
Approche multi-capteurs pour la cartographie par télédétection des ressources ligneuses en Wallonie : application à la commune de Paliseul

Auteur : Crutzen, Florine

Promoteur(s) : Lejeune, Philippe; 2524

Faculté : Gembloux Agro-Bio Tech (GxABT)

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

Année académique : 2016-2017

URI/URL : <http://hdl.handle.net/2268.2/2964>

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.

**APPROCHE MULTI-CAPTEURS POUR LA
CARTOGRAPHIE PAR TELEDETECTION DES
RESSOURCES LIGNEUSES EN WALLONIE:
APPLICATION A LA COMMUNE DE PALISEUL**

FLORINE CRUTZEN

**TRAVAIL DE FIN D'ETUDES PRESENTE EN VUE DE L'OBTENTION DU DIPLOME DE
MASTER BIOINGENIEUR EN GESTION DES FORETS ET DES ESPACES NATURELS**

ANNEE ACADEMIQUE 2016-2017

CO-PROMOTEURS: PR. PHILIPPE LEJEUNE & CORENTIN BOLYN

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.

APPROCHE MULTI-CAPTEURS POUR LA CARTOGRAPHIE PAR TELEDETECTION DES RESSOURCES LIGNEUSES EN WALLONIE: APPLICATION A LA COMMUNE DE PALISEUL

FLORINE CRUTZEN

**TRAVAIL DE FIN D'ETUDES PRESENTE EN VUE DE L'OBTENTION DU DIPLOME DE
MASTER BIOINGENIEUR EN GESTION DES FORETS ET DES ESPACES NATURELS**

ANNEE ACADEMIQUE 2016-2017

CO-PROMOTEURS: PR. PHILIPPE LEJEUNE & CORENTIN BOLYN

Remerciements

Je tiens à remercier sincèrement toutes les personnes qui m'ont aidée, de près ou de loin, dans la réalisation de ce mémoire et qui contribuent ainsi à la clôture de ces cinq années d'étude à Gembloux:

- le Professeur P. Lejeune, de la Faculté des sciences agronomiques de Gembloux et le doctorant Corentin Bolyn, co-promoteurs de mon travail, pour m'avoir orientée durant ces quelques mois et pour leur relecture constructive ;
- l'Axe Gestion des Ressources forestières (Département BIOSE), pour l'accueil dans ses locaux, la mise à disposition de matériel et de données. Merci aux collègues de l'unité pour la bonne ambiance, notamment aux pauses ;
- les membres du jury, pour leur lecture et l'intérêt porté au sujet de mon mémoire ;
- Stéphanie, doctorante à l'unité, pour sa relecture pertinente et critique, pour sa présence, son soutien et son grain de folie ;
- le SPW, l'IPRFW et le cantonnement de bouillon (M. Gigounon) pour leur collaboration, en me fournissant les données nécessaires ;
- Nicolas, docteur à l'unité, pour son aide précieuse en code R ainsi que ses idées innovantes pour toujours aller plus loin ;
- Coralie, technicienne à l'unité, pour la récolte de données de terrain, indispensables pour évaluer la qualité de mon travail ;
- Samuel, informaticien à l'unité, pour le soutien logistique ;
- Maman et Nico, mes précieux relecteurs, pour leur avis critique et la correction linguistique de mon travail ;
- Toute la promo GFEN 2017, pour ces cinq années formidables passées à Gembloux, dans une ambiance soudée, propre aux forestiers de Gembloux. Un merci plus particulier aux TFistes du bureau avec qui j'ai eu la chance de partager tellement de moments de folie!! (et de stress aussi...) :
- Max (ou plutôt tous les Max dans sa tête!), SS, Duff, Axel, Léo, Bastien, Claire, Morgane... ;
- Mes ami(e)s (gembloutois et autres) d'avoir été à mes côtés et de m'avoir encouragée à présenter mon mémoire en première session. Merci aussi à ceux qui m'ont permis de m'évader de temps en temps durant ces 6 derniers mois: Scotland Team & Forestry Versatility 2017 à Brno (IFSA) ;
- Mes proches qui me soutiennent dans mon parcours personnel et mon projet professionnel.

Résumé

L'objectif de ce travail est de mettre en place une méthode de cartographie forestière à l'aide d'une synergie de capteurs de résolutions spatiale, temporelle et radiométrique différentes. La carte est générée sur la commune de Paliseul, en Ardenne belge, et est destinée au gestionnaire forestier ainsi qu'à la filière bois, dans un but de caractérisation de la ressource ligneuse. Trois types de données ont été utilisées: des ortho-images (2012-2013), un Modèle Numérique de Hauteur (MNH) (2014) et de l'imagerie satellitaire du satellite Sentinel-2A (S2) (2016). Les données à très haute résolution spatiale (orthophotos et MNH) entrent dans un traitement de segmentation d'image sur eCognition. Ensuite, l'information S2 permet de caractériser les segments pour la classification par objet à l'aide du modèle Random Forest. La qualité des modèles et résultats a été évaluée grâce à une matrice de confusion. Le résultat final de la classification par objet est une carte de l'occupation du sol en forêt avec une précision du producteur moyenne de 73,6% (entre 47% et 100% selon la classe forestière) et une précision du consommateur moyenne de 62,1% (entre 3% et 95% selon la classe forestière). Ces résultats sont prometteurs, d'autant que les confusions entre les classes trouvent souvent une explication concrète sur le terrain. Les pistes d'amélioration sont nombreuses et concernent notamment la révision et évaluation de la qualité de la segmentation, la rectification de la légende de classification et l'amélioration de la base de données d'entraînement. L'acquisition de données mises à jour et synchrones permettrait d'augmenter la précision de la méthode actuelle mais également de bénéficier à de nouvelles applications. Le développement d'une cartographie dynamique pour le suivi régulier et précis de nos forêts est une perspective qui répondrait aux besoins de gestion forestière durable.

Abstract

The aim of this study is to develop a multi-sensor approach for the mapping and the characterisation of wood resources in a municipality of Wallonia, Paliseul. This cartography should help the forest manager and the wood sector for the characterisation of natural resources of our forests. We used different sources of Remote Sensing information: Aerial imagery (2012-2013), a Digital Elevation Model (DEM) (2014) and satellite imagery from Sentinel-2A (S2) (2016). Data with high spatial resolution (aerial imagery and DEM) were used in an image segmentation processing with eCognition. Then S2 information is needed to describe objects in an object-based classification with the Random Forest model. The accuracies of models and maps were evaluated. Final results of the cartography of the forest land cover with an object-based approach showed an average producer accuracy of 73,6% (between 47% and 100% depending on the forest class) and an average consumer accuracy of 62,1% (between 3% and 95% depending on the forest class). These results are promising especially because the class confusion can mostly be explained practically on the field. There are many potential solutions that might help us to improve this map. For example, we can improve and evaluate segmentation quality, revise the classification legend and improve our training data. Updated and ideally synchronous data should increase the accuracy of the current method but also facilitate new applications. More research is needed to develop a dynamic map for constant and precise forest monitoring in a context of sustainable forest management.

Table des matières

1	Introduction.....	6
1.1	Contexte	6
1.2	Télédétection.....	6
1.2.1	Télédétection appliquée à la foresterie	6
1.2.2	Les principaux produits de télédétection utilisés en foresterie.....	8
1.2.3	Les principales techniques de classification d'images.....	11
1.3	Cadre spécifique en Région wallonne	11
1.4	Objectif	12
2	Matériel et méthodes.....	12
2.1	Site d'étude	12
2.2	Données de télédétection utilisées.....	14
2.2.1	Sentinel-2A	14
2.2.2	Ortho-images.....	15
2.2.3	Modèle Numérique de Hauteur (MNH)	18
2.3	Traitement des données	19
2.3.1	Introduction.....	19
2.3.2	Préparation des données	21
2.3.3	Création d'indices spectraux	21
2.3.4	Création des segments	23
2.3.5	Calcul des variables descriptives (<i>metrics</i>) des objets.....	24
2.4	Classification.....	25
2.4.1	Introduction.....	25
2.4.2	Définition de la légende cartographique.....	26
2.4.3	Création des données d'entraînement.....	27
2.4.4	Classification Random Forest	29
2.4.5	Evaluation de la précision de la classification à l'aide d'une matrice de confusion.....	30
3	Résultats et discussions.....	32
3.1	Segmentations.....	32
3.2	Classifications	33
3.2.1	Modèle A	33
3.2.2	Modèle B	35
3.2.3	Approche par pixel sur le niveau 2	39
3.3	Comparaison des prédictions, références DNF et validations terrain	44
3.4	Avantages et inconvénients de la méthode	47

3.5	Pistes d'explication de la qualité du modèle.....	48
3.6	Pistes concrètes d'amélioration du modèle de classification	49
4	Perspectives.....	51
4.1	100% en libre accès	51
4.2	Déploiement de la méthode	51
4.3	Cartographie dynamique.....	51
4.4	Approche quantitative.....	51
4.5	Cartographie des arbres hors forêt	52
4.6	Suivi des coupes rases et des replantations.....	52
4.7	Caractérisation de l'habitat	52
5	Conclusion	54
	Bibliographie.....	55
	Annexes	61
	Annexe 1: Orthophoto 2012-2013 sur la commune de Paliseul ; source: SPW	61
	Annexe 2: Schéma général du traitement des données et de la classification.....	62

Sommaire des figures

Figure 1: Détermination de l'occupation du sol (sol nu, végétation et eau) sur base du signal spectral de la réflectance (réflexion du rayonnement solaire) ; source: http://remote-sensing.net/concepts.html	7
Figure 2: Réponses spectrales pour différentes classes d'occupation du sol ; source: Cipar <i>et al.</i> , 2004	7
Figure 3: Illustration des différentes bandes spectrales du capteur des satellites Sentinel-2. L'axe des abscisses correspond à la longueur d'onde (en nm). L'axe des ordonnées permet de préciser les résolutions spatiales de différentes bandes spectrales (10, 20, ou 60 m) ; source: ESA	9
Figure 4: Illustration du principe du LiDAR; d'après: Jensen, 2007	10
Figure 5: Carte des zones bioclimatiques de la Région wallonne et localisation de la zone d'étude ; source: Van der Perre <i>et al.</i> , 2015	13
Figure 6: Extraits des informations Sentinel-2A sur la zone d'étude pour les deux dates sélectionnées: 8 mai 2016 (au-dessus) et 20 juillet 2016 (en-dessous) à une résolution spatiale de 10 m et selon un niveau de correction L2A (FRE) ; source: ESA (corrigé par THEIA)	15
Figure 7: Extrait de la couche ortho-image 2012-2013 sur la zone d'étude. Une version vraie couleur est présentée au-dessus et une version fausse couleur en-dessous, mettant en évidence la végétation à une résolution spatiale de 25 cm et les problèmes de variabilité radiométrique; source: SPW	17
Figure 8: Illustration des différents modèles numériques: de terrain, de surface et de hauteur ; source: Dedry <i>et al.</i> , 2015	18
Figure 9: Extrait de la couche MNH LiDAR sur la zone d'étude mettant en évidence des hauteurs élevées en blanc et faibles en noir. La résolution spatiale du MNH permet de mettre en évidence les houppiers ; source: SPW	18
Figure 10: Schéma de la chaîne de traitement des données	20
Figure 11: Paramètres retenus pour la segmentation de Paliseul dans eCognition sur base d'essais et erreurs ainsi qu'une interprétation visuelle des résultats	24
Figure 12: Schéma des étapes de la classification supervisée	25
Figure 13: Choix des classes qui seront utilisées dans le modèle de classification selon un concept d'objets hiérarchisés	26
Figure 14: Extrait du parcellaire du Département de la Nature et des Forêts sur la zone d'étude. La délimitation des îlots est représentée par les polygones rouges pour l'année 2013 ; source: SPW	27
Figure 15: Illustration des imprécisions du parcellaire DNF. La délimitation des îlots est représentée par les polygones rouges pour l'année 2013 ; source: SPW	28
Figure 16: Nombre de segments de référence par classe pour l'entraînement des modèles de classification	29
Figure 17: Exemple théorique de matrice de confusion dans le cas de 2 classes. Calcul des précisions du producteur et du consommateur	31
Figure 18: Distribution des fréquences des aires des segments obtenus avec eCognition sur la commune de Paliseul	32
Figure 19: Extrait de la segmentation à l'aide d'eCognition sur une portion du territoire de la commune de Paliseul. Illustration du comportement de l'algorithme pour différentes classes d'occupation du sol ; source : SPW	33
Figure 20: Indices de Gini pour les 30 variables sélectionnées pour la classification "Forêt/Hors forêt" (modèle A)	34

Figure 21: Indices de Gini pour les 30 variables sélectionnées pour la classification sur le niveau 2 en forêt (modèle B)	35
Figure 22: Evaluation de la qualité "Out Of Bag" du modèle B de classification par classe en forêt. La précision du producteur est représentée en bleu et la précision du consommateur en rouge pour les classes Chêne, Douglas, Epicéa, Hêtre, Mélèze, Pin, Régénération et Sol nu forestier	36
Figure 23: Evaluation de la qualité du modèle B de classification par classe selon la validation des points d'échantillonnage terrain. La précision du producteur est représentée en bleu et la précision du consommateur en rouge pour les classes Chêne, Douglas, Epicéa, Hêtre, Mélèze, Pin, Régénération et Sol nu forestier	38
Figure 24: Superficies forestières prédites des classes du modèle B (Chêne/Douglas/Epicéa/ Hêtre/Mélèze/Pin/ Régénération/Sol nu forestier) sur l'ensemble de la commune de Paliseul	39
Figure 25: Indices de Gini pour les 30 variables sélectionnées pour le modèle de classification par pixel (niveau 2 en forêt)	40
Figure 26: Précision Out Of Bag du modèle de classification par pixel. La précision du producteur est représentée en bleu et la précision du consommateur en rouge pour les classes Chêne, Douglas, Epicéa, Hêtre, Mélèze, Pin, Régénération et Sol nu forestier	41
Figure 27: Précision des modèles de classification par objet et par pixel selon la validation des points d'échantillonnage terrain pour les classes Chêne, Douglas, Epicéa, Hêtre, Mélèze, Pin, Régénération et Sol nu forestier	43
Figure 28: Nombre de pixels (10 m de côté) prédit par classe forestière de niveau 2 (Chêne/Douglas/Epicéa/ Hêtre/Mélèze/Pin/ Régénération/Sol nu forestier) selon la classification par pixel sur la commune de Paliseul	43
Figure 29: Extrait des résultats de prédictions sur la zone d'étude selon la classification par pixel (voir le code couleur de la légende) et par objet (voir la classe prédite pour chaque segment) (niveau 2 en forêt). L'information du parcellaire est présentée dans les étiquettes blanches et délimitée en rose. Les données de validation terrain sont présentées dans les disques blancs et délimitées en rouge ; source: SPW	44
Figure 30: Extrait des résultats de prédictions sur la zone d'étude selon la classification par pixel (voir le code couleur de la légende) et par objet (voir la classe prédite pour chaque segment) (niveau 2 en forêt). L'information du parcellaire est présentée dans les étiquettes blanches et délimitée en rose. Les données de validation terrain sont présentées dans les disques blancs et délimitées en rouge. Etude des peuplements feuillus ; source: SPW	46
Figure 31: Classification selon le modèle B sur une zone forestière de Paliseul mettant en évidence un segment à cheval sur deux peuplements et classé en mélèze. Cet exemple illustre l'impact des imprécisions dans les segments sur la qualité de la classification ; source: SPW	47
Figure 32: Illustration des îlots sur le terrain : Ilot de chênes "pur" ; Source: F. Crutzen	47
Figure 33: Illustration de quelques éléments qui influencent la qualité finale de la carte.....	48
Figure 34: Classement des principales espèces ligneuses d'Europe occidentale d'après leur potentiel biologique (en % de la richesse spécifique maximale par groupe fonctionnel) ; source: Branquart et Dufrêne, 2005 in Branquart et Liégeois, 2005	53
Figure 35: Réponse du LiDAR haute fréquence (en moyenne 450 points par mètre carré) à travers le couvert forestier avec une résolution de 5 cm ; source: FlyThru.....	53

Sommaire des tableaux

Tableau 1: Résumé des dates d'acquisition des ortho-images sur la Région Wallonne et sur la commune de Paliseul ; source: Géoportail de la Wallonie, 2017	16
Tableau 2: Outils logiciels utilisés dans le cadre de ce travail.....	21
Tableau 3: Définition des différentes classes utilisées pour la classification.....	26
Tableau 4: Matrice de confusion générée par Random Forest pour le modèle A.....	34
Tableau 5: Matrice de confusion générée pour la validation du modèle A selon la grille (500 m)	34
Tableau 6: Matrice de confusion générée par Random Forest pour le modèle B (Chêne/Douglas/Epicéa/ Hêtre/Mélèze/Pin/ Régénération/Sol nu forestier).....	36
Tableau 7: Matrice de confusion de la validation des points d'échantillonnage terrain du modèle B.	37
Tableau 8: Matrice de confusion Out Of Bag du modèle de classification par pixel (Chêne/Douglas/Epicéa/ Hêtre/Mélèze/Pin/ Régénération/Sol nu forestier).....	40
Tableau 9: Matrice de confusion de la validation des points d'échantillonnage terrain du modèle de classification par pixel	42

1 Introduction

La première partie de ce travail vise à introduire le sujet d'étude en exposant le contexte global de la gestion forestière. Puis, l'utilisation de la télédétection en foresterie sera abordée en présentant la méthode de détection de la végétation, quelques produits de télédétection disponibles et le principe de classification d'image. Enfin, nous présenterons le contexte plus spécifique à ce travail et les objectifs.

1.1 Contexte

Depuis la nuit des temps, l'Homme vit en interaction avec son environnement ainsi qu'avec la forêt qui l'entoure. Celle-ci fournit de nombreux «services écosystémiques» qui sont repris selon trois catégories (Lindner *et al.*, 2010) :

- services économiques, qui concernent, par exemple, la production de bois et de produits forestiers non-ligneux tels que les baies et les champignons ;
- services écologiques, qui reprennent notamment le rôle d'accueil de la biodiversité et de soutien aux processus biogéochimiques (tels que le cycle du carbone et le cycle de l'eau) ;
- services sociaux et culturels, illustrés entre autres par le rôle récréationnel et ressourçant de nos forêts.

Dans un objectif de gestion durable de la forêt, il est primordial d'en caractériser la distribution spatiale et la nature. Cette information est traditionnellement disponible sous forme de cartes, de l'échelle mondiale jusqu'à l'échelle locale.

En plus de la cartographie, la ressource forestière peut également être caractérisée sur base d'inventaires de terrain réalisés généralement par voie d'échantillonnage. La caractérisation forestière est possible à différentes échelles spatiales. Des organismes tels que la FAO¹ réalisent un suivi à l'échelle mondiale². Les pays ou les régions disposent également d'un outil de monitoring de leurs forêts. En Région wallonne, l'Inventaire Permanent des Ressources Forestières de Wallonie (IPRFW) suit des unités d'échantillonnage, réparties selon une grille systématique à travers toute la forêt wallonne³. Et enfin, à l'échelle locale des propriétés forestières, les forestiers de terrain connaissent leurs peuplements et les gèrent à l'aide de parcellaires et d'inventaires «en plein».

Malgré les connaissances actuelles sur nos forêts, les demandes en informations sur les ressources forestières sont de plus en plus nombreuses et exigent des niveaux de précision sans cesse croissants. Les avancées technologiques, notamment en ce qui concerne la cartographie et les moyens de télédétection (diversité et spécificité des capteurs et des plateformes) fournissent des données précieuses, et parfois gratuites, pour une large gamme de résolutions spatiale, spectrale et temporelle. Ainsi, la télédétection peut répondre à la demande toujours croissante d'informations précises, à haute résolution spatiale et temporelle et ceci de façon relativement peu onéreuse (Renaud *et al.*, 2015).

1.2 Télédétection

1.2.1 Télédétection appliquée à la foresterie

La télédétection est « la technique qui, par l'acquisition d'images, permet d'obtenir de l'information sur la surface de la Terre sans contact direct avec celle-ci. La télédétection englobe tout le processus qui consiste à capter et à enregistrer l'énergie d'un rayonnement électromagnétique émis ou réfléchi, à traiter et à analyser l'information, pour ensuite mettre en application cette information»⁴. Les notions fondamentales en télédétection ne seront pas présentées dans ce document mais toute

¹ Organisation des Nations Unies pour l'alimentation et l'agriculture

² <http://www.fao.org/publications/sofo/2016/fr/> (consulté le 19 mai 2017)

³ <http://environnement.wallonie.be/dnf/inventaire/> (consulté le 19 mai 2017)

⁴ http://www.rncan.gc.ca/sites/www.rncan.gc.ca/files/earthsciences/pdf/resource/tutor/fundam/pdf/fundamentals_f.pdf (consulté le 17 mai 2017)

l'information nécessaire est disponible dans une publication du Centre canadien de télédétection⁵ ainsi que dans le livre de référence "*Remote Sensing of the Environment: An Earth Resource Perspective*" de Jensen (2007).

Depuis la Première Guerre mondiale, la télédétection est employée pour l'observation de la Terre. Du simple appareil photo embarqué à bord d'un avion aux satellites équipés de capteurs multi-spectraux, la télédétection est utilisée partout dans le monde, de manière complémentaire aux données de terrain. L'objectif est d'obtenir rapidement de l'information sur l'environnement sur une vaste zone. Durant la Seconde Guerre mondiale, on a découvert qu'adapter un film infrarouge à un appareil photo permettait de bien mieux détecter la végétation. En 1960, différentes végétations peuvent être discernées grâce à l'étude des longueurs d'onde. (ESA, 2017)⁶

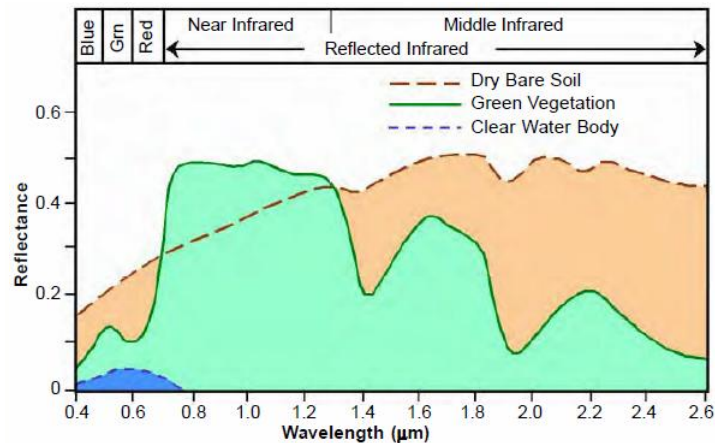


Figure 1: Détermination de l'occupation du sol (sol nu, végétation et eau) sur base du signal spectral de la réflectance (réflexion du rayonnement solaire) ; source: <http://remote-sensing.net/concepts.html>

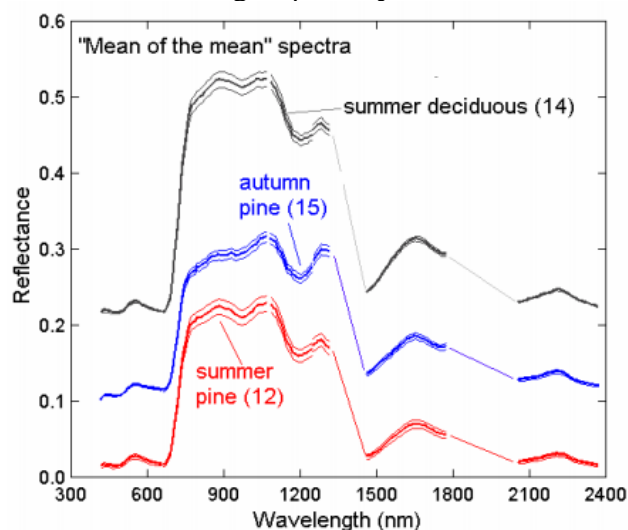


Figure 2: Réponses spectrales pour différentes classes d'occupation du sol ; source: Cipar *et al.*, 2004

La télédétection permet de différencier la végétation des autres occupations du sol (sol nu, eau, urbain) (Figure 1). Ceci permet l'étude du changement d'occupation du sol ou de l'environnement, d'origine naturelle ou anthropique. Pour la forêt, le suivi du morcellement, de la dégradation et du

⁵ http://www.rncan.gc.ca/sites/www.rncan.gc.ca/files/earthsciences/pdf/resource/tutor/fundam/pdf/fundamentals_f.pdf (consulté le 17 mai 2017)

⁶ http://www.esa.int/SPECIALS/Eduspace_FR/SEMUX1S7NWF_0.html (consulté le 22 avril 2017)

déboisement font partie des applications possibles à l'échelle régionale⁷. Localement, l'étude des coupes rases en milieu forestier est également possible.

L'analyse des réponses spectrales qui correspondent à l'infrarouge proche ($\lambda=0.8 \mu\text{m}$) et au rouge ($\lambda=0.66 \mu\text{m}$) permet de mettre en évidence la végétation active (Figure 1). Par ailleurs, divers indices sont basés sur le rayonnement électromagnétique dans ces gammes de longueurs d'onde. Le plus connu est le *Normalized Difference Vegetation Index* (NDVI) qui quantifie l'activité photosynthétique:

$$NDVI = \frac{\text{NearInfraRed} - \text{Red}}{\text{NearInfraRed} + \text{Red}}$$

Son évolution temporelle reflète la phénologie des types de végétation (Defries et Townshend, 1994). Wang J. *et al.* (2004) ont également montré un lien fort entre le NDVI et une estimation de la production ligneuse.

