Doctoral Dissertation
Doctoral Program in Civil Engineering (36th Cycle)

Innovative LiDAR-based solution for automatic assessment of forest parameters for estimating aboveground biomass and CO_2 storage

Claudio Spadavecchia

* * * * * *

Supervisors
Prof. Francesco Laio, Supervisor
Prof. Marco Piras, Co-Supervisor

Doctoral Examination Committee:
Dr. Frédéric Berger, Referee, Institut national de recherche en sciences et technologies pour l'environnement et l'agriculture
Prof. Harri Kaartinen, Referee, Finnish Geospatial Research Institute at the National Land Survey of Finland
Prof. Domenica Costantino, Politecnico di Bari
Prof. Vincenzo Di Pietra, Politecnico di Torino
Prof. Emanuele Lingua, Università degli Studi di Padova

Politecnico di Torino
July 2024

This thesis is licensed under a Creative Commons License, Attribution-Noncommercial - NoDerivative Works 4.0 International: see www.creativecommons.org. The text may be reproduced for non-commercial purposes, provided that credit is given to the original author.

I hereby declare that the contents and organization of this dissertation constitute my own original work and do not compromise in any way the rights of third parties, including those relating to the security of personal data.



Claudio Spadavecchia
Turin, July 2024

Summary

Over the past few decades, there has been a significant increase in the occurrence of natural hazards worldwide, including floods, droughts, fires, and landslides, impacting both urban and natural environments. The pivotal catalyst propelling climate change is the escalating concentration of greenhouse gases (GHG) in the Earth's atmosphere, predominantly characterized by the surge in carbon dioxide (CO2) levels. The intensification of atmospheric GHGs, attributed mainly to human activities such as industrial processes, deforestation, and the burning of fossil fuels, acts as a potent force amplifying the natural greenhouse effect and leading to global warming.

Monitoring natural resources is essential in guiding political decisions toward optimal practices for preserving these resources. This need is particularly critical as natural resources play a primary role in absorbing carbon dioxide; specifically, forests act as fundamental natural carbon reservoirs, absorbing carbon dioxide from the atmosphere during the process of photosynthesis. However, forest monitoring activities are traditionally time and resource-consuming. Nevertheless, nowadays, the scientific community can rely on the most recent technological innovations, which allow to acquire information about increasingly large areas with a greater level of detail and less time. Remote Sensing techniques need to be supported by automatic processing methodologies in order to manage the amount of data produced efficiently; different approaches have been developed depending on the data type.

This research focuses on the acquisition and processing of LiDAR point clouds, which represent a well-established acquisition technique in the field of Remote Sensing and provide detailed mapping of the scene. More specifically, the aim is to propose a methodology for extracting

forest parameters through LiDAR techniques and automated procedures. The final goal is to assess woody aboveground biomass and the carbon dioxide stocked in it according to different LiDAR acquisition techniques and evaluate temporal variations caused by climate changes or the natural growth of various tree species. The methodology is tested for three case studies: the first scenario concerns a multi-temporal analysis of aerial datasets before and after an ice storm occurred in the Dinaric Forest (Southwest Slovenia); the second scenario consists of a comparison of several terrestrial sensors in a post-forest fire scenario in the Italian Western Alps; the last case study delves into the processing of oblique point cloud datasets through a multitemporal approach under undisturbed conditions in the Hyytiälä boreal forest in Southern Finland.

Acknowledgement

I want to acknowledge the Geomatics Research Group of the Department of Environment, Land and Infrastructure Engineering (DIATI) at Politecnico di Torino, where I had the privilege of working daily. A particular thanks goes to Professor Marco Piras for allowing me to develop this research and expand my scientific knowledge.

Also, I have been privileged to visit the Remote Sensing and Photogrammetry at the Finnish Geospatial Research Institute (FGI) in 2022. I want to thank Professor Antero Kukko for the opportunity to collaborate with his research group and the research group leader, Eetu Puttonen, for allowing me access to their LiDAR dataset in the Hyytiälä forest. A special thanks goes to my friends in FGI.

Contents

List of Tables XI
List of Figures XV
List of Abbreviations XXIII
INTRODUCTION 1
1.1 Keywords 5
OVERVIEW OF CLIMATE CHANGE 7
2.1 What do we know about climate change? 10
2.2 How can we deal with climate change? 17
2.2.1 Monitoring of CO_2 storage capacity 20
2.3 The role of geomatics against climate change 22
TECHNIQUES AND SENSORS FOR FOREST MONITORING 25
3.1 Direct measurements 26
3.2 Remote Sensing 28
3.2.1 Passive and active sensors 32
Passive sensors 32
Active sensors 36
LiDAR 37
3.2.2 Platforms 44
Aerial 45
Terrestrial 49
Data fusion 51
3.2.3 Methods for georeferencing 53

Global Navigation Satellite System 54
Inertial Navigation System 58

LiDAR DATA PROCESSING, STATE OF THE ART 61

4.1 Data collection 62
4.2 Point cloud pre-processing 63
4.3 Point cloud segmentation 69

4.3.1 Segmentation in the forestry domain 77

4.4 Tree volume and aboveground biomass characterization 81
4.5 Carbon stock assessment 84
4.6 Results validation: methodologies and measures 84

CASE STUDIES AND MAIN WORKFLOW 91

5.1 Case study A: Slovenian Dinaric forest 91
5.2 Case study B: Italian Mompantero forest 94
5.3 Case study C: Finnish Hyytiälä forest 99
5.4 Processing workflow 104

5.4.1 Data collection and pre-processing 106
5.4.2 Individual Tree Detection 107
5.4.3 Multi-temporal detection of individual trees 110
5.4.4 Wood and leaf separation 111
5.4.5 Volume and AGB characterization and stocked CO_2 assessment 113
5.4.6 Case-specific tasks 114

CASE STUDY A: RESULTS AND DISCUSSION 117

6.1 General overview 118
6.2 Data processing 121

6.2.1 ITD 121

6.2.2 Wood and leaf separation 125
6.2.3 Error propagation in the allometric equations 128
6.2.4 Forest parameters (DBH and AGB) characterization 130
6.2.5 CO_2 assessment..... 135
6.3 Lesson learnt 137
CASE STUDY B: RESULTS AND DISCUSSION 141
7.1 General overview 142
7.2 Data processing 144
7.2.1 ITD..... 144
7.2.2 Wood and leaf separation 150
7.2.3 AGB characterization 153
7.2.4 CO_2 assessment..... 158
7.3 Lesson learnt 159
CASE STUDY C: RESULTS AND DISCUSSION 163
8.1 General overview 164
8.2 Data processing 166
8.2.1 ITD..... 166
8.2.2 Wood and leaf separation 172
8.2.3 Forest parameters (DBH and AGB) characterization 176
8.2.4 CO_2 assessment..... 181
8.3 Lesson learnt 183
CONCLUSIONS AND FUTURE VISION 187
REFERENCES..... 195
Appendix A 231
Appendix B 233
Appendix C 241
Appendix D 245

List of Tables

Table 1: Main characteristics of passive sensors. 36
Table 2: Main characteristics of LiDAR sensors. 44
Table 3: Case study A - Trees parameters summary. 92
Table 4: Case study B - Trees parameters summary. 99
Table 5: Case study C - Trees parameters summary. 101
Table 6: Summary of the proposed approaches for identifying tree woody and foliar parts. 111
Table 7: Features for identifying the woody and foliar parts of the trees in each case study and for the proposed approaches. 111
Table 8: Summary of the characteristics of the case studies analyzed. 115
Table 9: Parameters set for each case study. 115
Table 10: Summary of the biomass and DBH estimation and validation procedure - Case study A. 120
Table 11: Metrics of the segmentation procedure, PyCrown algorithm - Case study A. 123
Table 12: Metrics of the multitemporal detection of individual trees procedure - Case study A. 124
Table 13: Estimated and reference values of forest parameters - 2013 scenario. 131
Table 14: Statistical indexes for forest parameters - 2013 scenario. 132
Table 15: Estimated and reference values of forest parameters - 2014 scenario. 134
Table 16: Statistical indexes for forest parameters - 2014 scenario. 135

Table 17: Estimate of biomass loss following the disastrous event - Case study A. 135
Table 18: Estimated and reference values of stocked CO_2 - Case study A. 136
Table 19: Statistical indexes for stocked CO_2 - Case study A. 137
Table 20: Summary of the biomass estimation and validation procedure - Case study B. 143
Table 21: Metrics of the segmentation procedure, PyCrown algorithm - Case study B. 147
Table 22: Metrics of the segmentation procedure, FSCT algorithm - Case study B. 149
Table 23: Estimated and reference values of forest parameters - MLS point cloud. 155
Table 24: Statistical indexes for forest parameters - MLS scenario. 156
Table 25: Estimated and reference values of forest parameters - TLS scenario. 157
Table 26: Statistical indexes for forest parameters - TLS scenario. 157
Table 27: Estimated and reference values of stocked CO_2 - Case study B. 158
Table 28: Statistical indexes for stocked CO_2 - Case study B. 159
Table 29: Summary of the biomass estimation and validation procedure - Case study C. 166
Table 30: Metrics of the segmentation procedure, PyCrown algorithm - Case study C. 168
Table 31: Metrics of the segmentation procedure, FSCT algorithm - Case study C. 170
Table 32: Metrics of the multitemporal detection of individual trees procedure - Case study C. 171

Table 33: Estimated and reference values of forest parameters - 2020 point cloud. 178
Table 34: Statistical indexes for forest parameters - 2020 scenario. 179
Table 35: Estimated and reference values of forest parameters - 2021 scenario. 180
Table 36: Statistical indexes for forest parameters - 2021 scenario. 181
Table 37: Estimated and reference values of stocked CO_2 - Case study C. 182
Table 38: Statistical indexes for stocked CO_2 - Case study C. 183
Table 39: Summary of the results obtained in this study according to the case study and the task considered. ITD and multitemporal match were evaluated according to F1 score values; wood and leaf separation quality according to visual interpretation (* = awful result; ***** = excellent result). 191
Table 40: Estimated and reference values of stocked CO_2 - 2013 scenario. (Case study A). 241
Table 41: Statistical indexes for stocked CO_2 - 2013 scenario. (Case study A). 242
Table 42: Estimated and reference values of stocked CO_2 - 2014 scenario. (Case study A). 243
Table 43: Statistical indexes for stocked CO_2 - 2014 scenario. (Case study A). 243
Table 44: Estimated and reference values of stocked CO_2 - 2020 scenario. (Case study C). 245
Table 45: Statistical indexes for stocked CO_2 - 2020 scenario. (Case study C). 246
Table 46: Estimated and reference values of stocked CO_2 - 2021 scenario. (Case study C) 246

Table 47: Statistical indexes for stocked CO_2 - 2021 scenario. (Case study C). 246

List of Figures

Figure 1: Carbon dioxide concentration [ppm] in November 2019 [8]. 10
Figure 2: Temperature anomaly [°C] in 2022 compared to the 1951-1980 average [9]. 11
Figure 3: Global atmospheric CO_2 concentration [NOAA]. 12
Figure 4: Global causes of deforestation 2000-2018 [17]. 14
Figure 5: Regional differences in deforestation drivers (2000-2018) [17]. 14
Figure 6: Global land and ocean temperature anomalies from 1880 to 2022 [19]. In blue, anomalies below the 20th-century average of 13.9 °C; in red, anomalies above the 20th-century average. 15
Figure 7: Global surface temperature change registered and predicted according to different emission level scenarios [31]. 18
Figure 8: Schematization of a tree's subdivision of the aboveground and belowground biomass [46]. 22
Figure 9: Examples of diameter tape (on the left), dendrometric caliper (in the middle) and hypsometer (on the right) and their functioning. 26
Figure 10: Basic operating principle of remote sensing [63]. 30
Figure 11: Schematization of the electromagnetic spectrum. 33
Figure 12: Different kinds of resolution, with examples of lower and higher resolution data. Spatial resolution relates to pixel size, temporal resolution to observation frequency, radiometric resolution to the number of unique values, and spectral resolution to bandwidth in the electromagnetic spectrum [78]. 34

Figure 13: Scheme of the GSD in function of the factors of flight height, focal length and geometric resolution of the sensor. Modified from [79]. 35
Figure 14: Examples of LiDAR sensors. 37
Figure 15: Airborne laser scanning timeline [81]. 38
Figure 16: Example of a triangulation-based laser scanner. 40
Figure 17: Scanning mechanisms of LiDAR sensors and acquisition pattern of a moving sensor – From left to right: oscillating mirror, rotating mirror, palmer scanner, Risley prism. Modified from [83]. ... 42
Figure 18: Examples of a Sentinel satellite (on the left) and an aircraft for aerial survey [110] (on the right). 46
Figure 19: Examples of UAV (multirotor on the left and the right, fixed wing in the middle). 47
Figure 20: Schematization of the swath width related to flight altitude and sensor scan angle parameter [111]. 48
Figure 21: Examples of a Terrestrial Laser Scanning data acquisition (on the left) and a handheld mobile Laser Scanning System (on the right). 50
Figure 22: Examples of Hybrid RGB/LiDAR systems. 52
Figure 23: Example of data fusion: LiDAR and photogrammetric point cloud obtained through the integration of aerial and terrestrial platforms [142]. 53
Figure 24: GNSS scheme. 55
Figure 25: Scheme of pre-processing procedures of LiDAR point clouds. 64
Figure 26: Example of point cloud registration of two point clouds acquired in an urban environment [159]. 64
Figure 27: Examples of retroreflective targets (on the left) and edges of regular human-made objects (on the right, in the red circle) to be distinctively detected for the alignment procedure. 66
Figure 28: Georeferencing of ALS clouds [160]. 67

Figure 29: Scheme of segmentation approaches. 70
Figure 30: Example of single tree level segmentation at plot scale using the PyCrown algorithm [221]. 78
Figure 31: Example of single tree level segmentation at plot scale using the FSCT algorithm [222]. 79
Figure 32: Example of single tree level segmentation at plot scale using the Treeseg algorithm [229]. 80
Figure 33: Example of reconstructed tree structure [252]. 83
Figure 34: Case study A - Study area. EPSG: 3912 [268]. 93
Figure 35: Case study A - Point cloud (a) pre-ice storm acquisition (November 2013); (b) post-ice storm acquisition (April 2014). Color scale according to elevation. EPSG: 3912. 94
Figure 36: Case study B - Study area. EPSG: 32632 [271]. 95
Figure 37: Case study B – KAARTA Stencil 2 trajectory and point cloud. EPSG: 32632. 97
Figure 38: Case study B - Point cloud (a) MMS (KAARTA Stencil 2); (b) TLS (RIEGL VZ-400i). Color scale according to elevation. EPSG: 32632. 97
Figure 39: Case study B - RIEGL VZ-400i point cloud. EPSG: 32632. 98
Figure 40: Case study C - Study area. EPSG: 3067 [272]. 100
Figure 41: Case study C - LiPhe setup [273]. 101
Figure 42: Case study C - Point cloud (a) April 2020; (b) April 2021. Color scale according to elevation. EPSG: 3067. 102
Figure 43: Workflow of this study. 105
Figure 44: PyCrown main workflow summary: in green is the original workflow, and in light blue is the improvement proposed in this study. 108
Figure 45: Specific details about the procedure implemented in the PyCrown workflow. 109
Figure 46: FSCT complete workflow [222]. 110

Figure 47: Ground points of the study area A. EPSG:3912. 118
Figure 48: Error resulted during the ITD procedure - Case Study A. 121
Figure 49: Segmented point cloud at single tree level - (a) 2013 and (b) 2014 point cloud. EPSG:3912. 122
Figure 50: Treetops and crowns detection. The 2013 scenario is colored in red, and the 2014 scenario is colored in blue. EPSG:3912. 122
Figure 51: Segmented point cloud at single tree level with multitemporal match - (a) 2013 and (b) 2014 point cloud. The same trees are colored with the same colors; blue trees did not have a positive match. EPSG:3912. 124
Figure 52: Wood and leaf separation on the same tree: on the left, first approach; in the middle, second approach; on the right, third approach - 2013 scenario. 125
Figure 53: Wood and leaf separation on the same tree: on the left, first approach; in the middle, second approach; on the right, third approach. (a) leaf component; (b) wood component - 2013 scenario. 126
Figure 54: Wood and leaf separation on the same tree: on the left, first approach; in the middle, second approach; on the right, third approach - 2014 scenario. 127
Figure 55: Wood and leaf separation on the same tree: on the left, first approach; in the middle, second approach; on the right, third approach. (a) leaf component; (b) wood component - 2014 scenario. 127
Figure 56: Error propagation of Equation (16) on the left, and Equation (15) on the right for the AGB assessment. 129
Figure 57: Error propagation of Equation (19) for the DBH assessment. 130
Figure 58: Ground points of the study area B: (a) TLS point cloud; (b) MLS point cloud. EPSG:32632. 142

Figure 59: Accuracy of MLS point cloud compared with TLS dataset. On the left is the entire study area; on the right is a portion of individual trees [271]. 144
Figure 60: Segmented point cloud at single tree level with PyCrown algorithm - (a) MLS point cloud and (b) TLS point cloud. EPSG:32632. 145
Figure 61: Treetops and crowns detected with PyCrown algorithm. MLS scenario is colored in red; TLS scenario is colored in blue. EPSG:32632. 146
Figure 62: LiDAR acquisition of a single tree, (a) TLS point cloud; (b) MLS point cloud. 147
Figure 63: Segmented point cloud at single tree level with FSCT algorithm - (a) MLS point cloud and (b) TLS point cloud. EPSG:32632. 148
Figure 64: Treetops detected with FSCT algorithm. The MLS scenario is plotted in red circles; the TLS scenario is plotted in blue rhombus. EPSG:32632. 149
Figure 65: Wood and leaf separation on the same tree: on the left, first approach; in the middle, second approach; on the right, third approach - MLS point cloud. 150
Figure 66: Wood and leaf separation on the same tree: on the left, first approach; in the middle, second approach; on the right, third approach. (a) leaf component; (b) wood component - MLS point cloud. 151
Figure 67: Wood and leaf separation on the same tree: on the left, first approach; in the middle, second approach; on the right, third approach - TLS point cloud. 152
Figure 68: Wood and leaf separation on the same tree: on the left, first approach; in the middle, second approach; on the right, third approach. (a) leaf component; (b) wood component - TLS point cloud. 153

Figure 69: Error resulted during the QSM procedure - Case study B. 155
Figure 70: Ground points of the study area C: (a) 2020 point cloud; (b) 2021 point cloud. EPSG:3067. 165
Figure 71: Segmented point cloud at single tree level with PyCrown algorithm - (a) 2020 point cloud and (b) 2021 point cloud. EPSG:3067. 167
Figure 72: Treetops and crowns were detected with the PyCrown algorithm. The 2013 scenario is colored in red, and the 2021 scenario is colored in blue. EPSG:3067. 167
Figure 73: Segmented point cloud at single tree level with FSCT algorithm - (a) 2020 point cloud and (b) 2021 point cloud. EPSG:3067. 169
Figure 74: Treetops detected with FSCT algorithm. The 2013 scenario colored in red, and the 2021 scenario colored in blue. EPSG:3067. 169
Figure 75: Segmented point cloud at single tree level with multitemporal match. The same trees are colored with the same colors; blue trees did not have a positive match. EPSG:3067..... 171
Figure 76: Individual silver birch trees selected for wood and leaf separation procedure. EPSG:3067..... 173
Figure 77: Wood and leaf separation on the same tree: (I) first approach, without deviation; (II) first approach, with deviation; (III) second approach, without deviation; (IV) second approach, with deviation; (V) third approach - 2020 point cloud..... 174
Figure 78: Wood and leaf separation on the same tree: (I) first approach, without deviation; (II) first approach, with deviation; (III) second approach, without deviation; (IV) second approach, with deviation; (V) third approach. (a) leaf component; (b) wood component - 2020 point cloud. 174

Figure 79: Wood and leaf separation on the same tree: (I) first approach, without deviation; (II) first approach, with deviation; (III) second approach, without deviation; (IV) second approach, with deviation; (V) third approach - 2021 point cloud. 175
Figure 80: Wood and leaf separation on the same tree: (I) first approach, without deviation; (II) first approach, with deviation; (III) second approach, without deviation; (IV) second approach, with deviation; (V) third approach. (a) leaf component; (b) wood component - 2021 point cloud. 176

List of Abbreviations

1D	1-Dimensional
2D	2-Dimensional
3D	3-Dimensional
AGB	Aboveground biomass
ALS	Aerial Laser Scanning
AI	Artificial Intelligence
BIM	Building Information Modeling
CC	Climate Change
CD	Crown Diameter
CH_4	Methane
CHM	Canopy Height Model
CMIP5	Coupled Model Intercomparison Project Phase 5
CSF	Cloth Simulation Filter
CO_2	Carbon dioxide
DACCS	Direct Air Carbon Capture and Storage
DBH	Diameter at Breast Height
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DL	Deep learning
DSM	Digital Surface Model

DT	Decision Tree
DTM	Digital Terrain Model
ECEF	Earth-Centered, Earth-Fixed
EPA	Environmental Protection Agency
EPSG	European Petroleum Survey Group
EO	Earth Observation
EU	European Union
FAO	Food and Agriculture Organization of the United States
FGI	Finnish Geospatial Research Institute
FN	False Negative
FP	False Positive
FRA	Global Forest Resources Assessment
FSCT	Forest Structural Complexity Tool
GCM	General Circulation Models
GCP	Ground Control Points
GHG	Greenhouse gas
GIS	Geographic Information System
GLONASS	GLObal NAvigation Satellite System
GMM	Gaussian Mixture Model
GPS	Global Positioning System
GNSS	Global Navigation Satellite System
GSD	Ground Sampling Distance
HDBSCAN	Hierarchical DBSCAN

ICP	Iterative Closest Point
IMU	Inertial Measurement Unit
INS	Inertial Navigation System
IPCC	Intergovernmental Panel on Climate Change
IRNSS	Indian Regional Navigational Satellite System
ITD	Individual Tree Detection
***k*-NN**	*k*-Nearest Neighbors
LiDAR	Light Detection and Ranging
LULC	Land Use Land Cover
MEO	Medium Earth Orbit
ML	Machine Learning
MLS	Mobile Laser Scanning
MMS	Mobile Mapping System
MODIS	Moderate-resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
NET	Negative Emission Technologies
NIR	Near infraRed
N_2O	Nitrous oxide
NOAA	National Oceanic and Atmospheric Administration
P	Precision
PCA	Principal Component Analysis
PRN	Pseudo Random Noise
QSM	Quantitative Structure Model

QZSS	Quasi-Zenit Satellite System
RADAR	RAdio Detection And Ranging
RANSAC	RANdom SAmple Consensus
RBG	Red Blue Green
R	Recall
RF	Random Forest
RS	Remote Sensing
RT	Real Time
RTK	Real Time Kinematic
SAR	Synthetic Aperture Radar
SfM	Structure from Motion
SLAM	Simultaneous Localization and Mapping
SVM	Support Vector Machine
SW	Swath width
SWIR	Short-Wave InfraRed
TLS	Terrestrial Laser Scanner
ToF	Time of Flight
TP	True Positive
TS	Total Station
UAS	Uncrewed Aerial Systems
UAV	Uncrewed Aerial Vehicles
UNFCCC	United Nations Framework Convention on Climate Change
WMO	World Meteorological Organization

Chapter 1

INTRODUCTION

Nowadays, public opinion strongly discusses climate change and the actions to be taken to tackle it. Attention is raised due to the increasingly frequent news about disastrous flood events, earthquakes, extensive fires, prolonged periods of drought, the breaking of record high temperatures, etc. Consequences of these events have both short- and long-term impact: in the short term, human life may be endangered from the occurrence of the event; in the long term, critical issues may occur with the water and food supply, as well as there may be, by way of example but not limited to, transport, energy or ecosystem crises [1]. The goal of the scientific community is to raise awareness among the population and policymakers to encourage more environmentally friendly individual actions and policies and promote the latest scientific developments and innovative strategies for monitoring and analyzing climate change and adaptation techniques.

The leading cause of climate change is the increase in the concentration of greenhouse gases (GHG) in the atmosphere, mainly

carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), and fluorinated gases [2]. Although the greenhouse effect is a natural phenomenon fundamental to regulating the Earth's temperature and guaranteeing livable conditions, excessive concentrations of these gases disturb the natural balance of the atmosphere. This imbalance causes an increase in the frequency and intensity of the disastrous natural events previously mentioned. Various factors cause the rise of greenhouse gases in the atmosphere. Among these are deforestation, agriculture, and livestock; however, the most incisive factor is the combustion of fossil fuels, which humanity exploits to produce energy.

Various protocols and laws have been issued at national, European and international levels since the 1970s to reduce greenhouse gas emissions. In addition to man's greater attention to the Earth's health, the Earth implements natural strategies to deal with climate change. In particular, forests and oceans are the two largest CO_2 storage reservoirs: forests exploit the chlorophyll photosynthesis of plants and trees to fix carbon, producing biomass; oceans, on the other hand, exploit the photosynthesis of phytoplankton, which is then deposited on the seabed. Forests cover about 31% of the emerged lands [2] and can absorb about a third of the carbon dioxide emitted by burning fossil fuels each year [3]. However, due to deforestation caused by natural events and incorrect forest management by man, this capacity is shrinking; moreover, because of deforestation, new CO_2 is emitted into the atmosphere during the biomass decomposition phase. These reasons make it essential to improve the practices and strategies for sustainable and responsible use to be adopted in forest management. However, good resource management requires in-depth monitoring, which can allow us to quantify the carbon absorption capacity of forests more easily, evaluate their temporal variability, determine the mutual relationship with climate change, etc. Only in this way is it possible to know the phenomenon in

its entirety, which makes it possible to provide political decision-makers with all the information necessary to adopt the best strategies.

Monitoring the CO_2 emitted by the combustion of fossil fuels does not present particular challenges; it can be carried out with different strategies, the most precise of which is based on the use of special sensors for direct measurements; on the other hand, monitoring the CO_2 absorbed by trees with sufficient accuracy is an ambitious goal. As a matter of fact, not only is it necessary to quantify the biomass of forests, but it is also necessary to identify the relationship between the volume of trees and the carbon dioxide actually stored in the wood.

Traditionally, the calculation of the volume of a tree is carried out invasively or indirectly; in the first case, the tree is cut down, and its weight is measured; in the second case, some indirect parameters are measured, such as the Diameter at Breast Height (DBH) and the treetop height, which are related to the aboveground biomass (AGB) through allometric equations previously calibrated on control data. The advent of more advanced technologies based on remote sensing represents the new frontier of forest monitoring. Nowadays, it is possible to use techniques based on photogrammetry or LiDAR (Light Detection and Ranging) sensors to acquire the information necessary to estimate the volume of a tree with higher precision and, therefore, evaluate the stocked carbon quantity.

The ease of acquisition that the new instruments have brought makes using Artificial Intelligence (AI) essential for data processing. The amount of data that can be acquired differs from the capacity a man can put in place. Machine Learning (ML) and Deep Learning (DL) algorithms are currently the most powerful tools for analyzing large amounts of data, learning from past experiences, and providing autonomous predictions.

In the forestry literature, numerous studies exploit this data processing procedure [4], [5]; however, at present, there is no

standardized procedure to describe the automated forest monitoring process in its entirety, which comprehensively explains the data acquisition and processing procedure with the final goal of obtaining an estimate of the carbon stocked over time in a forest at the single tree level.

The study presented in this thesis aims to fill the scientific gap by proposing a semi-automated workflow for estimating forest parameters (i.e. aboveground biomass, tree elevation and diameter of the trunk) and assessing the carbon captured and stocked within the woody biomass at a single tree level starting from LiDAR data acquired through different techniques. Three datasets with different characteristics were taken into consideration (in terms of forest characteristics and acquisition method), and specific research questions were also investigated for each of them. Case study A and C address the issue of multitemporality, investigating the variation in biomass both as a consequence of the occurrence of disastrous events (case study A, with ALS data) and in undisturbed conditions (case study C, with oblique acquisitions); additionally, the dataset of case study A was also used to validate an improved procedure for segmentation at the level of a single tree, while that of case study C allowed to evaluate the most influenced features for the segmentation of wood and leaf points of the tree. Finally, case study B focuses more on testing different acquisition techniques (MLS and TLS) and defining which of these allows to correctly identify the greatest number of trees and calculate their woody biomass.

This research is contextualized within the *climate_change@polito* (*cc@polito*) project at the Department of Environment, Land and Infrastructure Engineering (DIATI), which aims to analyze and face the effects of climate change developing (i) innovative strategies for climate change monitoring and analysis, (ii) climate change mitigation techniques and (iii) adaptation solutions.

The thesis is structured as follows. In Chapter 2, the concept of climate change is introduced and deeply discussed; in particular, causes, effects and possible strategies to deal with it are highlighted. Moreover, a focus on forest monitoring and the aboveground biomass assessment for carbon stocking estimations is presented. Chapter 3 delves into sensors and equipment required in the forest monitoring field, describing the pros and cons of different acquisition techniques; greater emphasis will be placed on LiDAR functioning. In Chapter 4, a literature review of point cloud processing techniques and tools (i) for individual tree detection (ITD), (ii) aboveground biomass assessment and (iii) stocked carbon estimations are described. Chapter 5 introduces three case studies analyzed in this thesis by describing the datasets and the main workflow. Chapters 6, 7, and 8 describe the experiments conducted and discuss the results. Finally, in Chapter 9, conclusions are summarized, and future research topics are introduced.

1.1 Keywords

Climate Change (CC); Forest monitoring; LiDAR; Point clouds; Stocked CO_2 assessment; Individual Tree Detection (ITD).

Chapter 2

OVERVIEW OF CLIMATE CHANGE

According to the definition given by the Intergovernmental Panel on Climate Change (IPCC) [6], climate change (CC) refers to:

> *"a change in the state of the climate that can be identified (e.g., by using statistical tests) by changes in the mean and/or the variability of its properties and that persists for an extended period, typically decades or longer. Climate change may be due to natural internal processes or external forcings such as modulations of the solar cycles, volcanic eruptions and persistent anthropogenic changes in the composition of the atmosphere or land use."*

At the same time, the United Nations Framework Convention on Climate Change (UNFCC) [7] defines climate change in Article 2:

> *"Climate change means a change of climate which is attributed directly or indirectly to human activity that alters the composition of the global atmosphere and which is in addition to natural climate variability observed over comparable periods."*

The two definitions of climate change are complementary to each other. IPCC emphasises the physical meaning of climate change and how to determine it rigorously without dwelling too much on the description of the causes; on the contrary, the UNFCCC's definition emphasises the causes, particularly humanity's role. If, through the first definition, it could be deduced that natural and anthropogenic causes are equally influential and determining, the second firmly states that the leading causes are directly or indirectly attributable to human activity, which are also added causes linked to natural climatic variability. The coexistence of different definitions for the same phenomenon highlights how various aspects can be stressed depending on the perspective from which it is viewed (e.g., political, scientific, economic). From a scientific point of view, which is the approach with which this thesis research deals, it is essential to rigorously define climate change, identify its causes, and analyse its effects to accurately determine the cause-effect relationship and propose solutions to protect our climate.

As stated, the changes are attested by looking at the variation of statistical parameters relating to climatic properties over a sufficiently long period. More information about which are the parameters to consider is given in the definition of climate provided by the IPCC [6]:

> *"Climate in a narrow sense is usually defined as the average weather, or more rigorously, as the statistical description in terms of the mean and variability of relevant quantities over a period ranging from months to thousands or millions of years. The classical period for averaging these variables is 30 years, as*

> *defined by the World Meteorological Organization. The relevant quantities are most often surface variables such as temperature, precipitation and wind. Climate in a wider sense is the state, including a statistical description, of the climate system."*

This definition also emphasizes the importance of considering the proper time interval when discussing climate change. It is, therefore, necessary not to confuse climate change and weather. The latter is defined by the World Meteorological Organization (WMO) as natural events happening in short time intervals and specific places and times:

> *"Weather describes short-term natural events - such as fog, rain, snow, blizzards, wind and thunderstorms, tropical cyclones, etc. - in a specific place and time."*

Even the concept of global warming, often mistakenly interchanged with that of climate change, must be clear. The National Aeronautics and Space Administration (NASA) refers to global warming as:

> *"The long-term heating of Earth's surface observed since the pre-industrial period (between 1850 and 1900) due to human activities, primarily fossil fuel burning, which increases heat-trapping greenhouse gas levels in Earth's atmosphere. This term is not interchangeable with the term "climate change."*

Climate change refers to all the effects due to changes in the Earth's climate, while global warming relates only to the increase in the average temperature on Earth.

In the following paragraphs, a greater focus will be given to the causes of climate change and how human activities have affected it. Afterwards, the consequences of climate change will be explored in

depth. Finally, the knowledge about the actions undertaken (or to be undertaken) to mitigate and adapt to climate change will be deepened.

2.1 What do we know about climate change?

The Earth's climate is constantly changing over time. Indeed, the main factor responsible for climate change is the greenhouse effect. This natural phenomenon consists of taking advantage of greenhouse gases in the Earth's atmosphere to retain the heat of the Sun and keep the Earth's temperature stable. However, the increase of these gases in the atmosphere over time caused an exaggerated overheating worldwide, commonly referred to as global warming. Among the greenhouse gases that worsen global warming, the most effective is carbon dioxide (CO_2) due to the highest atmospheric concentrations. Methane (CH_4) is the second cause because, despite having a greater capacity to absorb solar energy, unlike CO_2, it lasts much less in the atmosphere; in much smaller concentrations, nitrous oxide (N_2O) and fluorinated gases are also present.

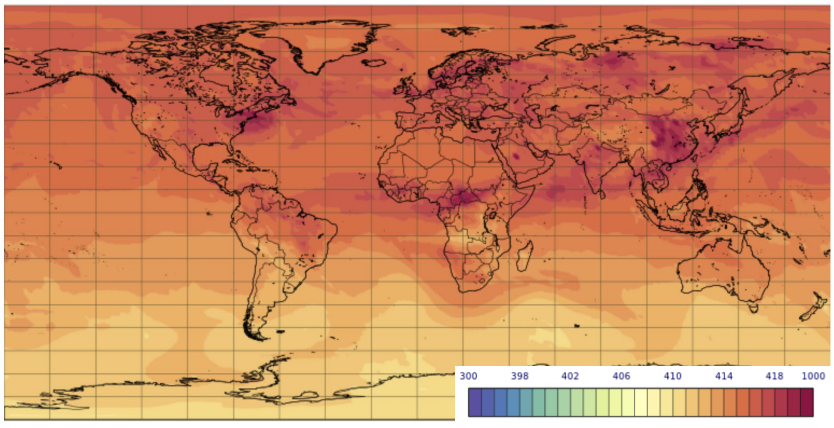


Figure 1: Carbon dioxide concentration [ppm] in November 2019 [8].

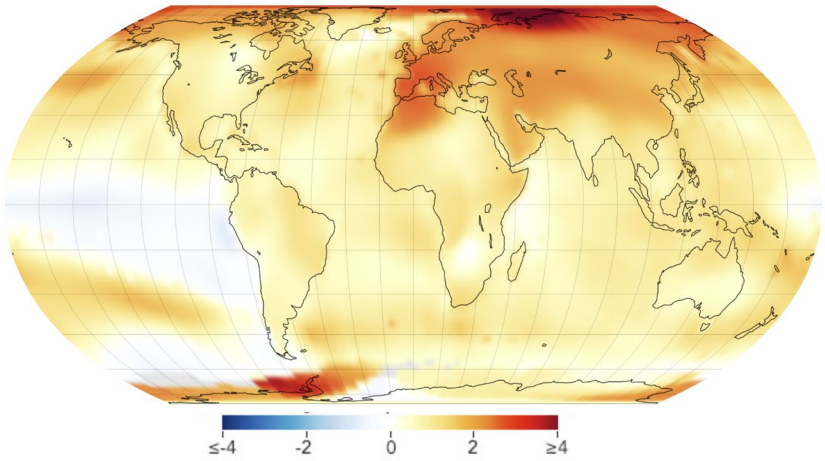


Figure 2: Temperature anomaly [°C] in 2022 compared to the 1951-1980 average [9].

The natural causes that can provoke the increase in the concentration of greenhouse gases are mainly (i) volcanic eruptions, (ii) anaerobic decomposition of organic matter, which releases methane, (iii) changes in the Earth's orbit and rotation, (iv) variations in Solar activity. Volcanic activities emit sulfur dioxide (SO2) and carbon dioxide into the atmosphere; nevertheless, conservative estimates of annual CO_2 emissions from volcanoes are less than 1% of CO_2 emissions from all anthropogenic activities [10]. Changes in Earth's orbit alter the way the Sun's energy is distributed with latitude and by season on Earth; these orbital changes were minimal over the last several hundred years and are insufficient to cause the observed magnitude of change in temperature [11]. The intensity of the sunlight that reaches the Earth's surface depends on solar energy. Satellite measurements have recorded minor variations in the solar energy emitted that can cause slight variations in the Earth's temperature but have not registered an increasing temperature trend that fully justifies long-term global warming [11].

Nevertheless, since the 19th century, the impact of human activities on climate has increased mainly due to the increase in the use of fossil fuels (i.e., coal, oil, and natural gas), which emit GHG into the atmosphere when burned. Since the beginning of the Industrial Revolution, the atmospheric CO_2 concentration has increased by more than 40%, with over half the increase occurring since 1970 [11]. Data published by the National Oceanic and Atmospheric Administration [12] from 1979 to the present shows an increase in the global atmospheric CO_2 concentration of about 80 parts per million (ppm), and nowadays, it settles around 420 ppm (Figure 3).

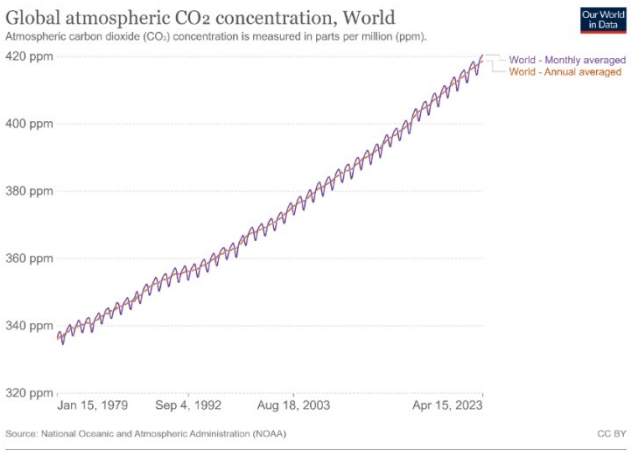


Figure 3: Global atmospheric CO_2 concentration [NOAA].

According to the United States Environmental Protection Agency (EPA), which annually develops an annual Inventory of U.S. Greenhouse Gas Emissions and Sinks [13], the sectors that cause the highest greenhouse gas emissions in the United States are transportation (28%), energy (25%) and industry (23%); the commercial and residential sectors and the agricultural sector have a lower impact (13% and 10% respectively). Regarding agriculture, specifically intensive farming, the

quantities of methane produced by animals through the decomposition of animal waste are far from negligible [14].

Furthermore, human activities like deforestation seriously endanger biodiversity and the Earth's natural capacity to absorb carbon. According to the Food and Agriculture Organization of the United States (FAO)'s Global Forest Resources Assessment [15], deforestation refers to:

> *"The conversion of forest to other land use independently whether human-induced or not."*

Forests (like oceans and soils), in the context of climate change, are commonly identified as natural carbon sinks, defined as "*any process, activity or mechanism which removes a greenhouse gas, an aerosol or a precursor of a greenhouse gas from the atmosphere*" [16]. Forests cover about 31% of the emerged lands [17]. However, the thesis aims to estimate the carbon dioxide absorbed by trees precisely. Rough studies at a global level have shown that forests absorb approximately two gigatons of carbon [18] and about a third of the carbon dioxide emitted by burning fossil fuels each year [3]. Deforestation seriously reduces the capacity of forests to stock carbon; its causes can be natural (e.g., floods, fires, storms, etc.) or anthropic. It is estimated [17] that almost 90% of deforestation worldwide is due to agricultural expansion; more specifically, 50% is linked to converting forests into cultivated land, and 40% is for livestock grazing. Urbanization affects deforestation only 6% (even if it is the primary driver in Europe, as shown in Figure 5), and the remaining 4% refers to the exploitation of wood resources (Figure 4). Between 1990 and 2020, 420 million hectares of forest were lost to deforestation (an area equivalent to that of the EU), mainly in South America, Congo and Southeast Asia. Furthermore, factors of natural origin also partially give rise to deforestation.

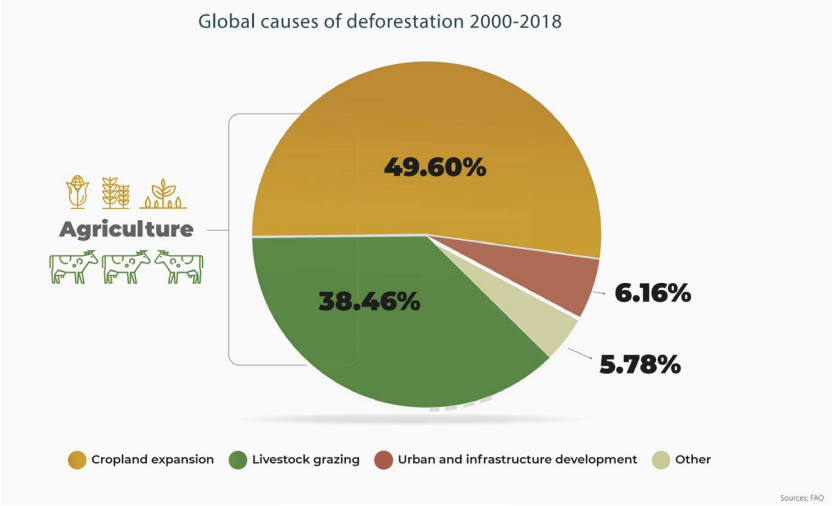


Figure 4: Global causes of deforestation 2000-2018 [17].

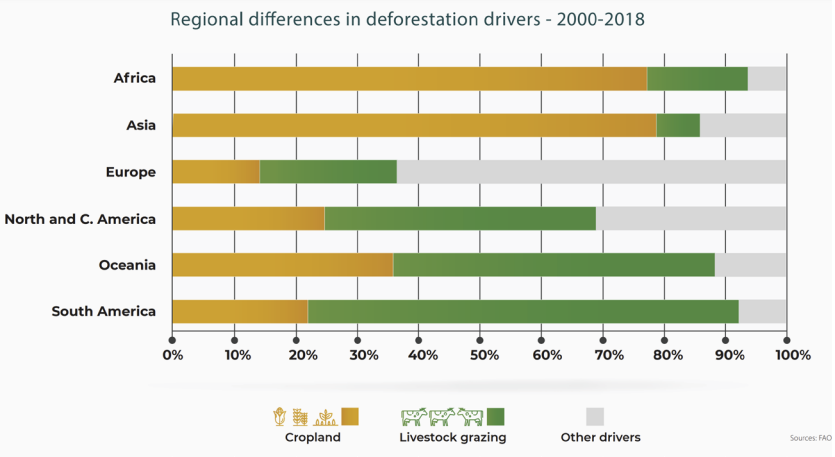


Figure 5: Regional differences in deforestation drivers (2000-2018) [17].

The main consequence of climate change is the increase of the average air and ocean temperature near the Earth's surface, mainly observed in the past one or two centuries. Time series about Earth's global average surface temperature collected by the National Oceanic

and Atmospheric Administration (NOAA) [17] clearly show a positive trend. Currently, the average annual temperature is about one °C higher than the average temperatures recorded in the twentieth century (Figure 6). While the 20th-century temperatures average settles at 13.9 °C, in 2016, the record of 14.9 °C was reached.

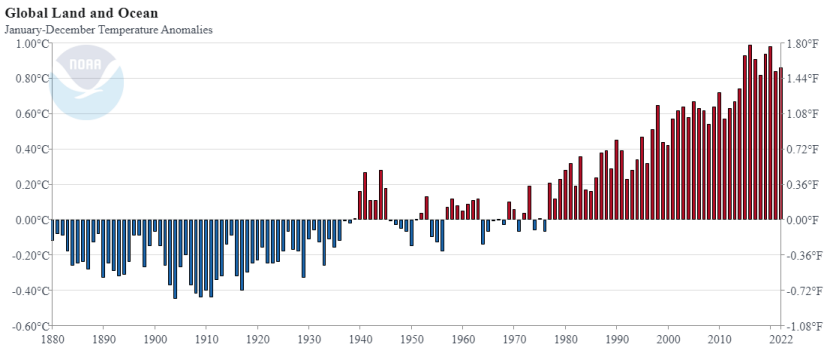


Figure 6: Global land and ocean temperature anomalies from 1880 to 2022 [19]. In blue, anomalies below the 20th-century average of 13.9 °C; in red, anomalies above the 20th-century average.

High temperatures and the increasing occurrence of heat waves have manifold consequences. Directly, high temperatures expose the world population to an increased risk of mortality: dehydration phenomena or heat strokes can be fatal; previous diseases, such as cardiac or respiratory diseases, can be more severe and recurring due to high temperatures [1]. Moreover, increasingly warmer climates facilitate the proliferation and spread of diseases carried by pathogens [20]; the increase in concentrations of pollutants in the air can cause inflammation of the airways and increase the susceptibility to infections, in particular in subjects suffering from respiratory diseases (e.g., asthma, bronchitis) [21]. Higher temperatures are also expected to cause changes in meteorological models and the occurrence of extreme events such as

torrential rains, prolonged droughts, wildfires, landslides, etc. It has been observed in recent years that the frequency and intensity of these events have increased [22]. Flood events or landslides can damage infrastructure and endanger the lives of millions of people, especially in Europe, where the high population density in areas with high seismic or flood risk can be a crucial factor [23], [24]. Wildfires break out and spread more efficiently due to high temperatures, putting biodiversity and human health at risk and favouring the release of CO_2 absorbed by plants and trees into the atmosphere [25].

Moreover, the reduction of wooded and vegetated areas causes a reduced capacity to absorb new carbon dioxide naturally. Also, surface water bodies absorb much of the heat from global warming; as a result, sea level rises, threatening coastal and island communities, affecting the stability of marine ecosystems, and modifying and reducing the availability of water resources water supply [26]. The increase of greenhouse gases in the atmosphere also affects oceans, which absorb them (it is estimated that the oceans today absorb around 25% of the CO_2 produced by human activities, corresponding to around two petagrams of carbon per year [27]). This phenomenon causes water acidification, endangering marine life [28]. Global warming is affecting the melting of glaciers, which conditions mountain and polar ecosystems and water supply in regions where glaciers are used as freshwater natural reservoirs [29]. Another consequence worth mentioning is the variation of the geographical distribution of climatic zones, which can (i) alter the distribution and abundance of plant and animal species, (ii) damage agriculture and livestock yields, and (iii) put in danger the freshwater availability [30].

2.2 How can we deal with climate change?

Tackling climate change is essential to preserve resources and global security for current generations and ensure good living conditions for future generations.

Unfortunately, nowadays, it is not possible to make deterministic, definitive predictions of how climate will evolve over the next century and beyond as it is with short-term weather forecasts [31]. The forces at play are disparate, and the current understanding of the climate system does not allow us to fully understand how these interact with each other and the environment; furthermore, the very magnitude of the forces at play is still being determined, both anthropic and natural. Nonetheless, the problem must be addressed first by raising awareness among the world population. At the same time, the best scientific, technological, and political solutions must be implemented. A lack of interest in tackling climate change would worsen global warming and its consequences, making it even more challenging to guarantee optimal living conditions for the entire world population.

Despite the uncertainty of the forces at play mentioned above and the difficulties in describing them with mathematical language, several studies have proposed climate models for modelling climate conditions and predicting future conditions [32]. Climate models, or General Circulation Models (GCMs), are based on well-documented physical processes to simulate the transfer of energy and materials through the climate system. Mathematical equations characterize how energy and matter interact in different parts of the ocean, atmosphere, and land. Once the model can correctly reproduce past climate conditions, it is used to predict future scenarios. These are achieved by varying complex parameters, which attempt to describe equally complex phenomena such as demographic trends, economic development, political choices, etc. [33].

In literature, several studies have been conducted on climate models [34], [35], [36], [37], [38]. Over time, improvements in resolution, the number of simulated processes, and better representations of the latter have been achieved [31]. Figure 7 shows the global mean temperature changes averaged across all Coupled Model Intercomparison Project Phase 5 (CMIP5) models relative to 1986–2005 for four different scenarios. They differ according to four predicted emission levels. According to this study, based on different possible energy policies and economic growth patterns, in 2100, the global surface temperature could increase between 1 and 4 °C.

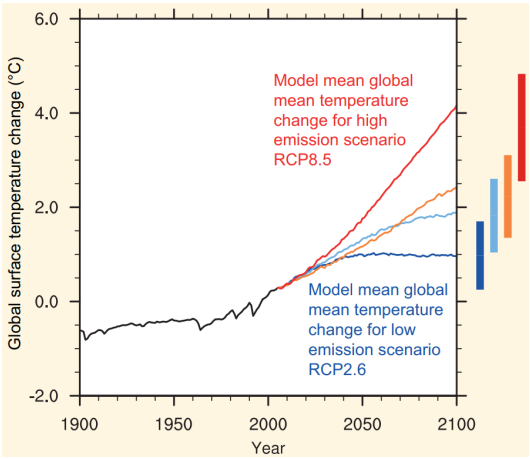


Figure 7: Global surface temperature change registered and predicted according to different emission level scenarios [31].

The gravity of the possible future scenarios has led decision-makers to act at international levels. The first international treaty that commits state parties to reduce greenhouse gas emissions is the Kyoto Protocol, ratified in 1997 and entered into force in 2005 [39]. More recently, studies have highlighted the need to limit the global temperature to 1.5°C above pre-industrial temperature to avert the worst impacts of climate

change and preserve a livable planet; this is the main objective set in the Paris Agreement, stipulated in 2015, between the members of the UNFCC [40]. To cope with global warming, we must mainly try to reduce greenhouse gas emissions, favouring (i) the transition towards renewable and low-carbon energy sources, (i) improving energy efficiency, and (iii) reducing consumption, especially at an agricultural and industrial level.

