A Benchmark of Lidar-Based Single Tree Detection Methods Using Heterogeneous Forest Data from the Alpine Space

Lothar Eysn, Markus Hollaus, Eva Lindberg, Frédéric Berger, Jean-Matthieu Monnet, Michele Dalponte, Milan Kobal, Marco Pellegrini, Emanuele Lingua, Domen Mongus And Norbert Pfeifer

Article

A Benchmark of Lidar-Based Single Tree Detection Methods Using Heterogeneous Forest Data from the Alpine Space

Lothar Eysn [1,*], Markus Hollaus [1], Eva Lindberg [1,2], Frédéric Berger [3], Jean-Matthieu Monnet [3], Michele Dalponte [4], Milan Kobal [5], Marco Pellegrini [6], Emanuele Lingua [6], Domen Mongus [7] and Norbert Pfeifer [1]

[1] TU Wien: Department of Geodesy and Geoinformation, Vienna University of Technology, Gußhausstraße 27–29, 1040 Vienna, Austria; E-Mails: Markus.Hollaus@geo.tuwien.ac.at (M.H.); Eva.Lindberg@slu.se (E.L.); Norbert.Pfeifer@geo.tuwien.ac.at (N.P.)

[2] SLU: Department of Forest Resource Management, Swedish University of Agricultural Sciences, 90183 Umeå, Sweden

[3] Irstea, UR EMGR Écosystèmes Montagnards, centre de Grenoble, F-38402 Saint-Martin-d'Hères, France; E-Mails: Frederic.Berger@irstea.fr (F.B.); Jean-Matthieu.Monnet@irstea.fr (J.-M.M.)

[4] FEM: Department of Sustainable Agro-Ecosystems and Bioresources, Research and Innovation Centre, Fondazione Edmund Mach, Via E. Mach 1, 38010 San Michele all'Adige (TN), Italy; E-Mail: Michele.Dalponte@fmach.it

[5] SFI: Department of Forestry and Renewable Forest Resources, Biotechnical Faculty, University of Ljubljana, Večna pot 83, 1000 Ljubljana, Slovenia; E-Mail: Milan.Kobal@bf.uni-lj.si

[6] TESAF: Department of Land, Environment, Agriculture and Forestry, University of Padova, Viale dell'Università 16, 35020 Legnaro (PD), Italy; E-Mails: Marco.Pellegrini@unipd.it (M.P.); Emanuele.Lingua@unipd.it (E.L.)

[7] UM-FERI: Faculty of Electrical Engineering and Computer Science, University of Maribor, Smetanova ulica 17, SI-2000 Maribor, Slovenia, E-Mail: Domen.Mongus@um.si

* Author to whom correspondence should be addressed; E-Mail: Lothar.Eysn@geo.tuwien.ac.at; Tel.: +43-158-801-122-49; Fax: +43-158-801-122-99.

Abstract: In this study, eight airborne laser scanning (ALS)-based single tree detection methods are benchmarked and investigated. The methods were applied to a unique dataset originating from different regions of the Alpine Space covering different study areas, forest

types, and structures. This is the first benchmark ever performed for different forests within the Alps. The evaluation of the detection results was carried out in a reproducible way by automatically matching them to precise *in situ* forest inventory data using a restricted nearest neighbor detection approach. Quantitative statistical parameters such as percentages of correctly matched trees and omission and commission errors are presented. The proposed automated matching procedure presented herein shows an overall accuracy of 97%. Method based analysis, investigations per forest type, and an overall benchmark performance are presented. The best matching rate was obtained for single-layered coniferous forests. Dominated trees were challenging for all methods. The overall performance shows a matching rate of 47%, which is comparable to results of other benchmarks performed in the past. The study provides new insight regarding the potential and limits of tree detection with ALS and underlines some key aspects regarding the choice of method when performing single tree detection for the various forest types encountered in alpine regions.

Keywords: single tree extraction; airborne laser scanning; forest inventory; comparative testing; co-registration; mountain forests; Alpine Space; matching

---

1. Introduction

The use of remote sensing data and related methods (*i.e.*, biomass estimation, delineation of forested areas) has become a standard in forest management [1–3]. Large area applications such as, for example, harvesting planning or forest stand mapping based on remote sensing products are now widely operational, especially in the northern European countries [4–6]. This development enables high-precision forest management, which is a prerequisite for the sustainable use of one of the key resources within the context of renewable energy. In mountainous regions, and particularly in the Alps, the use of forest resources in remote areas has decreased as valorization is hampered by accessibility constraints that prevent efficient mapping, management, and harvesting. To develop strategies to tackle this shortcoming, the research project NEWFOR (NEW technologies for a better mountain FORest timber mobilization) [7] was introduced to the Alpine Space program of the European Territorial Cooperation. The project aims at enhancing the wood supply chain within the Alpine Space (Alps core area and surrounding foothills/lowlands) to improve forest timber evaluation and mobilization using new remote sensing technologies such as airborne laser scanning (ALS), also referred to as Lidar.

When processing ALS remote sensing data for forest applications, area-based approaches [8–10] as well as single-tree based methods [11–13] can be found in the literature. Area-based methods provide statistically calibrated maps of forest stand parameters such as growing stock, stem density, and stand height, which are useful for large-area forest inventory and long term management planning. They can also be sufficient for harvesting planning in the case of simple forests such as plantations. Meanwhile, in complex alpine forests single tree information is highly valuable. Irregular stands are frequent, and silviculture is driven by the largest trees [14], which might also be used as intermediate support to optimize cable yarding. The spatial distribution of trees and their characteristics (height, species, crown size) are required inputs for growth simulation models [15], for the evaluation of the forest protection

effect against rockfalls [16], or even to identify trees with high biodiversity value [17]. Field inventories (FI) provide the required high level of local detailed information, but the high labor cost as well as accessibility constraint advocate for remote sensing solutions.

To gain such detailed information from ALS data, many studies on single tree detection were carried out by the research community [18–23]. Thus, many different methods are available for operational or scientific use. To gain deeper knowledge about the performance of different single tree detection methods, an international benchmark was carried out from 2005 to 2008 by the European Spatial Data Research Organization (EuroSDR), the International Society for Photogrammetry and Remote Sensing (ISPRS), and the Finnish Geodetic Institute (FGI). The study was published by Kaartinen *et al.* [24,25]. The benchmark was carried out using homogeneous ALS data and FI data from two study areas in southern Finland. A different benchmark based on ALS data from different types of forest was carried out in 2012 by Vauhkonen *et al.* [26] to test six different algorithms under different forest conditions. The investigated study areas are located in Norway, Sweden, Germany, and Brazil. Both studies had a great influence on understanding the performance of single tree detection based on ALS data. While the benchmark of Kaartinen *et al.* [24] focused on the performance of different methods using quite homogeneous dataset from one region, the benchmark of Vauhkonen *et al.* [26] focused on the effect of different study areas and forest types on the detection results of different algorithms. Vauhkonen *et al.* [26] used a very heterogeneous dataset which spans from a plantation of monospecies forest in Brazil to natural mixed forests in Europe.

For the Alps, the previously published benchmark results are only partly applicable as forests in Central Europe are different from forests in Northern Europe or Brazil. For this reason, the present study focuses on testing ALS-based single tree detection methods established in the Alpine Space. Based on a unique dataset covering different study areas, forest types, and structures from different regions in the Alpine Space, different methods are tested and analyzed in a clear and reproducible way. The focus is on the performance of the methods. Investigations on the effect of the heterogeneity of the ALS data (*i.e.*, different point density) on the detection results are not in the scope of this study. To the authors' best knowledge, this is the first benchmark ever performed for different forests within the Alps. This study is based on the single tree detection benchmark carried out within the project NEWFOR [7]. Detailed results are published in the project's final report [27]. The presented study summarizes the findings of the report and provides a discussion of the results.

