

Cartographie des unités d'occupation du sol du District d'Abidjan depuis le cloud Google Earth Engine, sur la base des images optiques Sentinel-2 et des algorithmes de Machine Learning

[Mapping of land use units in the District of Abidjan using Google Earth Engine cloud, based on Sentinel-2 optical images and Machine Learning algorithms]

NJEUGEUT MBIAFEU Amandine Carine², YOUAN TA Marc¹⁻², KAMENAN Satti Jean-Robert³, KOUAME Kouadio Armel¹, ASSOMA Tchimou Vincent², and JOURDA Jean Patrice¹⁻²

¹Unité de Formation et de Recherche des Sciences de la Terre et des Ressources Minières (UFR-STRM), Université Félix Houphouët-Boigny d'Abidjan-Cocody, Abidjan, Côte d'Ivoire

²Centre Universitaire de Recherche Appliquée en Télédétection (CURAT), Université Félix Houphouët-Boigny d'Abidjan-Cocody, Abidjan, Côte d'Ivoire

³EDP-INPHB Institut National Polytechnique Houphouët Boigny, Yamoussoukro, Côte d'Ivoire

Copyright © 2023 ISSR Journals. This is an open access article distributed under the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT: The objective of this study performed in the Abidjan District is to map land cover units using the Google Earth Engine (GEE) platform and Machine Learning algorithms such as Random Forest (RF), Support Vector Machine (SVM), Classification and Regression Tree (CART), Naive Bayes (NB), and Minimum Distance (MD). The data used include optical Multispectral Sentinel 2A satellite images with a 10-meter resolution, a 12.5-meter Alos Palsar digital terrain model (DTM) resampled to a 10-meter resolution, as well as cartographic data. The implemented methodology starts with the preprocessing and normalization of the composite image. The final composite image is created using eight spectral indices: NDVI, NDWI, MNDWI, VARI, SBI, SAVI, GCI, RGR, along with the first three bands of Principal Component Analysis and slope information. Subsequently, training and validation points are collected and coded based on image reflectance and ground truth data. The different classifiers SVM, RF, CART, MD, and ND are then trained and evaluated using various metrics such as confusion matrix, overall accuracy, producer's accuracy, consumer's accuracy (reliability), and Kappa coefficient. The classification performed with the RF algorithm achieved the highest overall accuracy of 83.28%, with a Kappa coefficient of 0.78. The statistics reveal that the Abidjan District is composed of 28.07% urban areas, 25.35% agricultural and other cultivated areas, 12.39% oil palm plantations, 10.05% rubber plantations, 4.66% banana plantations, 2.53% forests, 3.96% mangroves, 3.80% forest plantations (reforestation), and 9.2% water bodies in 2020. This study has led to an improved mapping of the distribution and proportions of land cover classes in the Abidjan District.

KEYWORDS: (Machine learning, Land cover, Google Earth Engine, Sentinel 2, Alos Palsar, Abidjan, Ivory Coast.

RESUME: L'objectif de cette étude menée dans le District d'Abidjan est de cartographier les unités d'occupation du sol en utilisant la plateforme Google Earth Engine (GEE) et des algorithmes de Machine Learning tels que la forêt aléatoire (RF), la machine à vecteurs de support (SVM), l'arbre de classification et de régression (CART), le naïve Bayes (NB) et la distance minimale (MD). Les données utilisées comprennent des images satellites optiques Multispectrale Sentinel 2A de 10 m de résolution, le modèle numérique de terrain (MNT) Alos Palsar de 12.5m rééchantillonné à une résolution de 10 mètres, ainsi que des données cartographiques. La méthodologie mise en œuvre débute par le prétraitement et la normalisation de l'image composite. L'image composite finale est créée en utilisant huit indices spectraux: le NDVI, le NDWI, le MNDWI, le VARI, le SBI, le SAVI, le GCI, le RGR, ainsi que les trois premières bandes de l'Analyse en Composantes Principales et des informations sur

les pentes. Ensuite, des points d'entraînement et de validation sont collectés et codifiés selon les réflectances de l'image et les vérités terrain. Les différents classifieurs SVM, RF, CART, MD et ND sont ensuite entraînés et évalués à l'aide de différentes métriques telles que la matrice de confusion, la précision globale, la précision du producteur, la précision du consommateur (fiabilité) et le coefficient de Kappa. La classification effectuée avec l'algorithme RF a obtenu la meilleure précision globale, atteignant 83,28%, avec un coefficient de Kappa de 0,78. Les statistiques révèlent que le District d'Abidjan est composé de 28,07% d'espace urbain, 25,35% d'aménagements agricoles et autres cultures, 12,39% de plantations de palmiers à huile, 10,05% de plantations d'hévéa, 4,66% de plantations de bananes, 2,53% de forêts, 3,96% de zones marécageuses, 3,80% de plantations forestières (reboisement) et 9,2% de plans d'eau en 2020. Cette étude a conduit à une meilleure cartographie de la répartition et des proportions des classes d'occupation du sol du District d'Abidjan.

MOTS-CLEFS: Machine learning, Occupation du sol, Google earth engine, sentinel 2, Alos polsar, Abidjan, Côte d'Ivoire.

1 INTRODUCTION

La ville d'Abidjan a connu une évolution galopante et par endroit à la limite anarchique ces trois dernières décennies [1]. Elle se retrouve actuellement en train de faire face à de nombreux problèmes tels que les catastrophes naturelles (glissements de terrain, inondations...) [2], [3], [4], [5], [6]. D'où l'intérêt d'avoir une carte détaillée des différents types d'occupation du sol qui le composent en vue d'aider à la gestion de ces risques. Depuis l'avènement des capteurs à haute résolution spatiale (HRS) et très haute résolution spatiale (THRS) dans les années 1990, les méthodes de traitement d'images, notamment la classification, ont connu un développement croissant [7], [8], [9], [10], [11], [12], [13].

Les méthodes de classification supervisées couramment utilisées basées sur une utilisation de support bureautique local, présentent de bons résultats mais restent limitées en termes d'exhaustivité des informations, d'automatisation et de paramètres pour améliorer les résultats. Le prétraitement des données et l'exécution des algorithmes sur un ordinateur local prennent beaucoup de temps en raison des limitations des ressources telles que la capacité de stockage, la vitesse de traitement et le temps d'analyse visuelle des images. Ce constat souligne l'intérêt d'explorer d'autres techniques et plateformes de classification d'images. Google Earth Engine (GEE) est une plateforme de traitement d'images basée sur le cloud qui offre un accès public à d'importantes collections de données satellitaires. Cette plateforme met à disposition des techniques d'apprentissage automatique pour la classification supervisée et non supervisée. Elle propose également de nombreuses fonctionnalités d'analyse d'images et permet un traitement rapide de vaste collection de données couvrant de vastes zones [14], [15], [16], [17]. De plus, dans la zone d'étude spécifique du District d'Abidjan, les algorithmes d'apprentissage automatique (MLA) ont été rarement utilisés pour différencier les unités d'occupation du sol à partir des données de télédétection multispectrales via la plateforme GEE.

Cette étude vise à explorer le potentiel de Google Earth Engine dans la classification supervisée des unités d'occupation du sol par les algorithmes de machine learning, en milieu urbain au niveau du District d'Abidjan, au Sud de la Côte d'Ivoire. Elle se fixe comme objectif général de comparer les cartes d'occupation du sol de 2020 du District d'Abidjan obtenues par entraînement des algorithmes de machine learning SVM, CART, RF, MD et NB depuis le cloud GEE. De façon spécifique, il s'agit d'exploiter les potentialités de la combinaison d'images optiques multispectrales Sentinel 2A et le MNT de haute résolution ALOS pour cartographier les unités d'occupation du sol à partir des AML SVM, CART, RF, MD et NB en 2020 et de sélectionner la meilleure classification en analysant les paramètres de validation et les statistiques des classes d'occupations du sols obtenues.

2 PRESENTATION DE LA ZONE D'ETUDE

Le District d'Abidjan, qui est l'objet de cette étude, est situé dans la partie méridionale de la Côte d'Ivoire, entre les latitudes 5°13' et 5°37' Nord et les longitudes 3°43' et 4°25' Ouest (Figure 1). Il a une superficie de 2137.43km² et est bordé au sud par l'océan Atlantique, au sud-ouest par le département de Dabou, au sud-est par le département de Grand-Bassam, au nord par le département d'Agboville, à l'ouest par le département de Grand-Lahou, et à l'est par le département d'Alépé.

Le District d'Abidjan est marqué par un climat chaud et humide. Il est caractérisé par un régime climatique de type équatorial de transition, connu sous le nom de climat Attiéen, selon une classification basée sur la durée des saisons, le déficit hydrique et des facteurs pédologiques à l'échelle de la Côte d'Ivoire [18].

En ce qui concerne la végétation du District d'Abidjan, elle se situe dans la région du domaine guinéen et se compose de deux secteurs distincts, à savoir le secteur ombrophile et le secteur littoral. On y trouve différents types de végétations, notamment les forêts sempervirentes denses et humides au nord, les forêts défrichées ombrophiles, les savanes littorales (pré-

lagunaires) ainsi que les forêts marécageuses et les mangroves [19]. La diversité végétale présente dans cette région est illustrée dans la Figure 2.

