Faculté des bioingénieurs

Application of LiDAR Data to Derive Key Forest Structure Metrics

Author	Orsolya Turi
Supervisors	Prof. Pierre Defourny Assoc.-Prof. Alexander Kmoch
Jurors	Dr. Mathieu Jonard Dr. Julien Radoux
Academic year	2024-2025
Programme	MSc Erasmus+ GEM - Geo-information Science and Earth Observation for Environmental Modelling and Management

Acknowledgements

I would like to thank my supervisors, Assoc.-Prof. Alexander Kmoch and Prof. Pierre Defourny, for their invaluable guidance and support throughout this thesis. I am also grateful to Mathieu Jonard and Mathilde Bonheure for providing the in situ dataset that made this research possible. Special thanks to Julien Radoux for initiating this thesis project and to Basile Boland for his helpful input and assistance during the process.

Table of Contents

Table of Figures

Table of Tables

List of Acronyms

ABA: Area-Based Approach

ADR: Area Difference Ratio

AGB: Aboveground Biomass

AGL: Above Ground Level

ALS: Airborne Laser Scanning

CBH: Canopy Base Height

CD: Crown Diameter

CE: Commission Error

CHM: Canopy Height Model

CSF: Cloth Simulation Filtering

DBH: Diameter at Breast Height

DR: Discrete Return

DTM: Digital Terrain Model

F: F1-score

FN: False Negative

FP: False Positive

FW: Full-Waveform

GNSS: Global Navigation Satellite System

GPS: Global Positioning System

GTR: Growing Tree Region

H: Height

IoU: Intersection over Union

IQR: Interquartile Range

ITS: Individual Tree Segmentation

LiDAR: Light Detection and Ranging

LM: Local Maxima

OE: Omission Error

P: Precision

P95: 95th Percentile

R: Recall

R^2: Coefficient of Determination

RF: Random Forest

RMSE: Root Mean Square Error

SD: Standard Deviation

SHAP: Shapley Additive exPlanations

TLS: Terrestrial Laser Scanning

TP: True Positive

1. Introduction

Forests are critical to global biodiversity, climate regulation, and the delivery of essential ecosystem services (Goldstein *et al.*, 2020). To manage these ecosystems sustainably, it is essential to monitor forest structure accurately and efficiently (Park *et al.*, 2021). This need has become even more urgent in the context of climate change, biodiversity loss, and evolving forest management goals. Traditional field-based methods for forest monitoring are time-consuming, costly, and limited in spatial extent (Aplin, 2005). In contrast, remote sensing techniques, and particularly LiDAR (Light Detection and Ranging), offer scalable and detailed alternatives (Hyde *et al.*, 2005). However, the question remains how accurately LiDAR can detect the structural characteristics of the forest.

Therefore, this research aims to extract key structural metrics from airborne LiDAR data and evaluate their accuracy against field measurements. A crucial step in this process is the delineation of individual trees, which must be reliable to ensure meaningful metric estimation. To address this, a novel segmentation method was also developed and tested. The results of this study are important not only for improving the understanding of LiDAR-based forest analysis but also for informing practical forest management, where accurate structural data can reduce reliance on field inventories, support ecological monitoring, and guide management decisions across larger areas.

This thesis is structured as follows: the Literature Review summarises key structural metrics relevant to forest management, examines how LiDAR is used to detect forest structure and introduces existing methods for LiDAR-derived metric calculation. It ends by outlining the research question and objectives. The Materials section describes the study area and datasets used. The remaining sections follow the classic structure: Methodology, Results, Discussion, and Conclusion.

2. Literature Review

2.1 Key Forest Structure Metrics for Management

As mentioned before, assessing forest structure is crucial for sustainable forest management, as it directly influences ecosystem processes such as nutrient cycling and biodiversity (Shugart *et al.*, 2010; Park *et al.*, 2021). Effective monitoring of forest structure relies on quantifying key metrics, including tree height, canopy base height (CBH), diameter at breast height (DBH), crown width, aboveground biomass (AGB), stem volume, tree density and tree species. These variables provide valuable insights into forest productivity, carbon storage, habitat quality, and stand dynamics, supporting data-driven forest management and conservation strategies (Morsdorf *et al.*, 2010; Nie *et al.*, 2017; Diószegi *et al.*, 2024).

Tree height is a fundamental metric for evaluating forest maturity, productivity, and habitat suitability. It serves as a primary input for estimating AGB and stem volume (Allouis *et al.*, 2013). CBH is another essential parameter, widely used for fire behaviour modelling (Luo *et al.*, 2018; Yang *et al.*, 2020), since it determines the vertical stratification of fuels, affecting the likelihood of crown fires (Alpine and Hobbs, 1994 as cited in Yang *et al.*, 2020). In addition, it serves as an indicator of foliage density and forest health (Luo *et al.*, 2018). DBH is among the most commonly used metrics in forestry, as it provides crucial information on tree volume, stand density, and forest productivity (Jucker *et al.*, 2017). It also serves as a key input for allometric equations to estimate biomass and stem volume (Allouis *et al.*, 2013).

Crown diameter and crown ratio are critical indicators of tree vigour, growth potential, and competitive status, and they influence resource availability through light interception (Yang *et al.*, 2020). Crown dimensions are particularly relevant in predicting DBH, as studies have shown strong correlations between crown and stem diameter. In addition, they can improve AGB estimation accuracy (Ploton *et al.*, 2016 as cited in Ni-Meister *et al.*, 2022). AGB is a critical metric for assessing forest productivity and carbon sequestration rates (He *et al.*, 2012). It serves as a key parameter in global carbon cycle studies and climate change mitigation efforts (Ni-Meister *et al.*, 2022). Stem volume, closely related to AGB, is used to evaluate timber yield and forest growth, informing sustainable harvesting plans.

Tree density is also a crucial metric in forest management, offering insights into stand structure, competition, and regeneration dynamics. High-density stands often experience intense competition for resources, leading to reduced growth rates among individual trees. Understanding tree density aids in making informed decisions, for instance, with thinning practices (Zhang *et al.*, 2024). Forest managers also benefit from information on individual tree species, as species composition significantly influences timber value, habitat quality, and susceptibility to environmental stressors (Vaughn *et al.*, 2012).

Based on their relevance for forest management and their detectability from LiDAR data, a subset of these structural metrics was selected for this study, as detailed in later sections. To understand how these parameters can be derived remotely, the next section introduces the fundamentals of LiDAR and its application in characterising forest structure.

2.2 LiDAR Fundamentals for Forest Structure Charaterisation

LiDAR is an active sensor that emits laser pulses and detects the return signal that is caused by scattering objects (like the canopy). The location of the intersected object can be obtained by measuring the travel time of the pulse with the scan angle (Wang and Kumar, 2019) and registering the sensor's position and orientation in space (Fernandez-Diaz and Carter, 2013). The intensity of the received signal provides additional information about the reflectance and surface properties of the scatterer (Wing *et al.*, 2012).

One of the advantages over optical remote sensing and photogrammetry is that LiDAR pulses can penetrate the canopy layer (Mallet and Bretar, 2009), hence it directly measures the 3D structure of forests (Lefsky *et al.*, 2009). It provides horizontal and vertical information on different canopy layers, including the understory (Ruiz *et al.*, 2018). LiDAR sensors can be classified based on the size of the covered area by one laser pulse (footprint), sensor platform, and type of system. The different LiDAR sensors are suitable for distinct purposes in forest characterisation. For instance, large footprint (> 10 m) LiDAR systems can measure common forest variables with high accuracy (e.g. canopy height) (Dubayah and Drake, 2000, as cited in Morsdof *et al.*, 2010), while small footprint (< 1 m) systems are suitable for forest detection to the level of single trees (Hyyppa *et al.*, 2001).

As for the sensor platforms, LiDAR systems can be ground-based (Terrestrial Laser Scanning; TLS), airborne (Airborne Laser Scanning; ALS), and operated on satellites (spaceborne) (Anderson *et al.*, 2016). TLS provides the most detailed and accurate data due to its proximity to the trees, making it particularly effective for capturing understory vegetation (Vierling *et al.*, 2013). However, its coverage is limited to small areas (Crespo-Peremarch *et al.*, 2018). In contrast, spaceborne LiDAR systems cover large regions but offer lower resolution due to the distance from the target area (Wulder *et al.*, 2012). ALS offers a balance between coverage and resolution. It can survey large areas, such as entire countries while maintaining high resolution (Crespo-Peremarch *et al.*, 2018).

Regarding the LiDAR system type, there are two options: discrete return (DR) and full waveform (FW) sensors (Hancock *et al.*, 2017). With DR the return signal is digitized into a small number of 3D coordinates, including the first and last return pulse (associated with the top of the canopy and the ground) (Wang and Kumar, 2019), and some intermediate signals that are likely the lower canopy or understory (Hall *et al.*, 2005 as cited in Hermosilla *et al.*, 2014; Wang and Kumar, 2019). It creates a 3D point cloud from the recorded information. As for FW LiDAR, it digitizes the entire return signal as a continuous wave and stores the wave's shape (Mallet and Bretar, 2009). Therefore, it records more scatterers and provides additional 1D information, where the wave's shape shows the characteristics of the intersected object, such as the spatial arrangement (Bruggisser *et al.*, 2017). Both DR and FW systems have advantages and limitations, which are important to consider in forest characterisation.

DR LiDAR is the most widely used type of system, valued for its simplicity and broad software support. However, it has limitations in capturing full vegetation structure due to blind spots between recorded points, which can hinder detection of the understory and lower canopy (Reitberger *et al.*, 2008; Hancock *et al.*, 2017). In contrast, FW LiDAR captures the entire return signal, offering more detailed information about forest structure (Wang and Kumar, 2019). This added complexity increases data volume and processing time, but also allows for a more complete characterisation of the canopy. While DR remains more practical in many applications, FW is becoming increasingly viable as processing tools improve (Mallet and Bretar, 2009).

This study uses point cloud data derived from FW LiDAR due to its availability and increased ability to characterise the total forest structure. The next section outlines the metrics selected for this study and reviews how they can be calculated from LiDAR data, along with the associated challenges and considerations reported in the literature.

2.3 Deriving Forest Structure Metrics from LiDAR

The selected metrics for this study are: tree height, CBH, crown diameter, DBH, and AGB. They were chosen based on their relevance to forest management (see Section 1.1), the availability of in situ validation data, and the possibility of calculating them from LiDAR data.

Forest structure metrics can be calculated using two main approaches: area-based and individual tree-based (Bouvier *et al.*, 2015). The area-based approach (ABA) aggregates forest attributes over a defined area and uses statistical relationships between LiDAR metrics and field plot data to estimate stand characteristics like mean canopy height, volume and stem density (Næsset, 2002). On the other hand, individual tree-based approaches focus on detecting and measuring individual trees from LiDAR point clouds. These methods provide detailed tree-level metrics, such as height, crown width, and species, which can be used to generate more precise inventories and enhance forest management practices (Lindberg *et al.*, 2010). Although ABAs have traditionally been more widely used due to the sparse point densities of older LiDAR datasets, improved data acquisition techniques are making individual tree calculations increasingly feasible (Ayrey *et al.*, 2017).

This research focuses on individual tree-based approaches, as they offer greater precision for forest management and can be aligned with traditional field-based inventories (Vastaranta *et al.*, 2009; Luo *et al.*, 2018). Nevertheless, this review introduces other methods to give a general overview of available calculations.

The first step in individual tree-based methods is tree segmentation, where individual trees are identified and delineated from the LiDAR point cloud. The accuracy of segmentation influences any further analysis; therefore, selecting a reliable method is essential. The next section introduces the available segmentation methods.

2.3.1 Individual Tree Segmentation (ITS)

Individual Tree Segmentation (ITS) requires the accurate detection of treetops, followed by crown delineation, to effectively segment individual trees from the surrounding canopy. ITS methods can be broadly classified into point cloud-based and CHM-based (Canopy Height Model) approaches, with some methods integrating both.

Point cloud-based methods segment individual trees directly from the point clouds by clustering the points into distinct objects (Duncanson *et al.*, 2014). These techniques provide high accuracy in detecting tree structure and can be effective in dense, multi-layered forests as well (Reitberger *et al.*, 2009; Vega *et al.*, 2014). However, their performance depends on point density and requires computationally intensive processing when segmenting large areas.

CHM-based methods rely on rasterised canopy surface (CHM) derived from LiDAR data, making them suitable for lower point-density datasets and computationally more efficient (Diószegi *et al.*, 2024). Treetops are detected using local maxima (LM) algorithms with fixed or variable window sizes (Yang *et al.*, 2020), followed by crown delineation through watershed segmentation (Ayrey *et al.*, 2017), region-growing algorithm (Popescu *et al.*, 2002), or other techniques. The accuracy of CHM-based segmentation relies on the created CHM. Many techniques are available, some require post-processing to smooth the detected canopy surface, reducing errors and artefacts (Mielcarek *et al.*, 2018; Stereńczak *et al.*, 2020), while others integrate these corrections in earlier steps (Khosravipour *et al.*, 2014).

CHM methods are the most common ITS methods, even though they face limitations in detecting small trees beneath the canopy, as they do not show up on the CHM. In addition, they struggle in heterogeneous forests where neighbouring trees may not appear as clearly separated local maxima (Reitberger *et al.*, 2009).

Several studies have shown that point cloud-based ITS methods generally outperform CHM-based approaches. For instance, Ayrey *et al.* (2017) reported tree detection rates of 66–89% using point cloud-based methods across different forest conditions, while CHM-based methods on the same plots achieved only 42–76%. Similarly, Vega *et al.* (2014) found that PTrees, a point cloud-based algorithm, achieved detection rates between 80% and 95% across three forest types. These values are higher than those typically reported for CHM-based approaches, which range from 43% to 96% depending on canopy structure and forest type (Solberg *et al.*, 2006; Yu *et al.*, 2014). This suggests that point cloud-based methods offer more consistent and accurate tree detection, especially in structurally complex forests.

Recent ITS methods increasingly integrate machine learning to improve segmentation accuracy. Some approaches apply machine learning to CHMs, training models to delineate individual tree crowns based on pixel patterns (Chang *et al.*, 2022). However, these methods inherit the limitations of CHM-based segmentation (Henrich *et al.*, 2024), such as difficulty detecting overtopped or suppressed trees. More advanced strategies directly apply 3D deep learning to point clouds, using networks trained on already segmented data to recognise and separate individual trees (Wilkes *et al.*, 2023; Henrich *et al.*, 2024). While these methods offer high accuracy, they are typically designed for high-resolution ALS or dense TLS data. Their main drawbacks include long processing times, dependence on large, high-quality training datasets, and limited applicability to lower-density datasets (Henrich *et al.*, 2024). Nevertheless, when sufficient data and computational resources are available, these methods offer significant performance improvements.

When selecting an ITS method, the primary goal is to achieve high segmentation accuracy. However, this must be balanced with practical considerations such as the computational cost of the method and the quality and resolution of the available LiDAR data. High-performing approaches like 3D deep learning may not be feasible for lower-density datasets or large-scale studies due to their processing demands and training data requirements. Beyond the technical approach, forest type plays a significant role in ITS performance (Hamraz *et al.*, 2016). Many existing methods were developed and tested primarily in coniferous or open-canopy forests, where tree spacing and crown shapes are more regular (Kaartinen *et al.*, 2012). In contrast, temperate deciduous forests pose greater challenges due to higher species diversity, more complex crown architecture, and denser canopies, making individual trees harder to detect (Vega *et al.*, 2014; Hamraz *et al.*, 2016). As such, segmentation methods must not only be accurate but also generalisable across diverse forest conditions (Vauhkonen *et al.*, 2012).

Once individual trees are delineated, key structural attributes can be extracted. The following sections introduce the selected structural metrics and methods used for their calculation from LiDAR data.

2.3.2 Tree Height

Depending on the available data and methodology, canopy height can be derived using different approaches. When working with FW LiDAR, canopy height can be calculated as the difference between the first and last return heights (Mallet and Bretar, 2009). However, this method is sensitive to errors in ground detection, particularly in uneven terrain or dense forests. Advanced techniques mitigate these challenges through pre-processing steps such as noise removal (Hyde *et al.*, 2005) and waveform smoothing (Hermosilla *et al.*, 2013). Alternatively, canopy height can be estimated from ITS, either by selecting the highest point within each segmented tree (Vega *et al.*, 2014) or by measuring the height of treetops detected from a CHM (Ayrey *et al.*, 2017; Diószegi *et al.*, 2024). However, the accuracy of these methods depends on the success of tree segmentation.

LiDAR-derived canopy height can be subject to both under- and overestimation due to various factors. It can be affected by errors in ground detection caused by complex terrain (Hyde *et al.*, 2005) or dense understory vegetation that obstructs ground returns (Lefsky *et al.*, 2009). Detecting the uppermost part of the canopy also presents challenges (Lefsky *et al.*, 2009). Pointed crowns, common in conifers, can cause LiDAR pulses to penetrate partially, causing underestimated height (Hyde *et al.*, 2005). In addition, with small footprint LiDAR, it is less likely to hit the apex of crowns, which leads to biases in complex and coniferous forests (Lim *et al.*, 2003).

2.3.3 Canopy Base Height (CBH)

CBH is the distance between the ground and the live canopy. It can be calculated using two main approaches: direct measurement and regression-based estimation.

Direct methods typically derive CBH by percentile ranking (Luo *et al.*, 2018) or using frequency statistics from the LiDAR data's vertical profile (Solberg *et al.*, 2006; Chamberlain *et al.*, 2021). These methods do not require field-based measurements, offering greater flexibility to be used in diverse forests (Stefanidou *et al.*, 2020). Nevertheless, these methods also depend heavily on the quality of point-cloud data and sensor capabilities (Yang *et al.*, 2020). For instance, CBH overestimation may occur due to insufficient LiDAR pulse penetration, particularly in dense canopies or complex terrain (Luo *et al.*, 2018).

Regression-based methods predict CBH through statistical models that establish relationships between field-measured CBH and other LiDAR-derived forest parameters, such as tree height and crown width (Popescu and Zhao, 2008). These models require high-quality field data and are often species- and site-specific, limiting their generalizability to other areas or forests with different conditions (Yang *et al.*, 2020).

Both approaches have their strengths and weaknesses. Direct methods are more flexible but may overestimate CBH in dense forests, while regression-based methods are more accurate but depend on high-quality field data and may not generalise well to other areas.

2.3.4 Crown Diameter

Crown diameter is the average width of a tree's canopy, typically measured across two perpendicular axes (North-South and East-West). It is commonly calculated after tree segmentation in LiDAR data, based on the delineated tree crowns (Wang *et al.*, 2008; Jucker *et al.*, 2017; Yang *et al.*, 2020). As a result, the accuracy of crown diameter estimation is highly dependent on the method used for ITS.

Various ITS approaches, such as those based on region-growing algorithms (Popescu *et al.*, 2002) or clustering (Duncanson *et al.*, 2014), can produce varying results in terms of crown delineation accuracy. Factors such as point density, canopy structure, and the presence of neighbouring trees can further complicate accurate segmentation (Stereńczak *et al.*, 2020), leading to errors in crown diameter estimation. For example, in dense or overlapping canopy areas, the method might struggle to distinguish individual trees, potentially underestimating or overestimating crown dimensions. Therefore, the reliability of crown diameter as a metric is closely tied to the performance and limitations of the ITS method employed.

2.3.5 Diameter at Breast Height (DBH)

DBH represents the stem diameter of a tree measured at 1.3 meters above the ground (Shugart *et al.*, 2010). While DBH cannot be directly measured from LiDAR (Jucker *et al.*, 2017), it is estimated after tree segmentation using allometric relationships or regression modelling (Jucker *et al.*, 2017; Ni-Meister *et al.*, 2022).

The power-law relationships suggest that tree height should scale with diameter. However, this fails to account for varying allometries between species, climate, and stand structure (Jucker *et al.*, 2017), and does not capture the asymptotic nature of height growth (Chave *et al.*, 2014). For instance, trees typically grow rapidly in height when young but invest more in diameter as they mature, leading to varying diameters for trees of similar height (King, 2005). Therefore, tree height on its own is not enough to predict DBH. Some studies have shown that crown diameter is strongly correlated with stem diameter even in large trees (Ploton *et al.*, 2016 as cited in Ni-Meister *et al.*, 2022). These studies indicate that incorporating both tree height and crown dimensions has the potential to develop robust models for estimating DBH, which can be accurate not only at a global scale but also at the local level (Yao *et al.*, 2012; Jucker *et al.*, 2017; Ni-Meister *et al.*, 2022).

2.3.6 Aboveground Biomass (AGB)

Most AGB estimation models rely on allometric equations, which traditionally use stem diameter (DBH) as the primary predictor (Jucker *et al.*, 2017). However, since DBH cannot be directly measured with LiDAR, alternative approaches focus on height-based models (Nie *et al.*, 2017). AGB shows a strong correlation with forest height, however, this relationship is not always satisfactory because tree growth dynamics lead to significant variation in DBH among trees of similar height (Jucker *et al.*, 2017; Ni-Meister *et al.*, 2022). Large trees, which store the majority of forest biomass, continue to grow in diameter even when height growth slows, making diameter estimation difficult (King, 2005).

To improve AGB estimates, researchers have begun incorporating crown dimensions into models (Jucker *et al.*, 2017; Ni-Meister *et al.*, 2022). Crown diameter is strongly linked to stem diameter, as trees invest in lateral expansion to maintain structural stability and hydraulic function (Hemery *et al.*, 2005). Studies have shown that models integrating both height and crown size perform better than height-only models, particularly for large trees (Jucker *et al.*, 2017; Ni-Meister *et al.*, 2022). While integrating multiple structural metrics has improved AGB estimation, challenges remain in developing models that are both accurate and widely applicable across different forest types.

2.4 Research Question and Objectives

This study investigates how accurately key individual tree-level structural metrics relevant to forest management can be derived from airborne LiDAR data in temperate deciduous forests. The focus is placed on upper canopy trees, as the available LiDAR dataset has limited capacity to detect lower canopy layers (see Section 2.3). Despite this constraint, upper canopy metrics remain highly valuable for both ecological assessment and forest management, as these trees drive most biomass accumulation, carbon storage, and canopy-level competition processes.

The broader aim of the research is to develop a workflow that supports large-scale, repeatable forest structure assessment using LiDAR data, with minimal reliance on in situ measurements. This transferability objective guided key methodological choices, such as prioritising direct metric extraction over regression modelling wherever feasible, and developing a segmentation approach that can be adapted across forest types.

Research Question

How accurately can LiDAR data capture individual tree-level structural metrics relevant to forest management according to different temperate deciduous forest structures?

Objectives

1. Metric Extraction and Validation

Derive individual tree-level structural metrics—tree height, diameter at breast height (DBH), crown base height (CBH), crown diameter, and aboveground biomass (AGB)—for upper-canopy trees using LiDAR data. Validate these metrics by comparing them against field-based measurements to assess accuracy.

2. Forest Condition Comparison

Compare the accuracy of LiDAR-derived metrics across forest stands with differing structural characteristics, including variation in species composition (pure vs. mixed), age structure (regular vs. irregular), and canopy density. Assess how these conditions influence metric performance.

3. Segmentation Method Development and Evaluation

Develop and evaluate a novel ITS method. The approach combines an existing point cloud-based segmentation algorithm (PTrees) with a Random Forest classifier to identify and iteratively correct segmentation errors. The method is designed to improve segmentation accuracy without requiring high computational cost or dense point cloud data, enhancing its usability in broader applications. To evaluate the performance of the new method, it was compared to a widely used ITS approach, serving as a benchmark.

By addressing these objectives, the study aims to improve the operational use of LiDAR for forest inventories, offering a workflow that balances accuracy and scalability, and has potential for application in diverse forest types with minimal ground-based input.

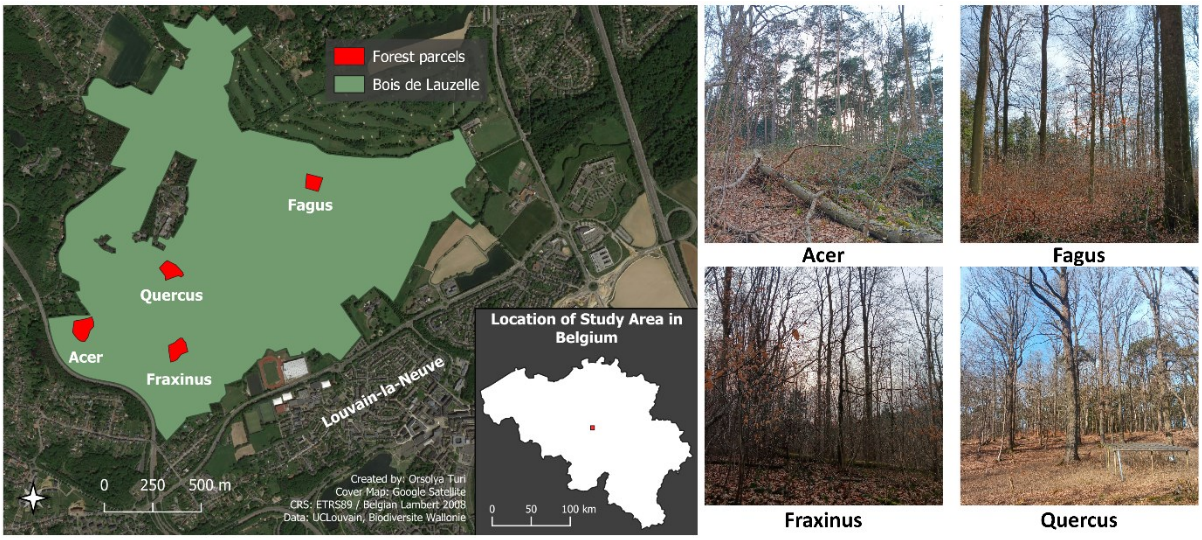
3. Materials

This section describes the datasets used in the study, including the in situ measurements and the airborne LiDAR data, along with the identified limitations, particularly concerning the spatial accuracy of the in situ tree locations.

3.1 Study Area and In Situ Data

The study was conducted in the Bois de Lauzelle, a peri-urban forest located near Ottignies-Louvain-la-Neuve in Belgium (Figure 1). Spanning approximately 200 hectares, this forest is owned and managed by the Université catholique de Louvain. Due to its affiliation with the university, the forest fulfils important educational and research functions (Bonheure 2022).

Our research benefits from the availability of in situ data collected within the forest, which serves as validation for LiDAR-derived forest structural metrics. The data was collected from four parcels (Figure 1), each representing temperate deciduous forest stands with varying age structures, species composition, and tree densities (Table 1).



Created by: Orsolya Turi
Cover Map: Google Satellite
CRS: ETRS89 / Belgian Lambert 2008
Data: UCLouvain, Biodiversite Wallonie

Acer Fagus

Fraxinus Quercus

Figure 1: Study area and location of parcels.

This diversity makes them suitable for addressing the objectives of this study. The parcels include[1]:

- ***Acer***: irregular (meaning uneven-aged) stand dominated by maple and beech
- ***Fagus***: regular (meaning even-aged) pure beech stand

---

[1] Parcel name refers to the dominant species.

- ***Fraxinus***: irregular stand dominated by ash and beech
- ***Quercus***: regular pure oak stand

Table 1: Characteristics of the forest parcels.

Parcel Name	***Acer***	***Fagus***	***Fraxinus***	***Quercus***
Area (ha)	1.95	1.68	1.56	1.68
In situ trees	253	99	275	177
Dominant species	*Acer pseudoplatanus* *Fagus sylvatica*	*Fagus sylvatica*	*Fraxinus excelsior* *Acer pseudoplatanus*	*Quercus petraea*
Density (trees/ha)	130	59	176	105
Dominant height[2] (m)	25.1	30.8	29.2	19.6
Age Structure	Irregular	Regular	Irregular	Regular
Slope (°)	24	0	14	13

The in situ dataset includes information for all trees that exceeded a DBH threshold of 6 to 10 cm, depending on the stand. For each tree, relative x and y coordinates (expressed with respect to a local reference point), total height (m), canopy base height (m), and girth at breast height (measured at 1.3 m; in cm) were recorded. Crown radii (m) in the four cardinal directions were also measured to estimate crown dimensions (Bonheure 2022). Aboveground biomass (kg C/ha) was estimated at the end of the growing season using species-specific allometric equations (Jonard *et al.*, 2020).

