
Cartographie de l'occupation du sol à très haute résolution via imagerie drone et satellitaire dans les montagnes du Burundi

Auteur : Stevens, Leo

Promoteur(s) : Lejeune, Philippe; Michez, Adrien

Faculté : Gembloux Agro-Bio Tech (GxABT)

Diplôme : Master en bioingénieur : sciences et technologies de l'environnement, à finalité spécialisée

Année académique : 2023-2024

URI/URL : <http://hdl.handle.net/2268.2/21560>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.



Cartographie de l'occupation du sol à très haute résolution via imagerie drone et satellitaire dans les montagnes du Burundi

LEO STEVENS

**TRAVAIL DE FIN D'ÉTUDES PRÉSENTÉ EN VUE DE L'OBTENTION DU DIPLÔME DE
MASTER EN SCIENCES ET TECHNOLOGIES DE L'ENVIRONNEMENT**

ANNÉE ACADÉMIQUE 2023-2024

CO-PROMOTEURS : PR. PHILIPPE LEJEUNE et DR. ADRIEN MICHEZ

MAITRE DE STAGE : DR. ALEXANDRE MAUGNARD

"Toute reproduction du présent document, par quelque procédé que ce soit, ne peut être réalisée qu'avec l'autorisation de l'auteur et de l'autorité académique de Gembloux Agro-Bio Tech".

"Le présent document n'engage que son auteur".



Cartographie de l'occupation du sol à très haute résolution via imagerie drone et satellitaire dans les montagnes du Burundi

LEO STEVENS

**TRAVAIL DE FIN D'ÉTUDES PRÉSENTÉ EN VUE DE L'OBTENTION DU DIPLÔME DE
MASTER EN SCIENCES ET TECHNOLOGIES DE L'ENVIRONNEMENT**

ANNÉE ACADÉMIQUE 2023-2024

CO-PROMOTEURS : PR. PHILIPPE LEJEUNE et DR. ADRIEN MICHEZ

MAITRE DE STAGE : DR. ALEXANDRE MAUGNARD

Résumé

L'érosion des sols et les inondations sont des problèmes majeurs au Burundi. D'importantes précipitations et un relief exacerbant l'érosion sont des facteurs contribuant à cela. Ils impactent des milliers de personnes, 90% de la population vivant de l'agriculture. C'est dans ce contexte que s'inscrit le Projet de Restauration et de Résilience du Paysage au Burundi financé par la Banque Mondiale. Celui-ci a pour but la restauration de la productivité des terres et l'amélioration de la gestion des eaux.

L'entreprise SHER intervient sur ce projet dans des bassins versants des communes de Isale et de Buhinyuza. Cette intervention se fait en partie par l'étude de l'occupation du sol. Ce mémoire s'inscrit dans cette étape et a pour but de proposer une nouvelle méthode de cartographie de l'occupation du sol. Pour ce faire, la télédétection via des images drones (DJI MAVIC 3M) et satellites (PlanetScope et Sentinel-2) est étudiée. Différents modèles de classification par RandomForest sont mis au point sur les différentes sources d'images. La précision de ces modèles au sein de différents bassins versants est étudiée. Afin d'évaluer la capacité des modèles à prédire l'occupation du sol de bassins n'ayant pas encore été étudiés, différentes combinaisons entre les sets d'entraînements de ces modèles sont mises au point.

La méthode mise au point permet une cartographie précise à environ 85% (images PlanetScope) et 65% (images Sentinel-2 et drone) au sein des bassins versants dont sont issues les données d'entraînement. Les prédictions effectuées sur d'autres bassins versants présentent des précisions de maximum 65%.

La méthode présente des avantages en terme de temps de travail et des sous-produits créés vis-à-vis de la méthode de cartographie actuellement utilisée par la SHER. L'utilisation d'images drones à très hautes résolution permet de minimiser la subjectivité lors de la photointerprétation.

Mots-clés : Drone, Satellite, Classification, RandomForest, Télédétection, Occupation du sol, Burundi, érosion.

Abstract

Soil erosion and flooding are major problems in Burundi. Heavy rainfall and a topography that exacerbates erosion contribute to these issues, impacting thousands of people, with 90% of the population relying on agriculture. This context frames the World Bank-funded Burundi Landscape Restoration and Resilience Project. The project's goals are to restore land productivity and improve water management.

The company SHER is involved in this project, focusing on the watersheds in the communes of Isale and Buhinyuza, partly through the study of land use. This thesis is part of that effort and aims to propose a new method for land use mapping. To achieve this, remote sensing using drone (DJI MAVIC 3M) and satellite (PlanetScope and Sentinel-2) images is studied. Various RandomForest classification models are developed using these different image sources, and the accuracy of these models within different watersheds is analyzed. To assess the models' ability to predict land use in unstudied watersheds, different combinations of training datasets are tested.

The developed method allows for precise mapping with approximately 85% accuracy (PlanetScope images) and 65% accuracy (Sentinel-2 and drone images) within the watersheds from which the training data originated. Predictions made for other watersheds show accuracies of up to 65%.

The method offers advantages in terms of labor time and the by-products created compared to the land mapping method currently used by SHER. The use of very high-resolution drone images minimizes subjectivity in photointerpretation.

Keywords : UAV, Satellite, Classification, RandomForest, Remote Sensing, Land Use, Burundi, Erosion.

Remerciements

En premier lieu, je tiens à remercier mes parents et mon frère pour le cadre de vie plein de rires et de joie dans lequel ils m'ont permis de grandir. Je vous dois mon amour de la nature et ma curiosité. Merci de me faire découvrir chaque jour mes qualités et mes défauts. Merci à mes grands-parents, oncles, tantes, cousins, cousines.

À Émilie, merci pour tout ces moments d'amour, de bonheur, de simplicité ou de faste culinaire qui m'ont permis d'envisager chaque jour sereinement.

Aux amis de l'Excentré, de l'Épicentre et autres perdus de Gembloux, merci pour ces années de fêtes, de blocus et de rigolade. J'espère vous revoir tous pour continuer toutes nos aventures. Courage aux étudiants, un jour ça s'arrête apparemment !

Asmani, Olivier, Mireille, Valéry, Dieudonné, Audrey, Théodomir, ... merci à vous tous pour ce mois magnifique passé au Burundi et l'accueil dont vous avez fait preuve. Vous êtes des personnes formidables, je ne pouvais pas rêver de mieux en partant vers l'inconnu. Je me réjouis déjà de vous revoir et de continuer à découvrir ce magnifique pays avec vous !

Merci à mes promoteurs Philippe Lejeune et Adrien Michez ainsi qu'à mon maître de stage Alexandre Maignard pour la patience, la présence et les conseils avisés. Plus globalement, merci au personnel de la SHER pour son accueil durant mon stage et ce mémoire.

Table des matières

Table des figures	iii
Liste des tableaux	iv
Nomenclature	vii
1 Introduction	1
1.1 Cadre physique du mémoire	1
1.1.1 Problématique de l'érosion des sols au Burundi	1
1.1.2 Les inondations au Burundi	2
1.1.3 L'occupation des sols et l'eau	2
1.1.4 La télédétection comme vecteur d'identification de l'occupation des sols	4
1.2 Cadre du mémoire en entreprise	6
2 Matériel et méthodes	8
2.1 Description des zones d'étude	8
2.2 Chaîne de traitements	10
2.2.1 Images satellites	13
2.2.2 Images drone	14
2.2.3 Photointerprétation	16
2.2.4 Les indices de végétation	20
2.2.5 Les textures	21
2.2.6 La classification par RandomForest	22
2.2.7 Évaluation des modèles	25
2.3 Méthode de cartographie de l'entreprise SHER	26
3 Résultats	27
3.1 Qualité des classifications de l'occupation du sol	27
3.1.1 Évaluation des modèles par bassin versant	27
3.1.2 Évaluation des modèles entraînés sur plusieurs bassins versants	38
3.2 Méthode de classification mise au point	44
4 Discussion	47
4.1 Le modèle à l'échelle du bassin versant	47
4.1.1 Le cas du bassin NYAM_01	49
4.1.2 Sentinel	49
4.1.3 Drone	50

4.2	Les modèles à l'échelle de la région	50
4.3	Les images PlanetScope	56
4.4	Comparaison des méthodes de classification	58
4.4.1	Méthode de cartographie de l'entreprise SHER	58
4.4.2	Comparaison avec le WorldCover de l'ESA	58
4.5	Sources d'images multitemporelles	60
5	Conclusion	62
	Bibliographie	63
6	Annexe	67
6.1	Annexe 1 : Manuel d'utilisation du drone Mavic 3M	67

Table des figures

1	Illustration de la technique de terrasse progressive.	3
2	Exemples de drones. De gauche à droite : drone multirotor DJI Mavic 3M, drone à aile fixe léger, drone à aile fixe lourd.	5
3	Localisation des communes de la zone d'étude à l'échelle nationale.	8
4	Localisation des bassins versants étudiés sur la commune de Isale (Burundi).	9
5	Localisation des bassins versants étudiés sur la commune de Buhinyuza (Burundi).	10
6	Flowchart de la chaine de traitements mise en place.	12
7	Traitements des images drones.	16
8	Illustration des différentes classes d'occupation de sol.	17
9	Photointerprétation sur le bassin versant GIKO_02 (commue de Isale, Burundi).	18
10	Présentation de la structure du code réalisé pour la classification.	23
11	Graphique de l'importance des variables dans le modèle entraîné sur KINY_03 (images PlanetScope).	30
12	Graphique de l'importance des variables dans le modèle entraîné sur NYAM_01 (images PlanetScope).	30
13	Carte d'occupation du sol du bassin versant Giko_02 générée grâce au modèle entraîné sur des images Planet du même bassin.	31
14	Carte d'occupation du sol du bassin versant Nyam_01 générée grâce au modèle entraîné sur des images Planet du même bassin.	32
15	Graphique de l'importance des variables du modèle entraîné sur des images Sentinel-2 sur le bassin versant GIKO_02.	34
16	Graphique de l'importance des variables du modèle entraîné sur des images Sentinel-2 sur le bassin versant MUHU_03.	35
17	Graphique de l'importance des variables du modèle entraîné sur des images Sentinel-2 sur le bassin versant NTAH_05.	35
18	Carte d'occupation du sol du bassin versant GIKO_02 générée grâce au modèle entraîné sur des images Sentinel du même bassin.	36
19	Carte d'occupation du sol du bassin versant NTAH_05 générée grâce au modèle entraîné sur des images Sentinel du même bassin.	37
20	Graphique des pentes par classe d'occupation de sol au sein de la commune de Buhinyuza.	48
21	Boxplot de l'indice NDVI pour chaque classe d'occupation de sol au sein des bassins versants de la commune de Isale (Image Planet).	51
22	Boxplot de l'indice NDVI pour chaque classe d'occupation de sol au sein des bassins versants de la commune de Isale (Image Sentinel).	52

23	Boxplot de la bande Red pour chaque classe d'occupation de sol au sein des bassins versants de la commune de Isale (Image Planet).	52
24	Boxplot de la bande Red pour chaque classe d'occupation de sol au sein des bassins versants de la commune de Isale (Image Sentinel).	53
25	Boxplots des pentes en fonction du bassin versant pour chaque occupation de sol sur la commune de Isale.	54
26	Boxplots de l'exposition en fonction du bassin versant pour chaque occupation de sol sur la commune de Isale.	55
27	Boxplots des pentes en fonction du bassin versant pour chaque occupation de sol sur la commune de Buhinyuza.	55
28	Boxplots de l'exposition en fonction du bassin versant pour chaque occupation de sol sur la commune de Buhinyuza.	56

Liste des tableaux

1	Caractéristiques topographiques des bassins versants.	10
2	Comparaison des bandes spectrales et des résolutions des satellites Sentinel-2 et PlanetScope	13
3	Résolution en fonction de l'altitude.	15
4	Surface et nombre de pixels de chaque classe pour GIKO_02	19
5	Surface et nombre de pixels de chaque classe pour MUHU_03	19
6	Surface et nombre de pixels de chaque classe pour NTAH_05	19
7	Surface et nombre de pixels de chaque classe pour KINY_03 et NYAM_01	20
8	Combinaison des couches de polygones dans les différents modèles composés.	24
9	Matrice de confusion binaire théorique	25
10	Précision, kappa et nombre de pixels par classe dans le set d'entraînement d'images Planet des modèles par bassin versant.	27
11	Matrice de confusion du modèle entraîné sur le bassin NTAH_05 (images PlanetScope).	28
12	Tableau statistique du modèle entraîné sur le bassin NTAH_05 (images PlanetScope).	28
13	Matrice de confusion du modèle entraîné sur le bassin NYAM_01 (images PlanetScope).	28
14	Tableau statistique du modèle entraîné sur le bassin NYAM_01 (images PlanetScope).	29
15	Matrice de confusion du modèle entraîné sur le bassin KINY_03 (images PlanetScope).	29

16	Tableau statistique du modèle entraîné sur le bassin KINY_03 (images PlanetScope).	29
17	Performances des modèles appliqués à d'autres bassins versants (images PlanetScope).	31
18	Précision, kappa et nombre de pixels par classe dans le set d'entraînement d'images Sentinel-2 des modèles par bassin versant.	32
19	Tableau statistique du modèle entraîné sur le bassin GIKO_02 (images Sentinel-2).	33
20	Matrice de confusion du modèle entraîné sur le bassin MUHU_03 (images Sentinel-2).	33
21	Tableau statistique du modèle entraîné sur le bassin MUHU_03 (images Sentinel-2).	33
22	Tableau statistique du modèle entraîné sur le bassin NTAH_05 (images Sentinel-2).	34
23	Performances des modèles appliqués à d'autres bassins versants (images Sentinel-2).	36
24	Résumé statistique du modèle entraîné sur le bassin versant GIKO_02 (images drones)	37
25	Résumé statistique de l'entraînement et de l'évaluation des modèles (images PlanetScope).	38
26	Tableau statistique obtenu à l'issue de l'entraînement du modèle 1 (images PlanetScope).	38
27	Matrice de confusion obtenue à l'issue de l'évaluation du modèle 1 (images PlanetScope).	39
28	Tableau statistique obtenu à l'issue de l'évaluation du modèle 1 (images PlanetScope).	39
29	Tableau statistique obtenu à l'issue de l'entraînement du modèle 4 (images PlanetScope).	40
30	Matrice de confusion obtenue à l'issue de l'évaluation du modèle 4 (images PlanetScope).	40
31	Tableau statistique obtenu à l'issue de l'évaluation du modèle 4 (images PlanetScope).	40
32	Résumé statistique de l'entraînement et de l'évaluation des modèles (images Sentinel-2).	41
33	Tableau statistique obtenu à l'issue de l'entraînement du modèle 1 (images Sentinel-2).	41
34	Tableau statistique obtenu à l'issue de l'évaluation du modèle 1 (images Sentinel-2).	42

35	Tableau statistique obtenu à l'issue de l'entraînement du modèle 2 (images Sentinel-2).	42
36	Matrice de confusion obtenue à l'issue de l'évaluation du modèle 2 (images Sentinel-2).	43
37	Tableau statistique obtenu à l'issue de l'évaluation du modèle 2 (images Sentinel-2).	43
38	Tableau statistique obtenu à l'issue de l'entraînement du modèle 3 (images Sentinel-2).	43
39	Tableau statistique obtenu à l'issue de l'évaluation du modèle 3 (images Sentinel-2).	44
40	Tableau statistique obtenu à l'issue de l'entraînement du modèle 4 (images Sentinel-2).	44
41	Nombre de pixels des classes identifiées par l'ESA présent dans les polygones de photointerprétation de chaque classe.	59

Nomenclature

FAO Food and Agriculture Organization of the United Nations

GIRE Gestion Intégrée des Ressources en Eau

GLCM Gray-Level Co-occurrence Matrix

GNSS Global Navigation Satellite System

Go Giga octet

Ha Hectare

MNS Modèle Numérique de Surface

MNT Modèle Numérique de Terrain

MS MultiSpectral

NDVI Normalized Difference Vegetation Index

NDWI Normalized Difference Water Index

NIR Near infrared

PRRPB Projet de Restauration et de Résilience du Paysage du Burundi

RF RandomForest

RGB Red, Green, Blue

SHER Société pour l'Hydraulique et l'Environnement et la Réhabilitation

SRTM Shuttle Radar Topographic Mission

TVI Transformed Vegetation Index

USLE Équation Universelle de Perte en Terre (Universal Soil Loss Equation)

VTOL Vertical Take Off and Landing

1 Introduction

1.1 Cadre physique du mémoire

1.1.1 Problématique de l'érosion des sols au Burundi

L'érosion des sols est à la base de différents problèmes dans notre société. Pour certains, elle représente la plus grosse menace pesant sur les sols ([Montanarella et al., 2016](#)) et selon la FAO, l'érosion hydrique et éolienne est la première des cinq causes de détérioration des sols ([FAO, 2015](#)). Longtemps ignorés par le monde scientifique, les sols remplissent un nombre important de services écosystémiques comme par exemple la production de biomasse (forestière et agricole entre autres), le stockage du carbone ou encore en stockant, filtrant et transformant des nutriments et de l'eau. Il est donc nécessaire de se prémunir des effets de leur détérioration ainsi que de limiter leur érosion ([FAO, 2015](#)).

Au Burundi, la lutte contre l'érosion des sols est menée depuis des dizaines d'années. L'altitude du pays varie entre 774 m et 2 670 m, certaines régions du pays présentent des reliefs importants dus à la présence de collines. Le climat est constitué de deux saisons des pluies et deux saisons sèches en alternance et les précipitations peuvent atteindre 2 000 mm/an dans certains endroits. Ces facteurs sont la source de phénomènes d'érosion hydrique observables sur certaines pentes ([République du Burundi, 2020](#)). Dès les années 1940, un programme de lutte contre le ruissellement est mis en place dans la région de Muramvya ([Bidou, 1990](#)). Un manque de connaissances théoriques à la base des réflexions engagées est cependant remarqué. Cela mènera au développement, dans les années 90, de diverses études visant à quantifier l'érosion mais surtout à déterminer l'origine de ce processus ainsi que les caractéristiques culturelles, anthropologiques, pédologiques et/ou climatologiques y étant liées ([Rishirumu-hirwa, 1994](#); [Bidou, 1990](#)).

À l'heure actuelle, le problème de l'érosion des sols au Burundi est toujours bien présent. À l'Est du pays, une perte par érosion de 4 tonnes (T) par hectare (ha) et par an est estimée. À l'Ouest, dans les collines dominant Bujumbura, cette perte est estimée à 100 T/ha et par an et peut atteindre les 150 T/ha et par an. Cette différence dans les quantités de pertes en sol s'explique par la variété des pentes rencontrées à l'échelle du pays, mais également par la diversité des types de sols et de leurs caractéristiques hydrologiques. Par exemple, les zones d'étude sont couvertes principalement par des Ferralsols, des Nitisols et des Vertisols. Ces sols sont caractérisés par des capacités de drainage, de rétention et des coefficients d'érodibilité différents ([ISRIC, 2024](#)).

Selon la FAO, 90% de la population burundaise est toujours active dans le domaine agri-

cole (FAO, 2005). L'érosion hydrique, en déplaçant les sols et en lessivant les nutriments s'y trouvant, a des impacts négatifs sur la productivité des terres agricoles (FAO, 2015). La population locale dépend donc de la capacité à préserver ces terres.

1.1.2 Les inondations au Burundi

Il a été vu au paragraphe précédent que les précipitations peuvent atteindre des quantités importantes au Burundi. En plus du phénomène d'érosion qui en découle, ces précipitations causent dans le pays de graves inondations. Avec le changement climatique, les précipitations sont amenées à augmenter dans cette région du monde. Par exemple, d'Octobre 2023 à Mars 2024, d'importantes inondations ont pu être observées en Afrique de l'Est, dues en partie au phénomène El Niño. Au Burundi, sur cette période, ce sont 184 000 personnes qui ont été impactées. Environ 30 000 d'entre elles se sont déplacées et 10% des cultures vivrières du pays ont été touchées (OCHA, 2024). Sur place, il est possible d'observer les effets de ces précipitations. Le lac Tanganyika, qui couvre la frontière Ouest sur plus de 120 kilomètres, voit son niveau augmenter depuis maintenant quelques années. À Bujumbura, capitale du pays située sur les côtes du lac, des quartiers se retrouvent sous eau, leurs habitants étant obligés de déménager et de se reloger, si ils en ont les moyens.

1.1.3 L'occupation des sols et l'eau

En plus de divers facteurs tels que la pente du terrain, le potentiel érosif d'une averse ou encore la localisation du terrain sur la pente, l'occupation du sol et la couverture de celui-ci influencent la propension d'un sol à l'érosion hydrique (FAO, 2015). L'équation universelle de perte en terre (USLE) identifie l'occupation des sols comme un facteur permettant de quantifier cette perte. L'équation est la suivante (Stone, 2000) :

$$A = R \times K \times LS \times C \times P$$

Où :

- A représente les pertes en terre annuelles ;
- R correspond au facteur de pluie et de ruissellement ;
- K représente le facteur d'érodibilité du sol ;
- LS représente le facteur de longueur et d'inclinaison de la pente ;
- C représente le facteur de culture ;
- P correspond au facteur de pratique de conservation.

Le facteur de culture permet de quantifier l'effet des modalités de gestion du sol sur l'érosion. Par exemple, une culture de maïs aura un facteur d'une valeur de 0,40 tandis qu'une

plantation d'arbres fruitiers aura une valeur égale à 0,10 (Stone, 2000).

Ce sont ces différents facteurs que la Société pour l'Hydraulique et l'Environnement et la Réhabilitation (SHER) tente de modifier en proposant des aménagements anti-érosion. Par exemple, un type d'aménagement réalisé dans le cadre du Projet de Restauration et de Résilience du Paysage du Burundi (PRRPB) est le terrassement progressif illustré en Figure 1. La technique consiste à planter des haies de plantes herbacées le long des courbes de niveau, à des distances dépendant de la pente et à creuser sous ces haies des fossés. La haie a comme effet de réguler la vitesse de ruissellement et de favoriser le dépôt des terres érodées sur la partie supérieure de la pente à ses pieds. Au fil du temps, la pente du terrain et sa longueur sont modifiées. Ces modifications ont pour effet d'augmenter l'infiltration de l'eau dans le sol et de faciliter le travail de la terre par la même occasion (Moeyersons, 1994).

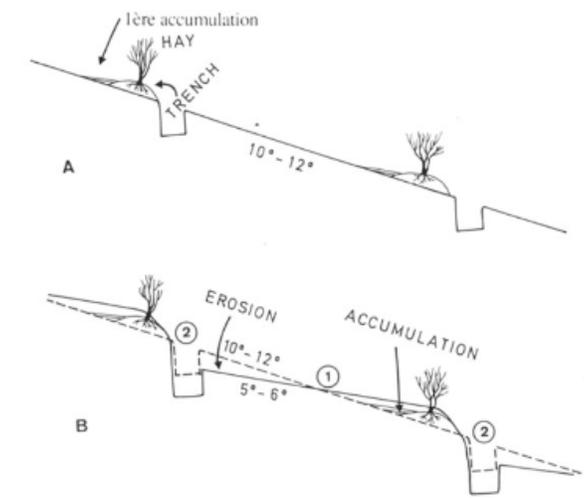


FIGURE 1 – Illustration de la technique de terrasse progressive (Moeyersons, 1994).

Il existe une grande variété d'aménagements destinés à la lutte antiérosive. Tous ne répondent cependant pas aux mêmes besoins ni à la même problématique. Certains conviennent à des terres agricoles, d'autres à des sols forestiers. Certains encore sont à privilégier sur des sols aux pentes faibles, à l'inverse d'autres. L'impact des aménagements sur l'écoulement des eaux varie également en fonction de ceux-ci (GISER, 2016). Il est donc important de caractériser le paysage avant d'envisager toute prise de décision.

Une partie de cette caractérisation passe par la détermination de l'occupation de sol. En effet, nous avons vu que celle-ci a un effet sur l'érosion en elle-même mais aussi sur la manière de lutter contre celle-ci. Parallèlement, l'occupation du sol a un impact sur les inondations. Les différents types de sol ainsi que les pratiques agricoles y étant appliquées modifient le ruissellement et l'infiltration. Des aménagements adaptant ces facteurs permettent donc de

minimiser l'impact et l'ampleur des inondations.

1.1.4 La télédétection comme vecteur d'identification de l'occupation des sols

L'utilisation d'images de télédétection à très haute et moyenne résolution spatiale se démocratise au fur et à mesure des années. Le terme de « très haute résolution spatiale » fait référence à des images dont la résolution varie de quelques centimètres à environ 1 mètre par pixel. Les images à moyenne résolution voient elles leur résolution varier entre 1 et 100 mètres par pixel. Cette démocratisation est due à l'accessibilité croissante des ressources à des prix plus abordables, à l'amélioration des capacités et de la qualité d'acquisition ainsi qu'au développement de nouveaux outils permettant de traiter les quantités croissantes de données acquises (Al-Najjar et al., 2019; Chen et al., 2020; Jiménez-Jiménez et al., 2021; Mountrakis & Heydari, 2023; Sefercik et al., 2021; Semeki Ngabinzeke et al., 2018; Toro & Tsourdos, 2018).

Diverses études ont démontré la pertinence de l'utilisation d'images à très haute résolution acquises par drones et satellites dans le cas de la cartographie de l'occupation du sol. Différentes méthodes de classification ont également été mises au point et leurs performances ont été étudiées (Rodriguez-Galiano et al., 2012; Al-Najjar et al., 2019; Mountrakis & Heydari, 2023). L'utilisation d'une source d'image au détriment d'une autre se justifie entre autre par la résolution attendue de l'étude, la possibilité d'effectuer des acquisitions via un opérateur et les fonds disponibles alloués à cette acquisition. À titre d'exemple, le projet WorldCover, initié par l'Agence Spatiale Européenne (ESA), a permis la création d'une carte d'occupation et de couverture du sol d'ampleur mondiale, reprenant onze classes d'occupation, pour une résolution de 10 mètres et une précision globale de 76,7%. Les images servant à l'élaboration de cette carte proviennent des missions Sentinel-1 et Sentinel-2 (Van De Kerchove, 2022).

1.1.4.1. Les différents programmes d'acquisition par satellite

Différents programmes d'acquisition d'images par satellite existent. Ceux-ci se différencient par la résolution des données acquises, le coût des données, la fréquence des acquisitions et le type de données acquises. Deux sources d'images satellites ont été étudiées dans le présent mémoire :

- La mission Sentinel-2, mise en place dans le cadre du programme de l'Union Européenne d'observation et de surveillance de la Terre, Copernicus. Deux satellites, Sentinel-2A et Sentinel-2B, permettent l'acquisition tous les 5 jours d'images d'une emprise au sol de 290 km pour une résolution variant de 10 à 60 m en fonction des bandes spectrales considérées. Ce sont 13 bandes spectrales qui sont acquises par les

satellites. L'orbite de ces satellites est dite héliosynchrone et son altitude est de 786 km. L'accès à ces données est libre.

- La mission PlanetScope, initiée par l'entreprise américaine Planet Labs, qui a pour objectif de fournir des images quotidiennes de la Terre à haute résolution. Cette mission repose sur une constellation de plusieurs centaines de petits satellites, appelés Doves. Ces satellites permettent l'acquisition quotidienne d'images couvrant l'ensemble de la surface terrestre avec une emprise au sol de 24 km par 24 km pour chaque image et une résolution spatiale allant jusqu'à 3 mètres. Les images produites comportent 8 bandes spectrales. Ces satellites orbitent à une altitude variant entre 475 et 525 km sur une orbite héliosynchrone. L'accès à ces données, provenant du secteur privé, est payant. Le géoréférencement, la calibration et la précision radiométrique des images PlanetScope n'est pas toujours optimale et peut poser problème lors de l'utilisation de ces images (NASA, 2020; Frazier & Hemingway, 2021).

1.1.4.2. Les drones

Au cours des dernières décennies, le marché du drone civil s'est fortement développé. Une diversité de forme et de gabarit existe, les drones pouvant être à aile fixe, multirotors, peser moins de 250 grammes ou des tonnes, . . . Certains drones sont destinés au suivi de l'environnement, à une application scientifique, d'autres répondent à des applications militaires ou mêmes de loisir. Différents modèles sont présentés sur la Figure 2.



FIGURE 2 – Exemples de drones. De gauche à droite : drone multiroteur DJI Mavic 3M, drone à aile fixe léger, drone à aile fixe lourd.

Le type de matériel à utiliser dépend donc des besoins mais aussi du budget disponible. Le type d'acquisition réalisé et la qualité et/ou résolution des images varient également en fonction du matériel choisi, certains drones embarquant de simples appareils photos, d'autres des capteurs multispectraux. La précision des données GPS acquises lors des missions a également un impact sur le traitement des images acquises.

Bien que les drones à aile fixe disposent d'une endurance plus importante que les drones multirotors, leur prise en main est plus compliquée. Ils ont également besoin d'un espace

plus important au décollage et à l'atterrissage, ce qui ne convient pas à tous les milieux. La manœuvrabilité des drones multirotors en fait des outils plus performants lorsque l'étude est la source d'un besoin de superposition entre les images. En effet, ceux-ci sont capables de vol stationnaire et d'effectuer des virages serrés ce qui n'est pas le cas des drones à ailes fixes. Un compromis existe, les VTOLs, de l'anglais « Vertical Take-off and Landing ». Ces engins sont conçus de manière à pouvoir décoller verticalement mais également à pouvoir voler horizontalement grâce à des forces des portances générées le long des ailes. Certains sont également équipés pour le vol stationnaire.

1.2 Cadre du mémoire en entreprise

Le PRRPB, mené entre 2018 et 2023 et financé par la Banque Mondiale, a pour objectif principal de « restaurer la productivité des terres dans des paysages dégradés ciblés et, en cas de crise ou d'urgence admissible, d'apporter une réponse immédiate et efficace à ladite crise ou urgence admissible. ». La zone d'intervention couvre deux communes, celle de Buhinyuza dans la Province de Muyinga située à l'Est du pays et celle de Isale (Isare) dans la Province de Bujumbura. Au total, ce sont 22 collines qui ont fait l'objet d'études et d'aménagements ([République du Burundi, 2020](#)). Le bureau d'études belge SHER (Groupe ARTELIA) a été mandaté sur ce projet afin d'y effectuer des études visant à terme à :

- La caractérisation des zones d'études ;
- L'analyse des problématiques liées à l'hydrologie sur ces zones d'études ;
- Le design d'éléments anti-érosifs adaptés aux besoins et spécificités des sites ;
- Le suivi de la réalisation de ces éléments.

Lors de l'étape de caractérisation des zones d'études, l'identification de l'occupation des sols (land use et land cover) est réalisée par des opérateurs sur base d'images satellites (Google Earth). Les zones d'occupations de sol sont numérisées manuellement par un opérateur chargé de tracer des polygones. Ensuite, un opérateur terrain est engagé pour réaliser une validation. Le projet n'étant pas isolé et l'entreprise travaillant régulièrement dans des pays aux caractéristiques climatiques et topographiques similaires, une réflexion interne a permis de faire naître un projet visant à améliorer ce processus d'identification via l'utilisation d'outil de télédétection et de sources de données différentes. Un budget réservé à ce projet a notamment permis l'acquisition d'un drone et d'une station GNSS.

Ce mémoire a pour objectif principal la mise au point d'une méthode de cartographie de l'occupation du sol par classification d'images satellites dans le cadre de la gestion intégrée des ressources en eau (GIRE) au Burundi. Cette classification sera basée sur l'utilisation

de sources d'images à très haute résolution ou de combinaisons de ces images (satellite et drone). Cette méthode aura pour but d'être précise, efficace et capable de remplacer la précédente tout en n'augmentant pas la charge de travail liée à l'identification de l'occupation des sols. La méthode de cartographie à mettre en place doit pouvoir répondre à différents besoins émanant de l'entreprise SHER, en prenant le contexte du PRRPB comme cas d'étude.

La nature des aménagements proposés lors de projets de GIRE par l'entreprise varie entre autres en fonction de l'occupation du sol. Des classes d'occupation de sol à identifier ont déjà été déterminées par l'entreprise, il est donc nécessaire que la méthode s'adapte à celles-ci.

Du point de vue d'une entreprise, il est nécessaire que la méthode proposée soit non seulement efficace mais également compétitive vis-à-vis de la méthode actuellement utilisée. Une comparaison des méthodes sera donc réalisée. Différentes sources d'images sont également utilisées lors de la mise au point de la méthode de classification. Il sera donc nécessaire de comparer ces sources afin de pouvoir orienter le choix de l'entreprise.

Enfin, l'utilisation d'un drone permet l'acquisition par l'entreprise de divers types de données : des orthomosaïques à très haute résolution et des modèles numériques de surface (MNS). Le présent mémoire a également pour but de mettre en avant certains intérêts liés à ce mode d'acquisition de données et les bénéfices marginaux pouvant en découler.

2 Matériel et méthodes

2.1 Description des zones d'étude

L'étude a été réalisée au Burundi, qui est un pays d'Afrique de l'Est au climat tropical chaud et humide. Deux saisons sèches (courte en janvier et février, longue de mai à août) alternent avec deux saisons humides (de septembre à novembre et de février à mai). L'étude est réalisée sur deux communes distinctes, observables sur la Figure 3.