The Earth has complex natural mechanisms that can influence the climate to achieve balance; the carbon sinks are the most effective way carbon is naturally stocked in oceans, forests and soil [41]. The oceans absorb about 25% of human carbon dioxide emissions. They can produce about 50% of the oxygen through the diffusion of carbon dioxide in the ocean's surface water and the oceanic photosynthesis of phytoplankton. However, this phenomenon can affect the balance of the marine ecosystem, as high concentrations of CO_2, which reacts in water to form carbonic acid, lowers the pH of the water (this process is commonly referred to as ocean acidification) [28], [42]. The role of forests has already been introduced in the previous Paragraph. Plants, through the natural process of photosynthesis, exploit solar energy as they grow to convert carbon dioxide and water into glucose and oxygen (O_2); in parallel, plants start the process of cellular respiration, through which they oxidize the sugars produced during photosynthesis to obtain energy. using oxygen and producing carbon dioxide. The net balance calculated as the difference between the CO_2 absorbed and that produced by a tree is generally positive (net CO_2 absorption of approximately 10-30 kg per year depending on the size, type of tree and growing conditions), and the carbon dioxide remains stored for hundreds of years. In addition to directly affecting climate change, they play equally crucial natural solutions in protecting urban environments from landslides, floods, and strong winds.

Finally, soil also stocks carbon in organic matter, which mainly derives from the decomposition of plants and other organic remains (e.g., leaves, branches). The organic matter is decomposed by soil microorganisms (bacteria, fungi) and invertebrates, and a significant portion of CO_2 is incorporated into the soil in the form of stable organic substances such as fulvic acid and humic acid. Furthermore, plants release organic compounds from their roots into the soil.

Furthermore, from a technological point of view, recent developments have been directed towards removing carbon dioxide in the atmosphere through Negative Emission Technologies (NET). The most direct way is the Direct Air Carbon Capture and Storage (DACCS) technique, through which the CO_2 in the air is extracted through chemical processes; other techniques involve afforestation, reforestation forest management, and wood utilization [43], [44], [45]. These approaches are developed as long-term techniques for storing atmospheric CO_2 (carbon sequestration).

2.2.1 Monitoring of CO_2 storage capacity

To know the effects of global warming in their entirety, quantify the damages or changes to the global balance, and lay the foundations for developing efficient mitigation and adaptation techniques, it is necessary to concentrate the first efforts on monitoring the phenomena involved and quantifying the quantities involved. The aspects to be explored in greater depth would be many and varied; however, the phenomenon on which this doctoral thesis focuses is the estimate of CO_2 captured by forests. Monitoring natural resources, particularly forestry resources, and the consequent estimate of the CO_2 they capture is a hotly debated topic. In the previous Paragraph, the importance of this natural resource concerning the increase in greenhouse gas concentration over time has already been described, and rough estimates were presented. However, it is essential to quantify more accurately the carbon stocking capacity of

forests and, in particular, of each individual tree. In this way, it is possible to (i) estimate with a greater degree of precision the benefits that forests perform as carbon sinks, (ii) monitor the validity of reforestation or afforestation practices, and more generally quantify the effects of mitigation policies implemented to combat and limit the effects of climate change, and (iii) quantify the loss of CO_2 storage capacity as a consequence of deforestation or destructive events. Natural hazards, in fact, not only put human lives at risk but can seriously damage natural resources and reduce the storage power of carbon sinks; among these, forests are the most sensitive to external disturbances. As an example in recent history, the Vaia storm hit North-Eastern Italy in October 2018. It caused the fall of millions of trees, destroying tens of thousands of hectares of alpine coniferous forests.

However, carbon dioxide absorbed by plants cannot be directly measured on a large scale; on the contrary, it is easier to obtain it indirectly through measurements and estimates regarding the volume of trees and, therefore, the aboveground biomass. The term "*aboveground biomass*" refers to the total amount of organic matter present in the parts of a plant that are above the ground, such as stems, branches, leaves, and fruits (Figure 8).

As anticipated, this measurement is of particular interest in ecology and forestry, as it provides crucial information on the amount of carbon stored in plants and adds information about the dynamics of the ecosystem. On the contrary, *belowground biomass* mainly refers to the roots (Figure 8). Belowground biomass is not the subject of this study, as the sensors used cannot acquire any information related to it.

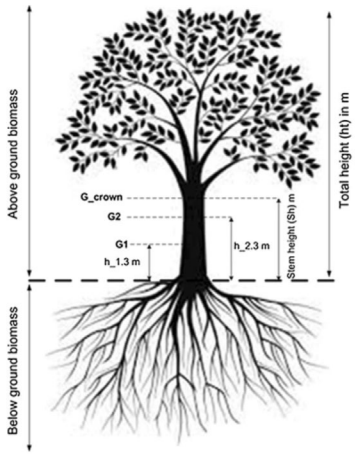


Figure 8: Schematization of a tree's subdivision of the aboveground and belowground biomass [46].

Studies on a global scale quantify the total worldwide biomass at approximately 550 gigatons of carbon (Gt C); of this, about 80% refers to plant biomass, of which in turn, 60% refers to aboveground biomass (320 Gt C), and approximately 24% to belowground biomass (130 Gt C) [47].

2.3 The role of geomatics against climate change

In recent years, the evolution of geomatics technologies and techniques combined with an automatic or semiautomatic data processing procedure provides the best approach to accurately assess the CO2 stocked in the aboveground biomass.

Geomatics focuses on obtaining, representing, understanding, handling, storing, and sharing georeferenced information, referring to data associated with a specific location in a selected reference system. Geospatial data can be acquired using different geomatics techniques, and they help support decision-makers; in particular, remote sensing data are mandatory nowadays in environmental monitoring since they

provide comprehensive knowledge of the Earth's surface. In the field of environmental monitoring and the fight against climate change, it supports the following actions:

- Monitor greenhouse gas emissions and stocking at a global, regional, or local level through specific sensors to support climate models;
- Monitor land use land cover (LULC) to assess the impact on carbon absorption and ecological balance;
- Monitor glaciers and sea levels, analyzing and quantifying the main consequences of global warming;
- Support natural resources management (e.g., forests, waterways, agricultural areas) and spatial planning, ensuring resilience to the effects of climate change and conservation of natural resources and human goods.

As detailed in Paragraph 3.2, remote sensing instruments and sensors can acquire data about the surface of the Earth; as we will see, some instruments, such as LiDAR, have the potential to penetrate through vegetation and reach the ground. However, it is not possible nowadays to acquire below-ground information in the field of remote sensing. As a matter of fact, only the visible part of the trees (trunk, branches and crown) has been analyzed, while the carbon stock capacity of roots and soil was not deepened. Different sensors and different methodologies can be used to monitor forests. For the purposes of this thesis, the focus is on using LiDAR instrumentation, which uses several automatic and semi-automatic data processing techniques. The main steps addressed are (i) the Individual Tree Detection, (ii) the separation between wood and leaf, (iii) the characterization of the aboveground biomass (AGB), and (iv) CO_2 assessment.

Chapter 3

TECHNIQUES AND SENSORS FOR FOREST MONITORING

Forest monitoring is an essential procedure encompassing all practices of systematically collecting, analysing, and evaluating forest-related information. This process is of fundamental importance for studying forest health, monitoring ecological dynamics, and managing forest resources sustainably, but above all, to cope with environmental challenges and climate change. The main parameters of interest in the forestry domain concern the characteristics of single trees (age, species, height, diameter, volume) as well as forest composition; additionally, from a biological point of view, several samplings can be carried out to perform biodiversity analysis and to individuate possible diseases or parasites. An in-depth analysis of the existing techniques for carrying out forest monitoring is proposed in the following paragraphs: specifically,

in Paragraph 3.2, in situ, traditional approaches are described, while in Paragraph 3.2, the more innovative and indirect methods based on remote sensing data are discussed.

3.1 Direct measurements

Traditional forest monitoring is carried out through forest inventory data collected in situ. Depending on the size of the area under examination, it may decide to carry out a complete inventory, in which all trees above a predetermined diameter are measured, or to apply a sampling strategy. Usually, systematic sampling is more efficient, in which the need for statistical precision determines the number of plots, and their size varies according to the expected number of measurements of the parameters of interest [48].



Figure 9: Examples of diameter tape (on the left), dendrometric caliper (in the middle) and hypsometer (on the right) and their functioning.

The primary information usually collected is related to the location, composition, distribution of the forest resource, and tree density over a given area. This basic information is necessary to outline general assessments at different levels of detail, such as harvest plans and the development of provincial or state-level strategies [49]. However, when analyzing single trees, the crucial information usually collected is their species, the Diameter at Breast Height (DBH), height, age, and health

state. Diameter tapes and dendrometric calipers are the most traditional tools for measuring the DBH, while the tree's height is traditionally determined with a hypsometer (Figure 9). Both diameter tapes and calipers usually have sub-centimeter precision when measuring trees with sufficiently regular circular geometries, while greater uncertainties and difficulties are observed when the trunk form is very irregular. It is specified that the DBH is measured at a height of approximately 1.30 m from the ground.

Ultrasonic or laser hypsometers are frequently employed in field measurement in forest scenarios because of their simplicity. Laser hypsometers rely on trigonometric principles and combine a laser rangefinder and a gravity clinometer; the first is used to measure distances from the point of view to the top and the bottom of objects, while the latter measures the angle between the lines to each. Laser hypsometers are often used when considering longer measures without obstacles between the instrument and the target. This condition, unfortunately, is not always verified, above all in dense forest scenarios [50]. For this reason, ultrasonic hypsometers have been developed: they use ultrasounds to measure distance, which improves hypsometer usability since it can measure the height even if obstacles (e.g., leaves or branches) cover the trajectory from the instrument to the treetop. On the other hand, unfortunately, the transmitting speed can easily be affected by air temperature and moisture content, affecting the accuracy of the measurement. Commercial hypsometers used in forestry reach accuracies of approximately 4 cm at a maximum distance of 1500-2000 meters, depending on the operating principle.

At an operational level, measuring tree heights as a function of forest density is not always possible. Furthermore, the volume and the biomass of a tree are particularly challenging to measure: in fact, the only direct way involves the tree being harvested and subsequently weighed; however, this strategy cannot be carried out on a large scale, as it would

effectively destroy the forests we intend to monitor and protect. These reasons led to the study and development of empirical equations that take into consideration the DBH and other forestry parameters, such as the distribution of trees [51], [52], [53]. Regarding the measurement of aboveground biomass (AGB), allometric equations are calibrated according to tree species considered and based on DBH and height measurements [54], [55], [56], [57], [58]. More details about the AGB assessment will be addressed in Paragraph 4.4, while the estimate of carbon dioxide stocked within the aboveground biomass will be explored in greater detail in Paragraph 4.5.

Direct measurements have the advantage of being very accurate concerning the forestry research domain in which they are used; they can be used to develop allometric equations to estimate more complex attributes without any direct measurement; they can also be stored, visualized and processed in a Geographic Information System. However, these traditional measurement tools are limited in their operability conditions: carrying out measurements in large areas for each tree within a large study area and possibly repeating them over time is highly time-consuming and requires a large workforce. Nowadays, since scientific interest is focused on regions of ever more significant extension and possibly in analyzing the temporal variability with even relatively narrow intervals, direct measurements are more commonly carried out on a sample basis and used mainly to validate the results obtained through remote sensing techniques [59] or to complement them [60].

3.2 Remote Sensing

The science that includes all territorial and environmental survey disciplines using the most recent instrumentation and informatic support is commonly referred to as geomatics. This science involves acquiring, managing, analyzing, and visualizing geographical data; it includes

several disciplines and technologies, including Remote Sensing (RS), which is [61]:

"the science and art of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area, or phenomenon under investigation."

In the modern age of technology, the ability to perceive and understand the world has been revolutionized by remote sensing. Nowadays, progress in the different RS technologies has led to an innovative and efficient approach to large-scale study of climate change.

The field of remote sensing evolved from the interpretation of aerial photographs to the analysis of satellite imagery and from local area studies to global analyses, with advances in sensor system technologies and digital computing. Nowadays, remote sensor systems can provide data from energy emitted, reflected, and/or transmitted from all parts of the electromagnetic spectrum [62]. The fundamental principle behind remote sensing is the interaction between incident radiation and the target of interest. This radiation includes visible light, infrared light, ultraviolet light, microwave light, and radio waves. By interpreting the information in these waves, remote sensing techniques enable insights about the Earth. Specifically, the energy source emits electromagnetic energy, which travels through the atmosphere, hits the target object, and interacts with it (depending on the properties of both the target and the electromagnetic energy). The scattered part travels again in the atmosphere; finally, the sensor collects and records it. The interaction can vary according to the target surface properties and the electromagnetic energy. Finally, according to the kind of sensor and support considered and the purposes for which they are used, data can be transmitted in real-time to an operational base or saved in internal memory and subsequently collected by the operator. A simplifying

schematization of the operating principle of remote sensing is shown in Figure 10.

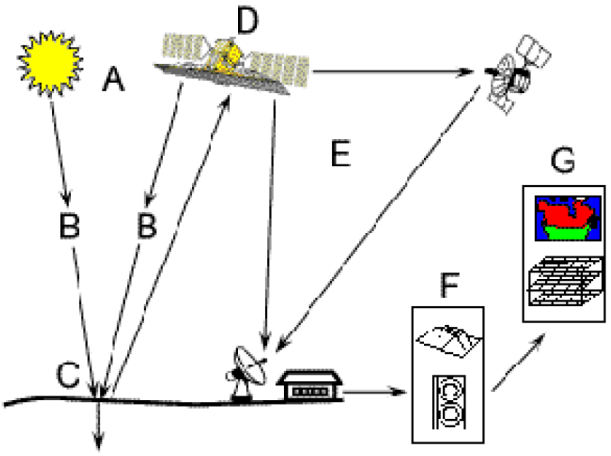


Figure 10: Basic operating principle of remote sensing [63].

The sensors used in the field of remote sensing are disparate. Anyway, the first classification in which they can be subdivided relates to the origin of the energy source. Sensors can be active or passive; in the first case, the sensor itself also produces electromagnetic energy, while in the latter, it relies on the energy of an external nature (e.g., the Sun). A further classification can be made if the distance from the object is considered, affecting the collected data's resolution and extension. The sensors used to acquire information can be equipped on aerial supports: low Earth orbit satellites, helicopters, airplanes, up to Uncrewed Aerial Systems (UAS), which rely on Uncrewed Aerial Vehicles (UAV); furthermore, they can be equipped on terrestrial supports, either mobile (as in the case of handheld sensors or when provided on cars) or fixed (for example mounted on tripods). Sensors and supports must be chosen according to the study's goal, considering the phenomena under examination, the properties requested to be investigated, and the level of detail. A widespread practice also involves using different sensors

simultaneously, integrating all the data collected to provide the most complete and detailed knowledge of the study area.

The forestry research field is one of the sectors in which the potential of remote sensing is fully exploited, allowing the acquisition of a considerable amount of data that would otherwise be impossible or highly expensive, time-consuming, and sometimes even risky. Unlike traditional field-based sampling, full-coverage RS data enables the production of maps of key forestry parameters, which are helpful for forest management purposes [64]. Images captured by satellites offer detailed spatial data with a high temporal revisit, allowing to deepen the variability of forests over time or the consequences of natural disasters immediately after (or meanwhile) these happened safely [65]; LiDAR sensors are usually involved for mapping forest structures with high accuracy with the goal of automatically assess forest attributes needed for forest management by experts in the sector; depending on the applications, the operability conditions, and the extension of the area under consideration, satellite, airborne or terrestrial data can be used. Like satellite-based acquisitions, LiDAR point cloud can be acquired multiple times; data processing helps detect even small changes, making this tool a valuable aid for detailed studies on biomass monitoring.

In the following subparagraph, knowledge about the sensors used in remote sensing will be deepened, emphasising active sensors and particularly LiDAR technology, which was explored and considered in this thesis. Subparagraph 3.2.2 will discuss the platform in more detail, while Subparagraph 3.2.3 will look at the positioning methods for the georeferencing point cloud processing.

3.2.1 Passive and active sensors

Passive sensors

Passive sensors are remote sensing systems that measure the naturally available reflected or remitted energy. Passive sensors rely on the illumination of the Sun since solar radiation is reflected from the surface. Energy can also be naturally emitted (such as thermal infrared), and it can be detected if the amount of energy is large enough to be recorded regardless of the presence of the Sun. Indeed, the power measured by passive sensors depends on several factors, such as the surface composition, physical temperature, surface roughness, and other physical characteristics of the detected object [66].

Among the most common passive sensors are traditional photographic cameras, multispectral imaging cameras, thermal infrared sensors, hyperspectral sensors, and radiometers. Traditional photographic cameras (RGB cameras) are the most common sensors used in remote sensing, mainly for applications like aerial photography and cartography; they can capture visible light, allowing for high-resolution images of the Earth's surface. However, the visible spectrum (wavelength of visible light in air ranges approximately from 400 nm to 700 nm) defines only a small part of the electromagnetic spectrum. Further information can be collected using thermal infrared sensors, which acquire in the Near InfraRed (NIR) portion of the spectrum between 700 nm and 1 mm, or with hyperspectral imaging sensors, which can collect information from across the electromagnetic spectrum (typically from visible to Short-Wave InfraRed (SWIR). Figure 11 shows the electromagnetic spectrum.

Passive sensor data are limited to reconstructing surfaces visible in the image data. For this reason, aerial images in high-density wooded areas provide a partial vision of the territory, lacking a description of the terrain and the lowest part of the trees, limiting themselves to a mere

description of the foliage and the tree crown distribution [67]. Nonetheless, numerous studies have been conducted in estimating forest parameters through aerial or satellite images [68], [69]. In addition, recent developments in photogrammetric processing, such as the SFM (Structure From Motion) technique, allow the reconstruction of a three-dimensional model of the surface of the study area, and it helps in the production of high-resolution orthophotos [70]. Thermal cameras can also be used in forestry, specifically for forest health monitoring [71] or forest fire detection [72]. In contrast, multispectral and hyperspectral cameras are primarily used for species classification and mapping tasks [73], [74], [75], detecting parasites [76], or estimating the aboveground biomass [77].

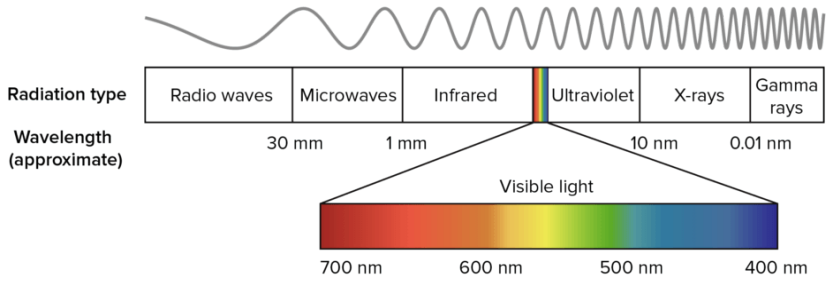


Figure 11: Schematization of the electromagnetic spectrum.

The reliability of passive sensors depends on the quality of the camera, which is understood as lens quality, stability, and resolution. The quality of a camera lens can be identified by several factors, such as its sharpness, clarity, color accuracy, and distortion cons. Image stabilization refers to techniques to reduce blurring associated with the motion of a camera during exposure.

The resolution of an image can be defined from a spatial, temporal, radiometric, and spectral point of view (Figure 12). The spatial resolution refers to the amount of detail that can be seen in that image and is

measured in terms of pixels per unit length; the temporal resolution indicates the frequency with which new images are acquired to represent a temporal sequence or the time elapsed between the acquisition of two consecutive images; the radiometric resolution indicates how many color shades or light intensity levels can be represented in each pixel of the image; finally, the spectral resolution indicates the ability of a sensor or device to capture information across various wavelength bands of light.

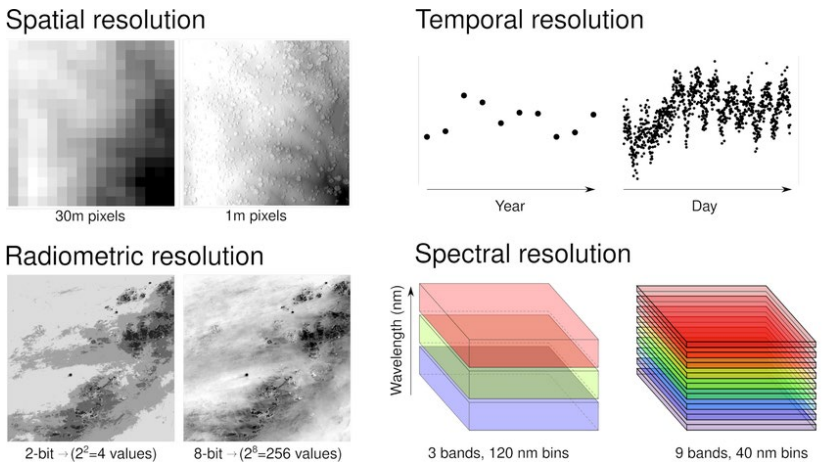


Figure 12: Different kinds of resolution, with examples of lower and higher resolution data. Spatial resolution relates to pixel size, temporal resolution to observation frequency, radiometric resolution to the number of unique values, and spectral resolution to bandwidth in the electromagnetic spectrum [78].

Photogrammetric surveys' relative accuracy (accuracy of individual features on a map concerning each other) increases with increasing overlap between images, with the quality of the pictures, and using Ground Control Points (GCPs). The relative accuracy can be roughly estimated to be between one and three times the Ground Sampling Distance (GSD). GSD is the distance between two consecutive pixel

centres, and it represents the size of the Earth's surface, which is covered by each pixel (schematized in Figure 13). It can be calculated as follows:

$$GSD = \frac{d * H}{h} \quad (1)$$

Where h refers to the focal length, d is the dimension of the sensor, and H is the distance between the sensor and the sensed object.

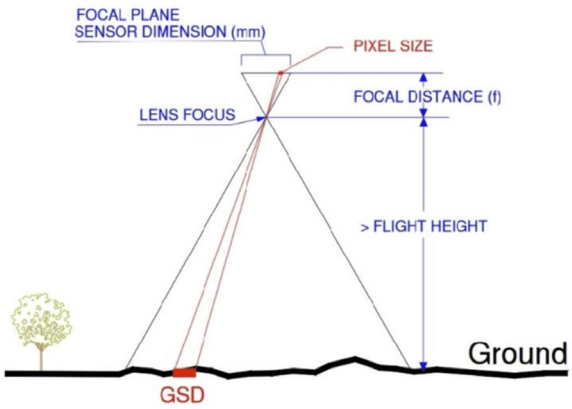


Figure 13: Scheme of the GSD in function of the factors of flight height, focal length and geometric resolution of the sensor. Modified from [79].

The GSD value is a requirement based on which to decide the instrument used and the acquisition distance. Better results are obtained by setting reduced distances from the object at the expense of acquisition times and vice versa.

RGB sensors are widely used because (i) they return high-resolution images, (ii) they are easily and intuitively interpretable, and (ii) they are economical. However, image quality can be affected by atmospheric and lighting conditions, and the chemical and physical properties of the surface cannot be acquired since the acquired electromagnetic spectrum is incomplete. On the other hand, thermal sensors are independent of

lighting conditions since they detect thermal radiation emitted by the surface, but (i) the spatial resolution is lower than RGB sensors, and (ii) they are not very sensitive to minor temperature variations. Finally, hyperspectral sensors are indispensable for studying the chemical composition of the analyzed object. Still, the drawbacks limit their development due to (i) the high cost of production and maintenance and (ii) the complexity of interpreting the data.

Table 1 summarizes the main characteristics of passive sensors.

Table 1: Main characteristics of passive sensors.

Passive sensor	Spectrum	Pros	Cons
RGB sensors	Visible spectrum (400 ~ 700 nm)	Low cost, easy to equip and process	Limited spectral information and light-sensitive
Thermal infrared sensors	NIR spectrum (700 nm ~ 1 mm)	Thermal analyzes and night applications	Limited spatial resolution
Hyperspectral sensors	Full spectrum	Detailed spectral analysis and material characterization	High costs, complex, and advanced data processing

Active sensors

Active sensors, like Radar or LiDAR, emit electromagnetic radiation to illuminate the object; they send a pulse of energy from the sensor to the scene and then receive the radiation reflected or backscattered from that object [11]. The main advantage of these sensors is that they can be used without an energy source and allow night-time acquisitions. Radio detection and ranging (Radar) sensors use electromagnetic waves (radio waves or microwaves) to map the position (fixed or moving) of objects. Since radio waves have a long wavelength, radar sensors are mainly

employed in mapping large objects or tracking movement. For this reason, this technology does not have significant applications in the forestry sector. However, in recent years, several studies have explored the functioning of SAR (Synthetic Aperture Radar) concerning forest studies. SAR is a particular type of Radar that uses the motion of the radar antenna over a target region to provide finer spatial resolution than conventional stationary beam-scanning radars. SAR output can be considered to monitor forest height [80] and evaluate its variation over time as it allows to frequently acquire data in all weather conditions (i.e. cloudy areas) and at all hours of the day. However, the characteristics of the radar waves and their interaction with the surface make the data acquired with the SAR not trivial to interpret; combined with high maintenance costs, such instruments are not as widely used as LiDAR. Furthermore, even if the spatial resolution is high considering the acquisition distance, it is still of the order of magnitude of a meter.

LiDAR

Light Detection and Ranging (LiDAR) sensors, or laser scanners (Figure 14 shows three examples of different LiDAR sensors), work according to the same operating principle as Radar, from which they differ because they rely on a LASER (Light Amplification by Stimulated Emission of Radiation) light to measure the distance of an object.



Figure 14: Examples of LiDAR sensors.

A light beam is produced by exciting a gain medium (a material of controlled purity, size, concentration, and shape); when excited, it emits photons amplified through a system of mirrors, creating a concentrated beam of light. In turn, the beam strikes a mirror with a known laser's angle of incidence. The laser propagates in space until it hits an object, is reflected, and is detected by the sensor, which finally measures the position in space of the detected object. Typically, the wavelength of the laser light is in the ultraviolet, visible or near-infrared spectrum. The output of laser scanner acquisitions is a highly accurate three-dimensional (3D) representation of the scene, commonly referred to as point cloud. As a matter of fact, laser scanners can capture millions of 3D points per second, acquiring not only geometric information but also the target's reflectance features by measuring the laser pulse's return strength.

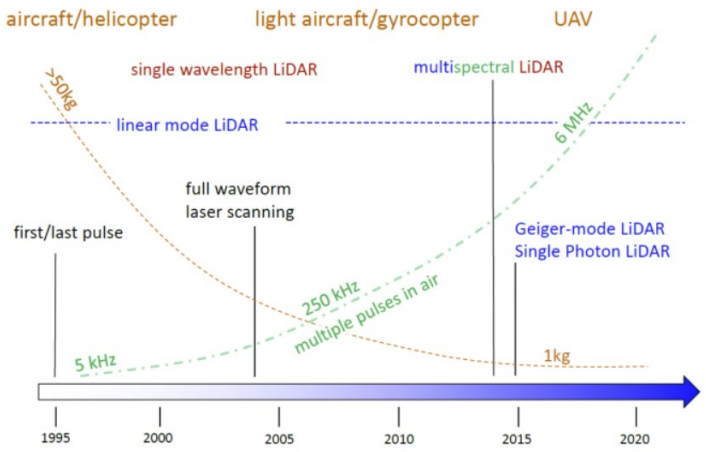


Figure 15: Airborne laser scanning timeline [81].

Figure 15 shows the timeline of the aerial laser scanner [81], which can be used to describe LiDAR generically. Since LiDAR entered the market in the 90s, the overall size and weight (initially in the order of

tens of kilos) have been significantly reduced, reaching weights of around a few kilos; this has allowed, as it will be further discussed, the possibility of equipping these instruments on increasingly lighter supports such as drones, or handheld/pedestrian LiDAR.

Not only has the size been reduced, but at the same time, the scan rate has increased by approximately 1000 times. Nowadays, laser scanners can emit up to 2 million shots per second, with multiple pulses simultaneously travelling in the air.

The first LiDAR sensors could record only the first and last pulses in a discrete way; therefore, for each single laser beam, only two echoes were recorded. Nevertheless, laser pulses can be totally or partially reflected according to the surface properties of the hit object; in some cases, the absorbed part of the laser pulses can penetrate the object, continue its trajectory, and hit multiple objects. In a forest environment, for example, several returns can be recorded relating to different layers of vegetation until the pulse hits an object impenetrable by the laser beam (e.g., the ground, a wall, a trunk). Discrete LiDAR systems identify peaks and record a point (*return*) at each peak location in the waveform curve, while more advanced systems can also record intermediate returns. Differently, full waveform LiDARs record a distribution of returned light energy: full-waveform LiDAR data are thus more complex to process; however, they acquire more information than discrete return LiDAR systems.

Traditional laser scanners operate in linear mode: in this case, they use high energy pulses (resulting in long laser pulses); one laser emitter and one receiver are mounted in the sensor. This approach allows for the acquisition of all the backscattering information (full-waveform). In the recent past, very sensitive LiDAR operate in Geiger Mode or Single Sensitive Photon mode, the advantage of which is a higher area coverage rate; however, due to their high sensitivity, these sensors are more subject to point cloud noise. Geiger mode LiDAR, operating with

medium energy and short laser pulses, is equipped with an array of receivers that only acquire the first echoes. On the other hand, Single Sensitive Photon LiDAR use low energy and short laser pulses; the emitted laser beam is divided into an array of partial beams (beamlets), each with a specific penetration capacity.

Another development that has occurred over time concerns the wavelength affected by the acquisition; from a single wavelength (typically near-infrared for topographical purposes), LiDAR can now acquire data in a multispectral domain, including visible wavelength. This particularity has allowed researcher of different domains to approach LiDAR for their purposes; for example, the green wavelength can penetrate the surface of water bodies, supporting bathymetric studies in shallow or deep water.

The position of the objects hit by the laser emitted by LiDAR is estimated by applying the principles of trigonometry (Triangulation Scanner) or direct measurements (Ranging Scanner).



Figure 16: Example of a triangulation-based laser scanner.

Triangulation scanners (Figure 16) consist of a laser emitter and a receiver separated by a priori known distance (baseline) on which the principle of triangulation is applied. By knowing the angle with which the laser beam is emitted after hitting the internal mirror and measuring the angle at which the receiver captures the laser, it is possible to estimate the target's position. These instruments are widely used in the industrial sector [82] since they have millimeter precision (short range). However,

since the baseline between the emitter and the receiver cannot exceed practical limits of handling, the range of these instruments is very low: as a matter of fact, the precision, which is inversely proportional to the distance to the object, exceeds one decimeter with distances greater than 20 meters.

Ranging Scanners LiDARs are the most commonly used in the geomatics and topographic domains as they achieve centimeter precision even for large ranges (the order of magnitude can vary from a few hundred meters to a few kilometers). In turn, they can be classified according to the kind of direct measurement, which can be related to the time of flight or the phase shift. Phase shift laser scanning systems measure the phase shift between the emitted laser and the same laser detected by the receiver after interacting with the target object. Subsequently, the phase difference is related to the distance between the scanning device and the object's surface according to a pre-defined proportion. Time of Flight (ToF) LiDAR systems determine the distance d to an object by precisely measuring the round-trip time Δt taken by a laser pulse to travel to the target and return to the receiver, dividing this quantity in half and multiplying it by the speed of light c in the air (2). Typically, phase measurement systems have a reduced range (a few hundred meters) compared to ToF ones since it is necessary for the return wave to have a specific energy to activate the measurement; however, they are faster and more accurate (at low ranges).

$$d = c * \frac{\Delta t}{2} \tag{2}$$

LiDAR can also be classified according to scanning mechanisms: the concentrated laser beam produced by the excitation of the gain medium can hit an oscillating or rotating mirror, a palmer scanner, or a Risley prism (Figure 17).

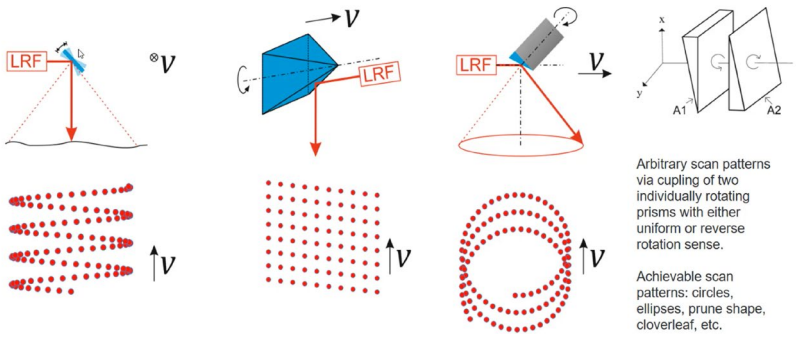


Figure 17: Scanning mechanisms of LiDAR sensors and acquisition pattern of a moving sensor – From left to right: oscillating mirror, rotating mirror, palmer scanner, Risley prism. Modified from [83].

In the case of an oscillating mirror, it oscillates between an initial and final position; when the LiDAR is equipped on a moving vehicle, the points acquired by a single laser are characterized by a meandering pattern, with a greater distribution at the end of the field of view, i.e. when the oscillating mirror slows down, stops, and reverses the direction of movement. The rotating mirror mechanism is based on the constant rotation speed of a polygonal mirror, which produces a homogeneous point distribution in the point cloud; on the other hand, this mechanism reduces the effective scan rate since the surface near the edges cannot be used to reflect the laser beam due to geometric reasons. In the palmer scanner mechanism, a tilted mirror (mounted obliquely with respect to the laser beam source) rotates with a constant rotation; it causes a conical movement of the laser beam, and a circular pattern can be observed in the point cloud. As in the case of the oscillating mirror, the density of points is greater at the border in this mechanism. The last mechanism (Risley prism) relies on two coaxial tilted mirrors, which can rotate with the same or different velocities and in the same or different directions; according to the combination of velocities speeds and directions, different scan patterns can be achieved.

Among the aspects not to be overlooked when deciding to use a LiDAR to conduct a survey are the resolution and scanning density. Resolution represents the number of points per unit length (or area), while density is the distance between two contiguous points. These two parameters describe the ability to detect and describe small geometric objects accurately; they are not constant (unless considering a sphere concentric to the instrument) and depend on the distance and inclination of the detected surface concerning the direction of the incident laser beam. Furthermore, the phenomenon of laser beam divergence must be taken into consideration. The divergence γ [mrad] measures how much the laser beam diverges at a specific distance R; it is essential to know the footprint of the laser beam when it hits the target by quantifying its diameter d as a function of the wavelength λ and the aperture diameter D (3):

$$d = \gamma * R = 2{,}44 * \frac{\lambda}{D} * R \tag{3}$$

For instance, modern medium-range terrestrial laser scanners have a beam divergence equal to 0.14 ÷ 0.35 mrad (resulting in 1.4 ÷ 3.5 cm laser footprints at 100 meters distance). Table 2 summarizes the main characteristics of LiDAR.

LiDAR sensors have applications in a wide range of fields, from geospatial mapping and classification in urban [84], [85] or natural environments [86], [87] to autonomous driving [88], archaeological and cultural heritage studies [89], underwater knowledge [90], [91] and forestry management. In the forestry domain, LiDAR plays a crucial and versatile role. At the same time, LiDAR can provide valuable information not accessible using field methods or optical remote sensing observations and has benefits in terms of speed of data acquisition, data accuracy, costs and coverage compared with traditional methods of acquiring the same information in the field [92]. The primary use is that

of forest monitoring and management: this can happen through the estimation of forest parameters and continuous investigations over time [93], [94], [95], [96], evaluating the evolution of restoration plantations [97], assessing deforestation effects [98], [99], detecting fires and fire damages [100], [101], or by mapping the vegetation structure for biodiversity analysis [102].

Table 2: Main characteristics of LiDAR sensors.

Sensor	Spectrum	Pros	Cons
LiDAR	Ultraviolet, visible or near-infrared spectrum	Speed and ease of data acquisition, high precision, 3D reconstruction, wide operability, vegetation penetration, variety of applications	High costs, weights and dimensions for very high-resolution sensors, sensitive to fog or rain, critical issues with reflective surfaces

LiDAR is a mighty sensor; however, it is crucial to correctly manage and implement the acquired data optimally in a standard reference system. It is, therefore, necessary to integrate this sensor with tools dedicated to addressing the positioning and orientation of the LiDAR. A complete integrated system involves the implementation of a Global Navigation Satellite System (GNSS) and an Inertial Navigation System (INS), the knowledge of which will be explored in more detail in Paragraph 3.2.3.

3.2.2 Platforms

*W*hen planning a survey, in addition to choosing which type of instrument to use, it is equally essential to identify the best way to use it: depending on the extension of the area under examination, the precision

required, the type of analysis to be carried out, it may be necessary use terrestrial (which in turn can be static or dynamic) or aerial support (satellites, airborne, drones).

Aerial

Since 1972, several hundred satellites have been launched into orbit around the globe for Earth Observation (EO) purposes. Nowadays, the diversity of EO technology allows for data acquisition through different sensors (e.g., optical, infrared, thermal, radar and LiDAR) at different spatial, spectral and temporal resolutions [103]. As a matter of fact, over time, several satellite constellations have been launched into orbit, such as Landsat (the first mission, launched by NASA in 1972), MODIS and Sentinel (Figure 18); they differ in the sensors equipped, the spectral bands acquired, the spatial and temporal resolution.

EO satellites monitor the Earth's land, ocean, atmosphere, cryosphere, and carbon cycle from space in real-time and constantly transmit that information to the ground [104]. Satellite imagery has been widely analyzed, among other objectives, for classification tasks [105], glacier melting assessment [106], and risk zone mapping [72]. Even though image resolution has improved significantly since the first satellite was inaugurated, satellite images are mainly used for medium or large-scale studies, while detailed analysis on a small scale is still out of reach. In more recent years, active instruments such as SAR and LiDAR have also been equipped on satellites; a considerable amount of studies have been conducted using satellite SAR data [105], [107], [108], while as regards LiDAR, this is not yet systematically used in the research field due to poor spatial resolution on the ground. The main applications of satellite LiDAR (such as the ICESat satellite, launched by NASA and operative from 2003 to 2010) refer to atmospheric investigations [109].

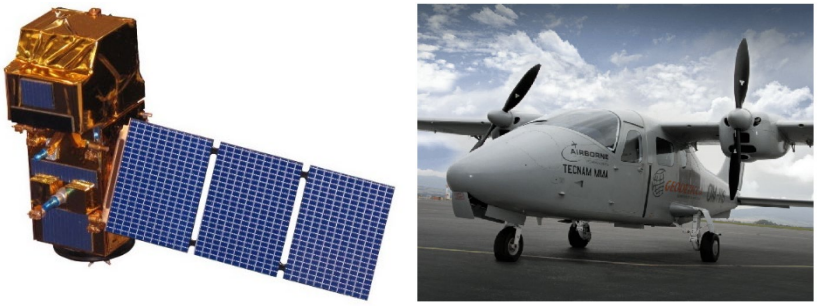


Figure 18: Examples of a Sentinel satellite (on the left) and an aircraft for aerial survey [110] (on the right).

Satellite aerial surveys allow the acquisition of a considerable amount of information and survey areas of substantial extension; on the other hand, spatial and temporal resolution can be limiting, so aircraft or Uncrewed Aerial Systems can be taken into account to acquire data. Compared to satellite images, these supports offer greater spatial resolutions and acquisition frequencies at the expense of the extension of the area, which is reduced locally. Flight planning is the first step to performing an aerial survey optimally. It aims, according to the expected final result, to (i) define the most suitable platform, (ii) choose the sensor (or sensors) to be used, and (iii) communicate the required flight characteristics. The extension of the study area, its location (if in an urban or rural area, if there are flight restrictions or if the area can be flown over without any limitation), the required payload (the aerial support must be able to support the weight of the sensor mounted on it), the resolution required, and the economic availability are factors that influence the decision to consider an aircraft or a UAV, or *drones*, for the aerial survey. Historically, the first platforms used to carry out aerial surveys were helicopters or aircraft; they allowed the survey of medium-sized areas using different types of sensors, which initially had more significant weight. When planning an airborne survey, it must be considered that the distance between parallel flight strips must be large

enough (150-500 m) to allow a safe turning maneuver; the distance can be reduced to 50-60 m when helicopters are considered.

Subsequently, with the advent of drones and the contemporary lightening of sensors, the latter has become more versatile for studies in very small areas (Figure 19).

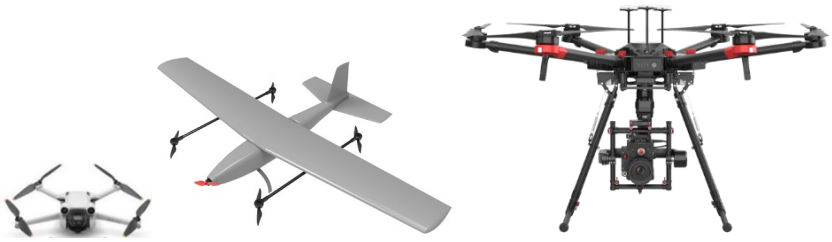


Figure 19: Examples of UAV (multirotor on the left and the right, fixed wing in the middle).

These are typically used in environmental and urban monitoring, cultural heritage, precision agriculture, and forestry. Drones can be distinguished according to (i) flight autonomy with a single battery (from a few minutes up to half an hour of autonomy for standard drones, while recent hybrid models rely on the combustion engine that charges the batteries during flight, extending the autonomy up to 2.5-3 hours of flight), (ii) the equipped sensor (commercial drones can equip standard RGB, thermal, multi-spectral cameras or LiDAR sensor, or they can have the predisposition to use additional sensors), (iii) the weight (depending on the weight and according to national legislation, heavier drones cannot be used in an urban environment or in the presence of civilians who are not informed about the flight mission), the type of construction (*fixed-wing drones*, which generate lift with their wings, consuming less energy to stay in the air, ensuring more excellent flight stability and increasing autonomy, but at the same time making maneuverability and the take-off and landing phase difficult, and preventing hovering, or

multirotor drones, which are more maneuverable, versatile, and easily transportable, at the expense of flight duration). Nevertheless, weather and lighting conditions should not be overlooked, as these can affect the dataset's quality or severely impact flight range and security.

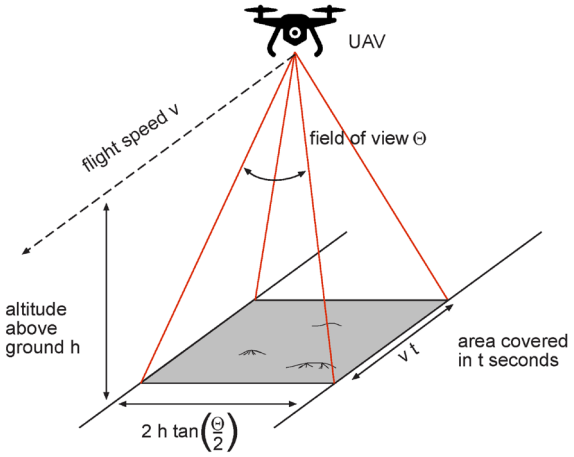


Figure 20: Schematization of the swath width related to flight altitude and sensor scan angle parameter [111].

When using photographic sensors in aerial surveys, it is also necessary to guarantee a certain percentage of overlap (generally between 60% and 80%) between contiguous images (longitudinally and transversely with respect to the direction of flight). This requirement is necessary to process the data using the Structure from Motion (SfM) technique, through which it is possible to estimate the 3D position of a point using two or more 2D images. Aerial photogrammetry is widely used in the forestry domain, mainly for estimating forest parameters and health monitoring [64], [112], [113]. On the other hand, the aerial acquisition technique through LiDAR is called Aerial Laser Scanning (ALS). The requirements for planning ALS surveys concern the overlap between adjacent flight strips to ensure complete coverage of the area

under examination, the quality assessment analysis, and the increased points density. With these sensors, the overlap is typically greater than 50%. Other parameters to be considered are the footprint diameter (previously described) and the swath width (Figure 20).

The swath width (SW) depends on the flight altitude (h) and the total scan angle (ψ) according to (4):

$$SW = 2 * h * \tan\left(\frac{\psi}{2}\right) \quad (4)$$

The mean point density can also be defined a priori as a relationship between the number of points acquired per second (equal to the effective scan rate) and the area covered per second as follows:

$$Point\ density = \frac{Nr.points\ per\ second}{Area\ per\ second} = \frac{Scan\ rate}{SW * v} \quad (5)$$

Where v represents the flight velocity.

Terrestrial

Terrestrial surveys are typically conducted to acquire information about natural or artificial objects with a limited spatial extension; due to the oblique and non-nadiral acquisition point perspective, they are not used to acquire terrain information [114], [115]. The distances at which data are obtained via a terrestrial survey are shorter than those from aerial; for this reason, the final output usually has higher accuracy and resolution [115], allowing the analysis of specific or localized phenomena in areas with limited extensions in a highly detailed manner.

Terrestrial photogrammetry, or close-range photogrammetry, in which acquisition distances reach a maximum of approximately 300 meters [116], is widely used in the architectural field for the monitoring

of assets of cultural and historical interest [117] or for 3D modelling [118], as well as for the monitoring of geological features and discontinuity characterization [119] or bridge deformation and geometry measurement [120]. In the forestry sector, this acquisition methodology has been used to estimate the DBH [121] and to reconstruct the stem surface [122], [123]. Its strengths lie mainly in this procedure's cost-effectiveness and ease of data management.



Figure 21: Examples of a Terrestrial Laser Scanning data acquisition (on the left) and a handheld mobile Laser Scanning System (on the right).

As with aerial acquisitions, LiDAR can also be used in terrestrial acquisitions. Although the sensors are more expensive than photogrammetric ones, they allow obtaining a greater quantity of data, even in difficult-to-access areas, in a shorter amount of time. Terrestrial acquisitions are called Terrestrial Laser Scanning (TLS) operations. They can be static, in which the instrument is positioned on a tripod during the acquisition phase, or dynamic, in which the instrument is equipped on a moving platform. In the latter case, when the final objective is to map an area, we refer to a Mobile Mapping System

(MMS), and when a LiDAR is used, we refer to that technique as Mobile Laser Scanning (MLS). When, in addition to mapping, an estimate of the sensor's position is performed through technological systems and algorithms, we talk about SLAM (Simultaneous Localization and Mapping) systems (Figure 21). Mobile Mapping systems represent a hybrid solution for acquiring medium/large-sized areas, in which the movement of the sensor itself compensates for the limited range and instrumental point of view. In this case, integration with other sensors (described in the following Paragraph) is essential to acquire data correctly.

In literature, it is possible to find a vast number of studies in which Terrestrial Laser Scanning (TLS) surveys are conducted in very disparate domains, such as cultural heritage [124], Earth science [115], forest inventories [125], [126], structures monitoring [127], [128], etc. MMS have broad applications in autonomous driving [129], [130], road quality assessment [131], [132], Building Information Modeling (BIM) generation [133], [134] and vegetation mapping and detection [135], [136], [137], [138], [139].

Data fusion

Very often, an integrated approach that combines different types of sensors or different platforms is mandatory to obtain a result with a greater degree of detail and precision or to achieve unattainable goals using a single type of sensor [140], [141], [142], [143], [144], [145], [146], [147]. Data integration from different sensors obtained with varying types of acquisition is a common practice in geoinformatics and environmental monitoring fields. As a matter of fact, since RGB, multispectral and SAR images, and LiDAR point cloud can capture different types of information about the scene, each of them can be used for a particular task. Still, none of them can capture the scene in such a complete way as to use a single type of data to replace the others [148],

[149], [150], [151]. For example, hyperspectral images can characterize objects' spectral and spatial characteristics. However, it is difficult to distinguish objects with similar spectral characteristics but different elevation information; compared with these images, LiDAR data have accurate three-dimensional information, which can classify objects using height information. Therefore, it dramatically improves classification results [143]. Integrating RGB images and LiDAR point clouds is widespread because combining visual information with three-dimensional data can significantly enhance the understanding and interpretation of the surrounding environment [142]. For this reason, instruments that integrate these two sensors are widespread on the market (Figure 22).



Figure 22: Examples of Hybrid RGB/LiDAR systems.

In the same way, aerial and terrestrial acquisition are very often integrated to obtain a complete description of scenarios in which some elements develop in elevation and for which terrestrial acquisitions fail to acquire data in the upper part. In contrast, the opposite happens with aerial acquisitions [142], [152]. In the forestry domain, data integration is often needed to achieve good forest parameters assessment at single tree level [153], [154], [155], [156]. Figure 23 shows the final result of

an urban environment scenario obtained with an integrated procedure between LiDAR and photogrammetric instrumentation with terrestrial and aerial acquisitions [142].

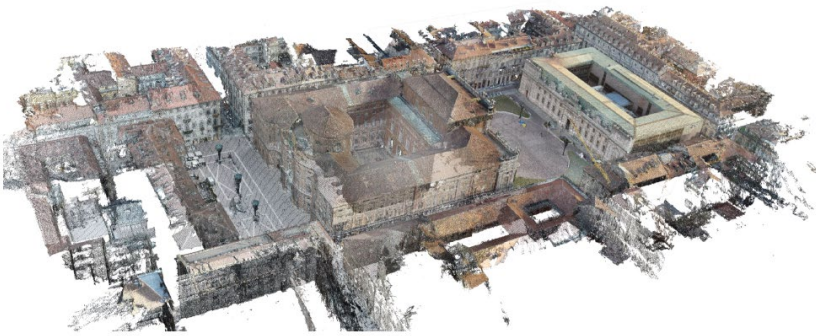


Figure 23: Example of data fusion: LiDAR and photogrammetric point cloud obtained through the integration of aerial and terrestrial platforms [142].

Furthermore, the importance of accurate planning should not be underestimated, as well as data standardization and the adoption of advanced analysis methodologies to integrate the different data types efficiently.

3.2.3 Methods for georeferencing

Geomatics surveys aim to acquire information in space. For this reason, aerial or terrestrial platforms comprise an integrated multisensor platform in which the acquisition instrument takes care of the data acquisition. However, it is equally essential to georeference the acquired data, that is, to associate geographical coordinates so that its position is known univocally on a map or a terrestrial surface. This process is critical to integrate data and enable spatial analyses. Indeed, a Global Navigation Satellite System (GNSS) and an Inertial Navigation System (INS) are needed to estimate the platform's exact position and orientation

to georeference the data acquired. As discussed in the following paragraphs, these two systems are complementary.

