone urbaine : l'onde émise en polarisation verticale va interagir fortement avec les murs du bâtiment et être renvoyée dans la même direction, avec la même polarisation. La surface lisse ne l'aura pas dépolarisée. Le coefficient de rétrodiffusion en polarisation VV sera donc important, et sera plus faible en polarisation croisée VH. En revanche, un arbre avec une canopée à orientation aléatoire renverra l'onde dans toutes les polarisations et donc la rétrodiffusion en polarisation VH sera plus importante que pour une cible artificielle (comme un bâtiment).

Ces facteurs composant le coefficient de rétrodiffusion permettent de caractériser un milieu.

4.2.4. PROPRIETES GEOMETRIQUES

Du fait de leur acquisition à l'oblique et de la mesure en distance, les images radar ont une géométrie particulière. En effet, la mesure en distance induit une distorsion des images qu'il faut bien appréhender pour interpréter des images RSO. Plusieurs effets inhérents à la technique de l'imagerie radar impacte la géométrie de l'image acquise : le repliement (foreshortening et layover) et l'ombre (Figure 17).

- Foreshortening : Les pentes faisant face au capteur apparaissent raccourcies de telle sorte qu'une montagne symétrique apparaîtrait dans l'image radar comme si elle était " penchée " vers le capteur, comme le montre la Figure 17 (a). La pente entre les points A et B sera raccourcie dans la zone d'image A′B′, en fonction de l'angle d'incidence du système et de l'angle de la pente ;
- Layover : Dans les zone où l'angle de la pente est supérieur à l'angle d'incidence, cet effet est exacerbé et crée un inversement des points dans l'image. La Figure 17 (b) illustre ce phénomène : les points B et C au sol ont créé une inversion du relief sur l'image (points C' et B') ;
- Ombre : Un point haut peut empêcher les ondes d'atteindre une surface, comme sur la Figure 17 (c). Cette surface ne renverra donc aucun signal et créera une ombre.

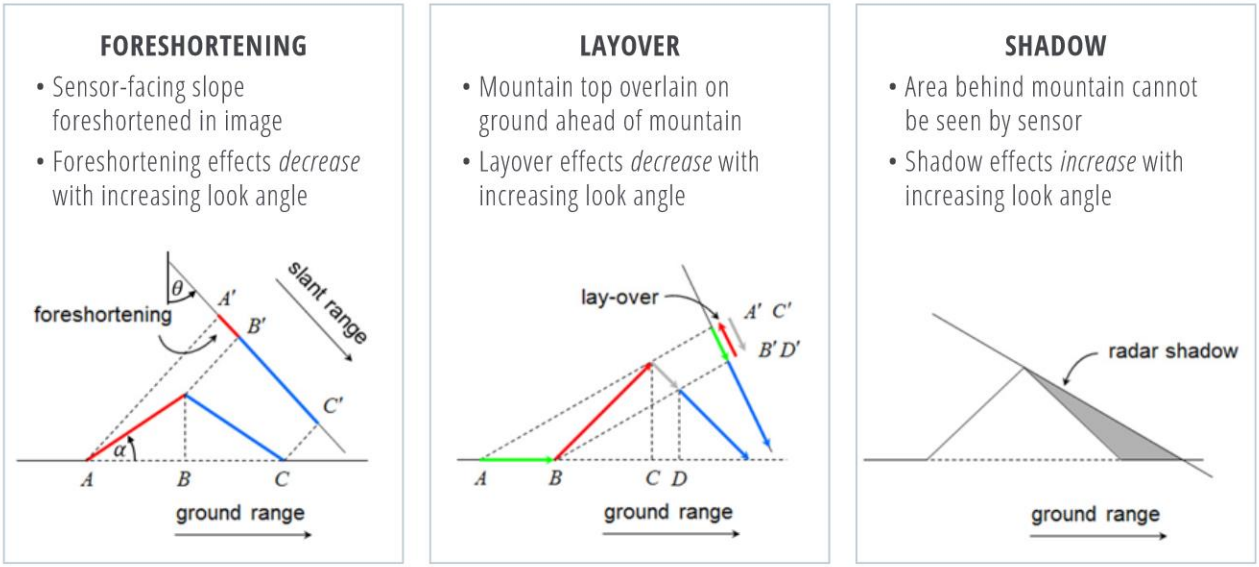


Figure 17 - Principales distorsions géométriques sur les images RSO avec leur dépendance à la géométrie d'acquisition : (a) le foreshortening de l'image, (b) le layover et (c) l'ombre. Copyright : (Flores-Anderson et al., 2019).

4.2.5. PROPRIETES RADIOMETRIQUES

En raison de l'aspect cohérent des ondes électromagnétiques utilisées par le radar, les images RSO sont affectées par un phénomène de chatoiement, appelé speckle en anglais, qu'on modélise par un bruit multiplicatif. Il est causé par la présence de diffuseurs élémentaires avec une distribution aléatoire dans une cellule de résolution. Ces diffuseurs sont les objets qui interagissent avec l'onde électromagnétique, et correspondent à ceux dont la taille est supérieure ou égale à la longueur de l'onde considérée. Par exemple, dans la cellule de résolution d'un terrain forestier, les diffuseurs peuvent être les feuilles, les tiges, les troncs, les objets au sol, etc. La somme cohérente des réponses de chaque diffuseur entraîne de fortes fluctuations de la rétrodiffusion totale d'une cellule de résolution à l'autre.

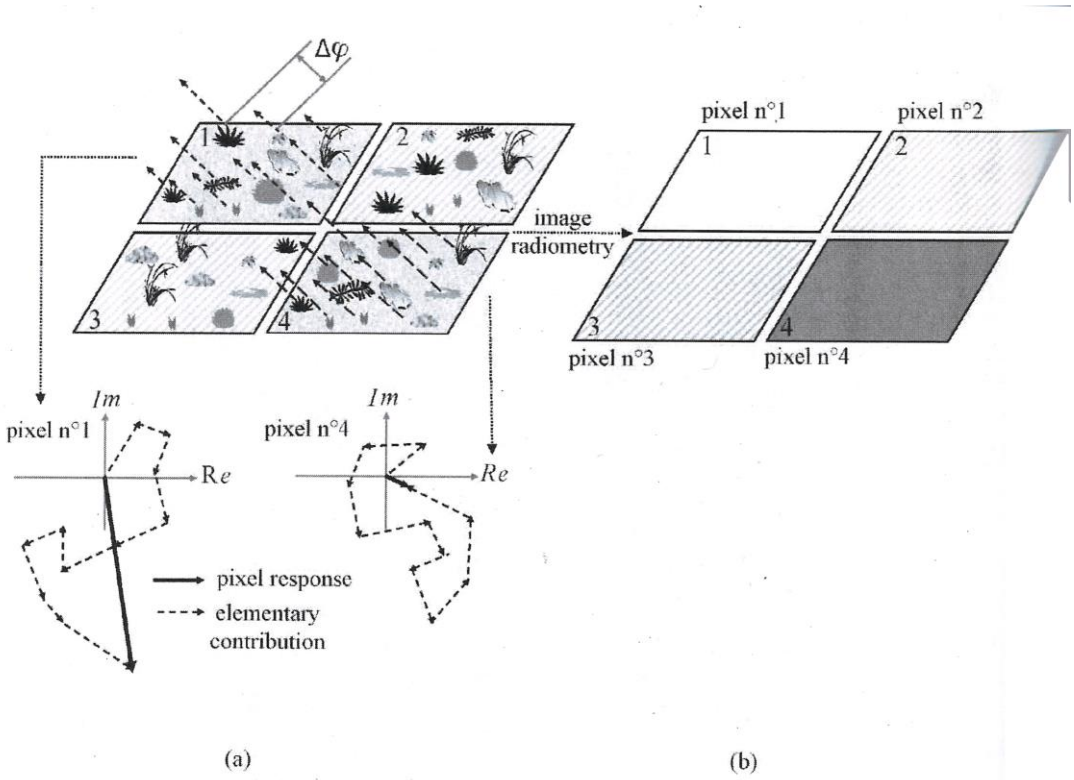


Figure 18 – (a) Origine physique du speckle : effet de la nature cohérente des ondes électromagnétiques utilisées pour le radar et combinaison de la contribution de chaque diffuseur dans chaque pixel. (b) Distribution radiométrique de l'image RSO en résultant.

Par conséquent, l'intensité et la phase dans l'image finale ne sont plus déterministes, mais suivent plutôt respectivement une distribution statistique exponentielle et uniforme (Oliver and Quegan, 2004). Il en résulte sur l'image un effet de bruit rendant la valeur individuelle d'un pixel difficilement exploitable.

La meilleure manière de réduire le speckle sur une zone homogène est de faire la moyenne en intensité des échantillons acquis par le radar. Ceci peut se faire par filtrage spatial et/ou temporel, suivant si l'on préfère préserver l'information spatiale ou temporelle. L'une des techniques les plus utilisée est le filtre de Lee (Lee, 1980) un filtre spatial (utilisé sur une image seule) qui travaille par

minimisation de l'erreur quadratique moyenne sur une fenêtre glissante de taille donnée. Ce filtre permet de lisser la texture d'une zone, mais dégrade la résolution spatiale et donc les contours. Le résultat de ce filtrage est visible sur la Figure 19 (avec une fenêtre de taille 11×11).

Le filtrage d'une image peut également s'effectuer au niveau temporel, par exemple, pour un pixel donné il est possible de faire la moyenne de ses valeurs à différentes dates. Comme on le voit dans la Figure 19, ce type de filtrage permet de conserver les contours d'une image. En revanche, l'information temporelle est perdue.

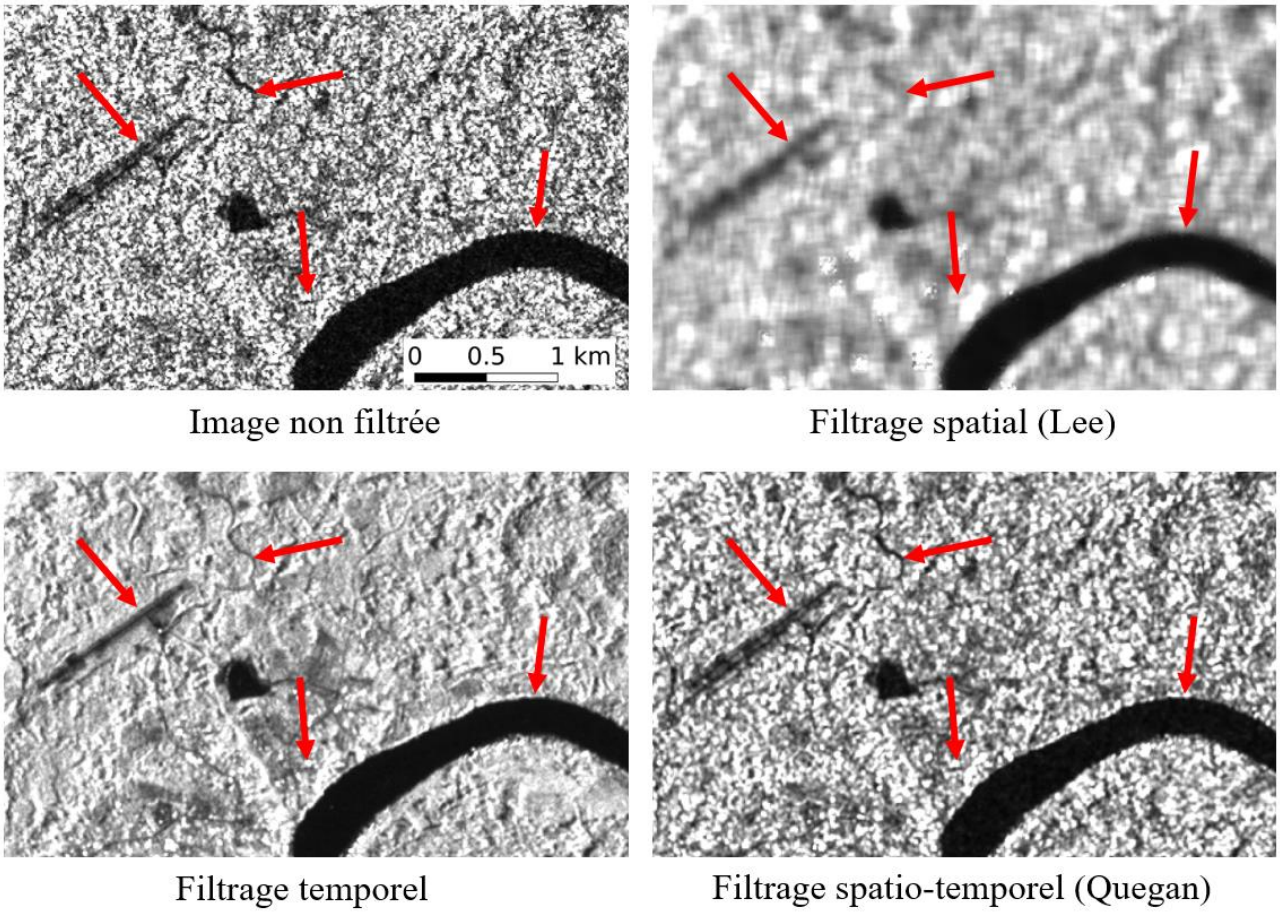


Figure 19 - Image Sentinel-1 avec différents types de filtrage. Les flèches rouges sont des repères pointant des zones d'intérêt de l'image.

Afin d'avoir un compromis entre préservation des contours et préservation de l'information à chaque acquisition, il est possible d'utiliser un filtre multi-canal, utilisant à la fois les informations spatiales, temporelles et polarimétriques ((Quegan and Yu, 2001)). Ce filtre opère en multipliant à chaque pixel, un coefficient définit comme la moyenne de la série temporelle de ce pixel sur la moyenne de son voisinage spatial. Il permet de réduire le speckle en préservant l'information spatiale.

Soit une série de M images I_i ($i \in (1,M)$). L'image J_k restaurée sera de la forme :

$$J_k = \frac{\hat{\sigma}_k}{M} \sum_{i=1}^{M} \frac{I_i}{\hat{\sigma}_i} \qquad 3.10$$

Où $\hat{\sigma}_k = \frac{1}{N} \sum_{i=1}^{k} I_i^{(k)}$ est le moyennage spatial sur N pixels de la fenêtre autour du pixel dans l'image k avec $I_i^{(k)}$les pixels de cette fenêtre. Le résultat de ce filtre sur une fenêtre de taille 3×3 est visible Figure 19. C'est le filtre que nous utiliserons sur toutes les images RSO dans cette thèse (voir section suivante).

4.3. SENTINEL-1

Dans cette étude, les images RSO du satellite Sentinel-1 sont utilisées. Depuis 2014, Sentinel-1 permet l'accès régulier à des images radar et couvre à présent toute la surface terrestre. Les prétraitements requis pour leur utilisation dans notre cas seront développés après la présentation de leurs caractéristiques.

Sentinel-1 (Torres et al., others, 2012) est un système spatial développé et lancé dans la cadre du programme européen Copernicus de la Commission Européenne, constituant un ensemble de satellites d'observation de la Terre et dont l'avantage principal est la fourniture de données diverses, gratuites, régulières et à l'échelle globale. Deux satellites nommés Sentinel-1A et 1B ont été lancés respectivement le 3 avril 2014 et le 25 mai 2016 depuis Kourou. Chaque satellite Sentinel-1 a un temps de revisite de douze jours. Grâce à un phasage des 2 satellites, certaines zones géographiques, dont la Guyane, bénéficient d'une revisite à 6 jours (voir Figure 20). Ils opèrent à environ 700 kilomètres d'altitude et suivent une orbite héliosynchrone, imageant en ascendant et descendant (par rapport au nord), avec une visée à droite.

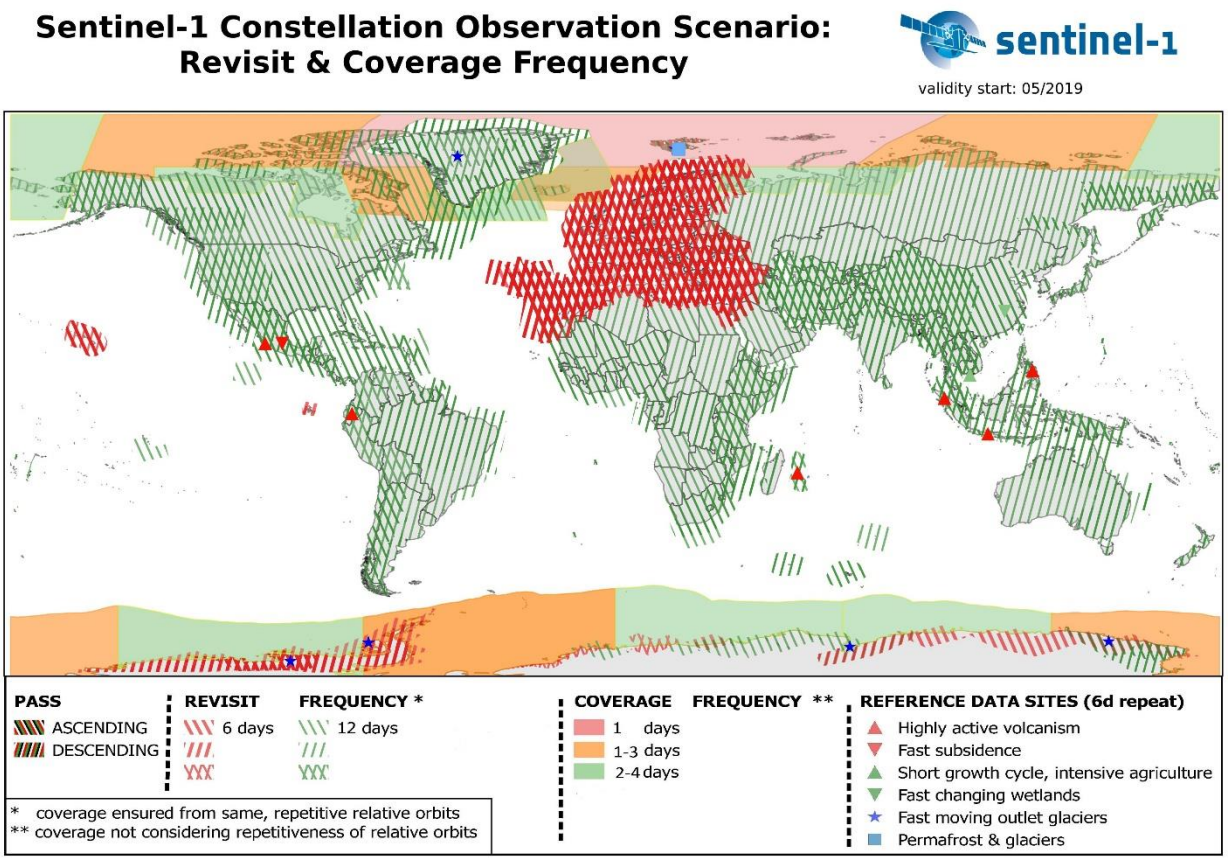


Figure 20 - Exemple de plan d'acquisition de Sentinel-1 qui contient les orbites imagées, le temps de revisite et les polarisations disponibles.

L'instrument principal de Sentinel-1 est un RSO en bande C (5.405 GHz) avec un angle d'incidence variant entre 29° et 46°. A cette fréquence, l'onde pénètre en moyenne de quelques dizaines de centimètres dans la matière végétale. Quatre modes de mesure sont disponibles : le mode Interferometric Wide-swath (IW), le mode Wave (WV), le mode Strip Map (SM) et le mode Extra Wide-swath (EW). Le mode dédié à l'observation des surfaces continentales est le mode IW. Les images IW de Sentinel-1 sont acquises dans 2 polarisations simultanément (VV et VH), avec une résolution native d'environ 3 mètres en distance et 22 mètres en azimut (géométrie radar). Plusieurs niveaux de produit sont disponibles : niveau-0 (échos radar bruts), niveau-1 Single Look Complex (SLC), niveau-1 Ground Range Detected (GRD), et niveau-2 Océan (OCN). Les travaux de cette thèse sont basés sur le traitement des données au format GRD. Ces images ont été géoréférencées et projetées en géométrie sol. Leur résolution est de 20×22 mètres (distance × azimut) et leurs pixels ont été ré-échantillonnés pour former des images avec une taille de pixel de 10×10 mètres.

Pour résumer, les images Sentinel-1 sont une source sans précédent dans l'histoire des images RSO, car elles fournissent pour la première fois :

- Des séries temporelles denses d'images radar (acquisitions tous les 6 à 12 jours suivant la

zone géographique) ;

- Une couverture globale des terres émergées ;
- Une résolution native d'environ 3×22 mètres.

Ces séries d'images exploitables indépendamment de la couverture nuageuse offrent une opportunité unique de suivre les déboisements systématiquement, tous les 6 ou 12 jours. De plus, la continuité des satellites Sentinel est assurée jusqu'au-delà de 2030, offrant la possibilité d'un suivi à long terme.

Les données utilisées dans l'étude

Dans cette thèse, nous avons utilisé les produits GRD de Sentinel-1 acquis sur la Guyane en mode IW entre le 8 novembre 2015 au 31 décembre 2020. La Guyane française est couverte par 4 orbites Sentinel-1 (Figure 21).

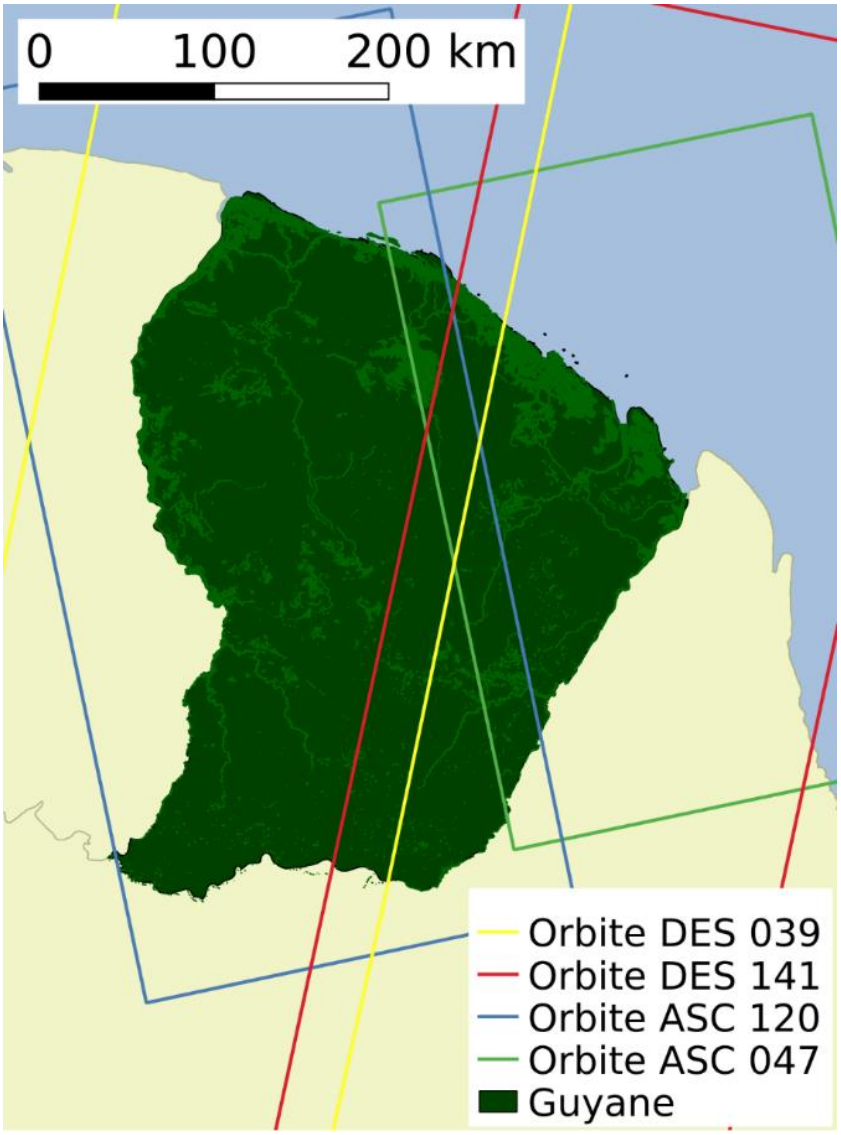


Figure 21 - Orbites Sentinel-1 couvrant la Guyane

Sur la zone d'étude, Sentinel-1A fournit des séries temporelles régulières depuis le 8 novembre 2015 avec un temps de revisite de 12 jours en orbites ascendantes et descendantes. A partir du 26 septembre 2016, le temps de revisite a été réduit de 12 à 6 jours en orbites ascendantes, avec l'acquisition des données par les 2 satellites (Sentinel-1A et 1B). Ceci représente un total de 2 204 produits Sentinel-1 acquis et utilisés sur la période 2015-2020.

Le prétraitement a été effectué par S1 Tiling (Koleck et al., 2019 ; https://gitlab.orfeo-toolbox.org/s1-tiling/s1tiling), une chaîne automatisée produite par le CNES et le CESBIO, et basée sur l'Orfeo Toolbox (https://www.orfeo-toolbox.org/), qui intègre les étapes suivantes :

- Le téléchargement des images : via la plateforme française PEPS (https://peps.cnes.fr/) qui les fournit librement et rapidement (produits disponibles 24h après leur acquisition) ;
- La calibration : conversion des digital numbers en γ^0 (coefficient de rétrodiffusion radar) ;
- La suppression des pixels bruités en bord d'image ;
- L'orthorectification : projection de l'image acquise pour qu'elle soit superposable à une carte donnée (ici le découpage est fait en tuile MGRS de 120×120 km identiques à celles définies pour les images Sentinel-2), avec prise en compte des déformations du terrain.

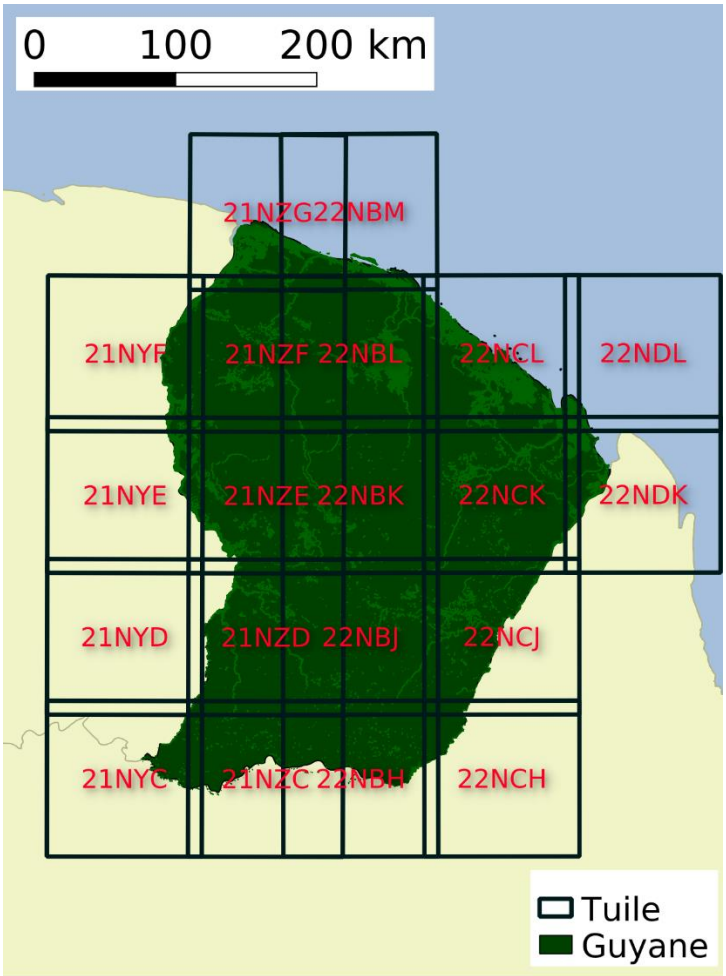


Figure 22 - Découpage de la Guyane en tuiles MGRS

Pour finir, un filtre spatio-temporel (ou multi-image) (Quegan and Yu, 2001) (voir paragraphe 4.2.5) a été appliqué sur une fenêtre de taille 3×3 et adapté à un contexte quasi-temps réel (pour filtrer l'image acquise à une date d, on utilise seulement les M images antérieures à cette date).

🔔 Ce qu'il faut retenir

Dans ce chapitre, nous avons vu que :

- La Guyane est un territoire majoritairement couvert de forêt tropicale humide ;
- Elle a une saison sèche et une longue saison humide entrecoupée d'un mois et demi plus sec ;
- Sur la Guyane, de nombreuses données spatialisées d'occupation et d'usage du sol sont disponibles ;
- Il n'existe pas d'information temporelle précise et exhaustive sur les déboisements observés ;
- Une campagne terrain a été mise en place afin de compléter le jeu de données existant et de combler les informations manquantes ;
- Les données Sentinel-1 sont relativement insensibles à la couverture nuageuse et permettent d'obtenir de nombreuses informations sur les cibles observées.
-

L'ensemble des données utilisées dans cette thèse est récapitulé dans le tableau ci-dessous.

Tableau 2 - Tableau récapitulatifs des ensembles de données utilisés dans cette thèse.

Données	Source	Type	Couverture spatiale	Couverture temporelle
Occupation du sol (inclut l'agriculture)	PAG (via https://www.geoguyane.fr/)	Vecteur (Polygone)	Parc Amazonien de Guyane	2005-2018
Chemins d'exploitation	ONF	Vecteur (Ligne)	Guyane	1960-2018
Parcs à grumes	ONF	Vecteur (Polygone)	Parcelles d'exploitation (non exhausif)	2016-2018

Données	Source	Type	Couverture spatiale	Couverture temporelle
Trouées d'abattage	ONF	Vecteur (Polygone)	Parcelles d'exploitation (non exhausif)	10/2017-11/2017
Exploitation aurifère	ONF/EMOPI (confidentielle)	Vecteur (Polygone)	Guyane	1950-04/2019
Occupation du sol	IGN (via https://geoservices.ign.fr/)	Vecteur (Polygone + ligne)	France	Jusqu'à maintenant
Titres miniers	Gouvernement français (via https://camino.beta.gouv.fr/)	Vecteur (Polygone)	France	1844-maintenant
Carte géologique	BRGM (via https://infoterre.brgm.fr/)	Vecteur (Polygone)	Guyane	-
Occupation du sol	ECOSEO	Raster	Guyane	2015
Alertes pertes forestières	Global Forest Watch (via https://www.globalforestwatch.org/)	Raster	Tropiques	2015-maintenant
Abattis	Campagne terrain réalisée dans le cadre de cette thèse	Vecteur (Polygone)	Maripasoula-Papaïchton (non exhaustif)	2019
Images RSO	Sentinel-1 (via https://peps.cnes.fr/)	Raster	Monde	2015-maintenant

Certaines de ces données ont été analysées a priori (voir chapitre 4), afin de définir les méthodes décrites dans le chapitre 5.

✋ Ma contribution

J'ai d'abord cherché toutes les données disponibles librement sur la Guyane sur internet. Je suis également allée au contact des acteurs locaux (ONF et EMOPI) en leur expliquant mon besoin afin d'obtenir des données de référence indépendantes de cette étude. Afin d'avoir des données permettant de compléter mon étude, j'ai présenté mon projet au Service de la Statistique et de la Prospective (SSP) du Ministère de l'Agriculture et de l'Alimentation qui m'a permis de mettre en

place la campagne terrain, avec un agent de la DAAF de Guyane. J'ai défini le protocole de la campagne terrain et fait les synthèses. J'ai également passé deux jours avec l'enquêteur pour visiter les abattis apparus en octobre 2019. Sur le plan informatique, j'ai contribué à la chaine de prétraitement S1tiling (Koleck et al., 2019) des données Sentinel-1, notamment en intégrant des conditions de téléchargement propres au projet et en ajoutant une fonction permettant de couper automatiquement les bords d'une image de manière automatique quand l'information contenue dedans n'est pas pertinente.

Chapitre 4 - Analyse et interprétation

Contenu

1. Analyse des données de référence 77

1.1. Hauteur des forêts 78

1.2. Topographie 79

1.3. Taille des surfaces déboisées 80

1.4. Caractéristiques géographiques 82

2. Interprétation des signatures temporelles Sentinel-1 sur les données de référence 87

Dans ce chapitre, les données présentées dans le chapitre 3 sont analysées en vue d'élaborer les méthodes utilisées dans le cadre de cette thèse et décrite dans le chapitre 5.

1. Analyse des données de référence

Dans un premier temps, les données de référence du PAG, de l'OAM et de l'ONF sur les années 2016 à 2018, ainsi que les données collectées lors de la campagne terrain ont été analysées pour en déduire leurs caractéristiques.

1.1. HAUTEUR DES FORETS

(Simard et al., 2011) a publié une carte mondiale de la hauteur des forêts en 2005, à une résolution de 1 km, et basée sur les données lidar de Geoscience Laser Altimeter System (GLAS) à bord d'ICESat. A partir de cette carte et de la couche forêt de la carte d'ECOSEO, on observe une hauteur moyenne de 36 mètres sur les forêts Guyanaises.

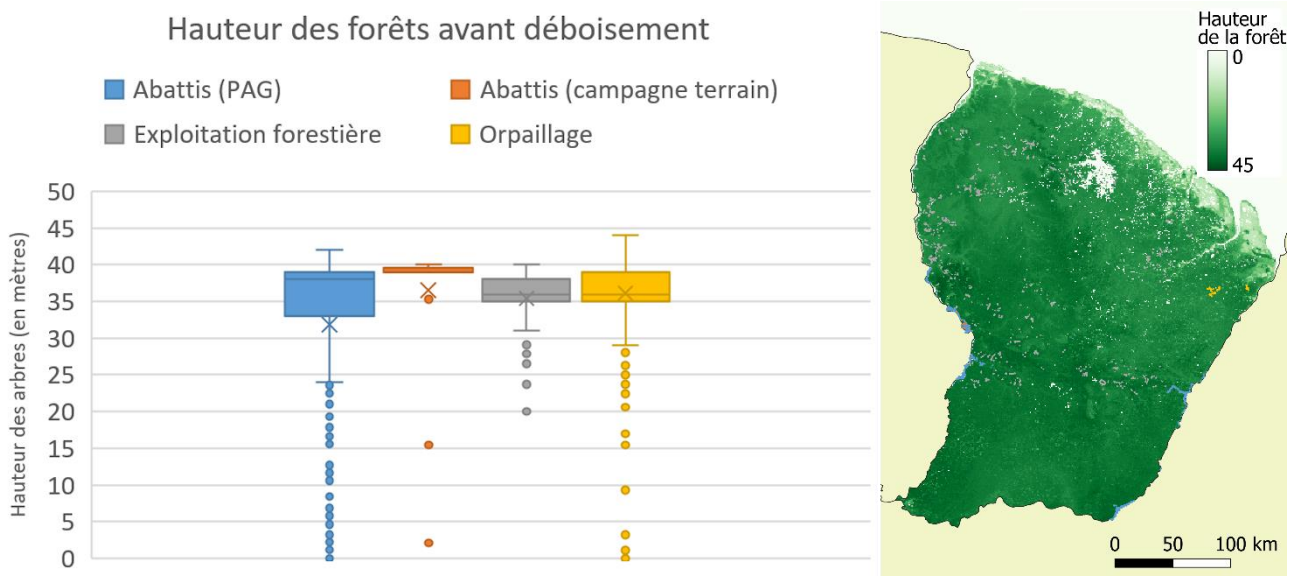


Figure 23 – Gauche : Boîtes à moustaches des hauteurs des arbres en mètres avant déboisement sur les parcelles déboisées présentes dans les jeux de données. Droite : Carte de hauteur de la forêt en mètres sur la Guyane (Simard et al., 2011), avec localisation des parcelles de référence.

La Figure 23 montre la hauteur des arbres sur les parcelles de référence avant déboisement. La hauteur médiane sur l'orpaillage et l'exploitation forestière est de 36 mètres. Elle augmente à 38 et 39 mètres sur les jeux de données relatifs aux abattis (qui se trouvent dans le PAG, dans la moitié sud de la Guyane). Les points ayant des valeurs particulièrement faibles proviennent d'endroits proches de cours d'eau ou de zones de non-forêt : la résolution spatiale de la carte à 1 km (en comparaison des parcelles déboisées), peut mener à des incertitudes élevées.

1.2. TOPOGRAPHIE

La carte des pentes de Guyane, dérivée du MNT SRTM à 30 mètres est visible sur la Figure 24 (à droite). Environ 12% de la surface de la Guyane a une pente supérieure à 15°. La distribution des pentes sur chaque jeu de données de référence est présentée à gauche de la Figure 24.

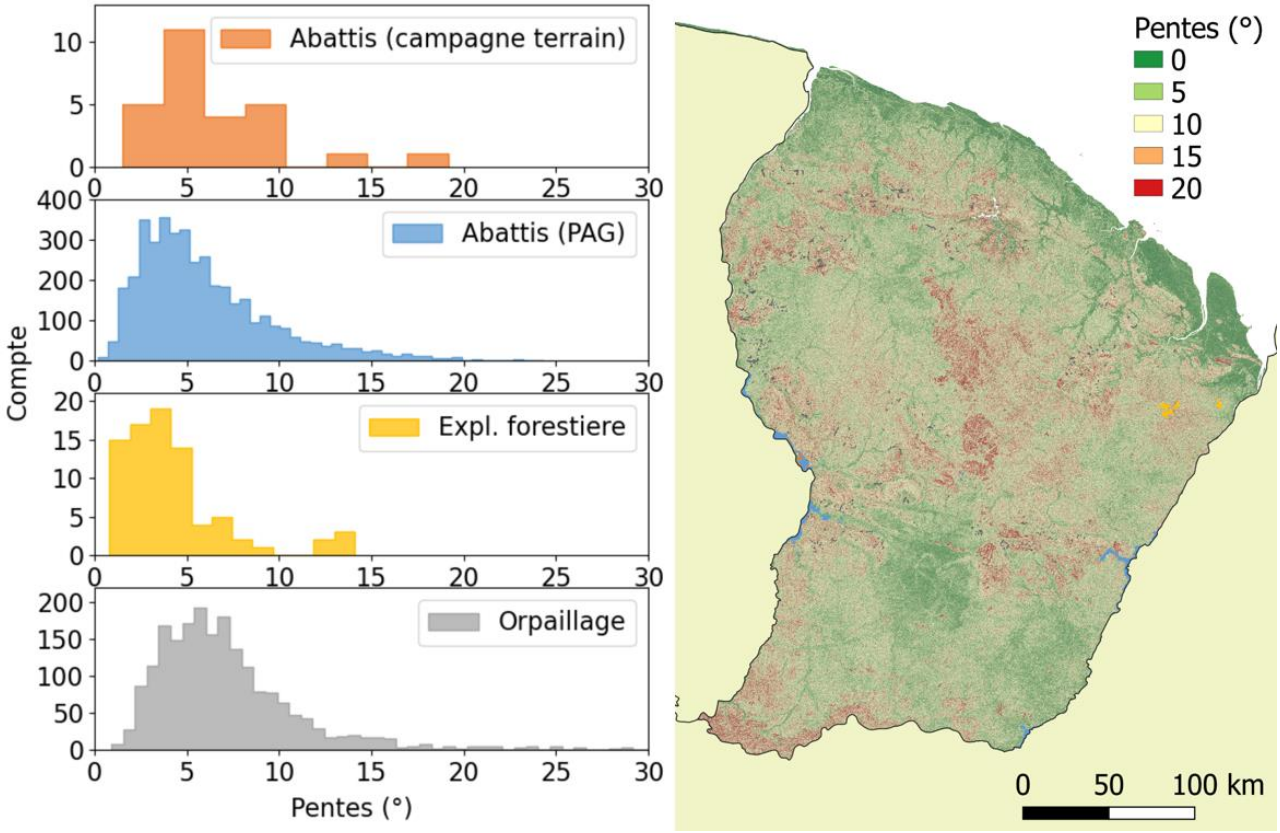


Figure 24 - Distribution des pentes sur les différents types de déboisements observés en Guyane (à gauche). Carte des pentes en degrés, avec localisation des parcelles de références (à droite). Les pentes sont dérivées du MNT SRTM à 30 mètres (http://dwtkns.com/srtm30m/).

L'exploitation forestière a lieu dans des zones peu inclinées (en moyenne 4.2°). Les abattis, eux, peuvent se trouver dans des zones un peu plus pentues (en moyenne 6° pour les abattis du PAG et 6.4° pour les abattis de la campagne terrain). Les parcelles d'orpaillage sont sur des pentes d'en moyenne 7.1°, et presque 5% d'entre elles sont au-dessus de 15°. En effet, les vallées sont généralement de bons sites pour l'exploitation aurifère.

1.3. TAILLE DES SURFACES DEBOISEES

Les abattis sont majoritairement de petites surfaces. Les données du PAG affichent une surface moyenne de 0.3 ha, tandis que celles de la campagne terrain donnent des surfaces moyennes de 0.6 ha. Cette différence est due à la présence de plusieurs polygones très petits dans le jeu de données du PAG (voir illustration Figure 25). Ces polygones proviennent de changements d'usage du sol sur plusieurs années autour des abattis existants, par exemple pour un agrandissement de l'abattis, ou une mise en jachère d'une partie de l'abattis avant ou après le reste de la parcelle.

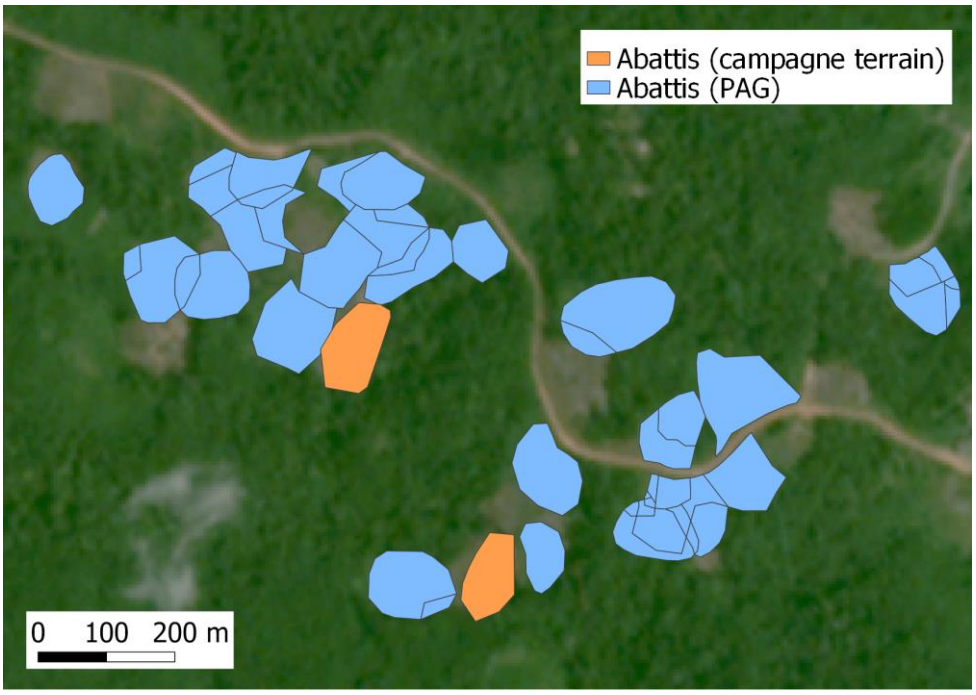


Figure 25 - Illustration de polygones de référence provenant de la campagne terrain à Maripasoula (en orange), et de polygones de référence provenant des données du PAG (en bleu). Le fond de carte provient d'une mosaïque d'images Planet (www.planet.com/nicfi).

La Figure 26 (haut) montre la distribution de la surface des polygones de chaque jeu de données. La forme des parcelles peut être exprimée par un indice de compacité ($\frac{périmètre^2}{aire}$). Celui-ci sera faible pour une structure compacte, et élevé pour une structure longiligne. Les abattis sont souvent de forme ronde ce qui s'exprime par un indice de compacité faible.

Dans les polygones d'exploitation forestière se trouvent des parcs à grumes et des débarcadères de 0.23 ha en moyenne, ainsi que des routes dont la surface est plus élevée, mais pour lesquelles la construction a pu prendre plus de temps. Les premiers ont un indice de compacité faible, tandis que les routes, objets longilignes, ont un indice de compacité élevé (le plus haut étant de 946.7). La surface moyenne des parcelles d'orpaillage est de 1.7 ha, et va jusqu'à 20.2 ha. Si les petites zones ont souvent un indice de compacité faible, les plus grandes zones, avec leurs formes sinueuses, affichent un indice de compacité plus fort, inférieur à ceux des routes, mais plus élevé

que celles des abattis. En revanche, comme pour les routes, il est difficile de savoir sur combien de temps les grandes parcelles ont été déboisées.

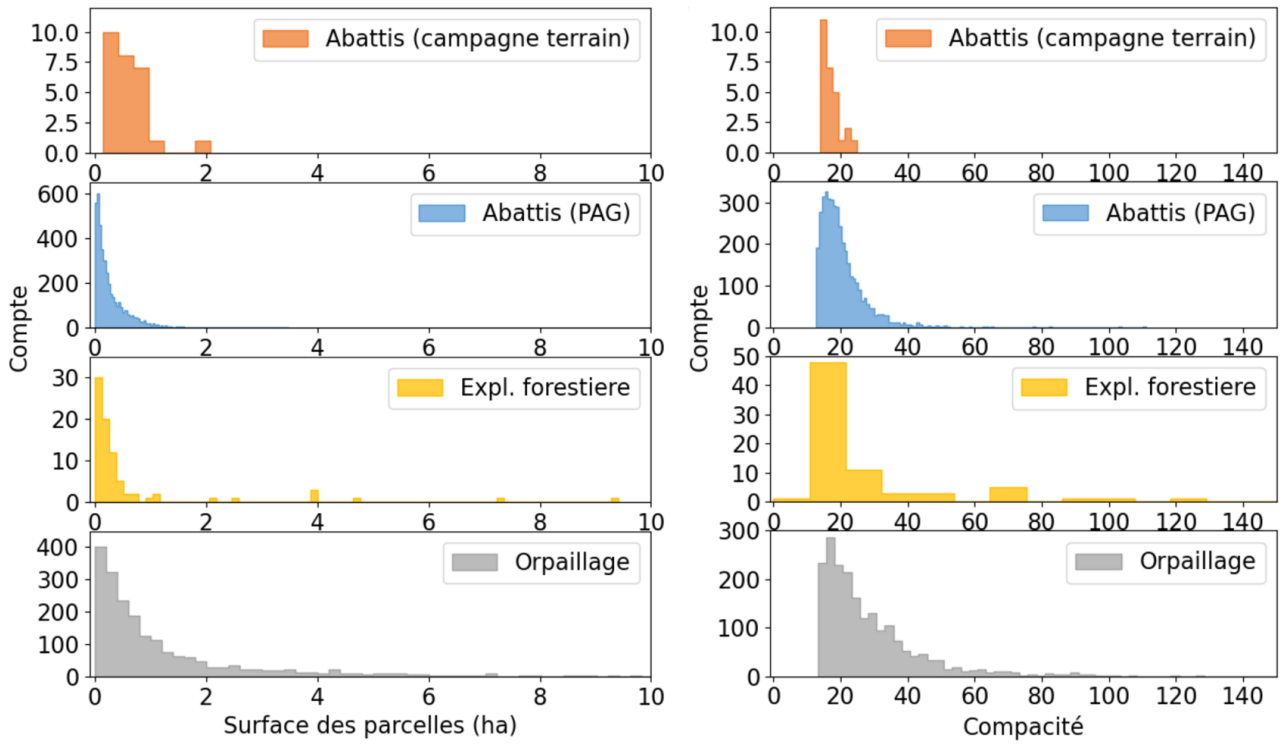


Figure 26 - Distribution des surfaces (en haut) et de la compacité (périmètre ² / aire) (en bas) des parcelles déboisées par jeu de données de référence.

1.4. CARACTERISTIQUES GEOGRAPHIQUES

La Guyane est un territoire dont les usages du sol après déboisement sont distincts géographiquement. Les Figure 27 à Figure 30, montrent l'analyse des jeux de données de référence par rapport à des indicateurs géographiques : habitations, surfaces hydrographiques, zones d'orpaillage passé, routes. Ceux-ci proviennent respectivement des couches « ZONE_D_HABITATION » et « SURFACE_HYDROGRAPHIQUE » de la BD TOPO, des données OAM datant d'avant 2016, et de la couche « TRONCON_DE_ROUTE » de la BD TOPO.

Sur la Figure 27, on voit que les abattis se situent souvent près des habitations (distance maximum de 10,5 km). Les plus éloignés sont ceux au bord de l'Oyapock, accessibles par le fleuve depuis Camopi. En excluant ces abattis, la distance maximum aux habitations est de 4.8 km. Les données liées à l'exploitation forestière se trouvent sur deux grandes zones : l'une en pleine forêt de Régina (3°9 N, 52°1 O), loin de toute habitation (minimum 22 km), et l'autre proche de Saint Georges à moins de 5 km des habitations de la ville. Il n'est donc pas évident de tirer des conclusions quant à la distance de l'exploitation forestière aux habitations. Pour finir, l'orpaillage semble indépendant de la distance aux habitations : on en trouve entre 0 et 35 km. Cependant, il est à noter que les zones d'habitation proches de sites d'exploitation aurifère ne sont autres que des villages d'orpailleurs montés près des grandes exploitations légales. On remarque aussi que l'orpaillage loin de ces zones d'habitation est proche d'autres zones d'orpaillage.

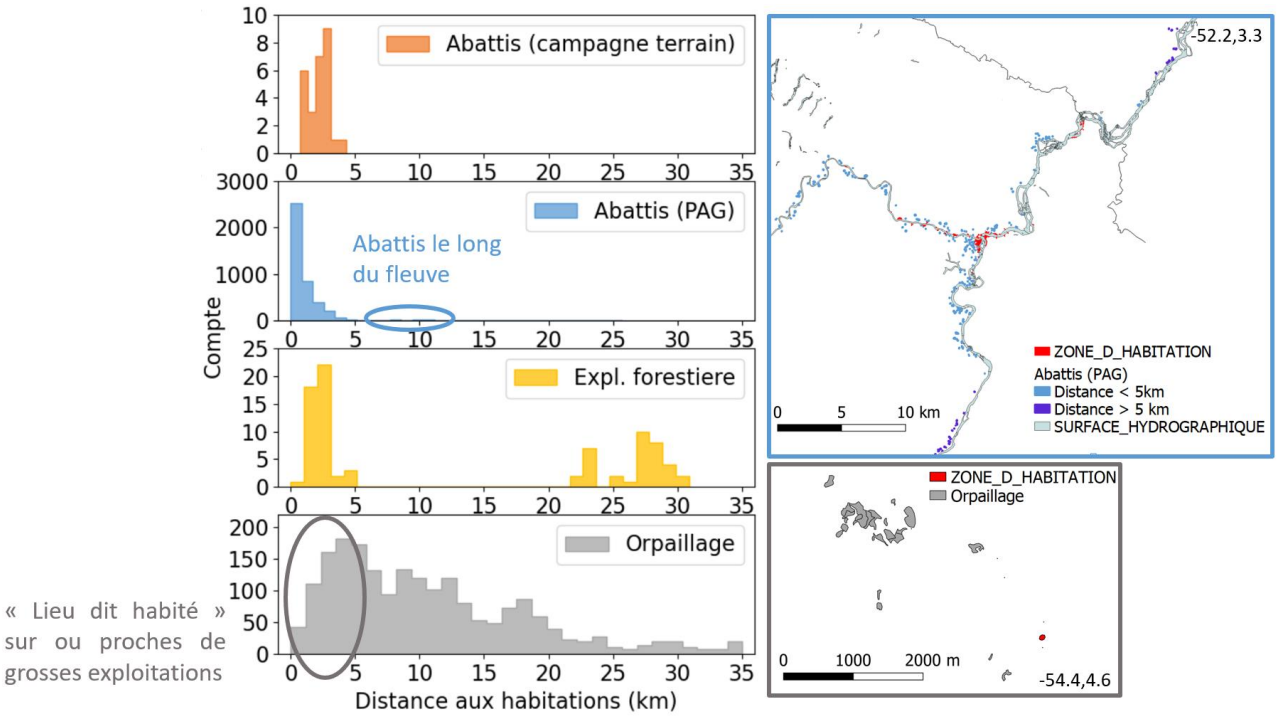


Figure 27 - Distance entre les polygones des jeux de données et la couche ZONE_D_HABITATION de la BD TOPO.

La Figure 28 montre l'histogramme des distances entre les parcelles des jeux de données et les surfaces hydrographiques données par la BD TOPO. On observe que la majorité des abattis et de l'orpaillage se trouvent proches des surfaces hydrographiques, utilisées comme voies de transport en Guyane. De plus, les villes et villages se forment souvent autour de points d'eau et l'orpaillage de Guyane est principalement alluvionnaire (nécessitant des surfaces en eau). Ainsi, les abattis se trouvent à une distance maximum de 2.5 km des zones d'eau (4.3 km si l'on ajoute les abattis de Saül). L'orpaillage, principalement à moins de 3 km, peut s'étendre jusqu'à 10.5 km. Nous émettons l'hypothèse que l'orpaillage se trouvant loin des surfaces hydrographiques est primaire (l'or est cherché dans la roche). L'exploitation forestière n'est pas particulièrement proche des surfaces hydrographiques, et semble en être indépendante (puisque n'en nécessitant pas pour ces activités).

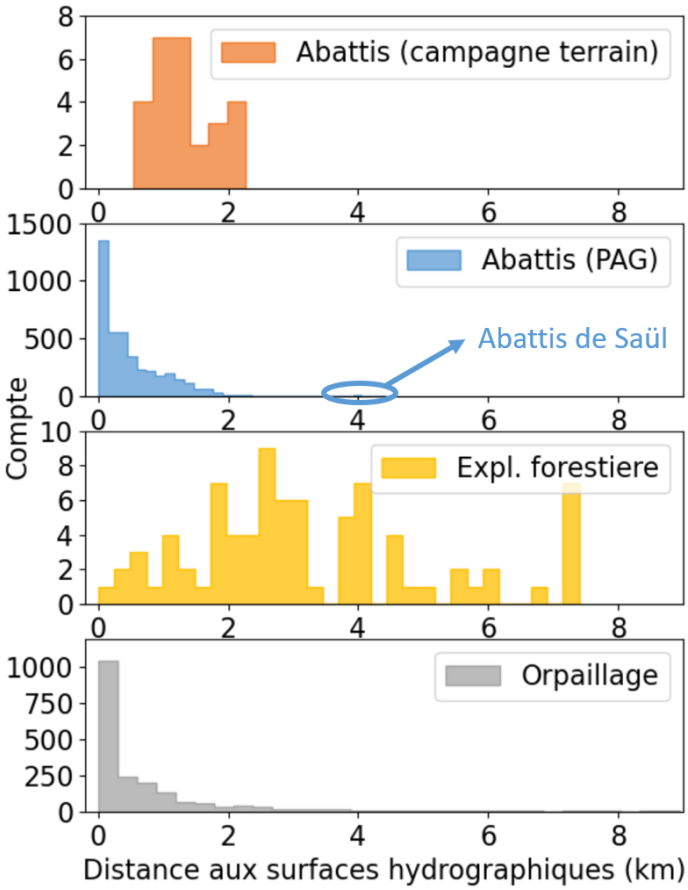


Figure 28 - Distance entre les polygones des jeux de données et la couche SURFACE_HYDROGRAPHIQUE de la BD TOPO.

La distance entre l'orpaillage passé et les parcelles des jeux de données est donnée Figure 29. Les nouvelles parcelles liées à l'exploitation aurifère se trouvent souvent très proches d'anciennes (à moins de 5 km). 56% de ces parcelles sont situées à moins de 200 mètres d'un ancien site aurifère. A l'inverse, les sites d'exploitation forestière présents dans le jeu de données étaient tous à plus de 13 km d'une zone d'orpaillage. Certains abattis de Maripasoula et Papaïchton sont à moins de 500 mètres d'une zone d'orpaillage. Ils représentent moins d'un pour cent du total des abattis sur le PAG.

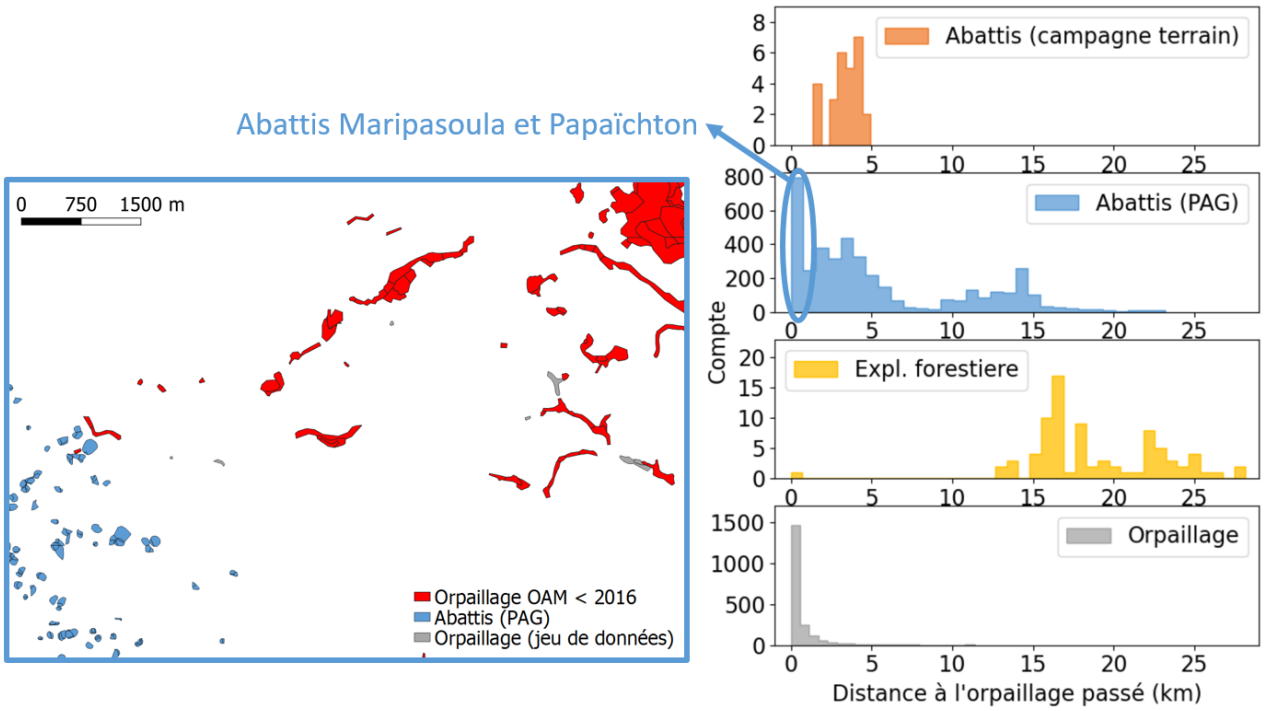


Figure 29 - Distance entre les polygones des jeux de données et les données de l'OAM antérieures à 2016.

La majorité des déboisements liés aux abattis sont très proches d'une route, à l'exception de ceux qui se trouvent proches de Lipo-Lipo (au sud de Maripasoula). Les parcelles d'exploitation forestière proches de Saint Georges sont proches d'une route, mais pas celles de la forêt de Régina. Des pistes ont pourtant été créées pour y accéder, mais elles ne sont pas incluses dans la couche de routes utilisée ici. Les orpailleurs semblent pouvoir se passer de ces voies de communication, l'orpaillage se trouvant à n'importe quelle distance d'une route. Il est possible de penser que les exploitants utilisent plus souvent le réseau hydrographique pour leurs déplacements.