Zones d'études au Burundi.

CRS : WGS 84 (EPSG : 32735)
Image OpenStreetMap

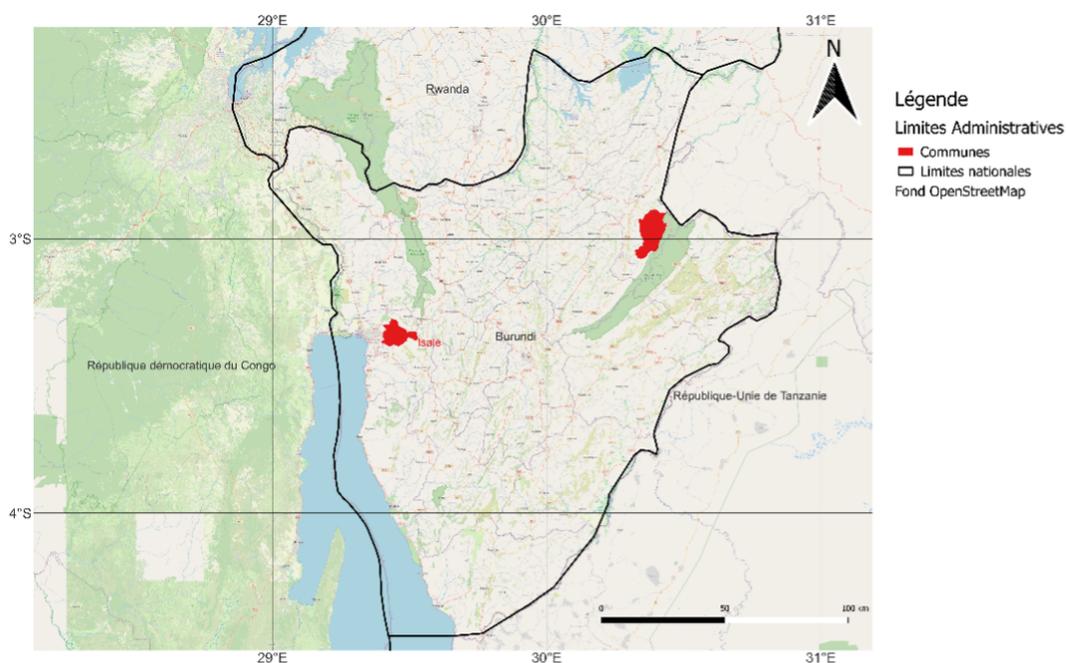


FIGURE 3 – Localisation des communes de la zone d'étude à l'échelle nationale.

La première, la commune de Isale, se trouve à l'Ouest du pays dans la province de Bujumbura Rural. La région est caractérisée par des reliefs accidentés, voire montagneux, aux altitudes variant entre 1000 et 2200 mètres. Des précipitations moyennes annuelles de 1200 mm y sont enregistrées. Les principaux types de sol rencontrés sont les Ferralsols et Nitisols (République du Burundi, 2020). Pour simplifier la compréhension, les zones de vol seront appelées bassin versant dans ce mémoire. Il est cependant important de noter que les zones sont en réalité des sous bassins versants ou des moitiés de bassins versants (flan de collines). Les bassins versants GIKO_02, MUHU_03 et NTAH_05 ont fait l'objet d'études dans cette région et sont présentés sur la Figure 4.

Zone d'étude de la commune de Isale (Burundi)

CRS : WGS 84 (EPSG : 32735)
Image Google Satellite

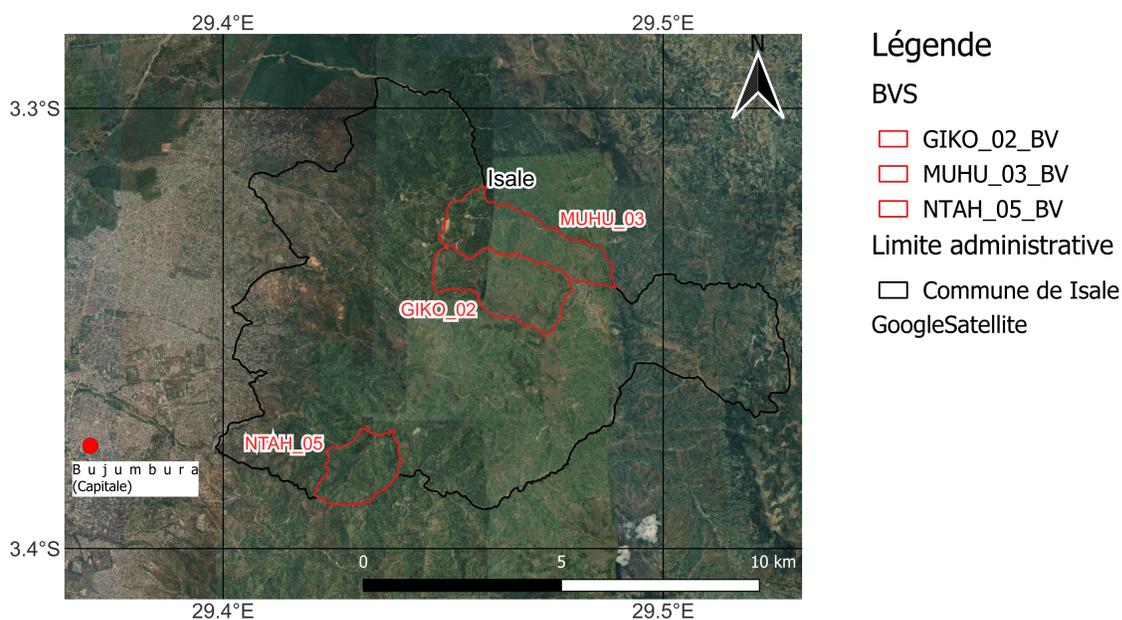


FIGURE 4 – Localisation des bassins versants étudiés sur la commune de Isale (Burundi).

La seconde zone est située à l'Est du pays, dans la province de Muyinga. Les deux bassins versants étudiés, NYAM_01 et KINY_03, se trouvent au sein de la commune de Buhinyuza. Le relief de cette région est moins accidenté et des précipitations moyennes annuelles de 925 mm sont enregistrées. Les principaux types de sols rencontrés sont les Vertisols et les Nitisols ([République du Burundi, 2020](#)). La Figure 5 présente la localisation de ces deux bassins versants.

Zone d'étude de la commune de Buhinyuza (Burundi)

CRS : WGS 84 (EPSG : 32735)
Image Google Satellite

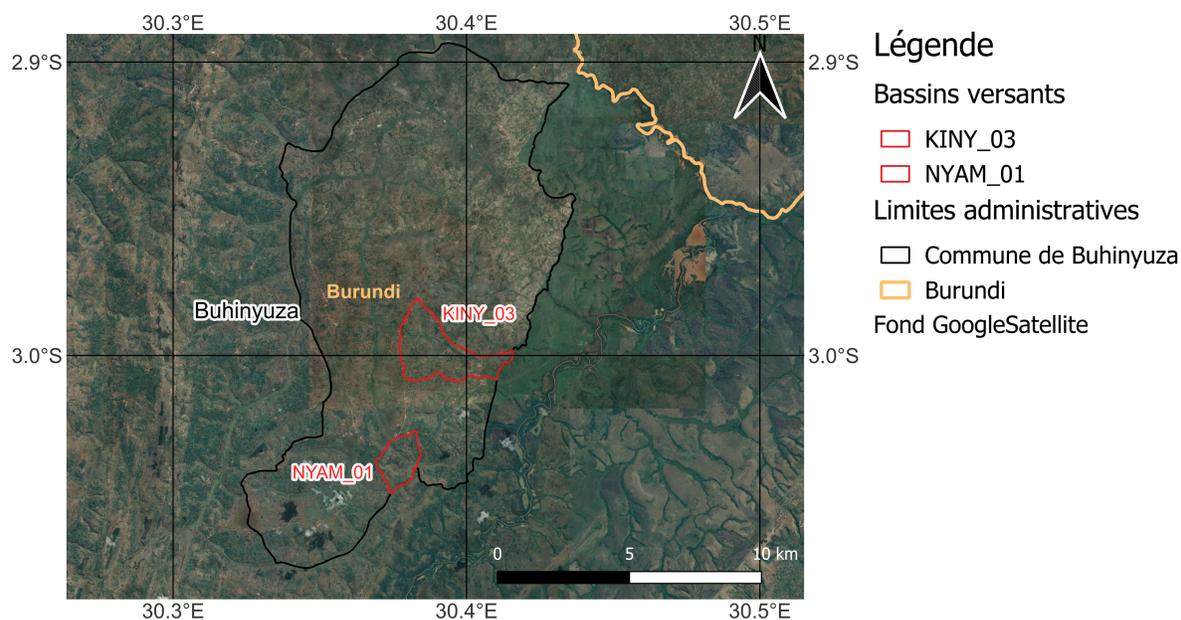


FIGURE 5 – Localisation des bassins versants étudiés sur la commune de Buhinyuza (Burundi).

Le Tableau 1 présente différentes caractéristiques des bassins versants.

TABLE 1 – Caractéristiques topographiques des bassins versants.

Commune	Bassin versant	Superficie [ha]	Pente Moyenne [°]	Altitude min [m]	Altitude max [m]
Isale	GIKO_02	455.9	21	1336	2024
Isale	MUHU_03	478.7	22	1136	1821
Isale	NTAH_05	281.8	27	978	1595
Buhinyuza	KINY_03	637.6	7	1357	1997
Buhinyuza	NYAM_01	236.4	10	1394	1563

Des différences de topographie entre les zones Est et Ouest sont clairement observables sur le Tableau 1. De cette différence découle également une différence dans les classes d'occupation de sol. Celle-ci seront présentées au paragraphe 2.2.3. mais nous noterons dès à présent la présence de la classe Marais au sein des bassins versants de la commune de Buhinyuza. Cette classe n'est pas présente sur les bassins de la commune de Isale.

2.2 Chaîne de traitements

La mise au point de la méthode de cartographie finale se base sur diverses sources de données, chacune d'entre elles devant subir divers traitements afin de les rendre exploitables.

La Figure 6 schématise l'approche développée. Les traitements seront détaillés dans la suite du document. L'ordinateur utilisé lors de l'ensemble des traitements présente les caractéristiques suivantes : Carte mère WS-C621E-SAGE Series 220298954500110 (Asus), CPU Intel(R) Xeon(R) Gold 6230R CPU @ 2.10GHz (26 cores), GPU NVIDIA GeForce GT 730 et RAM M393A2K40DB2-CVF 2x16GB + M393A8G40MB2-CVF 2x64GB.

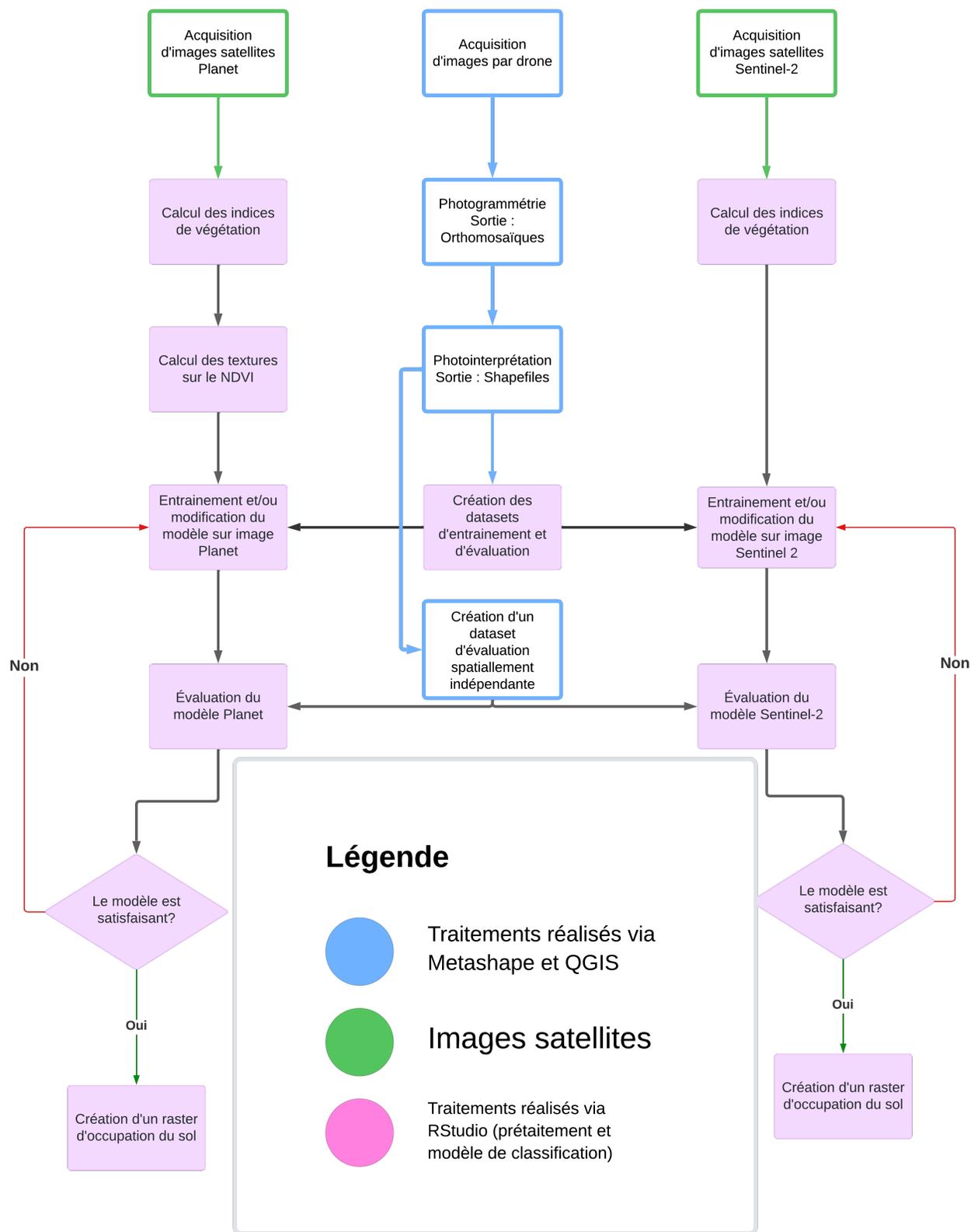


FIGURE 6 – Flowchart de la chaîne de traitements mise en place.

2.2.1 Images satellites

2.2.1.1. Acquisition

Ce mémoire étant réalisé en entreprise, la notion de gain en précision a dû être mise en relation avec l'évolution du prix d'acquisition des images. Pour cela, les images issues de deux missions d'acquisition différentes, Sentinel-2 et Planet Explorer, ont été utilisées. Les images Planet acquises proviennent du package de produits "analytic_8b_sr_udm2". Cela signifie que ces images, constituées de 8 bandes, ont été traitées pour corriger les effets de l'atmosphère et refléter les conditions de réflectance à la surface de la Terre. Ces images sont également harmonisées pour que les valeurs de bandes spectrales correspondent au mieux à celles des images Sentinel-2. Celles-ci proviennent du package "analytic" qui est calibré à la radiance au sommet de l'atmosphère (Planet Labs, 2024). Le Tableau 2 présente les différentes bandes spectrales disponibles pour chaque source de données satellite.

TABLE 2 – Comparaison des bandes spectrales et des résolutions des satellites Sentinel-2 et PlanetScope

Bande	Sentinel-2		PlanetScope	
	Longueur d'onde moyenne [nm]	Résolution [m]	Bande spectrale [nm]	Résolution [m]
Coastal Blue/Aerosol	443	60	431 - 452	3
Blue	490	10	465 - 515	3
Green	560	10	547 - 583	3
Green 1	-	-	546 - 586	3
Yellow	-	-	600 - 620	3
Red	665	10	650 - 680	3
Red Edge	705	20	697 - 713	3
Red Edge (2)	740	20	-	-
Red Edge (3)	783	20	-	-
Near InfraRed	842	10	845 - 885	3
Red Edge (4)	865	20	-	-
Water Vapour	945	60	-	-
SWIR - Cirrus	1375	60	-	-
SWIR (1)	1610	20	-	-
SWIR (2)	2190	20	-	-

Les images (PlanetScope et Sentinel) ont été acquises via le portail Planet Explorer. Les limites des bassins versants ainsi que de deux groupements de bassins versants relatifs aux deux communes de la zone d'études ont été importés sur la plateforme.

Pour les bassins versants situés sur la commune de Isale, les images Planet datent du 14 février 2024. Les images Sentinel 2 datent elles du 11 février 2024. Les images Planet acquises pour la commune de Buhinyuza datent du 14 mars 2024. Aucune image Sentinel-2 n'a été acquise sur la commune de Buhinyuza car les données disponibles entre le 1er décembre 2023 et le 31 avril 2024 ne pouvaient pas être utilisées lors de traitements en raison de la couverture nuageuse trop importante. Lorsque les modèles sont entraînés sur des images Sentinel-2,

seules les bandes Red, Green, Blue et NearInfrared sont utilisées.

Une fois commandées, les images sont disponibles au téléchargement sous format raster.

2.2.1.2. Traitements

Les images Planet sont fournies sous forme de fichiers rasters multibandes, chaque bande correspondant à un des spectres enregistrés. Lors de la commande, celles-ci sont coupées aux limites de la zone d'étude. Elles font également l'objet d'une harmonisation. Celle-ci a pour but d'ajuster les caractéristiques multispectrales des images selon un standard commun, les caractéristiques des capteurs et les conditions de capture pouvant varier. Aucun prétraitement supplémentaire ne sera réalisé dans le cadre de ce mémoire, si ce n'est de s'assurer de la correspondance de chaque bande raster avec une bande spectrale, les bandes raster n'étant pas nommées lors du téléchargement.

Les images Sentinel 2 sont fournies dans un dossier compressé. Ce dossier contient, en plus des métadonnées, un nombre de fichiers rasters correspondant au nombre de bandes spectrales commandées. L'étape de découpe décrite précédemment est également effectuée lors de la commande des images Sentinel 2. En vue de faciliter les traitements postérieurs, il convient de sélectionner les bandes nécessaires à l'identification de l'occupation de sol et d'assembler celles-ci en un raster multibandes. Il est également nécessaire de veiller au fait que les bandes n'ont pas toutes la même résolution. Dans le cadre de ce mémoire, ce sont les bandes dont la résolution est de 10 mètres qui ont été utilisées (bande : Red, Green, Blue et NIR).

2.2.2 Images drone

2.2.2.1. Acquisition

L'acquisition des images drones s'est faite sur les zones d'étude mentionnées ci-dessus entre le 22 février et le 24 mars 2024. Un drone, le DJI MAVIC 3M a été utilisé lors de la mission. Ce drone est de type multirotor et mesure 48 cm de large, 55 de long et 14 de haut pour un poids de 951 grammes. Il est équipé d'une caméra Red, Green, Blue (RGB) 4/3 CMOS de 20MP ainsi que de quatre caméras multispectrales 1/2.8" CMOS de 5MP. Les bandes multispectrales enregistrées sont le Green (G) (560 ± 16 nm), Red (R) (650 ± 16 nm), Red Edge (RE) (730 ± 16 nm) et Near-Infrared (NIR) (860 ± 26 nm). Pour la navigation, le drone est équipé d'un GPS différentiel et se connecte au Système Mondial de navigation par satellite (GNSS) via les systèmes GPS, Galileo et Beidou. Un module RTK est également connectable. Si le vol est effectué avec RTK, le système GLONASS est également accessible. Un capteur de lumière incidente est intégré au drone. Celui-ci permet la correction de la radiométrie en fonction des valeurs mesurées et l'adaptation de la Shutter speed.

Afin de déterminer les paramètres de vol, 3 jours de tests ont été nécessaires. En effet, ce modèle de drone permet d'ajuster plusieurs paramètres de vol ayant un impact sur la qualité de l'acquisition mais aussi sur le temps nécessaire à celle-ci. Différentes altitudes de vol ont été testées durant les premiers jours de vol sur une zone de 22 hectares sur le bassin versant de GIKO_02 situé sur la commune de Isale. Les temps de vol et la résolution en fonction de l'altitude déterminée lors de ces tests sont présentés sur le Tableau 3.

TABLE 3 – Résolution en fonction de l'altitude.

Altitude [m]	Temps de vol [min]	Résolution [cm/pixel]
60	18	1,6
100	10	2,7
200	4,5	5,3

Une altitude de vol de 200 mètres permet un compromis entre une bonne résolution (de l'ordre de 5 cm/pixel) et un temps d'acquisition acceptable.

La vitesse de vol a un impact sur la durée d'acquisition ainsi que sur la décharge des batteries. Elle peut également avoir un impact sur la qualité des images dans le cas d'une vitesse trop élevée. Des tests ont été effectués pour déterminer la vitesse maximale permettant une acquisition de qualité. La vitesse maximale du drone de 15 m/s a été choisie car elle permet une acquisition de qualité et de réduire le temps de vol au maximum.

Il a donc été choisi de réaliser l'ensemble des vols à une altitude de 200 mètres, à une vitesse de 15 m/s et d'utiliser le mode de suivi de l'altitude soit en temps réel soit sur base d'un Modèle Numérique de Terrain (MNT). L'angle de caméra était systématiquement NADIR (visée verticale). Après des tests effectués sur le terrain, la valeur ISO a été verrouillée sur 100. La Shutter speed était adaptée lors de la prise de chaque image grâce aux mesures effectuées par le capteur de lumière spectral.

2.2.2.2. Traitements photogrammétriques

Une chaîne de traitements photogrammétriques permet de transformer les images brutes en une orthoimage et un MNS (RGB et multispectrale (MS)). Le logiciel Agisoft Metashape v1.8.0. a été utilisé pour la réalisation des étapes de photogrammétrie. Les différents traitements ont été réalisés sur chaque bassin versant indépendamment et sont présentés sur la Figure 7. La résolution des orthoimages RGB varie entre 6 et 8 centimètres, pour le MS de 8 à 12 centimètres.

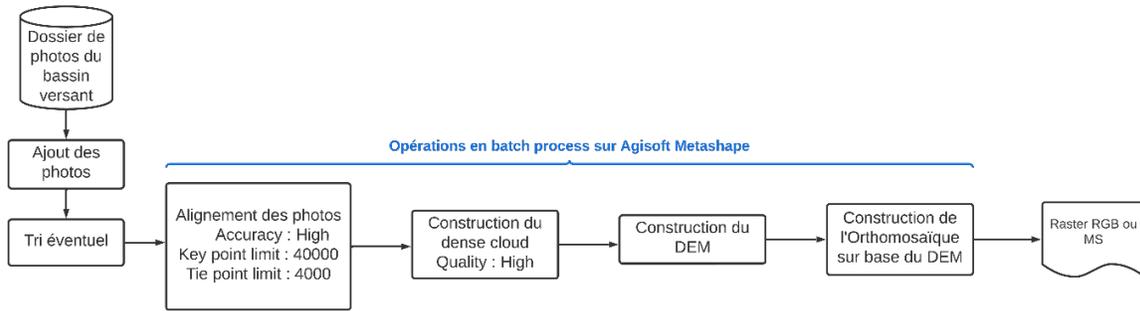


FIGURE 7 – Traitements des images drones.

Une fois les orthomosaïques extraites de Metashape, des traitements sont encore nécessaires afin de réaliser une classification sur celles-ci. Les rasters RGB et MS ayant été traités séparément et étant issus d'images différentes, ceux-ci n'ont pas la même emprise spatiale ni la même résolution et leurs pixels ne sont pas alignés. Lors de la classification, dont les étapes sont présentées au chapitre 2.2.6, il est nécessaire de disposer d'un raster duquel sont extraites les valeurs de différentes bandes en vue de calculer différents indices de végétation. Pour réaliser cette classification sur les images acquises par drone, il est donc nécessaire d'aligner le raster RGB sur le raster MS de chaque bassin. Le raster RGB aligné et le raster MS sont ensuite coupés aux limites du bassin versant. Ces étapes sont réalisées via l'outil QGIS. Un raster multibande comprenant l'ensemble des bandes RGB et MS est ensuite créé via un code sur RStudio.

2.2.3 Photointerprétation

La photointerprétation est réalisée sur les orthomosaïques. Elle a pour but de créer un set de données permettant d'entraîner un modèle de prédiction par classification supervisée. Les différentes classes d'occupation du sol sont :

- a) Boisement dense (BDEN)
- b) Boisement dégradé (BDEG)
- c) Culture pérenne (CP)
- d) Culture saisonnière (CS)
- e) Eau (EAU)
- f) Friche / Sol nu (FR)
- g) Habitation (HA)
- h) Marais (MA)
- i) Zone dégradée (ZD)

Celles-ci sont représentées sur la Figure 8.

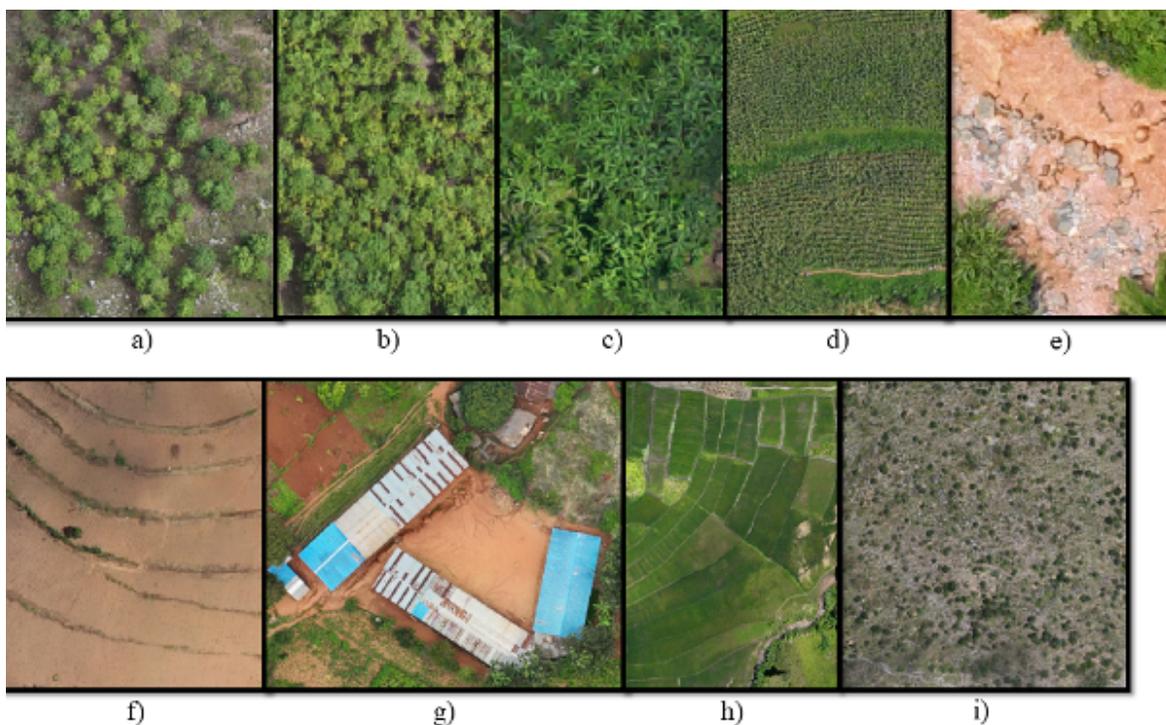


FIGURE 8 – Illustration des différentes classes d'occupation de sol. Légende : a) Boisement dégradé (BDEG), b) Boisement dense (BDEN), c) Culture pérenne (CP), d) Culture saisonnière (CS), e) Eau (EAU), f) Friche (FR), g) Habitation (HA), h) Marais (MA) et i) Zone dégradée (ZD).

Ces classes ne sont cependant pas systématiquement présentes sur chaque bassin versant. Ainsi, les bassins de la commune de Buhinyuza ne présentent pas la classe Eau tandis qu'à Isale, c'est la classe Marais qui est absente. La définition des classes a été réalisée en tenant compte des besoins de l'entreprise, l'occupation du sol conditionnant le type d'aménagement anti-érosion proposé. La photointerprétation est réalisée via le logiciel QGIS version 3.36.3.

2.2.3.1. Définition des zones d'entraînement et de validation

Pour réaliser cette photointerprétation, un projet Qgis a été créé pour chaque bassin versant. Une couche au format shapefile de polygones est créée. La photointerprétation par polygones consiste à tracer des polygones et à leur attribuer une valeur correspondant à une classe d'occupation du sol. Le nombre de polygones par classe ainsi que la surface qu'ils couvrent sont laissés libres, la surface et le nombre d'objets des différentes classes variant en fonction de celle-ci. Lors de la numérisation des objets de la classe Habitation, le contour précis des bâtiments est suivi. La Figure 9 illustre le résultat de la photointerprétation sur le bassin versant GIKO_02.

Photointerprétation sur le bassin versant versant GIKO_02 (Commune de Isale)

CRS : WGS 84 (EPSG : 32735)

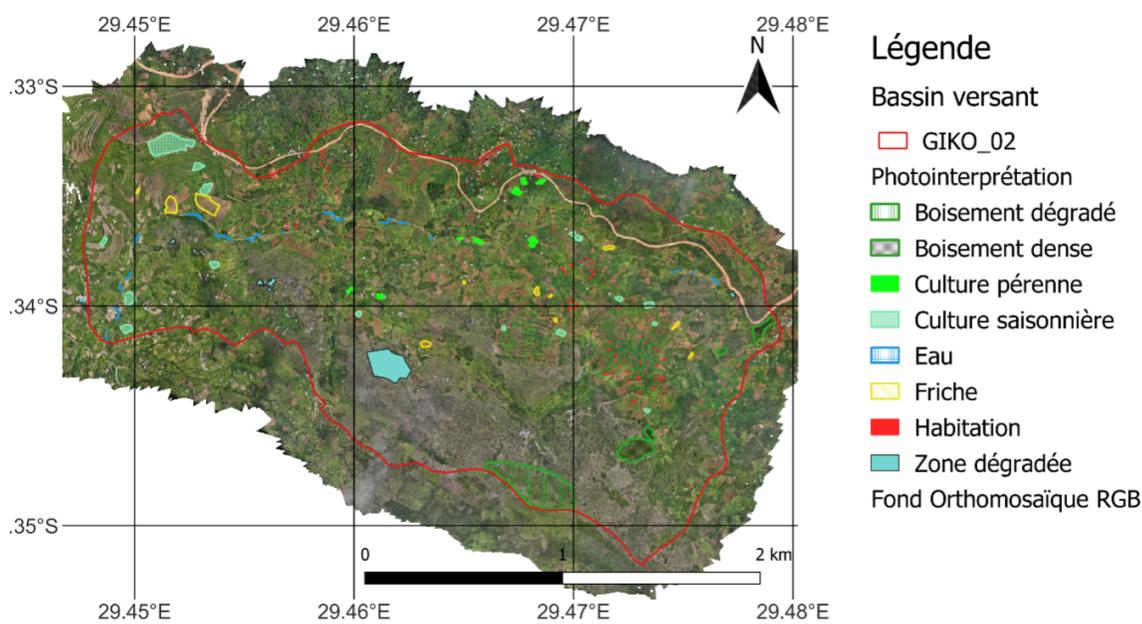


FIGURE 9 – Photointerprétation sur le bassin versant GIKO_02 (commune de Isale, Burundi).

Au final, une couche de polygones sera créée pour chaque bassin versant. La référence à ces couches se fera en parlant des polygones/de la couche du bassin (e.g. polygones/couche de GIKO_02). Afin de permettre l'évaluation de la capacité des modèles à prédire les classes sur des bassins versants différents de celui ou ceux dont sont issus les polygones d'entraînement, ces couches seront parfois regroupées en une seule lors de l'entraînement ou de l'évaluation des modèles. Pour chaque commune, un shapefile de polygones supplémentaire est créé dans le même but. Ces couches couvrent l'ensemble des bassins versants étudiés au sein de la commune concernée et sont constituées de polygones spatialement indépendants des polygones des couches citées précédemment. Elles ne seront utilisées que lors de l'évaluation de certains modèles. La référence à ces couches se fera en parlant des polygones/de la couche de la commune (e.g. polygones/couche de Isale). Les Tableaux 4, 5, 6 et 7 présentent la répartition des surfaces de chaque classe ainsi que le nombre de pixels compris dans chacune d'elles pour les différentes sources d'images au sein des couches de polygones créées.

