Afrique SCIENCE *29(5) (2026) 62 - 81*
ISSN 1813-548X, http://www.afriquescience.net

Performances du Stacking face aux modèles SVM, RF et XGBoost. Application au processus d'occupation du sol dans la commune d'Oyem et sa périphérie, Gabon

Luckin OBIANG OBIANG, Marius MASSALA MBINDZOUKOU* et Bruno NKOUMAKALI

Université Omar Bongo, Faculté des Lettres et Sciences Humaines, Département des Sciences Géographiques, Environnementales et Marines, Laboratoire de Géomatique, de Recherche Appliquée et de Conseil, BP 17043, Libreville, Gabon

(Reçu le 19 Août 2026 ; Accepté le 15 Octobre 2026)

* Correspondance, courriel : *massalamarius@gmail.com*

Résumé

La cartographie de l'occupation du sol constitue un outil fondamental pour la gestion territoriale. La présente étude propose une analyse comparative des algorithmes Machine à Vecteurs de Support (SVM), Forêt Aléatoire (FA), XGBoost (XGB) et de leur combinaison dans une approche d'Ensemble Learning (Stacking) pour classifier l'occupation du sol de la commune d'Oyem et de sa périphérie (Gabon) à partir d'une image Sentinel-2 de mars 2026. Trois configurations de variables sont évaluées : les bandes spectrales seules, les indices spectraux seuls (NDVI, EVI, SAVI, NDBI, MNDWI, BSI) et la combinaison de ces descripteurs. Les résultats montrent que le Stacking obtient le meilleur Kappa de l'étude (0,9664) avec les bandes spectrales seules, pour un gain toutefois modeste sur les modèles isolés (de 0,0013 à 0,0059). L'analyse indique que les bandes spectrales brutes constituent le descripteur le plus performant pour les quatre modèles. À l'inverse, les indices spectraux seuls entraînent une baisse systématique du Kappa (de 1,11 à 1,78 %), marquée par de fortes confusions sur la classe Eau. Enfin, la quantification issue du modèle optimal révèle un paysage largement dominé par les formations végétales, qui couvrent 94,63 % de la superficie, dont 14 270,69 ha. L'empreinte anthropique reste localisée le long des axes routiers, avec 729,95 ha (3,66 %) de zone urbanisée et 303,43 ha (1,52 %) de sol nu. Ces travaux fournissent une base méthodologique et un état de référence de l'occupation du sol d'Oyem et de sa périphérie, dont la transférabilité à d'autres zones du Gabon reste à tester.

Mots-clés : *occupation du sol, Sentinel-2, Géo-IA, ensemble Learning, Stacking, Oyem, Gabon.*

Abstract

Performance of Stacking compared with SVM, RF and XGBoost models. Application to the land-use process in the municipality of Oyem and its surrounding area, Gabon

Land cover mapping is a fundamental tool for territorial management. This study presents a comparative analysis of Support Vector Machine (SVM), Random Forest (RF) and XGBoost (XGB) algorithms, and of their combination through an Ensemble Learning (Stacking) approach, for land cover classification in the municipality of Oyem and its surrounding area (Gabon) using a Sentinel-2 image acquired in March 2026. Three variable configurations were evaluated: spectral bands alone, spectral indices alone (NDVI, EVI, SAVI, NDBI,

MNDWI, BSI), and the combination of all descriptors. The results show that Stacking obtains the best Kappa index of the study (0.9664) with spectral bands alone, although its gain over the individual models is modest (0.0013 to 0.0059). The analysis indicates that raw spectral bands are the best-performing descriptor for all four models. Conversely, spectral indices alone cause a systematic drop in Kappa (1.11 to 1.78%), characterized by severe confusion in the water class. Finally, the quantification derived from the optimal model reveals a landscape largely dominated by vegetation formations, which cover 94.63% of the total area, including 14,270.69 ha. The anthropogenic footprint remains localized along road networks, with 729.95 ha (3.66%) of urbanized area and 303.43 ha (1.52%) of bare soil. This research provides a methodological framework and a baseline land cover map of Oyem and its surrounding area, whose transferability to other areas of Gabon remains to be tested.

Keywords : *Land cover, Sentinel-2, Geo-AI, Ensemble Learning, Stacking, Oyem, Gabon.*

1. Introduction

L'occupation du sol constitue un élément fondamental dans l'analyse des territoires urbains et environnementaux [1]. À l'échelle mondiale, elle subit des mutations biophysiques profondes induites par la croissance démographique, l'expansion urbaine et l'intensification des activités humaines, caractérisées par une conversion accélérée des couvertures naturelles en surfaces anthropisées [2 - 4]. Cette dynamique est particulièrement prégnante en Afrique centrale où la fragmentation des strates forestières redéfinit les mosaïques paysagères et fragilise les équilibres écosystémiques [5]. Dans ce contexte régional, le Gabon se distingue par un couvert forestier qui s'étend sur environ 88 % de son territoire, ce qui en fait un pilier de la régulation climatique du Bassin du Congo [6]. Toutefois, l'émergence de nouveaux projets d'aménagement et la modernisation des infrastructures, impulsées par les réorientations politiques post-2023, induisent des transformations structurelles qui modifie l'organisation spatiale de plusieurs pôles urbains du pays. Cette initiative de modernisation urbaine nécessite ainsi une connaissance exhaustive et précise de l'occupation du sol pour une planification efficace. La télédétection et les systèmes d'information géographique (SIG) constituent des pistes de solutions pour répondre à ce besoin de précision. Les données issues des programmes Landsat de la NASA et Sentinel du programme Copernicus offrent des images multi-spectrales de haute qualité qui permettent d'identifier le bâti, les formations végétales, les plans d'eau ou les sols nus [7]. Cependant, la précision des cartes obtenues dépend fortement des méthodes de classification, des variables d'entrées et d'algorithmes utilisés pour la détection des classes d'occupation du sol [8,9]. Pour tenter de surmonter ces contraintes, les algorithmes d'apprentissage automatique tels que les Machines à Vecteurs de Support (SVM), les Forêts Aléatoires (FA) et XGBoost (XGB) sont de plus en plus utilisés en raison de leurs performances élevées face aux données complexes. Le SVM est privilégié pour son aptitude à définir des hyperplans de décision optimaux, même avec des échantillons limités [10]. L'algorithme XGB se distingue par sa vitesse de calcul et sa capacité à minimiser l'erreur résiduelle par l'optimisation du gradient [11]. Enfin, la FA reste reconnue pour sa stabilité face au bruit et sa résilience au sur apprentissage grâce à l'agrégation de multiples arbres de décision [10]. Parallèlement, les approches d'apprentissage d'ensemble (Ensemble Learning ou Stacking),fondées sur la combinaison de plusieurs modèles suscitent un intérêt croissant pour agréger les prédictions, compenser les biais individuels et maximiser la précision globale [12, 13]. Leur supériorité n'est toutefois pas systématique, car le gain du Stacking par rapport aux modèles de base dépend du contexte et des données utilisées [14]. Au-delà des algorithmes, les performances dépendent fortement des variables utilisées pour l'apprentissage [15]. Les bandes spectrales fournissent des informations radiométriques essentielles, tandis que les indices spectraux permettent de renforcer la discrimination thématique des classes. Les indices tels que l'Indice de Végétation par Différence Normalisée (NDVI), l'Indice

de Végétation Amélioré (EVI), l'Indice de Végétation Ajusté au Sol (SAVI), l'Indice de Bâti par Différence Normalisée (NDBI), l'Indice d'Eau par Différence Normalisée Modifié (MNDWI) et l'Indice de Sol Nu (BSI) sont largement utilisés afin d'améliorer l'identification de la végétation, des surfaces bâties, des plans d'eau et des sols nus [16]. Au Gabon, l'application de ces approches avancées d'apprentissage automatique (Machine Learning) et d'Ensemble Learning demeure embryonnaire malgré des travaux pionniers [17 - 19]. Par ailleurs, l'évaluation de ces modèles par certains auteurs repose souvent sur des validations croisées aléatoires qui ignorent l'autocorrélation spatiale des échantillons [17]. Des pixels voisins, très semblables se retrouvent alors à la fois dans les jeux d'apprentissage et de test, ce qui conduit à des précisions surestimées [20]. Or, la fiabilité des cartes d'occupation du sol est déterminante pour le suivi spatial et la gestion territoriale [21]. Au nord du Gabon, le choix des algorithmes et des variables les mieux adaptés à la cartographie de l'occupation du sol reste peu documenté. Il importe donc de l'établir à l'aide d'une démarche rigoureuse qui neutralise l'effet de l'autocorrélation spatiale. Dans cette perspective, l'objectif de cette étude est de comparer à partir d'une image Sentinel-2 de 2026, la capacité de généralisation des modèles SVM, FA et XGB pris isolément à celle d'un modèle d'ensemble de type Stacking, afin de cartographier l'occupation du sol de la commune d'Oyem et de sa périphérie. De manière spécifique, il s'agit (i) de déterminer si l'approche ensembliste surpasse les modèles isolés, (ii) d'évaluer l'impact de trois jeux de variables (bandes seules, indices seuls et combinaison bandes-indices) sur la précision de la classification afin d'identifier la configuration optimale et (iii) de quantifier les superficies de chaque classe d'occupation du sol.

2. Matériel et méthodes

2-1. Présentation de la zone d'étude

La zone d'étude (199,4274 km²) se situe au nord du Gabon (département du Woleu), entre 11°30'41'' E - 11°38'53'' E et 1°32'24'' N - 1°37'41'' N ***(Figure 1)***. Structuré par un axe routier majeur qui relie Libreville, Bitam et la Guinée équatoriale. Ce territoire regroupe plusieurs villages (Amvam, Alloum, Nkolayop, etc.). La topographie varie entre 610 et 837 mètres d'altitude avec une alternance de bas-fonds et d'interfluves ***(Figure 1)***.

