



TUSCIA UNIVERSITY

Dept. for Innovation in Biological, Agro-food and Forest systems (DIBAF)

in agreement with the **UNIVERSITY OF MOLISE**

Dept. of Bioscience and Territory (DiBT)

PhD course in

Science, Technology and Biotechnology for Sustainability- Cycle XXXIII

Curriculum: Forest Ecology

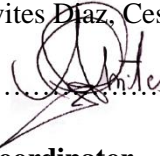
Dissertation title:

LiDAR as a tool for timber assortment assessment and characterization in mountain forests

(s.s.d. AGR/05)

PhD candidate:

Dr: Alvites Diaz, Cesar Ivan

.....


PhD Coordinator

Prof. Andrea Vannini

.....

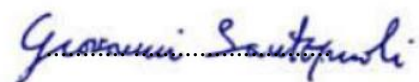

Thesis Supervisor

Prof. Bruno Lasserre

.....


Thesis Co-tutor

Prof. Giovanni Santopoli

.....


A.Y. 2019/20

Doctoral advisors

Tutor: Prof. Bruno Lasserre

Associate Professor
University of Molise - Department of Biosciences and Territory
Pesche (Is), Italy

Co-Tutor: Prof. Giovanni Santopuoli

Researcher
University of Molise - Department of Agricultural, Environmental and Food Sciences
Campobasso, Italy

Abstract

LiDAR as a tool for timber assortment assessment and characterization in mountain forests

Cesar Ivan Alvites Diaz (cesar.alvites@unimol.it)

Dept. of Bioscience and Territory, University of Molise, Pesche (IS), Italy.

The timber assortment estimation from forests offers socio-economic and environmental benefits to humans. Its accurate assessment supports the better allocation and use of timber and reduces timber waste. It is of primordial importance in the sustainable management of forests, and the conservation of biodiversity over time. In the past, the traditional non-destructive estimation method was commonly used, even if highly complex and with low performance, especially in natural and uneven-aged forests and old-intact forests. Now, in the present era, reliable and realistic representation of trees is possible through active remote sensing techniques such as Light Detection and Ranging (LiDAR). However, it has many operational and technical complexities (i.e., require expertise in data collection and processing), and the complexity of forest stand conditions is a further challenge (i.e., forest structure). These complexities are addressed in many studies, but a reliable method for extracting the timber assortment information using LiDAR data is still lacking. Hence, such a method would be extremely useful for valorising the timber resources and for promoting sustainable management activities.

The aim of the thesis was to develop a robust procedure for timber assortment estimation of trees in a mixed-species and multi-layered forest using LiDAR data. To achieve the objective, the thesis is divided into three sub-objectives. The first research aim was to provide an overview of the most recent approaches used in timber assortment estimation using LiDAR data, through a literature review; the second study aim was to provide a stepwise approach to assess the stem volume and carbon stock at single tree level using Airborne Laser Scanning (ALS) data, while the third study aim was to provide a stepwise approach to retrieve the timber assortment information from forest stands using Terrestrial Laser scanning (TLS) data. These studies were carried out in Bosco Pennataro, a Mediterranean forest belonging to the Central Apennine Mountains in Molise (Italy).

The first task on literature study proved that most research studies were focused on forest inventory (45.25 %) and forest productivity (23.46 %) topics, while few studies investigated the relation between the timber assortment and biodiversity conservation (7.26 %). Overall, the two most used LiDAR devices were ALS and TLS. Since 2010, machine learning algorithms have become essential for predicting, upscaling, modelling and classifying the LiDAR data as highlighted in many studies. The second task highlighted that the heterogeneity of forest structure (vertically and horizontally tree profiles) and the point cloud density are crucial for detecting single trees. The average detection rate was 48 % meaning that about half of trees were detected. The detection rate was higher in forests with heterogeneous structures regardless of the density of point clouds, reaching values between 0.49 and 0.65. The carbon stock was accurately predicted, with a bias ranging from -0.3 % to 1.5 %. The third task allowed to demonstrate that TLS greatly supports tree detection, recognizing 84.4% of trees, reconstructing 67% of detected trees and correctly quantifying 75% of merchantable logs.

It is worthy to note that the two procedures (second and third studies) proposed in this thesis were tested for the first time for forest monitoring, especially in mixed-species and multi-layered forests. The implemented approach for detecting

trees using ALS data was found slightly more accurate in comparison with a study carried out previously in the same study area; improvement was more visible in understory trees. The unsupervised algorithms used in the detection approach allowed to identify the trees without previous knowledge of tree position and in a fast way. The stepwise approach applied to TLS data proved to be efficient for extracting the timber assortments for many tree species, especially the *Q. cerris*. Moreover, this approach provided many insights into TLS data improvement and use for timber assortment assessment. For instance, timber-leaf discrimination in the forest is possible through machine learning even in high species richness conditions. This thesis highlights the usefulness of LiDAR data for accurately and directly representing the timber assortment resources, avoiding destructive methods like cutting and felling of trees. This LiDAR technique also supports sustainable forest planning and management.

Keywords: Sustainable forest management, remote sensing, Mediterranean forest, airborne laser scanning, terrestrial laser scanning.

Contents

<i>Acronyms</i>	6
<i>Tables and Figures</i>	8
<i>Extended abstract</i>	12
Research framework	12
Material and Methods	13
Results	14
Final remarks	17
1. CHAPTER 1- Introduction	19
2. CHAPTER 2 – Paper review and scientific papers	22
2.1. Challenges and opportunities for timber assortment evaluation through LiDAR. A review.	22
2.1.1. Preface	22
2.1.2. Abstract	22
2.1.3. Introduction	23
2.1.4. Material and Methods	25
2.1.5. Results	28
2.1.6. Discussion	39
2.1.7. Conclusion	42
2.1.8. Postface	43
2.2. Unsupervised algorithms to detect single trees in a mixed-species and multi-layered Mediterranean forest using LiDAR data.	45
2.2.1. Preface	45
2.2.2. Abstract	45
2.2.3. Introduction	45
2.2.4. Materials and Methods	47
2.2.5. Results	55
2.2.6. Discussion	64
2.2.7. Conclusion	66
2.2.8. Postface	67
2.3. A stepwise approach for deriving timber assortment of trees from Terrestrial Laser Scanning data	68
2.3.1. Preface	68
2.3.2. Abstract	68
2.3.3. Introduction	69
2.3.4. Materials and Methods:	71
2.3.5. Results	82
2.3.6. Discussion	94
2.3.7. Conclusion	98
2.3.8. Postface	99
3. CHAPTER 3 – CONCLUSION AND FUTURE PERSPECTIVES	100
4. LIST OF ARTICLES	102
5. ACKNOWLEDGEMENTS	103
6. REFERENCES	104

Acronyms

- **A:** Accuracy
- **ABA:** Area-Based Approach
- **ADS:** Field plot
- **AGB:** Above Ground Biomass
- **AGCC:** Above Ground Carbon Content
- **AIC:** Akaike Information Criterion
- **ALS:** Airborne Laser Scanning
- **ATLAS:** Advanced Topographic Laser Altimeter System
- **B:** Biodiversity
- **BA:** Basal Area
- **BEF:** Biomass Expansion Factor
- **C:** Climate Change
- **CHM:** Canopy Height Model
- **CIR:** Color Infrared
- **CLR:** Curve Length Ratio
- **CNPq:** Council of Technological and Scientific Development
- **CPA:** Canopy Projection Area
- **CS:** Carbon Stock
- **CSF:** Climate Smart Forestry
- **DBH:** Diameter at breast height
- **DBSCAN:** Density Based Spatial Clustering of Application with Noise
- **DEM:** Digital Elevation Model
- **Dmax:** Maximum End Diameter
- **Dmin:** Minimum End Diameter
- **DR:** Detection Rate
- **DSM:** Digital Surface Model
- **EVI:** Echidna(R) Validation Instrument
- **FalseNeg:** False Negative
- **FalsePos:** False Positive
- **FIVs:** Forest Inventory Variables
- **GEDI:** Global Ecosystem Dynamics Investigation
- **GEOBIA:** GEOgraphic Object-Based Image Analysis
- **GLAS:** Geoscience Laser Altimeter System
- **GLM:** Generalized Linear Model
- **GS:** Growing Stock
- **h:** Perpendicular distance between the highest convex curve and the straight line
- **I:** Inventory
- **ICESat:** Ice, Clouds, and Land Elevation Satellite
- **ITD:** Individual Tree Detection
- **k-NN:** k Nearest Neighbors
- **L:** Length of log
- **TreeTLS:** Number of stems extracted from TLS data
- **TruePos:** True Positive
- **TrueRStem:** Rates of RStem
- **TTv, Trunk:** Trunk Volume
- **TTv.log:** Log volume
- **LAI:** Leaf Area Index
- **LCR:** Locally Calibrated Regression
- **LiDAR:** Light Detection and Ranging
- **LMws:** Local Maxima Filter with Variable Window Size
- **L_n:** Local Neighbourhoods
- **MaB:** Man and Biosphere
- **MCGC:** Multi-Class Graph Cut
- **MinPts:** Minimum Number of Points
- **MLS:** Mobile Laser Scanning
- **MMAC:** Multi-Level Morphological Active Contour
- **MOLI:** Multi-footprint Observation LiDAR and Imager
- **MS-STI:** Multisource Single Tree Inventory
- **Mtry:** Number of Input Variables
- **Nodesize:** Minimum Size of Terminal Node
- **N^otree:** Number of Decision Trees
- **OBIA:** Object Based Image Analysis
- **OPALS:** Orientation and Processing of Airborne Laser Scanning
- **P:** Productivity
- **PAI:** Plant Area Index
- **PCA:** Principal Component Analysis
- **PHC:** Percent of the Tree Height Covered
- **PLS:** Portable Laser Scanning
- **PPM:** Parameter Prediction Method
- **r:** Coefficient of Correlation
- **R:** Review
- **RANSAC:** RANdom SAmple Consensus
- **RF:** Random Forests
- **RMSE:** Root Means Squared Error
- **RS:** Remote Sensing
- **R-squared:** Coefficient of Determination
- **RStem:** Reconstructed stem from TLS data
- **SD:** stem density
- **SFM:** Sustainable Forest Management
- **STR:** Straightness
- **SVR:** Support Vector Regression
- **SWA:** Spatial Wavelet Analysis
- **TAP:** Tapering
- **TDC:** Tree Crown Diameter
- **TH:** Tree Height
- **TH1:** First Attached Branch or Branch Union
- **THbase:** tree height at 50 cm from ground
- **Time for TD:** Time for Tree Detection
- **TLS:** Terrestrial Laser Scanning

- **Tw**: Left Truncated Distribution Truncated
- **VIF**: Variance Inflation Factor
- **Vis**: Vegetation Index
- **VOL**: Stem Volume
- **WBD**: Wood Basal Density
- **WFD**: Waveform Digitising
- **wsrf**: Weighted Subspace Random Forests for Classification

Tables and Figures

- TABLE 1 DESCRIPTION OF KEYWORDS USED TO RETRIEVE THE INFORMATION ABOUT LIDAR STUDIES. ADDITIONAL DETAILS ABOUT THE TERMS AND DEFINITIONS (A) ARE AVAILABLE AT [HTTP://WWW.FAO.ORG/FORESTRY/FRA2015/](http://www.fao.org/forestry/fra2015/)-----26
- TABLE 2 LITERATURE REVIEW STRINGS. ADVANCED DESCRIPTION OF CODES (SC) USED TO RETRIEVE THE INFORMATION ABOUT LIDAR STUDIES. THE FIXED WORDS ARE SHOWED IN ITALIC.-----26
- TABLE 3 NUMBER OF PAPERS COLLECTED BY DIFFERENT CODES. EXCLUDING THE REPEATED PAPERS, THE FINAL NUMBER OF PAPERS USED IN THE ANALYSIS WAS REPORTED FOR EACH RESEARCH CODE.-----28
- TABLE 4 FOREST-STAND CONDITION INFORMATION FOR EACH CLUSTER ANALYSED. REVIEW “R” PAPERS (11) WERE NOT CONSIDERED HERE -----30
- TABLE 5 MOST FREQUENT METHODS FOR PROCESSING LIDAR DATA. THE PRE-PROCESSING METHODS (1), TREE DETECTION APPROACHES (2) AND MODELLING OF THE TIMBER PROVISION VARIABLES (3) WERE DISPLAYED.-----37
- TABLE 6 SUMMARY OF FOREST STAND CHARACTERISTICS OF AIRBORNE LASER SCANNING (ALS) AND FOREST INVENTORY DATA PER EACH FIELD PLOT (ADS) AND COMPLEXITY CATEGORIES. THE AVERAGE POINT DENSITY (APD; POINTS M⁻²), AVERAGE POINT SPACING (APS; M), DIAMETER AT BREAST HEIGHT (DBH; CM) AND TREE HEIGHT (TH; M) WERE ESTIMATED PER ADS. THE STEM VOLUME (VOL; M³) AND CARBON STOCK (CS; TONS) WERE ESTIMATED PER HECTARE (HA.). THE NUMBER OF TREES (N°TREES; UNITS) WERE CALCULATED PER ADS AND HA. THE MEAN (*1) AND SUM (*2) AND STANDARD DEVIATION (*3) VALUES WERE SHOWED.-----56
- TABLE 7 TREE DETECTION RESULTS. NUMBER OF STEMS OBSERVED FROM REFERENCE DATA (TR; UNITS) AND NUMBER OF STEMS PREDICTED FROM ALS DATA (TALS; UNITS), TRUE POSITIVE (TRUEPOS; UNITS), FALSE POSITIVE (FALSEPOS; UNITS), FALSE NEGATIVE (FALSENEG; UNITS) AND DETECTION RATE (DR; %) FOR LOWER (LAYER1), INTERMEDIATE (LAYER2) AND UPPER (LAYER3) CANOPY LAYERS.-----59
- TABLE 8 SUMMARY STATISTICS OF THE FOREST INVENTORY VARIABLES ESTIMATED WITH THE RANDOM FORESTS ALGORITHM BY USING TOP-NINE METRICS FOR DIAMETER AT BREAST HEIGHT (DBH; CM), TREE HEIGHT (TH; M), AND STEM VOLUME (VOL; M³). THE NUMBER OF TREES (N°TREES; UNITS), COEFFICIENT OF DETERMINATION (R-SQUARED; 0-1) AND ROOT MEAN SQUARED ERROR (RMSE; CM, M AND M³) WERE DISPLAYED. THE OUTCOMES WERE DISPLAYED FOR ALL FOUR CATEGORIES (HIGHLY DIFFICULT, MODERATELY EASY, HIGHLY EASY AND MODERATELY DIFFICULT), WHICH WAS FURTHER DIVIDED BY LOWER (LAYER1), INTERMEDIATE (LAYER2) AND UPPER (LAYER3) CANOPY LAYERS.-----61
- TABLE 9 COMPARISON BETWEEN PREDICTED AND OBSERVED VALUES OF STEM VOLUME (VOL; M³) AND CARBON STOCK (CS; TONS) DERIVED FROM AIRBORNE LASER SCANNING (ALS) METRICS. THESE VALUES WERE FURTHERLY DISPLAYED BY EACH CATEGORY (HIGHLY DIFFICULT, MODERATELY EASY, HIGHLY EASY AND MODERATELY DIFFICULT) AND BY EACH FIELD PLOT (ADS). THE NUMBER OF TREES (N°TREES; UNITS) WAS CALCULATED PER ADS AND HA-1. ABSOLUTE (M³ AND TONS) AND PERCENT (%) VALUES OF BIAS AND ROOT MEANS SQUARED ERROR (RMSE) WERE EVEN DISPLAYED.-----62
- TABLE 10 TIMBER ASSORTMENT CHARACTERISTICS BASED ON GEOMETRY WOOD DEFECTS. THE STRAIGHTNESS (STR) AND DIAMETER MINIMUM OF ENDS LOGS (DMIN)

WERE DISPLAYED. “X” REPRESENTS THE VALUES NEEDED FOR STR AND DMIN CHARACTERISTICS.-----75

TABLE 11 SUMMARY OF FOREST-RELATED CHARACTERISTIC FROM ALL TREES OVER FIVE FIELD PLOTS (ADS). THE DIAMETER AT BREAST HEIGHT (DBH), TREE HEIGHT (TH), HEIGHT OF THE FIRST ATTACHED BRANCH OR BRANCH UNION (TH1), TREE STEM VOLUME (TSV) AND THE TREE SPECIES COMPOSITION (TSC) WERE SHOWED. -----83

TABLE 12 LOG QUALITY TRAITS FOR ALL FIVE STUDY AREAS (ADS). THE STRAIGHTNESS (STR) AND TAPERING (TAP) VARIABLES, AND THE LOG VOLUME (TTV.LOG) OF MERCHANTABLE AND NON-MERCHANTABLE LOGS WERE DISPLAYED. THE MEAN, STANDARD DEVIATION (SD ±) AND THE SUM WERE USED FOR EVALUATING THE ACCURACY.-----84

TABLE 13 SUMMARY OF TREE DETECTION RESULTS. OBSERVED TREES FROM FIELD DATA (TR), PREDICTED TREES FROM TLS DATA (TREETLS), TRUE POSITIVE (TRUEPOS), FALSE POSITIVE (FALSEPOS), FALSE NEGATIVE (FALSENEG), DETECTION RATE (DR), COMPLETENESS AND CORRECTNESS FOR EACH STUDY AREA (ADS). MEAN, STANDARD DEVIATION (SD ±) AND SUM WERE ALSO DISPLAYED. -----85

TABLE 14 DETECTION ACCURACY. OBSERVED TREES FROM FIELD DATA (TR, UNITS), TRUE POSITIVE (TRUEPOS, UNITS) AND DETECTION RATE (DR, %) WERE SHOWED FOR THE THREE DIFFERENT DIAMETER AT BREAST HEIGHT (DBH) INTERVALS. MEAN, STANDARD DEVIATION (SD ±) AND SUM WERE ALSO DISPLAYED. -----86

TABLE 15 STEM RECONSTRUCTION RESULTS. OBSERVED TREES FROM FIELD DATA (TR, UNITS), RECONSTRUCTED STEM FROM TERRESTRIAL LASER SCANNING DATA (RSTEM; UNITS) AND RATE OF RSTEM (TRUERSTEM, PERCENT) WERE DISPLAYED FOR FIVE DIFFERENT DIAMETER AT BREAST HEIGHT (DBH) INTERVALS AND IT IS SEPARATED BY EACH STUDY ARE (ADS). MEAN, STANDARD DEVIATION (SD ±) AND SUM WERE ALSO DISPLAYED. -----87

TABLE 16 LOG QUANTIFICATION RESULTS. THE NUMBER OF LOGS (N°LOGS, UNITS) AND THE LENGTH OF LOG (L.LOG, M) MEASUREMENTS WERE DISPLAYED FOR MERCHANTABLE AND NON-MERCHANTABLE LOGS. MEAN, STANDARD DEVIATION (SD ±) AND SUM WERE ALSO DISPLAYED.-----90

TABLE 17 LOG RESULTS SHOWED FOR TREE SPECIES. -----91

FIGURE 1 WORKFLOW OF THE PAPER COLLECTION, CLASSIFICATION AND ANALYSIS -----25

FIGURE 2 TREND OF THE ARTICLES PUBLISHED IN THE TIMEFRAME 2000-2018 (END JULY 2018).-----28

FIGURE 3 GEOGRAPHICAL REPRESENTATION OF LIDAR APPLICATIONS. THE BOX IN THE TOP SHOWS A GLOBAL OVERVIEW, WHILE THE BOX A SHOWS THE STUDIES CARRIED OUT IN NORTH AMERICA, WHILE BOX B SHOWS THE STUDIES CARRIED OUT IN ITALY. ---29

FIGURE 4 PAPER FREQUENCY BELONGING THE SIX CLUSTERS: ACCURACY (A), BIODIVERSITY (B), CLIMATE CHANGE (C), INVENTORY (I), PRODUCTIVITY (P), REVIEW (R).-----30

FIGURE 5 STACKED BAR GRAPH OF RS DEVICES USED IN THE PERIOD (2000-2018).-----32

FIGURE 6 LOCATION OF STUDY AREA BOSCO PENNATARO (RED TRIANGLE) AND LOCATION OF THE FIELD PLOTS (ADS). -----48

FIGURE 7 METHODOLOGICAL WORKFLOW APPLIED TO DERIVE THE CARBON STOCK AT THE SINGLE TREE LEVEL, USING AIRBORNE LASER SCANNING (ALS) DATA. THE ALS DATA

WAS CUT USING THE FIELD PLOT (ADS) BOX DIMENSIONS AND STRATIFIED INTO LOWER (LAYER1), INTERMEDIATE (LAYER2) AND UPPER (LAYER3) CANOPY LAYERS. THE ROOT MEAN SQUARED ERROR (RMSE) AND COEFFICIENT OF DETERMINATION (R-SQUARED) VALUES FOR STEM VOLUME PREDICTION WAS EVEN DISPLAYED-----50

FIGURE 8 WORKFLOW OF THE PROCESSING OF THE AIRBORNE LASER SCANNING (ALS) POINT CLOUD FOR EVERY CANOPY LAYER (LAYER1, LAYER2 AND LAYER3) WITHIN EACH FIELD PLOT (ADS). THE DIAMETER AT BREAST HEIGHT (DBH) AND TREE HEIGHT (TH) WERE CONSIDERED IN THE CATEGORIZATION STEP-----51

FIGURE 9 THE PROCESSING OF THE AIRBORNE LASER SCANNING (ALS) POINT CLOUD THROUGH DENSITY-BASED SPATIAL CLUSTERING OF APPLICATIONS WITH NOISE (DBSCAN) ALGORITHM. THE MINIMUM NUMBER OF POINTS (MINPTS) AND THE EPS NEIGHBORHOOD DISTANCE (NEPS(P)) THRESHOLDS WERE CONSIDERED.-----52

FIGURE 10 WORKFLOW OF THE PROCESSING FOR DETECTING THE TREES ACROSS THE THREE CANOPY LAYERS (I.E. LOWER LAYER: LAYER1, INTERMEDIATE LAYER: LAYER2 AND UPPER LAYER: LAYER3) FROM AIRBORNE LASER SCANNING (ALS) POINT CLOUD. THE MINIMUM NUMBER OF POINTS (MINPTS) AND EPS NEIGHBORHOOD DISTANCE (EPS) THRESHOLDS WERE USED FOR PROCESSING DENSITY-BASED SPATIAL CLUSTERING OF APPLICATIONS WITH NOISE (DBSCAN) AND K-MEANS. ----53

FIGURE 11 GRAPHICAL DISTRIBUTION OF THE FIELD PLOTS (ADS) ACCORDING TO THE AVERAGE POINT DENSITY (APD; POINTS M-2) AND THE STANDARD DEVIATION OF TREE HEIGHT (THSD; M) FOR EACH CATEGORY FROM A TO D GROUPS (A, HIGHLY DIFFICULT; B, MODERATELY EASY; C, HIGHLY EASY; D, MODERATELY DIFFICULT). 56

FIGURE 12 FOUR REPRESENTATIVE MAPS OF THE FOUR DIFFERENT AIRBORNE LASER SCANNING (ALS) POINT CLOUD COMBINATIONS (ONE PER CATEGORY). THE RED SQUARE SHOWED THE FIELD PLOT (ADS) BORDER; THE NUMBER OF TREES (N°TREES; UNITS) WAS SHOWED FOR EVERY CANOPY LAYER (I.E. LOWER LAYER: LAYER1, INTERMEDIATE LAYER: LAYER2 AND UPPER LAYER: LAYER3); THE TREE HEIGHT (TH; M) AND THE DIAMETER AT BREAST HEIGHT (DBH; CM) WERE EXPRESSED IN AVERAGE AND THE STANDARD DEVIATION (SD; ±) VALUES; THE TOP LETTERS REPORT THE CATEGORY LEVEL (I.E. A, HIGHLY DIFFICULT; B, MODERATELY EASY; C, HIGHLY EASY; D, MODERATELY DIFFICULT).-----58

FIGURE 13 COMPARISON BETWEEN PREDICTED VS. OBSERVED VALUES OF I) DETECTION RATE (DR; %), II) EUCLIDEAN DISTANCE (M), III) TREE CROWN OVERLAP (%), AND IV) TIME FOR TREE DETECTION (TIME FOR TD; SEC) FOR EACH CANOPY LAYER (LAYER1 “L1”, LAYER2 “L2”, AND LAYER3 “L3”) AND FOR EVERY CATEGORY (HIGHLY DIFFICULT, MODERATELY EASY, HIGHLY EASY AND MODERATELY DIFFICULT). THE AVERAGE VALUES OF THE TIME CONSUMING FOR DETECTING TREE (TIME FOR TD, SEC) BELONGING TO EACH PLOT WAS DISPLAYED FOR EACH CATEGORY. -----60

FIGURE 14 PREDICTED VALUES VS. OBSERVED FOREST INVENTORY VARIABLES. THE BOX A) SHOWS DIAMETER AT BREAST HEIGHT (DBH, CM); BOX B) SHOWS TREE HEIGHT (TH; M) AND BOX C) DISPLAYS STEM VOLUME (VOL, M³). THE NUMBER OF TREES (N°TREES; UNITS), NUMBER OF PREDICTORS (N°PREDICTORS; UNITS), COEFFICIENT OF DETERMINATION (R-SQUARED; 0-1) AND ROOT MEAN SQUARED ERROR (RMSE; CM, M AND M³) WERE EVEN REPORTED. -----61

FIGURE 15 A) SHOWS THE LOCATION OF THE FIELD PLOTS (ADS) IN RELATION WITH ITALY LANDSCAPE; B) DISPLAYS THE ADS IN RELATION WITH BOSCO PENNATARO; C) SHOWS ONE PICTURE OF THE TERRESTRIAL LASER SCANNING (TLS) DEVICE AND D) SHOWS A PICTURE OF BOSCO PENNATARO.-----72

FIGURE 16 GRAPHIC REPRESENTATION OF A TRUNK STRUCTURE AND MERCHANTABLE LOG FOR STRAIGHTNESS AND TAPERING ESTIMATION. THE “H” INDICATES THE HIGHEST CONVEX CURVE, THE MINIMUM AND MAXIMUM DIAMETER OF THE LOG ENDS (DMIN AND DMAX, RESPECTIVELY) WERE DISPLAYED. -----	75
FIGURE 17 METHODOLOGICAL APPROACH FOR TIMBER ASSORTMENT ESTIMATION USING TERRESTRIAL LASER SCANNING (TLS) DATA. GREY RECTANGLES INDICATE THE SUB- STEPS. -----	76
FIGURE 18 AN OVERVIEW OF THE SUB- STEPS FOR TIMBER-LEAVES DISCRIMINATION. THE INPUT TLS DATA (TOP LEFT), THE GEOMETRY-BASED CALCULATION (A), PREDICTOR VARIABLES SELECTION (B) AND TIMBER-LEAVES DISCRIMINATION (C) AND OUTPUT TIMBER POINT CLOUD (BOTTOM RIGHT). -----	77
FIGURE 19 AN EXPLICATIVE EXAMPLE ABOUT THE DISTRIBUTION OF CYLINDER ALONG THE TRUNK WITH AN INTERVAL OF 0.15 M (LEFT IMAGE) AND THE THREE CYLINDERS ALONG THE LOGS WITH AN INTERVAL OF 1.4 M (RIGHT IMAGE).-----	82
FIGURE 20 TREE POSITION AND SINGLE-SCAN LOCATION WITHIN EACH FIELD PLOT (ADS). THE BLACK SQUARES REPRESENT THE ADS BOUNDARIES, THE YELLOW SQUARES REPRESENT THE SINGLE-SCAN POSITIONS, THE COLOURED DOT POINTS REPRESENT THE OBSERVED TREE POSITION CLASSIFIED BY TREE SPECIES ON DIGITAL TERRAIN MODEL (DTM) BACKGROUND. -----	83
FIGURE 21 DETECTION RATE (DR) VALUES FOR INDIVIDUAL TREE SPECIES -----	86
FIGURE 22 PREDICTED VS. OBSERVED VALUES OF THE TWO LINEAR MODELS FOR THE DIAMETER AT BREAST HEIGHT (DBH). -----	87
FIGURE 23 RATE OF RECONSTRUCTED STEM DERIVED FROM TERRESTRIAL LASER SCANNING DATA (TRUERSTEM) WAS SHOWED FOR EACH TREE SPECIES -----	88
FIGURE 24 STEM CURVE PERFORMANCE. THE CURVE LENGTH RATIO (CLR; A) AND THE PERCENTAGE OF THE TREE HEIGHT COVERED (PHC; B) WERE DISPLAYED FOR ALL RECONSTRUCTED TREES. -----	89
FIGURE 25 PREDICTED VS OBSERVED VALUES OF THE LINEAR MODELS OBTAINED FOR MAXIMUM (DMAX) AND MINIMUM (DMIN) ENDS DIAMETERS. -----	90
FIGURE 26 PREDICTED VS OBSERVED VALUES OF THE LINEAR MODELS FOR THE VOLUME OF LOGS (TTV.LOG). -----	92
FIGURE 27 THE STRAIGHTNESS (STR) AND TAPERING (TAP) VALUES FOR MERCHANTABLE LOGS THE RED SQUARED INDICATES THE NEGATIVES VALUES OF TAP REPORTED FOR PREDICTED AND OBSERVED DATA. -----	93
FIGURE 28 CLASSIFICATION OF THE MERCHANTABLE LOGS INTO FIFTEEN ASSORTMENT TYPES BELONGING TO SAW-LOG PLUS, SAW-LOG, PULPWOOD, OTHER INDUSTRIAL ROUNDWOOD AND FUELWOOD ASSORTMENTS. -----	94

Extended abstract

Research framework

Forests are essential for life, besides providing economic, social and environmental benefits to humans, they ensure human wellbeing for decades through their essential ecosystem services. Such benefits, however, are often fostered by the implementation of Sustainable Forest Management (SFM) actions, maximizing roundwood production and at the same time ensuring the delivery of other important forest ecosystem services (FRA 2015). In the last decades, the demand for roundwood, namely pulpwood, saw-log, fuelwood and other roundwood categories, has increased, reaching 550 million m³ annually, representing a 40 % increment of total roundwood harvested (SoEF 2020).

Roundwood represents an important source of income for forest owners, and it is central to modern and sustainable forest management through its crucial role in long-term carbon storage. Despite the forest potency in climate change mitigation, the correct estimates of its standing timber assortments are poorly explored, particularly in natural forests. National forest inventory and national statistical survey are the most common tools supporting currently the assessment of forest resources and forest chains, including timber assortments. Nonetheless, forest inventory is the most accurate non-destructive method for assessing forest resources, even if it is time-consuming, requires well-trained operators but overall low efficiency in describing the upper part of canopy (West 2009; La Marca and Notarangelo, 2009). Moreover, the true information is strongly affected by hindering factors, as forest stand accessibility, terrain slope, tree species richness, stem straightness, and stem tapering (Pinto *et al.* 2004; Kankare *et al.* 2014). Therefore, the destructive method, including bucking and sawing processes of felled trees, was previously considered solely the most accurate method for assessing timber assortments (Holopainen *et al.* 2010; Jukka *et al.* 2010). But, an accurate and reliable method for assessing the timber assortments of standing trees is necessary to support and improve the quantification and description of timber assortments, forest productivity and for a more accurate carbon evaluation (FOREST EUROPE 2015).

Light Detection and Ranging (LiDAR) is a powerful active remote sensing method, enabling the characterization of the stem profiles and morphology using georeferenced points. Nowadays, the most representative LiDAR data used for forest monitoring are Airborne Laser Scanning (ALS), Terrestrial Laser Scanning (TLS) and spaceborne LiDAR data (Beland *et al.* 2019). Though LiDAR is well known for forest planning and implementing SFM actions, its use for assessing timber assortments is still limited and therefore more efforts are made to fill this gap (Vastaranta *et al.* 2014; Saarinen *et al.* 2019). The use of LiDAR data for timber assortment evaluation can offer many insights, such as, the valorisation of abandoned forests, accurate quantification of the carbon stock stored in trees and the choice of trees with a lesser ecological value. Therefore, a study aimed to highlight both pros and cons of using LiDAR for assessing timber assortments as well as studies focused to improve timber detection rate are very important to support SFM, to valorise forest resources, to promote forest productivity and to better quantify the carbon stored in the forest resources regardless the management aims. In this thesis, we first developed a literature review aimed to retrieve the most pertinent information of the recent approaches for timber assortment

evaluation using LiDAR data, then, we propose two stepwise approaches to extract the timber assortment information from standing trees using both ALS and TLS data:

The first study aims to provide an overview of the recent studies that used LiDAR data for timber assortment estimation between 2000 and 2018 using Elsevier's Scopus® search engine.

The second study aims to introduce the combined use of two unsupervised techniques for carbon stock assessment at the tree level in a mixed-species and multi-layered forest using ALS data.

The third study aims to introduce a stepwise approach for timber assortments estimation and classification using TLS data.

Material and Methods

Study 1: literature review

To reach the aim of the first study, we follow three steps: 1) paper collection, 2) paper clustering, and 3) paper analysis. Firstly, using 10 keywords organized in 12 queries, we collected many scientific papers and stored these in a database through Elsevier's Scopus® search engine, in a timeframe 2000-2018. Subsequently, all the papers were accurately scrutinized and classified into six thematic clusters based on their main and secondary aims. Finally, we compared the methods and outputs among all papers to highlight the temporal evolution of LiDAR applications for timber assortments estimation.

Study 2: tree detection and carbon quantification through ALS

To reach the aim of the second study, we developed and implemented a stepwise approach consisting of the following five steps: 1) pre-processing of the ALS data; 2) grouping and stratifying the point clouds of each field plot in four complexity levels; 3) tree detection and segmentation; 4) validation of the predicted tree crowns; and 5) prediction of forest inventory variables. The collection of ALS data was carried out in 2016 for 31 field plots (hereafter ADS) of 729 m². The study area is located in a mixed-species and multi-layered Mediterranean forest belonging to the Apennines Mountains, central Italy. First of all, the point cloud was clipped with the ADS boundaries. Subsequently, the field plots were categorized in four complexity levels (A, highly difficult; B, moderately easy; C, highly easy; D, moderately difficult) according to the three height variations and density of ADS point cloud. To better reflect the vertical stratification of tree crowns, each ADS point cloud was stratified into three layers (lower layer: small trees, intermediate layer: intermediate trees and upper layer: dominant and codominant trees). The pre-processing of each field plot point cloud was done on LAStools software. Seven accuracy parameters were used for validating the identification of trees, such as the detection rate (hereafter DR). The carbon stock prediction for each detected tree was assessed through Random Forests algorithm.

Study 3: Assessing timber assortments through TLS

A stepwise approach, consisting of the following four steps: 1) timber-leaves discrimination, 2) tree detection, 3) stem reconstruction, and 4) timber assortment estimation and classification was implemented. Theoretically, the first processing for obtaining the timber logs section from trunks was named bucking.

The bucking processing allows forest managers to divide the trunk timber section into several logs. The log is a desired timber section of the trunk axis with specific dimensions and length, and it can be classified into many types of assortments (i.e., saw-log). In our study, the trunk section was divided into merchantable ($2.5 \text{ m} \leq \text{length of log} \leq 3 \text{ m}$) and non-merchantable ($2.5 \text{ m} < \text{length of log}$) logs; the merchantable logs were classified into one out of 15 types of assortments. The collection of TLS data was carried out in 2018 and it covers five ADS of 729 m^2 in the same study area of the previous study. The five ADS include a total of 12 tree species and 178 trees, wherein 70 of them were considered large trees, meaning trees with a diameter at breast height (DBH) higher than 20 cm. Finally, for the reconstructed trees, the timber assortments were recognized and categorized into one out of the 15 types of merchantable logs.

Results

Study 1: literature review

A total of 179 papers were collected and stored in a database. The trend of the literature review highlights that there was a steady increment of publications from 2000 to 2016. The six clusters consisted in: Inventory (81 papers, 45.3 %), Productivity (42 papers, 23.5 %), Accuracy (24 papers, 13.41 %), Biodiversity (13 papers, 7.3 %), Climate Change (8 papers, 4.8 %) and Review (11 papers, 6.1 %). Most of them were carried out in North America (43.40 %), and in Europe (42.77 %). The number of papers was equal for mixed and pure forests, 85 and 83 respectively. Most papers (75 %) were based on the ALS data, while 15 % on spaceborne and 10 % on TLS, even if ALS and TLS resulted to be the most frequent devices used to assess timber assortments (i.e., inventory, and productivity). Since 2010, the integration between LiDAR and other remote sensing devices (satellite images) being increasingly used, particularly because that approach provides additional useful information (i.e., metrics) aimed to increase the performance of LiDAR data, especially when there is a low-quality point cloud (lower than 9 points m^{-2}) through the high quality of remote sensing images (2.4 m per pixel). Furthermore, the integration approach between ALS and TLS proved to be useful for timber assortment assessment at plot level. Since 2010, machine learning algorithms became crucial for predicting, upscaling and classifying the ALS data, supporting large-scale forest monitoring. The main challenge in using ALS data was the DR, especially for small trees in mixed-species and multi-layered forests. Studies using TLS data revealed that the skeletonization, voxel-based and cylinder-fitting can be effective methods for reconstructing the architecture of trees, even if a clear connection with timber assortment evaluation was missed.

Study 2: tree detection and carbon quantification through ALS

The DR was greater for the ADS included in the highly easy (65 %) and moderately difficult (49 %) categories. A lesser detection accuracy was found for the ADS of moderately easy (DR = 43 %) and highly difficult (36 %) categories. The overall detection accuracy was found to be more accurate for intermediate layer (DR = 54 %) in comparison with lower layer (DR = 42 %) and upper layer (DR = 49 %). Overall, 952 out of 2117 reference tree were detected. Significant values of coefficient of determination and root means squared error (RMSE) were showed for DBH (0.92; 4.03 cm), tree height (0.95; 1.33 m) and stem volume (0.82; 0.31 m^3). A small absolute bias for the carbon stock prediction was shown for all four categories,

ranging from -0.8 ('Highly easy'; relative bias = -0.46 %) to 0.1 tons' ha⁻¹ ('Moderately difficult'; relative bias = -0.88 %).

Study 3: Assessing timber assortments through TLS

Timber-leaves discrimination of TLS point clouds was highly accurate, reaching an average accuracy of 0.98. We were able to detect 151 out of 178 reference trees, accounting for 84.4 % (SD \pm = 4.7 %). An improved detection accuracy was observed for trees with a DBH higher to 30 cm due to all trees were correctly identified. As concern to stem reconstruction, we were capable to reconstruct 47 out of 70 large reference trees. The performance of the reconstruction proved to be accurate due to more than three-quarters of the trunk section was reconstructed for these 47 reconstructed trees, in detail 88.1% (SD = \pm 16.7 %). The stepwise approach allowed us to quantify 134 out of 179 reference merchantable logs, reaching an accuracy equal to 75 % (134 out of 179 reference merchantable logs). These 134 merchantable logs were classified in 11 out of 15 types of assortments. The classification of merchantable logs was more accurate for eight assortment types (i.e., some saw-log, pulpwood and other industrial roundwood), which was \pm 2 merchantable logs.

Final remarks

The results of this thesis revealed that the LiDAR data is a powerful source for assessing timber assortments from standing trees, regardless of the complexity of forest structure. This statement is supported by the power of LiDAR data to represent the vegetation structure rapidly, remotely and accurately. Over the years, the main hindering factors that conditioned the use of LiDAR for forest monitoring were gradually reduced, as for example, the availability of open-source LiDAR data for the whole Earth surface; the cost-effectiveness ratio is increasingly being suited; the availability of intuitive and free approaches for characterizing the trees (i.e. CloudCompare software, Computree, TreeLS), for tree detection (reFLex, MS-TSI algorithms) and for machine learning analysis (R packages) is even improved. Nonetheless, recent studies proved that handling LiDAR devices (i.e., portable laser scanning) can even be used for remote forest areas or areas where is not possible to flight for some reasons (e.g., ENAC restrictions).

