Geo-information Science and Remote Sensing
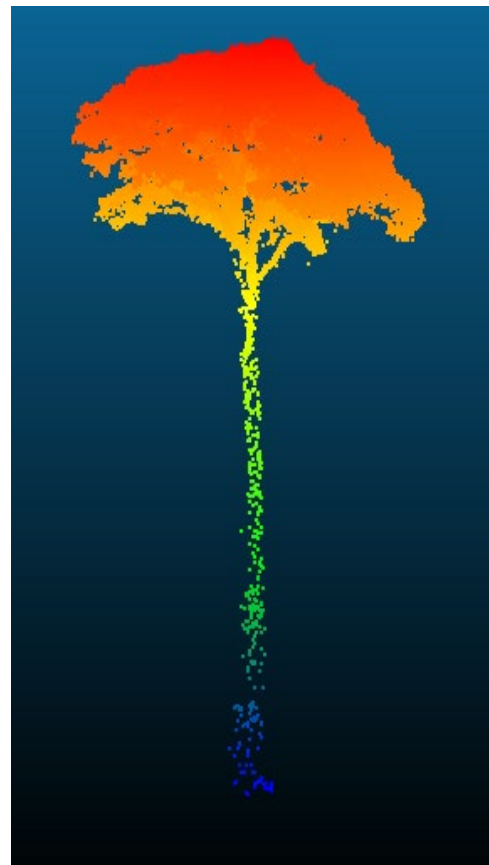
Thesis Report GIRS-2020-08

Assessment of tree segmentation algorithms based on UAV LiDAR data for Above Ground Biomass estimation

Leonardo Mauri

26 – 04 - 2020

WAGENINGEN
UNIVERSITY & RESEARCH

Assessment of tree segmentation algorithms on UAV LiDAR data for Above Ground Biomass estimation

Leonardo Mauri

Registration number 951020548010

<u>Supervisors</u>:

Benjamin Brede

A thesis submitted in partial fulfilment of the degree of Master of Science

at Wageningen University and Research Centre,

The Netherlands.

26 – 04 – 2020

Wageningen, The Netherlands

Abstract:

Above Ground Biomass (AGB) is recognized as a critical variable for understanding the global carbon cycle and represents a crucial feature for climate change policies. Remote Sensing techniques can provide consistent and large-scale estimation but need ground-measured estimates to calibrate and validate their algorithms. In this context, unmanned aerial vehicle LiDAR data (UAV-LS) can provide an accurate 3D representation of forest stands which can be exploited for the retrieval of AGB.

This study tested and compared three different segmentation algorithms on a tropical forest stand in French Guyana. The results were then validated with the stem map of the area and through a similarity analysis between identified crowns and crowns that were manually segmented from the LiDAR point cloud. Furthermore, Tree Height and Crown Diameter were extracted for each segmented tree and used to estimate AGB with the implementation of an allometric model. The obtained estimates were compared with reference estimates derived from a widely used global allometric equation.

The tested segmentation methods produced highly different numbers of trees and thus led to different AGB estimates. The AMS3D algorithm, the only one based on the whole point cloud, outperformed the other methods based on the Canopy Height Model for both tree matching and similarity analysis. The method also returned the most similar AGB estimates at plot level and for three of the four investigated DBH (Diameter at Breast Height) classes. Furthermore, AMS3D was also able to delineate understory features which are accurately represented in the UAV-LS data. This study demonstrates the poor result of CHM-based segmentation methods in tropical ecosystems and confirms the added value of processing the whole point cloud for accurately segmenting complex forest stands. The analyses also demonstrate the high-quality information contained in UAV-LS data that can be further exploited for a better characterization of forest structures.

Table of contents

List of abbreviations:

ABA: Area-based approaches
ACC: Accuracy
AGB: Aboveground biomass
AGBerror: Average Above Ground Biomass Error
AGBEst : Estimated Above Ground Biomass
AGBRef: Reference Above Ground Biomass
AGBRel_Error: Relative Above Ground Biomass Error
ALS: Airborne Laser Scanning
AMS3D: Adaptive Mean Shift 3D
AOI: area of study
CA: Crown Area
CD: Crown Diameter
CHM: Canopy Height Model
CHM_Diff: Difference CHM model (CHM_SF - CHM_St)
CHM_Sf: Spike-free Canopy Height Model
CHM_St: Standard Canopy Height Model
DBH: Diameter at breast height
ERR: Error Rate
EstTH: Estimated Tree Height
F1: F-score
FN: False Negative
FP: False Positive
Gamma: gamma value for the gaussian kernel
GNSS: Global Navigation Satellite System
ID: Identification code
ITC: Individual Tree Canopy
ITD: Individual tree-based approaches
LiDAR: Light Detection and Ranging
MaxDist: Maximum distance to merge two modes
MinMod: Minimum number of modes to create a cluster
PREC: Precision
QSM: Quantitative Structure Model
REC: Recall
RMSE: Root Mean Square Error
RSE: Residual Standard Error
TH: Tree Height
Thresh_D: Distance threshold
Thresh_H: Height threshold
TIN: Triangulate Irregular Network
TLS: Terrestrial Laser Scanning
TN: True Negative
TP: True Positive
UAV: Unmanned aerial vehicle
UAV-LS: Unmanned aerial vehicle- Laser Scanning
WD: Wood Density

Introduction

Forest ecosystems are valued globally for the services they provide to society (Pan et al., 2011). Besides wood-based products, which largely contribute to the world economy, other crucial services range from the conservation of soil, water and biological diversity, to the modulation of the carbon cycle (Wardle and Kaoneka, 1999). Furthermore, forests also influence climate through exchanges of energy, water, carbon dioxide and other chemical species, modulating complex interactions through which they can suppress or amplify anthropogenic climate change (Bonan, 2008). During their lifetime, forests can both store and emit large amounts of carbon which makes them a crucial variable for climate change mitigation and adaptation measures.

The importance of forest ecosystems, which cover about 30 % of our planet (Dalponte and Coomes, 2016), is widely recognized among the scientific community. As a result, many national and international programmes have been established to safely manage and protect all types of forests. One of the largest and most important, is the United Nations Collaborative Programme REDD+ which aims to reduce emission from forests through the implementation of the so-called payments for ecosystem services (PES), which promote the conservation of natural resources through economic incentives. Lately, to support the central role and conservation of such ecosystems, forest biomass has been recognized as a Global Climate Observing System (GCOS) Essential Climate Variable (ECV) and thus accurate biomass products are of great importance for forest management and climate mitigation (Duncanson et al. 2019). However, current knowledge of the distribution of biomass across the globe is limited and consistent data on changes over time at both the global and regional levels are even less available (Herold et al., 2019).

Tropical forests are among the most structurally complex and carbon-rich ecosystems in the world, and they store 40% of the total carbon stocked in forests globally(Gibbs et al., 2007). Although a number of forest inventories have been carried out here, there remain large areas in the tropics where such inventories are out of date, incomplete, or entirely lacking (Houghton, 2005). Consequently, the accurate estimation of structural characteristics (e.g., aboveground biomass, AGB) of tropical vegetation remains a major obstacle (Dubayah et al., 1997), resulting in large uncertainties. Furthermore, tropical forests play also a crucial role in climate change policies, with atmospheric analyses that define these ecosystems as carbon neutral or carbon sinks, highlighting their vulnerability to a warmer, drier climate, which may exacerbate global warming through a positive feedback (Bonan, 2008). For all the mentioned reasons, there is an urgent need for improved data sets that characterize the global distribution of AGB, especially in the tropics (Baccini et al. 2012).

Problem definition

AGB is defined as all biomass of living vegetation above the soil, and includes stems, stumps, branches, bark, seeds, and foliage (Eggleston et al., 2006). So far, different techniques have been developed to estimate it, each one presenting its own limitations in accuracy or feasibility.

True biomass measurements require the physical, destructive harvesting, drying, and weighing of trees. The process is extremely difficult in practice, and typically undesirable or in many cases logistically impossible (Duncanson et al. 2019). In general, the overall difficulty of acquiring in-situ measurements strongly highlights the need of new and advanced approaches to accurately measure AGB.

A common technique to estimate biomass at large scale is through the implementation of allometric equations. These models rely on empirical relationships between total tree biomass and vegetation attributes, among which, tree height (TH) and tree diameter at breast height (DBH) are the most frequently used. However, the quality of allometric models represents one of the most important limitations in assessing AGB (Chave et al. 2014). In fact, even if allometric equations have improved in recent years, they remain limited and they cannot always be applied to different regions. (Lau et al. 2019). These models, which also need to be supported by calibration data, require direct estimates of AGB and thus still involve destructive sampling of trees (Calders et al. 2015). Another important limitation to the final reliability of allometric models is caused by the presence of big trees, which store large quantities of carbon (Slik et al., 2013) and can therefore strongly affect the final estimation bias. In fact, due to the model's dependency on field sampling and the complexity of harvesting large trees, these are usually underrepresented. As a consequence, the resulting allometric reliability degrades as the DBH increases, causing a strong discrepancy in model estimates for large trees (Chave et al., 2001).

To overcome this limitation, researchers have explored different relationships at tree level which may allow to estimate biomass without the need of DBH measures. Among the tree attributes that have been tested so far, (Hemery et al., 2005) have found crown diameter (CD) and DBH to be strongly coupled, even in large trees, leading to new possibilities for estimating AGB. Based on this information Jucker et al. (2017) decided to compile a global dataset recording for each tree, DBH, CD, and TH. With the aim of developing new allometric models, the research group also collected AGB measurements by harvesting a portion of the inspected trees. The group showed that AGB was strongly related to TH and CD, developing a new allometric equation based solely on these parameters which was proved to estimate AGB with low Root Mean Square Error and bias values (RMSE = 1.70 Mg, bias = -4.3%). The discovery led to new possibilities in the field of forest biomass estimation, filling a gap between allometric equations and remote sensing techniques which in many cases are not able to measure DBHs.

Remote sensing is the science and practice of acquiring information about an object without actually coming into contact with it (Horning, 2008). In recent years, with development of environmental information science, remote sensing data have played an important role in many research fields (Liu, 2015). This is also the case of forestry, where remote sensing offers an efficient and economical mean for AGB monitoring by facilitating forest type and canopy density stratification (Tsitsi, 2016). The benefit of such techniques is confirmed by the numerous diverse technologies that can be exploited for biomass estimation, and by the level of accuracy that can be achieved.

Acknowledging the urgent need for improved, large scale, biomass estimations the European Space Agency is developing the first space system entirely conceived for AGB monitoring. The BIOMASS mission, whose launch is planned for 2022, will make use of P-band synthetic aperture radar (SAR) measurements to provide crucial information about the state of our forests (ESA,

n.d.). Nevertheless, it is essential that satellite missions are able to make use of accurate ground estimates of carbon stocks to train their inversion algorithms and validate their products (Duncanson et al., 2019). Although field measurements are expensive and time consuming, they are essential to provide the basis for extending local measurements to larger areas using remote sensing approaches (Zolkos et al., 2013). Among all remote sensing techniques, Light Detection And Ranging (LiDAR) represents a valuable alternative to destructive field measurements to provide plot-level AGB estimates in order to calibrate and validate large scale products. LiDAR has become a valuable source of information to assess vegetation canopy structure, in particular for complex forest canopies that limit manual and destructive sampling (Brede et al. 2017). LiDAR remote sensing is an active technique that calculates the distance of an object by recording the time during which the emitted laser pulse is reflected back to the sensor. In recent years, LiDAR has emerged as a valuable technology because of the directness of its approach to measuring canopy structure and relatively higher accuracies reported in the literature when compared to optical and radar approaches (Zolkos et al., 2013). Furthermore, several studies have found a strong correlation between LiDAR metrics and AGB (Drake et al., 2003) confirming its applicability for biomass estimation.

Three different categories of LiDAR technology can be distinguished based on the platform on which the sensors are installed: terrestrial laser scanning (TLS), airborne laser scanning (ALS) and recently also unmanned aerial vehicle laser scanning (UAV-LS). So far, TLS produces the most promising results, showing high accordance with destructive biomass measurements (Calders et al., 2015; Gonzalez de Tanago et al., 2018). The level of detail contained in TLS point clouds can be exploited in multiple ways to retrieve trees AGB. One approach makes use of the reliable extraction of tree parameters (DBH, TH, and CD), and their subsequent application for the development of new allometric models (Lau et al., 2019; Vaglio Laurin et al., 2019). Another commonly used technique estimates tree volume with the use of quantitative structure models (QSM) instead, and then converts volume to using basic density information (Calders et al., 2015; Gonzalez de Tanago et al., 2018).

On the other hand, ALS data have been widely used for generation of bare earth digital terrain models (DTM), as well as the estimation of forest inventory attributes (Hyyppä et al., 2008). ALS has also been proven to provide the required spatial detail and accuracy across large areas and a diverse range of forest types (White et al., 2016). In addition to that, recent technical advances in laser scanning made ALS one of the most promising technologies for the retrieval of forest information at different levels: individual tree, plot, and nationwide (Yu et al., 2010).

Nevertheless, the time consuming and expensive nature of TLS data acquisition, and the relatively low point density achievable with ALS systems (Brede et al., 2017) is causing scientists to look for valid alternatives, such as UAV-LS. As a matter of fact, recent improvements in small-scale technology have enabled the use of UAVs as an alternative remote sensing platform, offering a distinctive combination of very high resolution data captured at a significantly lower survey cost.(Wallace et al., 2012).

At present, there are two main techniques of LiDAR forest inventory that can be applied to ALS and UAV-LS data. Area-based approaches (ABA) retrieve forests structure at stand level, whereas individual tree-based approaches (ITD) extract tree parameters by individual tree detecting and modelling (Wang et al., 2019). ABA models are state of the art when it comes to estimating biomass from ALS data (Coomes et al., 2017), and they have been proven to produce reliable and consistent results (Wallace et al., 2014). When compared to tree-centric, ABA work better in estimating basal area and carbon stocks in tropical forests, probably due to strong limitations of the former methods that first need to identify and delineate individual trees (Aubry-Kientz et al., 2019). Previous studies have found the algorithm employed to be the main factor affecting the

accuracy of tree detection and delineation. Conversely, other studies suggested the accurate delineation of trees to be highly dependent on the properties of the forest, including stand density and spatial pattern of the trees (Wallace et al., 2014).

However, rapid technical advances in laser scanning has led to an increase in data point density which is expected to improve the accuracy of the ITD techniques (Yu et al., 2010). Coomes et al. (2017) presented several arguments supporting the idea that ITD approaches should be preferred over AGB area-based estimates. The first benefit of ITD methods is represented by the strong fundamental basis, conceptually similar to allometric equations, which these models are built upon. In addition, extracting and estimating AGB at tree level allows the mapping of small, high-value areas, the tracking of tree growth and death, and it finally reduces plot-dependent uncertainty. The promising nature of individual tree detection methods is being supported by the numerous recent studies that developed, tested, and compared tree-segmentation methods and approaches in diverse biomes and with different datatypes (Dalponte and Coomes, 2016; Ferraz et al., 2016; Wallace et al., 2014; Ferraz et al., 2012). Moreover, the fast and constant development of both high precision sensors and unmanned aerial vehicles will continue to provide high-quality data, which could further boost the application of tree-centric techniques for forestry applications as already suggested by the study conducted by Wallace et al. (2014) over an eucalyptus plantation. However, the added value of UAV high density point clouds, when implemented in existing segmentation algorithm, has yet to be further explored and proved, especially in dense tropical forests which could result in a massive number of returns. Finally, while ITD methods have already been widely tested in boreal (Dalponte et al., 2018; Maltamo et al., 2009; Paris et al., 2016; Persson et al., 2002) and temperate forests (Dalponte and Coomes, 2016; Koch et al., 2006; Popescu and Wynne, 2004), the accuracy and reliability of these algorithms in tropical regions have not extensively been examined yet, especially when aiming at tree attributes extraction for AGB estimation.

Research objectives and Research questions

The project explores the potential of individual tree segmentation techniques, applied on high density LiDAR data acquired by an UAV platform, for estimating AGB at plot level. In the project three different segmentation algorithms will be tested and their output will be used to estimate AGB. The segmentation results and AGB estimates will then be compared in order to evaluate which method performs the best in the area of study. This will be assessed through the following Research Questions:

RQ1: What tree segmentation algorithm delivers the most accurate segmentation of the area of study?

- **SRQ1**: What tree segmentation method returns the best statistics (Error Rate, Accuracy, Precision, Recall, and F-score) when comparing segmented trees with the stem map?

- **SRQ2**: What tree segmentation method returns the most similar tree crown shapes when compared to UAV-derived, manually segmented crowns?

RQ2: What tree segmentation algorithm returns the most similar AGB estimates to state-of-the-art ground-based estimates, at plot level and at different DBH classes, compared to reference AGB values in terms of RMSE and Bias?

Methodology

In the following section the various steps undertaken during the analysis are described. Throughout the project several software were used, namely: R language (Version 3.6.1; R Core Team, 2019) together with the Rstudio Integrated Development Environment (Version 1.2.1335; Rstudio Team, 2018), LAStools (Version 181108; Isenburg, M., 2018), ArcGIS (Version 10.7.1; ESRI, 2019), CloudCompare (Version 2.10.2 (Zephyrus); 2014), and Computree platform (Version 5.0; Computree Group, 2019).

Study area
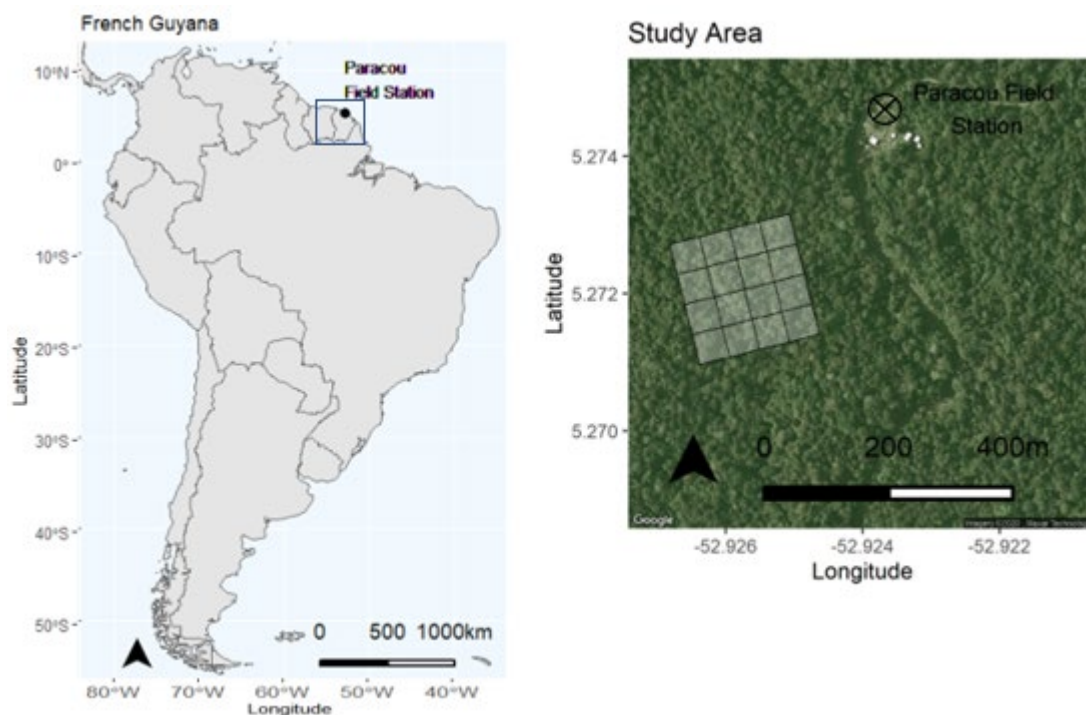
Figure 1: Map of South America highlighting French Guyana (black box) and the location of the study area (black dot), (b) Study area site, with indicated Paracou Field Station (cross), and South-East section of plot 6 (black-white polygon).

The study area is part of Paracou Experimental Station (https://paracou.cirad.fr/website), in the coastal part of French Guyana (Figure 1). The station, at N5°18' and W52°53' is approximately 100 km North-West of the capital Cayenne. The climate of the site is equatorial, characterized by two main climatic periods determined by the movement of the Intertropical Convergence Zone, a dry season from mid-August to mid-November and a rainy season, which is often interrupted by a short drier period between March and April. The mean annual rainfall recorded near Paracou station from 1979 to 2001 is 3,041 mm with a minimum during the month of September and a maximum in May. The mean annual temperature is 26 °C. The core area is covered with a lowland, old growth, tropical rain forest with an exceptional richness (over 750 woody species). According to previous studies the five most abundant families at the site are the *Lecythidaceae*, Fabaceae, *Chrysobalanaceae*, *Sapotaceae* and *Annonaceae* (Aubry-Kientz et al., 2019). A series of 16 permanent plots (fifteen with a surface of 6.25 ha plus one of 25 ha)

have been censused (DBH>10 cm) every 1-2 years since 1984. Nine of the plots were logged and subjected to human-induced disturbance in 1986.

Among the 16 plots, the object of the study will be a portion of Plot 6, which has not been subjected to any disturbance and it is set as control plot.

Data acquisition

To perform the analysis different types of data were implemented. The scientific organization of *CIRAD,* which manages Paracou Station, kindly provided the plots and subplots boundaries files. Furthermore, they also shared the stem map of plot 6 which contains location and additional information of each tree with a DBH greater than 10 cm. Plot corners were measured with centimetric accuracy using a total station, trees were then positioned within subplots (20 x 20 m) using a measuring tape. The location of new trees was later estimated based on their distance from neighbouring stems. All trees, which were located by cartesian coordinates of their trunks, have an estimates precision of +/- 2 m (Aubry-Kientz et al., 2019).

In addition to that, UAV-LiDAR data was acquired by flying the RIEGL RiCOPTER with VUX®-1UAV, an integrated UAV and LiDAR sensor system (Brede et al., 2017). The flight campaign was carried out between the 18th and the 20th of October 2019 and consisted of several flights. The cruise speed varied between 6 and 8 m/s and the flight height was set between 105 and 120 m above the ground. In total seven flights were performed over the full study area, of which six were successfully accomplished. Three flights were also conducted over a smaller portion of plot 6 but will not be included in the study. In order to acquire a more detailed and complete representation of the forest, the flights were performed following different patterns so that the laser beams could hit the vegetation features from several directions and better penetrate the canopies. In addition to change the principal flight direction, the pulse repetition of the laser scanner, which influences the laser penetration and number of returns, was also varied. A brief summary of all flights can be found in Table 1.