TABLE 4 – Surface et nombre de pixels de chaque classe pour GIKO_02

Classe	Surface [ha]	Nombre de pixels Planet	Nombre de pixels Sentinel-2
Boisement dégradé	6,01	6661	595
Boisement dense	2,40	2659	243
Culture pérenne	0,74	815	71
Culture saisonnière	3,14	3484	318
Eau	0,33	386	62
Friche/sol nu	1,54	1706	150
Habitation	0,36	386	137
Zone dégradée	2,90	3210	282
Marais	/	/	/

TABLE 5 – Surface et nombre de pixels de chaque classe pour MUHU_03

Classe	Surface [ha]	Nombre de pixels Planet	Nombre de pixels Sentinel-2
Boisement dégradé	0,89	987	92
Boisement dense	1,17	1300	121
Culture pérenne	0,96	1067	89
Culture saisonnière	0,91	998	96
Eau	0,80	888	88
Friche/sol nu	1,02	1122	102
Habitation	1,11	1234	211
Zone dégradée	3,62	4029	358
Marais	/	/	/

TABLE 6 – Surface et nombre de pixels de chaque classe pour NTAH_05

Classe	Surface [ha]	Nombre de pixels Planet	Nombre de pixels Sentinel-2
Boisement dégradé	0,81	897	78
Boisement dense	0,49	542	49
Culture pérenne	0,80	886	78
Culture saisonnière	1,34	1484	140
Eau	0,72	783	68
Friche/sol nu	1,09	1210	108
Habitation	0,70	782	175
Zone dégradée	0,59	658	64
Marais	/	/	/

TABLE 7 – Surface et nombre de pixels de chaque classe pour KINY_03 et NYAM_01

Classe	Surface [ha] KINY_03	Nombre de pixels Planet KINY_03	Surface [ha] NYAM_01	Nombre de pixels Planet NYAM_01
Boisement dégradé	1,60	1763	2,95	3269
Boisement dense	0,89	994	0,19	214
Culture pérenne	1,32	1463	1,39	1536
Culture saisonnière	1,98	2205	1,11	1230
Friche/sol nu	3,01	3319	0,86	955
Habitation	0,34	373	0,14	153
Zone dégradée	2,48	2743	0,48	282
Marais	2,95	3270	2,58	2852

2.2.4 Les indices de végétation

Les indices de végétation, issus des combinaisons mathématiques de différentes bandes spectrales, permettent l'analyse des couverts végétaux. Ils sont fréquemment utilisés comme facteurs lors de la classification d'utilisation du sol via télédétection. Les indices végétaux utilisés lors de différentes recherches ainsi que les bandes spectrales servant à les calculer varient en fonction des besoins et de l'importance que leur apporte le modèle développé (Chen et al., 2015; Wang et al., 2015; Rego et al., 2020; Sefercik et al., 2021). Les quatre indices suivants ont été considérés pour construire les modèles.

1. L'indice de végétation par différence normalisée (Normalized Difference Vegetation Index : NDVI) calculé comme suit :

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

Cet indice se base sur le fait que la végétation verte a tendance à absorber l'énergie issue de la bande rouge tandis que l'énergie de la bande NIR est plus susceptible d'être réfléchi. Le ratio entre ces deux bandes permet donc de mettre en avant la présence/l'absence de végétation en un point donné (Rousse, 1973).

2. L'indice d'eau par différence normalisée (Normalized Difference Water Index : NDWI) calculé comme suit :

$$NDWI = \frac{G - NIR}{G + NIR}$$

Le principe sur lequel se base le calcul de cet indice est que l'énergie de la bande verte est fortement réfléchi par l'eau à l'inverse de l'énergie de la bande NIR qui est elle absorbée (McFEETERS, 1996).

3. L'indice de végétation à rapport normalisé (Normalized Ratio Vegetation Index : NRVI) calculé comme suit :

$$NRVI = \frac{\frac{RED}{NIR-1}}{\frac{RED}{NIR+1}}$$

Cet indice se base sur le même principe physique que le NDVI mais permet de réduire les effets de variations dues aux conditions d'acquisition d'images ainsi qu'aux différences de caractéristiques de surface (Baret & Guyot, 1991).

4. L'indice de végétation transformé (Transformes vegetation index : TVI) calculé comme suit :

$$TVI = \sqrt{\frac{NIR - RED}{NIR + RED}} + 0.5$$

L'ensemble des indices ont été calculés grâce à la fonction « spectralIndices » de la librairie RStoolbox. Cette fonction, lorsque les bandes spectrales contenues dans le raster analysés lui sont renseignées, permet le calcul de l'ensemble des indices de végétation répertoriés dont le calcul est effectué sur base des bandes spécifiées. L'outil permet également une sélection des indices à calculer afin d'éviter le bruit que pourrait causer l'entraînement d'un modèle sur un trop grand nombre d'entre eux (Leutner et al., 2015). Cette sélection a été réalisée de manière empirique. Les premiers modèles entraînés l'ont été sur l'ensemble des indices calculables en fonction des bandes disponibles. Un graphique de l'importance des variables dans les modèles a ensuite été utilisé pour sélectionner celles-ci. Les effets de la sélection de différents indices sur les performances des modèles étaient ensuite analysés.

2.2.5 Les textures

La texture est définie par Haralick comme une mesure de la variation spatiale des intensités de gris des pixels au sein d'une image (Haralick, 1979). La texture est donc composée de trois éléments selon Hall-Beyer : « la différence tonale entre les pixels, la distance sur laquelle cette différence est mesurée et la directionnalité » (Hall-Beyer, 2017a). Afin d'obtenir une représentation numérique de la texture, nécessaire à l'implémentation de celle-ci comme facteur au sein de modèles de prédiction informatiques, la notion de matrice de cooccurrence de niveaux de gris (GLCM) a été mise au point par Haralick (Haralick, 1979; Hall-Beyer, 2017a). Différents descripteurs statistiques issus de cette matrice sont alors utilisés en fonction des besoins de l'étude.

L'impact positif de l'utilisation de GLCM et de leurs descripteurs sur les systèmes de classification et leur précision a été largement documenté. De plus, les textures sont moins sensibles à l'hétérogénéité de lumière et d'exposition étant donné qu'elles ne se concentrent que sur les relations locales entre pixels au lieu de considérer des valeurs individuelles absolues (Haralick et al., 1973; Clausi, 2002; Hall-Beyer, 2017b). Dans le cadre de ce mémoire, les statistiques liées à la GLCM sont calculées à partir du NDVI.

2.2.6 La classification par RandomForest

Le logiciel RStudio est utilisé pour réaliser les différents traitements de classification. Les éléments de base nécessaires à celle-ci sont les images satellites et les couches contenant les polygones de références utilisés dans les phases d'entraînement ou de validation des modèles de classification. Des éléments supplémentaires tels que la texture et la pente ont été intégrés au modèle au fur et à mesure de l'avancée du projet afin de répondre aux problèmes pouvant survenir lors de l'évaluation statistique du modèle précédent. L'objectif principal de ce mémoire étant la mise au point d'une méthode de classification, certains de ces éléments ne sont pas repris dans le modèle final. Ne seront présentés que les éléments ayant eu un impact positif sur le modèle.

Le classificateur utilisé dans le présent mémoire est le RandomForest (RF). Cette méthode de classification est de plus en plus utilisée de nos jours et permet une classification supervisée et non-paramétrique (Rodriguez-Galiano et al., 2012; Marais-Sicre et al., 2017). La Figure 10 présente le schéma du code RStudio créé pour la classification.

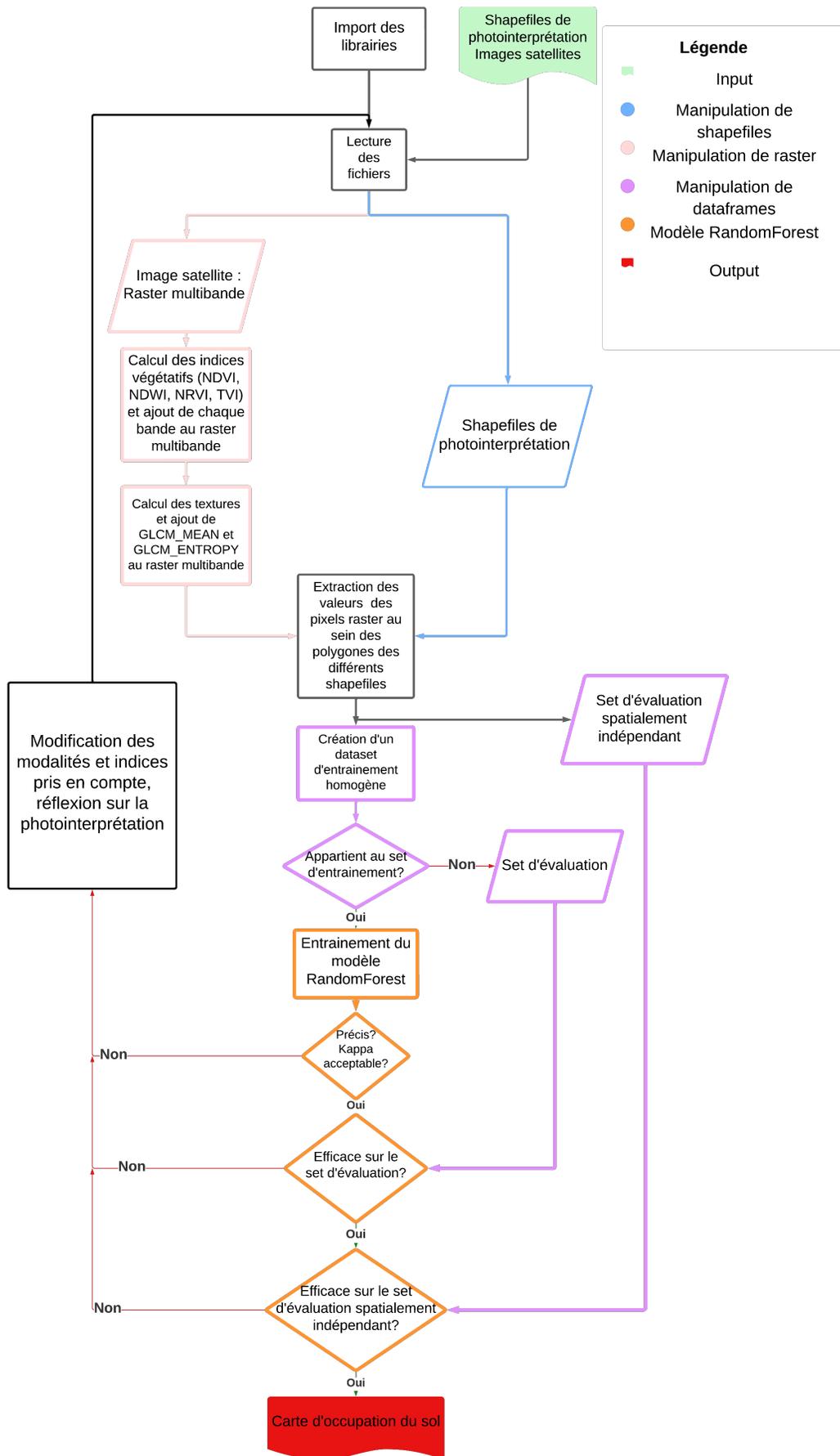


FIGURE 10 – Présentation de la structure du code réalisé pour la classification.

L'entraînement d'un modèle de classification via un modèle RF exige que les données d'entraînement du modèle présentent une certaine homogénéité quant au nombre d'entités considérées par classe. En effet, l'entraînement de modèles sur des sets de données présentant une ou des classes majoritaires et/ou une ou des classes minoritaires a pour effet de mener à une précision de près de 100 pourcents pour la/les classes majoritaires et une précision bien plus faible pour la/les classes minoritaires, on parle alors d'overfitting (He & Garcia, 2009; Millard & Richardson, 2015). Il a été vu sur les Tableaux 4, 5, 6 et 7 au chapitre 2.2.3 que la distribution des surfaces couvertes par les polygones de chaque classe dans les différents bassins versants n'est pas uniforme. Pour remédier à ce problème, avant l'entraînement du modèle, un nouveau set de données est créé. Pour se faire, le nombre d'observations de la classe minoritaire est calculé. Le nouveau set est formé par la sélection aléatoire d'entités de chaque classe, le nombre d'entités par classe sélectionnées étant égale à 70% du nombre d'observations de la classe minoritaire. Lors de l'extraction des valeurs de chaque variable (réalisée via la fonction `exact_extract()` du package `exactextract`) au sein des polygones de photointerprétation, la variable « coverage fraction » est créée. Elle indique la proportion de la surface de chaque pixel du raster incluse dans ce polygone. Pour entraîner les différents modèles, seules les valeurs provenant de pixels dont cette variable est supérieure à 50% est sélectionnée.

Afin d'évaluer la capacité de la méthode mise au point à être transposée, c'est-à-dire la possibilité d'utiliser un modèle dont les données d'entraînement sont issues d'un ou plusieurs bassins pour prédire l'occupation de sol sur un ou plusieurs autres bassins, différents modèles ont été entraînés et évalués. Les combinaisons des polygones de photointerprétation de plusieurs bassins sont présentées sur le Tableau 8. Aucune donnée Sentinel-2 n'étant disponible pour la région de Buhinyuza, seuls les quatre premiers modèles ont été créés avec ces images. Les images acquises par drone ne feront pas l'objet d'analyse visant à la prédiction sur d'autres bassins versants. En effet, l'intérêt de pouvoir transposer un modèle à un autre bassin versant est de ne pas avoir à réaliser l'acquisition d'images par drone nécessitant un opérateur terrain.

TABLE 8 – Combinaison des couches de polygones dans les différents modèles composés.

Entraînement sur les bassins versants :	Évaluation des prédictions sur le/les bassin(s) versant(s) :	Nom du modèle
GIKO_02 et MUHU_03	NTAH_05	1
MUHU_03 et NTAH_05	GIKO_02	2
GIKO_02 et NTAH_05	MUHU_03	3
GIKO_02, MUHU_03 et NTAH_05	ISALE	4
KINY_03 et NYAM_01	BUHINYUZA	5

2.2.7 Évaluation des modèles

L'évaluation de la précision des classifications est déterminée par différents indicateurs calculés via le logiciel RStudio. La matrice de confusion, la précision et le kappa sont des indicateurs de la qualité générale de la classification tandis que le reste des indicateurs décrivent la qualité de la classification pour chaque classe. Les indicateurs statistiques utilisés sont présentés à la suite de ce paragraphe.

La précision globale est le rapport entre le nombre de prédictions correctes et le nombre total de prédictions/d'entités. Le kappa de Cohen évalue l'accord entre les prédictions faites par un modèle et les valeurs de référence en prenant en compte le fait que certains accords peuvent survenir par hasard. Un kappa compris entre 0 et 0,2 notera un accord léger, entre 0,2 et 0,4 un accord juste, entre 0,4 et 0,6 un accord modéré, entre 0,6 et 0,8 un accord substantiel et entre 0,8 et 1 un accord presque parfait (McHugh, 2012; Warrens, 2015).

Une matrice de confusion compare les prédictions d'un modèle aux valeurs de référence. Elle permet de quantifier les vrais positifs (TP), les vrais négatifs (TN), les faux positifs (FP) et les faux négatifs (FN). Un vrai positif est une entité prédite comme appartenant à une classe et y appartenant réellement. Un vrai négatif est une entité prédite comme n'appartenant pas à une classe et n'y appartenant pas en réalité. Un faux positif est une entité prédite comme appartenant à une classe mais n'y appartenant pas réellement. Enfin, un faux négatif est une entité prédite comme n'appartenant pas à une classe mais y appartenant en réalité. Le Tableau 9 montre le schéma théorique d'une matrice de confusion issue d'une classification à deux classes.

TABLE 9 – Matrice de confusion binaire théorique

		Références	
		Positif	Négatif
Prédictions	Positif	Nb de TP	Nb de FP
	Négatif	Nb de FN	Nb de TN

- La sensibilité correspond au ratio : $\frac{TP}{TP+FN}$. Elle évalue la capacité à identifier correctement les entités appartenant à une classe.
- La spécificité évalue la capacité d'un modèle à identifier une instance comme n'appartenant pas à une classe. Elle est calculée comme suit : $\frac{TN}{TN+FP}$.
- La valeur prédictive positive mesure la part d'entités appartenant à une classe sur le nombre total d'entités prédites comme appartenant à cette classe. Elle est calculée

comme suit : $\frac{TP}{TP+FP}$.

- La valeur prédictive négative correspond au ratio : $\frac{TN}{TN+FN}$. Elle évalue la capacité du modèle à identifier comme négatif une entité l'étant réellement.
- Le taux de détection est le produit entre la sensibilité et la prévalence, la prévalence étant la proportion d'entités de chaque classe vis-à-vis du nombre total d'entités.
- La prévalence de détection correspond au ratio : $\frac{TP+FP}{N}$, N étant le nombre total d'entités. Comparée à la prévalence, elle indique si le modèle surestime ou sous-estime le nombre d'entités d'une classe.
- L'exactitude équilibrée correspond à la moyenne arithmétique entre la sensibilité et la spécificité.

2.3 Méthode de cartographie de l'entreprise SHER

Afin de pouvoir comparer les deux méthodes, la méthode actuellement utilisée par la SHER est présentée.

Sur base d'images récentes de Google Earth, un opérateur, familier aux occupations de sol présentes sur la zone d'étude, est chargé de numériser de manière manuelle ces zones d'occupation de sol. En cas d'incertitude, les données antérieures peuvent aider. Des travaux de validations sur le terrain sont ensuite menés. Ceux-ci permettent d'estimer avec quelle précision les travaux de numérisation décrivent l'occupation de sol. En terme de temps de travail, un opérateur est en mesure de numériser entre 500 et 1000 hectares par jour. La validation terrain permet de couvrir environ 500 hectares par jour et par opérateur sans compter le temps nécessaire à l'accès à la zone.

3 Résultats

3.1 Qualité des classifications de l'occupation du sol

3.1.1 Évaluation des modèles par bassin versant

Ce paragraphe décrit les résultats obtenus via les différents modèles entraînés sur un seul bassin versant. Un modèle a été entraîné sur chaque bassin versant et sur chaque source de données (Sentinel-2 et PlanetScope).

3.1.1.1 PlanetScope

L'effet de différentes variables et de différents indices sur la précision des modèles entraînés et prédits sur les images PlanetScope a été évalué. Les précisions globales et kappas pour chaque bassin versant sont présentés sur le Tableau 10.

TABLE 10 – Précision, kappa et nombre de pixels par classe dans le set d'entraînement d'images Planet des modèles par bassin versant.

Bassins	GIKO_02	MUHU_03	NTAH_05	KINY_03	NYAM_01
Précision globale	0.90	0.92	0.88	0.85	0.83
Kappa	0.89	0.91	0.86	0.83	0.80
Pixels par classe dans le set d'entraînement	544	739	459	839	204

Une capacité des modèles à prédire correctement l'occupation de sol au sein du bassin versant duquel proviennent les données d'entraînement est constatée pour chacun de ceux-ci. Une classification légèrement moins efficace est observée pour les bassins versants de la commune de Buhinyuza (KINY_03 et NYAM_01) bien que les performances des modèles restent toujours supérieures à 0,80.

Les données statistiques directement issues d'un modèle RF telles que les matrices de confusion et les tableaux de statistiques par classe donnent des informations plus complètes sur la capacité du modèle à identifier correctement chaque classe. Ceux présentant des résultats jugés moins bons sont présentés dans le Tableau 11 et dans le Tableau 12.

TABLE 11 – Matrice de confusion du modèle entraîné sur le bassin NTAH_05 (images PlanetScope).

Prédiction Référence	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
BDEG	443	11	9	3	0	0	0	0
BDEN	10	425	12	3	0	0	0	1
CP	5	14	422	5	1	5	8	0
CS	0	5	5	418	0	34	3	3
EAU	0	0	0	0	377	30	9	8
FR	1	0	1	19	58	323	14	19
HA	0	0	9	7	14	17	395	15
ZD	0	4	1	4	9	50	30	413

TABLE 12 – Tableau statistique du modèle entraîné sur le bassin NTAH_05 (images PlanetScope).

Classe	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
Sensibilité	0,965	0,926	0,919	0,911	0,821	0,704	0,861	0,900
Spécificité	0,993	0,992	0,988	0,984	0,985	0,965	0,981	0,970
Valeur prédictive positive	0,951	0,942	0,917	0,893	0,889	0,743	0,864	0,808
Valeur prédictive négative	0,995	0,989	0,989	0,987	0,975	0,958	0,980	0,985
Prévalence	0,125	0,125	0,125	0,125	0,125	0,125	0,125	0,125
Taux de détection	0,121	0,116	0,115	0,114	0,103	0,088	0,108	0,113
Prévalence de détection	0,127	0,123	0,125	0,128	0,116	0,119	0,125	0,139
Exactitude équilibrée	0,979	0,959	0,954	0,948	0,903	0,834	0,921	0,935

On observe dans le Tableau 10 que les bassins GIKO_02 et MUHU_03 ne présentent pas de réels problèmes. Pour ce qui est du bassin versant NTAH_05, des confusions plus importantes sont observées pour les classes Eau, Friche, Habitat et Zone dégradée. La valeur prédictive positive des classes Friche et Zone dégradée sont faibles par rapport à celles des autres classes. Le modèle a tendance à attribuer à ces classes des pixels appartenant en réalité à d'autres classes comme illustré sur la matrice de confusion. Les Tableaux 13, 14, 15 et 16 présentent un résumé statistiques obtenu à la suite de l'entraînement de modèles sur les bassins KINY_03 et NYAM_01.

TABLE 13 – Matrice de confusion du modèle entraîné sur le bassin NYAM_01 (images PlanetScope).

Prédiction Référence	BDEG	BDEN	CP	CS	FR	MA	ZD
BDEG	173	2	5	23	0	4	0
BDEN	4	188	16	3	0	9	0
CP	8	12	145	6	1	45	0
CS	13	0	9	157	8	18	0
FR	3	0	2	9	194	2	0
MA	2	2	27	6	0	126	0
ZD	1	0	0	0	1	0	204

TABLE 14 – Tableau statistique du modèle entraîné sur le bassin NYAM_01 (images PlanetScope).

Classe	BDEG	BDEN	CP	CS	FR	MA	ZD
Sensibilité	0,848	0,922	0,711	0,770	0,951	0,618	1,000
Spécificité	0,972	0,974	0,941	0,961	0,987	0,970	0,998
Valeur prédictive positive	0,836	0,855	0,668	0,766	0,924	0,773	0,990
Valeur prédictive négative	0,975	0,987	0,951	0,962	0,992	0,938	1,000
Prévalence	0,143	0,143	0,143	0,143	0,143	0,143	0,143
Taux de détection	0,121	0,132	0,102	0,110	0,136	0,088	0,143
Prévalence de détection	0,145	0,154	0,152	0,144	0,147	0,114	0,144
Exactitude équilibrée	0,910	0,948	0,826	0,865	0,969	0,794	0,999

TABLE 15 – Matrice de confusion du modèle entraîné sur le bassin KINY_03 (images PlanetScope).

Prédiction Référence	BDEG	BDEN	CP	CS	FR	MA	ZD
BDEG	726	4	4	17	0	41	46
BDEN	10	726	51	1	0	65	4
CP	18	48	673	21	0	79	6
CS	29	2	23	778	3	3	60
FR	1	0	0	6	822	2	20
MA	30	56	85	4	0	609	37
ZD	25	3	3	12	14	40	666

TABLE 16 – Tableau statistique du modèle entraîné sur le bassin KINY_03 (images PlanetScope).

Classe	BDEG	BDEN	CP	CS	FR	MA	ZD
Sensibilité	0,865	0,865	0,802	0,927	0,980	0,726	0,794
Spécificité	0,978	0,974	0,966	0,976	0,994	0,958	0,981
Valeur prédictive positive	0,866	0,847	0,796	0,866	0,966	0,742	0,873
Valeur prédictive négative	0,978	0,978	0,967	0,988	0,997	0,955	0,966
Prévalence	0,143	0,143	0,143	0,143	0,143	0,143	0,143
Taux de détection	0,124	0,124	0,115	0,133	0,140	0,104	0,113
Prévalence de détection	0,143	0,146	0,144	0,153	0,145	0,140	0,130
Exactitude équilibrée	0,922	0,920	0,884	0,952	0,987	0,842	0,887

La classe Marais présente dans la commune de Buhinyuza pose des problèmes. Pour les deux bassins versants, la sensibilité de cette classe est faible. Cela signifie que le modèle éprouve des difficultés à identifier les pixels appartenant à cette classe. Le modèle confond les marais avec la végétation dense/forestière, les cultures pérennes ou encore les boisements denses.

Dans une autre mesure, ce phénomène est également observé pour les classes Culture pérenne et Zone dégradée au sein du bassin KINY_03. Dans ce cas, les pixels de ces occupations sont parfois attribués aux classes Boisement dégradé, Culture saisonnière ou Marais.

Pour ce qui est du bassin versant NYAM_01, une sensibilité de 1 et une spécificité de 0,9884 sont observées pour la classe Zone dégradée. Ces résultats sont surprenants lorsque l'on compare les sensibilités et spécificités des autres classes d'occupation du sol.

L'importance des variables au sein des modèles varie d'un bassin versant à l'autre. Les Figures 11 et 12 ci-dessous illustrent les différences entre les bassins KINY_03 et NYAM_01 situés sur la commune de Buhinyuza.

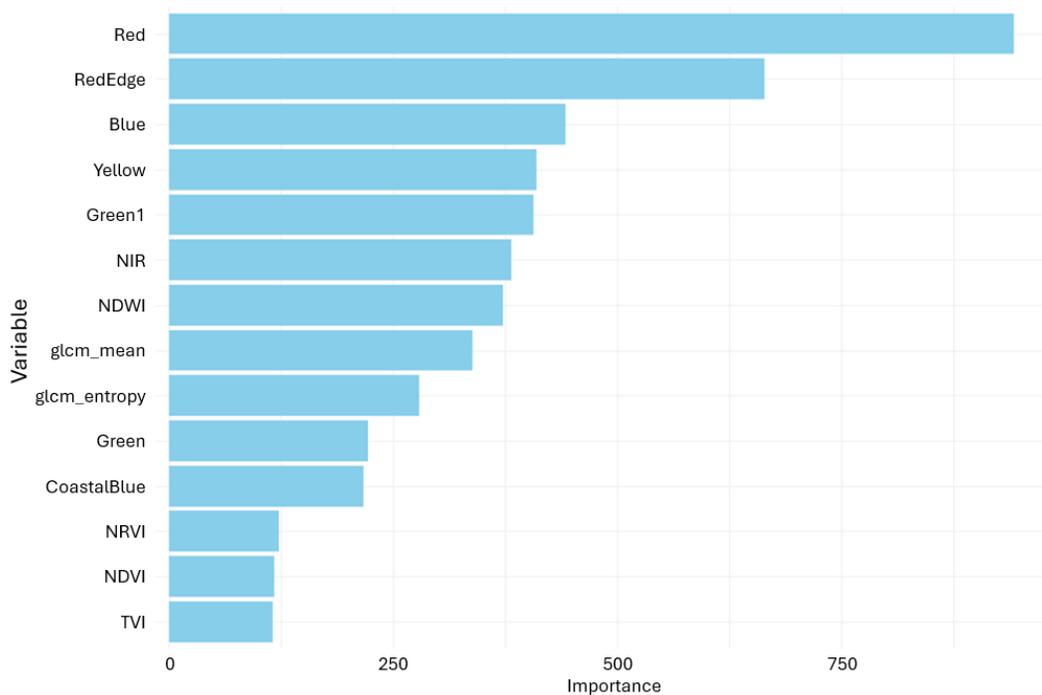


FIGURE 11 – Graphique de l'importance des variables dans le modèle entraîné sur KINY_03 (images Planet-Scope).

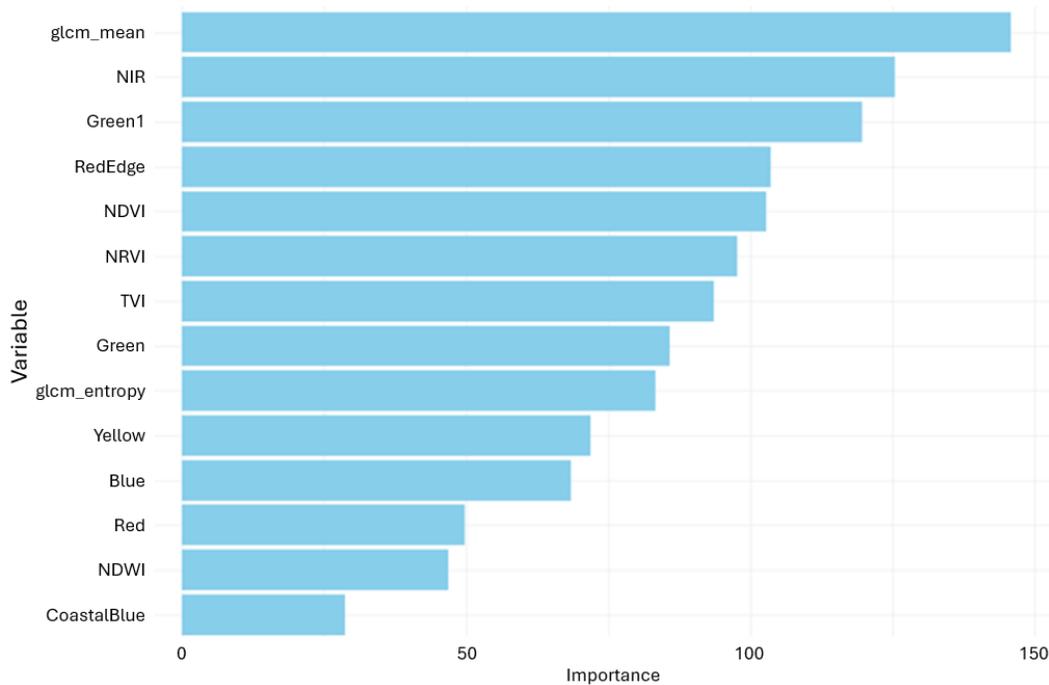


FIGURE 12 – Graphique de l'importance des variables dans le modèle entraîné sur NYAM_01 (images Planet-Scope).

TABLE 17 – Performances des modèles appliqués à d'autres bassins versants (images PlanetScope).

Bassin d'entraînement	GIKO_02	MUHU_03	NTAH_05
Bassin d'évaluation	MUHU_03	NTAH_05	GIKO_02
Précision	0.34	0.19	0.43
Kappa	0.23	0.05	0.29

Ces différences d'importance au sein des variables présage d'une difficulté des modèles à prédire les occupations de sol sur des bassins versants différents de ceux dont sont issues les données d'entraînement. Le Tableau 17 présente les performances de certains de ces modèles à prédire l'occupation du sol sur d'autres bassins versants au sein de la commune de Isale. Ces modèles ne permettent pas de prédictions fiables sur d'autres zones géographiques.

Les Figure 13 et Figure 14 présentent certaines des cartes obtenues en sortie directe des modèles.

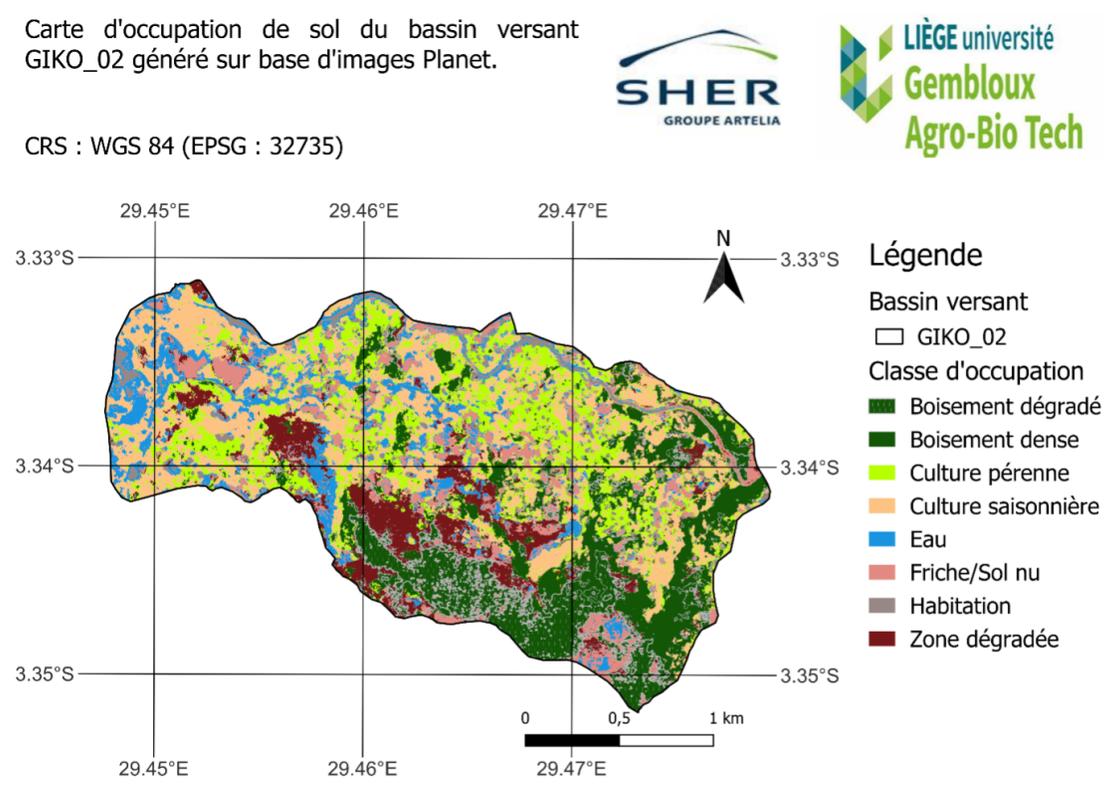


FIGURE 13 – Carte d'occupation du sol du bassin versant Giko_02 générée grâce au modèle entraîné sur des images Planet du même bassin.

