

## **Cartographie des infrastructures vertes de Lubumbashi, l'importance de la prise en compte des variations saisonnières de la signature spectrale**

**Auteur :** Estenne, Gaspard

**Promoteur(s) :** Bastin, Jean-François

**Faculté :** Gembloux Agro-Bio Tech (GxABT)

**Diplôme :** Master en bioingénieur : gestion des forêts et des espaces naturels, à finalité spécialisée

**Année académique :** 2021-2022

**URI/URL :** <http://hdl.handle.net/2268.2/15446>

---

### *Avertissement à l'attention des usagers :*

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

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

---

# Cartographie des infrastructures vertes de Lubumbashi, l'importance de la prise en compte des variations saisonnnières de la signature spectrale.

Gaspard Estenne

TRAVAIL DE FIN D'ÉTUDES PRÉSENTÉ EN VUE DE  
L'OBTENTION DU DIPLÔME DE MASTER BIOINGÉNIEUR EN  
GESTION DES FORÊTS ET DES ESPACES NATURELS

Année académique 2021 - 2022

PROMOTEUR: Jean-François Bastin

ENCADRANT: Harold Strammer



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

Le présent document n'engage que son auteur.

© Any reproduction of this document, by any means whatsoever, is only allowed with the authorization of the author and the academic authority of Gembloux Agro-Bio Tech.

This document reflects only the views of its author.



# Cartographie des infrastructures vertes de Lubumbashi, l'importance de la prise en compte des variations saisonnnières de la signature spectrale.

Gaspard Estenne

TRAVAIL DE FIN D'ÉTUDES PRÉSENTÉ EN VUE DE  
L'OBTENTION DU DIPLÔME DE MASTER BIOINGÉNIEUR EN  
GESTION DES FORÊTS ET DES ESPACES NATURELS

Année académique 2021 - 2022

PROMOTEUR: Jean-François Bastin

ENCADRANT: Harold Strammer

# Remerciements

J'aimerais tout d'abord remercier mes parents, sans qui je ne serais clairement pas là, pour leur soutien inconditionnel tout au long de mes études.

Je remercie ensuite Harold de m'avoir supporté pendant le mois passé ensemble à Lubumbashi, pour ses conseils avisés tout au long de mon mémoire et pour tous ces fantastiques moments passés ensemble.

Je remercie mon promoteur de m'avoir offert l'opportunité de réaliser ce mémoire, ainsi que pour sa patience et ses recommandations tout au long de ce travail.

Je remercie ma femme pour tout ce qu'elle a pu m'apporter dans ma vie, de l'intellectuel au culinaire en passant par la rigolade.

Je remercie ensuite mes soeurs, mes amis proches et mes cokotteurs.

J'aimerais finalement remercier le Prof Useni, Héritier, Ildephonse, Serge et Alain pour leur aide indispensable lors de la mission terrain.

Je remercie finalement tous les humains ayant relu ce travail pour le temps qu'ils y ont consacré et pour leurs remarques.

Le voyage réalisé dans le cadre du présent travail/stage a été rendu possible grâce au soutien financier de l'Académie de recherche et d'enseignement supérieur de la Fédération Wallonie-Bruxelles, Belgique, dans le cadre de sa politique de Coopération au développement.

# Abréviations

**AIV** Avec Indices de Végétation

**CEO** Collect Earth Online

**GEE** Google Earth Engine

**IV** Infrastructures Vertes

**NIR** Infra Rouges Proches (Near Infra Red)

**RDC** République Démocratique du Congo

**RGB** Red Green Blue

**RF** Random Forest

**RMSE** Écart Quadratique Moyen (Root Mean Square Error)

**SIV** Sans Indice de Végétation

**SVM** Support Vector Machine

**TFE** Travail de Fin d'Études

# Résumé

L'évolution dans le temps et dans l'espace des infrastructures vertes (IV) de la ville de Lubumbashi est rapide et complexe.

Chef lieu du Katanga en République Démocratique du Congo, Lubumbashi voit sa population croître fortement tout comme sa superficie urbanisée. Couplés à un exode rural conséquent, ces éléments induisent de nombreuses pressions anthropiques sur les espaces verts. Qui plus est, l'alternance de saisons sèches et humides résulte en des changements importants de land cover à l'année.

Très peu d'études se sont jusqu'à aujourd'hui intéressées au land cover urbain en Afrique sub-saharienne.

Dans le cadre de ce travail, des prédictions de land cover de Lubumbashi ont été effectuées grâce à 12 modèles utilisant des images et des prédicteurs différents. Les modèles ont été entraînés sur un jeu de données de land cover produit par photo-interprétation.

Sur base de photo-interprétation et de relevés terrain utilisés comme données observées, la précision des différents modèles testés a été calculée par l'intermédiaire des écarts quadratiques moyens (RMSE).

Les modèles utilisant comme prédicteur une harmonique de NDVI calculée sur 12 mois semblent produire de meilleurs résultats que les autres prédicteurs testés.

Cela est plus que probablement dû au dynamisme temporel et spatial accru de la zone d'étude causé principalement mais pas exclusivement par les diverses pressions anthropiques et par la saisonnalité.

Le choix de l'image parmi les 3 testées ne semble pas avoir une influence conséquente sur la précision des modèles.

L'utilisation de séries temporelles est relativement répandue dans la littérature scientifique pour cartographier des zones agricoles mais leur utilisation est très peu courante dans le cadre de zones urbaines.

# Abstract

The evolution in time and space of the green infrastructures of the city of Lubumbashi is rapid and complex.

As the capital of Katanga in the Democratic Republic of Congo, Lubumbashi's population is growing strongly, as is its urbanised area. Coupled with a significant rural exodus, these elements lead to numerous anthropogenic pressures on green spaces.

Moreover, the alternation of dry and wet seasons results in significant changes in land cover throughout the year. Until now, very few studies have focused on urban land cover in Sub-Saharan Africa. In the framework of this work, land cover predictions for Lubumbashi have been made using 12 models with different images and predictors.

The models were trained on a land cover dataset produced by photo-interpretation. Based on photo-interpretation and field surveys used as observed data, the accuracy of the different models tested was calculated using root mean square errors.

The models using a 12-month NDVI harmonic as a predictor seem to produce better results than the other predictors tested. This is most likely due to the increased temporal and spatial dynamism of the study area caused mainly but not exclusively by the various anthropogenic pressures and seasonality. The choice of the image among the 3 tested does not seem to have a significant influence on the accuracy of the models.

The use of time series is relatively widespread in the scientific literature for mapping agricultural areas, but their use is very rare for urbanized zones.

# Table des matières

<b>I</b>	<b>Introduction</b>	<b>10</b>
1	Infrastructures vertes urbaines . . . . .	10
i	Définition . . . . .	10
ii	Importance des IV urbaines . . . . .	12
iii	Un déficit d'études sur le sujet en Afrique sub-saharienne . . . . .	12
2	Population croissante . . . . .	13
3	Télédétection . . . . .	14
i	Land cover . . . . .	15
ii	Random Forest . . . . .	15
iii	Indices de végétation . . . . .	15
iv	Informations temporelles . . . . .	16
v	La télédétection dans le cadre des IV urbains . . . . .	16
4	Open access . . . . .	16
5	Objectifs & questions de recherches . . . . .	16
<b>II</b>	<b>Matériel et méthode</b>	<b>17</b>
1	Zone d'étude: Lubumbashi . . . . .	17
2	Organigramme . . . . .	19
3	Typologie de cover . . . . .	19
4	Photo-interprétation: jeu de données d'entraînement . . . . .	21
i	Carte préliminaire . . . . .	21
ii	QGIS . . . . .	22
iii	Collect Eart Online . . . . .	22
5	Mission Terrain . . . . .	23
i	Sélection des quadrats terrain . . . . .	23
ii	Protocole terrain . . . . .	25
6	Cartes de cover des IV de Lubumbashi . . . . .	28
i	Cartes de cover produites . . . . .	28
ii	Images utilisées . . . . .	28
iii	Prédicteurs utilisés . . . . .	29
iv	Données d'entraînement . . . . .	30
7	Analyse statistique statistique . . . . .	31
i	Valeurs observées: CEO . . . . .	31
ii	Valeurs observées: terrain . . . . .	32
<b>III</b>	<b>Résultats</b>	<b>33</b>
1	Comparaison entre la photo-interprétation et les inventaires terrain . . .	33
2	Valeurs observées: CEO . . . . .	34

i	RMSE absolu . . . . .	34
ii	RMSE relatif . . . . .	38
3	Valeurs observées: terrain . . . . .	41
i	RMSE absolu . . . . .	41
ii	RMSE relatif . . . . .	43
4	Cartes de cover . . . . .	45
i	Carte du cover ligneux . . . . .	45
ii	Carte du cover agriculture irriguée . . . . .	46
iii	Carte du cover agriculture sèche . . . . .	47
iv	Carte du cover herbacée sèche . . . . .	48
v	Carte du cover herbacée humide . . . . .	49
vi	Carte du cover sol nu perméable . . . . .	50
vii	Carte du cover bâti . . . . .	51
viii	Carte du cover eau végétalisée humide . . . . .	52
ix	Carte du cover eau végétalisée sèche . . . . .	53
x	Carte du cover eau . . . . .	54
<b>IV Discussion</b>		<b>55</b>
1	Clé typologique & mission terrain . . . . .	55
i	Évolution de la clé typologique . . . . .	55
ii	Comparaison entre le terrain et la photo-interprétation . . . . .	55
2	Interprétation des résultats - CEO . . . . .	56
i	Prédicteurs . . . . .	56
ii	Images satellites . . . . .	57
iii	RMSE absolu . . . . .	58
iv	RMSE relatif . . . . .	58
3	Interprétation des résultats - Terrain . . . . .	58
i	RMSE absolu . . . . .	59
ii	RMSE relatif . . . . .	59
4	Cartes de cover . . . . .	60
5	Critique du travail . . . . .	60
i	Photo-interprétation . . . . .	60
ii	Calcul de l'erreur . . . . .	60
iii	Ouverture . . . . .	60
6	Position de ce travail dans la littérature scientifique . . . . .	61
i	Séries temporelles . . . . .	61
ii	Cartes existantes . . . . .	61
<b>V Contribution personnelle</b>		<b>62</b>
<b>VI Conclusion</b>		<b>63</b>

<b>VI</b>	<b>Bibliographie</b>	<b>64</b>
<b>VII</b>	<b>Annexe</b>	<b>68</b>
1	Carte préliminaire . . . . .	68
2	Résultats CEO . . . . .	69
3	Première clé typologique . . . . .	70

## List of Figures

1	Distribution géographique de 217 études sur les services écosystémiques urbains (Haase et al., 2014) . . . . .	13
2	Évolution des pourcentages des populations urbaines et rurales entre 1950 et 2050 en RDC (United Nations, Department of Economic and Social Affairs, and Population Division, 2019) . . . . .	14
3	Localisation de la ville de Lubumbashi, Katanga, RDC. . . . .	17
4	Limites administratives fournies par la FAO pour la ville de Lubumbashi, Katanga, RDC. . . . .	18
5	Organigramme méthodologique . . . . .	19
6	Clé typologique de cover de la ville de Lubumbashi 2.0 . . . . .	20
7	Grille de 1000 points utilisés pour le placement des quadrats pour la photo-interprétation. . . . .	22
8	Exemple d'un quadrat CEO . . . . .	23
9	Grille utilisée pour le choix et la répartition des quadrats terrain . . . . .	24
10	Quadrats visités lors de la mission terrain . . . . .	25
11	Schéma d'un quadrat terrain avec les transects et les points de relevé de cover . . . . .	26
12	Schéma du protocole photo d'un quadrat terrain. . . . .	27
13	Graphique des valeurs de NDVI ajustée (rouge) et non ajustée (bleu) en fonction du temps. . . . .	30
14	Différence entre les valeurs de cover issues de la photo-interprétation (II.4.iii) et les valeurs issues des inventaires de terrain. . . . .	33
15	Erreur (%) (RMSE absolu) entre le cover prédit et observé (CEO) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées. . . . .	34
16	Erreur (RMSE relatif) entre le cover prédit et observé (CEO) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées. . . . .	38
17	Erreur (RMSE absolu) entre le cover prédit et observé (Terrain) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées. . . . .	41
18	Erreur (RMSE relatif) entre le cover prédit et observé (Terrain) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées. . . . .	43



19	Carte de prédiction du cover ligneux issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur.	45
20	Carte de prédiction du cover agriculture irriguée issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur. . . . .	46
21	Carte de prédiction du cover agriculture sèche issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur. . . . .	47
22	Carte de prédiction du cover herbacée sèche issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur. . . . .	48
23	Carte de prédiction du cover herbacée humide issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur. . . . .	49
24	Carte de prédiction du cover sol nu perméable issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur. . . . .	50
25	Carte de prédiction du cover bâti issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur. . . . .	51
26	Carte de prédiction du cover eau végétalisée humide issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur. . . . .	52
27	Carte de prédiction du cover eau végétalisée humide issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur. . . . .	53
28	Carte de prédiction du cover eau issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur. . . . .	54
29	Carte préliminaire - Isolation des zones urbanisées et des zones végétalisées de la ville de Lubumbashi. . . . .	68
30	Répartition des 49 000 points photo-interprétés dans les différents covers utilisés. . . . .	69
31	Clé typologique des IV de Lubumbashi créée avant la mission terrain. . . .	70

## List of Tables

1	Définition d'"IV": les longueurs et les formulations variées illustrent les difficultés à définir l'infrastructure verte au-delà des trois idées fondamentales de connectivité, de multifonctionnalité et de " verte " (Wright, 2011) . . .	11
2	Nombre de polygones par classe de cover . . . . .	21

3	Données de cover d'entraînement (CEO) pour les 5 premiers quadrats. . .	31
4	Erreur (%) (RMSE absolu) entre le cover prédit et observé (CEO) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées.	35
5	Moyennes des erreurs (RMSE absolu) entre le cover prédit et observé (CEO) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées. . . . .	36
6	Moyennes des erreurs (RMSE absolu) entre le cover prédit et observé (CEO) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées. . . . .	37
7	Erreur (RMSE relatif) entre le cover prédit et observé (CEO) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées.	39
8	Erreur (RMSE absolu) entre le cover prédit et observé (Terrain) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées.	42
9	Moyennes des erreurs (RMSE absolu) entre le cover prédit et observé (Terrain) pour les différentes combinaisons d'indices et de séries temporelles testées. . . . .	43
10	Moyennes des erreurs (RMSE absolu) entre le cover prédit et observé (Terrain) pour les différentes images testées. . . . .	43
11	Erreur (RMSE relatif) entre le cover prédit et observé (Terrain) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées.	44

# **I Introduction**

## **1 Infrastructures vertes urbaines**

Les IV urbaines jouent un rôle central, notamment au niveau écologique et social (Choi et al., 2021) mais leur présence a une influence beaucoup plus vaste.

### **i Définition**

Les IV urbaines peuvent être définies de plusieurs manières. De nombreuses définitions d'IV sont présentes dans la littérature scientifique. Ces dernières diffèrent de manière conséquente d'un cas d'étude à l'autre. Plusieurs définitions ont été rassemblées par Wright (Wright, 2011) en un tableau comparatif (Tableau 1).

Table 1: Définition d'IV: les longueurs et les formulations variées illustrent les difficultés à définir l'infrastructure verte au-delà des trois idées fondamentales de connectivité, de multifonctionnalité et de "verte" (Wright, 2011)

Source	Definition	
(Ahern 2007, p. 267)	“Green infrastructure is an emerging planning and design concept that is principally structured by a hybrid hydrological/drainage <i>network</i> , complementing and linking relict <i>green</i> areas with built infrastructure that provides ecological functions”.	Green infrastructure theory
(Benedict and McMahon 2002, p. 2)	“Our nation’s natural life support system – an <i>interconnected network</i> of waterways, wetlands, woodlands, wildlife habitats, and other natural areas; greenways, parks and other conservation lands; working farms, ranches and forests; and wilderness and other spaces that support native species, maintain natural ecological processes, sustain air and water resources, and contribute to the health and quality of life of America’s communities and people”.	
(DCLG 2008, p. 5, 2010, p. 25)	“‘Green infrastructure’ is a <i>network of multifunctional green</i> space, both new and existing, both rural and urban, which supports the natural and ecological processes and is integral to the health and quality of life of sustainable communities”.	Green infrastructure policy
(Natural England 2009, p. 7)	“Green Infrastructure is a strategically planned and delivered <i>network</i> comprising the broadest range of high quality <i>green</i> spaces and other environmental features”	
(Kambites and Owen 2006, p. 484)	“Green infrastructure is taken . . . to encompass <i>connected networks of multifunctional</i> , predominantly unbuilt, space that supports both ecological and social activities and processes”.	Linking theory and policy

Le tableau 1 offre un aperçu de la diversité de définitions d'IV existantes.

Benedict et McMahon (M. A. Benedict et al., 2012) les définissent comme: "Les espaces verts interconnectés (incluant les zones naturelles, les zones de conservations privées ou publiques, les terres exploitées avec une certaine valeur conservatoire, et autres espaces ouverts protégés) qui font l'objet d'une planification ou d'une gestion pour leurs ressources naturelles et pour les bénéfices associés qu'ils fournissent aux populations humaines".

Cette définition ne sera pas retenue dans le cadre de ce TFE bien qu'elle soit celle qui semble être le plus généralement admise. En effet, les IV de Lubumbashi, ne font pas toutes l'objet d'une planification et d'une gestion pour leur ressources naturelles.

La définition retenue dans le cadre de ce mémoire est la cinquième du tableau 1 (Kambites

and Owen, 2006).

## **ii Importance des IV urbaines**

Les infrastructures vertes ont une fonction centrale au sein des zones urbaines. Bien évidemment elles ont une influence sur les conséquences du changement climatique avec l'atténuation des îlots de chaleur (Sturiale et al., 2019) mais elles fournissent aussi de nombreux autres services écosystémiques (Coutts et al., 2015). Leur rôle essentiel concernant le maintien de la biodiversité a été prouvé maintes fois (Coutts et al., 2015; Choi et al., 2021; Sturiale et al., 2019; Kim et al., 2019). Néanmoins, les IV ont aussi un rôle à jouer à de nombreux autres niveaux:

- Santé humaine (Coutts et al., 2015)
- Régulation (Choi et al., 2021; Sturiale et al., 2019; Assessment, 2005)
- Culturel (Choi et al., 2021)
- Cohésion sociale (Nauman et al., 2011; Zölch et al., 2016)
- Tourisme (Nauman et al., 2011; Assessment, 2005)
- Récréation (Nauman et al., 2011; Assessment, 2005)

La planification de la gestion des IV est un élément central du développement durable des espaces urbains (Gill et al., 2007; Sturiale et al., 2019; Zölch et al., 2016). Ce travail pourrait être considéré comme la première étape vers cette planification.

## **iii Un déficit d'études sur le sujet en Afrique sub-saharienne**

Très peu d'études ont été menées sur le sujet des IV urbaines en Afrique sub-saharienne (Fig. 1). En effet, sur les 217 études recensées par Haase en 2014 ayant le thème des services écosystémiques urbains, seules 5 ont été effectuées en Afrique sub-saharienne et toutes ont pour zone d'étude la ville du Cap (Haase et al., 2014).

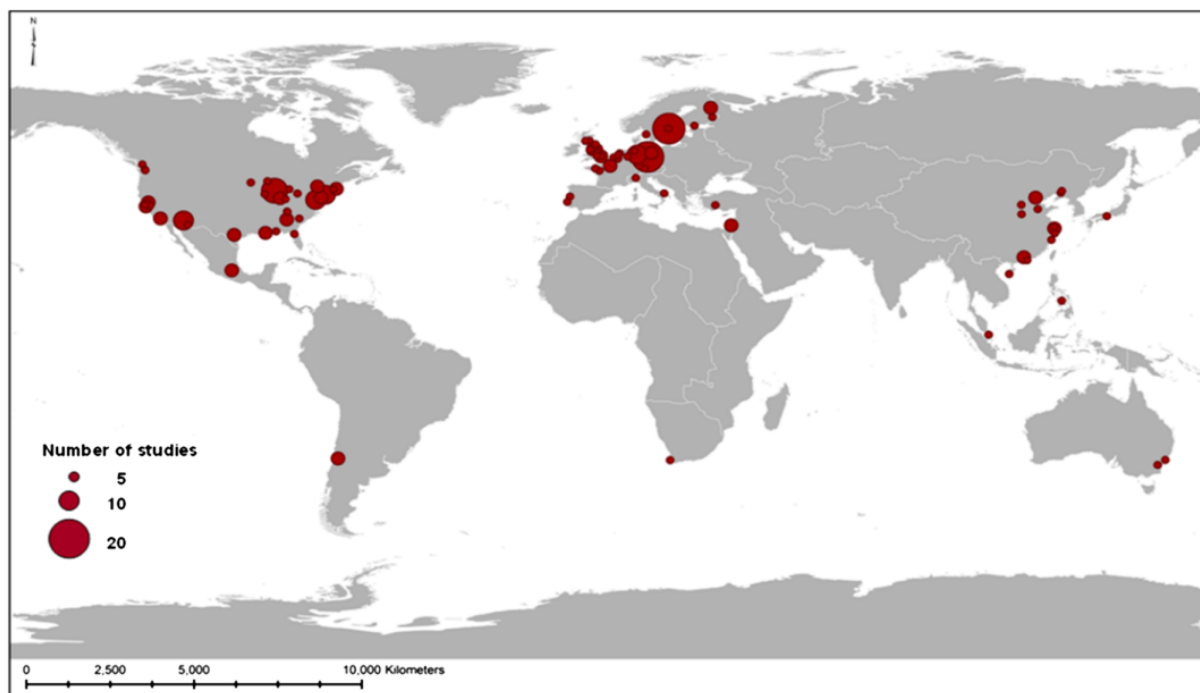


Figure 1: Distribution géographique de 217 études sur les services écosystémiques urbains (Haase et al., 2014)

En 2017, Carlos Bartsaghi Koc et ses collaborateurs publient un papier recensant les articles les plus cités de la littérature scientifique distinguant différents types d'IV. Dans les 85 articles recensés aucun n'a pour zone d'étude une ville africaine.