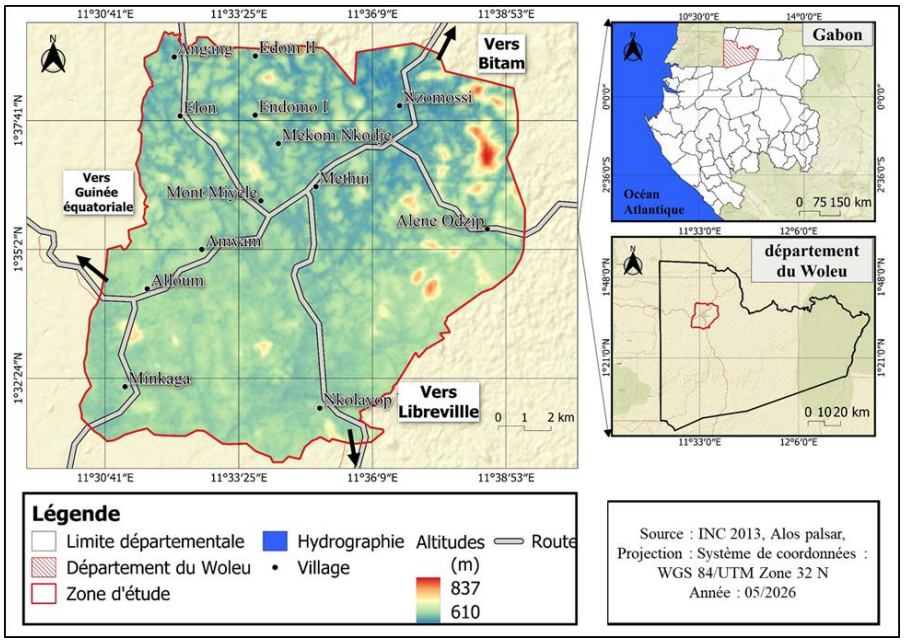


Figure 1 : *Localisation de la zone d'étude*

2-2. Démarche méthodologique générale

Le protocole technique de cette étude s'articule autour de plusieurs étapes essentielles ***(Figure 2)***. Il commence par l'acquisition des données Sentinel-2 de niveau 2A, leur prétraitement et leur traitement sur la plateforme Google Earth Engine (GEE). Ensuite, la classification et la validation statistique des résultats (matrice de confusion et indice Kappa) sont réalisées dans l'environnement Jupyter Notebook, avec le langage Python. L'étude se termine par la production d'une carte thématique sous QGIS.

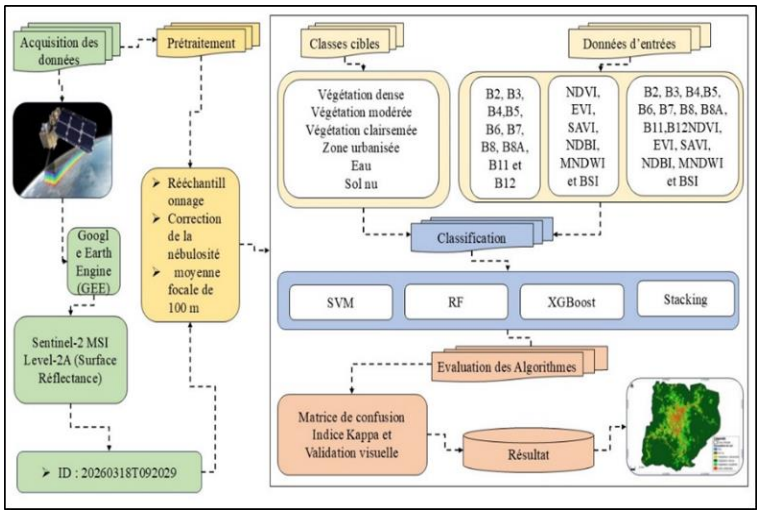


Figure 2 : *Résumé méthodologique de l'étude*

Cette étude mobilise plusieurs outils, dont chacun a une fonction précise dans l'acquisition, le traitement, la classification et la validation des données ***(Tableau 1)***.

Tableau 1 : *Outils et utilités dans l'étude*

Outil	Utilité
Google Earth Engine (GEE)	Plateforme cloud d'acquisition et de prétraitement des images satellites
Sentinel-2 (niveau 2A)	Source d'imagerie satellite (ID : 20260318T092029)
Python / Jupyter Notebook	Environnement de développement pour le Machine Learning
DBSCAN	Clustering spatial des données de vérité terrain
GroupKFold	Validation croisée des modèles
Random Forest,XGBoost,SVM	Modèles de classification (modèles de base du Stacking)
Régression logistique	Méta-modèle combinant les résultats du Stacking

2-3. Acquisition et prétraitement des données

L'acquisition, les prétraitements et les traitements des données Sentinel-2 sont effectués sur la plateforme cloud Google Earth Engine. L'étude exploite une image Sentinel-2 Level-2A (Surface Réflectance) du 18 mars 2026 (ID : 20260318T092029). Ce capteur (10 mètres) est choisi pour sa gratuité et sa capacité à discriminer le bâti fragmenté contrairement aux 30 mètres de Landsat. Le produit, est corrigé par l'algorithme Sen2Cor [22]. Il comprend 10 bandes spectrales ***(Tableau 2)***. Les bandes à 20 mètres dans cette étude sont rééchantillonnées à 10 mètres par interpolation bilinéaire afin d'harmoniser la géométrie de l'image [23].

Tableau 2 : *Caractéristiques des bandes spectrales Sentinel-2 retenues*

Bande	Nom	λ centrale (nm)	Résolution (m)	Utilité thématique
B2 / B3 / B4	Bleu / Vert / Rouge	490 / 560 / 665	10	Eau,sols nus,vigueur de la végétation
B5 / B6 / B7 / B8A	Bords rouges (1 à 4)	705 / 740 / 783 / 865	20	Suivi de la biomasse et santé des cultures
B8	Infrarouge proche	842	10	Croissance des feuilles / végétation dense
B11 / B12	SWIR 1 et 2	1 610 / 2 190	20	Humidité et stress hydrique du sol / couvert

Les nuages et ombres sont masqués manuellement (NaN), pour éviter le flou des filtres spatiaux. Les vides créés sont comblés par une interpolation géostatistique du plus proche voisin valide [24],via un réducteur de voisinage (ee.Reducer.first) sur un rayon de 550 m ***(Figure 3)***.

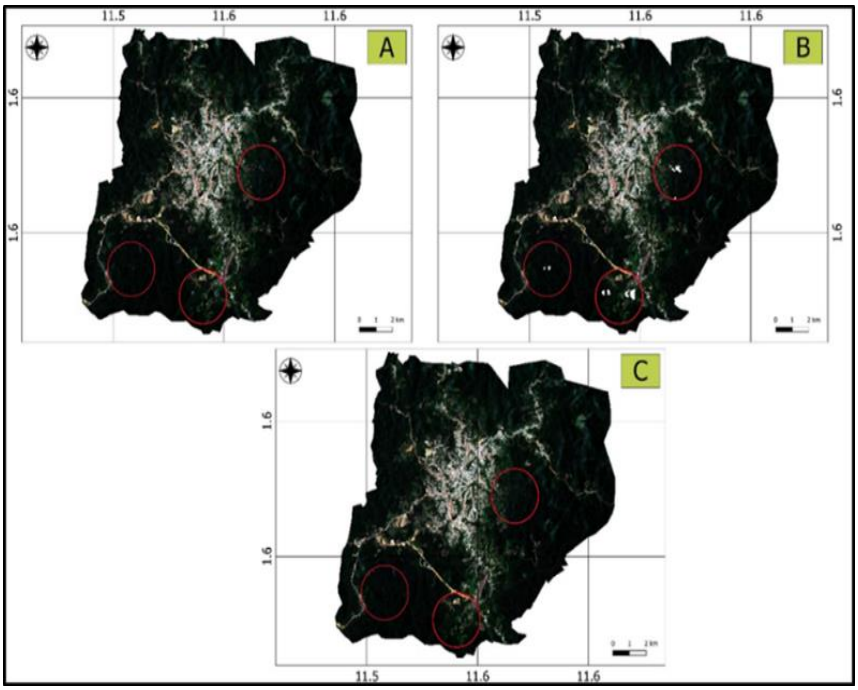


Figure 3 : *Compositions colorées (RGB) de la zone d'étude illustrant le processus de filtrage et de correction des données spectrales (A : Image initiale, B : Masquage des zones cibles, C : Image corrigée*

C'est immédiatement après cette étape d'interpolation que nous réalisons l'empilement complet des bandes (band stacking). Les bandes spectrales à 20 mètres de résolution sont rééchantillonnées à 10 mètres par interpolation bilinéaire, puis fusionnées avec les bandes natives à 10 mètres pour reconstituer une matrice corrigée et homogène de 10 bandes nettes. Il est ensuite procédé au calcul de six indices spectraux NDVI, EVI, SAVI, NDBI, MNDWI et BSI ***(Tableau 3)***.