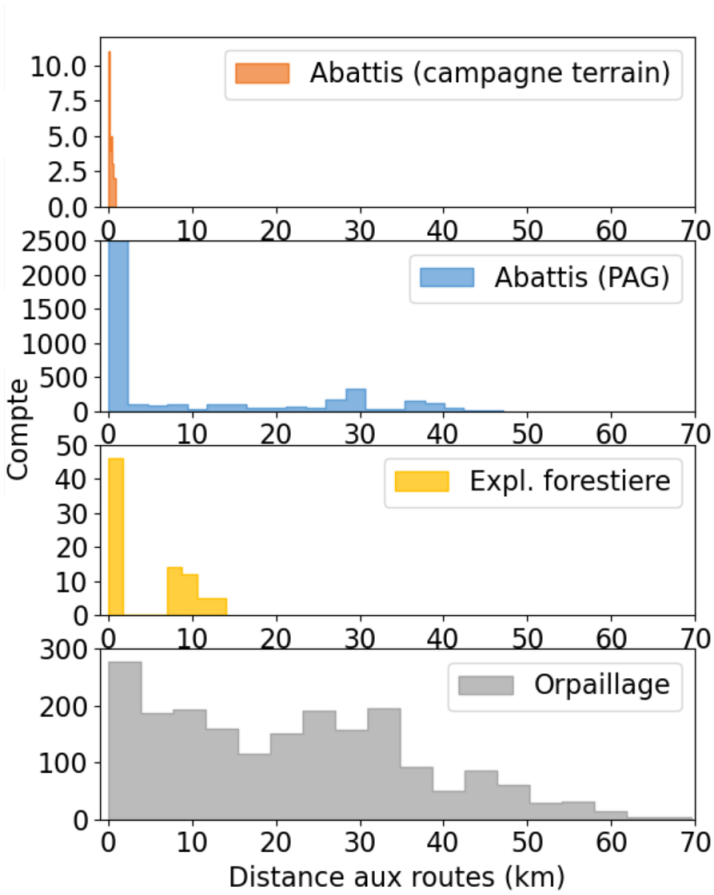


Figure 30 - Distance entre les polygones des jeux de données et la couche TRONCON_DE_ROUTE de la BD TOPO.

Ce travail est non-exhaustif, car il n'inclut que les abattis du PAG, qu'une partie de l'exploitation forestière en Guyane et qu'aucun déboisement n'a été traité sur le littoral. Cependant, il permet de mettre en évidence qu'une séparation des usages du sol après déboisement pourrait se faire à l'aide d'indicateurs géographiques. Nous avons également observé que certaines couches d'indicateurs pouvaient être affinées (notamment à l'aide de leur attribut) afin de mieux répondre au besoin.

En conclusion, d'après les données terrain, on observe que :

- **Les forêts de Guyane ont une hauteur majoritairement supérieure à 35 mètres ;**
- **Les déboisements ont lieu principalement sur des pentes inférieures à 15°. Les parcelles déboisées pour l'exploitation forestière sont sur des terrains plutôt plats, alors qu'une partie de l'orpaillage a lieu dans le fond de vallées, affichant ainsi des pentes plus raides ;**
- **Globalement, les petites parcelles déboisées affichent une faible compacité (inférieure à 30). Seules les grandes parcelles longilignes d'orpaillage et les routes liées à l'exploitation forestière affichent une compacité plus forte.**

De plus :

- **Les déboisements liés aux abattis se trouvent proche de zones d'habitation, sauf ceux qui sont très proches d'un fleuve (qui devient la voie d'accès). Ils sont généralement à distance des zones d'orpaillage, à l'exception de certains abattis de Maripasoula et Papaïchton ;**
- **Les données liées à l'exploitation forestière sont toutes à au moins 13 km d'un site d'orpaillage passé et n'ont pas l'air d'avoir de lien avec les surfaces hydrographiques ;**
- **Plus de la moitié de l'extraction aurifère a lieu à moins de 200 mètres d'un site déjà exploité, et se situe majoritairement proche de surfaces hydrographiques.**

2. Interprétation des signatures temporelles Sentinel-1 sur les données de référence

Comme cela est décrit dans le chapitre 2, la faible sensibilité des images RSO à la couverture nuageuse est un avantage non négligeable pour le suivi régulier des zones tropicales. La rétrodiffusion moyenne sur la forêt dense en Guyane est relativement constante, quel que soit le capteur, la polarisation et l'orbite (donc à fortiori l'angle d'incidence), comme le montre la Figure 31.

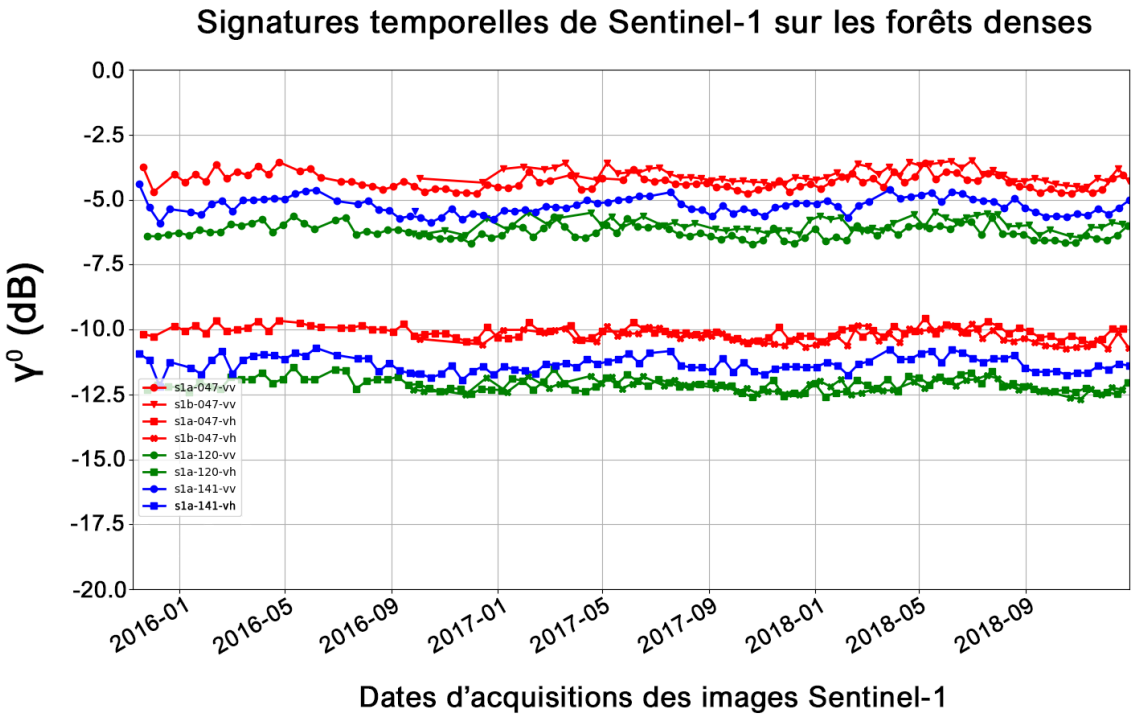


Figure 31 - Rétrodiffusion du signal radar sur la forêt dense en Guyane (en dB) pour 3 différentes orbites (numéros relatifs : 47, 120, 141). La légende, de la forme s1a-047-vv, donne le capteur, l'orbite et la polarisation des signaux.

La rétrodiffusion en bande C est déterminée par les caractéristiques physiques de la forêt (géométrie, caractéristiques diélectriques), mais aussi par des facteurs environnementaux tels que l'humidité du sol et la topographie, et par les conditions d'acquisitions de Sentinel-1 (Figure 32). Quand un déboisement apparaît, le signal varie en fonction du type et de l'intensité de la perturbation ainsi que son état a posteriori. Les parcelles nettoyées ou brûlées montrent souvent une baisse de la rétrodiffusion, alors que les troncs restés au sol peuvent présenter des valeurs de rétrodiffusion identiques, voire supérieures, à celles des forêts intactes, en raison du mécanisme de diffusion à double rebond tronc-sol.

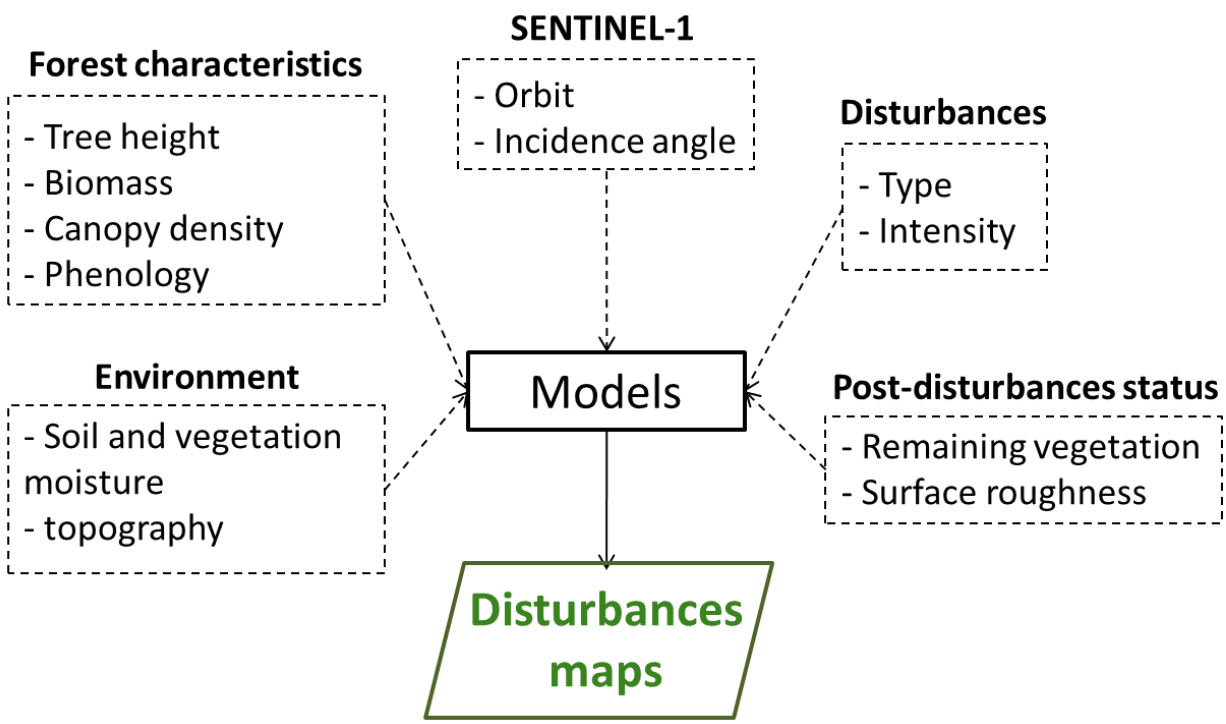


Figure 32 - Facteurs affectant le signal Sentinel-1, et donc la détection d'un déboisement. (Mermoz et al., 2018).

Exemples de séries temporelles sur différents déboisements

Dans cette section, nous utiliserons les données de terrain collectées à Maripasoula pour observer le comportement du signal radar sur ces déboisements (dus aux abattis). Les angles d'incidence sur cette zone sont assez similaires entre l'orbite ascendante et descendante (respectivement 37 et 39°).

Sur la parcelle de la Figure 33-Figure 34, le signal reste constant de janvier 2017 à mi-août 2019, traduisant une surface stable durant cette période. Ensuite, le signal chute dès l'apparition de la coupe et cette chute reste constante sur l'orbite ascendante (Figure 33, figure du haut à gauche). Sur cette parcelle, la pente varie de 6 à 13°, et le sommet de cette pente est placé côté capteur dans une configuration ascendante. De par cette géométrie, le signal rétrodiffusé par la parcelle sera plus faible que s'il était observé d'un autre point de vue comme en orbite descendante (Figure 33, figure du haut à droite). En orbite descendante, on observe une baisse de signal, moins constante et qui remonte sur la fin de l'année. La baisse est plus faible en polarisation VV qu'en VH. Cela pourrait provenir de la perte de la masse végétale en volume qui a fait fortement baisser le signal en VH, tandis que la conservation de plusieurs troncs verticaux empêche le signal VV de trop diminuer à cause du mécanisme de double-rebond sol-tronc plus fort en VV qu'en VH. Un reportage (*Notre*

part d'Amazonie[2]) montre cet abattis coupé depuis peu (2:48 min). On voit que dès la coupe, il y a relativement peu de masse végétale restante au sol (comparé à d'autres parcelles). Le pic de signal observé mi-octobre 2019 est probablement dû à une humidité du sol plus importante que d'habitude, induite par des précipitations. Lors de la visite du 10 octobre, la mise en culture avait déjà commencé (voir sur la Figure 34 en haut à droite). La pousse de cette végétation est un facteur qui peut expliquer la remontée du signal radar en fin d'année.

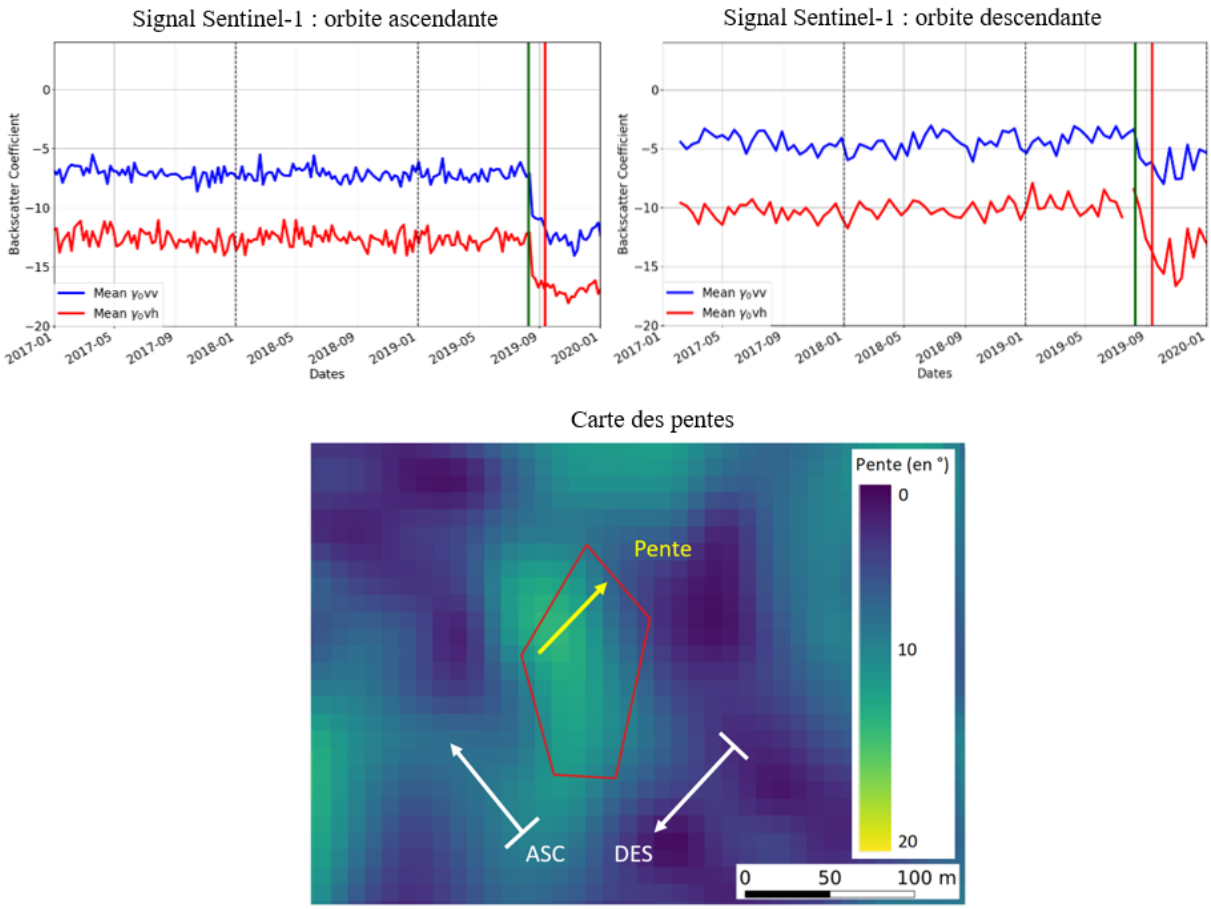


Figure 33 - Haut : signal Sentinel-1 rétrodiffusé en ascendant (gauche), et en descendant (droite). La ligne verticale verte montre le jour estimé de la coupe. La ligne verticale rouge montre la date du brûlis. Bas : image des pentes de la zone et direction des orbites ascendantes et descendantes. La flèche jaune représentant la pente pointe vers le point le plus bas.

[2] https://www.youtube.com/watch?v=eRYXduXtdqE&t=1s

Parcelle vue du haut

Parcelle vue du bas

Reportage dans lequel se trouve cet abattis

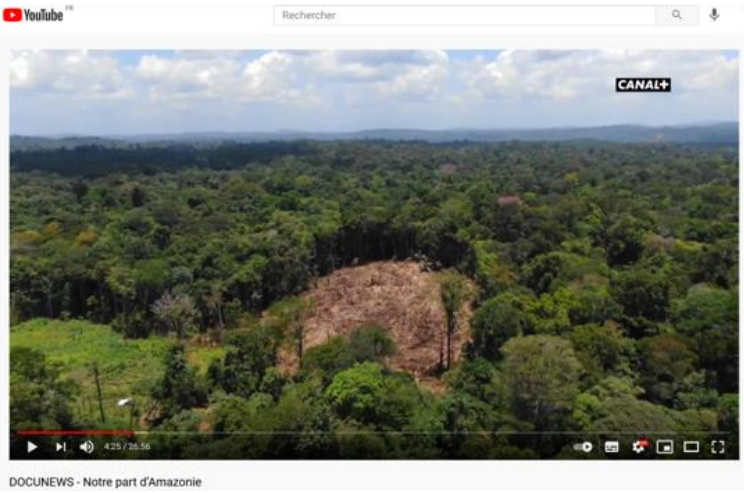


Figure 34 - Photos prises le 10/08/2019 lors de la campagne terrain du haut de la parcelle et du bas de la parcelle (image du bas). Le signal radar sur la parcelle est visible Figure 33. Cette parcelle apparaît juste après sa coupe dans le reportage Notre Part d'Amazonie.

Sur une autre parcelle, la Figure 35 montre une chute de signal plus précoce en VV qu'en VH en orbite ascendante et descendante. Juste déboisée, cette parcelle présentait une masse végétale au sol plus importante que la précédente (Figure 33-Figure 34), ce qui pourrait expliquer pourquoi le signal décroît plus lentement en polarisation VH. Les parcelles ayant une forte masse végétale au sol après le déboisement montrent souvent une décroissance progressive du signal. En effet, une forte masse végétale entraîne une réflexion diffuse et peut faire que le signal en VH reste élevé. Le signal décroit ensuite à mesure que cette masse végétale diminue et perd son humidité. Contrairement à la parcelle précédente, peu de troncs étaient encore debout, ce qui pourrait expliquer que le signal en VV descende plus. On note également que le signal remonte sur les deux polarisations à la fin de la saison sèche.

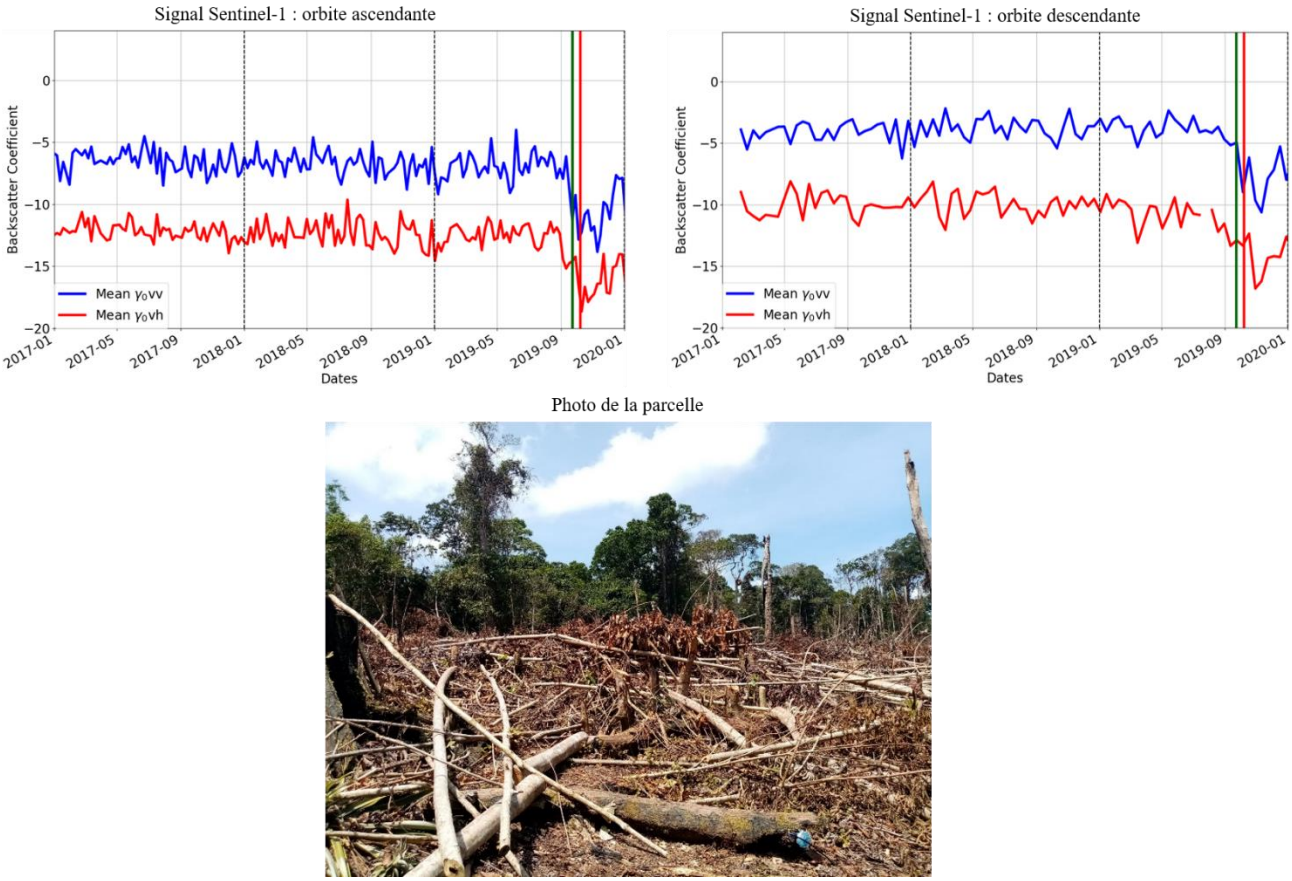


Figure 35 - Haut : signal Sentinel-1 rétrodiffusé en ascendant (gauche), et en descendant (droite). La ligne verticale verte représente un jour de visite sur la parcelle où le sous-bois avait commencé à être nettoyé (23/10/2019). La ligne verticale rouge représente un jour de visite sur la parcelle où les gros arbres étaient coupés (08/10/2019). Bas : photos prises de la parcelle le 10/08/2019 lors de la campagne terrain.

Dans certains cas, il arrive que le signal radar ne chute pas (ou très peu), comme la parcelle de la Figure 36. Celle-ci se trouve sur de fortes pentes (en moyenne 19°) et présente une forte masse végétale. En orbite ascendante, la baisse de signal est à peine perceptible, tandis qu'en orbite descendante, le signal montre une décroissance progressive.

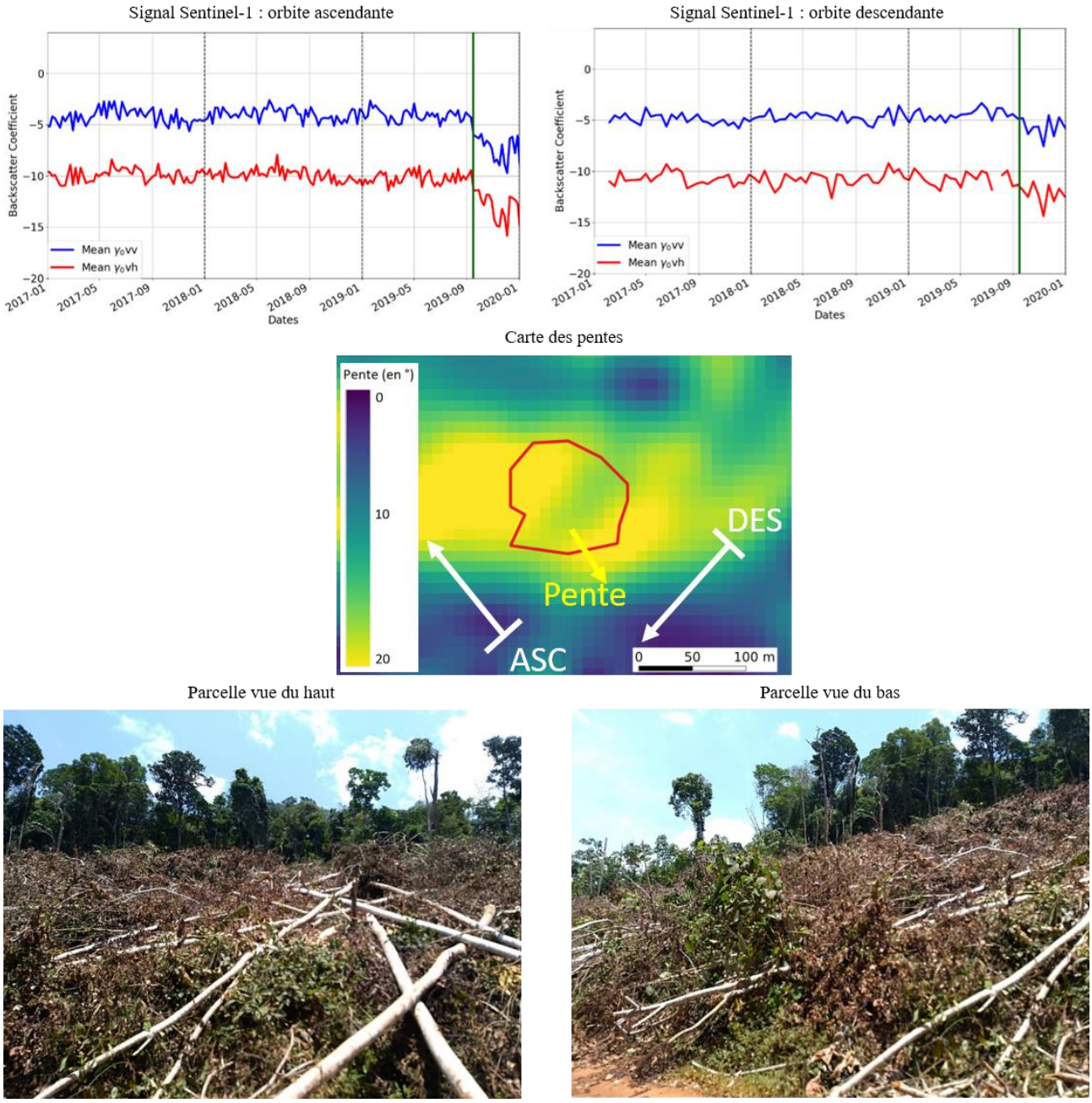


Figure 36 - Haut : signal Sentinel-1 rétrodiffusé en ascendant (gauche), et en descendant (droite). La ligne verticale verte représente un jour de visite sur la parcelle où la végétation avait commencé à sécher (16/09/2019). Milieu : image des pentes de la zone et direction des orbites ascendantes et descendantes. La flèche jaune représentant la pente pointe vers le point le plus bas. Bas : Photos prises du haut (gauche) et du bas (droite) de la parcelle le 16/09/2019 lors de la campagne terrain.

La majorité des méthodes de détection de la perte forestière se basant sur les images RSO font l'hypothèse que le signal diminue lorsqu'un déboisement se produit (Reiche et al., 2018a) ; (Lohberger et al., 2018). Ce n'est cependant pas toujours le cas, comme le montrent les exemples

ci-dessus. En effet, les effets environnementaux peuvent faire varier le signal : par exemple, des pluies peuvent entrainer une augmentation du signal. Pour contourner ce problème, la méthode de (Bouvet et al., 2018) se base sur la détection de nouvelles ombres dans l'image radar. Comme présenté dans le paragraphe 3.4, une ombre dans une image radar est une zone qui ne peut pas être atteinte par le signal radar en raison de la visée latérale du système (voir Figure 17). Cela arrive quand un objet, comme un arbre, crée un obstacle entre l'antenne radar et la zone observée. Ainsi, des ombres sont créées par les arbres aux lisières de forêt, selon la direction de visée du capteur : c'est un effet purement géométrique.

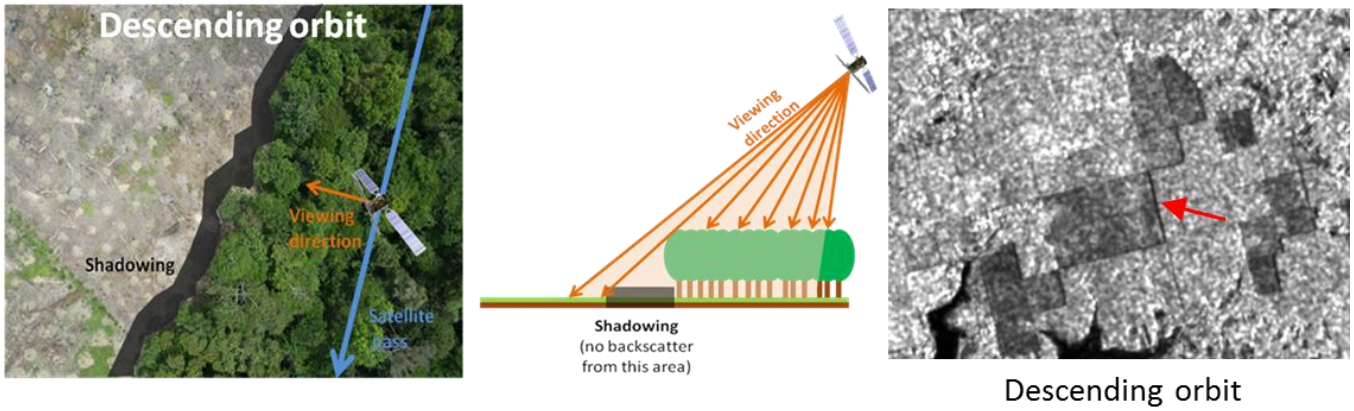


Figure 37 - Illustration de l'effet d'ombre du radar à synthèse d'ouverture (RSO) à la frontière entre les forêts et les zones déboisées, pour la configuration d'orbite descendante. (Image de fond arrière-plan : ConservationDrones.org/Orangutan Conservancy). (Bouvet et al., 2018).

Les ombres qui apparaissent sont caractérisées par une chute drastique de la rétrodiffusion dans la série temporelle du signal radar. Provenant d'un effet géométrique, cette diminution de la rétrodiffusion est persistante dans le temps et varie peu (car elle n'est pas sensible à l'humidité du sol, au statut post-perturbation, etc.) et les nouvelles ombres restent donc visibles pendant une longue période. Celles-ci sont facilement détectables sur des séries temporelles denses d'images radar, telles que celles de Sentinel-1.

Pour résumer cette section, nous avons vu que :

- **Le signal radar sur la forêt dense en Guyane est relativement constant ;**
- **L'analyse du signal radar sur les parcelles de référence a montré qu'un déboisement n'entraînait pas nécessairement une chute de signal ;**
- **Les ombres apparaissant en lisière des parcelles sont un indicateur fiable de déboisement (Bouvet et al., 2018).**

🔔 **Ce qu'il faut retenir**

Dans ce chapitre, nous avons vu que :

- Certaines données géographiques permettent de discriminer des causes de déboisement ;
- Le signal radar ne montre pas un motif régulier lors de déboisement.

🖈 **Ce que nous proposons**

Au regard des observations précédentes, nous nous baserons sur la méthode de détection de pertes forestières de Bouvet et al., 2018 qui se base sur un indicateur géométrique (et donc fiable) pour définir la présence d'un déboisement. Après des études préliminaires et améliorations, cette méthode sera appliquée et validée à l'échelle de la Guyane. Afin de déterminer les causes des déboisements détectés, nous proposerons une méthode basée sur un arbre de décision utilisant des données exogènes (disponibles à tout moment). Cette méthode sera également appliquée et validée à l'échelle de la Guyane.

Chapitre 5 - Méthodes et validation

Contenu

1. Détection du déboisement 96
1.1. Méthode initiale et choix des paramètres pour la Guyane 96
1.1.1. Choix des images 99
1.1.2. Prétraitements 105
1.1.3. Calcul du Radar Change Ratio 106
1.1.4. Détection des ombres 106
1.1.5. Détection des zones potentiellement déboisées 107
1.1.6. Discussion sur la résolution spatiale des produits recherchés 107
1.2. Améliorations apportées à la méthode initiale 109
1.2.1. Prise en compte de la dimension temporelle 109
1.2.2. Topographie 111
1.2.3. Indice de confiance 115
2. Méthode de validation et de comparaison de carte de déboisement 120
2.1. Méthode de validation spatiale 120
2.2. Méthode de comparaison temporelle 122
2.3. Données utilisées pour la validation 123
3. Caractérisation des déboisements 126

3.1. Idée générale de la méthode et concept généraux 126
3.1.1. Arbre de décision 126
3.1.2. Algorithme du Random Forest (RF) 129
3.2. Préparation : Définition d'indicateurs potentiels et création de la base d'apprentissage .. 131
3.3. Création du modèle 133
3.3.1. Importance des indicateurs 133
3.3.2. Nombre d'indicateurs et taille de l'arbre choisi 134
3.3.3. Arbre de décision choisi comme modèle pour la caractérisation des pertes forestières 134
3.4. Validation du modèle de caractérisation du déboisement 134

Ce chapitre décrit les méthodes développées pour le suivi des pertes forestières en quasi temps-réel sur la Guyane (méthode de détection et méthode de caractérisation du déboisement), ainsi que les méthodes de validation. Le choix et le développement de ces méthodes ont été motivés par les observations faites dans le chapitre 4.

1. Détection du déboisement

Cette partie présente la méthode de détection du déboisement en quasi temps-réel à partir des images radar Sentinel-1, introduite dans Bouvet et al., 2018, les paramètres choisis pour son utilisation sur la Guyane, ainsi que les améliorations qui lui ont été apportées dans le cadre de cette thèse.

1.1. METHODE INITIALE ET CHOIX DES PARAMETRES POUR LA GUYANE

L'analyse du signal radar sur plusieurs cas de parcelles déboisées (chapitre 4) a montré qu'en bande C, les déboisements n'étaient pas systématiquement caractérisés par une signature temporelle particulière (voir chapitre 4, section 2). Dans la plupart des cas, les zones déboisées montrent une baisse de signal, mais celle-ci peut être difficile à détecter et n'est pas nécessairement stable dans le temps. Se baser seulement sur cette baisse de signal pour détecter les déboisements entraînerait l'introduction de nombreuses fausses alarmes, car le signal sur la forêt peut connaître des variations similaires dues à des facteurs environnementaux ou à l'effet du speckle, même sur des données filtrées. Afin de contourner ce problème, la méthode de Bouvet et al., 2018 intègre la détection de

nouvelles ombres (effet expliqué dans la section 4 du chapitre 3) dans l'image radar. Cette méthode est composée de 3 étapes principales :

1- Détection des nouvelles ombres : c'est un indicateur fiable de déboisement, mais les ombres n'apparaissent qu'en lisière de forêt ;
2- Détection des zones potentiellement déboisées : caractérisées par des chutes de signal relativement faibles ;
3- Suppression des zones potentiellement déboisées qui ne sont pas intersectées par des ombres pour ne garder que les détections fiables, ce qui permet de limiter les fausses alarmes.

Les 2 premières étapes nécessitent une approche de détection de changement. De nombreuses études portant sur la détection de changement avec les images RSO sont basées sur le rapport des intensités entre deux dates (Mermoz and Le Toan, 2016, Lohberger et al., 2018, Joshi et al., 2015, Le Toan et al., 1997, Bouvet and Le Toan, 2011). Ce rapport se trouve être plus adapté aux caractéristiques statistiques des images RSO que la différence des intensités (Rignot and Van Zyl, 1993). La méthode de Bouvet et al., 2018 est basée sur l'utilisation du Radar Change Ratio (RCR) (Tanase et al., 2018) qui est un rapport entre la moyenne du signal pre- et post-perturbation. Le RCR a l'avantage d'être plus robuste aux variations de signal et moins sensible aux effets du speckle. Étant donnée une série temporelle à N dates, le changement entre la date d_i ($i \in (1,N)$) et d_{i+1} est mesuré par : $RCR_i = M_a/M_b$ où M_b est la moyenne du signal des X_b images avant la date d_i (inclus) et M_a est la moyenne du signal des X_a images après la date d_{i+1} (inclus) :

$$M_b = \frac{1}{X_b} \sum_{j=i-X_b+1}^{i} \gamma_j^0 \ et \ M_a = \frac{1}{X_a} \sum_{j=i+1}^{i+X_a} \gamma_j^0 \qquad 5.1$$

Avec γ_i^0 le signal d'un pixel à une date d_i.

Les ombres et les zones potentiellement déboisées sont détectées par seuillage sur cette mesure de changement, uniquement sur les zones forestières. Les ombres sont détectées avec un seuil α quand elles atteignent une surface minimale (Minimum Mapping Unit - MMU) de 4 pixels (0.04 ha), et les zones potentiellement déboisées avec un seuil β supérieur à α (donc moins restrictif) en atteignant un MMU de 20 pixels, comme détaillé et illustré ci-dessous. L'algorithme de détection complet est décrit Figure 38.

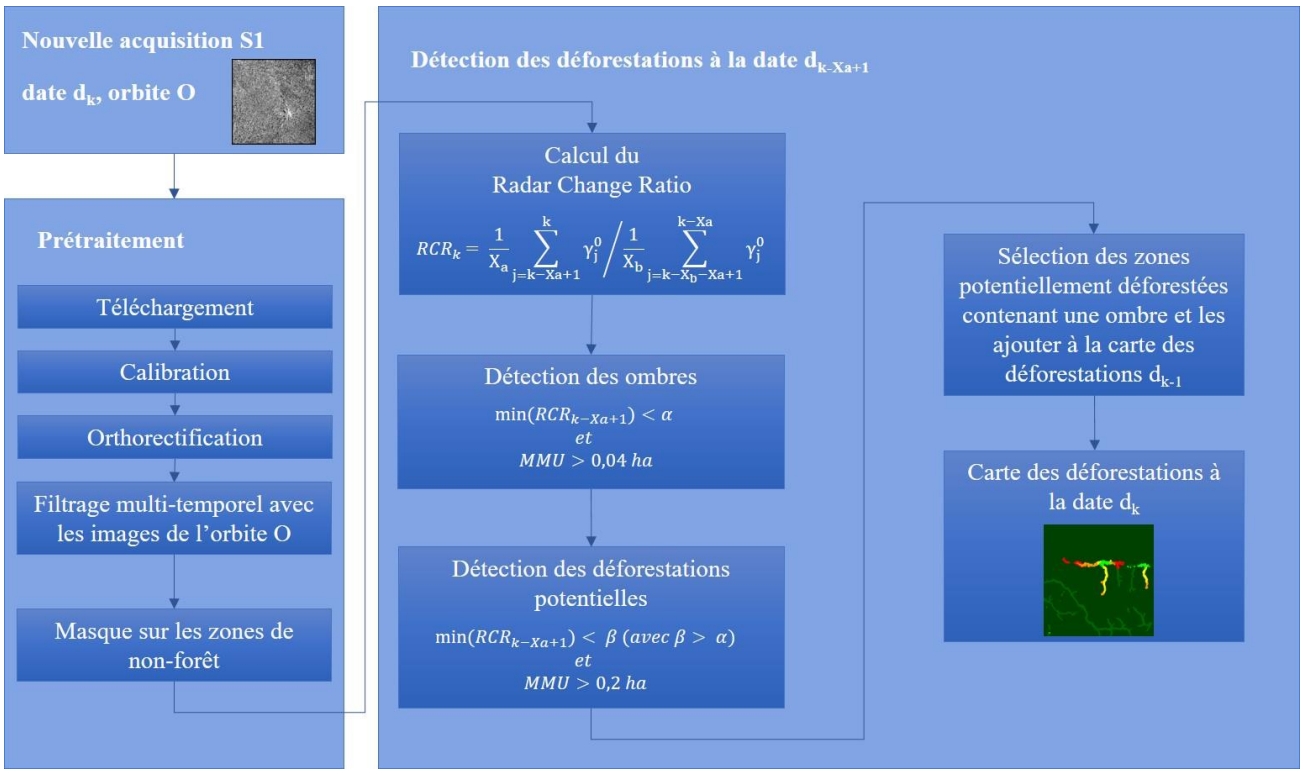


Figure 38 - Schéma décrivant l'algorithme de détection du déboisement produite à la date d_{k-Xa+1} à l'arrivée de l'image de la date d_k

La Figure 39 montre les effets de cet algorithme à l'échelle d'une parcelle. Sur le RCR, les baisses drastiques de signal induites par l'apparition d'une ombre en lisière de forêt sont trouvées avec un seuillage à -4.5 dB (en rouge sur la Figure 39-1).

Une zone déboisée entraîne souvent une légère baisse de signal sur l'ensemble de sa surface. Pour la trouver, un seuil moins contraignant est appliqué au RCR. Cependant ceci génère des fausses détections car une telle baisse de signal peut également venir d'autres facteurs, principalement environnementaux. La surface complète des zones potentiellement déboisées est donc obtenue, mais celles-ci sont peu fiables (en bleu sur la Figure 39-2).

En ne gardant que les potentielles zones déboisées qui contiennent une ombre, les fausses détections sont écartées et seules restent les surface réellement déboisées (en orange sur la Figure 39-3).

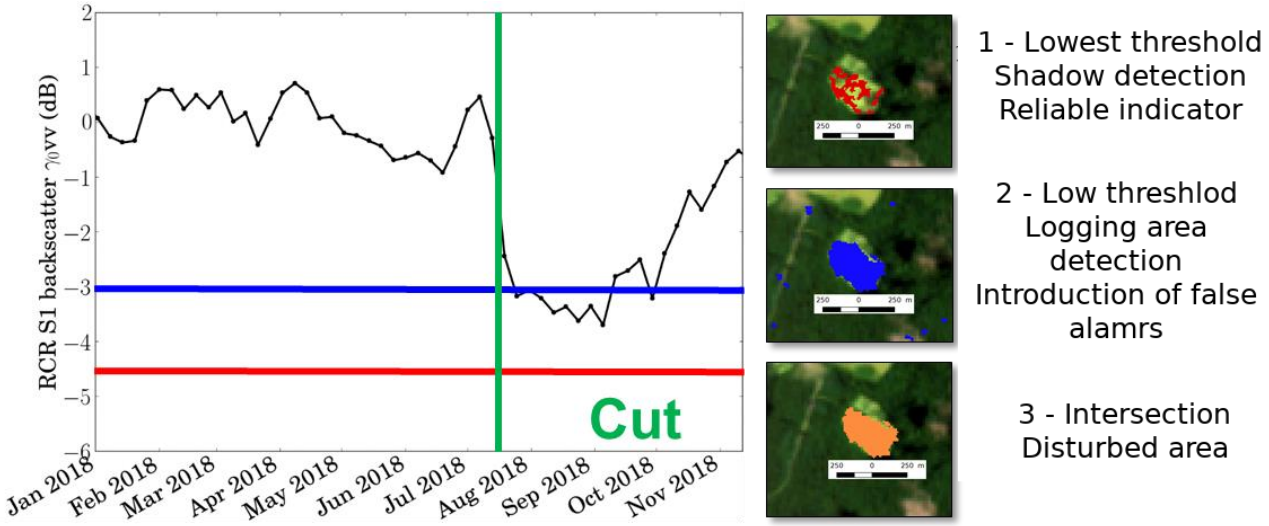


Figure 39 - Méthode de détection du déboisement basée sur Sentinel-1. A gauche, série temporelle du RCR d'un pixel d'une zone déboisée à laquelle les deux seuils sont appliqués. Dans cet exemple, ce pixel sera un déboisement potentiel. A droite, pixels détectés en utilisant deux seuils différents (rouge et bleu) et zone déboisée en orange (polygone bleu contenant un polygone rouge).

1.1.1. CHOIX DES IMAGES

1.1.1.1. POLARISATION

Comme vu dans le chapitre 3, les images Sentinel-1 sont disponibles sur les terres en polarisation VV et VH. En polarisation VV, la rétrodiffusion de surface est généralement plus élevée, et la polarisation VH discrimine d'avantage la rétrodiffusion volumique. La polarisation VH apparaît donc plus adaptée pour identifier une perte en volume (et donc des déboisements potentiels). Cependant, le signal sur les forêts en polarisation VV est plus fort que celui en VH, ce qui implique qu'à l'apparition d'une ombre (dans laquelle le signal est en théorie au niveau de l'équivalent de bruit sigma zéro), la baisse de signal en polarisation VV sera plus importante. Par ailleurs, il peut exister des ambiguïtés entre les forêts et les zones de non-forêt en VV, ce qui n'est pas gênant dans ce cas, l'algorithme n'étant appliqué que sur les zones de forêt grâce à l'utilisation d'un masque sur les zones de non-forêt (voir section 1.1.2).

La Figure 40 montre la fonction de densité de probabilité (PDF) de l'indicateur min(RCR) sur des données de perte de forêt (liées à l'orpaillage, l'exploitation forestière, l'urbanisation et l'agriculture) et des zones de forêt intacte. Ces données proviennent de photo-interprétation de mosaïques d'images Planet sur l'année 2020 et couvrent une surface cumulée de respectivement 433 et 64 338 ha (tuiles 22NCK et 22NCJ - se référer à la Figure 22, chapitre 3, section 4.2). Le terme "ombre" désigne les 20% de pixels de chaque parcelle ayant les valeurs min(RCR) les plus basses. En effet, nous posons l'hypothèse que les ombres apparaissant dans chaque parcelle représentent environ $^{1}/_{5}$ ème de leur surface, car celles-ci ne se trouvent qu'en lisière de parcelle déboisée.

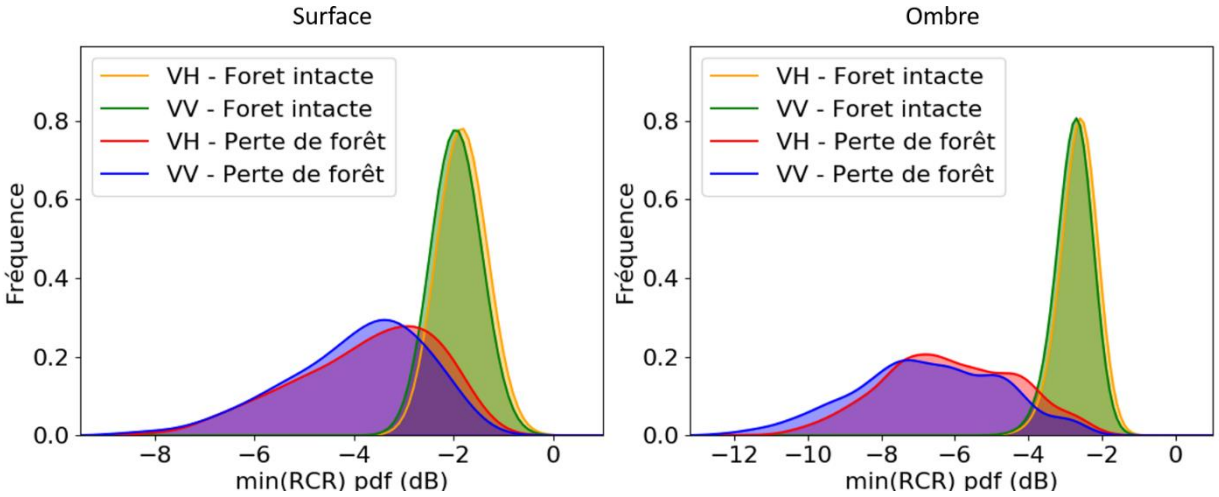


Figure 40 - Fonction de densité de probabilité (PDF) de l'indicateur min(RCR) en polarisations VV et VH pour des données de référence sur une surface totale de 64 771 ha (tuiles 22NCJ et 22NCK). Le terme "ombre" désigne les 20% de pixels de chaque parcelle ayant les valeurs min(RCR) les plus basses.

La Figure 40 montre :

- Une forte séparabilité du min(RCR) entre les forêts intactes et les zones de pertes forestières
- Que la séparabilité est plus forte avec la polarisation VV qu'avec la polarisation VH.

La Figure 40 montre également que sur la forêt intacte, le min(RCR) n'est pas égal à 0 dB. En effet, l'intensité au niveau du pixel peut varier au cours du temps (notamment en raison du speckle), ce qui peut entraîner des valeurs de RCR négatives.

L'aire sous la courbe (AUC) Receiver Operating Characteristics (ROC) décrivant la séparabilité du min(RCR) entre les forêts intactes et les zones de pertes forestières est de 0.93 en VV, contre 0.91 en VH (Figure 41), ce qui confirme ce qui est vu sur la Figure 40.

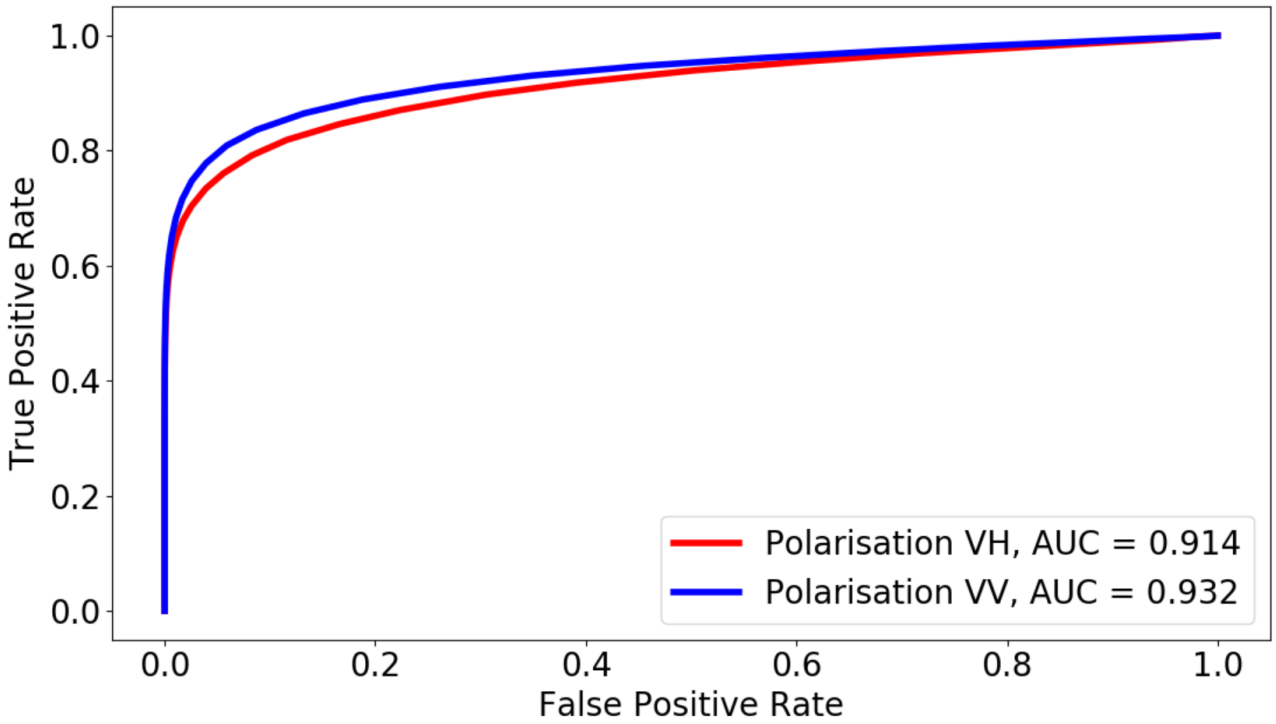


Figure 41 - Courbe ROC-AUC décrivant la séparabilité du min(RCR) entre les forêts intactes et les zones de pertes forestières en polarisation VH (courbe rouge) et VV (courbe bleue).

Cette étude succincte a montré que la polarisation VV était légèrement plus adaptée que la polarisation VH pour séparer le min(RCR) des forêts intactes et des zones de pertes forestières en Guyane. C'est pourquoi les images en polarisation VV sont utilisées dans le cadre de ces travaux. Cependant, une étude approfondie serait nécessaire pour connaître l'apport de la polarisation VH à la méthode développée ici. Des études sur d'autres sites (voir chapitre 7), ont montré que c'était la polarisation VH qui différenciait mieux les deux classes. Dans les deux cas, les résultats varient peu.

1.1.1.2. CAPTEUR

Deux satellites Sentinel-1 (S1A et S1B) sont actuellement en orbite. Comme le montre la Figure 42, il existe une légère variabilité entre le signal de S1A et S1B sur la forêt dense (zone de 38 589 ha), bien que ces acquisitions soient mesurées à la même heure. Même si les dates d'acquisitions des images S1A et S1B ne sont pas les mêmes, le signal de chaque capteur suit la même tendance au cours de l'année. La Figure 42 montre que :

- Cette variabilité est faible et inférieure à 0.87 dB (en prenant des acquisitions S1A et S1B consécutives)
- Il existe une légère saisonnalité du signal sur la forêt de Guyane, mais là encore elle est inférieure à 0.92 dB (différence maximum observée pour une même polarisation au cours d'une même année)

Utiliser les images provenant des deux plateformes permet de réduire la résolution temporelle à 6 jours. Nous avons donc décidé d'utiliser S1A et S1B conjointement, afin de réduire la résolution temporelle de nos détections, ce qui est crucial dans un contexte de détection rapide.

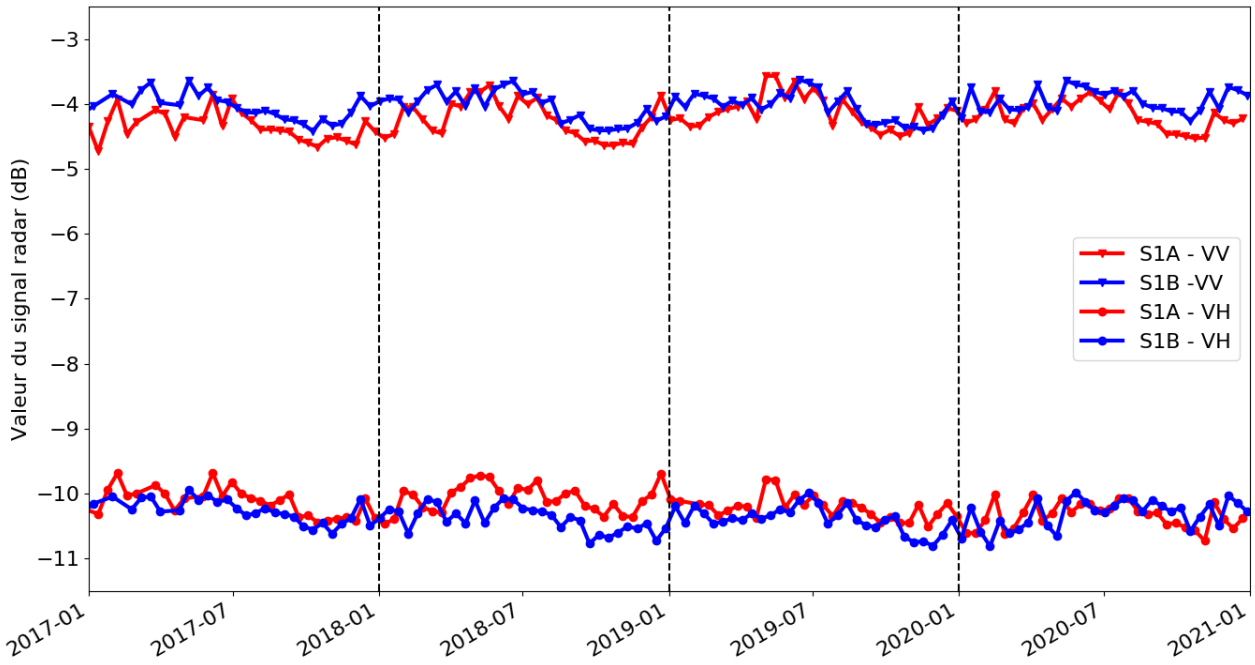


Figure 42 - Signal radar en VV et VH au cours de l'année 2017 pour les satellites S1A et S1B.

1.1.1.3. DIRECTION DE VISEE

Sur la Guyane, Sentinel-1 image en orbite ascendante (ASC), et descendante (DES). La direction de visée du radar est déterminante pour la détection des ombres. Comme le montre la Figure 43 :

- Suivant la configuration, une ombre peut apparaître seulement sur une orbite ;
- Les ombres les plus visibles sont celles qui sont parallèles à l'azimut (géométrie).

De plus, avoir plusieurs directions de visée est bénéfique pour augmenter la chance d'être dans une configuration où l'ombre est détectable, par exemple pour détecter avec une orbite donnée une zone qui se trouverait dans une ombre dans l'autre orbite. C'est pourquoi toutes les images disponibles en orbites ascendantes et descendantes sont utilisées ici.

Il est à noter que le RCR doit être calculé séparément pour chaque orbite, pour que les valeurs calculées soient cohérentes. Le min(RCR) sera, lui, calculé sur les séries temporelles, toutes orbites confondues.

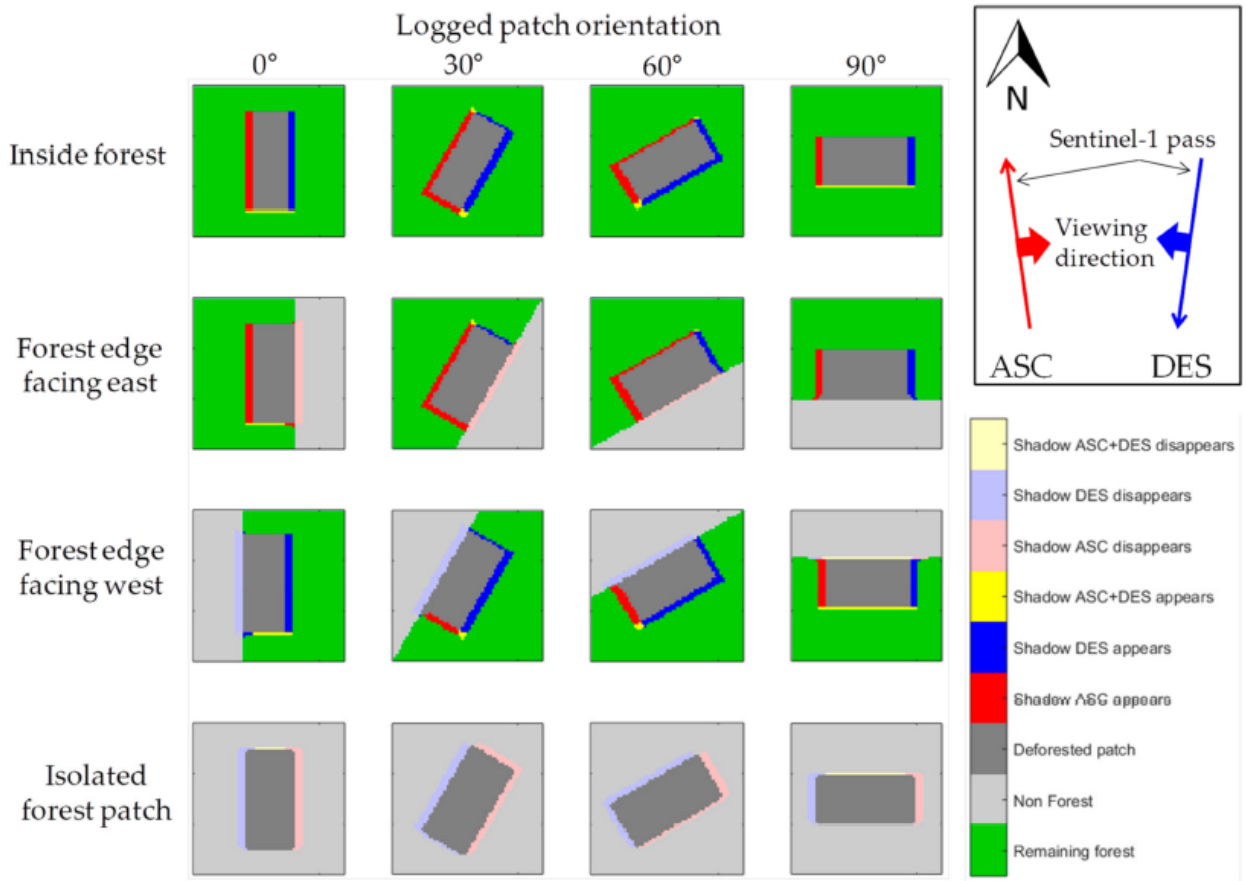


Figure 43 - Ombres qui apparaissent ou disparaissent lorsqu'une parcelle de forêt est déboisée, en fonction de l'orientation de l'orbite (ascendante ou descendante), de l'orientation de la parcelle, et de la configuration (parcelle à l'intérieur d'une forêt, à la lisière d'une forêt, ou isolée). La figure est simulée pour la configuration de Sentinel-1 (orientation vers la droite avec une inclinaison de l'orbite de 98 et 18) qui est résumée dans le cadre en haut à droite. (Bouvet et al., 2018)

1.1.1.4. ANGLE D'INCIDENCE

Au vu des passages des Sentinel-1 sur la Guyane, une cible peut appartenir à 1 ou 2 orbites dans une même direction de visée : une zone d'intersection existe entre 2 orbites parallèles consécutives (voir la Figure 20). Cette cible aura donc un angle d'incidence faible sur l'une des orbites, et élevé sur l'autre.