## 2 Population croissante

La population mondiale croît inexorablement, les modèles prédisent qu'elle atteindra 8.5 milliards en 2030 et 9.7 milliards en 2050 (United Nations, Department of Economic and Social Affairs, and Population Division, 2019).

Ce phénomène est d'autant plus prononcé en Afrique sub-saharienne où cette explosion démographique est couplée à une exode rurale massive (Teye, 2018). Ces deux facteurs induisent une expansion urbaine soutenue et parallèlement une pression anthropique conséquente sur les IV urbaines (Mensah, 2014).

Il est prédit que d'ici 2030 le taux d'augmentation de cover urbain des villes africaines sera le plus élevé au monde (Seto et al., 2012).

Actuellement, la République Démocratique du Congo (RDC) fait partie des pays ayant le cover urbain le moins étendu au monde (0.05%) mais l'Afrique sub-saharienne est considérée comme le sous-continent où la population urbaine aura la croissance la plus élevée à horizon 2050 (Angel et al., 2011). Parallèlement, l'Afrique possède la croissance urbaine la plus rapide au monde (Haase et al., 2014)

En ce qui concerne la RDC, la tendance est limpide. Le pays abrite une des populations

qui croît le plus rapidement au monde (United Nations, Department of Economic and Social Affairs, and Population Division, 2019) associée à un fort exode rural (Fig. 2).

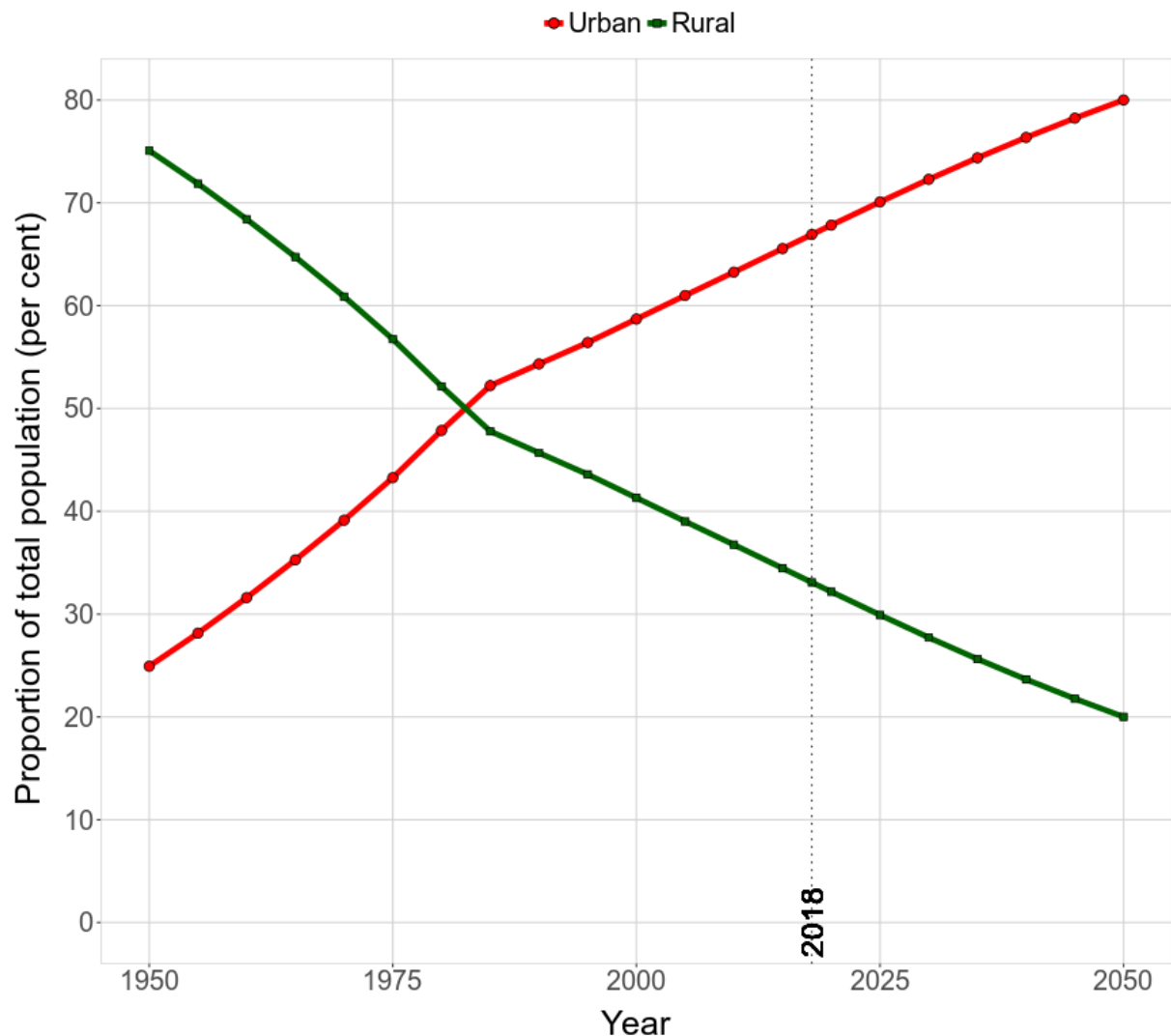


Figure 2: Évolution des pourcentages des populations urbaines et rurales entre 1950 et 2050 en RDC (United Nations, Department of Economic and Social Affairs, and Population Division, 2019)

Qui plus est, certaines sources semblent aussi indiquer que la RDC ne serait qu'au début de sa transition urbaine (Kessides, 2006).

La dernière estimation de la population de Lubumbashi date de 2020 et estime la population de la ville à 5 378 000 (INS, 2021).

### 3 Télédétection

La télédétection désigne l'acquisition de données à distance par l'intermédiaire d'un instrument habituellement transporté par un avion ou par un satellite (Fussell and Rundquist, 1986). Dans le cadre de ce TFE, l'utilisation d'images satellites permet de

travailler sur des grandes superficies et d’avoir accès à de multiples images datant de périodes différentes.

## **i Land cover**

Le land cover (couverture du sol) désigne ce qui est physiquement présent sur le sol (arbres, cultures, bâtiment, eau etc.) Cartographier le land cover en zones urbaines peut s’avérer être une tâche complexe. En effet ces régions présentent une hétérogénéité spatiale accrue ainsi qu’un dynamisme induits par la saisonnalité et par les activités anthropiques (Ghosh et al., 2014). Cependant, les informations relatives au cover et à son évolution peuvent s’avérer primordiales pour la prise de décision et pour la mise en place de plans de gestion des IV.

Le land cover joue un rôle central dans de nombreux aspects de la vie allant de la science et de l’économie jusqu’à la politique (Phan et al., 2020).

L’étude du land cover plutôt que du land use (classification) permet de produire des informations plus nuancées et potentiellement plus complètes. Typiquement, dans le cadre de ce travail, le fait de produire une carte par type de cover permet d’obtenir plusieurs informations pour chaque pixel (par exemple, 40% d’herbacée 30% de ligneux, 10% de sol nul etc.). En revanche, le travail en classification force l’algorithme à trancher et à choisir une classe d’occupation du sol pour chaque pixel.

Aussi, il est possible de convertir des cartes de cover en une carte d’occupation du sol alors que l’inverse n’est pas faisable.

## **ii Random Forest**

L’algorithme Random Forest (RF) est une variante de Bagging (Bootstrap Agregating) proposée par Breiman, 2001, qui a été utilisé pour produire toutes les cartes de ce travail. Cet algorithme de machine learning employé pour la classification (land use), la régression (land cover) et d’autres tâches, possède de multiples avantages. Il performe très bien sur de grands jeux de données, ne prend pas en compte le bruit dans les jeux de données (overfitting) (Guo et al., 2011), ou encore produit des résultats d’une qualité comparable à un Support Vector Machine (SVM) (Pal, 2005) ou à un boosting (Breiman, 2001).

## **iii Indices de végétation**

Les infrastructures vertes (IV) à Lubumbashi sont diverses et variées et occupent un espace important au sein de la ville.

Les indices de végétation, aussi divers et variés qu’ils soient, ont pour but d’aider à l’identification de couvert végétal et à son évolution dans le temps. Ils sont issus de la transformation d’images spectrales (Huete, 1988) utilisant au minimum deux bandes.



## **iv Informations temporelles**

La saisonnalité est particulièrement marquée à Lubumbashi (II.1). Cela induit des évolutions de cover significatives au cours de l'année, c'est pourquoi les informations temporelles ont le potentiel d'améliorer les prédictions de cover (Lambert et al., 2013; Adami et al., 2012). En effet, les différents covers étudiés ont le potentiel d'évoluer dans le temps de manière différente.

## **v La télédétection dans le cadre des IV urbains**

Le milieu urbain peut s'avérer être complexe à cartographier notamment à cause de la diversité de cover accrue et de leur imbrication spatiale.

De nombreuses études ont étudié le land cover urbain à travers le monde à l'aide de la télédétection. Schneider et ses collègues (Schneider, 2012) ont étudié l'évolution du land cover de 3 villes en Chine et ont testé différentes combinaisons d'algorithmes de machine learning, d'indices de végétation.

Ghosh, Sharma et Joshi ont étudié le cover de la ville de Delhi (Ghosh et al., 2014). Utilisant 9 types de cover, ils arrivent à la conclusion que l'utilisation d'images multi-temporelles est primordial pour distinguer les différents cover.

## **4 Open access**

Tous les outils et les données (logiciels, images satellites etc.) utilisés pour ce travail sont libres d'accès afin que la méthodologie appliquée soit facilement reproductible pour d'autres cas d'étude.

## **5 Objectifs & questions de recherches**

L'objectif de ce TFE est de tester l'impact du choix de l'image (saisonnalité), des indices de végétation et des séries temporelles sur la qualité de la prédiction du cover pour la ville Lubumbashi (RDC).

Les questions de recherche de ce TFE sont les suivantes:

- Le choix de l'image lié à la saisonnalité a-t-il une influence sur la précision de la prédiction de cover?
- L'utilisation d'indices de végétation comme prédicteur a-t-elle une influence sur la précision de la prédiction de cover?
- L'utilisation de séries temporelles comme prédicteur a-t-elle une influence sur la précision de la prédiction de cover?

## II Matériel et méthode

### 1 Zone d'étude: Lubumbashi

Ce travail a été effectué sur la ville de Lubumbashi, chef-lieu de la province du Katanga, située au sud de la RDC (11.66 S, 27.49 E) (Fig. 3).



Figure 3: Localisation de la ville de Lubumbashi, Katanga, RDC.

Plus précisément, la zone d'étude a été définie par les limites administratives de la ville (Fig. 4) fournies par la FAO.

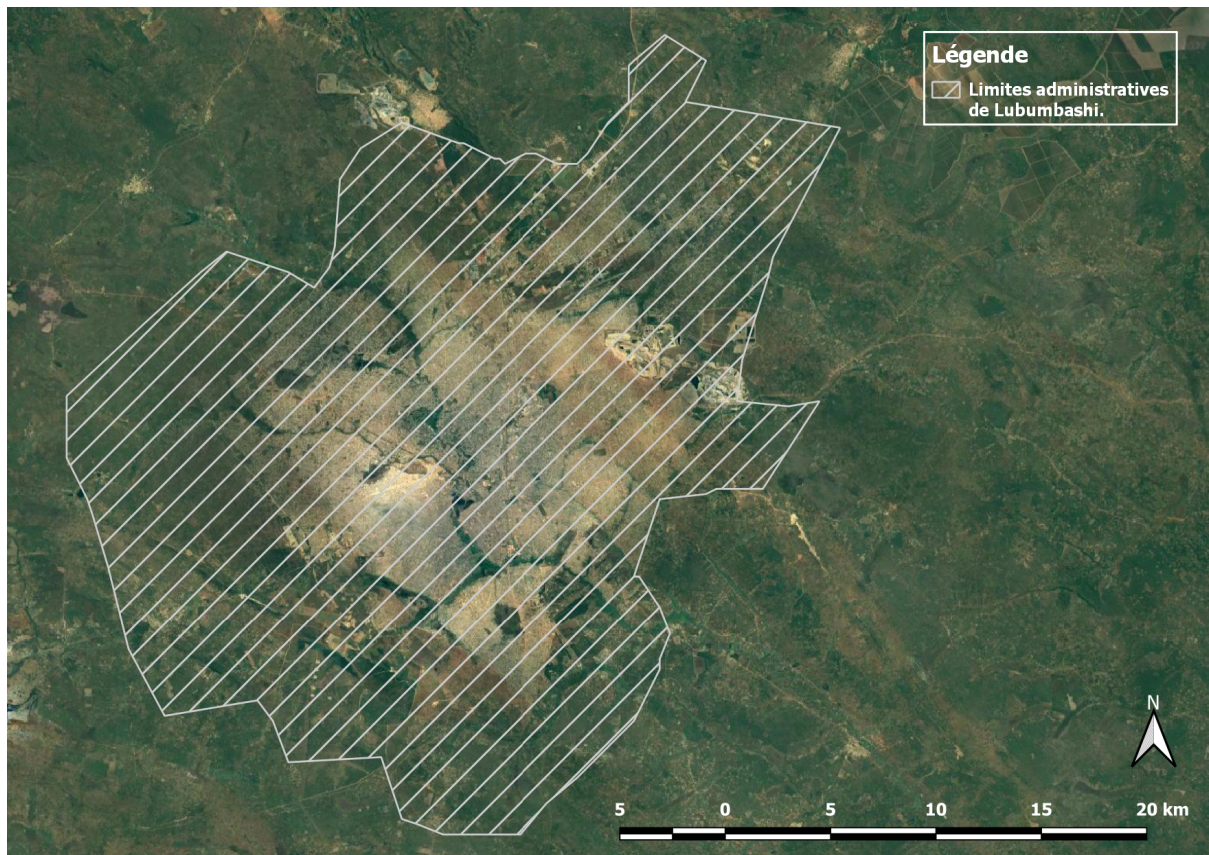


Figure 4: Limites administratives fournies par la FAO pour la ville de Lubumbashi, Katanga, RDC.

La zone d'étude a une superficie de 82 408 ha et son altitude varie entre 1200 et 1300 mètres (Sys, 1960).

Lubumbashi est la deuxième ville de la République Démocratique du Congo en termes de population (Chenge et al., 2010).

Le climat y est de type CW6 selon la classification de Köppen (Bultot, 1950; Kottek et al., 2006). Une saison sèche et une saison humide se partagent les mois de l'année, allant respectivement de mai à septembre et de novembre à mars, les mois d'octobre et d'avril étant considérés comme des mois de transition (Mulambi et al., 2013).

Le total des précipitations annuelles est de 1094 mm. La température annuelle moyenne est de 21.4°C (Useni Sikuzani, 2019).



## 2 Organigramme

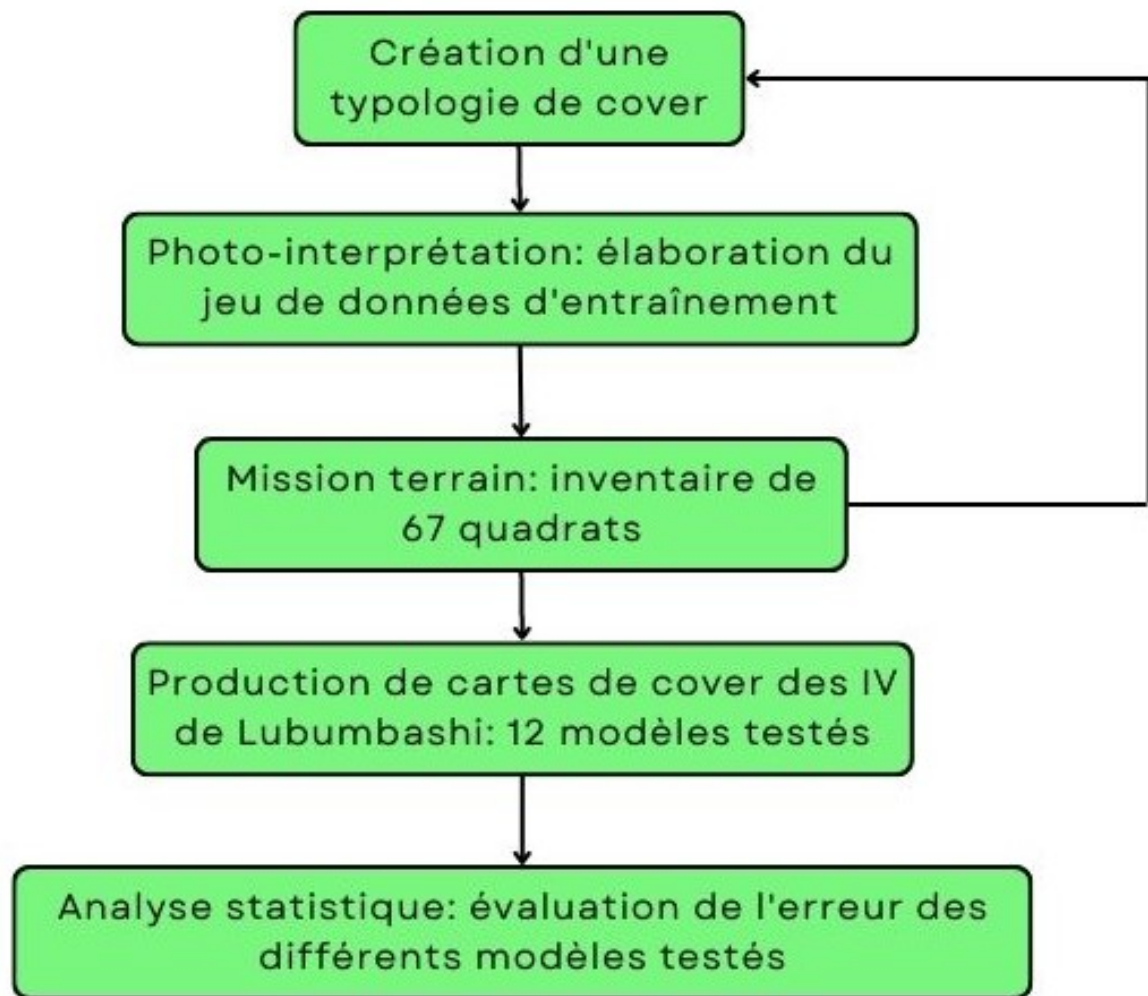


Figure 5: Organigramme méthodologique

La figure 5 représente l'organisation globale de la méthodologie de ce TFE. Les différents points sont développés ci-après.

## 3 Typologie de cover

La typologie de cover (Fig. 6) utilisée pour la photo-interprétation est le résultat d'une revue de la littérature à ce sujet (Andre, 2016; Bartesaghi-Koc et al., 2016; Munyemba et al., 2008; Mugo et al., 2020) couplée aux connaissances d'Harold Strammer des IV de la ville.

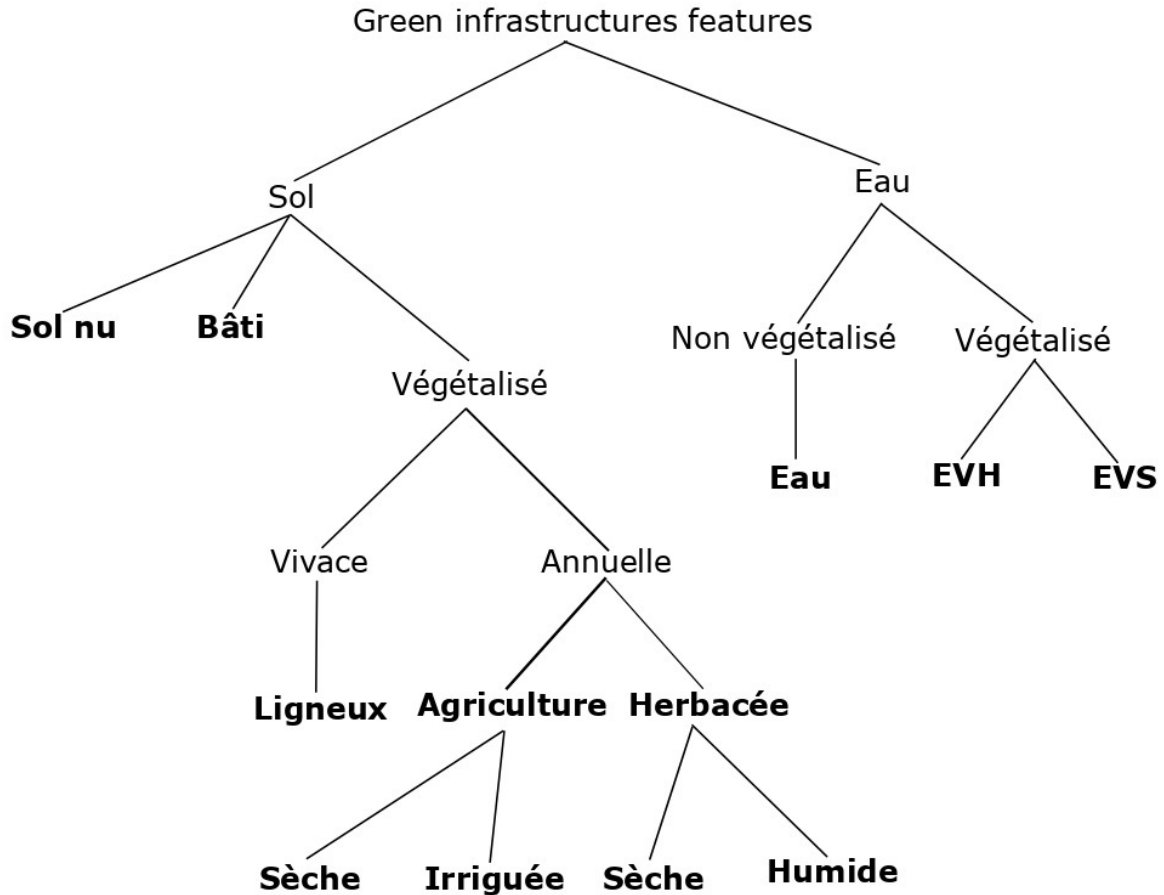


Figure 6: Clé typologique de cover de la ville de Lubumbashi 2.0

Le premier niveau de distinction de la clé (entre *Eau* et *Sol*) se fait selon que la surface est soit aquatique ou soit de la terre ferme.