The literature review revealed that the success of the use of ALS for timber assortment assessment was conditioned by the detection accuracy. More efforts aimed to increase the detection accuracy, especially for small trees (DBH \leq 20 cm), are still highly required, therefore, an alternative approach focused on detecting the trees through canopy layers can be useful to overcome this challenge.

We proposed, for the first time, an unsupervised tree detection approach for detecting trees in a mixed and multi-layered Mediterranean forest, as well as to use random forests for predicting the carbons stock of detected trees. It is worth noting that the investigated forest structure was complex and characterized by 13 different tree species. A stratification of the ALS point cloud was useful to simulate the vertical separation of strata. Enhanced detection accuracy was found in the ADS with significant structural heterogeneity, especially those covered with more than 30 points m⁻², reaching an average detection accuracy equal to 65 %. However, this detection accuracy decreases until 49 %, as the point density decreases. The accuracy

obtained in our study was slightly higher than the accuracy obtained from a similar study in the same forest stand condition in 2016 (Sačkov *et al.* 2016). Some strengths of our unsupervised approach reside in the capability for detecting trees without previous knowledge of stem position and in a fast way (about 65 sec per ADS). The obtained accuracy of carbon stock assessment for each detected tree was found to be comparable with the accuracy obtained for simple forest structure, as in pure forest stands of temperate forests (Popescu 2007). In the light of the above, our approach offered a considerable improvement for detecting the small trees, and it can be used for timber assortment assessment in forests with complex forest structures.

Unlike the ALS studies, to date, several approaches are available to automatically process the TLS data in order to characterize the trees (i.e., stem volume, trunk volume), specifically, the cylinder-fitting approach. Unfortunately, a study using the cylinder-fitting approach for timber assortment evaluation from standing trees is still required. In this regard, we propose, for the first time, a stepwise approach for timber assortment assessment through the cylinder-fitting approach. To reach this, accurate timber-leaf discrimination of points was achieved through Random forests algorithm. This discrimination proved to be useful to free the timber points from leaf points, and it allowed us to reconstruct the trunk using the cylinder-fitting algorithm. However, the challenge in reconstructing the trunk was given to the stem form, the presence of lianas, bulges and microhabitats. Nevertheless, the proportion of reconstructed reference trees was rather high (0.67) considering the heterogeneous structure and the richness of tree species. This accurate reconstruction was validated by the accurate prediction of trunk volume for detected trees (coefficient of determination = 0.91; RMSE = 0.03 m³). A good quantity of extracted logs from the trunk section was surveyed and classified into eleven different assortments. In summary, our stepwise approach allowed us to accurately quantify and classify the logs derived from standing trees using TLS data in mixed-species and multi-layered forests.

In conclusion, in this thesis, the literature review allowed us to identify the most important challenges and opportunities derived from LiDAR studies. As a result, we used some promising techniques for overcoming the challenges, as for example, the unsupervised algorithms for detecting the trees. The two stepwise approaches implementing for ALS and TLS data were tested for the first time for such targets, especially in mixed-species and multi-layered forests. Their application in other forests could provide better results than that obtained in our study, based on the heterogeneity of forest structure. As regards the ALS data, we discovered that the forest structural heterogeneity plays a slightly more important role than the point cloud in the occlusion of points in a tree detection step. In fact, in the down part of the canopy, where the point incidence is reduced, the detection accuracy decrease. As regards the TLS data, we discovered that optimal discrimination of timber-leaves points can facilitate the reconstruction of trees using the cylinder-fitting approach. However, the accuracy of that reconstruction was adjusted by the tree species and the stem defects (i.e., bulges, knots). Nevertheless, a study implementing our approach in other forest conditions can be useful to a deeper understanding of the potential of both stepwise approaches. The efficient use of timber resources in forests managed with conservative purposes could help to cater both the productivity and biodiversity, and at the same time it could allow managing the forest under SFM criteria. It is worth to note

that the timber resource in inner Italian forest areas is mainly used for energy purpose. Therefore, further efforts for valorising these forests can reduce the dependence from timber imported from extra EU countries and increasing the efficient use of our own timber resources. In addition, since the usability of this approach can result difficult to share with no expert, we thought that an intuitive, accurate and fast mobile app could be useful to overcome this challenge, so in future, it could be an interesting road to follows.

Final remarks

The results of this thesis revealed that the LiDAR data is a powerful source for assessing timber assortments from standing trees, regardless of the complexity of forest structure. This statement is supported by the power of LiDAR data to represent the vegetation structure rapidly, remotely and accurately. Over the years, the main hindering factors that conditioned the use of LiDAR for forest monitoring were gradually reduced, as for example, the availability of open-source LiDAR data for the whole Earth surface; the cost-effectiveness ratio is increasingly being suited; the availability of intuitive and free approaches for characterizing the trees (i.e. CloudCompare software, Computree, TreeLS), for tree detection (reFLex, MS-TSI algorithms) and for machine learning analysis (R packages) is even improved. Nonetheless, recent studies proved that handling LiDAR devices (i.e. portable laser scanning) can even be used for remote forest areas or areas where is not possible to flight for some reasons (e.g., ENAC restrictions).

The literature review revealed that the success of the use of ALS for timber assortment assessment was conditioned by the detection accuracy. More efforts aimed to increase the detection accuracy, especially for small trees ($DBH \leq 20$ cm), are still highly required, therefore, an alternative approach focused on detecting the trees through canopy layers can be useful to overcome this challenge.

We proposed, for the first time, an unsupervised tree detection approach for detecting trees in a mixed and multi-layered Mediterranean forest, as well as to use random forests for predicting the carbons stock of detected trees. It is worth noting that the investigated forest structure was complex and characterized by 13 different tree species. A stratification of the ALS point cloud was useful to simulate the vertical separation of strata. Enhanced detection accuracy was found in the ADS with significant structural heterogeneity, especially those covered with more than 30 points m^{-2} , reaching an average detection accuracy equal to 65 %. However, this detection accuracy decreases until 49 %, as the point density decreases. The accuracy obtained in our study was slightly higher than the accuracy obtained from a similar study in the same forest stand condition in 2016 (Sačkov *et al.* 2016). Some strengths of our unsupervised approach reside in the capability for detecting trees without previous knowledge of stem position and in a fast way (about 65 sec per ADS). The obtained accuracy of carbon stock assessment for each detected tree was found to be comparable with the accuracy obtained for simple forest structure, as in pure forest stands of temperate forests (Popescu 2007). In the light of the above, our approach offered a considerable improvement for detecting the small trees, and it can be used for timber assortment assessment in forests with complex forest structures.

Unlike the ALS studies, to date, several approaches are available to automatically process the TLS data in order to characterize the trees (i.e. stem volume, trunk volume), specifically, the cylinder-fitting approach.

Unfortunately, a study using the cylinder-fitting approach for timber assortment evaluation from standing trees is still required. In this regard, we propose, for the first time, a stepwise approach for timber assortment assessment through the cylinder-fitting approach. To reach this, accurate timber-leaf discrimination of points was achieved through Random forests algorithm. This discrimination proved to be useful to free the timber points from leaves points and it allowed us to reconstruct the trunk using the cylinder-fitting algorithm. However, the challenge in reconstructing the trunk was given to the stem form, the presence of lianas, bulges and microhabitats. Nevertheless, the proportion of reconstructed reference trees was rather high (0.67) considering the heterogeneous structure and the richness of tree species. This accurate reconstruction was validated by the accurate prediction of trunk volume for detected trees (coefficient of determination = 0.91; RMSE = 0.03 m³). A good quantity of extracted logs from the trunk section was surveyed and classified into eleven different assortments. In summary, our stepwise approach allowed us to accurately quantify and classify the logs derived from standing trees using TLS data in mixed-species and multi-layered forests.

In conclusion, in this thesis, the literature review allowed us to identify the most important challenges and opportunities derived from LiDAR studies. As a result, we used some promising techniques for overcoming the challenges, as for example, the unsupervised algorithms for detecting the trees. The two stepwise approaches implementing for ALS and TLS data were tested for the first time for such targets, especially in mixed-species and multi-layered forests. Their application in other forests could provide better results than that obtained in our study based on the heterogeneity of forest structure. As regards the ALS data, we discovered that the forest structural heterogeneity plays a slightly more important role than the point cloud in the occlusion of points in a tree detection step. In fact, in the down part of the canopy, where the point incidence is reduced, the detection accuracy decrease. As regards the TLS data, we discovered that optimal discrimination of timber-leaves points can facilitate the reconstruction of trees using the cylinder-fitting approach. However, the accuracy of that reconstruction was adjusted by the tree species and the stem defects (i.e. bulges, knots). Nevertheless, a study implementing our approach in other forest conditions can be useful to a deeper understanding of the potential of both stepwise approaches. The efficient use of timber resources in forests managed with conservative purposes could help to cater both the productivity and biodiversity and at the same time it could allow managing the forest under SFM criteria. It is worth to note that the timber resource in inner Italian forest areas is mainly used for energy purpose. Therefore, further efforts for valorising these forests can reduce the dependence from timber imported from extra EU countries and increasing the efficient use of our own timber resources. In addition, since the usability of this approach can result difficult to share with no expert, we thought that an intuitive, accurate and fast mobile app could be useful to overcome this challenge, so in future, it could be an interesting road to follow.

1. CHAPTER 1- Introduction

Forest can provide many goods and services to society, e.g., roundwood, no-timber forest products, genetic resources and biodiversity conservation and climate change mitigation (MEA 2005). Accurate roundwood information, from forests, became crucial to increase socio-economic and environmental benefits to forest owners and stakeholders; however, managing the forest towards sustainable forest management (SFM) became a crucial step (SoEF 2020). The European forests, accounting for 35 % of the total territory, provide about 550 million m³ annually of the marketed roundwood (SoEF 2020). In merchantable terms, the roundwood can be categorized in several timber assortments (i.e., saw-log, pulpwood, woodfuel) (SoEF 2020). Historically, the most accurate method for assessing timber assortment is the destructive one, because it considers the real log length and stem form (Jukka *et al.* 2010; Togni 2017). Moreover, such tendency was even supported by the challenges derived from the implementation of non-destructive methods, as forest inventory, which proved to be less accurate in the upper part of the canopy (West 2009). However, implementing the non-destructive methods might mean the timber assortment valorisation and the scheduling of SFM activities (Santopuoli *et al.* 2016). Therefore, an accurate and reliable non-destructive method could be crucial for timber assortment assessment of standing trees.

Since 2000, a representative non-destructive method, named LiDAR (Light Detection and Ranging), became a powerful tool for characterizing the structure of forests and trees through georeferenced points (Næsset 2002). The power of LiDAR was widely proved to accurately predict forest inventory variables, and therefore it was used for supporting local and national forest inventory (McRoberts and Tomppo, 2007; Næsset *et al.* 2011; Chirici *et al.* 2016) as well as, to schedule SFM activities aimed at promoting the biodiversity conservation (Barbati *et al.* 2014; Chirici *et al.* 2020) and the delivery of other important forest ecosystem services.

LiDAR's data were categorized into three types of sources: terrestrial, airborne and spaceborne LiDAR sources (Beland *et al.* 2019). The most implemented LiDAR devices were airborne laser scanning (ALS) and Terrestrial laser scanning (TLS) and spaceborne LiDAR data (Beland *et al.* 2019). Whereas the spaceborne LiDAR data is composed primarily by Earth satellite information, such as, GLAS (Geoscience Laser Altimeter System; <https://icesat.gsfc.nasa.gov/>), ATLAS (Advanced Topographic Laser Altimeter System; <https://icesat-2.gsfc.nasa.gov/>), GEDI (Global Ecosystem Dynamics Investigation; <https://gedi.umd.edu/>) and MOLI (Multi-footprint Observation Lidar and Imager; <https://www.wmo-sat.info/>)(Beland *et al.* 2019).

Some studies using ALS data demonstrated that many hectares of forests can be monitored by each conducted flight strip, and it can be at an affordable price and with a suit cost-effectiveness ratio (Montaghi *et al.* 2013; Kelly and DiTommaso, 2015). Some studies using TLS data proved that even if a small part of forests can be monitored by each fixed device mounted on the tripod, the description of the stem form of trees is assured (Liang *et al.* 2018). Studies combining the ALS with TLS proved to be more accurate than studies using solely the ALS data, however, the TLS data plays a crucial role in modelling the tree structure (Liu *et al.* 2017; Giannetti *et al.* 2018) Studies using spaceborne LiDAR data suggested that the tree structure and stem volume can be accurately described at worldwide level (Beland *et al.* 2019).

In the light of the above, the ALS and TLS are presented as effective data for timber assortment assessment (Kankare *et al.* 2014; Liang *et al.* 2018; Wan *et al.* 2019). For ALS studies, despite the capability for monitoring forest where is not possible to flight for some reasons is promising, the performance of ALS models proved to be less accurate in mixed than pure stands, especially in multi-layered forest (Wang *et al.* 2016, 2017). For TLS studies, despite many semi-automatic and automatic algorithms are available for stem form reconstruction, the success is strongly conditioned by many aspects, such as, operational (i.e., georeferencing, licenses), technical (i.e., automatic algorithms, huge quantity of data to manage), weather conditions, scan mode and forest stand structure (Wan *et al.* 2019).

The findings obtained in ALS and TLS allowed improving the management of forests based on SFM aims, as maintaining the forest resource and carbon cycles (Chen 2015), ensuring the forest health and vitality (Junttila 2014) and conserving forest biological diversity (Kelly and DiTommaso, 2015). Few studies using ALS data to determine the health status of forests were even reported, such as, fire, insect, and disease hazard issues (Smigaj *et al.* 2019). Studies using TLS to investigate ecological aspects were even displayed, e.g., tree species composition (Othmani *et al.* 2016), land cover classification (Walicka *et al.* 2019), the habitat quality of bird species (Michel *et al.* 2008), and tree-related microhabitats (Santopuoli *et al.* 2020; Rehush *et al.* 2019).

Despite the challenges and opportunities offering the LiDAR data to the forest science community, to date, a reliable and feasible approach using ALS data for timber assortment assessment was still necessary, especially at single tree level and for mixed-species and multi-layered forest (Silva *et al.* 2017). It is worth noting that, the use of ALS data can be affected by tree detection challenges, especially in multi-layered forests (Holopainen *et al.* 2010). Nevertheless, the use of TLS data, including many robust algorithms, i.e., TreeQSM (Raumonen *et al.* 2013), Simple-Tree (Hackenberg *et al.* 2015), requires further efforts for extracting the qualitative and quantitative description of timber assortment of standing trees (Saarinen *et al.* 2019; Chianucci *et al.* 2020).

Even if concrete advances to overcome the tree detection challenges were provided by many algorithms (i.e., reFLex algorithm, Multisource Single-Tree Inventory “MS-STI”, robust processing), a well-quality of the point cloud (higher than 30 points m⁻²), followed by a stratification approach were mandatory prerequisites (Sačkov *et al.* 2016; Hamraz *et al.* 2017). A robust algorithm using TLS data for reconstructing the stem form, named the cylinder-fitting approach (Pfeifer *et al.* 2004; Liang *et al.* 2018), was capable to reconstruct the trunk section of trees.

In the light of the foregoing, a study aimed at summarizing the most pertinent and available approaches for timber assortment assessment is still needed, and two stepwise approaches using ALS and TLS data for extracting the timber assortment information from standing trees were developed and described in this thesis. This study can be useful to optimize the use of timber sources from forests. In this thesis, three studies were described:

1. To gain insights from the recent studies using LiDAR data for timber assortment assessment, a literature review was proposed, namely “Challenges and opportunities for timber assortment evaluation through LiDAR. A review.”;

2. To overcome the principal hindering factor affecting the detection of trees into ALS studies, making the timber assortment estimation, an unsupervised method was proposed for detecting and segmenting the trees in mixed-species and multi-layered Mediterranean forests namely “Unsupervised algorithms to detect single trees in a mixed-species and multi-layered Mediterranean forest using LiDAR data”;
3. To extract the timber assortment from standing trees using TLS data in mixed-species and multi-layered Mediterranean forest, a stepwise procedure was proposed, namely “A stepwise approach for deriving timber assortment of trees from Terrestrial Laser Scanning data”.

2. CHAPTER 2 – Paper review and scientific papers

2.1. Challenges and opportunities for timber assortment evaluation through LiDAR. A review.

2.1.1. Preface

In the forest, quantifying and classifying accurately the timber assortments from standing trees results in timber valorisation and allocation. Light Detection and ranging (LiDAR) is a promising source suitable to reconstruct the stem form using point cloud. It can be categorized into three different sources: terrestrial, airborne and spaceborne LiDAR sources (Beland *et al.* 2019). The most implemented approaches for analysing LiDAR data are the area-based approach (ABA) and individual tree detection (ITD). The ABA was used for predicting several forest inventory variables and vegetation indices at plot level (Næsset *et al.* 2011; McRoberts *et al.* 2015). ITD was used for gathering stem form information from standing trees (Sackov *et al.* 2019; Wang *et al.* 2016). Nowadays, a better representation of the stem form was procured following ITD with respect to ABA. However, since there is a lot of information about ABA implementation, the processing used in ABA studies can even be useful for timber assortment assessment. This means that both approaches can provide valuable information to quantify and classify the timber assortment from standing trees using LiDAR data (White *et al.* 2014; Silva *et al.* 2017). Therefore, the objective of this study is to explore the challenges and opportunities derived from the implementation of LiDAR data for timber assortment assessment in the period 2000-2018.

2.1.2. Abstract

Timber assortment estimation became crucial information for retrieving economic and social benefits from forest stands. Nowadays, for timber assortment estimation, the most implemented non-destructive method is that traditional, while the most accurate method is that destructive. Light Detection and Ranging (LiDAR) is a powerful technology for depicting the tree profiles using georeferenced 3D point clouds. LiDAR gained more attention from forest researchers due to its capability to provide accurate tree structure representation, as well as the versatility to cope with many forest issues. Nevertheless, studies using LiDAR data for timber assortment assessment are still necessary, especially at single tree level. In this context, an overview of the methods available for implementing LiDAR data on timber assortment estimation can be useful firstly to implement LiDAR data for timber assortment assessment and secondly to promote the efficient extraction of the timber products, making it more sustainable the management of forests.

This paper aims to provide an overview of the recent studies that used LiDAR data for timber assortment estimation between 2000 and 2018 using Elsevier's Scopus® search engine. To reach this objective, the study follows three steps: 1) paper collection, 2) paper classification, and 3) paper analysis. A total of 179 papers were collected and grouped in six thematic forest topics: Inventory (45.25 %), Productivity (23.46 %), Accuracy (13.41 %), Biodiversity (7.26 %), Climate Change (4.47 %) and Review (6.15 %). Nearby, half of the papers were focused on gathering forest inventory information of forest stand. Almost one-quarter of papers have investigated forest productivity issues. Airborne and terrestrial LiDAR devices were

used for monitoring forest productivity issues. Though the airborne LiDAR devices reached to cover even large forest areas, detecting of trees remains a challenging step. Though the terrestrial LiDAR devices were capable to reconstruct the stems, the success of the collection of terrestrial LiDAR data depends on the operability, technical aspects and weather conditions. The integration of ALS with TLS proved to be an effective method for timber assortment prediction at plot level, however, among them, TLS data showed higher accuracy. The implementation of machine learning algorithms became effective for predicting, upscaling, modelling and classifying the LiDAR data in forest studies. However, studies using machine learning algorithms for timber assortment assessment were still limited, especially at single tree level. Further insights concerning the analysis of the combined use of LiDAR devices through machine learning algorithms can be useful information to promote accurate and faithful timber assortment estimation for large areas.

Keywords: Timber quality, timber provision, ALS, TLS, 3D modelling, saw-log, pulpwood.

2.1.3. Introduction

2.1.3.1. *Background:*

In recent decades, remote sensing techniques, especially the Light Detection and Ranging (LiDAR), became crucial and increasingly applied for monitoring forest ecosystems (Wulder *et al.*, 2008), and assessing many aspects of sustainable forest management (SFM). The huge versatility of the LiDAR data, due to the high accuracy and the amount of data provided, will further increase its implementation in forest management and planning.

To date, LiDAR data can be obtained by three types of sources, such as satellite, airborne and terrestrial devices (Van Leeuwen *et al.* 2011; Montaghi *et al.* 2013). The satellite LiDAR data fosters the monitoring of the large forest areas periodically, with a continuous global observation of the Earth, allowing the temporal analysis and monitoring of changes over time (<https://icesat.gsfc.nasa.gov/icesat/glas.php>). Airborne Laser Scanning (ALS) represents the most suitable LiDAR device at landscape and local scales used for forest inventory and research purposes as a decision support tool for forest planning and management due to it provides high-resolution point clouds (Næsset 2002; McRoberts *et al.* 2015; Chirici *et al.* 2020). Conversely, Terrestrial Laser Scanner (TLS) is a powerful device to provide very high-resolution point clouds at stand level, ensuring an accurate assessment of tree structure, even if several scans are necessary to cover a small forest area, resulting very expensive and time-consuming (Dassot *et al.* 2011; Liang *et al.* 2019; Saarinen *et al.* 2017).

In addition to LiDAR platform sources, the modelling of forest measurements can be derived from two main approaches, namely the area-based approach (ABA) and the individual-tree detection (ITD) approach (Montaghi *et al.* 2013). The ABA is the most common approach used for forest inventory, allowing the assessment of the forest inventory variables (FIVs), as the estimation of above-ground biomass, tree height, and the carbon stock (Næsset *et al.* 2011; N. d'Oliveira *et al.* 2012; White *et al.* 2014). Usually, LiDAR metrics (summary statistics of point clouds about the height measurements) are compared with reference data, to generate most of the common FIVs and forest structural indices (Vis) through regression models

(Næsset 2002; McRoberts *et al.* 2015) (Mura *et al.* 2016). Unlike the ABA, the ITD approach includes the detection and segmentation of single trees, providing accurate reproduction of tree characteristics, as dimension and defects, at the tree level (i.e. tree diameter, basal area, stem volume, tree structure). However, ITD is less implemented, particularly for irregular forests as mixed forests, due to the high tree density for which a very high-resolution point cloud is required (Karttinen *et al.* 2012; Wang *et al.* 2016). Though the implementation of both approaches resulted quite common for forest inventory assessment, depending on the availability of costs and well-trained personnel, some recent studies highlight that LiDAR can be very helpful to investigate particular forest management aspects, as for example forest biodiversity (Santopuoli *et al.* 2020; Rehush *et al.* 2018). Moreover, several authors stated that the integration of data from ABA and ITD allows obtaining better accuracy in forest inventory assessment (Tompalski *et al.* 2015; Shinzato *et al.* 2017).

Despite the use of LiDAR in forest monitoring and planning is increased over the years, with highly successful implementations, information about the hindering factors in using LiDAR technology is barely reported. This aspect is particularly important for supporting the assessment of timber assortments and fostering the active and sustainable management of forest ecosystems. For this reason, an overview concerning the LiDAR sources and methodological approaches should provide crucial information to take advantage in forest monitoring and planning. This is even more evident considering the rapidness of climate change scenario and loss of biodiversity as emerged by both European Union Forest and Biodiversity Strategies. This study aims to highlight the state of the art about the recent advantages and challenges concerning the use of LiDAR for assessing FIVs and timber assortments through a literature review between 2000 and 2018. To reach this goal, scientific papers were collected using 10 keywords concerning timber assortment in the Elsevier's Scopus® search engine. After a brief description of the most common applications of LiDAR in forestry, the paper describes the main approaches and LiDAR devices used over the years. The implementation of the literature review is explained in section 2, while results, discussion and conclusion are described in sections 3, 4 and 5 respectively.

2.1.3.2. *State-of-the-Art:*

Over the years, the contribution of LiDAR being increased consistently on SFM implementation, covering most of the criteria and indicators for SFM, such as forest cover, forest health and vitality, timber production, conservation of biodiversity, protective function (Kelly and Di Tommaso, 2015). The interest of such contributes resides in the capability of LiDAR data to provide accurate estimates of forest variables with respect to those collected through surveyed field data (Næsset 2002; McRoberts *et al.* 2015).

Beyond the most common FIVs, and the assessment of forest structure through the evaluation of the spatial distribution of trees and canopy, both vertically and horizontally, LiDAR was used to assess the habitat quality of forests (Clawges *et al.* 2008; Riedler *et al.* 2015) or also the forest biodiversity, through the evaluation of the leaf characteristics associated with the movement of Neotropical migrant songbird population (Goetz *et al.* 2010). Some authors demonstrated how point clouds were used for assessing ecological aspects, as tree species composition (Hollaus *et al.* 2009; Zhao *et al.* 2018), land cover

classification (Sasaki *et al.* 2012) and tree-related microhabitats (Santopuoli *et al.* 2020; Rehush *et al.* 2018). Few studies aimed to assess forest disturbances and diseases as damages caused by insects, pests and fire (Junttila 2014; Smigaj *et al.* 2019). Moreover, some authors used ALS data (Vastaranta *et al.* 2014; Xu *et al.* 2016) for assessing the merchantable timber retrieved from forests, mainly in pure conifer forest stands, in Scandinavia (White *et al.* 2014), and Brazil (Silva *et al.* 2017).

Considering the high versatility of LiDAR applications, and the increased operational implementation of LiDAR devices in forest management and planning, a well detailed and comprehensive review showing details about the LiDAR implementation, the pre-processing and processing methods, for predicting timber assortment could represent an important step to promote knowledge and to foster its implementation for assessing timber assortments estimates to support forest management and to schedule the forest activities in light of the recent threats/crisis concerned the climate change and biodiversity loss (Silva *et al.* 2017).

Considering the high versatility of LiDAR applications, and the increased operational implementation of LiDAR devices in forest management and planning, a well detailed and comprehensive review showing details about the LiDAR implementation, the pre-processing and processing methods, for predicting timber assortment could represent an important step to promote knowledge and to foster its implementation for assessing timber assortments estimates to support forest management and to schedule the forest activities in light of the recent threats/crisis concerned the climate change and biodiversity loss.

2.1.4. Material and Methods

The literature review was carried out through three steps: paper collection, paper classification, and paper analysis (Figure 1).

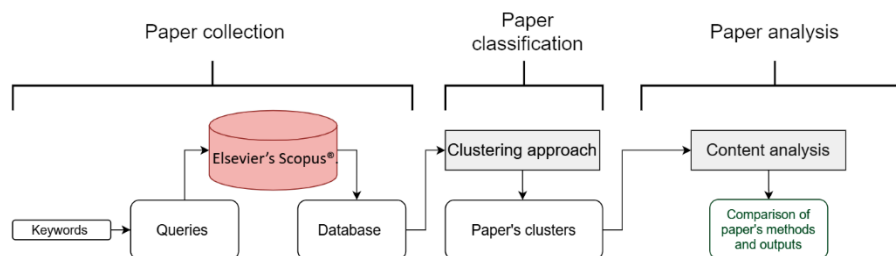


FIGURE 1 WORKFLOW OF THE PAPER COLLECTION, CLASSIFICATION AND ANALYSIS

2.1.4.1. Paper collection

The literature review was carried out in November 2018 using Elsevier’s Scopus® database. Ten keywords (Table 1) were used as input data to create 12 different queries, within which “Remote Sensing”, “LiDAR” and “Forest* or woodland” represented the fixed words for the codes (Figure 1; Table 2). Moreover, to give more emphasis on the timber assortments we use “timber”, “wood”, “branch” and “stem” as research words. Research papers were searched in three target sources “article title”, “abstract” and “keywords” in the period (2000-2018; July ends). The setting of the time frame is supported by the fact that the firsts LiDAR studies for mapping forest covers were done in the early 2000s (Wehr and Lohr, 1999). We did not use any restrictions, but non-English papers were excluded from the analysis.

TABLE 1 DESCRIPTION OF KEYWORDS USED TO RETRIEVE THE INFORMATION ABOUT LIDAR STUDIES. ADDITIONAL DETAILS ABOUT THE TERMS AND DEFINITIONS (A) ARE AVAILABLE AT [HTTP://WWW.FAO.ORG/FORESTRY/FRA2015/](http://www.fao.org/forestry/fra2015/)

Description of the keywords			
N ^o .	Keyword	Description	study
1	<i>Remote sensing (RS)</i>	It is the most important tool used for mapping forest covers through the reflected energy from the Earth (i.e. sunlight).	(McRoberts and Tomppo, 2007) https://earthdata.nasa.gov/learn/backgrounders/remotesensing
2	<i>LiDAR</i>	It is an active RS device, suitable to rapidly and directly depict the trees using 3D georeferenced points over different forest stands for large and small areas.	(Næsset 2002, 2011)
3	<i>Forest or Woodland</i>	'Forest' is the land covered by more than 0.5 ha with trees able to reach a minimum of 5 m of height, which possess a canopy cover of more than 10 %; while the "woodland" is a land covered by more than 0.5 ha with trees able to reach 5 m of height at maturity, which possess a canopy cover of 5 - 10 %; or land covered by a combined cover of shrubs, bushes, and trees above 10 %.	(FOREST EUROPE 2015) ^(a)
4	<i>Timber or Wood</i>	'Timber' and 'wood' are some of the most important goods provided by forests, and they play a role important in the wood supply chain.	(FOREST EUROPE and F.A.O. 2015)
5	<i>Stem or Branch</i>	'Stem' is the above-ground trunk of a vascular plant with similar anatomical properties, while 'branch' is the woody part of the tree that arising from a trunk.	(Hauglin <i>et al.</i> 2013; Saarinen <i>et al.</i> 2017)
6	<i>Hardwood or softwood</i>	'Hardwood' is commonly associated with deciduous stand (denser wood), while "softwood" is often associated with coniferous (less dense wood).	(Lim <i>et al.</i> 2003)
7	<i>Tree</i>	It indicates a tall plant composed of trunk and branches. Moreover, it is a principal component of both forest and woodland areas.	^(a)
8	<i>Quality or assortment</i>	'Quality' groups physical and chemical characteristics used for classifying wood based on specific wood features.; while the 'assortment' term is widely used to characterize the log of trees according to a merchantable approach.	^(a) (Jukka <i>et al.</i> 2010)
9	<i>Morphology</i>	This represents the physical form and external structure of trees. This word allowed us to collect papers having considered the morphology of the tree as objective.	(Antonarakis <i>et al.</i> 2009)
10	<i>Volume or merchant*</i>	These words allowed us to collect papers considering the wood in forest productivity and commercial terms as the target.	(Jozsa and Middleton, 1994)

TABLE 2 LITERATURE REVIEW STRINGS. ADVANCED DESCRIPTION OF CODES (SC) USED TO RETRIEVE THE INFORMATION ABOUT LIDAR STUDIES. THE FIXED WORDS ARE SHOWN IN ITALIC.

Abbreviation	Keywords used
SC1	<i>remote AND sensing*</i> ; <i>lidar</i> ; <i>forest*</i> OR <i>woodland</i> ; timber OR wood AND quality
SC2	<i>remote AND sensing*</i> ; <i>lidar</i> ; <i>forest*</i> OR <i>woodland</i> ; timber OR wood AND assortment*
SC3	<i>remote AND sensing*</i> ; <i>lidar</i> ; <i>forest*</i> OR <i>woodland</i> ; timber OR wood AND morphology
SC4	<i>remote AND sensing*</i> ; <i>lidar</i> ; <i>forest*</i> OR <i>woodland</i> ; timber OR wood AND volume
SC5	<i>remote AND sensing*</i> ; <i>lidar</i> ; <i>forest*</i> OR <i>woodland</i> ; stem OR branch AND volume
SC6	<i>remote AND sensing*</i> ; <i>lidar</i> ; <i>forest*</i> OR <i>woodland</i> ; stem OR branch AND morphology
SC7	<i>remote AND sensing*</i> ; <i>lidar</i> ; <i>forest*</i> OR <i>woodland</i> ; hardwood OR softwood AND merchant*
SC8	<i>remote AND sensing*</i> ; <i>lidar</i> ; <i>forest*</i> OR <i>woodland</i> ; tree AND morphology
SC9	<i>remote AND sensing*</i> ; <i>lidar</i> ; <i>forest*</i> OR <i>woodland</i> ; tree AND merchant*
SC10	<i>remote AND sensing*</i> ; <i>lidar</i> ; <i>forest*</i> OR <i>woodland</i> ; tree AND assortment*
SC11	<i>remote AND sensing*</i> ; <i>lidar</i> ; <i>forest*</i> OR <i>woodland</i> ; tree AND quality
SC12	<i>remote AND sensing*</i> ; <i>lidar</i> ; <i>forest*</i> OR <i>woodland</i> ; tree AND volume

2.1.4.2. *Papers classification*

After the collection phase, papers were accurately scrutinized and then classified in six clusters: Inventory (I); Productivity (P); Accuracy (A); Biodiversity (B); Climate Change (C); review (R), according to the main and secondary aims of articles.

1. **The cluster Inventory (I)** includes papers focused on monitoring forest area, over distinct forest-stand conditions (e.g. structure and forest management system). Usually, papers included in this cluster were based on ALS devices, in a singular way or in a combination with other RS data (Li *et al.* 2014), applying the ABA.
2. **The cluster Productivity (P)** includes papers centred on assessing the FIVs (i.e. tree height) and productivity issues (e.g. stem volume, above-ground biomass, carbon stock, saw-log volume, pulpwood volume), over distinct forest-stand conditions applying ITD approach.
3. **The cluster Accuracy (A)** includes all the papers focused on the description of the algorithms used for point clouds elaboration and on the comparison of the result accuracy obtained.
4. **The cluster Biodiversity (B)** includes the papers aimed to monitor the richness of tree species and habitat quality using LiDAR.
5. **The cluster Climate change (C)** includes papers focused to assess the relationship between forest structure and climate change responses and forest health, for instance, the combination of RS and LiDAR data was used for determining the occurrence of timely fire events in forest covers (Wulder *et al.* 2009).
6. **The cluster Review (R)** includes all review manuscripts found in the database. Papers included in this cluster were not used for the analysis of this paper because they had not a specific topic.

2.1.4.3. *Comparison of paper's methods and outputs*

Based on the study area reported in the papers a map indicating the geographic distribution of LiDAR studies was provided. Furthermore, for each cluster, the papers were compared among them to gain insights about the improvements of the methods and the findings carried out overtime. Narrative papers included in "R" cluster were not considered. The comparison was based on the forest-stand characteristics (i.e. coniferous, deciduous, plantations, mixed-forest) and the technical characteristics of LiDAR devices and platforms used, such as terrestrial, airborne and spaceborne platforms.

To a deeper timber assortment evaluation, the clusters "I" and "P" were further explored and analysed, highlighting the improvements obtained overtime (2000-2018), giving particular attention to the paper's aims, methods and analysis implemented for modelling the forest inventory and productivity variables. A chronological description of papers included in the cluster "I" allowed us to gain insights about the LiDAR developments focused on timber assortment estimation at plot level.

The cluster "P" was further explored and analysed, highlighting (i) pre-processing methods (ii), tree detection approaches (iii) and modelling the timber provision variables. A chronological description of the

papers of cluster “P” allowed us to gain insights about the LiDAR developments focused on timber assortment estimation at tree level. Therefore, the main gaps in timber assortment assessment, as highlighted in many papers, were considered to highlight the challenges, weaknesses and strengths to face up for further improving the use of LiDAR devices and data in forest management and planning.

2.1.5. Results

2.1.5.1. Literature review

A database composed of 301 papers was created using 12 different combinations of codes through Elsevier’s Scopus® engine. However, the final database included 179 papers, because 122 out of 301 were repeated in more than one query. The literature review highlights that since 2000 the number of publications per year increased until 2016. Thereafter, the trend slightly decreased (Figure 2). Considering that the literature collected in 2018 ends in July, we can observe that since 2014 the scientific community has published about 20 papers per year on the topic of assessing forest variables through LiDAR.

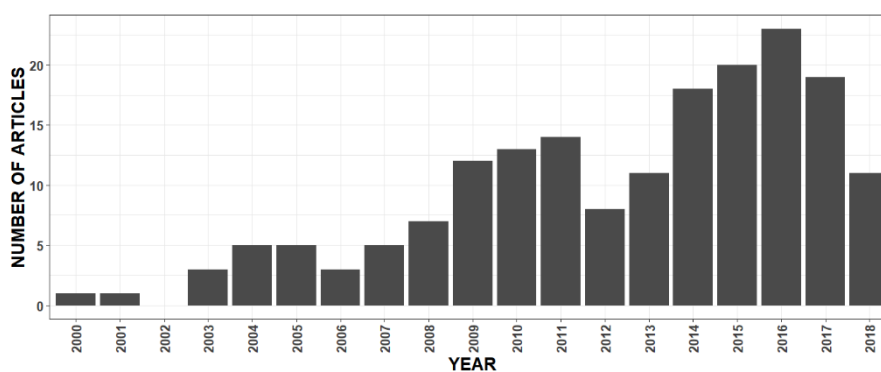


FIGURE 2 TREND OF THE ARTICLES PUBLISHED IN THE TIMEFRAME 2000-2018 (END JULY 2018).

Considering the number of retrieved papers, SC2 and SC9 were the codes that collected the lowest number of papers (Table 3). By contrast, SC12 was the code that allowed to retrieve the highest number of papers (100). More precisely, looking at the first four codes (SC1, SC2, SC3, and SC4), the keywords “volume” and “quality”, SC4 and SC1 respectively, allowed retrieving more papers than keywords “assortment” (SC2) and “morphology” SC3.

Similarly, looking at the last five codes, the keywords “merchant” (SC9) and “assortment” (SC10) were the keywords that retrieved the lowest number of papers, compared to keywords “volume” (SC12), and “quality” (SC11). Results show that the topics of timber assortment are poorly studied, while a more general term as forest volume is frequently explored.

TABLE 3 NUMBER OF PAPERS COLLECTED BY DIFFERENT CODES. EXCLUDING THE REPEATED PAPERS, THE FINAL NUMBER OF PAPERS USED IN THE ANALYSIS WAS REPORTED FOR EACH RESEARCH CODE.

Results of the paper collection		
Abbreviation	N° collected papers	N° faithful papers
SC1	13	13
SC2	2	1

SC3	3	1
SC4	41	35
SC5	54	40
SC6	10	1
SC7	6	1
SC8	20	16
SC9	2	1
SC10	3	1
SC11	47	33
SC12	100	36
TOTAL	301	179

From a geographical point of view, most of the LiDAR applications were carried out in North America (43.40 %), and in Europe (42.77 %), while few applications were observed in Oceania (5.03 %), Asia (6.29 %) and Africa (2.52 %) (Figure 3).

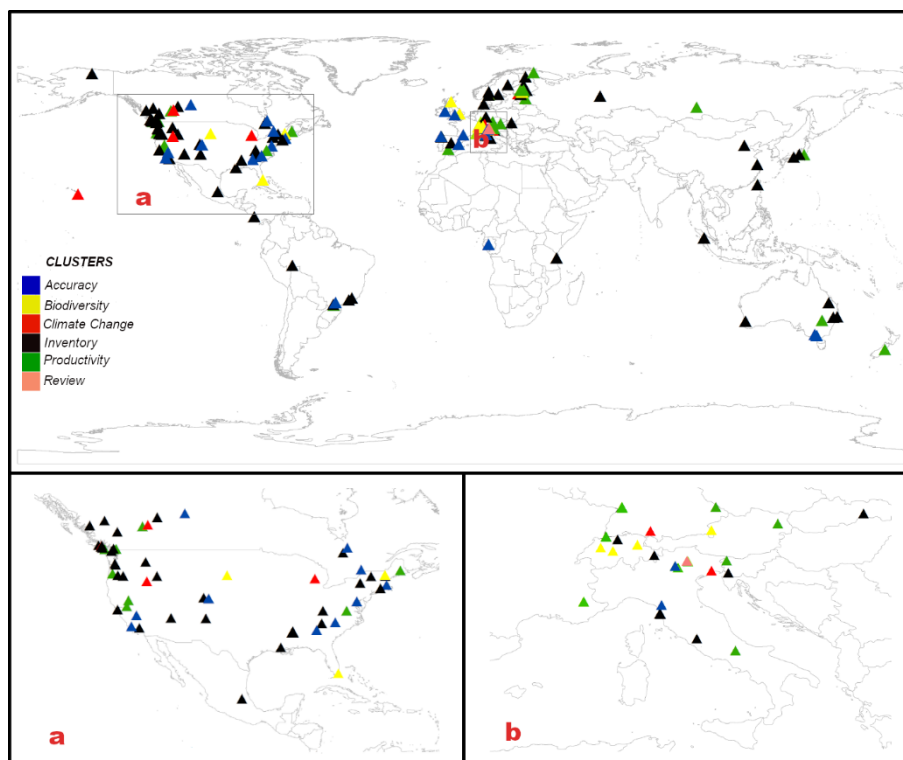


FIGURE 3 GEOGRAPHICAL REPRESENTATION OF LIDAR APPLICATIONS. THE BOX IN THE TOP SHOWS A GLOBAL OVERVIEW, WHILE THE BOX A SHOWS THE STUDIES CARRIED OUT IN NORTH AMERICA, WHILE BOX B SHOWS THE STUDIES CARRIED OUT IN ITALY.