On the one hand, GNSS provides a global position but can be subject to signal interruptions, multipath and atmospheric variations; on the other hand, INS provides estimates of position, velocity and orientation based on inertial sensors but may be subject to drift over time. By integrating GNSS and INS data, overall positioning errors can be reduced; additionally, during GNSS signal interruptions, the INS can provide reliable estimates. In this way, an uninterrupted and accurate position estimate can be obtained, which is fundamental, among other applications (e.g., autonomous driving, seamless positioning) for georeferencing collected data. Nevertheless, any interruptions in INS data is typically fatal, as the attitudes of the system cannot be determined.

The following paragraphs briefly mention the basic principles of these two systems.

Global Navigation Satellite System

GNSS (Global Navigation Satellite System) is a system of positioning and navigation based on the reception of radio signals between a constellation of artificial satellites and a receiver antenna. To date, there are four global operational GNSS constellations: GPS (the first constellation launched into orbit for positioning purposes by the USA in 1978), GLONASS (Russia), Galileo (Europe) and Beidou (China). There are also navigation satellite systems that operate at a regional level, such as QZSS (Japan) and IRNSS (India). GNSS relies on three main components (segments): a space segment, a control segment, and a user segment (Figure 24).

The space segment comprises MEO (Medium Earth Orbit) satellites in orbit around the Earth. It transmits radio signals, maintains an accurate time reference, and acquires information from the control segment. The latter continuously monitors the orbit of the individual satellites, sending

trajectory corrections input and transmitting each satellite's ephemeris (i.e., the position). Finally, the user segments are signal receivers equipped with an antenna that the end user uses for positioning or tracking purposes. Each receiver receives signals from all satellites above the horizon and not covered by obstacles (e.g., trees, buildings, mountains). The distance of each satellite from the receiver is calculated through the signals since this is proportional to the signal's propagation time. Finally, by knowing the satellites' ephemeris, the receiver's position is calculated through trigonometric relations.

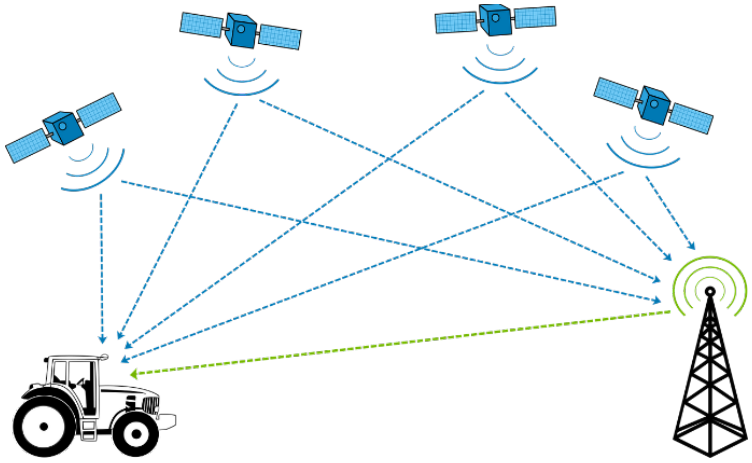


Figure 24: GNSS scheme.

The measurements with which GNSS positioning is obtained can be distance (pseudorange) or phase. Pseudorange (ρ) measurements rely on the measurement of the propagation time of the radio signal transmitted from the *j-th* satellite to the *i-th* receiver according to a trigonometric relationship (6):

$$\rho_i^j(t) = \int_{t_o}^{t} v \cdot dt = \\ = \sqrt{(x^j(t_0) - x_i)^2 + (y^j(t_0) - y_i)^2 + (z^j(t_0) - z_i)^2} \tag{6}$$

There are three quantities to determine (x_i, y_i, z_i) to which is added the time difference δt between the satellite clock and the receiver clock (7). For this reason, at least four satellites are needed to obtain a solution.

$$t_i = t^j + \delta t \tag{7}$$

The main advantages of using the pseudorange measurement method are the speed of position estimation (in real-time) and the ease of operability, as only one receiver is used. On the other hand, the main uncertainties concern the satellite's position calculated through the ephemeris and the signal's propagation speed in the atmosphere, including ionospheric and tropospheric delays and scintillations. Thus, the precision in this positioning method can be a few meters in good conditions, but it can drop to tens of meters when interferences disturb the signal.

The phase measurement method is widely used for topographical purposes, through which a centimeter accuracy is achieved. In this method, the phase difference between the wave transmitted by the satellite and that received by the receiver is measured, the frequency of which varies due to the Doppler effect due to the relative speed between the satellite and the receiver. The equations to solve are the following:

$$P_i^j(k) = \rho_i^j - cdT_i + cdt^j + \alpha_k I_i^j + T_i^j + m_{k,i}^j + \\ + E_i^j + cL_i^j + \varepsilon_{k,i}^j \tag{8}$$

$$\phi_i^j(k) = \rho_i^j - cdT_i + cdt^j - \alpha_k I_i^j + T_i^j + M_{k,i}^j + \\ + E_i^j + \lambda_i N_{k,i}^j + \varepsilon_i^j \tag{9}$$

Where:

- $P_i^j(k)$ is the pseudo-range on the *k-th* frequency;
- $\phi_i^j(k)$ is the carrier-phase on the *k-th* frequency;
- cdT_i is the bias of the receiver clock multiplied by the speed of light;
- cdt^j is the bias of the antenna clock multiplied by the speed of light;
- $\alpha_k I_i^j$ represents the ionospheric propagation delay, where α_k is a known parameter depending on the *k-th* frequency;
- T_i^j is the tropospheric propagation delay;
- $M_{k,i}^j$ is the multipath error;
- E_i^j is the ephemeris error;
- $N_{k,i}^j$ is the phase ambiguity that must be estimated as an integer value in order to achieve centimeter level of accuracy using two different receivers acquiring data from identical satellites;
- λ_i is the frequency length;
- $\varepsilon_{k,i}^j$ are random errors.

GNSS positioning can be done in Real Time Kinematic mode (RTK) or using post-processing techniques. Typically, this second approach is used when high-precision positions are required; in RTK mode, only centimetric accuracies are achieved.

GNSS is a handy tool in the geomatics field, as it allows to estimate an accurate position on a global scale with a worldwide satellite coverage; however, it is also characterized by some limitations. First of all, it can be limited (or in extreme cases absent) in severe environments,

i.e. in environments in which particular obstacles (e.g., urban buildings, tall vegetation, mountains or distinct land conformations) block or interfere with the signal; furthermore, radio or other interference can also cause inaccuracies in position estimation. In order to have sufficient accuracy, it is necessary to have a clear view of the sky and to program the data acquisition campaign in such a way as to have the greatest number of satellites possible on the horizon line. However, GNSS is often integrated with INS to address the limitations mentioned above.

Inertial Navigation System

The Inertial Navigation System (INS) is a self-contained navigation technique in which measurements provided by accelerometers and gyroscopes are used to track the position and orientation of an object relative to a known starting point, orientation and velocity [157]. Inertial navigation is based on the 1st Newton law, which states that, inside an inertial reference frame, it is possible to define a body's motion with constant mass, observing the external forces acting on it. The main component is the Inertial Measurement Unit (IMU), whose typical configurations include one accelerometer (which measures inertial accelerations as specific forces) and one gyroscope (which measures the rate of rotation) for each of the three principal axes.

Accelerometers measure specific forces f, while gyroscopes measure the rotation rate along specific axes ω. The INS keeps track of the movement of an object, quantifying the movement performed in an interval of time, starting from a known position. The kinematic relation between acceleration, velocity and position is a crucial principle in this process since velocity can be obtained by integrating acceleration measurements. In contrast, the position can be yielded by integrating the velocity. The second fundamental component is the central process unit, which elaborates the acquired data; it calculates the velocity and,

combining the information related to the direction of the movement, estimates the new position of the system.

INS can be classified according to how they are mounted on the object. *Strapdown* systems are rigidly mounted on the object and move integrally with it, while *gimballed* systems are mechanically isolated from the object. Contrary to strapdown systems, gimballed systems are more complex but offer better stabilization and performance under conditions of high acceleration or vehicle movements (such as in aircraft).

Depending on the IMU model, this may be affected by deterministic errors (manufacturing defects minimized by calibration procedure) and stochastic errors; as a result, the estimated position drifts after a few seconds. In order to limit this error, GNSS and INS are often both equipped on the platform and integrated. GNSS and INS integration is performed in three different hybridization architectures depending on the type of GNSS measurement used and the level of integration: loosely, tight or deeply integration. Loosely coupled architectures process GNSS and INS data separately, combining data at the final output level. The difference between INS and GNSS measurements in terms of position and velocity is used to estimate the INS error. Then, the INS navigation solution is corrected with the resulting INS error estimate [158]. Tightly coupled architectures integrate data deeper since the difference between the pseudo-range, carrier phase or Doppler shift measurements obtained by GNSS and INS are used [158]. Finally, deep integrating architectures integrate the sensors at the signal-tracking loop level. GNSS receiver and INS are not independent devices; on the contrary, the INS is an integral part of the GNSS.

There are numerous advantages to integrating the two positioning methods described. First of all, in this way, the sampling rate (i.e., the number of times per second that the chipset and satellites communicate to establish the device location) is higher. In fact, since GPS provides

absolute positions at larger time intervals while INS provides position and orientation data at much smaller intervals, their integration combines the long-term precision of GPS with the sensitivity and response speed of INS. When combining a higher sampling rate with a high-quality chipset, an adequate number of satellite constellations, a correct choice of the antenna selection and its orientation, and finally, with well-working embedded software and specialized filtering, a better accuracy in the position estimation can be achieved. Additionally, the integration guarantees excellent continuity and reliability of the service. The INS provides a continuous position even when the GPS signal is interrupted in particular scenarios (e.g., urban environments) or adverse weather conditions. The final information collected is the platform's position in an absolute reference system provided by GPS, together with the accurate trajectory and attitude that refers to the information acquired by the INS.

The applications are disparate but mainly refer to vehicle navigation. The integrated positioning approach has enabled developments in assisted navigation and autonomous driving of land road transport, as well as in air navigation, flight control and maritime navigation. Finally, it plays a crucial role in controlling mobile platforms such as drones or generic autonomous vehicles, allowing them to map the surrounding environment, plan optimal routes and move autonomously in dynamic and complex environments.

Chapter 4

LiDAR DATA PROCESSING, STATE OF THE ART

The ever-increasing ease of data acquisition, encouraged by technological advancement from a sensor point of view described in the previous Chapter, leads to an ever-increasing quantity, quality and type of data acquired. Therefore, nowadays, it is indispensable to develop automatic or semi-automatic data processing methodologies, which allow data to be processed without the direct aid of the operator, as this would be time-consuming, highly subject to subjective interpretation, and unreliable. This Chapter delves into point cloud collection and processing techniques, specifically focusing on standard practices and algorithms involved in forestry studies. The first paragraph describes all the procedures relating to pre-processing, which are always necessary when dealing with LiDAR data; the following paragraphs refer to the techniques used to process point clouds in a forest environment, describing segmentation algorithms and subsequently delving into

methods for estimating woody volume and biomass; finally, carbon dioxide estimates techniques are described.

4.1 Data collection

Data collection using LiDAR technology consists of all the practices required to acquire a complete dataset with all the information necessary to conduct in-depth analyses. It consists of planning, configuring the sensors and acquisition platforms, and the execution of the survey itself and all the preparatory activities. Below, the operative procedures are distinguished according to the type of acquisition (i.e., aerial, static terrestrial and cinematic terrestrial acquisition).

Planning starts by defining the objectives of the survey. Before data collection, several factors must be considered, including the geographic area of interest, the required point density, the desired spatial resolution and weather conditions. Planning must also consider potential obstacles, such as trees or buildings, that could affect the data quality collected. Furthermore, the area of interest must be precisely defined, analyzing every characteristic. This operation allows to choose the best type of acquisition and sensor to use in the survey phase.

In the case of aerial acquisitions, it is necessary to define the type of aerial support (plane, helicopter or drone) based on parameters such as flight altitude, platform speed, sensor field of view, overlap of scanning trajectories and other factors that affect the point cloud density. Last but not least, the economic aspect must also be taken into consideration in terms of price per hectare acquired. The trajectory of the acquisition platform must be planned in such a way as to cover the entire area of interest uniformly, minimizing the shadow areas and ensuring a homogeneous distribution of the LiDAR points. Furthermore, weather conditions should not be overlooked, as they can affect the data quality and on-site safety. No less important is checking local laws and

regulations regarding airspace and ensuring you have permits when necessary. Land acquisitions also require a well-structured planning phase. In the case of static terrestrial acquisitions, it is necessary to plan the scanning positions to ensure complete coverage of the survey area, optimizing visibility and minimizing shadow areas. In the case of dynamic acquisitions, it is necessary to plan the platform's trajectory based on the survey area's characteristics. The trajectory must cover the entire area of interest to collect data efficiently; furthermore, it is helpful to undertake a closed and possibly intertwined trajectory to ensure complete data capture by limiting the effects of the drift of the Inertial Navigation System.

An integrated survey with topographic and GNSS instrumentation allows the acquisition of information used in the pre-processing phase to register and georeference the point cloud. Planning and carrying out a topographic survey is also necessary, using more traditional instruments such as total stations and prisms to create a framework network. Then, GNSS receivers are placed to measure the coordinates of the total station position. The raw data is collected and subsequently processed.

4.2 Point cloud pre-processing

The pre-processing procedures of a LiDAR point cloud concern, first of all, registration; subsequently, it is necessary to georeference the resulting cloud, filter it and possibly color it by integrating photogrammetric sensors (Figure 25). Each of these operations, however, represents processes widely discussed in the literature. Nonetheless, in this thesis work, they have been summarized as preliminary operations and to place greater focus on the subsequent operations more closely linked to the thesis objectives.

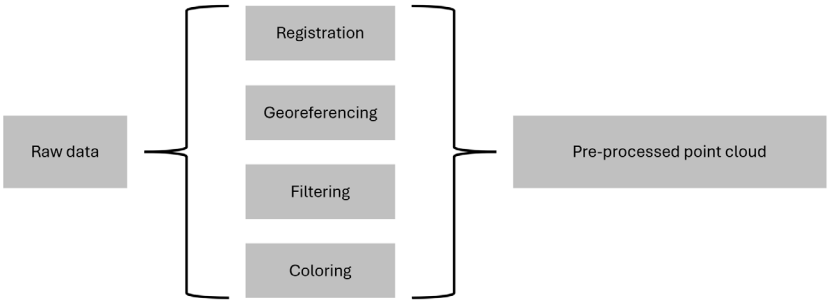


Figure 25: Scheme of pre-processing procedures of LiDAR point clouds.

Point cloud registration, or scan matching, consists of automatically, semi-automatically, or manually relatively aligning two (or more) point clouds to obtain a single coherent dataset in a specific coordinate system.

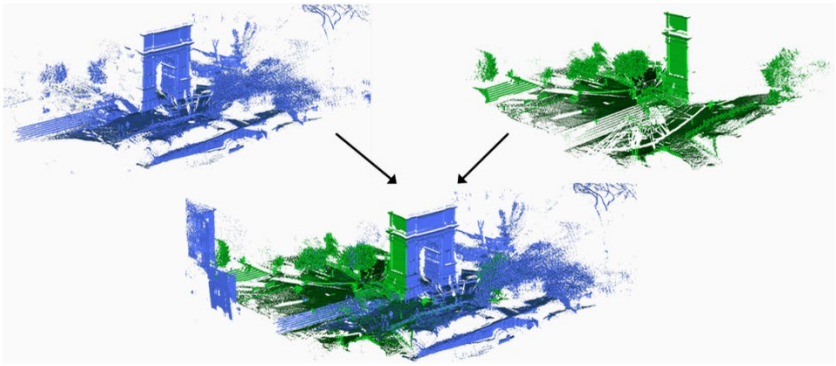


Figure 26: Example of point cloud registration of two point clouds acquired in an urban environment [159].

This procedure affects TLS acquisitions where external information such as GNSS is not used; for acquisitions carried out with integrated systems (GNSS and INS), the point cloud registration is not performed, and the coherent point cloud is obtained considering position and movement information for each scan through the georeferencing scan.

As a matter of fact, a LiDAR scanning survey of a generic study area requires carrying out several to sufficiently and satisfactorily detect every part of the area of interest. The alignment procedure is based on the use of the spatial rototranslation equation, which calculates the new position (x', y', z') of a point according to (10):

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R_{3x3}(\omega, \varphi, \kappa) * \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} \quad (10)$$

R_{3x3} is a 3x3 matrix expressed as a function of the rotation angles ω, φ, κ, each of which relates to the rotation around the x, y, and z axes, respectively, while $\Delta x, \Delta y$ and Δz are the translation components. The equation's solution requires estimating three rotation and three translation parameters; it is, therefore, essential to know the coordinates of at least three points in the two reference systems. Once the parameters have been estimated, the rototranslation equation is applied to register all the cloud points. The points considered to estimate the parameters must be spatially distributed (i.e., not aligned or concentrated in a small portion of the area). These points can be automatically or manually selected; usually, artificial retroreflective targets can be used (previously placed into the acquired scene) or elements characterizing the shapes and geometries of the detected objects (e.g., in an urban environment, edges of maintenance holes or architectural elements on facades). Their radiometric response allows retroreflective targets to be automatically detected and usually does not require manual selection. In order to have points in common between 2 subsequent scans, it is advisable to plan the survey to obtain an overlap of approximately 30%. Various alignment techniques have been developed and are present in the literature; among them, the Iterative Closest Point (ICP) is the most commonly used. This semi-automatic registration algorithm does not require the identification of reference points but selects them automatically; subsequently, for each

point, it identifies the closest point of the cloud to be registered, estimates the rototranslation parameters through a point-to-point mean quadratic distance metric minimization, and finally apply the transformation. This operation is iterated several times until convergence is reached. The advantages of this method lie in the simplicity and ease of implementation; furthermore, this method also works for unfiltered point clouds. Conversely, it is necessary to roughly align the point clouds to speed up the iterative process.

Subsequently, georeferencing aims to attribute geographical information to a given object with respect to a global coordinate system.



Figure 27: Examples of retroreflective targets (on the left) and edges of regular human-made objects (on the right, in the red circle) to be distinctively detected for the alignment procedure.

In the case of terrestrial point clouds acquired without GNSS sensors, it is necessary to carry out a topographic survey, in which the markers previously placed in the scene are essential for the point cloud co-registration phase. GNSS receivers and Total Station (TS) instruments are required in this operation. The topographic net consists of acquiring the position of the markers in local coordinates through the total station; the GNSS receiver is placed on the same tripod in order to know the acquisition position expressed in global coordinates. Finally, it is sufficient to apply a rigid rototranslation by replacing the local

coordinates of the total station with those described in the global reference system.

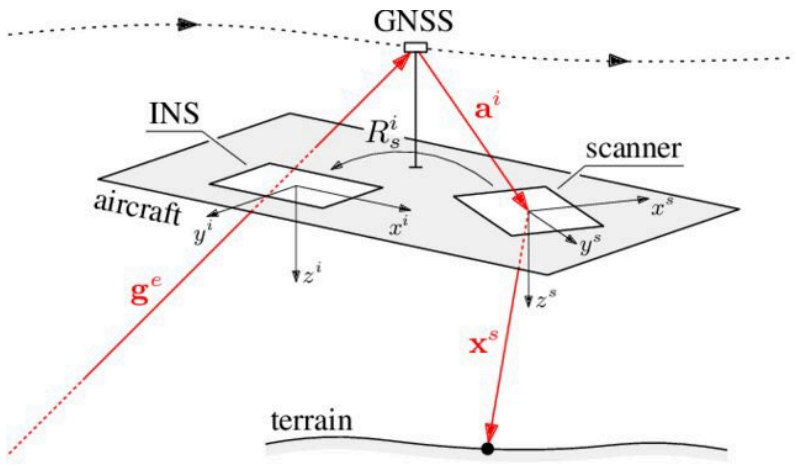


Figure 28: Georeferencing of ALS clouds [160].

On the other hand, in the case of aircraft-integrated systems, the information acquired from the LiDAR (position of the points in the point cloud in the local reference system), from the GNSS (position of the sensor) and the INS (trajectory of the sensor) is used simultaneously to georeference the cloud. The position of a point in a global reference system ($x^e(t)$) can be expressed according to (11):

$$x^e(t) = g^e(t) + R_n^e(t) * R_i^n(t) * (a^i + \\ + R_s^i * x^s(t)) \quad (11)$$

In which:

- $g^e(t)$ represents the coordinates of the GNSS receiver acquired by the sensor;

- $R_n^e(t)$ is the rotation matrix from the navigation system to the Earth-Centered, Earth-Fixed (ECEF) coordinate system;
- $R_i^n(t)$ is the rotation matrix describing the rotation from the inertial system to the navigation system;
- a^i is the level arm (offset between the GNSS antenna and the origin of the reference system of the scanner);
- R_s^i is the rotation matrix describing the rotation from the scanner system to the inertial system;
- $x^s(t)$ is the vector with the coordinates of the laser point in the scanner system.

The GNSS and INS data are processed through a Kalman filter to estimate the trajectory of the aerial platform in global coordinates [161].

After registering and georeferencing the point cloud, the filtering phase is essential to improve the data quality. A limitation of the survey using LiDAR instrumentation lies in the absence of a logical criterion in acquiring points; the laser scanner automatically acquires millions of 3D points only according to the methods dictated by constructive constraints. As a result, generally, there may be cases in which the points acquired in a given area are redundant (e.g., to describe flat surfaces, a few points are sufficient, while to describe discontinuities or more irregular objects, a more significant number would be needed); furthermore, moving elements (people, animals, objects subject to the force of the wind, etc.) can generate noise. For this reason, it is necessary to filter the cloud to reduce noise and remove undesired points; depending on the final objective to be achieved through the use of the point cloud, it is also possible to subsample the cloud in order to lighten it, and possibly speed up subsequent data processing procedures. Subsampling procedures can be random-based or spatial-based; in the second case, the 3D space is divided into regular cubic cells (voxel), for

each of which a single point is selected. The selection within each voxel can be random, or the point closest to the center of gravity can be chosen.

The final step of coloring is optional. It occurs when the laser scanner is combined with photographic equipment; integration can also occur within the instrument if hybrid sensors are used (Figure 22). Coloring occurs by assigning the RGB value of each raster pixel to each point of the point cloud with the same coordinates.

4.3 Point cloud segmentation

Segmentation is the procedure of dividing data into significant homogeneous areas to facilitate its graphic representation and processing. Specifically for point clouds, segmentation refers to the process of labelling each point of the point cloud so that points with similar characteristics are grouped in order to add important information to the dataset and develop subsequent, more specific analyses.

3D point cloud segmentation can be performed at different scales according to the level of detail required. When considering a scene, the identification of separate entities (e.g., terrain, buildings, trees, etc.) is commonly referred to as *semantic segmentation*; differently, *instance segmentation* refers to a semantic segmentation in which an object-level analysis is also performed, treating multiple objects of the same class as a single entity; finally, a further level of detail can be achieved through *part segmentation*, where more information about the composition of individual objects is exploited. Point cloud segmentation is a fundamental step to further process point clouds effectively; research areas that commonly need segmented point clouds concern autonomous driving [162], [163], object detection [164], classification tasks [165], and feature extraction [166]. Precisely for this study, the segmentation processing was applied in forest environments and divided into different steps in order to (i) identify the terrain and distinguish it from the rest of

the scene (i.e. trees), performing a semantic segmentation; (ii) detect individual trees (Individual Tree Detection) as instance segmentation; (iii) distinguish, for each tree, the woody part from the leaf part (part segmentation).

The point cloud segmentation research topic is hotly debated due to the complexity of the point cloud's structure. Performing manual segmentation is nowadays an unviable solution due to the large quantity and density of data; as it will be deepened in the next Paragraph, this operation is often carried out only during the validation phase of the automatic segmented point cloud when considering small portions of the dataset is suitable and less time-consuming. Countless methods and algorithms have been developed and discussed in the literature to segment point clouds. As already anticipated, a classification into which these can be divided concerns the level of detail to which they are applied. It can, therefore, be used to identify scenes, individual objects, or parts of objects. Segmentation methods can also be classified, considering the operational approaches; each technique has strengths that make it more suitable for carrying out specific segmentations in particular scenarios than others.

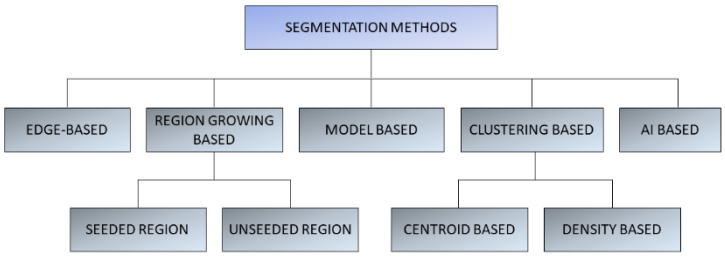


Figure 29: Scheme of segmentation approaches.

Edge-based segmentation methods identify the edge of objects and regions and then group the point clouds within the boundaries, assigning the same label. The edge detection is performed according to selected

features, often geometric, such as normal, curvatures and gradients. More specifically, the boundary of an object is assigned where changes in these parameters above a predefined threshold are detected. This approach is widespread [168], [169] in very highly dense point clouds; in fact, the main limitations of the edge-based segmentation method emerge when considering noisy point clouds or with insufficient or non-homogeneous point cloud density. Object edges are incorrectly identified in these cases, particularly for more complex shapes and object structures. Nonetheless, studies have been conducted using segmentation methods based on an edge-based approach for straight-line segment extraction, point cloud filtering, and ground breaklines extraction with ALS data [170].

Region-growing methods rely on the position of nearby points with similar properties to isolate objects with different features; typically, this approach is less sensitive to noise points and uneven point cloud density. This method can be further subdivided into *seeded-region-* or *unseeded-region*-based methods. Seeded-region methods require selecting seed points around which a region iteratively grows, adding neighbour points according to defined thresholds. The final region around each point is determined when no neighbour point respects the defined boundary conditions. Segmentation using this method is relatively fast and allows for setting the best parameters for region detection depending on the object or scene to be segmented; however, the result strongly depends on the starting seed points for generating the area. In fact, an incorrect choice of these points can result in obtaining an incorrectly segmented point cloud [167]. The choice of seed points is therefore crucial to get a satisfactory segmentation; these points are selected according to the best feature that identifies the object of the scene; surface features such as normal and curvatures are commonly considered [171]. In contrast, unseeded-region methods initially assign all points to a class; subsequently, considering one or more features as thresholds, they divide

the initial into smaller areas until these constraints are no longer respected. This approach has been used to define the geometry of buildings [172]; however, it is not commonly widespread as the interruption criterion of the subdivision of the point cloud is more difficult to identify, often causing over-segmentation [173].

Model-based methods group points with the same mathematical representation according to geometric shapes. The Random Sample Consensus (RANSAC) model belongs to this category [174], initially developed for image processing, and subsequently adapted in several point cloud studies [175], [176], [177]. Model-based methods are commonly used to detect plane shapes (e.g., roofs [178]) and spherical targets [175] or for feature extraction in the domain of cultural heritage [179]. In the forestry domain, this method has been applied to estimate forest parameters such as tree stems and height measurements [180], direct diameter measurements [181], allometric equations modelling based on shape detection [182] and more recently also the volumetric estimate of trees [183]. The model-based segmentation approach is robust with outliers and can easily help segment easy geometrical shapes; conversely, more complex geometries are not as easy to identify, and the model exponentially increases processing times to try to reach convergence.

For this reason, this method is not always the best in the case of complex natural environments except for the modeling of trunks or branches, which can easily be traced back to cylinders or truncated cones. The Cloth Simulation Filter (CSF) [181] can also be included among the model-based methods. This algorithm is also implemented in commercial software (e.g., Cloud Compare) and is used to identify points belonging to the terrain. This method is based on the physical principle of placing a cloth on the upside-down point cloud: when the cloth is sufficiently soft, it adapts perfectly to the distribution of points, allowing the Digital Surface Model (DSM) to be obtained; when, however, the

cloth has a certain rigidity, it only fits the ground, without following the distribution of points that describe elevated elements (e.g., houses or trees). In this way, the Digital Terrain Model (DTM) is obtained. The rigidity of the cloth is set through parameters; in fact, it is modeled through a grid that consists of particles with mass and interconnections, whose position is estimated over time. Among the methods developed to identify cloud points belonging to the terrain [184], [185], [186], [187], this is the most commonly used thanks to its fast processing and ease of application and implementation.

The segmentation approach can also be *clustering-based*; in this case, at first, attributes and features of the point cloud are calculated, and then the point cloud is segmented, combining this information and setting a threshold for each of them. In this way, it is possible to simultaneously consider different characteristics of the point cloud (geometric and spatial); however, the quality of the attributes is fundamental for a successful segmentation. Therefore, thresholds must be defined in a manner consistent with the characteristics of the objects to be segmented. Methods that rely on attributes prove to be a robust strategy for categorizing points into uniform regions, obtaining both flexible and precise results. Nevertheless, these approaches hinge on defining the neighbourhood relationships between points and the point density within point cloud data. A drawback of such methods is the time-consuming nature, especially when handling multidimensional attributes for extensive point cloud data [188]. This method is commonly used for extracting planar geometries, as demonstrated by [189], where each laser point defines a plane in the 3D attribute space, and laser points on the same planar surface intersect at the position in the attribute space that corresponds to the slopes and distance of the planar surface. Other studies have been based on the calculation of normal vectors [190] or magnitudes of surface normal [191]. The most popular clustering methods for point cloud segmentation are *centroid-based* and *density-*

based. In centroid-based cluster methods, each cluster is represented by a vector. The most common clustering approach of this type is the *k*-Means algorithm: the number of clusters is fixed to k, and the algorithm find k-cluster centers and assign points to the nearest cluster center minimizing the squared distances from the cluster. The advantages of centroid-based methods (particularly the *k*-Means) lie in the operational simplicity, speed and ease of implementation; however, some limitations should not be underestimated. First of all, it is necessary to know and define a priori the number of clusters into which divide the dataset. Furthermore, the procedure is iterative, and the initial points are randomly selected; in particular scenarios, therefore, the clustering result could vary by running the algorithm several times, losing coherence and robustness. On the other hand, density-based clustering methods rely on the number of points to define clusters, and they are more robust to outliers and noise. Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [192] and Hierarchical DBSCAN (HDBSCAN) [193] cluster methods fall into this class. They are relatively fast, do not require initialization, and the number of clusters does not need to be known a priori; unfortunately, they are mainly sensitive to changes in the density of different objects. Thus, they require a significant density change to detect borders correctly. Several studies have been conducted using *k*-Means for point cloud denoising and simplification [194] and segmentation [195], [196], [197], [198] in different scenarios, as well as processing examples based on the DBSCAN method are provided by [199], [200], [201].

In recent years, approaches based on artificial intelligence (AI) have been developed and spread rapidly. AI is a technology that enables machines to imitate various complex human skills [202]. Artificial intelligence algorithms can be macro-classified into Machine Learning (ML) and Deep Learning (DL). They allow us to identify patterns within datasets through empirical knowledge and make decisions based on the

information obtained. ML is an application of artificial intelligence that includes algorithms that analyse data, learn from them, and apply the acquired knowledge to make informed decisions. Various machine learning algorithms (models) have been developed and are massively implemented in almost every domain.

On the other hand, Deep Learning is a portion of ML; specifically, it can be seen as an improvement based on artificial neural networks, which are artificial connections built similarly to human brain neurons. ML algorithms generally need human intervention, while Deep Learning algorithms can autonomously learn from data through repetition, creating a network capable of making intelligent decisions independently. Examples of Machine Learning algorithms are *Support Vector Machine* (SVM), *Decision Tree* (DT), and *k-Nearest Neighbors* (*k*-NN). *k*-Nearest Neighbors algorithm [203] uses proximity to classify or predict point clustering based on the assumption that similar points are close to each other. This method is *supervised* (i.e. it needs a training set of a previously segmented point cloud); when the algorithm analyzes the new (unclassified) data, it classifies points based on their distance from the examples in the training set. In this approach, the choice of the *k* parameter is crucial: if it is too small, the classification is done with respect to the closest few points and could be affected by noise or outliers; on the contrary, if *k* value is too large, it contemporary increase the computational load and therefore the processing times, but it also results in greater generalization. However, it has achieved some success in the domain of point cloud segmentation [204, p. 3], [205], [206].

SVM models aim to find the class separation line that maximizes the margin between the classes. The margin is the minimum distance from the line to the points of the two classes. The area between the two margins is referred to as a hyperplane. It is necessary to define a minimal part of the training dataset (the support vectors) based on the definition of the separation line and the hyperplane. Support vectors are the values

of a class closest to the margins, which are classified with greater difficulty using different methods. It is therefore obvious that this method works very well when the clusters are well separated from each other, while it is less efficient if there is a lot of noise or if the clusters overlap. Although the method is memory efficient since it only needs a portion of the dataset, it is not suitable for segmenting large point clouds, and the training time is high. Studies involving SVM have been published concerning surface reconstruction [207], segmentation [208], [209], [210], and object detection [211]. Another family of Machine Learning algorithm used for segmentation tasks that is worthy of mention is the Decision Tree (so called DT). DT models comprise internal nodes representing a test on a feature, branches representing the result of the clustering, and leaf, the defined classes. The algorithm recursively divides the data into subsets according to input features and binary splits, ultimately leading to a tree-like structure; the partitioning process continues until a stopping criterion is met. The Random Forest (RF) algorithm is an ensemble learning technique that aggregates predictions from numerous individual decision trees to enhance overall performance and generalization. Operating as a bagging algorithm involves independently training multiple models and subsequently combining their predictions. Incorporating randomness during training, concerning both data and the features selected at each split, is designed to mitigate overfitting and bolster the model's robustness. Random forest models are usually highly high-performance and have a low probability of overfitting the dataset; furthermore, they are easier to interpret than other models. In the literature, this method is widely used to segment point clouds [212], [213], particularly in urban environments [214], [215].

Segmentation methods can be implemented in a more complex workflow, where clustering at different levels or in complex scenarios is carried out. Examples of methods integration are given by [216], [217], [218], [219].

4.3.1 Segmentation in the forestry domain

Several approaches and studies have been presented regarding point cloud segmentation in forest environments. The first step is to identify single trees to carry out single analyzes. The most commonly used methods in literature are described below.

PyCrown is a well-known open-source seeded-region growing segmentation method used in the forestry domain for Individual Tree Detection (ITD) [220]. It is fully developed in Python and detects single trees starting from seed points defined with spatial analysis raster-based. Specifically, the DTM and DSM are generated after detecting ground points; then, the Canopy Height Model (CHM) is calculated as a difference between the DSM and the DTM; subsequently, a spatial search is performed to find local maxima of the CHM which automatically detect the treetop position. Finally, the region-growing algorithms iteratively define the crown of each tree according to four different parameters, which are needed to check the height of neighbouring pixels and their distance from the seed points. As a matter of fact, according to the tree species, trees have different shapes. When the regions are finally defined, all their points are labelled as trees with unique labels. This approach is well-suited for segmenting coniferous forests since their conical shape allows treetops to be easily detected with a low rate of errors. At the same time, it may encounter more difficulties in segmenting broadleaves trees.

Furthermore, the density of the forest also plays an essential role in the accuracy of the result. When trees are close together, the canopies often interconnect with each other and the region growing algorithm fails to detect edges correctly; additionally, submissive trees (in the forestry domain, a submissive tree is defined as a tree whose top is lower than the treetop of surrounding trees) are not identified by the local maxima process since the upper part of other trees dominates them. Since treetop

detection is critical, this approach achieves better results for aerial point clouds. Figure 30 shows an example of a forest scenario segmented at a single tree level using the PyCrown algorithm.

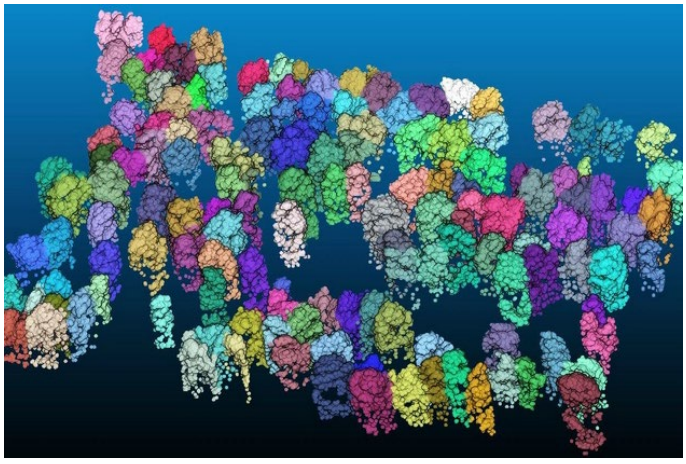


Figure 30: Example of single tree level segmentation at plot scale using the PyCrown algorithm [221].

A different Python-based open-source algorithm is the Forest Structural Complexity Tool (FSCT) [222]. It was developed to work mainly with terrestrial or high-resolution aerial data and can obtain various allometric information. However, it relies on a semantic segmentation workflow [223], which aims to work independently from the sensor used to acquire the point cloud dataset. As a matter of fact, this tool applies a complete workflow, which allows the identification of individual trees and the extraction of the main allometric parameters (i.e., DBH, height, volume). First, semantic segmentation is carried out to distinguish trees from terrain and low vegetation according to a deep learning technique, and the DTM is generated. Subsequently, coarse woody debris is further identified and excluded from further processing. To obtain a skeletonization of the tree, the stem point cloud is cut into

parallel horizontal slices, and they are then clustered using the HDBSCAN algorithm to get clusters of stems in each slice. The DBSCAN algorithm then converts the cluster skeletons into branch and stem segments. Finally, the clusters of stem segments are elaborated with a cylinder adaptation function, which allows the volume of the tree to be calculated and the DBH and height measurements to be extracted. This method is consistent, and different studies have integrated it [224], [225]; however, the leaf segmentation procedure is not implemented, and the results can significantly vary when considering leaf-on and leaf-off seasons [226].

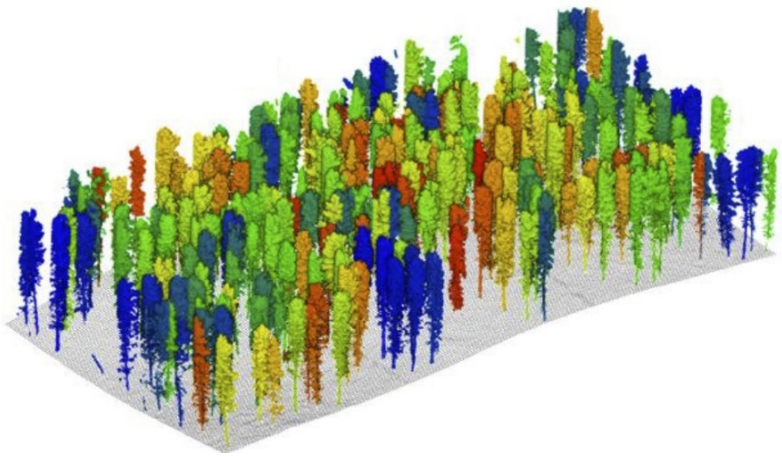


Figure 31: Example of single tree level segmentation at plot scale using the FSCT algorithm [222].

Other studies [227] apply a top-down approach to extract individual trees in urban scenarios. In contrast to the method described previously, MLS clouds are best suited to this approach. Firstly, trees are semantically segmented with a deep network approach; then, tree clusters are detected using Euclidean distance clustering. Next, a deep network is developed to detect individual trees, combining embedded

pointwise directions and detected tree centers. Treeseg [228] is implemented in C++ and nearly automatically extracts tree-level point clouds based on Euclidean clustering, principal component analysis, region-based segmentation, shape fitting, and connectivity testing (Figure 32).



Figure 32: Example of single tree level segmentation at plot scale using the Treeseg algorithm [229].

Individual trees can also be extracted from aerial or satellite images [230] using a Deep Learning approach [231], [232].

In recent years, thanks to technological and data processing advances, numerous studies have proposed different methodologies to carry out more in-depth analyzes in the forestry sector, and a particular objective is to identify the woody part of the tree and distinguish it from the leaf. The aim is to isolate the woody part in order to make more accurate considerations about the forest biomass. Considering the level of detail for this task, terrestrial point clouds are typically used, which can be carried out at closer distances, acquiring details that are difficult

to detect (primarily related to the lower part of the trees) with aerial acquisitions.

The approach proposed by [233] relies on a three-step procedure: at first, intensity values are used to segment the tree point cloud according to threshold values roughly; secondly, a *k*-NN algorithm detects misclassified leaf points according to their higher distance from actual wood points; finally, a further density-based consideration using voxelization was used to evaluate the point density on a larger scale, since points belonging to the same voxel are most likely to exhibit the same properties. Differently, [234] describes the TLSeparation Python library, through which the segmentation process is carried out using four approaches relating to geometric features and structural analysis. LeWoS open-source Matlab tool [235] detects wood and leaf based only on geometric information using a recursive point cloud segmentation and regularization procedure. In [236], a graph-based approach is tested, while [237] tests a Deep Learning network.

4.4 Tree volume and aboveground biomass characterization

Several applications require volumetric estimation of objects within scenarios acquired from point clouds: as an example, glaciers monitoring and change detection analysis [238], [239] or detailed quantifications of payload volume during excavations to optimize the number of earthmoving loaders required on construction sites [240].

In the forestry sector, the volumetric characterization of a tree is necessary to quantify the carbon dioxide stored within the woody part. However, this operation is not trivial; an accurate estimate is required to obtain reliable results. The simplest algorithms are based on the use of convex hulls [241]; however, these significantly overestimate the volume: gaps between branches are not identified and are included in the

calculated volume. More advanced studies have proposed using concave hulls [242] to overcome this problem. In this study, the point cloud is sliced in the direction orthogonal to the tree elevation, and the concave hull is calculated for each slide. Other studies applied a voxel-based approach, and the volumes of the voxels containing at least one point are summed to obtain the total tree volume [243]. However, this approach is highly time-consuming, and the final result depends on the voxel size chosen; moreover, incomplete point clouds underestimate the actual volume. The currently most used approach is based on cylinder fitting of the tree points in order to model the structure of the trunk and branches. This approach, also implemented in the FSCT tool previously described and deepened in several different studies [244], [245], [246], [247], allows for simplifying the structure of the tree skeleton while the missing parts are interpolated. In fact, some parts of the tree may have a point density that is insufficient to describe small branches or noise points. A cylinder fitting-based volumetric assessment of individual trees is also implemented in TreeQSM tool [248] written in a Matlab environment. It reconstructs the tree structure according to Quantitative Structure Models (QSM), i.e., topological ordered cylinder models of trees which cover the complete branching structure from the stem's base up to all tips [249] and that can be used to describe quantitatively its geometric and volumetric attributes. Figure 33 shows an example of the reconstructed tree structure.

The input point cloud must represent only individual trees with a low percentage of noise or leaves points, as well as ground or understory information; furthermore, all the tree parts need to be covered by a sufficiently dense point cloud to be reconstructed. Initial point cloud filtering is performed according to k-NN or point density-based approaches to exclude the cloud's most distant and isolated points. The cylinder fitting procedure relies on different parameters, which need to be optimized according to point density and tree structure; however,

some random elements cause minor variations in the reconstruction of the tree (and, therefore, of the allometric parameters); the algorithm developers, therefore, suggested running the same algorithm multiple times on each tree (at least five times) to estimate an average with about a few per cent of error. Several studies have applied and tested the accuracy of this method [250], [251].



Figure 33: Example of reconstructed tree structure [252].

The volume can also be monitored indirectly using empirical allometric equations expressed as a function of other forest parameters. They have historically been used based on parameters such as DBH and height, which, as already mentioned, are easily measurable in situ. These models need to be calibrated for the tree species under consideration and possibly need to take into account climatic conditions and any variables that can influence forest growth (e.g., forest density). More details about the allometric models will be given in Paragraph 4.5, as they will be used as a validation method in this study.

4.5 Carbon stock assessment

Estimating the CO_2 stored within the woody biomass requires further point cloud processing.

First of all, the volume calculated through one or more techniques described in the previous Paragraph must be calculated and multiplied by the density of the wood. Tabular density values of different types of wood are widely distributed and commonly used [253]; they are calculated by measuring the mass and volume of different samples and comparing these two values. Since wood is hygroscopic (i.e., it can gain or lose moisture depending on the environmental conditions in which it is found and consequently can expand or contract), its weight and volume are determined by specified moisture contents. Usually, this measurement is carried out in a laboratory with zero moisture content.

Once the biomass is calculated, estimating what percentage of CO_2 is stocked is necessary. Various studies consider that the carbon inside trees is equal to 50% of its weight [254], [255] and according to IPCC information; however, more in-depth studies [256] demonstrated that carbon content in woody biomass has an inevitable variability (between 47.6 and 52.5%) which depends on the species under consideration. However, this work only considered American tree species. Therefore, it does not add any knowledge regarding the case studies covered in this work.

4.6 Results validation: methodologies and measures

Results validation is an important procedure that needs to be implemented in segmentation workflows. It is required to define the quality of the information derived from the data processing. Reference outputs can be qualitative or quantitative. In the first case, data are

validated according to visual interpretation; on the contrary, a quantitative assessment requires ground truth data or differently processed data considered ground truth. The next sub-paragraphs present accuracy assessment approaches of different point cloud data processes.

Validating the point cloud segmentation procedure can be highly time-consuming. It requires a segmented reference point cloud; this is obtained by the operator carrying out the same segmentation procedure manually without the help of any automated procedure. For this reason, this operation depends on human sensitivity, which varies for each person; furthermore, for particularly complex scenarios such as those in natural environments, the difficulties to which the operator is subjected are more significant.

One of the metrics to evaluate segmentation results is accuracy; it measures how many times a model made a correct prediction across the entire dataset. However, it does not provide enough information in the case of class imbalance; for this reason, a more complete and widespread method is the *F1 score*. This metric combines two other metrics: *precision* and *recall*. The precision measures how many objects are automatically segmented correctly; recall measures how many objects in the scene are correctly identified by the model (detection rate). Precision is calculated as follows:

$$P = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{12}$$

Equation (13) shows how to calculate the recall:

$$R = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{13}$$

True Positives (TP) refer to objects correctly segmented, while False Positives (FP) are objects segmented even if they do not exist; FP is also

called commission errors. Finally, a False Negative (FN) is a real object that is not segmented and detected (omission error). FN indicate under-segmentation, while FP over-segmentation.

In the forestry sector, omission errors can occur when submitted trees are not identified or when trees very close to each other are considered single; on the contrary, commission errors are common when irregular tree structures or undergrowth fool the segmentation algorithm.

The F1 score combines precision and recall with harmonic mean as indicated in Equation (14):

$$F1\ Score = \frac{2 \times P \times R}{P + R} \tag{14}$$

This metric is commonly used for validating ITD segmentation based on point cloud [257], [258] or images dataset [259]; some studies [260], [261] have evaluated the variation of the F1 Score as a function of forestry parameters. Specifically, sparser forests or tree crops are more straightforward scenarios in which segmentation performs better; the metric's value can quickly decrease as the density increases.

Segmentation at the tree part level into the trunk and leaf can also be done by considering the F1 Score metric [262]. However, given the natural irregularity of such a complex natural element, this procedure may not be carried out accurately and may be extremely time-consuming. For this reason, the result of this segmentation can also be validated through visual interpretation [263], [264]. Although this procedure is equally operator-dependent, it is faster and allows the evaluation of the segmentation as a whole, having good feedback on whether the woody structure has been identified correctly or if some branches have been omitted.

It would be necessary to weigh the objects under examination to validate the results of the volumetric analyses and the biomass estimate. In fact, in the case of trees, the rigorous practice involves weighing them

after harvesting with a destructive approach. However, this cannot always be done and is not a standard or easy practice. In the case of environmental scenarios such as forestry ones, harvesting is often inconsistent with the purpose of monitoring and biodiversity protection. For this reason, the results obtained from remote sensing data processing are compared with those obtained through allometric or volumetric equations. Although these models only provide estimates which can be affected by error, using extraordinarily accurate and case-specific models for the scenarios under consideration can provide excellent comparison value. The principle underlying allometric models is that one part of the organism's growth ratio is proportional to that of another [50]. Numerous studies have been conducted on this topic in the literature [265]. For example, studies have been proposed with a generic validity on specific macro areas at a global or regional level [266]. More in detail, [266] directly relates the AGB [kg] of tropical trees to the DBH [cm], the Height [m] of the tree and the wood-specific gravity ρ [g/cm^3] as follows:

$$AGB_{Chave} = 0.0673 * (\rho * DBH^2 * H)^{0.976} \quad (15)$$

The equation proposed by [55] estimates the aboveground biomass [kg] of a tree through a relationship between the height H [m] and the Crown Diameter CD [m] as follows:

$$AGB_{Jucker} = \left(0.016 + \alpha_g\right) * (H * CD)^{\left(2.013+\beta_g\right)} * * exp[\frac{0.204^2}{2}] \quad (16)$$

α_g and β_g are parameters depending on the type of tree (angiosperm and gymnosperm trees). Since in the case study in question, the trees are gymnosperm, the values considered are the following:

$$\alpha_g = 0.093 \tag{17}$$

$$\beta_g = -0.223 \tag{18}$$

Furthermore, in this study, an equation has also proposed a relationship to estimate the DBH [cm] as a function of Height and Crown Diameter (19) when direct measures are not available:

$$DBH_{Jucker} = 0.557 * (H * CD)^{0.809} * \exp\left[\frac{0.056^2}{2}\right] \tag{19}$$

A recent study [54] developed volumetric models to estimate the volume of major tree species in Finland. These models will be considered to validate results related to case study C, as it will be deepened in Chapter 8. Models were obtained by fitting a volume and taper curve model specific to the typical Finnish boreal forests. The model is highly complete, as it has been calibrated in such a way as to correctly predict the woody volume of trees of all sizes, temperatures and soil types as predictors. The equation in question is the following:

$$AGB = logit^{-1}\left(a + b * DBH + c * h + d * \frac{1}{h} + e * DBH * h + f * \frac{1}{DHB * h} + \varepsilon_{\text{cluster}} + \varepsilon_{\text{plot}}\right) * \frac{\pi * DSH^2 * h}{40} + \varepsilon_{\text{trees}} \tag{20}$$

Where:

- The DSH [cm] is the Diameter at the Stump Height, and it is related to the DBH value according to Equation (21):

$$DSH = \frac{h}{h - 1.3} * DBH \tag{21}$$

- h is the tree height;
- Parameters a-c describe the temperature sum;
- Parameters d-f define the tree species under examination;
- $\varepsilon_{cluster}$, ε_{plot}, and ε_{tree} are respectively, the zero-mean cluster-, plot- and tree-level random effects that follow the standard assumption of mixed-effects models [267].