À partir de *Eau*, un choix s'offre à l'opérateur en fonction de la présence de végétation ou non. Si de la végétation est présente, les deux covers, *EVH* (roseaux) et *EVS* (roseaux brûlés) sont disponibles. S'il n'y a pas de végétation, l'opérateur est face à une étendue d'*Eau*.

De l'autre côté, à partir de *Sol* un triple choix s'offre à l'opérateur. *Végétalisé* implique que de la végétation est présente sur le sol. *Bâti* correspond à la présence d'une construction anthropique (bâtiment, asphalte ou autre) sur le sol. *Sol nu* implique qu'il n'y a ni végétation, ni construction anthropique.

À partir de *Végétalisé*, un choix est à faire entre *Vivace* et *Annuelle*. *Vivace* mène directement au cover *Ligneux*. *Annuelle* mène au choix entre *Agriculture* et *Herbacée*.

*Agriculture* se divise en *Agriculture sèche* et *Agriculture irriguée*.

*Herbacée* se divise entre *Herbacée sèche* et *Herbacée humide*.

Il est à noter que cette clé typologique est une version mise à jour de la clé initialement produite. En effet, lors de la mission terrain, plusieurs incohérences de la clé initiale ont été relevées et la clé a été adaptée en conséquence.

Dix covers distincts ont donc été retenus dans le cadre de ce TFE.

## 4 Photo-interprétation: jeu de données d'entraînement

### i Carte préliminaire

Google Earth Engine (GEE) est une interface utilisateur en ligne permettant d'effectuer toute une série de traitements sur des images satellites. Une série d'algorithmes de machine learning sont aussi mis à disposition de l'utilisateur.

GEE a permis la réalisation de cette étape de ce travail qui a été effectuée dans le but de produire une carte distinguant grossièrement les zones plus urbanisées des zones plus végétalisées de la ville de Lubumbashi. Pour ce faire, un algorithme de type Random Forest a été réalisé.

Le jeu de polygones d'entraînement a été produit sur base des connaissances de la ville d'Harold Strammer. Un total de 607 polygones (Tableau 2) ont été dessinés manuellement afin d'entraîner le modèle pour cette classification préliminaire.

Table 2: Nombre de polygones par classe de cover

Covers	Forêt	Sol nu	Zone humide	Agriculture	Herbacée	Eau	Urbain
Nombre de polygones	104	68	63	79	65	27	201

Une fois la classification faite, afin que la carte produite n'affiche que deux classes (urbanisée et végétalisée), les 6 classes suivantes ont été rassemblées:

- Forêt
- Sol nu
- Zone humide
- Agriculture
- Herbacée
- Eau

Le Random Forest a été exécuté sur une image Planet datant de mai 2021.

Le raster produit (VII. 1) a ensuite été importé sur QGIS.

## ii QGIS

QGIS est un logiciel de Système d'Informations Géographiques (SIG) libre d'accès. La version 3.16 (Hannover) a été utilisée dans le cadre de ce TFE.

Après une série de prétraitements, une grille de points a été générée.

Les points situés sur les zones urbaines selon la classification effectuée en II.4.i ont été supprimés afin de maximiser le nombre de points sur les zones classées comme étant des IV et ainsi de rendre la photo-interprétation plus efficace.

Ces points ont été placés selon une grille systématique, ils sont situés à 870 mètres les uns des autres. Un total de 1000 points ont été sélectionnés sur l'ensemble de la zone d'étude (Fig. 7).

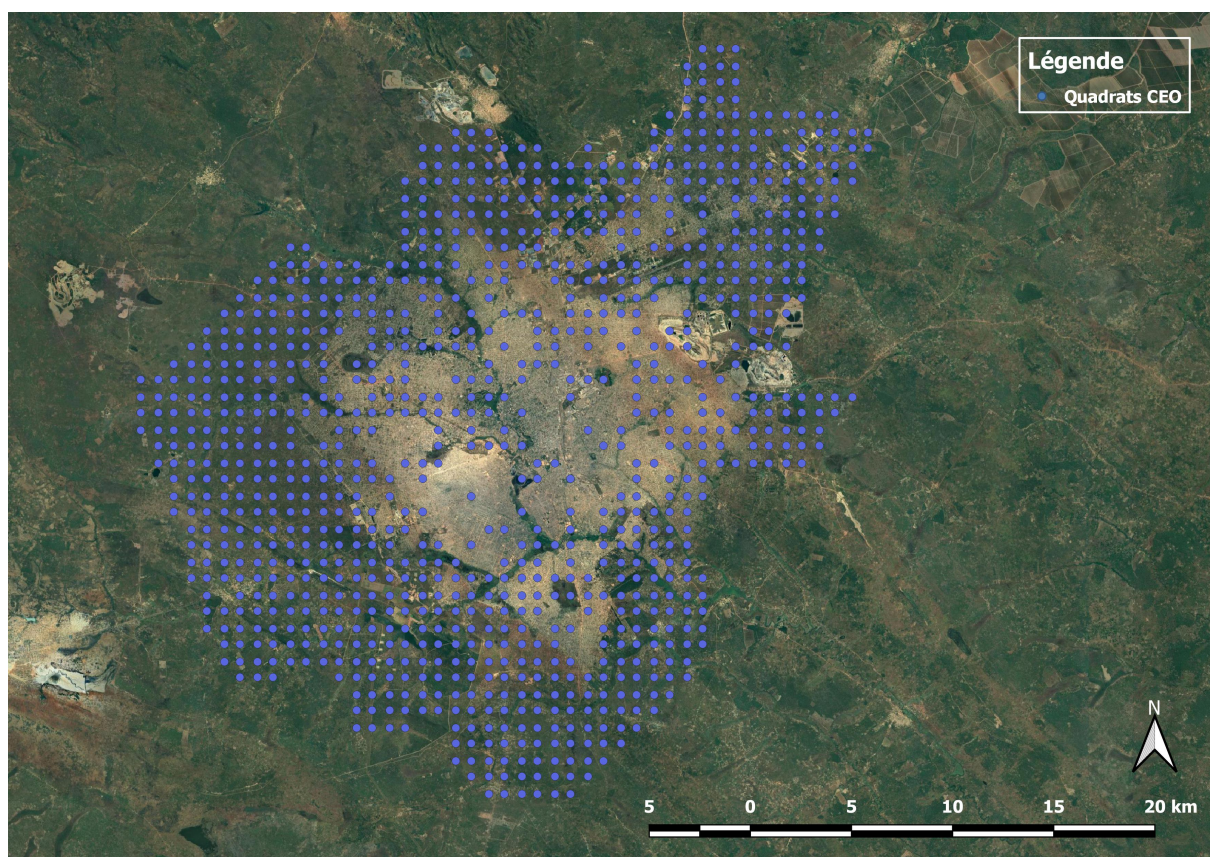


Figure 7: Grille de 1000 points utilisés pour le placement des quadrats pour la photo-interprétation.

## iii Collect Earth Online

Afin d'effectuer la photo-interprétation nécessaire à la création ultérieure de cartes, l'outil Collect Earth Online (CEO) a été utilisé (Bey et al., 2016). Cette version en ligne du logiciel Collect Earth permet de faire de la photo-interprétation facilement, et plus particulièrement, de travailler en "land cover". Cette méthode de travail s'oppose au "land use" (classification).



Avec l'importation de la grille de points 7 dans Collect Earth Online , des quadrats ont été créés et placés sur chaque point de la grille.

La photo-interprétation a donc été effectuée sur 1000 quadrats (Fig. 7) de 30 mètres de côté. Chaque quadrat est composé de 49 points espacés de 4,5 mètres (Fig. 8). Un total de 49000 points ont donc été photo-interprétés.

La grille de points (II.ii) a été créée de manière à avoir 1000 points (et donc 1000 quadrats), sachant que chaque quadrat serait composé de 49 points pour la photo-interprétation, CEO ayant une limite de maximum 50000 points à photo-interpréter.

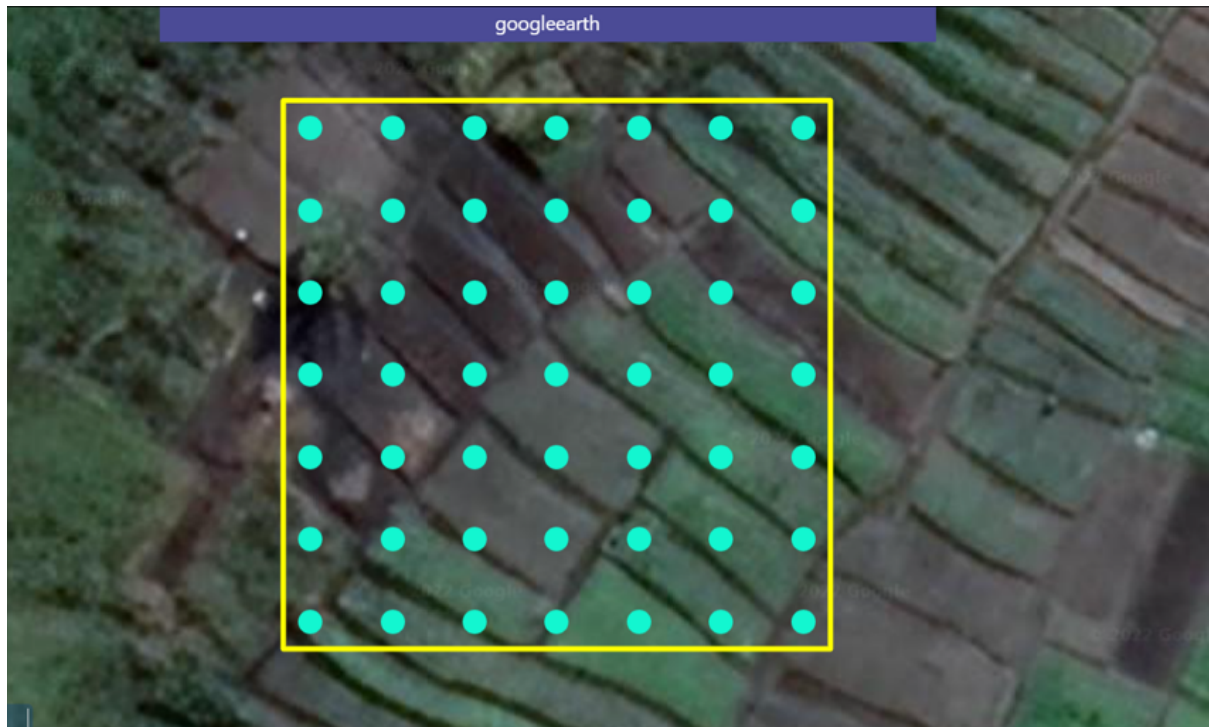


Figure 8: Exemple d'un quadrat CEO

Cette étape a été réalisée à l'aide d'images satellites Google Earth datant de mai 2021.

## 5 Mission Terrain

L'inventaire de terrain a été effectué durant les mois d'avril et de mai 2022. La photo-interprétation ayant été effectuée sur 1000 quadrats au sein de la ville et la mission terrain ayant une durée fixe de deux mois, une sélection de quadrats a dû être effectuée.

### i Sélection des quadrats terrain

Afin de répartir les quadrats visités de la manière la plus homogène possible, un échantillonnage stratifié a été effectué. Une grille (Fig. 9) a été produite sur QGIS afin d'optimiser la répartition géographique des quadrats au sein de la ville. Les carreaux de la grille font 5450 mètres de côté.



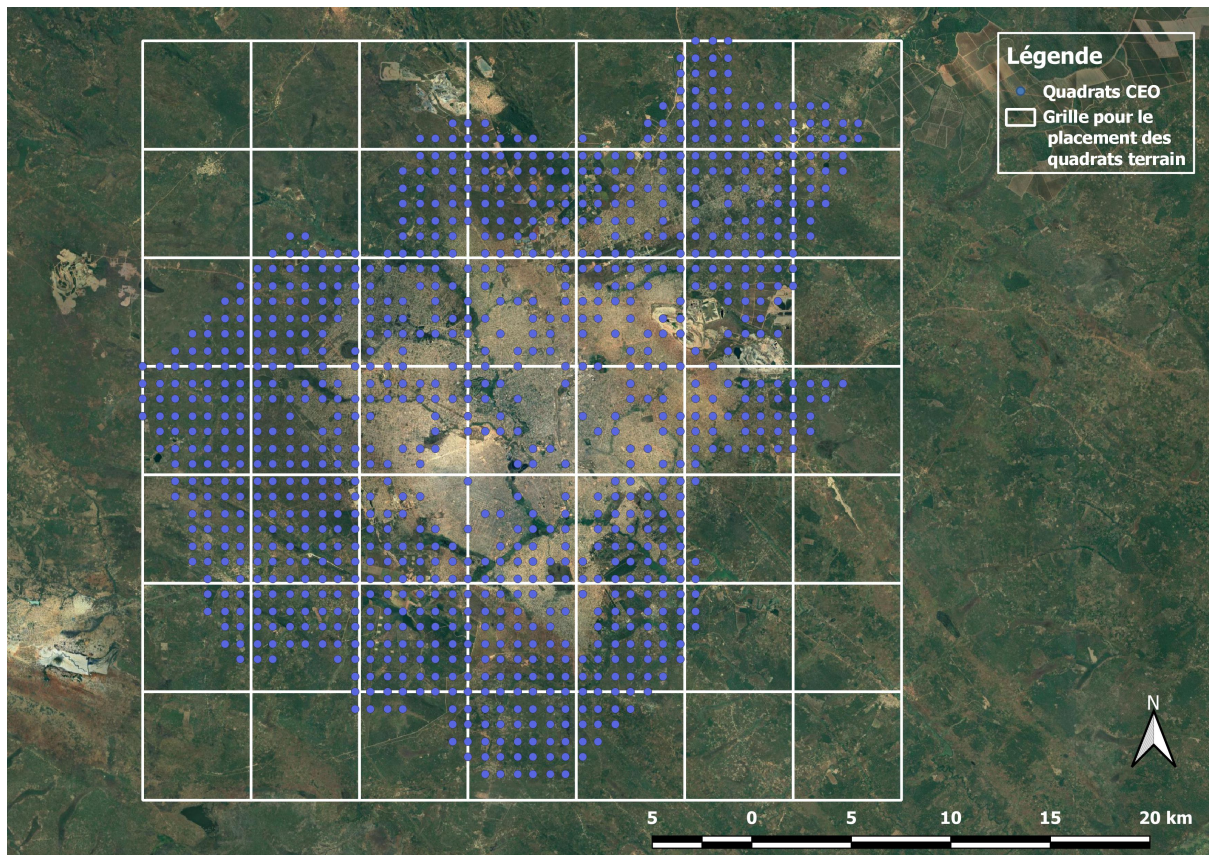


Figure 9: Grille utilisée pour le choix et la répartition des quadrats terrain

Un minimum de deux quadrats terrain pour chaque carreau de la grille a été sélectionné à l'exception des zones trop excentrées et inaccessibles.

Le choix des quadrats étudiés sur le terrain s'est avéré être assez complexe pour plusieurs raisons.

Tout d'abord, l'accessibilité des points est très variable étant donné mauvaise qualité de certaines routes (comprenez pistes).

Le premier critère de sélection des quadrats a donc été leur proximité avec une route goudronnée praticable avec une voiture standard.

Le second critère était le type de cover identifié lors de la photo-interprétation (II.4.iii). L'objectif était de se rendre sur des quadrats les plus diversifiés possibles et donc de faire les relevés sur un maximum de types de cover différents. Certains covers n'étaient pas présents au moment de la mission terrain, étant donné qu'elle a été effectuée à la fin d'une saison des pluies particulièrement humide.

Le troisième critère était tout simplement la possibilité d'accès au quadrat, par exemple un certain nombre de quadrats sont tombés dans des endroits dont l'accès est très complexe voire impossible (mine de Ruashi, enclos du zoo de Lubumbashi, propriété de l'ancien président etc.). Le temps nous étant compté, il a été décidé de ne pas introduire de demande d'autorisation d'accès à ces lieux et de se focaliser sur les endroits plus ac-



cessibles.

La mission terrain a permis de confronter la photo-interprétation avec la réalité terrain. En effet, certains quadrats visités lors des sorties terrain avaient déjà un cover différent de celui observé sur les images datant de mai 2021. Aussi, certains éléments observés lors de la photo-interprétation se sont avérés être difficilement identifiable en se basant uniquement sur les images satellites. La mission terrain a permis d'infirmer ou de confirmer les suppositions faites à leur sujet.

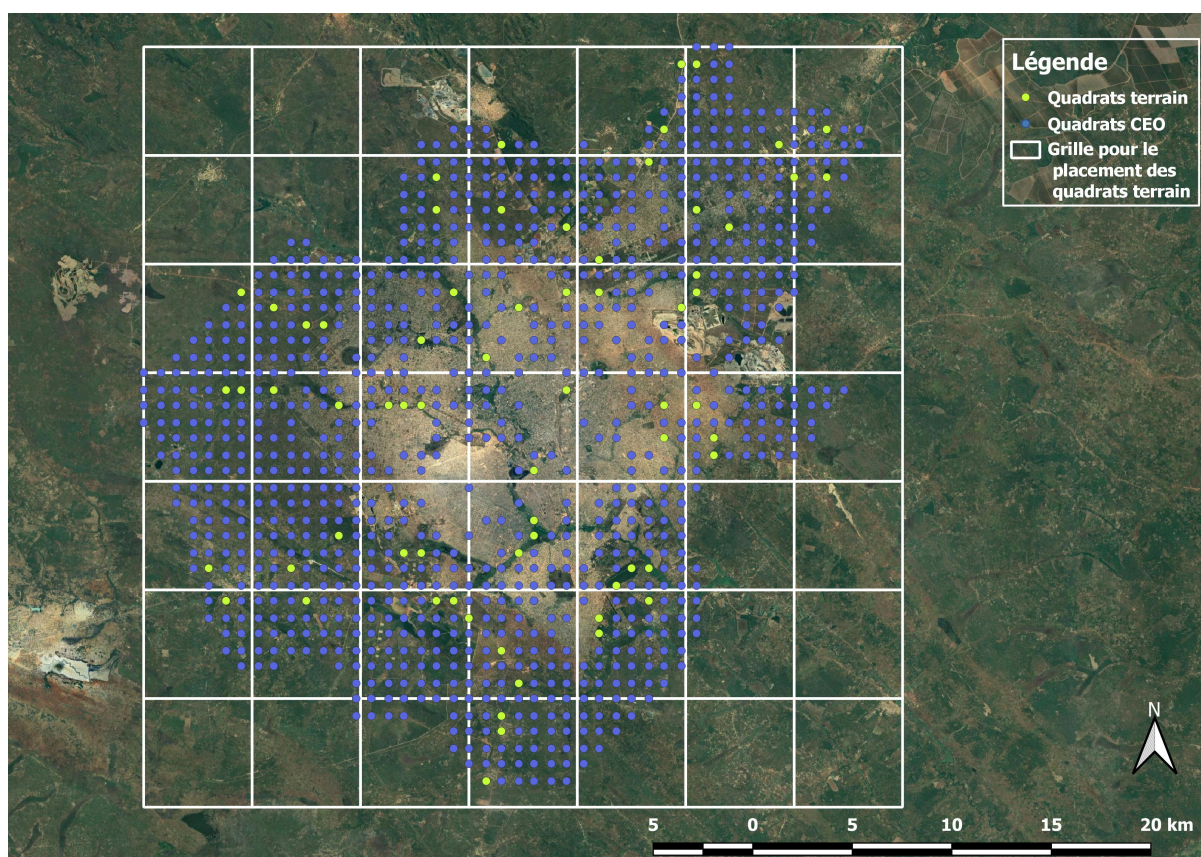


Figure 10: Quadrats visités lors de la mission terrain

La figure 10 représente en jaune les quadrats photo-interprétés et visités lors de la mission terrain et en bleu les quadrats uniquement photo-interprétés. Un total de 67 quadrats ont été échantillonnés, répartis au mieux sur la zone d'étude selon les différentes contraintes explicitées antérieurement

## ii Protocole terrain

Une fois les quadrats sélectionnés, le cover y fut estimé selon 6 transects (Fig. 11).

