Regional scale mapping and characterization of the poplar resource using two remote sensing data mining approaches

Auteur : Lacaille, Baptiste
Promoteur(s) : Lejeune, Philippe; Latte, Nicolas
Faculté : Gembloux Agro-Bio Tech (GxABT)
Diplôme : Master en bioingénieur : gestion des forêts et des espaces naturels, à finalité spécialisée
Année académique : 2021-2022
URI/URL : http://hdl.handle.net/2268.2/16312

LIÈGE université
Gembloux
Agro-Bio Tech

REGIONAL SCALE MAPPING AND CHARACTERIZATION OF THE POPLAR RESOURCE USING TWO REMOTE SENSING DATA MINING APPROACHES

BAPTISTE LACAILLE

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

ANNÉE ACADEMIQUE 2021-2022

CO-PROMOTEURS : PHILIPPE LEJEUNE ET NICOLAS LATTE

LIÈGE université
Gembloux
Agro-Bio Tech

REGIONAL SCALE MAPPING AND CHARACTERIZATION OF THE POPLAR RESOURCE USING TWO REMOTE SENSING DATA MINING APPROACHES

BAPTISTE LACAILLE

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

ANNÉE ACADEMIQUE 2021-2022

CO-PROMOTEURS : PHILIPPE LEJEUNE ET NICOLAS LATTE

Remerciements

Je voudrais tout d'abord exprimer ma reconnaissance à mes promoteurs des Travaux de Fin d'Etudes (TFE): Monsieur Philippe Lejeune, pour m'avoir donné l'opportunité d'approfondir sujet adjacent à la télédétection, sa supervision et sa disponibilité ainsi que Monsieur Nicolas Latte pour l'ensemble de ces conseils et aides précieuses, notamment avec la programmation en R et l'apprentissage profond.

Je tiens à remercier Monsieur Samuel Quevauvillers pour la formation accélérée de pilotage de drone et Monsieur Jérôme Périn pour son appui par l'Inventaire Forestier Permanent de la Région wallonne. Plus globalement, je souhaite remercier tout le personnel de l'*Axe Gestion des Ressources Forestières - Université de Liège - Gembloux Agro-Bio Tech (GxABT)*, pour m'avoir mis à disposition un bureau, leur écoute et leur gentillesse.

Un grand merci à mes parents, pour leur confiance et leur soutien inestimable, ainsi qu'à mes frères pour leurs encouragements.

Enfin, je désire exprimer ma gratitude à Emi', ta présence, ta positivité et tous les moments passés ensemble et ceux à venir ont été une source de motivation.

Abstract

In Wallonia, the regional forest resources are estimated from field methods that may present biases for estimating fast-growing forest species area like poplars, thus requiring support from remote sensing-based solutions.

The objectives of this master thesis concern the mapping and characterization of the poplar resource in the province of Hainaut. Specifically, it investigates (i) the potential of S2 super-resolution images (Latte and Lejeune, 2020) and (ii) the use of orthoimages through a deep learning-based approach to map the poplar resource, followed by (iii) the ability of an aerial photogrammetry CHM to characterize the latter.

The used methods are divided into two approaches: classification of super-resolved Sentinel-2 images using Random Forest algorithm (Breiman, 2001), semantic segmentation of orthoimages through a Deep Layer Aggregation (Yu et al., 2018) Neural Network. Both approaches involve 5 steps: data preparation, supervised learning, map production, height classification and accuracy assessment.

The results for the first approach map, with a F1-score of 0.923, is limited in detecting young poplar plantations and overestimates the poplar resource. Then, the second approach produced a map presenting great potential to detect poplar trees with an average accuracy of 1 m between the position of correctly predicted and observed poplars, but still contains many False Negatives, resulting in a F1-score of 0.653. Finally, poplar resource characterization shows for the first and second approach a respective ratio of properly identified height classes of 50% and 69%, these results are contrasted by poor ground truth data and a convincing visual assessment.

To conclude, the super-resolution of Sentinel-2 images seems to bring a higher accuracy compared to the poplar resource map made on Sentinel-2 images by (Bolyn, Latte, Colson, et al., 2020). Furthermore, a potential to map the poplar resource from orthoimages using a deep learning-based approach has been highlighted in this project, despite a low accuracy to be the subject of a management tool at this time. Lastly, although contrasting results, it would seem that aerial photogrammetry Canopy Height Model could be appropriate to characterize the poplar resource in this project, but would require field validation.

Keywords **—** Forest mapping, Forest characterization, Poplar resource, Random Forests, Deep Learning, Deep Layer Aggregation, Sentinel-2 super-resolution images, Orthoimages, Aerial photogrammetric canopy height model.

Résumé

En Wallonie, les ressources forestières régionales sont estimées à partir de méthodes de terrain qui peuvent comporter des lacunes dans l'estimation de la superficie des espèces forestières à croissance rapide telles que les peupliers, nécessitant de ce fait le soutien de méthodes basées sur la télédétection.

Ce mémoire de master a pour objectif la cartographie et la caractérisation de la ressource en peupliers dans la province du Hainaut. Plus précisément, il est étudié (i) le potentiel des images à super-résolution Sentinel-2 et (ii) l'utilisation d'ortho-images par une approche basée sur l'apprentissage profond pour cartographier la ressource populicole, puis (iii) la capacité d'un modèle numérique de hauteur de la canopée par photogrammétrie aérienne à caractériser cette dernière.

Les méthodes utilisées sont divisées en deux approches : classification d'images Sentinel-2 super-résolues (Latte and Lejeune, 2020) au moyen de l'algorithme de Forêt d'arbres de décision (Breiman, 2001), segmentation sémantique d'ortho-images par l'intermédiaire d'un réseau neuronal d'agrégation de couches profondes (Yu et al., 2018). Les deux approches impliquent 5 étapes : préparation des données, apprentissage supervisé, production de cartes, classification en hauteur et évaluation de la précision.

Les résultats de la carte issue de la première approche, avec un score-F1 de 0,923, est limitée dans la détection des jeunes plantations de peupliers et surestime la ressource en peupliers. Ensuite, la deuxième approche a produit une carte présentant un grand potentiel de détection des peupliers avec une précision moyenne de 1 m entre la position des peupliers correctement prédits et observés, mais contient encore beaucoup de faux négatifs, ce qui donne un score F1 de 0,653. Enfin, la caractérisation de la ressource en peupliers montre pour la première et la seconde approche un ratio respectif de classes de hauteur correctement identifiées de 50% et 69%, ces résultats sont contrastés par des données de terrain discutable et une évaluation visuelle convaincante.

En conclusion, l'amélioration de la super-résolution d'image Sentinelle-2 semble apporter une plus grande précision par rapport à la carte de la ressource en peupliers réalisée par (Bolyn, Latte, Colson, et al., 2020). Par ailleurs, un potentiel pour cartographier la ressource en peuplier à partir d'ortho-images en utilisant une approche basée sur l'apprentissage profond a été mis en évidence dans ce projet, malgré une faible précision pour faire l'objet d'un outil de gestion à l'heure actuelle. Finalement, bien que les résultats soient contrastés, il semblerait qu'un modèle numérique de hauteur par photogrammétrie aérienne pourrait être appropriée pour caractériser la ressource en peuplier dans ce projet, mais nécessiterait une validation sur le terrain.

Mots-clés — Cartographie forestière, Caractérisation forestière, Ressource en peuplier, Forêt d'arbres de décision, Apprentissage profond, Agrégation de couches profondes, Images Sentinel-2 à super-résolution, Orthoimages, Modèle numérique de hauteur par photogrammétrie aérienne

Table of Contents

List of Figures

List of Tables

Abbreviations and Acronyms

AI Artificial intelligence. v, I, VI, 5

ANN Artificial Neural Network. I–III

CHM Canopy Height Model. 4, 10, 39, *Glossary:* Canopy Height Model

CNN Convolutional Neural Network. III–VI, 5, 8, 16

CPU Central Processing Unit. II, 11, 31

CV Computer Vision. I, IV, VI, VII, 5, 15, 17

DAP Digital Aerial Photogrammetry. 4, 10, 19, 38, 40

DL Deep Learning. II, III, VI, 5–7, 11, 13, 31, 37–40

DLA Deep Layer Aggregation. 5, 14–16, 40

DNF Department of Nature and Forests. 3, 8, 10, 22, 27

DNN Deep Neural Network. II–IV, 6, 15, 38

DSM Digital Surface Model. 4, 10

DTM Digital Terrain Model. 4, 10, 19, 38, 40

EO Earth Observation. VI, 4

FAO Food and Agriculture Organization. v, 1

FC Forest Cover. 3, 6, 8, 14, 15, 29, 38

FI Forest Inventory. 3

FN False Negative. 20, 21, 24, 26, 30–33, 39

FP False Positive. 17, 20–22, 24, 27, 30, 31, 37–39

GDAL/ORG Geospatial Data Abstraction Library. 11

GIS Geographic Information System. v, 3, 8, 10, 11, 22, 27

GPU Graphics Processing Unit. II, 11, 16, 31

IoU Intersection over Union. 20

LIDAR Light Detection And Ranging. 4, 5, 7, 10, 19, 37–40

ML Machine Learning. I–III, VI, 5–7, 13, 14, 29, 31, 37, 38, 40

NIR Near-infrared. 4, 7–10, 40

OTB Orfeo ToolBox. 11, 15

PCHM photogrammetric Canopy Height Model. 5, 7, 10, 17, 19, 30, 31, 34, 36, 38–40

PS PlanetScope. 8

PSW Public Service of Wallonia. 8, 9

ReLU Rectified Linear Unit. IV, 16

RF Random Forest. 5, 14, 15, 17, 40

RGB Red, Green, Blue. 4, 6–9, 40

S2 Sentinel-2. 4–11, 13, 29, 37, 38, 40

SA Spatial Analysis. 11, 27

SRWC Short Rotation Woody Crop. 1, 3

SWIR Short-Wave Infrared. 8, 10, 13

TN True Negative. 20, 21, 29

TP True Positive. 20–22, 27–29, 31–33, 40

UAV Unmanned Aerial Vehicle. vi, 6

VRE Vegetation Red-Edge. 10

WFI Walloon Region Forest Inventory. i, ii, 3, 19–22, 25–28, 30, 34–36, 39

Glossary

Big data are digital data sets too large and complex to be conventionally exploited. This term became popular with the rise of the Internet and mobile technologies, allowing the centralization and massive storage of data, as well as the proliferation of data collection. This concept is inherently linked to data mining. III

Canopy Height Model is a landscape element height resulting from the subtraction of a Digital Surface Model (DSM) and a Digital Terrain Model (DTM). The former includes the height of both landscape elements and terrain and the latter only the terrain. iii

Data mining means the analysis of data in order to make it useful. This concept is useful for many sectors as well as for science and engineering, especially for revealing patterns, correlations, anomalies, statistical information and predictions of useful data. Various techniques and technologies surround data mining, such as data warehousing, data preparation, statistical software, Artificial intelligence (AI). v

Image interpretation is a data collection process that involves image observation to characterize a landscape. In a Geographic Information System (GIS), a regular grid of points is distributed over the area of interest, and each point is assigned a class (*e.g.*, forest, grassland, agricultural, or urban area), in order to quantify the distribution of the different classes in that area.. 10, 14, 21, 24, 25, 29–31

Orthoimage is an "orthorectified" aerial or satellite image, in other words, the image geometries have been corrected and harmonized. Indeed, depending on the position and shape of the sensor, the altitude of the platform on which the sensor is placed, as well as the relief, the image contains a series of irregularities that require processing to allow a regular mosaic of images and comparison between different dates. 4, 5, 7, 9–11, 13, 14, 18, 19, 21, 24, 26, 37–40

Poplar plantation is a component of the poplar resource that comes from plantations. The Food and Agriculture Organization (FAO) defines forest plantations as forest areas resulting from planted trees, with an even-aged and regularly spaced structure and composed of one or two species. The purpose of forest plantations is mainly economic, so they are managed for wood production (FAO, 2020b). 2, 6, 8, 11, 18, 30, 31, 34, 36, 38–40

Poplar resource is defined in the context of this master's thesis as all wooded areas of more than 3 m in height, whose canopy composition is at least 80% poplar and whose ground cover is at least 100 m^2 or consists of a group of at least 5 stems spaced from 6 m to 10 m apart. Thus, in relation to the FAO definitions of woodland and forest area, the poplar resource can be located both inside and outside forests. This definition does not include aspens, as single trees are not considered (FAO, 2020b). v, 2, 3, 5, 7, 8, 11, 14, 18, 19, 24–26, 28–31, 34, 37–40

Remote sensing refers to all remote data acquisition techniques on an object, including optical or height data. In earth observation, the sensor is placed on a platform that can be either an Unmanned Aerial Vehicle (UAV), an airplane, or a satellite, each of which has advantages and disadvantages in terms of spatial resolution (size of the pixel on the ground), spectral resolution (ability to distinguish electromagnetic radiation, corresponding to the smallest bandwidth recorded so implicitly to the number of spectral bands recorded), and temporal resolution (time delays between data acquisitions). The sensors can be passive (measuring natural radiations such as solar or infrared) or active (by emitting radiations and measuring the portion reflected back to the sensor). 3–5, 7, 11, 38

1 Introduction

1.1 Research context

1.1.1 State of the poplar resource

Planted Forests: poplars

Forests have a clear key function in biodiversity, wood and food production, carbon sequestration and fixation, human well-being and many other ecosystem services. Yet, despite a slowdown in the rate of net forest loss due to reduced deforestation in some countries and natural forest expansion or afforestation in others, net forest loss remains significant, with an estimated loss of 10 million hectares per year between 2015 and 2020 (FAO, 2020b).

Part of the solution to this problem, while taking into account the growing demand and consumption of wood products (FAO, 2021), can be found in planted forests. Indeed, the global potential production of industrial wood from planted forests was estimated at 1.2 billion m^3 in 2005, or about 2/3 of the total wood production that year (Carle and Holmgren, 2008). The potential production must be related to the area of planted forests. In fact, in 2020, 7% of the 4.06 billion hectares of forests in the world were planted, 45% of which were for production purposes, while the rest of the planted forests were not intensively managed (FAO, 2020a; Carle and Holmgren, 2008).

The high yield of woody biomass production from planted forests arises from the use of Short Rotation Woody Crops (SRWCs) based on fast-growing forest species such as poplars (*Populus sp.*), willows (*Salix sp.*), eucalyptus (*Eucalyptus sp.*), pines (*Pinus sp.*) among others (Zalesny Jr et al., 2019). The importance of poplar to satisfy a global demand for woody biomass was recognized at the end of the Second World War, which led to the creation of the International Poplar Commission (IPC) under the auspices of the Food and Agriculture Organization (FAO) (Hamrouni Berkaoui, 2021). Belgium is one of the 38 member countries constituting the IPC. Research in Sweden has shown that the stem-wood production yield of poplar plantations reaches 25 to 30 $m^3 ha^{-1} year^{-1}$ for a rotation period of 15 to 20 years (Adler et al., 2021).

The genus Populus

Poplars, of the genus Populus of the family *Salicaceae*[1], are divided taxonomically into 6 sections: *Abaso* (Mexican poplar), *Aigeiros* (Cottonwoods and black poplar), *Leucoides* (Swamp poplars), *Populus* (White poplars and aspens), *Tacamahaca* (Balsam poplars), *Turanga* (Arid and tropical poplars). These trees are medium and large size, with simple deciduous leaves and alternate phyllotaxis. This is a dioecious species, with separate floral and vegetative buds. They are dioecious species, having bracted flowers borne by catkins,

[1] family of dicotyledonous plants, which counts, with the Salix genus (willows, sallows and osiers), about 400 to 500 species

with separate floral and vegetative buds and tiny seeds attached to a cottony coma. Some characteristics enable to distinguish in more detail the different sections of the genus Populus (FAO-CABI, 2014). The wood of low density, soft, creamy white, with an uniform texture and diffuse pores, is versatile and traditionally used for peeling (light packaging or plywood), sawing (pallets and packaging boxes), pulpwood or energy wood (*Peuplier* 2022).

The natural range of poplars covers a large majority of the northern hemisphere, within the latitudinal and latitudinal limits of tree growth. In addition, the genus *Populus* has been widely planted in the world[2]. Hence, his presence in the southern hemisphere can be noted (FAO-CABI, 2014). Poplars are pioneer species of disturbed areas, characterized by a rapid growth rate and short life span compared to other species, due to hosting many diseases and pests. Some species of poplars from North America, Europe or Asia can reach a diameter exceeding 3 m and a height of 45 m (FAO-CABI, 2014). As interspecific and intraspecific hybridization is naturally common in the Populus genus section, the potential was quickly exploited by cloning in commercial cultivation (FAO-CABI, 2014). This practice increased sharply after the sequencing of the *P. trichocarpa* genome in 2004, it was the first tree and the third plant sequenced (Tuskan et al., 2006).

Poplars in Wallonia

Wallonia, like the rest of Belgium and France, naturally has three species of poplars, black poplar (P. nigra)[3], white poplar (P. alba)[4] and aspen (P. tremula)[5]. Poplar plantations are based on cultivars resulting from the crossing of North American (*P.trichocarpa* and *P.deltoïdes*), European (*P.nigra*) and Asian (*P.maximowiczii*) species, based on criterias such as the adaptation to soil characteristics, resistance to diseases and abiotic factors and growth (Carah Asbl, 2018).

In 2021, the poplar area has slightly decreased from 9,650 ha in 2008 to 7,700 ha, corresponding to 2% of the productive forest area of Wallonia (Alderweireld et al., 2015). The poplar resource is not insignificant, accounting for 7.7% of the annual hardwood harvest volume in the region (Office économique wallon du bois, 2021). In 2008, the structure of poplar stands was 45% even-aged, 41% coppice, and the remainder two-story stands, with a relatively good balance of area by age class, although a slowdown in planting can be observed (Alderweireld et al., 2015). In Wallonia in 2021, the forest area was allocated as follows: 51% by private and 49% by public. This trend was not observed for the poplar resource, which was held 86% by private owners and 14% by public owners. It is also the only decapitalized hardwood species with a standing volume of 1,687,355 m^3 and a harvesting rate is 117% of the annual increase in volume (Office économique wallon du bois, 2021).

The advantages of poplars are economic, because with little work (planting, training pruning, pruning and a few visits) and a short rotation of about 20 years, the owner can expect a net income of 3400€ to 6800€/ha excluding taxes (Office économique wallon du bois, 2021). In addition, the development of methods for adding value to hardwoods, particularly by heat treatment, allows poplar to have new uses in addition to the conventional ones mentioned in

[2] thanks to its rapid growth rate and ease of genetic manipulation
[3] associated with alluvial valleys
[4] common in the Mediterranean valleys and found in the rest of France and in Belgium
[5] present in the majority of forest areas

Section 1.1.1, such as siding (FAO, 2020a). The benefits of poplars are also environmental, as they capture nitrates and phosphorus upstream of alluvial areas, and diversify the landscape in agricultural areas, thus the fauna and flora (Interreg : Forêt-Pro-Bos, 2020). Furthermore, the plantation of poplars is regulated by Walloon legislation to limit potential negative effects on the environment (Dumont, 2018). Finally, poplars are part of the landscape, especially along the canals and on the edge of the meadows in the province of Hainaut, which gives them a social value (Interreg : Forêt-Pro-Bos, 2020).

1.1.2 National forest inventory

Regional Forest Inventory

Reliable data are essential for coherent decision-making, including national forest policy or forest resource management. The primary large-scale data sources are national Forest Inventory (FI) and forest condition monitoring networks (Travaglini et al., 2013). In Belgium, the FI is the responsibility of the 3 regions (Wallonia, Flanders and Brussels), the modern Walloon Region Forest Inventory (WFI) results from a harmonization of the methods carried out in 1994, but remains in constant evolution to integrate new variables related to biodiversity, damage monitoring or forest certification monitoring (Rondeux et al., 2016). The latter is based on a single-phase systematic sampling of approximately 11,000 sampling points composed of several concentric plots depending on the nature of the data collected (Alderweireld et al., 2015).

Evolution in needs

National FI can provide accurate and comprehensive estimates, but collecting forest data over large areas can be time-consuming and costly. In addition, the variables involved in FIs must meet evolving needs related to the multifunctionality of forests, from Forest Cover (FC) and productive function of forests to forest health, biodiversity, protective function of forests, social services, carbon balance, land use, biomass for energy, water supply, and so on (Koch, Dees, et al., 2008). Demand for large-scale forest data will continue to grow, for decision-making at the European Union level, or for reporting related to international agreements and activities under the Paris Agreement, the Convention on Biological Diversity (CBD) and the United Nations Framework Convention on Climate Change (UNFCCC 2015) (Vauhkonen et al., 2019). Fortunately, large-scale field inventories can be supported by the increasing acquisition of remote sensing data and the evolution of methods. This will be further detailed later in Section 1.2.1.

1.1.3 Unclear estimated poplar area

In connection with the concerns raised in Section 1.1.2, national or regional FIs might be inappropriate for some species, particularly for fast-growing species used for SRWC. Indeed, these species exhibit rapid change in standing volume and may also occur outside forests as defined by FAO (2020b). Uncertainty in the poplar resource area estimates reported in Section 1.1.1 is related to the fact that they are based on two data sources: the WFI and the Department of Nature and Forests (DNF) GIS data, which do not take into account the two aforementioned elements related to SRWC.

1.2 Mapping and characterization of the forest resource using remote sensing

In this section, the focus is on reviewing remote sensing data and state-of-the-art predictive methods suitable for use in this project, the literature search was therefore limited to large-scale applications.