Table 1: For each of the performed flights, principal flight pattern, height above ground level (agl), average flight speed and scanner pulse repetition

Principal flight pattern	Height agl (m)	Speed (m/s)	Pulse Repetition Rate (kHz)
N-S	110	6	550
E-W	105	7	550
NW-SE	115	6	550
NE-SW	105	6	550
N-S	120	6	300
N-S	120	6	100

Data Preprocessing

■ *Field data*
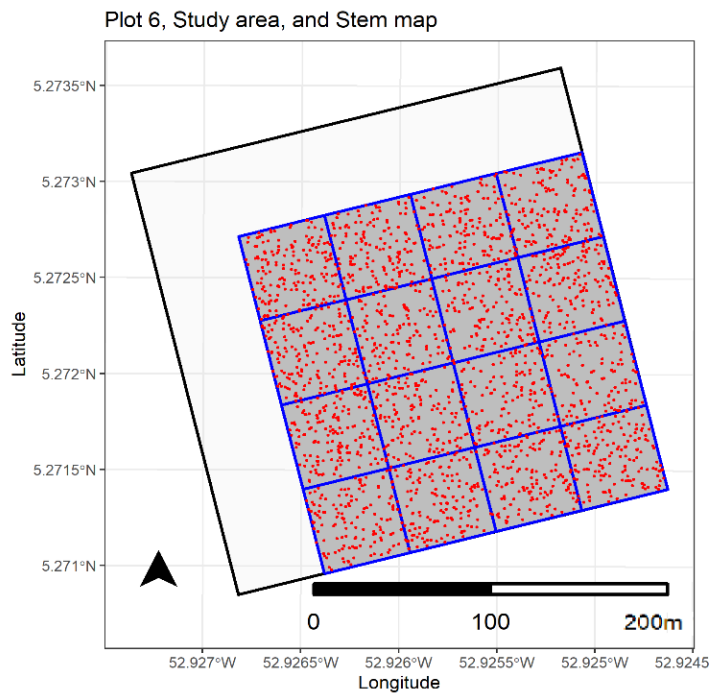
Figure 2: Boundaries of original Plot 6 (black), boundaries and quarter hectare sections of Area Of Interest (AOI, blue), stem map of AOI (red dots)

The analysis began with the preprocessing of the field data, namely the plot boundaries and the stem map. These data needed to undergo a series of steps in order to select only the required information corresponding to the area of study (AOI). After selecting plot 6 form the plots boundaries dataset, a 200 m x 200 m section in the South-East corner of the plot 6 was extracted and stems within the newly created plot were selected from the original stem map, Figure 2. To preserve the true shape of those trees whose crowns exceeded the plot boundaries each side of the AOI was increased with additional 10 m, returning a final square plot of 4.41 ha. To conclude the preprocessing of field data, the circumference measurements of each stem were converted to DBH values which is the commonly used measure to describe stem size.

As a result, the obtained stem map subset, based on the new AOI, returned a dataset of 2291 stems (Figure 2), summary statistics and distribution of DBH and Estimated Tree Height values can be found in Table 2 and Figure 3. Observing the histograms in Figure 3, it is clear that trees with small DBH are much more frequent than big trees with the most numerous class corresponding to tree with a DBH between 10 and 15 cm containing almost 60 % of the total stems. The count then constantly decreases registering only 62 trees with a DBH larger than 60 cm. The distribution of estimated height followed the same pattern with almost 60 % of the total trees estimated to be shorter than 20 m, which also corresponds to the mean value. It is important to notice that even though the first set of trees represented in the graph (12 m < TH < 14 m) shows a much lower number of observations, it is perfectly in line with the other data. In fact, the only reason for such a smaller count is due to the narrower range of estimated tree height which only extends from 13.47 m to 14 m.

Table 2: Number of stems in the Area Of Study and minimum, maximum, and mean value of Diameter at Breast Height (DBH), and Estimated Tree Height (EstTH)

N° stems	Min DBH (cm)	Mean DBH (cm)	Max DBH (cm)	Min EstTH (m)	Mean EstTH (m)	Max EstTH (m)
1291	10.03	22.58	101.38	13.47	20.52	45.34
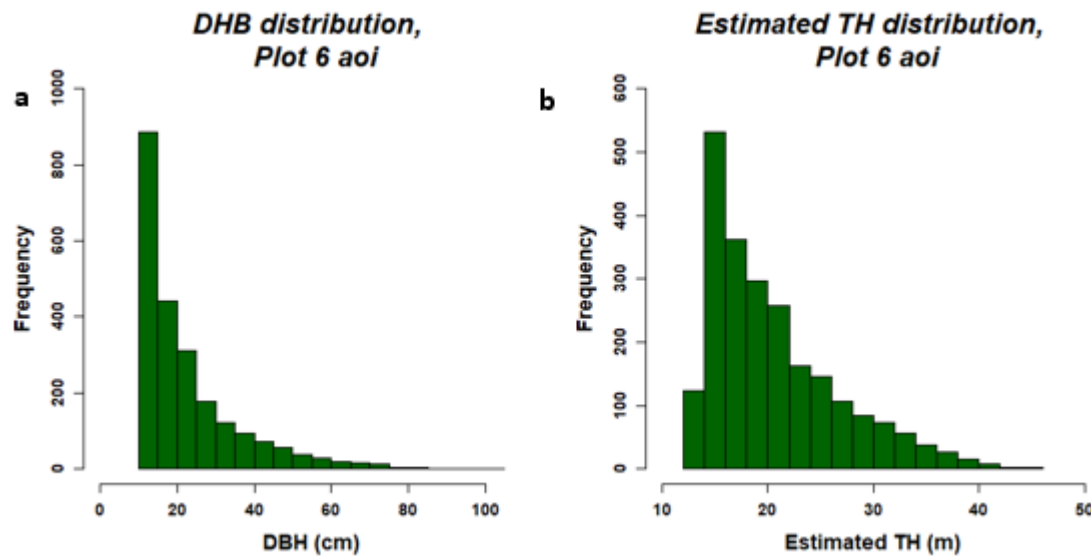
Figure 3: Diameter at Breast Height (DBH) and Estimated Tree Height (Estimated TH) distributions for stems in the Area Of Interest (AOI).

▪ *UAV data*

After the flight campaign the raw UAV data required to undergo a series of operations in order to obtain georeferenced point clouds which could be combined and exploited during the analysis.

The acquired UAV data were pre-processed by Benjamin Brede, a member of the UAV-LS campaign, who provided them in their final, ready to use version. For clarity a short description of these operations is provided.

In order to be implemented in the analysis the raw LiDAR data first needed to be converted into point clouds in RIEGL's *RiPROCESS* software which comes together with the RIEGL *UAV RiCOPTER* system. During this operation the targets are detected and Multiple Time Around Range ambiguities, which occur when one pulse is emitted before the previous one returns, are resolved. In order to align and match point clouds from different flight lines, the flight trajectories, need to be aligned and corrected with external GNSS data. The trajectory preprocessing was carried out in *POSpac Mobile Mapping Suite (V7.2)* software with GNSS data from a nearby permanent reference station, about 13 km away. Once the data were corrected and georeferenced, atmospheric noise, usually detected as unexpected returns close to the scanner, was removed and finally, the data were exported as LAZ files. All data are projected in Universal Transverse Mercator (UTM) system, zone 22N, EPSG: 32622.

The processed dataset composed of 70 point clouds, was then further processed to obtain the final point cloud data as described below.

- *Point cloud clipping*

The first operation performed on the georeferenced point clouds was intended to select and extract only the portion of data needed in the analysis. In order to do so each point cloud was clipped on the study area extent (200 m x 200 m SW section of plot 6). In addition to discard the unnecessary portion of data this operation also allowed to reduce the size of files, shortening the computation time of the following operations. With the same goal, after clipping the single flight line point clouds, the data were also divided into tiles of 50 m x 50 m. The tiling operation is very common when working with a large amount of points and it is also the conventional way to store and distribute point cloud data.

- *Data quality*

After having clipped the point clouds, the analysis proceeded with the data quality assessment. For each flight, the relative tiles were merged, and the points were classified as ground or non-ground. After the classification, ground points were used to create the Digital Elevation Model (DEM) of the area which represents the surface of the terrain. In addition to that, the Digital Surface Model (DSM), representation of all features on the ground, was also created making use of all points. The generation of these models was performed in order to compare the alignment of different flights and assess whether the co-registration between flight lines, and the georeferencing with the external GNSS data were successful. Finally, the point clouds of every flight were also compared with each other. The comparison was performed in *CloudCompare software (*Version 2.10.2 (Zephyrus); 2014) by measuring the distance between two point clouds at a time.

- *Final point cloud generation*

Once the alignment between point clouds of different flights was established, it was possible to proceed with the generation of the final point cloud. The first step consisted in combining tiles of all flights together that were then merged creating the final point cloud of the whole area of study. The newly created dataset was now composed of points acquired during all the different flights and thus presented a higher point density. Merging multiple point clouds can result in duplicate points which are points with the same exact coordinates (X, Y, Z). Since these points contained redundant information there was no need to keep them both, thus they were searched and one of the duplicates was eliminated. Finally, the newly created, full density, point cloud was exploited to generate the DEM and DSM of the AOI.

The reason behind the choice of merging the different point clouds was to exploit the information contained in each one of them to obtain the most detailed and accurate representation of vegetation features. In fact, each flight scanned the scene along only one direction and thus some of the objects might have not been hit by the laser pulses or have been hit only few times, producing a small number of returns that don't accurately reproduce the real shape of tree features. Exploiting point clouds acquired along several

flight trajectories increased the chance that objects were hit by enough points, returning a more faithful representation of submerged vegetation, now composed of a larger amount of points.

UAV Data Down-sampling and Height Normalization

Before proceeding with the analysis, the point cloud needed to be normalized to the ground. This step was required in order to compute the Canopy Height Model (CHM) and to be used as input of the AMS3D segmentation. Normalizing a point cloud means that the Z values, which now contain the absolute points height, have to be recalculated in respect to the ground. For a successful and accurate normalization, it is required to have a sufficient number of ground points so that a reliable ground surface can be constructed through TIN generation. Due to the limit size and thus ground points number, some tiles firstly had to be merged so that the normalization process could be successfully finalized. Furthermore, due to the large number of points contained in each tile, these could not be directly exploited to generate the CHM nor used as input of the AMS3D algorithm. In the first case the tool implemented to compute the spike free version of the CHM could not handle such high amount of points whereas the AMS3D algorithm was found to be so time consuming that segmenting the point cloud at its full density was not feasible. As a result the point density of the obtained tiles had to be decreased discarding some of the points. To better understand the thinning procedure, a brief description of the implemented tool is now provided. *Lasthin,* part of *LAStools* software *(Version 181108; Isenburg, M., 2018)*, works constructing a regular grid over the 3D point cloud and successively keeping one point in each grid cell, which was randomly selected. The level of down sampling is therefore uniquely controlled by the resolution of the grid, with larger cells that will result in lower point densities. The final cell size to be used was chosen after the assessment of different point densities on the AMS3D segmentation output (further details can be found in the relative section, AMS3D testing:).

Adaptive Mean Shift 3D segmentation

The first segmentation method to be tested was the Adaptive Mean Shift 3D or AMS3D (Ferraz et al., 2016). The algorithm takes a normalized point cloud as input and it segments the scene into individual point clusters which correspond to single trees. The version used in the study, implemented in the processing platform *Computree* (Version 5.0; Computree Group, 2019), where it is freely available (http://computree.onf.fr/?lang=en), will be briefly described in this paragraph. For a more detailed description of the method and its theoretical assumptions it is suggested to consult the works of Ferraz et al. (2012) and Ferraz et al. (2016).

The AMS3D algorithm, processing the entire point cloud, is capable of segmenting trees within every forest layer.

The feature that allows the algorithm to accurately segment and delineate trees of different shapes and sizes is represented by the so-called bandwidth model. As developed on *Computree* platform (Version 5.0; Computree Group, 2019), the model is based on allometric relationships between trees crown size and tree height, which determine the size of segmented structures.

The segmentation algorithm works in a two-step approach that assumes the point cloud to be a multi-modal distribution where each mode, defined as a local maximum in both density and

height, corresponds to a single location within an individual crown (Ferraz et al., 2016). During the first step, the algorithm, constructs a 3D-kernel (cylinder) around each point that iteratively moves towards dense regions until it converges to the local mode.

To do so, the authors designed a 3-D kernel profile as the products of two profiles, Figure 4, a Gaussian one for the horizontal dimension which searches for the local density maxima, and an Epanechnikov profile (vertical dimension) that deals with the local height maxima (Ferraz et al., 2012). To make the method more robust, the weighted distance is calculated as the product between the two kernels. As a result, the method can separate both horizontally and vertically interconnected crowns. In the first case, whereas the gaussian profile converge to the area where the two crowns are both present (high point density) the vertical profile force the algorithm to move towards the single treetops. Conversely, in the case of vertically connected crowns is the gaussian kernel which stabilize on the undercanopy tree apex, represented by a denser point region compared to the crown base of the higher tree (Ferraz et al., 2012). In the second step of the segmentation, all the points that converged towards the same mode, are now gathered together in 3D clusters which correspond to individual trees.

The definition of the kernel bandwidth has to be carefully evaluated since it strongly impacts the segmentation results. A small cylinder will result in the identification of several distinct modes (local basins of attraction), while a large one will aggregate small vegetation structures into larger ones (large basins of attraction) (Ferraz et al., 2012), causing respectively over or under-segmentation of the scene.

In the first version of the ASM3D, different bandwidth models where computed for different vegetation strata, with larger bandwidth for thicker layers. Unfortunately, this approach is not very suitable for tropical forest where the structural complexity does not make it possible to distinguish different forest layers. For this reason, a more advanced bandwidth model is required. Ferraz et al. (2016) developed a new model based on tree allometry equations based on a sample of large emergent trees. The 3D kernel dimension will then be calculated based on the linear regression models obtained for Crown Width and Tree Height for the horizontal dimension and between Crown Depth and Tree Height for the vertical dimension. For this research, there wasn't any need to calculate the allometric models, specific values of the intercept and slope of the bandwidth model were already calculated for Paracou Forest and they are provided on the AMS3D webpage (http://rdinnovation.onf.fr/projects/plugin-ams3d/wiki, personal account needed).
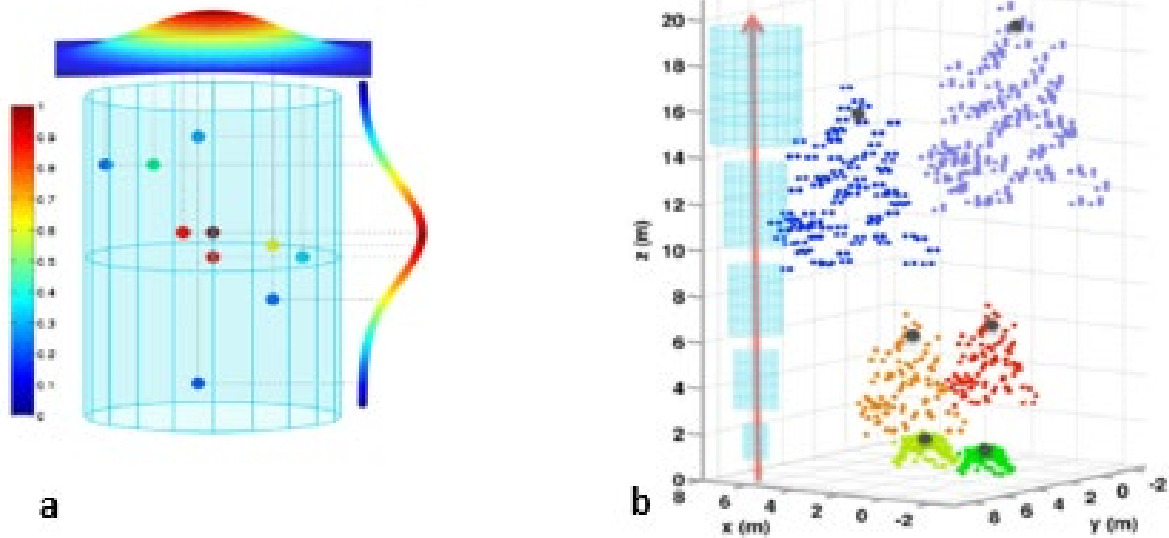
Figure 4: a) Horizontal (Gaussian profile, surface) and vertical (Epanechnikov profile, curve) kernel profiles. The point and colour bar indicates the weight of each point for the calculation of the kernel barycentre (Ferraz et al., 2012). b) Adaptive Mean Shift 3D segmentation of a simulated forest scene using a cylindrical kernel whose diameter and height are defined by the bandwidth model that depends on the lidar point's height (Ferraz et al., 2016)

To carry out the AMS3D segmentation, as implemented in *Computree* platform (Version 5.0; Computree Group, 2019), the user needs to manually set a list of parameters which are listed below:

− Intercept1 and coefficient1: values for the bandwidth diameter model. A coefficient value of about 0.2 is recommended for the area of study (Paracou, French Guyana). The intercept should be set to 0 so that for height equals to 0 the crown diameter is also 0.

− Intercept2 and coefficient2: values for the bandwidth height model. An intercept value of about 0.15 is recommended for the area of study (Paracou, French Guyana). As in the previous case, the intercept should be set to 0.

− Minimum cylinder diameter and height: these values can be kept small.

− Gamma for the Gaussian kernel: parameter to weight the points in the neighbour area. It was set to five in Ferraz et al. (2012).

− Maximum number of iterations: number of iterations to compute before moving to the next point, it has to be high enough to reach the mode convergence.

− Resolution of the grid for neighbouring points: size of the 3D grid used to search for neighbouring points.

− Resolution of the grid for neighbouring modes: size of the 3D grid used to search for neighbouring modes.

− Maximum distance between two modes to merge them: distance threshold for merging modes, first parameter of DBSCAN algorithm (used to merge the modes in the Computree version of AMS3D).

− Minimum number of modes to create a cluster: minimum number of points to form a single crown, second parameter of DBSCAN.

Here a brief guide on how to import, set, and run the AMS3D algorithm is presented, three steps need to be set, Figure 5:

- Point cloud loading: In the Steps window, select Load submenu, Points, Points, format LAS(V2), and select the LAS file.
- AMS3D segmentation: select Points submenu, Detect (Crown), and AMS3D MeanShift v2, set parameters for segmentation.
- Output file exporting: Select Export submenu, Points, Points + item, ASCII(X,Y,Z), then in the new window select the segmentation output, tick the boxes of both Group items, and finally tick the cluster (point scene) box.
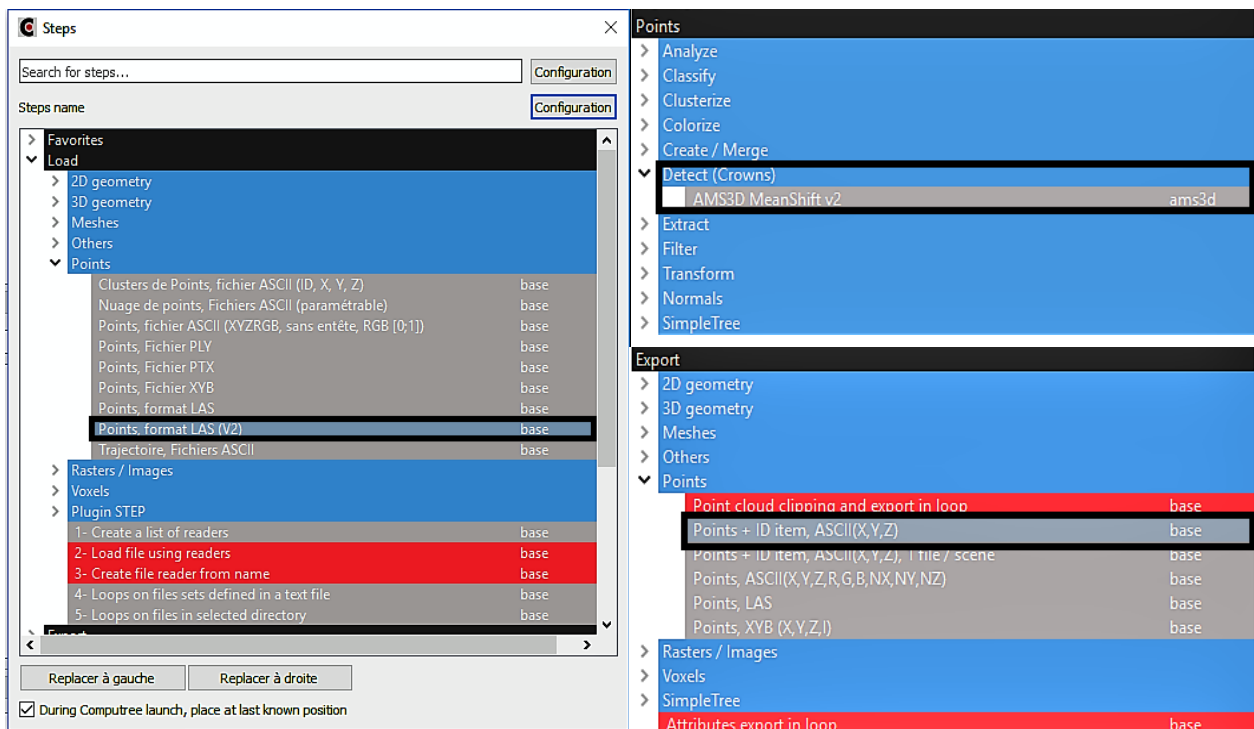

Figure 5: Illustration of required steps (surroundded in black) for running the Adaptive Mean Shift 3D segmentation on Computree platform.

- *Adaptive Mean Shift 3D Testing*
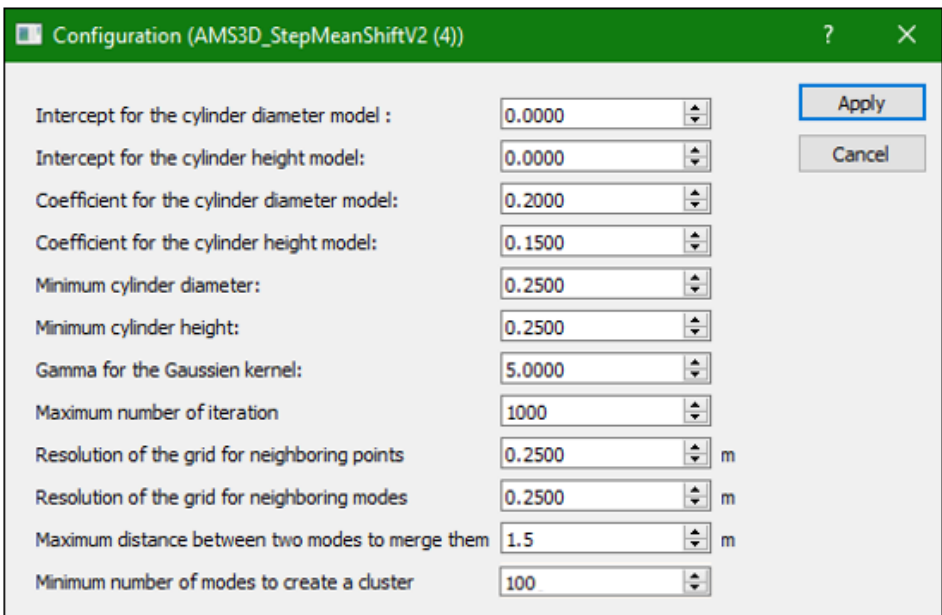
With the aim of finding the most appropriate parameters setting for the AMS3D segmentation, the algorithm was first run on a test tile experimenting with different values. After running preliminary tests (Tests 1-10) and visually assessing the results, three parameters, were identified to have an important influence on the results and were thus chosen for further testing (Tests 11-18). The analysis focused on the effect of the Gamma for the Gaussian kernel (Gamma), the maximum distance between two modes to merge them (MaxDist), and the minimum number of modes to create a cluster (MinMod). Table 3 presents an overview of the settings that were investigated. Furthermore, to inspect the influence that point density has on the final result, the segmentation was repeated on different down-sampled levels. The thinned tile versions, with the relative point densities, are presented in Table 7. The procedure to process the segmentation outputs will be presented in the next paragraph (AMS3D_ITCs_generation). Once the 49 outputs were generated, they were assessed following the same framework used to compare the three segmentation algorithms which will be described in detail in the Validation paragraph.

