Luís Filipe Rosa Simões

Master of Science

LiDAR based Biomass Estimation System for Forested Areas

Dissertation submitted in partial fulfillment
of the requirements for the degree of

Master of Science in
Electrotechnical and Computer Engineering

Adviser: José António Barata de Oliveira, Associate Professor, NOVA University of Lisbon

Co-adviser: Francisco Marques, Research Engineer, UNINOVA-CTS

Examination Committee

Chairperson:
Raporteurs:
Members:

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

November, 2020

LiDAR based Biomass Estimation System for Forested Areas

This document was created using the (pdf)LaTeX processor, based in the "novathesis" template[1], developed at the Dep. Informática of FCT-NOVA [2].
[1] https://github.com/joaomlourenco/novathesis [2] http://www.di.fct.unl.pt

Acknowledgements

Firstly, I would like to express my gratitude to my dissertation supervisor Prof. José de Oliveira Barata for allowing me to continue to develop my skills as a part of this amazing project. A big thanks to my co-supervisor Francisco Marques who pointed me in the right direction and made an effort to guide me trough the project during this difficult times.

To all the colleagues that helped throughout this year, by working together as a team, who were always ready to help, in particular those who accompanied me through all the robotics course.

Lastly, I would like to express my gratitude to my girlfriend and my family who always supported me, both financial and emotional, through this entire journey, and made this dream possible.

A warm thanks to my closest friends that FCT gave me: Filipe Antão, Francisco Cerveira, Flávio Silva, José Silva e Miguel Pato who always made me laugh and worked with me during these 5 amazing years.

Abstract

In continental Portugal, forest fires are considered the biggest and most serious cause of forest deterioration and therefore the introduction of forest management mechanisms and biomass monitoring are imperative for a better future. However, conducting field studies on a large scale is a very expensive and time-consuming task. Alternatively, through remote sensing via a LiDAR, it becomes possible to map, with high accuracy, forest parameters such as tree height, diameter at breast height or tree canopy length in order to carry out other relevant estimates such as above ground biomass.

In this sense, this dissertation aims to develop a system capable of, through algorithms and filters of point cloud processing, as statistical outlier removal, progressive morphological filters and region growing segmentation, extract in detail,a digital terrain model and correctly detect the number of trees in a given area, proceeding to the measurement of some interesting variables from the point of view of a forest inventory. Thus, testing data of different characteristics, our detection method obtained positive results, with all the average detection rates above 80 %.

Keywords: UAV, LiDAR, airborne remote sensing, wildfires, biomass, forest, mapping, monitoring, point cloud.

Resumo

Em Portugal continental, os incêndios florestais são considerados a maior e mais grave causa de deterioramento da floresta e por isso a introdução de mecanismos de gestão florestal e monitorização da biomassa são imperativos para um futuro melhor. No entanto, realizar estudos de campo em grande escala é uma tarefa muito dispendiosa e demorosa. Em alternativa, através da deteção remota por vias de um LiDAR torna-se possível mapear, com elevado rigor, parâmetros florestais como altura das arvores, diâmetro do tronco ou comprimento da copa da arvore de modo a proceder a outras relevantes estimações como a biomassa.

Neste sentido, esta dissertação teve como objetivo o desenvolvimento de um sistema capaz de, através de algoritmos e filtros de processamento de nuvens de pontos, como remoção de outliers estatístico, filtros morfologicos progressivos e segmentação por crescimento de regiões anexas , extrair com detalhe, um modelo digital do terreno e detetar corretamente o número de arvores numa determinada área, procedendo à medição de algumas variáveis interessantes do ponto de vista do inventário florestal. Assim, testando dados de diferentes características, o nosso método de deteção obteve resultados positivos, com todas as taxas deteção média superiores a 80 %.

Palavras-chave: UAV, LiDAR, sensorização remota, incêndios, biomassa, floresta, mapeamento, monitorização, nuvem de pontos.

Contents

List of Figures

List of Tables

Glossary

Basal Area	Cross-sectional area of a tree at breast height..
FARSITE	Fire growth simulation modelling system that uses spatial information on topography and fuels along with weather and wind data.
FlamMap	Fire analysis application that describes the potential fire behaviour..
Kd-tree	Space-partitioning data structure for organising points in a k-dimensional space..
Landsat	The Landsat program is the longest-running enterprise for acquisition of satellite imagery of Earth..
LiDAR	LiDAR is an active remote sensing technique that transmits lasers to an object and measures the distance between the sensor and the surface..
Nadir	Lowest point, that is vertically downward from the observer..
OpenCV	Open Computer Vision is an open source library for computer vision and machine learning..
PDAL	Point Data Abstraction is an open source library for translating and processing point cloud data..
Point curvature	A measure of how much the curve deviates from a straight line..
Point normal	The normal vector of a curve at a given point is perpendicular to the tangent vector at the same point..
Quickbird	High-resolution commercial Earth observation satellite..
RANSAC	Random Sample Consensus is an iterative method used to perform robust estimation of mathematical models..
ROS	Robot Operating System is a middle ware for development of robot software..

Voxel	Representation of a value in a three-dimensional space in a regular grid.
Watershed Algorithm	Watershed algorithms treats the image it operates upon like a topographic map, with the brightness of each point representing its height..

Acronyms

ABA	Area based approach.
AGB	Above Ground Biomass.
ALS	Airborne laser scanning (or scanner).
CHM	Canopy Height Model.
CMOS	Complementary Metal-Oxide-Semiconductor.
CV	Canopy volume.
DBH	Diameter at breast height.
DCM	Digital Canopy Model.
DEM	Digital elevation model.
DSM	Digital Surface Model.
DTM	Digital Terrain Model.
EMS	Electromagnetic Spectrum.
GIS	Geographic Information Systems.
GPS	Global Positioning System.
IMU	Inertial Measurement Unit.
ITD	Individual tree detection.
LAI	Leaf Area Index.
RMSE	Root mean square error.
RS	Remote Sensing.
TH	Tree height.

UAV Unmanned Aerial Vehicle.

VTOL Vertical-take-off-and-landing.

Chapter 1

Introduction

1.1 Rationale

Forest conservation requirements are changing rapidly within the context of a climate crisis. In recent years, climate change and the greenhouse effect has been increasingly discussed on the main stages of international politics with the aim of forcing a change of course. The scenario we live in is not sustainable and suggests an increase in the trend towards a greater occurrence of large forest fires. In general, wildfires have been causing a bigger impact in different regions of the globe, where, in the dry season, the vegetation is at an advanced desiccation process. In these conditions, forest fires deeply scar and reshape the revolving landscape. This uncontrolled combustion affect expressively, the different components of ecosystems, primarily the vegetation and the soil, causing immediate damage, such as the temporary absence of vegetation, degradation of the quality and the acceleration of erosion processes, putting the human lives and possessions in jeopardy.

Forest fires constitute an integral part of Mediterranean ecological communities, while also representing one of the main factors of ecosystem degradation. In continental Portugal, forest fires represent the most significant environmental issue and are often considered the main cause of deterioration of the superficial edaphic layer[21], alarming the authorities responsible for the need to implementation of new philosophies and methodologies for the management of forest spaces. However, the nature of the Portuguese climate and territory does not, in itself, justify neither the high number of ignitions, nor the vast burnt area over the past three decades. The media attention for this problem also brought more visibility to the sector, guaranteeing it a prominent place in the hierarchy of national priorities, since forest spaces represent more that 60% of the national territory.

Traditionally, the largest portion of biomass produced in the forest was collected by

the population and used as fuel or in animal activities. Over the years, the substitution of firewood for other sources of energy (gas, electricity, etc.), and the exodus of the rural population to urban areas led to a diminished maintenance of the national forest, breaking this natural balance generating large quantities of flammable fuel[28]. A comparative study of forest fire statistics conducted by ISA in regards to the Defense Plan Against Forest Fires between four Mediterranean countries allows two conclusions to be drawn: on one hand, the percentage of burnt area and the density of occurrences are significantly higher than the values recorded in any of the remaining four countries (Spain, France, Italy and Greece); on the other hand, while the other countries seem to have stabilised their values, Portugal presents a strong and worrying tendency to worsen the situation. In this way, forests represent a primary priority management target, regarding the preservation and conservation of these spaces, as well as the guarantee of their sustainability and long-term exploitation of the leisure, production and management functions of natural resources.

Accurate tree and forest biomass structure measurements are essential for a wide variety of operations, including climate change assessments, managements of natural resources, bio-energy production and finally biodiversity monitoring and conservation. Fires are known to be the main factor in forest degradation[8] due to its direct impact on its structure and dynamics, and has been shown that the consequences can be sustained for years after the event. Nonetheless, studying these longstanding repercussions of fires on forests in a efficient way remains a test because of the difficulty in collecting complex forest structure information over considerable areas.

In most countries, forest related activities are usually done by volunteers or by trained teams from various institutions, making it expensive and time consuming. Managing forest structure in wide-scale areas is a highly expensive and time consuming job which requires the collection of a large amount of data. One way to overcome this issue is with the use of remote sensing technologies in order to carry out such activities on a large scale and in record time.

The number of applications in which the use of drones has become useful is practically unlimited and is continuously growing. In 2007, the European Commission enumerated a set of development areas, including policing and security, control of the energy sector in the assessment of its infrastructure, climate monitoring and seismic events, communication and broadcasting and finally in fighting fires and forestry activities [40].The confidence shown by the governments to support this technology encourages researchers to work hard to develop algorithms and systems capable of demonstrating results.

Laser scanners and multi-spectral imagery has proven to be a revolutionary technologies offering forest management the needed spatial detail and accuracy across multiple applications and forest types. It has been proved in [30] that combining UAVs with both types of sensors provide an increase in performance of remote sensing platforms. The stability, security, autonomy and the number of sensors compatible with UAVs make this the ideal platform for wildfire prevention and post-fire monitoring programs.

1.2 Solution Prospect

This dissertation proposes a multi purpose biomass estimation system for forested areas by LiDAR sensors using small footprint UAVs in order to help understand the effects of wildfires and help the designated authorities by providing crucial information on the state and evolution of the forest before and after a fire. The system takes advantage of the resolution of the provided LiDAR data, as well as the type of forest and selects the right order in which the operations are performed in order to provide the user with a fast assessment of the state of the forest and the inherent biomass.

This system was designed to perform individual tree detection by applying a local-maxima seeded region growing in order to detect and segment possible tree clusters, and being able to extract measurements such as tree height and diameter at breast height in order to estimate biomass at plot levels. One of the main objectives of this dissertation was to try to quantify the damage that a fire does to the ecosystem and try to outline the burnt area in order to be able to apply economic models and estimate the economic impact, regarding the volume of burnt commercial wood. Despite being able to detect biomass changes, it was not possible to be tested due to the lack of resources.

Using a system of this nature and creating regular monitoring schedules, it is possible, in addition to studying and controlling the structure of the forest, to prevent and combat forest fires more quickly and efficiently.

The platform we intended to use for this study was composed of a DJI Matrice 210 rotary quad-copter fully equipped with the Velodyne's Puck LiDAR sensor (VLP-16) with an extended range of up to 100 metres, an on board computer (odroid) and access point system that allow us to communicate with the platform in real-time. From the UAV we are able to collect data from its position and orientation in the world, and collect detailed 3D point clouds with the LiDAR. A few trials were made in order to collect data, however, due to some system malfunctions and the on-going pandemic it was impossible for us to gather reliable in-house data.

1.3 Dissertation Outline

This dissertation is outlined as follows:

1. **Chapter 2** reviews the state of the art on remote sensing and LiDAR based systems as well as enlightening the reader with the possible applications and approaches on forestry activities.

2. **Chapter 3** gives an overview of the developed system and describes in depth the methods and strategies used during the investigation

3. **Chapter 4** goes over the experimental setup and explains the results achieved during the tests.

4. **Chapter 5** aggregates the conclusions extrapolated from this study and lists the future research and improvements on this topic.

Chapter 2

Background

This section surveys the state-of-the-art on remote sensing by providing insights on the technology and its working principle. Also demonstrates the two most common sensors used on forest related studies and highlights the two basic data processing approaches. A revision of the literature on LiDAR based forest remote sensing is made, as well as a brief observation on the economic impacts of wildfires.

2.1 Introductory Concepts on Remote Sensing

Remote Sensing (RS) can be defined as the science of acquiring and recording information about an object, area, or phenomenon, from a considerable distance (remotely) by specific instruments (sensors). Just like sensors, the human eye responds to the impulses of light reflected by the objects around. Those impulses, that vary from surface to surface, are the "data"which is then gathered and analysed by our mental computer allowing us to determine the type of object or environment around[11]. Transposing to science, these techniques allow us to take images of the earths surface in different wavelengths of the electromagnetic spectrum (EMS). These sensing devices record information about an object or surface by measuring the time it takes for the transmission of electromagnetic energy to reflect and radiate from the target surfaces back to the sensor, where the energy is transmitted from the RS platform, known as active RS. On the other hand, passive RS systems depend on external energy sources, such as the sun.

In short, a general RS operation workflow is as follows:

1. Emission of electromagnetic pulse
2. Transmission of energy from the source to the surface of the object
3. Interaction of the pulse with the surface in question

4. Re transmission of energy from the surface to the remote sensor

5. Sensor data output and processing.

6. Data interpretation and analysis.

Figure 2.1 gives an overview on electromagnetic remote sensing process divided into two main groups: data acquisition (steps 1 - 4) and data analysis (step 5 - 6).

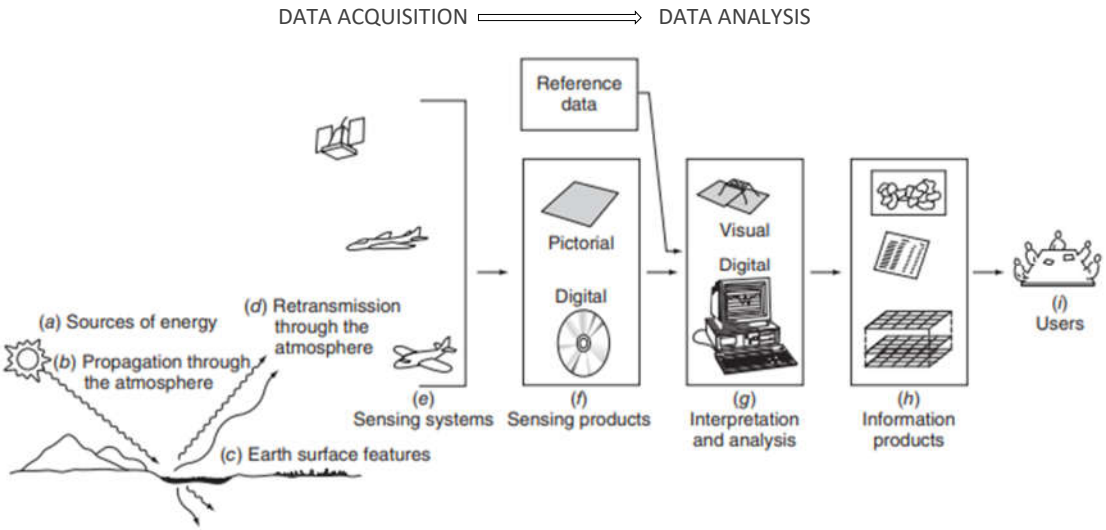


Figure 2.1: Electromagnetic remote sensing of earth resources. Adapted from [19].

When electromagnetic energy interacts with the Earth's surface, several portions of energy are reflected, absorbed, and/or transmitted across the surface. When applied the principle of conservation of energy with regards to the wavelength we can derive that this ratio of energy will vary depending on the surface material, shape and overall condition giving us the capability of differentiating features on the data retrieved. Also, different wavelengths will also make it possible to differentiate certain characteristics due to the variation of the transmitted energy depending on the wavelength. So features that seem indistinguishable in a spectral range can be quite different in another band.

Since most of remote sensing instruments work at a wavelength region where the reflected energy dominates, the reflectance properties of the target area are extremely important especially when concerning forestry activities. In figure 2.2 we see that the reflective properties of coniferous and deciduous trees in the visible range of the spectrum are practically the same, however when approaching the near-infrared (IR) range we can easily identify and separate both types of trees. The same rationale can be made for a number of other surfaces and materials. The spectral reflectante curves for common feature types are well known. The differences between soil, dry or green grass, concrete,

asphalt and sand, or even water, snow and clouds can be detected when varying the wavelength in most of the near-IR range, making it the go-to platform for a number of applications.

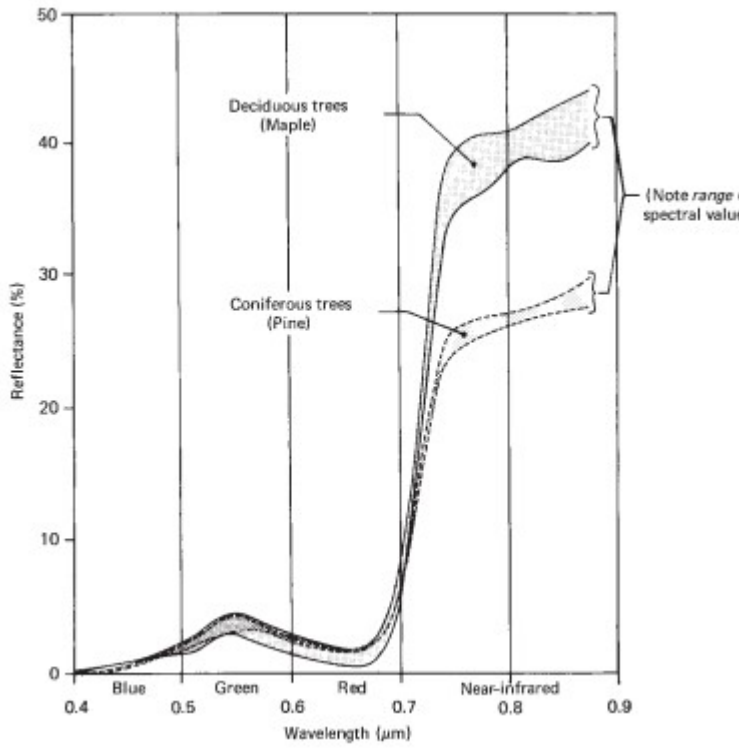


Figure 2.2: Reflectance properties of coniferous and deciduous trees. Adapted from [20]

To this day, more and more remote sensing systems are used to gather three-dimensional data of the earth surface, where Z represents the elevation data and X and Y represent the horizontal coordinates. This detailed description, when collected over a broader area is called *topography*. This means that it is also possible to retrieve the shape of objects or above ground infrastructures and biomass such as buildings in a city or tree tops in a forest. 3D data can be sourced from a variety of instruments, including photographic and multi spectral sensors, radar and LiDAR systems.

Like all branches of science it is imperative to have control data in order to serve as a reference when compared to the remotely sensed data. This reference data can take a number of different forms and can be derived from multiple sources. With regard to forestry applications, the most common form of reference data is the measurement of dendrometry parameters, or the investigation of species present in a given area of the forest. Reference data can serve as [20]:

1. Help in the analysis of the remotely sensed data.

2. For sensor calibration purposes.

3. To corroborate the information retrieved from the remote sensing platform.

However, this type of work can be very expensive and time consuming to collect, since certain field based measurements require specialized teams and on-foot procedures, or for example reflectance studies requiring intensive laboratory tests.

Despite the issues highlighted and with hyper/multi-spectral, and laser scanner sensors commercially available reaching a size and weight compatible with UAVs with small payload capability, makes remote sensing a viable and exciting field of study.

The next section will cover the capabilities of such platform, the two most used types of sensors as well as elaborate on the benefits of collaboration, coordination and cooperation of UAVs programmed to achieve one specific goal.

2.1.1 UAVs and Sensors: Capabilities and Technologies

As mentioned before, remote sensing is the technique of capturing information from a distance. In the past, RS was associated with satellite activities or manned air crafts with a set of sensors responsible for acquiring information. Nowadays, with the development of unmanned aerial vehicles and compatible sensors, higher risk operations, such as reconnaissance of hostile environments or places corrupted by natural disasters, have become safer and faster to deploy without the need of robust planning. Multi-rotor vehicles in particular offer higher versatility and flexibility with its Vertical-Take-Off-and-Landing (VTOL) capabilities, useful in confined spaces such as caves or ravines, both in rescue and mapping missions[31]. Additionally, these platforms are capable of making lower and slower flights in order to capture more detailed information with higher resolution.