The dataset used in this study is presented in Section 2. Section 3 provides a brief description of the tested methods as well as a detailed description of the matching procedure and evaluation of the results. In Section 4 results are presented in different levels of information while Section 5 contains the discussion. Finally, conclusions are drawn in Section 6.

2. Data and Materials

In total eight study areas in five alpine countries were investigated, representing different types of forest (Figure 1). For each study area, ALS data as well as detailed FI data were available. The ALS data and FI data used in the presented study were made public. The data are published on the NEWFOR website [28]. The dataset was somehow representative of the currently available operational data within the Alpine Space. This means that the data are heterogeneous as they originate from different sources

who acquired the data for different purposes. For example, the ALS data in Austria were collected with a focus on a nationwide terrain modeling campaign while the data in Slovenia were acquired with a focus on forestry applications. Detailed information about forest parameters is given in Table 1.
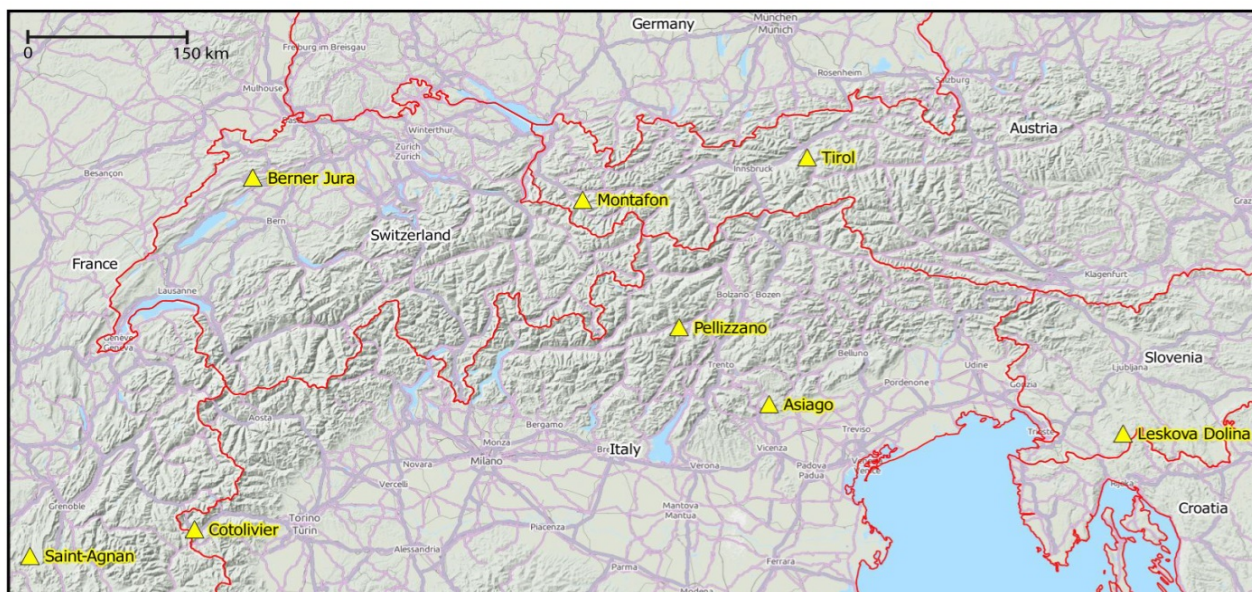
Figure 1. Study areas located within the Alpine Space.

Table 1. Acquisition parameters for the airborne laser scanning data and field inventory data.

Study Area	Country	Localization	Field Inventory			Airborne Laser Scanning		
			Nr. Plots	Total Size (ha)	Date	Date	Density (pts/m^2)	Sensor
Saint-Agnan	France	44°52' N 5°25' E	1	1.0	2010/7	2010/9	13	Riegl LMS-Q560
Cotolivier	Italy	45°2' N 6°46' E	3	0.4	2012/9	2012/7	11	Optech ALTM 3100
Berner Jura	Switzerland	47°9' N 7°4' E	1	0.1	2005	2011/4	5	Leica ALS 70
Montafon	Austria	47°4' N 9°58' E	1	0.3	2009/6	2011/9	22	Riegl LMS-Q560
Pellizzano	Italy	46°18 N 10°46' E	2	0.3	2013/8	2012/9	95–121	Riegl LMS-Q680i
Asiago	Italy	45°49' N 11°30' E	3	0.4	2012/10	2012/7	11	Optech ALTM 3100
Tyrol	Austria	47°23' N 11°44' E	3	1.2	2010/11	2008/7	4–10	Optech ALTM 3100
Leskova	Slovenia	45°39' N 14°28' E	4	0.8	2008/11	2009/10	30	Riegl LMS-Q560

2.1. ALS Data

The ALS data acquisition shows a mixture of different sensors and settings (Table 1). The acquired point densities vary from 5 points/m^2 in Switzerland up to 121 points/m^2 in Pellizzano, Italy. The mean flying heights above ground vary from 420 m to 800 m for the different flights. All flights were performed under snow-less conditions.

2.2. FI Data

For all study areas, fully calipered plots are available. Parameters such as, for example, stem locations, diameter at breast height (DBH), and tree heights as well as information about species composition and number of layers are provided for all plots (Table 2). Since the dataset originates from different institutions and inventory layouts, the size and shape of the inventory plots as well as the acquisition dates are different. Statistical measures are presented in Table 2, while the inventory dates of the FI data are presented in Table 1.

2.2.1. Positioning

The *in situ* absolute georeferencing of all plot locations was performed by GNSS measurements. The plots in Switzerland and Montafon were georeferenced with a total station or photogrammetry. In a post-processing step, each plot location was manually checked and co-registered to remote sensing data. To obtain an interpretable best fit of the tree pattern with the ALS Canopy Height Model (CHM), the tree pattern was visualized and manually moved in Quantum GIS 2.8.1 [29]. The CHM shows the local object heights and can be derived by subtracting the DTM from a surface model. A plot was only manually corrected if the initial position from the field survey showed gross errors. After manual co-registration, the estimated absolute planimetric accuracy of the plot location is ±2.0 m.

The stem positions were positioned relative to the given plot location with compass bearing and tape or laser range finding. The relative planimetric accuracy of positions varies from ±0.3 m to ±1.0 m, depending on the tools used. Vertex systems were used for measuring the tree heights. The vertical accuracy is expected to be ±1.0 m.

2.2.2. Classification

Four forest types were manually classified by interpreting the height distribution of trees in the FI data and considering the given meta-information. The classes are: (1) Single-Layered Mixed forest (SL/M); (2) Single-Layered Coniferous forest (SL/C); (3) Multi-Layered Mixed forest (ML/M); and (4) Multi-Layered Coniferous forest (ML/C).

Table 2. Statistical description of the forest plots. Only the three main species and species representing more than 5% of the stems are indicated. Corresponding Latin names: spruce (*Picea abies*), fir (*Abies alba*), beech (*Fagus sylvatica*), Scots pine (*Pinus sylvestris*), larch (*Larix decidua*), sycamore (*Acer pseudoplatanus*), elm (*Ulmus glabra*), and poplar (*Populus nigra*). The forest class correspond to single or multi-layered (SL or ML)/mixed or coniferous (M or C).