The results obtained with the different methods are finally compared to verify the reliability of one method rather than another. Since the CO_2 estimate cannot be verified in any way other than directly measuring it in a destructive way for the tree, verifying volumetric estimates plays a crucial role.

Chapter 5

CASE STUDIES AND MAIN WORKFLOW

In this chapter, the in-depth case studies in this thesis are introduced and described; furthermore, the data processing workflow is analyzed, emphasizing the objectives and main contributions.

5.1 Case study A: Slovenian Dinaric forest

The first proposed case study is localized in Southwest Slovenia, more precisely in the Dinaric Mountains area. In this area, the Dinaric forest is predominant: it is a European beech forest mainly made up of Silver fir (*Abies alba Mill.*), Norway spruce (*Picea abies Karst.*), and European beech (*Fagus sylvatica L.*) tree species. Airborne LiDAR acquisitions and an in-situ measurements campaign were conducted in November 2013 within a Life+ ManFor C.BD. project. This case study

was deepened based on a collaboration with the University of Ljubljana, particularly with Professor M. Kobal.

Figure 34 shows the study area location. The investigated area extends for about 4000 m^2 (35.7 meters radius). The LiDAR acquisition covered a more significant extension of the forest, but for the aim of this study, only a squared area (6400 m^2) circumscribed to the circular plot was considered. Within the in situ circular plot, 78 trees were manually detected, of which 9 were European beech and 63 were Silver fir. Smaller trees (DBH ≤10 cm) were not measured; concerning measured trees, information was acquired not only about their position (x and y coordinates) but also the DBH was measured, and the social status of each of them was annotated.

Additionally, the height of the trees was measured with the purpose of future local height curve development. Table 3 summarizes data acquired during the in-situ campaign. When considering the extension of the square area circumscribed to the circular area, the number of trees is 87.

Table 3: Case study A - Trees parameters summary.

Tree parameters	Features	Value
DBH [cm]	Minimum	10.2 cm
	Maximum	65.9 cm
	Mean	38.6 cm
Tree height [m]	Minimum	7.4 m
	Maximum	33.5 m
	Mean	24.5 m
Social status (%)	Dominant and	73,6 % (53 trees)
	Suppressed trees	26.4 % (19 trees)
Tree species composition (%)	Silver fir	87.5% (63 trees)
	European beech	12.5% (9 trees)

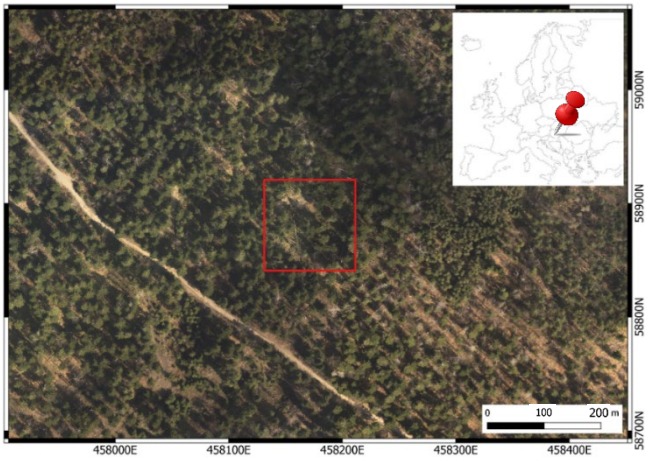
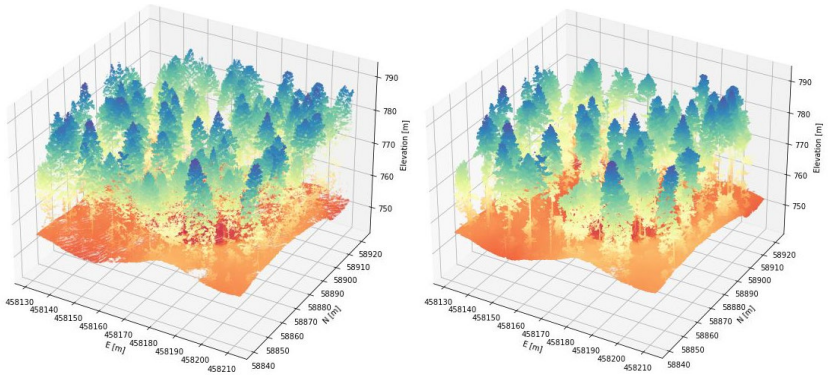


Figure 34: Case study A - Study area. EPSG: 3912 [268].

During the winter, in January-February 2014, an extreme ice storm occurred in this area, destroying more than half a million hectares of forests across Slovenia, including the study area [269]. Following the natural hazard, a new aerial LiDAR survey was carried out; a post-event in situ survey was also conducted, during which uprooted and damaged trees were identified. In the extent of the square area, 57 trees survived. In both measurements, a RIEGL LiteMapper-5600 system was equipped on an airborne, which kept a flying altitude of 700 m above ground level; the point clouds have an average point cloud density of about 150 pts/m^2 and 250 pt/m^2 in the 2013 scan and 2014 scan respectively. A RIEGL LMS-Q560 laser scanner, an IGI AEROcontrol-IId direct georeferencing system, and an IGI DigiCAM digital camera comprise the LiteMapper-5600 system. Figure 35 shows the point cloud data acquired before (a) and after (b) the ice storm. GNSS and INS data were acquired by the integrated platform equipped on the airborne.

The point cloud was georeferenced using the information collected by the integrated GNSS/INS system equipped on the airborne. The reference system chosen is the Slovenian National Grid (EPSG:3912).



(**a**) (**b**)

Figure 35: Case study A - Point cloud (a) pre-ice storm acquisition (November 2013); (b) post-ice storm acquisition (April 2014). Color scale according to elevation. EPSG: 3912.

5.2 Case study B: Italian Mompantero forest

The second study area is located in a highly sloped coniferous forest in the Susa Valley, in the Italian Western Alps, precisely in the municipalities of Mompantero and Bussoleno (Turin).

In October 2017, this area was affected by the most severe wildfires that occurred in the region during 2017. Overall, this wildfire burned almost 4000 hectares, and countless European Beeches (*Fagus sylvatica L.*) and Scots Pines (*Pinus sylvestris L.*) trees were destroyed [270]. The extension of the study area is approximately 70 hectares, and the tree vegetation consists almost solely of dense, even-aged *P. Sylvestris* stands. As Figure 36 shows, the survey was conducted in order to cover a portion of the forest with tree densities that are very different; as a matter of fact, the northern part of the study area was thinned out just before the forest fire [271].

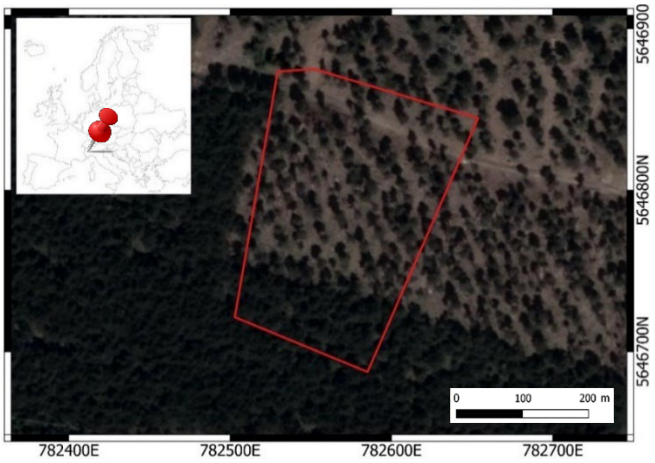


Figure 36: Case study B - Study area. EPSG: 32632 [271].

The survey was conducted using a terrestrial integrated sensors approach. A Mobile Mapping System (MMS) survey was conducted with the KAARTA Stencil 2 sensor (Figure 21); this platform combines a portable laser scanner with a video camera to automatically generate and register 3D point clouds. This sensor is low-cost and very manageable: these aspects make it very versatile, and multiple complex scenarios can be acquired. KAARTA Stencil 2 laser scanner has a maximum range of 100 meters, a horizontal and vertical field of view of 360° and 30°, respectively, and an accuracy of ± 30mm; the video camera acquires images with 640x360 pixels resolution and 50 Hz frame rate. The IMU system is required to produce a real-time 3D map of the scenario. At the same time, the SLAM algorithm leverages the acquired images to solve the localization problem, optimize the estimated trajectory, and produce a 3D point cloud of the examined area. SLAM algorithms solve the computational problem of the three-dimensional reconstruction of an environment, simultaneously keeping track of the position of the acquisition instrument within it. In fact, the localization problem consists of estimating the position and orientation of the agent

in time (2D or 3D) given the map. In contrast, the mapping problem involves building a map of the environment where the agent moves (2D or 3D) and notes the position. This approach is independent of external observations; it can be installed in portable instruments that can be easily used in challenging environments, producing 3D data quickly. From an analytical and algorithmic point of view, SLAM is approached as an estimation problem in which different acquisition and positioning systems are combined in a single multi-sensor system in order to benefit from the advantages or compensate for the disadvantages of each individual sensor. Although the problem presents itself as a chicken-and-egg problem, several algorithms, such as Particle Filter, Extended Kalman Filter, Covariance Intersection, and GraphSLAM, solve it approximately.

The survey was planned with a closed acquisition path to ensure a comprehensive area understanding. The acquisition started from the northern part of the area and continued downstream with a winding path, reaching the southern part; then, the path headed north again, intersecting the outward route multiple times until reaching the starting point.

The acquisition took approximately 13 minutes, covering a trajectory of around 550 meters, resulting in a point cloud of 143 million points with a point cloud density of approximately 500 pts/m^2 (Figure 38a). Specific settings for forest scenarios were set, including voxel size (0.4 m), point cloud resolution in the map file, point cloud resolution for scan matching and display (cornerVoxelSize equal to 0.4m, surfVoxelSize equal to 0.8m, surroundingVoxelSize equal to 0.6m), minimum point-to-point distance for mapping (1 m). In contrast, no restrictions on the planarity of motion were set.

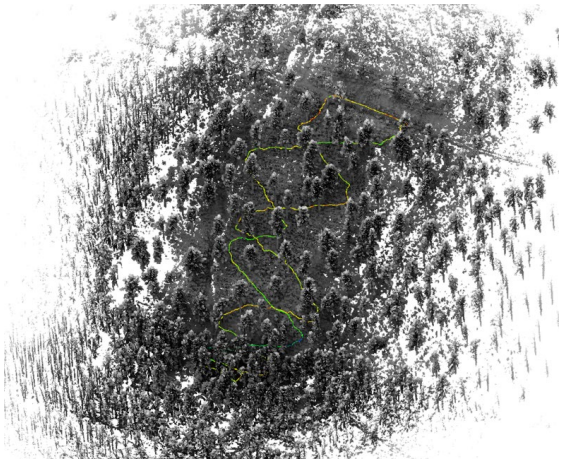
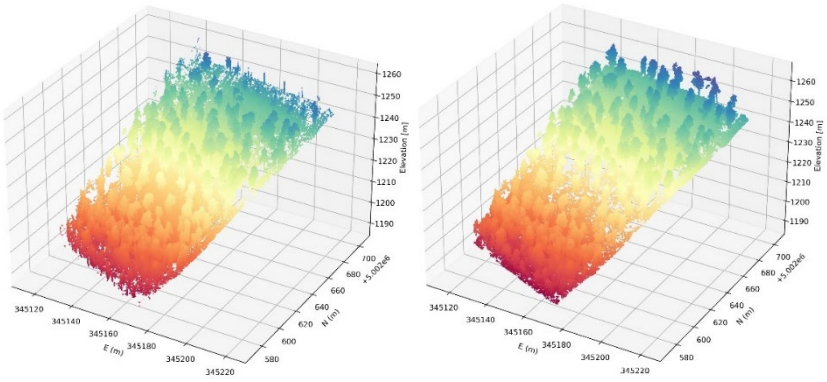


Figure 37: Case study B – KAARTA Stencil 2 trajectory and point cloud. EPSG: 32632.



(**a**) (**b**)

Figure 38: Case study B - Point cloud (a) MMS (KAARTA Stencil 2); (b) TLS (RIEGL VZ-400i). Color scale according to elevation. EPSG: 32632.

The point cloud registration was optimized through two procedures: the first reprocesses the alignment of the scans, simulating a lower acquisition speed, while the second optimizes the scan through the Loop Closure Tool, which forces the overlap of the starting and ending points.

The first procedure took approximately 40 minutes, while the Loop Closure took 30 minutes.

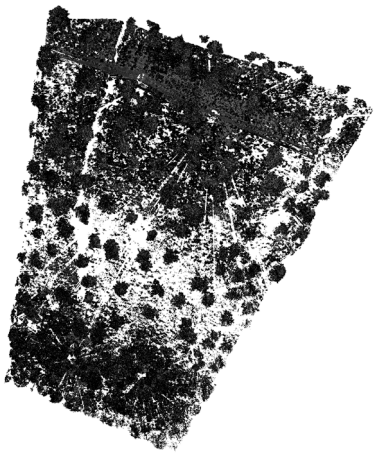


Figure 39: Case study B - RIEGL VZ-400i point cloud. EPSG: 32632.

A classic topographic survey conducted through the use of total stations and GNSS antennas was considered to measure the position of 13 reflective markers; these were used to georeference the MLS point cloud in the WGS84/UTM zone 32N (EPSG:32632) reference system.

Additionally, a TLS survey was conducted with a RIEGL VZ-400i LiDAR. 5 scans were needed to cover the entire extent of the study area and were registered using 13 reflective markers (Figure 39). The LiDAR was set with a resolution of one point every 6 millimeters at a distance of 10 meters; acquisition took approximately 4 hours, resulting in about 1.5 billion points (Figure 38b).

The previously described reflective markers were also used to fulfil the TLS scans' registration procedure and the entire dataset's georeferencing.

Table 4: Case study B - Trees parameters summary.

Tree parameters	Features	Value
Number of trees		242
DBH [cm]	Minimum	5,2 cm
	Maximum	45,7 cm
	Mean	22,4 cm
Height (m)	Minimum	2,3 m
	Maximum	24,1 m
	Mean	9,02 m
Tree species composition	European Beeches Scots Pines	

In this case study, no ground truth data were collected; as will be explored further in the Chapter, the purpose of this survey was initially to monitor the area after the fire. Table 4 summarizes the main tree parameter values obtained from manual analysis on the point cloud.

5.3 Case study C: Finnish Hyytiälä forest

The last case study proposed is located in a 61-year-old (2023) boreal forest in southern Finland. More precisely, it covers a portion of the Hyytiälä forest field station, a recognized forest educational and research center focused on long-term boreal forest zone monitoring (Figure 40). Access to this data and the research itself result from a collaboration with the Finnish Geospatial Research Institute (FGI), where I also had the pleasure of studying abroad during my doctoral studies and collaborating with Professor Antero Kukko and his research group.

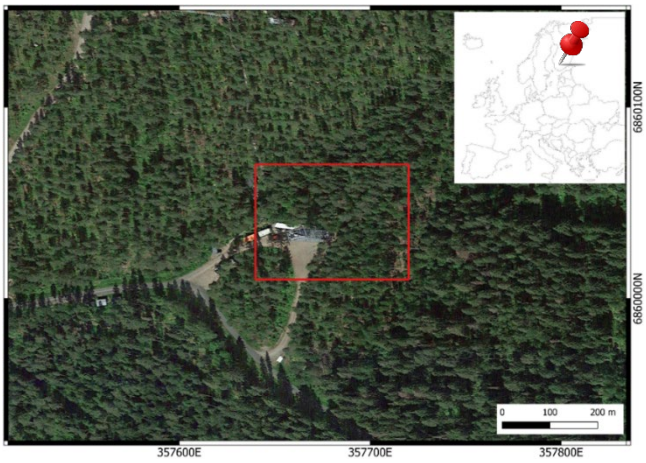


Figure 40: Case study C - Study area. EPSG: 3067 [272].

The forest is characterized by a predominance of coniferous boreal species, including Scots pine (*Pinus sylvestris L.*) and Norway spruce (*Picea abies H. Karst*), as well as deciduous trees, such as silver birch (*Betula pentula Roth*). It is a monitoring area equipped with disparate sensors to acquire various data types, including long-term field measurement sites, soil and atmospheric sampling, and continuous LiDAR time-series data collection. On top of a 35-meter-high tower, a RIEGL VZ-2000i is permanently equipped; this setup is named a phenological station and it is commonly referred to as LiPhe [273]. The laser scanner is tilted 60 degrees, and it provides an oblique perspective of the forest on a portion of almost 4.5 hectares, with an above-canopy view of trees, enabling a high spatial-resolution reconstruction of the study area. Setting parameters were set in order to acquire with a resolution of 1 point every one centimeter at a 100 m range, resulting in a very high-dense point cloud. In addition to the position of the points, point return number, intensity (expressed in dB), reflectance (dB) and the return pulse deviation were collected by the LiDAR sensor. The sensor is connected to a separate computer, which fulfils the role of

storage and data transmitter; data are automatically collected per hour, internally stored, and transmitted daily to a network storage. Moreover, local weather parameters are stored and linked to scans; subsequently, only the scans acquired during the best weather conditions are considered for further analysis (one per month) [273]. Each point cloud has a point density of approximately 114000 pts/m^2.



Figure 41: Case study C - LiPhe setup [273].

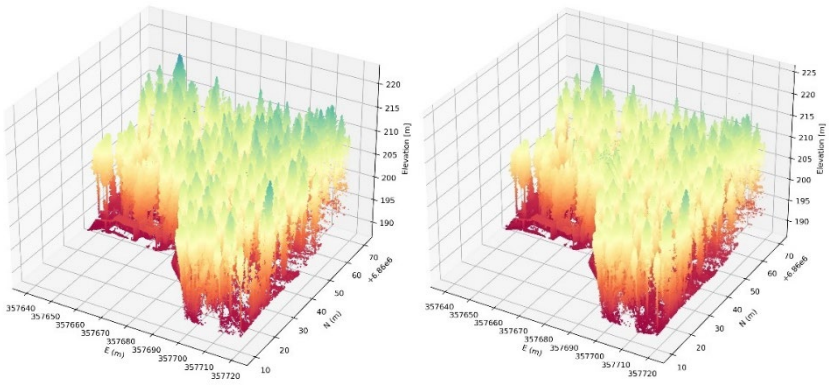
Table 5: Case study C - Trees parameters summary.

Tree parameters	**Features**	**Value**
Number of trees		184
DBH [cm]	Minimum	15,2 cm
	Maximum	29,4 cm
	Mean	22,1 cm
Height (m)	Minimum	14,4 m
	Maximum	25,1 m
	Mean	18,9 m
Tree species composition	Scots pine	
	Norway spruce	
	Silver birch	

For the purposes of the analyzes conducted in this thesis work, two different point clouds were considered (Figure 42): the first was acquired during April 2020, and the second during the same month of the following year (April 2021). The study area has a reduced extension compared to the entire dataset, extending for approximately 0.5 hectares.

In situ measurement campaigns were organized in the research forest field in relation to other research projects [274], [275]; the position of 184 trees was recorded within the study area, while no additional information about tree parameters was available. The coordinates of five ground control points were also collected for georeferencing purposes.

Table 5 summarizes the tree parameters obtained by manually processing the point cloud dataset.



(**a**) (**b**)

Figure 42: Case study C - Point cloud (a) April 2020; (b) April 2021. Color scale according to elevation. EPSG: 3067.

Oblique point clouds require rectification and georeferencing; since the scanning hood is tilted 60 degrees towards the ground, the original point clouds are consequently also tilted with respect to the ground. The point clouds were rectified with respect to a global reference system through a 3D passive rotation ($R_{\omega\varphi\kappa}$) in which a right-hand system was

defined with origin at the scan and z to up. The origin at the scan is located 30 m from the ground (tower), providing negative z values. The rotation parameters were ω=0, φ= 60 degrees and κ= 90 degrees. Subsequently, the georeferencing was based on a Helmert 3D transformation: a geometric transformation consisting of a rototranslation and scale variation in the plane to apply a change of reference system. The general equation of the 3D Helmert transformation can be expressed as follows:

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} sR_x & -sR_z & sR_y & T_x \\ sR_z & sR_x & -sR_x & T_y \\ -sR_y & sR_x & sR_x & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (22)$$

Where:

- X', Y', Z' are the transformed coordinates;
- X, Y, Z are the original coordinates;
- R_x, R_y, R_z are the rotation matrices around the x, y and z axes, respectively.
- s is the scale factor;
- T_x, T_y, T_z are the translation vectors.

The transformation parameters were computed in the least square adjustment based on the coordinates of five ground control points collected in the test area through a GNSS receiver in RTK mode (Finnish EPSG:3067 – ETRS89/ TM35FIN reference system).

No point cloud registration is necessary as the acquisitions are done through individual scans.

5.4 Processing workflow

The main objective of this study is to develop a semi-automated procedure for the complete processing of LiDAR point clouds acquired in a forest environment with the final aim of estimating the carbon captured and stocked within the woody biomass. A detailed complete workflow was developed (schematic in Figure 43) and applied in each case study. The crucial points of the processing are (i) the Individual Tree Detection, (ii) the separation between wood and leaf, (iii) the volume estimation and, therefore, the biomass calculation, and finally, (iv) the stocked CO_2 assessment. Each step of the workflow was carried out using different approaches and algorithms in order to (i) obtain more accurate estimates and (ii) evaluate any differences between datasets acquired with different sensors. Furthermore, a validation procedure was implemented for each step before moving on to the subsequent phases.

Specifying that the workflow mentioned above was optimized iteratively based on the objectives to be achieved and the case studies considered is essential. In fact, the case studies were conducted, explored and analyzed in a parallel approach. For this reason, as described in the following paragraphs, specific analyses relating to some of the main objectives were conducted only in specific case studies. Once their effectiveness had been ascertained and the fine-tuning procedure resulted in optimal results, these approaches were also consolidated for the remaining case studies without verifying the improvement they entail. Specifically, they concern (i) how to handle single-tree segmentation errors caused by over-segmentation during the ITD procedure (deepened in case study A) and (ii) the features taken into account to separate wood and leaf parts (more details are described in case study C).

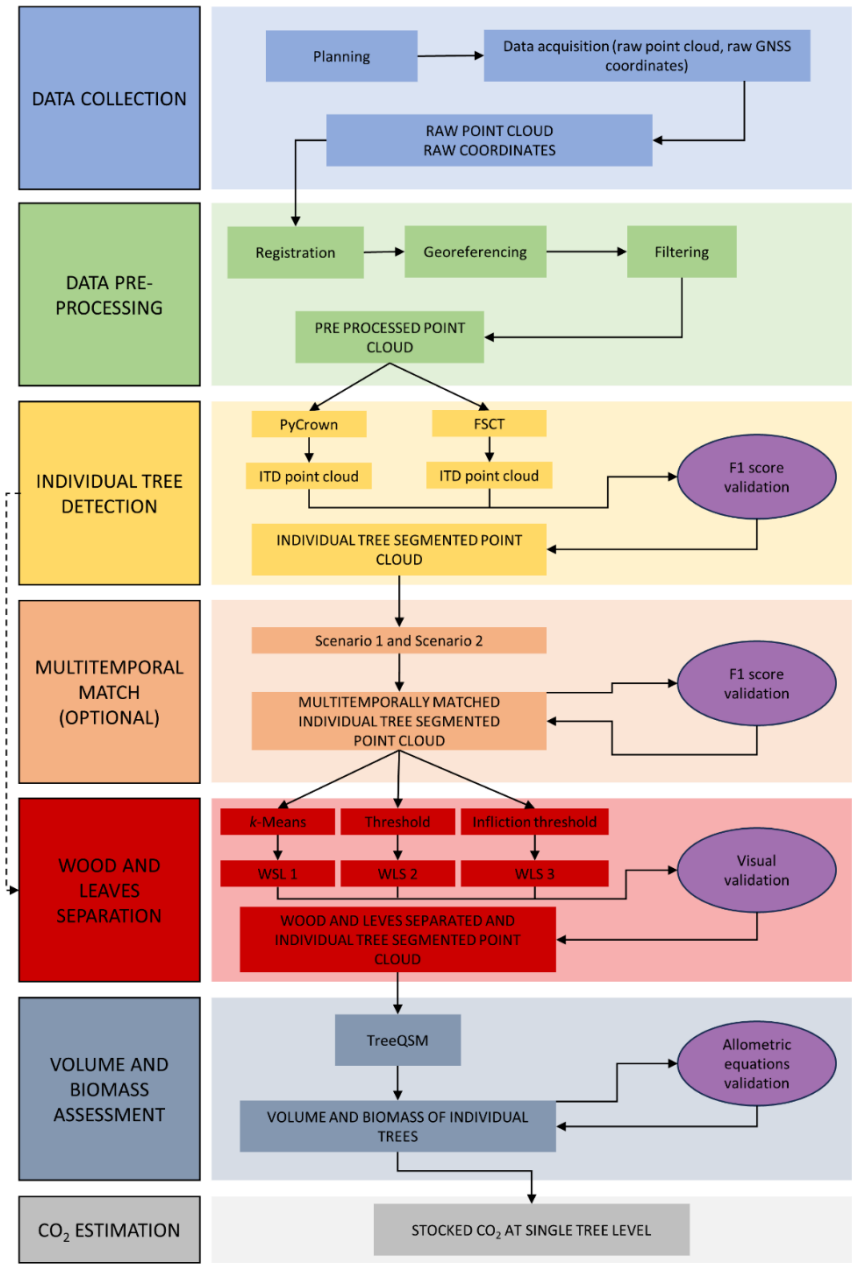


Figure 43: Workflow of this study.

5.4.1 Data collection and pre-processing

Since data collection is different depending on the type of planned acquisition, it has been explored in depth in the previous paragraphs. The registration and georeferencing phases were also discussed. Below, the filtering steps are explored in more detail.

The point cloud relating to case study A was not subsampled in order to preserve the information collected with a low point cloud density. In contrast, case study datasets B and C were subsampled to speed up subsequent data processing. Concerning the Italian study area (case study B), an additional motivation was to make the density of the point clouds acquired with different instrumentation comparable. To do this, the MLS point cloud was not modified, while the TLS one was reduced by discarding approximately four-fifths of the points. The very high density of the case study C dataset required subsampling of the point cloud. Therefore, the full-resolution point clouds were resampled based on a voxelization process, in which the closest point to the center of the voxels was kept. A voxel with a 5 cm size was considered, and it reduced the point cloud density to approximately 18000 pts/m^2. The latter dataset also required filtering noise relating to the tower's structure and fixing cables. This operation was conducted manually.

In this study, ground point detection can be included among the pre-processing operations. Although this operation can be considered a first data processing, it is treated in this subgroup of preliminary operations as this thesis aims not to delve into further details. The ground point detection was performed with the CSF algorithm. The parameters to be set are (i) the cloth resolution cloth, which refers to the grid size of cloth used to cover the terrain, (ii) the number of maximum iterations, and (iii) a classification threshold to classify the point clouds into ground and non-ground parts based on the distances between points and the simulated terrain, and finally (4) the cloth rigidness. The cloth resolution

was equal to 0.1 m, while the maximum iterations and the classification threshold were set as default values (500 and 0.5, respectively). The cloth rigidness was established according to the terrain conformations of the three case studies. A lower rigidness (equal to 2) was used in case studies A and B, while a higher value (equal to 3) was set in case study C. In fact, in the first two study areas, the conformation of the terrain is steep or slightly uneven, while the Finnish study area is flat.

5.4.2 Individual Tree Detection

Regarding the Individual Tree Detection, among the methodologies proposed in the literature, it was decided to compare the results obtained using two open-source Python libraries described in Paragraph 4.3.1: PyCrown [220] and FSCT [222]. PyCrown algorithm has been modified to improve segmentation performance, specifically reducing over-segmentation errors. In fact, as previously described, the single tree segmentation can be affected by (i) simple omission (the tree is entirely automatically undetected), (ii) under-segmentation (two or more trees are segmented as a single tree), and (iii) over-segmentation (a single tree is sub-segmented into two or more trees). The following procedure was implemented to limit over-segmentation. For each treetop identified through a local maxima-based approach on the CHM, a radius spatial search (1 meter) was added; in the event that the search produces a positive result and detects a different treetop in the search area, these two are merged, as are the two respective crowns. The usefulness of this improvement was tested in case study A. More details about the improvements in the code can be found in Appendix A, while Figure 45 schematizes the main implemented steps. As regards the parameters to be set when segmenting with PyCrown, there are four: three of them concern the elevation of the pixels close to the treetop, while one concerns the distance from it (crown radius). The first parameter defines the elevation of neighbouring pixels with respect to the treetop (default

parameter 0.7*seed height); the second parameter ensures that the height of the neighbouring pixel is greater than the mean of the current height of the crown (factor equal to 0.55); the third establishes that neighbouring pixels do not exceed a specific elevation (default value equal to 1.05*treetop height); finally the maximum radius of the crown is set by default to 10 meters.

Furthermore, an additional parameter can be set about the minimum height of a tree; therefore, it is possible to exclude the smaller trees from the analysis, which are often not even detected when in situ investigations are carried out. In this study, various attempts were made to vary the value attributed to the parameters to obtain the best possible segmentation. In the end, in all three case studies analyzed, the first parameter was decreased to a value equal to 0.45; the second was set equal to 0.50; the third was left unchanged (value equal to 1.05); finally, the additional parameter regarding the minimum height of the trees was set to 2 meters. The maximum crown radius was not fixed, but it was modified in each case study. As such, in the first case study, a value of 7 meters was set, while this was reduced in case B (maximum crown radius equal to 5 meters) and increased in case C (15 meters) according to the forest characteristics of each area. Figure 44 summarises the main workflow adopted in the PyCrown algorithm.

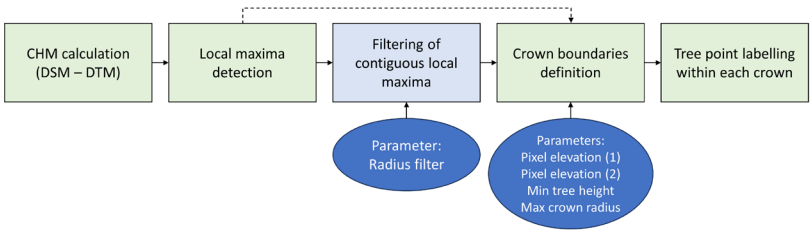


Figure 44: PyCrown main workflow summary: in green is the original workflow, and in light blue is the improvement proposed in this study.

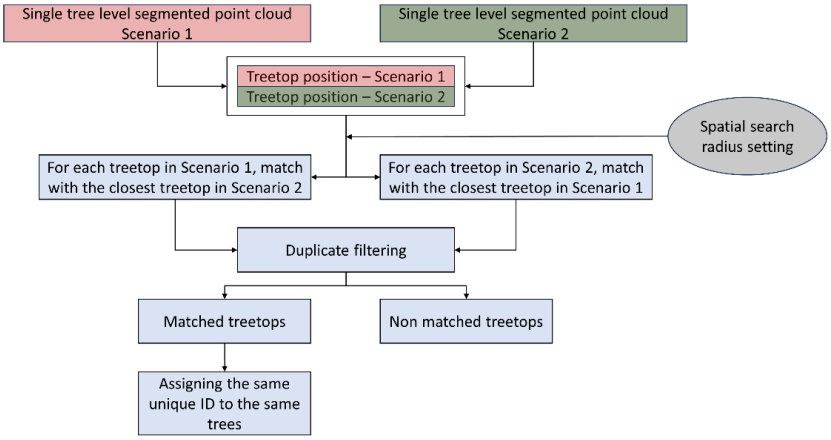


Figure 45: Specific details about the procedure implemented in the PyCrown workflow.

On the other hand, FSCT is a more complete tool that performs single tree-level segmentation and contextually allows the extraction of forest parameters such as DBH and volume (Figure 46). However, only the segmentation procedure was taken into account in this study. In fact, using this tool, it is impossible to carry out intermediate control and validation steps, which are crucial for the purposes of the analyses conducted in this thesis; furthermore, the entire procedure is costly in terms of the time and computational power required. For these reasons, no parameters have been changed, and the default ones are used.

The segmentation results obtained using PyCrown and FSCT were validated by comparing them with those obtained through manual segmentation. Validation was carried out to the best of interpretation ability using all available information (point cloud, CHM, in situ data when available), and it relies on the F1 score metric.

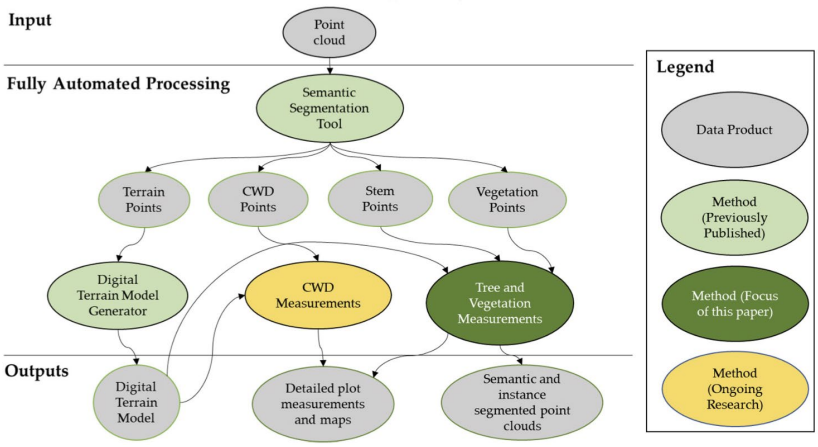


Figure 46: FSCT complete workflow [222].

5.4.3 Multi-temporal detection of individual trees

In case multi-temporal acquisitions were available (case studies A and C), an algorithm capable of identifying the same tree in different acquisitions was developed. The operating principle is based on the concept that the top of a tree does not substantially move from one acquisition to another; for each treetop identified in the point cloud at time t_0, a spatial search is carried out; if a treetop identified in the point cloud acquired at time instant t_1 is located within the search area, these trees are assigned the same tree ID.

As for the segmentation procedure, a validation based on a comparison with a manual multi-temporal point cloud match was performed, and the F1 score parameter was considered. Both point clouds were processed manually, and the same trees at different time instants were matched and then compared to the results of the automatic procedure.

5.4.4 Wood and leaf separation

Three approaches have been proposed and tested to identify the points relating to the woody part (trunks and branches) and leaf (Table 6). Specifying that this analysis was carried out only on trees that were segmented correctly is essential. In fact, over- or under-segmented trees were not considered either because this can affect the quality of the wood and leaf segmentation or because the final objective of estimating the volume of each individual tree cannot be correctly achieved in those cases.

Table 6: Summary of the proposed approaches for identifying tree woody and foliar parts.

Approach	Algorithms	Validation
Approach 1	*k*-Means	Visual interpretation
Approach 2	Threshold	
Approach 3	Infliction threshold	

Table 7 summarizes the features considered in each case study and approach.

Table 7: Features for identifying the woody and foliar parts of the trees in each case study and for the proposed approaches.

Case Study	Approach 1	Approach 2	Approach 3
Case Study A	x, y, z, density		Anisotropy, linearity, verticality, sphericity, curvature, PCA1
Case Study B			
Case Study C	x, y, z, density, intensity, reflectance		

The first methodology is based on the use of the *k*-Means segmentation algorithm. In each scenario, a single tree was used to train the algorithm, considering both geometric and radiometric features; subsequently, this classification was applied to the remaining trees. The

second approach instead distinguishes the two classes using a threshold value. This value was calculated uniquely for each tree and was defined as the median of the best-fitting skewed distribution of the point cloud selected features. The Logistic distribution, the Normal distribution, and the Laplacian distribution were considered, and the best one was chosen according to minimal mean square errors. The third method relies on the unsupervised infliction threshold method proposed by [276]. It employs six salient geometric features (local anisotropy, curvature, linearity, PCA1, verticality, and sphericity), which have been assessed to be robust in previous studies [233], [235]. This approach analyzes the distribution of each geometric attribute for individual trees according to Gaussian Mixture Model (GMM) fitting and curve fitting techniques. It defines threshold values as the infliction points of the fitted distributions. It initially performs the geometric feature extraction using eigenvalues and normalized eigenvalues. Subsequently, GMM were fitted to curvature, anisotropy, PCA1 and verticality features. Low curvature, anisotropy and verticality values define non-leaf components, while opposite values describe the wood part of the tree. Furthermore, a GMM with two components in 1D for each selected feature was fit to determine the centroids of the clusters to be used as threshold constraints. The next step is needed to extract the infliction points of the curve-fitted distribution of the considered attributes according to the sign change of the second derivative.

Regarding linearity and sphericity features, infliction values were detected using the previously calculated centroids to narrow down the range of values. Finally, the wood and leaf separation was performed based on these thresholds. Additionally, the DBSCAN algorithm was implemented to ensure the connectivity of the separated non-leaf component; moreover, isolated and clustered noise were eliminated to refine the results further.

The validation of the results obtained from the three proposed wood and leaf separation methods was carried out on visual interpretation. In fact, given the large amount of time required to segment leaf manually and the strong dependence on the skill and sensitivity of the operator, as well as on the quality of the dataset, it was decided to define the best-performing method by considering which of these would return the most continuous and regular woody structure possible, and which would recognize the most significant number of points relative to only to the larger trunks, but also to the thinner branches in the upper part of the tree.

5.4.5 Volume and AGB characterization and stocked CO_2 assessment

The volumetric estimate was carried out using the Quantitative Structure Model method implemented in the TreeQSM Matlab tool. Among all the outputs extracted, for this study, only volume calculated as the volume of cylinders and truncated cones that fit the woody part of the tree was considered. In fact, the ITD procedure has not been implemented in the workflow, although it performs automatic filtering for the separation of leaf and trunk. However, this is not valid if the tree has very thick foliage. The default parameters of the tool were considered for the analyses. Tabular density values from the literature were used to convert tree volume into aboveground biomass. Different values were considered depending on the geographical location of the case studies and the tree species.

Validation was carried out using allometric or volumetric equations. The parameters on which the equation depends (usually the tree height H and the DBH) were measured in situ. When no ground truth data is available, these parameters were obtained through manual processing of the point cloud: as regards the DBH, this was calculated by measuring the diameter of the portion of the point cloud between 1.20 m and 1.40

m; the height of the tree was instead obtained by considering the elevation of the highest point of the canopy in the point cloud. Different allometric equations have been considered, depending on the locality and tree species under consideration, and they will be described more carefully in the related chapters.

Finally, regarding the estimate of the CO_2 stored in each tree, the multiplication factor equal to 0.5 was applied to the aboveground biomass. No distinction was made based on species, as, to the best of my knowledge, no study has currently developed more accurate species-specific estimates. The results obtained at the single tree level were also extended to the plot area level through linear proportions. In this way it is possible to provide more accurate estimates about the reference stand; these values are in fact more significant for a 360-degree understanding of the phenomenon under analysis

Regarding data validation, the only non-invasive approach that could be implemented was to compare the results obtained with those calculated starting from the reference allometric equation used to validate the volumetric and biomass estimates.

5.4.6 Case-specific tasks

In addition to common objectives, each case study was also conducted in order to find answers to specific questions. These will be explored in depth in the relevant chapters; however, they are introduced below. In case study A, the multi-temporality of the dataset is exploited to quantify the damage to which the detected portion of the forest was subjected. The multi-temporality of the dataset also characterizes case study C; however, in this case, this occurred under undisturbed conditions, and the additional objectives are (i) to evaluate forest growth without external interventions (natural or human-caused) and (ii) to evaluate the usefulness of the oblique point cloud acquisition method for monitoring individual trees. Finally, case study B, characterized by a

multi-sensor approach, was developed to evaluate the same workflow using two different terrestrial acquisition techniques in order to evaluate, in case there were differences, which of the two was more suitable for the purposes of estimation of forestry parameters and stocked carbon dioxide.

Table 8: Summary of the characteristics of the case studies analyzed.

Case study	**Data type**	**Multitemporality**	**Multi-sensor**	**Specific goals**
A	ALS point cloud	Yes (pre and post ice-break)	No	Quantifying the damage of a disastrous event
B	TLS and MLS point cloud	No	Yes	Comparison between two different acquisition methodologies
C	Oblique point cloud	Yes (undisturbed conditions)	No	Evaluation of forest growth in undisturbed conditions through an innovative acquisition methodology

Table 8 summarizes the main characteristics of the datasets relating to the three case studies proposed in this study to clarify their understanding. Finally, Table 9 summarizes all the parameters of the different algorithms set for each case study.

Table 9: Parameters set for each case study.

Parameter	**Case Study**		
	A	**B**	**C**

Cloth rigidness - CSF		2	2	3
PyCrown parameters	1st height parameter	7	0.45	15
	2nd height parameter		0.5	
	3rd height parameter		1.05	
	Maximum crown radius		5	
	Minimum height tree		2	
ρ [g/cm^3]		0.45	0.8	0.51
% CO_2 in AGB			0.5	

Chapter 6

CASE STUDY A: RESULTS AND DISCUSSION

The first case study on which the analyzes were focused is located in a Slovenian Dinaric forest (Paragraph 5.1). Although the point cloud density is not extremely high when compared with that obtainable with more recent and developed sensors (about 150 points/m^2 in the 2013 acquisition and 250 points/m^2 in the following dataset), the peculiarity consists of the multi-temporality of the data in conditions of external disturbance. In fact, the two acquisitions were carried out before and after an ice storm that destroyed and damaged several hectares of forest. Moreover, ground truth data related to each tree's position, height, and DBH were collected before the hazard's occurrence. At the same time, the tree's health condition was also registered after the ice storm. Based on these singularities, the questions that sparked interest in this case study are the following:

- Is it possible to precisely quantify (i.e., at the level of a single tree) the damage to which a forest (or a portion of it) is subject using an airborne LiDAR dataset?
- Is it possible to improve the segmentation quality by reducing errors caused by over-segmentation?

6.1 General overview

In order to deepen the knowledge about the aforementioned topics, the processing workflow described in Paragraph 5.4 was applied to both the point clouds. Additionally, a change detection analysis was conducted in order (i) to identify the uprooted trees, (ii) to locate the survived ones, and (iii) to evaluate the total variation in volumetric, biomass, and stocked CO_2 terms. An estimate was also made of the loss of capacity to store carbon, both (i) following the uprooting of the trees and (ii) as a consequence of the damage to the surviving ones.

Figure 47 shows the terrain points automatically detected with the CSF algorithm. The territory is mainly flat; however, a slight depression is detected in the central part.

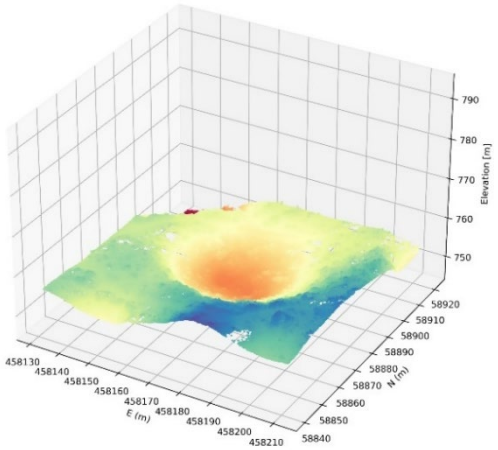


Figure 47: Ground points of the study area A. EPSG:3912.

After selecting the ITD algorithm that produces the best results, the same trees in the two acquisitions were identified, and the results were validated through the F1 score. In this scenario, and only for pre-event acquisitions, the improvement brought about by modifying the PyCrown algorithm to reduce over-segmentation (Figure 44) was quantified (Table 11). The F1 score was also calculated to evaluate the quality of identifying the same tree in the two different acquisitions.

Identification of trunk and leaf components was subsequently performed. Finally, the Quantitative Structure Model was applied for the volumetric estimate, from which the biomass and the carbon dioxide stocked were estimated. Biomass was additionally calculated by considering allometric equations (15) and (16) previously described in Paragraph 4.6; automatically extracted values of H and CD were considered, while the DBH was estimated according to Equation (19). Considering the forest composition in terms of species, the density value of Silver fir trees was used in this study; it was considered equal to 0.45 g/cm^3 according to a specific study conducted by [277] concerning Slovenian forests. The damage to which the portion of the forest was subjected was quantified (i) in absolute terms and (ii) considering only the damage suffered by the trees which, however, survived (i.e. not uprooted trees).

Reference aboveground biomass values were obtained with Equation (15), in which DBH and H considered values were measured in situ or extracted by manual point cloud processing. Equation (16) was not used as a reference, as ground truth crown diameter values were not available; besides, the tree crown is the part of the tree that varies most due to the damage of the ice storm, and using this parameter as a reference would be of little significance. In addition, DBH values assessed with Equation (19) were compared with ground truth data. Table 10 outlines the different tests undergone.

In the published work related to this case study [268], only a small portion of these analyzes has been published; besides, the workflow applied in that study has improved. For this reason, slight differences in the validation procedure of forest parameters can be observed.

Table 10: Summary of the biomass and DBH estimation and validation procedure - Case study A.

Tree parameter	Estimates	Validation values
AGB	TreeQSM approach	Equation (15) with in-situ measured H and DBH variables
	Equation (16) with derived H and CD variables	
	Equation (15) with derived H and DBH variables	
DBH	Equation (19) with derived H and DBH variables	Ground truth data

Additionally, to evaluate the reliability of allometric equations, an error propagation analysis was conducted to quantify the error in the tree parameter (DBH, AGB) assessment as an effect of incorrect input values of the dependent variables. The studies were carried out by introducing an error bias in the estimated crown diameter in Equations (16) and (19) and in the DBH in Equation (15). It was decided to consider only the error of these variables and to ignore any errors in estimating the height of the trees, assuming that this was obtained with sufficient precision by the local maxima algorithm. The error was then evaluated in percentage terms in the AGB estimate of Equation (19) due to an error in the prediction of the CD of (i) 0.1 m, (ii) 0.2 m, (iii) 0.5 m, and (iv) 1 m. The same error bias was introduced to evaluate the allometric Equation (16). Finally, similar considerations were made in the allometric equation for the AGB estimate (15), considering a systematic error in the estimate of the DBH of (i) 1 cm, (ii) 5 cm, and (iii) 10 cm.

6.2 Data processing

6.2.1 ITD

The Individual Tree Detection performed through the FSCT tool yielded no results. Unfortunately, during the processing phase, the error shown in Figure 47 is returned. The “ValueError: need at least one array to concatenate” error is due to the low density of points. This result was expected since this algorithm identifies trees with a procedure that analyzes the point cloud from the lower to the upper part of the trees; however, (i) the aerial approach of point cloud acquisition and (ii) the low density of points mainly located in the lower part of the trees do not allow the automatic identification of the trunks.

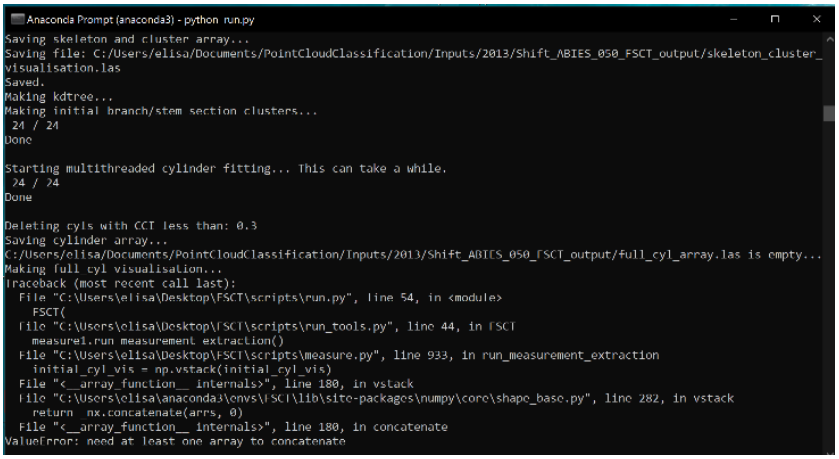
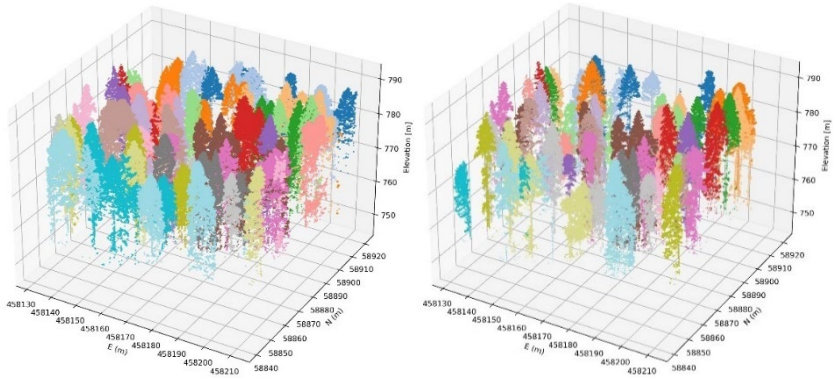


Figure 48: Error resulted during the ITD procedure - Case Study A.

The segmentation carried out through the PyCrown algorithm, on the other hand, produced positive results (Figure 49a-b). Individual trees are plotted using random colors.



(**a**) (**b**)

Figure 49: Segmented point cloud at single tree level - (a) 2013 and (b) 2014 point cloud. EPSG:3912.

Figure 50 compares the position of the treetops and the shape and extension of the crowns automatically extracted by the segmentation procedure in the same graph. In the 2013 scenario, 76 trees were identified, while they dropped to 69 in 2014. Of these, 62 and 49 were correctly identified, respectively.