2.1.5.2. Paper clusterization

The literature review highlights that forest inventory and forest productivity represent more than half of total papers published with 45.25 % and 23.46 % respectively (Figure 3). Few studies (13.41 %) deal with the improvement of LiDAR performances through statistic fundamentals, belonging to the accuracy cluster, and even fewer studies were reported for biodiversity (7.26 %) and climate change (4.47 %) clusters. As expected, literature reviews focused on LiDAR applications were rather limited (6.15 %).

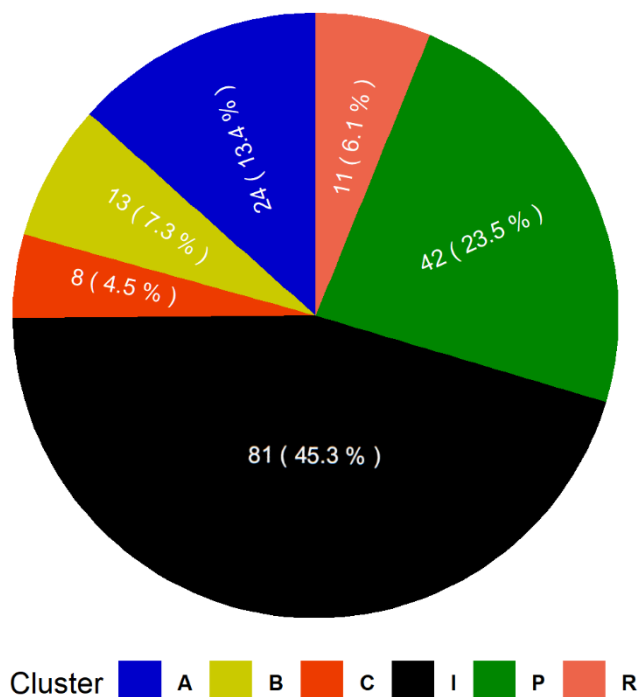


FIGURE 4 PAPER FREQUENCY BELONGING THE SIX CLUSTERS: ACCURACY (A), BIODIVERSITY (B), CLIMATE CHANGE (C), INVENTORY (I), PRODUCTIVITY (P), REVIEW (R).

Frequency of LiDAR application and forest stands.

As observed in this LiDAR review, the frequency of papers focused on mixed-species forest stands (85 out of 168) was similar to the number of papers carried out in pure forest stands (83 out 168) (Table 4). Among the pure forests, most LiDAR applications were carried out in conifers stands (62.65 %, 52 out of 83 studies), rather than in deciduous (19.27 %, 16 out of 83 studies) and forest plantations (18.07 %, 15 out of 83 studies).

Studies focused on forest productivity issues were more implemented in pure forest stands rather than in mixed forest stands (26 and 16 studies respectively), while inventory was slightly more implemented in mixed forest stands than pure forest stands (38 and 43 studies respectively). Biodiversity (4 vs. 9 studies, pure and mixed stands, respectively) and climate change (3 vs. 5 studies, pure and mixed stands, respectively) (Table 4). No difference was observed between mixed and pure stands for cluster “A” (50 % vs. 50%).

TABLE 4 FOREST-STAND CONDITION INFORMATION FOR EACH CLUSTER ANALYSED. REVIEW “R” PAPERS (11) WERE NOT CONSIDERED HERE

Cluster Abbreviation	Forest-stand condition				N° papers for mixed-forest	TOTAL
	N° papers for pure stands			Sub-total		
	Deciduous	Coniferous	Plantations			
I	6	28	4	38	43	81

P	7	12	7	26	16	42
A	2	6	4	12	12	24
B	1	3	0	4	9	13
C	0	3	0	3	5	8
R	-	-	-	-	-	-
TOTAL	16	52	15	83	85	168

LiDAR tools and RS devices

The results revealed that the first differences between airborne and spaceborne LiDAR devices were based on the extend of surveyed forest area and the amount of the FIVs and Vis investigated by authors. In particular, comparing airborne vs. spaceborne LiDAR studies, many FIVs (i.e. above ground biomass) and Vis were often monitored for varied hectares by each flight vs. solely two FIVs (i.e. tree height and tree volume) were often monitored for consecutive footprint laser points of 70m. As regards the difference between airborne and terrestrial LiDAR devices, it was based on the extend of surveyed forest areas, the distance measurement accuracy and tree structure characterization. More precisely, comparing airborne vs. terrestrial LiDAR devices, vary FIVs and Vis could be monitored for large forest areas by each flight at an accuracy higher than 4 points/m² (Wulder *et al.* 2012) vs. vary FIVs and the reconstruction of stem and branches could be monitored for small areas at an accuracy higher than ± 2 mm at a distance of 25 m (Dassot *et al.* 2011).

The results highlight that LiDAR applications were mostly implemented through airborne platforms (75 %), followed by spaceborne (15 %) and terrestrial (10 %) platforms. ALS devices proved to be the most important airborne LiDAR devices (84.66 %) against the other airborne devices, due to their capability to cover large forest areas and to cope with multiple objectives, as assessing SFM indicators, supporting forest management and planning, forest accessibility, contributing to forest monitoring and surveillance, with a yearly application equal to 13.7 papers in the timeframe 2000 - 2018 (Figure 5).

As regards the terrestrial platform, results highlighted that most LiDAR studies (68.2 %) were carried out through a fixed platform using TLS alone or in combination with a hyperspectral camera. However, the use of the mobile LiDAR platform (i.e. Portable Laser Scanning “PLS” and Mobile Laser Scanning “MLS”) was also rather common (31.8 %), particularly to assess the volume of stems and large branches through a voxel-based approach (Hosoi *et al.* 2013).

Regarding the spaceborne devices, GLAS mounted on-board NASA's ICESat satellite was the unique satellite LiDAR device used for monitoring forest covers (22.58 %). It is worth noting that even if a few studies were based on the satellite LiDAR platform, the major part of them proved a combined use of spaceborne LiDAR and other RS sources, as for example, the combination of LIDAR data with open-source satellite imagery as Landsat TM, ETM+ (Ke *et al.* 2010) and commercial satellite as Quickbird, and IKONOS data (Clawges *et al.* 2008).

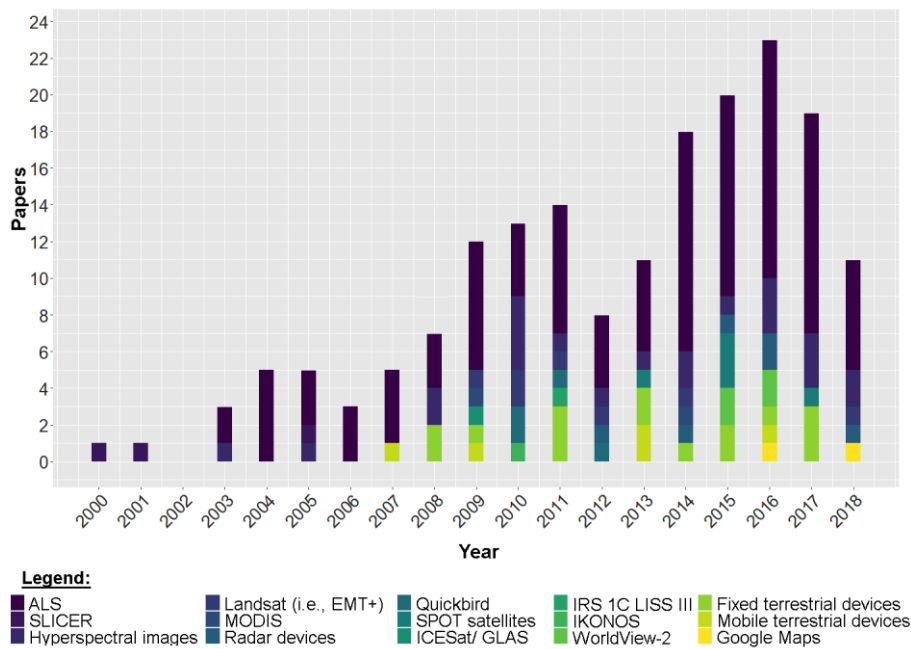


FIGURE 5 STACKED BAR GRAPH OF RS DEVICES USED IN THE PERIOD (2000-2018).

2.1.5.3. Diachronic analysis of LiDAR studies for assessing forest inventory variables.

Over the last two decades, researchers explored various FIVs (i.e. stem density “SD”, diameter at breast height “DBH”, tree height “TH”, basal area “BA”, stem volume “VOL”, above-ground biomass “AGB”, carbon stock “CS” and vary Vis (i.e. leaf area index “LAI” and plant area index “PAI”) using airborne (i.e. ALS) and spaceborne (i.e. GLAS) devices. More precisely, GLAS LiDAR data were used mainly to assess VOL and TH, while, ALS data were used to assess not only VOL and TH, but also SD, DBH, AGB, CS, LAI and PAI.

The first ALS study focused on modelling the FIV and VIs was implemented in a mixed and even-aged forest in Canada (Lim *et al.* 2003). ALS metrics (LiDAR-derived metrics) extracted from plots (49 plots of 400 m² were used to predict various FIVs. The accuracy reached, based on the coefficient of determination (R-squared), was 0.86 for TH, 0.63 for DBH, 0.87 for VOL, 0.85 for AGB, 0.82 for LAI, 0.76 for canopy openness and 0.86 for SD. In the subsequent year, Patenaude *et al.* (2004) used the canopy height model (CHM) as input data for quantifying the above-ground carbon content (AGCC) of woodland stands. The results were quantified at stand and woodland level. The coefficient of correlation between the AGCC prediction at plot level from both CHM (grid 20 * 20 m) and field-based approaches was equal to 0.85. In 2005, (Parker and Mitchel, 2005) focused to assess the influence of occlusion factors (i.e. quality of point cloud) for VOL prediction accuracy. In particular, the occlusion factors tested were: point density (low/high), processing of ALS data (smooth and unsmooth methods applied for canopy surface) and forest stand condition (richness of tree species). Best accuracy was found in the forest containing a low number of trees by using low quality point clouds (1 point/m²), and the point cloud was analysed by an unsmoothed

approach. The VOL prediction ($60.48 \text{ m}^3 \text{ 0.40 ha}^{-1}$) was slightly lower than the results obtained by field data ($63.08 \text{ m}^3 \text{ 0.40 ha}^{-1}$).

From 2005 to 2009, for the first time, LiDAR data were integrated with RS data for monitoring forest ecological aspects as the tree and shrub species composition (Hill and Thomson, 2005). The integrated use of ALS data and Airborne Hyperspectral Scanner (HyMap) image was implemented to determine the species composition of two ancient woodland areas (i.e. Monkey and Bevill's woodlands) in the United Kingdom (Hill and Thomson, 2005). To reach this objective, the unsupervised classifier algorithm, named ISODATA, was used for classifying the forest covers. Nevertheless, in 2008, Waser *et al.* (2008), published a paper that demonstrated that the combined use of Color-infrared (CIR) with ALS data can be an effective source for investigating the forest and woodland loss areas by applying a generalized linear model (GLM) method. The factors assessed to reach such an objective were: shrub encroachment and clear-cutting.

In both cases, the product derived from the point cloud, particularly DSM (Digital Surface Model) and CHM resulted strongly useful to classify forest cover maps. It is important to highlight that a pre-elaboration phase, focused on the co-registration and alignment of the LiDAR data with satellite imagery, was necessary and it represents a crucial hindering point for practical uses.

In 2009, Kim *et al.* (2009) used the ALS data to assess and to compare the AGB of living and dead trees in Grand Canyon National Park. Based on the intensity of points cloud, the authors were able to discriminate the pattern of points clouds from living to dead trees. The metrics extracted were separately analysed through a stepwise regression model. Higher accuracy was obtained in live (AGB; R-squared = 0.85; RMSE = 50 Mg ha^{-1}) respect to dead trees (AGB; R-squared = 0.79; RMSE = 42 Mg ha^{-1}). In the same year, based on the assumption that ALS data allows describing the vertical and horizontal vegetation structure, Kellner and Asner (2009) evaluated the applicability of ALS data for assessing forest disturbances in Costa Rica and Hawaii rainforest. In this study, the ALS data was firstly used to assess the canopy height, and then the areas with a tree height lower than 2 m were considered as disturbed vegetation (i.e., gap-size). The study allowed identifying 434501 gap sizes across all rainforests.

Since 2010, some studies have demonstrated that the integrated use of different sources of RS data can be advantageous to surpass the uncertainty associated with the quality of some of them. To support such a statement, Ke *et al.* (2010) combined low-quality ALS data ($<1 \text{ point/m}^2$) with well-quality RS data (Quickbird at 2.4m) to classify the forest species composition through machine learning algorithms. In this study, different metrics (i.e. spectral, topographic and ALS data) were analysed as input variables through the decision trees algorithm, reaching a kappa accuracy of 91.6%. Similarly, (Arroyo *et al.* 2010) applied the Object-based image analysis (OBIA) algorithm to map riparian forests in Australian tropical savannah areas. The overall accuracy reached by Arroyo *et al.* (2010) was equal to 85.6 %.

In Italy, in San Rossore regional park, (Maselli *et al.* 2011) combined low-quality of ALS data ($<9 \text{ points/m}^2$) with low-quality RS data (Landsat TM/ETM+; $30\text{m} \times 30\text{m}$ of pixel size) to map VOL. They use ALS metrics to predict VOL through machine learning algorithms calibrated by least-squared regression. Subsequently, they use Landsat TM/ETM+ data to upscale the VOL prediction from plots over to the large forest area using machine learning algorithms (i.e. locally calibrated regression "LCR" and k-Nearest Neighbors "k-NN"). They obtained more accurate maps with LCR (coefficient of correlation, $r = 0.88$; Root

means squared error “RMSE” = $48.9 \text{ m}^3 \text{ ha}^{-1}$) compared to k-NN ($r = 0.847$; $\text{RMSE} = 59.2 \text{ m}^3 \text{ ha}^{-1}$), while the least-squared regression applied for reference data was less accurate ($r = 0.626$).

Conversely, (Chen *et al.* 2012) combined low-quality ALS data ($<3 \text{ points/m}^2$) with well-quality RS data ($< 2.44 \text{ m}$ of pixel size) for assessing TH, AGB and VOL in forest ecosystems located in Boreal Shield Ecozone (Canada). The GEOgraphic Object-Based Image Analysis (GEOBIA) algorithm was implemented. Though the high quality of input data, the accuracy reached by GEOBIA was similar to the accuracy obtained in the previous studies ($R\text{-squared} = 0.85$; $\text{RMSE} = 3.37 \text{ m}$) for TH, ($R\text{-squared} = 0.85$; $\text{RMSE} = 39.48 \text{ Mg ha}^{-1}$) for AGB and ($R\text{-squared} = 0.85$; $\text{RMSE} = 52.59 \text{ m ha}^{-1}$) for VOL.

A different application of ALS data was recently performed by N. d’Oliveira *et al.* (2012), which focused to map the loss of forest cover due to the harvesting activities in Antimary State Forest (Brazil). They obtained an accuracy slightly lower than the previous studies (AGB, $R\text{-squared} = 0.72$ and VOL, $R\text{-squared} = 0.69$), and a marked difference of the AGB amounts was reported between impacted (225.3 Mg ha^{-1}) and non-impacted (233 Mg ha^{-1}) forest covers. Slightly better was the accuracy obtained by Tompalski *et al.* (2015), which focused to predict VOL ($R\text{-squared Adjusted} = 0.86$) and SD ($\text{RMSE} = 149 \text{ trees ha}^{-1}$ or 24.4%) in Vancouver Island forests.

From 2015 onward, the machine learning algorithms were frequently used to elaborate LiDAR data, due to authors recognized their power to improve the accuracy in prediction of FIVs. Particularly, an useful algorithm was the random forests (RF), which was implemented for example to upscale the prediction of the AGB and VOL from ALS transect to national forest cover in Canadian boreal forest (similar approach tested by Maselli *et al.* (2011)). The Bayesian linear model (Junttila *et al.* 2015; Kauranne *et al.* 2017) was also used for assessing FIVs in Taiga boreal forest (Russia), obtaining the accuracy in terms of RMSE equal to $0.28 \text{ m}^3 \text{ ha}^{-1}$ for VOL, 0.16 cm for DBH, and 0.07 m for TH.

Over the years, some authors (Ortiz-Reyes *et al.* 2015) focused to compare and select the best predictors models (i.e. multiple linear regression, non-linear regression, ratio estimators and traditional forest inventory “stratified sampling”) in order to improve and foster the prediction of FIVs using LiDAR data. Ratio estimator proved best fit predictions for VOL ($R\text{-squared} = 0.79$; $\text{RMSE} = 2.07 \text{ m}^3$), BA ($R\text{-squared} = 0.77$; $\text{RMSE} = 0.21 \text{ m}^2$) and crown coverage ($R\text{-squared} = 0.53$; $\text{RMSE} = 139.71 \text{ kg}$) and AGB ($R\text{-squared} = 0.76$; $\text{RMSE} = 1340.08 \text{ m}^2$). In addition, to reduce the huge point information using faithful approaches, (Moser *et al.* 2017) benchmarked three approaches aimed at reducing the number of predictor variables extracted from ALS data: Genetic algorithm combined with RMSE, Genetic algorithm combined with AIC, and RF algorithms. This study revealed that the Genetic algorithm with AIC (Akaike information criterion) showed a great capacity for selecting the predictor variables, ensuring fitted values in the prediction phase.

2.1.5.4. Diachronic analysis of LiDAR studies for assessing forest productivity variables.

Over the last two decades, researchers explored various forest productivity-related issues using airborne and terrestrial LiDAR devices, within which the most explored were VOL, AGB and CS using both devices, and the stem diameter and stem reconstruction using terrestrial LiDAR data.

Results show that the first ALS study focused on retrieving VOL and AGB at a single tree level was carried out in 2003 in pine and deciduous forest stands (Popescu *et al.* 2003). The authors implemented firstly the

local maxima filter with variable window size, “LM_{WS}” for detection of trees and then the region-growing algorithm supported by Erdas Field Guide for the segmentation of tree crowns. The values obtained for the tree crowns, tree crown diameter (TDC) represented the input data to predict VOL and AGB.

In 2004, Maltamo *et al.* (2004) proposed an alternative tree detection algorithm, named Arboreal Forest Inventory Tools software, for detecting trees within the temperate forest (Finland), and predicting the VOL and the SD. The differences compared to the previous study were: predictor variables (TCD vs. ALS metrics); objectives (VOL and AGB vs. VOL and SD); modelling approach (equations vs. parameter prediction method (PPM) and left-truncated distribution Truncated (Tw). This approach was able to identify between 309 and 694 out of 756 stems (large and small trees) with higher accuracy for VOL (RMSE_{PPM} = 16% and RMSE_{Tw} = 22.5%) compared to SD (RMSE_{PPM} = 49.2% and RMSE_{Tw} = 72.7%), respectively. Enhanced detection accuracy, over small and large trees (ranged between 30% and 90%), was reported for (Maltamo *et al.* 2004) compared to LM_{WS} method, because (Popescu *et al.* 2003) was focused on detecting large trees (15 % for the omission in detecting dominant and co-dominant trees).

In 2006, a study tested the capability of LM_{WS} and spatial wavelet analysis (SWA) for improving the accuracy of FIV assessment at a single tree level (Falkowski *et al.* 2006). The LM_{WS} approach analysed the variation of Z values to search the tree position and tree crown dimension (based on region-growing), while the SWA analysed the signal derived from the sums of simpler trigonometric functions to define the tree profile. Enhanced prediction accuracy was observed using the SWA compared to LM_{WS} approach (SWA: R-squared = 0.97 and RMSE = 2.64m, and LM_{WS}: R-squared = 0.97 and RMSE = 2.81m) and the crown dimension (SWA: R-squared = 0.86 and RMSE = 1.35m, and LM_{WS}: R-squared = 0.79 and RMSE = 1.66 m²) responses. Subsequently, (Chen *et al.* 2007) obtained better results implementing the LM_{WS} in combination with watershed segmentation (changed from region-growing method) for detecting trees in savannah woodland. The combined use suggested by Chen *et al.* (2007) resulted in more performance than the previous studies, particularly, for calculating the VOL and BA at a single tree level. In 2011, Lin *et al.* (2011) developed the algorithm MMAC (Multi-level morphological active contour) to further improve the detection accuracy of trees within 3 plots of 0.25 ha of mixed-species, plantation and coniferous stands, respectively. Though the algorithm allowed detecting 76 % of trees, the study highlighted a sensitivity for omission error (24%) rather than for commission error (13%).

In the same years, the use of TLS data for assessing forest production increased due to the high accuracy for tree architecture description. Results showed that the first study that used TLS points cloud was carried out in 2009 (Antonarakis *et al.* 2009) aimed to discriminate leaves from timber using geometry-based patterns (i.e. roughness). Other TLS studies using automatic algorithms for timber-leaves discrimination, stem diameter and stem reconstruction were also investigated, as for example, (Yao *et al.* 2011) demonstrated that the automatic TLS algorithm named find trunks, which was previously developed by Lovell *et al.* (2011), was able to automatically produce the stem position and stem diameter of standing trees in mixed-species stands using Echidna® validation instrument (EVI) TLS data. These outputs were used for calculating the AGB. Higher R-squared values were found in SD (0.90) and AGB (0.85) compared to DBH (0.48); (Côté *et al.* 2011) revealed that the L-Architect algorithm, which was based on the skeletonization approach, was able to automatically reconstruct the stem architecture in coniferous pure

forest stands. These two studies highlighted that the accuracy of the prediction models was conditioned by operational aspects and weather conditions. However, the first algorithm describes the tree structure characteristics of 42% of all trees within 20 m of the scan position, and the L-Architect algorithm reconstructs the stem, large and small branches of solely five stems.

Unlike the automatic TLS processing, innovative approaches aimed at improving the characterization of stems were also studied. For example, (Moorthy *et al.* 2011) tested an approach, named cross-sectional slicing, for describing the stem crown profile of trees through the separation of point cloud into several horizontal slices. The R-squared reached for LAI was ranged between 0.86 and 0.99 and for PAI was 0.76 (RMSE = 0.26 m² m⁻²); (Bremer *et al.* 2013) highlighted the importance of eigenvectors and eigenvalues patterns, from TLS and Mobile Laser Scanning (MLS) devices, in depicting the vegetation structure through the skeletonization approach; (Hosoi *et al.* 2013) highlighted that the grouping the points in a voxel space through a voxel-based approach (voxel size 0.13 cm³) can facilitate the reconstruction of the architecture of trees using Portable Laser Scanning (PLS). The VOL accuracy reached was higher for the stem and large branches (0.5 %, percentage error) than small branches (34 %, percentage error).

Since 2013 onward, some studies assessed the feasibility of the combined use of active RS data for enhancing the accuracy of models to assess VOL and AGB. For example, (Allouis *et al.* 2013) tested the accuracy of two LiDAR systems (i.e., discrete return and full-waveform data), to evaluate the prediction of the VOL and AGB. As regards the AGB assessment, the full-waveform data proved more fit predictions with respect to discrete return (mean error = -15 vs. mean error = 5, respectively), and a similar pattern was observed for the prediction of VOL (mean error = -4 vs. mean error = 2, respectively). The combination of ALS data with RS data offers a more accurate estimation of timber provision with respect to the results obtained through one type of data (Shinzato *et al.* 2017).

In 2017, as regards, TLS algorithms, (Saarinen *et al.* 2017) evaluated the automatic TLS algorithm, named cross-sections approach (also known as cylinder-fitting), which was already tested after 2004 (Pfeifer and Winterhalder, 2004), was able to automatically output the stem curve measurements and TH in mixed temperate forest using both TLS data (single scan and multiple-scan). Better accuracies were obtained using multiple-scans compared to single-scans, particularly, the rate of automatically derived diameters of multiple-scans (53.3 - 60.9 %) was higher than that observed using single-scan (44.1 %), however, the time consuming for collecting TLS single-scan (less than 10min.) was shorter than multiple-scans (less than 20min and between 20 - 30 min.)

From 2014 to 2018, as regards ALS algorithms, most studies introduced novel and versatile approaches for detecting trees in mixed-species and multi-layered forest stands. For example, focused on coniferous stands, Mikko *et al.* (2014) proposed an algorithm, named Multisource Single-Tree Inventory (MS-STI), to detect and to classify the stems and to predict the TH, DBH, VOL and timber assortments using ALS (< 9points/m²) and RS (0.25 m; UltracamXp) data. The reached RMSE for TH was 4.2 - 5.3%, for DBH was 10.0 - 19.9 %, for saw-log volume was 28.7 - 43.5 % and for pulpwood volume was 125.1 - 134.3 %. Moreover, focused on mixed-species and multi-layered forests, (Sačkov *et al.* 2016) proposed an algorithm, named reFLex, to detect the stems across several ALS horizontal slices obtained through the stratification approach of the point cloud (average 30 points/m²). This algorithm reached to detect 66 % of dominant, 48

% of the codominant, 18 % of the intermediate and 5 % of the suppressed stems. Similar stratification approach was applied to divide the ALS point cloud (average 50 points/m²) into overstory and multiple understory DSM-canopy layers (Hamraz *et al.* 2017). They reached to detect between 46 % and 68 % of trees for the understory layer. Even if the abovementioned studies suggested that the dense point cloud can increase the DR, the terrain conditions of plots (slope lower or equal to 50°) can also underestimate the TH of large trees (16.6 m) (Alexander *et al.* 2018). Branson *et al.* (2018) used an open-source derived from aerial and street view images of Google Maps™ for identifying the trees over the street through machine learning algorithms (i.e. convolutional neural networks). This study allowed detecting more than 70 % of trees across streets, 80 % of them was correctly matched with the tree species.

2.1.5.5. *Timber assortment estimation, algorithms*

The most common methods (Table 5) used for pre-processing raw LiDAR data were: a) open-source algorithms developed by the authors (Popescu *et al.* 2003; Chen *et al.* 2007; Allouis *et al.* 2013), b) commercial algorithms (i.e. Cyclone 5.5 software, TerraScan ® software, LASTools software and the FUSION v3.50 software) and c) open source for scientific purpose, as for example, OPALS (Opals Orientation and Processing of Airborne Laser Scanning data software).

Results revealed that the detection and segmentation of trees were mainly carried out through the combined use of LM_{WS} with region-growing and or watershed segmentation as well as the largest and robust processing.

From 2003 onward, the LM_{WS} algorithm was the most applied algorithm for detecting the tree position of dominant trees for deciduous, coniferous and mixed stands. The ALS grid sources used were the CHM, DSM and DEM (Digital Elevation Model) raster-based. For the segmentation approach, we observed many approaches for delineating the crowns (i.e. region-growing, watershed segmentation, eCognition software). Studies using algorithms able to automatically develop both steps were reported after 2010: reFLex algorithm, MS-STI approach, OPALS software, Arboreal Forest Inventory Tools of Arbonaut algorithm. As regards the modelling of timber provision, we noted a marked difference between the first and second decades of the period considered (2000-2018), which goes the use of solely least-squared regression (i.e. polynomial, exponential, logarithm) to even implement machine learning algorithms (Support Vector Regression “SVR”, Bayesian algorithm, K-NN, RF). This is because machine learning has proven a great capability to accurately predict the forest variables (García-Gutiérrez *et al.* 2015), as well as to infer, to upscale, to model, to classify these using even huge quantity of LiDAR information, without loss the accuracy of models (Maselli *et al.* 2011; Junttila *et al.* 2015; Kauranne *et al.* 2017). While TLS data was able to reconstruct many segments of trees (i.e. stem, large and small branches), the automatic and innovative approaches tested were: T-Architect, voxelization, skeletonization, cross-slicing algorithms, cylinder-fitting approach, convolutional neural network.

TABLE 5 MOST FREQUENT METHODS FOR PROCESSING LIDAR DATA. THE PRE-PROCESSING METHODS (1), TREE DETECTION APPROACHES (2) AND MODELLING THE TIMBER PROVISION VARIABLES (3) WERE DISPLAYED.

Methods applied for forest productivity purpose					
Device/ Year	Sensor	Methods			Study/ Journal
		(¹)	(²)	(³)	
ALS (2003)	AeroScan	The method explained by (Popescu <i>et al.</i> 2003).	LM _{WS} algorithm and tree crown algorithm.	Modelled through least-squared regression models.	(Popescu <i>et al.</i> 2003). “Can. J. Remote Sens.”
ALS (2010)	Optech ALTM 3033	TerraScan ® software	LM _{WS} algorithm and semi-variogram and tree-spacing approaches.	Modelled through equations.	(Tesfamichael <i>et al.</i> 2010) “Progress in Physical Geography”
ALS (2011)	ALS50	TerraScan ® software	MMAC algorithm		(Lin <i>et al.</i> 2011) “Photogramm. Eng. Remote Sens.”
ALS (2013)	RIEGL LMS-Q560	The method explained by the authors	The method explained by the authors	Modelled through least-squared regression models.	(Allouis <i>et al.</i> 2013) “IEEE J-STARS”
Terrestrial device (2015)	Nikon D5000 (Hyperspectral camera)	SFM-MVS software.	PhotoScan software (Agisoft LLC, St. Petersburg, Russia).		(Miller <i>et al.</i> 2015) “Urban Forestry and Urban Greening”
ALS (2016)	YellowScan Mapper	LASTools software.	reFLex algorithm.	Modelled through equations and least-squared regression models.	(Sačkov <i>et al.</i> 2016) “Forests”
ALS (2017)	ALS70 HP	TerraScan ® software and FUSION v3.50 software.	LM _{WS} in FUSION v3.50 and Near Table in ArcGIS (v10.1).	Modelled through k-medoids algorithm.	(Alexander <i>et al.</i> 2017) “Int J Appl Earth Obs Geoinformation”
ALS (2017)	ALS70-CM and Leica RCD30	The method explained by Sačkov <i>et al.</i> (2017).	eCognition and reFLex algorithms.	Modelled through equations and least-squared regression models.	(Sačkov <i>et al.</i> 2017) “Forests”
ALS (2017)	RIEGL LMS - Q680I	The method explained in Shinzato 2017.	PCA (Principal Component Analysis) tree detection method.	Modelled through Best Subset Selection and cross-validation approaches.	(Shinzato <i>et al.</i> 2017) “IForest”
ALS (2017)	ALS70 HP	FUSION v3.60.	LM _{WS} in FUSION v3.50 and watershed segmentation approach.		(Alexander <i>et al.</i> 2018) “Int J Appl Earth Obs Geoinformation”

2.1.5.6. Timber assortment estimation, strengths and weakness

This section highlights the most important strengths and weaknesses come out from forest productivity literature. First of all, terrestrial LiDAR devices provide a better reconstruction of trees rather than airborne LiDAR devices, especially the TLS devices, which demonstrated slight improvements with respect to other terrestrial devices (as for example PLS and MLS). Nevertheless, its practical usability is strongly conditioned by technical and handling, as well as site-specific characteristics, as for example, the number of scans influencing the quality of tree modelling are associated to the point cloud densities, weather conditions, richness of species, terrain slope, the density of trees and forest structure.

In dense forests with high coverage of branches and leaves, the prediction of TH is rather difficult (Saarinen *et al.* 2017). However, by reducing the distance between the TLS devices and tree stem target, these issues can be surpassed (Moorthy *et al.* 2011), despite the presence of hindering factors, weather conditions (i.e.

wind), affecting the usage of it (Côté *et al.* 2011). Significant improvements were obtained by combining the TLS and MLS data, where this last suggested to authors that the issue associated with the displacement of the fixed device can be overcome without losing the accuracy. Difficulties associated with processing suggested that the implementation of robust algorithms in modelling phases, such as object-based associated with skeletonization approach (Bremer *et al.* 2013), voxel-based approach (Hosoi *et al.* 2013) and the cylinder-fitting approach (Shinzato *et al.* 2017) resulted very useful for the stem reconstruction using TLS data.

From forest productivity literature, few hindering factors, mainly related to the identification of very small branches (i.e. branch very slender < 5mm and located between 20 and 30 cm of the final part of branches) and the low-quality TLS point clouds in the upper canopy part were highlighted (Miller *et al.* 2015). This means that the small trees are affected by noise point clouds, derived from assembled and co-registration problems, respect to large trees (Bremer *et al.* 2013). However, this issue can be surpassed by placing wooden stakes around single trees (Moorthy *et al.* 2011). It is important to note that these hindering factors represent important challenges, rather than true obstacles for forest inventory and practical applicability and usability. Terrestrial photographs, acquired from the hyperspectral camera (Nikon D5000, lens: AF-S NIKKOR 35 mm) can offer important advantages in the monitoring of mixed-forest stands due to the capacity to be easy to handle at low cost in inaccessible forest field areas (Miller *et al.* 2015).

Regarding the ALS advances, the development of a robust algorithm allowed authors to overcome the bottleneck phase affecting the correct calculation of timber assortment variables. The most frequent algorithms used for detecting and segmenting trees were the LM_{WS} algorithm associated with region-growing and watershed segmentation. However, these methods were focused mainly on large rather than on small trees. The tree detection algorithm associated with the stratifying approach of point cloud was used for identifying trees into several strata (Sačkov *et al.* 2016; Hamraz *et al.* 2017). In addition, the detection approach, the selection of best explicative predictor variables improving the performance of models (Ortiz-Reyes *et al.* 2015), as well as the modelling of such predictor variables using both least-squared regression and machine learning algorithms (i.e. SVR, Bayesian algorithm, K-NN, RF).

2.1.6. Discussion

The results revealed that since 2000 the interest shown by researchers for monitoring the timber assortments using LiDAR data is increased, especially in the period 2010 – 2016 (Figure 2), with more studies carried out in North America (43.40%) and Europe (42.77%) (Figure 3). This result was supported by the capability of LiDAR to rapidly, remotely, accurately depict the vegetation structure over distinct species, useful information from forest stands to support local and national forest inventory, ensuring the monitoring of timber provision in forest stands through sampling campaign (McRoberts and Tomppo, 2007; Vastaranta *et al.* 2014; Beland *et al.* 2019; Chirici *et al.* 2020). As regards the high interest of the scientific community from North America and Europe given to the timber assortment LiDAR issues, it could be supported by two assumptions: 1) the firsts LiDAR studies (in absolute terms) aimed at monitoring the forest areas using LiDAR data were conducted in the Soviet Union, Finland and North America (Montaghi *et al.* 2013; Kelly and Di Tommaso, 2015); 2) the LiDAR data for many states of North America (e.g., Ohio, Pennsylvania)

and Europe (e.g. Germany) is being increasingly available (<http://lidar.cr.usgs.gov/>; <https://www.geodaten.sachsen.de>). Thereafter, after the ends of December 2018, the LiDAR data has become completely available for the Earth's surface (<https://gedi.umd.edu/instrument/launch>).

The authors showed a poor interest in assessing timber assortment resources (23.5% of the total papers) compared to indicators of biodiversity (7.26 % of the total papers), as well as climate change effects on forest stands (4.47 % of the total papers) using LiDAR data (Figure 3). As far as concern the biodiversity, LiDAR data was used to describe the vertical and horizontal vegetation structure and tree species composition, two essential indicators used for assessing the habitat quality, occurrence of bird species. Regarding climate change impacts, LiDAR data were used to estimate forest cover loss or changes used for assessing the forest health after fire events, disease and other disturbance events. Though the methods for assessing the indicators of forest biodiversity and forest climate change are well documented from passive RS devices, due to a variety of useful information extracted from those devices (i.e. spectral reflectance, spectral indices, texture, spatial/temporal features) (Petrou *et al.* 2015), LiDAR data remains an essential and faithful data for extracting the tree structure and tree species composition from forest stands (Bergen *et al.* 2009). To date, despite the methods for assessing the forest biodiversity and forest climate change indicators using the tree structure as a proxy are well-documented (Ishii *et al.* 2004; Bohn and Huth, 2017), accurate characterization of trees from ALS data requires powerful tree detection methods (Hamraz *et al.* 2017; Sačkov *et al.* 2016), and well-quality of LiDAR data (i.e. 10-25 lidar maximum pulse size, the average of 15cm for waveform data) using spaceborne devices (Bergen *et al.* 2009). Moreover, another cause was associated with the cost of LiDAR acquisition (nearly half of the survey traditional inventory cost) (Kelly and Di Tommaso, 2015). However, nowadays, spaceborne LiDAR data are available at the worldwide level (at Earth level; <https://gedi.umd.edu/instrument/launch>), and airborne LiDAR data for various countries are already available (free source: <https://www.geodaten.sachsen.de>; commercial source: <http://www.pcn.minambiente.it/mattm/procedura-richiesta-dati-lidar-e-interferometrici-ps/>). Nevertheless, spaceborne LiDAR data might be most effective to retrieve forest inventory rather than timber assortment information. Therefore, in the future, it is expected the use of LiDAR data for extracting the timber products from forest stands, maintaining and conserving the biodiversity become more explored.

Papers included in the productivity cluster have highlighted that the most terrestrial LiDAR device used to retrieve accurate evaluations of the timber assortment was the TLS device. The results could be supported by the capability to rapidly and accurately depict the stem profiles, as well as the TLS algorithms (i.e. cylinder-fitting) available for extracting the forest productivity variables. The success or failure of the tree architecture representation from point clouds takes into account many factors, such as, the operational (i.e. sampling design, number of scans), technical (i.e. pre-processing and processing) and weather factors (i.e. wind) (Dassot *et al.* 2011). Along with these factors, some studies revealed that the automation of the TLS algorithm and the forest structure plays an important role in the processing phase, especially for stem reconstruction (Liang *et al.* 2018). Some of them allowed to reconstruct the trunk (i.e. cylinder-fitting), some of them allowed to reconstruct the whole stem architecture (i.e. skeletonization) and some of them allowed to extract automatically the VOL, AGB, CS of standing stems (L-Architect, Computree:

<http://computree.onf.fr/>, opalsForest OPALS package). Nevertheless, automatic and non-automatic approaches for timber assortment estimation (i.e. saw-log, pulpwood, or assortment types) from TLS data remain still required, especially for single trees. Therefore, an approach for timber assortment estimation becomes essential for extracting the timber products from forest stands, maintaining and conserving biodiversity.

The results revealed that, despite the capability for reconstructing the stem from TLS data, the ALS was the most implemented technique for investigating productivity-related issues. These results, however, have highlighted that the detection of trees (Table 5) remains a challenging step. Nevertheless, most of them, well-documented, were focused on large trees (dominant and co-dominant trees) for pure and mixed-species (Kaartinen *et al.* 2012; Wang *et al.* 2016), because of the detection accuracy decrease for small trees, especially in forests characterized by high structural heterogeneity. In this context, some authors have suggested that occlusion factors were mitigated for large and small trees in ALS plots with higher than 30 points m⁻² and dividing the ALS point cloud into varying slices (i.e. stratifying approach) (Sačkov *et al.* 2016; Hamraz *et al.* 2017).

