

Unsupervised algorithms to detect single trees in a mixed-species and multilayered Mediterranean forest using LiDAR data¹

Cesar Alvites, Giovanni Santopuoli, Mauro Maesano, Gherardo Chirici, Federico Valerio Moresi, Roberto Tognetti, Marco Marchetti, and Bruno Lasserre

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 assess 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 of each stand was further divided into 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 a root mean square error ranging from 0.14% to 1.48%.

Key words: tree detection, airborne laser scanning (ALS), forest structure, carbon stock, climate-smart forestry.

Résumé : La mesure précise du capital forestier en croissance est un prérequis à l'implantation de stratégies forestières intelligentes face au climat. Cette étude porte sur l'utilisation de données de balayage laser aéroporté pour évaluer le stock de carbone à l'échelle de l'arbre. Elle vise à démontrer que l'utilisation combinée de deux techniques non supervisées va améliorer la précision de l'estimation sur laquelle s'appuie un aménagement forestier durable. Sur la base de l'hétérogénéité de la hauteur des arbres et de la densité du nuage de points, nous avons classé 31 peuplements forestiers dans quatre catégories de complexité. Le nuage de points de chaque peuplement a par la suite été divisé en trois couches horizontales, ce qui améliore la précision de la détection de chacun des arbres pour lesquels nous avons calculé le volume et le stock de carbone. La précision moyenne de la détection des arbres était de 0,48. La précision était plus élevée pour les peuplements forestiers qui avaient une plus faible densité et une fréquence plus élevée de gros arbres, ainsi qu'un nuage de points dense (0,65). La prédiction du stock de carbone était plus élevée avec un biais allant de -0,3 à 1,5 % et un écart moyen quadratique (EMQ) entre 0,14 et 1,48 %. [Traduit par la Rédaction]

Mots-clés : détection des arbres, balayage laser aéroporté (BLA), structure de la forêt, stock de carbone, foresterie intelligente face au climat.

1. Introduction

In Europe, forests cover about 35% of the total land area (SoEF 2020) 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). However, traditional forest

inventory methods are time-consuming and require enormous efforts, particularly in multilayered 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 in the face of climate change in threatened forest ecosystems, such as those of Mediterranean mountains. Information about forest area, forest damage, 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

Received 1 December 2020. Accepted 5 June 2021.

C. Alvites and B. Lasserre. Dipartimento di Bioscienze e Territorio, Università degli Studi del Molise, Cda Fonte Lappone snc, Pesche (IS) 86090, Italy.

G. Santopuoli and R. Tognetti. Dipartimento di Agricoltura, Ambiente e Alimenti, Università degli Studi del Molise, Via De Sanctis snc, Campobasso 86100, Italy; Centro di Ricerca per le Aree Interne e gli Appennini (ArIA), Università degli Studi del Molise, Via De Sanctis snc, Campobasso 86100, Italy.

M. Maesano and F.V. Moresi. Department of Innovation in Biological, Agro-food and Forest Systems, University of Tuscia, Viterbo 01100, Italy.

G. Chirici. Dipartimento di Scienze e Tecnologie Agrarie, Alimentari, Ambientali e Forestali, Università degli Studi di Firenze, Via San Bonaventura 13, Firenze 50145, Italy.

M. Marchetti. Dipartimento di Bioscienze e Territorio, Università degli Studi del Molise, Cda Fonte Lappone snc, Pesche (IS) 86090, Italy; Centro di Ricerca per le Aree Interne e gli Appennini (ArIA), Università degli Studi del Molise, Via De Sanctis snc, Campobasso 86100, Italy.

Corresponding author: Cesar Alvites (email: cesar.alvites@unimol.it).

¹This Article is part of a collection of papers presented at the CLImate-Smart Forestry in MOuntain Regions (CLIMO) workshop held in Stará Lesná, Slovakia, 9–11 September 2019.

© 2021 The Author(s). Permission for reuse (free in most cases) can be obtained from copyright.com.

et al. 2021), while maintaining the full set of ecosystem services — in short, climate-smart forestry (CSF). Bowditch et al. (2020) ranked the importance of different 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). Since the early 2000s, the use of light detection and ranging (LiDAR) has considerably increased in the forest sector, particularly 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 (Chirici et al. 2016), and even at the single tree level for mixed (Mongus and Žalik 2015) and pure stands (Shao et al. 2019). The accuracy of prediction is higher for the individual-tree-based approach compared with 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 remains very challenging, particularly for trees belonging to the understory layers of multilayered and mixed forests (Sačkov et al. 2016).

We propose that ALS may allow quantification and monitoring of 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 have shown 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. 2019b). Accordingly, better results at the single tree level were obtained in regular forest structures, such as pure conifer stands or forest plantations (Delponte et al. 2015; Torresan et al. 2020). Indeed, the individual-tree-based approach is difficult to implement in natural and unmanaged forests because of the challenges for deriving single-tree-related forest inventory variables, as exemplified by Duncanson et al. (2014) in their study of mature secondary upland and floodplain forests of USA, and by Kandare et al. (2016) in their study of mountain forests of Italian Alps, and such single-tree-related information serve as an important bench mark 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 the reFLex algorithm, with both showing several limitations for detecting understory vegetation layers. K-means detected 46% of trees with height lower than 12 m and reflex detected 18% of all trees in intermediate and suppressed layers. To improve the detection accuracy, some authors suggested to split 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. 2017b). Further approaches, such as the RANSAC (RANdom SAmple Consensus) algorithm (Balsi et al. 2018) and the MCGC (Multi-Class Graph Cut) method (Williams et al. 2020), 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 of <30 cm. In particular, the RANSAC algorithm allowed detection of about 86% of trees in the overstory layer, while MCGC method allowed detection of approximately 30% of trees in the understory layer. Overall, the accuracy of the detection rate is higher for trees belonging to the top canopy 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 than do those of the overstory layer, they are crucial for the resilience and the

stability of forests, and thus contribute to mitigate the effects of climate change (Antos 2009) and ensure the continuity of forest regeneration and successional processes (Jules et al. 2008).

In particular, describing the vertical structure of multilayered 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, individual-tree-based approaches for delineating vertically heterogeneous canopies remain difficult in application, because of the requirement of site-specific parameters and the geometry of multi-canopy layers (Hamraz et al. 2017b). Developing a suitable method for fostering the segmentation of trees in a multilayered 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 to assess carbon stock at the tree level in a mixed-species and multilayered forest, using ALS data. To reach this objective, we 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. Materials and methods

2.1. Study area

The study area is located in Central Italy (Molise; $41^{\circ}42'N$, $14^{\circ}12'E$), namely Bosco Pennataro (Fig. 1). Bosco Pennataro is recognized as part of the core area of the Collemeluccio-Montedimezzo Alto Molise Man and Biosphere (MaB) Reserve and included in the Natura 2000 network. Bosco Pennataro is a mixed Mediterranean forest with 13 tree species, with 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 *Artemonio agrimonoides* – *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., and the average annual precipitation and temperature are $723.5\text{ mm}\cdot\text{year}^{-1}$ and $14.5\text{ }^{\circ}\text{C}$, respectively (<https://power.larc.nasa.gov>). The current management system is high forest with continuous canopy cover and uneven-aged mixed-species trees. The average stand density is about $700\text{ trees}\cdot\text{ha}^{-1}$, the growing stock is $385\text{ m}^3\cdot\text{ha}^{-1}$, of which $366\text{ m}^3\cdot\text{ha}^{-1}$ are living trees and $19\text{ m}^3\cdot\text{ha}^{-1}$ are standing dead trees (Santopuoli et al. 2019).

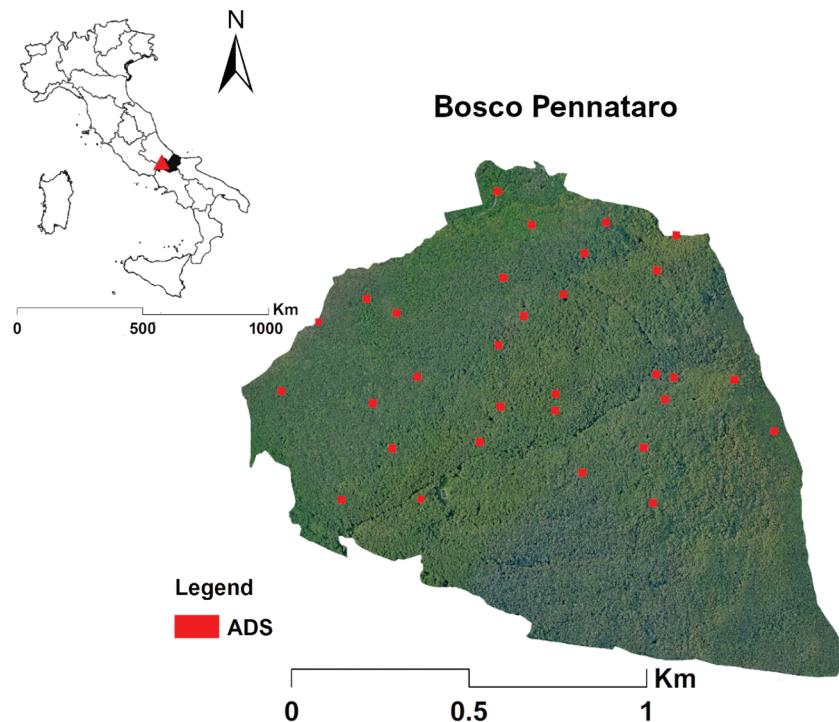
The absence of forestry interventions over the years has facilitated the conversion from even-aged to uneven-aged forest, supporting the shift of stand structure, from monolayer to multilayer.