To navigate the missions, UAVs must be equipped with different instruments that work together, such as Global Positioning Sensors (GPS), Inertial Measurement Unit (IMU), gyroscopes, accelerometers, cameras and laser scanners, to capture images and match each one with the pose of the platform and provide a detailed 3D map of the environment. Depending on the application different sensors can be used. Gas, smoke and ultraviolet flame detectors, infrared and thermal cameras, magnetic and radiation gauges, temperatures and humidity probes can all be adapted and useful in most applications, yet payload limits onboard small capacity UAVs represents a drawback [30]. So for the purpose of this work, only the required sensors for navigation were used, combined with a small footprint LiDAR and a multi-spectral camera.

2.1.1.1 LiDAR

LiDAR, or light detection and ranging is an active remote sensing system that uses lasers to measure, in high detail, elevation and characteristics of things like the ground, buildings or even entire forests. Just like a sonar uses sound waves to map the seabed, and radar uses radio waves to detect objects, LiDAR uses light pulses to gather information about the environment.

There are multiple ways to collect LiDAR data, either from the ground, air or even from space however airborne laser scanning (ALS) is the most commonly used and freely available[1]. In order to understand how a LiDAR system is used to calculate object heights in an ALS we need to understand the its four main components. First, is the aircraft which accommodates the LiDAR itself which uses a laser (either green or near infrared light) to scan the earth's surface as the platform flies. The next component is the GPS receiver that tracks the altitude and the coordinates of the aircraft allowing us to match each scan to a particular location on the ground. The third component of the LiDAR system is called an inertial measurement unit (IMU) that measures the platforms force, velocity and orientation, using a combination of accelerometers, gyroscopes, and magnetometers, in order to make the elevation calculation as accurate as possible. And finally an on-board computer that records all the useful information gathered by the measuring devices.

First we need to define to key terms in order to capture the essence of the system: a pulse and a return. A pulse is nothing more than a beam of energy emitted by the LiDAR laser. A return is the light that was reflected by the object and as been recorded by the LiDAR sensor. In short, bursts of light energy are transmitted through the atmosphere, reflected by the surface and return back to the LiDAR sensor. To get the height, the system records the time it takes for the pulse to travel to the surface and back, and then uses the known speed of light to calculate the distance between the top of the surface and the aircraft. Breaking down the calculation we have:

$$\text{Height} = \text{Travel time} \times \text{Speed of light} \times \frac{1}{2} \tag{2.1}$$

and this gives us the distance between the platform and the ground, but to calculate the actual ground elevation we need to take into account the aircraft's altitude, calculated using the data from the GPS receiver and subtract the height calculated earlier. But there are two more things to consider when calculating height, the first is the turbulence in the air that makes the aircraft rock a considerable amount. These movements are recorded by the IMU and should be considered for each LiDAR return; also some ALS systems have mobile scanners doing a sweeping movement, so while some light pulses travel vertically (at nadir level) most pulses leave the system at an off angle (off-nadir) and this discrepancy has to be considered when calculating elevation.

One key feature of LiDAR systems is the ability of a pulse to travel through some materials and deliver more information. The laser beams can travel between tree branches and leaves, all the way through the ground surface, producing multiple returns. These returns from within the forest canopy can tell us more about the forest structure, shape and density of trees, and even give important information on the lower vegetation. This makes it very useful in forestry, allowing the collection of more detailed data of the environment.

As mentioned, LiDAR data is recorded using a scanning sensor laser. The sweep width

and the overlap between parallel flight lines varies. Each laser beam registered generates a point with x, y and z coordinates, which can be converted in longitude, latitude and elevation. The set of LiDAR data, originated by recording information from different pulses, is called a point cloud. The density of this point cloud is specified a priori depending on the purpose of the flight. It is also possible to store another type information such as return information, intensity, time and altitude of the flight.

LiDAR systems can store data in two ways: (1) Discrete return and (2) full waveform. A discrete return LiDAR, when scanning through the forest for example, outputs the returns as individual hits for each branch the beam is reflected of, resulting in first, second, and N returns, finishing with a large and final pulse that represents the bare ground surface. On the other hand, full waveform LiDARs record the entire return as one uninterrupted wave, so, in order to classify each feature in the data you count the wave peaks, making it a discrete one. Despite the full waveform data being more complicated to analysed, the technology is moving towards this type of data acquisition.

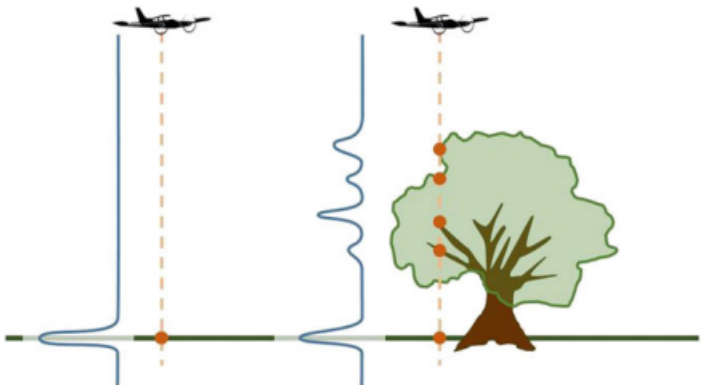


Figure 2.3: Illustration adapted from [47] showing the differences between full waveform LiDAR (full line - blue) against discrete waveform LiDAR (discrete points - orange).

The advantages of using LiDAR systems to complement or renew conventional photogrammetric procedures for terrain and surface mapping accelerated the development of high-performance scanning systems. Among their advantages, these systems allow researchers the freedom to collect surface data about arduous slopes and shadowed or inaccessible areas like caves or wells. Depending on the complexity of the surface and LiDAR resolution, the data sets can be extremely large and difficult to process, producing point clouds with a point density of 0.5 to 2 points per square meter, while detailed mapping of small areas where dense vegetation is present may require 10 to 50, or even more points per square meter.

The practical applications that derive from the use of the LiDAR systems are listed below:

- **Land Mapping** - Important in many planning and management activities, land mapping is one of the most reviewed applications of LiDAR remote sensing operations.

- **Atmospheric Applications** - As its able to detect particles in both water and air, LiDAR can be used to identify pollutants, as well as cloud profiling which is another filed of interest by researchers.
- **Biology and Conservation Applications** - Widely used for monitoring seasonal snow cover at mid to high altitudes and measure long-term changes to glaciers, or even collection details on ocean depth, composition and existing species.
- **Wildlife Ecology Applications** - Determination of the movement of individuals or groups of animals over time is often important in wildlife ecology and LiDAR can be used to directly contribute to this process.
- **Environmental and Natural Disaster Assessment** - Wildfires, storms, earthquakes, volcanic eruptions, all result in caos and LiDAR can be useful to survey and study those places where help is extremely needed.

These are just some examples of the most common applications of LiDAR systems which shows the role this system can have in the future of robotics, computer vision and by helping human beings in the most efficient way possible.

2.1.1.2 Multi and Hyper Spectral Devices

Multi spectral cameras, as the name implies, capture multiple images within the specific wavelength ranges across the electromagnetic spectrum from the ultraviolet, to the visible, to infra red bands, allowing us to obtain additional information that the human eye cannot see.

Before understanding how multi-spectral technology works, it is necessary to understand the operation of monochrome and colour cameras. A mono-chromatic camera is equipped with an image sensor that contains a 2D-array with light-sensitive pixels. These pixels are sensitive to most of the electromagnetic spectrum. In a monochromatic CMOS sensor, for example, each pixel is sensitive to light ranging from 400 nanometers to 1000 nanometers, covering all of the visible spectrum and near infra-red ranges. Thus, a monochrome image sensor captures light at all wavelengths, so it cannot discriminate between different colours or wavelengths, resulting in a black and white image.

Like monochrome cameras, a colour camera contains an image sensor with a two-dimensional array, however, in this type of cameras, this sensor is covered with a mosaic composed of different pigments that transmit the colours red, green and blue. These pigments together consist of a colour array mosaic or CFM. This mosaic is manufactured in a way that a quarter of the pixels see red, a quarter of the pixels see blue and the rest see green. Thus, a colour image can be thought of in three separate images, which when superimposed generate the colour image.

Spectral technology employs similar principles. To generate multi spectral images, instead of shaping pigments representing red, green or blue colours on the image sensor this

devices patterns the sensor with micro sized optical filters that have changeable colour and different transmission characteristics between them. The same way as a coloured image, a raw image captured with a multi spectral camera can be thought of as separate images each taken at a specific wavelengths with a well-defined bandwidth barrier such as those presented in the Table below.

Table 2.1: Multi and hyper spectral cameras - list of the different spectrums

Acronym	Description	Wave length
VIS	Visible portion of the spectrum	(380 - 800 nm)
VNIR	Visible and near-infrared spectrum	(400 - 1000 nm)
NIR	Near infra-red	(900 - 1700 nm)
SWIR	Short wavelength infra-red	(1000 - 2500 nm)
MWIR	Middle wavelength infra-red	(3 -8 μm)
LWIR	Long wavelength infra-red	(8 - 12.4 μm)

Up until a few years ago, you would find multi spectral cameras mainly in aerospace. The equipment was very large and very expensive. Today, the cameras have become significantly smaller, but they are still expensive, however its continuous development is important as the list of possible applications is vast:

- **Medicine** - Allow less invasive tests and diagnosis to be carried out, without the need for surgical interventions.
- **Environment** - Makes it possible to, more efficiently, locate the presence of contaminating elements in the different ecosystems, as in the case of plastics or oil spills in the oceans.
- **Agriculture** - Possibility to detect the condition of certain crops, as well as the presence of pests.
- **Forestry** - Allows mapping wooded land with greater efficiency and better results as well as carrying out different studies related to fire risk and its consequences.
- **Hydric sector** - Fundamental when analyzing water quality in both natural and urban environments.
- **Defense** - It has shown potencial in the detection of antipersonnel mines or the detection and fight against drug trafficking.

By combining both LiDAR systems and multi spectral cameras it is possible to study in a never before seen detail, as it offers a great variety of data which benefits all fields of science.

2.2 Data Processing Approaches

This topic will cover the two main techniques used in forestry activities, area-based approaches (ABA) and individual tree detection (ITD).

2.2.1 Area Based Approach

In general, area-based approaches have been targeted as the preferable method for LiDAR-based forestry activities because of its relatively high accuracy and lower point density, meaning more efficient data treatment when compared to ITD. Another important characteristic is that the resulting 3D point cloud of a sample plot consists of the heights for each laser beam reflected on a given surface. This contains information about the ground surface as well as the vertical distribution of vegetation. These parameters can be used to estimate inventory attributes such as diameter at breast height, or DBH, and canopy volume,or CV.

As pointed out by [13], a general ABA workflow is as follows:

- Collection of model calibration data through traditional field measurements;
- Establishment of empirical connections between the field-observed metrics and the LiDAR metrics, through means of a parametric or non parametric statistical methods;
- Using the previously created model, compute the forest plot against the LiDAR metrics on a variable grid cell size, and derive forest-level statistics.

In order to develop sturdiest prediction models, a number of studies have investigated different strategies that will be described next.

2.2.2 Individual Tree Detection

Individual tree detection methods are frequently used to detect individual tree crowns and proceed to its segmentation in order to predict attributes of interest using allometric models. As stated before, there are numerous methods proposed to delineate individual trees from ALS data, nonetheless, a general ITD workflow consists of tree detection, feature extraction, and estimation of tree attributes. In computer vision this translates to the flowchart presented in Figure 2.4. Detection of individual trees relies heavily on canopy height models, or CHM, interpolated from the ALS height data [38], although, point-based techniques could be used, both for tree detection and tree segmentation.

Literature has showed that the correct delineation and detection of individual of groups of trees is the key to success when it comes to ITD methods. Problems arise when there are groups of trees close to each other becoming challenging to discriminate individual ones, resulting in undetected trees and making the algorithm biased towards larger ones [33].

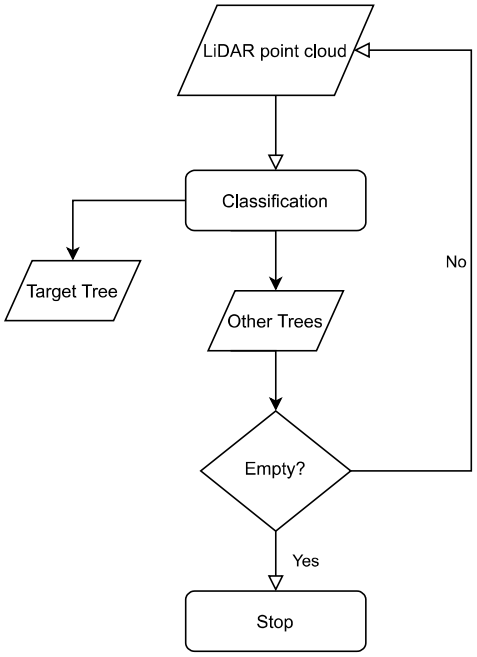


Figure 2.4: Individual Tree Detection flowchart.

A way to overcome such problems is to fuse both ABA and ITD techniques in order to achieve better results in the long run. The next section will give a brief overview of the literature regarding both methods for forestry applications, exploring the procedures researches have been using when performing all sorts of forest studies.

2.3 Previous Related Remote Sensing Efforts

On a global scale, forests are vulnerable to population growth and human activities that can cause deforestation, air pollution and consequently, climate change[39], the latter is expected to cause larger forest damage and tree mortality from direct and indirect causes, and in Portugal, in recent years more than a third of the territory was affected by damaging events [21]. There is an urgent need for quantitative data on forest health and monitoring systems that allow us to detect and locate harmful events. The next paragraphs will explain to the reader the studies carried out in the different forestry areas using the LiDAR system.

2.3.1 Estimation of Dendrometry Parameters and Tree Species Classification

Dendrometry [7] is a branch of botany that is concerned with the measurement of the different tree dimensions. The most important ones, when it comes to remote sensing and presented in Table 2.2.

Table 2.2: Important Tree Measurements

Attribute	Unit	Expected Measurement Accuracy
Tree Height (TH)	m	0.5 - 2 m
Diameter at Breast Height (DBH)	mm	5 - 10 mm
Upper Diameter	mm	5 - 10 mm
Height of crown base	m	0.2 - 0.4 m
Location	m	0.5 - 2 m
Basal Area (BA)	m^2	based on diameter accuracy
Leaf Area Index (LAI)		
Canopy Volume (CV)	m^3	10% - 20%
Biomass	$\frac{kg}{m^3}$	10% - 20%

Adapted from: Hyyppä, J. et al.,Remote Sensing of Forests from LiDAR and Radar,in Remote Sensing Handbook, Boca Raton, FL, CRC Press, pp. 397–427, 2015. as cited in [38]

These measurements, or variations of such, are of extremely importance to infer properties of greater interest, for instance, classification of species and overall health or quantity of commercial wood retrievable. However, the collection of this type of information by means of pure ground-based field techniques is remarkably time consuming and expensive [14]. Using LiDAR remote sensing technology, these activities can be performed rapidly and efficiently.

Early studies were mainly focused on retrieving tree height on a particular area for inventory purposes. With ALS systems, laser pulses hit tree tops, bushes and the ground below[3], after that, filtering techniques can be applied to separate the ground returns and vegetation returns (inliers) from the back scattered signals (outliers). If only the first pulses are recorded the height of the tree can be calculated by substracting the distance between the top of the canopy and the returns from the forest floor. Another way to calculate height is if both first and last returns are available. The difference between these two measurements is used to infer the vegetation distribution across the vertical plane [2].

There are numerous ALS parameters that can be optimised in order to increase the success of a given application.Today, ALS is becoming a standard technique in the mapping and monitoring of forest resources and so more researchers have devoted time to analyse

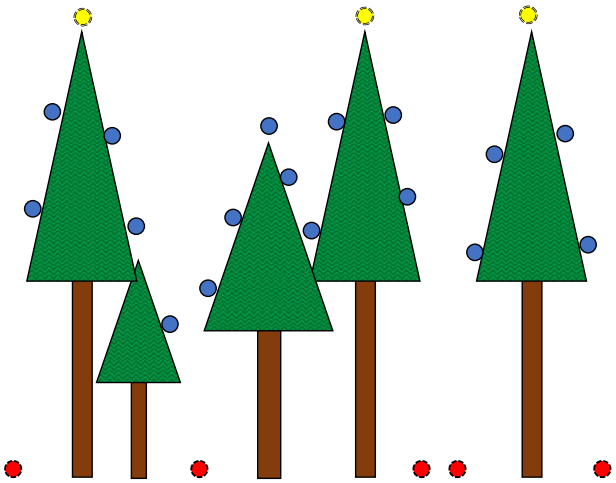


Figure 2.5: LiDAR's first and last returns example.

laser settings for such purposes, the most common being point and laser density, elevation and planimetric accuracy, operating range and plot area. Early studies, as in Magnussen et al., [23], for example, showed that the number of laser pulses per spatial unit is directly correlated to reliable LiDAR metrics, and since then, the minimum pulse density and plot areas have rarely been less than 0.1 $pulses/m^2$, and 200 m^2 respectively[25, 29].

Studies cited by [6] show that a combination of a TH metric and a canopy cover metric provides enough information for decisive prediction of forest standing volume. However, the study revealed that using solely TH metric from ALS as auxiliary information, can be of use when estimating the standing volume in a more practical way. Barbati et al. [2] also combined field measurements with ALS data for application with an ABA metrics. One of the objectives of this study was to estimate the standing volume of a *Pinus pinea L.* forest based on the fact that the volume of one tree is directly proportional to its height as demonstrated in [6]. The second part of the study was aimed at detecting gaps in order to improve the performance of the models for volume estimation. However, the possibilities go beyond that. If we look from a planning perspective, the detection of openings facilitates navigation through what was thought to be rougher terrain, and plan attack or escape routes in case of a wildfire.

Vastarant et al.[45] made use of ABA and ITD in conjuction with Random Forest (RF) algorithm to predict the tree state and create a map of defoliation. The results were significant as the plots were classified with an accuracy of 84.3percent. The effect of pulse density on area-based mapping was also taken into account, as seen in [23] and with variations between 0.5 and 20 $pulses/m^2$ the mapping accuracy was not extremely sensitive, varying up to a 12 percentage points, from 77 to 89 percent. Saarinen et al.[36] also used an ABA in conjunction with RF and nearest neighbours (NN) algorithms in order to classify, monitor and map riverine vegetation through ALS. The study achieved

an overall accuracy of 72.6 percent for vegetation cover classification which goes to show that ALS systems are one of the state-of-the-art technologies when it comes to mapping the environment.

Gatziolis et al. [10] used a voxel based approach in order to assess stem diameter and volume of individual trees in a point cloud. The method used in this study went from identifying which points belong to the targeted tree, to generating a tree representation using voxels and finally process them in order to extract stem and branch architecture. The researchers also managed to estimate the volume of commercial wood available in the error with a error of less than 2 and concluded that the precision obtained by using high-density scans is much higher when compared to those obtained using a standard density, as shown before.

Li et al. [18] took an interesting but effective approach on ITD by taking advantage of the relative horizontal spacing in between trees, which in general, is greater at the top than at the bottom. In short, starting at the top of a tree its possible to identify and grow a target tree by including points within a certain threshold and excluding the rest. The results obtained were better than expected for a mixed conifer forest on rugged terrain, with detection accuracy of 86 percent and 94 percent of the segmented trees were correct, when similar studies, cited by [18] reported accuracies between 60 and 80 percent.

More recently Yao et al. [49] in order to properly segment individual trees, used a normalized cut segmentation technique to a voxel representation of the forest area versus a watershed technique to consistently detect smaller trees which are not visible by local maxima in the CHM. Regarding species classification they considered different types of salient features, calculated with the help of LiDAR metrics, and a maximum-likelihood estimation method. Stem volume and DBH estimation were accomplished by deriving predictor parameters for regression analysis. The accuracy results of the classification between deciduous and coniferous trees were approximately 95 percent, when stem volume and DBH estimation provided a RMSE of 16 and 9 percent respectively, much better than previous studies at the time.

Overall both ABA and ITD, as reported by Peuhkurinen et al. [33] and Vastaranta et al. [44], achieved similar accuracy results when estimating the mean stand characteristics when compared to traditional field-work. The major differences among these methods are the bias of the estimates and the amount of field work needed. Low density data had a bigger impact on ITD results than in ABA. The quality of ABA is dependent on the quality and amount of field-work, since ITD only requires field-work for calibration purposes. Finally, the complementary properties of both approaches should be noted. In addiction to actual tree detection, data on individual trees can be generated by predicting diameter distributions with ABA, while combining both can be beneficial for reducing tree detection errors.