1.2.1 Remote sensing data

Remote sensing for local monitoring of forest resources using aerial imagery dates back to the middle of the 20^{th} century, followed by the dispatch in the 1970s of the first earth observation satellites[6] (Boyd and Danson, 2005). In the 1990s, an acceleration in the number of spatial programs for the EO occurred, in line with the awareness of the global changes (Hamrouni Berkaoui, 2021). The supply satellite images for EO has been widely diversified by a series of commercial programs, such as the French SPOT, the American Ikonos-2 and QuickBird, WolrdView and Pleiades (Hamrouni Berkaoui, 2021). Nevertheless, as mentioned in Section 1.1.2, an ever-increasing need for large-scale data processing was slowed by the cost of accessing commercial data. As a result, massive public and private investments have enabled the launch of free access programs, including the Copernicus land monitoring service (CLMS) (Copernicus, 2022). The CLMS, initiated by the European funding, cover all subjects related to remote sensing for EO. These services are based on the European satellites SPOT, PROBA-V and Sentinel (Ponomarenko and Zelentsov, 2021). The Sentinel satellite (ESA, 2022b) constellation has been deployed for the Copernicus program, two of which are of interest for forest resource monitoring, the Sentinel-1 and Sentinel-2 (S2) satellites. The former offering Synthetic-aperture radar (SAR) height data with a resolution of 5 m and the latter multispectral optical images.

The advantage of satellite images is that they have a high spectral resolution, generally multiband (about ten bands) to hyperspectral (hundreds of bands), a low revisit time and a large coverage of the scene (ESA, 2022a). These advantages may nevertheless be subject to certain limitations, such as low spatial resolution or dependence on cloud cover requiring complex image processing like image fusion through neural networks for satellite image fusion (Latte and Lejeune, 2020) or time series processing by domain adaptation (Hamrouni Berkaoui, 2021).

The Walloon region provides another type of valuable data source for large-scale mapping and characterization of the forest resource, namely orthoimages (PSW, 2022b) and Light Detection And Ranging (LIDAR) point clouds (PSW, 2022a). Unlike optical satellite images, orthoimages are characterized by a high spatial resolution (25 cm). They are acquired in the absence of cloud cover, but they have 4 bands (Red, Green, Blue (RGB) and Near-infrared (NIR)) and a revisit time of one year, which means that images from a given year may be outside of the growing season. Regarding the height data, aerial Canopy Height Models (CHMs), whether derived from a combination of the Digital Terrain Model (DTM) (from LIDAR) and Digital Surface Model (DSM) (from Digital Aerial Photogrammetry (DAP)) or from a full LIDAR CHM, have a higher spatial resolution than Sentinel-1 data. Nevertheless, a significant difference is observed between the two aerial height images. The LIDAR data with an average

[6] spatial Earth Observation (EO) Landsat program from National Aeronautics and Space Administration (NASA).

resolution of $0.8pts/m^2$ are direct measurements of height while the second obtained by crossing orthoimages creating a 3D relief, the latter is therefore less accurate. However, the high cost of LIDAR data acquisition reduces the temporal resolution, the last acquisition under this project took place between 2013 and 2014. In contrast, height data from the combination with photogrammetric Canopy Height Model (pCHM) has been produced for several years, including 2018 and 2019.

1.2.2 Artificial intelligence in remote sensing

Remote sensing involves many sciences, specifically physics, statistics, computer science and Artificial intelligence (AI). The latter is essential for forest resource map production, as predictive models are based on the learning domain (*see* Appendix A) of AI, and the tasks solved by these models applied to remote sensing data are derived from Computer Vision (CV) domain (*see* Appendix B) of AI.

As explained in Appendix A, the learning domain of AI is based on Machine Learning (ML), whereby the machine learns to solve a task for which it has not been explicitly configured. To map forest resources, supervised learning is required to tell the machine the labels to be predicted and the variables characterizing them. There are several algorithms commonly used for classification of remote sensing data, such as Support Vector Machines (SVMs), Naïive Bayes classifier, k-nearest neighbor classifier, remote sensing, as cited in (Kishore et al., 2016; C. Zhang et al., 2017).

Deep Learning (DL) is a subset of ML that differs in its automation. Namely, it does not require feature extraction before performing classification (Sandhya Devi, Vijay Kumar, and Sivakumar, 2021) and the learning process is iterative and nonlinear. Convolutional Neural Networks (CNNs) transform data by applying a convolutional filter. They are particularly well suited for multiband image processing and have been a game changer for the use of DL in CV (L. Zhang, L. Zhang, and Du, 2016). Among other CNN architectures, we can mention AlexNet, VGG, ResNet for image classification and U-Net, Deep Layer Aggregation (DLA), FastFCN, DeepLab, Mask R-CNN for object segmentation, as cited in (Ma et al., 2019; Wu, Q. Liu, and X. Liu, 2019).

Tasks involved in remote sensing data are divided into image preprocessing, change detection, accuracy assessment, and classification (Ma et al., 2019). The former includes image fusion, image registration and object segmentation, while the latter includes land cover and land use classification, scene classification and object recognition.

1.3 Remote sensing in support of the poplar resource inventory

Various research studies have been conducted for large-scale poplar resource mapping, mainly using ML approaches applied to satellite images. Highly accurate results were obtained for the mapping of hybrid poplar (P. deltoides) plantations in Sakary, Turkey, by using three algorithms: Random Forests (RFs), support vector machines (SVMs) and Adaboost (AdaB) applied to S2 images (Ozturk and Colkesen, 2020). In the same region, Kavzoglu, Tonbul, and Colkesen (2021) then performed a statistical comparison of atmospheric correction methods to increase the accuracy of poplar resource mapping. Another study conducted

in Tehran to map poplar plantations from S2 images by using an SVM algorithm has also obtained convincing results.

In Northern Italy, a study of the impact of spatial resolution for the mapping of poplar plantations was conducted between S2 images and aerial images from Unmanned Aerial Vehicle (UAV); the accuracy of the mapping was increased with the resolution (Chianucci et al., 2020), but the second data source does not allow a large-scale application.

Nearer to Wallonia, a complex S2 image processing approach for large-scale poplar grove mapping in France using ML classification has also shown good results which can be used by professionals in the sector (Hamrouni Berkaoui, 2021). In Wallonia, a mapping of the poplar FC by a ML approach has shown the interest of S2 images for forest mapping (Bolyn, Latte, Colson, et al., 2020).

A DL approach (fully connected neural network) compared to a more conventional ML approach (logistic regression) for the classification of poplar plantations in northern Italy revealed a superior potential for Deep Neural Networks (DNNs) (D'Amico et al., 2021). This was particularly true with the occurrence of non-linear patterns observed in massive datasets and the processing of massive data (D'Amico et al., 2021). A large-scale tree recognition method using a DNN U-Net architecture applied to very high resolution satellite images (GeoEye-1, 0.46 m/pixel and RGB images) to perform semantic segmentation was found to be effective in both detecting and differentiating Mongolian poplar (*Populus suaveolens*) and evergreen conifers (*Abies holophylla* , *Pinus koraiensis*) in the Primorsky region of the Russian Far East (Korznikov et al., 2021).

1.4 Research positioning

Poplars are important species for global and regional matters. In Wallonia, it is assumed that there is an inaccuracy in the estimation of the area of this resource. To overcome this uncertainty, remote sensing data are available within this Belgian region, namely: satellite and aerial images. Characterizing the predicted surfaces then provides a further insight into this woody resource.

1.4.1 Problematic and research questions

The problematic of this master's thesis concerns the study of different remote sensing data to support the inventory of the poplar resource in Wallonia. This problematic is divided into three main research questions:

- Using feature extraction developed by Hamrouni Berkaoui (2021) and Sentinel-2 super-resolution images by Latte and Lejeune (2020); can a finer spatial resolution of satellite images significantly improve the accuracy of poplar resource mapping?
- Could neural networks be used to map the poplar resource from another source of large-scale remote sensing data, namely orthoimages ? The last characterized by a lower spectral resolution but a high spatial resolution compared to the above-mentioned S2 super-resolution images.
- Can the high-quality photogrammetric Canopy Height Model (Michez et al., 2020) be used to characterize the poplar resource in the absence of available Light Detection And Ranging (LIDAR) data?

1.4.2 Objectives and approaches

The objectives of this master's thesis are to develop and compare two methods for mapping and characterizing the poplar resource on a regional scale, in the Province of Hainaut, part of the Walloon Region. The two approaches to meet these objectives are the following:

- A Machine Learning-based approach consisting of a pixel-based supervised classification of Sentinel-2 super-resolution images with a spatial resolution of 2.5 m and a spectral resolution of 10 bands.
- A Deep Learning-based approach, using orthoimage (Red, Green, Blue and Near-infrared) with a spatial resolution of 25 cm.

2 Material and Methods

2.1 Study area

Covering an area of 381,350 ha, the province of Hainaut was chosen to develop and compare two mapping and characterization methods of the poplar resource. Firstly, because this Belgian province, part of the Walloon Region, has a forest history closely linked to the poplar species. Secondly, the province of Hainaut has the largest area in 2008 with 5,400 ha, which corresponds to 55% of the region's poplar area and 10.2% of the province's total FC, even though it is the second smallest province in Wallonia according to the forest area and the FC rate[1] (Alderweireld et al., 2015). The Walloon poplar area has evolved between 2008 and 2021 from 9,800 ha to 7,700 ha (Office économique wallon du bois, 2021), but the trend mentioned above is still relevant. In fact, it is confirmed by both the DNF of the Public Service of Wallonia (PSW) GIS data with 318.3 ha of poplar plantations within the public productive forest area[2] in Wallonia in 2021, of which 56.5% are located in the province of Hainaut, and by observing the map of forest stands in Wallonia by remote sensing (Bolyn, Latte, Colson, et al., 2020).

Poplars are present in all natural regions of Wallonia, but are generally associated to agricultural landscapes with rich soils and good water supply (Interreg : Forêt-Pro-Bos, 2020); these conditions are found in the bioclimatic zones of the "Scaldian plains and valleys", "Hesbino-Brabançon" and "Sambre-et-Meuse and Condroz". The first, shown in red in Figure 1, covering the western and northern part of the province of Hainaut, is the most suitable for this tree species, thus explaining its abundance in this province.

2.2 Material

2.2.1 Remotely sensed data

Satellite images

Super-resolved satellite images were used in this master's thesis. These images were produced by fusing S2 and PlanetScope (PS) images through a method involving a CNN model based on residual learning, proposed by Latte and Lejeune (2020).

Funded by the European Copernicus project, the satellite images from the S2 constellation have a spatial resolution of 10 m, 20 m and 60 m, a spectral resolution of 13 bands across the visible, NIR and Short-Wave Infrared (SWIR) wavelengths and a temporal resolution of 5 days, (*see* ESA, 2022c). The PS constellation provides images with a pixel size of 3 m including 4 bands (RGB and NIR), acquired daily thanks to more than 180 nano-satellites,

[1] 13.9% of the area of the province is covered by forests, which is 53,150 ha, including 46,000 ha of productive forests

[2] 40% of the productive forest area of the province of Hainaut is public and 60% private

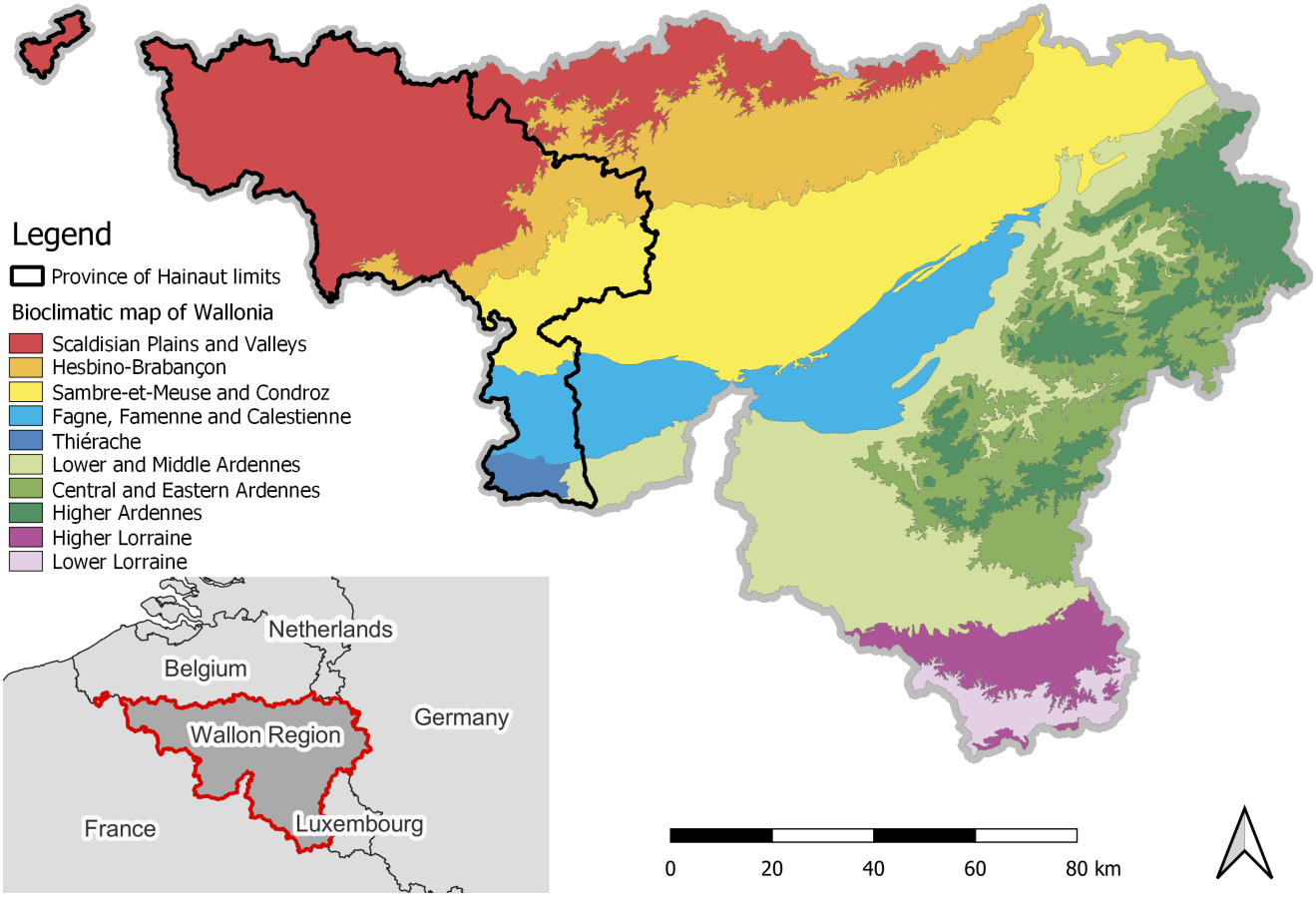


Figure 1: Bioclimatic map of Wallonia from the PSW (*see* Van-Der-Perre et al., 2015), with the administrative limits of the province of Hainaut in black, bounding the study area.

property of the private American company Planet Labs, (*see* Planet Labs PBC, 2021). The super-resolution images return the 10 bands from the 10 m and 10 m resolution S2 images, but with an enhanced pixel size at 2.5 m. These bands are listed in table 1.

Unfortunately, the above processing to super-resolve S2 images at a specific date includes gaps in the data corresponding to clouds and shadows. To address this, a multi-date mosaic of super-resolved images was produced over the 2018 growing season, based on a weighted average of pixels per date, for which the weighting is relative to the peak date of the growing season; the further away from that date, the lower the weighting assigned to pixels.

Aerial images

Another source of data comes from aerial images, namely a set of orthoimage covering Wallonia at different years, belonging to the PSW and freely available (*see* PSW, 2022b). The suitable aerial images for this study are those of the years 2016, 2018, 2019, 2020 and 2021; the others are either too old or acquired outside the plant growth period. It should be noted that the 2019 orthoimage was acquired on the most suitable dates for the growing season within the study area, from June 2, 2019 to August 24, 2019. These optical data with a spatial resolution of 25 cm and a spectral resolution of 4 bands (RGB and NIR) are

complemented by a DSM from DAP. A pCHM was then obtained by combining these DAP with the DTM from 2014 LIDAR using the approach developed by Michez et al. (2020).

Table 1: Spectral bands of the super-resolved Sentinel-2 images.

Band name	**Band ID**	**Central wavelength (nm)**	**Bandwidth (nm)**
Blue	B02	490	66
Green	B03	560	36
Red	B04	665	31
VRE	B05	705	16
VRE	B06	740	15
VRE	B07	783	20
NIR	B08	842	106
VRE	B08A	865	22
SWIR	B11	1,610	93
SWIR	B12	2,190	180

2.2.2 Reference dataset

The reference dataset consists of a set of points and polygons. The first composed of 25,951 point-marked poplars was generated from data collection within reference sampling unit, using Geographic Information System (GIS) tools by surveying the 2019 orthoimage. The second comes from previously produced points and the 2021 DNF GIS data. The latter is used by the forestry administration of the Walloon Region to prepare forest management plans and organize current management tasks (Crutzen, 2017). Two classes were considered in the DNF GIS data to form the reference polygons, "poplars" and "others"; the first one includes plots containing only poplars and the second one all other plots without poplars.

The area covered by the reference polygons is 269 ha for the "poplar" label (146 ha obtained from poplar pointing and 123 ha from the DNF GIS data) and 15,638 ha for the "other" label from the DNF GIS data.

2.2.3 Other datasets

An additional dataset was made from an image interpretation step performed on the whole woody mask applied to the province of Hainaut. The latter was produced by the mapping of trees outside of forests using height data from aerial LIDAR (*see* Bolyn, Latte, Fourbisseur, et al., 2020). The woody mask includes all wooded vegetation in the landscape, which after intersection with the 2018 CHM (*cf.*, section 2.2.1) whose spatial resolution has been reduced to 2.5 m, have a minimum height of 3 m and a minimum area of 6.25 m^2, corresponding to the area of one pixel. Thus, with a density of one point per square of 500 m side, 2,705 points in the 65,595 ha of the woody mask were assigned to the poplar or other class based on image interpretation of the 2018 orthoimage.

Finally, another dataset came from the plots of the Walloon regional forest inventory. Based on a proportion of poplar in the basal area of forest trees greater than or equal to 80%, plots

inventoried over the period 2014-2022 and a visual validation of the plots with respect to the 2019 orthoimage; 24 circular plots of 18 m radius were selected in the province of Hainaut. These data are supplemented by field measurements of the total height of some of the trees surveyed in the circular plots, and by point marking of all visually detectable trees from the 2019 orthoimage, for a total of 421 trees.

2.2.4 Hardware and software

The project ran on two different computers both with window 10 64 bits as explorer. Central Processing Unit (CPU) and Graphics Processing Unit (GPU) were respectively Intel®Core™with 20 cores for the first machine and 12 cores for the second and 64 GB of RAM for the first and 32 GB of RAM for the second. The DL and prediction by the neural network model was fully coded in R with four main packages: stars (Pebesma, Sumner, et al., 2022), sf (Pebesma, Bivand, Racine, et al., 2022), torch (Falbel, Luraschi, et al., 2022) and torchvision (Falbel, Regouby, and RStudio, 2022) by Nicolas Latte and executed on the first machine, the other processing was executed on the second machine and used the R packages: stars(Pebesma, Sumner, et al., 2022), raster (Hijmans et al., 2022), sf(Pebesma, Bivand, Racine, et al., 2022), sp (Pebesma, Bivand, Rowlingson, et al., 2018), randomForest (Cutler and Wiener, 2022), caret (Kuhn [aut et al., 2022) and in command line within R the Orfeo ToolBox (OTB) and Geospatial Data Abstraction Library (GDAL/ORG) software. Finally, data collection and GIS processing were performed using Qgis software.

2.3 Methods

Two approaches, each based on a different type of remote sensing data, have been developed to fulfill the research objectives related to large-scale mapping and characterization of the poplar resource. The first, a **classification-based approach**, implements a machine learning algorithm for pixel-based classification of the poplar resource from satellite images. The second is a **semantic segmentation-based approach** from aerial images using a DL algorithm followed by a Spatial Analysis (SA) of poplar plantation patterns. Both approaches are divided into 5 steps: **data preparation**, **supervised learning**, **map production**, resource characterization by **height classification** and **accuracy assessment**. It is important to note that the map production step by the semantic segmentation approach uses at some point a map produced in the classification approach. The process flowchart of the methodologies from data preparation to height classification is shown in Figure 2, and the evaluation of the results in Table 7.

2.3.1 Data preparation

Satellite images

The classification-based approach was performed on the **2018 S2 super-resolved images** (*cf.*, Section 2.2.1), within the **woody mask** (*cf.*, Section 2.2.3). The **features extracted** from these remote sensing data are based on the spectral indices selected in the thesis of Hamrouni Berkaoui (2021) concerning the automation of the poplar resource mapping in France, they are listed in Table 2. All variables (bands from the S2 super-resolved images and feature extraction) were **standardized**, *i.e.*, centered and reduced using the mean and standard deviation computed over all these variables, in order to limit troubles due to range differences.

