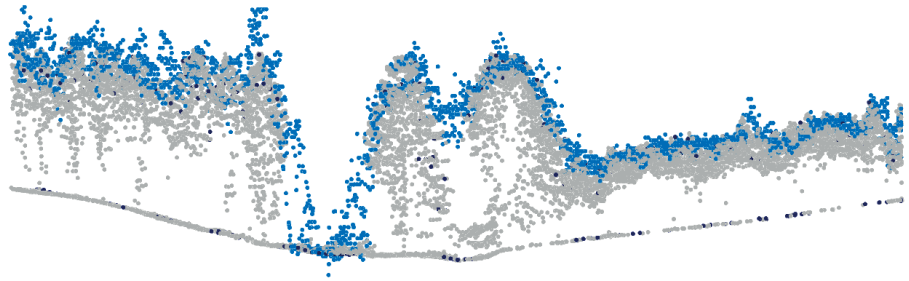
LIÈGE université
Gembloux
Agro-Bio Tech

Very High Resolution & 3D Remote Sensing Data for Supporting Forestry in Wallonia. What Resolution? What Scale? What Purpose?

Stéphanie Bonnet

Thesis submitted in fulfillment of the requirements for the degree
of Doctor (Ph.D.) in Agronomy and Bioengineering



COMMUNAUTÉ FRANÇAISE DE BELGIQUE
UNIVERSITÉ DE LIÈGE - GEMBLOUX AGRO-BIO TECH

Very High Resolution & 3D Remote Sensing Data for Supporting Forestry in Wallonia. What Resolution? What Scale? What Purpose?

Stéphanie BONNET

Thesis submitted in fulfillment of the requirements for the degree of Doctor (Ph.D.) in Agronomy and Bioengineering

Supervisor : Prof. P. Lejeune
Civil year 2018

Que dites-vous ? C’est inutile ?
Je le sais ! Mais on ne se bat pas
dans l’espoir d’un succès ! Non !
Non, c’est bien plus beau lorsque
c’est inutile !

Edmond Rostand

All that is gold does not glitter,
not all those who wander are lost.

JRR Tolkien

Abstract

BONNET Stéphanie (2018). *Very High Resolution & 3D Remote Sensing Data for Supporting Forestry in Wallonia. What Resolution? What Scale? What Purpose?* PhD Thesis. Université de Liège - Gembloux Agro-Bio Tech, Belgium. 152 p., 23 tabl., 50 fig.

The forest, which represents one-third of the Earth surface, has always been in close interaction with human societies. Forests provide many environmental services: soil and water resources protection, reduction of impact of gas emissions, biodiversity reservoir and conservation... Concerns about global changes add to social functions such as recreational activities or wood production and about forest management and silviculture practices. The need for information about the forest resource has never been greater to ensure a sustainable management. The description of the forest stands (location, extent, composition, structure ...) and their dendrometric characterization (dominant height, number of trees per hectare, mean quadratic circumference, basal area and volume per hectare ...) have become crucial prerequisites indispensable for planning. Remote sensing enables information acquisition over large areas and overcomes the capacity and accessibility limitations inherent in field data collection, which mobilizes significant human and financial resources. The Walloon forest is characterized by a great diversity of situations and preoccupations: type of owner, structure, composition, spatial distribution, stations, fragmentation, silvicultural practices. This diversity makes it particularly relevant to use remote sensing as an information acquisition tool. This thesis lies in this context. The main purpose was the implementation of different types of three-dimensional remote sensing data to determine how they could contribute properly to forest management in Wallonia. First, we considered the use of low-density and regional-scale Airborne Laser Scanning (ALS) data in combination with the data collected by the Walloon Regional Forest Inventory. This combination has shown its potential for quantifying coniferous forest attributes, coupled with regional allometry built on the inventory database. In parallel with the importance of coniferous forest, a major concern is the development of a close-to-nature forestry, promoting natural regeneration. The presence of gaps in the stands is therefore essential information. We focused on a robust mapping of gaps in broadleaved forests from high-density ALS data. To complete our progression in the use of three-dimensional data,

we focused on data acquired by drones. Our study showed their effectiveness in detecting (dominant, co-dominated, dominated) trees in coniferous stands as a basis for quantifying forest attributes. The potential use and relevance of several photogrammetric products (ortho- or rectified images, digital surface model, and correlation maps) have been tested to detect individual trees as a basis for forest characterization. We investigated the association of the individual tree detection approach and the area-based approach for the assessment of forest attributes with drone data. As a conclusion, we argue the importance of a strong interaction between the different kind of remote sensing data as indispensable support for forest managers.

Résumé

BONNET Stéphanie (2018). *Very High Resolution & 3D Remote Sensing Data for Supporting Forestry in Wallonia. What Resolution? What Scale? What Purpose?* PhD Thesis. Université de Liège - Gembloux Agro-Bio Tech, Belgium. 152 p., 23 tabl., 50 fig.

La forêt, qui représente un tiers de la surface de la Terre, a toujours été en étroite interaction avec les sociétés humaines. Les forêts fournissent de nombreux services environnementaux: protection des sols et des ressources en eau, réduction des impacts des émissions de gaz, réservoir de biodiversité et conservation ... Les inquiétudes relatives aux changements globaux s'ajoutent aux fonctions sociales telles que les activités récréatives ou la production de bois et aux préoccupations concernant la gestion des forêts et la sylviculture. Le besoin d'information sur la ressource forestière n'a jamais été aussi grand pour assurer une gestion durable. La description des peuplements forestiers (localisation, étendue, composition, structure ...) et leur caractérisation dendrométrique (hauteur dominante, nombre d'arbres par hectare, circonférence quadratique moyenne, surface terrière et volume par hectare ...) sont devenues des préalables indispensables à la planification. La télédétection permet une acquisition d'information sur des zones étendues et surmonte les limites de capacité et d'accessibilité inhérentes à la récolte de données sur le terrain, qui mobilise des ressources humaines et financières conséquentes. La forêt wallonne se caractérise par une grande diversité de situations et de préoccupations: type de propriétaire, structure, composition, répartition spatiale, stations, fragmentation, sylvicultures mises en place. Cette diversité rend particulièrement pertinente l'utilisation de la télédétection comme outil d'acquisition d'information. Cette thèse s'inscrit dans ce contexte. L'objectif principal est la mise en œuvre de différents types de données de télédétection tridimensionnelles afin de déterminer comment elles pourraient contribuer de manière appropriée à la gestion forestière en Wallonie. Nous avons premièrement considéré l'utilisation de données ALS (Airborne Laser Scanning) à basse densité et à l'échelle régionale en interaction avec les données collectées par l'inventaire forestier régional wallon. Cette combinaison a montré son potentiel pour la quantification des attributs de peuplements résineux, couplée à une allométrie régionale construite sur base des données de l'inventaire. En parallèle de l'importance de la forêt résineuse, une préoccupation majeure est le

développement d'une sylviculture proche de la nature, favorisant la régénération naturelle. La présence de trouées dans les peuplements est donc une information essentielle. Nous nous sommes dès lors concentrés sur une cartographie robuste des trouées en forêts feuillues à partir de données ALS haute densité. Pour terminer notre cheminement dans l'utilisation des données tridimensionnelles, nous nous sommes intéressés aux données acquises par les drones. Notre étude a montré leur efficacité à détecter les arbres (dominants, co-dominés, dominés) dans les peuplements résineux, comme base pour la quantification des attributs forestiers. L'utilisation potentielle et la pertinence de plusieurs produits photogrammétriques (images ortho- ou rectifiées, modèle numérique de surface, cartes de corrélation) ont été testées pour détecter des arbres individuels comme base pour la caractérisation forestière. Nous avons étudié l'association de l'approche de détection des arbres individuels et de l'approche par placette pour l'évaluation des attributs forestiers avec les données drone. En conclusion, nous soutenons l'importance d'une interaction forte entre les différents types de données de télédétection comme un appui indispensable aux gestionnaires forestiers.

Acknowledgements

"Who's the more foolish, the fool or the fool who follows him?"

Obi-Wan Kenobi

Les remerciements d'une thèse sont l'occasion d'évoquer les gens qui ont contribué, d'une manière ou d'une autre, à construire ce travail de longue haleine. La thèse est une aventure personnelle, mais c'est aussi (et surtout?) une opportunité de partager des tranches de vie (Y a un peu plus, j'vous l'mets quand même?). Écrire ces quelques lignes est également un moment de relâchement et de légèreté, bienvenu avant un atterissage parfois sous tension...

Pour commencer, je voudrais remercier le Professeur Jacques Rondeux qui, alors que j'étais encore Tfiste et bien éloignée de ses thématiques de recherche, m'a permis d'intégrer l'unité de Gestion des Ressources Forestières.

De manière pratico-pratique, c'est le Professeur Philippe Lejeune qui m'a donné la chance d'aller de l'avant dans le milieu de la télédétection forestière. Il y avait tout à construire à l'époque et cela a parfois été mouvementé. Nos orbites n'étaient jamais synchrones mais complémentaires, et cela m'a appris beaucoup...

Je remercie très chaleureusement les Professeurs Hugues Claessens, Jacques Hébert et Yves Brostaux, les autres piliers gembloutois de mon comité de thèse pour leurs commentaires avisés et leur soutien tout au long de ces nombreuses années. Merci également au Professeur Gilles Colinet, d'avoir accepté d'être président de mon jury.

Plus particulèrement, merci à Yves pour m'avoir donné l'opportunité de participer à la formation OpenSpat, qui fût très stimulante lors de la fin de thèse, mais qui m'a également permis de découvrir autrement de belles personnes. Au-delà de la thèse, je te remercie pour mon initiation aux jeux de société, d'être un bon partenaire de lourditude, pour ta lucidité et ton soutien dans bien des domaines, pour ton amitié en somme. May the Force be with you, always...

Mes remerciements vont également à Anne Joly. En tant que membre de mon jury, elle a apporté de nombreuses suggestions constructives à cette thèse, me permettant de l'améliorer, ce qui a été très motivant dans la dernière ligne droite.

> "Dans l'histoire d'une vie, on n'a jamais qu'un seul problème à résoudre, celui qui donne sens à notre existence et impose un style à nos relations."
>
> Boris Cyrulnik

Ces quelques lignes me donnent l'opportunité de remercier Richard Fournier de l'Université de Sherbrooke, également membre de mon jury et dont la contribution a été très précieuse pour l'amélioration de ce manuscript. Avant cela, il a accepté de m'accueillir pour un stage de trois mois dans son laboratoire en 2014. S'il ne s'en rend pas compte, je profite de ces remerciements pour lui dire que ce fût une expérience tellement enrichissante sur le plan personnel et professionnel. J'adresse également une pensée à son épouse, pour son accueil, sa bienveillance et le coeur qu'elle met dans les relations humaines!

I would also like to express my warmest thanks to Rachel Gaulton from Newcastle University for her patience and precious help. Without you, our gap mapping paper would not be what it is. I learned a lot thanks to you, from our meeting in Estonia and later, during our many exchanges to develop this paper! I hope to have the opportunity to renew this collaboration!

Un clin d'oeil féministe à Ludivine Lassois et Adeline Fayolle, merci à toutes les deux d'être de tels exemples (même sans vous en rendre compte) de caractère, d'authenticité et de réussite dans de nombreux domaines.

En descendant un peu plus bas dans la chaîne alimentaire de la recherche, je tiens à remercier chaleureusement mes collègues. Laurent Dedry, Jo Lisein, Gauthier Ligot, Jérôme Perin, Adrien Michez, Nicolas Latte, Peter Gaucher, Corentin Bolyn, Sébastien Bauwens,... De manière plus globale, merci à tous les collègues passés et présents pour toutes les bonnes tranches de rire, les soirées, sorties diverses et discussions parfois sans queue ni tête souvent graveleuses, tous ces exutoires bien nécessaires...

Il ne faut surtout pas oublier, ceux sans qui nos recherches n'auraient pas la base nécessaire, notre équipe technique active pour la récolte des données de terrain. Merci à Cédric Geerts, Alain Monseur, Thomas Baijot, Coralie Mengal, Amaury André, Benoît Mackels, Fred Henrotay, Adrien Schot, Julien Goijen de vous être toujours adaptés et rendus disponibles pour les campagnes de mesures parfois exigeantes.

Dans le même ordre d'idée, merci à Marie Fombona, Marie-Ange Golard et Jean-Yves De Vleeschouwer de faire tourner la boutique au quotidien.

Et parce qu'un jour les geeks hériteront du monde, je n'oublie pas de remercier les fous de barrettes de RAM : Mo Saidi, François Burnay, Astrid de Pierpont, Oliver de Thier, Mikhail Pitchugin, Ludwig Van Steen, Bernard Segers et Sam Quevauvillers (aka chaton). Une mention spéciale pour Tonton Bernard et Ludwig qui m'ont supporté (dans les deux sens du terme) lors de toutes ces heures de TP au SIG, et qui maintenant me cotoient dans leurs bureaux.

"Il est plus beau de transmettre aux autres ce qu'on a contemplé que de contempler seulement."

Saint Thomas d'Aquin

Au cours de ces années à Gembloux, j'ai eu la chance exceptionnelle, grâce à ma fonction d'assistante, de croiser le chemin de nombreux étudiants, certains sont même au fil du temps devenus des amis. Bac3, Ma1, Ma2, stagiaire, Tfiste, boulet, étoile montante, foireux, insolent, indépendant, dépendant, belette, chacun a apporté sa pierre à mon Inukshuk. En guise de clin d'oeil, je dirai que je pense souvent aux soupers et semaines forêt, aux salles informatiques mal aérées, aux vasques acrobatiques, aux affonds rocket, à Alf, aux attentions, aux "Touche pas à ça p'tit con", aux "Bonnet présidente", aux sextantes, aux petits chats qui lèvent la patte, aux pics de brochettes, au bain de feuilles mortes, aux "Ca va bien se passer", aux "Pas évident hein"... Sachez que je vous suis reconnaissante de ce que vous m'avez apporté. J'ai grandi, j'ai progressé, je me suis amusée dans et en dehors de mon travail avec et grâce à vous. Vous m'avez souvent motivée et rendue fière (euh parfois très énervée aussi...) et surtout vous m'avez permis de trouver un sens là où il ne semblait pas toujours y en avoir.

"On se demande parfois si la vie a un sens. Et on rencontre des êtres qui donnent un sens à la vie."

Brassaï

Cédric, Alain, Jean-Yves, merci pour nos correspondances délirantes et nos petits dîners italiens qui illuminaient mes semaines.

Cédric, je te suis reconnaissante pour ces années de collaboration et surtout de complicité. Tu es vraiment un cadeau!

Simon, Barbara et Manon, la Tropical Team, merci pour votre écoute inconditionnelle, vos conseils et votre enthousiasme sans borne...

Chez les Premières Nations du Nord-Ouest Pacifique, le colibri représente l'amitié, il nous apprend à apprécier la vie, à se concentrer sur la positivité. Les colibris sont très indépendants et semblent avoir une ressource d'énergie sans limite. Quel symbole pour toi! Merci pour ton amitié Mr Colibri. Et merci à la Meute pour cette amitié inattendue, laissant la part belle aux craquages...

Matthieu, cher ami, merci pour ton oreille attentive et compréhensive! Toujours là avec un bon mot, un sourire réconfortant et un bon jeu (de mots) de derrière les fagots...Vive les bébés koala!

Catherine, merci pour ton authenticité, ta présence et ton soutien! J'ai de la chance d'avoir fait ta connaissance, chère collègue de bureau qui a si souvent raison! Tu apportes une vraie plus-value à ma vie et j'espère que ce n'est que le début de notre amitié!

Blandine, jeune padawan, quelle recrue pour le côté obscur! Profite bien de ce qu'apporte la vie! Merci pour tous ces moments de complicité...

Isabelle, merci de m'avoir fait prendre conscience de ma résilience. Cette thèse c'est aussi grâce à toi...

Merci à mon Kern... Céline, Aurélie, Sandy, Anne-Lyse, François, Christelle (Popotte quoi!), Daniel. Difficile de résumer en quelques lignes ce que vous m'apportez...Vous êtes la famille que j'ai choisi, ma base, mes fondations. Depuis plus de 15 ans, nous avons partagé tant de choses, bonnes ou mauvaises, légères ou importantes, festives ou personnelles, une bonne partie de ce qui fait qui je suis aujourd'hui...

Merci à mes parents de croire en moi et d'être si fiers, depuis tout ce temps, en particulier, ces années gembloutoises.

Merci à Thibaut pour ce bout de chemin ensemble, pour le soutien ces derniers mois en particulier en ayant été un papa si présent...

Et enfin, merci à Line... C'est toi qui sera toujours mon nadir, et inconditionnellement ma plus belle réussite. Merci d'être mon Ground Control Point, mon petit soleil...

Contents

List of Figures

List of Tables

Abbreviations

3D	Tridimensional
ABA	Area-Based Approach
AIC	Akaike Information Criterion
ALS	Airborne Laser scanning
BCAA	Belgian Civil Aviation Authority
BIC	Bayesian Information Criterion
CBH	Circumference at Breast-Height
CHM	Canopy Height Model
CPI	Canopy Porosity Index
DBH	Diameter at Breast-Height
DCENN	Direction des Cours d'eau non navigables
dGPS	Differential Global Positioning System
DEM	Digital Elevation Model
DSM	Digital Surface Model
DTM	Digital Terrain Model
GIS	Geographic Information System
ESA	European Space Agency
GPR	Ground Points Ratio
GPS	Global Positioning System
ITD	Individual Tree Detection
LASER	Light Amplification by Stimulated Emission of Radiation
LiDAR	Light Detection And Ranging
PSP	Permanent Sample Plot
RFI	Regional Forest Inventory
RMSE	Root Mean Square Error
RS	Remote Sensing
SfM	Structure-from-Motion
SPW	Service Public de Wallonie
TLS	Terrestrial Laser Scanning
UAS	Unmanned Aerial System
UAV	Unmanned Aerial Vehicle

Chapter 1

INTRODUCTION

Chapter 1

Introduction

"Between two pines is a doorway to a new world."

John Muir

1.1 General Context

The forest, which represents one third of the Earth surface, has always been in close interaction with human societies. Forests provide many environmental services : soil and water resources protection, mitigation of local climate variation, reduction of impact of gas emissions, biodiversity reservoir and conservation [33, 57]. Therefore, current concerns about global change, deforestation and carbon sequestration are all issues facing the forest. These considerations must be taken into account in addition to the social functions in daily life for human beings as recreation or wood production, preoccupations about forest management and silviculture.

Forest management is all the necessary actions to reach the aims decided by the forest planning. To support the planning, an important descriptive and prospective phase is needed to define the different objectives, based on the initial situation. Silvicultural decision-making and definition of long-term goals for stands heavily rely on indicators of forest structure and density. Number of trees, basal area and volume are measures of stand density used to control level of

thinning and predict growth. Structure and diameter distribution, based on data collected at tree-level are also relevant indicators. Spatial variability of basal area at the stand-scale is useful for the forest manager, as a dendrometric indicator and a variable for forest typologies as silvicultural baseline. Basal area value can be a guideline for promoting regeneration of specific species.

The need for information about the forest resource has never been greater to ensure a sustainable management. The description of the forest areas (location, extent, composition, ...) and their dendrometric characterization have become crucial prerequisites for planning and are realized traditionally through field inventories. This work is carried out on potentially large surfaces and mobilizes significant human and financial resources. Remote sensing allows repeated acquisitions of information over extensive areas and overcomes the limits of capacity and accessibility inherent in field data prospecting. With the constant development of digital technologies, the use of remote sensing has become a central and indispensable player thanks to the development of efficient techniques adapted to the characterization of the forest resource and the development of new data, methodologies and tools.

This thesis lies in this context, with a specific focus on the Walloon forest case (Southern Belgium). Considering the broad issue of forest remote sensing, the purpose of this thesis was the use of several types data for the quantitative and qualitative characterization of forest stands considering different working scales, from local to regional. This introduction gives some key features about the Walloon forest resource and the need of information regarding the current concerns for forest managers. Some background is then provided about forest inventory and remote sensing data. Specificities of these data in the Walloon context are highlighted. Finally, the aims, structure and organization of the thesis are presented.

1.2 The Case of the Walloon Forest

1.2.1 *Key Features about the Resource*

The forest policy is handled at a regional level in Belgium (Wallonia, Flanders and Brussels Figure 1.1). Forest lies therefore within the Walloon government's competence. The forest is a major natural resource [76] in Wallonia (Figure 1.2 (a)), extending over 33% of the Region (556200 hectares divided into 475200 hectares of forest stands and 81000 hectares of non-productive areas as forest roads, firebreaks...) with 57% of broadleaved species and a proportion of 49% of public properties. The public forest (273700 hectares, 49%) is managed by the Forest Administration, a part of the regional Administration (Service Public de

Wallonie, SPW). The Walloon Region owns 67800 hectares and 194800 hectares belong to the municipalities. Norway spruce (*Picea abies* (L.) H. Karst.) stands represent 34% of the forest, oaks (*Quercus petraea* (Matt.) Liebl. and *Quercus robur* L.) stands 18% and 9% for beech (*Fagus sylvatica* L.). Oaks and Beech stands are together 57% of the broadleaved pubic forests while Norway Spruce and Douglas-fir represent 79% of the coniferous private forests (Figure 1.3). The public forest is composed of 1655 properties, with an average area of 170 hectares while the private forest is characterized by an average area of 3 hectares for 89790 properties.

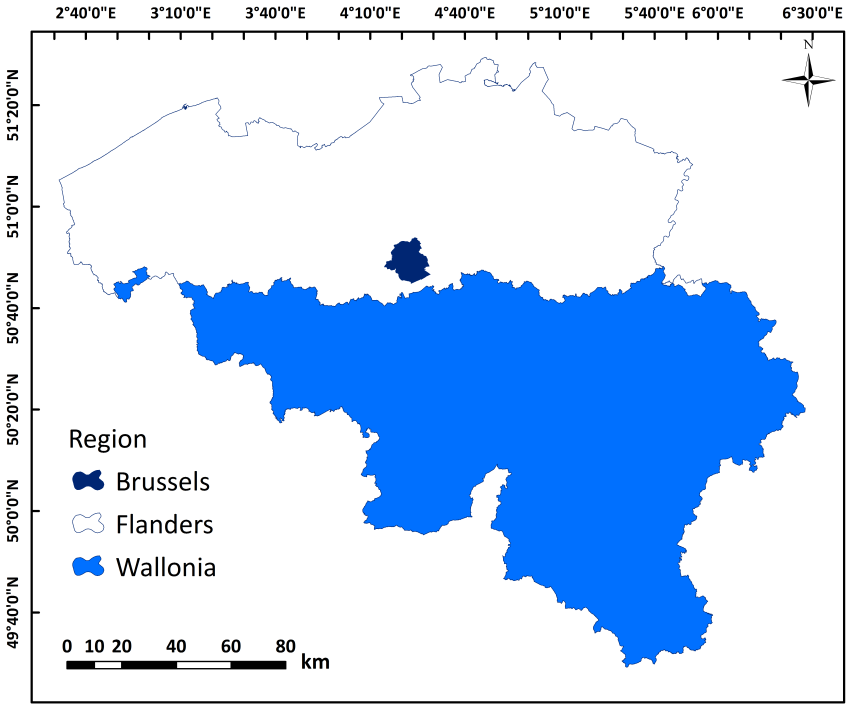


Figure 1.1: The forest policy is handled at a regional level in Belgium. The country is divided into three Regions: Wallonia, Flanders and Brussels.

There are geographical differences in the distribution of forest for example regarding public versus private properties or the types of stands (e.g. broadleaved versus coniferous). Environmental specificities are formalized with the concept of bioclimatic regions (Figure 1.2 (b) & (c)). The Walloon Region is splitted into 10 bioclimatic regions based on climatic data (temperatures, precipitation and solar radiation) and knowledge on species autecology [141]. The spatialization of the climatic indicators had led to an ecological stratification of the Walloon territory [141]. The extent and the composition of the stands have specificities according to the type of the owner, the province (Figure 1.4) or the bioclimatic regions. For example, the percentage of forest cover is 57,2% in Ardenne versus 23.5% in Condroz.

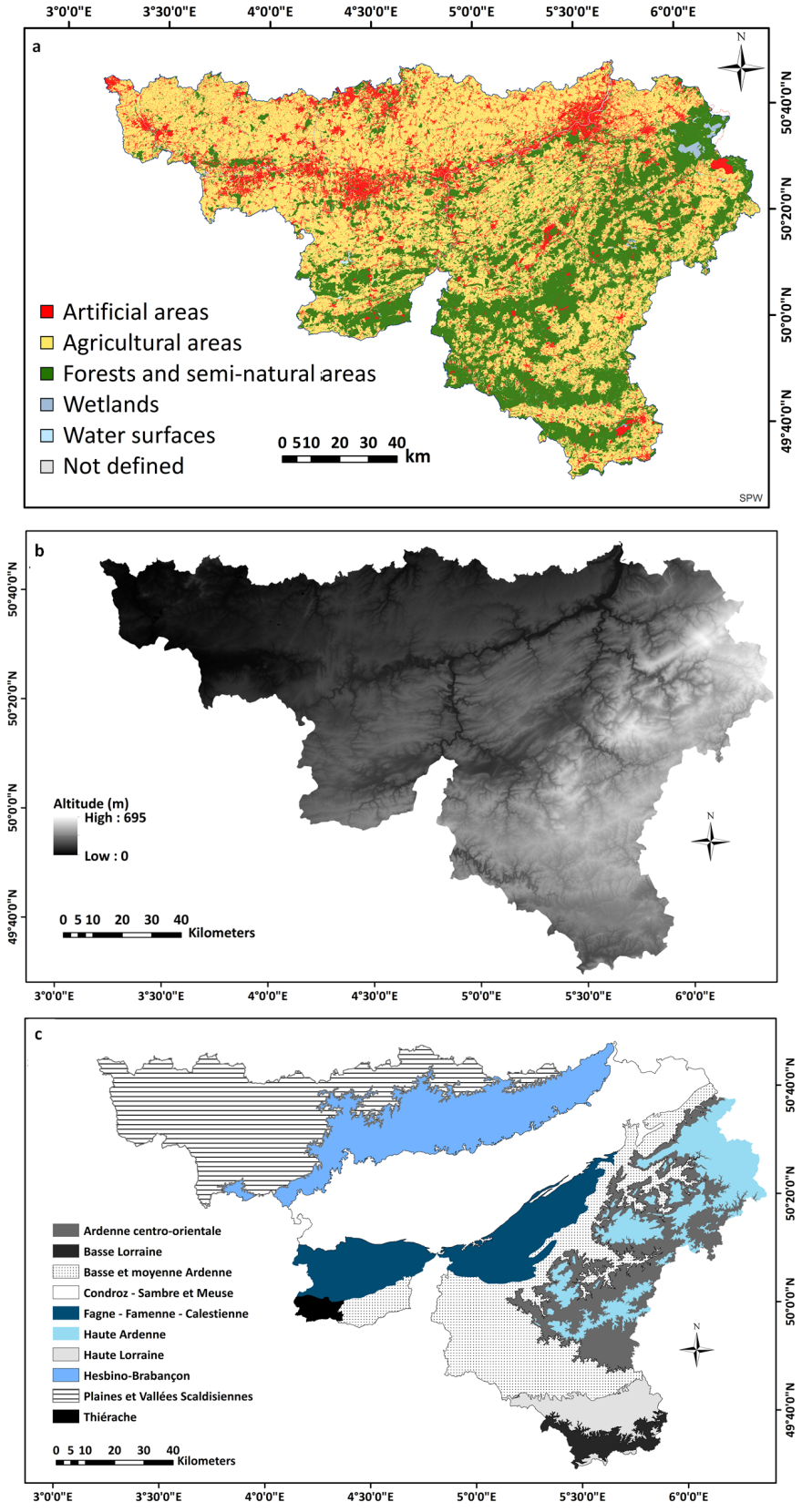


Figure 1.2: Land use/land cover map (a), gradient of altitude (b) and bioclimatic regions (c) of Wallonia. The forest is covers 33% of the Region which is splitted into 10 bioclimatic areas based on climatic & environmental factors and knowledge on species autecology.

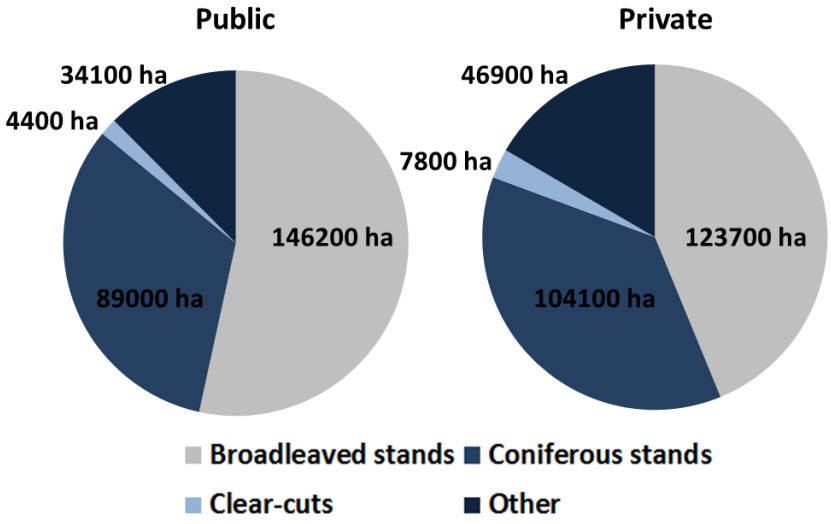


Figure 1.3: Proportion of forest types in Wallonia regarding the type of owner. The forest covers 33% of the Walloon territory (namely 556200 hectares divided into 475200 hectares of forest stands and 81000 hectares of non-productive areas (class "Other" as forest roads, firebreaks...).

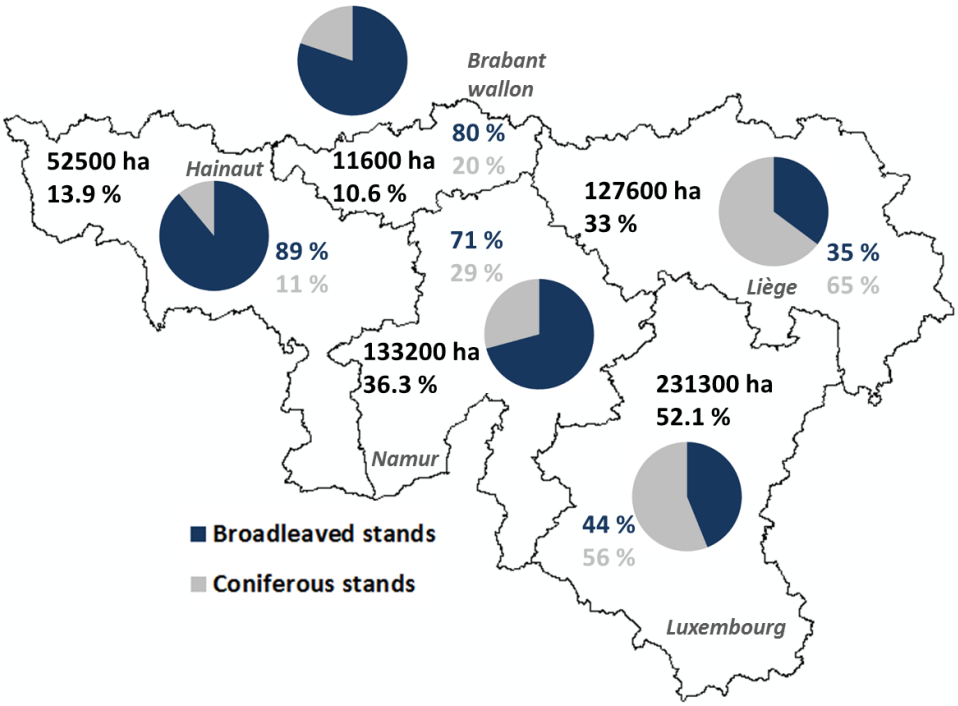


Figure 1.4: Proportion of forest types, forested area and forest cover percentage in Wallonia in each Walloon province (Brabant wallon, Hainaut, Liège, Luxembourg, Namur).

The broadleaved forest is diversified in terms of composition (Figure 1.5). Oaks are the most common ahead precious hardwoods species. The latter category includes mix of oaks and beech and pure or mixed stands of ash (*Fraxinus excelsior* L.), red oak (*Quercus rubra* L.), wild cherry (*Prunus avium* L.), elm (*Ulmus* sp.) or sycamore maple (*Acer pseudoplatanus* L.). Recent evolution shows a trend towards more diversification. Other broadleaved stands are composed with at least 50% deciduous trees but do not fit into other categories. Regarding coniferous trees, the Norway spruce stands are the majority despite a decline observed since the 1980s. Other coniferous stands are composed with at least 50% coniferous trees but do not fit into other categories. The evolution of the coniferous surfaces is more remarkable than that of the hardwoods [76]: strong decrease of pines (*Pinus* sp.), significant increase of Douglas-fir and disappearance of species outside the station following the application of the forest codes and Natura 2000 (mainly for the Norway spruce stands).

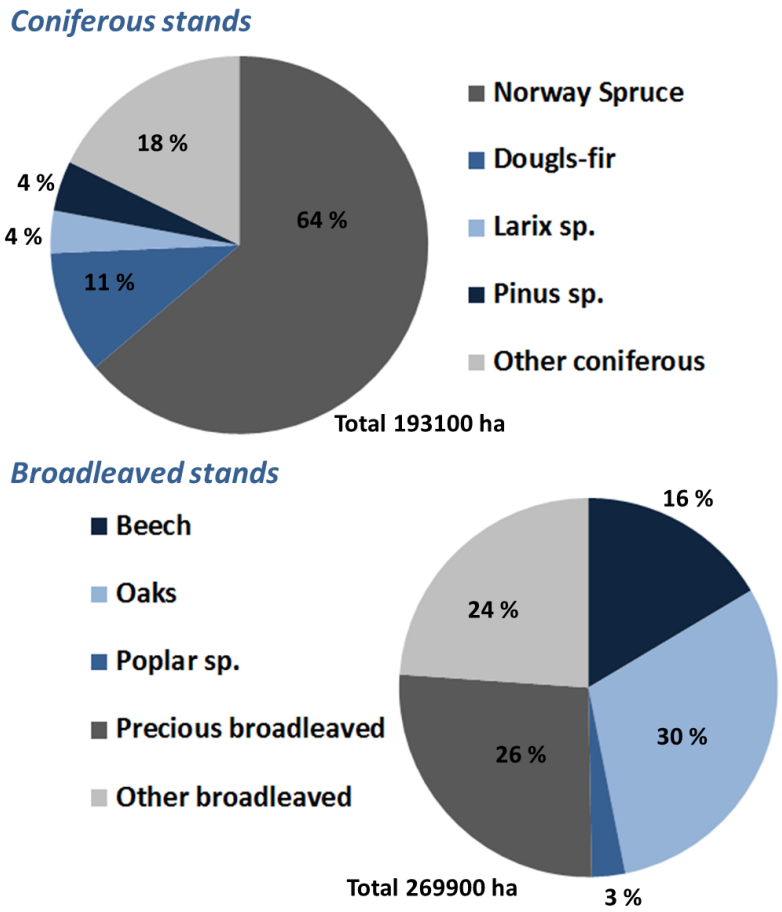


Figure 1.5: Proportion of the main species for coniferous or broadleaved stands. Precious hardwoods species include mix of oaks and beech and pure or mixed stands of ash, red oak, wild cherry, elm or sycamore maple. Other broadleaved or coniferous stands are composed with at least 50% deciduous ou coniferous trees but do not fit into other categories.

1.2.2 *Current Needs for Information*

The issues concerning the Walloon forest, its management and the threats it faces are varied. In order to highlight the specific context of Wallonia, and the need for information on the forest resource, this section briefly presents current major problems: diversification and transformation to uneven-aged stands, regeneration of broadleaved species (especially oaks), in the framework of resilience to global change [75, 161, 3].

The impact of climate change will be particularly marked by an increase in the growing season and the frequency and intensity of water stress or extreme weather events. Natural resource managers will have to adapt by favoring the resilience of stands by the choice of indigenous or acclimated species, by compositions, structures and biodiversity which are diversified. In this context, developing surveillance and characterization strategies for the Walloon forest is essential. A basic guideline must seek more resilient and stable ecosystems, for example by heading vertically stratified stands. Modeling and simulation tools can be used to evaluate the impact of different silvicultural factors and scenarios on stand evolution. These types of tool are based on a reliable quantitative assessment of the initial situation.

The Regional Forest Inventory (RFI) estimates more than 40.000 hectares of Norway spruce stands that will soon be regenerated in Wallonia. Often, natural regeneration is abundant and represents an opportunity to overcome clear-cutting, the reduction of the forest environment and the increasing risk of windfall in neighboring stands. At the same time, there is a clear trend to transform the even-aged forests into irregular stands, which are considered as more resilient to global changes. This transformation to uneven-aged stands is, however, largely unknown in the context of the Walloon forest. In order to compare different silvicultural scenarios to uneven-aged transformation, the monitoring of inventory plots with information on each individual tree by mean of remote sensing data, could be particularly interesting.

Different strategies exist for the transformation of coniferous forests to uneven-aged stands [129]. They are conditioned by the stage of development. The concept of precision forestry, involving the characterization of the resource at the tree level, therefore makes a lot of sense in the management of irregular forests. Standing capital, structure and composition are all guides for the forester in the procedure of transformation. Standing capital is often measured by the basal area, which can serve as a benchmark for determining the optimum of a stand. The structure and spatial distribution of the trees vertically (vertical stratification of heights) and horizontally (girth variability) are also elementary traits.

The dynamic of irregular stands is based on the natural regeneration of the forest. However, the high density of ungulates is an issue for the renewal of stands, especially those based on palatable species such as oak. Oak-Beech forests in Ardennes are more developed on acid soils and in conditions more favorable for the beech which becomes highly competitive and regenerates more effectively and is less depressed by deer. The oak allows the diversification of the beech forests and the improvement of the functioning of the ecosystem, in particular its resilience against the climatic and phytosanitary hazards. Unfortunately, the effective natural regeneration of oak, previously favored by its qualities and heritage value, is sporadic [3]. It is therefore important to look at gaps as potential areas for regeneration. The presence of gaps changes the amount of incident light reaching the ground and influences nutrient cycling and moisture availability, providing potential settlement sites for regeneration. The characterization of gaps is important to locate suitable conditions for the regeneration of target species. The study of forest gap dynamics is thus important for forest management. The development of remote sensing tools and in particular the use of aerial Light Detection And Ranging (LiDAR) technology offers opportunities for the forest domain in large-scale automatic gap detection.

1.3 Sources of Information

1.3.1 *Field Data*

Regional Forest Inventory

As a major natural resource, the Walloon forest must be managed sustainably and it's management has to be based on reliable and up-to-date data. The characterization of the Walloon forest is made through the RFI [127, 4], led by the Forest Administration. This structure is active since 1994 and allows establishing a complete inventory of the forest, theoretically, every 10 years (14 years in practice). The sampling method applied by the RFI is a single-phase, non stratified inventory using a systematic sampling design with Permanent Sample Plots (PSP). One-tenth of these PSP are remeasured each year (17 months in practice) according to a predefined scheme ensuring that the whole territory is uniformly covered [127]. The entire Walloon Region is covered by 33000 plots with 11000 in forest.

The sample units are concentric circular plots (Figure 1.6). Data are collected in productive forest land. Three main circular plots (radius of 4.5, 9 and 18 m radius) are used to measure living trees with a minimum circumference of respectively 20, 70 and 120 cm. Total height is measured for all hardwood species

and only for dominant trees for coniferous species. Four circular sub-units of 2.25 m radius are established for natural regeneration. More general variables are collected within a 36 m around the plot center: ownership status, topography, soil, structure, age, composition...

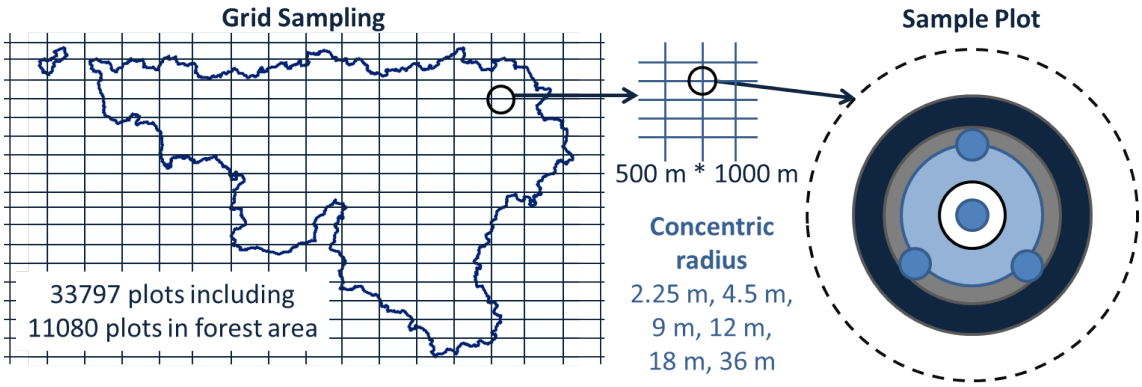


Figure 1.6: The inventory relies on a systematic sampling design with Permanent Sample Plots (PSP). The sample units are concentric circular plots. Data are collected in productive forest land. Three main circular plots (radius of 4.5, 9 and 18 m radius) are used to measure living trees with a minimum circumference of respectively 20, 70 and 120 cm. Total height is measured for all hardwood species and only for dominant trees for coniferous species. Four circular sub-units of 2.25 m radius are established for natural regeneration.

Data are collected in the field via an electronic field recorder allowing a first quality control. Plot position is now updated by Global Positioning System (GPS). The other tools used are a measuring tape and a vertex (Figure 1.7). Three reference trees were marked during the installation of the plot (first field visit) to allow the triangulation of the plot center in the next measurement campaign. Back to the office, the data are centralized in the main database. They are processed in order to have forest information as the proportion of forest types and structures, wood and biomass stocks, growth and the quality of the habitats. Collected, measured and processed data are available for any user in various forms as tables, graphics and maps. Detailed information on the methodology and results of the Walloon RFI can be found in [4, 76].

Forest Inventory for Planning

Forest management plans require reliable, descriptive and quantitative data on the resource. In order to help the managers, especially the Walloon Forest Administration, Gembloux Agro-Bio Tech has set up a methodology for carrying out systematic sampling inventories in forests for the preparation or revision of management plans [79]. In this context, a software has also been developed, including an office component as well as a field component [37].

Figure 1.7: Material for forest inventory: (a) Vertex for height, distance and angle measurement ; (b) a measuring tape for circumference ; (c) GPS for positioning and navigation.

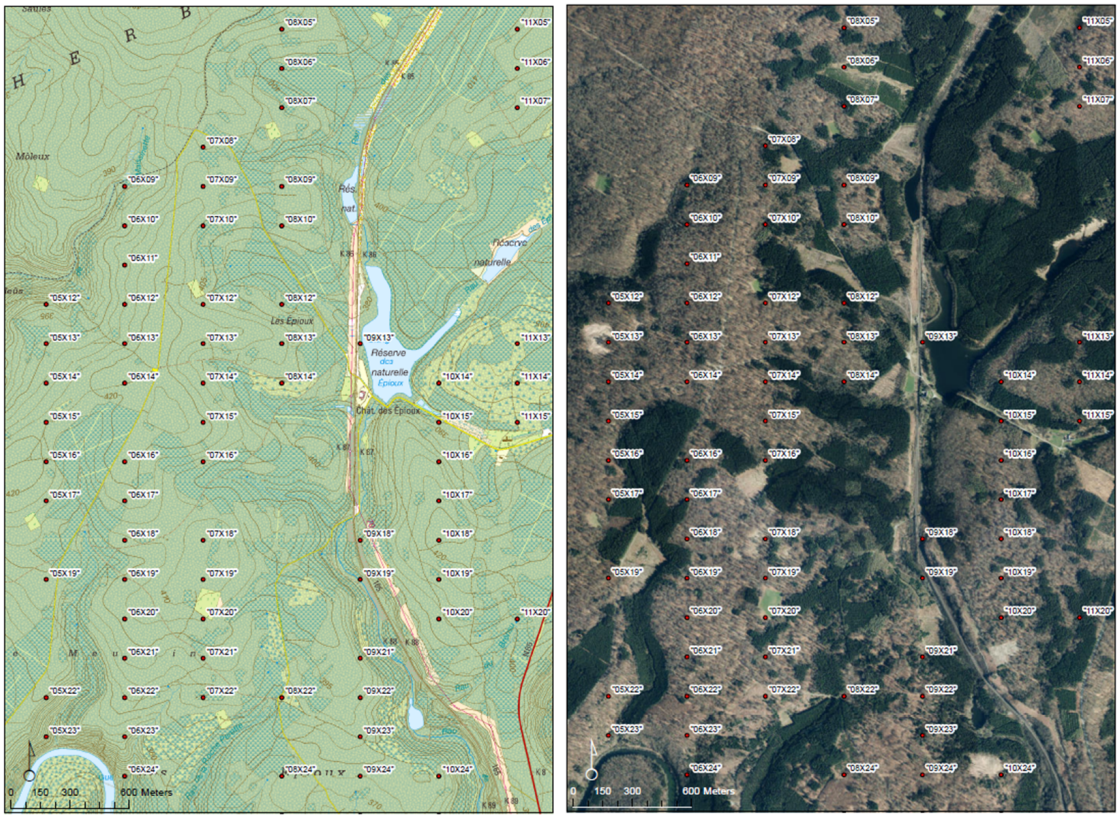


Figure 1.8: Grid example of Planning Forest Inventory. This type of inventory has a low sampling rate (< 10%) along a systematic sampling grid. The plot size is variable, defined by a minimum number of tree. The principle is to focus the effort of measurement on a relatively small set of trees (usually 10 to 15 trees) located within the plots. The low number of tree measurements is compensated by a greater rigor and an increased precision during the field data collection.

The main characteristics of this type of inventory are:

- a low sampling rate (< 10%) along a systematic sampling grid whose size allows the installation of 1 plot/ha at 1 plot/12.5 ha
- plots of variable size defined by a minimum number of tree (commonly 15 trees)
- accurate measurement of circumference at breast-height (CBH) of all trees and measurement of height of a selection of trees (to be defined according to the stands and objectives)
- the description of the trees including health status and the presence of damage.