Plot #	Study Area	Plot Size (ha)	Caliper Threshold (cm)	Stem Density (/ha)	Mean Height (m)	Basal Area (m²/ha)	Mean Diameter (cm)	Stand Density Index	Coniferous Proportion (%)	Main species	Forest Class
1	Saint-Agnan	1.00	7.5	359	17.1	32.6	30.1	485	56	Fir, beech	ML/M
2	Cotolivier	0.13	4.0	843	18.1	50.5	25.8	889	97	Scots pine, larch, and spruce	ML/C
3	Cotolivier	0.13	4.0	390	16.5	34.3	29.7	514	92	Scots pine and larch	SL/C
4	Cotolivier	0.13	4.0	175	12.9	15.5	24.2	166	59	Larch and sycamore	ML/M
5	Berner Jura	0.10	12.0	340	29.8	67.6	47.7	959	100	Spruce and fir	SL/C
6	Montafon	0.30	10.0	400	13.9	35.7	25.0	401	100	Spruce	ML/C
7	Pellizzano	0.13	5.0	374	25.6	60.1	40.9	823	100	Spruce, larch, and fir	SL/C
8	Pellizzano	0.13	5.0	1870	13.7	68.1	16.7	974	80	Larch, spruce, fir, sycamore, and poplar	ML/M
9	Asiago	0.13	5.0	708	23.6	48.9	29.5	921	100	Spruce and fir	SL/C
10	Asiago	0.13	5.0	851	16.9	56.2	23.7	779	80	Spruce, fir, and beech	ML/M
11	Asiago	0.13	5.0	1344	13.9	37.9	16.0	660	28	Spruce, fir, and beech	ML/M
12	Tyrol	0.40	10.0	317	36.7	59.8	46.6	864	100	Spruce	SL/C
13	Tyrol	0.40	10.0	260	22.0	35.3	39.0	530	29	Sycamore, beech, spruce, and fir	SL/M
14	Tyrol	0.40	10.0	390	23.6	50.5	37.0	733	23	Sycamore, beech, spruce, and pine	SL/M
15	Leskova	0.20	10.0	265	22.9	29.1	34.2	439	76	Fir, spruce, and beech	SL/M
16	Leskova	0.20	10.0	185	24.6	27.6	22.0	359	78	Fir, spruce, and beech	SL/M
17	Leskova	0.20	10.0	585	20.6	38.2	25.5	603	47	Fir, spruce, beech, sycamore, and elm	ML/M
18	Leskova	0.20	10.0	460	24.6	54.0	32.7	708	53	Fir, beech and sycamore	ML/M

3. Methods

The global workflow for the benchmark consisted of the following steps. For each plot the ALS data and rasterized DTMs at 0.5 m and 1 m resolution were provided to benchmark participants. Participants applied their fully-automated tree detection algorithms (Section 3.1) in order to output a list with tree coordinates and heights for each plot. For each participant, this tree list is compared to the FI data with an automated matching procedure (Section 3.2).

3.1. Methods of Participants

In total eight methods were applied to the benchmark dataset (Table 3). The tested methods were chosen as they are common in the Alpine Space and originate from different countries. Most methods rely on local maxima (LM) detection in a rasterized CHM, but also one point cloud-driven method was applied.

Table 3. Overview of the applied methods.

ID	Participant Name	Method	Raster/Point Cloud	Resolution of Raster (m)	Kernel Size (pixel)
1	Irstea	LM + Filtering	R	0.20	11 × 11
2	FEM	LM + Region Growing	R	0.50	5 × 5
3	SFI	LM + Multi CHM	R	NA	3 × 3
4	TESAF	LM + Watershed	R	0.50	3 × 3
5	SLU	Segmentation + Clustering	R + P	0.25	-
6	TU Wien	LM 3 × 3	R	1.00	3 × 3
7	TU Wien	LM 5 × 5	R	1.00	5 × 5
8	UM-FERI	Polyn. Fitting + Watershed	R	1.00	7 × 7

LM: Local Maxima with moving window. The full affiliations of the participants are given on the title page.

3.1.1. Method #1 (LM + Filtering)

The method [30,31] is based on LM filtering within a rasterized CHM. The algorithm consists of five sequential steps:

1. Calculation of rasterized products (0.2 × 0.2 m resolution) based on the ALS data. The DSM is computed by retaining the highest altitude value of the points located inside each pixel. A DTM is computed by resampling the provided DTM at 0.5 m resolution.
2. Non-linear filtering. Void pixels and artefacts in the DSM are removed with a closing filter. A disk of radius 4 pixels is used as structuring element.
3. Lowpass filtering. A smoothing filter, discrete approximation of a Gaussian kernel with sigma = 0.3 m, is applied to the DSM.
4. Maxima extraction. A LM filtering with sliding window of size 11 × 11 pixels is applied to extract the LM.
5. Maxima selection. Pixels that are a LM are retained if the value of the corresponding pixel in the CHM is superior to 7.5 m. The CHM is computed as the difference between the non-linear filtered DSM and the DTM.

The remaining maxima are the final tree top candidates. Corresponding coordinates are the pixel centers and heights, which are extracted from the CHM. Algorithm parameters (raster resolution, Gaussian kernel, and LM filtering size) were chosen in the framework of a previous study [31], by using an automatic training process designed to minimize the trade-off between omission and commission errors.

3.1.2. Method #2 (LM + Region Growing)

The method [32] exploits both a rasterized CHM and the ALS point cloud with normalized height. The CHM is computed by assigning each pixel the value of the 95th percentile of the elevations of the first return ALS points. A nearest neighbor interpolation is used for pixels with no corresponding ALS data. The following detailed steps are applied:

1. A low-pass (LP) filter is applied to the rasterized CHM. For the CHM, a spatial resolution of 0.5×0.5 m^2 is used. For the LP filter, a window of 3×3 pixels is used.
2. Seed points $S = \{s_1, \ldots, s_N\}$ are defined using a moving window approach. The central pixel of a 5×5 pixel moving window is a seed point if it is (a) the highest point inside the window and (b) higher than 2.5 m.
3. Initial regions are defined starting from the seed points, and a label map L is defined: $L_{i,j} = k$ if (i,j) is a seed point with index k, otherwise $L_{i,j} = 0$.
4. Region growing according to the following procedure:

a. consider a label map point $L_{i,j} \neq 0$ and take its neighbor pixels:
$$\{(i, j-1); (i-1, j); (i, j+1); (i+1, j)\};$$
b. a neighbor pixel (i', j') is added to the region n if:

$$\left(\text{dist}\left((i', j'), s_n\right) < \text{Dist}_{\text{Max}}\right) \& \left(\text{CHM}(i', j') > \left(\text{CHM}(s_n) * \text{Perc}_{\text{Tresh}}\right)\right) \& \left(L_{i', j'} \neq 0\right)$$
with $\text{Perc}_{\text{Tresh}} \in [0, 1]$.

c. iterate over all the pixels $L_{i,j} \neq 0$, and repeat until no pixels are added to any region.

5. From each region, extract the first return ALS points, and apply Otsu thresholding [33] to the normalized heights of the extracted points.
6. Take only the first return ALS points higher than the Otsu threshold and apply a 2D convex hull to these points;
7. The resulting polygons are the final tree crowns. The positions of the trees are defined as the position of the highest ALS point inside each crown. The height of the crown is defined as the 95th percentile of the first return ALS points inside the crown.

3.1.3. Method #3 (LM + Multi CHM)

The method is based on iterative CHM generation and LM detection within a moving window of 3×3 pixels for various CHMs. The method is fully automated and processes the data in two general steps, which are (A) sequential identification of potential trees and (B) filtering of the extracted potential trees. Step (A): The ALS point cloud is normalized to local heights by removing the terrain trend using a DTM. From the normalized point cloud, an initial CHM is created by assigning the 95th height

percentile within each raster cell. Based on this CHM, LM are detected and the found positions and heights are stored in a database. For the next iteration, points in the uppermost layer of the normalized ALS data are eliminated. The "eliminating" layer is defined as a band of 0.5 m below the current CHM. Based on the filtered data, a new CHM is created, LM are extracted, and the LM parameters are added to the database. This procedure is carried out sequentially until all points are removed from the normalized point cloud.