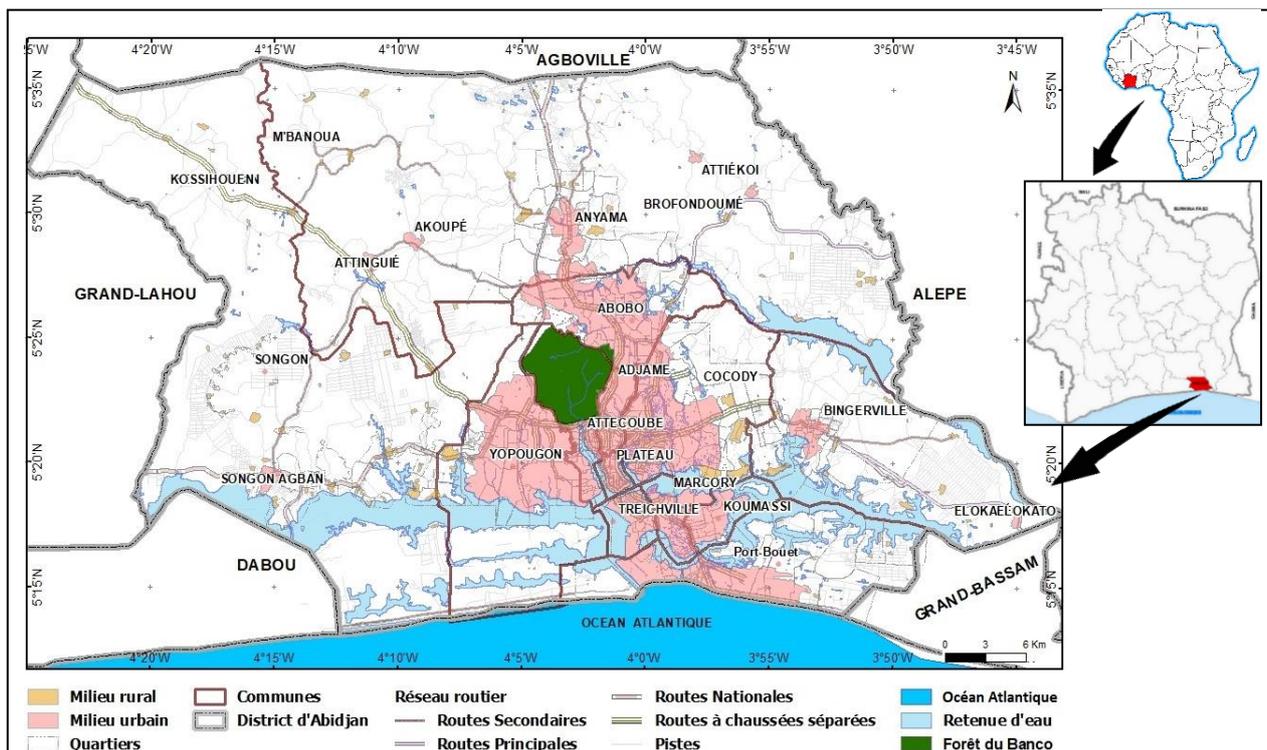


Fig. 1. Localisation de la zone d'étude

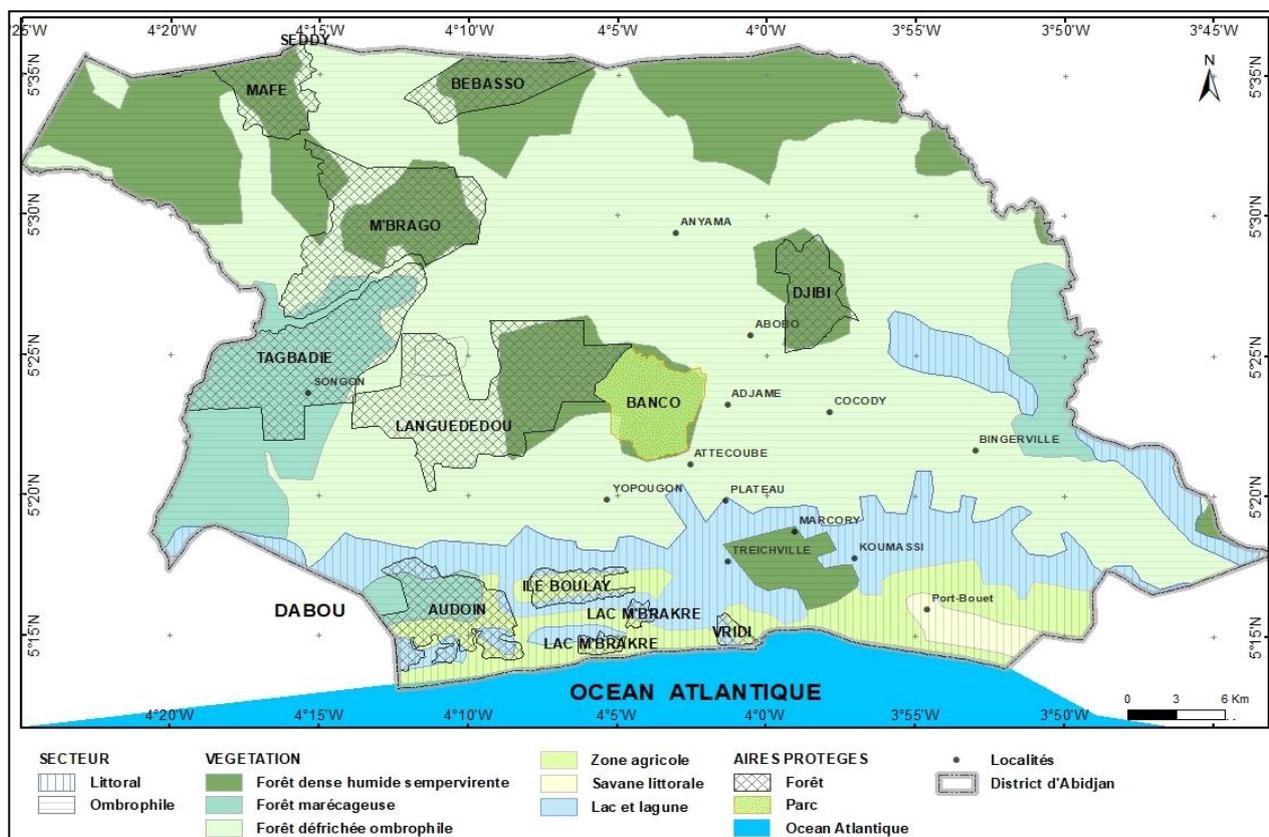


Fig. 2. Végétation du District d'Abidjan, (MCLAU, 2015)

3 MATERIELS ET METHODES

3.1 DONNEES ET MATERIEL

La cartographie des unités d'occupation du sol à partir de Google Earth Engine exploite les données d'imagerie optique provenant du capteur MSI de Sentinel-2. Les images de Sentinel-2 enregistrées pendant la période sèche et sans nuages ont un niveau radiométrique acceptable, ce qui permet une bonne identification des unités d'occupation du sol. Dans cette étude, GEE a été utilisé pour obtenir des données archivées de réflectance du sommet de l'atmosphère provenant de Sentinel-2 (ImageCollection ID: COPERNICUS/S2) offrant la meilleure qualité et un niveau de couverture nuageuse le plus bas possible (0 %). Pour une analyse optimale de la fenêtre temporelle, et la sélection des images de bonne qualité, la période d'acquisition a été fixé du 01 janvier au 31 Mars 2020. Quatre (04) scènes couvrant la zone ont été sélectionnées. Seules les sept bandes multispectrales (bleu, vert, rouge, proche infrarouge [NIR1 et NIR2], moyen infrarouge [MIR1 et MIR2] rééchantillonnées à 10 mètres) ont été utilisées (Tableau 1).

Tableau 1. *Caractéristiques des bandes Multi-Spectrales de (MSI) du satellites Sentinel-2 utilisées*

Scènes de sentinel-2	Canal	Résolution spatiale	Bande spectrale
T30NUL du 05/01/2020 T30NUM du 05/01/2020 T30NVL du 05/01/2020 T30NVM du 05/01/2020	Bleu	10 m	0.46 – 0.52 μm
	Vert	10 m	0.54 – 0.58 μm
	Rouge	10 m	0.65 – 0.68 μm
	Proche infra-rouge (PIR1)	10 m	0.78 – 0.90 μm
	Proche infra-rouge (PIR2)	20 m	0.85 – 0.87 μm
	Moyen infra-rouge 1 (MIR1)	20 m	1.57 – 1.66 μm
	Moyen infra-rouge 2 (MIR2)	20 m	2.10 – 2.28 μm

Le modèle numérique de terrain (DEM) PALSAR du satellite ALOS-1, avec une résolution élevée de 12,5 mètres, a été également utilisé pour extraire le relief et générer des produits dérivés tels que les pentes de la zone d'étude. Ces données d'archive sont disponibles au format Geotiff sur le site de l'Alaska Satellite Facility (ASF).

Les données administratives provenant de la base de données numérique géospatiale de la Côte d'Ivoire, réalisée en 2019 par le BNETD/CIGN, ont également servi pour extraire les contours du District d'Abidjan sur l'ensemble des données.

La cartographie des unités d'occupation du sol a été réalisée à l'aide du langage de programmation JAVA en utilisant le Code Editeur de Google Earth Engine et ses bibliothèques web cartographiques.

3.2 METHODES

La méthodologie utilisée pour cartographier les unités d'occupation du sol (OCS) se décline en cinq étapes essentielles. Tout d'abord, les images Sentinel-2 de l'année 2022 ont été prétraitées. Ensuite, les données d'entrée ont été préparées, notamment en effectuant une mission de collecte d'échantillons d'entraînement et de validation. Puis la classification supervisée des images Sentinel-2 à l'aide d'algorithmes d'apprentissage automatique (machine learning). Après la classification, une évaluation est réalisée pour déterminer la meilleure classification obtenue. Différents critères tels que la précision et l'exactitude sont pris en compte pour sélectionner la classification la plus fiable et représentative. Enfin, une analyse statistique de la carte d'occupation du sol est effectuée, permettant de tirer des conclusions et des informations pertinentes sur les différentes catégories d'occupation du sol. La figure 3 correspondante illustre cette méthodologie.

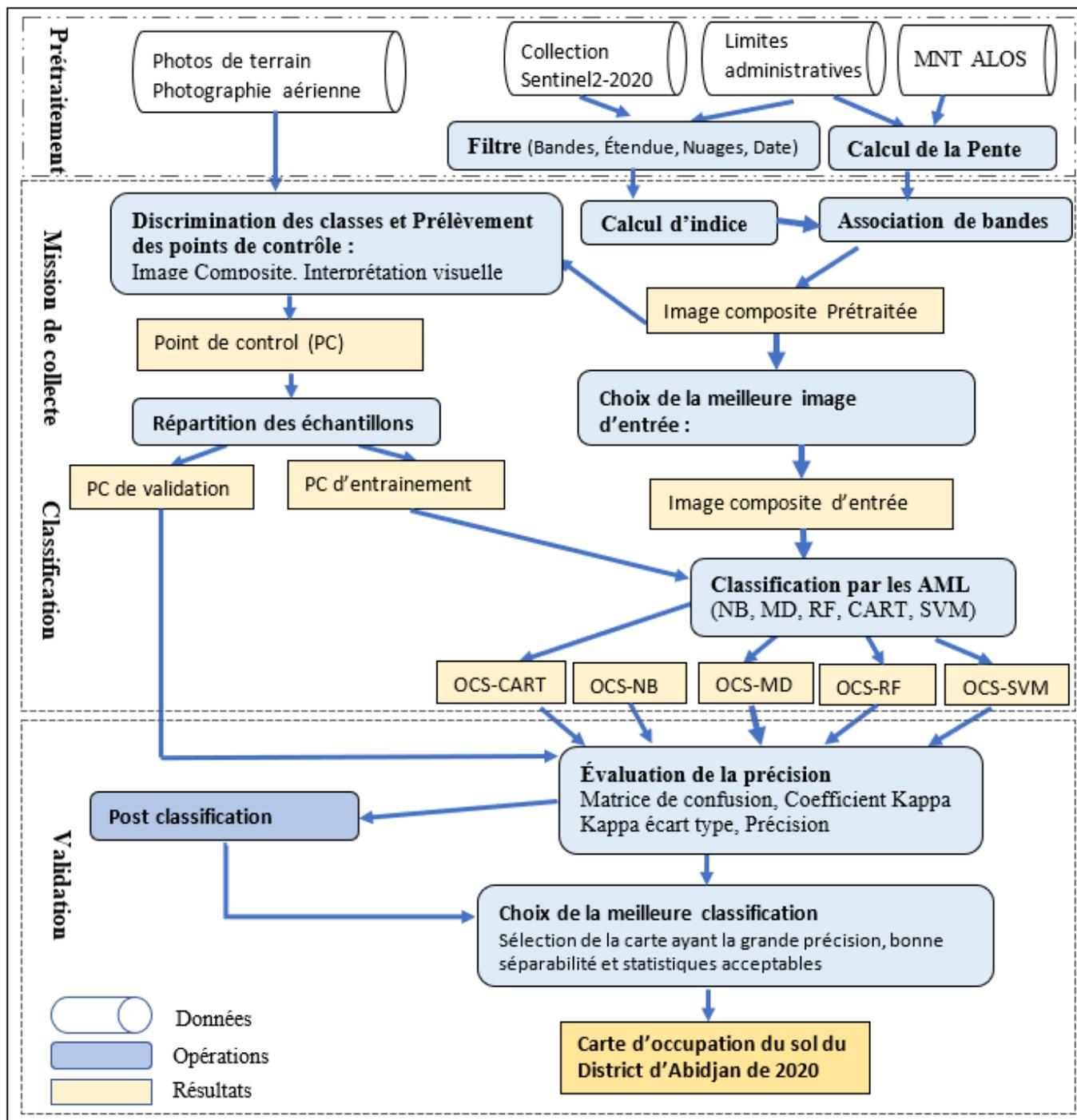


Fig. 3. Processus de classification des images multispectrales Sentine-2A par les algorithmes de Machine Learning depuis GEE

3.2.1 PRETRAITEMENT DES IMAGES

Les traitements préliminaires des images satellitaires, ont pour but de corriger certaines variations de la distribution des données causées par le décalage temporel dans l'acquisition des images. Cette variation s'explique en grande partie par les facteurs comme l'angle d'élévation du soleil, la distance terre-soleil, les conditions atmosphériques, la calibration des capteurs et la géométrie de visée qui affectent les valeurs numériques des pixels.

Le traitement préliminaire appliqué sur les images sentinel-2 a été essentiellement composé de filtres. Pour réduire l'étape de collecte d'images, un processus de filtrage est nécessaire. Dans ce travail, nous avons posé des conditions pour filtrer la

collection d'images à utiliser à l'aide de cinq variables: l'emplacement couvert, le type de produit, l'intervalle de temps, le pourcentage de nuage et la longueur d'onde des bandes.