Tree location and structural measurements were taken using Field-Map, a system developed by the Forest Ecosystem Research Institute (IFER, 2010). Field-Map integrates a laser rangefinder, an electronic clinometer, a digital compass, and a computer, enabling efficient georeferenced data collection. Tree heights and crown metrics were measured using a Vertex device, while circumference was measured with a tape. Slope was characterised with a clinometer and compass, and GPS was used to delineate the boundaries of the parcels. Data collection took place in July 2021, with the exception of the *Fagus* parcel, which was measured in 2020 by laboratory technicians (Bonheure 2022).

---

[2] Dominant height: average height of the 100 trees/ha with the biggest diameter at breast height (Bonheure, 2022).

Although the in situ dataset is generally reliable, some issues have to be highlighted. In the *Fraxinus* parcel, some trees were cut down after the field data was collected but before the LiDAR data was acquired (in January-March 2022), which causes mismatches between the two datasets.

Another problem was related to the tree coordinates. Only a single coordinate was originally provided per parcel without confirming whether it corresponded to the reference location for the relative tree coordinates. When these potential reference coordinates were used, many of the trees appeared outside their plot boundaries (Figure 2A). To resolve the issue, additional field measurements were carried out in February 2025 to estimate the true spatial origin (true reference coordinate) for each parcel.

3.2 In Situ Coordinate Estimation

Four easily identifiable trees were selected in each plot, and their positions were measured using a Trimble Geo7x differential GPS (dGPS), which offers a field accuracy of approximately 10 cm under optimal conditions and a calibration accuracy of 7.5 cm.

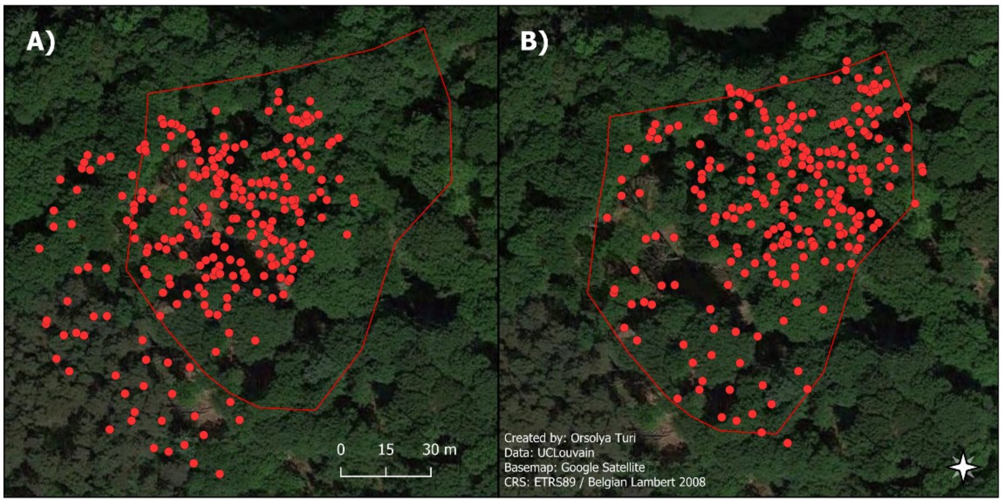


Figure 2: Tree locations in the Acer Parcel. A) Based on the provided potential reference coordinate. B) After applying the average coordinate shift derived from dGPS field measurements in February 2025.

The newly measured tree coordinates revealed inconsistencies: the required shift to align the trees correctly varied within the same parcel, sometimes by more than 5 meters. This indicates that the relative coordinates of some trees may have been recorded using different or imprecise reference locations. As a result, the spatial position of trees appears to be distorted, and no true reference coordinates were found. To improve the tree locations, an average shift was calculated based on the new measurements and applied to each parcel's potential reference coordinate (Figure 2B and Table 2). While this method helped to reduce the overall error, it could not fully correct the inconsistencies. This spatial mismatch makes it difficult to accurately link in situ measurements with LiDAR-derived tree data, introducing uncertainty into later parts of the analysis, especially when comparing individual trees between the two datasets. Attempts were also made to quantify the uncertainty in the field data tree positions (see Section 3.2.6).

Table 2: Average coordinate shifts applied per parcel. Spatial corrections (in meters) applied to the potential reference coordinate of each parcel based on GPS measurements. Values represent shifts in the latitude (north–south) and longitude (east–west) directions.

Parcel	***Acer***	***Fagus***	***Fraxinus***	***Quercus***
Latitude Shift (m)	+18.0	+3.6	+2.2	-17.7
Longitude Shift (m)	+35.5	-42.4	-18.9	-8.9

3.3 LiDAR Dataset and Point Cloud Properties

The LiDAR dataset used in this study was acquired between January and March 2022 through multiple overlapping flights over the Wallon region. Each flight stripe overlapped by approximately 60% to ensure full coverage. The acquisition was conducted using two RIEGL VQ-780 II-S FW LiDAR sensors to ensure higher point density (SPW, 2024).

Table 3: LiDAR data acquisition characteristics.

Characteristics	
Aircraft speed	314.8 km/h
Flying altitude AGL[3]	2400 m
Pulse repetition frequency	900 kHz
Pulse density (emitted)	6.8 pulses/m^2
Point density (minimum)	1.5 points/m^2
Footprint size	50 cm

Despite this setup, the resulting point cloud density remains relatively low (11.2 – 14.4 points/m^2 in the parcels), mainly due to the high flight altitude (Table 3). As a result, this study focuses only on the upper canopy layer, where point density is sufficient for extracting reliable structural metrics.

The raw LiDAR data was first processed by the data producer using RIEGL's proprietary software, RiAnalyze and RiWorld, which performed full waveform analysis and converted the raw echo data

---

[3] AGL: Above Ground Level

into Cartesian coordinates. Further processing, including point classification and projection, was carried out using the TerraScan and TerraModeler modules. Automatic classification was used to distinguish ground from non-ground points. After this, manual quality control was performed (SPW, 2024). Despite these steps, a significant number of points (mainly near the ground surface) remained unclassified, which had to be addressed later.

The final dataset is delivered in LAZ 1.2 format, containing only the processed 3D point cloud. It is projected in the ETRS89 / Belgian Lambert 2008 coordinates reference system. This pre-processed and classified point cloud forms the basis for all subsequent analyses in this study.

Although the sensors used during the data acquisition are capable of recording FW data, the acquisition was performed using the smart waveform setting. This mode stores only the most significant echo peaks, those that generated a return point, and discards the in-between return signal. As a result, the dataset does not contain FW data, and certain advanced analyses typically enabled by waveform information are not possible.

4. Methodology

This section outlines the processing workflow applied to the LiDAR dataset **during this research**. The workflow is divided into three main stages: **(1) Pre-processing**, **(2) Individual Tree Segmentation**, and **(3) Structural Metric Calculation** (Figure 3).

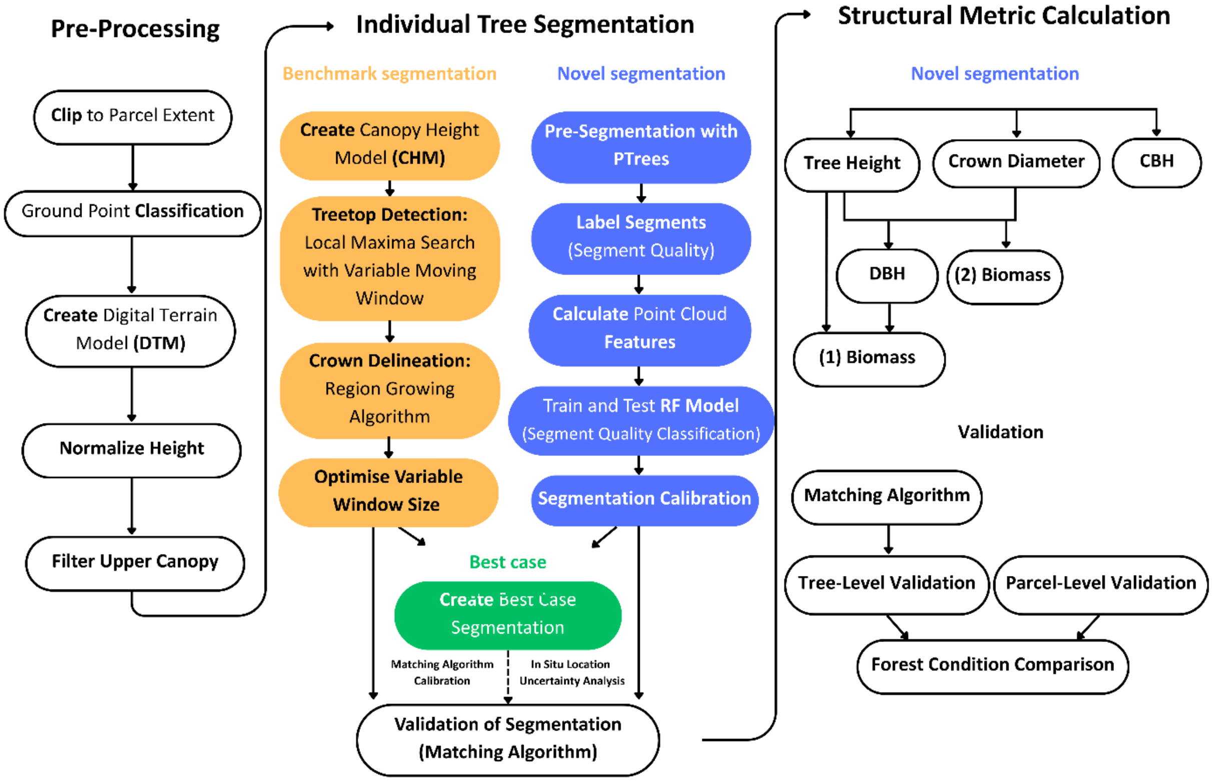


Figure 3: LiDAR data processing flowchart.

Each part is explained in detail in the following subsections. All processing was performed using a combination of **R** (R Core Team, 2023) (primarily the **lidR** library (Roussel *et al.*, 2020; Roussel, 2024) and its experimental extension **lidRplugins**), **LAStools** (Isenburg, 2022), and **QGIS** (QGIS Development Team, 2024). Use of lidRplugins required manual editing of outdated package dependencies. The ultimate aim of this workflow is to evaluate different segmentation approaches, extract individual tree metrics, assess their accuracy using field data, and compare accuracy across forest types.

4.1 LiDAR Pre-Processing

The LiDAR point cloud was first clipped to the extent of each forest parcel. A small buffer was added to include the trees near parcel edges fully. A high proportion of ground-level points were unclassified in the original dataset, around 12-30% depending on the parcel. These points were reclassified using Cloth Simulation Filtering (CSF). CSF works by turning the point cloud upside down and simulating a cloth falling over it (Zhang *et al.*, 2003). The cloth settles on the surface, and the points it touches are classified as ground.

From the classified ground points, a Digital Terrain Model (DTM) of 50 cm resolution was generated to represent bare-earth elevation. This DTM was then used to normalise the height of the point cloud, converting elevations above sea level into above-ground heights.

The final step of pre-processing involved filtering the point cloud to retain only upper canopy points. A grid was applied to each parcel, and the 95th percentile height was calculated within each cell to estimate the height of the tallest trees. The overall height range of the point cloud was then analysed in each cell, and only the top percentage of this range was kept. This allowed the exclusion of understory and lower canopy layers, which are poorly represented due to the low point density of the LiDAR dataset. In addition, cells with low 95th percentile height were removed to exclude areas with small or regenerating trees, such as those resulting from recent thinning or cutting. To further refine the filtered canopy layer, small isolated clusters were filtered out based on their size and distance from neighbouring clusters. This step helped eliminate residual noise. The final dataset focused on well-represented upper canopy trees.

4.2 Individual Tree Segmentation (ITS)

This section describes the methods used to segment individual trees from the pre-processed LiDAR point cloud and their validation. The segmentation process is a key step in the analysis, as it enables the extraction of tree-level structural metrics. Two segmentation approaches were evaluated: a benchmark method (CHM-based) and a novel method (point cloud-based with machine learning).

To evaluate the segmentation results, segmented crowns must be matched to field-measured trees. Due to known positional inaccuracies in the in situ dataset, a custom matching algorithm was developed. A best-case segmentation dataset was created to serve as a reliable reference to calibrate this algorithm. This dataset combines results from the benchmark and novel segmentation methods, supplemented with manual corrections using LAStools. Therefore, this represents the most accurate approximation of true tree positions and crowns. In addition to the calibration, the best-case segmentation was also used to analyse the spatial uncertainty of the in situ data.

4.2.1 Benchmark Segmentation Approach

A CHM-based segmentation method was implemented as a benchmark to evaluate the novel tree segmentation approach. The CHM was generated using the pit-free algorithm from lidR (Khosravipour *et al.*, 2014), which constructs a CHM surface by applying height thresholds to the point cloud and performing Delaunay triangulation on first returns. Large triangles were removed to reduce artefacts, and the highest surface value across all layers was retained, producing a smooth CHM without requiring post-processing.

Treetops were identified using local maxima search with a variable moving window that scales with tree height, assuming that taller trees have wider crowns (Eysn *et al.*, 2015; Dalponte and Coomes, 2016). Crown regions were delineated using a region-growing algorithm that expands from each treetop by comparing neighbouring CHM pixel heights. It is important to note that this segmentation method does not assign all LiDAR points to crowns; some remain unsegmented (Dalponte and Coomes, 2016).

To optimise performance, several combinations of moving window scaling parameters were tested. After validation, the configuration that yielded the best segmentation accuracy across all parcels was selected for the final benchmark segmentation dataset.

4.2.2 Novel Segmentation Workflow

As one of the central objectives of this research, a novel segmentation approach was developed to improve ITS in complex deciduous canopies. This method builds upon the PTrees algorithm, which is a point cloud-based segmentation approach (Vega *et al.*, 2014). Therefore, unlike CHM-based methods, PTrees is better suited for detecting trees that are overtopped, intertwined, or located in irregular canopy structures. In this study, PTrees was extended with a custom-built workflow involving Random Forest (RF) classification to evaluate and iteratively refine the quality of tree segments. The goal was to address common segmentation quality errors, such as oversegmentation (partial crowns) and undersegmentation (merged trees).

4.2.3 PTrees Algorithm

The PTrees segmentation algorithm consists of 4 main steps: (i) multi-scale segmentation, (ii) segment scoring, (iii) dynamic apex selection, and (iv) final segmentation (Vega *et al.*, 2014).

The only adjustable parameter is the k-value, which defines the size of the local neighbourhood used to detect tree apices and influences crown size.

The segmentation begins with a multi-scale segmentation step, where the algorithm runs several passes over the point cloud using a range of decreasing k-values (e.g., from large to small neighbourhoods) to identify potential tree apices. For each pass, points are processed from highest to lowest elevation. A point is considered a tree apex if it is higher than its k nearest neighbours; if not, it is assigned to an existing segment based on minimal shape deformation. This results in a set of crown segments at various scales.

Each segment is then scored based on crown size, circularity, apex-centeredness, and shape regularity. These scores are used in the dynamic apex selection step, where segmentations from different scales are compared. Higher-scoring apices from finer scales (smaller neighbourhood size) replace lower-quality segments from coarser scales. This process continues until the best-scoring apices across all scales are selected.

In the final segmentation step, tree crowns are grown from the selected apices. Since no new apices are allowed, unlabeled points are assigned to existing segments using neighbourhood sizes that expand from kmin to kmax, ensuring that all points are ultimately classified into a tree segment.

4.2.4 Random Forest Classification of Segment Quality

While PTrees delivers high-quality segmentation, some errors remain in our deciduous forest parcels, particularly oversegmentation (partial trees) and undersegmentation (merged trees). To detect these errors, RF model was trained to classify the quality of segments.

The training and testing data for the RF was generated through a pre-segmentation step, where PTrees was run on all parcels using fixed k-values (*k = 100* and *k = 60*) to intentionally produce a wide range of segmentation outcomes. Segments were then manually labeled into three classes: oversegmented (Class 0), correctly segmented (Class 1), and undersegmented (Class 2) (Figure 4). In total, 467 segments were labeled: 122 oversegmented, 191 correctly segmented, and 154 undersegmented.

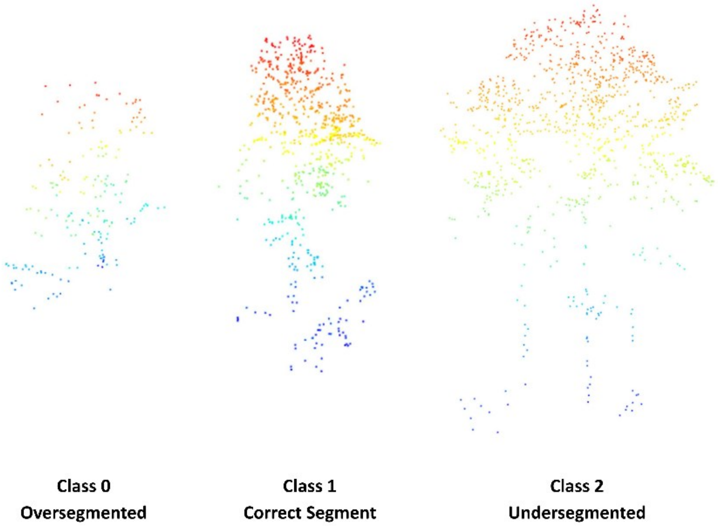


Figure 4: Segmentation quality classes.

Each segment was described using 15 features that captured its point cloud height distribution, point density, and geometric shape (see Section 4.2.7). These features were chosen to provide the RF model with the information needed to distinguish between the three quality classes based on structural characteristics. The dataset was randomly split, with 75% used for training and 25% for testing the RF model.

To understand which features most influenced the model's predictions, SHAP (SHapley Additive exPlanations) values were calculated on the test dataset using the iml R package (Molnar *et al.*, 2018; Molnar, 2019). SHAP values show how much each feature contributed to the model's prediction for a segment. They are calculated by checking how the prediction changes when a feature is added to different combinations of the other features. By averaging these effects across all segments it shows which features had the biggest influence on the model's decisions.

After the RF model training, an extra correction step was tested, where segments were interactively re-labeled if the model's predicted probabilities were nearly equal for two classes. The idea was to improve classification reliability by flagging uncertain predictions. However, this step only slightly improved accuracy and was not included in the final workflow.

4.2.5 Novel Segmentation Calibration

The final step of the novel ITS method was the segmentation calibration, an iterative process designed to improve the quality of segmentation based on RF feedback (Figure 5). The calibration process began with a new PTrees segmentation, using a high neighbourhood (k) value, different from the one applied during the RF training dataset preparation. This ensured that the new segmentation contained segments independent from those used for RF model development. Following this initial segmentation, each segment was classified using the RF model. Segments identified as correct (Class 1) were retained without further modification, while segments classified as oversegmented or undersegmented were flagged for adjustment.

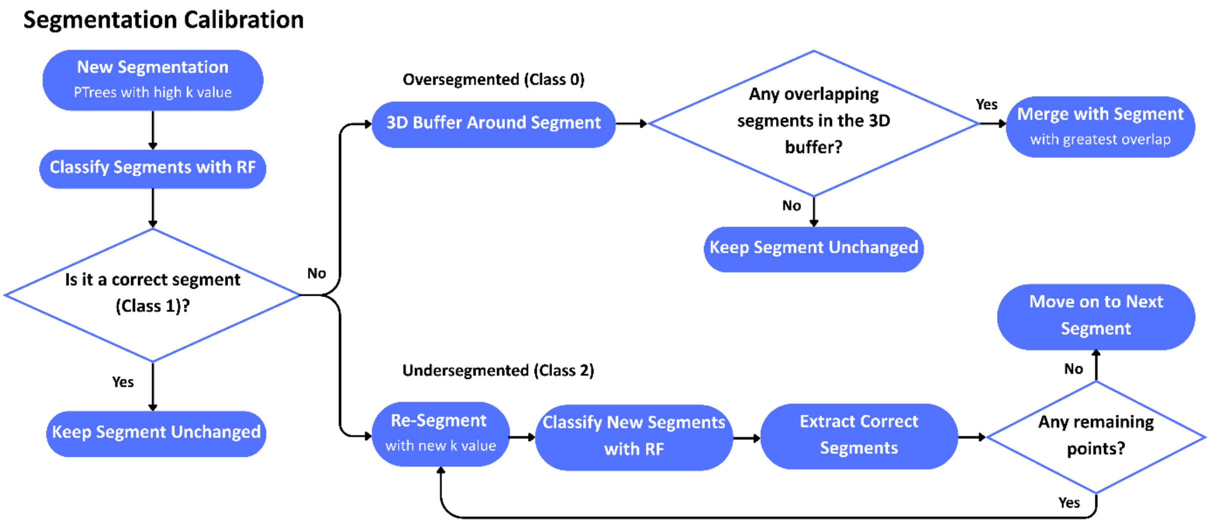


Figure 5: Iterative segmentation calibration process.

First, the oversegmented segments (Class 0) were corrected by spatially merging them with their nearest neighbouring segment. A 3D buffer (tested values: 10 cm, 30 cm, 50 cm) was applied around each of them to identify potential neighbours. If overlapping segments were found, the oversegmented segment was merged with the one showing the greatest overlap. If no overlap was detected, the segment was left unchanged.

Then, the undersegmented segments (Class 2) were re-segmented using different k-values (tested values: 100, 80, 60, 40, 30, 25, 20, 15, 10, 5; Vega *et al.*, 2014). After each iteration, the new segments were evaluated using the RF model. Correct segments were extracted and stored, while the remaining points were reprocessed in the next iteration. This loop continued until no further improvement was achieved.

To evaluate the impact of each processing step, validation metrics (e.g., F1-score, see Section 3.2.7) were calculated after each major stage, including initial segmentation, correction of oversegmented trees, and correction of undersegmented trees. This stepwise evaluation helped assess the contribution of each process to overall segmentation quality. In addition, calibration settings (e.g. 3D buffer size) were selected based on the combination that produced the highest average F1-score across segmentation outputs with different parameter settings.

4.2.6 In Situ Data Matching and Uncertainty Analysis

The last step of the ITS analysis was the validation of the segmentation results. This required matching the segments to the corresponding trees from the in situ dataset and then calculating the validation metrics (see Section 3.2.7) (Eysn *et al.*, 2015). Due to known positional inaccuracies in the field data, a custom matching algorithm was developed to associate each in situ tree with its most likely corresponding segment (Figure 6).

Before matching, two pre-processing steps were applied: (i) all segmented crowns were cleaned by removing outlier points, and (ii) the in situ tree dataset was filtered to include only trees visible in the upper canopy layer, consistent with the LiDAR dataset, which had already been filtered to represent only the upper canopy (see Section 3.1).

The matching algorithm was designed to account for the positional inaccuracies of the in situ data by using expanding search buffers around each in situ tree (Figure 6). Candidate segments within the buffer were evaluated based on three criteria (Eysn *et al.*, 2015):

1. Height similarity: The height of the field tree and segment must fall within an adaptive height difference threshold. Taller trees are allowed a larger height difference, while shorter trees require closer match. This approach accounts for the common tendency of LiDAR to underestimate the height of taller trees more (Mielcarek *et al.*, 2018).
2. Crown area similarity, measured using the area difference ratio (ADR):

$$ADR = \frac{|Area_{segment} - Area_{field}|}{Area_{field}} \qquad (1)$$

   Lower values indicate more similar crown sizes.

3. Crown shape similarity evaluated how well the shapes of the field tree and the segment overlap. First, the segment is shifted so that its center matches the field tree's position. Then the Intersection over Union (IoU) is calculated:

$$IoU = \frac{Area\ of\ Overlap}{Total\ Area\ Covered\ by\ Both\ Crowns} \qquad (2)$$

   Higher IoU means better shape alignment.

Each candidate segment receives a matching score, which combines the height difference, area difference ratio, and inverted IoU (so that lower scores indicate better matches). The best-scoring segment is selected, and one-to-one matching is enforced, meaning each tree and segment can only be matched once.

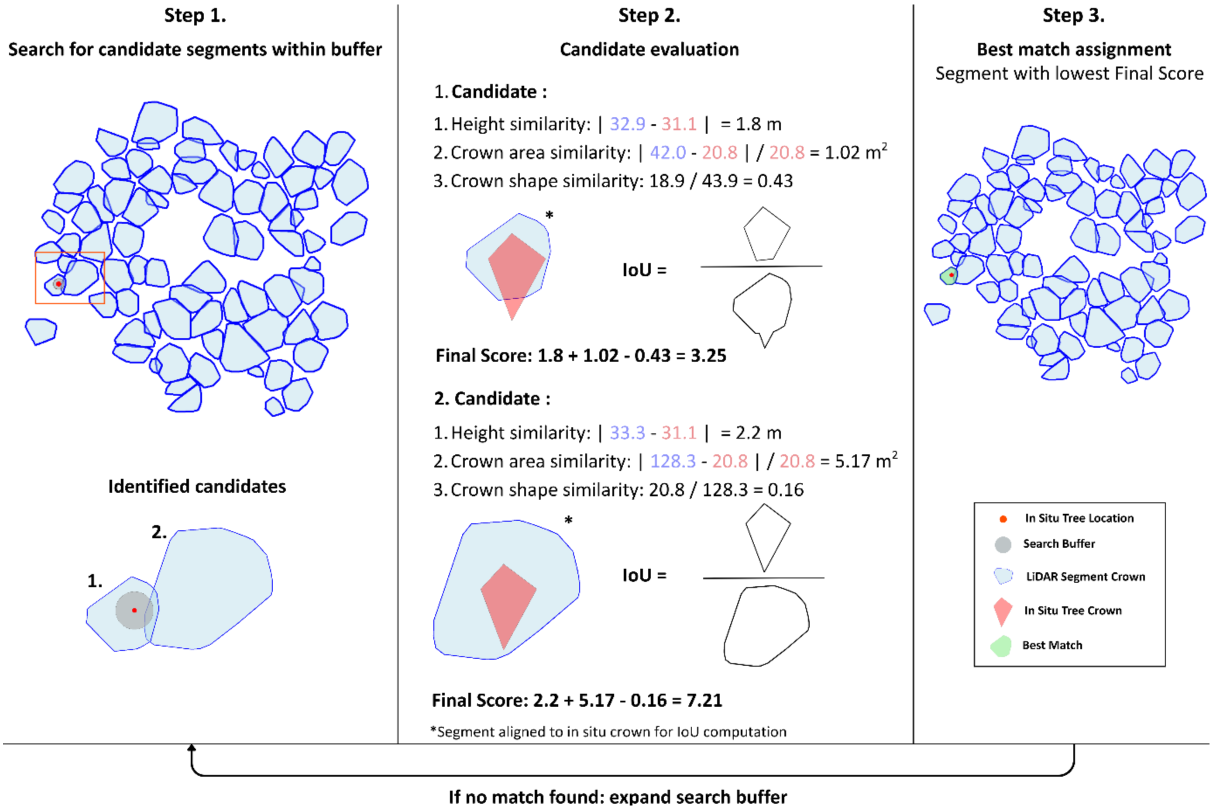


Figure 6: Illustration of the in situ–LiDAR segment matching algorithm. Step-by-step example showing how one in situ tree is evaluated against two candidate LiDAR segments using spatial proximity and height similarity criteria.

Before applying this method to the segmentation datasets, it was calibrated using the best-case segmentation, which represents the most accurate available approximation of true tree positions and crowns. The matching algorithm was calibrated using this dataset in two parcels (*Fagus* and *Acer*), where reliable in situ–segment pairs were known[4]. The goals were to:

- Calibrate the algorithm’s parameters (buffer sizes, height thresholds, scoring rules) for general application across parcels.
- Assess the spatial uncertainty of the in situ data by determining how many trees could be matched at different buffer sizes.
- Validate algorithm robustness by testing it on additional parcels (*Quercus* and *Fraxinus*).

---

[4] These known matches are referred to as ’reference matches’ in the text from now on.

Once calibrated, the matching procedure was applied to all segmentation outputs (benchmark and novel) and was also used later to validate structural metrics (on tree-level) by linking LiDAR-derived values to field measurements (see Section 3.3.4).