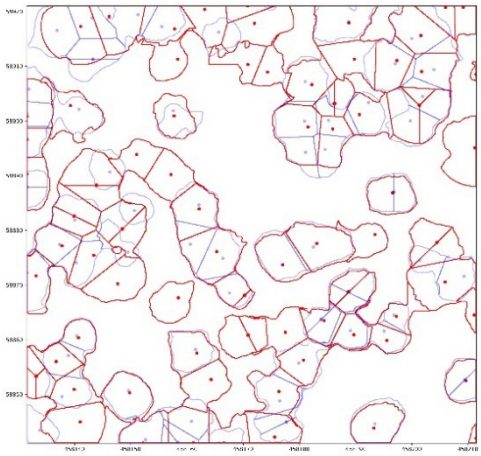


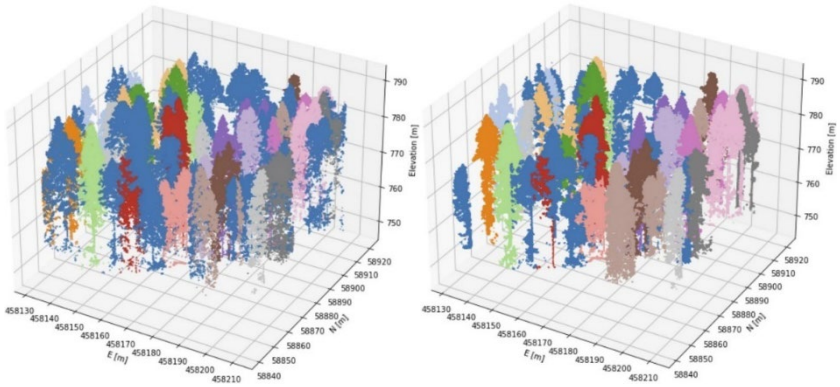
Figure 50: Treetops and crowns detection. The 2013 scenario is colored in red, and the 2014 scenario is colored in blue. EPSG:3912.

As shown in Table 11, the F1 score values are consistent with each other (76% in 2013 and 78% in 2014). In detail, this value appeared to be compatible with those proposed in the literature and obtained with different techniques and algorithms [278], [279]. No cases of simple omission were recorded, i.e. all the trees were detected; the main errors were due to under-segmentation and over-segmentation. The leading causes of erroneous segmentation were the high density of trees and the spatial heterogeneity typical of non-anthropized environments. Table 11 also shows the segmentation results of the 2013 scenario obtained through the original PyCrown code (first column). Comparing them with those obtained after implementing the improvement to limit over-segmentation (second column), a difference of about ten percentage points in the F1 score value is observed. This result was sufficient to extend this modification to the remaining scenarios and case studies without further analysis.

Table 11: Metrics of the segmentation procedure, PyCrown algorithm - Case study A.

Case study A - Metrics	**Individual Tree Detection**		
	2013 (Original code)	**2013 (Improved code)**	**2014**
$True_{Positive} + False_{Negative}$	87	87	57
$True_{Positive} + False_{Positive}$	85	76	69
$True_{Positive}$	58	62	49
Precision	0.68	0.82	0.71
Recall	0.67	0.71	0.86
F1 score	67%	76%	78%

The results of identifying the same tree in the two different multi-temporal acquisitions are shown in Figure 51.



(**a**) (**b**)

Figure 51: Segmented point cloud at single tree level with multitemporal match - (a) 2013 and (b) 2014 point cloud. The same trees are colored with the same colors; blue trees did not have a positive match. EPSG:3912.

Table 12: Metrics of the multitemporal detection of individual trees procedure - Case study A.

Case Study A - Metrics	**Multitemporal detection of individual trees**	
	2013	**2014**
$True_{Positive} + False_{Negative}$	57	57
$True_{Positive} + False_{Positive}$	43	41
$True_{Positive}$	34	34
Precision	0.79	0.83
Recall	0.60	0.60
F1 score	68%	69%

As shown in Table 12, out of 57 trees that survived the natural hazard (and therefore present in both LiDAR acquisitions), the multi-temporal correspondence was performed correctly for 34 trees in the point cloud dataset. The resulting F1 scores are 68% for the first point cloud and 69% for the second. It should be noted that the value related to the 2014 scenario cannot be higher than the F1 score value obtained in the

segmentation phase. As a matter of fact, it is not possible to correctly match trees if they are not adequately detected simultaneously in the two acquisitions. For this reason, the F1 score values related to the 2013 dataset cannot be compared with each other. However, the number of trees correctly identified in the two acquisitions is expected to be identical.

6.2.2 Wood and leaf separation

The results of the three proposed approaches for segmentation at the trunk and leaf levels are shown in Figure 52. Additionally, Figure 53 plots the woody and foliar components separately. To facilitate visualization and understanding, it was decided to show only a tree rather than the entire forest point cloud for illustrative purposes. However, the results shown on the selected tree are significant for all the processed trees.

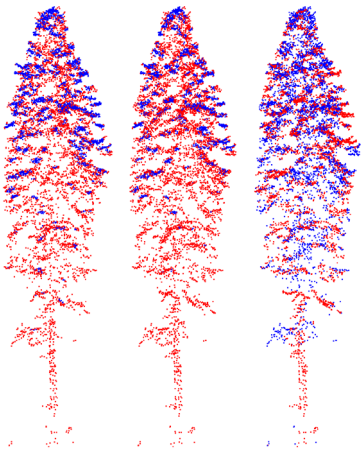
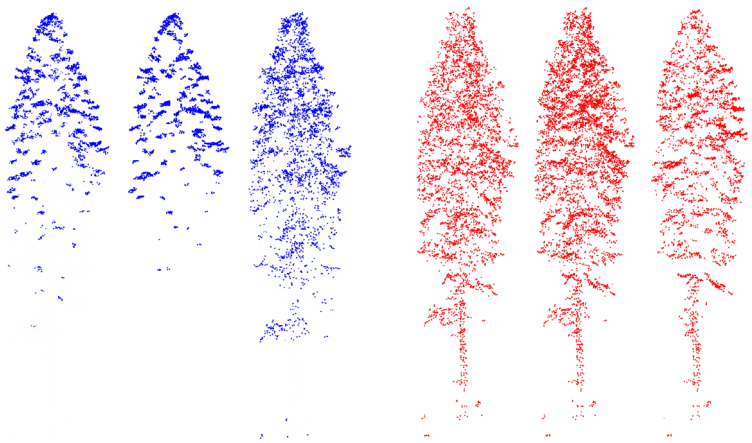


Figure 52: Wood and leaf separation on the same tree: on the left, first approach; in the middle, second approach; on the right, third approach - 2013 scenario.

Figure 53 shows the results more clearly, and more specifically, Figure 53b highlights the absence of a well-defined woody structure.



(**a**) (**b**)

Figure 53: Wood and leaf separation on the same tree: on the left, first approach; in the middle, second approach; on the right, third approach. (a) leaf component; (b) wood component - 2013 scenario.

The first two methods return results quite similar to each other, while with the third method, a greater number of points are labelled as leaf (blue points). From a visual interpretation, there are no doubts about correctly identifying the main trunk in the lower part of the tree; however, branches are more difficult to identify due to the low point density.

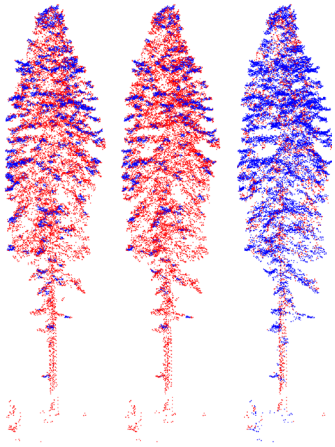
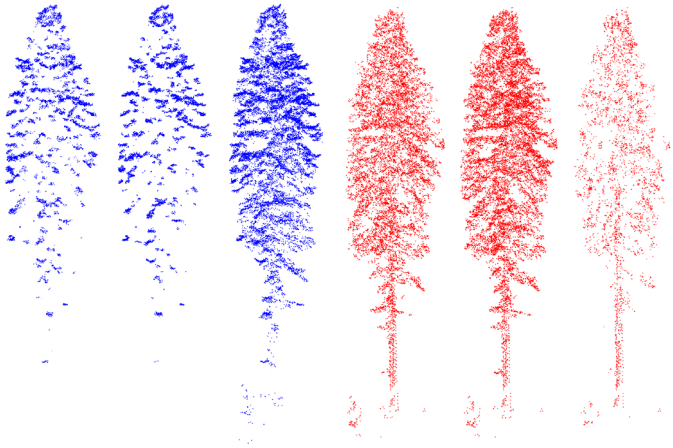


Figure 54: Wood and leaf separation on the same tree: on the left, first approach; in the middle, second approach; on the right, third approach - 2014 scenario.



(**a**) (**b**)

Figure 55: Wood and leaf separation on the same tree: on the left, first approach; in the middle, second approach; on the right, third approach. (a) leaf component; (b) wood component - 2014 scenario.

For comparative purposes, Figure 54 and Figure 55 show the segmentation results obtained on the same tree acquired in 2014 between the acquisitions that occurred at two different moments. Although the point cloud density is slightly higher, this is not enough to achieve better segmentation. The tree shown is one of those on which the storm had the most negligible impact: if the time instants are compared, only minor variations are observed.

In conclusion, it can be stated that the proposed wood and leaf segmentation methodologies are inadequate to complete this task in this scenario. However, it should be made clear that the point cloud type does not allow a clear identification of the woody structure, even through careful manual and visual segmentation. This peculiarity leads to the assumption that the validity of the wood and leaf separation data processing cannot be assessed based on the results obtained in this scenario.

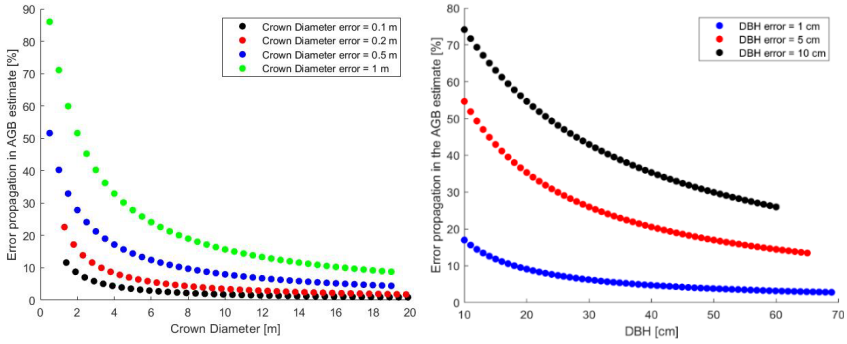
6.2.3 Error propagation in the allometric equations

The results of the estimation of the error propagation of Equations (16), (19) and (15) are detailed below. All the equations assumed an exponential trend in the percentage of error committed, and the propagation of the error in estimating the parameters increased as the size of the tree (crown diameter) decreased. When evaluating the error propagation in the AGB estimation through Equation (16), an error in the analysis of the crown diameter equal to one meter could lead to an overestimation of the biomass of the trees by up to 30%. The overestimation could be reduced and limited to 15% if the overestimation of the CD was halved (Figure 56a). Equation (15) is affected by error propagation with the fastest exponential growth compared to the previous equations (Figure 56b) since the analyzed parameter (the DBH, expressed in centimeters) is squared. For the smallest trees (with DBH between 10 and 20 cm), an error in the estimate of 1 cm caused an error

in terms of percentage ABG between 10 and 20%; if the DBH was overestimated by 10 cm, this error could exceed 70%. If we consider the medium/large trees, the percentage error of the ABG is reduced and becomes reasonable if the DBH is estimated with a maximum error of 1 cm. In comparison, it reached values higher than 15–20% for estimates affected by more significant uncertainty (5 cm).

With respect to Equation (19), it was observed (Figure 57) that the DBH estimate for trees with a crown described by a diameter between 5 and 14 m (such as those analyzed in the case in question) was affected by a maximum error propagation of about 15% when the error in the evaluation of the diameter of the canopy was one meter. Considering the errors in this estimate of less than one meter, the percentage of error committed was less than 10% for the trees in the study area.

These considerations highlight (i) the importance of having accurate estimates relating to basic forestry parameters but, at the same time, (ii) the substantial limitations to which allometric equations are subject.



(**a**) (**b**)

Figure 56: Error propagation of Equation (16) on the left, and Equation (15) on the right for the AGB assessment.

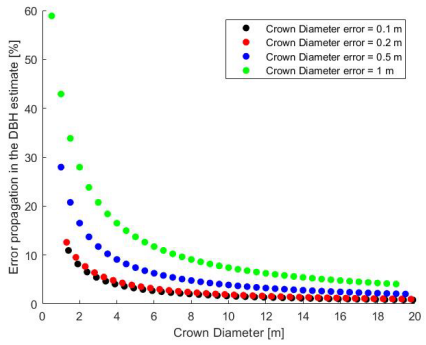


Figure 57: Error propagation of Equation (19) for the DBH assessment.

6.2.4 Forest parameters (DBH and AGB) characterization

The results related to the estimates of forest parameters are described in Table 13. In the case of pre-disaster lidar acquisition, the analysis was conducted on 29 correctly identified trees; among these, 14 were uprooted by the snowstorm, while 15 survived. Regarding the latter, the same analyzes were carried out using the point cloud acquired after the disastrous event.

Table 13: Estimated and reference values of forest parameters - 2013 scenario.

Tree ID	H [m]	CD [m]	AGB_{Jucker} [kg]	DBH_{Jucker} [cm]	DBH_{Ref} [cm]	AGB_{Chave} [kg]	$AGB_{TreeQSM}$ [kg]	AGB_{Ref} [kg]
104	24,6	7,9	1383	39,5	54,4	920	39938	1718
105	30,1	8,2	2130	48,1	60,8	1642	11314	2598
108	15,7	7,3	532	25,7	43,3	255	4102	708
113	24,3	8,4	1522	41,3	46,7	990	65115	1259
114	28,4	6,0	1103	35,7	37,0	868	31383	931
119	26,7	5,8	931	33,1	31,0	703	31970	620
120	22,8	8,2	1296	38,4	49,4	808		1323
128	27,5	8,3	1853	45,1	56,3	1329	33436	2050
129	25,4	5,7	817	31,2	35,0	598	28290	750
132	22,0	7,5	1044	34,8	35,1	644	3827	654
134	14,8	7,0	453	23,9	24,7	209	16028	224
135	22,2	5,7	638	27,9	29,0	422	40676	456
136	22,2	8,8	1417	40,0	43,3	853	17490	996
137	25,1	8,1	1502	41,0	60,8	1010	57785	2175
504	25,5	8,2	1591	42,1	46,8	1079	11538	1326
507	16,1	7,6	606	27,2	24,0	294	9887	230
509	30,0	10,4	3233	58,0	39,6	2366		1122
513	25,1	7,2	1223	37,4	33,1	844		665
514	26,9	5,9	954	33,4	33,2	725		715
515	26,9	8,6	1892	45,5	47,2	1326	62820	1419
516	27,6	5,9	1020	34,5	37,4	788		924
518	23,9	8,5	1504	41,1	46,8	966	21585	1246
522	25,1	5,7	795	30,8	33,0	577	43146	661
525	21,4	5,2	515	25,3	26,1	336		357
527	14,1	7,1	419	23,1	23,9	187	54342	200
530	22,7	7,4	1069	35,2	27,6	677	32760	421
531	22,9	7,8	1197	37,0	42,5	757	58703	990
532	14,4	7,9	541	25,9	33,1	239	33470	387
533	23,5	9,8	1891	45,5	46,8	1161	26811	1224

The results were subjected to statistical analyses: precisely, the RMSE and bias values were calculated. Table 14 and Table 16 refer to the 2013 and 2014 scenarios, respectively

As can be immediately observed, using an approach based on the Quantitative Structure Model encounters several problems. First of all, it cannot produce results for all the considered trees; furthermore, the estimates provided differ from the reference one by an order of magnitude (or more). These limitations are due to the impossibility of fitting wood points into hierarchical cylindrical elements, mainly due to the low point density and insufficient and incorrect woody and foliar point separation. However, when the algorithm manages to fit geometric shapes around the points, these are highly oversized, as they fit any point identified as wood. For these reasons, the results obtained from this model for estimating the AGB are not considered valid.

Equation (19) tends to underestimate the DBH of trees with a bias of about -6 % and an RMSE of 8 cm. On the contrary, Equation (16), although it depends, like Equation (19), on H and CD parameters, overestimates the AGB by approximately 42% (RMSE equal to 519 kg). As regards the estimate of the AGB through Equation (15), consistently with the estimates of the DBH of Equation (19) on which it depends, a slight underestimation is observed (about -9%, with RSME equal to about 452 kg). In fact, the DBH estimates are affected by negative bias (6.2%) and an RMSE equal to 8 cm.

Table 14: Statistical indexes for forest parameters - 2013 scenario.

Assessment	RMSE	BIAS
$\mathbf{AGB_{Jucker}}$	519 kg	42,5 %
$\mathbf{DBH_{Jucker}}$	8 cm	-6,2 %
$\mathbf{AGB_{Chave}}$	452 kg	-9,1 %
$\mathbf{AGB_{TreeQSM}}$	32040 kg	3797,9 %

Similar results are observed in the post-event scenario, showing that the best estimate of the AGB is confirmed to be obtained with Equation (15). In contrast, Quantitative Structure model estimates are completely misleading (Table 15). Consequently, the RMSE and bias values are also consistent (Table 16).

Generally speaking, the estimates obtained through equation (19) are affected by a higher error. This result was expected since these estimates are validated according to estimates assessed with Equation (15). However, comparing results obtained through different equations involves an intrinsic bias.

The estimation of forest parameters using the 2014 acquisition and the reference values considered are affected by more significant uncertainty. It is logical to think that the DBH of trees does not change significantly over a few months; however, since this is estimated as a function of the crown diameter and the tree height, different values can be obtained if the canopy has undergone significant variations. The same happens with respect to the use of allometric equations for biomass estimation. In the specific case where most of the trees have suffered major damage, these estimates are distorted, as they are generally valid only if the tree is undisturbed. However, for the purposes of this study, and especially in the absence of direct volumetric estimates carried out with QSM, considerations on allometric equations are the best tool available.

Table 15: Estimated and reference values of forest parameters - 2014 scenario.

Tree ID	H [m]	CD [m]	AGB_{Jucker} [kg]	DBH_{Jucker} [cm]	DBH_{Ref} [cm]	AGB_{Chave} [kg]	$AGB_{TreeQSM}$ [kg]	AGB_{Ref} [kg]
504	23	8,4	1387,1	40	46,8	865	9324	1199
507	13,23	6,6	331,9	21	24,0	143		190
509	25,45	7,0	1195,6	37	39,6	838	7153	955
513	20,1	6,6	709,4	29	33,1	420	2932	535
514	25,77	5,7	842,1	32	33,2	622	5729	686
515	10,76	14,1	896,8	33	47,2	281	12253	580
516	24,44	8,5	1572,7	42	37,4	1026	4288	821
518	27,04	7,0	1312,3	39	46,8	965	12164	1404
522	20,86	7,4	929,3	33	33,0	552	17456	551
525	17,99	7,3	685,8	29	26,1	366	14016	302
527	20,01	6,7	722,8	29	23,9	425	12995	282
530	12,27	9,0	509,4	25	27,6	194	11830	232
531	23,59	6,9	1022,5	34	42,5	678	3942	1018
532	20,69	7,9	1028,9	35	33,1	599	19334	550
533	12,05	7,7	370,0	22	46,8	143	4045	638

Table 16: Statistical indexes for forest parameters - 2014 scenario.

Assessment	RMSE	BIAS
$\mathbf{AGB_{Jucker}}$	337kg	55,2 %
$\mathbf{DBH_{Jucker}}$	9 cm	-8,7 %
$\mathbf{AGB_{Chave}}$	239 kg	-13,3 %
$\mathbf{AGB_{TreeQSM}}$	10141 kg	1798,1 %

Finally, Table 17 shows an estimate of the AGB loss referring to the analyzed trees. Of the 15 surviving trees, the loss is 16% (about 1950 kg), while the estimate is 34% (about 4200 kg). Roughly assuming that similar results can be applied to the remaining trees (for a total of 57 surviving trees and 30 uprooted trees), using a simple proportion, within the area in question, the reference biomass lost from damaged trees is approximately equal to 7400 kg, while that relating to uprooted trees is about 35300 kg (total of 43,000 kg). The estimates quantify 24,100 kg and 16,000 kg, respectively, for approximately 40,000 kg of AGB lost.

Table 17: Estimate of biomass loss following the disastrous event - Case study A.

Biomass loss	$\mathbf{AGB_{Chave}}$		Reference	
	[kg]	[%]	[kg]	[%]
14 uprooted trees	11250	100%	16461	100%
15 survived trees	4208	34%	1946	16%

6.2.5 CO_2 assessment

Table 18 shows the results relating to the estimate of the CO_2 stocked by the trees considered in the two different time instants through Equation (15), while Table 19 is related to statistical indexes. Complete results are reported in Appendix C, where the carbon dioxide assessment is based on AGB values estimated according to (16), and QSM output is

also listed. Since CO_2 was considered equal to half of the AGB, the bias is unchanged, while the RMSE value is halved.

Regarding the trees considered, the carbon dioxide preserved in the uprooted trees is approximately equal to 8232 kg, while that preserved in the portion of damaged trees is roughly equivalent to 973 kg. The estimates quantify these values at 5627 kg and 1997 kg, respectively. The cause of the differences, albeit minimal, is to be found in the changes in scenarios. The surviving trees suffered damage and biomass losses; however, the estimates' errors increased because the allometric equations are not designed to calculate the AGB under external damage.

Assuming that the damages thus quantified can be extended to the entire population of the forest under examination, a loss of carbon dioxide of between approximately 20 and 21.5 tons can be estimated.

Table 18: Estimated and reference values of stocked CO_2 - Case study A.

Tree ID	2013 scenario		2014 scenario	
	CO_{2Chave} [kg]	CO_{2Ref} [kg]	CO_{2Chave} [kg]	CO_{2Ref} [kg]
104	460	859		
105	821	1299		
108	128	354		
113	495	630		
114	434	466		
119	351	310		
120	404	662		
128	664	1025		
129	299	375		
132	322	327		
134	105	112		
135	211	228		
136	426	498		
137	505	1087		
504	540	663	433	600

Table 18: Estimated and reference values of stocked CO_2 - Case study A.

Tree ID	2013 scenario		2014 scenario	
	CO_{2Chave} [kg]	CO_{2Ref} [kg]	CO_{2Chave} [kg]	CO_{2Ref} [kg]
507	147	115	71	95
509	1183	561	419	478
513	422	333	210	267
514	362	358	311	343
515	663	709	140	290
516	394	462	513	411
518	483	623	482	702
522	289	331	276	276
525	168	179	183	151
527	93	100	212	141
530	339	211	97	116
531	378	495	339	509
532	120	194	300	275
533	581	612	72	319

Table 19: Statistical indexes for stocked CO_2 - Case study A.

Assessment	2013 scenario		2014 scenario	
	RMSE	BIAS	RMSE	BIAS
CO_{2Chave}	226 kg	-9,2 %	120 kg	-13,3 %

6.3 Lesson learnt

The analyzes of this first case study produced results of considerable interest. They are briefly summarized below, and major emphasis is placed on the results of greatest interest.

The segmentation procedure at the single tree level was performed sufficiently and consistently with the preciseness proposed by the literature. Furthermore, the efficiency of the improvement developed in

this study on the original workflow of the PyCrown algorithm was demonstrated. However, it was also highlighted that it is not possible to process aerial point clouds with low point density through a semantic segmentation approach due to the poor quality of the cloud in the lower part of the trees.

However, the distinction between woody and foliar points did not produce the desired results. The point cloud is not dense enough to obtain segmentation, enabling the woody structure to be uniquely identified. Nonetheless, it is essential to highlight that poor segmentation quality is also achieved through detailed visual interpretation of the point cloud. For this reason, the proposed methodologies need to be tested on point clouds with different characteristics (e.g., different acquisition perspectives or different instruments).

Consequently, the aboveground volumetric and biomass estimate through QSM is also fallacious due to a low-quality wood and leaf separation result. In fact, the algorithm is asked to identify cylindrical geometric structures around an insufficiently segmented point cloud at the part level. This limitation leads to incorrect results, as cylinders of larger dimensions than reality are automatically generated. Additionally, if there is undergrowth, this would increase the degree of error. However, analyzes have been performed directly through allometric equations. It is specified once again that these equations are not designed to monitor damage caused by external events; therefore, the estimates are certainly affected by an inherent error that cannot be quantified precisely.

Focusing on the results, it is estimated that approximately 40,000 - 43,000 kg of wood were uprooted, reducing the carbon-absorbing capacity of the analyzed forest (20,000 – 21,500 kg of carbon dioxide stocked in uprooted and damaged trees was lost). In this regard, there are two aspects to stress. First of all, post-event forest management is essential so that all actions aimed at minimizing the dispersion of CO_2 from dead trees back into the atmosphere caused by decomposition are

implemented. Furthermore, it must be considered that the damage suffered by individual trees affects their ability to grow and absorb carbon. Prolonged monitoring over time should be adopted to precisely quantify the resilience of trees and how quickly (if possible) they can grow again.

Finally, to answer the questions proposed at the beginning of the Chapter, it can be stated that aerial LiDAR data can be used to estimate the damage to which a forest is subject in the event of natural disasters. Although not rigorous, these estimates provide robust data on which to develop further monitoring and management analyses. Furthermore, it was demonstrated that the modification to the single tree identification procedure led to a substantial improvement in the final output, reducing misclassifications caused by over-segmentation.

Chapter 7

CASE STUDY B: RESULTS AND DISCUSSION

The second case study covers a 70-hectare coniferous forest in the North-West Italian Alps. The peculiarities of this case study are (i) the occurrence of a forest fire before the LiDAR acquisitions, (ii) the harvesting of trees in the upper portion of the study area, resulting in different tree densities, and (iii) the use of two LiDAR instruments and different terrestrial acquisition approaches.

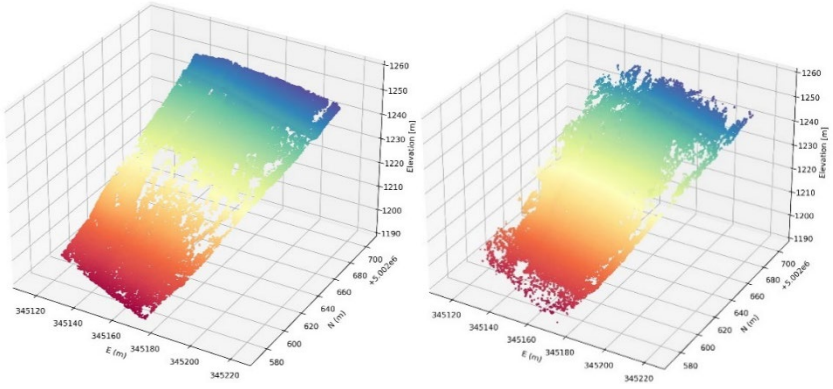
These characteristics directed the analyzes in order to:

- Test two different acquisition techniques, differing in characteristics and operating methods, to determine if one best suits this analysis.
- Identify individual trees in conditions of fire damage;
- Calculate the biomass (and stored CO_2) relating to surviving trees.

Contrary to case studies A and C, the point cloud dataset does not include multi-temporal analyses. Furthermore, there is no information about the allometric parameters collected in situ with traditional instruments.

7.1 General overview

Figure 58 shows the ground points segmented according to the CSF algorithm. Figure 52a shows that some ground portions were not acquired correctly with the TLS in the central part of the analysed area. On the other hand, concerning the MLS dataset, ground points are identified with greater continuity in the central part of the area. At the same time, a lack of data is found in the perimeter areas (Figure 58b). This highlights not only the importance of survey planning but also the operational difficulties, which can sometimes limit the quality of the data. In fact, the high slope of the section in question made movement in situ difficult, especially with the terrestrial LiDAR, due to its increased weight and size.



(**a**) (**b**)

Figure 58: Ground points of the study area B: (a) TLS point cloud; (b) MLS point cloud. EPSG:32632.

The aboveground biomass assessment was performed after selecting the best ITD and wood and leaf separation approaches. Estimated AGB values were validated according to Equation (15) results. However, since no in-situ data are available, DBH and H measures were entirely manually extracted from the point cloud. In this study, the density value of coniferous trees was set equal to 0.8 g/cm^3.

Table 20 summarizes the AGB estimation and validation procedure.

Table 20: Summary of the biomass estimation and validation procedure - Case study B.

Tree parameter	Estimates	Validation values
AGB	TreeQSM output	Equation (15) with in-situ measured H and DBH variables
	Equation (15) with derived H and DBH variables	

The datasets acquired with the different tools were compared in terms of accuracy. The Euclidean distance between the points was calculated using the commercial 3D Reshaper software. The analysis was conducted on the entire point cloud, and a limited portion was filtered to eliminate the undergrowth points and highlight the accuracy of individual trees (Figure 59). 70% of the points have an accuracy of less than 6.3 cm, and 85% have an accuracy of less than 12.5 cm. The figure also shows that the points with the least accuracy are mainly located in the area with the most significant forest density [271].

It should be emphasized that this type of error is also linked to the fact that the two acquisitions are not precisely contemporary and that the thinnest branches and leaves are subject to the effect of the wind, which modifies their position.

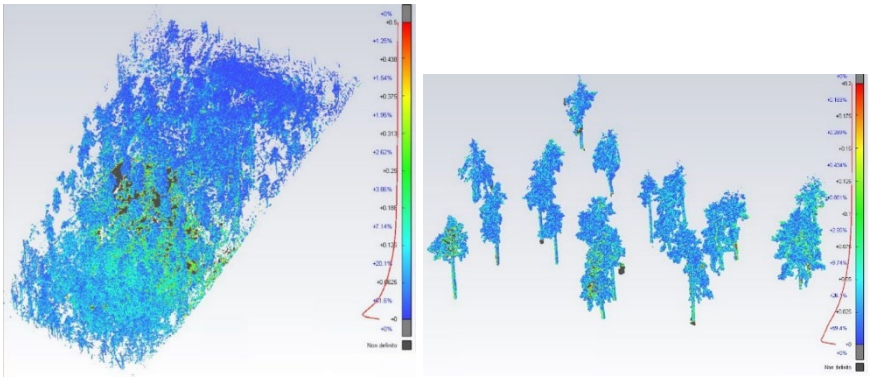
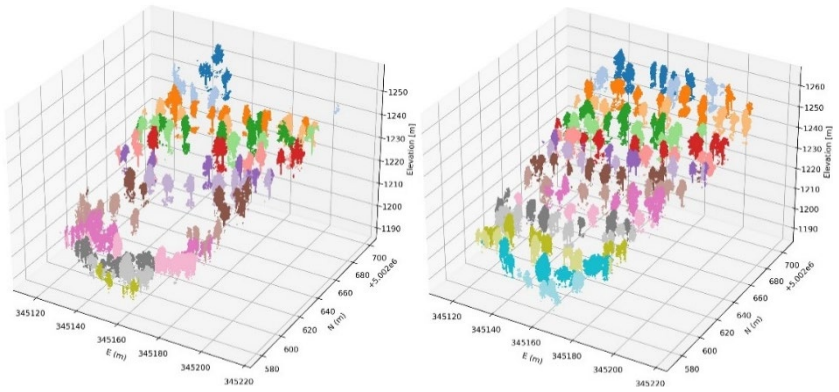


Figure 59: Accuracy of MLS point cloud compared with TLS dataset. On the left is the entire study area; on the right is a portion of individual trees [271].

7.2 Data processing

7.2.1 ITD

The two terrestrial LiDAR acquisitions produced point clouds that were positively analyzed with both proposed single-tree level segmentation methods. Figure 60 refers to results through the PyCrown segmentation, while Figure 63 refers to the FSCT segmentation results. It is essential to point out that the FSCT algorithm is particularly demanding from a computational point of view. For this reason, in order to avoid a lack of memory errors, it was necessary to divide the area under examination into five sub-areas, which were processed separately. The area boundaries were identified manually, taking care not to intersect the tree canopy. In fact, this would have led to segmentation errors due to the partial nature of the dataset. Finally, a single point cloud dataset merged the processed point clouds.



(**a**) (**b**)

Figure 60: Segmented point cloud at single tree level with PyCrown algorithm - (a) MLS point cloud and (b) TLS point cloud. EPSG:32632.

As can be easily observed by comparing Figure 60a and Figure 60b, missing trees can be detected in the MLS point cloud. They are mainly located in the central part of the study area, but some absent trees are also present in the upper part. This phenomenon is even more evident when comparing the position of the treetops and crowns of the trees identified for both datasets (Figure 61). The TLS point cloud appears to be better segmented; nonetheless, many trees are over-segmented or incorrectly segmented.

Analyzing the evaluation metrics, in fact, there is a very low F1 score with the MLS point cloud (33%); this value improves only by about 10 percentage points and stands at 44% relative to the TLS dataset (Table 21).

The results are unsatisfactory overall, and the segmentation method mentioned above cannot be considered good enough for single tree analysis. The leading causes of the poor results obtained are to be found: (i) in an incomplete description of the trees in the upper part (particularly concerning MLS point cloud), and (ii) in the consequences of forest fire.

As regards the first point, a clarifying example is represented in Figure 62. It shows the same tree acquired with the two proposed

techniques. Unfortunately, although the survey was carefully designed to survey the area comprehensively, some trees were not fully captured by the field view of the MLS LiDAR. The main causes are operational difficulty in a steeply sloping area and instrumental limitations. In this case, the region growing algorithm around the treetop (if identifiable) is not optimized, as the canopy is entirely absent; because of this, the algorithm fails to identify trees partially present in the point clouds.

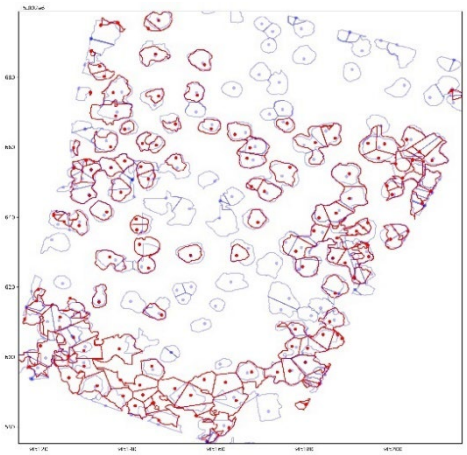


Figure 61: Treetops and crowns detected with PyCrown algorithm. MLS scenario is colored in red; TLS scenario is colored in blue. EPSG:32632.

At the same time, due to the consequences of the fire, most of the foliage and woody structures are destroyed. For the same reason discussed above, the region-growing algorithm encounters critical issues in individually segmenting a severely corrupted tree. It follows that an approach based on canopy and treetop identification is not the best solution in this scenario.

Table 21: Metrics of the segmentation procedure, PyCrown algorithm - Case study B.

Case study B - Metrics	**PYCROWN**	
	MLS	**TLS**
$True_{Positive} + False_{Negative}$	242	242
$True_{Positive} + False_{Positive}$	126	162
$True_{Positive}$	60	89
Precision	0,48	0,55
Recall	0,25	0,37
F1 score	33%	44%

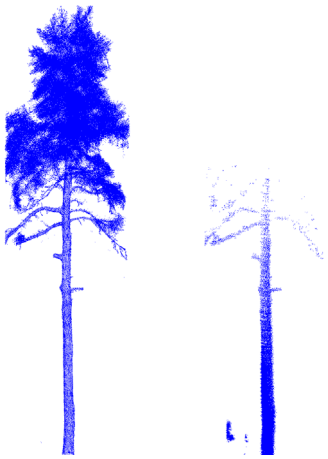
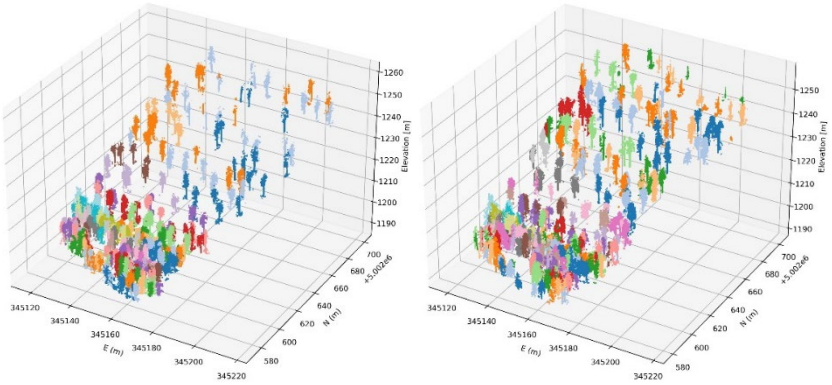


Figure 62: LiDAR acquisition of a single tree, (a) TLS point cloud; (b) MLS point cloud.

In Figure 63, the point clouds segmented with FSCT are shown. From a visual interpretation, the results are consistent with those obtained through PyCrown. As a matter of fact, some trees are not identified with the MLS point cloud.

A good match between the position of the treetops can be seen when taking Figure 64 into consideration. Nonetheless, the results are promising. The F1 score values are equal to 68% and 86%, respectively,

relative to the MLS and TLS point clouds. The quality of the MLS point cloud segmentation is slightly lower (but still acceptable for forestry analyses) due to the previously described point cloud characteristics.



(**a**) (**b**)

Figure 63: Segmented point cloud at single tree level with FSCT algorithm - (a) MLS point cloud and (b) TLS point cloud. EPSG:32632.

However, paying attention to the distribution of incorrectly segmented or completely missing trees is equally important. As a matter of fact, the latter is mainly concentrated in the lower part of the area, the one in which the distribution of trees is considerably greater. At the same time, isolated and not-too-close trees are easily identified. This result further confirms what was observed by [278] regarding the effect of forest density on ITD.

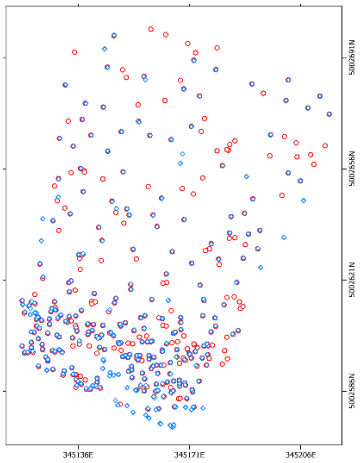


Figure 64: Treetops detected with FSCT algorithm. The MLS scenario is plotted in red circles; the TLS scenario is plotted in blue rhombus. EPSG:32632.

In conclusion, the best results in the terrestrial LiDAR point cloud case are obtained by processing the point cloud through the FSCT algorithm. For this reason, subsequent data processing was carried out considering the results obtained with the abovementioned method, while the individual trees detected according to the PyCrown algorithm were discarded.

Table 22: Metrics of the segmentation procedure, FSCT algorithm - Case study B.

Case study B - Metrics	**FSCT**	
	MLS	**TLS**
$True_{Positive} + False_{Negative}$	242	242
$True_{Positive} + False_{Positive}$	356	300
$True_{Positive}$	203	234
Precision	0,57	0,78
Recall	0,84	0,97
F1 score	68%	86%

7.2.2 Wood and leaf separation

Figure 65 and Figure 68 show the point cloud of an example single tree segmented into foliar and trunk parts. In general, satisfactory results are observed, as the woody structure can be clearly distinguished: not only the part relating to the central trunk but also secondary branches are identified. Concerning MLS point cloud, the use of the *k*-Means algorithm in the first segmentation method, however, is the one that performs worst among the three proposed methods. The density parameter added to the points' position is only sufficient to identify the central trunk (among other things, not wholly, since several points relating to it are misclassified as leaf class); as regards the branches, only the lower ones are partially identified.

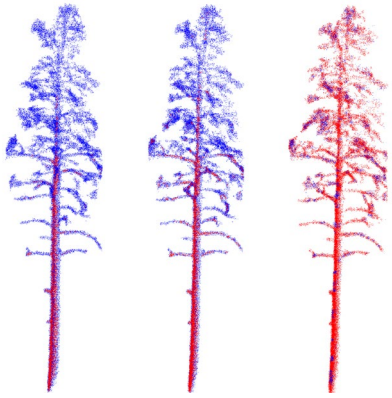
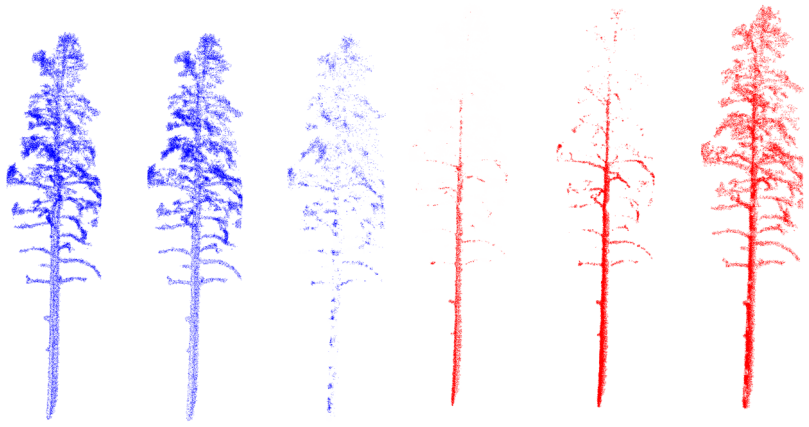


Figure 65: Wood and leaf separation on the same tree: on the left, first approach; in the middle, second approach; on the right, third approach - MLS point cloud.

The second approach returns slightly more robust results. In fact, a greater density of points is identified when detecting the trunk. In contrast, it is possible to define the boundaries of the lower branches with greater clarity and partially identify others. However, some woody parts

in the upper part of the tree, around the top, are identified, but the spatial continuity in elevation fundamental for QSM purposes is not guaranteed.

Finally, the third approach segments the point cloud more accurately and functionally for the purposes of this study.



(**a**) (**b**)

Figure 66: Wood and leaf separation on the same tree: on the left, first approach; in the middle, second approach; on the right, third approach. (a) leaf component; (b) wood component - MLS point cloud.

First, contrary to the first two approaches, only a few trunk and branch points are misclassified as leaf; furthermore, even the highest branches are satisfactorily identified. Although a good part of the leaves is identified (as can be clearly seen from Figure 66b), Figure 66a shows some leaves misclassified as wood, in particular in the upper part of the tree.

Concerning the TLS dataset, part segmentation results are consistent and coherent with what has been described (Figure 60); however, some minor differences must be pointed out and discussed.

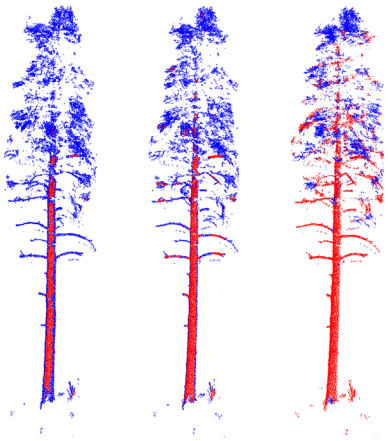
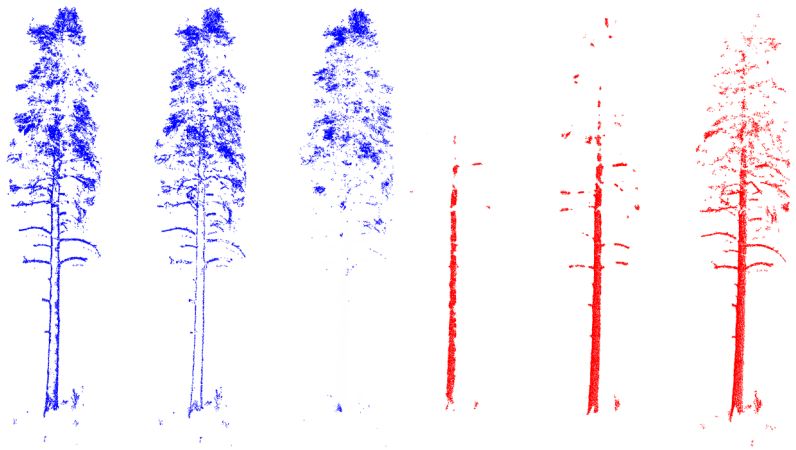


Figure 67: Wood and leaf separation on the same tree: on the left, first approach; in the middle, second approach; on the right, third approach - TLS point cloud.

Through all three approaches, but more so in the second and third, the trunk structure is classified wholly and correctly exhaustively. Through the first two methods, however, only a small percentage of points located in the most external part of the trunk are misclassified (Figure 68b). Contrary to what occurs with the MLS dataset, the branches are identified with greater difficulty: the first approach does not identify any of them, and the second identifies only some of them but in an unsatisfactory manner. At the same time, the third method is more exhaustive (Figure 68a). Additionally, fewer leaves are misclassified as trunks; however, since the density of points is lower in the higher part, the continuity between the trunks and branches is more intermittent.

Therefore, the segmentation outputs highlight how an approach based on the simple use of a k-Means algorithm is insufficient to segment the tree correctly. Generally, slightly better results are observed by first applying a threshold (approach 2). However, using only density as a discriminating parameter is not the best approach; the third approach, in which more features are considered, returns the best results. As expected,

the most significant uncertainties occur at the top of the trees. This aspect is not only dictated by the fact that the acquisitions were carried out with terrestrial approaches, although this could contribute. The leading cause lies in the structure of the trees themselves, as in the upper part, the branches have minimal dimensions, both in terms of diameter and length; furthermore, the quantity of leaves is usually greater.



(**a**) (**b**)

Figure 68: Wood and leaf separation on the same tree: on the left, first approach; in the middle, second approach; on the right, third approach. (a) leaf component; (b) wood component - TLS point cloud.

Regarding the volumetric estimate using Quantitative Structural Models, the woody part of the trees defined through the third approach was considered.

7.2.3 AGB characterization

To evaluate the AGB assessment quality, the weight estimated through QSM was compared with the reference weight calculated as expressed in Equation (18). For each LiDAR acquisition, the same 20

trees were considered; these were chosen from those correctly segmented individually and using the point cloud relating only to the woody part obtained as described in the previous Paragraph. Results are tabulated in Table 23 and Table 25.

The first observation and clarification that needs to be made concerns the reference values of height and DBH. As previously described (Paragraph 5.2), no in situ investigations were conducted. In order to have reference values, the point clouds were analyzed manually, extracting the forest parameters as follows: the DBH of the trees was calculated as the diameter of the circumference that best fit the points of the cloud between 1.2 m and 1.4 m above the ground; the height was calculated as the difference between the height of the highest and lowest point of each tree. However, Table 23 and Table 25 show different values between the two acquisitions. Specifically, heights are usually higher when calculated relative to the TLS point cloud; the same is observed with respect to the DBH. The type of acquisition causes this inconsistency. As previously discussed, the MLS point cloud is incomplete at the top of the trees due to operational limitations. On the other hand, the chosen trajectory of the MLS acquisition allows points that describe the trunks to be acquired homogeneously. Although TLS acquisitions were carried out from 5 different station points, they did not acquire complete information for some trees. For this reason, the points of the MLS cloud fit the circumference of the trees with less error percentage. However, to avoid distorting the results by comparing them with inappropriate (albeit correct) references, it was decided to use the forest parameters obtained for each point cloud.

Out of 20 trees and 40 AGB estimates (one for each dataset), QSM failed in the modelling procedure only in 3 cases (twice for the MLS dataset and once for the TLS dataset). The failure is related to three different trees, and it is caused by an insufficiently complete acquisition of each element of the trees (Figure 69).

```
Error using cell
NaN and Inf not allowed.

Error in cubical_partition (line 88)
  Partition = cell(N(1),N(2),N(3));

Error in cover_sets (line 91)
  [partition,CC] = cubical_partition(P,BallRad);

Error in treeqsm (line 275)
  cover1 = cover_sets(P,Inputs);

Error in TreeQSM_SingleTree (line 14)
QSM = treeqsm(P, inputs);
```

Figure 69: Error resulted during the QSM procedure - Case study B.

Table 23: Estimated and reference values of forest parameters - MLS point cloud.

Tree ID	**H [m]**	**DBH_{Ref} [cm]**	**AGB_{Ref} [kg]**	**AGB_{QSM} [kg]**
1	5,80	33,1	185	223
2	11,81	25,7	226	
3	8,34	23,8	138	213
4	12,03	26,1	237	613
5	13,17	28,5	307	
6	12,70	25,6	240	195
7	10,16	20,4	124	45
8	12,07	33,4	385	312
9	13,78	30,3	362	271
10	10,84	22,9	166	125
11	11,66	19,2	126	90
12	12,61	32,2	374	170
13	9,50	32,3	285	324
14	13,27	33,3	419	352
15	14,00	32,5	421	351
16	12,62	24,2	214	228
17	6,69	9,4	18	146
18	8,56	13,3	46	90
19	8,40	18,5	85	90
20	5,75	12,7	28	8

Table 23 summarizes reference and QSM AGB estimated values.

The results show an overestimation of the biomass weight (approximately equal to 13%) corresponding to an RMSE of roughly 120 kg when taking into account the MLS-acquired trees; lower errors can be recorded through the TLS acquisition (RMSE equal to 95 kg and overestimation BIAS of 6%).

Table 24: Statistical indexes for forest parameters - MLS scenario.

Assessment	RMSE	BIAS
$\mathbf{AGB_{QSM}}$	117 kg	39,8 %

The resulting overestimates (explicitly referring to MLS point cloud) must be contextualized. As a matter of fact, there is a substantial difference between the values obtained about tree number 17. If this value were excluded as an outlier, no significant change in the RMSE value would be achieved; however, the bias would drop to approximately 1%. This value would be consistent with the one calculated for the MLS dataset. Therefore, although the predicted and obtained values differ slightly, this difference must be sought in the validity of the allometric equations. Since these are more accurate only when the tree is growing steadily, one might expect the estimates obtained by modeling the wood structure after the tree has been damaged to be more reliable; furthermore, they are expected to provide greater values than allometric equations.

However, the overestimation observed with the TLS point cloud (Table 25 and Table 26) may be caused by the DBH values considered. Most likely, QSM is more accurate in estimating tree diameter; since the analysis is not limited to a height of 20 cm (from 1.20 to 1.40 m above the ground) but extends to the entire trunk, this can be reconstructed more precisely.

Table 25: Estimated and reference values of forest parameters - TLS scenario.

Tree ID	H [m]	DBH_{Ref} [cm]	AGB_{Ref} [kg]	AGB_{QSM} [kg]
1	11,49	30,5	307	354
2	11,46	28,5	268	287
3	10,68	24,3	183	248
4	13,80	36,9	532	447
5	13,35	31,2	372	295
6	14,04	27,7	309	296
7	11,37	25,8	219	305
8	12,24	31,9	356	161
9	14,25	35,5	510	576
10	11,49	25,5	216	192
11	13,62	19,9	157	154
12	11,58	23,9	192	146
13	10,74	26,3	215	235
14	14,43	40,8	677	538
15	13,25	31,6	378	492
16	10,90	35,4	390	
17	9,55	31,1	266	518
18	10,52	24,3	181	158
19	10,97	24,8	196	189
20	8,56	17,5	78	107

Generalizing, volumetric and aboveground biomass estimates are consistent and show high accuracy for forestry survey purposes, enough to improve and possibly replace the results of allometric equations.