Table 3: AMS3D parameter settings for Test 11-18, highlighted in yellow are the tested different parameter values (Test 14 was discarded during the analysis due to poor results).

Test	Gamma for the Gaussian kernel	Maximum distance between two modes to merge them	Minimum number of modes to create a cluster
Test 11	5	0.5	10
Test 12	5	1	10
Test 13	5	1.5	10
Test 15	1	1.5	10
Test 16	10	1.5	10
Test 17	5	1.5	2
Test 18	5	1.5	50

■ *Adaptive Mean Shift 3D AOI*

Once the most suitable parameters values were identified the AMS3D algorithm was run on the whole area of study with the parameter shown in Figure 6. The segmentation output contains for each point, the X, Y, Z coordinates and an *ID* value which identifies the tree. Nevertheless the algorithm's output could not be directly exploited for the AGB estimation, and thus several postprocessing operations were required. The first operation was to compute the Individual Tree Canopies polygons (ITCs), that are the ground projected areas of trees. The shape of these ITCs is very critical because it directly impacts the calculation of the tree's CD and thus the estimation of its biomass. For this reason it was crucial to compute the ITCs as accurately as possible, generating the polygons that best reproduced the crown shape. To do so, each tree cluster had to be processed to exclude those points that were found to be significantly distant from the rest, and that would have resulted in the generation of much larger ITCs. To automatically exclude these points the clusters' barycentres were calculated and then, the points whose distance from the barycentre was found to exceed a threshold based on the average point-barycentre distance, were removed. At this point the ITCs were generated computing the convex hull of the remaining points of each cluster. Once the ground-projected area was calculated the analysis proceeded with the extraction of the required tree attributes namely, TH and CD. The former was equalled to the height of the highest point of the cluster whereas for the calculation of CD more steps needed to be taken. In the study, considering the irregular shapes of ITCs, and in general the highly variable shape of crowns, it has been decided that consider the ITC shape as circles would have not returned an accurate measure of their diameter. It was instead decided that calculating the crown diameter as the average of four polygon axes could return a more accurate approximation. By doing so, in the case of crowns which were more elongated in one direction than another, the final CD could better estimate the mean polygon width. Once TH and CD were calculated, the postprocessing operations continued with the exclusion of those ITCs that did not intersect the original AOI (200 m x 200 m) and thus were not needed to be considered. Finally, ITCs with a surface area smaller than 1 m^2 or larger than 1000 m^2 were also excluded from the analysis since their size was not considered in line with the one of real crowns.

Figure 6: AMS3D parameters setting used to segment the point cloud of the Area Of Interest

Canopy Height Model generation

A CHM is the digital model representing the shape of vegetation features. Normally it is generated either keeping only the first LiDAR returns or computing the difference between the DSM and the DEM of the area. In this project however, considering the importance of the CHM for two of the applied segmentation methods, it has been decided to implement a more complex framework for generating the model. Based on the work of Khosravipour et al. (2016) the CHM was created implementing the Spike-free option of *las2dem* tool (Version 181108; Isenburg, M., 2018). The algorithm instead of considering only first LiDAR returns considers all returns, ignoring those points which will result in the formation of a pit. The resulting spike-free CHM was proved to work with different lidar point densities, and demonstrates a statistically significant improvement in the accuracy of tree detection (Khosravipour et al., 2016). In the generation process, the freeze constraint parameter plays an important role for the output result and it is related to the expected edge-length of the last-return TIN. As suggested on the webpage of the tool (https://rapidlasso.com/2016/02/03/generating-spike-free-digital-surface-models-from-lidar/), different values close to the triple of the pulse spacing were tested. The value of 0.2 was at the end chosen to better represent the area without oversimplifying the scene nor maintaining too many low elevation points.

In this research the spike free CHM (Sf_CHM) was first generated on single tiles and these were then assembled to form the model of the whole area. Lastly, the resulting output was also rasterized in order to be implemented in the Watershed segmentation. In addition to the Sf_CHM generation, also a standard CHM (St_CHM) was produced constructing a TIN based on solely the first returns. The two models were than compared to assess the difference between them, and check if the implemented framework was able to avoid the formation of pits in the Sf_CHM.

itcSegment segmentation

The second segmentation algorithm that was tested is the itcSegment as implemented in the *itcLiDAR* package (Dalponte, M., 2018). Contrary to AMS3D, the method segments the scene based on the CHM instead of the whole point cloud, and consequently it is not able to identify submerged trees and only detect those canopy that are visible in the CHM.

The delineation of individual tree crowns starts finding local maxima within the rasterized CHM, set them as treetops, and successively grow and delineate regions around these maxima following a multi-step procedure. Table 4 shows the function parameters with a brief description, while the complete segmentation process is further described below.

- In the first step the rasterized CHM is smoothed through the application of a mean-low-pass filter, which reduces the number of possible treetops.
- Local maxima are located using a variable moving window whose sides are constrained to vary between 9 and 15 pixels, according to the current cell height. A raster pixel is labelled as local maximum if it is higher than its neighbouring pixels, and higher than a pre-defined threshold, set to 10 meters.
- For each treetop, the heights of the neighbouring pixels are extracted. They are then added to the growing region if lower than the treetop, and within a user-defined range. The range is restricted by 2 thresholds, TRESHseed and TRESHCrown, which were

respectively set to 0.3 and 0.55. A pixel is added only if its height is greater than the treetop height multiplied by TRESHseed, and the mean region height multiplied by TRESHCrown.

- From each region the first LiDAR return are extracted
- Finally, a 2D convex hull is constructed around these points.

Contrary to AMS3D algorithm, the *itcLiDAR* function already returns the segmented trees in the form of ITC polygons whose height is already equalled to the one of the highest LiDAR point in the region. As a result, only CD values had to be retrieved, and they were calculated following the same framework developed for the ASM3D output. Also in this case ITCs that were not intersecting the original AOI were excluded from the analysis as well as ITCs with a surface area smaller than 1 m^2 while none of them were found to have an excessively large area.

Table 4: itcLiDAR function settings

Parameter	Value	Parameter use
Epsg	32622 (UTM Zone 22N)	Set coordinate system
Resolution	0.1 m	Set CHM resolution
MinSearchFilSize	9 pixels	Set minimum dimension of search window
MaxSearchFilSize	25 pixels	Set maximum dimension of search window
TRESHSeed	0.3	1st growing threshold. A pixel is added to the region if its height is greater than the treetop height multiplied by this value
TRESHCrown	0.55	2nd Growing threshold. A pixel is added to a region if its height is greater than the current mean height of the region multiplied by this value
minDIST	4 m	Minimum value of the crown diameter of a detected tree
maxDIST	40 m	Maximum value of the crown diameter of a detected tree
HeightTreshold	10 m	Minimum height of the trees
cw	2	Weighting exponent used to increase the contrast in the CHM used to detect the local maxima

Watershed segmentation

The last approach to delineate tree crowns that was tested is a marker-controlled inverse watershed segmentation. As for itcSegment algorithm the procedure started with smoothing the input CHM to reduce the possibility of locating additional treetops, which is common in the case of trees without a clear apex. This operation was performed applying a gaussian filter which reduced the pixels' height variability based on a moving window of 7 x 7 pixels. Once the input CHM has been smoothed it was possible to proceed with the delineation of crowns. In this case, the procedure implemented a variable window filter algorithm, as developed by Popescu and Wynne (2004) and executed by the *vmf* function of *ForestTools* R package (Plowright, A., 2018). The variable window size is calculated at any location as a function of the height of the current pixel. Other parameters, such as the minimum height to consider a pixel a possible treetop, the type of neighbouring window (Figure 7), and the maximum size of the variable window also needed to be specified. Before selecting the final parameters for the variable window (Table 5) several settings were tested and the resulting outputs were visually assessed to determine the most suitable ones. The testing particularly focused of the *winFun* parameter, which was found to have the most significant impact on the identification of treetops. To perform the segmentation, a marker – controlled watershed algorithm was applied to segment the area using the previously identified treetops as initial growing regions. The function mcwatershed, as implemented in *lidR* package (Roussel, J., 2019) takes a CHM and treetop locations as inputs and returns a raster with the individual segmented crowns. Furthermore, to avoid the understory vegetation to be added to tree regions, all those pixels with a height value smaller than 5 m were excluded from the segmentation. Once the segmentation was completed the obtained ITCs were again processed to retrieve the necessary information. TH values of each ITC were set equal to the corresponding treetop height, while CD were again calculated as the average of four ITCs axis. Finally, ITCs which did not intersect the AOI or had a surface area either smaller than 1 m^{2}or larger than 1000 m^2 were again excluded.

Table 5: vwf function parameters overview.

| Vwf parameter	Value	Parameter use
winFun	x * 0.06 + 0.5	The function to determine the size of the window at any given location based on the height of the current pixel (x).
minHeight	10 m	The minimum height value for a CHM pixel to be considered as a potential treetop
maxDiameter	100 pixels	Maximum window dimensions.
minWinNeigh	"Queen"	Define the smallest possible search window type (Figure 7).
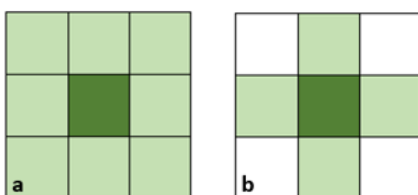
Figure 7: Queen's (a), and Rook's neighbourhood windows. In dark green the central pixel that is being processed, and in light green the pixels included for the identification of treetops.

Above Ground Biomass estimation

- ***Tree reference estimates***

In order to evaluate the AGB estimates of segmented ITCs, in absence of harvested-tree AGB measures, it has been decided to compare the obtained AGB values with "reference" estimates calculated on every stem. To maintain coherence with standard AGB estimates of the area of study, the same approach used by the local research group was followed, making use of the *BIOMASS* R package (Rejou-Mechain et al., 2017). AGB was estimated by applying a pantropical allometric model based on DBH, TH, and Wood Density (WD). The model, developed by Chave et al. (2014) (Equation 1), was implemented through the *computeAGB* function.

$$AGB = -0.0673 \times (WD * TH \times DBH^2)^{0.976}$$

Equation 1

Where WD (g/cm^3) is the wood density, TH (m) is the tree height, and DBH (cm) is the Diameter at Breast Height, parameters that were retrieved or estimated based on the information contained in the stem map.

First, the WD of each stem was retrieved by applying the *getWoodDensity* function, which estimates WD values based on Genus, Species, and Family information using the global wood density database (Chave et al., 2009).

After that, TH values were also estimated at stem level by applying an allometric model developed by Chave et al., 2014 (Equation 2). The model, implemented in the *retrieveH* function, requires the DBH of stems to be specified and can either take the coordinates of the plot or the geographic region to calculate the so-called environmental stress parameter (E), with the first option that was preferred returning a lower Residual Standard Error (RSE) value (0.243 m vs 5.285 m). The calculation of E (Equation 3) involves the use of three bioclimatic variables; namely, Temperature Seasonality (TS) which represents heat stress, Climatic Water Deficit (CWD), and Precipitation Seasonality (PS). The overall environmental stress which Paracou forest is subjected to was estimated to equal – 0.1077847.

Once all the required parameters were retrieved for each stem, AGB was finally estimated for all trees with a DBH greater than 10 cm.

$$ln(H) = 0.803 - E + 0.760 \ln(D) - 0.0340[ln(D)]^2$$

Equation 2

$$E = (0.178 \times TS - 0.938 \times CWD \times -6.61 \times PS) \times 10^{-3}$$

Equation 3

- *Tree derived estimates*

To estimate AGB for the automatically segmented trees the allometric equation developed in Jucker et al. (2016) was implemented. The equation is solely based on TH and CD values, which were already computed for each ITC. The allometric model, presented in Equation 4, was implemented by means of the *agb* function of *itcSegment* R package(Dalponte, M., 2018), setting the species group parameter to angiosperm and dividing the output value by 1000 to get estimates in Mg.

$$AGB = (0.016 + \alpha_G) + (TH \times CD)^{(2.013 + \beta_G)} \times exp(\frac{0.204^2}{2})$$ *Equation 4*

Where TH and CD are both expressed in meters. The terms αG, and βG are functional group-dependent parameters which account for the difference in scaling constant and scaling exponent between angiosperm and gymnosperm trees. In this study the most numerous tree species are found to be angiosperms and thus the values αG, and βG need to equal 0.

Validation

- ### *Segmentation algorithms validation*

To validate and compare the segmentation results of each algorithm, two procedures were developed. The first one aimed at validating the segmented ITCs based on the stem map by computing a confusion matrix with relative statistics. The second procedure was developed to assess ITCs similarity with reference crowns by calculating the Jaccard Index. Both methodologies are explained in further details in the next paragraphs.

- ### *Confusion matrix*

Due to the nature of the ITCs and stems datasets, the generation of the confusion matrix posed several problems that needed to be solved. In standard situations the computation of the confusion matrix provides a straightforward mean to assess the classification of two or more different classes based on true reference values. In image classification for example, each pixel is valuated as correctly or wrongly classified comparing it with the corresponding pixel in the true dataset. The reliability of such a comparison is guaranteed by the location information embedded in each pixel. In this study, such a comparison was not possible. While the ITCs features were in the form of polygons, the ground truth dataset was composed of simple points, so that a conventional assessment between the two datasets were not feasible. Furthermore, the uncertainty of stem locations added a further challenge when assessing the segmentation outputs. In order to overcome these limitations the comparison and match of segmented trees with the reference stems was performed with the introduction of different thresholds.

First, to evaluate if a tree was segmented in the correct location, the distance between the ITC centroid and the intersecting stems was calculated. If the distance was found to be smaller than a user-defined threshold (Thresh_D), the segmented tree was considered as a possible match with the corresponding stem. In this study the influence of three different distances was explored, namely 0.5 m, 1 m, and 2m. Conversely, since the LiDAR data had not been co-registered yet when testing the AMS3D algorithm, the distance thresholds was set to 4 meters.

Moreover, the validation of AMS3D ITCs posed an extra challenge for matching stems with the right crowns. The algorithm in fact, can also segments understory trees and thus produced overlapping polygons that may 'compete' for the same stem. For some stems it can therefore occur that the closest ITC does not represent the corresponding canopy, leading to a wrong match that should be avoided. To this end an extra constraint had to be set. Based on the available data, two tree attributes could be exploited to prevent erroneous association, TH and DBH. The former, was already obtained by applying Equation 2, while the latter could be estimated for each ITC based on CD and TH measures applying Jucker et al. (2016) DBH allometric model, Equation 5.

$$DBH = 0.557 \times (TH \times CD)^{0.809} \times exp\, 0.056^2/2 \hspace{2cm} \textit{Equation 5}$$

Both approaches were tested during the analysis and both showed high discrepancies between observed and estimated values. At the end, it was decided to estimate stem

height and use the difference with the ITC height as second constraint. The choice to exclude estimated DBH from the validation was driven by its high dependency on CD. Even though CD has been proved to be essential for differentiating trees of similar height but having substantially different trunk sizes (Jucker et al., 2016), the measure may already present a certain level of uncertainty due to the segmentation process. Furthermore, it was preferred to exclude the influence of ITCs from the computation of the second threshold. On the other hand, the estimation of tree height based on stem DBH was considered to only add a small uncertainty which corresponded to an RSE of 0.243 m. Nevertheless when deciding the most suitable threshold value to set, another consideration had to be made. As a matter of fact the AMS3D method, when segmenting a tree, might erroneously include points of adjacent canopies. These points, that can be higher than any other in the current tree cluster, add an extra source of error that should be considered.

As a result, the stems that were already found to be within the distance threshold were inspected comparing TH with the EstTH of the current stem. In order to limit the possibility of excluding the right match because of height difference, and exclude clearly erroneous stem-ITC matched, two thresholds (Thresh_H) of 10 and 15 m were tested. Finally, among the remaining possible matches that satisfied both constraints (Thresh_D and Thresh_H), the closest stem was paired with the ITC and labelled as True Positive (TP).

After that, unmatched ITCs were labelled as False Positive (FP) whereas unmatched stems were labelled as False Negative (FN). For the calculation of True Negatives (TN), another issue had to be overcome. In fact, due to the nature of data again, it was not possible to assess when the segmentation correctly had not segmented a tree at a certain location. To solve this problem it has been decided to randomly select sample locations within the plot that could be considered as points where no tree was present. This was achieved by selecting points that were at a certain distance from other stems. The distance was then set equal to the threshold previously used for validating the ITCs locations. Furthermore, to ensure an even distribution of stems (points with a tree) and random samples (point without a tree), the latter were selected in the same number as the stems (2291). FN ITCs were then tested for possible matches with sample points based again on the current distance threshold. At the end of the process, the unmatched sample points were labelled as TN. With this information available the computation of the confusion matrix was completed, and the analysis could proceed calculating several statistics (Equations 6-10). The singular steps taken to compute the confusion matrix and the associated statistics are presented below.

$$Error\ Rate\ (ERR) = \frac{FP+FN}{TP+TN+FP+FN}$$
Equation 6

$$Accuracy\ (ACC) = \frac{TP+TN}{TP+TN+FP+FN}$$
Equation 7

$$Precison\ (PREC) = \frac{TP}{TP+FP}$$
Equation 8

$$Recall\ (REC) = \frac{TP}{TP+FN}$$
Equation 9

$$F\ score\ (F1) = 2 * \frac{REC*PREC}{REC+PREC}$$

With TP = True Positive, TN = True Negative, FP = False Positive, and FN = False Negative.

The generation of the confusion matrix with the relative statistics was repeated on two ITCs subsets obtained by setting a TH value of 20 m as cut-off point. Dividing the identified crowns based on their height will provide further details regarding the accuracy of the algorithms and their ability to segment low and high vegetation.

As an example, Figure 8 illustrates four different situations where segmented ITCs are respectively assessed as TP (a), FP (b), TN (c), and FN (d). In the first case it can be observed that the validation algorithm was able to match seven different ITCs (red polygons) which partly or completely overlapped. The distance threshold allowed to pair most tree crowns with the corresponding stems (yellow dots), the height difference threshold made sure that no crowns were paired with a wrong stem. In case (b) instead, a FP crown is presented. Here there were three stems within or in the surroundings of the ITC, but none was matched. Two of them did not satisfy the distance constraint since they were further than 2 meters, the closest stem instead exceeded the height difference threshold of 15 m and hence was not matched either. In figure (c) a random point (light blue dot) is shown with a 2 m buffer around, as it is visible, no stem fell within the buffer and thus the point was labelled as TN. In the last case (d), a stem (yellow dot) is displayed with a 2 meters buffer around. Since none of the surrounding crowns barycentres are falling within the buffer the stem represented a FN, in other words, a tree that could not be segmented by the algorithm or matched during the validation process.
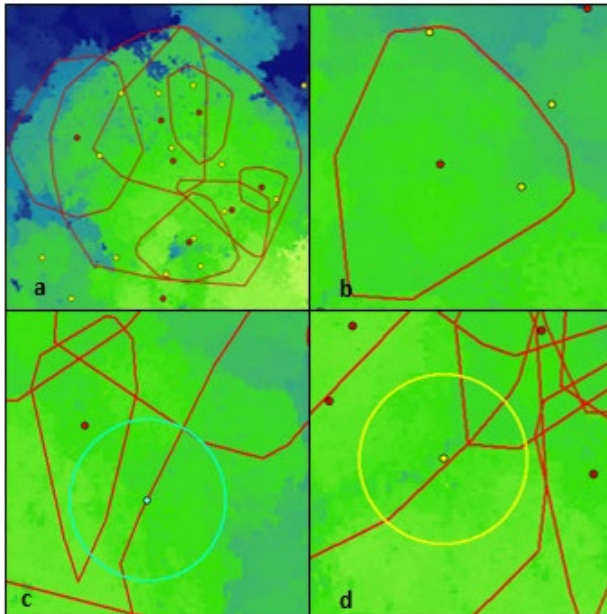
Figure 8: (a) True Positive: Individual Tree Canopies (ITCs) matched with a stem, (b) False Positive: ITC not matched with any stem (two stems in the upper part are further than 2 m while the stem on the right didn't matched the height difference constraint), (c) True Negative: random sampled point with no ITC centroid closer than 2 meters, (d) False Negative: stem with no ITC centroid closer than 2 m. Legend: ITCs = red polygons, ITC_centroids = red dots, stems = yellow dots, 2m_stem_buffer = yellow buffer, random_sample_point = light blue dot, 2m_random_sample_buffer = light blue buffer.

- *Jaccard Index*

To evaluate the accuracy of segmented crown shapes, a sample of five trees was generated for each of the following tree classes.

- Tree in the emergent canopy layer (ECT)
- Tree in the main canopy layer (MCT)
- Tree in the sub-canopy layer (SCT)

The reference trees were first manually segmented from the UAV point cloud and their convex hull (crown polygon) was then computed. The obtained crowns, shown in Figure 9, were then implemented for the calculation of the Jaccard index. The index, presented in Equation 11, tests the similarity between two crowns computing the ratio between their intersection and union. Following the approach of Aubry-Kientz et al. (2019), a crown was considered correctly segmented when the Jaccard Index exceeded the threshold value of 0.5. The presented methodology was only applied on those ITCs that had already been assessed as correctly matched. In this way the possibility of retrieving high Jaccard indices on wrongly identified crowns was avoided. On the other hand, it should be considered the possibility that crowns with a higher similarity might have been labelled as False Positive and thus not taken into consideration in the index computation. When more TPs intersected the reference crown, the one with the biggest overlap, assumed to be the corresponding ITC, was selected for the Jaccard Index calculation. Finally, in order to make the validation meaningful also for the raster-based segmentations, which cannot detect trees in the subcanopy layers, it was decided to segment low elevation trees which were still represented in the CHM.

$$J\,(A,B)\ =\frac{|A \cap B|}{|A \cup B|} \qquad\qquad\qquad \textit{Equation 11}$$
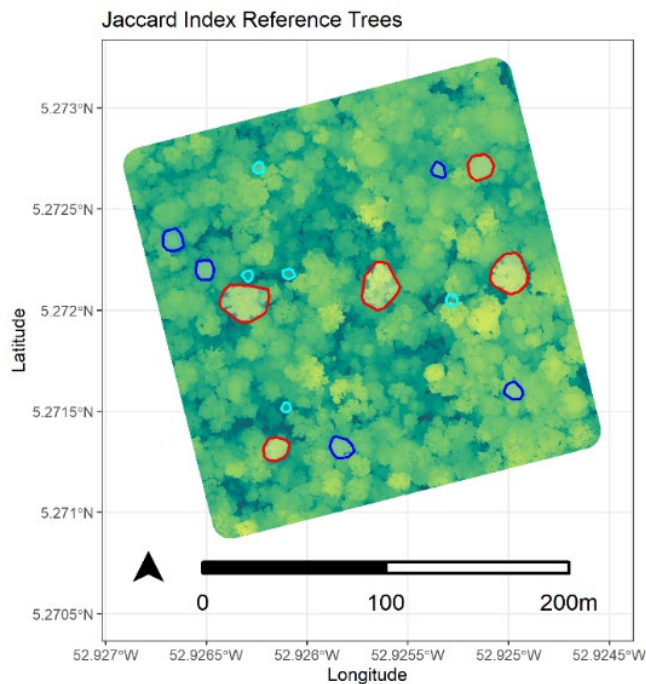
Figure 9 Reference, manually segmented ITCs. The colours indicate the three different classes. ECT: red, MCT: blue and SCT: light blue

- ***Above Ground Biomass estimates comparison***

To examine which segmentation algorithm produced the most reliable AGB estimates and to investigate their ability to predict AGB at tree level two different comparisons were performed. At first, we compared the obtained AGB estimates with the reference ones at plot level calculating absolute error (Equation 12) and relative error (Equation 13). The assessment can give a first indication whether the algorithms over- or underestimate the forest biomass. Comparing AGB at tree level instead is useful to test the prediction accuracy of the model itself and the segmentations performance for estimating biomass of small and large trees. In the latter comparison exclusively True Positive ITCs could be considered as the only ones with an associated stem. To test whether the adopted framework led to an over or underestimation of AGB based on tree size, ITCs and stems were divided into four DBH classes:

- DBH < 20 cm
- 20 cm < DBH < 40 cm
- 40 cm < DBH < 60 cm
- DBH > 60 cm

In order to compare the reference (AGBRef) and estimated (AGBEst) AGB values, two measures of average error were computed: the root mean square error (RMSE, in Mg), Equation 14, and the relative systematic error (or bias, in %), defined in Equation 15 .

$$AGB_{Error} = AGBEst - AGBRef \hspace{4cm} \text{\textit{Equation 12}}$$

$$AGB_{Rel\_Error} = \frac{AGBEst - AGBRef}{AGBEst + AGBRef} \hspace{3cm} \text{\textit{Equation 13}}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(AGBEst,i - AGBRef,i)^2}{n}} \hspace{2.5cm} \text{\textit{Equation 14}}$$

$$Bias = \ 1/N \ \times \ \sum_{i=1}^{n}\left(\frac{AGBEst,i - AGBRef,i}{AGBRef,i}\right) \times 100 \hspace{1cm} \text{\textit{Equation 15}}$$

Results

UAV Data pre-processing

As a result of the pre-processing operations performed on the raw data the obtained point cloud for the area of study is composed of 371,034,339 points distributed over 4 hectares and returns a final point density of 7502.31 points per square meter. Of all points, 1,666,070 have been classified as ground returns, which is about 0.5 % of the total point count. Additional information about the obtained point cloud dataset can be found in Table 6.