4.2.7 Segmentation Validation

Segmentation performance was evaluated by comparing matched segments to field-measured trees using standard detection metrics (Li *et al.*, 2012). A segment matched to an in situ tree was counted as a true positive (TP). A false positive (FP) occurred when a segment could not be matched to any in situ tree, typically caused by oversegmentation or spurious detections. A false negative (FN) occurred when a tree had no corresponding segment, typically due to missed detections or occlusion.

Five commonly applied metrics were used to quantify segmentation accuracy based on the number of TP, FP, and FN (Li *et al.*, 2012). Commission Error (CE) is the proportion of detected segments that do not correspond to real trees (Eq. 3), indicating the extent of oversegmentation or false detections. Omission Error (OE) reflects the proportion of in situ trees missed by the segmentation (Eq. 4). Precision (P) quantifies the fraction of correctly identified segments among all detected segments (Eq. 5), serving as a measure of detection accuracy. Recall (R) indicates the proportion of in situ trees that were successfully matched (Eq. 6), and therefore represents the detection rate. Finally, the F1-Score (F) combines precision and recall into a single metric by taking their harmonic mean (Eq. 7) and provides an overall measure of segmentation quality.

$$CE = \frac{FP}{TP + FP} \qquad (3) \qquad\qquad OE = \frac{FN}{TP + FN} \qquad (4)$$

$$P = \frac{TP}{TP + FP} \qquad (5) \qquad\qquad R = \frac{TP}{TP + FN} \qquad (6)$$

$$F = \frac{2 * P * R}{P + R} \qquad (7)$$

4.3 Structural Metric Estimation

This section describes the calculation and validation of key structural metrics from the segmented trees: tree height, CBH, crown diameter, DBH, and AGB. Metrics were only derived from the novel segmentation results, as an objective of this study was to evaluate the performance of the new segmentation approach. The metrics were grouped into two categories: those directly extracted from the LiDAR point cloud (tree height, CBH, crown diameter) and those estimated using regression models (DBH and AGB).

4.3.1 Direct Metric Calculation

Tree height was calculated by extracting the 95th percentile of point heights within the segment (Sumnall *et al.*, 2016A). This approach was chosen over using the absolute maximum point height, as it provides a more robust estimate that reduces the influence of outlier points or noise above the true crown surface.

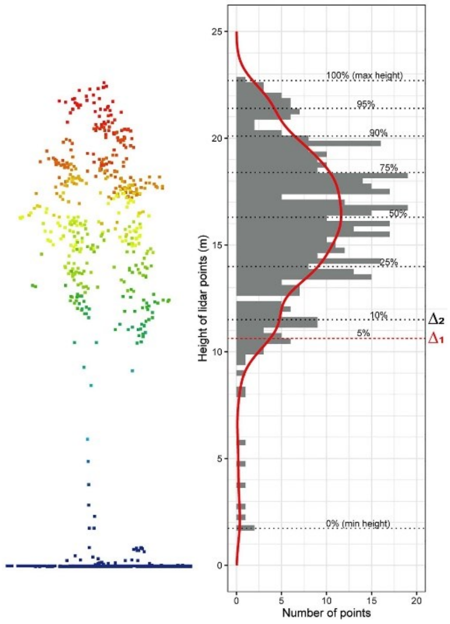


Figure 7: Estimation of individual tree CBH (Source: Chamberlain et al., 2021). By calculating the difference between successive 5th height percentiles within each crown (e.g., Δ1, Δ2, … Δ10) the CBHwas defined as the higher-end percentile corresponding to the largest height difference, indicating a transition from trunk to crown.

CBH was estimated by calculating height quantiles (tested values: 3%, 5%, 10%) for each segment and identifying the largest difference between adjacent values; the upper quantile of this difference was taken as the crown base height (Δ1 on Figure 7) (Solberg *et al.*, 2006; Chamberlain *et al.*, 2021). This method assumes that the start of the tree crown results in increased point cloud density, and the height of this density increase equals the CBH.

Crown diameter was calculated by first removing outlier points from each crown segment based on cluster size and distance thresholds (already mentioned in Section 3.2.6). This step was necessary to eliminate isolated points or small clusters incorrectly assigned to the segment, which could otherwise distort the crown size estimation. After cleaning, the maximum extent was measured along the north-south and east-west axes, and the mean of these two distances was taken as the final crown diameter (Jucker *et al.*, 2017; Yang *et al.*, 2020).

4.3.2 Model-Based Metric Calculation

DBH was estimated using a regression model based on the product of tree height (H) and crown diameter (CD) (Jucker *et al.*, 2017) (Eq. 8). Instead of modelling DBH separately as a function of height or crown size, the two variables were combined into a single predictor. This captures the joint influence of both metrics and helps avoid collinearity problems that arise when height and crown diameter are strongly correlated but entered separately into a model.

The regression model was built using binned in situ trees, grouping them by H × CD values and fitting a log–log linear model to the mean DBH and mean H × CD of each bin. Binning helps reduce the bias caused by having fewer emerging tall trees, and it minimises the influence of random measurement errors. Once established, this model was applied directly to predict DBH for each segment (Jucker *et al.*, 2017):

$$\ln(DBH) = \alpha + \beta \times ln(H \times CD) + \epsilon \qquad (8)$$

Where α nd β are the fitted model parameters, H is the tree height (in m), CD is the crown diameter (in m), and ϵ is the random error term (assumed to be normally distributed).

AGB was calculated for each tree using two approaches: (1) an allometric equation based on DBH and height (Eq. 9) (Jonard *et al.*, 2020), and (2) a regression model based on height (H) and crown diameter (Eq. 10) (CD) (Jucker *et al.*, 2017). Using both methods allowed for a comparison between a traditional approach (1) and one that avoids propagating errors from intermediate DBH estimations (2) (Jucker *et al.*, 2017; Ni-Meister *et al.*, 2022).

The first method estimated AGB using a species-specific allometric equation:

$$AGB = \alpha + \beta \cdot (DBH^2 \cdot H)^\gamma \qquad (9)$$

where DBH is the diameter at breast height (in cm), H is the tree height (in m), and α, β, and γ are species-specific coefficients derived from published literature (Jonard *et al.*, 2020; Guignabert *et al.*, 2024):

- **Beech**: α = 0.056, β = 292.8, γ = 0.966
- **Maple**: α = 5.32, β = 216.9, γ = 1.02
- **Oak:** α = 0.000, β = 263.4, γ = 0.969

Each forest parcel was assigned the appropriate species model: the *Fagus* parcel used the beech parameters, *Acer* used maple, *Quercus* used oak, and *Fraxinus* also used the beech model, since the majority of the trees were beech. This method relies on accurate DBH estimates from the LiDAR dataset (see above).

To reduce dependency on the estimated DBH and avoid error propagation, a second method was used to estimate biomass directly from tree height (H) and crown diameter (CD):

$$\ln(AGB) = \alpha + \beta \times ln(H \times CD) + \epsilon \qquad (10)$$

where H is the tree height (in m), CD is the crown diameter (in m), α and β are model coefficients, and ε is the residual error term.

The same binned regression approach used for DBH estimation was applied here, grouping trees by H × CD and fitting a log–log model to the bin means. This method reduces size-related bias and measurement noise, while providing a more direct AGB estimate that avoids propagating errors from the intermediate DBH calculation.

4.3.3 Regression Model Evaluation for DBH and AGB

To assess the accuracy of the regression models used for DBH and AGB, performance metrics were calculated based on field data (Jucker *et al.*, 2017; Ni-Meister *et al.*, 2022):

- Root Mean Square Error (RMSE): Measures the average prediction error, giving more weight to larger deviations.
- Bias: Indicates whether the model systematically overestimates or underestimates values.
- Coefficient of Determination (R^2): Reflects how much of the variability in field measurements is explained by the model.
- p-values for model coefficients: Indicate whether the predictors (e.g., H × CD) significantly contribute to the model.

In addition, model assumptions were checked to ensure validity: linearity between predictors and response, normality of residuals (via histograms and Q–Q plots), homoscedasticity (constant residual variance), and independence of residuals. The goodness-of-fit was also visually assessed by comparing predicted and observed log-transformed metric values (Jucker *et al.*, 2017).

4.3.4 Validation of Structural Metrics

To evaluate the accuracy and robustness of LiDAR-derived structural metrics, a multi-scale validation approach was applied, consisting of both tree-level and parcel-level assessments.

This two-level approach helps detect issues that might only be visible at one scale (Duncanson *et al.*, 2014).

At the tree level, only segments that were successfully matched one-to-one with field trees were included. This allows for a direct check of how accurate the LiDAR measurements are for individual trees. For each matched pair, LiDAR-derived values were compared against in situ measurements using three standard accuracy indicators (Crespo-Peremarch *et al.*, 2018): RMSE, bias, and R^2. These metrics were calculated for all parcels combined as well as per parcel to assess both overall performance and parcel-specific variation. Results were visualised using scatterplots showing the 1:1 line and fitted regression line (Chamberlain *et al.*, 2021).

At the parcel level, all segmented trees from LiDAR and all field-measured trees are included, without requiring a one-to-one match. Here, the focus is on comparing the overall distributions of the metrics across each forest stand. Two approaches were used:

1. Distribution Comparison: Kernel density plots were used to compare LiDAR- and field-based metric distributions for each parcel and across all parcels. These visualisations highlight differences in shape, spread, and range, which can indicate under- or over-detection of specific structural traits.

- Descriptive Statistics: Summary metrics (mean, standard deviation (SD), minimum, maximum, and quartiles) were computed per source, enabling comparison of central tendencies and variability between LiDAR and field data. This helps assess how well LiDAR captures not just the central tendency, but also the variability and extremes within each stand.

One of the objectives of this study was to examine how LiDAR accuracy and reliability vary under different forest conditions. To address this, the parcels were grouped into two forest types based on structural characteristics:

- **Group 1:** ***Acer*** and ***Fraxinus*** (irregular, mixed species, dense forest parcels)
- **Group 2:** ***Fagus*** and ***Quercus*** (regular, pure, less dense forest parcels)

Validation metrics (RMSE, bias, R^2) and distribution comparisons were repeated per group at both tree- and parcel-levels. This allowed assessment of how detection performance varies under different canopy structures and species mixtures.

5. Results

The structure of the results follows the three main sections of the methodology: (1) Pre-processing, (2) Individual Tree Segmentation, and (3) Structural Metric Results.

5.1 LiDAR Pre-Processing

5.1.1 Ground Classification

Applying Cloth Simulation Filtering (CSF) reduced the proportion of unclassified ground-level points from 12–30% to just 2.1–6.0% across parcels. Classification quality varied slightly by forest parcel. Minor misclassifications in the *Acer* parcel occurred in pits and open areas, however, this did not impact further analysis, since there were no trees in those areas. In *Fagus*, dense leaf litter led to slight elevation overestimates (10–40 cm). *Fraxinus* and *Quercus* showed good results, with most understory correctly not classified as ground (Figure 8).

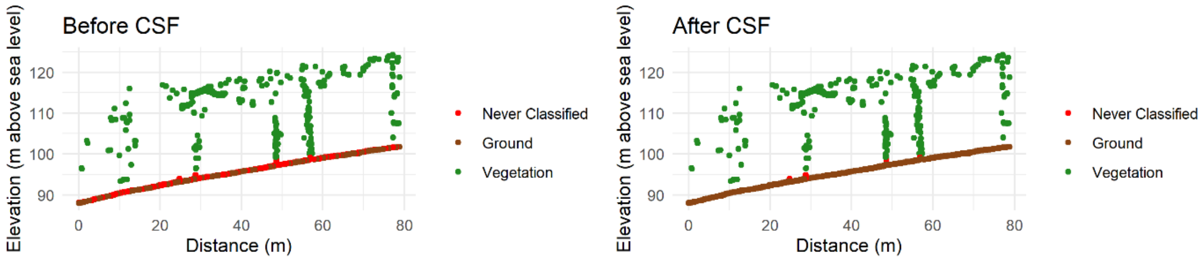


Figure 8: Example of point classification before and after CSF for a sample transect in the Quercus parcel.

5.1.2 Upper Canopy Filtering

To filter upper canopy trees, a local grid-based filtering approach was applied (Figure 9). Each parcel was divided into 5 m grid cells, and the 95th percentile height (P95) was calculated per cell to estimate local canopy tops. Only the top 60% of the height range in each cell was retained, effectively removing understory and lower canopy points, which are poorly captured in the low-density LiDAR dataset.

Cells with P95 values below 20 m were excluded to filter out areas dominated by small or regenerating trees. In a final cleaning step, small and spatially isolated clusters were removed as likely noise or low vegetation. This filtering process resulted in a cleaned, canopy-focused dataset prioritising structurally dominant trees.

The 5 m cell size was selected as a compromise between capturing local structural variation and preserving larger tree crowns that might otherwise be fragmented across multiple cells.

The top 60% threshold was chosen to exclude lower vegetation while retaining the upper canopy crowns, which is important for ensuring that metric calculations capture key crown characteristics.

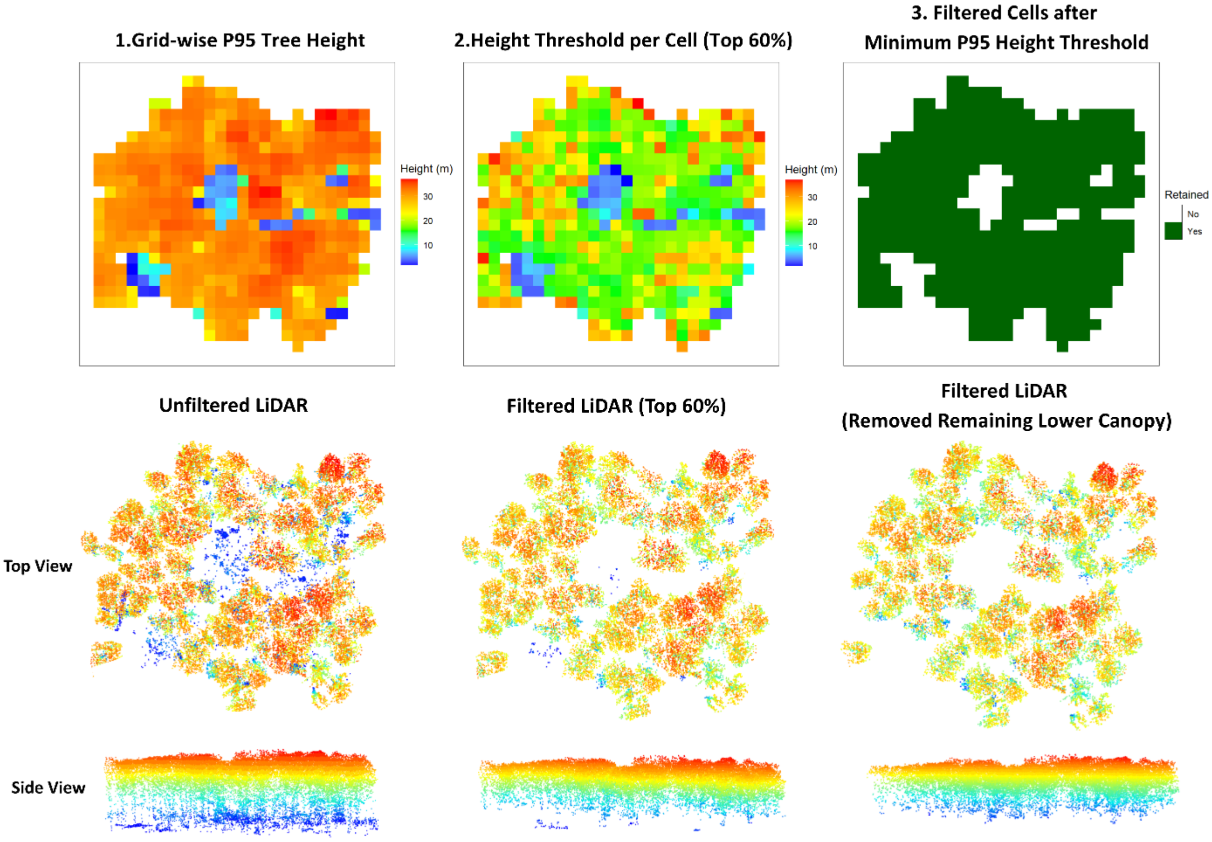


Figure 9: Stepwise filtering of upper canopy points using Fagus Parcel as an example. Top row: grid-based height representation of the parcel at each filtering stage used to guide point removal—initial P95 height per 5 m cell, height thresholds retaining the top 60%, and final mask after excluding low-canopy cells (P95 < 20 m). Bottom row: corresponding top and side views of the LiDAR point cloud at each filtering stage.

5.2 Individual Tree Segmentation (ITS)

Before comparing segmentation method performance, field-measured trees were aligned with their corresponding LiDAR-derived segments using the calibrated matching algorithm, designed to address known coordinate inaccuracies. This section first describes the calibration results based on best-case segmentation data, which also provided an estimate of positional uncertainty in the field data.

Following this, the section then focuses on each method individually: for the benchmark segmentation, the final parameter settings are presented; for the novel segmentation, results include the Random Forest (RF) classification performance, feature importance analysis, and a stepwise evaluation of how each processing stage influenced validation outcomes. Lastly, the performance of the two segmentation methods is compared (benchmark and novel).

5.2.1 Matching Algorithm Parameters

The matching algorithm combined structural and geometric criteria (height difference, crown area difference ratio, and IoU) to identify the best segment for each field tree. Parameter values were calibrated using visually confirmed reference matches in the *Fagus* and *Acer* parcels and then validated on *Fraxinus* and *Quercus* parcels.

Final parameters for the algorithm included a height-dependent threshold that defined the maximum allowed height difference between an in situ tree and a segment: if the in situ tree was taller than 29 m, a difference of up to 17 m was permitted; otherwise, the threshold was limited to 7 m. In addition, the scoring formula was adjusted for in situ trees taller than 40 m to reduce the penalty for high tree height difference. Without this correction, many tall trees would have been excluded from matching, since their height was usually more underestimated by the LiDAR (see Section 4.3.1) and difficult to measure from the ground. Buffer distances were fine-tuned per parcel to reflect varying forest conditions and in situ location accuracy:

- ***Acer***: 1, 2, 4, and 6 m
- ***Fagus***: 1, 2, and 4 m
- ***Fraxinus***: 1, 2, 3, and 5 m
- ***Quercus***: 1, 2, 4, and 5 m

These settings were applied across all segmentation methods to ensure consistent validation. The next section presents the best-case matching results that guided these parameter choices.

5.2.2 Matching Accuracy on Reference Segments

The table below summarises the matching algorithm outcomes using the best-case segmentation (Table 4). The number of in situ trees remaining after filtering (see Section 3.2.6) is especially important, as only these trees are included in the reference dataset for subsequent structural metric validation. While the number of reference matches varied across parcels, the algorithm correctly matched nearly all of them, confirming the effectiveness of the settings.

Overall, matching success was high across parcels. However, *Acer* and *Fraxinus* had a substantial number of unmatched in situ trees (58 and 44). This is likely due to the combination of dense canopy structure and the relatively low point density of the LiDAR dataset, which makes it difficult to detect all trees in such conditions. In addition, the filtering process may have been too permissive on the in situ side, resulting in a mismatch where some field trees remained in the dataset despite not being visible in the LiDAR-derived canopy layer, leading to less corresponding datasets. In contrast, *Quercus* showed many unmatched segments (27), suggesting that the in situ filtering may have excluded trees that were actually visible in the LiDAR data.

Table 4: Summary of matching results for the best-performing segmentation in each parcel. Includes the number of LiDAR-derived segments, in situ trees before and after filtering, known reference matches (available/correct), total matched trees, and counts of unmatched segments and in situ trees.

Parcel	***Acer***	***Fagus***	***Fraxinus***	***Quercus***
Number of Segments	88	78	122	101
In Situ Trees (Before/After Filtering)	253/137	99/80	275/153	177/82
Reference Matches (Available/Correct)	35/35	75/75	37/35	37/37
Total Matched Trees	87	76	109	74
Unmatched Segments	1	2	13	27
Unmatched In Situ Trees	58	4	44	8

In addition to guiding matching algorithm parameter selection, these results provided a basis for assessing in situ tree location uncertainty by analysing the buffer distance needed for successful matching, which is presented in the following section (Figure 10).

5.2.3 In Situ Tree Location Uncertainty

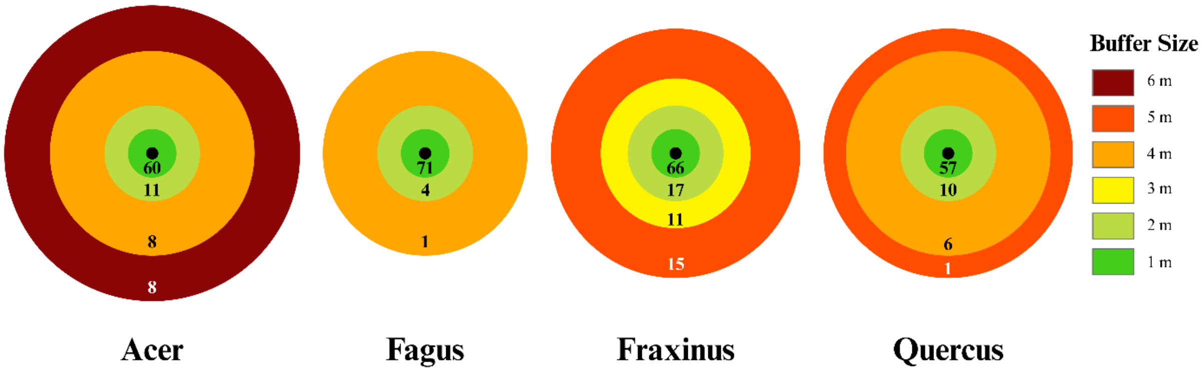


Figure 10: Number of in situ trees matched within each buffer distance for all parcels. Circles are color-coded by buffer size (green = 1 m to red = 6 m), with the number of matched trees indicated inside each circle. This visualises the sensitivity of tree matching to spatial uncertainty.

Out of 346 matched in situ trees, the majority were matched within 1–2 m, consistent with typical Global Navigation Satellite System (GNSS) accuracy under forest canopy (Figure 10). In

Fagus and *Quercus*, over 90% of trees fell within this range, indicating low spatial uncertainty. In contrast, *Acer* and *Fraxinus* required larger buffers (up to 6 m) for a notable portion of matches, likely due to denser canopy conditions and higher positional error. These differences underline the need for parcel-specific buffer calibration to ensure reliable matching.

5.2.4 Benchmark Segmentation Results

The benchmark segmentation (Figure 11) was based on a CHM generated using the pit-free algorithm with 10-meter height thresholds and a 1-meter spatial resolution. A maximum triangle edge length of 1.5 meters was used to balance detail and smoothness.

Treetop detection was optimised using a linear variable window size function, with the final parameters set to w(h) = 0.25 × h + 2, where h is the CHM pixel height. This parameterisation yielded the highest average F1-score across all parcels during validation.

Crown regions were delineated using a standard region-growing approach based on pixel height differences. Although effective in detecting dominant trees, this method tends to miss suppressed or smaller trees due to occlusion and the resolution limits of the CHM. It also does not segment all LiDAR points, as points outside crown polygons are excluded (grey areas on Figure 13).

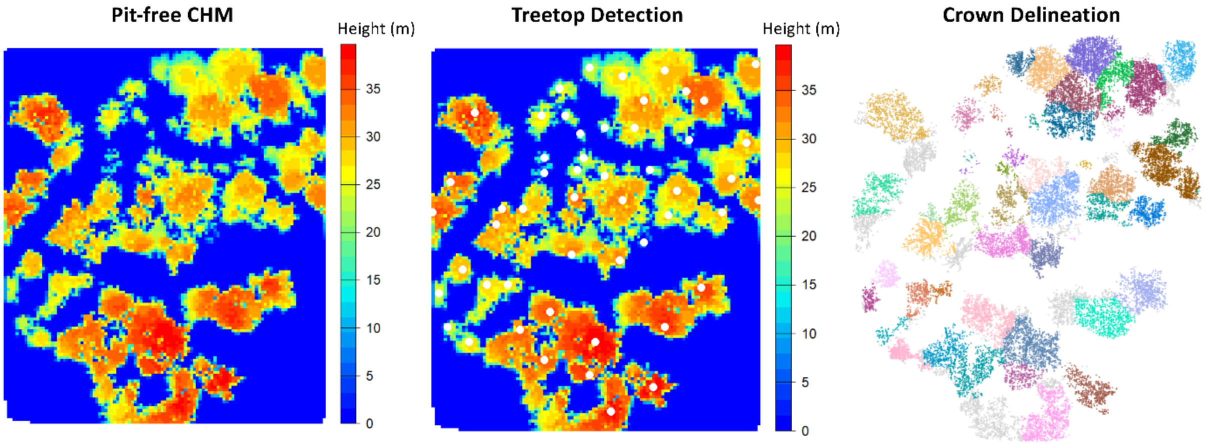


Figure 11: Benchmark segmentation workflow illustrated for the Acer parcel.

5.2.5 Novel Segmentation: Random Forest Classification Accuracy

A total of 467 segments were manually labeled into three quality classes: oversegmented (Class 0), correctly segmented (Class 1), and undersegmented (Class 2). Of these, 352 segments (75%) were used to train the RF model, while 115 segments (25%) were reserved as a test set. The classification results shown below are based on this independent test set and highlight the model's performance across all three classes (Figure 12). The model achieved 0.71 for precision, 0.73 for recall, and 0.72 for F1-score. These results indicate a moderate overall performance, suggesting that the model is reasonably effective at distinguishing between segmentation quality classes.

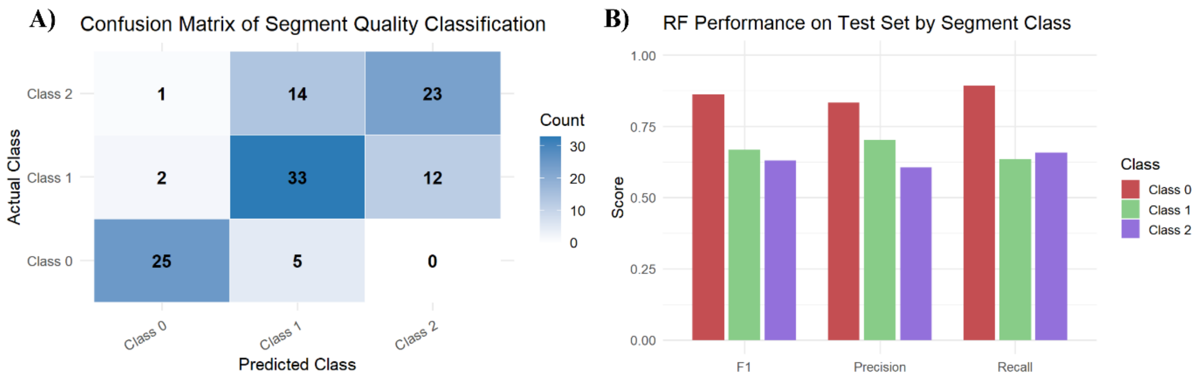


Figure 12: A) Confusion matrix showing the classification results of segment quality categories (Class 0 = oversegmented, Class 1 = correctly segmented, Class 2 = undersegmented) on the independent test dataset. B) Random Forest model performance metrics—F1-score, precision, and recall—broken down by segment class.

Performance was strongest for the oversegmented trees (Class 0), achieving a precision of 0.83 and recall of 0.89 (F1-score = 0.86) (Figure 12B). This suggests that these segments had more distinct point cloud characteristics that made them easier to classify compared to other classes.