L'étude de la réflectance peut permettre de discriminer des types de végétations ou d'essences. La distinction entre les feuillus et résineux est assez aisée (Figure 2). Cependant, la reconnaissance des essences feuillues (hêtre et chêne) reste plus compliquée étant donné leurs réponses spectrales très analogues (Figure 2).

Enfin, la télédétection est utilisée pour caractériser la ressource tant en termes de discrimination d'essences qu'en estimation de paramètres quantitatifs (dendrométriques).

1.2.2 Les principaux produits de télédétection utilisés en foresterie

La télédétection fournit une multitude de produits, possédant chacun des propriétés de résolution spatiale, temporelle et spectrale propre.

1.2.2.1 Les données satellites à basse résolution spatiale

MODIS⁸, Moderate Resolution Imaging Spectroradiometer, est un exemple d'instrument à basse résolution spatiale. Embarqué à bord des satellites Aqua et Terra de la NASA, il collecte des données sur la planète entière en 1 à 2 jours sur 36 bandes spectrales. Selon la bande concernée, la résolution spatiale varie de 250 m à 1000 m.

1.2.2.2 Les données satellites à moyenne résolution spatiale

Les capteurs à moyenne résolution, embarqués à bord de satellites, ont permis l'étude plus spécifique et le suivi de la végétation à l'échelle terrestre. Les applications liées aux capteurs à moyenne et haute résolution concernent la cartographie de l'occupation du sol (respectivement à l'échelle du globe ou régionale), l'estimation du rendement agricole, le suivi des feux de forêt, la surveillance des catastrophes naturelles (inondations, incendies...), le suivi des courants marins, la mesure de l'anthropisation (urbanisation, déforestation...) ou encore l'étude des interactions climat-végétation. (UVED, 2008)⁹

Par exemple, les missions Landsat fournissent des données à moyenne résolution (15 à 30 mètres). En 1972, le Landsat 1 (Earth Resources Technology Satellite de la NASA) embarque le 1er scanner multi-spectral. Il s'agit d'un capteur passif qui capte l'énergie lumineuse solaire et thermique infrarouge. Les données produites par le scanner sont des matrices numériques chorologiques, c'est-à-dire un enregistrement de données numériques par canal et transmises à la Terre par communication radio (ESA, 2017)¹⁰. Depuis plus de 40 ans, la série de satellites Landsat assure une continuité temporelle d'acquisition de données.

⁷ <http://e-cours.univ-paris1.fr/modules/uved/envcal/html/vegetation/role-teledetection/basse-resolution.html> (consulté le 9 mai 2017)

⁸ <https://modis.gsfc.nasa.gov> (consulté le 17 mai 2017)

⁹ <http://e-cours.univ-paris1.fr/modules/uved/envcal/html/vegetation/role-teledetection/basse-resolution.html> (consulté le 9 mai 2017)

¹⁰ http://www.esa.int/SPECIALS/Eduspace_FR/SEMUX1S7NWF_0.html (consulté le 22 avril 2017)

1.2.2.3 Les données satellites à haute résolution spatiale

Sentinel-2 est un complexe de 2 satellites jumeaux déphasés de 180° sur la même orbite (786 km d'altitude), Sentinel-2A (en orbite depuis le 23 juin 2015) et Sentinel-2B (mis en orbite le 7 mars 2017). Il s'agit d'une mission dirigée par le European Spatial Agency (ESA) dans le cadre du programme Copernicus (ESA, 2017)¹¹. L'objectif du programme Copernicus est la mise en place d'un système européen de monitoring de la Terre (European Commission, 2017)¹².

Sentinel-2A revisite la même cible tous les 10 jours à l'équateur avec une résolution spatiale de 10 m. Avec l'arrivée en orbite de Sentinel-2B, la fréquence de passage est doublée. Ceci augmente la probabilité d'obtenir des images dépourvues de perturbations nuageuses. Dès lors que l'on se rapproche des pôles, le recouvrement des fauchées est tel que le temps de revisite tombe à 2-3 jours. (ESA, 2017)¹¹

Les satellites Sentinel-2 possèdent un instrument multispectral (*MSI*) à bord et collectent des données sur 12 bandes spectrales (Figure 3) (ESA, 2012)¹³. Les capteurs ont une haute résolution spatiale et permettent une étude précise de l'occupation du sol et de son évolution (étude diachronique).

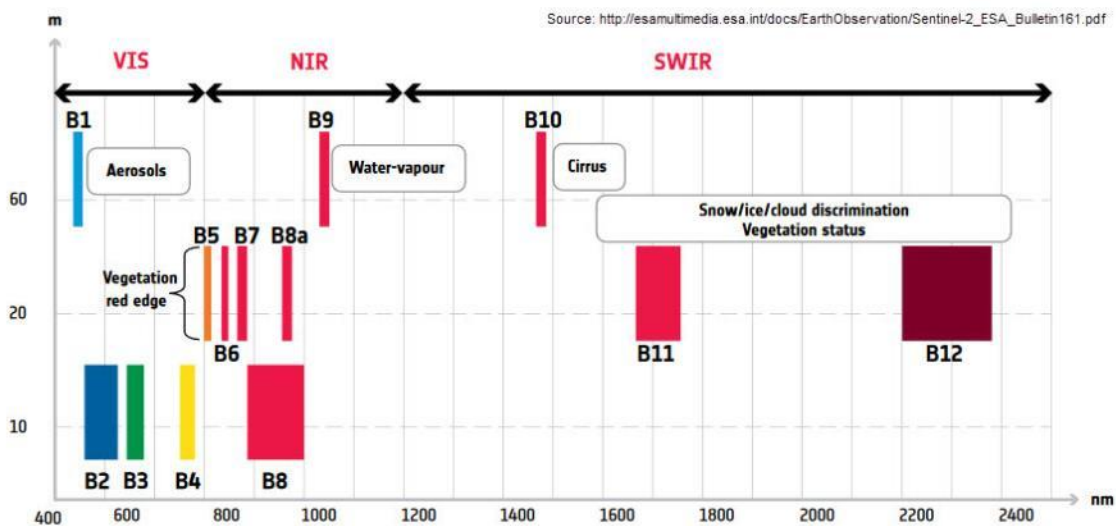


Figure 3: Illustration des différentes bandes spectrales du capteur des satellites Sentinel-2. L'axe des abscisses correspond à la longueur d'onde (en nm). L'axe des ordonnées permet de préciser les résolutions spatiales de différentes bandes spectrales (10, 20, ou 60 m) ; source: ESA

1.2.2.4 Les données satellite à très haute résolution

Enfin, il existe des imageries satellite à très haute résolution spatiale.

Par exemple, IKONOS est un satellite commercial lancé en 1999 qui fournit de l'information d'imagerie multispectrale à 1 mètre de résolution grâce au processus de *pansharpening*¹⁴. Avec les avancées technologiques, GeoEye, lancé en 2008, atteint moins de 0,5 mètres de résolution spatiale grâce à la même technique (*pansharpening*).

En plus d'une grande précision, les satellites à très haute résolution, principalement commerciaux, peuvent être programmés pour fournir des données sur une zone spécifique. Ceci est le cas des

¹¹ <https://sentinel.esa.int/web/sentinel/missions/sentinel-2> (consulté le 22 avril 2017)

¹² <http://ec.europa.eu/growth/sectors/space/copernicus/> (consulté le 22 avril 2017)

¹³ https://sentinel.esa.int/documents/247904/349490/S2_SP-1322_2.pdf (consulté le 22 avril 2017)

¹⁴ *Pansharpening* = combinaison de la haute résolution spatiale de la bande panchromatique avec la faible résolution spatiale de l'imagerie multispectrale.

images Pléiades (moins de 0,5 mètres de résolution spatiale). En effet, trois plans de programmation journaliers permettent de tenir compte de la demande. (Airbus Defense and Space, 2017)¹⁵

1.2.2.5 Les données Radar

Le programme Sentinel comporte également 2 satellites (1A et 1B) équipés de capteurs radar. Ils fournissent des données intéressantes nuit et jour. Bien que ce type de données soit potentiellement intéressantes pour la cartographie des ressources forestières (voir par exemple la publication de Dostalova *et al.*, 2017)¹⁶, elles requièrent une approche très spécifique qui ne sera pas abordée dans ce travail.

1.2.2.6 Les données d'imagerie aériennes

Longtemps, les données aériennes furent la seule source de données disponible. Traditionnellement utilisée en foresterie, pour la cartographie notamment, cette source de données avait perdu de son intérêt. Mais avec les avancées technologiques et les nouveaux logiciels de photogrammétrie¹⁷, le développement d'orthophotos sur base d'images aériennes permet d'extraire de l'information intéressante à très haute résolution. Par exemple, la stéréoscopie recrée des produits en trois dimensions à partir d'images en deux dimensions qui se chevauchent.

1.2.2.7 Les données LiDAR

Le LiDAR (Light Detection And Ranging) est un système actif, qui émet un faisceau laser et analyse la partie du signal qui revient au capteur (temps, distance, orientation...) (Figure 4). Cette technologie a été développée en 1960 mais ce n'est qu'en 1980 que ses informations sont utilisées pour des applications en forêt.

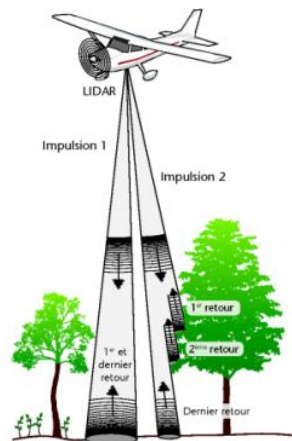


Figure 4: Illustration du principe du LiDAR; d'après: Jensen, 2007

Actuellement, la technologie se retrouve sur une grande diversité de plateformes, du LiDAR terrestre au satellite ICESat. Pour l'étude de la forêt à l'échelle locale, c'est la plateforme aérienne (avion ou hélicoptère) qui est souvent utilisée afin de survoler une zone et acquérir des données selon une densité de points définie.

Les informations qui peuvent être interprétées après un passage LiDAR sont par exemple la hauteur, la biomasse et la quantification de la ressource ligneuse. A l'avenir, les LiDAR multispectraux pourraient fournir de l'information spectrale pour les essences dominantes et dominées permettant

¹⁵ <http://www.intelligence-airbusds.com/fr/237-pleiades-imagerie-satellite-tres-haute-resolution> (consulté le 12 mai 2017)

¹⁶ <http://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/III-7/227/2016/isprs-annals-III-7-227-2016.pdf> (consulté le 25 mai 2017)

¹⁷ La photogrammétrie est "l'ensemble des techniques qui permettent de déterminer la forme, les dimensions, la position dans l'espace d'un objet à partir de photographies" (selon www.larousse.fr, consulté le 18 mai 2017).

de visualiser la forêt en 3 dimensions et «en couleurs» (Michez *et al.*, 2017). Un aspect plus écologique peut être étudié tel que la détection de la composition spécifique au moyen des hauteurs et de l'intensité du signal LiDAR (Donoghue *et al.*, 2007).

Cette technologie a cependant ses limites en termes de précision. Par exemple, la hauteur de l'apex d'un houppier sera sous-estimée puisqu'un point résume l'information d'un mètre carré. (Parkan, 2017)¹⁸

1.2.3 Les principales techniques de classification d'images

La classification d'image est un traitement automatisé qui vise à traduire l'information quantitative d'une image sous forme de variables discrètes. Il existe deux grandes familles de classification:

- la classification non supervisée est une approche non dirigée qui fonctionne sur un principe de *clustering*, c'est-à-dire que le jeu de données est séparé automatiquement en nuages (correspondant à des classes thématiques ou non) selon un processus itératif ;
- la classification supervisée, qui fera l'objet de ce travail, est une approche dirigée qui nécessite la définition préalable des classes thématiques avec la création d'un échantillon de données (d'entraînement et de validation). Un algorithme de classification, paramétrique ou non, est ensuite appliqué.

Au sein de chaque famille, on distingue également l'approche «par pixel» de l'approche «par objet»:

- la classification «par pixel» traite directement les pixels et les classe individuellement ;
- la classification «par objet» traite des groupements de pixels (objets). Cette approche nécessite au préalable un autre traitement d'image: la segmentation. Cette étape peut être définie comme étant « un processus de cloisonnement d'image en régions homogènes qui ne se recoupent pas ». De nombreux algorithmes existent, incluant la segmentation d'images industrielle et médicale (Pal et Pal, 1993).

En présence de données à haute résolution, la classification par objet est préférée (Lu et Weng, 2007). L'approche par objet serait plus précise car les capacités d'extraction de l'information sont considérablement multipliées: l'information spectrale contextualisée, la forme et la résolution spatiale (Blaschke, 2004).

1.3 Cadre spécifique en Région wallonne

En Région wallonne, la volonté d'investir dans les technologies relevant de la télédétection est observée pour la caractérisation du territoire notamment en matière de financement de campagnes d'acquisition d'orthophotos et de données LiDAR¹⁹ (en vue de produire des modèles numériques de surface et de terrain). De plus, les données Sentinel-2 fournissent une information à l'échelle régionale intéressante sur le plan des résolutions spatiale, spectrale et temporelle.

Mais les ambitions dépassent les limites de notre région, par exemple au travers de projets transfrontaliers/interrégionaux:

- Le projet Interreg V « Feel Wood - Forêt Pro Bos » concerne une zone de 62 000 km² à cheval sur le nord-est de la France et l'ouest de la Wallonie et la Flandre. Dans un cadre de valorisation de la forêt et du circuit court pour la filière bois, l'objectif de ce projet est l'identification des sources d'essoufflement de la filière. Pour cela, il est nécessaire d'évaluer finement la ressource et de porter une attention particulière au reboisement. Plus d'informations sont disponibles sur le net²⁰ ;

¹⁸ <https://sites.google.com/site/lidareforet/outils-et-methodes/interpretation-et-utilisation-des-produits-drivs> (consulté le 22 avril 2017)

¹⁹ Consulter le Géoportail de la Wallonie pour plus d'informations (<http://geoportail.wallonie.be>)

²⁰ <https://foretprobos.wordpress.com/> (consulté le 9 mai 2017)

<http://www.carah.be/projets/foret-pro-bos.html> (consulté le 1 juin 2017)

- Les projets Interreg VA Grande Region: «Regiowood I» (clôturé en 2011) et «Regiowood II» (2017-2019) réunissant les forestiers de la France, la Wallonie, l'Allemagne et le Luxembourg. Le projet « Regiowood II» fait partie de l'objectif 3 de l'axe prioritaire 2 : Assurer un développement respectueux de l'environnement et du cadre de vie²¹.

Ces projets, partiellement financés par l'Europe (*Fonds Européen de Développement Régional*), témoignent de la volonté des politiques de « décrire la ressource forestière à l'échelle de bassins d'approvisionnement », « répondre aux besoins en informations des gestionnaires publics ou privés» et enfin « dynamiser la filière bois, par une meilleure connaissance de la ressource»²².

En effet, au vu des faibles superficies forestières en Wallonie (environ 500 000 ha de surface productive) et de l'importance croissante du rôle multifonctionnel de nos forêts (forêts périurbaine, productive ou protégée), il est nécessaire de s'intéresser à la ressource disponible au-delà des frontières, en créant une base de données commune. Les résultats de ces projets ont également comme objectif d'aider les politiques à entamer une démarche transfrontalière commune de gestion durable de la ressource forestière. Cette démarche devra bénéficier directement à la filière bois qui pourra mieux évaluer les types d'essences et les volumes ligneux qui approvisionneront le marché dans les prochaines années.

1.4 Objectif

Partant du constat (i) du questionnement grandissant sur l'évolution de la ressource forestière, notamment dans le cadre du projet « Forêt Pro Bos » , (ii) de l'arrivée de nouvelle génération d'imagerie Sentinel-2 ainsi que (iii) de la disponibilité en données aériennes et 3D à très haute résolution (orthophotos et LiDAR), il nous a semblé pertinent d'approfondir la question de la plus-value d'une exploitation simultanée des ces différentes sources de données de télédétection pour améliorer la cartographie de nos ressources forestières. En effet, on peut penser que la combinaison de la très haute résolution spatiale des ortho-images avec l'information 3D issues des données LiDAR et l'information radiométrique multi-dates des images Sentinel-2 est susceptible d'ouvrir de nouvelles perspectives notamment en matière de cartographie des types de peuplements forestiers.

Notre travail tente donc de répondre à la question : L'approche multi-capteurs est-elle susceptible de permettre de cartographier, avec précision, les types de peuplements dans des massifs forestiers morcelés et hétérogènes tels que ceux de l'Ardenne belge?

A cette fin, nous utiliserons les données satellitaires Sentinel-2A, les orthophotos et le modèle numérique de hauteur dans une démarche multi-capteurs. Une classification «par objet» à l'aide du Random Forest nous permettra de cartographier les essences forestières de Paliseul. Les résultats seront ensuite validés sur base de données terrain. Et enfin, ceux-ci seront discutés et permettront d'envisager différentes perspectives.

2 Matériel et méthodes

La seconde partie de ce travail est dédiée à la présentation du matériel et des méthodes utilisées. Ce chapitre expose le site d'étude, les données utilisées, le traitement des données et la classification.

2.1 Site d'étude

Notre étude a été réalisée à l'échelle d'un territoire communal en zone forestière. Le choix s'est porté sur la commune de Paliseul, celle-ci faisant partie des sites pilotes du projet «Foret Pro Bos» dans lequel s'inscrit ce travail. Cette commune est située au sud-ouest de la province du Luxembourg, en basse et moyenne Ardenne (Figure 5).

²¹ <https://www.interreg-gr.eu/news/interreg-v-a-grande-region-21-nouveaux-projets-transfrontaliers-pour-la-grande-region/> (consulté le 9 mai 2017)

²² <https://foretprobos.wordpress.com/> (consulté le 9 mai 2017)

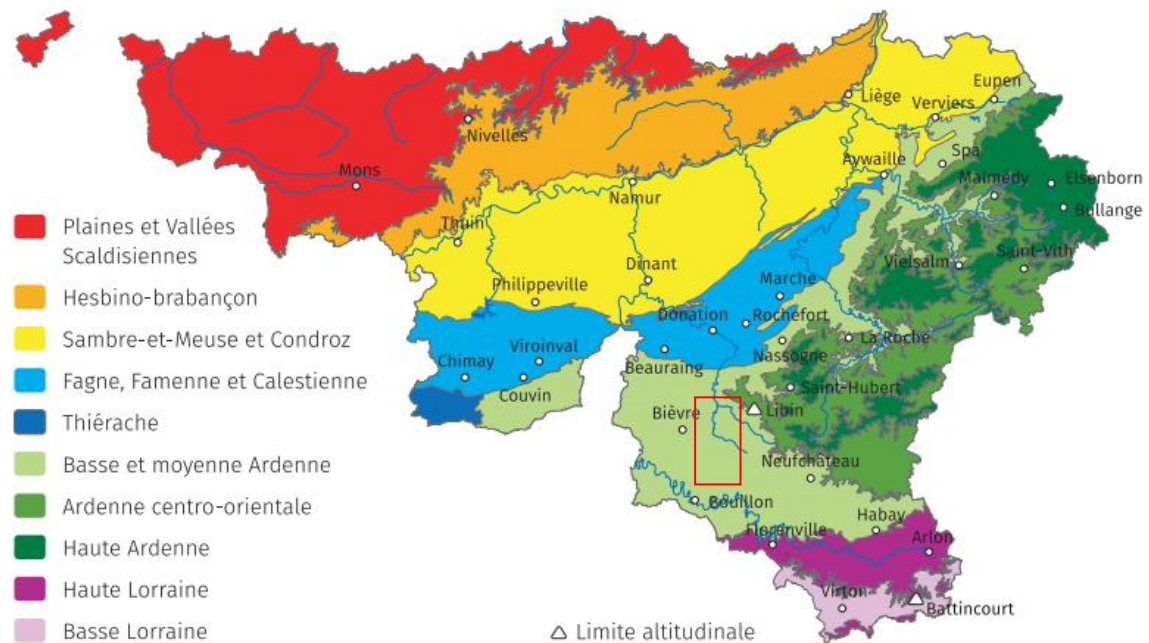


Figure 5: Carte des zones bioclimatiques de la Région wallonne et localisation de la zone d'étude ; source: Van der Perre *et al.*, 2015

Entre la Lesse et la Semois, Paliseul couvre une superficie de 11 295 ha²³. La forêt occupe près de la moitié du territoire communal. Environ 61% de cette forêt appartient à la commune, contre 37% à des propriétaires privés (IPRFW, 2017). Une vue aérienne sur la commune de Paliseul est disponible en annexe 1.

Très récemment, en avril 2017, Our, situé sur la commune de Paliseul, a été labellisé un des « Plus beaux villages de Wallonie »²⁴. Pour obtenir cette reconnaissance, les villages doivent répondre à une série de critères d'ordre paysager, urbanistique, architectural, ou encore d'accueil et d'animation (RTBF, 2017)²⁵. La commune a donc un intérêt à gérer durablement ses forêts et à disposer d'un outil de cartographie de son patrimoine forestier, mis à jour et dynamique.

²³ <http://www.si-paliseul.be> (consulté le 1 mars 2017)

²⁴ <http://www.beauxvillages.be/les-villages/our.htm?lng=fr> (consulté le 3 mai 2017)

²⁵ https://www.rtbef.be/info/regions/detail_montignies-sur-roc-et-our-nouveaux-plus-beaux-villages-de-wallonie?id=9587169 (consulté le 4 mai 2017)

2.2 Données de télédétection utilisées

Il est nécessaire de bien cibler les besoins de l'étude pour utiliser les données de résolution adéquate. De plus, l'accès à l'information dépend de l'origine des données et du cadre spécifique du travail. Pour notre étude cartographique de la forêt à l'échelle d'une commune, les produits de télédétection utilisés doivent être à haute résolution.

2.2.1 Sentinel-2A

Le site officiel de mise à disposition des données Sentinel-2 en libre accès est le HUB de l'agence spatiale européenne (ESA)²⁶. Au moment de la réalisation de ce travail, les images Sentinel-2A y sont distribuées sous la forme de scènes complètes (taille des fichiers de l'ordre de 7 Go) et en niveau de traitement L1C. Cela signifie qu'elles ont une correction radiométrique de type "*Top of atmosphere*" (TOA) et une correction géométrique de type ortho-rectification. Les images sont fournies en système de coordonnées projeté: UTM WGS84 ("epsg: 32631").

Le site Earth Explorer²⁷ permet également de télécharger les mêmes données mais sous la forme de tuiles de 100 km de côté et de manière plus simple et plus rapide. Depuis le mois de mai, le site THEIA²⁸ distribue pour la France et les régions limitrophes des images Sentinel-2 en niveau de traitement L2A. La version "*Surface Reflectance*" (SRE) de ce niveau de traitement reprend une correction radiométrique atmosphérique de type "*Top of canopy*" (TOC) et une ortho-rectification alors que la version "*Flat Reflectance*" (FRE) comporte une correction géométrique supplémentaire intégrant notamment une correction topographique.

Une limite principale à l'exploitation de l'information radiométrique des images satellite est la présence de couverture nuageuse. De février 2016 à février 2017, les images Sentinel-2 ne présentant ni nuages ni brouillard sur la zone d'étude ont été sélectionnées.

Dans la perspective d'une approche multi-dates, il est nécessaire de sélectionner les dates les plus pertinentes en termes de discrimination d'essences. Cela peut correspondre à des dates clés dans les stades phénologiques. Sur base de ce principe et des données disponibles, deux dates ont été retenues pour Paliseul: le 8 mai 2016 et le 20 juillet 2016 (en version FRE) (Figure 6). La première date tombe en tout début de végétation, la seconde en pleine période de végétation.

²⁶ <https://scihub.copernicus.eu/> (consulté en février 2017)

²⁷ <https://earthexplorer.usgs.gov/> (consulté en février 2017)

²⁸ <https://theia.cnes.fr/atdistrib/rocket/#/search?q=belgique> (consulté le 23 mars 2017)



Figure 6: Extraits des informations Sentinel-2A sur la zone d'étude pour les deux dates sélectionnées: 8 mai 2016 (au-dessus) et 20 juillet 2016 (en-dessous) à une résolution spatiale de 10 m et selon un niveau de correction L2A (FRE) ; source: ESA (corrigé par THEIA)

2.2.2 Ortho-images

Le Service Public de Wallonie (SPW) dispose de plusieurs couches ortho-images acquises à intervalle régulier (tous les trois ans depuis 2006). Ces couches résultent de l'ortho-rectification et du mosaïquage d'images aériennes comportant 4 bandes spectrales (rouge, vert, bleu et proche infrarouge). La résolution finale des ortho-images est de 25 cm, sauf pour la couverture 2006 où elle est de 50 cm. Les différentes couches ortho-images se présentent sous la forme de tuiles carrées de

2 km de côté et sont fournies en système de coordonnées projeté: Belge Lambert 72 ("epsg: 31370"). (Géoportail de la Wallonie, 2017)²⁹

La finalité première de ces images étant le contrôle agricole, le cahier des charges permet que les dates d'acquisition se situent avant la date de débourrement des essences feuillues (Politique Agricole Commune).

Le Tableau 1 présente les dates de survol de la commune de Paliseul pour les 3 dernières campagnes aériennes d'acquisition d'ortho-images.

Tableau 1: Résumé des dates d'acquisition des ortho-images sur la Région Wallonne et sur la commune de Paliseul ; source: Géoportail de la Wallonie, 2017

Orthophoto	Début de vol	Fin de vol	Date d'acquisition pour Paliseul
2009-2010	23-mai-09	07-juil-10	15-août-09
2012-2013	14-mai-12	08-juil-13	26-mai-12
2015	08-avr-15	17-juin-15	10-avr-15 & 15-avr-15

Ces images multi-spectrales présentent une variabilité radiométrique importante influencée notamment par les angles de prise de vue, l'orientation de la source d'éclairage et la présence de nuages. Ceci rend difficile l'utilisation de l'information spectrale des ortho-images dans une démarche de classification automatisée et à large échelle (Figure 7) (Dare, 2005).