For the field survey, we 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 random and uniform criterion. Based on one-per-stratum scheme, Bosco Pennataro was stratified into 50 strata. One squared plot (hereinafter ADS) of 529 m^2 per each 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 (Fig. 1).

2.2. Ground truth data

The forest-related characteristics within each ADS were collected in 2016, using 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

Fig. 1. Location of study area Bosco Pennataro (red triangle) and location of the field plots (ADS). This figure has been made using the Italian boundaries shapefile (<https://www.istat.it/>) through QGIS software, while Bosco Pennataro raster has been surveyed as part of the FRESHLIFE project “Demonstrating Remote Sensing integration in sustainable forest management” (LIFE14/IT000414). [Colour online.]



trees with a DBH ≥ 2.5 cm. The stem volume (VOL, m^3) 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 eq. 1:

$$(1) \quad AGB = GS \times BEF \times WBD \times A$$

where AGB is aboveground biomass (tons); GS is growing stock ($m^3 \cdot ha^{-1}$); BEF is biomass expansion factor, which is equal to 1.47; WBD is wood basal density ($t \text{ d.m. } m^{-3} \text{ f.v.}$), which is equal to 0.38; A is 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.

2.3. 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 ultralight vehicle able to collect three echoes per laser pulse, with an average point cloud density equal to 60 points· m^{-2} and accuracy equal to ± 5 cm ($\pm 50^\circ$ of Scan angle and pulse frequency of 20 kHz). However, most of the points belonged to the first echo. The ultralight vehicle flew at an altitude of 100 m above the 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 (Fig. 2).

2.3.1. Step 1 – Pre-processing of ALS data

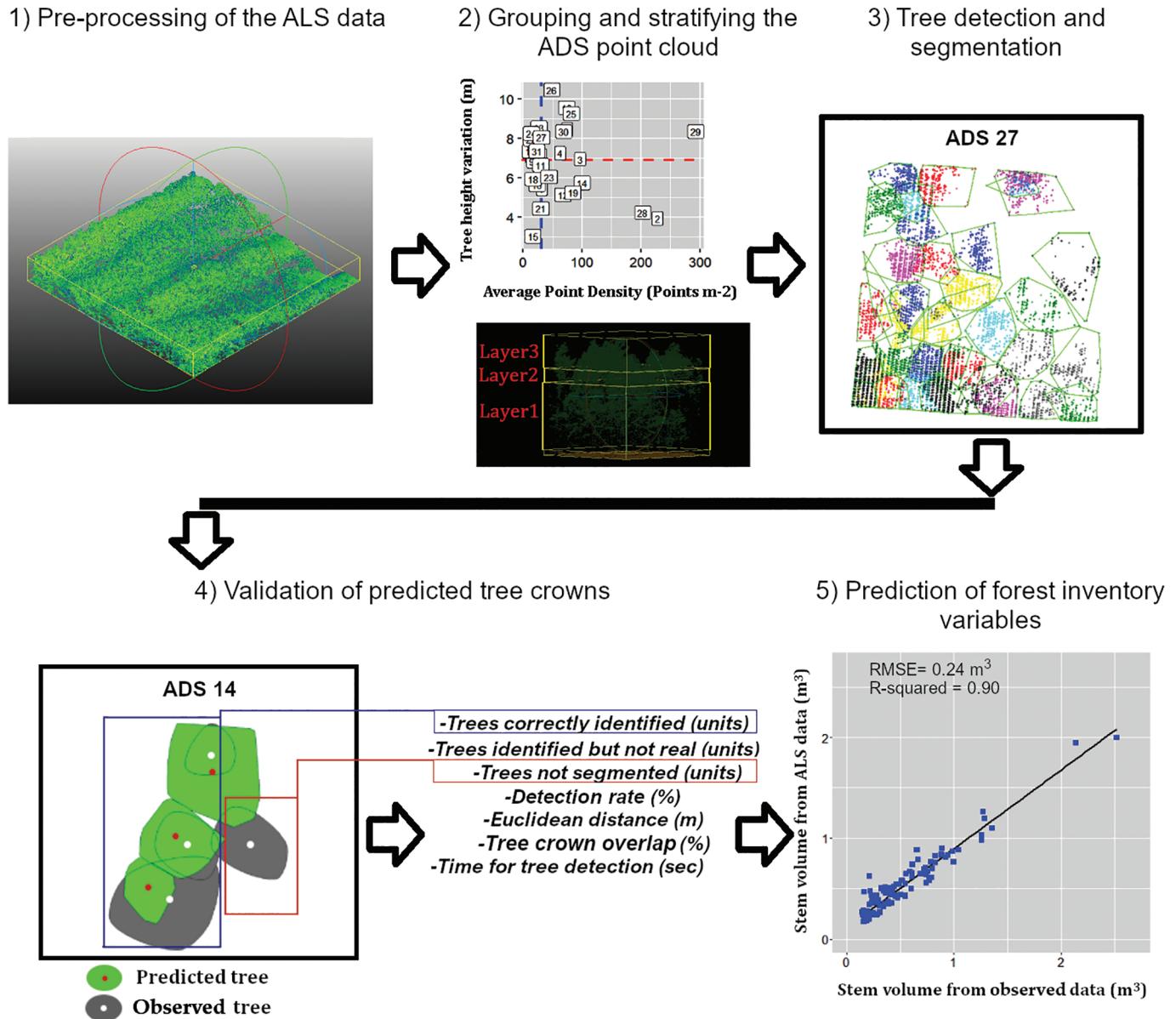
As part of the preprocessing step, the computing of ALS point cloud was carried out through several modules embedded in

LAStools software (www.rapidlasso.com). Initially, the raw ALS point cloud was classified in 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 ground surface, using “lasheight” module to derive a normalized aboveground point cloud source. The normalized aboveground 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 m^2 to 729 m^2 . The enlarged clipped point clouds for each ADS were used as input variables in the following steps.

2.3.2. Step 2 – Grouping and stratifying the ADS point cloud

To investigate factors influencing the accuracy of tree detection, due to the mixed-species and multilayered 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. 2017a, 2017b) and the standard deviation of surveyed tree heights (TH_{sd}) (Liang et al. 2019; Wang et al. 2019a), four groups containing uniform number of observations, i.e., ADS, were discriminated (Fig. 3). The value adopted as a threshold for APD was fixed at 31.02 points· m^{-2} , while for TH_{sd} the value was established at 6.879 m. Group A included the ADS that showed the lowest values of both APD and TH_{sd} ; group B included ADS with lowest values of APD and highest values of TH_{sd} ; group C included ADS with highest values of both APD and TH_{sd} ; group D included the ADS with highest values of APD and lowest values of TH_{sd} . The grouping process was achieved using “TreeLS” (available on

Fig. 2. 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 plots (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^2) values for tree volume prediction are also displayed. [Colour online.]



GitHub, <https://github.com/tiagodc/TreeLS>) and “stats” (authors, R Core Team, and contributors worldwide) R packages.

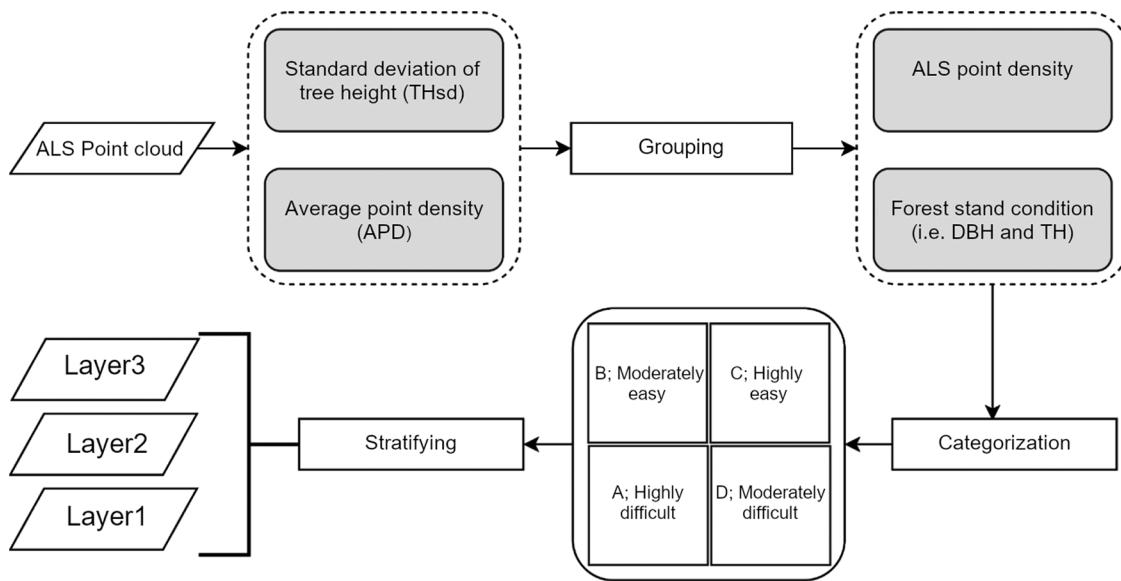
Moreover, the four groups were ranked in four complexity categories (“highly difficult”, “moderately difficult”, “highly easy”, “moderately easy”) (Liang et al. 2018, 2019; Wang et al. 2019b) to discriminate the accuracy of the detection approach within different forest structures (Fig. 3).