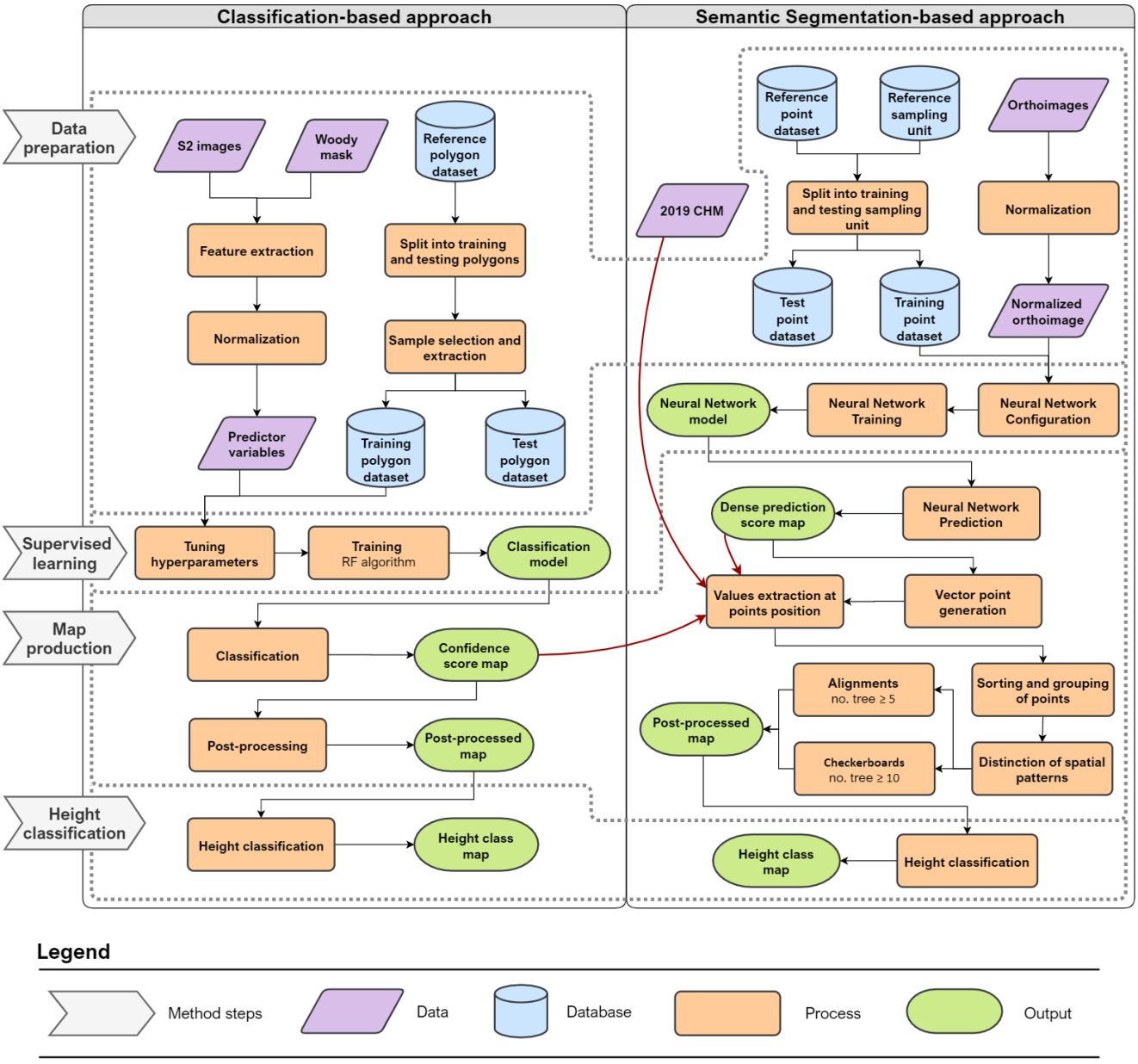


Figure 2: Process flowchart of the methods used for the mapping and characterization of the poplar resource in the province of Hainaut using remote sensing data, from data preparation to height classification.

Table 2: List of spectral indices extracted from super-resolved Sentinel-2 images, with their respective formulas and references.

Indice	Formula	Reference
Normalised Difference Vegetation Index	$NDVI = \frac{B08-B04}{B08+B04}$	((Rouse et al., 1974)
Simple ratio Moisture Index	$MSI = \frac{B11}{B08}$	(Hunt and Rock, 1989)
Simple ratio Disease Water Index 4	$DSWI_4 = \frac{B03}{B04}$	(Apan et al., 2003)
Normalised Pigment Chlorophyll ratio Index	$NPCRI = \frac{B04-B02}{B04+B02}$	(Peñuelas et al., 1994)
Normalised Burned Ratio Index	$NBRI = \frac{B08-B12}{B08+B12}$	(Key and Benson, 2006)
Shortwave Infrared Water Stress Index	$SIWSI = \frac{B08A-B11}{B08A+B11}$	(Fensholt and Sandholt, 2003)
Anthocyanin Reflectance Index	$ARI = \frac{1}{B03} - \frac{1}{B05}$	(Gitelson, Merzlyak, and Chivkunova, 2001)
Soil Adjusted Vegetation Index	$OSAVI = \frac{(1+0.16)*(B08-B04)}{(B08+B04+0.16)}$	(Rondeaux, Steven, and Baret, 1996)
Leaf Chlorophyll Index	$LCI = \frac{B08-B05}{B08+B04}$	(Datt, 1999)
Modified Chlorophyll Absorption in Reflectance Index	$MCARI = (B05 - B04) - 0.2 * (B05 - B03) * (\frac{B05}{B04})$	(Apan et al., 2003)
Red edge Index 2	$Red\,edge\,2 = \frac{B05-B04}{B05+B04}$	(Cloutis, 1996)
SWIR ratio	$SWIR\,ratio = \frac{B12}{B11}$	(Guerschman et al., 2009)
Poplar Index 1	$PI_1 = B11 - B12$	(Hamrouni Berkaoui, 2021)
Poplar Index 2	$PI_2 = B05 - (B11 + B12)$	(Hamrouni Berkaoui, 2021)
Poplar Index 3	$PI_3 = \frac{B05-(B11+B12)}{B05+(B11+B12}$	(Hamrouni Berkaoui, 2021)
Poplar Index 4	$PI_4 = B11 + B12$	(Hamrouni Berkaoui, 2021)

Reference polygon dataset

In order to train and test the ML algorithm, the reference polygon dataset (*cf.*, Section 2.2.2) was divided into a training and test dataset, from which the **sample selection** of pixels was performed. First, the reference polygon dataset was divided into classes (poplar and other), with approximately 80% of the geometry area intended for the training dataset and the rest for the test dataset. This first process prevents the influence of neighboring pixels, called spatial autocorrection, by splitting the reference dataset based on polygons [(Karasiak et al., 2019), as cited in Hamrouni Berkaoui (2021)]. In this way, training pixels come from training polygons, as well as test pixels come from test polygons. Then, the pixel **sampling strategy** consisted of selecting the same number of pixels for each class by limiting the number of pixels to the class containing the least number of pixels, here the poplar label. The number of pixels sampled for the classification is shown in Table 3.

Table 3: Number of sampled pixels for classification training and test.

	Label	Number of polygons	Available pixels	Kept pixels
Training	Others	5800	18,761,415	341,727
	Poplar	164	341,727	341,727
Test	Others	1933	6,260,266	89,093
	Poplar	54	89,093	89,093

Aerial images

The aerial images used for the semantic segmentation-based approach were the **2019 orthoimages** (*cf.*, Section 2.2.1) cropped to the study area. Since feature extraction is performed automatically by the DL algorithm, data preparation was only limited to **standardization** as in Section 2.3.1. Nevertheless, to avoid the consider potential outliers, the minimum and maximum values considered for standardization were respectively the 1^{st} and 99^{th} quantile.

Reference point dataset

As explained in Section 2.3.1, **sample selection** was done by point group within the reference point dataset (*cf.*, Section 2.2.2) to minimize spatial autocorrection. Indeed, the image interpretation process used to build the reference point dataset was done by sampling unit represented by polygons, the splitting of this reference dataset was done at the level of these polygons. The splitting consisted of approximately 80% of the geometric area for the training dataset and the rest for the test one. The number of reference points is shown in Table 4. The training point dataset was then rasterized, *i.e.*, a raster based on the 2019 orthoimages was built with a value of 1 for the pixels intersected by the points of reference, the rest was assigned a value of zero. It is important to underline that the effective area considered in the sample unit for learning is actually an area after applying a negative buffer of the patch size, in order to prevent the kernel bypassed areas outside the sample unit.

Table 4: Number of sampled points for neural network training and test.

	No. of points	**Proportion of points (%)**
Reference dataset	25,951	100
Training dataset	20,912	80.6
Test dataset	5,039	19.4

2.3.2 Supervised learning

Supervised learning for classification was performed by a RF algorithm while for semantic segmentation by a DLA architecture.

Random Forests

The RF algorithm is a ML based classifier widely used for FC type prediction (Kishore et al., 2016). This algorithm was preferred to others such as the Support Vector Machine (SVM), the Naïve Bayes classifier or k-nearest neighbors, because it has the best robustness to noise and **less sensitive to overfitting** (Sjöqvist, 2017; Sjöqvist, Längkvist, and Javed, 2020). In addition to the above-mentioned criteria, Hamrouni Berkaoui (2021) selected this algorithm for large-scale mapping of the poplar resource in France, due to a **lower processing time** than other algorithms, **performance little influenced by spatial or temporal scale** with the same hyperparameters, and a more **reliable evaluation observed** when test samples are independent from the training ones.

RFs proposed by Breiman (2001) is an ensemble supervised learning model, composed of decision trees for classification and regression. Decision tree training is based on an ensemble method called "bagging" or "bootstrap aggregation" which involves randomly sampling data from the training dataset to train the decision trees, then the final model is obtained by aggregating all these independent predictive models (see Figure 3). For lower correlation between decision trees, the RF algorithm additionally uses a feature randomness method by randomly sampling a number of features (variables) to train the classification model.

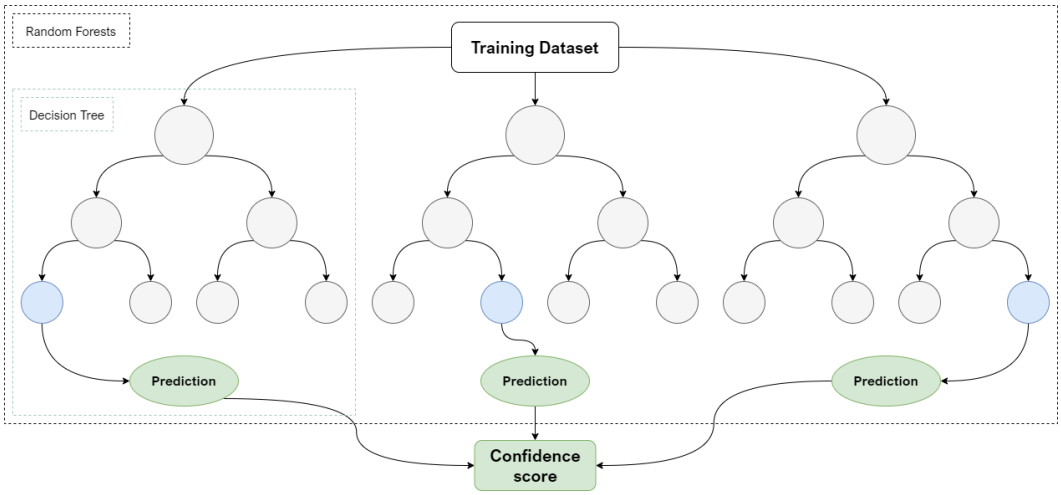


Figure 3: Illustration of the structure of a random forest algorithm and a decision tree.

The three main hyperparameters to set before training the RFs are: the maximum number of trees in the forest, the minimum node size for a split and the number of features tested at each node, which have been set with the following values, respectively 100, 25 and 12. These values were tuned using the R packages "randomForest" and "caret", training was then performed using the Shark Random Forests classification algorithm within OTB. The latter includes another hyperparameter to set the fraction of the training dataset for the calculation of the out of the bag (OOB) value which is a type of RF cross-validation method. The default value of 0.66 was kept, meaning that 0.66% of each bootstrap at the origin of the decision trees are used for performance evaluation and selection of variables of interest. However, the OOB technique is different from conventional validation method as the pixels used for validation in a decision tree can be randomly selected in another decision tree for training.

Deep Layer Aggregation

The DNN architecture tested before considering DLA was U-Net (Ronneberger, Fischer, and Brox, 2015), U-Net++ (Z. Zhou et al., 2020) and U-Net3+ (H. Huang et al., 2020). In fact, state-of-the-art results were observed in medical imaging but also for weed detection from crops (Rakhmatulin, Kamilaris, and Andreasen, 2021) or FC prediction (Bragagnolo, Silva, and Grzybowski, 2021; Giesen, 2022). Nevertheless, due to a high processing time to converge, which negatively affects the accuracy of the results, they were not kept. Therefore, a lighter detection architecture had to be considered for poplar detection, based on semantic segmentation and object-as-point detection method to limit the processing time compared to those producing a bounding box. X. Zhou, D. Wang, and Krähenbühl (2019) proposed a keypoint based approach for object detection named CenterNet, exhibiting higher speed than YOLOv3 with accuracy levels at Faster-RCNN-FPN. Four architectures were used as the backbone of the model: ResNet-18, ResNet101, DLA-34 and Hourglass-104. DLA-34 offered the best speed/accuracy trade-off. The performance evaluation was measured using the MS COCO dataset (Lin, Maire, et al., 2015). This trend has been confirmed by further research conducted on the CenterNet model (Xu, 2021; Zhao and Yan, 2021). The architecture has also been used and modified outside of the CenterNet model to perform CV tasks with accuracy equivalent to VGGNet, ResNet, and DenseNet-based architectures deployed on low performance devices such as a drone or laptop (F.-T. Wang et al., 2021).

For the reasons mentioned above, we chose the DLA-34 architecture based on DLA structure proposed by Yu et al. (2018). The goal of the study was to better fuse semantic and spatial information for recognition and localization tasks. Two DLA structures were then proposed (see Figure 4): iterative deep aggregation (IDA) focuses on merging resolutions and scales, and hierarchical deep aggregation (HDA) focuses on merging features from all modules and channels. The former is inspired by Feature pyramid networks (FPNs) proposed by Lin, Dollár, et al. (2017), while the latter is inspired by Densely connected networks (DenseNets) proposed by (G. Huang et al., 2018). IDAs differs from linear aggregation as in Fully Convolutional Networks (FCNs) proposed by Shelhamer, Long, and Darrell (2016), FPNs, and U-Nets by starting at the smallest and shallowest scale and then iteratively merges the larger and deeper scales. The study concludes that DLAs offer improved performance, parameter count, and memory utilization in comparison with competing architectures. The DLA34 network consists of basic blocks with 16, 32, 64, 128, 256, and 512 channels for stages 1 to 6 and an aggregation depth of 1, 2, 2, and 1 for stages 3 to 6.

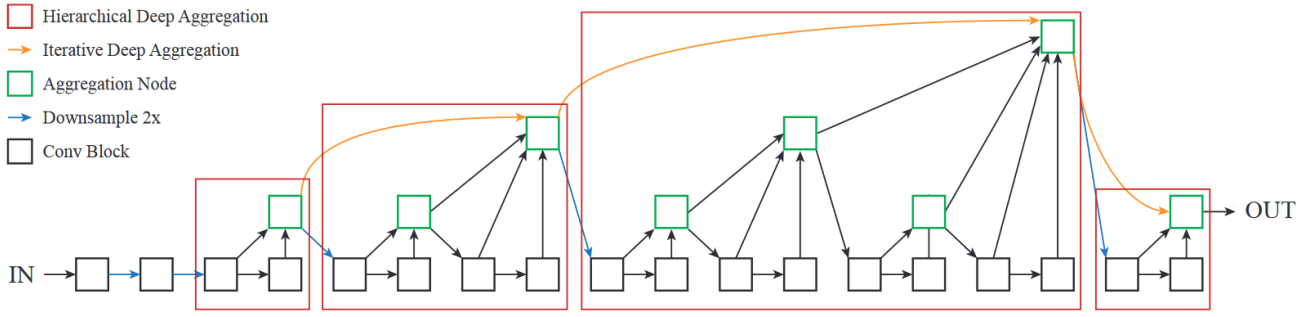


Figure 4: Illustration of the architecture of a DLA neural network with Hierarchical Deep Aggregation (HDA) and Iterative Deep Aggregation (IDA), from Yu et al. (2018).

The CNNs in this project used batch normalization layers and the Rectified Linear Unit (ReLU) activation function (see, Appendix A). The patch size was a 512 pixel square, the mini-batch size was 2 patches and the batch size was 2 mini-batches, the limit imposed by the GPU required the use of mini-batches in order to constitute a sufficient batch size for the descent gradient to converge to the global minimum without considering the local minimum. The loss function considered by gradient descent is a mean square error (MSE), whose equation is:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2 \tag{2.1}$$

where n is the number of pixels contained in a minibatch, y_i is the vector of observed values and $\tilde{y}_i$ is the vector of predicted values, the error being the difference between these last two. The training rate was constant at 0.0001, the maximum number of epochs was 1,000(*cf.*, Appendix X). Then, during the neural network training, data augmentation was performed by vertical and horizontal flip.

2.3.3 Map production

Classification-based approach

The classification model resulting from supervised learning by the RF algorithm produces a raster composed of a binary classification per pixel (other or poplar) with a confidence score per pixel corresponding to the proportion of votes for the majority class. Hence, the classified pixel values range from 0 to 1, respectively a large number of votes for the "other" class and a large number of votes for the "poplar" class.

In order to complete the map production, for a classification confidence score threshold chosen at 0.8, regularization was applied with a kernel radius of 22 pixels and an isolation threshold of 20 pixels, followed by sieving with a threshold of 80 pixels, *i.e.* a minimum area of 5 a.

Semantic Segmentation-based approach

The semantic segmentation was produced from the neural network prediction, resulting in a raster composed of a dense prediction of the "poplar" label for each pixel. The dense prediction task refers in CV to the prediction of a label for each pixel in the image (Sercu and Goel, 2016). The resulting raster from the semantic segmentation can then be understood as a density or heat map (Seddati et al., 2017; Zhao and Yan, 2021) with a pixel value ranging from 0 to 1, increasing progressively the closer the pixel is to the center of a poplar tree. The density prediction from the neural network was smoothed using a 2d Gaussian filter, whose equation is:

$$G_{2D}(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{2.2}$$

where the value of sigma (σ) was 1.5 and the window size was 11 defining x and y.

Post-processing of this semantic segmentation for poplar detection for map production involved point generation, point sorting, point grouping, and a threshold on the number of points per group.

First, vector points marking the position of poplars were generated by applying a local maximum detection algorithm, with a minimum dense prediction value of 0.3 and a minimum distance between points of 4 m as parameters.

Then, the sorting and grouping of points aims to drastically reduce the number of False Positives (FPs). This was done by first extracting the following maximum values in a 2 m buffer for each point: the confidence score of the classification (*cf.*, Section 2.3.5), the dense prediction score by the neural network model, and the height (2019 pCHM).

The point sorting was done at the level of a partition in 3,257 tiles of the province of Hainaut for a width of 1,254 m and a height of 1,027 m each; a variable was added to the points which is the average confidence score for the points of the tile. Based on this average confidence score per tile, different confidence scores and dense prediction scores were used to retain points (see Table 5). This method was necessary to not consider all the province of Hainaut

in a homogeneous way, by being less strict in retaining points in areas with the highest probability of poplars.

Table 5: Confidence score (c) and dense prediction score (d) used to retain a point according to the average confidence score value per tile (a).

For an average confidence score	**Values used to retain a point**
$a < 0.1$	$(c \geq 0.9 \,\&\, d \geq 0.6)$
$0.1 \leq a < 0.4$	$(c \geq 0.9 \,\&\, d \geq 0.5)$
$0.4 \leq a < 0.65$	$(c \geq 0.8 \,\&\, d \geq 0.4)$
$0.65 \leq a < 0.85$	$(c \geq 0.7 \,\&\, d \geq 0.35)$
$a \geq 0.85$	$(c \geq 0.9 \,\&\, d \geq 0.4)$

An algorithm originally designed by Bolyn, Lejeune, et al. (2019) and Bolyn, Latte, Fourbisseur, et al. (2020) was used for point grouping. The idea is to use the spatial characteristic of poplar plantations for a second point sorting. Points with a maximum distance of 10 m from each other and a maximum height difference of 8 m were considered to be part of the same group (a poplar plantation), thus eliminating isolated points or points with no spatial proximity (x, y, z) to these neighbors.

Finally, a minimum number of trees per group had to be considered to eliminate isolated groups too small to constitute a poplar plantation. However, aligned plantations, which are more difficult to identify, generally group fewer points, so in order not to eliminate them, it was necessary to differentiate the alignment and checkerboard spatial patterns. On the basis of polygons built from groups of points, two elements were considered, a ratio between the surface intersected by the polygon with a 5 m buffer around the boundaries of the cadastral parcels of the province of Hainaut and the total surface of the polygon as well as a ratio between the quartile 2 and 4 of a distance vector measured between the boundaries of the polygons. Indeed, alignment patterns are found along cadastral parcel boundaries and are characterized by a small distance ratio separating the internal boundaries of the polygon. Alignment patterns considered with a ratio based on cadastral boundaries of 0.6 and 0.15 based on internal distance consist of groups of at least 5 points while checkerboard patterns have at least 10 points.

2.3.4 Height classification

A height classification was used to characterize the poplar resource. By taking into account the growth of the forest species and in line with a study on the delineation of stand units (Koch, Straub, et al., 2009), the height classes listed in Table 6 were established. These height classes were informatively associated with a stand development stage based on the Koch, Straub, et al. (2009) and Kimmins (2003) work.

In areas classified as poplar by the classification-based approach, the height data from the photogrammetry applied to the 2018 orthoimage was low-pass filtered (average filter) with a square kernel of 15 pixels on each side and then classified according to the height classes

Table 6: Height classes (h) for characterization of poplar resource related to stand development stage.

Development stage	Height class
Sapling	$3 \leq h < 8\,m$
Pole	$8 \leq h < 16\,m$
Mature trees	$16 \leq h < 24\,m$
Old trees	$24\,m \geq h$

above. Before being polygonized, the height class raster was post-processed by regularization and sieving[3]. Regarding the data from the semantic segmentation-based approach, the average height of the points per group was classified by height classes. As a reminder, the height per point is derived from an extraction of the maximum height value provided by the 2019 pCHM[4], within a 2 m buffer around the point (*see* Section 2.3.3). The observed height class per sample plot of the WFI is the average of the dominant heights measured in the field within the plots.