²⁹ www.geoportail.wallonie.be/georeferentiel/orthophotos (consulté le 22 avril 2017)

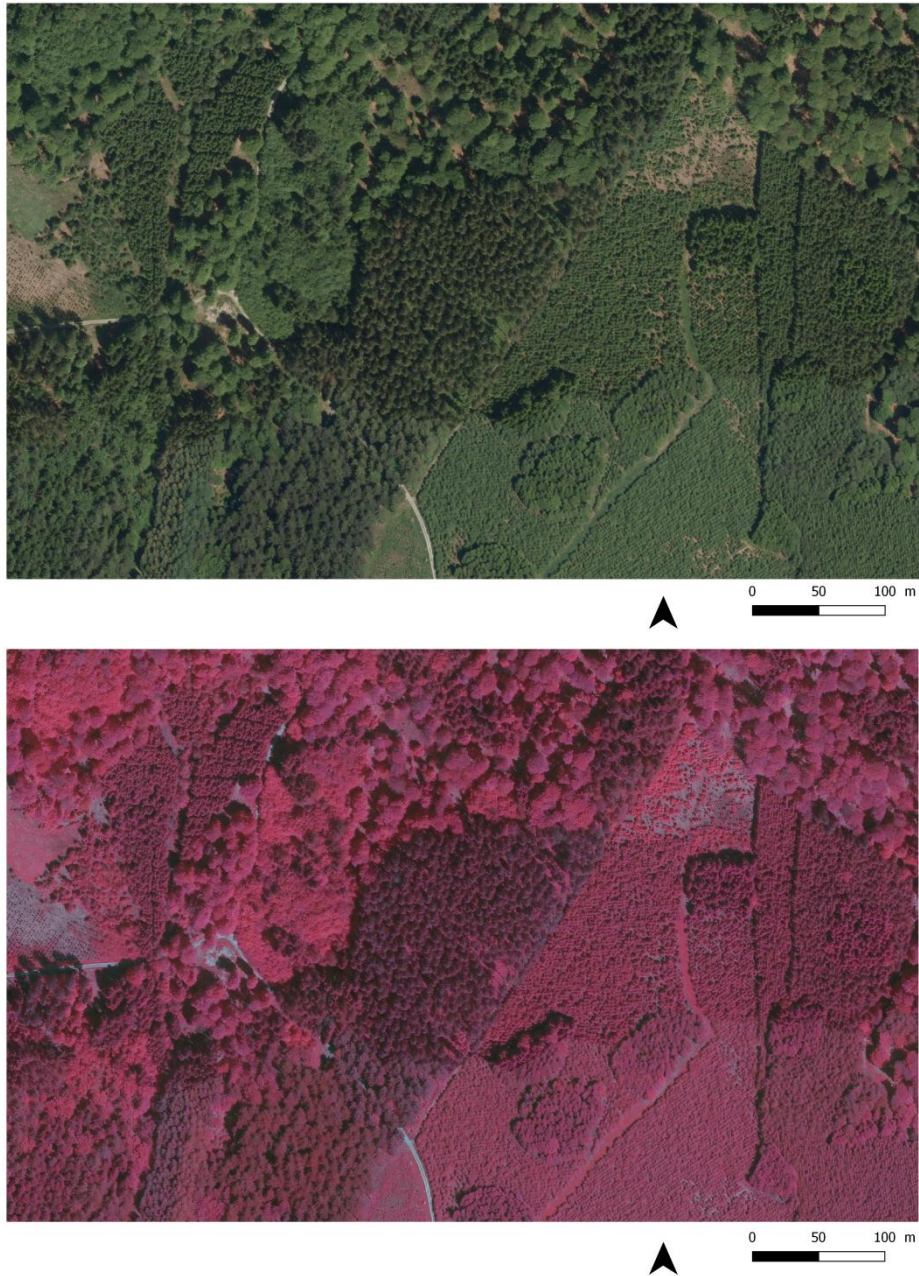


Figure 7: Extrait de la couche ortho-image 2012-2013 sur la zone d'étude. Une version vraie couleur est présentée au-dessus et une version fausse couleur en-dessous, mettant en évidence la végétation à une résolution spatiale de 25 cm et les problèmes de variabilité radiométrique; source: SPW

L'acquisition des données doit avoir lieu, idéalement, pendant la période de végétation lorsque les arbres ont leur feuillage bien développé. Malheureusement, les orthophotos de 2015 ne sont pas exploitables pour l'étude des essences puisqu'elles n'ont pas été prises à la bonne période. En outre, afin de limiter les ombres, les photos devraient être réalisées au zénith du soleil. Ce dernier critère n'est pour ainsi dire jamais respecté.

Afin de limiter le décalage temporel entre les différentes sources de données, il est justifié d'utiliser les orthophotos exploitables les plus récentes, à savoir celles de 2012-2013 (Figure 7). Les orthophotos ont soit été reprises physiquement du serveur cartographique du Département BIOSE³⁰

³⁰ Le département BIOSE (Ingénierie des biosystèmes) de Gembloux Agro-Bio Tech (ULg) comprend notamment l'Axe Gestion des Ressources forestières

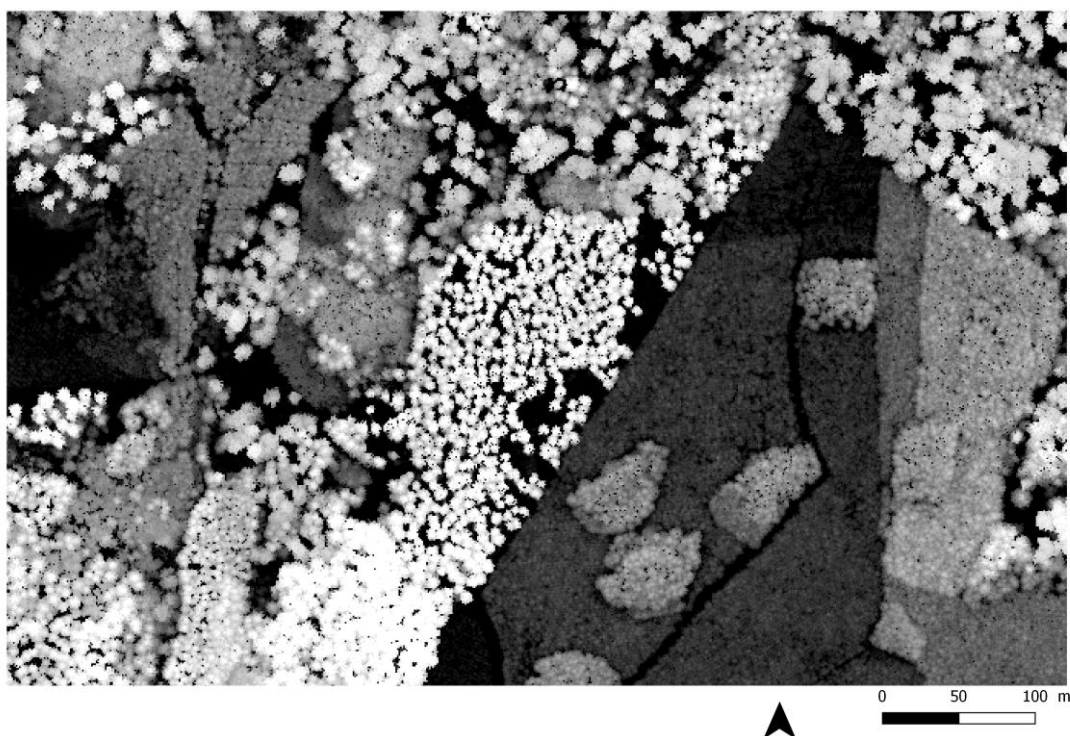
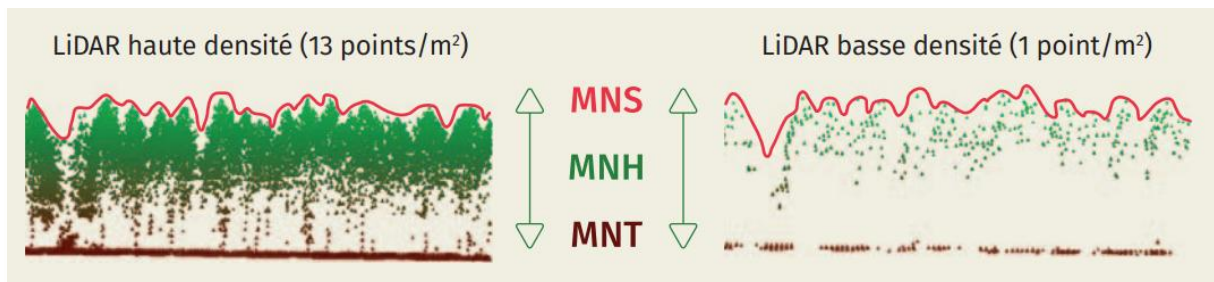
par tuile de 2 km sur 2 km, soit été visualisées via le Web Map Service à partir du géoportail du service public³¹.

2.2.3 Modèle Numérique de Hauteur (MNH)

Le Département de Géomatique du Service Public de Wallonie a commandé une couverture LiDAR aérien à basse densité (1 point/m²) sur toute la Région wallonne en 2013-2014. Les couches LiDAR sont fournies en système de coordonnées projeté: Belge Lambert 72 ("epsg: 31370").

Des couches Modèle Numérique de Terrain (MNT) et Modèle Numérique de Surface (MNS) à un mètre de résolution ont été produites par le prestataire (Figure 8). Un Modèle Numérique de Hauteur (MNH) a pu être calculé par soustraction de ces deux couches (Figure 8). Ce MNH était également disponible sur le serveur cartographique de BIOSE.

Le MNH en zone forestière permet de détecter les houppiers (Figure 9). Son utilisation permet d'améliorer la précision des inventaires (superficie et volume) et de la cartographie (Renaud *et al.*, 2015).



³¹ <http://geoportail.wallonie.be> (consulté de février à mai 2017)

2.3 Traitement des données

2.3.1 Introduction

La démarche adoptée pour exploiter les données présentées au point 2.2 et atteindre nos objectifs de cartographie est présentée de manière globale dans la Figure 10. L'annexe 2 reprend le schéma des étapes 2.3 et 2.4.

Plusieurs étapes ont été mises en œuvre, les principales sont:

- la préparation des données provenant de différentes sources ;
 - ↳ la reprojection des couches dans un système de coordonnées projeté commun ;
 - ↳ le découpage des couches aux limites de la commune ;
 - ↳ le rééchantillonnage des couches à des résolutions spatiales définies.
- la segmentation des images sur base des données à très haute résolution spatiale (les orthophotos et le MNH) ;
- la caractérisation des objets (segments produits à l'étape précédente) ;
 - ↳ la création d'indices « mono-date » et « multi-dates » dérivés des bandes spectrales de Sentinel-2A ;
 - ↳ la combinaison des bandes spectrales de Sentinel-2A, des indices et du MNH afin d'en calculer des variables descriptives (moyenne par objet).

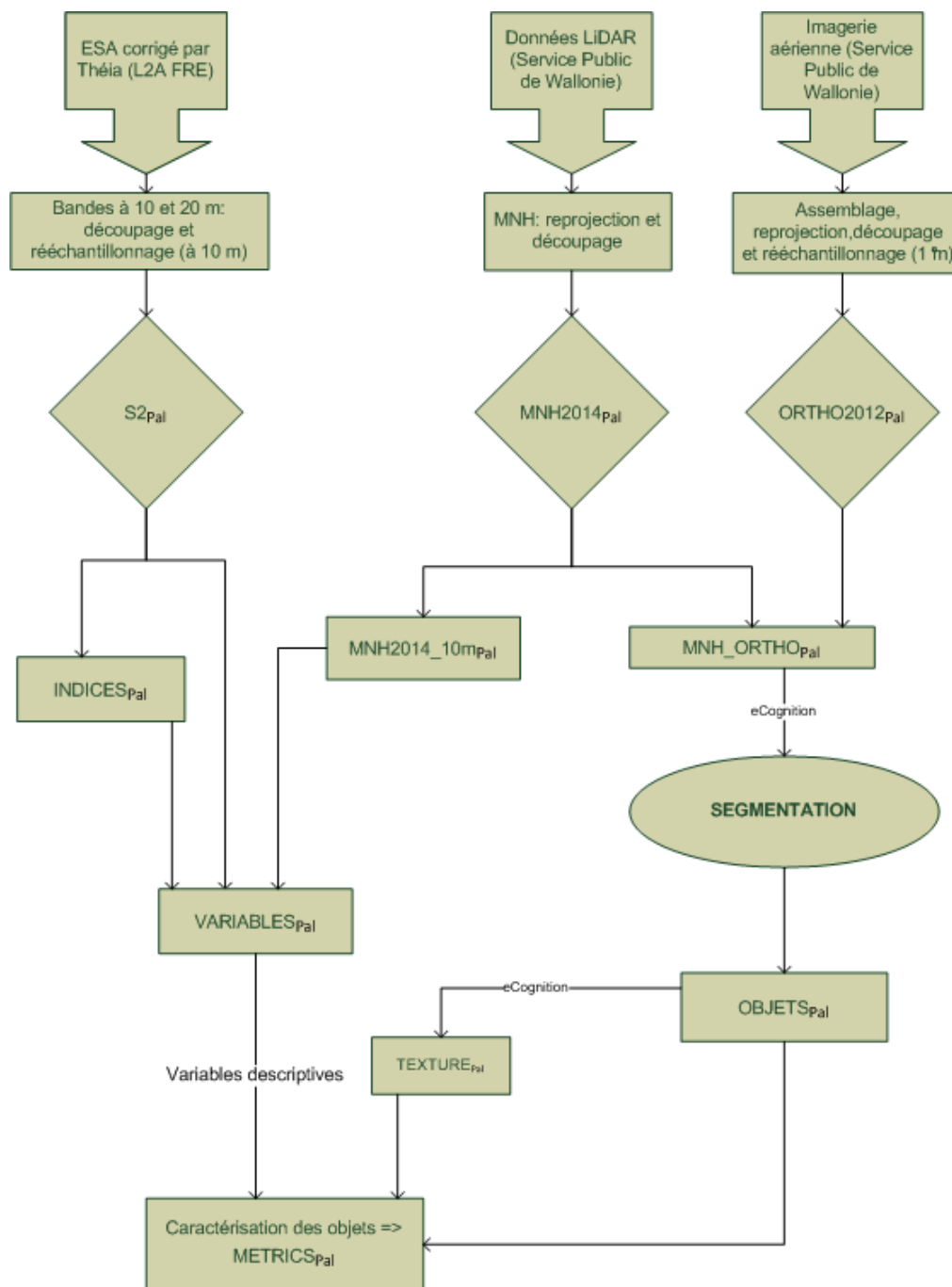


Figure 10: Schéma de la chaîne de traitement des données

Les logiciels utilisés sont présentés dans le Tableau 2.

Tableau 2: Outils logiciels utilisés dans le cadre de ce travail

Logiciel	Utilisation	Source	Accès	Informations diverses
Qgis	SIG	OSGeo ³²	Libre	
ArcGIS	SIG	ESRI ³³	Commercial	
Orfeo ToolBox (OTB)	Segmentation	OSGeo ³⁴	Libre	Développé par le CNES outil qui nécessite une formation de l'utilisateur
eCognition	Segmentation	Trimble Geospatial	Commercial	Algorithmes propriétaire («Boite noire»)
R	Ecriture de script en vue de créer des chaines de traitement	R Foundation	Libre	La communauté des utilisateurs de R partagent des informations, scripts et packages (via CRAN), notamment en matière d'algorithmes de traitement d'images.

2.3.2 Préparation des données

- MNH

Le MNH (créé par soustraction des MNS et MNT) a été reprojété dans le système de coordonnées WGS84 UTM31N ("epsg: 32631") puis découpé aux limites de la commune de Paliseul. Le résultat est stocké dans une couche baptisée: MNH2014_{Pal}.

- Ortho-images

Les tuiles de la couche orthophotos de 2012 couvrant le territoire de la commune de Paliseul ont été assemblées en une seule couche multi-bandes qui a été reprojétée en WGS84 UTM31N ("epsg: 32631"). Cette couche a ensuite été rééchantillonnée à une résolution de 1 m à l'aide de la commande Aggregate (opérateur « mean ») du package « Raster » de R en considérant l'emprise spatiale de la couche MNH2014_{Pal}. Le résultat est stocké dans un raster baptisé ORTHO2012_{Pal}.

Les couches MNH2014_{Pal} et ORTHO2012_{Pal} ont finalement été assemblées pour former un raster comportant 5 bandes à 1 m de résolution (MNH_ORTHO_{Pal}) qui sera utilisé dans l'étape de segmentation.

- Sentinel-2A

L'ensemble des bandes spectrales à 10 m et 20 m de résolution pour les deux dates a été rééchantillonné à 10 m et empilé en un raster multicouche aux limites de la commune (S2_{Pal}).

Le MNH2014_{Pal} a été rééchantillonné à une résolution de 10m en considérant l'emprise spatiale de S2_{Pal} (création d'une couche baptisée MNH2014_10m_{Pal}).

2.3.3 Création d'indices spectraux

Afin de mettre en évidence des spécificités de réponses spectrales en fonction de l'occupation du sol, on peut exploiter l'information synthétisée de rapports et de différences des bandes spectrales. Divers indices spectraux ont été calculés pour les deux dates des données Sentinel-2A sur base de l'analyse de divers articles traitant de l'utilisation d'indices pour la cartographie de la végétation. Ils ont ensuite été compilés dans une couche nommée INDICES_{Pal}.

³² https://live.osgeo.org/fr/overview/qgis_overview.html (consulté le 23 avril 2017)

³³ <http://www.esri.com/arcgis/about-arcgis> (consulté le 22 avril 2017)

³⁴ <https://www.orfeo-toolbox.org/> (consulté le 23 avril 2017)

$NDVI = \frac{NearInfraRed - Red}{NearInfraRed + Red}$	Indice 1 : Normalized Difference Vegetation Index (Tucker, 1979)
$GI = \frac{Green}{Red}$	Indice 2: Greenness Index (le Maire <i>et al.</i> , 2004)
$ChloGreen = \frac{NearInfraRed}{Green + Rededge1}$	Indice 3: Chlorophyll Green Index (Datt, 1999)
$gNDVI = \frac{NearInfraRed - Green}{NearInfraRed + Green}$	Indice 4: Green Normalized Difference Vegetation Index (Gitelson <i>et al.</i> , 1996)
$BAI = \frac{Blue - NearInfraRed}{Blue + NearInfraRed}$	Indice 5: Built-up Area Index (Shahi <i>et al.</i> , 2015)
$LAnthoC = \frac{Rededge3}{Green - Rededge1}$	Indice 6: Leaf Anthocyanid Content (Wulf et Stuhler, 2015)
$LCaroC = \frac{Rededge3}{Blue - Rededge1}$	Indice 7: Leaf Cartotenoid Content (Wulf et Stuhler, 2015)
$LChloC = \frac{Rededge3}{Rededge1}$	Indice 8: Leaf Chlorophyll Content (Wulf et Stuhler, 2015)
$MSI = \frac{SWIR1}{NearInfraRed}$	Indice 9: Moisture Stress Index (Vogelmann et Rock, 1985)
$NDrededgeSWIR = \frac{Rededge2 - SWIR2}{Rededge2 + SWIR2}$	Indice 10: Normalised Difference of Red-edge and SWIR2 (Radoux <i>et al.</i> , 2016)
$NDTI = \frac{SWIR1 - SWIR2}{SWIR1 + SWIR2}$	Indice 11: Normalized Difference Tillage Index (Van Deventer <i>et al.</i> , 1997)
$NDVire = \frac{NearInfraRed - Rededge1}{NearInfraRed + Rededge1}$	Indice 12: Red-edge Normalized Difference Vegetation Index (Gitelson <i>et al.</i> , 1996)
$NDWI1 = \frac{NearInfraRed - SWIR1}{NearInfraRed + SWIR1}$	Indice 13: Normalized Difference Water Index 1 (Gao, 1996)
$NDWI2 = \frac{Green - NearInfraRed}{Green + NearInfraRed}$	Indice 14: Normalized Difference Water Index 2 (Gitelson <i>et al.</i> , 1996)
$NHI = \frac{SWIR1 - Green}{SWIR1 + Green}$	Indice 15: Normalized Humidity Index (Lacaux <i>et al.</i> , 2007)
$RPA = Red + Rededge1 + Rededge2 + Rededge3 + NearInfraRed$	Indice 16: Red-edge Peak Area (Radoux <i>et al.</i> , 2016 ; Filella et Penuelas, 1994)
$RS = Red - SWIR1$	Indice 17: Red SWIR1 (Jacques <i>et al.</i> , 2014)
$RTVcore = 100 * (NearInfraRed - Rededge1) - 10 * (NearInfraRed - Green)$	Indice 18: Red-edge Triangular Vegetation Index (Chen <i>et al.</i> , 2010)
$SAVI = \frac{NearInfraRed - Red}{NearInfraRed + Red + 0,5} * 1,5$	Indice 19: Soil Adjusted Vegetation Index (Huete, 1988)
$BR1 = \frac{Blue}{Rededge1}$	Indice 20: Simple Blue and Red-edge 1 Ratio (le Maire <i>et al.</i> , 2004)
$BR2 = \frac{Blue}{Rededge2}$	Indice 21: Simple Blue and Red-edge 2 Ratio (Lichtenthaler <i>et al.</i> , 1996)
$BR3 = \frac{Blue}{Rededge3}$	Indice 22: Simple Blue and Red-edge 3 Ratio (Radoux <i>et al.</i> , 2016)
$NB = \frac{NearInfraRed}{Blue}$	Indice 23: Simple Ratio Nearinfrared and Blue (Blackburn, 1998)
$NG = \frac{NearInfraRed}{Green}$	Indice 24: Simple Ratio Nearinfrared and Green (le Maire <i>et al.</i> , 2004)
$NR = \frac{NearInfraRed}{Red}$	Indice 25: Simple Ratio Nearinfrared and Red (Blackburn, 1998)
$NR1 = \frac{NearInfraRed}{Rededge1}$	Indice 26: Simple Nearinfrared and Rededge1 Ration (Datt, 1999)
$NR2 = \frac{NearInfraRed}{Rededge2}$	Indice 27: Simple Nearinfrared and Rededge2 Ratio (Radoux <i>et al.</i> , 2016)

$$NR3 = \frac{NearInfraRed}{Rededge3} \quad \text{Indice 28: Simple Nearinfrared and Redededge3 Ratio (Radoux et al., 2016)}$$

$$STI = \frac{SWIR1}{SWIR2} \quad \text{Indice 29: Soil Tillage Index (Van Deventer et al., 1997)}$$

$$WBI = \frac{Blue-Red}{Blue+Red} \quad \text{Indice 30: Water Band Index (Domenech et Mallet, 2014)}$$

$$GEMI = n * (1 - 0,25 * n) - \frac{Red-0,125}{1-Red} \quad \text{où } n = \frac{2*(NearInfraRed^2 - Red^2) + 1,5*NearInfraRed + 0,5*Red}{NearInfraRed + Red + 0,5}$$

Indice 31: Global Environmental Monitoring Index (Pinty et Verstraete, 1992)

$$norm_G = \frac{Green}{NearInfraRed + Red + Green} \quad \text{Indice 32: Normalised Green (Sripada et al., 2006)}$$

$$norm_{NIR} = \frac{NearInfraRed}{NearInfraRed + Red + Green} \quad \text{Indice 33: Normalised Infrared (Sripada et al., 2006)}$$

$$norm_R = \frac{Red}{NearInfraRed + Red + Green} \quad \text{Indice 34: Normalised Red (Sripada et al., 2006)}$$

En outre, pour tester la pertinence des indices « multi-dates » pour traduire la dynamique de la végétation, nous avons considéré deux indices supplémentaires, correspondant respectivement à la variation du NDVI entre les deux dates (*deltaNDVI*) et la moyenne des NDVI des deux dates (*meanNDVI*) :

$$deltaNDVI = NDVI_{D2} - NDVI_{D1} \quad \text{Indice 35: NDVI Difference}$$

$$meanNDVI = \frac{NDVI_{D1} + NDVI_{D2}}{2} \quad \text{Indice 36: NDVI Mean}$$

Les couches $S2_{pal}$, $MNH2014_10m_{pal}$ et $INDICES_{pal}$ ont finalement été assemblées pour former un raster baptisé $VARIABLES_{pal}$.

2.3.4 Création des segments

La couche d'entrée de la segmentation est MNH_ORTHO_{pal} .

eCognition est un logiciel très répandu pour la segmentation d'image (Benz et al. 2004, Wang L. et al. 2004). L'algorithme utilisé, «*Multiresolution segmentation*», est basé sur l'agrégation de pixels «*bottom-up*» ou «*region growing*». Il associe entre eux des petits objets homogènes pour en former de plus grands et recalcule d'homogénéité de ce nouvel objet par rapport à ses voisins. L'algorithme continue l'agrégation jusqu'à rencontrer les critères définis par l'utilisateur: des propriétés de couleurs (valeurs spectrales), de forme et d'échelle (Darwish et al., unknown).

La segmentation multi-échelle proposée par eCognition permet également d'attribuer un poids aux différentes bandes en entrée. Ceci permet de donner plus d'importance à certaines gammes spectrales ou à la hauteur.

L'évaluation de la segmentation se fait en regard des objectifs. Dans le cadre de ce travail, il faut des segments qui permettent de mettre en évidence un à plusieurs houppiers d'une même essence. Ainsi, différentes combinaisons ont été testées de façon non exhaustive. Les paramètres de segmentation pour Paliseul ont été retenus sur base d'essais et d'erreurs et d'interprétations visuelles sur une tuile test (2 km sur 2 km) (Figure 11).