Tableau 3 : *Équations des indices thématiques*

Nom	Equations	References
NDVI	$\frac{B_8 - B_4}{B_8 + B_4}$	[25]
EVI	$2.5 \cdot \frac{B_8 - B_4}{B_8 + 6 \cdot B_4 - 7.5 \cdot B_2 + 1}$	[26]
SAVI	$\frac{(B_8 - B_4) \cdot (1+L)}{B_8 + B_4 + L}$ (avec L = 0.5)	[27]
NDBI	$\frac{B_{11} - B_8}{B_{11} + B_8}$	[28]
MNDWI	$\frac{B_3 - B_{11}}{B_3 + B_{11}}$	[29]
BSI	$\frac{(B_{11} + B_4) - (B_8 + B_2)}{(B_{11} + B_4) + (B_8 + B_2)}$	[30]

Ensuite, deux blocs de données sont structurés. Le FI (6 indices spectraux seuls) et FIBA (10 bandes interpolées + 6 indices, soit 16 bandes). Les compositions colorées sous QGIS (B8-B4-B3 pour la végétation ; B11-B8-B4 pour l'eau ; B4-B3-B2 pour le bâti) guident la sélection visuelle des régions d'intérêt ***(Figure 4)***. La classification est de type supervisé pixel à pixel et repose sur six classes d'occupation du sol : végétation dense, végétation modérée, végétation clairsemée, eau, zone urbanisée et sol nu.

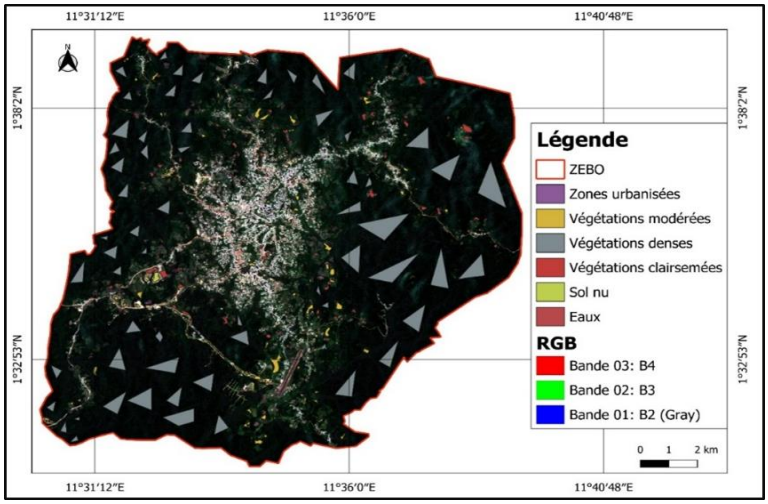


Figure 4 : *Distribution spatiale des échantillons*

2-4. Classification et validation

Pour neutraliser le biais d'autocorrélation propre aux données géospatiales transversales, une validation croisée spatiale par blocs (Group-KFold à 5 plis) est adoptée, car elle prévient la surestimation des prédictions [31, 32]. L'algorithme de partitionnement spatial DBSCAN (esp = 30, min_samples = 1) est d'abord appliqué sur les coordonnées géographiques (Map X, Map Y) pour regrouper les pixels contigus en entités géographiques distinctes (parcelles). Le jeu de données est ensuite scindé de telle sorte que l'intégralité des pixels d'une même parcelle soit affectée soit exclusivement à l'apprentissage, soit exclusivement au test, ce qui interdit toute fuite d'information spectrale. Les quatre configurations algorithmiques exécutées sous Python sont les suivantes :

- **Forêt Aléatoire (FA) :** paramétrée avec 500 arbres de décision.
- **Machine à Vecteurs de Support (SVM) :** configurée avec un noyau à fonction de base radiale (RBF, C = 10, gamma = 'scale') couplé à une normalisation standardisée impérative par pli (Standard Scaler).
- **XGBoost (XGB) :** ajusté avec un taux d'apprentissage de 0,1, une profondeur maximale de 6 niveaux et 500 estimateurs.
- **Stacking (Ensemble Learning) :** déploiement d'un Stacking Classifier qui fusionne les trois modèles de base précédents (en conservant exactement les mêmes hyper paramètres) via un méta-modèle de Régression Logistique contrôlé par une validation croisée interne à 5 plis.

L'évaluation s'appuie quantitativement sur le calcul de la Précision Globale (OA) et de l'indice Kappa ***(Tableau 4)***, interprété selon l'échelle de Landis et Koch [33]. Elle se complète par une validation cartographique visuelle par zooms sous QGIS.

Tableau 4 : *Échelle de concordance du coefficient Kappa de Landis et Koch*

Valeur du coefficient Kappa (κ)	**Qualité de l'accord (Fiabilité)**
< 0,20	Très faible-Mauvaise
0,21 - 0,40	Faible-Médiocre
0,41 - 0,60	Modérée
0,61 - 0,80	Bonne-Forte
0,81 - 1,00	Excellente-Accord presque parfait

3. Résultats

3-1. Évaluation comparative des modèles et pouvoir prédictif des variables

Les performances globales des différents algorithmes de Machine Learning s'avèrent globalement excellentes selon les configurations de variables testées ***(Tableau 5)***. Avec les bandes spectrales seules, tous les modèles sont excellents avec des précisions globales (OA) supérieure à 98,3 % et Kappa supérieur à 0,960. Le Stacking domine ce scénario (OA : 98,59 % ; Kappa : 0,9664), suivi de près par XGBoost (OA : 98,54 % ; Kappa : 0,9651) et le SVM (OA : 98,50 % ; Kappa : 0,9647), tandis que la Forêt Aléatoire reste légèrement en retrait (OA : 98,35 % ; Kappa : 0,9605). L'écart entre le meilleur et le moins bon modèle dans cette configuration est faible soit : 0,24 point d'OA et de 0,0059 en Kappa. Avec les indices spectraux seuls, les performances baissent pour tous les modèles : l'OA se situe entre 97,74 % (FA) et 98,06 % (SVM), et le Kappa entre 0,9460 et 0,9540. Dans cette configuration, le SVM devance légèrement le Stacking (Kappa de 0,9540 contre 0,9522), suivi de XGBoost (0,9479) et de la FA (0,9460). Avec la combinaison des bandes et des indices, les performances des quatre modèles se rapprochent. Le SVM, XGBoost et le Stacking obtiennent des OA quasi identiques (98,49 %, 98,50 % et 98,48 %) et des Kappa compris entre 0,9641 et 0,9644. La FA reste en retrait (OA : 98,33 % ; Kappa : 0,9600).

Tableau 5 : *Synthèse comparative des performances globales des modèles*

Configuration des variables	Métrique	SVM	FA	XGBoost	Ensemble Learning
Bandes spectrales seules	OA (%)	98,50	98,35	98,54	98,59
	kappa	0,9647	0,9605	0,9651	0,9664
Indices spectraux seuls	OA (%)	98,06	97,74	97,82	97,99
	kappa	0,9540	0,9460	0,9479	0,9522
Bandes + Indices	OA (%)	98,49	98,33	98,50	98,48
	kappa	0,9644	0,9600	0,9642	0,9641

3-2. Analyse des confusions thématiques

L'analyse des matrices de confusion met en évidence la supériorité globale du Stacking face aux modèles isolés dans la configuration des bandes spectrales seules ***(Tableau 6)***. Pour la classe eau, XGBoost (863 pixels) et le Stacking (854) sont les plus stables, tandis que le SVM subit de fortes commissions (131 pixels) confondu avec la végétation clairsemée. Pour le sol nu, le SVM et le Stacking dominent en nombres de pixels pour le sol nu (5425 pixels) avec des confusions non négligeables pour les deux entre les classes sol nu et la végétation clairsemée (133 pour le SVM et 126 pour le Stacking), alors que la FA commet le plus d'omissions vers l'urbain (163 pixels) en raison de la ressemblance spectrale entre toitures et pistes latéritiques. Le pôle majeur de confusion oppose la végétation dense à la végétation modérée. Le Stacking minimise ce chevauchement (243 et 241 pixels) contre 316 et 347 pixels pour la FA. La végétation clairsemée enregistre une confusion secondaire vers la strate modérée (198 pixels pour le Stacking,209 pour XGBoost). Enfin, l'extraction des zones urbanisées est dominée par un grand nombres de pixels par XGBoost (4752) et le Stacking (4680), dont les omissions vers la végétation clairsemée (667-745 pixels) traduisent l'effet de pixel mixte inhérent au tissu bâti d'Oyem.