Correctly segmented crowns (Class 1) were more challenging to identify, with lower recall (0.64) and precision (0.70), while undersegmented crowns (Class 2) showed the weakest precision (0.61) but slightly better recall (0.66) (Figure 12B), indicating that larger merged crown regions often overlapped with multiple true crowns. The lower performance for these two classes suggests that the model frequently confused correctly segmented and undersegmented trees (Figure 12A). This confusion is likely due to the diverse forest conditions represented in the dataset, where crown sizes and point densities vary considerably. In such cases, distinguishing between a slightly oversized crown and a correctly segmented one becomes difficult

5.2.6 Novel Segmentation: Feature Importance in Segment Classification

The RF model was trained using 15 features grouped into three categories: (1) point cloud height distribution, (2) point cloud density and coverage, and (3) point cloud shape features. The first category captures the vertical structure of the crown segment:

- Mean Height, Median Height, Minimum Height: Basic statistical summaries of point elevations.
- Height SD, Height Interquartile Range (IQR): Vertical variability; useful to detect uneven structures.
- 10th and 90th Height Percentiles: Indicators of vertical crown spread.
- Points Above 10m (%): Helps identify tall or multi-layered crowns.
- Height Gini Coefficient: Captures the inequality in height distribution within a segment. A higher Gini value indicates a greater difference between lower and upper returns, potentially reflecting uneven crown shapes.
- Canopy Relief Ratio: A relative measure of crown height complexity, higher values suggest a peaked or complex crown profile, while lower values indicate flatter crowns.
- Vertical Compactness: Quantifies how tightly LiDAR points are vertically clustered. Segments with many layers or tall, narrow structures typically have lower compactness, while homogeneous or dense crowns appear more vertically compact.

The second category reflects how dense and complete the crown segment is:

- Point Count, Point Density: Total number of LiDAR returns and their concentration.
- Planar Point Density: Represents how dense the point cloud is when projected onto a 2D plane. It helps assess how fully the crown footprint is sampled horizontally, important for distinguishing fragmented or incomplete segments.

The third category describes the 2D geometry of crowns:

- Crown Convex Hull Area: Surface area of the convex shape enclosing the segment; useful to distinguish between small and large crowns or partial oversegments.

The SHAP value summary below shows the relative importance of these features in the RF's predictions, calculated across all test samples (Figure 13). The analysis revealed that Minimum Height and Crown Convex Hull Area were the most influential features across all classes. These features were especially important for identifying oversegmented crowns (Class 0), which are characterised by small, tightly clustered point clouds located near the canopy top. In addition, Height Gini Coefficient, Height SD, and Point Count played significant roles, highlighting that variation in vertical structure and the density of points within a segment are key factors in distinguishing between segmentation classes.

Although some features had lower mean SHAP values individually, attempts to remove them from the model resulted in reduced accuracy, suggesting that feature interactions play a significant role. Therefore, all 15 features were retained to preserve model performance and account for subtle dependencies between structural and shape descriptors.

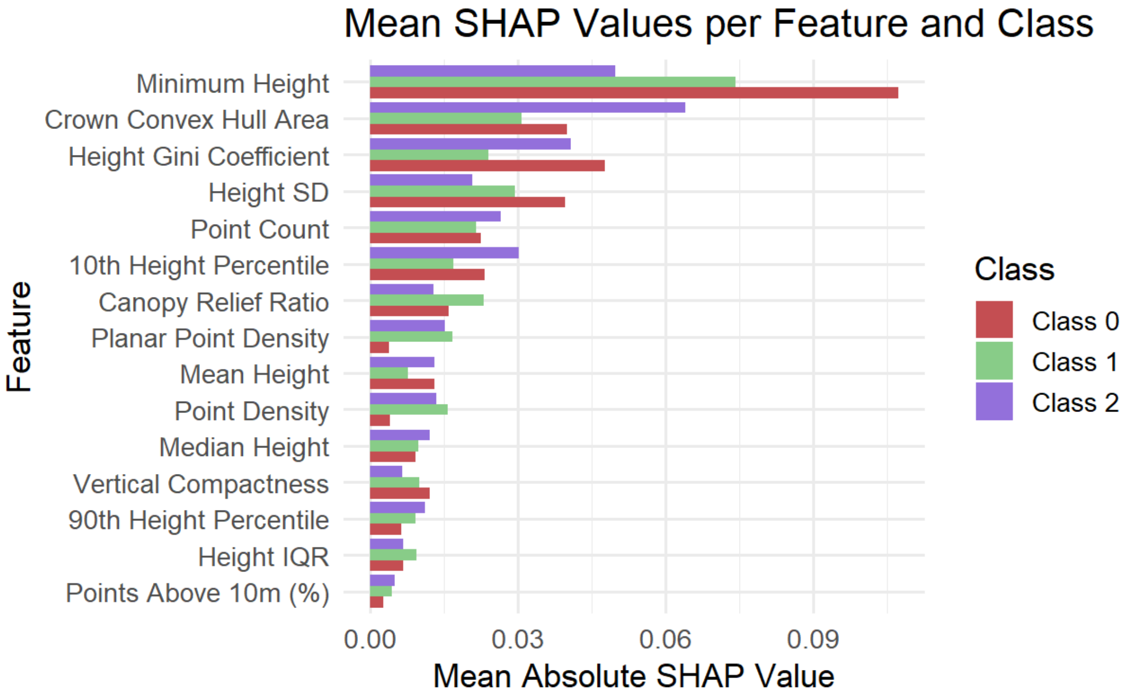


Figure 13: SHAP value summary showing the average contribution of each feature to the RF model predictions by class.

5.2.7 Stepwise Performance Evaluation of the Novel Segmentation Workflow

To evaluate the effectiveness of each step in the novel segmentation, a step-by-step validation was conducted focusing on three stages of the workflow: initial segmentation, oversegmentation correction, and undersegmentation correction (Figure 14).

The initial PTrees segmentation (k = 40 and 80) produced 376 segments, resulting in an F1-score of 0.65. Precision was moderate (P = 0.72), recall was lower (R = 0.60), and OE reached 40%, indicating that many field trees were not detected.

During the oversegmentation correction phase, the RF model identified 101 oversegmented segments, of which 98 were successfully merged with a 30 cm 3D buffer. This step significantly reduced false positives from 105 to 36, increasing precision to 0.87, the highest among all stages. However, this came at the cost of recall (R = 0.54), slightly lowering the number of true positives. Despite this trade-off, the F1-score improved to 0.67, confirming better segment quality.

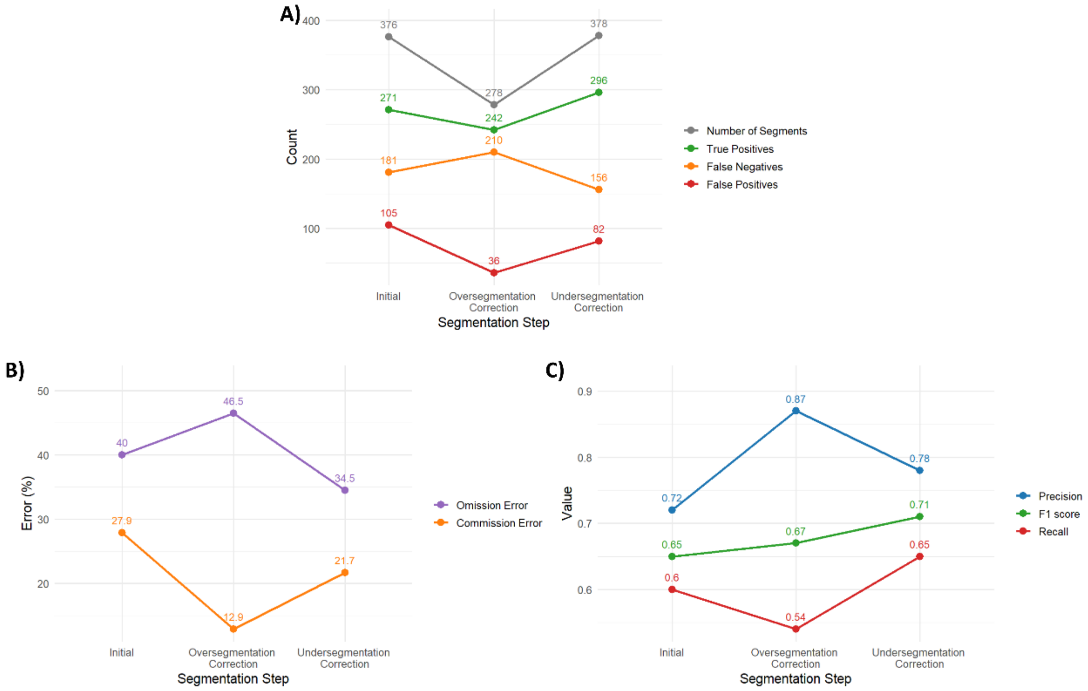


Figure 14: Step-by-step evaluation of the novel individual tree segmentation (ITS) workflow. (A) Raw detection counts per step, showing the number of segments, true positives (TP), false positives (FP), and false negatives (FN), (B) Corresponding commission and omission errors (%) across steps, (C) Accuracy metrics (precision, recall, F1-score) for each major step.

The final undersegmentation correction step identified 67 undersegmented segments and applied adaptive re-segmentation using k values: 100, 80, 60, 40, 30, 20, 5. This process generated 100 new segments, including 54 true positives and 46 false positives, increasing the overall segment count to 378. Importantly, recall improved from 0.54 to 0.65, and OE dropped to 34.5%, pushing the final F1-score to 0.71, the highest across all stages.

In summary, the step-by-step corrections improved different aspects of segmentation: oversegmentation correction refined precision by reducing false detections, while undersegmentation correction enhanced recall by recovering missed trees.

5.2.8 Segmentation Performance Overview

The novel segmentation method consistently outperformed the benchmark across parcels, reaching an F1-score of 0.71, compared to 0.63 for the benchmark, along with lower OE and CE (Table 5). This suggests that the novel approach was more effective at both detecting true trees (fewer missed detections, hence lower OE) and avoiding false or redundant segments (fewer oversegmentations or spurious detections, hence lower CE).

Table 5: Performance comparison between the benchmark and novel segmentation methods combined across all parcels. Metrics include the number of segments, true positives (TP), false negatives (FN), false positives (FP), commission error (CE), omission error (OE), precision (P), recall (R), and F1-score.

Method	**Segments**	**TP**	**FN**	**FP**	**CE (%)**	**OE (%)**	**P**	**R**	**F1-score**
Benchmark	339	251	201	88	26.0	44.5	0.73	0.56	0.63
Novel	378	296	156	82	21.7	34.5	0.78	0.65	0.71

Some clear differences emerged between parcels (Figure 15, Appendix Table A1). *Fagus* showed the best segmentation quality for both methods, with the novel approach achieving an F1-score of 0.87. *Acer* was the most challenging parcel: the benchmark method had a particularly low recall of 0.36, meaning it failed to detect a large proportion of actual trees, resulting in an F1-score of just 0.52. The novel method improved detection performance, raising recall to 0.55 and the F1-score to 0.66. In *Fraxinus*, both methods struggled, with OE exceeding 47%, which reflects a high number of undetected in situ trees, and F1-scores remaining below 0.65. *Quercus* showed a contrasting trend: the benchmark segmentation resulted in excessive oversegmentation, with a CE of 46%, meaning nearly half the detected segments were not real trees. The novel method reduced this to 35%, indicating better precision, while maintaining high recall (both at 0.89), thus successfully detecting most trees without increasing false positives.

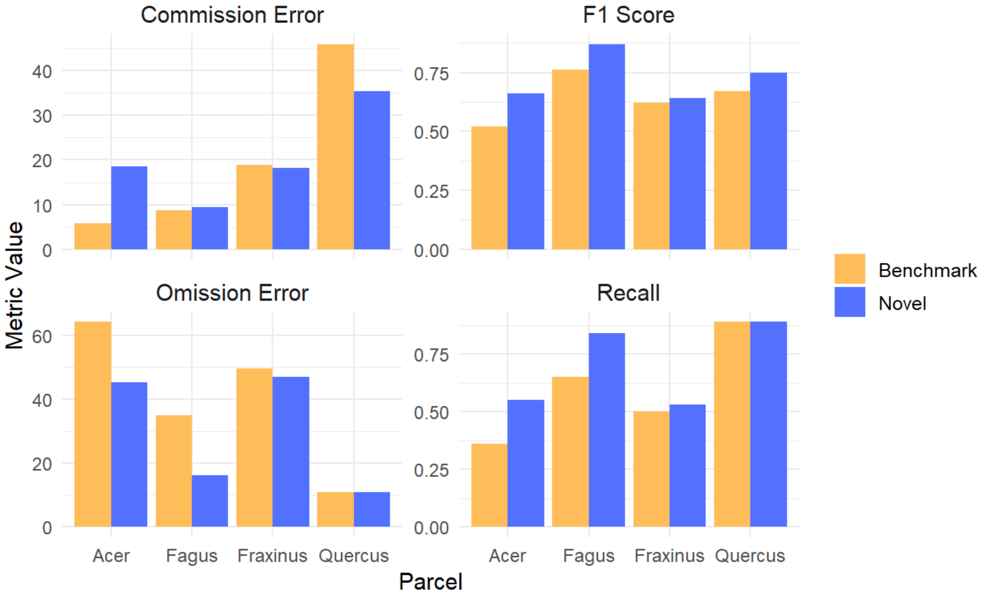


Figure 15: Comparison of segmentation performance metrics (F1-score, Recall, Omission Error, and Commission Error) for the benchmark and novel methods per parcel.

Segmentation performance was clearly influenced by forest structure: parcels in Forest Condition Group 2 (*Fagus* and *Quercus*), characterised by regular, pure, and less dense stands, generally yielded better results (Figure 15). These structural conditions likely facilitated more accurate crown delineation and reduced segmentation errors.

The novel segmentation method consistently improved tree detection by capturing more true trees and reducing oversegmentation in most parcels. Although it still produced some false positives (CE between 9% and 35%), the overall gains in recall and F1-score show it to be a more accurate and reliable alternative to the benchmark. A visual comparison in the *Fagus* parcel highlights these improvements (Figure 16).

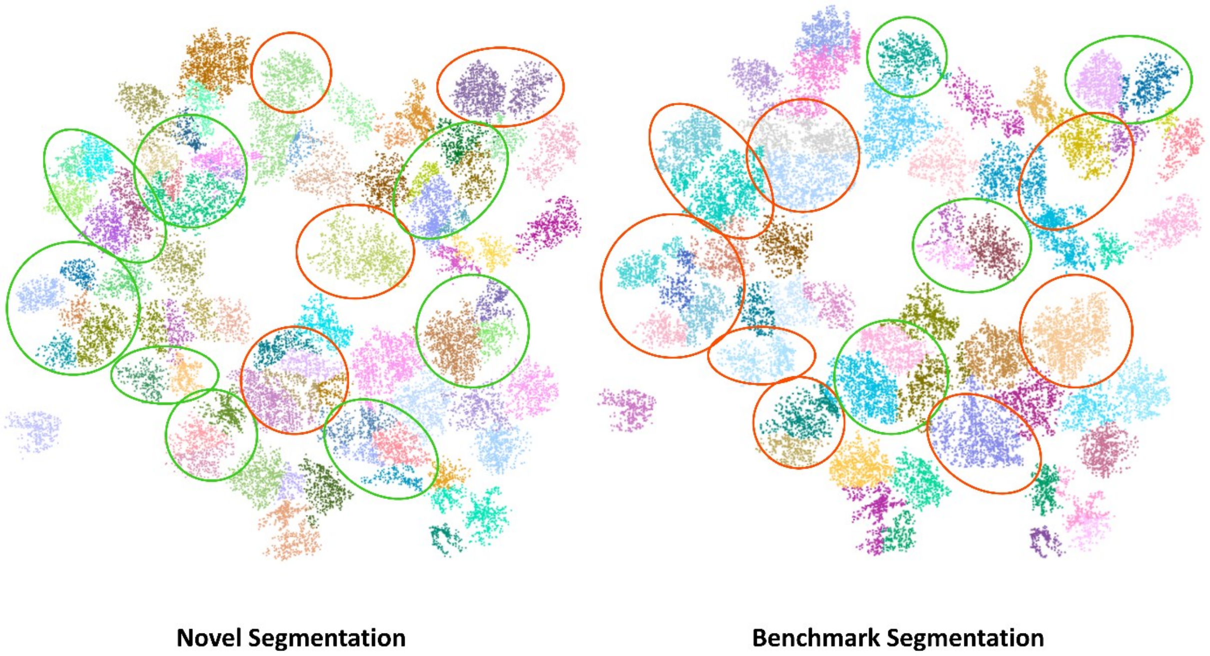


Figure 16: Comparison of novel and benchmark segmentation results on the Fagus parcel. Red circles indicate areas where the given method performed worse than the other, while green circles show areas where it performed better. Unmarked areas represent locations where both methods were either equally accurate or equally incorrect, based on comparison with the best-case segmentation and in situ reference data. Differences in point cloud composition due to outlier removal (see Section 3.2.6) explain why some points appear in only one segmentation. The color of the point cloud indicates the created segments.

5.3 Structural Metric Accuracy

This section presents the validation results of the LiDAR-derived structural metrics. The evaluation is structured metric by metric, and results are shown at different spatial scales and stratified by forest condition. Three main validation outputs are included in the main text:

1. Tree-level validation across all parcels (combined), showing scatterplots and accuracy metrics (RMSE, bias, R^2)
2. Tree-level comparison between forest condition groups, with scatterplots and metrics

3. Parcel-level comparison between forest condition groups, based on density plots and descriptive statistics for LiDAR segments and field trees.

The mentioned forest condition groups are Group 1 (*Acer*, *Fraxinus*): characterised by irregular, mixed-species, and dense stands and Group 2 (*Fagus*, *Quercus*): regular, pure, and less dense parcels. Additional detailed results are included in the Appendix to maintain clarity and focus in the main text. This choice was made because parcel-by-parcel and forest group results were very similar in most cases. These supplementary materials include:

- Tree-level scatterplots and metrics per parcel (Appendix Figure A1-6, A-D)
- Parcel-level density plots (Appendix Figure A1-6, I) and statistics combining all parcels (Appendix Table A2)
- Parcel-level distribution comparisons per parcel (Appendix Figure A1-6, E-H)
- Summary Table of main results from all metrics (Appendix Table A3)

5.3.1 Tree Height

Tree height was calculated as the 95th percentile of the normalised point cloud within each segment. This metric showed relatively strong agreement with field data at the tree-level (RMSE: 4.23 m, Bias: –1.60 m, R^2: 0.58), indicating a general tendency to underestimate tree height, particularly for the tallest individuals (Figure 17).

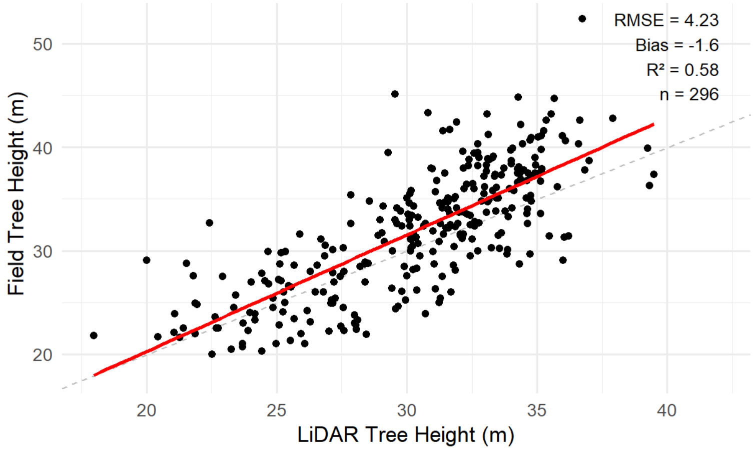


Figure 17: Tree-level comparison between LiDAR-derived and field-measured tree height across all parcels. Red line indicates the best-fit linear regression based on observed values, while the black line represents the 1:1 theoretical relationship. RMSE, bias, R^2, and sample size (n) are reported to summarise model performance.

When stratified by forest condition, tree-level agreement varied between the two groups (Figure 18A-B). In Group 1 (irregular, mixed, dense stands), performance was weaker (RMSE:

4.46 m, Bias: –1.61 m, R²: 0.48), while Group 2 (regular, pure, less dense stands) showed stronger alignment (RMSE: 3.97 m, Bias: –1.59 m, R²: 0.68). These results suggest greater reliability of LiDAR height estimates in more structurally homogeneous forest conditions (like Group 1).

At the parcel level, these differences were also reflected in the distributional patterns (Figure 18C-D). In Group 1, LiDAR-derived height distributions were narrower and more peaked, lacking representation of both smaller and taller trees compared to field data. In Group 2, the LiDAR and field-based distributions were more closely aligned in shape, though the LiDAR estimates still slightly underrepresented the tallest individuals. Descriptive statistics supported these visual patterns. In Group 1, the field data exhibited greater variability, with a wider overall range and higher maximum tree heights (up to 55.9 m) compared to LiDAR (max 39.5 m), emphasising LiDAR's underestimation of tall trees in more structurally complex forests. In Group 2, although both LiDAR and field data had more similar central tendencies, field-based measurements still reached higher maximum values (48.3 m vs. 37.0 m).

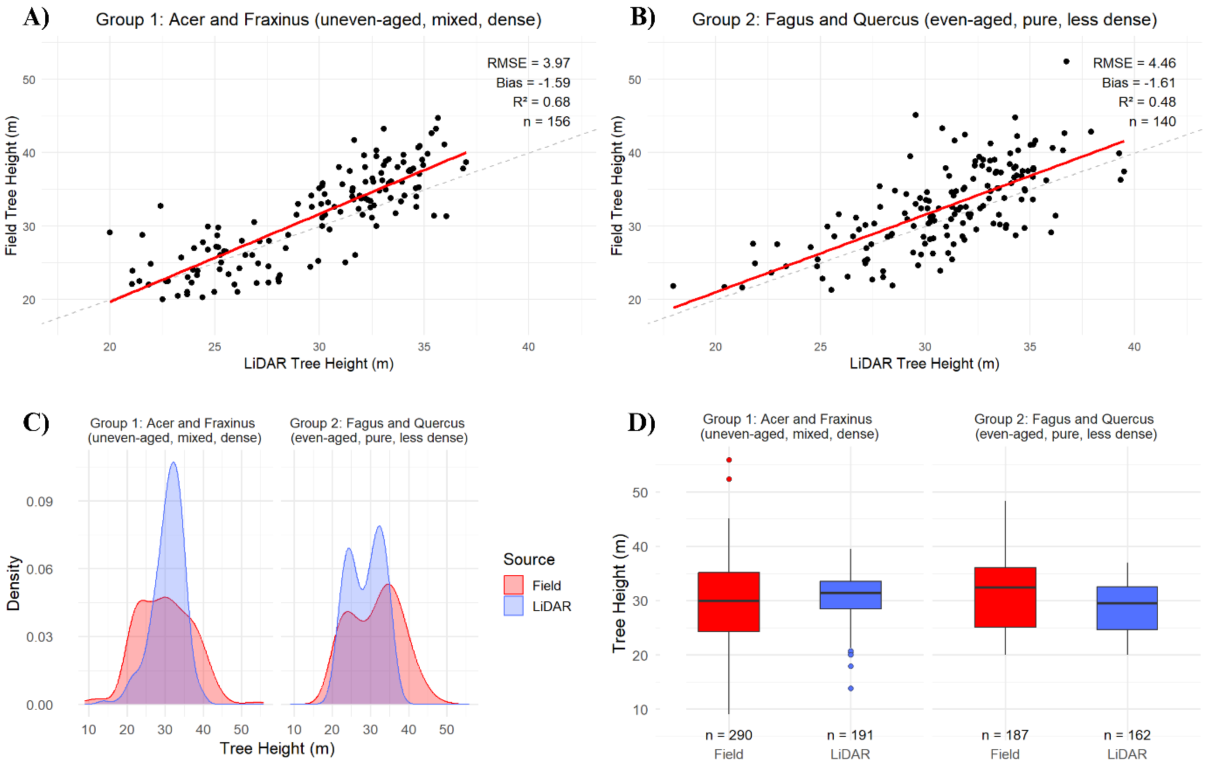


Figure 18: Comparison of LiDAR-derived and field-measured tree height by forest condition group. ***(A–B)*** *Tree-level scatterplots for each group, with the red line showing the best-fit linear regression and the black line indicating the 1:1 theoretical relationship.* ***(C)*** *Density plots comparing the distribution of LiDAR-derived and field-measured values.* ***(D)*** *Boxplots showing the distribution of LiDAR and field values per group.*

These results demonstrate that LiDAR-based height estimates tend to underestimate tall individuals, particularly in complex, multi-layered stands. Performance improves in more uniform stands.

5.3.2 CBH

CBH was estimated using a quantile-based method that identifies the largest vertical gap between adjacent percentiles in the height distribution of each segment. The method assumes that a notable increase in point density marks the beginning of the tree crown. Three quantile separations (3%, 5%, and 10%) were tested, with the 3% interval producing the lowest error.

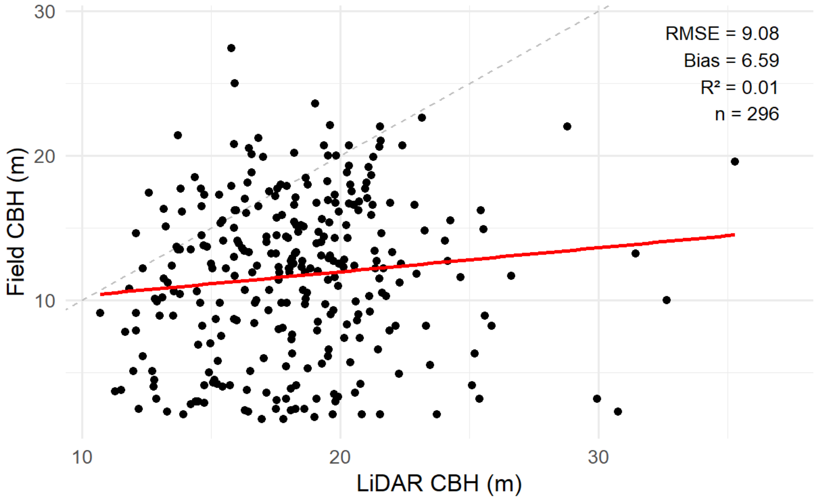


Figure 19: Tree-level comparison between LiDAR-derived and field-measured CBH across all parcels. Red line indicates the best-fit linear regression based on observed values, while the black line represents the 1:1 theoretical relationship. RMSE, bias, R², and sample size (n) are reported to summarise model performance.

At the tree level, overall accuracy was very low (RMSE: 9.08 m, Bias: 6.59 m, R^2: 0.01), with consistent overestimation of CBH compared to field data (Figure 19). Both forest condition groups exhibited similar patterns (Figure 20A-B). In Group 1, CBH estimates had an RMSE of 8.80 m and a positive bias of 6.53 m (R^2: 0.05), while Group 2 showed comparable results (RMSE: 9.32 m, Bias: 6.64 m, R^2: 0.00). No correlation was observed between LiDAR and field values.

At the parcel level, density plots showed that LiDAR-derived CBH values were shifted toward higher values relative to the field data, with lower CBH values largely absent from the LiDAR distribution (Figure 20C-D). While field measurements included minimum values close to 0–1 m, LiDAR estimates did not fall below approximately 10.5–11 m in either group. Boxplots reflected this offset: LiDAR medians were approximately 18 m in both groups, whereas field medians were closer to 11–12 m.

The absence of low CBH values in LiDAR estimates induced by the filtering of the upper 60% of canopy points during the pre-processing stage clearly limits the detection of lower canopy structures. No substantial difference in CBH estimation accuracy was observed between the two groups.