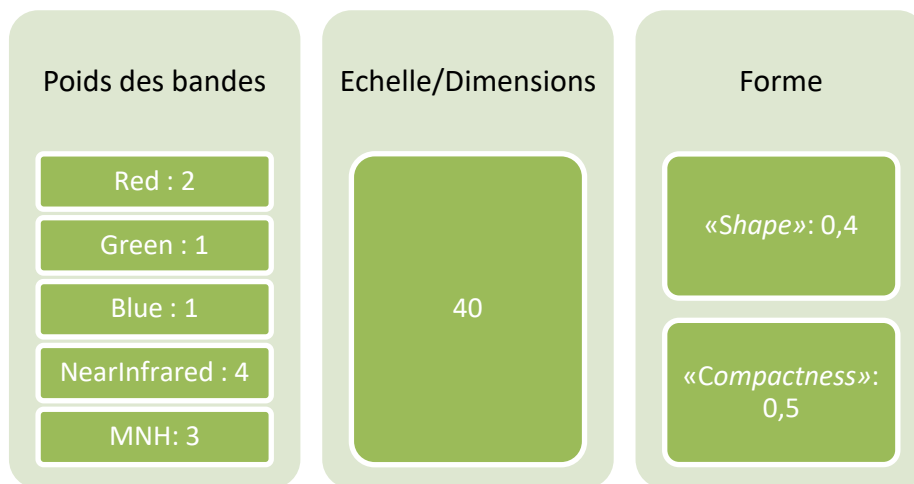


Figure 11: Paramètres retenus pour la segmentation de Paliseul dans eCognition sur base d'essais et erreurs ainsi qu'une interprétation visuelle des résultats

Délimitées aux limites de la commune, la couche *shapefile* de segments sera utilisée pour les étapes suivantes et est baptisée OBJETS_{Pal}.

2.3.5 Calcul des variables descriptives (*metrics*) des objets

Une fois les segments créés, leur caractérisation se fait sur base des informations compilées des pixels.

Des informations sur la valeur spectrale moyenne des objets ainsi que des paramètres de distribution (écart-type) sont calculables. Une procédure de calcul et d'extraction des différentes variables descriptives des objets (*metrics*) a été développée dans R. Les «Moyennes» de la couche VARIABLES_{Pal} ont été extraites par segment (OBJETS_{Pal}).

La texture des objets peut également être étudiée. Elle peut être définie comme la «répétition spatiale d'un même motif dans différentes directions de l'espace»³⁵ ou «quantification de la variation spatiale des tons d'une image qui échappe à toute définition précise du fait qu'elle soit liée à la perception (Hay *et al.*, 1996)». Elle peut être calculée directement par segment lors de l'export du *shapefile* avec eCognition (TEXTURE_{Pal}). La méthode utilisée est GLCM *Entropy*³⁶ sur la bande proche infrarouge de l'orthophoto (MNH_ORTHO_{Pal}).

Les variables descriptives et la texture composent la couche METRICS_{Pal}.

³⁵ http://www.tsi.telecom-paristech.fr/pages/enseignement/ressources/beti/textures_Fourier/texture.html (consulté le 25 mai 2017)

³⁶ Le GLCM (*Gray Level Co-occurrence Matrix*) est une méthode qui permet d'extraire quatre paramètres statistiques de texture différents. L'entropie, la corrélation, l'homogénéité³⁶ et l'uniformité³⁶. L'entropie indique la quantité d'information nécessaire pour la compression. Elle informe par la même occasion de l'information contenue dans l'image transmise et de la perte d'information (Mohanaiah *et al.*, 2013).

2.4 Classification

2.4.1 Introduction

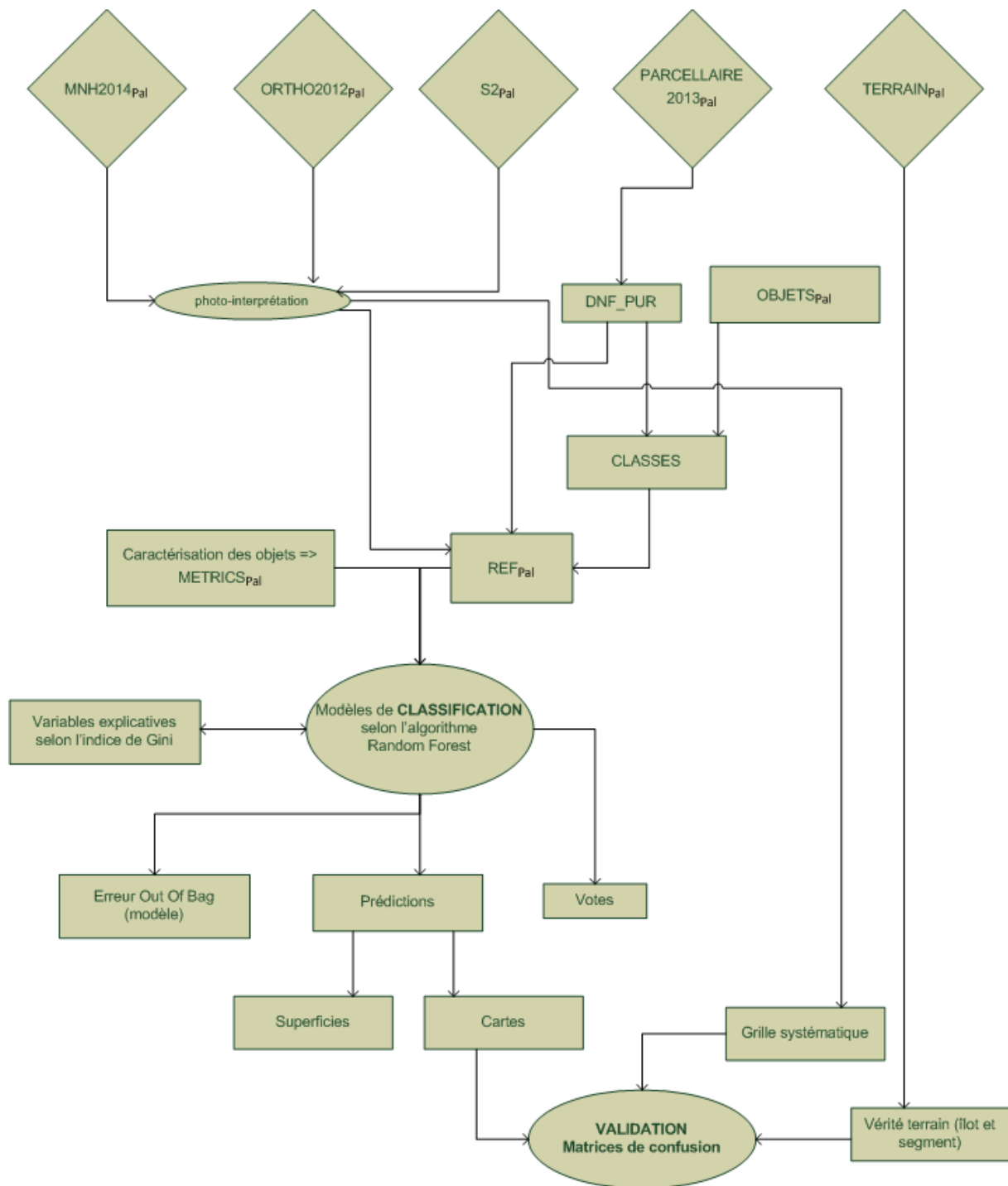


Figure 12: Schéma des étapes de la classification supervisée

En cartographie de l'occupation du sol, la classification est une étape automatisée qui permet de traduire des informations spectrales sous la forme de classes thématiques basées sur la vérité terrain.

La classification supervisée comporte quatre grandes étapes: la définition de la légende, la création de données d'entraînement, la création et application du modèle de classification et la validation (Figure 12).

2.4.2 Définition de la légende cartographique

Le choix des classes est la première étape à considérer dans une classification. Ces classes doivent, autant que possible, répondre à l'objectif poursuivi. Elles doivent en outre être mutuellement exclusives et exhaustives (Congalton et Green, 1999). Le degré de détail considéré doit être en adéquation avec les capacités de différenciation offertes par les données utilisées dans la classification. Les classes doivent également être clairement définies³⁷.

Les classes qui ont été utilisées dans les différents modèles de classification sont reprises dans la Figure 13 (CLASSES) et sont définies dans le Tableau 3.

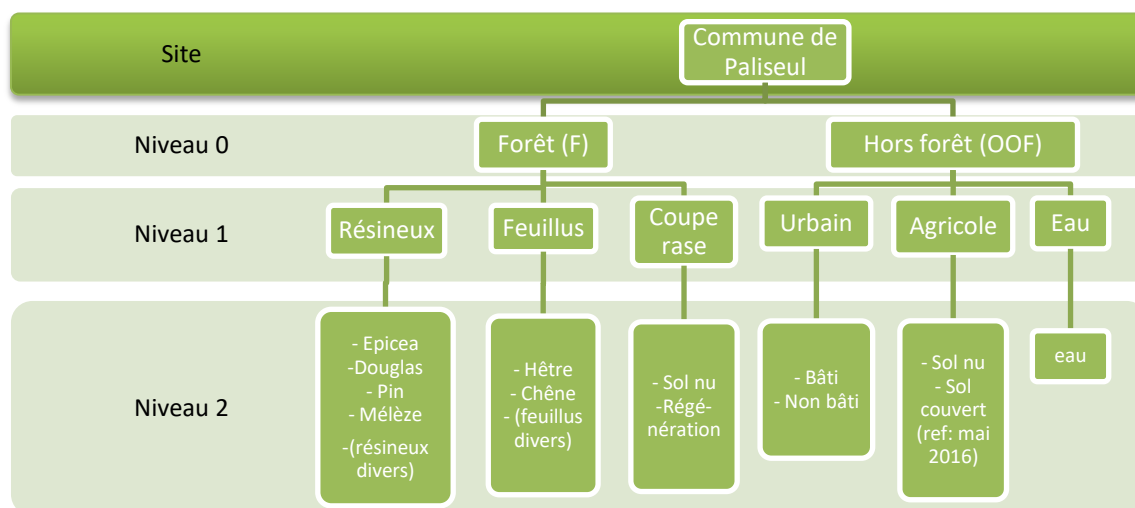


Figure 13: Choix des classes qui seront utilisées dans le modèle de classification selon un concept d'objets hiérarchisés

Tableau 3: Définition des différentes classes utilisées pour la classification

Classe	Définition
Forêt	Zone végétalisée d'une superficie de plus de 0,5 ha et d'une hauteur de plus de 7 m ou zone récemment exploitée
Hors forêt	Zone non forestière correspondant au milieu urbain artificialisé, aux zones agricoles et aux pièces d'eau
Résineux	Zone forestière occupée par des essences résineuses (conifères), constituée de peuplements "purs" selon le DNF
Feuillus	Zone forestière occupée par des essences feuillues, constituée de peuplements "purs" selon le DNF
Coupe rase/Mise à blanc	Zone forestière pour laquelle le peuplement a été récemment exploité. La zone peut être du sol nu ou en cours de re-végétalisation (régénération naturelle ou plantation)

Les mélanges ne sont pas abordés explicitement en tant que classe à part entière dans le cadre de ce travail. Le modèle de classification est entraîné uniquement sur des classes d'essences pures, les peuplements mélangés ne pourront donc pas être détectés dans la cartographie.

³⁷ https://servirglobal.net/Portals/0/Documents/Articles/ChangeDetectionTraining/Module3_LC_Classification_presentatio n.pdf (consulté le 28 mai 2017)

Par ailleurs, nous avons considéré deux manières pour aborder le cas des peuplements mélangés :

- les votes dans l'approche Random Forest «par objet» qui évaluent, par segment, les probabilités d'appartenance à chaque classe. Le but est de tester s'il est possible de discriminer les essences en mélange en fonction des décisions principales des arbres du Random Forest, et pas uniquement du choix majoritaire ;
- la réalisation d'une classification supervisée «par pixel» sur les données raster à 10 m de résolution. On peut imaginer que la classification indépendante des pixels à l'échelle de 10 m pourrait mettre en évidence les différentes essences en mélange.

2.4.3 Création des données d'entraînement

La classification supervisée requiert des données de référence pour entraîner le modèle de classification.

- Source de données d'entraînement

Afin de créer un jeu de données de référence, certains segments ont été sélectionnés sur base du parcellaire DNF et classés par photo-interprétation.

Le parcellaire du DNF³⁸ (Figure 14) est la couche cartographique de base utilisée par l'administration forestière de la Région wallonne dans le cadre de la préparation des plans d'aménagement et dans ses missions de gestion quotidienne des forêts publiques. Elle contient la délimitation des « îlots » constitutifs des forêts publiques. L'îlot est une unité spatiale élémentaire, homogène quant à l'habitat et au peuplement forestier (en zone forestière). La table d'attributs renseigne notamment sur les essences présentes et leur importance relative au sein de l'îlot et leur date de plantation. La date de référence de la couche que nous avons utilisée est 2013.

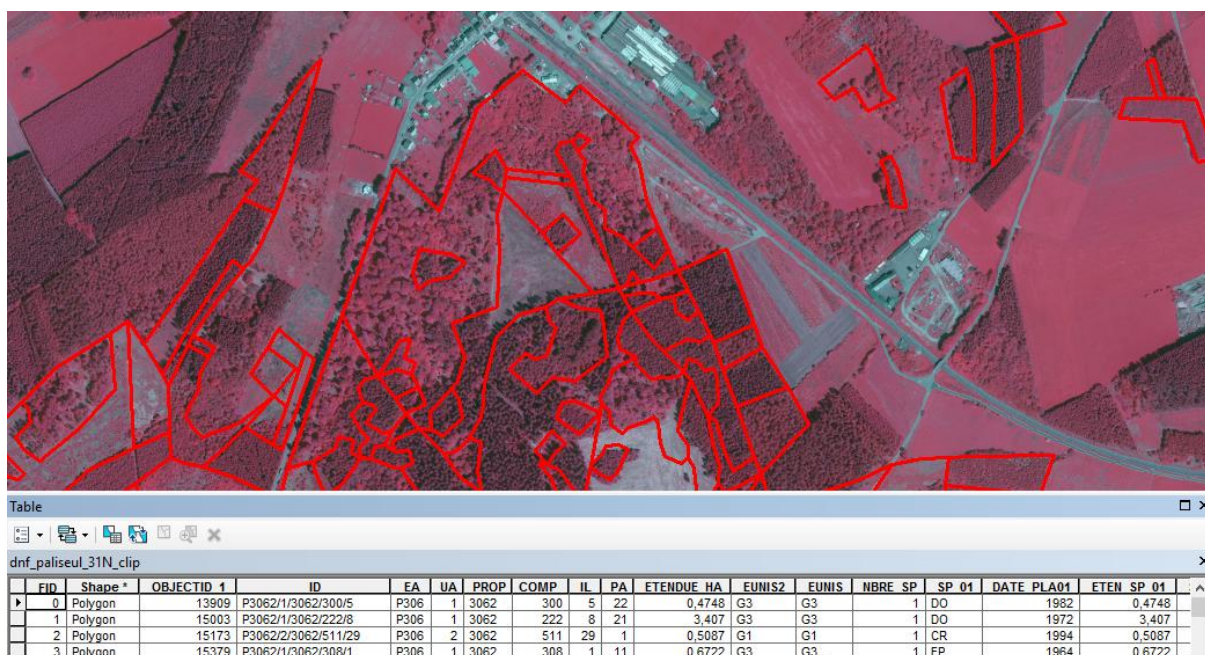


Figure 14: Extrait du parcellaire du Département de la Nature et des Forêts sur la zone d'étude. La délimitation des îlots est représentée par les polygones rouges pour l'année 2013 ; source: SPW

La qualité tant géométrique (Figure 15) qu'attributaire de cette couche est variable. Sans avoir pu réaliser une vérification complète des données pour le territoire de Paliseul, les quelques vérifications que nous avons réalisées sur le terrain et sur base de photo-interprétations nous ont permis de mettre en évidence un certain nombre d'erreurs.

³⁸ Le Département de la Nature et des Forêts fait partie du DGO3, Direction Générale Opérationnelle, Agriculture, Ressources naturelles et Environnement, du Service Public de Wallonie.

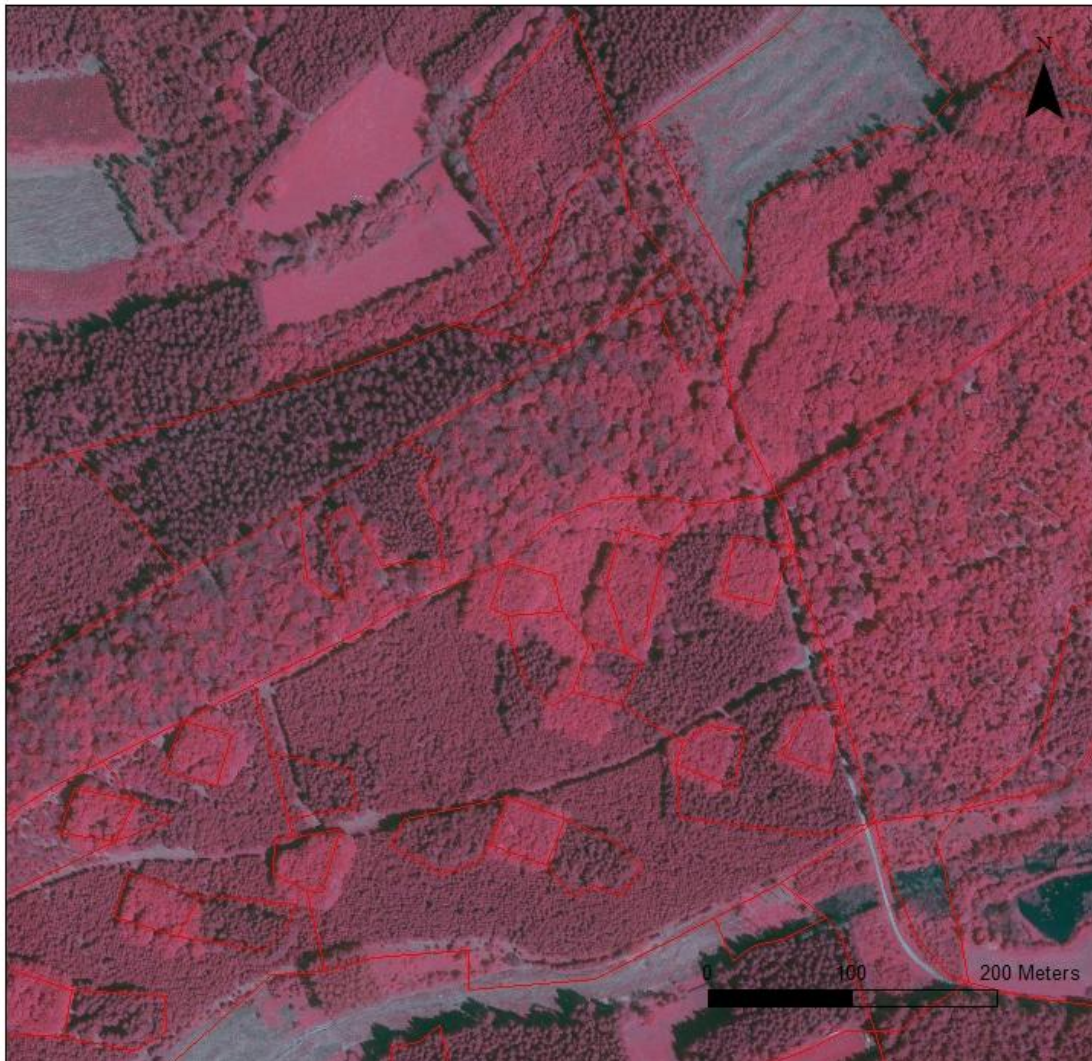


Figure 15: Illustration des imprécisions du parcellaire DNF. La délimitation des îlots est représentée par les polygones rouges pour l'année 2013 ; source: SPW

- Création des données d'entraînement

Les segments d'entraînement ont été sectionnés dans les peuplements «purs»³⁹ référencés par le DNF (DNF_PUR). Une attention toute particulière a été portée à la fiabilité des informations du parcellaire en favorisant les segments homogènes (sur base des orthophotos 2016, des orthophotos 2012-2013, des données Sentinel-2 des deux dates et du MNH) qui sont totalement à l'intérieur de l'îlot DNF.

Dans le but de la cartographie des peuplements forestiers, un minimum de 100 observations de référence par classe de niveau 2 ont été photo-interprétées. Nous obtenons ainsi 1 894 observations en forêt et 500 observations hors forêt (Figure 16). Ces informations constituent la couche d'entraînement, nommée REF_{pal}.

³⁹ Un îlot est considéré "pur" lorsqu'une seule essence est référencée dans le parcellaire du DNF (80%).

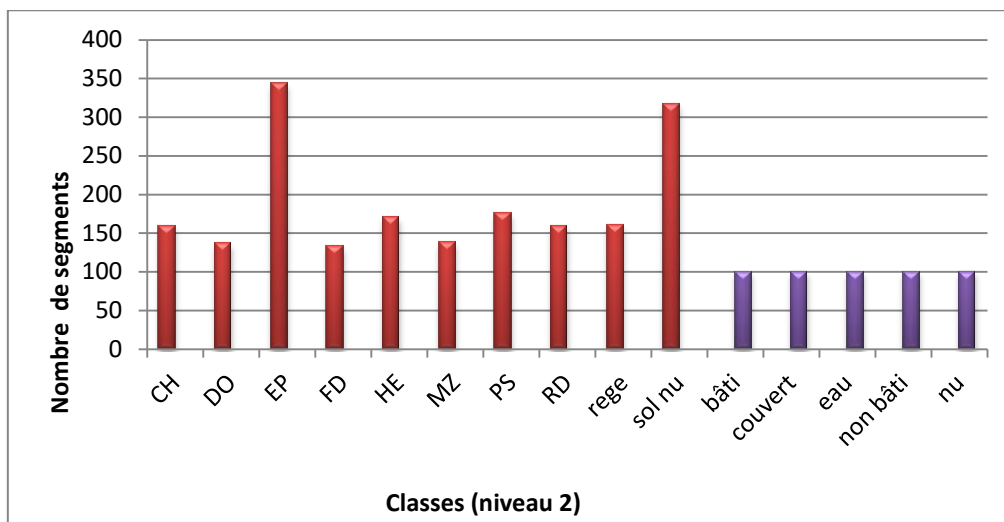


Figure 16: Nombre de segments de référence par classe pour l'entrainement des modèles de classification

2.4.4 Classification Random Forest

- Random Forest

L'algorithme de classification que nous avons utilisé est le Random Forest⁴⁰ proposé par Breiman (2001). L'algorithme crée n arbres associés chacun à un échantillon des données d'entrainement. Chaque arbre va ensuite sélectionner aléatoirement m variables et les associer de façon à créer le meilleur modèle. Enfin, sur de nouvelles données, chaque arbre pronostique une classe et le modèle prédit la classe de l'objet selon un vote à majorité. (Liaw et Wiener, 2002)

Les avantages du Random Forest sont, notamment :

- la puissance de l'algorithme même lorsque le nombre de variables m est élevé en regard du nombre d'observations d'entrainement;
- l'auto-évaluation de la précision du modèle, appelée précision «Out of bag» (OOB). Elle est calculée par chaque arbre à l'aide d'une matrice de confusion réalisée sur la portion de données d'entrainement non-utilisée pour la création du modèle ;
- la nature non-paramétrique du modèle, sa haute précision de classification et sa capacité à déterminer des importances de variables (Svetnik *et al.*, 2003 ; Waske et Braun, 2009).

Néanmoins, une limite du Random Forest est qu'il ne possède pas de classe «inconnu» ou «indéfini» et est forcé d'attribuer une classe à chaque observation sur base de la réponse spectrale la plus proche. De plus, il ne prédira que la classe majoritaire même si le risque de confusion est élevé.

- Modalités

Dans le cadre de ce travail, nous avons utilisé principalement la classification supervisée «par objet». L'objectif est une cartographie sur le niveau 2, à l'échelle des peuplements forestiers.

Diverses modalités ont été testées, de façon itérative par essai-erreur, lors de tests préliminaires. Un équilibre a été trouvé afin d'améliorer la qualité du modèle tout en limitant la complexité des traitements:

⁴⁰ package "randomForest" dans le logiciel R

Pour plus d'informations : https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm#workings (consulté le 18 mai 2017)

- les données d'entraînement (REF_{pai}) ont été équilibrées par le bas ;
- le nombre d'arbres (n) a été fixé à 2 000 ;
- les autres paramètres du Random Forest ont été utilisés par défaut ;
- les 30 meilleures variables selon l'indice de Gini maximal sont sélectionnées sur base d'un premier modèle Random Forest. Ensuite, le modèle final est créé en ne tenant compte que de cette sélection de variables.

L'indice d'importance de Gini se base d'une part sur la fréquence de sélection de la variable en question dans les arbres de décision aléatoires, et d'autre part sur la valeur discriminatoire totale de la variable (Menze et al., 2009). Certaines études montrent que la précision globale du modèle augmente lorsque les variables sont sélectionnées sur base de l'importance de l'indice de Gini (Menze et al., 2007 ; Jiang et al., 2009). D'autres concluent sur une précision non significativement différente mais mettent en avant le fait d'utiliser moins de variables en entrée du modèle (Li et al., 2005). Menze B. et al. (2009) suggèrent tout de même que la technique est une approche qui peut précéder une autre méthode de sélection de variables.

Suite à cela, deux modèles de classification « par objet » ont été retenus.

↳ Modèle A

Le modèle A concerne la prédiction sur le niveau 0 (Forêt/Hors forêt) sur toute la commune et validé selon la grille de 500 mètres. Cette première classification constitue un masque Forêt/Hors forêt.