2.3.2 Wildfire and Biomass Fuel Assessment

In recent years, wildfires have been the most important natural source of disruption in Mediterranean ecosystems, infrastructures and human lives. There was an abrupt increase in the number of ignition sources, causing an atypical increase in exposure and in the recurrence of uncontrolled fires. To avoid large scale fire and smoke damage, prompt response and accurate fire detection is critical in order to minimise the destruction that fires may cause due to their rapid propagation and combustion cycle.

In general, the remote assessment of active fire characteristics can be grouped into two main application branches:

- The detection of actively burning areas using by optical and thermal imagery combined;
- Estimation of the energy radiated from fire as it burns with thermal imagery.

However, detecting fire from ALS data is a relatively new research subject and instead of observing the flames, LiDAR systems are mainly used to detect the resulting smoke plume, which is much larger and higher, making it easier to evaluate.

Utkin et al. [42] developed a LiDAR based system capable of tracing smoke-plume evolution and detection of the origin location under unfavourable conditions. The technique used was based on a previous study [41] published by the group in which the signal-to-noise ratio (SNR) was related to the LiDAR metrics, smoke plume and atmospheric conditions. They also concluded that the distance to the fire plays a significant role in the correct detection of the smoke plume as a reliable system must achieve a SNR equal or greater than 5. The experiment also showed the possibility of detecting smoke plumes as early as 40 seconds after the fire started.

In 2014 [43], the same author provided experimental proof of excellent scalability of LiDAR fire detection techniques that can be implemented at different price points in various segments of the fire surveillance market. Despite being terrestrial focused, the same idea is applied to airborne LiDAR systems.

A recent study by Price et al. [35] explored the potential of LiDAR technology to map fire fuel hazard throughout large forest areas prone to fires. According to the authors this study was focused on a type of vegetation and litter important for crown fire propagation. The researchers were able to clearly discriminate the percentage of fuel present in three classes of vegetation height, near-ground (0.5 - 4 m), lower (4 - 15 m) and upper (15 - 45 m) canopy fuels, however, when estimating the fuel hazard present in the area they concluded that time-since-fire is a poor predictor of fuel accumulation since the resurgence of vegetation post-fire e heavily dependent on fire severity, atmospheric conditions and location.

Mutlu et al. [26] took a different approach and managed to accomplish accurate estimates of surface fuel parameters and created a fuel map by processing LiDAR data

using the height bin approach [34] and multispectral imagery as inputs for fire simulation software (FARSITE). The results derived from this study showed that LiDAR-derived models were able to assess fuel models with high accuracy and provide fire perimeters and fire growth area. The results were yet compared to Quickbird-derived model showing a big discrepancy which can cause problems when applied to real scenarios where accurate information is needed.

2.3.3 Post-Fire Recovery and Forest Health Monitoring

Wildfires play a major role in driving vegetation changes and can cause important environmental losses, where the dominant species lacks efficient regeneration mechanisms. The speed and extent of recovery depends on fire severity, timing and ecosystem. Nature has equipped many plants capable of quickly recovering from fire, many shrubs and grass readily sprout from underground root structures after a fire, however post disturbance vegetation management strategies of burned areas are in need.

The remote assessment of post-fire effects can be broadly divided into:

- Burned area and perimeter methods;
- Methods that assess a surface change caused by fire, such as cover or fuel.

This class of study is the most subject to data-fusion. Most of the research reviewed in the matter reported cases of LiDAR derived metrics coupled together with multi spectral, thermal and satellite imagery. An example of that is [24], who combined an aerial LiDAR and multi spectral imagery with the objective of detecting and mapping regeneration types in a Mediterranean forest, based on a classification model using remote sensing variables. Both sets of data were manipulated using the same grid cell size in order to facilitate matching and metric extraction from the two sources. The classification algorithm used was, again, RF, since the literature showed it has good performance when dealing with this sort of data [36].The study achieved good results, with classification accuracy of up to 79 percent between the 5 types of regeneration in the study area.

Kane, V.R., et al. [16] studied the fire effects on forest spatial gaps and structure using data from airborne LiDAR together with Landsat fire severity measurements. The author analysed different burn ratios, over two bands of the spectrum (near and mid infrared) in order to estimate fire severity across the area. LiDAR data collection and processing was common to many other studies. Researches concluded that as fire severity increased, the total canopy area decreased while the number of agglomerates increased indicating progressive fragmentation of remaining canopy into smaller clumps, and bigger gaps in between vegetation.

Sato et al. [37] managed to assess post-fire biomass changes using only a LiDAR in Amazonian forests. The goal of the study was to quantify the impact of wildfires on forest height and biomass, 10 years after the event. First a forest inventory was made to estimate the current levels of vegetation and biomass. Drifting away from typical

methods to predict above ground biomass (AGB) that are based on regression models as a way to establish a correlation between LiDAR metrics and field measures, in this study the FUSION software was used to extract information such as number of returns, height distribution and intensity. Similarly to other studies, the data was then processed in a grid layout and the results clearly demonstrate the persistent loss of biomass up to 10 years after the fire.

Regarding forest health, Solberg et al. [39] combined LiDAR and hyper-spectral data sources to integrate a variety of important measures for forecasting the health status of vegetation. Like most of the studies here reviewed the first part consisted of gathering information on tree detection and positioning by searching for the local maxima in the data set. LiDAR data helped in calculating the foliar mass by estimating the projections of the tree crowns onto the horizontal plane. Only then, and with the help of a hyper-spectral imagery techniques were able to model chlorophyll concentrations in the forest canopy layer, which when coupled with the canopy volume and mass it is possible to derive any type of forest damage. The results were not clear, however the author remains a strong advocate of LiDAR based remote sensing and considers it a suitable tool for future forest health monitoring services.

In 2012, White et al. [48] evaluated the use of metrics derived from pre and post fire LiDAR and muti spectral data in order to detect fire effects and measure the extent of the damage impacted by the Lockheed fire. While preliminary, this study revealed that combining field-measured data with remotely derived metrics can extend our understanding of how forest react to disturbance and suggests many way of data analysis.

2.4 Economic impact of wildfire events

In recent years, the phenomenon of forest fires in the Mediterranean regions has been studied from multiple perspectives, yet the primary focus is on prevention and combat studies, tasks that are the responsibility of the designated authorities. However, studies focused on the social and economic side have had very few contributions, which is strange given that the most immediate effects of forest fires are usually of an economic nature.

It is easy to determine with some rigour the value of the areas affected by the fire, however it is difficult to accumulate this value as a fundamental support for the protection and maintenance of animal species and the soil itself. In the same way, it is difficult to account for many of the forms of traditional economy associated with the forest, which are destroyed by fires, such as subsistence agriculture, beekeeping and the collection of wild fruits and aromatic plants. On the other hand, these events favour the development of pests that often not only destroy burnt trees that have not yet been removed but also attack those that did not burn but are on the edge of the fire and end up being affected, resulting in their death.

In Portugal, every year a high number of forest fires reduces to ash several thousand hectares of forest and unfortunately it is unknown the actual values of the volumes of

wood burned, however, estimates made by the Forestry District of Coimbra reveal that in recent years more than 60 million cubic metres of wood have been reduced to dust, which is equivalent to more than 1500 million euros. This numbers are astronomical and only a small part of this amount has been returned to forest owners. This problem leads to the lack of raw materials causing closure of sawmills and wood processing plants, leading workers to unemployment [28]

Knowing the importance of its effects, so vast and causing so much damage, should be a strong reason that contributes to the regular maintenance and preservation of our forests. It is urgent that all countries, collectively, demonstrate efforts to prevent, to the best of their ability, the occurrence of fires by developing regular prevention and combat procedures in order to stop the resurgence and evolution of small-scale fires in unstoppable fires. This way, the problem would be greatly reduced.

2.5 Discussion

This chapter provided a brief explanation of the concept of remote sensing and the theory inherent in that technology. The two methodologies, ABA and ITD were explained in detail and accompanied by studies related to the theme. It was concluded that ABA is considered to be the most cost-efficient, due to the use of lower point densities, instead of the ITD that needs higher resolutions in order to be able to identify all parts that compose of a tree. However, for calibration, ABA needs a lot of real data, while ITD only needs a few field measurements.

Studies demonstrate that ITD provides a means to measure the true diameter of a tree, as well as the height distribution along the crown, which is important for forest planning, simulations and model optimisation. With ABA, the characteristics of the forest are estimated based on the population and other measures would have to be predicted. However, it is always important to test both methods as the combination of both proves beneficial in reducing detection and estimation errors.

Continuous monitoring of forests, before and after a fire is critical in order to quantify lifelong impacts, not only on above ground biomass and forest structure but also to isolate changes in ecosystem biodiversity. To this extent, LiDAR technology has shown promising results when collecting surface materials to be processed in a more faster way than the traditional data acquisition techniques. In this research the effectiveness of using LiDAR point cloud data at a forestry level were debated by reviewing previously conducted studies. These forest applications were mainly focused on measuring forest features, performing inventory and tree species classification as well as detecting smoke and managing forest fires, and finally planning forest operations which is useful in areas with high tree density and with difficult access.

Chapter 3

Methodology

This chapter gives a few pointers on key concepts that help the reader understand the decisions made and also a more in depth description of the methods and algorithms developed throughout the realisation of this study. Section 3.1 paints a brief picture on the current systems available for forest inventory and its disadvantages and provides a general overview of the software architecture implemented. Section 3.2 goes through the different types of point cloud data and explains the extensive pre-processing required, starting with data conversion and going through the ground extraction and an explanation on the outlier removal filter. In Sections 3.3 and 3.4 we go over the tree top and tree trunk detection algorithms and list the possible problems that can hinder our approach. To conclude, Section 3.5 reminds the reader of the importance of forest inventories, lists the attributes that are perceivable from the data acquisition and reviews some of the literature regarding key tree measurements and the procedure used for above ground biomass estimation, so it can correctly estimate and map an extensive area.

3.1 General Overview

In this Section we present a brief overview of the software architecture developed, based on open-source libraries for use with a LiDAR system and the ROS framework with the intent of performing a fast and efficient forest inventory and biomass mapping.

To this day, most forest software analysis tools require powerful systems with high processing capabilities and most of them are either paid or require lots of training. We are beginning to see a lot of concern from the governments in regards with the forest structure and maintenance in order to protect the environment and the habitats from wildfire events. However, most forestry activities still rely heavily on large human fire patrols and eye measurements to collect data. These patrols require specific training and

tools to take most measurements and can only cover a small percentage of ground when compared with a UAV. Another example would be the police patrols after a wildfire, taking days trying to delineate and estimate the burned area. In Portugal most forest structure data is sealed and can't be accessed, which hinders even more all the forest related studies.

Forest inventory and mapping studies are commonly inserted on a two-stage procedure using on-site field plots and remotely sensed data. The LiDAR system emits pulses of light energy from the platform to the ground using a laser and the on-board computer records the time it took for the pulse to travel to the ground, reflect and return back to the sensor. These returns, bounce of the tree stems and provide more information from within the canopy and tell us more about what's happening inside the forest, and the ground below.

Figure 3.1 illustrates the step by step approach we implemented in order to define a biomass map.

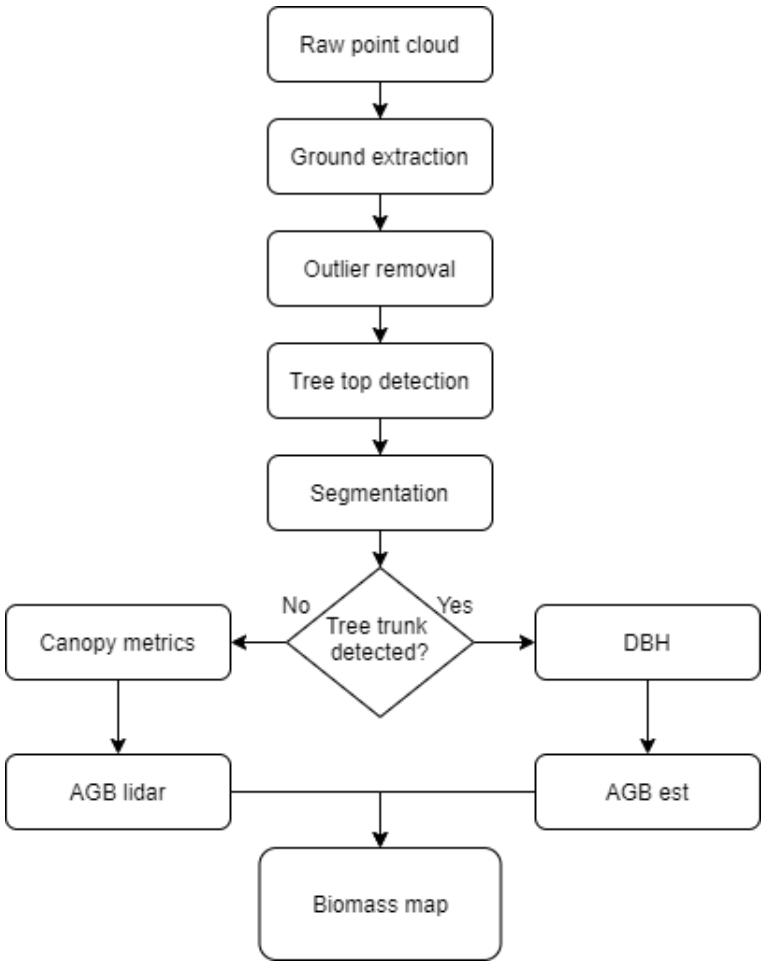


Figure 3.1: Flowchart of the general system workflow for tree segmentation and biomass mapping of the forest point cloud.

This study is meant to accurately predict the above-ground biomass in a certain area, and for that a few hindering factors have to be taken into account, such as overlapping trees and under story vegetation. The designed system was developed with that in mind and meant to overcome such problems with high degree of accuracy.

3.2 Data Pre-processing

A rigorous pre-processing step for the LiDAR point cloud is required in order to efficiently derive quantitative information from the data set. The amount of data that a UAV - LiDAR system combo can gather in a short flight is extraordinary and so a few steps have to be taken in order to lessen the time it takes to process high amounts of information.

As mentioned earlier, the main reason that surveyors are choosing LiDAR mapping system instead of photogrametry is the hability to obtain palpable data under high complexity environments, such as over populated forests. This brings up the question of how can the point density of the point clouds produced by a LiDAR system affect the estimates of biophysical parameters such as those presented in Table 2.2.

There are multiple ways to increase point density. The first option is by increasing the number of laser beams generated by the sensor, thus, scanners such as the Velodyne VLP-32C with 32 laser-beams will generate point clouds with double the point density as the 16 beam counter part VLP-16, under the same conditions, however the more laser beams the more expensive the system is. Three additional variables at our disposal that directly affect the output's point density are *(1)* height of flight, where the lower the platform's height is, the greater is the point density, *(2)* UAV speed, where the lower the platform's speed is, the greater is the point density and *(3)* LiDAR frequency setting, where the higher the laser frequency is the greater the point density is. This three degrees of freedom are critical in planning flight missions and can gives some margin of manoeuvre when collecting data in different conditions, which can be beneficial when high detailed data sets are needed. Studies [22] show however that in some cases high point density is not detrimental for certain studies, and their results show no significant decrease in measurement accuracy when using a lower point density. In this case, time of flight (TOF) was not a priority, yet in some high risk situation a trade off has to be made.

The system developed in this work takes into account the nature of the given data set in order to choose the best combination of algorithms. If the data set has lower resolution it is unnecessary to try and retrieve specific information on the leafs or tree trunk due to the incapability of correctly detecting and delineating such features, and so processing time can be saved. In order to extract certain parameters, such as height and biomass, [22] showed that reduced point density could deliver reasonable estimation results. If the data set has higher resolution, the amount of data possible to extract is bigger and the algorithms require more time to process. As shown in [32], Figure 3.2 demonstrates the obvious differences in detail when using high density data sets. Despite this differences, a pre-processing routine is always needed.

Table 3.1: Common point densities and applications.

Point Density	**Application**
0.5 - 1 pts/m^2	Basic surface model and forest inventory
1 - 2 pts/m^2	Flood modelling
2 - 5 pts/m^2	Multi-purpose data sets
5 - 10 pts/m^2	Basic 3D models
10 + pts/m^2	Detailed 3D city models

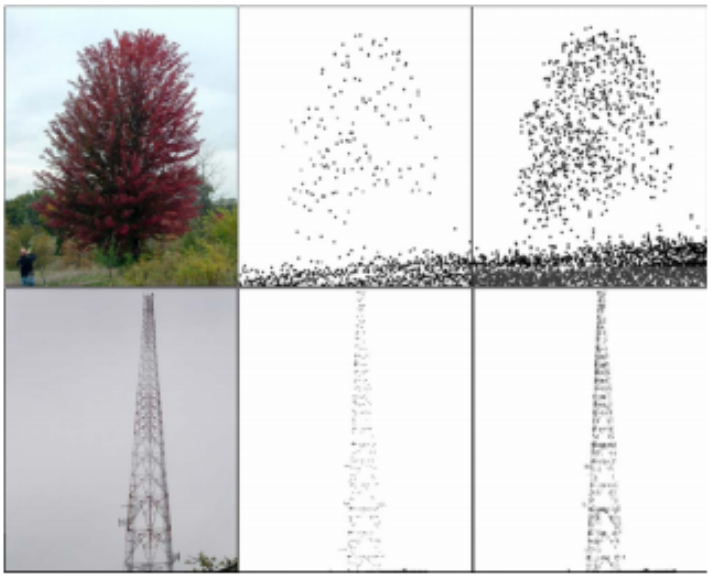


Figure 3.2: Comparison of point density for a tree and an electrical tower in point clouds generated using a discrete-return LiDAR data (middle) and full-waveform LiDAR data(right). Adapted from [32].

Starting with format conversion, we begin by using the PDAL library for translating and processing our point cloud data. The most common format used in the LiDAR industry for point cloud data is the LAS format. Being a binary file, it becomes easier to be archived or imported. Each LAS data file contains a header block with important information about the LiDAR survey, followed by the individual records for each laser pulse collected by the sensor. For our processing needs, a more user friendly file format was required in order to work directly with the PCL library. The ability of PCD files to store organised point cloud data is of extreme importance in areas such as robotics and computer vision. That coupled with the support off all primitives data types allows the point cloud data to be very flexible. With a single command, the PDAL tool is capable of converting all of our LAS files into PCD files for later use.

After the conversion, the raw point cloud is loaded onto the system and a spatial grid

```
# .PCD v0.7 - Point Cloud Data file format
VERSION 0.7
FIELDS x y z normal_x normal_y normal_z curvature
SIZE 4 4 4 4 4 4 4
TYPE F F F F F F F
COUNT 1 1 1 1 1 1 1
WIDTH 3283
HEIGHT 1
VIEWPOINT 0 0 0 1 0 0 0
POINTS 3283
DATA ascii
1.068 0.094720997 -0.016231 -0.33680952 -0.5904243 -0.73345655 0.021525998
1.061 -0.028772 0.060295001 -0.56758565 -0.38228491 -0.72918081 0.0050898455
1.0623 -0.040527001 0.061046999 -0.55886567 -0.084654853 -0.8249259 0.000421467
1.0662 -0.062633999 0.068122998 -0.64745134 -0.36214662 -0.67056441 0.0028368302
```

Figure 3.3: Snippet of a PCD file format with forest data.

is applied in order to divide the point cloud into smaller plots with a specific pre-defined rectangular grid size to lessen the computation time it takes to process.

3.2.1 Ground identification and extraction

Almost all geographic analysis tools have the ability to model, visualise and extract some sort of representation of the elevation representing the surface of the earth for multiple applications. In scientific literature there are three common terms that need to be clarified: digital elevation model (DEM), digital terrain model (DTM) and digital surface model (DSM). A digital surface model represents the earth's surface with all surrounding objects on it, whereas the digital terrain model represents a topographic model of the bare ground, without any objects nor buildings; whilst digital elevation model is generally used as a collective term for both DSMs and DTMs. A clear difference between a DSM and a DTM is highlighted in Figure 3.4, where the presence of objects in the data set is easily visible in the DSM whereas in the DTM only the ground is present.

The extraction of digital terrain models has proven to be an extremely challenging and time consuming task with photogrammetry technologies. Since laser-scanners can gather data from objects, buildings, vehicles and most importantly penetrate the vegetation, we can now acquire high detailed DTMs by differentiating and classifying the LiDAR data sets as ground or non ground features.

In forestry, in order to detect and individualise trees, some sort of ground extraction from the point clouds is needed, otherwise it would be difficult to visualise and estimate key parameters, such as height.