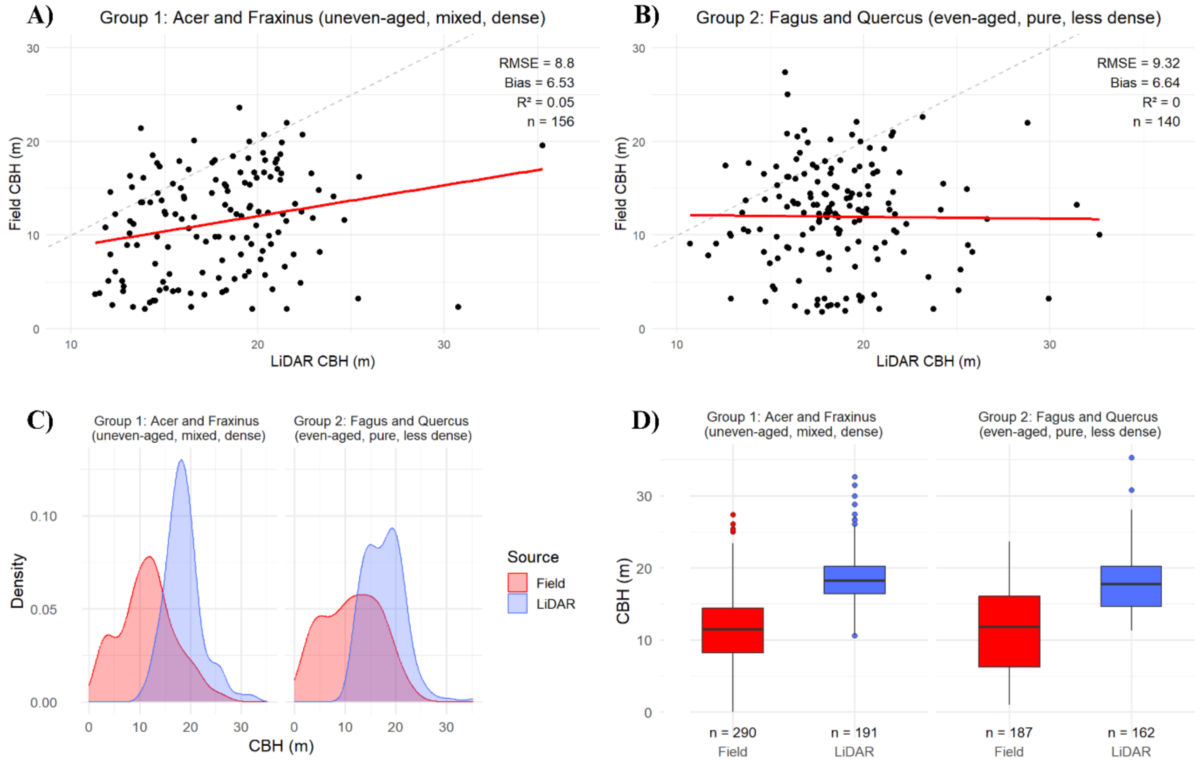


Figure 20: Comparison of LiDAR-derived and field-measured CBH by forest condition group. ***(A–B)*** *Tree-level scatterplots for each group, with the red line showing the best-fit linear regression and the black line indicating the 1:1 theoretical relationship.* ***(C)*** *Density plots comparing the distribution of LiDAR-derived and field-measured values.* ***(D)*** *Boxplots showing the distribution of LiDAR and field values per group.*

5.3.3 Crown Diameter

Crown diameter was estimated by calculating the average extent of each segmented crown in the north–south and east–west directions. At the tree level (Figure 21), the results showed limited accuracy. The overall RMSE was 4.45 m, which is substantial given that the average crown diameter in the field data ranged between 9–10 m. The bias was –0.92 m, indicating a slight underestimation, and the R² was 0.11, reflecting very low explanatory power.

This pattern was consistent across both forest condition groups (Figure 22A-B). In Group 1, the RMSE was 4.67 m with a bias of –0.35 m and R² of 0.03. In Group 2, the RMSE was slightly lower at 4.24 m, but the bias increased to –1.44 m, and R² improved modestly to 0.20. While

Group 2 showed stronger correlation (higher R^2), the systematic underestimation was more pronounced (higher bias). The low accuracy at the tree level could be influenced by segmentation errors.

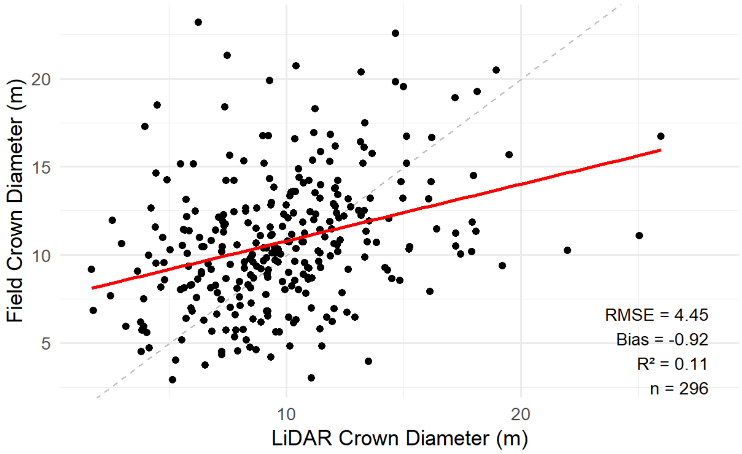


Figure 21: Tree-level comparison between LiDAR-derived and field-measured crown diameter across all parcels. Red line indicates the best-fit linear regression based on observed values, while the black line represents the 1:1 theoretical relationship. RMSE, bias, R², and sample size (n) are reported to summarise model performance.

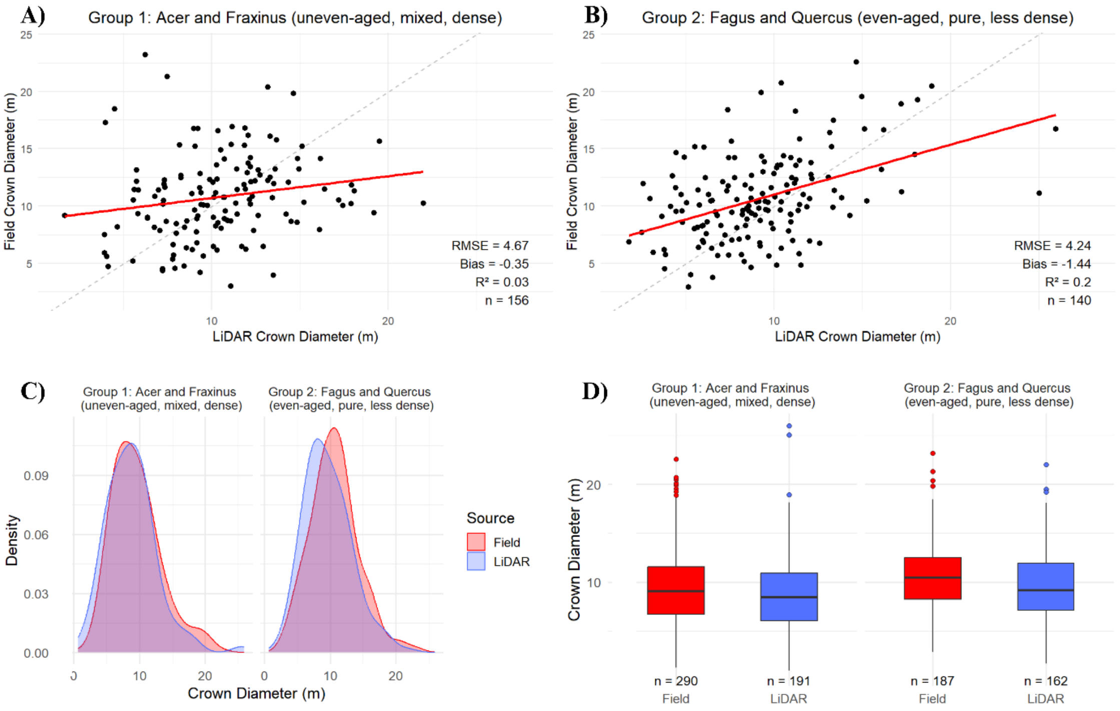


Figure 22: Comparison of LiDAR-derived and field-measured crown diameter by forest condition group. ***(A–B)*** *Tree-level scatterplots for each group, with the red line showing the best-fit linear regression and the black line indicating the 1:1 theoretical relationship.* ***(C)*** *Density plots comparing the distribution of LiDAR-derived and field-measured values.* ***(D)*** *Boxplots showing the distribution of LiDAR and field values per group.*

In contrast, results at the parcel level (Figure 22C-D), showed much stronger agreement between LiDAR-derived and field crown diameters. Density plots revealed similar distributions, with only a slight left shift in LiDAR estimates, consistent with the observed negative bias.

Boxplots supported these observations: in Group 1, LiDAR estimates had a median of 9.22 m (IQR = 4.79 m), compared to 10.51 m (IQR = 4.23 m) in the field data. In Group 2, LiDAR-derived crowns had a median of 8.51 m (IQR = 4.85 m), closely matching field measurements with a median of 9.11 m (IQR = 4.82 m).

The strong agreement between LiDAR and field data at the stand level, despite poor accuracy at the individual tree level, suggests that the method captured overall crown size patterns well but struggled to correctly identify or outline individual tree crowns. This discrepancy may be due to systematic errors in the matching process used to link LiDAR segments to field trees. If that is the case, it could affect the interpretation of all other tree-level metrics, since accurate matching is essential for reliable individual comparisons.

5.3.4 DBH

DBH was estimated using a log–log linear regression model with the predictor variable combining tree height and crown diameter. To reduce bias from the uneven distribution of tree sizes, the model was fitted using binned field data. The resulting model was:

$$\ln(DBH) = -0.658 + 0.782 \times ln(H \times CD) \qquad (11)$$

Both coefficients were significant ($p < 0.0001$), and the model showed good fit on the log-transformed scale (Appendix Figure A7), with an R^2 of 0.97 and RMSE of 0.07. Bias was negligible both in the log and original scales (nearly 0), and diagnostic plots confirmed that assumptions of linearity, normality, and homoscedasticity were met (Appendix Figure A7).

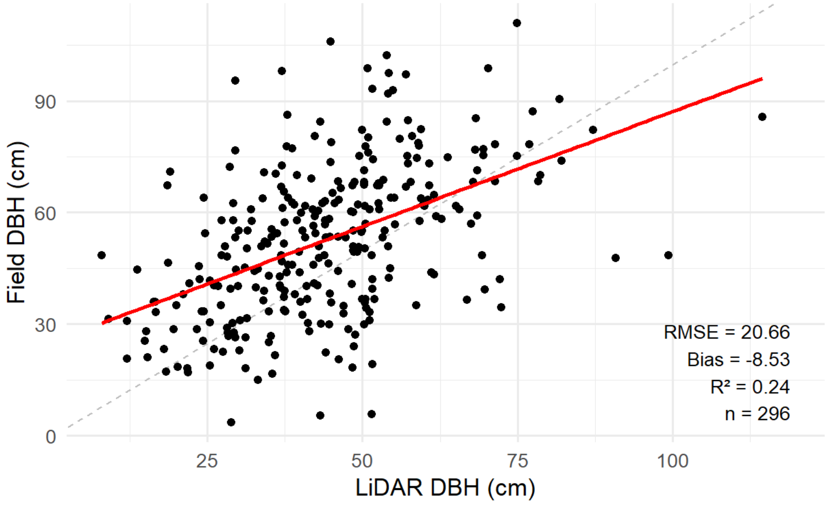


Figure 23: Tree-level comparison between LiDAR-derived and field-measured DBH across all parcels. Red line indicates the best-fit linear regression based on observed values, while the black line represents the 1:1 theoretical relationship. RMSE, bias, R^2, and sample size (n) are reported to summarise model performance.

This regression model was then applied to the LiDAR-derived height and crown diameter values. The resulting tree-level DBH predictions showed moderate accuracy, with an RMSE of 20.66 cm and a bias of –8.53 cm, indicating systematic underestimation (Figure 23). The R² was 0.24, suggesting that only a small portion of the variability in observed DBH could be explained at the individual tree level.

Model performance varied between forest groups (Figure 24A-B). In Group 1, the tree-level RMSE was 22.1 cm with a bias of –11.96 cm and an R² of 0.19. In Group 2, performance improved slightly, with an RMSE of 19.28 cm, bias of –5.45 cm, and an R² of 0.27.

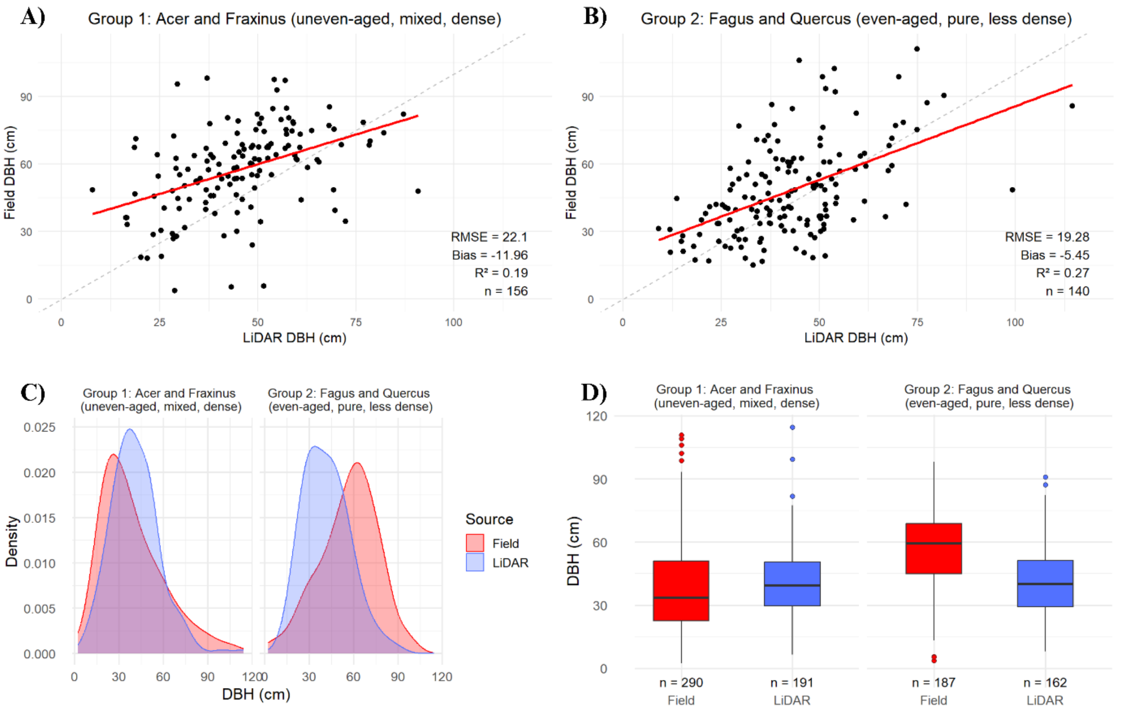


Figure 24: Comparison of LiDAR-derived and field-measured DBH by forest condition group. ***(A–B)*** *Tree-level scatterplots for each group, with the red line showing the best-fit linear regression and the black line indicating the 1:1 theoretical relationship.* ***(C)*** *Density plots comparing the distribution of LiDAR-derived and field-measured values.* ***(D)*** *Boxplots showing the distribution of LiDAR and field values per group.*

At the parcel level, the comparison between LiDAR-derived and field DBH values showed a different pattern (Figure 24C-D). In Group 2, LiDAR-based DBH remained consistently lower than field measurements, with a mean of 41.7 cm compared to 56.7 cm. The interquartile range (29.3–51.2 cm) overlapped only with the lower portion of the field distribution (44.8–68.7 cm), indicating stronger underestimation for larger trees. In contrast, in Group 1, the mean DBH from LiDAR (40.9 cm) was very close to the field mean (39.2 cm). However, the model overestimated DBH for smaller trees and underestimated it for larger ones, reducing the spread of predicted values. This resulted in a narrower distribution in the LiDAR estimates compared to the field data.

The discrepancy between tree- and parcel-level accuracy likely results from segmentation and matching errors, not model performance. While Group 2 showed better tree-level metrics, Group 1 aligned more closely at the parcel level, suggesting individual tree mismatches affected tree-level accuracy.

5.3.5 AGB: Allometric Equation and Regression Model

AGB was estimated using two methods: (1) a species-specific allometric equation based on DBH and height (Jonard *et al.*, 2020), and (2) a regression model using tree height and crown diameter (Jucker *et al.*, 2017). The goal of applying both methods was to compare a conventional approach dependent on DBH (1) with an alternative that avoids error propagation from intermediate DBH estimates (2).

The regression model was calibrated using binned log–log linear regression, with height (H) and crown diameter (CD) as the predictors. The resulting model was:

$$\ln(AGB) = -3.693 + 1.845 \times ln(H \times CD) \qquad (12)$$

The model explained 97% of the variation in log-transformed AGB ($R^2 = 0.97$, RMSE = 0.17) with negligible bias and no major assumption violations, confirming strong, unbiased predictive performance (Appendix Figure A7).

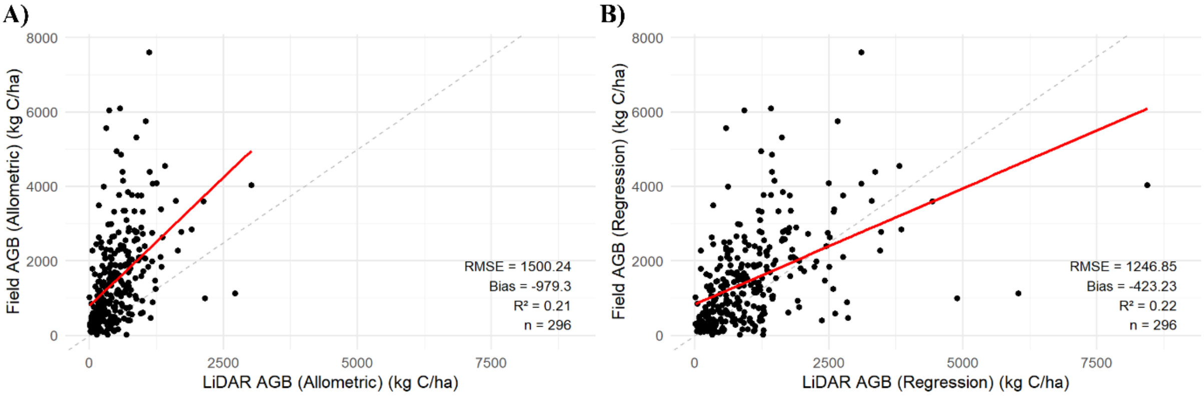


Figure 25: Tree-level comparison between LiDAR-derived and field-measured AGB across all parcels: A) calculated from allometric equation, B) calculated from regression model. Red line indicates the best-fit linear regression based on observed values, while the black line represents the 1:1 theoretical relationship. RMSE, bias, R^2, and sample size (n) are reported to summarise model performance.

When comparing the two methods on tree-level, the regression approach outperformed the allometric one (Figure 25). The regression model achieved a lower RMSE (1246.86 vs. 1500.24

kg C/ha), a smaller bias (-423.23 vs. -979.3 kg C/ha), and a slightly higher R^2 (0.22 vs. 0.21), suggesting it provided more accurate and consistent AGB estimates (Figure 25B).

At the group level, performance differences became more apparent (Figure 26). The allometric method performed slightly better in Group 1 than Group 2 ($R^2 = 0.24$ vs. 0.18), but still exhibited substantial underestimation and high bias in both (Figure 26A-B). The regression method again performed better overall, especially in Group 2 ($R^2 = 0.24$ vs. 0.19), showing reduced bias (-286.43 kg C/ha) compared to Group 1 (-575.68 kg C/ha) (Figure 26C-D). These results suggest that more homogeneous stands (Group 2) allowed the regression model to perform more reliably.

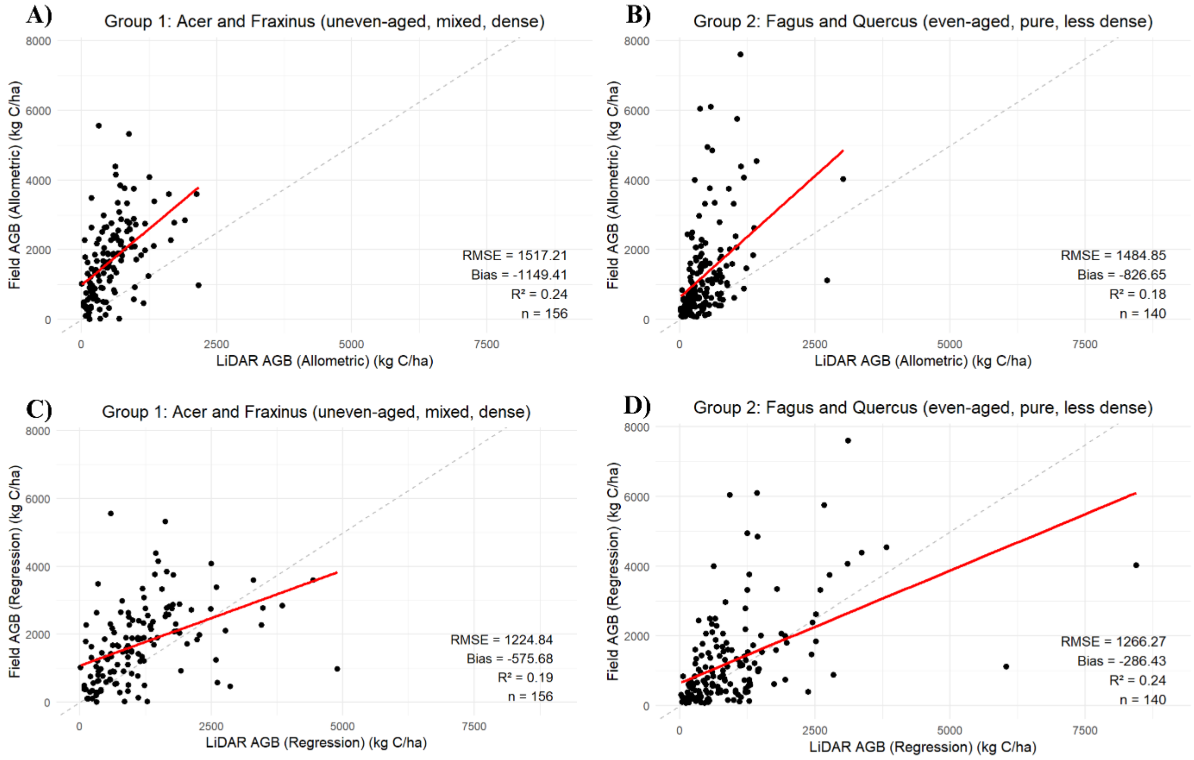


Figure 26: Comparison of LiDAR-derived and field-measured AGB by forest condition group. Tree-level scatterplots for forest condtion groups, with the red line showing the best-fit linear regression and the black line indicating the 1:1 theoretical relationship: ***(A–B)*** *for allometric equation calculated AGB,* ***(C-D)*** *for regression modelled ABG.*

At the parcel level, both methods consistently underestimated AGB compared to field measurements (Figure 27). Density plots revealed that the allometric approach missed high-AGB trees entirely, showing a compressed distribution skewed toward lower values (Figure 27A). The regression model improved this slightly, particularly in Group 1, where its distribution more closely followed the field data, though underestimation of peak values remained a limitation (Figure 27B). Boxplots confirmed this trend: interquartile ranges (IQR) for both LiDAR-based methods were narrower than those of the field data (Figure 27C-D), but the regression approach showed wider IQRs and a closer match to observed central tendencies (Figure 27D).

Descriptive statistics reinforce these patterns. For Group 1, the mean AGB from LiDAR was 448.07 kg C/ha (allometric) and 940.44 kg C/ha (regression), compared to 929.41 kg C/ha from the field. In the structurally less complex parcels (Group 2), the allometric method produced a mean of 475.43 kg C/ha, while the regression method improved this to 958.77 kg C/ha, still underestimating compared to the field mean of 1691.87 kg C/ha. Overall, while both methods struggled to fully capture the highest biomass values, the regression model outperformed the traditional DBH-based approach, especially in Group 1.

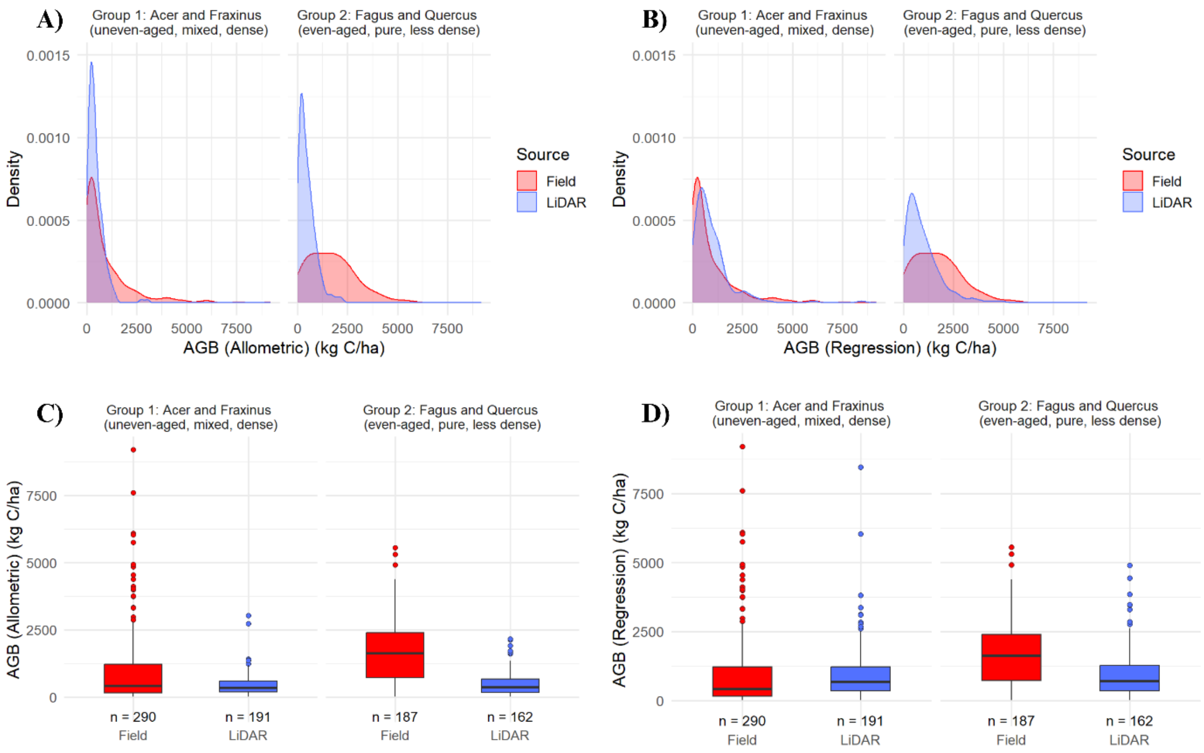


Figure 27: Comparison of LiDAR-derived and field-measured AGB (allometric equation/regression modelled) by forest condition group. ***(A–B)*** *Density plots comparing the distribution of LiDAR-derived and field-measured values.* ***(C-D)*** *Boxplots showing the distribution of LiDAR and field values per group.*

6. Discussion

This discussion reflects on how well LiDAR data captures individual tree-level structural metrics in temperate deciduous forests, addressing the study's main research question and objectives. It begins with an evaluation of the novel segmentation method developed, followed by an assessment of the accuracy of the extracted metrics and factors influencing their performance. The influence of forest conditions on metric accuracy is then explored, and the discussion concludes with broader implications, transferability, and future prospects.

6.1 Evaluation of Novel ITS Approach

This study introduced a novel ITS approach that builds upon an existing point cloud-based clustering algorithm (PTrees) (Vega *et al.*, 2014) and incorporates an iterative correction procedure guided by a machine learning model. The RF model classifies the created segments based on their quality to identify missegmented crowns (specifically, oversegmented and undersegmented ones), then attempts to correct them. Oversegmented segments are merged with their closest neighbouring crown based on 3D buffer overlap, while undersegmented segments are reprocessed using alternative clustering parameters that alter the local neighbourhood size.

The advantage of this approach lies in its balance between accuracy and feasibility. Most existing machine learning applications in ITS are either applied to CHM-based 2D crown delineation (Chang *et al.*, 2022; Henrich *et al.*, 2024), which inherit the limitations of CHM methods, or require large amounts of labeled data and high-density point clouds to support deep learning architectures (Wilkes *et al.*, 2023; Henrich *et al.*, 2024). In contrast, the proposed method can function with lower-density ALS data, making it suitable for large-scale applications where such data is more accessible. In addition, it retains the strengths of point cloud-based methods, such as the ability to resolve overtopped trees and separate overlapping crowns (Vega *et al.*, 2014; Hamraz *et al.*, 2016).

The performance of the RF model was key to the success of this iterative correction process. The model achieved an overall F1-score of 0.72, with strong performance in detecting oversegmented crowns (precision = 0.83, recall = 0.89). This is likely due to their distinctive features: small segment size, high vertical position, and limited spatial extent, which were captured by predictors such as minimum height and crown convex hull area. Correct and undersegmented crowns were more difficult to distinguish, likely due to the diverse forest composition and varying tree heights in the study area. These classes showed moderate and similar performance, often being confused with each other.