Carte d'occupation de sol du bassin versant NYAM_01 généré sur base d'images Planet.



CRS : WGS 84 (EPSG : 32735)

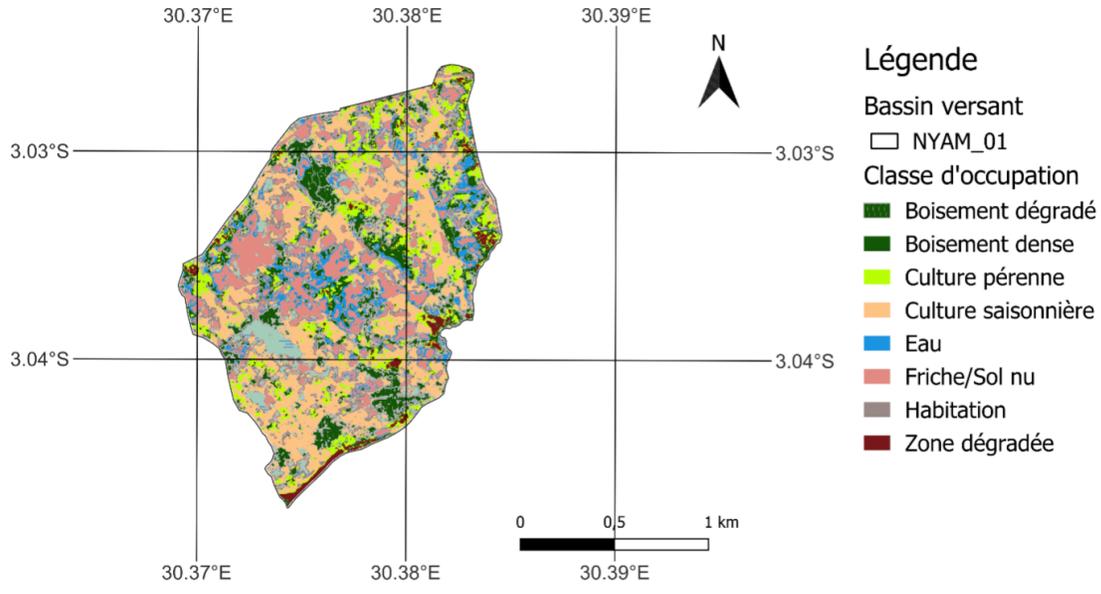


FIGURE 14 – Carte d'occupation du sol du bassin versant Nyam_01 générée grâce au modèle entraîné sur des images Planet du même bassin.

3.1.1.2. Sentinel-2

Les précisions et kappas issus de ces modèles, propres à chaque bassin versant, sont présentés sur le Tableau 18. Les modèles sont moins performants que ce qui a pu être observé préalablement sur les modèles entraînés sur des images PlanetScope. Les précisions et kappas des différents modèles sont du même ordre de grandeur malgré des différences entre les nombres de pixels par classe dans les sets d'entraînement. Il semble donc que cette baisse de la performance des modèles ne soit pas due à ce facteur.

TABLE 18 – Précision, kappa et nombre de pixels par classe dans le set d'entraînement d'images Sentinel-2 des modèles par bassin versant.

Bassins	GIKO_02	MUHU_03	NTAH_05
Précision globale	0.65	0.68	0.64
Kappa	0.60	0.63	0.59
Pixels par classe dans le set d'entraînement	155	112	83

Le résumé statistique du modèle relatif aux données Sentinel-2 sur le bassin versant GIKO_02 présente la difficulté que celui-ci a à identifier certaines classes. La sensibilité des classes Culture pérenne, Culture saisonnière et Habitation sont inférieures à 0,6. La classe Habitation est celle dont la sensibilité est la plus faible (0,486). Aucune sensibilité n'est supérieure à 0,8. Les spécificités sont-elles toutes supérieures à 0,9. Le modèle est donc plus performant

lorsqu'il s'agit de discriminer que de classer. Le résumé statistique est présenté sur le Tableau 19.

TABLE 19 – Tableau statistique du modèle entraîné sur le bassin GIKO_02 (images Sentinel-2).

Classe	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
Sensibilité	0,798	0,759	0,574	0,559	0,619	0,688	0,486	0,720
Spécificité	0,965	0,960	0,950	0,953	0,934	0,967	0,928	0,944
Valeur prédictive positive	0,767	0,729	0,621	0,631	0,571	0,751	0,489	0,647
Valeur prédictive négative	0,971	0,965	0,940	0,938	0,945	0,956	0,927	0,959
Prévalence	0,125	0,125	0,125	0,125	0,125	0,125	0,125	0,125
Taux de détection	0,100	0,095	0,072	0,070	0,077	0,086	0,061	0,090
Prévalence de détection	0,130	0,130	0,116	0,111	0,135	0,115	0,124	0,139
Exactitude équilibrée	0,882	0,859	0,762	0,756	0,777	0,828	0,707	0,832

Le résumé statistique du modèle relatif au bassin versant MUHU_03 permet d'identifier des classes pour lequel celui-ci est moins performant. Les sensibilités des classes Culture saisonnière, Friche et Habitation sont proches de 0,6. Aucune classe n'a de sensibilité supérieure à 0,81. Les spécificités des différentes classes sont elles toutes supérieures à 0,9. Cette différence entre sensibilité et spécificité s'observe sur la matrice de confusion qui présente une asymétrie. Le résumé statistique est présenté aux Tableau 20 et Tableau 21.

TABLE 20 – Matrice de confusion du modèle entraîné sur le bassin MUHU_03 (images Sentinel-2).

Référence	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
Prédiction BDEG	235	63	11	0	0	26	20	14
Prédiction BDEN	31	236	24	21	0	12	0	0
Prédiction CP	15	8	227	44	8	28	18	0
Prédiction CS	0	14	45	218	7	44	26	13
Prédiction EAU	1	0	5	15	271	12	41	16
Prédiction FR	30	15	15	27	12	187	14	28
Prédiction HA	18	0	9	11	36	10	204	20
Prédiction ZD	6	0	0	0	2	17	13	245

TABLE 21 – Tableau statistique du modèle entraîné sur le bassin MUHU_03 (images Sentinel-2).

Classe	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
Sensibilité	0,699	0,702	0,676	0,649	0,807	0,557	0,607	0,729
Spécificité	0,943	0,963	0,949	0,937	0,962	0,940	0,956	0,984
Valeur prédictive positive	0,637	0,728	0,652	0,594	0,751	0,570	0,662	0,866
Valeur prédictive négative	0,956	0,958	0,953	0,949	0,972	0,937	0,945	0,962
Prévalence	0,125	0,125	0,125	0,125	0,125	0,125	0,125	0,125
Taux de détection	0,087	0,088	0,084	0,081	0,101	0,070	0,076	0,091
Prévalence de détection	0,137	0,121	0,129	0,137	0,134	0,122	0,115	0,105
Exactitude équilibrée	0,821	0,833	0,812	0,793	0,884	0,748	0,781	0,857

Pour le modèle issu du bassin versant NTAH_05, des sensibilités inférieures à 0,6 sont observées pour les classes Culture pérenne, Friche, Habitation et Zone dégradée. Comme vu précédemment, les spécificités sont supérieures à 0,9. Le Tableau 22 présente le résumé statistique relatif à ce modèle.

TABLE 22 – Tableau statistique du modèle entraîné sur le bassin NTAH_05 (images Sentinel-2).

Classe	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
Sensibilité	0,827	0,751	0,506	0,651	0,695	0,546	0,546	0,590
Spécificité	0,956	0,966	0,947	0,950	0,962	0,940	0,937	0,930
Valeur prédictive positive	0,728	0,757	0,578	0,648	0,724	0,564	0,555	0,547
Valeur prédictive négative	0,975	0,964	0,931	0,950	0,972	0,935	0,935	0,941
Prévalence	0,125	0,125	0,125	0,125	0,125	0,125	0,125	0,125
Taux de détection	0,103	0,094	0,063	0,081	0,087	0,068	0,068	0,074
Prévalence de détection	0,142	0,124	0,109	0,126	0,120	0,121	0,123	0,135
Exactitude équilibrée	0,892	0,858	0,727	0,800	0,828	0,743	0,742	0,760

Une performance plus faible des modèles est donc observée, principalement pour les classes : Culture saisonnière, Culture pérenne et Habitation.

L'ordre d'importance des variables dans les différents modèles semble plus uniforme que pour les images Planet. Les graphiques d'importance des variables sont présentés sur les Figure 15, Figure 16 et Figure 17.

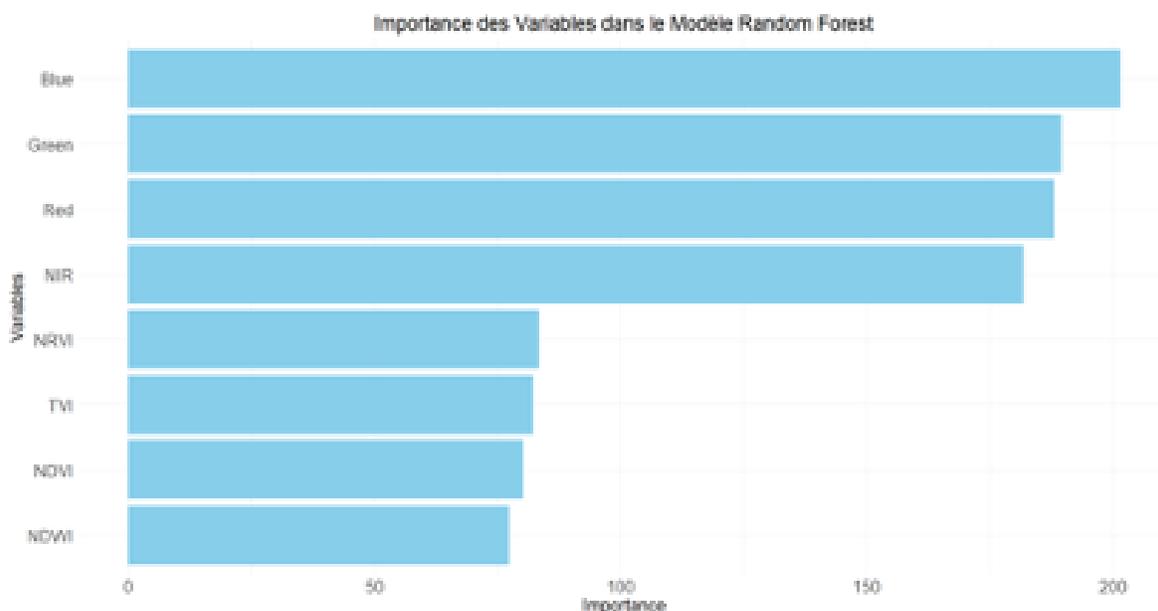


FIGURE 15 – Graphiques de l'importance des variables du modèle entraîné sur des images Sentinel-2 sur le bassin versant GIKO_02.

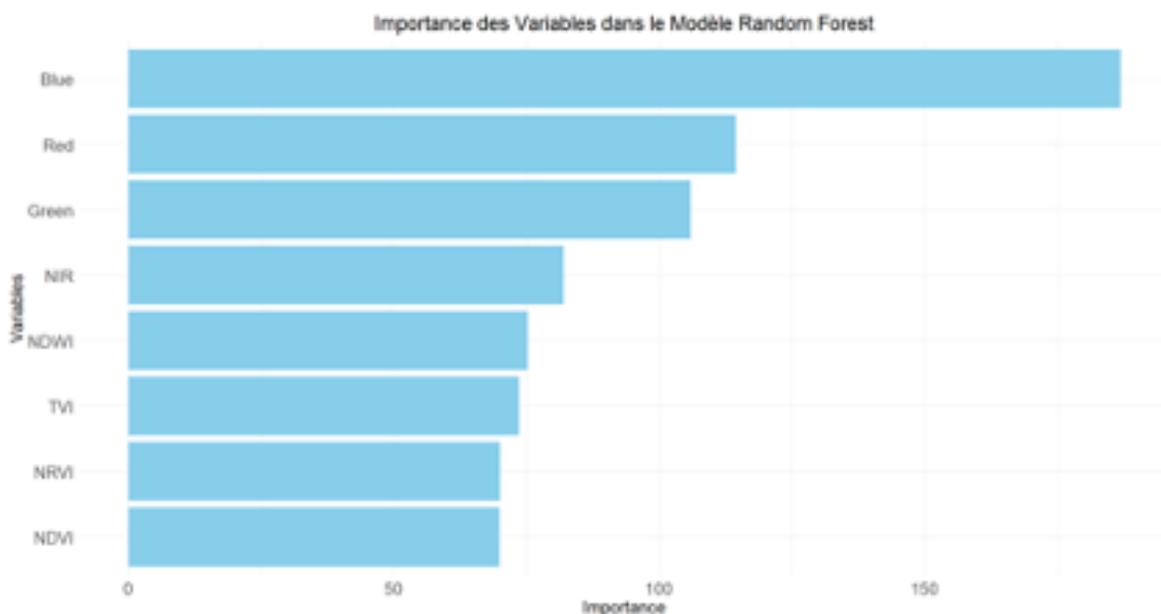


FIGURE 16 – Graphiques de l'importance des variables du modèle entraîné sur des images Sentinel-2 sur le bassin versant MUHU_03.

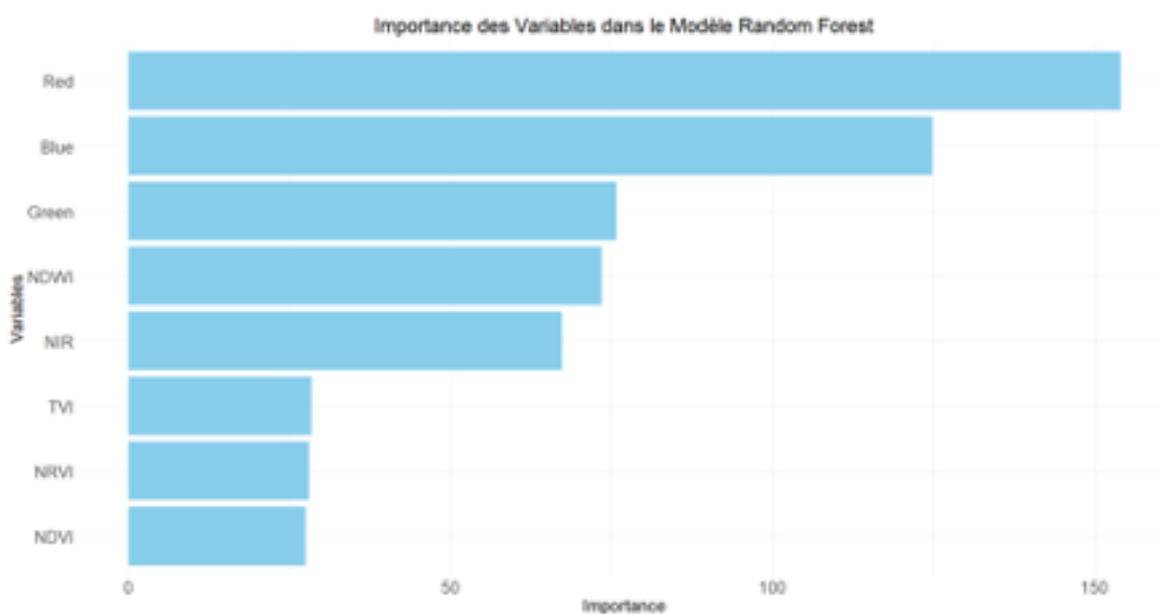


FIGURE 17 – Graphiques de l'importance des variables du modèle entraîné sur des images Sentinel-2 sur le bassin versant NTAH_05.

Le Tableau 23 reprend les précisions et kappas issus de l'évaluation de la capacité des modèles à prédire l'occupation de sol sur certains bassins versants. Ces modèles ne présentent pas de bonnes performances lorsqu'ils sont évalués sur les polygones de photointerprétation d'autres bassins versants, comme c'était le cas avec les images Planet.

TABLE 23 – Performances des modèles appliqués à d'autres bassins versants (images Sentinel-2).

Bassin d'entraînement	GIKO_02	MUHU_03	NTAH_05
Bassin d'évaluation	MUHU_03	NTAH_05	GIKO_02
Précision	0.28	0.26	0.36
Kappa	0.17	0.11	0.25

Certaines cartes obtenues lors de la prédiction sur l'ensemble du bassin versant d'entraînement du modèle sont présentées sur les Figures 18 et 19 ci-dessous.

Carte d'occupation de sol du bassin versant GIKO_02 généré sur base d'images Sentinel.



CRS : WGS 84 (EPSG : 32735)

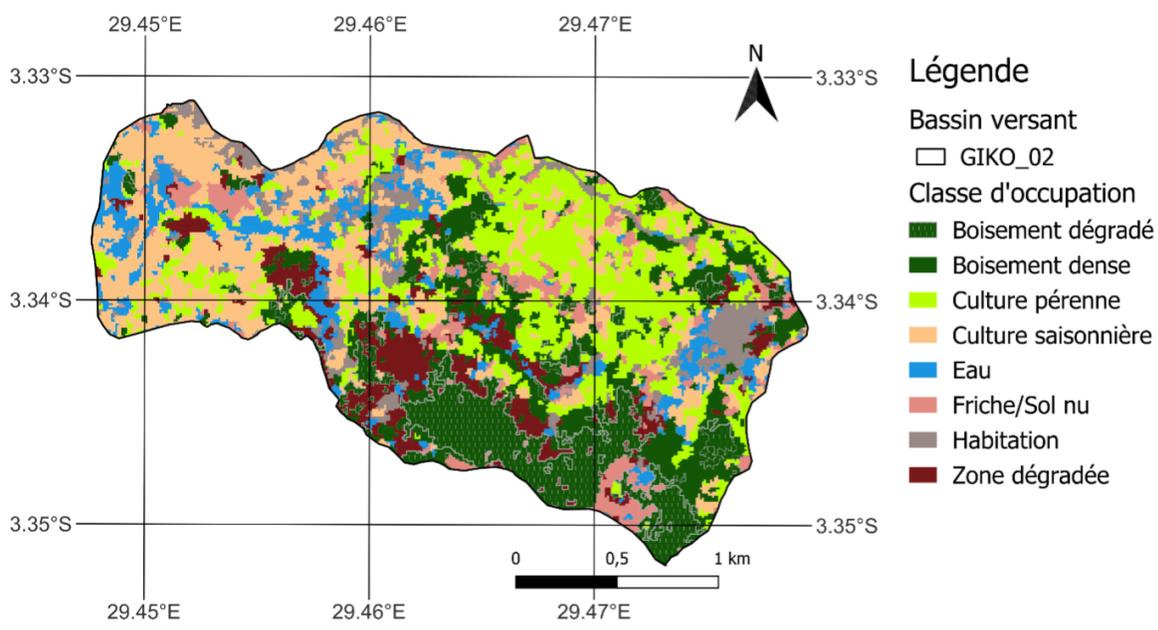


FIGURE 18 – Carte d'occupation du sol du bassin versant GIKO_02 générée grâce au modèle entraîné sur des images Sentinel du même bassin.

Carte d'occupation de sol du bassin versant NTAH_05 généré sur base d'images Sentinel.



CRS : WGS 84 (EPSG : 32735)

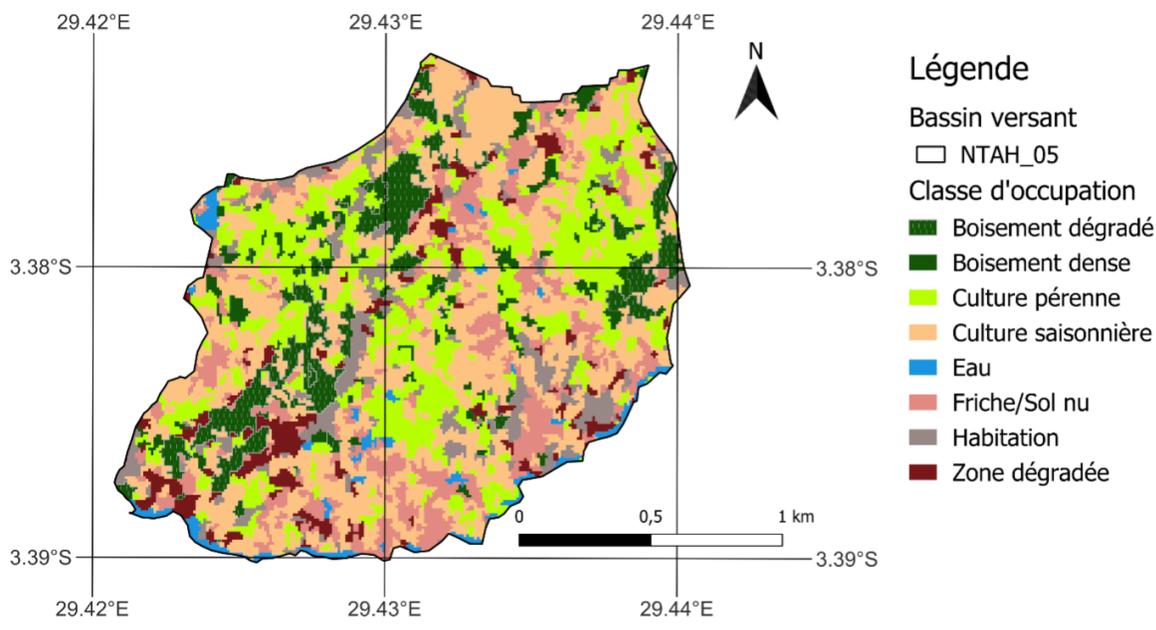


FIGURE 19 – Carte d'occupation du sol du bassin versant NTAH_05 générée grâce au modèle entraîné sur des images Sentinel du même bassin.

3.1.1.3. Drone

Une seule classification a pu être réalisée sur les données drone sur le seul bassin versant GIKO_02. Le modèle développé est basé sur l'utilisation des bandes spectrales acquises par drone et des indices de végétation calculés préalablement. La texture n'a pas été utilisée car les calculs requis n'ont pas pu être réalisés, faute de temps.

La précision et le kappa du modèle développé valent respectivement 0,673 et 0,626. Le Tableau 24 présente le résumé statistique obtenu lors de l'entraînement de ce modèle. Des sensibilités particulièrement faibles sont observées pour les classes BDEG, BDEN et CS.

TABLE 24 – Résumé statistique du modèle entraîné sur le bassin versant GIKO_02 (images drones)

Classe	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
Sensibilité	0,426	0,488	0,706	0,512	0,835	0,907	0,858	0,649
Spécificité	0,939	0,925	0,954	0,922	0,977	0,979	0,989	0,942
Valeur prédictive positive	0,498	0,483	0,686	0,485	0,840	0,858	0,914	0,615
Valeur prédictive négative	0,920	0,927	0,958	0,930	0,977	0,987	0,980	0,949
Prévalence	0,125	0,125	0,125	0,125	0,125	0,125	0,125	0,125
Taux de détection	0,053	0,061	0,088	0,064	0,104	0,113	0,107	0,081
Prévalence de détection	0,107	0,126	0,129	0,132	0,124	0,132	0,117	0,132
Exactitude équilibrée	0,683	0,707	0,830	0,717	0,906	0,943	0,923	0,796

3.1.2 Évaluation des modèles entraînés sur plusieurs bassins versants

Comme précisé en fin du paragraphe 2.2.6., des modèles entraînés sur les données issues de différents bassins versants ont été évalués. Pour rappel, les différentes combinaisons ainsi que les noms utilisés pour discuter des modèles sont présentés dans le Tableau 8.

3.1.2.1. PlanetScope

Les précisions globales et kappas obtenus lors de l'entraînement et de l'évaluation des différents modèles sur les images PlanetScope sont présentés dans le Tableau 25.

TABLE 25 – Résumé statistique de l'entraînement et de l'évaluation des modèles (images PlanetScope).

Modèle	Précision (En- traînement)	Kappa (En- traînement)	Pixels Classe (Entraîne- ment)	Bassin d'évaluation	Précision (Évaluation)	Kappa (Évaluation)
Modèle 1	0,900	0,886	1548	NTAH_05	0,208	0,071
Modèle 2	0,897	0,883	1348	GIKO_02	0,384	0,270
Modèle 3	0,898	0,883	1348	MUHU_03	0,384	0,292
Modèle 4	0,892	0,876	2352	ISALE	0,6437	0,365
Modèle 5	0,801	0,772	777	ISALE	0,623	0,532

En premier lieu, on constate que les modèles affichent tous une précision et un kappa relatifs à l'entraînement supérieurs à 0,8 hormis pour le dernier modèle dont le kappa vaut 0,77. Ensuite, des précisions et kappas largement inférieurs à 0,8 sont observés suite à l'évaluation des modèles sur des polygones indépendants. Les valeurs minimales sont issues du modèle 1 tandis que le modèle 4 et le modèle 5 affichent les valeurs de précision et de kappa maximales.

TABLE 26 – Tableau statistique obtenu à l'issue de l'entraînement du modèle 1 (images PlanetScope).

Classe	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
Sensibilité	0,943	0,941	0,862	0,880	0,920	0,899	0,857	0,902
Spécificité	0,994	0,987	0,978	0,984	0,982	0,981	0,987	0,995
Valeur prédictive positive	0,956	0,910	0,848	0,886	0,879	0,868	0,904	0,960
Valeur prédictive négative	0,992	0,992	0,980	0,983	0,989	0,986	0,980	0,986
Prévalence	0,125	0,125	0,125	0,125	0,125	0,125	0,125	0,125
Taux de détection	0,118	0,118	0,108	0,110	0,115	0,112	0,107	0,113
Prévalence de détection	0,123	0,129	0,127	0,124	0,131	0,129	0,119	0,118
Exactitude équilibrée	0,968	0,964	0,920	0,932	0,951	0,940	0,922	0,948

Le Tableau 26 présente le résumé statistique de l'entraînement du modèle 1. Le modèle 1, entraîné sur les bassins GIKO_02 et MUHU_03, présente des sensibilités légèrement inférieures aux autres pour les classes Culture pérenne, Culture saisonnière, Friche et Habitation.

Ces valeurs restent cependant supérieures à 0,85.

Les statistiques relatives à l'évaluation de la capacité du modèle à prédire l'occupation du sol sur le bassin NTAH sont présentées sur les Tableaux 27 et 28. Premièrement, il est constaté que, hormis pour la classe Habitation, les sensibilités sont toutes presque nulles ou nulles. Pour la classe Habitation, celle-ci est élevée tandis que cette fois-ci, la spécificité est faible. Les classes à faible sensibilité sont confondues avec la classe Habitation. Ce phénomène peut être observé sur la matrice de confusion qui présente une asymétrie flagrante.

TABLE 27 – Matrice de confusion obtenue à l'issue de l'évaluation du modèle 1 (images PlanetScope).

Prédiction Référence	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
BDEG	68	0	0	1	0	0	0	0
BDEN	0	0	0	0	0	0	0	0
CP	38	6	4	0	0	0	0	0
CS	163	430	668	466	0	84	0	0
EAU	8	6	0	8	15	49	1	62
FR	21	0	0	2	0	0	0	0
HA	85	196	451	1100	743	1251	1515	793
ZD	645	18	0	353	390	334	4	7

TABLE 28 – Tableau statistique obtenu à l'issue de l'évaluation du modèle 1 (images PlanetScope).

Classe	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
Sensibilité	0,066	0,000	0,004	0,241	0,013	0,000	0,997	0,008
Spécificité	1,000	1,000	0,995	0,833	0,985	0,997	0,454	0,809
Valeur prédictive positive	0,986	NaN	0,083	0,257	0,101	0,000	0,247	0,004
Valeur prédictive négative	0,903	0,934	0,887	0,821	0,885	0,828	0,999	0,896
Prévalence	0,103	0,066	0,112	0,193	0,115	0,172	0,152	0,086
Taux de détection	0,007	0,000	0,000	0,047	0,002	0,000	0,152	0,001
Prévalence de détection	0,007	0,000	0,005	0,181	0,015	0,002	0,614	0,175
Exactitude équilibrée	0,533	0,500	0,499	0,537	0,499	0,499	0,726	0,408

Les Tableaux 29, Tableau 30 et Tableau 31 présentent les résumés statistiques de l'entraînement et de l'évaluation du modèle 4. Pour rappel, ce modèle est entraîné sur les couches de polygones GIKO2, MUHU_03 et NTAH_05 et évalué sur les polygones indépendants de la couche de Isale.

TABLE 29 – Tableau statistique obtenu à l’issue de l’entraînement du modèle 4 (images PlanetScope).

Classe	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
Sensibilité	0,942	0,936	0,892	0,902	0,893	0,844	0,858	0,867
Spécificité	0,992	0,988	0,983	0,984	0,979	0,976	0,980	0,993
Valeur prédictive positive	0,946	0,918	0,882	0,892	0,861	0,835	0,861	0,945
Valeur prédictive négative	0,992	0,991	0,985	0,986	0,985	0,978	0,980	0,981
Prévalence	0,125	0,125	0,125	0,125	0,125	0,125	0,125	0,125
Taux de détection	0,118	0,117	0,112	0,113	0,112	0,105	0,107	0,108
Prévalence de détection	0,125	0,128	0,127	0,126	0,130	0,126	0,125	0,115
Exactitude équilibrée	0,967	0,962	0,938	0,943	0,936	0,910	0,919	0,930

TABLE 30 – Matrice de confusion obtenue à l’issue de l’évaluation du modèle 4 (images PlanetScope).

Prédiction Référence	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
BDEG	54	42	9	3	0	0	0	67
BDEN	5	155	246	97	0	0	0	0
CP	0	101	643	29	4	1	26	0
CS	0	4	3	277	0	8	0	0
EAU	0	0	0	0	73	41	0	0
FR	0	0	0	106	11	172	0	6
HA	0	0	17	35	17	11	76	0
ZD	3	0	0	2	0	2	0	169

TABLE 31 – Tableau statistique obtenu à l’issue de l’évaluation du modèle 4 (images PlanetScope).

Classe	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
Sensibilité	0,871	0,513	0,700	0,505	0,695	0,732	0,745	0,698
Spécificité	0,951	0,843	0,899	0,992	0,983	0,946	0,967	0,997
Valeur prédictive positive	0,309	0,308	0,800	0,949	0,640	0,583	0,487	0,960
Valeur prédictive négative	0,997	0,927	0,839	0,878	0,987	0,972	0,989	0,969
Prévalence	0,025	0,120	0,365	0,218	0,042	0,093	0,041	0,096
Taux de détection	0,021	0,062	0,256	0,110	0,029	0,068	0,030	0,067
Prévalence de détection	0,070	0,200	0,320	0,116	0,045	0,117	0,062	0,070
Exactitude équilibrée	0,911	0,678	0,800	0,749	0,839	0,839	0,856	0,848

Pour ce qui est de l’entraînement, il est constaté que les classes Friche, Habitation et Zone dégradée ont une sensibilité plus faible que celles des autres classes. Elles sont cependant supérieures à 0,80. Le modèle identifie donc moins de vrais positifs pour ces classes. Les valeurs prédictives positives des classes Culture pérenne, Culture saisonnière, Eau, Friche et Habitation sont inférieures à 0,90. La proportion de vrais positifs parmi les prédictions positives est donc plus faible pour ces classes.

Au niveau de l’évaluation, une asymétrie dans la matrice de confusion est à nouveau remarquée bien qu’elle soit moins marquée que sur le modèle 1. Des valeurs de sensibilité variant entre 0,5 et 0,75 sont observées hormis pour la classe Boisement dégradé. Celle-ci est ce-

pendant caractérisée par une valeur prédictive positive d'environ 0,3. Cela démontre que le modèle surestime la présence de cette classe, ce qui se marque lorsque l'on compare la prévalence à la prévalence de détection.

3.1.2.2. Sentinel-2

Les précisions et kappas obtenus suite à l'entraînement et à l'évaluation des quatre modèles sur bases d'images Sentinel-2 sont présentés sur le Tableau 32. Les performances liées à l'entraînement sont moins bonnes que celles observées lors de l'entraînement sur des images Planet. L'ordre de grandeur des statistiques relatives à l'évaluation est similaire à celui issu des modèles Planet. Une amélioration est notée lorsque les 3 bassins versants de la région sont pris en compte.

TABLE 32 – Résumé statistique de l'entraînement et de l'évaluation des modèles (images Sentinel-2).

Modèle	Précision (Entraînement)	Kappa (Entraînement)	Nombre de pixels par classe	Bassin(s) d'évaluation	Précision (Évaluation)	Kappa (Évaluation)
Modèle 1	0,659	0,611	358	NTAH_05	0,312	0,179
Modèle 2	0,641	0,590	236	GIKO_02	0,266	0,167
Modèle 3	0,639	0,587	319	MUHU_03	0,533	0,440
Modèle 4	0,640	0,588	522	ISALE	0,515	0,414

Le Tableau 33 renseigne sur les statistiques d'entraînement du modèle 1. Une performance globalement faible est constatée, certaines classes présentant des sensibilités inférieures à 0,65. Aucune d'entre elles ne présente de sensibilité supérieure à 0,8.

TABLE 33 – Tableau statistique obtenu à l'issue de l'entraînement du modèle 1 (images Sentinel-2).