Tableau 6 : *Matrices de confusion issues de la classification par bandes spectrales*

SVM							
Classes Réelles \ Prédites	Eau	Sol nu	Végét. Clairsemée	Végét. Dense	Végét. modérée	Zone urb	Total
Eau	867	0	5	2	17	0	891
Sol nu	19	5425	53	0	1	93	5591
Végét. clairsemée	131	133	10614	4	208	140	11230
Végét. dense	2	0	0	109512	232	0	109746
Végét. modérée	7	0	145	179	12347	10	12688
Zone urb	1	23	775	0	0	4655	5454
Total	1027	5581	11592	109697	12805	4898	145600
FA							
Classes Réelles \ Prédites	Eau	Sol nu	Végét. clairsemée	Végét. dense	Végét. modérée	Zone urb	Total
Eau	852	0	4	2	32	1	891
Sol nu	31	5319	76	0	2	163	5591
Végét. clairsemée	0	165	10829	6	203	27	11230
Végét. dense	8	0	0	109422	316	0	109746
Végét. modérée	6	0	239	347	12096	0	12688
Zone urb	1	30	741	0	5	4677	5454
Total	898	5514	11889	109777	12654	4868	145600

XGBoost							
Classes Réelles \ Prédites	Eau	Sol nu	Végét. clairsemée	Végét. dense	Végét. modérée	Zone urb	Total
Eau	863	0	3	3	21	1	891
Sol nu	0	5413	63	0	10	105	5591
Végét. clairsemée	4	126	10801	3	209	87	11230
Végét. dense	5	0	0	109428	313	0	109746
Végét. modérée	4	0	226	247	12211	0	12688
Zone urb	1	31	667	0	3	4752	5454
Total	877	5570	11760	109681	12767	4945	145600
Stacking							
Classes Réelles \ Prédites	Eau	Sol nu	Végét. clairsemée	Végét. dense	Végét. modérée	Zone urb	Total
Eau	854	0	4	2	30	1	891
Sol nu	0	5425	70	0	1	95	5591
Végét. clairsemée	1	126	10826	3	198	76	11230
Végét. dense	4	0	0	109499	243	0	109746
Végét. modérée	6	0	182	241	12259	0	12688
Zone urb	1	26	745	0	2	4680	5454
Total	866	5577	11827	109745	12733	4852	145600

Source : Sentinel-2 Level-2A, Année : 05/2026

À l'inverse, l'analyse basée sur les indices spectraux seuls révèle une baisse globale des performances par rapport aux bandes brutes ***(Tableau 7)***. Pour la classe eau, tous les modèles enregistrent de fortes confusions vers les zones urbanisées, qui oscillent entre 332 pixels (SVM) et 356 pixels (XGBoost), en raison de la confusion spectrale entre l'eau et les ombres portées du bâti d'Oyem. Pour le sol nu, le SVM obtient le meilleur score diagonal en terme de pixels (5456 pixels), tandis que le Stacking affiche des omissions résiduelles vers l'urbain (86 pixels). La confusion majeure oppose la végétation dense à la végétation modérée. Le SVM et le Stacking minimisent cette distorsion (431-440 pixels), contre 499 pour XGBoost et 517 pour la FA. La végétation clairsemée enregistre des omissions diffuses vers la strate modérée (340 pixels, Stacking) et les zones urbanisées (129 pixels). Enfin, l'identification des zones urbanisées restent robustes sous XGBoost (4801 pixels) et le Stacking (4775 pixels), mais leurs omissions se dispersent massivement vers la végétation clairsemée (481 pour le XGBoost, 484 pour la FA,547 pixel pour le SVM et 563 pour le Stacking), traduisant l'effet de pixel mixte inhérent au tissu bâti végétalisé d'Oyem.

Tableau 7 : *Matrices de confusion des classifications basées sur les indices spectraux*

SVM							
Classes Réelles \ Prédites	Eau	Sol nu	Végét. clairsemée	Végét. dense	Végét. modérée	Zone urb	Total
Eau	503	0	5	3	48	332	891
Sol nu	0	5456	77	0	0	58	5591
Végét. clairsemée	12	51	10712	20	360	75	11230
Végét. dense	3	0	2	109310	431	0	109746
Végét. modérée	10	0	141	416	12121	0	12688
Zones urb	106	125	547	0	4	4672	5454
Total	634	5632	11484	109749	12964	5137	145600
FA							
Classes Réelles \ Prédites	Eau	Sol nu	Végét. clairsemée	Végét. dense	Végét. modérée	Zone urb	Total
Eau	494	0	8	3	44	342	891
Sol nu	0	5375	122	0	0	94	5591
Végét. clairsemée	0	131	10607	18	333	141	11230
Végét. dense	1	0	5	109223	517	0	109746
Végét. modérée	11	0	226	588	11860	3	12688
Zones urb	23	188	484	0	5	4754	5454
Total	529	5694	11452	109832	12759	5334	145600
XGBoost							
Classes Réelles \ Prédites	Eau	Sol nu	Végét. clairsemée	Végét. dense	Végét. modérée	Zone urb	Total
Eau	488	0	9	3	35	356	891
Sol nu	0	5418	96	0	0	77	5591
Végét. clairsemée	1	119	10581	16	339	174	11230
Végét. dense	3	0	9	109235	499	0	109746
Végét. modérée	11	0	258	520	11899	0	12688
Zones urb	36	130	481	0	6	4801	5454
Total	539	5667	11434	109774	12778	5408	145600
Stacking							
Classes Réelles \ Prédites	Eau	Sol nu	Végét. clairsemée	Végét. dense	Végét. modérée	Zone urb	Total
Eau	490	0	5	2	53	341	891
Sol nu	0	5425	80	0	0	86	5591
Végét. clairsemée	0	59	10686	16	340	129	11230
Végét. dense	2	0	1	109303	440	0	109746
Végét. modérée	9	0	175	504	11997	3	12688
Zones urb	10	102	563	0	4	4775	5454
Total	511	5586	11510	109825	12834	5334	145600

Source : Sentinel-2 Level-2A, Année : 05/2026

L'examen de la configuration fusionnée à travers la matrice de confusion confirme que l'adjonction des indices aux bandes spectrales n'apporte aucune synergie scientifique particulière ***(Tableau 8)***. Pour la classe eau, les modèles restent stables, avec XGBoost qui domine (865 pixels bien classés). Pour le sol nu, SVM (5435 pixels) et Stacking (5427 pixels) surpassent la FA, qui omet 164 pixels vers les zones urbanisées en raison des confusions radiométriques persistantes entre pistes en terre et toitures. La confusion entre végétation dense et végétation modérée reste supérieure au scénario des bandes brutes. Le Stacking dévie 235 pixels denses vers la strate modérée, contre 226 pour le SVM. Pour la végétation clairsemée, le Stacking (10669 pixels) enregistre des omissions diffuses vers la végétation modérée (211 pixels) et les zones urbanisées (249 pixels). Enfin, les zones urbanisées restent robustes sous XGBoost (4760 pixels) et le Stacking (4669 pixels), mais leurs omissions vers la végétation clairsemée demeurent élevées (756 pixels pour le Stacking, 746 pour le SVM), ce qui confirme que la fusion sature les algorithmes par redondance informationnelle et fait des bandes spectrales brutes l'approche la plus efficiente pour cartographier Oyem.

Tableau 8 : *Matrices de confusion issues des combinaisons de variables spectrales et d'indices*

SVM							
Classes Réelles \ Prédites	Eau	Sol nu	Végét. clairsemée	Végét. dense	Végét. modérée	Zone urb	Total
Eau	849	0	5	2	34	1	891
Sol nu	1	5435	55	0	1	99	5591
Végét. clairsemée	176	101	10575	3	215	160	11230
Végét. dense	2	0	0	109518	226	0	109746
Végét. modérée	6	0	134	181	12362	5	12688
Zones urb	3	43	746	0	0	4662	5454
Total	1037	5579	11515	109704	12838	4927	145600
FA							
Classes Réelles \ Prédites	Eau	Sol nu	Végét. clairsemée	Végét. dense	Végét. modérée	Zone urb	Total
Eau	850	0	4	1	35	1	891
Sol nu	5	5329	92	0	1	164	5591
Végét. clairsemée	0	166	10780	5	233	46	11230
Végét. dense	4	0	0	109435	307	0	109746
Végét. modérée	7	0	232	349	12100	0	12688
Zones urb	1	43	736	0	6	4668	5454
Total	867	5538	11844	109790	12682	4879	145600
XGBoost							
Classes Réelles \ Prédites	Eau	Sol nu	Végét. clairsemée	Végét. dense	Végét. modérée	Zone urb	Total
Eau	865	0	4	3	18	1	891
Sol nu	0	5416	69	0	2	104	5591
Végét. clairsemée	5	143	10756	5	214	107	11230
Végét. dense	5	0	0	109430	311	0	109746
Végét. modérée	5	0	236	259	12188	0	12688
Zones urb	0	32	659	0	3	4760	5454
Total	880	5591	11724	109697	12736	4972	145600
Stacking							
Classes Réelles \ Prédites	Eau	Sol nu	Végét. clairsemée	Végét. dense	Végét. modérée	Zone urb	Total
Eau	845	0	4	2	39	1	891
Sol nu	0	5427	64	0	1	99	5591
Végét. clairsemée	2	96	10669	3	211	249	11230
Végét. dense	2	0	0	109509	235	0	109746
Végét. modérée	6	0	169	241	12272	0	12688
Zones urb	1	26	756	0	2	4669	5454
Total	856	5549	11662	109755	12760	5018	145600

Source : Sentinel-2 Level-2A, Année : 05/2026

Une validation visuelle permet de comparer qualitativement les trois meilleurs résultats de classification obtenus pour chaque configuration de variables ***(Figure 5)***. Cette validation confronte l'image satellitaire originale présentée en colonne A à trois classifications thématiques distinctes, à savoir l'Ensemble Learning sur bandes spectrales (B), le SVM sur indices spectraux (C) et le SVM sur combinaisons (D), appliquées à trois environnements précis : une zone routière en milieu forestier (ligne 1), un tissu urbain dense avec plan d'eau (ligne 2) et une infrastructure aéroportuaire (ligne 3). Cette analyse visuelle montre que la méthode C (indices spectraux) souffre d'une forte fragmentation spatiale. Sur les axes routiers et la piste aéroportuaire, elle engendre un effet « poivre-et-sel » agressif qui bruite les infrastructures par de faux pixels d'eau. À l'inverse, les méthodes B (Stacking - bandes) et D (SVM - combinaisons) offrent des tracés continus, rectilignes et fidèles à la géométrie du sol. La divergence majeure s'observe sur le plan d'eau principal. Les méthodes B et D extraient le lac de manière homogène, tandis que la méthode C mite le centre du bassin par des pixels

urbains et végétaux. Cet écueil visuel confirme la mauvaise séparabilité spectrale des surfaces hydriques lorsque les bandes brutes sont exclues au profit des seuls indices. La confrontation quantitative et visuelle confirme la supériorité absolue de la méthode B. En évitant la sur-segmentation de la méthode C et les pixels mixtes de la méthode D, elle valide le meilleur compromis avec un Kappa de 0,9664. Ce modèle optimal est donc retenu pour quantifier de manière fiable les superficies d'occupation du sol d'Oyem.