The principle is to focus the effort of measurement on a relatively small set of trees located within the plots. The low number of tree measurements is compensated by a greater rigor and an increased precision during the field data collection. Similar to RFI, the measuring tape, vertex and GPS are commonly used. Reference trees can be selected in the field in the case of permanent plots to ease re-measurements. Amongst the results of this type of inventory, the main synthetic dendrometric parameters are crucial: number of stems, basal area, volume per hectare, dominant height. Several other information can be mentioned: a characterization of the structure and composition (number stems by size class and species), estimates of regeneration cover, estimates of the frequency of defects or damage to trees in the forest or natural regeneration, estimates of growth increment in circumference or volume (in case of remeasuring in the plots of a previous inventory). Planning inventories can cover several thousand hectares of forest. This survey can be implemented by teams of two operators. The speed of execution of the measurement is highly variable and depends mainly (1) on the size of the sampling grid that influences the travel times between plots and (2) the nature and diversity of the collected variables.

1.3.2 *Remote Sensing Data*

The variety of remote sensing data available has become very large as the technology evolves but data can be synthetically distinguish following a typology based on 5 characteristics [56, 82, 64].

- **Platform**: satellite, plane, helicopter, microlight plane, Unmanned Aerial Vehicle (UAV). The type of platform impacts the extent of the area of data acquisition.

- **Sensor**: passive or active, visible or multispectral or hyperspectral. The sensor is dedicated to specific applications.
- **Product**: images or point clouds
- **Resolution**: from low to very high and formalized in cm or point density. The resolution is inversely proportional to the size of detection for object of interest.
- **Acquisition frequency**: few days with satellite constellation or to be planned with plane. This characteristic is highly dependent of the financial resources and the speed of change of the object study.

The following sections present in more details data available and used for research purposes in Wallonia, in terms of product type.

Imagery

The most common remote sensing data are aerial or satellite images. These images are acquired by sensors installed on plane or satellite platforms. These sensors record the electromagnetic radiation emitted or reflected by the surface of the globe [64]. There are different types of sensors, producing several types of images that can be used in a large number of environmental applications. These types of image are derived from passive sensors, as opposed to active sensors. A passive or optical system records the natural radiation emitted or reflected by an object as a result of sunlight (the main source of electromagnetic energy received on Earth). In the case of an active system, such as LiDAR, there is emission of energy by the sensor itself to the object; the system records the signal returned by the target [82]. Active systems do not depend on sunshine, so they are less sensitive to weather conditions. The characteristics of the sensors and platforms (sensitivity, flight altitude, orbit, ...) rule the properties of the images and their use. These are defined by several types of resolution: spatial, spectral, radiometric and temporal.

The spatial resolution corresponds to the size of the pixels that constitute the basic elements of an image. In general, aerial images have a finer spatial resolution (between 10 cm and 1 m) than satellite images (between 1 m and 1 km) [82]. The size of the pixel of an image determines how the landscape is understood. Smaller pixels make it possible to capture finer levels of detail. Unlike larger pixels will gradually erase the details present in a scene. The sensors record the radiation emitted or reflected from the Earth's surface for certain wavelength ranges, termed spectral bands. These correspond to certain portions of the electromagnetic spectrum, some of which being part of the visible spectrum. The

spectral resolution corresponds to the number and the size of the spectral bands recorded by the sensor (Figure 1.9). The radiometric resolution is the ability of the sensor to distinguish differences in reflectance; the sensitivity of the sensor to the signal strength. The greater the radiometric resolution, the more accurate the image will be. Radiometric resolution is quantified in number of bits, typically in the range of 8 to 16 bits. Finally, the temporal resolution corresponds to the frequency of acquisition of an image at a given location. This notion is mainly used in the case of satellite images, satellites being designed to fly over the same places at regular intervals (from 1 to 16 days) [64].

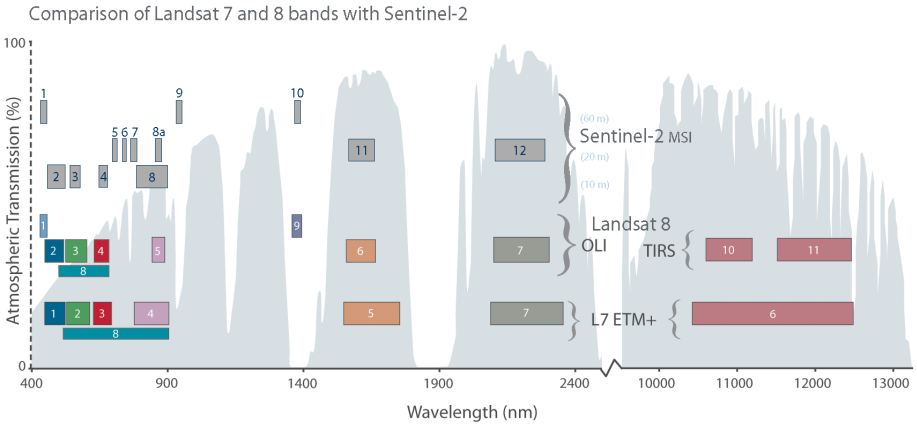


Figure 1.9: The 13 spectral bands of Sentinel-2 (compared with the spectral bands of Landsat 7 and 8) overlapped with atmosphere transmission. The spatial resolution depends on the spectral band: 10 m (bands 2, 3, 4, and 8) , 20 m (bands 5, 6, 7, 8a, 11, and 12) and 60 m (bands 1, 9, and 10).

➛ Satellite images

Thanks to their orbit, satellites allow a repetitive and continuous coverage of the surface of the Earth. Since decades, Earth observation satellites have provided images for the study of the terrestrial biosphere, with variable temporal, spatial, spectral and radiometric resolutions. A major limitation to the exploitation of satellite images is the presence of cloud cover. Landsat images (30 m resolution) are available free of charge. Commercial satellites provide images with a resolution of less than 10 m but with payment (Ikonos, Quickbird, Pléïades, ...). Recently, a compromise emerged with a new type of satellite data: Sentinel-2. The launch of the two Sentinel-2 satellites is an opportunity to enhance forest characterization on a large areas. The multispectral 13-band sensors produce high-quality images (Figure 1.10) at a 5-day equatorial temporal resolution. Such an availability of free data is unprecedented and will substantially promote research in this topic. On 28 February 2008, the European Union (EU) and European Space Agency

(ESA) signed an agreement over the creation of the COPERNICUS program. The aim of this program is to provide earth surface monitoring services [44] (Land, Atmosphere & Marine Monitoring, Climate Change, Security & Emergency Management Services). Sentinel-2 on-board the passive Multi-Spectral Instrument (MSI) providing 13 spectral bands [136, 45] with a Focal Of View (FOV) of 290 km tiled at 100km^2 in UTM/WGS84 projection [136].

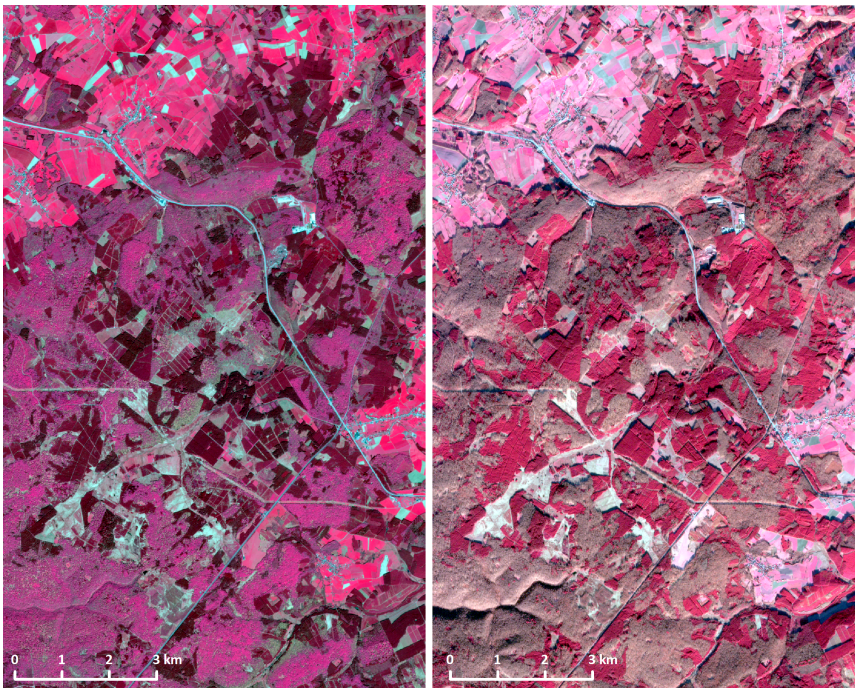


Figure 1.10: Example of two Sentinel-2 images (August 2015 and May 2016 for the same area). The infrared false color composition (Near infrared (red), green (blue), red (green)) gives a good contrast between coniferous, broadleaved, agricultural areas and roads. The two dates allow to visualize variation in land cover and phenology of broadleaved species.

➛ Aerial images

Aerial images are photographs commonly taken vertically from an aircraft with a highly-accurate camera. This kind of images are acquired approximately at 3000 to 4000 meters altitude and can be used for the creation of ortho-images, through a process called "orthorectification" often followed by mosaicking and final tiling. The mission for the acquisition of the aerial photos is planned according to the characteristics of the sensor and the desired result as well as the characteristics of the aircraft. The cameras used are metric, that means the geometric and chromatic deformations of the lens are well measured and known. An important parameter during the acquisition is the longitudinal and side overlap which is determining for the further processing of the images especially photogrammetry.

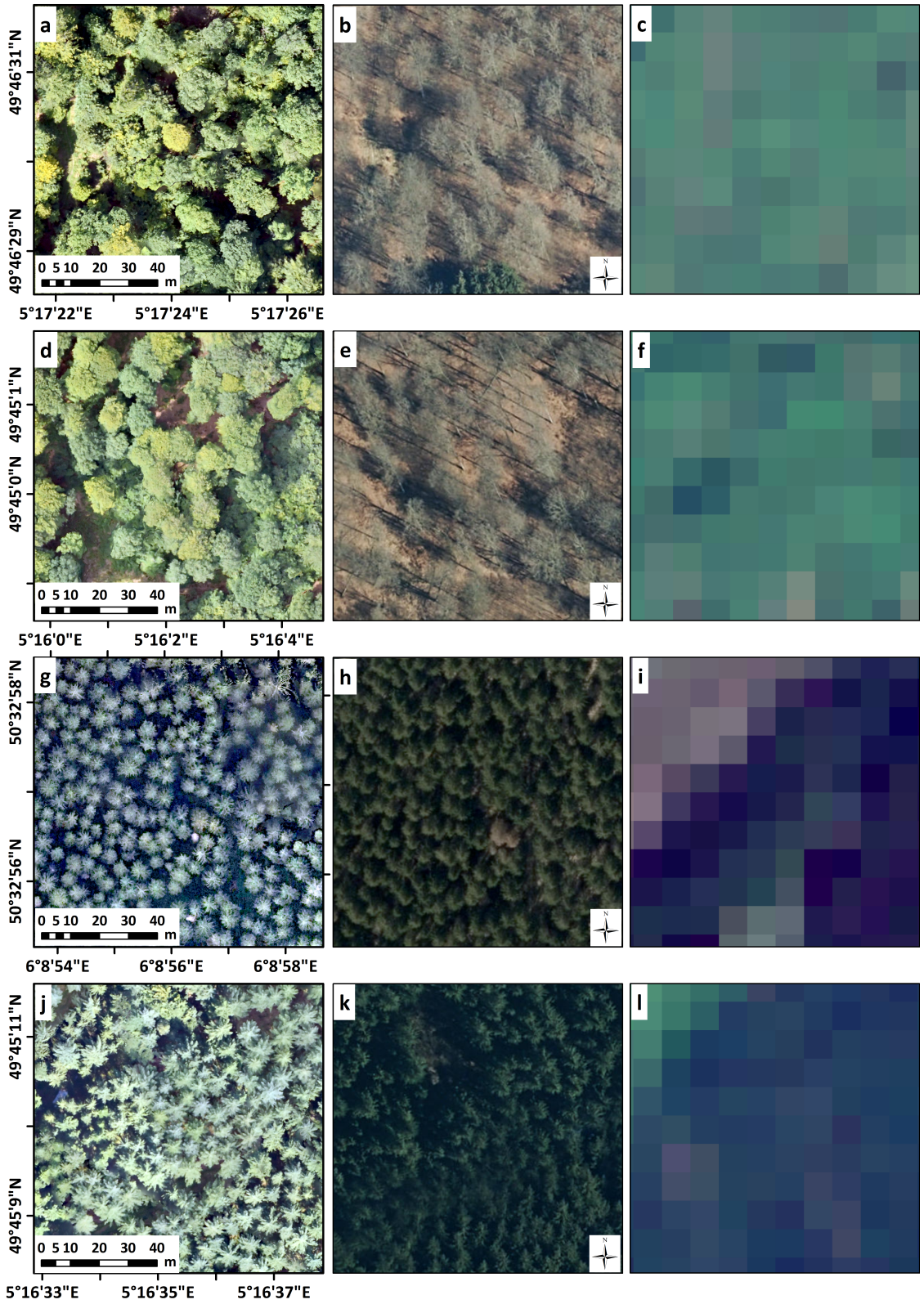


Figure 1.11: The four main species of the Walloon forest are oak (a, b, c), beech (d, e, f), Norway spruce (g, h, i) and Douglas-fir (j, k, l). These four species are illustrated with three types of images and resolutions, from left to right for the same area: UAS 10 cm image, aerial 25 cm image and Sentinel-2 10 m image.

➛ **Unmanned Aerial System**

A drone, also called UAV or Unmanned Aerial System (UAS), is an aircraft without pilot and remotely controlled, with a certain autonomy of flight [43] thanks to the on-board navigation instruments: GPS, inertial station (measurement of the orientation), compass, barometer (measurement of the flight height). Drones navigate with a flight plan composed of a set of navigation points. UAS are able to carry on board sensors relevant for environmental monitoring (Figure 1.12). Shooting must be planned with a large overlap between images (65 to 90 %). The quality of the result is influenced by many flight parameters such as the altitude and the distance between two flight lines, which will condition the lateral overlap of the images. The endurance of flying an electric drone rarely exceeds 1 to 2 hours; it is the most limiting factor for forestry use (except legal issues). The use has to be considered for applications at local level [123].

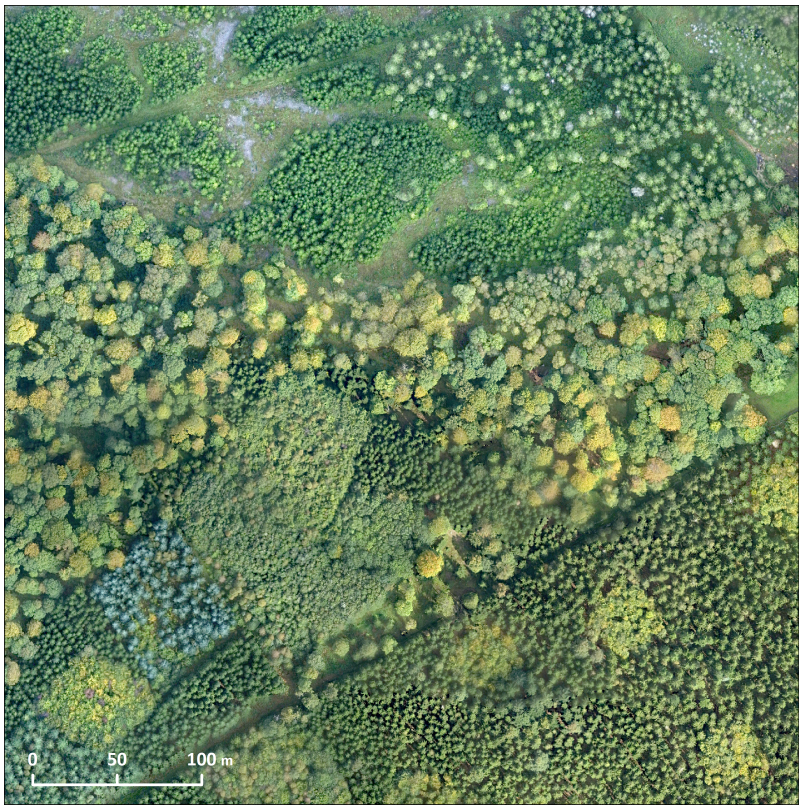


Figure 1.12: Example of Unmanned Aerial System (UAS) image at 25 cm resolution in East Belgium. The high spatial resolution allows to easily distinguish individual tree crowns and the spectral variations between trees.

There are two types of drones: fixed-wing and rotor drones. Fixed-wing drones have better flight autonomy and higher flight speed than rotor wing UAS, resulting in a larger surface area for equivalent flight endurance [83]. However, they have less workability and are less versatile. Rotor wing UAS

are able to perform stationary flights, to board larger payloads and are less sensitive to changes in the center of gravity of the platform. In addition, these allow a take-off and a landing in the vertical. The need for a landing area without obstacle is a limiting factor in the deployment of UAS in the forest environment.

The UAS imagery is very specific due to the characteristics of the acquisition platform. Drones fly at low altitude, thus the footprint of a raw image is limited (between 50 and 300 meters in width) and the resolution is sub-centimetric to decimetric. Cameras embedded on drones involve significant distortions due to lens deformation [1]. The quality of the images also suffers from a spun effect, from the movement of the drone, or from blur caused by the vibrations. Besides, the changes of relief of the flown scene lead to many omissions (hidden parts), called occlusions. As a result, to cover an area of interest, it is necessary to acquire a considerable number of images with a high overlap between them [43]. This overlap between images allows, by means of homologous points, to orient them relative to each other and to correct the distortions of the sensor. The sensors embedded in the drone are not calibrated and are generally of poor quality. In addition, the orientation of the camera at the time of shooting is slightly oblique, and very variable between two successive shots. The inertial unit does not provide precise external orientation data (position and orientation of the camera at the time of shooting).

The figure 1.11 shows extracts of three types of images for the four main species of the Walloon forest.

Point Clouds

A point cloud is a set of points with their position defined in a three-dimensional (3D) coordinate system (x, y, z) and representing the external surface of an object. Point clouds may be created by Airborne Laser Scanning (ALS, also called aerial LiDAR) or photogrammetry from overlapping images. An important issue about point cloud acquisition is the visibility of the target objects, without consideration of the method (ALS or photogrammetry): the occlusion results in area with no data. The resolution of point cloud datasets is quantified by the point density (number of points by square meter or mean distance between two points). Obviously, a higher density leads to a better representation of the complexity of an object. Figure 1.13 illustrates a 2D projection of an ALS (lighter blue) and a photogrammetric (darker blue) point clouds for the same area. The specificities of these two types of clouds will be described hereafter.

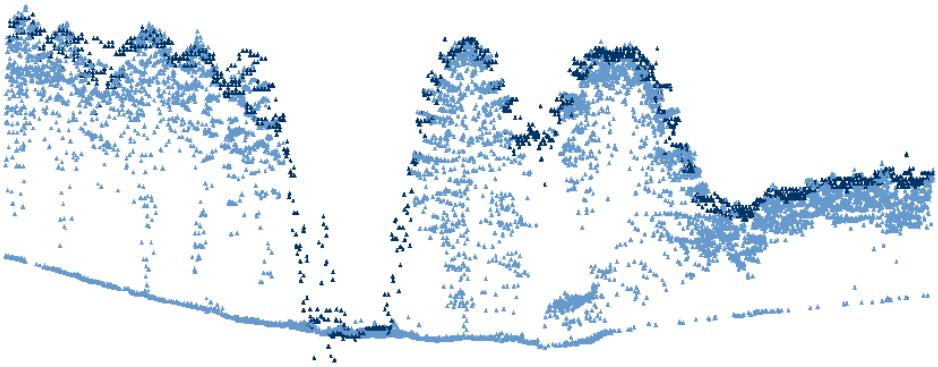


Figure 1.13: Example of two point cloud types. The dark blue points come from a photogrammetric workflow. This cloud allows a characterization of the canopy top. The light blue is an ALS point cloud with a high point density allowing a better vertical description of the forest structure.

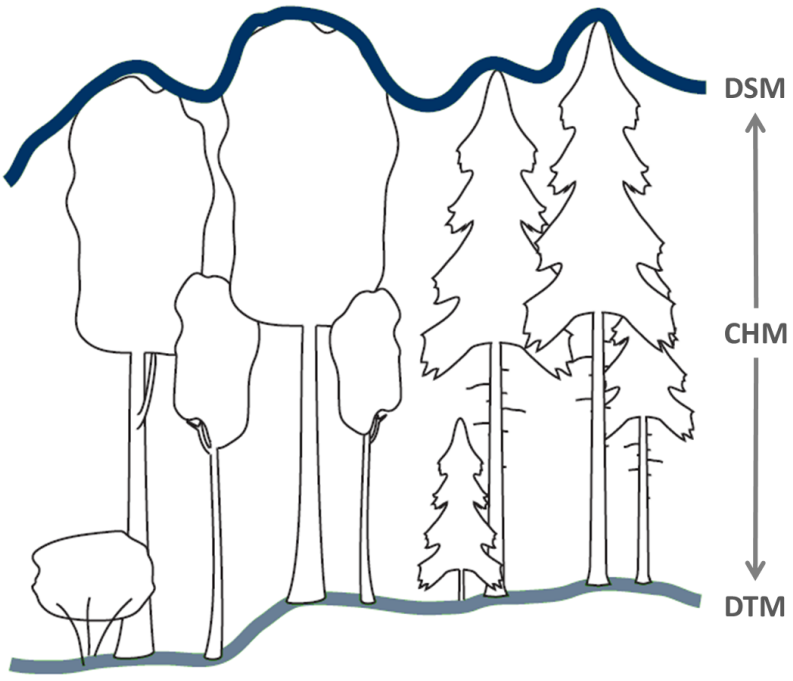


Figure 1.14: A Canopy Height Model (CHM) is a very relevant product regarding forest assessment with remote sensing. A CHM can be created directly from an ALS point cloud or by subtracting a Digital Terrain Model (DTM) from a Digital Surface Model (DSM). A DSM is a raster representing the altitude of the object top and a DTM is the ground altitude. A DSM can be created from photogrammetric or ALS point cloud. Adapted from [20].

The interest of the point clouds lies in their 3D information, in particular in the z coordinate, the elevation, which makes possible to obtain data about height of the objects. The information contained in a 3D point cloud can be summarized in 2D data as two rasters: the Digital Surface Model (DSM) and the Digital Terrain Model (DTM). The first one represents the altitude of the top of the objects (tree tops for example), the second one is the altitude of the ground. Finally, a Canopy Height Model (CHM) can be computed as the difference between the DSM and the DTM (Figure 1.14).

➛ **Airborne Laser Scanning**

The data from ALS technology consists of a representation of the Earth's surface as 3D point clouds where each point corresponds to the point of impact of the LASER (Light Amplification by Stimulated Emission of Radiation) beam with an object. ALS data is a special case of LiDAR data acquired from aerial platforms (there is also satellite and terrestrial LiDAR data but it is beyond the scope of this introduction). LiDAR is an active sensor, based on the emission of a LASER pulse by a transmitter installed on an air platform (plane or helicopter) and the recording, after interception with the terrestrial surface, of the return signal (echo) by a receiver [9, 92].

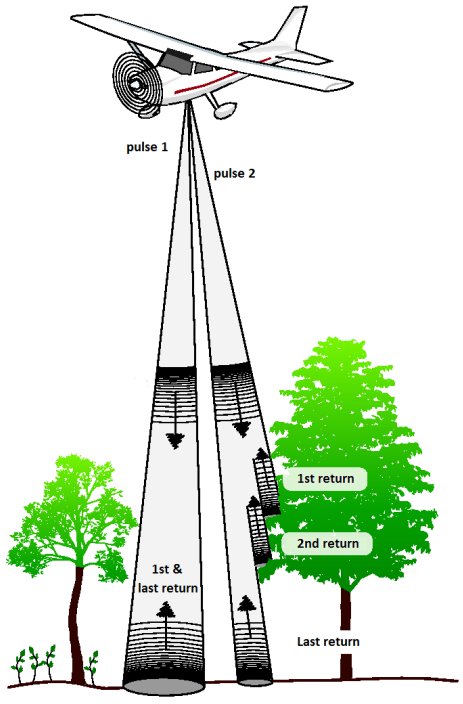


Figure 1.15: Principle of ALS data acquisition (modified from [22]). A pulse is emitted by the sensor in the plane. A single return is recorded if the pulse directly hits the ground. Several returns can be recorded if the pulse intercepts obstacle as trees.

The basic principle of LiDAR systems is based on the determination of the distance between the target and the platform thanks to the time taken by the LASER beam to intercept the target and return to the receiver [64]. The transformation of this raw "distance" into coordinates (x, y, z) involves especially a Differential Global Positioning System (dGPS) and an Inertial Measurement Unit (IMU). Many LASER systems have the property of being able to record several echoes for the same transmitted pulse. Multi-echo systems typically collect the first and the last echo [92]. The Figure 1.15 illustrates the principle of LiDAR data acquisition: a LASER pulse sent by the transmitter to the ground can generate one or more echoes in return if "objects" are located on the path of the impulse (stones, trees, buildings,...); on the other hand, a single return will be recorded if the impulse reaches the ground directly [92]. This feature is interesting in the forest area since it is possible to collect information on the vertical structure of a stand.

➛ **Photogrammetry - Structure-from-Motion**

Because of their high operational flexibility, UAS can deliver very fine spatial resolution data at specific moments defined by the end users [5]. Flying low and slow, a small UAS with an on-board optical sensor can acquire images of natural areas (e.g. forest canopies) and provide spectral information. By their their versatile acquisition and high overlap at local scale, UAS imagery is characterized by a high level of information redundancy allowing to accurately model the canopy surface at a very high spatial and temporal resolution [87]. The concomitant improvements of image processing techniques make todays' novel "Structure-from-Motion" photogrammetry (SfM) workflow (Figure 1.16), operational with UAS imagery [31]. The SfM photogrammetry approach deals automatically with a collection of unordered overlapping images from an uncalibrated camera [155]. The SfM algorithms aim to produce a sparse 3D point cloud and camera orientations by automatically determining scene geometry, camera calibration, camera position and orientation from an unordered overlapping collection of images [132]. When used with UAS imagery, this modern photogrammetry technique delivers without difficulty two major final products of the photographed area for the user : a 3D model of the relief and an orthophotomosaic. UAS imagery as basis for dense point-cloud generation on a forested area allows a cost-effective alternative to ALS acquisition.

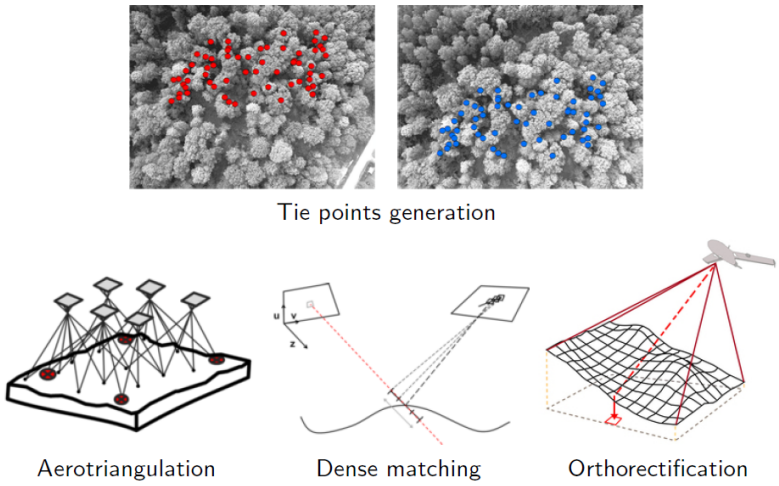


Figure 1.16: Principle of photogrammetric workflow on UAS images

1.3.3 *Specificities of Data Available in Wallonia*

In Wallonia, there is a clear growing ambition to develop the use of remote sensing for the characterization of the natural resource. Gembloux Agro-Bio Tech conducts research in collaboration with the Forest Administration and Non Navigable Watercourses Administration to lay the basis for future operational use at various scales. In this context, several non-satellite data are acquired by the SPW and used transversally by different departments or universities. In parallel, Gembloux Agro-Bio Tech supports drone data acquisition. The SPW finances data acquisition such as orthophotos and ALS and promotes access to data via a geoportal.

➛ ALS and orthoimages

The Public Service of Wallonia regularly produces aerial views of the Walloon territory. In Wallonia, ortho-images have been produced initially for agricultural applications in the framework of Common Agricultural Policy. Collections of orthophotos of the Walloon territory are available for the following years: 1971, 1994-2000, 2006-2007, 2009-2010, 2012-2013, 2015 and 2016. The primary purpose of these images being the agricultural control, the specifications allow that the acquisitions are made before the budburst date of broadleaved species.

In terms of spatial and spectral resolutions, regional ortho-images could be compared to Pléiades images. The advantages of the latter are a potential daily revisit time, a reduced computer space and a clear sky probability of less than 20%. Their spatial resolution (2 m/50 cm), although close, is less (25 cm) than the aerial images which are acquired once a year and fully preprocessed data provided by the SPW for the users. The acquisition period

extends over several months. A major interest of the regional ortho-images datasets is their availability for research purposes at a null or low marginal cost thanks the collaboration between the SPW and universities. Besides, recurring acquisitions are planned by the SPW for years to come.

A first ALS dataset was acquired in 2011 by the Walloon Non Navigable Watercourses Administration in the Houille Watershed. The acquisition was part of the Interreg project "Floods Houille- P2IH", and consisted in two LIDAR high density (40 points/m^2) surveys of the Houille basin (flights in March and July). The study area corresponds to the Houille transboundary watershed. A regional LIDAR dataset was acquired by the SPW between 2013 and 2014 (1 point/m^2), with a flight in December 2012). The objective of the LiDAR acquisition was to develop a DTM. In order to favor the presence of ground points, the acquisition was mainly carried out between December and March.

➛ Sentinel-2 images

Sentinel-2 images are interesting data in terms of spatial, spectral and temporal resolutions considering the extent of the data, the weekly revisit time, the free data policy and the user's autonomy to download the data. These data are a great opportunity to conduct studies at an international scale, with neighboring countries as in the framework of two Interreg projects. These projets, partly financed by European Regional Development Fund, aim to directly benefit the wood industry by better assessing the types of species and wood volumes that will supply the market in the coming years. The Interreg V project "Feel Wood - Pro Bos Forest" concerns an area of 62 000 km^2 straddling the Northeastern France and Western Wallonia and Flanders. In a framework of valorization of the forest and short circuit for the wood sector, the objective of this project is the identification of the sources shortness of breath of the die. The Interreg VA "Regiowood II" (2017-2019) bringing together foresters from France, Wallonia, Germany and Luxembourg, as a part of Objective 3 of Priority Axis 2: Ensure development respectful of the environment and the living environment.

➛ UAS data

The use of drones offers the chance to acquire customized very high resolution data, as needed. A specificity of the acquisition of these data is the legislation which surrounds it. In Belgium, the Royal Decree of April 10, 2016 regulates the use of remotely piloted aircraft systems (drones) in Belgian airspace. In addition to the safety of other aircraft and persons and objects on the ground, the decree also aims to guarantee the privacy of the

citizen. A risk-based approach is applied to set proportionate requirements for the different types of operations with drones. There are 5 types of exploitation/classes, based on the weight of the drone, the height of the flight, the type of terrain overflown (private model aircraft terrain, other), the distance from view, the distance around obstacles, the pilot age. Private use or use as model aircraft are not allowed for commercial or professional purposes. Each flight has to be notified to the Belgian Civil Aviation Authority (BCAA) before take-off.

Figure 1.17: Three types of drones are used in Gembloux Agro-Bio Tech. (a) The Gatewing X100 used for the data acquisition in chapter 4 ;(b) An Multirotor, octocopter; (c) A fixed wing UAS.

The Unit of Forest Management of Gembloux Agro-Bio Tech - ULiège makes numerous acquisitions of data by UAS (Figure 1.17) for several scientific researches in partnership with the Forest Administration and the Walloon Non Navigable Watercourses Administration. In this particular context, the type of operation falls within the class 1a because of the flight height, the weight of the drone and non visual flight rules. In that case, the requirements of the decree are: registration of the drone at the BCAA, a remote pilot license according the class, an operations manual and risk assessment, or class 1 operations, the prior authorization of the BCAA.

1.4 Aims and structure of the thesis

As developed in the previous sections, the Walloon forest is characterized by a great diversity of situations and concerns: type of owner, structure, composition, spatial distribution, stations, fragmentation, silviculture. Supporting the different forest policy directions (resilience and versatility to climate change, preservation of biodiversity, support wood production, keep balance between broadleaved and coniferous stands and between forestry and game populations, certification of forest management plans) requires a good knowledge about the forest resource. This framework makes it particularly interesting and useful to use remote sensing as a support to management. Moreover, if we consider more precisely the information collected on the forest by the Walloon RFI, several basic data are par-

ticularly crucial and consistent to link to remote sensing data as the four main stand attributes: top height, number of trees per hectare, basal area and volume per hectare.

The diversity of remote sensing data goes together with the variety of possible applications involving different spatial resolutions and working scales. It is important to be aware of the possibilities of the different data in order to better understand the needs according to the objective and the expectations of a forest manager towards remote sensing. Especially, 3D data (as ALS or photogrammetric point clouds) are really relevant regarding forest characterization and quantification.

➨ Given the different types and resolutions of 3D data and their interest for forest applications, what are their relevant uses to acquire information on forest stands, in the context of sustainable forest management in Wallonia?

Considering this broad issue, the purpose of this thesis is to use several types of remote sensing data for the quantitative and qualitative characterization of forest stands considering different working scales, from local to regional (Figure 1.18): UAV data, high and low density ALS. Our generic question is hinged on three specific research questions corresponding to different working scales determined by the spatial resolution and forest types specific issues:

- ➛ How to combine ALS data and the Regional Forest Inventory to assess forest attributes of coniferous stands at a regional-scale?
- ➛ How reliable is a canopy gap mapping in broadleaved uneven-aged stands using ALS data?
- ➛ What is the potential of UAV imagery to characterize coniferous stands at local-scale?

These specific questions are dealt with three chapters (chapters 2 to 4) built on three scientific papers listed below. We decide to implement complementary approaches by considering coniferous versus broadleaved forests, forest stands versus canopy gaps and forest attributes assessment versus mapping. The datasets used in this thesis are presented in each chapter specific to each scale of interest. The second chapter is a manuscript in preparation, complementary to [38] while chapters 3 and 4 are the identical transcriptions of the published papers (listed above). Last chapter is a general discussion and provides several perspectives.

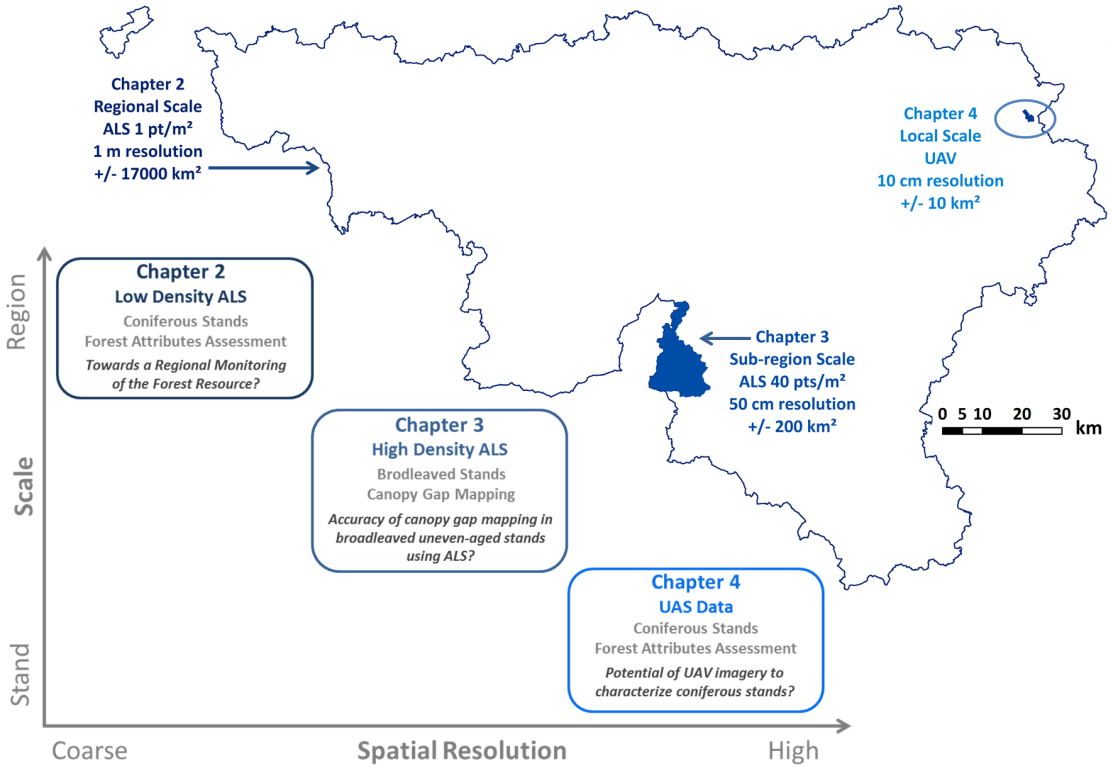


Figure 1.18: The three research chapters are related to three working scales in line with the study area.

- Stéphanie **BONNET**, Gauthier LIGOT, Sébastien BAUWENS, Philippe LEJEUNE and Adeline FAYOLLE. Monitoring of Coniferous Stands by Combination of ALS and Regional Forest Inventory Datasets. In prep. *(chapter 2 - Monitoring of Coniferous Stands with a Regional ALS Dataset)*

- Stéphanie **BONNET**, Rachel GAULTON, François LEHAIRE and Philippe LEJEUNE. Canopy Gap Mapping from Airborne Laser Scanning: An Assessment of the Positional and Geometrical Accuracy. In: Remote Sensing (2015), 7, 11267-11294. *(chapter 3 - Canopy Gaps mapping in Broadleaved Forests with ALS)*

- Stéphanie **BONNET**, Jonathan LISEIN and Philippe LEJEUNE. Comparison of UAV Photogrammetric Products for Tree Detection and Characterization of Coniferous Stands. In: International Journal of Remote Sensing (2017), 38(19) , 5310 - 5337. *(chapter 4 - Forest Attributes Assessment of Coniferous Stands with UAV Data)*

Around the main papers, several ancillary co-authored papers are listed as they contribute to the discussion. These latter are of scientific or popularised nature (written in French in this latter case).

- Stéphanie **BONNET**. Un modèle numérique de canopée pour l'estimation de la hauteur dominante des peuplements résineux en Région wallonne. In: Forêt Wallonne (2009), 98, 53-59.

- Stéphanie **BONNET**, François TOROMANOFF, François FOURNEAU and Philippe LEJEUNE. Principes de base de la télédétection et ses potentialités comme outil de caractérisation de la ressource forestière. I. Images aériennes et satellitaires. In: Forêt Wallonne (2011), 114, 45-56.

- Stéphanie **BONNET**, François TOROMANOFF, Sébastien BAUWENS, Adrien MICHEZ, Laurent DEDRY and Philippe LEJEUNE. Principes de base de la télédétection et ses potentialités comme outil de caractérisation de la ressource forestière – II. LiDAR aérien. In: Forêt Wallonne (2013), 124, 28-41.

- Jonathan LISEIN, Marc PIERROT-DESEILLIGNY, Stéphanie **BONNET** and Philippe LEJEUNE. A Photogrammetric Workflow for the Creation of a Forest Canopy Height Model from Small Unmanned Aerial System Imagery. In: Forests (2013), 4, 922–944.

- Laurent DEDRY, Olivier DE THIER, Jérôme PERIN, Adrien MICHEZ, Stéphanie **BONNET** and Philippe LEJEUNE. FORESTIMATOR: un plugin QGIS d'estimation de la hauteur dominante et du Site Index de peuplements résineux à partir de LiDAR aérien. In: Revue Française de Photogrammétrie et de Télédétéction (2015), 211-212, 119-127.

- Adrien MICHEZ, Sébastien BAUWENS, Stéphanie **BONNET** and Philippe LEJEUNE. Characterization of Forests with LiDAR Technology. In: Land Surface Remote Sensing in Agriculture and Forest (2016), eds. Nicolas BAGHDADI, Mehrez ZRIBI, 331-362.

- Corentin Bolyn, Adrien MICHEZ, Peter GAUCHER, Philippe LEJEUNE and Stéphanie **BONNET**. Regional Mapping of Forest and Forest Types by Supervised Per-Pixel Classification on Sentinel-2 Imagery. In: BASE (2018), submitted.

Chapter 2

MONITORING OF CONIFEROUS STANDS WITH A REGIONAL ALS DATASET

Chapter 2

Monitoring of Coniferous Stands with a Regional ALS Dataset

> "On dit souvent que les arbres nous empêchent de voir la forêt ; il est tout aussi juste de dire qu'on ne voit pas les arbres à cause de la forêt."
>
> Ilya Ehrenbourg

Airborne Laser Scanning (ALS) can be efficiently used in forestry, providing three-dimensional (3D) information to get insight of forest characterization. The aim of this chapter is to examine the use of low density but very high spatial resolution data available at regional-scale to help for the forest monitoring. The integration of data from the Regional Forest Inventory (RFI) is also considered

as an important reference source about silvicultural knowledge. This chapter is based on two types of data available at the scale of the Walloon Region: rasters at 1 m resolution based on low density ALS data, and field data coming from the plot network of the RFI of Wallonia.

Regarding this chapter, we highlight three complementary steps aiming to describe qualitatively and quantitatively the forest stands: (i) the construction of a individual height - circumference at breast-height (H-C) allometry for Norway spruce and Douglas-fir at regional-scale with the aim to be used with remote sensing data; (ii) the analysis of the spatial (environmental factors: altitude or slope) and structural (age and stand density factors) variabilities of this H-C allometry; (iii) the combination of Area-Based Approach (ABA), Individual Tree Detection (ITD) and allometry for forest attributes assessment.

2.1 Introduction

Forest ecosystems provide important services to society, sustainable management and adapted policies are essential to maintain their ecological and socio-economic functions. Forest managers and policy makers must consider the relationships between forest functions and ecosystem characteristics and the evolution of forests in order to manage forests and make regional decisions [48, 85]. Sustainable forest management requires an accurate and regular quantification of the resource: height distribution, stem number, top-height, basal area and stem volume are essential data for the forest managers. For decades, field inventories have been used to better characterize forest. National inventories exist in countries around the world and have evolved over time to adapt to users' needs [139]. Field inventories provide timely and accurate estimates of forest resources and their evolution at a large scale. Nevertheless, this method is time-consuming and quite expensive. Therefore, technological innovation is becoming crucial to improve the efficiency of measurements and estimations while making the production of inventory data simpler [98].

For many years, remote sensing data are usually used by forest managers to help in the resource characterization. Remote sensing decreases the cost of data acquisition, increases the area it is possible to cover without sampling, and enables the production of map layers that give precious information about the distribution of forest resources. These complement sample-based procedures in the field and today are commonly used by researchers and managers [96]. Optical remote sensing (aerial or satellite) has increased the possible applications through image classification in forestry. Airborne Laser Scanning (ALS) data and digital photogrammetry technologies have made it possible to use 3D data, such as Canopy Height Models (CHM) when investigating forests. These technologies

improve the accuracy of image classification of forest classes [153]. The ability of ALS to capture the 3D structure of vegetation cover has promoted research on forest environments characterization. During the last two decades, ALS has been widely studied for quantification, mapping and monitoring of forest resources [58, 165, 164]. The accuracy of dendrometric variables derived from LiDAR data (especially concerning height) can be considered comparable to that obtained by inventory methods [105, 95]. Thanks to the experience acquired in forestry, the use of ALS has progressed out of the realm of research and is beginning to be used operationally to support forest management in several countries [157, 106, 94, 162].

Two fundamental approaches are recognized in ALS forest applications : area-based approach (ABA) and Individual Tree Detection (ITD) approach. Many publications are focused on boreal forests because methods were initially developed in boreal stands. Using ALS data in ITD approach consists mainly in detecting individual tree tops and predicting the related attributes by mean of allometric models. The tree tops detection can be followed or replaced by tree crown delineation. A major issue to be considered is that all trees cannot be detected regarding the detection algorithm, the forest conditions, especially the stand density and the spatial pattern [144]. Basically, the ABA does not detect each tree top and consider globally the whole point cloud or the entire raster within the plot extent. Several variables are extracted from the vertical distribution of the return to describe the vegetation structure.

Top height is a measure commonly used in forest management. This is a relevant indicator of the level of development that can be used to estimate the productivity of even-aged stands. However, its measurement in the field is relatively time-consuming and expensive. For several years, ALS has been known for its ability to accurately estimate tree height. In their study, [38] developed and presented *Forestimator*, a tool for characterizing the forest resource. It combines a model exploiting ALS data with top height growth models [109, 108] to provide updated top height and Site Index estimates for the two main coniferous species planted in Wallonia. As *Forestimator* was focused only on top height, further work is needed to strengthen the assessment of the other elementary attributes as number of stems, basal area and volume.

King [68] defines the allometry as the relationship between size and shape of trees. Allometry is an important issue in forestry as it influences the forest structure and function. In forestry, relations between height and diameter or circumference at breast-height (DBH, CBH) are crucial because measuring height can be difficult and costly; CBH being privileged because easily accessible in the field. Therefore H-C model can be used to reduce data acquisition costs by predicting height from CBH. From a RS point of view, the easiest variable to

measure is height and not circumference. The use of allometric models can allow the estimation of diameter from height extraction.

The aim of this study is to evaluate the combination, at regional-scale, of low density ALS data and H-C allometry based on RFI data for the monitoring of coniferous stands by the assessment of important forest attributes (number of stems, basal area & volume per hectare and quadratic mean of the CBH); the top height assessment developed in [38], was considered as achieved and not longer addressed in this study. We implement a simple and versatile workflow based on these types of data acquired at regional-scale. For this purpose, we structured our approach in three steps : (i) the construction of a H-C allometry at regional-scale (RFI data) with the aim to be used with remote sensing data ; (ii) the analysis of the spatial (environmental factors) and structural (stand density factors) variabilities of this H-C allometry; (iii) the combination of ABA, ITD and allometry for forest attributes assessment.

2.2 Material

2.2.1 *ALS Data*

A regional small-footprint ALS dataset was acquired at the average point density of 0.8 points/m^2. Survey flights were realized from December 12, 2012 to April 21, 2013 and from December 20, 2013 to March 9, 2014. The survey covered Wallonia with a regional Digital Terrain Model (1 m of Ground Sampling Distance, XY resolution). Based on the raw ALS point cloud, a Digital Terrain Model (DTM) and a Digital Surface Model (DSM) were generated for the whole region with a 1 m resolution. A canopy height model (CHM) with 1 m resolution was produced by the subtraction of the DTM from The DSM (Figure 2.1).