L'angle d'incidence θ influe sur la mesure du signal radar. En « far range » (angle d'incidence élevé), la résolution spatiale est plus fine qu'en « near range » (angle d'incidence faible) : il est donc possible de distinguer des cibles plus rapprochées. De plus, pour une configuration géométrique donnée, la taille des ombres est plus importante en « far range ». En effet :

$$largeur\ de\ l'ombre = \tan(\theta) \times Ht \qquad 5.2$$

Avec Ht la hauteur de la forêt (voir Figure 44).

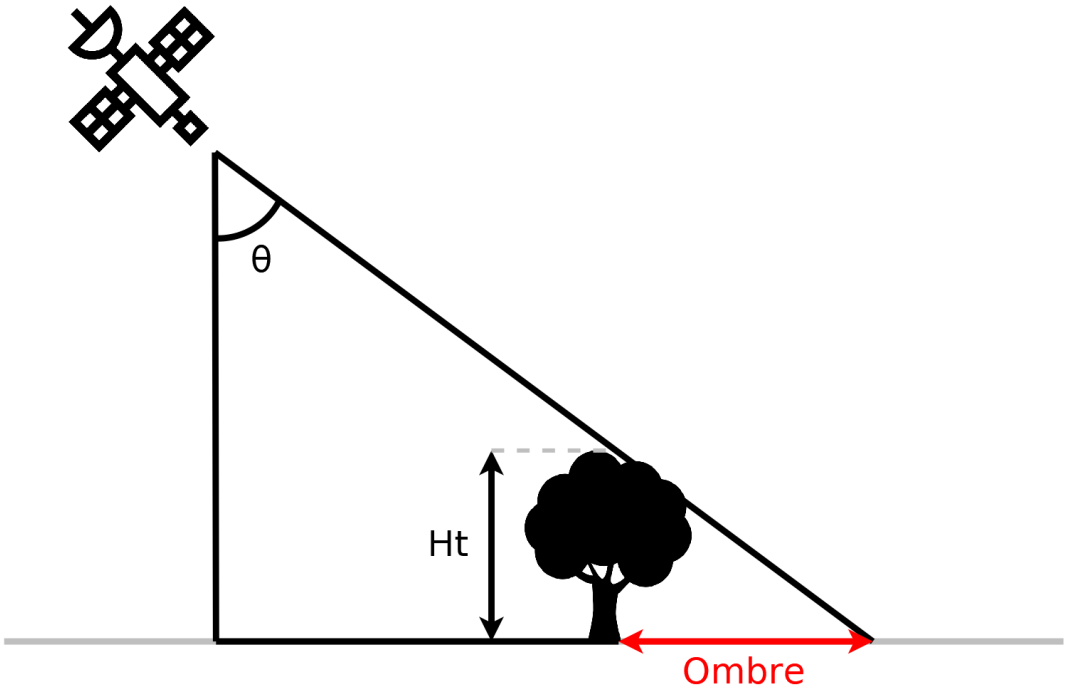


Figure 44 – Schéma de la relation entre l'angle d'incidence θ, la hauteur d'un arbre Ht et la taille d'une ombre projetée sur sol plat.

La largeur de l'ombre est indiquée sur la Figure 45 pour des d'angles d'incidence allant de 29° à 46° (angles d'incidence possibles pour Sentinel-1) et pour des hauteurs d'arbres comprises entre 0 et 40 mètres, dans le cas d'un terrain plat. Cela montre que les ombres sont plus grandes à angle d'incidence élevé. Quand deux orbites de même direction de visée se superposent, les pixels en « far range » sont donc privilégiés.

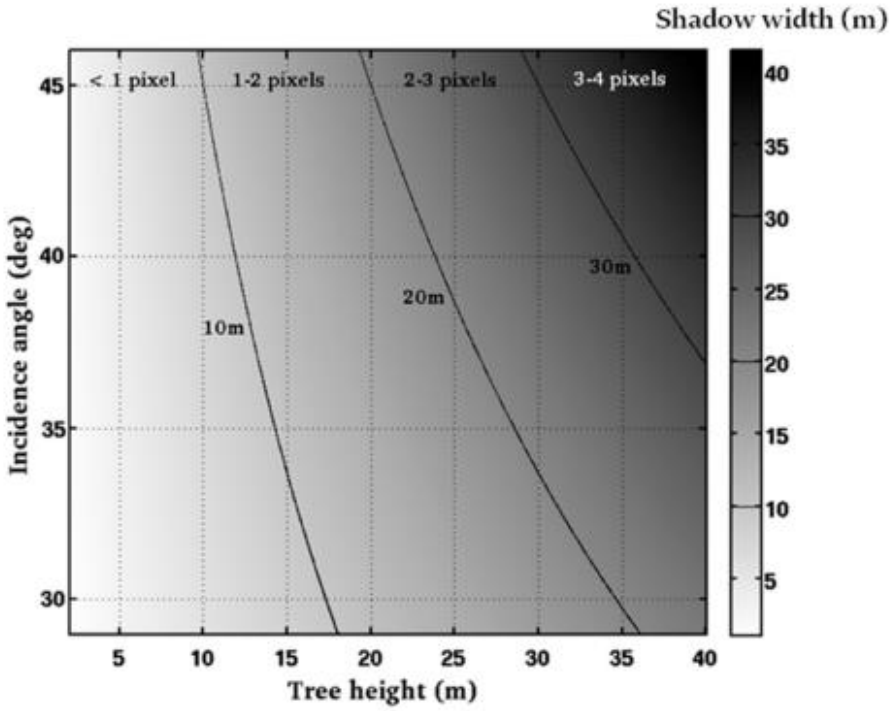


Figure 45 - Taille de l'ombre en fonction de la hauteur de la forêt et de l'angle d'incidence de Sentinel-1. Le nombre de pixels correspondant est basé sur une taille de pixel de 10 mètres. (Bouvet et al., 2018)

Pour récapituler les images retenues et traitées pour la détection rapide du déboisement en Guyane seront donc :

- **Les images en VV car elles montrent une séparabilité du min(RCR) entre les zones déboisées et les zones de forêt intacte légèrement meilleure qu'en VH ;**
- **Les images de S1A et S1B dans une même série temporelle afin d'obtenir une résolution temporelle plus fine ;**
- **Les orbites ASC et DES, afin de maximiser les chances de détection d'ombre ;**
- **Les séries temporelles en far range (quand deux orbites de même direction de visée se superposent) pour leur meilleure résolution et la taille plus importante des ombres.**

1.1.2. PRETRAITEMENTS

Une nouvelle image Sentinel-1 acquise à la date d_k, est d'abord téléchargée, calibrée, orthorectifiée et filtrée comme décrit dans la partie 4.3 du chapitre 3. Cette image est ensuite masquée pour ne garder que les zones de forêts. L'algorithme détecte les changements en se basant sur une hypothèse de forêt a priori et l'appliquer sur des zones de non-forêt pourrait introduire des fausses alarmes. Le masque utilisé est construit à partir de la carte d'occupation du sol produite dans le cadre d'ECOSEO (présentée chapitre 3, section 2.2.4), en agrégeant toutes les classes autres que « Forêt ».

1.1.3. CALCUL DU RADAR CHANGE RATIO

Le calcul du RCR dépend du choix de X_b et X_a (nombre d'images pre- et post-perturbation sur lesquelles la moyenne sera calculée). Plus X_b et X_a seront élevés, plus le RCR sera robuste (au speckle et aux variations exceptionnelles telles que de fortes pluies par exemple). X_b peut être choisi aussi élevé que possible. Toutes les images acquises avant la date considérée seront donc sélectionnées afin de réduire la sensibilité de la détection des changements aux effets environnementaux et du speckle affectant la rétrodiffusion.

En revanche, plus X_a sera élevé, plus le délai de détection sera important, réduisant l'aspect rapide de la détection. Un compromis entre robustesse et délai de détection est donc à trouver. X_a = 3 est suffisant en termes de réduction du speckle (Bouvet et al., 2018). Sur une orbite à 6 jours, X_a = 3 entraîne un délai de détection de 12 jours, passant à 24 jours dans les régions pour lesquelles les fréquences d'observations sont de 12 jours (voir figure 19 du chapitre 3). De plus, la Guyane connaît parfois des épisodes pluvieux locaux tellement importants qu'ils font chuter drastiquement le signal sur une date donnée (phénomène illustré dans le chapitre 6). Nous décidons donc de garder X_a = 3 qui permet à la méthode d'être robuste lors de ces épisodes. Le chapitre 6, abordera cependant la précision des détections en fonctions du X_a utilisé.

Le minimum du RCR est ensuite calculé sur la série temporelle des RCR.

1.1.4. DETECTION DES OMBRES

La détection des ombres se fait en appliquant un seuil α sur le min(RCR). En théorie, la valeur du signal radar sur une ombre devrait être au niveau de l'équivalent de bruit sigma zéro lié au bruit thermique de l'instrument (en dessous de -22 dB pour Sentinel-1). En pratique, ces pixels ne sont généralement pas purs. Étant contaminés par les zones autour de la lisière, leur valeur est plus élevée que l'équivalent de bruit sigma zéro mais reste largement inférieure aux valeurs de rétrodiffusion sur la forêt.

Pour définir automatiquement le seuil α qui sépare le mieux le min(RCR) des classes « déboisée » et « forêt intacte », il est possible d'utiliser les statistiques dérivées des échantillons de référence. Il a été montré que le seuil optimal pour séparer deux classes de ratio d'intensité moyen r_A et r_B ($r_A < r_B$) a une expression complexe qui implique des paramètres qui ne peuvent pas être estimés dans le cas général (comme la proportion de chaque classe). Cependant, celui-ci peut être approximé par $r_0 = \sqrt{r_A r_B}$ (calcul en valeur naturelle) (Bouvet et al., 2010). Dans notre cas, en approximant le min(RCR) par un rapport d'intensité et en définissant les ombres comme 20% des pixels de chaque parcelle ayant les valeurs min(RCR) les plus basses, on obtient r_A= -7.07 dB, r_B= -

2.78 dB, ce qui conduit à r_0= -4.9 dB (en utilisant les mêmes données que celles décrites dans le paragraphe 1.1.1 de ce chapitre).

Pour produire la carte sur la Guyane entière, nous décidons donc de fixer α = -4.5 dB (cf Figure 39), car il est proche du r_0 trouvé dans cette étude et des seuils trouvés dans des études similaires (Bouvet et al., 2018), (Hirschmugl et al., 2020) sur d'autres territoires avec une forêt dense (le but étant de converger vers des seuils globaux). Par ailleurs, dans le paragraphe 1.2.3, l'influence du seuil α choisi sur la précision des alertes sera étudiée (la détection des ombres étant l'indicateur fiable de la présence d'un déboisement). Nous avons décidé d'écarter les détections inférieures à 0.04 ha (4 pixels) afin d'éviter les fausses alarmes liées au speckle (voir paragraphe 1.1.6), car Sentinel-1 a une résolution de 20 mètres, ce qui couvre 2×2 pixels de 10 mètres.

1.1.5. DETECTION DES ZONES POTENTIELLEMENT DEBOISEES

Sur une zone déboisée, le signal dépend de plusieurs caractéristiques techniques (la polarisation, l'angle d'incidence) et environnementales (type de coupe, humidité du sol, la végétation restante au sol). En suivant la même logique que ci-dessus et en définissant l'intérieur des parcelles déboisées comme les 80% de pixels de chaque parcelle ayant les valeurs min(RCR) les plus hautes, on obtient r_A= -3.93 dB, r_B= -1.94 dB, ce qui conduit à r_0= -2.9 dB.

Pour produire la carte sur la Guyane entière, nous avons décidé de fixer β = -3 dB (cf Figure 39). Le choix du MMU pour les zones potentiellement déboisées a été décidé à 20 pixels (0.2 ha) (voir paragraphe 1.1.6).

1.1.6. DISCUSSION SUR LA RESOLUTION SPATIALE DES PRODUITS RECHERCHES

Pour les utilisateurs, la résolution spatiale idéale est souvent « la meilleure possible ». Cette section décrit la résolution à laquelle on peut s'attendre en se basant sur des produits Sentinel-1 IW-GRD pour détecter des déboisements.

Le nombre de vue équivalent (ENL) décrit le degré de moyennage d'une image radar dû à son traitement initial et caractérise la réduction du speckle. Un ENL important indique une faible variabilité des intensités due au speckle pour une réflectivité donnée de la scène. Il peut être évalué sur une surface homogène de l'image radar par (Li et al., 1983) :

$$ENL = \frac{\mu^2}{\sigma^2} \qquad 5.3$$

Avec μ l'intensité moyenne de la zone et σ son écart-type.

Dans notre cas, après traitement du speckle tel qu'appliqué dans l'algorithme, les images obtenues passent d'un nombre de vues estimé de 4.4 (ENL initial des images GRD) à 10. Ce

nombre peut sembler faible pour utiliser ces images dans des applications car cela pourrait introduire le dépassement des seuils α et β au niveau du pixel à cause du speckle. C'est pourquoi la détection est faite au niveau d'un agrégat de pixels et non au pixel lui-même, d'où l'utilisation de MMU dans l'algorithme.

En imagerie radar, il est difficile d'évaluer la résolution spatiale avec laquelle on travaille après avoir appliqué des traitements. La résolution des produits GRD utilisées pour cette thèse est approximativement 20×22 mètres. Ces cellules de résolutions sont échantillonnées pour former des pixels de 10 mètres. L'application d'un filtrage de speckle utilisant une fenêtre de 3×3 pixels et une pile temporelle dégrade donc cette résolution. On peut essayer d'estimer la résolution finale après filtrage. Avec une fenêtre de filtrage de 3×3 pixels (qui représente une fenêtre de filtrage d'environ 1.5×1.35 cellule de résolution), on peut penser que les données filtrées ont une résolution spatiale d'environ 30 mètres. Il est donc possible d'imaginer qu'un déboisement soit détectable à partir de 0.09 ha. Cependant, définir une surface minimale de détection (en absolu) est compliqué car un déboisement montrant un fort contraste sur une petite surface sera aussi visible qu'un déboisement entraînant un faible contraste mais sur une surface plus grande. Les observations faites en appliquant la méthode montrent qu'en dessous de 0.2 ha, la plupart des détections sont des fausses alarmes. C'est pourquoi nous avons défini le MMU à 0.2 ha.

Pour résumer, les paramètres retenus dans l'algorithme sont :

- **X_b = N – X_a avec N le nombre d'images de la série temporelle, et X_a le nombre d'images sélectionnées pour calculer les valeurs post-perturbation ;**
- **X_a = 3 pour la carte produite à l'échelle de la Guyane. Ceci permet d'avoir un délai de 12 jours pour les déboisements détectés sur une série temporelle provenant d'une orbite en ascendant et 24 jours en descendant. La précision en fonction de X_a sera étudiée dans le chapitre 6 ;**
- **α = -4.5 dB pour la carte produite à l'échelle de la Guyane. La section 1.2.3 présentera un indice de confiance basé sur la valeur de α.**
- **β = -3 dB pour la carte produite à l'échelle de la Guyane ;**
- **Un MMU de 4 pixels pour la détection des ombres ;**
- **Un MMU de 20 pixels (0.2 ha) pour les surfaces déboisées.**

1.2. AMELIORATIONS APPORTEES A LA METHODE INITIALE

La méthode initiale décrite dans Bouvet et al., 2018 a été améliorée au cours de la thèse. Ces améliorations sont décrites dans la section suivante.

1.2.1. PRISE EN COMPTE DE LA DIMENSION TEMPORELLE

L'identification et l'analyse des fausses alarmes produites par l'algorithme initial a permis de mettre en lumière un artefact lié au calcul du min(RCR). Dans l'algorithme initial, le minimum du RCR était calculé sur la totalité de sa série temporelle. L'augmentation de cette série temporelle peut mener à un agrégat de pixels dont la valeur de min(RCR) passe en dessous du seuil α à des dates différentes, notamment à cause du speckle (voir paragraphe 1.1.6). Si cet agrégat atteint une taille minimum de 4 pixels, il sera identifié comme « ombre » et une fausse détection en résultera. La Figure 46 est un exemple de fausse détection due à ce problème. La série temporelle du RCR des pixels formant « l'ombre » détectée est montrée Figure 47.

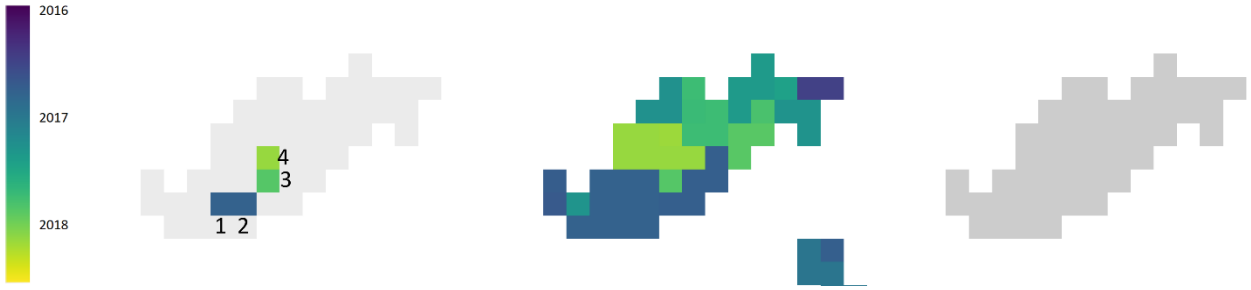


Figure 46 – Détection d'une fausse alarme due à un agrégat de pixels dont le min(RCR) passe en dessous du seuil α à différentes dates. La figure à gauche représente les pixels formant l'ombre détectées (min(RCR)<α). La série temporelle du RCR de chaque pixel formant l'ombre est disponible dans la Figure 47. La couleur représente la date et le fond gris représente l'emprise de la détection finale. La figure du milieu représente les pixels formant la zone potentiellement déboisée (min(RCR)<β). La détection finale (la fausse alarme) est représentée sur la figure de droite.

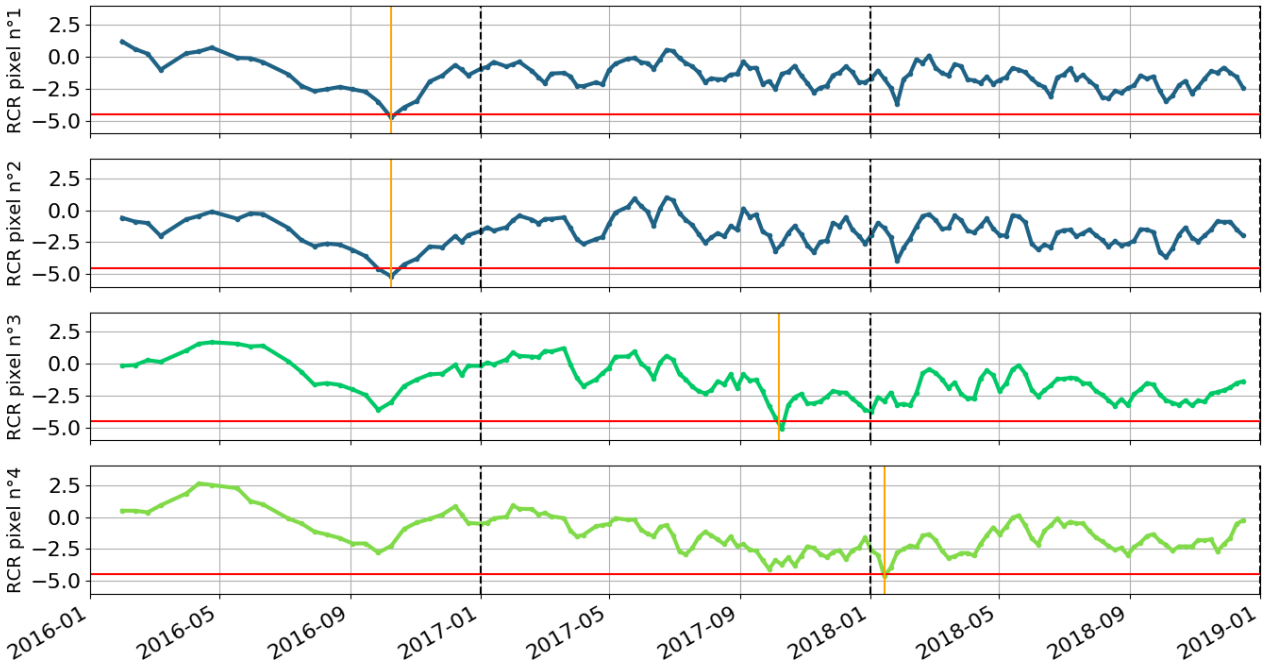


Figure 47 – RCR calculé sur chaque pixel du cas d'étude de la Figure 48 au cours du temps. Les lignes verticales oranges représentent le min(RCR) de la série et la ligne verticale rouge représente le seuil α=-4.5 dB.

La Figure 47 montre que chaque pixel est passé en dessous du seuil α de façon indépendante, et que ces pixels ne devraient pas être associés.

En théorie, la solution pourrait être d'agréger seulement les pixels dont la date du min(RCR) est identique, mais en pratique, les pixels d'une « vraie ombre » peuvent passer le seuil α à des dates proches mais différentes (l'ombre pouvant augmenter à mesure de la perturbation).

Idéalement, l'agrégat devrait se faire avec les pixels dont le min(RCR) est inférieur à α sur une fenêtre glissante des 6 derniers mois (durée définie de manière empirique suite aux observations). En pratique, la mise en place d'une telle méthode à grande échelle nécessite une grande capacité de stockage et pour le minimiser, seul le minimum global du RCR de la série est stocké. Il a donc été décidé d'agréger les pixels si la date de leur min(RCR) était inférieure à 6 mois. Cependant, cette solution comporte toujours une faille : le min(RCR) est le minimum global de la série et il n'est pas forcément égal au minimum local de la fenêtre glissante des 6 derniers mois. Si pour une certaine raison le RCR d'un pixel est à -6 dB en 2016, et qu'il passe à -5 dB en 2018 lors d'un déboisement (dans lequel il serait dans l'ombre), ce pixel ne sera, à tort, pas comptabilisé. Ce problème a encore peu été rencontré en pratique, mais il est à prendre en compte pour le développement opérationnel de la méthode sur de nombreuses années.

1.2.2. TOPOGRAPHIE

Dans (Bouvet et al., 2018), le territoire étudié était localisé sur le côté Est des Andes au Pérou où la topographie est relativement faible, ne nécessitant pas de prendre en compte les effets topographiques. En revanche, en Guyane, 12% du territoire a une pente supérieure à 15°.

Sur un terrain de pente ρ, la largeur d'une ombre est définie par :

$$largeur\ de\ l'ombre = \tan(\theta) \times \mathrm{Ht} \times \frac{\cos(\theta)}{\cos(\theta - \rho)} \qquad 5.4$$

Avec θ l'angle d'incidence et Ht la hauteur de la forêt (voir Figure 49).

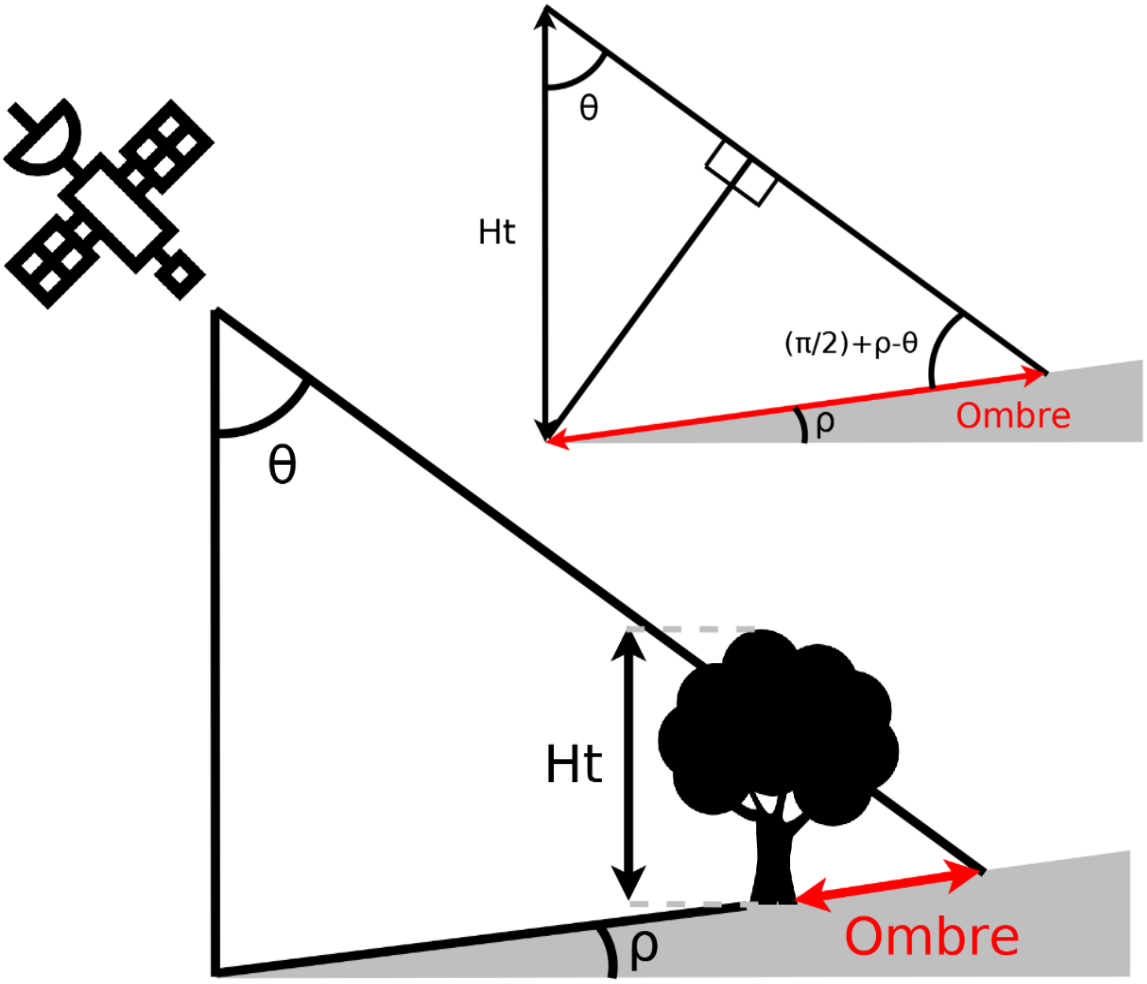


Figure 49 - Schéma de la relation entre l'angle d'incidence θ, la hauteur d'un arbre Ht et la taille d'une ombre projetée sur sol avec une pente ρ.

En prenant une taille moyenne de forêt à 35 mètres et avec l'angle d'incidence de Sentinel-1 variant entre 29° et 46°, la largeur de l'ombre est de 17 mètres minimum (pour ρ = 30° et θ = 29°).

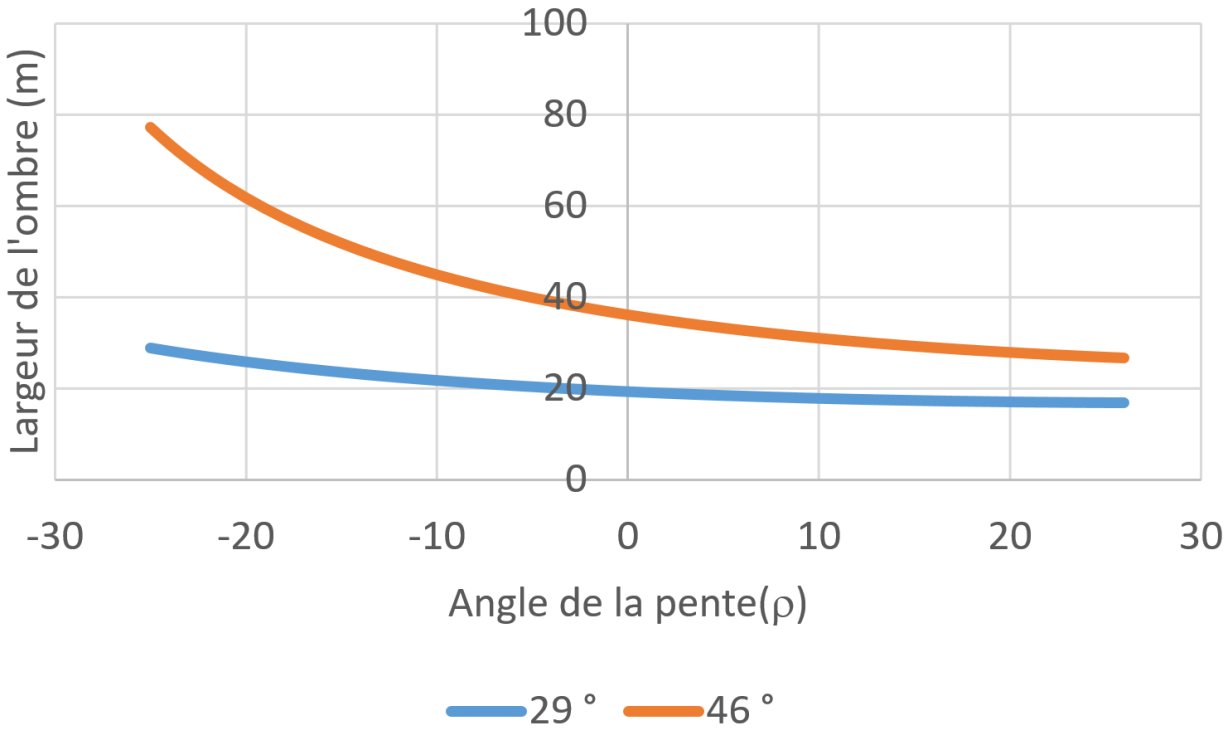


Figure 50 - Largeur de l'ombre radar (en mètres) en fonction de la pente au sol pour des angles d'incidence de 29° (bleu) et 46° (orange), avec une hauteur de forêt fixée à 35 mètres.

La topographie influe également sur la cellule de résolution au sol. Pour une pente ρ, la résolution sol en distance sera donnée par (adaptation de la formule 3.5) :

$$R_{d,s} = \frac{c}{2B\sin(\theta - \rho)} \qquad 5.5$$

La Figure 51 montre l'influence d'une pente ρ pour les données GRD de Sentinel-1 (en prenant $\frac{c}{2B\sin(\theta)}$ = 20 mètres pour θ = 37.5° (moyenne des angles d'incidence)).

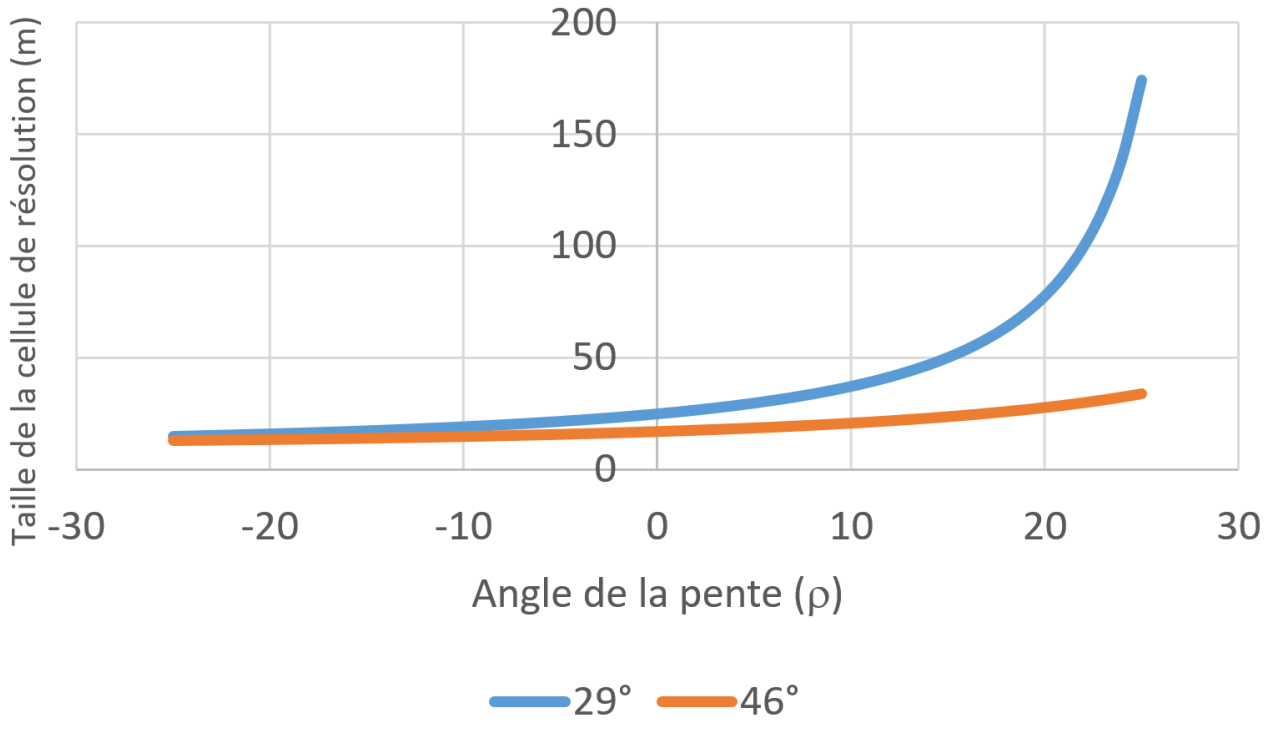


Figure 51 – Taille de la cellule de résolution (en mètres) en fonction de la pente au sol pour des angles d'incidence de 29° (bleu) et 46° (orange).

La cellule de résolution augmente fortement quand la pente augmente (en particulier après 10° pour un angle d'incidence de 29°), augmentant également la radiométrie. En revanche, à 10°, la pente à peu d'effet sur la taille de l'ombre. La proportion d'ombre dans une cellule de résolution devient donc plus faible à mesure que la pente augmente, ce qui complique sa détection.

La Figure 52 montre que la concentration de fausses alarmes augmente à mesure que la pente augmente, de la même façon que la taille de la cellule de résolution. Cette étude a été réalisée sur une zone de 99 194 ha (sur la tuile 22NCL, voir la Figure 22 du chapitre 3), et compte la proportion de fausses alarmes (au pixel) après 33 mois de détections cumulées (de janvier 2016 à septembre 2018) calculées avec la méthode initiale.

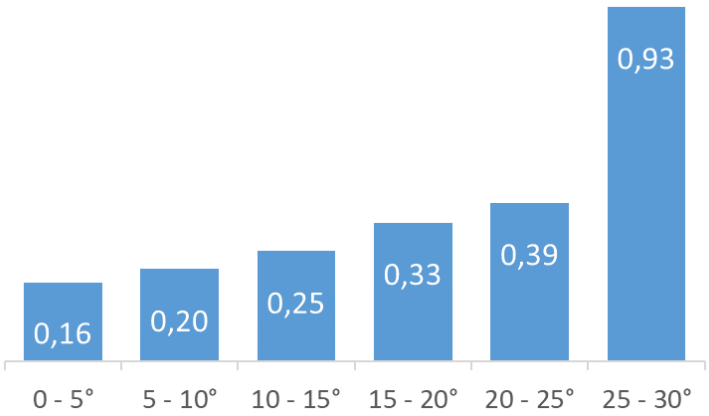


Figure 52 - Pourcentage de fausses détections en fonction de la topographie pour des plages de pentes données

Pour éviter les fausses alarmes liées à une fausse détection d'ombre (indicateur de fiabilité), seules les détections d'ombre situées dans une zone dont la pente est inférieure à 15° sont gardées.

Etant donné que la méthode est basée sur des seuils appliqués sur un rapport de rétrodiffusion (RCR) et non directement sur la valeur de rétrodiffusion elle-même, la topographie aura peu d'impact sur la détection d'une surface déboisée (en omettant la taille de la zone). C'est pourquoi, le seuil sur la pente n'est pas appliqué pour la détection de zones de déboisements potentiels, afin d'éviter certains artefacts. Cela permet de conserver les déboisements qui commenceraient au bas d'une colline et dont l'ombre aurait été détectée de manière fiable, comme le montre la Figure 53. Appliquer un masque sur les zones potentiellement déboisées ou sur les détections a posteriori pourrait supprimer une partie d'une zone déboisée s'étendant sur une pente.

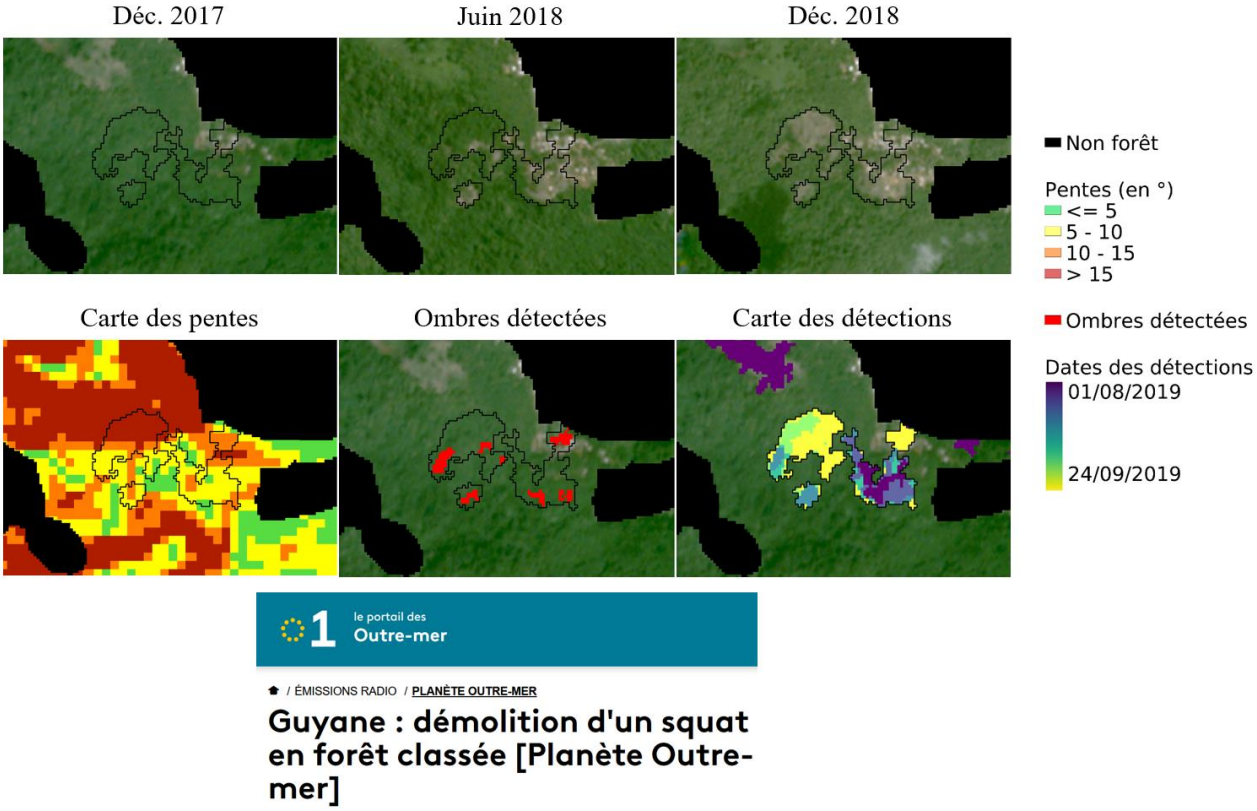


Figure 53 - Détection d'un déboisement dans une zone de forte topographie. Les images du haut montrent l'évolution du déboisement mis en évidence par les contours noirs. Les fonds de carte sont des mosaïques biannuelles sans nuages d'images Planet (www.planet.com/nicfi). Les ombres (image bas milieu) ont toutes été détectées dans une pente < 15° (image bas gauche). Il est possible d'observer la temporalité de ce déboisement qui commence par la partie inférieure puis remonte (image bas droite), comme le montrent également les images optiques. Ce déboisement est celui du mont Fortuné et est due à la création de logements illégaux (Guyane : démolition d'un squat en forêt classée[3]).

Enfin, le fait d'avoir des orbites ascendantes et descendantes permet d'avoir deux vues d'un pixel. Un pixel situé dans une ombre « fixe » de l'image radar dans une direction d'orbite, sera visible dans l'autre direction d'orbite.

[3] https://la1ere.francetvinfo.fr/emissions-radio/planete-outre-mer/guyane-demolition-squat-foret-classee-planete-outre-mer-756199.html

1.2.3. INDICE DE CONFIANCE

Les différents utilisateurs de cartes de déboisement peuvent avoir des besoins différents. Certains voudront connaître toutes les parcelles déboisées, privilégiant les détections et favorisant ainsi les fausses alarmes (si par exemple la vérification par un expert via photo-interprétation optique est possible, si la carte est croisée avec d'autres résultats, ou si l'aire à surveiller est limitée), et d'autres voudront pouvoir être sûrs de la détection, minimisant ainsi les fausses alarmes au détriment des détections (dans le cas par exemple, de grands territoires tels que l'ensemble de la Guyane où une détection déclencherait une intervention sur le terrain sans vérification possible au préalable). Afin de répondre à ces différents besoins, il est intéressant d'ajouter un indice de confiance aux cartes produites.

L'indicateur de fiabilité de la méthode de détection est l'apparition d'ombres. Celles-ci sont détectées par seuillage du RCR. Plus le seuil utilisé sera bas, moins il y aura de détections, mais meilleure sera leur précision. C'est pour cela que l'indice de confiance proposé se base sur les valeurs des pixels constituant les ombres. Le but est de définir une fonction approximant la probabilité qu'une détection soit bonne en fonction de la valeur des pixels de son ombre.

Pour arriver à cet objectif, 5 étapes ont été nécessaires :

- Produire des cartes de déboisement à différents seuils ;
- Définir une zone d'étude sur laquelle aucun changement n'a eu lieu de façon certaine pour connaître les taux de fausses alarmes en fonction des seuils ;
- Définir la probabilité qu'une détection soit correcte à chaque seuil, pour une zone « type » en Guyane ;
- Trouver une fonction approximant cette probabilité par seuil ;
- Définir un indice de confiance en fonction des valeurs contenues dans les pixels « ombres » d'une détection.

Des cartes de déboisement basées sur la méthode ont été générées à chaque acquisition Sentinel-1 sur l'année 2017, avec des valeurs de seuil α différentes. Les seuils appliqués variaient entre -4 et -8 dB avec un pas de 0.2 dB, pour un total de 21 seuils.

La zone d'étude (cumul des zones où il n'y a pas de changement) a été définie en utilisant les données d'occupation du sol décrites dans le chapitre 3, et la photo-interprétation de mosaïques biannuelles Planet (www.planet.com/nicfi). Ces zones de forêt intacte sont non-connexes et ont été

sélectionnées à plusieurs endroits (Figure 54) pour une superficie cumulée de 35 864 ha sur la tuile 21NZE (voir Figure 22 du chapitre 3).

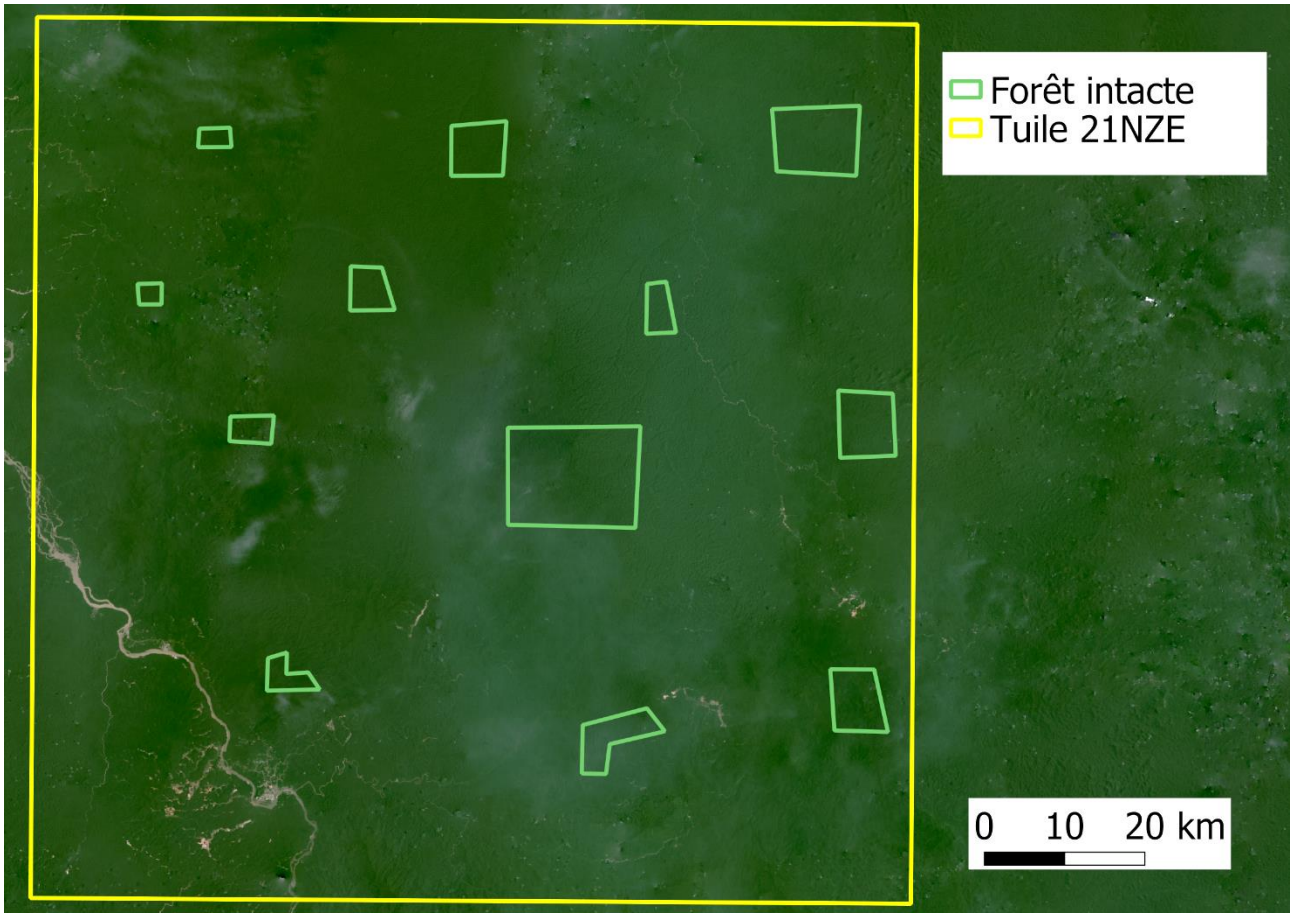


Figure 54 - Les polygones verts représentent la zone de forêt intacte sur laquelle les fausses détections sont comptabilisées.

Les zones définies ont permis de calculer le nombre de fausses alarmes sur une zone de 35 864 ha sur une année (nombre de détections dans une zone qui n'a pas été déboisée) en fonction des différents seuils α. En théorie, le nombre de fausses alarmes est constant sur une forêt « homogène » et varie seulement en fonction de ce seuil α.

Le nombre total de détections $Tot(\alpha) = FA(\alpha) + BD(\alpha)$ (avec $BD(\alpha)$ le nombre de bonnes détections pour un seuil α) dépend des déboisements en eux-mêmes (s'il n'y a pas de déboisement, il n'y aura pas de bonnes détections à fortiori), et donc de la zone et de la période de l'année. Afin d'obtenir que la probabilité P qu'une détection soit correcte dépende seulement du nombre de fausses alarmes, nous fixons $BD(\alpha) = BD$, avec BD constant pour tout seuil. Pour fixer la valeur de BD, nous avons pris la valeur moyenne des surfaces déboisées détectées avec la méthode entre les années 2016 et 2018 (4 073 ha) sur la Guyane (8 380 000 ha), et nous l'avons rapporté à la surface de la zone d'étude (35 864 ha), ce qui donne BD = 17.4 ha pour une année sur la zone d'étude.

La probabilité P qu'une détection soit correcte pour un seuil α donné a été définie telle que :

$$P(\alpha) = 1 - \frac{FA(\alpha)}{Tot(\alpha)} \qquad 5.6$$

La Figure 55 montre la proportion de bonnes détections en fonction du seuil utilisé. Il est à noter qu'aucune détection n'est faite dans la zone de forêt intacte avec un seuil en dessous de -5.6 dB.

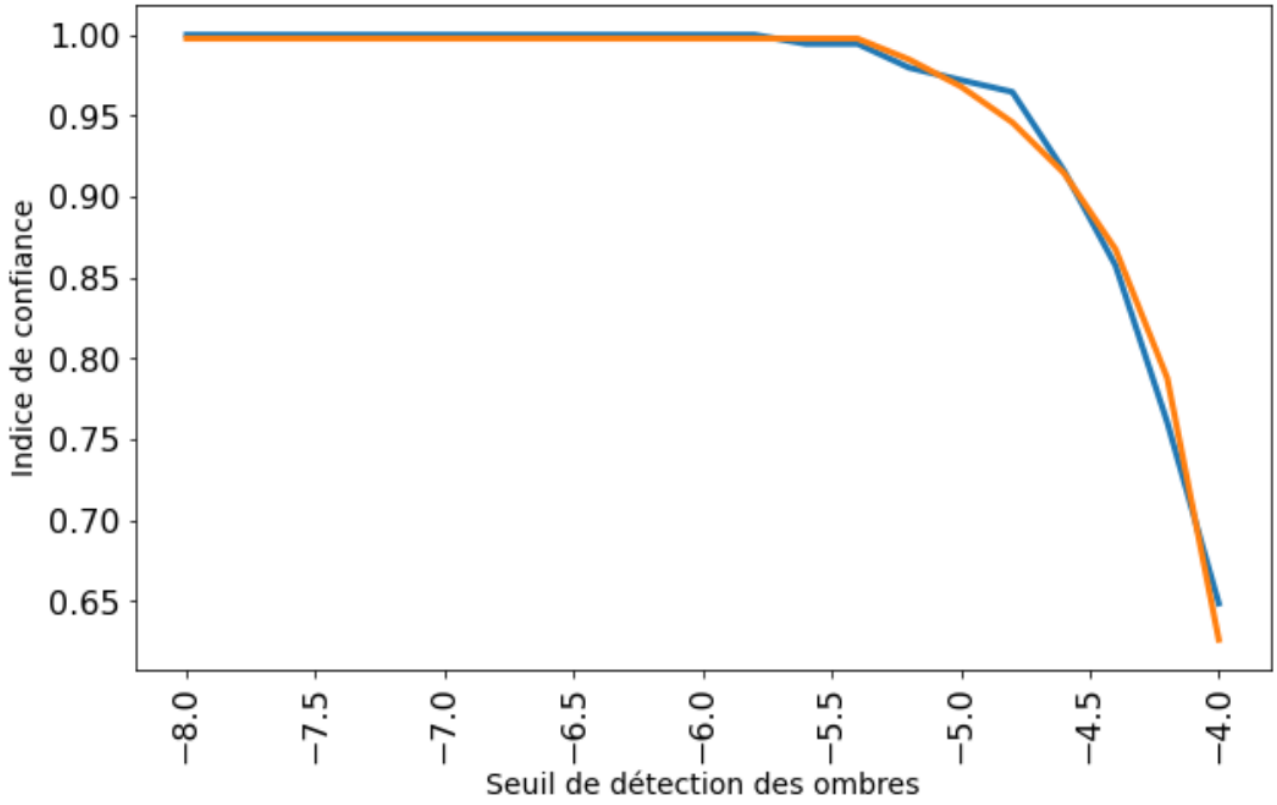


Figure 55 - Indice de confiance $P(\alpha)$ en fonction du seuil de détection des ombres α. La courbe bleue est la courbe trouvée de façon expérimentale. La courbe orange est la fonction d'approximation.

L'approximation faite à partir de ces observations donne une fonction F (courbe orange sur la Figure 55) telle que :

$$\mathrm{F}(x) = \begin{cases} 1.1 + \dfrac{1}{5.4725x + 19.78} \; pour \; x > -5.4 \\ 1 \; sinon \end{cases} \qquad 5.7$$

Pour finir, il faut définir le seuil minimal à partir duquel la détection (et donc l'ombre) serait apparue (nommé ici SMD). La sélection d'une ombre dépend de la valeur de min(RCR) de ses pixels, d'une composante spatiale (MMU = 4 pixels), et d'une composante temporelle. Dans un cadre idéal, la recherche du SMD pourrait être défini par une recherche sur chaque pixel de la manière suivante :

Soit P = [p1, p2, …, pn] l'ensemble des pixels de l'ombre et pk = (vk, dk) avec vk la valeur de min(RCR) du pixel pk et dk la date associée à vk.

Trier P dans l'ordre croissant des vk

Initialiser i = MMU = 4 et SMD = no data

Tant que SMD == no data :

Trouver la plus petite composante connexe C parmi {pj (j ∈ 1, i) | $abs(dj - di) \leq 183$} # Correspond à la fenêtre de 6 mois (183 jours) pour l'agrégation des pixels formant une ombre

Si taille(C) >= 4, SMD = vi, sinon i = i+1

Pour des raisons de temps de calcul, cette recherche est approximée par :

$$g(P) = SMD\ tel\ que \begin{cases} SMD = percentile_{50}(P)\ si\ nombre\ d'éléments\ de\ P < 9 \\ SMD = percentile_{25}(P)\ si\ nombre\ d'éléments\ de\ P < 20 \\ SMD = percentile_{12}(P)\ sinon \end{cases} \quad 5.8$$

Ces valeurs définies de façon expérimentale visent à se rapprocher de la valeur réelle de SDM, en se basant sur l'hypothèse que la proportion des pixels effectivement associés à une ombre dans P diminue quand la taille de P augmente. En effet, un pixel sur lequel un déboisement a entraîné un sol nu peut avoir un signal (et donc un min(RCR)) très faible et donc être inclut dans P, sans qu'il ne résulte d'un effet d'ombre.

L'indice de confiance d'une détection sera donc défini au niveau objet par F(SMD) donnée par l'équation 5.7. Cet indice, contrairement aux détections en soit, est évolutif. Le niveau de confiance d'un pixel défini à 80% pourra augmenter si les caractéristiques incluses dans P y contribuent.

Cette amélioration ayant été développée à la fin de la thèse dans le cadre d'un projet (TropiSCO[4]), les résultats qui en découlent sur la Guyane ne sont pas disponibles, car rejouer la détection des déboisements sur plusieurs années demande de grandes ressources en calcul et en stockage.

[4] https://www.spaceclimateobservatory.org/fr/tropisco-amazonie

Pour résumer, les améliorations apportées à la méthode initiale sont :

- **La meilleure prise en compte de la dimension temporelle ;**
- **L'intégration de la topographie ;**
- **La définition d'un indice de confiance pour les détections.**

2. Méthode de validation et de comparaison de carte de déboisement

La méthode définie ci-dessus a été appliquée à l'échelle de la Guyane. Les mesures utilisées couramment pour évaluer les performances d'une carte de changement sont rappelées en annexe (1). Cette partie présente la méthode choisie pour valider les cartes de déboisement à grande échelle. Une validation temporelle en absolue n'étant pas possible à grande échelle par manque de données adéquates, une méthode de comparaison des cartes au niveau temporel est définie.

2.1. METHODE DE VALIDATION SPATIALE

L'objectif est d'évaluer spatialement la précision des détections produites par la méthode présentée ci-dessus à l'échelle de la Guyane. Afin d'évaluer la sensibilité de la méthode à différents motifs spatiaux et à différentes temporalités, la validation a été effectuée séparément pour 3 types de déboisement (agriculture de subsistance, exploitation forestière et orpaillage). Notez qu'à des fins de comparaison, les données produites par GLAD-FA (Hansen et al., 2016) présentées dans le chapitre 3 ont été évaluées avec la même méthode. La validation a été réalisée sur 2017 et 2018 uniquement, en raison de la disponibilité de GLAD-FA.

La stratification est un partitionnement des sites de validation en groupes homogènes mutuellement exclusifs (chaque unité évaluée appartient à une seule strate). Le plus souvent, les strates sont découpées avec les mêmes classes que celles des cartes évaluées (ici zone déboisée et forêt intacte). Le but de cette stratification est d'améliorer l'estimation de la précision de chaque classe. La strate "zone déboisée" a été formée par les données de référence collectées auprès d'organismes français (voir section 2.3 : données exploitation forestière, données orpaillage, données abattis). La strate "forêt intacte" a été prise comme le complément de la strate « zone déboisée » sur le site de validation. Les zones de non-forêt ont été masquées a priori à l'aide de la carte forêt/non-forêt décrite dans la section 5.1 (produite à partir de la donnée d'ECOSEO). Dans le contexte de faible déboisement de la Guyane, utiliser un protocole d'échantillonnage aléatoire stratifié permet une représentation statistique suffisante de chaque classe de la carte (déboisement / forêt intacte) (Olofsson et al., 2014).

Dans chaque strate, une sélection aléatoire des échantillons a donc été effectuée. Afin de déterminer le nombre d'échantillons total optimal pour un échantillonnage aléatoire stratifié, (Cochran, 1977) fournit la formule suivante :

$$\mathrm{n} = \frac{(\sum \mathrm{W_i S_i})^2}{\left[\mathrm{S}(\hat{\mathrm{O}})\right]^2 + (1/\mathrm{N})\Sigma \mathrm{W_i S_i^2}} \approx \left(\frac{\sum W_i S_i}{S(\hat{O})}\right)^2 \qquad 5.9$$

où W_i est la proportion de la surface cartographiée pour la « classe i », $S_i = \sqrt{U_i(1 - U_i)}$ est l'écart type de la strate i avec U_i la précision (user's accuracy) attendue, $S(\widehat{O})$ est l'erreur standard de la précision globale estimée, et N est le nombre d'unités dans la zone de validation. Comme N est généralement grand (ici > 10 millions de pixels), le terme $(1 / N)\Sigma W_i S_i^2$ peut être ignoré.