It has been shown in the literature that the combination of the two most basic morphological operations such as dilation and erosion generates opening and closing operations that can be used to filter and extract measurements for buildings and trees from LiDAR point cloud data. In image processing, the dilation algorithm adds pixels to the perimeter

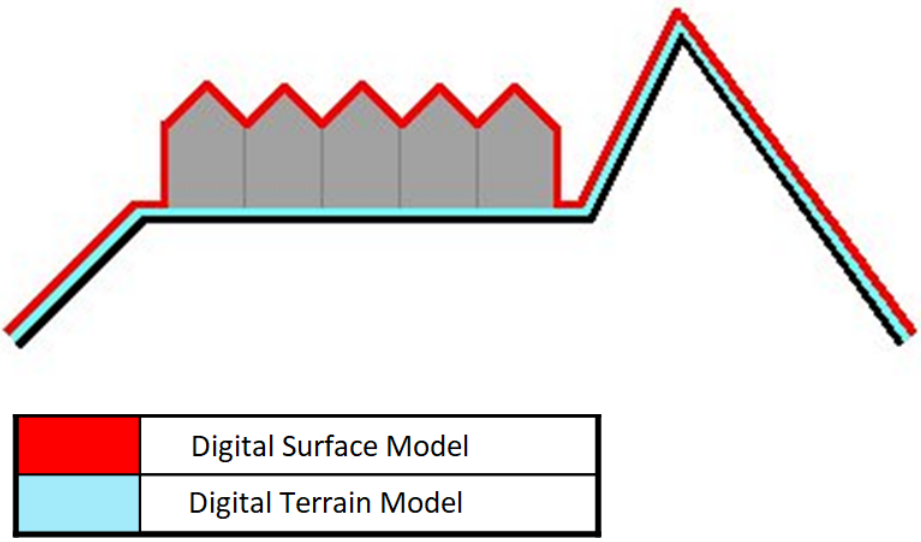


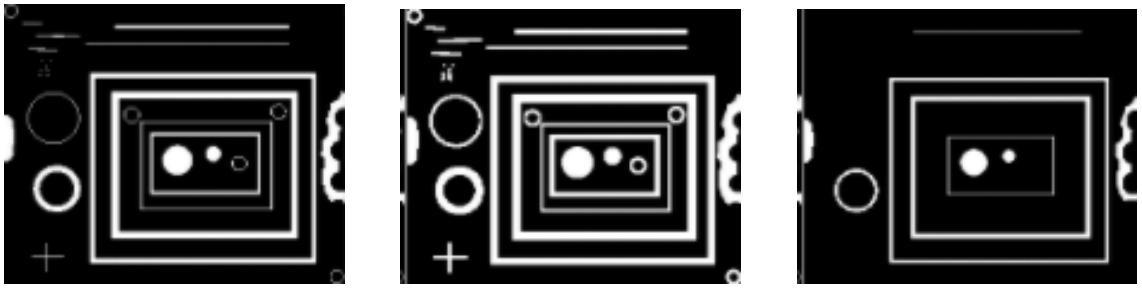
Figure 3.4: Illustration representing a Digital Surface Model that includes buildings, trees and other objects and a Digital Terrain Models representing the bare ground.

of an object in the image making the objects more visible and full, while erosion removes pixels on the object perimeter, resulting in the removal of islands and other small objects in order to focus on the substantive objects remaining. The number of pixels added or removed from the image vary according to a structuring element used during the process and can be tailored to a specific goal. The rules for both operations are:

- Dilation - The value of the output pixel, corresponds to the maximum value of all pixels in the neighbourhood.
- Erosion - The value of the output pixel, corresponds to the minimum value of all pixels in the neighbourhood.

By combining both operations in a certain order we can restore and recover images to the maximum extent, or smooth the contours of a distorted image and fuse narrow breaks and eliminate holes in the data. These operations are called opening and closing. The opening operation is accomplished by performing an erosion of the data set followed by a dilation, whilst a closing operation is achieved by first performing a dilation and an erosion after. The combination of opening and closing operations is generally used to clean up the data by eliminating features from the data set.

Common filtering methods based on those morphological operation rely on a base fixed window size. This means that most of the non ground objects are not identified correctly due to the its different sizes. The selection of the filtering window size is a problem. If the window is too small only small objects such as lamp posts, cars and trees will be removed, leaving buildings in the data set. On the other hand, if the window size is too big, the filter tends too remove more than it should resulting in for example flattened dunes. An attempt to overcome this problem is explained by [4] and consisted



(a) Original image (b) Dilation output (c) Erosion output

Figure 3.5: Results of application of the two basic morphological operators.(a) original image with no prior processing. (b) image after dilation operation - note the enlarged size of the objects present in the image. (c) image after erosion operation - most small objects are now eliminated.

of applying the morphological operations with varying window sizes, starting from the smallest, where each point is given a weight according to the window size (smaller the window, smaller the weight) if it was identified as a ground return. This resulted in a better derived DTMs however it does not improve in classifying ground and non ground points.

In order for our system to be able to extract DTMs and classify ground and non grounds points in urban areas, with buildings, cars, trees and also mountainous areas, where the main non ground features are vegetation we used the implementation of an automated progressive morphological filter (PMF), developed by [50] to classify the raw lidar pointcloud data into ground and non-ground returns.

In general, the first layer is derived by applying an opening filter with a window of size l to the data. The biggest non-ground features such as buildings remain because their magnitude is larger than l, while trees, lamp posts, or other features smaller than l are removed. For the ground, again, when a certain feature is smaller than l, they are cut off and replaced by the minimum elevation found in that place. In the next few iterations the window size is increased and another opening operation is applied, resulting in the smoother surface. This time if the building measurements are smaller than l they are removed and replaced by the minimum elevation of the previous surface. By performing this gradually the PMF is capable or eliminating buildings, trees and other features from the data set resulting in a 2-D flag array, where each value is classified between ground and non ground return. Pseudo-code regarding the previous explanation is described in Algorithm 1 where a description of the input parameters, output parameters and order of operation.

Input parameters such as window size and elevation changes are critical in order to achieve good results. In urban areas, where large non ground features exist, such as buildings an exponentially increasing window fares better, in order to reduce the number of iterations necessary to eliminate them from the data set:

$$w_k = 2b^k + 1 \tag{3.1}$$

where w_k is the window size, $k = 1, 2, ..., M$ and b is the initial window size given by the user.

For natural landscapes such as forests and mountainous sites where most features revolve around gradually changing topographics a linearly increasing window size is the most obvious choice, since it preserves the environment better:

$$w_k = 2kb + 1 \tag{3.2}$$

In this conditions the there is no need to study the elevation threshold dh_{max} and its normally defined as the largest elevation difference of the point cloud. The parameter's value for the PMF are presented in Chapter 4, as well as the results of the filtering process.

The second part of the algorithm is presented below in Algorithm 2 and is meant to iterate over the filtered data and whenever a point exists in the matrix of LiDAR points, an evaluation of the homologous flag is made and according to its value (check if it is zero) we classify the point as ground or non ground point.

Algorithm 2: Progressive morphological filtering - part II

```
for i = 0 to m do
    for j = 0 to n do
        if A'[i,j](x) > 0 and A'[i,j](y) > 0 then
            if flag[i,j] = 0 then A'[i,j] is a ground point;
            else A'[i,j] is a non ground point;
        end
    end
end
```

Subsequently to the point cloud classification process, we transfer the indices obtained by the PMF and use an extractor filter on the PCL library in order to negate and remove the subset of the identified ground points from the original point cloud, in order to separate and generate a digital terrain model, from the above ground returns, or vegetation, to be used for individual tree detection and forest inventory.

3.2.2 Outlier Removal

In 1980, Hawkins [12] classified an outlier as "an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism". Despite the robustness of new LiDAR systems, measurement errors can

Algorithm 1: Progressive morphological filtering - part I

Input : A LiDAR point cloud data, where each point is represented by (x, y, z) coordinates
Cell size, **c**
Initial window size, **b**, to be used in Equations 3.2 or 3.1
Maximum window size, $\mathbf{max}_{window}$
Slope, **s**
Initial distance, $\mathbf{dh}_0$
Maximum distance, $\mathbf{dh}_{max}$

Output : Two sub sets of point clouds, one representing the ground returns and the other representing the vegetation returns.

begin
  Determine limits, **x** and **y** of the data set;
  Determine the number of rows **m** and columns **n**;
  Create a 2-D array $\mathbf{A}[\mathbf{m}, \mathbf{n}]$ for the points within the point cloud. If more than one point fall into the same cell, the one with the lowest elevation value is selected. If the cell is empty, interpolate elevation value from its neighbours and negate x and y coordinates to differentiate;
  Initialize elements of a 2-D integer array $\mathbf{flag}[\mathbf{m}, \mathbf{n}]$ with 0;
  Determine series of $\mathbf{w}_k$ with Equations 3.2 or 3.1, ensuring $\mathbf{w}_k \leq \mathbf{max}_{window}$;
  Set $\mathbf{dh}_T = \mathbf{dh}_0$;
  foreach $\mathbf{w}_k$ **do**
    for $i = 0$ **to** m **do**
      $\mathbf{P}_i = \mathbf{A}[\mathbf{i};]$; /* $\mathbf{A}[\mathbf{i};]$ represents a row of points at row i in A */
      $Z \leftarrow \mathbf{P}_i$; /* Assign elevation values from $\mathbf{P}_i$ to 1-D elevation array */
      $\mathbf{Z}_f = erosion(Z, \mathbf{w}_k)$;
      $\mathbf{Z}_f = dilation(Z, \mathbf{w}_k)$;
      $\mathbf{P}_i \leftarrow \mathbf{Z}_f$; /* Replace z coordination values of $\mathbf{P}_i$ with treated values from $\mathbf{Z}_f$ */
      $\mathbf{A}[\mathbf{i};] = \mathbf{P}_i$; /* Put the filtered row of points $\mathbf{P}_i$ back to row i of array A */
      for $j = 0$ **to** n **do**
        if $Z[j] - Z_f[j] > dh_T$ **then**
          $flag[i, j] = w_k$
        end
      end
    end
    if $dh_T > dh_{max}$ **then** $dh_T = dh_{max}$;
    else $dh_T = s(w_k - w_{k-1})c + dh_0$;
  end
end

occur and the resulting point cloud may contain sparse outliers due to external factors that can corrupt the results even more, complicating the estimation of point cloud characteristics, such as surface normals or curvature changes. In point clouds from airborne laser scanners especially the appearance of outliers is a very common problem. Unwanted objects like overhead power lines, birds and even dust particles can come across the laser beam and reflect it generating incongruities in our data.

To evaluate our data, i.e. know something is not right or if it is far from the normal situation, we can measure the distance between the new observation and the rest of the data set (observed earlier), and judge the closeness of this new data point to the historical data set. In many applications, if we have fair confidence in the normality of the historical data set, a low distance would show the normality of new observation.

There are several methods that can be used for outlier removal, however we settled on the statistical outlier removal (SOR). This filter assumes that the distance between a certain point and its neighbours is normally distributed. Surprisingly the algorithm isn't very complicated and the step by step process can be found in the illustration below in Figure 3.6.

After the setup of the filter and parameter selection, a first pass is done to find the points (K) nearest neighbours and compute the mean and standard deviation of the distances from each point in the point cloud to their neighbours. A control threshold is calculated and the average distance is evaluated based on the sigma rule. If the result is not within the N standard-deviation from the mean, the point is flagged. On the second pass a sweep of all points is made, in order to eliminate all the points in the cloud which have been marked (in the first pass).

While the algorithm will in effect eliminate points which the distance to its nearest neighbours follows any statistical distribution, the remaining parameters, μ and σ, have only meaningful effect when applied to a data set that has a normal (Gaussian) distribution.

In this study, only the parameter (K) was considered, since the algorithm assumes that the data set has at a minimum (K) nearest neighbours for any given point in the point cloud; while the assumption of normality was not assessed. However, and even though the filter being designed mainly for indoor data sets (due to its sensitivity to density changes), the results obtained were good and in compliance with the expected.

3.3 Tree Top Detection and Segmentation

Tree top location is the basis of most forest inventory studies as it is critical for extracting key parameters such as tree height, or to delineate crown space, and possibly even distinguish different tree species. In previous work, the tree tops were usually determined by selecting the pixel with greater brightness in a certain individual tree crown in the image or, in 3D data by finding the highest point in a point cloud of a single tree. This worked relatively well, however, collecting and providing measurements for statistical analysis

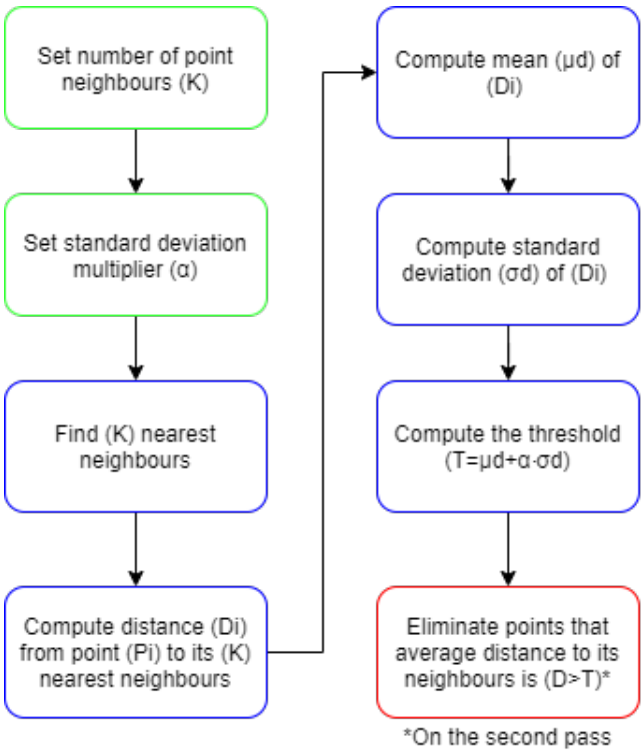


Figure 3.6: Walk through of the statistical outlier removal filter used for point cloud filtering. Highlighted in green - input parameters, in blue - nearest neighbour location and auxiliary calculations, and in red - outlier extraction.

based solely on these assumptions opens space for the introduction of serious estimation errors. Especially when the tree crown is large, considerable local maximas can be found on a single crown surface [18].

In a forest environment it is easy to imagine that there is horizontal spacing in between each tree, and that each of those gaps increases as we move further up into the tree and is reflected and more pronounced at the top of the tree. Our method takes advantage of this condition and according to the resolution conditions of the data set tries to segment each tree in a top to bottom approach, starting by identifying possible tree top locations and growing a region based on those seeds and fully build an individual tree. A bottom up approach was defined as well for high detailed data sets in order to start the process with a fully defined tree trunk, when possible, this way by locating the local minimas, and start growing the tree naturally from the base to the top.

One problem we could predict for both approaches was the classification of the points at lower to middle levels, as the spacing in between tress decreases and tree branches extend and overlap each other. However, since the main goal of this study was to develop a system capable of performing some sort of biomass estimation at plot level this will not be of great impact in the results.

Our method is a hybrid approach that combines local maxima and/or minima filtering

and region growing segmentation in order to separate individual trees into different clusters. The process starts by applying a simple local maxima (or inversely a grid minimum, depending on the nature of the data set) detection filter on the vegetation point cloud in order to select the possible candidates for tree top detection within a given window size. After a full sweep of the data set these points were marked as seeds and uploaded into the region growing algorithm to grow each segment.

This algorithm is very straight forward and by default, it starts by examining the input data and sorting the points by their curvature value. Once the cloud is sorted, and until there are no unlabelled points in the data set, the algorithm picks the point with the minimum curvature value, marks as a seed and begins the region growth. In our case, we fed the algorithm with the local maximas previously detected, however the overall functioning is remains the same as is as follows:

- For every point available in the seed set, the algorithm finds its point neighbours using a kd-tree.

- The normal angle of every neighbour is then tested against the normal of the current seed point. If the difference is less than a smoothness threshold, the point is suggested to be in the same cluster and added to the region.

- After that the curvature value is tested. If the curvature is less than a curvature threshold, the point is marked as a seed for the algorithm to continue the growth of that cluster, using the new point.

Once the seeds set is empty, the process is repeated from the beginning until there are no unlabelled points in the data set.

This way, we manage to get good results either with a high resolution data set and low resolution one, however we noted a slight bias towards larger trees as it becomes increasingly difficult to classify points at the lower level because the spacing between trees decreases, particularly for overlapping trees. In theory, one way to solve this was by using a Min-cut based segmentation to detect smaller trees that were represented by the local maxima filter. By computing the clusters centroid and its radius the algorithm tries to detect the edges of an object and divide the point cloud into two sets: foreground and background. The problem with this method is the need for the accurate object position in space, to try and separate it from the surroundings. Despite the effort, we couldn't observe any improvement in the results and so it was discarded. Since the main goal was to estimate AGB at plot level, smaller trees wont have greater impact and the mapping will not be affected.

3.4 Tree Trunk Detection

One of the main objectives in computer vision is the application of theories and models for scene reconstruction, object recognition and even 3D pose estimation. In forestry disciplines, when extracting forest parameters from LiDAR point clouds, the most common technique is to fit circles or cylinders to the individual data sets corresponding to the tree trunk or stems directly from the point cloud as a way to measure DBH, tree trunk density or volume. Methods such as least square adjustment, Hough transform and convex hull, have all showed promising results in order to help perform in depth forest analysis.

Our approach was based on the random sample consensus method (RANSAC) presented by [9] due to the ability of interpreting and smoothing the data sets containing a significant proportion of outliers. This method, when provided significant point coverage, is able to reconstruct a scene by interpolating a surface from points and fit the data to the desired mathematical model. Developed by the computer vision community, this method is different from conventional robust estimation techniques that use large amounts of data in order to provide an initial solution, and then proceeding to try and remove outliers. Instead, RANSAC generates candidate solutions by using the minimum possible number of points to estimate the elemental model and only then it proceeds to increase the data set with consistent data points. A common example would be the task of fitting a circle to a set of 2D data points, where the RANSAC algorithm would select three points (the required to define a circle), compute its centre and radius and check the compatibility of the remaining points to see if they are close enough to the model. If so, a smoothing technique is applied to improve the estimation.

The basic algorithm is summarised as follows:

Algorithm 3: Random sample consensus method - summarised

1: Select at random the minimum number of points to be tested ;
2: Compute for the selected model parameters;
3: Determine how many points from the input sub set fit the limits;
4: If the ratio of inliers to total points in the set exceeds a predefined threshold, the model was found and terminate;
5: Otherwise, repeat steps 1 - 4, a maximum of MAX times.

Figure 3.7 demonstrates a simple example of fitting a sphere to a set of 3D point observations. As explained before, RANSAC attempts to exclude the outliers and find a linear model that only uses inliers in its calculations. This is done by fitting this model to random samples of the data set and eventually returning the best fit. As expected, a subset only containing inliers will give the best model, so RANSAC keeps on computing and gathering random samples until it can find the best fit. In practice, there is no guarantee that a sub sample of inliers will be randomly selected, as the probability depends on the ratio of inliers in the dataset as well as the input parameters.

One key aspect of the RANSAC method is its ability to perform robust estimation for

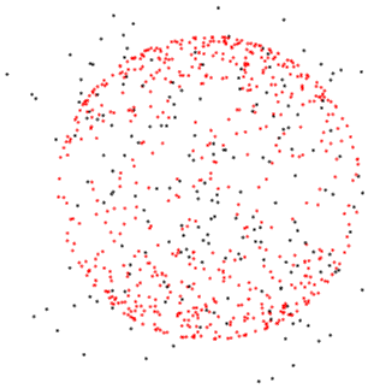


Figure 3.7: Display of the indices of the original point cloud (black) and the ones that satisfy the chosen model (sphere in red). The same can be done for a simple line or a cylinder to model tree trunks.

a number of model parameters, with a high degree of accuracy, even in the presence a relevant number of outliers, however it is not always able to find the optimal set even for moderately contaminated sets and presents poor performance when the number of inliers is less that 50% of the data sets population. A big disadvantage of RANSAC is that there is no limit on the time it can take to compute for a certain model, barring exhaustion. However, if we limit the number of iterations, the solution obtained may not be best, and it may not even fit the data properly. In this way RANSAC offers a trade-off where the higher the number of iterations is, the probability of a reasonable model being produced increases. In urban environments where exists the necessity of identifying more than one object with different model parameterisations, RANSAC may fail to find a fit, as it was designed to only estimate one model for a particular data set. In such cases other methods such as Hough transform can be used, however in forestry environments this does not pose a problem.