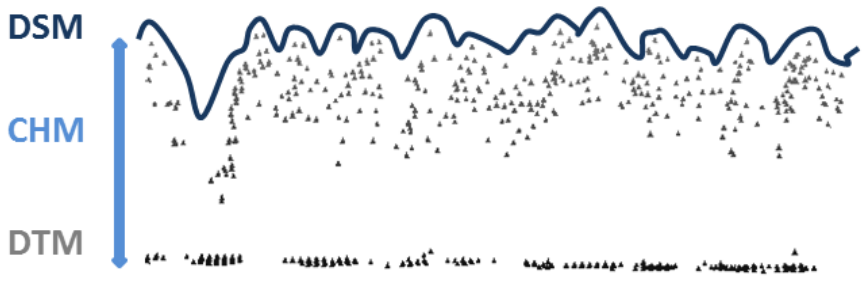


Figure 2.1: The CHM is computed by subtracting the Digital Terrain Model (DTM) from the Digital Surface Model (DSM), both created from the low density ALS point cloud at regional-scale. The points are classified beforehand to differentiate the ground from the vegetation and then create the DTM and the DSM.

2.2.2 *Forestimator*

Forestimator is a forest management tool based on a top prediction model with ALS data aiming to the assessment of Site Index in coniferous stands. The estimates provided by this model are then coupled with pre-existing top height growth models [109, 108] that allow for updating the top height over time and estimating the Site Index. This tool was validated for stands of Norway spruce and Douglas-fir on a geographic area corresponding to Wallonia. An easy access to this tool is ensured by its integration into a QGIS Open Source software plug-in. The tool, validated across this region, operates in the QGIS open source environment and may be considered as a Web Processing Service. The GIS environment allows the user to benefit from all the functionalities of such software (import/export of data, spatial analysis, creation of layout...), to enhance the data provided by this service. Free and easy to use, this plug-in is intended for both forest managers and scientists.

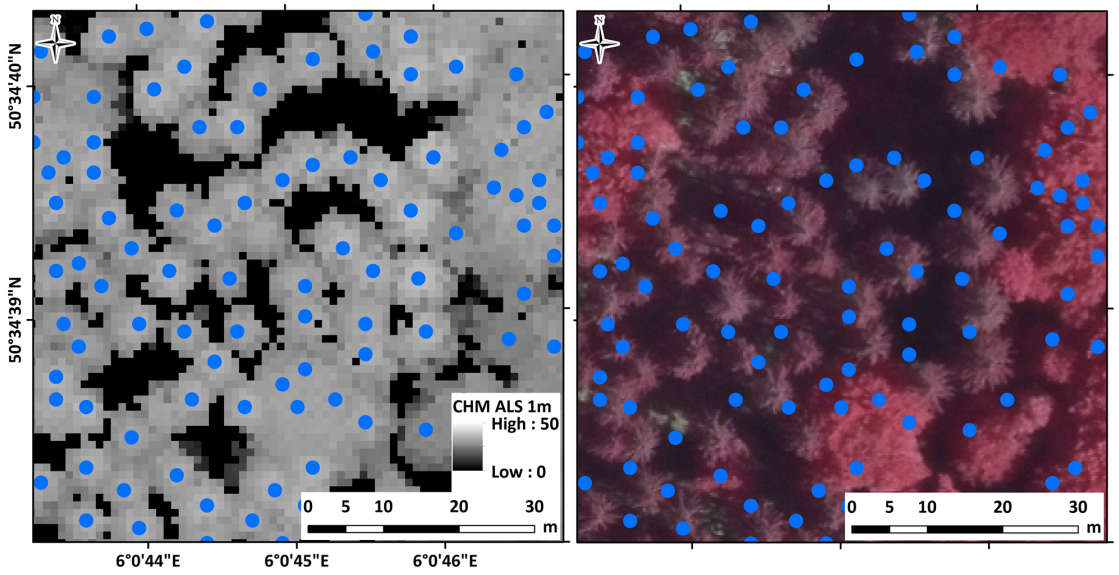


Figure 2.2: Example of local maxima extracted from the CHM (ALS, 1m) used in the plug-in *Forestimator*, overlaid on the CHM and on orthophoto IR.

The top height assessment is built on the detection of dominant trees from the CHM derived from the regional ALS data presented previously. The dominant trees were detected from the CHM using the `CanopyMaxima` function within FUSION software [97]. This algorithm is based on the assumption that the local maxima of a CHM generally correspond to the apex of a tree [121, 120]. The `CanopyMaxima` function was used with a fixed window of 5 m and considering a threshold of 4 m height below which no maximum is detected. Then, a model is developed from these local maxima to estimate the top height of plots or stands.

2.2.3 *Regional Forest Inventory Data*

The sampling method applied by the RFI is a single-phase, non stratified inventory using a systematic sampling design with permanent sample plots (PSP) at the intersections of a 1000 m (east-west) x 500 m (north-south) grid. One-tenth of these PSP are remeasured each year according to a predefined scheme ensuring that the whole territory is uniformly covered [127]. The entire Walloon Region is covered by 33000 plots with 11000 in forest. The sample units are concentric circular plots. Data are collected in productive forest land. Three main circular plots (radius of 4.5, 9 and 18 m radius) are used to measure living trees with a minimum CBH of respectively 20, 70 and 120 cm. Total height is measured for all hardwood species and only for dominant trees for coniferous species. Stand variables are collected within a 36 m around the plot center especially about structure and composition. More details about the methodology of the Walloon RFI can be found in [4]. For the purpose of this study, two datasets were built from the RFI plots network.

➛ Dataset for H-C allometry for coniferous species

This dataset consisted in reference data for the construction of a H-C allometry at regional scale. Specifically, we focused on Norway Spruce (*Picea abies* (L.) H. Karst.) and Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco). In Wallonia (southern region of Belgium), Norway spruce, Douglas-fir and mixture of these two species represent 82 % of coniferous stands and 33 % of the total forest area [77, 109]. This stand type is well-established in Belgium. The silviculture of these two species is widely based on high density plantations (2000 - 3000/ha), establishing even-aged stands [77, 109].

We selected all trees measured (CBH and total height) in plots where Norway spruce or Douglas-fir accounted for more than 80% of the total basal area. This dataset is composed of 8784 Norway spruce and 906 Douglas-fir stands, the main characteristics of these sampled stands are presented in Table 2.2. The methodology of the RFI ensures that this dataset represents all site conditions and the wide range of top-height, age and densities encountered. If several dates are available in the database, the most recent data are kept. A total of 9690 trees in 2525 plots make the dataset. Environmental variables were computed for each plot and allocated to the trees.

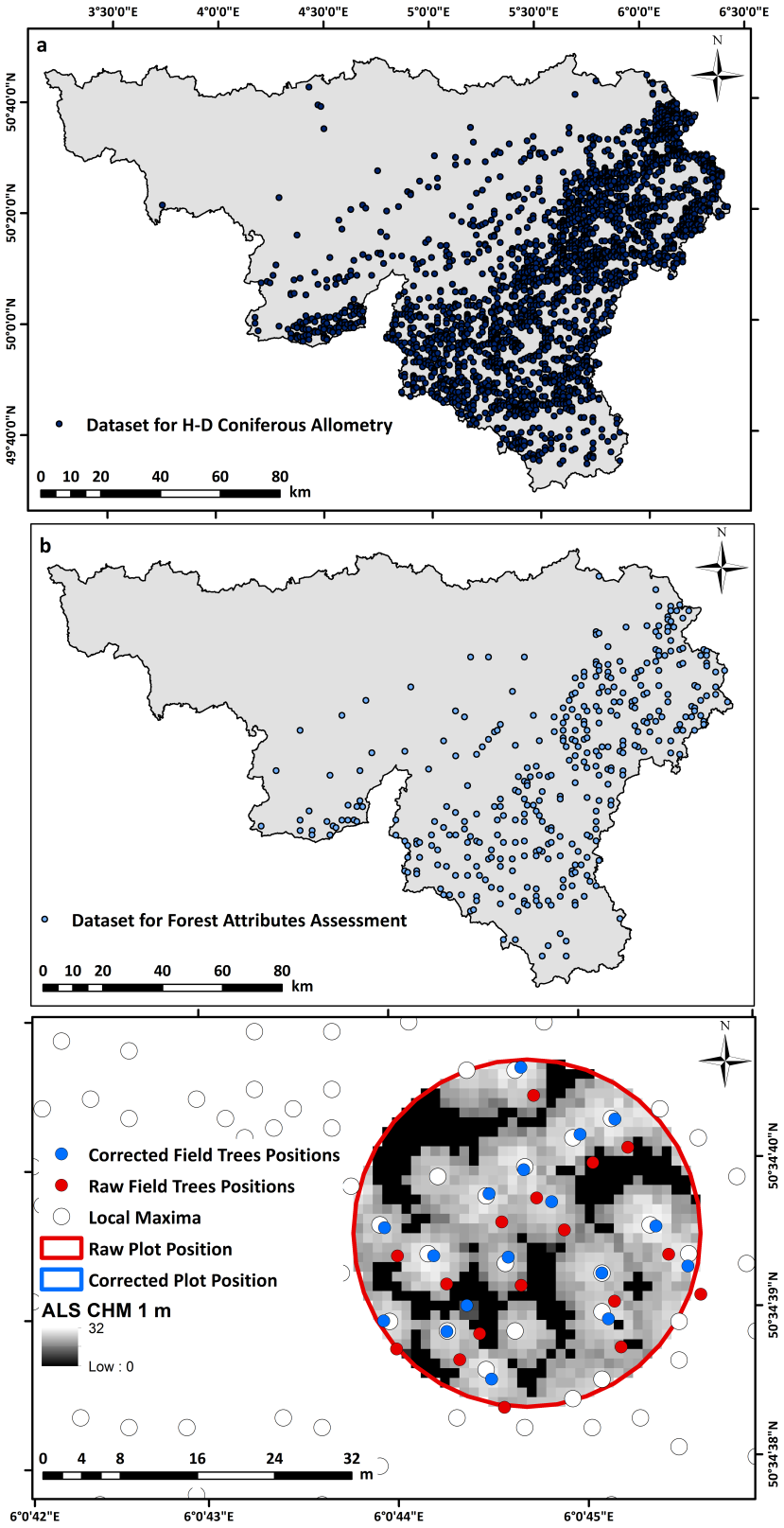


Figure 2.3: Two datasets are built from the RFI database according to the three main goals of our study: (a) H-D coniferous allometry; (b) forest attributes assessment. (c) is an example of the fine repositioning of the trees and plots used with local maxima and ALS CHM at 1 m resolution.

➛ **Dataset for forest attributes assessment in combination with ALS**

For the assessment of the forest attributes with ALS CHM data, we selected plots synchronous with ALS data. An excellent spatial match between RS data and field data is crucial as we wanted to consider ITD approach to use H-C allometry. A substantial phase of check and correction of the spatial location of trees and plots based on the ALS CHM and local maxima.

The correction of the position of the trees and plots was done by photo-interpretation based on a semi-automatic co-registration procedure. The local maxima extracted from the CHM were selected in a buffer of 70 m around the plot center. A pattern recognition algorithm, under development at Gembloux Agro-Bio Tech, was used to find the field trees positions amongst the local maxima. Several solutions can be identified by the algorithm and displayed on a GIS. The plot position was finally chosen by photo-interpretation. The operator/photo-interpreter has the CHM, aerial images with high resolution and the local maxima displayed in background on a GIS to ease the positioning. After this co-registration step, 179 plots (5874 trees) were available (Table 2.3).

Table 2.1: Statistics of circumference at breast-height (CBH) and total height of the trees selected from the RFI database for the allometry fitting.

	Min	Mean	Max	St.Dev.
Two species - 9690 trees				
Total Height (m)	5.3	25.99	50	5.8
CBH (cm)	20	120.2	290	36.5
Norway Spruce - 8784 trees				
Total Height (m)	5.3	25.9	41.5	5.5
CBH (cm)	20	119.8	262	35.8
Douglas-fir - 906 trees				
Total Height (m)	5.6	27	50	7.9
CBH (cm)	22	123.9	290	42.6

Table 2.2: Main attributes of the plots corresponding to the trees selected for the H-C allometry from the RFI database, in even-aged stands composed of Norway spruce, Douglas-fir or mixture of the two species.

	Min	Mean	Max	St.Dev.
Top Height (m)	0	23.4	52.8	11.2
N (N/ha)	10	755	5001	640
Basal Area (m^2/ha)	1.2	38.6	101	12.2
Volume (m^3/ha)	6.9	445.4	1526.3	175.3
Quadratic Mean of Girth (cm)	25.7	96.7	256.3	32.6
Altitude (m)	77.6	426.3	677.7	106.2
Slope (°)	1.4	7.6	40.1	5.7

Table 2.3: Main stand attributes of the plots selected from the RFI database, composed of Norway spruce, Douglas-fir or mixture of the two species for the forest attributes assessment.

	Min	Mean	Max	St.Dev.
Top Height (m)	0	28.3	41.5	5.1
N (/ha)	89.3	467.7	2222.3	311.5
Basal Area (m^2/ha)	14.1	38.1	73.3	9.6
Volume (m^3/ha)	158.2	480.9	1109.2	150.8
Quadratic Mean of Girth (cm)	49.3	112.1	184.3	27.9
Altitude (m)	128.4	439.9	656.5	98.8
Slope (°)	2.1	7.1	34.2	5.7

2.3 Methodology

2.3.1 *H-C allometry*

The RFI dataset was used to fit tree-level allometric models linking the CBH to the total height. Using this model will allow us to estimate the CBH of trees detected with ALS data (local maxima) and the CHM value. The generalized non-linear model was used to find the final equation. The vector of starting parameters was set with a initial linear model fitting. The function `gnls()` from the nlme package [115] was used in R [124]. Heteroscedasticity was addressed using a variance power function. The function `nlme()` was also used to introduce a random effect on the coefficient a and test mixed models as trees belonging to a same plot are not independent.

Several factor were studied and added in our analysis as influence on the model : slope, altitude, density (number of trees and basal area) and age of the stand. Slope and altitude were computed from the DTM created from ALS data. Top height was computed from the CHM coming from ALS data. Number of trees, basal area and age came from the RFI database. The Age variable has the advantage of not requiring additional field measurement, if the year of planting is known. The AIC was used to analyze the explanatory power of a series of models. The best model was used to estimate CBH of trees for the plots selected for the forest attributes assessment.

Table 2.4: Several power-law models were tested for the H-C allometry for Norway Spruce and Douglas-fir with the gnls/nlme functions. The coefficients were tested species-specific or not. CBH is the tree diameter at breast-height, H the individual height. Altitude and slope were extracted from the DTM at the plot level. Basal area, number of trees, the quadratic mean of girth (Cqm) and age come from the RFI database.

	Model
Model 1	$CBH \sim a * H^b$
Model 2	$CBH \sim a * H^b * Altitude^c * Slope^d$
Model 3	$CBH \sim a * H^b * BasalArea^f * NumbTrees^g$
Model 4	$CBH \sim a * H^b * Altitude^c * Slope^d * BasalArea^f * NumbTrees^g$
Model 5	$CBH \sim a * H^b * Altitude^c * Slope^d * Cqm^h$
Model 6	$CBH \sim a * H^b * Altitude^c * Slope^d * Age^i$

2.3.2 *Forest Attribute Assessment*

Two alternative approaches are usually used to assess forest inventory attributes with remote sensing data: (1) Area-Based approach relies on the extraction of variables at plot-scale from point clouds or CHM while (2) individual tree detection approach is obviously based on the detection if not delineation of trees/crowns. These approaches, although traditionally opposed, are in fact complementary. In order to take advantage of their respective strengths, we choose to combine them (Figure 2.4). We identified a limited number of variables to reduce the process of predictor selection among numerous variables, to limit over-fitting and consider the low number of plots [65]. Table 2.5 presents the sixteen selected variables representing the horizontal and vertical structures computed at the plot level, from the local maxima and from the use of our allometric models.

From the CHM, the 95^{th} percentile of the CHM within the plot (CHM_{p95}) was computed as a first variable related to the whole plot area. The canopy cover (CC_p) is the proportion of canopy inside the plot, where canopy is identified by a height threshold of 5 m on the CHM. It is an indicator of the horizontal distribution of the crowns area, highly significant for coniferous sites [23]. The volume below the canopy top, normalized by the plot area ($Vol_{CHM\text{-}X}$), is computed at the plot level from the CHM and is considered as a measure the structural stand complexity. The number of trees (N_{Trees}) is the number of local maxima classified as true positives and normalized by the area of the plot. The maximum height (h_{max}) and the coefficient of variation of height (h_{cv}) are computed from the local maxima within the plot. The aggregation index of Clark & Evans ($Aggreg_{als}$, Equation 4.1) describes the location diversity of the trees, based on the average distance between trees ($\bar{d}_{Trees}$) and ranges between 0 and 2.1419. V_XN_{Trees} and $V_X\_N_{Trees}$ are mixed variables, defined respectively as the product and ratio of $Vol_{CHM\text{-}X}$ and

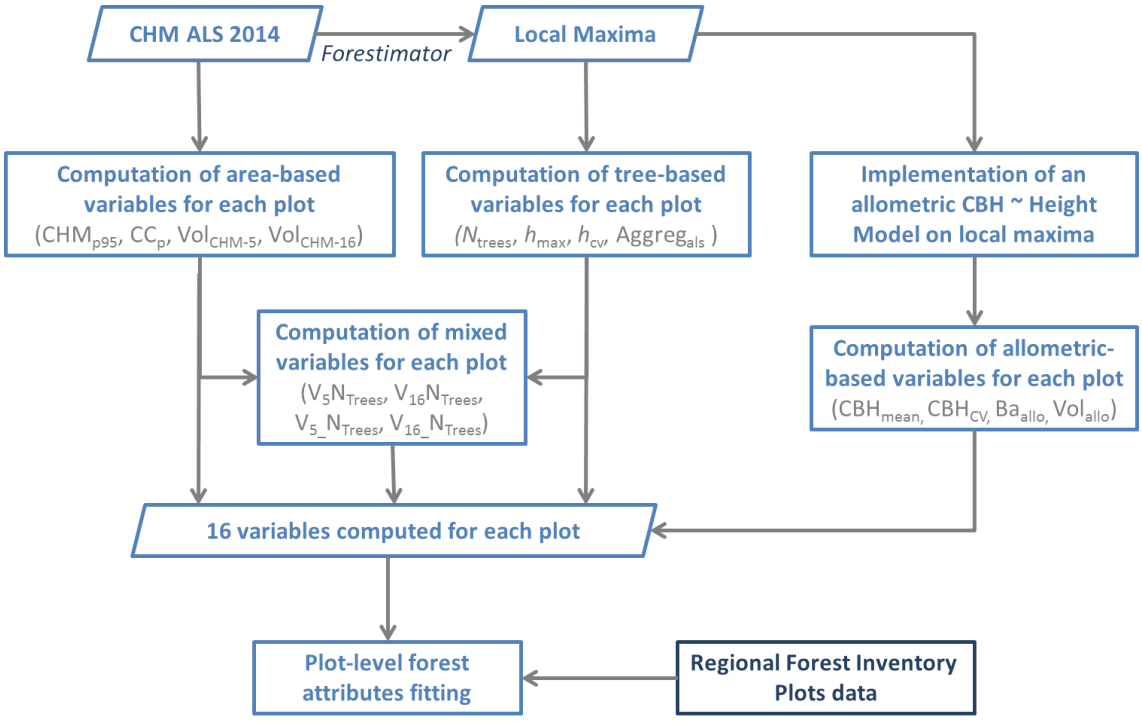


Figure 2.4: Synthesis of the sources of variables for the models fitting of forest attributes at plot level. The variables are explained in Table 4.5.

N_{Trees}. CHM_{p95}, h_{max}, and h_{cv} are directly related to the height of trees and the variability. N_{Trees} and CC_{p} describe the density of the plot (in particular, CC_{p} is related to the proportion of gaps within a plot). $\text{Vol}_{\text{CHM-5}}$ and VN_{Trees} are considered as a proxy for the amount of biomass. The two variables CBH_{mean} and CBH_{CV} are respectively, the mean and the coefficient of variation of the CBH estimated for each detected tree by the allometric model. In the same way, BA_{allo} is the sum of the basal area of the detected trees, computed from the estimated CNH and scaled per hectare. VOl_{allo} is the sum of the individual volume of the detected trees, computed from the CBH and the height with the equation of volume mensuration found in [32], and scaled per hectare.

$$Aggreg_{\text{als}} = \frac{\bar{d}_{\text{Trees}}}{\left(\frac{1}{2 * \sqrt{\frac{N_{\text{Trees}}}{\text{Area}_{\text{Plot}}}}}\right)} \tag{2.1}$$

To assess the forest inventory attributes (number of trees per hectare, volume and basal area per hectare, quadratic mean of girth), we choose to use simple linear regression with a parsimonious selection of variables. Model selection was based on the best subset regression analysis (*regsubsets* in the R package "leaps" [90]) and with the bayesian information criterion (BIC). Field inventory data were used as reference. Assumptions for the linear regression (homoscedasticity, independence of errors, normality of error distribution, absence of bias) were checked (with several tests using the *car* and *lmtest* packages [46, 167]) to ensure model

validity. The multicollinearity was also tested with the variance inflation factor (VIF). A VIF value of less than 2 was considered the threshold for no collinearity. The quality of the regression models was assessed using the R^2 and the Root Mean Square Error (RMSE) values.

Table 2.5: Sixteen variables are selected for construction of forest attribute models. These variables are from the local maxima detection, or are computed at plot-level (ABA), or are a combination of local maxima detection and ABA.

Variable Name	Description	Type
N_{Trees}	Number of trees (classified as true positives local maxima) within the plot	ITD
h_{max}	local maximum with maximum height within the plot	ITD
h_{cv}	Coefficient of variation of local maxima height within the plot	ITD
$\text{Aggreg}_{\text{als}}$	Clark & Evans aggregation index computed from local maxima	ITD
CHM_{p95}	95^{th} percentile of CHM within the plot	ABA
CC_{p}	Canopy cover is the percentage of plot area covered by canopy (canopy is identified by a height threshold of 5 m on the CHM)	ABA
$\text{Vol}_{\text{CHM-5}}$	Volume between canopy (defined by CHM) and 5-m-height threshold, normalized by plot area	ABA
$\text{Vol}_{\text{CHM-16}}$	Volume between canopy (defined by the CHM) and 16-m-height threshold, normalized by plot area	ABA
$\text{V}_5\text{N}_{\text{Trees}}$	Product of $\text{Vol}_{\text{CHM-5}}$ and N_{Trees}	Mixed
$\text{V}_{16}\text{N}_{\text{Trees}}$	Product of $\text{Vol}_{\text{CHM-16}}$ and N_{Trees}	Mixed
$\text{V}_5\_\text{N}_{\text{Trees}}$	Ratio of $\text{Vol}_{\text{CHM-5}}$ and N_{Trees}	Mixed
$\text{V}_{16}\_\text{N}_{\text{Trees}}$	Ratio of $\text{Vol}_{\text{CHM-16}}$ and N_{Trees}	Mixed
CBH_{mean}	Mean of the estimated CBH for the detected trees	ITD (Allo)
CBH_{CV}	Coefficient of variation of the estimated CBH for the detected trees	ITD (Allo)
BA_{allo}	Sum of the Basal Area of the detected trees scale per hectare	ITD (Allo)
Vol_{allo}	Sum of the Volume of the detected trees scale per hectare	ITD (Allo)

2.4 Results

2.4.1 *H-C allometry*

The construction of a H-C allometry has led to several models (Tables 2.6, 2.7 & 2.8 and Figures 2.5 & 2.6) to estimate individual tree CBH with individual tree height extracted from a CHM and with the contribution of several other variables. These ancillary variables are describing the environment of the plot in which are the trees (altitude and slope) or describing the surrounding stands (basal area, number of stems, quadratic mean of circumference and age). The coefficients of the different models presented in Tables 2.6 are significant. Some models are species-specific (one line per species). Model 2 was used for the forest attributes assessment to compute variables at plot-scale. The inclusion of ancillary information allows to improve the quality of allometries. The reported %RMSE (Table 2.8) are good.

Table 2.6: Several models were fitted with the gnls function in R for Norway Spruce (NS), Douglas-fir (DF) or both together (All or All (rd) for random effect). Besides individual height (H) extracted from the CHM, ancillary variables are used to describe the environment of the plot in which are the trees (altitude and slope) or describing the surrounding stands (basal area, number of stems, quadratic mean of circumference and age).

	Species	**Model**
		Tree information only
model 1	All	$CBH = 2.45 * H^{1.19}$
	NS	$CBH = 2.2 * H^{1.22}$
	DF	$CBH = 3.85 * H^{1.05}$
		Integration of environmental variables
model 2a	All	$CBH = 1.14 * H^{1.22} * Altitude^{0.13} * Slope^{-0.04}$
model 2b	All	$CBH = 0.91 * H^{1.21} * Altitude^{0.16}$
	All (rd)	$CBH = 0.75 * H^{1.24} * Altitude^{0.17}$
model 2c	All	$CBH = 2.55 * H^{1.21} * Slope^{-0.06}$
	All (rd)	$CBH = 2.25 * H^{1.25} * Slope^{-0.06}$
		Integration of structural variables (needing a preliminary estimation by ABA for example)
model 3a	All	$CBH = 42.35 * H^{0.5} * BasalArea^{0.3} * NTrees^{-0.27}$
	NS	$CBH = 52.53 * H^{0.45} * BasalArea^{0.33} * NTrees^{-0.29}$
	DF	$CBH = 11.56 * H^{0.8} * BasalArea^{0.13} * NTrees^{-0.12}$
model 3b	All	$CBH = 12.32 * H^{0.92} * NTrees^{-0.12}$
	NS	$CBH = 12.83 * H^{0.93} * NTrees^{-0.13}$
	DF	$CBH = 7.2 * H^{0.97} * NTrees^{-0.06}$
model 3c	All	$CBH = 2.98 * H^{1.21} * BasalArea^{-0.07}$
model 4	All	$CBH = 25.43 * H^{0.56} * Altitude^{0.05} * Slope^{-0.02} * BasalArea^{0.27} * NTrees^{0.25}$
	NS	$CBH = 32.38 * H^{0.51} * Altitude^{0.05} * Slope^{-0.02} * BasalArea^{0.3} * NTrees^{-0.27}$
	DF	$CBH = 6.9 * H^{0.82} * Altitude^{0.07} * Slope^{-0.02} * BasalArea^{0.14} * NTrees^{-0.11}$
model 5a	All	$CBH = 21.09 * H^{0.54} * Cqm^{0.51}$
	All (rd)	$CBH = 22.43 * H^{0.52} * Cqm^{0.53}$
	NS	$CBH = 24.46 * H^{0.49} * Cqm^{0.55}$
	DF	$CBH = 8.59 * H^{0.81} * Cqm^{0.23}$
model 5b	All	$CBH = 13.48 * H^{0.6} * Altitude^{0.05} * Slope^{-0.02} * Cqm^{0.47}$
	NS	$CBH = 9.58 * H^{0.55} * Altitude^{0.05} * Slope^{-0.02} * Cqm^{0.51}$
	DF	$CBH = 6.14 * H^{0.84} * Altitude^{0.05} * Slope^{-0.02} * Cqm^{0.21}$
		Integration of stand age
model 6a	All	$CBH = 2.48 * H^{0.88} * Age^{0.26}$
	NS	$CBH = 2.47 * H^{0.77} * Age^{0.34}$
	DF	$CBH = 3.4 * H^{0.77} * Age^{0.29}$
model 6b	All	$CBH = 1.8 * H^{0.94} * Altitude^{0.06} * Slope^{-0.05} * Age^{0.23}$

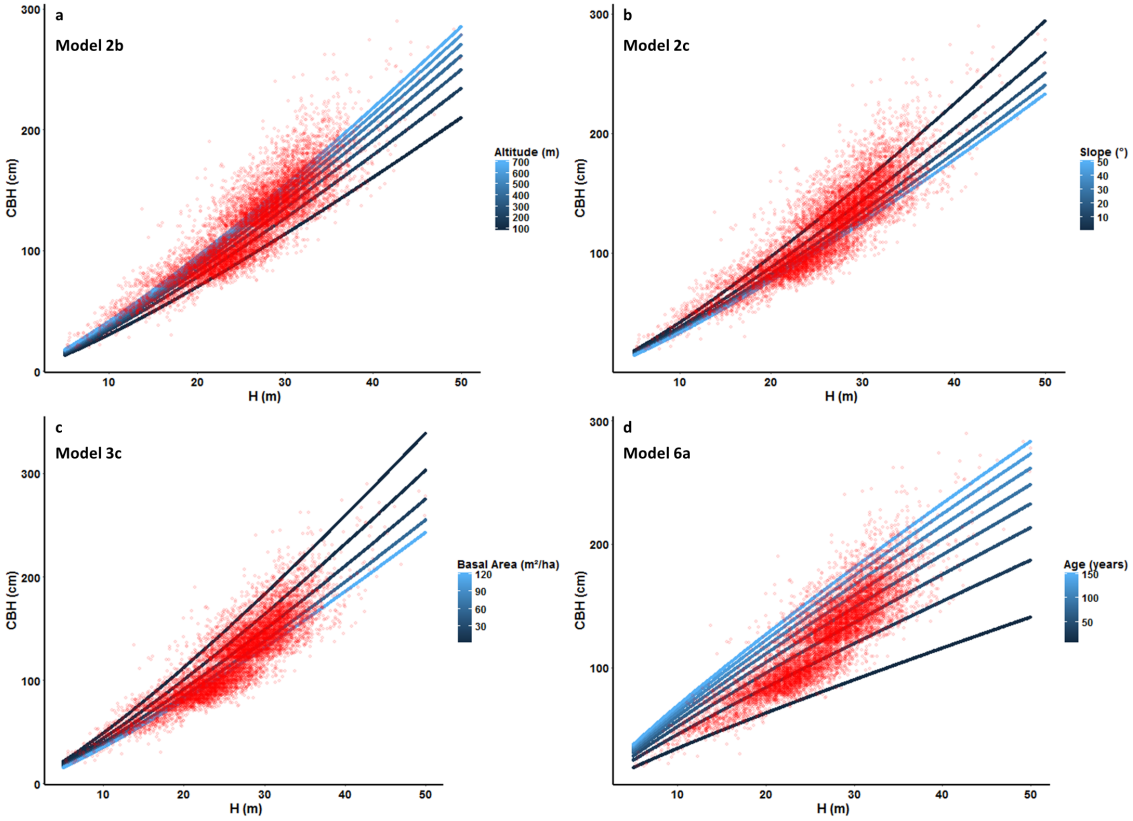


Figure 2.5: Selection of allometric models. The red points correspond to the tree data extracted from the Regional Forest Inventory database (total height in m and breast-height circumference in cm) for the two coniferous species (Norway spruce and Douglas-fir). The four panels are a selection of four allometric models (reference of the model in Table 2.6) without distinction of species and with integration of one ancillary variable: altitude for (a), slope for (b), basal area for (c) and age for (d). The blue gradient lines are the model curves with different values of the ancillary variables.

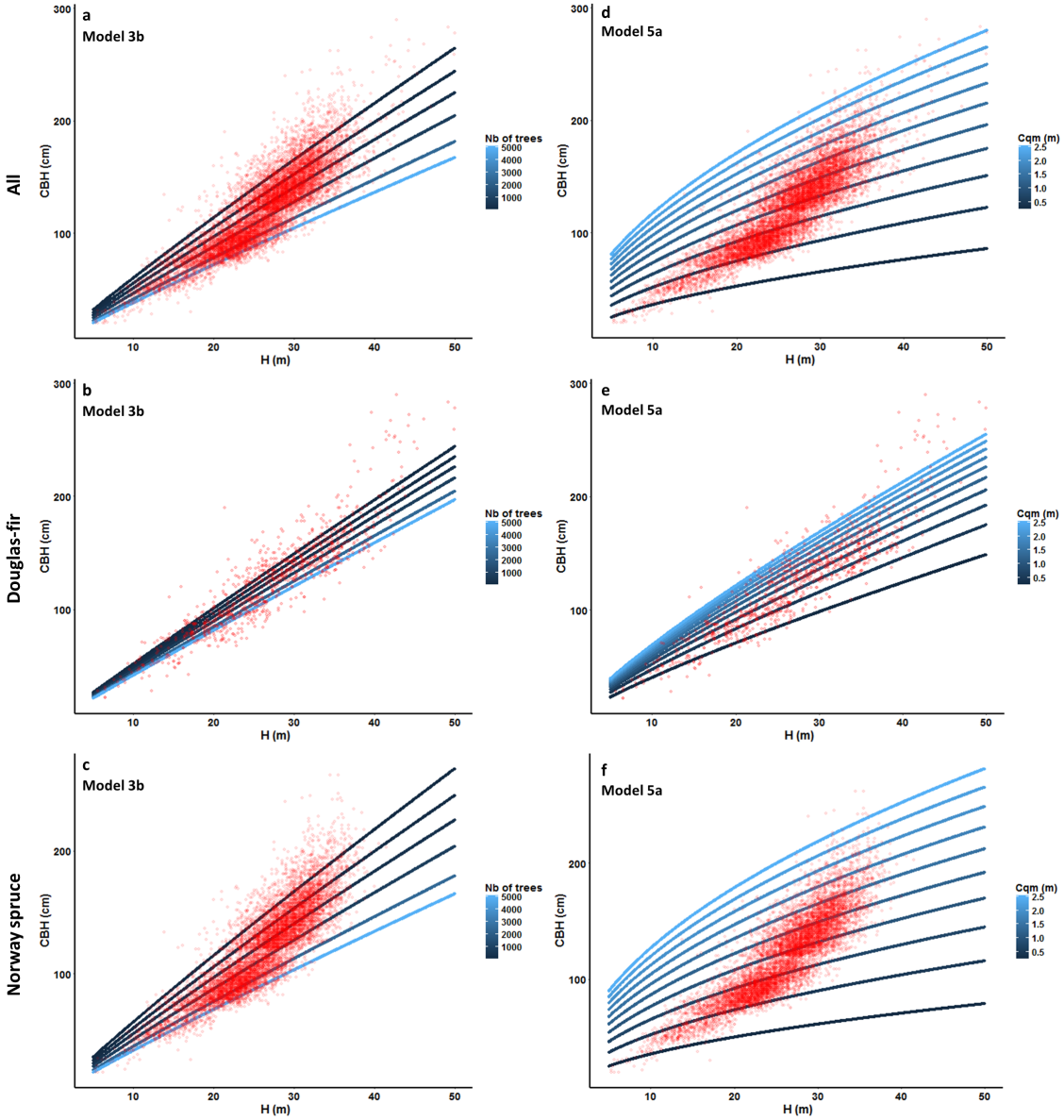


Figure 2.6: Selection of two species-specific allometric models with integration the number of stems (left side) or quadratic mean of circumference (Cqm, right side) as structural ancillary variable. The red points correspond to the tree data extracted from the Regional Forest Inventory database (total height in m and breast-height circumference in cm) for both Norway spruce and Douglas-fir (a) and (b), for Douglas-fir only (c) and (d), and for Norway spruce only (e) and (f) . The details of the models are found in Table 2.6 following the reference in the panel. The blue gradient lines are the model curves with different values of the ancillary variables.

Table 2.7: Comparison of the H-C models based upon AIC and deltaAIC which is the difference between the AIC for the current model and the lower AIC.

Model	Type	AIC	deltaAIC
model 1	All	82369.87	5675.8
	Species-specific	82265.01	5570.9
model 2a	All	81171.83	4477.8
model 2b	All	81442.73	4748.7
	All (random)	81819.86	5125.8
model 2c	All	81715.64	5021.6
	All (random)	82210.29	5516.2
model 3a	All	77421.61	727.5
	Species-specific	77062.37	368.3
model 3b	All	79690.16	2996.1
	Species-specific	79563.63	2869.6
model 3c	All	82058.45	5364.4
model 4	All	77033.07	339.0
	Species-specific	76694.07	0
model 5a	All	77468.14	774.1
	All (random)	77891.66	1197.6
	Species-specific	77136.04	442.0
model 5b	All	77074.40	380.3
	Species-specific	76759.22	65.1
model 6a	All	79349.48	2655.4
	Species-specific	78819.57	2125.5
model 6b	All	78617.23	1923.2

2.4.2 *Forest Attribute Assessment*

In this study, we focused on the assessment of four forest attributes to characterize coniferous plots (volume, basal area and number of stem per hectare, quadratic mean of circumference) with a low number of selected predictors (Table 4.5). A model with ABA variables only and a model with all types of variables were fitted for each attribute (Table 2.9 and Figure 2.8). The inclusion of ABA, tree and allometric variables improved undeniably the results. However, especially in the case of Cqm, the ABA model gives good results which is encouraging in an iterative approach to obtain a preliminary estimate of this attribute, which could be injected as complementary variable in an allometric model. The presented models respect the assumptions for linear regression, except for the number of stems which is the most problematic model. It can be explained by the high omission error of detected trees coming from *Forestimator* (Fig 2.7).

In order to improve our basal area model, we tested the integration of Cqm and N/ha previously estimated by our combined ITD + ABA approach. This strategy allows to globally enhance the model quality by increasing the R^2 from 39% to 44% and the %RMSE from 19.4% to 18.6%.

Table 2.8: Comparison of the H-C models based upon bias, RMSE and %RMSE. The p-Value comes from a paired t-test used to compare the observed and predicted CBH values.

Model	**Type**	**Bias** (cm)	**p-Value**	**RMSE** (cm)	**%RMSE**
model 1	All	-0.01	0.97	17.5	14.5%
	Species-specific	0	0.98	17.4	14.4%
model 2a	All	0	0.99	16.3	13.5%
model 2b	All	0	0.99	16.5	13.7%
	All (random)	-0.01	0.39	16.5	13.7%
model 2c	All	0	0.99	16.8	14%
	All (random)	-0.09	0.45	16.8	14%
model 3a	All	0	0.99	13.4	11.1%
	Species-specific	0	0.99	13.2	10.9%
model 3b	All	0	0.98	15.4	12.8%
	Species-specific	-0.01	0.97	15.3	12.7%
model 3c	All	-0.01	0.97	17.2	14.3%
model 4	All	0	0.99	13.12	10.9%
	Species-specific	0	0.99	12.9	10.7%
model 5a	All	0	0.99	13.5	11.2%
	All (random)	0	0.96	13.5	11.2%
	Species-specific	0	0.99	13.3	11%
model 5b	All	0	0.99	13.2	11%
	Species-specific	0	0.99	12.9	10.8%
model 6a	All	-0.01	0.96	15.6	13%
	Species-specific	-0.01	0.97	15.1	12.6%
model 6b	All	-0.02	0.91	14.9	12.4%

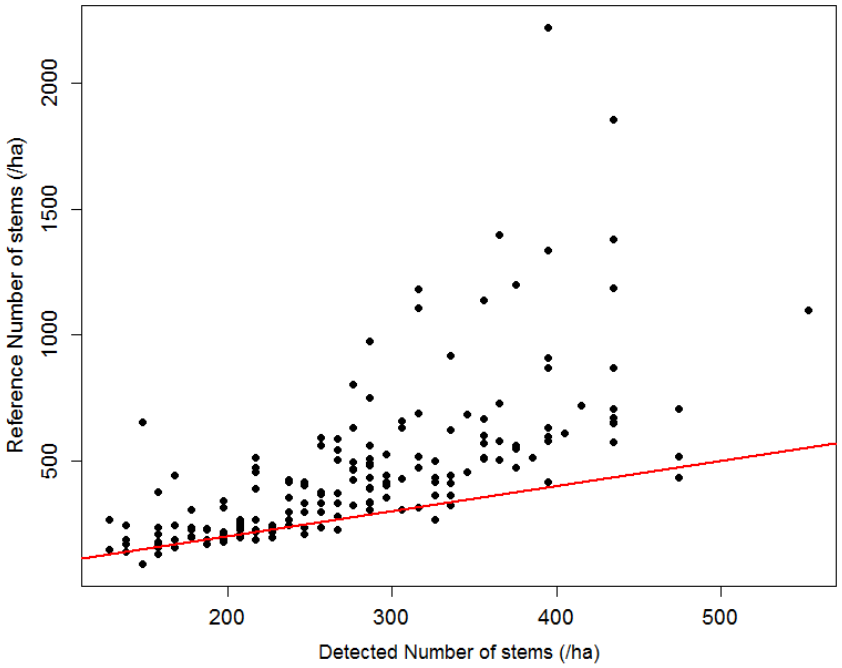


Figure 2.7: Comparison of the number of stems detected by *Forestimator* [38] with the reference number of stems for each plot selected from the RFI database and synchronous with the ALS dataset.

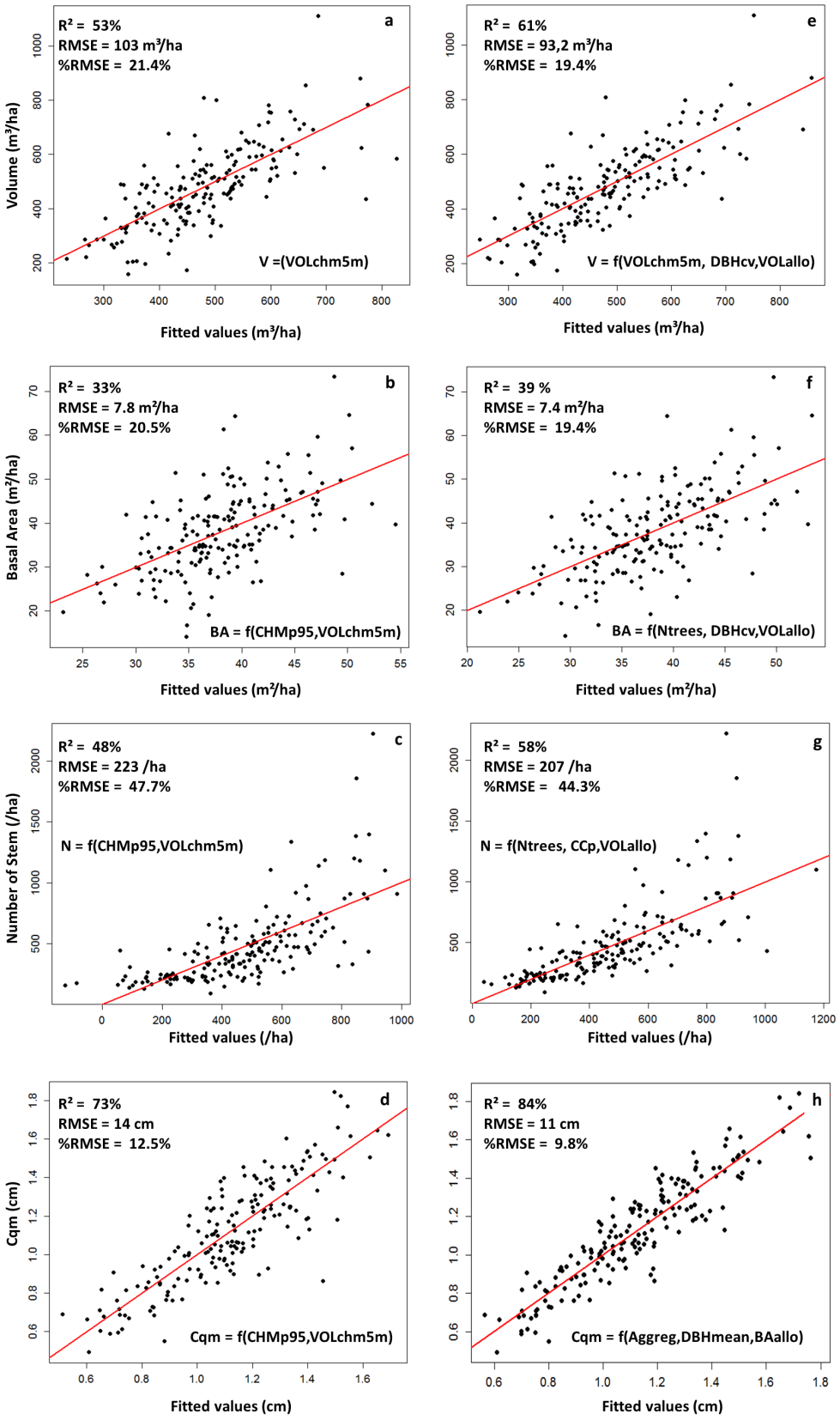


Figure 2.8: Observed versus fitted values for the forest attributes models. Comparison between area-based only (a to d) and mixed (e to h) approaches.

Table 2.9: Performance and variables selected for each forest attribute model and comparison between ABA only with all predictors (mixed approach).

Attributes	Type	Variables	Intercept	Coefficients
Volume	ABA	$Vol_{CHM\text{-}5}$	91.94	30.16
(m^3 ha^{-1})	All	$Vol_{CHM\text{-}5}$, CBH_{CV}, Vol_{allo}	135.09	12.19; 0.73; 0.45
Basal area	ABA	$Vol_{CHM\text{-}5}$, CHM_{p95}	35.36	2.43; -1.06
(m^2 ha^{-1})	All	N_{Trees}, CBH_{CV}, Vol_{allo}	3.43	0.05; 0.15; 1.6
Number of trees	ABA	$Vol_{CHM\text{-}5}$, CHM_{p95}	1835.80	-74.79; 49.85
(ha^{-1})	All	N_{Trees}, CC_p , Vol_{allo}	-213.62	2.13;3.09; -0.45
Quadratic mean	ABA	$Vol_{CHM\text{-}5}$, CHM_{p95}	-0.39	-0.03; 0.07
of girth (cm)	All	$Aggreg_{als}$, CBH_{mean}, BA_{allo}	-0.39	0.17; 0.01; -0.01

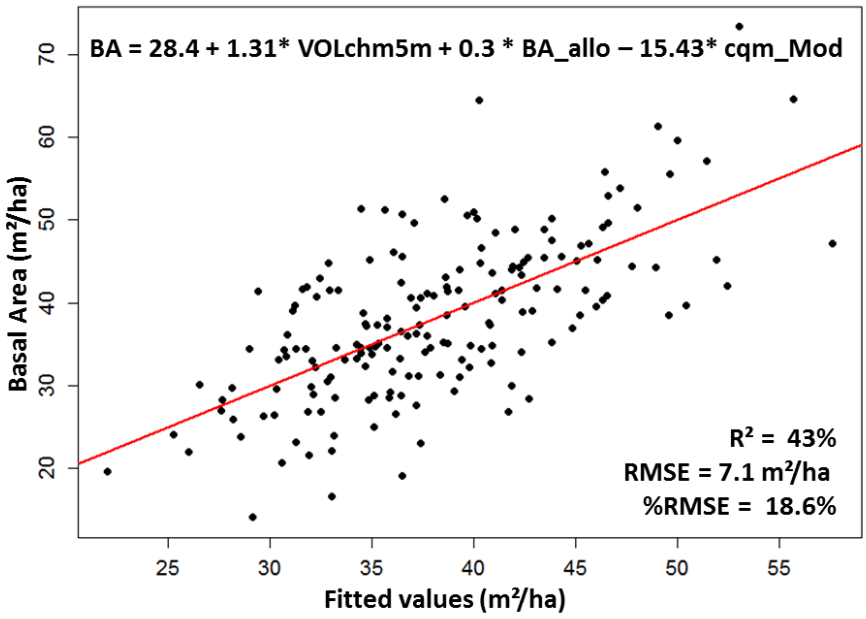


Figure 2.9: Observed versus fitted value for the basal area estimation model that integrates the estimation of the quadratic mean of breast-height circumference (Cqm) as an estimator.