↳ Modèle B

Le modèle B concerne la prédiction sur les classes forestières du niveau 2. Le modèle est prédit uniquement sur les segments forestiers obtenus avec le masque du modèle A. Le résultat est validé selon les données récoltées sur le terrain.

Un modèle de classification « par pixel » sur le niveau 2 a également été réalisé. La validation est réalisée par extraction des pixels contenus dans les segments validés sur le terrain.

2.4.5 Evaluation de la précision de la classification à l'aide d'une matrice de confusion

Une fois que le Random Forest a attribué une classe de prédiction à chaque segment, il est nécessaire d'évaluer la qualité du modèle ainsi que celle de la prédiction. Une approche simple, efficace et objective est la création d'une matrice de confusion qui confronte des données de validation indépendantes (dont la classe est connue) à l'échantillon correspondant de données issues d'une prédiction (Figure 17). Les observations sur la diagonale sont correctement prédites, les autres sont des erreurs.

Par classe, on distingue deux types de précisions :

- la précision du producteur (ou la sensibilité) est la proportion d'observations de références correctement prédites (par classe) ;
- la précision du consommateur (ou la précision) est la proportion d'observations prédites correctement classées (par classe).

Données de validation		Référence terrain		Précision du consommateur
		Classe 1	Classe 2	
Prédiction du modèle	Classe 1	A	C	$\frac{A}{A + C}$
	Classe 2	B	D	$\frac{D}{B + D}$
Précision du producteur		$\frac{A}{A + B}$	$\frac{D}{C + D}$	

Figure 17: Exemple théorique de matrice de confusion dans le cas de 2 classes. Calcul des précisions du producteur et du consommateur

A partir d'une matrice de confusion, on peut également évaluer la qualité générale du modèle de classification selon la précision totale et l'indice de Kappa (Rosenfield et Fitzpatrick-Lins, 1986):

$$\text{Précision totale} = \frac{A + D}{A + B + C + D}$$

$$\text{Kappa} = \frac{\text{Précision totale} - \text{Précision attendue}}{1 - \text{Précision attendue}}$$

$$\text{où } \text{Précision attendue} = \frac{(D+C)*(D+B)+(A+C)*(A+B)}{\text{total*total}}$$

Le package "randomForest" propose une première estimation de la qualité du modèle via la construction d'une matrice de confusion qui exprime l'erreur «Out of bag» (Rodriguez-Galiano *et al.*, 2012). La précision de chaque arbre est évaluée sur une partie des données d'entraînement. Le modèle n'est donc pas validé sur un jeu de données externe, neutre et représentatif. Ceci est particulièrement vrai pour ce travail car les données d'entraînement ne sont que des données de peuplements purs.

Dans notre cas d'étude, l'erreur « OOB » tend à surestimer la précision du modèle. Une estimation réaliste de la précision implique d'utiliser un jeu de données indépendant de celui qui a servi à construire le modèle.

L'origine des données utilisées pour la phase de validation de la carte dépend du niveau de classification considéré. Pour la validation du modèle « Forêt/Hors forêt », nous avons considéré une grille de points systématiques (1 point par 25 ha, soit 445 points) couvrant la zone d'étude et dont le statut a été défini par photo-interprétation. Pour la validation du niveau 2 forestier, nous avons confronté la carte à la vérité terrain. 526 segments en forêt publique ont été visités en trois jours. Cette validation de la carte avec des points d'échantillonnage terrain est coûteuse mais primordiale. Elle permet de détecter les sources d'erreurs, de se rendre compte de la dynamique forestière et de la variabilité locale.

3 Résultats et discussions

La troisième partie de ce travail est consacrée à la présentation des résultats, en lien avec l'objectif de l'étude. Dans un souci de concision, seuls les résultats finaux, découlant de tests itératifs, sont présentés dans ce document. Nous aborderons la segmentation, la classification par objet et la classification par pixel. La qualité des modèles et des résultats sera évaluée et les approches seront comparées. Nous discuterons de la méthodologie et de pistes concrètes d'amélioration.

3.1 Segmentations

La segmentation réalisée avec eCognition, en considérant les valeurs de paramètre du point 2.3.4, a généré 84 467 segments sur l'ensemble de la commune. Ceux-ci présentent une superficie moyenne de 1 315 m² et un écart-type de 1 626 m². Ceci correspond à un coefficient de variation de 123 % et des superficies extrêmes de 0,12 m² et 27 019 m² (Figure 18).

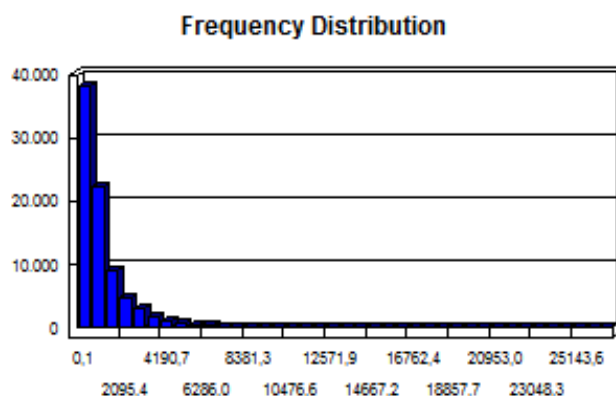


Figure 18: Distribution des fréquences des aires des segments obtenus avec eCognition sur la commune de Paliseul

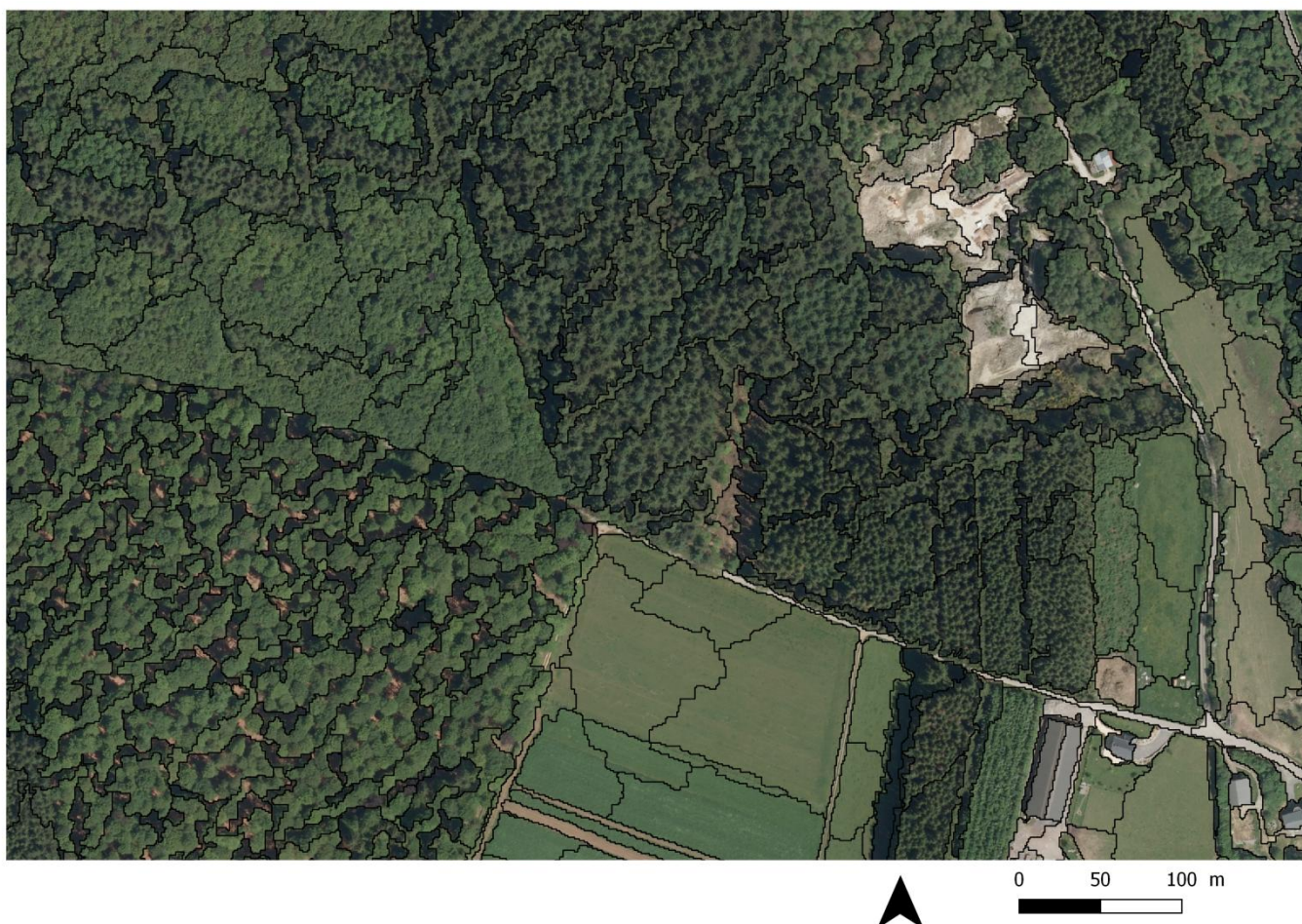


Figure 19: Extrait de la segmentation à l'aide d'eCognition sur une portion du territoire de la commune de Paliseul. Illustration du comportement de l'algorithme pour différentes classes d'occupation du sol ; source : SPW

On observe que la taille des segments est liée à l'homogénéité de la classe d'occupation du sol (Figure 19). Les zones agricoles sont segmentées grossièrement alors que les zones urbaines et forestières sont segmentées de façon plus fine. Au sein de la forêt, les peuplements denses sont délimités par de plus grands objets que les zones hétérogènes présentant du sol nu, des ombres et des houppiers.

3.2 Classifications

3.2.1 Modèle A

Les 5 variables intervenant de manière prépondérantes dans le modèle sont la hauteur moyenne (mn_mnh), la moyenne de la bande bleue du 20 juillet 2016 (mn_D2B), la moyenne de la bande rouge de cette même date (mn_D2R), la moyenne de l'indice vert normalisé du 8 mai 2016 (mn_norm_G_D1) et la moyenne de la bande verte du 20 juillet 2016 (mn_D2G). La sixième variable qui ressort est la texture. (Figure 20)

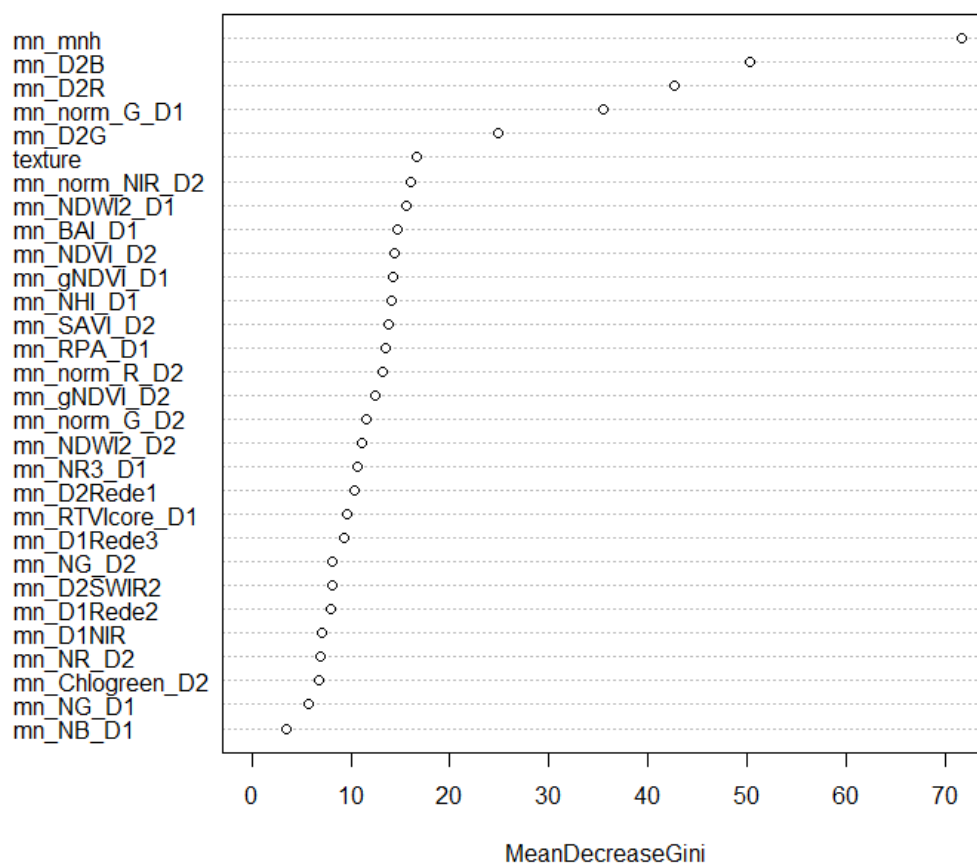


Figure 20: Indices de Gini pour les 30 variables sélectionnées pour la classification "Forêt/Hors forêt" (modèle A)

Tableau 4: Matrice de confusion générée par Random Forest pour le modèle A

Modèle A (Random Forest)	Forêt	Hors forêt (OOF)	Précision du consommateur
Forêt	497	4	0,992
Hors forêt (OOF)	4	497	0,992
Précision du producteur	0,992	0,992	

La matrice de confusion indique que la précision du producteur et du consommateur est de 99,2% pour les deux classes du modèle A (Tableau 4). L'erreur "Out of Bag" du modèle A est de 0,8%. On peut conclure de ces résultats que la proportion de segments mal classés est faible. Elle concerne probablement des segments en lisière.

Tableau 5: Matrice de confusion générée pour la validation du modèle A selon la grille (500 m)

Modèle A (grille 500)	Forêt	Hors forêt (OOF)	Précision du consommateur
Forêt	223	10	0,957
Hors forêt (OOF)	11	201	0,948
Précision du producteur	0,953	0,953	

La validation externe de la carte (modèle A), à l'aide des 445 observations photo-interprétées, évalue la précision totale de la prédiction à 95,28% et un indice de Kappa de 90,5%. L'erreur du consommateur plus grande pour la classe hors forêt concerne des segments prédits en forêt alors que la référence est du hors forêt (Tableau 5). Cela représente probablement les erreurs de détection correcte des zones de lisière ou des sols nus forestiers classés comme des sols agricoles.

3.2.2 Modèle B

Les variables les plus prépondérantes, selon l'indice de Gini, sont la moyenne de la bande verte pour les deux dates (mn_D1G et mn_D2G) et la moyenne de la bande Rededge1 pour le 20 juillet 2016 (mn_D2Rede1). La quatrième variable est un indice normalisé dérivé de la bande Green pour le 8 mai 2016 (mn_norm_G_D1). La texture n'intervient pas dans les 30 variables sélectionnées et la hauteur moyenne apparait dans les 15 premières variables (contre la première pour le modèle A) (Figure 21). En effet, l'importance de la hauteur est plus grande pour distinguer la forêt du hors forêt, plutôt que pour discriminer les essences les unes des autres. Immitzer *et al.* (2016) ont prouvé l'importance des bandes Rededge, SWIR et Blue de Sentinel-2 pour la cartographie de la végétation. Dans ce travail, pour le modèle B, il semble que ce soient les bandes Green, Rededge et SWIR qui ressortent.

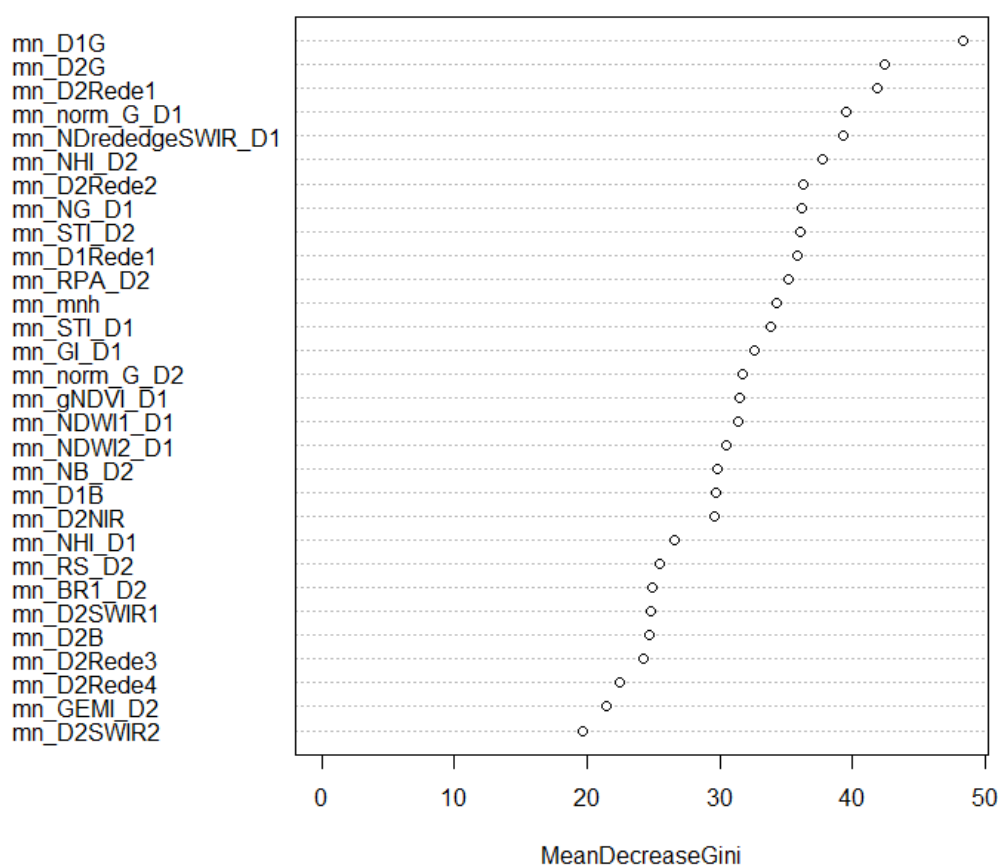


Figure 21: Indices de Gini pour les 30 variables sélectionnées pour la classification sur le niveau 2 en forêt (modèle B)

L'erreur Out of Bag du modèle B est de 13,41%. Les précisions par classe varient entre 73% et 94%. Les classes avec les précisions du producteur et du consommateur les plus faibles sont la classe de Régénération et de Sol nu (Figure 22). La matrice de confusion permet toutefois de nuancer le résultat car les erreurs de prédiction de la classe Régénération se font au profit de la classe Sol nu, et *vice versa* (Tableau 6). Les deux classes représentent des coupes rases qui ne constituent qu'une seule classe thématique en termes de gestion forestière.

Tableau 6: Matrice de confusion générée par Random Forest pour le modèle B (Chêne/Douglas/Epicéa/ Hêtre/Mélèze/Pin/ Régénération/Sol nu forestier)

Modèle B (Random Forest)	CH	DO	EP	HE	MZ	PS	régénération	sol nu	Précision du consommateur
CH	118	0	0	11	0	3	1	4	0,861
DO	0	124	7	0	3	3	0	0	0,905
EP	0	12	119	0	1	2	3	0	0,869
HE	6	0	0	124	3	1	3	0	0,905
MZ	0	2	0	0	129	5	1	0	0,942
PS	2	0	1	8	1	122	2	1	0,890
régénération	1	3	1	2	1	5	100	24	0,730
sol nu	1	0	0	1	0	3	19	113	0,825
Précision du producteur	0,922	0,879	0,930	0,850	0,935	0,847	0,775	0,796	

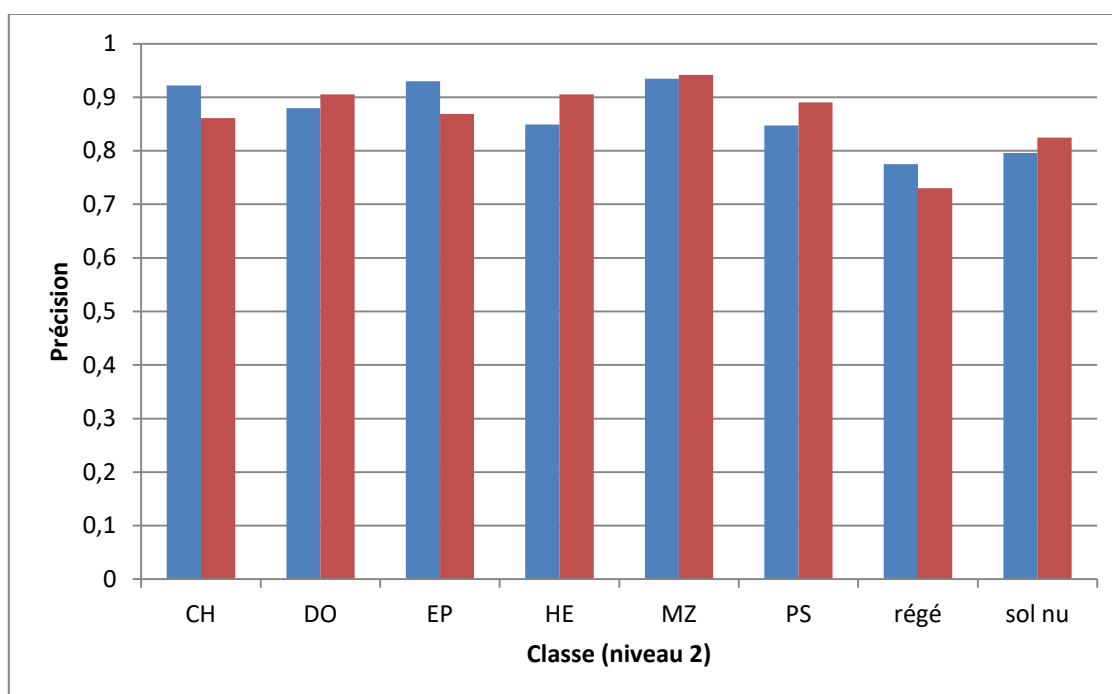


Figure 22: Evaluation de la qualité "Out Of Bag" du modèle B de classification par classe en forêt. La précision du producteur est représentée en bleu et la précision du consommateur en rouge pour les classes Chêne, Douglas, Epicéa, Hêtre, Mélèze, Pin, Régénération et Sol nu forestier

La validation à l'aide des points d'échantillonnage terrain donne une matrice de confusion à partir de laquelle on peut évaluer les précisions par classe (Tableau 7). On peut considérer qu'il s'agit des précisions les plus représentatives de la qualité réelle du modèle et de la carte. Les précisions sont globalement moindres par rapport à celles estimées sur les données d'entraînement.

Tableau 7: Matrice de confusion de la validation des points d'échantillonnage terrain du modèle B

Modèle B: validation des points d'échantillonnage terrain	AB	BO	CH	CH-HE	CH/HE	DO	DO/AB	DO/EP	DO/HE	EP	EP/AB	EP/DO	EP/DO/HE	EP/HE	HE	HE/DO	HE/do/ep	MZ	PS	PS/HE	Régénération	sol nu	Précision du consommateur	
AB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CH	0	0	32	1	3	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0,820
CH-HE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CH/HE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DO	2	0	0	0	0	31	0	5	0	1	1	48	0	0	0	0	0	0	0	0	0	0	0	0,352
DO/AB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DO/EP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DO/HE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EP	1	0	0	0	0	1	0	1	0	34	1	5	0	0	0	0	0	0	0	0	0	0	0	0,791
EP/AB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EP/DO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EP/DO/HE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EP/HE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HE	0	0	2	0	0	1	0	0	0	1	0	0	0	0	90	0	0	0	1	0	0	0	0	0,947
HE/DO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HE/do/ep	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MZ	0	3	1	0	0	3	0	0	1	5	0	2	1	4	55	1	1	28	0	0	1	0	0	0,264
PS	6	2	1	0	0	12	1	0	1	12	0	0	0	0	42	0	0	6	3	1	5	0	0	0,033
PS/HE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
régénération	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	24	0	0	0,923
sol nu	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	31	0	0,838
Précision du producteur	0	0	0,889	0	0	0,646	0	0	0	0,641	0	0	0	0	0,471	0	0	0,823	0,75	0	0,667	1	0	

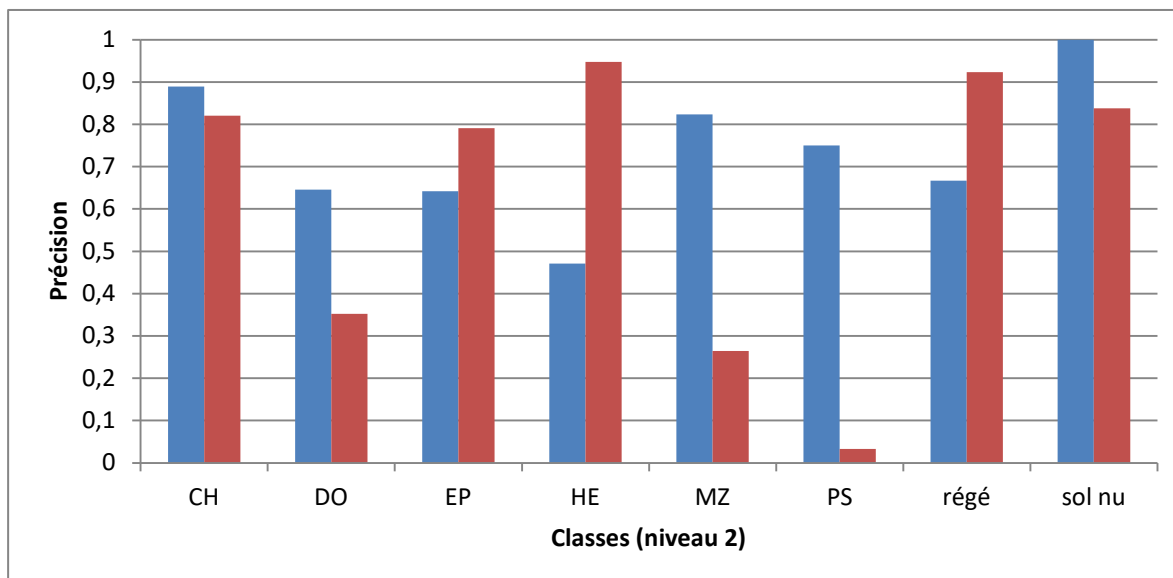


Figure 23: Evaluation de la qualité du modèle B de classification par classe selon la validation des points d'échantillonnage terrain. La précision du producteur est représentée en bleu et la précision du consommateur en rouge pour les classes Chêne, Douglas, Epicéa, Hêtre, Mélèze, Pin, Régénération et Sol nu forestier

La précision totale est de 51,9% avec un indice de Kappa de 45,6%.