2.3.5 Accuracy assessment

The results of the two approaches to mapping and characterizing the poplar resource in the province of Hainaut, as well as the methods and datasets used to evaluate them, are summarized in Table 7. These methods (confusion matrix, proportion of labels, false positives and ratio of properly labeled height classes) are detailed below, with respect to the datasets mentioned in Table 7, described in the Sections 2.2.2 and 2.2.3.

Table 7: Overview of the data produced by the two approaches for the mapping and characterization of the poplar resource int the province of Hainaut, with the method and dataset used to evaluate these results.

Approach	Output data	Evaluation method	Dataset
Classification-based	Confidence score map	Confusion matrix	Test polygons
		Proportion of labels	Image interpretation points
			WFI plots
	Post-processed map	Confusion matrix	Test polygons
		Proportion of labels	Image interpretation points
			WFI plots
	Height class map	Ratio of properly labelled height classes	WFI plots
Semantic segmentation-based	Dense prediction score map	Confusion matrix	Test points + WFI plots
		False positives	DNF GIS field
	Post-processed map	Confusion matrix	Test points + WFI plots
		False positives	DNF GIS field
		Hausdorff Distance	Test points + WFI plots
	Height class map	Ratio of properly labelled height classes	WFI plots

Confusion matrix

As the name suggests, this evaluation method is based on a confusion matrix. In the case of binary classification or object detection according to the approach considered in this

[3] a threshold below which a group of pixels is not taken into account
[4] combination of DTM from the 2014 LIDAR and a DAP applied to the 2019 orthoimage

project, the latter allows identifying correct predictions from erroneous ones. Table 8 shows a confusion matrix for a binary classification per pixel.

Table 8: Composition of a confusion matrix for a binary classification, according to the predicted class for each pixel compared to the observed one, 4 classification results are obtained: True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN).

		Produced labels	
		Other	Poplar
Reference labels	Other	True Negative *(TN)*	False Positive *(FP)*
	Poplar	False Negative *(FN)*	True Positive *(TP)*

For the classification-based approach, the confusion matrices are derived from the pixels produced and observed in the test polygon dataset. Confusion matrices for the semantic segmentation-based approach are obtained from an Intersection over Union (IoU) between points in the test point dataset combined with those in the WFI sample plots and those predicted. Before combining the two databases mentioned above, two WFI plots located on the training point dataset were removed, leaving 22 plots containing 399 points, for a combined database of 5,437 points (5,038 + 399). The IoU (*see* Figure 5), is used to allow the computation of evaluation metrics for object detection by a bounding box. The observed points and those predicted by the segmentation-based approach were then surrounded by a circular buffer of 3 m radius, this buffer acting as a bounding box. A True Positive (TP) class was assigned to a predicted bonding box, if the ratio of the area intersected with the observed bonding box to the area of the union of these two bonding boxes was at least 0.3. The points observed in the semantic segmentation approach considered as positive differ from the background of the image which is then negative, this background not being predicted; the confusion matrix of this approach does not contain True Negatives (TNs).

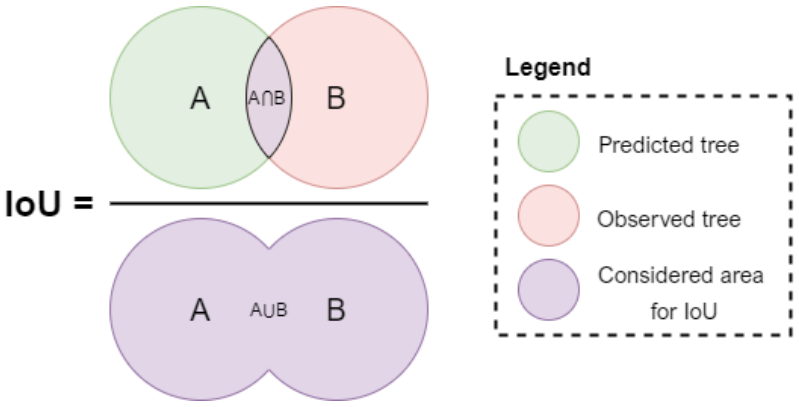


Figure 5: Equation of the Intersection over Union (IoU) between a predicted point and an observed point respectively bounded by a bounding box A and B.

The evaluation metrics computed on the basis of the above-mentioned confusion matrices are shown in Table 9. Regarding the "poplar" label, precision is the ratio of correct predic-

tions to the total number of predictions, recall is similar but considers the total number of observations as divisor, F1-score is the harmonic mean between precision and recall. Finally, the kappa refers to Cohen's kappa, indicating how much better the classification is than a simple random classification based on the frequency of each class. It is important to note that the latter includes TNs in its equation. This evaluation metric can therefore only be applied to the confusion matrix of the classification-based approach, because the semantic segmentation-based approach does not contain TNs.

Table 9: Evaluation metrics with the defining equation used in this project, where k is the confidence score threshold, pop refers to the poplar label, p_o is the observed agreement, p_e is the expected agreement, TP, FP, TN, and FN are the number of True Positive, False Positive, True Negative, and False Negative respectively from the confusion matrix.

Evaluation metric	**Formula**	**Evaluation metric**	**Formula**
Kappa	$\kappa = \frac{p_o - p_e}{1 - p_e}$	Precision	$P_{pop}(\tau) = \frac{TP}{TP+FP}$
F1-score	$F1_{pop}(\tau) = \frac{2 * P_{pop}(\tau) * R_{pop}(\tau)}{P_{pop}(\tau) + R_{pop}(\tau)}$	Recall	$R_{pop}(\tau) = \frac{TP}{TP+FN}$

$$N = TP + FP + TN + FN$$

$$p_o = \frac{TP + TN}{N}$$

$$p_e = \left(\frac{TP + FP}{N} * \frac{TP + FN}{N}\right) + \left(\frac{TN + FP}{N} * \frac{TN + FN}{N}\right)$$

Proportion of labels

The distribution between "poplar" and "other" labels within the woody mask was measured by an image interpretation step, a set of points evenly spread over the mask were interpreted from the 2018 orthoimage (*see*, Section 2.2.1). Then, the values of the classification confidence score map and the map resulting from the post-processing of the latter were extracted at the points used for image interpretation. By this way, it is possible to compare the distribution of labeled points by the image interpretation with the predictions of the classification-based approach. The area per label is obtained by interpolating the proportion of labels in the woody mask and the area of the latter.

The proportion of labels was also observed in the WFI sample plots, where all pixels are assumed to be classified with a "poplar" label. To make the evaluation objective and independent, two WFI plots intersecting the training polygon dataset were removed, so that in the remaining 22 plots.

False Positives

To better evaluate changes in the number of FPs between the dense prediction score map and the post-processed map from the semantic segmentation-based approach, the predicted points in the province of Hainaut were intersected with the DNF GIS data. These data had been used to supplement the reference polygon dataset used by the classification-based approach (*see* Section 2.2.3). Only DNF GIS data excluding any poplar presence were used, for an area of 15,628 ha. Therefore, the predictions from the semantic segmentation-based approach intersected with these data are FPs.

Hausdorff Distance

To quantify the location error of the points marked as detected poplars, the Hausdorff distance (Nutanong, Jacox, and Samet, 2011; Ribera et al., 2019) was applied between the TPs produced and those predicted observed. This distance is the average of the nearest distances measured between the predictions and the observations mentioned above. Observations come from the test point dataset combined with the WFI WFI sample plot database, the latter containing 399 points after removing those crossing the training point dataset.

Ratio of properly labelled height classes

A ratio of properly labeled height classes was calculated from the height class maps (*cf.*, Section 2.3.4) and the average dominant tree heights measured within the WFI sample plots. This ratio is therefore calculated on a pixel basis for the classification-based approach and on a poplar (point) basis for the semantic segmentation-based approach. For the second approach, only the correct predictions, *i.e.* the TPs are compared with the respective observation.

3 Results

3.1 Classification-based approach

3.1.1 Confidence score map

The confidence score map from the supervised pixel-based classification within the woody mask is a raster consisting of pixels with a spatial resolution of 2.5 m with values ranging from 0.00 to 1.00. A value close to zero indicates a large number of votes for the other class and vice versa for the poplar class.

However, in order to evaluate the distribution of predicted labels as well as their relative confidence score within the test polygon dataset, Figure 6 shows the percentage of confidence scores. Each label ranked from 50 to 100, from a low number of votes to a high number of votes for the majority class. It can be observed that the balanced distribution of the two labels in the test polygon dataset seems to be satisfied, with a majority of high confidence scores. The evaluation metrics that will be mentioned later will help to confirm this trend or not, as well as to verify if the predictions are reliable.

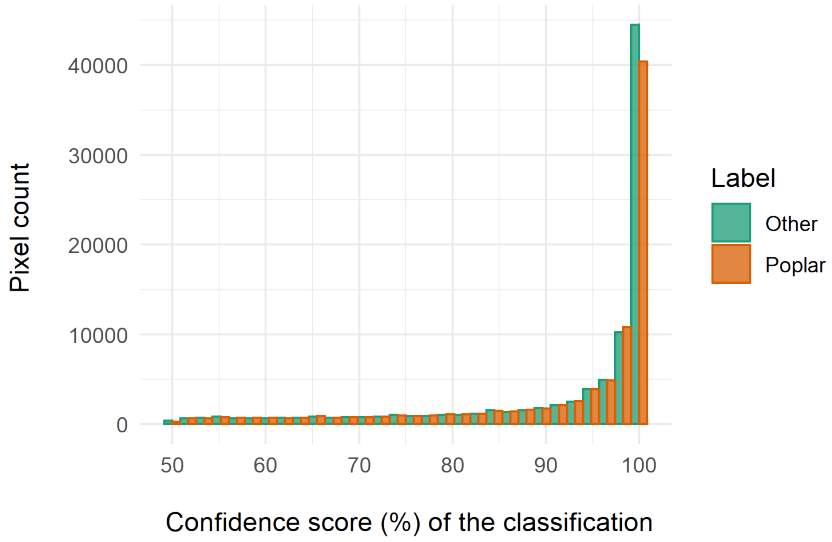


Figure 6: Histogram of the number of pixels per confidence score (proportion of votes for the majority class) in the test polygon dataset from pixel-based supervised classification of super-resolved Sentinel-2 images using the Random Forest algorithm.

The evaluation metrics of the confusion matrix obtained based on a threshold (from 0.50 to 1.00 with an interval of 0.05) applied to the classification confidence score of the 178,186 pixels (2*89,093 pixels) in the test polygon dataset are shown in Figure 7. The performance of

the kappa and F1-score remain stable but decrease slightly until a threshold on the classification confidence score of 0.80 is reached. Beyond the mentioned threshold, the performance drops. Indeed, the F1-score being the harmonic mean of precision and recall for the poplar label; increasing the threshold slightly improves the former but significantly deteriorates the latter. In other words, increasing the threshold has no significant impact on the number of FPs but significantly increases the number of False Negatives (FNs).

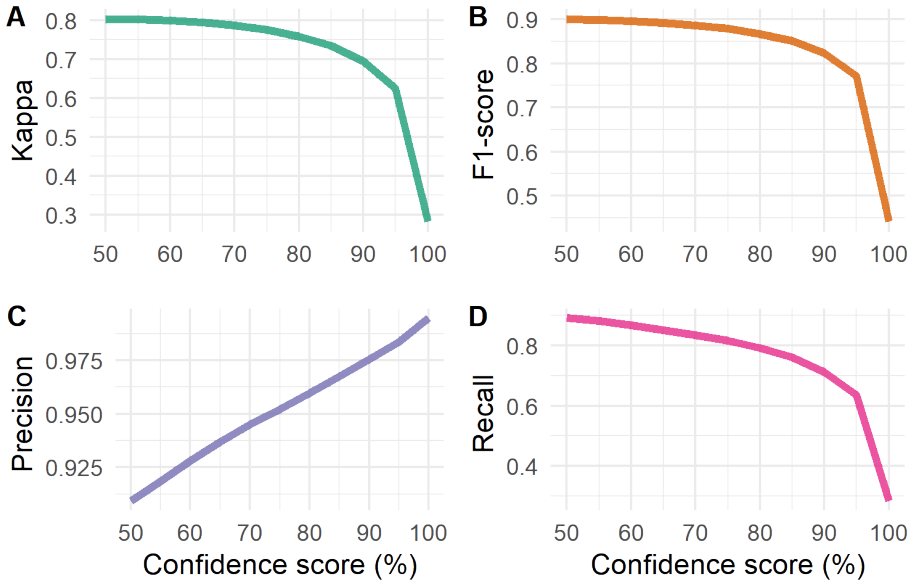


Figure 7: Evaluation metrics computed from the confusion matrix in the test polygon dataset, based on a threshold applied to the classification confidence score. The evaluation metric in A is the Cohen's kappa, then respectively from B to D the F1-score, Precision, and Recall for the predicted labels as poplar.

Based on the 2,705 points used for image interpretation, the distribution between the two classes derived from the observation (reference) and the values derived from the confidence score are presented in Table 10. The area of the poplar resource observed by interpretation of the 2018 orthoimage, is 5,900 ha. The predicted areas of the poplar resource based on the confidence score map decrease when the threshold increases. Finally, a threshold between 0.90 and 0.95 should be considered to tend towards the observed surface.

Table 10: Estimation of the area (in ha) per label from the points used for image interpretation n and extraction of the confidence score map values at these points for confidence score thresholds ranging from 0.50 to 1,00 with an interval of 0.05. The estimated areas per label are obtained by extrapolating the proportion occupied per label from the 65,595 ha of the woody mask.

Label	**Estimated area (ha) and proportion (%) of pixel per label in the image interpretation points**											
	Reference	Confidence level of the classification										
		0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	100
Other	59,702 *(91.02)*	42,485 *(64.77)*	44,110 *(67.25)*	45,783 *(69.80)*	47,335 *(72.16)*	49,299 *(75.16)*	50,973 *(77.71)*	53,373 *(81.37)*	55,507 *(84.62)*	57,835 *(88.17)*	60,551 *(92.31)*	64,504 *(98.34)*
Poplar	5,893 *(8.98)*	23,110 *(35.23)*	21,485 *(32.75)*	19,812 *(30.20)*	18,260 *(27.84)*	16,296 *(24.84)*	14,622 *(22.29)*	12,222 *(18.63)*	10,088 *(15.38)*	7,760 *(11.83)*	5,044 *(7.69)*	1,091 *(1.66)*

The distribution of labels within the 24 WFI sample plots of 24,429 m^2 is detailed in Table 11. For a reference distribution of 100% for the label "poplar", the distribution for this label was maximal for the threshold of 0.50.

Table 11: Proportion of pixels (in %) by predicted labels in Walloon Region Forest Inventory sample plots database, according to thresholds on the classification confidence score ranging from 0.50 to 1.00 with an interval of 0.05.

Label	**Proportion (%) of pixels per label in the Walloon region Forest Inventory plots**											
	Reference	Confidence level of the classification										
		0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	100
Other	0	19.93	21.26	22.46	23.94	25.65	27.29	29.36	31.76	35.55	42.18	78.33
Poplar	100	80.07	78.74	77.54	76.06	74.35	72.71	70.64	68.24	64.45	57.82	21.67

3.1.2 Post-processed map

Following the assessment of the above classification confidence score map, a threshold of 0.80 was selected for the post-processing step. Indeed, this threshold was the optimal balance between a high kappa, a high F1-score, while being close to the reference poplar area indicated by the image interpretation. This should not be too high to remain close to the reference distribution within the WFI plots. The cover occupied by the poplar resource estimated by the classification-based approach in the province of Hainaut is **10,201 ha** (*cf.*, Appendix C.1).

The evaluation metrics computed from the post-processed confidence score map with a 0.80 threshold are: kappa=0.856, F1-score=0.923, precision=0.987 and recall=0.867. An improvement of these evaluation metrics is thus observed, for the same confidence score threshold before post-treatment the kappa was 0.758 and the F1-score 0.867. The estimated area for the "poplar" label from the image interpretationpoints was 10,236 ha, with a distribution of 83.52% for the "other" label and 16.48% for the "poplar" label. A decrease of 1,986 ha was observed after post-treatment, but the poplar resource remains overestimated compared to the reference area. Finally, the distribution of the "poplar" label within the WFI plots was 83.52%, increased by 11.80% compared to the non-post-processed map.

3.1.3 Height class map

The mapping of the poplar resource was followed by a characterization through classifying it into height classes, the results obtained are presented in Figure 8.

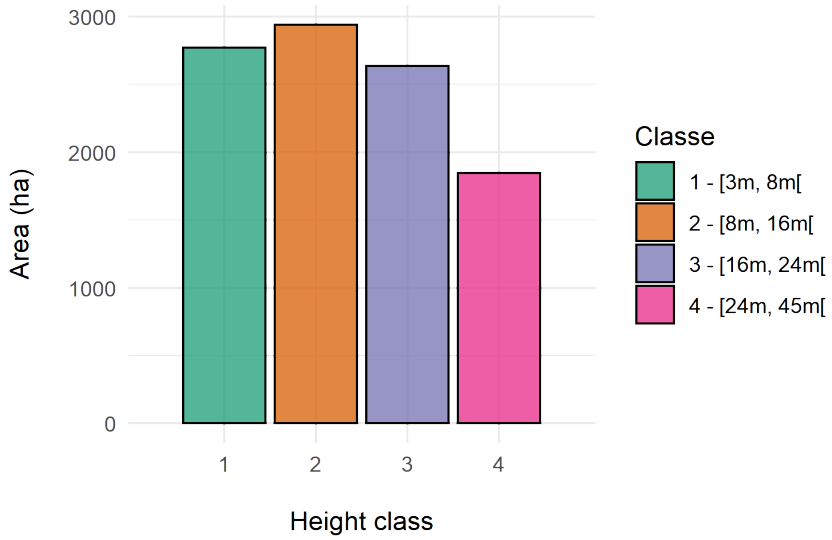


Figure 8: Histogram of the occupied area by height class in the poplar resource area estimated by the post-processed map in the classification-based approach. The height classes from 1 to 4 correspond respectively to the stand development stages: Sapling, Pole, Mature Trees and Old trees.

Within the WFI plots, the ratio of properly labeled height classes was only 50.4%. The number of pixels per predicted and observed height class is indicated in Table 12. Because of FN and the absence of field-measured height, 19 were considered, for a total of 14,325 pixels.

Table 12: Number of predicted and observed pixels by height class in the Walloon Region Forest Inventory sample plots.

Height class	No. of predicted pixels	No. of observed pixels
1	374	0
2	5152	2962
3	2562	4180
4	6237	7183

3.2 Semantic Segmentation-based approach

3.2.1 Dense prediction score map

The dense prediction score map is a vector layer of points obtained from the semantic segmentation of the 2019 orthoimage. This map is composed of a set of 1,548,152 points which mark the position of poplars.

Within the test sample units and the WFI sample plots, the evaluation metrics computed from the confusion matrices obtained from the 5,437 (5,038 + 399) points is shown in Figure 9. The decrease in these values suggests that increasing the threshold eliminates many TPs and thus strongly alters the evaluation metrics. This finding motivated the post-processing according to the method proposed in Section 2.3.3.

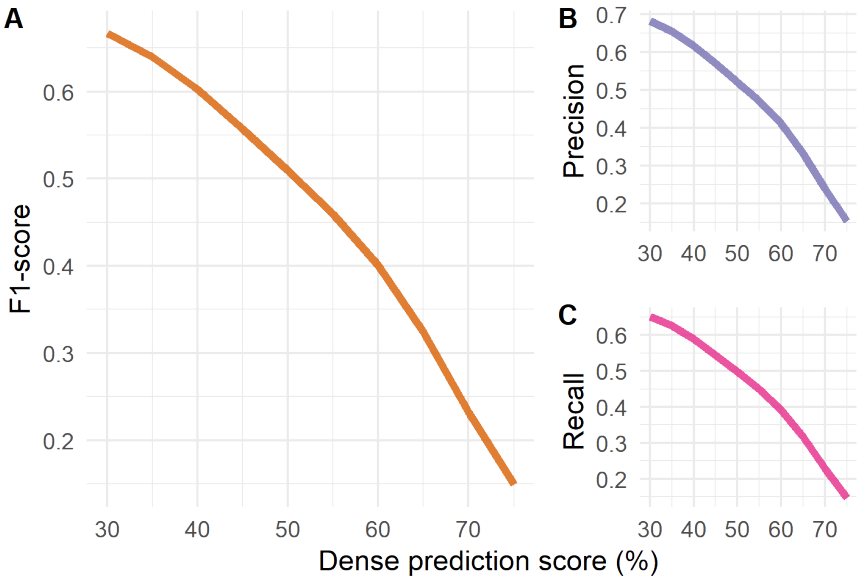


Figure 9: Evaluation metrics computed from the test point dataset merged with the Walloon Region Forest Inventory sample plots, based on a threshold applied to the dense prediction score. The evaluation metrics from A to C, respectively, are F1-score, precision and recall for labels predicted as poplar.

3.2.2 Post-processed map

The first sort applied to the dense prediction score reduced the number of points by 44.68%, *i.e.*, obtaining a vector layer of 856,469 points. Full post-processing returns 11,595 point groups, for a total of 527,528 polars, which represents a 65.93% decrease in the number of points compared to the dense prediction score map. The evaluation metrics for these results are 0.653, 0.776, and 0.564 for F1-score, precision, and recall, respectively. However, a lower recall indicates that some of the TPs have also been removed.

In light of these results, post-processing using the confidence score and dense prediction score and followed by a point SA; significantly reduced the number of points while maintaining a stable F1-score, thus helping to eliminate a significant number of FPs.

To better measure the significance of the number of removed FPs, intersected points between the two maps and the 15,628 ha of poplar-free areas from the DNF GIS data was considered. An evolution of the number of points from 98,058 to 18,189 was observed, *i.e.* a reduction of 81.45% of the number of FPs.