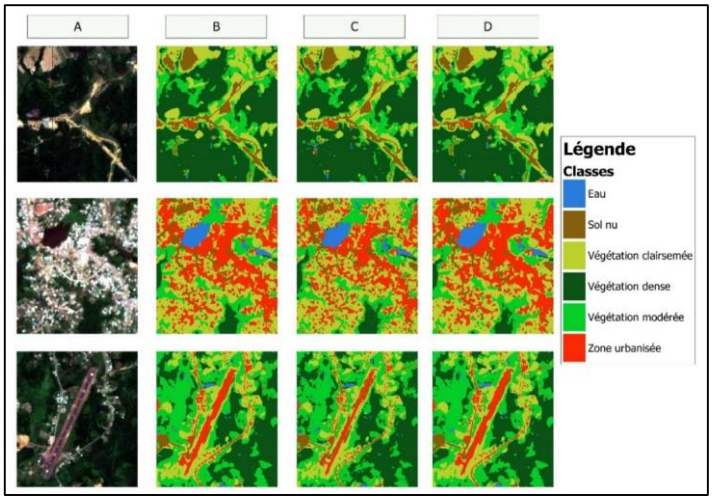


Figure 5 : *Comparaison visuelle des résultats de classification thématique*

3-3. Occupation du sol

L'évaluation de la répartition quantitative de l'occupation du sol de la zone d'étude met en lumière la domination écrasante de la matrice forestière sur l'ensemble du territoire ***(Tableau 9)***. Les superficies occupées par les différentes classes ainsi que leurs proportions respectives révèlent ainsi la prépondérance de la végétation dense face aux emprises anthropiques et aux plans d'eau ***(Tableau 9)***.

Tableau 9 : *Répartition des classes d'occupation du sol de la zone d'étude*

Classes d'occupation du sol	Superficie (ha)	Pourcentage (%)
Végétation dense	14 270,69	71,56
Végétation modérée	2 408,95	12,08
Végétation clairsemée	2 192,73	10,99
Zone urbanisée	729,952	3,66
Sol nu	303,427	1,52
Eau	36,992	0,19
Total	19 942,74	100

Source : Sentinel-2 Level-2A, Année : 05/2026 ; Projection : Système de coordonnées : WGS 84/UTM Zone 32 N

Le paysage, est saturé à 94,63 % par les formations végétales, dominées par la végétation dense (71,56 %), devant les strates modérée (12,08 %) et clairsemée (10,99 %). L'emprise anthropique reste très localisée et contenue, la zone urbanisée couvrant 3,66 % du territoire et le sol nu 1,52 %. Enfin, les surfaces en eau occupent une part totalement marginale avec seulement 0,19 % de la superficie globale d'Oyem ***(Figure 6)***.

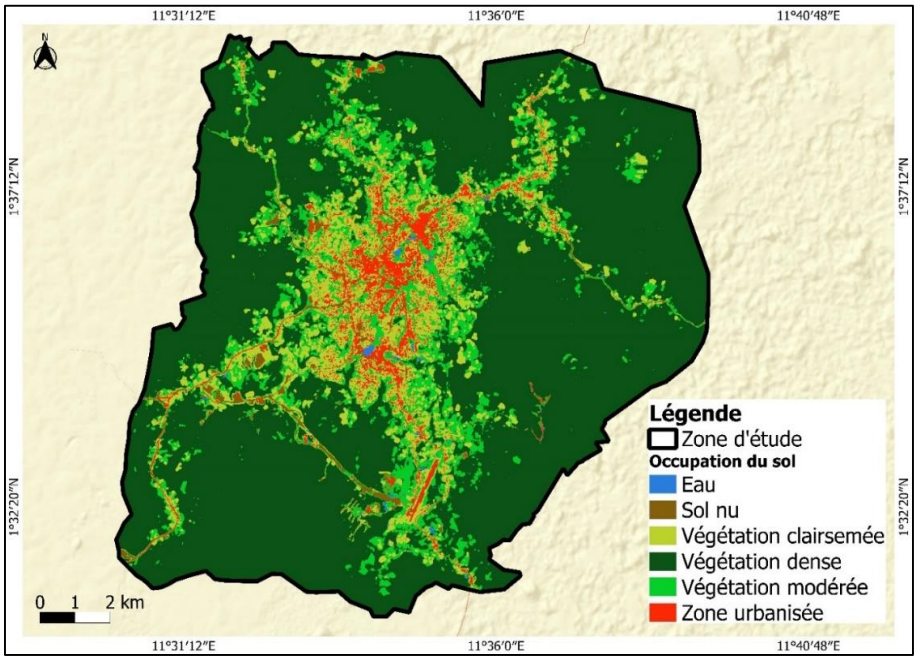


Figure 6 : *Occupation du sol issue de la relation Ensemble Learning-bandes spectrales de la ville d'Oyem et ses périphéries*

3-4. Synthèse de la robustesse des classifieurs selon les descripteurs

L'évaluation comparative des performances montre que l'utilisation des descripteurs spectraux influence de manière significative la robustesse des modèles de classification ***(Tableau 10).*** Les bandes spectrales servent de référence, car les quatre modèles obtiennent leur meilleur Kappa soit 0,9605 pour la Forêt Aléatoire, 0,9647 pour le SVM,0,9651 pour XGBoost et 0,9664 pour le Stacking, qui atteint la valeur maximale du tableau. Ces valeurs correspondent toutes à un accord presque parfait selon l'échelle de Landis et Koch et l'écart entre le meilleur et le moins bon modèle reste faible (0,0059). Lorsque les indices spectraux sont utilisés seuls, tous les modèles perdent en performance avec des baisses comprises entre -0,0107 et -0,0172. XGBoost enregistre la plus forte dégradation (-0,0172, soit -1,78 %), suivi de la Forêt Aléatoire (-0,0145 ; -1,51 %) et de le Stacking (-0,0142 ; -1,47 %), tandis que le SVM résiste le mieux (-0,0107 ; -1,11 %). Cette stabilité relative place le SVM en tête dans cette configuration (0,9540) devant le Stacking (0,9522), XGBoost (0,9479) et la Forêt Aléatoire (0,9460). Dans cette configuration, de notre étude, nous observons un écart entre les modèles qui s'élargit à 0,0080. Ce qui indique que les indices seuls rendent les performances plus dépendantes de l'algorithme utilisé. Cette baisse s'explique par le fait que les indices résument l'information de plusieurs bandes en un seul rapport, ce qui réduit les informations dont disposent les modèles pour distinguer les classes, notamment la classe eau. À l'inverse, avec la combinaison des bandes et des indices, les performances restent très proches de celles des bandes seules. Le Kappa diminue de -0,0003 pour le SVM (-0,03 %), de -0,0005 pour la Forêt Aléatoire (-0,05 %), de -0,0009 pour XGBoost (-0,09 %) et de -0,0023 pour le Stacking (-0,24 %), ce dernier étant le plus pénalisé. Ces pertes, toutes inférieures à 0,25 % montrent que l'ajout des indices n'apporte aucun gain, car ils sont calculés à partir des mêmes bandes et ne fournissent donc pas d'information nouvelle. Dans cette configuration, les quatre modèles se resserrent encore (écart de 0,0044), le SVM obtient le meilleur Kappa (0,9644), très proche de ceux de XGBoost (0,9642) et du Stacking (0,9641). La Forêt Aléatoire présente quant à elle le Kappa le plus bas dans les trois configurations. Au finale, le Stacking est le meilleur modèle qu'avec les bandes spectrales seules, alors que le SVM se montre le plus robuste face au changement de descripteurs. Les bandes spectrales constituent donc le descripteur le plus performant, quel que soit le modèle utilisé ***(Tableau 10)***.

Tableau 10 : *Évolution de la robustesse (Indice Kappa) des classifieurs SVM, FA, XGB et Ensemble Learning selon les descripteurs utilisés*

Algorithme	Variables	Indice Kappa	Variation brute (Δ)	Évolution relative (%)
XGBoost	Bandes spectrales (T5)	0,9651	Réf.	Réf.
XGBoost	Indices spectraux (T6)	0,9479	-0,0172	-1,78 %
XGBoost	Bandes + Indices (T7)	0,9642	-0,0009	-0,09 %
SVM	Bandes spectrales (T5)	0,9647	Réf.	Réf.
SVM	Indices spectraux (T6)	0,9540	-0,0107	-1,11 %
SVM	Bandes + Indices (T7)	0,9644	-0,0003	-0,03 %
Stacking	Bandes spectrales (T5)	0,9664	Réf.	Réf.
Stacking	Indices spectraux (T6)	0,9522	-0,0142	-1,47 %
Stacking	Bandes + Indices (T7)	0,9641	-0,0023	-0,24 %
FA	Bandes spectrales (T5)	0,9605	Réf.	Réf.
FA	Indices spectraux (T6)	0,9460	-0,0145	-1,51 %
FA	Bandes + Indices (T7)	0,9600	-0,0005	-0,05 %

Source : Sentinel-2 Level-2A, Année : 05/2026 ; Projection : Système de coordonnées : WGS 84/UTM Zone 32 N

4. Discussion

La mise en place de procédures précises de cartographie et de suivi est fondamentale pour distinguer les systèmes d'occupation du sol et assurer une gestion durable des ressources territoriales [34 - 37]. Dans ce cadre, l'utilisation de modèles d'apprentissage automatique couplés aux données satellitaires se développe à l'échelle mondiale pour répondre à la complexité des paysages [38, 39]. Cette étude mobilise des machines à vecteurs de support (SVM), des forêts aléatoires (FA) et XGBoost (XGB), ainsi qu'une approche d'Ensemble Learning (Stacking). Ces modèles sont entraînés sur des échantillons de référence issus de l'interprétation visuelle de compositions colorées Sentinel-2 et validés par une validation croisée spatiale. Les résultats de cette étude présentent un double intérêt. Sur le plan méthodologique, ils indiquent la combinaison d'algorithmes et de descripteurs à retenir pour cartographier une zone forestière tropicale à partir de Sentinel-2. Sur le plan opérationnel, ils fournissent un état de référence de l'occupation du sol d'Oyem et de sa périphérie, utile à l'aménagement de ce territoire.