Certaines classes sont très problématiques. Les précisions du consommateur sont très faibles pour les classes Pin, Mélèze et Douglas. Cela signifie qu'une grande proportion de données prédites dans ces classes est erronée. La classe Hêtre est celle avec la meilleure précision du consommateur mais la plus grande erreur du producteur. Cela signifie que les peuplements prédits en hêtre sont corrects mais qu'une proportion non négligeable de hêtraie est prédite dans une autre classe. (Figure 23)

La précision du consommateur pour les classes feuillues (Chêne et Hêtre) est au-delà de 80% pour le modèle B (Figure 23). On peut en conclure que le modèle B arrive à bien distinguer ces deux essences. A titre de comparaison, Immitzer *et al.* (2016) avaient obtenu une précision totale de 65% pour la classification d'essences feuillues et résineuses.

La matrice de confusion permet de détecter les erreurs et de formuler des pistes sur leur origine (en associant également diverses informations de terrain récoltées). Par exemple, la faible précision du producteur pour la classe Hêtre est liée à une grande proportion de segments de hêtre prédits en Pin et la faible précision du consommateur dans la classe Douglas est liée à la présence sur le terrain de mélange Epicéa/Douglas (Tableau 7).

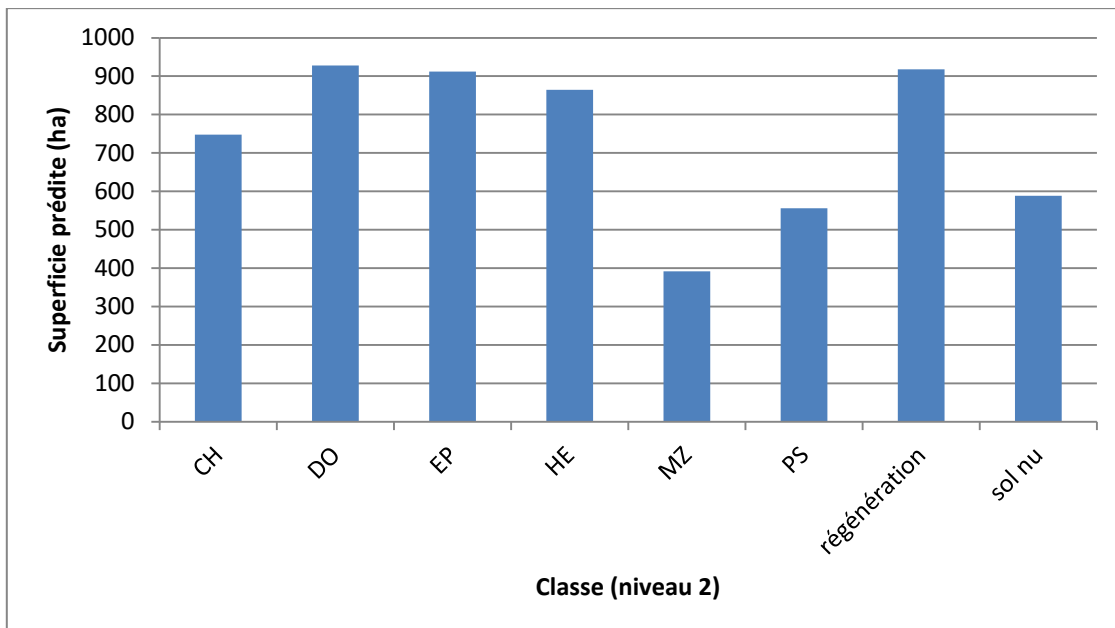


Figure 24: Superficies forestières prédites des classes du modèle B (Chêne/Douglas/Epicéa/Hêtre/Mélèze/Pin/ Régénération/Sol nu forestier) sur l'ensemble de la commune de Paliseul

Le résultat du modèle B permet d'estimer une superficie par classe forestière (Figure 24). Ce résultat est intéressant pour la filière bois. Toutefois, les résultats des superficies sont très probablement surestimés puisque le modèle ne sait prédire que sur les classes d'entraînement (quelques essences d'intérêt économique). Un grand nombre d'essences et de cas particuliers, tels que les mélanges, ne savent pas être détectés avec les classes de la légende actuelle. Or, dans ces cas, l'algorithme Random Forest a tout de même attribué la classe dont la réponse spectrale est la plus proche. On ne peut donc pas utiliser ce résultat brut pour estimer la ressource ligneuse de la commune.

3.2.3 Approche par pixel sur le niveau 2

Les variables prépondérantes pour le modèle de classification par pixel sont la hauteur, ainsi que la bande verte pour les deux dates. L'indice de Gini est globalement très élevé et atteint des valeurs au-delà de 800 pour l'approche par pixel contre moins de 50 pour l'approche par objet. Les deux approches ont plus de 60% de variables en commun. (Figure 25)

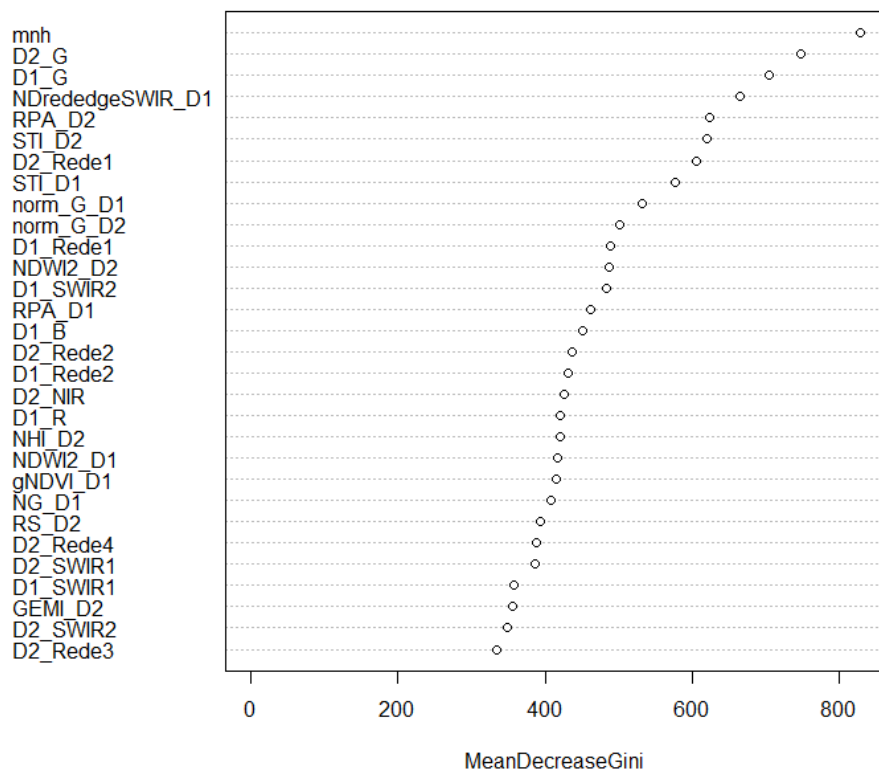


Figure 25: Indices de Gini pour les 30 variables sélectionnées pour le modèle de classification par pixel (niveau 2 en forêt)

Tableau 8: Matrice de confusion Out Of Bag du modèle de classification par pixel (Chêne/Douglas/Epicéa/ Hêtre/Mélèze/Pin/ Régénération/Sol nu forestier)

Approche par pixel (Random Forest)	CH	DO	EP	HE	MZ	PS	Régénération	sol nu	précision du consommateur
CH	1897	2	5	113	1	72	7	6	0,902
DO	3	1836	72	44	75	39	27	7	0,873
EP	2	121	1849	10	46	47	14	14	0,879
HE	53	26	10	1941	20	26	18	9	0,923
MZ	3	40	45	29	1931	37	13	5	0,918
PS	46	5	27	43	41	1903	18	20	0,905
régénération	13	19	11	24	3	43	1804	186	0,858
sol nu	2	4	3	5	5	29	189	1866	0,887
précision du producteur	0,94	0,89	0,91	0,88	0,91	0,87	0,86	0,883	

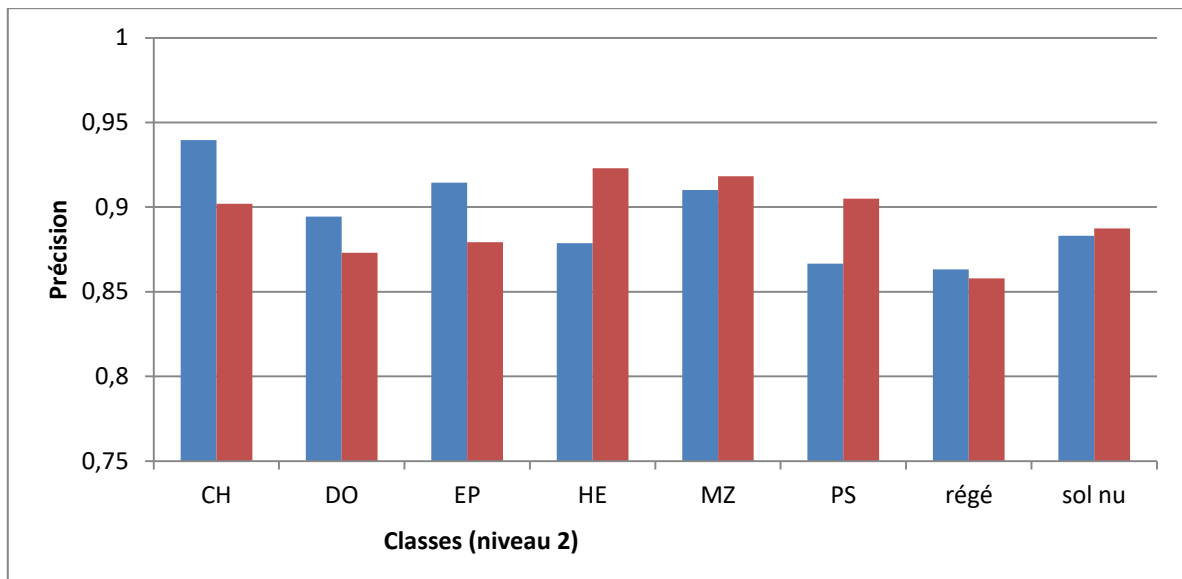


Figure 26: Précision Out Of Bag du modèle de classification par pixel. La précision du producteur est représentée en bleu et la précision du consommateur en rouge pour les classes Chêne, Douglas, Epicéa, Hêtre, Mélèze, Pin, Régénération et Sol nu forestier

L'erreur Out Of Bag du modèle par pixel est de 10,68%. Cette erreur est plus faible que celle du modèle B (approche par objet). Les précisions Out of Bag par classe sont toutes supérieures à 85% (Figure 26). La précision Out Of Bag du modèle par pixel est meilleure que celle du modèle B (par objet).

Comme pour la classification par objet, la matrice de confusion indique que les confusions principales se font entre les classes de régénération et sol nu (Tableau 8). Ces deux classes font partie du niveau 1 «Coupe rase», une classe très dynamique et hétérogène.

Tableau 9: Matrice de confusion de la validation des points d'échantillonnage terrain du modèle de classification par pixel

Approche par pixel (validation des points d'échantillonnage terrain)	AB	BO	CH	CH-HE	CH/HE	DO	DO/AB	DO/EP	DO/HE	EP	EP/AB	EP/DO	EP/DO/HE	EP/HE	HE	HE/DO	HE/do/ep	MZ	PS	PS/HE	régé	sol nu	Précision du consommateur (approche par pixel)		
AB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
BO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
CH	0	0	19	0	3	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0,76
CH-HE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CH/HE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DO	1	3	12	0	0	25	0	5	0	6	1	38	0	0	9	0	0	2	0	0	0	0	0	0	0,245
DO/AB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DO/EP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DO/HE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EP	2	1	2	0	0	9	0	1	0	31	0	14	1	2	45	1	1	5	1	1	2	0	0	0	0,260
EP/AB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EP/DO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EP/DO/HE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EP/HE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HE	3	1	1	1	0	8	1	0	2	7	1	2	0	0	100	0	0	6	1	0	4	0	0	0	0,725
HE/DO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HE/do/ep	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MZ	2	0	1	0	0	2	0	0	0	7	0	1	0	2	27	0	0	19	2	0	4	1	0	0	0,28
PS	1	0	0	0	0	4	0	1	0	2	0	0	0	0	5	0	0	2	0	0	12	2	0	0	0
PS/HE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
régé	0	0	1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	12	14	0	0	0,414
sol nu	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	14	0	0	0,875
Précision du producteur (approche par pixel)	0	0	0,528	0	0	0,5	0	0	0	0,585	0	0	0	0	0,5	0	0	0,6	0	0	0,33	0,452	0	0	

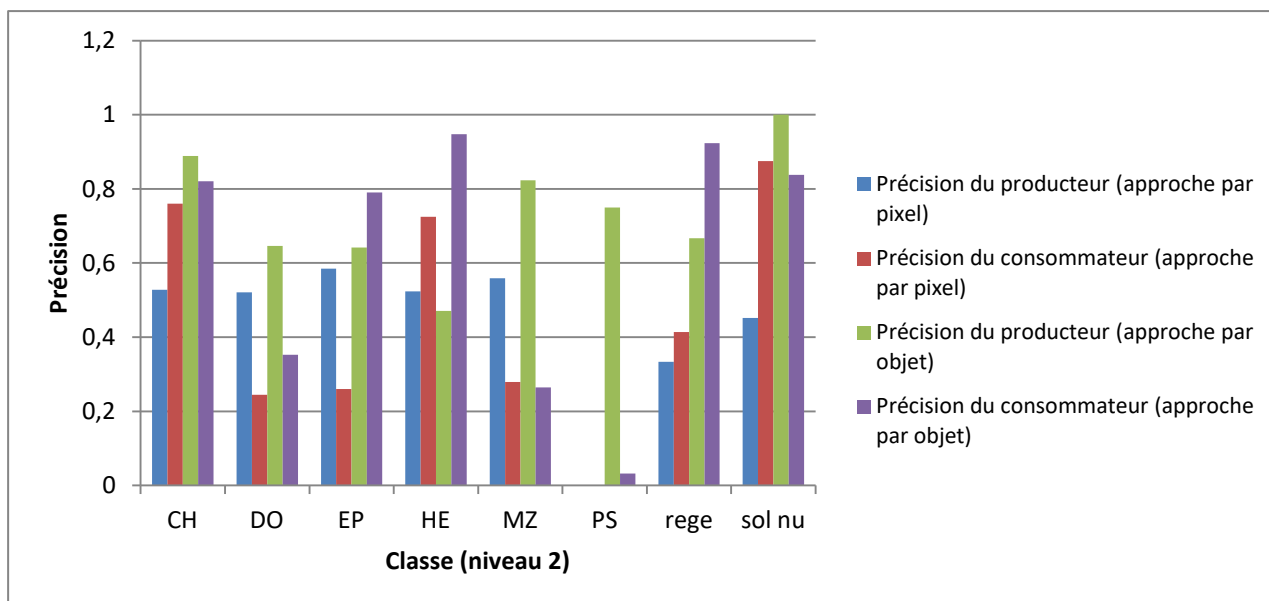


Figure 27: Précision des modèles de classification par objet et par pixel selon la validation des points d'échantillonnage terrain pour les classes Chêne, Douglas, Epicéa, Hêtre, Mélèze, Pin, Régénération et Sol nu forestier

En ce qui concerne la validation des points d'échantillonnage terrain, le modèle de classification par pixel est moins précis que la classification par objet (Figure 27). Les confusions entre les classes pour l'approche par pixel sont reprises dans le Tableau 9. Par exemple, la faible précision du consommateur pour les classes Douglas et Epicéa est liée au fait que les points d'échantillonnage terrain présentent un mélange d'Epicéa/Douglas. La faible précision du producteur de la classe Régénération est liée à des zones de régénération prédites en classe Pin.

Le modèle de classification par pixel prédit, sur le masque forestier du modèle A, une majorité de pixels en classe Epicéa, Douglas et Hêtre (Figure 28). L'approche par objet prédisait plus de régénération et moins d'épicéa (Figure 24).

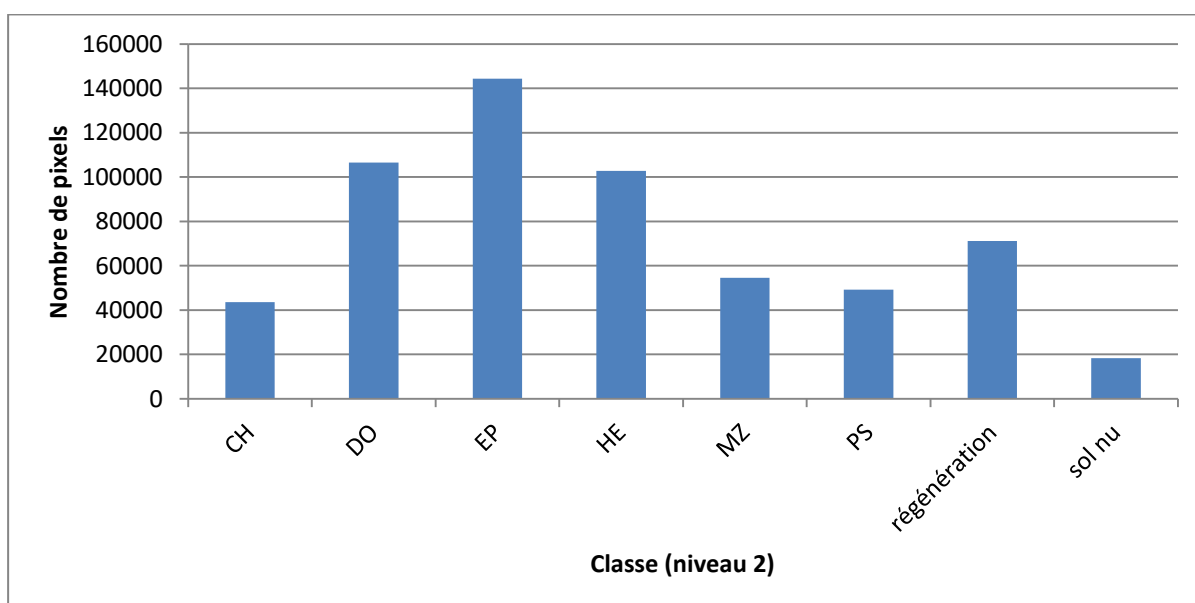


Figure 28: Nombre de pixels (10 m de côté) prédit par classe forestière de niveau 2 (Chêne/Douglas/Epicéa/ Hêtre/Mélèze/Pin/ Régénération/Sol nu forestier) selon la classification par pixel sur la commune de Paliseul

3.3 Comparaison des prédictions, références DNF et validations terrain

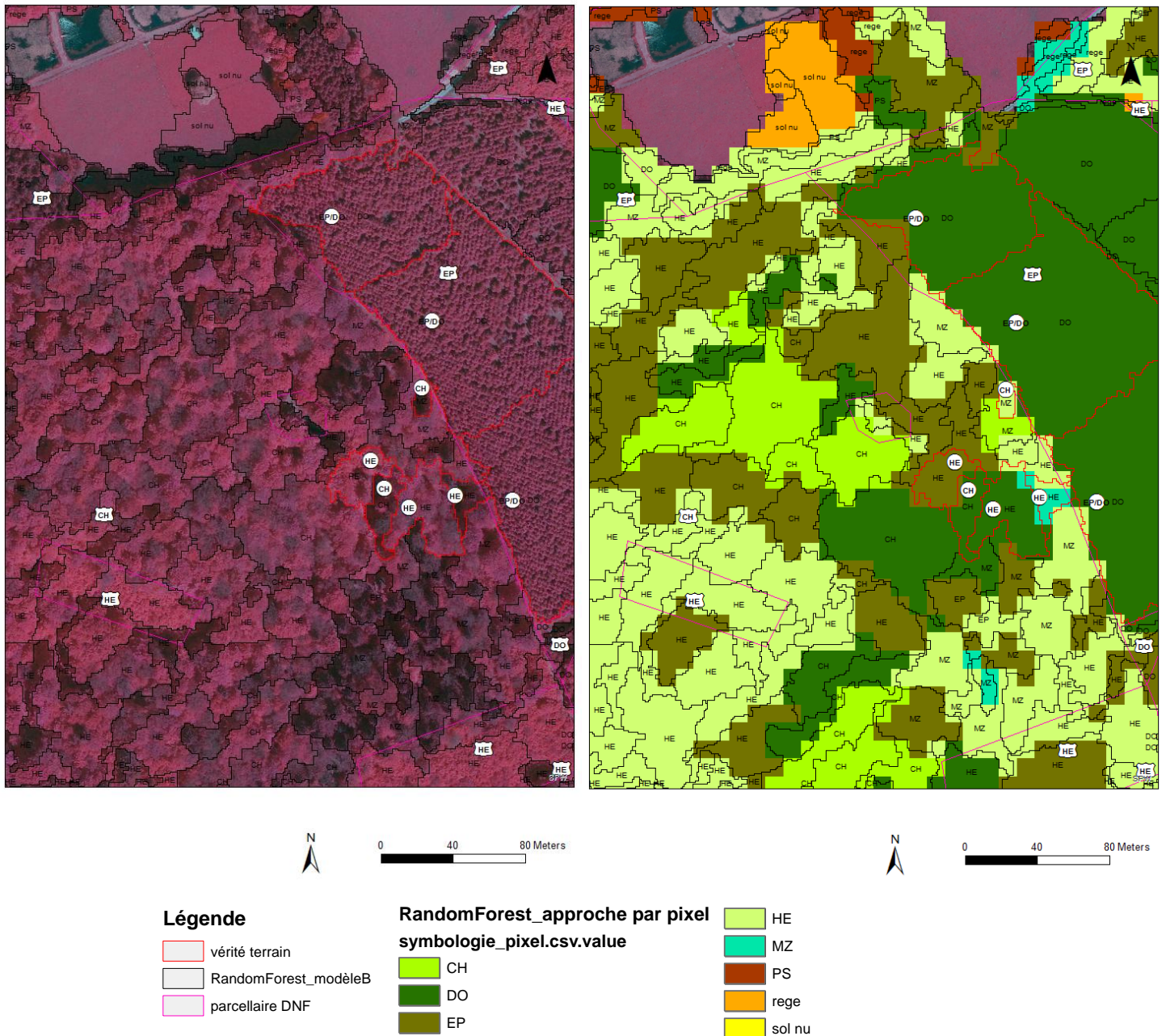


Figure 29: Extrait des résultats de prédictions sur la zone d'étude selon la classification par pixel (voir le code couleur de la légende) et par objet (voir la classe prédite pour chaque segment) (niveau 2 en forêt). L'information du parcellaire est présentée dans les étiquettes blanches et délimitée en rose. Les données de validation terrain sont présentées dans les disques blancs et délimitées en rouge ; source: SPW

L'analyse visuelle de quelques zones tend à confirmer que la classification par pixel n'est pas précise. En effet, sur base de l'orthophoto, on peut identifier des peuplements feuillus sur la partie gauche de la Figure 29. La classification par objet a tendance à bien les identifier alors que la classification par pixel prédit la mauvaise classe (Douglas et Epicéa).

- Cas des mélanges résineux

Le modèle B prédit les mélanges résineux, de type Epicéa-Douglas, quasi toujours en Douglas (Figure 29). Ceci est peut-être lié aux proportions réelles des houppiers, suite à une mortalité spécifique ou à des éclaircies dirigées sur les Epicéa, par exemple, ou à la hauteur plus importante du Douglas par rapport à l'Epicéa. Une autre piste serait que la réponse spectrale du Douglas est plus significative,

c'est-à-dire qu'elle influence fort la moyenne à l'échelle du segment. L'hypothèse d'un changement de proportion d'essences ou de densité n'a pas été confirmée sur le terrain.

L'analyse des probabilités d'attribution de classe par segment (votes du Random Forest) a été réalisée pour environ 50 segments. Le mélange typique de plantation Epicéa-Douglas n'a pas montré de schéma de votes clair et systématique permettant de mettre en évidence les deux essences en mélange. Pour certains segments, la première essence est prédite à 50% et la seconde à 45%. Pour d'autres segments, la seconde essence renseignée dans le mélange selon le DNF n'est prédite qu'à moins de 10%.

L'analyse de la prédiction de la classification par pixel a montré une homogénéité totale dans les pixels extraits par segment. Au vu de ces résultats, les peuplements mélangés ne savent pas être prédits avec l'approche par pixel (Figure 29, à droite).

- Cas des mélanges feuillus

L'analyse de peuplements feuillus a mis en évidence que le modèle B détecte plusieurs essences sur un îlot (hêtraie-chênaie) alors que le parcellaire ne renseigne qu'une hêtraie pure (Figure 30). On s'attend donc à ce que le modèle confonde la classification de ces deux essences ou que la source d'information soit incorrecte. Sur le terrain, le modèle B a su aussi détecter des essences non renseignées dans le parcellaire. Dans un cas précis, il s'agissait de l'ancien peuplement de chênes, dont certains individus sont toujours sur pied (Figure 30). Néanmoins, la qualité de la détection des houppiers individuels dépend fortement de la qualité de la segmentation. L'évaluation de la segmentation sera abordée dans les pistes d'amélioration de ce travail. A ce stade, il est important de souligner que la segmentation des houppiers d'essences forestières, sous nos latitudes, est une problématique de recherche en soi assez difficile à traiter.