Finally, the Hausdorff distance measured between the post-processed map and the test point dataset merged with the WFI plots is 1.06 m. This measurement indicates that the location of TPs provides an excellent accuracy for forestry use.

3.2.3 Height class map

The height class map from the semantic segmentation-based approach, shown in Figure 10, highlights a significant imbalance between classes, especially for the height class one, the sapling development stage.

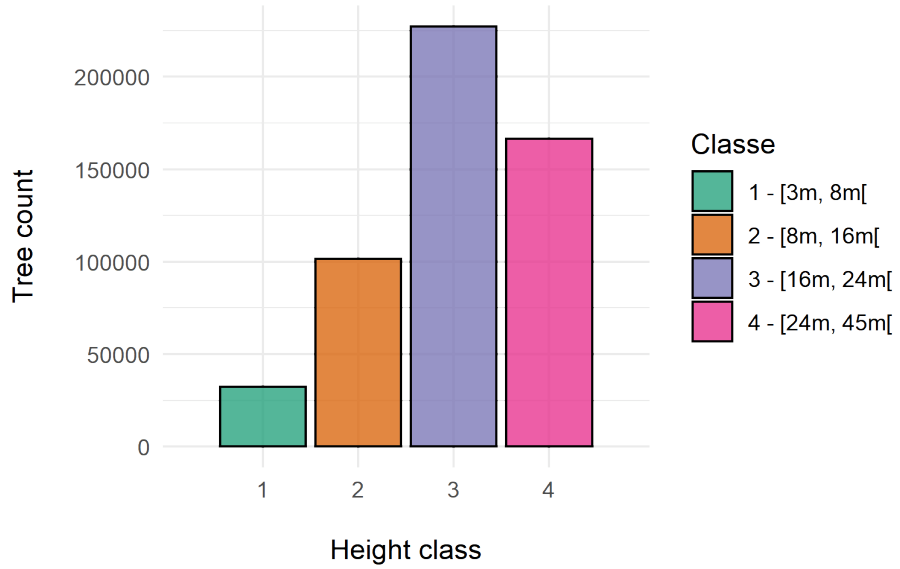


Figure 10: Histogram of the number of poplars by height class among the poplar resource estimated by the post-processed map of the semantic segmentation-based approach. Height classes 1 to 4 correspond respectively to the stand development stages: Sapling, Pole, Mature Trees and Old trees.

The ratio of properly labeled height classes among the 204 poplars successfully detecting the reference one (*i.e.* TPs) in 20 WFI plots was 0.6863. As explained in Section 3.1.3, not all plots were considered due to missing height measurement or TPs.

Table 13: Number of produced and observed points by height class in the Walloon Region Forest Inventory sample plots.

Height class	**No. of predicted points**	**No. of observed points**
1	13	0
2	18	82
3	86	35
4	87	87

4 Discussion

The results addressing the objectives of this master's thesis are discussed in this section to answer the research questions raised in Section 1.4.1. The discussion is divided into 4 parts: poplar resource mapping, poplar resource characterization, main personal contributions and finally perspectives.

4.1 Poplar resource mapping

4.1.1 Classification-based approach

Added value of super-resolved Sentinel-2 images

Due to a large number of pixels in the reference dataset and the splitting of the pixels constituting the latter by polygons to prevent spatial autocorrection effects (*cf.* Section 2.3.1), it allows in principle, to consider these values in a reliable way. Keeping this in mind, we can then argue the results of the post-processed map of the poplar resource from the classification-based approach are satisfactory. In fact, a kappa of 0.856 indicates that a large part of the TPs and TNs pixels are not due to randomness and that a F1-score of 0.923 induced an important proportion of TPs.

To better appreciate the accuracy of these results, the evaluation metrics obtained by Bolyn, Latte, Colson, et al. (2020) and Hamrouni Berkaoui (2021) for their poplar resource mapping will be presented. It should be considered that these were obtained with different datasets for different conditions. For the "poplar" label, the first obtained a F1-score, a precision and a recall of 0.746, 0.843, 0.792 respectively, for the pilot site in Wallonia and of 0.570, 0.831, 0.676, for the pilot site in France. This map was produced from S2 images in northern France and in Wallonia. The second obtained a F1-score ranging from 0.895 to 0.993 by local S2 image classification and a F1-score ranging of 0.900 and 0.970, after domain adaptation and active learning on the scale of the French territory. It would appear that the super-resolution process has added value to the accuracy of the poplar resource mapping and on the other hand the additional elaboration of ML approaches can potentially further improve this map.

Overestimation of the poplar resource

As seen above, the poplar resource map from the classification-based approach is relatively reliable. Correct predictions are observed for the poplar stand at the old tree stage and is not influenced by the spatial pattern as observed at points A and D in Figure 11. However, these results are contrasted by those of the image interpretation, indicating a difference of 4,308 ha (42% of the predicted poplar resource area) between the predicted and observed areas. Consequently, the results of this map overestimate the area of poplar FC.

As raised by Bolyn, Latte, Colson, et al. (2020), areas of mixed stands can bias the prediction (*see* Point B in Figure 11). Actually, forest stands not considered as poplar plantations can contain natural poplars, which can lead to a classification of these pixels as "poplar", therefore being FPs. Some species can also present spectral similarities with poplar, for instance willows of the same family which are generally found in the same bioclimatic zone. Surprisingly, we observe confusions with softwoods, similar to pines or larches by the interpretation of the images as in Point E in Figure 11.

The process of image interpretation is not indisputable, old poplar plantations or areas of non-pure poplar may have been confused with the label "other" by the operator. This may slightly decrease the observed poplar resource.

Finally, despite an overestimation of the poplar resource, 16.5% of the 3,423 pixels intersected by the WFI plots were not predicted as "poplar" label. The recall obtained from the confusion matrix is 0.867, *i.e.* 13% of the poplars observed in the test dataset were not predicted. Nevertheless, the field data collection period of the considered WFI plots ranges from 2014 to 2022, so the ground truth might be biased by the temporal difference. Alternatively, it could mean that the evaluation metrics overestimate the accuracy of the map.

Limitations for young poplar plantation detection

Even though the poplar resource is overestimated, it does have FNs as discussed above. These are particularly noticed at the Sapling stage, as shown at point C and on the left side of the canal at point E in Figure 11. There are two reasons for this limitation, one is that the tree crowns are not extended enough to be reflected by the pixel, the other is that the woody mask does not take them into account.

Dependence on the woody mask

The classification-based approach to mapping the poplar resource depends on the woody mask. Consequently, the latter has a significant influence on the predicted poplar area within the province of Hainaut, as well as on the detection of young poplar plantations. The quality of the woody mask is itself dependent on the accuracy of the height data. In the case of the 2018 pCHM, small crowns may potentially not be detected.

It should be pointed out that we consider the poplar resource from a height of 3 m, thus not considering the juvenile stage of development. Not detected young poplars at point C in Figure 11, may also be excluded from the definition of the poplar resource as considered in this project.

4.1.2 Semantic segmentation-based approach

Accurate location of detected objects

The post-processed map of the poplar resource predicted by the semantic segmentation-based approach has a remarkable accuracy in localizing poplar trees. In fact, an average distance of 1 m is obtained between the TPs predictions and those observed. Since the distance is only calculated between the TPs and observations, it is precisely a derivation of the Hausdorff distance. This accuracy of localization is confirmed by image interpretation of the post-processed vector layer as observed at point A in Figure 12.

Underestimation of the poplar resource

It was identified that the recall value for this map is 0.56, which implies a large number of FNs, thus underestimating the poplar resource. While poplar plantations at a young stage of development (young sapling) are sometimes partially detected (*see* point B in Figure 12), they are rarely fully detected. Even more pronounced is the poor detection of plantations at an advanced stage of development (Old Trees) (*see* point F in Figure 12). We also observe that spatial patterns influence the quality of the prediction, as alignments were difficult to detect in this project. The main difficulty for the model is to detect alignments with a young or old stand development stage (*see* point C in Figure 12). Finally, post-processing the dense prediction score map to reduce FPs, may remove TPs as well, thus contributing to an underestimation of the poplar resource.

Sensitivity to the reference dataset

A Deep Learning-based approach requires a larger reference dataset than a Machine Learning-based approach (L. Zhang, L. Zhang, and Du, 2016), making the latter more sensitive. Poplar alignments, rarer in the landscape of the province of Hainaut compared to checkerboard spatial patterns, were consequently less frequent in the reference dataset, thus probably under-trained. This necessity of the Deep Learning-based approach for a large reference dataset is a limitation of its use compared to the Machine Learning-based approach.

Additional data dependency

Although post-processing has removed many FPs, some still remain (*see* points B and E in Figure 12). However, this step depends on additional data, namely: a 2019 pCHM and the confidence score map. Their quality is crucial for post-processing and has been essential in improving the accuracy of the map.

Specific knowledge

The implementation of the semantic segmentation approach was possible thanks to Nicolas Latte, having specific skills for DL. This level of knowledge was necessary for the implementation of this type of approach, contrary to an approach based on classification. Moreover, even with a powerful GPU and CPU, the DL approach requires considerably more time than a ML-based approach.

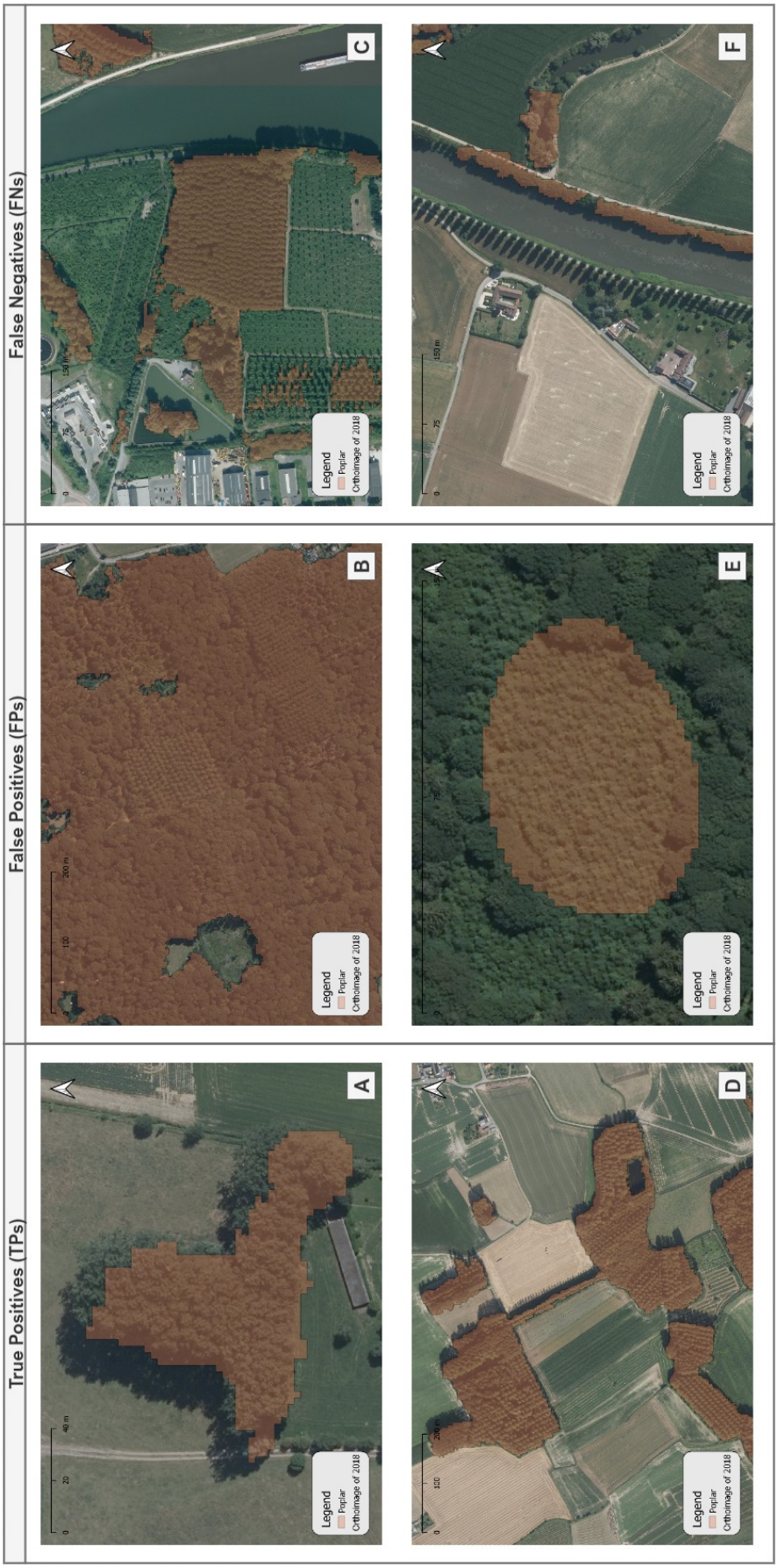


Figure 11: Illustration of TPs, TPs and FNs from the Confidence score map, of the classification-based approach.

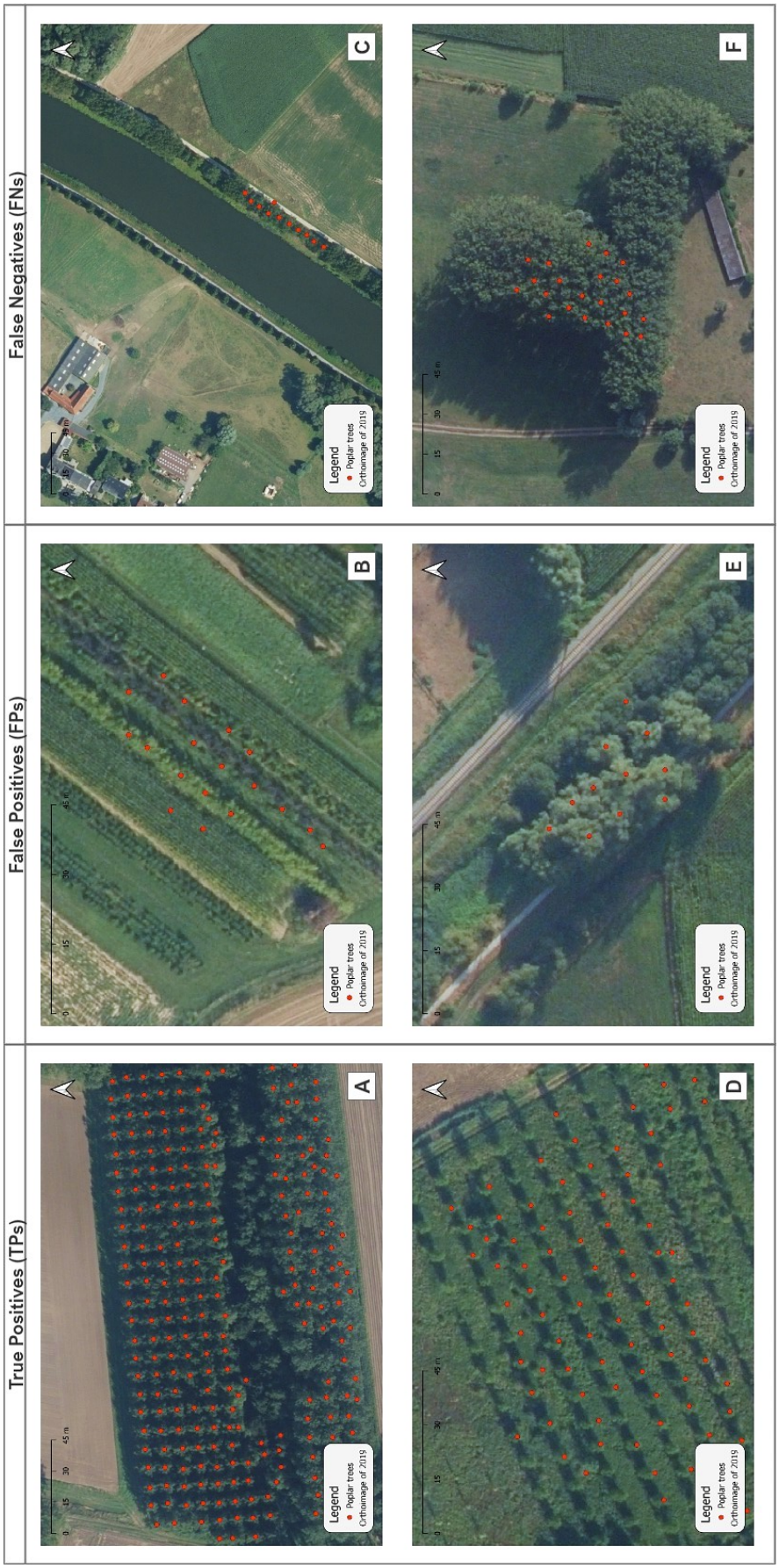


Figure 12: Illustration of TPs, TPs and FNs from the Dense prediction score map, of the semantic segmentation-based approach.

4.2 Poplar resource characterization

Poplar resource characterization using the classification-based approach and the semantic segmentation-based approach is distinguished by processing at the pixel level for one and the object level for the other. The height class map for the first approach is derived from the pixel processing of the 2018 pCHM (low-pass filter, height classification, and post-processing) and for the other the tree height of a group of poplars is the average of the height extraction of the 2019 pCHM per poplar in the group.

4.2.1 Classification-based approach

Pixel-based map

The poor rate of properly labeled height classes in the WFI plots (50%), may call into doubt the height class map derived from the classification-based approach. The situation is a bit more complex, as observed heights per WFI plots is an average of a few dominant tree heights. The approach considering pCHM pixels treats the entire poplar area after applying an average filter, so some of the differences may come from this observation.

Canopy Height Model accuracy

Despite the above discussion, we observe that the pCHM seems to be effective in characterizing stands by height. We can see at points A and D in Figure 14 that poplar plantations have been distinguished by their height classes. However, limitations to the pCHM accuracy appear to be observed at isolated poplar trees or alignments (*see* point C in Figure 14) and at the edges of poplar plantations (*see* point D in Figure 14).

WFI plots

The results presented in Section 3.1.3 are conditioned by the small area covered by the WFI plots and the fact that the dominant heights measured in the field were spread between 2014 and 2022. Figure 13 shows in green the difference in height between the predicted average height per year from the 2018 pCHM and the observed average height per year; in orange the observed average height per year. Two findings seem to appear: first, although not present at both ends (2014 and 2022), the height difference (green) appears to increase the further away from 2018, and second, the height difference (green) appears to be larger for high observed height (orange) is large. The second hypothesis is favored in this case, because significant differences in height (green) is still observed for 2018.

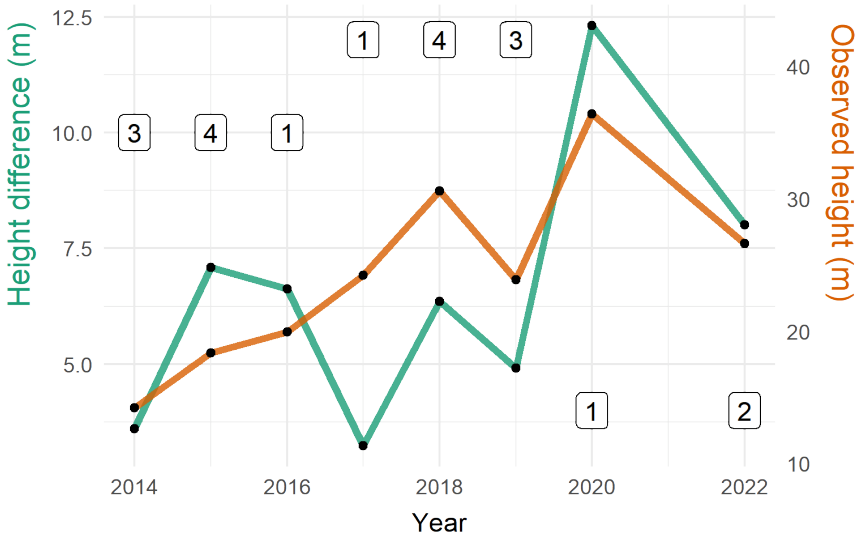


Figure 13: The difference in average height per pixel between predictions and observations (green) and average observed height per pixel (orange) from 2014 to 2022, except for 2021 containing no field height measurements. The observed height corresponds to the average of the dominant height measured per Walloon Region Forest Inventory plot and the boxes correspond to the number of plots considered per year.