Classe	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
Sensibilité	0,770	0,729	0,616	0,541	0,698	0,622	0,625	0,672
Spécificité	0,957	0,947	0,948	0,941	0,956	0,954	0,948	0,959
Valeur prédictive positive	0,720	0,665	0,630	0,567	0,693	0,658	0,631	0,700
Valeur prédictive négative	0,967	0,961	0,945	0,935	0,957	0,946	0,946	0,953
Prévalence	0,125	0,125	0,125	0,125	0,125	0,125	0,125	0,125
Taux de détection	0,096	0,091	0,077	0,068	0,087	0,078	0,078	0,084
Prévalence de détection	0,134	0,137	0,122	0,119	0,126	0,118	0,124	0,120
Exactitude équilibrée	0,864	0,838	0,782	0,741	0,827	0,788	0,786	0,816

Les résultats obtenus lors de l'évaluation du modèle 1 sont présentés dans le Tableau 34. Les sensibilités des différentes classes sont inférieures à 0,3 à l'exception de la classe Habitation ayant une sensibilité d'environ 0,75. Cela se fait au détriment de la spécificité de cette classe qui ne vaut que 0,65. Le même phénomène a été observé lors de l'évaluation du modèle correspondant entraîné avec les images Planet.

TABLE 34 – Tableau statistique obtenu à l’issue de l’évaluation du modèle 1 (images Sentinel-2).

Classe	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
Sensibilité	0,218	0,084	0,281	0,265	0,263	0,053	0,744	0,131
Spécificité	0,982	0,978	0,915	0,922	0,844	0,997	0,650	0,895
Valeur prédictive positive	0,500	0,171	0,271	0,410	0,201	0,789	0,375	0,092
Valeur prédictive négative	0,938	0,952	0,918	0,859	0,885	0,833	0,900	0,926
Prévalence	0,077	0,051	0,102	0,170	0,129	0,175	0,220	0,076
Taux de détection	0,017	0,004	0,029	0,045	0,034	0,009	0,164	0,010
Prévalence de détection	0,033	0,025	0,105	0,110	0,170	0,012	0,437	0,107
Exactitude équilibrée	0,600	0,531	0,598	0,594	0,554	0,525	0,697	0,513

Le résumé statistique issu de l’entraînement du second modèle est présenté au Tableau 35. Des sensibilités faibles sont à nouveau constatées pour l’ensemble des classes. Les classes Culture pérenne, Friche et Zone dégradée ont les sensibilités les plus faibles.

TABLE 35 – Tableau statistique obtenu à l’issue de l’entraînement du modèle 2 (images Sentinel-2).

Classe	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
Sensibilité	0,757	0,756	0,586	0,667	0,641	0,468	0,669	0,585
Spécificité	0,948	0,963	0,950	0,948	0,941	0,923	0,949	0,968
Valeur prédictive positive	0,675	0,747	0,627	0,645	0,609	0,464	0,652	0,721
Valeur prédictive négative	0,965	0,965	0,941	0,952	0,948	0,924	0,953	0,942
Prévalence	0,125	0,125	0,125	0,125	0,125	0,125	0,125	0,125
Taux de détection	0,095	0,094	0,073	0,083	0,080	0,058	0,084	0,073
Prévalence de détection	0,140	0,126	0,117	0,129	0,132	0,126	0,128	0,101
Exactitude équilibrée	0,853	0,860	0,768	0,807	0,791	0,695	0,809	0,776

Les résultats obtenus lors de l’évaluation du modèle sur le bassin versant GIKO_02 sont présentés sur les Tableau 36 et Tableau 37. Aucune sensibilité n’est supérieure à 0,32 si ce n’est pour la classe Culture Pérenne qui présentait pourtant une sensibilité faible lors de l’entraînement du modèle. Une confusion entre les classes Boisement dégradé et Friche, observable sur la matrice de confusion, cause une baisse de la spécificité de cette dernière classe. Une confusion importante est observée entre les classes BDEG et FR, CP et CS, EAU et ZD.

TABLE 36 – Matrice de confusion obtenue à l'issue de l'évaluation du modèle 2 (images Sentinel-2).

Référence	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
BDEG (Prédiction)	199	78	9	27	0	5	10	49
BDEN (Prédiction)	6	51	15	19	2	1	17	2
CP (Prédiction)	11	89	83	177	15	5	45	34
CS (Prédiction)	1	1	11	130	25	17	21	9
EAU (Prédiction)	32	31	7	13	41	65	12	129
FR (Prédiction)	352	58	19	64	56	83	40	59
HA (Prédiction)	0	16	9	53	21	7	68	44
ZD (Prédiction)	71	3	2	7	9	76	10	56

TABLE 37 – Tableau statistique obtenu à l'issue de l'évaluation du modèle 2 (images Sentinel-2).

Classe	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
Sensibilité	0,296	0,156	0,536	0,265	0,243	0,320	0,305	0,147
Spécificité	0,911	0,974	0,851	0,961	0,885	0,732	0,939	0,922
Valeur prédictive positive	0,528	0,451	0,605	0,605	0,124	0,114	0,312	0,239
Valeur prédictive négative	0,794	0,892	0,854	0,945	0,910	0,937	0,867	0,867
Prévalence	0,251	0,122	0,181	0,122	0,063	0,097	0,083	0,143
Taux de détection	0,074	0,019	0,083	0,049	0,015	0,031	0,025	0,021
Prévalence de détection	0,141	0,042	0,172	0,080	0,123	0,273	0,081	0,087
Exactitude équilibrée	0,604	0,565	0,693	0,613	0,564	0,526	0,622	0,535

Le modèle 3, entraîné sur les bassins versants GIKO_02 et NTAH_05, présente des résultats similaires aux modèles précédents. Les sensibilités sont faibles, les spécificités élevées. Les classes Culture saisonnière, Friche et Habitation ont des sensibilités particulièrement faibles. Ces résultats sont présentés sur le Tableau 38.

TABLE 38 – Tableau statistique obtenu à l'issue de l'entraînement du modèle 3 (images Sentinel-2).

Classe	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
Sensibilité	0,778	0,735	0,611	0,568	0,697	0,520	0,574	0,627
Spécificité	0,962	0,957	0,941	0,949	0,944	0,952	0,941	0,941
Valeur prédictive positive	0,744	0,712	0,596	0,616	0,689	0,609	0,583	0,602
Valeur prédictive négative	0,944	0,939	0,956	0,948	0,933	0,939	0,946	0,934
Prévalence	0,125	0,125	0,125	0,125	0,125	0,125	0,125	0,125
Taux de détection	0,097	0,092	0,076	0,071	0,087	0,065	0,072	0,078
Prévalence de détection	0,131	0,129	0,115	0,115	0,136	0,107	0,123	0,130
Exactitude équilibrée	0,870	0,846	0,776	0,759	0,820	0,736	0,758	0,784

Le Tableau 39 présente les indicateurs statistiques issus de l'évaluation de la capacité du modèle 3 à prédire l'occupation sur le bassin MUHU_03. Seules les classes Boisement dégradé et Zone dégradée ont une sensibilité supérieure à 0,6. Toutes les classes ont une spécificité supérieure à 0,9.

TABLE 39 – Tableau statistique obtenu à l’issue de l’évaluation du modèle 3 (images Sentinel-2).

Classe	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
Sensibilité	0,750	0,455	0,560	0,424	0,600	0,425	0,577	0,800
Spécificité	0,948	0,827	0,901	0,958	0,965	0,919	0,976	0,962
Valeur prédictive positive	0,310	0,247	0,743	0,735	0,536	0,370	0,625	0,632
Valeur prédictive négative	0,992	0,924	0,800	0,859	0,973	0,934	0,970	0,983
Prévalence	0,030	0,111	0,338	0,215	0,063	0,101	0,066	0,076
Taux de détection	0,023	0,051	0,189	0,091	0,038	0,043	0,038	0,061
Prévalence de détection	0,073	0,205	0,255	0,124	0,071	0,116	0,061	0,096
Exactitude équilibrée	0,849	0,641	0,730	0,691	0,782	0,672	0,776	0,881

Le Tableau 40 renseigne sur les performances à l’entraînement du modèle 4, portant sur les trois bassins versants de la région de Isale. À nouveau, les sensibilités de toutes les classes sont faibles tandis que leurs spécificités sont élevées. De meilleurs performances sont observées pour les classes Boisement dégradé et Boisement dense, sans toutefois présenter de sensibilité supérieure à 0,8.

TABLE 40 – Tableau statistique obtenu à l’issue de l’entraînement du modèle 4 (images Sentinel-2).

Classe	BDEG	BDEN	CP	CS	EAU	FR	HA	ZD
Sensibilité	0,77075	0,72989	0,58046	0,58748	0,68455	0,54662	0,60856	0,6092
Spécificité	0,95868	0,9491	0,94846	0,93687	0,93934	0,94837	0,94636	0,96105
Valeur prédictive positive	0,72711	0,67196	0,61669	0,57072	0,61716	0,60197	0,61843	0,6908
Valeur prédictive négative	0,96697	0,96093	0,94056	0,94082	0,95422	0,93607	0,94421	0,9451
Prévalence	0,125	0,125	0,125	0,125	0,125	0,125	0,125	0,125
Taux de détection	0,09634	0,09124	0,07256	0,07344	0,08557	0,06833	0,07607	0,07615
Prévalence de détection	0,1325	0,13578	0,11766	0,12867	0,13865	0,11351	0,123	0,11023
Exactitude équilibrée	0,86471	0,83949	0,76446	0,76218	0,81194	0,74749	0,77746	0,78512

3.2 Méthode de classification mise au point

Nous avons vu que les modèles créés sur un bassin versant ne sont pas applicables à d’autres bassins versants. Il est donc nécessaire, pour l’instant, d’acquérir des images par drone sur chaque bassin versant d’intérêt. Afin de comparer la méthode mise au point avec celle actuellement utilisée par l’entreprise, une estimation des besoins à chaque étape de la nouvelle méthode est nécessaire. Ces étapes peuvent être réalisées par un seul opérateur, certaines simultanément.

La création des plans de vols via l’outil QGIS a demandé un jour de travail pour l’ensemble des bassins versants, c’est à dire une surface d’environ 2100 hectares. Cette étape est décrite en Annexe 6.1.

L'acquisition d'images satellites demande très peu de temps. L'acquisition d'images par drone est par contre elle chronophage. Le rapport « surface couverte par jour » dépend cependant du relief de la zone et du nombre de batteries disponibles. Lors de l'échantillonnage réalisé au Burundi, il aura fallu 2 heures de vol au total pour couvrir les 280 hectares du bassin versant NTAH (pente moyenne de 27%) tandis que les 235 hectares de NYAM (pente moyenne de 10%) ont été couverts en 1 heure. En moyenne, ce sont 200 hectares par heure de vol qui ont pu être couverts. Ce chiffre ne prend pas en compte le temps nécessaire à l'accès à la zone. Le nombre de batterie a constitué un facteur limitant lors de l'échantillonnage sur le terrain. En moyenne, une batterie permettait de couvrir 100 hectares. Disposant de 7 batteries, une acquisition théorique de 700 hectares par jour semble possible.

Ce chiffre ne reflète cependant pas la réalité. Du fait de la superposition des plans de vol, la surface nécessaire à l'échantillonnage d'un bassin versant est supérieure à la surface réelle du bassin. Par exemple, le bassin NTAH de 280 hectares a été couvert par trois vols pour une surface totale de 335 hectares. Le nombre de plans de vol nécessaires pour couvrir un bassin versant dépend à nouveau du relief. Le bassin NYAM, moins accidenté, de 230 hectares a pu être couvert en un seul vol. Plusieurs plans de vol pour une même surface demandent aussi des atterrissages et décollages plus fréquents, ces éléments participant à une consommation accrue.

Les conditions météorologiques ont également une incidence sur le temps d'acquisition. Si elles le permettent, ce sont environ 450 hectares par jour qui peuvent être couverts grâce aux sept batteries, en environ deux heures et quart. Le nombre de batteries ne permet pas d'augmenter la surface couverte. En prenant en compte les trajets jusqu'aux zones de vol, une demi-journée permet de traiter cette surface. L'autre moitié de la journée est destinée à la recharge des batteries.

Les traitements photogrammétriques permettent le traitement d'environ 500 hectares par jour. Ces traitements ne demandent cependant par la présence d'un opérateur et peuvent être réalisés en parallèle de l'acquisition. L'importance de l'acquisition d'images multispectrales n'a pas encore été discutée. Dans le cas où celles-ci ne seraient pas acquises, 1000 hectares par jour pourraient être traités.

La photointerprétation a nécessité un jour de travail pour traiter 1000 hectares. C'est lors de la réalisation de cette étape qu'un biais de subjectivité peut survenir. Il est cependant aisé de fixer des règles de photointerprétation, les orthomosaïques ayant une résolution de 5 à 10 centimètres.

L'entraînement et la prédiction du modèle ne demandent eux qu'environ 5 minutes pour traiter 1000 hectares lorsque des images PlanetScope ou Sentinel-2 sont utilisées. Pour un modèle réalisé via des images drones à la résolution maximale, il faut compter environ 8 heures pour 1000 hectares. À nouveau, cette étape ne demande pas de réel investissement de la part d'un opérateur.

Cette méthode demande donc le travail d'un opérateur durant environ une journée et un quart pour traiter 450 hectares, soit 2,8 jours de travail d'un opérateur pour 1000 hectares.

L'acquisition d'images par drone génère des sous-produits pouvant être valorisés par l'entreprise. Lors des traitements de photogrammétrie un modèle numérique de surface est généré. Les MNS générés lors du projet présentaient une résolution d'environ 50 cm/pixel. Lors de la photointerprétation, seules les orthomosaïques issues du traitement des images RGB sont utilisées car celles-ci suffisent à différencier les différentes classes d'occupation du sol. Cependant, l'acquisition et le traitement d'images multispectrales ont également été réalisés.

L'ajout des prises de vue multispectrales multiplie le nombre total d'images acquises par 5, 4 bandes multispectrales étant acquises. Ainsi, pour couvrir les 335 hectares de plan de vol du bassin NTAH 05, 271 images RGB et 1084 images multispectrales ont été acquises. En terme de mémoire, les images RGB occupent en moyenne 7860 kilooctets contre 9320 pour les images multispectrales. Une mémoire totale de 2,13 Gigaoctets (Go) est occupée par les images RGB et 10,1 Go sont occupés par les images multispectrales pour ce bassin versant. L'acquisition d'images multispectrales a donc un impact considérable sur le stockage à prévoir, la quantité nécessaire variant du simple au quintuple.

Pour couvrir l'ensemble des zones d'études (2100 hectares), ce sont environ 1050 Go de stockage qui ont été nécessaires. Environ 200 de ceux-ci sont occupés par des images RGB, le reste par des images multispectrales. Il est donc estimé qu'un stockage de 0,1 Go à l'hectare est nécessaire pour l'acquisition RGB, un stockage de 0,5 Go/ha pour une acquisition RGB et multispectrale (l'acquisition multispectrale ne pouvant se faire indépendamment d'une acquisition RGB). Ces chiffres sont cependant à considérer prudemment, la surface des plans de vol nécessaire à la couverture d'un bassin versant dépendant de la topographie et de l'opérateur les traçant tandis que le nombre d'images prises par hectare dépend de l'altitude de vol et du pourcentage de recouvrement choisi.

4 Discussion

4.1 Le modèle à l'échelle du bassin versant

Les modèles entraînés sur des images Planet ne portant que sur un seul bassin versant ont tous démontré une capacité à prédire l'occupation de sol au sein de ceux-ci avec des précisions et kappas supérieures à 80%. Sur la commune de Isale, celles-ci sont proches des 90%. Les modèles constituent donc une alternative précise à la méthode mise en place par l'entreprise en terme de précision et de temps de travail par opérateur à l'hectare.

La différence du nombre de pixels par classe entre les sets d'entraînement des bassins KINY_03 (839 pixels par classe) et NYAM_01 (204 pixels par classe) n'a pas d'impact sur la qualité des prédictions. Il reste néanmoins préférable de disposer de sets de données complets et au sein desquels chaque classe est suffisamment représentée en vue de la classification par RandomForest (He & Garcia, 2009).

Dans le chapitre 3.1.1.1., l'impact de la présence de la classe Marais sur les modèles est observé. Cette classe représente une des particularités de l'occupation de sol de la région de Muyinga. Il est donc nécessaire de pouvoir la distinguer. Dans cette région au relief plus plat, les marais sont situés le long des cours d'eau, dans les anciens lits de ceux-ci. Parfois celui-ci n'en est pas distinguable. Ils sont principalement cultivés, les agriculteurs y plantant du riz. Une des caractéristiques de ces marais est donc que la pente y est faible (Figure 20).

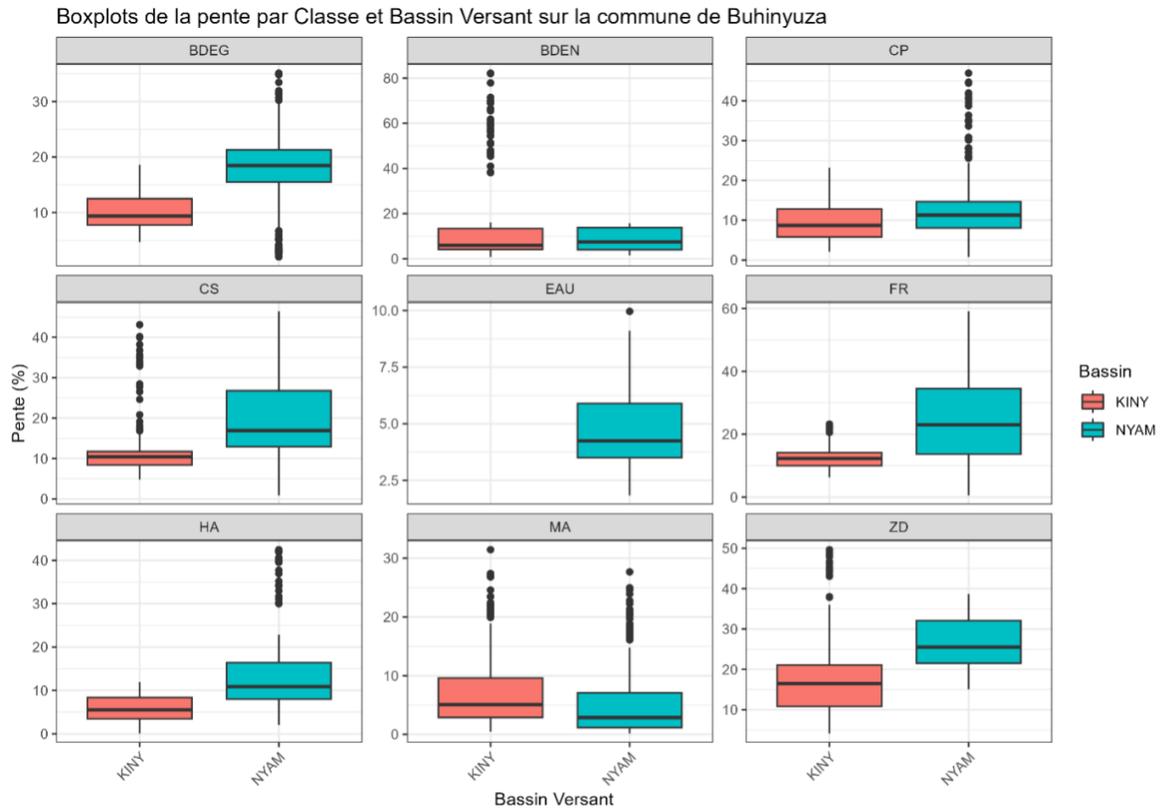


FIGURE 20 – Graphique des pentes par classe d’occupation de sol au sein de la commune de Buhinyuza.

Le graphique présenté est issu du calcul des pentes sur un MNT fourni par l’entreprise qui présente une résolution de 10 mètres par pixel. Il est constaté que, au sein de l’espace interquartile, les pentes pour la classe Marais ne dépassent pas les 10%. Les médianes sont quant à elles proches ou inférieures à 5%. À l’inverse, les pentes des autres classes étudiées ont tendance à être supérieures à 10%, les médianes n’étant pas inférieures à 5%, à l’exception de la classe Eau. Les zones d’eau et de marais se retrouvant par définition dans le même type d’endroits, ce critère ne permet pas de les distinguer.

Il semble donc envisageable d’intégrer la pente comme critère d’identification de la classe Marais au modèle de prédiction sur la commune de Buhinyuza afin de permettre une meilleure identification de cette classe. Pour ce faire, les valeurs de pente sur chaque pixel pourraient être extraites et cette variable intégrée au modèle telle quelle. Ce critère pourrait cependant amener du bruit, la pente pouvant ne pas être un critère d’identification efficace pour les autres classes. Dans ce cas, une variable d’identification binaire pourrait être mise en place, celle-ci valant 1 lorsque la pente est inférieure ou égale à 5° et 0 lorsque ce n’est pas le cas par exemple.

4.1.1 Le cas du bassin NYAM_01

Lors de l'entraînement et de l'évaluation du modèle sur le bassin versant NYAM_01, il a été constaté que la classe d'occupation de sol « Zone dégradée » était décrite par une sensibilité égale à 1 et une spécificité égale à 0,989. Comparé aux résultats obtenus par les autres classes pour ces deux indicateurs, il semble qu'un problème de surapprentissage survienne.

Après révision de l'échantillonnage effectué sur le bassin NYAM_01, il s'avère qu'un seul polygone s'est vu assigner la classe Zone dégradée. Le bassin ne couvrant que 236 hectares, une seule zone de cette classe a pu y être identifiée. Le fait qu'il n'y ait qu'un polygone de cette classe a pour conséquence que les seules données d'entraînement pour celle-ci proviennent d'une seule zone géographique. Les caractéristiques de cette zone risquent donc d'être moins variées et non transposables que celles des autres classes, étant représentées par plusieurs polygones d'origines géographiques plus variées. Un surapprentissage peut alors survenir. Le classificateur risque d'être nettement plus compétent lors de l'identification de cette classe aux caractéristiques claires, au détriment des autres classes.

La classe étant sous représentée à l'échelle du paysage du bassin versant, il n'est pas possible d'augmenter le nombre de polygones dans le set d'entraînement. Une solution envisageable pourrait être le suréchantillonnage. Cela consiste à augmenter le nombre d'observations d'une ou de plusieurs classes minoritaires. Répéter simplement les exemples n'apporterait pas de diversités à l'échantillon, le risque de surapprentissage n'en serait que plus élevé (Chen, 2004; He & Garcia, 2009; Chawla et al., 2011-06-09). Des méthodes d'interpolation permettraient de créer de nouveaux échantillons artificiels aux caractéristiques différentes de celles déjà observées. Il faudrait alors intégrer, à la suite de l'étape de photointerprétation, des méthodes telles que la SMOTE avant de réaliser la classification (Chawla et al., 2011-06-09).

4.1.2 Sentinel

Les modèles entraînés sur des images Sentinel à l'échelle du bassin versant ne sont pas aussi précis que ceux entraînés sur des images Planet. Leur précision varie entre 0,64 et 0,68. Ces performances moins bonnes sont probablement dues à la résolution des images Sentinel. Des éléments de petite taille tels que les habitations ont vu leur sensibilité fortement baisser. Il en va de même pour les classes dont les dispositions spatiales sont entremêlées, par exemple les classes Culture pérenne et saisonnière. Il apparaît que la résolution des images Sentinel est insuffisante pour capturer les valeurs spécifiques à de petits éléments ou à des éléments variant spatialement. Ces éléments sont en quelque sorte noyés dans les données.

Le nombre plus faible de pixels disponibles, dû à une résolution plus faible, ne semble pas impacter la qualité des modèles. Le modèle entraîné sur NTAH_05, avec 83 pixels par classe lors de l'entraînement, a une précision et un kappa similaires à ceux du modèle entraîné sur GIKO_02 avec 155 pixels par classe.

4.1.3 Drone

Une seule classification a été réalisée sur base de données de photogrammétrie. La précision de ce modèle n'est pas élevée. Le résumé statistique présenté au Tableau 24 démontre cependant que cette précision est faible à cause des performances du modèle sur trois classes (BDEG, BDEN et CS). Les classes CP et ZD ont-elles des sensibilités moyennes. Des bonnes performances sont observées pour les autres classes.

Lors de la mise au point des modèles sur base des images Sentinel et PlanetScope, il est apparu que l'ajout des textures au modèle permettait d'améliorer les capacités de ceux-ci à distinguer les classes constituées d'objets semblables mais aux densités variables tels que des arbres (présents dans les classes BDEG, BDEN et CP) ou des plantes herbacées (caractérisant les zones CS et ZD). Dans le cadre des projets réalisés par la SHER, la distinction entre ces classes est cependant nécessaire. L'ajout des textures au modèle améliorerait probablement ses performances lors de l'identification de ces différentes classes. Il est toutefois nécessaire de noter que le temps de calcul nécessaire à la création des rasters de texture est important lorsqu'il s'agit de traiter des données denses en informations telles que les orthomosaïques. De plus, le temps nécessaire à l'entraînement du modèle et à la prédiction sur une zone d'intérêt augmente également avec la quantité de données. L'entraînement du modèle sur le bassin GIKO_02 a demandé environ une journée.

4.2 Les modèles à l'échelle de la région

Il a été vu au chapitre 3.1.2. que des modèles entraînés sur la combinaison de bassins versants ne permettaient pas de prédire avec précision l'occupation de sol sur d'autres bassins versants, pour les deux sources de données considérées. Dans ce chapitre, une analyse des causes de ce manque de précision sera réalisée.

Lors des premiers entraînements, les images Planet ou Sentinel relatives aux différents bassins versants ont été acquises séparément. Certaines d'entre elles, utilisées à l'évaluation, n'ont pas été prises par les satellites aux mêmes dates que celles utilisées pour l'entraînement. La précision des modèles entraînés et évalués sur ces images étant faible, de nouvelles images ont été commandées, celles-ci couvrant l'ensemble des bassins d'une même commune à une

même date. Les paramètres de prise d'images sont donc les mêmes sur les bassins versants. Cette mesure a permis d'améliorer la précision des modèles mais n'a pas complètement résolu le problème.

Une analyse et une comparaison des valeurs des indices calculés sur chaque classe au sein des bassins versants a été réalisée sur les images Planet et Sentinel. Les Figure 21, Figure 22, Figure 23 et Figure 24 présentent une partie des résultats graphiques obtenus.

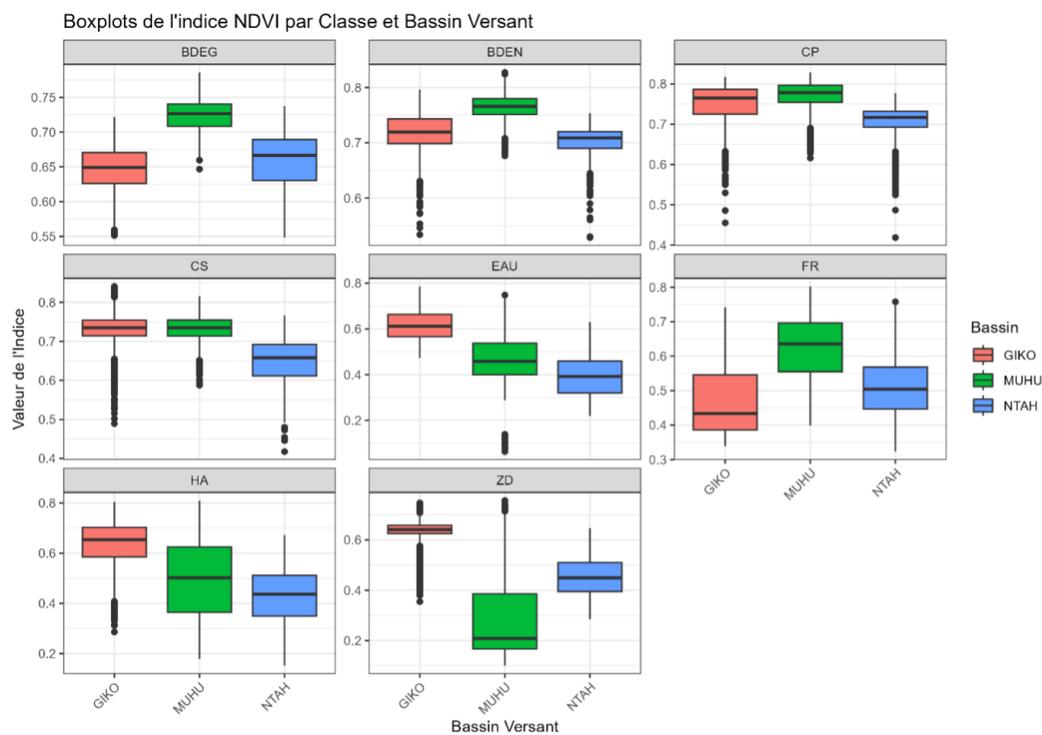


FIGURE 21 – Boxplot de l'indice NDVI pour chaque classe d'occupation de sol au sein des bassins versants de la commune de Isale (Image Planet).

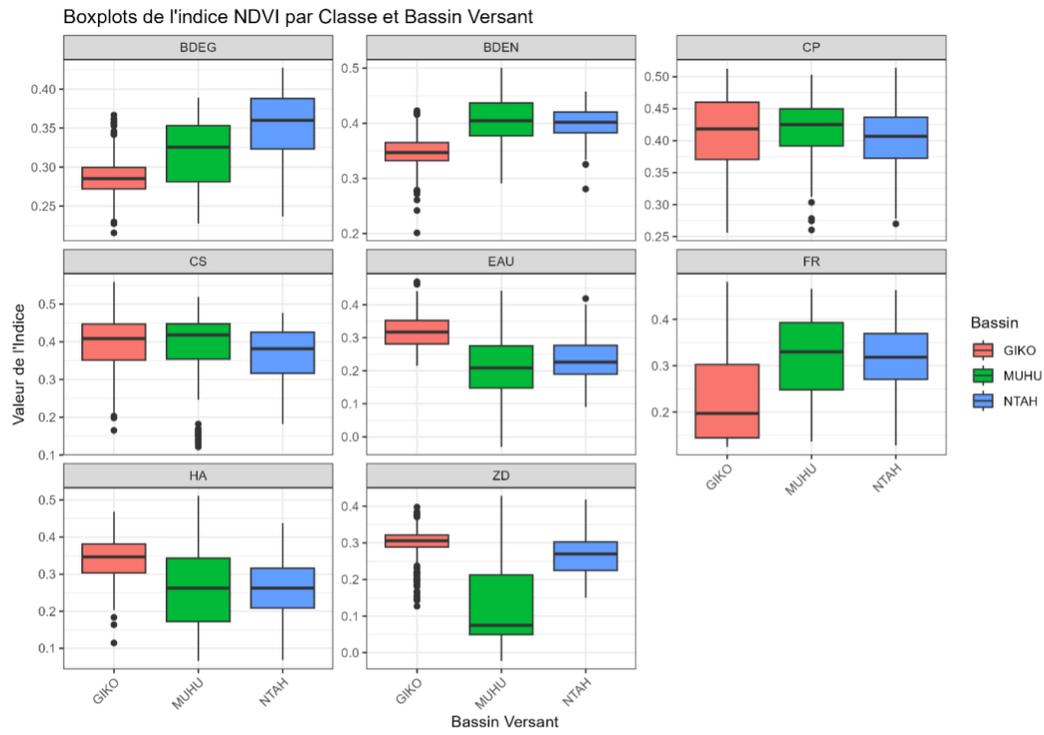


FIGURE 22 – Boxplot de l'indice NDVI pour chaque classe d'occupation de sol au sein des bassins versants de la commune de Isale (Image Sentinel).