3.5 Detectable Attributes

Forest inventory is defined as a systematic collection of data and forest information to be used for assessment or analysis. This data is extremely important in forest management as it can help predict, monitor and evaluate the state of a forest during every stage of its development and as a control for post fire assessment systems. When taking forest inventory, the most important aspects to take into account are: species, height, DBH and crown span. From these we can then calculate numerous quantities.

Once an individual tree is fully segmented we can start extracting relevant direct measurements:

3.5.1 Tree Attributes

1. Position: Outputs the (x, y, z) coordinates of the tree base in the Cartesian coordinates system.

2. DBH: Determines the diameter at breast height of the detected tree trunk from a subset of point between 1.25 and 1.35 meters above the tree base.

3. Height: Calculates the tree height by subtracting the ground height at the tree position from the highest point of the tree.

4. Cloud length: Outputs the distance between the two furthest points of the cloud, useful for trees that are inclined.

5. Tree points: Outputs the number of points representing a single tree.

3.5.2 Crown Attributes

1. Crown centroid: Outputs the (x, y, z) coordinates of the tree base in the Cartesian coordinates system.

2. Crown height: Calculates the crown height by measuring the vertical distance between maximum and minimum point in the crown cloud.

3. Crown volume and surface area: Calculates the crown volume and surface area based on the current literature.

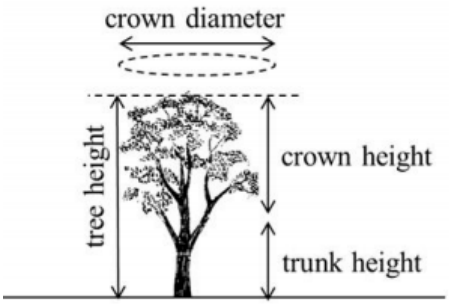


Figure 3.8: Example of tree features and how to measure them. Adapted from [46]

3.5.3 Above ground biomass estimation

The main goal of this study, beyond tree detection and parameter extraction was to perform an above ground biomass estimation and mapping in order to detect biomass changes in the forest structure after a wildfire event. For that we needed to extract the features enumerated in Table 2.2.

Many LiDAR derived measurements can be used to classify and assess above ground biomass. Figure 3.9 shows a diagram of all the possible tree characteristics that can be used for such study.

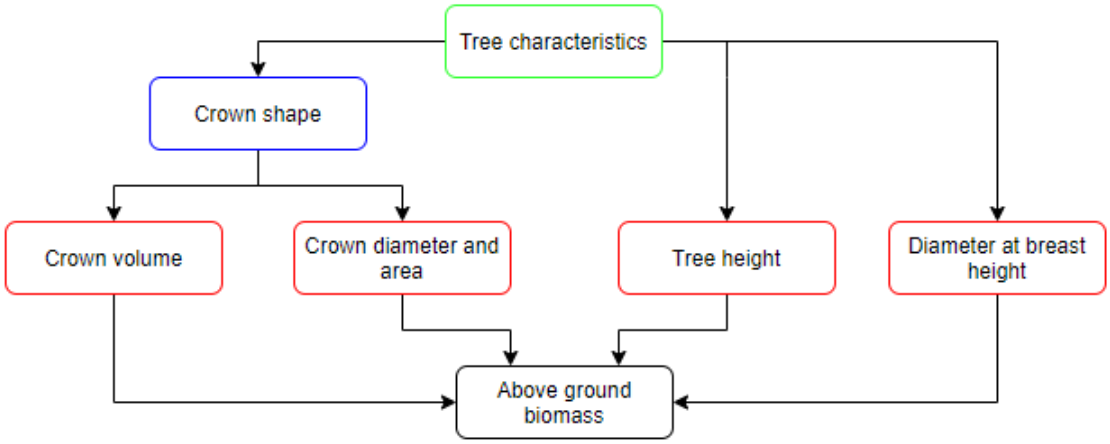


Figure 3.9: Flow diagram of all possible tree measurements to be used for AGB estimation.(Blue - shape approximation; Red - direct measurements).

As stated earlier in Section 3.2, point cloud density very important and in some data sets it is impossible to extract relevant information from tree because the laser couldn't penetrate the canopy. when that happens and the resulting vegetation point cloud has lower resolution the detailed segmentation of the tree trunk is not possible. So we had to estimate the AGB and used the equation developed by [37] that represents 69% of all forest types with a root mean square error of 4.27 $kg.m^{-}2$:

$$AGB_{LiDAR} = 0.36\mu_z^{1.16} z_{75}^{0.78} z_{10}^{-0.18} k_z^{0.41} \tag{3.3}$$

where AGB_{LiDAR} is the above ground biomass model in $kg.m^{-2}$, μ_z is the mean of all height returns in m, z_{75} and z_{10} are the third quartile and tenth percentile of height and $k_z$is the kurtosis of the distribution of all height returns. From this, it was possible to build biomass maps with a 25m x 25m and 50m x 50m resolution grid.

On the other hand, when the resulting tree cloud is of high detail, by modelling the tree trunk as a cylinder and cutting a slice of the tree trunk at approximately 1.3 meters high it is possible to measure DBH and with the work developed by [5] calculate the biomass of each individual tree using Equation (3.4):

$$AGB_{est} = 0.0673 \times (\rho D^2 H)^{0.976} \tag{3.4}$$

where AGB_{est} is the above ground biomass in kg, D is the diameter at breast height in cm, H is the total tree height in m and ρ is the wood density in g/cm^3. From this equation we can estimate the above ground biomass for an entire plot of, for example, 50 by 50 meters

and build a biomass map of the entire region, which gives us some knowledge regarding the state of the forest and the adjacent ecosystems.

A correlation between canopy volume and above ground biomass showed that the geometry of the tree canopy are directly related to tree growth and can be used estimate biomass, and even assess its health. General methods for calculating canopy volume use a predefined formula that takes into account the crown shape[46]:

$$CanopyVolume = CanopyHeight \times (CrownDiameter) \times ShapeMultiplier \tag{3.5}$$

The shape multiplier varies according to the shape of the tree crown, which are typically approximated to a spheroid, ellipsoid, circular cone or a cylinder. Since we couldn't compute the shape approximation, and be performing a visual assessment of the test plots we decided to set the shape multiplier to 0.3927 which is the multiplier for a paraboloid, and compared the results, using the formula used by [17]:

$$CanopyVolume = \frac{1}{3} \times \pi \times TreeHeight \times CrownWidth \times CrownLength \tag{3.6}$$

Seeing that we couldn't collect ground truth, this way we can provide a comparison between canopy volume a biomass maps in order to provide a more accurate estimation.

3.6 Discussion

In this chapter we covered the methods we used to process the LiDAR forest point clouds in order to detect segment individual trees for forest inventory purposes. The LiDAR system was reviewed and the method descripted, starting with data pre processing where we gave a brief explanation on the topic of LiDAR point density and the advantages and disadvantages of high and low point resolution data sets. A differentiation between digital elevation models and digital surface models was clarified and dilation/erosion operations performed in order to extract them were explained as well as the pseudo code for the progressive morphological filter that makes use of them. Despite the advancements, LiDAR data still can capture high amounts of outliers due to environment conditions which have to be removed; in this work we used a statistical outlier removal filter. Tree top detection was performed by applying a local maxima filter and coupled with the region growing algorithm we were able to segment and extract individual trees either in high or low resolution data sets. A few attempts at separating smaller and overlapping trees were made using min-cut segmentation however the results were not sufficient and it was discarded. Tree trunk segmentation was performed with RANSAC method where we managed to extract the tree trunk and by cutting a horizontal slice from the subset we are able to measure DBH which is used for calculation of AGB. A list with a description of the detectable attributes is presented to give the reader an overview of the system output. In the next chapter the results will be presented, with detailed explanations and illustrations that show the tests made throughout the development of the system.

Chapter 4

Results

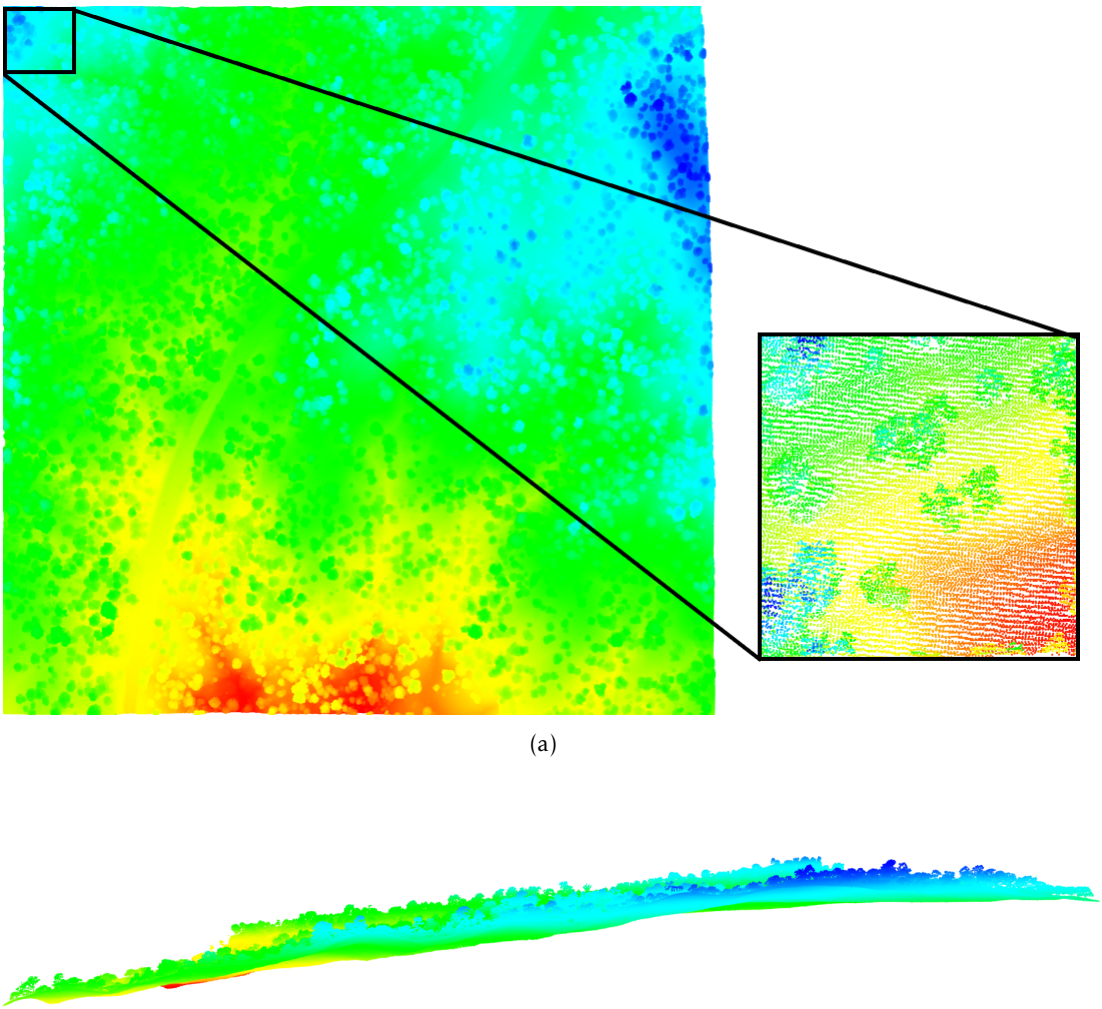
This chapter goes through all the materials we used for the study, presents the parametrisation of the algorithms and the thought process behind those, and finally, gives an in depth analysis with detailed figures of the results obtained during the experimental trials.

4.1 Experimental Setup

The proposed system was developed and implemented in the C++ programming language and made fully compliant with the Robot Operating System (ROS) framework. The system makes use of three libraries that work together to give a more general and capable tool for analysis of 3D forest point clouds:

1. PDAL library, used for translating and manipulating point cloud data;
2. PCL library, used for filtering, feature estimation and segmentation of point clouds;
3. OpenCV library, used for computer vision and image processing.

The low density materials used for this study were provided by the National Ecological Observatory Network and according to the documentation, collected using two Optech ALTM Gemini for discrete return LiDAR point cloud. The flying altitude was 1000 meters above ground level (AGL) and the density of the LiDAR data was approximately 4 points per square meter,across a 100 ha study area. The point cloud has 6.6 million points and covers an area of 1000 meters by 1000 meters, resulting in an average point density of 6 $points/m^2$. Due to the computational power necessary to manipulate a file this large, a spatial grid was applied and the data set sub divided into smaller plots to shorten the amount of processing time and to make the tests run smoother and faster. To get the



(a)

(b)

Figure 4.1: Data set A: Original low resolution point cloud out sourced from [27] that covers an area of 1000 by 1000 meters. (a) Top view of the original point cloud with a zoomed in 50 by 50 meter sub set prior to any process. (b) Side view of the original point cloud with the slope in evidence.

desired mapping resolution the grid was divided into sub plots of 50 by 50 and 25 by 25 meters.

This point cloud has a considerable ground elevation difference of about 120 meters which is why the height normalisation is imperative. In addition, the region contains several objects that are easily detected to the naked eye as outliers, such as roads, cars and light poles. The data set presents low density and as a result tree trunks do not appear and cannot be detected and so a canopy only model was used to extract LiDAR metrics. The data set is provided in the LAS format, which requires translation for the PCD format in order to process it. To do that we used the PDAL library.

In order to test the full capabilities of our system, high density materials were out-sourced and were gathered by a mobile laser scanner that covered a 20 meter by 45 meter

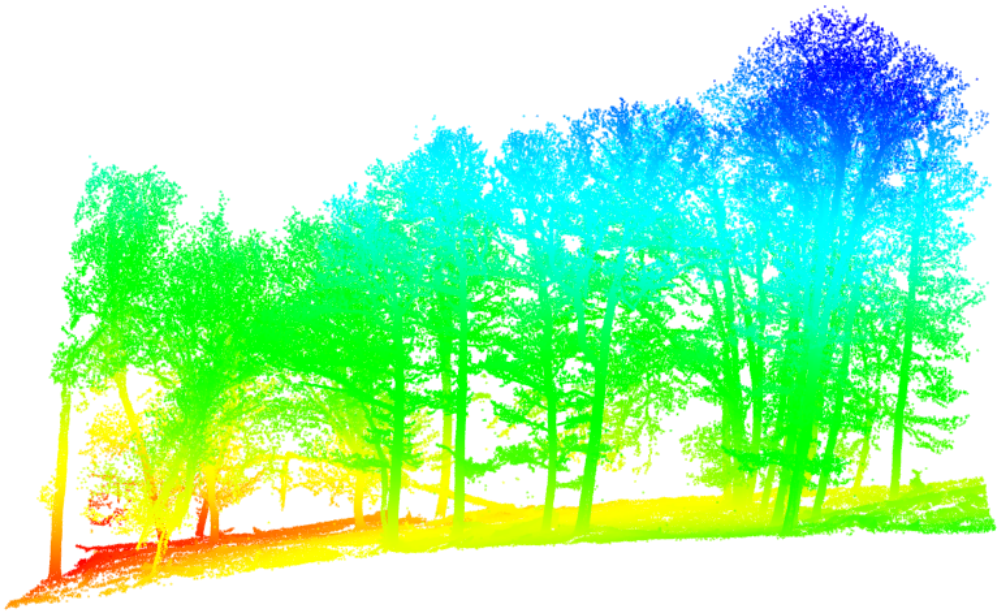


Figure 4.2: Data set B: Original high resolution point cloud out sourced from [15] that covers an area of 20 by 45 meters.

area and has 474.269 points resulting in an average point density of 500 $points/m^2$. This point cloud illustrates a high complexity forest, with trees varying in sizes, from small to bigger trees, dead trees and low vegetation and a ground elevation difference of 16 meters. The data shows high detail and we can clearly see the delineation of a tree trunk which is suitable for the testing of our algorithms 4.2. The reference tree composition complied of 26 trees and measured a mean height of 19.25 meters and a mean diameter at breast height of 68.7 centimetres.

In each of the given plots we defined the correct detection of a tree as "success"and divided this number by the number of trees detected manually. We performed this operation for every plot in the data set and calculated the average success rate of the algorithm. Furthermore, in order to evaluate our model, the coefficient of determination, or R^2, is used here to demonstrate the so called goodness-of-fit by measuring it's strength.

4.2 Progressive Morphological Filter

The progressive morphological filter was tested on both data sets to ensure its filtering capabilities on different environments. The filtering parameters are listed in Table 4.1. The selection of the parameters is critical for mixed environments such as an urban environment with large forest surroundings where the filtering has to be more precise. Since we are using only forested data sets, this has more margin for error and so some of the parameters were accomplished based on a trial and error approach. The window

Table 4.1: Progressive morphological filter parameters for ground extraction.

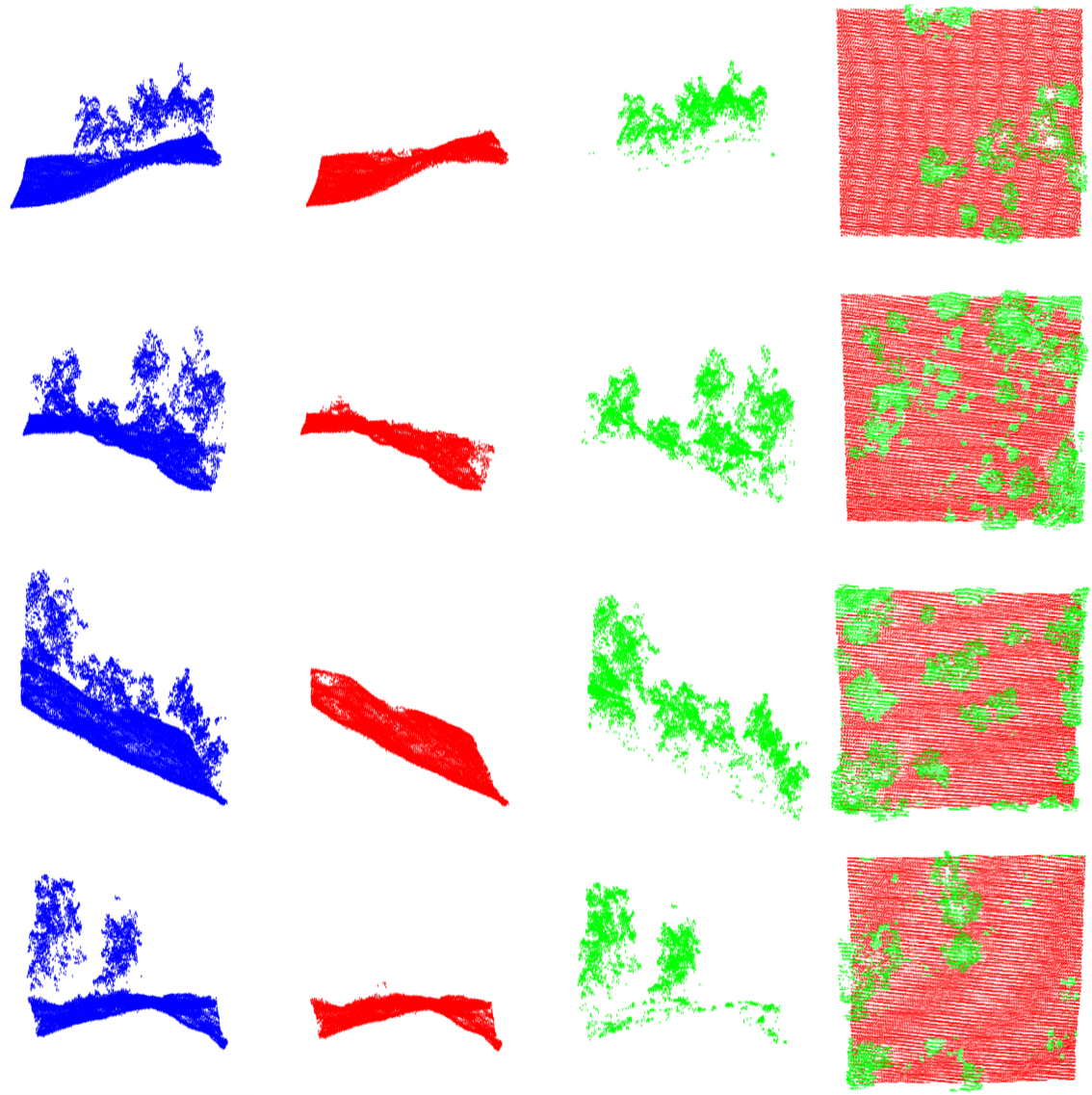
Variable	Value
Maximum window size	20
Initial distance	0.5 (f)
Maximum distance	3.0 (f)
Slope	1.0 (f)

size is computed in an exponential manner using Equation (3.1) where the maximum window size was left default, to be able to extract buildings in case there were any. Initial distance was set up as 0.5 meters, which is approximately equal to the elevation accuracy described for the LiDAR in use (ranging from 5 to 45 centimetres). It is easy to picture that there is an abrupt change in elevation when a tree is encountered, in regards to the adjacent ground points, and so, the height threshold for this purpose was set to 3.0 meters. The terrain slope are relatively steep and so the slope value was set to 1.0.