Afin de calculer le nombre d'échantillons total optimal dans notre cas, l'erreur standard prise pour la précision globale est de 0.01. Il a été montré, d'après l'expérience acquise dans le cadre d'efforts similaires de cartographie des changements, que les erreurs de commission sont relativement fréquentes pour les classes de changement, tandis que les classes stables sont plus précises (Olofsson et al., 2011, 2010). Les précisions définies pour les classes « déboisées » et « forêt intacte » sont donc fixées à 0.7 et 0.9 respectivement. La proportion de zones déboisées est d'environ 0.07%. Étant donnés ces chiffres, l'équation 5.9 suggère d'utiliser environ 900 points. Pour évaluer séparément les 3 types de déboisement, ce chiffre passerait à 2 700 échantillons nécessaires au total et vérifier ces points un à un (voir suite de la méthode) aurait pris trop de temps. De plus, l'évaluation globale (tout type de déboisement confondu) aurait été basée sur 2 700 points (au lieu des 900 suggérés). Utiliser 500 points par type de déboisement (pour un total de 1500 points en considérant les 3) représente un bon compromis entre nombre de points nécessaires et coût en temps acceptable.

Les 500 points par type de déboisement ont été alloués en suivant les recommandations de (Olofsson et al., 2020, 2014), de sorte que 50 points soient sélectionnés dans la strate « déboisée » et 450 points dans la strate « forêt intacte ». L'unité d'évaluation spatiale utilisée est un point avec une zone tampon de 10 mètres. L'étiquetage des points sélectionnés a été généré par l'interprétation visuelle de séries temporelles d'images optiques à partir des sites internet du *Sentinel Hub* (https://apps.sentinel-hub.com/eo-browser/) et de *Google Earth* (www.google.fr/intl/fr/earth/). La classe prédite du point est la classe couvrant la majorité de sa surface. Sur les 1 500 points générés aléatoirement et interprétés, 1 487 points ont été retenus : 5 points de « forêt intacte » ont été supprimés et 5 ont finalement été classés comme déboisés. Quatre points « déboisés » étaient déjà déboisés au début de l'étude ou non déboisés à la fin de l'étude et ont donc été supprimés. Enfin, 4 points ont été supprimés car il était impossible de les interpréter en raison de l'absence d'images optiques sans nuages.

Conformément aux recommandations de bonnes pratiques (Olofsson et al., 2014) la matrice de confusion devrait être rapportée en termes de proportion de surfaces estimées plutôt qu'en termes de nombre d'échantillons en utilisant les équations (4) dans (Olofsson et al., 2014). Utiliser cet estimateur sans biais permet d'estimer les surfaces ajustées avec les erreurs, et leurs incertitudes.

Cependant, comme mentionné dans Olofsson et al., 2020, les effets des erreurs d'omission sont exacerbés dans une situation où la superficie de changement est faible par rapport à la superficie totale de la zone. En Guyane française, la proportion de changement forestier est extrêmement faible (environ 0,07% par rapport à la forêt intacte) et l'utilisation de la proportion des surfaces estimées pour l'évaluation de la précision ne serait donc pas adaptée. Comme les effets de l'omission sont réduits avec le poids de la strate forestière, Olofsson et al., 2020 recommande de diviser la strate forestière en un petit sous ensemble qui contient toutes les erreurs d'omission et un grand sous ensemble qui est exempt d'erreurs d'omission. Pour y parvenir, une approche illustrée dans la littérature consiste à utiliser une strate tampon « proche déboisement » pour atténuer les effets des erreurs d'omission (Arévalo et al., 2020 ; Bullock et al., 2020 ; Potapov et al., others, 2017 ; Tyukavina et al., 2013). Cette solution est basée sur l'hypothèse que les changements sont spatialement proches les uns des autres. Par conséquent, les erreurs d'omissions des changements se produiraient à proximité des changements cartographiés. Aussi, les zones cartographiées « forêt intacte » seraient peu susceptibles de contenir des erreurs d'omission, car elles se trouveraient à des distances plus importantes des changements cartographiés. Or, afin qu'une strate tampon soit efficace pour réduire le poids des erreurs d'omission, sa taille (et donc son poids) doit rester relativement faible.

La Guyane française présente un déboisement de faible densité. Le déboisement due à l'orpaillage peut être éloignée de plus de 10 km de tout autre déboisement. Par conséquent, l'inclusion des erreurs d'omission dans une strate tampon nécessiterait une strate de grande taille, ce qui la rendrait peu efficace. L'approche tampon n'est donc pas indiquée dans le cas de déboisements à très faible densité. En raison de la situation spécifique de la Guyane française (surface de forêt intacte extrêmement importante par rapport à la surface de changements forestiers, et déboisement à faible densité), nous avons décidé de rapporter la matrice de confusion en termes de nombre d'échantillons.

2.2. METHODE DE COMPARAISON TEMPORELLE

Les dates de déboisement associées aux données de référence ne sont pas suffisamment précises pour permettre une validation temporelle fiable des dates de détection estimées (cf. section 3.3). De plus, la difficulté d'interpréter visuellement les images RSO en raison du speckle et la faible récurrence des images optiques exploitables n'ont pas permis de construire un jeu de données de validation temporelle fiable, à large échelle et donc de calculer une incertitude temporelle réelle. Afin d'évaluer le potentiel des méthodes pour le suivi du déboisement en quasi-temps réel, les dates

estimées par chaque méthode (méthode de détection Sentinel-1 et GLAD-FA) ont été comparées. Cela concerne les détections faites sur 2017 et 2018 comme pour la validation spatiale.

Les polygones de référence utilisés pour la validation spatiale et détectés par les deux méthodes ont servi de base de comparaison (soit 41.6% de la quantité totale de données de référence). Un polygone est considéré comme détecté par une méthode s'il est intersecté par une de ses détections. Pour chaque méthode, la valeur de date estimée de déboisement assignée à chaque polygone de référence est la médiane des dates de détection de tous les pixels contenus dans ce polygone.

En Guyane, Sentinel-1 fournit une image exploitable tous les 6 jours dans le passage ascendant, et 12 jours dans le passage descendant. Par conséquent, un déboisement se produira à un maximum de plus ou moins 3 jours d'une image exploitable de Sentinel-1 et l'ensemble du territoire est vu 3 fois tous les 12 jours. Avec le temps de revisite de Sentinel-1 et en prenant 3 images post-perturbation (X_a = 3, voir section 5.1), le délai d'alerte de la méthode présentée ci-dessus est de 12 jours.

En théorie, Landsat a une résolution temporelle de 8 jours (en combinant Landsat-7 et Landsat-8), donc un déboisement se produira à un maximum de plus ou moins 4 jours d'une acquisition. En pratique, la disponibilité des observations est limitée par les stratégies d'acquisition de Landsat et par la couverture nuageuse. Le délai d'alerte de GLAD-FA est donc variable, avec un minimum de 8 jours (2 observations nécessaires pour confirmer une alerte).

2.3. DONNEES UTILISEES POUR LA VALIDATION

Pour la validation spatiale et la comparaison temporelle, des données de référence provenant de différents facteurs de pertes forestières ont été utilisées. Ces données, décrites dans le chapitre 3, sont toutes issues d'organismes français et utilisées comme référence pour le suivi des forêts en Guyane. Toutes les données de l'OAM, du PAG et de l'ONF disponibles pour 2017 et 2018 et sur des parcelles de plus de 0.2 ha ont été utilisées. Leur nombre et leurs statistiques descriptives après cette sélection sont présentées dans le Tableau 3.

Tableau 3 - Nombre et taille des polygones de référence pour les différents types de déboisement en Guyane.

	Nombre de zones	**Surface moyenne (ha)**	**Ecart type des surfaces (ha)**	**Surface totale (ha)**
Agriculture	602	0.47	0.35	285.7
Exploitation forestière	82	1.31	3.61	107.8
Orpaillage	1183	1.46	2.1	1731

Les données de référence sur l'agriculture sont produites annuellement par le PAG. Sur une zone de 423 585 ha (Figure 56 gauche, en orange), 665 polygones correspondent à une nouvelle perte de forêt : étiquetés « forêt ancienne » jusqu'en 2016 (inclus) et « abattis-brulis » en 2017 ou 2018. Après application du MMU de 0.2 ha, il reste 602 polygones, totalisant 285.7 ha déboisés (0.067% de la zone de 423 585 ha).

Les données de référence des routes nouvellement construites, des parcs à grumes et des trouées d'abattage pour l'exploitation forestière ont été mises à disposition par l'ONF. Sur une zone de 26 630 ha (Figure 56 gauche, en vert), 82 polygones ont été sélectionnés, totalisant 107.8 ha de perte de forêt (0.40%).

Les données sur l'orpaillage sont collectées en continu par l'OAM. En 2017-2018, 1 183 polygones ont été numérisés. Ces polygones sont répartis sur une zone de 5.8 Mha (Figure 56 gauche, en jaune), totalisant 1 731 ha de perturbations (0.03%).

L'urbanisation n'a pas été évaluée car il n'existe pas de données de qualité équivalente à celles utilisées pour les autres types de déboisements.

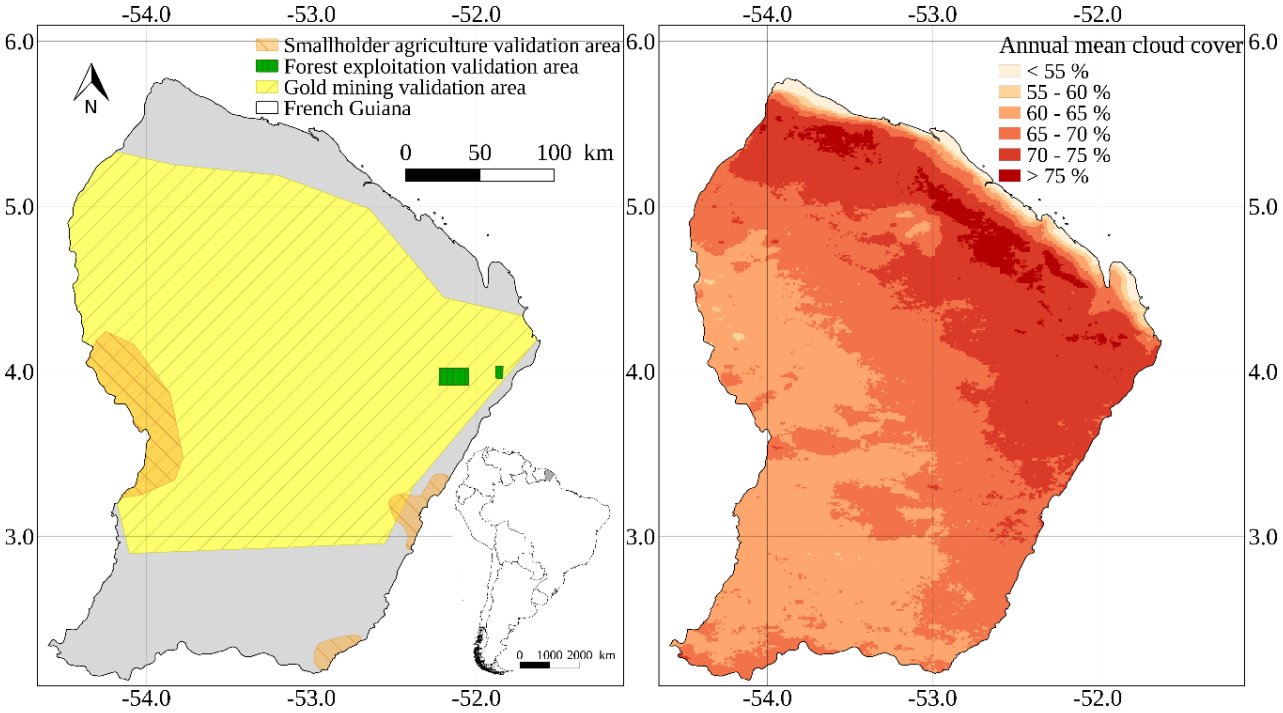


Figure 56 - À gauche : Guyane française avec les zones de validation pour chaque type de déboisement (orange : agriculture, vert : exploitation forestière, jaune : orpaillage) ; à droite : couverture nuageuse moyenne annuelle de la Guyane française (Crédits : Imagery produced by the NASA Earth Observations team based on data provided by the MODIS Atmosphere Science Team, NASA Goddard Space Flight Center.).

Pour résumer :

- **Les méthodes de validation spatiale des changements forestiers recommandées dans la littérature ne sont pas applicables pour la Guyane car les pertes forestières peuvent y être très faibles relativement à la surface de forêts totale ;**
- **La validation spatiale se fera avec des points provenant d'un échantillonnage aléatoire stratifié et les strates seront définies par des données produites et utilisées opérationnellement par les acteurs locaux. La matrice sera reportée en nombre d'échantillons ;**
- **Une validation temporelle à large échelle n'est pas possible en raison du manque de données adaptées ;**
- **La méthode développée pour détecter le déboisement sera comparée au niveau spatial et temporel, à GLAD-FA, basé sur des données optiques, souvent utilisé comme référence.**

3. Caractérisation des déboisements

Une fois les déboisements détectés, il faut pouvoir définir leur cause rapidement pour que le système d'alerte soit efficace. L'objectif est donc de proposer un algorithme pour labelliser les zones déboisées détectées avec leur cause. Cet algorithme devra être automatique, fonctionner en temps réel, à grande échelle et avec des données disponibles pour les acteurs locaux. Cette section présente la méthode utilisée pour développer le modèle.

3.1. IDEE GENERALE DE LA METHODE ET CONCEPT GENERAUX

L'idée générale pour faire cette caractérisation, est de créer un arbre de décision à partir d'indicateurs géographiques (exemple : distance aux routes), morphologiques (exemple : surface) et temporel (date de la coupe), et permettant de définir les facteurs directs des déboisements détectés. Pour cela, des méthodes d'apprentissage automatique (ou machine learning en anglais) sont utilisées. La priorité est mise sur la réduction du nombre d'indicateurs et sur la réduction du nombre de règles de décision nécessaires, afin que le modèle produit puisse être répliqué et utilisé facilement par les acteurs locaux. Le processus de création de cet arbre se déroule en 3 étapes :

- Définition des indicateurs potentiellement utiles. Les indicateurs géographiques sont construits à partir des données présentées dans le chapitre 3. Les indicateurs morphologiques et l'indicateur temporel sont extraits des détections produites avec Sentinel-1 ;
- Réduction du nombre d'indicateurs pour sélectionner ceux qui ont le plus d'importance et éliminer ceux qui pourraient être redondants ;
- Production d'un arbre de décision à partir des indicateurs sélectionnés : cet arbre devra minimiser le nombre d'indicateurs utilisés.

La fouille de données (ou data mining) au sens large est une technique qui considère tout le processus d'extraction de connaissances à partir d'un grand nombre de données. Les algorithmes d'apprentissage automatique permettent de résoudre un problème de manière automatique, en se servant des informations contenues dans les données utilisées pour l'entraînement. La section suivante présente les arbres de décision et leur technique d'apprentissage automatique, ainsi que le concept de random forest, utilisé ici pour la sélection des indicateurs pertinents.

3.1.1. ARBRE DE DECISION

Un arbre de décision est un algorithme déroulant une séquence de tests basés sur des indicateurs (variables explicatives) afin de prédire un résultat. Pour cette étude, nous nous

intéresserons aux arbres de classification CART (Breiman et al., 1984), qui permettent de faire des prédictions qualitatives (prédire des classes). Le nœud racine et les nœuds internes sont associés à un test binaire, et les nœuds terminaux sont associés à une prédiction.

3.1.1.1. CONSTRUCTION D'UN ARBRE DE DECISION

Un arbre de décision peut être construit de manière automatique à partir d'un ensemble d'apprentissage Q tel que $Q=\{X_i = y_i\}$ ($i \in 1, N$) avec N le nombre d'échantillons dans l'ensemble d'apprentissage, $X_i = (x_{i,1}, x_{i,2}, x_{i,3}, \ldots, x_{i,k})$ les k variables explicatives, et y_i la classe associée à la variable i.

Au départ, toutes les observations sont placées dans le nœud racine. La division du nœud t à l'aide d'une règle binaire δ de la forme $x_{i,j} \leq \alpha$ ($j \in 1, k$), partitionne l'ensemble des données Q_t du nœud t dans deux nœuds fils (Figure 57) tel que :

- L'ensemble $Q_{t(g)}$ du nœud de gauche t_g est défini par : $Q_{t(g)}(\delta) = \{(X, y) | \, x_j \leq \alpha \}$
- L'ensemble $Q_{t(d)}$ du nœud de droite t_d est défini par : $Q_{t(d)}(\delta) = Q_t \backslash Q_{t(g)}$

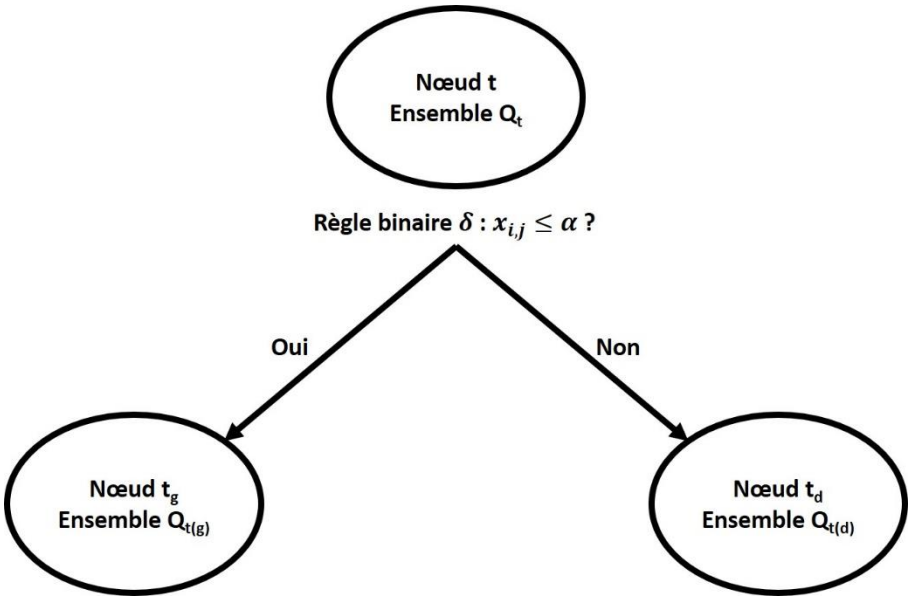


Figure 57 - Illustration de la division d'un nœud t et de son ensemble Q_t à l'aide d'une règle binaire δ.

Pour définir la règle associée à un nœud, on teste plusieurs règles δ avec différents $x_{i,j}$ et différentes valeurs α, et on choisit la règle δ^* qui minimise l'hétérogénéité des classes dans les nœuds fils en utilisant une fonction d'impureté H() :

$$\delta^* = argmin_\delta \left(\frac{N(t_g, \delta)}{N(t)} H\left(Q_{t(g)}(\delta)\right) + \frac{N(t_d, \delta)}{N(t)} H\left(Q_{t(d)}(\delta)\right) \right) \qquad 5.10$$

Où $\mathrm{N}(\mathrm{t_g}, \delta)$ est le nombre d'observations finissant dans le nœud $\mathrm{t_g}$ avec la règle δ, $\mathrm{N}(\mathrm{t_d}, \delta)$ le nombre d'observations finissant dans le nœud $\mathrm{t_d}$ avec la règle δ et $\mathrm{N}(\mathrm{t})$ le nombre d'observations dans le nœud t.

La fonction d'impureté utilisée pour un ensemble Q_t sera celle de Gini :

$$H(Q_t) = \sum_l p_{tl}(1 - p_{tl}) \qquad 5.11$$

Avec $p_{tl} = \frac{1}{N(t)} \sum_{y \in Q_t} I(y = l)$ la proportion d'observations de la classe l dans le nœud t.

L'apport A du nœud t sera évalué par la variation de l'hétérogénéité générée par δ^* :

$$A(\delta^*) = \frac{\mathrm{N(t)}}{\mathrm{N}} \times \left(H(Q_t(\delta^*)) - \frac{\mathrm{N}(\mathrm{t_g}, \delta)}{\mathrm{N(t)}} H\left(\mathrm{Q}_{t(g)}(\delta^*)\right) - \frac{\mathrm{N}(\mathrm{t_d}, \delta)}{\mathrm{N(t)}} H\left(\mathrm{Q}_{t(d)}(\delta^*)\right) \right) \qquad 5.12$$

Les nœuds suivants sont construits par récursivité, en utilisant l'ensemble de données qui leur est associé. Il est à noter qu'une variable explicative peut être utilisée plusieurs fois pour des tests différents.

Le critère d'arrêt de l'arbre de décision peut être basé :

- Soit sur le nombre N(t) minimal d'observations requises dans un nœud t ;
- Soit si l'apport d'une nouvelle règle est inférieur à un certain seuil ;
- Soit si la taille maximale fixée pour l'arbre est atteinte.

Dans un nœud terminal, la classe dominante (majoritaire) donne la prédiction.

Une fois l'arbre construit, classer une observation se fait par descente de l'arbre en partant de la racine vers les feuilles. A chaque nœud, l'observation est évaluée par un test qui détermine où elle continue son chemin.

3.1.1.2. AVANTAGES ET INCONVENIENTS

Les avantages des arbres de décision sont multiples :

- Ils sont facilement compréhensibles, interprétables et peuvent être visualisés ;
- Ils nécessitent peu de préparation des données (normalisation, création de variables fictives, etc…) ;
- Ils ont un coût d'utilisation logarithmique (faible) pour classifier une nouvelle donnée et peuvent répondre à des problèmes en temps-réel.

En revanche :

- Au vu de leur construction, on comprend que les arbres de décision sont des algorithmes gloutons : ils vont chercher la meilleure règle à chaque nœud sans retour en arrière. La solution n'atteindra donc pas nécessairement un optimum global, et l'association d'autres indices que ceux utilisés dans l'arbre construit pourrait donner des résultats équivalents ou meilleurs ;
- Ils ont tendance à s'adapter de manière excessive aux données comportant un grand nombre de caractéristiques (phénomène de sur-apprentissage), avec un risque de généralisation difficile (à d'autres données, d'autres dates ou d'autres endroits). Pour éviter ce problème, il est possible de fixer le nombre minimal d'observations requises pour diviser un nœud, son apport minimal ou fixer la profondeur maximale de l'arbre ;
- Ils peuvent être instables car de petites variations dans les données peuvent entraîner la génération d'arbres complètement différents. Les changements produits près de la racine auront plus d'impact sur l'arbre en résultant. On dit que les arbres de décision produisent des estimateurs de variance élevée.

Nous avons choisi d'utiliser un arbre de décision comme modèle prédictif car il sera facilement compréhensible et interprétable, qu'il ne nécessite pas de préparation de la donnée a priori pour être utilisé et qu'il peut répondre à un problème en temps réel sans nécessiter de gros moyens de calcul. Ces caractéristiques sont essentielles au vu des besoins de la méthode et pour son appropriation par les acteurs locaux. Cependant, le résultat obtenu ne sera pas nécessairement optimal.

3.1.2. ALGORITHME DU RANDOM FOREST (RF)

Pour répondre au problème de variance élevée des arbres de décision, plusieurs approches dont les random forest (RF) (Breiman, 2001), ou forêts aléatoires sont apparues. Les RF sont basées sur l'apprentissage de plusieurs arbres de décision entraînés sur des sous-ensembles de données légèrement différents, et en intégrant une part de hasard pour le choix des indicateurs dans les règles de décision. La moyenne de ces estimateurs permettra de réduire la variance.

Les RF sont une amélioration des arbres de décision et en diffèrent sur les 3 points suivants :

- Les RF sont formées de plusieurs arbres de décision donc le résultat est donné par la classe prédite qui aura obtenu un vote majoritaire en sa faveur, ou en moyennant les probabilités de chaque classe sur chaque arbre et en prenant la plus forte ;

- La base d'apprentissage Q' de chaque arbre est un sous-ensemble de la base d'apprentissage initiale Q. Les N' échantillons qui sont utilisés dans Q' sont tirés aléatoirement avec remise ;
- Les variables explicatives disponibles pour établir les règles de décision de chaque nœud sont un sous-ensemble des variables explicatives disponibles initialement Xi = ($x_{i,1}$, $x_{i,2}$, $x_{i,3}$, …, $x_{i,k}$). Elles sont tirées aléatoirement avec remise.

La variance est donc réduite d'un rapport $\frac{1}{nombre\ d'arbres\ de\ la\ RF}$. Le niveau de hasard introduit dans chaque arbre permet de réduire leur corrélation.

Pour prédire l'importance des indicateurs, il est préférable d'utiliser les RF. Le pouvoir prédictif d'un indicateur peut être estimé en se basant sur l'apport de ses nœuds associés, et sur la proportion attendue d'échantillons qu'il va traiter. Utiliser les RF permet de réduire la variance d'une telle estimation. C'est pourquoi elles peuvent être utilisées pour la sélection de variables explicatives (Louppe, 2014). Cependant, le calcul de l'importance basé sur l'impureté a deux inconvénients. D'une part, il peut être biaisé : l'importance est calculée sur des statistiques dérivées des données d'apprentissage et n'informe donc pas nécessairement sur les indicateurs les plus importants pour faire de bonnes prédictions sur un autre ensemble de données. D'autre part, il favorise les caractéristiques à cardinalité élevée, c'est-à-dire les caractéristiques ayant de nombreuses valeurs uniques.

Calculer l'importance des caractéristiques **par permutation** est une alternative qui ne souffre pas de ces défauts. Elle est définie comme étant la diminution du score d'un modèle lorsque les valeurs d'un indicateur sont mélangées de façon aléatoire (Breiman, 2001). Cette procédure rompt la relation entre l'indicateur et la cible, et la diminution du score indique la dépendance du modèle à cet indicateur. Une faible diminution du score impliquera que l'indicateur a peu d'importance dans le modèle. Cette technique a l'avantage d'être agnostique vis-à-vis du modèle et peut être calculée de nombreuses fois avec différentes permutations sur les indicateurs. De plus, l'importance par permutation peut être calculée sur un ensemble de données autre que l'ensemble d'apprentissage. L'utilisation de l'ensemble de validation pour ce calcul permet de mettre en évidence les indicateurs qui contribuent le plus au pouvoir de généralisation du modèle. En revanche, le calcul de l'importance par permutation est biaisé dans le cas de variables explicatives corrélées. En effet, lorsque deux caractéristiques sont corrélées et qu'une d'entre elles est permutée, le modèle aura toujours accès à la caractéristique par le biais de sa caractéristique corrélée. Il en résultera une valeur d'importance inférieure pour les deux caractéristiques, alors qu'elles pourraient être importantes.

Les RF permettent d'avoir des résultats plus fiables et plus stables que les arbres de décision à la suite de l'entrainement. Cependant, dans notre étude, les RF ne serviront pas de modèle prédictif final car nous privilégions l'interprétabilité du modèle. En revanche, elles seront utilisées pour définir l'importance des indicateurs dans notre problème.

3.2. PREPARATION : DEFINITION D'INDICATEURS POTENTIELS ET CREATION DE LA BASE D'APPRENTISSAGE

En apprentissage automatique, les modèles ont besoin d'une base d'observations différentes pour l'apprentissage (données servant à construire le modèle), la validation (données servant à évaluer les prédictions du modèle en vue de le modifier si besoin, ou pour comparer deux modèles) et le test (données utilisées pour évaluer la précision du modèle finalement choisi).

Chaque base est composée d'observations disjointes, afin d'éviter de biaiser l'évaluation du modèle avec des données qui serviraient à l'apprentissage. Chaque observation est de type (X_i, y_i) avec $X_i = (x_{i,1}, x_{i,2}, x_{i,3}, \ldots, x_{i,k})$, les x_i sont les variables explicatives du modèle, et les y_i sont les classes de chaque observation.

Les indicateurs (X_i) définis pour notre étude ont été choisis suite à l'analyse du chapitre 4. Ils ont été construits à l'aide des données décrites dans le chapitre 3, de manière à être exhaustifs sur le territoire. Certaines couches de données ont été construites (voir annexe (2)) dans le but de représenter des phénomènes particuliers comme l'agriculture vivrière ou les exploitations agricoles. Au total, 51 indicateurs ont été définis dont : 42 géographiques (distance à des usages ou occupations du sol particuliers, élévation et pente), 8 morphologiques (dérivés des formes des détections), et 1 temporel (jour de la coupe dans l'année). La liste complète de ces indicateurs est disponible en annexe (2).

Les observations (détections Sentinel-1) ont été labellisées (yi) avec les classes suivantes : orpaillage, agriculture, urbanisation, exploitation forestière, autres (causes naturelles). Ces classes ont été choisies en se basant sur les différents usages du sol répertoriés en Guyane après une coupe anthropique :

- Orpaillage ;
- Urbanisation : infrastructures et bâtiments ;
- Agriculture : cultures herbacées et ligneuses, agriculture itinérante ;
- Exploitation forestière : chemins liés à l'exploitation, parcs à grumes.

La classe « autre » a pour but d'intégrer toutes les pertes de couvert forestier liées à des causes naturelles comme les chablis.

Trois jeux de données ont été préparés pour l'entraînement, la validation et le test. Les observations utilisées dans la base de données pour l'**apprentissage** sont dérivées des déboisements détectés en 2016. Concrètement, une observation correspond à un pixel sur laquelle un déboisement a été détectée en 2016. Les variables explicatives (X_i) associées à ce pixel ont été dérivées : de l'état de la variable observée au début de 2016 dans le cadre des indicateurs géographiques ; de la parcelle déboisée détectée (objet) dans laquelle s'intègre le pixel (pixels connexes déboisés à une même date) ou de la parcelle déboisée détectée connexe (pixels connexes déboisés quelle que soit la date) pour les indicateurs morphologiques ; de la date de détection du pixel pour l'indicateur temporel.

Les données de **validation** sont dérivées des détections 2017 et de l'état de la variable début 2017. De même, les données de **test** sont dérivées des détections 2018 et de l'état de la variable début 2018.

Ce découpage a été fait dans le but d'avoir des variables distinctes pour chaque étape, mais également pour illustrer la prédiction du modèle en temporel.

Les observations ont été labélisées (y_i), en utilisant les couches suivantes (si leurs informations spatiale et temporelle concordaient) :

- Orpaillage : OAM, agriPAG : OR, Titres miniers Camino ;
- Agriculture : zones agricoles comme définies dans les indicateurs ;
- Exploitation forestière : Parcs, parcelles d'exploitation, trouées d'abattage.

Ces assignations automatiques ont été complétées par de la photo-interprétation d'images optiques. Pour cela, les mosaïques biannuelles « sans nuage » disponibles via le plug-in Planet Explorer (www.planet.com/nicfi) ont été utilisées. Ces données sont libres d'accès. Au final, 35% des zones ont été labélisées automatiquement et 65% ont été labélisées par photo-interprétation. Le Tableau 4 présente le nombre d'observations recueillies dans chaque base pour chaque classe.

Tableau 4 - Nombre d'observations (pixels) disponibles par classe

	Apprentissage	Validation	Test
Orpaillage	57 347	66 952	68 473
Agriculture	162 632	102 429	113 605
Urbanisation	40 846	15 961	22 295
Exploitation forestière	22 413	6 642	8 926
Autre	3 114	1 620	2 567

Afin qu'un arbre de décision ne soit pas biaisé avec des classes dominantes, il est préférable d'équilibrer l'ensemble de données lors de l'apprentissage. Ceci peut être effectué en sélectionnant un nombre égal d'échantillons par classe, ou en pondérant les échantillons par le poids de chaque classe. Dans le cadre de l'étude, c'est la deuxième option qui a été adoptée.

3.3. CREATION DU MODELE

Toutes les variables explicatives ne sont pas nécessairement importantes pour prédire les classes prédites. Une variable non informative peut avoir des effets néfastes sur la précision du modèle. Effectuer une sélection de variables (et donc réduire la dimensionnalité) au préalable a deux avantages. D'abord, le modèle aura de meilleures chances de trouver des caractéristiques discriminantes pour s'approcher d'une solution globale. L'erreur de prédiction devrait donc être réduite. D'autre part, un modèle avec moins d'indicateurs est plus interprétable, et l'utiliser nécessite moins de travail (préparation des données d'entrées). Le but est donc d'obtenir le meilleur résultat avec le moins d'indicateurs possible.

La sélection des indicateurs se fera en 2 étapes :

- Définir l'importance des indicateurs pour expliquer les classes prédites ;
- Calculer les performances d'un arbre de décision en fonction du nombre de variables descriptives en entrée. Les variables seront insérées une par une en fonction de leur importance.

3.3.1. IMPORTANCE DES INDICATEURS

Afin de définir les indicateurs les plus utiles pour caractériser les pertes forestières, nous utilisons un algorithme RFE (Recursive Feature Elimination), avec une mesure d'importance par permutation comme critère de rang. RFE élimine récursivement les variables les plus faibles.

Gregorutti et al., 2014 a montré que cette approche permettait de conserver les avantages de la mesure d'importance par permutation en réduisant l'effet des variables corrélées.

3.3.2. NOMBRE D'INDICATEURS ET TAILLE DE L'ARBRE CHOISI

Une fois les indicateurs classés par importance, le but est de trouver un arbre de décision qui maximise les résultats et minimise le nombre d'indicateurs utilisés. Pour cela, on a choisi d'entraîner des arbres de profondeur 3, 5, 7, et 10 en utilisant les j indicateurs ayant la plus grande importance, pour chaque valeur $j \in (1, k)$. La métrique utilisée pour évaluer le score de chaque modèle est le kappa (décrit dans l'annexe 1). Les échantillons du Tableau 4 ont tous été utilisés, pondérés par le poids de chaque classe.

3.3.3. ARBRE DE DECISION CHOISI COMME MODELE POUR LA CARACTERISATION DES PERTES FORESTIERES

Afin de caractériser au mieux les pertes forestières en ayant une méthode qui peut être facilement prise en main par les acteurs locaux, l'arbre sélectionné comme modèle prédictif sera celui qui aura le plus grand kappa et le plus petit nombre d'indicateurs. A nombre d'indicateurs et kappa égaux, l'arbre sélectionné sera celui qui sera le moins profond, afin d'avoir un arbre plus interprétable d'une part, et pour minimiser le sur-apprentissage d'autre part.

Un indice de confiance est associé à chaque prédiction. Il est défini par la proportion d'observations classées identiquement dans le nœud terminal qu'il a atteint.

3.4. VALIDATION DU MODELE DE CARACTERISATION DU DEBOISEMENT

L'arbre de décision produit pour caractériser le déboisement sera évalué avec une matrice de confusion contenant ses 5 classes. Les mesures présentées dans l'annexe 1 sont donc valables pour ce modèle. La précision globale peut être biaisée par des classes majoritaires. Étant donné le fort déséquilibre de notre jeu de données (voir section 3.2) le F-score sera utilisé comme métrique d'évaluation de chaque classe.

Pour résumer :

- **Nous avons choisi d'utiliser un arbre de décision comme modèle prédictif pour définir les causes des déboisements en Guyane ;**
- **Afin d'obtenir un arbre de décision le plus performant possible, nous avons défini l'importance des indicateurs possibles pour la discrimination des causes de déboisement à l'aide d'une méthode de Random Forest.**

🔔 **Ce qu'il faut retenir**

Dans ce chapitre, nous avons décrit la méthode de détection des pertes forestières (Bouvet et al., 2018) et nous avons défini les paramètres et les images à utiliser pour l'appliquer sur la Guyane. Les améliorations apportées à cette méthode dans le cadre de cette thèse ont été décrites : meilleure prise en compte de la dimension temporelle, gestion de la topographie et définition d'un indice de confiance.

Les méthodes recommandées et couramment utilisées pour valider ce type de produit étant peu adaptées dans le contexte de la Guyane, nous avons décidé d'utiliser une validation avec échantillonnage stratifié, et une matrice rapportée en nombre d'échantillons. La stratification est basée sur les données de référence fournit par des organismes locaux et utilisées de manière opérationnelle.

Nous avons proposé une méthode de caractérisation des déboisements basée sur des indicateurs géographiques, morphologiques et temporel utilisés dans un arbre de décision. Les indicateurs géographiques sont des distances à des usages ou occupation du sol particuliers. Les indicateurs morphologiques sont liés à la forme du déboisement ou de la parcelle à laquelle il se rattache. L'indicateur temporel correspond au jour de coupe dans l'année. La méthode d'analyse de ces indicateurs, dont le but est de déduire ceux qui sont les plus pertinents pour répondre au problème posé, est décrite dans ce chapitre, ainsi que la méthode de sélection du modèle prédictif final.

La chapitre 6 présentera les résultats obtenus à partir de ces méthodes.

✋ **Ma contribution**

La méthode Bouvet et al., 2018 était déjà développée et publiée au début de ma thèse. Celle-ci avait été testée sur une zone du Pérou d'environ 600 000 ha. Pendant ma thèse, j'ai développé une chaine de traitement permettant d'appliquer cette méthode à large échelle et à laquelle j'ai également intégré les améliorations décrites dans ce chapitre. Cette chaine de traitement a permis d'étudier les résultats de cette méthode à l'échelle de plusieurs territoires, et sur plusieurs années (voir chapitre 7). Aujourd'hui cette chaine de traitement est opérationnelle et tourne de manière automatique dans le cadre d'un projet de détecion des pertes forestières tropicales nommé TropiSCO[5].

J'ai également participé au développement du logiciel de prétraitement S1tiling en intégrant une fonction permettant de couper automatiquement les bords d'une image quand l'information qu'ils contiennent n'est pas pertinente.

J'ai proposé et étudié l'utilisation d'indicateurs géographiques, morphologiques et temporel pour caractériser les déboisements en quasi temps-réel comme cela est décrit dans ce chapitre. J'ai fait ce choix car attendre l'acquisition d'une image optique exploitable, ou une longue série temporelle d'images Sentinel-1 permettant de caractériser le déboisement ne me semblait pas pertinent quant au besoin de rapidité. J'ai défini cette méthode et ces indicateurs en dialoguant avec les acteurs locaux (en particulier avec l'État-major de l'Orpaillage et de la Pêche Illicite, car la caractérisation de l'orpaillage a été le premier besoin soulevé).

[5] https://www.spaceclimateobservatory.org/tropisco-amazonia

Chapitre 6 - Résultats

Contenu

1. Résultats – Détections des pertes forestières 138
1.1. Validation spatiale 139
1.2. Comparaison spatiale avec un produit optique 141
1.3. Evaluation temporelle 147
1.4. Comparaison temporelle 149
1.4.1. Agriculture vivrière 149
1.4.2. Exploitation forestière 151
1.4.3. Orpaillage 152
1.5. Réduction du délai de détection 153
1.6. Observations 156
1.6.1. Taux annuel de changements forestiers 156
2. Résultats – Caractérisation des déboisements 158
2.1. Création du modèle 158
2.1.1. Importance des indicateurs 158
2.1.2. Nombre d'indicateurs et taille de l'arbre choisi 161
2.2. Validation du modèle 164
2.2.1. Résultats globaux 165
2.2.2. Analyse de la confusion des classes 165
2.2.3. Carte de confiance 167

Dans ce chapitre, les produits générés en utilisant les méthodes décrites dans le chapitre 5 sont présentés et validés.

1. Résultats – Détections des pertes forestières

Cette partie présente les résultats obtenus en utilisant la méthode de détection des déboisements basée sur Sentinel-1 et la méthode de validation, présentées au chapitre 5. Cette méthode a été appliquée à l'échelle de la Guyane entre les années 2016 et 2019. La Figure 58 présente la carte produite, ainsi que des zooms de 8×12 km mettant en évidence les différentes tailles et distributions des zones déboisées.

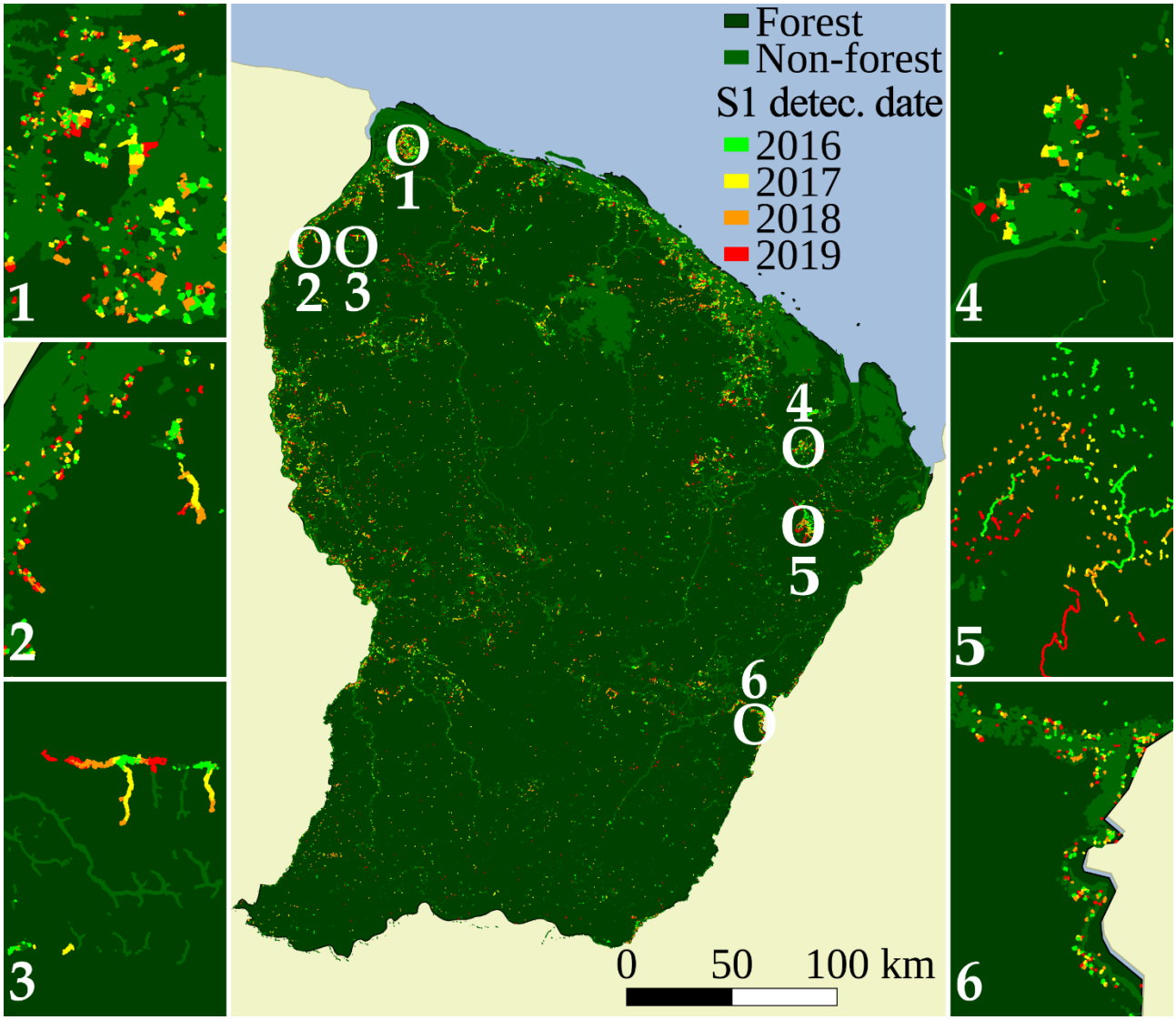


Figure 58 – Pertes forestières sur l'ensemble de la Guyane à partir de la méthode développée entre 2016 et 2019 à 10 mètres de résolution. Les images sur les bords montrent des zones de 8×12 km, mettant en évidence les différentes tailles et distributions des zones perturbées.

Sur cette figure, il est possible d'observer différentes pratiques agricoles. Les petites parcelles plutôt circulaires présentes dans les zones 2 et 6 correspondent à des abattis-brûlis, principalement pratiqués par les communautés Wayana, Wayãpi, Teko et Businengé. Les cultures fruitières des Hmong se manifestent par les parcelles de forme carrée un peu plus grandes visibles

dans la zone 1. Les activités minières légales se manifestent par de grandes formes linéaires, comme dans les zones 2 et 3. La zone 5 est une zone d'exploitation forestière : les détections linéaires fines sont les routes créées et les petites parcelles détectées avec une forte densité traduisent des ouvertures de la canopée dans la zone.

La Figure 59 montre des détections de pertes forestières résultant de la création d'abattis (à gauche) et d'orpaillage (à droite). Il est intéressant d'observer l'évolution temporelle progressive sur l'orpaillage par rapport aux abattis dont la surface n'évolue plus beaucoup après 6 à 12 jours (une ou deux acquisitions Sentinel-1).

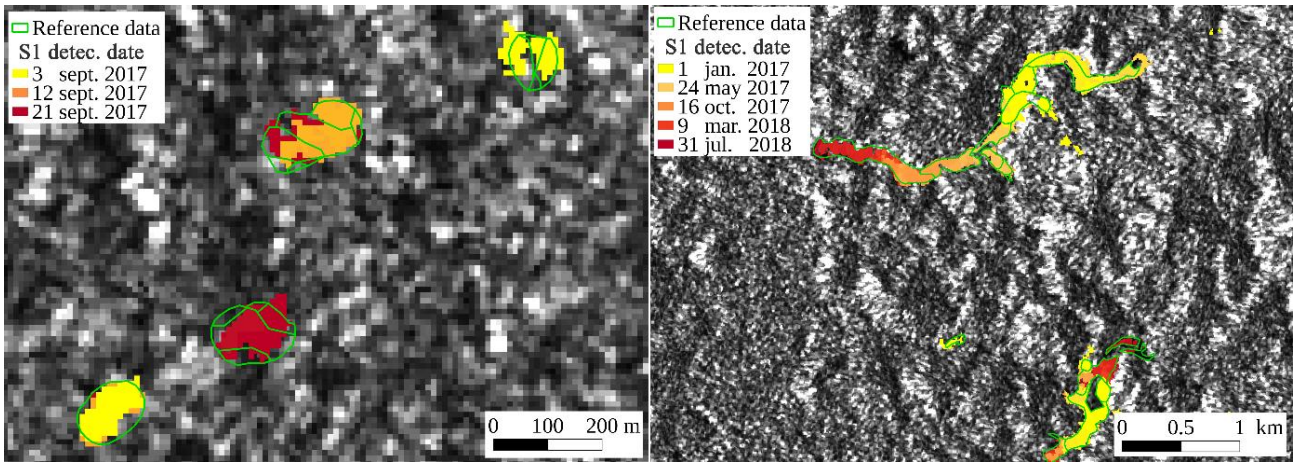


Figure 59 - Cartes des détections de pertes forestières résultant de la création d'abattis (à gauche) et d'orpaillage (à droite). Les couleurs indiquent la date de coupe détectée par l'algorithme. Les lignes vertes sont les limites des polygones de référence. La carte de gauche est centrée sur 3°40 N et 54°3O. Arrière-plan : Image RSO Sentinel-1.

1.1. VALIDATION SPATIALE

La carte de pertes forestières produite montre des performances élevées (Tableau 5) : tout type de perturbations confondues, le rappel est de 81.5% pour la classe « déboisée » et de 99.6% pour la classe « forêt intacte ». Ces valeurs sont très satisfaisantes pour un système d'alerte. Les précisions respectives pour ces classes sont de 96% et 98%, et la précision globale est de 97.8%.

Tableau 5 - Validation de la méthode développée avec les données Sentinel-1 (S1) (RSO) et de GLAD-FA (optique).

	Agriculture vivrière		**Exploitation forestière**		**Orpaillage**		**Global (3 types de pertes forestières combinées)**	
	Méthode basée sur S-1	GLAD-FA (Landsat)	Méthode basée sur S-1	GLAD-FA (Landsat)	Méthode basée sur S-1	GLAD-FA (Landsat)	Méthode basée sur S-1	GLAD-FA (Landsat)
Précision « déboisée » (%)	97.7	100	91.1	95.2	100	100	96.0	98.1
Précision « forêt intacte » (%)	98.9	93.3	97.1	92.9	98.0	94.3	98.0	93.5
Rappel « déboisée » (%)	89.4	31.9	75.9	37.0	80.0	40.0	81.5	36.6
Rappel « forêt intacte » (%)	99.8	100	99.0	99.7	100	100	99.6	99.9
Précision globale (%)	98.8	93.5	96.5	93.0	98.2	94.6	97.8	93.7

L'agriculture vivrière est le type de perturbations avec la plus grande proportion de perturbations détectées (rappel de 89.4%). Celle-ci s'apparente à de petites parcelles (moyenne de 0.47 ha et écart-type de 0.35 ha, voir chapitre 5, Tableau 3). Ces événements correspondant à des changements courts dans le temps permettent de détecter les ombres d'un seul tenant, et se produisent durant la saison sèche, c'est-à-dire quand les conditions météorologiques sont optimales pour observer une diminution du signal sur la surface. En effet, une augmentation du signal est souvent due à des précipitations et celles-ci sont rares en saison sèche. De plus, le brûlis entraîne une diminution de la végétation restante au sol, et donc de la rétrodiffusion.

Sur les zones d'exploitation forestière, la valeur du rappel est de 75.9%. Parmi les erreurs d'omission, un tiers concernaient des parcs à grumes de formes rectangulaires présents sur le bord d'une route déjà existante. Cette configuration fait que l'ombre a moins de chance d'être détectée car elle n'apparait que d'un côté. Les deux tiers d'erreurs d'omission restantes sont liés aux routes. Ces éléments longs et fins (environ 20 mètres) sont à la limite de ce qui peut être détecté avec un capteur ayant une résolution de 20 mètres. De plus, la largeur d'une ombre varie suivant la direction de visée du satellite. Une route construite perpendiculairement à la trajectoire de l'orbite aura une ombre peu importante. Par conséquent, les routes construites dans la direction nord-sud sont mieux

détectées par Sentinel-1 que les routes ouest-est. 43% de la surface totale des routes provenant des données de référence ont été détectées.
La détection des déforestations liées à l'orpaillage (rappel) est de 80%. Les sites d'exploitation aurifère peuvent être petits ou grands (moyenne de 1.46 ha et écart-type de 2.1 ha, voir chapitre 5, Tableau 3) et ont souvent une forme sinueuse car ils suivent les rivières (et, plus rarement, une forme circulaire). Des erreurs d'omission ont été constatées sur des formes sinueuses, notamment lorsqu'elles sont fines (environ 20 mètres), à l'image des routes.

La précision de la classe déboisée pour tous ces types de pertes forestières est supérieure à 91%, et atteint 100% pour l'orpaillage. Cela signifie que le nombre de fausses alarmes est faible. La classe de forêt intacte a une précision et un rappel supérieurs à 97% dans tous les cas.

1.2. COMPARAISON SPATIALE AVEC UN PRODUIT BASE SUR DES DONNEES OPTIQUES

Le produit GLAD-FA a été validé de la même façon que la carte produite avec la méthode basée sur les données Sentinel-1 afin de pouvoir en faire une comparaison avec un produit basé sur des données optiques. GLAD-FA est généré de manière automatique et à large échelle. Il est basé sur les données Landsat à 30 mètres de résolution, et a une surface de détection minimum de 0.09 ha. Pour rappel, la méthode basée sur les données Sentinel-1 GRD dégradées à environ 30 mètres de résolution après le filtrage a une surface de détection minimum d'environ 0.2 ha (voir discussion chapitre 5, section 1.1.6).

D'une manière générale, GLAD-FA détecte moins de déboisements que la méthode basée sur Sentinel-1 (cf. Tableau 5, rappel de 36.6% contre 81.5% respectivement). Cela s'explique en grande partie par la grande sensibilité des images Landsat à la couverture nuageuse, qui est le facteur limitant pour l'exploitation des données optiques sur la Guyane.

La précision globale de GLAD-FA est de 93.7%. L'avantage de ce produit est qu'il fait très peu d'erreurs de commission (1.9%), au détriment du nombre de détections.

Les rappels de GLAD-FA sur l'exploitation forestière et l'orpaillage pour la classe déboisée sont de 37% et 40% respectivement. Sur la première, les erreurs d'omission étaient également réparties entre les routes et les parcs à grumes.

Sur l'exploitation aurifère, les deux tiers des omissions concernaient des parcelles de petite surface, ou des formes longues et minces ressemblant aux routes. L'autre tiers concernait les grandes exploitations qui ont une forme sinueuse mais large, et qui étaient partiellement détectées.

Sur ces grands sites d'exploitation aurifère, la forêt est remplacée par un sol nu et par des bassins d'eau qui peuvent rester visibles pendant plusieurs années. Sur ces sites, la moitié des points de validation « déboisés » correctement détectés par les données Sentinel-1 et manqués par les données Landsat se sont produits pendant la saison humide, et l'autre moitié pendant la saison sèche. Ces déboisements laissant place à un sol nu ou à des bassins d'eau, et restant « visibles » plusieurs mois sur les images optiques, montrent qu'en plus d'apporter des délais temporels de détection, la persistance de la couverture nuageuse peut également rendre difficile la présence de deux images optiques exploitables dans une fenêtre de 180 jours sur certains sites, malgré la présence de la saison sèche.

Plus surprenant, GLAD-FA détecte seulement 31.9% des déboisements liés à l'agriculture de subsistance, alors que la méthode basée sur Sentinel-1 en détecte 89.4%. Ce type de pertes forestières se produit presque seulement pendant la saison sèche, donc avec une couverture nuageuse moindre. Le nombre d'acquisitions optiques par ciel clair devrait donc être voisin de celui des acquisitions Sentinel-1 pendant cette période. Cependant, bien que réduite, la couverture nuageuse est encore présente, et affecte la performance de GLAD-FA (qui nécessite au moins deux acquisitions sur lesquelles le déboisement est visible dans un délai de 180 jours pour lever une alerte). En effet, la présence de nuages dans les quelques images disponibles post-déboisement (notamment pour les parcelles abattues dans la deuxième moitié de la saison sèche (octobre)) diminue leurs chances d'être « vues » deux fois avant la fin de la saison sèche.

Il est à noter qu'en utilisant les données Sentinel-2, en complément, ou à la place des données Landsat, la probabilité d'avoir une acquisition optique sans nuage serait plus importante, car celui-ci à une résolution temporelle de 5 jours, contre 8 jours pour Landsat 7 et 8 combinés. Cela montre cependant que sur des territoires tels que la Guyane, la couverture nuageuse peut bloquer les détections, y compris en saison sèche.

D'autre part, sur d'autres parcelles d'agriculture vivrière non-détectées, il est possible de détecter visuellement le déboisement sur deux images Landsat séparées de moins de 180 jours, mais sans que la méthode ne lève d'alertes (voir exemple Figure 60). Il est difficile de connaître la raison de ces omissions, mais d'une manière générale, les méthodes automatiques et opérant à large échelle sont souvent moins précises que les méthodes adaptées très localement, ou nécessitant une intervention humaine.

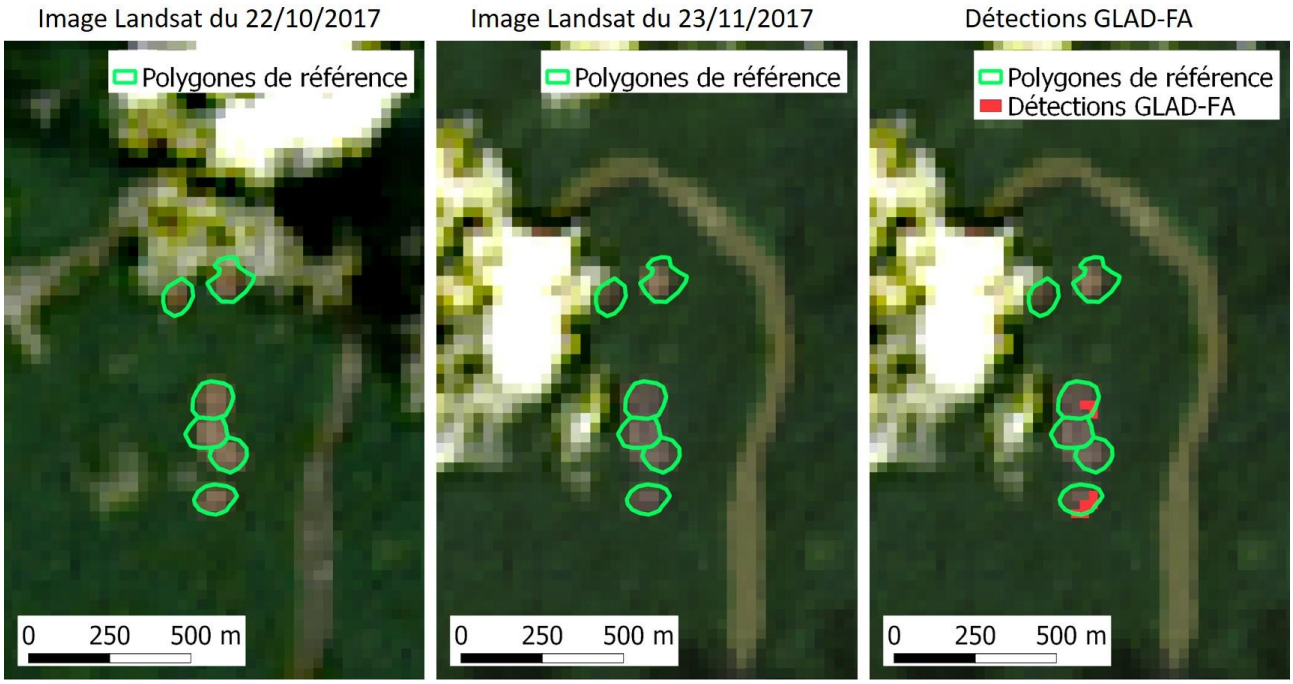


Figure 60 – Images Landsat du 22/10/2017 (gauche) et du 23/11/2017 (milieu) sur lesquelles on voit un abattis récemment coupé, et détections GLAD-FA associées en rouge (droite). Cartes centrées sur 3°2 N et 52°4 O.

De manière qualitative, il est possible d'observer que la forme des surfaces détectées est globalement moins précise avec GLAD-FA, comme le montre la Figure 61. Une analyse plus poussée et pas à pas de la méthode GLAD-FA permettrait d'en comprendre les raisons. Cependant, cette étude n'a pas été faite dans le cadre de cette thèse, car elle nécessiterait l'implémentation de la méthode et l'analyse de chacune de ses étapes intermédiaires, ce qui représente un coût en temps trop important.