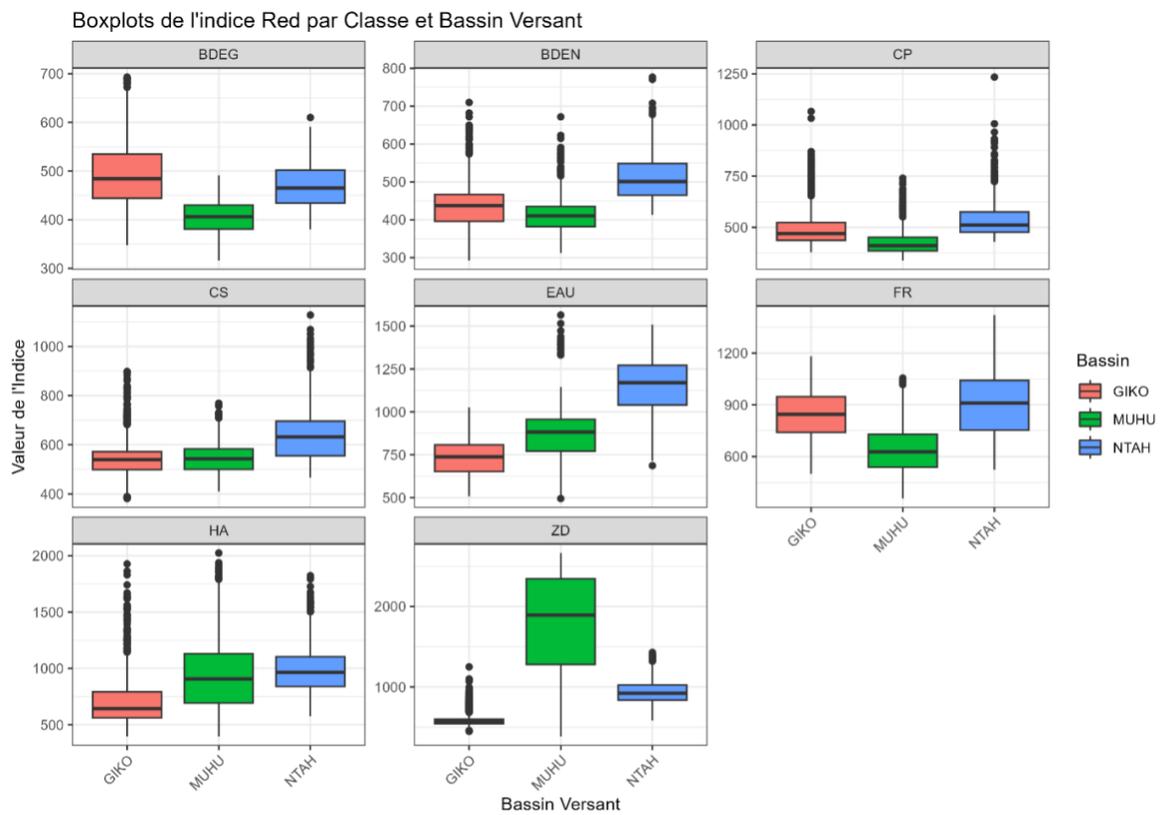


FIGURE 23 – Boxplot de la bande Red pour chaque classe d'occupation de sol au sein des bassins versants de la commune de Isale (Image Planet).

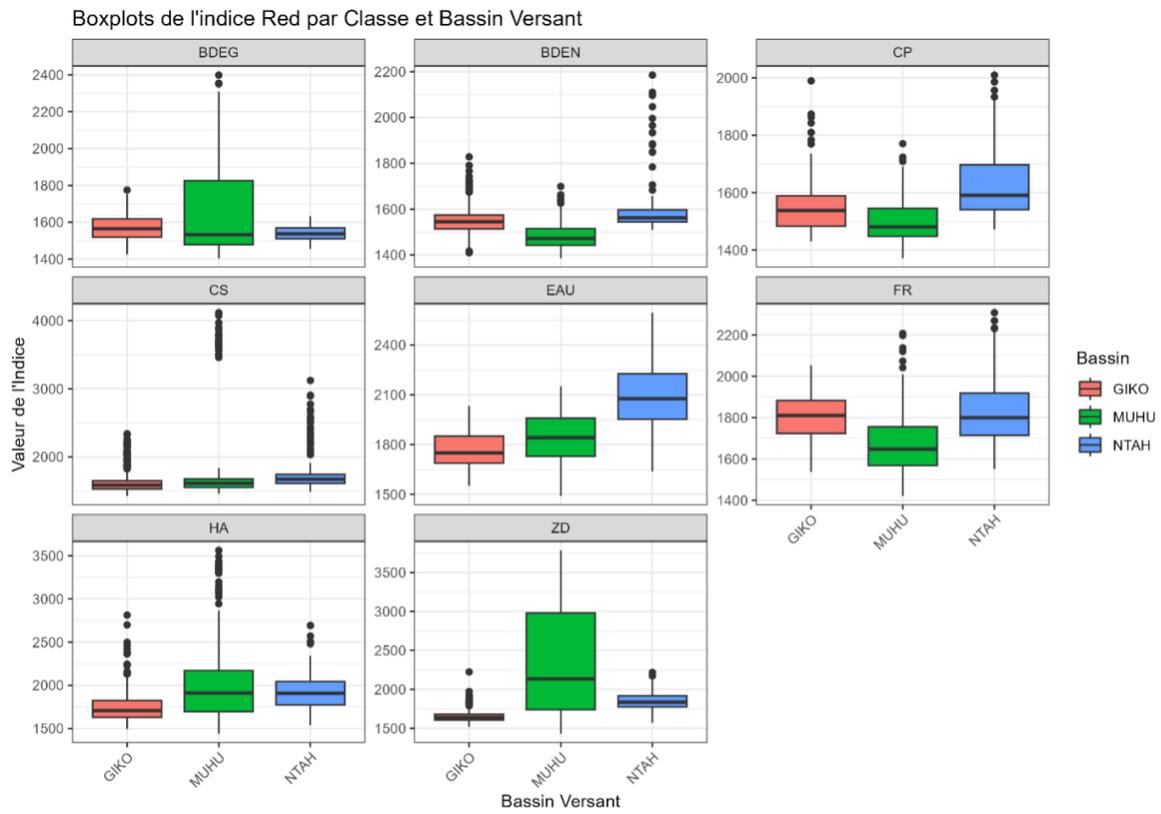


FIGURE 24 – Boxplot de la bande Red pour chaque classe d’occupation de sol au sein des bassins versants de la commune de Isale (Image Sentinel).

Il est observé sur ces figures que pour la plupart des occupations du sol, la distribution des valeurs de bande varie en fonction du bassin versant. Les difficultés éprouvées par un modèle à prédire l’occupation du sol sur un bassin versant sont en partie expliquées par le fait que celui-ci est entraîné sur des données n’ayant pas les mêmes caractéristiques pour une même classe d’occupation. Ceci explique aussi l’augmentation de la précision des modèles lorsque ceux-ci sont entraînés sur plusieurs bassins versants. En plus d’accroître le nombre d’entités reprises dans les sets de données d’entraînement, la variance des valeurs d’indices au sein des classes est mieux prise en compte.

Sur le Tableau 32, il est observé que le modèle 3 Sentinel, entraîné sur les bassins GIKO_02 et NTAH_05, a des performances similaires au modèle 4 Sentinel entraîné sur les 3 bassins versants de la commune d’Isale. L’ajout au set de données d’entraînement des données issues du bassin MUHU_03 n’a pas d’effet sur ces performances. L’ordre d’importance des variables de ces deux modèles ne varie que pour les trois variables les moins importantes.

Selon différentes recherches, l’environnement et les conditions météorologiques influencent la réflectance mesurée et donc les valeurs des différentes bandes spectrales. L’humidité de l’atmosphère et du sol, l’angle d’incidence des rayons, la pente et l’exposition des surfaces

sont des facteurs impactant celle-ci (Phillips et al., 2012; Zhao et al., 2023).

Une étude des pentes et expositions des classes d'occupation de sol au sein des différents bassins versants a été réalisée. Les Figure 25, Figure 26, Figure 27 et Figure 28 présentent les résultats graphiques obtenus.

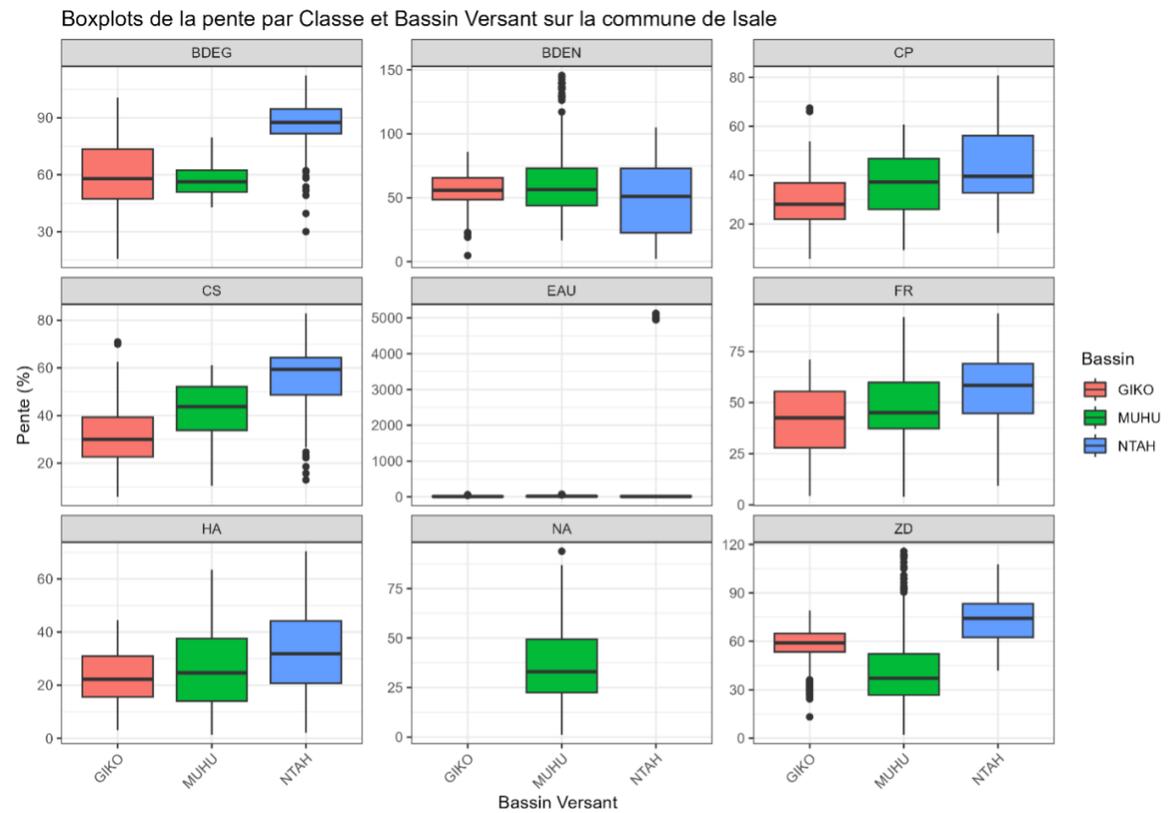


FIGURE 25 – Boxplots des pentes en fonction du bassin versant pour chaque occupation de sol sur la commune de Isale.

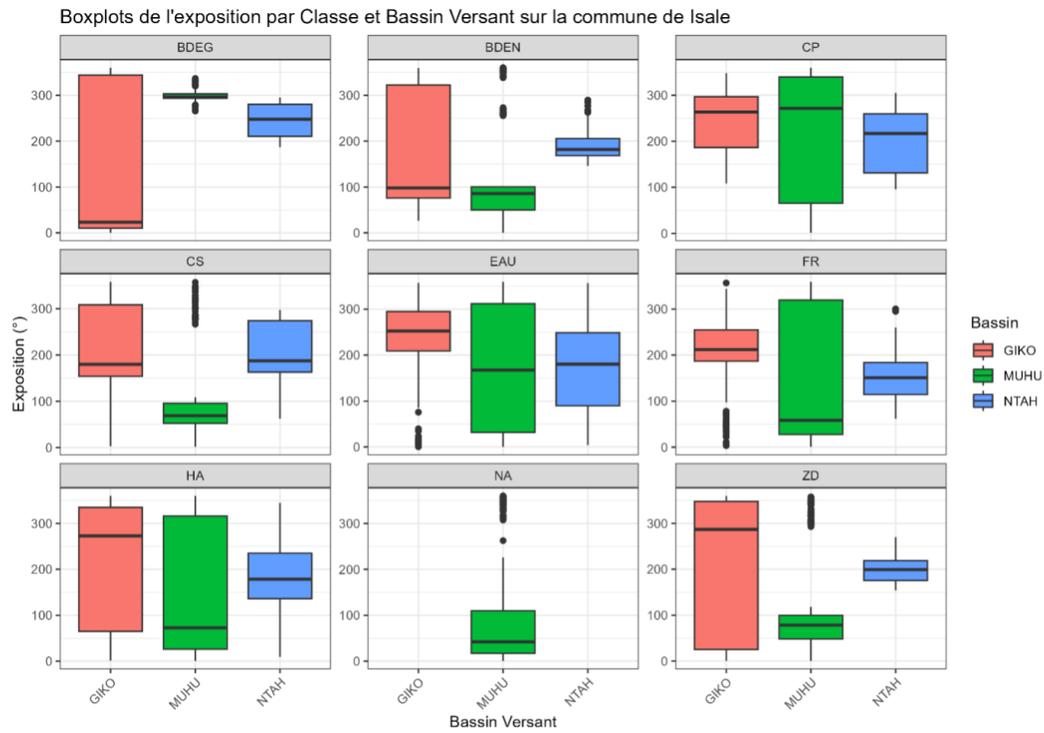


FIGURE 26 – Boxplots de l'exposition en fonction du bassin versant pour chaque occupation de sol sur la commune de Isale.

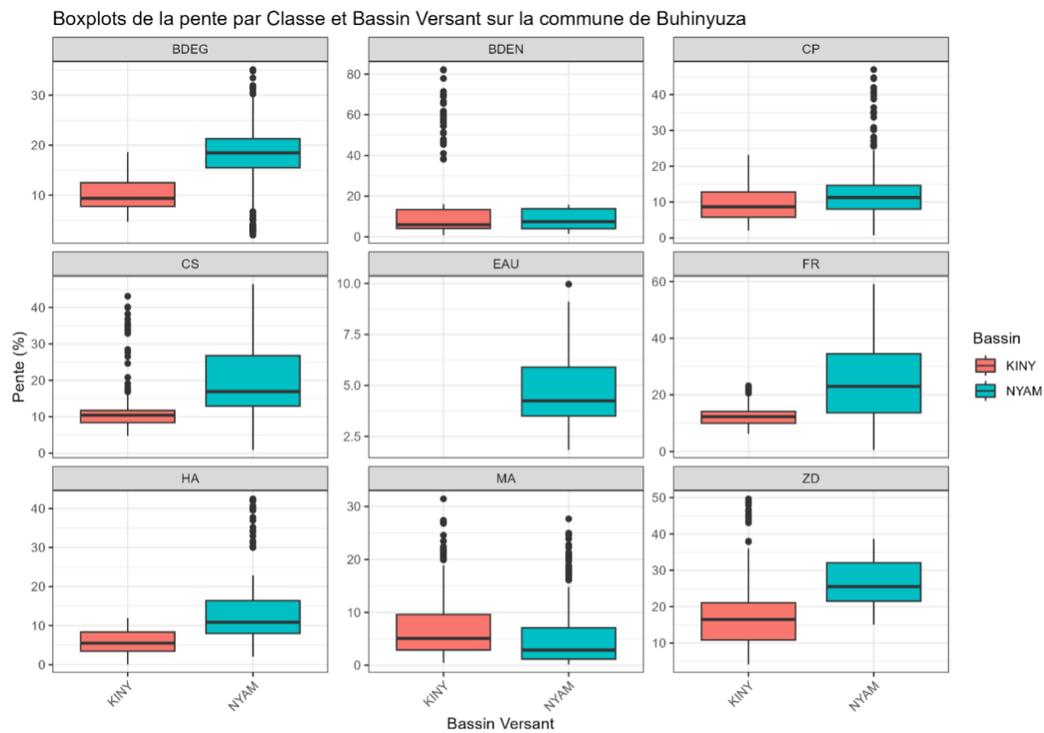


FIGURE 27 – Boxplots des pentes en fonction du bassin versant pour chaque occupation de sol sur la commune de Buhinyuza.

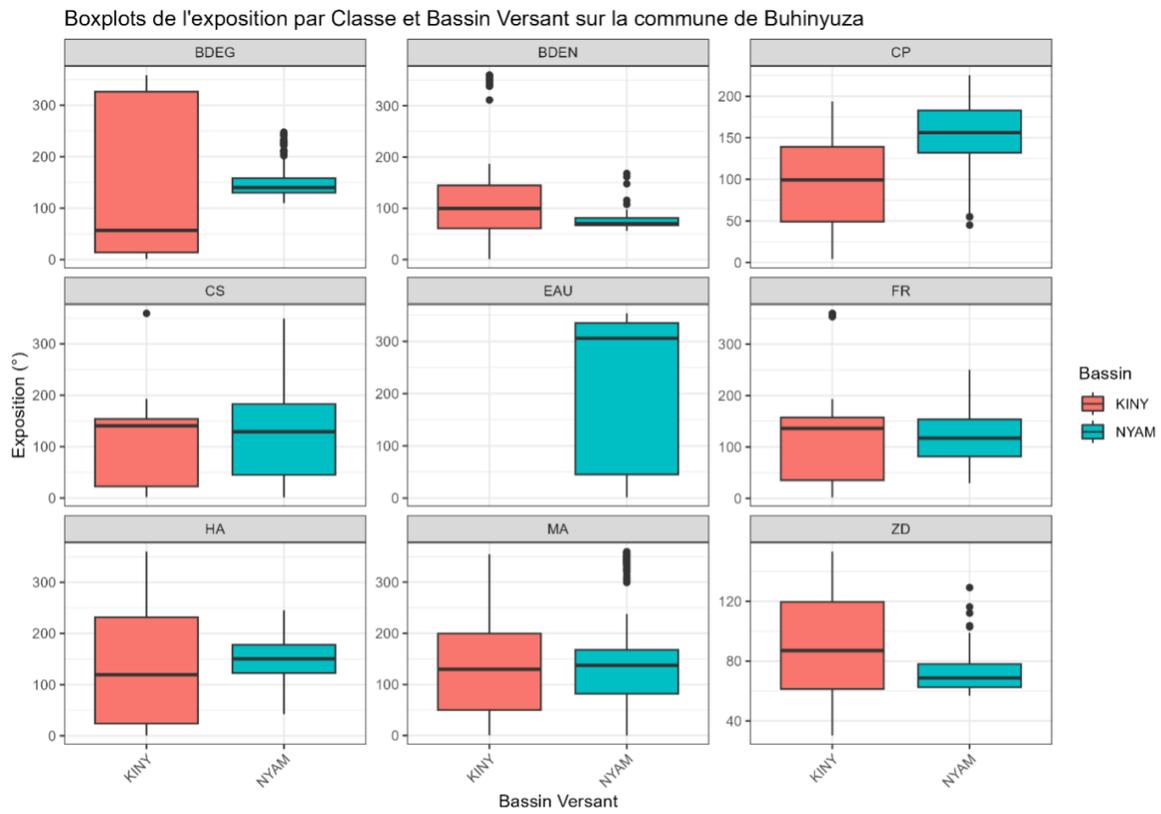


FIGURE 28 – Boxplots de l’exposition en fonction du bassin versant pour chaque occupation de sol sur la commune de Buhinyuza.

Il est constaté que les pentes et expositions diffèrent en fonction du bassin versant, et ce pour la plupart des classes. Ce phénomène est particulièrement visible sur les graphiques relatifs à la commune de Isale. Ce constat est logique, la commune présentant un relief plus important. Des méthodes ont été développées afin de corriger les effets de l’environnement sur la réflectance, elles ne semblent cependant pas encore totalement abouties (Zhao et al., 2023).

4.3 Les images PlanetScope

Il a été vu à la section 1.1.4.1 que les données issues de la mission PlanetScope sont parfois considérées comme n’étant pas prêtes à l’utilisation lors de leur téléchargement.

D’un point de vue géométrique, il a été démontré que les produits PlanetScope ayant subi la correction la plus rigoureuse effectuée par l’entreprise pouvaient présenter des erreurs de positionnement horizontaux pouvant aller d’environ 5 mètres à 19 mètres (Houborg & McCabe, 2018; Frazier & Hemingway, 2021). Il est donc envisageable que dans certains cas il n’y ai pas de correspondance entre certains pixels et les polygones tracés lors de la photointerprétation. De plus, un filtre est appliqué sur la fraction de la surface des pixels couverte par les

polygones. Il est possible que certains pixels n'aient pas été affectés par le filtre à cause d'une imprécision géométrique. Ces pixels ont alors pu être source de bruit lors de la classification. Des classes aux objets de petites surfaces comme les habitations ou les rivières auront été particulièrement affectées par cette imprécision.

Pour ce qui est de la radiométrie, une différence entre les plages de NDVIs mesurées par les satellites de la société Planet Labs et celles mesurées par les satellites Landsat a été constatée par [Houborg & McCabe \(2018\)](#) au cours d'une saison de croissance du maïs en culture irriguée en Arabie Saoudite. La plage mesurée par les doves Planet avait une amplitude plus faible (de 0,15 à 0,5) que la plage mesurée par les satellites Landsat (de 0,2 à 0,92). Ils ont également démontré que les données acquises par les satellites Planet divergeaient de plus en plus au fur et à mesure que le couvert végétal augmentait. Les images Planet ont donc tendance à lisser et minimiser la mesure et se montrent plus performante pour les mesures sur sol nu/à couvert faible. Le paysage du Burundi est verdoyant, le sol n'étant nu que lorsque la terre a été défrichée. Il se peut donc que l'utilisation d'images PlanetScope n'ayant pas subi de traitement suite à la commande ne soit pas une solution optimale à la classification d'images. Ces images subissent pourtant des traitements de normalisation et d'harmonisation (vis-à-vis des images Sentinel-2). Ceux-ci ont principalement pour effet d'améliorer la concordance entre les différentes prises de vue ainsi qu'entre les différentes générations de satellites. La concordance et la précision radiométrique avec les données Sentinel-2 n'est cependant que raisonnable selon l'entreprise PlanetLabs elle-même ([Kington, 2018](#)).

Dans leur étude, [Houborg & McCabe \(2018\)](#) utilisent la « méthode d'amélioration spatio-temporelle facilitée par CubeSat » (CESTEM) afin de normaliser les données radiométriques Planet acquises par plusieurs doves aux caractéristiques et horaires de passage différents. Le but est ici de rendre ces données cohérentes vis-à-vis des données Landsat. Le projet RapidAI4EO, financé par l'Union Européenne, a lui pour buts, entre autre, « d'établir les bases de la prochaine génération d'applications de surveillance rapide des terres, en particulier pour la gamme de produits Copernicus Land Use Land Cover (LULC) » et « de démontrer la fusion des images haute résolution de Copernicus avec des images très haute résolution de tiers ». Pour ce faire, le projet vise à la « création du corpus d'entraînement en apprentissage automatique spatio-temporel de télédétection le plus complet et dense à ce jour, combinant Sentinel-2 avec des images multispectrales harmonisées de Planet à très haute résolution de 3 mètres et à haute cadence sur 500 000 emplacements en Europe, et en rendant ces ensembles de données open source pour le bénéfice de toute la communauté de la télédétection » ([RapidAI4EO, 2024](#)). À nouveau, la CESTEM est évoquée afin de rendre cette fusion possible.

En proposant une méthode de "fusion spectro-spatiale directe d'images multisatellites mul-

tenseurs à l'aide de réseaux neuronaux convolutifs récurrents", [Latte & Lejeune \(2020\)](#) ont également réalisé la normalisation de données Planet via des images Sentinel, avec succès.

Il semble donc que des alternatives permettant le traitement à grande échelle des images Planet voient le jour. Celles-ci pourraient permettre l'utilisation d'images Planet plus homogènes permettant la création de modèles performants lors de prédictions sur des zones dont ne proviennent pas les données d'entraînement.

4.4 Comparaison des méthodes de classification

4.4.1 Méthode de cartographie de l'entreprise SHER

La méthode actuellement utilisée par la SHER est limitée par la validation terrain en terme de surface traitée par jour. Afin d'augmenter le nombre d'hectares traités, de nouveaux opérateurs peuvent cependant être engagés. Pour traiter 1000 hectares, le travail d'un opérateur durant 3 à 4 jours sera nécessaire à l'acquisition, la numérisation et la validation de cette numérisation. La méthode développée est donc moins chronophage.

Un des inconvénients de cette méthode provient du fait que différents opérateurs peuvent être la source d'interprétations différentes lors de la photointerprétation. Augmenter le nombre d'opérateurs en vue d'un traitement plus rapide de grandes surfaces peut donc amener un biais de subjectivité. De plus, cette méthode est limitée par la résolution des images Google Earth qui est variable. Il peut parfois s'avérer difficile de différencier précisément deux occupations de sol en certains endroits. Sur ce point, la photointerprétation réalisée sur les ortomosaïques, dont la résolution est meilleure, permet la mise en place de règles précises de photointerprétation. Un des atouts de la méthode actuellement utilisée est cependant la multitemporalité des sources de photointerprétation.

Actuellement, la méthode développée au cours de ce mémoire présente donc des avantages et des désavantages vis-à-vis de la méthode appliquée par la SHER. Les sous-produits générés lors de l'acquisition par drone peuvent être valorisés et les modèles de prédictions sont améliorables.

4.4.2 Comparaison avec le WorldCover de l'ESA

Pour rappel, la mission WorldCover de l'ESA propose une carte d'occupation du sol (11 classes) sur l'ensemble de la planète.

Sur la zone d'étude, tous bassins confondus, six classes d'occupation du sol sont répertoriées selon l'ESA : sol nu, bâti, cultures, prairie, couvert arbustif et couvert arboré (en anglais : bare/sparse vegetation, built-up, cropland, grassland, shrubland, tree cover). La carte permet donc moins de distinction. Une comparaison entre la carte de l'ESA et les polygones de photointerprétation a été réalisée afin de représenter la confusion entre les classes identifiées par les différents systèmes. Le Tableau 41 présente les proportions des classes identifiées par l'ESA composant les polygones de chaque classe identifiées dans le cas présent.

TABLE 41 – Nombre de pixels des classes identifiées par l'ESA présent dans les polygones de photointerprétation de chaque classe.

Classe	Sol nu / végétation épars	Bâti	Culture (annuelle)	Prairie	Couvert arbustif	Couvert arboré
BDEG	0	0	6	1193	22	254
BDEN	0	3	1	95	16	562
CP	0	0	34	43	146	503
CS	0	0	282	683	88	46
EAU	0	4	51	70	1	60
FR	0	0	214	559	102	44
HA	0	83	12	13	5	70
MA	0	0	202	203	36	541
ZD	193	29	0	739	10	207

Les classes prairies, couvert arbustif et couvert arboré constituent la plupart des observations réalisées. Une confusion entre les classes Boisement dégradé, Culture saisonnière, Friche et Zone dégradée et la classe prairie est observée. Cette confusion est due à la définition de la classe prairie. Selon [Van De Kerchove \(2022\)](#), « cette classe inclut toute zone géographique dominée par des plantes herbacées naturelles (plantes sans tige ou pousses persistantes au-dessus du sol et dépourvues de structure ferme définie) : prairies, steppes, savanes, pâturages, avec une couverture de 10% ou plus, indépendamment des différentes activités humaines et/ou animales, telles que le pâturage, la gestion sélective des incendies, etc. Des plantes ligneuses (arbres et/ou arbustes) peuvent être présentes à condition que leur couverture soit inférieure à 10%. Elle peut également contenir des terres agricoles non cultivées (sans récolte/période de sol nu) durant l'année de référence. ». Les classes BDEG, FR et ZD peuvent être caractérisées par endroit par un couvert arbustif ou arboré à la couverture inférieure à 10%. Les zones de CS identifiées peuvent également, selon les cycles de rotation des cultures, correspondre à des terres agricoles non cultivées. Pour rappel, la photointerprétation a été réalisée sur base d'orthomosaiques issues d'images acquises sur une période d'un mois environ. La rotation des cultures n'est donc pas prise en compte.

Près de 60% des pixels de la classe couvert arboré sont repris dans les zones de Boisement

dégradé et dense et de Culture pérenne. Une correspondance entre les deux méthodes de classification est donc observée. La méthode mise en place dans le cadre de ce mémoire permet une distinction plus fine. Une confusion entre cette classe et la classe Marais est cependant observée. La carte proposée par l'ESA contient une classe « Herbaceous wetland » dont la définition correspond à celle de la classe Marais. Celle-ci n'est cependant pas identifiée au sein de la zone d'étude. Une des limitations connues de la carte WorldCover est le fait de confondre les classes « herbaceous wetland » et les terres de culture irriguées (Van De Kerchove, 2022). Il est effectivement constaté que les observations de la classe culture annuelle correspondent principalement aux classes Culture saisonnière, Friche et Marais. Le lien établi entre les deux premières classes est lui logique ; la carte WorldCover se basant sur une évolution temporelle, des terres identifiées comme friche ont pu être des cultures saisonnières et donc être classées comme culture.

Des confusions ont été observées entre les deux classifications. La carte proposée par l'ESA ne permet pas de distinction entre certaines classes identifiées par l'entreprise et dont la détection est nécessaire à l'efficacité des solutions de GIRE proposées. La carte WorldCover a pour but de donner une information sur l'ensemble de la planète. À l'inverse, il a été vu que les différents modèles proposés sont précis sur des zones de la taille des bassins versants. Cette carte peut toutefois servir de base à l'appréhension des zones d'études, certaines correspondances entre les méthodes ayant été observées.

4.5 Sources d'images multitemporelles

Un des aspects n'ayant pas été pris en compte lors de l'élaboration de la méthode de classification est la disponibilité des données Planet et Sentinel à une multitude de dates. Plusieurs études font état des performances de modèles de cartographie par classification entraînés sur base de données acquises à des dates différentes (Valman et al., 2024; Mountrakis & Heydari, 2023; Wang et al., 2015). L'utilisation d'images satellites multitemporelles permet de prendre en compte différents stades phénologiques présents sur un même type d'occupation du sol. Cet élément pourrait permettre de mieux intégrer les variations pouvant survenir dans des classes comme les Cultures Saisonnières. Une distinction entre les classes dont la couverture herbeuse varie (Culture Saisonnière), varie peu (Zone dégradée) ou n'existe pas (Friche/Sol nu) pourrait être faite plus aisément. De plus, cela permet d'augmenter la quantité de données pour une même classe lorsque la zone de travail ne permet pas la création de suffisamment d'entités de certaines classes lors de la photointerprétation.

Pour ce qui est de la gestion des risques d'inondations, une analyse temporelle permet la

détection de la présence momentanée de masses d'eau ([Valman et al., 2024](#)). En marge des classes identifiées dans le présent mémoire, les modèles construits pourraient également servir à la télédétection de zones inondables si un aspect multitemporel y était ajouté.

5 Conclusion

En conclusion, il apparaît que les modèles développés permettent une classification efficace au sein des zones dont sont issues les données d'entraînement. Ces modèles sont plus performants lorsque les images issues de la mission PlanetScope sont utilisées. Les modèles ne permettent cependant pas la prédiction de l'occupation de sol sur des zones éloignées des régions d'acquisition de données. Des traitements à effectuer en amont de l'entraînement des modèles sur les différentes sources de données ont été proposés. Ceux-ci pourraient permettre d'améliorer la capacité des modèles à être transposés spatialement.

L'entraînement des modèles ne s'est fait que sur base d'images acquises à une date précise. Il a été vu que d'autres cartes d'occupation des sols sont développées sur base d'un suivi temporel. Celui-ci permet de comprendre les changements survenant au niveau du paysage. Il est possible qu'une compréhension de ces changements, mise en lien avec les pratiques culturelles des régions d'intérêt aide à l'amélioration des modèles.

Pour ce qui est de l'aspect travail en entreprise, la méthode d'acquisition et de traitement d'images mise en place est moins chronophage que celle préalablement utilisée. Elle permet également de minimiser le biais de subjectivité apporté par l'opérateur chargé de la numérisation. Certaines classes créées afin de s'adapter aux incertitudes liées à la résolution de certaines images Google Earth ont aussi pu être supprimées. La classification est donc plus claire et efficace. La précision des prédictions peut cependant s'avérer faible.

Bibliographie

- Al-Najjar H.A.H. et al., 2019. Land cover classification from fused DSM and UAV images using convolutional neural networks. 11(12) :1461. doi :10.3390/rs11121461. Number : 12 Publisher : Multidisciplinary Digital Publishing Institute.
- Baret F. & Guyot G., 1991. Potential and limitations of vegetation indices for LAI and APAR assessment. 104 :88–95.
- Bidou J., 1990. Érosion des sols au burundi. 3(1) :2–20. Publisher : Persée - Portail des revues scientifiques en SHS.
- Chawla N.V., Bowyer K.W., Hall L.O. & Kegelmeyer W.P., 2011-06-09. SMOTE : Synthetic minority over-sampling technique. doi :10.1613/jair.953.
- Chen C., 2004. Using random forest to learn imbalanced data.
- Chen J. et al., 2015. Global land cover mapping at 30 m resolution : A POK-based operational approach. 103 :7–27. doi :10.1016/j.isprsjprs.2014.09.002.
- Chen P.C., Chiang Y.C. & Weng P.Y., 2020. Imaging using unmanned aerial vehicles for agriculture land use classification. 10(9) :416. doi :10.3390/agriculture10090416. Number : 9 Publisher : Multidisciplinary Digital Publishing Institute.
- Clausi D.A., 2002. An analysis of co-occurrence texture statistics as a function of grey level quantization. 28(1) :45–62. doi :10.5589/m02-004. Publisher : Taylor & Francis _eprint : <https://doi.org/10.5589/m02-004>.
- FAO, 2005. AQUASTAT profil de pays – burundi.
- FAO, 2015. Status of the world's soil resources.
- Frazier A.E. & Hemingway B.L., 2021. A technical review of planet smallsat data : Practical considerations for processing and using PlanetScope imagery. 13(19) :3930. doi :10.3390/rs13193930. Number : 19 Publisher : Multidisciplinary Digital Publishing Institute.
- GISER, 2016. Bonnes pratiques pour la gestion du risque de ruissellement en zone rurale.
- Hall-Beyer M., 2017a. GLCM Texture : A Tutorial v. 3.0 March 2017. doi :10.13140/RG.2.2.12424.21767.
- Hall-Beyer M., 2017b. Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales. 38(5) :1312–1338. doi :10.1080/01431161.2016.1278314.