2.5 Discussion

ALS data processing makes it possible to detect and individualize trees and estimate their total height and thus estimate top-height and density (number of tree per hectare) of stands. These data, combined with allometric models could be sufficient to estimate fairly precisely other important dendrometric parameter such as total basal area or total stand volume. Our approach aimed at testing a method for combining ALS and RFI data to monitor the coniferous forest resource at regional-scale. The first step was the construction of a H-C allometry based on regional dataset and enhanced with environmental and structural variables. The purpose of this allometry was to be used with local maxima detected on an ALS CHM, to enrich the set of variables for ITD and to benefit of the potential of a tree-level approach. With this study, we wanted to test the potential of the data used in the *Forestimator* plug-in with the underlying idea to complete this tool. The goal was to take advantage of an existing tool and methodology that was originally only dedicated to the top height assessment. Our research has led us to highlight several key points presented below.

2.5.1 *RFI dataset - Representativity for allometry*

The dataset from the RFI which is the basis of the development of allometries can raise questions. Indeed, in view of the methodology of the regional inventory, only the height-circumference pairs of the three dominant trees are measured, which could represent a problem of representativeness. However, considering that the allometry will be used on the local maxima detected from the CHM, so on data assimilated to dominant trees, there is always a consistency. A second point considering the representativeness of the data is the proportion between Norway spruce and Douglas-fir. Although overall, an improvement of the allometries is observed with models specific to each species, the coefficients are close to those of the Norway spruce, which represent 91% of the dataset. The evolution of forest areas will therefore be a factor to take into account.

2.5.2 *H-C Allometry*

Five types of models were tested species-specific or not. Introducing the top height variable was tested but never significant, alone or combined with other variables. The allometries are characterized by an error-ranging from 10.7% to 14.5%, with a variable degree of complexity. The integration of ancillary variables describing the surrounding plot had improved the quality, for example the best model was the species-specific model 4 which used four ancillary variables (10.7% %RMSE and lowest AIC and RMSE).

The integration of ancillary variables characterizing the plot (slope, altitude, stand density or development stage) had improved the quality of the results. When there is convergence, considering the species-specific model improves slightly the results. The greater proportion of Norway spruce explain probably this fact, as the coefficients of the All-species models are close to the Norway spruce models.

Altitude and slope have a low impact (low power coefficient value). On equal height, slope has a negative impact on circumference. Adding structural variables about the stand change the type of the relationship (power value lower than 1). Models with the age variable are intermediate in terms of quality. Taking the age into account has the advantage to improve the model but without difficulty. Indeed, age is a information easily available for even-aged stands (planting date). The use of structural variables (N/ha, BA and Cqm) is dependent on the possibility to extract these information reliably, for example by an ABA which would be seen as a preliminary step in an iterative workflow. The introduction of a random effect on the coefficient a has had few impact on the coefficients and quality of the models 2b, 2c, 5a.

The different plot variables can be interpreted empirically: the altitude is related to the fertility of the stand, the Cqm and age are a measure of the development stage and slope can be read as a corrective factor of the height estimated by ALS, and an indicator of the growing condition. The effect of the structural variables is higher for Norway spruce than for the Douglas-fir. This could be explained by the greater diversity of stand conditions for Norway spruce. This latter is wider spread and for longer including in unsuitable situations.

2.5.3 *RFI dataset - Tree positioning*

An important step but particularly time-consuming was the repositioning of the inventory data in relation to the CHM. The field data were synchronous with the ALS data but although positioned at the GPS, the quality of the positioning was rather poor, probably because of the device used. In order to improve the location of trees mapped in the field, a step of co-registration with the local maxima from the CHM has been implemented. A pattern recognition algorithm was used to compare the arrangement of the trees with that of the local maxima, in order to be able to translate the trees.

One difficulty encountered was that the RFI plots are concentric and the lack of measurement for all the trees. On the other hand, the presence of undetected trees is also a source of confusion. The matching of identification is very dependent on the detection of local maxima. A total of 179 plots were repositioned, making it possible to have a good dendrometric diversity of the dataset. A large search radius had been fixed (70 m around the theoretical plot center) to manage the

potential important uncertainty of the plot location. The algorithm frequently offers many possible solutions for matching local maxima and field trees. This phase of the work must therefore be considered semi-automatic, more as a tool for facilitating work, but always involving expertise and human interpretation, which is indispensable.

2.5.4 *Forest attributes assessment*

Forest attributes models at plot-scale were implemented based on a low number of predictors to follow a parsimonious approach and avoid over-fitting. We compared two approaches for forest attributes assessment: ABA and ITD. For both, models were fitted with variables computed from the CHM only for ABA and for CHM and local maxima extracted from CHM for ITD.

Several variables were chosen as predictors in forest attributes models either in ABA or ITD. This study used only a few numbers of variables (by comparison to the possible metrics computed from ALS datasets). The chosen variables were similar to those use in [21] with UAS data and closely related to the attributes of interest. The accuracy was significantly higher for models combining the different kind of variables (plot, tree, mixed, allometric).

The volumetric variable $Vol_{CHM\text{-}5}$ is the most frequent predictor in the different models. The models provided R^2 ranging from 33% to 73% for ABA and for 39% to 84% ITD. The assessment of the number of stems is the weakest model, especially considering high densities. The best model is obtained for the quadratic mean of circumference (R^2 = 84% and %RMSE = 9.8%). The integration of variables coming from ITD and allometric model allowed to improve the quality of fitting for the four forest attributes.

The assessment of Cqm gave the best result either for ABA only or mixed approach. The use of this structural attributes seems henceforth a good compromise when considering an allometric model with ancillary variables.

The model for N/ha is troublesome for high density of stems (but not so high if we consider values greater than 500 trees per ha). Omissions concern few dominant trees which certainly contribute mostly to the wood volume. These tree detection problems were not felt in the previous work about top height in *Forestimator*. Errors due to omissions have more impact on small areas but should be smoothed over larger areas of interest (stands instead of plots, for example). A stratified method for assessment with an a priori diagnosis of stand density could mitigate such issue. This limitation of the method comes from probably because of the resolution of data :a kind of saturation of the CHM limiting the tree detection. But the type of dataset in this study is easy to use, regionally available at low cost, meaningful for low density (aged) stands.

2.5.5 *General considerations*

When comparing ABA and ITD, an advantage of the latter is the possibility of a more complete and integrated approach including species discrimination. Also, a parallel can be drawn between ITD and allometries classically used in forest science. In this study, we change the paradigm and we developed allometries for using with remote sensing data instead of field data. An individual tree approach is particularly interesting in the concept of precision forestry, for example in the context of uneven-aged stand transformation, while an ABA must also be seen as a mapping approach.

In an operational perspective, the use of simple and versatile models is attractive, especially for successive ALS dataset acquisitions. The success of the appropriation of a method by the forest managers depends on the relative simplicity, in order to make accept the interest of this method and ease the knowledge transfer.

The integration of the age in the allometries allows to improve the quality of the models. In contrast to the other tested and stand-related ancillary variables, this data has the advantage of being easy to obtain, particularly in the case of coniferous stands with a known planting date. The age is already used in *Forestimator* and allows to update the top height and to compute the Site Index through its use with ALS data.

The integration of an allometric model and our approach to estimating forest attributes could be integrated into *Forestimator* relatively easily. This would provide a more complete characterization of coniferous plots. In addition, the integration of other models of growth or simulation could be interesting to deploy more easily the cartographic dimension and the accessibility of this type of approach.

Forest attribute models have been developed at the scale of inventory plots. Their use must therefore be set up on the same scale but can be considered in different ways. First, the implementation of these models can be handled with a network of existing plots, for example to update the data. Secondly, an artificial network of plots can be created on an area of interest that one wishes to describe. Finally, a map approach in full to characterize stands deserves to be validated. A division into pixels (from 10 to 30 m of resolution for example) of stands polygons coupled with our approach would make it possible to analyze the spatial variation of these forest attributes.

Our study has further strengthened our confidence that an iterative and stratified approach is necessary for a reliable assessment of forest attributes. As well as to promote a better cooperation with the RFI. Especially, a reflexion about the methodology of the field data collection should be undertook to adapt the in-situ

data to remote sensing considerations. As such, the concept of concentric plots is a feature that complicates synchronization with remote sensing data.

We didn't use information on crown diameter because we didn't want to manually digitalizing crown based on CHM or aerial images. We wanted to implement a simple but versatile approach at regional-scale as we considered the resolution of the CHM was too weak compared with other types of data. Besides, tree crown delineation could be computationally intensive and limited to the upper canopy even with 3D point clouds if the point density is low. The required cost and effort may not be worth it.

2.6 Conclusions

In this study, we developed and tested allometric equations combined with the use of ALS data. Our approach aimed at testing a method for combining ALS and RFI data to monitor the coniferous forest resource. The first key point was the construction of a H-C allometry based on regional data and enhanced with environmental and structural variables. The purpose of this allometry was to be used with local maxima detected on an ALS CHM, to compute attributes at tree-level. Five types of models were tested species-specific or not. The allometries are characterized by a %RMSE ranging from 10.7% to 14.5%, with a variable degree of complexity (Tables 2.6 & **??**). The integration of ancillary variables characterizing the plot (slope, altitude, stand density or development stage) had improved the quality of the results.

Then forest attributes models at plot-scale were implemented based on a low number of predictors to follow a parsimonious approach and avoid over-fitting. We compared two approaches for forest attributes assessment: ABA and ITD. For both, models were fitted with variables computed from the CHM only for ABA and for CHM and local maxima extracted from CHM for ITD. The H-C allometry was also applied to the detected trees to enrich the set of variables for ITD and to benefit of the potential of a tree-level approach. The volumetric variable $Vol_{CHM\text{-}5}$ is the most frequent predictor in the different models. The models provided R^2 ranging from 33% to 73% for ABA and for 39% to 84% for combined ABA and ITD. The assessment of the number of stem is the weakest model, especially when considering high density stands. The best model is obtained for the quadratic mean of circumference (R^2 = 84% and %RMSE = 9.8%). The integration of variables coming from ITD and allometric model allowed to improve the quality of fitting for the four forest attributes.

The present study has only investigated forest attributes assessment in coniferous stands. In Belgium, these stands are mainly even-aged resulting from plantations. By contrast, broadleaved forests are uneven-aged (resulting mostly

from natural regeneration) which makes individual tree modeling more complex. The advantage of working on coniferous stands implies a low incidence of the date of acquisition of ALS data, unlike broadleaved stands. Further work is needed to developed relevant and reliable approach in such kind of forest. In the same way, an advantage of ITD is the potential for a more complete workflow including species discrimination.

Chapter 3

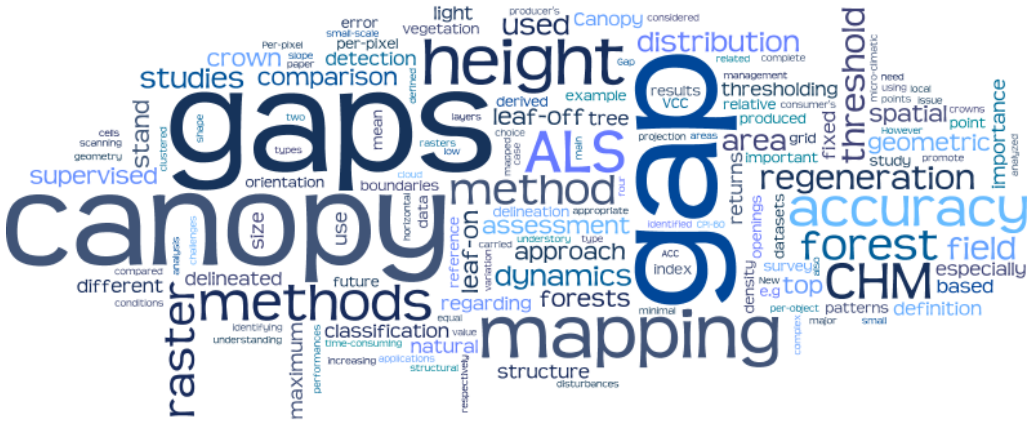
CANOPY GAP MAPPING IN BROADLEAVED FORESTS WITH ALS

Chapter 3

Canopy Gap Mapping in Broadleaved Forests with ALS

"Il n'y a pas de lumière sans ombre."

Louis Aragon



This chapter aims to estimate and map canopy gap with ALS data. Canopy gaps are small-scale openings in forest canopies which offer suitable micro-climatic conditions for tree regeneration. Field mapping of gaps is complex and time-consuming. Several studies have used Canopy Height Models (CHM) derived from ALS to delineate gaps but limited accuracy assessment has been carried out, especially regarding the gap geometry. This study offers a new prospect on canopy gap mapping which is a major issue in sustainable forestry to promote natural regeneration.

The ability of ALS to describe the height and the forest structure, even under the canopy is a huge advantage compared to other remote sensing data as photogrammetric CHM. The raster data used in this research are produced from high density ALS points cloud with a 50 cm resolution at a watershed-scale. The precision and the richness of the ALS data depend on the point cloud density, determined by the flight parameters. In this framework, we investigated three mapping methods based on raster layers produced from ALS leaf-off and leaf-on datasets: thresholding, per-pixel and per-object supervised classifications with Random Forest. In addition to the CHM, other metrics related to the canopy porosity were tested. The Geometric Accuracy was analyzed with the gap area, main orientation, gap shape-complexity index and a quantitative assessment index of the matching with reference gaps polygons.

We highlight four originalities of our work: (i) a detailed geometric accuracy assessment of the gaps which was missing in previous studies, especially in broad-leaved forests; (ii) the use of leaf-off and leaf-on datasets; (iii) a comparison of several methods including some which have not been previously used for gap mapping (e.g. supervised classification with random forest and image segmentation) and (iv) the analysis of the influence of the forest stand type on the methods performances.

3.1 Introduction

Defined as small-scale openings in forest canopy, gaps are created by management activities (thinning) or natural disturbances, such as windthrow [39] or natural mortality. These openings alter micro-climatic conditions, increasing plant diversity and allowing regeneration by increasing light levels [40, 130]. Because of their high regeneration potential, gaps are important for foresters seeking to promote a nature-based silviculture, by emulating natural disturbances.

Spatial characteristics (size, shape, distribution, orientation) of gaps in forests are of central importance in understanding regeneration, ecosystem dynamics, species diversification and distribution [16, 66, 91, 128, 147, 154]. [73] highlighted the importance of gaps with different spatial properties for seed establishment and determining the future forest structure. [84] proposed horizontal heterogeneity due to presence and distribution of gaps as a mechanism for conserving forest biodiversity by maintenance of stand structural complexity. Gap shape, orientation and size variation affect the gap dynamics, mainly through the variation in the light environment and moisture levels [84, 122]. In his study, [55] used gap shape metrics derived from Unmanned Aerial Vehicles (UAV) images to assess the plant diversity in forests.

Reliable survey methods are an important prerequisite for the analysis of spatial patterns of gaps [66]. However, mapping and characterising canopy gaps is a complex issue. Delineation of gaps in the field is difficult, time-consuming and there is a lack of consensus about canopy gap definition and methods for describing regeneration [14, 147, 166, 130]. Besides, field survey is not always reliable because of the difficulty of the ocular evaluation of the exact limits of the gap [149].

A commonly used definition of a gap was introduced by [26] who defined it by the canopy drip-line, *i.e.*, the vertical projection of the external borders of the tree crown with an height greater than 2 m. This definition is pragmatic in the field but can lead to an underestimation of the area affected by the gap disturbance. Thereafter, [128] introduced the concept of expanded gap which extends the gap to the base of the trunks. But, this concept is not easily adapted for remote sensing purposes as the position of trunks is hard to identify from above [50]. The vegetation height is a common criterion to define the closing of a canopy gap. As Airborne Laser Scanning (ALS) has become a common technique to assess forest height, the use of Canopy Height Models (CHM) derived from ALS to delineate gaps has been explored in a number of studies [149, 66, 147, 7].

3.1.1 *Canopy Gap Mapping with ALS : A Background*

Applying a fixed height threshold on an ALS CHM is the most common method of canopy gap delineation, based on the definition of [26]. In most cases, studies differ by the value of the threshold and a constraint of minimal area. [7] and [24] quantified gap-size frequency distribution in Southern Peruvian Amazonia by mapping gaps with a height threshold of, respectively, 1 to 20 m and 2 m with a minimum area of 2 m^2. [149] and [147] chose a fixed 5 m height with a minimal gap size of 5 m^2. [149] focused on gap dynamics, with two co-registered CHM (1998 and 2003), whilst [147] identified understory vegetation types. [66] studied patch dynamics, stand structure and spatial patterns in Pacific Northwest forests by analysing the spatial distribution of gaps, delineated as 3 m maximum height and larger than 9 m^2. In 2004, [72] post-processed delineated gaps with GIS functions (shrinking) to isolate gaps connected by corridors and to refine gap shape and boundaries.

A second method of canopy gap mapping is the use of a relative height threshold. Gaps are then defined depending on the neighbourhood stand height. This approach requires typically, the creation of a canopy top raster. The gaps are identified by comparison of the original CHM with the canopy top raster. In mangrove forests, [169] used opening and closing operations on a CHM to produce a surface representing the top of the canopy. Gap grid cells were detected after a

black top-hat transformation using a relative height threshold (0.65). According to the author, this approach gave better results than the fixed height method. However, no complete accuracy assessment was performed against field reference data. Instead, only a visual comparison of the gap boundaries was carried out, using a slope raster derived from the CHM and based on the assumption that the slopes along gap boundaries are steep. Such slope rasters were also previoulsy used in [72] for helping to identify the fixed height threshold to apply. In the study of [50], a canopy top raster was produced from a ALS CHM by applying a moving window and setting the maximum value to the grid cell center. Then, CHM grid cells with heights inferior to 66% of the corresponding height of the canopy top raster were classified as gap. [14] use a similar approach in Nothofagus spp. forest in New Zealand with a DSM produced from aerial photographs. In this case, gaps were delineated as areas showing a difference of 12 m between the two rasters.

To improve gap extraction, [50] mapped gaps directly from the point cloud by initially identifying the canopy area. Their processing method was initiated by the identification of local maxima corresponding to tree tops. Then, the returns with a relative height of less than 66% of canopy height were removed to retain only canopy returns, which were then clustered to a local maximum. The clustered canopy returns were merged and delineated to retrieve gap extent. This method avoids interpolation to a CHM and makes use of more complete height information from the return points. After comparison with the raster method, the authors demonstrated a small increase of accuracy compared to CHM methods, especially for low density ALS data, but the point cloud method is much more computationally expensive.

This latter approach favored a delineation of tree crowns and considers that gaps are found ”by default“. Intrinsically, gaps and crowns are complementary as their distribution defines the canopy structure. ALS sensors are able to capture canopy structure wall-to-wall and are used for the estimation of canopy structural variables as Vertical Canopy Cover (VCC), Angular Canopy Closure (ACC) or Leaf-Area index (LAI) [69, 70, 71]. For example, the VCC is the total crown projection area (defined by the outermost perimeter of the crown on the horizontal plane) divided by the stand area. In this case, small gaps inside the crown are considered to be a part of the crown. VCC and ACC are related to the penetration of light through the canopy and are considered as indirect measurement of light regimes [63].

3.1.2 *Aims of the Study*

ALS makes detection and analysis of canopy gaps easier over large areas [134]. But further work is needed to assess the potential of ALS for accurate gap mapping, especially regarding retrieval of gap area and shape, with the importance of these characteristics outlined previously.

In this context, we analyse the potential of ALS for canopy gap mapping in uneven-aged broadleaved stands with a focus on geometric accuracy of gap retrieval. For the purposes of this research, three mapping methods based on raster datasets produced from high density leaf-off and leaf-on ALS data are tested: Thresholding, supervised classification of individual grid cells ("per-pixel" classification) and per-object supervised classification.

In this study, priority is given to the raster-based approach, more commonly used and more suitable for use over extensive areas. Point-based delineation is not tested here due to the computational costs and as it produces only relatively small accuracy increases. Unlike in previous studies, we tested the use of ALS rasters other than CHM. We analyse if alternative metrics could improve gap delineation, especially in terms of retrieval of gap shape and borders. These rasters aim to capture the three-dimensional structure of points cloud and to describe the canopy porosity. For the thresholding method, several threshold values were tested. For the per-pixel and per-object approaches, we used supervised classification by Random Forest [25]. The maps produced are compared in terms of detection quality (through a confusion matrix) and geometric quality (through a comparison with gaps mapped in the field).

The influence of the season is also considered by the use of both leaf-off and leaf-on ALS datasets. Such a comparison of the effectiveness of several mapping methods on both leaf-on and leaf-off ALS data sets, is novel for canopy gap delineation. Besides, no previous studies have attempted to validate gap geometry reliability. Furthermore, we investigate a new approach by which to classify forest stands according to their canopy opening degree prior to gap mapping, to provide a means to identify the optimal gap delineation method.

3.2 Data Sets

3.2.1 *Study Site*

The study area ($\approx 18,000$ ha, Figure 3.1) was located in the watershed of the Houille river near the village of Gedinne (Wallonia, Belgium). Forests, which cover approximately 60% of the area, are composed of uneven-aged broadleaved stands with oak (*Quercus petraea* (Matt.) Liebl. and *Quercus robur* L.) and European

beech (*Fagus sylvatica* L.) and even-aged coniferous stands dominated by Spruce (*Picea abies* (L.) H. Karst.) and Douglas-fir stands (*Pseudotsuga menziesii* (Mirb.) Franco). Four compartments (management units of the Forest Service) were selected by a field visit and by photo-interpretation according to two main criteria: predominantly broadleaved stands and exhibiting varying degrees of canopy closure (Figure 3.1).

Forty-two stands compose the four compartments of the study area. The limits of the stands were available from the Forest Service but the quality of the borders were checked by photo-interpretation with ortho-images and corrected with ArcGIS 10.1. Tree regeneration is present in the selected stands with a variable cover (ranging from a few seedling stems to large seedling cover). The regenerating species are oak, European beech and birch (*Betula pendula Roth*). The dominant herbaceous species are *Pteridium aquilinum (L.) Kuhn, Rubus fruticosus L., Rubus idaeus L., Deschampsia flexuosa (L.) Trin, Luzula luzuloides (Lam.) Dandy & Willm* and *Vaccinium myrtillus L.*

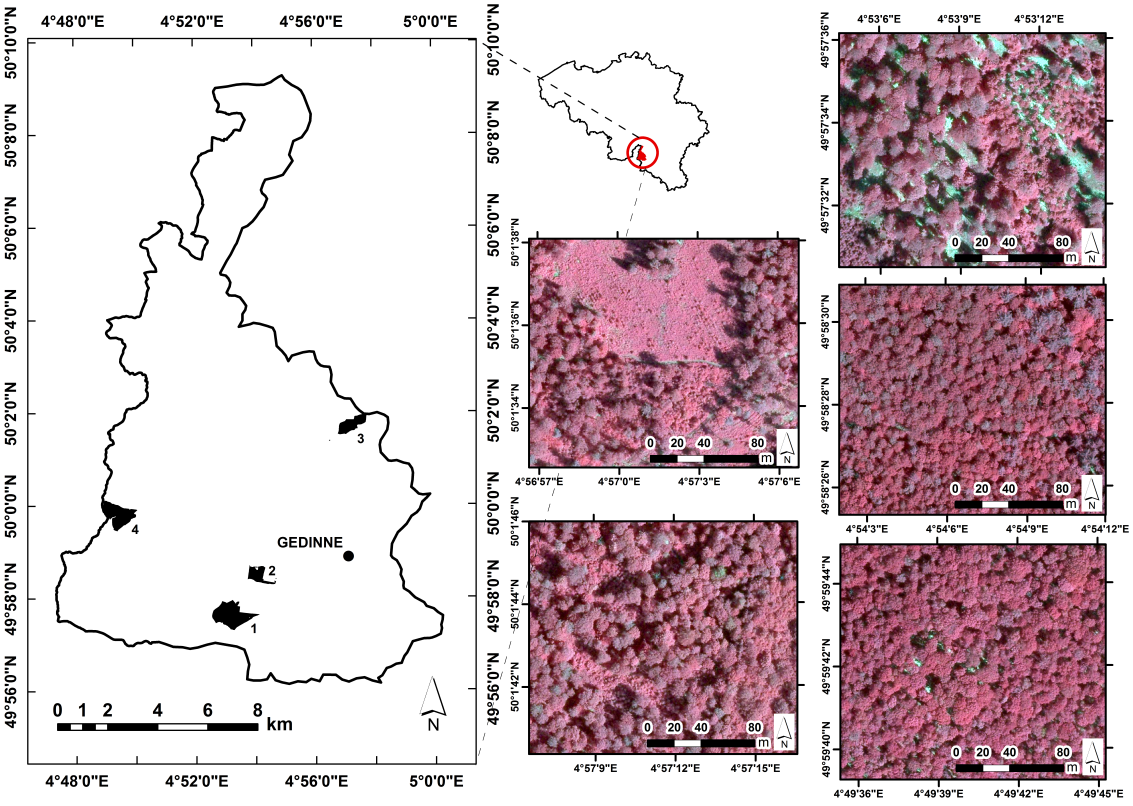


Figure 3.1: Boundary of the watershed of the Houille river, near the Belgium/France border. The four sites (compartments numbered one to four) are scattered in the study area and their areas are respectively 113 ha, 39 ha, 49 ha and 85 ha. Several types of stands with different degrees of canopy closure are illustrated by a subset of an infrared false color aerial image acquired in 2009 (50 cm resolution).

3.2.2 *Canopy Gap Definition*

For the purposes of this study, a clear and simple definition was necessary to identify gaps in the field. A definition similar to [26] was used. Throughout the study, we defined gaps as openings in the canopy with a minimum area of 50 m^2, a minimum width of 2 m and a maximum vegetation height of 3 m. The vegetation level value of 3 m is considered to indicate well-established regeneration [81] and is a critical height for the survival of regeneration, especially against ungulate predation. Most previous studies showed that ALS is able to detect very small size gaps (around 5 to 10 m^2). We consider gaps as priority areas for promoting natural regeneration, especially for oak which has survival difficulties below beech [3]. [27] concluded that the minimum opening area for oak regeneration is 70 m^2 and the optimal is between 100 to 120 m^2. Moreover, small gaps (<50 m^2) close very fast due to the crown growth of surrounding trees [107]. For this reason a minimum gap size of 50 m^2 is selected.

3.2.3 *Field Data*

Two field campaigns were conducted in August 2012 (in leaf-on conditions). First, a systematic grid (50 m × 50 m) was created covering each of the four study sites. Each point was visually interpreted in the field as “gap” or “not gap” areas (with a distinction of forest tracks and new gaps made by recent exploitation). A total of 1140 points were visited (295 of which fell in gaps). These data were used for the assessment of the gap detection quality. The corrected position of the theorical center of each 50 m × 50 m cell was established using a dGPS SXBlue II during the field visit.

Secondly, 39 canopy gaps (55 m^2 to 2400 m^2) were mapped in the field with a dGPS, electronic compass and laser rangefinder, with post-processing of the GPS data to obtain a planimetric error of, on average, 20 cm. The gap boundaries were delimited by the projection of the external borders of the crowns. Several points were recorded along the gap where the gap boundary was judged to significantly change its orientation [50]. These reference gaps were used to assess the geometric quality of the delineated gaps.

3.2.4 *ALS Data*

The ALS datasets were captured in March (leaf-off) and July (leaf-on) 2011 by a Riegl LMS-Q680 (300 KHz), with a maximum scan angle of 15 °. First and last returns were recorded and classified by the data provider as water, building, high vegetation (returns with a height >1.5 m), ground and unclassified, according to the LAS format standard. Intensity information was also available. The point

clouds were organised by 500 m × 500 m tiles. The average number of returns per 1 m^2 grid cell is 40 (average range between 1 and 238) and the average number of first returns is 21 per m^2 (average range between 1 and 190). The overall planimetric accuracy of the ALS dataset is 25 cm.

3.2.5 *Pre-Processing of ALS Data*

ALS data pre-processing was carried out using *Las2Pix*, a software developed by the Unit of Forest and Nature Management (University of Liège-Gembloux Agro-Bio Tech). This software is based on LASTools [61] and allows the production of raster metrics from the pre-classified ALS point cloud according to a number of different statistical functions and constraints and at a given spatial resolution. Each grid cell is assigned a value calculated from points falling inside the cell (e.g., maximum height for single returns only). Raster metrics were produced for each tile and thereafter, these were merged into one raster per compartment, at 50 cm resolution. This spatial resolution was chosen as a good compromise between the point density and the precision of the gap mapping (considering the minimal gap size of 50 m^2).

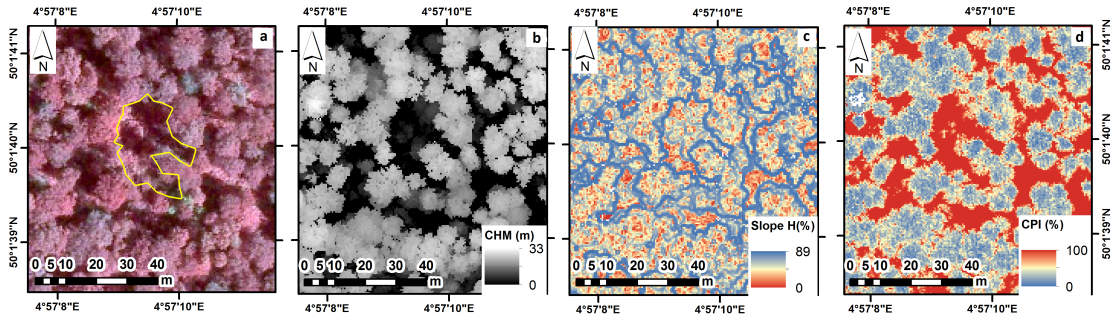


Figure 3.2: Metrics are produced from the raw airborne laser scanning (ALS) data. The yellow polygon corresponds to a field mapped gap over an infrared aerial image (a); The Canopy Height Model (CHM, b), the slope of the CHM (Slope-H, c) and the Canopy Porosity Index (CPI, d) are illustrated.

The CHM was generated as the height of the highest return in each grid cell. The Canopy Porosity Index (CPI) is the percentage of points below 3 m-high. This metric is inversely proportional to the amount of leaves, branches and stems, and considered as a measure of the degree of opening of the canopy (Figure 3.2). This is the complementary of the Canopy Closure as defined in [66]. As with the CPI, the Ground Point Ratio (GPR) is proportional to the porosity of the canopy and is calculated as the percentage of points classified as ground in each grid cell. The CPI and GPR are calculated on the basis of all returns. The Height-Scaled Crown Openness Index (HSCOI) was developed by [78] and is based on the voxel concept, calculated within a moving window. This index provides a quantitative

measure of the relative penetration of ALS returns into the canopy. Three other metrics were produced describing the canopy height: the standard deviation of height (Std-H), the coefficient of variation of height (CV-H) and the slope of the CHM (Slope-H). [169] used this latter metric to assess the quality of automatically delineated gaps boundaries, as many gap boundaries are well captured by rings of pixels with large slope values (Figure 3.2). Slope-H is tested here as a potential approach to refine gap boundaries. Finally, an intensity metric was generated as the mean intensity of single returns, to characterize the spectral response of the high vegetation (Mean-I-S).

3.3 Methodology

3.3.1 *Gap Mapping Methods*

The three methods implemented in this study for gap detection (thresholding, per-pixel supervised classification and per-object supervised classification) are described in the following sections (Figure 4.4).

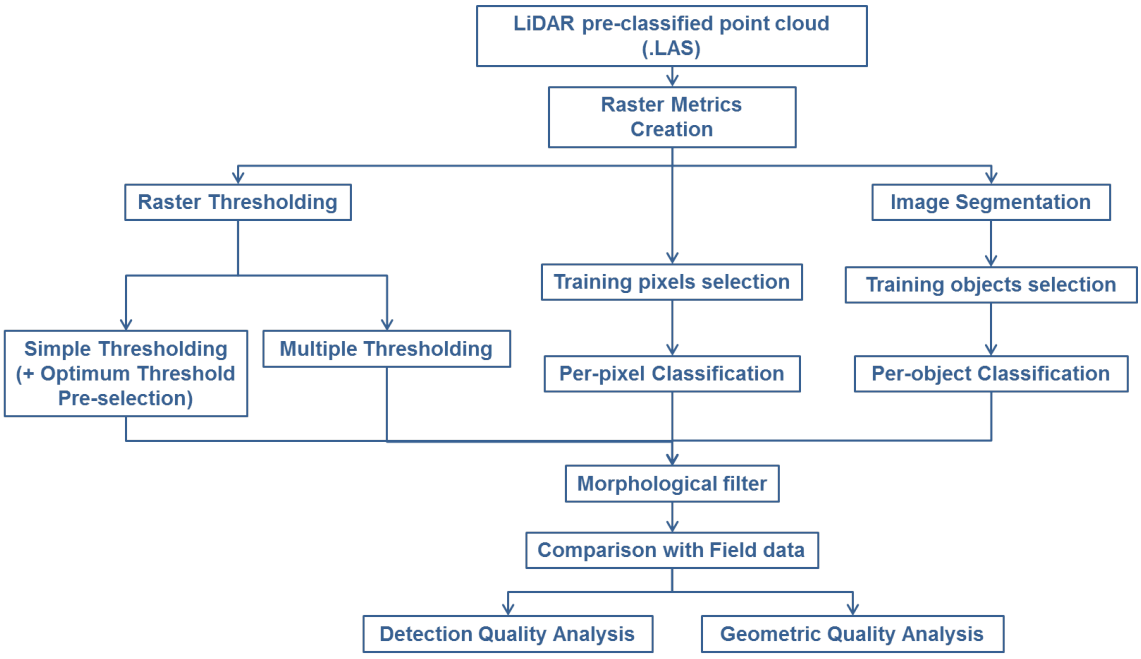


Figure 3.3: Synthesis of the mapping methods comparison. Three methods (thresholding, per-pixel and per-object supervised classifications) are compared. Per-pixel and per-object classifications are implemented with the Random Forest algorithm in R. Quality assessment is by a confusion matrix (detection quality) and by comparison with field-mapped reference gaps (geometric accuracy).

Thresholding

LiDAR canopy gap mapping is basically to discriminate canopy areas and non-canopy areas (corresponding to low vegetation or ground areas). In the simple thresholding approach, a fixed threshold value was applied to a single raster layer, either CHM, CPI or GPR to delineate gap areas. Three different metrics to identify non-canopy pixels were selected : using a fixed height threshold on a CHM, using a proportion of points classified as ground (GPR) or using a proportion of points below a fixed height threshold, considering presence of low vegetation and not only ground points (CPI). Several threshold values were tested for these three rasters and the result of the simple thresholding approach was a binary raster output (canopy or gap):

- CHM: gaps are identified as grid cells with a height below 3 m or 5 m;
- CPI and GPR: gaps are identified as grid cells with a CPI/GPR above a threshold. A specific step to determine the optimum threshold values for CPI and GPR (respectively for leaf-off and leaf-on) was implemented to avoid a systematic test of several values for each datasets. This threshold selection is presented in Section 3.2.1.

Secondly, multiple thresholding was implemented by adding three thresholded binary raster layers (CHM, CPI and Slope-H). With this particular combination of metrics, we aimed to improve gap delineation from the most commonly used CHM thresholding approach by the complementary use of other LiDAR information, especially in refining the boundaries using slope information and reducing the impact of small areas of high understorey or regeneration, using the CPI. The slope and CPI thresholds were set after a trial-error visual assessment and two combinations were tested:

- Combination1: CHM-3m + CPI-75% +Slope-H-75%
- Combination2: CHM-5m + CPI-55% +Slope-H-60%

The binary rasters Slope-H-75% and Slope-H-60% participate in identifying gaps areas by detecting the boundaries of gaps as grid cells with a slope above 75% and 60% respectively (slope of the CHM is usually steep along the gap border). The results of the two combinations had grid cell values between 0 and 3 and thus were reclassified into binary rasters (value 0 as canopy and values 1, 2 or 3 as gap).

Per-Pixel Supervised Classification

For the supervised classification process, 100,000 training points were selected (equally distributed in gap or non-gap classes) for leaf-off and leaf-on datasets. These reference points were chosen based on field data and photo-interpretation, and were spread across the four study sites. Corresponding grid cell values were extracted for four rasters metrics (CHM, CPI, Std-H, I-S) for each reference point. We used the package randomForest [80] in R software [124] to construct a Random Forest (rF). The rF algorithm was applied with ntree = 1000 and mtry = 4. Variable importance was calculated to analyse which metrics were the most discriminatory. The out-of-bag error (OOB error) was derived as a robust way to assess the global accuracy of the final classification by the rF algorithm. After this training step, the constructed Rf was applied to the full extent of the four sites.

Random Forests [25] are a very efficient and popular algorithm for classification [52]. The OOB error is used to estimate the prediction error and to assess the importance of each variable (VarImp) which allows sorting of predictors. There are two objectives regarding variable selection by the VarImp measure: to find variables related to the response for interpretation purposes and to find a small number of variables that are sufficient for a good prediction. Calculation of the OOB is based on a bootstrap sample of the data [80, 52].

Per-Object Supervised Classification

A segmentation was conducted in eCognition [10, 13], based on two raster metrics (CHM and CPI) for both leaf-off and leaf-on datasets. The Multiresolution Segmentation Algorithm [11] was used to extract small homogeneous areas, typically segments of tree crowns or gaps. The scale parameter was set to 10 for a shape/color ratio of 0.5 and a compactness/smoothness ratio of 0.5. These objects were exported from eCognition in polygon shapefile format, with the following object attributes: mean of the CHM, mean CPI, mean I-S, ratio of mean object I-S to total scene average I-S (Ratio to Scene Intensity, RtS Int.), calculated as the mean intensity of the object divided by the mean intensity of the entire scene), standard deviation of the CHM, standard deviation of the CPI and standard deviation of the I-S. The four compartment shapefiles were merged in ArcGIS. We selected training objects from among the eCognition segments for the four study sites: 530 gap objects and 530 canopy objects for the leaf-off and leaf-on data sets. We applied the rF algorithm with ntree = 1000 and mtry = 4 and calculated the variable importance to analyse which metrics were most discriminatory.

Morphological Filtering

A post-processing stage was then implemented (with ArcGIS tools) to obtain the best possible match between the ALS mapping approaches and the definition of a gap stated above. Thin corridors and small spaces between trees (width inferior to 2 m) or holes within crowns, were eliminated by using the *Shrink* tool (an erode function) which reduces the gap areas by 1 m around the classified gap area by replacing gap pixels with the most frequently occurring value in the local neighborhood. This step is followed by the application of the *Expand* tool (a dilate function) which enlarges the classified raster gaps by 1 m (to return remaining gap areas to the initial size before shrinking). We reduced the pixelization and eliminated remaining trees within gaps by applying the *Focal Statistics* tool with a circular neighborhood of 1.5 m radius and retaining the most frequently occurring value (majority filter). Following this initial post-processing, the binary rasters were converted into polygon shapefiles and forest roads were removed manually with the *Erase* tool. Finally, gap polygons with an area $<$50 m^2 were removed (*Select*). These filters were applied consistently for all gap mapping methods.

3.3.2 *Analysis of Mapping Quality*

Gap Detection

For the CPI and GPR thresholding, a preliminary step is applied to select an optimum threshold for both leaf-off and leaf-on datasets. For this purpose, Receiver Operating Characteristic (ROC) curves are produced, by means of the ROCR package [131] in R. ROC curves are used to visualise classifier performance, to compare and evaluate prediction models, by changing the threshold value used for the detection to computing, for each case, the True Positive Rate (TPR) and the False Positive Rate (FPR). As a trade-off is present between the TPR and the FPR, ROC curves help identify the optimum threshold value. To apply this technique, we used the same training data as for the training step of Random Forests (3.1.2).

In order to compare the gap detection efficiency of the thresholding and supervised classification methods, a quality assessment was performed using a confusion matrix (or error matrix) to derive global/overall, producer and consumer/user accuracies for the gap class [82]. CA measures errors of commission specific to the user's map. The producer's accuracy (PA) measures errors of omission. As reference data, we used the 295 gap points of the systematic 50 m $\times$ 50 m grid and 295 randomly selected points from among the 845 canopy (non-gap) points.

Influence of Stand Type on Gap Detection

As ALS metrics can be used to describe the canopy cover, a range of metrics were derived to characterize the opening degree of the forest stands and identify "forest types" in our study area. Each stand was described by nine variables derived from five ALS metrics: mean and standard deviation of the HSCOI, mean and standard deviation of the CPI, mean and standard deviation of the CHM, mean of CV-H, the percentage of the stand area with a height of < 3 m (P-3m, based on the CHM) and the percentage of stand area with CPI $\geq 75\%$ (P-CPI75). To simplify comparison, this analysis was conducted with leaf-off data only. Before the analysis, a negative buffer (5 m) was applied to limit the border effect. The hierarchical clustering function *hclust* in the R software [124] was used to group the stands, using the euclidean distance and the Ward method.

Subsequently, the quality of the gap mapping was examined in relation to the forest type, to analyze if it is relevant to choose a gap mapping method based on stand openness characteristics. The confusion matrices of all mapping methods for the leaf-off dataset were analysed and compared for each cluster using PA, CA and the Kappa Index (instead of GA). Kappa Index (KI) is a standard statistical measure of the difference between observed agreement and the chance agreement between reference and classified data [82]. KI is of relevance when the number of reference points for the different classes is not identical, as is the case within the forest types.

Gaps Geometry

The 39 field mapped gaps were used to evaluate the geometric quality of the ALS delineated gaps. ALS gap polygons were selected which intersected the reference gaps. To assess the geometric accuracy, the reference gaps were compared to the ALS delineated gaps using four geometric indicators. The area, the shape complexity and the Index D describe the shape and size of the gap, which influences its functioning. Finally, the main direction of the gap is the global orientation and is considered as a descriptor of the light environment. The following measures were used:

- Gap Shape Complexity Index [72]

$$GSCI = \frac{perimeter}{2 \times \sqrt{area \times \pi}} \tag{3.1}$$

  GSCI is a measure of the relative complexity of shape; it compares the shape of the gap to a circle. It is computed as the ratio of a gap's perimeter to the perimeter of a circular gap of the same area. An increasing value of GSCI is an increasing shape complexity [72, 55].

- Area of the gap (m^2)
- Main direction (degrees): This is the azimuth of the longest line within the boundaries of the polygon without crossing edges. This line is created by the *geom.polygonfetch* command of *Geospatial Modelling Environment*. The azimuth value ranges between 0 and 180 degrees.
- Index D: derived from Oversegmentation and Undersegmentation (Equations (3.2) to (3.4)). This index is used in several studies to assess the accuracy of object-based image segmentation [103, 30], to determine best segmentation parameters. Index D is a distance which varies between 0–1 and is a quantitative assessment of the goodness of polygon matching. Index D has to be minimized and a good balance between oversegmentation and undersegmentation has to be found to optimize the results. In this analysis, polygons for which the ratio of the intersected area between ALS and field gaps was a minimum 10% were retained (compared to 50% in [103] and [30]).

$$Oversegmentation = 1 - \frac{Area_{Intersection}}{Area_{reference}} \tag{3.2}$$

$$Undersegmentation = 1 - \frac{Area_{Intersection}}{Area_{ALS}} \tag{3.3}$$

$$IndexD = \sqrt{\frac{Oversegmentation^2 + Undersegmentation^2}{2}} \tag{3.4}$$

3.4 Results

In Section 4.1, the gap detection results are presented, while Section 4.2 focuses on the influence of the stand type on gap mapping. Finally, Section 4.3 is dedicated to the assessment of the geometric accuracy.

3.4.1 *Gaps Detection Accuracy*

Optimum Threshold Selection for CPI and GPR

To select an optimum threshold for CPI and GPR data, ROC curves were used with a training dataset. Two types of graphics are analysed : changes in the TPR and FPR according to the threshold values (Figure 3.4) and changes in the global accuracy according the threshold value (Figure 3.5). Based on the ROC curves, the following values for CPI and GPR were selected : CPI-60 for the leaf-off dataset, CPI-50 for the leaf-on dataset, GPR-60 for the leaf-off dataset. These thresholds were the only CPI and GPR simple thresholding methods retained in the

subsequent analysis. The optimum CPI and GPR thresholds in leaf-off conditions are the same, but GPR performance is systematically lower. CPI and GPR global accuracy curves (Figure 3.5) are very similar for the leaf-off datasets and for CPI leaf-on for thresholds >50%. Leaf-off accuracies are superior in general to leaf-on, which can be explained by the difference in vegetation density. Decreasing the GPR threshold does not really improve the quality of gap detection. This tendency seems to highlight the importance of the low vegetation influence on the results. Based on the low gap detection accuracy, GPR thresholding is not considered in further analysis for leaf-on data sets.