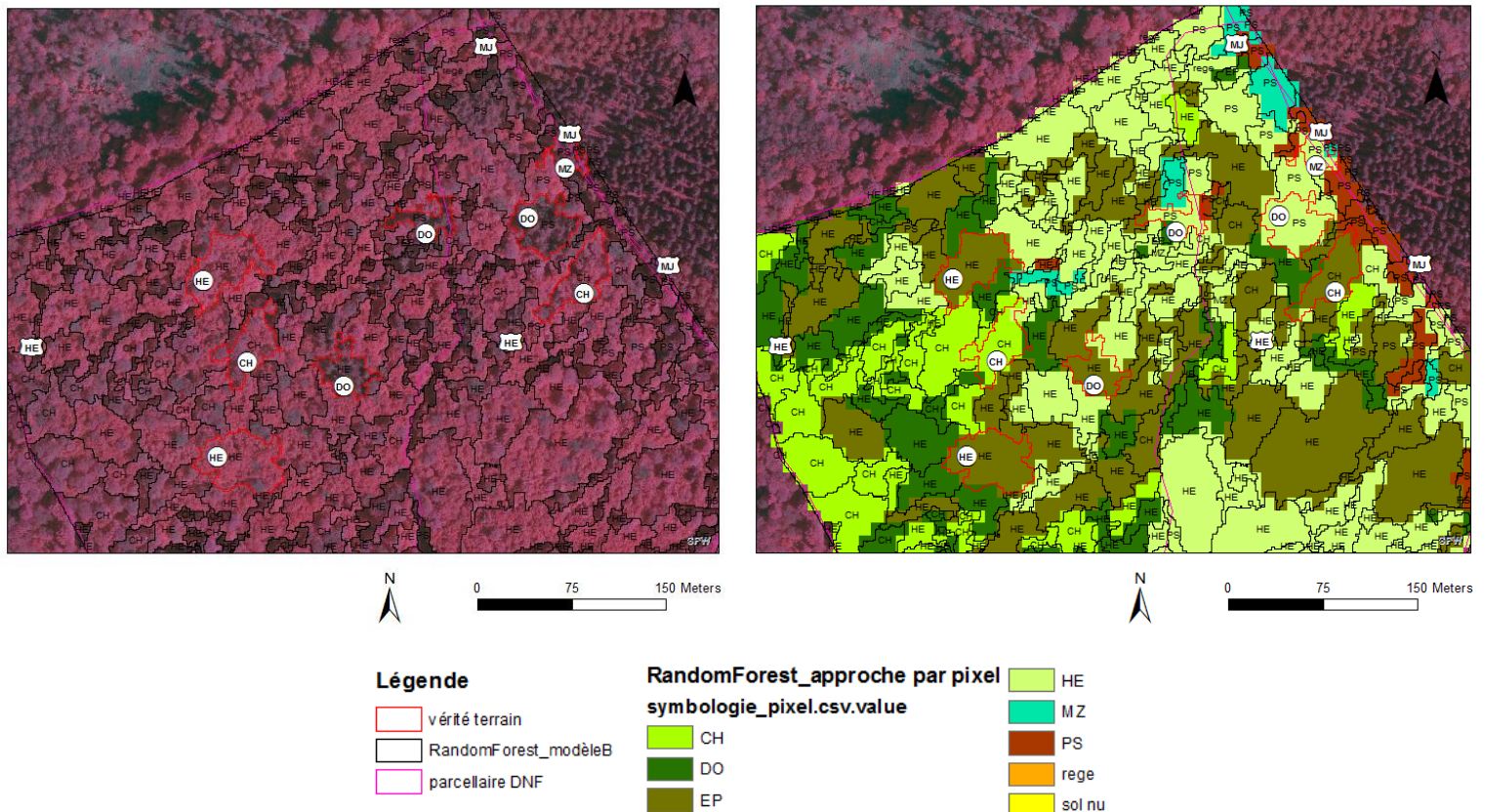


Figure 30: Extrait des résultats de prédictions sur la zone d'étude selon la classification par pixel (voir le code couleur de la légende) et par objet (voir la classe prédite pour chaque segment) (niveau 2 en forêt). L'information du parcellaire est présentée dans les étiquettes blanches et délimitée en rose. Les données de validation terrain sont présentées dans les disques blancs et délimitées en rouge. Etude des peuplements feuillus ; source: SPW

- Cas de peuplements ouverts

Beaucoup de segments «problématiques» sont classés en Pins ou en Mélèze (faible précision du consommateur pour ces classes). Ces essences sont caractérisées par un couvert très léger et un espacement entre les individus. Il pourrait s'agir en réalité de segments qui reprennent des trouées ou des zones fort éclaircies, voire des feuillus qui n'ont pas encore débourré ou avec un couvert très léger.

La validation des points d'échantillonnage terrain a mis en évidence que les segments classés en Pins sont de vieux hêtres fort espacés. La proportion de végétation herbacée et arbustive visible par télédétection est donc plus grande. La réponse spectrale du sous-étage peut influencer la moyenne au sein du segment et est probablement significative pour la classification. Par ailleurs, l'utilisation d'images Sentinel-2 avec des acquisitions hors période de végétation permettrait peut-être de différencier le Pin du Hêtre.

- Cas de mauvaise segmentation

Si le segment correspond aux limites d'un houppier, la réponse spectrale correspondra à celle d'un peuplement pur. A contrario, si un segment se retrouve à cheval sur plusieurs houppiers d'essences différentes, la moyenne des pixels fausse la prédiction. De la sorte, un segment qui sera à moitié sur un peuplement de douglas et l'autre moitié en hêtre pourrait être classé «mélèze» (Figure 31).

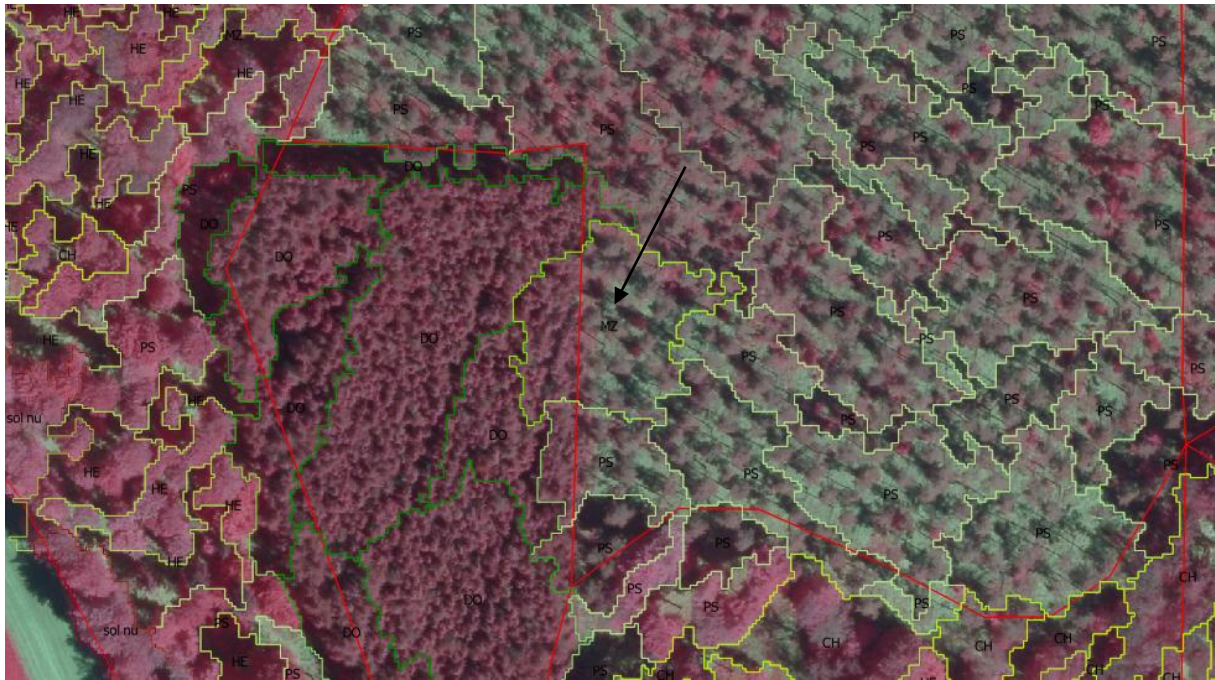


Figure 31: Classification selon le modèle B sur une zone forestière de Paliseul mettant en évidence un segment à cheval sur deux peuplements et classé en mélèze. Cet exemple illustre l'impact des imprécisions dans les segments sur la qualité de la classification ; source: SPW

3.4 Avantages et inconvénients de la méthode

- Les données multi-capteurs

L'approche multi-capteurs combine les différentes résolutions et qualités spectrales et spatiales disponibles.

Dans le cadre de ce travail, nous avons utilisé les données auxquelles nous avons accès gratuitement. Néanmoins, en fonction du contexte de l'étude, l'accès à ce panel de données n'est pas systématique et peut être coûteux.

L'utilisation des données de télédétection disponibles pour ce travail a ses limites. En effet, le modèle, basé sur des images Sentinel-2, des orthophotos et un MNH, ne considère finalement que de l'information en deux dimensions pour la segmentation et la classification. Il détecte la végétation et les houppiers des essences dont le couvert est visible depuis le ciel aux dates de prise de vue (8 mai 2016 et 20 juillet 2016). Il lui est impossible de détecter les arbres de sous-étage qui sont référencés ou non dans les inventaires de terrain (Figure 32).



Figure 32: Illustration des îlots sur le terrain : îlot de chênes "pur" ; Source: F. Crutzen

- L'approche par objet

Dans le cadre de ce travail, la qualité de la segmentation a été évaluée par photo-interprétation et par essai-erreur. Aucune évaluation objective de la qualité des segments n'a été réalisée. Ceci est problématique car tous les résultats découlent de la qualité des segments. Cette piste sera abordée en 3.6.

Partant de l'hypothèse d'une segmentation de qualité, l'information spectrale est résumée tout en tenant compte d'une certaine variabilité au sein de chaque classe. Le calcul de variables statistiques au sein des segments améliore la qualité finale de la classification, ce qui n'est pas réalisable avec une approche par pixel (Benz *et al.*, 2004). L'approche par segment permet aussi d'éviter l'«effet poivre et sel» rencontré lorsque chaque pixel est classé indépendamment⁴¹.

Cependant, l'information pour un segment est extraite des pixels de 10 m dont le centroïde tombe à l'intérieur du segment. Ces pixels ne sont pas tous représentatifs de l'objet. Si la segmentation crée des segments à cheval sur plusieurs houppiers, parfois d'essences différentes, la réponse spectrale est biaisée par rapport aux classes d'entraînement.

- L'approche par pixel

Les figures 29 et 30 comparent les résultats des prédictions des deux approches. Conforté par les résultats de validation des modèles sur base des données de terrain (Figure 27), on peut confirmer que le résultat de l'approche par pixel est moins précis que celui de l'approche par segment. De plus, l'approche par pixel ne permet pas de détecter les essences en mélange.

3.5 Pistes d'explication de la qualité du modèle

Pour rappel, l'objectif de ce travail est une cartographie forestière à l'échelle des essences d'un peuplement. Ceci est possible avec une approche «segments» et en se focalisant sur les essences d'intérêt économique pour lesquelles la superficie occupée est suffisante. La précision globale de la carte générée (modèle B) varie en fonction des classes (62%-73%).

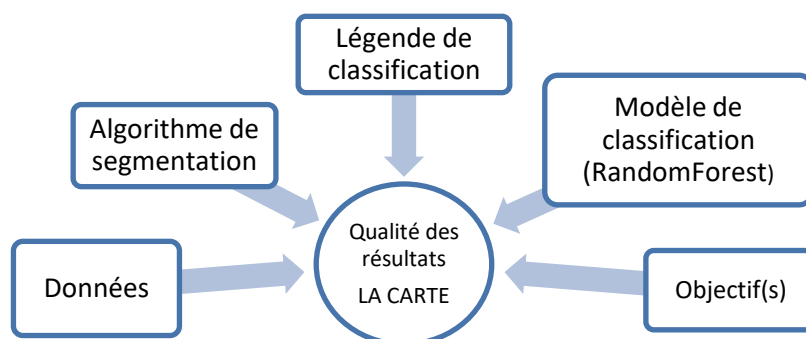


Figure 33: Illustration de quelques éléments qui influencent la qualité finale de la carte

Les erreurs se propagent tout au long du processus (Lu et Weng, 2007). La précision de la carte finale est donc liée à la précision de chaque étape, notamment le choix des données pertinentes, la qualité des données initiales, la préparation des données, la qualité de la segmentation (évaluation de la sur- ou sous-segmentation), le calcul de variables, le choix des classes, la création du jeu de données de référence et de validation (effet de l'observateur), la précision des outils et les limites des modèles. La proportion d'erreur acceptée dépendra également des objectifs d'utilisation de la carte (Figure 33).

⁴¹

https://servirglobal.net/Portals/0/Documents/Articles/ChangeDetectionTraining/Module3_LC_Classification_presentation.pdf (consulté le 28 mai 2017)

3.6 Pistes concrètes d'amélioration du modèle de classification

A l'issue de ce travail, il est possible de formuler des pistes concrètes d'amélioration du travail en lien avec la méthodologie actuelle. Les pistes d'amélioration sont, entre autres :

- d'améliorer la segmentation en intégrant plus de variables dans la couche d'entrée en testant l'apport des indices de type NDVI, par exemple ;
- d'exploiter le contexte des objets en créant des segments imbriqués dans eCognition ;
- d'automatiser les combinaisons de paramètres de segmentation (eCognition) ;
- d'évaluer la qualité de la segmentation de façon objective en évaluant la sur- ou sous-segmentation des objets par rapport aux segments de référence digitalisés selon les objectifs de l'étude. On sélectionnera le modèle non significativement différent de la référence. (Lu et Weng, 2007 ; Gestion des Ressources Forestières et des Milieux Naturels - Unité de Géomatique - ULg - Gembloux Agro-Bio Tech, 2010) ;
- de revoir la définition des classes forestières en fonction des objectifs. Une piste d'amélioration peut être, notamment, l'ajout des classes Abies et Bouleaux. De plus, il peut être intéressant d'entraîner le modèle Random Forest sur des peuplements mélangés puisqu'aucune technique n'a permis de les mettre en évidence dans ce travail. Pour cela, les données d'entraînement doivent être revues également en fonction des nouvelles classes ;
- d'utiliser plus de deux dates pour la classification. L'augmentation du nombre de dates, avec des images Sentinel-2A et 2B de qualité, devrait permettre de disposer de dates en période de végétation très spécifiques aux essences. D'autres images Sentinel-2 de dates hors période de végétation pourraient aider à créer un masque sur les essences sempervirentes / décidues. Immitzer *et al.* (2016) soulignaient déjà le haut potentiel des données multi-temporelles pour la classification d'image de végétation ;
- d'améliorer la sélection de variables en combinant plusieurs méthodes dont la sélection des variables selon l'indice de Gini (Menze *et al.*, 2009). Afin de ne retenir que les variables les plus explicatives et les moins corrélées entre elles, l'analyse en composantes principales, la régression ou encore l'utilisation de VSurf sont des pistes qui ont été envisagées, voire examinées mais qui n'ont pas abouti par manque de temps ;
- d'améliorer la qualité de données de référence (pour l'entraînement et la validation du modèle) sur base d'un plus grand échantillon de données de terrain ;
- d'avoir à notre disposition des données ortho-images et MNH acquises en même temps que les images Sentinel. Cela permettrait d'augmenter la qualité de la segmentation et de la classification en associant toutes les informations instantanées à la définition d'une classe.

- Cas particulier des ombres

Les ombres apparaissent sur les prises de vues des orthophotos. En effet, la plateforme aérienne (avion ou drone) survole la zone à une altitude bien plus faible qu'un satellite. Par ailleurs, sa fauchée est beaucoup moins large et le recouvrement de la zone se fait en plusieurs passages avec un angle de prise de vue différent. Les vols se font toute la journée et l'orientation du soleil varie en créant des ombres maximales en début et en fin de journée. Dès lors, la segmentation va discerner les ombres en tant qu'objets à part entière puisqu'elles ont une réponse spectrale assombrie. A cela il faut ajouter les problèmes de *mosaïquage* des images qui ont été prises à différentes périodes (Figure 7 et Figure 19).

La classification, elle, est basée sur les images satellites dans le but de s'affranchir de ces problèmes. En effet, le satellite recouvre une grande zone en un seul passage. De plus, sous nos latitudes, Sentinel-2 passe en fin d'avant-midi. On peut également supposer que les effets des ombres sont limités lorsque l'acquisition a lieu au zénith solaire (en été). Voici trois pistes pour traiter les ombres:

- On peut entraîner le modèle de classification (i) à reconnaître les ombres et (ii) à attribuer l'ombre à son «propriétaire». Beaucoup d'algorithmes existent et permettent de détecter les ombres de façon efficace (Dare, 2005). Lorsque l'on connaît l'heure et la date de prise de vue, on peut connaître l'orientation du soleil et l'origine de l'ombre ;

- Si l'on compte classifier sur base des données de l'orthophoto, il peut être nécessaire de corriger la radiométrie des ombres. Tout d'abord, la détection des zones sombres est possible sur base d'un seuil de clarté (*brightness*).

$$Brightness = \frac{Green + Red + Blue}{3}$$

Ensuite, on peut éclaircir ces pixels, c'est-à-dire ajouter des unités de valeur spectrale en rouge, vert et bleu. Ainsi, on peut supposer que les ortho-images sont rectifiées et que la réponse spectrale correspond à l'occupation du sol réelle. Dare (2005) s'est concentré sur le milieu urbain. Il montre que les ombres créent de gros problèmes de classification et que la correction radiométrique est la seule solution pour les éliminer. Le succès de cette correction est variable mais permet, avec la détection automatique des ombres, d'augmenter l'interprétabilité des images ;

- Pour la gestion des ombres projetées hors de la forêt, une solution qui a été utilisée dans le présent travail est la création d'un masque Forêt/Hors forêt. Une classification à deux niveaux peut aider à limiter les erreurs. Un masque efficace peut être réalisé sur base d'un Modèle Numérique de Hauteur actualisé à haute résolution ainsi que d'indices de végétation et de paramètres de forme et de texture pour discerner la forêt des autres végétations.

Dans ce travail, les ombres ont été considérées comme un problème à régler. Mais, elles peuvent aussi être une source d'information. Certains chercheurs utilisent les ombres dans des algorithmes pour créer une segmentation à l'échelle du houppier (Warner *et al.*, 2006).

- Cas particulier des coupes rases

Le problème majeur lié aux mises à blanc est leur caractère temporaire et dynamique. En effet, l'exploitation forestière peut coucher un peuplement en l'espace de quelques jours, voire quelques heures. Puis, assez rapidement, une végétation spontanée ou une plantation d'essences de production occupera ce sol. A chaque stade, la réponse en télédétection est très différente. Premièrement, la hauteur chute instantanément d'environ 20-30 mètres à 0 mètres puis augmente avec la croissance de la végétation en place. Ensuite, la réponse spectrale qui était bien spécifique à une essence devient celle d'un sol nu puis celle d'un jeune mélange d'herbacées et ligneux divers. Enfin, la texture, qui pouvait être typique d'un peuplement mature d'une essence, devient un signal peu exploitable.

La disponibilité actuelle en données de télédétection n'est pas suffisante pour détecter les mises à blanc de manière efficace. Les données proviennent de différentes dates et sont obsolètes pour détecter certaines classes d'occupation du sol. Ceci nous oblige à utiliser des informations contradictoires pour définir une classe homogène. Il est très clair qu'une mise à blanc qui a eu lieu entre 2012 et 2016 aura une ortho-image correspondant au peuplement en 2012 sur pied encore en place et un MNH associé au peuplement toujours présent ou pas en 2014. Les réponses spectrales des images Sentinel-2 datent de mai et juillet 2016. A titre d'exemple, le cas d'une mise à blanc entre 2014 et 2016 est illustré en 2.2 dans l'extrait de la zone d'étude présenté en Figure 6, Figure 7 et Figure 9. Ces données non synchrones amènent de l'hétérogénéité dans la définition de la classe «Coupe rase». Ceci explique également pourquoi les classes «Régénération» et «Sol nu» sont les plus erronées.

Diverses pistes de solutions existent mais toutes omettent à un moment ou à un autre une partie des zones mises à blanc. Par exemple, un masque de seuil de MNH. La seule solution correcte est de travailler avec des données récoltées au même moment par toutes les sources pour refléter l'état réel de la forêt en un temps T. La cartographie qui pourra être produite dans ce cas sera donc valable pour ce temps T.

4 Perspectives

Cette partie aborde l'enjeu des outils en libre accès, du déploiement de la méthode actuelle et des perspectives liées à de nouvelles données. Il est également question du développement d'autres applications d'approche quantitative, de cartographie des arbres hors forêt et de cartographie dynamique.

4.1 100% en libre accès

Une perspective de ce travail est l'implémentation d'outils en libre accès pour créer une méthodologie totalement indépendante de logiciels commerciaux.

La segmentation pourrait être réalisée avec l'algorithme Meanshift⁴² disponible sur OrfeoToolBox (OTB), au lieu d'eCognition. OTB, un outil en libre accès qui évolue en continu grâce aux packages développés par des utilisateurs. OTB a été envisagé dans le cadre de ce travail mais est toutefois peu intuitif et son utilisation nécessite un temps de formation. De plus, il ne permet pas d'attribuer un poids aux bandes. Il faut donc trouver un artéfact: par exemple, redéfinir les gammes des unités radiométriques de chaque bande de sorte qu'une différence absolue soit plus ou moins rapidement mise en évidence.

4.2 Déploiement de la méthode

La méthodologie mise au point peut également être généralisée à d'autres communes pilotes dans d'autres régions, à condition d'acquérir suffisamment de données de référence (parcellaire DNF ou données de terrain). Il serait également intéressant de voir comment le modèle développé pour Paliseul se comporte ailleurs en Ardenne ou en Région wallonne. Il serait peut-être nécessaire de revoir et adapter la classification en fonction des essences dominantes de la nouvelle zone étudiée.

4.3 Cartographie dynamique

Si l'on vient à disposer des données de parcellaire mises à jour, de nouvelles orthophotos réalisées en période de végétation, d'un nouveau Modèle Numérique de Hauteur et d'une base de données d'images Sentinel-2, on peut imaginer créer des séries saisonnières à multi-annuelles, en fonction du phénomène d'étude. Cela peut conduire à une cartographie très dynamique qui révélerait la dynamique spatio-temporelle de l'écosystème «forêt». Un monitoring de la forêt permettrait de mettre en évidence les effets du changement climatique (Lindner *et al.*, 2010), la détection de problèmes sanitaires, la caractérisation de l'ampleur de chablis, la gestion des mises à blanc et des problèmes d'érosion qui y sont liés, la replantation ou la mise en culture, voire même la détection de plantations de sapins de Noël.

4.4 Approche quantitative

L'approche de ce travail se veut qualitative avec la création d'une cartographie des essences forestières à l'échelle des peuplements. Un autre sujet de recherche très actuel est la caractérisation quantitative de la ressource ligneuse à l'aide de la télédétection. De nouvelles applications utilisent un LiDAR à plus haute résolution afin de générer un modèle en trois dimensions de la forêt. Ceci permet notamment d'estimer un grand nombre de variables dendrométriques, par exemple, le volume sur pied et la surface terrière. Le LiDAR est plus puissant pour estimer la biomasse et le stock carbone en forêt contrairement aux images hyper-spectrales qui offrent de meilleurs résultats dans la différenciation des essences (Koch, 2010). Ainsi, la combinaison des deux méthodes permettrait d'estimer un volume sur écorce par essence.

⁴² Plus d'information sur le Meanshift dans l'article de Comaniciu et Meer (2002).

4.5 Cartographie des arbres hors forêt

L'objectif de ce travail est la cartographie des essences forestière. Néanmoins, certaines catégories d'arbres en dehors des forêts méritent une attention toute particulière: les arbres remarquables d'intérêt patrimonial, les plantations de peupliers, les arbres le long des voies de transport (autoroutes et voies de chemin de fer, par exemple). Afin de les cartographier de manière efficace, on pourrait créer un masque sur l'indice de photosynthèse (NDVI) ainsi qu'un masque de forme pour les séparer de la "forêt" à proprement parler. Il faudrait également tester une segmentation à l'échelle des houppiers et entraîner le modèle sur de nouvelles classes et d'autres données.

4.6 Suivi des coupes rases et des replantations

L'étude qualitative et quantitative de la dynamique forestière est nécessaire à une bonne gestion de la ressource. Comme il a déjà été évoqué en 4.3, la mise à jour des données et la création de séries multi-annuelles pourrait rendre cela possible et permettre la création d'un outil à l'intention du gestionnaire. Cet outil pourrait également être utile aux politiques qui se sont lancés dans une campagne d'aide financière à la replantation. En effet, on observe depuis quelques années qu'une proportion non négligeable de zones forestières, dans les petites propriétés privées, ne sont pas replantées après exploitation (Portail de la Wallonie, 2016)⁴³. Une carte dynamique des peuplements forestiers pourrait mettre en évidence des zones de fragmentation de l'habitat.