Table 6: Main characteristics for the point cloud of the Area Of Study

Attribute	Min	Max
X	265148	1314654
Y	-152120	897390
Z	551721	799224
Intensity	0	62966
Return number	1	7
Number of returns	1	7
Classification	-74	71
Point density (pts/m^2)	All returns: 7502.31	Last returns: 4441.31
Point spacing (m)	All returns: 0.01	Last returns: 0.02
Coordinate System	WGS 84 / UTM zone 22N, epsg = 32622	
Number of points	371,034,339	
Ground points	1,666,070	

- ▪ ***Data down-sampling and normalization***

During the point density testing, seven down-sampled versions of the original test tile were obtained. Resulting point densities and percentages of original points can be found in Table 7. Before proceeding to down-sample the whole point cloud and continue with the analysis, the AMS3D segmentation was run on each thinned version of the test tile and the results were analysed to identify at which point density the algorithm performed the best. A grid size of 0.05 m was at the end identified as the most suitable option and the down-sample of the whole point cloud was performed. The thinning process generated a point cloud composed of 19,022,940 points which corresponded to 5.13 % of the original point count. The point density dropped from 7502.31 to 384.64 (pts/m^2) for all returns and from 4441.31 to 231.32 (pts/m^2) for only last returns. The spacing between individual returns increased to 0.05 m (grid size) for all returns and 0.07 m for last returns only. Figure 10 presented a horizontal crown section before and after the down-sampling process (grid size 0.05) and gives an idea about the inevitable loss of information when the point density is reduced. At full point density tree branches are clearly distinguishable and well represented. In the thinned version instead, only few main branches are still visible, however the crown boundaries seem to still fairly represent the original ones. This is of extreme importance for the segmentation methods, especially for those based on the CHM. In fact, when rasterizing the point cloud it

is important that points at the borders of crowns are still present so that the corresponding pixels will get assigned with an accurate height value. On the other hand, the lack of points within the canopy does not represent a problem since those pixels will be calculated based on higher points which represent the top of the canopy and that are still present

Table 7: Point density (pts/m2), and percentage of remaining original points for each down-sampled test performed on Tile 286400_583100.Higlighted in yellow the down-sample level used to thin the final dataset.

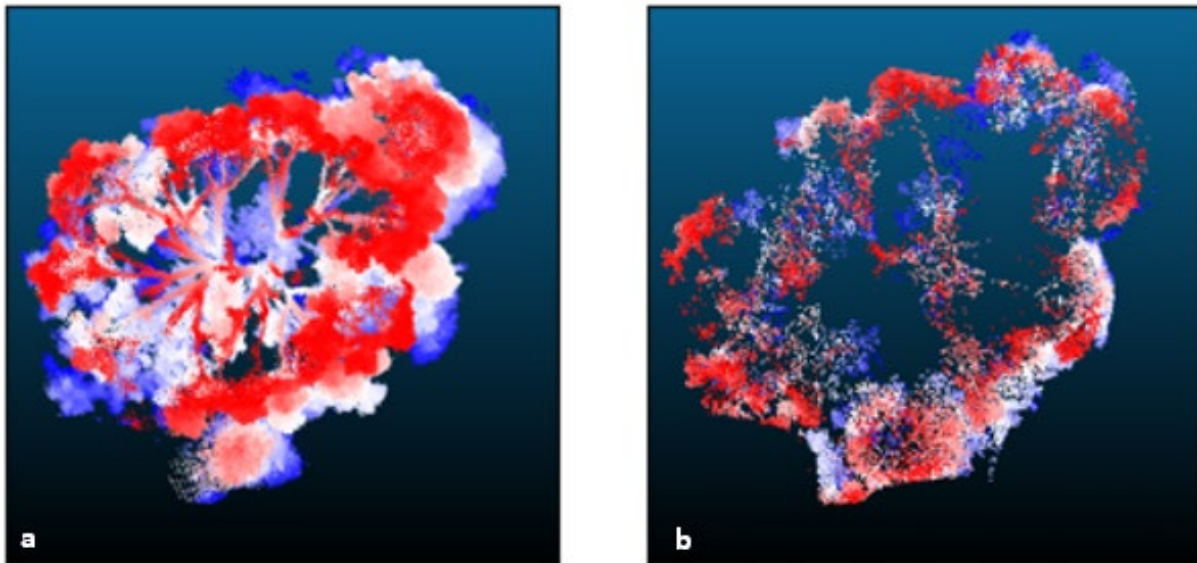| Test Tile	Grid size	Points density (pts/m²)	Percentage of original points (%)
286400_583100		9,489	100
Test_0.0125	0.0125	1922.82	75
Test_0.025	0.025	1054.15	41
Test_0.05	0.05	350.89	13
Test_0.1	0.1	93.43	3.6
Test_0.15	0.15	42.08	1.6
Test_0.2	0.2	23.87	0.9
Test_0.3	0.3	10.71	0.4

Figure 10: Horizontal section of an emergent canopy at full point density (a), and at a point density of 350.89 pts/m² (b).

Adaptive Mean Shift 3D Segmentation

- ***AMS3D testing***

When testing the AMS3D algorithm, the analysis focused on the effect of three parameters. The first one to be inspected was the maximum distance to merge two modes (MaxDist), which was found to control the aggregation of individual clusters. Small MaxDist values prevent the algorithm to merge points that converge to nearby modes. On the other hand, high values cause the grouping of individual clusters into larger ones. Merging different modes can help to connect different regions of the same tree, avoiding over-segmentation of the scene. However, excessively large thresholds result in the combination of far clusters which might belong to different trees, causing under-segmentation. The MaxDist parameter should hence be carefully selected considering the nature of the forest, the point cloud density, and the research goal. In fact, in the presence of a dense understory layer, the algorithm will be more prone to identify many modes segmenting a high number of trees. The same result might result from very high-density point clouds. Finally, depending on the goal of the project, one might decide to set a large MaxDist value, under-segment the lower strata, but avoid over-segmentation of big emergent canopies, which are well-known to store large quantities of carbon (Slik et al., 2013) and thus may be more relevant for biomass estimation.

The second parameter, the gamma value for the Gaussian profile (Gamma) has a completely different effect on the segmentation, affecting the algorithm efficiency for distinguishing adjacent, interconnecting canopies. Low values will widen the Gaussian profile increasing its variance and thus assigning higher weights to points that are further away in the horizontal dimension. Consequently, the computation of the bandwidth barycentre will be more affected by far points and the algorithm will struggle to separate adjacent crowns. On the other hand, high Gamma values will narrow the Gaussian kernel preventing the barycentre to move in the horizontal direction favouring the identification of a larger number of modes and hence the under-segmentation of big trees.

The last parameter tested was the minimum number of modes to form a cluster (MinMod), which can be thought as the minimum number of points to segment a tree. This parameter strongly depends on the point cloud density and it is expected to primarily influence down-sampled tiles where single trees are composed of few points. Small MinMod values will prevent the segmentation of clusters of too few points whereas high values will cause the fusion of different trees due to their insufficient number of points.

After running the AMS3D algorithm on the test tiles and retrieving the ITC polygons, the effect of the above-mentioned parameters was assessed comparing different attributes of the obtained outputs.

- *Number of ITCs*

The number of segmented trees gives a first indication about the degree of the segmentation. A large number of ITCs may be an indication of over segmentation while a small number may suggest that trees have been under-segmented or clustered together. Segmenting the scene with a point density of 10.71 pts/m2 and a MinMod value of 100 (Test 18_03) returned the least number of ITCs, on the contrary, the largest number was observed when segmenting a high density point cloud (1922.82 pts/m2) with the smallest MaxDist value (Test 11_00125).

When assessing the influence of the three parameters it was clear that MaxDist had a greater influence than the other two parameters. Figure 11 (a), shows that increasing the distance from 0.5 m to 1.5 m caused a reduction of ITCs which dropped from about 3700 to about 700. Conversely, an increase of Gamma (Figure 11 (b)), as expected, led the algorithm to segment a greater number of trees. Following the same pattern of the first graph, also the MinMod parameter resulted in fewer trees when set to larger values, but this time the difference was not as pronounced as in the first case (Figure 11 (c)). Furthermore, observing the different point colours it is interesting to notice how tiles that were heavily down-sampled corresponded to smallest number of segmented ITCs.
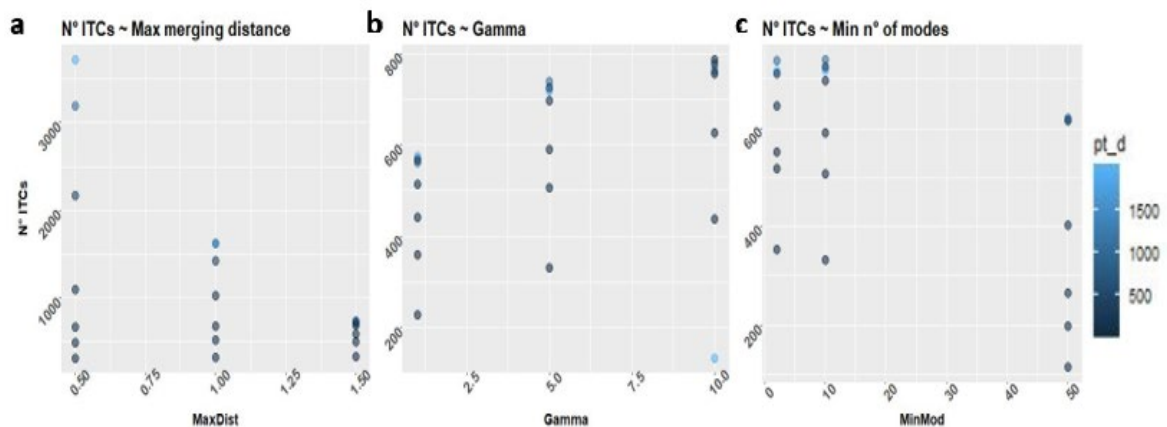
Figure 11: Graphs showing the relationship between the number of segmented crowns (ITCs) and the three parameters explored during the Adaptive Mean Shift 3D algorithm testing, (a) Maximum distance to merge two modes (MaxDist), (b) gamma value for the gaussian kernel (Gamma), and (c) Minimum number of modes to create a cluster (MinMod). Different point colours show the point density of the down-sampled tests.

- *Understory points*

Every test grouped a large number of low elevation points together, producing a large cluster that extended over the whole area of study. The size of this big cluster, which was composed of mainly understory points, was found to primarily depend on the MaxDist parameter. The clustering of points from all over the plot can be explained by the fact that in low forest layers, vegetation features form a dense, homogeneous set of points. Here the points lead to modes that are close to each other and thus, are more prone to be combined into single clusters when increasing the merging threshold. When inspecting the relationship between the understory points, expressed as percentage of the total points number, three different patterns were identified. In Figure 12 (a), the three densest

versions of the test tile are shown (point density of 1922.82 pts/m2, 1054.15 pts/m2, 350.89 pts/m2) and they depict a clear pattern between the MaxDist parameter and the resulting number of that populate the understory cluster. For higher values of MaxDist the total number of points that have been grouped also increases, with test_13_0.05 (MaxDist = 1.5 m and Down-sample percentage = 13.7%) having the densest understory cluster which contain about 10 % of the total amount of point. The second graph, Figure 12(b), illustrates the atypical case of test_13_0.1. Here the percentage of understory points dropped from about 8% to about 4% when MaxDist increased from 0.5 m to 1 m, and then rose again up to about 9 % for a MaxDist of 1.5 m. At first, the low percentage of understory points of the second case might be interpreted as an error. However, when visually assessing the three cases it was found that at this value of the merging threshold, additional large point clusters were generated. Therefore, the typical big understory group counted a smaller number of points and thus the lower percentage. Once MaxDist was increased again to 1.5 m, these groups were merged again with the big cluster, that now consisted of about 10% of the total number of points. The last graph, Figure 12 (c), which referred to the three tests with the lowest point densities, depict a completely opposite pattern compared to the first case. Here, the percentage of understory points decreases with the increase of MaxDist. As in the previous case, with higher MaxDist values more and more big clusters were segmented in the lower forest layer. As a result, the big understory group presents, in percentage, less points than the cases with a smaller merging threshold.
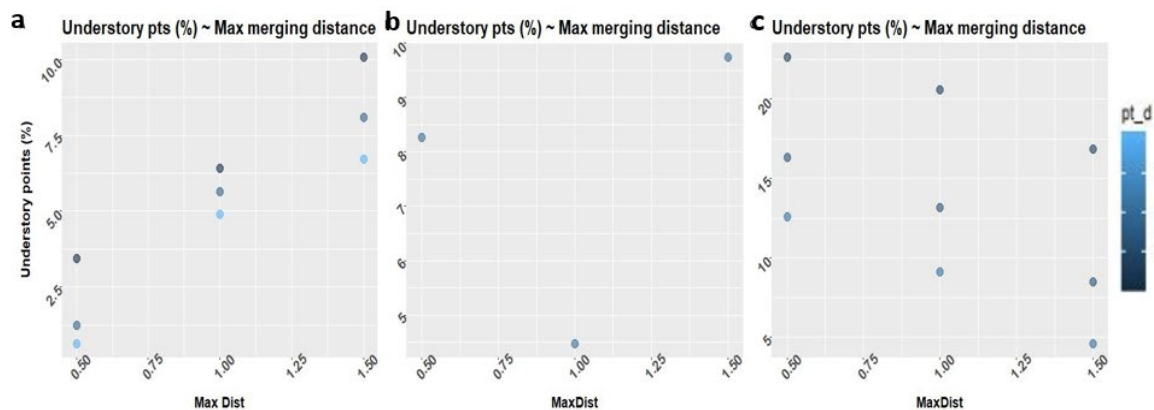
Figure 12: Graphs showing the relationship between the percentage of points that have been segmented as a single, large understory cluster and the Maximum distance to merge two modes parameter (MaxDist). The point data referred to the segmentation outputs of test 13 obtained at different down-sampling levels. (a) test tiles at a point density of 1922.82 pts/m², 1054.15 pts/m², and 350.89 pts/m². (b) test tile at a point density of 93.43 pts/m², and (c) test tiles at a point density of 42.08 pts/m², 23.87 pts/m², and 10.71 pts/m².
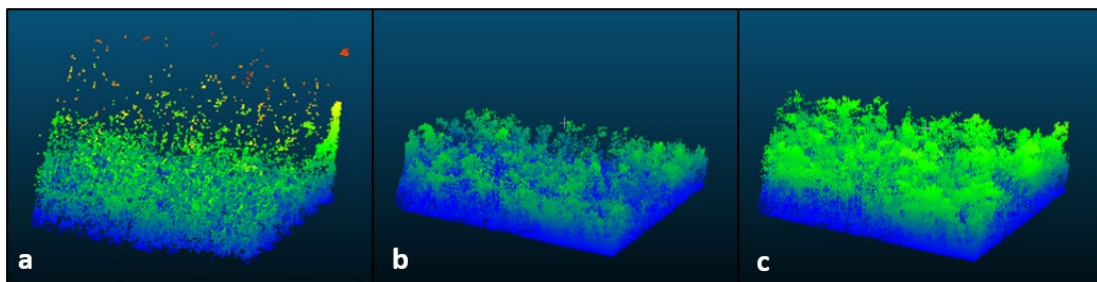
Figure 13:Points that were clustered together with different values set for the Maximum distance to merge two modes parameter (MaxDist) at a point density of 1054.15 pts/m². (a) MaxDist of 0.5 m, (b) MaxDist of 1 m, and (c) MaxDist of 1.5 m.

- *Point density*

The AMS3D segmentation was originally developed to segment ALS point clouds, which present much lower point densities than the ones obtained with UAV systems. The algorithm, however, does not evaluate the density function itself but normalized local gradients, thus the point density at which the crowns are sampled has only a slight impact on the identification of individual vegetation features (Ferraz et al., 2012). Nevertheless, it was decided to test this assumption to evaluate the influence of increasing point densities on tree segmentation results. It was found that the amount of points does affect the number of segmented ITCs, especially at low MaxDist values, as it shown in Figure 11.

The effect of point density on the segmentation outputs is particularly clear when visually inspecting singular trees. Figure 14, shows how the AMS3D algorithm, segmented the same tree at increasing point densities. As it is noticeable, in each figure the tree includes part of other trees and miss some other sections, but what is more striking is the general representation of the tree itself. While at low densities only the crow is represented, as the number of returns gradually increases the shape of the tree becomes more defined and new elements that were not visible before, now become clear. From a point density of about 350 pts/m^2 (Figure e) the trunks start to be visible and help to distinguish the multiple tree in the scene. The reason why at lower point densities only the upper part of the crows was segmented is again related with the clustering of low elevation point in one big cluster, referred as understory cluster in the previous sections. On the other hand it is also true that denser point clouds are indeed able to segment a tree from the trunk up to the canopy apex but also identify part of the same canopy as separated cluster, leading to new challenges for finding the best compromise. Nevertheless, the added value of high point densities is evident as more points enable the algorithm to easily differentiate adjacent trees.

Figure 14: Segmentation of the same tree obtained with test 18 at different point density levels. (a) 10.71 pts/m², (b) 23.87 pts/m², (c) 42.08 pts/m², (d) 93.43 pts/m², (e) 350.89 pts/m², (f) 1054.15 pts/m², and (g) 1922.82 pts/m².

- *Confusion matrix and statistics*

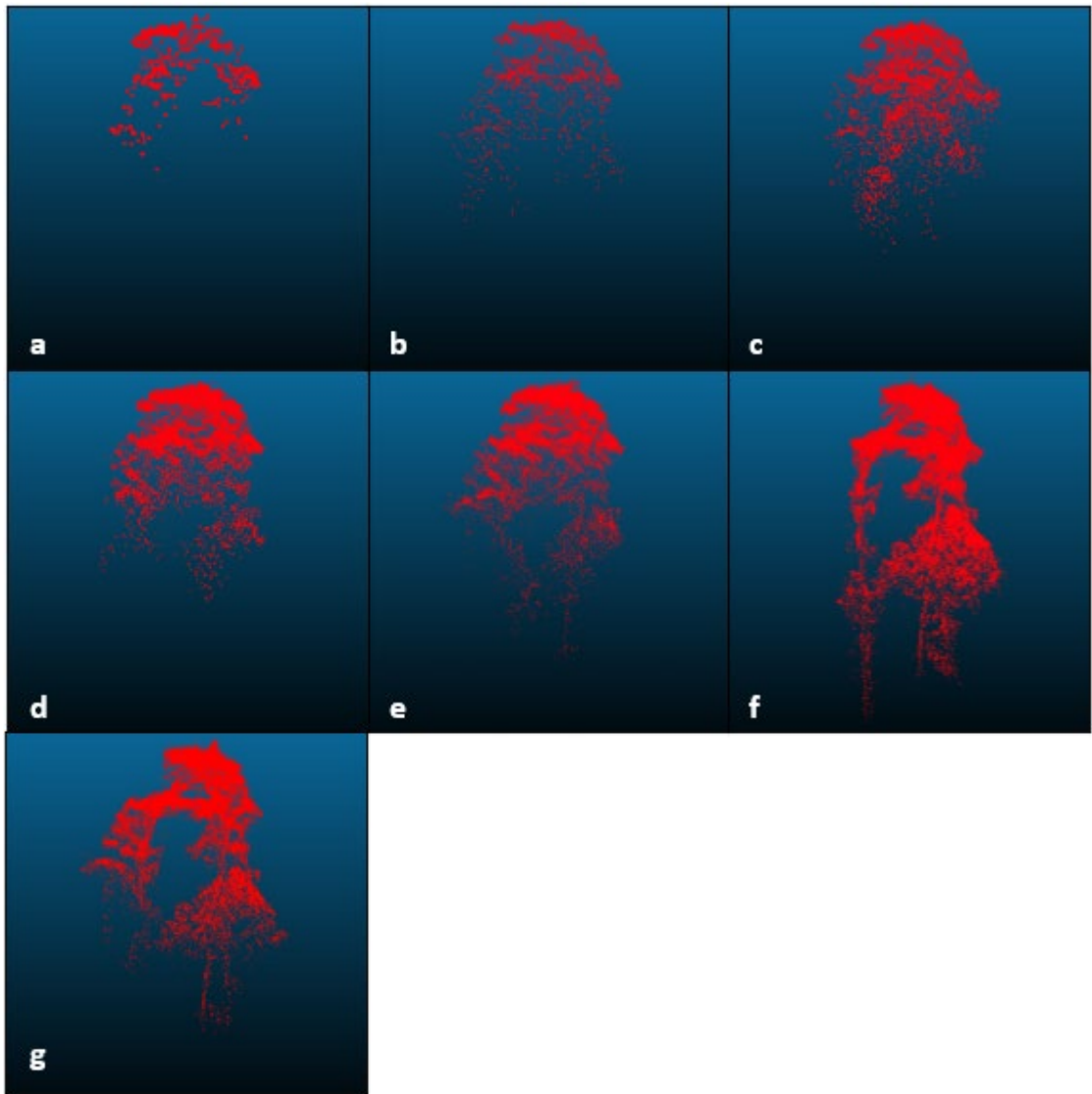To better compare the different tests outcomes, the results were also assessed with the computation of a confusion matrix and the calculation of the relative statistics. A first indication of the test performance can be inferred by examining the number of correctly segmented trees, nevertheless the information is only partially meaningful. In fact, a large number of TP may also indicate a severe over-segmentation of the area, whit many ITCs that could be matched with a stem. For this reason the final comparison and assessment of different tests was based on a more robust indicator which is the F-score. The latter is a well-known measure of test accuracy which considers both precision and recall and varies between 0 (worst performance) and 1 (perfect performance). In addition to evaluate the segmentations based on the whole ITC dataset, two additional datasets were created based on TH. Computing the confusion matrix and statistics on the two new subsets helps to better understand the segmentations effectiveness in the lower part of the canopy (TH < 20 m) and in the higher layer (TH > 20 m). Complete version of the confusion matrices can be found in Appendix 2.