Step (B): All detected LM in the database are sorted by decreasing heights. The highest LM is considered a detected tree. For each following LM, the LM is considered a detected tree if there is no detected tree within a 2D distance of 2 m as well as a 3D distance of 5 m.

3.1.4. Method #4 (LM + Watershed)

The presented method [34] is based on the method published in Koch *et al.* [35]. In a first step, a rasterized CHM with a spatial resolution of 0.5×0.5 m^2 is generated from the ALS data. The next processing step consists of a CHM surface smoothing using a Gaussian kernel filter. Focal statistics with a LM detection algorithm are used to extract potential tree tops from the smoothed CHM. The identified trees are then analyzed through a conditional script that considers a minimum distance and height difference from the nearest trees in order to identify and delete potential false positives. The coordinates of the found tree tops are then used as seed points in a watershed algorithm run on the inverse CHM in order to delineate single tree crowns and to generate polygon features with associated information on the canopy area. The method is fully automated and has been implemented as a workflow of geoprocessing tools within the software ESRI ArcGIS 10.2 [36].

3.1.5. Method #5 (Segmentation + Clustering)

In the presented method [37], the delineation is done by segmentation of a correlation surface model followed by ellipsoidal tree model clustering of the ALS data in 3D. The aim of the segmentation is to establish one segment for each tree in the topmost canopy layer. The segmentation method is based on geometric tree crown models and raster maps with 0.25 m cells. For each raster cell, an ellipsoid surface is calculated from different generalized ellipsoids. Correlation coefficients are calculated between the height of the ellipsoid surfaces and the height of the ALS data within the horizontal model radius. The correlation surface (CS) is defined as the highest correlation coefficient for each raster cell. The CS is smoothed and delineated with watershed segmentation.

The aim of the clustering is to establish one cluster for each tree in the topmost canopy layer as well as one cluster for each tree below. The algorithm is based on k-means clustering using ellipsoidal tree crown models. The clustering is done in two steps. In the first step, the ALS data are assigned to different clusters based on the Euclidian distance between the ALS data and the cluster centers. In the second step, an ellipsoid surface is fitted to each cluster and the ALS data are re-assigned to the different clusters based on a distance derived from the ellipsoid surface. Two categories of clusters are defined: Fixed clusters corresponding to trees already identified by segmentation of the CS and additional flexible clusters corresponding to trees below the topmost canopy layer. The top of the clusters is defined as the maximum height above the ground of the ALS data assigned to each cluster. The horizontal position of

the fixed clusters is defined as the horizontal position of the top of the cluster, while the horizontal position of the flexible clusters is defined as the horizontal position of the cluster center.

3.1.6. Method #6 (LM 3 × 3)

The method is published in Eysn *et al.* [38]. First, a DSM is processed based on the ALS point cloud using a land-cover-dependent derivation approach [39]. This approach makes use of the strengths of different algorithms for generating the final DSM by using surface roughness information to combine two DSMs, which are calculated based on (i) the highest echo within a raster cell, and (ii) moving planes interpolation. Second, a CHM with a spatial resolution of 1×1 m^2 is derived by subtracting the DTM from the DSM. The base products are derived consistently for all study areas using the OPALS [40] software. Finally, the positions and heights of single trees are determined from the CHM using an LM filter based on a moving window (MW) approach. If the center pixel of the MW is an LM, a potential tree is detected. Only detected positions with a CHM height greater than 3 m are considered. A circular kernel with a diameter of three pixels is used (indicator in the assigned name is "3 × 3"). For all detected trees, the position and tree height are stored.

3.1.7. Method #7 (LM 5 × 5)

The workflow is exactly the same as described in Section 3.1.6. Instead of a circular kernel with a diameter of three pixels, a circular kernel with a diameter of five pixels (indicator in the assigned name is "5 × 5") is used for the moving window.

3.1.8. Method #8 (Polynomial Fitting + Watershed)

First, a rasterized DSM is derived from the ALS data. A resolution of 0.5×0.5 m^2 is used if the point density is greater than 10 pts/m^2; otherwise the spatial resolution is set to 1.0×1.0 m^2. The height of a grid cell is defined by the highest points within the cell, while inverse distance weighting interpolation is used for defining the heights of the cells with no contained points. The DSM is normalized to a CHM by subtracting a DTM. Morphological opening and closing are performed in order to remove possible outliers. Best fitting second-degree polynomials are estimated in the 7×7 pixel neighborhood of each grid-cell, using the least squares method. Its factors are used to detect concave neighborhoods (potential tree-tops). This approach is also known as local fitting surfaces (LoFS) [41]. Watershed regions are then estimated based on concave markers. A region-adjacency graph is constructed over the obtained regions. Finally, region merging is performed based on geometric attributes of the regions (height, area, and shape compactness) with priorities defined by the measured similarity of best fitting polynomials. In all the test-cases, the same attribute-thresholds are used for identifying regions (nodes of the graph) that were merged. They are defined as follows: Regions with areas smaller than 5 m^2 or heights smaller than 0.2 m are merged under the conditions that the resulting region after the merging does not exceed 150 m^2, and its compactness does not exceed the value of π. The positions of treetops are defined by the highest points of smoothed CHM (Gaussian filter with standard deviation = 1.0 m) within the corresponding regions, while the original height of the treetop is used for defining the height of the tree.

3.2. Tree Matching Process

A fully automated tree matching procedure for linking the different detection results to the reference FI data is established and applied. Compared to manual interpretation by experienced human interpreters, this methodology enables a clear, objective, and reproducible testing.

3.2.1. Input Data

The input data for the automated tree matching are:

1. Resulting single tree data from benchmark participants (hereinafter referred to as "Test");
2. Forest Inventory data of the study areas (hereinafter referred to as "Reference"); and
3. Area of Interest of the study areas (hereinafter referred to as "AoI").

3.2.2. Implementation of the Matching Algorithm

The matching procedure is performed in three general steps (Figure 2). The detailed workflow is presented in Figure 3. The matching between Test trees and corresponding Reference trees is implemented in Python 2.7.8 [42].
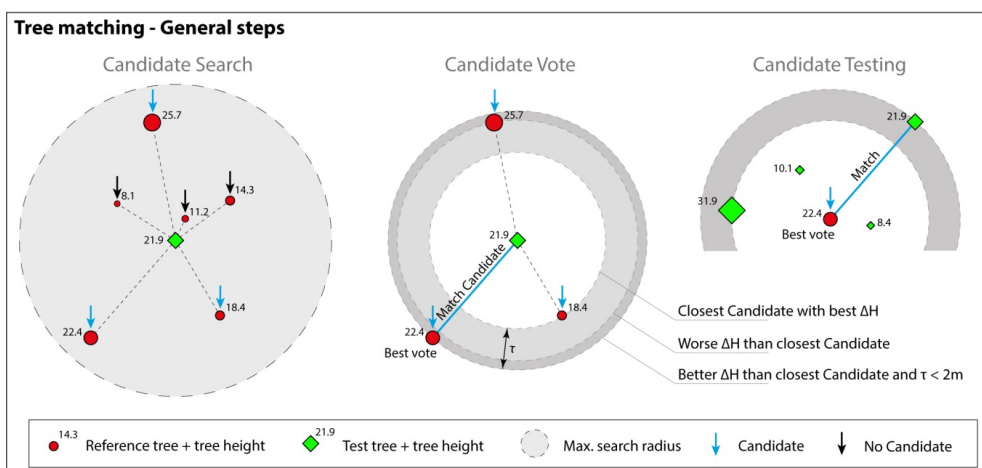