4-1. Précision des modèles et importance de l'Ensemble Learning

Le premier objectif de cette étude consiste à comparer les capacités de généralisation des algorithmes SVM, FA et XGBoost à celle de leur intégration au sein d'un modèle d'Ensemble Learning, en utilisant l'indice Kappa comme juge de paix statistique afin de comparer l'ensemble des configurations testées et d'identifier celle qui surpasse les autres. Parmi tous les modèles et toutes les configurations, le Stacking associé aux bandes spectrales obtient le meilleur avec un Kappa de 0,9664. Il surpasse XGBoost (0,9651 ; +0,0013), le SVM (0,9647 ; +0,0017) et la Forêt Aléatoire (0,9605 ; +0,0059). Ce résultat s'explique par le principe même du Stacking ou le méta-modèle exploite la complémentarité des frontières de décision des classifieurs de base. Dans cette étude, elle exploite le Boosting de XGBoost qui corrige séquentiellement les erreurs de

classification, l'optimisation de l'hyperplan séparateur du SVM et le Bagging de la Forêt Aléatoire. En fusionnant ces prédictions, le Stacking compense les biais propres à chaque algorithme. Ce résultat concorde avec la littérature récente, selon laquelle les architectures ensemblistes surpassent fréquemment les modèles uniques sur des paysages complexes et hétérogènes [13,40,41]. Il convient toutefois de noter que cet avantage est lié aux bandes spectrales. Car, avec les indices spectraux seuls et avec la combinaison des descripteurs, le SVM obtient un Kappa légèrement supérieur à celui du Stacking (0,9540 contre 0,9522, puis 0,9644 contre 0,9641). Par ailleurs, les modèles pris isolément présentent des vulnérabilités localisées. Le SVM est plus sensible aux confusions de la classe eau, car il y affecte par commission 131 pixels de végétation clairsemée. La Forêt Aléatoire affiche le Kappa le plus bas (0,9605) et omet 316 pixels de végétation dense au profit de la végétation modérée, contre 243 pour le Stacking. XGBoost, dont le Kappa (0,9651) est très proche de celui du Stacking, reste le modèle isolé le plus performant avec les bandes spectrales. Cette hiérarchie des performances est comparable à celle rapportée dans des travaux sur la classification des sols [42],de la végétation des marais [43] et des risques d'incendies de forêt [44]. Pour la zone d'Oyem et sa périphérie, le Stacking associé aux bandes spectrales est donc la configuration qui a donné le meilleur Kappa.

4-2. Impact des jeux de variables

En réponse à la question de l'influence des descripteurs sur la qualité de la cartographie (objectif 2), les résultats montrent que les bandes spectrales seules offrent les meilleures performances pour l'ensemble des modèles testés. Ce résultat rejoint ceux de travaux menés à l'échelle nationale pour discriminer les structures forestières [19] et cartographier les zones agricoles [17, 18],selon lesquels les bandes spectrales seules suffisent à discriminer les classes. La littérature en télédétection souligne pourtant fréquemment l'utilité d'indices spectraux comme le NDVI ou le SAVI pour capturer les signatures phénologiques [45]. Or, dans cette étude, ni les indices spectraux seuls ni leur combinaison avec les bandes n'améliorent la précision par rapport aux bandes spectrales brutes. Avec les indices spectraux seuls, le Kappa baisse pour tous les modèles -0,0107 (-1,11 %) pour le SVM, -0,0142 (-1,47 %) pour le Stacking (Ensemble Learning), -0,0145 (-1,51 %) pour la Forêt Aléatoire et -0,0172 (-1,78 %) pour XGBoost. Le Kappa de l'Ensemble Learning passe ainsi de 0,9664 à 0,9522. Cette baisse, modérée au regard de précisions globales très élevées est systématique. Elle s'explique par une perte d'information ou chaque indice résume plusieurs bandes en un seul rapport, ce qui prive les modèles de réflectances qui discriminent certaines classes. Nous l'observons notamment avec la classe eau où 332 à 356 de ses pixels sont classés en zone urbanisée avec les indices, contre 0 ou 1 avec les bandes. Ce résultat diffère de l'intérêt accordé aux indices spectraux dans des travaux comparant indices et techniques de classification [46], dont le contexte (type de couverture, capteur) diffère toutefois de celui de cette étude. Avec la combinaison des bandes et des indices, les performances restent très proches de celles des bandes seules, sans les dépasser. Le Kappa diminue de -0,0003 (-0,03 %) pour le SVM, de -0,0005 (-0,05 %) pour la FA, de -0,0009 (-0,09 %) pour XGBoost et de -0,0023 (-0,24 %) pour l'Ensemble Learning, dont le Kappa passe de 0,9664 à 0,9641. Ces pertes sont faibles mais constantes, ce qui indique que l'ajout des indices n'apporte pas d'information utile. Les indices étant calculés à partir des mêmes bandes, les variables ajoutées sont en grande partie redondantes. Elles accroissent la dimension du problème sans améliorer la séparation des classes. Ce constat s'oppose à l'idée selon laquelle l'ajout d'indices calculés aux réflectances d'origine améliore l'exactitude des classifieurs [47]. De ce fait, les bandes spectrales brutes constituent l'approche la plus efficiente pour la zone d'étude. Elles donnent le meilleur Kappa pour les quatre modèles avec 10 variables, contre 16 pour la combinaison bandes-indices.

4-3. Quantification des superficies et fiabilité thématique

En réponse au dernier objectif de cette étude, la quantification des superficies permet de dresser l'état de l'occupation du sol d'Oyem et de sa périphérie. Les résultats montrent la coexistence d'un noyau forestier dominant et de zones anthropisées localisées. L'omniprésence des formations végétales, qui totalisent 94,63 % de la superficie cartographiée, rejoint la persistance du bloc forestier septentrional du Gabon [48], caractéristique de la résilience à grande échelle des écosystèmes du bassin du Congo [49]. La forte proportion de végétation dense (71,56 %) indique un couvert largement fermé. Les classes de végétation modérée (12,08 %) et clairsemée (10,99 %) traduisent des perturbations locales, comme l'agriculture vivrière qui fragmentent le paysage sans suffire à indiquer, à elles seules, une déforestation massive. Ce constat s'aligne avec les travaux au Gabon, selon qui l'anthropisation ancienne a eu peu d'impact sur la dynamique forestière globale, l'exploitation forestière constituant le principal risque à long terme [49]. Une gestion rationnelle du massif forestier apparaît donc essentielle à son développement durable. L'organisation spatiale des zones urbanisées (3,66 %) et des sols nus (1,52 %) révèle la signature géographique de l'anthropisation. Leur concentration le long du réseau routier et des axes d'échange traduit un développement linéaire, typique des fronts de pénétration en zone tropicale et rejoint les travaux montrant le rôle structurant des infrastructures de transport dans l'étalement du bâti et des sols mis à nu [50]. Le sol nu, minoritaire en surface (303,43 ha) correspond à des ouvertures du milieu lié aux aménagements ou aux rotations culturales, hypothèse qu'une analyse multi-dates permet de vérifier. Les surfaces en eau ferment ce bilan avec une part très marginale de 0,19 % (36,99 ha). Cet inventaire constitue ainsi un état de référence de l'occupation du sol, offrant aux gestionnaires une base pour planifier le développement régional et pour suivre l'évolution du couvert.

4-4. Avancées, limites et transférabilité

Cette étude apporte plusieurs avancées, en lien avec les objectifs fixés. Premièrement, elle compare sur une même zone et avec un même protocole, trois algorithmes (SVM, FA et XGBoost) et leur combinaison par Stacking, ce qui permet de situer l'apport de l'approche ensembliste par rapport aux modèles isolés. Elle applique ainsi une approche de Stacking à la cartographie de l'occupation du sol dans le nord du Gabon où ces approches restent peu informées. Deuxièmement, elle évalue l'effet de trois descripteurs (bandes spectrales, indices spectraux et leur combinaison) et montre que les bandes spectrales brutes suffisent pour obtenir les meilleures performances, ce qui simplifie la chaîne de traitement. Troisièmement, elle fournit une quantification des superficies de chaque classe d'occupation du sol pour la commune d'Oyem et sa périphérie, utilisable comme état de référence. Sur le plan méthodologique, l'avancée principale est un cadre d'évaluation associant le partitionnement spatial par DBSCAN et une validation croisée spatiale par blocs (Group-KFold à 5 plis). Contrairement aux approches fondées sur un découpage purement aléatoire des pixels, qui tendent à surestimer les métriques en raison de l'autocorrélation spatiale intra-parcellaire. Le protocole de cette étude limite fortement la fuite d'information entre les données d'apprentissage et de validation. Les précisions globales (OA) et les indices Kappa (κ) obtenus fournissent ainsi une estimation plus réaliste de la capacité des algorithmes à faire face à l'hétérogénéité des paysages d'Oyem et ses environs. La méthode présente toutefois des limites. La forte nébulosité des régions équatoriales restreint la disponibilité d'images optiques sans nuages, ce qui perturbe la régularité des observations et altère la qualité radiométrique du signal [18, 51, 52]. L'interpolation sous Google Earth Engine comble les zones masquées (nuages et ombres), mais elle ne garantit pas une reconstitution fidèle des pixels réels de ces zones. L'étude repose en outre sur une seule image (18 mars 2026), qui ne permet pas de prendre en compte la saisonnalité de la végétation. Enfin, les modèles entraînés restent propres à leur zone d'apprentissage. Leur application à une autre province du Gabon ou autre pays exige un réentraînement ou une adaptation des échantillons de référence. Le cadre méthodologique (prétraitement, validation spatiale, comparaison des descripteurs) est en revanche reproductible dans d'autres contextes, sa transférabilité restant à tester.