In Table 8 the five tests that obtained the highest F-score when assessing the whole ITCs dataset are shown. Four out of the five best results were obtained by test 18 (MaxDist = 1.5, gamma = 5, MinMod = 50), which generated the best statistics compared to the other tests. The best results were obtained with a point density of 42.08 pts/m^2 and reached a F-score of 0.4563.

Table 8: Statistics of the five best tests of the Adaptive Mean Shift 3D segmentation algorithm based on the obtained F– scores for the whole dataset.

Test	Point density (pts/m^2)	Error Rate	Accuracy	Precision	Recall	F-Score
Test_18	42.08	0.5679	0.4321	0.4023	0.5271	0.4563
Test_18	10.71	0.5357	0.4643	0.6000	0.3399	0.4340
Test_18	23.87	0.5661	0.4339	0.4394	0.4286	0.4339
Test_15	10.71	0.5897	0.4103	0.3904	0.4384	0.4130
Test_18	93.43	0.6292	0.3708	0.3077	0.6108	0.4092

When considering the segmentation of only high trees (Table 9), test 18 confirmed to be one of the most accurate, generating the first and third best statistics, with test 15 which returned the second highest F-score. It is interesting to notice that for high trees the best segmentations have been obtained with point densities of more than 1000 pts/m^2. A possible explanation is that at high densities the upper canopies, which contain most of the points, are finer represented and thus lead to a more accurate delineation of big trees.

Table 9: Statistics of the five best tests of the Adaptive Mean Shift 3D segmentation algorithm based on the obtained F– scores for the high-trees subset (Tree Height > 20 m).

Test	Point density (pts/m²)	Error Rate	Accuracy	Precision	Recall	F-Score
Test_18	1054.15	0.5119	0.4881	0.5272	0.4778	0.5013
Test_13	1054.15	0.5278	0.4722	0.4854	0.4926	0.4890
Test_18	1922.82	0.5237	0.4763	0.5109	0.4631	0.4858
Test_17	1054.15	0.5315	0.4685	0.4806	0.4877	0.4841
Test_12	1054.15	0.5482	0.4518	0.4375	0.5172	0.4740

A similar situation occurs when considering the subset composed of short trees (TH< 20 m), with test 18 that was again found to be the most effective. As shown in Table 10, in this case a lower point density (350 pts/m²) produced the best results, however, a significant decrease of about 0.2 was observed in the F-scores. The poorer results corroborate that in general, the segmentation of understory vegetation still represents a major challenge.

In conclusion, the analysis of the obtained confusion matrices gave an unequivocal indication on the most effective settings for segmenting the test tile. Test 18 outperformed other tests for both low and high vegetation. It is thus clear that a higher value of MinMod can significantly improve the segmentation avoiding the creation of small cluster of points and the consequent over-segmentation of the area.

Table 10: Statistics of the five best tests of the Adaptive Mean Shift 3D segmentation algorithm based on the obtained F– scores for the low-trees subset (Tree Height < 20 m).

Test	Point density (pts/m²)	Error Rate	Accuracy	Precision	Recall	F-Score
Test_18	350.89	0.7129	0.2871	0.2162	0.4483	0.2917
Test_18	1054.15	0.7201	0.2799	0.2082	0.4483	0.2844
Test_15	1054.15	0.7262	0.2738	0.2018	0.4483	0.2783
Test_18	1922.82	0.7268	0.2732	0.2005	0.4286	0.2732
Test_18	93.43	0.6820	0.3180	0.2589	0.2857	0.2717

- *Jaccard Index:*

The computation of the Jaccard index, despite the limited number of reference crowns, returns a first indication of the congruency of segmented canopies. Table 11 lists, for each reference tree, the five highest Jaccard indices and the corresponding tests. In the case of the emergent canopy (ECT), test 15, 18, 12, and 16 generated the same ITC and thus obtain an equal Jaccard index of 0.8467, which indicates an excellent similarity. These tests are characterized by a MaxDist of 1.5 m except for test 12 where the parameter was set equal to 1 m, indicating that low MaxDist values may not appropriate for segmenting large crowns. The overlap then dropped to about 0.58 for both the remaining reference crowns, indicating the general higher complexity of accurately segmenting trees in the lower layers. It is interesting to notice that also in this case larger MaxDist values produced the best results.

Table 11: List of five matched segmented ITCs with the highest Jaccard index for each reference tree.

ITC test	Point density (pts/m²)	Reference tree	Jaccard Index
15	1054.15	ECT	0.8467
18	1054.15	ECT	0.8467
12	1054.15	ECT	0.8467
16	1054.15	ECT	0.8467
13	23.87	ECT	0.8378
15	350.89	MCT	0.5852
18	93.43	MCT	0.5840
18	350.89	MCT	0.5679
17	42.08	MCT	0.5582
13	10.71	SCT	0.5876
17	10.71	SCT	0.5876
12	10.71	SCT	0.5876
15	10.71	SCT	0.5876
18	10.71	SCT	0.5876

- *Visual comparison*

Finally, as in the case of point density, outputs of different tests were also compared at tree level. The same individual tree was extracted from the point clouds and the segmentation outcomes were visually inspected. In general, none of the test was able to segment the investigated tree without including parts of adjacent trees or segmenting section of the crown as different clusters. Inspecting the single trees shown in Figure 15, it is interesting to notice how multiple tests produced quite similar clusters, with only test 15 (d) and 16 (e) that shows higly different results. These tests were respectively designed to experiment with a lower and a higher value of gamma, which plays an important role when differentiating adjacent trees. The tests confirm that at low gamma values, far points in the horizontal direction still get a high weight in the computation of the

bandwidth barycenter. As a result, the algorithm struggles to separate canopies that interpenetrate each others, which is confirmed by the big cluster in **Error! Reference source not found.** (d) that comprends parts of several canopies.

On the other hand, varying MaxDist and MinMod parameters does not generate any appreciatable difference in the segmentation.
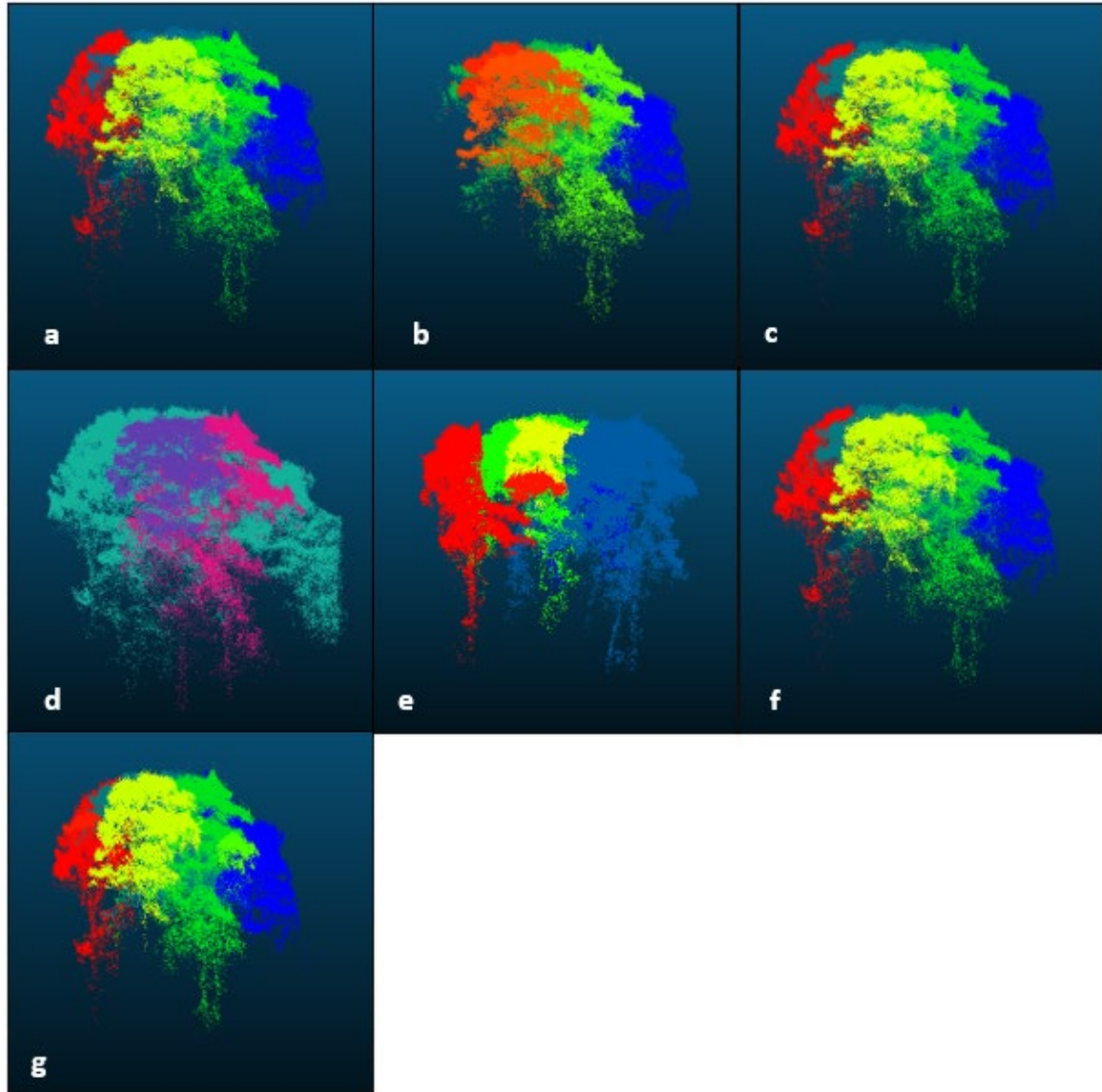
Figure 15: Extraction of a sample tree and the adjacent canopies obtained with different test setting at a point density of 1054.15 pts /m². (a) Test 11, (b) Test 12, (c) Test 13, (d) Test 15, (e) Test 16, (f) Test 17, and (g) Test 18.

- *Adaptive Mean Shift 3D AOI*
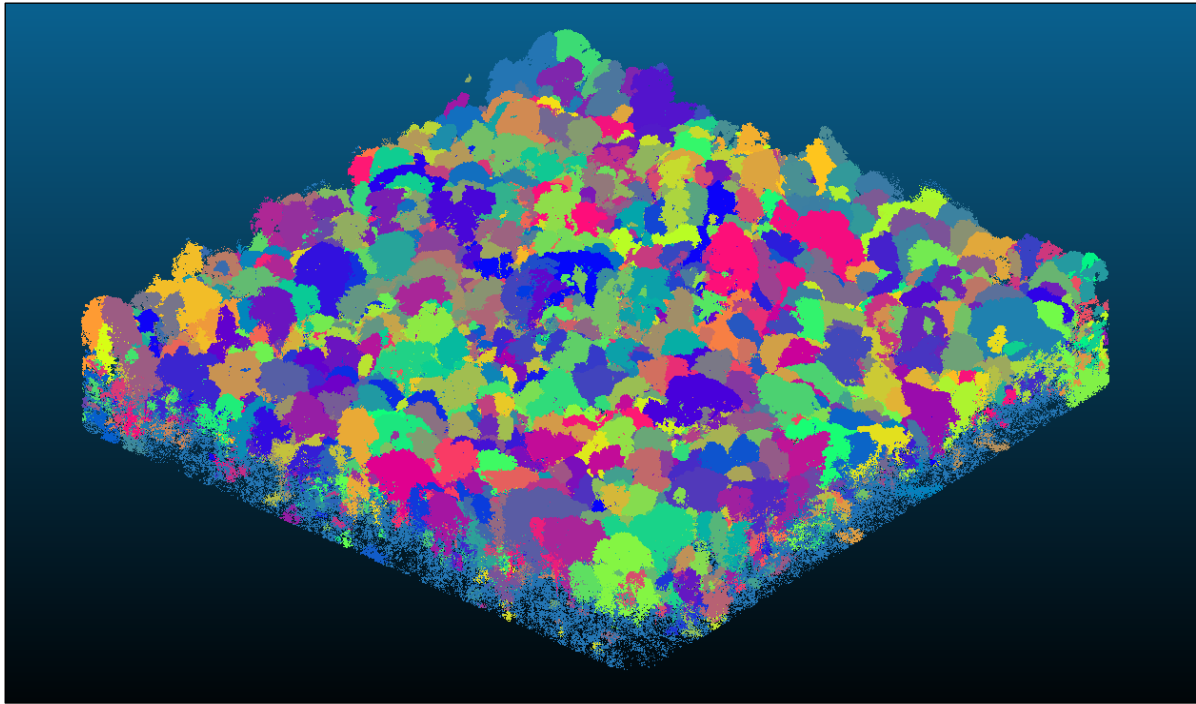
Figure 16: Segmented scene of Area Of Study obtained with the AMS3D segmentation algorithm.

The segmentaiton of whole area (10 m buffer included), shown in Figure 16, generated a total of 6723 ITCs of which, 967 have been excluded from the analysis since they were not intersecting the original area of study (4 ha, no buffer) or because their size was not congrous with a canopy. Figure 17 shows the big understory cluster (a) and others big group of points (b) that were removed before proceeding with the segmentation assessment. The remaining 5756 segmented polygons, which are about 2.5 times more than the reference stems, ranged from very small trees, with a minimum Crown Area of 1.03 m^2 and minimum Tree Height of 0.540 m, up to large trees with an area of over 500 m^2 and a height equal to 45.21 m. Summary statistics of the obtained ITCs can be found in Table 12.

The understory cluster (Figure 17 (a)), whose area corresponded to the whole plot surface, contained 694,634 points with an average height of 6.93 m. In addition to the latter cluster, the segmentation also produced the formation of other big groups of points (Figure 17 (b)) that were removed during the post-processing operations.These large set of points ranged from an area of 1215.86 m^2 to 4305.09 m^2 and comprended points up to 13.95 m. The exclusion of such clusters prevented the chance of matching them with a stem but on the other hand, it negatively affected the segmentation results removing a large number of points and thus understory trees. Compared to the segmentation results obtained by Kientz et al. (2019), which are an average over six different plots in Paracou, the obtained segmentation produced more ITCs (6723 vs 2564), with in general smaller crowns (mean area of 24.78 m^2 vs 32.46m^2).

Table 12: Summary statistics of Crown Area, Crown Diameter, and Tree Height of Adaptive Mean Shift 3D algorithm-segmented trees in the Area Of Study.

Summary statistic	Crown Area (CA)	Crown Diameter (CD)	Tree Height (TH)
Min	1.03 m^2	0.81 m	0.54 m
Mean	24.78 m^2	4.68 m	17.82 m
Max	690.316 m^2	29.21 m	45.21 m
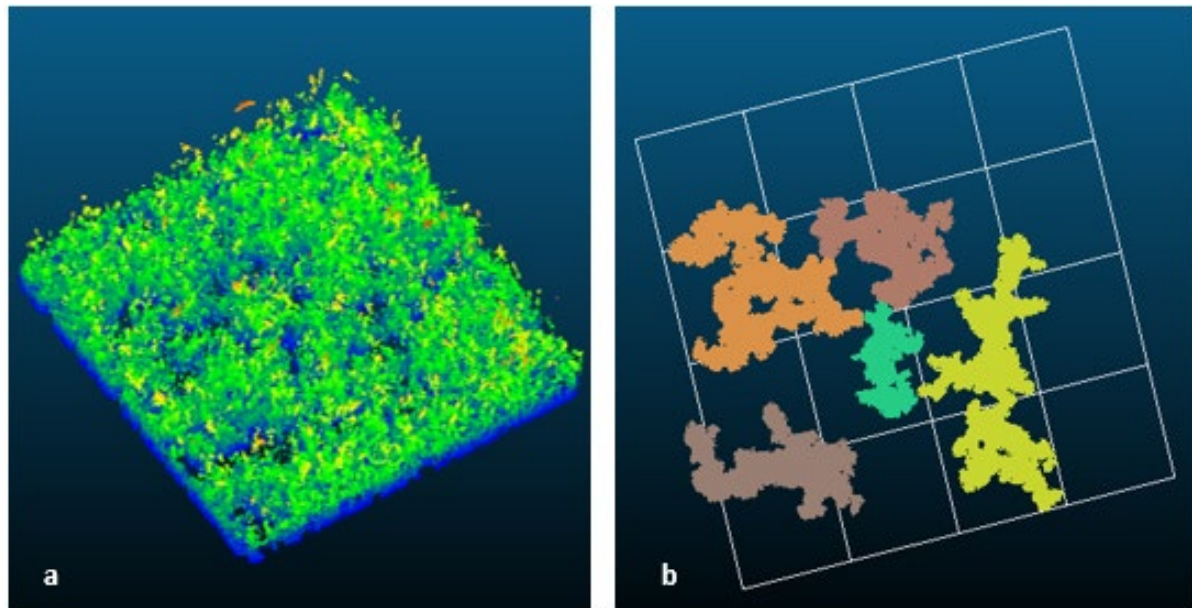
Figure 17: On the left the big, understory cluster of points is shown, on the right instead, the big clusters that were removed together with the understory points are plotted with the plot6 AOI and the quarter hectares boundaries.
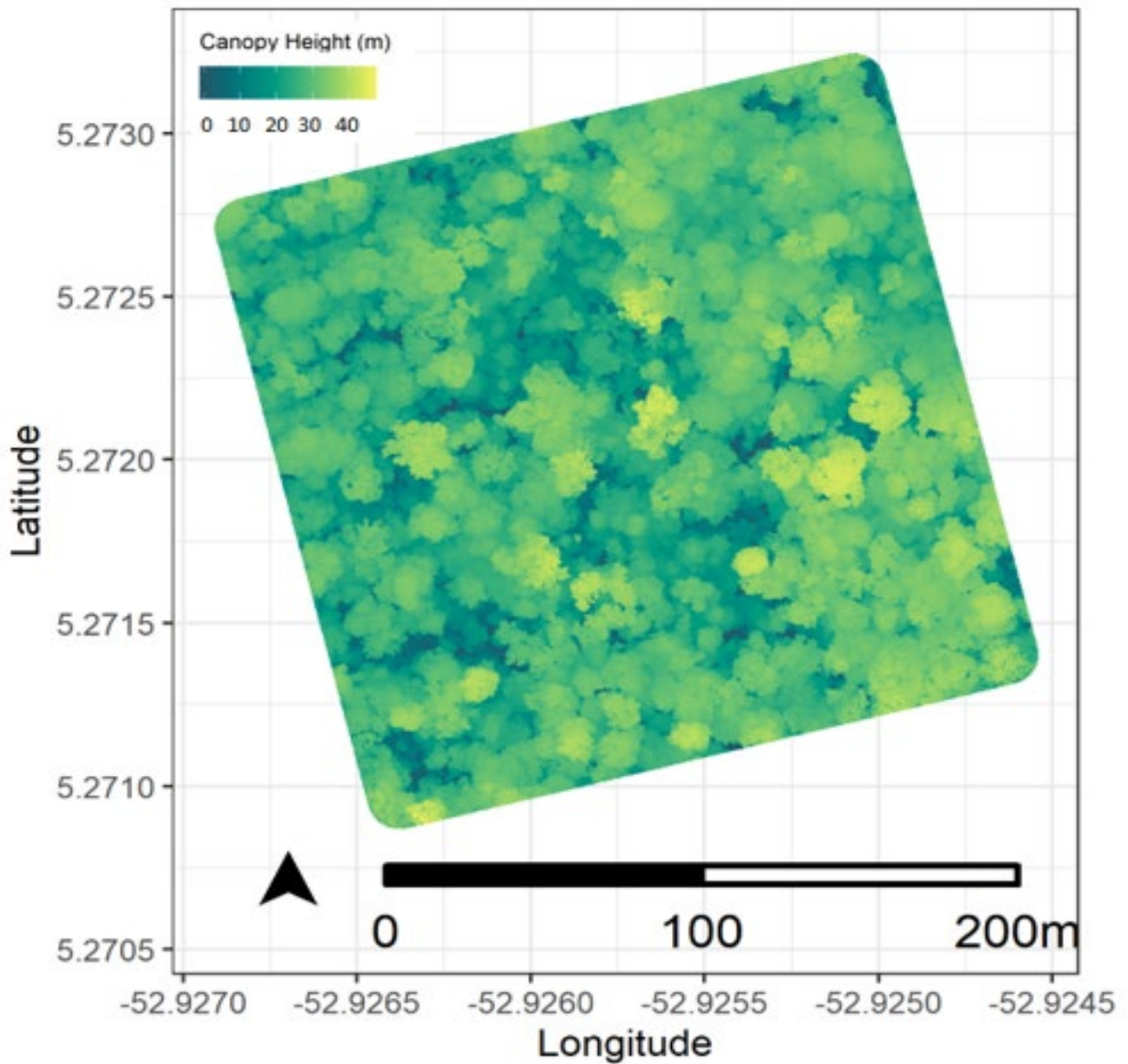
Figure 18: Spike-free Canopy Height Model of the Area Of Study.