Table 26: Statistical indexes for forest parameters - TLS scenario.

Assessment	RMSE	BIAS
AGB_{QSM}	95 kg	5,9 %

7.2.4 CO_2 assessment

Table 27 summarizes CO_2 estimates, while Table 28 shows statistical indexes. Since the correction factor for carbon dioxide assessment is a constant value, the same considerations made for the latter are also valid for stocked carbon dioxide. Therefore, also in this case, excluding tree 17 from the analyzes as it is an outlier, the bias of the MLS cloud decreases up to 1%.

Table 27: Estimated and reference values of stocked CO_2 - Case study B.

Tree ID	MLS		TLS	
	CO_{2QSM} [kg]	CO_{2Ref} [kg]	CO_{2QSM} [kg]	CO_{2Ref} [kg]
1	112	92	221	154
2		113	180	134
3	106	69	155	92
4	306	118	280	266
5		154	185	186
6	98	120	185	155
7	22	62	191	110
8	156	192	101	178
9	136	181	360	255
10	62	83	120	108
11	45	63	97	79
12	85	187	92	96
13	162	143	147	108
14	176	210	336	338
15	176	211	308	189
16	114	107		195
17	73	9	324	133
18	45	23	99	90
19	45	43	118	98
20	4	14	67	39

Table 28: Statistical indexes for stocked CO_2 - Case study B.

Assessment	MLS		TLS	
	RMSE	**BIAS**	**RMSE**	**BIAS**
CO_{2QSM}	58 kg	40 %	68 kg	6 %

The total estimate of carbon dioxide stored within the 20 trees analyzed is approximately equal to 1,9 tons and 3,5 tons, respectively, on MLS and TLS point cloud-based analysis. Based on all the considerations about the two datasets, it is reasonable to assume that the second estimate is more reliable. This leads to the conclusion that terrestrial point cloud acquisitions must be programmed with extreme care in order to minimize errors. Furthermore, MLS acquisition techniques in the operating methods explored in this study seem unsuitable for the purposes pursued in this thesis work; nevertheless, further studies should be conducted considering handheld LiDARs with wider fields of view which could offer better results.

7.3 Lesson learnt

This case study focused on comparing two terrestrial data acquisition methodologies, evaluating the pros and cons of data acquisition in a forest environment subject to forest fire damage.

The raster-based approach for segmentation at the single tree level encounters some critical issues. Added to the difficulty of processing datasets with poor quality in the upper part of the tree is a further degree of difficulty caused by the damage to which the trees are subject. On the contrary, the results obtained via semantic segmentation are acceptable and consistent with the values proposed by the literature. Using the latter procedure, an F1 score of 86% is obtained relative to the TLS point cloud. This value is the highest among all the case studies analyzed

(86%). The MLS dataset also performs well; however, limitations related to mobile acquisitions caused the F1 score to drop to 68%.

Separation of wood and leaf points has been performed with promising results. All the proposed approaches identify the main trunk; however, using the third approach, it is possible to precisely delineate the larger and smaller branches (if these are located in the lower part of the tree or are bare of leaves). It should be specified that the foliar part of the trees was damaged due to the fire, facilitating their separation from the woody part.

As a result of good wood and leaf segmentation, it was possible to perform the volumetric modeling of trees through QSM in almost all of the cases analysed. The comparison with the values obtained from allometric equations is not excellent. However, taking into account the limited validity of these equations for quantifying the biomass of damaged trees, these results can be considered valid.

In conclusion, in order to answer the questions posed at the beginning of the Chapter, it is possible to segment individual trees in post-fire scenarios through semantic segmentation. Furthermore, the quality of the terrestrial datasets allows to conduct detailed analyzes of the biomass and, consequently, the CO_2 stocked in the wood through geometric modeling. However, further investigations are needed to validate the results obtained through this approach; although the allometric equations provide a rough estimate, they have a lower value in this case study due to external events that have seriously damaged the trees. It is estimated that the carbon dioxide stored in the trees in the study area is between 23 and 42 tons.

Finally, as already mentioned, this case study highlighted the pros and cons of static and dynamic terrestrial acquisitions in the forestry domain. By analyzing the quality of the point cloud in its entirety and considering the evaluation metrics and indexes, the TLS approach provides the best results. However, it should not be forgotten that this

type of acquisition is characterized, on the other hand, by greater operational difficulty and processing times. On the contrary, the MLS point cloud has a lower quality, affecting the accuracy of the final results; nevertheless, these are still acceptable for large-scale studies in forestry. Furthermore, the greater simplicity in the operational phase in scenarios with a particular degree of difficulty means that this approach can be considered a fair compromise.

Chapter 8

CASE STUDY C: RESULTS AND DISCUSSION

Data acquisitions occurred within the Hyytiälä forest field station, which is commonly used for educational and research purposes. The instrument used for forest monitoring is a RIEGL VZ-2000i terrestrial LiDAR; however, the point cloud acquisition is not conventional. The laser scanner is permanently installed on a 35-meter-high tower, constantly acquiring information; its position and its orientation in the space allow to obtain an oblique point cloud with an innovative and unique perspective (more details about the setup can be found in Paragraph 5.3).

The analysis concerned two LiDAR point clouds acquired within a year of each other (2020 and 2021) during the same month (April). The month was chosen because, in this period, the snow accumulated during the winter months should have melted entirely, and there should be no

accumulations of it on tree branches. Additionally, the leaf-off season helps in the wood and leaf separation procedure.

Contrary to case study A, no major disturbance events occurred during the year under examination. This means that the only changes expected to be observed are an increase in biomass following the growth of the trees (as a matter of fact, the forest is only 63 years old); however, some weaker branches can have broken due to the excessive snow load sustained during the winter season. The main aspects to be investigated in this case study are:

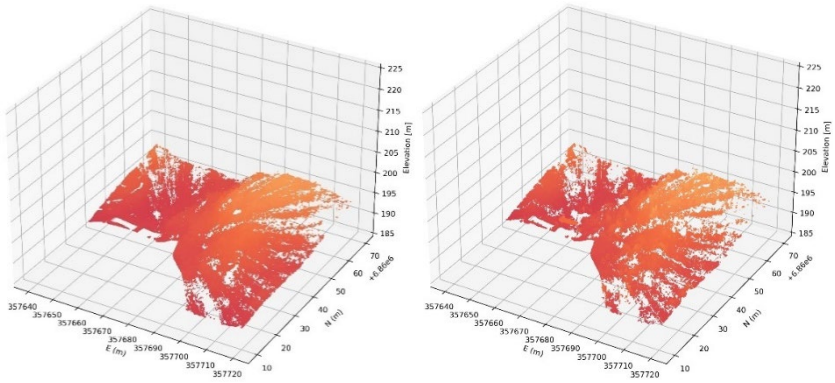
- The actual usefulness of an innovative type of acquisition such as the oblique one;
- The evaluation of the temporal evolution of CO_2 in substantially undisturbed conditions;
- Establish which features most influence the segmentation of wood and leaf.

8.1 General overview

Ground truth data regarding the position of the trees within the area in question were acquired through in situ surveys, while no additional data is available. Segmented ground points are shown in Figure 70.

Some features of these outputs can already be discussed. First of all, it can be observed that the single point of view, albeit with a high point density, generates shadow cones caused by the trunks and larger branches, which hinder the penetration of a certain quantity of laser beams. Obviously, the cones become larger and more numerous as they move further away from the station point. Secondly, but no less critical, the quantity of leaves increases with the distance from the LiDAR, making the point cloud further scattered in the perimeter areas of the study area.

Concerning the ITD task, the maximum extension of the tree crown was set equal to 15 meters. Once again, the F1 score was used (i) to evaluate the results obtained in ITD procedure and (ii) to validate the multi-temporal matching procedure.



(a) **(b)**

Figure 70: Ground points of the study area C: (a) 2020 point cloud; (b) 2021 point cloud. EPSG:3067.

The same methodologies described in the other case studies were adopted regarding the separation of trunk and leaf. However, contrary to previous case studies, in this study, further analyzes were conducted regarding the features to be considered. Additional tests concern the contribution of the deviation as an additional feature [272].

The volumetric estimates obtained via QSM were finally compared with those obtained from the Finland empirical volumetric equation [54] detailed in Paragraph 4.6. As ground truth data (DBH and tree height) were unavailable, they were extracted from a manual interpretation of the point cloud, limited to the trees on which these analyses were conducted; DSH values were estimated according to Equation (21). Table 29 summarizes the validation procedure.

Table 29: Summary of the biomass estimation and validation procedure - Case study C.

Tree parameter	Estimates	Validation values
AGB	TreeQSM output	Equation (15) with in-situ measured H and DBH variables
	Equation (15) with derived H and DBH variables	
DBH	Equation (19) with derived H and DBH variables	Manually extracted values

8.2 Data processing

8.2.1 ITD

The Individual Tree Detection procedure was performed with both algorithms explored in this study (Figure 71 and Figure 73 refer to PyCrown and FSCT outputs, respectively). In fact, the particular type of acquisition is to be considered, for all intents and purposes, a middle ground between terrestrial and aerial acquisitions. For this reason, it is expected that segmentation at the single tree level can be performed successfully with both proposed algorithms.

On the downside, the forest density is high (approximately equal to 2000 trees/ha). This peculiarity could affect the reliability of the results; additionally, even through a visual interpretation, the single tree detection procedure is not totally easy.

Figure 72 emphasizes the difficulty of correctly segmenting very close trees, for which the crowns often intersect each other. Nonetheless, the results are encouraging.

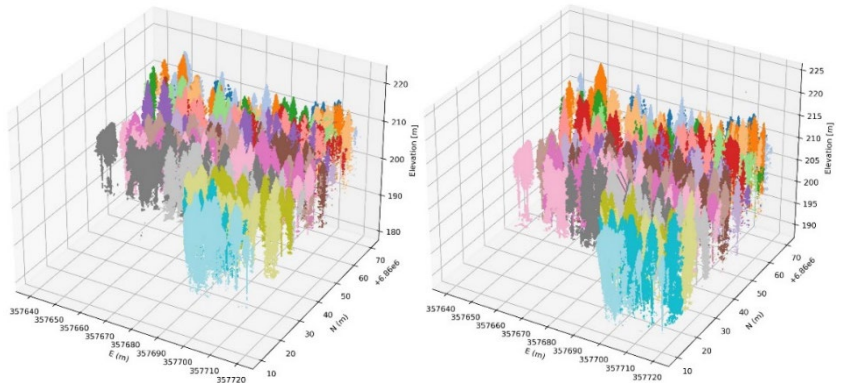


Figure 71: Segmented point cloud at single tree level with PyCrown algorithm - (a) 2020 point cloud and (b) 2021 point cloud. EPSG:3067.

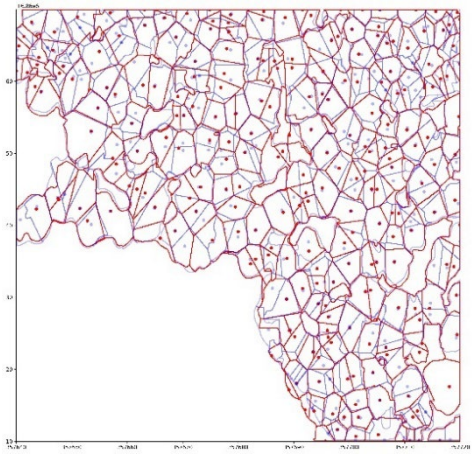


Figure 72: Treetops and crowns were detected with the PyCrown algorithm. The 2013 scenario is colored in red, and the 2021 scenario is colored in blue. EPSG:3067.

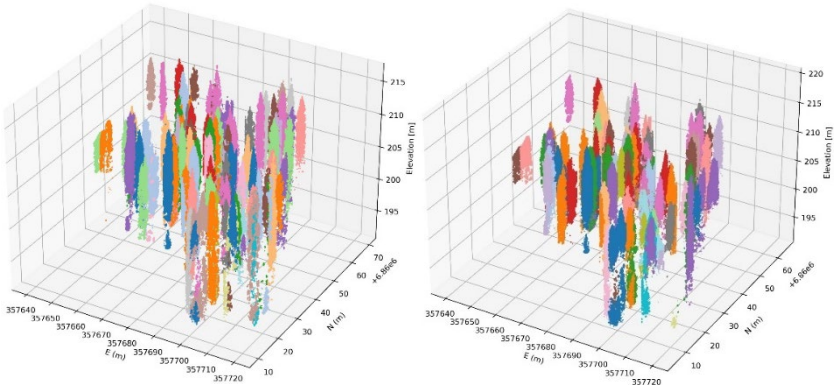
The F1 score values (Table 30) are equal to 65% and 53% for the time instants considered. The difference (27 trees not correctly identified relating to the second dataset) is related to an increase in over- and under-

segmentation errors. Although the scenario should be approximately unchanged, these differences could be caused by minimal differences between the two acquisitions, such as different foliar coverage.

Table 30: Metrics of the segmentation procedure, PyCrown algorithm - Case study C.

Case study C - Metrics	**PYCROWN**	
	2020	**2021**
$True_{Positive} + False_{Negative}$	184	184
$True_{Positive} + False_{Positive}$	241	235
$True_{Positive}$	138	111
Precision	0,57	0,47
Recall	0,75	0,60
F1 score	65%	53%

As has already been highlighted in the previous Chapter, also in this case study, it was necessary to divide the point cloud into five areas of reduced extension in order to process them using the FSCT algorithm and then subsequently merge them again into a single dataset. The results, however, are not optimal. Numerous trees are not identified; this aspect can be observed by comparing the FSCT output of different datasets (Figure 73a and Figure 73b) and by comparing PyCrown and FSCT results (Figure 71 and Figure 73). The predicted positions of the trees further confirm this phenomenon (Figure 72 and Figure 74).



(a) **(b)**

Figure 73: Segmented point cloud at single tree level with FSCT algorithm - (a) 2020 point cloud and (b) 2021 point cloud. EPSG:3067.

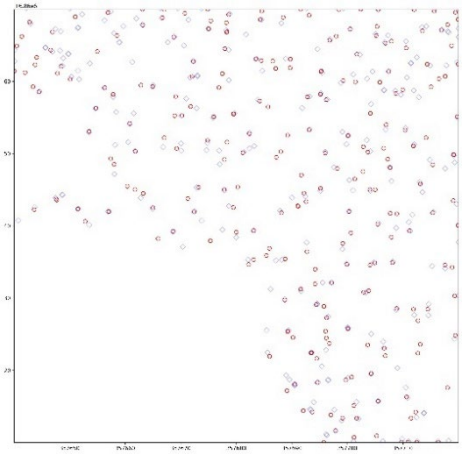


Figure 74: Treetops detected with FSCT algorithm. The 2013 scenario colored in red, and the 2021 scenario colored in blue. EPSG:3067.

In this case, the F1 score evaluation metrics are 59% and 56%, respectively (Table 31). The different approach the FSCT algorithm takes compared to PyCrown leads to several discussion points. First of all, this approach is more consistent with canopy variations; for this

reason, there is only a difference of 8 unidentified trees between the two datasets. Furthermore, the method is highly precise, with values approximately equal to 80%. On the other hand, recall is very low because many trees are not identified (recall value lower than 50%).

Table 31: Metrics of the segmentation procedure, FSCT algorithm - Case study C.

Case study C - Metrics	**FSCT**	
	2020	**2021**
$True_{Positive} + False_{Negative}$	184	184
$True_{Positive} + False_{Positive}$	107	94
$True_{Positive}$	86	78
Precision	0,80	0,83
Recall	0,46	0,42
F1 score	59%	56%

As predicted, the two methods behave similarly towards oblique point cloud processing. However, it was decided to proceed with subsequent processing, taking into consideration the results obtained through the PyCrown algorithm. This choice was motivated by the fact that it was deemed more appropriate to consider a methodology that allows identifying a greater number of trees to be more valid, albeit with a greater quantity of trees identified incorrectly. In this way (subject to the exclusion of incorrect trees), more consistent results are obtained as they refer to a larger sample.

The procedure of identifying the same tree at different time instants produced the results shown in Figure 75. Although no natural disturbance occurred in the time interval between the two acquisitions, the multi-temporal identification between individual trees does not concern all correctly identified trees. This phenomenon emphasizes that (i) different errors in terms of single tree detection and (ii) small variations in the scenario affect the final result.

Approximately 45% of all trees were correctly matched in both the LiDAR acquisitions, resulting in an F1 score equal to 57% and 56%). F1 score would reach approximately 70% if only correctly segmented trees were considered. In fact, as already pointed out in Paragraph 6.2, only correctly identified matched trees can result in a correct multi-temporal match.

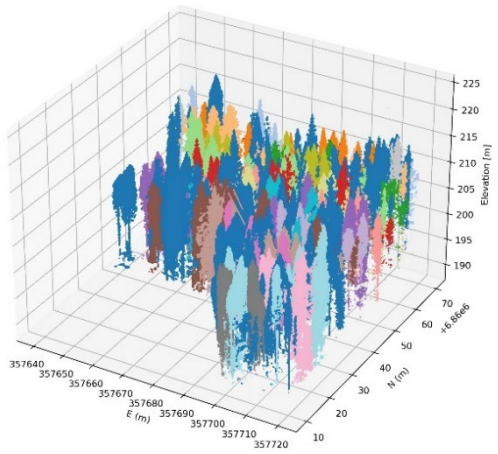


Figure 75: Segmented point cloud at single tree level with multitemporal match. The same trees are colored with the same colors; blue trees did not have a positive match. EPSG:3067.

Table 32: Metrics of the multitemporal detection of individual trees procedure - Case study C.

Case Study A - Metrics	**Multitemporal detection of individual trees**	
	2013	**2014**
$True_{Positive} + False_{Negative}$	184	184
$True_{Positive} + False_{Positive}$	142	151
$True_{Positive}$	84	93
Precision	0,59	0,62
Recall	0,46	0,51
F1 score	52%	56%

The reliability of this procedure is slightly lower than that obtained in Chapter 6 due to the significant difference in forest density (
Table 32).

8.2.2 Wood and leaf separation

The analyzes regarding the separation between trunk and leaf conducted in case study C are broader than those proposed in previous case studies. In fact, regarding approaches 1 and 2 (Table 6), in addition to the features already taken into consideration (Table 7), it was also decided to consider the deviation feature, which indicates how much the points in the cloud deviate from their average position. Therefore, the segmentation results performed through the first two approaches will be twofold: in one case, the deviation will also be considered, while in the other, it will not.

The analyzes were performed on 20 trees. All silver birch trees were chosen to be located homogeneously within the area under examination. The choice of tree species was dictated by the will to focus on one of the most widespread species in the Finnish national territory and northern Europe; additionally, it allowed species-specific discrepancies to be avoided. As regards the homogeneous spatial distribution, it was defined to investigate if any particularities can be caused by the distance and the specific perspective related to the single station point. Figure 76 shows the distribution of the selected trees in the study area.

Figure 77 and Figure 78 show the results of wood and leaf separation on a tree acquired during the first point cloud acquisition, while Figure 79 and Figure 80 refer to the same tree acquired during April 2021.

As a first point of discussion, it is noted that deviation is not a feature that affects the part segmentation result. The results obtained considering the deviation (trees II and IV in Figure 77 and Figure 78) are identical to those obtained without considering them (trees I and III in the same

figures). Furthermore, the segmentation conducted with approaches 1 and 2 yielded similar results. Generally speaking, the second approach correctly identifies a more significant number of woody points, while they are misclassified in the first approach. However, the woody structure is not fully identifiable. The main trunk is partially misclassified into leaf; as regards the branches, these are classified slightly better with the second approach.

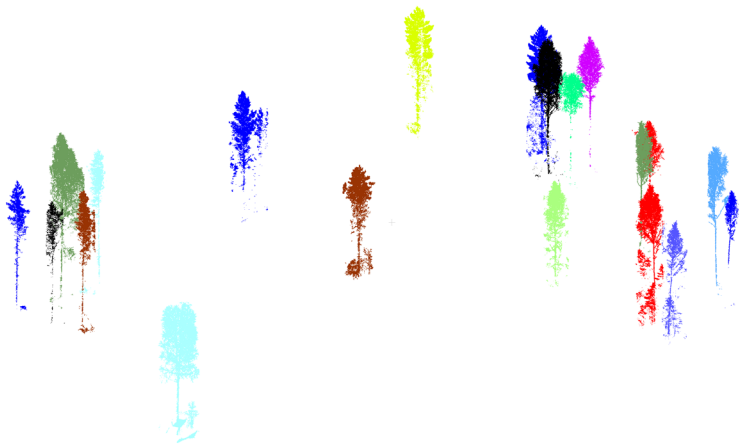
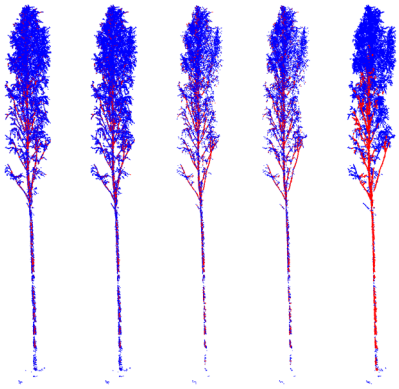


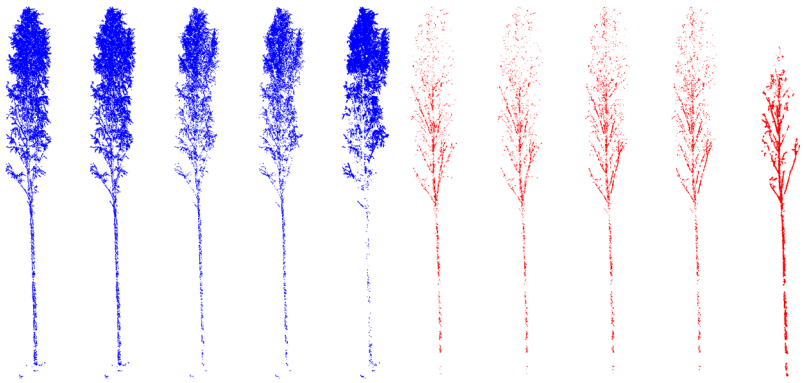
Figure 76: Individual silver birch trees selected for wood and leaf separation procedure. EPSG:3067.

On the contrary, the segmentation performed through the third method returns the best results. The woody structure in the lower part of the tree (trunk and branches) is correctly identified in its entirety, except for the smaller branches.



(I) (II) (III) (IV) (V)

Figure 77: Wood and leaf separation on the same tree: (I) first approach, without deviation; (II) first approach, with deviation; (III) second approach, without deviation; (IV) second approach, with deviation; (V) third approach - 2020 point cloud.



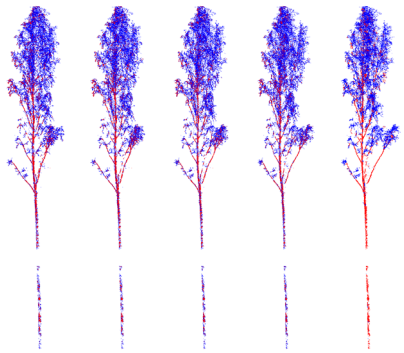
(a) (b)

(I) (II) (III) (IV) (V) (I) (II) (III) (IV) (V)

Figure 78: Wood and leaf separation on the same tree: (I) first approach, without deviation; (II) first approach, with deviation; (III) second approach, without deviation; (IV) second approach, with deviation; (V) third approach. (a) leaf component; (b) wood component - 2020 point cloud.

More outstanding critical issues are observed in the upper part of the foliage (Figure 78a), where the branches' size decreases and, simultaneously, the quantity of leaves increases.

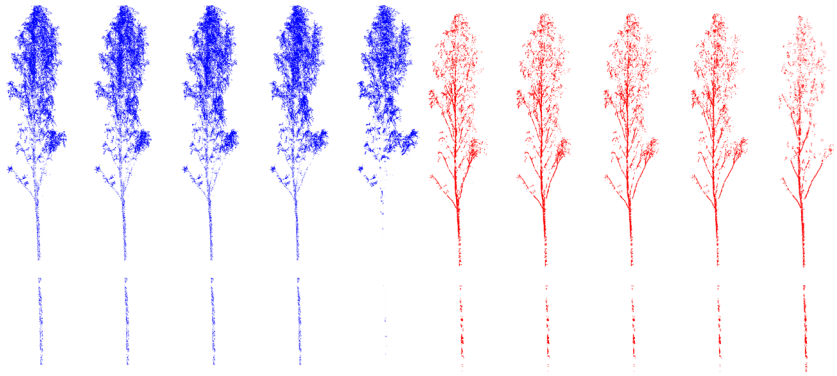
However, for modeling purposes using QSM algorithms, it is preferable not to have woody points scattered and disconnected from the main structure. This result causes an underestimation of the total volume; nevertheless, it represents a very low percentage and can be neglected. Figure 78 clarifies what has been stated so far.



(I) **(II)** **(III)** **(IV)** **(V)**

Figure 79: Wood and leaf separation on the same tree: (I) first approach, without deviation; (II) first approach, with deviation; (III) second approach, without deviation; (IV) second approach, with deviation; (V) third approach - 2021 point cloud.

For comparison, the segmentation results obtained on the same tree acquired during April 2021 are shown in Figure 79 and Figure 80. Segmented trees are similar and completely comparable. It can be observed that the foliage is less lush; therefore, a greater number of points are classified as branches. Some of these points are misclassified, as they are disconnected from the woody structure and would appear to be part of the crown.



(**a**) (**b**)

(**I**) (**II**) (**III**) (**IV**) (**V**) (**I**) (**II**) (**III**) (**IV**) (**V**)

Figure 80: Wood and leaf separation on the same tree: (I) first approach, without deviation; (II) first approach, with deviation; (III) second approach, without deviation; (IV) second approach, with deviation; (V) third approach. (a) leaf component; (b) wood component - 2021 point cloud.

Comparing the two datasets, it can also be noted that during the year between the two acquisitions, an obstruction caused a lack of data around the central part of the main trunk. Although this peculiarity, as will be discussed in the next Paragraph, does not particularly influence the aboveground biomass estimate, it once again points out the limitations of a single observation point.

8.2.3 Forest parameters (DBH and AGB) characterization

Estimates of forest parameters of 20 trees were explored (Table 33). Of these, only 12 were correctly identified and automatically matched between the two multi-temporal acquisitions (Table 35). Therefore, only these were considered in order to compare the estimates with those relating to the previous year's dataset. The diameter estimates obtained through Equation (19) were compared with manually measured values

after carefully interpreting the point cloud. The results are excellent. Considering the RMSE and Bias indexes, these are equal to approximately 5,4 cm and -10.9%, respectively. The empirical equation underestimates the diameter more than what happened in case study A; however, the value of RMSE is lower (Table 14). Similar values are obtained by analyzing the point cloud acquired in 2021, where the underestimation is approximately 8%, and the RMSE value slightly increases to 7.3 cm (Table 36).

Table 33: Estimated and reference values of forest parameters - 2020 point cloud.

Tree ID	H [m]	CD [m]	DBH_{Jucker} [cm]	DBH_{Ref} [cm]	Vol_{Kangas} [L]	Vol_{QSM} [L]	AGB_{Kangas} [kg]	AGB_{Chave} [kg]	AGB_{QSM} [kg]
1	20,49	3,9	20,6	17,7	255	170	130	187	87
2	20,1	3,0	16,2	27,7	588	111	300	439	57
3	25,15	5,2	30,4	29,4	786	897	401	614	457
4	18,06	3,4	16,5	28,8	581	926	296	427	472
5	19,05	4,4	21,2	25,8	493	344	251	363	175
6	20,68	3,8	20,1	22	389	292	199	288	149
7	18,83	4,2	20,2	24,1	429	564	219	314	288
8	18,72	4,8	22,4	22,8	384	604	196	280	308
9	16,5	4,6	19,7	19,9	266	716	136	190	365
10	19,92	5,1	25,0	21,5	361	440	184	265	224
11	19,43	4,3	21,3	15,2	182	101	93	132	52
12	15,96	4,6	19,1	28,7	523	66	267	376	34
13	15,92	3,2	14,2	19	236		121	168	
14	16,01	5,1	20,6	25,3	412	116	210	295	59
15	19,14	3,4	17,4	18,8	270	217	138	196	111
16	20,25	3,7	19,2	20,2	325	455	166	239	232
17	20,95	4,0	21,2	21	360	248	184	266	126
18	17,33	1,7	9,3	15,7	175	30	89	125	15
19	14,44	2,9	12,1	15,7	150	180	77	105	92
20	18,79	4,0	19,3	22,1	377	360	192	277	184

Regarding biomass estimates, the following were taken into consideration:

- the values obtained by multiplying the volumetric component obtained through the allometric equation (20 calibrated for the Finnish trees and the tree density (for simplicity, they will be called Kangas results);
- the values obtained in the same way, but taking into account the output of the QSM algorithms (QSM results);
- the AGB was estimated using the generic allometric equation (15), referred to as Chave results.

QSM results were validated according to Kangas and Chave's results (Table 34 and Table 36). As expected, the best validation is obtained with respect to the results obtained through Kangas' equation (RMSE 119 kg, bias -5,2% in the first acquisition, RMSE 111 kg and bias -5,1 % in the second dataset).

The results, although consistent with those obtained in case study A and case study B (considering TLS data), are the best among all the case studies presented.

Table 34: Statistical indexes for forest parameters - 2020 scenario.

Assessment	Reference	RMSE	BIAS
$\mathbf{AGB_{QSM}}$	AGB_{Kangas}	119 kg	-5,2 %
	AGB_{Chave}	162 kg	-33,8 %
$\mathbf{DBH_{Jucker}}$	$DBH_{Reference}$	5,4 cm	-10,9 %

The difference with respect to values obtained through the Chave equation places greater emphasis on the importance of using equations that are as case-specific as possible. In fact, with a generic equation, the assessment has little robustness.

Table 35: Estimated and reference values of forest parameters - 2021 scenario.

Tree ID	H [m]	CD [m]	DBH_{Jucker} [cm]	DBH_{Ref} [cm]	Vol_{Kangas} [L]	Vol_{QSM} [L]	AGB_{Kangas} [kg]	AGB_{Chave} [kg]	AGB_{QSM} [kg]
2	19,03	3,6	18,1	27,7	563	118	287	416	60
4	18,34	2,6	13,4	28,8	588	940	300	433	479
5	20,14	4,8	23,8	25,8	515	380	263	383	194
8	19,95	5,8	27,6	22,8	404	599	206	298	305
10	20,59	5,3	26,1	21,5	371	461	189	274	235
11	18,07	3,8	18,2	15,2	171	176	87	123	90
12	14,72	4,0	15,9	28,7	490	100	250	347	51
15	18,62	5,0	23,0	18,8	264	274	135	191	140
16	20,93	3,9	20,7	20,2	334	481	170	247	245
17	20,61	3,3	17,8	21	355	356	181	262	182
18	15,11	2,6	11,6	15,7	156	80	80	110	41
19	18,20	2,7	13,9	15,7	183	164	93	132	84

At the same time, these results are fundamental to state that, without specific allometric equations and using a dataset acquired correctly, the approach through the construction of QSM represents a valid procedure for deepening knowledge about tree biomass. Furthermore, the dataset's quality (i.e., the point cloud density and points homogeneous distribution) is also fundamental for modeling the tree structure itself. In fact, in this scenario, the errors in the construction of cylindrical and conical geometric shapes are reduced: the analysis of 20 trees (12 of which were described by two different datasets) produced only one modeling error (related to tree number 13); additionally, no unreasonable overestimation was obtained.

Finally, it is underlined that a partial cause of the underestimation of the biomass can be referred to as the upper branches of the canopy that were not identified during the wood and leaf separation procedure.

Table 36: Statistical indexes for forest parameters - 2021 scenario.

Assessment	Reference	RMSE	BIAS
$\mathbf{AGB_{QSM}}$	AGB_{Kangas}	111 kg	-5,1 %
	AGB_{Chave}	150 kg	-33,8 %
$\mathbf{DBH_{Jucker}}$	$DBH_{Reference}$	7,3 cm	-7,9 %

8.2.4 CO_2 assessment

The estimates of the CO_2 absorbed by the analyzed trees were conducted only with respect to the AGB estimates resulting from QSM analysis and Equation (20); the results of Equation (15) were excluded based on the considerations made in the previous Paragraph. Table 37 summarizes the results, while Table 38 summarizes the evaluation indexes. Appendix D reports the complete results.

Focusing on the trees analyzed in the two datasets, an increase in CO_2 stocked is observed (approximately equal to 1%). Higher growth

rates are observed on smaller trees; however, these have a low influence on the total absorbed CO_2. Indeed, the order of magnitude of the total amount of carbon dioxide stocked in the study area makes minor irrelevant variations.

Table 37: Estimated and reference values of stocked CO_2 - Case study C.

Tree ID	2020 acquisition		2021 acquisition	
	$CO_{2Kangas}$ [kg]	CO_{2QSM} [kg]	$CO_{2Kangas}$ [kg]	CO_{2QSM} [kg]
1	65	43		
2	150	28	144	30
3	200	229		
4	148	236	150	240
5	126	88	131	97
6	99	74		
7	109	144		
8	98	154	103	153
9	68	183		
10	92	112	95	118
11	46	26	44	45
12	133	17	125	26
13	60			
14	105	30		
15	69	55	67	70
16	83	116	85	123
17	92	63	91	91
18	45	8	40	20
19	38	46	47	42
20	96	139		

Therefore, an accurate estimate of the absorption capacity of the forest under examination is proposed using the QSM approach. In fact, with a simple proportion, it can be estimated that the area considered,

which includes 184 trees, absorbed approximately 16 thousand kilos of CO_2.

Regarding validation, the same considerations were conducted regarding the estimate of aboveground biomass, which can be applied to estimate stocked carbon dioxide. The underestimation, which, as already seen, can be caused by an intrinsic error of the reference volumetric equation and by the volume relating to the small branches not modeled, is negligible.

Table 38: Statistical indexes for stocked CO_2 - Case study C.

Assessment	2020 acquisition		2021 acquisition	
	RMSE	BIAS	RMSE	BIAS
CO_{2QSM}	60 kg	-2,6 %	55 kg	-5,1 %

8.3 Lesson learnt

The last case study analyzed and proposed in this thesis work is based on a particular and innovative dataset type. Several talking points were developed; the main lessons learned are briefly summarized below.

Both proposed single tree-level segmentation methodologies (raster-based and semantic segmentation) positively process the dataset, albeit with different characteristics. PyCrown identifies more trees; however, false positives are also higher. On the contrary, FSCT is more precise, but more trees are omitted. Although the values of the evaluation metric are comparable between the two algorithms, it was decided to continue with the processing results through PyCrown, assuming that it is preferable to identify the greatest number of trees possible at the expense of a greater number of misclassifications. Although the acquisitions were carried out with the same characteristics and referred to the same study area, some trees are classified correctly only in one of the two datasets. This peculiarity is due to forest characteristics (e.g., foliar cover, uneven

growth, particular atmospheric conditions), which should be explored in greater depth.

Regarding wood and leaf separation, the results observed in case study B were confirmed. The first two proposed approaches mainly identify the main trunk, while the third outlines the woody structure more thoroughly. Additional testing about the deviation feature has shown that it does not affect the final result. Since the wood and leaf segmentation procedure was tested for the first time on case study C, it was not tested on previous proposed studies.

Biomass assessment using QSM algorithms produced excellent results: they are comparable with those estimated through a case-specific volumetric equation, while a greater uncertainty results from the comparison with a generic allometric equation. However, case-specific allometric equations require considerable field study with relative validation through field biomass measurements; for this reason, these are not widely spread. The tested approach, however, proved to be highly valid and capable of replacing empirical estimates. However, the requirement is that the quality of the point cloud is excellent. The tree must, in fact, be totally acquired and described through a high-density point cloud. Minor inaccuracies in the modeling of the tree are observed around the smaller branches, which are often omitted. In any case, they have a relatively small biomass, and it is not incorrect to neglect them. It should be explored further whether the single perspective characteristic of this dataset could also cause this underestimation. In fact, some branches may be entirely in the shadow cone caused by larger branches or other trees. A quick visual interpretation, however, led to the rejection of this option.

The DBH estimates are affected by greater uncertainty but still acceptable in the forestry sector. The most significant degree of uncertainty derives from using a general and not case-specific allometric equation, as in the case of the AGB.

To conclude, the innovative acquisition methodology proposed by the Finnish research center represents an excellent solution for monitoring a portion of the forest over time. The fixed station point allows you to analyze point clouds with the same characteristics; the high acquisition density allows detailed analyses; multitemporality allows in-depth studies on forest growth. In one year, an increase of approximately 1% was recorded in the monitored trees; however, more significant gains were measured for younger trees with a faster growth rate. On the other hand, some limitations are caused by the single perspective, which sometimes does not allow for a complete view.

Chapter 9

CONCLUSIONS AND FUTURE VISION

This doctoral thesis aims to investigate LiDAR sensors and relative methodologies to perform forest parameters assessment at a single tree level with a semi-automatic procedure. The reasons that led to this research were the need to fill the scientific gap regarding a well-defined workflow, which, starting from data acquisition and through different processing phases, results in assessing the aboveground biomass and the carbon dioxide stocked by the wooden component of trees. In fact, from an in-depth study of the literature on point cloud processing technologies, methodologies and algorithms, the need to generalize and automate this working methodology emerged. The final goal was to provide an innovative procedure for estimating the carbon stocked within the woody biomass and evaluating the effects of changes over time.

The approach pursued in this work was developed based on three case studies that were different from each other in terms of forest

composition and characteristics. The LiDAR datasets were acquired according to different methodologies (terrestrial, aerial, oblique) and were subsequently processed through semi-automatic and artificial intelligence-based algorithms. Specifically, the topics relating to the automatic identification of individual trees, the multi-temporal identification of the same trees, the wood and leaf separation, and finally, the volumetric and biomass characterization were addressed. A unique workflow was proposed for processing algorithms and approaches for testing different data processing methodologies. The workflow proposed in this study represents a unique, complete and integrated process that describes each step to obtain the desired results.

The ITD was investigated through two completely different methodologies (Paragraphs 4.3.1 and 5.4.2), which yielded interesting results. It has been shown that raster-based algorithms (PyCrown [220], which was tested in this study) are more efficient than algorithms that perform semantic segmentation (FSCT [222]) when it comes to processing low-density aerial point clouds. (case study A) or when using oblique datasets in particular high-density scenarios (case study C). On the contrary, FSCT has proven to be the best approach when processing terrestrial point clouds, whether TLS or MLS (case study B). In Paragraph 6.2.1, the usefulness of the improvement introduced in this study regarding the PyCrown algorithm was also demonstrated, which increases the F1 score from 67% to 76%, and it was subsequently applied to the other case studies. In general, the F1 score values obtained with the segmentation algorithm best suited to each case study are in line with those proposed in the literature (from 76% to 78% in case study A through PyCrown algorithm, 68% and 86% relating to point cloud MLS and TLS of case study B respectively through FSCT algorithm, and finally 65% and 53% in case study C through PyCrown). In the last case study proposed, the values are slightly lower (but still acceptable in the forest domain to obtain rough estimates) due to the high forest density

and differ from each other due to varied forest characteristics (e.g., foliar cover, uneven growth, particular atmospheric conditions), to be deepened.

Regarding the datasets characterized by acquisitions that occurred at different time instants (case studies A and C), the multi-temporal identification procedure of the same tree was carried out as described in Paragraph 5.4.3. The results are acceptable (F1 score equal to 69% in case study A and between 56% and 59% in case study C) and made it possible to evaluate the temporal variation of forest parameters in conditions of external disturbance (ice storm in case study A), and under undisturbed conditions over an annual time interval (case study C). Results allow to state that LiDAR datasets could be used for large-scale forest monitoring purposes in undisturbed conditions and scenarios of pre and/or post-external disturbance, both human and natural.

The wood and leaf separation procedure was carried out using the approaches described in Paragraph 5.4.4. The methods differ in terms of the algorithms and the features considered. Evaluation of the results was based on visual interpretation. The procedure did not produce the desired results in case study A, for which the density of points was insufficient. On the contrary, the results for case studies B and C are satisfactory; in particular, the approach based on the unsupervised infliction threshold method has proven to produce the best results. Indeed, unlike the first two approaches, which mainly identify the main trunk, the woody structure can be identified almost wholly. Furthermore, in case study C, it was proven that the deviation feature does not affect the quality of the segmentation, so this was not explored in depth in the previous case studies.

Regarding the DBH, it was estimated through generic allometric equations and the results were validated using in situ measurements; the AGB, on the other hand, was estimated using both allometric equations (in case study C, reference was made to an equation derived explicitly to

have greater validity in the tree species of the Finnish territory) and through Quantitative Structure Models (Paragraph 5.4.5). In the absence of reference data, the results obtained through the various proposed methods were compared. The DBH estimates are affected by underestimation for case studies A and C, with RMSEs of approximately 8 - 9 cm for the first case study and 5,4 – 7,3 cm for the last one. No in-depth studies were carried out on case study B, for which ground truth data are unavailable. Concerning the AGB assessment, the results allowed for the deepening of several conclusions. The approach based on Quantitative Structure Models requires a complete and sufficiently dense description of the entire tree; for this reason, the dataset of case study A (for which the density of the point cloud is insufficient, and few points describe the part of the tree below the canopy) does not allow this approach to be used, making the results obtained unreliable (it overestimates of approximately one order of magnitude compared to the AGB assessed through allometric equations). Estimates obtained through allometric equations generally underestimate biomass compared to values calculated through the same equations in which the values of the variables are related to ground truth data (underestimates between -9.1% and -13.3%). The terrestrial acquisitions conducted in case study B showed overestimates in the AGB values obtained through the QSM approach of approximately 117 kg in the case of MLS point cloud and approximately 95 kg for TLS data, with percentages of roughly 40% and 6%, respectively. However, it should be noted, as extensively discussed in Chapter 7, that (i) reference values have been calculated using allometric equations, which do not account for any damages that trees may undergo due to external disturbance events, and that (ii) incomplete or partial tree acquisitions may cause the difference observed between the two types of datasets. Finally, case study C is the one where the QSM approach works best; indeed, compared to the estimated values through the specific equation calibrated for Finnish forests, slight

underestimations of about 5% are shown, corresponding to approximately 110 - 120 kg in both multi-temporal acquisitions. In this particular case, where no major disturbance event has occurred, it can be assumed that such estimates may be considered more accurate than those obtained through an empirical equation, net of minimum values not estimated due to the small size of the smaller and peripheral branches (Chapter 8).

Table 39: Summary of the results obtained in this study according to the case study and the task considered. ITD and multitemporal match were evaluated according to F1 score values; wood and leaf separation quality according to visual interpretation (* = awful result; ***** = excellent result).

Task		**Case study**		
		A	**B**	**C**
ITD	PyCrown	76% ÷ 78%	33% ÷ 44%	65% ÷ 53%
	FSCT	68% ÷ 69%	68% ÷ 86%	59% ÷ 56%
Multitemporal match		69%	No	56% ÷ 59 %
Wood and leaf separation		*	****	****
DBH	RMSE	8 ÷ 9 cm	No	5,4 ÷ 7,3 cm
	BIAS	-6,2% ÷ 8,7%		-10,9% ÷ -7,9%
AGB	RMSE	452 kg ÷ 239 kg	117 kg ÷ 95 kg	119 kg ÷ 111 kg
	BIAS	-9,1% ÷ -13,3%	39,8% ÷ 5,9%	-5,2% ÷ -5,1%
CO_2 variation (time)		-26% ÷ -34%	No	+1%
CO_2 variation (acquisition methodologies)		No	MLS < TLS 46%	No

Based on the characterization of the AGB, it was possible to estimate the CO_2 absorbed by the trees considered and extend this analysis to the areas investigated in their entirety. In case study A, the CO_2 loss was estimated to be approximately equal to 20 and 21,5 tons according to allometric equations (in percentage terms, loss between approximately

26% and 34%), while in case study C, an increase in stored CO_2 of approximately 1% was assessed. Finally, in case study B, a difference was observed in the estimate of CO_2 absorbed of roughly 1.6 tons relative to the 20 trees considered (approximately equal to 46% of the CO_2 estimated relative to the TLS dataset). Last but not least, the economic aspect and cost/benefit considerations must be taken into account. Generally speaking, handheld LiDARs are low cost sensors, while TLS sensors are more expensive. The formers are more manageable and easier to use, but they have a smaller field of view than the latter, which on the contrary produce a cloud of points that is usually denser. Instead, relying on companies that provide the product of a lidar acquisition from an airplane or helicopter has a variable cost depending on the extension of the area, but does not require the presence in situ of operators (in conditions of natural disasters it is essential to safeguard their safety); however, it may not be the most suitable choice if the final goal is to constantly monitor a large area with multiple acquisitions.

A summary of the main results obtained in this study is summarized in Table 39.

In conclusion, this study demonstrated the usefulness and versatility of LiDAR point clouds for forest monitoring purposes. LiDAR has definitely emerged increasingly in recent years within the Remote Sensing domain. Nowadays, it stands out as the most complete approach for acquiring high-precision 3D data in various conditions and scenarios. Also, the enormous amount of data that LiDAR acquires makes it essential to develop automated or semi-automated data-processing methodologies.

However, further aspects should be investigated. Firstly, the validation procedure must be independent of human intervention to reduce the processing time. Furthermore, a notable improvement would be achieved through an innovative ITD approach to eliminate misclassifications and work effectively with all datasets. Also, this

approach can be adopted to study forests with similar characteristics, choosing the sensors and the operational procedures most suitable for the case study; nevertheless, the validity of this workflow for substantially different forests should be explored. For trees which grow mainly in height, whose treetop is easily identifiable, as well as for low forest densities, such that the crowns of the trees do not intersect each other, or there are no submissive trees, the proposed workflow is optimal. However, it may not be the right approach when considering trees with a globular or expanded shape, whose crown does not univocally define a treetop; further difficulties would be caused by higher forest densities. In these cases, further segmentation approaches should be explored.

REFERENCES

[1] A. Crimmins *et al.*, "The Impacts of Climate Change on Human Health in the United States: A Scientific Assessment," U.S. Global Change Research Program, 2016. doi: 10.7930/J0R49NQX.

[2] T. S. Ledley, E. T. Sundquist, S. E. Schwartz, D. K. Hall, J. D. Fellows, and T. L. Killeen, "Climate change and greenhouse gases," *EoS Transactions*, vol. 80, no. 39, pp. 453–458, Sep. 1999, doi: 10.1029/99EO00325.

[3] P. Friedlingstein *et al.*, "Global Carbon Budget 2020," *Earth Syst. Sci. Data*, vol. 12, no. 4, pp. 3269–3340, Dec. 2020, doi: 10.5194/essd-12-3269-2020.

[4] J. S. Estrada, A. Fuentes, P. Reszka, and F. Auat Cheein, "Machine learning assisted remote forestry health assessment: a comprehensive state of the art review," *Front. Plant Sci.*, vol. 14, p. 1139232, Jun. 2023, doi: 10.3389/fpls.2023.1139232.

[5] A. Hamedianfar, C. Mohamedou, A. Kangas, and J. Vauhkonen, "Deep learning for forest inventory and planning: a critical review on the remote sensing approaches so far and prospects for further applications," *Forestry: An International Journal of Forest Research*, vol. 95, no. 4, pp. 451–465, Aug. 2022, doi: 10.1093/forestry/cpac002.

[6] Intergovernmental Panel On Climate Change (Ipcc), *Climate Change 2022 – Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, 1st ed. Cambridge University Press, 2023. doi: 10.1017/9781009325844.

[7] "UNFCCC, 1992: United Nations Framework Convention On Climate Change. United Nations, FCCC/INFORMAL/84 GE.05-62220 (E) 200705, Secretariat of the United Nations Framework

Convention on Climate Change, Bonn, Germany, 24 pp., unfccc.int/resource/docs/convkp/conveng.pdf."

[8] "Copernicus 'New high-quality CAMS maps of carbon dioxide surface fluxes obtained from satellite observations'. [Online] Accessed: March 1, 2024. https://atmosphere.copernicus.eu/new-high-quality-cams-maps-carbon-dioxide-surface-fluxes-obtained-satellite-observations."

[9] "NASA Earth Observatory. '2022 Tied for Fifth Warmest Year on Record' Accessed: Feb 15, 2024. [Online]. https://earthobservatory.nasa.gov/images/150828/2022-tied-for-fifth-warmest-year-on-record."

[10] D. W. Fahey *et al.*, "Ch. 2: Physical Drivers of Climate Change. Climate Science Special Report: Fourth National Climate Assessment, Volume I," U.S. Global Change Research Program, 2017. doi: 10.7930/J0513WCR.

[11] The National Academy of Sciences and The Royal Society, *Climate Change: Evidence and Causes: Update 2020*. Washington, D.C.: National Academies Press, 2020, p. 25733. doi: 10.17226/25733.

[12] "NOOA. 'Trends in Atmospheric Carbon Dioxide.' Accessed: Dic. 20, 2023. [Online]. Available: https://gml.noaa.gov/ccgg/trends/global.html."

[13] L. Hockstad and L. Hanel, "Inventory of U.S. Greenhouse Gas Emissions and Sinks." Environmental System Science Data Infrastructure for a Virtual Ecosystem, 2018. doi: 10.15485/1464240.

[14] M. N. Hayek and S. M. Miller, "Underestimates of methane from intensively raised animals could undermine goals of sustainable development," *Environ. Res. Lett.*, vol. 16, no. 6, p. 063006, Jun. 2021, doi: 10.1088/1748-9326/ac02ef.

[15] "FAO. 'Global Forest Resources Assessment (FRA)'. Accessed: Dec. 21, 2023. [Online]. Available: https://www.fao.org/3/I8661EN/i8661en.pdf."

[16] Intergovernmental Panel On Climate Change, *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on*

Climate Change, 1st ed. Cambridge University Press, 2023. doi: 10.1017/9781009157896.

[17] *FRA 2020 Remote Sensing Survey*. FAO, 2022. doi: 10.4060/cb9970en.

[18] T. A. M. Pugh *et al.*, "Role of forest regrowth in global carbon sink dynamics," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 116, no. 10, pp. 4382–4387, Mar. 2019, doi: 10.1073/pnas.1810512116.