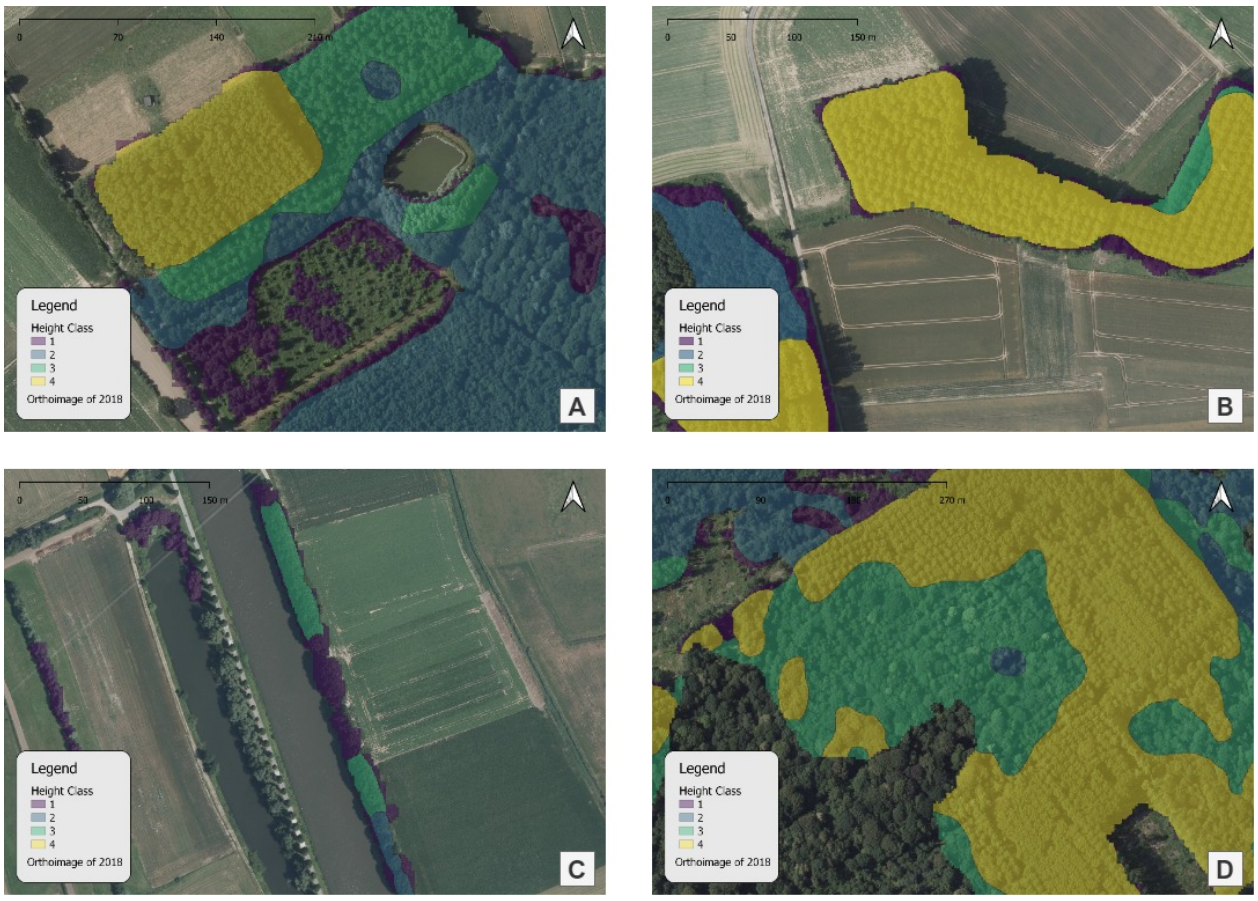


Figure 14: Case example from the height class map of the classification-based approach.

4.2.2 Semantic segmentation-based approach

Object-based map

It was pointed out in Section 3.2.3 that the accuracy of the ratio of properly labeled height classes was nearly 70%. This better ratio can be explained by more relevant predicted heights compared to those observed in the WFI plots, which are also average dominant heights.

Canopy Height Model accuracy

The height class map, from this approach, using the 2019 pCHM, appears to be sufficient to distinguish poplar plantations, as it can be seen at points A and C in Figure 16. In addition, the edge effect observed for the pixel-based height class map (classification-based approach) is considerably less present with an object-based approach (*see* point B in Figure 16), but some imprecision of the pCHM is still observed on alignments as shown in point C in Figure 16.

WFI plots

Related to Section 4.2.1, the difference in height between the average height predicted per year from the 2019 pCHM and that observed per year within the WFI plots (green) does not appear to increase as the years get further away from 2019. The hypothesis stating that the difference in height (green) is higher as the observed height (orange) increases, does not appear to be true in this case as well. The only observed fact is that the height difference (green) reaches a peak, for both height class maps, at years 2015, 2018, and 2020.

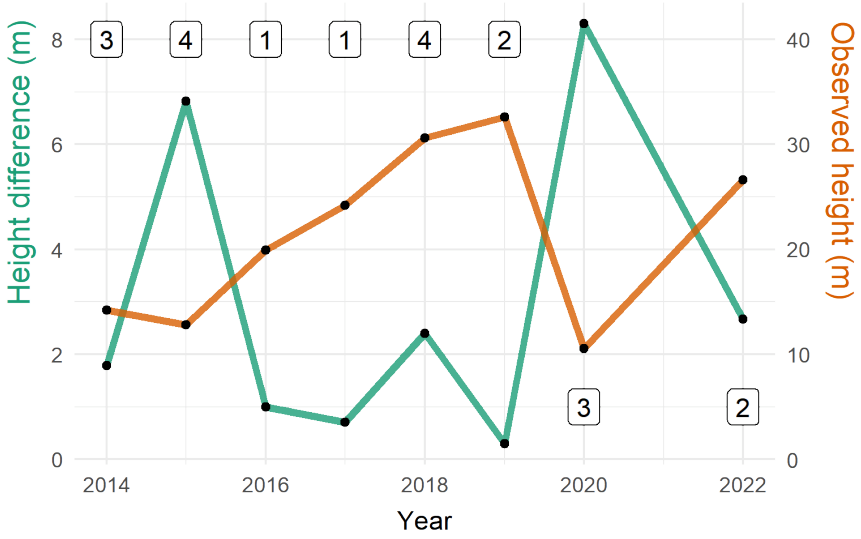


Figure 15: The difference in average height per tree between predictions and observations (green) and average observed height per tree (orange) according to the years 2014 to 2022, except for 2021 containing no field height measurements. The observed height corresponds to the average of the dominant height measured per Walloon Region Forest Inventory plot and the boxes correspond to the number of plots considered per year.

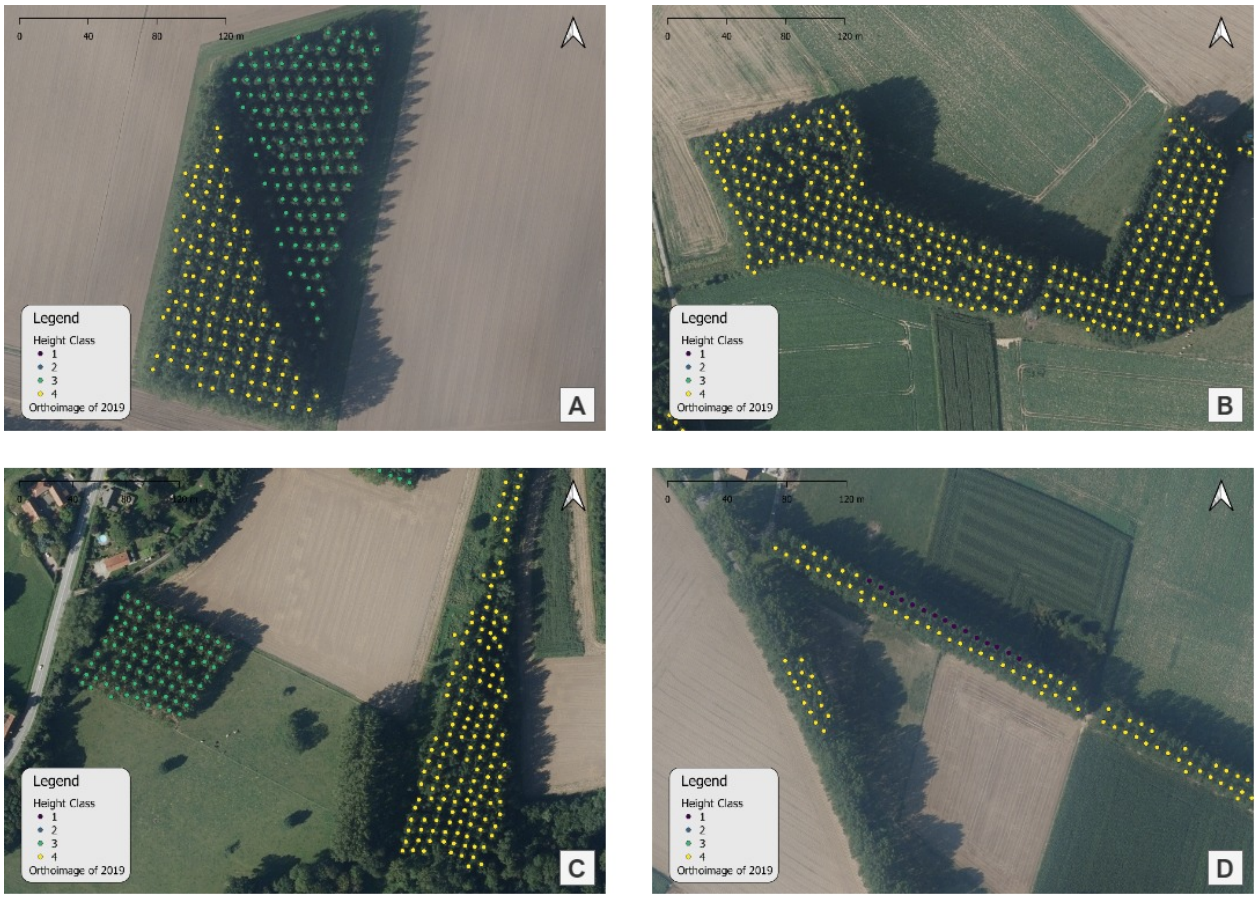


Figure 16: Case example from the height class map of the semantic segmentation-based approach.

4.3 Personal contributions

The activities conducted in the framework of this master's thesis concern the study of two types of data to map the poplar resource in the province of Hainaut, namely S2 super-resolution images and orthoimages, followed by an alternative to LIDAR data to characterize this resource.

A first cartographic approach using ML from the sub-mentioned satellite images allowed to evaluate the added value provided by an improvement of the spatial resolution of S2 images. Considering different datasets, a higher accuracy was obtained compared to the maps produced by Bolyn, Latte, Colson, et al. (2020) from S2 images without super-resolution.

A second cartographic approach has implemented DL in order to leverage orthoimages to estimate the poplar resource. To this end, a reference dataset was collected to feed the neural network to produce a semantic segmentation. These results highlighted the potential of orthoimages for large-scale poplar detection.

These results were nevertheless marred by a large number of FPs, so a method was adapted from Bolyn, Latte, Fourbisseur, et al. (2020) initially to identify and classify groups of trees outside of forests from LIDAR data to consider groups of poplars and remove those isolated

or having no regular spatial pattern in checkerboard or alignment. This method required the use of height data from the 2019 pCHM. Beforehand, an initial sorting was performed based on the confidence score map resulting from the approach using super resolved S2 images and the prediction dense score map or semantic segmentation result. FPs were significantly reduced, improving the accuracy of the map produced from the orthoimages.

Finally, a pCHM-based poplar resource height characterization derived from a combination of the 2014 LIDAR data DTM and the 2018 or 2019 DAP was applied for both maps. The characterization of the poplar resource was more accurate using the map resulting from the segmentation-based approach because the height extraction is tied to the position of the poplars, thus avoiding potential inaccuracy of the pCHM, especially at the edges or alignments of poplar plantations.

4.4 Perspectives

4.4.1 Sentinel-2 super resolution images

This work highlighted the improved accuracy of the poplar resource map thanks to the super-resolution of the S2 images. It would be possible to produce a new map of the main forest stand types in Wallonia, thus extending the classification to a larger area and to multiple forest species. This new map would confirm for other forest species the improvement in accuracy with that of the spatial resolution of satellite images.

It could also be considered to use a DL approach for poplar resource mapping by classification of super-resolved S2 images and compare it to results obtained by a ML approach. Such a study would measure the potential of a complex algorithm combined with complex images and assess whether the improved spatial resolution of S2 images is now sufficient for this type of algorithm.

The classification method of remote sensing data, after having been largely access on ML tends to reconsider DL, after improving the processing time and the results originally thanks to the AlexNet architecture (Shafaey et al., 2019). Land cover and land use classification (Pelletier, Webb, and Petitjean, 2019) and FC classification (primary dry forest, secondary dry plantation forest) in Indonesia (Miranda, Mutiara, and Wibowo, 2019), both from S2 images showed promising results. Although forest species classification is a more complex problem than those performed in the previous work, it could be considered to use DNNs to map the poplar resource through a S2 super-resolution image classification.

Nevertheless, a high processing time, the need for large datasets and a powerful hardware are limitations to the use of DL, especially as these limitations are accentuated by large-scale studies. At this point, (Z.-H. Zhou and Feng, 2017) proposed a ML algorithm, inspired by DL and decision trees, the Deep Forest Classifier, which deserves special attention. A study to map wetlands using Sentinel-1 and S2 imagery shows good results with this algorithm (Jamali et al., 2021). It might be interesting to compare the classification of super-resolved S2 images for poplar resource mapping by a conventional ML method, with a DL method and finally with the Deep Forest classifier.

4.4.2 Orthoimages

The potential of orthoimages for poplar resource mapping using a DL approach proved promising but littered with FPs and FNs. The first one has been greatly reduced by a post-processing method that seems to be effective, so limiting the number of FNs should be considered. The following avenues can be considered for this purpose:

- Quantitative improvement of the reference dataset by pointing poplars within manually delineated poplar plantations with new LIDAR data acquired for the years 2021 and 2022.
- Qualitative improvement of the reference dataset, by considering a greater proportion of alignments as well as poplars at a young and old development stage.
- Further push the neural network configuration, such as using rotation and brightness change during data augmentation.

The method investigated in this project to map the poplar resource from orthoimages relies on the confidence score map in order to reduce the number of FPs, it could be interesting to conduct further studies to avoid this dependence on satellite images during post-processing.

4.4.3 Photogrammetric canopy height model

The findings reported in this master thesis regarding the characterization of the poplar resource by pCHM should be considered with caution due to poor ground truth data. Further study to measure the capability of pCHM data as an alternative to LIDAR CHM for forest resource height characterization should include field data collection.

Characterization of the poplar resource was more accurately achieved through the use of an object-based map, as presented in Section 3.2.3. The height classes in the map resulting from the classification-based approach could then be revised by averaging the local maximum of the original pCHM by polygons. These local maximums, which can be seen as dominant tree heights would then be more relevant to compare with the WFI plots data.

5 Conclusion

Map production and characterization of the poplar resource in the province of Hainaut are the core components of this master's thesis. Specifically, the work investigates: the mapping potential of the Sentinel-2 (S2) super-resolution images (2.5 m spatial resolution and 10 bands) using a Machine Learning-based classification approach Random Forests (RFs)) and the potential of orthoimages (0.25 m spatial resolution and RGB and NIR bands) using a Deep Learning-based semantic segmentation approach (Deep Layer Aggregation (DLA)). Following these detections of the poplar resource, the feasibility of a characterization based on the classification into height classes of photogrammetric Canopy Height Models (pCHMs) is explored. The latter resulting from the combination of a Digital Terrain Model (DTM) (generated from Light Detection And Ranging (LIDAR) data) and of a Digital Aerial Photogrammetry (DAP).

First, the accuracy assessment of the map produced from the super-resolved S2 images indicates a kappa of 0.856 and an F1-score of 0.923, for a precision of 0.987 and a recall of 0.867. Although different datasets must be considered, "poplar" label obtained the highest F1-score of 0.746 by classification of S2 images (Bolyn, Latte, Colson, et al., 2020). The results from this master's thesis suggest an improvement in the accuracy of poplar resource detection by super-resolving these satellite images

Second, with an average distance of 1 m between the location of correctly predicted poplars (True Positives (TPs)) and observed poplars, leveraging orthoimage to map the poplar resource through Deep Learning (DL) is undeniable. Nevertheless, with an F1-score of 0.653, a precision of 0.776 and a recall of 0.564, substantial imperfections are observed. This approach has notable shortcomings in detecting poplar plantations at young (young sapling) or old (mature trees) development stages and those with a spatial pattern in alignment.

Third, 50% of the pixels and 69% of the poplars (points) were properly characterized according to height class. However, due to the lack of adequate ground truth data, these results must be measured with caution. poplar resource characterization is more accurate starting from the object-based height class map from the semantic segmentation-based approach; as it considers the precise location of poplars, contrary to a pixel-based height class map (from the classification-based approach) that considers the edge effect and landscape elements (*e.g.*, alignments) attenuation by the low-pass filter. In the context of this project, since poplar resource characterization is visually correct and the accuracy assessment of results from the object-based height class map (semantic segmentation-based approach) is reasonable, pCHM would be a suitable alternative to LIDAR.

References

Adler, A. et al. (June 2021). "Variation of growth and phenology traits in poplars planted in clonal trials in Northern Europe—implications for breeding". en. In: *BioEnergy Research* 14.2, pp. 426–444. ISSN: 1939-1234, 1939-1242. DOI: 10.1007/s12155-021-10262-8.

Alderweireld, M. et al. (2015). *Inventaire forestier wallon, résultats de 1994 - 2012*. fr. SPW Editions. SPW, DGO3, DNF, Direction des Ressources Forestières, Jambes.

Apan, A. et al. (Sept. 2003). "Formulation and assessment of narrow-band vegetation indices from EO1 Hyperion imagery for discriminating sugarcane disease". In: *Proceedings of the Spatial Sciences Conference*.

Bolyn, C., N. Latte, V. Colson, et al. (Nov. 2020). "Une carte des principaux types de peuplements forestiers de Belgique et du Nord de la France". French. In: *Forêt.Nature* 156, pp. 48–57. ISSN: 1372-8903.

Bolyn, C., N. Latte, A. Fourbisseur, et al. (2020). "Cartographie et caractérisation des arbres hors forêt à l'aide de la technologie LiDAR". French. In: *Forêt.Nature* 155, pp. 34–46. ISSN: 1372-8903.

Bolyn, C., P. Lejeune, et al. (2019). "Automated classification of trees outside forest for supporting operational management in rural landscapes". en. In: *Remote Sensing* 11.10, p. 1146. ISSN: 2072-4292. DOI: 10.3390/rs11101146.

Boyd, D. S. and F. M. Danson (2005). "Satellite remote sensing of forest resources: three decades of research development". en. In: *Progress in Physical Geography: Earth and Environment* 29.1, pp. 1–26. ISSN: 0309-1333. DOI: 10.1191/0309133305pp432ra.

Bragagnolo, L., R. V. da Silva, and J. M. V. Grzybowski (May 2021). "Amazon forest cover change mapping based on semantic segmentation by U-Nets". en. In: *Ecological Informatics* 62, p. 101279. ISSN: 1574-9541. DOI: 10.1016/j.ecoinf.2021.101279.

Breiman, L. (2001). "Random forests". en. In: *Machine Learning* 45.1, pp. 5–32. ISSN: 1573-0565. DOI: 10.1023/A:1010933404324.

Carah Asbl (2018). *Peuplier et populiculture 2.0. : pour une culture pérenne et responsable des peupleraies*.

Carle, J. and P. Holmgren (2008). "Wood from planted forests A global outlook 2005-2030". In: *Forest Products Journal* 58.12, pp. 6–18.

Chianucci, F. et al. (Dec. 2020). "Influence of image pixel resolution on canopy cover estimation in poplar plantations from field, aerial and satellite optical imagery". en. In: *Annals of Silvicultural Research* 46.1. ISSN: 2284354X. DOI: 10.12899/asr-2074.

Cloutis, E. A. (Aug. 1996). "Review Article hyperspectral geological remote sensing: evaluation of analytical techniques". In: *International Journal of Remote Sensing* 17.12, pp. 2215–2242. ISSN: 0143-1161. DOI: 10.1080/01431169608948770.

Copernicus (2022). *Copernicus - Land*.

Crutzen, F. (June 2017). “Approche multi-capteurs pour la cartographie par télédétection des ressources ligneuses en Wallonie : application à la commune de Paliseul”. PhD thesis. Liege University (ULiege).

Cutler, F. o. b. L. B. a. A. and R. p. b. A. L. a. M. Wiener (May 2022). *randomForest: Breiman and Cutler's Random Forests for Classification and Regression*.

D'Amico, G. et al. (2021). “A deep learning approach for automatic mapping of poplar plantations using Sentinel-2 imagery”. en. In: *GIScience and Remote Sensing* 58.8, pp. 1352–1368. ISSN: 1548-1603. DOI: 10.1080/15481603.2021.1988427.

Datt, B. (Jan. 1999). “A New Reflectance Index for remote sensing of chlorophyll content in higher plants: Tests using eucalyptus leaves”. en. In: *Journal of Plant Physiology* 154.1, pp. 30–36. ISSN: 0176-1617. DOI: 10.1016/S0176-1617(99)80314-9.

Dumont, S.-P. (2018). “Peuplier et populiculture, deuxième partie”. In: *Silva Belgica*, pp. 16–27.

ESA (2022a). *Newcomers Earth Observation Guide | ESA Business Applications*.

– (2022b). *Sentinel Online*.

– (2022c). *Sentinel-2 MSI Technical Guide*.

Falbel, D., J. Luraschi, et al. (June 2022). *torch: Tensors and Neural Networks with 'GPU' Acceleration*.

Falbel, D., C. Regouby, and RStudio (Jan. 2022). *torchvision: Models, Datasets and Transformations for Images*.

FAO (Oct. 2020a). *Activities related to poplar and willow cultivation and utilization 2016-2019 : Belgium*. Tech. rep. Italy, Rome.

– (2020b). *Global Forest Resources Assessment 2020*. en. Rome. ISBN: 978-92-5-132581-0. DOI: 10.4060/ca8753en.

– (Apr. 2021). *Évaluation des ressources forestières mondiales 2020*. fr. FAO. ISBN: 978-92-5-134306-7. DOI: 10.4060/ca9825fr.

FAO-CABI (2014). *Poplars and willows: trees for society and the environment*. en. Ed. by J. G. Isebrands and J. Richardson. Boston, MA : Rome: CABI ; FAO. ISBN: 978-1-78064-108-9.

Fensholt, R. and I. Sandholt (Sept. 2003). “Derivation of a Shortwave Infrared Water Stress Index from MODIS Near and Shortwave Infrared Data in a Semiarid Environment”. In: *Remote Sensing of Environment* 87, pp. 111–121. DOI: 10.1016/j.rse.2003.07.002.

Giesen, L. (2022). “Deep learning semantic segmentation of tree stock in South Africa using satellite images”. en. PhD thesis.