For the data set A, we detected 4.749.768 ground points representing 72% of all points in the data set and 1.860.061 were classified as vegetation. This happens due to the altitude and area covered by the airborne laser scanner. If the point density were to be increased, this ratio might decrease significantly, since the number of points per tree would be larger. Figure 4.3 shows 4 different sub plots from the original point cloud, with the ground returns in red and vegetation returns in green, and a top view after the PMF, separated for further processing, where it is easy to point out individual tree crowns in each data set, however it is difficult do distinguish if there are overlapping ones.

As for the data set B, the same ratio of points does not apply. Since the number of points per square meter is increased, the number of points in each tree is exponentially bigger, and so we managed to identify about 57.858 ground points, which is approximately 12% of the number of points in the data set, which means that the remaining 426.337 points comprehend the vegetation and some outliers, which will be removed next.

As expected, in Figure 4.3 it is relatively easy to see that the resulting vegetation point cloud has large number of outliers that can negatively impact our segmentation results, leading to over estimation of trees. The same can be said regarding figure 4.4 which present outliers at lower levels, due to the presence of dead trees and bushes. All of this can be cleaned up and removed by applying a statistical outlier removal filter, as will be shown in the next section.



(a) (b) (c) (d)

Figure 4.3: Results of the application of the progressive morphological filter to four different sub sets of the original 1000 by 1000 meter point cloud.(a) Original sub sets with no prior processing. (b) Ground returns classified by the PMF. (c) Vegetation returns resulting of the subtraction of the DTM from the original sample. (d) Top view with the ground returns (red) separated from the vegetation returns (green).

Figure 4.4: Original high resolution point cloud of figure 4.2. From left to right: Original point cloud, ground returns classified by the progressive morphological filter and again, vegetation cloud, resulting from the subtraction of the ground returns from the original sub set.

4.3 Statistical Outlier Removal

To test the statistical outlier removal we began by performing a visual assessment of the vegetation point clouds and counted the possible inliers of the data sets to make sure that the filter parametrization is done properly. By measuring the mean distance of a point to its neighbours and plotting the results we concluded that the spikes in the data were not normal, meaning that the point that caused that disturbance is too far from the rest of the points and is probably an outlier. Table 4.2 lists the parameters for the filter. The number of neighbours to analyse for each point was set to 50 which we concluded was the sweet spot for a good filtering and time saving, and the standard deviation multiplier to 1. This means that all points that have a distance larger than 1 standard deviation of the mean distance to the point in question, will be marked as outliers and later removed.

Table 4.2: Statistical outlier removal filter parameters for outlier detection and removal.

Variable	Value
Number of Neighbours	50
Standard deviation multiplier	1

After the filtering process we calculated again the mean distance of a point to its neighbours, plotted the results and concluded that the points that previously damaged the results were now eliminated, resulting in a more condensed data set. The number of outliers in each data set directly correlates to the its size, as the number of points increase, the number of outliers removed increased in the same manner. Figure 4.5 (a) shows the differences in the data sets before and after the filtering process, where before we could see excessive spikes in the data, after the application of the filter the results were much

more clean and the small islands disappeared. Figure 4.5 (b) shows a top view with the outliers highlighted in red. Once those were removed the segmenting procedure was much easier, and the time it takes to do so is decreased.

For the data set B the same procedure was taken, however, due to the high point density we experimented with the filter parameters to optimise the number of outliers detected. We increased the number of neighbours to be tested from 50 to 100 and reduced the deviation multiplier in order to reduce the number of point while maintaining the key features. This resulted in increased processing time and a small output difference, with an extra 1776 points eliminated. When doing this we began to have problems in the latter stages when trying to segment the individual trees, and so we limited the variables to their default of 50 neighbours.

4.4 Local Maximas and Region Growing

The segmenting process started with the search of the local maximas in the data set. The filter setup is simple and only requires the input radius for the search. This input is important since the radius determines the number of local maximas that can be found in a determined region. This can cause over estimation errors if the window is too small, detecting more than one local maxima for a single large tree crown, and under estimation errors if the window is too big, finding only one maxima for more than one tree. This method as proved to be bias towards larger trees, and so we only used it to feed the region growing algorithm and enrich the seeds it gathered by itself.

Table 4.3: Local maxima and region growing parameters for individual tree segmentation.

Variable	**Value (data set A)**	**Value (data set B)**
Local Maxima Radius	6 metres	-
Search Method	Tree	Tree
Normal Search	50	15
Minimum Cluster Size	50	1200
Maximum Cluster Size	20.000	90.000
Number of Neighbours	20	40
Smoothness Threshold	5	0.5
Curvature Threshold	1.0	1.0

The manual assessment of the vegetation clouds was key in order to correctly choose the parameters. We can see that with the lower point density, the amount of points in each cluster in very small comparatively to the data set B. This means that we have to limit the each cluster size to stop that isolated points that remained after the outlier removal from

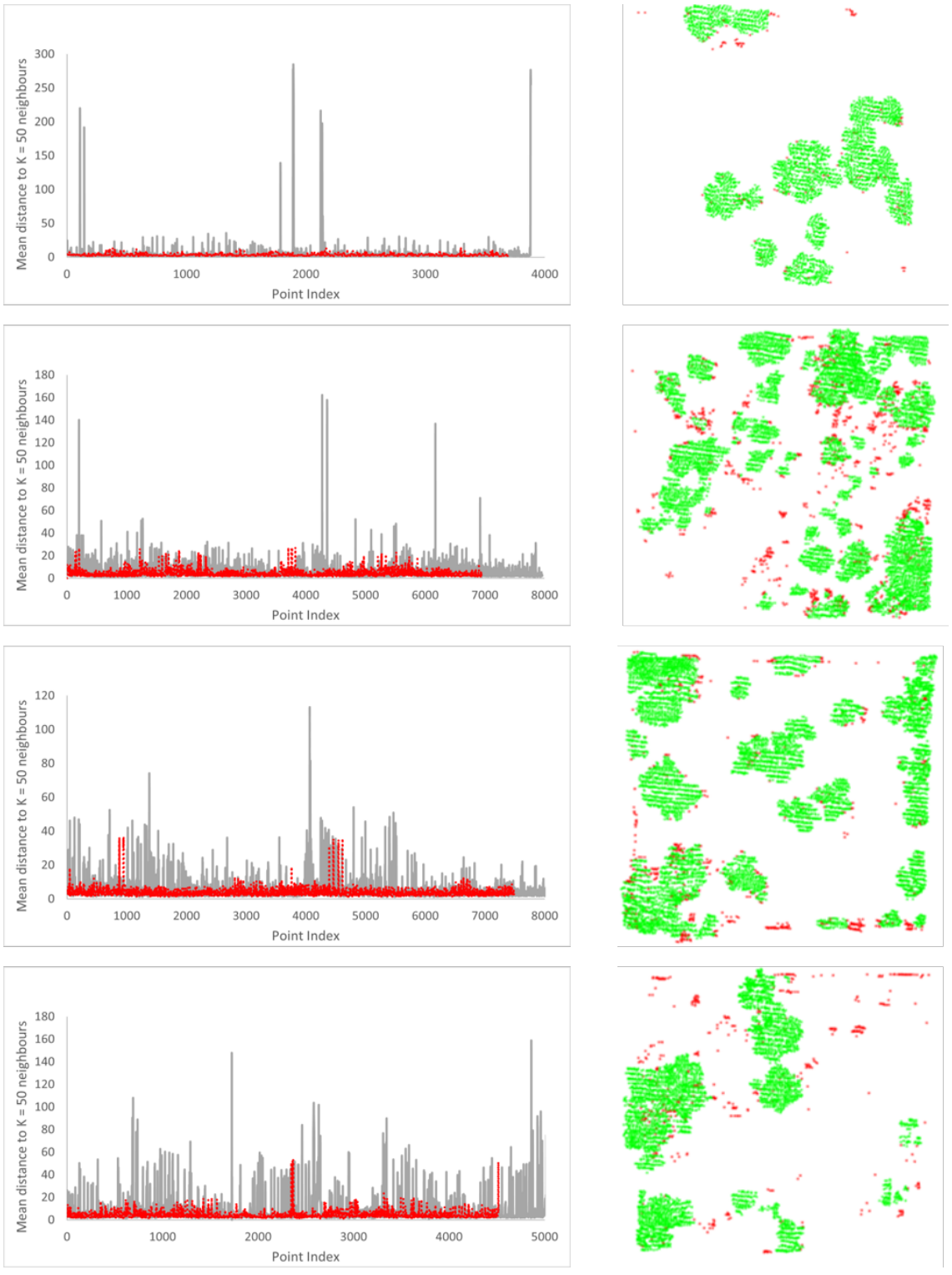
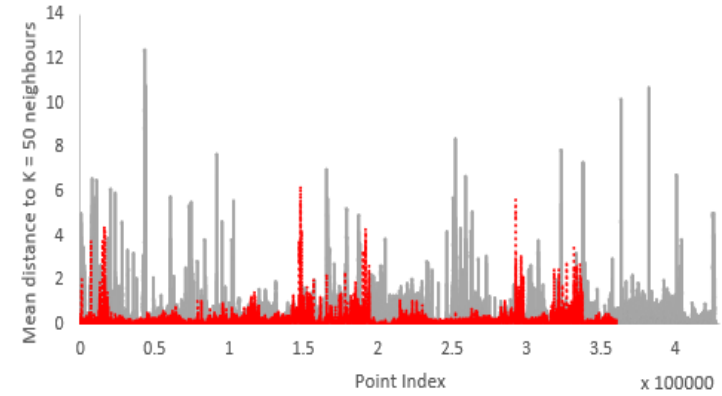
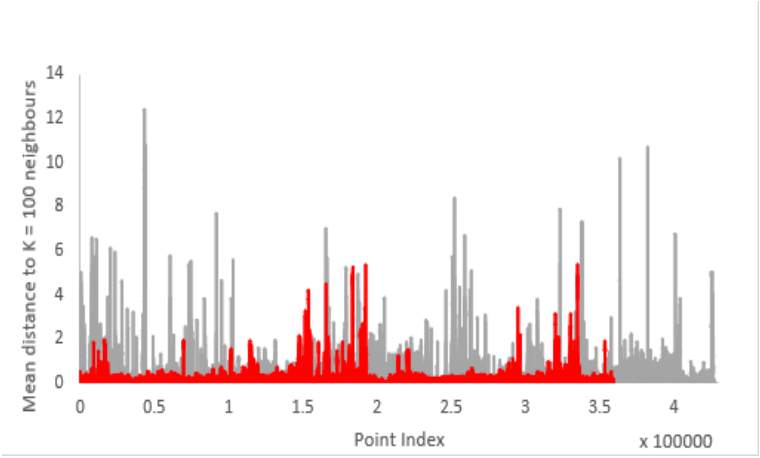


Figure 4.5: Results of the application of the statistical outlier removal filter to four different sub sets of the original 1000 by 1000 meter point cloud. For the 4 plots here shown an average of 600 points were removed. (a) Plot of the mean distance of a point to (K) = 50 nearest neighbours - full line (grey) represents the point cloud before the SOR filter applied; dotted line (red) represents the filter output without the presence of sparse outliers (b) Top view with the vegetation returns (green) separated from the outliers in the sub set returns (red).



(a)



(b)

Figure 4.6: Plot of the mean distance of a point to (K) nearest neighbours - full line (grey) represents the point cloud before the SOR filter applied; dotted line (red) represents the filter output without the presence of sparse outliers (a) K = 50 nearest neighbours. (b) K = 100 nearest neighbours. The difference between the two is explained in 1776 points removed, however the number of spikes in the data set remained.

being classified as a small cluster leading to errors, functioning as a second layer filter of some sorts. With that in mind, we setup the filter as shown in Table 4.3, where the minimum and maximum cluster sizes reflect the amount of points we think are enough for segmenting an individual tree. As expected, the data set with higher density has to be parametrized accordingly, since we concluded that the resulting vegetation point cloud had about 88% of the total points.

In forestry, the point clouds are very unorganised, and sometimes its hard to segment clusters with different characteristics, even if the distance between them is decreased and a few more tests have to be considered. The first one is the smoothness threshold that is responsible for testing the deviation between point normals. In the low density data set, since the points are spaced out, the deviation between the normals is higher, and so we set them to 5 degrees. Whereas in the high density data set, points are much more close

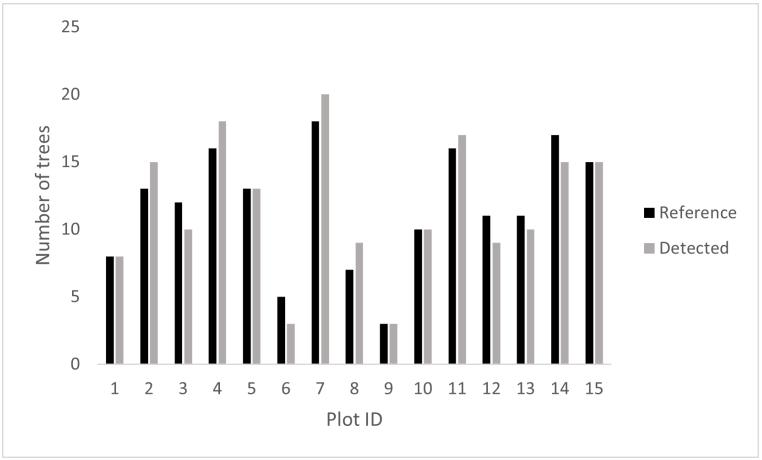
to each other and the normal deviation is much lower, so the angle is tested for a smaller threshold. If this deviation is computed to be less than the limit then the current point is added to the cluster. In some cases, the normal deviation can be small and the point belong to another cluster, and so its curvature is tested. Varying this parameter didn't seem to make any difference so it remained default.

Table 4.4: Detection and segmentation algorithm over a sample of 15 plots with different levels of complexity, category discriminated.

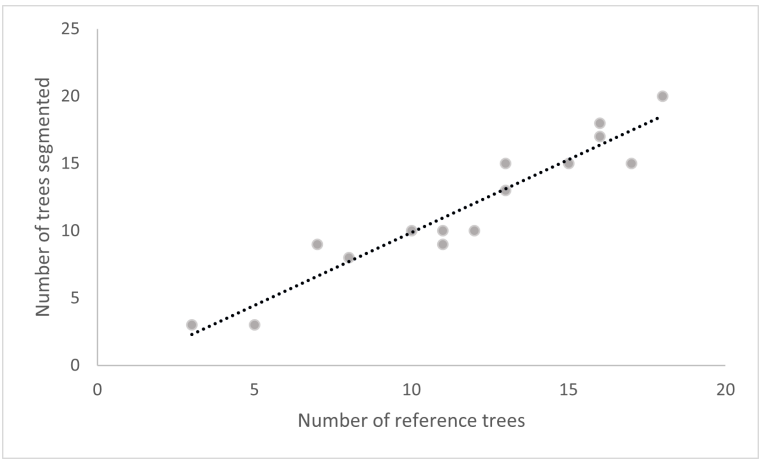
Plot id	**Category**	**Truth**	**Detected**
1	A	8	8
2	A	13	15
3	A	12	10
4	A	16	18
5	A	13	13
6	B	5	3
7	B	18	20
8	B	7	9
9	B	3	3
10	B	10	10
11	C	16	17
12	C	11	9
13	C	11	10
14	C	17	15
15	C	15	15

Due to the lack of ground truth data we began by manually assessing the low density data set and selected 15 plots of a 25 by 25 metres grid layout with different characteristics and categorised them between low (category A), medium (category B) and high (category C) complexities, ranging from a few spaced trees to conglomerates of trees with varying sizes. Table 4.4 shows the results of the detection and segmentation algorithm in such conditions, allowing us to properly calibrate the system and validate its results.

The algorithm presented obtained an average success rate of 88%, a mean absolute error of 1.2 and a corresponding R^2 value of 0.88, representing a good fit over the 15 test plots. The impossibility to detect the tree trunk in the low density data set hinders this approach since it becomes increasingly difficult to separate trees close together only from the tree crown.



(a)



(b)

Figure 4.7: Detection and segmentation algorithm over a sample of 15 plots with different levels of complexity. Comparison between reference and detected trees in each different plot.

In the data set B, we managed to identify 22 out of 26 trees present in the area, which represents a success rate of about 85%. When trees have overlapping branches the algorithm keeps on growing the region, showing 2 trees in the same cluster. Despite being a problem when performing forest inventory, if both tree trunks are visible, when estimating the above ground biomass from DBH it will be corrected. In Figure 4.9 it is possible to see in some areas two or more trees segmented in the same colour, meaning that the algorithm thinks they belong to the same cluster, resulting in under estimation errors. The results here present don't have much scientific meaning since there was only one available data set in this conditions, however the purpose of this was to try and identify the tree trunks, which is shown in the next section.

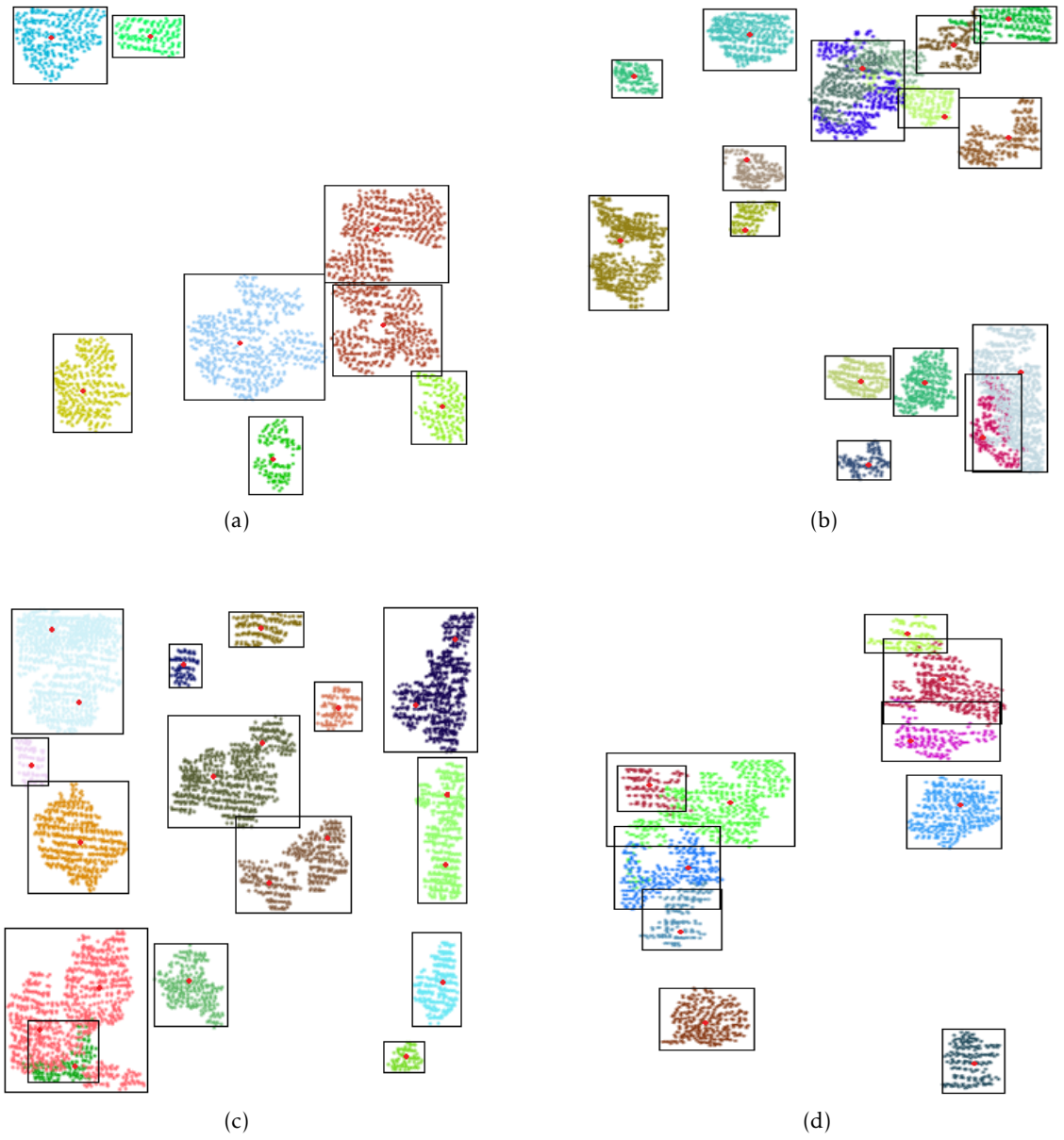


Figure 4.8: Detection and region growing segmentation algorithm of 4 sub sets of different categories. (a) Plot id - 1 (cat. A) - 8/8 trees detected. (b) Plot id - 11 (cat. C) - 17/16 trees detected (over estimation error). (c) Plot id - 14 (cat. C) - 15/17 trees detected (under estimation error). (d) Plot id - 10 (cat. B) - 10/10 trees detected.