The integration of ALS data with the RS data gained more attention from researchers for forest monitoring. This greater attention was supported by the capability of spectral resources from RS data to reinforce the low-quality of point clouds (<10 points/m²) become an effective approach for characterizing the trees and to cope with distinct forest ecosystem issues (Wulder *et al.* 2012). Conversely, despite the significant contribution received from the integration of ALS with RS data, the implementation of it on timber assortment estimation remains still a critical challenge. Unlike the combined use of ALS with RS data, the combined use of ALS with TLS data provided an accurate estimation of the timber assortments and stem distribution at large scales, however, in this combination, the role of TLS data became essential (Kankare *et al.* 2014), because TLS was able to depict small and large profiles and vary automatic algorithms are available (Liu *et al.* 2017).

The results revealed that the implementation of machine learning algorithms is being increasingly tested for analysing the LiDAR data. The implementation was supported by the power of machine learning for running many functions (i.e. inferring, upscaling, modelling, classifying), as well as offering better-fitted timber assortments prediction using even huge quantity of point density compared to least-squared regression. The strengths of machine learning reside on the strategies used for processing LiDAR data, as for example, conceptually, RF, is an ensemble method and it is composed by a combination of several tree predictors, which are composed by random, independent and equally distributed values within a dataset, practically, RF is capable to process the LiDAR data through a decision trees approach, and to select the most explicative predictor variables, from a training data, measuring the variable importance (Breiman 2001); conceptually, K-NN is a non-parametric method and it allows to run the classification and the regression analysis of a dataset: the classification of single vector (with an specific label) is based on the "K" nearest training reference labels within a dataset; the regression of single value is inferred as the weighted average values based on the closely training reference value within a dataset; practically, K-NN

predicts the environmental parameter value for every pixel using the weighted average of the nearest "K" observing parameter values (Maselli *et al.* 2011). Nevertheless, the capability to make better use of the reduced number of predictor variables compared to the least-squared approach was even highlighted (Hauglin *et al.* 2013). A systematic benchmarking of fourteen algorithms, between machine learning (i.e. SVR, KNN, RF) and least-squared regression (i.e. liner regression) methods, were used for calculating various FIVs using LiDAR data. The results revealed that better R-squared and RMSE results were obtained from machine learning algorithms, especially the SVR algorithm (García-Gutiérrez *et al.* 2015). Despite the countless benefits provided by machine learning algorithms, its use for timber assortment estimation is still limited, especially in forests characterized by mixed-species and multi-layered (Vastaranta *et al.* 2014; Silva *et al.* 2017). For this reason, studies that deal with the implementation of machine learning algorithms for timber assortment assessment can be useful to improve the use of LiDAR data for forest monitoring (Vastaranta *et al.* 2014).

In addition, we noted that a considerable part of productivity studies (about 69 % of "P" literature) benefited from the funds provided by the NATIONAL programs (i.e. Institute, University and or Ministry): e.g., MOST fund of the Ministry of Science and Technology (Taiwan) and National Council of Technological and Scientific Development (CNPq), and EUROPEAN programs (i.e. LIFE programs): e.g., FRESH LIFE and ForestSAFE projects. More precisely, two types of collaboration strategies were adopted for those studies: 1) the scientific collaboration with expert researchers of pre-processing and processing of LiDAR data, and with researchers having LiDAR raw data (about 83 % of "P" literature); 2) the scientific collaboration between two or more institutional entities (i.e. University, Research Institute) geographically distant between them were also adopted into productivity studies (about 17 % of "P" literature) e.g., "NASA" – "National Key Basic Research Development Program of China" and "FRESH LIFE project Italy" – "Slovak Research and Development Agency". Analysing both collaboration strategies, the implementation of the first strategy could be supported by the fact that one LiDAR strip raw data can be used for several forest issues, and the implementation of the second strategy could be supported by the efficiency and synergism of methods aimed at accurately analysing the LiDAR data.

2.1.7. Conclusion

This paper provides an overview of the recent advances in using LiDAR data for timber assortment estimation. This review outlines the importance of the LiDAR data for tree structure characterization, as well as to cope with several forest inventory issues.

There was a greater interest by research to use LiDAR data for retrieving and gathering forest inventory information from forest stands (45.3 %), followed by forest productivity information (23.5%). The poor interest from researchers for timber assortment assessment maintaining and conserving the biodiversity was supported by the proportion of studies found out in "B" (7.26 %), and the studies focused on monitoring the climate change effects on trees were even limited (6.15 %). These results highlighted that to sustainably manage the forests using LiDAR techniques is still necessary for greater efforts.

The two implemented LiDAR devices used for timber assortment assessment were airborne (i.e. ALS) and terrestrial (i.e. TLS) devices. As concerns the ALS data, based on the assumption of the timber assortment prediction increase, as tree detection accuracy increase, several tree detection algorithms were tested (i.e. reFLex, MS-STI). However, the best tree detection algorithms tested on forests characterized by mixed-species and multi-layered require specific point cloud data (>30 points/m²) and processing (stratifying approach). The best approach for extracting the timber assortment from standing trees was the TLS. TLS is a powerful source to reconstruct the stems from standing trees through automatic approaches (i.e. cylinder-fitting). However, the success or failure of the TLS implementation for forest monitoring takes into account some limitations, such as, operational and technical aspects, and weather conditions. The integration of ALS with TLS proved to be an effective method for timber assortment prediction at plot level, however, among them, TLS data played an essential role in the accuracy of this study. The implementation of machine learning algorithms became effective for predicting, upscaling, modelling and classifying the LiDAR data in forest studies. However, studies using machine learning algorithms for timber assortment assessment were still limited, especially at the single tree level.

This study outlines that the success of LiDAR implementation for timber assortment assessment was followed by optimal flight campaign strategies, which were translated into the good quality of point cloud data, robust tree detection methods, and machine learning approaches.

In conclusion, for small areas, the TLS data became the most effective method for timber assortment estimation, while for large areas, the integration of ALS with TLS using machine learning algorithms became the most effective method for timber assortment estimation. Further investigation deals with the processing of the combined use of ALS with TLS data through machine learning can be useful information for revalorising the stems in mixed and heterogeneous forest stands. Despite the poor interest of researchers showed on forest issues gaps (i.e. biodiversity and climate change effects), it could become the tendency in the future, in the light of the worldwide policies implemented to protect the planet.

2.1.8. Postface

In this study, we developed an overview of the most implemented approaches using LiDAR data for timber assortment assessment. Nearly the half of the studies were focused on assessing inventory information from forests stand based on ABA and nearby one-quarters of studied studies explored productivity indicators based on ITD that gained the least attention by the scientific community for connecting timber assortment with biodiversity and climate change indicators. This amount is worrisome in climate change context and SFM context, also because there is the useful information of how such matter can be resolved. This study highlighted that the most recommended LiDAR data for timber assortment assessment were ALS and TLS. The main challenge of ALS was the detection accuracy, especially in mixed-species and multi-layered forests. While the main challenges provided by TLS studies were associated with operational, technical, weather conditions, forest stand structure. The promising approaches capable to overcome the challenges observed for studies were the stratification associated with the clustering approach for ALS data and cylinder-fitting approaches for TLS data, respectively. A further study testing alternative tree detection

LiDAR as a tool for timber assortment assessment and characterization in mountain forests

approaches, cylinder-fitting and machine learning approaches on LiDAR data can result effective for timber assortment assessment.

2.2. **Unsupervised algorithms to detect single trees in a mixed-species and multi-layered Mediterranean forest using LiDAR data.**

2.2.1. Preface

In the forest, an accurate and reliable method for detecting the trees on ALS data became essential for optimizing the timber assortment extraction from forest stands, as well as for implementing SFM actions. Some studies using ALS data for monitoring the forests proved that the accuracy of the total forest inventory variable increases, as the detection rate increases. However, nowadays, enhanced findings were showed by studies using a well quality of ALS point cloud and within homogenous forest stands (Vastaranta *et al.* 2014; Sačkov *et al.* 2016; Hamraz *et al.* 2017). As a result, a greater detection accuracy was found for trees belonging to the upper and the intermediate layers (Wang *et al.* 2016, 2017; Liang *et al.* 2019) and a lesser detection accuracy was found on trees belonging to the lower layer (Vastaranta *et al.* 2014; Sačkov *et al.* 2016; Hamraz *et al.* 2017). Although the role of the trees belonging to the lower layer can be judged on timber provision terms, the importance of these lies in the capability in ensuring forest regeneration, succession and stability (Jules *et al.* 2008; Antos 2009). In the light of the foregoing, a tree detection method for detecting trees over the lower, intermediate and upper layers can be beneficial to better characterize the trees. In this context, this study introduces a stepwise approach using two unsupervised algorithms for detecting the trees across the strata in mixed-forest and multi-layered Mediterranean forests.

2.2.2. Abstract

Accurate measurement of forest growing stock is a prerequisite for implementing Climate-Smart Forestry strategies. This study deals with the use of Airborne Laser Scanning data to detect carbon stock at the tree level. It aims to demonstrate that the combined use of two unsupervised techniques will improve the accuracy of estimation supporting sustainable forest management. Based on the heterogeneity of tree height and point cloud density, we classified 31 forest stands into four complexity categories. The point cloud for each stand was further cut in three horizontal layers, improving the accuracy of tree detection at tree level for which we calculated volume and carbon stock. The average accuracy of tree detection was 0.48. The accuracy was higher for forest stands with lower tree density and higher frequency of large trees, as well as a dense point cloud (0.65). The prediction of carbon stock was higher with a bias ranging from -0.3 % to 1.5 % and the RMSE ranging from 0.14 % to 1.48 %.

Keywords: Tree detection; Airborne Laser Scanning (ALS), Forest structure, Carbon stock, Natural forest, Climate-Smart Forestry.

2.2.3. Introduction

In Europe, forests cover about 33 % of the total land area (FOREST EUROPE 2015) and play a significant role in climate change mitigation thanks to their capacity to remove carbon dioxide from the atmosphere and to store carbon in timber (Nabuurs *et al.* 2018). Improving the storage of carbon through mitigation

techniques and the adaptation of forest ecosystems to climate change, namely managing the forest in a responsible way, supporting the provision of socio-economic and environmental benefits, requires advanced knowledge and continuous update of forest inventory data (Lindner and Karjalainen, 2007; IPCC 2014). However, traditional forest inventory methods are time-consuming and require enormous efforts, particularly in multi-layered forests or poorly accessible forest areas, like those in mountain areas. In these environments, time-efficient and accurate techniques are required to facilitate data acquisition, particularly to provide timely forest management responses facing climate change in threatened forest ecosystems, such as those of Mediterranean mountains. Information about forest area, forest damages, tree species composition, growing stock, and carbon stock is increasingly important to develop climate change mitigation and adaptation strategies for the management of forest ecosystems (Santopuoli *et al.* 2020b), while maintaining the full set of ecosystem services, in short Climate-Smart Forestry, CSF. Bowditch *et al.* (2020) ranked sustainable forest management indicators to assess CSF, based on their usefulness to monitor forest adaptation and mitigation. Among others, growing stock and carbon stock were considered highly important for CSF. In the last decades, several studies focused on the use of remote sensing for assessing forest growing stock and carbon stock (Chirici *et al.* 2008; Brosofske *et al.* 2014). Since the early 2000s, the use of Light Detection and Ranging (LiDAR) has considerably increased in the forest sector, particularly the Airborne Laser Scanning (ALS), which is a sensor mounted on aerial vehicles (Næsset 1997). ALS provides advantages in the prediction of forest inventory variables at different scales, from the landscape to the stand levels (Montaghi *et al.* 2013; McRoberts *et al.* 2015; Chirici *et al.* 2016), and even at single tree level (Kandare *et al.* 2014; Mongus and Žalik, 2015; Shao *et al.* 2018). The accuracy of prediction is higher for the individual tree-based approach compared to the area-based approach, as demonstrated for example by Yu *et al.* (2010). Despite the increased use of ALS devices for assessing forest inventory variables, the individual tree-based approach remaining very challenging, particularly for trees belonging the understory layers of multi-layered and mixed forests (Kaartinen *et al.* 2012; Sačkov *et al.* 2016; Balsi *et al.* 2018).

We propose that ALS may allow quantifying and monitoring smartness indicators in response to rapidly changing environmental conditions while collecting detailed information on stand productivity, tree health, and species diversity from forest patches. Nevertheless, studies using ALS data to characterize mixed forests showed that the identification of single trees is strongly influenced by forest structure, such as tree species composition, tree height stratification, and stand density (Liang *et al.* 2019; Wang *et al.* 2019). Accordingly, better results at single tree level were obtained in regular forest structures, such as pure conifer stands or forest plantations (Dalponte *et al.* 2015; Torresan *et al.* 2020). Indeed, natural and unmanaged forests represent a hard sampling problem for single tree detection through ALS data, due to the challenges for deriving single tree-related forest inventory variables (Duncanson *et al.* 2014; Kandare *et al.* 2016; Liang *et al.* 2019), which serve as an important benchmark for CSF.

Recently, many approaches have been developed to exploit ALS point clouds for detecting single trees. Kandare *et al.* (2014) and Sačkov *et al.* (2016) used respectively the K-means algorithm and reFLex algorithm, showing several limitations for detecting understory vegetation layers. Both methods detected

about 46 % of trees with height lower than 12 m through K-means and 18 % of all trees in intermediate and suppressed layers through reFLex. To improve the detection accuracy, some authors suggested splitting the point clouds into several tiles simulating the vertical distribution of trees in the forests, obtaining a higher detection accuracy for trees in the understory layers (68 %) (Hamraz *et al.* 2017). Further approaches, such as RANSAC (RANdom SAmple Consensus) algorithm (Balsi *et al.* 2018) and MCGC (Multi-Class Graph Cut) (Williams *et al.* 2019), have been used for tree segmentation with interesting results for trees belonging to large diameter classes (> 30 cm), but with uncertain results for trees with a diameter at breast height < 30 cm. In particular, RANSAC algorithm allowed detecting about 86 % of trees in the overstory layer, while MCGC method allowed detecting approximately 30 % of trees in the understory layer. Overall, the accuracy of the detection rate is higher for trees belonging to the top canopy, rather than for those in the understory vegetation. We hypothesize that the combined use of the clustering approach and the stratification of point clouds may improve the accuracy of results, even with low-density ALS point clouds. Though trees of the understory layer contribute less to the forest carbon sink in comparison with those of the overstory layer, they are crucial for the resilience and the stability of forests, thus contributing to mitigate the effects of climate change (Jules *et al.* 2008; Antos 2009) and ensuring the continuity of forest regeneration and successional processes.

In particular, describing the vertical structure of multi-layered stands, such as the Mediterranean mountain forests that are characterized by a complex stratification of canopy layers and a mixture of tree species, is a difficult task. Despite their continuous improvement, single-tree based methods for delineating vertically heterogeneous canopies remain of difficult application, because of the requirement of site-specific parameters and the geometry of multi-canopy layers (Hamraz *et al.* 2017; Sačkov *et al.* 2016). Developing a suitable method for fostering the segmentation of trees in a multi-layered mixed forest through remote sensing techniques is crucial to support CSF, particularly with the objectives of reducing the loss of biodiversity and increasing the adaptation of trees facing climatic changes.

In this study, we combined, for the first time, two unsupervised techniques to identify individual trees in order to assess carbon stock at the tree level in a mixed-species and multi-layered forest, using ALS data. To reach this objective, we firstly focused on the identification of single trees and subsequently showed the changes in the accuracy of detection rate across the three canopy layers. The successful use of these unsupervised techniques in combination might provide a great contribution in monitoring forest ecosystems and collecting CSF indicators.

2.2.4. Materials and Methods

2.2.4.1. Study area

The study area is located in Central Italy (Molise; 41°42' N, 14° 12' E), namely Bosco Pennataro (Figure 1). Bosco Pennataro is recognized as part of the core area of the Man and Biosphere (MaB) Reserve of Collemeluccio-Montedimezzo Alto Molise and included in the Natura 2000 network. Bosco Pennataro is a mixed Mediterranean forest with 13 tree species, Turkey oak (*Quercus cerris* L.; 40 %), European beech (*Fagus sylvatica* L.; 21 %), and Italian maple (*Acer obtusatum* Mill.; 9.6 %) being the most frequent ones

(Santopuoli *et al.* 2019). The natural forest community is *Aremonio agrimonioidis - Quercetum cerridis* (Biondi *et al.* 2010), classified as Oak-hornbeam according to the European Forest Type (Barbati *et al.* 2014). The mean altitude of Bosco Pennataro is about 930 m a.s.l., while the average annual precipitation and temperature are 723.5 mm year⁻¹ and 14.5 °C, respectively (<https://power.larc.nasa.gov>). The current management system is a high forest with continuous canopy cover and uneven-aged mixed species trees. The average stand density is about 700 trees ha⁻¹, the growing stock is 385 m³ ha⁻¹ of which 366 m³ ha⁻¹ are living trees and 19 m³ ha⁻¹ are standing dead trees (Santopuoli *et al.* 2019). The absence of forestry interventions over the years has facilitated the conversion from even-aged to the uneven-aged forest, supporting the shift of stand structure, from monolayer to multilayer.

The field survey used the one-per-stratum stratified sampling scheme (Barabesi *et al.* 2012). This sampling strategy partitions a region into several equal-size strata and selects one portion for each stratum based on a random and uniform criterion. Based on a one-per-stratum scheme, Bosco Pennataro was stratified into 50 strata and one squared field plot (hereafter ADS) of 529 m² per stratum was randomly selected and considered for the ALS study. Since the ALS strips covered only partially Bosco Pennataro, we selected the ADS covered by ALS data, and 31 out of 50 ADS were selected (Figure 6).

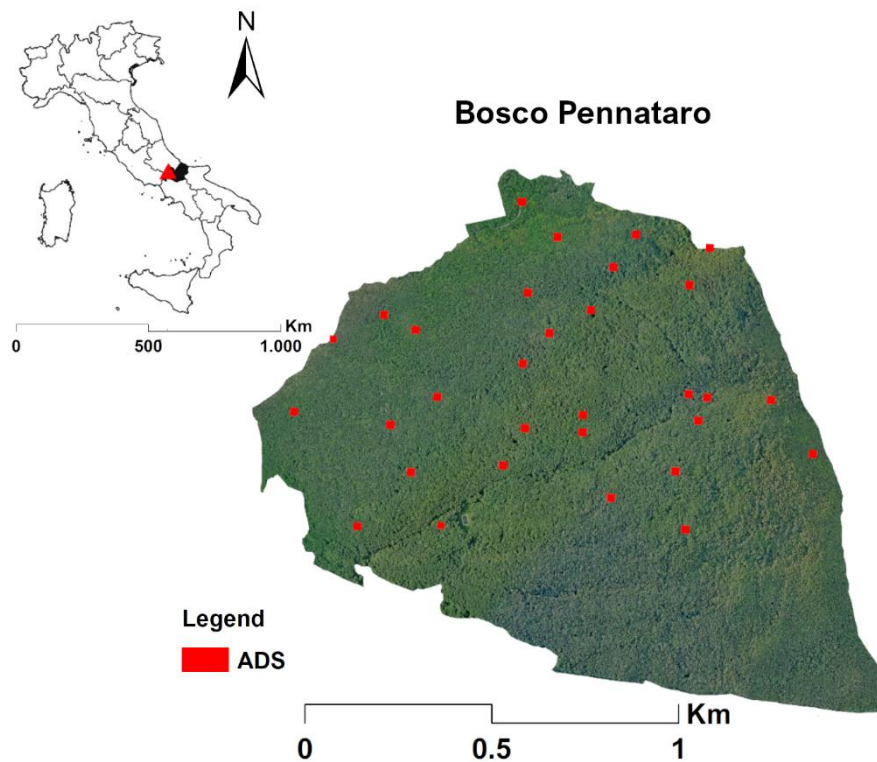


FIGURE 6 LOCATION OF STUDY AREA BOSCO PENNATARO (RED TRIANGLE) AND LOCATION OF THE FIELD PLOTS (ADS).

2.2.4.2. Ground truth data

The forest-related characteristics within each ADS were collected in 2016, using the Field-Map technology (<https://www.fieldmap.cz/>). The sampled parameters were: tree position, tree crown area, tree species, tree

height (TH, m), height of the first branch insertion (I, m), and diameter at breast height (DBH, cm) for all trees with a DBH ≥ 2.5 cm. The stem volume (VOL, m³) was calculated through allometric equations developed for the Italian tree species (Tabacchi *et al.* 2011) and used in the National Forest Inventory. The carbon stock stored in stems and large branches with diameter ≥ 5 cm (CS, tons) was calculated by multiplying the aboveground biomass (AGB, tons) by 0.5 (Federici *et al.* 2008), following the equation (1):

$$AGB = GS * BEF * WBD * A \quad (\text{eq. 1})$$

where:

AGB — aboveground biomass, (tons);

GS — growing stock (m³ ha⁻¹);

BEF — biomass expansion factor, which is equal to 1.47;

WBD — wood basal density (t d.m. m⁻³ f.v.), which is equal to 0.38;

A — forest area occupied by a specific forest category (ha⁻¹).

According to Federici *et al.* (2008), “other broadleaved” forest category was used for BEF and WBD values.

2.2.4.3. *ALS data collection and analysis*

The ALS data were collected in June 2016, in leaf-on forest canopy condition, by Oben S.r.l. company (<https://www.oben.it/sito/>). The LiDAR sensor (YellowScan Mapper) was mounted on an ultra-light vehicle able to collect 3 echoes per laser pulse, with an average point cloud density equal to 60 points m⁻² and accuracy equal to ± 15 cm ($\pm 50^\circ$ of Scan angle and pulse frequency of 20 kHz), however, most points belonged to the first echo. The ultra-light vehicle flew at an altitude of 100 m above ground level.

In this study, a step-by-step methodological approach was implemented, consisting of the following five steps: 1) pre-processing of the ALS data; 2) grouping and stratifying the ADS point clouds; 3) tree detection and segmentation; 4) validation of the predicted tree crowns; and, 5) prediction of forest inventory variables (Figure 7).

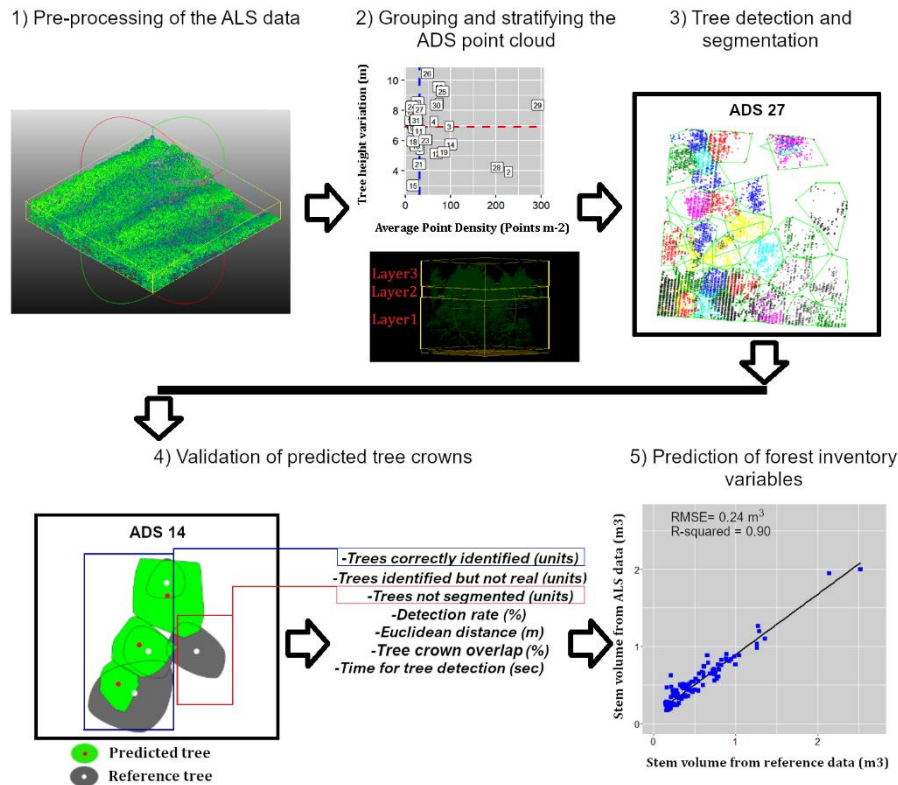


FIGURE 7 METHODOLOGICAL WORKFLOW APPLIED TO DERIVE THE CARBON STOCK AT THE SINGLE TREE LEVEL, USING AIRBORNE LASER SCANNING (ALS) DATA. THE ALS DATA WAS CUT USING THE FIELD PLOT (ADS) BOX DIMENSIONS AND STRATIFIED INTO LOWER (LAYER1), INTERMEDIATE (LAYER2) AND UPPER (LAYER3) CANOPY LAYERS. THE ROOT MEAN SQUARED ERROR (RMSE) AND COEFFICIENT OF DETERMINATION (R-SQUARED) VALUES FOR STEM VOLUME PREDICTION WAS EVEN DISPLAYED

Step 1 - Pre-processing of ALS data

As part of the preprocessing step, the computing of ALS point cloud was running through several modules embedded in LAsTools software (www.rapidlasso.com). Initially, the raw ALS point cloud was classified in the ground and non-ground strata using the “lasground” module, then, the points marked as outlier were filtered using “lasheight” to generate a point cloud classified and cleaned. The generated point cloud was height normalized, based on the ground surface, using “lasheight” module to derive a normalized above-ground point cloud source. The normalized above-ground point cloud was clipped based on the ADS dimension using “lasclip” module. To include the crowns of the edge trees, the areas of ADS were enlarged with a buffer of 2 m, shifting from 529 m² to 729 m². The enlarged clipped point clouds for each ADS were used as input variables in the following steps.

Step 2 - Grouping and stratifying the ADS point cloud

To investigate factors influencing the accuracy of tree detection, due to the mixed-species and multi-layered characteristics of forest stands, the ADS point clouds were split in four groups (A, B, C, and D) according to the forest stand condition (i.e. tree height variation) and the point clouds density. This step was necessary to classify different complexity levels of the forest stand in more homogeneous groups, according to the

sample probability distribution theory (Barabesi *et al.* 2012). Based on the mean values of both the average point density (APD) (Hamraz *et al.* 2017; Hamraz *et al.* 2017a) and the standard deviation of surveyed tree heights (THsd) (Wang *et al.* 2019a; Liang *et al.* 2019), four groups containing the uniform number of observations, i.e., ADS, were discriminated (Figure 3). The value adopted as a threshold for APD was fixed at 31.02 points m⁻², while for THsd the value was established at 6.879 m. Group A included the ADS that showed the lowest values of both APD and THsd; group B included ADS with lowest values of APD and highest values of THsd; group C included ADS with the highest values of both APD and THsd; group D included the ADS with highest values of APD and lowest values of THsd. The grouping process was achieved using “TreeLS” (available on GitHub, <https://github.com/tiagodc/TreeLS>) and “stats” (authors, R Core Team, and contributors worldwide) R packages.

Moreover, the four groups were ranked in four complexity categories (“highly difficult”, “moderately difficult”, “highly easy”, “moderately easy”) (Liang *et al.* 2018, 2019; Wang *et al.* 2019) to discriminate the accuracy of the detection approach within different forest structures (Figure 8).

In detail, the ADS characterized by the highest number of trees with a higher frequency of small trees (DBH ≤ 20 cm), as for example ADS of group A and D, fall in the categories “highly difficult” and “moderately difficult” respectively, though with differences in the APD values, which were 21.9 points m⁻² for ADS of “highly difficult” and 106.6 points m⁻² for ADS of the “moderately difficult”. Conversely, the ADS belonging groups B and C, characterized by the lowest number of trees with a higher frequency of large trees (DBH > 20 cm), were in the categories “moderately easy” and “highly easy”, respectively, with values of APD equal to 100.3 points m⁻² for “highly easy” category and 19.75 points m⁻² for “moderately easy” category. Therefore, the ALS and forest stand conditions preserved the structural heterogeneity between ADS, while maintaining the structural homogeneity within categories, which supports the assumption that an appropriate sample probability distribution of ADS was sampled (Barabesi *et al.* 2012).

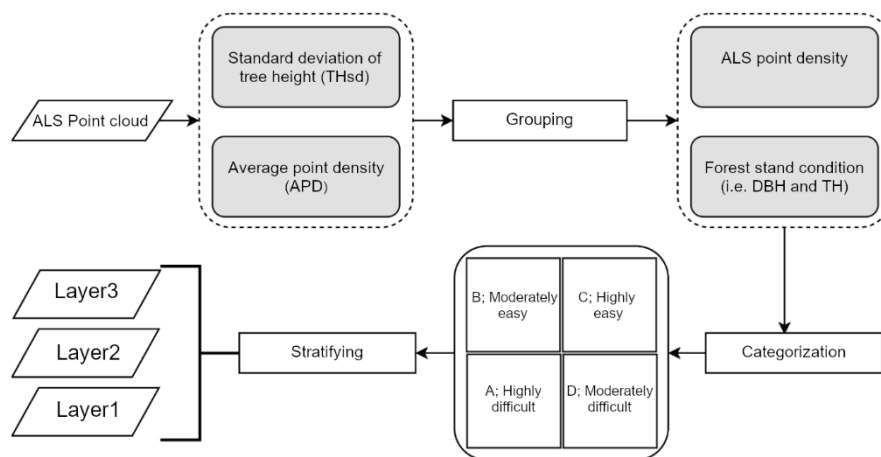


FIGURE 8 WORKFLOW OF THE PROCESSING OF THE AIRBORNE LASER SCANNING (ALS) POINT CLOUD FOR EVERY CANOPY LAYER (LAYER1, LAYER2 AND LAYER3) WITHIN EACH FIELD PLOT (ADS). THE DIAMETER AT BREAST HEIGHT (DBH) AND TREE HEIGHT (TH) WERE CONSIDERED IN THE CATEGORIZATION STEP

Thereafter, to simulate the vertical stratification of the forest stands, each ADS point cloud was split in three canopy layers, from suppressed to top canopy trees. Layer1 representing the vegetation of the suppressed trees, Layer2 representing the subdominant trees and Layer3 representing the dominant and codominant trees. The splitting procedure based on the vertical distribution of the tree heights, namely 33th (Layer1), 66th (Layer2), and 99th (Layer3) percentiles (Figure 3), was done using “lascanopy” module available on LAStools software. The resulted from tiled point clouds were used as input data for the tree detection and segmentation (step 3).

Step 3 - Tree detection and segmentation

To detect the stem position and to segment the stem and crown of every single tree, the combined use of Density-Based Spatial Clustering of Applications with Noise (DBSCAN) (Ester *et al.* 1996) and K-means, was implemented. DBSCAN is an unsupervised clustering algorithm able to discover the clusters, the noise and the outliers in a database, with poor knowledge of arbitrary shapes. Conceptually, the Density-Based clustering approach is referred to a set of points (p) belonging to a database (D); $p \in D$. The DBSCAN algorithm strives to estimate the quantity of points (p) around each point in a database (D) based on a Euclidean distance measurement called Eps-neighborhood distance. The Eps-neighborhood of each point, named $N_{Eps}(p)$, can be derived following the equation:

$$N_{Eps}(p) = \{q \in D \mid dist(p, q) \leq Eps\} \quad (eq. 2)$$

Where p and $d \in D$, $dist.$ is the distance. In density-based clustering, p is located within the Eps-neighborhood distance. Nevertheless, the size of $N_{Eps}(p)$ around each point relies on a specific minimum number of points used to form a dense region, called MinPts.

$N_{Eps}(p)$ and MinPts are mandatory thresholds to classify the point dispersion into core, border and noise points (Ester *et al.* 1996; Smits *et al.* 2012). The core point consists of a high density of points based on MinPts ($N_{Eps}(p) \geq MinPts$); the border is a point out of the core point but easy to be reachable ($p \in N_{Eps}(q)$); the noise point is an isolated point far away from the core point (Figure 9). To define the core, border and noise points, the DBSCAN algorithm plays an internal validation based on the density-reachability and density-connectivity (Figure 4) (Ester *et al.* 1996; Smits *et al.* 2012).

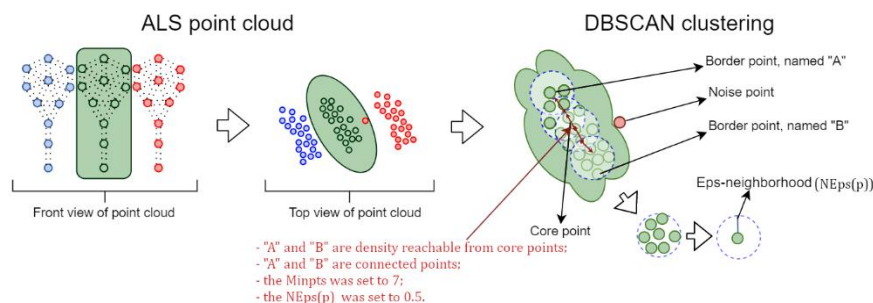


FIGURE 9 THE PROCESSING OF THE AIRBORNE LASER SCANNING (ALS) POINT CLOUD THROUGH DENSITY-BASED SPATIAL CLUSTERING OF APPLICATIONS WITH NOISE (DBSCAN) ALGORITHM. THE MINIMUM NUMBER OF POINTS (MINPTS) AND THE EPS NEIGHBORHOOD DISTANCE (NEPS(P)) THRESHOLDS WERE CONSIDERED.

K-means is an unsupervised clustering algorithm able to partition a database into K clusters for N dimensions, with high intra-class similarities, based on the concept that the K parameter has to be set (Hartigan 1975; Hartigan and Wong, 1979). The K-means equation is:

$$j = \sum_{j=1}^K \sum_{i=1}^n \left\| X_i^{(j)} - C_j \right\|^2 \quad (\text{eq. 3})$$

Where j is the K-means function, “ K ” is the number of clusters, n is the number of cases, X is a case j and C is a centroid for cluster j .

To retrieve the value of “ K ” cluster from all horizontal strata in order to run the partition of the K-means processing, the DBSCAN was applied for each horizontal stratum (i.e., Layer1, Layer2 and Layer3) over all ADS point clouds.

K-means algorithm allowed us to delineate the tree crown boundary of detected tree positions, using the K number of clusters derived by DBSCAN findings (Figure 10).

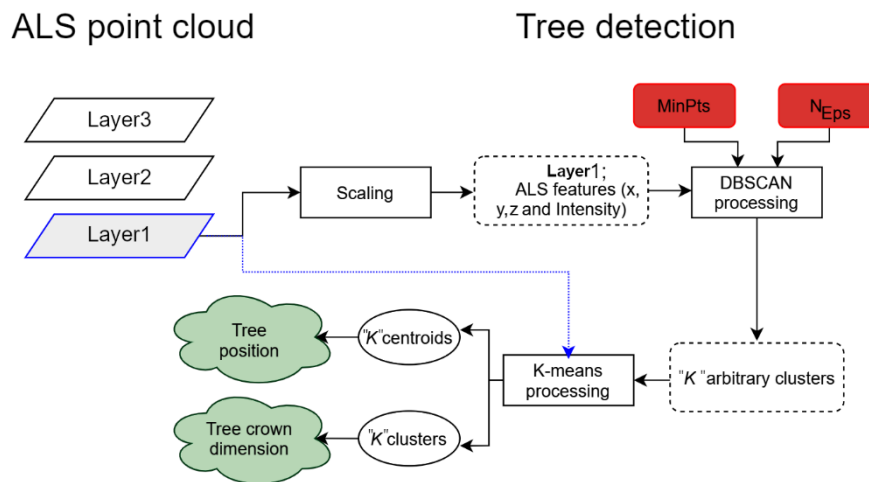


FIGURE 10 WORKFLOW OF THE PROCESSING FOR DETECTING THE TREES ACROSS THE THREE CANOPY LAYERS (I.E. LOWER LAYER: LAYER1, INTERMEDIATE LAYER: LAYER2 AND UPPER LAYER: LAYER3) FROM AIRBORNE LASER SCANNING (ALS) POINT CLOUD. THE MINIMUM NUMBER OF POINTS (MINPTS) AND EPS NEIGHBORHOOD DISTANCE (EPS) THRESHOLDS WERE USED FOR PROCESSING DENSITY-BASED SPATIAL CLUSTERING OF APPLICATIONS WITH NOISE (DBSCAN) AND K-MEANS.

Since the MinPts and NEps(p) were pre-requisites to run DBSCAN algorithm, we manually calculated these two values (Ferrara *et al.* 2018). In particular, the "MinPts" was set to 7 and the "NEps(p)" was set to 0.5 (Figure 5). The analysis was developed in R software, through the "TreeLS" package (available on GitHub, <https://github.com/tiagodc/TreeLS>), "dbscan" and the "kNNdist" function (Hahsler *et al.* 2016).

Since the "K" number of clusters was provided by DBSCAN processing (MacQueen 1967; Kandare *et al.* 2016), the number of K-means clusters was the same. Each K-means cluster was composed by the "K" centroids (tree position) and "K" clusters (tree crown dimension). To remove the noise contained in the predicted tree clusters, we used Mahalanobis distance using R packages "TreeLS", "akmeans" (Kwak 2014), "rgdal" (Bivand and Rowlingson, 2016) and "rLiDAR" (Silva *et al.* 2015). ALS metrics were extracted for each true detected tree through the "lascanopy" module implemented in LAStools software. The point cloud data for each potential stem were exported and validated in the following step.

Step 4 - Validation of predicted tree crowns

The validation accuracy of the DBSCAN and K-means results was carried out following the most used accuracy parameters in ALS detection studies (Kandare *et al.* 2014; Vastaranta *et al.* 2014; Sačkov *et al.* 2016). More precisely, the accuracy of the tree position and tree crown delineation was achieved by comparing the reference data (tree position, tree crown dimension from field survey) with the predicted data (centroid of stems, tree crowns from ALS data) through the Euclidean distance, with a tolerance value of three meters, as reference values to validate the detection accuracy. Specific accuracy parameters were:

- True-positive (TruePos; units), representing the correctly identified tree.
- False-positive (FalsePos; units) was the commission error, representing the trees that could not be associated with any surveyed tree (i.e., identified but not real).
- False-negative (FalseNeg; units) was the omission error, representing the non-segmented tree.
- Percent tree crown overlap (TREE CROWN OVERLAP; %), as the parameter indicating the difference between the isolated reference and predicted crown segment.
- Distance between the predicted centroid of the crown segment and the centroid of the reference crown (Euclidean distance; m). Euclidean distance was applied to determine the distance between the predicted and reference centroid crown segments.
- Detection Rate (DR; %), reporting the relationship between the TruePos and the reference stem.
- Time for tree detection (Time for TD; sec), reporting the time-consuming in analyzing each sampled area of 729 m².

Step 5 - Prediction of forest inventory variables

To predict different forest inventory variables, for trees that were previously identified, the Random Forests algorithm was applied. Random Forests algorithm allowed us to achieve regression tree classification based

on decision trees (Breiman 2001), as being widely used to handle a high number of factors and for reducing the overfitting (Shi *et al.* 2018).

The Random Forests parameters used for the prediction were (a) “Ntree”, the number of decision trees to be used during the prediction phase; (b) “Mtry”, the number of input variables for splitting at each tree nodes; and (c) “nodesize”, the minimum size of terminal nodes (Belgiu and Drăgu, 2016).

In this study, the three forest inventory variables (i.e., DBH, TH, and VOL) for each layer (i.e., Layer1, Layer2 and Layer3, and Layer1-Layer3) within each category (i.e., highly difficult, moderately difficult, highly easy, moderately easy) were predicted using the ALS metrics (Top-nine) of its corresponding TruePos. The whole predicted models amount to 48: 16 out of 48 corresponding to DBH, 16 out of 48 corresponding to TH and 16 out of 48 corresponding to VOL. Furthermore, to investigate the performance of models using the ALS metrics (Top-nine) given to the total TruePos, we calculated the forest inventory variables (i.e., DBH, TH, and VOL) using the merged information of categories; the whole predicted models were three, one per forest inventory variable.