- Haralick R., 1979. Statistical and structural approaches to texture. 67(5) :786–804. doi : 10.1109/PROC.1979.11328. Conference Name : Proceedings of the IEEE.
- Haralick R.M., Shanmugam K. & Dinstein I., 1973. Textural features for image classification. SMC-3(6) :610–621. doi :10.1109/TSMC.1973.4309314. Conference Name : IEEE Transactions on Systems, Man, and Cybernetics.
- He H. & Garcia E.A., 2009. Learning from imbalanced data. 21(9) :1263–1284. doi : 10.1109/TKDE.2008.239. Conference Name : IEEE Transactions on Knowledge and Data Engineering.
- Houborg R. & McCabe M.F., 2018. A cubesat enabled spatio-temporal enhancement method (CESTEM) utilizing planet, landsat and MODIS data. 209 :211–226. doi :10.1016/j.rse.2018.02.067.
- ISRIC, 2024. ISRIC — world soil information.
- Jiménez-Jiménez S.I., Ojeda-Bustamante W., Marcial-Pablo M. & Enciso J., 2021. Digital terrain models generated with low-cost UAV photogrammetry : Methodology and accuracy. 10(5) :285. doi :10.3390/ijgi10050285.
- Kington J., 2018. Scene level normalization and harmonization of planet dove imagery.
- Latte N. & Lejeune P., 2020. PlanetScope radiometric normalization and sentinel-2 super-resolution (2.5 m) : A straightforward spectral-spatial fusion of multi-satellite multi-sensor images using residual convolutional neural networks. Publisher : Multidisciplinary Digital Publishing Institute MDPI.
- Leutner B., Horning N., Schwalb-Willmann J. & Mueller K., 2015. RStoolbox : Remote sensing data analysis. doi :10.32614/CRAN.package.RStoolbox. Institution : Comprehensive R Archive Network Pages : 1.0.0.
- Marais-Sicre C., Fieuzal R. & Baup F., 2017. Apport des images satellites multi-spectrales (optique et radar) pour la classification des surfaces en herbe. (215) :25–41. doi :10.52638/rfpt.2017.311. Number : 215.
- McFEETERS S.K., 1996. The use of the normalized difference water index (NDWI) in the delineation of open water features. 17(7) :1425–1432. doi :10.1080/01431169608948714. Publisher : Taylor & Francis _eprint : <https://doi.org/10.1080/01431169608948714>.
- McHugh M.L., 2012. Interrater reliability : the kappa statistic. 22(3) :276–282.

- Millard K. & Richardson M., 2015. Remote sensing | free full-text | on the importance of training data sample selection in random forest image classification : A case study in peatland ecosystem mapping.
- Moeyersons J., 1994. Les essais récents de lutte anti-érosive au rwanda. doi :10.3406/caoum.1994.3505. Publisher : Persée - Portail des revues scientifiques en SHS.
- Montanarella L. et al., 2016. World's soils are under threat. 2(1) :79–82. doi :10.5194/soil-2-79-2016. Publisher : Copernicus GmbH.
- Mountrakis G. & Heydari S.S., 2023. Harvesting the landsat archive for land cover land use classification using deep neural networks : Comparison with traditional classifiers and multi-sensor benefits. 200 :106–119. doi :10.1016/j.isprsjprs.2023.05.005.
- NASA, 2020. Commercial SmallSat data acquisition program pilot evaluation report.
- OCHA U.N.O.f.t.C.o.H.A., 2024. Burundi : Plan de réponse inondations et montée du lac tanganyika (effets el-niño).
- Phillips R.L. et al., 2012. Mixed-grass prairie canopy structure and spectral reflectance vary with topographic position. 50(5) :914–928. doi :10.1007/s00267-012-9931-5.
- Planet Labs ., 2024. Scenes product bundles reference.
- RapidAI4EO, 2024. About – RapidAI4eo.
- Rego F.C., Cadima I.S. & Strand E.K., 2020. A log-ratio vegetation index (LRVI) for discrimination and classification of remote sensing data. 28(1) :45–70. doi :10.1051/silu/20202801045. Place : Oieras Publisher : EDP Sciences.
- Rishirumuhirwa T., 1994. Facteurs anthropiques de l'érosion dans les montagnes et hauts plateaux au burundi, rwanda et zaïre. 47(185) :23–34. doi :10.3406/caoum.1994.3502. Company : Persée - Portail des revues scientifiques en SHS Distributor : Persée - Portail des revues scientifiques en SHS Institution : Persée - Portail des revues scientifiques en SHS Label : Persée - Portail des revues scientifiques en SHS Publisher : Presses Universitaires de Bordeaux.
- Rodriguez-Galiano V.F. et al., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. 67 :93–104. doi :10.1016/j.isprsjprs.2011.11.002. Place : Amsterdam Publisher : Elsevier B.V.
- Rousse, 1973. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation.

- République du Burundi ., 2020. Projet de restauration et de résilience du paysage du burundi.
- Sefercik U.G. et al., 2021. LAND COVER CLASSIFICATION PERFORMANCE OF MULTISPECTRAL RTK UAVs. XLVI-4-W5-2021 :489–492. doi :10.5194/isprs-archives-XLVI-4-W5-2021-489-2021. Conference Name : ISPRS TC IV (WG IV-1)
The 6th International Conference on Smart City Applications - 27–29 October 2021, Virtual Safranbolu, Turkey Publisher : Copernicus GmbH.
- Semeki Ngabinzeke J. et al., 2018. Une méthode simple et rapide pour l'évaluation de statistiques d'occupation du sol à l'aide d'images à très haute résolution acquises par mini-drone. 335(1) :15–23. doi :10.19182/bft2018.335.a31497. Publisher : Centre de Coopération Internationale en Recherche Agronomique Pour le Développement.
- Stone R.P., 2000. Équation universelle des pertes en terre (USLE).
- Toro F.G. & Tsourdos A., 2018. UAV sensors for environmental monitoring.
- Valman S.J. et al., 2024. An AI approach to operationalise global daily PlanetScope satellite imagery for river water masking. 301 :113932. doi :10.1016/j.rse.2023.113932.
- Van De Kerchove R., 2022. WorldCover product user manual.
- Wang J. et al., 2015. Mapping global land cover in 2001 and 2010 with spatial-temporal consistency at 250 m resolution. 103 :38–47. doi :10.1016/j.isprsjprs.2014.03.007.
- Warrens M.J., 2015. Five ways to look at cohen's kappa. 05(4). doi :10.4172/2161-0487.1000197.
- Zhao H. et al., 2023. The impact and correction of sensitive environmental factors on spectral reflectance measured in situ. 15(22) :5332. doi :10.3390/rs15225332. Number : 22 Publisher : Multidisciplinary Digital Publishing Institute.

6 Annexe

6.1 Annexe 1 : Manuel d'utilisation du drone Mavic 3M

1. Introduction

L'ensemble des manipulations détaillant les étapes de montage du drone, d'entretien du matériel et relatives à l'utilisation des batteries sont disponibles dans le manuel d'utilisateur fourni par DJI¹. Les préparations préalables au vol, les différents points nécessitant une attention particulière et les étapes à suivre le vol sont détaillées dans la suite de ce document.

2. Préparation du vol

2.2. Travail effectué sur Qgis et RStudio

Avant de pouvoir envisager le vol, la première étape consiste à définir les plans de vol sur Qgis et de convertir celui-ci au format kml sur RStudio, le format Shapefile n'étant pas pris en compte par DJI. Il faut donc définir les zones d'intérêt et ensuite importer les fichiers sur la télécommande du drone.

2.2.1. Tracé des plans de vol

Il est d'abord nécessaire de créer une couche shapefile de polygones via Qgis. Il est à noter que ces couches doivent être créées dans le référentiel utilisé par DJI, à savoir le EPSG 4326 – WGS 84. Une fois la couche créée, tracer les polygones correspondant aux différents vols à effectuer sur la zone d'étude.

Dans le cadre du mémoire, des polygones dont la surface varie entre 140 et 200 hectares ont été privilégiées. Cette gamme de surface consomme de 2 à 2,5 batteries pour un temps de vol variant de 35 à 50 minutes pour une altitude de vol de 200m à une vitesse de 15m/s. Il est à noter que les surfaces de vol ont également été limitées pour éviter les problèmes de perte de signal liés à une trop grande distance entre les points les plus éloignés du plan de vol et le point de décollage. Cette distance dépend du relief de la zone et de la densité du couvert autour de la télécommande. Les batteries ne devant pas subir de décharge trop poussée, il faut veiller à ne pas prévoir des distances de plus de 1000m dans le plan de vol afin de pouvoir envisager un retour rapide en cas de besoin. À vitesse maximale (15 m/s), il faut un peu plus d'une minute pour couvrir cette distance, les manœuvres d'atterrissage et la reprise en mode manuel du drone peuvent également durer une minute. Ces manœuvres consomment plus de batteries qu'un vol régulier. Un retour d'une telle distance peut donc consommer une dizaine de pourcents de batterie ce qui n'est pas négligeable.

1. https://dl.djicdn.com/downloads/DJI_Mavic_3_Enterprise/20221216/DJI_Mavic_3M_User_Manual-EN.pdf

Avec 7 batteries disponibles dans le cadre de ce mémoire, il est donc envisageable de survoler 400 à 500 hectares par jour en fonction des conditions météorologiques. Un vent fort a pour effet de ralentir le drone et d'augmenter la consommation de batterie nécessaire au maintien de la trajectoire. La pluie empêche de voler, elle endommage les composants du drone. Les nuages, lorsque leur altitude est faible, empêchent l'acquisition d'images de qualité.

Dans le cas où plusieurs plans de vol sont à prévoir pour le survol d'une zone d'étude, il est nécessaire de veiller à ce que ces plans de vol se superposent afin d'assurer un recouvrement des images acquises entre les différents plans de vol. La taille de la zone de recouvrement est fonction de la largeur du drone et donc de l'altitude de vol. Pour une altitude de 200m, un recouvrement de 100m a été choisi et a permis de reconstituer des scènes d'ensemble satisfaisantes lors du traitement des images.

Les principaux éléments à prendre en compte pour délimiter les polygones relatifs aux différents vols sont le relief et la facilité d'accès aux diverses zones, notamment aux points de décollage.

Au niveau du relief, il est nécessaire de pouvoir tracer une ligne directe entre le drone et la télécommande sans qu'un élément du paysage ne vienne la couper, au risque de perdre le signal avec le drone et donc d'interrompre la mission. Si la zone d'étude présente des reliefs importants, il faut découper celle-ci en zones au sein desquelles le signal pourra être gardé. Les points hauts, si accessibles, sont à privilégier comme point de décollage. Ceci permet notamment d'optimiser la taille de la zone de vol. L'accessibilité de ces points est déterminée via des images Google Satellite et OpenStreetMap et discutée avec les opérateurs de terrain. Si le point de décollage se situe sur une crête, il faut veiller à ce qu'il n'y ait pas de perte de signal entre la télécommande et le drone, lorsque le drone descend vers le fond de vallée. Pour éviter cela, il peut être nécessaire de se déplacer le long de la crête. La Figure 1 illustre les plans de vol pour le bassin versant KINY_03 dans la commune de Buhinyuza. Le chevauchement entre les plans de vol, les chemins d'accès aux points de décollage prévus ainsi que ces points de décollage sont observables.

Plans de vol couvrant le bassin versant Kiny 03 dans la commune de Buhinyuza.



Figure 1 : Illustration des plans de vol couvrant le bassin versant KINY_03.

Plans de vol réalisés pour le survol du SBV NTAH 05 dans la commune de Isale. Illustration de l'impact du relief sur la structure des plans de vol et le choix des points de décollage.

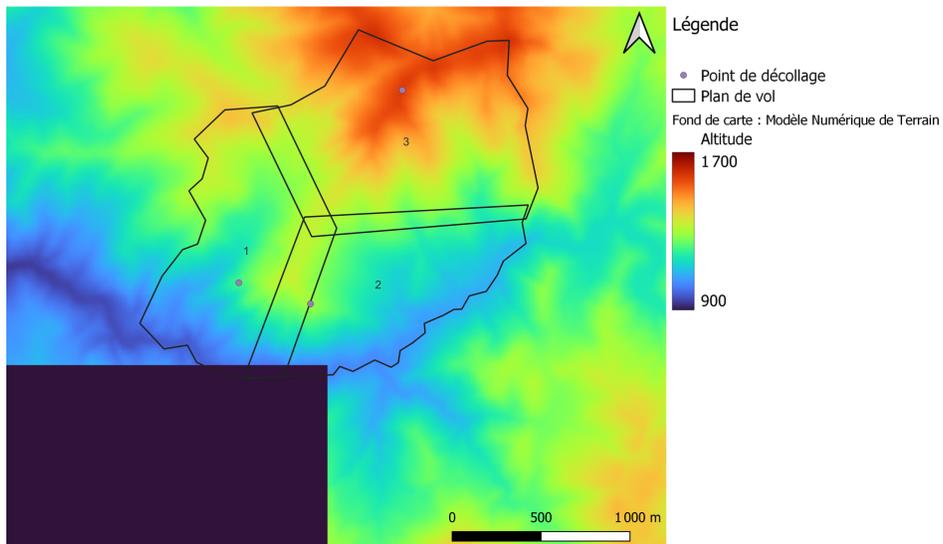


Figure 2 : Illustration des plans de vol réalisés sur le bassin NTAH_05, mis en lien avec le relief de la zone.

La Figure 2 illustre l'impact du relief sur la délimitation des zones de vols. Le tracé de la zone 3 est influencé par la présence d'une crête marquée en son milieu permettant de dominer le vol du drone à 360°. Il a cependant été nécessaire de se déplacer le long de cette crête lorsque le drone descendait vers la jonction entre les plans de vol 1 et 2. Il est à noter que les dénivelés important rencontrés sur ce bassin versant ont causés une consommation plus importante en terme de batterie, le drone volant à une altitude constante par rapport au sol (200m).

Les zones 1 et 2 sont tracées de manière à couper la crête afin de considérer des flans de colline différents et d'éviter d'avoir à se déplacer d'un coté à l'autre de la crête à chaque passage du drone.

2.1.2. Importation des plans de vol et du MNT dans la télécommande

Il est nécessaire de séparer les différents polygones relatifs aux plans de vol de la couche des plans de vols créée de manière à gérer chaque plan de vol séparément sur le terrain. Pour cela, l'outil « Séparer une couche vecteur » de Qgis est utilisé.

Afin que les plans de vol individuels puissent être lu sur la télécommande du drone, une transformation des couches shapefile générées précédemment en fichier kml est nécessaire. Pour cela, le code R nommé « Code_generate_kml_v2 » est utilisé (Figure 3). Ce code est directement fonctionnel et ne demande pas d'autre modification que la spécification du chemin d'accès du dossier d'intérêt à la ligne 8. Les fichiers kml peuvent alors être exporter par câble USB/USB-C ou USB-C/USB-C sur la télécommande du drone.

```

1 library(sf)
2 library(sp)
3 library(dplyr)
4 library(raster)
5 library(spatialEco)
6
7 # Update plots (put them in a specific folder. Ensure that these are polygons without inner rings or cuts)
8 plots <- list.files("C:/TFE/prepa_burundi/plan_vol_burundi/NTAH_05/SHP", pattern = "\\_shp$", full.names = TRUE)
9 # keep only the shapefiles
10
11
12 for(j in c(1:length(plots))){
13   file_name <- gsub(".shp", "", plots[j])
14   plot <- st_read(plots[j]) %>%
15     st_as_sf() %>%
16     #dplyr::select(Label, geometry) %>%
17     st_transform(crs = 4326)
18
19   df.coords <- extract.vertices(plot) %>%
20     st_coordinates() %>%
21     as.data.frame()
22   df.coords <- df.coords %>%
23     bind()
24   data.frame(Z = rep(0, nrow(df.coords))) %>%
25     mutate(coords = paste0(X, ", ", Y, ", ", Z))
26
27   vect_coord <- paste(df.coords$coords, collapse = ", ") %>%

```

Figure 3 : Illustration des modifications à apporter au code "Code_generate_kml_v2".

Il est également nécessaire d'importer un MNT pour assurer le suivi d'altitude lors du vol. Ce MNT doit être importé au format raster TIF. Il est à noter que la commande du drone n'est pas en mesure de traiter des fichiers raster d'un volume trop important .

Pour chaque plan de vol individuel, il convient d'appliquer un tampon de 50 à 100m lors du découpage du MNT afin de permettre le dépassement des limites de la zone par le drone (décrit plus loin).

Il est également à noter que le suivi d'altitude peut être géré par le drone sans avoir à importer un MNT grâce à des capteurs dont le drone est équipé.

2.2. Gestion des fichiers sur la télécommande

La suite des manipulations est directement réalisée via la télécommande du drone. Pour allumer celle-ci, il faut appuyer une fois court puis une fois long sur le bouton de démarrage (idem pour l'ensemble des appareils DJI, drone compris). Lorsque la télécommande démarre, elle émet un petit bruit et affiche l'écran d'accueil (Figure 4). L'application DJI Pilot 2 est automatiquement lancée mais n'est pas utile dans la suite des manipulations (Figure 5). Les différents éléments rencontrés sur ces deux écrans sont décrits aux pages 47 et 57 du manuel. Pour l'utilisation générale de la télécommande, se référer à ce manuel. La suite des traitements consiste à se rendre dans l'application « Files ». Elle correspond à l'explorateur de fichier. Il est possible d'y créer des dossiers dans lesquels stocker des fichiers. Pour chaque zone d'étude, créer un dossier. Ensuite, importer les fichiers kml des plans de vol créés précédemment. Dans le cas où un MNT est utilisé, celui-ci peut être stocké à l'endroit désiré de manière similaire.

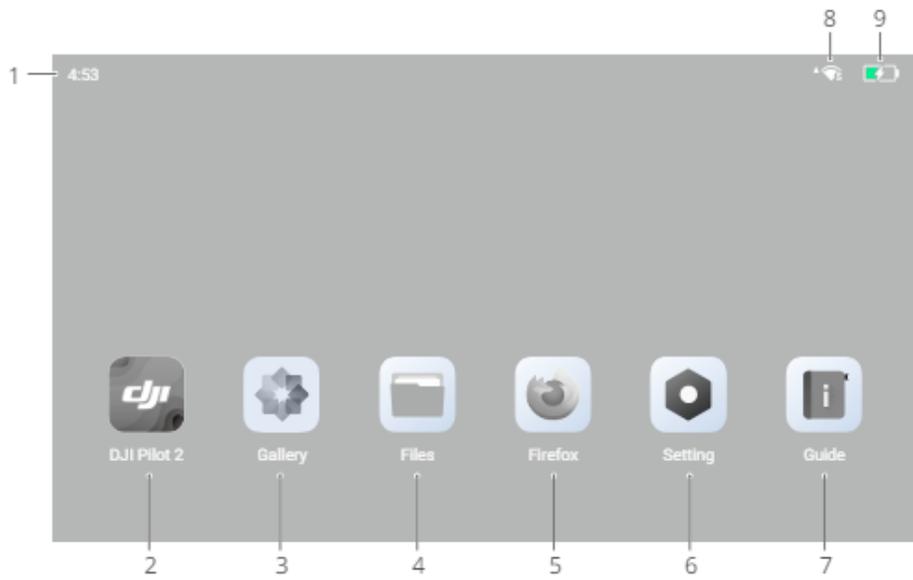


Figure 4 : Ecran d'accueil de la télécommande.

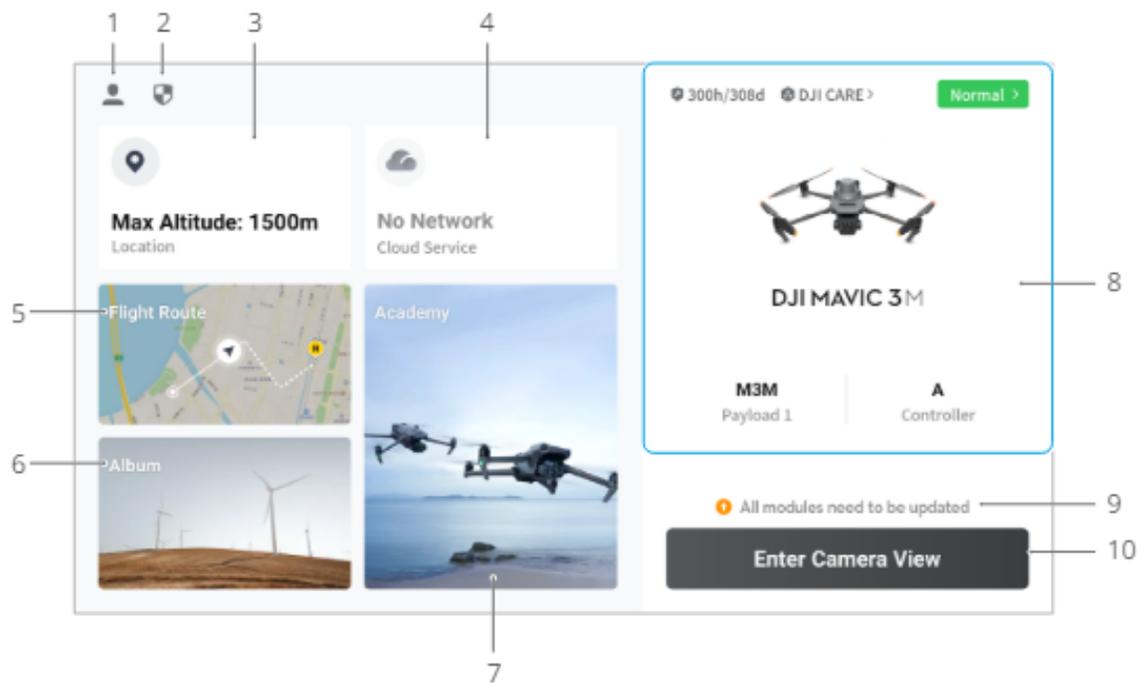


Figure 5 : Application DJI Pilot 2 de la télécommande.

2.3. Paramètres généraux de vol

Préalablement au vol, un ensemble de paramètres doivent être encodés à partir de l'application « DJI Pilot 2 » de la télécommande. Il s'agit des paramètres qui concernent le contexte de vol et peuvent rester les mêmes pour toutes une série de vols. Ils permettent de définir la hauteur maximale de vol, l'altitude de return-to-home (RTH) et les distances de freinage horizontales avant et arrière et verticale maximum en cas de détection d'obstacles. L'utilisation potentielle d'un module RTK peut également être gérée via ce menu. Pour accéder aux menus permettant de définir ces paramètres, qu'un drone soit connecté ou non à la télécommande, cliquer sur « Accéder à la Vue caméra » sur l'interface de base de DJI Pilot 2 (Figure 6) puis sur les trois points en haut à droite de la vue caméra (Figure 7 et Figure 8).

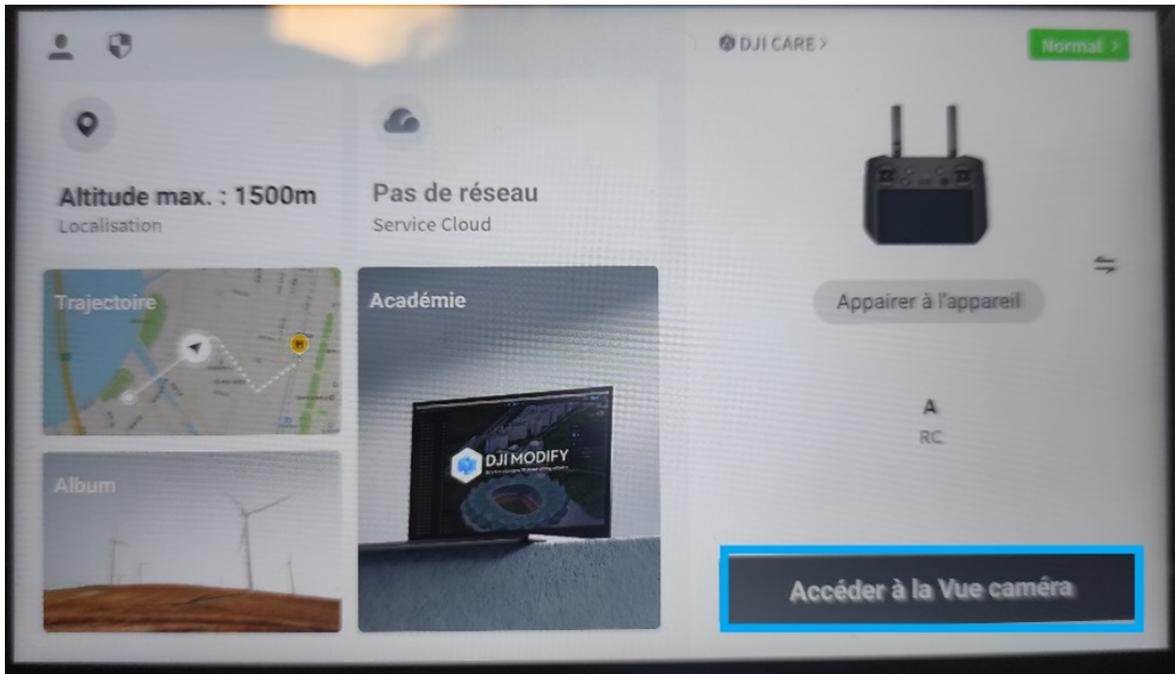


Figure 6 : Accès à la vue caméra.



Figure 7 : Accès aux menus des paramètres de vol.

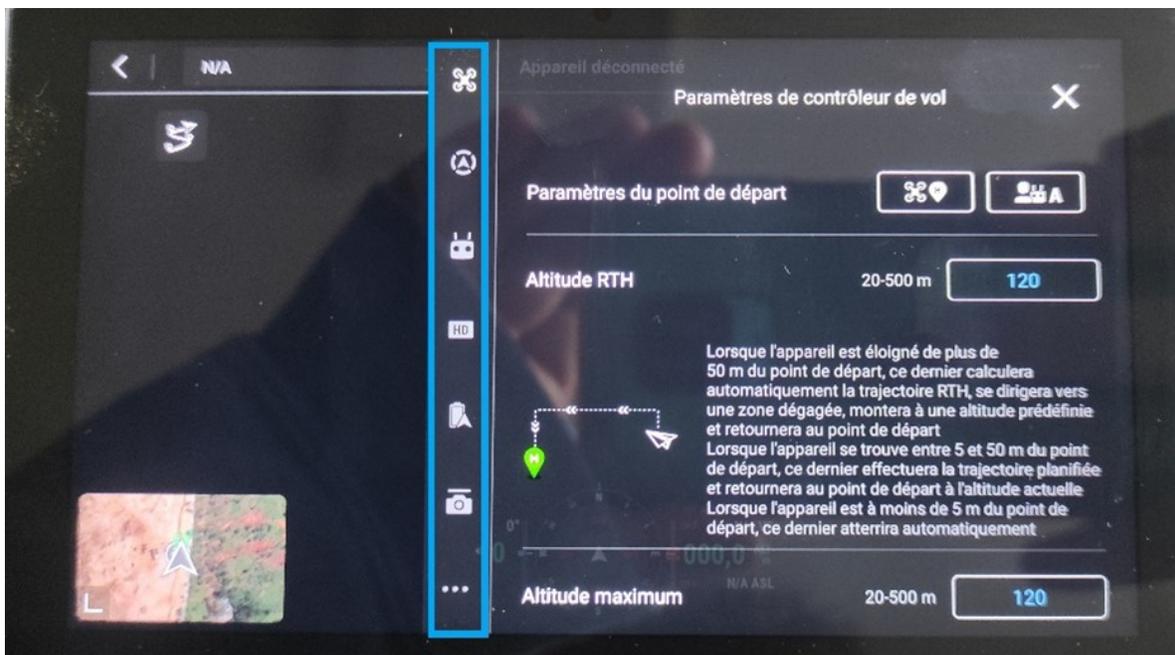


Figure 8 : Menu pour gérer les paramètres de vol.

Le premier volet concerne les paramètres de contrôleur de vol. On y règle les paramètres de :

- Point de départ : rien à modifier
- Mode de vol : autoriser les modes de vol multiples et sélectionner le mode N parmi les options A/N/S (Altitude : le drone maintient une altitude constante mais dérive au vent, utile en cas de problème de boussole / Normal : le drone reste fixe si aucune instruction n'est envoyée par la commande / Sport : plus de mobilité)
- RTH : Garder les préréglages. Modifier l'altitude au besoin. En mode suivi de terrain, il peut arriver que l'altitude de vol maximum doive être augmentée afin de permettre le RTH, ce sera signalé lors de la création des plans de vol.
- Altitude Maximum : Modifier au besoin et être attentif aux réglementations (120m en Belgique).
- Distance limite : à activer ou non en fonction des besoins et de la visibilité, entrer une distance manuellement.
- Statut capteurs : pas d'intervention, donne l'état de ceux-ci.
- Réglage du gain et de l'exposition : rien à modifier.
- Action Perte du signal : important de paramétrer à « Retour au point de départ ».

Le reste des paramètres ne demandent pas d'attention particulière.

Le second volet permet de paramétrer l'action à effectuer en cas de détection d'obstacles par les différents capteurs. Seul le freinage, paramètre par défaut a été utilisé.

Le troisième volet concerne les paramètres de la télécommande. Aucune modification n'a été apportée. Le mode 2 est le mode utilisé habituellement. Il est possible d'étalonner la télécommande et de personnaliser les touches au besoin.

Le quatrième volet concerne des paramètres caméra. Aucune modification n'a été apportée.

Le cinquième volet donne des informations sur l'état du drone. Il est possible de fixer le niveau de batterie pour lequel une alerte est émise ou bien le RTH automatique est déclenché.

Le sixième volet concerne les réglages de la nacelle. Aucune modification n'a été apportée.

Le septième volet permet les réglages du vol RTK. Activer le positionnement RTK en cas de besoin et suivre les indications du manuel de l'antenne (Station mobile D-RTK 2).

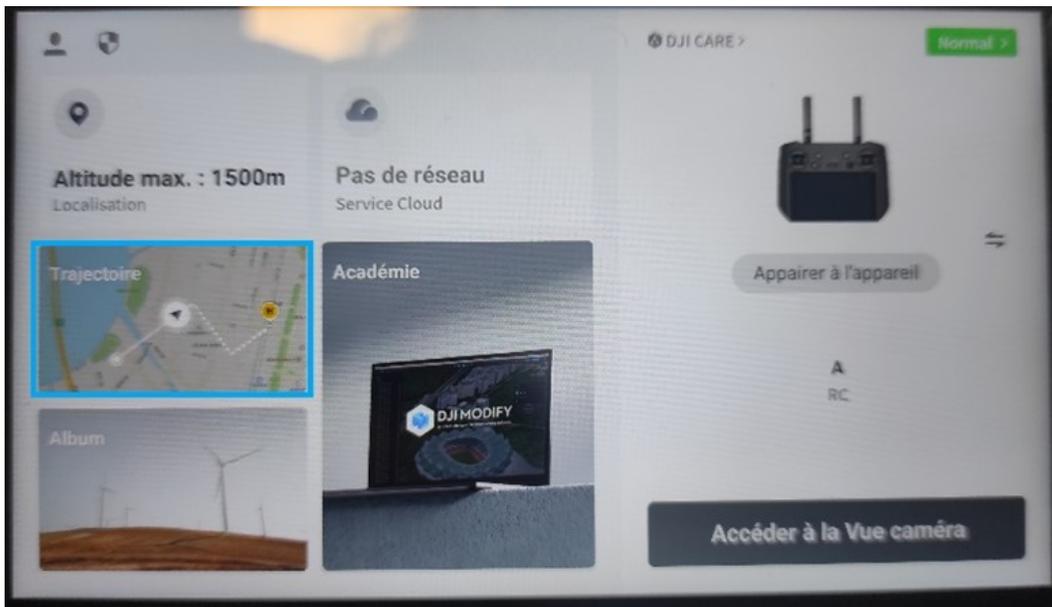


Figure 9 : Fenêtre "Trajectoire" de l'application DJI Pilot 2.

Le huitième volet reprend des paramètres plus généraux dont la modification n'est pas nécessaire.

2.4. Programmation du vol

Une fois les paramètres généraux vérifiés et modifiés, la programmation du vol peut commencer. Pour cela, se rendre sur DJI Pilot 2 et cliquer sur la fenêtre « Trajectoire » (Figure 9). Dans cet onglet apparaissent l'ensemble des plans de vol créés et toujours sauvegardés (Figure 10). Pour créer un nouveau plan de vol, cliquer sur l'icône + . Un choix sera alors à faire entre « Créer une trajectoire » ou « Importer l'itinéraire (KMZ/KML) ». Si un fichier KML est disponible, choisir d'importer l'itinéraire. Ceci renvoie vers le gestionnaire de fichier ou il est possible de choisir le plan de vol fichier dans le dossier où il est stocké. Une fois le choix du fichier confirmé, trois choix se présentent . Je n'ai travaillé que sur des trajectoires de zones. C'est l'outil à utiliser dans le cadre d'acquisition d'images sur un bassin versant. Pour lancer un plan de vol, cliquer dessus. Une nouvelle interface s'ouvre montrant une carte et les limites du plan de vol (Figure 11). Pour modifier les paramètres propres au plan de vol, cliquer sur la flèche à droite de l'intitulé du vol et ensuite sur le symbole de crayon affichée (Figure 12 et Figure 13).