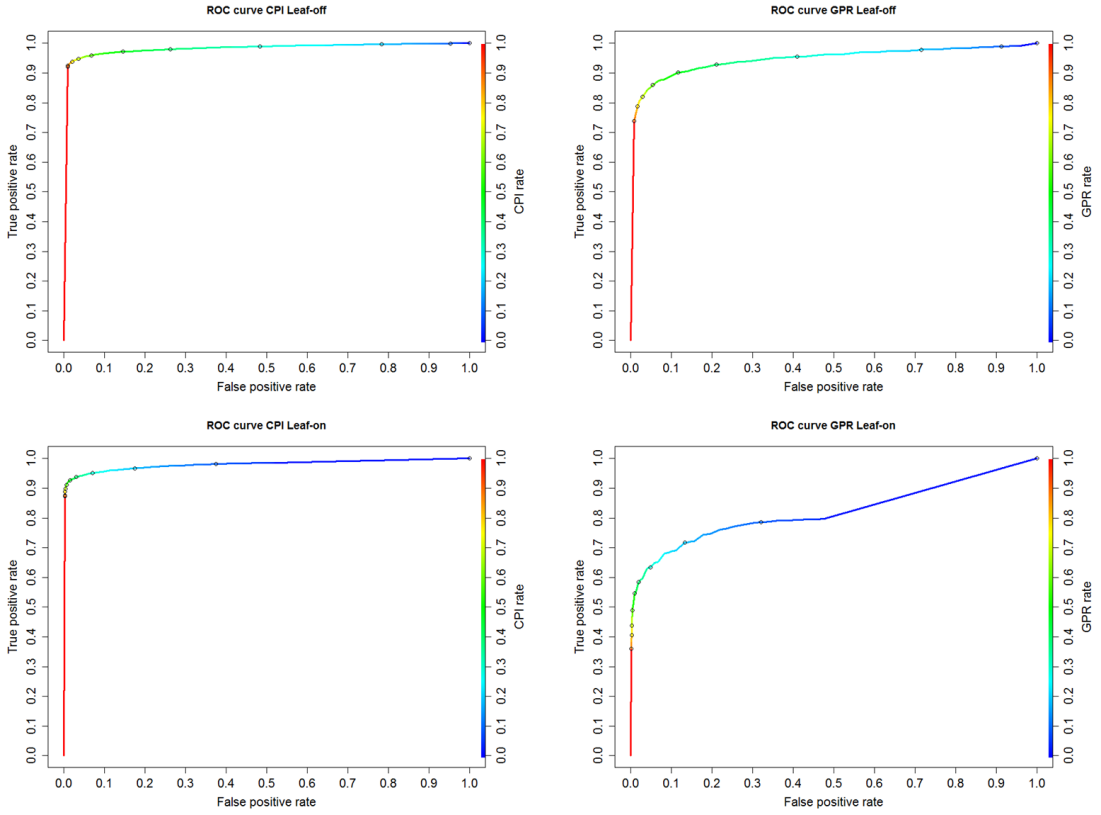


Figure 3.4: Receiver Operating Characteristic (ROC) curves for Canopy Porosity Index (CPI) and Ground Point Ratio (GPR) performance, for thresholds between 0% and 100%. The dots correspond to threshold values between 0 to 100% by 10%-step. Below a certain value, a plateau of True Positive Rate (TPR) is observed with an increasing of False Positive Rate (FPR).

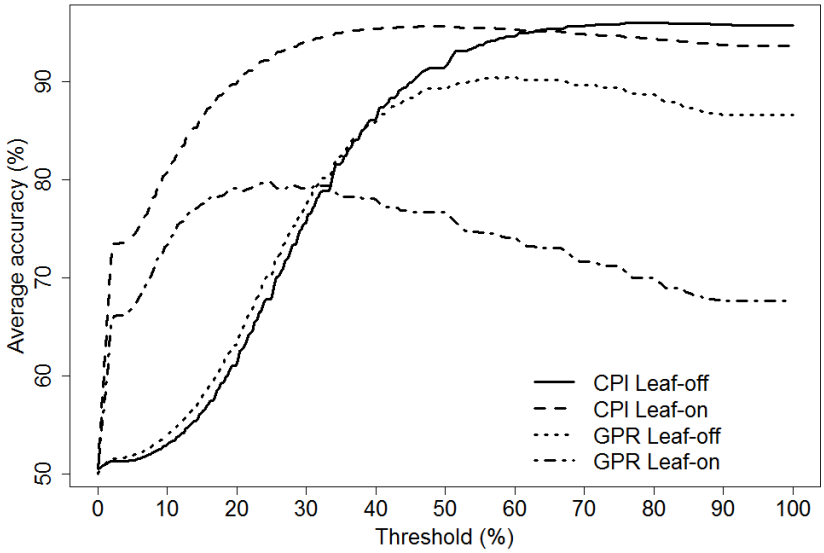


Figure 3.5: Comparison of the global accuracy obtained from CPI and GPR thesholding according to the threshold applied (0% to 100%)

Summary of the Confusion Matrices

The results of the gap detection accuracy assessment for the nine selected methods, for the two seasons, are summarized in Table 3.1. These results allow a number of comparisons to be made: leaf-on *vs.* leaf-off data, thresholding *vs.* supervised classification, per-object *vs.* per-pixel, CHM *vs.* other ALS metrics. Based on global accuracy, the three best methods are: per-pixel classification in leaf-on and leaf-off conditions, CPI-60 for leaf-off data and combination1 multiple thresholding in leaf-on conditions. CHM-5m thresholding also produces high global accuracy in leaf-off conditions, but has a significantly lower producer's accuracy, indicating larger omission errors. The specific results of supervised classifications are presented in Table 3.2 and Figure 3.6 and discussed further in Sections 3.4.1 and 3.4.1.

Leaf-on *vs.* Leaf-off

Global accuracies are, in general, greater for leaf-off conditions. Leaf-off producer's accuracies lie between 52% and 94% against 19% and 86% for leaf-on conditions. Leaf-off producer's accuracies are around 10% greater than leaf-on for all methods (Table 3.1). In consequence, omission errors are lower in winter conditions. The difference between leaf-off and leaf-on data is lower for the consumer's accuracy. The commission errors are small (7 to 22%, excluding Combination2 multiple thresholding) meaning that areas identified as gaps are reliable, but are generally underestimated from leaf-on data.

For the per-pixel and per-object supervised classifications (Table 3.2), the accuracy results (OOB, global, producer and consumer accuracies) are similar for leaf-off and leaf-on data. The VarImp parameter (Figure 3.6) highlights Mean-I-S and CPI as important for the leaf-on dataset when using per-pixel classification, whereas the CHM is of relatively greater importance for leaf-off data. This contrast is lower for the leaf-off dataset. No such differences in the importance of intensity are observed for per-object classification between the two seasons.

Table 3.1: Accuracy assessment of gap detection (summarized from the confusion matrix) for the different methods in leaf-off and leaf-on conditions.

	Leaf-off			Leaf-on		
Type	**Global acc. (%)**	**Producer acc. (%)**	**Consumer acc. (%)**	**Global acc. (%)**	**Producer acc. (%)**	**Consumer acc. (%)**
Simple thresh.						
CHM-3m	78	60	93	72	50	90
CHM-5m	81	66	93	77	60	91
CPI-50				78	63	91
CPI-60	82	73	88			
GPR-60	76	52	89			
Multiple thresh.						
Combination1	79	81	78	80	71	87
Combination2	62	94	58	73	86	68
Supervised classif.						
Per-pixel	81	73	88	81	73	90
Per-object	79	77	81	80	72	86

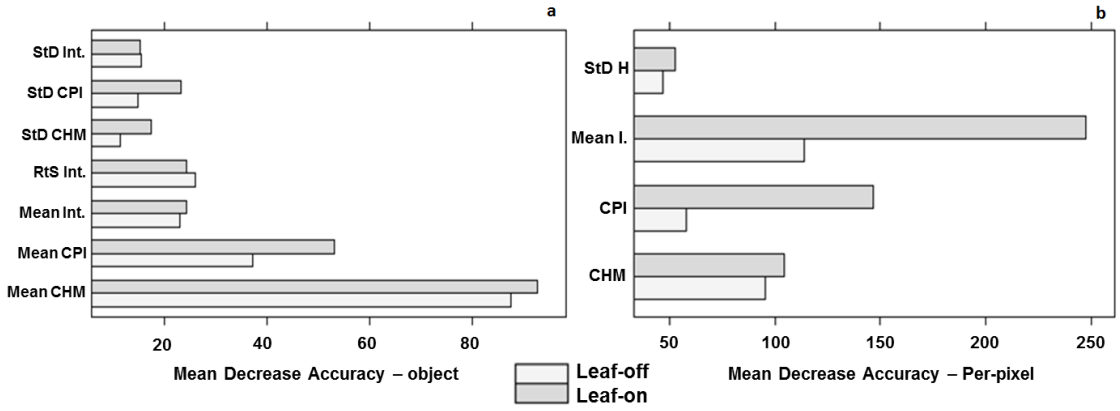


Figure 3.6: The mean decrease accuracy is a way to measure the variable importance for a) per-object and b) per-pixel classifications. This indicator allows the sorting of the variables according to their discriminatory abilities.

Thresholding *vs.* Supervised Classification

Gap detection by CHM thresholding shows an average producer's accuracy but consumer's accuracy is very high, more so than for per-pixel or per-object classification. On average, the discrepancy between consumer's and producer's accuracies is greater for thresholding-based methods, but global accuracy remains relatively similar and good for all methods, except for GPR leaf-on thresholding. Based on global accuracy, CPI-60 is the best performing method, but high and similar accuracy is obtained from per-pixel classification.

Per-Object *vs.* Per-Pixel

The OOB accuracies for per-pixel and per-object classifications are given in Table 3.2; the values are similar and very high. The importance of variables involved in the supervised classification are shown in Figure 3.6. Owing to the definition of objects (group of several grid cells), more variables can be utilised in the per-object classification process, as the mean and the standard deviation of the grid cell values can be calculated. The most important variable for leaf-off and leaf-on per-object classification is the mean of the CHM, followed by the mean CPI. The variable importance analysis for per-pixel classification identifies a different order of the variable importance. The Mean-I-S is the most discriminating variable for the per-pixel classification in both leaf-off and leaf-on data sets.

Table 3.2: The out-of-bag accuracy is an evaluation of the global accuracy of the final classification. The results are given for the per-object and the per-pixel supervised classifications for both leaf-off and leaf-on conditions.

Per-Object		Per-Pixel	
Leaf-off	Leaf-on	Leaf-off	Leaf-on
98.1%	96.3%	97.0%	97.2%

Table 3.3: Amount of improvement over Canopy Height Model (CHM)-5m thresholding for the three best methods selected based on gap detection accuracy. The leaf-off results are compared with leaf-off CHM thresholding and leaf-on with leaf-on CHM thresholding.

	Per-Pixel Leaf-on	CPI-60 Leaf-off	Combination1 Leaf-on
Global accuracy	4%	1%	3%
Producer accuracy	13%	7%	11%
Consumer accuracy	-1%	-5%	-4%

CHM Thresholding *vs.* other Metrics

CHM thresholding did not produce the best accuracies in this comparison, particularly in terms of the omission error (CHM-3m had a low producer's accuracy). Use of a higher threshold for the CHM (5 m) improves the gap detection, with an increase of 6 to 10% for producer's accuracy. The GPR has poor performance compared to CPI and CHM metrics due to a high omission rate. CPI-60 in leaf-off conditions provides the highest global accuracy and the best balance between omission and commission errors.

Table 3.3 presents the improvement in gap detection of the three best performing alternative methods over CHM thresholding (per-pixel classification of leaf-on data, CPI-60 thresholding of leaf-off data and Combination1 multiple thresholding of leaf-on data). The leaf-off results are compared with leaf-off and leaf-on with leaf-on for all methods. The consumer's accuracies vary less than producer's accuracies between the methods and are best for the CHM thresholding. Alternatives to CHM thresholding provide higher global accuracy, but these are relatively small improvements (1%–4% increase in global accuracy). The most important improvement is the decrease in omission error.

3.4.2 *Influence of Stand Type on Gap Detection*

The hierarchical clustering of the 42 stands produced a dendrogram from which three groups were extracted. Analysis of these clusters shows a separation according to the opening degree of the canopy. The following three stand types are identified: opened canopy with isolated trees only (cluster 1), closed canopy (cluster 2) and heterogeneous canopy (cluster 3). Two main variables amongst the nine are particularly useful to differentiate the three clusters: the P-3m and P-CPI75 (Figure 4.6). The two variables are highly correlated and thus only one is sufficient to discriminate the three groups. The choice may depend on the forest type, the recruitment characteristics or the ALS data. The P-3m is the easiest indicator to use as only a CHM is needed.

Figure 3.8 shows the Kappa index, consumer and producer accuracies for each method for each of the three stand types. The Kappa index shows the greatest differences between the methods. The discrepancies in accuracy between the methods varies according to the stand type: commission errors are more stable between methods, but are higher for open canopy (consumer accuracies were greater than 80% in general for heterogeneous and closed canopies), whilst omission errors are more variable between methods for closed canopy. Heterogeneous canopies have the highest Kappa index for most methods. Regarding Kappa index, the best methods for each stand type are : combination1 multiple thresholding for closed and open canopies and CPI-60 thresholding for heterogeneous canopies.

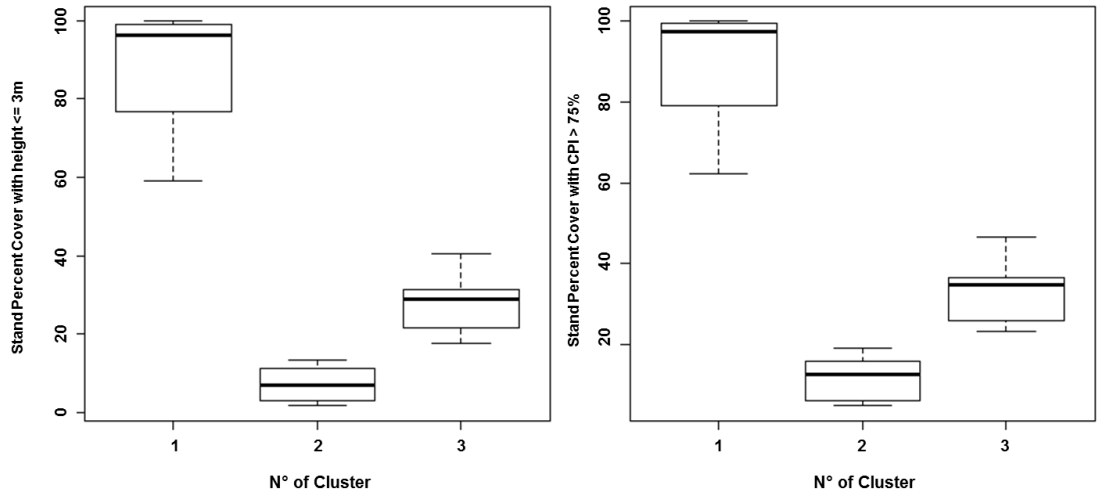


Figure 3.7: Three clusters are identified from the hierarchical clustering, corresponding to three stand types: open canopy with isolated trees only (cluster 1), closed canopy (cluster 2) and heterogeneous canopy (cluster 3). These types are well defined by two indicators: the percent of stand area with height < 3 m and the percent stand area with CPI $\geq 75\%$.

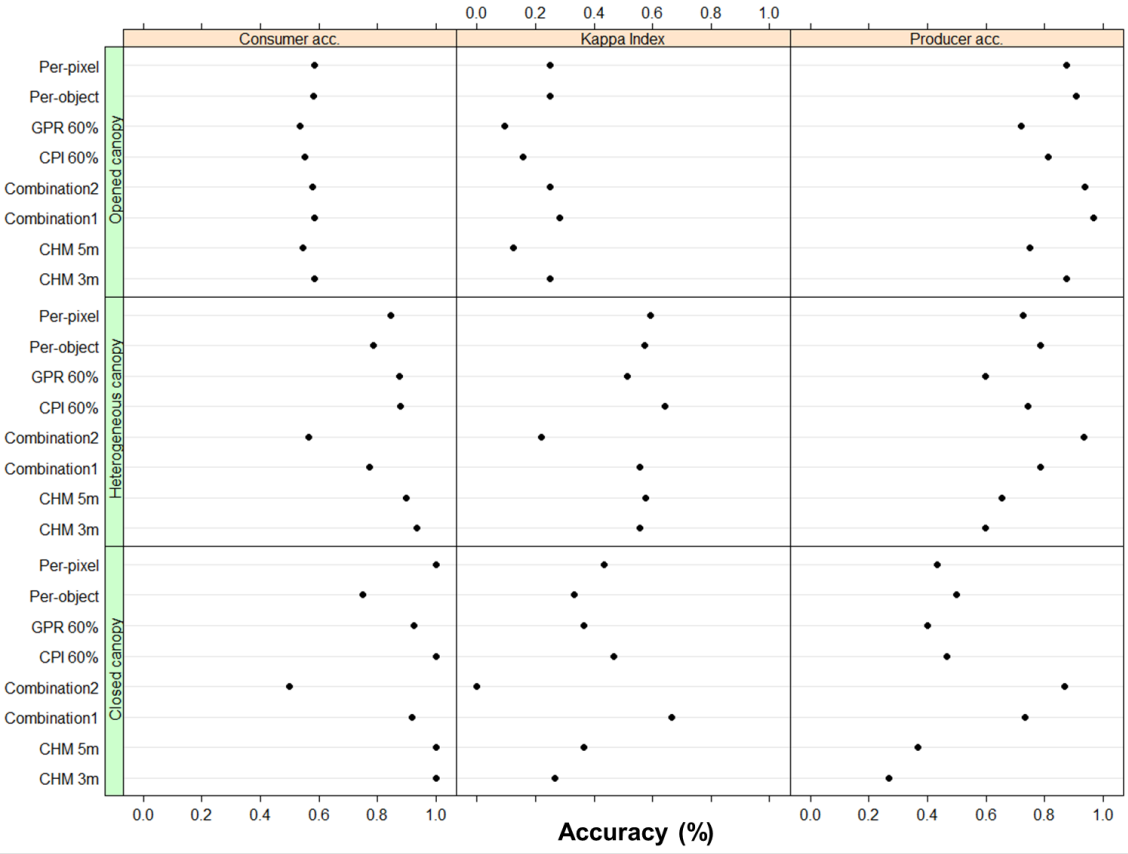


Figure 3.8: Kappa Index, producer's and consumer's accuracies of the three stand types for each gap detection method using the leaf-off dataset.

The analysis of the Kappa index (in this case, for leaf-off data only) shows that the best global methods may be different from the best methods for specific stand types. The three best leaf-off methods when considering all stands (Table 3.1) are per-pixel classification, CPI-60 CHM-5m thresholding. Amongst these three methods, per-pixel classification performs most consistently across all stand types, but CPI thresholding is best in heterogeneous and closed stands. CHM thresholding (5 m threshold) gives lower accuracies, with notably higher omission errors, especially in closed canopy. In such closed stand types, all these methods are significantly outperformed by Combination1 multiple thresholding, which performed less well for leaf-off data in the global case.

3.4.3 *Gaps Geometry Accuracy*

This section will provide a synthesis of the accuracy of the different methods in retrieving characteristics of gap geometry. Figure 3.9 illustrates four typical gaps with example delineations from the three methods giving the best detection results. A particularly poor result is observed in Figure 3.9 G and K, where the ALS delineation resulted in large and very branched gaps, which encompassed several reference gaps (becoming connected by corridors). Such results appear in stands where openings in the canopy are small or medium and spread throughout the stand, forming a network of connected openings. The gaps delineated in this case are more complex in shape than isolated ones and may be less well fitted to the definition of a gap used in this study and to accurate field delineation. To overcome this issue, which may bias comparison of geometric accuracy, highly branched and interconnected gaps were therefore eliminated from the subsequent analysis of geometric accuracy by applying a selection based on the GSCI (GSCI$<$ 5). In the following sections, ALS gaps are compared to reference gaps based on the error, computed as the difference between the reference value and the ALS value; except for the case of Index D, which is intrinsically a comparison with reference data (Figure 3.10).

GSCI

Regarding GSCI, the weakest results are obtained for multiple thresholding : Combination2 for leaf-off data produces no gaps with GSCI $<$5 and Combination1 has the largest mean error and high extreme values (for both leaf-off and leaf-on). The two CHM thresholding methods are quite similar with a tendency to overestimate shape-complexity in comparison to field data, as for all the methods. The best performance is from per-pixel classification of leaf-off data.

Gap Area

CPI thresholding and per-pixel classification perform well for both seasons. Multiple and GPR thresholding show the worst results, especially for the leaf-on dataset. The average errors are acceptable but, despite the GSCI-based filtering, there are still extreme values, mainly representing large overestimates of gap area (except for CHM thresholding in leaf-on conditions).

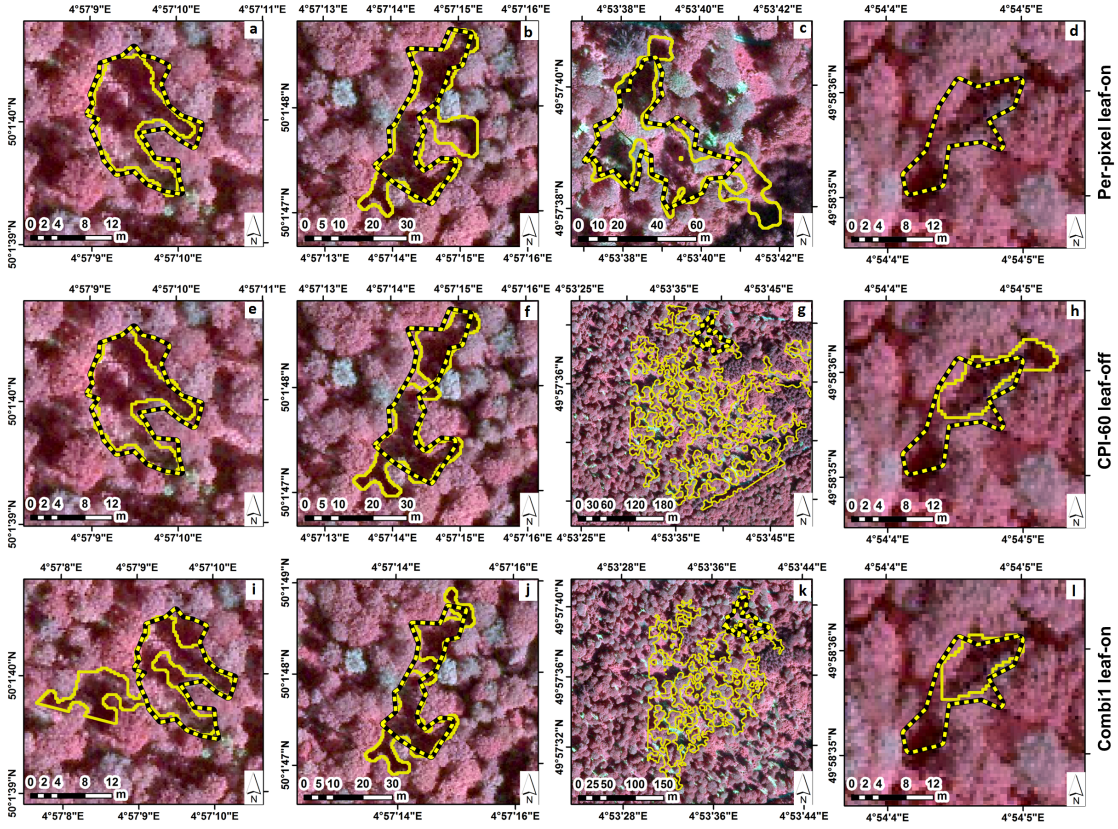


Figure 3.9: Four examples of canopy gap mapping results (yellow polygons) are compared with the the reference gaps (black-yellow dashed polygons) overlain on a near-infrared composite (50 cm, year 2009): (A) to (D) are per-pixel classification of leaf-on data; (E) to (H) are CPI-60 thresholding of leaf-off data and (I) to (L) are Combination1 multiple thresholding of leaf-on data. These four cases are typical of results obtained in this study: a good correspondence between the field and ALS gap ((A) and (E)); a good coverage of the reference gap but in several parts ((B), (F), (J)), a weaker result resulting in a large and very branched airborne laser scanning (ALS) gap ((C), (I) and especially (G) and (K)) corresponding to several separate reference gaps, which is typical of sparse or heterogeneous canopy cover; and small gaps missed or partially missed in more closed canopy ((D), (H), (L)).

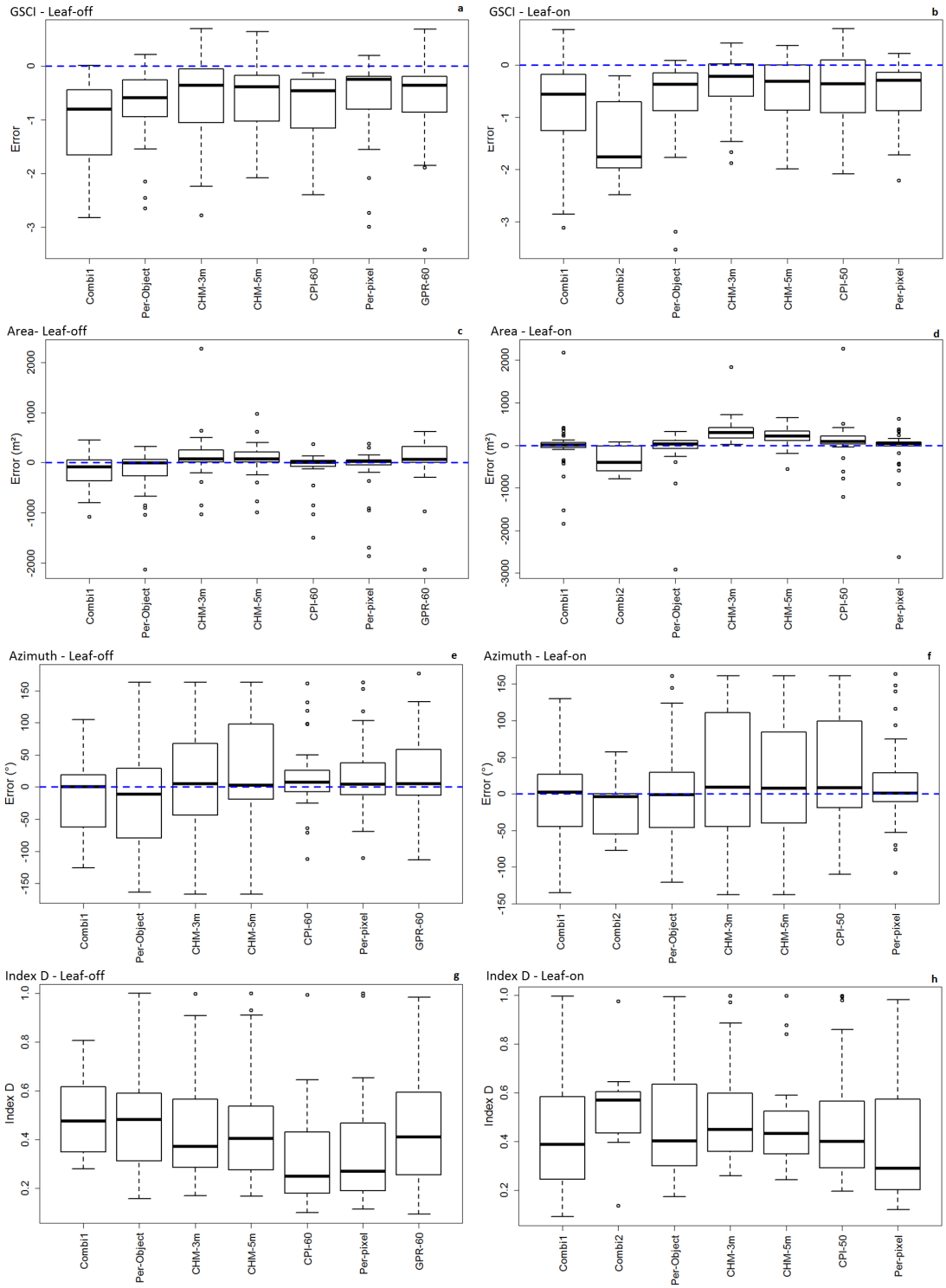


Figure 3.10: The geometric accuracy of the selected mapping methods is assessed by four indicators : gap-shape complexity index (a,b), area (c,d), main direction (e,f) and Index D (g,h). The error is computed as the difference between the reference value and the ALS-value. The results are summarized in boxplots for the two seasons (left for leaf-off and right for leaf-on).

Main Direction

For both leaf-off and leaf-on datasets, and all methods, the average error in azimuth is close to 0, but there is large dispersion or errors with extreme error values (up to 150%). The best results are obtained from CPI-60 thresholding of leaf-off data and per-pixel classification of leaf-on data. The standard deviation of errors for the azimuth is lower for the leaf-off dataset, suggesting more consistent prediction of gap orientation.

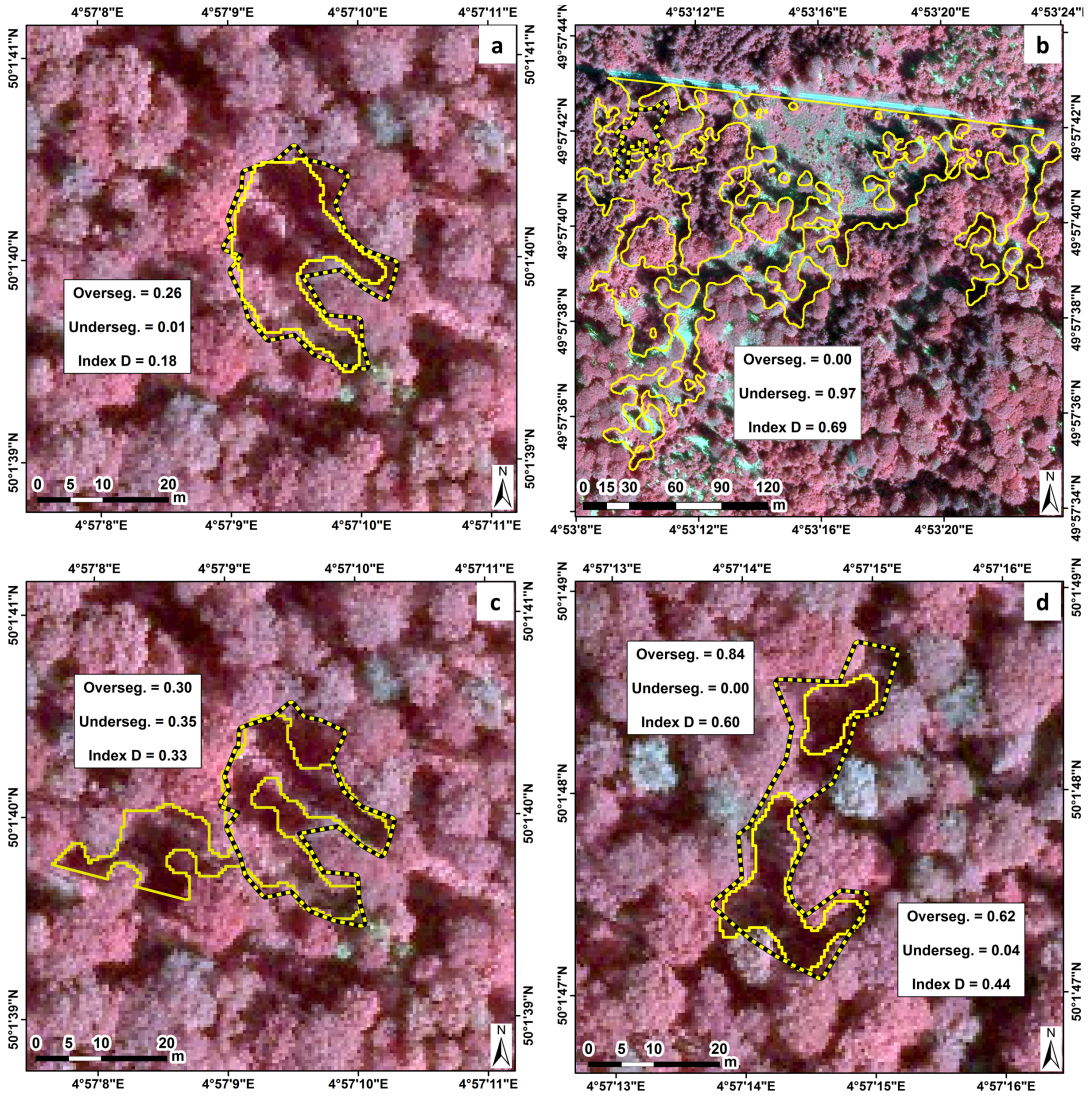


Figure 3.11: Examples of oversegmentation and undersegmentation values for a sample of gaps. (a) large value for undersegmentation corresponds to an ALS gap overestimating the reference gap (e.g., very large and branched ALS gap, b). This situation is often associated with a small oversegmentation value because the reference area is often entirely covered by the ALS-derived gap. Small values of undersegmentation occur when a reference gap is covered by several small ALS gaps (c) or a single ALS gap but with a good coverage and a satisfactory global shape (a). (d) is an intermediate case.

Index D

A good match between reference and ALS delineated polygons corresponds to a low Index D and indicates a good balance between oversegmentation and undersegmentation (Figure 3.11). Index D is, in general, lower for leaf-off datasets, and CPI-60 thresholding from leaf-off data has the smallest Index D value with a good oversegmentation/undersegmentation ratio. In leaf-on conditions, per-pixel classification is the most effective method. Undersegmentation is higher in leaf-off conditions, which confirms the tendency to delineate excessively large gaps from leaf-off data. The multiple thresholding methods generally produced the poorest results for this index, although Combination1 performed better in leaf-on conditions. The CHM-3m and 5m thresholding show similar results for leaf-off and leaf-on datasets.

3.5 Discussion

Three methods are highlighted as producing the best gap detection results: per-pixel supervised classification of leaf-on data, Combination1 multiple thresholding of leaf-on data and CPI-60% simple thresholding of leaf-off data. The global accuracies achieved using these best methods (81%–82%) compare favourably with those of past studies (e.g., 73.1% of field and ALS gap length matched along transects in [149] or 77%–88% global accuracy in [50] using a computationally expensive point cloud based approach). The choice between methods implies trade-offs dependent on the study requirements, *i.e.*, more reliable gap detection but with a risk of missing some gaps or more detected gaps with a risk of false detection. For example, multiple thresholding methods provide the highest producer's accuracy (greater than 80 % for leaf-off data, 90 % for leaf-on), but at the cost of higher commission errors. These methods also have variable performances regarding geometric accuracy. Independently of the gap detection accuracy, CPI-60 thresholding of leaf-off data and per-pixel classification of leaf-on data give acceptable results for geometric accuracy by most measures, unlike Combination1 multiple thresholding of leaf-on data which produces a poor final map in terms of GSCI and Index D.

Optimum thresholds for CPI and GPR were identified through ROC curves, but it is acknowledged that the thresholds used are derived empirically. Further research to examine their physical basis would be valuable. However, the identified thresholds appear reasonable, as within a correctly identified gap, substantial numbers of returns > 3 m in height (e.g., a CPI of up to 40%) may result from very sparse but high understorey or regeneration, small within-gap tree crowns or mixed canopy and gap cover within pixels along the gap boundary.

In general, better results were obtained from most methods in leaf-off conditions (with the exception of multiple thresholding and per-object classification) regarding gap detection and geometry. The high point densities of the ALS datasets used in this study allow good characterisation of tree crowns, even in leaf-off conditions, and the absence of leaves decreases occlusions and improves representation of the ground. In deciduous forest, leaf-off data may therefore be generally preferable to leaf-on data unless point density is low, but the transferability of this finding to other forest types and datasets needs further examination. GPR thresholding performed worse for leaf-on data than CPI thresholding, whilst CPI and GPR leaf-off results are more similar. These findings likely indicate an effect of the low understorey vegetation presence in summer on gap detection (few returns penetrating to the ground in leaf-on conditions even within gaps) with GPR-based methods. The effect of the season is also highlighted in the per-pixel classification. The variable importance shows a relatively greater importance of CHM for leaf-off data versus Mean-I-S and CPI for leaf-on data. This is probably due to the higher contrast in reflectance between elements in leaf-on conditions (e.g., ground *vs.* leaves).

Despite the relative technical simplicity of the method, thresholding is not clearly surpassed by supervised classification. Given the complexity of implementation and the similar performance compared to other simpler methods, classification methods may not always be the most efficient choice. Image segmentation used eCognition and was intentionally simple but the use of such software requires a learning process by the user and remains a black box approach. The simpler per-pixel approach also outperforms the per-object classifier based on global accuracies from the confusion matrix. The preparation of the training samples is time-consuming for both supervised classification methods.

The results of this study also suggest that CHM thresholding is outperformed by other methods (see Table 3.3). However, choice of threshsold is also important if using a CHM. An increase in the height threshold (3 m to 5 m) improves the detection. This is probably due to better representation of the gap border, with closer correspondence to the canopy drip line, but the results of the 3m threshold could also be influenced by high understorey vegetation. The CHM-based method also performs less consistently between stand types (high omission errors in closed canopies), although this may be improved by the use of relative rather than fixed height thresholds, as proposed by, for example, [169] and [50]. Methods based on CPI thresholding from leaf-off data and per-pixel supervised classification based on ALS-derived raster layers are, according to our results, the most accurate and reliable, (*i.e.*, using canopy porosity index, rather than a CHM, decreases the under-estimation of canopy openings). However, it should be noted that CHMs have other very useful applications in forestry and are valuable ALS products.

The creation of a CHM has other advantages for a forest manager, such as allowing height estimation and a CHM is therefore more multi-purpose. The CHM is also likely to be less sensitive to ALS point density than the CPI. Moreover, CHMs can also be produced from a photogrammetric DSM by subtracting an existing DTM. Photogrammetry can be, in some cases, more cost-effective than ALS. The relatively small accuracy gains over the use of CHM thresholding may therefore not always justify the use of other metrics, such as CPI, for gap mapping.

The results of the geometric accuracy assessment indicate poor agreement with field data in many cases, especially in terms of gap shape. In this case, a complementary explanation could lie not only in the gap mapping method, but also in the field survey, during which challenges arise when mapping complex gaps or interconnected gap networks. It is likely that neither the field mapping or the ALS-derived gaps fully represent the true gap geometry. Whilst gaps have been field surveyed with a good degree of accuracy, subtle details of gap boundaries may be better represented by ALS, resulting in more complex shapes. To some extent this is to be expected, due to the high resolution of the ALS data relative to field mapping and some differences between reference and ALS gaps may therefore reflect this issue. Even though many errors were larger than would be accounted for by this factor, future studies are needed to investigate the possibility of using more advanced approaches, such as terrestrial laser scanning, to improve validation. The increasing use of unmanned aerial vehicles also provides a potential alternative method for mapping gaps [54]. The relatively easy acquisition of very high-resolution images and the possibility for programming repeated acquisitions [87] should foster the development of gap dynamic studies.

Particular challenges were encountered in obtaining comparable ALS and field delineations, in terms of gap geometry, for the heterogeneous stand type. However, a more detailed analysis reveals that two subgroups composed this stand type, with the same global gap coverage but differing spatial distribution: one with larger open areas, more isolated and well defined and a second where openings are distributed more diffusely within the stands. Extensive networks of interconnected gaps are a source of error, due to challenges in both gap definition and the field data collection. The chosen definition of gaps in this study fits better with the concept of isolated gaps. When considering issues such as light availability for regeneration, alternative conceptual approaches to examining gap geometry, size distribution and spatial arrangement in heterogeneous or open stands are likely to be required, moving away from the concept of isolated gaps (which is appropriate in closed canopies) to an 'open matrix model' [59, 15]. ALS mapping of the distribution of gap space has potential to facilitate this over larger spatial extents.

3.6 Conclusions

Gap mapping is a topic of major importance for forest managers and ecologists. A reliable mapping of canopy gaps is the basis for numerous research and management applications (for example gap dynamics, gap distribution at the scale of the landscape, habitat modelling and silvicultural applications). This paper has investigated canopy gap mapping through four types of methods (simple thresholding, multiple thresholding, per-pixel supervised classification and per-object supervised classification) applied to leaf-off and leaf-on datasets (with an average density of 40 returns per m^2), taking into account the mapping accuracy in terms of both gap detection accuracy (590 validation points) and geometric quality (analyzed for 39 reference gaps mapped in the field).

Our findings demonstrate that mapping of canopy gaps with ALS is not trivial, especially regarding errors in gap areas (maximum error up to 2000 m^2) and geometry (e.g., mean Index D = 0.43 for Combination1). This study indicates that ALS can be used to map canopy gaps in uneven-aged broadleaved forests with gap detection accuracies of up to 82%. New methods, based on thresholding of CPI and per-pixel supervised classification of ALS-derived raster layers, show particular promise for obtaining accurate gap detection (82% for CPI-60 leaf-off and 81% for Per-pixel leaf-on) and geometric representation (GSCI mean error 16° and 13° for CPI-60 leaf-off and Per-pixel leaf-on respectively and main direction mean error –0.7 and –0.55). But taken together, the results suggest the importance of and difficulty in identifying an optimum mapping method for all circumstances. The choice of method is generally likely to be a trade-off between producer's and consumer's accuracy or between detection and geometric accuracy, e.g., GPR-60 Leaf-on global and consumer's accuracy equal to 76% and 89% but producer's accuracy equal to 52%. Beyond the performances of the specific methods presented here, we argue the critical need for future ALS-based gap studies to consider the geometric accuracy of their results, even if ALS remains the best option for gap mapping over large extents.

Further challenges arise regarding the selection of an appropriate definition of canopy gaps, compatible with remote sensing, and in achieving a meaningful comparison of ALS and field delineations. Despite the challenges, understanding the differences in gap geometries and gap size distributions obtained from traditional field survey methods and from ALS delineation is vital, if conclusions regarding gap dynamics and spatial patterns from studies utilising these different approaches (in different locations or at different time periods) are to be compared and understood.

The most appropriate approach to mapping of canopy gaps will depend on the forest type and the application. This implies the adaptation of methods, for example in the choice of threshold, of morphological filters and in the methodology for the field survey. Moreover, as the effect of the understory vegetation has important implications, future studies should examine the assessment and characterisation of low vegetation and regeneration with ALS data. The effects of ALS acquisition parameters such as scanning angle, point density and derived raster resolution were not assessed in this paper, but also need to be further explored [148, 50].

Chapter 4

FOREST ATTRIBUTES ASSESSMENT OF CONIFEROUS STANDS WITH UAV DATA

Chapter 4

Forest Attributes Assessment of Coniferous Stands with UAV Data

"A fool sees not the same tree that a wise man sees."

William Blake

The use of UAS has opened a new era for remote sensing and forest management by allowing the acquisition of images with a centimetric resolution. Except legal constraints, the flight are done on-demand by the user. The resolution and the data acquisition flexibility make it a tool suitable for issues that require information up to tree scale. However, the technical constraints (mainly battery life and flight control) involve flights over small areas (about 1000 hectares).

The management and processing of the data require expertise and important technical resources that make difficult to extent over large areas. As Single-Tree Detection is an important issue in sustainable precision forestry, this study develops an application of UAV data processing for forestry. We propose a comprehensive workflow to detect trees and assess forest attributes using UAV imagery, from data acquisition to model construction.

We implement a Local Maxima Detection (LMD) to identify the tree tops, based on a fixed-radius mobile window in a Canopy Height Model (CHM) and images produced from UAV surveys. To compare the contribution of different photogrammetric products, we analysed the local maxima detected from the CHM, based on three image types and a combination of the LM from the CHM and each image type. A filtering process of false positives was also implemented, using a supervised classification. Based on the LMD combined with an area-based approach, we constructed models to assess top height, number of stems, basal area, volume, and individual tree height.

Specifically, we highlight four key points which form the originality of our work: (i) the synergy between the CHM, the correlation map and the imagery to improve the local maxima detection; (ii) The implementation of a process to filter the false positives during the local maxima detection; (iii) the proposal of a complete methodology from the data acquisition to forest attributes assessment and (iv) the use of a deliberately limited number of predictors to rationalize the variables selection for the forest attributes assessment.

4.1 Introduction

With the advent of small Unmanned Aerial Systems (UAS), robotic and geomatic technologies have established a new paradigm of aerial remote sensing [31]. Because of their high operational flexibility, UAS can deliver very fine spatial resolution data at specific moments defined by the end users [5]. Flying low and slow, a small UAS with an on-board optical sensor can acquire images of natural areas (e.g. forest canopies) and provide spectral information. The concomitant improvements of image processing techniques make todays' novel "Structure-from-Motion" photogrammetry workflow operational with UAS imagery [31].

While traditional photogrammetry requires a metric camera, a specific image block configuration, and a substantial amount of work by the operator, the Structure-from-Motion photogrammetry approach deals automatically with a collection of unordered overlapping images from an uncalibrated camera [155]. When used with UAS imagery, this modern photogrammetry technique delivers without difficulty two major final products of the photographed area for the user : a three-dimensional (3D) model of the relief and an orthophotomosaic.

A recent view of the literature shows a diversity of applications of drones within the natural resource management framework. [55] and [54] extract canopy gaps from UAS images (1) to assess the floristic biodiversity of the forest understorey from spatial characteristics of gaps and (2) to analyse their spatial patterns related to the spatio-temporal dynamics in several managed and unmanaged forests. The article of [150] is a first attempt of elephants census by UAS survey while [83] browse deeper the strengths and drawbacks of UAS for wildlife monitoring. [114] use UAS data to produce very high resolution digital elevation model in the domain of soil conservation for which the flexibility of data acquisition constitutes a leverage after major rainfalls to model relief changes. [2] compare UAS-based cameras for purposes of land-cover classification, by object-based image analysis. [86] and [101] explore time series of UAS imagery to discriminate tree species and health conditions in deciduous and riparian forests using supervised classification methods. Following the same idea, [99] develop a workflow to map three riparian invasive species based on spectral and textural variables computed from UAS images.

The development of drones as a source of very high resolution data is an opportunity for forest resource management and give more emphasis to precision forestry, focusing on plot-level and on individual trees [41]. Sustainable forest management requires an accurate and regular quantification of resources. Height distribution, stem number, top height (or dominant height), basal area, and stem volume are key data for forest managers. Sampling inventories are generally undertaken to estimate these variables through a limited number of plots because their installation in the field is very expensive and time-consuming. Inventory at the tree level is highly relevant for forest structure characterization, whether for silvicultural purposes or for the modeling of forest ecosystems (e.g. mapping of habitat quality, vegetation biomass, or regeneration spots).

Very high resolution images and 3D model of the relief are well suited to the forest manager's needs for high-quality mapping [47]. Indeed, the spatial resolution of low oblique vantage aerial images allows for a single-tree inventory of the forest stands, particularly interesting for silvicultural and harvesting operations. As canopy height measurements are the basis for extracting trees and stands attributes, UAS data could be used, at a local scale, to substitute or update airborne laser scanning (ALS) and its proved ability for describing the 3D structure of forest [165, 60, 133, 93]. However, this opportunity remains strongly linked to the availability of an existing DTM to compensate for the lack of data at ground level and especially under the canopy [157].