- La condition sur l'emplacement des images a été imposée par la limite administrative du District d'Abidjan, l'objet de cette étude, que nous avons importée dans l'environnement GEE à partir de l'outil « Table Upload » du module Asset. Ensuite l'extraction des portions couvertes par la limite de la zone, sur la collection obtenue a été réalisée grâce à la commande clip ()
- La condition sur le type de produit s'est faite par la sélection des images Sentinel-2 de réflectance de surface (BOA). Nous avons utilisé des images de niveau 2A, qui sont déjà corrigées atmosphériquement. Dans l'environnement de l'éditeur de code GEE, les images Sentinel-2 de niveau 2A ont été importées en tant que collection d'images (c'est-à-dire un ensemble d'images du moteur Google Earth) grâce à la commande ee.ImageCollection (COPERNICUS/S2_SR)
- La condition sur la période a été fixée, allant du 01 janvier au 31 Mars 2020 dans GEE grâce à la commande filterDate (start, end)
- La condition sur la couverture nuageuse a été faite en imposant une valeur seuil de couverture nuageuse égale à zéro, pour s'assurer d'avoir des données disponibles sur l'ensemble de la zone. Nous avons utilisé la valeur du pourcentage de pixels nuageux, stockée sous forme de métadonnées pour chaque image, pour sélectionner et extraire uniquement les images avec une couverture nuageuse égale à 0% dans la zone d'étude, grâce à la commande filterMetadata ('CLOUD_COVERAGE_ASSESSMENT')
- Enfin la condition sur les bandes à utiliser, s'est faite en tenant compte de la résolution spectrale (longueur d'onde) et la résolution spatiale (valeur du plus petit pixel au sol) de nos différentes bandes. Nous avons appliqué un filtre sur les bandes pour extraire uniquement les bandes du multispectral, rééchantillonnées à une résolution spatiale de 10m (B2, B3, B4, B8, B11, B12], jugées largement suffisantes pour l'identification des unités d'occupation du sol de la zone d'étude, grâce à la commande select (bande)

Ces images prétraitées depuis Google earth engine sont ensuite utilisées pour cartographier l'occupation du sol par les algorithmes de machine Learning (AML)

3.2.2 PREPARATION DES DONNEES D'ENTREES POUR LA CLASSIFICATION PAR LES ALGORITHMES DE MACHINES LEARNING

3.2.2.1 EXTRACTION DES VARIABLES D'ENTREES

Différents paramètres qui mettent mieux en évidence certaines informations ont été extraits de nos images. Il s'agit de l'analyse en composante principale (ACP), des indices spectraux, et des produits dérivés du MNT. Les compositions colorées avec les bandes brutes (ETM4-5-3) et les néocanaux de l'analyse en composantes principales (ACP1-2-3) réalisées en complément des indices ont permis de:

- caractériser les différents types de végétation par rapport à l'intensité de leur activité photosynthétique, leur recouvrement au sol et leur humidité ou niveau de stress hydrique pendant la saison sèche. Il a été ainsi possible de discriminer les forêts et les zones agricoles
- discriminer les différents types d'occupation du sol;
- sélectionner les sites à visiter et orienter sur le terrain;
- choisir des parcelles d'entraînement pour la classification dirigée;
- et choisir les parcelles de validation pour l'élaboration de la matrice de confusion

3.2.2.1.1 L'ANALYSE EN COMPOSANTES PRINCIPALES (ACP)

En appliquant l'analyse en composantes principales (ACP), l'information contenue dans les différentes bandes va être réduite, en un nombre plus restreint de composantes. La composition colorée des trois premières composantes de l'ACP a permis d'augmenter le contraste entre les différentes unités d'occupation sol.

Seule les trois premières bandes (pc1, pc2, pc3) qui regroupent le maximum d'informations ont été retenues.

3.2.2.1.2 INDICES SPECTRAUX

Pour différencier les différents types d'unités d'occupation du sol, les indices spectraux appropriés ont été calculés. Les principaux indices utilisés dans la présente étude sont donnés ci-dessous:

- $NDVI = (PIR - R) / (PIR + R)$ Rouse et al. (1974);
- $GCI = image1.expression('float(((NIR) / (GREEN)) - 1)');$
- $RGR = image.expression('float((RED) / (GREEN))');$
- Indice normalisé de différence des eaux (NDWI): $NDWI = (NIR - SWIR) / (NIR + SWIR);$
- Indice de brillance du sol (SBI): $SBI = 0.4328 (Green) + 0.6490 (Red) + 0.4607 (NIR);$
- Indice de végétation ajusté au sol (SAVI): $SAVI = ((NIR - Red) / (NIR + Red + L)) \times (1 + L);$
- Indice visible de résistance atmosphérique (VARI): $VARI = (Green - Red) / (Green + Red - Blue);$
- Indice d'eau de différence normalisée modifiée (MNDWI): $MNDWI = (Green - SWIR) / (Green + SWIR);$
- Indice de Brillance (IB): $IB = A = \sqrt{NIR^2 + Red^2};$
- Indice d'humidité de différence normalisée (NBRI): $NBRI = (NIR - SWIR1) / (NIR + SWIR1)$

3.2.2.1.3 PRODUITS DERIVES DU MNT

La distribution des unités d'occupation du sol est fortement influencée par la topographie. Le relief est aussi un facteur important puisque certaines cultures sont disposées en fonction de l'altitude. Afin de prendre en compte le relief, nous avons ainsi utilisé le MNT du satellite PolarisPro de 12.5m de résolution spatiale rééchantillonnée à 10m, ainsi que les pentes dérivées de ce MNT.

La figure 4 présente les images issues des différentes opérations de calcul d'indices et transformation d'images ayant servi à l'identification des classes d'entité d'occupation du sol du District d'Abidjan.

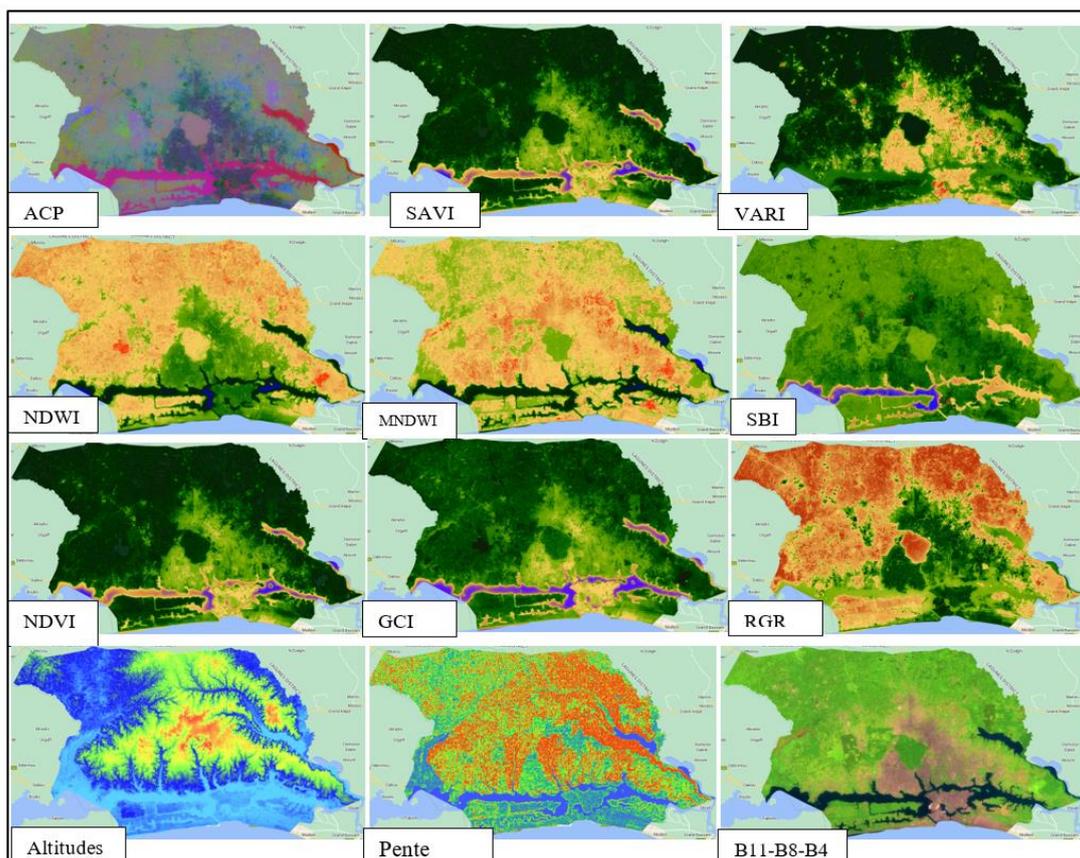


Fig. 4. Aperçu des paramètres d'entrées du classifieur, générés à partir des images multispectrales Sentinel-2 A depuis GEE, caractérisant l'état de surface du District d'Abidjan

3.2.2.2 NORMALISATION DES BANDES

Pour l'apprentissage automatique, il est recommandé de normaliser ou de standardiser les caractéristiques des bandes. Ce principe consiste à mettre à l'échelle les caractéristiques dans la même tranche de valeur. La normalisation des bandes fait référence à un processus visant à harmoniser les valeurs des différentes bandes d'une image, généralement dans le domaine

de la télédétection ou de l'imagerie satellitaire. L'objectif est de réduire les variations de luminosité et de contraste entre les bandes, afin de faciliter l'analyse et l'extraction d'informations. Cela permet d'obtenir une représentation plus cohérente et équilibrée des informations contenues dans chaque bande, facilitant ainsi les analyses ultérieures telles que la classification, la détection des changements, ou l'extraction de caractéristiques.

Une fonction de normalisation qui consiste à diviser toutes les bandes par la valeur de la bande max a été implémentée dans GEE, et appliquée à l'image composite grâce à la syntaxe:

normalized = image.subtract (mins).divide (maxs.subtract (mins)).

3.2.3 DISCRIMINATION, VERIFICATION TERRAIN ET NOMENCLATURE DES CLASSES D'OCCUPATION DU SOL

La composition colorée (MIR1-PIR-MIR2) retenue, associée aux indices, aux photos de terrain et aux prises de vue aérienne de 2020, qui ont fait l'objet d'une observation directe, nous ont permis de catégoriser et nommer les différentes classes d'occupation du sol dans District d'Abidjan.

En superposant, dans l'interface de QGIS Desktop, les différentes couches MSI (11,8,12) et les prises de vues aériennes de très haute résolution, 296 points de contrôle ont été collectés manuellement au total et étiquetés par interprétation visuelle. L'identité de ces différents points sera ensuite justifiée par une connaissance approfondie de la zone et les vérités terrains (enquêtes in situ).

La collecte des données de terrain pour la validation s'est faite du 26 février au 02 Mars 2022 et a porté sur 4 sites principaux présentant des échantillons représentatifs de l'ensemble de la zone d'étude (Figure 5).

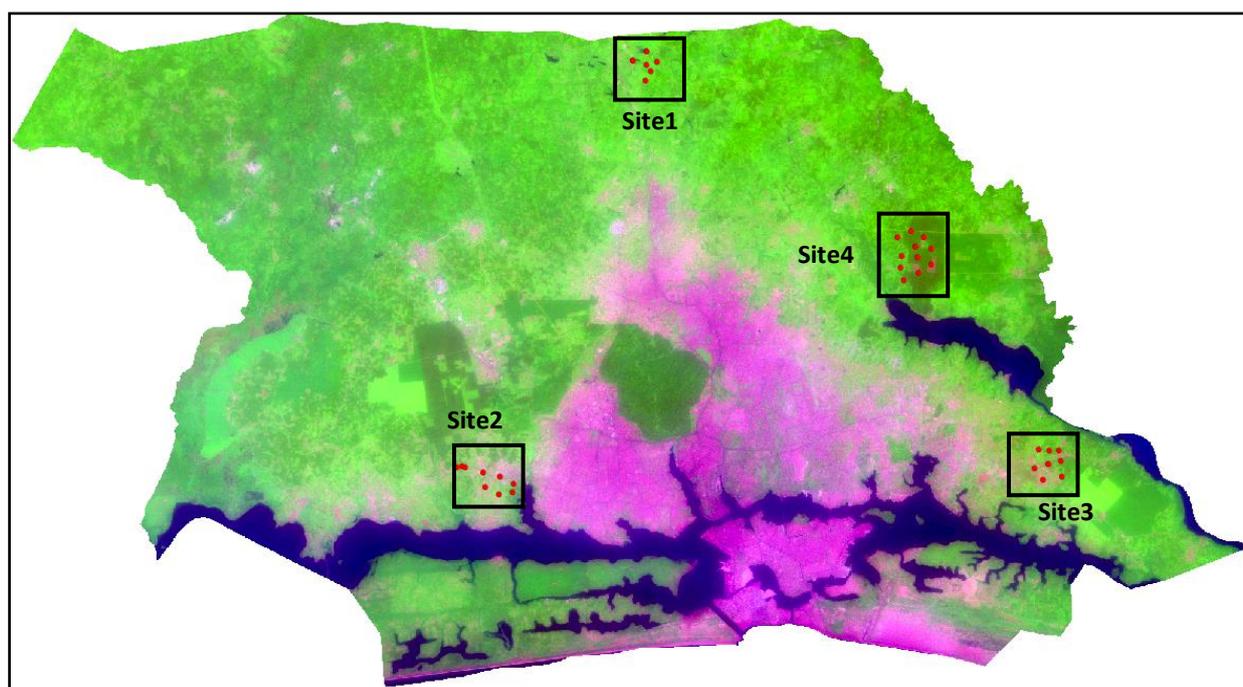


Fig. 5. Sites d'échantillonnage pour les points de contrôle (vérité terrain)

La visite a consisté à:

- localiser et identifier les points à vérifier (unités d'occupation du sol) à l'aide d'une application de collecte de données mobiles développées à partir de kobotoolbox;
- faire une description de l'unité d'occupation du sol (catégorie, type) visité, conformément aux questionnaires du formulaire;
- densifier les échantillons par la collecte de nouveaux échantillons;
- relever de façon automatique la position géographique, et la description exacte de l'unité d'occupation du sol vérifié, pour mieux consolider la validation

La nomenclature des unités d'occupation du sol du district d'Abidjan s'est appuyée sur la base de données du projet national de surveillance spatiale des terres (SST) en Côte d'Ivoire, définie parfaitement compatible au Land Cover Classification System

(LCCS) et conforme au GIEC. Le LCCS est un système de classification des couvertures terrestres utilisé pour catégoriser et cartographier les différents types de couverture des terres, tels que les forêts, les zones urbaines, les terres agricoles et les zones humides. Neuf (09) classes d'occupation du sol parfaitement différenciables ont été retenues pour la cartographie des unités d'occupation du sol du District d'Abidjan. Il s'agit (1) des Forêt; (2) des Forêts marécageuses; (3) des Plantations forestières; (4) des Formations arbustives; (5) des Hévées; (6) des Palmiers; (7) des Aménagements agricoles et autres cultures; (8) des Cours d'eau; et (9) des Habitats/Sols nus (Tableau 2).