In detail, the ADS characterized by the highest number of trees, especially for small trees (DBH \leq 20 cm), as for example ADS of group A and D, were 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, especially for 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 homo-geneity within categories, which supports the assumption that an appropriate sample probability distribution of ADS was sampled (Barabesi et al. 2012).

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 sub-dominant trees, and Layer3 representing the dominant and codominant trees. The splitting procedure based on the vertical distribution of the tree heights, namely 33rd (Layer1), 66th (Layer2), and 99th (Layer3) percentiles (Fig. 3), was done using “lascanopy”

Fig. 3. Workflow of the processing of the airborne laser scanning (ALS) point cloud for each 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.



module available on LAStools software. The resulted tiled point clouds were used as input data for the tree detection and segmentation (step 3).

2.3.3. Step 3 – Tree detection and segmentation

To detect the stem position and to segment stem and crown of each 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 $N_{\text{Eps}}(p)$, can be derived from following the equation:

$$(2) \quad N_{\text{Eps}}(p) = [q \in D | \text{dist}(p, q) \leq \text{Eps}]$$

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

$N_{\text{Eps}}(p)$ 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 ($N_{\text{Eps}}(p) \geq \text{MinPts}$); the border is a point out of the core point but easy to be reachable ($p \in N_{\text{Eps}}(q)$); the noise point is an isolated point far away from the core point (Fig. 4). To define the core, border, and noise points, the DBSCAN algorithm plays an internal validation based on the density reachability and density connectivity (Fig. 4) (Ester et al. 1996; Smits et al. 2012).

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 and Wong 1979). The K-means equation is

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

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 three canopy layers to run the partition of the K -means processing, the DBSCAN was applied for each canopy layer (i.e., Layer1, Layer2, and Layer3) of 31 ADS point clouds.

K-means algorithm allowed us to delineate the tree crown boundaries of detected trees, using the “ K ” number of clusters derived by DBSCAN findings (Fig. 5).

Since the MinPts and $N_{\text{Eps}}(p)$ 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 “ $N_{\text{Eps}}(p)$ ” was set to 0.5 (Fig. 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. 2019).

Since the “ K ” number of clusters was provided by DBSCAN processing (Kandare et al. 2016; MacQueen 1967), 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 truly 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.

2.3.4. Step 4 – Validation of predicted tree crowns

The validation accuracy of the DBSCAN and K-means results was carried out following the most commonly used accuracy parameters in ALS detection studies (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

Fig. 4. 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 ($N_{Eps}(p)$) thresholds were considered. [Colour online.]

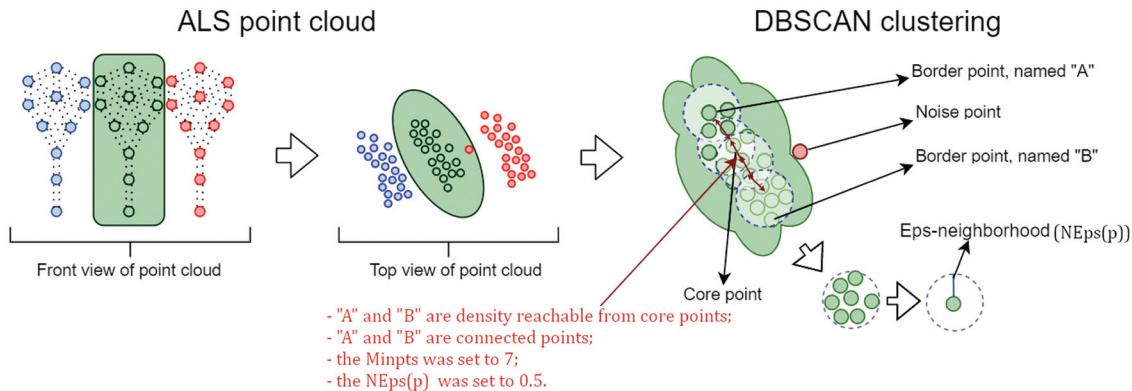
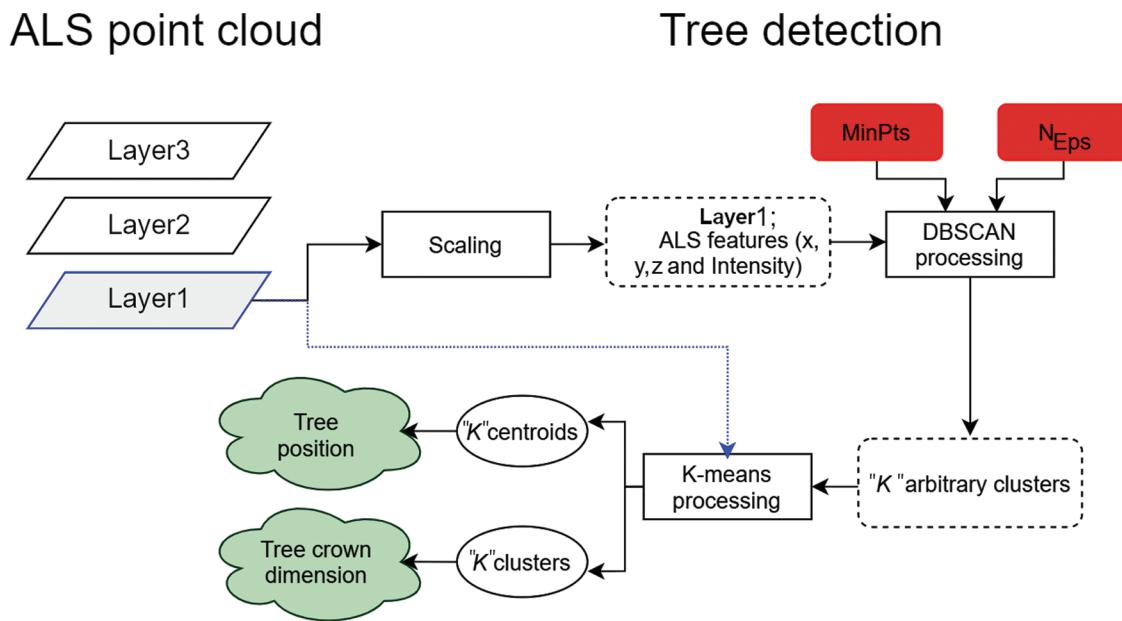


Fig. 5. Workflow of the processing for detecting the trees across the three canopy layers (i.e., lower layer, Layer1; intermediate layer, Layer2; and upper layer, Layer3) from airborne laser scanning (ALS) point cloud. The minimum number of points (MinPts) and Eps neighborhood distance (N_{Eps}) thresholds were used for processing Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and K-means. [Colour online.]



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 to 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 ADS of 729 m² using a personal compute Inter® Core™ i7-7500U CPU, 2.70 GHz and 8.00 GB RAM.

2.3.5. 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 (i) “Ntree”, the number of decision trees to be used during the prediction phase; (ii) “Mtry”, the number of input variables for splitting at each tree nodes; and (iii) “nodesize”, the minimum size of terminal nodes (Belgiu and Drăguț 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

Table 1. Summary of forest stand characteristics of airborne laser scanning (ALS) and forest inventory data per each field plot (ADS) according to complexity category.