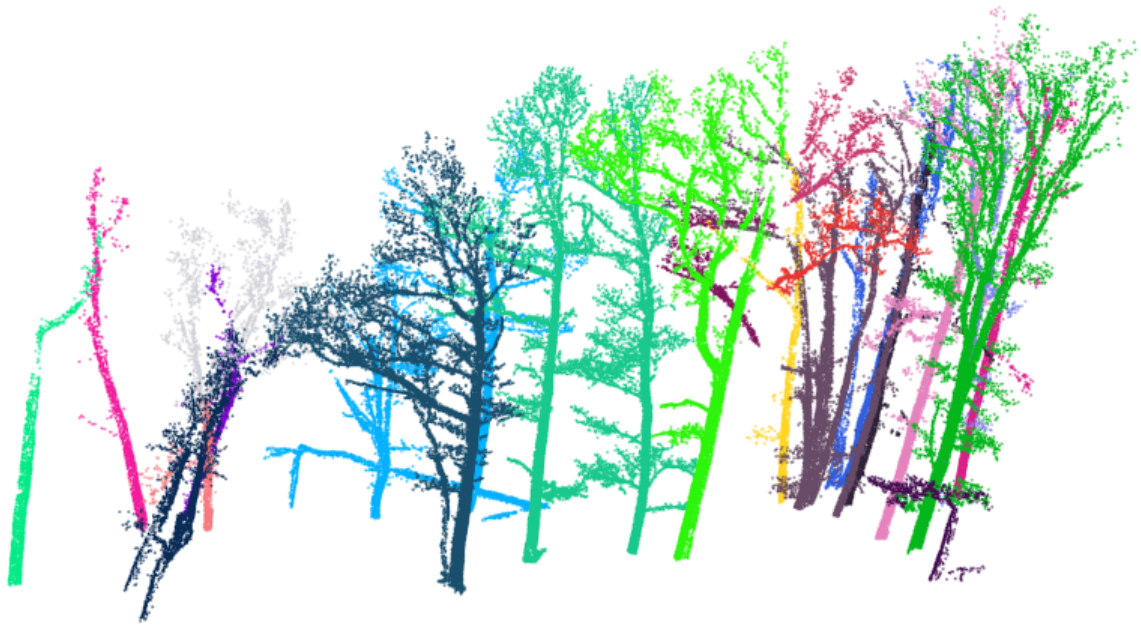


Figure 4.9: Detection and region growing segmentation algorithm for the high point density data set. 22 out 26 trees were correctly identified and will be later processed for tree trunk recognition.

4.5 Tree Trunk Detection

The detailed process of identifying and extracting the tree trunk using the random sample consensus algorithm was described in chapter 3, section 3.4. After individualising each tree cluster, we modelled the tree trunk as a cylinder and extract it using RANSAC. Doing so, from the 22 detected trees we were able to separate the trunk from all the trees. The parameters for tree trunk extraction were the following:

Table 4.5: Random sample consensus parameters for tree trunk delineation.

Variable	Value
Model Type	Cylinder
Method Type	RANSAC
Normal Distance Weight	0.1
Maximum Iterations	10000
Distance Threshold	0.7
Radius Limits	0.0 - 0.5

We are imposing a distance threshold (used to determine when a data point fits the model) from each inlier point to the model no greater than 7 centimetres. In addition,

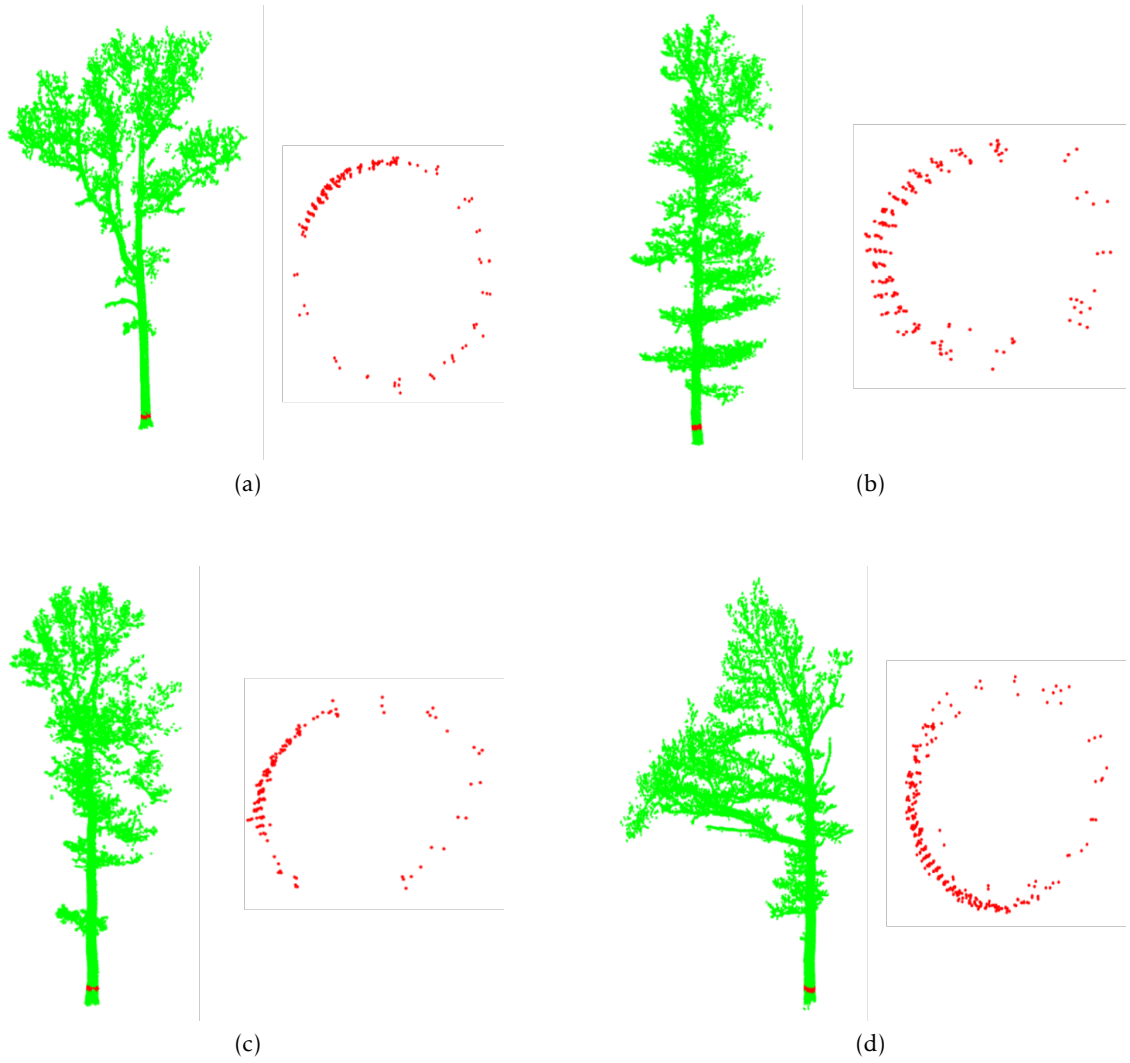


Figure 4.10: Horizontal slice of the tree trunk using after tree trunk detection by RANSAC. Front view with the tree and the horizontal cut highlighted in red and top view of the slice.

we set the surface normals influence to a weight of 0.1, since the point density is higher and the points a close to each other, and we limit the radius of the cylindrical model to be smaller than 50 centimetres.

Figure 4.10 shows the full tree cluster, and a slice of the segmented tree trunk in between 1.2 and 1.4 metres. From the trees detected,we measured an average of 24.000 points per tree and only an average of 250 points were used for the horizontal profile cut. From this we are able to extract, not only the full tree height by scanning the limits of the point cloud, but to measure the diameter at breast height, which is the main measures taken by foresters.

This worked relatively well, however, we noted that if a tree is at an angle, the DBH is not measured at the right height. This might be a problem in the future, however,

one way to overcome it is by analysing the angle of the normal of the points in the tree trunk. This way the height at which the slice is made can be corrected and the DBH can be measured correctly.

4.6 Above Ground Biomass Estimation

Once all the segmentation is done we can begin to extract measurements from the data sets, useful for our application.In order to test the biomass present on a given plot, the equation used takes advantage of the height returns of the canopy, resulting in fewer errors of estimation.

The biomass for the whole plot was then computed using Equation (3.3) and a biomass map of the region was drawn to offer an easier assessment of the current state of the forest. The results obtained showed that the biomass is directly correlated with the height and crown volume, since it primarily accounts for the number of returns in the tree canopy. The presence of outliers was corrected and the data optimised, resulting in an R^2 of 0.92. The average tree height was around 14 metres and the above ground biomass of 345 kg/m^2.

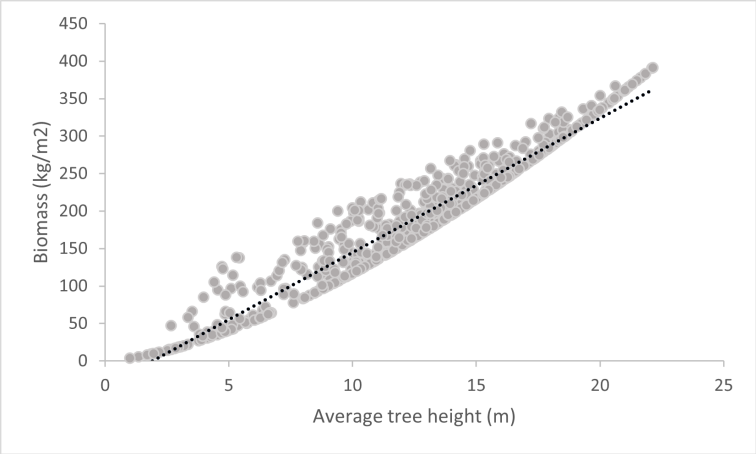
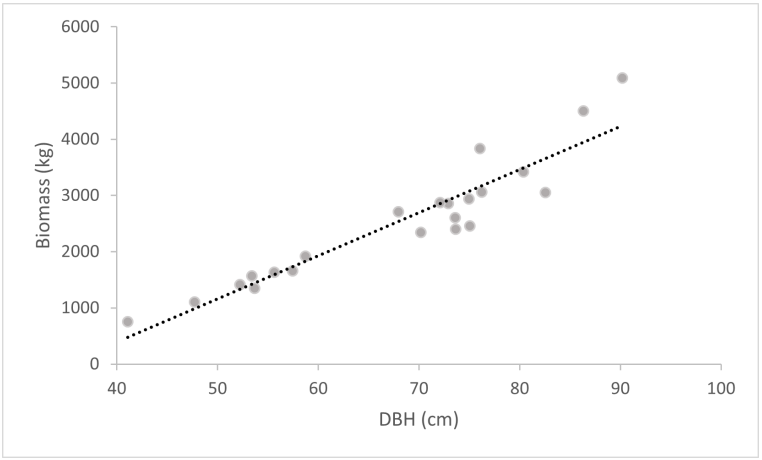


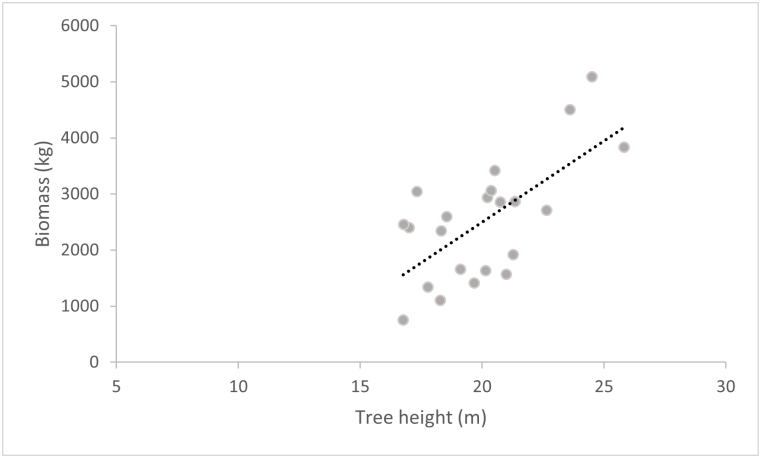
Figure 4.11: Scatter plot of above ground biomass in $kg.m^{-2}$ in relation to the average tree height on a given grid plot size, of the entire area using Equation (3.3).

Figure 4.11 illustrates the correlation between the average tree height on a grid cell and the above ground biomass per square meter.

For the data set B, since we managed to detect the tree trunk, it was possible to extract the DBH and calculate the AGB directly, as its done commonly. Because the species of the trees was not available, an average tree density ($\rho = 500kg/m^3$) was used for all trees and a comparison between the DBH and the tree biomass was plotted in Figure 4.12a, resulting in a R^2 of 0.8761, versus an R^2 of 0.4434 when compared to the tree height, as in 4.12b, meaning that the estimation of AGB solely based on tree height is not an effective estimation method. The average height stands at around 20 metres and an



(a)



(b)

Figure 4.12: Scatter plot of the biomass of individual reference trees (22/26) expressed in *kg* in relation (a) to its tree trunk diameter at breast height and (b) in relation to tree height.

average diameter at breast height at 67.8 centimetres whilst the total biomass available was around 55.5 metric tonnes.

In the end a biomass map was drawn in order to quickly identify the state of the forest, and in the future serve as a reference when comparing pre and post fire biomass changes or to accompany and evaluate the progression of the forest along the years (Figure 4.13a). In order to compare the biomass results to the average tree height (Figure 4.13b and the average crown volume (Figure 4.13c) in each plot, maps with the same resolution were computed. As it was shown in Figure 4.12a since the AGB is calculated from the height returns of the tree crown, the similarities are more pronounced. When comparing the biomass and volume maps it is clear that crown volume, in this case, does not directly correspond in the same proportion to AGB.

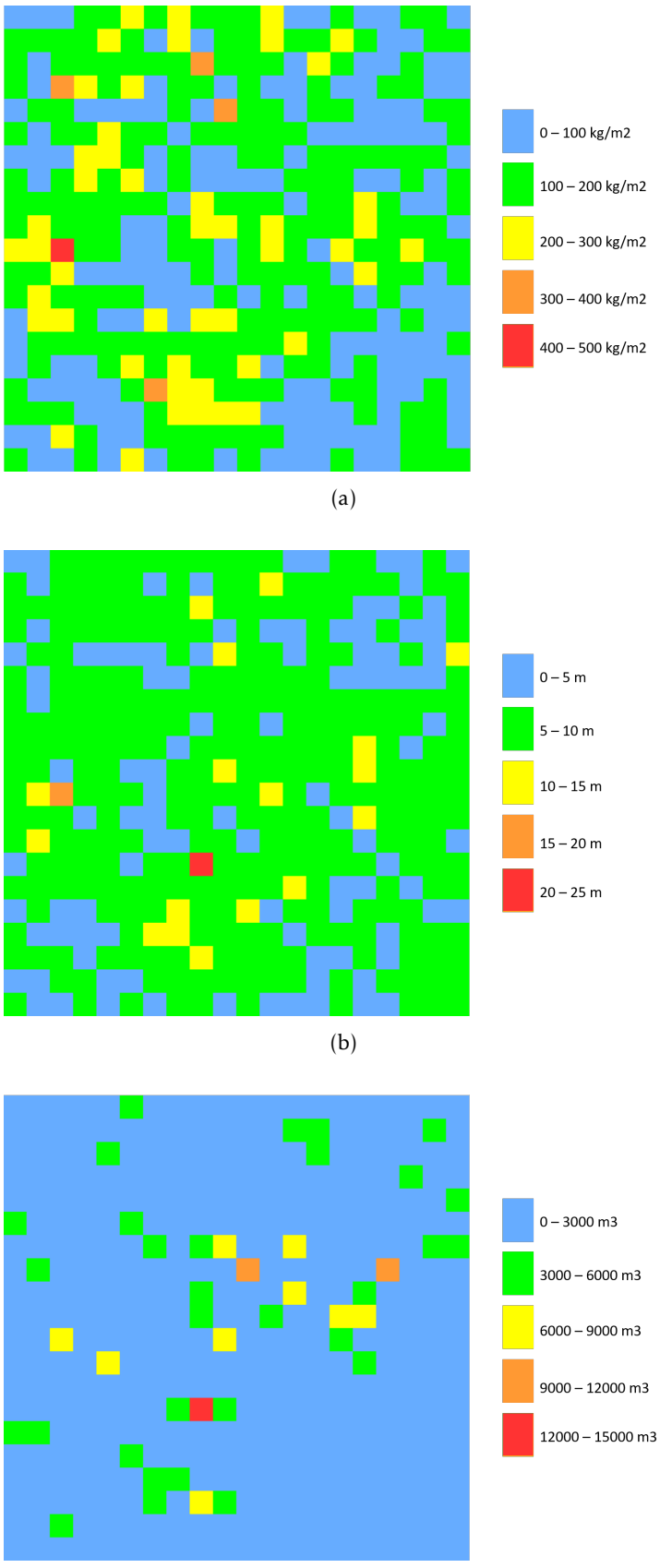


Figure 4.13: 50 by 50 metre maps of the data set A. (a) Corresponds to above ground biomass map (b) to an average tree height map and (c) to the average volume map, for each plot.

4.7 Discussion

This study addresses the methods of airborne LiDAR-based remote sensing for above-ground biomass estimation, at plot and individual tree level, in which the detection of individual trees is the primary issue. It is well known that the success of individual tree detection approaches highly depend on many factors, such as the LiDAR processing method, the forest characteristics and the LiDAR acquisition parameters, so, we started by investigating the effects of LiDAR point density on the estimation of biophysical tree parameters, useful for forest inventory and presented the differences in point cloud processing for both types of data. Firstly, we explored the potential of an open source 3D point cloud processing library by using the a variety of filters and operations in order to minimise the estimation errors.

Similarly to on-site techniques, our method assesses above ground biomass at the individual tree level, taking into account the under-story and lower vegetation. By applying the Equations (3.4) and (3.3) we are able to convert, LiDAR forest metrics estimates such as height percentiles, tree density or crown volume into AGB.

Our approach was based on a local-maxima seeded region growing algorithm which works relatively well with both low and high density 3D forest point clouds, and manages to detect, with an high degree of accuracy, most trees in the data sets, despite the variation in tree heights. One problem we encountered while testing the system was it's inability to correctly separate trees with overlapping branches or tree crowns. Due to the nature of the environments, the resulting point clouds are very unorganised as it become increasingly hard to predict the state of the forest and the possible tree location in the subset without some sort of ground truth data. The differentiation of dead and live trees, as can be seen in Figure 4.2 is also something to take into account and can be further studied in order to reduce the estimation errors. Despite that, our tree detection method showed positive results, with all average detection rates greater than 80% and the values for R^2 above 0.85.

Table 4.6: Summary of the results - Data set A.

Succ. Rate	**MAE**	**Mean Height (m)**	**Average AGB (kg/m^2)**	R^2
88%	1.2	14	345	0.92

Table 4.7: Summary of the results - Data set B.

Succ. Rate	**Mean Height (m)**	**Mean DBH (cm)**	**AGB (T)**	R^2
85%	20	67.8	55.5	0.87

From the study of both data sets we can conclude that estimating AGB from height

returns solely is not as reliable as estimating it from DBH. As expected, Figure 4.12 shows that when estimating AGB, DBH is better measure to be taken as it is much more compliant, since the height of a tree doesn't necessarily mean that has enough wood density to make an impact of biomass. Another conclusion is that in order to gather specific measurements, point density poses a significant difference however there is always a trade off, meaning that with higher resolutions, the harder it is for the system to process it, taking longer amounts of time. This results heavily depend, not only on the platform used for the bathymetry but mainly on the type of forest, although, despite the lack of ground truth we were able to get fair results when comparing to other similar studies and provide a suitable system for mapping AGB in a more efficient and rapid manner, although the methods used require adjustments for the different kinds of forest types and data density.