Tableau 2. Nomenclature des classes d'occupation du sol dans le District d'Abidjan, conformément projet national de surveillance spatiale des terres (SST)

Catégorie	Classe	Description
Forêt	Forêt dense	Formation naturelle de type primaire, peuplement fermé avec des arbres et des arbustes atteignant différentes hauteurs entre 5 et 50 m
	Forêt marécageuse	Formations forestières (couverture supérieure à 30 %) établies sur des sols hydromorphes à proximité de cours d'eau
	Plantation forestière	Parcelles boisées ou régénération environnementale; hauteur des arbres supérieure à 5 m et taux de couverture supérieure à 30 %.
	Formation arbustive	Formation végétale constituée d'arbustes, de moins de 5 m de haut
Cultures	Hévéa	Plantations d'hévéas de 30 m de haut; à feuilles caduques
	Palmier	Plantations d'espèces de palmiers: famille des cocotiers de 5 à 25 m de haut, surmontées d'une large couronne de conifères.
	Aménagements agricoles et autres cultures	Superficies des autres terres avec des cultures indifférenciées et des vergers divers, y compris les cultures de plaine et les jeunes jachères âgées de 3 à 7 ans.
Eau	Cours/Plan d'eau	Plan d'eau et cours d'eau: Groupes de zones de dépression contenant de l'eau sous la forme d'un réservoir et d'un flux d'eau interconnecté (lagune, réservoir) et Réseau hydrographique linéaire interconnecté, se dirigeant vers un plan d'eau
Urbain	Habitats/sols nus	Zones résultant des activités humaines, autres qu'agricoles, pour l'urbanisation et les constructions diverses. Terrains dépourvus de couverture végétale naturelle (herbacée, arbustive, boisée), ne constituant pas une zone cultivée.

Une mission de validation terrain nous a permis d'attribuer à chaque unité d'occupation du sol ayant une tâche homogène au sol de 10m (résolution spatiale de l'image sentinel2), sa réflectance spectrale sur l'image.

La superposition de l'image sentinel2 prétraitée, aux prises de vues aériennes et aux photos de terrain, nous a permis de différencier et de valider les classes d'OCS illustrées à la figure 6.



Fig. 6. Identification et Nomenclature des classes d'occupation du sol dans le District d'Abidjan

3.2.4 CODIFICATION ET REPARTITION DES POINTS D'ENTRAINEMENT ET DE VALIDATION

Pour améliorer la qualité des résultats de la classification finale, le choix des points d'entraînement et de validation est l'une des étapes les plus délicates de tout le processus [20]. La méthode de classification utilisée est basée sur les pixels; par conséquent, les points de contrôle (entraînement et validation) représentent les classes auxquelles ils appartiennent, évitant ainsi les problèmes de pixels mélangés. Le nombre total d'échantillons implique la précision attendue et l'incertitude de l'estimation. En raison de la possibilité de biais dans nos données d'échantillon, le nombre de points d'échantillonnage a été constamment ajusté, afin de réduire autant que possible les erreurs de classification.

Afin d'implémenter nos points de contrôle dans l'environnement GEE, nous avons procédé ensuite à une codification des étiquettes des différentes classes à 2 caractères (Tableau 3).

Tableau 3. Échantillon de données sélectionné dans cette étude

Occupation du sol	Codification	Échantillons
Forêt	Fo	17
Forêt marécageuse	FM	29
Plantation forestière	PF	40
Formation arbustive	FA	29
Hévéa	He	49
Palmier	Pa	35
Aménagements agricoles et autres cultures	AG	45
Eau	Ea	23
Urbain	Ur	29

Les pseudocodes sont créés en attribuant un code de type numérique à la classe correspondante qui prend la valeur texte.

Avec les algorithmes d'apprentissage, les 296 échantillons collectés sont divisés en ensemble de données d'entraînement et en ensemble de données de validation. Cela facilite l'utilisation des données d'échantillon aussi efficacement que possible en l'absence de données de validation indépendantes.

Afin de diviser les données en deux ensembles, la fonction `randomColumn` de GEE a été utilisée pour générer un nombre aléatoire dans tous les points d'échantillonnage, les valeurs aléatoires étant comprises entre (0~1). Par conséquent, toutes les données de l'échantillon génèrent une valeur aléatoire supplémentaire:

- 70% des échantillons sont retenus comme l'ensemble d'entraînement grâce à la commande `ee.Filter.lt ('random', 0.7)`
- et le reste de valeurs des échantillons 30% sont pris comme l'ensemble de validation grâce à la commande `ee.Filter.gte ('random', 0.7)`

Enfin, l'ensemble d'entraînement a été utilisé pour l'entraînement, et l'ensemble de validation a été utilisé pour valider la performance des différents algorithmes.

3.2.5 CLASSIFICATION DES UNITES D'OCCUPATION DU SOL PAR LES ALGORITHMES MACHINE LEARNING

L'étape de classification proprement dite consiste à exécuter le scripts Java de classification des données d'entrée par les algorithmes de ML dans l'éditeur de code de GEE et la validation de la classification. Le code java développé est organisé selon une hiérarchie à respecter:

- le jeu de données à utiliser générer précédemment à l'étape du choix de la meilleure image d'entrée. L'image composite (PC1, PC2, PC3, B2, B3, B4, B8, B11, B12, NDVI, NDWI, MNDWI, VARI, SBI, SAVI, GCI, RGR, Pente) est implémentée dans le script comme données d'entrée classifieur;
- la collection de caractéristiques contenant toutes les données d'entraînement étiquetées avec des codes correspondant aux classes d'occupation du sol. Les points de contrôle d'entraînement et de validation sont ensuite introduits dans le script. Pour la formation du modèle, des échantillons au format «.shp » ont été créés pour les 9 classes. Les 70% de ces échantillons ont été utilisés pour entraîner les classifieurs et 30% ont servi pour la validation. Lors du processus d'entraînement, la colonne contenant les codes numériques sont indiqués à l'algorithme comme les classes référentielles d'arbres aléatoires;
- le paramétrage des différents algorithmes de classification;
- la légende implémentée avec les codes couleurs propre à chaque classe, pour une meilleure visualisation des résultats;
- l'application de la légende du LCCS permet à tous ceux qui veulent utiliser la carte de comprendre les éléments de la légende et leurs définitions en consultant le guide LCCS [21]. Google Earth engine a permis d'implémenter la légende de toutes les classes à cartographier. Les critères de sémiologie sont définis de telle sorte que toutes les classes soient clairement visible et interprétable
- les statistiques et graphes pour l'évaluation quantitative et qualitative des différentes classifications ont été générées: matrice de confusion, PG, Kappa, PU, PP, superficie,
- le modèle est ensuite sauvegardé pour la deuxième phase du processus de l'algorithme

Après avoir exécuté le script de classification, l'image est rapidement classifiée par une définition préliminaire des différents classificateurs (RF, SVM, CART, MD, NB). Un processus de validation est ensuite appliqué aux différentes cartes obtenues.

3.2.6 CHOIX DE LA MEILLEURE CLASSIFICATION

Le choix de la meilleure classification s'est fait en comparant les valeurs des paramètres d'estimation de la précision obtenues pour chaque classification et par l'analyse de la séparabilité des classes.

3.2.6.1 ÉVALUATION DE LA CLASSIFICATION

Différentes métriques nous ont permis de mieux juger de la qualité de nos prédictions de classification: Matrice de confusion, Précision globale, Précision du producteur, Précision du consommateur (fiabilité) et le Coefficient de Kappa).

- **Matrice de confusion**

La matrice de confusion de la classification offre une mesure quantitative de la qualité de l'échantillonnage et la séparabilité des classes. Cette évaluation de la qualité du résultat est utile pour comprendre l'origine des confusions et ajuster les efforts de photo-interprétation.

Pour chaque algorithme de classification effectué, la matrice de confusion a été réalisée depuis GEE grâce à l'algorithme `ee.Classifier.confusionMatrix ()`.

- **Précision du producteur (PP)**

La Précision du producteur (producer's accuracy) est la probabilité pour un pixel de la classe C_k sur le terrain d'être bien classé en C_k sur la carte, dans quelle mesure la classification a prédit chaque classe. La fonction `confusionMatrix ().producersAccuracy ()`, nous a permis de déterminer la PP des différentes classifications;

- **Précision de l'utilisateur (PU)**

Précision de l'utilisateur (user's accuracy) est la probabilité pour un pixel classé en C_k sur la carte d'appartenir à C_k sur le terrain. L'algorithme `confusionMatrix ().consumersAccuracy ()`, nous a permis de déterminer la PU des différentes classifications;

- **Précision globale (PG)**

La PG est le pourcentage total de classification, donné par le rapport entre le nombre d'unités correctement classées et leur nombre total [22]. La PG a été générée automatiquement depuis GEE grâce à l'algorithme `confusionMatrix ().accuracy ()`;

- **Coefficient Kappa**

Le coefficient kappa, permet d'évaluer globalement:

- la fiabilité des résultats de la classification par rapport aux données de référence;
- et la performance de la classification par rapport à l'assignation aléatoire

L'algorithme `confusion Matrix ().kappa ()`, nous a permis de déterminer le coefficient Kappa des différentes classifications;

3.2.6.2 SEPARABILITE DES ALGORITHMES

La séparabilité des classes se caractérise soit par des confusions ou des différenciations des éléments de classes différentes. Afin de montrer plus spécifiquement les différences de classification des algorithmes SVM, RF, CART, MD et NB, dans différentes régions et classes, trois sous-régions de 5×5 km (a, b, et c) ont été sélectionnées au hasard à partir des résultats de classification des trois meilleurs algorithmes afin d'examiner les erreurs de classification des pixels des différentes unités d'occupation du sol obtenus.

Parallèlement, afin de valider l'exactitude des cinq algorithmes dans l'identification des pixels, les classes d'occupation du sol obtenues sont comparées aux fonds de cartes libres de très haute résolution (Google Map).

Grâce à l'analyse visuelle des erreurs de classification des trois sous-régions différentes, nous pourrions déterminer quel algorithme distingue au mieux les pixels des différentes classes.

4 RESULTATS ET DISCUSSION

4.1 RESULTATS

4.1.1 PRETRAITEMENT DE L'IMAGE COMPOSITE

La Figure 7 présente le résultat du prétraitement de l'image composite Sentinel-2A, pour la classification de l'OCS à partir de GEE. Les intervalles de valeurs des bandes résultantes normalisés sont dans la même frange et la résolution de toutes les bandes est de 10m.

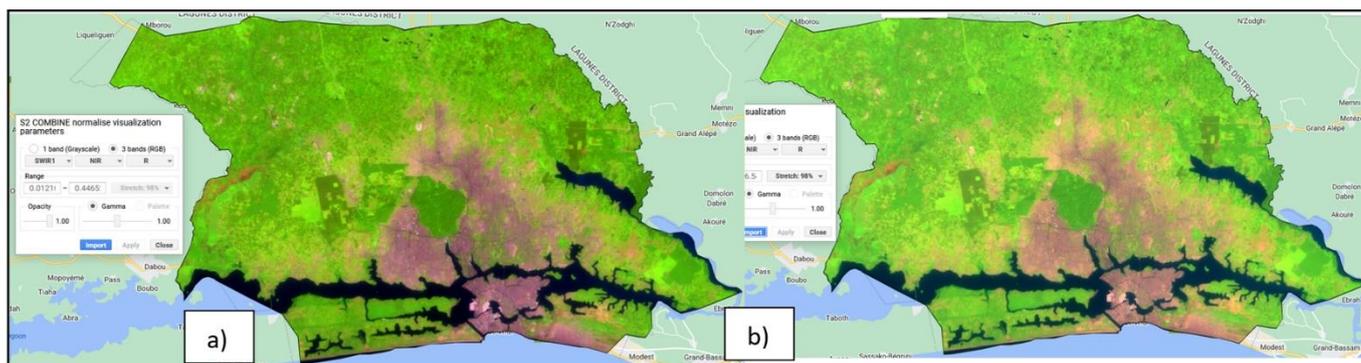


Fig. 7. Visualisation sur GEE du composite normalisé Mosaic Sentinel-2A (a) et le composite Mosaic Sentinel-2A (b) en composition colorée fausse couleur (SWIR1, NIR, R)

4.2 IMPORTANCE DES BANDES

La Figure 8 présente l'apport de chaque bande dans la classification supervisée des unités d'occupation du sol par les algorithmes de Machine Learning (Exemple du RF).