Several studies confirm the potential of UAV to monitor structural properties of forest as new inventory tool. [87] developed a photogrammetric workflow with Structure-from-Motion approach on UAS images combined with LiDAR digital ter-

rain model to assess top height in deciduous plots and compared results with ALS data. [123] highlighted the determination of the applicability of drone technology as accurate and timely forest inventory information at local scale. They highlight the following major strengths of UAS : adaptive planning, high project customization and rapid implementation of data acquisition even under challenging weather conditions, and the high spatial and temporal resolution of data. In a conifer-dominated boreal forest, they assess Lorey's mean and dominant height, stem number, basal area, and volume. The relevance of spectral information in forest attributes modelling is evaluated but results appear not conclusive. Similarly, [140] extract 3D, spectral, and textural features from orthoimages and canopy height data to assess dendrometric attributes of plots following an important step of variable selection with genetic algorithm method for two consecutive years and two different platforms and cameras. Two spatial resolutions were tested (20 and 50 cm) but the effect on the forest variables estimations was small. They note an important variation of quality between the two years, explained by changes in weather and/or solar illumination conditions, identified as a stumbling issue.

Although the great potential of UAS for ecology is arising in the literature, environmental remote sensing with UAS is still relatively new. There is still limited understanding of how the data are acquired, processed, and used [160]. The rise of UAS expands the possibilities and flexibility of remote-sensing data acquisition in forestry, involving not only new prospects for quantification of the resource but also new challenges regarding the data variability. Some questions remain regarding the adequate processing of drone imagery for utilization in forestry. Some standard of traditional photogrammetry still needs to be revisited in order to better fit both characteristics of low-altitude and low-quality imagery and the specificity of the forest structure. In this context, the objective of this study is to implement a comprehensive workflow to detect tree tops and assess forest attributes using UAS imagery and photogrammetry, from data acquisition to model construction. Finally, we outline several issues requiring further improvement and highlight several recommendations.

4.2 Material

4.2.1 *Study Site*

The study area, named Hattlich, is located in the high plateau of the eastern Hertogenwald near the city of Eupen (in eastern Belgium, close to the German border, Figure 4.1). It occupies an area of about 3000 ha with altitudes ranging between 500 and 650 m. This estate is a large forest area where coniferous stands are the major forest type (66%). Spruce stands (*Picea abies* (L.) H. Karst.)

are predominant (64.7%), and most of them are more than 50 years old. The silviculture of spruce in this area is widely based on high-density plantations (2000–3000 ha^{-1}) establishing even-aged stands [109] with successive thinning steps from 25 years. These stands are now involved in a process of transformation to an uneven-aged structure. The objective of such uneven-aged transformation is to promote a silviculture close-to-nature, more favourable to multifunctionality and shift to the concept of continuous cover forestry. The area covered during UAS surveys represents approximately 800 ha of the estate.

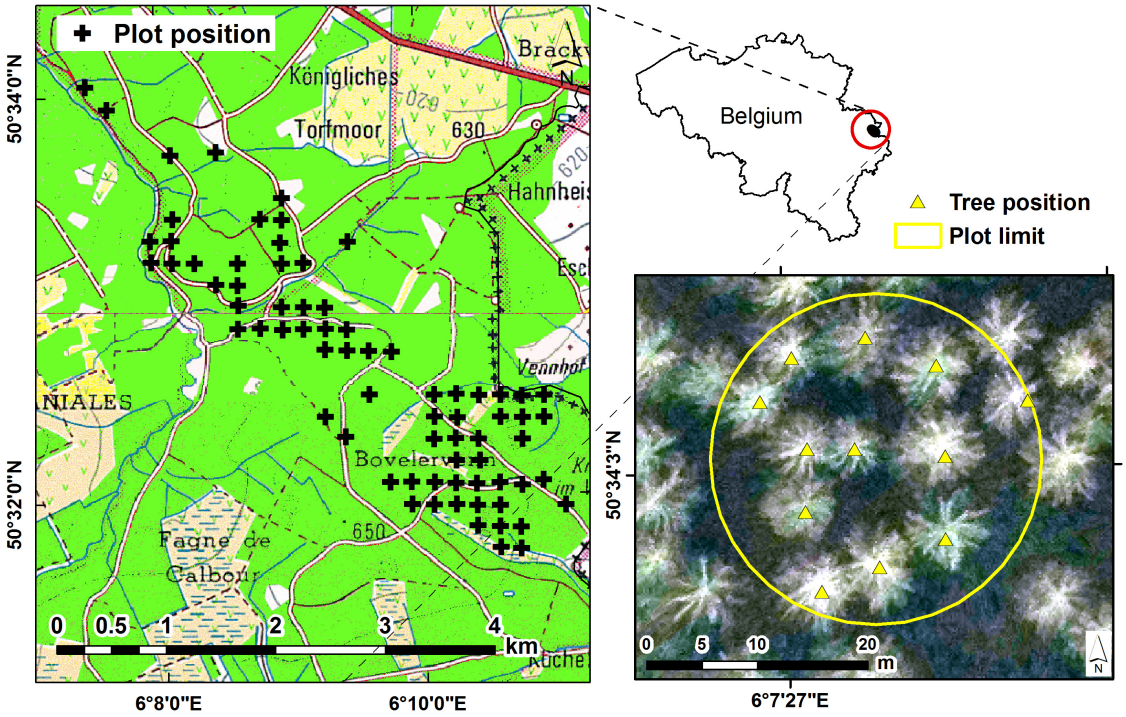


Figure 4.1: Study site localization and illustration of one plot. The plot is spread along a systematic 200 × 200 m grid. The study site is a large forest area where coniferous stands are the major forest type with predominant spruce stands

4.2.2 *Field Data Inventory*

A forest inventory was carried out in order to compare trees detected with UAS data with a field reference and to assess forest attributes (data are summarized in Table 4.1). The forest inventory is based on variable-area circular plots (n = 72) installed on a systematic 200 m × 200 m grid. The plot radius was adapted to ensure a minimum number of 15 trees per plot. The maximum plot radius was fixed at 18 m (0.1 ha) in the case of low-density stands. The plot centre was positioned by GPS (SX Blue II with a nominal horizontal accuracy < 60 cm with DGPS, see more details in the technical reference manual [51]. Each tree

in the plot was measured (diameter at breast height, dbh) and localized, based on distance and azimuth from the plot centre. The basal area and volume were computed based on the dbh of each tree and scaled per hectare, using species-specific allometric equations for the wood volume [32]. Heights were measured only for a subset of trees (about three per plot) to compute the top height, which was defined as the mean height of the 100 largest trees for an area of 1 ha [88].

A social status was assigned to each tree, based on dbh_i/dbh_{Dom}, the ratio between its dbh and the dominant dbh (computed for each plot as the mean dbh of the 100 largest trees for an area of 1 ha). Three classes were defined: dominant ($dbh_i/dbh_{Dom} > 0.9$), co-dominant ($0.75 < dbh_i/dbh_{Dom} < 0.9$), and dominated tree ($dbh_i/dbh_{Dom} <= 0.75$).

Checking of plot and tree positions was done with ortho-images and a Canopy Height Model (CHM) to improve the matching between field data and aerial data, followed by a second field visit for ambiguous cases. To carefully consider the trees bordering plot, we also considered trees within a 2-m buffer around the plot (extended plot).

Table 4.1: Dendrometric characteristics of 72 plots. Initially based on high-density plantations of established even-aged stands, this area is now involved in a process of transformation to uneven-aged structure.

	Minimum	Mean	Maximum	SD
Plot radius (m)	8	14.4	18	2.5
Number of trees (ha^{-1})	147	300	945	164
Top height (m)	16.7	24.7	27.9	3.8
Basal area (m^2ha^{-1})	21.3	35.1	59.2	8.6
Volume (m^3ha^{-1})	183.9	438.8	844.4	136.4

4.2.3 *Aerial Data Acquisition and Processing*

UAS Surveys

The Gatewing X100 small UAS is a professional fixed wing drone dedicated to rapid mapping (technical characteristics detailed in Table 4.2). The Gatewing X100 flies with a small positive pitch angle, which causes a low-oblique vantage acquisition. The flight plans were prepared with the software Quickfield® [49] on a rugged tablet computer prior to the aerial survey. The software allows the preparation of several adjacent flights with overlap when the surface of the site implies more than one survey. Flights are fully automatic from take-off to landing and complete stop. The sensors were two similar RGB compact camera from Ricoh (GR3 and GR4 still camera, 10 megapixel Charged Coupled Device, 6 mm focal length or 28 mm in 35 mm equivalent focal length).

Table 4.2: Characteristics of small fixed wing UAS Gatewing X100

Wingspan	100 cm
Weight	2.2 kg
Cruise speed	80 km h^{-1}
Flight height	100–750 m
Maximum flight duration	40 min
Launch	Catapult
Landing	Belly landing

Surveys were performed in the autumn of 2014 at a flight altitude of 200 m above ground level. Flight plans were ordered from north to south, and acquisitions were carried out by a duo of professional remote pilots. Eight flights (in 4 days) were required in order to capture the entire forest estate of Hattlich: the first flight on 1 October 2014, three successive flights on 2 October 2014, the southern part of the forest estate on 28 October (two flights) and 20 November(the last two flights).

The four first flights were performed with the Ricoh GR4 camera. Then, a contact failure affecting the triggering cable obliged the pilots to change both triggering cable and the camera. The last four flights thus used the Ricoh GR3 camera.

An images side and forward overlap of 80% was set up. The weather was cloudy and without direct sunlight. The camera was used in manual mode, and exposure settings (ISO and shutter speed) were set before each take-off according to the light conditions. Because of late season, the duo of remote pilots made a great effort to operate close to the solar noon, large shadows cast by the trees were visible during most of the survey.

The surveys resulted in a total of 3341 individual low oblique aerial images, covering an area of 930 ha. The average ground sample distance of these images was 6.76 cm $pixel^{-1}$, and the swath width on the ground was about 250 m.

Photogrammetric Workflow and Outputs

The photogrammetric open source toolbox `MICMAC` [113] was used to generate a 3D canopy surface model and output images suited for mapping purposes, starting with the raw image block. The workflow is synthesized in Figure 4.2 .

All the images were first handled as a unique image block for the computation of image orientation by aerotriangulation. First, potential image pairs are determined based on the GPS data (Oriconvert tool of `MICMAC`), for the extraction of the tie points (Tapioca tool of `MICMAC`). Even though the compact cameras used in this study were precalibrated in the laboratory, the internal parameters are still refined during the bundle block adjustment (self-calibrating bundle block adjustment [112]). We used embedded GPS information to achieve constrained

bundle block adjustment which supports both tie point and embedded GPS observations (Campari tool of MICMAC), in order to remove non-linear distortions that may otherwise taint the photogrammetric models [163].

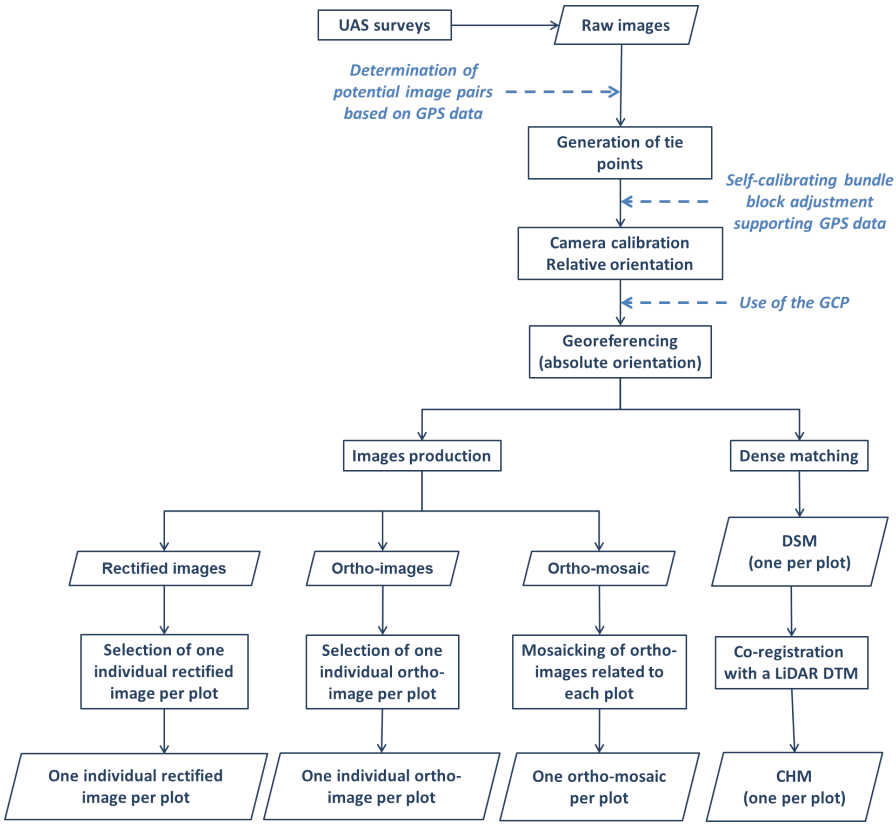


Figure 4.2: Synthesis of the photogrammetric workflow

Image orientation succeeded for 3321 images but failed for 20 images. This failure occurred for images showing very large regular spruce stands. Indeed, the repetitive texture of even-aged stands hindered the process of determining tie points between two images. Moreover, these images were located in an area with important variation in the relief, involving less overlap between images.

Then, georeferencing was achieved by using ground control points (GCP) (coordinate system: Belgian Lambert 72, EPSG 31370). Nine GCP (road crossing, curb, road markings) were identified, in parallel, on the UAS images and on a reference regional orthophotos coverage (non-UAS, 2012, 0.25 m GSD) to determine their planimetric coordinates. The altimetric positions were extracted from a LiDAR-DSM.

Then, image dense matching was operated in image geometry for each overlapping image pair. Automated dense matching algorithms use image similarity measures (here, the normalized cross-correlation score) to establish a correspondence between homologous windows (matching windows) through the image pair and to compute a stereomodel. We used the per image matching tools from the

MICMAC suite in order to generate the DSM for each individual forest plot. The dense point cloud is generated at a resolution of 1:4 of the raw images. For example, if the flight height above the canopy level is 180 m, the image ground sample distance is 6 cm pixel^{-1} and a photogrammetric DSM of 24 cm pixel^{-1} is computed. In addition to the photogrammetric DSM, a map of the normalized cross-correlation score, expressing the similarity of the images at each pixel of the 3D model, was produced. For further details on the photogrammetric procedure used in this study, we refer the reader to "The Grand-Leez dataset" of the MICMAC user guide [111].

In parallel to the photogrammetric DSM generation, we produced images for each plot to proceed the tree detection approach. True ortho-rectification is promoted by the remote-sensing community (see e.g. [67]), which consists of removing object relief displacements by utilizing a fine DSM during the ortho-rectification process, but true ortho-rectification applied on UAS images causes many artefacts in forested areas. In consequence, we decided to implement image rectification on single images. To complete our study, we decided to include also orthorectification on single image to test three different levels of photogrammetric treatment and to analyse the quality of the tree detection regarding the type of output: individual rectified image, individual ortho-image and ortho-mosaic.

The image rectification process involves the projection of the image to a reference horizontal plane. Although assimilation of the canopy surface to a horizontal plane is an extreme simplification impacting the geometric quality of the georeferenced image, the rectified image has fewer artefacts and appears more natural to the human eye than a true ortho-rectified image (Figure 4.3). During the image rectification process, radial displacement due to lens distortion is corrected by using the camera calibration. The process of image rectification is sufficient to remove scale and perspective distortions [160]. Rectified images are thus suitable for mapping, as displacement due to relief occurs mainly on the images edge which present a larger incident angle. For each plot, several rectified images were available, thanks to the high rate of overlap during the flights. Selection of these images was conducted in two steps. First, we kept images with their centre closest to the plot centre. Second, a visual selection was done amongst the remaining images by an operator to keep only a single image that shows the lowest deformation of trees and is as sharp as possible.

In parallel to the rectification process, ortho-rectification was applied to the manually selected images (used for rectification) to produce single individual ortho-images per plot. Besides, ortho-image mosaics were produced for each plot from the multiple overlapping images, which are stitched together based on the lowest angle of view of each image. The orthorectification and the mosaicking were performed separately for each forest plot at a resolution of 1:2 full resolution.

Finally, a CHM was computed for each plot as the difference between a LiDAR DTM and the UAS photogrammetric DSM at the DSM resolution. This DTM was produced with a spatial resolution of 1 m from ALS data with a density of 1 ground classified point per m^2 on the whole Southern Belgium territory. Alignment of DSM and DTM was performed by a co-registration process using ground controls points.

As the flight altitude was constant, change in terrain elevation impacted the distance between the UAS and forest canopy. Therefore, the overlap decreases if the terrain elevation increases, and the GSD worsened. Thus, because of the steep relief, changes in elevation impacted the individual image resolution, which varied from 4.4 to 8.5 cm $pixel^{-1}$. The resulting CHM of 1:4 full resolution thus ranged from 18 to 35 cm $pixel^{-1}$. Table 4.3 summarizes the flight heights and resulting resolutions of the different photogrammetric products. Figure 4.3 shows an example of these photogrammetric products for one plot, and the corresponding 3D point cloud.

Table 4.3: Flight heights and resulting resolutions of different photogrammetric products. The steep relief involves variations of elevation which impact the individual image resolution and the resulting CHM resolution (ranged from 18 to 35 cm).

	Minimum	Mean	Maximum
Flight height (m)	129	179	250
Mean raw resolution (cm)	4.4	6.1	8.5
CHM resolution (cm)	18.0	25.5	35.0
Rectified image resolution (cm)	9.2	12.7	18.2
Ortho-image resolution (cm)	9.0	12.8	17.5
Ortho-mosaic resolution (cm)	9.0	12.8	17.5

4.3 Methodology

4.3.1 *Tree Detection*

We implemented a comprehensive tree top detection algorithm, which takes the best advantage of the different photogrammetric products. This tree top detection algorithm followed a multi-source strategy, based on a local maxima detection approach (Figure 4.4). The identification of the tree tops is based on the detection of local maxima from a CHM and UAS imagery. Local maxima detection is simple, versatile and easy to implement on a large number of data types and in an operational context. We chose to work on a small moving window, not only promoting the detection of the greatest number of trees but also involving detection of a large number of false positives. To eliminate these false positives, a classification step

is supervised by the implementation of random forest. The objective is to determine whether a local maximum is actually a tree or whether it is a false positive, based on spectral, 3D, or neighbourhood variables.

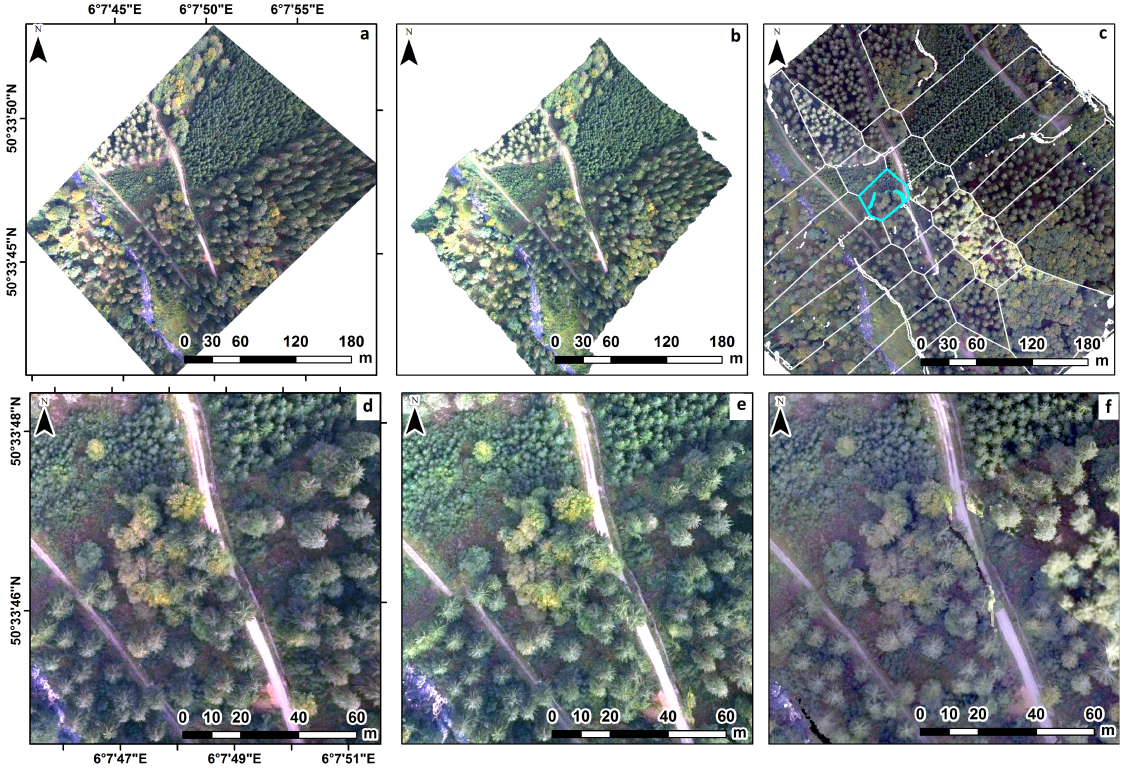


Figure 4.3: Three different photogrammetric products are computed and compared for the local maxima detection: rectified image (a); individual ortho-image (b) and ortho-photo mosaic. The illustrated area shows a variety of stand types within the study site. White lines on c represent the boundaries between the different ortho-photo making up the mosaic and the blue line corresponds to the part of the individual ortho-photo image (b) used in the mosaic. d, e and f are a selection within each image product. The rectification and true-orthorectification differ mainly on the image border but not on the nadir part (centre of the image). Trees looked more natural to human eyes in rectified images. As we worked with a high overlap rate and we cropped the data at the plot scale, corresponding to only a small part of each image, close to the nadir, discrepancies between the three images are less important.

Local Maxima Detection

Before the local maxima detection, a Gaussian filter was applied to both CHM and image products. The Gaussian filter is a low-pass filter, preserving boundaries and edges, reducing high-frequency components. The aim was to remove micro-peaks of the CHM (coming from the DSM), specific to the photogrammetric process and bringing noise during local maxima detection [104, 145]. The Gaus-

sian filter was implemented using the Focal Function (fun = mean) using a weight matrix computed by the Focalweight function (type = Gauss, sigma = 1.5). The filter consisted in an average weighted mean of the neighbour pixels. The weight is defined by a Gaussian window whose size is variable, depending on the CHM resolution.

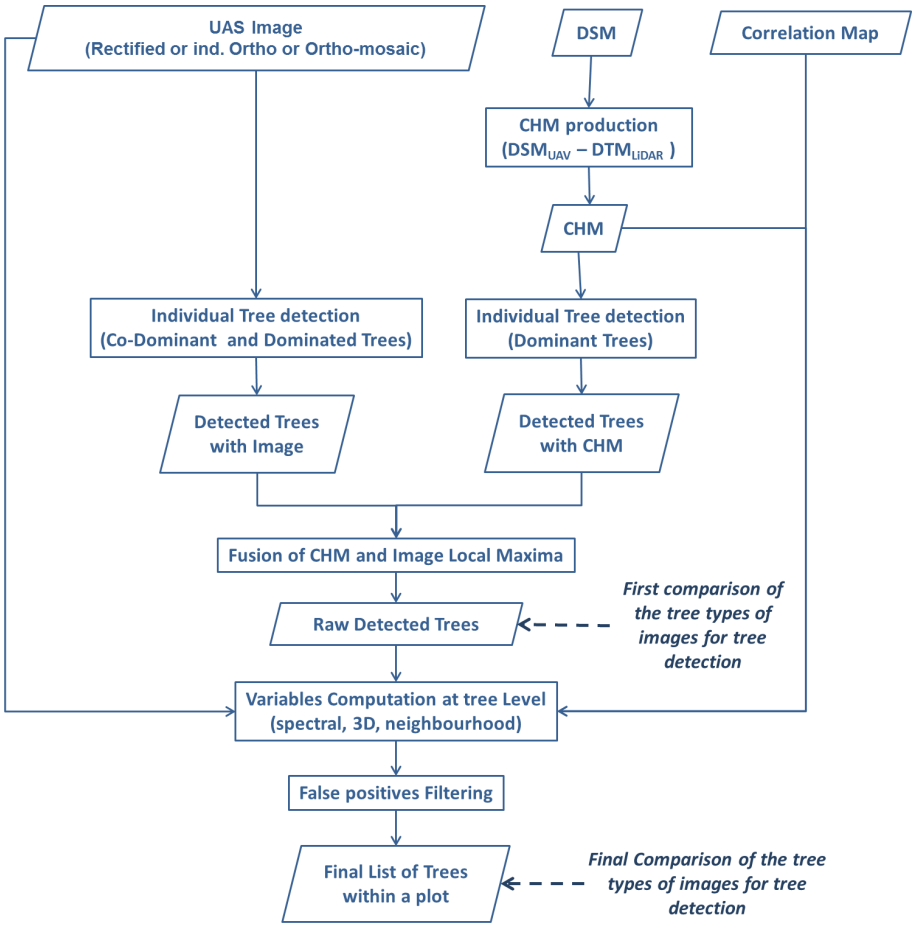


Figure 4.4: Synthesis of the methodology workflow: the comprehensive tree top detection process take the best advantage of the different photogrammetric products (images, DSM, and correlation map) to extract tree tops that will be used with the CHM to assess forest attributes.

The local maxima detection consisted of identification of the tree tops within each plot based on the local maxima detection within a fixed radius (1.5 m) moving window from UAS images and CHM. The whole analysis was implemented in the R software [124] plot by plot. The local maxima detection result gave the coordinates (x_i, y_i) and height (h_i) of each estimated tree top location within a plot.

In an attempt to compare the contributions of the different photogrammetric products, we analysed the local maxima obtained from the CHM, rectified images, individual ortho-rectified images, ortho-image mosaic, and the combination of the CHM with each of the tree image products. Dominated tree tops are not considered

local maxima on the CHM because of the smoothing of the canopy relief inherent in photogrammetric DSM [87]. The tree detection with the UAS images was supposed to be complementary to CHM (which mainly detects dominant trees) and improve the detection of co-dominant and dominated. When merging local maxima from CHM and local maxima from image, duplicates (defined as a tree top detected by both CHM and the images, within a distance of 1.5 m) were removed, considering the local maxima from CHM as the primary.

We focused on the omission and true-positive rates as indicators of the detection quality (omission has to be minimized and true positives maximized). For this purpose, the local maxima were categorized as true positives or false positives by comparing with trees mapped in the field. False positives were also taken into account but not considered a decisive factor as we implemented a filtering process in our workflow.

Local Maxima Filtering

After the local maxima detection, our aim was to classify local maxima as true or false positives using a supervised classification by random forest. For each local maxima, variables were computed from several data in order to prepare for the detection of false positives. Table 4.4 presents the tested variables. These variables are either extracted from the tree top positions or computed within a 1.5 m window (considered the central part of the crown) and based on 3D information, spectral information, and distances to neighbours. Among the 3D information, statistics from the surface model are included, as well as information from the normalized cross-correlation score.

Random forests [80] are a very efficient algorithm for classification. The Out-Of-Bag error (OOB) is used to analyse the prediction performance and to assess the importance of each variable (VarImp) which allows sorting of predictors. Calculation of the OOB is based on a bootstrap sample of the data. The VSURF package [53] allows the selection of variables based on variable importance ranking. The objective is to remove irrelevant variables and select important variables (1) for interpretation purposes and (2) for identification of a sufficient parsimonious set of variables for prediction purposes. This set of identified prediction variables will be used to construct the random forests. The strategy of VSURF is based on a recursive elimination of variables. Three parameters are especially important when using VSURF: `nfor` is the number of random forests constructed during the process, `ntree` is the number of trees in each forest, and `nmj` (number of mean jumps) allows the discrimination by relevant or noise variables. In this study, `nfor` was set to 20, `ntree` to 1000, and `nmj` to 4. The other parameters were set to default values.

Table 4.4: Several variables are used for false-positive filtering by a random forest process (supervised classification). These variables are computed from the CHM, image, and correlation map. The values are extracted either from the local maxima position or a statistic is computed within a 1.5 m window around the local maxima. Information about the neighbor local maxima are also taken into account for the false-positive diagnosis.

Variable name	Description
Type_{LM}	Local maximum type, corresponding to source (CHM, image type)
CHM	
h_{LM}	Height of local maximum extracted from CHM
h_{mean}	Mean height within 1.5 m window around local maximum
h_{max}	Maximum height within 1.5 m window around local maximum
h_{SD}	Standard deviation of height within 1.5 m window around local maximum
Image	
r_{LM}	Normalized red reflectance value of local maximum extracted from image $R_R/(R_R+R_G+R_B)$
r_{mean}	Mean normalized red reflectance value within 1.5 m window around local maximum
g_{LM}	Normalized green reflectance value of local maximum extracted from image $R_G/(R_R+R_G+R_B)$
g_{mean}	Mean normalized green reflectance value within 1.5 m window around local maximum
b_{LM}	Normalized blue reflectance value of local maximum extracted from image $R_B/(R_R+R_G+R_B)$
b_{mean}	Mean normalized blue reflectance value within 1.5 m window around local maximum
GR_{LM}	Green and red reflectance values ratio R_G/R_R of local maximum extracted from image
GR_{mean}	Mean green and red reflectance values ratio R_G/R_R within 1.5 m window around local maximum
GB_{LM}	Green and blue reflectance values ratio R_G/R_B of local maximum extracted from image
GB_{mean}	Mean Green and blue reflectance values ratio R_G/R_B within 1.5 m window around local maximum
RB_{LM}	Red and blue reflectance values ratio R_R/R_B of local maximum extracted from image
RB_{mean}	Mean red and blue reflectance values ratio R_R/R_B within 1.5 m window around local maximum
Bright_{LM}	Brightness value of local maximum extracted from image $(R_R+R_G+R_B)/3$
Bright_{mean}	Mean brightness within 1.5 m window around local maximum
NRBVI_{LM}	Normalized Red Blue Vegetation Index of local maximum$(R_R-R_B)/(R_R+R_B)$
NRBVI_{mean}	Mean NRBVI within 1.5 m window around local maximum
NGBVI_{LM}	Normalized Green Blue Vegetation Index of local maximum$(R_G-R_B)/(R_G+R_B)$
NGBVI_{mean}	Mean NGBVI within 1.5 m window around local maximum
NGRVI_{LM}	Normalized Green Red Vegetation Index of local maximum $(R_G-R_R)/(R_G+R_R)$
NGRVI_{mean}	Mean NGRVI within 1.5 m window around local maximum
Correlation map	
c_{mean}	Mean correlation value within 1.5 m window around local maximum
c_{SD}	Standard deviation of correlation value within 1.5 m window around local maximum
Neighbourhood	
Only the two nearest neighbour(N_x with x = 1 or 2) of the current local maximum are considered	
d_{N_x}	Distance between current local maximum and neighbour
Type_{N_x}	Local maximum type, corresponding to source (CHM, image type)
$h_{max N_x}$	Maximum height within 1.5 m window around neighbour
$h_{mean N_x}$	Mean height within 1.5 m window around neighbour
$h_{rel N_x}$	Relative height (ratio between neighbour height and current local maximum height)
$r_{mean N_x}$	Mean red reflectance value within 1.5 m window around neighbour
$g_{mean N_x}$	Mean green reflectance value within 1.5 m window around neighbour
$b_{mean N_x}$	Mean blue reflectance value within 1.5 m window around neighbour
$c_{mean N_x}$	Mean correlation value within 1.5 m window around neighbour

Half of the plots (36) were randomly selected for variable selection and classification training. Each local maximum within the training plots was identified as a true positive or false positive by comparison with field trees by an operator. Then, a validation of the filtering procedure was performed with all the trees in the other 36 plots. The latter were used for construction and validation of forest attribute models.

4.3.2 *Forest Attribute Assessment*

Two alternative approaches are usually used to assess forest inventory attributes with remote-sensing data: (1) Area-Based approach relies on the extraction of variables at plot-scale from point clouds or CHM while (2) individual tree detection approach is obviously based on the detection if not delineation of trees/crowns. These approaches, although traditionally opposed, are in fact complementary. In order to take advantage of their respective strengths, we choose to combine them.

Linear regression with a parsimonious selection of variables extracted from the detected trees and CHM was considered to assess forest inventory attributes (individual tree height, number of trees per hectare, top height, and volume and basal area per hectare). A deliberately limited number of predictors were used to rationalize the variables selection for the forest attributes assessment. Model selection was based on the best subset regression analysis (*regsubsets* in the R package "leaps" [90]) and with the bayesian information criterion (BIC). Field inventory data were used as reference : the 36 plots not used previously for local maxima filtering for top height, number of trees, basal area and volume and the trees with individual height measure within these plots (n = 102). The quality of the regression models was assessed using the R^2 and the Root Mean Square Error (RMSE) values. A cross validation was also implemented with the package *bootstrap* in R [138]. The data are split into five groups to analyse the quality of the results.

Individual tree height estimation was based on variables calculated at tree-level: h_{LM}, h_{mean}, h_{max} and h_{SD} (see Table 4.4). For the stand level attributes, we adopted a mixed approach combining area-based and tree-based variables. We identified a limited number of variables to reduce the process of predictor selection among numerous variables, to limit over-fitting, and consider the low number of plots [65]. Table 4.5 presents the variables computed at the plot level and from the local maxima to construct the forest stand attribute models.

The 12 selected variables represent the horizontal and vertical structure and were tested for the fitting of forest attribute models. The number of trees (N_{Trees}) is the number of local maxima classified as true positives and normalized by the area

of the plot. The maximum height (h_{max}) and the coefficient of variation of height (h_{cv}) are computed from the local maxima within the plot. The aggregation index of Clark & Evans ($Aggreg_{uas}$, Equation 4.1) describes the location diversity of the trees, based on the average distance between trees ($\bar{d}_{Trees}$) and ranges between 0 and 2.1419. The 95^{th} percentile of the CHM within the plot (CHM_{p95}), the canopy cover (CC_p; canopy is identified by a height threshold of 5 m on the CHM), and the volume below the canopy top, normalized by the plot area ($Vol_{CHM\text{-}X}$), are computed at the plot level from the CHM. $V_X N_{Trees}$ and $V_X\_N_{Trees}$ are mixed variables, defined respectively as the product and ratio of $Vol_{CHM\text{-}X}$ and N_{Trees}. CHM_{p95}, h_{max}, and h_{cv} are directly related to the height of trees and the variability. N_{Trees} and CC_p describe the density of the plot (in particular, CC_p is related to the proportion of gaps within a plot). $Vol_{CHM\text{-}5}$ and $V_X\_N_{Trees}$ are considered to be a proxy for the amount of biomass.

$$Aggreg_{uas} = \frac{\bar{d}_{Trees}}{\left(\frac{1}{2 * \sqrt{\frac{N_{Trees}}{Area_{Plot}}}}\right)} \tag{4.1}$$

Area Assumptions for the linear regression (homoscedasticity, independence of errors, normality of error distribution, absence of bias) were checked (with several tests using the *car* and *lmtest* packages [46, 167]) to ensure model validity. The multicollinearity was also tested with the variance inflation factor (VIF). A VIF value of less than 2 was considered the threshold for no collinearity.

Table 4.5: Twelve variables are selected for construction of forest attribute models. These variables are from the local maxima detection, or are computed at plot-level (ABA), or are a combination of local maxima detection and ABA.

Variable Name	Description	Type
N_{Trees}	Number of trees (classified as true positives local maxima) within the plot	Tree
h_{max}	Local maximum with maximum height within the plot	Tree
h_{cv}	Coefficient of variation of local maxima height within the plot	Tree
$Aggreg_{uas}$	Clark & Evans aggregation index computed from local maxima	Tree
CHM_{p95}	95^{th} percentile of CHM within the plot	Plot
CC_p	Canopy cover is the percentage of plot area covered by canopy (canopy is identified by a height threshold of 5 m on the CHM)	Plot
$Vol_{CHM\text{-}5}$	Volume between canopy (defined by CHM) and 5-m-height threshold, normalized by plot area	Plot
$Vol_{CHM\text{-}16}$	Volume between canopy (defined by the CHM) and 16-m-height threshold, normalized by plot area	Plot
$V_5 N_{Trees}$	Product of $Vol_{CHM\text{-}5}$ and N_{Trees}	Mixed
$V_{16} N_{Trees}$	Product of $Vol_{CHM\text{-}16}$ and N_{Trees}	Mixed
$V_5\_N_{Trees}$	Ratio of $Vol_{CHM\text{-}5}$ and N_{Trees}	Mixed
$V_{16}\_N_{Trees}$	Ratio of $Vol_{CHM\text{-}16}$ and N_{Trees}	Mixed

4.4 Results

4.4.1 *Tree Detection*

Local Maxima Detection

Results of the local maxima detection are illustrated in Figures 4.5 and 4.6. The lowest false-positive rate is obtained with CHM only, but with poor results regarding omission. Worse results are obtained with rectified images. Individual ortho-rectified images and ortho-image mosaic combined with CHM give similar results. The best true positives and omission rates are obtained with a combination of CHM and imagery. For this reason, we decided to further investigate only these three methods for the false-positive filtering. The effect of the variation of the resolution between the plots is analysed in Figure 4.7. The detection performances of each plot regarding the resolution are presented in terms of percentages of omissions, true and false positives (before filtering) for CHM and mosaic.

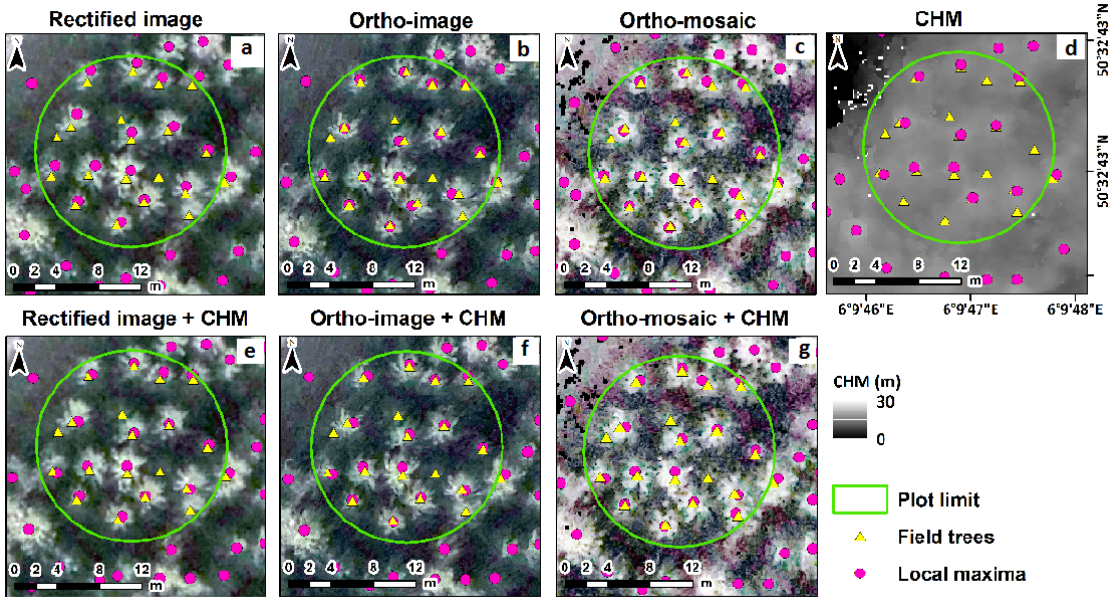


Figure 4.5: Example of local maxima detection results for the seven tested data sources. In this plot, eight trees are not detected with the CHM (d). Amongst the missing trees, four are identified with the rectified images (a) and individual ortho-rectified images (b) and five with the ortho-image mosaic (c), showing the benefit of the imagery in the tree top detection process (e,f,g).

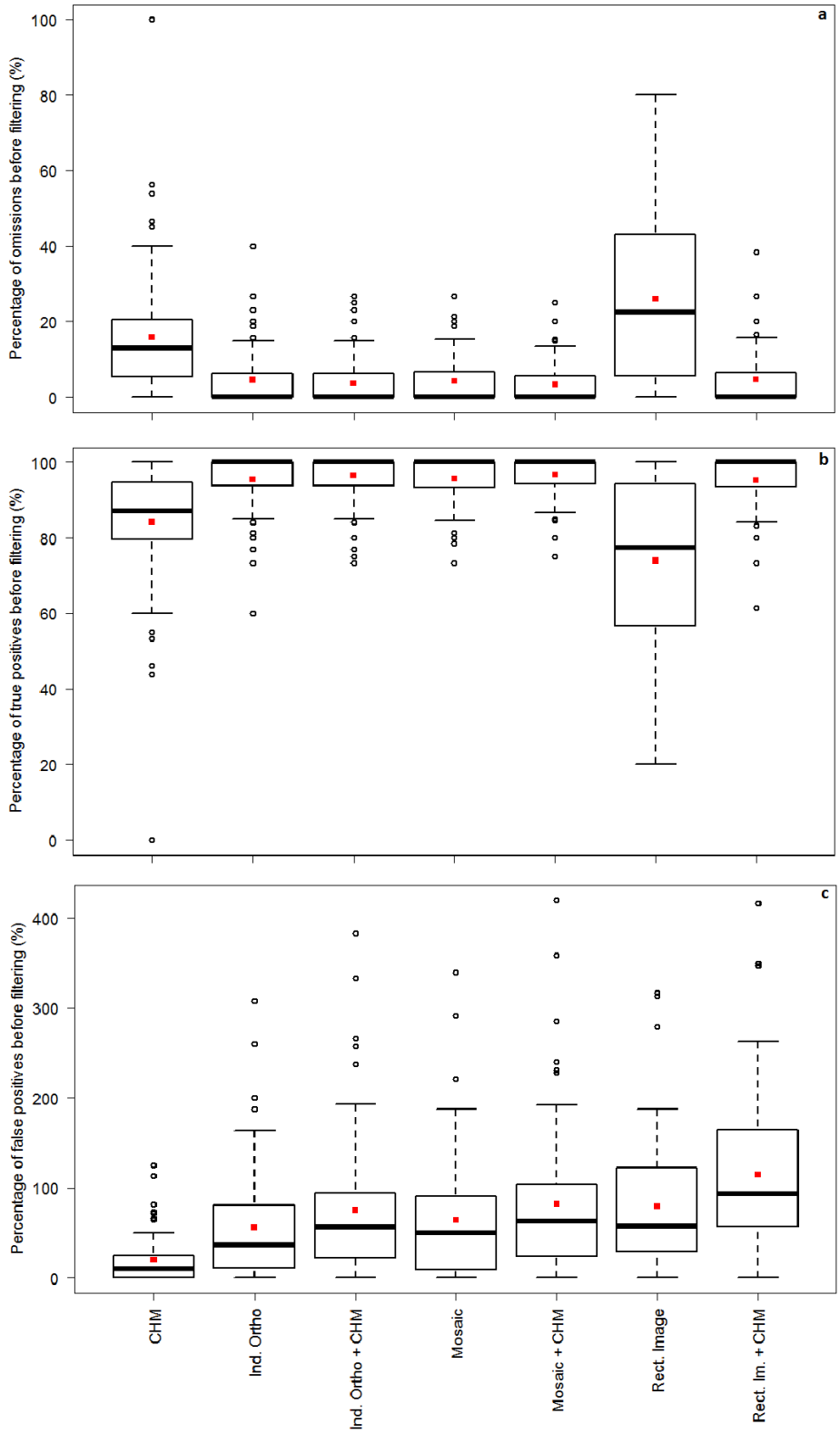


Figure 4.6: Detection performance of local maxima for the 72 plots by comparing results for the different data sources considered regarding omissions rate (a), true positives rate (b), and false positives (c). These results precede the false positives filtering. CHM gave lowest false-positive results, but with higher omission, showing the benefit of combining CHM and imagery. Rectified images alone gave the worst results.

Filtering of Local Maxima

Trees from the training plots were used for the variable selection and the classification process. Several variables, either from CHM or imagery, were selected by VSURF for the three combinations of data (Table 4.6). These variables were then used for training the random forests. The resulting random forests were applied to the local maxima of the validation plots. The quality of the random forests (OOB errors) and the corrected rates of omission, true positives, and false positives are also presented in Table 4.6. Both CHM and spectral information contribute to the detection of false positives. The false-positive rate is dramatically reduced after the filtering and the results are overall good even if intermediate false positives results are poor.

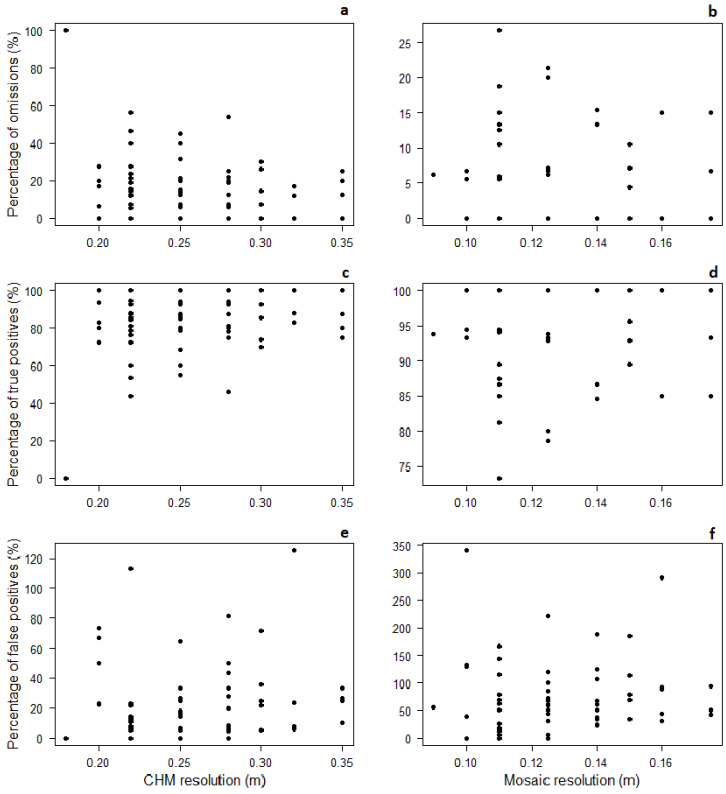


Figure 4.7: Detection performance of local maxima for the 72 plots by comparing results for the different resolutions for mosaic and CHM regarding omissions (a, b), true positives (c, d) and false positives (e, f) rates before the false-positive filtering.

Figure 4.8 shows an example of the resulting local maxima after the filtering process. Considering the detection performance after filtering (omission, true positives, OOB), the combination of mosaic and CHM is the best for extracting local maxima. Further analysis will only take into account the local maxima derived from CHM and mosaic.