The Random Forests models were implemented using the randomForest package in R (Liaw and Wiener, 2002). The setting of the Random Forests algorithm was implemented by "Ntree" as 1000, "Mtry" as 3-4, and node size as 5. The validation of these models was developed by the coefficient of determination (R-squared; 0-1) and root mean square error (RMSE; cm, m, m³) for the number of trees examined (N^otrees; units), using the “stats” (authors, R Core Team and contributors worldwide) and “usdm” (Naimi 2017) R packages.

Moreover, the CS was predicted using as input the VOL from ALS data for each canopy layer. Validation was done by comparing the predicted vs. observed CS amount for each ADS.

2.2.5. Results

Bosco Pennataro is characterized by a heterogeneous forest structure; among the ADS, the number of trees ranged between 453 and 3698 trees ha⁻¹, the mean DBH ranged between 9.9 cm and 26.9 cm, the mean TH ranged between 8.2 m and 23.1 m, and the stem volume ranged from 183 m³ ha⁻¹ (carbon amount = 51.1 tons ha⁻¹) to 633.9 m³ ha⁻¹ (carbon amount = 177 tons ha⁻¹). The heterogeneity of the forest stand, due to both vertical stratification and DBH variability, as well as the stand density, impacted the point density and spacing of ALS point clouds that varied from 12.13 points m² to 292.9 points m² (Table 6)

2.2.5.1. ADS groups and ALS point clouds layers

The clusterization of the surveyed ADS in four distinct groups allowed us to assess forest inventory variables in this mixed-species and multilayered Mediterranean forest correctly. Though the number of ADS for each group was similar (Figure 11), ADS showed a varying pattern across the complex forest structure (Table 6). Stand density was high among ADS of difficult categories, ranged between 1724 trees ha⁻¹ and 1542 trees ha⁻¹. Moreover, these ADS presented high standard deviation values (985 trees ha⁻¹ and 840 trees ha⁻¹), compared to those of easy categories (between 339 trees ha⁻¹ and 262 trees ha⁻¹).

Additionally, the easy categories were characterized by a great number of big trees compared to the difficult categories and, as a consequence, by high values of assessed forest inventory variables, i.e., DBH, TH, VOL, and CS.

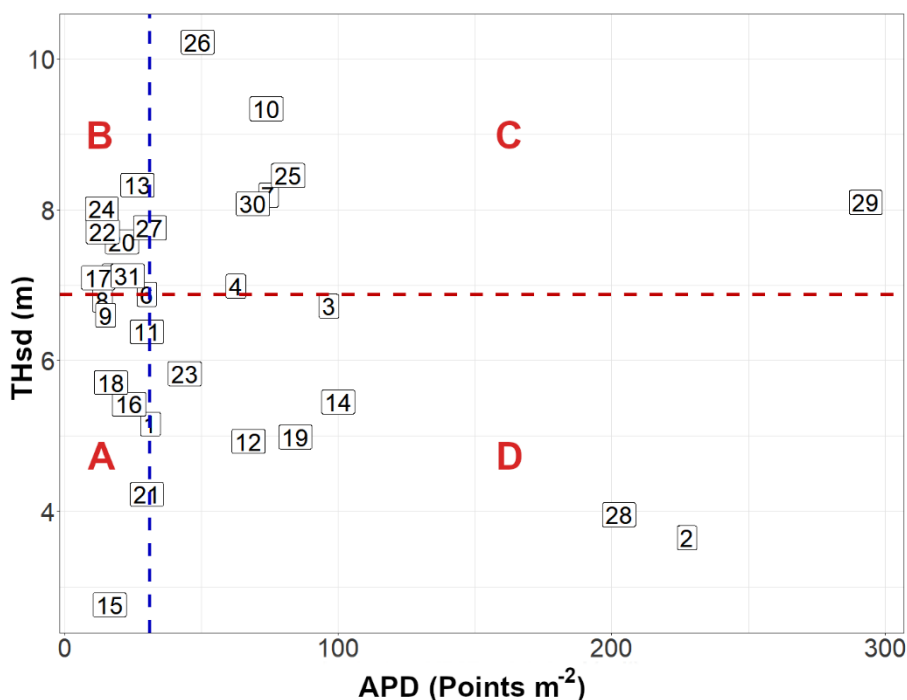


FIGURE 11 GRAPHICAL DISTRIBUTION OF THE FIELD PLOTS (ADS) ACCORDING TO THE AVERAGE POINT DENSITY (APD; POINTS M⁻²) AND THE STANDARD DEVIATION OF TREE HEIGHT (THSD; M) FOR EACH CATEGORY FROM A TO D GROUPS (A, HIGHLY DIFFICULT; B, MODERATELY EASY; C, HIGHLY EASY; D, MODERATELY DIFFICULT).

Results showed a greater variability among ADS of the difficult categories rather than among ADS of the easy categories, allowing us to state that the heterogeneity of forest structure impacted on the detection of single trees.

TABLE 6 SUMMARY OF FOREST STAND CHARACTERISTICS OF AIRBORNE LASER SCANNING (ALS) AND FOREST INVENTORY DATA PER EACH FIELD PLOT (ADS) AND COMPLEXITY CATEGORIES. THE AVERAGE POINT DENSITY (APD; POINTS M⁻²), AVERAGE POINT SPACING (APS; M), DIAMETER AT BREAST HEIGHT (DBH; CM) AND TREE HEIGHT (TH; M) WERE ESTIMATED PER ADS. THE STEM VOLUME (VOL; M³) AND CARBON STOCK (CS; TONS) WERE ESTIMATED PER HECTARE (HA.). THE NUMBER OF TREES (N°TREES; UNITS) WERE CALCULATED PER ADS AND HA. THE MEAN (*1) AND SUM (*2) AND STANDARD DEVIATION (*3) VALUES WERE SHOWN.

Category	ADS	ALS data		Forest inventory data					
		ADS				Ha			
		APD*1 (Points m ²)	APS*1 (m)	N°trees*2 (units)	DBH*1 (cm)	TH ¹ (m)	N°trees*2 (units)	VOL*2 (m ³)	CS*2 (ton)
High	6	30	0.2	32	18	17.3	604	220.3	61.5
	8	13.8	0.3	52	15.7	13.7	981	307.8	86

	9	14.9	0.3	47	19	16.3	887	353.4	98.7
	11	30	0.2	86	15.2	12.9	1623	341.7	95.4
	15	16.3	0.3	196	9.9	8.2	3698	183	51.1
	16	23.4	0.2	121	12.1	9.4	2302	272.1	76
	18	16.8	0.2	95	12.9	9.7	1792	277.3	77.4
	21	30	0.2	101	14	11.4	1906	329	91.9
	(* ¹)	21.9	0.2	91	14.6	12.4	1724	285.6	79.8
	(* ²)			730			13792		
	(* ³)	7.3	0.1	52	3	3.3	985	60	16.8
Moderately easy	5	17.2	0.2	34	20	18.2	642	325.9	91
	13	26.5	0.2	83	16.5	14.6	1566	488.3	136.4
	17	12.1	0.3	58	18.1	13.3	1094	485.9	135.7
	20	20.8	0.2	31	21.1	16	585	308.8	86.3
	22	13.9	0.3	70	17.1	15.6	1321	633.9	177
	24	13.5	0.3	42	18.6	16.2	792	406.4	113.5
	27	31	0.2	54	13.5	11.1	1019	408.1	114
	31	23	0.2	63	17.1	14	1189	435.2	121.5
	(* ¹)	19.75	0.24	54	17.75	14.88	1026	436.56	121.93
	(* ²)			435			8208		
(* ³)	6.78	0.05	18	2.32	2.14	339	103.06	28.77	
Highly easy	4	62.5	0.1	49	17.8	15.3	925	357.2	99.8
	7	74.4	0.1	36	20	13.3	679	450.2	125.7
	10	73.7	0.1	37	23.1	16.9	698	528.5	147.6
	25	81.6	0.1	32	20.7	21.3	623	430.4	120.2
	26	48.6	0.1	24	25.8	23.1	453	477.1	133.3
	29	292.9	0.1	60	16.1	15.4	1132	400	111.7
	30	68.7	0.1	60	16.5	13.9	1132	413.6	115.5
	(* ¹)	100.3	0.1	43	20	17	806	436.7	122
	(* ²)			298			5642		
	(* ³)	85.6	0	14	3.6	3.8	262	55.5	15.5
Moderately difficult	1	31.3	0.2	33	20.5	18.5	623	249	69.6
	2	227.6	0.1	120	12.8	9.8	2264	270.8	75.6
	3	96.6	0.1	35	26.9	17.1	660	537.5	150.1
	12	67.1	0.1	140	10.6	9.1	2642	286.8	80.1
	14	99.9	0.1	91	11.6	9.8	1717	344.7	96.3
	19	84.3	0.1	50	20.1	14.3	943	295.4	82.5
	23	43.8	0.2	53	18	12.3	1000	390.7	109.1
	28	202.7	0.1	132	10.7	10.8	2491	220.2	61.5
	(* ¹)	106.66	0.13	82	16.4	12.71	1542	324.39	90.6
	(* ²)			654			12340		
(* ³)	71.41	0.05	45	5.92	3.57	840	101.34	28.29	

The number of trees across the three canopy layers was rather similar, from Layer1 to Layer3, with a relatively low presence of trees in the Layer2 (Figure 12). Therefore, the discrimination of trees was similar also across different ADS.

However, the distinction of crowns across the three canopy layers was facilitated in ADS of easy compared to difficult categories. For this reason, the poor presence of stems, more accentuated in ADS of the “slightly easy” and “moderately easy” categories, was a contributing factor that enabled the discrimination of single trees (Figure 12B and Figure 12C); while the high values of stand density created an overlapping effect

among tree crowns, which slightly hindered the detection of trees, particularly for the intermediate layers. (Figure 12A and Figure 12D).

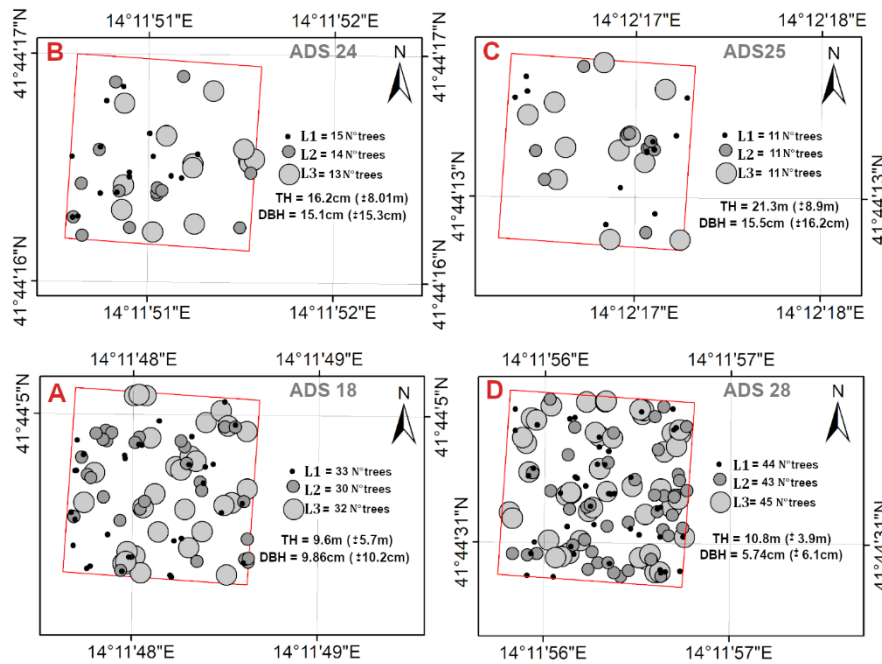


FIGURE 12 FOUR REPRESENTATIVE MAPS OF THE FOUR DIFFERENT AIRBORNE LASER SCANNING (ALS) POINT CLOUD COMBINATIONS (ONE PER CATEGORY). THE RED SQUARE SHOWED THE FIELD PLOT (ADS) BORDER; THE NUMBER OF TREES (N° TREES; UNITS) WAS SHOWN FOR EVERY CANOPY LAYER (I.E. LOWER LAYER: LAYER1, INTERMEDIATE LAYER: LAYER2 AND UPPER LAYER: LAYER3); THE TREE HEIGHT (TH; M) AND THE DIAMETER AT BREAST HEIGHT (DBH; CM) WERE EXPRESSED IN AVERAGE AND THE STANDARD DEVIATION (SD; \pm) VALUES; THE TOP LETTERS REPORT THE CATEGORY LEVEL (I.E. A, HIGHLY DIFFICULT; B, MODERATELY EASY; C, HIGHLY EASY; D, MODERATELY DIFFICULT).

2.2.5.2. Tree detection

We detected 952 out of 2117 reference trees, reaching an average detection rate of 48 % (Table 7), with a moderate uniformity/similarity across the three layers ($SD = \pm 12.5$). Our tree detection approach was more sensitive to the omission error, 1165 out of 2117 reference trees, than to the commission error, 795 out of 2117 reference trees. Better results in terms of the detection rate were obtained in ADS belonging to the ADS of groups B and C (easy categories) rather than in those of groups A and D (difficult categories). The detection rate was 36 % ($SD = \pm 7.3$) for ADS of the “highly difficult” category, identifying 261 out of 730 trees. The detection rate was 49 % ($SD = \pm 19.2$) for ADS of “moderately difficult” category, identifying 215 out of 435 trees. The detection rate for ADS of “moderately easy” category was 43 % ($SD = \pm 7.8$), identifying 282 out of 654 trees. The detection rate for ADS of “highly easy” category reached 65 % ($SD = \pm 7.0$), identifying 194 out of 298 trees.

The detection rate values were more accurate for trees of the Layer2 (54 %, $SD = \pm 13.7$) than for trees of the Layer1 (42 %, $SD = \pm 7.8$) and Layer3 (49 %, $SD = \pm 15.4$).

The detection of trees in ADS with the lowest point density, corresponding to the “highly difficult” and “moderately easy” categories, was affected by the occlusion effects from subdominant, codominant and dominant to suppressed trees; a better performance was obtained for trees of the Layer2 and Layer3. Whereas, an opposite pattern was observed in ADS with a higher point density, “highly easy” and “moderately difficult” categories. Hence, the point density influenced the occlusion effects from large to small tree crown dimension in tree detection, regardless the forest structure.

The highest value of the commission error was found for the ADS of the “highly easy” category, which was 123 %, (367 FalsePos), while ranging between 21 % and 25 % in the remaining three categories

Similarly, the highest value of the omission error was found in the “highly easy” category, which was 135 %, (104 FalseNeg), while the omission error for the other three categories ranged between 51 % and 64 %. The best and worst compromise between commission and omission errors were found in ADS of “moderately easy” (106 and 220 out of 435 surveyed stems) and “highly easy” (367 and 104 out of 298 surveyed stems), respectively.

The sensitivity variation of our algorithm for commission and omission errors was rather small among the three canopy layers, which ranged from 44 % to 58 % for FalsePos and from 46 % to 58 % for FalseNeg.

TABLE 7 TREE DETECTION RESULTS. NUMBER OF STEMS OBSERVED FROM REFERENCE DATA (TR; UNITS) AND NUMBER OF STEMS PREDICTED FROM ALS DATA (TALS; UNITS), TRUE POSITIVE (TRUEPOS; UNITS), FALSE POSITIVE (FALSEPOS; UNITS), FALSE NEGATIVE (FALSENEG; UNITS) AND DETECTION RATE (DR; %) FOR LOWER (LAYER1), INTERMEDIATE (LAYER2) AND UPPER (LAYER3) CANOPY LAYERS.

Tree detection results							
Categories	Canopy layers	TR (units)	Tree detection				DR (%)
			T _{ALS} (units)	TruePos (units)	FalsePos (units)	FalseNeg (units)	
Highly difficult	Layer1	245	124	69	55	176	28
	Layer2	237	176	101	75	136	43
	Layer3	248	147	91	56	157	37
	Sum	730	447	261	186	469	
	Mean & SD (±)						36 (7.3)
Moderately easy	Layer1	144	54	40	14	104	28
	Layer2	141	120	78	42	63	55
	Layer3	150	147	97	50	53	65
	Sum	435	321	215	106	220	
	Mean & SD (±)						49 (19.2)
Highly easy	Layer1	99	178	63	115	36	64
	Layer2	96	213	70	143	26	73
	Layer3	103	170	61	109	42	59
	Sum	298	561	194	367	104	
	Mean & SD (±)						65 (7.0)
Moderately difficult	Layer1	218	166	108	58	110	50
	Layer2	213	147	97	50	116	46
	Layer3	223	105	77	28	146	35
	Sum	654	418	282	136	372	
	Mean & SD (±)						43 (7.8)
Layer1	Sum	706	522	280	242	426	
	Mean & SD (±)						42 (17.5)

Layer2	<i>Sum</i>	687	656	346	310	341
	<i>Mean & SD (±)</i>					
Layer3	<i>Sum</i>	724	569	326	243	398
	<i>Mean & SD (±)</i>					
Total	<i>Sum</i>	2117	1747	952	795	1165
	<i>Mean & SD (±)</i>					

The estimation of the crown position displayed similar values for all four categories, ranging between 1.73 m and 2.55 m (Figure 13II). The similarities were also observed among the three canopy layers, particularly for ADS of the “highly easy” category, within which the most homogeneous values were observed. On the contrary, small differences were observed between Layer3 and Layer1 or layer2 in the remaining categories. Although the observed crown dimension was not completely covered by the predicted tree crown dimension, the average overlap value was 57%; this was moderately consistent across ADS (SD = ± 11) (Figure 13III), within which Layer2 was the most accurate.

Time required in detecting the trees, using combined unsupervised algorithms, was faster in the ADS with the lowest (21.9 points m⁻²) point density in comparison with those with the highest (19.7 points m⁻²) (Figure 13IV).

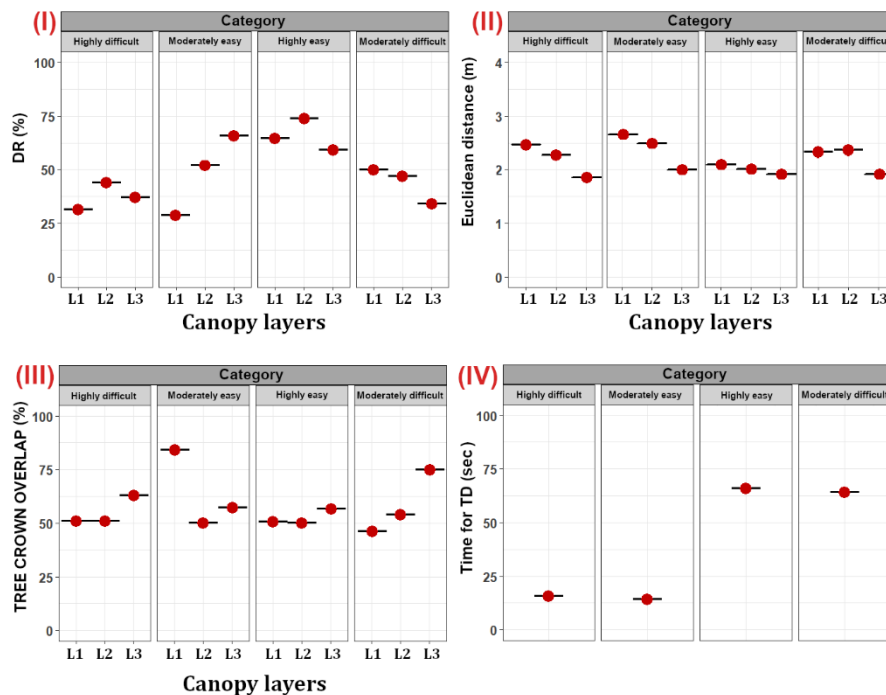


FIGURE 13 COMPARISON BETWEEN PREDICTED VS. OBSERVED VALUES OF I) DETECTION RATE (DR; %), II) EUCLIDEAN DISTANCE (M), III) TREE CROWN OVERLAP (%), AND IV) TIME FOR TREE DETECTION (TIME FOR TD; SEC) FOR EACH CANOPY LAYER (LAYER1 “L1”, LAYER2 “L2”, AND LAYER3 “L3”) AND FOR EVERY CATEGORY (HIGHLY DIFFICULT, MODERATELY EASY, HIGHLY EASY AND MODERATELY DIFFICULT). THE AVERAGE VALUES OF THE TIME CONSUMING FOR DETECTING TREE (TIME FOR TD, SEC) BELONGING TO EACH PLOT WAS DISPLAYED FOR EACH CATEGORY.

2.2.5.3. Forest inventory variables

Comparing the predicted vs. observed data from correctly detected trees, corresponding to 952 trees, we found significant values of the coefficient of determination and the RMSE for DBH (0.92; 4.03 cm), TH (0.95; 1.33 m) and VOL (0.82; 0.31 m³), respectively (Figure 14).

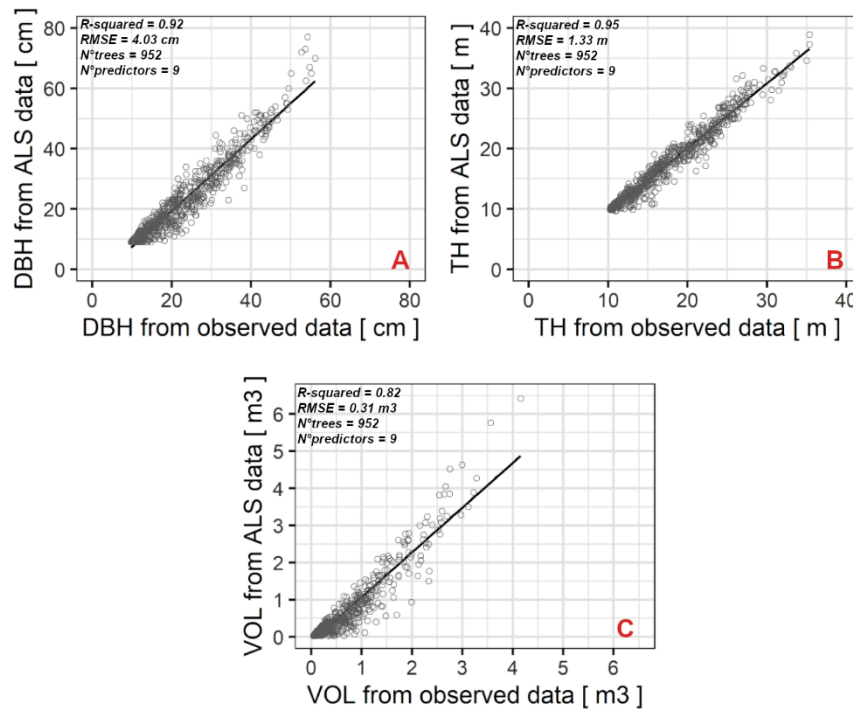


FIGURE 14 PREDICTED VALUES VS. OBSERVED FOREST INVENTORY VARIABLES. THE BOX A) SHOWS DIAMETER AT BREAST HEIGHT (DBH, CM); BOX B) SHOWS TREE HEIGHT (TH; M) AND BOX C) DISPLAYS STEM VOLUME (VOL, M³). THE NUMBER OF TREES (N°TREES; UNITS), NUMBER OF PREDICTORS (N°PREDICTORS; UNITS), COEFFICIENT OF DETERMINATION (R-SQUARED; 0-1) AND ROOT MEAN SQUARED ERROR (RMSE; CM, M AND M³) WERE EVEN REPORTED.

Despite the different quantities of trees analyzed (TruePos), slight differences in terms of coefficient of determination between predicted vs. observed across categories were observed. However, the categories were less accurate for DBH (N° trees = 261 and 215; R-squared = 0.9) belonging to ADS of the “highly difficult” and “moderately easy” categories; whereas, for TH (N° trees = 215; R-squared = 0.93) and VOL (N° trees = 215; R-squared = 0.89), this was the case for the ADS belonging to the “moderately easy” category. Therefore, the categories with smaller point densities (in absolute terms) were slightly less accurate (Table 8).

TABLE 8 SUMMARY STATISTICS OF THE FOREST INVENTORY VARIABLES ESTIMATED WITH THE RANDOM FORESTS ALGORITHM BY USING TOP-NINE METRICS FOR DIAMETER AT BREAST HEIGHT (DBH; CM), TREE HEIGHT (TH; M), AND STEM VOLUME (VOL; M³). THE NUMBER OF TREES (N°TREES; UNITS), COEFFICIENT OF DETERMINATION (R-SQUARED; 0-1) AND ROOT MEAN SQUARED ERROR (RMSE; CM, M AND M³) WERE DISPLAYED. THE OUTCOMES WERE DISPLAYED FOR ALL FOUR CATEGORIES (HIGHLY DIFFICULT,

MODERATELY EASY, HIGHLY EASY AND MODERATELY DIFFICULT), WHICH WAS FURTHER DIVIDED BY LOWER (LAYER1), INTERMEDIATE (LAYER2) AND UPPER (LAYER3) CANOPY LAYERS.

		Linear regression											
Category	Statistic measurements	DBH (cm)				TH (m)				VOL (m ³)			
		Layer1	Layer2	layer3	Layer1 - Layer3	Layer1	lyer2	Layer3	Layer1 - Layer3	Layer1	Layer2	Layer3	Layer1 - Layer3
Highly difficult	N°trees	69	101	91	261	69	101	91	261	69	101	91	261
	R-squared	0.91	0.91	0.89	0.9	0.92	0.91	0.91	0.95	0.93	0.87	0.91	0.9
	RMSE	0.9	2.25	3.8	3.62	1.05	1.01	1.03	1.14	0.01	0.05	0.27	0.2
Moderately easy	N°trees	40	78	97	215	40	78	97	215	40	78	97	215
	R-squared	0.91	0.89	0.88	0.9	0.92	0.92	0.89	0.93	0.78	0.89	0.87	0.89
	RMSE	1.44	2.43	4.25	4.59	0.7	0.93	1.1	1.38	0.02	0.08	0.37	0.35
Highly easy	N°trees	63	70	61	194	63	70	61	194	63	70	61	194
	R-squared	0.82	0.89	0.91	0.91	0.82	0.93	0.9	0.95	0.8	0.89	0.88	0.9
	RMSE	1.77	3.26	4.28	4.63	1.87	1.07	1.25	1.56	0.04	0.12	0.56	0.4
Moderately difficult	N°trees	108	97	77	282	108	97	77	282	108	97	77	282
	R-squared	0.88	0.9	0.89	0.91	0.86	0.97	0.89	0.95	0.81	0.86	0.91	0.9
	RMSE	1.38	3.3	5.15	4.08	1.16	0.61	1.52	1.21	0.02	0.15	0.38	0.24

We observed that the best and worst accuracies were found in the ADS of the “moderately easy” and “highly difficult” categories, based on the fitted prediction for stem volume (RMSE = 0.14 % and bias = 0.1 %) and carbon stock (RMSE = 1.48 % and bias = 1.5 %) variables (Table 9). However, we note that the “moderately difficult” category offered better performances than the “highly easy” category. Therefore, ADS with a higher number of trees with a higher frequency of small trees were less affected by the performance of the models in terms of bias and RMSE values. Moreover, the bias and RMSE in the case of “moderately easy” and “moderately difficult” categories suggested that the ADS with a higher point density associated and higher share of trees with the predominance of large trees might solve issues associated with uncertainties. It is worth noting that the prediction of stem volume was weakly related to the tree detection accuracy.

TABLE 9 COMPARISON BETWEEN PREDICTED AND OBSERVED VALUES OF STEM VOLUME (VOL; M3) AND CARBON STOCK (CS; TONS) DERIVED FROM AIRBORNE LASER SCANNING (ALS) METRICS. THESE VALUES WERE FURTHERLY DISPLAYED BY EACH CATEGORY (HIGHLY DIFFICULT, MODERATELY EASY, HIGHLY EASY AND MODERATELY DIFFICULT) AND BY EACH FIELD PLOT (ADS). THE NUMBER OF TREES (N°TREES; UNITS) WAS CALCULATED PER ADS AND HA-1. ABSOLUTE (M3 AND TONS) AND PERCENT (%) VALUES OF BIAS AND ROOT MEANS SQUARED ERROR (RMSE) WERE EVEN DISPLAYED.

Stem volume and carbon stock prediction					
N°trees	VOL (m ³ ha ⁻¹)	CS (tons ha ⁻¹)	VOL	CS	VOL and CS

Category	ID	ADS	Ha ⁻¹	Observed	Predicted	Observed	Predicted	Bias (m ³ ha ⁻¹)	Bias (tons.ha ⁻¹)	Bias %*(Relative bias %)	RMSE (%)	
Highly difficult	6	29	547	185.5	227	51.8	63.4					
	8	22	415	253	245.2	70.7	68.5					
	9	34	641	304.1	305.1	84.9	85.2					
	11	28	528	143.5	146.9	40.1	41					
	15	40	755	34.2	32.4	9.6	9.1					
	16	40	755	105.1	102.6	29.4	28.7					
	18	31	585	119.8	131.8	33.5	36.8					
	21	37	698	175.1	148.9	48.9	41.6					
	<i>Sum</i>	261	4924									
	<i>Mean</i>				<i>165</i>	<i>167.5</i>	<i>46.1</i>	<i>46.8</i>				
<i>Accuracy</i>								-2.4	-0.7	1.5(-0.14)	1.48	
Moderately easy	5	30	566	322	325.2	89.9	90.8					
	13	38	717	277.2	310.8	77.4	86.8					
	17	20	377	306.2	284.1	85.5	79.3					
	20	17	321	200.1	212.3	55.9	59.3					
	22	43	811	471.7	493.3	131.7	137.8					
	24	19	358	304	308.4	84.9	86.1					
	27	16	302	256.4	227.5	71.6	63.5					
	31	32	604	338.9	318.3	94.7	88.9					
	<i>Sum</i>	215	4056									
	<i>Mean</i>				<i>309.6</i>	<i>310</i>	<i>86.5</i>	<i>86.6</i>				
<i>Accuracy</i>								-0.4	-0.1	0.1(0.46)	0.14	
Highly easy	4	34	641	192.4	198.7	53.7	55.5					
	7	20	377	417.1	324	116.5	90.5					
	10	28	528	388.7	372.2	108.6	104					
	25	23	434	392.3	420.6	109.6	117.5					
	26	21	396	441.6	462.3	123.3	129.1					
	29	44	830	304.3	371.1	85	103.7					
	30	24	453	158.4	164.6	44.2	46					
	<i>Sum</i>	194	3659									
<i>Mean</i>				<i>327.8</i>	<i>330.5</i>	<i>91.6</i>	<i>92.3</i>					
<i>Accuracy</i>								-2.7	-0.8	0.8(-0.46)	0.83	
Moderately difficult	1	22	415	172.7	188.2	48.2	52.6					
	2	40	755	29.7	44.3	8.3	12.4					
	3	25	472	406.5	407.6	113.6	113.8					
	12	32	604	68.5	58	19.1	16.2					
	14	25	472	170.6	177.9	47.7	49.7					
	19	41	774	265.5	253.9	74.2	70.9					
	23	22	415	159.7	147.6	44.6	41.2					
	28	75	1415	119.8	111.4	33.5	31.1					
	<i>Sum</i>	282	5322									
	<i>Mean</i>				<i>174.1</i>	<i>173.6</i>	<i>48.6</i>	<i>48.5</i>				
<i>Accuracy</i>								0.5	0.1	-0.3(-0.88)	0.3	

2.2.6. Discussion

2.2.6.1. Tree detection

Results revealed that the joint use of DBSCAN and K-means allowed detecting nearly half of the trees identified through ALS data in the studied multi-layered and mixed-species Mediterranean mountain forest. Enhanced detection accuracy was obtained in forest ADS with higher heterogeneity of tree height, regardless of stand density. This approach may improve monitoring of forest dynamics related to tree growth and surveying of tree mortality due to forest disturbance. Indeed, mixed-species and multi-layered forests in Mediterranean mountains are complex systems and the assessment of their 3D full structure is of importance for reducing uncertainties in the collection of reference data. In particular, consistent ALS monitoring of forest changes may allow deriving new indicators of CSF related to vertical and horizontal forest attributes (Bodwitch *et al.* 2020; Santopuoli *et al.* 2020b).

Though the detection was challenging for trees of the lower layer, results obtained here were somewhat encouraging in comparison with those reported by other authors. For example, Sačkov *et al.* (2016) showed accuracy values from 24 % (all trees) to 36 % (trees higher than 16 m) and 48 % (trees higher than 21 m). Similarly, Duncanson *et al.* (2014) reported values from 21 % for suppressed trees to 70 % for dominant trees, and Hamraz *et al.* (2017) observed that the accuracy of tree detection decreased from dominant to suppressed trees and highlighted that a dense point cloud was required for a satisfactory detection. The LiDAR point clouds used here had an average of 60 points m², ranging between 21 to 106 points m². Nevertheless, the choice to split the point clouds into three canopy layers allowed us to improve the overall detection accuracy, supporting the use of ALS data for monitoring forest inventory variables and smart forestry indicators at a large scale. This aspect is crucial to support forest managers with a monitoring tool for well-timed and spatial-explicit forest inventory data, and appears promising for implementing smart management strategies to reduce operating costs (Torresan *et al.* 2021)

Our study revealed that the point density, the forest stand conditions (Hamraz *et al.* 2017; Kandare *et al.* 2016; Williams 2019), and the site-specific parameters, e.g., species composition and forest structure (Sackov *et al.* 2016; Liang *et al.* 2019) impacted the identification of trees, as well as the detection rate, and commission and omission errors. Therefore, the density of ALS point clouds would represent one important limitation of unsupervised techniques for detecting single trees, which failed for values below the threshold of 30 points m². In particular, the detection accuracy was further worsened in ADS of this Mediterranean mountain mixed-species and multi-layered forest with high values of stand density (1542 trees ha⁻¹). Beyond the stand density, the presence of large trees was advantageous in the identification processes using our unsupervised approach. Therefore, the detection was more accurate for those ADS with higher average values of DBH and TH, namely veteran trees (Santopuoli *et al.* 2020a).

It is important to note that, though the detection accuracy was higher for trees belonging to the intermediate and upper layers, a better compromise between omission and commission errors was found for the lower layer (Table 2). This apparent contradiction was probably related to the higher stand density inducing commission errors but avoiding omission errors, due to the clustering approach and the Mahalanobis

filtering the outliers. Dense forest stands may hinder the correct separation between nearby trees (Kaarinen *et al.* 2012; Dalponte *et al.* 2015; Wang *et al.* 2016). This means that the ALS point density and the forest structure may play a complementary role in identifying and segmenting trees using point cloud sources for multi-layered as well as for two-layered mixed-species forests (Torresan *et al.* 2020).

The detection performance was improved by the evaluation of the crown radius, which allowed us to obtain good results (ranging between 1.73 m and 2.55 m), somehow better than those reported in the literature. For example, 2 m was the value reported by Shao *et al.* (2018), 2.5 m by Balsi *et al.* (2018), 3.5 m by Mongus and Žalik (2015), and 5 m by Sačkov *et al.* (2016). Contrary to what was revealed by these authors, for which the values of Euclidean distance decreased from the upper to the lower layers, we demonstrated that the detection accuracy could be relatively constant across the three canopy layers. Tree crown overlap ranged between 47.26 % and 82.51 % (more stable values were obtained in ADS of the “highly easy” and “highly difficult” categories), supporting the hypothesis that an optimum performance for identifying and segmenting trees could be expected for multi-layered mixed-species forests of this type.

2.2.6.2. *Forest inventory variables*

The approach implemented in this study allowed us to predict three forest inventory variables, namely DBH, TH, and VOL, reaching the accuracy in coefficient of determination of about 0.92 for DBH, 0.95 for TH, and 0.82 for VOL. Though the feasibility of the prediction approach was tested in four complexity levels, there were no substantial differences in the prediction accuracy among all categories. Such versatility of the Random Forests approach increased the prediction performance of forest inventory variables and was proved promising for collecting CSF indicators. It is worth noting that ALS data analyzed by means of canopy layers might describe thoroughly the forest inventory variables for trees within every canopy layer, especially for trees of intermediate and lower layers.

The performance of VOL models was more accurate using the information of whole TruePos (Layer1-Layer3) compared to the TruePos of the upper layer (Layer3), based on the RMSE measurements found in all four categories. More accurate prediction of DBH and VOL was observed in the “highly difficult” category, whereas, for TH the fitted prediction was observed in all four categories. The effect of the quantity of TruePos on the performance of models was mitigated by the bootstrap approach of the Random Forests algorithm, as supported by almost all RMSE values across the three canopy layers.

As expected, the performance of models based on RMSE values declined from Layer1 to Layer3 for DBH and VOL; however, this pattern was moderately smoothed for TH. This means that the estimation of DBH and VOL for intermediate and dominant trees was a challenging task, when the stratification approach was applied; whereas, the prediction for TH was rather accurate for all three canopy layers.

Here, the stand structural heterogeneity and the ALS point density represented the most hindering factors for the prediction, though results were satisfactory and higher than those reported in similar studies. Indeed, the accuracy obtained for the prediction of DBH in this study was higher than in Sačkov *et al.* (2016, 2019), who reported R-squared equal to 0.71 for mixed-species forest stands, and 0.78 for deciduous and 0.72 for

coniferous forests. Yet, for the prediction of VOL, other studies reported lower values of accuracy (Sačkov *et al.* 2016; Alberti *et al.* 2013).

The prediction accuracy for carbon stock was more accurate in the ADS with high ALS point density (“moderately difficult”; bias = -0.1) and low ALS point density (“moderately easy”; bias = -0.1), but with a more homogeneous forest structure. Therefore, in the prediction of forest inventory variables, a low ALS density would represent an issue in areas with relatively homogeneous forest structures. Results obtained for the stem volume (where input data to derive the carbon stock was ranged between 0.89 and 0.90 of R-squared) were in line with those observed by other authors: Popescu (2007) showed higher R-squared values for above-ground biomass in mature stands of loblolly pine, ranging between 0.88 and 0.93, whereas Allouis *et al.* (2012) reported higher R-squared values of above-ground biomass in individual black pine trees, ranging between 0.87 and 0.91.

Accurate predictions of carbon stock could be expected in all the four categories considered here. However, the bias in prediction (minimum bias = -0.3 % and maximum bias = 1.5 %) could be associated with other factors, e.g., understory vegetation, standing deadwood, terrain slope, site aspect, and species richness (Næsset and Gobakken, 2008; Yu *et al.* 2010; White *et al.* 2014).

Overall, the accuracy of tree detection and carbon stock accurateness assessment resulted to be more sensitive to point density than heterogeneity of forest structure (Table 7; Table 9). This means that further efforts focused on improving the quality of points will be beneficial to better exploit the potential of tested algorithms. We found many weak points during the ALS processing. For example, the ADS point clouds characterized by lower, altered and irregularly-spaced densities were hard to be processed by DBSCAN algorithm; fixed values of minPts and Eps-neighborhood became disadvantageous for identifying the trees in ADS from difficult categories; the ADS with dense points are time-consuming. These weakness points suggested that DBSCAN algorithm was sensitive to the quality of point cloud and fixed minPts and Eps-neighborhood values (Ahmad and Dang, 2015). Nevertheless, careful consideration in operational activities could be beneficial to overcome part of these issues, especially before the collection phase: 1) forest canopy structure (changing from leaf-on to leaf-off) (Shao *et al.* 2018); 2) flight strips (changing from 0 % to more than 50 % of overlapped flight strips) (Liang *et al.* 2019) and 3) ALS sensor (changing from 3 echoes to 4-15 echoes) (Kandare *et al.* 2016; Hamraz *et al.* 2017). Since our ultra-light vehicle flew at an altitude of 100 m above ground level, we hypothesized this flying height was good enough. In conclusion, the quality of the point cloud may vary depending on the ALS sensor returns, operational aspect and forest structure, therefore, the potential of our algorithm can also be affected.