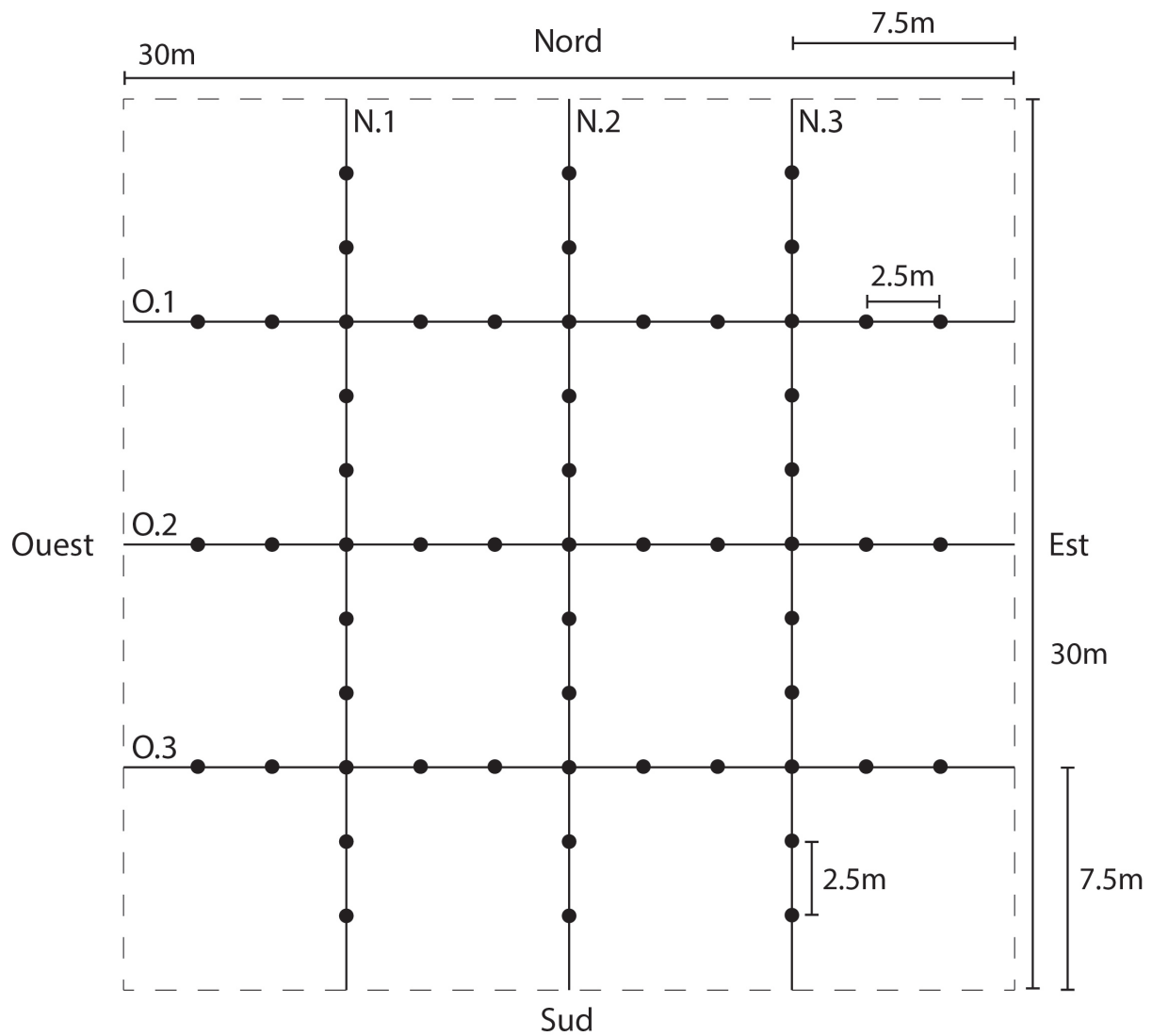


Figure 11: Schéma d'un quadrat terrain avec les transects et les points de relevé de cover

Les transects sont distants de 7.5 mètres. Le long de chaque transect, le cover est estimé sur une surface d'un mètre carré tous les 2.5 mètres. Une série de photos ont aussi été prises pour chaque quadrat visité (Fig. 12):

- 4 photos aux angles dirigées vers le centre
- 4 photos au milieu des arêtes dirigées vers l'extérieur du quadrat
- 4 photos prises du centre du quadrat vers les 4 points cardinaux
- des photos complémentaires

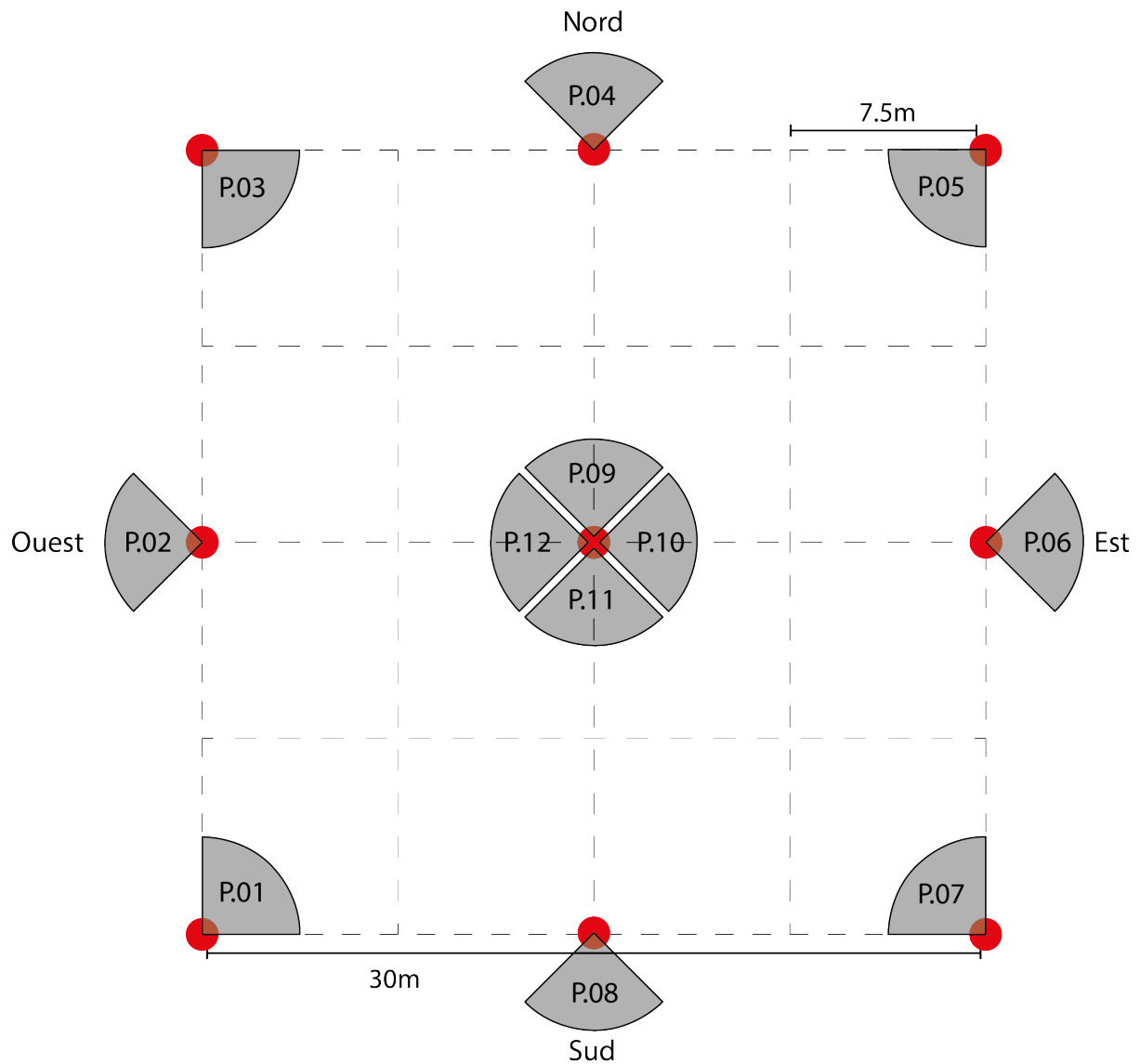


Figure 12: Schéma du protocole photo d'un quadrat terrain.

Les numéros présents sur la figure 12 correspondent à l'ordre de prise des photos.

Sur le terrain, la première étape était de repérer l'angle sud-ouest du quadrat, la prise de photo se faisant (dans la mesure du possible) dans le même sens et dans le même ordre, à partir du coin sud-ouest donc et dans le sens horaire (Fig. 12).

Le quadrat était localisé à l'aide d'une combinaison d'outils:

- Un GPS
- Google Maps
- les éléments notables du paysages (termitières, arbres remarquables, etc.)

Une fois le quadrat trouvé, il était tracé à l'aide d'un hexadécamètre et de boussoles. Lors du traçage du quadrat, les points de départ et d'arrivée des transects étaient marqués.

Les transects étaient ensuite parcourus et le relevé de cover effectué.

## 6 Cartes de cover des IV de Lubumbashi

### i Cartes de cover produites

Une multitude de cartes de cover ont été produites grâce à l'outil Google Earth Engine. Ces cartes ont été produites grâce à un algorithme de type Random Forest dont l'output était réglé en mode "Regression", s'opposant au mode "Classification". Ce type de Random Forest permet de créer une carte par type de cover avec pour chaque pixel un pourcentage de cover associé. Cette méthode permet de potentiellement obtenir plusieurs informations de cover continues pour chaque pixel contrairement à la classique classification qui ne fournit qu'une information discrète par pixel.

Cet algorithme a été utilisé sur plusieurs images issues de différentes périodes (saison sèche, saison humide) et avec différents types de prédicteur. Toutes les images utilisées pour produire les cartes de cover sont des images Planet. Ces images annoncent 3.7 mètres de résolution mais sont en réalité plus proches de 4.7 mètres. Planet met à disposition une image par mois depuis septembre 2020 de toute la zone tropicale. Quatre bandes sont disponibles, RGB et NIR.

### ii Images utilisées

Il a été décidé d'utiliser des images Planet malgré leur nombre de bandes relativement limité car leur résolution est fine. L'aspect urbain de la zone d'étude induit une hétérogénéité et une complexité spatiale conséquente justifiant la préconisation de la résolution fine sur le nombre de bandes.

Les images utilisées sont les suivantes:

- Image Planet du mois de mai 2021 (saison humide)
- Image Planet du mois de novembre 2021 (saison sèche)
- Image Planet médiane sur 6 ans (2016 à avril 2022)

Cette dernière est la médiane de 28 images Planet; deux par an de 2016 à 2020 (une en juin et une en décembre) et une par mois de septembre 2020 à mai 2022. Cela aurait dû faire un total de 29 images mais l'image de janvier 2022 n'était pas disponible.

Il aurait pu être préférable d'utiliser des images d'avril pour la saison humide mais la couverture nuageuse était trop importante pour cette image, c'est pourquoi l'image du mois de mai a été sélectionnée.

De même, l'image du mois d'octobre 2021 est de moins bonne qualité (moins contrastée), que l'image du mois de novembre, c'est pourquoi cette dernière a été choisie.

### iii Prédicteurs utilisés

Différents prédicteurs ont été utilisés dans le cadre de ce travail dont une série d'indices de végétation. D'autres indices auraient pu être pertinents mais le nombre de bandes restreint des images Planet limite les indices calculables.

**NDVI** Le NDVI (Normalized Difference Vegetation Index) (Reed et al., 1994; DeFries et al., 1995) est peut-être l'indice de végétation le plus populaire. Son efficacité a été prouvée maintes fois. (Aburas et al., 2015) Il se calcule de la manière suivante (Reed et al., 1994):

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

**MNDWI** Le MNDWI (Modified Normalized Difference Water Index) est un indice mettant en évidence les étendues d'eau. Il se calcule de la manière suivante (Xu, 2006):

$$\text{MNDWI} = \frac{\text{GREEN} - \text{NIR}}{\text{GREEN} + \text{NIR}}$$

**SAVI** Le SAVI (Soil Ajusted Vegetation Index) est un indice qui a pour but de minimiser les influences de la luminosité du sol à partir des indices spectraux de végétation impliquant les longueurs d'onde du rouge et du proche infrarouge (NIR). Il se calcule de la manière suivante (Huete, 1988):

$$\text{SAVI} = 1.5 * \left( \frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED} + 0.5} \right)$$

**Différence de NDVI** La différence de NDVI résulte de la soustraction des valeurs de NDVI (Aburas et al., 2015) pour chaque pixel à la saison humide aux valeurs de NDVI pour chaque pixel à la saison sèche. Cette manipulation permet d'obtenir une information sur l'évolution du NDVI pour chaque pixel entre la saison sèche et la saison humide.

$$\text{NDVI}_{\text{Diff}} = \text{NDVI}_{\text{saison humide}} - \text{NDVI}_{\text{saison sèche}}$$

**Harmonique ajustée de NDVI** Le concept d'harmonique ajustée est basé sur les travaux de Fourier selon lesquels on peut approximer une fonction par une somme de fonctions trigonométriques.

L'harmonique ajustée de NDVI a été calculée sur base de 12 cartes de NDVI, une par mois, d'avril 2021 à avril 2022 inclus (l'image du mois de janvier n'étant pas disponible).

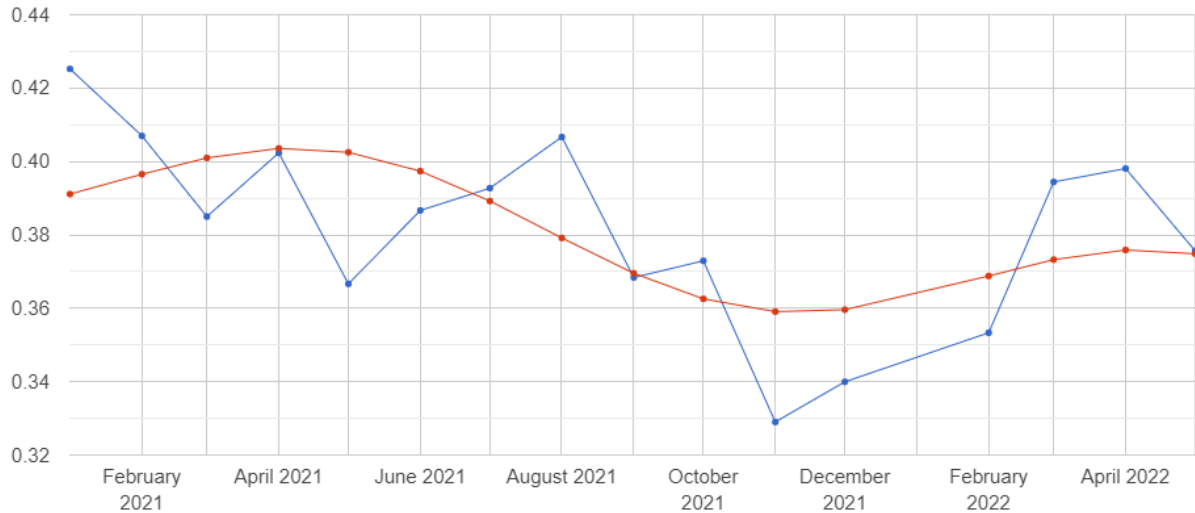


Figure 13: Graphique des valeurs de NDVI ajustée (rouge) et non ajustée (bleu) en fonction du temps.

Différents cover exhibent des motifs distincts de variation de NDVI en fonction du temps qui ont souvent de fortes caractéristiques périodiques (Jakubauskas et al., 2000; Jin et al., 2018).

Les combinaisons de prédicteurs testées sont les suivantes:

**SIV** signifie *Sans Indices de Végétation*. Pour cette prédiction, seules les bandes RGB et NIR ont été utilisées.

**AIV** signifie *Avec Indices de Végétation*. Cet acronyme caractérise les modèles pour lesquels le NDVI, le SAVI et le MNDWI ont été calculés et utilisés en plus des bandes RGB et NIR.

**NDVIDiff** signifie *Différence de NDVI*. Pour ces modèles, la différence de NDVI a été utilisée en tant que prédicteur en plus des bandes utilisées pour SIV.

**Harmo** signifie *Harmonique de NDVI*. Pour ces modèles, une courbe harmonique de NDVI ajustée a été utilisée en tant que prédicteur en plus des bandes RGB et NIR.

Au total, 12 associations entre les images et les prédicteurs ont été testées.

Les RF ont été exécutés avec 150 arbres, la précision ne semblant pas augmenter significativement au-delà de 100 arbres. (Jin et al., 2018).

#### iv Données d'entraînement

Les données d'entraînement utilisées pour entraîner le modèle sont les 1000 quadrats photo-interprétés (II.4.iii). À chaque quadrat est associée les valeurs de cover (%) calculées

sur base des 49 points photo-interprétés.

Table 3: Données de cover d’entraînement (CEO) pour les 5 premiers quadrats.

ID quadrat	Longi. centre	Lati. centre	Herb. sèche (%)	Ligneux (%)	Agri. irriguée (%)	EVH (%)	EVS (%)	Agri. sèche (%)	Eau (%)	SNP (%)	Herb. humide (%)	Bâti (%)
1	27,299	-11,640	14,290	0	0	0	0	85,710	0	0	0	0
2	27,299	-11,648	93,880	6,120	0	0	0	0	0	0	0	0
3	27,299	-11,655	69,390	28,570	0	0	0	0	0	2,040	0	0
4	27,299	-11,663	71,430	26,530	0	0	0	0	0	2,040	0	0
5	27,307	-11,640	75,510	12,240	0	0	0	0	0	12,240	0	0

Le tableau 3 présente les données issues de la photo-interprétation pour les 5 premiers quadrats. Ce sont ces valeurs de cover pour les 1000 quadrats qui ont été utilisées pour entraîner les RF effectués par la suite.

## 7 Analyse statistique statistique

### i Valeurs observées: CEO

Les cartes de cover produites ont été téléchargées et analysées sur R. Les valeurs des pixels des carte produites en II.6 situés sur les quadrats photo-interprétés en II.4.iii ont été extraites. La médiane des valeurs prédites a alors été calculée afin d’avoir des valeurs prédites homologues des valeurs observées.

Ces valeurs prédites ont été comparées aux valeurs de cover observées (II.4.iii) par l’intermédiaire de l’écart quadratique moyen (RMSE):

$$RMSE_{\text{absolu}} = \sqrt{\sum_{i=1}^N \frac{(\text{Observé-Prédit})^2}{N}}$$

Le RMSE absolu exprime l’erreur en pourcentage, a comme borne inférieure 0% (prédiction complètement juste) et comme borne supérieure 100% (prédiction complètement fausse).

En effet, l’écart quadratique moyen est directement corrélé aux valeurs de cover prédites des pixels. Étant donné que ces dernières varient significativement d’un cover à l’autre, il semble pertinent de pondérer ces valeurs de RMSE par la moyenne de chaque cover pour obtenir le RMSE relatif.



$$RMSE_{\text{relatif}} = \frac{RMSE_{\text{absolu}}}{moyenne_{\text{cover}}}$$

Le RMSE relatif exprime l'erreur sans unité, a comme borne inférieure 0 (prédiction complètement juste) et comme borne supérieure  $\infty$  (prédiction complètement fausse).

Cette mesure permet d'avoir une meilleure appréciation de la performance des modèles pour les différents cover étudiés.

Les manipulations développées ci-dessus ont été effectuées sur chaque combinaison de prédicteurs et d'images étudiées (II.6.ii & II.6.iii).

## **ii Valeurs observées: terrain**

Les mêmes manipulations ont été effectuées à partir des données récoltées sur le terrain. Les valeurs des pixels des carte produites en II.4.6 situés sur les quadrats visités lors de la mission terrain (II.5.i ont été extraites. La médiane des valeurs prédites a alors été calculée afin d'avoir des valeurs prédites homologues des valeurs observées.

Les RMSE absolus et relatifs ont été calculés à partir des valeurs prédites et des valeurs observées sur le terrain.

### III Résultats

Les résultats présentés ici sont les résultats finaux de ce TFE. Les différents résultats intermédiaires sont disponibles en annexe.

#### 1 Comparaison entre la photo-interprétation et les inventaires terrain

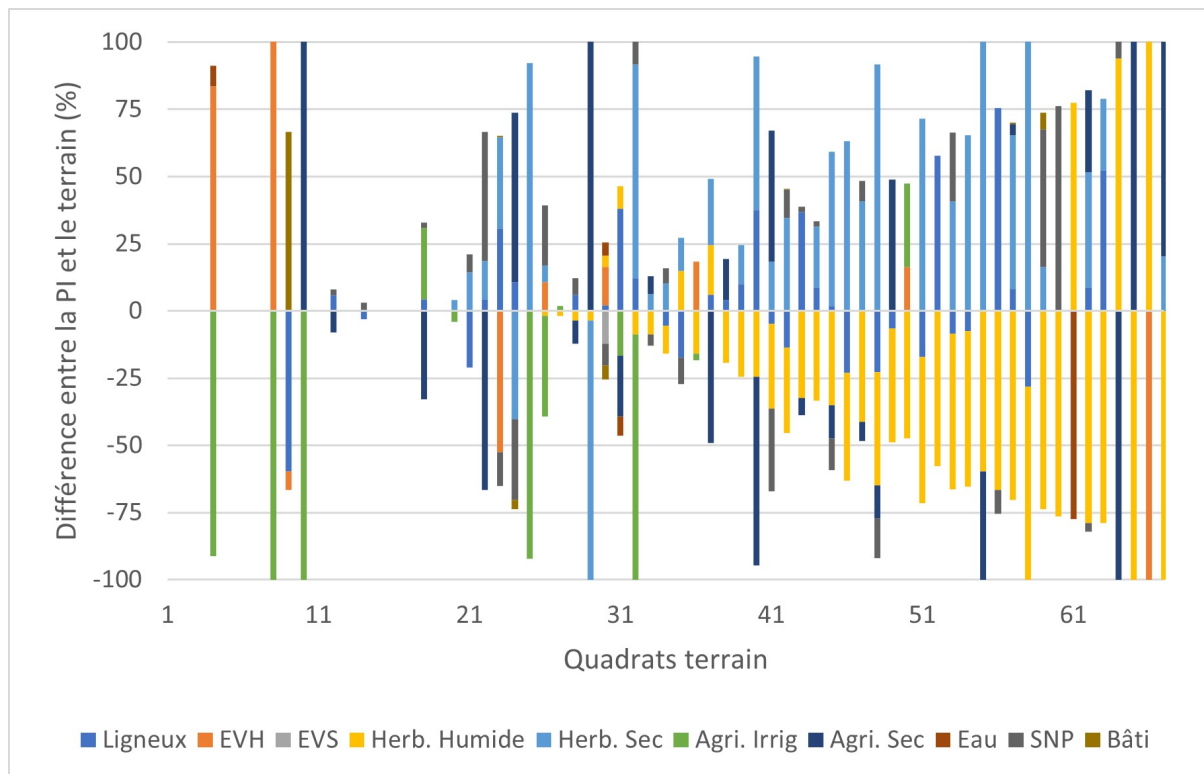


Figure 14: Différence entre les valeurs de cover issues de la photo-interprétation (II.4.iii) et les valeurs issues des inventaires de terrain.

Sur la figure 14, chaque valeur en abscisse correspond à un quadrat. Comme 67 quadrats ont été inventoriés sur le terrain, il y a 67 valeurs en abscisse.

En ordonnée, les valeurs supérieures à 0% sont celles de la photo-interprétation. Les valeurs inférieures à 0% sont celles issues des inventaires de terrain.

Il est tout d'abord à noter que pour 12 des 67 quadrats, la photo-interprétation et le terrain correspondent parfaitement (colonnes vierges).

Aussi, beaucoup de superficie considérée comme herbacée sèche lors de la photo-interprétation est en réalité de l'herbacée humide.

## 2 Valeurs observées: CEO

Les résultats suivants ont été obtenus en utilisant les données issues de la photo-interprétation sur CEO (II.4.iii) comme valeurs observées.