Cette figure met en évidence l'apport de chaque indice dans la cartographie des unités d'occupation du sol par les AML, en particulier les bandes du multispectral, les produits dérivés du MNT et les indices spectraux: B + G + R + NIR + SWIR1 + SWIR2 + NDVI + NDWI + SBI + SAVI, VARI + GCI + RGR + ACP1 + ACP2 + ACP3 + pente + relief.

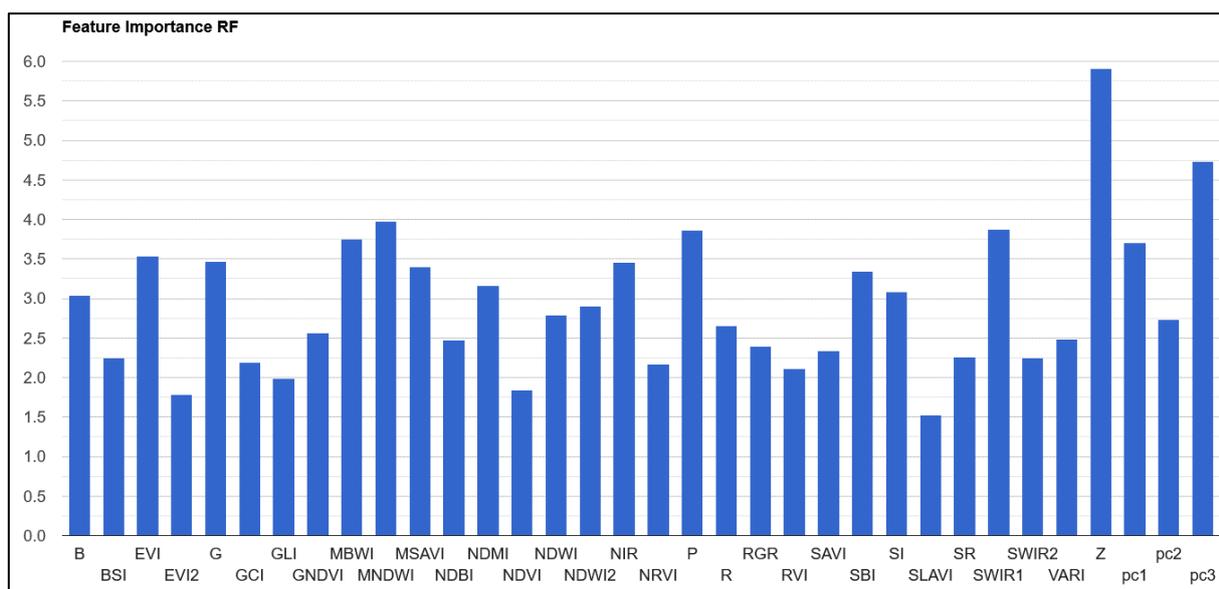


Fig. 8. Importance des bandes dans la classification par l'AML SVM depuis GEE

4.2.1 CLASSIFICATION SUPERVISEE DE L'OCCUPATION DU SOL DU DISTRICT D'ABIDJAN PAR LES AML SVM RF, CART, MD ET NB

Les résultats cartographiques montrent que neuf (09) classes d'unités d'occupation du sol, Forêt (F), Plantations Forestières ou reboisement (PF), Forêt Marécageuse (FM), Bananeraie (B), Hévée (H), Palmier à huile (PH), Aménagements Agricoles et autres cultures (AA), Eau (E), et Milieu Urbain (MU) sont représentées dans des proportions différentes dans le District d'Abidjan en fonction de l'algorithme choisie (Figure 9).

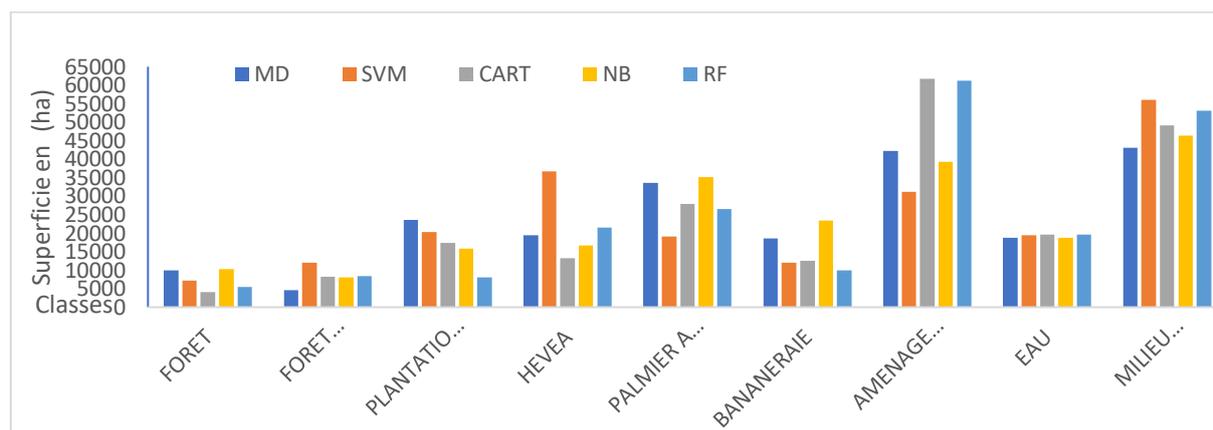


Fig. 9. Répartition des superficies de classe d'occupation du sol en 2020 dans le District d'Abidjan en fonction des algorithmes SVM, RF, CART, MD et NB

L'analyse visuelle des cartes produites à partir des images Sentinel-2 et de la classification par les algorithmes de machine learning RF, CART, SVM, MD et NB nous renseigne sur la répartition spatiale des catégories d'occupation du sol dans le District d'Abidjan en 2020.

Les cartes ci-après (Figure 10), illustrent les résultats issus de la classification supervisée par les algorithmes SVM, RF, CART, MD et NB des images Sentinel 2A en 2020, dans le District d'Abidjan.

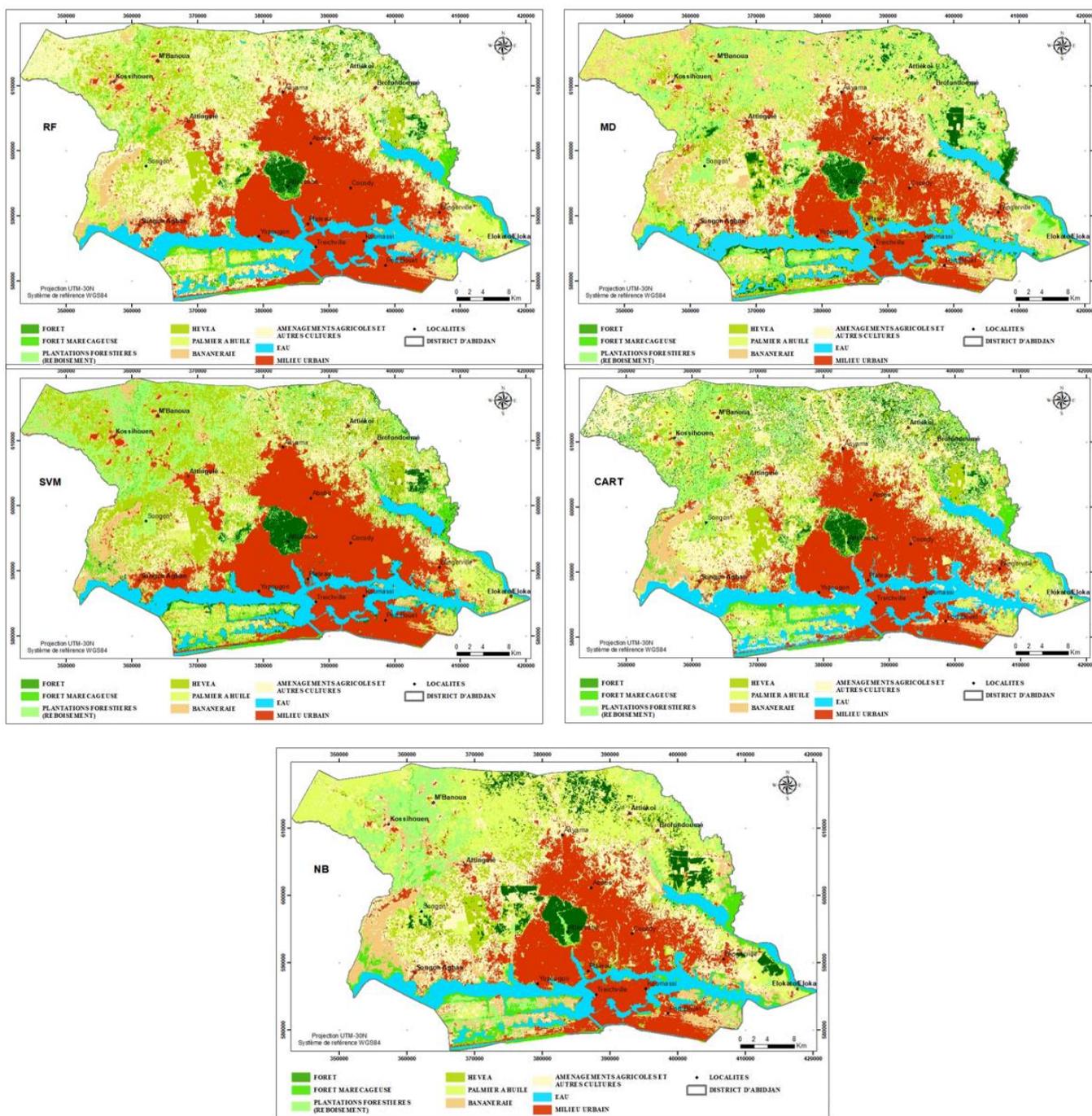


Fig. 10. Carte d'occupation du sol du District d'Abidjan de 2020, par les algorithmes de Machine Learning SVM, CART, RF MD

Les Aménagements Agricoles et autres cultures sont dominantes dans le District d'Abidjan en exploitant les AML RF et CART avec des proportions respectivement égales 28.6% et 28.8%. Quant aux Milieu Urbain (MU), elles dominent dans la zone d'étude en exploitant les algorithmes MD, SVM et NB avec des proportions respectivement égales 20.1%, 26.2%, 21.7%. Ses proportions avoisines les proportions de la classe Milieu Urbain obtenues par les AML RF et CART respectivement égales à 24.8 % et 23.0% (Tableau 4).

Tableau 4. Statistiques des classes d'occupation du sol du District d'Abidjan en 2020, obtenus par les AML CART, RF, SVM, MD et NB

Classe	Superficie en ha					Proportion en %				
	MD	SVM	CART	NB	RF	MD	SVM	CART	NB	RF
FORET	9927.4	7213.4	4018.3	10228.9	5400.1	4.6	3.4	1.9	4.8	2.5
FORET MARECAGEUSE	4538.1	11958.8	8156.2	8039.7	8464.2	2.1	5.6	3.8	3.8	4.0
PLANTATIONS FORESTIERES	23614.7	20305.2	17412.0	15869.1	8130.5	11.0	9.5	8.1	7.4	3.8
HEVEA	19379.9	36708.1	13278.5	16776.3	21487.4	9.1	17.2	6.2	7.8	10.0
PALMIER À HUILE	33630.1	19045.3	27939.9	35102.9	26510.3	15.7	8.9	13.1	16.4	12.4
BANANERAIE	18609.4	12007.7	12620.0	23439.3	9947.6	8.7	5.6	5.9	11.0	4.7
AMENAGEMENTAGRICOLAS	42288.5	31136.4	61670.2	39228.1	61207.2	19.8	14.6	28.8	18.3	28.6
EAU	18844.7	19385.3	19570.2	18718.7	19672.7	8.8	9.1	9.2	8.8	9.2
MILIEU URBAIN	43021.3	56093.9	49188.9	46451.1	53034.1	20.1	26.2	23.0	21.7	24.8
TOTAL	213854	213854	213854	213854	213854	100	100	100	100	100

• **AML RF**

Les estimations obtenues à partir de l'algorithme RF indiquent que le District d'Abidjan est dominé par la la classe Aménagement agricole qui représente 28.6% (61207.2 ha) de sa superficie totale. Il renferme 2.5% (5400.1ha) de Forêt, 3.8% (8130.5 ha) de Plantations Forestières ou reboisement; 4% (8464.2 ha) de Forêt Marécageuse, 4.7% (9947.6 ha) de Bananeraie, 10% (21487.4 ha) d'Hévéa, 12.4% (26510.3 ha) de Palmier à huile, 9.2% (19672.7 ha) d'Eau; et 24.8 (53034.1 ha) de Milieu Urbain.