The generated Spike-free Canopy Height Model (CHM_Sf), Figure 18, clearly shows the natural variability of the area of study. While few emergent trees were well-defined and easily detectable, the rest of the canopy shows a fairly constant height which makes it more complicated to identify and discern adjacent crowns. Observing the CHM_Sf some other patterns are clearly visible. While the canopy in the East portion of the plot appear very dense and on average higher than the rest of the area, the vegetation in the central section looks much sparser and shorter. The area, which stretches from South-East to North-West, interestingly follows the course of a small creek that crosses the forest in the same direction. The presence of the stream, which is located at the bottom of a small valley may be somehow related to the presence of a sparser vegetation. On the other hand, when inspecting the CHM_Sf in more detail, few more observations can be made. Figure 19 shows a portion of the spike-free CHM (a) compared to the

39

standard CHM (b) and the difference between the two (c). The standard CHM presents numerous dark pixels that cause noise in the raster. These pits, clearly visible as dark pixels, are produced when, in the case of LiDAR data, laser beams penetrate the canopy and are reflected by lower objects or the ground. When converting a point cloud to raster, due to the high resolution of the raster (0.10 m in this case) some pixels get assigned with a low Z value even if they are located within a high canopy. As a consequence, the final raster looks noisy and would negatively affect those segmentation algorithms that rely on the CHM. Compared to the standard CHM, the spike-free version looks smoother and the crown looks more defined. However the vegetation features do not appear simplified and result in a fair representation of the canopy.
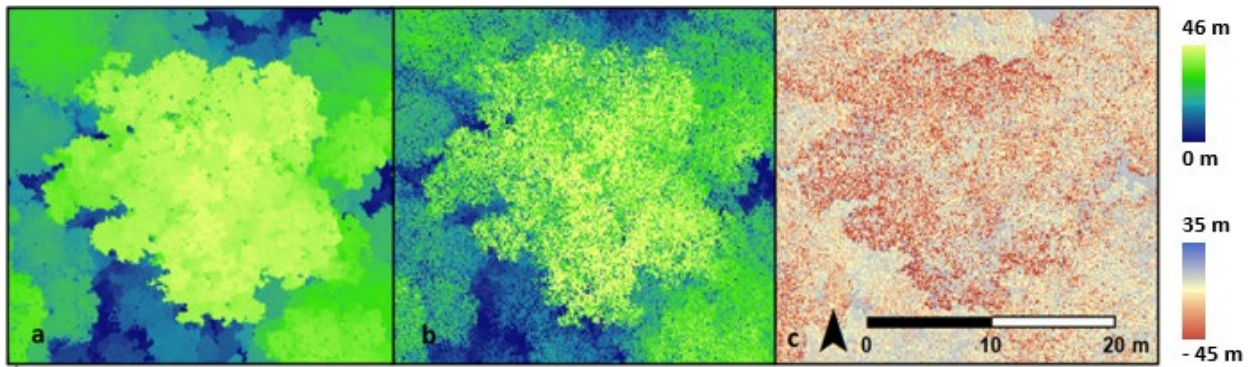
Figure 19: Visual comparison of the Spike-free CHM (a) and the standard CHM (b), and their difference (c) for a test tree.

To better quantify the difference between the two models, the raster histograms and summary statistics were generated and are presented in Figure 20 and Table 13. While the minimum and maximum values of the two CHM versions differ of few centimetres, the quantiles and the median values help getting more insights of the difference between the two. In fact, the standard CHM presents many more low-elevation pixels, which can be thought as pits in the raster. The difference raster instead, is useful to have an indication of height variation between same pixels in the two rasters. Figure 10 (c) and Table 9 show that pits reached a depth of 45 m, indicating that the LiDAR laser beam probably reached the ground in that specific location. On the other hand, the presence of spikes can also occur but, as the histogram show, it is less common.
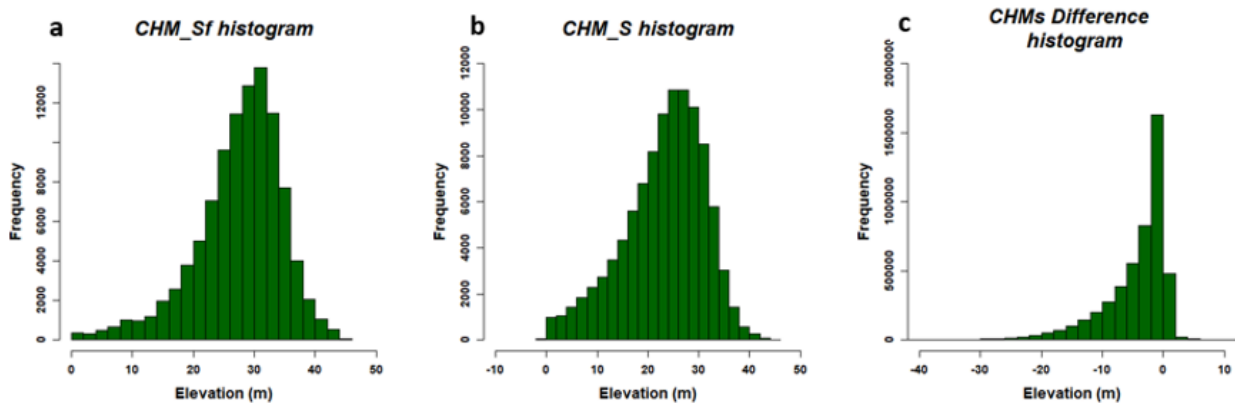
Figure 20: Frequency histograms of pixels height of the Spike-free Canopy Height model (a), the standard Canopy Height Model (b), and the model of their difference (c).

Statistic	CHM_Sf	CHM_St	CHM_Diff
Min	0.00	- 0.04	- 45.44
1st Quantile	23.92	18.41	- 6.46
Median	28.63	24.32	- 2.57
3rd Quantile	32.30	28.95	- 0.53
Max	46.33	46.21	35.40

Generating the spike-free model on single tiles, as shown in Figure 21, cause some of the points at the edge to have lower heights compared to the rest of the scene. When the tiles were merged and the whole scene was successively converted to raster, these points caused the formation of pixels with low values at the edges of tiles. These height differences in pixel height, produced few artefacts in the final raster which are detectable as straight lines, Figure 22 (a). These artefacts in some point reached the depth of 1 meter, as it is shown by the vertical profile of a sample transect, Figure 22 (b).
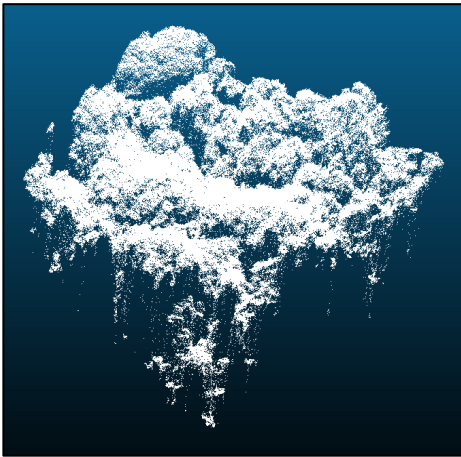
Figure 21: Spike-Free point cloud of a test tile showing the low points at the edges.
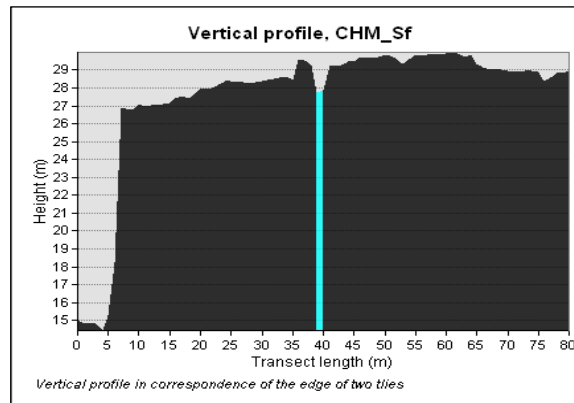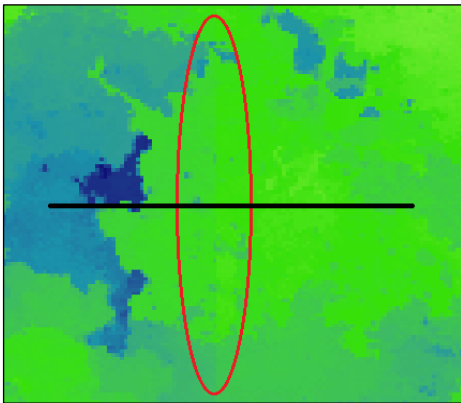
Figure 22: Transect over a artefact section (a) and vertical profile of the transect with the location of the artefact highlighted in light blue (b).
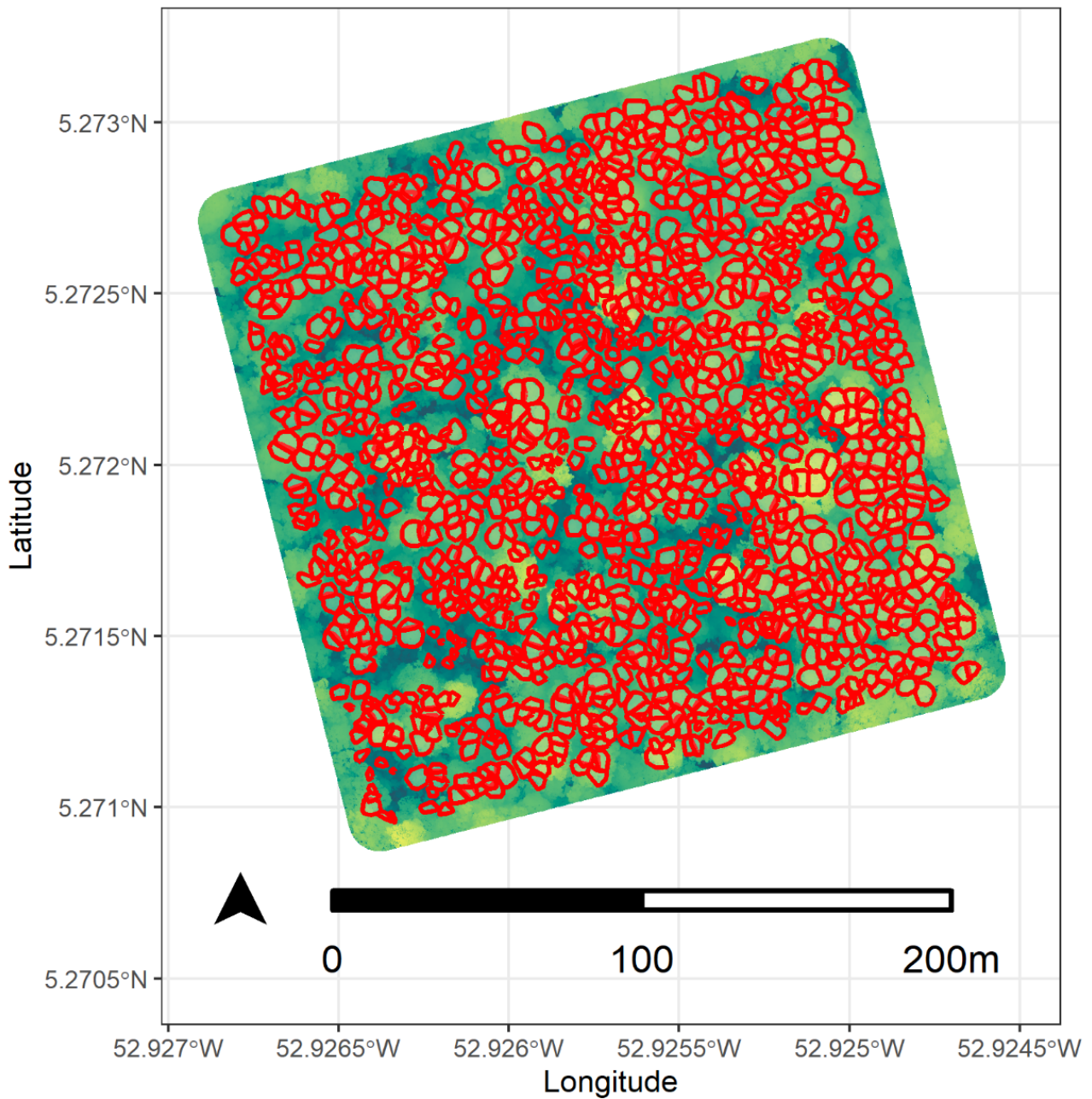
Figure 23: itcSegment-derived Individual Tree Canopies (red polygons) displayed above the Spike-free Canopy Height Model (CHM_Sf)

The itcSegment algorithm generated in total 1858 ITCs, 346 of which were removed because not intersecting the plot, returning 1512 trees within the AOI (Figure 23). As already explained in the Methodology section, the algorithm operates on the CHM, and thus, it is only able to identify tree crowns located in the upper canopy. This is confirmed by the fact that the average TH of segmented canopies equals to 32.150 m, while the lowest identified tree had a height on 12.090 m. For such high trees it would be normal to expect large canopies as well, but this is not the case. As a matter of fact, as presented by the ITCs summary statistics, (

Table 14)it is evident how the segmentation generally produced small crowns, with a maximum area of 49.220 m², corresponding to a diameter of 7.872 m. As a result, large trees appear over-segmented as it is clear in Figure 24 where a large, emergent canopy has been divided in multiple smaller polygons. Furthermore, it is also important to mention that itcSegment was not able to completely segment the scene, segmenting only 27.84 % of the total 4 ha area.

Table 14: Summary statistics of Crown Area, Crown Diameter, and Tree Height of itcSegment-segmented trees in the Area Of Study.

Summary statistic	Crow Area (m²)	Crow Diameter (m)	Tree Height (m)
Min	1.02	1.11	12.09
Mean	15.55	4.17	32.15
Max	49.22	7.87	45.03
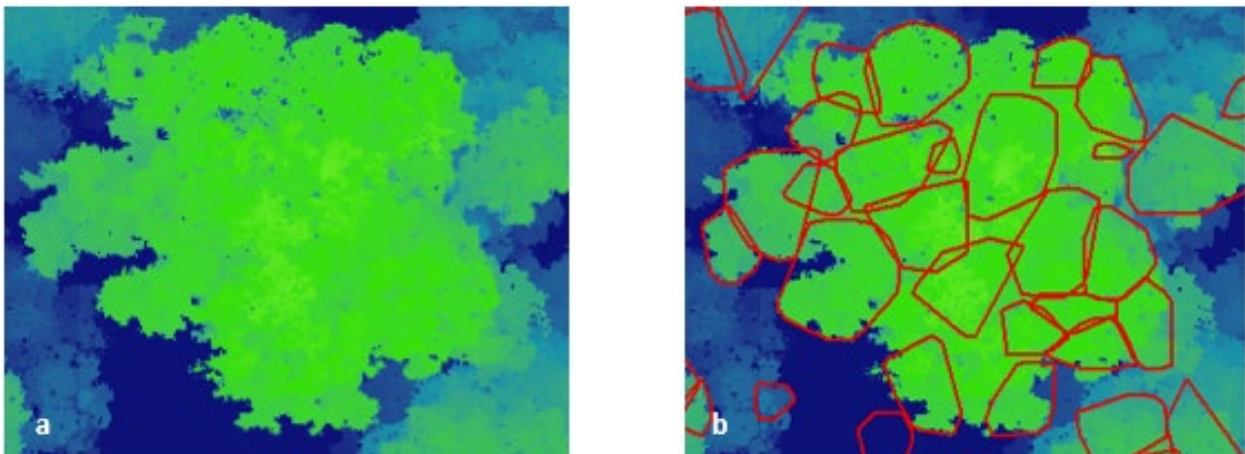
Figure 24: (a) Portion of the Spike-free Canopy Height Model showing an emergent canopy tree, and (b) itcSegment-segmented trees (red polygons).

Watershed segmentation:
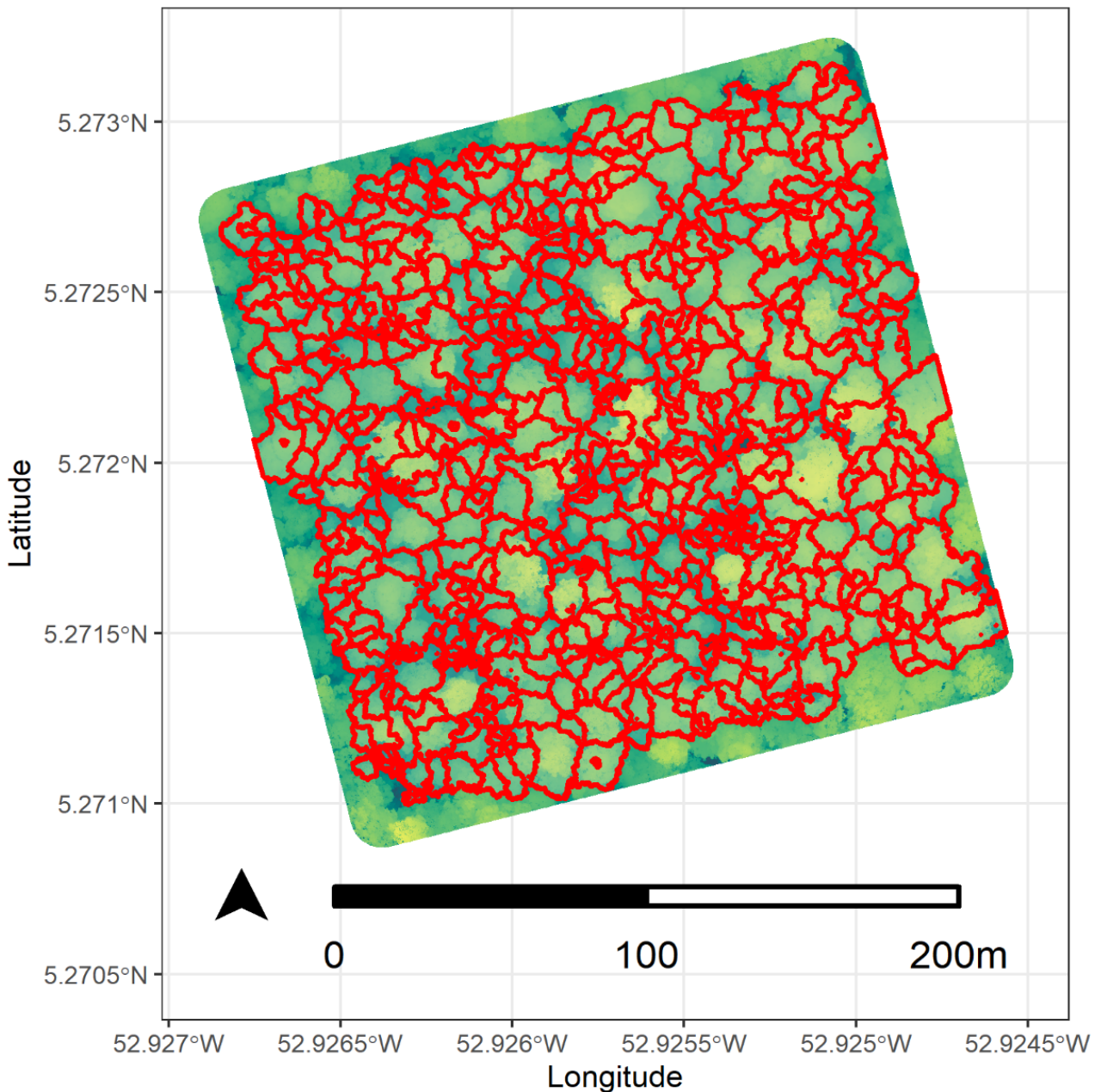
Figure 25: Watershed-derived Individual Tree Canopies (red polygons) displayed above the Spike-free Canopy Height Model (CHM_Sf).

The last method that was tested is a marker-based watershed segmentation. It segmented 810 individual canopies, 180 were filtered out since not intersecting the study area, or because they had an area smaller than 1 m^2 (Figure 25). Table 15 briefly presents the segmented ITCs summary statistics. As it can be noticed, the watershed algorithm, which also relies on the CHM, produced large crowns, and segmented a higher portion of the AOI than itcSegment. It is interesting to notice that, compared to the other CHM-based algorithm, the watershed method segmented crows with comparable heights but much larger area with a maximum area of 503.86 m^2. The reason of such big crowns (mean CA = 63.68 m^2), lies in the algorithm itself, which is programmed to segment the whole area, excluding only those pixels with a height smaller than a user-defined

threshold. In the study the threshold was set to 5 m in order to exclude understory features. Consequently, some of the segmented ITCs extended all the way to the border of the CHM resulting in unnatural shapes with straight sides. In general, the algorithm undersegmented the scene producing large ITCs as already explained, altough some crows were divided into multiple polygons, Figure 26 (a). In fact, the main driver of the watershed segmentation is the capability of the algorithm to accurately locate tree tops. When more than one apex is found within a crown, the tree will result divided into multiple ITCs, on the other hand, if less tree tops than present are identified, the segmentation will return large polygons which includes multiple trees, Figure 26 (b). Despite the application of a gaussian smoothing filter to avoid the localization of spurius maxima, multiple treetops were found within the same canopy. This represents one of the main limitation of the algorithm, especially in tropical forests where trees, due to the lack of apical dominance are prone to cause the identification of incorrect treetops.

Table 15: Summary statistics of Crown Area, Crown Diameter, and Tree Height of Watershed-segmented trees in the Area Of Study.

Summary statistic	Crow Area (m^2)	Crow Diameter (m)	Tree Height (m)
Min	1.09 m^2	1.11 m	11.91 m
Mean	63.68 m^2	7.74 m	31.04 m
Max	503.86 m^2	25.22 m	45.03 m
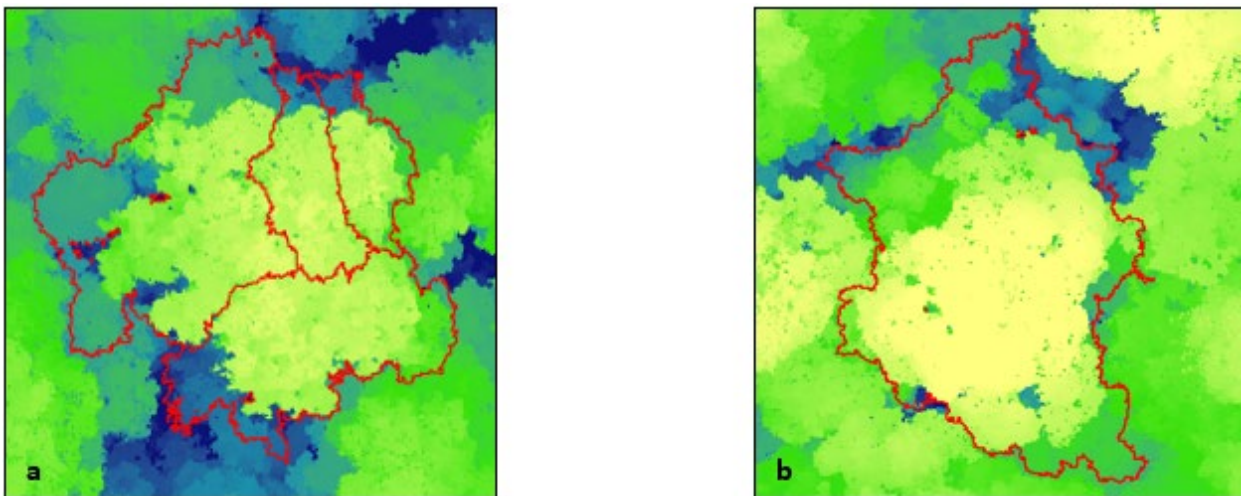
Figure 26: A portion of the Spike-free Canopy Height Model showing an emergent canopy that was over-segmented (a), and different emergent canopies that were instead under-segmented (b).

Algorithms comparison

Validation thresholds selection

As already presented in the relative section (Segmentation algorithms validation), the developed framework had to implement two user-defined thresholds to assess the segmentation results, namely Thresh_D, and Thresh_H. The choice of these two parameters can strongly affect the validation effectiveness and thus must be carefully made.

In the study, to examine the influence of the buffer size when searching for possible ITC-stem matches, three different radii were implemented, respectively of 0.5 m, 1 m, and 2 m. As shown in, the selection of Thresh_D significantly impacted the validation results returning a larger number of TP for larger thresholds. Small values in fact limit the number of possible matches with the chance that also correct ITC will be excluded. On the other hand, large values, increasing the allowed centroid-stem distance can lead to wrong matches. In order to choose the most appropriate threshold, the segmentation results were visually analysed searching for ITCs that accurately represented a crown, and the distance between the ITC centroid and the corresponding stem was measured. For some of the inspected ITCs this distance was found to vary between 1.5 m and 2 m. Additionally, considering the uncertainty related to the location of stems, that was found to be +/- 2 m (Aubry-Kientz et al., 2019), the final threshold was set equal to 2 m.