2.2.7. Conclusion

This study aimed to improve the use of ALS data for the prediction of forest inventory variables in mixed-species and multi-layered forests of Mediterranean mountain environments. Such a development might represent an important advance for the estimation of forest characteristics and the collection of CSF indicators, as well as to monitor the dynamics of these complex forest ecosystems over time.

The most important limitation faced in this study was the ALS point density. Using very low point density, the detection of single trees was challenging, as for those stands with less than 30 points m^{-2} . ADS primarily composed of big trees would be less problematic. In this latter case, we obtained more than 65 % of detection accuracy, regardless of the canopy layers. Nevertheless, to detect trees in forest areas where small trees are abundant, a denser point cloud would be required. The stratification approach adopted in this study, minimized the negative impacts due to the low point density and the heterogeneity of forest structure, stressing the usefulness of ALS data for assessing forest inventory variables and climate-smart forestry indicators. However, the heterogeneity of forest structure could be an important hindering factor when using ALS in the understory layer, especially in forest areas with poor ALS densities (>30 points m^{-2}). The occlusion effect of ALS point in tree detection could be caused by highly overlapped crowns, hindering the detection of trees. It is worth noting that the unsupervised technique implemented in this study allowed us to obtain satisfactory accuracy for a forest ecosystem characterized by heterogeneous canopy profiles and big tree sizes.

The application of unsupervised algorithms for detecting single trees in a mixed-species and multi-layered Mediterranean forest through LiDAR data was proved feasible in support of actively measuring and monitoring of complex mountain forest ecosystems. The stratification of ALS point clouds might represent a valid alternative to simulate the vertical distribution of trees in stands with heterogeneous structures, allowing forest operators to detect and monitor a large number of trees.

2.2.8. Postface

In this study, a stepwise approach composed of two unsupervised algorithms and a machine learning algorithm was tested for carbon stock assessment at a single tree level. The detection approach using two unsupervised algorithms was tested for the first time in forests, especially in mixed-species and multi-layered forests were presented. This unsupervised method allowed us to detect nearby the half of reference trees and a considerable part of them belonging to trees from lowest and intermediate strata. A point of strength was the capability to identify the trees without previous knowledge of the number of trees. This is important because this information may prove to be expensive and time-consuming. We noted that our stepwise approach better works in the forest with high tree height variation, especially in forests covered by a point density of more than 30 points m^{-2} . It is worth noting that most points belonged to the first return and that the collection of ALS data was done in leaf-on condition. The carbon stock estimation resulted to be accurate for all forest field plots, in fact, the bias varied from -0.3 % to 1.5 %. Further studies testing our stepwise approach can be useful to deeper understand the potential of detecting trees with the better condition of ALS collection.

2.3. A stepwise approach for deriving timber assortment of trees from Terrestrial Laser Scanning data

2.3.1. Preface

Providing accurate and reliable approaches for extracting the timber assortment information from tree stands became crucial in the forestry sector, especially in managed forests (Marchetti *et al.* 2018; SoEF 2020). Recent studies using TLS have reported a great capacity for reconstructing the trees of the upper, intermediate and lower canopy layers, however, its accuracy was conditioned by the operational (i.e. sampling design), the technical (i.e. automatically), the weather condition (i.e. wind) and the forest structure (i.e. stem density) (Dassot *et al.* 2011; Liang *et al.* 2018; Wan *et al.* 2019). In 2004, the cylinder-fitting approach was introduced as a useful tool for measuring the diameter at breast height of trunks on TLS point cloud; while now this approach proved to be efficient for reconstructing the trunk section on TLS point cloud, namely stem curve. However, to date, the use of the cylinder-fitting approach for retrieving quantitative and qualitative information of the logs belonging to the trunk section on TLS point cloud has not been investigated. In this context, this study introduces a stepwise approach for timber assortment assessments using TLS point cloud in mixed tree-species and multi-layered Mediterranean forests

2.3.2. Abstract

Forest ecosystems represent an important source of income for landowners and at the same time an important source of ecosystem services for society. Quantitative and qualitative information about timber assortments is particularly important to support sustainable forest management, representing a crucial prerequisite for active forest management. To date, the most accurate methods for assessing the timber assortments available within forest stands are destructive, and the development of an effective method for deriving these estimates on standing trees is highly needed. This is particularly more evident for mixed forests, which are often subject to the conflict between conservation and productive functions.

This study aims to introduce a stepwise approach for timber assortment estimation and classification using TLS data. The proposed approach is consisting of four steps: a) timber-leave discrimination, b) tree detection, c) stem reconstruction, and d) timber assortment estimation and classification. The study was carried out in a mixed tree-species and multi-layered Mediterranean forest, observing 178 trees of twelve different species, wherein 70 out of 178 were large trees, with a diameter at breast height higher than 20 cm.

Results indicate that the discrimination between timber and leaves reached 0.98 for accuracy using Random Forest algorithm. The overall detection rate was 84.40 % ($SD_{\pm} = 4.7$ %), particularly, all trees with a diameter at breast height higher than 30 cm were correctly identified. Among the detected trees, the most frequent species detected were *A. lobelii*, *S. torminalis*, *F. excelsior*, *Q. cerris*, *A. campestre* and *F. sylvatica* (higher than 84.3%) tree species. 47 out of 70 large trees from observed data were correctly reconstructed. These 47 trees provided 179 merchantable logs and 40 non-merchantable logs from observed data. We quantified 134 out of 179 merchantable logs and 34 out of 40 non-merchantable logs. The 179 merchantable

logs were classified into 15 assortment types, which were even categorized into five assortment groups (i.e. saw-log plus, saw-log, pulpwood, other industrial roundwood and fuelwood). The most accurate timber assortment assessment assessed were saw-log and other industrial roundwood and someone else of the other assortment types. The abovementioned results support the feasibility of this stepwise approach for calculating the timber assortment of standing trees, hence it can be used to valorise the timber resource from forests, even those with a high richness of species and structural heterogeneity.

Keywords: Timber quality, stepwise procedure, 3D modelling, point cloud, multi-layered, mixed-forest, LiDAR.

2.3.3. Introduction

At Pan-European level, the roundwood from forests, since offering socio-economic and environmental benefits to forest owners and stakeholders, is an essential source for the forestry production chain. Roundwood products can be subdivided into industrial roundwood (wood in the rough) and woodfuel sources (wood used for energy purpose) (FOREST EUROPE 2015). Over the years, roundwood production in Europe has been growing, reaching a maximum of almost 550 million m³ annually (SoEF 2020), making the Europe's forests one of the main producers of industrial roundwood (Proskurina *et al.* 2019); however, many European countries faced difficulties in monitoring the woodfuel production (FOREST EUROPE 2015). The reasons might be found in the lower market value of woodfuel respect to the roundwood, the dimension of logs which usually are lower than logs used for industrial roundwood, such as coppice forests or the branches and other minute parts of large trunks, the lack of management plans or equivalent for small and private forests, due to the fragmented forest ownership.

Therefore, innovative and accurate methods for roundwood assessment are necessary to promote the faithful classification of the assortments from standing trees in order to support sustainable forest management through the revalorisation of timber from forest resources (Gazull and Gautier, 2014).

Several approaches have been developed worldwide, however, accurate timber assortment measurements have to consider many tree factors as stem tapering, stem curve and stem diameter (Kankare *et al.* 2014), to accurately reconstruct the stem profile and the stem volume (West 2009). Moreover, stem curve allows reconstructing the diameters at different heights through consecutive cylinders' models (Lassasenaho 1982). Other approaches focused on the assessment of the stem profile factor, as (Togni 2017) proposed some sets of mathematical equations to classify the timber assortments from both coniferous and broadleaves stems focusing on the description of geometry wood defects (i.e. stem straightness, and stem tapering); (Lassasenaho 1982) which proposed the use of polynomial functions to derive the stem curves and stem volume from coniferous and broadleaves stems (i.e. pine, spruce and birch); (Tabacchi *et al.* 2011) that proposed the usage of equations for accurate estimation of the aboveground biomass estimation of Italian tree species.

Hence, accurate measurement of standing trees became a crucial pre-requisite to manage forests (La Marca and Notarangelo, 2009), to define and schedule forestry interventions and harvesting activities (Nosenzo

2007), in light of sustainable forest management (SFM) criteria, as well as to maintain or enhance the provision of multiple forest ecosystem services. Over the last decades, accurate measurement of the forest timber resources was obtained using LiDAR (Light Detection and Ranging) data at national (i.e. national forest inventories) and regional levels (McRoberts and Tomppo, 2007; Chirici *et al.* 2020). This means that LiDAR technique can be an effective tool in monitoring the forest timber resources supporting forest inventory and the implementation of SFM at different geographical levels.

Among LiDAR data, in the last two decades, Terrestrial Laser Scanning (TLS) has gained more attention among scientists due to the high accuracy reconstruction of stem architecture from 3D point cloud data (Dassot *et al.* 2011; Disney *et al.* 2019). The high versatility of TLS point cloud, and the development of new approaches and models to automatically elaborate TLS point cloud (Liang *et al.* 2018), fostered the use of TLS for several issues in distinct topics, as for example, in geomorphology to reduce the uncertainties associated with the georeferencing of TLS data (Walicka *et al.* 2019); in forest biodiversity to assess the relationships between forest structure and habitat quality (Michel *et al.* 2008), to assess the accuracy for the reconstruction of tree characteristics (Bournez *et al.* 2017; Othmani *et al.* 2016; Torresan *et al.* 2018); the classification of tree species composition (Othmani *et al.* 2013) (Lin and Herold, 2016) and the discrimination between timber and leave (Ferrara *et al.* 2018; Wang *et al.* 2017).

As reported in many studies, assessing tree characteristics, as the stem profile, through TLS data is easier in forest plantations rather than in natural forests (Liang and Hyypä, 2011; Kankare *et al.* 2013; Liang *et al.* 2018). Robust cylinder-fitting methods are the most common approach to assess the stem profile (Pfeifer *et al.* 2004; Liang *et al.* 2018). Such approach is based on the subdivision of point cloud in several horizontal slices, and by the recognition of the stem position into the horizontal slices and then detecting the closest cylinders from each stem (Lukács *et al.* 1997; Liang *et al.* 2014; Wang *et al.* 2016b; Pitkänen *et al.* 2019a). As supported by an international study comparing about 18 algorithms, to validate the stem curve performance, fitted cylinder prediction (consecutive cylinders along the stem from the ground; $0.50\text{m} \pm 0.015\text{m}$), the curve length ratio (CLR) and percent of the tree height covered (PHC) are often tested (Liang *et al.* 2018).

Otherwise, to systematically compare the accuracy, such as completeness and correctness of stem reconstruction, (Liang *et al.* 2018) tested the accuracy of 18 algorithms for tree characteristic evaluation using two scan modes (single-scan and multi-scans) and three structural complexity levels as driving factors. Particularly, the advantage of the tested TLS algorithms lies in the capacity to automatically model the TLS data ($\geq 80\%$), in the versatility for processing huge point clouds through recent approaches (i.e. raster-based, voxel-based and point-based). By contrast, the limitations were associated with the characterization of small trees and the occlusion effects associated with the poor density from single-scan (Wan *et al.* 2019). Moreover, it was demonstrated that several models focused to reconstruct stem profile e.g., TreeQSM (Raumonen *et al.* 2013), Simple-Tree (Hackenberg *et al.* 2015), TreeLS (de Conto *et al.* 2017), 3Dforest (Trochta *et al.* 2017), Computree (Torresan *et al.* 2018), L-Architect (Côté *et al.* 2018), treeSeg (Burt *et al.* 2019), were affected by forest stand factors (i.e. stem density, dense small branches, the

presence of leaves) and methods for aligning and assembling the TLS scans were missing. However, a modular program, named OPALS (Orientation and Processing of Airborne Laser Scanning data), offers the possibility to overcome part of the abovementioned issues enabling the accurate description of the stem profile (<https://opals.geo.tuwien.ac.at>) (Wang *et al.* 2016, 2016a).

Recently, few studies have tried to retrieve the timber assortment using TLS data, even less were focused on natural forests (Murphy *et al.* 2010; Mengesha *et al.* 2015; Sun *et al.* 2016; Stovall *et al.* 2018; Saarinen *et al.* 2019; Chianucci *et al.* 2020). This means that a procedure suitable to overcome the hindering factors in order to optimize the use of TLS data for timber assortment estimation became crucial to implement SFM strategies.

This study aims to introduce a stepwise procedure to quantify and classify the timber assortments from standing trees using TLS data in mixed-forest and multi-layered forests. Stem detection and reconstruction were achieved through a robust cylinder-fitting approach. The stepwise approach follows four steps: a) timber-leave discrimination, b) tree detection, c) stem reconstruction, and d) timber assortment estimation and classification.

2.3.4. Materials and Methods:

2.3.4.1. Study area

The study was carried out in Bosco Pennataro (Figure 15 A-B), a Mediterranean mixed tree species forest located in Molise region (Central of Italy, 41°42' N, 14° 12' E). Bosco Pennataro is characterized by the high tree species richness and heterogeneous forest structure (Fig 15C-D). The tree species composition of Bosco Pennataro includes *Q. cerris* (40%), *F. sylvatica* (21%), *A. obtusatum* Mill. (9.6%), and other broadleaves tree species (Santopuoli *et al.* 2019). Due to the large number of broadleaves species with predominance of *Q. cerris*, the forest community of Bosco Pennataro is classified as oak–hornbeam forest type (Barbati *et al.* 2014). Bosco Pennataro, besides belonging to the network of Natura 2000 sites, is worldwide recognized as a core area of the Man and Biosphere (MaB) reserve of Collemeluccio-Montedimezzo Alto Molise. Due to the recognized ecological importance, the forest was historically managed for productive purposes as an even-aged forest with natural regeneration, while in the last 50 years the harvesting activities were very limited and focused to prevent fire. As a result, currently, the forest is characterized by high structural heterogeneity, both vertical and horizontal as well as a high value of biodiversity (i.e. microhabitat) (Santopuoli *et al.* 2019a).

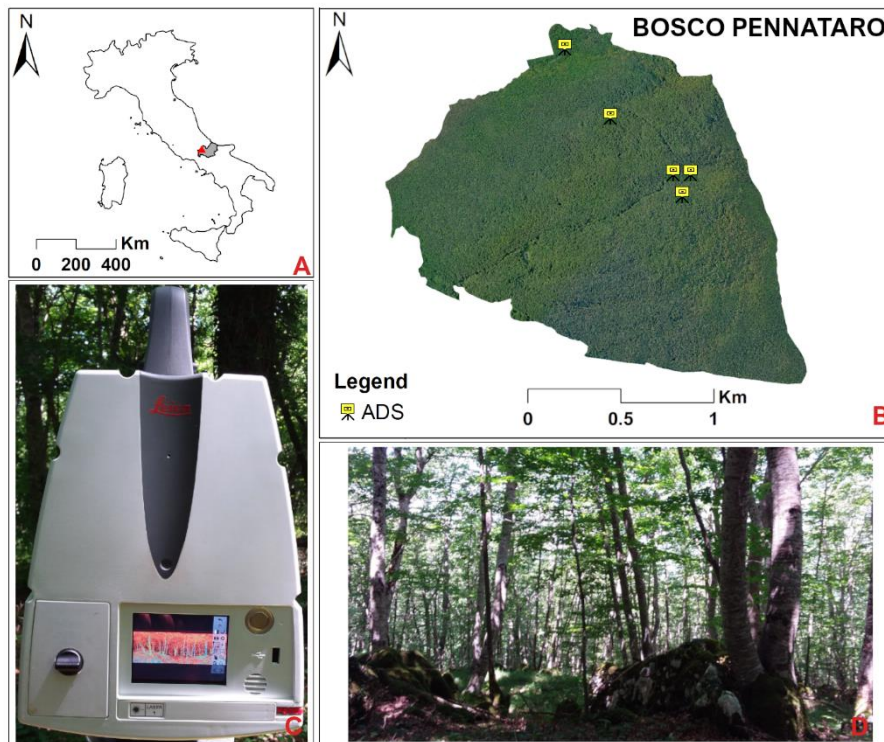


FIGURE 15 A) SHOWS THE LOCATION OF THE FIELD PLOTS (ADS) IN RELATION TO ITALY LANDSCAPE; B) DISPLAYS THE ADS IN RELATION WITH BOSCO PENNATARO; C) SHOWS ONE PICTURE OF THE TERRESTRIAL LASER SCANNING (TLS) DEVICE AND D) SHOWS A PICTURE OF BOSCO PENNATARO.

2.3.4.2. Ground truth field data

The sampling was carried out in 2016 for five square plots (hereafter ADS) of 529 m² (23m * 23m) within Bosco Pennataro. All trees with a diameter at breast height (DBH) ≥ 2.5 cm were measured through the Field-Map tool (<https://www.fieldmap.cz/>). The sampled forest-related characteristics surveyed were: DBH, tree height (TH), the height of the first attached branch or branch union (TH1), stem position, canopy projection area (CPA), tree species and tree vitality. Moreover, the stem volume (TSv) was calculated through allometric equations implemented for Italian tree species in the National Forest Inventory (Tabacchi *et al.* 2011).

2.3.4.3. Terrestrial Laser Scanning data

Terrestrial Laser Scanning (TLS) data were collected using a Leica ScanStation P30/40 device (<https://leica-geosystems.com/it-it/>) in July 2018. The laser scanning system for Leica ScanStation P30/40 is an Ultra-high-speed time-of-flight enhanced by Waveform Digitising (WFD) technology. Leica ScanStation P30/40 is a Laser 3D scanner suitable for collecting 1 million points per second for a wide range of up to 270 m. The horizontal and vertical field-of-view of Leica ScanStation P30/40 was 360° and 290°, respectively. The distance measurement accuracy for the objects was equal to ± 2 mm. A total of 178 single trees, divided into five ADS, were scanned using the Leica ScanStation P30/40 (Figure 15B-C). The average of TLS multiple-scans collected among the ADS was 9 (SD = ± 1.4 TLS multiple-scans).

The TLS multiple-scans were co-registered through an automatic Leica Geosystems processing (<https://leica-geosystems.com/>). This co-registration was supported by the recorded geographic coordinates taken into random positions derived from the GPS Trimble GeoXT mounted on a Hurrigan Antenna. To minimize the distance between two sets of points (one set = one TLS single-scan), in an iterative way, caused by a systematic residual error from the co-registration, a rigid transformation was computed on it through an Iterative closest points (ICP) algorithm (Besl and McKay, 1992). To reach this transformation, the TLS multiple-scans were further pre-processed, particularly, the alignment, co-registration and assembling were running followed batch scripts in OPALS modular program (Glira *et al.* 2015; Fuad *et al.* 2018), especially the OpalsICP batch scripts (<https://opals.geo.tuwien.ac.at/html/stable/ModuleICP.html>). As a result, five merged TLS multiple-scans were generated. Finally, to optimize the point cloud depicting of trees located at the edge of each ADS, the ADS point cloud from each ADS was enlarged from 529 m² to 729 m². To reach this, the merged TLS multiple-scans were imported, clipped and exported using a box dimension equal to 729 m² (27*27m) through opalsImport, opalsAlgebra and opalsExport batch scrips running in OPALS modular program (<https://opals.geo.tuwien.ac.at/html/stable/index.html>). The five TLS multiple-scans including geographic coordinates (i.e. x, y, z) and intensity feature data were used as the input source in the following subsequent steps.

2.3.4.4. *Ground truth TLS data*

Some tree measurements were manually measured using TLS data through CloudCompare software. The useful trunk section used for retrieving tree measurements was ranged between the ground (0.50 m, named THbase) and the first attached branch or branch union (TH1). The trees considered for retrieving tree measurements were trees with more than 20 cm of DBH (Nosenzo 2007; Jukka *et al.* 2010; Togni 2017; Liang *et al.* 2018). Moreover, in order to optimize the valorisation of the trunk section (in commercial terms), each trunk section was divided into merchantable logs (2.5 m ≤ length of log ≤ 3 m) and non-merchantable logs (2.5 m < length of log) (Nosenzo 2007; Jukka *et al.* 2010) (Figure 16). Based on such statements, several measurements from merchantable and non-merchantable logs were manually extracted using CloudCompare software (<http://www.danielgm.net/cc/>). These measurements were: the maximum and minimum end diameters (Dmax and Dmin) and the length of log (L). Along with these measurements, the log volume (hereafter TTV.log) was estimated using the Dmax, Dmin and L through the Smalian formula, following the equation (1).

$$TTV_{log} = \frac{(Dmin^2 + Dmax^2)}{8} * \pi * L \quad (\text{eq. 1})$$

Where:

TTV.log – log volume, (m³);

Dmax – maximum diameter of ends log, (m);

Dmin –minimum diameter of ends log, (m);

L –length of log, (m); and

$\pi - 3.1416$.

Furthermore, the trunk volume (hereafter TTV. Trunk) was estimated by summing the TTV.log belonging to the same tree, which includes the merchantable and non-merchantable logs. In addition, to optimize the characterization of the high-quality logs, merchantable logs, the straightness (STR) and tapering (TAP) characteristics of these logs were calculated through the equations (2 and 3) (Figure 16) (Togni 2017) . The stem diameter information (Dmin and Dmax) was scaled from m to cm.

$$STR = \frac{h}{L} \quad (\text{eq. 2})$$

$$TAP = \frac{(Dmax - Dmin)}{L} \quad (\text{eq. 3})$$

Where:

STR– straightness of logs, (cm/m);

TAP– tapering of logs, (cm/m);

L – length of log, (m);

Dmax – maximum diameter of ends log, (cm);

Dmin – minimum diameter of ends log, (cm); and

h – perpendicular distance (90°) between the highest convex curve and the straight line between small (Dmin) and large (Dmax) ends logs.

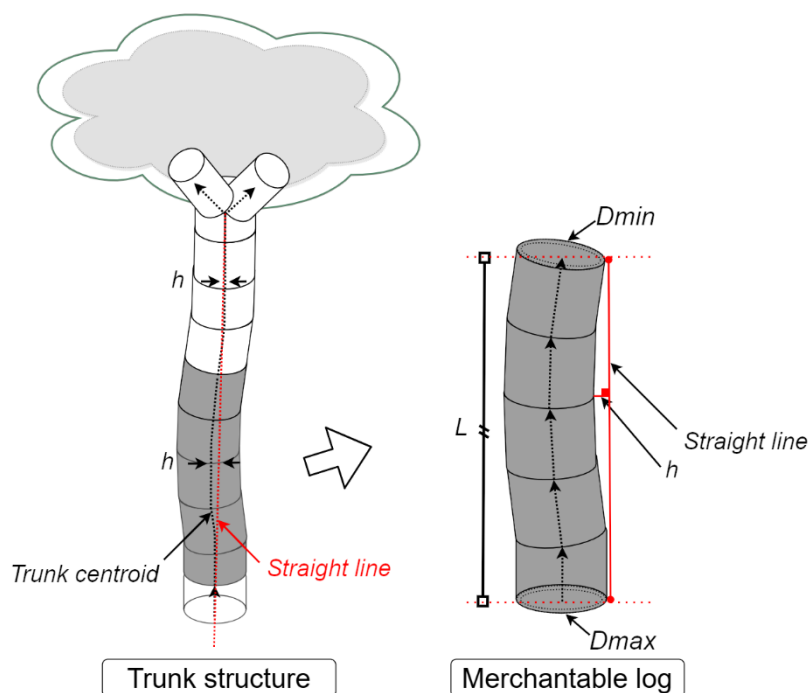


FIGURE 16 GRAPHIC REPRESENTATION OF A TRUNK STRUCTURE AND MERCHANTABLE LOG FOR STRAIGHTNESS AND TAPERING ESTIMATION. THE “H” INDICATES THE HIGHEST CONVEX CURVE, THE MINIMUM AND MAXIMUM DIAMETER OF THE LOG ENDS (D_{MIN} AND D_{MAX}, RESPECTIVELY) WERE DISPLAYED.

Lastly, to classify the merchantable logs into several types of assortments, based on their geometric defects (Togni, 2017). Before classifying the logs, we generated fifteen types of assortments using the STR and D_{min} thresholds described by Togni (2017) (Table 10). In detail, 3 of them representing the type “A”, 3 of them representing the type “B” and so on. Then, every merchantable log was classified into one out of 15 assortment types.

TABLE 10 TIMBER ASSORTMENT CHARACTERISTICS BASED ON GEOMETRY WOOD DEFECTS. THE STRAIGHTNESS (STR) AND DIAMETER MINIMUM OF ENDS LOGS (D_{MIN}) WERE DISPLAYED. “X” REPRESENTS THE VALUES NEEDED FOR STR AND D_{MIN} CHARACTERISTICS.

Timber assortment							
ID	Timber assortment	Abbreviation	Advanced forest-related				
			STR		D _{min}		
			Type	Thresholds (cm/m)	Type	Thresholds (cm/m)	
1	Saw-log plus	1	A+	A	$x \leq 2$	Higher	$0.30 \leq x$
2		A0	A	$x \leq 2$	Medium	$0.2 < x \leq 0.30$	
3		A-	A	$x \leq 2$	Lower	$x < 0.20$	
4	Saw-log	1	B+	B	$2 < x \leq 3.4$	Higher	$0.30 \leq x$
5		B0	B	$2 < x \leq 3.4$	Medium	$0.2 < x \leq 0.30$	
6		B-	B	$2 < x \leq 3.4$	Lower	$x < 0.20$	
7	Pulpwood	1	C+	C	$3.4 < x \leq 5$	Higher	$0.30 \leq x$
8		C0	C	$3.4 < x \leq 5$	Medium	$0.2 < x \leq 0.30$	
9		C-	C	$3.4 < x \leq 5$	Lower	$x < 0.20$	
10	Other industrial roundwood	1	D+	D	$5 < x \leq 6.6$	Higher	$0.30 \leq x$
11		D0	D	$5 < x \leq 6.6$	Medium	$0.2 < x \leq 0.30$	
12		D-	D	$5 < x \leq 6.6$	Lower	$x < 0.20$	

13		1	Fuelwood+	Fuelwood	$6.6 < x$	Higher	$0.30 \leq x$
14	Fuelwood	2	Fuelwood0	Fuelwood	$6.6 < x$	Medium	$0.2 < x \leq 0.30$
15		3	Fuelwood-	Fuelwood	$6.6 < x$	Lower	$x < 0.20$

2.3.4.5. *TLS analysis*

In this study, a stepwise approach for deriving the timber assortments of trees using TLS data was implemented. The four steps are a) timber-leaves discrimination; b) tree detection; c) stem reconstruction and d) timber assortment estimation (Figure 17)

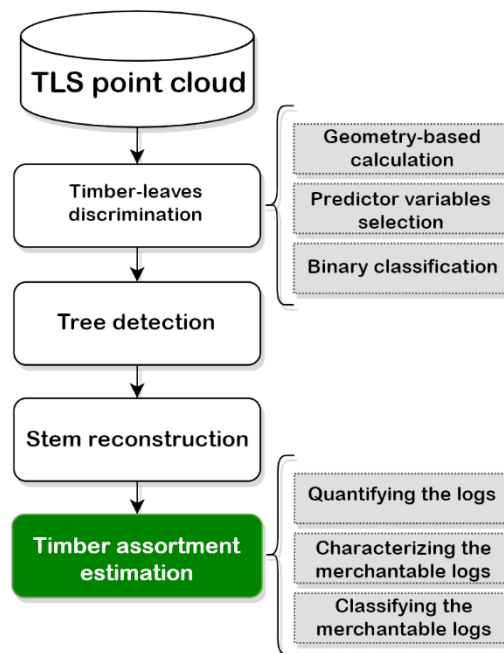


FIGURE 17 METHODOLOGICAL APPROACH FOR TIMBER ASSORTMENT ESTIMATION USING TERRESTRIAL LASER SCANNING (TLS) DATA. GREY RECTANGLES INDICATE THE SUB-STEPS.

Step 1- Timber-wood point clouds discrimination

To discriminate the timber from leaf point clouds, a binary classification of the TLS point clouds was implemented through Random Forest (RF) algorithm. The binary classification processing consists of three sub-steps: a) geometry-based calculation, b) predictor variables selection and c) binary classification (Figure 18).

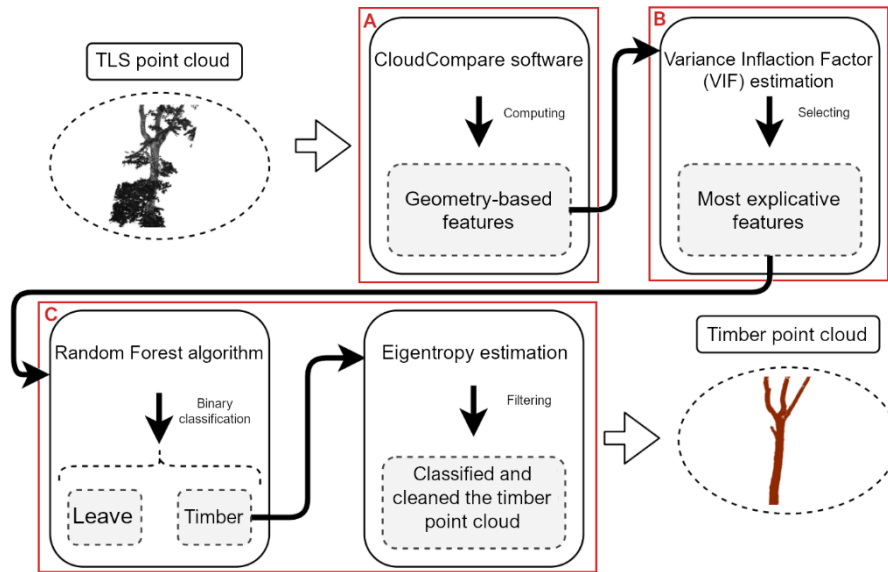


FIGURE 18 AN OVERVIEW OF THE SUB-STEPS FOR TIMBER-LEAVES DISCRIMINATION. THE INPUT TLS DATA (TOP LEFT), THE GEOMETRY-BASED CALCULATION (A), PREDICTOR VARIABLES SELECTION (B) AND TIMBER-LEAVES DISCRIMINATION (C) AND OUTPUT TIMBER POINT CLOUD (BOTTOM RIGHT).

Sub-step 1.1. Geometry-based calculation

To optimize the extraction of information from the point clouds an optimal local neighbourhood (hereafter L_n) was searched using a tool named “compute geometric features” embedded in CloudCompare open source software (<http://www.danielgm.net/cc/>). The “ L_n ” values allowed us to characterize the local surface and local point density variation within the point cloud, optimizing and facilitating the description of the point clouds (Hackel *et al.* 2016). The “compute geometric features” is a tool embedded in CloudCompare software. It allowed us to detect the contours of several surface orientations using specific “ L_n ” values, commonly called geometry-based features (Weinmann *et al.* 2014, 2015; Hackel *et al.* 2016).

To find the optimal “ L_n ” values in every ADS point cloud, 10 % of each ADS point cloud was computed using four distinct “ L_n ” values (0.03 m, 0.05 m, 0.07 m and 0.09 m) through “compute geometric features” tool. The value 0.07 m proved to be effective to characterize 92% of point clouds for each ADS (Belton and Lichti, 2006).

Geometry-based extraction: Subsequently, eighteen geometry-based features (i.e. roughness, mean curvature, Gaussian curvature, Gaussian normal change rate, number of neighbors, surface density, volume density, sum of eigenvalues, omnivariance, eigentropy, anisotropy, planarity, linearity, first “PCA1” and second principal component “PCA2”, surface variation, sphericity and verticality) were automatically generated, using the 0.07 m of “ L_n ” in “compute geometric features” CloudCompare tool (Hackel *et al.* 2016; Abu Alasal *et al.* 2014). The geometry-based and point cloud information were used in the subsequent sub-step.

Sub-step 1.2. Predictor variables selection

To reduce the quantity of geometry-based information without losing its contribution into models, the most explicative geometry-based features, named predictor variables, were selected through the variance inflation factor (VIF) score. The VIF score can find the correlation, collinearity and multicollinearity among variables, increasing significantly the contribution of each feature. Several studies highlighted that significant predictor variables were obtained customizing the VIF score as 5 of value tolerance (Neter *et al.* 1996; Zuur *et al.* 2010). Based on the results of previous studies, the VIF values higher than 5 derived from the geometry-based features were discarded. The VIF score quantification was running through the “vifstep” function implemented in usdm R package (Naimi 2015). The generated data including eight geometry-based features (i.e. anisotropy, sum of eigenvalues, Gaussian curvature, mean curvature, PCA2, roughness, verticality, volume density) and the five TLS point clouds were used as input data in the subsequent sub-step.

Sub-step 1.3. Binary classification

To classify the 3D point cloud as timber and leave classes, a binary classification was running through RF algorithm, because it was faster, easy-to-use and more accurate than other machine learning approaches in classifying the point cloud (Wang *et al.* 2017). RF is an algorithm able to build multiple decision trees from randomly input training data for accurate classification and regression (Breiman 2001).

In this study, RF classifies the point cloud into timber and leaf labels using eight geometry-based features through a package embedded in R, named Weighted Subspace Random Forest for Classification (wsrf), (Geiß *et al.* 2015; Zhao *et al.* 2017). The parameters set out for binary classification through wsrf were "Ntree" as 2500, "Mtry" as 3-4, and node size as 5.

Lastly, to remove the noise points, we implemented a filtering approach using a geometry-based, named eigenentropy. We used eigenentropy geometry-based values for filtering processing because it was assessed useful for better characterizing the surface noise on point clouds (Weinmann *et al.* 2015). This filtering approach considers the extreme values of eigenentropy geometry-based, ranging between 0.03 (25th percentile) and 0.80 (75th percentile), as noise points (Weinmann *et al.* 2014, 2015; Hackel *et al.* 2016). In our study, to predict the eigentropy values of the point clouds, we followed the same processes described in Sub-step 2.1. (Geometry-based calculation); while to remove the extreme eigenentropy geometry-based values, we imported, removed and exported the point clouds in R software (no packages are required).

To validate the classification accuracy, we assessed the sensitivity, the specificity and the accuracy measurements, as (Wang *et al.* 2017). The sensitivity represents the percentage of point clouds correctly identified (true positive), the specificity represents the percentage of point clouds correctly excluded (true negative) and the accuracy represents the proportion of true positive values. These statistics measurements were computed through pROC R package (Robin *et al.* 2020).

Step 2- Tree detection

This step aimed to find the potential stem position and to derive the DBH of these using the timber TLS point cloud through a raster-based approach embedded in OPALS modular program. OPALS is a powerful modular program consisting of several modules (derived from grouped packages). It is capable to process several types of LiDAR data, from airborne and terrestrial platforms, across several topics, as for example, forestry topics, named opalsForest (<https://opals.geo.tuwien.ac.at>). Inside opalsForest package, there are many algorithms for calculating the stem diameter distribution, tree height, tree crown gaps, and others forest variables (https://opals.geo.tuwien.ac.at/html/stable/pkg_opalsForest.html).

In this study, some modules from several packages (i.e. opalsForest) were used for detecting and estimating the trees using TLS point cloud. Prior detecting the tree position, the TLS grids, named Digital Terrain Model (DTM; 0.05 m) and Digital Surface Model (DSM; 0.20 m), were generated for each ADS using opalsDSM and opalsGrid OPALS modules. The normalization of the point cloud data, using both TLS grids, was running through opalsCell OPALS module. For each ADS, a thick horizontal slice between 1m and 2m above the ground was extracted from the normalized point cloud. Each thick horizontal slice was divided into 9 thin horizontal slices. These thin horizontal slices were further cut in several voxel-based (0.01m³). The statistic information (i.e. sum, mean and maximum) from the points contained in each voxel-based was achieved by running opalsAlgebra OPALS module to generate several zones by each horizontal slice. The several zones from the horizontal slices were used as input data for detecting and estimating the stem position and DBH using a least-squared cylinder-fitting approach implemented in opalsDBH OPALS module (<https://opals.geo.tuwien.ac.at/html/nightly/ModuleDBH.html>). To run the tree detection and DBH estimation, the TLS data were analysed by several OPALS modules e.g., opalsImport, opalsCell, opalsAlgebra and opalsExport, and the final product was composed by tree position and predicted DBH, exported as .txt and .LAS formats.

The accuracy parameters assessed for evaluating the tree detection and DBH results were: True-positive (TruePos; units) representing the correctly identified trees; the false-positive (FalsePos; units) representing the commission error; the false-negative (FalseNeg; units) representing the omission error, the detection rate and completeness (DR and completeness; percentage) representing the relationship between TruePos and observed tree and the correctness (correctness; percentage) representing the relationship between TruePos and number of stems extracted from TLS data (TreeTLS).

Step 3- Stem reconstruction

The stem reconstruction, corresponding to the trunk section of detected trees, was based on a cylinder-fitting approach embedded in opalsDBH OPALS module. OpalsDBH is one of many forestry modules, embedded in OPALS modular program. This module was initially developed for estimating the DBH, but it recently has been adjusted for measuring the stem diameters at different height levels (<https://opals.geo.tuwien.ac.at/html/nightly/ModuleDBH.html>), namely stem curve (Liang *et al.* 2018). Theoretically, the cylinder and cone geometric shapes drawn on cross-section horizontal slice were measured using a least-squares cylinder-fitting approach implemented in opalsDBH OPALS module (Lukács *et al.* 1997). Practically, the shape of cylinders drawn on each cross-section horizontal slice were measured through a cylinder-fitting approach. The reconstruction of tree trunk was done measuring the

cross-sections consecutive belong to the same stem position, based on the Euclidean distance between predicted vs. observed stem position. The consecutive cylinders distributed along the tree trunk were separated by 0.15 m between two consecutive cylinders (Liang *et al.* 2012).

In this study, the stem position, the DBH and the TLS point clouds labelled as timber are the sources required for running the opalsDBH OPALS module. In addition, we set some mandatory parameters “trace”, representing the option to enable the search of cylinders in one stem axis; “overlap”, representing the percentage of overlapping patches values between traced cylinders; and “patchLength”, representing the length of the shift vector between two consecutive patches. Particularly, the “trace”, “patchLength” and “overlap” parameters were set to 1-1, 0.5 and 0.8, respectively. However, the tracing of cylinders drawn on the cross-sections is stopped if the stem characteristics showed some incongruences: 1) the axis change ($>10^\circ$); 2) the change between two consecutive stem radius ($>50\%$); 3) trunk consecutive overlap ($>60\%$) and 4) distance between two consecutive cylinders ($>50\%$ or “patchLength”). To understand the quantity of reconstructed stems and the proportion of the trunk section was covered by stem curve measurements, four validation parameters were tested:

- Reconstructed stem from TLS data (RStem; units), representing the quantity of reconstructed stems;
- RStem rates (TrueRStem; %), representing the relationship between RStem and observed trees;
- Curve length ratio (CLR, %), representing the relationship between proportion of the stem length covered by the extracted stem curve from TLS data and that obtained from observed data (Liang *et al.*, 2018);
- Percent of the tree height covered (PHC, %), representing the relationship between proportion of the stem length covered by the extracted stem curve from TLS data and the tree height from observed data (Liang *et al.*, 2018).

It is worth to noting that, the extracted stem curve was forced to stop at the TH1 due to inaccurate estimation were expected after this point. Since TH1 was adjusted, these outcomes cannot be used to compare these results to the results of other studies.

Nevertheless, Dmin and Dmax measurements for each trunk stem were validated through statistic measurements. In particular, the coefficient of determination (R-squared; 0-1) and root mean square error (RMSE; m, m³) obtained from linear regression models including and excluding were considered. These measurements were implemented using “stats” (authors, R Core Team and contributors worldwide) and “usdm” (Naimi 2015) R packages. The stem curve information was used as input data in the subsequent step.