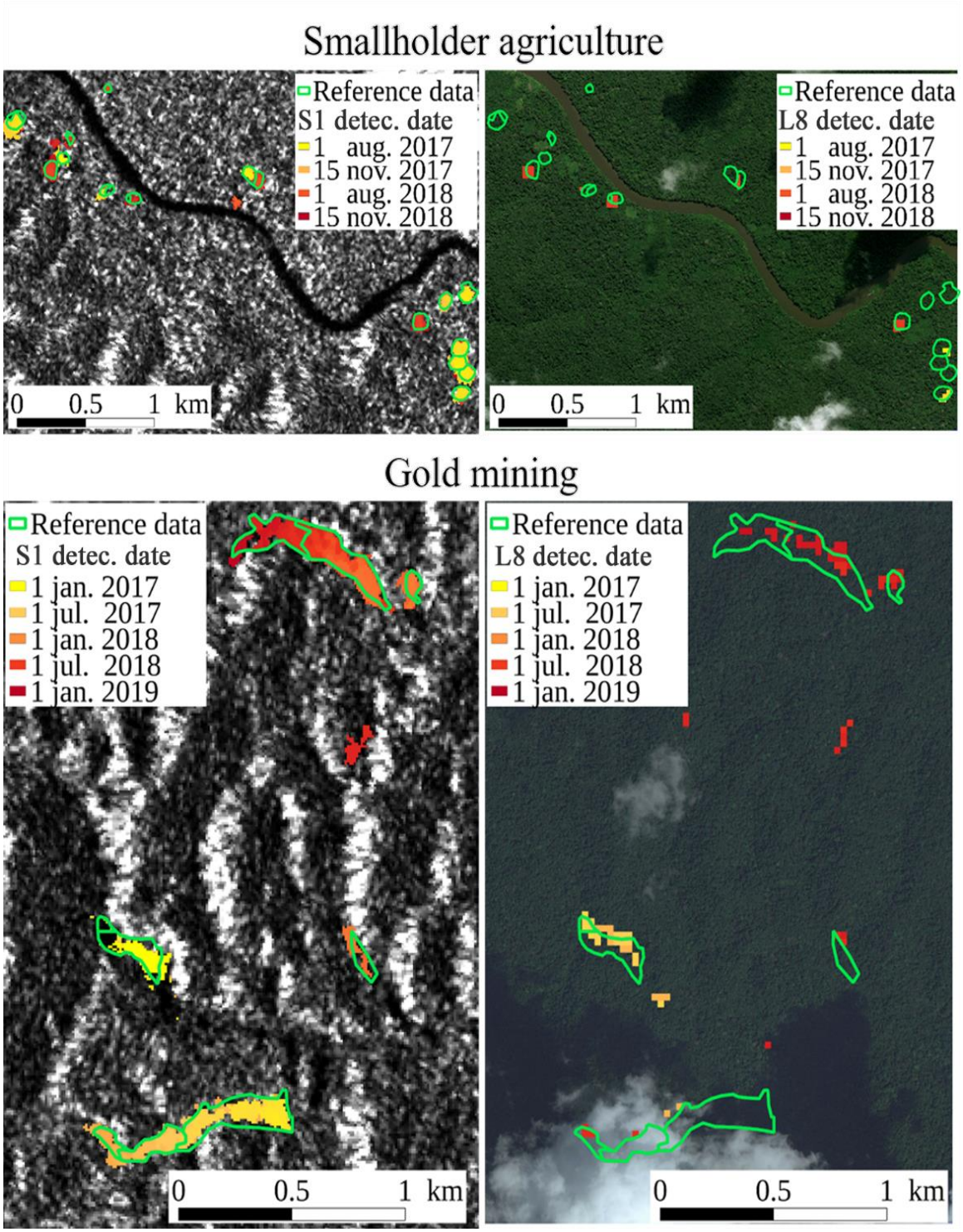


Figure 61 - Comparaison spatiale des cartes de perte de forêt basée sur Sentinel-1 (S1) (RSO) (à gauche - Fond : image Sentinel-1) et de GLAD-FA basé sur Landsat (L7/8) (optique) (à droite - Fond : image optique aérienne Bing). Les données de référence sont dessinées en vert. Les deux cartes supérieures sont centrées sur 3°12 N et 52°25 O.

Pour illustrer et évaluer ce phénomène, la Figure 62 montre l'histogramme du pourcentage de la surface détectée par chaque méthode pour chaque polygone de référence. La Figure 62 montre que la méthode basée sur Sentinel-1 tend à délimiter plus complètement les zones perturbées par rapport à GLAD-FA, confirmant l'évaluation qualitative.

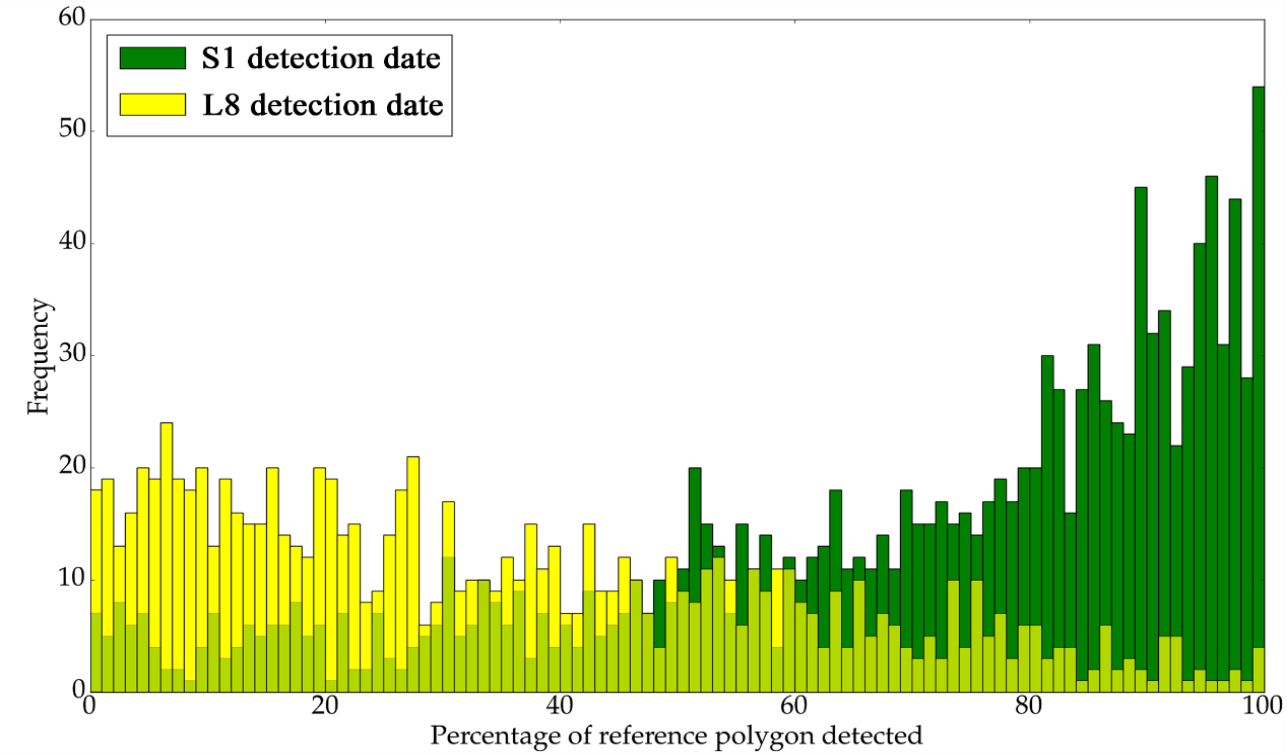


Figure 62 - Histogramme du pourcentage de polygones de référence couverts par les deux méthodes.

Au total, les détections basées sur Sentinel-1 et sur Landsat ont respectivement intersecté 1339 et 867 polygones de référence. 783 polygones de référence ont été détectés à la fois par les deux systèmes. Cela signifie que 84 polygones de référence ont été recoupés uniquement par des détections GLAD-FA et 556 uniquement par le produit basé sur Sentinel-1, ce qui met en évidence la complémentarité de ces deux méthodes de détection.

Pour résumer, il apparaît que le facteur limitant l'exploitation des données optiques pour le suivi du déboisement en Guyane soit la persistance de la couverture nuageuse. Le système optique omet plus de déboisement (rappel inférieur à 40%) par rapport au système radar (rappel supérieur à 75%), et ce quel que soit leur type. En effet, la méthode basée sur les données Sentinel-1 détecte mieux les pertes forestières, notamment grâce à sa résolution temporelle plus fine.

Certaines omissions de GLAD-FA sont inhérentes à la méthode. Cette méthode est automatique et opère à large échelle, et le traitement automatique de grandes zones est souvent moins précis que des méthodes appliquées localement ou nécessitant une intervention humaine.

Cependant, certaines omissions sont liées au capteur pour 3 raisons :

- **La résolution à 30 mètres de Landsat ne permet pas toujours de détecter les objets longs et fins tels que les routes (comme c'est le cas pour les données Sentinel-1) ;**
- **Sur certaines parcelles déboisées en saison humide, il y a assez de végétation qui repousse (avant la saison sèche) pour que l'optique confonde cette végétation avec des forêts, ce qui ne lui permet pas de détecter ces déboisements ;**
- **On pourrait s'attendre à ce que les déboisements récents ou laissant place à un sol nu soient très bien détectés en saison sèche car ils sont bien visibles sur les images optiques sans nuages. Cependant, il arrive que certains d'entre eux ne soient pas « vus » au moins deux fois sur les images optiques dans un délai de 180 jours à cause de la couverture nuageuse (notamment les abattis de fin de saison), ce qui empêche une méthode comme GLAD-FA de confirmer les alertes associées.**

Par ailleurs, GLAD-FA fait moins de fausses détections que la méthode basée sur Sentinel-1, même si la précision des détections reste élevée pour les deux systèmes (respectivement précision supérieure à 95.2% et supérieure à 91.1%).

Il est à noter que cette analyse concerne des parcelles déboisées relativement petites par rapport à celles que l'on retrouve dans des paysages avec des taux de déboisement élevés. Par conséquent, les résultats spatiaux des détections basées sur Landsat et présentés ici sont sûrement moins bons que sur d'autres types de déboisements portant sur des surfaces plus grandes.

1.3. EVALUATION TEMPORELLE

Les données temporelles existantes liées aux pertes forestières ne sont pas assez précises pour évaluer en absolu les méthodes de détection rapide. De ce fait, la date des détections des systèmes en quasi-temps réel est rarement validée dans les études. Pourtant, l'aspect temporel est primordial pour ce type de produit. C'est pourquoi nous avons mis en place une campagne terrain pour collecter des données temporelles de coupe (voir chapitre 3, section 3). Sur toutes les parcelles visitées, 27 étaient des abattis nouvellement créés (anciennement forêt dense) et seul l'un d'entre eux n'a pas été détecté par la méthode basée sur Sentinel-1.

Les dates de coupes ont été estimées en fonction d'intervalles de coupe définis par photo-interprétation d'images Sentinel-2 et des observations faites sur le terrain. Ces intervalles varient suivant l'exploitabilité des données optique sur la zone avant et après la coupe. Les parcelles ont été classées en 3 catégories suivant l'intervalle donné par les images Sentinel-2 : 5 jours, 10 jours ou supérieur à 1 mois (principalement le cas des parcelles coupées en début de saison sèche vers août, car le délai entre deux images exploitables est plus important à cette période qu'en milieu de saison sèche). La Figure 63 montre les dates de coupe observées par Sentinel-1 sur les 26 parcelles visitées comparées à la date de coupe déduite des observations faites sur place et des intervalles Sentinel-2. Il est possible d'observer que la date de coupe de la plupart des observations datées « précisément » (dans un intervalle de 5 jours) est estimée par Sentinel-1 dans un délai de 6 jours après la date « terrain ». Cela signifie que ces coupes sont visibles sur les images Sentinel-1 dès la première acquisition, malgré la végétation encore peu sèche restante au sol. Certaines parcelles sont détectées avant la date déduite sur le terrain : ceci provient d'une mauvaise estimation de la date « terrain » et les détections restent incluses dans l'intervalle défini par les images Sentinel-2.

Les parcelles dont la date déduite est dans un intervalle de plus de 30 jours ont un RSME de 18.8 jours. Celles ayant un intervalle de 5 ou 10 jours ont respectivement un RMSE de 3.42 et 4.71 jours. Si l'on considère toutes les parcelles, le délai médian est de 3.5 jours.

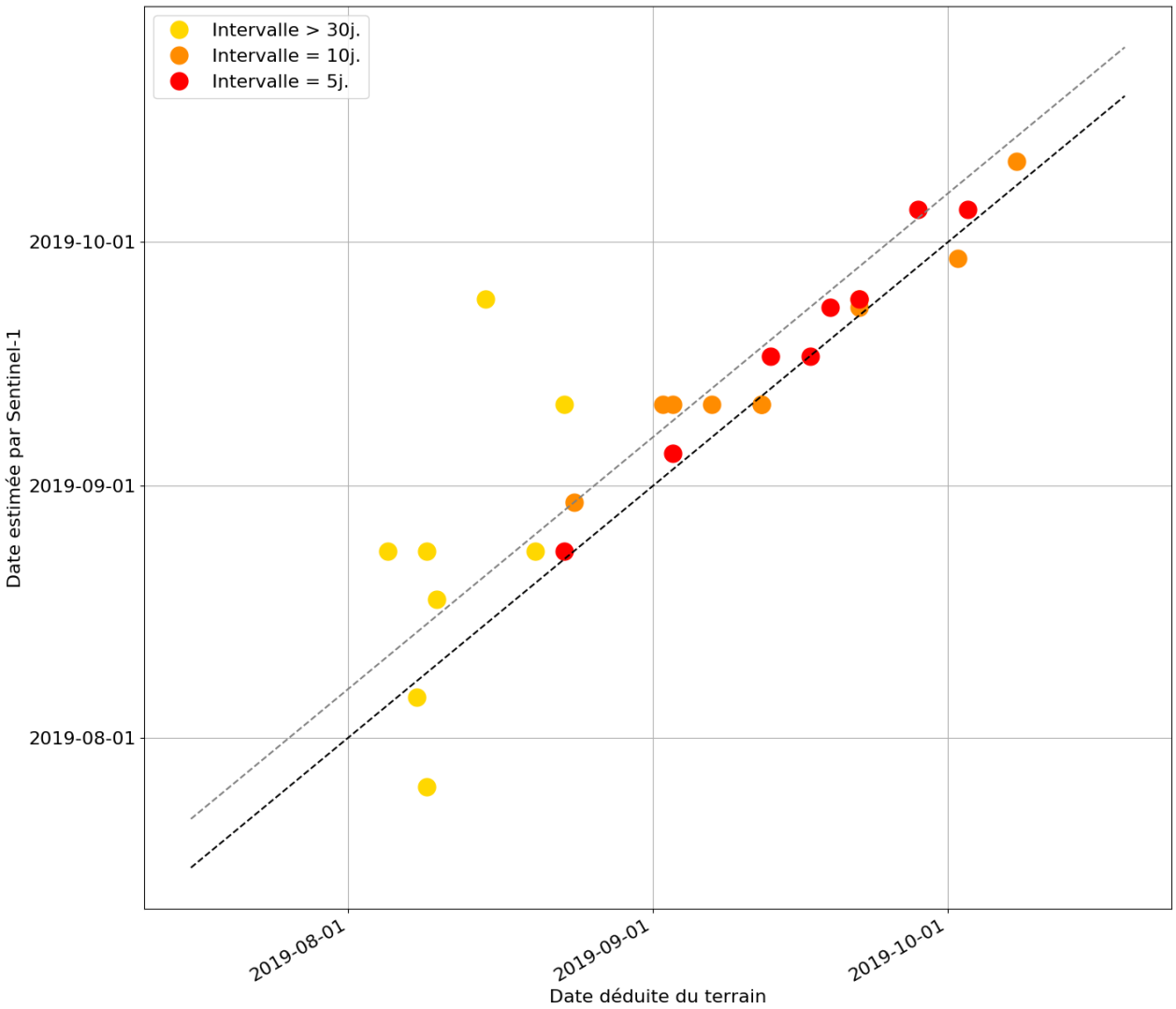


Figure 63 – Dates de coupe observées par la méthode basée sur Sentinel-1 des 26 parcelles visitées sur le terrain, comparées à la date de coupe déduite des observations faites sur place et des images Sentinel-2. La ligne en pointillés noire représente une date observée par Sentinel-1 égale à une date déduite du terrain. La ligne en pointillés grise représente une date observée par Sentinel-1 6 jours après la date de terrain déduite : ce délai peut être attribué aux acquisitions de Sentinel-1.

La taille du jeu de données de référence ne permet pas de conclure sur le retard de détection de la méthode à large échelle. Cependant, ces observations ont montré que la coupe apparaissait sur la première acquisition Sentinel-1 suivant la date « terrain » pour celles contenues dans des intervalles réduits. Cela signifie qu'en saison sèche, la méthode est capable de détecter une coupe « immédiatement » même s'il reste de la végétation au sol.

1.4. COMPARAISON TEMPORELLE

Afin d'évaluer les résultats temporels à plus large échelle, une comparaison entre les dates de coupe estimées par le système basé sur Sentinel-1 et le système basé sur Landsat GLAD-FA a été effectuée. L'objectif est de déterminer quel produit est le plus proche des dates réelles de coupe. Il existe des causes bien connues qui peuvent conduire à un retard de détection, par exemple, la couverture nuageuse pour les données optiques, et la végétation restante au sol pour les données radar. En revanche, aucune cause identifiée ne peut conduire à des détections anticipées (avant l'occurrence réelle du déboisement). Par conséquent, nous posons l'hypothèse que les détections les plus précoces sont plus proches de la véritable date de coupe. Pour cette comparaison temporelle, seuls les polygones de référence détectés par les deux capteurs sont utilisés.

1.4.1. AGRICULTURE VIVRIERE

Sur l'agriculture vivrière, 410 polygones ont été détectés par les données Sentinel-1, 243 polygones ont été détectés par les données Landsat, et 210 polygones ont été détectés conjointement par les deux méthodes.

Dans la Figure 64, le graphique de gauche représente le nombre de polygones détectés mensuellement (en vert pour les données Sentinel-1 et en jaune pour les données Landsat) : les deux algorithmes détectent le déboisement pendant la saison sèche (de juillet à novembre), ce qui est cohérent avec le calendrier de l'agriculture de subsistance. Pendant la saison sèche, la couverture nuageuse est réduite, tout comme le décalage temporel entre les systèmes.

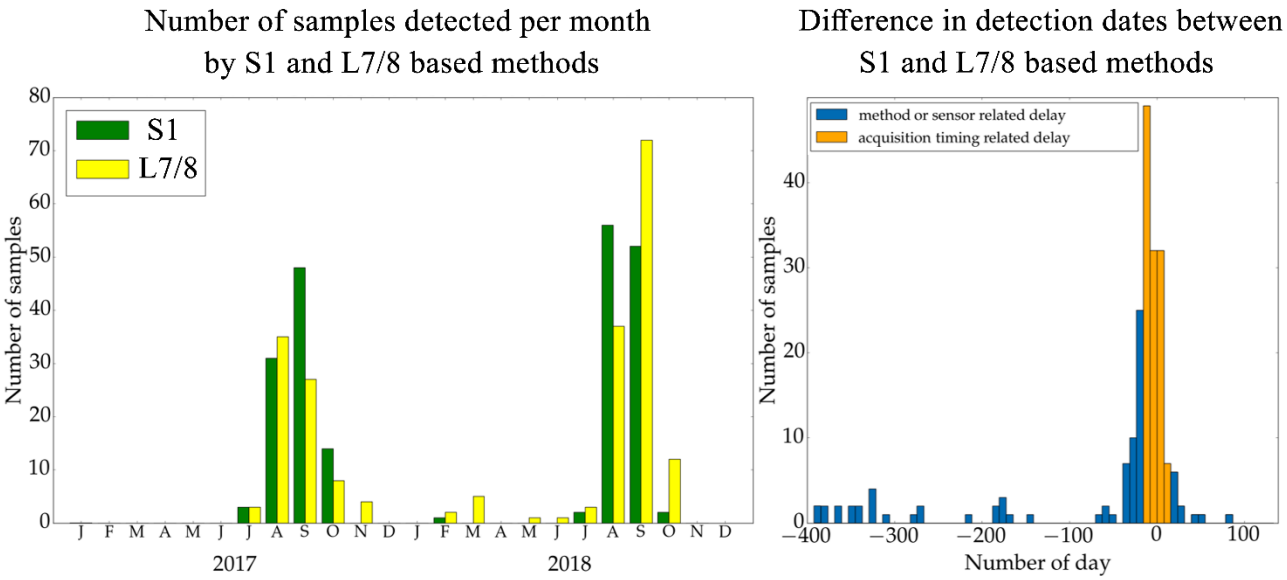


Figure 64 – Agriculture vivrière : nombre d'échantillons de référence détectés par mois pour les systèmes d'alertes forestières Sentinel-1 (S1) (vert) et Landsat (L7/8) (jaune) (gauche). Histogramme de la différence entre les dates de détection des polygones de référence, où les valeurs négatives indiquent une détection plus précoce par le système radar (droite) : en orange les délais pouvant provenir de l'écart entre les acquisitions de chaque capteur ; en bleu, les délais liés aux différences inhérentes aux capteurs ou aux méthodes utilisées.

Le graphique de droite de la Figure 64 montre que les différences de dates entre le produit basé sur les données Sentinel-1 et celui basé sur les données Landsat sont généralement distribuées autour de zéro (avec une moyenne de 20 jours plus tôt pour Sentinel-1). La partie positive de l'abscisse montre les polygones détectés en premier par les données Landsat, et vice versa.

Dans environ 50% des cas, le délai des détections peut être expliqué par le décalage d'acquisition de chaque capteur par rapport à l'événement, plutôt que par les limitations intrinsèques de chacun (en orange sur la Figure 64). Nous considérons ici la répétitivité de Sentinel-1 (12 jours) et de Landsat (16 jours).

Les parcelles détectées un an plus tôt par le système basé sur Sentinel-1 (10%) sont des parcelles exploitées pour la première fois en 2017 et détectées par les données Landsat seulement pendant la saison sèche suivante, lors de la deuxième coupe de la parcelle. Celles qui sont détectées environ 200 jours avant par le produit basé sur Sentinel-1 correspondent à un retard de détection de GLAD-FA dû à la couverture nuageuse : exploitées fin 2017, elles n'ont été détectées que pendant le petit été de mars 2018.

Au final, le système optique détecte les polygones en retard par rapport au système radar, bien que cette différence temporelle ne soit pas significative pour ce type de déboisement. Ces détections proches temporellement, dans un contexte favorable, permettent d'être confiant quant à l'estimation de la date de coupe fournie par les deux produits.

1.4.2. EXPLOITATION FORESTIERE

Sur l'exploitation forestière, 51 et 32 polygones ont été détectés respectivement par les données radar et optiques. Vingt-huit polygones ont été détectés conjointement par les deux systèmes. La plupart de ces polygones étaient liés à des coupes survenues en 2017, année sur laquelle davantage de données de référence étaient disponibles. La majorité des alertes ont été levées par les deux systèmes pendant la saison sèche et le petit été de mars (Figure 65). Cela correspond au moment où se produisent la plupart des activités créant des ouvertures détectables dans la canopée.

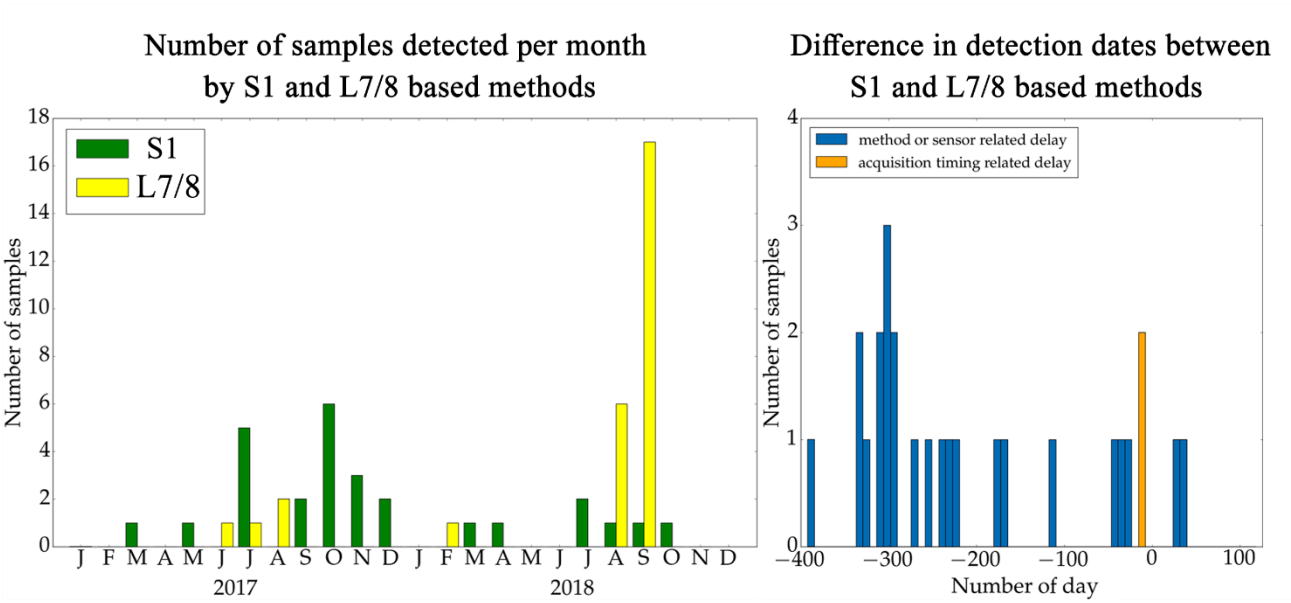


Figure 65 - Exploitation forestière : nombre d'échantillons de référence détectés par mois pour les systèmes d'alertes forestières Sentinel-1 (S1) (vert) et Landsat (L7/8) (jaune) (gauche). Histogramme de la différence entre les dates de détection des polygones de référence, où les valeurs négatives indiquent une détection plus précoce dans le système radar (droite) : en orange les délais pouvant provenir de l'écart entre les acquisitions de chaque capteur ; en bleu, les délais liés aux différences inhérentes aux capteurs ou aux méthodes utilisées.

Plus de 90% des zones déboisées sont d'abord détectées par le système basé sur le radar (Figure 65). Pour deux parcelles seulement, le retard de GLAD-FA s'explique par la fréquence d'acquisitions des images Landsat. Pour les autres, le retard est attribuable au capteur en lui-même : la forêt de Regina/Saint-Georges (lieu de validation de l'exploitation forestière) fait partie des zones les plus ennuagées de Guyane, avec une couverture nuageuse supérieure à 65% sur l'année (chapitre 5, Figure 56). Ceci explique le retard plus important des détections optiques sur ce site de validation par rapport à celui de l'agriculture.

1.4.3. ORPAILLAGE

En ce qui concerne l'extraction aurifère, 878 polygones ont été détectés par le système basé sur Sentinel-1, et 592 par les données Landsat. 545 polygones de référence ont été détectés conjointement par les deux produits.

L'orpaillage a lieu toute l'année, à un rythme qui varie en fonction des conditions. En effet, l'exploitation illégale dépend du niveau d'eau à la fois pour le fret fluvial et pour le fonctionnement des pompes utilisées pour l'orpaillage alluvionnaire. Or, ces sites peuvent se retrouver noyés lorsque le niveau d'eau monte pendant les fortes pluies d'avril à juin. Pour l'orpaillage légal, le démarrage des chantiers est plus facile en saison sèche, bien que les équipements permettent également de travailler en saison humide.

Sans surprise, les données radar détectent ces pertes forestières tout au long de l'année (Figure 66, gauche), tandis que les données Landsat les détectent presque toutes en saison sèche. Ceci est cohérent avec la capacité de chacun des capteurs.

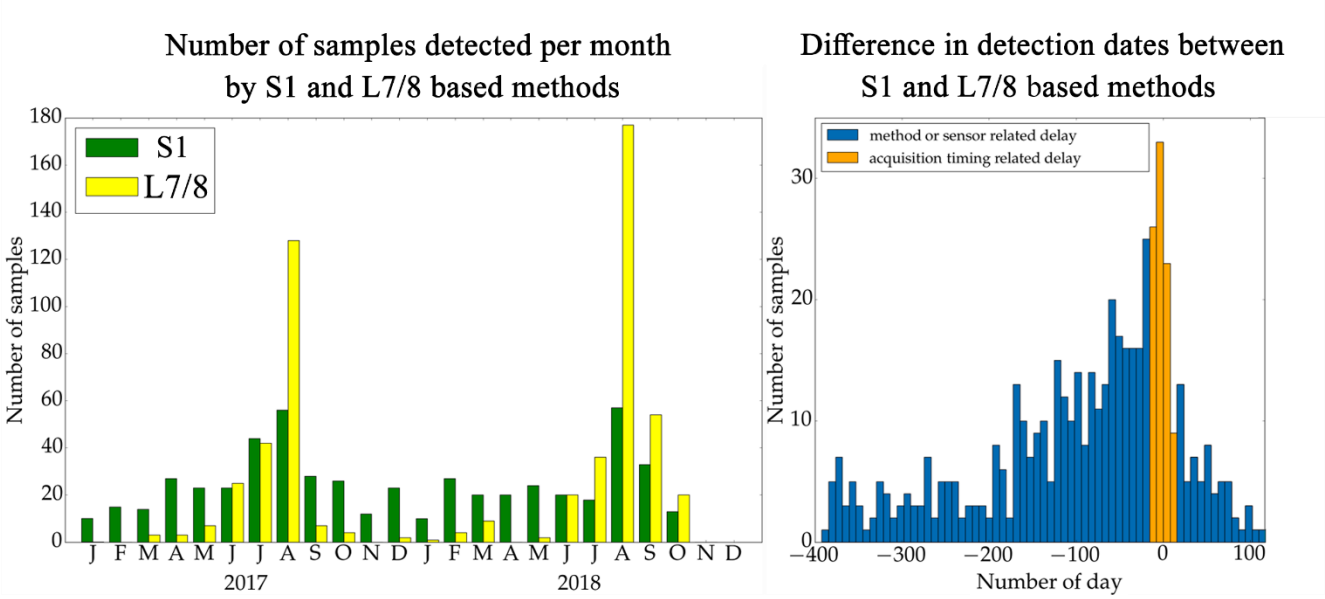


Figure 66 - Orpaillage : nombre d'échantillons de référence détectés par mois pour les systèmes d'alertes forestières Sentinel-1 (S1) (vert) et Landsat (L7/8) (jaune) (gauche). Histogramme de la différence entre les dates de détection des polygones de référence, où les valeurs négatives indiquent une détection plus précoce dans le système radar (droite) : en orange les délais pouvant provenir de l'écart entre les acquisitions de chaque capteur ; en bleu, les délais liés aux différences inhérentes aux capteurs ou aux méthodes utilisées.

La Figure 66 (droite) montre une bien meilleure précision temporelle des détections via l'imagerie radar : 80% des polygones de référence sont détectés en premier par la méthode basée sur Sentinel-1, et 40% d'entre eux sont détectés par les données Landsat avec un retard de plus de 3 mois. La majorité de ces retards est due aux limitations du capteur optique liées à la couverture nuageuse, ce retard augmentant pendant la saison humide.

Parmi les polygones de référence, 13% de ceux qui sont détectés en premier par le produit basé sur Sentinel-1 peuvent être dus au délai lié aux acquisitions de chaque capteur. Ce chiffre est de 23% pour les polygones détectés en premier par les données Landsat.

De manière globale, les données radar estiment mieux les dates de coupes des perturbations forestières et sont plus adéquates pour lancer des alertes rapidement. Pour résumer :

- **Les observations réalisées dans des conditions de détection optimales pour les deux jeux de données produits ont montré un bon accord entre les dates estimées par les systèmes radar et optique ;**
- **Dans les périodes où la couverture nuageuse est importante, l'imagerie radar est la seule à pouvoir faire des détections.**

Ces exemples mettent en évidence les avantages de l'imagerie radar par rapport à l'imagerie optique dans les tropiques, en particulier dans un contexte de quasi-temps réel, le facteur limitant de l'imagerie optique étant essentiellement lié à la couverture nuageuse.

1.5. REDUCTION DU DELAI DE DETECTION

Le but de cette étude est d'évaluer si le délai de détection causé par le nombre d'images utilisées après la date de détection qu'on cherche à estimer (X_a, voir chapitre 5, section 1.1.3) choisi peut être abaissé sans perdre en précision : c'est-à-dire quantifier le nombre de fausses alarmes qui apparaîtraient en diminuant X_a.

La qualité du RCR peut être évaluée dans un premier temps par sa variation sur la forêt. La moyenne et l'écart-type du RCR pour un pixel calculé sur l'année 2020 sont donnés dans le Tableau 6, moyennés sur plusieurs pixels de forêt intacte en fonction de X_a. Ces pixels ont été choisis par interprétation visuelle d'images biannuelles Planet (www.planet.com/nicfi) et couvrent plusieurs zones réparties dans les tuiles 22NCJ et 22NCK (voir chapitre 3, Figure 22), pour une surface totale de 65 338 ha.

Tableau 6 - Moyenne de la moyenne et de l'écart-type du RCR

X_a	Moyenne	Ecart-type
1	-0.21	1.21
2	-0.09	0.86
3	-0.05	0.68

Le Tableau 6 montre que l'écart-type double entre $X_a = 3$ et $X_a = 1$. Cela laisse présager que $X_a = 1$ est trop instable pour obtenir une information fiable. La Figure 67 montre le RCR de 8 pixels tirés aléatoirement dans la couche de forêt intacte. On observe que plusieurs pixels passent le seuil β (égal à -3 dB et en bleu sur la Figure 67) à plusieurs reprises. Ce seuil qui permet de détecter les zones potentiellement déboisées peut également permettre de détecter des fausses alarmes, mais la proportion de fausses alarmes apparait très importante et régulière. Ces fausses alarmes risquent soit de s'agréger à des vraies détections, déformant ainsi la surface réelle de la parcelle déboisée, soit de former une fausse détection avec une ombre détectée à cause de la variabilité du RCR. De plus, les évènements exceptionnels comme les grosses pluies (qui peuvent parfois apparaitre sur les images RSO), peuvent faire chuter le signal drastiquement sur une image et seront détectés car le signal post-perturbations ne sera calculé que sur cette image. Pour l'ensemble de ces raisons, nous décidons d'écarter la solution $X_a = 1$ pour se concentrer sur $X_a = 2$ et $X_a = 3$.

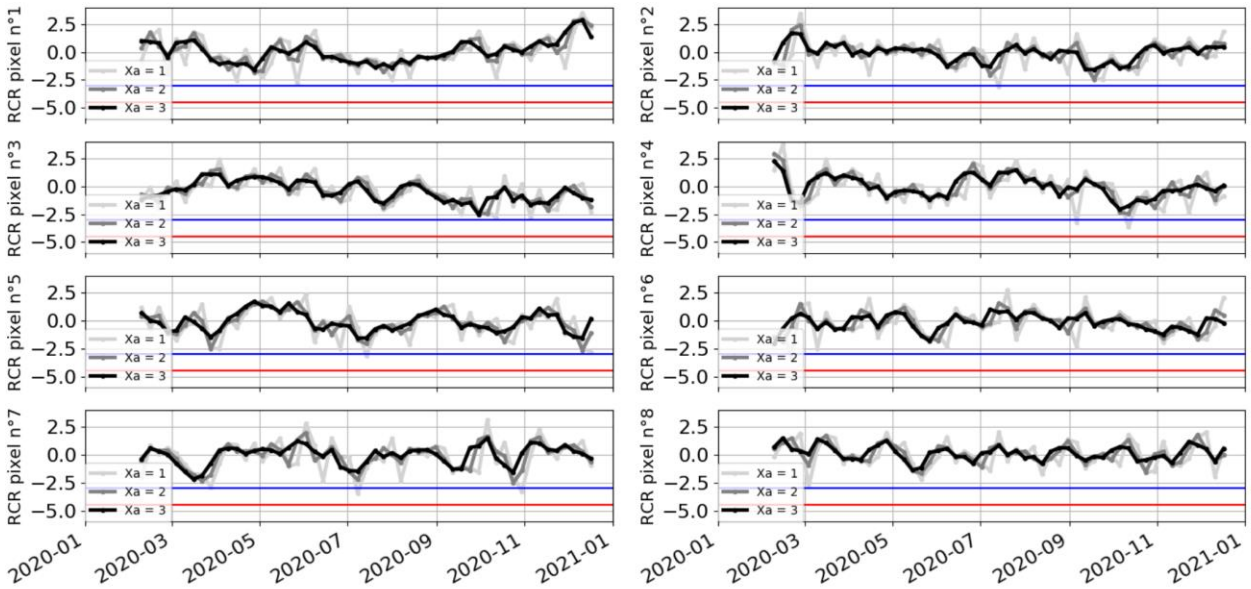


Figure 67 – RCR en fonction de X_a sur des pixels de forêt intacte tirés aléatoirement.

Sur les détections cumulées produites sur 2020, la proportion de fausses alarmes (au pixel) dans la zone de forêt intacte est de 0.009% pour $X_a = 3$ et 0.237% pour $X_a = 2$ (rapport ×25). La détection de fausses alarmes avec $X_a = 2$ revêt plusieurs explications. Une partie d'entre elles est due à un évènement pluvieux important présent à la date du 18/10/2020 comme le met en évidence la Figure 68.

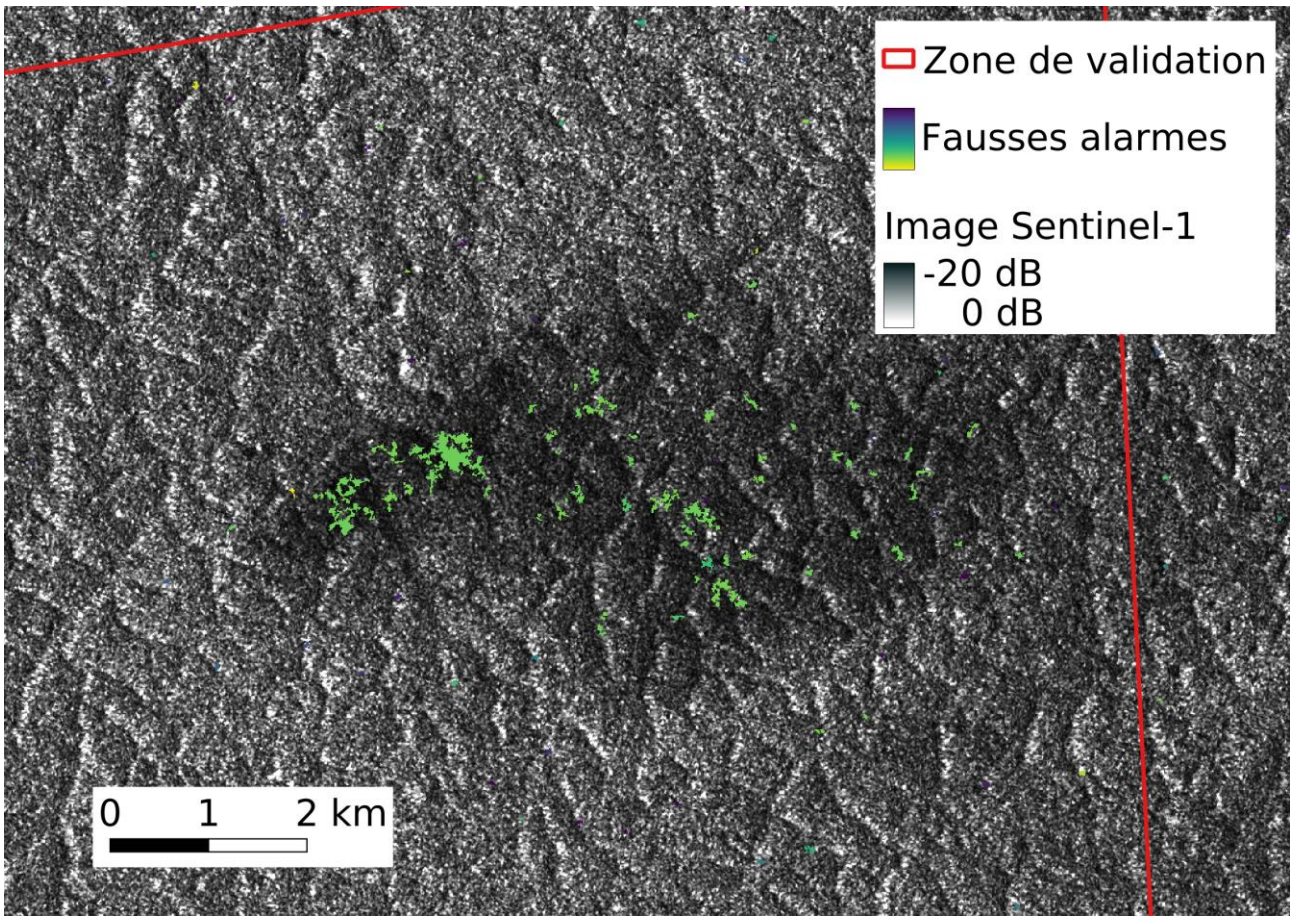


Figure 68 – Fausses alarmes dues à la présence d'un évènement pluvieux important sur l'image Sentinel-1 du 18/10/2020. Cet évènement pluvieux important est caractérisé par la masse noire visible au milieu de l'image.

Dans cette étude, 41% des fausses détections étaient dues à des évènements exceptionnels de ce type qui couvraient eux-mêmes une surface de 1 728 ha (2.64% de la zone d'étude totale). En enlevant ces fausses détections au calcul, la proportion de fausses détections avec $X_a = 2$ est de 0.14%, pratiquement 2 fois plus que pour $X_a = 3$.

Prendre X_a supérieur à 3 n'est pas envisagé car comme vu dans le chapitre 4, le signal sur une surface déboisée ne reste pas nécessairement bas dans le temps, et une valeur de X_a trop grande pourrait nuire à la baisse de signal post-perturbation dans la moyenne. De plus, le délai de détection d'allongerait, ce qui est contraire au besoin de détection rapide.

En résumé, réduire $X_a = 3$ ne permettrait pas d'avoir des résultats convenables de manière automatique (détection de gros nuages). Il est cependant envisageable de lever une alerte à $X_a = 2$ et de la confirmer à $X_a = 3$.

1.6. OBSERVATIONS

1.6.1. TAUX ANNUEL DE CHANGEMENTS FORESTIERS

La Figure 69 montre une comparaison des taux annuels de changement de la couverture forestière sur la Guyane à partir de la méthode basée sur Sentinel-1, GLAD-FA et des données de pertes forestières annuelles de GFW. Le taux annuel de changement est calculé en utilisant l'équation (4) de Puyravaud, 2003 comme suit :

$$r = \left(\frac{1}{t2 - t1}\right) \times ln(A2 - A1) \qquad 6.1$$

Avec A1 et A2 sont les surfaces de la forêt pour les années t1 et t2, respectivement. Le taux r est négatif lorsque la forêt perdue est supérieure à la forêt gagnée. Notez que le système basé sur Sentinel-1 et GLAD-FA produisent tous deux des détections dans un contexte de quasi-temps réel, alors que les données de GFW sont un produit annuel. Ce dernier a donc pour objectif de quantifier l'ampleur du changement sur une année et ainsi d'obtenir une vue exhaustive des surfaces déboisées contrairement à GLAD-FA.

Les détections basées sur Sentinel-1 montrent une diminution du taux de déboisement entre 2016 et 2017 (respectivement 5 619 et 3 169 ha déboisés) et une stabilisation de 2017 à 2018 (3 430 ha déboisés) (en vert dans la Figure 69). Les mêmes tendances se retrouvent dans les données de GFW. Comme nous l'avons vu dans la section 1.2, le système basé sur le radar a soulevé plus de d'alertes que le système basé sur l'optique, ce qui peut être attribué aux conditions de couverture nuageuse.

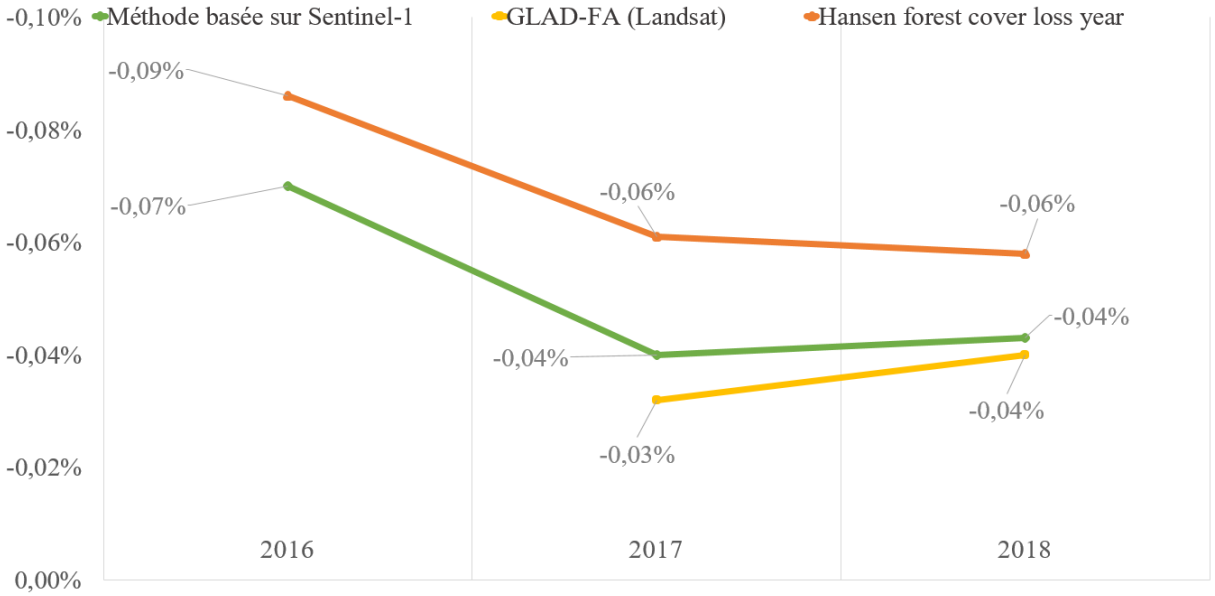


Figure 69 - Taux de variation du couvert forestier en Guyane, donné par différentes méthodes. La méthode basée sur Sentinel-1 est en vert, GLAD-FA en jaune et l'année de perte de couvert forestier de (Hansen et al., 2013) en orange.

Ce graphique illustre bien le bon accord global des différentes estimations : elles suivent toutes la même tendance.

1.6.2. SURFACE DES PARCELLES DETECTEES

La Figure 70 montre les pertes forestières estimées à partir des données Sentinel-1 pour différentes tailles de parcelles en Guyane. Ceci souligne l'importance des petites zones déboisées, y compris celles provenant de l'orpaillage illégal et de l'agriculture vivrière. Les petites parcelles (inférieures à 2 ha) contribuent à la moitié de la surface de la perte forestière.

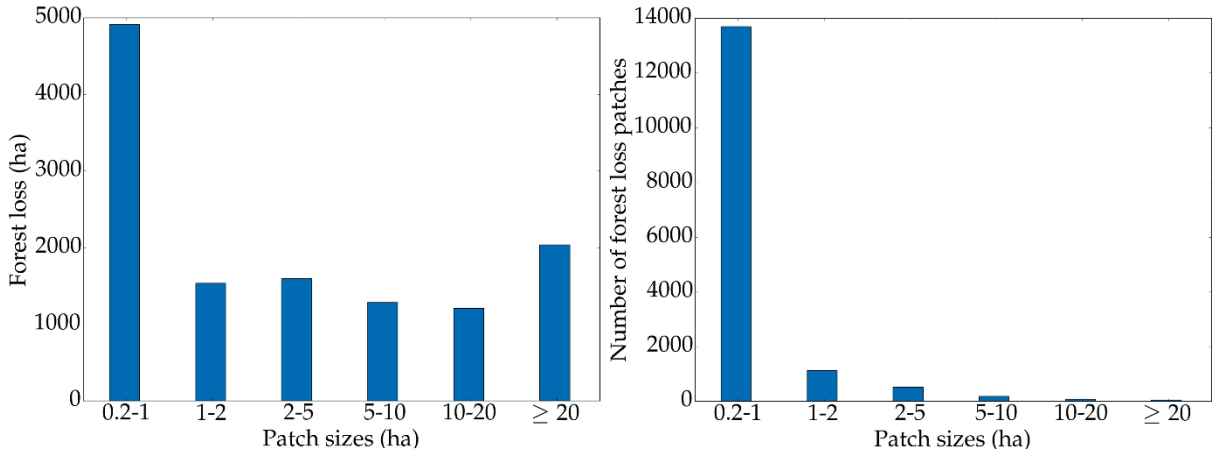


Figure 70 - Surface déboisée (ha) d'après les tailles de parcelles (gauche). Nombre de parcelles déboisées d'après leur surface (droite). Ces données proviennent des détections effectuées avec les données Sentinel-1 en Guyane sur 2016, 2017 et 2018.

Ce résultat souligne la nécessité de disposer d'un système de détection des pertes forestières à haute résolution, comme cela a été avancé dans (Kalamandeen et al., 2018).

2. Résultats – Caractérisation des déboisements

Cette partie présente le modèle (arbre de décision) obtenu pour caractériser le déboisement en quasi-temps réel en Guyane, en suivant la procédure décrite dans le chapitre 5. Cet arbre a été appliqué et validé à l'échelle de la Guyane sur les déboisements de l'année 2018.

2.1. CREATION DU MODELE

2.1.1. IMPORTANCE DES INDICATEURS

Dans un premier temps, nous observons pour chaque classe les indicateurs qui permettent de les distinguer. Un indicateur sera utile s'il permet de différencier les variables d'une classe par rapport aux autres. Les boîtes à moustaches et les graphiques en violon sont des représentations graphiques de données statistiques. Les premières permettent de visualiser rapidement les caractéristiques d'un ensemble de données (médiane, moyenne, etc.), tandis que les deuxièmes informent sur leur distribution. Les boîtes à moustaches et les graphiques en violon des variables de chaque classe sur chaque indicateur sont disponibles en annexe. Grâce à cela, il est par exemple possible de déduire que la finesse (aire/périmètre^2) n'est pas un indicateur utile, car les données de toutes les classes ont à peu près la même distribution pour cet indicateur (voir Figure 71). En revanche, on voit que la distance aux abattis est susceptible de distinguer des autres les classes agriculture et urbanisation.

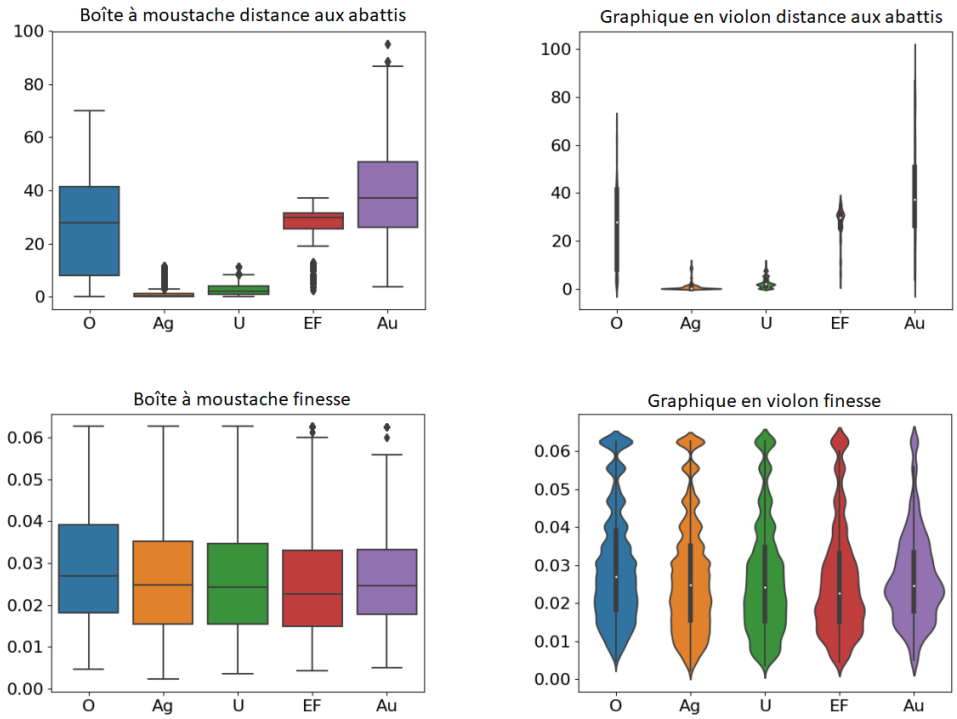


Figure 71 – Boîtes à moustache (gauche) et graphiques en violon (droite) de la distance aux abattis en km (haut) et de la finesse (bas), par type de déboisement (O : Orpaillage ; Ag : Agriculture ; U : Urbanisation ; EF : Exploitation forestière ; Au : Autre)

Les indicateurs ont été classés en fonction de leur importance relative pour différencier les facteurs de déboisement les uns des autres. La Figure 72 montre les 3 indicateurs les plus importants d'après les RF produites pour chaque classe.

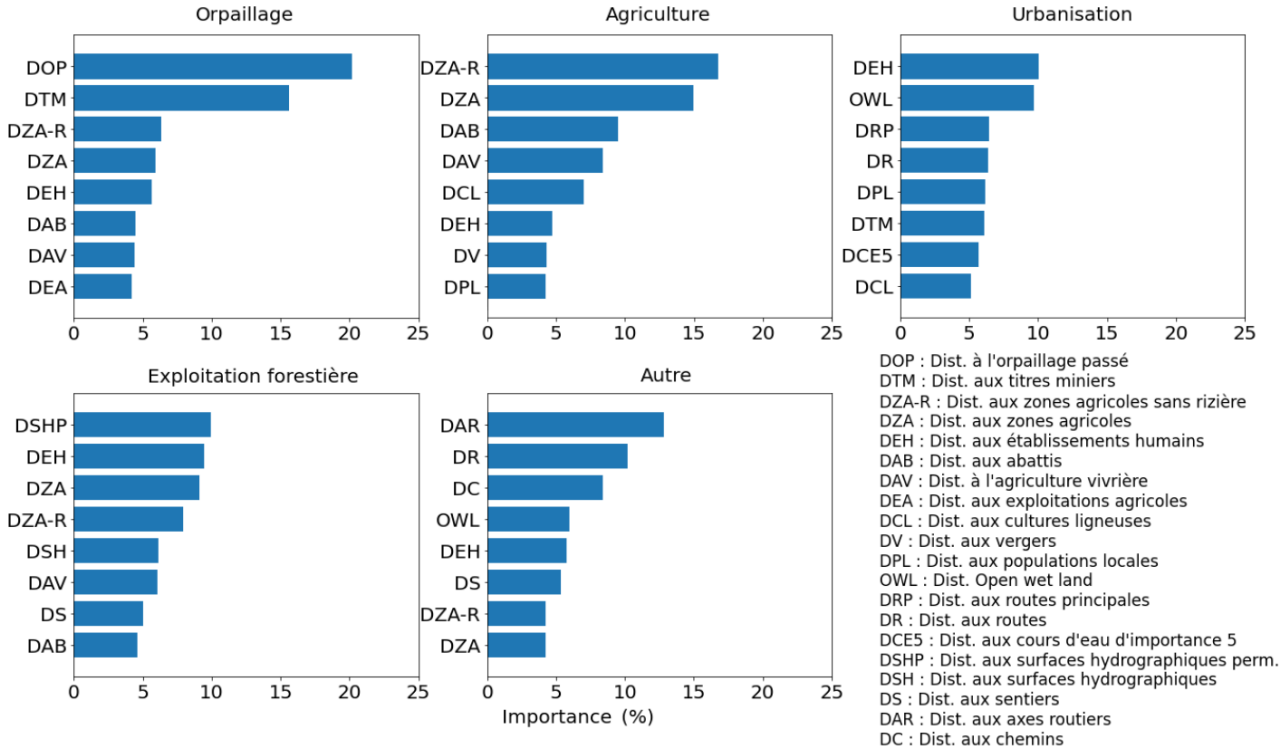


Figure 72 - Importance des indicateurs (les 8 plus importants) calculée sur des RF pour chaque classe.

La classe orpaillage se distingue très clairement de toutes les autres avec les indicateurs « distance à l'orpaillage passé » et « distance aux titres miniers » (voir Figure 72 pour leurs définitions). En effet, l'orpaillage légal a lieu exactement à l'emplacement des titres miniers et l'orpaillage illégal se trouve souvent près d'un autre site d'orpaillage (qui peut ne plus être exploité), les bons filons d'or étant revisités. La « distance aux zones agricoles hors rizières » permet de distinguer l'exploitation aurifère de l'agriculture et l'urbanisation.

Les indicateurs liés à l'agriculture « distance aux zones agricoles hors rizières », « distance aux zones agricoles » et « distance aux abattis passés » les séparent bien du reste, car comme pour l'orpaillage, les nouvelles parcelles agricoles se trouvent souvent dans des zones déjà exploitées. On remarque tout de même que la distribution des observations liées à l'urbanisation sont proches de celles liées à l'agriculture dans ces indicateurs.

L'indicateur « distance aux établissements humains » distingue particulièrement l'urbanisation des autres classes, même si encore une fois, la différence avec l'agriculture est moins importante que celle avec les autres classes. La couche « open wetland » d'ECOSEO est également intéressante pour distinguer l'urbanisation : c'est parce qu'elle décrit surtout des surfaces en eau

proches du littoral et donc de la majorité des populations. Pour finir, l'urbanisation a souvent lieu plus proche des routes principales que les autres facteurs de déboisement.

Les indicateurs importants ressortis pour l'exploitation forestière sont la « distance aux établissements humains », la « distance aux surfaces hydrographiques permanentes » et la « distance aux zones agricoles ». Ils n'apparaissent pas comme étant des caractéristiques propres à l'exploitation, et ont dû être sélectionnées car ils permettaient d'écarter certaines variables (la « distance aux établissements humains » et la « distance aux zones agricoles » écartent agriculture et urbanisation et la « distance aux surfaces hydrographiques permanentes » écartent tout sauf « autre »).

Les déboisements naturels se produisent sur tout le territoire. La plupart du temps, leur distribution (sur les distances) est très étalée. Les indicateurs importants ressortis pour cette classe sont la « distance aux axes routiers », la « distance aux chemins » et la « distance aux routes ». Ils indiquent une présence d'activités anthropiques et ont donc été sélectionnés car ils distinguent ces activités anthropiques des déboisements naturels.

Cette première analyse a été faite avec l'importance donnée par les RF ne prenant pas en compte la corrélation entre les indicateurs : elle permet donc d'avoir un point de vue des indicateurs importants pour chaque classe mais n'a pas pour but de répondre au problème (minimisation des indicateurs pour le modèle prédictif).

Comme l'objectif est d'avoir un arbre de décision traitant toutes les classes à la fois, un classement des indicateurs importants a été fait, cette fois, en prenant en compte toutes les classes en même temps en utilisant la méthode décrite chapitre 5, section 3.3 qui permet de réduire la redondance de l'information contenue dans les indicateurs.

Les 8 indicateurs trouvés comme étant les plus importants sont listés, dans l'ordre, dans le Tableau 7 :

Tableau 7 - Rang d'importance des indicateurs pour la caractérisation des déboisements en Guyane

1 – Distance à l'orpaillage passé
2 – Distance aux zones agricoles (hors rizières)
3 – Distance aux cultures ligneuses
4 – Distance aux établissements humains
5 – Distance aux exploitations agricoles
6 – Distance aux routes principales
7 – Distance aux titres miniers

8 – Distance aux rizières
9 – Distance aux surfaces hydrographiques artificielles
10 – Distance à la ceinture de roche verte du Paramaca
11 – Distance aux surfaces hydrographiques intermittentes
12 – Distance aux cours d'eau 4
13 – Distance aux surfaces hydrographiques naturelles

Parmi ces indicateurs, on retrouve les 3 indicateurs discriminants pour l'orpaillage (« distance à l'orpaillage passé », « distance aux zones agricoles hors rizières » et « distance aux titres miniers »). Pour l'agriculture, seule reste la distance aux zones agricoles (hors rizières) : les autres indicateurs importants pour cette classe ont été supprimés car ils lui étaient fortement corrélés (respectivement r = 0.99 et 0.87) et l'information était redondante. Sont aussi ressorties la distance aux établissements humains (importantes pour l'urbanisation et l'exploitation forestière) et la distance aux routes principales (importante pour l'urbanisation).

Les distances liées aux infrastructures routières importantes pour la classe « autre » ne sont pas ressorties, probablement parce qu'elles permettaient d'écarter les classes anthropiques, distinguées par d'autres indicateurs dans ce cas général.