• **AML CART**

Les estimations obtenues à partir de l'algorithme CART indiquent que District d'Abidjan renferme plus de 28.8% soit 61670.2 ha d'Aménagement agricole. 1.9% (4018.3 ha) de Forêt, 3.8% (17412.0 ha) de Plantations Forestières ou reboisement; 4.0% (8156.2 ha) de Forêt Marécageuse, 5.9% (12620.0 ha) de Bananeraie, 6.2% (13278.5 ha) d'Hévéa, 13.1% (27939.9 ha) de Palmier à huile, 9.2% (19570.2 ha) d'Eau; et 23.0 (49188.9ha) de Milieu Urbain. La classe dominante est la classe Aménagement agricole.

• **AML SVM**

Les estimations obtenues à partir de l'algorithme SVM indiquent que District d'Abidjan est dominé par le Milieu Urbain soit 26.2% (56093.9 ha) de sa superficie totale. On observe également 3.4% (7213.4ha) de Forêt, 9.5% (20305.2 ha) de Plantations Forestières ou reboisement; 5.6% (11958.8 ha) de Forêt Marécageuse, 5.6% (12007.7ha) de Bananeraie, 17.2% (36708.1ha) d'Hévéa, 8.9% (19045.3 ha) de Palmier à huile, 9.1% (19385.3 ha) d'Eau; et 14.6% (31136.4 ha) d'Aménagement agricole.

• **AML MD**

Comme pour l'AML SVM, les estimations obtenues à partir de l'algorithme MD indiquent que District d'Abidjan est dominé par le Milieu Urbain soit 20.1% (56093.9 ha) de sa superficie totale. Il de renferme 4.6% (9927.4 ha) de Forêt, 11% (23614.7ha) de Plantations Forestières ou reboisement; 2.1% (4538.1 ha) de Forêt Marécageuse, 8.7% (18609.4 ha) de Bananeraie, 9.1% (19379.9 ha) d'Hévéa, 15.7% (33630.1 ha) de Palmier à huile, 8.8% (18844.7 ha) d'Eau; et 19.8% (42288.5 ha) d'Aménagement agricole.

• **AML NB**

Les estimations obtenues à partir de l'algorithme NB indiquent que District d'Abidjan renferme plus de 18.3% soit 39228.1 ha d'Aménagement agricole. 4.8% (10228.9 ha) de Forêt, 7.4% (15869.1 ha) de Plantations Forestières ou reboisement; 3.8% (8039.7ha) de Forêt Marécageuse, 11% (23439.3 ha) de Bananeraie, 7.8% (16776.3 ha) d'Hévéa, 16.4% (35102.9 ha) de Palmier à huile, 8.8% (18718.7 ha) d'Eau; et 21.7 (46451.1 ha) de Milieu Urbain qui est la classe dominante.

4.2.2 CHOIX DE LA MEILLEURE CLASSIFICATION

Le choix de la meilleure classification s'est fait en comparant les valeurs des paramètres d'estimation de la précision obtenues pour chaque classification et par l'analyse de la séparabilité des classes.

4.2.2.1 VALIDATION DES CLASSIFICATIONS

Les résultats de la classification ont été soumis à une évaluation par l'analyse de la précision du producteur, la précision de l'utilisateur, la précision globale, le coefficient Kappa et la séparabilité des classes.

- **Precision du producteur et de l'utilisateur**

Les résultats de la classification ont été soumis à une évaluation de la précision après classification. La précision de l'utilisateur et du producteur de la classification par les algorithmes SVM, RF, CART, MD et NB ont été générées (Tableau 5).

Tableau 5. Évaluation de la précision de classification des unités d'occupation du sol de 2020 par les algorithmes de machine learning SVM RF, CART, MD et NB

Classes	MD		NB		RF		SVM		CART	
	PP	PU	PP	PU	PP	PU	PP	PU	PP	PU
F	89%	44%	78%	50%	70%	78%	100%	80%	83%	63%
PF	20%	9%	20%	11%	63%	100%	88%	88%	71%	63%
FM	30%	100%	20%	40%	78%	64%	64%	88%	86%	100%
B	79%	69%	36%	71%	78%	90%	83%	83%	87%	100%
H	30%	100%	50%	83%	63%	83%	86%	75%	83%	63%
PH	14%	17%	50%	28%	86%	63%	100%	92%	50%	50%
AA	38%	27%	38%	33%	71%	71%	67%	71%	45%	63%
E	100%	100%	92%	100%	100%	100%	93%	100%	100%	100%
MU	90%	96%	90%	100%	100%	94%	97%	95%	97%	92%

Un examen attentif du tableau XXII a révélé quelques confusions entre certaines classes:

- Les classes Eau et Milieu urbain, sont parfaitement classifiées avec l'utilisation des 5 algorithmes SVM, RF, CART, MD et NB (Précision >81%);
- Les classes correspondant à la végétation sont mieux classifiées par les classifieurs SVM, RF et CART. Par contre, il existe des confusions entre les classes de végétation pour les algorithmes MD et NB. On observe des erreurs d'omission les plus élevées allant respectivement de 31% à 91% et 50% à 72% pour MD et MB, soit une précision de l'utilisateur qui va de 69% à 9% et 50% à 28%. Également des erreurs de commission les plus élevées allant respectivement de 86% à 62% et 50% à 89% pour MD et MB, soit une précision du producteur qui va de 14% à 38% et 50% à 11%

Malgré une différence légère entre la précision du producteur et de l'utilisateur (Figure 11), les pixels ont été bien classifiés pour chacune des classes définies par les algorithmes CART, SVM et RF (valeur >60%), et moyennement précises pour les algorithmes MD et NB.

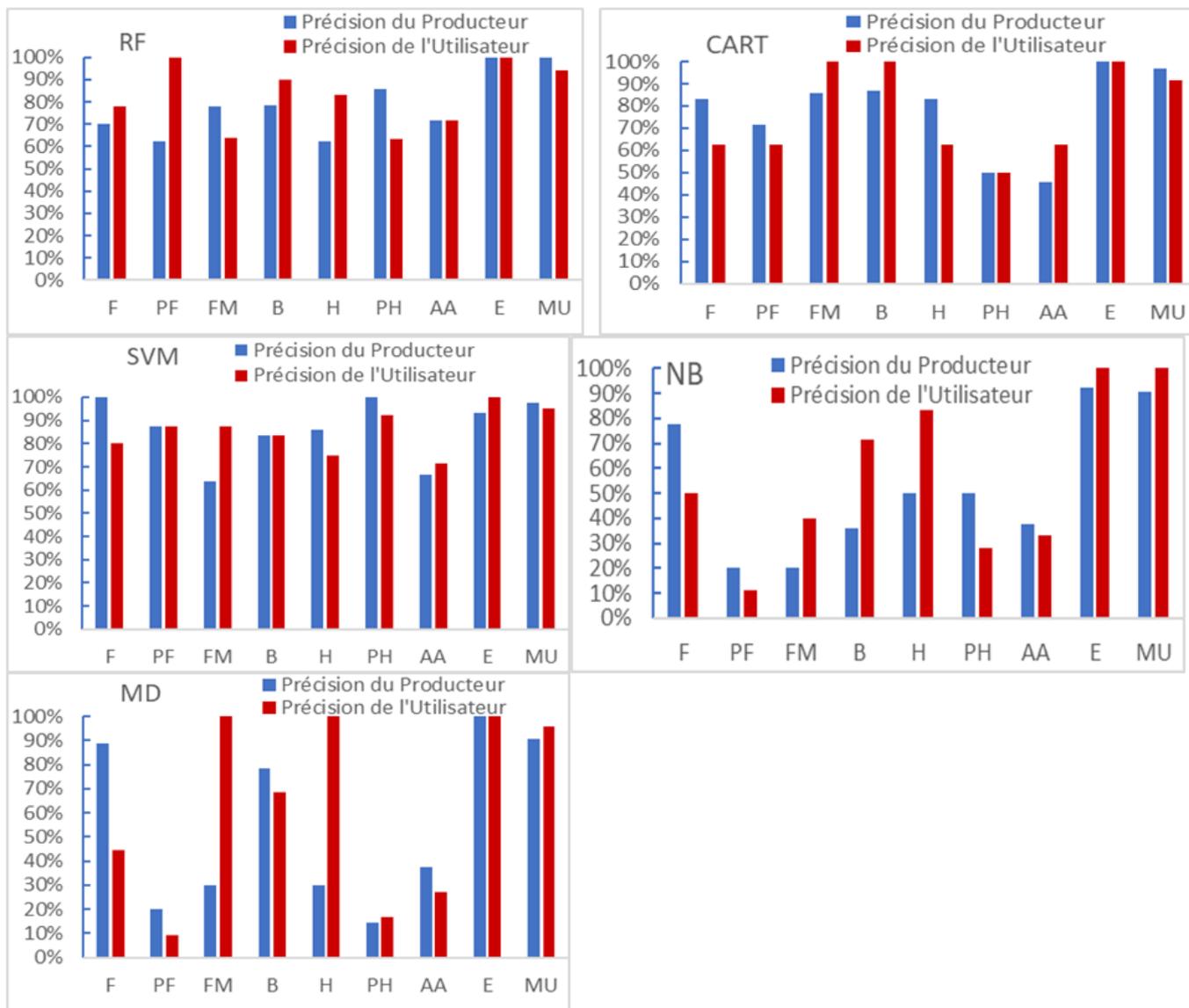


Fig. 11. Précision des producteurs vs Précision de l'utilisateur de la classification supervisée des unités d'occupation du sol du District d'Abidjan de 2020 à partir des images Sentinel 2A par les algorithmes de machine Learning SVM RF CART MD et NB

• Précision globale et Kappa

Le Tableau 6 présente la précision globale de la classification par les algorithmes de machine Learning RF, CART, SVM, MD, NB, qui correspond au total de pixels bien classifiés par rapport au total de pixels dans la matrice d'erreur.

Tableau 6. Évaluation de la précision de classification des différents algorithmes par la Précision globale et le coefficient Kappa

Algorithme	Kappa	Précision globale %
RF	0.94	95.54
SVM	0.91	94.10
CART	0.88	86.25
MD	0.68	72.60
NB	0.66	71.00

4.2.2.2 SEPARABILITE DES ALGORITHMES

La superposition des cartes d'occupation du sol obtenues à partir des 3 meilleurs algorithmes SVM, CART, et RF avec d'autres sources externes de spatio-carte de haute résolution reflétant la réalité terrain nous a permis de mieux apprécier nos résultats de classification (Figure 12).

On observe une meilleure concordance d'informations avec les résultats du classifieur RF.

La carte d'occupation du sol qui reflète le mieux la réalité terrain est la carte d'occupation qui résulte de la classification à partir de l'AML RF.

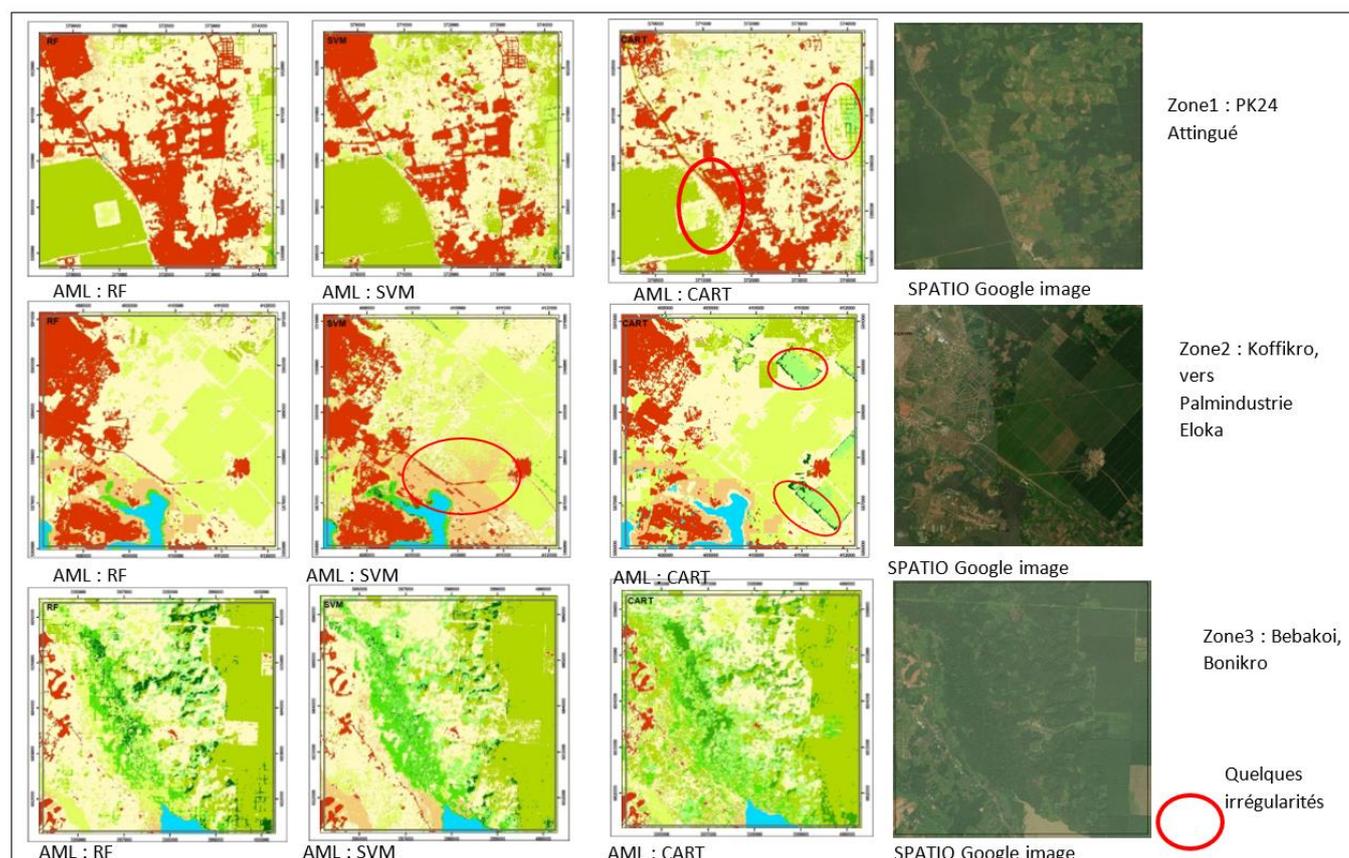


Fig. 12. Séparabilité des classes d'OCS obtenues à partir des algorithmes de AML SVM, RF, CART et Spatio-images

4.3 DISCUSSION

Les images générées par le capteur MSI de Sentinel-2, à différents niveaux de traitement, ont prouvé leur efficacité dans plusieurs études de classification supervisée de l'occupation du sol [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35].