The same approach was also followed for the selection of the second threshold. The ITCs and stems that were found to satisfy the distance constraint were examined and the discrepancy between the two height measures was calculated. After the inspection of several a threshold of 15 m was determined to be more suitable than the other tested value of 10m.

The obtained statistics, relative to all segmented ITCs, based on the tested Thresh_D and Thresh_H values are respectively presented in Table 16: and Table 17. Examining the F-scores values it is clear how larger Thresh_D values resulted in a more consistent improvement than when increasing Thresh_H, confirming the higher influence of the distance threshold than the one-off Thresh_H. It is also interesting to notice that all the three algorithms showed the greatest improvement in F1 values when increasing Thresh_D from 1 m to 2 m which can be explained by the higher chance to match ITCs with a stem. A complete version of the confusion matrix can be found in Appendix 3.

Table 16: Validation statistics for each segmentation algorithm obtained at different Distance Thresholds.

Segmentation algorithm	Distance Threshold	ERR	ACC	PREC	REC	F1
AMS3D_	0.5	0.7711	0.2289	0.0332	0.0834	0.0475
AMS3D_	1	0.7563	0.2437	0.1122	0.2815	0.1604
AMS3D_	2	0.7042	0.2958	0.2378	0.5967	0.3400
itc_Segment_	0.5	0.6190	0.3810	0.0304	0.0201	0.0242
itc_Segment_	1	0.6066	0.3934	0.1171	0.0773	0.0931
itc_Segment_	2	0.5838	0.4162	0.2784	0.1838	0.2214
Watershed_	0.5	0.5589	0.4411	0.0238	0.0065	0.0103
Watershed_	1	0.5554	0.4446	0.0905	0.0249	0.0390
Watershed_	2	0.5272	0.4728	0.3333	0.0917	0.1438

Table 17: Validation statistics for each segmentation algorithm obtained at different Height Thresholds.

Segmentation algorithm	Height Threshold	ERR	ACC	PREC	REC	F1
AMS3D_	15	0.7378	0.2622	0.2079	0.5216	0.2973
AMS3D_	10	0.7042	0.2958	0.2378	0.5967	0.3400
itc_Segment_	15	0.6325	0.3675	0.1812	0.1196	0.1441
itc_Segment_	10	0.5838	0.4162	0.2784	0.1838	0.2214
Watershed_	15	0.5555	0.4445	0.2127	0.0585	0.0917
Watershed_	10	0.5272	0.4728	0.3333	0.0917	0.1438

Segmentations comparison

Among the tested algorithms, AMS3D was the only one that segments the whole 3D point cloud. itcSegment and Watershed, often referred to as raster-based methods, first need to convert the 3D data into a Canopy Height Model and then perform the segmentation onto the obtained raster. The different nature of AMS3D allows sub-canopy features identification while in the other cases only trees belonging to the upper forest layer, and hence visible in the CHM model, can be extracted. As a result, the AMS3D was able to segment a considerably higher number of ITC and represents a more valid option when operating in complex, dense, multi-layered forest as it is the case of the French Guyanese tropical forest. Although the raster-based methods are significantly faster than AMS3D, they produced scarcer results, confirming their unsuitability on complex forests. Several studies have proved that CHM-based algorithms have higher accuracy in coniferous forests rather than deciduous or mixed ones (Duncanson et al., 2014; Kwak et al., 2007; Vega et al., 2014). The main limitation of such algorithms was found to be the identification of tree apices which strongly affect the number and size of the final segmented ITCs. Despite the implementation of smoothing strategies to eliminate spurious maxima in the CHM, and the use of variable moving windows to search for local maxima, both algorithms were not able to accurately detect treetops. On the other hand, the AMS3D algorithm does not rely on treetops identification but segments the 3D point cloud in an iterative process, clustering points that converge to the same mode, which represents the tree apex. The method, which is also able to distinguish intersecting canopies, was found to return the overall best results in terms of matched stems, segmented crowns similarity, and finally AGB estimation.

Regarding the implementation of the three algorithms, some additional remarks are necessary. While the itcSegment directly generated the ITC polygons, the watershed segmentation returned the segmented scene in raster format that had to be converted into polygons. The conversion however did not have any remarkable repercussion on the ITCs shapes. In the case of AMS3D instead, further postprocessing operations were required.

The three tested algorithms produced very different numbers of segmented trees, with AMS3D segmentation producing almost four times more ITCs than itcSegment and nine times more than Watershed segmentation. The AMS3D algorithm also returned the highest percentage of True Positives, matching 59.67 % of stems in the area of study, whereas the other two methods respectively matched only 18.38 % and 9.17 % of stems. Another interesting consideration regards the percentage of segmented area (sum of ITC surface areas), which is in line with the different algorithm approaches. The AMS3D, which performs the segmentation on the whole point cloud, and thus can identify trees in the under canopy, segmented a surface of about 3.5

times the area of study. The watershed algorithm instead, segmented a slightly bigger surface than the original plot. This can be explained by the fact that the segmentation is carried out on the whole CHM with the exception for pixels lower than 5 m. At the same time some canopies extended beyond the plot borders, returning a final percentage of 100.31%. Different results were obtained with the itcSegment algorithm which only managed to segment 58.8% of plot surface.

When assessing the three methods based on the Jaccard Index computation, AMS3D again clearly outperformed the other two methods with a mean value of 0.3192, which is anyway lower than 0.5, threshold value to consider a crown as correctly segmented. Among the fifteen reference trees, AMS3D matched 3 of them, itcSegment 0, and watershed algorithm 2. Compared to the congruence analysis performed by Kientz et al. (2019), the percentage of crowns with a match are considerably lower for both AMS3D and itcSegment, which it can be partly explained by the different size of reference datasets (15 vs 1598 manually segmented trees). The information is summarized in Table 18.

Table 18: For the three segmentation algorithm, number of segmented trees (N°ITCs), percentage of correctly segmented trees (True Positive), percentage of segmented area (Area Segmented), mean Jaccard Index, and percentage of crowns with a Jaccard Index > 50% (Crowns with a match).

Segmentation algorithm	N° ITCs	True Positive (%)	Area Segmented	Mean Jaccard Index	Crowns with a match (Jaccard Index > 50 %)
AMS3D	5756	59.67	356.65%	0.3192	20.00 %
ItcSegment	1512	18.38	58.80 %	0.1329	0.00 %
Watershed	630	9.17	100.31 %	0.1899	13.33 %

Comparing the confusion matrices and statistics of each segmentation confirmed again the higher performances achieved by AMS3D algorithm, which outperformed the other two methods based on F-score values. Despite the highest F-score, obtained when considering the whole dataset, reached a value of 0.34, the value is still low and indicates rather poor results when matching ITCs and reference stems. In general, all methods returned better results when assessing the whole dataset, with superior statistics for high trees (TH > 20 m) than short ones (TH < 20 m). ItcSegment and Watershed, the two CHM-based algorithms presented highly different results for the two subsets, with very low F-scores for the subcanopy layer. On the other hand, the AMS3D returned similar values for the two datasets, with an F-score of 0.299 for short trees and 0.322 for taller ones, confirming the ability of the algorithm in segmenting subcanopy vegetation. Observing Table 19 it can be noticed that in most of the cases AMS3D obtained the highest Error rate and the lowest Accuracy and Precision values. Nevertheless, when considering Recall and F-score it returned the highest values among the three methods. The reason of that lies in the number of identified ITCs which is far higher than the number of reference stems. On the other hand, the higher percentage of ITCs assessed as True Positive and thus the fewer not-matched stems (False Negative) led to higher Recall values and consequently F-scores.

Segmentation algorithm	Subset	Error Rate	Accuracy	Precision	Recall	F-Score
AMS3D	All_Data	0.7042	0.2958	0.2378	0.5967	0.3400
AMS3D	TH_u20	0.6617	0.3383	0.2382	0.4020	0.2992
AMS3D	TH_o20	0.5513	0.4487	0.3569	0.2933	0.3220
itc_Segment	All_Data	0.5838	0.4162	0.2784	0.1838	0.2214
itc_Segment	TH_u20	0.5021	0.4979	0.1852	0.0022	0.0043
itc_Segment	TH_o20	0.5821	0.4179	0.2803	0.1816	0.2204
Watershed	All_Data	0.5272	0.4728	0.3333	0.0917	0.1438
Watershed	TH_u20	0.4996	0.5004	0.4545	0.0065	0.0129
Watershed	TH_o20	0.5277	0.4723	0.3266	0.0851	0.1350

In addition to compare the algorithms based on the validation results, a visual inspection of the different outputs was also performed. Figure 27 shows how one emergent tree with several smaller trees in the surroundings were segmented. The Figure confirms the unique ability of AMS3D to detect and distinguish understory features whereas itcSegment and Watershed, computing the segmentation on the CHM, group all points that fall below an ITC in the same cluster. Figure 27 (d,e,f) clearly shows that the main canopy was oversegmented in all the three cases, with several ITCs polygons within the canopy. ItcSegment clearly performed the worst, dividing the tree canopy in several small ITCs. The watershed algorithm instead produced bigger ITCs reducing the oversegmentation of the scene but including parts of adjacent trees. Conversely AMS3D was the algorithm which oversegmented the less the emergent canopy, detecting and distinguishing also submerged trees, even though a level of oversegmentation is also present in the lower layer.
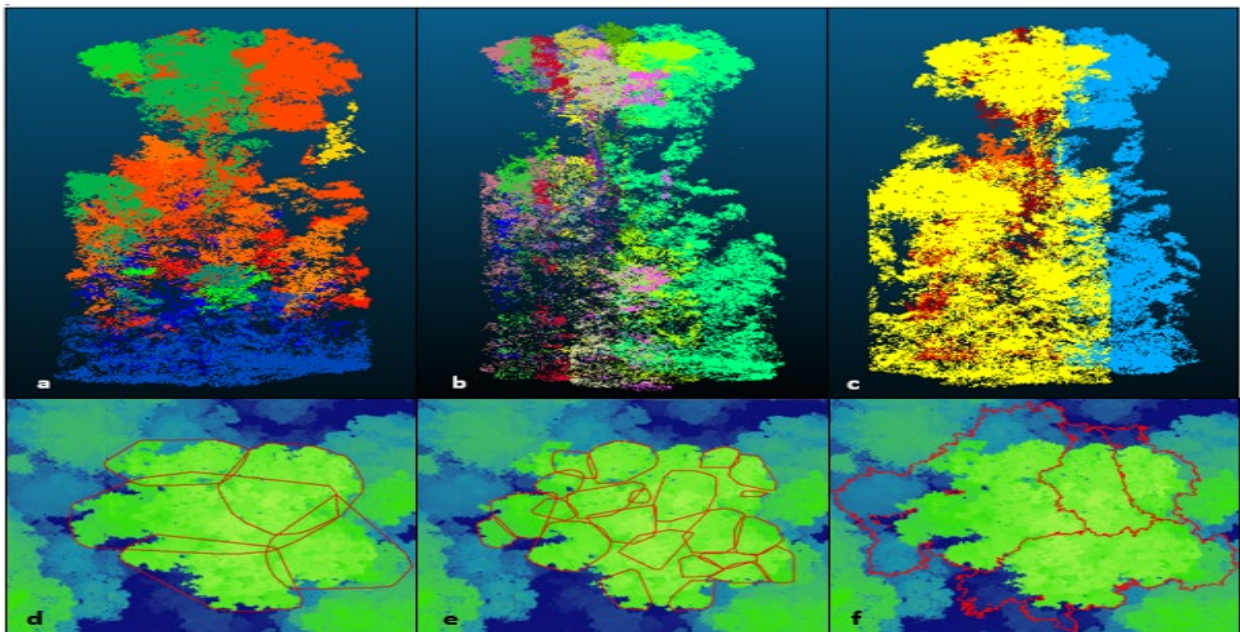

Figure 27: Segmentation of one emergent tree and adjacent trees obtained by the three different segmentation algorithms. Tree point cloud, and crown polygons segmented with: Adaptive Mean Shift 3D algorithm (a, d), itcSegment algorithm (b, e), and Watershed algorithm (c, f).

Above Ground Biomass comparison

At first, the amount of Above Ground Biomass (AGB) was estimated for the whole AOI and compared among algorithms. The three methods produced highly different AGB estimates, from 554.17 Mg of itcSegment up to 2169.7 Mg of AMS3D. Comparing the predicted and reference estimates gave a first indication of under- or overestimation at plot scale. The itcSegment algorithm was the one that deviated most from the -67.29%. Both watershed and AMS3D performed better. In the first case the algorithm caused an underestimation of 49.59%, whereas the AMS3D overestimated plot AGB of 28.05%, producing the best results. Estimated, reference and error values are summarized in Table 20.

Table 20: Estimated Above Ground Biomass (AGB_Est), reference Above Ground Biomass, absolute error (ABG_Err), and relative error (AGB_Rel_Err) for the whole Area Of Study obtained based on the different segmentation algorithm outputs.

Segmentation Algorithm	AGB_Est (Mg)	AGB_Ref (Mg)	AGB_Err (Mg)	AGB_Rel_Err (%)
AMS3D	2169.7	1694.37	475.33	28.05
itcSegment	554.17	1694.37	-1140.2	-67.29
Watershed	854.04	1694.37	-840.33	-49.59

Besides comparing the AGB estimates at plot scale, the AGB similarity was also tested at tree level. For this purpose only those ITCs that were correctly matched with a stem (True Positives) were considered. Furthermore, to better analyse the ability of tested algorithms when segmenting trees of different sizes, AGB was compared at four different DBH classes. The obtained results, shown in Table 21, depict a clear picture where AMS3D returned the most similar AGB estimates for three out of four DBH classes. That was the case for trees with DBH smaller than 40 cm and for ones with DBH larger than 60 cm. For the remaining class (DBH between 40 and 60 cm) the watershed algorithm returned the most comparable estimates to the reference data. In the case of AMS3D, the retrieved estimates overestimate the amount of AGB of 93.14 % for the smallest DBH class, and 10.99 % for trees with DBH between 20 and 40 cm. Conversely, the estimates returned a Bias of -77 % for large trees (DBH > 60 cm). On the other hand, the watershed-derived AGB estimates underestimate the biomass of trees with a DBH between 40 and 60 cm of 30 %.

When interpreting the results, the prediction accuracy of the two allometric models should be carefully considered. While the allometric model developed by Jucker et al. (2016) was proved to return accurate estimates at tree level with a Bias of only -4.3 %, the pantropical model of Chave et al. (2014), tested on the same dataset, was found to estimate biomass with a higher Bias of 27.7 %. Based on this information the ITCs-derived AGB estimates can be expected to present lower values when compared to the reference ones. Nevertheless, the obtained results did not suggest such a relationship between estimated and reference AGB. This can be explained by the nature of the implemented allometric model which completely relies on the accurate extraction of tree parameters. As a matter of fact, each of the tested segmentation methods presented a certain level of over- or under-segmentation which directly affects the extraction of tree parameters, introducing a large level of uncertainty in AGB estimates. Even though the inclusion of crown diameter in the allometric equation contributes to reduce the problem of

over-segmentation at plot level (Dalponte and Coomes, 2016), inaccurate CD values strongly affect the estimates at tree level, eliminating all the advantages of tree-based AGB estimation. In addition to that, another drawback of inaccurate segmentation is represented by the high level of uncertainty that is introduced by the wrong detection of subcanopy trees, which may represent one of the most important error to correct in order to accurately measure AGB (Aubry-Kientz et al., 2019).

Table 21: Root Mean Square Error (RMSE) and percentual Bias (Bias) of Above Ground Biomass estimated on the different segmentation algorithm outputs, calculated based on different Diameter at Breast Height classes (DBH class) For each DBH class the best results (based on Bias) are highlighted in yellow.

Segmentation Algorithm	DBH class (cm)	RMSE (Mg)	Bias (%)
AMS3D	DBH < 20	0.4406	93.14
	20 < DBH < 40	1.4541	10.99
	40 < DBH < 60	2.4271	-46.75
	DBH > 60	5.9169	-77.09
itcSegment	DBH < 20	0.1945	101.60
	20 < DBH < 40	0.6189	-35.41
	40 < DBH < 60	2.4501	-80.54
	DBH > 60	6.8646	-90.62
Watershed	DBH < 20	1.3882	615.80
	20 < DBH < 40	1.9324	152.21
	40 < DBH < 60	1.9785	-30.42
	DBH > 60	6.3217	-81.66

Discussion and recommendations

Validation process

In order to validate the segmentation outputs and compare the performance of the tested algorithms a multi-step framework was developed as previously presented in the [Segmentation algorithms validation](#) section.

The first part of the validation process exploited the information contained in the stem map, assessing the segmented ITCs based on their location and height. The main advantage of the developed algorithm lies in its flexibility which allows the user to easily adjust the two thresholds in order to better define the validation process based on the structural features of the forest. In addition to that, the implementation of a second threshold based on tree height provides a simple but effective method to distinguish stacked trees, making the algorithm suitable for multi-layered forests as well. Nevertheless the algorithm is not error free and its accuracy largely depends on the threshold choice. The two thresholds should be carefully selected based on the nature of the input data and its level of uncertainty. It is thus suggested to visually assess the results and gradually calibrate the parameters before performing the final validation.

In the second part of the algorithm validation, the segmented crowns were instead assessed based on their similarity with reference, manually segmented crowns. The calculation of the Jaccard Index provides a meaningful and easy-to-implement approach to evaluate the accuracy of crowns delineation and it does not present significant drawbacks. To obtain a more exhaustive indication of the algorithm performance, despite the time-consuming nature of the operation, it is suggested to create a large reference dataset so that a consistent number of crowns can be evaluated, and more significant results can be obtained. Furthermore, when testing a segmentation algorithm based on the 3D point cloud, even though submerged trees are difficult to identify and extract, their inclusion in the validation is strongly recommended since it can unearth valuable insights regarding the segmentation of subcanopy layers.

Segmentation comparison

As implemented in the current study, among the tested algorithms, AMS3D was found to perform the most accurate segmentation of the forest stand. The results confirm the benefit of performing the segmentation on the 3D point cloud, which allows the identification of subcanopy trees. The segmentation on the whole point cloud is however computational intense, time expensive and can still produce a significant over-segmentation of large crowns. When working with large 3D datasets is suggested to down-sample the point cloud based on the project goals and requirements.

On the other hand, even though CHM-based algorithms are considerably faster and required less data postprocessing than the first method, they can only segment emergent canopies, failing in the identification of all those trees that are not represented in the CHM. The use of such algorithms is therefore unsuitable and should be avoided for complex forests where many trees is found in subcanopy layers.

Above Ground Biomass estimation

The estimation of AGB of single trees can strongly boost our knowledge of forest biomass providing unprecedent information at tree level. Nevertheless, at present the implementation of such approach still presents significant limitations which mainly depend on the segmentation accuracy. In fact, even though the implemented allometric equation, developed by Jucker et al's (2017), was found to accurately predict AGB of single trees with a Bias of only -4.3 %, the results obtained in the current study returned much larger uncertainties. This uncertainty is caused by the inaccurate segmentation of canopies which directly affects the extraction of tree attributes and thus the estimation of biomass. The high dependency of AGB estimation on the segmentation result confirms again that at present area-based method should still be preferred. Nevertheless, with the development of new and more efficient segmentation algorithms AGB estimation at tree level is expected to become the next state of the art technique.

Conclusion

The aim of the present research was to test the feasibility of individual tree segmentation methods on UAV-LS data to estimate AGB in a tropical multi-layered forest. The study compared the performance of three different algorithms, confirming the well-known limitations of ITD approaches for AGB estimation. The results also showed the promising results of 3D methods for segmenting a complex tropical forest, with the AMS3D algorithm that outperformed the other two. When assessing the results with the stem map the latter correctly segmented almost 60 % of trees whereas the second-best method only matched about 20 % of reference trees. The AMS3D also obtained the highest result when assessing crown similarity with manually segmented ones, correctly delineating 20 % of the reference trees.

The results clearly showed the advantage of algorithms based on the whole point clouds rather than segmenting solely on the CHM raster. During the study the influence of increasing point densities on AMS3D method was also tested and was proved to significantly improve trees detection and delineation. Regarding the estimation of forest biomass, the estimates obtained on AMS3D outcomes returned the smallest errors when compared to reference AGB values at plot level. When considering different tree sizes the algorithm also returned the smallest bias for three out of four DBH classes. The results of this research suggest that high-density UAV LiDAR data can improve tree detection and in general forest structure retrieval. Moreover, the flexibility of UAV platforms indicates that forest inventories based on such platforms may substitute other, more expensive and time-consuming field approaches. In addition to that, they may further boost the development of segmentation algorithms and make individual tree approaches a more valuable option to area-based methods for AGB estimation at small scales.

References

Aubry-Kientz, M., Dutrieux, R., Ferraz, A., Saatchi, S., Hamraz, H., Williams, J., Coomes, D., Piboule, A., Vincent, G., 2019. A Comparative Assessment of the Performance of Individual Tree Crowns Delineation Algorithms from ALS Data in Tropical Forests. Remote Sens. 11, 1086. https://doi.org/10.3390/rs11091086

Bonan, G.B., 2008. Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. Science 320, 1444–1449.

Brede, B., Lau, A., Bartholomeus, H., Kooistra, L., 2017. Comparing RIEGL RiCOPTER UAV LiDAR Derived Canopy Height and DBH with Terrestrial LiDAR. Sensors 17, 2371. https://doi.org/10.3390/s17102371

Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P., Herold, M., Culvenor, D., Avitabile, V., Disney, M., Armston, J., Kaasalainen, M., 2015. Nondestructive estimates of above-ground biomass using terrestrial laser scanning. Methods Ecol. Evol. 6, 198–208. https://doi.org/10.1111/2041-210X.12301

Chave, J., Coomes, D., Jansen, S., Lewis, S.L., Swenson, N.G., Zanne, A.E., 2009. Towards a worldwide wood economics spectrum. Ecol. Lett. 12, 351–366. https://doi.org/10.1111/j.1461-0248.2009.01285.x

Chave, J., Riéra, B., Dubois, M.-A., 2001. Estimation of Biomass in a Neotropical Forest of French Guiana: Spatial and Temporal Variability. J. Trop. Ecol. 17, 79–96.