The RF model relied on 15 features describing segment geometry, height distribution, and point density. While some features were similar, they each captured slightly different structural aspects of the point cloud. SHAP analysis showed that none of the features could be excluded without reducing classification performance, suggesting that their combined contribution was essential. Still, future iterations of this approach could benefit from incorporating additional or alternative features that describe less explored aspects of crown structure to further refine classification accuracy.

Each step of the novel segmentation workflow contributed to the final result. The initial PTrees segmentation yielded an F1-score of 0.65 but suffered from high OE (40%), indicating a substantial number of missed trees. The oversegmentation correction step, informed by the RF model and 3D buffer merging, greatly improved precision by reducing false positives from 105 to 36, raising the F1-score to 0.67. The final stage (addressing undersegmentation) applied adaptive re-segmentation, successfully recovering many previously missed trees and reducing OE to 34.5%. This improved recall and increased the overall F1-score to 0.71, confirming that the stepwise corrections had a cumulative positive effect.

When compared to the benchmark segmentation, the novel method clearly outperformed across all key metrics: precision increased from 0.73 to 0.78, recall from 0.56 to 0.65, and F1-score from 0.63 to 0.71. Both OE and CE were reduced, underscoring the effectiveness of the iterative correction mechanism.

A parcel-wise analysis revealed that segmentation accuracy varied with forest structure in agreement with previous research (Dalponte and Coomes, 2016). The best performance was observed in regular, pure stands with lower density, such as *Fagus* (F1 = 0.87) and *Quercus* (F1 = 0.75). These conditions likely facilitated better crown separation, uniform canopy structures, and fewer overlapping layers (Yao *et al.*, 2012). In contrast, denser, mixed, and irregular parcels like *Acer* and *Fraxinus* showed lower recall (0.55 and 0.53, respectively), pointing to challenges in detecting trees in complex forest layers. While *Quercus* showed high recall (0.89), it also had elevated CE (35.4%), likely due to local oversegmentation in denser sections and the presence of oversegmented pine trees, which had higher point densities than surrounding broadleaf species that likely caused the segmentation error.

Beyond forest characteristics, segmentation accuracy was influenced by several methodological factors. First, RF classification errors, particularly between correct and undersegmented classes, could lead to inappropriate corrections. Second, the decision to focus only on upper canopy trees required filtering both the LiDAR and field data accordingly. If these filters were mismatched, the validation process could be skewed by either increasing false negatives or false positives. In addition, the location uncertainty in the in situ data required a custom matching algorithm, which, while validated to some extent, was primarily tested on visually obvious matches and might not generalise perfectly. Lastly, the use of moderate-density ALS data remains a limiting factor, as finer forest structures may not be fully resolved at this resolution (Henrich *et al.*, 2024).

Despite these limitations, the performance of the novel method is encouraging when placed in context with other ITS research. Ayrey *et al.* (2015), using a layer-stacking approach on denser point clouds, reported an F1-score of 0.81 in dense mixed woods, while Vega *et al.* (2014) and Hamraz *et al.* (2016) achieved F1-scores of 0.80–0.82 in similarly dense forests with higher point densities (11–25 pts/m^2). Given that the current method operated on lower-density ALS data and achieved an F1-score of 0.71, its performance is competitive.

Future improvements could focus on expanding and diversifying the RF training dataset to improve classification robustness across forest types. In addition, a supplementary segmentation step using alternative methods could be tested on the remaining undersegmented crowns. A key direction would be testing this method across broader regions and more varied forest conditions, including coniferous stands and multi-layered forests with significant understory.

The need for a trained RF model remains a limitation, as transferability may depend on forest structure and point cloud characteristics. Nevertheless, this study presents a flexible and scalable framework for ITS using ALS data that bridges the gap between high-performance but data-intensive methods and more accessible operational approaches. Based on the obtained ITS results, individual structural metrics were extracted and compared to field data, as discussed in the following section.

6.2 Accuracy of LiDAR-Derived Forest Metrics

This study aimed to assess how accurately key forest structural metrics can be derived from airborne LiDAR data. Overall, the results show a mixed but encouraging picture. Tree height was the best-performing metric, with a moderate R^2 (0.58), relatively low RMSE (4.23 m), and a small bias, although tall trees tended to be underestimated. Crown diameter also showed promising results, particularly at the parcel level, where LiDAR-derived distributions closely matched those from field data despite a low R^2 (0.11). These two metrics form the foundation for the indirect estimation of DBH and AGB, and their relatively strong performance highlights the potential of LiDAR to support accurate structural modelling. In fact, the allometric and regression models developed using field-based relationships between crown diameter and height achieved good internal fit ($R^2 > 0.97$), suggesting that if those input variables are well captured from LiDAR, DBH and AGB can also be reliably estimated.

While LiDAR-derived DBH showed moderate tree-level agreement ($R^2 = 0.24$) and some underestimation, its distributional alignment at the parcel level was encouraging. AGB estimates, calculated both through allometric models and regression, showed fair accuracy ($R^2 \approx 0.21$–0.22), with consistent slight underestimations. The only metric with clearly poor performance was canopy base height (CBH), which exhibited large overestimations and weak correlation with field values ($R^2 = 0.01$).

One important pattern across metrics was the generally low R^2 values, despite some metrics demonstrating reasonable RMSE or strong agreement in distributions. This suggests that while LiDAR-derived metrics can capture general trends and overall structural patterns, they often fail to accurately explain tree-level variation. A likely contributor to this discrepancy is the tree matching algorithm used to link field-measured and LiDAR-derived trees (Maltamo *et al.*, 2011; Luo *et al.*, 2018). Particularly for metrics like crown diameter, strong agreement at the parcel level contrasted with weaker tree-level performance, suggesting mismatches between corresponding individuals in the two datasets. Such mismatches would reduce apparent accuracy at the individual level without necessarily reflecting poor detection (Yang *et al.*, 2020).

Beyond matching uncertainties, several general factors influenced metric accuracy. First, the point density of the LiDAR dataset was relatively low, especially in the lower canopy, reducing precision for metrics requiring detailed vertical information (e.g., CBH) (Sumnall *et al.*, 2016B;

Stefanidou *et al.*, 2020). Second, forest structure and condition played a role, with variation in tree density, species composition, and crown complexity affecting detectability (Hyde *et al.*, 2005; Maltamo *et al.*, 2011; Mielcarek *et al.*, 2018). Third, segmentation errors (Luo *et al.*, 2018), such as inaccurate crown delineation or missed trees, also impacted the accuracy.

The following section provides a metric-by-metric discussion, exploring each metric's performance in more detail, identifying specific limitations, comparing the results with previous research, and offering suggestions for improving estimation where applicable.

6.2.1 Tree Height

Tree height was the most accurately estimated structural metric in this study, with a moderate R^2 (0.58), an RMSE of 4.23 m, and a consistent underestimation, particularly for taller trees. This is evident from the fact that LiDAR-derived tree heights did not exceed 39.5 m, while field measurements reached up to 55.9 m, suggesting that the tallest trees were often not fully captured.

Underestimation of tree height is a well-documented issue in LiDAR studies (Næsset and Økland, 2002). One common cause is errors in ground detection, particularly in dense understory or complex terrain, where laser pulses may fail to reach the true ground, resulting in an elevated digital terrain model (DTM) and underestimated normalised heights (Hyde *et al.*, 2005). While this factor likely contributed in some parcels with dense vegetation, the magnitude of its impact (typically 10–40 cm) is insufficient to fully explain the larger underestimation observed here.

Another major factor is the failure of LiDAR pulses to reach the actual treetop. This is especially problematic in deciduous forests, where laser pulses can penetrate the upper canopy before producing a return, causing the recorded first echo to originate from below the apex (Næsset and Økland, 2002; Hill *et al.*, 2002). The underestimation was more pronounced for taller trees, consistent with findings from Mielcarek *et al.* (2018), and may also reflect limitations in field measurement accuracy for large, complex crowns. Field height measurements in natural deciduous forests are prone to error due to irregular crown structures (Lim *et al.*, 2001; Maltamo *et al.*, 2011) and limited visibility of treetops, especially under leaf-on conditions, as was the case for this study (Yao *et al.*, 2020).

As mentioned before, matching errors between LiDAR and field trees likely also affected accuracy, as misaligned pairs could distort the height comparison (Maltamo *et al.*, 2011). Furthermore, structural complexity (e.g species composition, stand density) can reduce the likelihood of laser returns from the true treetop (Maltamo *et al.*, 2011).

Compared to previous studies, the RMSE observed here (4.23 m) is slightly lower than that reported by Mielcarek *et al.* (2011) for CHM-derived heights in mixed forests (5.8 m RMSE), though their models achieved much higher R^2 values (0.93–0.98). Accuracy reported for FW LiDAR is even higher, for instance, with RMSE as low as 1.6 m in similar forest conditions (Vastaranta *et al.*, 2009). These differences are likely due to higher point densities and the use of waveform data, both of which improve treetop detection and ground modelling.

To improve height estimation, using LiDAR with a higher pulse density would be beneficial. A denser point cloud increases the probability of capturing true treetops, enhances ground point detection, and supports better segmentation of individual trees.

6.2.2 CBH

CBH estimation showed the weakest performance among all structural metrics, with a very low R^2 (0.01), an RMSE of 9.08 m, and a positive bias of 6.59 m, indicating substantial overestimation. Several factors likely contributed to this poor accuracy.

A key limitation was segmentation error, which has been shown to significantly affect individual tree–based LiDAR metrics (Lu *et al.*, 2014). Another major factor was the calculation method, which relies on vertical point distribution to detect the height of the lowest substantial vegetation layer. This approach is sensitive to point density and requires sufficient points in the lower canopy (Sumnall *et al.*, 2016B; Stefanidou *et al.*, 2020). In this study, the use of upper canopy filtering, necessary to improve segmentation in low-density data, further reduced the already sparse understory information, severely limiting CBH detectability.

In addition, forest type likely played a role. Many CBH methods, especially those based on regression, are developed for coniferous forests, where crown structures are more regular and vertically aligned (Stefanidou *et al.*, 2020; Yang *et al.*, 2020; Chamberlain *et al.*, 2021). In deciduous forests, irregular and complex crown forms, along with overlapping vertical strata, complicate direct CBH detection (Luo *et al.*, 2018).

Compared to other studies, the results here are notably lower. For example, Yang *et al.* (2020) achieved an RMSE of 1.77 m using mixed linear regression modelling in coniferous forests, while Luo *et al.* (2018) reported RMSE of 0.88 m and R^2 of 0.62 in mixed forests using a direct approach. Chamberlain *et al.* (2021) used the same direct calculation as this study, achieving an RMSE of 1.25 m and R^2 of 0.4 in coniferous forest. These studies demonstrate that direct CBH estimation methods can be accurate, but their success relies on sufficient data resolution.

In this study, the upper canopy filtering, though necessary for segmentation, likely played a central role in reducing CBH accuracy. Future improvements could involve acquiring higher-density LiDAR data to capture more vertical detail and allow for segmentation without excluding understory points.

6.2.3 Crown Diameter

Crown diameter estimation yielded low accuracy, with $R^2 = 0.11$, an RMSE of 4.45 m, and a slight underestimation (bias = –0.92 m). Despite the low tree-level correspondence, the overall distribution of crown diameters closely matched the field data, suggesting that structural patterns were reasonably well captured at the stand level.

One possible source of error is the matching algorithm, as indicated by inconsistencies between tree- and parcel-level results. Another likely contributor is segmentation error, since undersegmentation tends to inflate crown diameter estimates, while oversegmentation leads to underestimation (Lu *et al.*, 2014).

Comparative studies focusing specifically on crown diameter are limited, as many report alternative crown metrics such as crown area, canopy cover, or volume (Vastaranta *et al.*, 2009; Nie *et al.*, 2017; Yao *et al.*, 2020). Some studies incorporate crown diameter into models but do not report its standalone accuracy (Dalponte & Coomes, 2016; Jucker *et al.*, 2017; Sumnall *et al.*, 2023). One relevant comparison is Zhang *et al.* (2015), who applied a convex hull–based method similar to the one used here and achieved $R^2 = 0.7$ and RMSE = 3.58 m in mixed forests, highlighting the potential of this approach under improved conditions.

To enhance crown diameter estimation, improvements in individual tree segmentation are essential. In addition, using in situ reference data with lower positional uncertainty could improve tree matching and more reliably assess accuracy at the individual level.

6.2.4 DBH

DBH estimation exhibited low to moderate accuracy, with an R^2 of 0.24, RMSE of 20.66 cm, and a negative bias of –8.53 cm, indicating a general underestimation. This result is expected, as DBH was indirectly derived from both tree height and crown diameter, which themselves were underestimated. Accordingly, all sources of error affecting those two metrics, particularly segmentation inaccuracies and tree matching issues, also propagated to DBH (Dalponte and Coomes, 2016; Yao *et al.*, 2020).

Moreover, the biological variability in deciduous forests complicates DBH estimation. Unlike coniferous trees, which generally exhibit strong apical dominance and more uniform growth forms, deciduous species show greater crown plasticity in response to competition. This flexibility in crown structure makes it harder to model DBH from structural traits (Hulshof *et al.*, 2015; Jucker *et al.*, 2017).

In terms of comparison, Jucker *et al.* (2017) used the same approach, predicting DBH from height and crown diameter, and reported RMSE = 9.7 cm across a global forest dataset, which proves that this approach is efficient. Other studies have achieved even lower RMSEs using more detailed crown metrics or limited their analysis to coniferous forests. For instance, Sumnall *et al.* (2022) used crown size and local attributes in coniferous stands and achieved RMSE = 3.13 cm, $R^2 = 0.62$. Similarly, Yao *et al.* (2020), using FW data and additional crown metrics, reported RMSE = 4.26 cm, $R^2 = 0.87$.

Although the results of this research fall short of these benchmarks, such comparisons should be interpreted with caution, given the lower point density, lack of FW data, and deciduous forest context in this study. The relatively poor performance likely comes from previously identified issues in height and crown diameter estimation, compounded by the difficulty of capturing structural variability in diverse, mixed-species forests.

6.2.5 AGB

The AGB results show that both the allometric and regression-based approaches underestimated aboveground biomass, particularly in larger trees. The allometric method, using LiDAR-derived DBH and height, resulted in an RMSE of 1500.24 kg C/ha, a bias of 979.30 kg C/ha, and an R^2 of 0.21. The regression model, using height and crown diameter, showed slightly better performance, with an RMSE of 1246.86 kg C/ha, a bias of

423.23 kg C/ha, and an R² of 0.22. The underestimation is largely explained by the compounded effects of segmentation errors and the underestimated input metrics (height, DBH, and crown diameter). Since taller trees were more underestimated in height, their contribution to total biomass was particularly affected, which is critical as these individuals disproportionately drive stand-level AGB.

To compare these results with literature values typically expressed in Mg C/ha, it is necessary to divide the values in kg by 1000 (e.g., 1500.24 kg C/ha = 1.50 Mg C/ha). This places the results slightly higher than Jucker *et al.* (2017), who reported RMSE of 1.70 Mg and bias of 30.1% for allometric calculation (Chave *et al.*, 2014), and RMSE of 1.78 Mg (bias: 4.3%) for regression modelling using the same input variables. In contrast, Nie *et al.* (2017) reported a plot-level RMSE of 5.99 Mg C/ha using FW data. While differences in data quality and forest type limit direct comparison, the results suggest that the regression model is preferable given its lower RMSE and bias. The limitation of this approach is that it either requires in situ data or known reliable relationship between height and crown diameter.

6.3 Influence of Forest Structure on Accuracy

Forest condition, particularly species composition, stand structure, and density, was expected to influence the accuracy of LiDAR-derived forest metrics, with more reliable results anticipated in structurally simpler, compositionally homogeneous stands (Maltamo *et al.*, 2011; Mielcarek *et al.*, 2018). This expectation was grounded in previous research showing that LiDAR performs better in stands where trees are more uniform in shape and size, the canopy is less vertically layered, and occlusion is reduced (Maltamo *et al.*, 2011; Mielcarek *et al.*, 2018). In such simplified structures, crown stratification is clearer, allowing better delineation of individual crowns and more accurate estimation of metrics such as tree height, crown diameter, crown base height, and DBH (Næsset, 2002; Sumnall *et al.*, 2016b; Stefanidou *et al.*, 2020). Additionally, field-based reference measurements tend to be more accurate in these stands due to reduced crown overlap and better visibility (Lim *et al.*, 2001), which improves validation quality. In contrast, dense, mixed, and irregular stands were expected to lead to more segmentation errors due to overlapping crowns and high structural heterogeneity, as well as reduced point penetration to lower canopy layers.

The results partially confirmed expectations. Tree height was more accurately estimated in the regular, pure, and less dense stands (Group 2), supporting the idea that simpler structure improves metric extraction. DBH also showed better performance in these stands at the tree level. In contrast, CBH showed no difference between stand types, likely due to overall poor performance. Crown diameter was similarly accurate across both groups on the parcel level, suggesting it was not strongly influenced by forest structure. For DBH and AGB, parcel-level results were unexpectedly better in the more complex, mixed, and dense stands (Group 1). In conclusion, forest condition clearly influenced height and tree-level DBH accuracy, but had no visible effects on other metrics due to the noise in the results.

6.4 Broader Implications, Transferability, and Future Prospects

This study demonstrated that LiDAR data, even with relatively low point-density and without FW information, can be effectively used to estimate key forest structural metrics, offering valuable insight into tree- and stand-level structure. While the accuracy varied across metrics, the results are promising and reinforce LiDAR's potential as a scalable tool for forest monitoring. Structural attributes derived in this study are not only useful for forest management, but also serve as critical inputs for ecological models. For example, models like HETEROFOR require detailed structural metrics, including DBH, tree height, crown dimensions, and CBH, which can feasibly be derived from LiDAR to simulate tree growth, competition, and carbon dynamics under changing conditions (Jonard *et al.*, 2020). Similarly, structural data support applications such as fire risk modelling (Hermosilla *et al.*, 2014), forest health monitoring (Parada-Díaz *et al.*, 2022), storm impact assessments (Nothdurft *et al.*, 2021), or tracking change due to disease or climate variability (Li *et al.*, 2024). With the future integration of species identification, LiDAR could also contribute to biodiversity assessments (Acebes *et al.*, 2021).

The metrics derived in this study, particularly DBH, tree height, and biomass, are also directly applicable to forest economics. They support pre-harvest inventory, timber volume estimation, carbon credit and stand productivity assessments, which are key for planning harvests and maximising economic returns. LiDAR enables scalable, cost-effective forest valuation, reducing reliance on field surveys (Wulder et al., 2012).

This research was designed to be transferable across forest types and scales by minimising reliance on in situ data and prioritising direct LiDAR-based metric extraction. However, for broader application, especially in mixed forests, further validation and adjustments are required. In particular, regression models used for DBH and AGB would need to be tested and potentially recalibrated in forests with differing species compositions or structural attributes (Ni-Meister et al., 2022). For coniferous forests, segmentation would require a newly trained Random Forest model, as point cloud characteristics differ significantly. For instance, crown size, canopy shape, and vertical profiles vary, and point density is much higher in winter conditions compared to leaf-off deciduous forest. Likewise, new regression models for DBH and AGB could either be species-specific (one-to-one model for deciduous and coniferous forests) or trained using combined deciduous–coniferous datasets. An efficient strategy would be to classify coniferous and deciduous areas using satellite imagery and apply the corresponding models accordingly, although this still depends on having appropriate in situ reference data.

Future improvements could also include incorporating understory layers in the analysis. While this would increase segmentation complexity, it would allow more complete structural profiling, particularly benefiting CBH estimation and potentially enabling quantification of understory vegetation (Crespo-Peremarch *et al.*, 2018). This is ecologically relevant, as the understory serves key roles in habitat provisioning, nutrient cycling, and regeneration processes (Wing *et al.*, 2012; Royo and Carson, 2006). In addition, expanding the scope of LiDAR-based metrics to include crown volume, stem volume, and species classification would further enhance ecosystem assessment capabilities (Yang *et al.*, 2020). These additions would not only support forest management but also provide critical inputs for broader ecosystem and carbon modelling frameworks, contributing to a more holistic understanding of forest dynamics.

7. Conclusion

This study aimed to evaluate how accurately LiDAR data can capture individual tree-level structural metrics relevant to forest management in temperate deciduous forests. To address this, a novel segmentation method was developed and tested, in which an existing algorithm was extended by incorporating a Random Forest model to identify incorrectly segmented trees, followed by an iterative correction process. Key structural metrics (tree height, DBH, crown base height, crown diameter, and aboveground biomass) were then extracted and validated against field measurements across forest stands with varying structural conditions.

The results demonstrated that even with relatively low point-density LiDAR data and no access to full waveform information, it is possible to estimate structural metrics with moderate to fair accuracy, particularly for tree height and crown diameter. However, metrics such as crown base height and DBH were less accurate, often showing systematic under- or overestimation. The performance of LiDAR-derived metrics was influenced by forest structure: even-aged and pure stands generally yielded better results at the tree level, while more structurally complex stands occasionally produced more accurate parcel-level metrics.

The developed segmentation method showed improved delineation compared to some standard approaches and offers a practical and scalable solution for individual tree detection in similar forest contexts.

Still, limitations such as the filtering of the understory vegetation and only focusing on deciduous forest constrained the analysis. The methods developed could be further refined by integrating more accurate in situ data, higher-density LiDAR and testing on more diverse forest conditions.

Overall, this work highlights both the potential and the limitations of LiDAR for operational forest monitoring. Future research should focus on improving metric estimation in complex stands, testing the approach in other forest types, and integrating species identification to broaden its usability. In addition, this study highlights the need for specific in situ data collection to validate LiDAR estimation. Despite its constraints, the research provides a meaningful step toward cost-effective, LiDAR-based forest structure analysis that supports both ecological understanding and forest economic planning.

Bibliography

Allouis, T., Durrieu, S., Vega, C., Couteron, P., 2013. Stem Volume and Above-Ground Biomass Estimation of Individual Pine Trees From LiDAR Data: Contribution of Full-Waveform Signals. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 6, 924–934. https://doi.org/10.1109/JSTARS.2012.2211863

Anderson, K., Hancock, S., Disney, M., Gaston, K.J., 2016. Is waveform worth it? A comparison of LiDAR approaches for vegetation and landscape characterization. Remote Sensing in Ecology and Conservation 2, 5–15. https://doi.org/10.1002/rse2.8

Aplin, P., 2005. Remote sensing: ecology. Progress in Physical Geography: Earth and Environment 29, 104–113. https://doi.org/10.1191/030913305pp437pr

Ayrey, E., Fraver, S., Kershaw, J.A., Kenefic, L.S., Hayes, D., Weiskittel, A.R., Roth, B.E., 2017. Layer stacking: A novel algorithm for individual forest tree segmentation from LiDAR point clouds. Canadian Journal of Remote Sensing. 43(1): 16-27. 43, 16–27. https://doi.org/10.1080/07038992.2017.1252907

Bonheure, Mathilde, 2022. Effects of stand characteristics and silvicultural treatment on forest resilience to storms in the context of climate change.Faculty of Bioengineering, Catholic University of Louvain, 2022. Prom.: Ponette, Quentin; Jonard, Mathieu. http://hdl.handle.net/2078.1/thesis:35756

Bouvier, M., Durrieu, S., Fournier, R.A., Renaud, J.-P., 2015. Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. Remote Sensing of Environment 156, 322–334. https://doi.org/10.1016/j.rse.2014.10.004

Bruggisser, M., Roncat, A., Schaepman, M.E., Morsdorf, F., 2017. Retrieval of higher order statistical moments from full-waveform LiDAR data for tree species classification. Remote Sensing of Environment 196, 28–41. https://doi.org/10.1016/j.rse.2017.04.025

Chamberlain, C.P., Sánchez Meador, A.J., Thode, A.E., 2021. Airborne lidar provides reliable estimates of canopy base height and canopy bulk density in southwestern ponderosa pine forests. Forest Ecology and Management 481, 118695. https://doi.org/10.1016/j.foreco.2020.118695

Chang, L., Fan, H., Zhu, N., Dong, Z., 2022. A Two-Stage Approach for Individual Tree Segmentation From TLS Point Clouds. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 15, 8682–8693. https://doi.org/10.1109/JSTARS.2022.3212445

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M.S., Delitti, W.B.C., Duque, A., Eid, T., Fearnside, P.M., Goodman, R.C., Henry, M., Martínez-Yrízar, A., Mugasha, W.A., Muller-Landau, H.C., Mencuccini, M., Nelson, B.W., Ngomanda, A., Nogueira, E.M., Ortiz-Malavassi, E., Pélissier, R., Ploton, P., Ryan, C.M., Saldarriaga, J.G., Vieilledent, G., 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. Global Change Biology 20, 3177–3190. https://doi.org/10.1111/gcb.12629

Crespo-Peremarch, P., Tompalski, P., Coops, N.C., Ruiz, L.Á., 2018. Characterizing understory vegetation in Mediterranean forests using full-waveform airborne laser scanning data. Remote Sensing of Environment 217, 400–413. https://doi.org/10.1016/j.rse.2018.08.033

Dalponte, M., Coomes, D.A., 2016. Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. Methods in Ecology and Evolution 7, 1236–1245. https://doi.org/10.1111/2041-210X.12575

Diószegi, G., Molnár, V.É., Nagy, L.A., Enyedi, P., Török, P., Szabó, S., 2024. A new method for individual treetop detection with low-resolution aerial laser scanned data. Model. Earth Syst. Environ. 10, 5225–5240. https://doi.org/10.1007/s40808-024-02060-w

Duncanson, L.I., Cook, B.D., Hurtt, G.C., Dubayah, R.O., 2014. An efficient, multi-layered crown delineation algorithm for mapping individual tree structure across multiple ecosystems. Remote Sensing of Environment 154, 378–386. https://doi.org/10.1016/j.rse.2013.07.044

Eysn, L., Hollaus, M., Lindberg, E., Berger, F., Monnet, J.-M., Dalponte, M., Kobal, M., Pellegrini, M., Lingua, E., Mongus, D., Pfeifer, N., 2015. A Benchmark of Lidar-Based Single Tree Detection Methods Using Heterogeneous Forest Data from the Alpine Space. Forests 6, 1721–1747. https://doi.org/10.3390/f6051721

Fernandez-Diaz, J., Carter, W., 2013. Understanding Waveform Digitizing and Waveform Data Processing NCALM’s Data Users White Paper NCALM-WP-2013-01. https://doi.org/10.13140/RG.2.2.15608.55042

Goldstein, A., Turner, W.R., Spawn, S.A., Anderson-Teixeira, K.J., Cook-Patton, S., Fargione, J., Gibbs, H.K., Griscom, B., Hewson, J.H., Howard, J.F., Ledezma, J.C., Page, S., Koh, L.P., Rockström, J., Sanderman, J., Hole, D.G., 2020. Protecting irrecoverable carbon in Earth's ecosystems. Nat. Clim. Chang. 10, 287–295. https://doi.org/10.1038/s41558-020-0738-8

Guignabert, A., Jonard, M., Messier, C., André, F., de Coligny, F., Doyon, F., Ponette, Q., 2024. Adaptive forest management improves stand-level resilience of temperate forests under multiple stressors. Science of The Total Environment 948, 174168. https://doi.org/10.1016/j.scitotenv.2024.174168

Hamraz, H., Contreras, M.A., Zhang, J., 2016. A robust approach for tree segmentation in deciduous forests using small-footprint airborne LiDAR data. International Journal of Applied Earth Observation and Geoinformation 52, 532–541. https://doi.org/10.1016/j.jag.2016.07.006

Hancock, S., Anderson, K., Disney, M., Gaston, K.J., 2017. Measurement of fine-spatial-resolution 3D vegetation structure with airborne waveform lidar: Calibration and validation with voxelised terrestrial lidar. Remote Sensing of Environment 188, 37–50. https://doi.org/10.1016/j.rse.2016.10.041

He, Q.-S., Cao ,Chun-Xiang, Chen ,Er-Xue, Sun ,Guo-Qing, Ling ,Fei-Long, Pang ,Yong, Zhang ,Hao, Ni ,Wen-Jian, Xu ,Min, Li ,Zeng-Yuan, and Li, X.-W., 2012. Forest stand biomass estimation using ALOS PALSAR data based on LiDAR-derived prior knowledge in the Qilian Mountain, western China. International Journal of Remote Sensing 33, 710–729. https://doi.org/10.1080/01431161.2011.577829