	ADS	ALS data		Forest inventory data					
		APD (points·m ⁻²)	APS (m)	No. trees (ADS ⁻¹)	DBH (cm)	TH (m)	No. trees (ha ⁻¹)	VOL (m ³ ·ha ⁻¹)	CS (tons·ha ⁻¹)
Highly difficult	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	1 623	341.7	95.4
	15	16.3	0.3	196	9.9	8.2	3 698	183	51.1
	16	23.4	0.2	121	12.1	9.4	2 302	272.1	76
	18	16.8	0.2	95	12.9	9.7	1 792	277.3	77.4
	21	30	0.2	101	14	11.4	1 906	329	91.9
	Total	—	—	730	—	—	13 792	—	—
Moderately easy	Mean	21.9	0.2	91	14.6	12.4	1 724	285.6	79.8
	SD	7.3	0.1	52	3	3.3	985	60	16.8
	5	17.2	0.2	34	20	18.2	642	325.9	91
	13	26.5	0.2	83	16.5	14.6	1 566	488.3	136.4
	17	12.1	0.3	58	18.1	13.3	1 094	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	1 321	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	1 019	408.1	114
	31	23	0.2	63	17.1	14	1 189	435.2	121.5
Highly easy	Total	—	—	435	—	—	8 208	—	—
	Mean	19.75	0.24	54	17.75	14.88	1 026	436.56	121.93
	SD	6.78	0.05	18	2.32	2.14	339	103.06	28.77
	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	1 132	400	111.7
	30	68.7	0.1	60	16.5	13.9	1 132	413.6	115.5
Moderately difficult	Total	—	—	298	—	—	5 642	—	—
	Mean	100.3	0.1	43	20	17	806	436.7	122
	SD	85.6	0	14	3.6	3.8	262	55.5	15.5
	1	31.3	0.2	33	20.5	18.5	623	249	69.6
	2	227.6	0.1	120	12.8	9.8	2 264	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	2 642	286.8	80.1
	14	99.9	0.1	91	11.6	9.8	1 717	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	1 000	390.7	109.1
	28	202.7	0.1	132	10.7	10.8	2 491	220.2	61.5
	Total	—	—	654	—	—	12 340	—	—
	Mean	106.66	0.13	82	16.4	12.71	1 542	324.39	90.6
	SD	71.41	0.05	45	5.92	3.57	840	101.34	28.29

Note: 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 (No. trees) were calculated per ADS and Ha. The mean, total, and standard deviation values are shown.

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^2 ; 0–1) and root mean square error (RMSE; cm, m, m³) for the number of examined trees (No. trees; 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.

Table 2. Tree detection results.

Category	Canopy layers	TR	Tree detection				
			T _{ALS}	TruePos	FalsePos	FalseNeg	DR (%)
Highly difficult	Layer1	245	124	69	55	176	28
	Layer2	237	176	101	75	136	43
	Layer3	248	147	91	56	157	37
	Total	730	447	261	186	469	—
	Mean (\pm SD)	—	—	—	—	—	36 (\pm 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
	Total	435	321	215	106	220	—
	Mean (\pm SD)	—	—	—	—	—	49 (\pm 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
	Total	298	561	194	367	104	—
	Mean (\pm SD)	—	—	—	—	—	65 (\pm 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
	Total	654	418	282	136	372	—
	Mean (\pm SD)	—	—	—	—	—	43 (\pm 7.8)
Layer1	Total	706	522	280	242	426	—
Layer1	Mean (\pm SD)	—	—	—	—	—	42 (\pm 17.5)
Layer2	Total	687	656	346	310	341	—
Layer2	Mean (\pm SD)	—	—	—	—	—	54 (\pm 13.7)
Layer3	Total	724	569	326	243	398	—
Layer3	Mean (\pm SD)	—	—	—	—	—	49 (\pm 15.4)
Total	Total	2117	1747	952	795	1165	—
Total	Mean (\pm SD)	—	—	—	—	—	48 (\pm 12.5)

Note: Number of stems observed from reference data (TR; units) and number of stems predicted from ALS data (T_{ALS}; 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.

3. 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 and 26.9 cm, the mean TH ranged between 8.2 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 on the point density and spacing of ALS point clouds that varied from 12.13 to 292.9 points·m⁻² (Table 1).

3.1. ADS groups and ALS point cloud 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. The similar number of ADS for each group allowed us to better characterize the structural heterogeneity among ADS groups (Fig. 6; Table 1). Stand density was high in ADS of difficult categories, ranging between 1724 and 1542 trees·ha⁻¹. Moreover, these ADS presented high standard deviation values (985 and 840 trees·ha⁻¹), compared with those of easy categories (between 339 and 262 trees·ha⁻¹). Additionally, the easy categories were characterized by a great number of big trees compared with the difficult categories and, as a consequence, by high values of assessed forest inventory variables, i.e., DBH, TH, VOL, and CS.

Fig. 6. Graphical distribution of the field plots (ADS) according to the average point density (APD; points·m⁻²) and the standard deviation of tree height (TH_{sd}; m) for each category from A to D groups (A, highly difficult; B, moderately easy; C, highly easy; D, moderately difficult). [Colour online.]

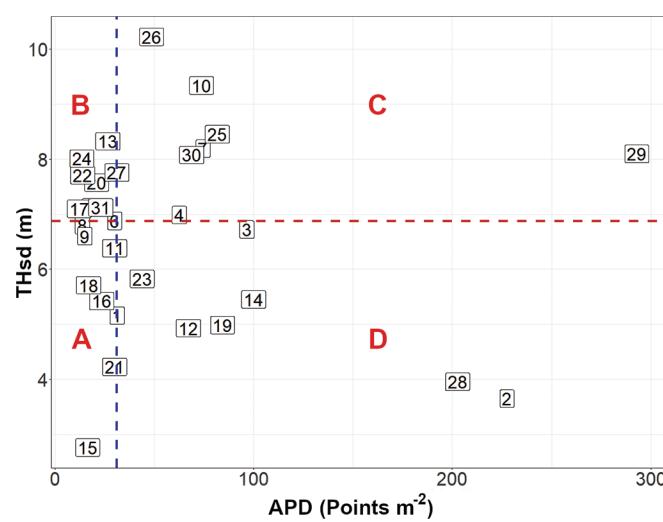
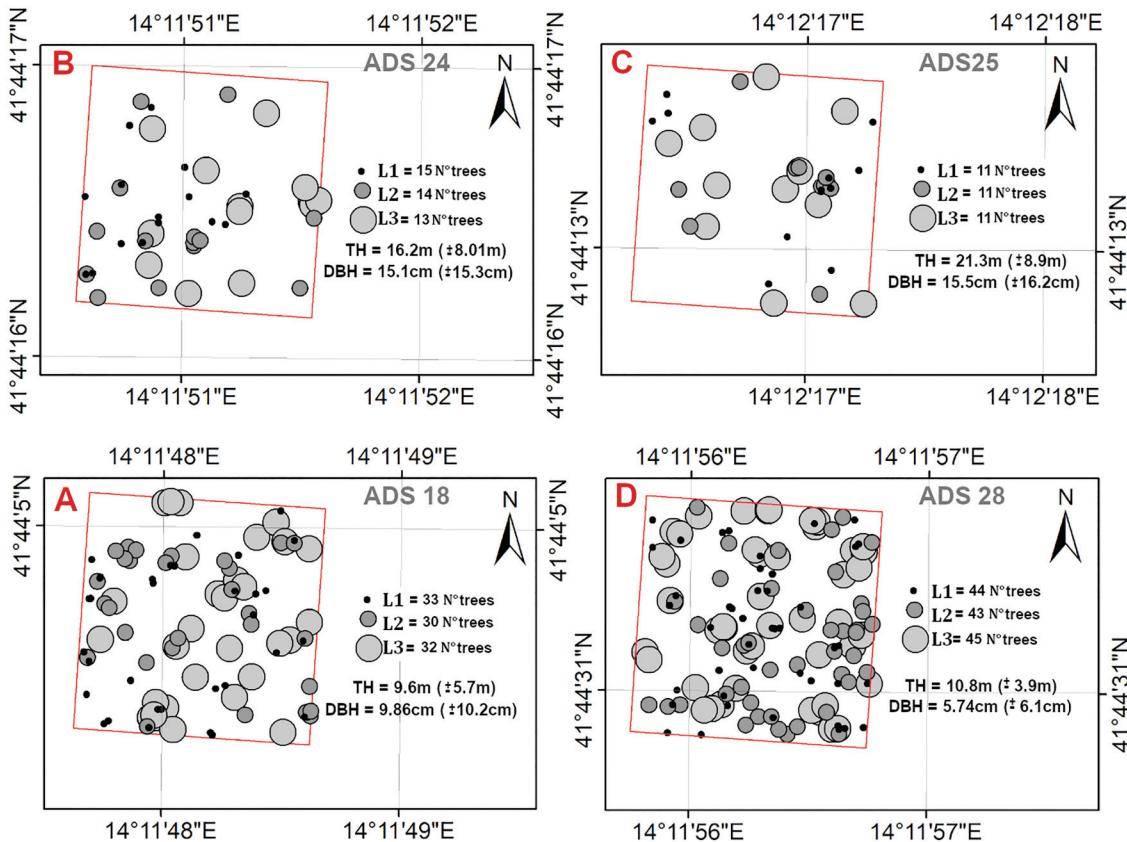


Fig. 7. Four representative maps of the four different airborne laser scanning (ALS) point cloud combinations (one per category). The red square showed the field plots (ADS) border; the number of trees (No. trees) is shown for every canopy layer (i.e., lower layer, Layer1 (L1); intermediate layer, Layer2 (L2); and upper layer, Layer3 (L3)); the tree height (TH; m) and the diameter at breast height (DBH; cm) are 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). [Colour online.]