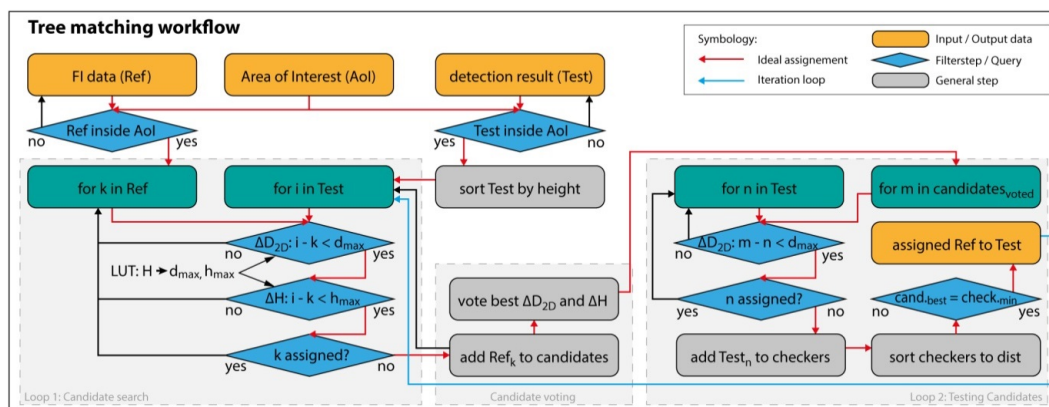**Figure 2.** Basic steps of the tree matching workflow.

Figure 3. Workflow of the matching procedure.

3.2.3. Candidate Search

Initially, all input data inside an AoI are selected for the matching procedure. The AoI is defined by the surveyed area within the FI. The Test trees are sorted by tree height and the matching procedure starts from the highest Test tree. For a Test tree the restricted nearest neighbor Reference trees within a defined neighborhood are determined. Ideally the neighborhood is defined by the crown area of the Test tree. As this information is not available, circular buffering is used as the 2D neighborhood. In contrast to other methods, which use the nearest neighbor with the smallest 2D or 3D Euclidean distance as a match [26,43], only restricted nearest neighbors are matched. This method introduces a neighborhood criterion ΔD_{2D} and a height criterion ΔH. Both need to be fulfilled to make a tree a candidate for being matched/assigned (Table 4). Additionally, already assigned neighboring Reference trees cannot become candidates. ΔD_{2D} checks the horizontal distance between Test and Reference tree, while ΔH checks the tree height difference. The thresholds of ΔD_{2D} and ΔH vary depending on the tree height of the Test tree. The motive is the following.

The locations of trees are, in case of a terrestrial survey, measured at the bottom of the stem while most detection methods detect the position of the stem at the tree top. If a tree is tilted or shows a curved stem, the locations of tree top and the bottom of the stem differ. This effect normally increases with increasing tree heights. Therefore, with increasing tree heights the ΔD_{2D} criterion is increased up to a value of 5 m. The value of 5 m was chosen to align with the thresholds applied in Kaartinen *et al.* [24]. Additionally, the ΔD_{2D} criterion considers positional errors from the FI survey as well as inaccuracies originating from the tree detection.

The height accuracy of trees measured in a terrestrial forest inventory is believed to be decreasing with increasing tree height. Therefore, when comparing a terrestrially measured tree height to an automatically detected one, this effect should be considered. This is accounted for by increasing the ΔH value with increasing tree heights (Table 4).

Table 4. Height and neighborhood criteria for the candidate search. H_{Test}: Height of Test tree; ΔH: Height difference between Test and Reference; ΔD_{2D}: 2D distance between Test and Reference.

Criterion	Height Test	Distance Test
1	$H_{Test} \leq 10$ m and $\Delta H < 3$ m	$\Delta D_{2D} < 3$ m
2	10 m $< H_{Test} \leq 15$ m and $\Delta H < 3$ m	$\Delta D_{2D} < 4$ m
3	15 m $< H_{Test} \leq 25$ m and $\Delta H < 4$ m	$\Delta D_{2D} < 5$ m
4	$H_{Test} > 25$ m and $\Delta H < 5$ m	$\Delta D_{2D} < 5$ m

The thresholds presented in Table 4 were empirically found by testing different settings on a subset of the dataset while visually interpreting the quality of the matching results. The resulting values are applied to all datasets within this benchmark.

3.2.4. Candidate Voting

Since multiple trees can become candidates in the candidate searching process, the selected candidates are ranked depending on their ΔH and ΔD_{2D} value. Starting from the nearest candidate, all other local

candidates are tested for a better ΔH. If a candidate shows a better ΔH and its ΔD_{2D} is at a maximum of 2.5 m greater than the initial candidate's ΔD_{2D}, the candidate becomes the new best voted candidate (Figure 2). The value of 2.5 m (half of the maximum possible ΔD_{2D}) is introduced to spatially limit possible candidate jumps. This feature is helpful if candidates are clustered and the best fitting tree inside this cluster should be found.

3.2.5. Candidate Testing

Since the tree matching process is more than an isolated problem of matching one Test tree against a group of Reference trees, all other Test trees in the surrounding need to be considered. This is performed by checking the best voted candidate against the surrounding Test trees. If the previously best voted Test tree is the closest tree with the best height difference, these two trees are finally matched.

3.2.6. Products of the Matching Process

The outputs of the matching process are qualitative and quantitative statistical parameters as well as vector layers which can be displayed in a Geographical Information System (Figure 4). The following statistical parameters and vector layers are provided.

Detailed parameters: For each method and plot

- Number of extracted trees N_{Test} and number of Reference trees N_{Ref}
- Number of matched trees N_{Match} and commission errors N_{Com}. $N_{Com}+N_{Match}=N_{Test}$
- Extraction rate → Total number (N_{Test}) or rate (N_{Test}/N_{Ref}) of extracted Test trees by a method
- Matching (assignment) rate → Total number (N_{Match}) or rate (N_{Match}/N_{Ref}) of matched trees
- Commission rate → Total number (N_{Com}) or rate (N_{Com}/N_{Test}) of Test trees that could not be matched
- Omission rate → Total number ($N_{Om}=N_{Ref}-N_{Match}$) or rate (N_{Om}/N_{Ref}) of Reference trees that could not be matched
- H_{Mean} → Mean of horizontal modulus of matching vectors (2D vector between Test and Reference)
- V_{Mean} → Mean of tree height differences (ΔH between matched Test and Reference)

Global parameters: Using detailed parameters of multiple plots or methods:
- RMS_{extr} → Root Mean Square of extraction rates
- RMS_{ass} → Root Mean Square of matching rates
- RMS_H → Root Mean Square of H_{Mean} values
- RMS_V → Root Mean Square of V_{Mean} values
- RMS_{Com} → Root Mean Square of commission rates
- RMS_{Om} → Root Mean Square of omission rates

The results of the matching process are presented in different levels of information. Exploring the detection results at the method level is displayed in Section 4.2. Section 4.3 shows the results for different forest types. The overall performance of the benchmark is presented in Section 4.4.
In addition to the derived statistical values per plot, the matching rates in different height layers are derived. The tree matches are sorted to the height layers defined by the intervals (in meters) [2, 5],

[5, 10], [10, 15], [15, 20], and [20–∞], which were also used in the benchmark of Kaartinen *et al.* [24]. The matching results in different height layers are derived to get a better understanding of how the different methods perform in understory vegetation.