Step 4- Timber assortment estimation

In this step, to extract the timber assortment information from point cloud (labelled as timber), we follow three sub-steps: 1) quantifying the logs (i.e. merchantable or non-merchantable logs); 2) characterizing the

merchantable logs, based on the STR and TAP measurements; and, 3) classifying each merchantable logs into one out of 15 types of assortment.

Sub-step 4.1. Quantifying the logs

In this sub-step, the log quantification was based on stem curve output and it was ran using many R packages. Prior to start with the quantification approach, the stem curve obtained in previous step was further pre-processed, particularly, we first calculated the Euclidean distance between two consecutive cylinder positions from down to up; we second grouped the consecutive cylinders included in each trunk section into two cylinders' groups using the accumulated Euclidean distance values as predicted length of log. This means that the large cylinders' groups corresponding to merchantable logs and the small cylinders' groups corresponding to non-merchantable logs (the length for logs and the accumulated Euclidean distance values were considered equivalent measurements); we third quantified the number of cylinders' groups for both merchantable and non-merchantable types. It is worth noting that, to remove the unnecessary cylinder information from merchantable logs to use in the subsequent sub-step, we reduced the number of cylinders included in merchantable logs from infinite to three cylinders by everyone, therefore, the "first cylinder", the "second cylinder" and the "third cylinder" representing maximum, central and minimum cylinder position were left (Figure 19) (Liang et al., 2012). To run the quantification approach, the "stats" (authors, R Core Team, and contributors worldwide), the "dplyr" (Wickham and Francois, 2016) and "usdm" (Naimi, 2015) R packages were implemented. The validation was carried out comparing predicted vs. observed merchantable and non-merchantable measurements. The mean, standard deviation ($SD \pm$) and sum of the quantity of log and length of logs for each log types were used as the evaluation criteria.

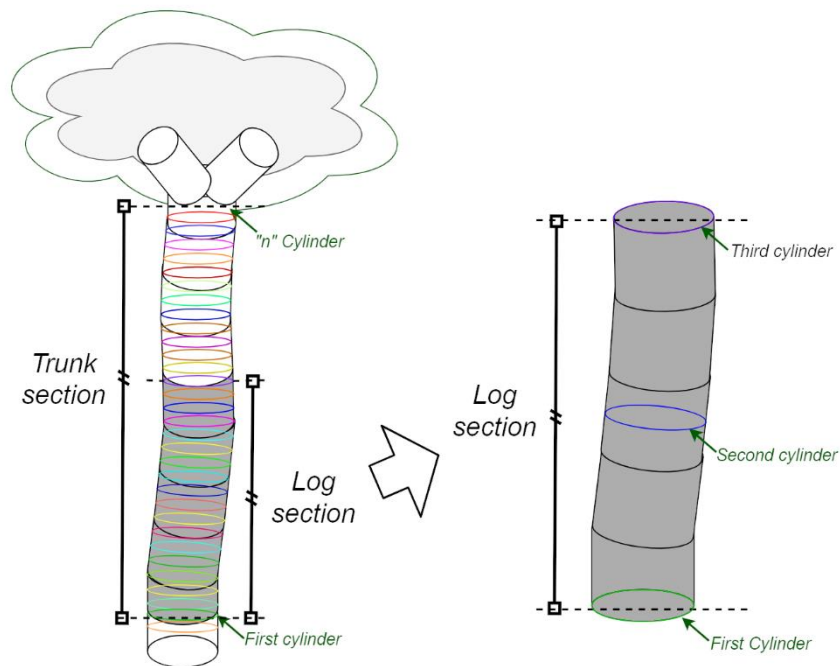


FIGURE 19 AN EXPLICATIVE EXAMPLE ABOUT THE DISTRIBUTION OF CYLINDER ALONG THE TRUNK WITH AN INTERVAL OF 0.15 M (LEFT IMAGE) AND THE THREE CYLINDERS ALONG THE LOGS WITH AN INTERVAL OF 1.4 M (RIGHT IMAGE).

Sub-step 4.2. Merchantable log characterization

For extracting the STR and TAP patterns of the solely merchantable log, the two equations (2 and 3) described from (Togni 2017) were adjusted and implemented using the information of the three cylinders through R functions. These R functions used the “stats” (authors, R Core Team and contributors worldwide) and “dplyr” (Wickham and Francois, 2016) R packages. For the matches, the accuracy of the predicted STR and TAP was evaluated with respect to the observed data. The mean and the standard deviation ($SD\pm$), the bias and the RMSE were used as the evaluation criteria.

Sub-step 4.3. Merchantable log classification

For classifying the merchantable logs into one of 15 timber assortments, the STR and Dmin outcomes provided by the previous sub-step were used here. Particularly, we used the threshold limits (Table 10), described by Togni (2017), for classifying those logs. To run this classification, we implemented a function classifying the logs using the STR and Dmin measurements. To reach this, we used the “stats” (authors, R Core Team and contributors worldwide) and the “dplyr” (Wickham and Francois, 2016) R packages. For the matches, the accuracy of the predicted number of merchantable logs by class was evaluated with respect to the observed data. The bias and the RMSE were used as the evaluation criteria.

2.3.5. Results

The samples plots were characterized by a huge tree species richness and structural heterogeneity (Table 11). The most frequent tree species, ranging from 5 to 9, were *F. sylvatica* (28.7%), *F. excelsior* (14%), *U.*

carpinifolia (12.4%), and *Q. cerris* (11.8%) (Figure 20). Among the ADS, the structural heterogeneity was supported by the DBH ranging between 0.16 m and 0.26 m, the TH ranging between 13.28 m and 23.10 m and the TSv ranging between 0.31 m³ and 1.05 m³ per ADS, as well as by stem density (Table 11). However, the stem density may be divided into complexity levels: low (<500 trees ha⁻¹), moderate (500-900 trees ha⁻¹) and high (>900 trees ha⁻¹). The randomly locations of the all single-scan positions, ranging between 7 and 10 single-scans, were also showed (Figure 20).

TABLE 11 SUMMARY OF FOREST-RELATED CHARACTERISTIC FROM ALL TREES OVER FIVE FIELD PLOTS (ADS). THE DIAMETER AT BREAST HEIGHT (DBH), TREE HEIGHT (TH), HEIGHT OF THE FIRST ATTACHED BRANCH OR BRANCH UNION (TH1), TREE STEM VOLUME (TSV) AND THE TREE SPECIES COMPOSITION (TSC) WERE SHOWED.

Field data								
ADS	N°trees ADS ⁻¹ (N° trees ha ⁻¹)	Complexity level	Description	Forest-related characteristics			TSC (units)	
				DBH (m)	TH (m)	TH ₁ (m)		
1	33 (623)	moderate	Mean	0.2	18.52	7.94	13.2	7
			SD (+/-)	0.09	5.16	3.35		
			Sum					
2	36 (679)	moderate	Mean	0.2	13.28	6.05	23.86	9
			SD (+/-)	0.19	8.19	2.86		
			Sum					
3	52 (981)	high	Mean	0.16	13.72	7.12	16.32	8
			SD (+/-)	0.13	6.79	3.51		
			Sum					
4	33 (623)	moderate	Mean	0.21	21.27	9.86	22.81	9
			SD (+/-)	0.14	8.97	3.84		
			Sum					
5	24 (453)	low	Mean	0.26	23.1	10.57	25.29	5
			SD (+/-)	0.15	10.22	6.06		
			Sum					
Sum	178 (3358)							

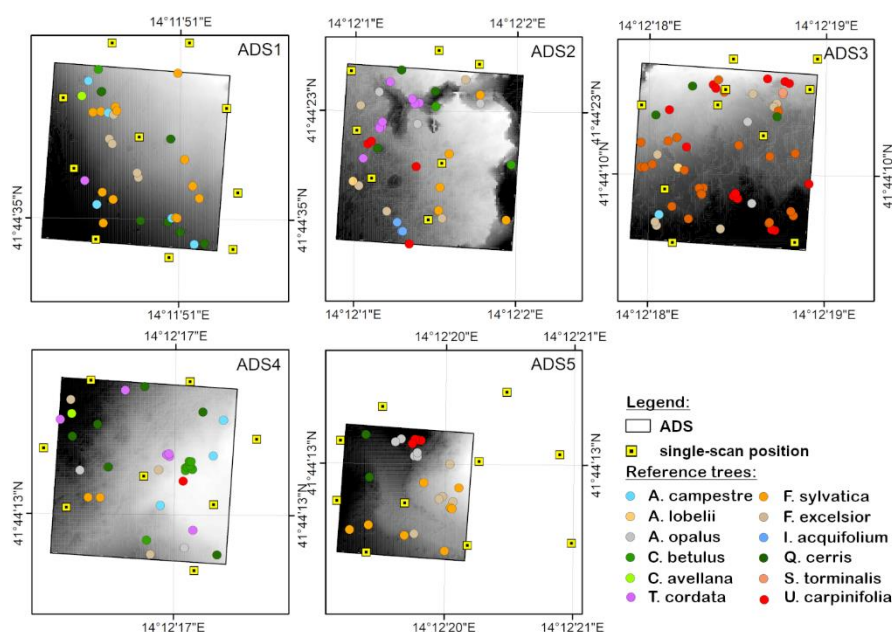


FIGURE 20 TREE POSITION AND SINGLE-SCAN LOCATION WITHIN EACH FIELD PLOT (ADS). THE BLACK SQUARES REPRESENT THE ADS BOUNDARIES, THE YELLOW SQUARES REPRESENT THE SINGLE-SCAN POSITIONS, THE

COLOURED DOT POINTS REPRESENT THE OBSERVED TREE POSITION CLASSIFIED BY TREE SPECIES ON DIGITAL TERRAIN MODEL (DTM) BACKGROUND.

In our study, more than one-third of observed trees, corresponding to 70 out of 178 observed trees, were suited for timber assortment assessment based on (DBH > 0.20 m). Overall, the most frequent species of them were the *F. sylvatica* (32.9 %; 23 trees), *Q. cerris* (21.4 %; 15 trees), *F. excelsior* (14.3 %; 10 trees) and *A. opalus* (11.4 %; 8 trees) and other four broadleaved species (20%; 14 trees). 70 observed trees provide 306 and 79 observed merchantable and non-merchantable logs (Table 12). Based on the STR, ranging between 1.4 cm m⁻¹ and 2.9 cm m⁻¹, and the TAP, ranging between 1.1 cm m⁻¹ and 1.8 cm m⁻¹, measurements, most merchantable logs goes from slightly to strongly contorted due to the stem profile description. The volume stored in merchantable logs was ten times higher than that stored in non-merchantable logs.

TABLE 12 LOG QUALITY TRAITS FOR ALL FIVE STUDY AREAS (ADS). THE STRAIGHTNESS (STR) AND TAPERING (TAP) VARIABLES, AND THE LOG VOLUME (TTV.LOG) OF MERCHANTABLE AND NON-MERCHANTABLE LOGS WERE DISPLAYED. THE MEAN, STANDARD DEVIATION (SD ±) AND THE SUM WERE USED FOR EVALUATING THE ACCURACY.

ADS		Log quality traits						
		STR (cm m ⁻¹)		TAP (cm m ⁻¹)		TTV.log (m ³)		
Type	ADS	N°logs	Mean	SD (±)	Mean	SD (±)	Sum	
Merchantable	1	88	2.9	1.9	1.5	0.7	7.2	
	2	45	1.6	1.3	1.8	1.0	10.9	
	3	35	1.4	1.1	1.6	0.4	6.7	
	4	56	1.8	1.4	1.1	0.5	12.9	
	5	82	1.6	0.9	1.2	0.4	12.4	
	Mean			1.8	1.3	1.4	0.6	10.0
	Sum	306						
Non-merchantable	1	30	2.1	4.0	1.3	2.6	1.1	
	2	13	1.7	2.5	0.9	1.3	1.2	
	3	11	1.4	1.7	0.5	1.2	0.7	
	4	8	2.1	2.9	1.0	1.4	0.4	
	5	17	1.3	2.6	1.0	1.3	1.2	
	Mean			1.7	2.7	0.9	1.6	0.9
	Sum	79						

2.3.5.1. Timber-leaves discrimination

The results revealed that the Random Forest algorithm was able to accurately discriminate the timber from leave points in mixed-species and heterogeneous stand structure, as supported by the similar accuracy (0.98), sensitivity (0.98) and specificity (0.98) values obtained in all five ADS. Despite the optimal capability for classifying the point clouds, the upper part of canopy height (affecting taller trees) and the presence of shrubs and or lianas (understory layer) have favoured the occurrence of noise points. In the light of the foregoing, the timber-leave discrimination was slightly influenced by the quality of point clouds and forest structure in mixed-species and multi-layered forests.

2.3.5.2. Tree detection and DBH estimation

We detected 151 out of 178 observed trees, reaching average detection rate accuracy equal to 84.4%, with high uniformity/similarity across the ADS, based on the standard deviation values ($SD = \pm 4.7\%$) (Table 13). The results revealed that the tree detection approach was more sensitive to the commission error (84 as the sum of FalsePos) than the omission error (27 as the sum of FalseNeg), which was also supported by the different patterns of the completeness (84.4%) and correctness (66.9%) accuracies (Table 13). It is worth noting that, although the average detection accuracy was 84.4 %, it was increased for large trees with a DBH > 0.30 m, reaching an average detection accuracy equal to 100 %.

The detection accuracy was rather variables among the three complexity levels of stem density. Particularly, as concerns the ADS belonging to moderate complexity levels, we detected 87 out of 102 observed trees, reaching an accuracy ranged between 80.6% and 90.9%. About the ADS belonging to the high complexity level, we detected 45 out of 52 observed trees, supporting an accuracy equal to 86.6%. About the ADS belonging to low stem density level, we detected 19 out of 24 observed trees, reaching an accuracy equal to 79.2% (Table 13). In the light of above, the detection accuracy increased in moderate and high complexity levels of stem density.

TABLE 13 SUMMARY OF TREE DETECTION RESULTS. OBSERVED TREES FROM FIELD DATA (TR), PREDICTED TREES FROM TLS DATA (TREETLS), TRUE POSITIVE (TRUEPOS), FALSE POSITIVE (FALSEPOS), FALSE NEGATIVE (FALSENEG), DETECTION RATE (DR), COMPLETENESS AND CORRECTNESS FOR EACH STUDY AREA (ADS). MEAN, STANDARD DEVIATION ($SD \pm$) AND SUM WERE ALSO DISPLAYED.

Tree detection results								
ADS	TR	Tree detection measurements						
		TreeTLS	TruePos	FalsePos	FalseNeg	DR (%)	Completeness (%)	Correctness (%)
1	33	45	30	15	3	90.9	90.9	66.7
2	36	54	29	25	7	80.6	80.6	53.7
3	52	71	45	26	7	86.5	86.6	63.4
4	33	36	28	8	5	84.8	84.9	77.8
5	24	26	19	7	5	79.2	79.2	73.1
Sum	178	232	151	81	27			
Mean	36	46	30	16	5	84.4	84.4	66.9
SD (+/-)	10.2	17.2	9.4	9.0	1.7	4.7	4.7	9.3

The detection accuracy decreases, as the tree species composition increase because the best and worst detection accuracy was found in ADS including five (ADS5; DR = 90.9 %) and seven (ADS1; DR = 79.2 %) tree species. Conversely, the detection accuracy varies from 80.6 % to 86.5 % in ADS include seven, eight and nine different tree species (ADS2, ADS3 and ADS4). Nevertheless, a great detection rate accuracy was found in *A. lobelii*, *S. torminalis*, *F. excelsior*, *Q. cerris*, *A. campestre* and *F. sylvatica* (> 84.3 %) compared to other six broadleaved stems (50-77 %) (Figure 21). Although the detection accuracy increases in ADS with a least number of species (< 5 tree species), it is related to tree species (*A. lobelii*). (Figure 21)

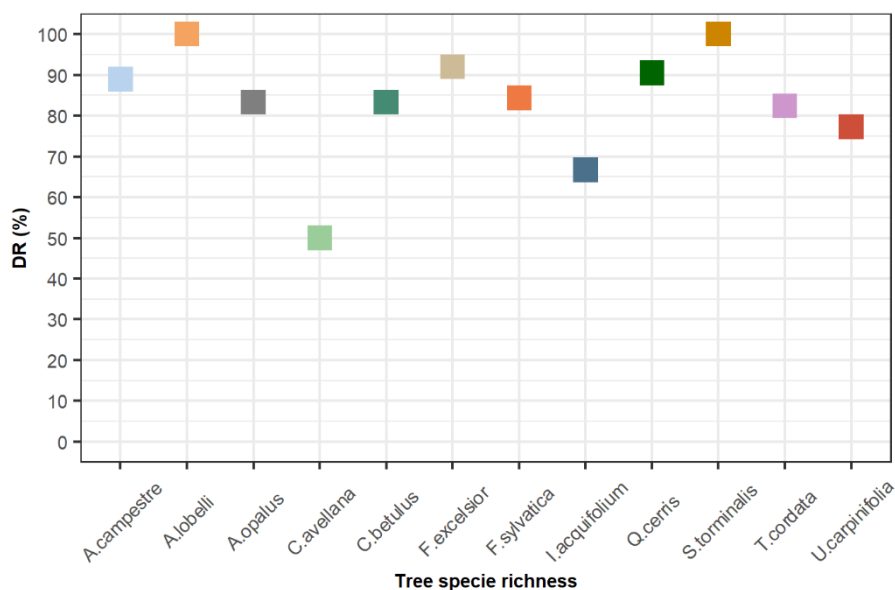


FIGURE 21 DETECTION RATE (DR) VALUES FOR INDIVIDUAL TREE SPECIES

The detection accuracy increases, as the stem dimension increases, and therefore, a great capacity for detecting large trees (DBH > 0.30 m) was proved by using the cylinder-fitting approach (DR = 100 %). This statement was more evident for ADS showing moderate stem densities (ADS 1, 2 and 4) compared to ADS showing lowest and highest stem densities (ADS 3 and 5) (Table 14).

TABLE 14 DETECTION ACCURACY. OBSERVED TREES FROM FIELD DATA (TR, UNITS), TRUE POSITIVE (TRUEPOS, UNITS) AND DETECTION RATE (DR, %) WERE SHOWED FOR THE THREE DIFFERENT DIAMETERS AT BREAST HEIGHT (DBH) INTERVALS. MEAN, STANDARD DEVIATION (SD ±) AND SUM WERE ALSO DISPLAYED.

		Tree detection results														
		ADS1			ADS2			ADS3			ADS4			ADS5		
DBH interval (m)		TR	TD	DR (%)	TR	TD	DR (%)	TR	TD	DR (%)	TR	TD	DR (%)	TR	TD	DR (%)
1 st	<0.1]	3	2	66.7	17	12	70.6	22	17	77.3	5	4	80	4	3	75
2 nd	[0.1-0.2]	17	15	88.2	5	3	60	17	16	94.1	15	12	80	3	1	33.3
3 rd	[0.2-0.3]	7	7	100	6	6	100	4	3	75	7	6	85.7	9	7	77.8
4 th	[>0.3]	6	6	100	8	8	100	9	9	100	6	6	100	8	8	100
	Sum	33	30		36	29		52	45		33	28		24	19	
	Mean			88.7			82.7			86.6			86.4			71.5
	SD (±)			15.7			20.5			12.3			9.4			27.8

Comparing the predicted vs. observed DBH data from correctly detected trees, corresponding to 151 trees (Table 15), we found better predictions accuracy in the linear regression model excluding the outlier data (R-squared = 0.84; RMSE = 0.02 m) than that including outlier data (R-squared = 0.67; RMSE = 0.08 m) (Figure 22). The cleaned outliers, mainly overestimated, belong to both large and small stems (different DBH patterns), hence, a moderate accuracy for predicting the DBH was showed for all stems.

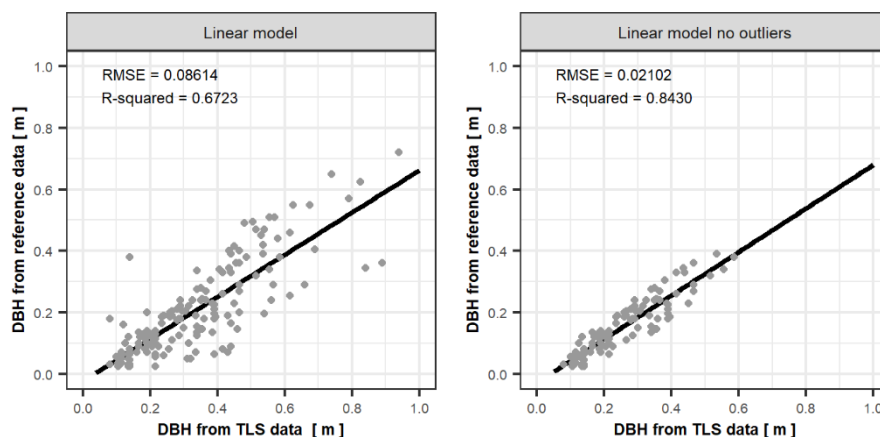


FIGURE 22 PREDICTED VS. OBSERVED VALUES OF THE TWO LINEAR MODELS FOR THE DIAMETER AT BREAST HEIGHT (DBH).

2.3.5.3. Stem reconstruction

We reconstructed 47 out of 70 observed trees using TLS data through a cylinder-fitting approach, reaching an average stem reconstruction accuracy equal to 67.2%, with low similarity/uniformity among the ADS (SD = ±14.86%) (Table 15).

The stem reconstruction accuracy was rather variables among the three complexity levels of stem density. Particularly, as concerns the ADS belonging to moderate complexity levels, we reconstructed 28 out of 40 observed trees, reaching a stem reconstruction accuracy ranged between 53.8 % and 84.6 %. About the ADS belonging to the high complexity level, we reconstructed 7 out of 13 observed trees, supporting a stem reconstruction accuracy equal to 53.8 %. About the ADS belonging to low stem density level, we reconstructed 12 out of 17 observed trees, reaching a stem reconstruction accuracy equal to 70.6 % (Table 15).

Although the reconstruction accuracy was equal to 67.2 %, an enhanced stem reconstruction was found for *Q. cerris* (66.7%), *A. opalus* (41.7%) and *F. excelsior* (40%) because the other five tree species showed a reconstruction accuracy lower than 26 % (Figure 23).

TABLE 15 STEM RECONSTRUCTION RESULTS. OBSERVED TREES FROM FIELD DATA (TR, UNITS), RECONSTRUCTED STEM FROM TERRESTRIAL LASER SCANNING DATA (RSTEM; UNITS) AND RATE OF RSTEM (TRUERSTEM, PERCENT) WERE DISPLAYED FOR FIVE DIFFERENT DIAMETERS AT BREAST HEIGHT (DBH) INTERVALS AND IT IS SEPARATED BY EACH STUDY ARE (ADS). MEAN, STANDARD DEVIATION (SD ±) AND SUM WERE ALSO DISPLAYED.

		Stem reconstruction results						
ADS	Description	DBH interval (m)					Sum	Total %
		1 st [0.2-0.3]	2 nd [0.3-0.4]	3 rd [0.4-0.5]	4 th [0.5-0.6]	5 th [>0.6]		
ADS1	TR	7	6				13	
	RStem	6	4				10	
	TrueRStem							76.9
ADS2	TR	6	1	3	2	2	14	
	RStem	2	1	1	2	1	7	

	TrueRStem						50
ADS3	TR	4	3	6			13
	RStem	1	2	4			7
	TrueRStem						53.8
ADS4	TR	7	1	3	2		13
	RStem	5	1	3	2		11
	TrueRStem						84.6
ADS5	TR	9	6	0	1	1	17
	RStem	6	5	0	1	0	12
	TrueRStem						70.6
Sum	TR	33	17	12	5	3	70
	RStem	20	13	8	5	1	47
Mean and SD (\pm)	TrueRStem	60.6	76.5	66.7	100	33.3	67.2 (14.9)

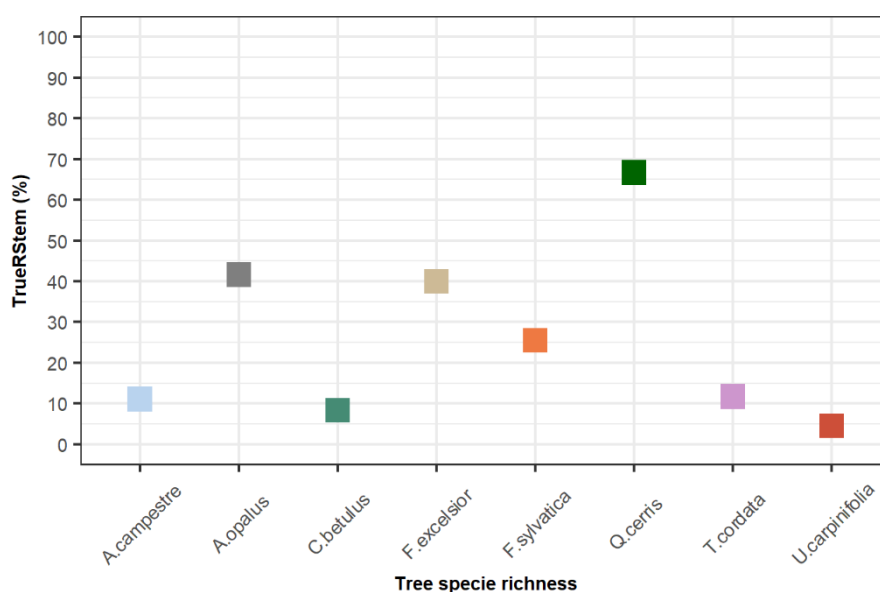


FIGURE 23 RATE OF RECONSTRUCTED STEM DERIVED FROM TERRESTRIAL LASER SCANNING DATA (TRUERSTEM) WAS SHOWN FOR EACH TREE SPECIES

The capability of the cylinder-fitting approach for detecting and tracing the cylinders along each trunk was accurately proved by the stem curve patterns. These patterns were supported by the high CLR and PHC accuracies, 88.1% (SD = \pm 16.7 %) and 35.4% (SD = \pm 11.3 %) (Figure 24A-B), respectively.

As regards the trunk section described by the cylinders, more than three-quarters of all stems, corresponding to 39 reconstructed stems, were entirely described by fitted cylinders (CLR > 80 %) and remaining stems, corresponding to 8 reconstructed trees, were solely partially described by cylinders, ranged between 26.4 % and 76.3 % (Figure 24A). As regards the proportion of tree height described by fitting cylinders, nearby three-quarters of all stems, corresponding to 34 reconstructed stems, were partially described by fitted cylinders (30 % >PHC > 68.3 %) and the remaining stems, corresponding to 13 reconstructed trees, were barely described by cylinders, ranged between 12.9 % and 29.9 % (Figure 24B). Despite the small PHC values support the idea that trunk section with productivity aptitudes was short, the main cause that stopped

the detection of the cylinder along each stem was derived from the height of the first attached branch or branch union (Figure 24A-B).

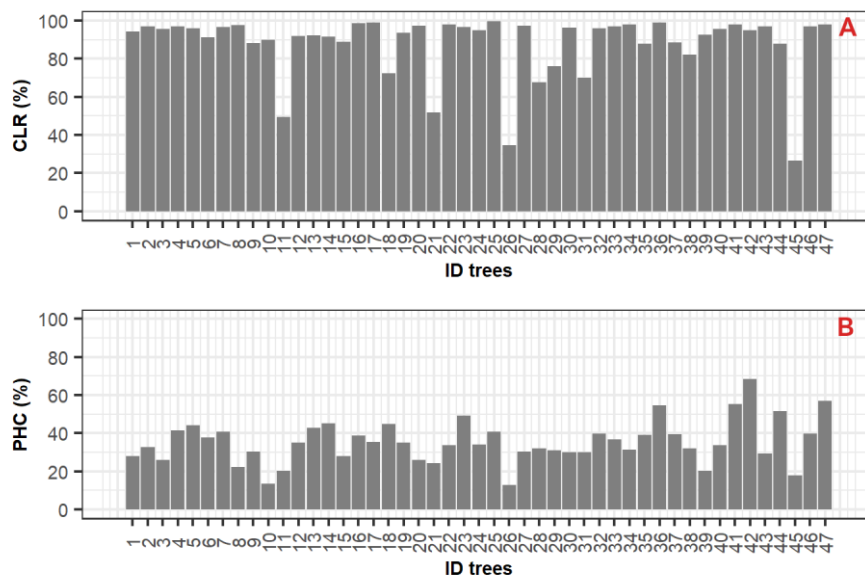


FIGURE 24 STEM CURVE PERFORMANCE. THE CURVE LENGTH RATIO (CLR; A) AND THE PERCENTAGE OF THE TREE HEIGHT COVERED (PHC; B) WERE DISPLAYED FOR ALL RECONSTRUCTED TREES.

Comparing the predicted vs. observed values for Dmax and Dmin data from correctly reconstructed trees, corresponding to 47 trees, we found better predictions accuracy in the linear regression model excluding outlier for both Dmax (R-squared = 0.86; RMSE = 0.03 m) and Dmin (R-squared = 0.89; RMSE = 0.03 m), respect to that including the outliers for both Dmax (R-squared = 0.60; RMSE = 0.08 m) and Dmin (R-squared = 0.56; RMSE = 0.08 m) (Figure 25). As regards the Dmax outliers, the eleven cleaned outliers, that fostered a slight Dmax under/overestimation, were values exceeding 0.49 m. As regards the Dmin outliers, the three cleaned outliers, that fostered a slight Dmin underestimation, were values exceeding 0.61 m. Therefore, fitted predictions for both Dmax and Dmin were found for values lower than 0.49 m and 0.61 m, respectively.

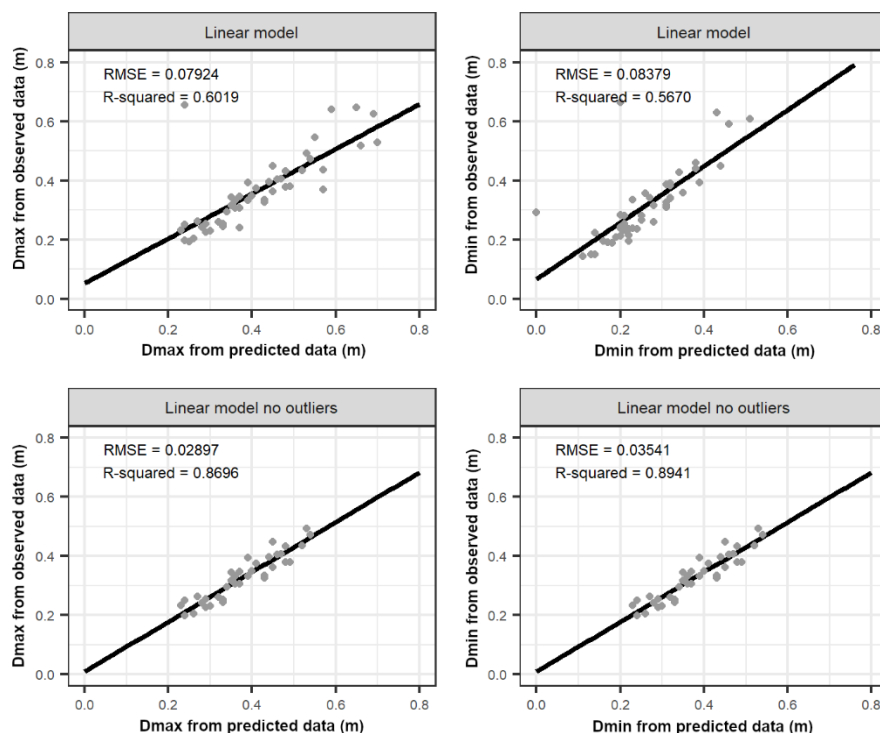


FIGURE 25 PREDICTED VS OBSERVED VALUES OF THE LINEAR MODELS OBTAINED FOR MAXIMUM (DMAX) AND MINIMUM (DMIN) ENDS DIAMETERS.

2.3.5.4. Timber assortment estimation

Log quantification:

Since the timber assortment assessment was based on stem reconstruction outcomes, we used the reconstructed trees, corresponding to 47 reconstructed trees, for timber assortment assessment. In our study, 47 observed trees provided 179 merchantable and 40 non-merchantable logs. More than three-quarters of both type of logs were quantified, particularly, 134 out of 179 merchantable logs and 34 out of 40 non-merchantable logs were quantified. (Table 6). Comparing the predicted vs. observed length of logs, we noted that predicted data was larger than observed data, and such difference was similar for merchantable (2.5m vs. 2.78) and non-merchantable (1.35m vs. 1.62) logs (Table 16).

TABLE 16 LOG QUANTIFICATION RESULTS. THE NUMBER OF LOGS (N°LOGS, UNITS) AND THE LENGTH OF LOG (L.LOG, M) MEASUREMENTS WERE DISPLAYED FOR MERCHANTABLE AND NON-MERCHANTABLE LOGS. MEAN, STANDARD DEVIATION (SD ±) AND SUM WERE ALSO DISPLAYED.

Log section results								
Log section	Observed data				Predicted data			
	N°logs	L.log		Sum	N°logs	L.log		Sum
		Mean	SD(±)			Mean	SD(±)	
Merchantable	179	2.5	0	447.5	134	2.78	0.12	372.51
Non-merchantable	40	1.35	0.69	53.90	34	1.62	0.57	54.99

Considering the quantity of predicted and observed logs for each tree species, we observed a similar quantification accuracy for both log types. The quantification accuracy obtained for merchantable logs based on bias and RMSE were 5.6 and 8.3 units per tree species, respectively. The quantification accuracy obtained for non-merchantable logs based on bias and RMSE were 0.8 and 1.7 units per tree species, respectively. We also noted that an enhanced quantification accuracy was found for *C. betulus* (1 out of 1), *F. sylvatica* (36 out of 42), *A. opalus* (11 out of 14), *F. excelsior* (34 out of 45), *U. carpinifolia* (3 out of 4), since more than three-quarters of logs were correctly matched between observed and predicted data (Table 17).

TABLE 17 LOG RESULTS SHOWED FOR TREE SPECIES.

Log results for tree species							
Tree species	Merchantable (Units)			Non-merchantable (Units)			
	Observed data	Predicted data	Accuracy	Observed data	Predicted data	Accuracy	
1 <i>Q. cerris</i>	62	43		10	11		
2 <i>F. sylvatica</i>	42	36		12	10		
3 <i>F. excelsior</i>	45	34		10	6		
4 <i>A. campestre</i>	3	2		0	0		
5 <i>A. opalus</i>	14	11		5	4		
6 <i>T. cordata</i>	8	4		1	1		
7 <i>U. carpinifolia</i>	4	3		1	1		
8 <i>C. betulus</i>	1	1		1	1		
Sum	179	134		40	34		
bias			5.6			0.8	
RMSE			8.3			1.7	

Comparing the predicted vs. observed TTV.log data, from quantified 134 merchantable logs, we found better predictions accuracy in the linear regression model excluding the outlier data (R-squared = 0.92; RMSE = 0.03 m³) than that including the outlier data (R-squared = 0.77; RMSE = 0.06 m³) (Figure 26). As far as concerns the outliers, the twelve cleaned outliers, that fostered a slight TTV.log overestimation, were values exceeding 0.33 m³. Therefore, fitted predictions were observed for TTV.log values to be inferior 0.33 m³.

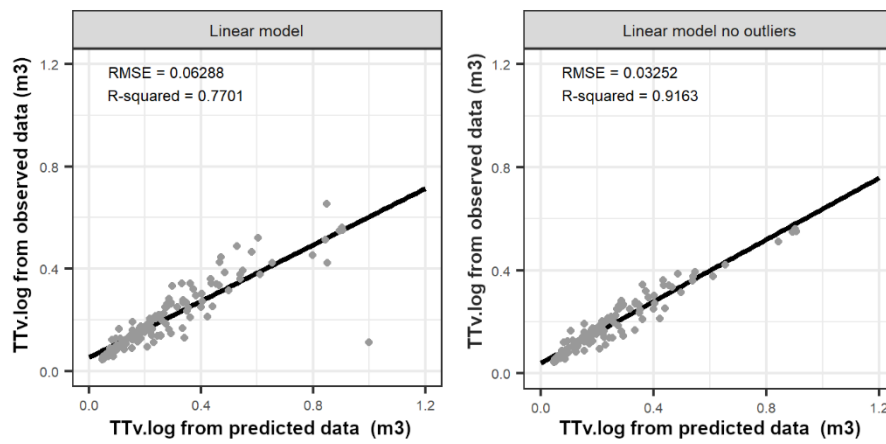


FIGURE 26 PREDICTED VS OBSERVED VALUES OF THE LINEAR MODELS FOR THE VOLUME OF LOGS (TTV.LOG).

Log characterization

The characterization of logs proved to be more accurate for STR with respect to TAP measurements, based on the accuracy findings found for these, particularly, the STR bias value was 0.77 cm m⁻¹ and the TAP bias value was 1.69 cm m⁻¹(Figure 27). Conversely, the TAP and STR patterns resulted to be rather similar among them based on the standard deviation values for these: TAP (SD = ±1.79 cm m⁻¹) and STR (SD = ±1.73 cm m⁻¹). Despite the poor accuracy obtained comparing predicted with observed TAP data, in several log cases the negative and positive patterns found for observed data were even reported for predicted data.

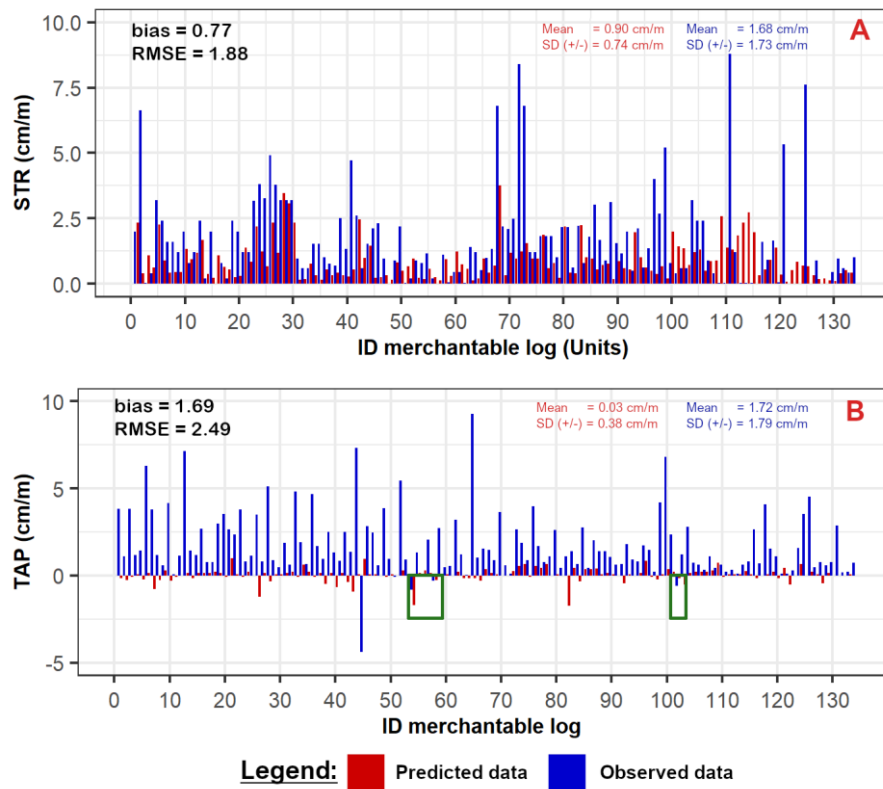


FIGURE 27 THE STRAIGHTNESS (STR) AND TAPERING (TAP) VALUES FOR MERCHANTABLE LOGS THE RED SQUARED INDICATES THE NEGATIVES VALUES OF TAP REPORTED FOR PREDICTED AND OBSERVED DATA.