Le code de formes utilisé pour tous les graphiques est le suivant:

- Humide  $\mapsto$  Carré
- Sèche  $\mapsto$  Triangle
- Médiane  $\mapsto$  Rond

Le code couleurs utilisé pour tous les graphiques est le suivant:

- SIV  $\mapsto$  Orange
- AIV  $\mapsto$  Vert
- NDVIDiff  $\mapsto$  Rouge
- Harmonique NDVI  $\mapsto$  Bleu

### i RMSE absolu

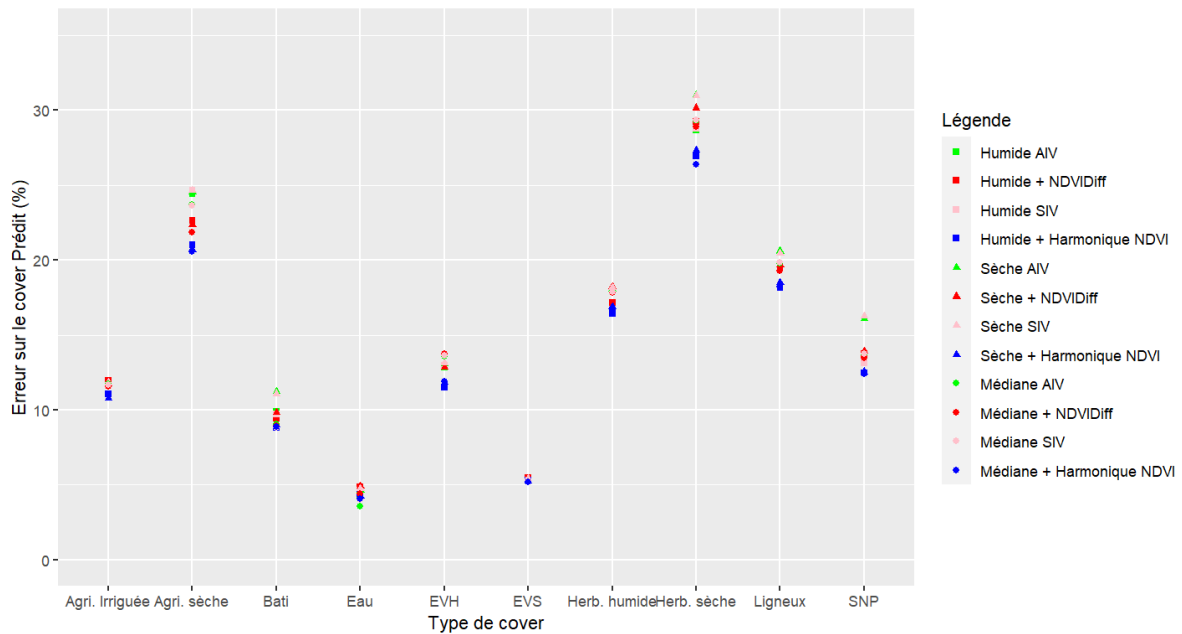


Figure 15: Erreur (%) (RMSE absolu) entre le cover prédit et observé (CEO) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées.

Table 4: Erreur (%) (RMSE absolu) entre le cover pr dit et observ  (CEO) pour les diff rentes combinaisons d'images, d'indices et de s ries temporelles test es.

	Humide				S�che				Mediane 6 ans			
	AIV	NDVIDiff	SIV	Harmo	AIV	NDVIDiff	SIV	Harmo	AIV	NDVIDiff	SIV	Harmo
Ligneux	18,21	19,42	18,31	18,15	20,59	19,66	20,46	18,52	19,81	19,30	19,81	18,23
Herb. s�che	28,64	29,24	28,78	26,94	30,99	30,15	30,96	27,31	29,31	28,92	29,35	26,35
Herb. humide	16,91	17,16	16,72	16,43	18,07	18,20	18,18	16,89	17,85	17,80	17,86	16,64
Agri. s�che	24,40	22,65	24,66	21,00	24,54	22,36	24,63	20,67	23,71	21,85	23,65	20,56
Agri. irrigu�e	11,80	11,98	11,63	11,07	11,83	11,72	11,86	10,78	11,73	11,58	11,69	11,05
EVH	12,85	12,98	12,97	11,52	13,09	12,83	13,12	11,82	13,55	13,76	13,64	11,91
EVS	5,35	5,47	5,33	5,33	5,41	5,26	5,41	5,23	5,35	5,19	5,37	5,20
Eau	4,19	4,89	4,65	4,39	4,65	4,91	4,74	4,25	3,59	4,40	4,03	4,05
SNP	13,14	13,73	13,15	12,49	16,10	13,92	16,27	12,54	13,67	13,46	13,73	12,39
B�ti	9,92	9,44	9,72	8,83	11,21	9,83	11,09	8,99	9,05	8,93	8,85	8,89

La figure 15 et le tableau 4 repr sentent l'erreur absolue effectu e sur la pr diction de cover pour les diff rentes combinaisons test es.

Le minimum de chaque ligne est écrit en bleu et le maximum en rouge.

Les différents cover prédits ayant des bornes différentes, cette partie des résultats se concentrera sur une comparaison intra cover, c'est-à-dire une comparaison des performances des différentes combinaisons d'images et de prédicteurs testées pour chaque cover.

Les valeurs de RMSE issues des prédictions utilisant une image datant de la saison sèche sont les plus élevées pour plusieurs cover:

- Bâti
- Eau
- Herbacée humide
- Herbacée sèche
- Ligneux
- SNP

À l'inverse, les valeurs de RMSE issues des prédictions utilisant l'harmonique de NDVI comme prédicteur sont les plus faibles pour plusieurs covers:

- Agriculture irriguée
- Agriculture sèche
- EVH
- Herbacée humide
- Herbacée sèche
- Ligneux
- SNP

Table 5: Moyennes des erreurs (RMSE absolu) entre le cover prédit et observé (CEO) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées.

Combinaison	AIV	NDVIDiff	SIV	Harmo
Moyenne des valeurs	14,98	14,70	15,02	13,61

La moyenne des valeurs de RMSE produite avec l'harmonique de NDVI en prédicteur toutes images confondues est plus faible que les moyennes pour les autres prédicteurs.

Table 6: Moyennes des erreurs (RMSE absolu) entre le cover prédit et observé (CEO) pour les différentes combinaisons d’images, d’indices et de séries temporelles testées.

Image	Humide	Sèche	Médiane
Moyenne des valeurs	14,36	14,98	14,40

La moyenne est légèrement plus élevée pour l’image datant de la saison sèche que pour les deux autres images testées.

## ii RMSE relatif

Les résultats suivants ont été obtenus en utilisant les données issues de la photo-interprétation sur CEO (II.4.iii) comme valeurs observées. Les différents cover ayant des bornes différentes, le RMSE relatif a été calculé afin qu'une comparaison inter-cover soit possible. Cette partie des résultats se concentrera donc sur une comparaison des performances des différents cover étudiés.

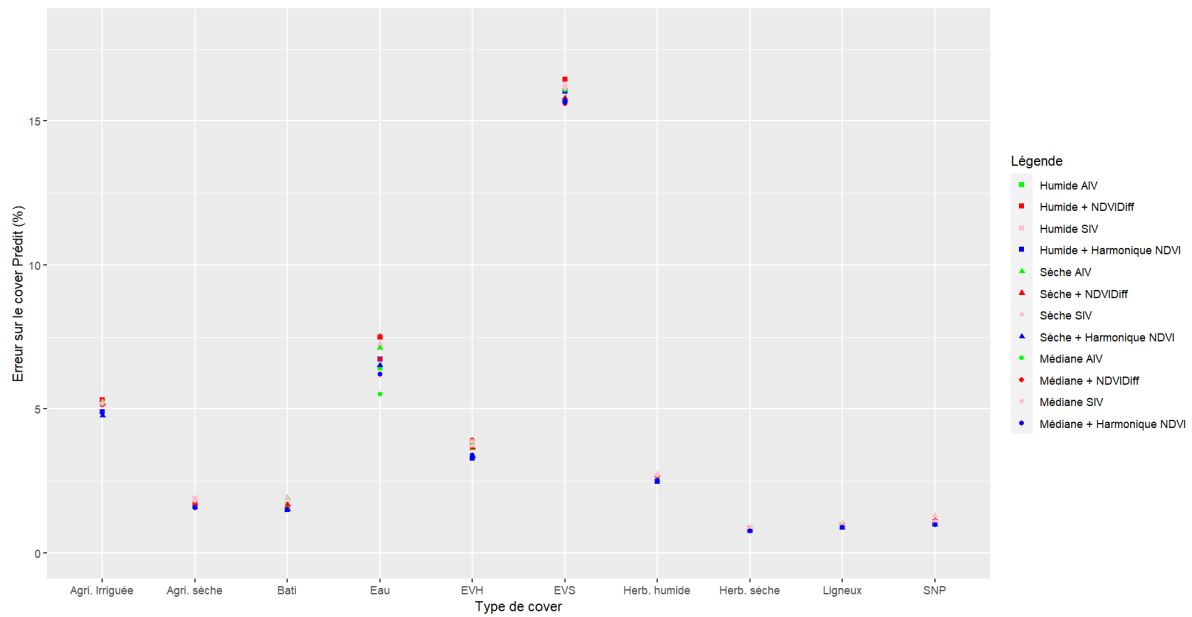


Figure 16: Erreur (RMSE relatif) entre le cover prédit et observé (CEO) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées.

Table 7: Erreur (RMSE relatif) entre le cover prédit et observé (CEO) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées.

	Humide				Sèche				Mediane 6 ans			
	AIV	NDVIDiff	SIV	Harmo	AIV	NDVIDiff	SIV	Harmo	AIV	NDVIDiff	SIV	Harmo
Ligneux	0,90	0,96	0,91	0,90	1,02	0,97	1,01	0,92	0,98	0,95	0,98	0,90
Herb. sèche	0,82	0,84	0,83	0,77	0,89	0,87	0,89	0,78	0,84	0,83	0,84	0,76
Herb. humide	2,55	2,59	2,53	2,48	2,73	2,75	2,74	2,55	2,70	2,69	2,70	2,51
Agri. sèche	1,87	1,73	1,89	1,61	1,88	1,71	1,88	1,58	1,81	1,67	1,81	1,57
Agri. Irriguée	5,23	5,31	5,15	4,91	5,24	5,19	5,26	4,77	5,20	5,13	5,18	4,90
EVH	3,66	3,69	3,69	3,28	3,73	3,65	3,73	3,36	3,86	3,92	3,88	3,39
EVS	16,08	16,45	16,01	16,03	16,26	15,80	16,26	15,72	16,07	15,60	16,14	15,64
Eau	6,42	7,49	7,12	6,73	7,13	7,52	7,27	6,51	5,50	6,74	6,17	6,21
SNP	1,05	1,09	1,05	0,99	1,28	1,11	1,29	1,00	1,09	1,07	1,09	0,99
Bâti	1,68	1,60	1,65	1,50	1,90	1,67	1,88	1,53	1,54	1,52	1,50	1,51

La figure 16 représente l'erreur relative effectuée sur la prédiction de cover pour les différentes combinaisons testées.



Il est tout d'abord remarquable qu'à l'exception du cover EVS, toutes les valeurs de RMSE relatifs de la figure 16 sont inférieurs à 10%. Les covers suivant présentent des valeurs de RMSE inférieures à 5%:

- Agriculture sèche
- Bâti
- Eau végétalisée humide
- Herbacée humide
- Herbacée sèche
- Ligneux
- Sol nu perméable

Les valeurs de RMSE pour le cover EVS sont bien supérieures aux valeurs de RMSE des autres covers.

### 3 Valeurs observées: terrain

Les résultats suivants ont été obtenus en utilisant les données récoltées sur le terrain comme données observées.

#### i RMSE absolu

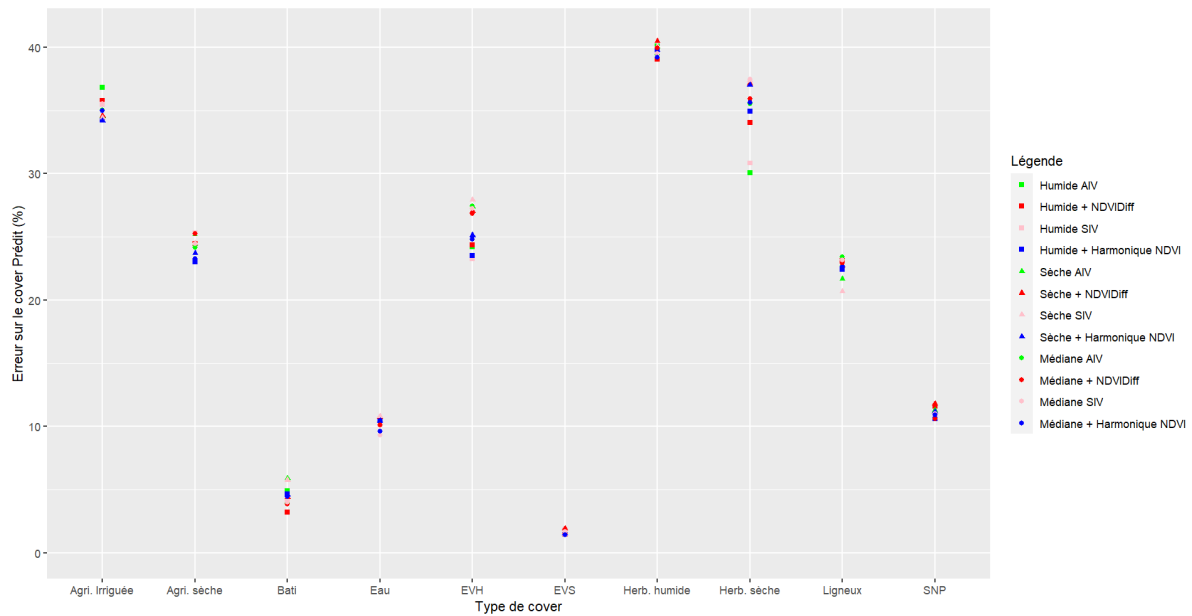


Figure 17: Erreur (RMSE absolu) entre le cover prédit et observé (Terrain) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées.

La figure 17 représente l'erreur absolue effectuée sur la prédiction de cover pour les différentes combinaisons testées.

Table 8: Erreur (RMSE absolu) entre le cover prédit et observé (Terrain) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées.

	Humide				Sèche				Médiane 6 ans			
	AIV	NDVIDiff	SIV	Harmo	AIV	NDVIDiff	SIV	Harmo	AIV	NDVIDiff	SIV	Harmo
Ligneux	22,51	23,15	22,41	22,43	21,69	22,66	20,69	22,79	23,44	22,94	23,19	22,59
Herb. sèche	30,09	34,03	30,84	34,92	37,39	37,12	37,37	37,02	35,55	35,93	37,48	35,61
Herb. humide	40,11	39,06	39,86	39,79	40,06	40,48	40,11	39,79	39,38	39,93	39,49	39,22
Agri. sèche	25,27	24,44	25,31	23,03	24,47	23,73	24,51	23,70	24,15	25,24	24,45	23,25
Agri. Irriguée	36,84	35,76	35,46	34,29	34,52	34,58	34,45	34,20	35,05	35,01	35,54	34,96
EVH	24,22	24,34	23,26	23,52	27,38	27,04	27,95	25,13	27,46	26,85	27,24	24,83
EVS	1,67	1,49	1,86	1,51	1,52	1,93	1,50	1,52	1,43	1,54	1,67	1,43
Eau	10,38	10,64	10,16	10,44	10,78	10,75	10,80	10,48	9,63	10,14	9,31	9,60
SNP	10,62	11,59	10,64	10,59	11,28	11,80	11,13	11,18	10,95	10,63	11,01	10,92
Bâti	4,91	3,22	4,20	4,64	5,88	4,40	5,77	4,61	3,90	3,86	4,02	4,48

Tout d'abord, il est à noter que les valeur de RMSE (Fig. 17) se situent dans le même ordre de grandeur que les valeur de RMSE issues de la comparaison avec le CEO (Fig.

15)

Le minimum de chaque ligne est écrit en bleu et le maximum en rouge.

Table 9: Moyennes des erreurs (RMSE absolu) entre le cover prédit et observé (Terrain) pour les différentes combinaisons d'indices et de séries temporelles testées.

Combinaison	AIV	NDVIDiff	SIV	Harmo
Moyenne des valeurs	21,08	21,14	21,06	20,75

Une fois de plus, les modèles utilisant l'harmonique de NDVI comme prédicteur sont en moyenne plus précis que les autres.

Table 10: Moyennes des erreurs (RMSE absolu) entre le cover prédit et observé (Terrain) pour les différentes images testées.

Image	Humide	Sèche	Médiane
Moyenne des valeurs	20,59	21,35	21,08

Les modèles ayant utilisé l'image datant de la saison humide affichent une valeur moyenne de RMSE plus faible que les autres modèles. La valeur moyenne de RMSE pour les modèles utilisant l'image médiane est supérieure aux valeurs pour les deux autres images testées.

## ii RMSE relatif

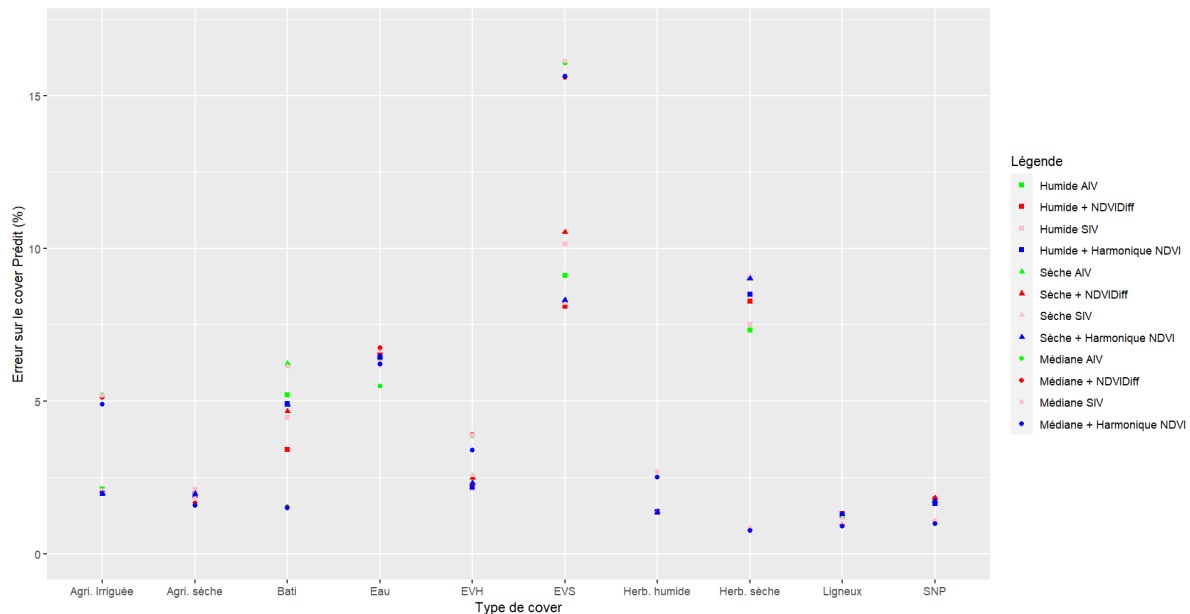


Figure 18: Erreur (RMSE relatif) entre le cover prédit et observé (Terrain) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées.

La figure 18 représente l'erreur relative effectuée sur la prédiction de cover pour les différentes combinaisons testées. Le cover EVS présente des valeurs plus élevées pour l'image médiane que pour les deux autres images testées (tableau 11).

Table 11: Erreur (RMSE relatif) entre le cover prédit et observé (Terrain) pour les différentes combinaisons d'images, d'indices et de séries temporelles testées.

	Humide				Sèche				Médiane 6 ans			
	AIV	NDVIDiff	SIV	Harmo	AIV	NDVIDiff	SIV	Harmo	AIV	NDVIDiff	SIV	Harmo
Ligneux	1,28	1,32	1,27	1,27	1,23	1,29	1,18	1,30	0,98	0,95	0,98	0,90
Herb. sèche	7,32	8,28	7,50	8,49	9,09	9,03	9,09	9,01	0,84	0,83	0,84	0,76
Herb. humide	1,39	1,35	1,38	1,38	1,38	1,40	1,39	1,38	2,70	2,69	2,70	2,51
Agri. sèche	2,10	2,03	2,11	1,92	2,04	1,97	2,04	1,97	1,81	1,67	1,81	1,57
Agri. Irriguée	2,13	2,07	2,05	1,98	1,99	2,00	1,99	1,98	5,20	5,13	5,18	4,90
EVH	2,23	2,25	2,15	2,17	2,53	2,49	2,58	2,32	3,86	3,92	3,88	3,39
EVS	9,11	8,11	10,14	8,22	8,31	10,53	8,21	8,30	16,07	15,60	16,14	15,64
Eau	6,39	6,55	6,26	6,43	6,64	6,62	6,65	6,45	5,50	6,74	6,17	6,21
SNP	1,65	1,80	1,65	1,64	1,75	1,83	1,73	1,73	1,09	1,07	1,09	0,99
Bâti	5,21	3,41	4,46	4,93	6,24	4,66	6,12	4,89	1,54	1,52	1,50	1,51

Les erreurs de prédiction pour les cover Herb. sèche, EVS, Eau et Bâti sont bien plus élevée que pour les autres cover.

## 4 Cartes de cover

Les cartes de cover suivantes ont été produites sur base de l'image de la saison humide de 2021 avec l'harmonique ajustée de NDVI en prédicteur.