Gitelson, A. A., M. N. Merzlyak, and O. B. Chivkunova (July 2001). “Optical properties and nondestructive estimation of anthocyanin content in plant leaves”. In: *Photochemistry and Photobiology* 74.1, pp. 38–45. ISSN: 0031-8655. DOI: 10.1562/0031-8655(2001)074<0038:opaneo>2.0.co;2.

Guerschman, J. et al. (May 2009). “Estimating fractional cover of photosynthetic vegetation, non-photosynthetic vegetation and bare soil in the Australian tropical savanna region up-

scaling the EO-1 Hyperion and MODIS sensors". In: *Remote Sensing of Environment* 113, pp. 928–945. DOI: 10.1016/j.rse.2009.01.006.

Hamrouni Berkaoui, Y. (2021). "Développement d'une approche opérationnelle pour l'identification automatique des peupleraies à large échelle par télédétection hyper-temporelle : De l'adaptation de domaine à la création d'un indice spectral dédié". fr. PhD thesis. Institut National Polytechnique de Toulouse.

Hijmans, R. J. et al. (Aug. 2022). *raster: Geographic Data Analysis and Modeling*.

Huang, G. et al. (Jan. 2018). *Densely Connected Convolutional Networks*. arXiv:1608.06993 [cs]. DOI: 10.48550/arXiv.1608.06993.

Huang, H. et al. (Apr. 2020). *UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation*. DOI: 10.48550/arXiv.2004.08790.

Hunt, E. R. and B. N. Rock (Oct. 1989). "Detection of changes in leaf water content using Near- and Middle-Infrared reflectances". en. In: *Remote Sensing of Environment* 30.1, pp. 43–54. ISSN: 0034-4257. DOI: 10.1016/0034-4257(89)90046-1.

Interreg : Forêt-Pro-Bos (2020). *Peuplier, environnement et climat, une essence irremplaçable*.

Jamali, A. et al. (Oct. 2021). "Deep Forest classifier for wetland mapping using the combination of Sentinel-1 and Sentinel-2 data". In: *GIScience & Remote Sensing* 58.7. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/15481603.2021.1965399, pp. 1072–1089. ISSN: 1548-1603. DOI: 10.1080/15481603.2021.1965399.

Karasiak, N. et al. (2019). "Statistical stability and spatial instability in mapping forest tree species by comparing 9 years of satellite image time series". en. In: *Remote Sensing* 11.21. Publisher: Multidisciplinary Digital Publishing Institute, p. 2512. ISSN: 2072-4292. DOI: 10.3390/rs11212512.

Kavzoglu, T., H. Tonbul, and I. Colkesen (Nov. 2021). "Evaluation of atmospheric correction methods for Sentinel-2 imagery in the spectral identification of poplar (Populus deltoides Bartr.) species". en. In: Can Tho University, Can Tho city, Vietnam, p. 8.

Key, C. and N. Benson (Jan. 2006). "Landscape Assessment: Ground measure of severity, the Composite Burn Index; and Remote sensing of severity, the Normalized Burn Ratio." In: *FIREMON: Fire Effects Monitoring and Inventory System*. RMRS-GTR-164. USDA Forest Service, Rocky Mountain Research Station, Ogden, UT, LA 1–51.

Kimmins, J. (Nov. 2003). "Old-growth forest: An ancient and stable sylvan equilibrium, or a relatively transitory ecosystem condition that offers people a visual and emotional feast? Answer - It depends". In: *Forestry chronicle* 79.3, pp. 429–440. DOI: 10.5558/tfc79429-3.

Kishore, R. R. et al. (Dec. 2016). "Comparative accuracy of different classification algorithms for forest cover type prediction". In: pp. 116–123. DOI: 10.1109/APWC-on-CSE.2016.029.

Koch, B., M. Dees, et al. (July 2008). "Forestry applications". In: *Advances in Photogrammetry, Remote Sensing and Spatial Information Sciences: 2008 ISPRS Congress Book*, pp. 439–465. ISBN: 978-0-415-47805-2. DOI: 10.1201/9780203888445.ch32.

Koch, B., C. Straub, et al. (Feb. 2009). "Airborne laser data for stand delineation and information extraction". In: *International Journal of Remote Sensing* 30.4, pp. 935–963. DOI: 10.1080/01431160802395284.

Korznikov, K. A. et al. (Jan. 2021). "Using U-Net-Like deep convolutional neural networks for precise tree recognition in very high resolution RGB (Red, Green, Blue) satellite images". en. In: *Forests* 12.1. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute, p. 66. ISSN: 1999-4907. DOI: 10.3390/f12010066.

Kuhn [aut, M. et al. (Aug. 2022). *caret: Classification and Regression Training*.

Latte, N. and P. Lejeune (2020). "PlanetScope radiometric normalization and Sentinel-2 super-resolution (2.5 m): A Straightforward spectral-spatial fusion of multi-satellite multi-sensor images using residual convolutional neural networks". en. In: *Remote Sensing* 12.15, p. 2366. ISSN: 2072-4292. DOI: 10.3390/rs12152366.

Lin, T.-Y., P. Dollár, et al. (Apr. 2017). *Feature Pyramid Networks for Object Detection*. DOI: 10.48550/arXiv.1612.03144.

Lin, T.-Y., M. Maire, et al. (Feb. 2015). *Microsoft COCO: Common Objects in Context*. DOI: 10.48550/arXiv.1405.0312.

Ma, L. et al. (June 2019). "Deep learning in remote sensing applications: A meta-analysis and review". en. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 152, pp. 166–177. ISSN: 0924-2716. DOI: 10.1016/j.isprsjprs.2019.04.015.

Michez, A. et al. (2020). "Can regional aerial images from orthophoto surveys produce high quality photogrammetric Canopy Height Model? A single tree approach in Western Europe". en. In: *International Journal of Applied Earth Observation and Geoinformation* 92, p. 102190. ISSN: 1569-8432. DOI: 10.1016/j.jag.2020.102190.

Miranda, E., A. B. Mutiara, and E. a. W. C. Wibowo (Dec. 2019). "Forest Classification Method Based on Convolutional Neural Networks and Sentinel-2 Satellite Imagery". en. In: *International Journal of Fuzzy Logic and Intelligent Systems* 19.4. Publisher: Korean Institute of Intelligent Systems, pp. 272–282. DOI: 10.5391/IJFIS.2019.19.4.272.

Nutanong, S., E. Jacox, and H. Samet (May 2011). "An Incremental Hausdorff Distance Calculation Algorithm." In: *PVLDB* 4.8, pp. 506–517. DOI: 10.14778/2002974.2002978.

Office économique wallon du bois (2021). *PanoraBois Wallonie edition 2021*.

Ozturk, M. Y. and I. Colkesen (Nov. 2020). "Mapping of poplar tree growing fields with machine learning algorithms using multi-temporal Sentinel-2A imagery". In.

Pebesma, E., R. Bivand, E. Racine, et al. (July 2022). *sf: Simple Features for R*.

Pebesma, E., R. Bivand, B. Rowlingson, et al. (Jan. 2018). *sp: Classes and Methods for Spatial Data*.

Pebesma, E., M. Sumner, et al. (July 2022). *stars: Spatiotemporal Arrays, Raster and Vector Data Cubes*.

Pelletier, C., G. I. Webb, and F. Petitjean (July 2019). "Deep Learning for the Classification of Sentinel-2 Image Time Series". en. In: *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*. Yokohama, Japan: IEEE, pp. 461–464. ISBN: 978-1-5386-9154-0. DOI: 10.1109/IGARSS.2019.8900123.

Peñuelas, J. et al. (May 1994). "Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves". en. In: *Remote Sensing of Environment* 48.2, pp. 135–146. ISSN: 0034-4257. DOI: 10.1016/0034-4257(94)90136-8.

Peuplier (2022). fr-FR. Publication Title: Hout Info Bois.

Planet Labs PBC (Sept. 2021). *Satellite Imagery and Archive*. Publication Title: Planet.

Ponomarenko, M. R. and V. A. Zelentsov (Aug. 2021). "Forest monitoring and analysis based on Earth observation data services". en. In: *IOP Conference Series: Earth and Environmental Science*. Vol. 806, p. 012003. DOI: 10.1088/1755-1315/806/1/012003.

PSW (2022a). *Nuage de points LIDAR 2013-2014*. en.

– (2022b). *Orthophotos : la Wallonie vue du ciel*. en.

Rakhmatulin, I., A. Kamilaris, and C. Andreasen (Jan. 2021). "Deep neural networks to detect weeds from crops in agricultural environments in real-time: A review". en. In: *Remote Sensing* 13.21. Publisher: Multidisciplinary Digital Publishing Institute, p. 4486. ISSN: 2072-4292. DOI: 10.3390/rs13214486.

Ribera, J. et al. (June 2019). "Locating Objects Without Bounding Boxes". en. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, CA, USA: IEEE, pp. 6472–6482. ISBN: 978-1-72813-293-8. DOI: 10.1109/CVPR.2019.00664.

Rondeaux, G., M. Steven, and F. Baret (Feb. 1996). "Optimization of soil-adjusted vegetation indices". en. In: *Remote Sensing of Environment* 55.2, pp. 95–107. ISSN: 0034-4257. DOI: 10.1016/0034-4257(95)00186-7.

Rondeux, J. et al. (Nov. 2016). "National Forest Inventories : Assessment of Wood Availability and Use - Chapter 8 - Belgium (Wallonia)". In: *National Forest Inventories: Assessment of Wood Availability and Use*. Journal Abbreviation: National Forest Inventories: Assessment of Wood Availability and Use, pp. 159–179. ISBN: 978-3-319-44014-9. DOI: 10.1007/978-3-319-44015-6_8.

Ronneberger, O., P. Fischer, and T. Brox (May 2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. DOI: 10.48550/arXiv.1505.04597.

Rouse, J. W. et al. (Jan. 1974). "Monitoring vegetation systems in the Great Plains with ERTS". In: NTRS Author Affiliations: Texas A&M Univ. NTRS Report/Patent Number: PAPER-A20 NTRS Document ID: 19740022614 NTRS Research Center: Legacy CDMS (CDMS).

Sandhya Devi, M. R., V. Vijay Kumar, and P. Sivakumar (Dec. 2021). "A review of image classification and object detection on machine learning and deep learning techniques". In: *2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 1–8. DOI: 10.1109/ICECA52323.2021.9676141.

Seddati, O. et al. (Oct. 2017). "Towards Good Practices for Image Retrieval Based on CNN Features". In: pp. 1246–1255. DOI: 10.1109/ICCVW.2017.150.

Sercu, T. and V. Goel (Dec. 2016). *Dense Prediction on Sequences with Time-Dilated Convolutions for Speech Recognition*. arXiv:1611.09288 [cs]. DOI: 10.48550/arXiv.1611.09288.

Shafaey, M. A. et al. (2019). "Deep Learning for Satellite Image Classification". en. In: *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2018*. Ed. by A. E. Hassanien et al. Advances in Intelligent Systems and Computing. Cham: Springer International Publishing, pp. 383–391. ISBN: 978-3-319-99010-1. DOI: 10.1007/978-3-319-99010-1_35.

Shelhamer, E., J. Long, and T. Darrell (May 2016). *Fully Convolutional Networks for Semantic Segmentation*. en.

Sjöqvist, H. (2017). "Classifying forest cover type with cartographic variables via the support vector machine, naive bayes and random forest classifiers". en. PhD thesis. Örebro University.

Sjöqvist, H., M. Längkvist, and F. Javed (2020). "An analysis of fast learning methods for classifying forest cover types". In: *Applied Artificial Intelligence* 34.10, pp. 691–709. ISSN: 0883-9514. DOI: 10.1080/08839514.2020.1771523.

Travaglini, D. et al. (Mar. 2013). "Large-Scale Pan-European Forest Monitoring Network: A Statistical Perspective for Designing and Combining Country Estimates". In: *Developments in Environmental Science*. Vol. 12. Journal Abbreviation: Developments in Environmental Science, pp. 105–135. ISBN: 978-0-08-098222-9. DOI: 10.1016/B978-0-08-098222-9.00007-8.

Tuskan, G. A. et al. (Sept. 2006). "The Genome of Black Cottonwood, Populus trichocarpa (Torr. & Gray)". In: *Science* 313.5793. Publisher: American Association for the Advancement of Science, pp. 1596–1604. DOI: 10.1126/science.1128691.

Van-Der-Perre, R. et al. (2015). "La carte bioclimatique de Wallonie : un nouveau découpage écologique du territoire pour le choix des essences forestières". In: *Forêt.Nature* 135, pp. 47–58.

Vauhkonen, J. et al. (2019). "Harmonised projections of future forest resources in Europe". en. In: *Annals of Forest Science* 76.3, p. 79. ISSN: 1297-966X. DOI: 10.1007/s13595-019-0863-6.

Wang, F.-T. et al. (Mar. 2021). "DLA+: A Light Aggregation Network for Object Classification and Detection". en. In: *Machine Intelligence Research* 18.6. Publisher: Machine Intelligence Research, pp. 963–972. ISSN: 2731-538X. DOI: 10.1007/s11633-021-1287-y.

Wu, H., Q. Liu, and X. Liu (2019). "A review on deep learning approaches to image classification and object segmentation". en. In: *Computers, Materials & Continua* 60.2, pp. 575–597. ISSN: 1546-2226. DOI: 10.32604/cmc.2019.03595.

Xu, N. (2021). "Semi-supervised training with one-stage object detection". eng. PhD thesis.

Yu, F. et al. (June 2018). "Deep Layer Aggregation". en. In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT: IEEE, pp. 2403–2412. ISBN: 978-1-5386-6420-9. DOI: 10.1109/CVPR.2018.00255.

Zalesny Jr, R. S. et al. (2019). "Positive water linkages of producing short rotation poplars and willows for bioenergy and phytotechnologies". en. In: *WIREs Energy and Environment* 8.5. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wene.345, e345. ISSN: 2041-840X. DOI: 10.1002/wene.345.

Zhang, C. et al. (2017). "An up-to-date comparison of state-of-the-art classification algorithms". en. In: *Expert Systems with Applications* 82, pp. 128–150. ISSN: 0957-4174. DOI: 10.1016/j.eswa.2017.04.003.

Zhang, L., L. Zhang, and B. Du (June 2016). "Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art". In: *IEEE Geoscience and Remote Sensing Magazine* 4.2, pp. 22–40. ISSN: 2168-6831. DOI: 10.1109/MGRS.2016.2540798.

Zhao, K. and W. Q. Yan (2021). "Fruit detection from digital images using CenterNet". en. In: *Geometry and Vision*. Ed. by M. Nguyen, W. Q. Yan, and H. Ho. Communications in Computer and Information Science. Cham: Springer International Publishing, pp. 313–326. ISBN: 978-3-030-72073-5. DOI: 10.1007/978-3-030-72073-5_24.

Zhou, X., D. Wang, and P. Krähenbühl (Apr. 2019). *Objects as points*. DOI: 10.48550/arXiv.1904.07850.

Zhou, Z.-H. and J. Feng (Aug. 2017). "Deep Forest: Towards an alternative to deep neural networks". en. In: *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*. Melbourne, Australia: International Joint Conferences on Artificial Intelligence Organization, pp. 3553–3559. ISBN: 978-0-9992411-0-3. DOI: 10.24963/ijcai.2017/497.

Zhou, Z. et al. (Jan. 2020). *UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation*. DOI: 10.48550/arXiv.1912.05074.

Appendices

A Artificial Intelligence concepts related to learning

A.1 Artificial Intelligence

This concept dating back to the 1950s and first mentioned at the Dartmouth College Conference in 1956, Artificial intelligence (AI) was defined by John McCarthy as "*the science and engineering of making intelligent machines, especially intelligent computer programs*". In other words, it is about making machines and programs able to mimic human intelligence, by analyzing its environment and designing a plan of action to maximize the chances of solving the task (Kersting, 2018). Impacting many sectors, AI can be divided into different interconnected fields such as Machine Learning (ML), Artificial Neural Networks (ANNs), Computer Vision (CV), natural language processing, robotics, speech recognition and expert systems (Collins et al., 2021). The different subsets of AI related to learning used in this Master's thesis, as well as their relationships are shown in Figure 17.

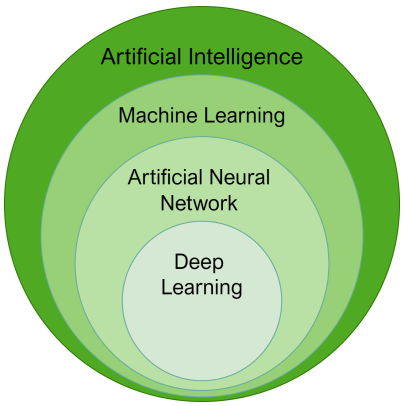


Figure 17: Relationship between Artificial intelligence learning subsets (Alkabbani et al., 2021; Alzubaidi et al., 2021).

A.2 Machine Learning

A subset of AI, ML is a system that learns from data to perform tasks for which it has not been explicitly programmed. Conventionally, learning can be **supervised** for classification (prediction of discrete values) and regression (prediction of continuous values) tasks, **unsupervised** for clustering, association and dimension reduction tasks,or by **reinforcement**, as well as variants of these types of ML, such as active learning or semi-supervised learning. In CV, supervised learning can be done on three categories of processing units: image, pixel or object level.

A.3 Artificial Neural Networks

ANN is a brain-inspired ML model. Specifically, it is a group of interconnected nodes that allow learning by layer in a non-linear way. An ANN is composed of three network layers: input layer (*e.g.*, an image), hidden layer and output layer (*e.g.*, categories), each of them containing nodes, also called artificial neurons or computational units (*cf.*, Figure 18), with edges to connect them to the inputs and outputs (Park and Lee, 2016; Alzubaidi et al., 2021). The parameters related to a node are usually called **weights** and can be weights (or gains) at each input as well as a offset value (or bias), then the **activation function** will "break" the linearity of the output (*cf.*, Section X) (Cresson, 2020). Conventional ANNs are shallow, *i.e.* they contain only one hidden layer, yet the latter becomes deep when it contains multiple hidden layers as illustrated in Figure 19, making it much more complex and heavy to support by the Central Processing Unit (CPU) and the Graphics Processing Unit (GPU). Deep ANNs are then called Deep Neural Networks (DNNs) and are the backbone of Deep Learning (DL).

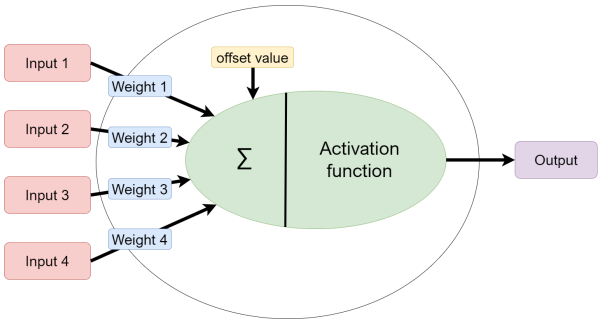


Figure 18: Node structure in an Artificial Neural Network (Alkabbani et al., 2021; Dastres and Soori, 2021).

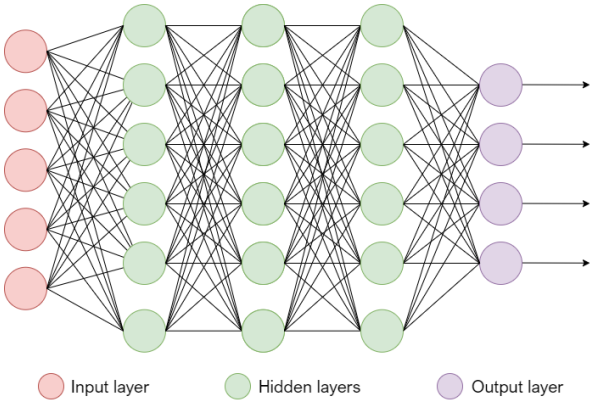


Figure 19: Deep Neural Network architecture (Alkabbani et al., 2021; Dastres and Soori, 2021).

A.4 Deep Learning

DL is a specific ML method that integrates DNN or ANN in successive layers to learn data iteratively. The model obtained by this learning induces the activation of billions of neurons in response to the input signals, thus providing a solution to the target problem. The main difference between DL and ML is its automation, since it does not require an operator to pre-process, extract features and select features before image classification. In addition, the latter is more efficient in terms of processing geospatial data, given its ability to identify spatial dependency by combining spectral and vector information, whereas ML algorithms are limited to identifying spectral signatures (Nikparvar and Thill, 2021; Zuo et al., 2015; Alzubaidi et al., 2021).

The neural network learning process works as follows: by **backpropagation**, parameters such as input **weights** of an artificial neuron are adjusted in order to activate or not the latter and to obtain an output of the neural network as close as possible to the training case. The backpropagation process is therefore iterative and returns to the input weights an adjustment to reduce the error. As a reminder, the **activation function** allows to keep or not a node (assign a value between 0 and 1) within a neural network by adding a non-linear property to the network. In short, a linear classification as in ML is faster but faces limits when two classes are very close in spectral range, but evolve in a different environment, a non-linear approach will be more likely to solve this kind of problem. The error considered at the output layer and backward through the neural network during the backpropagation process is computed using a **loss function** (also called cost function). A key point is to find the smallest error. To achieve this, a **gradient descent** algorithm is used to locate the global minimum of the loss function. In DL, usually big data are used, implying a partition of the dataset for training the model. The tiles or **patches** making up the complete dataset are divided into **batches**. An **epoch** refers to the running of all batches, therefore a full dataset, the number of **iterations** corresponds to the number of batches needed to complete an epoch.

Hyperparameters are values defined before the training process and determine how the neural network learns and affects its structure. Among those, the **learning rate** defining the adjustment time of the neural network parameters at each backward. A too small value increases the time needed to converge the descent gradient to the global minimum, a too high value makes the behavior of the loss function chaotic thus the descent gradient fails to converge. Other important hyperparameters are related to the **batch size**, or the **number of epochs**.