For all levels of information, the obtained qualitative and quantitative parameters are plotted in two different bar graphs. One bar graph focuses on the different rates found in the matching process, while the other focuses on the spatial accuracy. An example is presented in Figure 5.
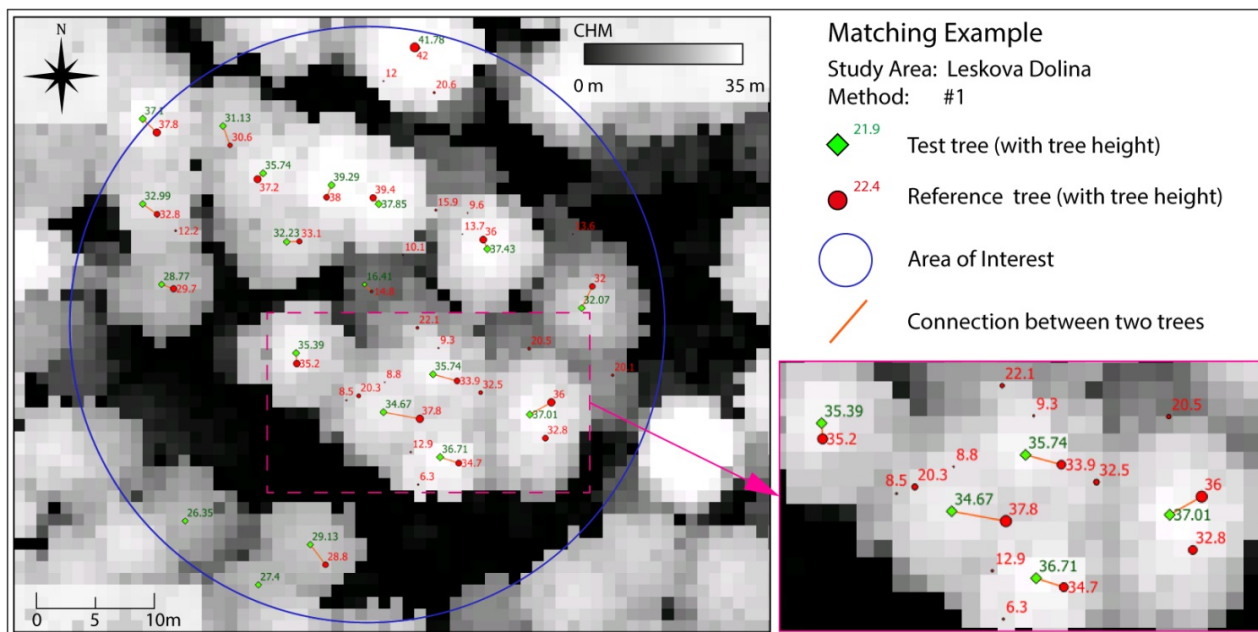
Figure 4. Matching example visualized in Quantum GIS. The detected Test trees (green diamonds), Reference trees (red disks), Area of Interest (blue circle), and the matched connections (orange lines) are displayed together with a height-coded CHM.
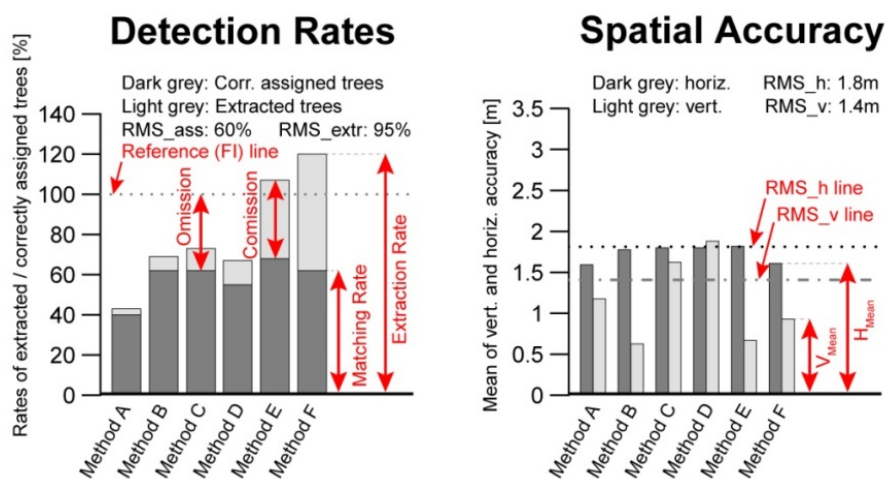
Figure 5. Bar graph examples for detection rates and spatial accuracy.

3.2.7. Validation of the Matching Procedure

The results of the automatic matching are validated by visually interpreting randomly selected matching results in Quantum GIS 2.8.1 [29]. The output vector layers of the matching process are

visualized as an overlay of the CHM. An experienced human interpreter classifies the matching results into four classes: Correctly Assigned (True Positive), Correctly not Assigned (True Negative), Wrongly Assigned (False Positive), and Wrongly not Assigned (False Negative). Descriptive measures are derived and presented.

4. Results

All participants were able to apply their method to the provided dataset. All submitted detection results were checked using the matching procedure described in Section 3.2. In total, 168 results consisting of 10987 detected potential tree positions were investigated.

4.1. Validation of the Matching Procedure

A subset of 699 Test trees, randomly selected from the submitted results, was manually interpreted and classified. The resulting error matrix and descriptive measures are presented in Table 5. An example of a visualized matching result displayed in Quantum GIS 2.8.1 [29] is presented in Figure 4. Nearby trees with matching tree heights get correctly connected in most cases. From the validated tree sample, only 3% were wrongly treated within the matching procedure. The obtained quality of the matching process shows an overall accuracy of 97% and a Kappa of 0.94.

Table 5. Error matrix and descriptive measures for the matching quality check.

| Matching Result | Reference—Manual Interpretation | | | |
	Match	No Match	*Totals*	User's Accuracy
Match	307	8	*315*	97%
No match	14	370	*384*	96%
Totals	*321*	*378*	*699*	
Producer's accuracy	96%	98%		
Overall accuracy: 97%		**Kappa: 0.94**		

4.2. Matching Results at Method Level

The matching results per method indicate how well a method performed for all study areas. In Table 6, the resulting statistical values are summarized.

The highest extraction rate (RMS_{extr}: 154%) was obtained by Method #6 (LM 3 × 3), while the lowest rate was found by Method #1 (LM + Filtering), showing a value of 51%. Regarding the matching rates, the highest rate (RMS_{ass}: 54%) was found by Method #6. In contrast, the lowest rate was obtained by Method #7 (LM 5 × 5). Speaking about incorrect detections, the highest commission rate (RMS_{Com}) with a value of 113% was produced by Method #6. The best RMS_{Com} rate with a value of 9% was found for Method #1. The highest omission rate (RMS_{Om}) was found for Method #7, which missed 63% of the given reference trees. The lowest and therefore best omission rates (RMS_{Om}: 51%) were found for Methods #5 and #6.

In the spatial accuracy section, the best positional accuracy with a RMS_H of 1.6 m was obtained by Methods #1, #4, and #6. The best height accuracy with a RMS_V value of 0.7 m was found for Method #3.

Table 6. Summarized detection results per method: RMS of selected indicators for all plots.

ID	Method	RMS$_{extr.}$ (%) Extraction Rate	RMS$_{ass.}$ (%) Matching Rate	RMS$_{Com}$ (%) Commission Rate	RMS$_{Om}$ (%) Omission Rate	RMS$_H$ (m) Height Accuracy	RMS$_V$ (m) Planar Accuracy
1	LM + Filtering	51	45	9	59	1.6	0.9
2	LM + Region Growing	57	43	20	61	1.8	1.2
3	LM + Multi CHM	101	46	61	57	1.7	0.7
4	LM + Watershed	86	49	49	55	1.6	1.1
5	Segment. + Clustering	139	53	95	51	1.7	1.0
6	LM 3 × 3	154	54	113	51	1.6	0.9
7	LM 5 × 5	52	41	16	63	1.8	1.1
8	Polyn. Fitting + Watersh.	54	44	13	59	1.8	1.1

Table 7 shows the matching results (*RMS$_{ass}$* values) in different height layers. For the layers 2–5 m and 5–10 m, Method #5 shows the best performance with values of 15% and 17%, respectively, while all other methods detected only half or even a quarter of the trees. For the layer from 10 to 15 m, the clear lead of Method #5 gets lost as Methods #3 and #6 show comparable values. For the uppermost two layers, the performance difference between the different methods is reduced. In the uppermost layer greater than 20 m, *RMS$_{ass}$* values from 66% to 82% are found.