### i Carte du cover ligneux

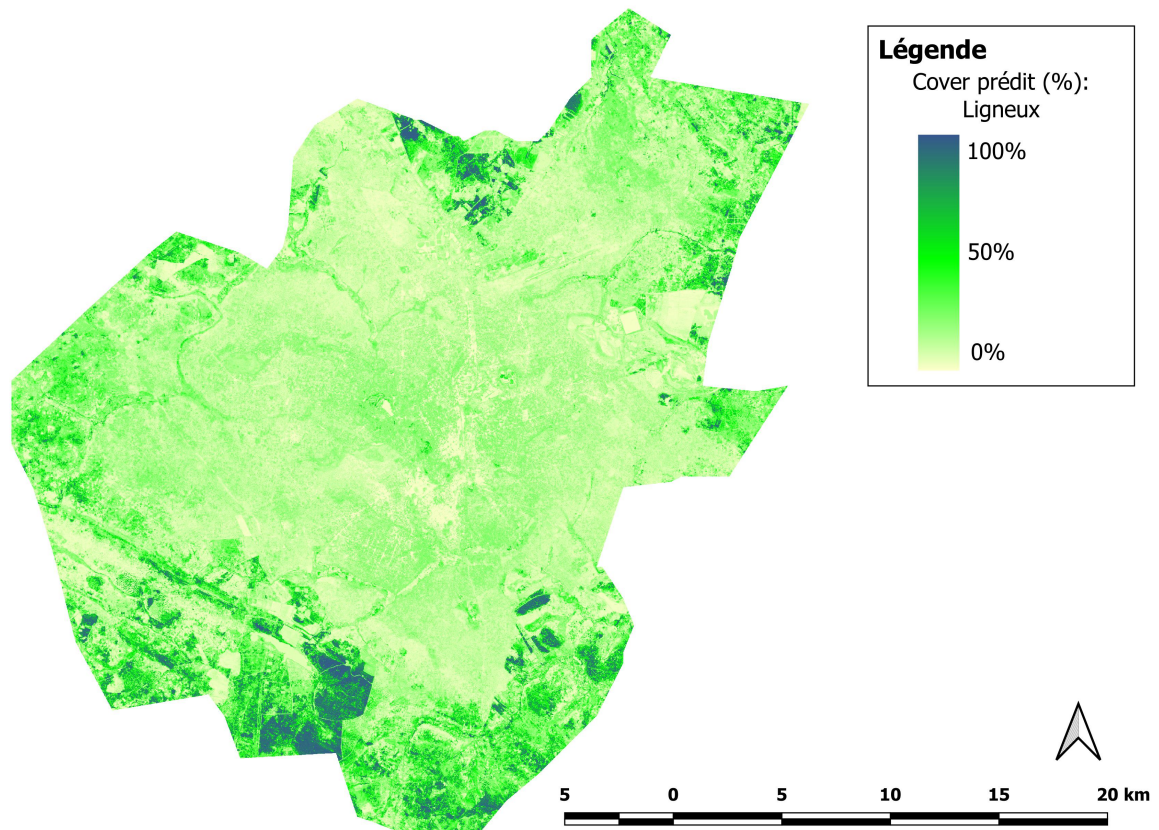


Figure 19: Carte de prédiction du cover ligneux issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur.

## ii Carte du cover agriculture irriguée

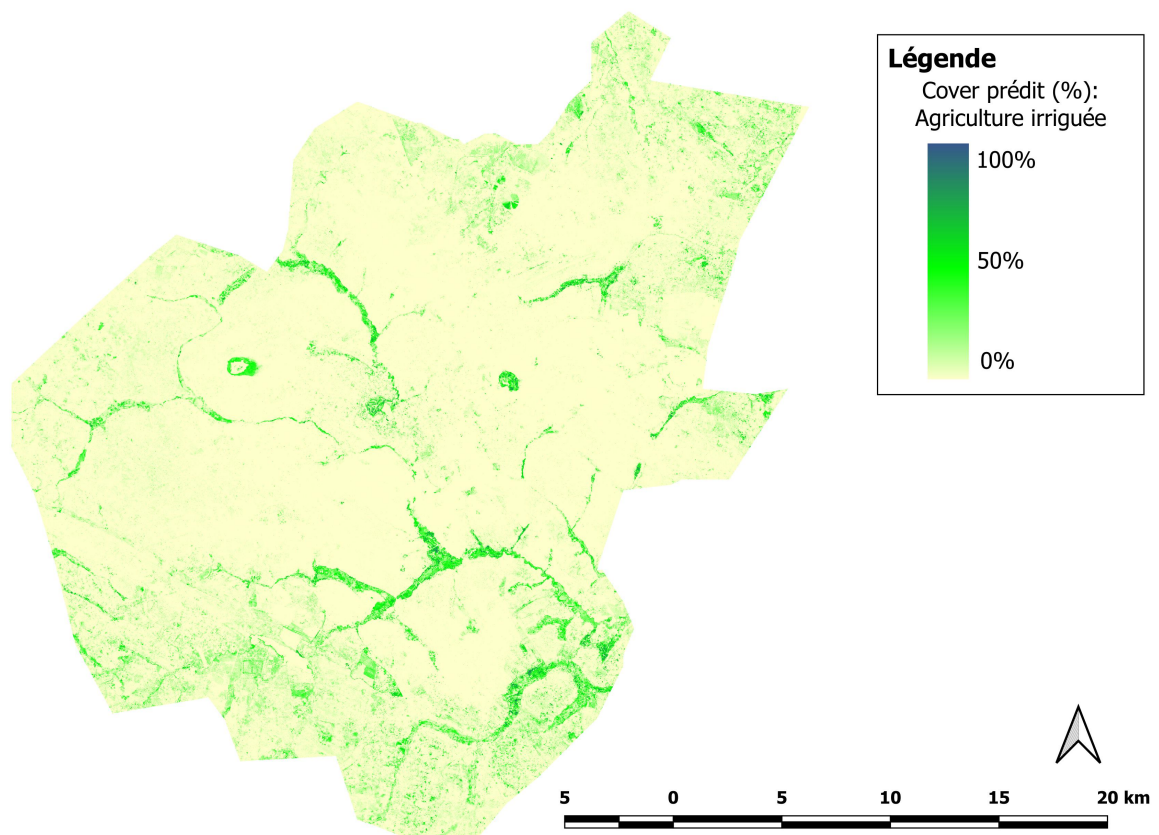


Figure 20: Carte de prédiction du cover agriculture irriguée issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur.

### iii Carte du cover agriculture sèche

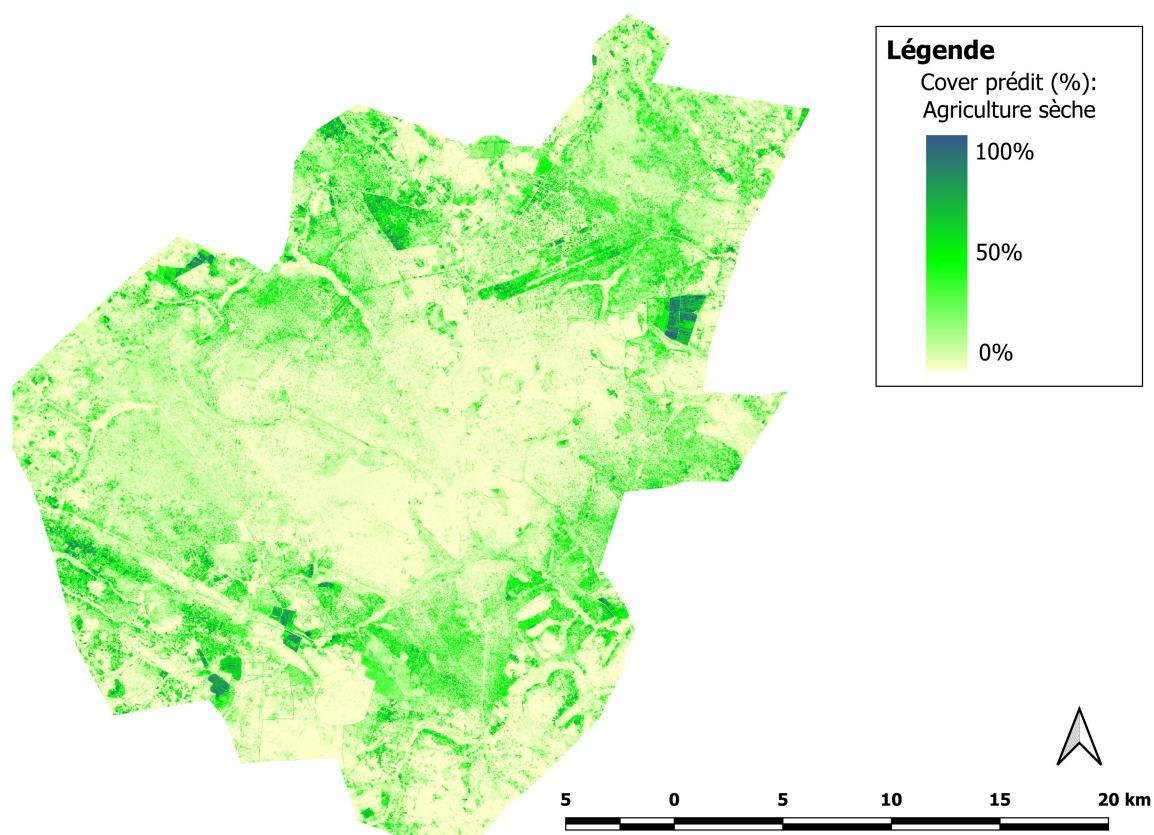


Figure 21: Carte de prédiction du cover agriculture sèche issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur.



#### iv Carte du cover herbacée sèche

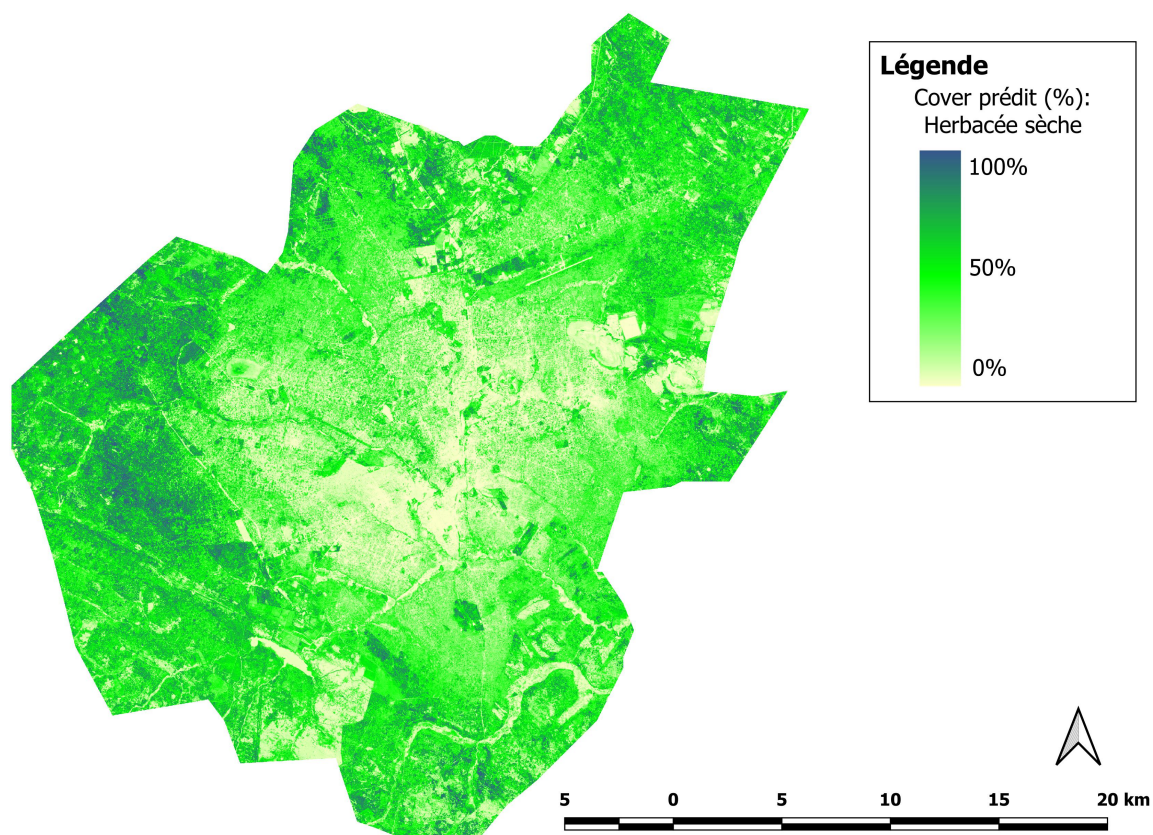


Figure 22: Carte de prédiction du cover herbacée sèche issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur.

## v Carte du cover herbacée humide

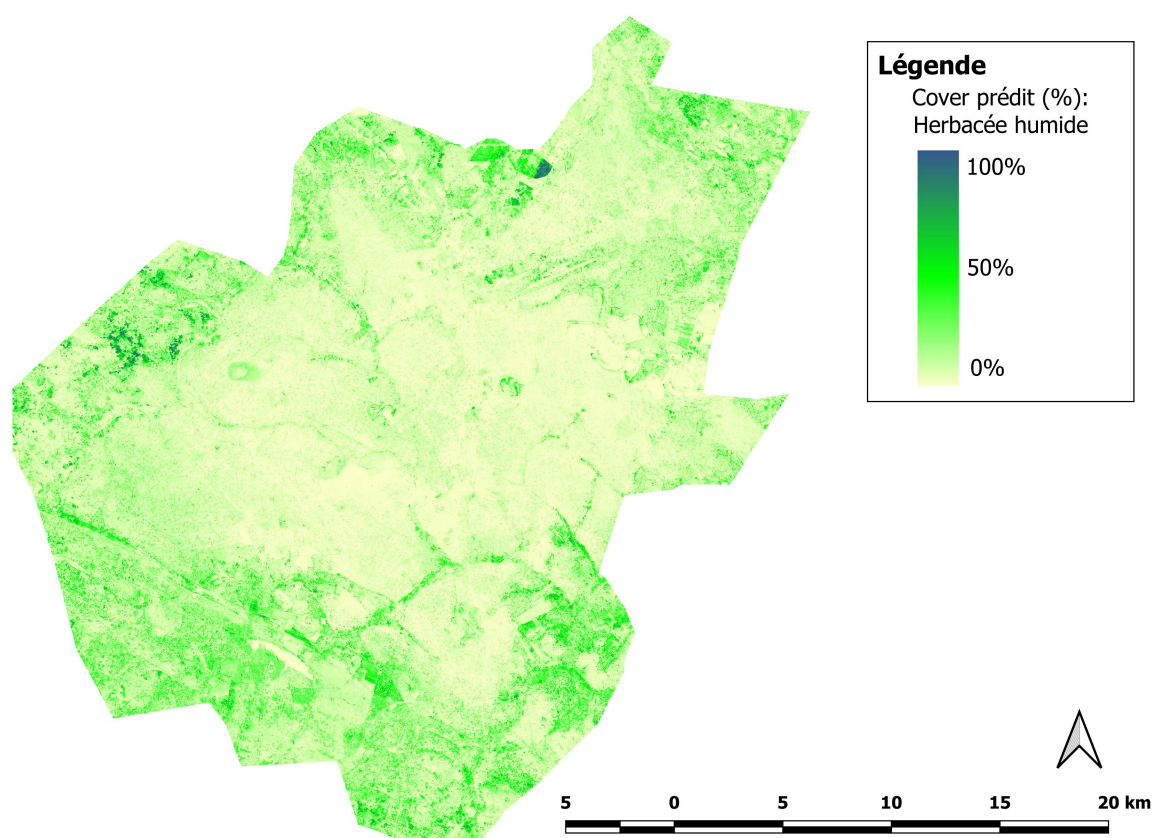


Figure 23: Carte de prédiction du cover herbacée humide issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur.

## vi Carte du cover sol nu perméable

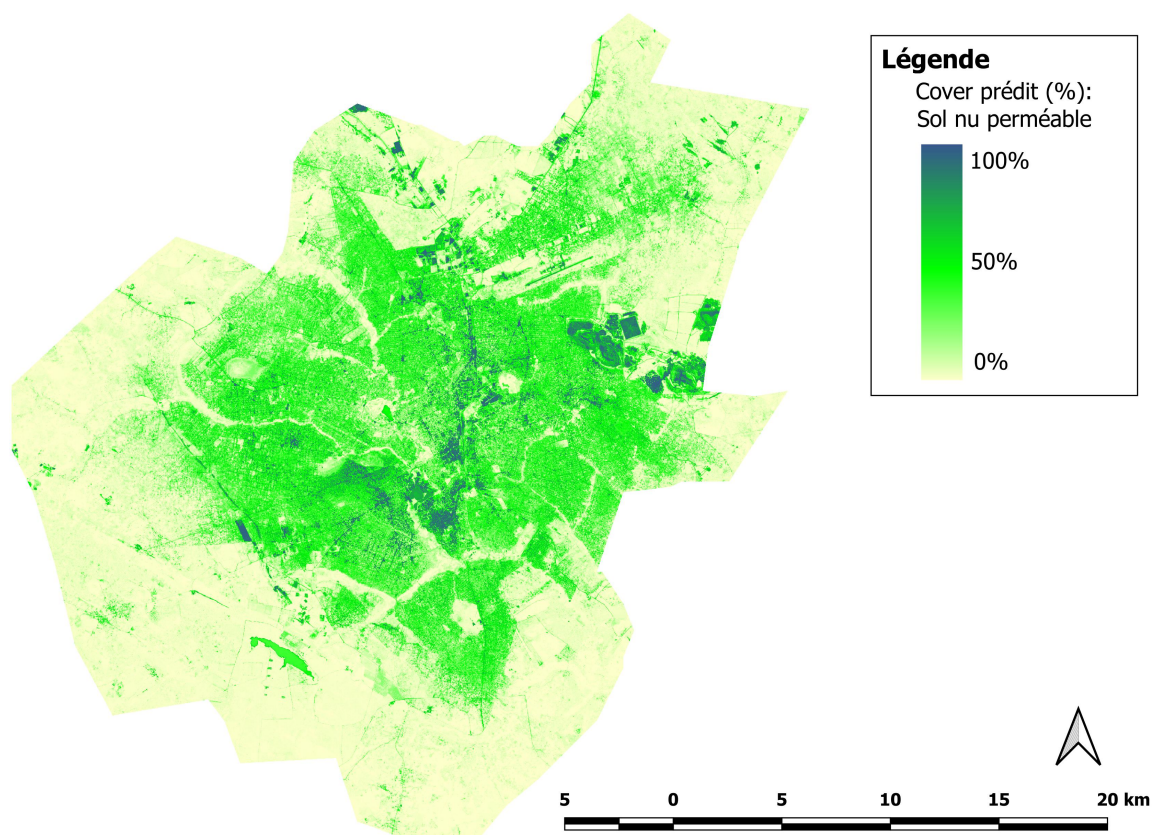


Figure 24: Carte de prédiction du cover sol nu perméable issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur.

## vii Carte du cover bâti

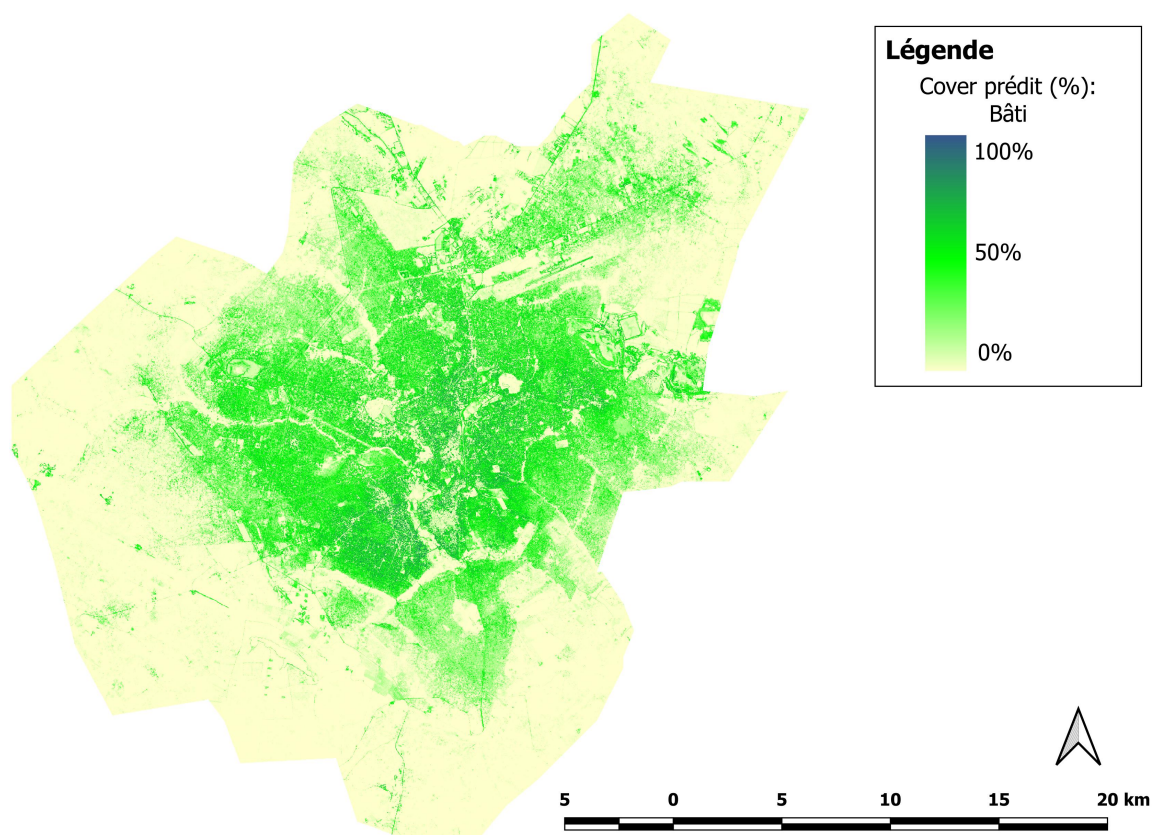


Figure 25: Carte de prédiction du cover bâti issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur.