5. Conclusion

Cette étude compare les performances des modèles SVM, FA et XGBoost à celles d'un Stacking, selon trois jeux de variables, pour cartographier l'occupation du sol d'Oyem et de sa périphérie à partir d'une image Sentinel-2 de 2026. Les trois objectifs spécifiques sont atteints. D'une part, le Stacking obtient le meilleur Kappa (0,9664) avec les bandes spectrales seules, mais son avantage sur XGBoost (0,9651), le SVM (0,9647) et la FA (0,9605) reste faible et ne se confirme pas avec les indices seuls ou la combinaison bandes-indices, où le SVM est légèrement supérieur. D'autre part, les bandes spectrales brutes constituent le meilleur jeu de variables pour tous les modèles. Les indices seuls dégradent le Kappa (Stacking : 0,9522) à cause des confusions sur la classe eau et leur ajout aux bandes n'apporte aucun gain. Enfin, la classification optimale à la période de la prise de l'image montre un territoire dominé par la végétation (94,63 %), dont 71,56 % de végétation dense (14 270,69 ha), avec une emprise anthropique limitée (zone urbanisée : 729,95 ha, soit 3,66 % ; sol nu : 303,43 ha, soit 1,52 %) concentrée le long des axes routiers. La méthode fournit ainsi une cartographie fiable de référence. Elle gagnerait à être complétée par des images multi-dates, des données radar Sentinel-1 et une validation sur données de terrain, avant son extension à l'ensemble du territoire gabonais.

Références

[1] - K. ENNOURI, S. SMAOUI et M. A. TRIKI, "Detection of urban and environmental changes via remote sensing", Circular Economy and Sustainability, 1 (4) (2021) 1423 - 1437, https://doi.org/10.1007/s43615-021-00035-y

[2] - P. AZIZI, A. SOLTANI, F. BAGHERI, S. SHARIFI et M. MIKAEILI, "An Integrated Modelling Approach to Urban Growth and Land Use/Cover Change", Land, 11 (10) (2022) 1715, https://doi.org/10.3390/land11101715

[3] - G. A. AFUYE, L. NDUKU, A. M. KALUMBA, C. A. G. SANTOS, I. R. ORIMOLOYE, V. N. OJEH, K. H. THAMAGA et P. SIBANDZE, "Global trend assessment of land use and land cover changes: A systematic approach to future research development and planning», Journal of King Saud University - Science,36 (7) (2024) 103262, https://doi.org/10.1016/j.jksus.2024.103262

[4] - N. T. TUAN, "Longitudinal Assessment of Land Use Change Impacts on Carbon Services in the Southeast Region, Vietnam", Geographies, 5 (4) (2025) 62, https://doi.org/10.3390/geographies5040062

[5] - P. K. NYONGO et J. N. KIMENGSI, "Socio-Ecological Outcomes of Forest Landscape Mutations in the Congo Basin: Learning from Cameroon", Land, 14 (8) (2025) 1644, https://doi.org/10.3390/land14081644

[6] - MINISTÈRE DES EAUX ET FORÊTS, DE LA MER, DE L'ENVIRONNEMENT, CHARGÉ DU PLAN CLIMAT ET DU PLAN D'AFFECTATION DES TERRES, "Evaluation environnementale et sociale stratégique de la stratégie nationale de réduction des émissions pour le secteur forestier et l'élaboration des instruments de gestion environnementale et sociale y afférents", Rapport, (2022), https://echosdeleco.com/articles/178 (Mai 2026)

[7] - V. NASIRI, A. DELJOUEI, F. MORADI, S. M. M. SADEGHI et S. A. BORZ, "Land Use and Land Cover Mapping Using Sentinel-2, Landsat-8 Satellite Images, and Google Earth Engine: A Comparison of Two Composition Methods", Remote Sensing, 14 (9) (2022) 1977, https://doi.org/10.3390/rs14091977

[8] - Y. YANG, D. YANG, X. WANG, Z. ZHANG et Z. NAWAZ, "Testing Accuracy of Land Cover Classification Algorithms in the Qilian Mountains Based on GEE Cloud Platform", Remote Sensing, 13 (24) (2021) 5064, https://doi.org/10.3390/rs13245064

[9] - C. SOVANN, S. OLIN, A. MANSOURIAN, S. SAKHOEUN, S. PREY, S. KOK et T. TAGESSON, "Importance of Spectral Information, Seasonality, and Topography on Land Cover Classification of Tropical Land Cover Mapping", Remote Sensing, 17 (9) (2025) 1551, https://doi.org/10.3390/rs17091551

[10] - C. A. RAMEZAN, T. A. WARNER, A. E. MAXWELL et B. S. PRICE, "Effects of Training Set Size on Supervised Machine-Learning Land-Cover Classification of Large-Area High-Resolution Remotely Sensed Data", Remote Sensing, 13 (3) (2021) 368, https://doi.org/10.3390/rs13030368

[11] - P. ZHANG, Y. JIA et Y. SHANG, "Research and application of XGBoost in imbalanced data", *International Journal of Distributed Sensor Networks*, 18 (6) (2022) 155013292211069, https://doi.org/10.1177/15501329221106935

[12] - C. PENG, Z. XIE et X. JIN, "Using Ensemble Learning for Remote Sensing Inversion of Water Quality Parameters in Poyang Lake", Sustainability, 16 (8) (2024) 3355, https://doi.org/10.3390/su16083355

[13] - F. SHI, X. GAO, R. LI et H. ZHANG, "Ensemble Learning for the Land Cover Classification of Loess Hills in the Eastern Qinghai–Tibet Plateau Using GF-7 Multitemporal Imagery", Remote Sensing, 16 (14) (2024) 2556, https://doi.org/10.3390/rs16142556

[14] - M. R. SUBEDI, C. PORTILLO-QUINTERO, N. E. MCINTYRE, S. S. KAHL, R. D. COX, G. PERRY et X. SONG, "Ensemble machine learning on the fusion of sentinel time series imagery with high-resolution orthoimagery for improved land use/land cover mapping", Remote Sensing, 16 (15) (2024) 2778, https://doi.org/10.3390/rs16152778

[15] - S. CUYPERS, A. NASCETTI et M. VERGAUWEN, "Land Use and Land Cover Mapping with VHR and Multi-Temporal Sentinel-2 Imagery", Remote Sensing, 15 (10) (2023) 2501, https://doi.org/10.3390/rs15102501

[16] - Y. LI, Z. ZHAO, Y. XIN, A. XU, S. XIE, Y. YAN et L. WANG, "How Are Land-Use/Land-Cover Indices and Daytime and Nighttime Land Surface Temperatures Related in Eleven Urban Centres in Different Global Climatic Zones?", Land, 11 (8) (2022) 1312, https://doi.org/10.3390/land11081312

[17] - M. M. MASSALA, B. NKOUMAKALI et L. OBIANG OBIANG, "Occupation du sol et identification des cultures agricoles à partir des algorithmes d'apprentissage automatique", Perspectives Plurielles, Tome II, Série 1 (3) (2026) 159 - 175, https://hal.science/hal-05755207

[18] - L. OBIANG OBIANG, "Analyse comparative des approches semi-automatique et d'apprentiss l'identification des zones agricoles familiales à Lébamba et ses environs", Mémoire de master, Libreville, (2026)

[19] - A. WAŚNIEWSKI, A. HOŚCIŁO, B. ZAGAJEWSKI et D. MOUKÉTOU-TARAZEWICZ, "Assessment of Sentinel-2 Satellite Images and Random Forest Classifier for Rainforest Mapping in Gabon", Forests, 11 (9) (2020) 941, https://doi.org/10.3390/f11090941

[20] - N. KARASIAK, J.-F. DEJOUX, C. MONTEIL et D. SHEEREN, "Spatial dependence between training and test sets: another pitfall of classification accuracy assessment in remote sensing", Machine Learning, 111 (7) (2022) 2715 - 2740, https://doi.org/10.1007/s10994-021-05972-1

[21] - I. DIÉDHIOU, C. MERING, O. SY et T. SANÉ, "Cartographier par télédétection l'occupation du sol et ses changements: application à l'analyse de la dynamique des paysages forestiers sénégambiens entre 1972 et 2016", EchoGéo, 54 (2020), https://doi.org/10.4000/echogeo.20510

[22] - M. Á. OBREGÓN, G. RODRIGUES, M. J. COSTA, M. POTES et A. M. SILVA, "Validation of ESA Sentinel-2 L2A Aerosol Optical Thickness and Columnar Water Vapour during 2017–2018", Remote Sensing, 11 (14) (2019) 1649, https://doi.org/10.3390/rs11141649