Table 7. Root mean square of matching rate per method in different heights.

| ID	Method	RMS$_{ass.}$ 2–5 m	RMS$_{ass.}$ 5–10 m	RMS$_{ass.}$ 10–15 m	RMS$_{ass.}$ 15–20 m	RMS$_{ass.}$ > 20 m
1	LM + Filtering	0%	3%	16%	35%	72%
2	LM + Region Growing	0%	5%	15%	30%	72%
3	LM + Multi CHM	0%	3%	32%	46%	68%
4	LM + Watershed	4%	7%	20%	36%	76%
5	Segment + Clustering	15%	17%	32%	45%	76%
6	LM 3 × 3	4%	6%	28%	44%	82%
7	LM 5 × 5	2%	4%	14%	24%	66%
8	Polyn. Fitting + Watersh	2%	9%	16%	40%	73%

4.3. Matching Results by Forest Type

The detailed matching results by forest type indicate how well the different methods performed for different forest types. A graphical preparation of the matching results sorted to forest type is presented in Figure 6. In Table 8, the statistical parameters are summarized.

The highest extraction rate (*RMS$_{extr}$*) of 142% was found for single-layered mixed forests, while the lowest rate of 55% was found for multi-layered coniferous forests (ML/C).

For the matching rates, the highest *RMS$_{ass}$* rate of 86% was found for single-layered coniferous forests. The lowest matching rate (47%) was found for single-layered mixed forests.
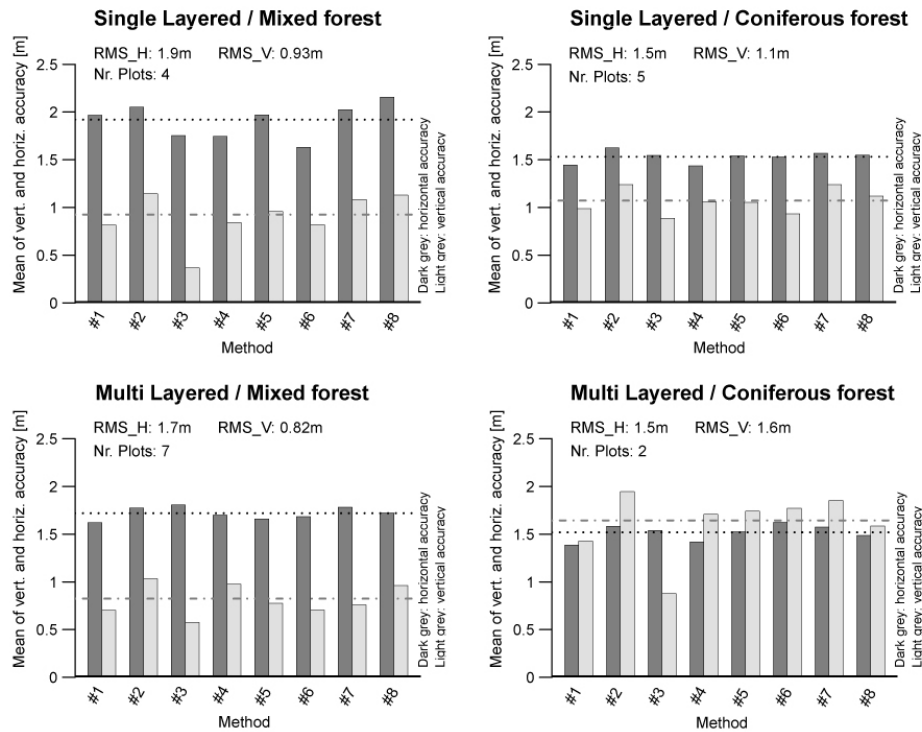
Detection Rates
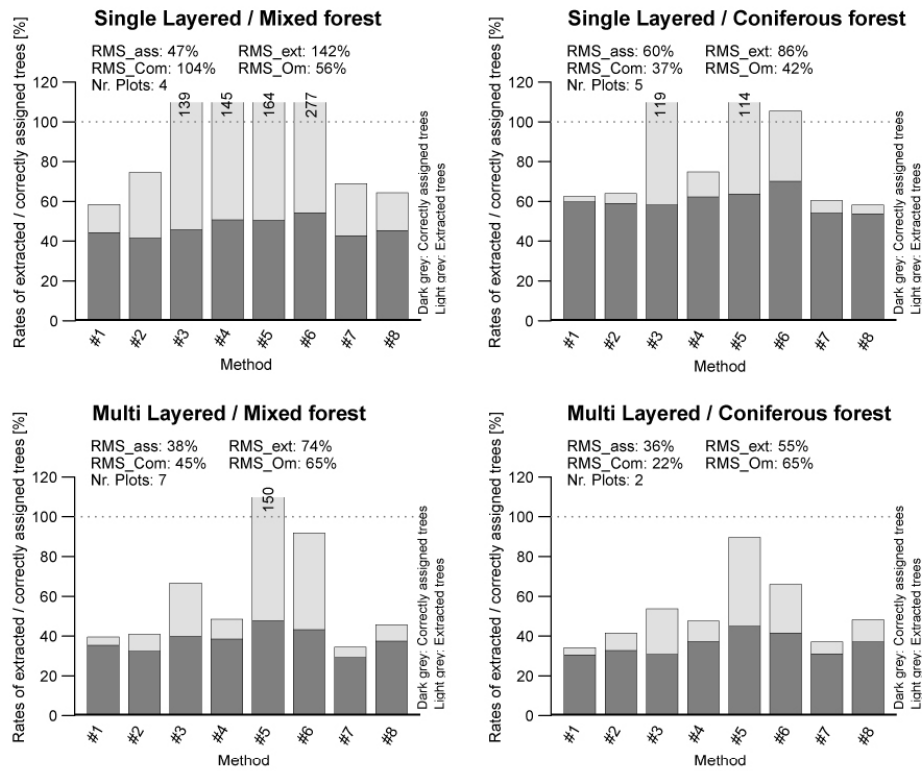
Spatial Accuracy

Figure 6. Bar graphs of detection rates and accuracies of the different forest types.

The highest commission rate (RMS_{Com}) of 104% was found for single-layered mixed forests. The lowest RMS_{Com} rate was found for multi-layered coniferous forests with 22%.

The highest omission rate (RMS_{Om}) was found for both types of multi-layered forests, with a value of 65%. The lowest rate was found for single-layered coniferous forests with a value of 37%.

In the spatial accuracy section, the best positional accuracy with a RMS_H of 1.5 m was obtained for coniferous forests. The best height accuracy with a RMS_V value of 0.8 m was found for multi-layered mixed forests.

Table 8. Summarized matching results by forest type—statistical parameters. Forest type: single or multi-layered (SL or ML)/mixed or coniferous (M/C).