Les indicateurs apparus intéressants dans le cas général sont la « distance aux cultures ligneuses » et la « distance aux rizières ». Ces indicateurs montrent une distribution différente pour chaque classe. Les rizières sont très localisées sur environ 67 km^2 au nord de la Guyane. Elles représentent donc la localisation de chaque facteur sur le territoire par rapport au nord. Par exemple, sur les graphiques en violon, le petit site d'exploitation forestière proche du nord apparait en premier, puis vient le site plus important situé dans la forêt de Régina/Saint Georges.

2.1.2. NOMBRE D'INDICATEURS ET TAILLE DE L'ARBRE CHOISI

Le score de chaque arbre de décision entraîné avec les k plus importants indicateurs pour chaque valeur de $k \in (1, 51)$ est montré Figure 73 pour des arbres de profondeur 3, 5, 7 et 10.

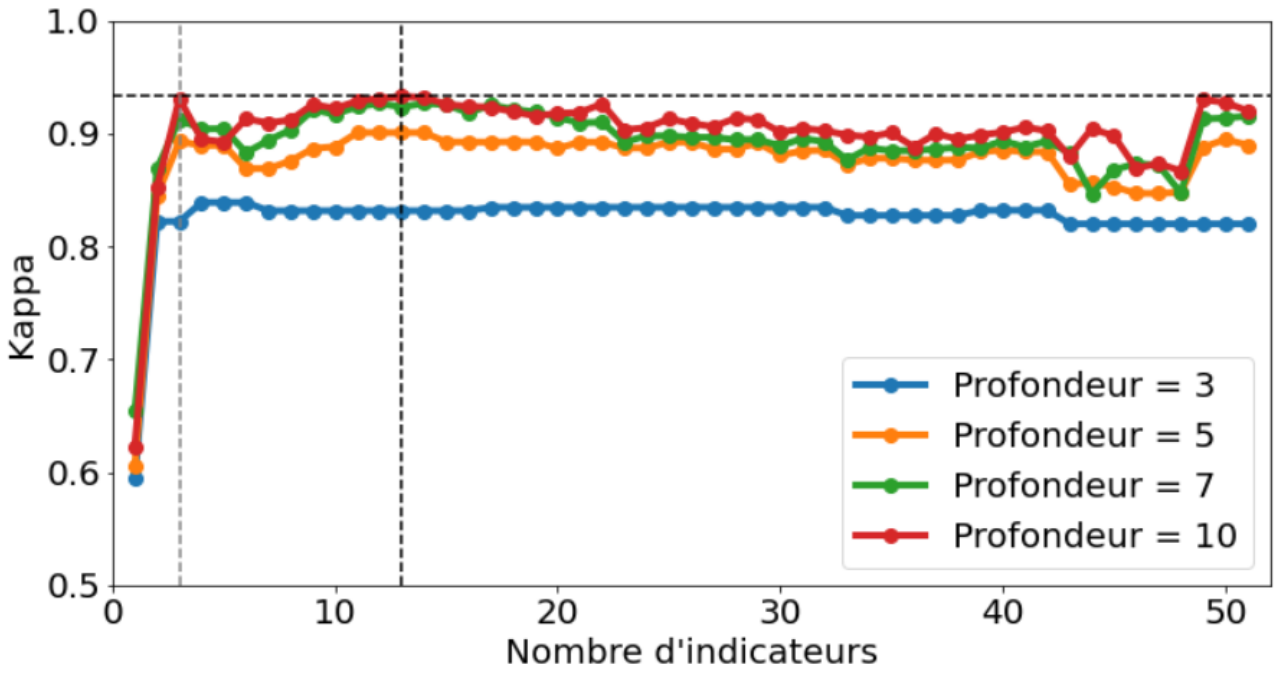


Figure 73 - Score kappa en fonction du nombre d'indicateurs fournis à l'arbre de décision pour son apprentissage, pour 4 profondeurs d'arbres différentes. Les lignes verticales montrent le nombre d'indicateurs avec les scores les plus élevés pour un arbre de profondeur de 10 (grise : 3 indicateurs – kappa = 0.931, noire : 13 indicateurs – kappa = 0.9328). La ligne noire verticale montre le score le plus élevé pour un arbre de profondeur 10

Le score augmente systématiquement avec la profondeur de l'arbre. Pour une profondeur de 10, le score maximum de 0.9328 est atteint en donnant 13 indicateurs les plus importants (voir Tableau 7) à l'arbre de décision pour son apprentissage. En utilisant seulement les 3 indicateurs les plus importants, le score est de 0.931 : le gain est significatif pour la préparation des variables descriptives par les acteurs locaux, et le résultat global en est peu altéré.

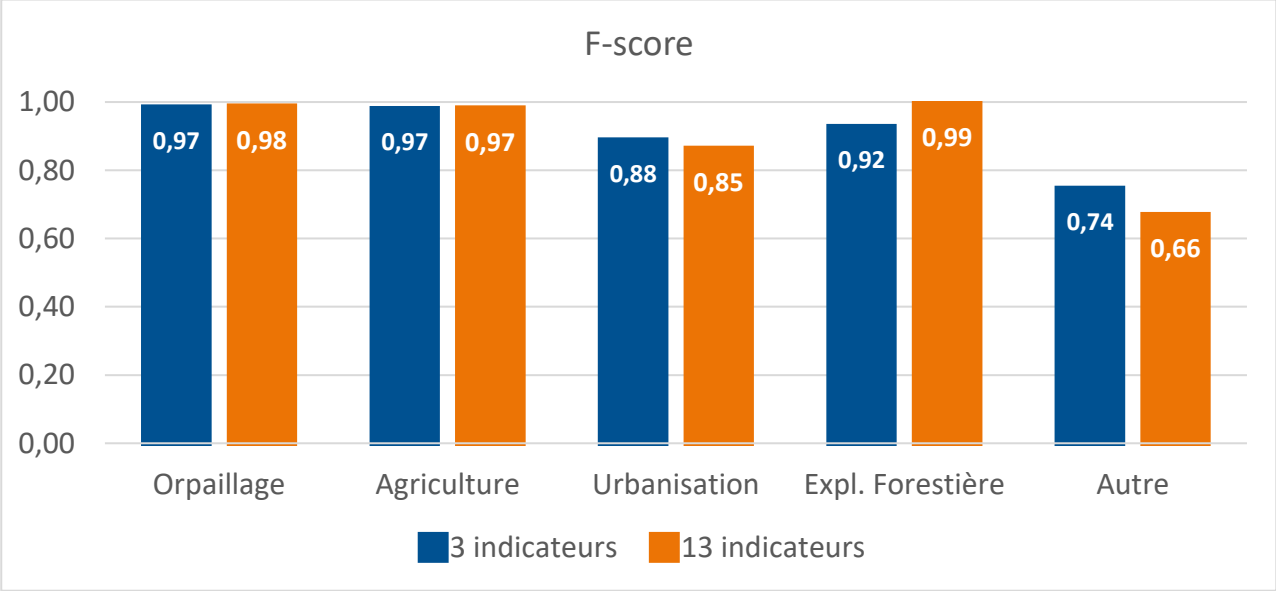


Figure 74 - F-score par classe en fonction pour un nombre d'indicateurs de 3 et de 13.

La Figure 74 montre le F-score de chaque classe pour le modèle utilisant les 3 meilleurs indicateurs et celui utilisant les 13 meilleurs.

Celui-ci est équivalent pour l'orpaillage et l'agriculture et montre une bonne classification de ces deux classes (F-score supérieur à 0.97).

Pour l'urbanisation et la classe « autre », le F-score de l'arbre avec 3 indicateurs est plus important que celui utilisant 13 indicateurs : 0.88 et 0.85 respectivement pour la première classe et 0.74 et 0.66 pour la deuxième. L'urbanisation montre une précision et un rappel plus élevés pour l'arbre utilisant 3 indicateurs (0.88 pour les deux), contre 0.84 et 0.87 pour l'arbre avec 13 indicateurs. Cela indique que l'urbanisation est mieux discriminée par l'arbre à 3 indicateurs. La classe autre, elle, montre une précision plus élevée (0.7 contre 0.53) pour l'arbre avec 3 indicateurs, mais un rappel plus faible (0.78 contre 0.88). Cela signifie que l'arbre avec 13 indicateurs prédit beaucoup la classe « autre » et trouve beaucoup de vrais échantillons de cette classe, mais il est erroné la moitié du temps.

Pour l'exploitation forestière, c'est l'arbre avec 13 indicateurs qui montre le F-score le plus élevé (0.99 contre 0.92). Sa précision et son rappel (0.99 et 0.98) sont plus élevés que ceux de l'arbre à 3 indicateurs (0.91 et 0.93), ce qui indique qu'il distingue mieux cette classe.

Pour une profondeur égale à 10, l'arbre à 3 indicateurs a besoin de 182 règles de décisions contre seulement 134 pour celui à 13 indicateurs.

En conclusion, après avoir étudié et expliqué l'importance des indicateurs pour discriminer les classes, nous avons montré que :

- **Utiliser les 13 indicateurs les plus importants avec une profondeur d'arbre de 10 donne l'arbre de décision avec le meilleur résultat (kappa = 0.9328) ;**
- **Quel que soit le nombre d'indicateurs utilisés, la profondeur d'arbre augmente le score, bien que l'écart se réduise à mesure qu'elle augmente ;**
- **L'arbre utilisant seulement 3 indicateurs atteint un kappa de 0.931, des F-scores équivalents à celui à 13 indicateurs pour l'orpaillage et l'agriculture, et plus élevés pour l'urbanisation et la classe « autre » ;**
- **L'arbre à 13 indicateurs nécessite moins de règles de décisions pour arriver à son résultat final.**

Nous choisissons de continuer avec l'arbre à 3 indicateurs (« distance à l'orpaillage passé », « distance aux zones agricoles (hors rizières) » et « distance aux cultures ligneuses ») qui répond mieux au problème (obtenir un bon score en minimisant le nombre d'indicateurs).

Nous utiliserons une profondeur de 10 (et donc 182 règles de décision) car la complexité supérieure du modèle est nécessaire pour avoir des résultats satisfaisants. De plus, pour utiliser un arbre de décision, il est plus facile d'automatiser les règles de décision plutôt que la préparation d'indicateurs.

2.2. VALIDATION DU MODELE

Le modèle a été appliqué sur les données 2018 afin d'être validé. Tous les pixels d'un déboisement ont été classifiés, puis la classe majoritaire a été assignée à l'objet. La Figure 75 montre les classes prédites par l'arbre de décision (sur les zones où l'on connaît les classes de référence) sur toute la Guyane.

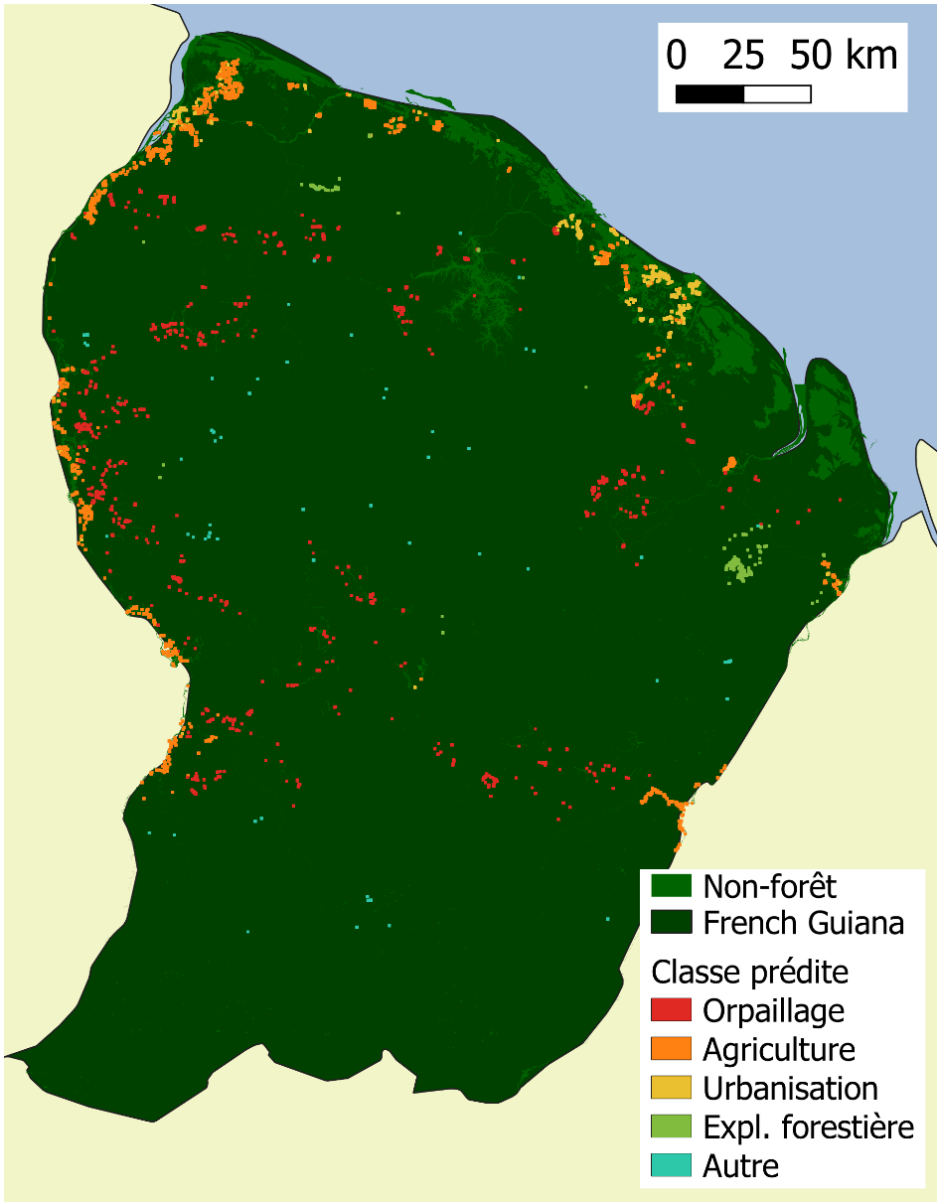


Figure 75 - Classes des zones déboisées prédites par le modèle à l'échelle de la Guyane sur l'année 2018 (pour les zones ayant des références)

2.2.1. RESULTATS GLOBAUX

Le modèle atteint un score kappa de 0.92 sur les données de 2018. Le Tableau 8 montre la précision, le rappel et le F-score atteints pour chaque classe.

Tableau 8 – Score de validation du modèle prédictif pour la caractérisation du déboisement

	Précision (%)	Rappel (%)	F-score (%)
Orpaillage	99	96	97
Agriculture	97	97	97
Urbanisation	85	88	87
Expl. Forestière	94	96	95
Autre	81	71	76

Sur l'orpaillage, l'agriculture et l'exploitation forestière, la précision et le rappel sont supérieurs à 94%. Les classes étant les plus confondues sont l'urbanisation et la classe « autre ». La première à une précision de 85% et un rappel de 88%. La deuxième a une précision de 81% et seulement 71% des entités de cette classe sont détectées. Il est à noter que la construction d'un arbre de décision spécifique à chaque type de déboisement répondant à un besoin défini (par exemple détection seule de l'orpaillage) pourrait donner de meilleurs résultats en utilisant des indicateurs plus adaptés à chaque cas.

2.2.2. ANALYSE DE LA CONFUSION DES CLASSES

La Figure 76 met en évidence les erreurs de classification de chaque classe.

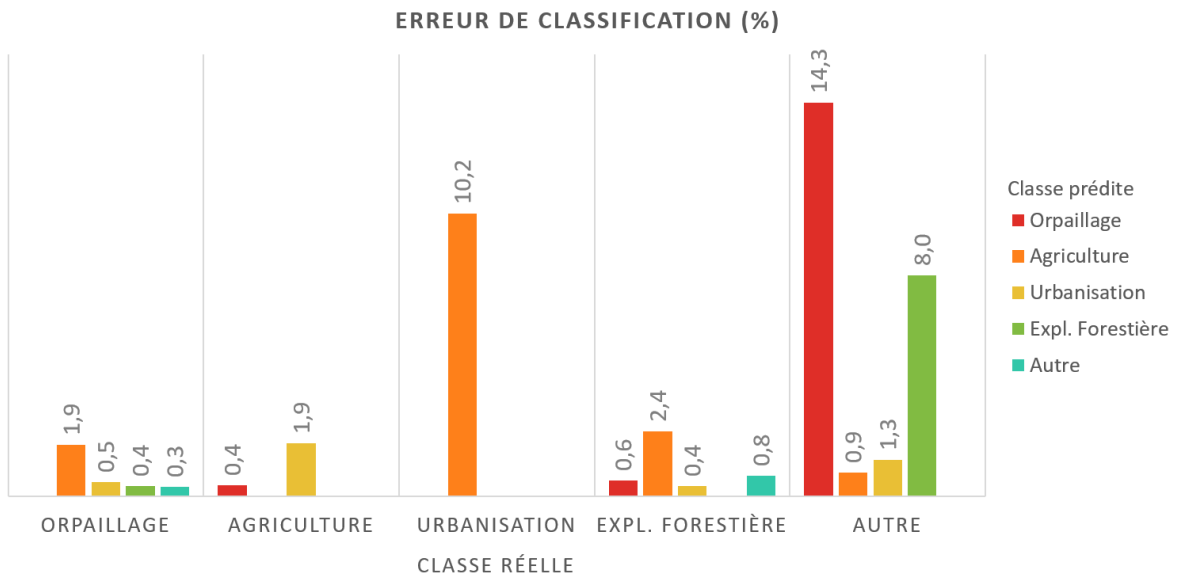


Figure 76 - Confusion en fonction de chaque classe (en pourcentage). Le titre de chaque sous-graphe donne la classe réelle. La classe prédite est donnée par les étiquettes de couleur.

Peu de déboisements dus à l'orpaillage ont une classe prédite différente de la leur : 0.5% sont classés en urbanisation, 0.4% en exploitation forestière et 0.3% en « autre ». L'agriculture est la classe qui est la plus prédite à la place de l'orpaillage (1.9%) : ce chiffre est principalement dû à une mauvaise classification de la nouvelle route entre Cacao et le site d'orpaillage de Devez. Celle-ci avait comme référence « orpaillage » mais a été classée « agriculture » (voir Figure 77-A). Les autres zones ainsi mal classées sont des petits sites d'orpaillage proche d'abattis dans le PAG.

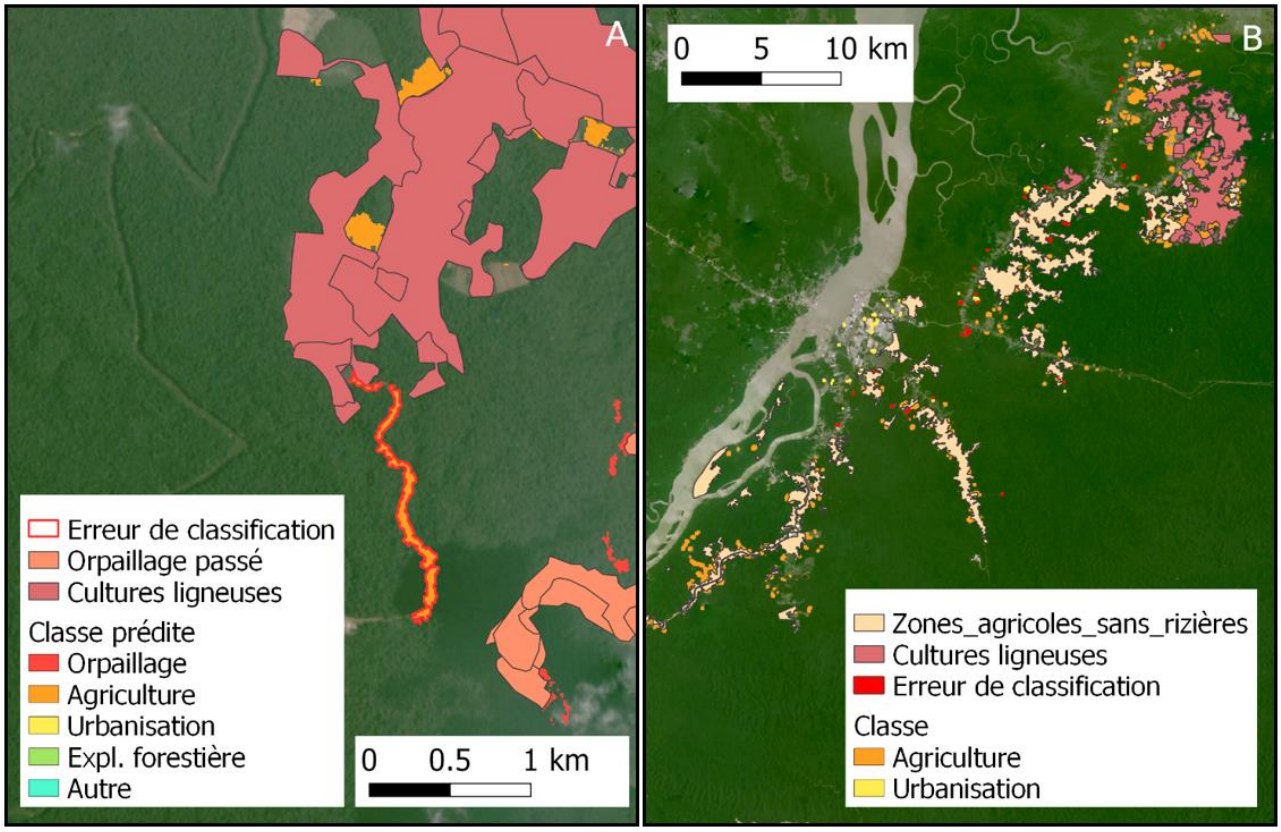


Figure 77 – Exemples d'erreur de classification de zones déboisées. A : route construite pour l'exploitation aurifère classée comme de l'agriculture (entre Cacao et la mine de Devez). B : confusions entre agriculture et urbanisation aux alentours de Saint Laurent du Maroni.

L'agriculture est principalement confondue avec l'urbanisation (1.9%) et cette dernière prend uniquement l'étiquette agriculture quand elle est mal classifiée (10.2%). Ceci est dû à la proximité entre les zones agricoles et les zones de développement humain. L'urbanisation mal classifiée se trouve majoritairement près de bâtiments agricoles ou de parcelles exploitées. La zone proche de Saint Laurent du Maroni est particulièrement sujette à ces confusions (voir Figure 77-B).

L'exploitation forestière est confondue le plus souvent avec des déboisements liés à l'agriculture (2.4%). Ces erreurs proviennent d'une zone au nord de Saint Georges, proche d'une zone agricole en comparaison des autres sites d'exploitation forestière. La classe « autre » a été beaucoup classée en tant qu'orpaillage (14.3%). Ceci s'est produit quand la zone déboisée été très proche de sites anciennement orpaillés. La classe « autre » a également été confondue avec

l'exploitation forestière (8%). Cela a eu lieu à plusieurs endroits du territoire, quand la zone était à une certaine distance des indicateurs (ni trop près, ni trop loin).

2.2.3. CARTE DE CONFIANCE

Le Tableau 9 montre l'indice de confiance pour les pixels de chaque classe (indice défini par la proportion d'observations classées identiquement dans le nœud terminal qu'il a atteint). La meilleure confiance appartient aux zones déboisées pour l'orpaillage (moyenne de 98.6% et écart-type de 11.6%). Vient ensuite celle des zones liées à l'agriculture et à l'exploitation forestière avec respectivement une moyenne de 97.6 et 96.4% et un écart-type de 12.2 et 7.6%. Connaissant une grande variation dans leur indice de confiance, les classes urbanisation et autre sont celles avec les indices de confiance les plus faibles (avec respectivement une moyenne de 94.5 et 95.6% et un écart-type de 17.4 et 20.5%).

Tableau 9 – Moyenne et écart-type des indices de confiance des pixels de chaque classe.

Classe	Moyenne	Ecart-type
Orpaillage	98.6	11.6
Agriculture	97.6	12.2
Urbanisation	94.5	17.4
Expl. Forestière	96.4	7.6
Autre	95.6	20.5

Ces résultats montrent qu'il est possible d'exploiter un modèle simple, fiable et utilisable en temps-réel pour déduire la cause des déboisements en Guyane, en particulier pour l'orpaillage, l'agriculture et l'exploitation forestière.

🔔 Ce qu'il faut retenir

Dans ce chapitre nous avons validé et analysé les résultats des méthodes décrites dans le chapitre 5 appliquées sur la Guyane. Pour résumer, nous avons d'abord évalué les cartes produites à partir des données Sentinel-1. Puis, nous avons montré et quantifié le fait que ces détections soient plus précises au niveau temporel mais aussi au niveau spatial que les détections de GLAD-FA basées sur les données optiques Landsat et souvent prises comme référence. La majorité de ces résultats ont donné lieu à une publication dans le journal *Remote Sensing of Environment* (Ballère et al., 2021a).

Concernant la méthode de détection du déboisement, nous avons vu qu'au niveau spatial :

- Le système basé sur les données radar, développé dans cette thèse, a une précision supérieure à 90% et un rappel supérieur à 75% sur les pertes forestières de tous types ;
- La détection spatiale des parcelles déboisées est particulièrement bonne sur les parcelles coupées en saison sèche et donnant place à un sol relativement lisse ;
- La détection spatiale est plus compliquée dans le cas d'objets longs et fins comme les routes, notamment quand ceux-ci sont perpendiculaires à la trajectoire des orbites ;
- La méthode GLAD-FA, basée sur des données optiques, se montre largement moins performante que la méthode basée sur les données Sentinel-1 : GLAD-FA omet plus de déboisement (rappel inférieur à 40% quel que soit le type de changement) ;
- Cette plus faible performance de GLAD-FA vient en grande partie de l'inaccessibilité des acquisitions optiques en raison de la persistance de la couverture nuageuse sur la Guyane (la contrainte imposant qu'une zone déboisée doit être « vues » deux fois par Landsat dans le délai fixé par la méthode, 180 jours, est trop forte pour le peu d'acquisitions exploitables) ;
- Il est nécessaire d'avoir un système de détection des pertes forestières à haute résolution spatio-temporelle pour pouvoir suivre le déboisement en Guyane.

Au niveau temporel :

- Des observations basées sur des données de terrain ont montré qu'en saison sèche, la méthode basée sur le radar et utilisée dans cette thèse peut détecter une coupe « immédiatement » même s'il reste de la végétation au sol ;
- En comparaison du système optique basé sur les données Landsat, le système radar basé sur les données Sentinel-1 permet de mieux estimer les dates de coupes des perturbations forestières et est plus adéquat pour lancer des alertes rapidement, notamment parce qu'il est le seul à pouvoir faire des détections dans les périodes où la couverture nuageuse est importante. Sur le cas de l'orpaillage, plus d'un tiers des détections basées sur l'otique montrent un retard de plus de 3 mois.

Concernant la méthode d'identification des causes des déboisements, après avoir étudié et expliqué l'importance des indicateurs pour discriminer les classes, nous avons montré que :

- Il est possible d'utiliser des indicateurs provenant de données exogènes pour définir la cause des déboisements sur un territoire comme la Guyane de manière automatique, en temps-réel avec un modèle simple nécessitant peu de préparation de données pour l'utiliser ;
- L'orpaillage, l'agriculture et l'exploitation forestière obtiennent un F-score supérieur à 95% et sont donc particulièrement bien reconnus.

Il est important de souligner que la création d'un modèle adapté à la discrimination d'un seul type de déboisement parmi les autres pourrait donner de meilleurs résultats, avec un nombre d'indicateur aussi faible. Ainsi, en spécialisant le modèle pour un besoin utilisateur, on pourrait améliorer encore la qualité de la classification.

Le chapitre 7 présentera les besoins des acteurs de Guyane recensés dans le cadre de cette thèse, ainsi que des applications concrètes de la méthode de détection des déboisements et des perspectives d'amélioration des méthodes et d'usage des cartes produites.

✋ Ma contribution

J'ai traité toutes les images sur la Guyane en utilisant la chaine de traitement développée dans cette thèse, sur l'ensemble du territoire et sur plusieurs années. Chaque début d'année de la thèse, les détections de l'année précédente ont été mises en ligne sur un site web (www.cesbiomass.net) afin d'être accessible et téléchargeable par les personnes qui le souhaitaient. Malheureusement, celui-ci n'est plus fonctionnel depuis l'incendie des serveurs OVH qui l'hébergeaient.

J'ai réalisé l'ensemble des validations et des analyses montrées dans ce chapitre.

Les résultats de la méthode de détection en Guyane sur les années 2016, 2017 et 2018 ont fait l'objet d'une publication scientifique (Ballère et al., 2021a), et j'ai présenté les résultats intermédiaires lors d'un oral en conférence (Ballère et al., 2019), et d'un poster aux journées de la recherche à l'IGN 2019 (Ballère et al., 2019). Les résultats liés à la caractérisation des déboisements ont été présentés à l'oral au vEGU (Ballère et al., 2021c), ainsi qu'aux journées de la recherche de l'IGN 2021 (Ballère et al., 2021b). La thèse a été présentée dans sa globalité sous forme de vidéo lors d'une journée thématique sur la télédétection (Ballère, 2021).

Chapitre 7 - Valorisations et perspectives

Contenu

1. Besoins utilisateurs 170
2. Exemple d'applications en Guyane 173
2.1. Détection des dommages causés par les barges d'orpaillage aux rives du Maroni 173
2.2. Détections de sites d'orpaillage clandestins 176
3. Utilisation des images Sentinel-1 à l'échelle globale 177
3.1. Géneralisation de la méthode de détection du déboisement 177
3.2. Opérationnalisation 180
3.2.1. Confiance dans la détection 180
3.2.2. Délai d'alerte et retards de détection 180
3.2.3. Délais de production 181
4. Perspectives 181
4.1. Perspectives d'amélioration 181
4.1.1. Combinaison de données multi-capteurs 181
4.1.2. Evolution dynamique du système de caractérisation 182
4.2. Perspectives d'utilisation 183

Dans le cadre de cette thèse, de nombreux contacts avec les acteurs forestiers guyanais ont été initiés. La première partie de ce chapitre expose les besoins utilisateurs recensés en Guyane. Dans la deuxième partie, sont exposés des exemples d'applications concrètes dans lesquelles les cartes développées dans cette thèse peuvent avoir un intérêt. Ensuite, nous discutons de l'utilisation des méthodes développées dans cette thèse à l'échelle de l'ensemble des forêts tropicales. Enfin, les perspectives d'amélioration des méthodes et de leurs utilisations sont exposées.

1. Besoins utilisateurs

Les cartes de déboisement sont utiles à plusieurs acteurs des forêts. Plusieurs points sont à prendre en compte pour définir leurs besoins :

- La périodicité : la périodicité est le pas temporel auquel l'utilisateur a besoin du produit. Plus elle sera longue, plus il existera des données à utiliser sur la période. Cela permet soit d'utiliser les données les plus adaptées à son problème, soit d'augmenter une série temporelle et ses statistiques (afin d'améliorer le filtrage par exemple) pour obtenir un résultat de meilleure qualité ;
- La rapidité de mise à disposition : certains utilisateurs n'auront pas besoin d'une mise à disposition très rapide, comme les scientifiques étudient l'historique des déboisements. Par exemple, les données PRODES (Valeriano et al., 2004) de l'INPE (utilisées dans de nombreuses études sur les changements forestiers) sont accessibles seulement plusieurs mois après la fin de la période analysée (d'août à août et disponibles en janvier/février). La rapidité de mise à disposition dépend de deux facteurs. Le premier est la disponibilité de la ressource : la donnée Sentinel-1 est accessible quelques heures après son acquisition (avec une co-registration rapide dans un premier temps, et améliorée au bout de 24 heures après réception de sa trace d'orbite précise). Le deuxième est le délai de confirmation d'une alerte. Plus il y a de données disponibles pour confirmer une alerte, plus la confiance dans ces alertes augmentera ;
- La surface minimale de détection : la plupart du temps, plus un objet est petit, plus il est susceptible d'être une fausse alarme, rendant la détection moins fiable ;
- Le compromis entre précision et rappel : certains utilisateurs auront besoin de détections très fiables quitte à en avoir peu, alors que d'autres préfèreront pouvoir contrôler chaque déboisement possible, quitte à ce qu'il soit faux.

La section suivante présente les besoins utilisateurs recensés en termes de cartes de déboisement.

Les cartes de déboisement en quasi-temps réel en Guyane intéressent principalement les personnes en charge du suivi des activités illégales (ONF – Office National des forêts, EMOPI – Etat-Major de l'Orpaillage et de la pêche illicite), et les ONG (WWF – World Wide Fund for Nature) dans le but d'interpeller les citoyens sur des sujets d'actualité.

L'ONF est en charge du suivi des défriches illégales liée à l'urbanisation en Guyane. Dans ce contexte, une détection hebdomadaire des nouveaux déboisements avec un seuil de détection minimum entre quelques centaines de m² (optimal) et 0.5 ha serait idéal.

La Guyane a son propre système d'information géographique (l'OAM – Observatoire de l'Activité Minière), dédié au suivi de l'activité minière, et partagé entre tous les acteurs de la LCOI

(Lutte Contre l'Orpaillage Illégal) (Linarès and André, 2020). Il répond à 2 enjeux principaux : renseigner et orienter le volet répressif de la LCOI ; évaluer la pression environnementale de l'activité minière (légale et illégale). L'ONF participe à l'OAM, sur les deux enjeux, principalement par l'usage de la télédétection (repérage et délimitation des déboisements, repérage de la turbidité) et par les missions héliportées. Dans le cadre de la LCOI, l'ONF relève les surfaces déboisées supérieures à 0.1 ha avec des données optiques. Les surfaces télé-détectées sont ensuite survolées en hélicoptère pour être confirmées et les informations sont ensuite remontées dans l'OAM (Linarès and André, 2020). Cette opération étant très couteuse, l'ONF privilégiera la confiance dans les détections plutôt que leur nombre. Pour qu'un système de détection rapide puisse compléter ce suivi de manière efficace, la détection doit être fiable et arriver le plus tôt possible. Le système sera toujours utile à 15 jours mais commencera à ne plus l'être à partir de 30 jours. Dans le cadre du suivi des surfaces exploitées, l'ONF produit chaque année le bilan environnemental qui rapporte les surfaces déboisées et impacts aux cours d'eau. Pour ce volet, les cartes doivent être précises spatialement mais une information annuelle est suffisante.

Les cartes de déboisement annuelles intéressent le ministère de l'agriculture pour le recensement des terres en Guyane et pour surveiller l'évolution des abattis, et le WWF pour dresser des rapports sur l'évolution du déboisement à destination des citoyens.

Au niveau du Service de la Statistique et de la Prospective (SSP) du Ministère de l'Agriculture et de l'Alimentation, la surface minimale qu'il serait nécessaire de prendre en compte pour les créations d'abattis ou défrichements supplémentaires adjacents à un abattis existant se situe entre 0.3 et 0.5 ha. C'est souvent à partir de ce seuil que se distingue une exploitation agricole (quelques cultures de rente contribuant à l'alimentation des marchés locaux) d'un jardin familial uniquement destiné à l'alimentation du foyer. Dans le cadre d'une utilisation annuelle visant à établir la surface agricole utilisée au cours de l'année, un compromis entre erreurs de commission et erreurs d'omission semble idéal : les unes pouvant compenser les autres. Ainsi, la surface totale détectée (valeur qui les intéresse) sera vraisemblablement très proche de la réalité, permettant ainsi des calculs théoriques de production sur la base de rendements moyens. Cependant, si l'on ne peut être certain de la compensation d'une erreur par l'autre, ou dans le cadre d'études plus fines sur un secteur donné avec une volonté d'exhaustivité de la détection, des erreurs de commission ont vraisemblablement moins d'impact car elles permettent d'identifier l'ensemble des sites de production, et donc potentiellement le nombre de ménages impliqués dans l'activité agricole.

Les cartes produites à partir des données Sentinel-1 comme présentées dans cette thèse répondent à la plupart des besoins exposés, notamment avec la mise en place d'un indice de confiance sur les détections pour que chaque utilisateur puisse avoir un produit qui lui convienne. Dans le cas de la lutte contre l'orpaillage illicite, les cartes produites paraissent intéressantes mais insuffisantes pour avoir une vue exhaustive de l'activité, notamment parce qu'en utilisant les produits Sentinel-1 GRD, la détection de zones inférieures à 0.2 ha n'est pas certaine. Cependant, celles-ci sont intéressantes à intégrer, notamment pour compléter les systèmes optiques existants (qui eux peuvent détecter des déboisements en dessous de 0.2 ha) pendant la saison humide.

2. Exemple d'applications en Guyane

Les cartes de déboisement produites sur la Guyane en temps-réel ont de nombreuses applications. Cette section présente des exemples d'application dans lesquelles ces cartes ont un intérêt.

2.1. DETECTION DES DOMMAGES CAUSES PAR LES BARGES D'ORPAILLAGE AUX RIVES DU MARONI

Le Suriname, comme l'ensemble du plateau des Guyanes, est touché depuis plus de trente ans par le fléau de l'orpaillage qui entraîne la destruction du couvert forestier et la pollution des cours d'eau en impactant fortement l'environnement et les populations locales. L'orpaillage y est dit primaire quand l'or est directement extrait de la roche-mère, alors qu'il est alluvionnaire quand il est exploité dans le lit des cours d'eau.

L'orpaillage alluvionnaire est le plus répandu au Suriname, où il prend une forme particulière dans les fleuves et rivières de taille importante. Des barges, véritables mini-usines flottantes, draguent les sédiments du fond des cours d'eau qui passent ensuite sur une série de rampes de lavage pour en extraire un concentré aurifère. L'or est finalement extrait par amalgame avec le mercure, hautement toxique (environ 1.5 kg par kg d'or). Outre cette pollution, le procédé est particulièrement dévastateur pour les cours d'eau dans lesquels de grandes quantités d'eau boueuse qui nuisent à la flore et la faune aquatique sont déversées. Les populations locales sont également directement touchées par cette pollution, le fleuve constituant leur bassin de vie : elles s'y nourrissent et s'y lavent.

Figure 78 - Photo des barges sur le Maroni (gauche), et les dégâts qu'elles entrainent sur les berges (droite). Source : Clément Villien, WWF Guyane.

Ces barges (Figure 78), interdites en Guyane mais pas au Suriname, font régulièrement polémique sur le Maroni, fleuve transfrontalier entre les deux territoires. Depuis début 2019, on assiste à un nouveau développement de cette pratique qui s'attaque désormais aux rives surinamaises. Celles-ci, normalement recouvertes de forêt, sont défrichées et la couche supérieure du sol est arasée à la pelle mécanique. Les barges entrent ensuite en action depuis le fleuve en creusant la terre des berges à la recherche d'or. En plus des boues rejetées dans le fleuve en grandes quantités, cette pratique entraine le déboisement et l'affaissement des rives et les rend ainsi vulnérables à l'érosion dans une région où la pluviométrie est très importante. C'est le lit même du cours d'eau qui s'en trouve fortement modifié par endroits.

Figure 79 - Images Sentinel-2 de berges intactes le 23/08/2018 (à gauche) et défrichées par les barges le 01/12/2019 (à droite) sur le Maroni (3°71 N, 54° O). Les images proviennent de EO Browser (https://apps.sentinel-hub.com/eo-browser/).

Cette pratique peut être suivie en quasi temps-réel et tout au long de l'année grâce aux cartes de déboisement produites dans cette thèse, contrairement aux systèmes de surveillance optique limités par la couverture nuageuse importante.

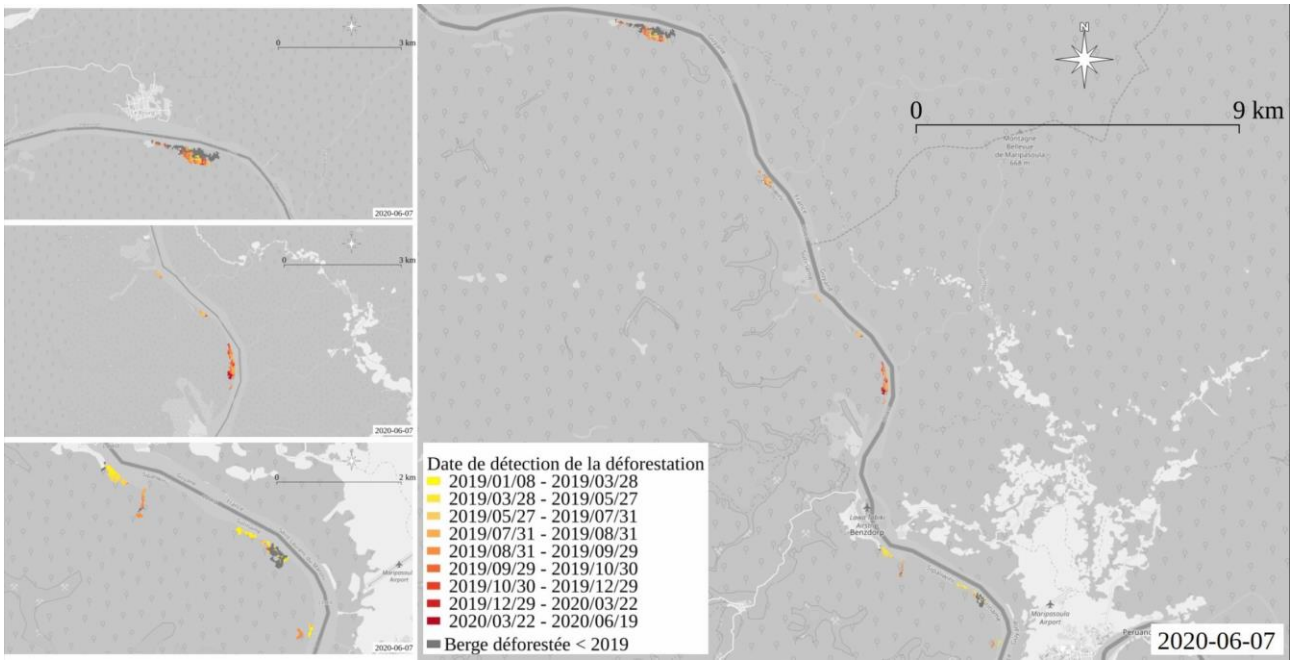


Figure 80 - Détection des berges défrichées à l'aide de Sentinel-1. La couleur correspond à la date du déboisement.

Comme on peut le voir sur la Figure 80, plusieurs chantiers de ce type se sont développés en 2019 entre Maripasoula et Papaïchton (environ 25 km de fleuve). Les barges attaquent les rives et creusent de véritables chenaux sur plusieurs dizaines de mètres. Sur cette seule zone, on estime une perte de plus de 36.7 ha de forêt seulement en 2019.

Ces pratiques, régulièrement dénoncées par les autorités françaises et les organisations locales, devraient disparaître à terme. En août 2020, le nouveau gouvernement surinamais s'est en effet engagé, dans un accord de coopération avec la France, à ne pas délivrer de nouvelles licences d'exploitations pour ces usines flottantes et ne plus renouveler celles arrivant à expiration[6]. Le respect de cet accord constituerait une avancée pour le fleuve Maroni et ses habitants : le suivi précis que l'on peut faire grâce à Sentinel-1 permettra de vérifier si ces annonces sont bien suivies d'actions.

[6] Saint-Laurent : un accord de coopération conclu entre la France et le Suriname - https://la1ere.francetvinfo.fr/guyane/ouest-guyanais/des-accords-de-cooperation-sanitaire-868066.html

2.2. DETECTIONS DE SITES D'ORPAILLAGE CLANDESTINS

En Guyane, les exploitants de mines aurifères sont tenus de respecter des règlementations, comme la nécessité de détenir les permis d'exploitation, la non-utilisation du mercure et l'obligation de restaurer la zone exploitée. En outre, les mines déviant des cours d'eau naturels doivent s'assurer que la turbidité des cours d'eau en aval n'augmente pas quand les sédiments sont relâchés. Le 17 août 2019, des agents de la compagnie des guides de Guyane ont effectué des prélèvements en aval de la crique Nelson (4°99 N, 52°91 O), révélant une forte augmentation de la turbidité en aval de la crique et accusant l'exploitant de la mine[7]. Quelques mois plus tard, ce dernier est revenu faire des prélèvements sur les lieux. Proche de son exploitation, la turbidité s'est révélée être inférieure aux seuils critiques fixés par la loi. En revanche, à 3 km en aval de la crique (où elle rejoint la crique Couy), le cours d'eau était pollué et la turbidité y était particulièrement élevée. A proximité se trouvaient des signes de la présence de sites d'orpaillage clandestins sur lesquels l'opérateur de la crique Nelson a rejeté la faute.

Dans ce type de situation, la temporalité est importante : les orpailleurs clandestins étaient-ils présents lors du prélèvement initial, ou sont-ils seulement arrivés lorsque l'exploitant a effectué ses propres analyses, plusieurs mois plus tard ?

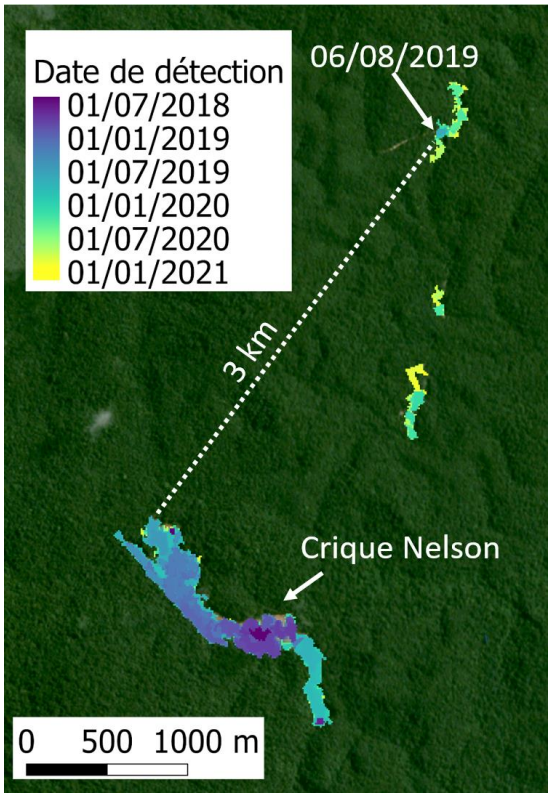


Figure 81 - Détection des pertes forestières via Sentinel-1 sur l'exploitation de la crique Nelson et en aval. Image centrée en 5°N 52°9O

[7] Crique Nelson : l'opérateur minier se défend d'être à l'origine de la pollution - https://la1ere.francetvinfo.fr/guyane/crique-nelson-operateur-minier-se-defend-etre-origine-pollution-760399.html

Les détections faites à partir de Sentinel-1 (Figure 81) montrent qu'une partie de la forêt a été déboisée à 3 km en aval de l'exploitation le 6 août 2019, soit 11 jours avant le prélèvement montrant une augmentation de la turbidité dans l'eau. Ceci n'établit pas la culpabilité de l'un ou l'autre des acteurs, mais la fine échelle temporelle des cartes produites à partir des données Sentinel-1 en Guyane permet de définir de manière factuelle la présence des orpailleurs clandestins au moment des prélèvements effectués par la compagnie des guides.

3. Utilisation des images Sentinel-1 à l'échelle globale

La méthode de détection des déboisements utilisée dans cette thèse est bien adaptée pour l'estimation des pertes forestières en Guyane. Nous discuterons ici de la généralisation de cette méthode sur d'autres territoires.

3.1. GENERALISATION DE LA METHODE DE DETECTION DU DEBOISEMENT

Divers facteurs affectent le signal RSO et donc la détection des changements forestiers, notamment les paramètres d'acquisition du système RSO, les facteurs environnementaux, le type de perturbation et les caractéristiques de la forêt.

En ce qui concerne les caractéristiques de la forêt, la méthode a été développée pour détecter des déboisements sur les forêts tropicales primaires (hauteur élevée des arbres, forte densité de la canopée et faible saisonnalité). En effet, la première condition pour distinguer une ombre en lisière de forêt sur une image RSO est d'avoir des arbres d'une hauteur suffisante (car la taille de l'ombre est fonction de la hauteur de l'arbre). Cette méthode n'est donc pas destinée à des forêts ayant une hauteur inférieure à 15 mètres pour un terrain plat (voir chapitre 5, 1.1.4). Dans les forêts tropicales humides denses et peu saisonnières, le signal radar varie peu dans le temps. En revanche, dans une forêt clairsemée, le sol et ses variations associées dues aux effets environnementaux (pluies, phénologie, variations de rugosité, etc.) contribueront au signal RSO. Par conséquent, appliquer la méthode sur une forêt clairsemée pourrait conduire à un taux de fausses alarmes élevé. De plus, la phénologie dans des forêts décidues peut introduire des changements dans le signal.

Le type de déboisement est un autre facteur qui peut affecter la rétrodiffusion des images Sentinel-1. La détection d'une coupe sélective n'est souvent pas aussi précise que celle d'une coupe à blanc et peut être caractérisée par une forte densité de petites détections dans la zone perturbée. Les lois de certains pays peuvent également influencer la détection. Par exemple, la législation brésilienne n'autorise plus la coupe de plusieurs grands arbres (comme le noyer du Brésil). Or, dans

une zone déboisée, les arbres restants peuvent entrainer une rétrodiffusion de volume, rendant la détection plus difficile car la chute de signal due au déboisement n'est pas aussi marquée que lors d'une coupe rase de tous les arbres.

Les résultats de détection dépendent aussi du nombre de directions de visée RSO disponibles. En Guyane, les orbites ascendantes et descendantes sont toutes deux disponibles. Dans les régions avec une seule direction de visée RSO (voir la Figure 19 du chapitre 3), moins d'ombres seront susceptibles d'être détectées.

La méthode Bouvet et al., 2018 a été testée avec succès sur plusieurs sites à travers la forêt tropicale. La méthode a légèrement évolué entre chaque étude, ainsi que les paramètres utilisés (seuils et images). Il est important de noter que la méthode de validation diffère suivant les sites en fonction des données disponibles, des moyens alloués et de la configuration des sites.

En **Amazonie**, la méthode a initialement été développée sur une zone de 600 000 ha autour de la ville de Yurimaguas au **Pérou,** entre octobre 2014 et septembre 2017 (Bouvet et al., 2018). Cette zone est principalement touchée par l'agriculture vivrière (riz, papaye et légumes), bien que l'agriculture s'y développe également de façon industrielle (palmiers à huile et cacao). La validation de la carte produite sur cette zone a été effectuée à partir d'une matrice de confusion rapportée en nombre de pixels, et basée sur des polygones de référence sélectionnés automatiquement. La précision atteinte est de 99.4% et 96.6% respectivement pour les classes « déboisée » et « forêt intacte », et le rappel de 80.3% et 99.9%. Cela signifie que sur ce site, 4/5ème des déboisements effectifs ont été détectés, mais qu'il y a eu très peu de fausses détections.

Sur la **Guyane** (territoire de 8 380 000 ha), différents types de déboisements ont été évalués sur les années 2016, 2017 et 2018 (Ballère et al., 2021a). La validation était basée sur une matrice de confusion rapportée en nombre d'échantillons, eux-mêmes déterminés avec une méthode d'échantillonnage aléatoire stratifié. La précision et le rappel atteints sont respectivement de 97.7% et 89.4% pour les petites parcelles agricoles, 91.1% et 75.9% pour l'exploitation forestière (routes et places de dépôt), et 100% et 80% pour l'orpaillage (grandes exploitations légales et petites exploitations illégales).

Le **Suriname** connait les mêmes types de déboisement que ceux de la Guyane. La carte produite sur ce territoire (travaux non-publiés) a été validée sur l'année 2017, à partir d'une matrice de confusion rapportée en nombre de pixels, et basée sur des polygones de référence sélectionnés d'après des cartes de déboisement existantes, disponibles sur www.gonini.org. La précision obtenue sur les 1704 ha de déboisements évalués (61 812 ha de forêt intacte) est de 99.9% et le rappel de 78.3%.

Au **Brésil**, la méthode a été testée sur un ensemble de 12 000 000 ha dans l'Etat du Para dans le cadre d'un projet visant à évaluer un programme de lutte contre le déboisement mis en place auprès d'exploitants agricoles de la région (travaux non-publiés). La validation a été réalisée à partir d'une matrice de confusion rapportée en nombre de pixel et des polygones de références réalisés par photo-interprétation, sur les détections de 2018 et 2019. La détection des déboisements dans cette zone a atteint une précision de 99.9%, et un rappel de 66.1%. La détection des coupes sélectives présentes dans les polygones de référence était moins bonne : dans la majorité des cas, l'ombre est détectée, mais les arbres restants debout induisent une baisse trop faible ou inexistante du signal radar pour la détection des surfaces déboisées.

Dans le **Bassin du Congo**, une étude sur un site de 648 000 ha situé sur les provinces de Ngouine, Ogooue-Maritime et Moyen-Ogooue, a permis d'évaluer la méthode en fonction de la taille des parcelles déboisées entre 2015 et 2017 (Hirschmugl et al., 2020). Les parcelles déboisées dans cette région concernent les exploitations minières au nord, des grandes plantations de palmiers à huile dans le sud, des activités agricoles le long des routes et rivières, et de l'exploitation forestière, caractérisée par de très petites perturbations et par la construction de routes. Sur les parcelles inférieures à 25 ha, la précision atteinte était de 80% et le rappel de 61%. Pour les parcelles supérieures à 25 ha, la précision était de 100%, et le rappel de 37.6%. La sélection des échantillons de validation a été faite avec un protocole d'échantillonnage aléatoire stratifié, et la matrice a été rapportée en termes de proportion de surfaces estimées. La faible proportion de détection des parcelles supérieures à 25 ha provient du fait que la méthode appliquée sur cette zone utilise une enveloppe convexe sur les ombres détectées pour reconstruire la surface déboisée : dans le cas de grandes parcelles, les ombres détectées à leur lisière sont éloignées les unes des autres, rendant leur association plus compliquée.

En **Asie du Sud-Est**, le projet SOFT financé par l'ESA (https://www.globeo.net/soft) a mis en évidence la capacité de la méthode à détecter avec fiabilité les déboisements au **Vietnam**, au **Laos** et au **Cambodge**, avec une étude à l'échelle des 3 pays (environ 75 Mha), sur 3 années (2018 à 2020) (Mermoz et al., 2021). Les forêts de ces pays sont déboisées au profit de plantations (hévéa, acacia), d'infrastructures routières et d'exploitation agricole. La validation était basée sur une matrice rapportée en termes de proportion de surfaces estimées, et dont les échantillons ont été sélectionnés par échantillonnage aléatoire stratifié. La précision des déboisements est de 95% et le rappel de 89.8%. Les coupes sont principalement des coupes rases : celles-ci sont caractérisées par une baisse de signal importante sur toute la surface déboisée, ce qui explique que l'algorithme

obtienne de tels résultats. Cette étude a notamment montré la capacité du système à opérer dans des zones avec une forte topographie comme le Nord-Vietnam et le Nord-Laos.

Ces études réalisées à travers les 3 grands bassins de forêt tropicales (Amazonie, Bassin du Congo et Asie du Sud-Est) aux usages variés montrent la capacité de ce système à être généralisé et à opérer comme un système d'alerte en cas de déboisement, quelle qu'en soit la cause, notamment sur les zones de forte topographie et sur les zones de petites tailles (sur lesquelles il y a un besoin pointé par la littérature). Il est donc possible d'envisager un suivi des pertes de forêts tropicales avec ce système.

3.2. OPERATIONNALISATION

Nous discutons ici du potentiel du système d'alerte pour une utilisation opérationnelle.

3.2.1. CONFIANCE DANS LA DETECTION

Dans la majorité des cas, un système d'alerte opérationnel doit avoir un faible taux de fausses alarmes. Un taux de fausses alarmes important serait acceptable pour surveiller un petit territoire et si aller sur le terrain vérifier ne coûtait pas cher, mais cela est rarement le cas en forêt tropicale. Dans toutes les études utilisant cette méthode, la précision des détections est supérieure à 95% pour un MMU de 0.2 ha (excepté pour l'étude sur le Gabon où le MMU était fixé à 0.04 ha et la précision était de 80%), ce qui implique une grande confiance dans les résultats. La précision augmente avec le MMU, tout comme la confiance. Par ailleurs, la mise en place d'un indice de confiance pourra permettre à un opérateur de fixer le taux de détection et de fausse alarme en fonction de ses besoins.

3.2.2. DELAI D'ALERTE ET RETARDS DE DETECTION

Pour les activités illégales ou surveillées, le délai entre la perturbation et la détection doit être le plus court possible. L'utilisation d'images radar, indépendantes des conditions météorologiques présente l'avantage de garantir la disponibilité d'acquisitions exploitables. De plus, la méthode de détection est basée sur un indicateur fiable s'affranchissant des effets de l'environnement affectant les données radar. Trois images sont nécessaires pour déclencher une alerte avec confiance. Les acquisitions Sentinel-1 permettent donc de confirmer un déboisement 12 jours après sa première détection pour les orbites à 6 jours, ou 24 jours après pour les orbites à 12 jours. Si le déboisement est détecté dès la première acquisition Sentinel-1 après son apparition, le délai de détection global est compris entre 12 jours (déboisement survenant le jour d'une acquisition

Sentinel-1) et 18 jours (déboisement survenant le lendemain d'une acquisition Sentinel-1) ; et entre 24 et 36 jours pour les orbites à 12 jours. Cette rapidité d'alerte est particulièrement utile pour surveiller les conversions de forêt illégales (urbanisation illégale, exploitation aurifère illégale, etc.)

Dans des cas moins urgents, des estimations précises de la date de déboisement sont également nécessaires, comme par exemple dans les études liées à la temporalité des déboisements. Aujourd'hui, seuls les systèmes basés sur les images RSO sont capables de fournir une information temporelle aussi fine à large échelle.

3.2.3. DELAIS DE PRODUCTION

Une image Sentinel-1 est disponible 24 heures maximum après son acquisition sur la plateforme française d'exploitation des produits Sentinel (PEPS) et nécessite moins de 30 minutes pour être téléchargée, filtrée et traitée afin d'obtenir une carte de déboisement à l'aide des algorithmes développés. Il est donc possible de fournir des mises à jour du produit de déboisement le lendemain de chaque acquisition.

De par sa précision élevée, son adaptabilité aux besoins de chaque utilisateur et ses délais de production raisonnables, la méthode de détection du déboisement étudiée dans cette thèse a le potentiel pour être un système d'alerte opérationnel efficace et complémentaire de ceux existants.

4. Perspectives

Cette partie décrit les perspectives d'amélioration des méthodes, et leurs perspectives d'usage.

4.1. PERSPECTIVES D'AMELIORATION

4.1.1. COMBINAISON DE DONNEES MULTI-CAPTEURS

L'analyse des résultats de la méthode du déboisement (chapitre 5) a montré que le système d'alerte basé sur Sentinel-1 était plus performant que le système basé sur les données optiques Landsat, tant sur le plan spatial (respectivement rappel supérieur à 75% et inférieur à 40% et précision supérieure à 90% pour les deux systèmes) que temporel (un tiers des parcelles sont détectées par les données optiques avec un retard supérieur à 3 mois pour des déboisements annuels). Cependant, certaines parcelles déboisées se sont révélées être détectées seulement par imagerie optique (presque 5% des polygones de référence), que ce soit à cause de la topographie ou des effets environnementaux qui empêchaient la baisse du signal radar. De la même façon, les

données optiques ont permis la détection de certaines parcelles avant les données radar (20% des parcelles pour des déboisements annuels), bien que le retard reste inférieur à 4 mois, que ce soit en raison des dates d'acquisition des images, ou parce que ce signal radar n'a pas suffisamment chuté au moment de la coupe. Le système optique est ainsi complémentaire du système radar afin d'obtenir une vue plus exhaustive des détections et une rapidité de détection optimale. Il est également à noter que les images optiques Sentinel-2 ont une répétitivité de 5 à 10 jours sur tout le globe et une résolution spatiale de 10 mètres sur les bandes RGB et PIR. Elles peuvent ainsi compléter les détections RSO de la même façon que les images Landsat.