Table 4.6: Comparison of detection performance for the three combinations of CHM and imagery. Omission (Om), true positives (TP) and false positives (FP) rates are computed compared to the number of trees in the field. The sum of true positives and omission rate gives 100%. Before filtering results are raw results with the false positives identifications. The variables used in the random forest are listed in descending order of importance. After filtering, results refer to the removal of false positives from the validation plot using the selected random forest variables. A decrease in the true positives rate is observed after filtering, as some true positives are misclassified as false positives during the classification process.

	CHM + Rect. image	**CHM + Ortho**	**CHM + Mosaic**
Before filtering			
Om. rate (%)	4.8	3.5	3.4
TP rate (%)	95.2	96.5	96.6
FP rate (%)	114.5	75.3	81.9
Var. selection			
Variables	h_{SD}, c_{mean}, h_{relN_1}, h_{LM}, h_{mean}, $Bright_{mean}$, d_{N_2}, $Type_{N_1}$, h_{meanN_1}, h_{meanN_2}	h_{SD}, $Bright_{mean}$, c_{mean}, d_{N_1}, h_{LM}, h_{relN_1}, h_{maxN_1}	$NGRVI_{mean}$, $NGBVI_{mean}$, $Bright_{mean}$, h_{SD}, d_{N_2}, GR_{mean}, h_{relN_1}
OOB error (%)	8.5	5.5	5.5
After filtering			
Om. rate (%)	9.2	13.9	10.8
TP rate (%)	85.9	86.1	89.2
FP rate (%)	14.1	2.6	7.8

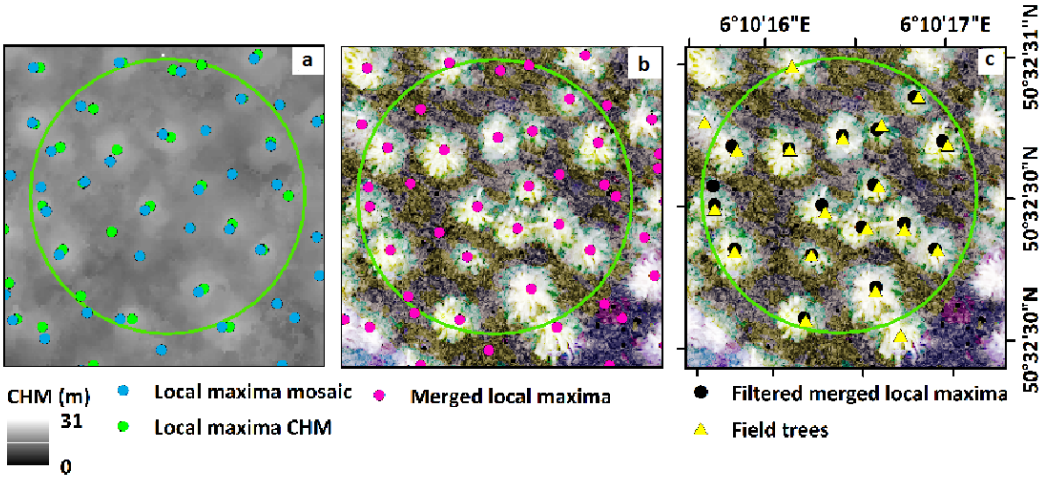


Figure 4.8: Example of intermediate and final results for one plot: local maxima produced from CHM and Image (a) merged local maxima with duplicates removing (b) final results after removing local maxima outside the plot and false positives filtering (c). Several trees missed by the CHM are identified by imagery and wrong detection is corrected during the filtering process.

The detection performance was analysed in terms of social status for the combination mosaic and CHM (see Table 4.7). A percentage 48.2 of the missed trees were dominated trees, 28.7% co-dominant, and 23.1% were dominant. The omission rate was 17.7% but it was partially balanced by false positives. Of the dominant trees, 89.8% were detected, 85.9% for the co-dominant as well as 64.6% of the dominated ones. Amongst the detected trees, 81.2% were dominant or co-dominant and only 18.8% were dominated.

Table 4.7: Detection performance in terms of social status after false-positive filtering. The first row is the distribution of the trees amongst the social status. The second row is the percentage of detected trees for each social status. The third is the distribution of the detected trees by social status. The fourth row is the distribution of the omitted trees by social status. The fifth and sixth rows are the detecting and omitting rate by social status.

	Social status		
%	**Dominant**	**Co-dominant**	**Dominated**
Distribution of number of trees	40.1	35.9	24.0
Detecting rate for each social status	89.8	85.9	64.6
Distribution of the detected trees	43.8	37.4	18.8
Distribution of the omitted trees	23.1	28.7	48.2
Detected trees by total number of trees	36.1	30.8	15.5
Omitted trees by total number of trees	4.1	5.1	8.5

4.4.2 *Forest Attribute Assessment*

In this study, we focused on the prediction of four stand-level forest attributes (number of stems, top height, basal area, and volume) and one tree-level attribute (individual tree height). At the stand-scale, four models with only a single variable and four models with maximum of three variables were fitted and compared to assess the contribution of variables in reducing noise and improving model performance. One model was constructed for individual tree height. The variables for the single models were those most closely related to the considered attributes (e.g. N_{Trees} for the number of stems). Table 4.8 and Figure 4.9 present the selected variables and the performance of the models in terms of the R^2 and RMSE during the model fitting and the cross-validation.

The R^2 results during cross validation are very close for the number of trees, basal area and volume (71%, 70% and 69%) and better for top height (83%). Regarding % RMSE, the performance was more contrasted: top height gave also the best result, followed by basal area (16.2%), volume (20.1%), and number of trees (28.3%)). According the different tests, the required assumptions of linear regression are all met. The use of several variables improves the quality of the fours models, especially for the assessment of the tree number.

Table 4.8: Performance and variables selected for each forest attribute model. A parsimonious predictor selection was done among a limited number of variables. Cross-validation results are less optimistic than fitting, but are still quite good. Contribution of each selected predictor is significant.

Attributes	Variables	Intercept	Coefficients	R^2 (%)/ R^2 CV (%)	RMSE/ RMSE CV	% RMSE/ % RMSE CV
Top height (m)	h_{max}	9.28	0.666	78/75	1.8/1.9	6.5/6.8
	h_{max}, N_{Trees},$Aggreg_{uas}$	6.04	0.568; -0.012; 5.802	86/83	1.4/1.6	5.0/5.7
Number of trees (ha^{-1})	N_{Trees}	-48.25	1.729	71/62	84/98	28/32.7
	N_{Trees}, Vol_{CHM-5}	67.93	2.51; -6.44×10^{-6}	81/71	71/85	23.7/28.3
Basal area (m^2 ha^{-1})	$V_{16}N_{Trees}$	24.51	1.499×10^{-6}	70/70	5.8/5.7	16.5/16.2
	$V_{16}N_{Trees}$, CC_p	3.177	1.412×10^{-6}; 0.234	76/70	5.1/5.7	14.5/16.2
Volume (m^3 ha^{-1})	$V_{16}N_{Trees}$	274.4	2.306×10^{-5}	72/68	84/89	19.1/20.3
	$V_{16}N_{Trees}$, CC_p	89.13	2.221×10^{-5}; 2.042	73/69	81/88	18.5/20.1
Ind. height (m)	h_{max-LM}	8.764	0.73	71/70	1.9/2	6.9/7.2

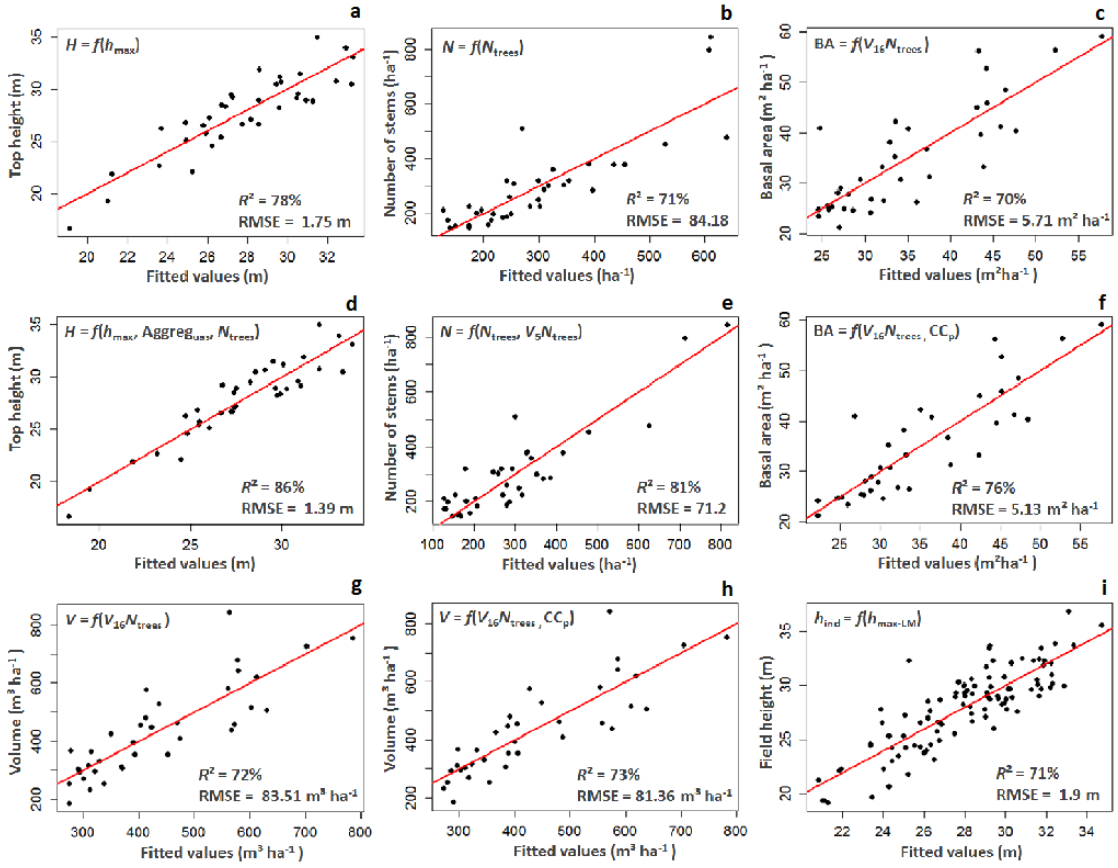


Figure 4.9: Comparison of fitted values with reference values (field data) for the four stand models (a, d: top height; b, e: number of stems; c, f: basal area; g, h: volume; n = 36) and individual tree height model (i; n = 102).

Six predictors appear to be highly significant for the models: h_{max}, N_{Trees} & $Aggreg_{uas}$ (tree-level variables), $V_5 N_{Trees}$ & CC_p (ABA variables), and $V_{16} N_{Trees}$ (mixed variable).

4.5 Discussion

Our study provides additional support for promoting the relevance of UAS data in the framework of forest resource assessment. Below, we highlight several key points from UAS data acquisition to forest attributes estimation.

4.5.1 *UAS data acquisition*

Relief variation

Significant variations in relief at the study site caused lack of overlap between images and variation in resolution of the photogrammetric products. However, our results were not affected by the between-plots variation of resolution. Our methodology, based on simple techniques, has proved to be robust and versatile and the procedure also managed the changes of cameras (Ricoh GR3 and GR4). Besides, the pixel size variations were very tenuous and the resolution remains excellent and seems sufficient in relation to the application.

Light conditions

Imagery is strongly impacted by variation in illumination due to season, weather and shadows. These latter reduce the spectral quality of images and can hinder classification or 3D reconstruction [12, 159, 152, 102]. [123] faced with difficult weather conditions due to snow patches alternating with dark forest involving saturation and high contrast. [87] concluded that shadows conduct to overestimation of the CHM (as occlusion and smoothing), while in [152], the predicted canopy cover is lower with Structure-from-motion than with ALS. Several ideas emerge to solve the shadow issue. For example, [86] keep only sunlit part of crowns for species discrimination. In their study, [102] tested a method to detect and remove shadowed pixels in UAS images with Maximum Likelihood and Support Vector Machine classifiers.

In this research, the geometry of mosaics is suitable for mapping, but they suffer from heterogeneity in the image exposures because of illumination differences during the flights. We suspected that this variation in illumination would negate the local maxima detection. This is the reason why we tested the local maxima detection on individual ortho and rectified images, in order to get rid of the variation in illumination present amongst the image block. However, as shown by

the comparison of individual ortho-image and mosaic, the variation of illumination of our images block didn't impact the local maxima detection. Our choice of implementing apex detection instead of crown delineation minimizes probably the impact of shadows. Moreover, the spectral heterogeneity could have an effect on the classification of false positives but their identification was effective. The use of spectral indices and normalized spectral bands helped to manage this problem.

Recommendations

In our experience, major improvements rely on the standardization and optimization of UAS acquisition in order to create the finest conditions and to mitigate negative effects on the accuracy. We thus recommend a standardization of the spatial resolution, use of a unique camera, choice of the flight date/hours to minimize the shadows, and reduction of the number of flight days (thanks to the evolution of the drone platform or batteries).

In our opinion, we consider that the diffuse light conditions are the most suitable for a manual adjustment of the exposure and the low dynamics of the compact sensors. A constant cloudy cover is therefore suitable for obtaining radiometrically equalized images, and these conditions are largely preferable to full sun at noon solar. The management of shadows is challenging for the optimization of flight planning and can be a shortcoming for the flexibility of UAS by reducing the acquisition window.

We argue on the importance of a reflection on the implementation of data acquisition protocols and data processing adapted to forest areas could allow the good production of maps of canopy relief from drone imagery [35].

4.5.2 *Field Data Collection*

Regarding the field data collection, a point of discussion is the variation of plot radius which is relatively common in traditional forest inventory. Larger and fixed plot radius would have been more relevant, especially to decrease the edge effect and the instability of the tree detection results. Positioning of plots and trees remains also a delicate issue under the canopy, especially for the pairing of local maxima and field trees. Although we used dGPS, we took the necessary time to carefully validate the tree centre position by photo-interpretation and in the field if necessary.

4.5.3 *Detection of Local Maxima*

The accuracy of tree detection is known to be sensitive to the settings of algorithms and can vary depending on the forest type and structure. Therefore, we decided to use a simple but widely used algorithm: the local maxima detection with a fixed size window on both CHM and imagery. The small size of the window (1.5 m) ensures to extract as many trees as possible thus, we considered that all visible trees in images and CHM were identified. Hidden trees or other deficiencies in image or CHM production may lead to no or wrong detection. A consequence of the small window size was a high rate of false positives, which were removed by a cleaning step. The false-positive filtering is based on supervised classification with random forest that can easily handle many variables. The variables highlighted came from CHM and imagery.

Omissions result from either no visibility of a tree (e.g. dominated trees, deficiency of the data) or a classification error. The omission rate is low but it increases after the false-positive filtering, due to the classification error. The detection rate of dominant trees was promising whereas forest management focuses on dominant trees: 48.1% of the omitted trees were dominated while only 23.1% were dominant; 35.1% of the dominated were omitted against 10.2% of the dominant, which is encouraging in a forest management perspective, where dominant trees are the more important.

Comparison of the detection results for the different combinations of data showed that the best performance was achieved with the combination of CHM and mosaic (89.2% true positives, 10.8% omission, and 7.8% false positives). Regarding the results and given that the use of individual rectified and ortho-rectified images requires manual selection of the sharpest images from the many available for a plot, this image type may not be suited for local maxima detection. In order to overcome the constraints of the manual handling of selection of individual images, the local maxima detection could be applied to all individual images (rectified or ortho-rectified) related to a plot. The groups of local maxima of each individual images would then merged and post-processed to remove duplicates before false-positive filtering.

Wrong detection (omission or false positives) can be explained partially by the smoothing of the CHM or by the imbrication of neighbour crowns. Crowns are generally wider and less individualized in photogrammetric CHM, and tree tops are better represented in an ALS CHM [87]. The smoothing involved in the photogrammetric process influences the local maxima detection and the computation of metrics such cc_p and volume variables. Compared to ALS data, photogrammetric data tend to smooth the rendering of 3D structure, namely the gaps and spaces between trees [158].

4.5.4 *Forest Attributes Assessment*

Area-based approaches to assess forest attributes are marked by the use of many inter-correlated metrics affecting the variable selection and undermining the model fitting process [28, 23]. This trend is often observed with ALS data, which are intrinsically very rich. Thus, we chose to construct parsimonious models, considering the low number of plots (half of the plots were used for the training of supervised classification) and to avoid over-fitting. Nevertheless, we decided to keep four variables from each of the three types (tree, plot, and mixed) to test the potential and relevance of our data. The variables were generated to be closely linked to the four major forest stand attributes (top height, number of stems, basal area, and volume, scaled per hectare).

The quality of the different forest attribute models is rather promising in terms of the fitting and cross-validation results. Our results are in line (or better) with previous studies carried out with traditional aerial images [18, 62, 135, 143, 116, 156] or with UAS data [87, 123, 140]. Table 4.9 summarizes the results of these studies in terms of %RMSE, working scale and overlap, for purposes of comparison. The contribution of "tree" or "mixed" variables made it possible to improve the models of dendrometric parameters. In our view, the good performance of our results emphasize the relevance of our approach.

Table 4.9: Summary of results (%RMSE) of previous studies on forest attributes assessment with aerial or UAS data.

	Forward/side overlap (%)	**Scale**	**Top height (%)**	**Stem number (%)**	**Basal area (%)**	**Volume (%)**
This study	80/80	plot	5.7	28.3	16.2	20.1
[18]	60/30	stand	8.8	-	14.9	13.1
[62]	70/30	plot	18.2	-	36.2	40.4
[87]	75/75	plot	8.4	-	-	-
[135]	65/30	plot	-	-	35.29	37.9
[143]	70/30	plot	-	-	23.62	24.5
[116]	30/30	plot	10.8	-	28.1	26.8
[123]	95/80	plot	3.5	39.2	15.4	14.5
[140]	80/80	plot	-	-	23.87	26.12
[156]	60/20	plot	14	-	37.7	36.9

The three types of variables (tree, plot and mixed) appear to be determinants for construction of the models, showing the relevance of combining area-based and individual tree approaches. The horizontal structure is described by $\text{Aggreg}_{\text{uas}}$ and CC_{p}, while h_{max} is related to the canopy top as a variable highly correlated with the site index. The amount of biomass is highlighted through $\text{Vol}_{\text{CHM-5}}$ and $V_{16}N_{\text{Trees}}$. This latter metric gives more emphasis to the crown volume. Most of the variables are correlated with the attributes (positive coefficients) except N_{Trees} and $\text{Vol}_{\text{CHM-5}}$, which have negative coefficients for top height and number of stems, respectively.

The involvement of $Aggreg_{uas}$ is particularly relevant in the framework of uneven-aged transformation. A high value of $Aggreg_{uas}$ corresponds to a random spatial arrangement of trees, differing from the regular pattern found in plantations.

4.5.5 *Computation Time*

The present article has not investigated the issue of computing time. But in comparison with previous studies, our methodology being technically straightforward allows an easy implementation for a low computational time. Regarding the various photogrammetric products, the use of rectification instead of true orthorectification should more reduce the computation time. However, overall results were not in favour of rectification. As already outlined above, the step of manual selection of the sharpest image was a drawback for efficiency. In our methodology, the identification of false positives, although largely automatically performed by different criteria of distance and positioning, required the intervention of an operator to ensure the quality of the learning data set for the supervised classification. In a general perspective, the time for data processing is an important topic, especially with a view to promote the operational use of UAS data. The photogrammetric workflow could be time-consuming and requires high-performance equipment. The use of tools as `MICMAC` is complex due to the possibilities of fine tunings and requires the involvement of reference persons with significant technical expertise and/or the development of softwares dedicated to end users. To our best knowledge, [42] create the first open-source GIS application based on `MICMAC`, bringing data generation (point clouds, orthophotographs, DSM) to a wider user community, mainly thanks to the direct integration in a GIS software). The constant evolution of the algorithms, especially in open source softwares, makes it difficult to estimate the impact on computation time. The degree of automation is not really an impediment in a research framework but will become a key issue at the time of a more operational/commercial deployment of the methods.

4.5.6 *General Reflections*

Despite a suboptimal data acquisition, our simple and flexible method has yielded good results and shows great potential for a large panel of applications.

We believe that our methodology could be adapted for other types of data, involving the production of adapted metrics for false-positive filtering and forest attributes assessment. For example, the principle of false-positive filtering could be transposed to variables extracted from ALS data, benefiting from the finer 3D information in the point cloud. However, the fine resolution of UAS data appears to us as determinant for improving dominated tree detection, where lower image resolution would probably be a weakness. Future work should test this approach

on other forest types and conditions. Also other modalities need to be explored using a variable size window or fixed-radius window depending on previous identification of the stand type (e.g. by species, structure, or age). [137] studied the performance of high altitude photogrammetric CHM in detecting individual trees and the inconsistency of detection between different forest types (mature stands, managed forest with varying species). They used watershed segmentation on smoothed CHM produced from high-altitude aerial images to delineate crowns, which were the basis for the reconstruction of height distribution. In this latter study, the effect of CHM preprocessing was also tested by implementing several smoothing filter and regarding the stand type.

In this article, we chose to implement local maxima detection as a robust and versatile method. Another branch of single-tree detection methods focuses on tree crown delineation. Similarly to our study, this kind of approaches brings additional information at tree-level, e.g. related to crown shape or size, which can be aggregated at plot level, as new variables for assessment of forest attributes. For instance, in [6], the integration of variables coming from crown segmentation improved the performance of forest attributes models, especially for volume and basal area. Tree detection from ALS data directly from the point cloud seems a promising alternative method. [146] present a multi-scale dynamic point cloud segmentation dedicated to forest tree extraction from LiDAR data. Raw elevation data are used in place of height normalized data to preserve the crown geometry and apex identification is maximized with the use of absolute maxima instead of the local maxima approach. At the end of the process, each tree is normalized using the elevation of the highest point of the crown as unique ground reference. In their study, [89] implemented a bottom-up method based on both spatial arrangement of points and the intensity. Tree trunks are first extracted based on the intensity and topological relationships between points. The allocation of the other points relied then on a complex procedure involving several thresholds on 2D and 3D distance between points. Unlike [146], this study is characterized by a classical height normalization and a high level of parametrization.

4.6 Conclusions

In this study, we investigated a novel methodology using UAS image-based products to detect individual trees, applied to a large forest area. Our research highlighted the potential use of several photogrammetric products (ortho- or rectified images, DSM, correlation maps) for forest characterization. The core of our work was the use of three types of UAS images to apply an individual tree detection approach combined with a random forest supervised classification to remove false positives. Despite a suboptimal data acquisition, our simple and flexible

method has yielded good results and shows great potential for application.

Our findings demonstrated that UAS images and CHM can be used efficiently to identify individual trees in forest inventory plots (forest attribute assessment up to 83%, true-positive detection of 89%, omission rate of 11%), in the particular context of coniferous stands in transformation from even-aged to uneven-aged stands. These findings add to a growing body of literature on the promising use of UAS data in forestry. We investigated combining the individual tree approach (without crown delineation) and area-based approach for forest attribute assessment with UAS data. Considering the particular silvicultural context of the present study area, multi-temporal data acquisition would be relevant for monitoring the uneven-aged transformation.

Some questions remain regarding the adequate processing of drone imagery for utilization in forestry. Some standard of traditional photogrammetry still needs to be revisited in order to better fit both characteristics of low-altitude and low-quality imagery and the specificity of the forest structure. Indeed, the forest canopy surface is a complex layer that is difficult to model by means of a digital surface model or orthophotomosaic. Obviously, before accurate measurements can be made based on aerial images, raw images have to be corrected of their distortions [126]. Relief-based and lens-based distortion removal is performed by all modern photogrammetric software, but changes in the settings of the processing chain can produce considerably different results.

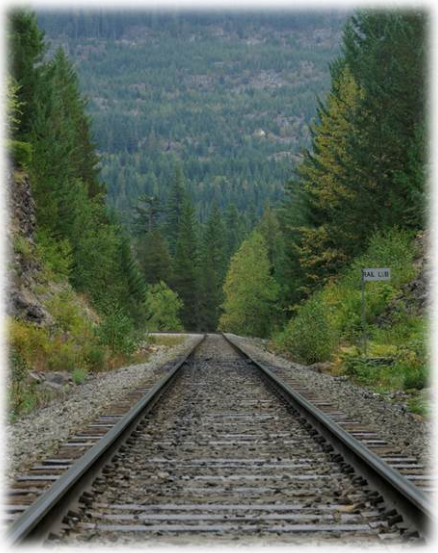
Further challenges arise regarding the potential and advantages of UAS data. As drones permit us to control acquisition costs and to acquire very high resolution images at the moment chosen by the user, they could potentially be used for interesting applications such as species discrimination, determination of stand types, identification of areas with tree regeneration, or monitoring silvicultural transformation process, thus improving forest inventory. In the era of Big Data, the real challenge lies perhaps more in the exploitation of the data to identify the gain to be derived from the drones for the managers of natural resources than in the pure technical issues.

Chapter 5

DISCUSSION & CONCLUSION

Chapter 5

Discussion & Conclusion



"Toute science crée une nouvelle ignorance."

Henri Michaux

5.1 In a nutshell

The main purpose of this thesis was the implementation of different types of 3D remote sensing data at contrasted working scales and to determine how it could support forest management in Wallonia. ALS and UAS data are both assets to characterize forest resources. In practical terms, chapter 2 analyzed the potential of RFI data to construct allometric equations and how ALS data could be used in combination with RFI data to assess forest attributes. Chapter 3 focused on the potential of ALS to map forest canopy gaps in uneven-aged broadleaved stands. Chapter 4 examined the performance of UAS data to assess dendrometric attributes in coniferous stands, using individual tree detection and comparing several photogrammetric outputs. This final chapter summarizes the main achievements of the chapters 2 to 4. Figure 5.1 synthesizes the main strengths and weaknesses of these three chapters and Table 5.1 presents the data used in the thesis. General considerations are discussed and research perspectives are also highlighted.

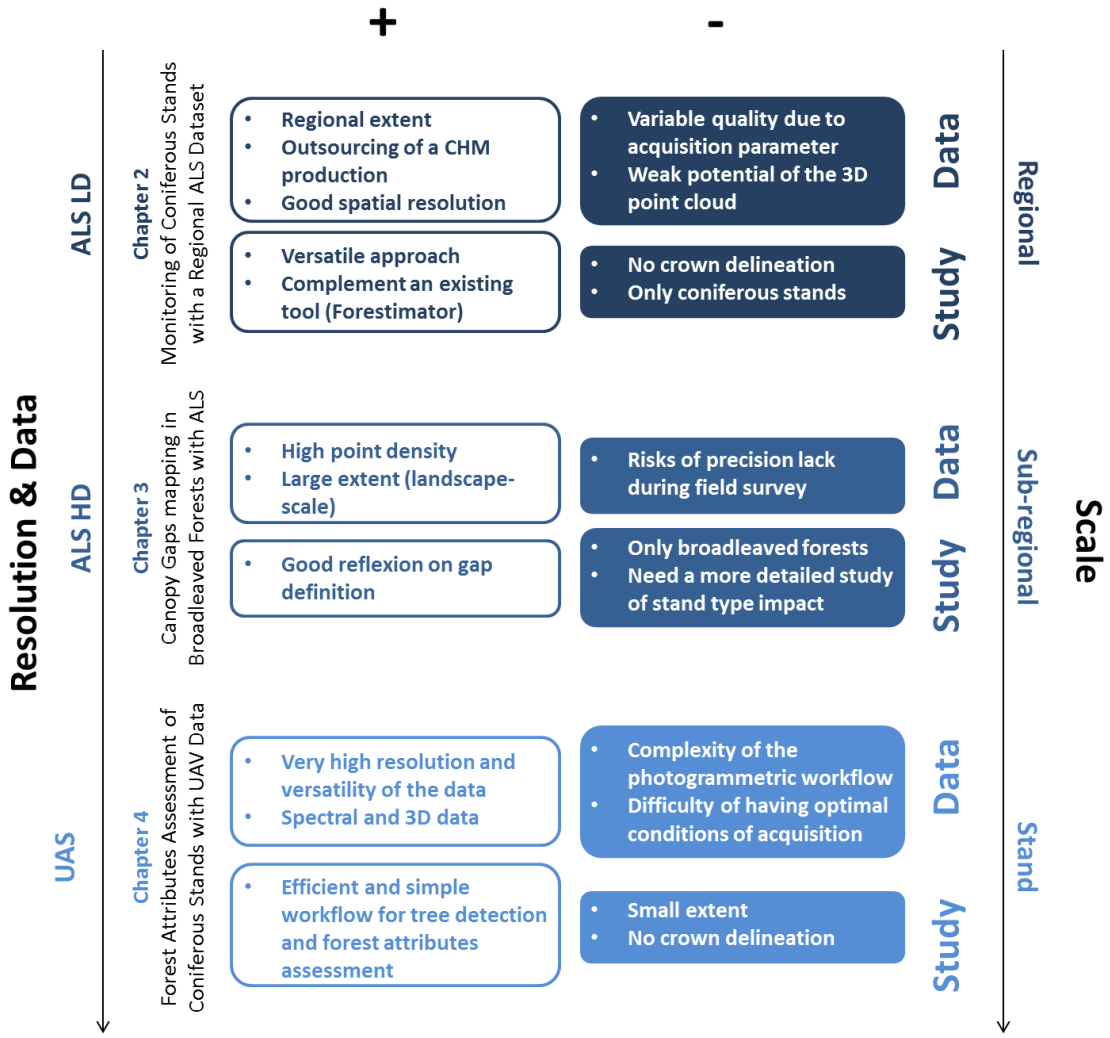


Figure 5.1: Synthesis of the main strengths and weaknesses of the data and methods used in the studies presented in chapters 2 to 4, in relation to the scale of work and the type of data

Table 5.1: Summary of the remote sensing data implemented in this thesis and their main characteristics and use.

	Chapter 2	Chapter 3	Chapter 4
Type	ALS low density	ALS high density	UAV images & CHM
Resolution	1 m	50 cm	9 - 35 cm
Extent	17000 km^2	180 km^2	30 km^2
Date	Dec. 2012 - Apr. 2013	Mar. & Jul. 2011	Oct. 2014
Owner	SPW	SPW	Gembloux Agro-Bio Tech
Use	Tree detection & modelling	Mapping	Tree detection & modelling

5.1.1 *How to combine ALS and RFI data to assess forest attributes of coniferous stands at regional-scale?*

This question was dealt by the implementation of a H-C allometry based on RFI data and its combination with low density ALS data for the monitoring of coniferous stands. We particularly aimed to test other uses of the data available in *Forestimator* plug-in to enhance this tool with new applications. The goal was to take advantage of an existing tool and methodology that was originally only dedicated to the top height assessment.

The first step of our approach was the construction of a H-C allometry based on regional data and enhanced with environmental and structural variables. The purpose of this allometry was to be used with local maxima detected on an ALS CHM, to compute attributes at tree-level. Six types of models were tested. They were species-specific or not. The allometries are characterized by a %RMSE ranging from 10.7% to 14.5%, with a variable degree of complexity (Tables 2.6, 2.7 & 2.8). The integration of variables characterizing the plot (slope, altitude, stand density or development stage) had improved the quality of the results. Especially, the introduction of the stand age in the allometries allowed to improve the quality of the models. This data has the advantage of being easy to obtain, particularly in the case of coniferous stands with a known planting date.

Then forest attributes models at plot-scale were implemented. They were functions of a limited number of predictor to avoid model over-fitting. We compared two approaches of forest attributes assessment: ABA only and ABA combined with ITD. Both models were fitted with variables computed from the CHM only for ABA and for CHM and local maxima extracted from CHM for ITD. The H-C allometry was also applied on the detected trees to enrich the set of variables for ITD. The models provided R^2 ranging from 33% to 73% for ABA-based models and for 39% to 84% ABA/ITD-based models. The weakest model is the one predicting the number of stem that particularly provide poor predictions of dense stands. The best model is the one predicting the quadratic mean of circumference (R^2 = 84% and %RMSE = 9.8%). The integration of variables coming from ITD and allometric model allowed to improve the quality of fitting for the four forest attributes.

In this study, we decided to develop allometries to assess tree girth from height information coming from remote sensing data (instead of field data). Trying to build a synergy between RFI and ALS data has further strengthened our confidence that an iterative and stratified approach is necessary for a reliable assessment of forest attributes. Especially, a reflexion about the methodology of the field data collection should be undertook to adapt the in-situ data to remote sensing considerations.

5.1.2 *How reliable is a canopy gap mapping in broadleaved uneven-aged stands using ALS data?*

To address this question, chapter 3 has investigated canopy gap mapping through a comparison of four types of methods (simple thresholding, multiple thresholding, per-pixel supervised classification and per-object supervised classification) applied to leaf-off and leaf-on high density ALS datasets taking into account the mapping accuracy in terms of both gap detection accuracy and a detailed geometric accuracy assessment in the field which was missing in previous studies, especially in broad-leaved forest.

It first appeared that the definition a gap was ultimately a key issue to consider. We began to construct a clear and simple definition to identify gaps in the field. Throughout the study, we defined gaps as openings in the canopy with a minimum area of 50 m^2, a minimum width of 2 m and a maximum vegetation height of 3 m. The vegetation level value of 3 m is considered to indicate well-established regeneration [81] and is a critical height for the survival of regeneration, especially against ungulate predation.

Three methods were identified as producing the best detection results: per-pixel supervised classification of leaf-on data, a multiple thresholding of leaf-on data (Combination1) and a simple thresholding of leaf-off data (CPI-60%). The highest global accuracy achieved was 82%. In general, better results were obtained from most methods in leaf-off conditions (with the exception of multiple thresholding and per-object classification) regarding gap detection and geometry. The high point densities of the ALS datasets used in this study allowed good characterization of tree crowns, even in leaf-off conditions, and the absence of leaves decreases occlusions and improves representation of the ground. In deciduous forest, leaf-off data may therefore be generally preferable to leaf-on data unless point density is low.

Despite its relative technical simplicity, thresholding is not clearly surpassed by supervised classification. But the results of this study showed that the very common CHM thresholding was outperformed by other methods (see Table 3.3). However, the creation of a CHM has other advantages for a forest manager, as it allows other applications also shown in this thesis.

Our results suggested the importance of identifying an optimum mapping method regarding the stand type. It is likely that neither the field mapping or the ALS-derived gaps fully represent the true gap geometry. Whilst gaps have been field surveyed with a good degree of accuracy, subtle details of gap boundaries may be better represented by ALS, resulting in more complex shapes.

5.1.3 *What is the potential of UAV imagery to characterize coniferous stands at local-scale?*

To answer this question, Chapter 4 investigated a methodology using UAS image-based products to identify trees. The use and relevance of several photogrammetric products (ortho- or rectified images, DSM, correlation maps) was tested to detect individual trees as a basis for forest characterization. We investigated combining the ITD approach and ABA for forest attributes assessment. The geometry of mosaics suffered from heterogeneity in the image exposures because of illumination differences during the flights. Also, significant variations in relief at the study site caused lack of overlap between images and a variable resolution of the photogrammetric products.

The local maxima detection method was implemented with a small size of the window (1.5 m) to ensure extracting as many trees as possible. Hidden trees or other deficiencies in image or CHM production may lead to missing or wrong detection. A consequence of the small window size was a high rate of false positives, which were removed by a supervised classification with random forest. The best performance of tree detection was achieved with the combination of CHM and mosaic (89.2% true positives, 10.8% omission, and 7.8% false positives). The detection rate of dominant trees was promising: 48,1% of the omitted trees were dominated while only 23.1% were dominant; 35,1% of the dominated were omitted against 10.2% of the dominant, which is encouraging in a forest management perspective, where dominant trees are often the most important ones. Omissions resulted from either no visibility of a tree or a classification error.

About forest attributes, three types of variables (tree, plot, and mixed) were generated to be closely linked to the four major forest attributes (top height, number of stem, basal area & volume). Considering the low number of plots, we chose to construct parsimonious models and we decided to keep only four variables of each type. The three types of variables appear to be determinants for model fitting, showing the relevance of combining area-based and individual tree approaches. The contribution of "tree" or mixed "tree" & "plot" variables made it possible to improve the models of dendrometric parameters.

Our good results demonstrated that UAS data can be used efficiently to identify trees, in the particular context of coniferous stands in transformation from even-aged to uneven-aged stands and despite a sub-optimal acquisition. The variation of illumination of our images did not impact the tree detection. Our choice of implementing apex detection instead of crown delineation minimizes probably the impact of shadows. Besides, our methodology being technically straightforward allows an easy implementation with a low computational time, which is an important issue to promote the operational use of UAS data.

5.2 3D Remote Sensing for Supporting Forestry in Wallonia

Wallonia is a medium-sized region whose forest resource, representing one third of the territory, is characterized by a large fragmentation, an anthropized structure, a great diversity of forest species and different management styles. Its extent facilitates the acquisition or availability of complete coverage in a short time, which is an opportunity to promote the development of remote sensing. The appropriation of remote sensing tools by foresters in an operational context must begin with simple but reliable approaches. Accuracy and reliability are the best guarantors of the relevance of integrating remote sensing into the forest manager's toolbox. However, remote sensing is not the panacea for characterization of the resource. Each data has its limitations, depending on its specific characteristics (Figure 5.2). For a given objective, it is necessary to think about the appropriate data and methods.

The forest, its resources and the various stakeholders are increasingly solicited. Taking advantage of all available tools to describe the forest resource is essential to better provide guidance for data collection and various actions in the field. The underlying goal is not to contemplate the forest only through images. But the noble purpose should be to take advantage of the remote sensing to go further, save time and improve forest monitoring efficiency. Given the diversity of available data and in particular 3D data, look at the construction of a typology seems relevant and necessary. Two ways of thinking are possible: either in terms of the availability of initial information, the forest and its limits, or in terms of users' needs. From the users' point of view, considerations of silviculture and regeneration concern a local scale, requiring a fairly fine level of detail, along with very high spatial resolution and regular updates for monitoring. At the scale of a forest estate, the needs are in the context of forest management, whereas the evaluation of the forest resource and its evolution at the regional level is intended to guide forest policies. These different spatial scales go together with different time scales.

Remote sensing researches must therefore focus on developing methods of producing information with added value. These methods must be robust, repeatable, adapted to different scales and, as far as possible, based on open access or low cost data. Then, integration of different approaches will meet different needs. That was the guiding principle behind the three working scales of this thesis, considered and conceived as complementary. Chapters 2 and 4 implement relatively similar approaches, albeit with their own data specificities (Table 5.2). Some methods are adaptable from one work scale to another.

Table 5.2: Comparison of the study area, data and model used in chapters 2 and 4. These two chapters develop similar approaches for estimating forest attributes but from different data and scales. The top height model performance values come from [38]; NS is for Norway spruce and DF for Douglas-fir.

	Chapter 2	**Chapter 4**
Extent	17000 km^2	30 km^2
Plot number	179	72
Resolution	1 m	9 - 35 cm
Data range		
Top height (m)	0 - 28.3	16.7 - 27.9
Tree number (/ha)	89 - 2222	147 - 945
Basal area (m^2/ha)	14.1 - 73.3	21.3 - 59.2
Volume (m^3/ha)	158.2 - 1109.2	183.9 - 844.4
R^2		
Top height	95 %/97 % (NS/DF)	86 %
Tree number (/ha)	58 %	81 %
Basal area (m^2/ha)	43 %	76 %
Volume (m^3/ha)	61 %	73 %
RMSE		
Top height (m)	1.04/0.93 (NS/DF)	1.39
Tree number (/ha)	207	71.2
Basal area (m^2/ha)	7.1	5.13
Volume (m^3/ha)	93.2	81.4

As seen in chapter 2, a quantitative approach is possible at regional level. Quantify the forest resource must be considered by stand type but it is important to be aware that the precision obtained will be lower than with higher resolution data at local scale. The RFI is a very complete tool to draw up a diagnosis of the Walloon forest over a ten years horizon. This type of regional or national inventory is worth enhancing with remote sensing approaches, as seen in the Scandinavian countries or in North America. The richness and diversity of the RFI data makes it a very valuable source of reference data for the development of methods for estimating forest resources by remote sensing. But the particular case of concentric plots is sometimes difficult to handle when looking at the detection of individual trees because it complicates synchronization with remote sensing data. Besides, a detection error on a small area will cause great differences once scaled par hectare. However, it is important to realize that we move towards more hybrid methods where remote sensing is a source of data for the characterization of the resource as well as the RFI. Just as the height becomes as easy to measure by remote sensing as the circumference are in the field. We are witnessing a more general paradigm shift where remote sensing is a source of information for the RFI and vice versa.

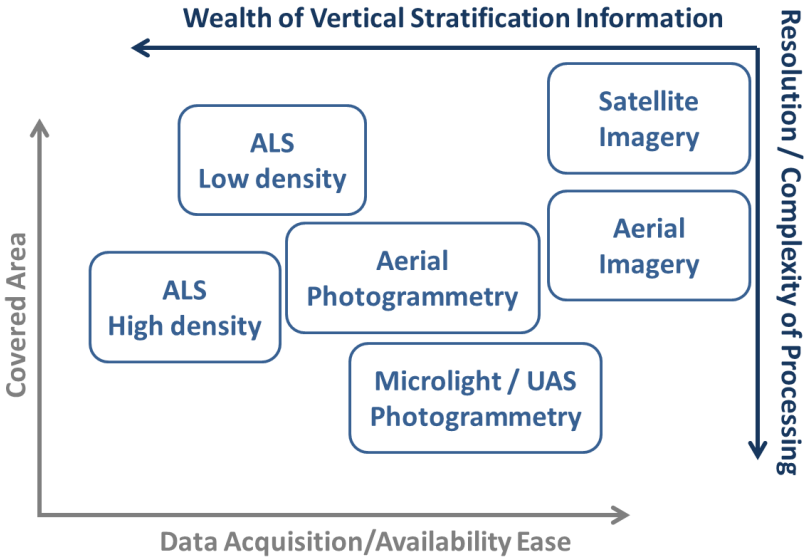


Figure 5.2: Relevance, scale, availability of the data.

In line with the close-to-nature forestry, the transformation of even-aged to uneven-aged stands is growing, especially in the Walloon coniferous forests. This trend will logically give more emphasis to individual tree and require a new kind of data which are relevant at tree-level. In this context of precision forestry, the advent of civil drones has opened a new era but their use is not obvious. Their apparent ease of use goes together with constraints of data processing. In addition, the data acquisition phase can be tricky and flight conditions are difficult to optimize and adding the issues related to the legislation.

In this thesis, we have always worked with a priori information on the delimitation of forest areas. It is important to evaluate the quality of this delimitation because it is often tainted with error and inaccuracy. Having reliable cartographic information identifying forest areas or stand types would make it possible to target the tools used according to a forest stratification. It could be the starting point for regional and integrated approach. Indeed, it seems reasonable to think that the production of a regional forest map is an essential prerequisite, regardless of the type of owner. The advent of Sentinel-2 data in particular and the availability of regional datasets acquired by the public administration must be the basis of such kind of work. Different methodological choices can be envisaged, ranging from per-pixel to per-object classification, combining the spectral information of Sentinel-2 images with 3D information from CHM ALS or photogrammetric data produced from regional aerial images. The Sentinel-2 images have a resolution of 10 to 20 m while the regional aerial data are of the order of one meter and therefore offer a higher level of accuracy. The notion of hierarchy or stratification of applications and data implemented is therefore important. Regional aerial images have the disadvantage of having a high radiometric heterogeneity, which compli-

cates their use and pleads in favor of satellite images. On the other hand, the 3D and multi-temporal information could make it possible to create a forest map with a follow-up of harvests and recruitments. The identification of the different forest types within the forest matrix, in particular by the species or structures discrimination, is the natural progression in the characterization of the resource. This second phase of the resource description process would make it possible to deploy specific approaches related to certain types of stands or stages of development (Figure 5.3). Then, a reflection of the needs related to the specificities of the stands to which one is interested must be led.

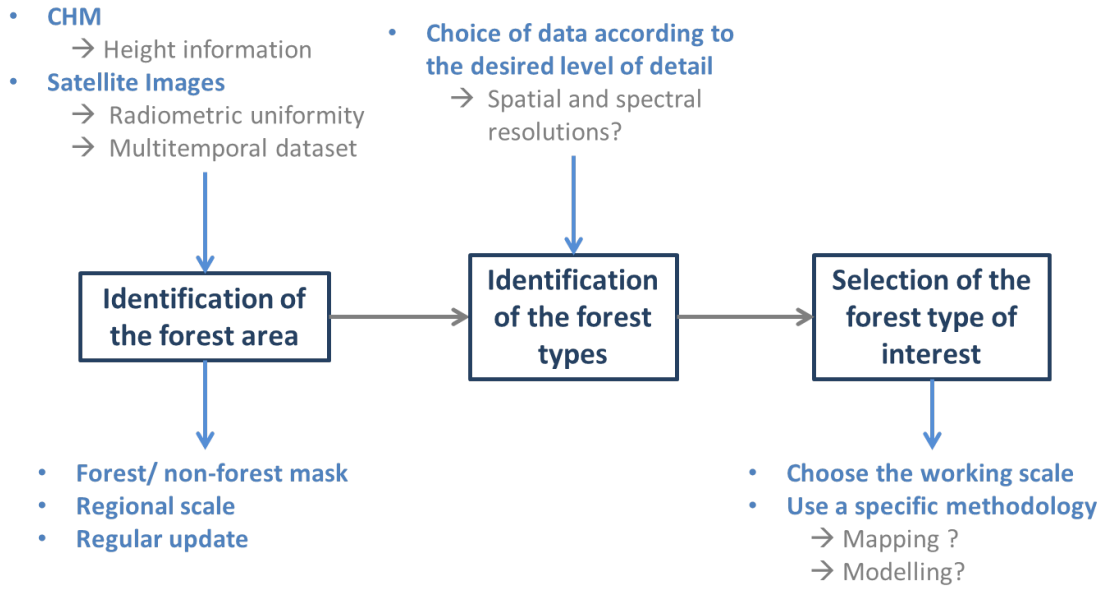


Figure 5.3: Towards an integrated monitoring

As a complement to mapping approach, the generalization of the estimation of forest attributes at larger scales can be addressed. In this study, forest attributes were assessed at the plot scale (about 10 ares). The next methodological step could be wall-to-wall estimation of forest attributes, according to a grid whose pixel size is of the same order of magnitude as inventory plots.