[23] - N. Q. MINH, N. T. T. HUONG, P. Q. KHANH, L. P. HIEN et D. T. BUI, "Impacts of Resampling and Downscaling Digital Elevation Model and Its Morphometric Factors: A Comparison of Hopfield Neural Network, Bilinear, Bicubic, and Kriging Interpolations", Remote Sensing, 16 (5) (2024) 819, https://doi.org/10.3390/rs16050819

[24] - S. MANJUNATHA et M. M. PATIL, "Interpolation Techniques in Image Resampling", *International Journal of Engineering and Technology*, 7 (3.34) (2018) 567 - 570, https://doi.org/10.14419/ijet.v7i3.34.19383

[25] - C. J. TUCKER, "Red and photographic infrared linear combinations for monitoring vegetation", Remote Sensing of Environment, 8 (2) (1979) 127 - 150, https://doi.org/10.1016/0034-4257(79)90013-0

[26] - A. HUETE, K. DIDAN, T. MIURA, E. P. RODRIGUEZ, X. GAO et L. G. FERREIRA, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices", Remote Sensing of Environment, 83 (1-2) (2002) 195 - 213, https://doi.org/10.1016/S0034-4257(02)00096-2

[27] - A. R. HUETE, "A soil-adjusted vegetation index (SAVI)", Remote Sensing of Environment, 25 (3) (1988) 295 - 309, https://doi.org/10.1016/0034-4257(88)90106-X

[28] - Y. ZHA, J. GAO et S. NI, "Use of normalized difference built-up index in automatically mapping urban areas from TM imagery", International Journal of Remote Sensing, 24 (3) (2003) 583 - 594, https://doi.org/10.1080/01431160304987

[29] - H. XU, "Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery", International Journal of Remote Sensing, 27 (14) (2006) 3025- 3033, https://doi.org/10.1080/01431160600589179

[30] - K. ROUIBAH et M. BELABBAS, "Applying Multi-Index approach from Sentinel-2 Imagery to Extract Urban Area in dry season (Semi-Arid Land in North East Algeria)", *Revista de Teledetección*, 56 (2020) 89, https://doi.org/10.4995/raet.2020.13787

[31] - P. TZIACHRIS, M. Nikou, V. Aschonitis, A. Kallioras, K. Sachsamanoglou, M. Fidelibus et E. Tziritis "Spatial or Random Cross-Validation? The Effect of Resampling Methods in Predicting Groundwater Salinity with Machine Learning in Mediterranean Region", Water, 15 (12) (2023) 2278, https://doi.org/10.3390/w15122278

[32] - C. I. ALVAREZ, A. GOVIND, A. M. BOLLAS, K. WAHA et A. LABBACI, "Seasonal agricultural vulnerability in semi-arid Morocco: combining remote sensing and farmer knowledge to inform climate adaptation", Climatic Change, 179 (4) (2026) 64, https://doi.org/10.1007/s10584-026-04160-1

[33] - J. R. LANDIS et G. G. KOCH, "A one-way components of variance model for categorical data ", Biometrics, 33 (4) (1977) 671, https://doi.org/10.2307/2529310

[34] - S. GADAL et G. MOZGERIS, "Advances of Remote Sensing in Land Cover and Land Use Mapping", Remote Sensing, 17 (12) (2025) 1980, https://doi.org/10.3390/rs17121980

[35] - K. ABIDA, M. BARBOUCHI, K. BOUDABBOUS, W. TOUKABRI, K. SAAD, H. BOUSNINA et T. SAHLI CHAHED," Sentinel-2 data for land use mapping: comparing different supervised classifications in semi-arid areas", Agriculture, 12 (9) (2022) 1429, https://doi.org/10.3390/agriculture12091429

[36] - P. SZAREK-IWANIUK, A. DAWIDOWICZ et A. SENETRA,"Methodology for Precision Land Use Mapping towards Sustainable Urbanized Land Development", International Journal of Environmental Research and Public Health, 19 (6) (2022) 3633, https://doi.org/10.3390/ijerph19063633

[37] - P. SZAREK-IWANIUK, "A comparative analysis of spatial data and land use/land cover classification in urbanized areas and areas subjected to anthropogenic pressure for the example of Poland", Sustainability, 13 (6) (2021) 3070, https://doi.org/10.3390/su13063070

[38] - K. KRYZIA, A. RADZIEJOWSKA, J. ADAMCZYK et D. KRYZIA," Analysis of Using Machine Learning Application Possibilities for the Detection and Classification of Topographic Objects", ISPRS International Journal of Geo-Information, 15 (2) (2026) 59, https://doi.org/10.3390/ijgi15020059

[39] - J. TAMÁS, A. LOUIS, Z. Z. FEHÉR et A. NAGY, "Land Cover Mapping Using High-Resolution Satellite Imagery and a Comparative Machine Learning Approach to Enhance Regional Water Resource Management", Remote Sensing, 17 (15) (2025) 2591, https://doi.org/10.3390/rs17152591

[40] - R. LI, X. GAO et F. SHI, "A Framework for Subregion Ensemble Learning Mapping of Land Use/Land Cover at the Watershed Scale", Remote Sensing, 16 (20) (2024) 3855, https://doi.org/10.3390/rs16203855

[41] - P. BARMAN, S. MUSTAK, M. KUFFER et S. K. SINGH, "Transfer-Ensemble Learning: A Novel Approach for Mapping Urban Land Use/Cover of the Indian Metropolitans", Sustainability, 15 (24) (2023) 16593, https://doi.org/10.3390/su152416593

[42] - F. ZHU, C. ZHU, W. LU, Z. FANG, Z. LI et J. PAN, "Soil Classification Mapping Using a Combination of Semi-Supervised Classification and Stacking Learning (SSC-SL)", Remote Sensing, 16 (2) (2024) 405, https://doi.org/10.3390/rs16020405

[43] - H. YAO, B. Fu, Y. Zhang, S. Li, S. Xie, J. Qin, D. Fan et E. Gao "Combination of Hyperspectral and Quad-Polarization SAR Images to Classify Marsh Vegetation Using Stacking Ensemble Learning Algorithm", Remote Sensing, 14 (21) (2022) 5478, https://doi.org/10.3390/rs14215478

[44] - Y. LI, G. LI, K. WANG, Z. WANG et Y. CHEN, "Forest Fire Risk Prediction Based on Stacking Ensemble Learning for Yunnan Province of China", Fire, 7 (1) (2023) 13, https://doi.org/10.3390/fire7010013

[45] - A. D. SAKTI, A. I. FAUZI, F. N. WILWATIKTA, Y. S. RAJAGUKGUK, S. A. SUDHANA, L. F. YAYUSMAN, L. N. SYAHID, T. SRITARAPIPAT, J. A. PRINCIPE, N. T. Q. TRANG, E. SULISTYAWATI, I. UTAMI, C. W. ARIEF et K. WIKANTIKA," Multi-Source Remote Sensing Data Product Analysis: Investigating Anthropogenic and Naturogenic Impacts on Mangroves in Southeast Asia", Remote Sensing, 12 (17) (2020) 2720, https://doi.org/10.3390/rs12172720

[46] - A. J. STERN, C. S. T. DAUGHTRY, E. R. HUNT et F. GAO, "Comparison of Five Spectral Indices and Six Imagery Classification Techniques for Assessment of Crop Residue Cover Using Four Years of Landsat Imagery", Remote Sensing, 15 (18) (2023) 4596, https://doi.org/10.3390/rs15184596

[47] - Y. ZHANG, W. YANG, Y. SUN, C. CHANG, J. YU et W. ZHANG, "Fusion of Multispectral Aerial Imagery and Vegetation Indices for Machine Learning-Based Ground Classification", Remote Sensing, 13 (8) (2021) 1411, https://doi.org/10.3390/rs13081411

[48] - F. NZIGOU BOUCKA, C. V. OBAME, F. MANFOUMBI, A. NZUE MBA, M. NGUI ONDO, V. OVONO et A. MAMBIMBA NDJOUNGUI, "Cartographie de l'occupation du sol du Gabon en 2015 - Changements entre 2010 et 2015", Revue Française de Photogrammétrie et de Télédétection, 223 (2021) 118 - 128, https://doi.org/10.52638/rfpt.2021.567

[49] - B. PEYROT, "Incidences écologiques, anthropiques et paléoécologiques sur l'évolution des forêts du Gabon : Essai de synthèse", Cahiers d'Outre-Mer, 61 (241-242) (2008) 111 - 144, https://doi.org/10.4000/com.3672

[50] - G. ZHAO, X. ZHENG, Z. YUAN et L. ZHANG, "Spatial and temporal characteristics of road networks and urban expansion", Land, 6 (2) (2017) 30, https://doi.org/10.3390/land6020030

[51] - Y. LIAO, S. LI, M. GAO, S. LI, W. QIN, Q. XIONG, C. LIN, Q. CHEN and P. TAO, "Optical and SAR image registration in equatorial cloudy regions guided by automatically point-prompted cloud masks", Remote Sensing, 17 (15) (2025) 2630, https://doi.org/10.3390/rs17152630

[52] - C. HUANG, C. ZHANG, Y. HE, Q. LIU, H. LI, F. SU, G. LIU et A. BRIDHIKITTI, "Land cover mapping in cloud-prone tropical areas using Sentinel-2 data: integrating spectral features with NDVI temporal dynamics", Remote Sensing, 12 (7) (2020) 1163, https://doi.org/10.3390/rs12071163