4.7 Caractérisation de l'habitat

Sur base des cartes de sols et des pentes mais également de la caractérisation qualitative et quantitative de la ressource ligneuse, on peut identifier des zones forestières dont l'exploitation n'est pas rentable. De plus, une cartographie des essences sur base d'information de télédétection (objet de ce travail) peut nous renseigner sur l'accueil potentiel de la biodiversité (Figure 34). En effet, de nombreuses espèces d'insectes, d'oiseaux, de champignons et de lichens sont les hôtes d'essences ligneuses bien spécifiques.

⁴³ <http://www.wallonie.be/fr/actualites/la-wallonie-favorise-le-reboisement-de-nos-forets> aspect politique (consulté le 3 mai 2017)

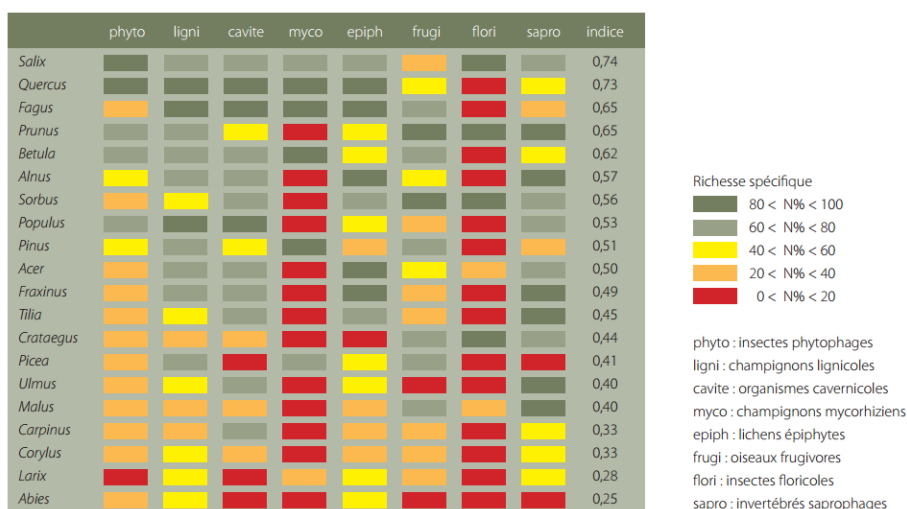


Figure 34: Classement des principales espèces ligneuses d'Europe occidentale d'après leur potentiel biologique (en % de la richesse spécifique maximale par groupe fonctionnel) ; source: Branquart et Dufrêne, 2005 in Branquart et Liégeois, 2005

Enfin, on peut encore tirer des informations extrêmement utiles de l'approche quantitative avec le LiDAR. En effet, on peut identifier la structure verticale du couvert, estimer le volume sur pied et les dimensions des arbres. Du point de vue de l'accueil de la biodiversité, les arbres de plus grosses dimensions ont plus de probabilité de développer des cavités, micro-habitats indispensables à près d'une espèce forestière sur quatre (Branquart E. et Liégeois S., 2005). De plus, le LiDAR haute densité (450 points/m² pour une résolution de 5 cm) permettrait d'identifier les troncs au sol et ainsi de tenter une estimation d'un volume de bois mort au sol (Figure 35).

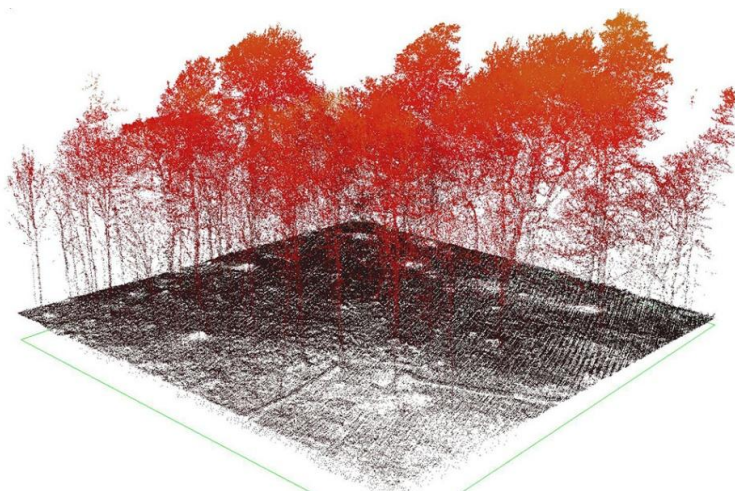


Figure 35: Réponse du LiDAR haute fréquence (en moyenne 450 points par mètre carré) à travers le couvert forestier avec une résolution de 5 cm ; source: FlyThru⁴⁴

Cette application pourrait participer à la modélisation des habitats d'espèces forestières (Vierling *et al.*, 2008). Ainsi, on peut éventuellement détecter et sélectionner des sites qui pourraient potentiellement être dédiés à la biodiversité et assurer à la forêt un rôle multifonctionnel. Néanmoins, cette technologie LiDAR haute densité est peu accessible en raison de son coût. Le Service Public de Wallonie n'a pas les moyens de fournir cette donnée sur le territoire wallon.

⁴⁴ <http://www.flythru.co.uk/gallery-lidar/10-large.jpg>

5 Conclusion

L'objectif de ce travail est la cartographie forestière d'une commune de Wallonie (Paliseul) par télédétection selon une approche multi-capteurs. La carte est générée à destination du gestionnaire forestier et de la filière bois, dans un but de caractérisation de la ressource ligneuse. Cette méthodologie, alliant les données multi-spectrales provenant des ortho-images, les données de hauteurs issues du LiDAR et les images du satellite Sentinel-2A, tire profit des résolutions des différentes sources de données. L'approche de classification par objet à l'aide du Random Forest est un outil puissant. Le résultat final est une carte précise de l'occupation du sol avec une erreur du producteur moyenne estimée à 73,6% (entre 47% et 100% selon la classe forestière) et une erreur du consommateur moyenne estimée à 62,1% (entre 3% et 95% selon la classe forestière). Les confusions entre les classes peuvent être expliquées sur le terrain et concernent principalement des peuplements ouverts avec une végétation de sous-bois ou des peuplements mélangés. Les pistes d'amélioration sont nombreuses et concernent notamment l'acquisition de données synchrones et mises à jour ainsi qu'une révision de la légende cartographique et un meilleur choix des zones d'entraînement. L'évaluation de la qualité de la segmentation et son amélioration comme la sélection de variables pour la classification sont également des étapes à repenser. La perspective d'une cartographie dynamique pour le suivi régulier et précis de nos forêts est un défi très prometteur. Le développement d'autres applications utiles à la gestion durable du territoire et de la ressource est également une voie d'avenir.

Bibliographie

- BENZ U.C. et al., 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information, *ISPRS Journal of Photogrammetry and Remote Sensing*, 58, 23p.
- BLACKBURN G.A., 1998. Quantifying Chlorophylls and Carotenoids at Leaf and Canopy Scales: An Evaluation of Some Hyperspectral Approaches. *Remote Sensing of Environment*, 66, 273-285
- BLASCHKE T., 2004. Object-based contextual image classification built on image segmentation. In: Advances in Techniques for Analysis of Remotely Sensed Data, 2003, *IEEE*, 113-119
- BRANQUART E. & DUFRENE M., 2005. Les arbres, de puissants révélateurs de la biodiversité forestière. In : Branquart É. et al. (eds). Résumé des interventions de la journée d'étude « Gestion forestière et biodiversité », Gembloux, 23 mars 2005
- BRANQUART E. & LIÉGEOIS S., 2005. Normes de gestion pour favoriser la biodiversité dans les bois soumis au régime forestier (complément à la circulaire n° 2619). Ministère de la Région wallonne, Direction générale des Ressources naturelles et de l'Environnement, 84p.
- BREIMAN L., 2001. Random forests. *Machine Learning*, 45(1), 5-32
- CHEN P.f. et al. 2010. New Index for Crop Canopy Fresh Biomass Estimation. *Spectroscopy and Spatial Analysis*, 30, 512-517
- COMANICIU D. & MEER P., 2002. Mean Shift: A Robust Approach Toward Feature Space Analysis. *IEEE Trans. PAMI*, 24 (5), 603-619
- CONGALTON R. & GREEN K., 1999. Basic analysis techniques. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*.
- CIPAR J. et al., 2004. Distinguishing Between Coniferous and Deciduous Forests Using Hyperspectral Imagery. *IEEE*, 2348-2351
- DARE P.M., 2005. Shadow Analysis in High-Resolution Satellite Imagery of Urban Areas. *Photogrammetric Engineering & Remote Sensing*, 71 (2), 169-177
- DARWISH et al., unknown. Image segmentation for the purpose of object-based classification. *eCognition*, 3p.
- DATT B., 1999. Visible/near infrared reflectance and chlorophyll content in Eucalyptus leaves. *International Journal of Remote Sensing*, 20, 2741-2759
- DEDRY L. et al. 2015. ForEstimator : un nouvel outil cartographique pour mieux connaître la forêt wallonne. *Forêt.Nature*, 135, 42-46
- DEFRIES R. & TOWNSHEND J., 1994. NDVI-derived land cover classifications at a global scale. *International Journal of Remote Sensing*, 15(17), 3567-3586
- DOMENECH E. & MALLET C., 2014. Change Detection in High resolution land use / land cover geodatabases (at object level). *EuroSDR official publication*
- DONOGHUE D. et al., 2007. Remote sensing of species mixtures in conifer plantations using LiDAR height and intensity data. *Remote Sensing of Environment*, 110, 509-522

- DOSTALOVA A. et al., 2016. Forest area derivation from Sentinel-1 data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3(7), 227-233
- FILELLA I. & PENUELAS J., 1994. The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. *International Journal of Remote Sensing*, 15, 1459-1470
- GAO B.c., 1966. NDWI - A normalised difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58, 257-266
- GITELSON A.A. et al., 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*, 58, 289-298
- Gestion des Ressources Forestières et des Milieux Naturels - Unité de Géomatique (ULg - Gembloux Agro-Bio Tech), 2010. Télédétection spatiale de l'occupation du sol en Wallonie. 202p.
- HAY G.J. et al., 1996. An object-specific image texture analysis of H-resolution forest imagery. *Remote Sensing of Environment*, 55, 108-22
- HUETE A.R., 1988. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25, 295-309
- IMMITZER M. et al., 2016. First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote Sensing*, 8 (3), 27p.
- JACQUES D.C. et al., 2014. Monitoring dry vegetation masses in semi-arid areas with MODIS SWIR bands. *Remote Sensing of Environment*, 153, 40-49
- JENSEN J.R., 2007. *Remote Sensing of the Environment: An Earth Resource Perspective*, 2nd ed. ed.; Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2007
- JIANG R. et al., 2009. A random forest approach to the detection of epistatic interactions in case-control studies. *BMC Bioinformatics*, 10(1), 65p.
- KOCH B., 2010. Status and future of laser scanning, synthetic aperture radar and hyperspectral remote sensing data for forest biomass assessment. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65, 581-590
- LACAUX J. et al., 2007. Classification of ponds from high-spatial resolution remote sensing: Application to Rift Valley Fever epidemics in Senegal. *Remote Sensing of Environment*, 106, 66-74
- LE MAIRE G. et al., 2004. Towards universal broad leaves chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. *Remote Sensing of Environment*, 89, 1-28
- LIAW A. & WIENER M., 2002. Classification and Regression by randomForest. *R News*, 2(3), 18-22
- LICHTENTHALER H.K. et al., 1996. Detection of Vegetation Stress Via a New High Resolution Fluorescence Imaging System. *Journal of Plant Physiology*, 148, 599-612
- LINDNER M. et al., 2010. Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems. *Forest Ecology and Management*, 259, 698-709

- LU D. & WENG Q., 2007. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 18(5), 823-870
- MENZE B.H. et al., 2007. Multivariate feature selection and hierarchical classification for infrared spectroscopy: serum-based detection of bovine spongiform encephalopathy. *Anal Bioanal Chem*, 387(5), 1801-1807
- MENZE B.H. et al., 2009. A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinformatics*, 10 (1), 213-228
- MICHEZ A. et al., 2017. Caractérisation des forêts à l'aide de la technologie LiDAR. IN: BAGHADADI N. & ZRIBI M., 2017. Observation des surfaces continentales par télédétection I: Agriculture et Forêt. *Collection système terre - environnement: Série Télédétection pour l'observation des surfaces continentales*. ISTE Editions Ltd., 321-348
- MOHANAIAH P. et al., 2013. Image Texture Feature Extraction Using GLCM Approach. *International Journal of Scientific and Research Publications*, 3 (5), 5p.
- PAL R. & PAL K., 1993. A review on image segmentation techniques. *Pattern Recognition*, 26 (9), 1277-1294
- PINTY B. & VERSTRAETE M.M., 1992. GEMI: a non-linear index to monitor global vegetation from satellites. *Vegetation*, 101, 15-20
- RADOUX J. et al., 2016. Sentinel-2's Potential for Sub-Pixel Landscape Feature Detection. *Remote Sensing*, 8, 87-93
- RENAUD J-P. et al., 2015. Apport de modèles numériques de hauteur à l'amélioration de la précision d'inventaires statistiques forestiers. *Revue Française de Photogrammétrie et de Télédétection*, 211-212, 43-51
- RODRIGUEZ-GALIANO V.F. et al., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, 93-104
- ROSENFELD G. H. & FITZPATRICK-LINS K., 1986. A coefficient of agreement as a measurement of thematic classification accuracy. *Photogrammetric Engineering and Remote Sensing*, 52, 223-227
- SHAHI K. et al., 2015. A novel spectral index to automatically extract road networks from WorldView-2 satellite imagery. *The Egyptian Journal of Remote Sensing and Space Science*, 18, 27-33
- SRIPADA R.P. et al., 2006. Aerial Color Infrared Photography for Determining Early In-Season Nitrogen Requirements in Corn. *Agronomy Journal*, 98, 968
- SVETNIK V. et al., 2003. Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling. *J. Chem. Inf. Comput. Sci.*, 43 (6), 1947-1958
- TUCKER C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8, 127-150

VAN DER PERRE R. et al., 2015. La carte bioclimatique de Wallonie : un nouveau découpage écologique du territoire pour le choix des essences forestières, *Forêt.Nature*, 135, 47-58

VAN DEVENTER A.P. et al., 1997. Using Thematic Mapper data to identify contrasting soil plains and tillage practices. *Photogrammetric Engineering and Remote Sensing*, 63, 87-93

VIERLING K. et al., 2008. Lidar: shedding new light on habitat characterization and modeling. *Frontiers in Ecological Environment*, 6(2), 90–98

VOGELMANN J.E. & ROCK B.N., 1985. Spectral Characterization of Suspected Acid Deposition Damage in Red Spruce (*Picea rubens*) Stands from Vermont. *Proc. of Airborne Imaging Spectrometer Data Anal Workshop, United States*, 51-55

WANG J. et al., 2004. Relations between NDVI and tree productivity in the central Great Plains. *International Journal of Remote Sensing*, 25 (16), 3127–3138

WANG L. et al., 2004 Comparison of IKONOS and QuickBird images for mapping mangrove species on the Caribbean coast of panama. *Remote Sensing of Environment*, 91, 432–440.

WARNER T. et al., 2006. Segmentation and classification of high resolution imagery for mapping individual species in a closed canopy, deciduous forest. *Science in China: Series E Technological Sciences* 2006, 49 (1), 128—139

WASKE B. & BRAUN M., 2009. Classifier ensembles for land cover mapping using multitemporal SAR imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64, 450-457

WULF H. & STUHLER S., 2015. Sentinel-2: Land Cover, Preliminary User Feedback on Sentinel-2A Data. In: Proceedings of Sentinel-2A Experts Users Technical Meeting, Frascati, Italy, 29-30 September 2015

Sites internet

Aibus Defense and Space, 2017. Images satellites Pléiades - La constellation Très Haute Résolution. Disponible sur:
<http://www.intelligence-airbusds.com/fr/237-pleiades-imagerie-satellite-tres-haute-resolution>
(consulté le 12 mai 2017)

BARRE A. & DOLLET L. Analyse de texture par Fourier. Disponible sur:
http://www.tsi.telecom-paristech.fr/pages/enseignement/ressources/beti/textures_Fourier/texture.html (consulté le 25 mai 2017)

European Commission, 2017. Copernicus: The European Earth Observation Programme. Disponible sur:
<http://ec.europa.eu/growth/sectors/space/copernicus/> (consulté le 22 avril 2017)

EuropeanSpaceAgency (ESA), 2017. Copernicus Open Access Hub. Disponible sur:
<https://scihub.copernicus.eu> (consulté en février 2017)

EuropeanSpaceAgency (ESA), 2017. Histoire de l'observation de la Terre. Disponible sur:
http://www.esa.int/SPECIALS/Eduspace_FR/SEMUX1S7NWF_0.html (consulté le 22 avril 2017)

EuropeanSpaceAgency (ESA), 2017. Sentinel-2 mission. Disponible sur:
<https://sentinel.esa.int/web/sentinel/missions/sentinel-2> (consulté le 22 avril 2017)

EuropeanSpaceAgency (ESA), 2012. Sentinel-2, ESA's Optical High-Resolution Mission for GMES Operational Services. Disponible sur:
https://sentinel.esa.int/documents/247904/349490/S2_SP-1322_2.pdf (consulté le 22 avril 2017)

ESRI, 2017. About ArcGIS. Disponible sur:
<http://www.esri.com/arcgis/about-arcgis> (consulté le 22 avril 2017)

FAO, 2017. Situation des forêts du monde 2016. Disponible sur:
<http://www.fao.org/publications/sofo/2016/fr/> (consulté le 19 mai 2017)

FLYTHRU. LiDAR through vegetation. Disponible sur:
<http://www.flythru.co.uk/gallery-lidar/10-large.jpg> (consulté le 28 avril 2017)

Géoportail de la Wallonie, 2017. Disponible sur:
www.geoportail.wallonie.be/georeferentiel/orthophotos (consulté le 22 avril 2017)

INTERREG VA Grande Region | Großregion, 2016. Comité de sélection INTERREG : 21 nouveaux projets transfrontaliers pour la Grande Région. Disponible sur:
<https://www.interreg-gr.eu/news/interreg-v-a-grande-region-21-nouveaux-projets-transfrontaliers-pour-la-grande-region/> (consulté le 9 mai 2017)

Interreg V France-Wallonie-Flandre Feel Wood, unknown. Forêt Pro Bos: Valoriser les circuits courts pour la filière bois – Bosbouw van korte omloop. Disponible sur:
<https://foretprobos.wordpress.com/> (consulté le 9 mai 2017)

Les Plus Beaux Villages de Wallonie a.s.b.l., 2017. Our. Disponible sur:
<http://www.beauxvillages.be/les-villages/our.htm?lng=fr> (consulté le 3 mai 2017)

NASA. MODIS Moderate resolution imaging spectroradiometer. Disponible sur:
<https://modis.gsfc.nasa.gov> (consulté le 17 mai 2017)

OrfeoToolBox, 2017. CookBook: Segmentation. Disponible sur:
https://www.orfeo-toolbox.org/CookBook/Applications/app_Segmentation.html (consulté le 22 avril 2017)

OrfeoToolBox, 2017. Open Source processing of remote sensing images. Disponible sur:
<https://www.orfeo-toolbox.org/> (consulté le 23 avril 2017)

OSGeoLive, 2017. Quantum GIS (QGIS). Disponible sur:
https://live.osgeo.org/fr/overview/qgis_overview.html (consulté le 23 avril 2017)

PARKAN M., 2017. LiDAR et Forêts. Disponible sur:
<https://sites.google.com/site/lidaretforet/outils-et-methodes/interpretation-et-utilisation-des-produits-drivs> (consulté le 22 avril 2017)

Portail de la Wallonie, 2016. La Wallonie favorise le reboisement de nos forêts. Disponible sur:
<http://www.wallonie.be/fr/actualites/la-wallonie-favorise-le-reboisement-de-nos-foretsaspect-politique> (consulté le 3 mai 2017)

Ressources naturelles Canada. Notions fondamentales de télédétection. Disponible sur:
http://www.rncan.gc.ca/sites/www.nrcan.gc.ca/files/earthsciences/pdf/resource/tutor/fundam/pdf/fundamentals_f.pdf (consulté le 17 mai 2017)

RTBF, 2017. Montignies-sur-Roc et Our, nouveaux «Plus Beaux villages de Wallonie». Disponible sur:
https://www.rtbef.be/info/regions/detail_montignies-sur-roc-et-our-nouveaux-plus-beaux-villages-de-wallonie?id=9587169 (consulté le 4 mai 2017)

Service public de Wallonie, 2017. IPRFW. Disponible sur:
<http://environnement.wallonie.be/dnf/inventaire/> (consulté le 19 mai 2017)

Syndicat d'initiative de Paliseul asbl. Disponible sur:
<http://www.si-paliseul.be> (consulté le 1 mars 2017)

THEIA (CNES), 2017. Disponible sur:
<https://theia.cnes.fr/atdistrib/rocket/#/search?q=belgique> (consulté le 23 mars 2017)

The R Foundation, 2017. The R Project for Statistical Computing. Disponible sur:
<https://www.r-project.org/> (consulté le 22 avril 2017)

Trimble Geospatial, 2017. eCognition essentials. Disponible sur:
<http://www.ecognition.com/> (consulté le 22 avril 2017)

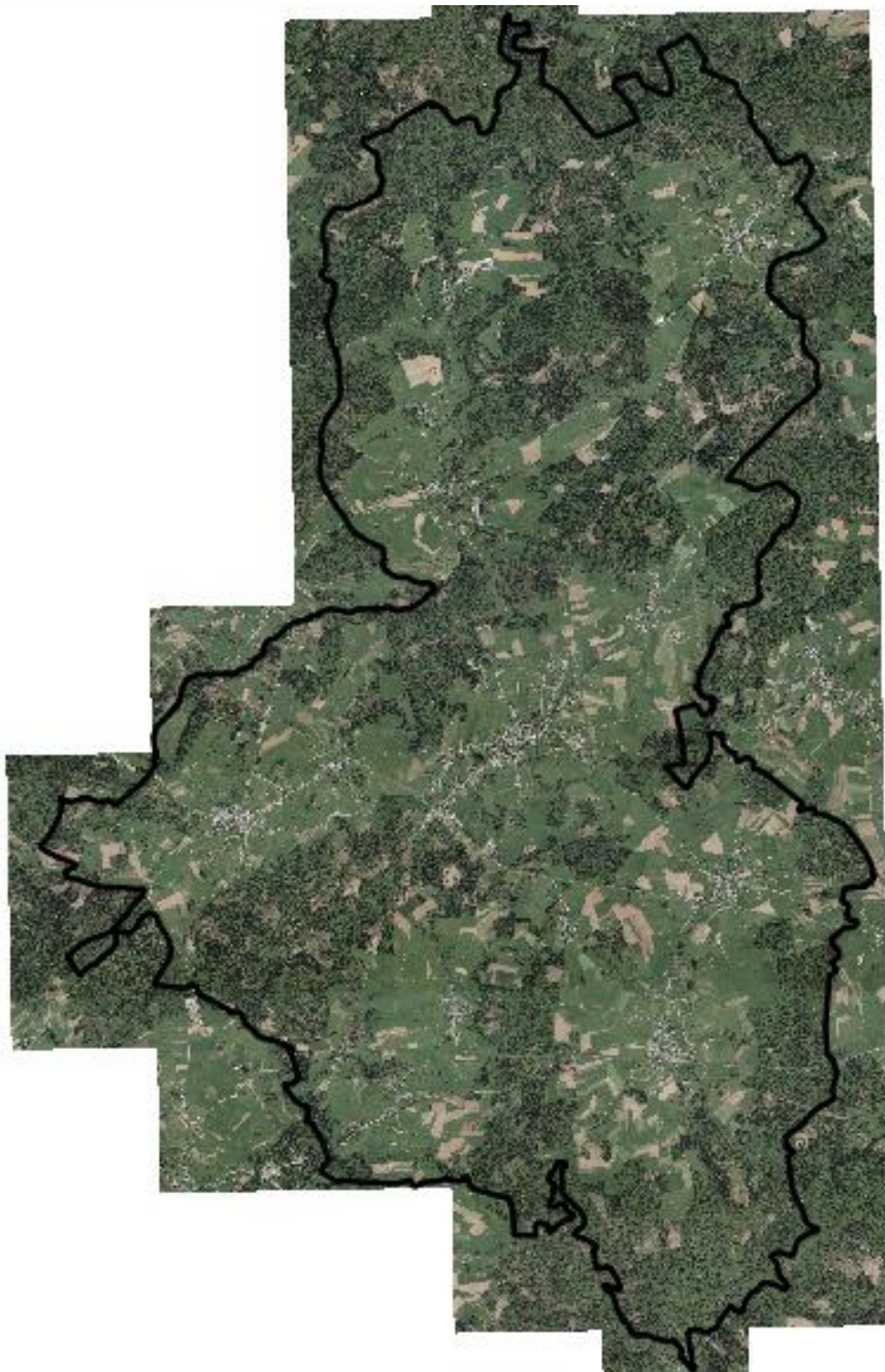
USFS Geospatial Technology and Applications Center (GTAC). Land Cover Classification Techniques supervised classification and random forests. Disponible sur:
https://servirglobal.net/Portals/0/Documents/Articles/ChangeDetectionTraining/Module3_LC_Classification_presentation.pdf (consulté le 28 mai 2017)

USGS, 2017. Earthexplorer. Disponible sur:
<https://earthexplorer.usgs.gov/> (consulté en février 2017)

UVED, 2008. ENVAL - Suivi de l'environnement par télédétection : environnement terrestre
Disponible sur:
<http://e-cours.univ-paris1.fr/modules/uvved/envcal/html/vegetation/role-teledetection/basse-resolution.html> (consulté le 9 mai 2017)

Annexes

Annexe 1: Orthophoto 2012-2013 sur la commune de Paliseul ; source: SPW



Annexe 2: Schéma général du traitement des données et de la classification