Le produit de pertes du couvert forestier pourrait également être complété par la détection des repousses forestières, ce qui permettrait d'avoir une vue exhaustive de la dynamique du couvert forestier. Les données en bande-L ALOS-PALSAR pourraient répondre à ce besoin. Elles ne sont cependant disponibles gratuitement qu'annuellement. Plusieurs systèmes en bande-L devraient voir le jour dans les années à venir. En 2023, NISAR (projet NASA/ISRO) devrait fournir des images tous les 6 jours à une résolution de 3 à 10 mètres. Tandem-L (projet du DLR) et ROSE-L (projet de l'ESA) devraient également être des SAR imageurs en bande-L, respectivement en orbite en 2024 et 2028.

4.1.2. EVOLUTION DYNAMIQUE DU SYSTEME DE CARACTERISATION

Le système de caractérisation du déboisement étudié dans cette thèse a pour but de répondre à une contrainte de temps-réel en contexte tropical caractérisé par une forte couverture nuageuse. C'est pourquoi les images optiques n'ont pas été utilisées. Cependant, celles-ci peuvent être très efficaces pour relever la cause de déboisement. Il serait intéressant d'étudier l'apport de l'arrivée d'une image optique exploitable pour la caractérisation automatique. Il faudrait dans ce cas développer une chaine de traitement permettant la sélection des images optiques exploitables. Ceci s'avère délicat : si les nuages épais peuvent être facilement détectés, la détection de leur ombre est plus délicate tout comme la détection des voiles nuageux ou brumeux.

De la même manière, les séries temporelles d'images RSO évoluent différemment suivant l'usage du sol. Ainsi, une série temporelle conséquente post-perturbation, pourrait apporter une information sur l'usage du sol (remplaçant l'imagerie optique si celle-ci tarde à arriver). L'information multi-échelle des déboisements parait également appropriée au regard des structures détectables dans les cartes : l'exploitation forestière révèle de grandes routes autour desquelles la densité de petites détections est élevée, les abattis sont souvent proches les uns des autres, etc. Il serait également intéressant de tester plus d'indicateurs morphologiques, comme les indices de

Miller ($\frac{4\pi \times aire}{périmère^2}$), de Morton ($\frac{aire}{\pi(0.5 \times étalement)^2}$) et de Gravélius ($\frac{périmètre}{2\sqrt{\pi \times aire}}$) qui pourraient peut-être mieux discriminer les structures associées à chaque déboisement. Ces informations supplémentaires pourraient permettre d'obtenir une confiance plus élevée dans le modèle prédictif, et une meilleure classification. La méthode a montré que les indicateurs géologiques faisaient sens au regard des activités forestières en Guyane. Le lieu des déboisements est souvent lié à des commodités (voies de communication), à une géologie particulière ou à des conditions climatiques spécifiques. Sans appliquer directement le modèle prédictif produit sur la Guyane, il serait intéressant d'approfondir l'utilisation de ces indicateurs sur d'autres territoires.

4.2. PERSPECTIVES D'UTILISATION

L'opérationnalisation de la méthode de détection du déboisement proposée pourrait intéresser différents acteurs tels que les acteurs nationaux et les organisations à but non lucratif. Les premiers pour la gestion des aires protégées, la gestion des concessions agricoles et le recensement des terres, et les seconds pour la sensibilisation du grand public.

Par exemple, ce système pourrait compléter la lutte contre les activités illégales (orpaillage illégal et défrichement illégal) en Guyane car il est exploitable toute l'année, automatisé et coûte moins cher qu'un suivi par photo-interprétation. La carte de confiance associée aux résultats permettra de sélectionner seulement les détections les plus fiables. Cependant, pour être utile dans ce cas, la détection de la perte de forêt doit être caractérisée comme de l'orpaillage. Les acteurs ayant une connaissance et une utilisation plus courante des systèmes optiques, les cartes obtenues à partir des images Sentinel-1 pourraient donc être introduites progressivement, d'abord pour confirmer les détections optiques, avant leur utilisation propre. L'identification en temps-réel des zones où la biodiversité est menacée à cause des perturbations anthropiques devrait être un outil utile pour les programmes de gestion. Par exemple, cela pourrait servir à la gestion des trames verte et bleue, avec notamment le suivi de l'habitat et/ou du déboisement illégal : une image actualisée permettrait un suivi des corridors, dans le respect (aussi) des préconisations données aux orpailleurs. La connaissance fine des zones et de l'intensité des déboisements dans les aires protégées (problème endémique en Guyane) peut contribuer à mieux encadrer et développer ces aires protégées, et de les organiser en « systèmes de refuge » (Simmons et al., 2018).

En outre, en plus d'être un produit d'alerte, ce système pourra également être utilisé pour évaluer la quantité de changement pour un rapport ou une analyse approfondie, car il montre des précisions élevées. Il peut donc être utilisé par les scientifiques, par exemple, pour la quantification

du déboisement, en particulier dans les régions toujours nuageuses. Autre exemple : le déboisement est une variable dans les modèles de niches écologiques (Chavy et al., others, 2019) permettant de comprendre et prédire les zones à risques pour certaines maladies, et dans l'évaluation de l'empreinte humaine sur la biodiversité (de Thoisy et al., 2010). Avoir une échelle spatiale et temporelle fine, permet de confronter des inventaires animaliers aux déboisements pour mettre en évidence leurs conséquences. L'étude (Guégan et al., 2020) a montré les mécanismes entre forêts et émergences des zoonoses, la fragmentation des forêts ayant un rôle très important à cause du contact rapproché entre populations forestières et non-forestières. Pour faire avancer ces études, il est nécessaire de porter une attention particulière aux évènements précis menant à des foyers de maladies à l'échelle locale et mondiale. Ceci serait rendu possible par la précision spatiale et temporelle de nos cartes, et leur généralisation possible à large échelle.

🔔 Ce qu'il faut retenir

Dans ce chapitre, nous avons vu que :

- Il apparaît que les cartes produites avec les données Sentinel-1 présentées dans cette thèse répondent à la plupart des besoins des utilisateurs en Guyane, ou peuvent compléter les systèmes existants ;
- Il est possible d'opérationnaliser la méthode présentée dans cette thèse pour répondre à un besoin de détection en quasi-temps réel ;
- La méthode de détection du déboisement a été appliquée avec succès sur plusieurs territoires des trois grands bassins de forêt tropicale (Amazonie, Asie de Sud-Est et Bassin du Congo) ;
- Les cartes de détection produites avec Sentinel-1 pourraient être compléter par des cartes produites à partir d'autres données compensant leurs faiblesses ;
- Il serait intéressant de tester l'utilisation d'indicateurs géographiques pour la caractérisation des déboisements sur des territoires similaires à la Guyane, car les déboisements sont souvent liés à des commodités (voies de communication), à une géologie particulière ou à des conditions climatiques spécifiques ;
- Les cartes produites en temps-réel peuvent compléter les systèmes de surveillance existants, mais peuvent également faire avancer les études nécessitant des informations spatiales et temporelles fines à large échelle.

✋ Ma contribution

J'ai recensé les besoins des utilisateurs en Guyane ainsi que dans les projets sur lesquels j'ai travaillé. J'ai travaillé sur plusieurs cas d'application avec le WWF Guyane, dont celui sur les berges du Maroni, qui a fait l'objet d'un article sur le blog du CESBIO (Ballère, 2020).

En plus de celle sur la Guyane, j’ai réalisé les études et validation de la méthode de détection des déboisements sur le Suriname et sur le Brésil, dans le cadre d’une étude avec des économistes de l’INRAE (Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement) visant à évaluer les programmes REDD+. J’ai participé aux études sur le Pérou et sur l’Asie du Sud-Est qui ont donné lieu à des publications (Bouvet et al., 2018) et (Mermoz et al., 2021).

Chapitre 8 - Conclusion

L'objectif général de cette thèse était de développer des méthodes pour le suivi continu de la forêt tropicale basées sur les données de télédétection satellitaire, et des données cartographiques exogènes. En raison de leur insensibilité à la couverture nuageuse, qui est particulièrement persistante en Guyane, l'accent a été mis sur l'exploitation des données radar à synthèse d'ouverture Sentinel-1 de Copernicus, disponibles systématiquement tous les 6 à 12 jours à l'échelle globale. Leur accès libre et gratuit les rend particulièrement adaptées pour être utilisées de manière effective dans des applications opérationnelles. De la même manière, les données cartographiques décrivant une occupation ou un usage du sol peu changeant peuvent être utilisées à tout moment dans des applications de suivi en continu. Dans cette étude, l'objectif spécifique était de fournir des outils générant des alertes en continu de déboisements labélisés par leur cause. Ces informations sont nécessaires à la lutte contre les défriches illicites, et à l'estimation et au suivi du déboisement sur les zones tropicales.

La première partie de cette thèse a introduit l'importance de la forêt tropicale, notamment pour la biodiversité et l'environnement et décrit les conséquences et les menaces liées au déboisement. Les causes et les pratiques de déboisement des forêts tropicales ont ensuite été exposées, ainsi que les systèmes existants permettant de les surveiller. Cela a permis d'identifier les besoins qui ne sont actuellement pas couverts par les systèmes d'observation actuels. De manière générale, au niveau spatial, le suivi des parcelles supérieures à 3 ha est bien maîtrisé, mais les pratiques de coupe évoluent et ces dernières tendent à être de plus en plus petites. Au niveau temporel, la détection des activités se déroulant en saison humide (majoritairement les activités illégales) est souvent tardive. En effet, la plupart des systèmes sont basés sur de l'imagerie optique sensible à la couverture nuageuse persistante dans les zones tropicales, bien que la rapidité soit un facteur clé dans la lutte contre le déboisement illicite. De plus, l'identification continue des causes de déboisements détectés est primordiale pour pouvoir agir en conséquence, mais encore peu abordé dans les études.

Le site d'étude utilisé dans cette thèse pour étudier les méthodes de suivi en continu du déboisement est l'ensemble de la Guyane. Ce territoire présente des pratiques, des formes, des temporalités et des causes de déboisement diversifiées, permettant l'analyse des méthodes sur des coupes différentes. De plus, c'est un territoire bien suivi par télédétection, et les données cartographiques existantes liées à l'occupation des sols sont nombreuses.

La première partie de ce travail de thèse a consisté à analyser les différentes données de déboisement, ainsi que le signal radar renvoyé par Sentinel-1 sur des parcelles déboisées. Cela a montré que certaines données géographiques permettaient de discriminer des causes de déboisement, et que le signal radar ne montrait pas un motif régulier lors d'un déboisement.

Au regard de ces observations, nous avons choisi de travailler sur une méthode de détection de pertes forestières utilisant les données Sentinel-1, et basée sur un indicateur géométrique (et donc fiable) de l'image radar pour établir la présence d'un déboisement. Cette méthode a été appliquée et validée à l'échelle de la Guyane sur deux années, après avoir défini ses paramètres applicables en Guyane et apporté des améliorations (prise en compte de la dimension temporelle, de la topographie et création d'un indice de confiance).

Au niveau spatial, les résultats de la méthode de détection du déboisement ont montré une précision supérieure à 90% pour tous les types de déboisements évalués (agriculture, exploitation forestière, orpaillage) et un rappel supérieur à 75%. La détection, particulièrement bonne pour les parcelles coupées pendant la saison sèche et donnant place à un sol relativement lisse, était plus compliquée dans le cas d'objets longs et fins comme les routes, notamment quand ceux-ci sont perpendiculaires à la trajectoire des orbites.

Ce produit a été comparé à un système optique basé sur les données Landsat et souvent pris comme référence dans les études de détection continue des déboisements. Le système optique a montré un rappel inférieur à 40% quel que soit le type de changement. Les omissions étaient en partie liées au capteur. Certaines alertes ne pouvaient pas être confirmées car les parcelles déboisées n'étaient pas « visibles » deux fois sur des images optiques Landsat dans le délai fixé par la méthode en raison de la couverture nuageuse (que ce soit parce le déboisement était éphémère, ou par manque d'image exploitable, même en saison sèche). Bien que ce résultat soit valable uniquement pour la méthode GLAD-FA basé sur les images Landsat, il souligne l'apport des données Sentinel-1 au niveau spatial pour des produits d'alertes de déboisement dans des territoires comme la Guyane.

Ce produit a été comparé à un système : GLAD-FA faisant référence dans la communauté. Ce produit basé sur les données optiques Landsat a montré de faibles performances avec un rappel inférieur à 40% pour le déboisement. Cette faible performance s'explique en grande partie par la présence persistante de la couverture nuageuse, rendant l'exploitation des données optiques très problématique sur la zone. Cette limitation a été observée même en saison sèche, pendant laquelle une couverture nuageuse moindre qu'en saison humide mais toujours présente gêne l'exploitation des données optiques. Ces résultats soulignent l'apport des données Sentinel-1 au niveau spatial pour des produits d'alertes de déboisement dans des territoires comme la Guyane.

Il a également été montré la nécessité d'avoir un système détectant les surfaces déboisées inférieures à 3 ha, et pouvant compléter les systèmes de détections actuels.

Au niveau temporel, la méthode n'a pas pu être validée en absolu à l'échelle de la Guyane, car il n'existe pas de données adaptées (avec une date de coupe assez précise). Cependant, des observations basées sur des données de terrain ont montré qu'en saison sèche, les données Sentinel-1 pouvait détecter une coupe « immédiatement » même s'il reste de la végétation au sol. En comparaison du système optique GLAD-FA, les données Sentinel-1 permettent de mieux estimer les dates de coupes forestières et de lancer des alertes rapidement, notamment parce qu'elles sont les seules à pouvoir faire des détections dans les périodes où la couverture nuageuse est importante. Sur le cas de l'orpaillage, plus d'un tiers des détections basées sur l'otique montrent un retard de plus de 3 mois. Par ailleurs, il a été établi que le lancement d'une alerte avec certitude nécessitait un délai de 12 jours avec la méthode basée sur les données Sentinel-1.

Bien qu'en général les données Sentinel-1 seules permettent une meilleure détection (spatiale et temporelle) des déboisements tropicaux que les données Landsat seules, localement, certaines surfaces déboisées sont détectées plus vite (ou seulement) par les données optiques. Ceci montre que la fusion de capteurs (ou de cartes basées sur des capteurs différents) pourrait améliorer les détections de déboisements.

Pour que l'information de coupe puisse être utilisée dans un système de lutte contre le déboisement, celle-ci doit être complétée par la caractérisation de sa cause le plus rapidement possible. Nous avons proposé un modèle permettant de caractériser les déboisements en se basant sur des indicateurs géographiques (distances à des usages ou occupation du sol particuliers) car ceux-ci évoluent peu en Guyane et peuvent donc être utilisés dans des applications temps-réel. Celui-ci a été appliqué et validé à l'échelle de la Guyane sur l'année 2018.

L'orpaillage, l'agriculture et l'exploitation forestière ont obtenu un F-score supérieur à 95% ce qui montre qu'ils sont particulièrement bien reconnus. Les principales confusions étaient l'assignation de l'étiquette « agriculture » à la classe « urbanisation » et de la classe « orpaillage » à la classe « autre » qui représente les perturbations forestières naturelles. Ces résultats montrent qu'il est possible d'exploiter un modèle simple, fiable et utilisable en temps-réel pour déduire la cause des déboisements en Guyane.

Il serait intéressant de tester l'utilisation d'indicateurs géographiques pour la caractérisation des déboisements sur des territoires similaires à la Guyane, car les déboisements sont souvent liés à des commodités (voies de communication), à une géologie particulière ou à des conditions climatiques spécifiques.

Pour finir, cette thèse a bénéficié de nombreux contacts avec les acteurs en Guyane. Il apparaît que les cartes produites avec les données Sentinel-1 répondent à la plupart de leurs besoins, ou peuvent compléter les systèmes existants. Par ailleurs, il a été défini qu'il était possible d'opérationnaliser ces méthodes pour répondre à un système de détection en quasi-temps réel. De plus, des travaux sur la méthode de détection du déboisement similaires à ceux de cette thèse ont été menés avec succès sur plusieurs territoires des trois grands bassins de forêt tropicale (Amazonie, Asie de Sud-Est et Bassin du Congo) démontrant l'efficacité de cette méthode à l'échelle des forêts tropicales. Ces cartes produites en continu peuvent compléter les systèmes de surveillance existants, mais peuvent également nourrir les études nécessitant des informations spatiales et temporelles fines à large échelle.

Bibliographie

Alvarez-Berríos, N.L., Aide, T.M., 2015. Global demand for gold is another threat for tropical forests. Environ. Res. Lett. 10, 014006.

Arévalo, P., Olofsson, P., Woodcock, C.E., 2020. Continuous monitoring of land change activities and post-disturbance dynamics from Landsat time series: A test methodology for REDD+ reporting. Remote Sens. Environ. 238, 111051.

Armenteras, D., Cabrera, E., Rodríguez, N., Retana, J., 2013. National and regional determinants of tropical deforestation in Colombia. Reg. Environ. Change 13, 1181–1193.

Asner, G.P., Tupayachi, R., 2017. Accelerated losses of protected forests from gold mining in the Peruvian Amazon. Environ. Res. Lett. 12, 094004.

Ballère, M., 2021. Suivi de la forêt tropicale en quasi temps-réel à partir de données de télédétection, Journée Thématique du PNTS 2021.

Ballère, M., 2020. Orpaillage sur le fleuve frontalier entre la Guyane et le Suriname : télédétection des dommages causés par les barges aux rives du Maroni, https://labo.obs-mip.fr/multitemp/orpaillage-sur-le-fleuve-frontalier-entre-la-guyane-et-le-suriname-teledetection-des-dommages-causes-par-les-barges-aux-rives-du-maroni/.

Ballère, M., Bouvet, A., Mermoz, S., Koleck, T., Frison, P.-L., 2019. Détection quasi temps-réel de la déforestation en Guyane à partir des séries temporelles Sentinel-1, Journée de la recherche de l'ING 2019.

Ballère, M., Bouvet, A., Mermoz, S., Le Toan, T., Koleck, T., Bedeau, C., André, M., Forestier, E., Frison, P.-L., Lardeux, C., 2021a. SAR data for tropical forest disturbance alerts in French Guiana: benefit over optical imagery. Remote Sens. Environ. 252, 112159.

Ballère, M., Koleck, T., Mermoz, S., Bouvet, A., Shapiro, A., Frison, P.-L., Lardeux, C., Le Toan, T., 2019. Early Warning of forest disturbance using Sentinel-1 and Sentinel-2 data in order to improve the response. Study case: gold mining in Guiana Shield, Living Planet Symposium 2019.

Ballère, M., Le Toan, T., Frison, P.-L., Koleck, T., Mermoz, S., Bouvet, A., 2021b. Identification en quasi-temps réel des causes de déforestation en Guyane française, Journée de la recherche de l'ING 2021.

Ballère, M., Mermoz, S., Bouvet, A., Koleck, T., Le Toan, T., 2021c. Near-real-time identification of the drivers of deforestation in French Guiana, vEGU 2021.

Bank, W., 2012. Artisanal mining in critical ecosystems: A look at Gabon, Liberia, and Madagascar.

Baraloto, C., Alverga, P., Quispe, S.B., Barnes, G., Chura, N.B., da Silva, I.B., Castro, W., da Souza, H., de Souza Moll, I.E., Chilo, J.D.A., others, 2015. Effects of road infrastructure on forest value across a tri-national Amazonian frontier. Biol. Conserv. 191, 674–681.

Bosc, P.-M., others, 2018. Understanding smallholders in oil palm cultivation: a case study from Sumatra, in: Achieving Sustainable Cultivation of Oil Palm Volume 2. Burleigh Dodds Science Publishing, pp. 381–402.

Boucher, D., Elias, P., Lininger, K., May-Tobin, C., Roquemore, S., Saxon, E., others, 2011. The root of the problem: what's driving tropical deforestation today? Root Probl. Whats Driv. Trop. Deforestation Today.

Bouvet, A., Le Toan, T., 2011. Use of ENVISAT/ASAR wide-swath data for timely rice fields mapping in the Mekong River Delta. Remote Sens. Environ. 115, 1090–1101.

Bouvet, A., Le Toan, T., Floury, N., Macklin, T., 2010. An end-to-end error model for classification methods based on temporal change or polarization ratio of SAR intensities. IEEE Trans. Geosci. Remote Sens. 48, 3521–3538.

Bouvet, A., Mermoz, S., Ballère, M., Koleck, T., Le Toan, T., 2018. Use of the SAR shadowing effect for deforestation detection with Sentinel-1 time series. Remote Sens. 10, 1250.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Breiman, L., Friedman, J., Olshen, R., Stone, C., 1984. Classification and regression trees–crc press. Boca Raton Fla.

Bullock, E.L., Woodcock, C.E., Olofsson, P., 2020. Monitoring tropical forest degradation using spectral unmixing and Landsat time series analysis. Remote Sens. Environ. 238, 110968.

Caballero Espejo, J., Messinger, M., Román-Dañobeytia, F., Ascorra, C., Fernandez, L.E., Silman, M., 2018. Deforestation and forest degradation due to gold mining in the Peruvian Amazon: A 34-year perspective. Remote Sens. 10, 1903.

Canadell, J.G., Raupach, M.R., 2008. Managing Forests for Climate Change Mitigation. Science 320, 1456–1457. https://doi.org/10.1126/science.1155458

Chanthalath, X., others, 2017. Assessing the socioecological perspectives of eucalyptus cultivation and plantation expansion in Laos. Open Access Libr. J. 4, 1.

Chavy, A., Ferreira Dales Nava, A., Luz, S.L.B., Ramírez, J.D., Herrera, G., Vasconcelos dos Santos, T., Ginouves, M., Demar, M., Prévot, G., Guégan, J.-F., others, 2019. Ecological niche modelling for predicting the risk of cutaneous leishmaniasis in the Neotropical moist forest biome. PLoS Negl. Trop. Dis. 13, e0007629.

Cochran, W., 1977. Sampling Techniques Third Edition. pdf, 428.

Colchester, M., 1997. Guyana: fragile frontier. Race Cl. 38, 33–56.

Costenbader, J., Broadhead, J., Yasmi, Y., Durst, P.B., 2015. Drivers affecting forest change in the greater mekong subregion (GMS): an overview. FAO Rome Italy.

Curtis, P.G., Slay, C.M., Harris, N.L., Tyukavina, A., Hansen, M.C., 2018. Classifying drivers of global forest loss. Science 361, 1108–1111.

Dávalos, E., 2016. New answers to an old problem: Social investment and coca crops in Colombia. Int. J. Drug Policy 31, 121–130.

DeFries, R.S., Rudel, T., Uriarte, M., Hansen, M., 2010. Deforestation driven by urban population growth and agricultural trade in the twenty-first century. Nat. Geosci. 3, 178–181.

De Kom, J.F., van der Voet, G.B., de Wolff, F.A., 1998. Mercury exposure of maroon workers in the small scale gold mining in Suriname. Environ. Res. 77, 91–97.

Demaze, M.T., Manusset, S., 2008. L'agriculture itinérante sur brûlis en Guyane française: la fin des durabilités écologique et socioculturelle? Cah. D'Outre-Mer Rev. Géographie Bordx. 61, 31–48.

De Thoisy, B., Richard-Hansen, C., Goguillon, B., Joubert, P., Obstancias, J., Winterton, P., Brosse, S., 2010. Rapid evaluation of threats to biodiversity: human footprint score and large vertebrate species responses in French Guiana. Biodivers. Conserv. 19, 1567–1584.

Diele-Viegas, L.M., Rocha, C.F.D., 2020. Why releasing mining on Amazonian indigenous lands and the advance of agrobusiness is extremely harmful for the mitigation of world's climate change? Comment on Pereira et al.(Environmental Science & Policy 100 (2019) 8–12). Environ. Sci. Policy 103, 30–31.

Diniz, C.G., de Almeida Souza, A.A., Santos, D.C., Dias, M.C., da Luz, N.C., de Moraes, D.R.V., Maia, J.S., Gomes, A.R., da Silva Narvaes, I., Valeriano, D.M., others, 2015. DETER-B: The new Amazon near real-time deforestation detection system. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8, 3619–3628.

Eisler, R., 2004. Biogeochemical, health, and ecotoxicological perspectives on gold and gold mining. CRC press.

Évaluation des ressources forestières mondiales 2020, 2021. . FAO.

Finer, M., Novoa, S., Weisse, M.J., Petersen, R., Mascaro, J., Souto, T., Stearns, F., Martinez, R.G., 2018. Combating deforestation: From satellite to intervention. Science 360, 1303–1305.

Flores-Anderson, A.I., Herndon, K.E., Thapa, R.B., Cherrington, E., 2019. The SAR Handbook: Comprehensive Methodologies for Forest Monitoring and Biomass Estimation.

Forkuor, G., Ullmann, T., Griesbeck, M., 2020. Mapping and monitoring small-scale mining activities in Ghana using sentinel-1 time series (2015–2019). Remote Sens. 12, 911.

Forte, J., 1999. Karikuri: the evolving relationship of the Karinya people of Guyana to gold mining. New West Indian Guid. West-Indische Gids 73, 59–82.

Friedlingstein, P., Jones, M.W., O'sullivan, M., Andrew, R.M., Hauck, J., Peters, G.P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., others, 2019. Global carbon budget 2019. Earth Syst. Sci. Data 11, 1783–1838.

Giglio, L., Schroeder, W., Justice, C.O., 2016. The collection 6 MODIS active fire detection algorithm and fire products. Remote Sens. Environ. 178, 31–41.

Godar, J., Tizado, E.J., Pokorny, B., 2012. Who is responsible for deforestation in the Amazon? A spatially explicit analysis along the Transamazon Highway in Brazil. For. Ecol. Manag. 267, 58–73.

González, J., Cubillos, A., Chadid, M., Cubillos, A., Arias, M., Zúñiga, E., Joubert, F., Pérez, I., Berrío, V., 2018. Caracterización de las principales causas y agentes de la deforestación a nivel nacional período 2005–2015. Programa ONU-REDD Colomb.

Gregorutti, B., Michel, B., Saint-Pierre, P., 2014. Correlation and variable importance in random forests. arXiv :1310.5726.

Guégan, J.-F., Ayouba, A., Cappelle, J., de Thoisy, B., 2020. Forests and emerging infectious diseases: unleashing the beast within. Environ. Res. Lett. 15, 083007. https://doi.org/10.1088/1748-9326/ab8dd7

Hansen, M.C., Krylov, A., Tyukavina, A., Potapov, P.V., Turubanova, S., Zutta, B., Ifo, S., Margono, B., Stolle, F., Moore, R., 2016. Humid tropical forest disturbance alerts using Landsat data. Environ. Res. Lett. 11, 034008.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. Science 342, 850–853. https://doi.org/10.1126/science.1244693

Hayashi, M., Nagatani, I., Watanabe, T., Tadono, T., Miyoshi, H., Watanabe, M., Koyama, C., Shimada, M., Ogawa, T., Ishii, K., others, 2017. Development of Early Warning System Using ALOS-2/PALSAR-2 Data to Detect and Prevent Deforestation, in: AGU Fall Meeting Abstracts. p. B31D–2013.

Hirschmugl, M., Deutscher, J., Sobe, C., Bouvet, A., Mermoz, S., Schardt, M., 2020. Use of SAR and optical time series for tropical forest disturbance mapping. Remote Sens. 12, 727.

Hosonuma, N., Herold, M., De Sy, V., De Fries, R.S., Brockhaus, M., Verchot, L., Angelsen, A., Romijn, E., 2012. An assessment of deforestation and forest degradation drivers in developing countries. Environ. Res. Lett. 7, 044009. https://doi.org/10.1088/1748-9326/7/4/044009

Houghton, R.A., Nassikas, A.A., 2017. Global and regional fluxes of carbon from land use and land cover change 1850–2015. Glob. Biogeochem. Cycles 31, 456–472.

Ickowitz, A., Slayback, D., Asanzi, P., Nasi, R., 2015. Agriculture and deforestation in the Democratic Republic of the Congo: A synthesis of the current state of knowledge.

Jelsma, I., Woittiez, L.S., Ollivier, J., Dharmawan, A.H., 2019. Do wealthy farmers implement better agricultural practices? An assessment of implementation of Good Agricultural Practices among different types of independent oil palm smallholders in Riau, Indonesia. Agric. Syst. 170, 63–76.

Joshi, N., Mitchard, E.T., Woo, N., Torres, J., Moll-Rocek, J., Ehammer, A., Collins, M., Jepsen, M.R., Fensholt, R., 2015. Mapping dynamics of deforestation and forest degradation in tropical forests using radar satellite data. Environ. Res. Lett. 10, 034014.

Kalamandeen, M., Gloor, E., Mitchard, E., Quincey, D., Ziv, G., Spracklen, D., Spracklen, B., Adami, M., Aragão, L.E., Galbraith, D., 2018. Pervasive rise of small-scale deforestation in Amazonia. Sci. Rep. 8, 1–10.

Keenan, R.J., Reams, G.A., Achard, F., de Freitas, J.V., Grainger, A., Lindquist, E., 2015. Dynamics of global forest area: Results from the FAO Global Forest Resources Assessment 2015. For. Ecol. Manag. 352, 9–20.

Kissinger, G., Herold, M., De Sy, V., 2012. Drivers of deforestation and forest degradation: a synthesis report for REDD+ policymakers. Citeseer.

Koleck, T., Ballère, M., Marie-Sainte, W., 2019. A Multipurpose Open Source Processing Chain for Sentinel-1 Time Series. Proc. Living Planet Milan Italy 13–17 May 2019.

Laing, T., 2019. Small man goes where the large fears to tread: Mining in Guyana: 1990-2018. Resour. Policy 63, 101426.

Lardeux, C., 2008. Apport des données radar polarimétriques pour la cartographie en milieu tropical (PhD Thesis). Université Paris-Est.

Lardeux, C., Kemavo, A., Rageade, M., Rahm, M., Frison, P.-L., Rudant, J.-P., 2019. Mise en øeuvre d'outils open source pour le suivi opérationnel de l'occupation des sols et de la déforestation à partir des données Sentinel radar et optique: études de cas en Guyane et au Togo. Rev. Fr. Photogrammétrie Télédétection 59–70.

Laurance, W.F., Carolina Useche, D., Rendeiro, J., Kalka, M., Bradshaw, C.J.A., Sloan, S.P., Laurance, S.G., Campbell, M., Abernethy, K., Alvarez, P., Arroyo-Rodriguez, V., Ashton, P., Benítez-Malvido, J., Blom, A., Bobo, K.S., Cannon, C.H., Cao, M., Carroll, R., Chapman, C., Coates, R., Cords, M., Danielsen, F., De Dijn, B., Dinerstein, E., Donnelly, M.A., Edwards, D., Edwards, F., Farwig, N., Fashing, P., Forget, P.-M., Foster, M., Gale, G., Harris, D., Harrison, R., Hart, J., Karpanty, S., John Kress, W., Krishnaswamy, J., Logsdon, W., Lovett, J., Magnusson, W., Maisels, F., Marshall, A.R., McClearn, D., Mudappa, D., Nielsen, M.R., Pearson, R., Pitman, N., van der Ploeg, J., Plumptre, A.,

Poulsen, J., Quesada, M., Rainey, H., Robinson, D., Roetgers, C., Rovero, F., Scatena, F., Schulze, C., Sheil, D., Struhsaker, T., Terborgh, J., Thomas, D., Timm, R., Nicolas Urbina-Cardona, J., Vasudevan, K., Joseph Wright, S., Carlos Arias-G., J., Arroyo, L., Ashton, M., Auzel, P., Babaasa, D., Babweteera, F., Baker, P., Banki, O., Bass, M., Bila-Isia, I., Blake, S., Brockelman, W., Brokaw, N., Brühl, C.A., Bunyavejchewin, S., Chao, J.-T., Chave, J., Chellam, R., Clark, C.J., Clavijo, J., Congdon, R., Corlett, R., Dattaraja, H.S., Dave, C., Davies, G., de Mello Beisiegel, B., de Nazaré Paes da Silva, R., Di Fiore, A., Diesmos, A., Dirzo, R., Doran-Sheehy, D., Eaton, M., Emmons, L., Estrada, A., Ewango, C., Fedigan, L., Feer, F., Fruth, B., Giacalone Willis, J., Goodale, U., Goodman, S., Guix, J.C., Guthiga, P., Haber, W., Hamer, K., Herbinger, I., Hill, J., Huang, Z., Fang Sun, I., Ickes, K., Itoh, A., Ivanauskas, N., Jackes, B., Janovec, J., Janzen, D., Jiangming, M., Jin, C., Jones, T., Justiniano, H., Kalko, E., Kasangaki, A., Killeen, T., King, H., Klop, E., Knott, C., Koné, I., Kudavidanage, E., Lahoz da Silva Ribeiro, J., Lattke, J., Laval, R., Lawton, R., Leal, M., Leighton, M., Lentino, M., Leonel, C., Lindsell, J., Ling-Ling, L., Eduard Linsenmair, K., Losos, E., Lugo, A., Lwanga, J., Mack, A.L., Martins, M., Scott McGraw, W., McNab, R., Montag, L., Myers Thompson, J., Nabe-Nielsen, J., Nakagawa, M., Nepal, S., Norconk, M., Novotny, V., O'Donnell, S., Opiang, M., Ouboter, P., Parker, K., Parthasarathy, N., Pisciotta, K., Prawiradilaga, D., Pringle, C., Rajathurai, S., Reichard, U., Reinartz, G., Renton, K., Reynolds, G., Reynolds, V., Riley, E., Rödel, M.-O., Rothman, J., Round, P., Sakai, S., Sanaiotti, T., Savini, T., Schaab, G., Seidensticker, J., Siaka, A., Silman, M.R., Smith, T.B., de Almeida, S.S., Sodhi, N., Stanford, C., Stewart, K., Stokes, E., Stoner, K.E., Sukumar, R., Surbeck, M., Tobler, M., Tscharntke, T., Turkalo, A., Umapathy, G., van Weerd, M., Vega Rivera, J., Venkataraman, M., Venn, L., Verea, C., Volkmer de Castilho, C., Waltert, M., Wang, B., Watts, D., Weber, W., West, P., Whitacre, D., Whitney, K., Wilkie, D., Williams, S., Wright, D.D., Wright, P., Xiankai, L., Yonzon, P., Zamzani, F., 2012. Averting biodiversity collapse in tropical forest protected areas. Nature 489, 290–294. https://doi.org/10.1038/nature11318

Leblois, A., Damette, O., Wolfersberger, J., 2017. What has driven deforestation in developing countries since the 2000s? Evidence from new remote-sensing data. World Dev. 92, 82–102.

Lee, J.-S., 1980. Digital image enhancement and noise filtering by use of local statistics. IEEE Trans. Pattern Anal. Mach. Intell. 165–168.

Le Quéré, C., Raupach, M.R., Canadell, J.G., Marland, G., Bopp, L., Ciais, P., Conway, T.J., Doney, S.C., Feely, R.A., Foster, P., Friedlingstein, P., Gurney, K., Houghton, R.A., House, J.I., Huntingford, C., Levy, P.E., Lomas, M.R., Majkut, J., Metzl, N., Ometto, J.P., Peters, G.P., Prentice, I.C., Randerson, J.T., Running, S.W., Sarmiento, J.L., Schuster, U., Sitch, S.,

Takahashi, T., Viovy, N., van der Werf, G.R., Woodward, F.I., 2009. Trends in the sources and sinks of carbon dioxide. Nat. Geosci. 2, 831–836. https://doi.org/10.1038/ngeo689

Le Toan, T., Ribbes, F., Wang, L.-F., Floury, N., Ding, K.-H., Kong, J.A., Fujita, M., Kurosu, T., 1997. Rice crop mapping and monitoring using ERS-1 data based on experiment and modeling results. IEEE Trans. Geosci. Remote Sens. 35, 41–56.

Lewis, S.L., Edwards, D.P., Galbraith, D., 2015. Increasing human dominance of tropical forests. Science 349, 827–832. https://doi.org/10.1126/science.aaa9932

Li, F.-K., Croft, C., Held, D.N., 1983. Comparison of several techniques to obtain multiple-look SAR imagery. IEEE Trans. Geosci. Remote Sens. 370–375.

Li, G., Lu, D., Moran, E., Calvi, M.F., Dutra, L.V., Batistella, M., 2019. Examining deforestation and agropasture dynamics along the Brazilian TransAmazon Highway using multitemporal Landsat imagery. GIScience Remote Sens. 56, 161–183.

Linarès, S., André, M., 2020. La lutte contre l'orpaillage illégal en Guyane. Stratégie et engagement de l'État. Société Géologique Fr. REVUE OFFICIELLE DE LA SOCIÉTÉ GÉOLOGIQUE DE FRANCE.

Linarès, S., Joubert, P., Gond, V., 2008. Contre l'orpaillage clandestin. Espac. Nat.

Lohberger, S., Stängel, M., Atwood, E.C., Siegert, F., 2018. Spatial evaluation of Indonesia's 2015 fire-affected area and estimated carbon emissions using Sentinel-1. Glob. Change Biol. 24, 644–654.

Louppe, G., 2014. Understanding random forests: From theory to practice. ArXiv Prepr. ArXiv14077502.

Maître, H., Adragna, F., Ducrot, D., Garello, R., Le Caillec, J., Le Hégarat-Mascle, S., Lopès, A., Nicolas, J., Tupin, F., Trouvé, E., 2001. Traitement des images de RSO. Hermès.

Massonnet, D., Souyris, J.-C., 2008. Imaging with synthetic aperture radar. EPFL press.

Mbow, H.-O.P., Reisinger, A., Canadell, J., O'Brien, P., 2017. Special Report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems (SR2). Ginevra IPCC.

Mermoz, S., Ballère, M., Bouvet, A., Koleck, T., Lardeux, C., Le Toan, T., 2018. Near-real time forest disturbances detection in the Amazonian wet forest using Sentinel-1 images. ForestSAT 2018 Symp.

Mermoz, S., Bouvet, A., Koleck, T., Ballère, M., Le Toan, T., 2021. Continuous Detection of Forest Loss in Vietnam, Laos, and Cambodia Using Sentinel-1 Data. Remote Sens. 13(23), 4877.

Mermoz, S., Le Toan, T., 2016. Forest disturbances and regrowth assessment using ALOS PALSAR data from 2007 to 2010 in Vietnam, Cambodia and Lao PDR. Remote Sens. 8, 217.

Miller, J.R., Lechler, P.J., Bridge, G., 2003. Mercury contamination of alluvial sediments within the Essequibo and Mazaruni river basins, Guyana. Water. Air. Soil Pollut. 148, 139–166.

Milodowski, D., Mitchard, E., Williams, M., 2017. Forest loss maps from regional satellite monitoring systematically underestimate deforestation in two rapidly changing parts of the Amazon. Environ. Res. Lett. 12, 094003.

Mittermeier, R.A., Myers, N., Thomsen, J.B., da Fonseca, G.A.B., Olivieri, S., 1998. Biodiversity Hotspots and Major Tropical Wilderness Areas: Approaches to Setting Conservation Priorities. Conserv. Biol. 12, 516–520. https://doi.org/10.1046/j.1523-1739.1998.012003516.x

Mora, B., 2018. User needs assessment for forest change early warning systems.

Murillo-Sandoval, P.J., Hilker, T., Krawchuk, M.A., Van Den Hoek, J., 2018. Detecting and attributing drivers of forest disturbance in the Colombian andes using landsat time-series. Forests 9, 269.

Nicolau, A.P., Flores-Anderson, A., Griffin, R., Herndon, K., Meyer, F.J., 2021. Assessing SAR C-band data to effectively distinguish modified land uses in a heavily disturbed Amazon forest. Int. J. Appl. Earth Obs. Geoinformation 94, 102214.

OBS Habitat, 2015. L'urbanisation spontanée en Guyane : recensement du phénomène en 2015.

Oliver, C., Quegan, S., 2004. Understanding synthetic aperture radar images. SciTech Publishing.

Olofsson, P., Arévalo, P., Espejo, A.B., Green, C., Lindquist, E., McRoberts, R.E., Sanz, M.J., 2020. Mitigating the effects of omission errors on area and area change estimates. Remote Sens. Environ. 236, 111492.

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 148, 42–57.

Olofsson, P., Kuemmerle, T., Griffiths, P., Knorn, J., Baccini, A., Gancz, V., Blujdea, V., Houghton, R., Abrudan, I., Woodcock, C., 2011. Carbon implications of forest restitution in post-socialist Romania. Environ. Res. Lett. 6, 045202.

Olofsson, P., Torchinava, P., Woodcock, C.E., Baccini, A., Houghton, R.A., Ozdogan, M., Zhao, F., Yang, X., 2010. Implications of land use change on the national terrestrial carbon budget of Georgia. Carbon Balance Manag. 5, 1–13.

Ordway, E.M., Naylor, R.L., Nkongho, R.N., Lambin, E.F., 2017. Oil palm expansion in Cameroon: Insights into sustainability opportunities and challenges in Africa. Glob. Environ. Change 47, 190–200.

Pacheco, P., Mo, K., Dudley, N., Shapiro, A., Aguilar-Amuchastegui, N., Ling, P., Anderson, C., Marx, A., 2021. Deforestation fronts: Drivers and responses in a changing world. WWF Gland Switz.

Palmer, C.J., Validum, L., Loeffke, B., Laubach, H.E., Mitchell, C., Cummings, R., Cuadrado, R.R., 2002. HIV prevalence in a gold mining camp in the Amazon region, Guyana. Emerg. Infect. Dis. 8, 330.

Perz, S., Brilhante, S., Brown, F., Caldas, M., Ikeda, S., Mendoza, E., Overdevest, C., Reis, V., Reyes, J.F., Rojas, D., others, 2008. Road building, land use and climate change: prospects for environmental governance in the Amazon. Philos. Trans. R. Soc. B Biol. Sci. 363, 1889–1895.

Phompila, C., Lewis, M., Ostendorf, B., Clarke, K., 2017. Forest cover changes in Lao tropical forests: physical and socio-economic factors are the most important drivers. Land 6, 23.

Piotrowski, M., 2019. Nearing the tipping point: Drivers of deforestation in the Amazon Region. Inter-Am. Dialogue Wash. WA USA.

Potapov, P., Siddiqui, B., Iqbal, Z., Aziz, T., Zzaman, B., Islam, A., Pickens, A., Talero, Y., Tyukavina, A., Turubanova, S., others, 2017. Comprehensive monitoring of Bangladesh tree cover inside and outside of forests, 2000–2014. Environ. Res. Lett. 12, 104015.

Puyravaud, J.-P., 2003. Standardizing the calculation of the annual rate of deforestation. For. Ecol. Manag. 177, 593–596.

Quegan, S., Yu, J.J., 2001. Filtering of multichannel SAR images. IEEE Trans. Geosci. Remote Sens. 39, 2373–2379.

Rahm, M., Jullian, B., Lauger, A., De Carvalho, R., Vale, L., Totaram, J., Cort, K., Djojodikromo, M., Hardjoprajitno, M., Neri, S., others, 2015. Monitoring the impact of gold mining on the forest cover and freshwater in the Guiana Shield. Ref. Year 20.

Rahm, M., Smartt, T., Totaram, J., Sukhu, B., Thornhill-Gillis, D., Amin, N., Thomas, R., Sookdeo, C., Paloeng, C., Kasanpawiro, C., Moe Soe Let V, Hoepel I, Pichot L, Bedeau C, Farias P, Carvalho R, Weber JL, Lardeux C, 2020. Mapping land use land cover change in the Guiana shield from 2000 to 2015. ECOSEO Proj.

Reiche, J., Hamunyela, E., Verbesselt, J., Hoekman, D., Herold, M., 2018a. Improving near-real time deforestation monitoring in tropical dry forests by combining dense Sentinel-1 time series with Landsat and ALOS-2 PALSAR-2. Remote Sens. Environ. 204, 147–161.

Reiche, J., Verhoeven, R., Verbesselt, J., Hamunyela, E., Wielaard, N., Herold, M., 2018b. Characterizing tropical forest cover loss using dense Sentinel-1 data and active fire alerts. Remote Sens. 10, 777.

Reymondin, L., Paz-García, P., Tello, J.J., Coca-Castro, A., Bautista, O., Menzinger, B., 2019. User Manual. Training Manual Terra-i.

Rignot, E.J., Van Zyl, J.J., 1993. Change detection techniques for ERS-1 SAR data. IEEE Trans. Geosci. Remote Sens. 31, 896–906.

Rojas Briceño, N.B., Barboza Castillo, E., Maicelo Quintana, J.L., Oliva Cruz, S.M., Salas López, R., others, 2019. Deforestation in the Peruvian Amazon: indexes of Land Cover/Land Use (LC/LU) changes based on GIS. Bol. Asoc. Geógrafos Esp.

Seymour, F., Harris, N.L., 2019. Reducing tropical deforestation. Science 365, 756–757.

Simard, M., Pinto, N., Fisher, J.B., Baccini, A., 2011. Mapping forest canopy height globally with spaceborne lidar. J. Geophys. Res. Biogeosciences 116.

Simmons, B.A., Law, E.A., Marcos-Martinez, R., Bryan, B.A., McAlpine, C., Wilson, K.A., 2018. Spatial and temporal patterns of land clearing during policy change. Land Use Policy 75, 399–410.

Sloan, S., Alamgir, M., Campbell, M.J., Setyawati, T., Laurance, W.F., 2019. Development corridors and remnant-forest conservation in Sumatra, Indonesia. Trop. Conserv. Sci. 12, 1940082919889509.

Smith, B.T., McCormack, J.E., Cuervo, A.M., Hickerson, M.J., Aleixo, A., Cadena, C.D., Pérez-Emán, J., Burney, C.W., Xie, X., Harvey, M.G., Faircloth, B.C., Glenn, T.C., Derryberry, E.P., Prejean, J., Fields, S., Brumfield, R.T., 2014. The drivers of tropical speciation. Nature 515, 406–409. https://doi.org/10.1038/nature13687

Stern, N., Stern, N.H., 2007. The economics of climate change: the Stern review. cambridge University press.

Tanase, M.A., Aponte, C., Mermoz, S., Bouvet, A., Le Toan, T., Heurich, M., 2018. Detection of windthrows and insect outbreaks by L-band SAR: A case study in the Bavarian Forest National Park. Remote Sens. Environ. 209, 700–711.

Tegegne, Y.T., Lindner, M., Fobissie, K., Kanninen, M., 2016. Evolution of drivers of deforestation and forest degradation in the Congo Basin forests: Exploring possible policy options to address forest loss. Land Use Policy 51, 312–324.

The State of the World's Forests 2020. In brief, 2020. . FAO and UNEP.

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Floury, N., Brown, M., others, 2012. GMES Sentinel-1 mission. Remote Sens. Environ. 120, 9–24.

Tyukavina, A., Stehman, S., Potapov, P., Turubanova, S., Baccini, A., Goetz, S., Laporte, N., Houghton, R., Hansen, M., 2013. National-scale estimation of gross forest aboveground

carbon loss: a case study of the Democratic Republic of the Congo. Environ. Res. Lett. 8, 044039.

Ulaby, F., Moore, R., Fung, A., 1982. Microwave remote sensing: Active and passive. Volume 2-Radar remote sensing and surface scattering and emission theory.

Valeriano, D.M., Mello, E.M., Moreira, J.C., Shimabukuro, Y.E., Duarte, V., Souza, I., Santos, J., Barbosa, C.C., Souza, R., 2004. Monitoring tropical forest from space: the PRODES digital project. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 35, 272–274.

Weisse, M., Petersen, R., Sargent, S., Gibbes, S., 2017. Places to Watch: Identifying High-Priority Forest Disturbance from Near–Real Time Satellite Data.

Wheeler, D., Hammer, D., Kraft, R., Steele, A., 2014. Satellite-based forest clearing detection in the Brazilian Amazon: FORMA, DETER, and PRODES. World Resour. Inst. Wash. DC USA.

WRI - Global Forest Watch Dashboard, 2020.

WWF, 2018. LUTTE CONTRE L'ORPAILLAGE ILLÉGAL EN GUYANE : orientations pour une efficacité renforcée.

Zeppetello, L.R.V., Parsons, L.A., Spector, J.T., Naylor, R.L., Battisti, D.S., Masuda, Y.J., Wolff, N.H., 2020. Large scale tropical deforestation drives extreme warming. Environ. Res. Lett. 15, 084012.

Annexes

1. Mesures communes d'évaluation

Il est possible d'évaluer une carte de changement à la façon d'un modèle de classification supervisée (en rapportant les observations attendues et les observations observées). La matrice de confusion est un outil utile pour mettre en forme ces observations sous la forme matricielle. Les classes réelles (données de références) sont ici les lignes, tandis que les classes prédites sont les colonnes. Chaque cellule contient le nombre d'observations de la classe de la ligne associée prédites comme étant de la classe de la colonne associée. La matrice de confusion utilisée pour l'évaluation des détections des pertes forestières est présentée dans le Tableau 10.

Tableau 10 - Matrice de confusion associée à la détection de la perte de forêt

		Classe estimée	
		Perte de forêt (1)	Forêt intacte (2)
Classe réelle	Perte de forêt (1)	Perte de forêt bien détectée $n_{1,1}$	Perte de forêt non-détectée (omission) $n_{1,2}$
	Forêt intacte (2)	Forêt intacte détectée comme perte de forêt (commission) $n_{2,1}$	Forêt intacte non-détectée comme perte de forêt $n_{2,2}$

Les éléments diagonaux contiennent les pixels associés à la bonne classe : la précision de la méthode sera d'autant meilleure que sa matrice de confusion sera proche d'une matrice diagonale. Les éléments hors diagonale représentent ici les erreurs d'omission et de commission. Il est maintenant possible de définir les termes de précision, rappel et précision globale.

La **précision** (ou user's accuracy) de la classe i (i∈(1, N) avec n le nombre de classe) est la proportion d'observations i correctement classées par rapport au nombre total d'observations classées i :

$$précision_i = \sum_{j=1}^{N} \frac{n_{i,i}}{n_{i,j}}$$

Elle correspond à la pertinence de l'information remontée, mais ne donne pas la quantité d'information remontée : une précision de 1 pour la classe i signifie que tous les éléments prédits i sont vrais. En revanche, tous les éléments effectivement i n'auront peut-être pas été prédits i.

Le taux de **commissions** (ou de fausses alarmes) est donné par :

$$commission_i = 1 - précision_i$$

Le **rappel** (ou producer's accucary) de la classe i est la proportion d'observations i correctement classées par rapport au nombre total d'observations effectivement i :

$$rappel_i = \sum_{j=1}^{N} \frac{n_{i,i}}{n_{j,i}}$$

Il est complémentaire de la précision et correspond à la quantité d'information effectivement remontée : un rappel de 1 indiquera que tous les points appartenant effectivement à la classe i ont été identifiés comme appartenant à la classe i. En revanche, il peut y avoir des points n'appartenant pas à la classe i qui ont été prédits i.

Le taux d'**omissions** est donné par :

$$omission_i = 1 - rappel_i$$

Le **F-score** est la moyenne harmonique de la précision et du rappel :

$$F - score_i = \frac{2 \times précision_i \times rappel_i}{précision_i + rappel_i}$$

Si la précision ou le rappel baisse, le F-score diminuera également.

La **précision globale** (ou overall accuracy) est la proportion d'éléments correctement classés par rapport au nombre total d'éléments évalués n :

$$précision\ globale = \sum_{j=1}^{N} \frac{n_{i,i}}{n}$$

La précision globale peut être trompeuse au moment de l'interprétation des résultats si les classes sont disproportionnées.

La statistique **kappa** compare une précision observée avec une précision attendue. Ce score est utilisé pour évaluer un classificateur en prenant en compte le hasard (accord avec un classificateur aléatoire), ce qui signifie qu'il est généralement moins trompeur que l'utilisation simple de la précision globale. En effet, une précision observée de 80% est beaucoup moins impressionnante

avec une précision attendue de 75% qu'avec une précision attendue de 50%. Celui-ci est calculé de la manière suivante :

$$kappa = \frac{po - pa}{1 - pa}$$

Avec po la précision globale observée et pa la précision globale attendue. Ce score est compris entre -1 et 1, et peut être interprété de cette manière :

Accord	Kappa
Excellent	> 0.81
Bon	0.61 - 0.8
Modéré	0.41 - 0.6
Faible	0.21 - 0.4
Mauvais	0 – 0.2
Très mauvais	< 0

2. Description des indicateurs créés pour la caractérisation du déboisement en Guyane

Les indicateurs utilisés pour le développement d'un arbre de décision visant à caractériser les déforestations en temps-réel sont présentés dans le Tableau 11 :

Ces indicateurs ont été construits de manière à être exhaustifs sur la Guyane. Aucun indicateur associé à l'exploitation forestière n'a été créé car il n'était pas possible de le faire globalement sur le territoire avec les données dont nous disposions.

Processing

Les descripteurs géométriques (surface, périmètre) ont été calculés via Qgis. Le calcul des distances s'est fait en utilisant gdal proximity sur les rasters binaires déduits de la sélection des polygones dans les fichiers vecteurs associés.

Tableau 11 - Indicateurs utilisés pour le développement d'un arbre de décision visant à caractériser les déforestations en temps-réel en Guyane, avec leurs sources. La source est décrite de la manière suivante : NOM_DE_LA_COUCHE : liste des conditions sur les classes sélectionnées (séparées par une virgule). DI signifie : date à laquelle on veut l'état de la variable.

Indicateur	Source	Remarques
Distance à l'orpaillage	OAM : date_obs < DI, deb_expl < DI, date_ref < DI AgriPAG : Classe « OR »	On écarte la couche Ecoseo qui est une image de 2015 de la couche OAM. On considère toutes les zones qui ont un jour été de l'orpaillage
Distance aux titres miniers	Camino : date_debut < DI	
Distance aux abattis	AgriPAG : TYPE[DI] = AF, A1, A2 ou AG ECOSEO : ECO_N2_TX = Shifting agriculture	
Distance aux vergers	BD TOPO - ZONE_DE_VEGETATION : Nature = Vergers, DATE_CREAT < DI	Entre 2016 et 2018, seulement 2 entités diffèrent

Indicateur	Source	Remarques
Distance aux cultures ligneuses	BD TOPO - ZONE_DE_VEGETATION : Nature = Vergers, DATE_CREAT < DI ECOSEO : ECO_N2_TX = Agriculture, permanent crops (Woody crops)	Entre 2016 et 2018, seulement 2 entités diffèrent
Distance aux rizières	BD TOPO - ZONE_DE_VEGETATION : Nature = Rizières	Aucune entité ne diffère entre 2016 et 2018. La classe ECOSEO Herbaceous crops est égale à la classe rizière de la BD TOPO.
Distance aux pâturages	AgriPAG : TYPE[DI] = PA	
Distance aux bâtiments agricoles	BD TOPO - BATIMENTS : USAGE1 = Agricole, DATE_CREAT < DI	
Distance aux zones agricoles	AgriPAG : TYPE[DI] = AF, A1, A2, AG ou PA BD TOPO - BATIMENTS : USAGE1 = Agricole, DATE_CREAT < DI BDTOPO- BD TOPO - ZONE_DE_VEGETATION : Nature = Vergers ou Rizières ECOSEO : ECO_N2_TX = Agriculture, permanent crops (Woody crops)	

Indicateur	Source	Remarques
Distance aux zones agricoles sans rizières	AgriPAG : TYPE[DI] = AF, A1, A2, AG ou PA BD TOPO - BATIMENTS : USAGE1 = Agricole, DATE_CREAT < DI BDTOPO- BD TOPO - ZONE_DE_VEGETATION : Nature = Vergers ECOSEO : ECO_N2_TX = Agriculture, permanent crops (Woody crops) ou Shifting agriculture	
Distance aux zones agricoles vivrières	AgriPAG : TYPE[DI] = AF, A1, A2, AG ou PA ECOSEO : ECO_N2_TX = Shifting agriculture	
Distance aux exploitations agricoles	BD TOPO - BATIMENTS : USAGE1 = Agricole, DATE_CREAT < DI BDTOPO- BD TOPO - ZONE_DE_VEGETATION : Nature = Vergers, Rizières ECOSEO : ECO_N2_TX = Agriculture, permanent crops (Woody crops)	

Indicateur	Source	Remarques
Distance aux habitations population locale	BD TOPO - ZONE_D_HABITATION : Fictif = Non, Nature != Habitat temporaire	Couche séparée manuellement avec connaissance des experts. Vu l'hétérogénéité entre la BD TOPO et la classe Settlements d'ECOSEO et la couche HA d'AgriPAG, nous decidons de garder définir les habitations des populations locales avec la BD TOPO seulement. Aucune entité ne diffère entre 2016 et 2018.
Distance aux établissements humains	ECOSEO : ECO_N2_TX = Settlements	Aucune entité ne diffère entre 2016 et 2018 car ECOSEO est une image de l'année 2015.
Distance aux routes	BD TOPO - TRONCON_ROUTE : Nature = Route, Etat = En service	La classe infrastructures d'ECOSEO intègre des routes d'importances différentes (par rapport à la BD TOPO), mais n'est pas exhaustive. Nous décidons donc de ne pas l'utiliser. La classe Voierie d'AgriPAG est inclus dans la BD TOPO, donc on ne l'utilise pas. La date d'apparition n'est pas disponible et celle de confirmation, pas toujours. La date de création des routes est souvent antérieure à 2016 et celles postérieures à 2016 sont souvent spatialement proches des autres. Les différences notables concernent la route de l'Est et la D6. Après

Indicateur	Source	Remarques
		vérification, ces routes existaient avant 2016. Aucune entité ne diffère entre 2016 et 2018.
Distance aux routes principales	BD TOPO - TRONCON_ROUTE : Nature = Route, Etat = En service, Importance = 2 ou 3	Pour la même raison que pour les routes, aucune entité ne diffère entre 2016 et 2018.
Distance aux chemins	BD TOPO - TRONCON_ROUTE : Nature = Chemin ONF - Routes_ONF : STATUT = Chemin d'exploitation ferme à la circulation publique, DATE_CREAT = DI	Mêmes observations pour les chemins. Les entités qui diffèrent entre 2016 et 2018 proviennent de la couche de l'ONF.
Distance aux axe routiers (routes + chemins)	BD TOPO - TRONCON_ROUTE : Nature = Route ou Chemin ONF - Routes_ONF : STATUT = Chemin d'exploitation ferme à la circulation publique, DATE_CREAT = DI	Les entités qui diffèrent entre 2016 et 2018 proviennent de la couche de l'ONF.

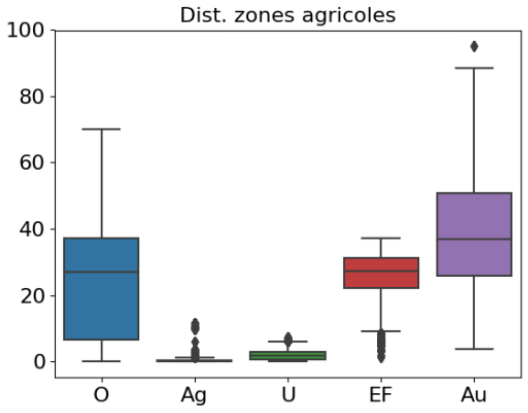
Indicateur	Source	Remarques
Distance aux sentiers	BD TOPO - TRONCON_ROUTE : Nature = Sentier	Les sentiers présentent la même similarité mais un groupe à l'est a été numérisé après 2016. Le sentier de la Reine et le sentier de la Savane Roche font partis des exceptions mais il est difficile de vérifier leur date d'apparition. Dans ces conditions, nous gardons la même couche de sentiers pour 2016, 2017 et 2018.
Distance à la Green Belt	BRGM - GEO500K_CART_81_S_FGEOL_2972 : DESCR = Ceinture de roches vertes (Paramaca)	Couche fixe.
Distance aux cours d'eau par catégorie	BD TOPO - COURS_D_EAU : Importance *	Date d'apparition pas disponible. On suppose que les surfaces en eau n'évoluent pas beaucoup dans le temps à une échelle annuelle. Aucune entité ne diffère entre 2016 et 2018.
Distance aux cours d'eau par catégories cumulées	BD TOPO - COURS_D_EAU : Importance *	Aucune entité ne diffère entre 2016 et 2018.
Distance aux surfaces hydrographiques permanentes	BD TOPO - SURFACE_HYDROGRAPHIQUE : Persistance = Permanant	Aucune entité ne diffère entre 2016 et 2018.