5.3 Perspectives

5.3.1 *General trends*

Remote sensing is an evolving discipline and new trends are emerging regularly. The democratization of the very high resolution and high density data, will allow to gradually eliminate the inherent limitations of the Earth observation data. Regional aerial image coverage remains dependent on strong constraints that slow down data access and the spatial resolution of satellite images (10-20

m for Sentinel-2 images) is too low for some analyzes; even if Sentinel-2 images are of great interest to support species discrimination and wall-to-wall forest map production to complete quantitative approaches.

Thanks to technical developments, high density or very high resolution data will progressively become the norm. For example, high density ALS datasets like those in chapter 3 should be available at regional-scale. In this context, we can therefore question the relevance of drone data, considering, for example, the small size of covered areas. This question is reinforced by the rise of the microlights which represent a real opportunity for sub-regional acquisitions, covering areas up to several thousands of hectares. Until now, drones do not have sufficient autonomy to cover such surfaces. The minimum altitude of microlights flight being higher than for a drone, for the same sensor, the resolution will be slightly lower but similar. This relative decrease in the resolution is compensated by a much greater autonomy of flight and a much larger covered surface as well as by a significant reduction of the weight constraints making it possible to load several sensors simultaneously.

Acquisition flexibility remains one of the main assets of drones (apart from the legal constraints that are important). The mobilization of microlight or the order of aerial data (images or ALS) require a huge logistic and planning. From our point of view, the use of drones keeps its meaning in local applications in order to answer very specific issues requiring the acquisition of data very quickly at very high resolution and when the combination of spectral data and 3D are necessary.

5.3.2 *The challenge of data fusion*

The principle of data fusion is to combine multiple sources of data to improve the performance and acquire an enhanced dataset for new applications [168]. This concept can correspond to several realities, such as the combination of lower resolution multispectral data with higher resolution but panchromatic data. Beyond combining different spatial resolutions, the objective is increasingly to take advantage of complementary spectral or temporal characteristics. With the rapid development of new sensors and platforms, data fusion is a growing discipline, especially in the era of big data.

Due to the lack of ground elevation information (*e.g.* a DTM), photogrammetric 3D data needs to be supplemented with another data to allow the computation of height. For example, a photogrammetric DSM has been combined with an ALS DTM in [87] to estimate heights of forest plots with UAS images or in [100] for the characterization of riparian forest on a regional scale through a wall-to-wall monitoring with regional aerial data. In most cases, and this is especially true in forest, the ground topography is considered constant over time. The advantage

of this type of approach lies in the regular availability of aerial image datasets or in the relative ease of acquisition of UAS or microlight data. For example in Wallonia, regular acquisitions of aerial images by the regional administration are a great opportunity since this type of images can be used in many ways by several departments, allowing a reduction of acquisition costs.

The relevance of LiDAR data for quantification and characterization of the forest resource is well established, especially considering the numerous studies conducted on the use of small footprint ALS. However, forest ecosystems are complex, and ALS data also has to be enhanced with other types of data such as multispectral, hyperspectral or RaDAR data. For example, [74], the combination of Worldview-2 multispectral images and high density LiDAR data to distinguish two species of Pinus was tested and compared with the two individual datasets. Improved classification results were obtained by integrating the variables of crown height and structure from LiDAR and vegetation and texture indices from satellite imagery. In [34], tree crowns were delineated with ALS point cloud and species were identified based on hyperspectral data in Italian forest. Also, [8] showed promising results in terms of detailed characterization of complex ecosystems, such as the species composition in Amazonian rain forests. In addition, the combination of RaDAR and optical data has more and more sense and is going to expand with the availability of Sentinel-1 and Sentinel-2 data. The tools and applications developed in the context of the Copernicus program, coupled with revisit time, resolution and free access to data are all factors of success for the development of such data and their integration in many Earth observation projects.

Another way to combine data is to implement multi-temporal approaches, either to identify changes or to take advantage of the seasonality for example. In their study, [125] used diachronic photogrammetric CHM to quantify top height and volume changes after storms in broadleaved forests. Similarly, pre- and post-fire LiDAR datasets were used by [17] to estimate the proportion of fallen trees due to root damages. The study was based on the used of statistical metrics showing a strong relationship with structural variables. To analyze the advantage of having seasonal spectral variations, [19] investigated the classification potential of Sentinel-2 for forest lands mapping. Two cloud-free S2 scenes (August 2015 and May 2016) were used and a series of spectral indices were computed. First a forest/non-forest map of the Belgian Ardenne ecoregion was produced. Secondly, tree species discrimination was tested. These first investigations on species discrimination were encouraging with an OA of 88.9%, even with the simplified per-pixel approach.

These studies, among many others, highlight the importance of building bridges between applications and different types of data, while technical limitations are becoming less of a problem. The difficulty is to avoid being overwhelmed by the

multitude of data available or the relative ease of collecting remote sensing data, but first of all to think about the relevance of their use. Especially, more emphasis need to be given to the effect of different forest types and conditions.

In the same vein, data fusion can refer to the use of other types of geographical data (not necessarily derived from remote sensing): altitude, slope, aspect, forest stands delineation, soil, moisture or diverse environmental factors... These data combined with information from the "Fichier Écologique des Essences" [29] and forest attributes derived from remote sensing data suggest future considerations about integrated decision support tools. In the light of the above, the main developments mentioned only make sense if a joint reflection between the researchers and the forest administration and even the forest industry is built concretely. Beyond the funding of research projects, a real commitment is needed to modernizing IT and mapping architecture that are still lagging behind.

5.3.3 *Examples of specific research perspectives*

Characterization & dynamics of regeneration inside a gap

While the third chapter of this thesis focused on canopy gap mapping with ALS, studying regeneration in canopy gaps with remote sensing data is still a challenge. The ability of ALS is largely unexplored in seedling stands; further work is needed about the assessment of regeneration potential in canopy gaps, especially the development of indicators of the presence/abundance of regeneration. The combination of 3D and very high resolution spectral data could offer the opportunity to focus on two major themes: modeling the availability of light and the characterization & dynamics of regeneration.

Chapter 3 has demonstrated the ability of ALS data to describe the shape and size of gaps with several attributes such as area, gap shape complexity index, main direction,... In addition, ALS offers the possibility of obtaining a good quality DTM. Having such data enable the quantification, at the gap level, of biotic factors that may affect regeneration process (elevation, slope, aspect,...). The use of other topographic index to quantify the relief variation could be considered.

A first methodological approach would be to evaluate the potential of a gap to accommodate regeneration by evaluating a set of indicators related to the installation of regeneration in and around the gap. Starting from a gap mapping, a descriptive phase based on remote sensing data should be set up to evaluate light, nutrient and moisture conditions. The description of the gap's attributes could be done for each gap, or by considering a spatial division of each gap (according to a grid), enabling the spatialization of variables whose values could changes according to the position in the gap. The vegetation can be described at several levels by using the CHM to extract the height of vegetation or detect individual trees: cover

and height of vegetation present inside the gap, description of isolated trees within the gap (height, position of the tree relative to the borders, crown cover, size,...), characterization of the stand around the gap (height variation, distribution, presence of large or tall tree, species). A tricky challenge would be the discrimination of herbaceous vegetation from woody or semi-woody vegetation in order to assess the regeneration already in place.

The study of light availability is particularly relevant when considering regeneration issues. Having 3D data, such as ALS point clouds, offers interesting perspectives for analyzing the porosity of the canopy, understanding the forest structure and studying the pulse trajectories. The density of the point cloud would be a particularly important parameter to take into account. The integration of the terrestrial LiDAR could mitigate the use of the only aerial data, but this seems not very feasible on very large areas.

Construction of an objective typology of the forest structure with TLS & ALS

As ALS, Terrestrial Laser Scanning (TLS) consists in 3D point clouds but acquired from the ground in a static way. TLS provides detailed and complex data to finely visualize trees and crowns and the overall forest structure. TLS is intended for acquisition of data at local-scale and is particularly used for research and development as a new tool for forest inventory [110, 142, 36]. By their differences (point density, area of acquisition, data management) and their similarities (ability to describe the 3D structure of the forest), ALS and TLS are very complementary. From a research perspective, TLS can be considered as a good substitute to classical forest inventory to collect field data.

TLS data offers the possibility to extract tree maps and individual tree measures but also many other variables computed from the 3D point clouds, outperforming the classical field data method. Thanks to its richness of information, TLS allows us to capture the forest structure and to rethink it through a typology based on objective and quantitative indicators. A typology is a conceptual framework useful to synthesize a very complex information about forest structure. Its quantitative description has became an important element of modern sustainable forest management. Silvicultural practices can modify structure as well as an existing structure can guide silvicultural practices.

Attributes related to the structure can be computed in two ways from TLS data. First, in an idea close to the use of classical field data, variables can be computed from the tree map. These variables are spatial or non-spatial indicators. This quantification approach of forest structure are explained in more details in [117, 119, 118, 151]. Secondly, proxies of the structure can be extracted from the point cloud. These latter are remote sensing metrics (raster or point clouds) to be linked to a strong forest or environmental meaning. As example, several

proxies [117, 119, 118] aiming to quantify the structure could be the basis for the construction of an innovative typology by the use of data mining methods: indices of spatial heterogeneity based on tree position (agregation, mean of angles, mingling,...), indices based on tree attributes variation (Gini, parameters of height and DBH distributions, dominance,...), point cloud variables (computed from the 3D distribution of the points), raster variables (CHM, rumple, HSCOI, landscape metrics).

5.4 Conclusion

The constant evolution of remote sensing data is an opportunity for natural resources managers and especially for foresters. The Walloon forest has high economic, ecological and social importance to the Region. The specificity of the owners, the diversity of the stand types and the imbrication of the forest in the landscape, make its characterization and quantification important for a sustainable and long-term management.

Considering the interest of 3D data and the challenge of a regional characterization, we took into account the use of this type of data at the regional scale in interaction with the wealth of the RFI. The integration of 3D from low-density ALS has shown its potential for the quantification of forest attributes thanks to a regional allometry built on the basis of the RFI. Close-to-nature forestry, favoring natural regeneration, is a topic of major concern for foresters. The presence of canopy gaps is therefore an essential information. In consequence, we focused on a robust mapping of gaps in broadleaved forests from high-density ALS data. Data acquired by drone has shown their effectiveness for a fine tree detection (dominant, co-dominated, dominated) in coniferous stands stands, as a basis for the quantification of forest attributes.

The main purpose of this thesis was the implementation of different types and resolution of remote sensing data considering contrasted working scales and to determine how it could support forest management in Wallonia. However, caution must be taken because not all data are suitable for all uses and the prospects are numerous. It is necessary to identify the needs or challenges to be addressed and the methodological constraints which are related to. This is where the work and expertise of the forester makes more sense than ever. The meaning of this thesis was not to replace field work. It was rather to test tools to enlighten it by useful methods allowing to free up time so that the forester's expertise can be expressed where it is most useful. As a conclusion, we argue the importance of a strong synergy between the different sources of data to enable remote sensing to deploy its everyday potential as an indispensable support to forest managers.

Bibliography

[1] J. S. Aber, I. Marzolff, and J. Ries. *Small-format aerial photography: Principles, techniques and geoscience applications*. Elsevier, 2010.

[2] O. S. Ahmed, A. Shemrock, D. Chabot, C. Dillon, G. Williams, R. Wasson, and S. E. Franklin. Hierarchical land cover and vegetation classification using multispectral data acquired from an unmanned aerial vehicle. *International Journal of Remote Sensing*, pages 1–16, 2017.

[3] M. Alderweireld, G. Ligot, N. Latte, and H. Claessens. Le chêne en forêt ardennaise, un atout à préserver. *Forêt Wallonne*, 109:10–24, 2010.

[4] M. Alderweireld, J. Rondeux, N. Latte, J. Hébert, and H. Lecomte. Belgium (Wallonia). In *National Forest Inventories*, pages 159–179. Springer, 2016.

[5] K. Anderson and K. J. Gaston. Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Frontiers in Ecology and the Environment*, 11(3):138–146, 2013.

[6] A. C. André, J.-P. Renaud, C. Véga, A. Munoz, J. Bock, and L. Saint-André. Apport de variables issues de la segmentation d'arbres sur données lidar aéroporté pour l'estimation des variables dendrométriques de placettes forestières. *Revue Française de Photogrammétrie et Télédétection*, 211–212:53–61, 2015.

[7] G. P. Asner, J. R. Kellner, T. Kennedy-Bowdoin, D. E. Knapp, C. Anderson, and R. E. Martin. Forest Canopy Gap Distributions in the Southern Peruvian Amazon. *PloS one*, 8(4):e60875, 2013.

[8] G. P. Asner, D. E. Knapp, J. Boardman, R. O. Green, T. Kennedy-Bowdoin, M. Eastwood, R. E. Martin, C. Anderson, and C. B. Field. Carnegie Airborne Observatory-2: Increasing science data dimensionality via high-fidelity multi-sensor fusion. *Remote Sensing of Environment*, 124:454–465, 2012.

[9] P. Axelsson. Processing of laser scanner data: algorithms and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2):138–147, 1999.

[10] M. Baatz, C. Hoffmann, and G. Willhauck. Progressing from object-based to object-oriented image analysis. In *Object-Based Image Analysis*, pages 29–42. Springer, 2008.

[11] M. Baatz and A. Schäpe. Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation. In J. Strobl, T. Blaschke, and G. Griesebner, editors, *Angewandte Geographische Informationsverarbeitung XII*, pages 12–23. Wichmann, Heidelberg, Germany, 2000.

[12] E. Baltsavias, A. Gruen, H. Eisenbeiss, L. Zhang, and L. T. Waser. High-quality image matching and automated generation of 3d tree models. *International Journal of Remote Sensing*, 29(5):1243–1259, 2008.

[13] U. C. Benz, P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of photogrammetry and remote sensing*, 58(3):239–258, 2004.

[14] H. Betts, L. Brown, and G. Stewart. Forest canopy gap detection and characterisation by the use of high–resolution digital elevation models. *New Zealand Journal of Ecology*, 29(1):95–103, 2005.

[15] G. A. Blackburn, Z. Abd Latif, and D. S. Boyd. Forest disturbance and regeneration: a mosaic of discrete gap dynamics and open matrix regimes? *Journal of Vegetation Science*, 25(6):1341–1354, 2014.

[16] G. A. Blackburn and E. Milton. Filling the gaps: remote sensing meets woodland ecology. *Global Ecology and Biogeography Letters*, 5(4/5):175–191, 1996.

[17] I. Bohlin, H. Olsson, J. Bohlin, and A. Granström. Quantifying post-fire fallen trees using multi-temporal lidar. *International Journal of Applied Earth Observation and Geoinformation*, 63:186–195, 2017.

[18] J. Bohlin, J. Wallerman, and J. E. Fransson. Forest variable estimation using photogrammetric matching of digital aerial images in combination with a high-resolution DEM. *Scandinavian Journal of Forest Research*, 27(7):692–699, 2012.

[19] C. Bolyn, A. Michez, P. Gaucher, P. Lejeune, and S. Bonnet. Regional mapping of forest and forest types by supervised per-pixel classification of sentinel-2 imagery. *Biotechnologie, Agronomie, Société et Environnement*, submitted, 2018.

[20] S. Bonnet. Un modèle numérique de canopée pour l'estimation de la hauteur dominante des peuplements résineux en Région wallonne. *Forêt Wallonne*, 98:53–59, 2009.

[21] S. Bonnet, J. Lisein, and P. Lejeune. Comparison of UAS photogrammetric products for tree detection and characterization of coniferous stands. *International Journal of Remote Sensing*, 38(19):5310–5337, 2017.

[22] S. Bonnet, F. Toromanoff, S. Bauwens, A. Michez, L. Dedry, and P. Lejeune. Principes de base de la télédétection et ses potentialités comme outil de caractérisation de la ressource forestière–II. LiDAR aérien. *Forêt Wallonne*, (124):28–41, 2013.

[23] M. Bouvier, S. Durrieu, R. A. Fournier, and J.-P. Renaud. Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. *Remote Sensing of Environment*, 156:322–334, 2015.

[24] D. S. Boyd, R. A. Hill, C. Hopkinson, and T. R. Baker. Landscape–scale forest disturbance regimes in southern peruvian amazonia. *Ecological Applications*, 23(7):1588–1602, 2013.

[25] L. Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[26] N. V. Brokaw. The definition of treefall gap and its effect on measures of forest dynamics. *Biotropica*, 14(2):158–160, 1982.

[27] M. Bruciamacchie, G. Grandjean, and F. Jacobée. Installation de régénérations feuillues dans de petites trouées en peuplements irréguliers. *Revue Forestière Française*, 6:639–652, 1994.

[28] Q. Chen. LIDAR remote sensing of vegetation biomass. *Remote sensing of natural resources*, 399:399–420, 2013.

[29] H. Claessens, E. Bifolchi, S. Cordier, A. Desjonquères, F. Ridremont, S. Bythell, A. De Bont, S. Iboukassene, C. Vincke, and Q. Ponette. Le nouveau fichier écologique des essences. pourquoi et comment? *Forêt Wallonne*, (129):60–70, 2014.

[30] N. Clinton, A. Holt, J. Scarborough, L. Yan, and P. Gong. Accuracy assessment measures for object–based image segmentation goodness. *Photogrammetric engineering and remote sensing*, 76(3):289–299, 2010.

[31] I. Colomina and P. Molina. Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 92:79–97, June 2014.

[32] P. Dagnelie, R. Palm, J. Rondeux, and A. Thill. *Tables de cubage des arbres et des peuplements forestiers*. Les Presses agronomiques de Gembloux, 1999.

[33] G. C. Daily, S. Alexander, P. R. Ehrlich, L. Goulder, J. Lubchenco, P. A. Matson, H. A. Mooney, S. Postel, S. H. Schneider, D. Tilman, and G. M. Woodwell. *Ecosystem Services: Benefits Supplied to Human Societies by Natural Ecosystems*, volume 2. Ecological Society of America, 1997.

[34] M. Dalponte and D. A. Coomes. Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. *Methods in ecology and evolution*, 7(10):1236–1245, 2016.

[35] J. P. Dandois, M. Olano, and E. C. Ellis. Optimal altitude, overlap, and weather conditions for computer vision UAV estimates of forest structure. *Remote Sensing*, 7(10):13895–13920, 2015.

[36] M. Dassot, T. Constant, and M. Fournier. The use of terrestrial LiDAR technology in forest science: application fields, benefits and challenges. *Annals of Forest Science*, 68(5):959–974, 2011.

[37] O. De Thier, J. Lisein, and P. Lejeune. IFA: un logiciel simple pour la réalisation et le traitement d'inventaires forestiers d'aménagement. *Forêt Wallonne*, (129):44–45, 2014.

[38] L. Dedry, O. De Thier, J. Perin, A. Michez, S. Bonnet, and P. Lejeune. Forestimator: un plugin Qgis d'estimation de la hauteur dominante et du site index de peuplements résineux à partir de lidar aérien. *Revue Française de Photogrammétrie et de Télédétection*, (211-212):119–127, 2015.

[39] T. Degen, F. Devillez, and A.-L. Jacquemart. Gaps promote plant diversity in beech forests (Luzulo–Fagetum), North Vosges, France. *Annals of Forest Science*, 62:429–440, 2005.

[40] D. Dobrowolska and T. Veblen. Treefall–gap structure and regeneration in mixed Abies alba stands in central Poland. *Forest Ecology and Management*, 255(8):3469–3476, 2008.

[41] M. Drauschke, J. Bartelsen, and P. Reidelstuerz. Towards UAV-based Forest Monitoring. In *Proceedings of the Workshop on UAV-based Remote Sensing Methods for Monitoring Vegetation*, pages 21–32, Cologne, Germany, 2014. Geographisches Institut der Universität zu Köln - Kölner Geographische Arbeiten.

[42] L. Duarte, A. C. Teodoro, O. Moutinho, and J. A. Gonçalves. Open-source GIS application for UAV photogrammetry based on MicMac. 38(8-10):3181–3202, 2017.

[43] H. Eisenbeiss. *UAV photogrammetry*. PhD thesis, 2009.

[44] European Commission. *Copernicus Europe's eyes on Earth*. 2015.

[45] European Space Agency. Sentinel-2a MSI Spectral Responses.xlsx, 2015.

[46] J. Fox and S. Weisberg. *An R Companion to Applied Regression*. Sage, Thousand Oaks CA, second edition, 2011.

[47] A. Fritz, T. Kattenborn, and B. Koch. Uav-Based Photogrammetric Point Clouds - Tree Stem Mapping in Open Stands in Comparison to Terrestrial Laser Scanner Point Clouds. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, (2):141–146, Aug. 2013.

[48] E. Führer. Forest functions, ecosystem stability and management. *Forest Ecology and management*, 132(1):29–38, 2000.

[49] Gatewing. *X100 UAS - X100 System Manual*, 2012.

[50] R. Gaulton and T. J. Malthus. LiDAR mapping of canopy gaps in continuous cover forests: A comparison of canopy height model and point cloud based techniques. *International Journal of Remote Sensing*, 31(5):1193–1211, 2010.

[51] Geneq Inc. *SXBlue II GPS Series - Technical Reference Manual (v 2.3)*, 2012.

[52] R. Genuer, J.-M. Poggi, and C. Tuleau-Malot. Variable selection using random forests. *Pattern Recognition Letters*, 31(14):2225–2236, 2010.

[53] R. Genuer, J.-M. Poggi, and C. Tuleau-Malot. *VSURF: Variable Selection Using Random Forests*, 2014. R package version 0.8.1.

[54] S. Getzin, R. S. Nuske, and K. Wiegand. Using unmanned aerial vehicles (UAV) to quantify spatial gap patterns in forests. *Remote Sensing*, 6(8):6988–7004, 2014.

[55] S. Getzin, K. Wiegand, and I. Schöning. Assessing biodiversity in forests using very high-resolution images and unmanned aerial vehicles. *Methods in Ecology and Evolution*, 3(2):397–404, 2012.

[56] M.-C. Girard and C.-M. Girard. *Traitement des données de télédétection*. Dunod, 1999.

[57] M. C. Hansen, S. V. Stehman, and P. V. Potapov. Quantification of global gross forest cover loss. *Proceedings of the National Academy of Sciences*, 107(19):8650–8655, 2010.

[58] A. T. Hudak, J. S. Evans, and A. M. Stuart Smith. LiDAR utility for natural resource managers. *Remote Sensing*, 1(4):934–951, 2009.

[59] H. Hytteborn and T. Verwijst. Small-scale disturbance and stand structure dynamics in an old-growth Picea abies forest over 54 yr in central Sweden. *Journal of Vegetation Science*, 25(1):100–112, 2014.

[60] J. Hyyppä, H. Hyyppä, X. Yu, H. Kaartinen, A. Kukko, and M. Holopainen. Forest inventory using small-footprint airborne LiDAR. In *Topographic laser ranging and scanning: Principles and processing*, pages 335–370. CRC/Taylor & Francis, Book News, Inc.: Portland, OR, USA, 2009.

[61] M. Isenburg. LAStools - efficient tools for LiDAR processing (version 150406, academic). *obtained from http://rapidlasso.com/LAStools*, 2015.

[62] J. Järnstedt, A. Pekkarinen, S. Tuominen, C. Ginzler, M. Holopainen, and R. Viitala. Forest variable estimation using a high-resolution digital surface model. *ISPRS Journal of Photogrammetry and Remote Sensing*, 74:78–84, 2012.

[63] S. Jennings, N. Brown, and D. Sheil. Assessing forest canopies and understorey illumination: canopy closure, canopy cover and other measures. *Forestry*, 72(1):59–74, 1999.

[64] J. Jensen. *Remote Sensing of the Environment - An Earth Resource Perspective (Second Edition)*. Pearson Prentice Hall, Upper Saddle River, 2007.

[65] V. Junttila, T. Kauranne, A. O. Finley, and J. B. Bradford. Linear models for airborne-laser-scanning-based operational forest inventory with small field sample size and highly correlated LiDAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 53(10):5600–5612, 2015.

[66] V. R. Kane, R. F. Gersonde, J. A. Lutz, R. J. McGaughey, J. D. Bakker, and J. F. Franklin. Patch dynamics and the development of structural and spatial heterogeneity in Pacific Northwest forests. *Canadian Journal of Forest Research*, 41(12):2276–2291, 2011.

[67] P. Kempeneers, L. Bertels, K. Vreys, and J. Biesemans. Geometric Errors of Remote Sensing Images Over Forest and Their Propagation to Bidirectional Studies. *IEEE Geoscience and Remote Sensing Letters*, 10(6):1459–1463, Nov. 2013.

[68] D. A. King. Allometry and life history of tropical trees. *Journal of tropical ecology*, 12(1):25–44, 1996.

[69] L. Korhonen, K. T. Korhonen, M. Rautiainen, and P. Stenberg. Estimation of forest canopy cover: a comparison of field measurement techniques. *Silva Fennica*, 40(4):577–588, 2006.

[70] L. Korhonen, I. Korpela, J. Heiskanen, and M. Maltamo. Airborne discrete-return LIDAR data in the estimation of vertical canopy cover, angular canopy closure and leaf area index. *Remote Sensing of Environment*, 115(4):1065–1080, 2011.

[71] L. Korhonen and F. Morsdorf. Estimation of canopy cover, gap fraction and leaf area index with airborne laser scanning. In M. Maltamo, E. Næsset, and J. Vauhkonen, editors, *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*, volume 27 of *Managing Forest Ecosystems*, pages 397–417. Springer Science & Business Media, Netherlands, 2014.

[72] S. Koukoulas and G. Blackburn. Quantifying the spatial properties of forest canopy gaps using LiDAR imagery and GIS. *International Journal of Remote Sensing*, 25(15):3049–3072, 2004.

[73] S. Koukoulas and G. Blackburn. Spatial relationships between tree species and gap characteristics in broad-leaved deciduous woodland. *Journal of Vegetation Science*, 16(5):587–596, 2005.

[74] C. B. Kukunda, J. Duque-Lazo, E. González-Ferreiro, H. Thaden, and C. Kleinn. Ensemble classification of individual pinus crowns from multispectral satellite imagery and airborne LiDAR. *International Journal of Applied Earth Observation and Geoinformation*, 65:12–23, 2018.

[75] C. Laurent, D. Perrin, D. Bemelmans, M. Carnol, H. Claessens, C. De Cannière, L. François, E. Gerard, J.-C. Grégoire, M. Herman, et al. Le changement climatique et ses impacts sur les forêts wallonnes. recommandations aux décideurs et aux propriétaires et gestionnaires. 2009.

[76] H. Lecomte. La forêt wallonne en quelques chiffres. In P. Blerot, editor, *Le grand livre de la forêt*, pages 37–46. Forêt.Nature, Belgium, 2017.

[77] H. Lecomte, P. Florkin, J. Morimont, and M. Thirion. La forêt wallonne, état de la ressource à la fin du 20ème siècle. *Les caractéristiques du milieu forestier. Division de la Nature et des Forêts, Direction des Ressources forestières, Cellule Inventaire des Ressources forestières de Wallonie*, 2006.

[78] A. Lee and R. Lucas. A LiDAR-derived canopy density model for tree stem and crown mapping in australian forests. *Remote Sensing and Environment*, 111(4):493–518, 2007.

[79] P. Lejeune and J. Rondeux. Inventaire forestier de gestion: proposition d'un cadre méthodologique et technique. *Forêt Wallonne*, (73):22–29, 2004.

[80] A. Liaw and M. Wiener. Classification and regression by randomForest. *R news*, 2(3):18–22, 2002.

[81] G. Ligot, P. Balandier, A. Fayolle, P. Lejeune, and H. Claessens. Height competition between Quercus petraea and Fagus sylvatica natural regeneration in mixed and uneven-aged stands. *Forest Ecology and Management*, 304:391–398, 2013.

[82] T. Lillesand, R. W. Kiefer, and J. Chipman. *Remote sensing and image interpretation (Sixth edition)*. John Wiley & Sons, 2008.

[83] J. Linchant, J. Lisein, J. Semeki, P. Lejeune, and C. Vermeulen. Are unmanned aircraft systems (UASs) the future of wildlife monitoring? A review of accomplishments and challenges. *Mammal Review*, 45(4):239–252, 2015.

[84] D. Lindenmayer, J. Franklin, and J. Fischer. General management principles and a checklist of strategies to guide forest biodiversity conservation. *Biological conservation*, 131(3):433–445, 2006.

[85] D. B. Lindenmayer, C. R. Margules, and D. B. Botkin. Indicators of biodiversity for ecologically sustainable forest management. *Conservation biology*, 14(4):941–950, 2000.

[86] J. Lisein, A. Michez, H. Claessens, and P. Lejeune. Discrimination of deciduous tree species from time series of unmanned aerial system imagery. *PloS one*, 10(11):e0141006, 2015.

[87] J. Lisein, M. Pierrot-Deseilligny, S. Bonnet, and P. Lejeune. A photogrammetric workflow for the creation of a forest canopy height model from small unmanned aerial system imagery. *Forests*, 4(4):922–944, 2013.

[88] F. Loetsch, F. Zöhrer, and K. Haller. *Forest inventory. Vol. II*. 1973.

[89] X. Lu, Q. Guo, W. Li, and J. Flanagan. A bottom-up approach to segment individual deciduous trees using leaf-off LiDAR point cloud data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 94:1–12, 2014.

[90] T. Lumley and A. Miller. *leaps: regression subset selection*, 2009. R package version 2.9.

[91] D. Malcolm, W. Mason, and G. Clarke. The transformation of conifer forests in Britain - regeneration, gap size and silvicultural systems. *Forest Ecology and Management*, 151(1-3):7–23, 2001.

[92] C. Mallet and F. Bretar. Full-waveform topographic lidar: State-of-the-art. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(1):1–16, 2009.

[93] M. Maltamo, E. Næsset, and J. Vauhkonen. Forestry applications of airborne laser scanning. *Concepts and case studies - Manag For Ecosys*, 27:464, 2014.

[94] M. Maltamo, P. Packalén, E. Kallio, J. Kangas, J. Uuttera, J. Heikkilä, et al. Airborne laser scanning based stand level management inventory in Finland. In *Proceedings of SilviLaser 2011, 11th International Conference on LiDAR Applications for Assessing Forest Ecosystems, University of Tasmania, Australia, 16-20 October 2011.*, pages 1–10. Conference Secretariat, 2011.

[95] M. Maltamo, P. Packalén, A. Suvanto, K. Korhonen, L. Mehtätalo, and P. Hyvönen. Combining ALS and NFI training data for forest management planning: a case study in Kuortane, Western Finland. *European Journal of Forest Research*, 128(3):305–317, 2009.

[96] G. McDermid, R. Hall, G. Sanchez-Azofeifa, S. Franklin, G. Stenhouse, T. Kobliuk, and E. LeDrew. Remote sensing and forest inventory for wildlife habitat assessment. *Forest Ecology and Management*, 257(11):2262–2269, 2009.

[97] R. J. McGaughey. FUSION/LDV: Software for LiDAR data analysis ad visualization. *USDA Forets Service. Pacific Northwest Research Station, University of Washington, Seattle, WA, USA*, March 2014 - FUSION Version 3.42, 2014.

[98] R. McRoberts and E. Tomppo. Remote sensing support for national forest inventories. *Remote Sensing of Environment*, 110(4):412–419, Oct. 2007.

[99] A. Michez, H. Piégay, L. Jonathan, H. Claessens, and P. Lejeune. Mapping of riparian invasive species with supervised classification of unmanned aerial system (UAS) imagery. *International Journal of Applied Earth Observation and Geoinformation*, 44:88–94, 2016.

[100] A. Michez, H. Piégay, P. Lejeune, and H. Claessens. Multi-temporal monitoring of a regional riparian buffer network (>12,000 km) with LiDAR and photogrammetric point clouds. *Journal of Environmental Management*, 2017.

[101] A. Michez, H. Piégay, J. Lisein, H. Claessens, and P. Lejeune. Classification of riparian forest species and health condition using multi-temporal and hyperspatial imagery from unmanned aerial system. *Environmental monitoring and assessment*, 188(3):1–19, 2016.

[102] A. S. Milas, K. Arend, C. Mayer, M. A. Simonson, and S. Mackey. Different colours of shadows: classification of UAV images. *International Journal of Remote Sensing*, 38(8-10):3084–3100, 2017.

[103] M. Möller, L. Lymburner, and M. Volk. The comparison index: A tool for assessing the accuracy of image segmentation. *International Journal of Applied Earth Observation and Geoinformation*, 9(3):311–321, 2007.

[104] J.-M. Monnet, E. Mermin, J. Chanussot, and F. Berger. Tree top detection using local maxima filtering: a parameter sensitivity analysis. In *10th International Conference on LiDAR Applications for Assessing Forest Ecosystems (Silvilaser 2010)*, page 9p, Freiburg, Germany, 2010.

[105] E. Næsset. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research*, 19(2):164–179, 2004.

[106] E. Næsset. Airborne laser scanning as a method in operational forest inventory: status of accuracy assessments accomplished in Scandinavia. *Scandinavian Journal of Forest Research*, 22(5):433–442, 2007.

[107] T. Nagel, M. Svoboda, T. Rugani, and J. Diaci. Gap regeneration and replacement patterns in an old-growth Fagus–Abies forest of Bosnia–Herzegovina. *Plant Ecology*, 208(2):307–318, 2010.

[108] J. Perin, O. De Thier, H. Claessens, P. Lejeune, and J. Hébert. Nouvelles courbes de productivité harmonisées pour le douglas, l'épicéa et les mélèzes en Wallonie. *Forêt Wallonne*, 129(mars/avril):26–41, 2014.

[109] J. Perin, J. Hébert, Y. Brostaux, P. Lejeune, and H. Claessens. Modelling the top-height growth and site index of norway spruce in Southern Belgium. *Forest ecology and management*, 298:62–70, 2013.

[110] G. Petrie and C. K. Toth. Terrestrial laser scanners. *Topographic Laser Ranging and Scanning Principles and Processing*, pages 87–128, 2009.

[111] M. Pierrot-Deseilligny. MicMac documentation: *MicMac, Apero, Pastis and Other Beverages in a Nutshell!*, 2016.

[112] M. Pierrot-Deseilligny and I. Clery. Apero, an open source bundle adjusment software for automatic calibration and orientation of set of images. In *Proceedings of the ISPRS Symposium, 3DARCH11*, 2011.

[113] M. Pierrot-Deseilligny and N. Paparoditis. A multiresolution and optimization-based image matching approach: An application to surface reconstruction from SPOT5-HRS stereo imagery. *Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(1/W41), 2006.

[114] N. Pineux, J. Lisein, G. Swerts, C. Bielders, P. Lejeune, G. Colinet, and A. Degré. Can DEM time series produced by UAV be used to quantify diffuse erosion in an agricultural watershed? *Geomorphology*, 280:122–136, 2017.

[115] J. Pinheiro, D. Bates, S. DebRoy, D. Sarkar, and R Core Team. *nlme: Linear and Nonlinear Mixed Effects Models*, 2017. R package version 3.1-131.

[116] D. G. Pitt, M. Woods, and M. Penner. A comparison of point clouds derived from stereo imagery and airborne laser scanning for the area-based estimation of forest inventory attributes in Boreal Ontario. *Canadian Journal of Remote Sensing*, 40(3):214–232, 2014.

[117] A. Pommerening. Approaches to quantifying forest structures. *Forestry*, 75(3):305–324, 2002.

[118] A. Pommerening. Evaluating structural indices by reversing forest structural analysis. *Forest Ecology and Management*, 224(3):266–277, 2006.

[119] A. Pommerening and S. Murphy. A review of the history, definitions and methods of continuous cover forestry with special attention to afforestation and restocking. *Forestry*, 77(1):27–44, 2004.

[120] S. C. Popescu and R. H. Wynne. Seeing the trees in the forest. *Photogrammetric Engineering & Remote Sensing*, 70(5):589–604, 2004.

[121] S. C. Popescu, R. H. Wynne, and R. F. Nelson. Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size. *Computers and electronics in agriculture*, 37(1):71–95, 2002.

[122] T. L. Poulson and W. J. Platt. Gap light regimes influence canopy tree diversity. *Ecology*, 70(3):553–555, 1989.

[123] S. Puliti, H. O. Ørka, T. Gobakken, and E. Næsset. Inventory of Small Forest Areas Using an Unmanned Aerial System. *Remote Sensing*, 7(8):9632, 2015.

[124] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015. ISBN 3-900051-07-0.

[125] J.-P. Renaud, C. Vega, S. Durrieu, J. Lisein, S. Magnussen, P. Lejeune, and M. Fournier. Stand-level wind damage can be assessed using diachronic photogrammetric canopy height models. *Annals of Forest Science*, 74(4):74, 2017.

[126] D. Rocchini and A. Di Rita. Relief effects on aerial photos geometric correction. *Applied Geography*, 25(2):159–168, 2005.

[127] J. Rondeux, C. Sanchez, and N. Latte. Belgium (Walloon Region). In T. Erkki, T. Gschwantner, M. Lawrence, and R. E. McRoberts, editors, *National Forest Inventories-Pathways for Common Reporting*, pages 73–87. Springer Science & Business Media, Netherlands, 2010.

[128] J. R. Runkle. Patterns of disturbance in some old–growth mesic forests of Eastern North America. *Ecology*, 63(5):1533–1546, 1982.

[129] C. Sanchez. Synthèse des méthodes d'irrégularisation des pessières pour la wallonnie. vers une sylviculture durable, des écosystemes diversifiés et des revenus soutenus., 2012.

[130] S. Schliemann and J. Bockheim. Methods for studying treefall gaps: A review. *Forest Ecology and Management*, 261(7):1143–1151, 2011.

[131] T. Sing, O. Sander, N. Beerenwinkel, and T. Lengauer. Rocr: visualizing classifier performance in R. *Bioinformatics*, 21(20):3940–3941, 2005.

[132] N. Snavely, S. M. Seitz, and R. Szeliski. Modeling the world from internet photo collections. *International Journal of Computer Vision*, 80(2):189–210, Nov. 2008.

[133] B. St-Onge, P. Treitz, and M. A. Wulder. Tree and canopy height estimation with scanning lidar. In *Remote Sensing of Forest Environments*, pages 489–509. Springer, 2003.

[134] B. St-Onge, U. Vepakomma, J.-F. Sénécal, D. Kneeshaw, and F. Doyon. Canopy Gap Detection and Analysis with Airborne Laser Scanning. In M. Maltamo, E. Næsset, and J. Vauhkonen, editors, *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*, pages 419–438. Springer Science & Business Media, Netherlands, 2014.

[135] C. Straub, C. Stepper, R. Seitz, and L. T. Waser. Potential of UltraCamX stereo images for estimating timber volume and basal area at the plot level in mixed European forests. *Canadian Journal of Forest Research*, 43(8):731–741, 2013.

[136] Suhet and B. Hoersch. *Sentinel-2 User Handbook*. European Space Agency, 2015.

[137] T. Tanhuanpää, N. Saarinen, V. Kankare, K. Nurminen, M. Vastaranta, E. Honkavaara, M. Karjalainen, X. Yu, M. Holopainen, and J. Hyyppä. Evaluating the Performance of High-Altitude Aerial Image-Based Digital Surface Models in Detecting Individual Tree Crowns in Mature Boreal Forests. *Forests*, 7(7):143, 2016.

[138] R. Tibshirani and F. Leisch. *bootstrap: Functions for the Book "An Introduction to the Bootstrap"*, 2013. R package version 2012.04-1.

[139] E. Tomppo, T. Gschwantner, M. Lawrence, and R. E. McRoberts, editors. *National Forest Inventories*. Springer Netherlands, Dordrecht, 2010. DOI: 10.1007/978-90-481-3233-1.

[140] S. Tuominen, A. Balazs, H. Saari, I. Pölönen, J. Sarkeala, and R. Viitala. Unmanned aerial system imagery and photogrammetric canopy height data in area-based estimation of forest variables. *Silva Fennica*, 49(5), 2015.

[141] R. Van der Perre, S. Bythell, P. Bogaert, H. Claessens, F. Ridremont, C. Tricot, C. Vincke, and Q. Ponette. La carte bioclimatique de Wallonie: un nouveau découpage écologique du territoire pour le choix des essences forestières. *Forêt.Nature*, (135):47–58, 2015.

[142] M. Van Leeuwen and M. Nieuwenhuis. Retrieval of forest structural parameters using LiDAR remote sensing. *European Journal of Forest Research*, 129(4):749–770, 2010.

[143] M. Vastaranta, M. A. Wulder, J. C. White, A. Pekkarinen, S. Tuominen, C. Ginzler, V. Kankare, M. Holopainen, J. Hyyppä, and H. Hyyppä. Airborne laser scanning and digital stereo imagery measures of forest structure: Comparative results and implications to forest mapping and inventory update. *Canadian Journal of Remote Sensing*, 39(5):382–395, 2013.

[144] J. Vauhkonen, M. Maltamo, R. E. McRoberts, and E. Næsset. Introduction to forestry applications of airborne laser scanning. In M. Maltamo, E. Næsset, and J. Vauhkonen, editors, *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*, volume 27 of *Managing Forest*

Ecosystems, pages 1–16. Springer Science & Business Media, Netherlands, 2014.

[145] C. Véga and S. Durrieu. Multi-level filtering segmentation to measure individual tree parameters based on Lidar data: Application to a mountainous forest with heterogeneous stands. *International Journal of Applied Earth Observation and Geoinformation*, 13(4):646 – 656, 2011.

[146] C. Véga, A. Hamrouni, S. El Mokhtari, J. Morel, J. Bock, J.-P. Renaud, M. Bouvier, and S. Durrieu. PTrees: A point-based approach to forest tree extraction from lidar data. *International Journal of Applied Earth Observation and Geoinformation*, 33:98–108, 2014.

[147] M. Vehmas, P. Packalen, M. Maltamo, and K. Eerikainen. Using airborne laser scanning data for detecting canopy gaps and their understory type in mature boreal forest. *Annals of Forest Science*, 68:825–835, 2011.

[148] U. Vepakomma, D. Kneeshaw, and B. St-Onge. Interactions of multiple disturbances in shaping boreal forest dynamics: a spatially explicit analysis using multi–temporal lidar data and high–resolution imagery. *Journal of Ecology*, 98:526–539, 2010.

[149] U. Vepakomma, B. St-Onge, and D. Kneeshaw. Spatially explicit characterization of boreal forest gap dynamics using multi–temporal lidar data. *Remote Sensing of Environment*, 112(5):2326–2340, 2008.

[150] C. Vermeulen, P. Lejeune, J. Lisein, P. Sawadogo, and P. Bouché. Unmanned aerial survey of elephants. *PloS One*, 8(2):e54700, 2013.

[151] K. von Gadow, C. Y. Zhang, C. Wehenkel, A. Pommerening, J. Corral-Rivas, M. Korol, S. Myklush, G. Y. Hui, A. Kiviste, and X. H. Zhao. Forest structure and diversity. In *Continuous cover forestry*, pages 29–83. Springer, 2012.

[152] L. Wallace, A. Lucieer, Z. Malenovskỳ, D. Turner, and P. Vopěnka. Assessment of Forest Structure Using Two UAV Techniques: A Comparison of Airborne Laser Scanning and Structure from Motion (SfM) Point Clouds. *Forests*, 7(3):62, 2016.

[153] L. Waser, C. Ginzler, M. Kuechler, E. Baltsavias, and L. Hurni. Semi-automatic classification of tree species in different forest ecosystems by spectral and geometric variables derived from Airborne Digital Sensor (ADS40) and RC30 data. *Remote Sensing of Environment*, 115(1):76–85, 2011.

[154] A. S. Watt. Pattern and process in the plant community. *The Journal of Ecology*, 35:1–22, 1947.

[155] M. J. Westoby, J. Brasington, N. F. Glasser, M. J. Hambrey, and J. M. Reynolds. Structure-from-Motion photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, 179:300–314, 2012.

[156] J. C. White, C. Stepper, P. Tompalski, N. C. Coops, and M. A. Wulder. Comparing ALS and image-based point cloud metrics and modelled forest inventory attributes in a complex coastal forest environment. *Forests*, 6(10):3704–3732, 2015.

[157] J. C. White, M. A. Wulder, A. Varhola, M. Vastaranta, N. C. Coops, B. D. Cook, D. Pitt, and M. Woods. A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. *The Forestry Chronicle*, 89(6):722–723, 2013.

[158] J. C. White, M. A. Wulder, M. Vastaranta, N. C. Coops, D. Pitt, and M. Woods. The utility of image-based point clouds for forest inventory: A comparison with airborne laser scanning. *Forests*, 4(3):518–536, June 2013.

[159] K. Whitehead and C. H. Hugenholtz. Remote sensing of the environment with small unmanned aircraft systems (UASs), part 1: A review of progress and challenges 1. *Journal of Unmanned Vehicle Systems*, 2(3):69–85, 2014.

[160] K. Whitehead, C. H. Hugenholtz, S. Myshak, O. Brown, A. LeClair, A. Tamminga, T. E. Barchyn, B. Moorman, and B. Eaton. Remote sensing of the environment with small unmanned aircraft systems (UASs), part 2: scientific and commercial applications 1. *Journal of Unmanned Vehicle Systems*, 2(3):86–102, 2014.

[161] G. Wilhelm and H. Rieger. *Stratégie QD: une gestion de la forêt basée sur la qualité et les cycles naturels*. Institut pour le Développement Forestier, 2017.

[162] M. Woods, D. Pitt, M. Penner, K. Lim, D. Nesbitt, D. Etheridge, and P. Treitz. Operational implementation of a LiDAR inventory in Boreal Ontario. *The Forestry Chronicle*, 87(4):512–528, 2011.

[163] C. Wu. Critical configurations for radial distortion self-calibration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 25–32, 2014.

[164] M. Wulder, N. Coops, A. Hudak, F. Morsdorf, R. Nelson, G. Newnham, and M. Vastaranta. Status and prospects for LiDAR remote sensing of forested ecosystems. *Canadian Journal of Remote Sensing*, 39(sup1):S1–S5, 2013.

[165] M. A. Wulder, C. W. Bater, N. C. Coops, T. Hilker, and J. C. White. The role of LiDAR in sustainable forest management. *The Forestry Chronicle*, 84(6):807–826, 2008.

[166] S. Yamamoto, N. Nishimura, T. Torimaru, T. Manabe, A. Itaya, and K. Becek. A comparison of different survey methods for assessing gap parameters in old-growth forests. *Forest Ecology and Management*, 262(5):886–893, 2011.

[167] A. Zeileis and T. Hothorn. Diagnostic checking in regression relationships. *R News*, 2(3):7–10, 2002.

[168] J. Zhang. Multi-source remote sensing data fusion: status and trends. *International Journal of Image and Data Fusion*, 1(1):5–24, 2010.

[169] K. Zhang. Identification of gaps in mangrove forests with airborne LIDAR. *Remote Sensing of Environment*, 112(5):2309–2325, 2008.