The 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.

The number of trees across the three canopy layers was rather similar, from Layer1 to Layer3, with a relative low presence of trees in the Layer2 (Fig. 7). 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 with difficult categories. For this reason, the poor presence of trees, more accentuated in ADS of the “slightly easy” and “moderately easy” categories, was a contributing factor that enabled the discrimination of single trees (Figs. 7B and 7C); 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 (Fig. 7).

3.2. Tree detection

We detected 952 out of 2117 reference trees, reaching an average detection rate of 48% (Table 2), with a moderate uniformity or 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 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.

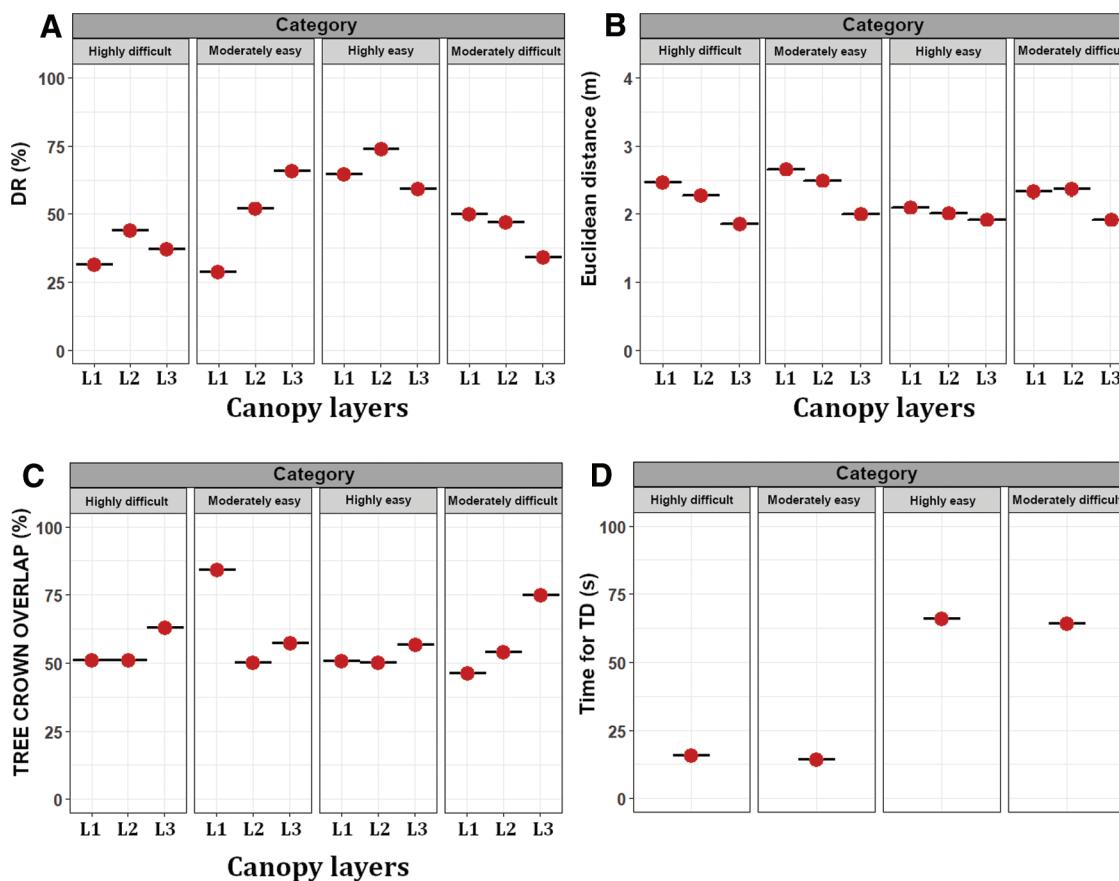
The detection rate values were more accurate for trees of the Layer2 (54%, $SD = \pm 13.7$) than for trees of the Layer1 (42%, $SD = \pm 7.8$) and Layer3 (49%, $SD = \pm 15.4$).

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 tree crowns; a better performance was obtained for trees of the Layer2 and Layer3. However, 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 compromises between commission and omission errors were found in ADS of “moderately easy”

Fig. 8. Comparison between predicted versus observed values of (A) detection rate (DR; %), (B) Euclidean distance (m), and (C) tree crown overlap (%) for each canopy layer (Layer1, L1; Layer2, L2; and Layer3, L3) and for every category (highly difficult, moderately easy, highly easy, and moderately difficult). (D) The average values of the time consumed for tree detection (time for TD, s) for each study area are displayed for each category. [Colour online.]



(106 and 220 out of 435 stems surveyed) and “highly easy” (367 and 104 out of 298 stems surveyed), respectively.

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

The estimation of the crown position displayed similar values for all four categories, ranging between 1.73 and 2.55 m (Fig. 8B). 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 amongst ADS ($SD = \pm 11$) (Fig. 8C), within which the Layer2 was the most accurate.

Time required in detecting the trees, using combined unsupervised algorithms, was faster in the ADS with lowest (ranged between 19.7 and 21.9 points·m⁻²) point density in comparison with those with the highest (ranged between 100.3 and 106.6 points·m⁻²) (Fig. 8D).

3.3. 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) (Fig. 9).

Despite the different quantity of trees analyzed (TruePos), slight differences in terms of coefficient of determination between

predicted vs. observed across categories were observed (Table 3). However, the categories were less accurate for DBH (No. of trees = 261 and 215; $R^2 = 0.9$) belonging to ADS of the “highly difficult” and “moderately easy” categories; whereas, for TH (No. of trees = 215; $R^2 = 0.93$) and VOL (No. of trees = 215; $R^2 = 0.89$), this was the case for the ADS belonging to the “moderately easy” category. Therefore, the categories with smaller point density (in absolute terms) were slightly less accurate.

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 4). However, we note that the “moderately difficult” category offered better performances than the “highly easy” category. Therefore, ADS with 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 and with a higher 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.

4. Discussion

4.1. Tree detection

Results revealed that the joint use of DBSCAN and K-means allowed detecting nearly half of the trees identified through ALS

Fig. 9. Predicted versus observed values for forest inventory variables: (A) diameter at breast height (DBH; cm), (B) tree height (TH; m), and (C) displays stem volume (VOL; m³). The number of trees (No. trees), number of predictors (No. predictors), coefficient of determination (R^2 ; 0–1), and root mean squared error (RMSE; cm, m, or m³) are also reported.