La méthode la plus adaptée pour classifier les unités d'occupation du sol de la zone d'étude consiste à utiliser une classification supervisée basée sur des zones d'entraînement déterminées par la photo-interprétation et une mission de terrain, car cela permet une proximité avec le terrain et la possibilité de travailler avec des personnes ayant une grande connaissance de la zone étudiée [24], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46].

Bien que l'ensemble des classes reprenne la quasi-totalité des types d'occupation du sol de la zone, nous avons choisi délibérément de limiter le nombre de classes à neuf (09) selon leur proximité spectrale, une approche qui a été prouvée efficace par plusieurs auteurs, afin d'éviter une complexité excessive de la carte [28], [34], [36], [41], [44].

Au niveau de la méthodologie utilisée pour classifier les unités d'occupation du sol dans le District d'Abidjan, le choix s'est porté sur l'utilisation des algorithmes de machine learning et du cloud computing grâce à la technologie de Google Earth Engine. Cela permet de réaliser des traitements à petite et grande échelle d'étude, tout en économisant du temps et de l'espace de stockage. Les différentes étapes de la méthodologie incluent l'acquisition des mosaïques de données, le prétraitement des

données, l'entraînement du modèle, le post-traitement des classifications et l'export sous différents formats (shapefile, raster ou web). En utilisant les algorithmes de machine learning couplés à la plateforme de cloud computing Google Earth, [15] ont pu caractériser les changements de la couverture terrestre et de la dynamique forestière au Togo sur une période allant de 1985 à 2020, en se basant sur des images Landsat. De même, à grande échelle, [47] ont réussi à classier avec précision cinq grands types de cultures mondiales et à définir leurs stades de croissance en utilisant les bandes étroites hyperspectrales optimales de Earth Observing-1 Hyperion sur la plateforme Google Earth Engine. Ces exemples démontrent la pertinence de cette technique dans le processus de cartographie des unités d'occupation du sol en milieu urbain, notamment dans le District d'Abidjan qui présente un environnement complexe et hétérogène.

L'ajout des indices spectraux à l'image composite constituée des bandes multispectrales de Sentinel-2 en tant que données d'entrée a considérablement réduit les erreurs de prédiction et amélioré les performances du modèle de Machine Learning. En se limitant à l'utilisation des bandes du multispectral rééchantillonnées à 10m, la précision globale atteignait 84,85%. Toutefois, l'ajout des indices spectraux et l'analyse en composantes principales (ACP) a permis une légère amélioration de la précision, la faisant passer à 93,2%. Enfin, en incluant également la pente, la précision a encore augmenté pour atteindre 94,30%. Cette approche a permis de conclure que l'utilisation du MNT est approprié pour discriminer les classes de végétation des autres classes d'occupation du sol ayant des caractéristiques spectrales similaires dans une classification. [48] ont utilisé une approche similaire en utilisant une image composite constituée des indices NDVI, NDBI, NDWI ainsi que l'ACP pour mieux cartographier les unités d'occupation du sol à partir d'images Sentinel-2 sur le site Ramsar d'Azagny dans le sud de la Côte d'Ivoire. Cette méthode a également été employée par [49] pour cartographier le couvert végétal dans une zone de contact entre forêt et savane dans le département de Toumodi, en utilisant les images Sentinel-2 ainsi que les indices de brillance (IB), d'humidité (NDWI) et le NDVI. Cependant, l'utilisation du modèle numérique de terrain (MNT) pour différencier les unités d'occupation du sol qui ont des signatures spectrales proches s'est avérée pertinente [50].

La précision cartographique globale de 95,54% et l'indice de Kappa de 0,94 obtenus grâce à l'algorithme RF ont permis de discriminer avec succès les différents types d'occupation du sol lors de la validation. Ces résultats soulignent le potentiel de la combinaison des données Sentinel-2 avec le MNT Alos Palsar pour une cartographie détaillée. En comparant les résultats obtenus dans cette nouvelle étude avec ceux obtenus à partir de l'image LANDSAT-7 ETM+ dans d'autres études [1], qui ont obtenu un indice kappa de 0,81 en 2017, on constate que cette nouvelle étude fournit une cartographie plus détaillée et plus précise.

En examinant la matrice de confusion classe par classe, il ressort que les classes de végétation cultivée en général et la classe "Végétation naturelle en régénération (jachère)", qui ont été fusionnées avec la classe habitat/sol nu, présentent les plus grandes augmentations d'erreurs. Ces erreurs sont attribuées à la difficulté de la photo-interprétation à séparer certains objets. Par exemple, il n'est pas toujours facile de distinguer clairement la limite entre les classes "sol nu et jachère" et "forêt galerie et cours d'eau". De même, les cultures mixtes peuvent poser des difficultés. Cela dénote de la complexité et de la forte hétérogénéité des formations observées sur le territoire ivoirien soulevées par [7], [25], [28], [34], [36], [41], [44], [48], [51]

5 CONCLUSION

Cette étude visait à cartographier les unités d'occupation du sol par l'utilisation de la télédétection à très haute résolution, la technologie Google Earth Engine et les algorithmes de machine learning à l'échelle du District d'Abidjan.

Les images multispectrales optiques Sentinel-2 de 2020 ont été utilisées pour cartographier de manière précise les unités d'occupation du sol dans le District d'Abidjan en utilisant des algorithmes de machine learning, le cloud computing et des bases de données de référence libres.

Dans l'ensemble les classes Eau et Milieu urbain ont été parfaitement classifiées avec l'utilisation des 5 algorithmes SVM, RF, CART, MD et NB (Précision >81%). Les classes correspondant à la végétation ont été mieux classifiées par les classifieurs SVM, RF et CART. Par contre, il existe des confusions entre les classes de végétation pour les algorithmes MD et NB. Les erreurs d'omission les plus élevées ont été obtenues en utilisant les algorithmes MD et MB allant respectivement de 31% à 91% et 50% à 72%, soit une précision de l'utilisateur qui va de 69% à 9% et 50% à 28%. Pareil pour les erreurs de commission les plus élevées allant de 86% à 62% et 50% à 89%, soit une précision du producteur qui va de 14% à 38% et 50% à 11%. Malgré une différence légère entre la précision du producteur et de l'utilisateur, les pixels ont été bien classifiés pour chacune des classes définies par les algorithmes CART, SVM et RF (valeur >60%), et moyennement précises pour les algorithmes MD et NB. La méthode de classification Random Forest a démontré son efficacité en permettant de cartographier neuf (09) catégories d'occupation et d'usage des terres avec une précision globale de 95,54% et un indice kappa de 0,94.

L'analyse de la répartition des classes d'occupation du sol obtenues à partir de l'AML RF, permis de relever que les unités d'occupation du sol sont réparties comme suit dans le District d'Abidjan: 28,07% de milieu urbain, 25,35% d'aménagements agricoles et autres cultures, 12,39% de palmiers à huile, 10,05% d'hévéa, 4,66% de bananeraie, 2,53% de forêts, 3,96% de

forêts marécageuses, 3,80% de plantations forestières (reboisement) et 9,2% d'eau. Ces statistiques indiquent que les aménagements agricoles dominent le milieu rural du District d'Abidjan.

La carte produite par cette étude s'avère être un outil très essentiel et utile à la gestion des catastrophes naturelles. En effet elle constitue une pièce d'entre importance à la modélisation de ces phénomènes

REMERCIEMENTS

Nos remerciements aux institutions qui nous ont aidé dans l'accomplissement du travail présenté:

- Le Centre Universitaire de Recherche et d'Application en Télédétection (CURAT/UFHB);
- Le Programme GMES de l'union Africain sur la thématique des inondations (MIFMASS),
- Et le Bureau National d'Etude Technique et de Développement (BNETD/CIGN)

REFERENCES

- [1] G. Adon, J.-B. Kassi, et C. Yoboue, « Dynamics of Land Use in the City of Abidjan from 1986 to 2017: Contribution of Remote Sensing and GIS », *Journal of Environmental & Analytical Toxicology*, vol. 08, janv. 2018, doi: 10.4172/2161-0525.1000573.
- [2] D. A. ALLA, « Cartographie des zones à risques d'inondation, d'érosion et de mouvement de terrain dans la ville d'Abidjan, PNUD, MESUDD. », 2013.
- [3] D. A. ALLA, « Risques naturels dans l'agglomération d'Abidjan (Côte d'Ivoire) », Thèse de Doctorat d'État ès - Sciences Humaines, Université Félix Houphouët-Boigny, UFR des Sciences de l'Homme et de la Société, Institut de Géographie Tropicale, Abidjan, Côte d'Ivoire, 2013.
- [4] N. Aghui et J. Biemi, Géologie et hydrogéologie des nappes de la région d'Abidjan et risques de contamination. Ann. in série C (sciences), 20: Univ Nat. de Côte d'Ivoire, 1984.
- [5] C. Hauhouot, « Analyse du risque pluvial dans les quartiers précaires d'Abidjan. Etude de cas à Attécoubé », *Geo-Eco-Trop*, n° 32, p. 75-82, 2008.
- [6] S. DIALLO *et al.*, « Dynamique de l'occupation du sol du bassin ivoirien de la lagune Aghien », *2019 ISSR Journals*, vol. 26, n° 1, p. 203-217, avr. 2019.
- [7] Y. K. A. Bio *et al.*, « Étude de la dynamique forestière du sud-ouest de la Côte d'Ivoire par télédétection optique (moyenne et haute résolution spatiale) : contribution des images MODIS ET LANDSAT. 2ème Conférence Internationale GEOFORAFRI, Abidjan, 26 – 28 Janvier 2016, Communication orale. », 2016.
- [8] G. M. Foody, « Status of land cover classification accuracy assessment », *Remote Sensing of Environment*, vol. 80, n° 1, p. 185-201, avr. 2002, doi: 10.1016/S0034-4257(01)00295-4.
- [9] L. S. Macarringue, É. L. Bolfe, et P. R. M. Pereira, « Developments in Land Use and Land Cover Classification Techniques in Remote Sensing: A Review », *JGIS*, vol. 14, n° 01, p. 1-28, 2022, doi: 10.4236/jgis.2022.141001.
- [10] D. Lu, « The potential and challenge of remote sensing-based biomass estimation », *International Journal of Remote Sensing*, vol. 27, n° 7, p. 1297-1328, avr. 2006, doi: 10.1080/01431160500486732.
- [11] D. Lu, P. Mausel, M. Batistella, et E. Moran, « Land-cover binary change detection methods for use in the moist tropical region of the Amazon: a comparative study », *International Journal of Remote Sensing*, vol. 26, n° 1, p. 101-114, janv. 2005, doi: 10.1080/01431160410001720748.
- [12] Z. Lv, T. Liu, P. Zhang, J. Atli Benediktsson, et Y. Chen, « Land Cover Change Detection Based on Adaptive Contextual Information Using Bi-Temporal Remote Sensing Images », *Remote Sensing*, vol. 10, n° 6, p. 901, juin 2018, doi: 10.3390/rs10060901.
- [13] B. Liu, W. Song, Z. Meng, et X. Liu, « Review of Land Use Change Detection—A Method Combining Machine Learning and Bibliometric Analysis », *Land*, vol. 12, n° 5, p. 1050, mai 2023, doi: 10.3390/land12051050.
- [14] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, et R. Moore, « Google Earth Engine: Planetary-scale geospatial analysis for everyone », *Remote Sensing of Environment*, vol. 202, p. 18-27, déc. 2017, doi: 10.1016/j.rse.2017.06.031.
- [15] A. Kombate, F. Folega, W. Atakpama, M. Dourma, K. Wala, et K. Goïta, « Characterization of Land-Cover Changes and Forest-Cover Dynamics in Togo between 1985 and 2020 from Landsat Images Using Google Earth Engine », *Land*, vol. 11, n° 11, Art. n° 11, nov. 2022, doi: 10.3390/land11111889.
- [16] N. Sidhu, E. Pebesma, et G. Câmara, « Using Google Earth Engine to detect land cover change: Singapore as a use case », *European Journal of Remote Sensing*, vol. 51, n° 1, p. 486-500, janv. 2018, doi: 10.1080/22797254.2018.1451782.