Log classification:

The results highlighted that the log classification resulted to be fit for eleven out of 14 assortment types. It was based on the absolute lower values of bias and RMSE accuracy, -1.36 and 7.13 logs per type of assortment, respectively (Figure 28). Comparing predicted with observed log classification findings, we observed that 8 out of 11 assortment types proved to be more accurate based on the variation of merchantable logs (± 2). These assortment types belong mainly to saw-log and other industrial roundwood and someone else of other assortment types, particularly, A-, B-, B0, B+, C-, D-, D+ and Fuelwood-. The principal high-quality assortment, namely A+, was found to be strongly overestimated (predicted = 67 vs. observed = 43) (Figure 28). Among the assortment types that did not include some merchantable logs, corresponding to C+, Fuelwood+ and Fuelwood0, one was empty using observed data too and the other

two were assortment less abundant or low-quality (C+ = 2 logs, Fuelwood+ = 1 log and Fuelwood0 = 7 logs).

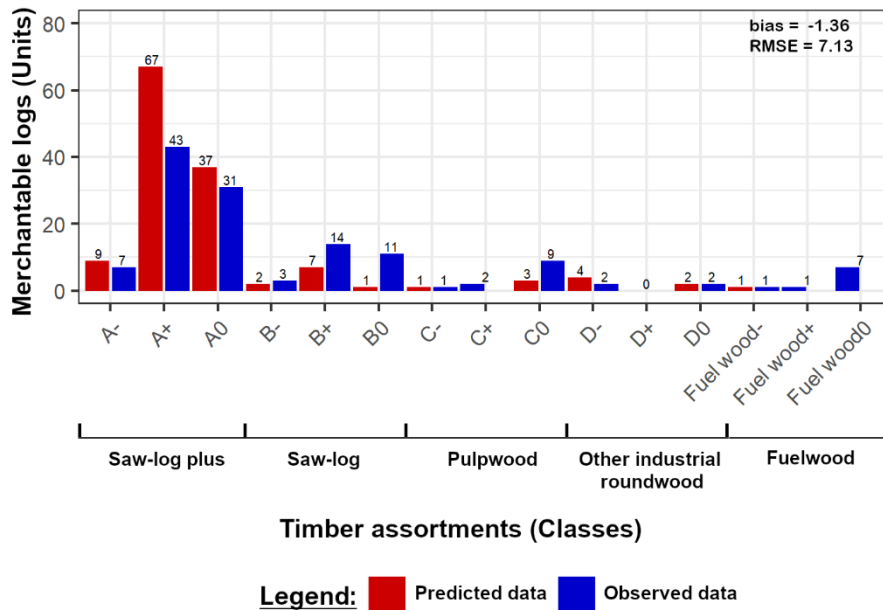


FIGURE 28 CLASSIFICATION OF THE MERCHANTABLE LOGS INTO FIFTEEN ASSORTMENT TYPES BELONGING TO SAW-LOG PLUS, SAW-LOG, PULPWOOD, OTHER INDUSTRIAL ROUNDWOOD AND FUELWOOD ASSORTMENTS.

2.3.6. Discussion

2.3.6.1. Separating the timber from point clouds

Results display that occlusion factors as trees in the understory layers, trunks, branches and leaves hinder the discrimination of timber from leaves points. This is even more worsened by the abundance of lianas, as a naturalness indicator of this forest (Santopuoli *et al.* 2019; Vicari *et al.* 2019). Although the eigenentropy thresholds allowed us to remove the noise points that prompted the timber-leaf discrimination, more efforts are necessary to better classify small branches (< 0.01m), especially those situated in the upper portion of the canopy. (Ma *et al.* 2016; Vicari *et al.* 2019) faced similar challenges in the timber-leave discrimination, and they associated it with the quality of point cloud, particularly, the point spacing, the density and the incidence angle uncertainties; other study indicated that these challenges can be also associated with the shaded effect from large to small stems derived from the pre-processing issues (i.e. assembling among scans) (Vicari *et al.* 2019). In our study, the shaded effect in the discrimination approach seems to be a plausible justification, but it was aggravated in study areas with high richness species and structural heterogeneity stands. Despite the challenges found in timber-leave discrimination, our findings were in line with that reported in some studies (Ma *et al.* 2016; Wang *et al.* 2017; Vicari *et al.* 2019). However, the two main differences between our study and other studies were the number of predictor variables (our study = 8 predictor variables vs. literature = 10 predictor variables) and tree species composition (our study = 9 vs. literature = less than 3 tree species) (Ma *et al.* 2016; Wang *et al.* 2017; Vicari *et al.* 2019).

We observed that the standard value of “Ln” has influenced the interpretation of neighbouring points, in particular, the occurrence of noise points resulted to be most abundant in some trees, despite of these were removed using the filtering approach. Therefore, timber-leaf discrimination can be improved using variable values of “Ln” in accordance with its point cloud quality, and it can also result beneficial in interpreting the neighbouring points (Lari and Habib, 2012; Weinmann *et al.* 2015). Nevertheless, the combined use of RF algorithm with a filtering approach allowed us to separate the timber from leaves points and to generate appropriate input data for tree detection and stem reconstruction subsequent steps. Similar strategies were even tested to improve the performance of the binary classification approach, for example, (Vicari *et al.* 2019) proposed a stepwise approach for timber-leaves discrimination following four steps: majority filter, feature filter, cluster filter and path filter; (Tao *et al.*, 2015) proposed an approach using the spatial distribution of the point neighbourhoods for separating the leaves from timber points.

2.3.6.2. Tree detection

It is worth highlighting that our study was carried out in a mixed and multilayer Mediterranean forest, within which the main management aim is biodiversity conservation through very limited harvesting activities in the last 50 years. The results revealed that the tree species composition and forest structure have slightly influenced the detection accuracy of trees using TLS data, reaching an average DR (DR = 84.4 %) (Table 13). Enhanced detection accuracy was found in the forest with more than 500 tree ha⁻¹, for six tree species (i.e. *A. lobelii*), and large trees (DBH > 0.30 m). Conversely, we observed a limited capacity of the cylinder-fitting algorithm for detecting small trees, however, this challenge was even found for eighteen automatic and semi-automatic algorithms (Liang *et al.* 2018), and in this study, this challenge was related to the incomplete definition of the cylinder of stems. The main hindering factors influencing the incomplete definition of the cylinder of stems were shadow effects from large to small trees, poor point density, assembling errors, shadow effects from branches to trunk, stem straightness, non-circular shape and tree species composition (Liang *et al.*, 2018). Along with these hindering factors, secondary factors, such as, lianas' and shrubs' occurrence, and also the terrain pendency, can affect the detection accuracy of small trees (Liang *et al.* 2012; Olofsson *et al.* 2014; Koreň *et al.* 2017). However, in our study, since the commission error was superior to omission error, the assumption of a shadow effect from large to small trees, shadow effects from branches to trunk, stem straightness became plausible.

Despite the slightly different forest stand conditions amongst the ADS, our results were higher and or in line than the results observed in studies using TLS single-scan and TLS multiple-scans in other contexts. For example, TLS single-scans in forests characterized by mixed dry broadleaved species, (Reddy *et al.* 2018) reached to detect more than 70 % of observed trees in plots with 450 trees ha⁻¹ (vs. ADS 5; DR = 79.2 %; ADS with high stem density) through a circle-fitting approach; in forest characterized by *Pinus spp.*, *P. abies* and *B. pendula* (Liang *et al.* 2012) reached to detect the 73 % of observed trees in plots with 509 - 1432 trees ha⁻¹ (vs. ADS1,2,4; DR = 80.6 - 90.9 %; ADS with low and moderate stem density) through an approach using the flatness, direction and shape features for detecting trees. As regards the multiple-scans: in forest characterized by *P. abies*, *Pinus spp.* and *B. pendula*, (Olofsson *et al.* 2014) reached to detect the 87 % of observed trees in plots with 358 - 1042 trees ha⁻¹ (vs. DR = 84.4 %) through a cylinder-fitting

approach using the TLS features (i.e. flatness saliency) for detecting trees; in forest characterized by *C. betulus*, *F. sylvatica*, *P. menziessi*, *P. abies*, *Quercus spp.* and *P. sylvestris*, (Bauwens *et al.* 2016) reached to detect the 93% (SD \pm 8%) of observed trees in plots with 113 - 1344 trees ha⁻¹ (vs. DR = 84.4 %) through a cylinder-fitting approach.

Nevertheless, comparing our tree detection method with other fourteen algorithms using multiple-scans, our completeness accuracy (84.4%) was in line with the best five algorithms (ranged between 76 % and 88%) (Liang *et al.* 2018), despite the forest stand condition difference (i.e. 4 vs. 12 tree species and monolayer vs. multi-layered). However, the correctness obtained from our study (66.9 %) was lower than the results from fourteen algorithms (ranged between 50 - 95%) (Liang *et al.* 2018).

The accuracy of DBH detection was affected by the occlusion and shadow effects from the bark roughness, stem straightness and non-circular shape of trunks, liana's presence and non-circular shape of trunks, despite the powerful capability of the cylinder-fitting approach (Liang *et al.*, 2018, 2019). Such hindering factors are often caused by technical (i.e. collection-georeferencing) and operational aspects (distance between tree position and TLS scanner, number of scans) (Saarinen *et al.* 2017; Liang *et al.* 2019). However, in our study, the automatic approaches applied allowed to overcome part of these (i.e. OpalsICP for georeferencing). The RMSE values showed by our linear models including and excluding outliers (ranged between 0.01 m and 0.086 m) were comparable with the results obtained in other studies from about 14 TLS algorithms (0.053m - 0.074m SD \pm 0.057m – 0.072m) (Liang *et al.* 2018). Similarly, our DBH responses in RMSE terms were comparable with the result showed for (Reddy *et al.* 2018), ranging between 0.01 m and 0.05 m, for (Kankare *et al.* 2016), ranging between 0.063 m and 0.147 m and for (Bauwens *et al.* 2016) was 0.013 m, despite the forest structure in Bosco Pennataro was rather complex.

2.3.6.3. Stem reconstruction

Results revealed that the use of a cylinder-fitting approach allowed us to reconstruct nearby three-quarters of detected trees from TLS point cloud in mixed-species and heterogeneous stand structures (Table 14 and Table 15). The enhanced stem reconstruction values were found in the ADS of moderate and high complexity density levels, and with lower than 5 tree species, especially the *Q. cerris*. This is because despite *Q. cerris* presents logs defects, such as, the variation in straight and the presence of insect holes (Musat *et al.* 2017), these defects were overcome required by our cylinder-fitting approach due to the stem diameter and straightness were computed. It is important to underline the tree species with a high dimension were the most frequently reconstructed (*A. opalus*, *F. excelsior*, *F. sylvatica*). This can be supported by the genetic traits of each tree species, represented by the stem straightness and irregularities of the trunks (i.e. knots, bulges). In the light of the above, the stem reconstruction accuracy becomes the challenge for stem with irregular stem profile (i.e. stem straightness) and it can be further worsened by the presence of lianas around stem axis, or showing trunk-branch crossing or some microhabitats type (Grinwall *et al.* 2005; Kankare *et al.* 2016; Koreň *et al.* 2017; Santopuoli *et al.* 2019).

A good proportion of the stems were reconstructed based on the stem curve pattern, which was supported by the CLR and PHC outcomes, particularly the average of CLR was 88.1 % (SD = \pm 16.7 %) and PHC was

35.4 % (SD = ± 11.3 %) (Figure 24). Our CLR and PHC results were comparable with the results from thirteen algorithms using multiple-scans, which ranged between 74 % and 87 % (Liang *et al.* 2018) and ranged between 56 % and 94 %, CLR and PHC, respectively. Despite the optimal capability of cylinder-fitting approach for extracting the stem curve, the stem curve accuracy decrease as the tree height increase, especially at the height of first attached branch and in the branch union (Wang *et al.* 2016a; Liang *et al.* 2018) due to the point density in the upper canopy decrease (Figure 24).

In regards to the stem diameter, the Dmin and Dmax resulted to be more accurate in linear regression models excluding the outliers, which are frequent belong to the large stem diameters. This result could be supported by the three assumptions: 1) high frequency of irregular forms, compared to the cylinder, for large trees respect to the small trees, due to the presence of lianas, non-circular forms, presence of knots and or bulges, presence of microhabitats (Bienert *et al.* 2007; Koreň *et al.* 2017; Kankare *et al.* 2013; Rehush *et al.* 2018); 2) low quality of the TLS point clouds which fostered a shadow effect of the cylinders on the cross-sections (Olofsson *et al.* 2014; Pitkänen *et al.* 2019), and, 3) the manual gathered of diameters using CloudCompare software which increased the uncertainties of the cylinder measurements (Kankare *et al.* 2013; Olofsson *et al.* 2014). All the above-mentioned assumptions are strictly dependent on the forest structures, even if the latter could be improved with experience and well-trained staff.

2.3.6.4. Timber assortment

Results revealed that the quantification and classification were based on stem reconstruction outcomes (Table 16-17). Forty-seven detected trees have provided 219 logs, 179 merchantable logs and 40 non-merchantable logs. More than three-quarters of merchantable and non-merchantable logs were quantified using the cylinder position and cylinder dimension from the stem curve (Table 16). We noted that some logs from observed data were “missing”; most of these “missing” logs were triggered from the trunks that not were completely covered by the stem curve (CLR patterns; Figure 24). This error, however, was associated with the irregular stem form (i.e. stem straightness) and the irregularities on the bark (i.e. geometry defects: knots, bulges, microhabitats) (West 2009; Liang and Hyypä, 2011).

As regards the log volume (TTv.log), the best prediction was shown for the linear regression model excluding the outliers (R-squared = 0.91; RMSE = 0.03 m³; Figure 16). This was supported by comparing our results with that obtained in other pure stands, despite the forest covers studied were characterized by the tree species richness and heterogeneous stand structure. For example, our results were comparable with results reported for a study focused on *P. sylvestris* and *P. abies*, in which the accuracy for stem volume was 0.83 (R-squared), using the cylinder position and dimension as input data, and 0.94 (R-squared) using a stem model approach (Kankare *et al.* 2013). Conversely, our accuracy was lower than the accuracy found in *Pinus* spp. and *P. abies* stems due to the R-squared and RMSE was 0.98 (RMSE = 0.02-0.03m³) and in our study was 0.89 (RMSE = 0.03m³) (Liang *et al.* 2014), probably due to the different tree species composition and complexity of forest structure.

In our study, the log quality of the merchantable logs from observed data was characterized by the crooked logs with several bends based on STR and TAP measurements. Particularly, the STR and TAP patterns

derived from observed data showed high variability/uniformity among the merchantable logs and this variability was more marked for TAP compared to STR pattern. The variability of the STR and TAP patterns however might be associated with the morphological traits and these were based on genetic and physiologic factors and our plots including eight tree species (Cowell 2004; West 2009). Along with genetic and physiologic factors, some secondary factors might also influence the TAP and STR patterns, such as, the stem form, stem density, edaphic condition, bark irregularities, as well as by the manual approach implemented for characterising the logs (Cowell 2004; West, 2009). Nevertheless, the variability of STR and TAP might be even influenced by the manual procedure used for measuring the advanced forest-related measurements in CloudCompare software (Henning and Radtke, 2006), and it became hard in trees with bark irregularities or infested by lianas (Liang and Hyypä, 2011; Wan *et al.* 2019). Hence, in our study, we assume that the variation of the advanced forest-related measurements can be linked to the management of forests, due to the structural heterogeneity played a crucial role in the accuracy of the advanced forest-related measurements, even if the manual approach used for estimating these from TLS data became crucial too.

As regards the classification of the merchantable logs, 134 out of 179 merchantable logs were classified in one of the 15 assortment types. The whole predicted logs were classified in 11 assortment types, so eleven out of 15 assortment types were correctly matched between predicted and observed data. The classification of merchantable logs was more accurate for eight assortment types. These 8 assortment types were included in saw-log and other industrial roundwood and someone else of other assortment types (i.e. A-, B-, B0, B+, C-, D-, D+ and Fuelwood-), which was ± 2 merchantable logs. Since the STR and Dmin were mandatory pre-requisites for classifying the merchantable logs, their accuracy affected the log quality. Therefore, if the dimension of each cylinder is overestimated, the log tapering will be overestimated too. In this context, the inaccuracy of the measurement of cylinders can promote a hindering factor. In our study, some of them can be the tree species richness, architecture of stems, irregularities of the bark (i.e. knots, lianas, bulges, microhabitats) (West 2009; Cowell 2004). Along with these hindering factors, the stem form (i.e. neiloidic, parabolic) and their eccentricities can even play an important role in describing the stem curve (Puletti *et al.* 2019; Luoma *et al.* 2019).

2.3.7. Conclusion

This study provides a stepwise procedure for extracting the timber assortment of standing trees using TLS data in forests characterized by a high tree species richness and heterogeneous stand structure. Results display that the approach provides significant insights for mixed and multi-layered forests allowing assessing the timber assortments for productive forests, but also to characterize timber volume within forests with conservative management aims, as carbon storage in old-growth forests or within protected areas. Furthermore, from a productive point of view, it represents a suitable approach to assess the timber assortment types within abandoned forests allowing their revalorization and utilization supporting the revalorization of socio-economic function within marginal mountain areas.

The stepwise approach offered the possibility to extract the timber assortments from standing trees of eight tree species using TLS data, especially the *Q. cerris*. From technical point of view, an accurate and realistic timber-leave discrimination was made through RF algorithm. The cylinder-fitting approach allowed us to reconstruct the trunk of stems, however, the main challenges making difficult the reconstruction were the stem form, presence of shrubs, lianas and microhabitat. Since our approach was tested, for the first time, in Mediterranean forest, especially in mixed-species and multi-layered forests, the comparison with other similar studies has not been possible. However, our approach proved to be a useful source for valorising the timber resource in an accurate way, even considering the conservative purpose of forest stands. The most accurate timber assortment assessment was found for saw-log and other industrial roundwood and someone else of other assortment types. The implementation of our approach for analysing TLS data could serve to better select the trees to be fell and cut, making more efficient the harvesting activities, therefore, it could ensure the timely and accurate forest decision towards the SFM. Further investigation to increase the knowledge about the applicability of this approach in other forest stand conditions could be useful for future studies.

2.3.8. Postface

This study proposes, for the first time, a stepwise approach for retrieving the qualitative and quantitative information of the timber assortment on TLS point cloud in mixed-species and multi-layered forests. Our stepwise approach proved to be efficient for separating the timber from leave points. This outcome allowed us to better identify and reconstruct the trunk of trees. All trees with a DBH higher than 30 cm were correctly identified. Moreover, the cylinder-fitting approach proved to be accurate in the reconstruction of eight trees species with a DBH higher 20 cm, especially for *Q. cerris*. The cylinder-fitting approach proved to be competitive in comparison with several similar algorithms, it allowed us to accurately reconstruct the trunk section of trees, regardless the stem form, presence of bulges, microhabitats. Our approach was proved to be accurate for quantifying and classifying the logs from the reconstructed trees, in fact, more than three-quarters were quantified and classified. However, the accuracy in classification was well performed for eight out of 15 types of assortment, but it can be associated with many aspects, such as, the observed data that has been manually estimated. Our approach represents a starting point for valorising the forest managed with conservative purposes.

3. CHAPTER 3 – CONCLUSION AND FUTURE PERSPECTIVES

The main objective of this thesis was to develop one or more approaches analysing LiDAR data for timber assortment assessment. LiDAR is a promising source for monitoring the forest because it now is available at worldwide (spaceborne LiDAR data is free data); the usability and versatility of ALS joint to the cost-effectiveness is increasingly being encouraging.

Literature review allowed a deeper understanding of the usability and versatility of LiDAR data for timber assortment evaluation. Five conclusions may be drawn from the literature review. First, there is great interest for retrieving forest inventory information and there is less interest for connecting inventory with biodiversity purposes; second, the most recommended LiDAR data to use for timber assortment investigation were the ALS and TLS, ALS because allows covering several hectares by each flight and TLS because was able to reconstruct the architecture of trees, regardless of the dimensions; third, better simulation of the vertical stratification of trees was favoured by the vertical splitting of the point cloud; fourth, the clustering is revealed as a promising approach for detecting the trees in forest covered by low point densities. Fifth, trees belonging to intermediate or lower canopy layers are a challenge to identifying using ALS data, and an approach for quantifying and classifying the timber assortment using TLS data is still required.

As outlined by the previous conclusions, a prerequisite to better use of ALS data for timber assortment assessment was associated with a reliable and accurate tree detection approach. We proposed a stepwise approach for carbon stock assessment at single tree level in mixed-species and multi-layered forests. Such an approach was subdivided into tree detection and carbon stock approaches. Four conclusions may be drawn from the stepwise approach used for analysing the ALS data. First, the combined use of two unsupervised algorithms proved to be appropriate for detecting trees in heterogeneous forest structure; second, the detection accuracy using our tree detection approach is favoured by the heterogeneity of forest structure and high point density (>30 points m^{-2}); third, our approach allowed to identify the trees without previous information of the tree position; fourth, a fit prediction of the carbon stock was achieved for all detected trees, therefore, the timber assortment assessment at single tree level can even be derived using our stepwise approach. Since one algorithm is used for detecting trees, named DBSCAN, better detection accuracy can be expected in forest areas with dense points.

The findings obtained using our stepwise approach for timber assortment assessment from TLS data provided many insights. In particular, the assessment of the timber resource from the productive forest, the selection of the trees with a lesser ecological weight for cutting process and the balance between productive and ecological aspects can be integrated in the planning and management of forest. Moreover, five conclusions may be drawn from the stepwise approach used for analysing the TLS data. First, accurate timber-leaves discrimination favoured the reconstruction of dominant trees species; second, most of the large trees (> 20 cm of DBH) were accurately identified; third, the cylinder-fitting approach proved to be

powerful for reconstructing the trunk section, especially the *Q. cerris*; fourth, the stem form, presence of shrubs, lianas and microhabitats were the main challenges hindering the stem reconstruction step; fifth, our approach proved to be more accurate in quantifying and classifying most of appreciated 15 assortment types, such as, pulpwood and other industrial roundwood. Since the main challenges found in this study were associated with forest stand conditions, we suggested that the forest stand condition (i.e. leaf-off canopy conditions) can favoured the stem reconstruction using our stepwise approach. Our approach better work in veteran trees, this is a significant outcome, because this approach is focused on trees with a greater timber volume.

In conclusion, the thesis provides two stepwise approaches using ALS and TLS data for timber assortment assessment and some suggestions for the better use of LiDAR data. As for example, about the ALS, the challenge associated with the vertical stratification of strata was overcome using a stratification approach, the challenge with the relationship between point density and forest structure was overcome in a heterogeneous forest stand, better conditions of the forest canopy condition (i.e. leaf-on) and the occlusion from large to small trees can facilitate the timber assortment assessment; about the TLS, the challenges were mainly associated with the stem form, irregularity of bark, microhabitats, however, these aspects can be used for discarding the trees, based on the assumption that tree with a greater productive aptitude holds a lesser ecological aptitudes. It is worth highlighting that the paper review offers a pool of approaches for analysing ALS data, and as many suggestions for increasing the accuracy of the ALS survey.

More efforts should be concentrated on connecting biodiversity and climate change indicators with inventory LiDAR campaign, because forests play a vital role in life and human wellbeing now and in future. In the light of the above, the finding of this thesis proved to be useful to accurately valorise the timber resource taking into account the conservative management of forests. The implementation of the approaches for analysing the LiDAR data can give many social and environmental benefits, as well as can facilitate the selection of trees to be fell and cut, making timber use more efficient. Further studies to increase the knowledge of the versatility of our approaches can be useful to deeper understand the potential under different forest stand conditions. The implementation of these methods on the rainforest and tropical forest can be beneficial to valorise the trees with high ecological aptitudes.

4. LIST OF ARTICLES

Paper I-III

- I. **Alvites Cesar**¹, Santopuoli Giovanni^{2,3}, Maesano Mauro⁴, Moresi Federico Valerio⁴, Roberto Tognetti, Marchetti Marco, Lasserre Bruno. Unsupervised algorithms to detect single trees in a mixed-species and multi-layered Mediterranean forest using LiDAR data. 2021. *Canadian Journal of Forest Research*.

Author contribution: Conceptualization, **A.C.** and S.G.; Data collection, S.G.;

Revision of literature, **A.C.**, S.G., M.M, M.F.V., R.T., M.M. and L.B.; Methodological development and data analysis, **A.C.** and S.G., writing and original draft preparation, **A.C.**; Analysis of results, C.A. and S.G.; Writing the manuscript, **A.C.** and S.G.; Review of manuscript, S.G., M.M, M.F.V., R.T., M.M. and L.B.

- II. **Alvites Cesar**, Santopuoli Giovanni, Marchetti Marco, Lasserre Bruno. Challenges and opportunities for timber assortment evaluation through LiDAR. A review. (submitted to Elsevier journal, International Journal of Applied Earth Observations and Geoinformation, deadline by 31 March 2022).

Author contribution: Conceptualization, **A.C.**, S.G. and B.L.; methodology, **A.C.** and S.G., writing and original draft preparation, **A.C.** and S.G, Revision of literature, **A.C.**, S.G., M.M., L.B.; Writing the paper review, **A.C.**, S.G.; Review of paper review, S.G., M.M., L.B.

- III. **Alvites Cesar**, Santopuoli Giovanni, Hollaus Markus, Pfeifer Norbert, Mauro Maesano, Moresi Federico Valerio, Marchetti Marco, Lasserre Bruno. Terrestrial Laser Scanning for quantifying timber assortments from standing trees in a mixed and multi-layered Mediterranean forest. 2021. Remote sensing.

Author contribution: Conceptualization, **A.C.**, S.G. and B.L.; Data collection, **A.C.**, S.G., M.M., M.F.V.; Revision of literature, **A.C.**, S.G., H.M., P.N.; Methodological development and data analysis, **A.C.**, H.M., P.N.; writing original draft preparation, **A.C.** and S.G., writing review and editing, G.S. and H.M., P.N., M.M, and M.F.V., supervision, M.M. and B.L.

5. ACKNOWLEDGEMENTS

Ph. D experience represented an important step in my life. I believe it helped me to increase my strengths and reduce weaknesses points, and to understand that behind every success there is a sacrifice.

Foremost, I would like to express my deep and sincere gratitude to my research supervisors, prof. Bruno Lasserre and researcher Giovanni Santopuoli from the University of Molise, for continuous supports given to my Ph. D study, for their motivation and sharing with me their experience, for the opportunity to learn to do research. Their guidance on the path of research helped me to overcome every obstacle during my Ph.D. experience and writing the thesis. Many thanks to Prof. Marco Marchetti and Prof. Roberto Tognetti, because besides being the first two persons to believe in my capabilities, they helped me to view the forests with a global and sustainable vision. In the light of the above, it was a great privilege to pursue my dream together with them.

Then, I would like to thank the Tuscia University, Department for Innovation in Biological, Agro-food and Forest systems (DIBAF), and the University of Molise, Department of Bioscience and Territory (DiBT) for the possibility to follow this Ph.D. course. Moreover, many thanks to Prof. Andrea Vannini (as Coordinator of my Ph.D. course), Prof. Maria Laura Carranza, and Prof. Anna Barbati (as Members of my Ph.D. Commission), for their constructive criticisms. I thank the reviewers of my thesis, Prof. Gherardo Chirici and researcher Ugo Chiavetta, who provided valuable comments that significantly improved the thesis.

I would also like to thank all components of the Ecogefor team, e.g. Daniela, Serena, Marco, Lorenzo, Marco di Cristofaro, Elena, Zaib-un, Elio and all for their friendship, for café time, for lunchtime, for grateful moments in every congress/workshop, for the beer after the work in forests and for many times of sharing.

And, last but not least, many thanks to my family: my mom for the life, my sister and brother for their suggestions. My special thanks go to my girlfriend Nadia for love and her support during these three years, and even for allowing me to find a new family.

6. REFERENCES

1. Antos J. (2009). Understory plants in temperate forests. *For. For. Plants.* 1: 262-279.
2. Barbati A, Marchetti M, Chirici G, Corona P. (2014). European forest types and forest Europe SFM indicators: tools for monitoring progress on forest biodiversity conservation. *Forest Ecology and Management.* 321: 145-157.
3. Beland M, Parker G, Sparrow B, Harding D, Chasmer L, Phinn S, Antonarakis A, Strahler A. (2019). On promoting the use of lidar systems in forest ecosystem research. *Forest Ecology and Management.* 450: 117484.
4. Chen Q. (2015). Modeling aboveground tree woody biomass using national-scale allometric methods and airborne lidar. *ISPRS Journal of Photogrammetry and Remote Sensing.* 106: 95-106.
5. Chianucci F, Puletti N, Grotti M, Ferrara C, Giorcelli A, Coaloa D, Tattoni C. (2020). Nondestructive tree stem and crown volume allometry in hybrid poplar plantations derived from terrestrial laser scanning. *Forest Science.* 66(6): 737-746.
6. Chirici G, Giannetti F, McRoberts RE, Travaglini D, Pecchi M, Maselli F, Chiesi M, Corona P. (2020). Wall-to-wall spatial prediction of growing stock volume based on Italian National Forest Inventory plots and remotely sensed data. *International Journal of Applied Earth Observation and Geoinformation.* 84: 101959.
7. Chirici G, McRoberts RE, Fattorini L, Mura M, Marchetti M. (2016). Comparing echo-based and canopy height model-based metrics for enhancing estimation of forest aboveground biomass in a model-assisted framework. *Remote Sensing of Environment.* 174: 1-9.
8. Dassot M, Constant T, Fournier M. (2011). The use of terrestrial LiDAR technology in forest science: application fields, benefits and challenges. *Annals of forest science.* 68(5): 959-974.
9. FOREST EUROPE. (2015). State of Europe's forests 2015.
10. SoEF. (2020). Summary for Policy Makers State of Europe's Forest 2020. (Vol. 4, pp. 64–75).
11. Giannetti F, Puletti N, Quatrini V, Travaglini D, Bottalico F, Corona P, Chirici G. (2018). Integrating terrestrial and airborne laser scanning for the assessment of single-tree attributes in Mediterranean forest stands. *European Journal of Remote Sensing.* 51(1): 795-807.
12. Hackenberg J, Spiecker H, Calders K, Disney M, Raunonen P. (2015). SimpleTree—an efficient open source tool to build tree models from TLS clouds. *Forests.* 6(11): 4245-4294.
13. Hamraz H, Contreras MA, Zhang J. (2017). Vertical stratification of forest canopy for segmentation of understory trees within small-footprint airborne LiDAR point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing.* 130: 385-392.
14. Holopainen M, Vastaranta M, Rasinmäki J, Kalliovirta J, Mäkinen A, Haapanen R, Melkas T, Yu X, Hyyppä J. (2010). Uncertainty in timber assortment estimates predicted from forest inventory data. *European Journal of Forest Research.* 129(6): 1131-1142.
15. Jukka M, Teppo P, Harri K, Tapio W, Erkki V. Timber Assortment Recovery Models for Southern Finland. *Baltic Forestry.* 16(1):102–12.

16. Jules MJ, Sawyer JO, Jules ES. (2008). Assessing the relationships between stand development and understory vegetation using a 420-year chronosequence. *Forest Ecology and Management*. 255(7): 2384-2393.
17. Junttila S. (2014). Effect of forest health and structure to the relative surface temperature captured by airborne thermal imagery: Case study in Norway Spruce-dominated stands in Southern Finland. Master's Thesis." (August).
18. Kankare V, Vauhkonen J, Tanhuanpää T, Holopainen M, Vastaranta M, Joensuu M, Krooks A, Hyypä J, Alho P, Viitala R. (2014). Accuracy in estimation of timber assortments and stem distribution—A comparison of airborne and terrestrial laser scanning techniques. *ISPRS Journal of Photogrammetry and Remote Sensing*. 97: 89-97.
19. Kelly M, Di Tommaso S. (2015). Mapping forests with Lidar provides flexible, accurate data with many uses. *California Agriculture*. 69(1): 14-20.
20. La Marca O, Notarangelo G. (2009). Influenza dei diradamenti sulla produzione legnosa in una fustaia di cerro (*Quercus cerris* L.) in Italia meridionale. *Forest@-Journal of Silviculture and Forest Ecology*. 6(3): 173.
21. Liang X, Wang Y, Pyörälä J, Lehtomäki M, Yu X, Kaartinen H, Kukko A, Honkavaara E, Issaoui AE, Nevalainen O, Vaaja M. (2019). Forest in situ observations using unmanned aerial vehicle as an alternative of terrestrial measurements. *Forest Ecosystems*. 6(1): 1-16.
22. Liang X, Hyypä J, Kaartinen H, Lehtomäki M, Pyörälä J, Pfeifer N, Holopainen M, Brolly G, Francesco P, Hackenberg J, Huang H. (2018). International benchmarking of terrestrial laser scanning approaches for forest inventories. *ISPRS journal of photogrammetry and remote sensing*. 144: 137-179.
23. Liu L, Pang Y, Li Z, Si, L, Liao S. (2017). Combining airborne and terrestrial laser scanning technologies to measure forest understorey volume. *Forests*. 8(4): 111.
24. Marchetti M, Motta R, Pettenella D, Sallustio L, Vacchiano G. (2018). Forests and forest-wood system in Italy: towards a new strategy to address local and global challenges. *Forest@ - Rivista Di Selvicoltura Ed Ecologia Forestale*. 15.
25. McRoberts RE, Næsset E, Gobakken T. (2015). Optimizing the k-Nearest Neighbors technique for estimating forest aboveground biomass using airborne laser scanning data. *Remote Sensing of Environment*. 163:13-22.
26. McRoberts RE, Tomppo EO. (2007). Remote sensing support for national forest inventories. *Remote sensing of environment*. 110(4): 412-419.
27. MEA. (2005). *Ecosystems and human well-being (Vol. 5)*. Washington, DC: Island Press. Pp. 1–155.
28. Michel P, Jenkins J, Mason N, Dickinson KJ, Jamieson IG. (2008). Assessing the ecological application of lasergrammetric techniques to measure fine-scale vegetation structure. *Ecological Informatics*. 3(4-5): 309-320.

29. Montaghi A, Corona P, Dalponte M, Gianelle D, Chirici G, Olsson H. (2013). Airborne laser scanning of forest resources: An overview of research in Italy as a commentary case study. *International Journal of Applied Earth Observation and Geoinformation*. 23:288-300.
30. Mura M, McRoberts RE, Chirici G, Marchetti M. (2016). Statistical inference for forest structural diversity indices using airborne laser scanning data and the k-Nearest Neighbors technique. *Remote Sensing of Environment*. 186:678-686.
31. Næsset E, Gobakken T, Solberg S, Gregoire TG, Nelson R, Ståhl G, Weydahl D. (2011). Model-assisted regional forest biomass estimation using LiDAR and InSAR as auxiliary data: A case study from a boreal forest area. *Remote Sensing of Environment*. 115(12): 3599-3614. <https://doi.org/10.1016/j.rse.2011.08.021>
32. Næsset E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote sensing of environment*. 80(1): 88-99.
33. Othmani AA, Jiang C, Lomenie N, Favreau JM, Piboule A, Voon LFLY. (2016). A novel computer-aided tree species identification method based on burst wind segmentation of 3d bark textures. *Machine Vision and Applications*. 27(5): 751-766.
34. Pfeifer N, Gorte B, Winterhalder D. (2004). Automatic reconstruction of single trees from terrestrial laser scanner data. In *Proceedings of 20th ISPRS Congress (Vol. 35, pp. 114-119)*. Istanbul: ISPRS.
35. Pinto I, Usenius A, Song T, Pereira H. (2005). Sawing simulation of maritime pine (*Pinus pinaster* Ait.) stems for production of heartwood containing components. *Forest products journal*. 55(4): 88-96.
36. Popescu SC. (2007). Estimating biomass of individual pine trees using airborne lidar. *Biomass and Bioenergy*. 31(9): 646-655.
37. Raumonon P, Kaasalainen M, Åkerblom M, Kaasalainen S, Kaartinen H, Vastaranta M, Holopainen M, Disney M, Lewis P. (2013). Fast automatic precision tree models from terrestrial laser scanner data. *Remote Sensing*. 5(2): 491-520.
38. Rehush N, Abegg M, Waser LT, Brändli U-B. (2018). Identifying Tree-Related Microhabitats in TLS Point Clouds Using Machine Learning. *Remote Sens*. 10, 1735.
39. Saarinen N, Kankare V, Pyörälä J, Yrttimaa T, Liang X, Wulder MA, Holopainen M, Hyypä J, Vastaranta M. (2019). Assessing the effects of sample size on parametrizing a taper curve equation and the resultant stem-volume estimates. *Forests*. 10(10):1–19.
40. Sačkov I, Santopuoli G, Bucha T, Lasserre B, Marchetti M. (2016). Forest inventory attribute prediction using lightweight aerial scanner data in a selected type of multilayered deciduous forest. *Forests*. 7(12): 307. <https://doi.org/10.3390/f7120307>
41. Sačkov I, Kulla L, Bucha T. (2019). A comparison of two tree detection methods for estimation of forest stand and ecological variables from airborne LiDAR data in central european forests. *Remote Sensing*. 11(12): 1431.

42. Santopuoli G, di Cristofaro M, Kraus D, Schuck A, Lasserre B, Marchetti M. (2019). Biodiversity conservation and wood production in a Natura 2000 Mediterranean forest. A trade-off evaluation focused on the occurrence of microhabitats. *iForest-Biogeosciences and Forestry*. 12(1): 76.
43. Santopuoli G, Di Febbraro M, Maesano M, Balsi M, Marchetti M, Lasserre B. (2020). Machine Learning Algorithms to Predict Tree-Related Microhabitats using Airborne Laser Scanning. *Remote Sens*. 12: 2142.
44. Silva CA, Klauberg C, Hudak AT, Vierling LA, Jaafar WS, Mohan M, Garcia M, Ferraz A, Cardil A, Saatchi S. (2017). Predicting stem total and assortment volumes in an industrial Pinus taeda L. forest plantation using airborne laser scanning data and random forest. *Forests*. 8(7): 254. <https://doi.org/10.3390/f8070254>
45. Smigaj M, Gaulton R, Suárez JC, Barr SL. (2019). Canopy temperature from an Unmanned Aerial Vehicle as an indicator of tree stress associated with red band needle blight severity. *Forest Ecology and Management*. 433: 699-708.
46. Togni M. (2017). Classificazione commerciale del legname grezzo tondo: regole per la classificazione visuale dei tronchi. Ente Regionale per i Servizi All'Agricoltura e Alle Foreste. 1:1-97.
47. Vastaranta M, Saarinen N, Kankare V, Holopainen M, Kaartinen H, Hyyppä J, Hyyppä H. (2014). Multisource single-tree inventory in the prediction of tree quality variables and logging recoveries. *Remote Sensing*. 6(4): 3475-3491. <https://doi.org/10.3390/rs6043475>
48. Walicka A, Pfeifer N, Jóźków G, Borkowski A. (2019). TLS POINT CLOUD REGISTRATION FOR DETECTING CHANGE IN INDIVIDUAL ROCKS OF A MOUNTAIN RIVER BED. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*. ISPRS Archives 42: 1149-54
49. Wan P, Wang T, Zhang W, Liang X, Skidmore AK, Yan G. (2019). Quantification of occlusions influencing the tree stem curve retrieving from single-scan terrestrial laser scanning data. *Forest Ecosystems*. 6(1): 1-13.
50. Wang Y, Hyyppä J, Liang X, Kaartinen H, Yu X, Lindberg E, Holmgren J, Qin Y, Mallet C, Ferraz A, Torabzadeh H. (2016). International benchmarking of the individual tree detection methods for modeling 3-D canopy structure for silviculture and forest ecology using airborne laser scanning. *IEEE Transactions on Geoscience and Remote Sensing*. 54(9): 5011-5027. <https://doi.org/10.1109/TGRS.2016.2543225>
51. West PW. (2009). *Tree and forest measurement*. Berlin: Springer. Jun 12.
52. White JC, Wulder MA, Buckmaster G. (2014). Validating estimates of merchantable volume from airborne laser scanning (ALS) data using weight scale data. *The Forestry Chronicle*. 90(3): 378-385. <https://doi.org/10.5558/tfc2014072>