## viii Carte du cover eau végétalisée humide

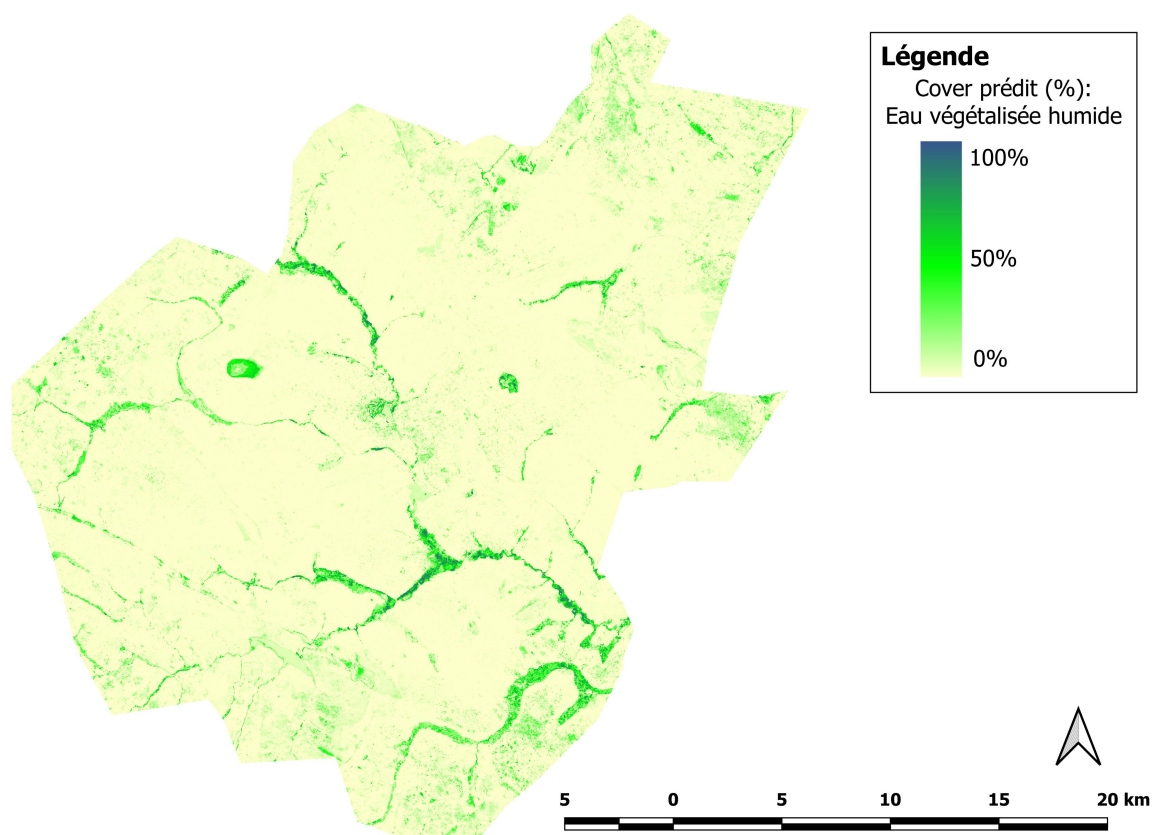


Figure 26: Carte de prédiction du cover eau végétalisée humide issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur.

## ix Carte du cover eau végétalisée sèche

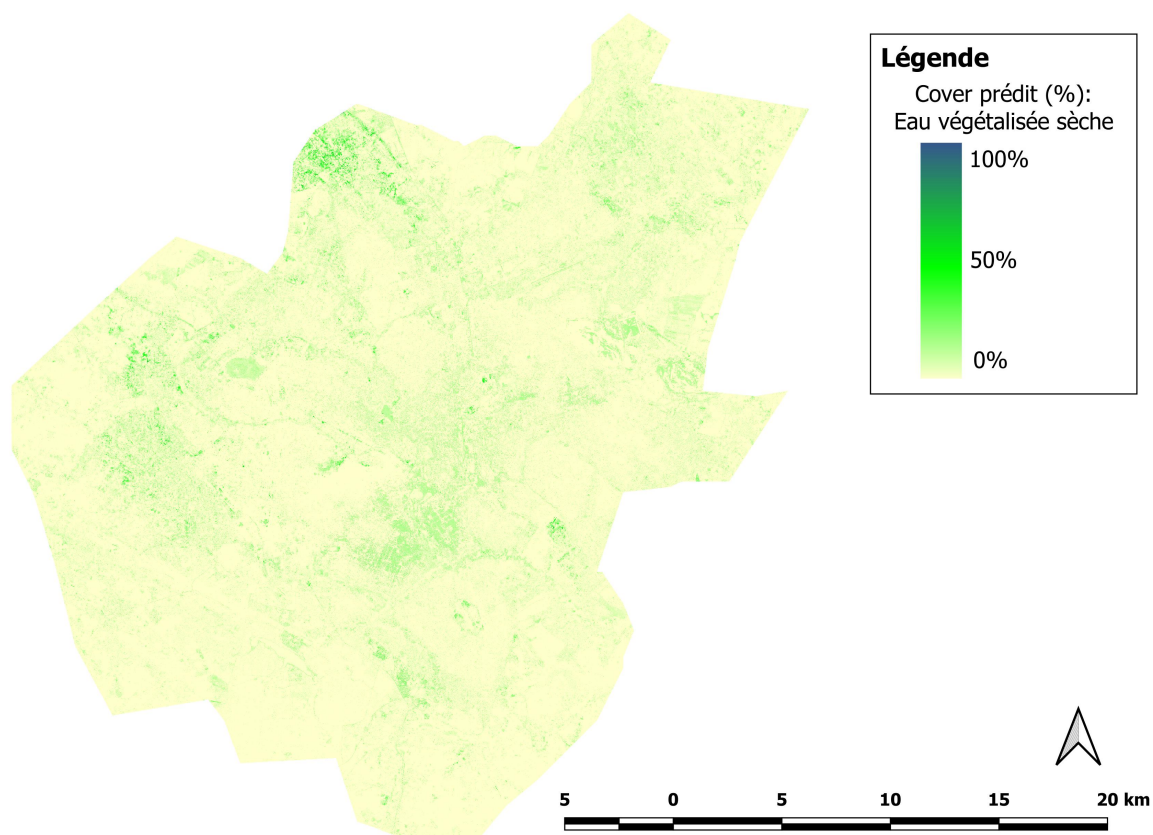


Figure 27: Carte de prédiction du cover eau végétalisée humide issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur.

## x Carte du cover eau

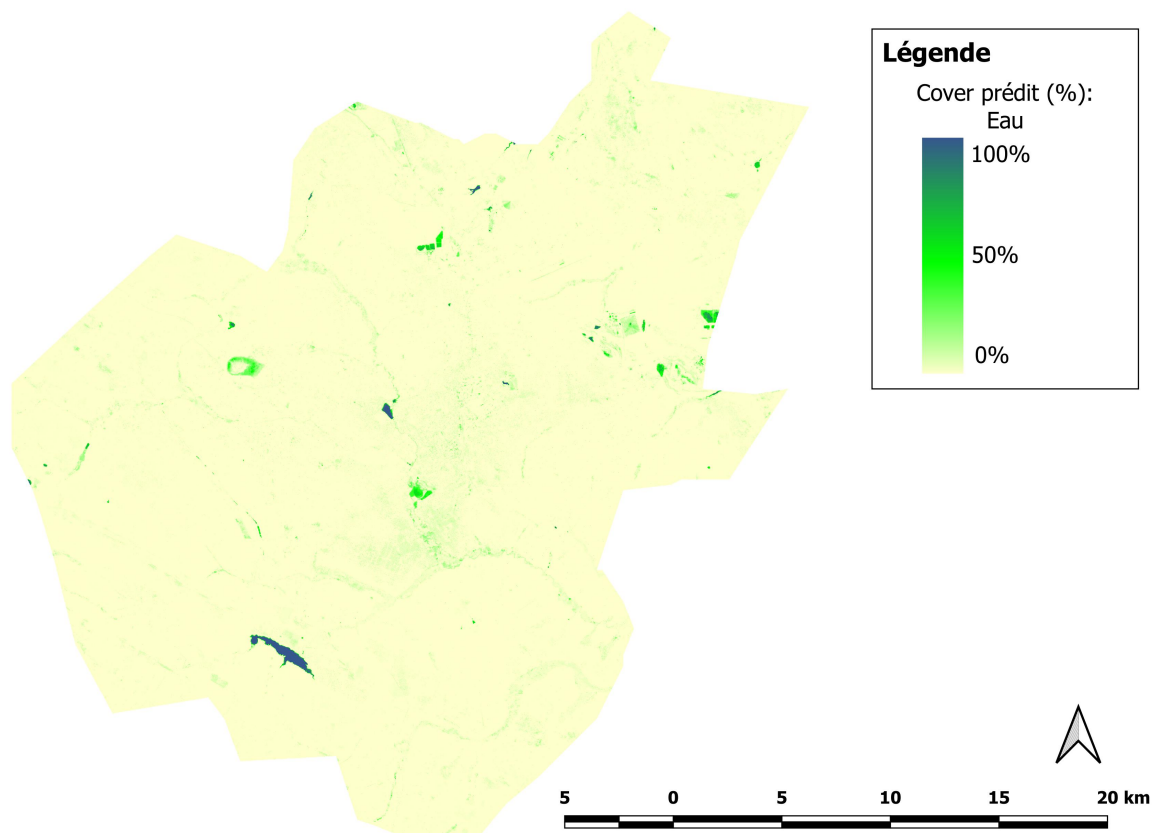


Figure 28: Carte de prédiction du cover eau issue du modèle utilisant l'image datant de la saison humide et l'harmonique de NDVI comme prédicteur.

# IV Discussion

## 1 Clé typologique & mission terrain

### i Évolution de la clé typologique

La clé typologique initialement produite (VII.3) a été modifiée suite aux premiers inventaires de terrain. De nombreux éléments de cette clé présentaient des incohérences face à la réalité du terrain. Par exemple, le critère discriminant de hauteur entre les cover ligneux et herbacée ne fonctionnait pas dans la mesure où certaines herbacées (*Tithonia diversifolia* par exemple) mesurent régulièrement 4 mètres de haut.

Aussi, certains cover pourraient être adaptés comme notamment EVS. En effet, très peu de superficie correspond à la description de ce cover. Cette dernière pourrait donc être par exemple élargie à toutes les zones brûlées.

Les inventaires de terrain nous ont aussi permis de préciser le cover de certaines zones pour lesquelles les images satellites seules n'ont pas suffi.

Peu de clé typologiques d'IV urbaines existent pour des villes africaines. Bien qu'ayant été créée spécifiquement pour Lubumbashi, la clé (Fig. 6) est vraisemblablement applicable à de nombreuses autres villes africaines. Cependant, ces dernières doivent être parcourues par un réseau hydrographique sans quoi plusieurs types de cover ne seraient plus pertinents.

### ii Comparaison entre le terrain et la photo-interprétation

La figure 14 compare les valeurs de cover pour 67 quadrats entre la photo-interprétation et le terrain.

Plusieurs raisons expliquent le fait que seulement 12 quadrats ne présentent aucune différence.

Tout d'abord, les images utilisées pour la photo-interprétation datent de mai 2021, une année complète s'est donc écoulée jusqu'au moment de l'inventaire terrain. Cela induit que pendant un an, des pressions anthropiques et naturelles ont pu modifier le cover des quadrats. Typiquement, les conversions du cover ligneux au cover herbacée humide (quadrat 56 par exemple) sont plus que probablement dues à l'Humain. Parallèlement, la conversion d'herbacée sèche en agriculture irriguée du quadrat 25 est elle aussi probablement le résultat de pression anthropique.

Ensuite, certaines différences sont dues à une erreur lors de la photo-interprétation. En effet, certains cover se confondent très facilement par exemple, le quadrat 8 photo-interprété comme du cover EVH est en réalité une rizière et par conséquent, de l'agriculture irriguée. La rizière et les roseaux sont pratiquement identiques en vue aérienne.

Finalement, certaines mauvaises interprétations ont été simplement causées par un



manque de connaissance du terrain. Le quadrat 10 a été interprété comme de l'agriculture sèche alors qu'il est en réalité couvert par de l'agriculture irriguée. Lors de la photo-interprétation, les puits présents sur le champ n'ont pas été identifiés comme tels.

La confusion entre les cover herbacée sec et herbacée humide est liée d'une part à l'année écoulée entre la photo-interprétation et la mission terrain et d'autre part à difficulté de différencier ces cover uniquement sur base d'image satellite. De plus, la saison des pluies précédant la mission terrain aurait été particulièrement humide. Par conséquent, certaines zones couvertes d'herbacées sèches en mai 2021 et interprétées correctement sont potentiellement devenues des zones couvertes d'herbacées humides en mai 2022.

## 2 Interprétation des résultats - CEO

Les divers résultats présentés semblent démontrer une meilleure performance des modèles utilisant comme prédicteur l'harmonique de NDVI.

### i Prédicteurs

Le dynamisme à la fois temporel et spatial de la zone d'étude est induit par la saisonnalité et par les diverses pressions anthropiques. Il est probablement la principale cause des meilleurs résultats produits par les modèles ayant comme prédicteur l'harmonique de NDVI.

Aucun des trois autres prédicteurs testés à savoir, AIV, SIV et NDVIDiff n'apparaît comme ayant systématiquement une influence remarquable sur la précision du modèle.

**AIV & SIV** Les modèles utilisant les indices de végétation comme prédicteur n'ont pas systématiquement une précision supérieure aux modèles SIV. Cela pourrait être causé par la complexité spatiale accrue de la zone d'étude et par le fait que ces indices de végétation n'ont d'impact que sur les zones végétalisées or la zone d'étude est en partie urbanisée. Il est aussi possible que les indices de végétation perturbent plus les modèles plus qu'autre chose. Le cover Ligneux par exemple, les valeurs de RMSE issues des modèles SIV sont majoritairement plus faibles que les valeurs de RMSE des modèles AIV.

**NDVIDiff** Malgré la dimension temporelle de ce prédicteur (différence entre la saison humide et la saison sèche) les modèles l'utilisant comme prédicteur ne semblent pas performer particulièrement bien.

Bien que ce soit une information en plus pour le modèle, il est possible que la différence de NDVI l'induisse en erreur. Cette dernière est produite à partir de deux images par an. Au vu de la précision de ces modèles, il est probable que cette information temporelle soit trop peu dense dans le temps. Cela amène à penser que le dynamisme de cover de la zone

d'étude serait si intense que deux informations par an ne suffiraient pas à le caractériser. Ce résultat confirme aussi le fait que le dynamisme temporel et spatial de la zone d'étude n'est pas exclusivement dû à la saisonnalité mais aussi aux activités anthropiques. En effet, de nombreuses pressions anthropiques n'ayant aucun lien avec la saisonnalité ont une influence sur les fluctuations de cover. Par exemple, l'extraction de minerai ou de gravier illégale. La pratique de cette activité intermittente est régie par les variations des prix du marché. L'activité complémentaire des locaux pratiquant le mining illégal lorsque les prix sont bas est bien souvent l'agriculture, il y aurait donc probablement une corrélation négative entre les prix du marché du mining et la superficie de cover agricole.

**Harmonique de NDVI** Selon les résultats obtenus, l'harmonique de NDVI semble être le meilleur prédicteur parmi ceux testés. Le fait d'utiliser douze images réparties sur treize mois pour créer l'harmonique de NDVI permet d'avoir un suivi beaucoup plus continu de l'évolution du cover.

La dimension temporelle de ce prédicteur semble faire toute la différence étant donné que les modèles utilisant AIV et NDVIDiff comme prédicteurs n'offrent pas une précision aussi élevée.

Les différents covers étudiés auraient donc chacun une évolution temporelle bien distincte comme spéculé plus haut.

Ces résultats confirment donc l'importance des séries temporelles pour la prédiction de cover de Lubumbashi.

## ii Images satellites

L'image ne semble pas avoir une forte influence sur la précision du modèle si l'on compare les valeurs de RMSE pour tous les cover confondus.

Cependant, la moyennes des RMSE des modèles utilisant l'image de la saison humide (Tableau 6) est légèrement plus faible que la moyenne des RMSE des modèles utilisant les deux autres images. Cela semble logique dans la mesure où la photo-interprétation (II.4.iii) a été effectuée sur une image satellite datant de la saison humide.

La précision des modèles basée sur la saison sèche présentent des résultats meilleurs qu'attendu. En effet, plusieurs éléments jouent en leur défaveur:

- La poussière
- Un plus faible contraste général

Pendant la saison sèche, dû aux faibles précipitations, les zones de sol nu s'étendent et créent de la poussière souflée par le vent qui masque la couleur originale de nombreux éléments du paysage. Cette dernière se dépose sur les bâtiments, sur les routes, sur les arbres etc, rendant leur distinction sur base d'images satellites plus compliquée. Malgré cela, ces modèles ne présentent pas des performances si mauvaises, en particulier pour

certain covers.

### iii RMSE absolu

**Minimums et maximums** Les minimums et maximums mis en évidence dans le tableau 4 fournissent des informations intéressantes.

Le RMSE ayant une valeur minimale pour l'agriculture irriguée est issu de la prédiction utilisant une image de la saison sèche et l'harmonique de NDVI comme prédicteur. Cela pourrait s'expliquer par le fait que pendant la saison sèche, l'agriculture irriguée ressort plus sur l'image satellite étant donné que pendant la saison sèche le sol nu s'étend.

Similairement, le sol nu est le moins bien prédit par le modèle utilisant l'image datant de la saison sèche et aucun indice de végétation (SIV). La photo-interprétation ayant été effectuée sur des images datant de la saison humide, il est logique que le RMSE soit plus élevé pour la prédiction sur une image datant de la saison sèche. Qui plus est, comme mentionné plus haut, la poussière vient tromper le modèle en se déposant sur diverses surfaces.

Les covers variant le plus selon les saisons à savoir les deux covers agricoles, les deux covers herbacés et le sol nu ont tous une valeur de RMSE minimale pour des modèles utilisant l'harmonique de NDVI comme prédicteur. Cela confirme la place centrale de l'information temporelle pour cas d'étude.

### iv RMSE relatif

La figure 16 montre une valeur de RMSE très élevée pour le cover EVS. Cela s'explique par le fait que ce cover occupe une place très réduite et que donc très peu de données sont disponibles (VII.2.) pour entraîner le modèle. Le même mécanisme se retrouve pour le cover eau.

Les covers ligneux, herbacée sèche, SNP et bâti présentent des valeurs de RMSE relatives très faibles. Cela est probablement induit par deux éléments. Tout d'abord le fait que la prédiction pour ces covers est globalement exacte. Ensuite, le fait que la moyenne pour ces covers doit être relativement élevée et comme le RMSE est divisé par la moyenne du cover prédit, le RMSE est faible.

## 3 Interprétation des résultats - Terrain

En utilisant les valeurs de terrain (prélevées en avril et en mai 2022) comme valeurs observées, les valeurs de RMSE se situent dans le même ordre de grandeur (Fig. 17 & Fig. 18) que lors de la comparaison avec les données du CEO (Fig. 15 & Fig. 16).

Ce résultat suggère que le relevé terrain et la photo-interprétation sont de qualité équivalente.

Cela étant, ils se différencient sur deux aspects majeurs:

D'abord, le terrain est un jeu de données de 67 quadrats au total, alors que la photo-

interprétation est un jeu de données de 1000 quadrats.

Ensuite, le relevé terrain a été effectué en avril et en mai 2022 alors que les images utilisées pour faire tourner les modèles datent de 2021. En effet, le cover de plusieurs quadrats visités pendant la mission terrain avait déjà évolué entre les images de 2021 et la mission terrain (III.1).

## **i RMSE absolu**

Certains covers présentent des valeurs nettement plus élevées de RMSE entre la figure 15 et la figure 17, notamment, herbacée humide, agriculture irriguée et eau végétalisée humide. Cela est probablement dû au fait que la mission terrain a été effectuée après une saison des pluies très intense donc que les covers humides sont plus présents que ce qui était attendu.

Il est aussi important de rappeler que le terrain nous a permis de bien mieux comprendre la nature et l'agencement des différents covers étudiés. Par exemple, l'espèce exotique *Tithonia diversifolia* était considérée comme faisant partie du cover Ligneux en raison de sa hauteur souvent conséquente (4 mètres). La mission terrain nous a permis de réaliser que cette espèce est en fait une espèce annuelle. Cette espèce est donc passée du cover ligneux au cover herbacée humide.

**Minimums et maximums** La valeur de RMSE minimale pour le cover ligneux est celle du modèle utilisant l'image de la saison sèche et aucun indice de végétation. Cela pourrait s'expliquer par le contraste accentué de la saison sèche pour le cover ligneux comme expliqué plus haut.

Les cover agricoles sont mieux prédit par les modèles utilisant l'harmonique de NDVI probablement à cause de leur forte périodicité.

Assez étonnement, les valeurs de RMSE issues des prédictions utilisant l'image de la saison sèche sont plus élevées que celles utilisant l'image de la saison sèche contrairement à ce qui est observable dans le tableau 4.

## **ii RMSE relatif**

La figure 18 et le tableau 11 montrent des valeurs de RMSE issues de l'image médiane différentes de celles issues des deux autres images pour presque tous les covers. Cela est dû à des valeurs de cover prédites moyennes différentes pour l'image médiane, le RMSE absolu étant divisé par le cover prédit moyen pour obtenir le RMSE relatif. Les valeurs de cover prédites par les modèles utilisant l'image médiane sont donc très différentes de celles prédites par les modèles utilisant les autres images dans la mesure où ces différences de valeurs de RMSE ne sont pas présentes dans le tableau 8.

## 4 Cartes de cover

Les différentes cartes de cover présentes dans la section III.4 (Fig. 19 à Fig. 28) fournissent de nombreuses informations intéressantes.

Les différents covers semblent globalement correspondre à la réalité de terrain, en tout cas aucun ne semble être complètement aberrant. Par exemple, il est limpide que les cover prédits EVH (Fig. 26) et agriculture irriguée (Fig. 20) suivent à la trace les zones aquatiques présentent sur la zone d'étude.

Cependant, la carte de cover eau (Fig. 28) ne présente pas du tout le même motif que les cartes de cover EVH (Fig. 26) et agriculture irriguée (Fig. 20). Cela peut s'expliquer de part le fait que les ces zones humides sont fortement végétalisées et donc que l'eau n'est que peu visible sur les images satellites.

## 5 Critique du travail

### i Photo-interprétation

Une critique qui peut être émise sur de nombreux papiers de télédétection utilisant la photo-interprétation pour entraîner les modèles est que la photo-interprétation comprend une part de subjectivité.

Cela est d'autant plus problématique lorsque plusieurs individus s'attellent à cette tâche comme cela a été le cas pour ce mémoire.

### ii Calcul de l'erreur

Les RMSE ont été calculés afin de quantifier l'erreur de différents modèles testés pour les différents covers étudiés. Cependant, la quantité d'information varie drastiquement entre les différents covers, en effet certains sont bien plus présents que d'autres. À titre d'exemple, herbacée sec et ligneux sont les plus présents sur les 1000 quadrats photo-interprétés et EVS et eau sont les moins présents (VII.2). Il semble donc logique d'assumer que l'erreur estimée par le RMSE pour les covers les moins représentés est moins précise que pour les covers les plus représentés.