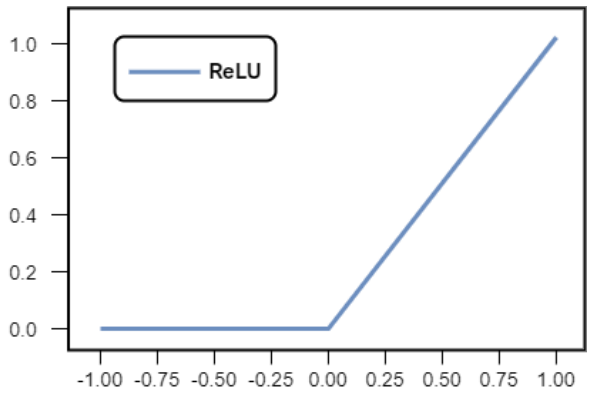
During the training process, a middle ground between **underfitting** and **overfitting** must be found. The last occurs when the model learns from noise and details, especially in the case of non-linear models more flexible. To prevent overfitting, some techniques can be mentioned, such as **early stopping**, **data augmentation** or **dropout**. The first technique is to stop learning before overfitting, then data augmentation increases the dataset by adding new samples with different brightness, after flipping or rotation, finally dropout deliberately reduces the number of nodes thus increasing the number of iterations needed to converge the model. A comprehensive description of DL is found in (Mishra, Reddy, and Pathak, 2021; Rakhmatulin, Kamilaris, and Andreasen, 2021; Zhao and Yan, 2021).

In a non-exhaustive way, the DNN architecture constituting the backbone of the DL algorithm can be based on Convolutional Neural Networks (CNNs) or Recurrent Neural Networks

(RNNs) during supervised learning or on self-organizing maps or auto-encoders during unsupervised learning. Both RNNs and CNNs are used in CV, the former processes information sequences such as temperature, time series, while the latter is used for image classification and object recognition. The following section only describes CNNs, as other architecture types are not relevant to this project.

A.5 Convolution Neural Networks

CNNs are a type DNN architecture for which state-of-the-art results are obtained in the field of image recognition, object detection and semantic segmentation (Alzubaidi et al., 2021; Cresson, 2020; Ma et al., 2019); they are therefore dominant for the processing of satellite or aerial images with DNN. A convolution acts as a filter on the data that passes through the network, in the case of CV process the input images by weighting them to highlight patterns, objects, shapes, *etc.*. CNNs are composed of four types of layers, namely **convolution layers**, **ReLU layers**, **pooling layers** and **fully connected layers**. The first one performs a convolution to create several smaller images with features, the result is called feature maps. The second one is a Rectified Linear Unit (ReLU) activation function to add non-linarity by giving a zero value to the negative value resulting from the convolution figure 20. The third one receives as input the rectified feature maps and performs a down-sampling operation to reduce the spatial dimension while preserving the important features. In the end of the process the fully-connected layer get as input a flattened pooling result in order to recognize and classify the images or objects. As shown in figure 21, a deep CNN architecture includes several convolution, ReLU, and pooling layers before the fully connected layer to implement image classification by the final algorithm.



$$f(x) = \max(0, x)$$

Figure 20: Rectified Linear Unit (ReLU) activation function (Albawi, Mohammed, and Al-Zawi, 2017).

The entire processing through a CNN involves a **kernel** with a certain width and height with the same depth as the input image and moves from the top left to the bottom right in order to process the entire image.

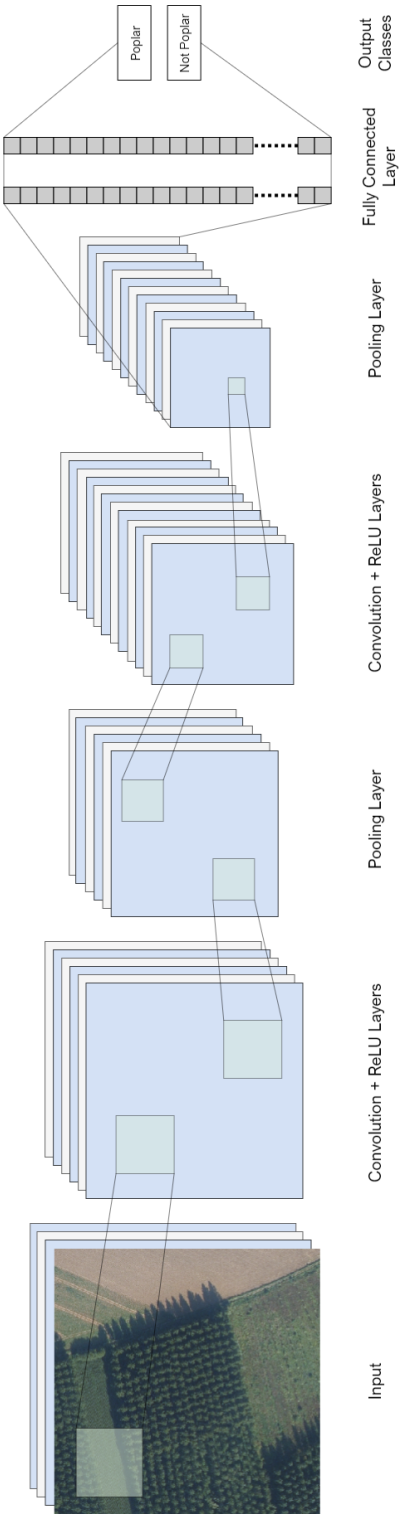


Figure 21: Example of Convolutional Neural Network architecture for aerial image classification (Alkabbani et al., 2021; Ma et al., 2019).

B Overview of Computer Vision tasks

CV is a field of AI, dedicated to the reading and understanding of images, namely an array of pixels containing values. Automation of image analysis is achieved by coupling CV with the learning field of AI (cf., Appendix A). Therefore, the progress of CV performance depends on that of AI learning. Research in the 2000s was accessed on improving features for ML, then following the foundation of CNNs and the successful achievement of the AlexNet model in 2012, the development of DL algorithms for CV has gained momentum (Alzubaidi et al., 2021). The applications of the CV are multiple, whether they are military, allowing the automation of industries, smart cities, medical imaging, intelligent transportation, Earth Observation (EO) among others. The following four task types are relevant to distinguish in this project: classification, localization, detection, and segmentation (Alzubaidi et al., 2021). Before describing them, it is necessary to understand that these tasks are based on **recognition** or object recognition resulting from prediction or prediction density, by the backbone of the algorithm to perform the CV task, this backbone is usually a CNN-based architecture for image processing (see Figure 22).

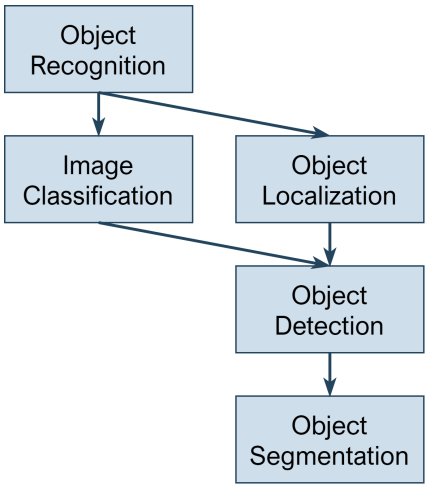


Figure 22: Overview of Computer Vision tasks.

Classification, the basis of CV, categorizes a set of data into classes (Wu, Q. Liu, and X. Liu, 2019). The input data, an image, pixel or recognized object, is labeled and, depending on the model, is attached with an evaluation measure such as probability, point density, loss. **Localization** or object localization is the process that locates an object in an image to be surrounded by a single bounding box. **Detection** or object detection is a combination of classification and localization. This type of task locates and classifies multiple objects which may have multiple occurrences with consideration of the background (see Figure 23).

Segmentation or object segmentation deals with the edges of target objects in the image, they are delimited by a pixel mask. This pixel-based operation includes **semantic segmentation** and **instance segmentation** (Wu, Q. Liu, and X. Liu, 2019). The former is a separation of the image into parts with different semantics, the pixels of the class of interest are distinguished from the background (Giesen, 2022). The latter is an extension of the detection task, required to process and refine the edges of the objects (Wu, Q. Liu, and X. Liu, 2019). Therefore, the detection task enables instance segmentation to differentiate instances unlike semantic segmentation.

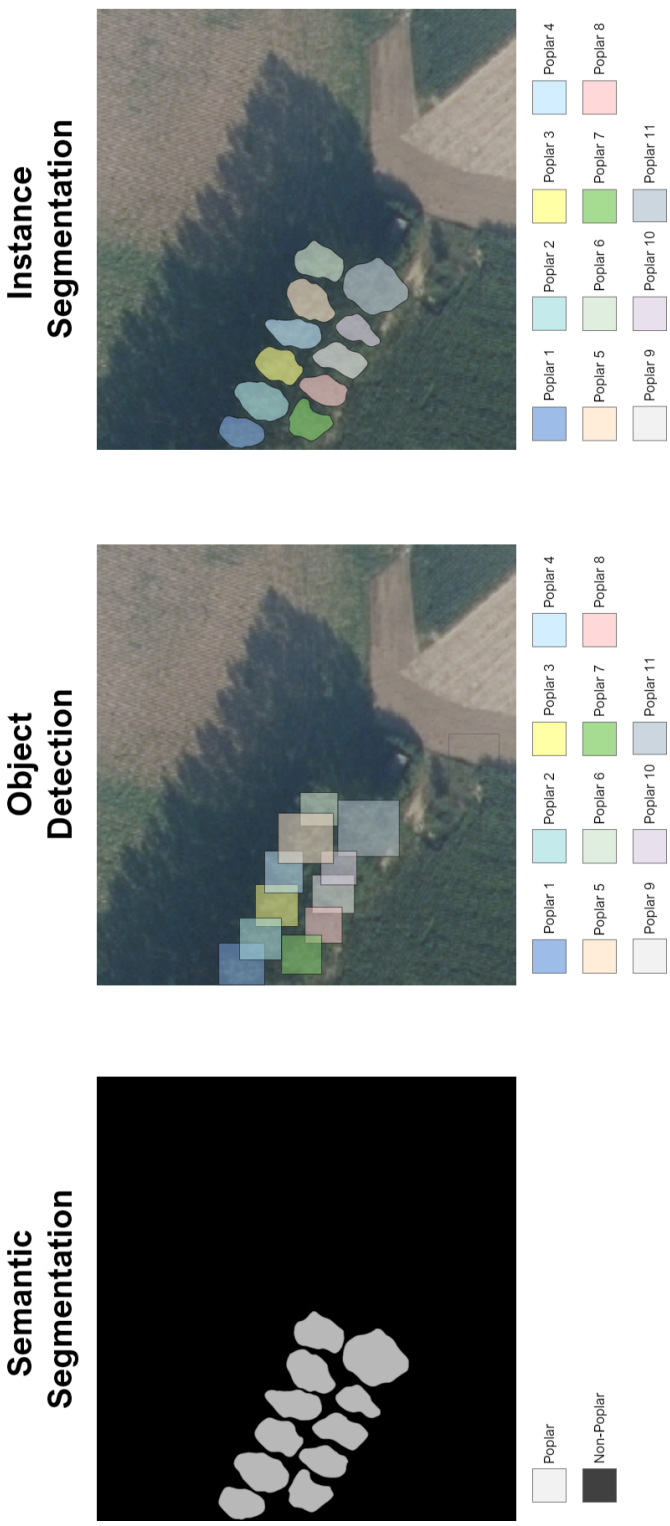
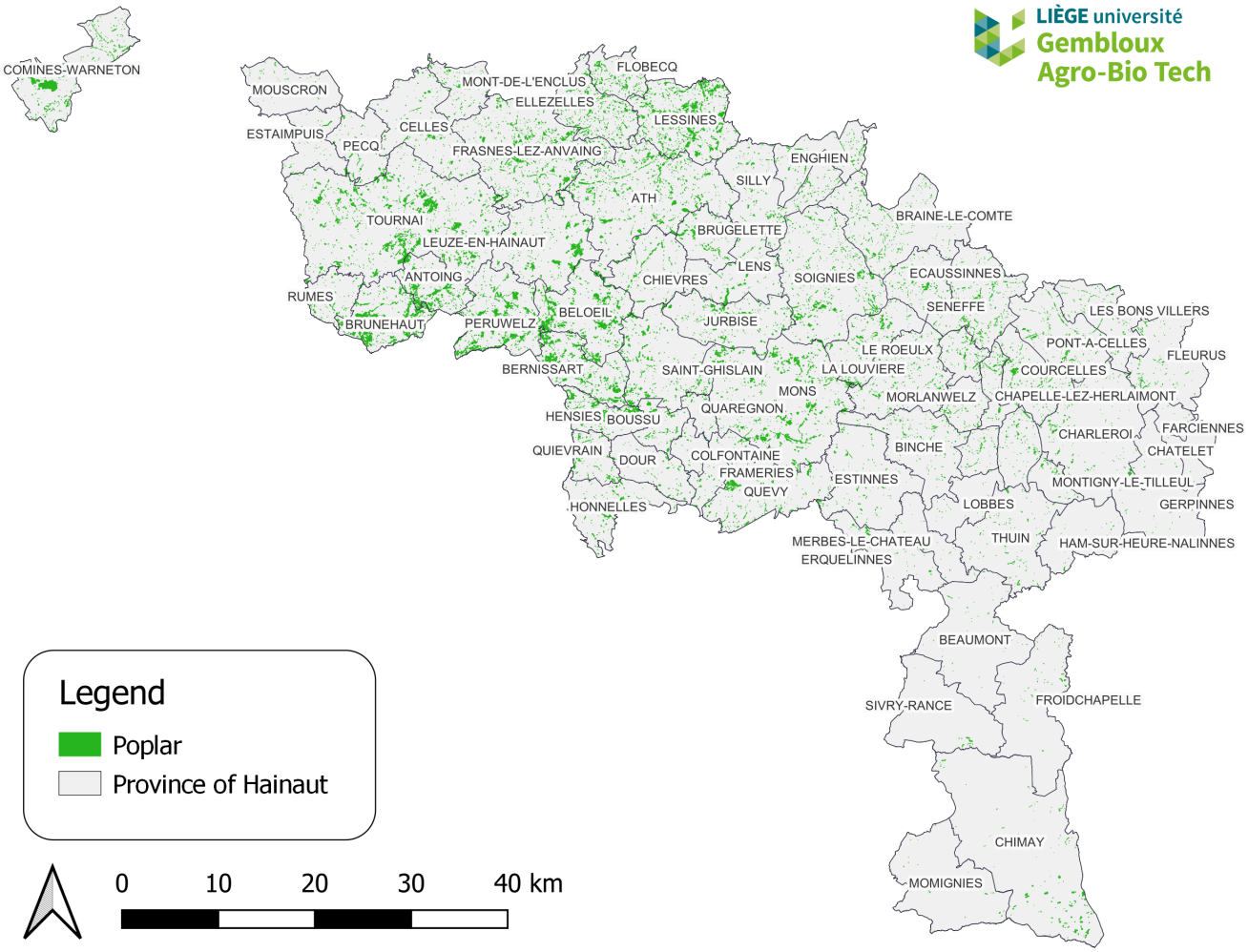


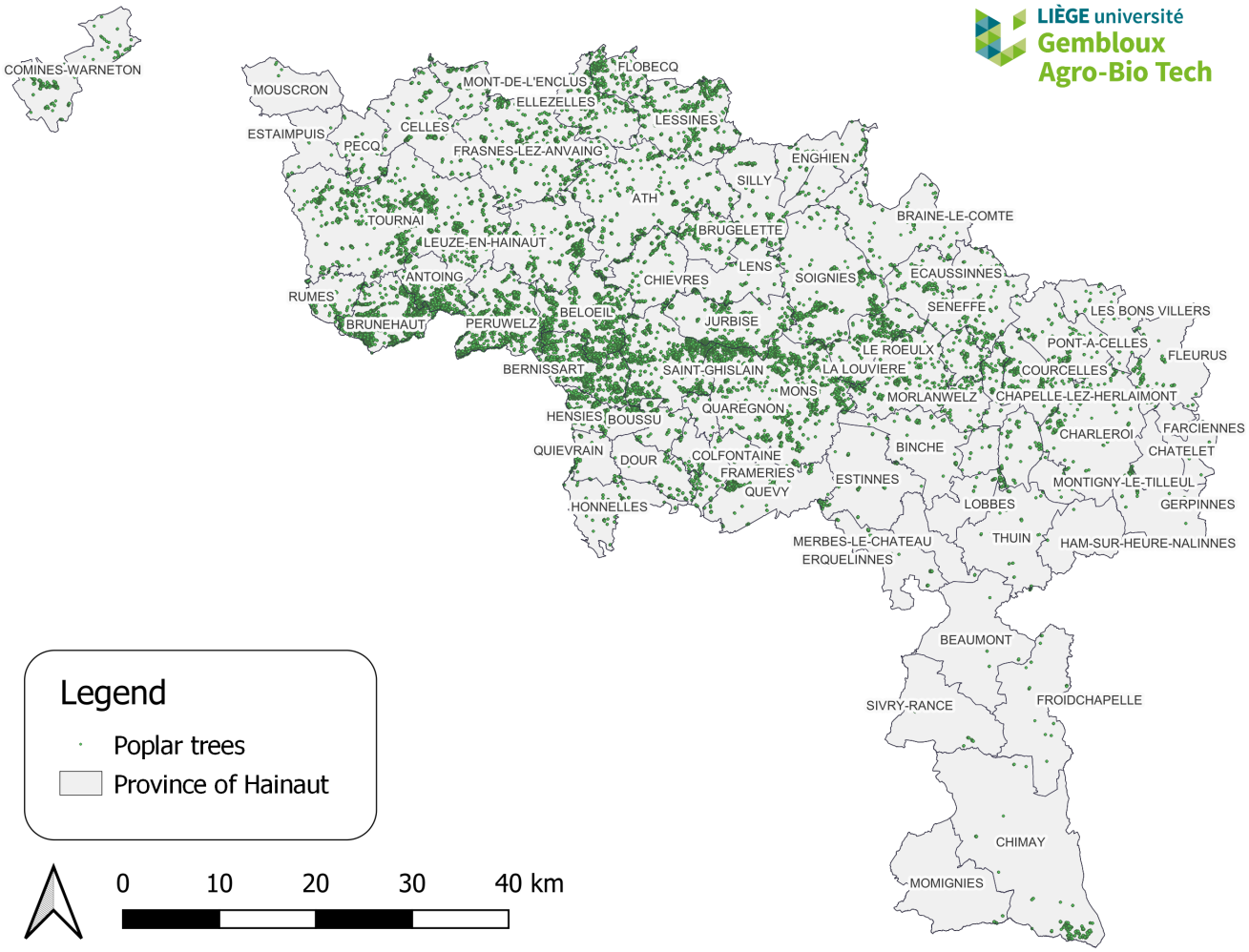
Figure 23: Illustration of three Computer Vision tasks: semantic segmentation, object detection and instance segmentation.

C Mapping of the poplar resource in the province of Hainaut

C.1 Map produced by the classification-based approach



C.2 Map produced by the semantic segmentation-based approach



Appendix References

Albawi, S., T. A. Mohammed, and S. Al-Zawi (Aug. 2017). "Understanding of a convolutional neural network". en. In: *2017 International Conference on Engineering and Technology (ICET)*. Antalya: IEEE, pp. 1–6. ISBN: 978-1-5386-1949-0. DOI: 10.1109/ICEngTechnol.2017.8308186.

Alkabbani, H. et al. (Apr. 2021). "Machine Learning and Metaheuristic Methods for Renewable Power Forecasting: A Recent Review". In: *Frontiers in Chemical Engineering* 3, p. 665415. DOI: 10.3389/fceng.2021.665415.

Alzubaidi, L. et al. (2021). "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions". In: *Journal of Big Data* 8.1, p. 53. ISSN: 2196-1115. DOI: 10.1186/s40537-021-00444-8.

Collins, C. et al. (2021). "Artificial intelligence in information systems research: A systematic literature review and research agenda". en. In: *International Journal of Information Management* 60, p. 102383. ISSN: 0268-4012. DOI: 10.1016/j.ijinfomgt.2021.102383.

Cresson, R. (2020). *Deep learning for remote sensing images with open source software*. CRC Press. ISBN: 978-1-00-302085-1.

Dastres, R. and M. Soori (2021). "Artificial Neural Network Systems". In: *International Journal of Imaging and Robotics* 21.2, pp. 13–25. ISSN: 2231–525X.

Kersting, K. (2018). "Machine learning and artificial intelligence: Two fellow travelers on the quest for intelligent behavior in machines". In: *Frontiers in Big Data* 1. ISSN: 2624-909X. DOI: 10.3389/fdata.2018.00006.

Mishra, R. K., G. Y. S. Reddy, and H. Pathak (Apr. 2021). "The understanding of deep learning: A comprehensive review". en. In: *Mathematical Problems in Engineering* 2021. Publisher: Hindawi, e5548884. ISSN: 1024-123X. DOI: 10.1155/2021/5548884.

Nikparvar, B. and J.-C. Thill (2021). "Machine learning of spatial data". en. In: *ISPRS International Journal of Geo-Information* 10.9. Publisher: Multidisciplinary Digital Publishing Institute, p. 600. ISSN: 2220-9964. DOI: 10.3390/ijgi10090600.

Park, S. R. and J. Lee (Sept. 2016). *A Fully Convolutional Neural Network for Speech Enhancement*. arXiv:1609.07132 [cs]. DOI: 10.48550/arXiv.1609.07132.

Zuo, Z. et al. (June 2015). "Convolutional recurrent neural networks: Learning spatial dependencies for image representation". In: *2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Boston, MA, USA: IEEE, pp. 18–26. ISBN: 978-1-4673-6759-2. DOI: 10.1109/CVPRW.2015.7301268.