Hemery, G.E., Savill, P.S., Pryor, S.N., 2005. Applications of the crown diameter–stem diameter relationship for different species of broadleaved trees. Forest Ecology and Management 215, 285–294. https://doi.org/10.1016/j.foreco.2005.05.016

Henrich, J., van Delden, J., Seidel, D., Kneib, T., Ecker, A.S., 2024. TreeLearn: A deep learning method for segmenting individual trees from ground-based LiDAR forest point clouds. Ecological Informatics 84, 102888. https://doi.org/10.1016/j.ecoinf.2024.102888

Hermosilla, T., Ruiz, L.A., Kazakova, A.N., Coops, N.C., Moskal, L.M., 2013. Estimation of forest structure and canopy fuel parameters from small-footprint full-waveform LiDAR data. Int. J. Wildland Fire 23, 224–233. https://doi.org/10.1071/WF13086

Hermosilla, T., Coops, N.C., Ruiz, L.A., Moskal, L.M., 2014. Deriving pseudo-vertical waveforms from small-footprint full-waveform LiDAR data. Remote Sensing Letters 5, 332–341. https://doi.org/10.1080/2150704X.2014.903350

Hill, R.A., Gaveau, D.L.A., and Spendlove, M. 2002. Accuracy issues in the assessment of deciduous woodland canopy height by airborne laser scanning: a case study. In Proceedings of the ForestSAT 2002 Conference, Edinburgh, UK, 5–9 August 2002. [CD-ROM]. Forestry Commission England and Forestry Commission Scotland, Edinburgh, UK. (PDF) The accuracy of estimating individual tree variables with airborne laser scanning in a boreal nature reserve. Available from: https://www.researchgate.net/publication/237866238_The_accuracy_of_estimating_individual_tree_variables_with_airborne_laser_scanning_in_a_boreal_nature_reserve [Accessed May 18 2025]

Hulshof, C.M., Swenson, N.G., Weiser, M.D., 2015. Tree height–diameter allometry across the United States. Ecology and Evolution 5, 1193–1204. https://doi.org/10.1002/ece3.1328

Hyde, P., Dubayah, R., Peterson, B., Blair, J.B., Hofton, M., Hunsaker, C., Knox, R., Walker, W., 2005. Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems. Remote Sensing of Environment 96, 427–437. https://doi.org/10.1016/j.rse.2005.03.005

Hyyppa, J., Kelle, O., Lehikoinen, M., Inkinen, M., 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. IEEE Transactions on Geoscience and Remote Sensing 39, 969–975. https://doi.org/10.1109/36.921414

IFER. 2010. Field-Map, a tool to help collect and process field data. Monitoring and mapping solutions. https://fieldmap.cz/download/FM_catalogue_fr.pdf

Isenburg, M. (2022). LAStools - Efficient LiDAR Processing Software (Version 2022). Rapidlasso GmbH. http://rapidlasso.com/lastools

Jonard, M., André, F., de Coligny, F., de Wergifosse, L., Beudez, N., Davi, H., Ligot, G., Ponette, Q., Vincke, C., 2020. HETEROFOR 1.0: a spatially explicit model for exploring the response of structurally complex forests to uncertain future conditions – Part 1: Carbon fluxes and tree dimensional growth. Geoscientific Model Development 13, 905–935. https://doi.org/10.5194/gmd-13-905-2020

Jucker, T., Caspersen, J., Chave, J., Antin, C., Barbier, N., Bongers, F., Dalponte, M., van Ewijk, K.Y., Forrester, D.I., Haeni, M., Higgins, S.I., Holdaway, R.J., Iida, Y., Lorimer, C., Marshall, P.L., Momo, S., Moncrieff, G.R., Ploton, P., Poorter, L., Rahman, K.A., Schlund, M., Sonké, B., Sterck, F.J., Trugman, A.T., Usoltsev, V.A., Vanderwel, M.C., Waldner, P., Wedeux, B.M.M., Wirth, C., Wöll, H., Woods, M., Xiang, W., Zimmermann, N.E., Coomes, D.A., 2017. Allometric equations for integrating remote sensing imagery into forest monitoring programmes. Global Change Biology 23, 177–190. https://doi.org/10.1111/gcb.13388

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., Næsset, E., Pitkänen, J., Popescu, S., Solberg, S., Wolf, B.M., Wu, J.-C., 2012. An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. Remote Sensing 4, 950–974. https://doi.org/10.3390/rs4040950

Khosravipour, A., Skidmore, A., Isenburg, M., Hussin, Y., 2014. Generating Pit-free Canopy Height Models from Airborne Lidar. Photogrammetric Engineering & Remote Sensing 80, 863–872. https://doi.org/10.14358/PERS.80.9.863

King, D.A., 2005. Linking tree form, allocation and growth with an allometrically explicit model. Ecological Modelling 185, 77–91. https://doi.org/10.1016/j.ecolmodel.2004.11.017

Lefsky, M., Cohen, W., Parker, G., Harding, D., 2009. Lidar Remote Sensing for Ecosystem Studies. BioScience 52, 19–30. https://doi.org/10.1641/0006-3568(2002)052[0019:LRSFES]2.0.CO;2

Li, H., Hiroshima, T., Li, X., Hayashi, M., Kato, T., 2024. High-resolution mapping of forest structure and carbon stock using multi-source remote sensing data in Japan. Remote Sensing of Environment 312, 114322. https://doi.org/10.1016/j.rse.2024.114322

Li, W., Guo, Q., Jakubowski, M.K., Kelly, M., 2012. A new method for segmenting individual trees from the lidar point cloud. Photogrammetric Engineering and Remote Sensing 78, 75–84. https://doi.org/10.14358/PERS.78.1.75

Lim, K., Treitz, P., Groot, A., and St-Onge, B. 2001. Estimation ofindividual tree heights using LIDAR remote sensing. In Pro-ceedings of the 23rd Annual Canadian Symposium on RemoteSensing, Québec, Que., 20–24 August 2001. CASI, Ottawa, Ont.pp. 243–250 *(PDF) The* *accuracy of estimating individual tree variables with airborne laser scanning in a boreal nature reserve*. Available from: https://www.researchgate.net/publication/237866238_The_accuracy_of_estimating_individual_tree_variables_with_airborne_laser_scanning_in_a_boreal_nature_reserve [Accessed May 18 2025].

Lim, K., Treitz, P., Baldwin, K., Morrison, I., Green, J., 2003. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. Canadian Journal of Remote Sensing 29, 658–678. https://doi.org/10.5589/m03-025

Lindberg, E., Holmgren, J., Olofsson, K., Wallerman, J., Olsson, H., 2010. Estimation of tree lists from airborne laser scanning by combining single-tree and area-based methods. International Journal of Remote Sensing 31, 1175–1192. https://doi.org/10.1080/01431160903380649

Lu, X., Guo, Q., Li, W., Flanagan, J., 2014. A bottom-up approach to segment individual deciduous trees using leaf-off lidar point cloud data. ISPRS Journal of Photogrammetry and Remote Sensing 94, 1–12. https://doi.org/10.1016/j.isprsjprs.2014.03.014

Luo, L., Zhai, Q., Su, Y., Ma, Q., Kelly, M., Guo, Q., 2018. Simple method for direct crown base height estimation of individual conifer trees using airborne LiDAR data. Opt. Express, OE 26, A562–A578. https://doi.org/10.1364/OE.26.00A562

Mallet, C., Bretar, F., 2009. Full-waveform topographic lidar: State-of-the-art. ISPRS Journal of Photogrammetry and Remote Sensing 64, 1–16. https://doi.org/10.1016/j.isprsjprs.2008.09.007

Maltamo, M., Mustonen, K., Hyyppä, J., Pitkänen, J., Yu, X., 2004. The accuracy of estimating individual tree variables with airborne laser scanning in a boreal nature reserve. Can. J. For. Res. 34, 1791–1801. https://doi.org/10.1139/x04-055

Mielcarek, M., Stereńczak, K., Khosravipour, A., 2018. Testing and evaluating different LiDAR-derived canopy height model generation methods for tree height estimation. International Journal of Applied Earth Observation and Geoinformation 71, 132–143. https://doi.org/10.1016/j.jag.2018.05.002

Molnar, C., Casalicchio, G., Bischl, B., 2018. iml: An R package for Interpretable Machine Learning. Journal of Open Source Software 3, 786. https://doi.org/10.21105/joss.00786

Molnar, C., 2019. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable. https://christophm.github.io/interpretable-ml-book/

Morsdorf, F., Mårell, A., Koetz, B., Cassagne, N., Pimont, F., Rigolot, E., Allgöwer, B., 2010. Discrimination of vegetation strata in a multi-layered Mediterranean forest ecosystem using height and intensity information derived from airborne laser scanning. Remote Sensing of Environment 114, 1403–1415. https://doi.org/10.1016/j.rse.2010.01.023

Næsset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sensing of Environment 80, 88–99. https://doi.org/10.1016/S0034-4257(01)00290-5

Næsset, E., Økland, T., 2002. Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. Remote Sensing of Environment 79, 105–115. https://doi.org/10.1016/S0034-4257(01)00243-7

Nie, S., Wang, C., Zeng, H., Xi, X., Li, G., 2017. Above-ground biomass estimation using airborne discrete-return and full-waveform LiDAR data in a coniferous forest. Ecological Indicators 78, 221–228. https://doi.org/10.1016/j.ecolind.2017.02.045

Ni-Meister, W., Rojas, A., Lee, S., 2022. Direct use of large-footprint lidar waveforms to estimate aboveground biomass. Remote Sensing of Environment 280, 113147. https://doi.org/10.1016/j.rse.2022.113147

Nothdurft, A., Gollob, C., Kraßnitzer, R., Erber, G., Ritter, T., Stampfer, K., Finley, A.O., 2021. Estimating timber volume loss due to storm damage in Carinthia, Austria, using ALS/TLS and spatial regression models. https://doi.org/10.48550/arXiv.2106.08912

Parada-Díaz, J., Fernández López, Á.B., Gómez González, L.A., del Arco Aguilar, M.J., González-Mancebo, J.M., 2022. Assessing the Usefulness of LiDAR for Monitoring the Structure of a Montane Forest on a Subtropical Oceanic Island. Remote Sensing 14, 994. https://doi.org/10.3390/rs14040994

Park, S.-H., Jung, H.-S., Lee, S., Kim, E.-S., 2021. Mapping Forest Vertical Structure in Sogwang-ri Forest from Full-Waveform Lidar Point Clouds Using Deep Neural Network. Remote Sensing 13, 3736. https://doi.org/10.3390/rs13183736

Popescu, S.C., Wynne, R.H., Nelson, R.F., 2002. Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size. Computers and Electronics in Agriculture 37, 71–95. https://doi.org/10.1016/S0168-1699(02)00121-7

Popescu, S.C., Zhao, K., 2008. A voxel-based lidar method for estimating crown base height for deciduous and pine trees. Remote Sensing of Environment 112, 767–781. https://doi.org/10.1016/j.rse.2007.06.011

QGIS Development Team (2024). QGIS Geographic Information System (Version 3.28.10). Open Source Geospatial Foundation. http://qgis.org

R Core Team (2023). _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Reitberger, J., Krzystek, P., Stilla, U., 2008. Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees. International Journal of Remote Sensing 29, 1407–1431. https://doi.org/10.1080/01431160701736448

Reitberger, J., Schnörr, Cl., Krzystek, P., Stilla, U., 2009. 3D segmentation of single trees exploiting full waveform LIDAR data. ISPRS Journal of Photogrammetry and Remote Sensing 64, 561–574. https://doi.org/10.1016/j.isprsjprs.2009.04.002

Roussel J, Auty D (2024). Airborne LiDAR Data Manipulation and Visualization for Forestry Applications. R package version 4.1.2, https://cran.r-project.org/package=lidR.

Roussel, J.-R., Auty, D., Coops, N.C., Tompalski, P., Goodbody, T.R.H., Meador, A.S., Bourdon, J.-F., de Boissieu, F., Achim, A., 2020. lidR: An R package for analysis of Airborne Laser Scanning (ALS) data. Remote Sensing of Environment 251, 112061. https://doi.org/10.1016/j.rse.2020.112061s

Royo, A., Carson, W.P., 2006. On the formation of dense understory layers in forests worldwide: consequences and implications for forest dynamics, biodiversity, and succession. Canadian Journal of Forest Research 36, 1345–1362. https://doi.org/10.1139/x06-025

Ruiz, L.Á., Recio, J.A., Crespo-Peremarch, P., Sapena, M., 2018. An object-based approach for mapping forest structural types based on low-density LiDAR and multispectral imagery. Geocarto International 33, 443–457. https://doi.org/10.1080/10106049.2016.1265595

Shugart, H.H., Saatchi, S., Hall, F.G., 2010. Importance of structure and its measurement in quantifying function of forest ecosystems. Journal of Geophysical Research: Biogeosciences 115. https://doi.org/10.1029/2009JG000993

Solberg, S., Naesset, E., Bollandsas, O.M., 2006. Single Tree Segmentation Using Airborne Laser Scanner Data in a Structurally Heterogeneous Spruce Forest. Photogrammetric Engineering & Remote Sensing 72, 1369–1378. https://doi.org/10.14358/PERS.72.12.1369

SPW (Service Public de Wallonie), 2024. *Généralités et débats – Suivi du LiDAR#3 Workshop*. [online] Available at: https://geoportail.wallonie.be/files/PDF/LiDAR24/Presentations_Generalite_Debats.pdf [Accessed 19 May 2025].

Stefanidou, A., Gitas, I.Z., Korhonen, L., Stavrakoudis, D., Georgopoulos, N., 2020. LiDAR-Based Estimates of Canopy Base Height for a Dense Irregular Structured Forest. Remote Sensing 12, 1565. https://doi.org/10.3390/rs12101565

Stereńczak, K., Kraszewski, B., Mielcarek, M., Piasecka, Ż., Lisiewicz, M., Heurich, M., 2020. Mapping individual trees with airborne laser scanning data in an European lowland forest using a self-calibration algorithm. International Journal of Applied Earth Observation and Geoinformation 93, 102191. https://doi.org/10.1016/j.jag.2020.102191

Sumnall, M.J., Hill, R.A., Hinsley, S.A., 2016A. Comparison of small-footprint discrete return and full waveform airborne lidar data for estimating multiple forest variables. Remote Sensing of Environment 173, 214–223. https://doi.org/10.1016/j.rse.2015.07.027

Sumnall, M.J., Peduzzi ,Alicia, Fox ,Thomas R., Wynne ,Randolph H., and Thomas, V.A., 2016B. Analysis of a lidar voxel-derived vertical profile at the plot and individual tree scales for the estimation of forest canopy layer characteristics. International Journal of Remote Sensing 37, 2653–2681. https://doi.org/10.1080/01431161.2016.1183833

Sumnall, M.J., Albaugh, T.J., Carter, D.R., Cook, R.L., Hession, W.C., Campoe, O.C., Rubilar, R.A., Wynne, R.H., Thomas, V.A., 2023. Estimation of individual stem volume and diameter from segmented UAV laser scanning datasets in Pinus taeda L. plantations. International Journal of Remote Sensing 44, 217–247. https://doi.org/10.1080/01431161.2022.2161853

Vastaranta, M., Holopainen, M., Reija, H., Yu, X., Melkas, T., Hyyppä, J., Hyyppä, H., 2009. Comparison between an area-based and individual tree detection method for low-pulse density ALS-based forest inventory. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 38, 147–151.

Vaughn, N.R., Moskal, L.M., Turnblom, E.C., 2012. Tree Species Detection Accuracies Using Discrete Point Lidar and Airborne Waveform Lidar. Remote Sensing 4, 377–403. https://doi.org/10.3390/rs4020377

Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkänen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K.M., Lien, V., Packalén, P., Gobakken, T., Koch, B., Næsset, E., Tokola, T., Maltamo, M., 2012. Comparative testing of single-tree detection algorithms under different types of forest. Forestry: An International Journal of Forest Research 85, 27–40. https://doi.org/10.1093/forestry/cpr051

Vega, C., Hamrouni, A., El Mokhtari, S., Morel, J., Bock, J., Renaud, J.-P., Bouvier, M., Durrieu, S., 2014. PTrees: A point-based approach to forest tree extraction from lidar data. International Journal of Applied Earth Observation and Geoinformation 33, 98–108. https://doi.org/10.1016/j.jag.2014.05.001

Vierling, L.A., Xu, Y., Eitel, J.U.H., Oldow, J.S., 2013. Shrub characterization using terrestrial laser scanning and implications for airborne LiDAR assessment. Canadian Journal of Remote Sensing. https://doi.org/10.5589/m12-057

Wang, K., Kumar, P., 2019. Characterizing relative degrees of clumping structure in vegetation canopy using waveform LiDAR. Remote Sensing of Environment 232, 111281. https://doi.org/10.1016/j.rse.2019.111281

Wang, Y., Weinacker, H., Koch, B., 2008. A Lidar Point Cloud Based Procedure for Vertical Canopy Structure Analysis And 3D Single Tree Modelling in Forest. Sensors 8, 3938–3951. https://doi.org/10.3390/s8063938

Wilkes, P., Disney, M., Armston, J., Bartholomeus, H., Bentley, L., Brede, B., Burt, A., Calders, K., Chavana-Bryant, C., Clewley, D., Duncanson, L., Forbes, B., Krisanski, S., Malhi, Y., Moffat, D., Origo, N., Shenkin, A., Yang, W., 2023. TLS2trees: A scalable tree segmentation pipeline for TLS data. Methods in Ecology and Evolution 14, 3083–3099. https://doi.org/10.1111/2041-210X.14233

Wing, B.M., Ritchie, M.W., Boston, K., Cohen, W.B., Gitelman, A., Olsen, M.J., 2012. Prediction of understory vegetation cover with airborne lidar in an interior ponderosa pine forest. Remote Sensing of Environment 124, 730–741. https://doi.org/10.1016/j.rse.2012.06.024

Wulder, M.A., White, J.C., Nelson, R.F., Næsset, E., Ørka, H.O., Coops, N.C., Hilker, T., Bater, C.W., Gobakken, T., 2012. Lidar sampling for large-area forest characterization: A review. Remote Sensing of Environment 121, 196–209. https://doi.org/10.1016/j.rse.2012.02.001

Yang, Z., Liu, Q., Luo, P., Ye, Q., Sharma, R.P., Duan, G., Zhang, H., Fu, L., 2020. Nonlinear mixed-effects height to crown base model based on both airborne LiDAR and field datasets for *Picea crassifolia* Kom trees in northwest China. Forest Ecology and Management 474, 118323. https://doi.org/10.1016/j.foreco.2020.118323

Yao, W., Krzystek, P., Heurich, M., 2012. Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. Remote Sensing of Environment 123, 368–380. https://doi.org/10.1016/j.rse.2012.03.027

Yu, X., Litkey, P., Hyyppä, J., Holopainen, M., Vastaranta, M., 2014. Assessment of Low Density Full-Waveform Airborne Laser Scanning for Individual Tree Detection and Tree Species Classification. Forests 5, 1011–1031. https://doi.org/10.3390/f5051011

Zhang, C., Zhou, Y., Qiu, F., 2015. Individual Tree Segmentation from LiDAR Point Clouds for Urban Forest Inventory. Remote Sensing 7, 7892–7913. https://doi.org/10.3390/rs70607892

Zhang, K., Chen, S.-C., Whitman, D., Shyu, M.-L., Yan, J., Zhang, C., 2003. A progressive morphological filter for removing nonground measurements from airborne LIDAR data. IEEE Transactions on Geoscience and Remote Sensing 41, 872–882. https://doi.org/10.1109/TGRS.2003.810682

Zhang, W., Xi, Y., Brandt, M., Ren, C., Bai, J., Ma, Q., Fensholt, R., 2024. Stand structure of tropical forests is strongly associated with primary productivity. Commun Earth Environ 5, 1–9. https://doi.org/10.1038/s43247-024-01984-6

Appendix

Figures

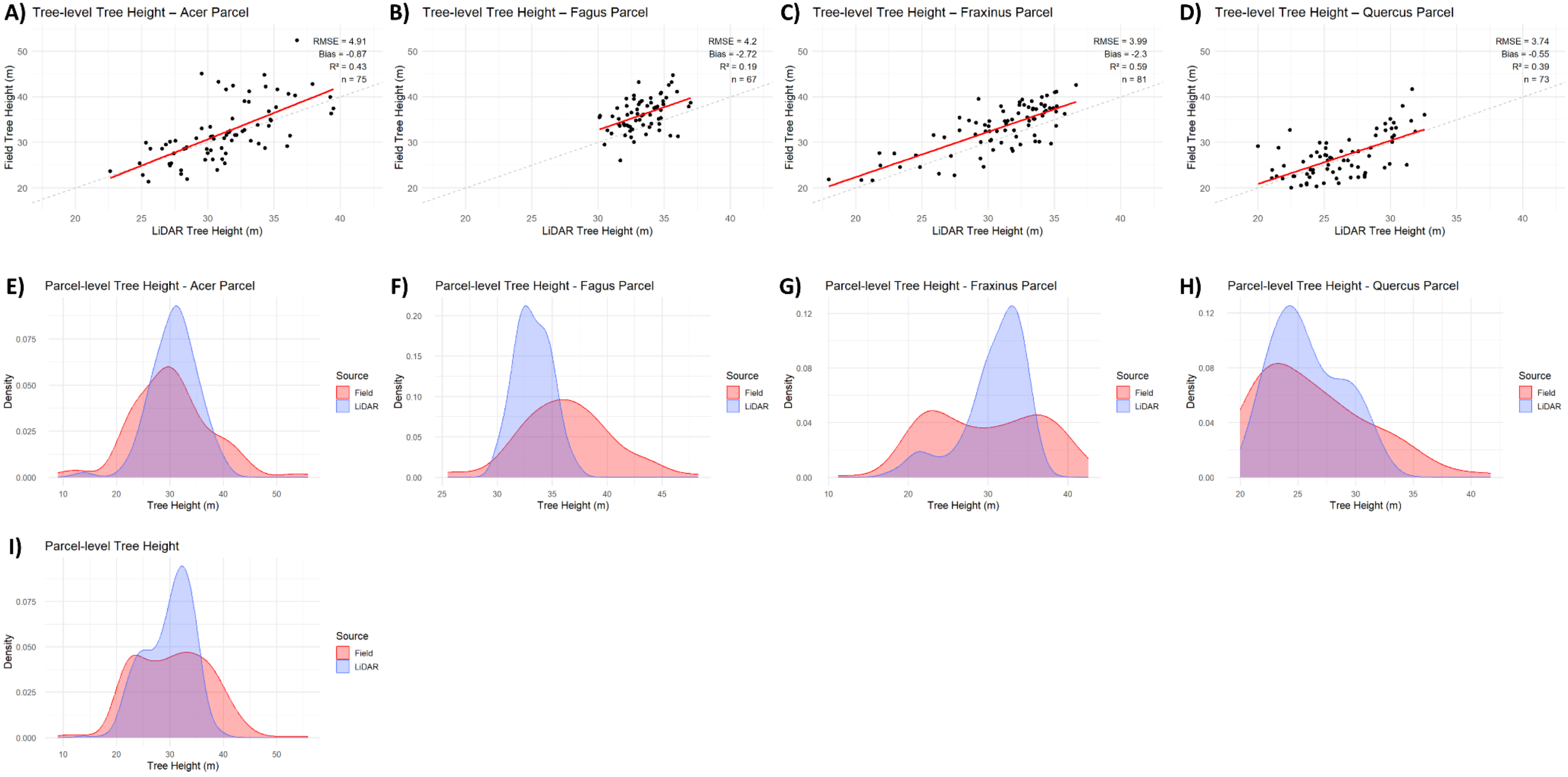


Figure A 1: Comparison between LiDAR-derived and field-measured height at tree and parcel levels. (A–D) Tree-level scatterplots per parcel (Acer, Fagus, Fraxinus, Quercus); red line shows the best-fit regression, black line indicates the 1:1 theoretical relationship. RMSE, bias, R², and sample size (n) are reported. (E–H) Parcel-level density plots comparing the distribution of field and LiDAR-derived values per parcel. (I) Combined parcel-level density plot comparing overall LiDAR and field distributions.

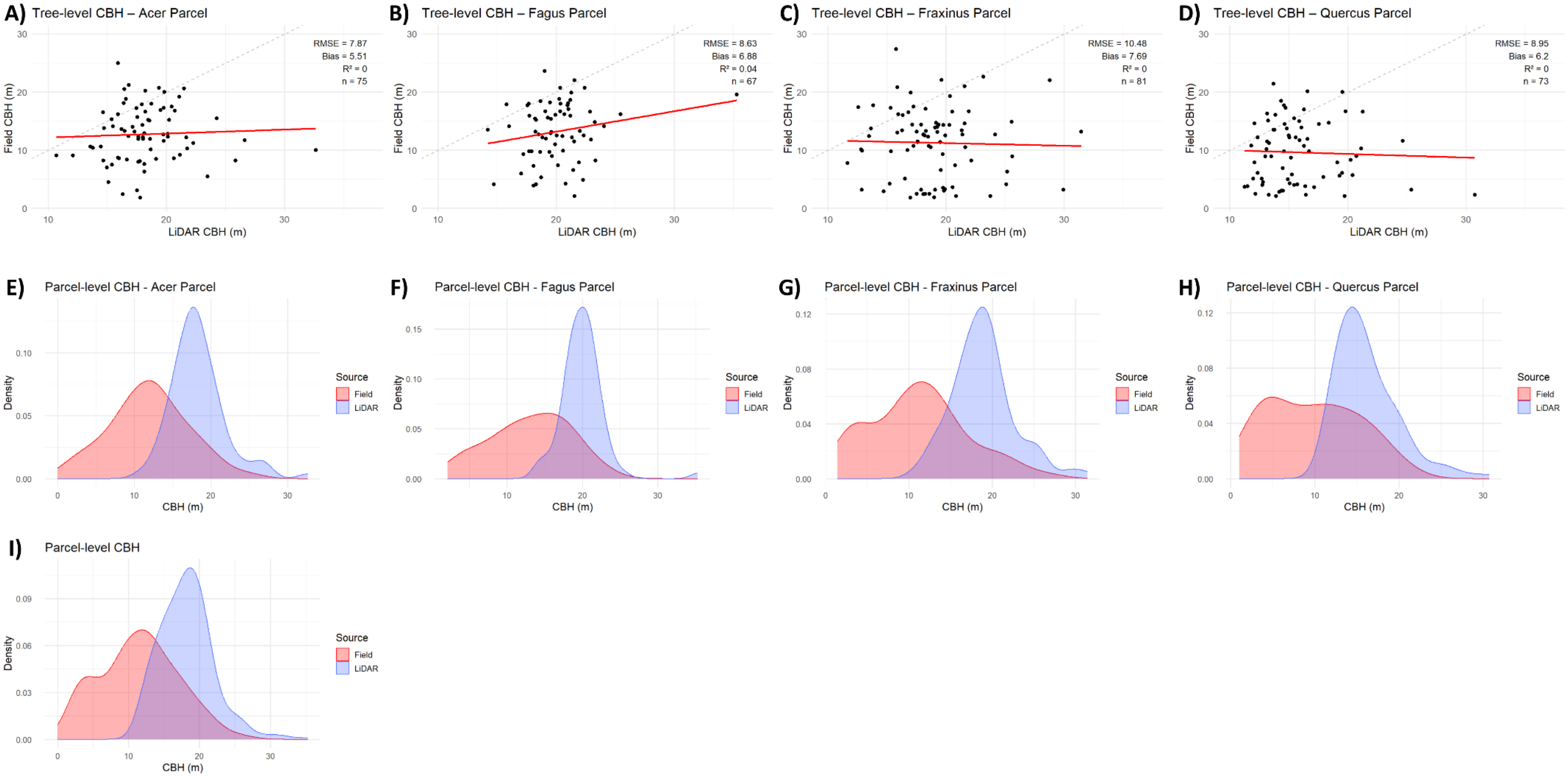


Figure A 2: Comparison between LiDAR-derived and field-measured CBH at tree and parcel levels. (A–D) Tree-level scatterplots per parcel (Acer, Fagus, Fraxinus, Quercus); red line shows the best-fit regression, black line indicates the 1:1 theoretical relationship. RMSE, bias, R², and sample size (n) are reported. (E–H) Parcel-level density plots comparing the distribution of field and LiDAR-derived values per parcel. (I) Combined parcel-level density plot comparing overall LiDAR and field distributions.