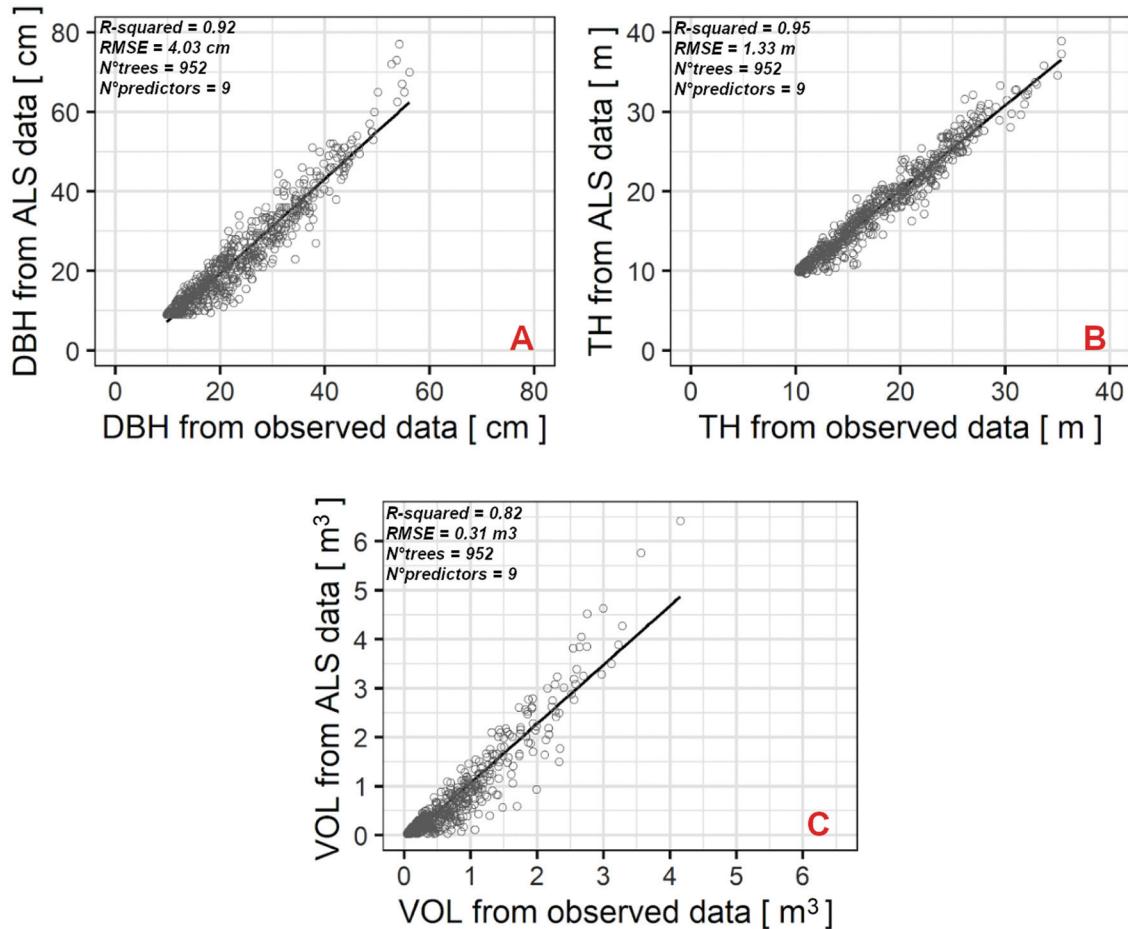


Table 3. Summary statistics of the forest inventory variables estimated with the Random Forests algorithm by using top-nine metrics for diameter at breast height (DBH; cm), tree height (TH; m), and stem volume (VOL; m³).

Category	Statistic measurement	Linear regression											
		DBH (cm)				TH (m)				VOL (m ³)			
Highly difficult	Layer1	Layer2	Layer3	Layer1-Layer3	Layer1	Layer2	Layer3	Layer1-Layer3	Layer1	Layer2	Layer3	Layer1-Layer3	
	No. of trees	69	101	91	261	69	101	91	261	69	101	91	261
	R^2	0.91	0.91	0.89	0.9	0.92	0.91	0.91	0.95	0.93	0.87	0.91	0.9
Moderately easy	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
	No. of trees	40	78	97	215	40	78	97	215	40	78	97	215
	R^2	0.91	0.89	0.88	0.9	0.92	0.92	0.89	0.93	0.78	0.89	0.87	0.89
Highly easy	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
	No. of trees	63	70	61	194	63	70	61	194	63	70	61	194
	R^2	0.82	0.89	0.91	0.91	0.82	0.93	0.9	0.95	0.8	0.89	0.88	0.9
Moderately difficult	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
	No. of trees	108	97	77	282	108	97	77	282	108	97	77	282
	R^2	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

Note: The number of trees (No. trees), coefficient of determination (R^2 ; 0–1), and root mean squared error (RMSE; cm, m and m³) are presented for all four categories (highly difficult, moderately easy, highly easy, and moderately difficult), and further divided by lower (Layer1), intermediate (Layer2), and upper (Layer3) canopy layers.

Table 4. Comparison between predicted and observed values of stem volume (VOL; m³) and carbon stock (CS; tons) derived from airborne laser scanning (ALS) metrics.

Stem volume and carbon stock prediction											
Category	ID	No. of trees		VOL (m ³ ·ha ⁻¹)		CS (tons·ha ⁻¹)		VOL Bias (m ³ ·ha ⁻¹)	CS Bias (tons·ha ⁻¹)	VOL and CS Bias (%)	RMSE (%)
		ADS	ha ⁻¹	Observed	Predicted	Observed	Predicted				
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	—	—	—	—
	Total	261	4924	—	—	—	—	—	—	—	—
Moderately easy	Mean	—	—	165	167.5	46.1	46.8	—	—	—	—
	Accuracy	—	—	—	—	—	—	-2.4	-0.7	1.5	1.48
Highly 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	—	—	—	—
	Total	215	4056	—	—	—	—	—	—	—	—
Moderately difficult	Mean	—	—	309.6	310	86.5	86.6	—	—	—	—
	Accuracy	—	—	—	—	—	—	-0.4	-0.1	0.1	0.14
Moderately difficult	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	—	—	—	—
	Total	194	3659	—	—	—	—	—	—	—	—
	Mean	—	—	327.8	330.5	91.6	92.3	—	—	—	—
	Accuracy	—	—	—	—	—	—	-2.7	-0.8	0.8	0.83

Note: The data are divided by category (highly difficult, moderately easy, highly easy, and moderately difficult) and by field plot (ADS). The number of trees (No. trees) was calculated per ADS and per hectare. Absolute (m³ and tons) and percent (%) values of bias and root mean squared error (RMSE) are presented.

data in the studied multilayered and mixed-species Mediterranean mountain forest. Enhanced detection accuracy was obtained in forest stands 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 multilayered 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 (Bowditch et al. 2020; Santopuoli et al. 2021).

Though the detection was challenging for trees of the under-story layer, results obtained here were somewhat encouraging in comparison with those reported by other authors (Table 2). For example, Sackov 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. (2017b) observed that the accuracy of tree detection

decreased from dominant to suppressed trees and highlighted that a dense point clouds was required for a satisfactory detection. The LiDAR point clouds used here had an average of 60 points·m⁻², ranging between 21 and 106 points·m⁻². Nevertheless, the choice to split the point clouds into three 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 (Hamraz et al. 2017b), the forest stand conditions e.g., stand density (Kandare et al. 2016) and DBH classes (Williams et al. 2020), and the site-specific parameters, e.g., tree species composition (Liang et al. 2019) and forest structure (Sačkov et al. 2016) impacted the identification of trees, as well as the detection rate, and commission and omission errors. As a consequence, 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·m⁻². In particular, the detection accuracy was further worsened in ADS of this Mediterranean mountain mixed-species and multilayered 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. 2020).

It is important to note that, though the detection accuracy was higher for trees belonging to Layer2 and Layer3, a better compromise between omission and commission errors was found in the Layer1 (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 of outlier. Dense forest stands may hinder the correct separation between nearby trees (Dalponte et al. 2015). 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 multilayered 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. (2019), 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 revealed by these authors, for which the values of Euclidean distance decreased from Layer3 to Layer1, 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 multilayered mixed-species forests of this type.

4.2. 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 (Fig. 9). Though the feasibility in the prediction approach was tested in four complexity levels, there were not substantial differences in the prediction accuracy among them. Such a 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 subdividing the ALS data in three

canopy layers might describe thoroughly the forest inventory variables for trees within each canopy layer, especially for trees of Layer2 and Layer1.

The performance of VOL models was more accurate using the information of whole TruePos (Layer1–Layer3) compared with the TruePos of the Layer3, based on the RMSE measurements found in all four categories (Table 3). 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 an R^2 value of 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 (Alberti et al. 2013; Sačkov et al. 2016).

The prediction accuracy for carbon stock was more accurate in the ADS with high ALS point density (“moderately difficult”; bias = -0.3%) and low ALS point density (“moderately easy”; bias = 0.1%), but with 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 structure. Results obtained for the stem volume (where input data to derive the carbon stock was ranged between 0.89 and 0.90 of R^2) were in line with those observed by other authors. For example, Popescu (2007) showed higher R^2 values for aboveground biomass in mature stands of loblolly pine, ranging between 0.88 and 0.93, whereas Allouis et al. (2013) reported higher R^2 values of aboveground 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 tree species richness (White et al. 2014).

Overall, the accuracy of tree detection and carbon stock assessment was found to be more sensitive to point density than to heterogeneity of forest structure (Tables 2 and 4). This means that further efforts focused on improving the quality of points will be beneficial for better exploiting the potential of tested algorithms. In particular, we found many weak aspects during the ALS processing. For example, the ADS point clouds characterized by lower, altered, and irregularly spaced densities were hard to process by DBSCAN algorithm; fixed values of MinPts and Eps-neighborhood became disadvantageous for identifying the trees in ADS from difficult categories; analyzing the ADS with dense point clouds is time consuming. These weaknesses 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 of several operational activities could be beneficial to overcome some of these issues, especially before the collection phase: (i) forest canopy structure (changing from leaf-on to leaf-off) (Shao et al. 2019), (ii) flight strips (changing

from 0% to more than 50% of overlapped flight strips) (Liang et al. 2019), and (iii) ALS sensor (changing from 3 echoes to 4–15 echoes) (Kandare et al. 2016; Hamraz et al. 2017b). Since our ultra-light vehicle flew at an altitude of 100 m above the 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.

5. Conclusion

This study aimed to improve the use of ALS data for the prediction of forest inventory variables in mixed-species and multilayered 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 found for those stands with fewer than 30 points·m⁻². ALS composed primarily 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 CSF 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 profile and big tree size.

The application of unsupervised algorithms for detecting single trees in a mixed-species and multilayered 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 structure, allowing forest operators to detect and monitor a large number of trees.