In the end we were able to draw a biomass map representative of a data set of huge proportions, which can be used as a temporal snapshot of the environment as is, and can be later processed in order to detect biomass changes and forest growth evolution after a wildfire.

Chapter 5

Conclusions and Future Work

5.1 Conclusion

This dissertation covers the main factors that make LiDAR bathymetry one of the best remote sensing platforms for collecting surface information more efficiently than the existing remote sensing techniques available in the market. By providing horizontal and vertical information at high resolutions, managing forest structure in wide-scale areas becomes an easier task since it requires the collection of a large amounts of data which are more expensive and time consuming when taking into account the normal methods. With ALS systems, forest attributes such as canopy height, canopy volume and diameter at breast height can be directly measured from the point cloud and used to model AGB, fuel availability and simulate fire behaviour. Here we presented a system capable of performing individual tree detection from 3D LiDAR point clouds with high degree of accuracy.

When it comes to estimating forest metrics with LiDAR technology there are two main methodologies: individual tree detection approaches and area-based approaches. Both approaches, while different in workflow, share the use of empirical models to establish connections between the field-observed metrics and the LiDAR acquired ones[13]. Although there are no specific studies on the application of this technology for soil cover assessment in Portugal, the bibliography consulted allowed to outline the objectives and identify the different study variables.

As described throughout this dissertation, in order to take relevant forest measurements, the airborne LiDAR data has to go through several stages, starting with preprocessing, visual validation, algorithm parametrisation and finally tree detection and segmentation. All of this should be complemented with real on field surveys to corroborate the results, however we couldn't fulfil this objective, so we randomly selected 15

test plots, categorised them by complexity and manually counted the number of trees in order to test the capabilities of the detection software.

To perform a distinction and detect individual trees, we first needed to filter the data sets and classify points as ground and non ground points. After testing other methods we settled on using a progressive morphological filter that uses basic dilation and erosion operation in order to detect objects a remove them from the set. This method despite performing well, can introduce omission errors when generating DTMs.

After the classification process, the data set was cleaned using a statistical outlier removal in order to remove islands, small tree clusters that do not have an impact on the biomass and a few isolated points that can be caused by travelling birds, light polls or cars.

In both data sets we applied a local-maxima seeded region growing algorithm that managed to detect the trees in all of our test plot with an accuracy greater than 80%. From there, by isolating each tree cluster it was possible to analyse and compute the features listed in 3.5. These results were positive and are summarised in Section 4.7.

In order to ease the visualisation of the results, a biomass map of the region is computed. The addition of a tool like this allows direct monitoring at the municipal scale, allowing each municipalities to inventory their spaces and plan the necessary interventions within the scope of exploration, conservation or enjoyment of these areas.

There are still many challenges to be overcome with the use of the LiDAR sensor for forestry applications, especially the improvement and development of methodologies that can enable its application in extensive areas, on the other hand, this results, when comparing to similar studies indicate that this system has room to be further developed and good potential for use in other forested areas. Due to the limitation of real data collected in field, the effectiveness of the system cannot be completely tested, however we believe that the objectives of this dissertation were accomplished and that it can provide help by quantify changes in forest biomass caused by wildfires and evaluate its progression of the years.

5.2 Future Work

There are a few possible areas where the system can be further improved. The first is by collecting in-house data and acquire real ground truth in order to validate the results and giving the system the best possible conditions for a more accurate estimation. This would also help in developing a system based on machine learning that could help parametrise the filters used. If not possible, a comparison between this system and all the available LiDAR processing tools would be interesting to see. The ability of predicting the tree species would be of great value as well, since it is important for forest inventory and management.

Further development in the detection of overlapping trees needs to be made, as it is increasingly difficult to detect individual trees in a highly complex and highly condensed

forest.

An obvious development would be the introduction of a multi-spectral camera to allow the system to differentiate between live and dead trees according to their reflectance values. This would be a great improvement since it would become feasible to study and map the actual fuel present in a given area and predict the behaviour of a wildfire.

Another interesting development would be the design of an interface capable of connecting a fire simulator program, such as FlamMap, in order to visualise and pin point the possible causes and effects of a wildfire in a 3D LiDAR point cloud. A further study on the effects of a wildfire and the intrinsic biomass changes that occur can also be of value as well as an in depth study of the economic impact of a fire regarding the commercial wood available in a given area.

Bibliography

[1] A. E. Akay, H. Oğuz, I. R. Karas, and K. Aruga. "Using LiDAR technology in forestry activities." In: *Environmental Monitoring and Assessment* 151.1-4 (2009), pp. 117–125. ISSN: 01676369. DOI: 10.1007/s10661-008-0254-1.

[2] A. Barbati, G. Chirici, P. Corona, A. Montaghi, and D. Travaglini. "Area-based assessment of forest standing volume by field measurements and airborne laser scanner data." In: *International Journal of Remote Sensing* 30.19 (2009), pp. 5177–5194. ISSN: 13665901. DOI: 10.1080/01431160903023017.

[3] T. Brandtberg. "Classifying individual tree species under leaf-off and leaf-on conditions using airborne lidar." In: *ISPRS Journal of Photogrammetry and Remote Sensing* 61.5 (2007), pp. 325–340. ISSN: 09242716. DOI: 10.1016/j.isprsjprs.2006.10.006.

[4] A. C. Carrilho, M Galo, and R. C. Dos Santos. "STATISTICAL OUTLIER DETECTION METHOD FOR AIRBORNE LIDAR DATA." In: (2018). DOI: 10.5194/isprs-archives-XLII-1-87-2018. URL: https://doi.org/10.5194/isprs-archives-XLII-1-87-2018.

[5] J. Chave, M. Réjou-Méchain, A. Búrquez, E. Chidumayo, M. S. Colgan, W. B. Delitti, A. Duque, T. Eid, P. M. Fearnside, R. C. Goodman, M. Henry, A. Martínez-Yrízar, W. A. Mugasha, H. C. Muller-Landau, M. Mencuccini, B. W. Nelson, A. Ngomanda, E. M. Nogueira, E. Ortiz-Malavassi, R. Pélissier, P. Ploton, C. M. Ryan, J. G. Saldarriaga, and G. Vieilledent. "Improved allometric models to estimate the aboveground biomass of tropical trees." In: *Global Change Biology* 20.10 (2014), pp. 3177–3190. ISSN: 13652486. DOI: 10.1111/gcb.12629.

[6] P. Corona and L. Fattorini. "Area-based lidar-assisted estimation of forest standing volume." In: *Canadian Journal of Forest Research* 38.11 (2008), pp. 2911–2916. ISSN: 00455067. DOI: 10.1139/X08-122.

[7] M. Court-Picon, C. Gadbin-Henry, F. Guibal, and M. Roux. "Dendrometry and morphometry of Pinus pinea L. in Lower Provence (France): Adaptability and variability of provenances." In: *Forest Ecology and Management* 194.1-3 (2004), pp. 319–333. ISSN: 03781127. DOI: 10.1016/j.foreco.2004.02.024.

[8] FAO. "International Handbook on Forest Fire Protection - Technical guide for the countries of the Mediterranean basin." In: (2009), pp. 1–163. URL: http://www.fao.org/forestry/27221-06293a5348df37bc8b14e24472df64810.pdf.

[9] M. A. Fischler and R. C. Bolles. "RANSAC1981.pdf." In: *Graphics and Image Processing* 24.6 (1981), pp. 381–395. ISSN: 00010782.

[10] D. Gatziolis, S. Popescu, R. Sheridan, and N. W. Ku. "Evaluation of terrestrial LiDAR technology for the development of local tree volume equations Demetrios Gatziolis." In: *Processing* September (2010), pp. 197–205.

[11] P. J. Gibson, W. Contributions, T. O. The, T. From, and C. H. Power. *Introductory remote sensing: principles and concepts*. Vol. 38. 07. 2001, pp. 38–3925–38–3925. ISBN: 9780415170246. DOI: 10.5860/choice.38-3925.

[12] D. M. Hawkins. *Identification of Outliers*. 1980. ISBN: 9789401539944. DOI: 10.1007/978-94-015-3994-4.

[13] R. Hayashi, A. Weiskittel, and J. A. Kershaw. "Influence of Prediction Cell Size on LiDAR-Derived Area-Based Estimates of Total Volume in Mixed-Species and Multicohort Forests in Northeastern North America." In: *Canadian Journal of Remote Sensing* 42.5 (2016), pp. 473–488. ISSN: 17127971. DOI: 10.1080/07038992.2016.1229597.

[14] R. Hill, S. A. Hinsley, and Richard K. Broughton. "Forestry Applications of Airborne Laser Scanning: Chapter 17 Assessing habitats and organism-habitat relationships by ALS." In: *Springer* 27.January (2014), pp. 63–88. ISSN: 1568-1319. DOI: 10.1007/978-94-017-8663-8. URL: http://link.springer.com/content/pdf/10.1007/978-94-017-8663-8.pdf{\%}5Cnhttp://link.springer.com/10.1007/978-94-017-8663-8.

[15] T. S. T. R. Institute. *3D forest*. URL: https://www.3dforest.eu/.

[16] V. R. Kane, M. P. North, J. A. Lutz, D. J. Churchill, S. L. Roberts, D. F. Smith, R. J. McGaughey, J. T. Kane, and M. L. Brooks. "Assessing fire effects on forest spatial structure using a fusion of landsat and airborne LiDAR data in Yosemite national park." In: *Remote Sensing of Environment* 151 (2014), pp. 89–101. ISSN: 00344257. DOI: 10.1016/j.rse.2013.07.041. URL: http://dx.doi.org/10.1016/j.rse.2013.07.041.

[17] J. W. Karl, J. V. Yelich, M. J. Ellison, and D. Lauritzen. "Estimates of Willow (Salix Spp.) Canopy Volume using Unmanned Aerial Systems." In: *Rangeland Ecology and Management* 73.4 (2020), pp. 531–537. ISSN: 15507424. DOI: 10.1016/j.rama.2020.03.001. URL: https://doi.org/10.1016/j.rama.2020.03.001.

[18] W. Li, Q. Guo, M. K. Jakubowski, and M. Kelly. "A new method for segmenting individual trees from the lidar point cloud." In: *Photogrammetric Engineering and Remote Sensing* 78.1 (2012), pp. 75–84. ISSN: 00991112. DOI: 10.14358/PERS.78.1.75.

[19] T. M. Lillesand, R. W. Kiefer, and J. W. Chipman. *Remote Sensing and Image Interpretation*. 7. 2015. ISBN: 9788578110796. DOI: 10.1017/CBO9781107415324.004. arXiv: arXiv:1011.1669v3.

[20] T. M. Lillesand, R. W. Kiefer, and J. W. Chipman. *Remote Sensing and Image Interpretation*. 7. 2015. ISBN: 9788578110796. DOI: 10.1017/CBO9781107415324.004. arXiv: arXiv:1011.1669v3.

[21] M. Lowman, S. Devy, and T. Ganesh. "Treetops at risk: Challenges of global canopy ecology and conservation." In: *Treetops at Risk: Challenges of Global Canopy Ecology and Conservation* (2013), pp. 1–444. DOI: 10.1007/978-1-4614-7161-5.

[22] S. Luo, J. M. Chen, C. Wang, X. Xi, H. Zeng, D. Peng, and D. Li. "Effects of LiDAR point density, sampling size and height threshold on estimation accuracy of crop biophysical parameters." In: *Optics Express* 24.11 (2016), p. 11578. ISSN: 1094-4087. DOI: 10.1364/oe.24.011578.

[23] S. Magnussen and P. Boudewyn. "Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators." In: *Canadian Journal of Forest Research* 28.7 (1998), pp. 1016–1031. ISSN: 00455067. DOI: 10.1139/x98-078.

[24] S. Martín-Alcón, L. Coll, M. De Cáceres, L. Guitart, M. Cabré, A. Just, and J. R. González-Olabarría. "Combining aerial LiDAR and multispectral imagery to assess postfire regeneration types in a Mediterranean forest." In: *Canadian Journal of Forest Research* 45.7 (2015), pp. 856–866. ISSN: 12086037. DOI: 10.1139/cjfr-2014-0430.

[25] J. E. Means, S. A. Acker, B. J. Fitt, M. Renslow, L. Emerson, and C. J. Hendrix. "Predicting forest stand characteristics with airborne scanning lidar." In: *Photogrammetric Engineering and Remote Sensing* 66.11 (2000), pp. 1367–1371. ISSN: 00991112.

[26] M. Mutlu, S. C. Popescu, and K. Zhao. "Sensitivity analysis of fire behavior modeling with LIDAR-derived surface fuel maps." In: *Forest Ecology and Management* 256.3 (2008), pp. 289–294. ISSN: 03781127. DOI: 10.1016/j.foreco.2008.04.014.

[27] N. N. E. O. Network). *Working with LiDAR forest data*. URL: https://www.neonscience.org(accessed26Jan2020).

[28] A. Nunes, L. Lourenço, A. Gonçalves, and A. Vieira. "Três décadas de incêndios florestais em Portugal : incidência regional e principais fatores responsáveis." In: *Cadernos de Geografia* 32 (2013), pp. 133–143. ISSN: 0871-1623.

[29] P. Packalen, J. L. Strunk, J. A. Pitkänen, H. Temesgen, and M. Maltamo. "Edge-Tree Correction for Predicting Forest Inventory Attributes Using Area-Based Approach With Airborne Laser Scanning." In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8.3 (2015), pp. 1274–1280. ISSN: 21511535. DOI: 10.1109/JSTARS.2015.2402693.

[30] G. Pajares. "Overview and current status of remote sensing applications based on unmanned aerial vehicles (UAVs)." In: *Photogrammetric Engineering and Remote Sensing* 81.4 (2015), pp. 281–329. ISSN: 00991112. DOI: 10.14358/PERS.81.4.281.

[31] G. Pajares. "Overview and current status of remote sensing applications based on unmanned aerial vehicles (UAVs)." In: *Photogrammetric Engineering and Remote Sensing* 81.4 (2015), pp. 281–329. ISSN: 00991112. DOI: 10.14358/PERS.81.4.281.

[32] C. E. Parrish and R. D. Nowak. "Improved Approach to LIDAR Airport Obstruction Surveying Using Full-Waveform Data." In: *Journal of Surveying Engineering* 135.2 (2009), pp. 72–82. ISSN: 0733-9453. DOI: 10.1061/(asce)0733-9453(2009)135:2(72).

[33] J. Peuhkurinen, L. Mehtätalo, and M. Maltamo. "Comparing individual tree detection and the areabased statistical approach for the retrieval of forest stand characteristics using airborne laser scanning in Scots pine stands." In: *Canadian Journal of Forest Research* 41.3 (2011), pp. 583–598. ISSN: 00455067. DOI: 10.1139/X10-223.

[34] S. C. Popescu and K. Zhao. "A voxel-based lidar method for estimating crown base height for deciduous and pine trees." In: *Remote Sensing of Environment* 112.3 (2008), pp. 767–781. ISSN: 00344257. DOI: 10.1016/j.rse.2007.06.011.

[35] O. F. Price and C. E. Gordon. "The potential for LiDAR technology to map fire fuel hazard over large areas of Australian forest." In: *Journal of Environmental Management* 181 (2016), pp. 663–673. ISSN: 10958630. DOI: 10.1016/j.jenvman.2016.08.042. URL: http://dx.doi.org/10.1016/j.jenvman.2016.08.042.

[36] N. Saarinen, M. Vastaranta, M. Vaaja, E. Lotsari, A. Jaakkola, A. Kukko, H. Kaartinen, M. Holopainen, H. Hyyppä, and P. Alho. "Area-based approach for mapping and monitoring riverine vegetation using mobile laser scanning." In: *Remote Sensing* 5.10 (2013), pp. 5285–5303. ISSN: 20724292. DOI: 10.3390/rs5105285.

[37] L. Y. Sato, V. C. F. Gomes, Y. E. Shimabukuro, M. Keller, E. Arai, M. N. Dos-Santos, I. F. Brown, and L. E. O. e.Cruz de Aragão. "Post-fire changes in forest biomass retrieved by airborne LiDAR in Amazonia." In: *Remote Sensing* 8.10 (2016), pp. 1–15. ISSN: 20724292. DOI: 10.3390/rs8100839.

[38] J. Shan and C. K. Toth. *Topographic Laser Ranging and Scanning: Principles and Processing*. Taylor & Francis, 2018. ISBN: 9781498772273.

[39] S. Solberg, E. Naesset, H. Lange, and O. Bollandsas. "Remote Sensing of Forest Health." In: *International Archives of Photogrammery, Remote Sensing and Spatial Information Sciences* XXXVI - 8/ (). DOI: 10.5772/8283.

[40] C. Torresan, A. Berton, F. Carotenuto, S. F. Di Gennaro, B. Gioli, A. Matese, F. Miglietta, C. Vagnoli, A. Zaldei, and L. Wallace. "Forestry applications of UAVs in Europe: a review." In: *International Journal of Remote Sensing* 38.8-10 (2017), pp. 2427–2447. ISSN: 13665901. DOI: 10.1080/01431161.2016.1252477. URL: http://dx.doi.org/10.1080/01431161.2016.1252477.

[41] A. B. Utkin, A. V. Lavrov, L. Costa, F. Simões, and R. Vilar. "Detection of small forest fires by lidar." In: *Applied Physics B: Lasers and Optics* 74.1 (2002), pp. 77–83. ISSN: 09462171. DOI: 10.1007/s003400100772.

[42] A. B. Utkin, A. Fernandes, F. Simões, A. Lavrov, and R. Vilar. "Feasibility of forest-fire smoke detection using lidar." In: *International Journal of Wildland Fire* 12.2 (2003), pp. 159–166. ISSN: 10498001. DOI: 10.1071/WF02048.

[43] A. B. Utkin, F. Piedade, V. Beixiga, P. Mota, and P. Lousã. "Scalable lidar technique for fire detection." In: *Second International Conference on Applications of Optics and Photonics* 9286 (2014), p. 92860D. ISSN: 1996756X. DOI: 10.1117/12.2060254.

[44] M. Vastaranta, M. Holopainen, X. Yu, R. Haapanen, T. Melkas, J. Hyyppä, and H. Hyyppä. "Individual tree detection and area-based approach in retrieval of forest inventory characteristics from low-pulse airborne laser scanning data." In: *Photogrammetric Journal of Finland* 22.2 (2011), pp. 1–13.

[45] M. Vastaranta, T. Kantola, P. Lyytikäinen-Saarenmaa, M. Holopainen, V. Kankare, M. A. Wulder, J. Hyyppä, and H. Hyyppä. "Area-based mapping of defoliation of scots pine stands using airborne scanning LiDAR." In: *Remote Sensing* 5.3 (2013), pp. 1220–1234. ISSN: 20724292. DOI: 10.3390/rs5031220.

[46] N. Verma, D. Lamb, N. Reid, and B. Wilson. "Comparison of Canopy Volume Measurements of Scattered Eucalypt Farm Trees Derived from High Spatial Resolution Imagery and LiDAR." In: *Remote Sensing* 8 (May 2016), p. 388. DOI: 10.3390/rs8050388.

[47] W. Wagner, M. Hollaus, C. Briese, and V. Ducic. "3D vegetation mapping using small-footprint full-waveform airborne laser scanners." In: *International Journal of Remote Sensing* 29.5 (2008), pp. 1433–1452. ISSN: 13665901. DOI: 10.1080/01431160701736398.

[48] R. White and B. Dietterick. "Use of LiDAR and multispectral imagery to determine conifer mortality and burn severity following the lockheed fire." In: ... *Imagery To Determine Conifer Mortality* ... (2012), pp. 667–675. URL: http://wwwtest.fs.fed.us/psw/publications/documents/psw{_}gtr238/psw{_}gtr238{_}667.pdf.

[49] W. Yao, P. Krzystek, and M. Heurich. "Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data." In: *Remote Sensing of Environment* 123 (2012), pp. 368–380. ISSN: 00344257. DOI: 10.1016/j.rse.2012.03.027. URL: http://dx.doi.org/10.1016/j.rse.2012.03.027.

[50] K. Zhang, S. C. Chen, D. Whitman, M. L. Shyu, J. Yan, and C. Zhang. "A progressive morphological filter for removing nonground measurements from airborne LIDAR data." In: *IEEE Transactions on Geoscience and Remote Sensing* 41.4 PART I (2003), pp. 872–882. ISSN: 01962892. DOI: 10.1109/TGRS.2003.810682.