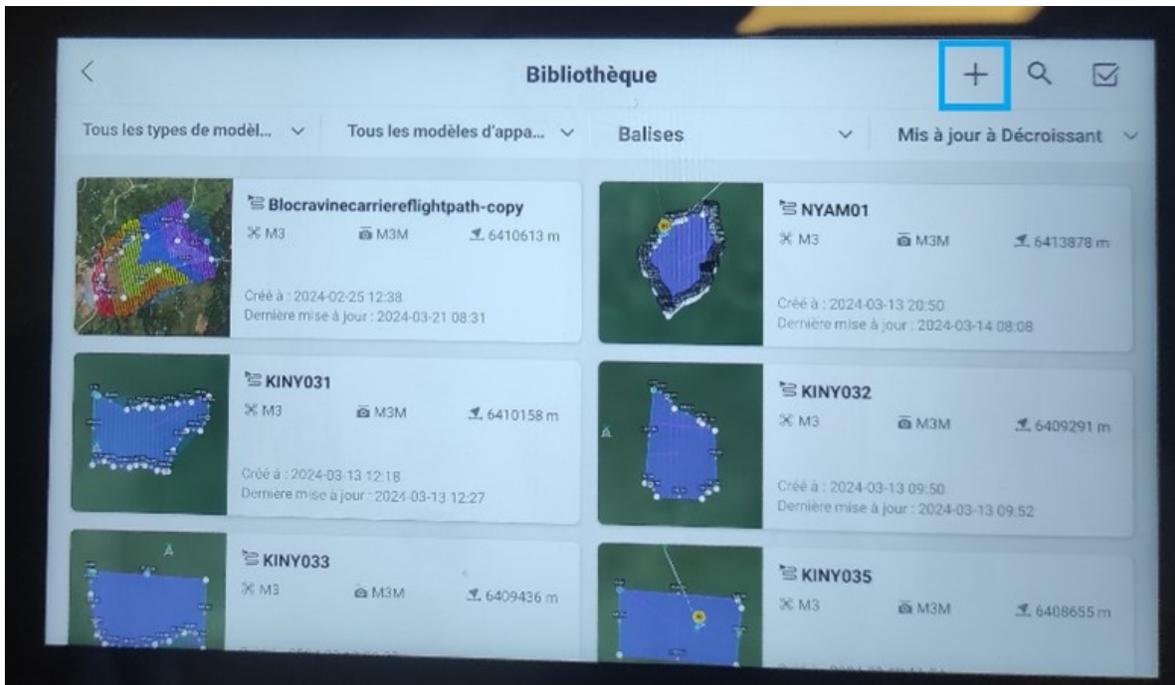


Figure 10 : Plans de vol sauvegardés sur la télécommande.

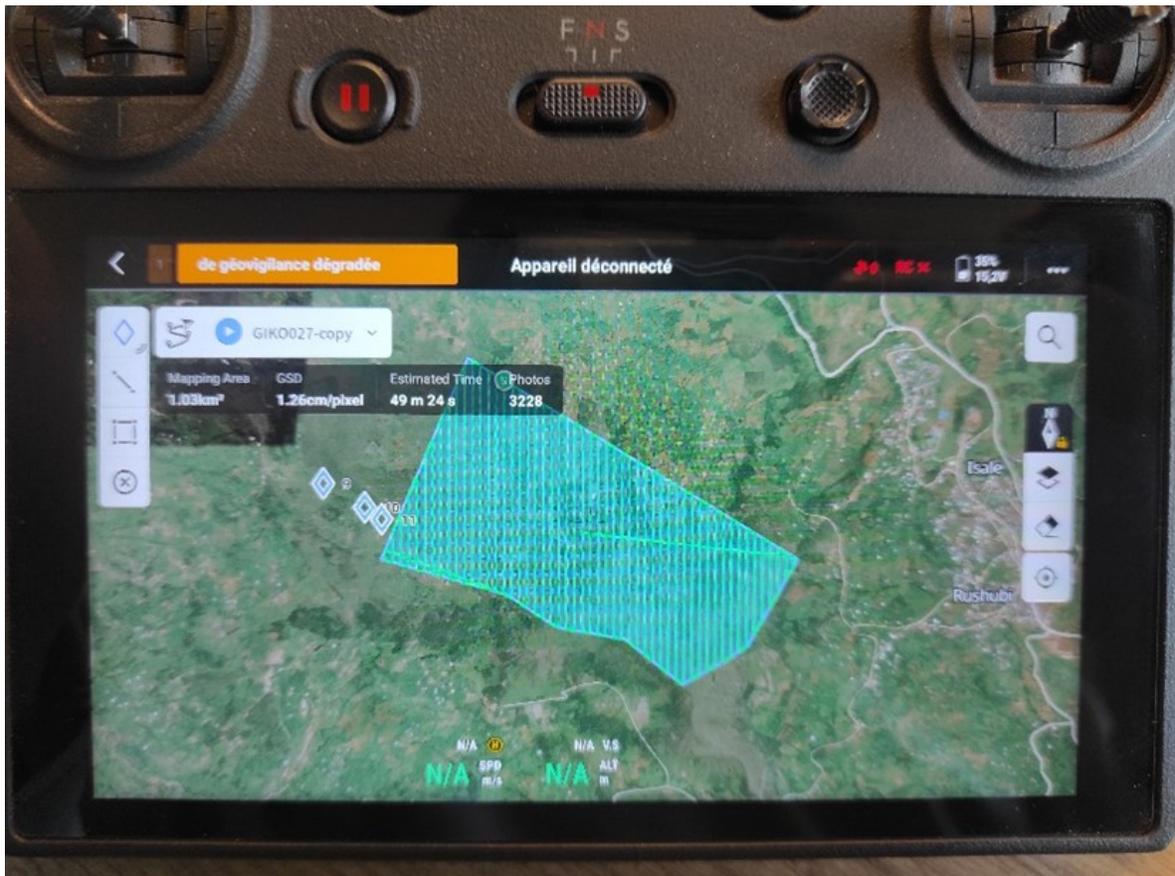


Figure 11 : Vue du plan de vol.

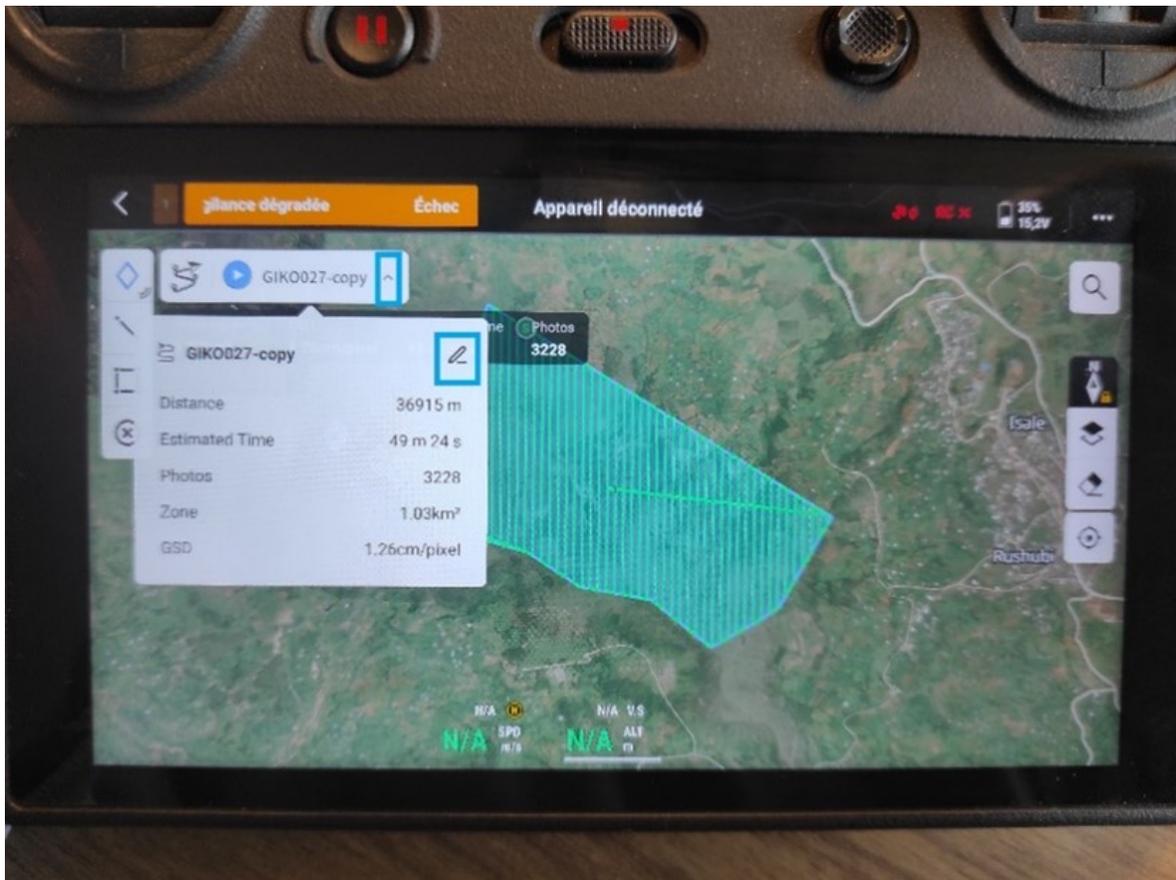


Figure 12 : Accès aux paramètres du plan de vol.

Il est alors possible/nécessaire de :

- Changer le nom de la trajectoire.
- Modifier le type de caméras : il est nécessaire ici de choisir la caméra fournie avec le drone. Dans le cas présent, sélectionner la gamme M3E et le modèle M3M. Si des images multispectrales sont nécessaires, cocher le sigle MS (affiché en bleu lorsque sélectionné). Cliquer sur « OK » une fois ces étapes réalisées (Figure 14).

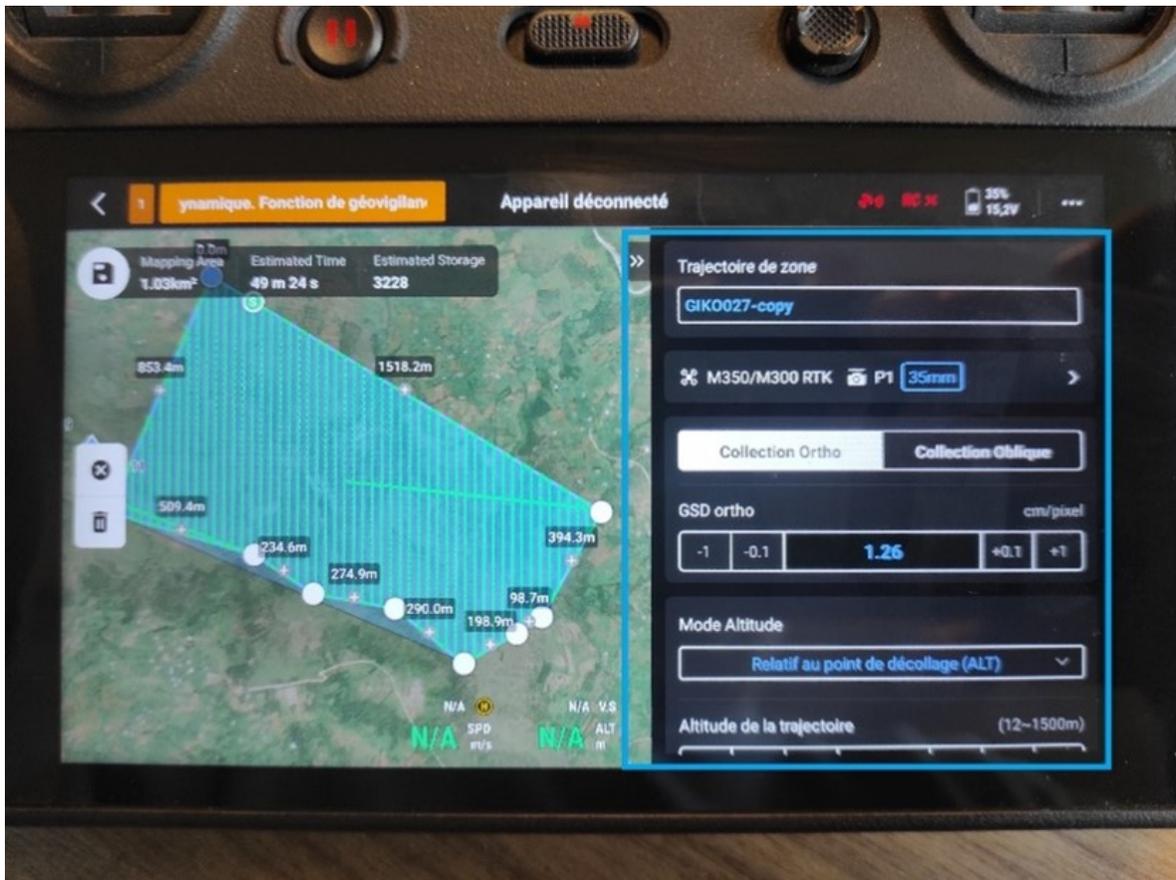


Figure 13 : Accès aux paramètres du plan de vol.



Figure 14 : Choix de la caméra.

- Choisir le type de collection. Dans le cas de ce mémoire, seule la collection « ortho » a été utilisée. Cela signifie que la nacelle est constamment orientée vers le bas, à la verticale.
- La résolution (GSD) ortho donne la résolution obtenue selon les paramètres d'altitudes décrits par après.
- Régler le mode altitude (Figure 15). Trois modes sont proposés :
 - Relatif au point de décollage : le drone volera à une altitude par rapport au niveau de la mer constante sur un plan horizontal.
 - Altitude : ne pas utiliser, semblable au mode précédent, l'altitude étant directement choisie par rapport au niveau de la mer et non du point de décollage, moins précis sans RTK.
 - AGL : Above Ground Level, paramètre le plus intéressant pour une photogrammétrie uniforme. Ce mode permet le suivi, en temps réel ou non, du relief à une hauteur de vol désirée. S'il est choisi, le suivi en temps réel peut être activé en cliquant sur l'icône à droite (bleue quand activée).

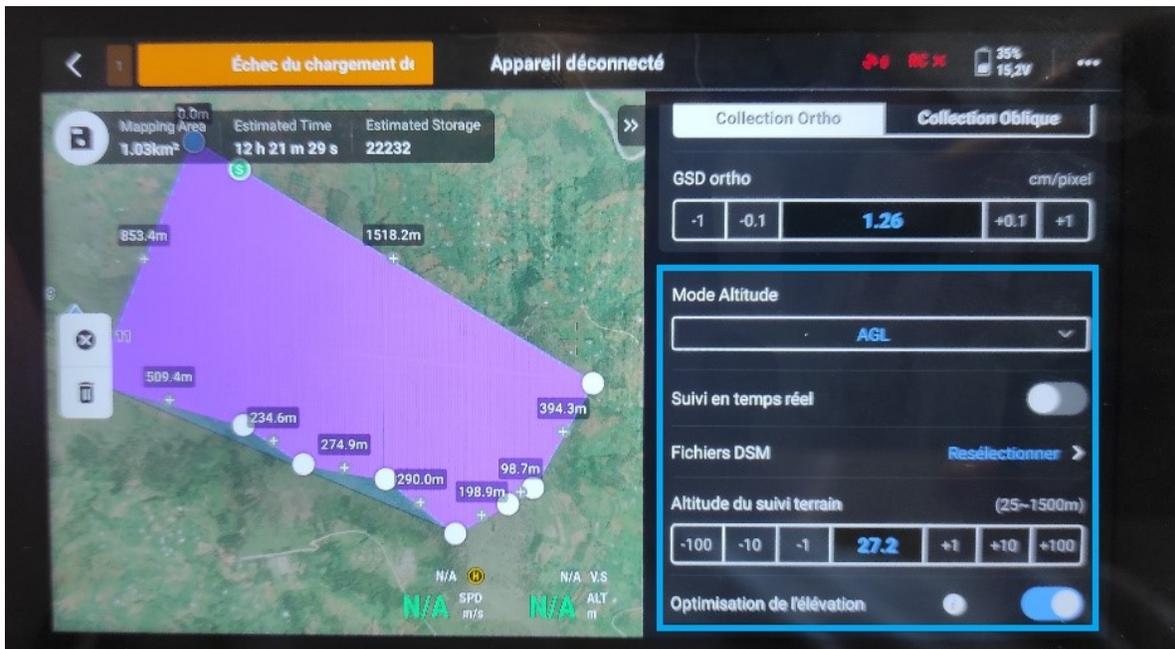


Figure 15 : Mode Altitude.

Si le suivi d'altitude est opéré par rapport à un MNT ou MNS, il est possible de le sélectionner parmi les fichiers importer ou parmi une liste reprenant les fichiers déjà utilisés (Figure 16).

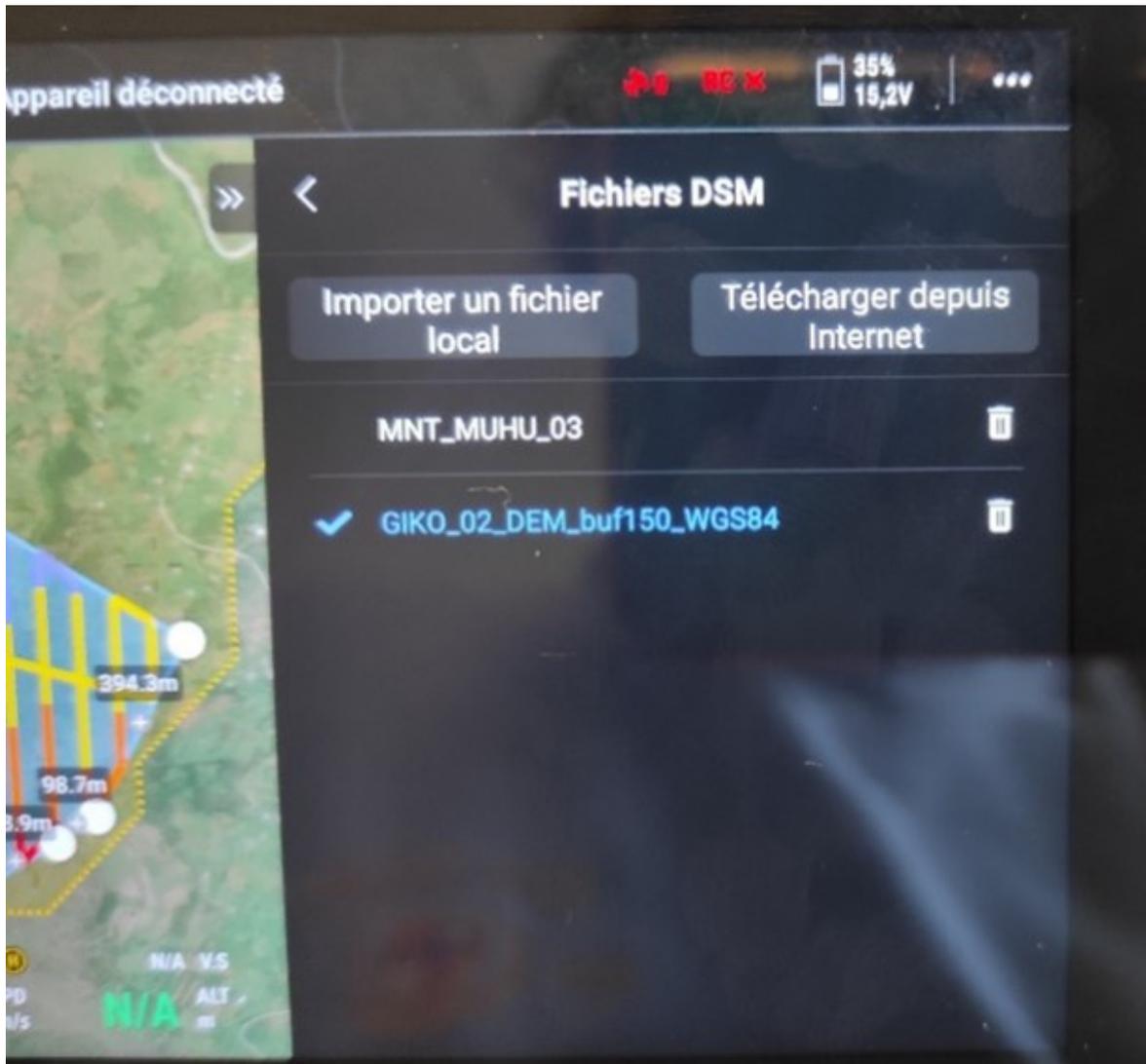


Figure 16 : Choix du MNT/MNS à utiliser.

- L'altitude de suivi de terrain peut alors être sélectionnée. Le Tableau 1 présente certaines résolutions prédites en fonction de l'altitude pour le drone MAVIC 3M.

Résolution en fonction de l'altitude.

Altitude [m]	Résolution [cm/pixel]
60	1,6
100	2,7
150	4
200	5,3

- · Régler l'altitude de décollage sûre : altitude par rapport au point de décollage que le drone devra atteindre avant de faire un mouvement autre que vertical. Ce paramètre dépend de la présence d'arbres ou de bâtiments autour du point de décollage (Figure 17).



Figure 17 : Réglage de l'altitude décollage sûre.

- Régler la vitesse de trajectoire : une vitesse de 15m/s a été choisie dans le cadre de ce mémoire.
- Régler l'angle de trajectoire : ce paramètre permet de régler l'orientation des lignes de trajectoire suivie par le drone. Si le relief est important, il est conseillé d'orienter ces lignes le plus parallèlement possible aux lignes de niveau afin d'éviter autant que possible de faire monter et descendre le drone, ce qui consomme de la batterie. Si la zone à couvrir est longue, faire revenir le drone le plus souvent possible près du point de décollage permet de garder un visuel sur celui-ci et également d'interrompre la tâche en vue d'un changement de batterie lorsque celui-ci se trouve près du point de départ. Les figures 17 et 18 illustrent deux trajectoires aux orientations différentes. La couleur des lignes de trajectoires varie en fonction des altitudes renseignées par le DSM selon une échelle donnée sur la gauche de l'écran. Dans le cas présent, une colline est située au sud, les lignes de trajectoire illustrée sur la Figure 18 permettent d'éviter au drone de devoir monter et descendre de fortes pentes à chaque aller-retour, contrairement à ce qui peut être observé sur la Figure 17.

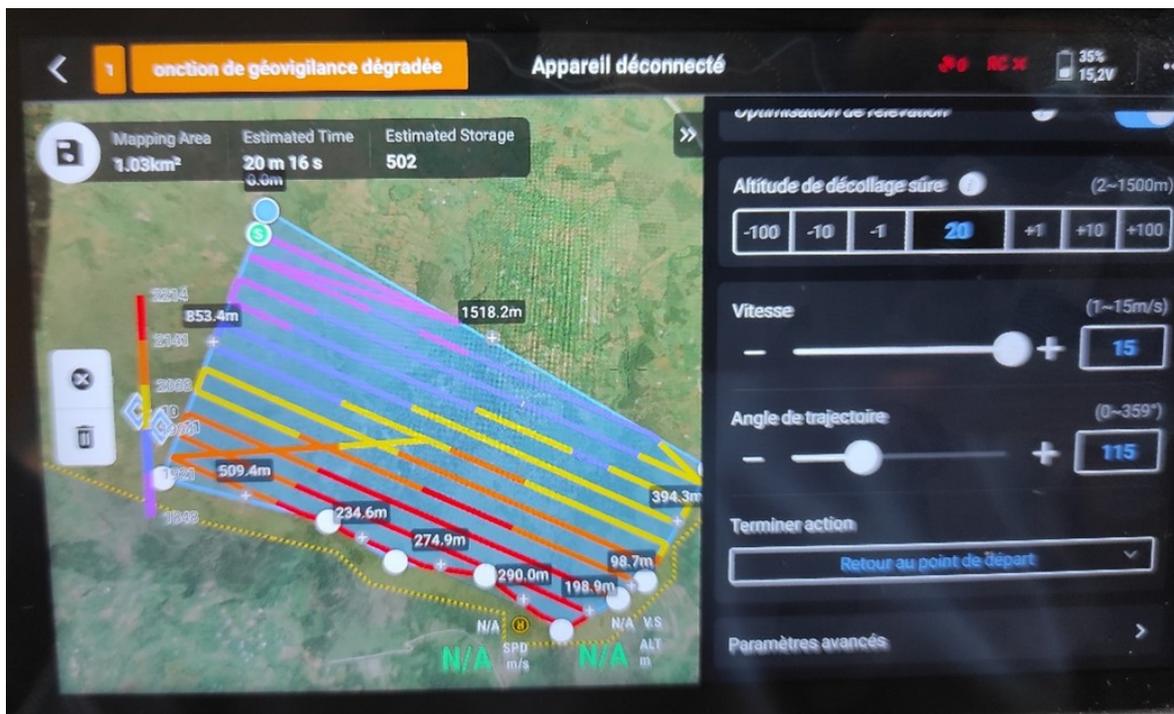


Figure 18 : Orientation des trajectoires de vol suivies par le drone.

- Choisir comment terminer l'action, le retour au point de départ étant à privilégier. Lorsque la tâche sera terminée, une signal sonore indiquera que le retour au point de départ est initié.

Les paramètres avancés sont important à régler. Ils concernent :

- Les taux de superposition latérale et frontale : détermine le taux de recouvrement entre chaque prise de vue. Pour réaliser la photogrammétrie, une superposition minimale de 70% est recommandée. Il est à noter que le nombre de prises de vue augmente avec le taux de recouvrement. Un taux de recouvrement de 80% a été utilisé dans le cadre de ce mémoire,
- La marge : correspond à une distance tampon appliquée au contour de la zone d'étude permettant d'éviter les effets de bords inhérents à la photogrammétrie. La marge à choisir dépend de la fauchée des photos et donc de l'altitude de vol. Une marge de 50m pour une altitude de vol de 200m a donné des résultats satisfaisant dans le cadre de ce mémoire. Il faut être attentif à ce que cette marge reste inférieure au tampon appliqué lors de la découpe du MNT dans la cas d'un vol en suivi de terrain.
- Le mode photo : choix de la méthode d'intervalle choisie pour la prise de photo. Selon les conseils donnés en cours à Gembloux, choisir la prise d'intervalle à distance.
- Angle de caméra personnalisé : ne pas modifier.

Lorsque ces étapes sont réalisées, le plan de vol peut être enregistré en cliquant sur l'icône

en haut à gauche de l'écran. À côté de celle-ci sont indiqués : la surface de vol, le temps de vol estimé et le nombre de prises de vue estimé. Une fois enregistré, retourner à l'interface de base de vol.

3. Vol

Ce chapitre détaille les précautions et étapes à suivre pour réaliser un vol préalablement programmé en mode automatique. Le vol manuel est lui simplement lancé en accédant à la vue caméra (voir section 2.3) et en utilisant la télécommande. Il est décrit à la section 4.

3.1. Point de décollage

La première chose à faire est de repérer un endroit propice au décollage. Cette endroit se situera en hauteur si le relief est important et dans une zone suffisamment dégagée. Une surface la plus plane possible de 2m² sera nécessaire au décollage et à l'atterrissage. Un petit hélicoptère en tissu ou autre est également à prévoir.

3.2. Préparation du drone

Une fois le point de décollage installé :

- Ouvrir la valise du drone.
- Sortir celui-ci et en déplier les bras.
- **Enlever le cache de la nacelle.**
- Insérer une batterie chargée et s'assurer qu'une carte SD est présente.
- Fixer les ailes et s'assurer que celles-ci sont bien fixées (attention au code couleur et au sens de rotation).
- Démarrer la télécommande (après en avoir déplié les antennes) et le drone en appuyant court puis long respectivement sur le bouton de démarrage et le bouton de la batterie.
Veiller à ce que la nacelle soit libre de mouvement.
- Le drone et la télécommande doivent se connecter automatiquement. Le modèle du drone s'affiche alors dans l'onglet supérieur droit de l'application DJI Pilot 2. Si ce n'est pas le cas, toujours dans cet onglet, cliquer sur « appairer l'appareil » et suivre les indications.
- Poser le drone au centre de l'hélicoptère si ce n'est pas déjà fait. Un signal sonore doit indiquer que le point de départ est mis à jour.

3.3. Vol

Pour lancer le vol :

- Se rendre sur l'application DJI Pilot 2.
- Sélectionner la trajectoire à réaliser dans l'onglet Trajectoire.
- Sur l'interface, vérifier que le contact entre le drone et la télécommande est établi.

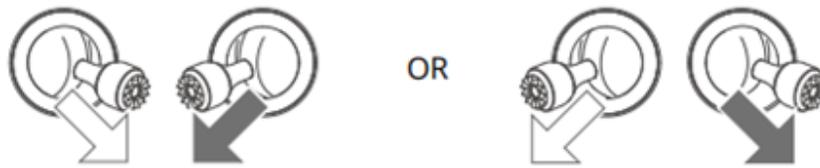
- Vérifier les paramètres de vol comme lors de leur création à la section 2.3 si besoin.
- Cliquer sur l'icône Start à gauche du nom de la tâche.
- Lire les vérifications d'avant décollage, s'assurer de l'état de charge des batteries (drone et télécommande), de l'espace de stockage disponible, ...
- Charger la mission de vol (envoi de la mission de vol au drone)
- S'assurer que personne ne soit trop proche du drone
- Lancer la mission de vol. **Attention le drone décollera automatiquement juste après.** Ne pas s'en approcher.

Pendant le vol :

- L'écran de la télécommande peut être mis en veille grâce à un cliqué court sur le bouton de démarrage. La vue caméra ou une vue GPS peuvent être choisies. La vue GPS permet de suivre la progression de la tâche et la position du drone par rapport au point de décollage. La vue caméra permet elle de s'assurer de la visibilité en cas de présence de nuages et également de la qualité des prises de vue.
- Un signal sonore est émis à chaque prise de vue. Il arrive cependant que ce signal ne soit plus émis lorsque la batterie est changée durant la réalisation de la tâche.
- Par sécurité, garder le drone en visuel tant que possible.
- Vérifier également régulièrement le niveau de la batterie du drone ainsi que la qualité du signal. Pour rappel, si celui-ci est perdu, la tâche sera interrompue. Il faut donc se déplacer avant la perte totale du signal. Si celui-ci est perdu, la tâche est interrompue et le drone revient au point de décollage.

Changement de batterie/carte SD :

- Lorsque la batterie du drone est trop faible, ramener le drone en mettant la tâche en pause au préalable. Si cela n'est pas fait, pour le niveau de batterie défini dans les paramètres du vol (voir section 2.3), celui-ci engagera le RTH et annoncera l'action par un signal sonore. La tâche sera mise en pause automatiquement. Dans le cas où la carte SD est remplie, un signal sonore retenti, mettre la tâche en pause et ramener le drone.
- Une fois que le drone a atterri, l'éteindre en appuyant court puis long sur le bouton de la batterie. Il n'est pas nécessaire d'éteindre la télécommande.
- Retirer la batterie/carte SD et en réinsérer une chargée. L'état de charge des batteries peut être observée en effectuant un cliqué court sur le bouton de celles-ci.
- Allumer le drone.
- Une fois celui-ci connecté, relancer le vol comme précédemment. Veiller à sélectionner l'action **Reprendre** si vous voulez continuer la tâche en court.



3.4. Après le vol

Lorsque la tâche est finie, le drone enclenche le RTH. Il est possible de le ramener manuellement afin de lui éviter la montée jusqu'à l'altitude de RTH. Lors de l'atterrissage, veiller à ce que les personnes présentes s'éloignent de la zone. En effet, le drone freine ses déplacements lorsque des obstacles sont détectés et peut refuser l'atterrissage lorsque trop d'obstacles sont présents. Il faut ensuite :

- Éteindre le drone et la télécommande.
- Détacher les ailes.
- Fixer le cache à la nacelle.
- Plier les bras.
- Ranger le drone et la télécommande dans la mallette.

4. Vol manuel

Le vol manuel comprend les mêmes étapes de préparation que le vol automatique. La différence s'opère une fois le drone connecté, monté et posé sur l'héliport.

À partir de l'interface DJI Pilot 2, se rendre sur la vue caméra. Les vérifications générales préalables au vol s'affichent, vérifier l'ensemble des paramètres et fermer si tout est bon. La vue de la caméra s'affiche alors sur l'écran de la télécommande.

Pour enclencher la rotation des hélices, orienter simultanément les joysticks vers le bas et vers le centre/l'extérieur.

5. Mesures de sécurité en vol

- Surveiller la météo et ses changements, le drone tolère mal la pluie ou un passage dans un nuage.
- Utiliser un petit héliport pour décoller et atterrir, la poussière endommage le drone.

Mode 2

Left Stick



Up



Down



Turn Left

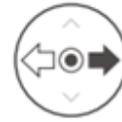
Turn Right

Right Stick

Forward



Backward



Left

Right

- Rester en mode Normal tant que possible. En cas de doute ou de problème autour du point de décollage, ce mode permet de simplement lâcher les joysticks et de maintenir le drone en vol stationnaire.