Acknowledgements

The study was supported by the FRESH LIFE project “Demonstrating Remote Sensing integration in sustainable forest management” (LIFE14/IT000414). We thank Markus Hollaus for helpful feedback aimed at improving the quality of the manuscript. This study was finalized in the framework of the COST (European Cooperation in Science and Technology) Action CLIMO (Climate-Smart Forestry in Mountain Regions – CA15226) and financially supported by the EU Framework Programme for Research and Innovation HORIZON 2020.

References

- Ahmad, P.H., and Dang, S. 2015. Performance evaluation of clustering algorithm using different datasets. *Int. J. Adv. Res. Comput. Sci. Manage. Stud.* **8**: 167–173. Available from www.ijarcsms.com [accessed 28 March 2021].
- Allouis, T., Durrieu, S., Véga, C., and Couteron, P. 2013. Stem volume and above-ground biomass estimation of individual pine trees from LiDAR data: contribution of full-waveform signals. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **6**(2): 924–934. doi:[10.1109/JSTARS.2012.2211863](https://doi.org/10.1109/JSTARS.2012.2211863).
- Antos, J. 2009. Understory plants in temperate forests. In *Forests and forest plants*. Department of Biology, University of Victoria, Victoria, BC. pp. 262–279.
- Alberti, G., Boscutti, F., Pirotti, F., Bertacco, C., De Simon, G., Sigura, M., et al. 2013. A LiDAR-based approach for a multi-purpose characterization of Alpine forests: an Italian case study. *iForest*, **6**(3): 156–168. doi:[10.3832/ifor0876-006](https://doi.org/10.3832/ifor0876-006).
- Balsi, M., Esposito, S., Fallavollita, P., and Nardinocchi, C. 2018. Single-tree detection in high-density LiDAR data from UAV-based survey. *Eur. J. Remote Sens.* **51**(1): 679–692. doi:[10.1080/22797254.2018.1474722](https://doi.org/10.1080/22797254.2018.1474722).
- Barabesi, L., Franceschi, S., and Marcheselli, M. 2012. Properties of design-based estimation under stratified spatial sampling with application to canopy coverage estimation. *Ann. Appl. Stat.* **6**(1): 210–228. doi:[10.1214/11-AOAS509](https://doi.org/10.1214/11-AOAS509).
- Barbati, A., Marchetti, M., Chirici, G., and Corona, P. 2014. European Forest Types and Forest Europe SFM indicators: tools for monitoring progress on forest biodiversity conservation. *For. Ecol. Manage.* **321**: 145–157. doi:[10.1016/j.foreco.2013.07.004](https://doi.org/10.1016/j.foreco.2013.07.004).
- Belgiu, M., and Drăguț, L. 2016. Random forest in remote sensing: a review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **114**: 24–31. doi:[10.1016/j.isprsjprs.2016.01.011](https://doi.org/10.1016/j.isprsjprs.2016.01.011).
- Biondi, E., Blasi, C., Burrasciano, S., Casavecchia, S., Copiz, R., Del Vico, E., et al. 2010. Manuale italiano di interpretazione degli habitat (Direttiva 92/43/CEE) [Italian guidelines for habitat interpretation (Directive 92/43/EE)]. MATTM. Available from <http://vnr.unipg.it/habitat/> [accessed 15 March 2021].
- Bivand, R., and Rowlingson, B. 2016. Package ‘rgdal’. R Package. Available from <https://cran.r-project.org/web/packages/rgdal/index.html> [accessed 15 January 2019].
- Bowditch, E., Santopuoli, G., Binder, F., del Rio, M., La Porta, N., Kluvankova, T., et al. 2020. What is Climate-Smart Forestry? A definition from a multinational collaborative process focused on mountain regions of Europe. *Ecosyst. Serv.* **43**(1): 101113–101113. doi:[10.1016/j.ecoser.2020.101113](https://doi.org/10.1016/j.ecoser.2020.101113).
- Breiman, L. 2001. Random forests. *Machine Learning*, **45**: 5–32. doi:[10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324).
- Chirici, G., Barbati, A., Corona, P., Marchetti, M., Travaglini, D., Maselli, F., and Bertini, R. 2008. Non-parametric and parametric methods using satellite images for estimating growing stock volume in alpine and Mediterranean forest ecosystems. *Remote Sens. Environ.* **112**(5): 2686–2700. doi:[10.1016/j.rse.2008.01.002](https://doi.org/10.1016/j.rse.2008.01.002).
- Chirici, G., McRoberts, R.E., Fattorini, L., Mura, M., and 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 Sen. Environ.* **174**: 1–9. doi:[10.1016/j.rse.2015.11.010](https://doi.org/10.1016/j.rse.2015.11.010).
- Dalponte, M., Reyes, F., Kandare, K., and Gianelle, D. 2015. Delineation of individual tree crowns from ALS and hyperspectral data: a comparison among four methods. *Eur. J. Remote Sens.* **48**(1): 365–382. doi:[10.5721/EuJRS20154821](https://doi.org/10.5721/EuJRS20154821).
- Duncanson, L.I., Cook, B.D., Hurt, G.C., and Dubayah, R.O. 2014. An efficient, multi-layered crown delineation algorithm for mapping individual tree structure across multiple ecosystems. *Remote Sens. Environ.* **154**: 378–386. doi:[10.1016/j.rse.2013.07.044](https://doi.org/10.1016/j.rse.2013.07.044).
- Ester, M., Kriegel, H.P., Sander, J., and Xu, X. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. *KDD-96 Proc.* **96**(34): 226–231. doi:[10.1016/B978-044452701-1.00067-3](https://doi.org/10.1016/B978-044452701-1.00067-3).
- Federici, S., Vitullo, M., Tulipano, S., De Lauretis, R., and Seufert, G. 2008. An approach to estimate carbon stocks change in forest carbon pools under the UNFCCC: the Italian case. *iForest*, **1**(2): 86–95. doi:[10.3832/ifor0457_010086](https://doi.org/10.3832/ifor0457_010086).
- Ferrara, R., Virdis, S.G., Ventura, A., Ghisu, T., Duce, P., and Pellizzaro, G. 2018. An automated approach for wood-leaf separation from terrestrial LiDAR point clouds using the density based clustering algorithm DBSCAN. *Agric. For. Meteorol.* **262**: 434–444. doi:[10.1016/j.agrformet.2018.04.008](https://doi.org/10.1016/j.agrformet.2018.04.008).
- Hahsler, M., Piekenbrock, M., and Doran, D. 2019. dbscan: Fast density-based clustering with R. *J. Stat. Soft.* **91**(1): 1–30. doi:[10.18637/jss.v091.i01](https://doi.org/10.18637/jss.v091.i01).
- Hamraz, H., Contreras, M.A., and Zhang, J. 2017a. Forest understory trees can be segmented accurately using sufficiently dense airborne laser scanning point clouds. *Sci. Rep.* **7**(1): 1–9. doi:[10.1038/s41598-017-07200-0](https://doi.org/10.1038/s41598-017-07200-0).
- Hamraz, H., Contreras, M.A., and Zhang, J. 2017b. Vertical stratification of forest canopy for segmentation of understory trees within small-footprint airborne LiDAR point clouds. *ISPRS J. Photogramm. Remote Sens.* **130**: 385–392. doi:[10.1016/j.isprsjprs.2017.07.001](https://doi.org/10.1016/j.isprsjprs.2017.07.001).
- Hartigan, J.A., and Wong, M.A. 1979. Algorithm AS 136: a k-means clustering algorithm. *J. R. Stat. Soc. Series C (Appl. Stat.)*, **28**(1): 100–108. doi:[10.2307/2346830](https://doi.org/10.2307/2346830).
- Jules, M.J., Sawyer, J.O., and Jules, E.S. 2008. Assessing the relationships between stand development and understory vegetation using a 420-year chronosequence. *For. Ecol. Manage.* **255**(7): 2384–2393. doi:[10.1016/j.foreco.2007.12.042](https://doi.org/10.1016/j.foreco.2007.12.042).
- Kandare, K., Dalponte, M., Gianelle, D., and Chan, J.C.W. 2014. A new procedure for identifying single trees in understory layer using discrete LiDAR data. In *Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium*, Quebec City, Que., 13–18 July 2014. IEEE. pp. 1357–1360. doi:[10.1109/IGARSS.2014.6946686](https://doi.org/10.1109/IGARSS.2014.6946686).
- Kandare, K., Örka, H.O., Chan, J.C.W., and Dalponte, M. 2016. Effects of forest structure and airborne laser scanning point cloud density on 3D delineation of individual tree crowns. *Eur. J. Remote Sens.* **49**(1): 337–359. doi:[10.5721/EuJRS20164919](https://doi.org/10.5721/EuJRS20164919).