### iii Ouverture

Il aurait probablement été intéressant de tester une série temporelle d'images RGB, sans NDVI (Adami et al., 2012). Cela aurait potentiellement permis de quantifier l'impact de ce dernier dans la performance du modèle. En toute logique, le modèle utilisant comme prédicteur une série temporelle sans NDVI serait moins performant qu'un modèle comme celui testé ici mais ce n'est pas sûr. Par exemple, dans le cadre de ce mémoire, les modèles AIV ne sont pas systématiquement plus précis que les modèles SIV.

Bien d'autres prédicteurs auraient pu être testés comme par exemple un GLCM (texture) et un modèle numérique de terrain (Jin et al., 2018) mais aussi d'autres algorithmes que le RF comme par exemple des réseaux de neurones convolutifs.

## **6 Position de ce travail dans la littérature scientifique**

### **i Séries temporelles**

La littérature scientifique regorge d'articles utilisant des séries temporelles (de NDVI ou tout simplement de RGB) pour la cartographie de zones agricoles.

En effet, de par leur très forte saisonnalité, les cultures sont toutes indiquées pour faire bon usage d'informations temporelles.

En 2002, Jakubauskas et ses collaborateurs utilisaient déjà les harmoniques de NDVI comme prédicteur pour l'identification de différents types de culture. Les cultures de soja, de maïs et de luzerne montrent des motifs de variation saisonnières de NDVI distincts (Jakubauskas et al., 2000).

En 2012, Adami et ses collègues se servaient des séries temporelles (RGB) afin d'évaluer le land use change en production de canne à sucre au Brésil et plus précisément pour identifier ce qui était présent avant la canne à sucre.

En 2017, toujours au Brésil, Bellon et ses collègues utilisaient des séries temporelles de NDVI pour identifier différents systèmes de culture.

L'année suivante, à quelques milliers de kilomètres de là, en Chine, Jin et ses collaborateurs publient un article explicitant l'importance des séries temporelles de NDVI pour la cartographie de cover agricole.

Tous ces articles utilisent des séries temporelles comme prédicteur pour des zones très majoritairement agricoles à travers le monde.

Ce TFE se distingue de ces travaux par l'application de séries temporelles sur une zone urbaine. De part les signatures temporelles des différents covers, les résultats semblent être concluants.

### **ii Cartes existantes**

L'agence Spatiale Européenne (ESA) a récemment dévoilé sa dernière carte de cover à l'échelle mondiale. La quasi-totalité notre zone d'étude est catégorisée comme Bâti sur cette carte alors que la réalité est bien plus nuancée que ça.

Cela est probablement dû au fait que cette carte a été créée à échelle mondiale et non pas pour la ville de Lubumbashi. Effectivement une part non négligeable de la zone d'étude de ce travail de fin d'étude peut être qualifiée d'urbanisée mais cette urbanisation est spatialement complexe et très dynamique dans le temps. Qui plus est, elle est bien moins drastique que ce que nous pouvons observer dans les pays du nord. Par exemple, la

bétonnisation y est beaucoup moins importante et une plus grande place est donc laissée à des IV tel que le sol nu.

## V Contribution personnelle

La quasi totalité de ce travail a été effectuée par Gaspard Estenne, Harold Strammer ayant été d'une aide précieuse lors du premier mois de la mission de terrain. Les inventaires de terrain du second mois ont été effectués par Gaspard Estenne.

L'entièreté de la recherche bibliographique, des analyses et de l'interprétation a été effectuée par Gaspard Estenne.

## VI Conclusion

La prédiction de land cover urbain est une tâche complexe. Dans le cas qui nous occupe aujourd'hui, cette complexité est accentuée par le dynamisme spatial accru de la zone d'étude et par la diversité élevée de cover présents.

Ce TFE offre une première proposition de clé typologique d'IV urbains, étape clé dans le cadre du monitoring des IV urbains et des services écosystémiques qui y sont associés. Cette clé doit potentiellement être adaptée, mais elle offre une base pour de futurs travaux.

Selon les résultats associés aux différentes images testées, cette dernière semble avoir moins d'influence sur la qualité de la prédiction que le prédicteur, cependant plus l'image utilisée pour la prédiction est proche de l'image utilisée pour la photo-interprétation, meilleur est le résultat.

L'harmonique de NDVI semble être le meilleur prédicteur parmi ceux testés pour prédire le cover de la ville de Lubumbashi. Ce travail met donc en évidence la nécessité de prendre en compte les séries temporelles pour décrire correctement le land cover de Lubumbashi, mais aussi potentiellement d'autres villes africaines.

Il aurait pu aussi être intéressant de tester une série temporelle d'images RGB afin de mieux comprendre le rôle du NDVI dans cette prédiction.



## VII Bibliographie

- 1 Maher Milad Aburas et al. “Measuring Land Cover Change in Seremban, Malaysia Using NDVI Index”. en. *Procedia Environmental Sciences* 30 (2015), pp. 238–243.
- 2 Marcos Adami et al. “Remote Sensing Time Series to Evaluate Direct Land Use Change of Recent Expanded Sugarcane Crop in Brazil”. en. *Sustainability* 4.4 (2012), pp. 574–585.
- 3 J Ahern. “Green infrastructure for cities: The spatial dimension”. en (2007), p. 17.
- 4 Marie Andre. “Landscape ecological consequences of the (sub)urbanization process in an African city: Lubumbashi (Democratic Republic of Congo)”. en. *BIO TECH* (2016), p. 229.
- 5 Shlomo Angel, Jason Parent, and Alexander Blei. “The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050”. en. *Progress in Planning* 75.2 (2011), pp. 53–107.
- 6 Millennium Ecosystem Assessment, ed. *Ecosystems and human well-being: wetlands and water synthesis: a report of the Millennium Ecosystem Assessment*. en. Washington, DC: World Resources Institute, 2005.
- 7 Carlos Bartesaghi-Koc, Paul Osmond, and Alan Peters. “A Green Infrastructure Typology Matrix to Support Urban Microclimate Studies”. en. *Procedia Engineering* 169 (2016), pp. 183–190.
- 8 Beatriz Bellón et al. “A Remote Sensing Approach for Regional-Scale Mapping of Agricultural Land-Use Systems Based on NDVI Time Series”. en. *Remote Sensing* 9.6 (2017), p. 600.
- 9 Mark Benedict and Edward McMahon. “Green Infrastructure: Smart Conservation for the 21st Century” (2002).
- 10 Mark A. Benedict and Edward T. McMahon. *Green Infrastructures Linking Landscapes and Communities*. Island Press. 2012.
- 11 Adia Bey et al. “Collect Earth: Land Use and Land Cover Assessment through Augmented Visual Interpretation”. en. *Remote Sensing* 8.10 (2016), p. 807.
- 12 Leo Breiman. “Random Forests”. *Machine Learning* (2001), p. 28.
- 13 F. Bultot. “Carte des régions climatiques du Congo Belge: établi d’après les critères de Köppen” (1950).

- 14 Mukalenge Chenge et al. “La carte sanitaire de la ville de Lubumbashi, République Démocratique du Congo Partie I : problématique de la couverture sanitaire en milieu urbain congolais”. fr. *Global Health Promotion* 17.3 (2010), pp. 63–74.
- 15 Changsoon Choi, Pam Berry, and Alison Smith. “The climate benefits, co-benefits, and trade-offs of green infrastructure: A systematic literature review”. en. *Journal of Environmental Management* 291 (2021), p. 112583.
- 16 Departement for Communities and Local Government. “Planning policy statement 12: creating strong safe and prosperous communities through local spatial planning.” (2008).
- 17 Christopher Coutts and Micah Hahn. “Green Infrastructure, Ecosystem Services, and Human Health”. en. *Int. J. Environ. Res. Public Health* (2015), p. 31.
- 18 Ruth DeFries, Matthew Hansen, and John Townshend. “Global discrimination of land cover types from metrics derived from AVHRR pathfinder data”. en. *Remote Sensing of Environment* 54.3 (1995), pp. 209–222.
- 19 Natural England. “Green Infrastructure Guidance”. *Natural England* (2009), p. 107.
- 20 Jay Fussell and Donald Rundquist. “On Defining Remote Sensing”. en. *PHOTOGRAMMETRIC ENGINEERING* (1986), p. 5.
- 21 Aniruddha Ghosh, Richa Sharma, and P.K. Joshi. “Random forest classification of urban landscape using Landsat archive and ancillary data: Combining seasonal maps with decision level fusion”. en. *Applied Geography* 48 (2014), pp. 31–41.
- 22 S E Gill et al. “Adapting Cities for Climate Change: The Role of the Green Infrastructure”. en. *CLIMATE CHANGE AND CITIES* 33.1 (2007), p. 19.
- 23 Li Guo et al. “Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests”. en. *ISPRS Journal of Photogrammetry and Remote Sensing* 66.1 (2011), pp. 56–66.
- 24 Dagmar Haase et al. “A Quantitative Review of Urban Ecosystem Service Assessments: Concepts, Models, and Implementation”. en. *AMBIO* 43.4 (2014), pp. 413–433.
- 25 A.R Huete. “A soil-adjusted vegetation index (SAVI)”. en. *Remote Sensing of Environment* 25.3 (1988), pp. 295–309.
- 26 INS. “Annuaire Statistique RDC 2020”. *Institut National de la Statistique*. Groupe de la Banque Africaine de Développement (2021), p. 433.
- 27 Marc Jakubauskas and David Legates. “HARMONIC ANALYSIS OF TIME-SERIES AVHRR NDVI DATA FOR CHARACTERIZING US GREAT PLAINS

- LAND USE/LAND COVER”. International Archives of Photogrammetry and Remote Sensing (2000).
- 28 Yuhao Jin et al. “Land-cover mapping using Random Forest classification and incorporating NDVI time-series and texture: a case study of central Shandong”. en. *International Journal of Remote Sensing* 39.23 (2018), pp. 8703–8723.
  - 29 Carol Kambites and Stephen Owen. “Renewed prospects for green infrastructure planning in the UK 1”. en. *Planning Practice and Research* 21.4 (2006), pp. 483–496.
  - 30 Christine Kessides. *La transition urbaine en Afrique subsaharienne*. Cities Alliance. 2006.
  - 31 Donghyun Kim and Seul-Ki Song. “The Multifunctional Benefits of Green Infrastructure in Community Development: An Analytical Review Based on 447 Cases”. en. *Sustainability* 11.14 (2019), p. 3917.
  - 32 Markus Kottek et al. “World Map of the Köppen-Geiger climate classification updated”. en. *Meteorologische Zeitschrift* 15.3 (2006), pp. 259–263.
  - 33 Jonas Lambert et al. “Monitoring forest decline through remote sensing time series analysis”. en. *GIScience & Remote Sensing* 50.4 (2013), pp. 437–457.
  - 34 Collins Adjei Mensah. “Urban Green Spaces in Africa: Nature and Challenges”. en. *International Journal of Ecosystem* (2014), p. 11.
  - 35 Robinson Mugo et al. “Quantifying Land Use Land Cover Changes in the Lake Victoria Basin Using Satellite Remote Sensing: The Trends and Drivers between 1985 and 2014”. en. *Remote Sensing* 12.17 (2020), p. 2829.
  - 36 Mpundu Mubemba Mulambi et al. “Teneurs en éléments traces métalliques dans les sols de différents jardins potagers de la ville minière de Lubumbashi et risques de contamination des cultures potagères”. fr (2013), p. 12.
  - 37 K. F. Munyemba et al. “Occupation des sols dans le cône de pollution à Lubumbashi” (2008).
  - 38 Sandra Nauman et al. *Design, implementation and cost elements of Green Infrastructure projects*. Tech. rep. 2011.
  - 39 M. Pal. “Random forest classifier for remote sensing classification”. en. *International Journal of Remote Sensing* 26.1 (2005), pp. 217–222.
  - 40 Thanh Noi Phan, Verena Kuch, and Lukas W. Lehnert. “Land Cover Classification using Google Earth Engine and Random Forest Classifier—The Role of Image Composition”. en. *Remote Sensing* 12.15 (2020), p. 2411.

- 41 Bradley C. Reed et al. “Measuring phenological variability from satellite imagery”. en. *Journal of Vegetation Science* 5.5 (1994), pp. 703–714.
- 42 Annemarie Schneider. “Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach”. en. *Remote Sensing of Environment* 124 (2012), pp. 689–704.
- 43 Karen C. Seto, Burak Güneralp, and Lucy R. Hutya. “Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools”. en. *Proceedings of the National Academy of Sciences* 109.40 (2012), pp. 16083–16088.
- 44 Luisa Sturiale and Alessandro Scuderi. “The Role of Green Infrastructures in Urban Planning for Climate Change Adaptation”. en (2019), p. 24.
- 45 C. Sys. “Principles of soil classification in the Belgian Congo.” Transactions 7th int. Congr. Soil Sci. Vol.4 (1960), pp.112–118.
- 46 Joseph Teye. “Urbanization and migration in Africa”. en (2018), p. 6.
- 47 United Nations, Department of Economic and Social Affairs, and Population Division. *World population prospects Highlights, 2019 revision Highlights, 2019 revision*. en. OCLC: 1110010089. 2019.
- 48 Hannah Wright. “Understanding green infrastructure: the development of a contested concept in England”. en. *Local Environment* 16.10 (2011), pp. 1003–1019.
- 49 Hanqiu Xu. “Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery”. en. *International Journal of Remote Sensing* 27.14 (2006), pp. 3025–3033.
- 50 Teresa Zölch et al. “Using green infrastructure for urban climate-proofing: An evaluation of heat mitigation measures at the micro-scale”. en. *Urban Forestry & Urban Greening* 20 (2016), pp. 305–316.

## VIII Annexe

### 1 Carte préliminaire

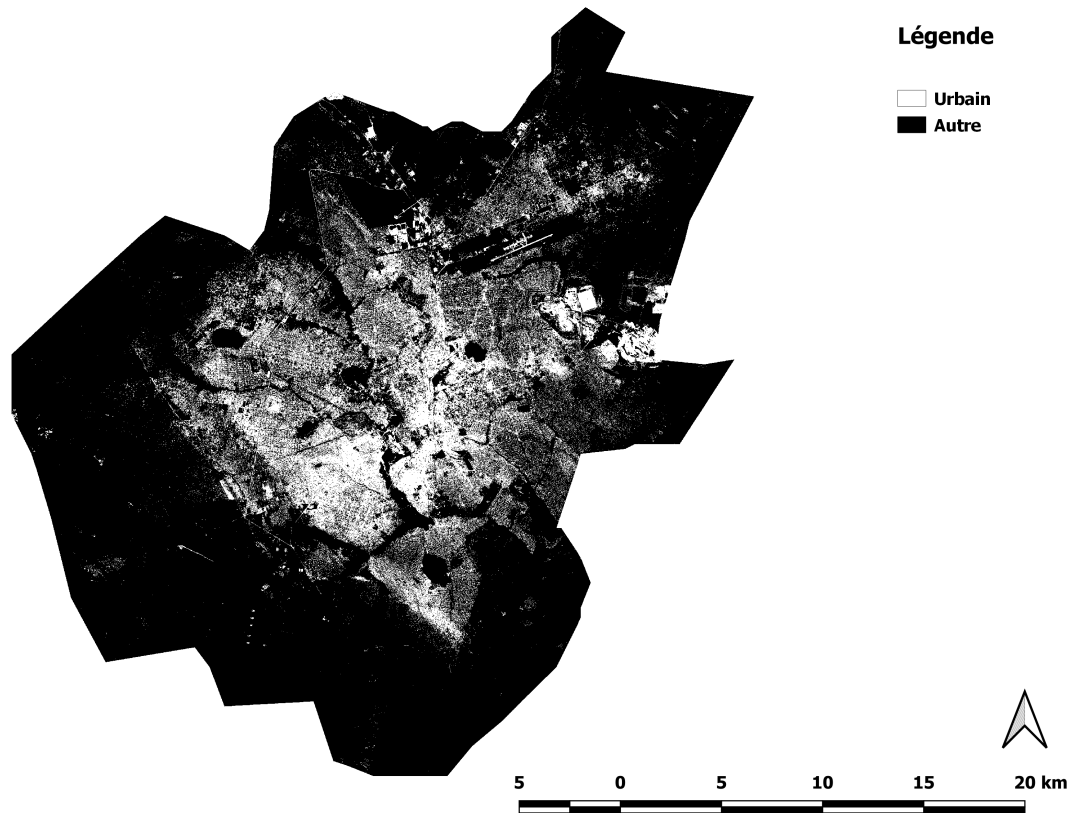


Figure 29: Carte préliminaire - Isolation des zones urbanisées et des zones végétalisées de la ville de Lubumbashi.

Cette carte a été utilisée pour sélectionner les points servant à placer les quadrats pour la photo-interprétation.

## 2 Résultats CEO

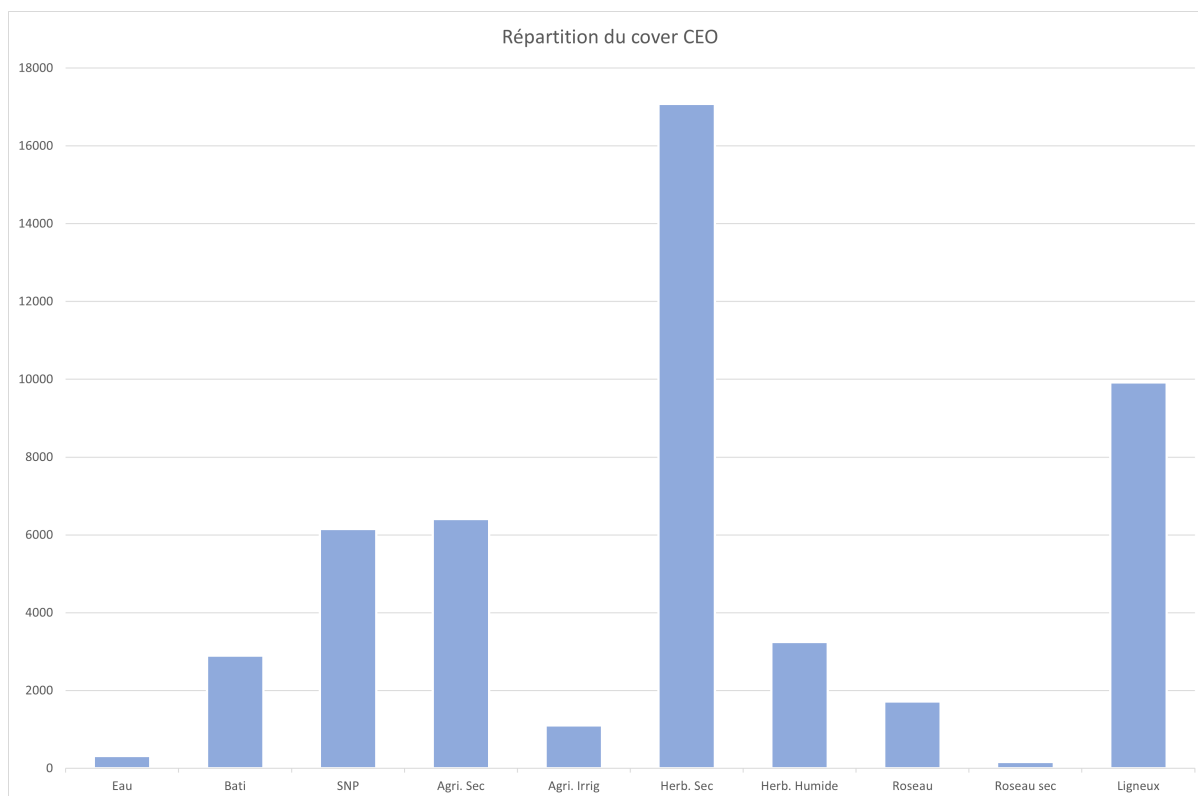
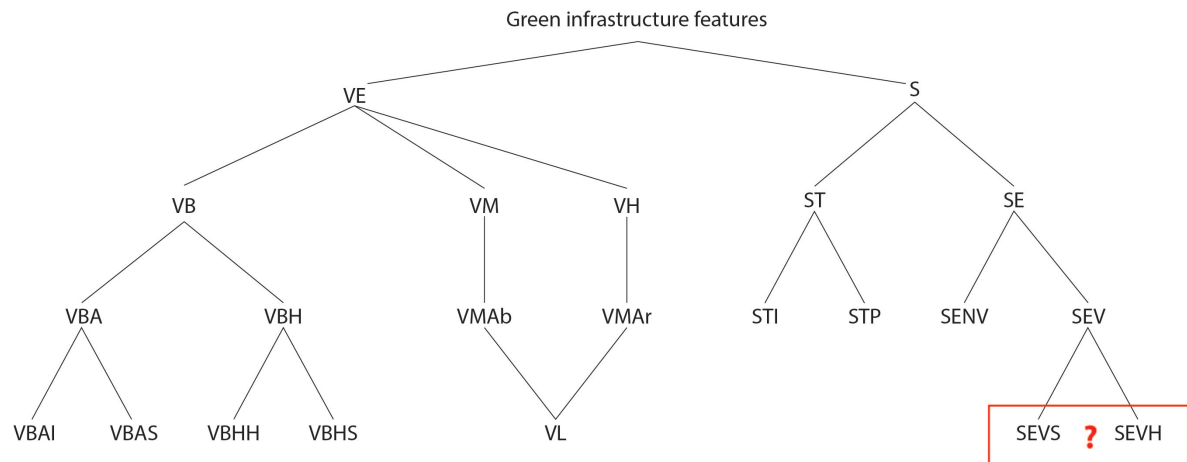


Figure 30: Répartition des 49 000 points photo-interprétés dans les différents covers utilisés.

### 3 Première clé typologique



VE : Végétation, S : Sol, VB : Basse, VM : Moyenne, VH : Haute, VBA : Agriculture, VBH : Herbacée, VMAb : Arbuste, VMAr : Arbres, VBAI : Irriguée, VBAS : Sèche, VBHH : «Humide», VBHS : Sèche «savane», VL : Ligneux.  
 ST : Terre, SE : Eau, STI : Imperméable, STP : Perméable, SEV : Végétale Roseaux, SENV : Non végétale, SEVS : Sec, SEVH : Humide «vivant».

Figure 31: Clé typologique des IV de Lubumbashi créée avant la mission terrain.