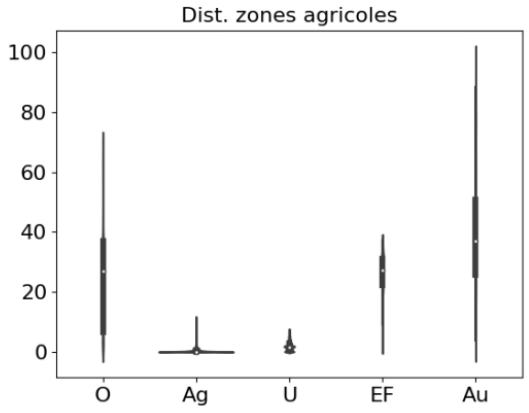
Indicateur	Source	Remarques
Distance aux surfaces hydrographiques intermittentes	BD TOPO - SURFACE_HYDROGRAPHIQUE : Persistance = Intermittent	Aucune entité ne diffère entre 2016 et 2018.
Distance aux surfaces hydrographiques naturelles	BD TOPO - SURFACE_HYDROGRAPHIQUE : Origine = Naturelle	Aucune entité ne diffère entre 2016 et 2018.
Distance aux surfaces hydrographiques artificielles	BD TOPO - SURFACE_HYDROGRAPHIQUE : Origine = Artificielle	Aucune entité ne diffère entre 2016 et 2018.
Distance aux surfaces hydrographiques	BD TOPO - SURFACE_HYDROGRAPHIQUE	Aucune entité ne diffère entre 2016 et 2018.
Distance aux zones humides ouvertes	ECOSEO : Open wet land	Aucune entité ne diffère entre 2016 et 2018 car la couche ECOSEO est une image de 2015.
Distance aux eaux intérieures	ECOSEO : Inland water bodies	Aucune entité ne diffère entre 2016 et 2018 car la couche ECOSEO est une image de 2015.
Elévation	SRTM	Couche fixe
Pente	Dérivé SRTM	Couche fixe
Surface	Dérivé des cartes de déforestations produites avec Sentinel-1	
Surface connexe	Dérivé des cartes de déforestations produites avec Sentinel-1	
Périmètre	Dérivé des cartes de déforestations produites avec Sentinel-1	

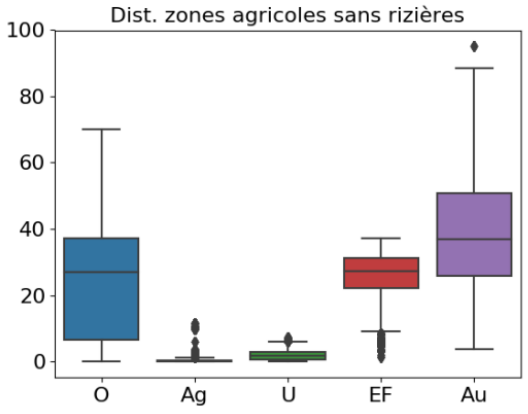
Indicateur	Source	Remarques
Périmètre connexe	Dérivé des cartes de déforestations produites avec Sentinel-1	
Compacité (périmètre^2/aire)	Dérivé des cartes de déforestations produites avec Sentinel-1	
Compacité connexe	Dérivé des cartes de déforestations produites avec Sentinel-1	
Finesse (aire/périmètre^2)	Dérivé des cartes de déforestations produites avec Sentinel-1	
Finesse connexe	Dérivé des cartes de déforestations produites avec Sentinel-1	
Moment de la coupe	Dérivé des cartes de déforestations produites avec Sentinel-1	

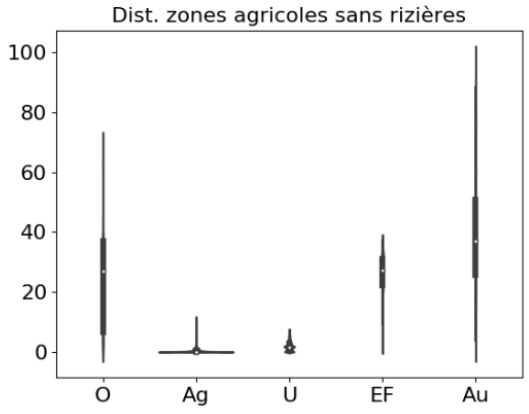
3. Visualisations graphiques des indicateurs de caractérisation par type de déboisement

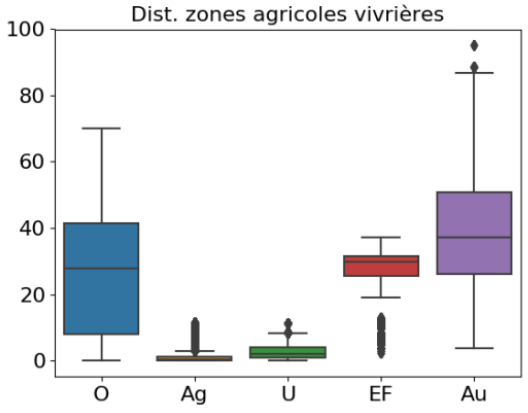
Boîtes à moustaches et graphiques en violon des 51 indicateurs testés en regard des 5 classes (orpaillage, agriculture, urbanisation, exploitation forestière et autre).

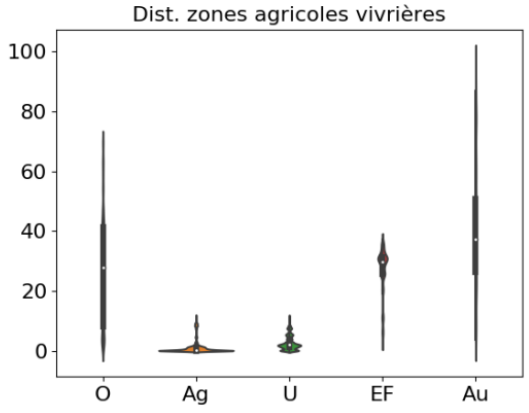


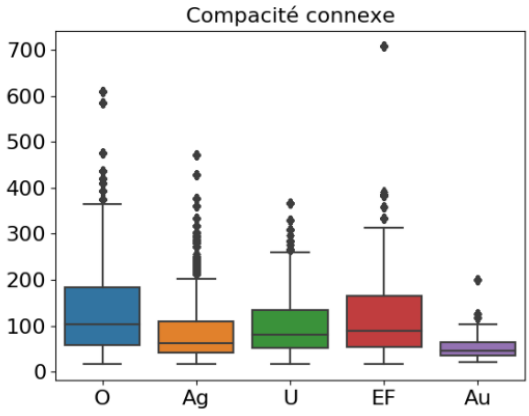


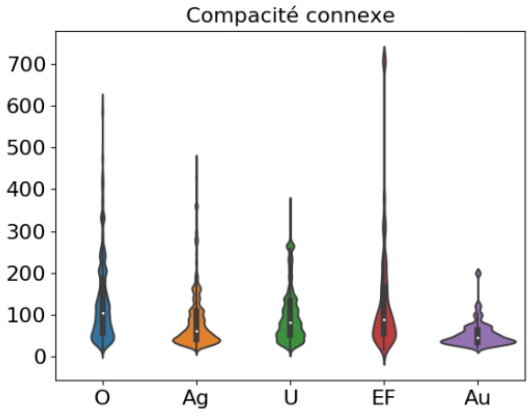


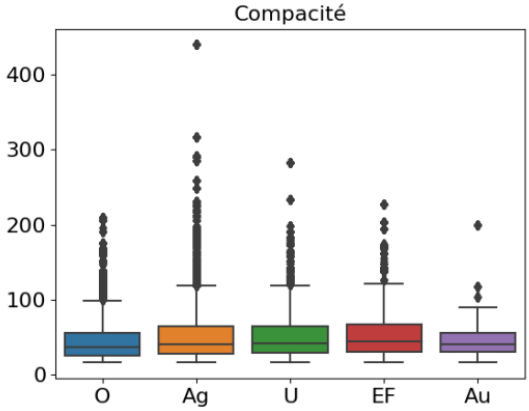


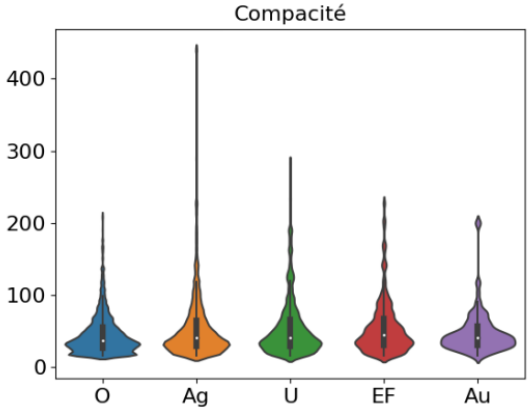


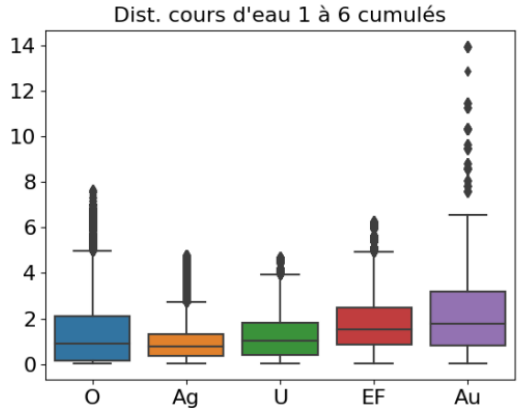


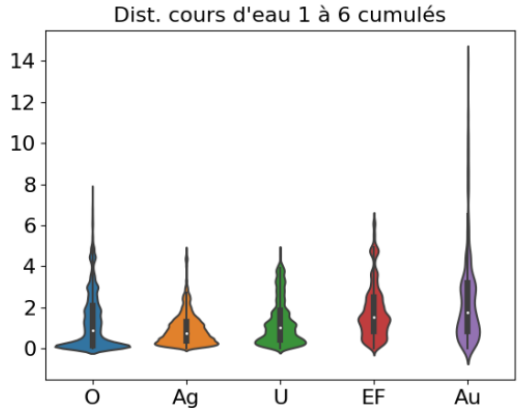

Compacité connexe
700
600
500
400
300
200
100
0
O
Ag
U
EF
Au

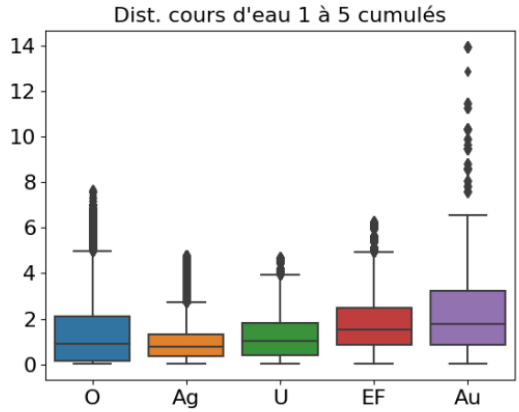

Compacité connexe
700
600
500
400
300
200
100
0
O
Ag
U
EF
Au

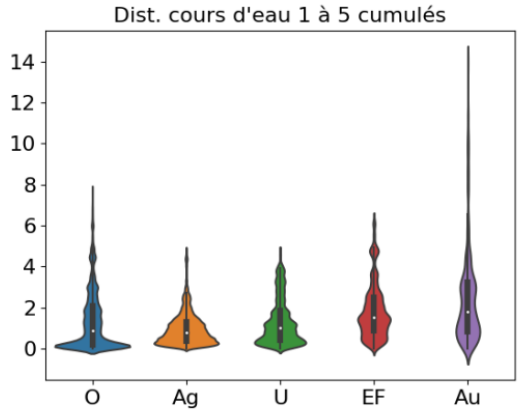

Compacité
400
300
200
100
0
O
Ag
U
EF
Au

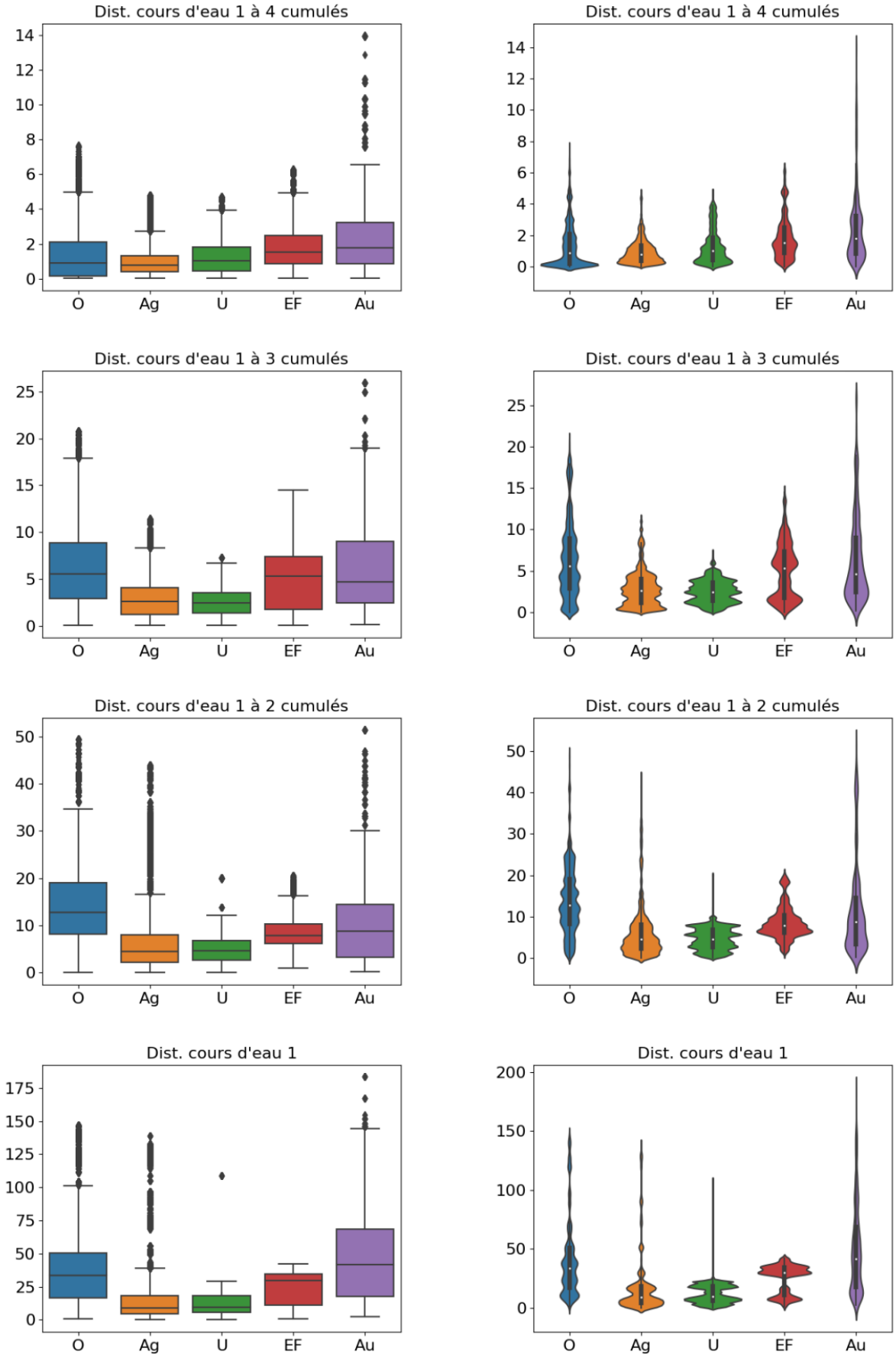

Compacité
400
300
200
100
0
O
Ag
U
EF
Au

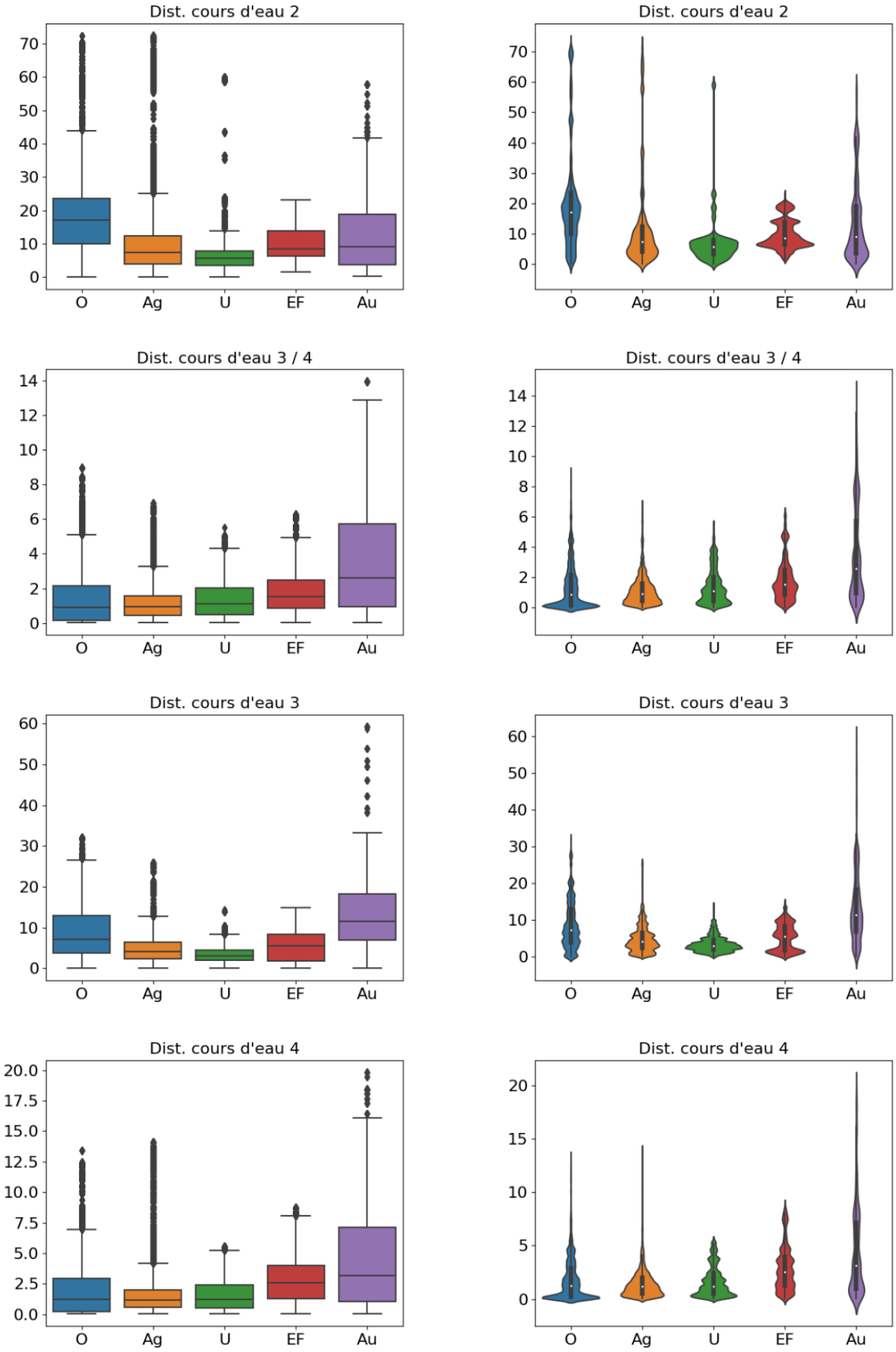

Dist. cours d'eau 1 à 6 cumulés
14
12
10
8
6
4
2
0
O
Ag
U
EF
Au

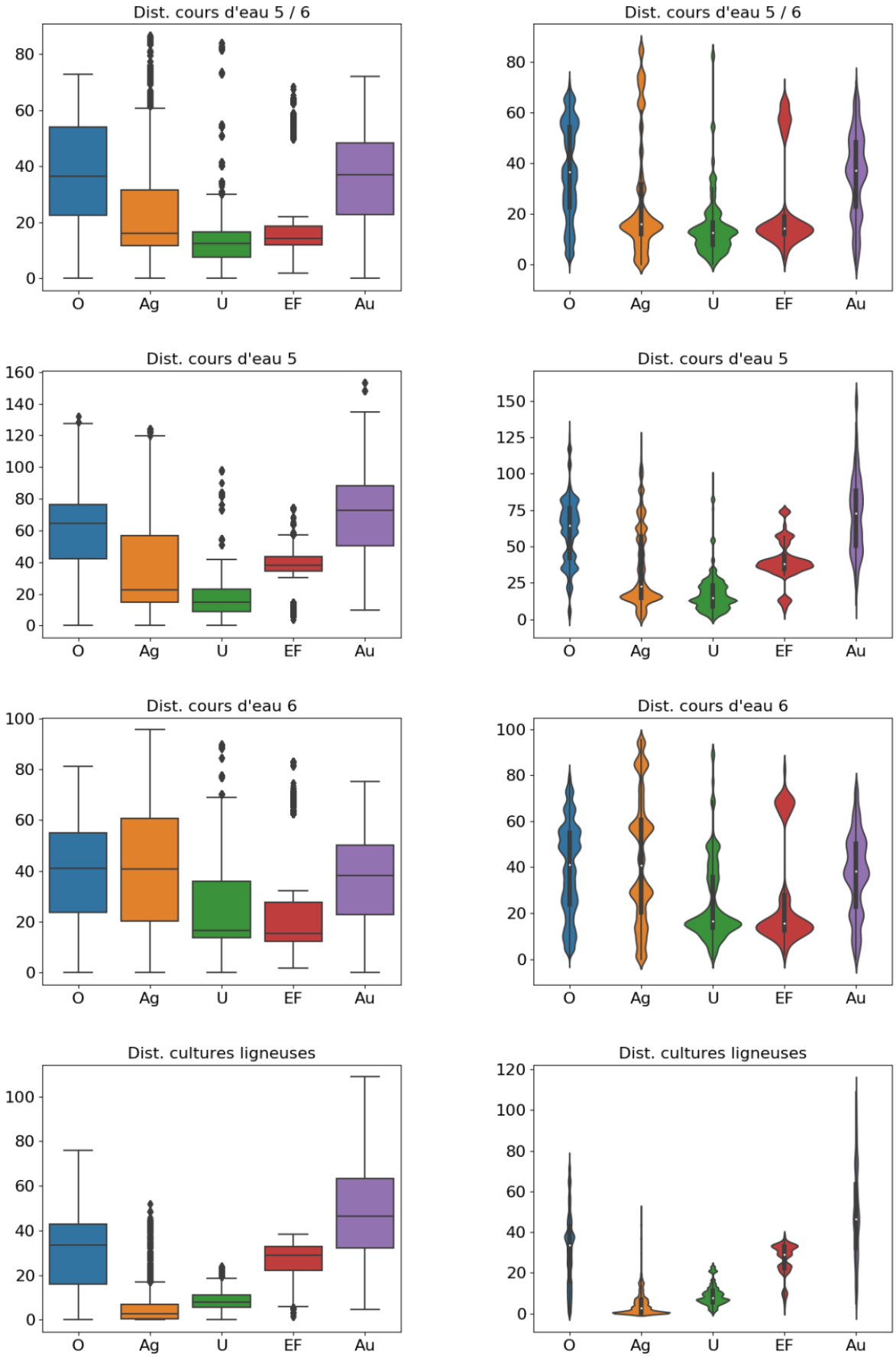

Dist. cours d'eau 1 à 6 cumulés
14
12
10
8
6
4
2
0
O
Ag
U
EF
Au

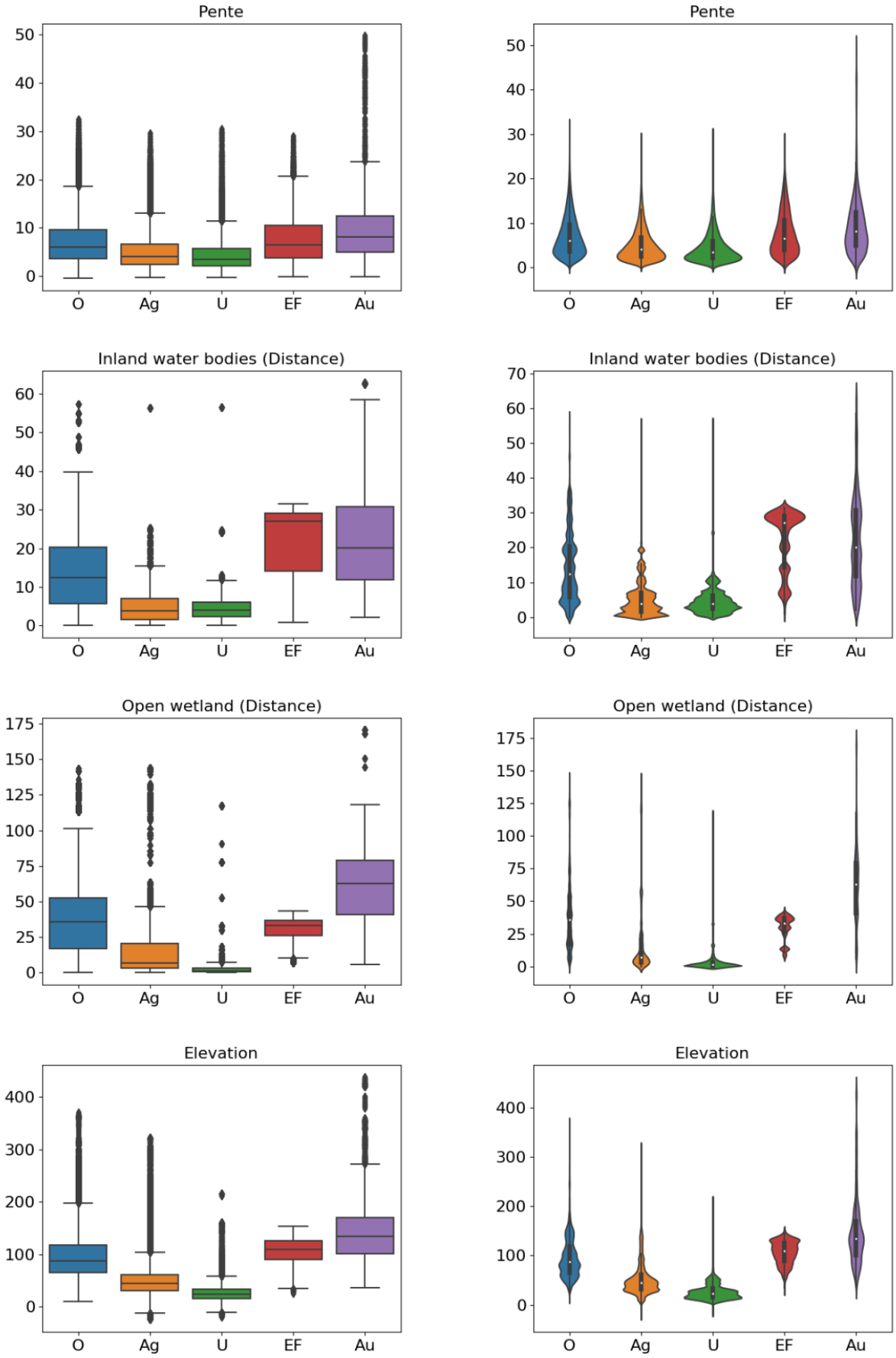

Dist. cours d'eau 1 à 5 cumulés
14
12
10
8
6
4
2
0
O
Ag
U
EF
Au

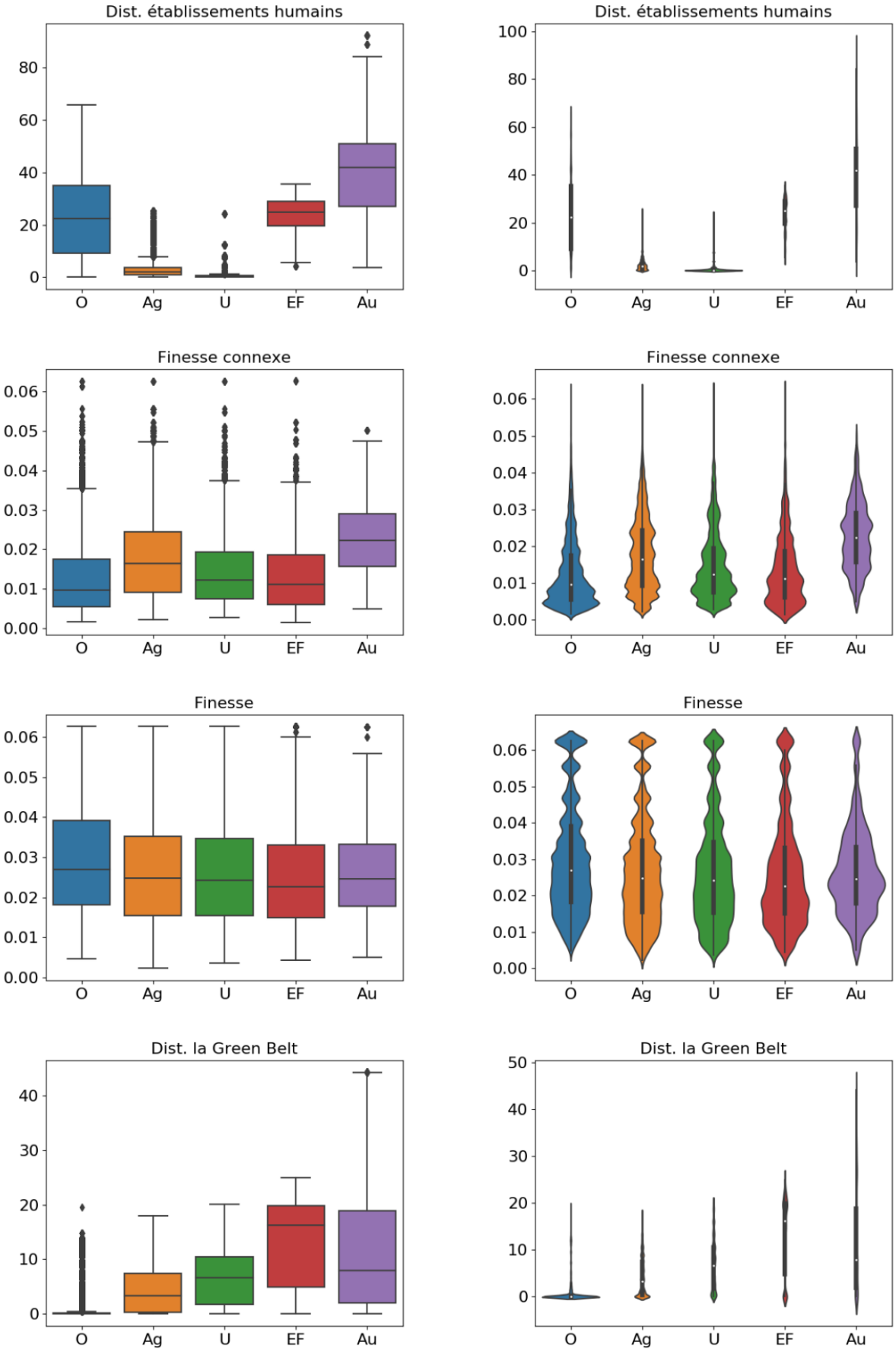

Dist. cours d'eau 1 à 5 cumulés
14
12
10
8
6
4
2
0
O
Ag
U
EF
Au

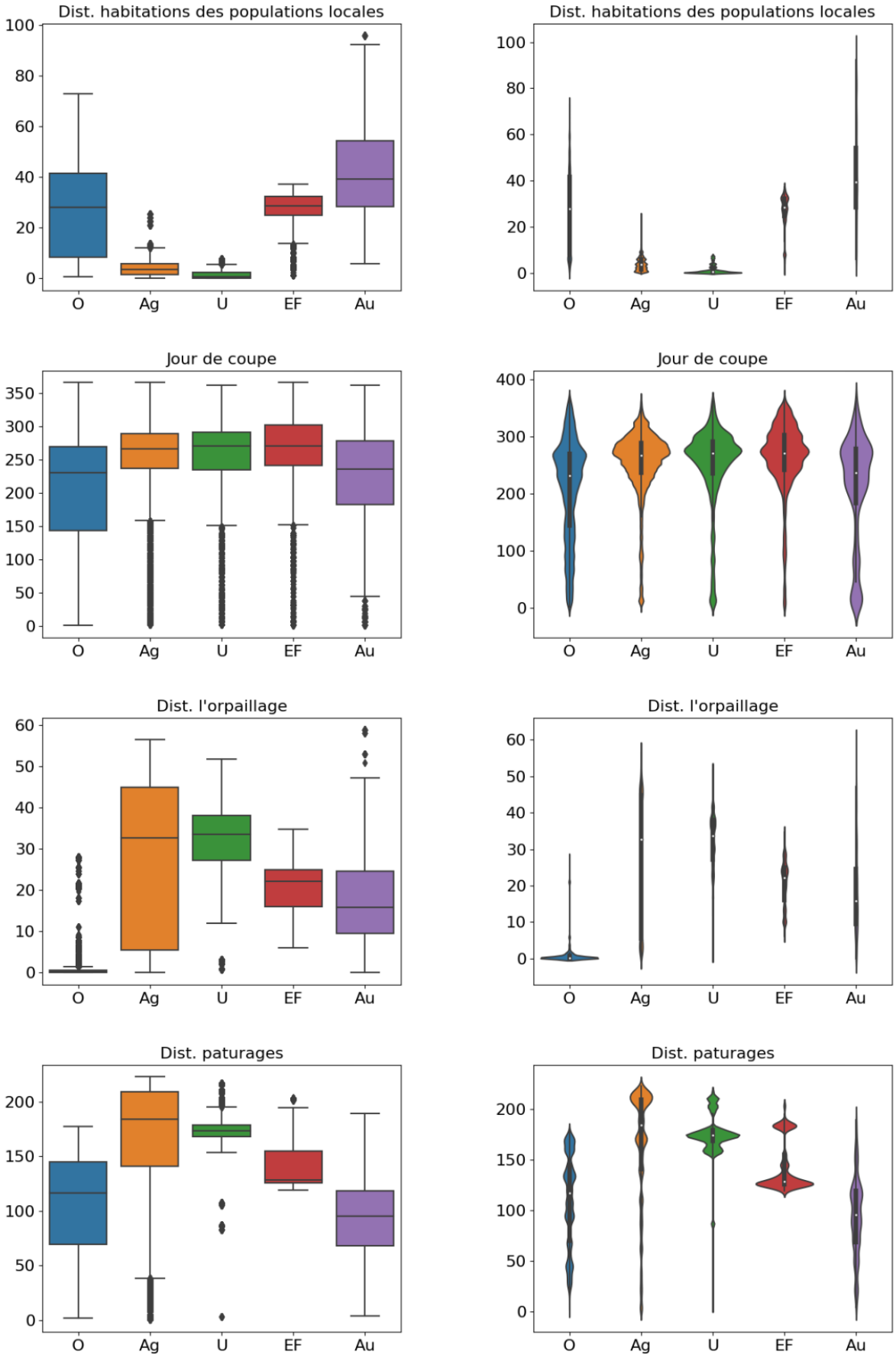

Dist. cours d'eau 1 à 4 cumulés
Dist. cours d'eau 1 à 4 cumulés
Dist. cours d'eau 1 à 3 cumulés
Dist. cours d'eau 1 à 3 cumulés
Dist. cours d'eau 1 à 2 cumulés
Dist. cours d'eau 1 à 2 cumulés
Dist. cours d'eau 1
Dist. cours d'eau 1
O
Ag
U
EF
Au

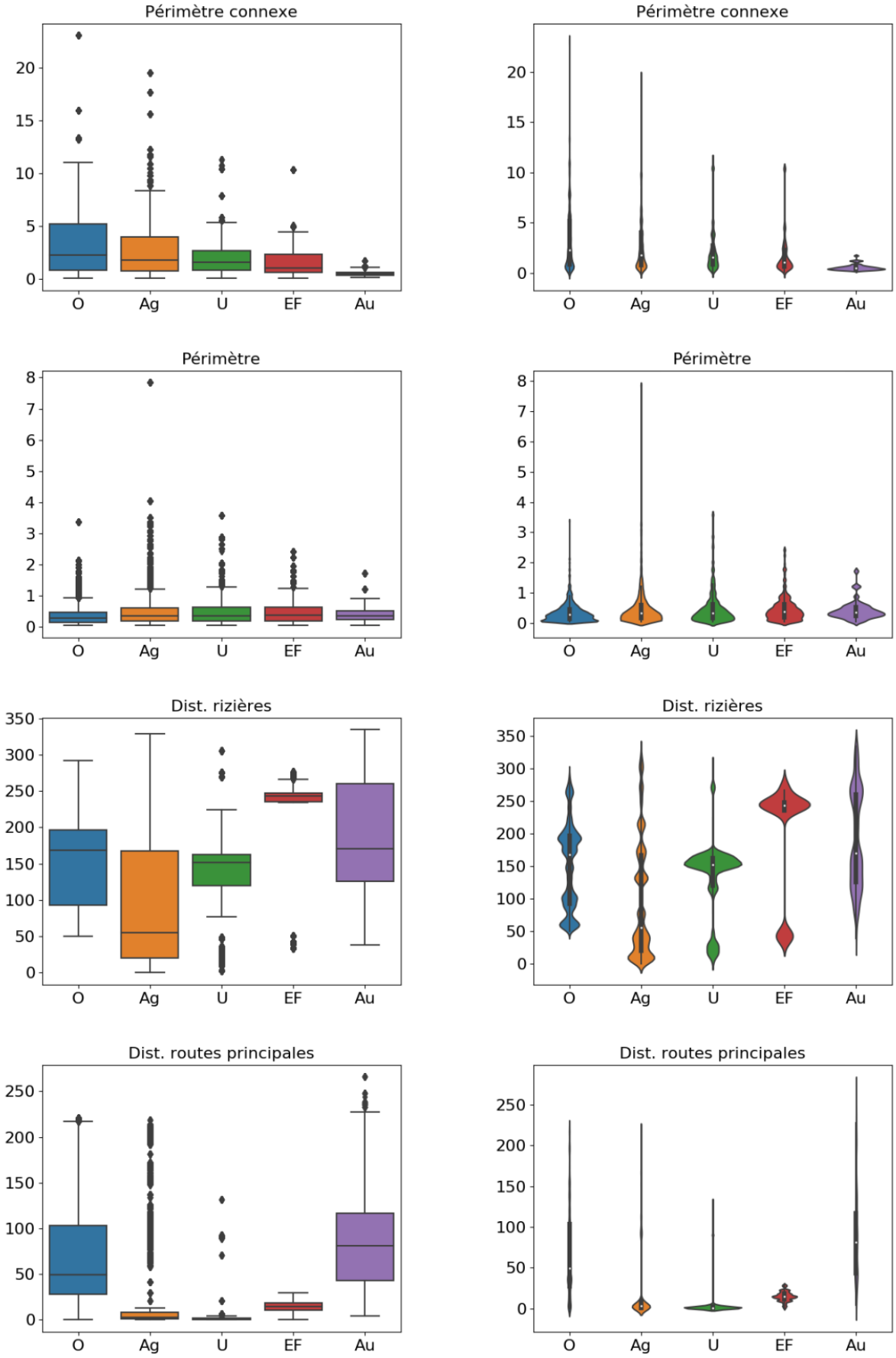

Dist. cours d'eau 2
Dist. cours d'eau 2
Dist. cours d'eau 3 / 4
Dist. cours d'eau 3 / 4
Dist. cours d'eau 3
Dist. cours d'eau 3
Dist. cours d'eau 4
Dist. cours d'eau 4
O
Ag
U
EF
Au

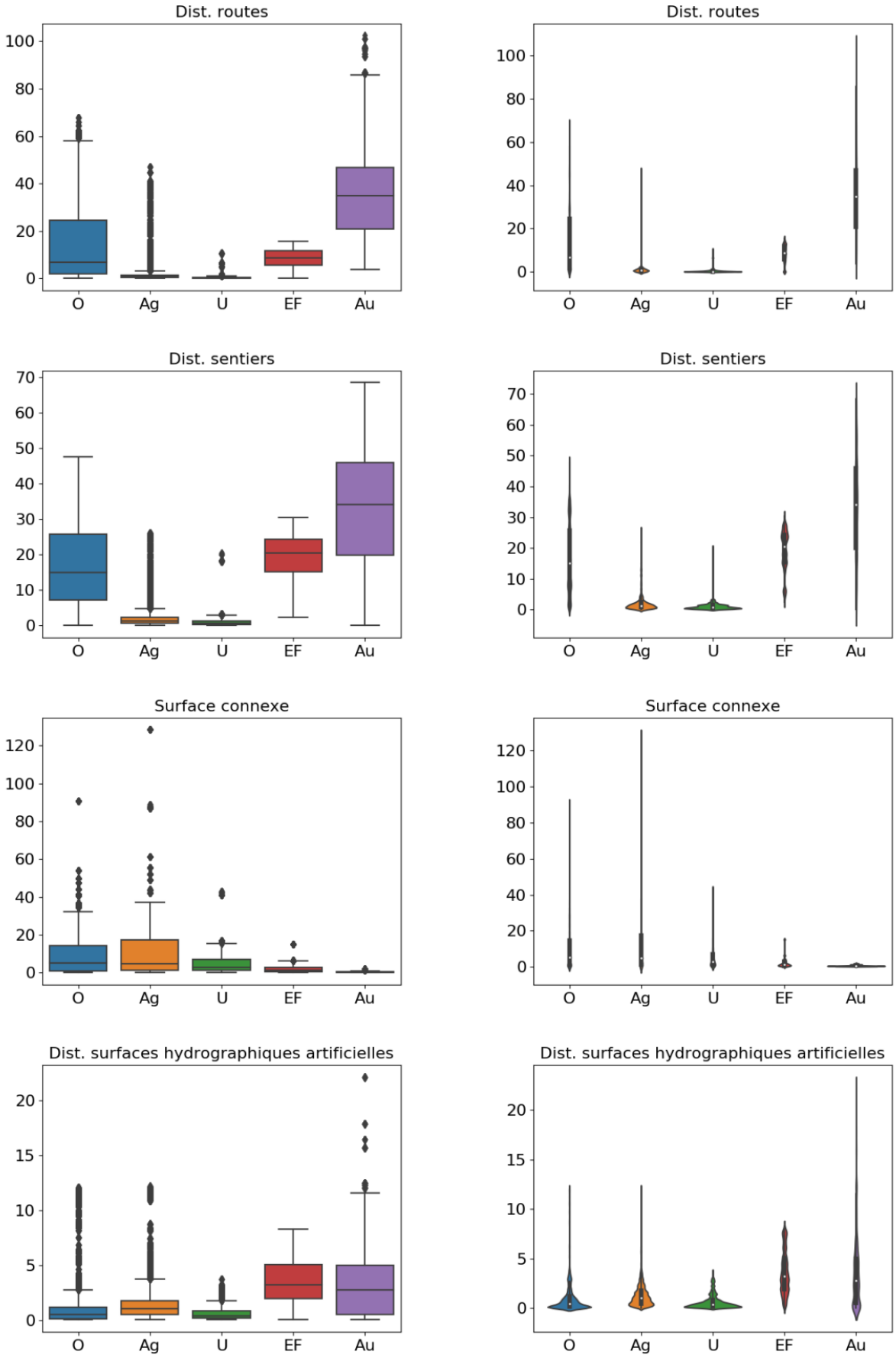

Dist. cours d'eau 5 / 6
Dist. cours d'eau 5 / 6
Dist. cours d'eau 5
Dist. cours d'eau 5
Dist. cours d'eau 6
Dist. cours d'eau 6
Dist. cultures ligneuses
Dist. cultures ligneuses
O
Ag
U
EF
Au

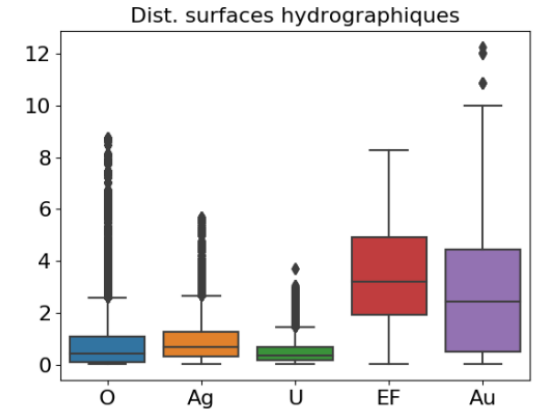

Pente
Pente
Inland water bodies (Distance)
Inland water bodies (Distance)
Open wetland (Distance)
Open wetland (Distance)
Elevation
Elevation
O
Ag
U
EF
Au

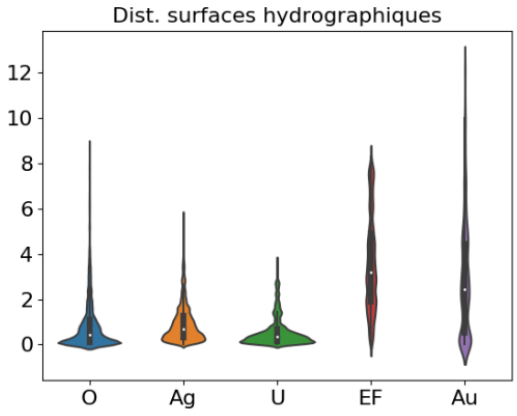

Dist. établissements humains
Dist. établissements humains
Finesse connexe
Finesse connexe
Finesse
Finesse
Dist. la Green Belt
Dist. la Green Belt
O
Ag
U
EF
Au

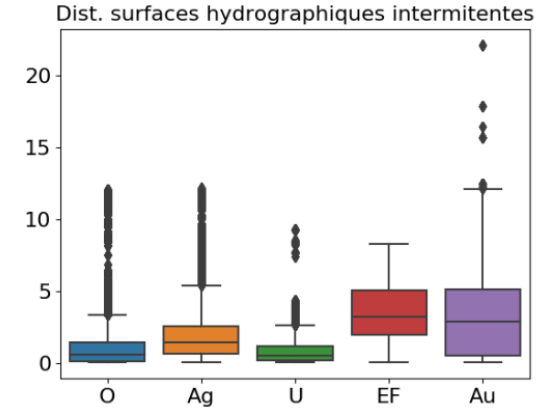

Dist. habitations des populations locales
Dist. habitations des populations locales
Jour de coupe
Jour de coupe
Dist. l'orpaillage
Dist. l'orpaillage
Dist. paturages
Dist. paturages
O
Ag
U
EF
Au

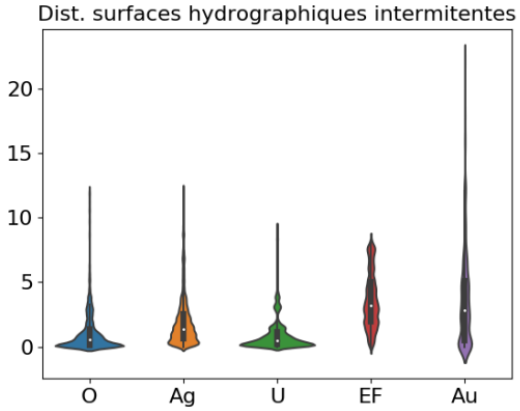

Périmètre connexe
Périmètre connexe
Périmètre
Périmètre
Dist. rizières
Dist. rizières
Dist. routes principales
Dist. routes principales
O
Ag
U
EF
Au

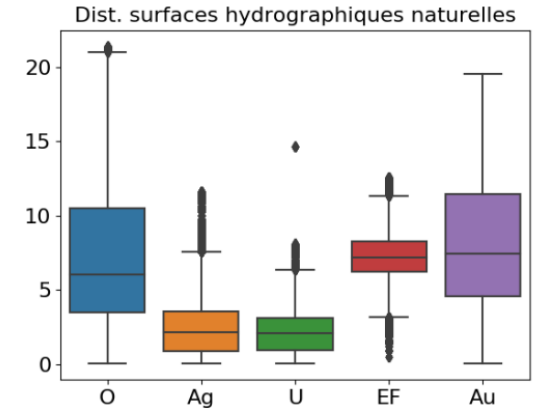

Dist. routes
Dist. routes
Dist. sentiers
Dist. sentiers
Surface connexe
Surface connexe
Dist. surfaces hydrographiques artificielles
Dist. surfaces hydrographiques artificielles
O
Ag
U
EF
Au

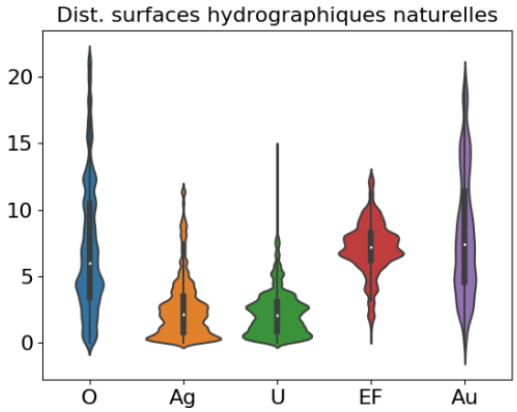

Dist. surfaces hydrographiques
12
10
8
6
4
2
0
O
Ag
U
EF
Au

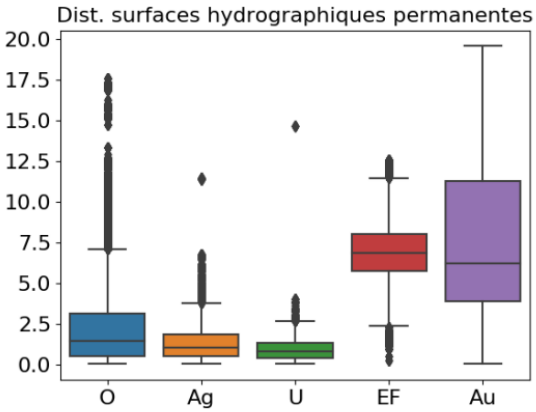

Dist. surfaces hydrographiques
12
10
8
6
4
2
0
O
Ag
U
EF
Au

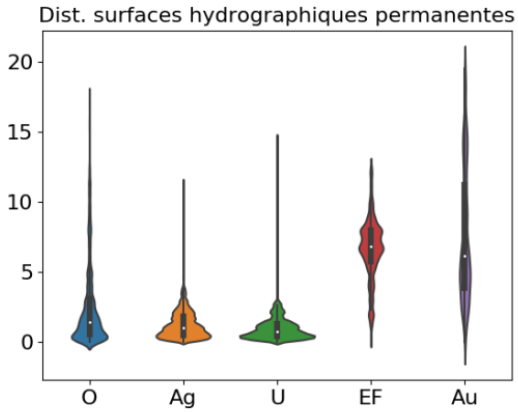

Dist. surfaces hydrographiques intermitentes
20
15
10
5
0
O
Ag
U
EF
Au

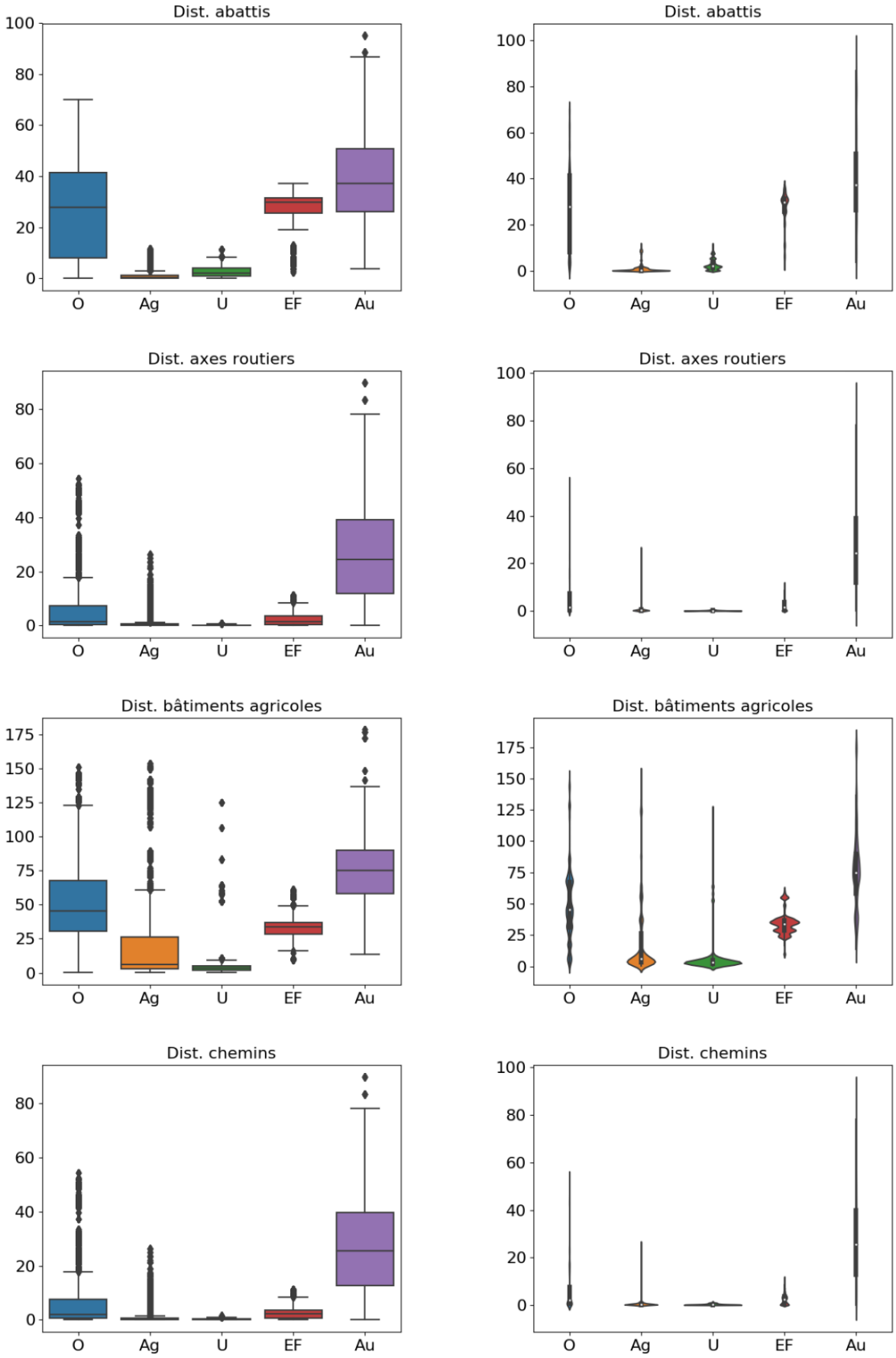

Dist. surfaces hydrographiques intermitentes
20
15
10
5
0
O
Ag
U
EF
Au


Dist. surfaces hydrographiques naturelles
20
15
10
5
0
O
Ag
U
EF
Au


Dist. surfaces hydrographiques naturelles
20
15
10
5
0
O
Ag
U
EF
Au


Dist. surfaces hydrographiques permanentes
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
O
Ag
U
EF
Au


Dist. surfaces hydrographiques permanentes
20
15
10
5
0
O
Ag
U
EF
Au


Dist. abattis
Dist. abattis
Dist. axes routiers
Dist. axes routiers
Dist. bâtiments agricoles
Dist. bâtiments agricoles
Dist. chemins
Dist. chemins
O
Ag
U
EF
Au

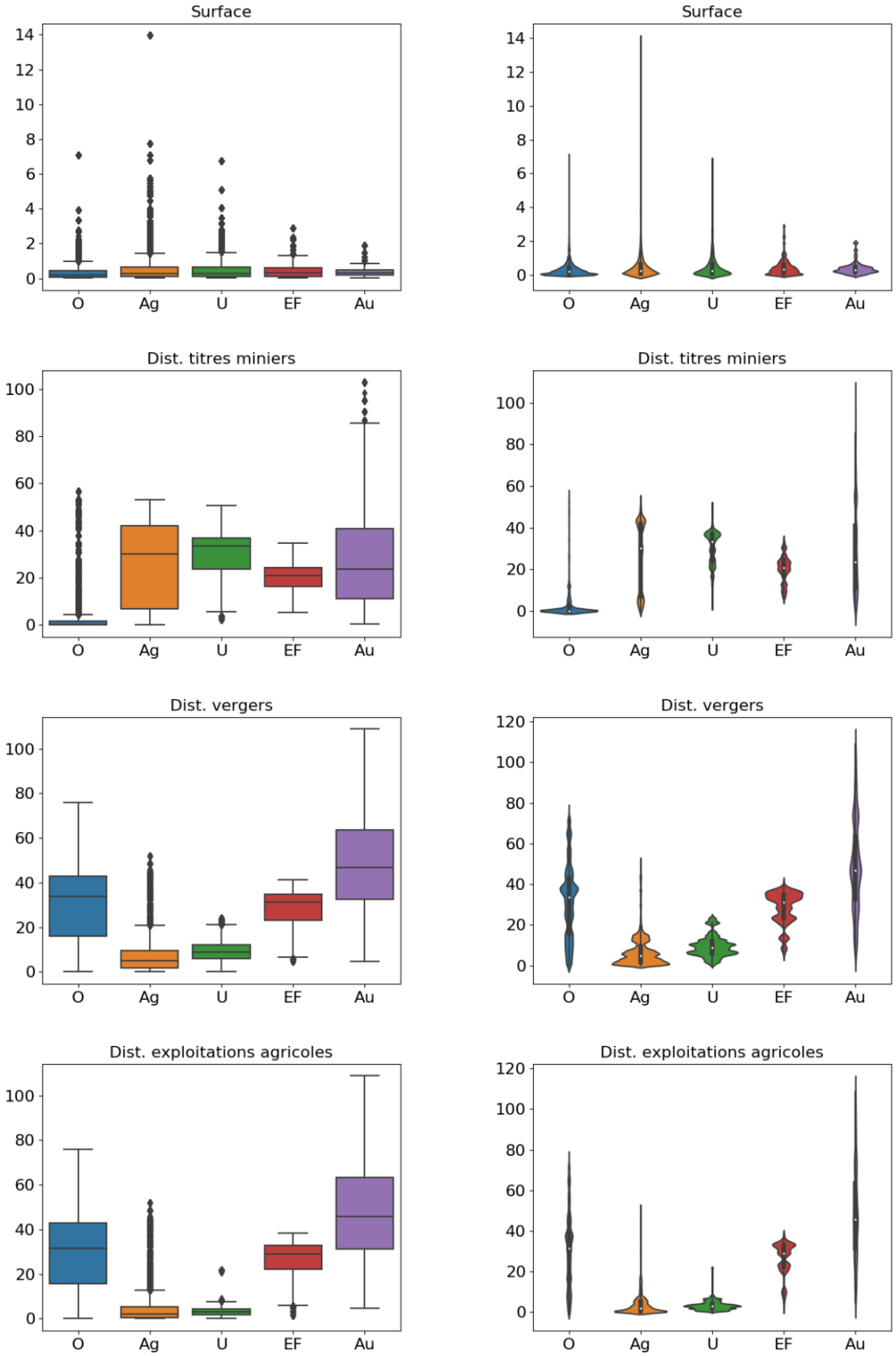


Figure 82 - Boîtes à moustache (gauche) et graphiques en violon (droite) des 51 indicateurs exploités par classe (O : Orpaillage ; Ag : Agriculture ; U : Urbanisation ; EF : Exploitation forestière ; Au : Autre). Les distances et les périmètres sont exprimés en km, l'élévation est donnée en mètres et les surfaces sont exprimées en ha.