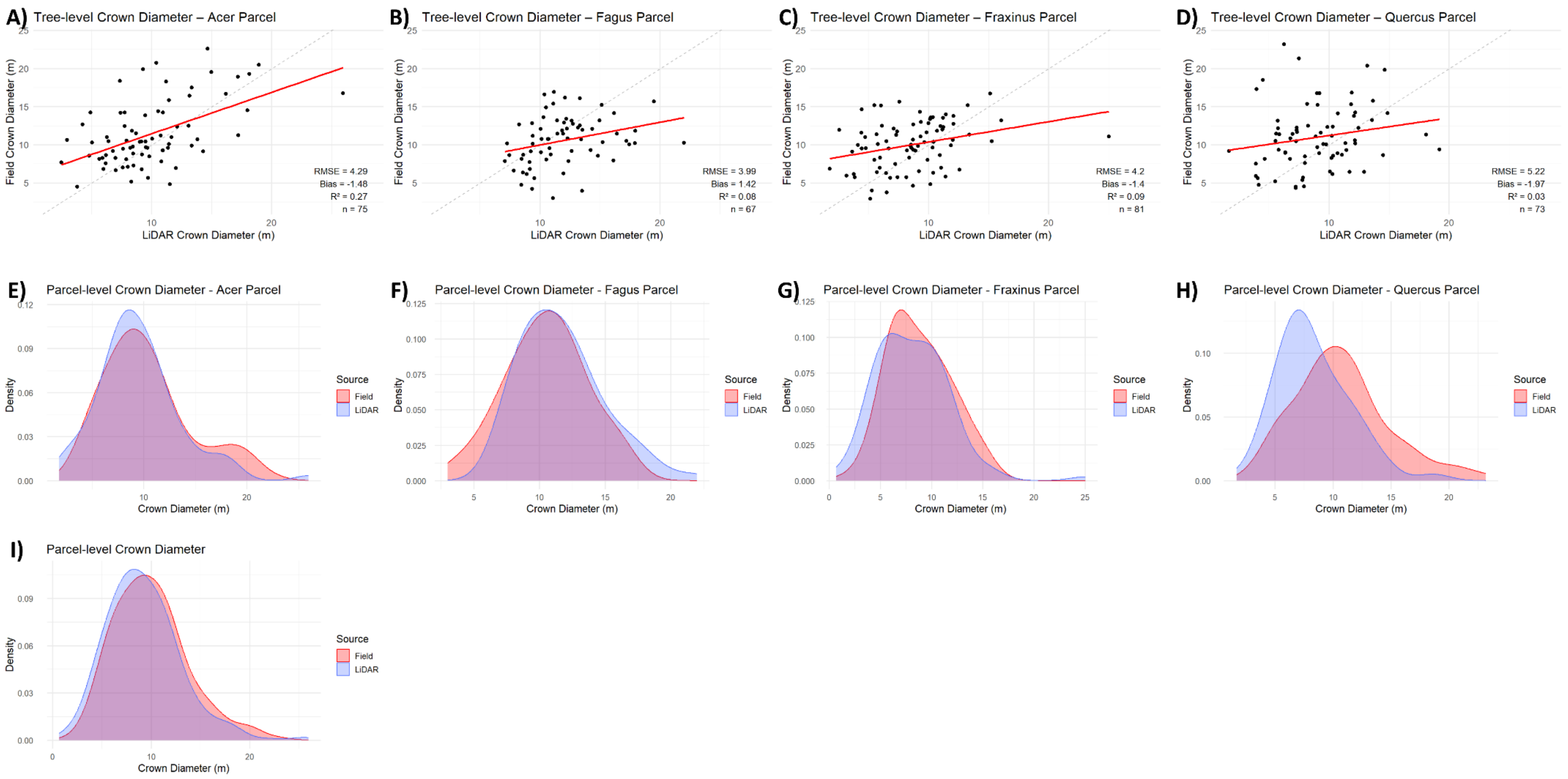


Figure A 3: Comparison between LiDAR-derived and field-measured crown diameter at tree and parcel levels. (A–D) Tree-level scatterplots per parcel (Acer, Fagus, Fraxinus, Quercus); red line shows the best-fit regression, black line indicates the 1:1 theoretical relationship. RMSE, bias, R², and sample size (n) are reported. (E–H) Parcel-level density plots comparing the distribution of field and LiDAR-derived values per parcel. (I) Combined parcel-level density plot comparing overall LiDAR and field distributions.

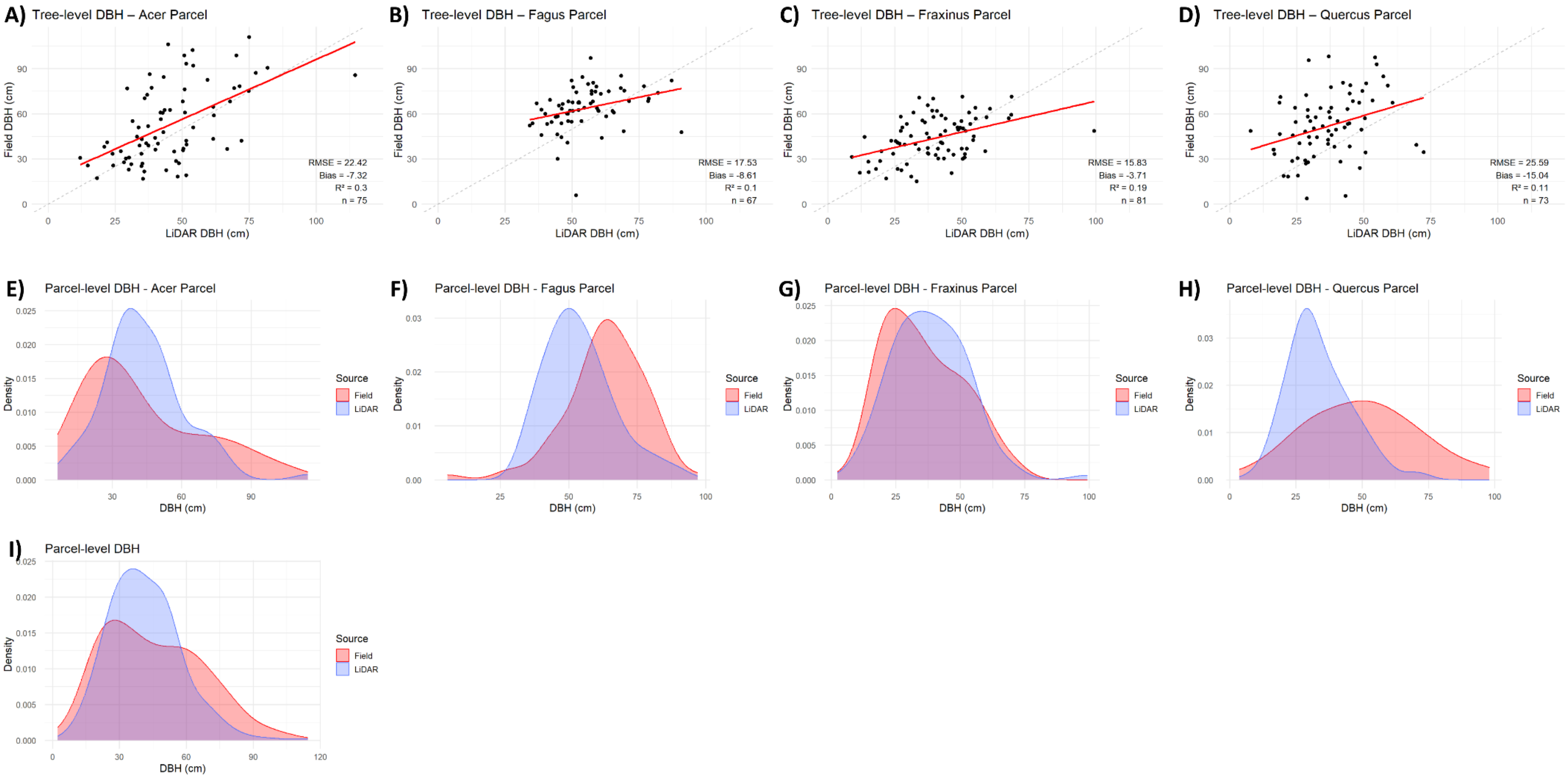


Figure A 4: Comparison between LiDAR-derived and field-measured DBH at tree and parcel levels. (A–D) Tree-level scatterplots per parcel (Acer, Fagus, Fraxinus, Quercus); red line shows the best-fit regression, black line indicates the 1:1 theoretical relationship. RMSE, bias, R², and sample size (n) are reported. (E–H) Parcel-level density plots comparing the distribution of field and LiDAR-derived values per parcel. (I) Combined parcel-level density plot comparing overall LiDAR and field distributions.

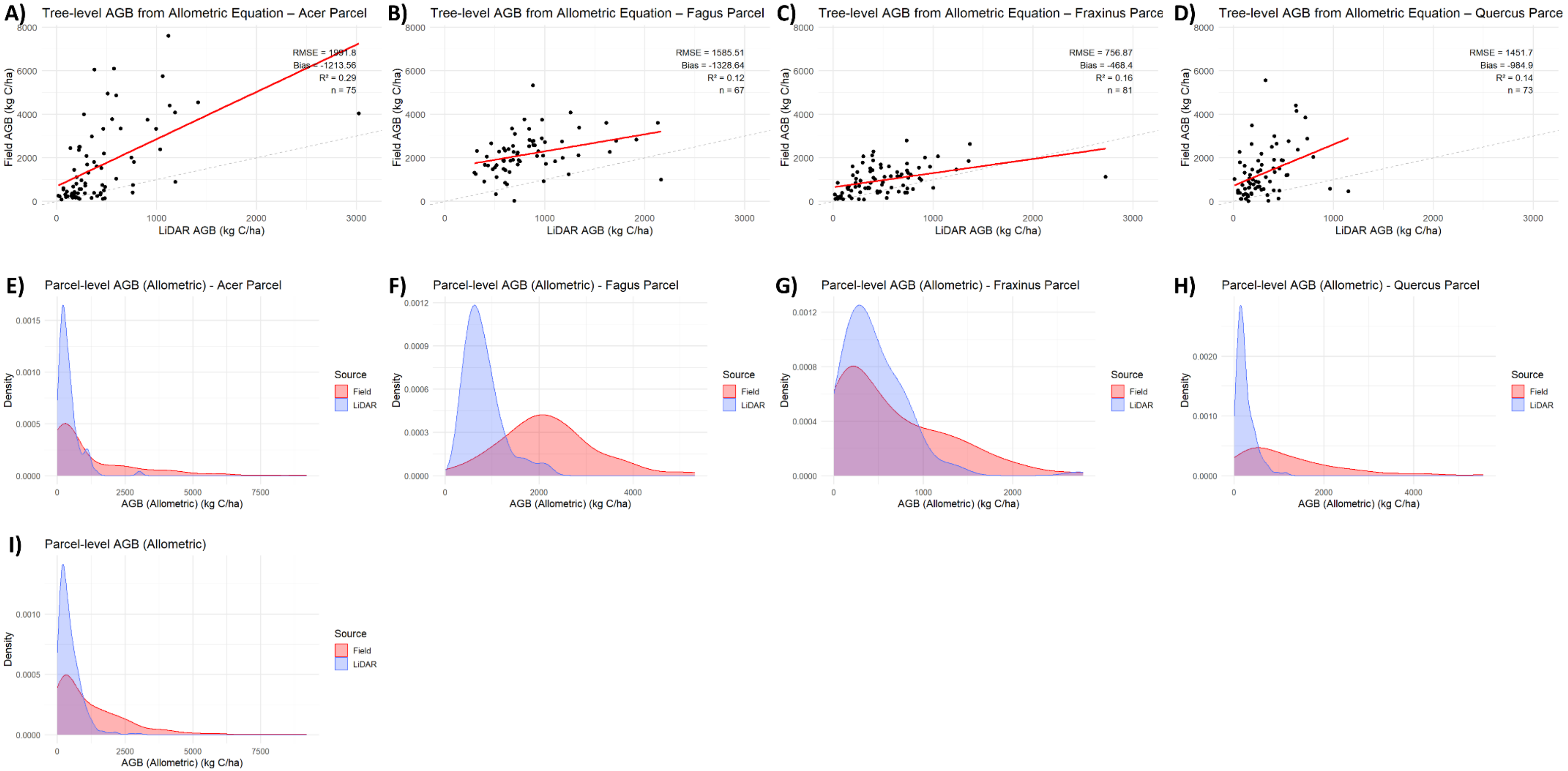


Figure A 5: Comparison between LiDAR-derived and field-measured AGB from allometric equation at tree and parcel levels. (A–D) Tree-level scatterplots per parcel (Acer, Fagus, Fraxinus, Quercus); red line shows the best-fit regression, black line indicates the 1:1 theoretical relationship. RMSE, bias, R², and sample size (n) are reported. (E–H) Parcel-level density plots comparing the distribution of field and LiDAR-derived values per parcel. (I) Combined parcel-level density plot comparing overall LiDAR and field distributions.

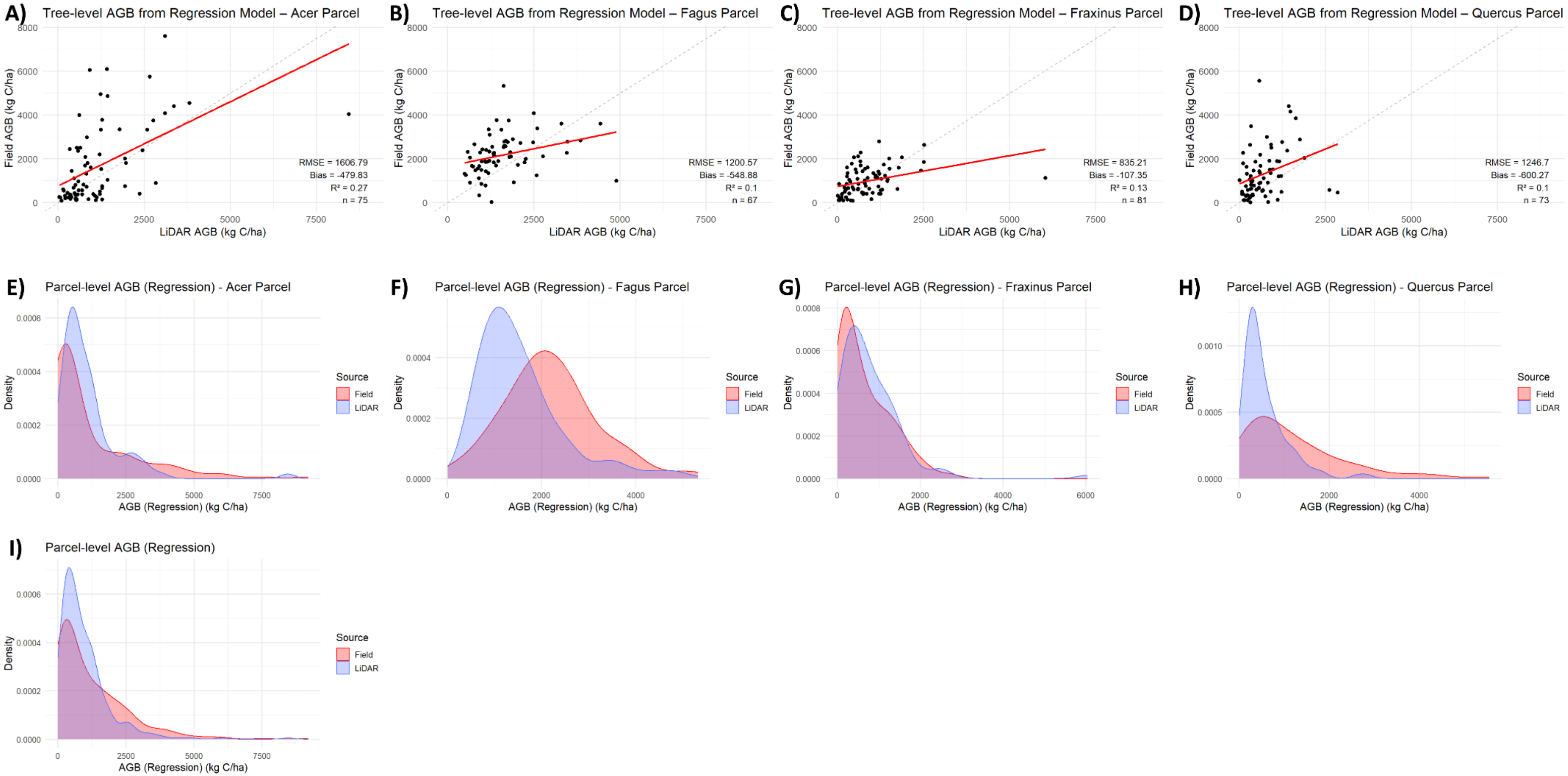


Figure A 6: Comparison between LiDAR-derived and field-measured AGB from regression model at tree and parcel levels. (A–D) Tree-level scatterplots per parcel (Acer, Fagus, Fraxinus, Quercus); red line shows the best-fit regression, black line indicates the 1:1 theoretical relationship. RMSE, bias, R², and sample size (n) are reported. (E–H) Parcel-level density plots comparing the distribution of field and LiDAR-derived values per parcel. (I) Combined parcel-level density plot comparing overall LiDAR and field distributions.

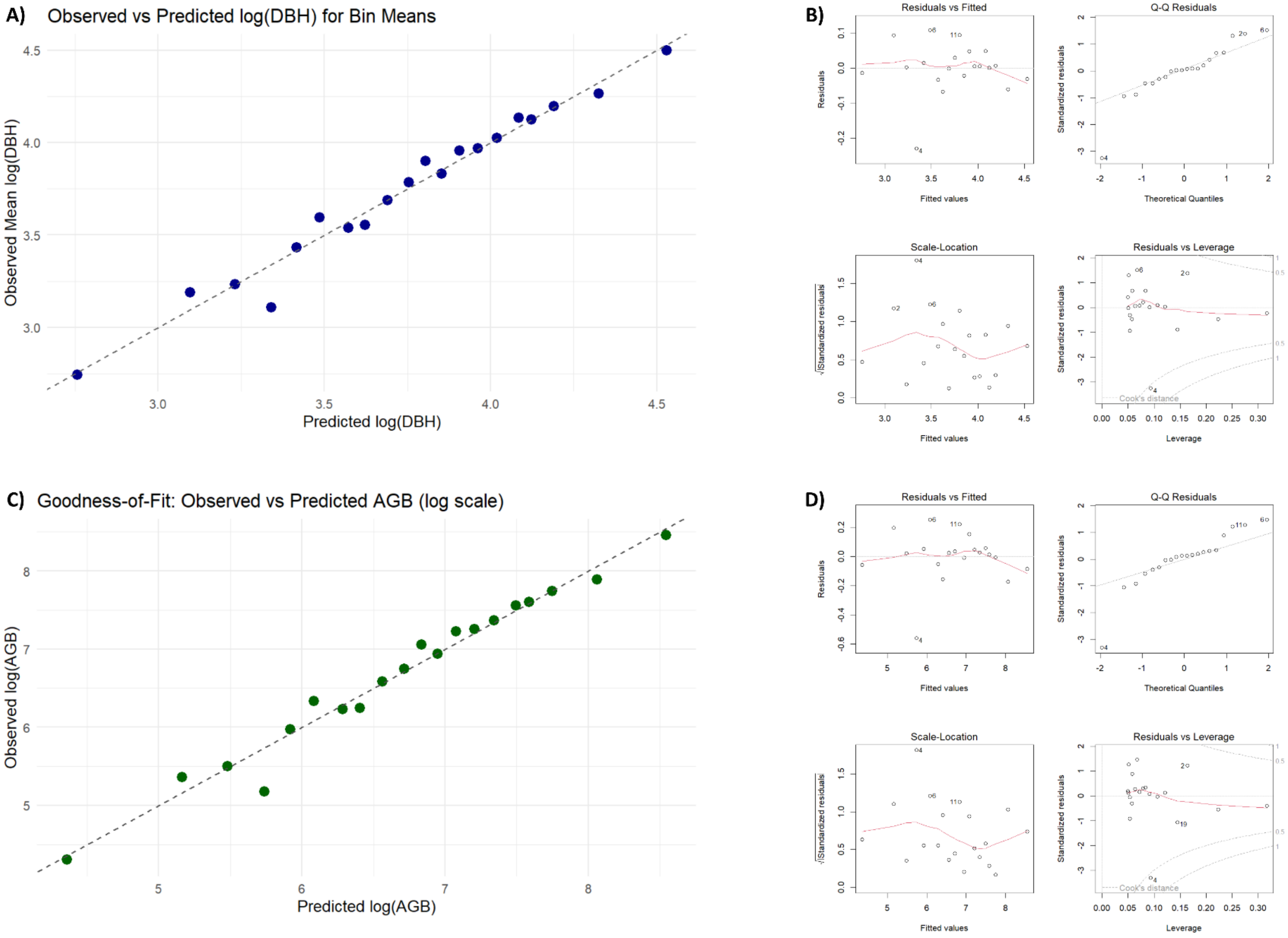


Figure A 7: Regression model evaluation for DBH (A–B) and AGB (C–D). (A and C) Goodness-of-fit plots comparing predicted and observed log-transformed values using binned data. (B and D) Standard regression diagnostics: Residuals vs. Fitted to assess non-linearity and homoscedasticity (top left), Normal Q-Q to check normality of residuals (top right), Scale-Location to examine spread of residuals (bottom left), Residuals vs. Leverage to identify outliers (bottom right).

Tables

Table A 1: Segmentation performance comparison between Novel and Benchmark segmentation methods per forest parcel. The introduced validation metrics: True Positives (TP), False Negatives (FN), False Positives (FP), Commission Error (CE), Omission Error (OE), Precision (P), Recall (R), and F1-score.

Parcel	**Segmentation Method**	**TP**	**FN**	**FP**	**CE (%)**	**OE (%)**	**P**	**R**	**F1-score**
Acer	Benchmark	49	88	3	5.8	64.2	0.94	0.36	0.52
	Novel	75	62	17	18.5	45.3	0.82	0.55	0.66
Fagus	Benchmark	52	28	5	8.8	35.0	0.91	0.65	0.76
	Novel	67	13	7	9.5	16.3	0.91	0.84	0.87
Fraxinus	Benchmark	77	76	18	19.0	49.7	0.81	0.50	0.62
	Novel	81	72	18	18.2	47.1	0.82	0.53	0.64
Quercus	Benchmark	73	9	62	45.9	11.0	0.54	0.89	0.67
	Novel	73	9	40	35.4	11.0	0.65	0.89	0.75

Table A 2: Parcel-level descriptive statistics for LiDAR- and field-derived tree metrics across all parcels. Metrics include tree height, crown base height (CBH), crown diameter, diameter at breast height (DBH), and aboveground biomass (AGB). AGB is shown as both allometrically calculated and regression-estimated values. Reported statistics are the mean, maximum, minimum, standard deviation (SD), and the 25th (P25), 50th (P50), and 75th (P75) percentiles.

Metric	Dataset	Mean	Max	Min	SD	P25	P50	P75
Tree Height (m)	LiDAR	29.78	39.50	13.85	4.39	26.33	30.49	33.11
Tree Height (m)	Field	30.46	55.90	9.00	6.97	24.50	30.50	35.80
CBH (m)	LiDAR	18.17	35.28	10.58	3.79	15.40	18.10	20.25
CBH (m)	Field	11.46	27.40	0.00	5.52	7.88	11.65	15.30
Crown Diameter (m)	LiDAR	9.20	25.95	0.66	3.72	6.54	8.91	11.36
Crown Diameter (m)	Field	9.94	23.19	1.30	3.78	7.03	9.56	12.00
DBH (cm)	LiDAR	41.34	114.48	6.45	15.94	29.52	39.91	50.85
DBH (cm)	Field	45.51	110.90	2.29	22.17	27.10	42.84	61.89
AGB (Allometric) (kg C/ha)	LiDAR	461.78	3026.86	4.39	402.98	178.71	354.87	642.80
AGB (Allometric) (kg C/ha)	Field	1202.68	9198.48	3.71	1287.00	239.09	747.61	1832.85
AGB (Regression) (kg C/ha)	LiDAR	949.63	8446.96	9.56	911.78	345.22	703.37	1245.55
AGB (Regression) (kg C/ha)	Field	1202.68	9198.48	3.71	1287.00	239.09	747.61	1832.85

Table A 3: Summary of accuracy metrics for LiDAR-derived structural variables compared to field data. Accuracy is reported using RMSE, bias, and R². Notes highlight trends in performance across forest conditions, particularly between Group 1 (irregular/mixed stands) and Group 2 (even-aged/pure stands), based on tree- and parcel-level comparisons.

Metric	RMSE	Bias	R^2	Notes on Parcel-level Trends / Forest Conditions
Tree Height	4.23 m	-1.60 m	0.58	Moderate accuracy; underestimated tall trees. / Group 2 (even-aged) performed better; more accurate distributions.
CBH	9.08 m	+6.59 m	0.01	Poor accuracy; overestimation and lack of low values. / No meaningful difference between Group 1 and 2.
Crown Diameter	4.45 m	-0.92 m	0.11	Strong agreement with field data at parcel level. / Group 2 slightly closer to field median; both groups performed well.
DBH	20.66 cm	-8.53 cm	0.24	Mixed results; general underestimation; narrow LiDAR distribution. / **Tree-level:** Group 2 better, **Parcel-level:** Group 1 better; closer mean and distribution.
AGB (allometric)	1500.24 kg C/ha	-979.30 kg C/ha	0.21	Fair accuracy; slight underestimation overall. / Group 1 slightly better; estimates closer to field values in dense stands.
AGB (regression)	1246.86 kg C/ha	-423.23 kg C/ha	0.22	Similar to allometric, slight understimation. **/** **Tree-level:** Group 2 better, **Parcel-level:** Group 1 better; better distributional match.

Application of LiDAR Data to Derive Key Forest Structure Metrics by **Orsolya Turi**

Abstract

This study investigates the accuracy of airborne LiDAR in deriving individual tree-level structural metrics relevant to forest management in temperate deciduous forests. Five key metrics—tree height, diameter at breast height (DBH), crown base height (CBH), crown diameter, and aboveground biomass (AGB)—were assessed across four structurally diverse parcels in the Bois de Lauzelle (Belgium), using both LiDAR and field data.

A novel individual tree segmentation (ITS) workflow was developed by extending an existing algorithm (PTrees). The method incorporates a Random Forest model to identify segmentation errors and applies an iterative correction process to address over- and undersegmentation. The workflow was benchmarked against a standard CHM-based method and evaluated using a custom in situ matching algorithm designed to handle spatial uncertainty in field data. The novel segmentation approach outperformed the benchmark in all parcels, reducing both omission and commission errors.

Structural metrics were extracted from the corrected segments and validated with in situ data. Results showed promising accuracy for tree height (RMSE = 4.23 m, $R^2 = 0.58$), though tall trees tended to be underestimated. Crown diameter showed low correlation ($R^2 = 0.11$) but aligned well with field distributions. DBH estimates had mixed results (RMSE = 20.66 cm, $R^2 = 0.24$), with a consistent underestimation trend. AGB, estimated via allometric equations and regression models, reached similar performance ($R^2 = 0.21$–0.22), with slight underestimation and varying accuracy across forest conditions. The only metric showing poor estimation was CBH (RMSE = 9.08 m, $R^2 = 0.01$), largely due to overestimation and a lack of low values.

The findings confirm that moderate-density airborne LiDAR, combined with calibrated segmentation and validation, can produce usable structural metrics. The method offers a scalable and transferable approach to forest inventory and monitoring, with strong potential to support data-driven forest management in structurally diverse temperate forests.

Keywords: LiDAR, Individual Tree Segmentation, Forest Structure, Temperate Deciduous Forest

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Faculté des bioingénieurs
Croix du Sud, 2 bte L7.05.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/agro**