[19] NOAA National Centers for Environmental Information, "Monthly Global Climate Report for Annual 2022, published online January 2023, retrieved on December 4, 2023 from https://www.ncei.noaa.gov/access/monitoring/monthly-report/global/202213."

[20] K. D. Lafferty, "The ecology of climate change and infectious diseases," *Ecology*, vol. 90, no. 4, pp. 888–900, Apr. 2009, doi: 10.1890/08-0079.1.

[21] G. D'Amato, L. Cecchi, M. D'Amato, and I. Annesi-Maesano, "Climate change and respiratory diseases," *European Respiratory Review*, vol. 23, no. 132, pp. 161–169, Jun. 2014, doi: 10.1183/09059180.00001714.

[22] V. Iglesias *et al.*, "Risky Development: Increasing Exposure to Natural Hazards in the United States," *Earth's Future*, vol. 9, no. 7, p. e2020EF001795, Jul. 2021, doi: 10.1029/2020EF001795.

[23] Z. W. Kundzewicz, I. Pińskwar, and G. R. Brakenridge, "Changes in river flood hazard in Europe: a review," *Hydrology Research*, vol. 49, no. 2, pp. 294–302, Apr. 2018, doi: 10.2166/nh.2017.016.

[24] S. L. Gariano and F. Guzzetti, "Landslides in a changing climate," *Earth-Science Reviews*, vol. 162, pp. 227–252, Nov. 2016, doi: 10.1016/j.earscirev.2016.08.011.

[25] C. A. Williams, H. Gu, R. MacLean, J. G. Masek, and G. J. Collatz, "Disturbance and the carbon balance of US forests: A quantitative review of impacts from harvests, fires, insects, and droughts," *Global and Planetary Change*, vol. 143, pp. 66–80, Aug. 2016, doi: 10.1016/j.gloplacha.2016.06.002.

[26] K. Ostad-Ali-Askar, R. Su, and L. Liu, "Water resources and climate change," *Journal of Water and Climate Change*, vol. 9, no. 2, pp. 239–239, Jun. 2018, doi: 10.2166/wcc.2018.999.

[27] A. J. Watson *et al.*, "Revised estimates of ocean-atmosphere CO2 flux are consistent with ocean carbon inventory," *Nat Commun*, vol. 11, no. 1, p. 4422, Sep. 2020, doi: 10.1038/s41467-020-18203-3.

[28] H. S. Findlay and C. Turley, "Ocean acidification and climate change," in *Climate Change*, Elsevier, 2021, pp. 251–279. doi: 10.1016/B978-0-12-821575-3.00013-X.

[29] R. Hock and M. Huss, "Glaciers and climate change," in *Climate Change*, Elsevier, 2021, pp. 157–176. doi: 10.1016/B978-0-12-821575-3.00009-8.

[30] X. Zhang and X. Yan, "Spatiotemporal change in geographical distribution of global climate types in the context of climate warming," *Clim Dyn*, vol. 43, no. 3–4, pp. 595–605, Aug. 2014, doi: 10.1007/s00382-013-2019-y.

[31] "Collins, M., R. Knutti, J. Arblaster, J.-L. Dufresne, T. Fichefet, P. Friedlingstein, X. Gao, W.J. Gutowski, T. Johns, G. Krinner, M. Shongwe, C. Tebaldi, A.J. Weaver and M. Wehner, 2013: Long-term Climate Change: Projections, Commitments and Irreversibility. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA."

[32] P. N. Edwards, "History of climate modeling," *WIREs Climate Change*, vol. 2, no. 1, pp. 128–139, Jan. 2011, doi: 10.1002/wcc.95.

[33] S. L. Grotch and M. C. MacCracken, "The Use of General Circulation Models to Predict Regional Climatic Change," *J. Climate*, vol. 4, no. 3, pp. 286–303, Mar. 1991, doi: 10.1175/1520-0442(1991)004<0286:TUOGCM>2.0.CO;2.

[34] C. C. Ibebuchi, "On the representation of atmospheric circulation modes in regional climate models over Western Europe," *Intl Journal of Climatology*, vol. 43, no. 1, pp. 668–682, Jan. 2023, doi: 10.1002/joc.7807.

[35] M. Vrac, S. Thao, and P. Yiou, "Changes in temperature–precipitation correlations over Europe: are climate models reliable?," *Clim Dyn*, vol. 60, no. 9–10, pp. 2713–2733, May 2023, doi: 10.1007/s00382-022-06436-5.

[36] M. Pourali, M. R. Kavianpour, B. Kamranzad, and M. J. Alizadeh, "Future variability of wave energy in the Gulf of Oman using a high resolution CMIP6 climate model," *Energy*, vol. 262, p. 125552, Jan. 2023, doi: 10.1016/j.energy.2022.125552.

[37] J. H. Richter *et al.*, "Response of the Quasi-Biennial Oscillation to a warming climate in global climate models," *Quart J Royal Meteoro Soc*, vol. 148, no. 744, pp. 1490–1518, Apr. 2022, doi: 10.1002/qj.3749.

[38] E. Guilyardi *et al.*, "Understanding El Niño in Ocean–Atmosphere General Circulation Models: Progress and Challenges," *Bull. Amer. Meteor. Soc.*, vol. 90, no. 3, pp. 325–340, Mar. 2009, doi: 10.1175/2008BAMS2387.1.

[39] "Kyoto Protocol to the United Nations Framework Convention on Climate Change, Dec. 10, 1997, 2303 U.N.T.S. 162."

[40] "Paris Agreement to the United Nations Framework Convention on Climate Change, Dec. 12, 2015, T.I.A.S. No. 16-1104."

[41] N. Gruber *et al.*, "The oceanic sink for anthropogenic CO_2 from 1994 to 2007," *Science*, vol. 363, no. 6432, pp. 1193–1199, Mar. 2019, doi: 10.1126/science.aau5153.

[42] K. J. Kroeker *et al.*, "Impacts of ocean acidification on marine organisms: quantifying sensitivities and interaction with warming," *Global Change Biology*, vol. 19, no. 6, pp. 1884–1896, Jun. 2013, doi: 10.1111/gcb.12179.

[43] Committee on Developing a Research Agenda for Carbon Dioxide Removal and Reliable Sequestration *et al.*, *Negative Emissions Technologies and Reliable Sequestration: A Research Agenda*. Washington, D.C.: National Academies Press, 2019, p. 25259. doi: 10.17226/25259.

[44] H. K. Jeswani, D. M. Saharudin, and A. Azapagic, "Environmental sustainability of negative emissions technologies: A review," *Sustainable Production and Consumption*, vol. 33, pp. 608–635, Sep. 2022, doi: 10.1016/j.spc.2022.06.028.

[45] J. C. M. Pires, "Negative emissions technologies: A complementary solution for climate change mitigation," *Science of The Total Environment*, vol. 672, pp. 502–514, Jul. 2019, doi: 10.1016/j.scitotenv.2019.04.004.

[46] A. Chabi, S. Lautenbach, V. O. A. Orekan, and N. Kyei-Baffour, "Allometric models and aboveground biomass stocks of a West African Sudan Savannah watershed in Benin," *Carbon Balance Manage*, vol. 11, no. 1, p. 16, Dec. 2016, doi: 10.1186/s13021-016-0058-5.

[47] D. E. Reichle, "Ecosystem productivity," in *The Global Carbon Cycle and Climate Change*, Elsevier, 2023, pp. 197–232. doi: 10.1016/B978-0-443-18775-9.00007-3.

[48] C. Vidal, I. A. Alberdi, L. Hernández Mateo, and J. J. Redmond, Eds., *National Forest Inventories*. Cham: Springer International Publishing, 2016. doi: 10.1007/978-3-319-44015-6.

[49] M. A. Wulder, "RESOURCE ASSESSMENT | GIS and Remote Sensing," in *Encyclopedia of Forest Sciences*, Elsevier, 2004, pp. 997–1001. doi: 10.1016/B0-12-145160-7/00161-7.

[50] S. Liang and J. Wang, Eds., *Advanced remote sensing: terrestrial information extraction and applications*, Second edition. Amsterdam: Academic Press, 2020.

[51] X. Zhang, A. Duan, J. Zhang, and C. Xiang, "Estimating Tree Height-Diameter Models with the Bayesian Method," *The Scientific World Journal*, vol. 2014, pp. 1–9, 2014, doi: 10.1155/2014/683691.

[52] H. Temesgen and K. V. Gadow, "Generalized height–diameter models—an application for major tree species in complex stands of interior British Columbia," *Eur J Forest Res*, vol. 123, no. 1, pp. 45–51, Apr. 2004, doi: 10.1007/s10342-004-0020-z.

[53] L. Mehtätalo, S. de-Miguel, and T. G. Gregoire, "Modeling height-diameter curves for prediction," *Can. J. For. Res.*, vol. 45, no. 7, pp. 826–837, Jul. 2015, doi: 10.1139/cjfr-2015-0054.

[54] A. Kangas, T. P. Pitkänen, L. Mehtätalo, and J. Heikkinen, "Mixed linear and non-linear tree volume models with regional parameters to main tree species in Finland," *Forestry: An International Journal of Forest Research*, vol. 96, no. 2, pp. 188–206, Mar. 2023, doi: 10.1093/forestry/cpac038.

[55] T. Jucker *et al.*, "Allometric equations for integrating remote sensing imagery into forest monitoring programmes," *Glob Change Biol*, vol. 23, no. 1, pp. 177–190, Jan. 2017, doi: 10.1111/gcb.13388.

[56] J. Návar, "Allometric equations for tree species and carbon stocks for forests of northwestern Mexico," *Forest Ecology and Management*, vol. 257, no. 2, pp. 427–434, Jan. 2009, doi: 10.1016/j.foreco.2008.09.028.

[57] A. Komiyama, S. Poungparn, and S. Kato, "Common allometric equations for estimating the tree weight of mangroves," *J. Trop. Ecol.*, vol. 21, no. 4, pp. 471–477, Jul. 2005, doi: 10.1017/S0266467405002476.

[58] T. G. Cole and J. J. Ewel, "Allometric equations for four valuable tropical tree species," *Forest Ecology and Management*, vol. 229, no. 1–3, pp. 351–360, Jul. 2006, doi: 10.1016/j.foreco.2006.04.017.

[59] T. Mengesha, M. Hawkins, and M. Nieuwenhuis, "Validation of terrestrial laser scanning data using conventional forest inventory methods," *Eur J Forest Res*, vol. 134, no. 2, pp. 211–222, Mar. 2015, doi: 10.1007/s10342-014-0844-0.

[60] H. Mäkelä, H. Hirvelä, T. Nuutinen, and L. Kärkkäinen, "Estimating forest data for analyses of forest production and utilization possibilities at local level by means of multi-source National Forest Inventory," *Forest Ecology and Management*, vol. 262, no. 8, pp. 1345–1359, Oct. 2011, doi: 10.1016/j.foreco.2011.06.027.

[61] T. M. Lillesand and R. W. Kiefer, *Remote sensing and image interpretation*, 4th ed. New York: John Wiley & Sons, 2000.

[62] J. Estes, K. Kline, and E. Collins, "Remote Sensing," in *International Encyclopedia of the Social & Behavioral Sciences*, Elsevier, 2001, pp. 13144–13150. doi: 10.1016/B0-08-043076-7/02526-2.

[63] ""Natural Resources Canada. 'Fundamentals of Remote Sensing.' Accessed: Dec. 11, 2023. [Online]. Available: https://natural-resources.canada.ca/sites/www.nrcan.gc.ca/files/earthsciences/pdf/resource/tutor/fundam/pdf/fundamentals_e.pdf."

[64] J. Iglhaut, C. Cabo, S. Puliti, L. Piermattei, J. O'Connor, and J. Rosette, "Structure from Motion Photogrammetry in Forestry: a Review," *Curr Forestry Rep*, vol. 5, no. 3, pp. 155–168, Sep. 2019, doi: 10.1007/s40725-019-00094-3.

[65] C. B. Engel, S. D. Jones, and K. J. Reinke, "Real-Time Detection of Daytime and Night-Time Fire Hotspots from Geostationary Satellites," *Remote Sensing*, vol. 13, no. 9, p. 1627, Apr. 2021, doi: 10.3390/rs13091627.

[66] L. Zhu, J. Suomalainen, J. Liu, J. Hyyppä, H. Kaartinen, and H. Haggren, "A Review: Remote Sensing Sensors," in *Multi-purposeful Application of Geospatial Data*, R. B. Rustamov, S. Hasanova, and M. H. Zeynalova, Eds., InTech, 2018. doi: 10.5772/intechopen.71049.

[67] J. Uuttera, A. Haara, T. Tokola, and M. Maltamo, "Determination of the spatial distribution of trees from digital aerial photographs," *Forest Ecology and Management*, vol. 110, no. 1–3, pp. 275–282, Oct. 1998, doi: 10.1016/S0378-1127(98)00292-8.

[68] J. Bohlin, J. Wallerman, and J. E. S. Fransson, "Forest variable estimation using photogrammetric matching of digital aerial images in combination with a high-resolution DEM," *Scandinavian Journal of Forest Research*, vol. 27, no. 7, pp. 692–699, Oct. 2012, doi: 10.1080/02827581.2012.686625.

[69] D. J. N. Young, M. J. Koontz, and J. Weeks, "Optimizing aerial imagery collection and processing parameters for drone-based individual tree mapping in structurally complex conifer forests," *Methods Ecol Evol*, vol. 13, no. 7, pp. 1447–1463, Jul. 2022, doi: 10.1111/2041-210X.13860.

[70] M. Alonzo, H.-E. Andersen, D. Morton, and B. Cook, "Quantifying Boreal Forest Structure and Composition Using UAV Structure from Motion," *Forests*, vol. 9, no. 3, p. 119, Mar. 2018, doi: 10.3390/f9030119.

[71] M. Smigaj, R. Gaulton, S. L. Barr, and J. C. Suárez, "UAV-BORNE THERMAL IMAGING FOR FOREST HEALTH MONITORING: DETECTION OF DISEASE-INDUCED CANOPY TEMPERATURE INCREASE," *Int. Arch. Photogramm.*

Remote Sens. Spatial Inf. Sci., vol. XL-3/W3, pp. 349–354, Aug. 2015, doi: 10.5194/isprsarchives-XL-3-W3-349-2015.

[72] M. Sadi, Y. Zhang, W.-F. Xie, and F. M. A. Hossain, "Forest Fire Detection and Localization Using Thermal and Visual Cameras," in *2021 International Conference on Unmanned Aircraft Systems (ICUAS)*, Athens, Greece: IEEE, Jun. 2021, pp. 744–749. doi: 10.1109/ICUAS51884.2021.9476865.

[73] J. Shin, W. Seo, T. Kim, J. Park, and C. Woo, "Using UAV Multispectral Images for Classification of Forest Burn Severity—A Case Study of the 2019 Gangneung Forest Fire," *Forests*, vol. 10, no. 11, p. 1025, Nov. 2019, doi: 10.3390/f10111025.

[74] E. Belcore, M. Pittarello, A. M. Lingua, and M. Lonati, "Mapping Riparian Habitats of Natura 2000 Network (91E0*, 3240) at Individual Tree Level Using UAV Multi-Temporal and Multi-Spectral Data," *Remote Sensing*, vol. 13, no. 9, p. 1756, Apr. 2021, doi: 10.3390/rs13091756.

[75] S. E. Franklin, O. S. Ahmed, and G. Williams, "Northern Conifer Forest Species Classification Using Multispectral Data Acquired from an Unmanned Aerial Vehicle," *photogramm eng remote sensing*, vol. 83, no. 7, pp. 501–507, Jul. 2017, doi: 10.14358/PERS.83.7.501.

[76] E. Honkavaara *et al.*, "USING MULTITEMPORAL HYPER- AND MULTISPECTRAL UAV IMAGING FOR DETECTING BARK BEETLE INFESTATION ON NORWAY SPRUCE," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XLIII-B3-2020, pp. 429–434, Aug. 2020, doi: 10.5194/isprs-archives-XLIII-B3-2020-429-2020.

[77] V. González-Jaramillo, A. Fries, and J. Bendix, "AGB Estimation in a Tropical Mountain Forest (TMF) by Means of RGB and Multispectral Images Using an Unmanned Aerial Vehicle (UAV)," *Remote Sensing*, vol. 11, no. 12, p. 1413, Jun. 2019, doi: 10.3390/rs11121413.

[78] A. L. Mahood *et al.*, "Ten simple rules for working with high resolution remote sensing data," *Peer Community Journal*, vol. 3, p. e4, Jan. 2023, doi: 10.24072/pcjournal.223.

[79] M. R. Galán and R. S. Bard, Eds., *Studies in Archaeometry: Proceedings of the Archaeometry Symposium at NORM 2019, June 16-19, Portland, Oregon, Portland State University. Dedicated to the Rev. H. Richard Rutherford, C.S.C., Ph.D*. Archaeopress Publishing Ltd, 2020. doi: 10.2307/j.ctv1zckxdm.

[80] H. Aghababaei *et al.*, "Forest SAR Tomography: Principles and Applications," *IEEE Geosci. Remote Sens. Mag.*, vol. 8, no. 2, pp. 30–45, Jun. 2020, doi: 10.1109/MGRS.2019.2963093.

[81] "GIM International 'Recent Developments in Airborne Lidar.' Accessed: Dic. 19, 2023. [Online]. Available: https://www.gim-international.com/content/article/recent-developments-in-airborne-lidar-2."

[82] G. A. Idrobo-Pizo, J. M. S. T. Motta, and R. C. Sampaio, "A Calibration Method for a Laser Triangulation Scanner Mounted on a Robot Arm for Surface Mapping," *Sensors*, vol. 19, no. 8, p. 1783, Apr. 2019, doi: 10.3390/s19081783.

[83] N. Pfeifer, G. Mandlburger, and P. Glira, "Laserscanning," in *Handbuch der Geodäsie*, W. Freeden and R. Rummel, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 1–51. doi: 10.1007/978-3-662-46900-2_44-1.

[84] J. Kim, J. Jeong, Y.-S. Shin, Y. Cho, H. Roh, and A. Kim, "LiDAR configuration comparison for urban mapping system," in *2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, Jeju: IEEE, Jun. 2017, pp. 854–857. doi: 10.1109/URAI.2017.7992845.

[85] W. Y. Yan, A. Shaker, and N. El-Ashmawy, "Urban land cover classification using airborne LiDAR data: A review," *Remote Sensing of Environment*, vol. 158, pp. 295–310, Mar. 2015, doi: 10.1016/j.rse.2014.11.001.

[86] W. Wagner, M. Hollaus, C. Briese, and V. Ducic, "3D vegetation mapping using small-footprint full-waveform airborne laser scanners," *International Journal of Remote Sensing*, vol. 29, no. 5, pp. 1433–1452, Mar. 2008, doi: 10.1080/01431160701736398.

[87] M. Machala and L. Zejdová, "Forest Mapping Through Object-based Image Analysis of Multispectral and LiDAR Aerial Data,"

European Journal of Remote Sensing, vol. 47, no. 1, pp. 117–131, Jan. 2014, doi: 10.5721/EuJRS20144708.

[88] Y. Li and J. Ibanez-Guzman, "Lidar for Autonomous Driving: The Principles, Challenges, and Trends for Automotive Lidar and Perception Systems," *IEEE Signal Process. Mag.*, vol. 37, no. 4, pp. 50–61, Jul. 2020, doi: 10.1109/MSP.2020.2973615.

[89] Ø. D. Trier, J. H. Reksten, and K. Løseth, "Automated mapping of cultural heritage in Norway from airborne lidar data using faster R-CNN," *International Journal of Applied Earth Observation and Geoinformation*, vol. 95, p. 102241, Mar. 2021, doi: 10.1016/j.jag.2020.102241.

[90] A. Filisetti, A. Marouchos, A. Martini, T. Martin, and S. Collings, "Developments and applications of underwater LiDAR systems in support of marine science," in *OCEANS 2018 MTS/IEEE Charleston*, Charleston, SC: IEEE, Oct. 2018, pp. 1–10. doi: 10.1109/OCEANS.2018.8604547.

[91] V. Mitra, Chia-Jiu Wang, and S. Banerjee, "Lidar detection of underwater objects using a neuro-SVM-based architecture," *IEEE Trans. Neural Netw.*, vol. 17, no. 3, pp. 717–731, May 2006, doi: 10.1109/TNN.2006.873279.

[92] M. Beland *et al.*, "On promoting the use of lidar systems in forest ecosystem research," *Forest Ecology and Management*, vol. 450, p. 117484, Oct. 2019, doi: 10.1016/j.foreco.2019.117484.

[93] K. Zhao, J. C. Suarez, M. Garcia, T. Hu, C. Wang, and A. Londo, "Utility of multitemporal lidar for forest and carbon monitoring: Tree growth, biomass dynamics, and carbon flux," *Remote Sensing of Environment*, vol. 204, pp. 883–897, Jan. 2018, doi: 10.1016/j.rse.2017.09.007.

[94] M. A. Wulder *et al.*, "Lidar sampling for large-area forest characterization: A review," *Remote Sensing of Environment*, vol. 121, pp. 196–209, Jun. 2012, doi: 10.1016/j.rse.2012.02.001.

[95] L. Cao *et al.*, "Estimation of forest biomass dynamics in subtropical forests using multi-temporal airborne LiDAR data," *Remote Sensing of Environment*, vol. 178, pp. 158–171, Jun. 2016, doi: 10.1016/j.rse.2016.03.012.

[96] M. Dalponte, T. Jucker, S. Liu, L. Frizzera, and D. Gianelle, "Characterizing forest carbon dynamics using multi-temporal lidar data," *Remote Sensing of Environment*, vol. 224, pp. 412–420, Apr. 2019, doi: 10.1016/j.rse.2019.02.018.

[97] D. R. A. Almeida *et al.*, "Monitoring the structure of forest restoration plantations with a drone-lidar system," *International Journal of Applied Earth Observation and Geoinformation*, vol. 79, pp. 192–198, Jul. 2019, doi: 10.1016/j.jag.2019.03.014.

[98] W.-L. Lu, J. J. Little, A. Sheffer, and H. Fu, "Deforestation: Extracting 3D Bare-Earth Surface from Airborne LiDAR Data," in *2008 Canadian Conference on Computer and Robot Vision*, Windsor, Ont.: IEEE, May 2008, pp. 203–210. doi: 10.1109/CRV.2008.36.

[99] J. F. Weishampel, J. N. Hightower, A. F. Chase, and D. Z. Chase, "Use of Airborne LiDAR to Delineate Canopy Degradation and Encroachment along the Guatemala-Belize Border," *Tropical Conservation Science*, vol. 5, no. 1, pp. 12–24, Mar. 2012, doi: 10.1177/194008291200500103.

[100] A. B. Utkin, A. V. Lavrov, L. Costa, F. Simões, and R. Vilar, "Detection of small forest fires by lidar," *Applied Physics B: Lasers and Optics*, vol. 74, no. 1, pp. 77–83, Jan. 2002, doi: 10.1007/s003400100772.

[101] M. Alonzo, D. C. Morton, B. D. Cook, H.-E. Andersen, C. Babcock, and R. Pattison, "Patterns of canopy and surface layer consumption in a boreal forest fire from repeat airborne lidar," *Environ. Res. Lett.*, vol. 12, no. 6, p. 065004, May 2017, doi: 10.1088/1748-9326/aa6ade.

[102] X. Guo, N. C. Coops, P. Tompalski, S. E. Nielsen, C. W. Bater, and J. John Stadt, "Regional mapping of vegetation structure for biodiversity monitoring using airborne lidar data," *Ecological Informatics*, vol. 38, pp. 50–61, Mar. 2017, doi: 10.1016/j.ecoinf.2017.01.005.

[103] B. Ferreira, R. G. Silva, and M. Iten, "Earth Observation Satellite Imagery Information Based Decision Support Using Machine Learning," *Remote Sensing*, vol. 14, no. 15, p. 3776, Aug. 2022, doi: 10.3390/rs14153776.

[104] Q. Zhao *et al.*, "An Overview of the Applications of Earth Observation Satellite Data: Impacts and Future Trends," *Remote Sensing*, vol. 14, no. 8, p. 1863, Apr. 2022, doi: 10.3390/rs14081863.

[105] K. Johansen, N. C. Coops, S. E. Gergel, and Y. Stange, "Application of high spatial resolution satellite imagery for riparian and forest ecosystem classification," *Remote Sensing of Environment*, vol. 110, no. 1, pp. 29–44, Sep. 2007, doi: 10.1016/j.rse.2007.02.014.

[106] S. Tak and A. K. Keshari, "Investigating mass balance of Parvati glacier in Himalaya using satellite imagery based model," *Sci Rep*, vol. 10, no. 1, p. 12211, Jul. 2020, doi: 10.1038/s41598-020-69203-8.

[107] T. Toutin and L. Gray, "State-of-the-art of elevation extraction from satellite SAR data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 55, no. 1, pp. 13–33, Feb. 2000, doi: 10.1016/S0924-2716(99)00039-8.

[108] F. Cigna *et al.*, "Exploiting satellite SAR for archaeological prospection and heritage site protection," *Geo-spatial Information Science*, pp. 1–26, Jul. 2023, doi: 10.1080/10095020.2023.2223603.

[109] H. J. Zwally *et al.*, "ICESat's laser measurements of polar ice, atmosphere, ocean, and land," *Journal of Geodynamics*, vol. 34, no. 3–4, pp. 405–445, Oct. 2002, doi: 10.1016/S0264-3707(02)00042-X.

[110] Ľ. Kovanič, P. Blistan, M. Štroner, R. Urban, and M. Blistanova, "Suitability of Aerial Photogrammetry for Dump Documentation and Volume Determination in Large Areas," *Applied Sciences*, vol. 11, no. 14, p. 6564, Jul. 2021, doi: 10.3390/app11146564.

[111] M. Storch, T. Jarmer, M. Adam, and N. De Lange, "Systematic Approach for Remote Sensing of Historical Conflict Landscapes with UAV-Based Laserscanning," *Sensors*, vol. 22, no. 1, p. 217, Dec. 2021, doi: 10.3390/s22010217.

[112] T. R. H. Goodbody, N. C. Coops, and J. C. White, "Digital Aerial Photogrammetry for Updating Area-Based Forest Inventories: A Review of Opportunities, Challenges, and Future Directions," *Curr Forestry Rep*, vol. 5, no. 2, pp. 55–75, Jun. 2019, doi: 10.1007/s40725-019-00087-2.

[113] A. Michez, H. Piégay, J. Lisein, H. Claessens, and P. Lejeune, "Classification of riparian forest species and health condition using multi-temporal and hyperspatial imagery from unmanned aerial system," *Environ Monit Assess*, vol. 188, no. 3, p. 146, Mar. 2016, doi: 10.1007/s10661-015-4996-2.

[114] V. Stojakovic, "Terrestrial photogrammetry and application to modeling architectural objects," *Facta Univ Arch Civ Enge*, vol. 6, no. 1, pp. 113–125, 2008, doi: 10.2298/FUACE0801113S.

[115] J. Telling, A. Lyda, P. Hartzell, and C. Glennie, "Review of Earth science research using terrestrial laser scanning," *Earth-Science Reviews*, vol. 169, pp. 35–68, Jun. 2017, doi: 10.1016/j.earscirev.2017.04.007.

[116] P. R. Wolf and B. A. Dewitt, *Elements of photogrammetry: with applications in GIS*, 3rd ed. Boston: McGraw-Hill, 2000.

[117] J. Moyano, J. E. Nieto-Julián, D. Bienvenido-Huertas, and D. Marín-García, "Validation of Close-Range Photogrammetry for Architectural and Archaeological Heritage: Analysis of Point Density and 3D Mesh Geometry," *Remote Sensing*, vol. 12, no. 21, p. 3571, Oct. 2020, doi: 10.3390/rs12213571.

[118] S. Pal Singh, K. Jain, and V. R. Mandla, "Virtual 3D Campus Modeling by Using Close Range Photogrammetry," *AJCEA*, vol. 1, no. 6, pp. 200–205, Dec. 2013, doi: 10.12691/ajcea-1-6-9.

[119] W. C. Haneberg, "Using close range terrestrial digital photogrammetry for 3-D rock slope modeling and discontinuity mapping in the United States," *Bull Eng Geol Environ*, vol. 67, no. 4, pp. 457–469, Nov. 2008, doi: 10.1007/s10064-008-0157-y.

[120] R. Jiang, D. V. Jáuregui, and K. R. White, "Close-range photogrammetry applications in bridge measurement: Literature review," *Measurement*, vol. 41, no. 8, pp. 823–834, Oct. 2008, doi: 10.1016/j.measurement.2007.12.005.

[121] M. Mokroš *et al.*, "Evaluation of Close-Range Photogrammetry Image Collection Methods for Estimating Tree Diameters," *IJGI*, vol. 7, no. 3, p. 93, Mar. 2018, doi: 10.3390/ijgi7030093.

[122] P. Surový, A. Yoshimoto, and D. Panagiotidis, "Accuracy of Reconstruction of the Tree Stem Surface Using Terrestrial Close-

Range Photogrammetry," *Remote Sensing*, vol. 8, no. 2, p. 123, Feb. 2016, doi: 10.3390/rs8020123.

[123] Y. Dong, G. Fan, Z. Zhou, J. Liu, Y. Wang, and F. Chen, "Low Cost Automatic Reconstruction of Tree Structure by AdQSM with Terrestrial Close-Range Photogrammetry," *Forests*, vol. 12, no. 8, p. 1020, Jul. 2021, doi: 10.3390/f12081020.

[124] L. J. Sánchez-Aparicio, S. Del Pozo, L. F. Ramos, A. Arce, and F. M. Fernandes, "Heritage site preservation with combined radiometric and geometric analysis of TLS data," *Automation in Construction*, vol. 85, pp. 24–39, Jan. 2018, doi: 10.1016/j.autcon.2017.09.023.

[125] X. Liang *et al.*, "Terrestrial laser scanning in forest inventories," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 115, pp. 63–77, May 2016, doi: 10.1016/j.isprsjprs.2016.01.006.

[126] P. J. Watt and D. N. M. Donoghue, "Measuring forest structure with terrestrial laser scanning," *International Journal of Remote Sensing*, vol. 26, no. 7, pp. 1437–1446, Apr. 2005, doi: 10.1080/01431160512331337961.

[127] H. S. Park, H. M. Lee, H. Adeli, and I. Lee, "A New Approach for Health Monitoring of Structures: Terrestrial Laser Scanning," *Computer aided Civil Eng*, vol. 22, no. 1, pp. 19–30, Jan. 2007, doi: 10.1111/j.1467-8667.2006.00466.x.

[128] C. Wu, Y. Yuan, Y. Tang, and B. Tian, "Application of Terrestrial Laser Scanning (TLS) in the Architecture, Engineering and Construction (AEC) Industry," *Sensors*, vol. 22, no. 1, p. 265, Dec. 2021, doi: 10.3390/s22010265.

[129] K. Wong, Y. Gu, and S. Kamijo, "Mapping for Autonomous Driving: Opportunities and Challenges," *IEEE Intell. Transport. Syst. Mag.*, vol. 13, no. 1, pp. 91–106, 2021, doi: 10.1109/MITS.2020.3014152.

[130] F. Tsushima, N. Kishimoto, Y. Okada, and W. Che, "CREATION OF HIGH DEFINITION MAP FOR AUTONOMOUS DRIVING," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XLIII-B4-2020, pp. 415–420, Aug. 2020, doi: 10.5194/isprs-archives-XLIII-B4-2020-415-2020.

[131] H. Kawamura, Y. Nagata, T. Tokuno, T. Ishida, T. Mizutani, and J. Yamashita, "Pavement Evaluation Method Using MMS," in *Proceedings of the RILEM International Symposium on Bituminous Materials*, vol. 27, H. Di Benedetto, H. Baaj, E. Chailleux, G. Tebaldi, C. Sauzéat, and S. Mangiafico, Eds., in RILEM Bookseries, vol. 27. , Cham: Springer International Publishing, 2022, pp. 1455–1463. doi: 10.1007/978-3-030-46455-4_185.

[132] R. Ravi, D. Bullock, and A. Habib, "Pavement Distress and Debris Detection using a Mobile Mapping System with 2D Profiler LiDAR," *Transportation Research Record*, vol. 2675, no. 9, pp. 428–438, Sep. 2021, doi: 10.1177/03611981211002529.

[133] C. Thomson, G. Apostolopoulos, D. Backes, and J. Boehm, "Mobile Laser Scanning for Indoor Modelling," *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. II-5/W2, pp. 289–293, Oct. 2013, doi: 10.5194/isprsannals-II-5-W2-289-2013.

[134] M. Sgrenzaroli *et al.*, "INDOOR MOBILE MAPPING SYSTEMS AND (BIM) DIGITAL MODELS FOR CONSTRUCTION PROGRESS MONITORING," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XLIII-B1-2022, pp. 121–127, May 2022, doi: 10.5194/isprs-archives-XLIII-B1-2022-121-2022.

[135] N. Saarinen *et al.*, "Area-Based Approach for Mapping and Monitoring Riverine Vegetation Using Mobile Laser Scanning," *Remote Sensing*, vol. 5, no. 10, pp. 5285–5303, Oct. 2013, doi: 10.3390/rs5105285.

[136] A. Jaakkola *et al.*, "A low-cost multi-sensoral mobile mapping system and its feasibility for tree measurements," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 6, pp. 514–522, Nov. 2010, doi: 10.1016/j.isprsjprs.2010.08.002.

[137] F. Monnier, B. Vallet, and B. Soheilian, "TREES DETECTION FROM LASER POINT CLOUDS ACQUIRED IN DENSE URBAN AREAS BY A MOBILE MAPPING SYSTEM," *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. I–3, pp. 245–250, Jul. 2012, doi: 10.5194/isprsannals-I-3-245-2012.

[138] R. C. Lindenbergh, D. Berthold, B. Sirmacek, M. Herrero-Huerta, J. Wang, and D. Ebersbach, "AUTOMATED LARGE

SCALE PARAMETER EXTRACTION OF ROAD-SIDE TREES SAMPLED BY A LASER MOBILE MAPPING SYSTEM," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XL-3/W3, pp. 589–594, Aug. 2015, doi: 10.5194/isprsarchives-XL-3-W3-589-2015.

[139] X. Liang, J. Hyyppa, A. Kukko, H. Kaartinen, A. Jaakkola, and X. Yu, "The Use of a Mobile Laser Scanning System for Mapping Large Forest Plots," *IEEE Geosci. Remote Sensing Lett.*, vol. 11, no. 9, pp. 1504–1508, Sep. 2014, doi: 10.1109/LGRS.2013.2297418.

[140] R. Wang, "3D building modeling using images and LiDAR: a review," *International Journal of Image and Data Fusion*, vol. 4, no. 4, pp. 273–292, Dec. 2013, doi: 10.1080/19479832.2013.811124.

[141] H. Guan, J. Li, M. Chapman, F. Deng, Z. Ji, and X. Yang, "Integration of orthoimagery and lidar data for object-based urban thematic mapping using random forests," *International Journal of Remote Sensing*, vol. 34, no. 14, pp. 5166–5186, Jul. 2013, doi: 10.1080/01431161.2013.788261.

[142] N. Grasso, C. Spadavecchia, V. Di Pietra, and E. Belcore, "LiDAR and SfM-MVS Integrated Approach to Build a Highly Detailed 3D Virtual Model of Urban Areas:," in *Proceedings of the 9th International Conference on Geographical Information Systems Theory, Applications and Management*, Prague, Czech Republic: SCITEPRESS - Science and Technology Publications, 2023, pp. 128–135. doi: 10.5220/0011760800003473.

[143] S. Kahraman and R. Bacher, "A comprehensive review of hyperspectral data fusion with lidar and sar data," *Annual Reviews in Control*, vol. 51, pp. 236–253, 2021, doi: 10.1016/j.arcontrol.2021.03.003.

[144] M. Pardini *et al.*, "Early Lessons on Combining Lidar and Multi-baseline SAR Measurements for Forest Structure Characterization," *Surv Geophys*, vol. 40, no. 4, pp. 803–837, Jul. 2019, doi: 10.1007/s10712-019-09553-9.

[145] Y. Yan, Z. Mao, J. Wu, T. Padir, and J. F. Hajjar, "Towards automated detection and quantification of concrete cracks using integrated images and lidar data from unmanned aerial vehicles,"

Struct Control Health Monit, vol. 28, no. 8, Aug. 2021, doi: 10.1002/stc.2757.

[146] T. Mikita, P. Janata, and P. Surový, "Forest Stand Inventory Based on Combined Aerial and Terrestrial Close-Range Photogrammetry," *Forests*, vol. 7, no. 12, p. 165, Jul. 2016, doi: 10.3390/f7080165.

[147] J. L. Lerma, S. Navarro, M. Cabrelles, and V. Villaverde, "Terrestrial laser scanning and close range photogrammetry for 3D archaeological documentation: the Upper Palaeolithic Cave of Parpalló as a case study," *Journal of Archaeological Science*, vol. 37, no. 3, pp. 499–507, Mar. 2010, doi: 10.1016/j.jas.2009.10.011.

[148] A. Dietmaier, G. J. McDermid, M. M. Rahman, J. Linke, and R. Ludwig, "Comparison of LiDAR and Digital Aerial Photogrammetry for Characterizing Canopy Openings in the Boreal Forest of Northern Alberta," *Remote Sensing*, vol. 11, no. 16, p. 1919, Aug. 2019, doi: 10.3390/rs11161919.

[149] J. Liao, J. Zhou, and W. Yang, "Comparing LiDAR and SfM digital surface models for three land cover types," *Open Geosciences*, vol. 13, no. 1, pp. 497–504, May 2021, doi: 10.1515/geo-2020-0257.

[150] J. Torres-Sánchez *et al.*, "Mobile terrestrial laser scanner vs. UAV photogrammetry to estimate woody crop canopy parameters – Part 2: Comparison for different crops and training systems," *Computers and Electronics in Agriculture*, vol. 212, p. 108083, Sep. 2023, doi: 10.1016/j.compag.2023.108083.

[151] L. Cao, H. Liu, X. Fu, Z. Zhang, X. Shen, and H. Ruan, "Comparison of UAV LiDAR and Digital Aerial Photogrammetry Point Clouds for Estimating Forest Structural Attributes in Subtropical Planted Forests," *Forests*, vol. 10, no. 2, p. 145, Feb. 2019, doi: 10.3390/f10020145.

[152] A. Fryskowska, P. Walczykowski, P. Delis, and M. Wojtkowska, "ALS and TLS data fusion in cultural heritage documentation and modeling," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XL-5/W7, pp. 147–150, Aug. 2015, doi: 10.5194/isprsarchives-XL-5-W7-147-2015.

[153] M. N. Bazezew, Y. A. Hussin, and E. H. Kloosterman, "Integrating Airborne LiDAR and Terrestrial Laser Scanner forest parameters for accurate above-ground biomass/carbon estimation in Ayer Hitam tropical forest, Malaysia," *International Journal of Applied Earth Observation and Geoinformation*, vol. 73, pp. 638–652, Dec. 2018, doi: 10.1016/j.jag.2018.07.026.

[154] C. Paris, D. Kelbe, J. Van Aardt, and L. Bruzzone, "A Novel Automatic Method for the Fusion of ALS and TLS LiDAR Data for Robust Assessment of Tree Crown Structure," *IEEE Trans. Geosci. Remote Sensing*, vol. 55, no. 7, pp. 3679–3693, Jul. 2017, doi: 10.1109/TGRS.2017.2675963.

[155] D. Panagiotidis, A. Abdollahnejad, and M. Slavík, "3D point cloud fusion from UAV and TLS to assess temperate managed forest structures," *International Journal of Applied Earth Observation and Geoinformation*, vol. 112, p. 102917, Aug. 2022, doi: 10.1016/j.jag.2022.102917.

[156] R. Zhou *et al.*, "Improving Estimation of Tree Parameters by Fusing ALS and TLS Point Cloud Data Based on Canopy Gap Shape Feature Points," *Drones*, vol. 7, no. 8, p. 524, Aug. 2023, doi: 10.3390/drones7080524.

[157] S. A. Berrabah and Y. Baudoin, "GPS data correction using encoders and inertial navigation system (INS) sensors," in *Using Robots in Hazardous Environments*, Elsevier, 2011, pp. 269–282. doi: 10.1533/9780857090201.2.269.

[158] N. Boguspayev, D. Akhmedov, A. Raskaliyev, A. Kim, and A. Sukhenko, "A Comprehensive Review of GNSS/INS Integration Techniques for Land and Air Vehicle Applications," *Applied Sciences*, vol. 13, no. 8, p. 4819, Apr. 2023, doi: 10.3390/app13084819.

[159] ""ETHzürich. 'Automatic registration of partially overlapping terrestrial laser scanner point clouds' Accessed: Dec. 13, 2023. [Online]. Available: https://prs.igp.ethz.ch/research/completed_projects/automatic_registration_of_point_clouds.html."

[160] P. Glira, N. Pfeifer, C. Briese, and C. Ressl, "RIGOROUS STRIP ADJUSTMENT OF AIRBORNE LASERSCANNING DATA

BASED ON THE ICP ALGORITHM," *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. II-3/W5, pp. 73–80, Aug. 2015, doi: 10.5194/isprsannals-II-3-W5-73-2015.

[161] M. Hebel and U. Stilla, "Simultaneous Calibration of ALS Systems and Alignment of Multiview LiDAR Scans of Urban Areas," *IEEE Trans. Geosci. Remote Sensing*, vol. 50, no. 6, pp. 2364–2379, Jun. 2012, doi: 10.1109/TGRS.2011.2171974.

[162] D. Zermas, I. Izzat, and N. Papanikolopoulos, "Fast segmentation of 3D point clouds: A paradigm on LiDAR data for autonomous vehicle applications," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, Singapore: IEEE, May 2017, pp. 5067–5073. doi: 10.1109/ICRA.2017.7989591.

[163] X. Xie, H. Wei, and Y. Yang, "Real-Time LiDAR Point-Cloud Moving Object Segmentation for Autonomous Driving," *Sensors*, vol. 23, no. 1, p. 547, Jan. 2023, doi: 10.3390/s23010547.

[164] B. Wang, M. Zhu, Y. Lu, J. Wang, W. Gao, and H. Wei, "Real-Time 3D Object Detection From Point Cloud Through Foreground Segmentation," *IEEE Access*, vol. 9, pp. 84886–84898, 2021, doi: 10.1109/ACCESS.2021.3087179.

[165] M. A. Uy, Q.-H. Pham, B.-S. Hua, D. T. Nguyen, and S.-K. Yeung, "Revisiting Point Cloud Classification: A New Benchmark Dataset and Classification Model on Real-World Data," 2019, doi: 10.48550/ARXIV.1908.04616.

[166] Z. Wang *et al.*, "A Multiscale and Hierarchical Feature Extraction Method for Terrestrial Laser Scanning Point Cloud Classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 53, no. 5, pp. 2409–2425, May 2015, doi: 10.1109/TGRS.2014.2359951.

[167] E. Grilli, F. Menna, and F. Remondino, "A REVIEW OFPOINT CLOUDS SEGMENTATION AND CLASSIFICATION ALGORITHMS," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XLII-2/W3, pp. 339–344, Feb. 2017, doi: 10.5194/isprs-archives-XLII-2-W3-339-2017.

[168] C. Mineo, S. G. Pierce, and R. Summan, "Novel algorithms for 3D surface point cloud boundary detection and edge reconstruction," *Journal of Computational Design and Engineering*, vol. 6, no. 1, pp. 81–91, Jan. 2019, doi: 10.1016/j.jcde.2018.02.001.

[169] S. M. Ahmed, Y. Z. Tan, C. M. Chew, A. A. Mamun, and F. S. Wong, "Edge and Corner Detection for Unorganized 3D Point Clouds with Application to Robotic Welding," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Madrid: IEEE, Oct. 2018, pp. 7350–7355. doi: 10.1109/IROS.2018.8593910.

[170] H. Ni, X. G. Lin, and J. X. Zhang, "APPLICATIONS OF 3D-EDGE DETECTION FOR ALS POINT CLOUD," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XLII-2/W7, pp. 277–283, Sep. 2017, doi: 10.5194/isprs-archives-XLII-2-W7-277-2017.

[171] K. Klasing, D. Althoff, D. Wollherr, and M. Buss, "Comparison of surface normal estimation methods for range sensing applications," in *2009 IEEE International Conference on Robotics and Automation*, Kobe: IEEE, May 2009, pp. 3206–3211. doi: 10.1109/ROBOT.2009.5152493.

[172] J. Chen and B. Chen, "Architectural Modeling from Sparsely Scanned Range Data," *Int J Comput Vis*, vol. 78, no. 2–3, pp. 223–236, Jul. 2008, doi: 10.1007/s11263-007-0105-5.

[173] A. Nguyen and B. Le, "3D point cloud segmentation: A survey," in *2013 6th IEEE Conference on Robotics, Automation and Mechatronics (RAM)*, Manila, Philippines: IEEE, Nov. 2013, pp. 225–230. doi: 10.1109/RAM.2013.6758588.

[174] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981, doi: 10.1145/358669.358692.

[175] R. Schnabel, R. Wahl, and R. Klein, "Efficient RANSAC for Point-Cloud Shape Detection," *Computer Graphics Forum*, vol. 26, no. 2, pp. 214–226, Jun. 2007, doi: 10.1111/j.1467-8659.2007.01016.x.

[176] Z. Li and J. Shan, "RANSAC-based multi primitive building reconstruction from 3D point clouds," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 185, pp. 247–260, Mar. 2022, doi: 10.1016/j.isprsjprs.2021.12.012.

[177] M. Ghahremani *et al.*, "Direct and accurate feature extraction from 3D point clouds of plants using RANSAC," *Computers and Electronics in Agriculture*, vol. 187, p. 106240, Aug. 2021, doi: 10.1016/j.compag.2021.106240.

[178] B. Xu, W. Jiang, J. Shan, J. Zhang, and L. Li, "Investigation on the Weighted RANSAC Approaches for Building Roof Plane Segmentation from LiDAR Point Clouds," *Remote Sensing*, vol. 8, no. 1, p. 5, Dec. 2015, doi: 10.3390/rs8010005.

[179] S. Yang, M. Hou, and S. Li, "Three-Dimensional Point Cloud Semantic Segmentation for Cultural Heritage: A Comprehensive Review," *Remote Sensing*, vol. 15, no. 3, p. 548, Jan. 2023, doi: 10.3390/rs15030548.

[180] K. Olofsson, J. Holmgren, and H. Olsson, "Tree Stem and Height Measurements using Terrestrial Laser Scanning and the RANSAC Algorithm," *Remote Sensing*, vol. 6, no. 5, pp. 4323–4344, May 2014, doi: 10.3390/rs6054323.

[181] K. Kuželka, M. Slavík, and P. Surový, "Very High Density Point Clouds from UAV Laser Scanning for Automatic Tree Stem Detection and Direct Diameter Measurement," *Remote Sensing*, vol. 12, no. 8, p. 1236, Apr. 2020, doi: 10.3390/rs12081236.

[182] A. Singh, S. K. P. Kushwaha, S. Nandy, and H. Padalia, "NOVEL APPROACH FOR FOREST ALLOMETRIC EQUATION MODELLING WITH RANSAC SHAPE DETECTION USING TERRESTRIAL LASER SCANNER," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XLVIII-4/W4-2022, pp. 133–138, Oct. 2022, doi: 10.5194/isprs-archives-XLVIII-4-W4-2022-133-2022.

[183] A. Singh, S. K. P. Kushwaha, S. Nandy, and H. Padalia, "An approach for tree volume estimation using RANSAC and RHT algorithms from TLS dataset," *Appl Geomat*, vol. 14, no. 4, pp. 785–794, Dec. 2022, doi: 10.1007/s12518-022-00471-x.

[184] W. Huang *et al.*, "A Fast Point Cloud Ground Segmentation Approach Based on Coarse-To-Fine Markov Random Field," *IEEE Trans. Intell. Transport. Syst.*, vol. 23, no. 7, pp. 7841–7854, Jul. 2022, doi: 10.1109/TITS.2021.3073151.

[185] Z. Shen, H. Liang, L. Lin, Z. Wang, W. Huang, and J. Yu, "Fast Ground Segmentation for 3D LiDAR Point Cloud Based on Jump-Convolution-Process," *Remote Sensing*, vol. 13, no. 16, p. 3239, Aug. 2021, doi: 10.3390/rs13163239.

[186] Z. Cheng, G. Ren, and Y. Zhang, "Ground Segmentation Algorithm Based on 3D Lidar Point Cloud," in *Proceedings of the 2018 International Conference on Mechanical, Electrical, Electronic Engineering & Science (MEEES 2018)*, Chongqing, China: Atlantis Press, 2018. doi: 10.2991/meees-18.2018.4.

[187] M. Zhang, D. D. Morris, and R. Fu, "Ground Segmentation Based on Loopy Belief Propagation for Sparse 3D Point Clouds," in *2015 International Conference on 3D Vision*, Lyon, France: IEEE, Oct. 2015, pp. 615–622. doi: 10.1109/3DV.2015.76.

[188] Z. Lari, A. Habib, and E. Kwak, "AN ADAPTIVE APPROACH FOR SEGMENTATION OF 3D LASER POINT CLOUD," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XXXVIII-5/W12, pp. 103–108, Sep. 2012, doi: 10.5194/isprsarchives-XXXVIII-5-W12-103-2011.

[189] G. Vosselman and S. Dijkman, "3D building model reconstruction from point clouds and ground plans," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXIV, Jan. 2001.

[190] S. Filin and N. Pfeifer, "Neighborhood Systems for Airborne Laser Data," *photogramm eng remote sensing*, vol. 71, no. 6, pp. 743–755, Jun. 2005, doi: 10.14358/PERS.71.6.743.

[191] C. Kim, A. F. Habib, and P. Mrstik, "NEW APPROACH FOR PLANAR PATCH SEGMENTATION USING AIRBORNE LASER DATA," 2007. [Online]. Available: https://api.semanticscholar.org/CorpusID:42366087

[192] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, in KDD'96. Portland, Oregon: AAAI Press, 1996, pp. 226–231.

[193] L. McInnes, J. Healy, and S. Astels, "hdbscan: Hierarchical density based clustering," *JOSS*, vol. 2, no. 11, p. 205, Mar. 2017, doi: 10.21105/joss.00205.