- [17] B. DeVries, C. Huang, J. Armston, W. Huang, J. Jones, et M. Lang, « Rapid and robust monitoring of flood events using Sentinel-1 and Landsat data on the Google Earth Engine », *Remote Sensing of Environment*, vol. 240, janv. 2020, doi: 10.1016/j.rse.2020.111664.
- [18] M. Eldin, « Le climat de la Côte d'Ivoire. In Le milieu naturel de Côte d'Ivoire, Mémoire ORSTOM (pp. 73-108), Paris ORSTOM. - References - Scientific Research Publishing », 1971. [https://www.scirp.org/S\(351jmbntvnsjt1aadkposzje\)/reference/ReferencesPapers.aspx?ReferenceID=1809522](https://www.scirp.org/S(351jmbntvnsjt1aadkposzje)/reference/ReferencesPapers.aspx?ReferenceID=1809522) (consulté le 25 août 2021).
- [19] MCLAU, « LE PROJET DE DEVELOPPEMENT DU SCHEMA DIRECTEUR D'URBANISME DU GRAND ABIDJAN (SDUGA) », Abidjan, Côte d'Ivoire, SCHEMA DIRECTEUR D'URBANISME DU GRAND ABIDJAN ET AUTRES TRAVAUX DU PROJET VOLUME II, mars 2015.
- [20] L. Ma, L. Cheng, M. Li, Y. Liu, et X. Ma, « Training set size, scale, and features in Geographic Object-Based Image Analysis of very high resolution unmanned aerial vehicle imagery », *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 102, p. 14-27, avr. 2015, doi: 10.1016/j.isprsjprs.2014.12.026.
- [21] L. Jansen et A. Di Gregorio, *Land Cover Classification System (LCCS): Classification Concepts and User Manual*. 2000.
- [22] R. G. Congalton et K. Green, *Assessing the accuracy of remotely sensed data: principles and practices*, Third edition. Boca Raton London New York: CRC Press, 2019.
- [23] V. Thiérion, P. A. Herrault, A. Vincent, J. Inglada, et D. Sheeren, « Utilisation des séries temporelles d'images Sentinel-2 pour la cartographie de l'occupation du sol dans un contexte de modélisation de la biodiversité », présenté à Colloque PAYOTTE 2017, oct. 2017, p. 31 p. Consulté le: 25 février 2023. [En ligne]. Disponible sur: <https://hal.inrae.fr/hal-02738272>.
- [24] K. D. KOUASSI, B. H. Kouadio, D. A. ALLA, et T. V. ASSOMA, « Cartographie de l'occupation du sol par une nouvelle approche de classification orientée objet à partir de l'image sentinel-2a de la commune d'Abobo (Abidjan, Côte d'Ivoire). », *International Journal of Engineering Science Invention (IJESI)*, vol. 7, n° 1, p. 28-38, 2018.
- [25] O. G. K. KOUASSI, F. K. N'GUESSAN, et G. W. KOUKOUNGON, « Analyse cartographique et statistique de l'état de l'occupation du sol dans la sous-préfecture de Bondoukou (nord-est, Côte d'Ivoire) », *Regard sud*, 4 mai 2022. <https://regardsuds.org/analyse-cartographique-et-statistique-de-letat-de-loccupation-du-sol-dans-la-sous-prefecture-de-bondoukou-nord-est-cote-divoire/> (consulté le 27 décembre 2022).
- [26] D. Lessire, « Cartographie de types de forêt et de l'occupation du sol en République Centrafricaine à l'aide de télédétection optique (Sentinel-2) », 2019.
- [27] F. N. Kabemba *et al.*, « Couvert forestier et distribution de *Pan paniscus* dans la Réserve Naturelle de Sankuru, RD Congo », présenté à Conférence OSFACO : Des images satellites pour la gestion durable des territoires en Afrique, mars 2019. Consulté le: 25 février 2023. [En ligne]. Disponible sur: <https://hal.science/hal-02189507>.
- [28] V. J. Sokeng *et al.*, « Suivi par télédétection des affectations des terres pour la promotion d'une agriculture intégrée au développement forestier en Côte d'Ivoire », présenté à Conférence OSFACO : Des images satellites pour la gestion durable des territoires en Afrique, mars 2019. Consulté le: 25 février 2023. [En ligne]. Disponible sur: <https://hal.science/hal-02189403>.
- [29] F. Bioresita, A. Puissant, A. Stumpf, et J.-P. Malet, « Fusion of Sentinel-1 and Sentinel-2 image time series for permanent and temporary surface water mapping », *International Journal of Remote Sensing*, juin 2019, Consulté le: 2 octobre 2021. [En ligne]. Disponible sur: <https://www.tandfonline.com/doi/abs/10.1080/01431161.2019.1624869>.
- [30] K. A. Stéphane, A. Roland, et E. M. Jalal, « Utilisation d'une image satellite LANDSAT 8 pour la cartographie de l'occupation des sols dans la ville de Bondoukou et ses environs en Côte d'Ivoire », vol. 30, n° 1, p. 10, 2020.
- [31] A. M. Abdi, « Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data », *GIScience & Remote Sensing*, vol. 57, n° 1, p. 1-20, janv. 2020 doi: 10.1080/15481603.2019.1650447.
- [32] C. Qiu, L. Mou, M. Schmitt, et X. X. Zhu, « Fusing Multiseasonal Sentinel-2 Imagery for Urban Land Cover Classification With Multibranch Residual Convolutional Neural Networks », *IEEE Geosci. Remote Sensing Lett.*, vol. 17, n° 10, p. 1787-1791, oct. 2020, doi: 10.1109/LGRS.2019.2953497.
- [33] Y. Konko, B. Afelu, et K. Kokou, « Potentialité des données satellitaires Sentinel-2 pour la cartographie de l'impact des feux de végétation en Afrique tropicale : application au Togo », *Bois et Forêts des Tropiques*, vol. 347, p. 59, mars 2021, doi: 10.19182/bft2021.347.a36349.
- [34] T. Ouattara, F. Kouamé, C. Zo-Bi, R. Vaudry, et C. Grinand, « Changements d'occupation et d'usage des terres entre 2016 et 2019 dans le Sud-Est de la Côte d'Ivoire : impact des cultures de rente sur la forêt », *BOIS & FORETS DES TROPIQUES*, vol. 347, p. 91-106, mars 2021, doi: 10.19182/bft2021.347.a31868.
- [35] B. Mertens et C. Pinet, « Accompagnement de pays d'Afrique Centrale et de l'Ouest dans l'utilisation de l'imagerie satellitaire pour la conception, la mise en oeuvre et le suivi des politiques publiques d'aménagement durable du territoire, de 2010 à 2020 », *Revue Française de Photogrammétrie et de Télédétection*, vol. 223, p. 81-87, nov. 2021, doi: 10.52638/rfpt.2021.565.

- [36] H. D. N'Da, E. K. N'Guessan, M. E. Wajda, et K. Affian, « Apport de la télédétection au suivi de la déforestation dans le Parc National de la Marahoué (Côte d'Ivoire) », *Teledetection*, vol. 8, n° 1, p. 17, 2008.
- [37] T. D. Soro *et al.*, « Hydroclimatologie et dynamique de l'occupation du sol du bassin versant du Haut Bandama à Tortiya (Nord de la Côte d'Ivoire) », *Vertigo - la revue électronique en sciences de l'environnement*, n° Volume 13 Numéro 3, Art. n° Volume 13 Numéro 3, déc. 2013, doi: 10.4000/vertigo.14468.
- [38] G. Soro *et al.*, « Apport de la télédétection à la cartographie de l'évolution spatio-temporelle de la dynamique de l'occupation du sol dans la région des Lacs (Centre de la Côte d'Ivoire) », *Afrique Science: Revue Internationale des Sciences et Technologie*, vol. 10, n° 3, Art. n° 3, nov. 2014, doi: 10.4314/afsci.v10i3.
- [39] G. Ake, K. Kouame, A. Koffi, et J. Jourda, « Cartographie des zones potentielles de recharge de la nappe de Bonoua (sud-est de la Côte d'Ivoire) », *rseau*, vol. 31, n° 2, p. 129-144, 2018, doi: 10.7202/1051696ar.
- [40] G. E. Ake et B. H. Kouadio, « Apport des SIG à la délimitation des périmètres de protection autour de la prise d'eau de la SODECLI à Aboisso, sud - Est de la Côte d'Ivoire », 2019, n° 15, Art. n° 15, 2019.
- [41] A. E. N'Guessan, Y. L. Akpa, N. O. Yao, et J. N. Kassi, « Cartographie de la dynamique du couvert végétal de la forêt Classée d'Agbo 1 Côte d'Ivoire », *Agronomie Africaine*, vol. 31, n° 1, Art. n° 1, mai 2019, doi: 10.4314/aga.v31i1.
- [42] M. Y. Ta *et al.*, « Diagnostic of hydrous state of drain in the mid-dry season of kohodio watershed using multispectral images of Landsat generation from 1986 to 2018 », *International Journal of Innovation and Applied Studies*, vol. 27, n° 1, Art. n° 1, août 2019.
- [43] V. T. Assoma, N. A. Yao, J. S. Dio, et J.-P. Jourda, « Apport de la télédétection et d'un SIG à la cartographie des changements de l'occupation du sol dans le bassin versant de la Lobo en Côte d'Ivoire », *Physio-Géo. Géographie physique et environnement*, n° Volume 16, Art. n° Volume 16, janv. 2021, doi: 10.4000/physio-geo.12654.
- [44] M. A. C. NJEUGEUT, T. M. YOUAN, V. M. SOROKOBY, T. V. ASSOMA, M. G. ADJA, et J. P. JOURDA, « Dynamique d'occupation du sol du bassin versant de la volta, par la méthode de l'arbre de décision, à partir des images multispectrales de la génération Landsat de 1990 à 2020 », *International Journal of Engineering Science Invention (IJESI)*, vol. 10, n° 4, p. 34-45, avr. 2021.
- [45] D. Akkari, « L'apport du système d'information géographique (SIG) dans la définition des zones de potentiel hydrique dans le bassin versant Abou Ali (Liban Nord) », *Journal of Alpine Research | Revue de géographie alpine*, n° 110-4, Art. n° 110-4, janv. 2022, doi: 10.4000/rga.10015.
- [46] S. M. Bachir *et al.*, « Télédétection et système d'information géographique pour la cartographie du parcellaire situé en bordure de rivière de la région de Man (Ouest de la Côte d'Ivoire). », p. 2.
- [47] I. Aneece et P. Thenkabail, « Accuracies Achieved in Classifying Five Leading World Crop Types and their Growth Stages Using Optimal Earth Observing-1 Hyperion Hyperspectral Narrowbands on Google Earth Engine », *Remote Sensing*, vol. 10, n° 12, Art. n° 12, déc. 2018, doi: 10.3390/rs10122027.
- [48] K. Aka, ¹ R., H. N, D. Crystel, et N. Bohoussou, « Étude comparative de Sentinel-2 et Landsat-8 Oli à l'évaluation de l'occupation du sol du site Ramsar d'Azagny, Sud de la Côte d'Ivoire », *Afrique Science Revue Internationale des Sciences et Technologie*, vol. VOL.20, N*5 (2022), p. 48-63, avr. 2022.
- [49] H. N'da, C. Kaudjhis, et K. S. Dahan, « Dynamique spatio-temporelle des feux de 2001 à 2019 et dégradation du couvert végétal en zone de contact forêt-savane, Département de Toumodi, Centre de la Côte d'Ivoire », *Afrique Science Revue Internationale des Sciences et Technologie*, vol. 19, p. 94-113, sept. 2021.
- [50] C. B. Tiesse, E. N. Wandan, et H. D. N'da, « Apport De La Teledetection Pour Le Suivi SpatioTemporel De L'occupation Du Sol Dans La Region Montagneuse Du Tonkpi (Cote D'Ivoire) », *ESJ*, vol. 13, n° 15, p. 310, mai 2017 doi: 10.19044/esj.2017.v13n15p310.
- [51] T. Ouattara *et al.*, « Suivi des Terres et de la Déforestation par Télédétection Spatiale et Aérienne à l'Est et au Sud-Est de la Côte d'Ivoire », 2022. doi: 10.13140/RG.2.2.33587.71200.