Coomes, D.A., Dalponte, M., Jucker, T., Asner, G.P., Banin, L.F., Burslem, D.F.R.P., Lewis, S.L., Nilus, R., Phillips, O.L., Phua, M.-H., Qie, L., 2017. Area-based vs tree-centric approaches to mapping forest carbon in Southeast Asian forests from airborne laser scanning data. Remote Sens. Environ. 194, 77–88. https://doi.org/10.1016/j.rse.2017.03.017

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

Dalponte, M., Frizzera, L., Ørka, H.O., Gobakken, T., Næsset, E., Gianelle, D., 2018. Predicting stem diameters and aboveground biomass of individual trees using remote sensing data. Ecol. Indic. 85, 367–376. https://doi.org/10.1016/j.ecolind.2017.10.066

Drake, J.B., Knox, R.G., Dubayah, R.O., Clark, D.B., Condit, R., Blair, J.B., Hofton, M., 2003. Above-ground biomass estimation in closed canopy Neotropical forests using lidar remote sensing: factors affecting the generality of relationships. Glob. Ecol. Biogeogr. 12, 147–159. https://doi.org/10.1046/j.1466-822X.2003.00010.x

Duncanson, L., Armston, J., Disney, M., Avitabile, V., Barbier, N., Calders, K., Carter, S., Chave, J., Herold, M., Crowther, T.W., Falkowski, M., Kellner, J.R., Labrière, N., Lucas, R., MacBean, N., McRoberts, R.E., Meyer, V., Næsset, E., Nickeson, J.E., Paul, K.I., Phillips, O.L., Réjou-Méchain, M., Román, M., Roxburgh, S., Saatchi, S., Schepaschenko, D., Scipal, K., Siqueira, P.R., Whitehurst, A., Williams, M., 2019. The Importance of Consistent Global Forest Aboveground Biomass Product Validation. Surv. Geophys. 40, 979–999. https://doi.org/10.1007/s10712-019-09538-8

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 Sens. Environ. 154, 378–386. https://doi.org/10.1016/j.rse.2013.07.044

Eggleston, S., Buendia, L., Miwa, K., Ngara, T., Tanabe, K., n.d. 2006 IPCC Guidelines for National Greenhouse Gas Inventories.

ESA - Biomass [WWW Document], 2019. URL https://www.esa.int/Applications/Observing_the_Earth/The_Living_Planet_Programme/Earth_Explorers/Biomass (accessed 3.9.20).

Ferraz, A., Bretar, F., Jacquemoud, S., Gonçalves, G., Pereira, L., Tomé, M., Soares, P., 2012. 3-D mapping of a multi-layered Mediterranean forest using ALS data. Remote Sens. Environ. 121, 210–223. https://doi.org/10.1016/j.rse.2012.01.020

Ferraz, A., Saatchi, S., Mallet, C., Meyer, V., 2016. Lidar detection of individual tree size in tropical forests. Remote Sens. Environ. 183, 318–333. https://doi.org/10.1016/j.rse.2016.05.028

Gibbs, H.K., Brown, S., Niles, J.O., Foley, J.A., 2007. Monitoring and estimating tropical forest carbon stocks: making REDD a reality. Environ. Res. Lett. 2, 045023. https://doi.org/10.1088/1748-9326/2/4/045023

Gonzalez de Tanago, J., Lau, A., Bartholomeus, H., Herold, M., Avitabile, V., Raumonen, P., Martius, C., Goodman, R.C., Disney, M., Manuri, S., Burt, A., Calders, K., 2018. Estimation of above-ground biomass of large tropical trees with terrestrial LiDAR. Methods Ecol. Evol. 9, 223–234. https://doi.org/10.1111/2041-210X.12904

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

Herold, M., Carter, S., Avitabile, V., Espejo, A.B., Jonckheere, I., Lucas, R., McRoberts, R.E., Næsset, E., Nightingale, J., Petersen, R., Reiche, J., Romijn, E., Rosenqvist, A., Rozendaal, D.M.A., Seifert, F.M., Sanz, M.J., De Sy, V., 2019. The Role and Need for Space-Based Forest Biomass-Related Measurements in Environmental Management and Policy. Surv. Geophys. 40, 757–778. https://doi.org/10.1007/s10712-019-09510-6

Horning, N., 2008. Remote Sensing, in: Jørgensen, S.E., Fath, B.D. (Eds.), Encyclopedia of Ecology. Academic Press, Oxford, pp. 2986–2994. https://doi.org/10.1016/B978-008045405-4.00237-8

Houghton, R.A., 2005. Aboveground forest biomass and the global carbon balance. Glob. Change Biol. 11, 945–958.

Hyyppä, J., Hyyppä, H., Leckie, D., Gougeon, F., Yu, X., Maltamo, M., 2008. Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. Int. J. Remote Sens. 29, 1339–1366. https://doi.org/10.1080/01431160701736489

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., 2016. Allometric equations for integrating remote sensing imagery into forest monitoring programmes. Glob. Change Biol. 23, 177–190. https://doi.org/10.1111/gcb.13388

Khosravipour, A., Skidmore, A.K., Isenburg, M., 2016. Generating spike-free digital surface models using LiDAR raw point clouds: A new approach for forestry applications. Int. J. Appl. Earth Obs. Geoinformation 52, 104–114. https://doi.org/10.1016/j.jag.2016.06.005

Koch, B., Heyder, U., Weinacker, H., 2006. Detection of Individual Tree Crowns in Airborne Lidar Data. Photogramm. Eng. Remote Sens. 72, 357–363. https://doi.org/10.14358/PERS.72.4.357

Kwak, D.-A., Lee, W.-K., Lee, J.-H., Biging, G.S., Gong, P., 2007. Detection of individual trees and estimation of tree height using LiDAR data. J. For. Res. 12, 425–434. https://doi.org/10.1007/s10310-007-0041-9

Lau, A., Calders, K., Bartholomeus, H., Martius, C., Raumonen, P., Herold, M., Vicari, M., Sukhdeo, H., Singh, J., Goodman, R., 2019. Tree Biomass Equations from Terrestrial LiDAR: A Case Study in Guyana. Forests 10, 527. https://doi.org/10.3390/f10060527

Liu, P., 2015. A survey of remote-sensing big data. Front. Environ. Sci. 3. https://doi.org/10.3389/fenvs.2015.00045

Maltamo, M., Peuhkurinen, J., Malinen, J., Vauhkonen, J., Packalén, P., Tokola, T., 2009. Predicting tree attributes and quality characteristics of Scots pine using airborne laser scanning data. Silva Fenn. 43. https://doi.org/10.14214/sf.203

Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L., Canadell, J.G., Ciais, P., Jackson, R.B., Pacala, S.W., McGuire, A.D., Piao, S., Rautiainen, A., Sitch, S., Hayes, D., 2011. A Large and Persistent Carbon Sink in the World's Forests. Science 333, 988–993. https://doi.org/10.1126/science.1201609

Paris, C., Valduga, D., Bruzzone, L., 2016. A Hierarchical Approach to Three-Dimensional Segmentation of LiDAR Data at Single-Tree Level in a Multilayered Forest. IEEE Trans. Geosci. Remote Sens. 54, 4190–4203. https://doi.org/10.1109/TGRS.2016.2538203

Persson, A., Holmgren, J., Soderman, U., 2002. Detecting and measuring individual trees using an airborne laser scanner. Photogramm. Eng. Remote Sens. 68, 925–932.

Popescu, S.C., Wynne, R.H., n.d. Seeing the Trees in the Forest: Using Lidar and Multispectral Data Fusion with Local Filtering and Variable Window Size for Estimating Tree Height. Photogramm. Eng. 16.

Slik, J.W.F., Paoli, G., McGuire, K., Amaral, I., Barroso, J., Bastian, M., Blanc, L., Bongers, F., Boundja, P., Clark, C., Collins, M., Dauby, G., Ding, Y., Doucet, J.-L., Eler, E., Ferreira, L., Forshed, O., Fredriksson, G., Gillet, J.-F., Harris, D., Leal, M., Laumonier, Y., Malhi, Y., Mansor, A., Martin, E., Miyamoto, K., Araujo-Murakami, A., Nagamasu, H., Nilus, R., Nurtjahya, E., Oliveira, Á., Onrizal, O., Parada-Gutierrez, A., Permana, A., Poorter, L., Poulsen, J., Ramirez-Angulo, H., Reitsma, J., Rovero, F., Rozak, A., Sheil, D., Silva-Espejo, J., Silveira, M., Spironelo, W., ter Steege, H., Stevart, T., Navarro-Aguilar, G.E., Sunderland, T., Suzuki, E., Tang, J., Theilade, I., van der Heijden, G., van Valkenburg, J., Van Do, T., Vilanova, E., Vos, V., Wich, S., Wöll, H., Yoneda, T., Zang, R., Zhang, M.-G., Zweifel, N., 2013. Large trees drive forest aboveground biomass variation in moist lowland forests across the tropics: Large trees and tropical forest biomass. Glob. Ecol. Biogeogr. 22, 1261–1271. https://doi.org/10.1111/geb.12092

Tsitsi, T., B., 2016. Remote sensing of aboveground forest biomass: A review. Trop. Ecol. 57, 125–132.

Vaglio Laurin, G., Ding, J., Disney, M., Bartholomeus, H., Herold, M., Papale, D., Valentini, R., 2019. Tree height in tropical forest as measured by different ground, proximal, and remote sensing instruments, and impacts on above ground biomass estimates. Int. J. Appl. Earth Obs. Geoinformation 82, 101899. https://doi.org/10.1016/j.jag.2019.101899

Vega, C., Hamrouni, A., Mokhtari, S.E., Morel, J., Bock, J., Renaud, J.-P., Bouvier, M., Durrieu, S., 2014. PTrees: A point-based approach to forest tree extraction from lidar data. Int J Appl Earth Obs Geoinformation. https://doi.org/10.1016/j.jag.2014.05.001

Wallace, L., Lucieer, A., Watson, C., Turner, D., 2012. Development of a UAV-LiDAR System with Application to Forest Inventory. Remote Sens. 4, 1519–1543. https://doi.org/10.3390/rs4061519

Wallace, L., Lucieer, A., Watson, C.S., 2014. Evaluating Tree Detection and Segmentation Routines on Very High-Resolution UAV LiDAR Data. IEEE Trans. Geosci. Remote Sens. 52, 7619–7628. https://doi.org/10.1109/TGRS.2014.2315649

Wang, Y., Lehtomäki, M., Liang, X., Pyörälä, J., Kukko, A., Jaakkola, A., Liu, J., Feng, Z., Chen, R., Hyyppä, J., 2019. Is field-measured tree height as reliable as believed – A comparison study of tree height estimates from field measurement, airborne laser scanning and terrestrial laser scanning in a boreal forest. ISPRS J. Photogramm. Remote Sens. 147, 132–145. https://doi.org/10.1016/j.isprsjprs.2018.11.008

Wardle, P., Kaoneka, A.R., 1999. Perceptions and concepts of the importance of forests, in: World Forests, Society and Environment. Springer, pp. 43–56.

White, J.C., Coops, N.C., Wulder, M.A., Vastaranta, M., Hilker, T., Tompalski, P., 2016. Remote Sensing Technologies for Enhancing Forest Inventories: A Review. Can. J. Remote Sens. 42, 619–641. https://doi.org/10.1080/07038992.2016.1207484

Yu, X., Hyyppä, J., Holopainen, M., Vastaranta, M., 2010. Comparison of Area-Based and Individual Tree-Based Methods for Predicting Plot-Level Forest Attributes. Remote Sens. 2, 1481–1495. https://doi.org/10.3390/rs2061481

Zolkos, S.G., Goetz, S.J., Dubayah, R., 2013. A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. Remote Sens. Environ. 128, 289–298. https://doi.org/10.1016/j.rse.2012.10.017

Appendix 1

Table of Content of the zip file accompanying the thesis report

- Report (PDF)
- Midterm & Final presentation (PPTX)
- Input & Output Datasets
- Tables
- R Scripts
- Literature (PDFs)

Appendix 2

AMS3D Testing:

- **Confusion Matrix**

Table 22: Confusion matrix for each AMS3D test.

Test	Point density (pts/m²)	TP	FN	FP	TN	ERR	ACC	PREC	REC	F1
Test_11	1922.82	178	25	3541	87	0.9308	0.0692	0.0479	0.8768	0.0908
Test_11	1054.15	177	26	3015	87	0.9201	0.0799	0.0555	0.8719	0.1043
Test_11	350.89	166	37	2002	87	0.8896	0.1104	0.0766	0.8177	0.1400
Test_11	93.43	144	59	955	87	0.8145	0.1855	0.1310	0.7094	0.2212
Test_11	42.08	128	75	540	87	0.7410	0.2590	0.1916	0.6305	0.2939
Test_11	23.87	110	93	388	87	0.7094	0.2906	0.2209	0.5419	0.3138
Test_11	10.71	97	106	221	87	0.6399	0.3601	0.3050	0.4778	0.3724
Test_12	1922.82	168	35	1454	87	0.8538	0.1462	0.1036	0.8276	0.1841
Test_12	1054.15	176	27	1447	87	0.8486	0.1514	0.1084	0.8670	0.1928
Test_12	350.89	161	42	1267	87	0.8407	0.1593	0.1127	0.7931	0.1974
Test_12	93.43	152	51	880	87	0.7957	0.2043	0.1473	0.7488	0.2462
Test_12	42.08	131	72	551	87	0.7408	0.2592	0.1921	0.6453	0.2960
Test_12	23.87	112	91	415	87	0.7177	0.2823	0.2125	0.5517	0.3068
Test_12	10.71	98	105	227	87	0.6422	0.3578	0.3015	0.4828	0.3712
Test_13	1922.82	151	52	567	87	0.7223	0.2777	0.2103	0.7438	0.3279
Test_13	1054.15	160	43	579	87	0.7158	0.2842	0.2165	0.7882	0.3397
Test_13	350.89	151	52	574	87	0.7245	0.2755	0.2083	0.7438	0.3254
Test_13	93.43	139	64	558	87	0.7335	0.2665	0.1994	0.6847	0.3089
Test_13	42.08	133	70	458	87	0.7059	0.2941	0.2250	0.6552	0.3350

Test_13	23.87	115	88	392	87	0.7038	0.2962	0.2268	0.5665	0.3239
Test_13	10.71	100	103	232	87	0.6418	0.3582	0.3012	0.4926	0.3738
Test_15	1922.82	129	74	447	87	0.7069	0.2931	0.2240	0.6355	0.3312
Test_15	1054.15	139	64	423	87	0.6830	0.3170	0.2473	0.6847	0.3634
Test_15	350.89	132	71	436	87	0.6983	0.3017	0.2324	0.6502	0.3424
Test_15	93.43	124	79	391	87	0.6902	0.3098	0.2408	0.6108	0.3454
Test_15	42.08	122	81	320	87	0.6574	0.3426	0.2760	0.6010	0.3783
Test_15	23.87	101	102	260	87	0.6582	0.3418	0.2798	0.4975	0.3582
Test_15	10.71	89	114	139	87	0.5897	0.4103	0.3904	0.4384	0.4130
Test_16	1922.82	19	184	115	87	0.7383	0.2617	0.1418	0.0936	0.1128
Test_16	1054.15	158	45	606	87	0.7266	0.2734	0.2068	0.7783	0.3268
Test_16	350.89	154	49	623	87	0.7360	0.2640	0.1982	0.7586	0.3143
Test_16	93.43	153	50	634	87	0.7403	0.2597	0.1944	0.7537	0.3091
Test_16	42.08	149	54	607	87	0.7369	0.2631	0.1971	0.7340	0.3107
Test_16	23.87	137	66	490	87	0.7128	0.2872	0.2185	0.6749	0.3301
Test_16	10.71	116	87	323	87	0.6688	0.3312	0.2642	0.5714	0.3614
Test_17	1922.82	151	52	566	87	0.7220	0.2780	0.2106	0.7438	0.3283
Test_17	1054.15	159	44	578	87	0.7166	0.2834	0.2157	0.7833	0.3383
Test_17	350.89	150	53	561	87	0.7215	0.2785	0.2110	0.7389	0.3282
Test_17	93.43	138	65	507	87	0.7177	0.2823	0.2140	0.6798	0.3255
Test_17	42.08	128	75	423	87	0.6985	0.3015	0.2323	0.6305	0.3395
Test_17	23.87	126	77	393	87	0.6881	0.3119	0.2428	0.6207	0.3490
Test_17	10.71	105	98	248	87	0.6431	0.3569	0.2975	0.5172	0.3777
Test_18	1922.82	149	54	469	87	0.6891	0.3109	0.2411	0.7340	0.3630
Test_18	1054.15	159	44	462	87	0.6729	0.3271	0.2560	0.7833	0.3859
Test_18	350.89	150	53	466	87	0.6865	0.3135	0.2435	0.7389	0.3663
Test_18	93.43	124	79	279	87	0.6292	0.3708	0.3077	0.6108	0.4092
Test_18	42.08	107	96	159	87	0.5679	0.4321	0.4023	0.5271	0.4563
Test_18	23.87	87	116	111	87	0.5661	0.4339	0.4394	0.4286	0.4339
Test_18	10.71	69	134	46	87	0.5357	0.4643	0.6000	0.3399	0.4340

- *Jaccard Index*

Table 23: Jaccard Indices for each test at different point densities and for the different reference tree classes.

Test	Point density (pts/m²)	ID_ECT	Jacc_Ind_ECT	ID_MCT	Jacc_Ind_MCT	ID_SCT	Jacc_Ind_SCT
11	1922.82	4390	0.671074	5740	0.142707	3298	0.443498
11	1054.15	3681	0.144557	4877	0.137729	2789	0.431573
11	350.89	2587	0.155593	3403	0.119653	2069	0.548266
11	93.43	2509	0.83184	3035	0.157388	2245	0.412606
11	42.08	1839	0.817615	2165	0.132855	1705	0.468055
11	23.87	877	0.837844	1113	0.516795	745	0.428808
11	10.71	699	0.64996	837	0.300395	605	0.587576
12	1922.82	3527	0.671074	4199	0.464936	2931	0.443498
12	1054.15	11149	0.84648	11839	0.446982	10593	0.431573
12	350.89	6811	0.155593	7311	0.119653	6439	0.548266
12	93.43	1537	0.83184	2001	0.157388	1283	0.412606
12	42.08	3477	0.817615	3781	0.132855	3337	0.468055
12	23.87	3785	0.837844	4037	0.516795	3645	0.428808
12	10.71	2915	0.64996	3025	0.526837	2823	0.587576
13	1922.82	3455	0.671074	3789	0.464936	3175	0.443498
13	1054.15	9648	0.84648	9984	0.446982	9382	0.431573
13	350.89	6602	0.799812	6868	0.141464	6424	0.548266
13	93.43	3586	0.83184	3906	0.157388	3410	0.412606
13	42.08	5078	0.165754	5262	0.179524	4938	0.468055
13	23.87	4901	0.101016	5115	0.516795	4765	0.428808
13	10.71	5787	0.64996	5897	0.526837	5691	0.587576
15	1922.82	2737	0.654138	2959	0.377996	2521	0.217102
15	1054.15	2293	0.600389	2539	0.371279	2099	0.203923
15	350.89	1549	0.163673	1853	0.585162	1409	0.208485
15	93.43	5027	0.18231	5223	0.546027	4887	0.203736
15	42.08	6135	0.308368	6295	0.179524	6009	0.066951
15	23.87	7393	0.599489	7537	0.533774	7295	0.26438
15	10.71	8007	0.637204	8089	0.524147	7931	0.113066
16	1922.82	20805	0.08101	20887	0.119146	0	0
16	1054.15	0	0	0	0	20385	0.458083
16	350.89	11326	0.728866	11644	0.567909	11146	0.500368
16	93.43	6726	0.824776	7010	0.58397	6540	0.451404
16	42.08	1337	0.168107	1507	0.179524	1133	0.434415
16	23.87	12696	0.817623	13008	0.063468	12550	0.525196
16	10.71	8804	0.597819	9034	0.037068	8744	0.061727
17	1922.82	3779	0.671074	4119	0.464936	3481	0.443498
17	1054.15	13674	0.84648	14024	0.446982	13394	0.431573
17	350.89	10242	0.799812	10526	0.141464	10054	0.548266
17	93.43	8827	0.83184	9189	0.157388	8623	0.412606
17	42.08	8076	0.165754	8352	0.244561	7876	0.468055
17	23.87	1403	0.837844	1709	0.083683	1177	0.428808
17	10.71	14932	0.64996	15322	0.09728	14702	0.587576

18	1922.82	12904	0.671074	13214	0.464936	12656	0.443498
18	1054.15	16284	0.84648	16574	0.446982	16036	0.431573
18	350.89	1425	0.799812	1637	0.141464	1387	0.039515
18	93.43	10284	0.83184	10438	0.547725	10176	0.412606
18	42.08	2272	0.165754	2376	0.558174	2210	0.468055
18	23.87	13485	0.837844	13585	0.516795	13433	0.428808
18	10.71	13807	0.64996	13833	0.526837	13769	0.587576

Appendix 3

Algorithms validation:

- **Confusion Matrix**

Table 24: Confusion matrix and statistics obtain on each segmentation algorithm output and on different subsets of data.

Segmentation	Subset	TP	FN	FP	TN	ERR	ACC	PREC	REC	F1
AMS3D	All_Data	1367	924	4382	862	0.7042	0.2958	0.2378	0.5967	0.34
AMS3D	TH_u20	921	1370	2945	1285	0.6617	0.3383	0.2382	0.402	0.2992
AMS3D	TH_o20	672	1619	1211	1631	0.5513	0.4487	0.3569	0.2933	0.322
itc_Segment	All_Data	421	1870	1091	1690	0.5838	0.4162	0.2784	0.1838	0.2214
itc_Segment	TH_u20	5	2286	22	2284	0.5021	0.4979	0.1852	0.0022	0.0043
itc_Segment	TH_o20	416	1875	1068	1697	0.5821	0.4179	0.2803	0.1816	0.2204
Watershed	All_Data	210	2081	420	2033	0.5272	0.4728	0.3333	0.0917	0.1438
Watershed	TH_u20	15	2276	18	2283	0.4996	0.5004	0.4545	0.0065	0.0129
Watershed	TH_o20	195	2096	402	2041	0.5277	0.4723	0.3266	0.0851	0.135

- *Jaccard Index*

Table 25: Jaccard Indices obtained from each segmentation algorithm for all three reference datasets.

Segmentation	ITC_class	ID	Jaccard Index
AMS3D	ECT	17985	0.1064
AMS3D	ECT	21077	0.7198
AMS3D	ECT	14807	0.2006
AMS3D	ECT	13599	0.0945
AMS3D	ECT	21703	0.7792
AMS3D	MCT	11735	0.1718
AMS3D	MCT	16293	0.7976
AMS3D	MCT	20645	0.2064
AMS3D	MCT	21477	0.4421
AMS3D	MCT	11789	0.7280
AMS3D	SCT	20857	0.1610
AMS3D	SCT	14959	0.1815
AMS3D	SCT	14317	0.1469
AMS3D	SCT	14571	0.0120
AMS3D	SCT	15473	0.0412
itcSegment	ECT	884	0.0066
itcSegment	ECT	1042	0.0803
itcSegment	ECT	876	0.0616
itcSegment	ECT	-	-
itcSegment	ECT	1522	0.1418
itcSegment	MCT	1202	0.0479
itcSegment	MCT	242	0.2056
itcSegment	MCT	548	0.0025
itcSegment	MCT	1512	0.3181
itcSegment	MCT	1062	0.3160
itcSegment	SCT	948	0.1494
itcSegment	SCT	-	-
itcSegment	SCT	-	-
itcSegment	SCT	-	-
itcSegment	SCT	-	-
Watershed	ECT	364	0.0098
Watershed	ECT	-	-
Watershed	ECT	354	0.3042
Watershed	ECT	673	0.0037
Watershed	ECT	76	0.3462
Watershed	MCT	226	0.7650
Watershed	MCT	651	0.0064
Watershed	MCT	529	0.0096
Watershed	MCT	-	-
Watershed	MCT	280	0.5980
Watershed	SCT	-	-
Watershed	SCT	-	-
Watershed	SCT	615	0.0316
Watershed	SCT	286	0.0017
Watershed	SCT	47	0.0131