[194] S. M. Su Benyue Ma Jinyu, Peng Yusheng, "Algorithm for RGBD Point Cloud Denoising and Simplification Based on K-means Clustering," *Journal of System Simulation*, vol. 28, no. 10, p. 2329, 2016.

[195] A. Saglam, H. B. Makineci, Ö. K. Baykan, and N. A. Baykan, "Clustering-Based Plane Refitting of Non-planar Patches for Voxel-Based 3D Point Cloud Segmentation Using K-Means Clustering," *TS*, vol. 37, no. 6, pp. 1019–1027, Dec. 2020, doi: 10.18280/ts.370614.

[196] R. A. Kuçak, E. Özdemir, and S. Erol, "THE SEGMENTATION OF POINT CLOUDS WITH K-MEANS AND ANN (ARTIFICAL NEURAL NETWORK)," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XLII-1/W1, pp. 595–598, May 2017, doi: 10.5194/isprs-archives-XLII-1-W1-595-2017.

[197] Y. Mu, G. Zhou, and H. Wang, "CANOPY LIDAR POINT CLOUD DATA K-MEANS CLUSTERING WATERSHED SEGMENTATION METHOD," *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. VI-3/W1-2020, pp. 67–73, Nov. 2020, doi: 10.5194/isprs-annals-VI-3-W1-2020-67-2020.

[198] I. Davidson and S. S. Ravi, "Clustering With Constraints: Feasibility Issues and the *k* -Means Algorithm," in *Proceedings of the 2005 SIAM International Conference on Data Mining*, Society for Industrial and Applied Mathematics, Apr. 2005, pp. 138–149. doi: 10.1137/1.9781611972757.13.

[199] T. Czerniawski, B. Sankaran, M. Nahangi, C. Haas, and F. Leite, "6D DBSCAN-based segmentation of building point clouds for planar object classification," *Automation in Construction*, vol. 88, pp. 44–58, Apr. 2018, doi: 10.1016/j.autcon.2017.12.029.

[200] C. Wang, M. Ji, J. Wang, W. Wen, T. Li, and Y. Sun, "An Improved DBSCAN Method for LiDAR Data Segmentation with Automatic Eps Estimation," *Sensors*, vol. 19, no. 1, p. 172, Jan. 2019, doi: 10.3390/s19010172.

[201] H. Chen, M. Liang, W. Liu, W. Wang, and P. X. Liu, "An approach to boundary detection for 3D point clouds based on DBSCAN clustering," *Pattern Recognition*, vol. 124, p. 108431, Apr. 2022, doi: 10.1016/j.patcog.2021.108431.

[202] H. Sheikh, C. Prins, and E. Schrijvers, "Artificial Intelligence: Definition and Background," in *Mission AI*, in Research for Policy. , Cham: Springer International Publishing, 2023, pp. 15–41. doi: 10.1007/978-3-031-21448-6_2.

[203] E. Fix and J. L. Hodges, "Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties," *International Statistical Review / Revue Internationale de Statistique*, vol. 57, no. 3, p. 238, Dec. 1989, doi: 10.2307/1403797.

[204] N. Luo, Y. Wang, Y. Gao, Y. Tian, Q. Wang, and C. Jing, "k NN-based feature learning network for semantic segmentation of point cloud data," *Pattern Recognition Letters*, vol. 152, pp. 365–371, Dec. 2021, doi: 10.1016/j.patrec.2021.10.023.

[205] A. Mahdaoui and E. H. Sbai, "3D Point Cloud Simplification Based on *k* -Nearest Neighbor and Clustering," *Advances in Multimedia*, vol. 2020, pp. 1–10, Jul. 2020, doi: 10.1155/2020/8825205.

[206] W. Lin, W. Fan, H. Liu, Y. Xu, and J. Wu, "Classification of Handheld Laser Scanning Tree Point Cloud Based on Different KNN Algorithms and Random Forest Algorithm," *Forests*, vol. 12, no. 3, p. 292, Mar. 2021, doi: 10.3390/f12030292.

[207] Lianwei Zhang, Tingbo Hu, Xiaolin Liu, Yan Li, Tao Wu, and Hangen He, "Surface reconstruction from cloud points based on Support Vector Machine," in *2008 IEEE International Conference on Automation and Logistics*, Qingdao, China: IEEE, Sep. 2008, pp. 377–381. doi: 10.1109/ICAL.2008.4636179.

[208] J. Zhang, X. Lin, and X. Ning, "SVM-Based Classification of Segmented Airborne LiDAR Point Clouds in Urban Areas," *Remote Sensing*, vol. 5, no. 8, pp. 3749–3775, Jul. 2013, doi: 10.3390/rs5083749.

[209] M. Caputo, K. Denker, M. O. Franz, P. Laube, and G. Umlauf, "Support Vector Machines for Classification of Geometric Primitives in Point Clouds," in *Curves and Surfaces*, vol. 9213, J.-

D. Boissonnat, A. Cohen, O. Gibaru, C. Gout, T. Lyche, M.-L. Mazure, and L. L. Schumaker, Eds., in Lecture Notes in Computer Science, vol. 9213. , Cham: Springer International Publishing, 2015, pp. 80–95. doi: 10.1007/978-3-319-22804-4_7.

[210] C. Zhang *et al.*, "A method for organs classification and fruit counting on pomegranate trees based on multi-features fusion and support vector machine by 3D point cloud," *Scientia Horticulturae*, vol. 278, p. 109791, Feb. 2021, doi: 10.1016/j.scienta.2020.109791.

[211] T. Mizutani *et al.*, "Automatic detection of delamination on tunnel lining surfaces from laser 3D point cloud data by 3D features and a support vector machine," *J Civil Struct Health Monit*, Oct. 2023, doi: 10.1007/s13349-023-00731-3.

[212] M. Zeybek, "Classification of UAV point clouds by random forest machine learning algorithm," *Turkish Journal of Engineering*, vol. 5, no. 2, pp. 48–57, Apr. 2021, doi: 10.31127/tuje.669566.

[213] L. Liao *et al.*, "A Supervoxel-Based Random Forest Method for Robust and Effective Airborne LiDAR Point Cloud Classification," *Remote Sensing*, vol. 14, no. 6, p. 1516, Mar. 2022, doi: 10.3390/rs14061516.

[214] M. Bassier, B. Van Genechten, and M. Vergauwen, "Classification of sensor independent point cloud data of building objects using random forests," *Journal of Building Engineering*, vol. 21, pp. 468–477, Jan. 2019, doi: 10.1016/j.jobe.2018.04.027.

[215] M. Mohamed, S. Morsy, and A. El-Shazly, "Improvement of 3D LiDAR point cloud classification of urban road environment based on random forest classifier," *Geocarto International*, vol. 37, no. 27, pp. 15604–15626, Dec. 2022, doi: 10.1080/10106049.2022.2102218.

[216] D. Xu, F. Li, and H. Wei, "3D Point Cloud Plane Segmentation Method Based on RANSAC And Support Vector Machine," in *2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, Xi'an, China: IEEE, Jun. 2019, pp. 943–948. doi: 10.1109/ICIEA.2019.8834367.

[217] P. Raktrakulthum and C. Netramai, "Vehicle classification in congested traffic based on 3D point cloud using SVM and KNN," in *2017 9th International Conference on Information Technology and*

Electrical Engineering (ICITEE), Phuket: IEEE, Oct. 2017, pp. 1–6. doi: 10.1109/ICITEED.2017.8250451.

[218] H. Ni, X. Lin, and J. Zhang, "Classification of ALS Point Cloud with Improved Point Cloud Segmentation and Random Forests," *Remote Sensing*, vol. 9, no. 3, p. 288, Mar. 2017, doi: 10.3390/rs9030288.

[219] A. K. Aijazi, A. Serna, B. Marcotegui, P. Checchin, and L. Trassoudaine, "Segmentation and Classification of 3D Urban Point Clouds: Comparison and Combination of Two Approaches," in *Field and Service Robotics*, vol. 113, D. S. Wettergreen and T. D. Barfoot, Eds., in Springer Tracts in Advanced Robotics, vol. 113. , Cham: Springer International Publishing, 2016, pp. 201–216. doi: 10.1007/978-3-319-27702-8_14.

[220] J. Zörner, J. Dymond, J. Shepherd, and B. Jolly, "PyCrown - Fast raster-based individual tree segmentation for LiDAR data." Landcare Research NZ Ltd, 2018. doi: 10.7931/M0SR-DN55.

[221] ""GitHub. 'PyCrown - Fast raster-based individual tree segmentation for LiDAR data.' Accessed: Feb. 07, 2024. [Online]. Available: https://github.com/manaakiwhenua/pycrown?tab=readme-ov-file."

[222] S. Krisanski *et al.*, "Forest Structural Complexity Tool—An Open Source, Fully-Automated Tool for Measuring Forest Point Clouds," *Remote Sensing*, vol. 13, no. 22, p. 4677, Nov. 2021, doi: 10.3390/rs13224677.

[223] S. Krisanski, M. S. Taskhiri, S. Gonzalez Aracil, D. Herries, and P. Turner, "Sensor Agnostic Semantic Segmentation of Structurally Diverse and Complex Forest Point Clouds Using Deep Learning," *Remote Sensing*, vol. 13, no. 8, p. 1413, Apr. 2021, doi: 10.3390/rs13081413.

[224] X. Xu, F. Iuricich, K. Calders, J. Armston, and L. De Floriani, "Topology-based individual tree segmentation for automated processing of terrestrial laser scanning point clouds," *International Journal of Applied Earth Observation and Geoinformation*, vol. 116, p. 103145, Feb. 2023, doi: 10.1016/j.jag.2022.103145.

[225] F. Tupinambá-Simões, A. Pascual, J. Guerra-Hernández, C. Ordóñez, T. De Conto, and F. Bravo, "Assessing the Performance of

a Handheld Laser Scanning System for Individual Tree Mapping—A Mixed Forests Showcase in Spain," *Remote Sensing*, vol. 15, no. 5, p. 1169, Feb. 2023, doi: 10.3390/rs15051169.

[226] M. J. Allen, S. W. D. Grieve, H. J. F. Owen, and E. R. Lines, "Tree species classification from complex laser scanning data in Mediterranean forests using deep learning," *Methods Ecol Evol*, vol. 14, no. 7, pp. 1657–1667, Jul. 2023, doi: 10.1111/2041-210X.13981.

[227] H. Luo, K. Khoshelham, C. Chen, and H. He, "Individual tree extraction from urban mobile laser scanning point clouds using deep pointwise direction embedding," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 175, pp. 326–339, May 2021, doi: 10.1016/j.isprsjprs.2021.03.002.

[228] A. Burt, M. Disney, and K. Calders, "Extracting individual trees from lidar point clouds using *treeseg*," *Methods Ecol Evol*, pp. 2041–210X.13121, Dec. 2018, doi: 10.1111/2041-210X.13121.

[229] M. Boni Vicari, "On leaf and wood separation from Terrestrial LiDAR data," PhD Thesis, 2019.

[230] C. Xiao, R. Qin, X. Huang, and J. Li, "Individual Tree Detection from Multi-View Satellite Images," in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Valencia: IEEE, Jul. 2018, pp. 3967–3970. doi: 10.1109/IGARSS.2018.8518040.

[231] M. P. Ferreira *et al.*, "Individual tree detection and species classification of Amazonian palms using UAV images and deep learning," *Forest Ecology and Management*, vol. 475, p. 118397, Nov. 2020, doi: 10.1016/j.foreco.2020.118397.

[232] A. A. dos Santos *et al.*, "Assessment of CNN-Based Methods for Individual Tree Detection on Images Captured by RGB Cameras Attached to UAVs," *Sensors*, vol. 19, no. 16, p. 3595, Aug. 2019, doi: 10.3390/s19163595.

[233] J. Sun *et al.*, "Wood–Leaf Classification of Tree Point Cloud Based on Intensity and Geometric Information," *Remote Sensing*, vol. 13, no. 20, p. 4050, Oct. 2021, doi: 10.3390/rs13204050.

[234] M. B. Vicari, M. Disney, P. Wilkes, A. Burt, K. Calders, and W. Woodgate, "Leaf and wood classification framework for terrestrial

LiDAR point clouds," *Methods Ecol Evol*, vol. 10, no. 5, pp. 680–694, May 2019, doi: 10.1111/2041-210X.13144.

[235] D. Wang, S. Momo Takoudjou, and E. Casella, "LeWoS: A universal leaf-wood classification method to facilitate the 3D modelling of large tropical trees using terrestrial LiDAR," *Methods Ecol Evol*, vol. 11, no. 3, pp. 376–389, Mar. 2020, doi: 10.1111/2041-210X.13342.

[236] Z. Tian and S. Li, "Graph-Based Leaf–Wood Separation Method for Individual Trees Using Terrestrial Lidar Point Clouds," *IEEE Trans. Geosci. Remote Sensing*, vol. 60, pp. 1–11, 2022, doi: 10.1109/TGRS.2022.3218603.

[237] R. Liu, L. Ren, and F. Wang, "3D Point Cloud of Single Tree Branches and Leaves Semantic Segmentation Based on Modified PointNet Network," *J. Phys.: Conf. Ser.*, vol. 2074, no. 1, p. 012026, Nov. 2021, doi: 10.1088/1742-6596/2074/1/012026.

[238] V. Ulrich, J. G. Williams, V. Zahs, K. Anders, S. Hecht, and B. Höfle, "Measurement of rock glacier surface change over different timescales using terrestrial laser scanning point clouds," *Earth Surf. Dynam.*, vol. 9, no. 1, pp. 19–28, Jan. 2021, doi: 10.5194/esurf-9-19-2021.

[239] A. Malekian, D. Fugazza, and M. Scaioni, "3D Surface Reconstruction and Change Detection of Miage Glacier (Italy) from Multi-date Archive Aerial Photos," in *Computational Science and Its Applications – ICCSA 2022 Workshops*, vol. 13379, O. Gervasi, B. Murgante, S. Misra, A. M. A. C. Rocha, and C. Garau, Eds., in Lecture Notes in Computer Science, vol. 13379. , Cham: Springer International Publishing, 2022, pp. 450–465. doi: 10.1007/978-3-031-10545-6_31.

[240] J. Guevara, T. Arevalo-Ramirez, F. Yandun, M. Torres-Torriti, and F. A. Cheein, "Point cloud-based estimation of effective payload volume for earthmoving loaders," *Automation in Construction*, vol. 117, p. 103207, Sep. 2020, doi: 10.1016/j.autcon.2020.103207.

[241] L. Korhonen, J. Vauhkonen, A. Virolainen, A. Hovi, and I. Korpela, "Estimation of tree crown volume from airborne lidar data using computational geometry," *International Journal of Remote*

Sensing, vol. 34, no. 20, pp. 7236–7248, Oct. 2013, doi: 10.1080/01431161.2013.817715.

[242] Z. Yan, R. Liu, L. Cheng, X. Zhou, X. Ruan, and Y. Xiao, "A Concave Hull Methodology for Calculating the Crown Volume of Individual Trees Based on Vehicle-Borne LiDAR Data," *Remote Sensing*, vol. 11, no. 6, p. 623, Mar. 2019, doi: 10.3390/rs11060623.

[243] A. Bienert, C. Hess, H.-G. Maas, and G. Von Oheimb, "A voxel-based technique to estimate the volume of trees from terrestrial laser scanner data," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XL–5, pp. 101–106, Jun. 2014, doi: 10.5194/isprsarchives-XL-5-101-2014.

[244] A. Nurunnabi, Y. Sadahiro, and R. Lindenbergh, "ROBUST CYLINDER FITTING IN THREE-DIMENSIONAL POINT CLOUD DATA," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XLII-1/W1, pp. 63–70, May 2017, doi: 10.5194/isprs-archives-XLII-1-W1-63-2017.

[245] B. Paláncz *et al.*, "A robust cylindrical fitting to point cloud data," *Australian Journal of Earth Sciences*, vol. 63, no. 5, pp. 665–673, Jul. 2016, doi: 10.1080/08120099.2016.1230147.

[246] A. Nurunnabi, Y. Sadahiro, R. Lindenbergh, and D. Belton, "Robust cylinder fitting in laser scanning point cloud data," *Measurement*, vol. 138, pp. 632–651, May 2019, doi: 10.1016/j.measurement.2019.01.095.

[247] V. Méndez, J. R. Rosell-Polo, R. Sanz, A. Escolà, and H. Catalán, "Deciduous tree reconstruction algorithm based on cylinder fitting from mobile terrestrial laser scanned point clouds," *Biosystems Engineering*, vol. 124, pp. 78–88, Aug. 2014, doi: 10.1016/j.biosystemseng.2014.06.001.

[248] P. Raumonen and M. Åkerblom, "InverseTampere/TreeQSM: Version 2.4.1." Zenodo, May 11, 2022. doi: 10.5281/ZENODO.6539580.

[249] J. Hackenberg and J.-D. Bontemps, "Improving quantitative structure models with filters based on allometric scaling theory," *Appl Geomat*, vol. 15, no. 4, pp. 1019–1029, Dec. 2023, doi: 10.1007/s12518-023-00537-4.

[250] J. Gonzalez De Tanago *et al.*, "Estimation of above-ground biomass of large tropical trees with terrestrial LiDAR," *Methods Ecol Evol*, vol. 9, no. 2, pp. 223–234, Feb. 2018, doi: 10.1111/2041-210X.12904.

[251] K. Calders *et al.*, "Nondestructive estimates of above-ground biomass using terrestrial laser scanning," *Methods Ecol Evol*, vol. 6, no. 2, pp. 198–208, Feb. 2015, doi: 10.1111/2041-210X.12301.

[252] "GitHub. 'FSCT.' Accessed: Feb. 01, 2024. [Online]. Available: https://github.com/InverseTampere/TreeQSM."

[253] "The Engineering ToolBox (2004). Wood Species - Densities. [online] Available at: https://www.engineeringtoolbox.com/wood-density-d_40.html [Accessed 19 12 2023]."

[254] S. Proietti *et al.*, "Carbon footprint of an olive tree grove," *Applied Energy*, vol. 127, pp. 115–124, Aug. 2014, doi: 10.1016/j.apenergy.2014.04.019.

[255] N. Kumar, Nishtha, and J. Kumar, "Estimation of carbon sequestration by some selected dominant tree species of Chandigarh University, Gharuan, Mohali," *Materials Today: Proceedings*, p. S221478532301252X, Mar. 2023, doi: 10.1016/j.matpr.2023.03.210.

[256] S. H. Lamlom and R. A. Savidge, "A reassessment of carbon content in wood: variation within and between 41 North American species," *Biomass and Bioenergy*, vol. 25, no. 4, pp. 381–388, Oct. 2003, doi: 10.1016/S0961-9534(03)00033-3.

[257] W. Li, Q. Guo, M. K. Jakubowski, and M. Kelly, "A New Method for Segmenting Individual Trees from the Lidar Point Cloud," *photogramm eng remote sensing*, vol. 78, no. 1, pp. 75–84, Jan. 2012, doi: 10.14358/PERS.78.1.75.

[258] H. Fu, H. Li, Y. Dong, F. Xu, and F. Chen, "Segmenting Individual Tree from TLS Point Clouds Using Improved DBSCAN," *Forests*, vol. 13, no. 4, p. 566, Apr. 2022, doi: 10.3390/f13040566.

[259] J. Giménez-Gallego, J. D. González-Teruel, M. Jiménez-Buendía, A. B. Toledo-Moreo, F. Soto-Valles, and R. Torres-Sánchez, "Segmentation of Multiple Tree Leaves Pictures with Natural Backgrounds using Deep Learning for Image-Based

Agriculture Applications," *Applied Sciences*, vol. 10, no. 1, p. 202, Dec. 2019, doi: 10.3390/app10010202.

[260] M. Larsen, M. Eriksson, X. Descombes, G. Perrin, T. Brandtberg, and F. A. Gougeon, "Comparison of six individual tree crown detection algorithms evaluated under varying forest conditions," *International Journal of Remote Sensing*, vol. 32, no. 20, pp. 5827–5852, Oct. 2011, doi: 10.1080/01431161.2010.507790.

[261] S. M. A. Jeronimo, V. R. Kane, D. J. Churchill, R. J. McGaughey, and J. F. Franklin, "Applying LiDAR Individual Tree Detection to Management of Structurally Diverse Forest Landscapes," *Journal of Forestry*, vol. 116, no. 4, pp. 336–346, Jun. 2018, doi: 10.1093/jofore/fvy023.

[262] T. Miao *et al.*, "Automatic stem-leaf segmentation of maize shoots using three-dimensional point cloud," *Computers and Electronics in Agriculture*, vol. 187, p. 106310, Aug. 2021, doi: 10.1016/j.compag.2021.106310.

[263] S. Li, L. Dai, H. Wang, Y. Wang, Z. He, and S. Lin, "Estimating Leaf Area Density of Individual Trees Using the Point Cloud Segmentation of Terrestrial LiDAR Data and a Voxel-Based Model," *Remote Sensing*, vol. 9, no. 11, p. 1202, Nov. 2017, doi: 10.3390/rs9111202.

[264] S. M. Krishna Moorthy, K. Calders, M. B. Vicari, and H. Verbeeck, "Improved Supervised Learning-Based Approach for Leaf and Wood Classification From LiDAR Point Clouds of Forests," *IEEE Trans. Geosci. Remote Sensing*, vol. 58, no. 5, pp. 3057–3070, May 2020, doi: 10.1109/TGRS.2019.2947198.

[265] S. Latifah, A. Purwoko, K. S. Hartini, and K. A. Fachrudin, "Allometric models to estimate the aboveground biomass of forest: A literature review," *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 1122, no. 1, p. 012047, Mar. 2021, doi: 10.1088/1757-899X/1122/1/012047.

[266] J. Chave *et al.*, "Improved allometric models to estimate the aboveground biomass of tropical trees," *Glob Change Biol*, vol. 20, no. 10, pp. 3177–3190, Oct. 2014, doi: 10.1111/gcb.12629.

[267] L. Mehtätalo and J. Lappi, *Biometry for Forestry and Environmental Data: with Examples in R*, 1st ed. Boca Raton, FL : CRC Press, 2020. | Series: Chapman & Hall/CRC applied

environmental statistics: Chapman and Hall/CRC, 2020. doi: 10.1201/9780429173462.

[268] C. Spadavecchia, E. Belcore, M. Piras, and M. Kobal, "An Automatic Individual Tree 3D Change Detection Method for Allometric Parameters Estimation in Mixed Uneven-Aged Forest Stands from ALS Data," *Remote Sensing*, vol. 14, no. 18, p. 4666, Sep. 2022, doi: 10.3390/rs14184666.

[269] T. A. Nagel, D. Firm, D. Rozenbergar, and M. Kobal, "Patterns and drivers of ice storm damage in temperate forests of Central Europe," *Eur J Forest Res*, vol. 135, no. 3, pp. 519–530, Jun. 2016, doi: 10.1007/s10342-016-0950-2.

[270] D. Vacha, G. Mandrone, D. Morresi, and M. Garbarino, "Mapping Post-fire Monthly Erosion Rates at the Catchment Scale Using Empirical Models Implemented in GIS. A Case Study in Northern Italy," in *Progress in Landslide Research and Technology, Volume 1 Issue 1, 2022*, K. Sassa, K. Konagai, B. Tiwari, Ž. Arbanas, and S. Sassa, Eds., in Progress in Landslide Research and Technology. , Cham: Springer International Publishing, 2023, pp. 99–112. doi: 10.1007/978-3-031-16898-7_6.

[271] C. Spadavecchia, E. Belcore, N. Grasso, and M. Piras, "A FULLY AUTOMATIC FOREST PARAMETERS EXTRACTION AT SINGLE-TREE LEVEL: A COMPARISON OF MLS AND TLS APPLICATIONS," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XLVIII-1/W1-2023, pp. 457–463, May 2023, doi: 10.5194/isprs-archives-XLVIII-1-W1-2023-457-2023.

[272] C. Spadavecchia, M. B. Campos, M. Piras, E. Puttonen, and A. Shcherbacheva, "WOOD-LEAF UNSUPERVISED CLASSIFICATION OF SILVER BIRCH TREES FOR BIOMASS ASSESSMENT USING OBLIQUE POINT CLOUDS," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XLVIII-1/W2-2023, pp. 1795–1802, Dec. 2023, doi: 10.5194/isprs-archives-XLVIII-1-W2-2023-1795-2023.

[273] M. B. Campos *et al.*, "A Long-Term Terrestrial Laser Scanning Measurement Station to Continuously Monitor Structural and Phenological Dynamics of Boreal Forest Canopy," *Front. Plant Sci.*, vol. 11, p. 606752, Jan. 2021, doi: 10.3389/fpls.2020.606752.

[274] K. Machacova, E. Vainio, O. Urban, and M. Pihlatie, "Seasonal dynamics of stem N2O exchange follow the physiological activity of boreal trees," *Nat Commun*, vol. 10, no. 1, p. 4989, Nov. 2019, doi: 10.1038/s41467-019-12976-y.

[275] D. Chen *et al.*, "The trend of the oxidants in boreal forest over 2007–2018: comprehensive modelling study with long-term measurements at SMEAR II, Finland," Gases/Atmospheric Modelling/Troposphere/Chemistry (chemical composition and reactions), preprint, Apr. 2020. doi: 10.5194/acp-2020-128.

[276] A. Shcherbcheva, M. B. Campos, X. Liang, E. Puttonen, and Y. Wang, "UNSUPERVISED STATISTICAL APPROACH FOR TREE-LEVEL SEPARATION OF FOLIAGE AND NON-LEAF COMPONENTS FROM POINT CLOUDS," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XLVIII-1/W2-2023, pp. 1787–1794, Dec. 2023, doi: 10.5194/isprs-archives-XLVIII-1-W2-2023-1787-2023.

[277] J. Diaci, "Silver Fir Decline in Mixed Old-Growth Forests in Slovenia: an Interaction of Air Pollution, Changing Forest Matrix and Climate," in *Air Pollution - New Developments*, A. Moldoveanu, Ed., InTech, 2011. doi: 10.5772/17962.

[278] K. Ma *et al.*, "Performance and Sensitivity of Individual Tree Segmentation Methods for UAV-LiDAR in Multiple Forest Types," *Remote Sensing*, vol. 14, no. 2, p. 298, Jan. 2022, doi: 10.3390/rs14020298.

[279] L. Qiu, L. Jing, B. Hu, H. Li, and Y. Tang, "A New Individual Tree Crown Delineation Method for High Resolution Multispectral Imagery," *Remote Sensing*, vol. 12, no. 3, p. 585, Feb. 2020, doi: 10.3390/rs12030585.

Appendix A

Modified PyCrown algorithm code to reduce oversegmentation

```
# INPUT DATA
points = gpd.read_file(outFolder+"\PyCrown\\tree_location_top_cor.shp")
crows = gpd.read_file(outFolder+"\PyCrown\\tree_crown_poly_raster.shp")

#ASSOCIATE THE SAME ID TO TREETOP AND CROWN OF THE SAME TREE
joined=gpd.sjoin(points, crows, how='inner', op='intersects')
ID=joined['index_right']
ID_list=ID.tolist()
points['DN']=ID
points.to_file(outFolder+"\PyCrown\\treetops.shp")

# SPATIAL SEARCH (RADIUS = 1 METER)
points['buffer'] = points.geometry.buffer(1)
points['included'] = False  # Initialize flag variable
result = gpd.GeoDataFrame()
for i, point in points.iterrows():
   if not point['included']:  # Skip point if already included
      mask = (points.within(point.buffer))
      temp = gpd.GeoDataFrame(points[mask])
      temp['New_DN'] = point['DN']
      result = result.append(temp)
      points.loc[mask, 'included'] = True  # Update flag variable for included points
result = result.reset_index(drop=True)  # Reset index for result GeoDataFrame
result = result[result['included'] == False]

# RELATE OLD ID WITH NEW ONES
Old_ID=result['DN'].values
New_ID=result['New_DN'].values
IDs = np.column_stack((Old_ID, New_ID))
```

```
    # IMPORT THE ORIGINAL POINT CLOUD DATASET
    file = laspy.file.File(outFolder+"\PyCrown\\Final_Trees.las", mode="r")
    x = (file.x).reshape(-1,1)
    y = (file.y).reshape(-1,1)
    z = (file.z).reshape(-1,1)
    trees = (file.trees).reshape(-1,1)
    returnnumber = (file.return_num).reshape(-1,1)
    File = np.hstack((x, y, z, trees, returnnumber))

    # SUBSTITUTE OLD ID WITH NEW ONES
    newID=[]
    for x in range(len(File)):
       for j in range(len(IDs)):
          if File[x][3]==IDs[j][0]:
               NewClass=IDs[j][1]
               newID.append(NewClass)
    newID=np.array(newID)

    # SAVE A NEW POINT CLOUD FILE WITH IMPROVED SEGMENTATION
    header = laspy.header.Header()
    outfile = laspy.file.File(outFolder+"\PyCrown\\Final_Trees_NO_Overseg.las", mode="w",
header=header)
    outfile.define_new_dimension(name = "trees", data_type = 9, description = "Trees")
    outfile.define_new_dimension(name = "return_num", data_type = 9, description =
"return_num")
    xmin = np.floor(np.min((file.x).reshape(-1,1)))
    ymin = np.floor(np.min(y))
    zmin = np.floor(np.min(z))
    outfile.header.offset = [xmin,ymin,zmin]
    outfile.header.scale = [0.001,0.001,0.001]
    outfile.x=(file.x).reshape(-1)
    outfile.y=(file.y).reshape(-1)
    outfile.z=(file.z).reshape(-1)
    outfile.trees=newID.reshape(-1)
    outfile.return_num=returnnumber.reshape(-1)
    outfile.close()
```

Appendix B

Modified PyCrown algorithm code to reduce oversegmentation

```
# INPUT FILE AND REFERENCE SYSTEM SETTING
System=*** #Reference System (EPSG)
inputFolder = r"C:\..."
outFolder=r"C:\..."
Trees_PRE = gpd.read_file(inputFolder+"\ABIES_050\\PRE\\PyCrown\\tree_location_top_cor.shp")
Trees_POST = gpd.read_file(inputFolder+"\ABIES_050\\POST\\PyCrown\\tree_location_top_cor.shp")
RadiusPRE = gpd.read_file(inputFolder+"\ABIES_050\\PRE\\PyCrown\\tree_crown_poly_smooth.shp").Radius
RadiusPRO = gpd.read_file(inputFolder+"\ABIES_050\\POST\\PyCrown\\tree_crown_poly_smooth.shp").Radius

# EXTRACT TREETOP POSITION
x=Trees_PRE['geometry'].x
y=Trees_PRE['geometry'].y
x_list = x.tolist()
y_list = y.tolist()
RadiusPRE_list= RadiusPRE.tolist()
IDindex_PRE = len(x)
ID_PRE_list=np.arange(IDindex_PRE).tolist()
Trees_PRE_Tops=(x_list, y_list, RadiusPRE_list, ID_PRE_list)
Trees_PRE_NewTops = list(zip(*Trees_PRE_Tops))
Trees_PRE_NewTops_array=np.array(Trees_PRE_NewTops)

xx=Trees_POST['geometry'].x
yy=Trees_POST['geometry'].y
xx_list = xx.tolist()
yy_list = yy.tolist()
```

```
RadiusPOST_list= RadiusPOST.tolist()
IDindex_POST = len(xx)
ID_POST_list=np.arange(IDindex_POST).tolist()
Trees_POST_Tops=(xx_list, yy_list, RadiusPOST_list, ID_POST_list)
Trees_POST_NewTops = list(zip(*Trees_POST_Tops))
Trees_POST_NewTops_array=np.array(Trees_POST_NewTops)

# MATCH THE CLOSEST "POST" TREETOP TO EACH "PRE" TREETOP
delta=0.1
pairings_PRE = {}
pairings_POST = {}
distances = {}
pairings = {}

Trees_POST_NewTops.sort()
def dist_squared(a, b):
   return (a[0] - b[1])**2 + (a[1] - b[0])**2
for idx, point in enumerate(Trees_PRE_NewTops):
   contenders = Trees_POST_NewTops[bisect(Trees_POST_NewTops,(point[0]-point[2]-
delta, 0)) : bisect(Trees_POST_NewTops,(point[0]+point[2]+delta, 0))] #point[2] è il raggio di
ricerca!
   contenders = list(map(lambda p: (p[1], p[0], p[2], p[3]), contenders))
   contenders.sort()
   contenders = contenders[bisect(contenders,(point[1]-point[2]-delta, 0)) :
bisect(contenders,(point[1]+point[2]+delta, 0))]
   matches = [(dist_squared(point, p2), point, p2) for p2 in contenders if
(dist_squared(point, p2) <= point[2]**2)]
   if matches:
      pairings_PRE[idx] = min(matches)[1]
      pairings_POST[idx] = min(matches)[2]
      distances[idx] = (min(matches)[0])**0.5
      pairings[idx] = min(matches)

same_PRE = list(pairings_PRE.items())
same_PRE_array = np.array(same_PRE)
data_array_PRE=same_PRE_array[:,1]
IDPRE=[k for x,y,z,k in data_array_PRE]
same_x_PRE = [x for x,y,z,k in data_array_PRE]
same_y_PRE = [y for x,y,z,k in data_array_PRE]
same_PREist=list(zip(IDPRE, same_x_PRE, same_y_PRE))
same_PRE_array = np.array(same_PRE_list)
same_POST = list(pairings_POST.items())
```

```
    same_POST_array = np.array(same_POST)
    data_array_POST=same_POST_array[:,1]
    IDPOST=[k for x,y,z,k in data_array_POST]
    same_x_POST = [y for x,y,z,k in data_array_POST]
    same_y_POST = [x for x,y,z,k in data_array_POST]
    same_POST_list=list(zip(IDPOST, same_x_POST, same_y_POST))
    same_POST_array = np.array(same_POST_list)

    MatchPRE=[]
    for i in range(len(Trees_PRE_NewTops_array)):
        for j in range(len(same_PRE_array)):
            if ((Trees_PRE_NewTops_array[i][0]==same_PRE_array[j][1]) and
(Trees_PRE_NewTops_array[i][1]==same_PRE_array[j][2])):
                MatchPRE=np.append(MatchPRE,Trees_PRE_NewTops_array[i], axis=0)
    MatchPRE=MatchPRE.reshape(-1,4)
    MatchPRE[:,[0, 3]] = MatchPRE[:,[3, 0]]
    MatchPRE[:,[1, 3]] = MatchPRE[:,[3, 1]]
    MatchPRE[:,[2, 3]] = MatchPRE[:,[3, 2]]

    MatchPOST=[]
    for i in range(len(Trees_POST_NewTops_array)):
        for j in range(len(same_POST_array)):
            if ((Trees_POST_NewTops_array[i][0]==same_POST_array[j][1]) and
(Trees_POST_NewTops_array[i][1]==same_POST_array[j][2])):
                MatchPOST=np.append(MatchPOST,Trees_POST_NewTops_array[i], axis=0)
    MatchPOST=MatchPOST.reshape(-1,4)
    MatchPOST[:,[0, 3]] = MatchPOST[:,[3, 0]]
    MatchPOST[:,[1, 3]] = MatchPOST[:,[3, 1]]
    MatchPOST[:,[2, 3]] = MatchPOST[:,[3, 2]]

    ## MATCHING ARRAY (IDPRE, XPRE, YPRE, IDPOST, XPOST, YPOST)
    Match=np.hstack((MatchPRE[:,:3],MatchPOST[:,:3]))

    # MATCH THE CLOSEST "PRE" TREETOP TO EACH "POST" TREETOP
    pairings_PRE_inv = {}
    pairings_POST_inv = {}
    distances_inv = {}
    pairings_inv = {}

    Trees_PRE_NewTops.sort()
    def dist_squared(a, b):
        return (a[0] - b[1])**2 + (a[1] - b[0])**2
```

```
    for idx, point in enumerate(Trees_POST_NewTops):
        contenders = Trees_PRENewTops[bisect(Trees_PRE_NewTops,(point[0]-point[2]-
delta, 0)) : bisect(Trees_PRE_NewTops,(point[0]+point[2]+delta, 0))] #point[2] è il raggio di
ricerca!
        contenders = list(map(lambda p: (p[1], p[0], p[2], p[3]), contenders))
        contenders.sort()
        contenders = contenders[bisect(contenders,(point[1]-point[2]-delta, 0)) :
bisect(contenders,(point[1]+point[2]+delta, 0))]
        matches = [(dist_squared(point, p2), point, p2) for p2 in contenders if
(dist_squared(point, p2) <= point[2]**2)]
        if matches:
          pairings_PRE_inv[idx] = min(matches)[2]
          pairings_POST_inv[idx] = min(matches)[1]
          distances_inv[idx] = (min(matches)[0])**0.5
          pairings_inv[idx] = min(matches)


    same_PRE = list(pairings_PRE_inv.items())
    same_PRE_array = np.array(same_PRE)
    data_array_PRE=same_PRE_array[:,1]
    IDPRE=[k for x,y,z,k in data_array_PRE]
    same_x_PRE= [y for x,y,z,k in data_array_PRE]
    same_y_PRE = [x for x,y,z,k in data_array_PRE]
    same_PRE_list=list(zip(IDPRE, same_x_PRE, same_y_PRE))
    same_PRE_array = np.array(same_PRE_list)

    same_POST = list(pairings_POST_inv.items())
    same_POST_array = np.array(same_POST)
    data_array_POST=same_POST_array[:,1]
    IDPOST=[k for x,y,z,k in data_array_POST]
    same_x_POST = [x for x,y,z,k in data_array_POST]
    same_y_POST = [y for x,y,z,k in data_array_POST]
    same_POST_list=list(zip(IDPOST, same_x_POST, same_y_POST))
    same_POST_array = np.array(same_POST_list)

    MatchPRE_inv=[]
    for i in range(len(Trees_PRE_NewTops_array)):
        for j in range(len(same_PRE_array)):
            if ((Trees_PRE_NewTops_array[i][0]==same_PRE_array[j][1]) and
(Trees_PRE_NewTops_array[i][1]==same_PRE_array[j][2])):
                MatchPRE_inv=np.append(MatchPRE_inv,Trees_PRE_NewTops_array[i],
axis=0)
```

```
    MatchPRE_inv=MatchPRE_inv.reshape(-1,4)
    MatchPRE_inv[:,[0, 3]] = MatchPRE_inv[:,[3, 0]]
    MatchPRE_inv[:,[1, 3]] = MatchPRE_inv[:,[3, 1]]
    MatchPRE_inv[:,[2, 3]] = MatchPRE_inv[:,[3, 2]]

    MatchPOST_inv=[]
    for i in range(len(Trees_POST_NewTops_array)):
        for j in range(len(same_POST_array)):
            if ((Trees_POST_NewTops_array[i][0]==same_POST_array[j][1]) and
(Trees_POST_NewTops_array[i][1]==same_POST_array[j][2])):
                MatchPOST_inv=np.append(MatchPOST_inv,Trees_POST_NewTops_array[i],
axis=0)
    MatchPOST_inv=MatchPOST_inv.reshape(-1,4)
    MatchPOST_inv[:,[0, 3]] = MatchPOST_inv[:,[3, 0]]
    MatchPOST_inv[:,[1, 3]] = MatchPOST_inv[:,[3, 1]]
    MatchPOST_inv[:,[2, 3]] = MatchPOST_inv[:,[3, 2]]

    ## MATCHING ARRAY (IDPRE, XPRE, YPRE IDPOST, XPOST, YPOST)
    Match_inv=np.hstack((MatchPRE_inv[:,:3],MatchPOST_inv[:,:3]))

    # DELETE DUPLICATES MATCHES
    index_inv=[]
    for i in range(len(Match_inv)):
        for j in range(len(Match_inv)):
            if Match_inv[i][0]==Match_inv[j][0] and i!=j:
                index_inv=np.append(index_inv,Match_inv[i][:], axis=0)
    index_inv=index_inv.reshape(-1,6)
    index_inv = np.unique(index_inv, axis=0)

    corretto=[]
    for i in range(len(index_inv)):
        for j in range(len(Match)):
            if index_inv[i][0]==Match[j][0]:
                corretto=np.append(corretto,(Match[j][:])).reshape(-1,6)
    corretto=np.unique(corretto,axis=0)

    Match_inc = np.array([x for x in Match_inv if x not in index_inv])

    Match_new = np.vstack((Match_inc, corretto))
    ID = len(Match_new)
    ID=np.arange(ID).reshape(ID,1)
    ID=ID+5000
```

```
## FINAL ARRAY OF MATCHING TREETOPS (ID5000, IDPRE, XPRE, YPRE, IDPOST, XPOST, YPOST)
Match_fin=np.hstack((ID, Match_new))

# FINAL ARRAY OF NON MATCHING TREETOPS

## PRE SCENARIO ARRAY  (ID1000, IDPRE, XPRE, YPRE)
index_nomatchPRE=[]
NoMatchPRE_pre=Trees_PRE_NewTops_array
for i in range(len(Trees_PRE_NewTops_array)):
   for j in range(len(Match_fin)):
      if ((Trees_PRE_NewTops_array[i][0]==Match_fin[j][2]) and (Trees_PRE_NewTops_array[i][1]==Match_fin[j][3])):
         index_nomatchPRE=np.append(index_nomatchPRE,[i], axis=0)
index_nomatchPRE=index_nomatchPRE.reshape(-1,1)
index_nomatchPREindex_nomatchPRE.astype(int)
NoMatchPRE_pre=np.delete(Trees_PRE_NewTops_array, index_nomatchPRE, axis=0)
ID_NoMatchPRE=len(NoMatchPRE_pre)
ID_NoMatchPRE=np.arange(ID_NoMatchPRE).reshape(ID_NoMatchPRE,1)
ID_NoMatchPRE=ID_NoMatchPRE+1000
NoMatchPRE = np.hstack((ID_NoMatchPRE, NoMatchPRE_pre[:,3:]))
NoMatchPRE = np.hstack((NoMatchPRE, NoMatchPRE_pre[:,:2]))

## POST SCENARIO ARRAY (ID2000, IDPOST, XPOST, YPOST)
index_nomatchPOST=[]
NoMatchPOST_pre=Trees_POST_NewTops_array
for i in range(len(Trees_POST_NewTops_array)):
   for j in range(len(Match_fin)):
      if ((Trees_POST_NewTops_array[i][0]==Match_fin[j][5]) and (Trees_POST_NewTops_array[i][1]==Match_fin[j][6])):
         index_nomatchPOST=np.append(index_nomatchPOST,[i], axis=0)
index_nomatchPOST=index_nomatchPOST.reshape(-1,1)
index_nomatchPOST=index_nomatchPOST.astype(int)
NoMatchPOST_pre=np.delete(Trees_POST_NewTops_array, index_nomatchPOST, axis=0)
ID_NoMatchPOST=len(NoMatchPOST_pre)
ID_NoMatchPOST=np.arange(ID_NoMatchPOST).reshape(ID_NoMatchPOST,1)
ID_NoMatchPOST=ID_NoMatchPOST+2000
NoMatchPOST = np.hstack((ID_NoMatchPOST, NoMatchPOST_pre[:,3:]))
NoMatchPOST = np.hstack((NoMatchPOST, NoMatchPOST_pre[:,:2]))
```

```
    ## ASSIGN THE SAME ID TO SAME TREES AND SAVE THE OUTPUT
    inFile_PRE_Final = laspy.file.File(inputFolder+"\ABIES_050\\PRE\\PyCrown\\Final_Trees_NO_Overseg.las", mode="r")
    inFile_POST_Final = laspy.file.File(inputFolder+"\ABIES_050\\POST\\PyCrown\\Final_Trees_NO_Overseg.las", mode="r")

    trees_PRE = (inFile_PRE_Final.trees).reshape(-1,1)
    x_PRE = (inFile_PRE_Final.x).reshape(-1,1)
    y_PRE = (inFile_PRE_Final.y).reshape(-1,1)
    z_PRE= (inFile_PRE_Final.z).reshape(-1,1)
    FilePRE = np.hstack((x_PRE, y_PRE, trees_PRE))
    #(ID(5000 e 1000), IDPRE, XPRE, YPRE)
    MatchNoMatchPRE= np.vstack((Match_fin[:,:4], NoMatchPRE))

    trees_POST = (inFile_POST_Final.trees).reshape(-1,1)
    x_POST = (inFile_POST_Final.x).reshape(-1,1)
    y_POST = (inFile_POST_Final.y).reshape(-1,1)
    z_POST = (inFile_POST_Final.z).reshape(-1,1)
    FilePOST = np.hstack((x_POST, y_POST, trees_POST))
    MatchNoMatchPOST= np.hstack((Match_fin[:,:1], Match_fin[:,4:]))
    #(ID(5000 e 2000), IDPOST, XPOST, YPOST)
    MatchNoMatchPOST= np.vstack((MatchNoMatchPOST, NoMatchPOST))

    classificationPRE=[]
    for x in range(len(FilePRE)):
       for j in range(len(MatchNoMatchPRE)):
          if FilePRE[x][2]==MatchNoMatchPRE[j][1]:
               NewClass=MatchNoMatchPRE[j][0]
               classificationPRE.append(NewClass)

    classificationPOST=[]
    for x in range(len(FilePOST)):
       for j in range(len(MatchNoMatchPOST)):
          if FilePOST[x][2]==MatchNoMatchPOST[j][1]:
               NewClass=MatchNoMatchPOST[j][0]
               classificationPOST.append(NewClass)

    header = inFile_PRE_Final.header
```

```
        outfile = laspy.file.File(outFolder+"\ABIES_050\\PRE\\PyCrown\\Final_Trees_NO_Overseg_matched.las", mode="w", header=header)
        outfile.x=x_PRE.reshape(-1)
        outfile.y=y_PRE.reshape(-1)
        outfile.z=z_PRE.reshape(-1)
        outfile.trees=np.array(classificationPRE).reshape(-1).astype(int)
        outfile.close()

        header = inFile_POST_Final.header
        outfile = laspy.file.File(outFolder+"\ABIES_050\\POST\\PyCrown\\Final_Trees_NO_Overseg_matched.las", mode="w", header=header)
        outfile.x=x_POST.reshape(-1)
        outfile.y=y_POST.reshape(-1)
        outfile.z=z_POST.reshape(-1)
        outfile.trees=np.array(classificationPOST).reshape(-1).astype(int)
        outfile.close()
```

Appendix C

Table 40: Estimated and reference values of stocked CO_2 - 2013 scenario. (Case study A).

Tree ID	$CO_{2Jucker}$ [kg]	CO_{2Chave} [kg]	CO_{2QSM} [kg]	$CO_{2Reference}$ [kg]
104	692	460	19969	859
105	1065	821	5657	1299
108	266	128	2051	354
113	761	495	32558	630
114	551	434	15692	466
119	465	351	15985	310
120	648	404		662
128	926	664	16718	1025
129	409	299	14145	375
132	522	322	1914	327
134	226	105	8014	112
135	319	211	20338	228
136	709	426	8745	498
137	751	505	28892	1087
504	796	540	5769	663
507	303	147	4943	115
509	1617	1183		561
513	611	422		333
514	477	362		358
515	946	663	31410	709
516	510	394		462
518	752	483	10793	623

Table 40: Estimated and reference values of stocked CO_2 - 2013 scenario. (Case study A).

Tree ID	$CO_{2Jucker}$ [kg]	CO_{2Chave} [kg]	CO_{2QSM} [kg]	$CO_{2Reference}$ [kg]
522	397	289	21573	331
525	258	168		179
527	210	93	27171	100
530	535	339	16380	211
531	598	378	29351	495
532	271	120	16735	194
533	946	581	13406	612

Table 41: Statistical indexes for stocked CO_2 - 2013 scenario. (Case study A).

Assessment	RMSE	BIAS
$CO_{2Jucker}$	260 kg	42,5 %
CO_{2Chave}	226 kg	-9,2 %
CO_{2QSM}	17987 kg	4814,8 %

Table 42: Estimated and reference values of stocked CO_2 - 2014 scenario. (Case study A).

Tree ID	$CO_{2Jucker}$ [kg]	CO_{2Chave} [kg]	CO_{2QSM} [kg]	$CO_{2Reference}$ [kg]
504	694	433	4662	600
507	166	71		95
509	598	419	3577	478
513	355	210	1466	267
514	421	311	2864	343
515	448	140	6127	290
516	786	513	2144	411
518	656	482	6082	702
522	465	276	8728	276
525	343	183	7008	151
527	361	212	6497	141
530	255	97	5915	116
531	511	339	1971	509
532	514	300	9667	275
533	185	72	2023	319

Table 43: Statistical indexes for stocked CO_2 - 2014 scenario. (Case study A).

Assessment	RMSE	BIAS
$CO_{2Jucker}$	169 kg	55,2 %
CO_{2Chave}	120 kg	-13,3 %
CO_{2QSM}	5249 kg	1933,7 %

Appendix D

Table 44: Estimated and reference values of stocked CO_2 - 2020 scenario. (Case study C).

Tree ID	CO_{2Chave} [kg]	$CO_{2Kangas}$ [kg]	CO_{2QSM} [kg]
1	94	65	44
2	220	150	29
3	307	200	229
4	214	148	236
5	182	126	88
6	144	99	75
7	157	109	144
8	140	98	154
9	95	68	183
10	133	92	112
11	66	46	26
12	188	133	17
13	84	60	
14	148	105	30
15	98	69	56
16	120	83	116
17	133	92	63
18	63	45	8
19	53	38	46
20	139	96	92

Table 45: Statistical indexes for stocked CO_2 - 2020 scenario. (Case study C).

Assessment	Reference	RMSE	BIAS
CO_{2QSM}	$CO_{2Kangas}$	60 kg	-2,6 %
	CO_{2Chave}	80 kg	-32,1%

Table 46: Estimated and reference values of stocked CO_2 - 2021 scenario. (Case study C)

Tree ID	CO_{2Chave} [kg]	$CO_{2Kangas}$ [kg]	CO_{2QSM} [kg]
2	208	144	30
4	217	150	240
5	192	131	97
8	149	103	153
10	137	95	118
11	62	44	45
12	174	125	26
15	96	67	70
16	124	85	123
17	131	91	91
18	55	40	20
19	66	47	42

Table 47: Statistical indexes for stocked CO_2 - 2021 scenario. (Case study C).

Assessment	Reference	RMSE	BIAS
CO_{2QSM}	$CO_{2Kangas}$	55 kg	-5,1 %
	CO_{2Chave}	75 kg	33,9 %