- Kwak, J. 2014. Akmeans: adaptive K-means algorithm based on threshold. R package version 1.1. Available from <https://cran.r-project.org/web/packages/akmeans/index.html> [accessed 15 October 2019].
- Liang, X., Hyppä, J., Kaartinen, H., Lehtomäki, M., Pyörälä, J., Pfeifer, N., et al. 2018. International benchmarking of terrestrial laser scanning approaches for forest inventories. *ISPRS J. Photogram. Remote Sens.* **144**: 137–179. doi:[10.1016/j.isprsjprs.2018.06.021](https://doi.org/10.1016/j.isprsjprs.2018.06.021).
- Liang, X., Wang, Y., Pyörälä, J., Lehtomäki, M., Yu, X., Kaartinen, H., et al. 2019. Forest in situ observations using unmanned aerial vehicle as an alternative of terrestrial measurements. *For. Ecosyst.* **6**(1): 1–16. doi:[10.1186/s40663-019-0173-3](https://doi.org/10.1186/s40663-019-0173-3).
- Liaw, A., and Wiener, M. 2002. Classification and regression by randomForest. *R News*, **2**(3): 18–22. Available from <https://cran.r-project.org/web/packages/randomForest/randomForest.pdf> [accessed 18 October 2020].
- Lindner, M., and Karjalainen, T. 2007. Carbon inventory methods and carbon mitigation potentials of forests in Europe: a short review of recent progress. *Eur. J. For. Res.* **126**(2): 149–156. doi:[10.1007/s10342-006-0161-3](https://doi.org/10.1007/s10342-006-0161-3).
- MacQueen, J. 1967. Some methods for classification and analysis of multivariate observations. In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, Statistical Laboratory University of California, Davis, Calif., 21 June – 18 July 1965, and 27 December 1965 – 7 January 1966. University of California Press, Berkeley, Calif. pp. 281–297.
- Mongus, D., and Žalik, B. 2015. An efficient approach to 3D single tree-crown delineation in LiDAR data. *ISPRS J. Photogram. Remote Sens.* **108**: 219–233. doi:[10.1016/j.isprsjprs.2015.08.004](https://doi.org/10.1016/j.isprsjprs.2015.08.004).
- Nabuurs, G.J., Verkerk, P.J., Schelhaas, M., González-Olabarria, J.R., Trasobares, A., and Cienciala, E. 2018. Climate-Smart Forestry: mitigation impact in three European regions. From Science to Policy 6. European Forest Institute, Joensuu, Finland. pp. 1–32.
- Naimi, B. 2017. R package usdm: uncertainty analysis for species distribution models. Available from <https://cran.r-project.org/web/packages/usdm/usdm.pdf> [accessed 18 October 2020].
- Næsset, E. 1997. Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS J. Photogram. Remote Sens.* **52**(2): 49–56. doi:[10.1016/S0924-2716\(97\)83000-6](https://doi.org/10.1016/S0924-2716(97)83000-6).
- Popescu, S.C. 2007. Estimating biomass of individual pine trees using airborne lidar. *Biomass Bioenergy*, **31**(9): 646–655. doi:[10.1016/j.biombioe.2007.06.022](https://doi.org/10.1016/j.biombioe.2007.06.022).
- Sačkov, I., Santopuoli, G., Bucha, T., Lasserre, B., and Marchetti, M. 2016. Forest inventory attribute prediction using lightweight aerial scanner data in a selected type of multilayered deciduous forest. *Forests*, **7**(12): 307. doi:[10.3390/f7120307](https://doi.org/10.3390/f7120307).
- Sačkov, I., Kulla, L., and 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 Sens.* **11**(12): 1431. doi:[10.3390/rs11121431](https://doi.org/10.3390/rs11121431).
- Santopuoli, G., di Cristofaro, M., Kraus, D., Schuck, A., Lasserre, B., and 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*, **12**(1): 76–84. doi:[10.3832/ifor2617_011](https://doi.org/10.3832/ifor2617_011).
- Santopuoli, G., Di Febbraro, M., Maesano, M., Balsi, M., Marchetti, M., and Lasserre, B. 2020. Machine learning algorithms to predict tree-related microhabitats using airborne laser scanning. *Remote Sens.* **12**: 2142. doi:[10.3390/rs12132142](https://doi.org/10.3390/rs12132142).
- Santopuoli, G., Temperli, C., Alberdi, I., Barbeito, I., Bosela, M., Bottero, A., et al. 2021. Pan-European sustainable forest management indicators for assessing Climate-Smart Forestry in Europe. *Can. J. For. Res.* [This issue]. doi:[10.1139/cjfr-2020-0166](https://doi.org/10.1139/cjfr-2020-0166).
- Shao, G., Shao, G., and Fei, S. 2019. Delineation of individual deciduous trees in plantations with low-density LiDAR data. *Int. J. Remote Sens.* **40**(1): 346–363. doi:[10.1080/01431161.2018.1513664](https://doi.org/10.1080/01431161.2018.1513664).
- Shi, Y., Wang, T., Skidmore, A.K., and Heurich, M. 2018. Important LiDAR metrics for discriminating forest tree species in Central Europe. *ISPRS J. Photogram. Remote Sens.* **137**: 163–174. doi:[10.1016/j.isprsjprs.2018.02.002](https://doi.org/10.1016/j.isprsjprs.2018.02.002).
- Silva, C.A., Crookston, N.L., Hudak, A.T., and Vierling, L.A., 2015. rLiDAR: an R package for reading, processing and visualizing LiDAR (Light Detection and Ranging) data. Available from <http://cran.r-project.org/web/packages/rLiDAR/index.html> [accessed 15 January 2019].
- Smits, I., Prieditis, G., Dagis, S., and Dubrovskis, D. 2012. Individual tree identification using different LiDAR and optical imagery data processing methods. *Biosyst. Inf. Technol.* **1**: 19–24. doi:[10.11592/bit.121103](https://doi.org/10.11592/bit.121103).
- SoEF. 2020. Summary for Policy Markers: State of Europe's Forest 2020 (SoEF 2020). Vol. 4. Ministerial Conference on the Protection of Forests in Europe – FOREST EUROPE Liaison Unit Bratislava, Zvolen, Slovak Republic. pp. 64–75.
- Tabacchi, G., Di Cosmo, L., and Gasparini, P. 2011. Aboveground tree volume and phytomass prediction equations for forest species in Italy. *Eur. J. For. Res.* **130**(6): 911–934. doi:[10.1007/s10342-011-0481-9](https://doi.org/10.1007/s10342-011-0481-9).
- Torresan, C., Carotenuto, F., Chiavetta, U., Miglietta, F., Zaldei, A., and Gioli, B. 2020. Individual tree crown segmentation in two-layered dense mixed forests from UAV LiDAR data. *Drones*, **4**(2): 10. doi:[10.3390/drones4020010](https://doi.org/10.3390/drones4020010).
- Torresan, C., Benito Garzón, M., O'Grady, M., Robson, T.M., Picchi, G., Panzacchi, P., et al. 2021. A new generation of sensors and monitoring tools to support climate-smart forestry practices. *Can. J. For. Res.* [This issue]. doi:[10.1139/cjfr-2020-0295](https://doi.org/10.1139/cjfr-2020-0295).
- Wang, Y., Lehtomäki, M., Liang, X., Pyörälä, J., Kukko, A., Jaakkola, A., et al. 2019a. Is field-measured tree height as reliable as believed – A comparison study of tree height estimates from field measurement, airborne laser scanning and terrestrial laser scanning in a boreal forest. *ISPRS J. Photogram. Remote Sens.* **147**: 132–145. doi:[10.1016/j.isprsjprs.2018.11.008](https://doi.org/10.1016/j.isprsjprs.2018.11.008).
- Wang, Y., Pyörälä, J., Liang, X., Lehtomäki, M., Kukko, A., Yu, X., et al. 2019b. In situ biomass estimation at tree and plot levels: What did data record and what did algorithms derive from terrestrial and aerial point clouds in boreal forest. *Remote Sens. Environ.* **232**: 111309–111315. doi:[10.1016/j.rse.2019.111309](https://doi.org/10.1016/j.rse.2019.111309).
- White, J.C., Wulder, M.A., and Buckmaster, G. 2014. Validating estimates of merchantable volume from airborne laser scanning (ALS) data using weight scale data. *For. Chron.* **90**(3): 378–385. doi:[10.5558/fc2014072](https://doi.org/10.5558/fc2014072).
- Williams, J., Schönlieb, C.B., Swinfield, T., Lee, J., Cai, X., Qie, L., and Coomes, D.A. 2020. Three-dimensional segmentation of trees through a flexible Multi-Class Graph Cut algorithm (MCGC). *IEEE Trans. Geosci. Remote Sens.* **58**: 754–776. doi:[10.1109/TGRS.2019.2940146](https://doi.org/10.1109/TGRS.2019.2940146).
- Yu, X., Hyppä, J., Holopainen, M., and Vastaranta, M. 2010. Comparison of area-based and individual tree-based methods for predicting plot-level forest attributes. *Remote Sens.* **2**(6): 1481–1495. doi:[10.3390/rs2061481](https://doi.org/10.3390/rs2061481).