Type	Nr. Plots	$RMS_{extr.}$	$RMS_{ass.}$	RMS_{Com}	RMS_{Om}	RMS_H	RMS_V
SL/M	4	142%	47%	104%	56%	1.9 m	0.9 m
SL/C	5	86%	60%	37%	42%	1.5 m	1.1 m
ML/M	7	74%	38%	45%	65%	1.7 m	0.8 m
ML/C	2	55%	35%	22%	65%	1.5 m	1.6 m

4.4. Overall Performance

In this section, the overall performance of the matching results of all eight methods is put together. A graphical preparation of the result is presented in Figure 7. The overall matching rate RMS_{ass} shows a value of 47%. This means that statistically 47% of all available Reference trees could be successfully matched. For the extraction rate RMS_{ext}, a value of 95% was found. The commission error and omission error show values of 60% and 57%, respectively.
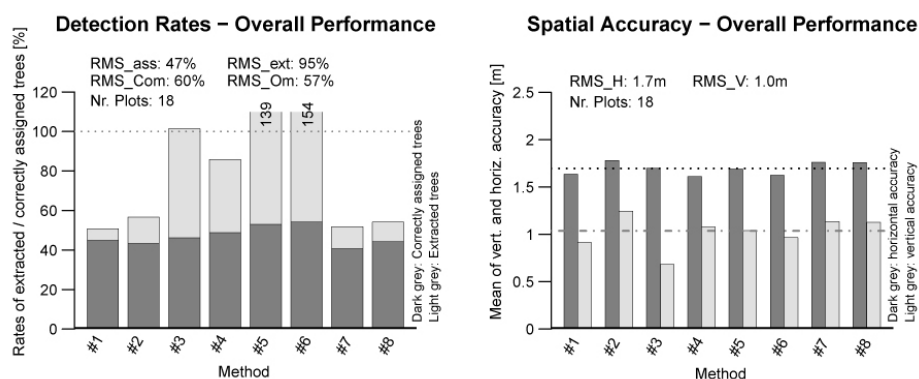
Figure 7. Overall performance of the tested methods.

5. Discussion

5.1. Input Data

For some study areas, the time gap between the ALS flight and the field survey is quite important. In the case of a field survey conducted after the ALS flight, the following errors can be expected:

- false omission errors due to tree growth in diameter (small trees reaching the caliper threshold between the ALS and the field surveys);

- false commission errors due to tree removal;
- false commission errors combined with false omission errors due to tree growth in height, which exceeds the matching threshold.

The opposite errors can be expected when the ALS flight is made afterwards. The time gap is two years in Montafon and Tyrol and six years in Berner Jura. It is one year or less in the other study areas. In mountain areas, tree growth is quite slow so that this factor has a limited impact on the matching process. Indeed the height growth of dominant trees during two years is likely to be lower than the vertical accuracy, which is already handled by the matching process. Besides, the proportion of trees reaching the caliper threshold in between the ALS and field survey is very low, except for young, dense stands. In all plots, the mean diameter exceeds the diameter threshold by more than 10 cm. Unless the diameter distribution is bimodal, the proportion of trees with a diameter close to the threshold is small. In Berner Jura, the forest is a dense, mature stand with no understory, so that the caliper threshold and height growth are not an issue. Regarding tree harvesting or mortality, no changes were reported by the partners who provided the data.

The difference in caliper threshold might also influence the matching results. The smallest trees represent a minor proportion of basal area but a major proportion of omission errors. Foresters usually set the caliper threshold by taking into consideration the inventory objective, the forest structure, and the required field effort. In this benchmark, the caliper thresholds were not *a posteriori* set to the highest value (12 cm in Berner Jura) because this would not be suitable to describe some young, dense plots. When comparing the detection results at the forest type level, it should nevertheless be taken into consideration.

Accurate georeferencing of the input data in a pre-processing step is a prerequisite for correctly comparing remotely sensed data with FI ground truth data. Three main error sources exist when comparing these data. These error sources are (*A*) positional errors due to inaccurate measurements (georeferencing); (*B*) irregularities of the local forest (*i.e.*, tilted trees, complex crowns); and (*C*) errors originating from the detection algorithm. To minimize error source *A*, forest inventory data is ideally acquired by using a survey grade GNSS system or a total station. A positional check by manually co-registering the inventory data to the remote sensing data in a post-processing step is necessary in most cases. Automated co-registration methods [44–46] can help to support the co-registration process. The reported performance of automated methods varies from 68% up to 92% and depends mainly on the input data and the variability of the forest stands [45]. In the presented study, error *A* was minimized by using manually co-registered datasets, with an estimated planar accuracy of ±2.0 m. and ±1.0 m for the vertical accuracy. The errors *B* and *C* are handled within the matching procedure by allowing flexible matching connections with a search radius depending on tree heights. The used FI data are heterogeneous due to different acquisition guidelines and methods. Therefore the data can be considered as imperfect compared to a local standardized FI. However, imperfect data should not have disadvantaged one of the detection methods as all participants faced the same conditions.

The presented matching procedure enables interpreter-independent and reproducible results in a short amount of time. The automatic matching took a few minutes while the manual interpretation within the validation process of the subset took several hours. The Overall Accuracy of 97% indicates that the matching procedure worked sufficiently for the presented dataset.

5.2. Matching Results

In general, the vertical structure of the forest (vertical distribution of tree heights) seems to have a major impact on the detection/matching results of the different methods. This finding is also reported by Vauhkonen *et al.* [26]. The more vertically distributed the trees are, the lower the matching rates are. The matching rates in different height layers indicate that especially in the lower height layers more advanced methods like point cloud-driven 3D clustering (Method #5) can detect more trees than methods that rely on local maximum detection based on a rasterized canopy height model. This finding was also reported in the benchmark study of Kaartinen *et al.* [24]. Method #5 achieved the highest number of small trees extracted.

Matching results combining a high matching rate with a low commission rate indicate a good matching result. The best detection result was obtained within the study area of Berner Jura, which consists of an old forest stand with high trees and no understory vegetation. The lowest detection result was obtained within plot #5 in Pellizzano, which consists of a multi-layered forest with a high amount of trees in different height layers. Only 15% of trees smaller than 10 m could be correctly extracted by the best performing method. In a summarized view, the results show that multi-layered forests are challenging for all tested methods. Maybe new methods as e.g., presented in Vega *et al.* [47] or Kandare *et al.* [48] will help to improve this issue in the future. Vega *et al.* [47] reported an overall performance of 75% for mixed multi-layered mountainous forest in the French Alps, with a 58% detection rate in the dominated tree layer.

Regarding the detection of small trees, it can be assumed that small trees in subdominant layers theoretically get mapped more efficiently at a higher ALS point density. This can be linked to a potentially higher canopy penetration rate. Kaartinen and Hyyppä [25] and Reitberger *et al.* [49] conclude that the laser point density has less impact on the individual tree detection. In contrast Wallace, Lucieer, and Watson [43] conclude that high point densities are more significant for single tree detection than the method used. Wallace *et al.* studied a very young, planted stand of Eucalyptus trees using high-resolution UAV Lidar data, which is not comparable to the dataset of the Alpine Space. Therefore, it is assumed that the different point densities given in the presented study should not significantly influence the detection results. Within the pilot area Pellizzano, a maximum ALS point density of 121 pts/m^2 is given and the inventory data show a high vertical distribution of trees. Even at this high point density, only the worst detection result of all tested areas could be obtained. All study areas have point densities higher than 10/m^2, except Berner Jura and Tyrol. Both of these have plots of single-layered, mature stands so that the smaller point density is probably not a limiting factor for tree detection. However, investigating the effect of different point density on the detection results was not in the scope of this study.

5.3. Matching Results per Method

The best ratio between a high matching rate and a low commission rate was found for Method #1, which consists of a local maximum search in a canopy height model using a moving window approach. The fact that the algorithm parameters were automatically optimized for this purpose on an independent dataset seems to be an advantage compared to methods where parameters are set according to the user

appreciation. In the lower height layers up to 10 m tree height, only up to 3% of the extracted trees could be correctly matched. Since the method relies on a rasterized canopy height model and filtering of trees smaller than 7.5 m, this rather low value was expected. For the spatial accuracy, the obtained values for the location are comparable to the results of the other tested methods. For the height component, the second best value was achieved with a RMS value of 0.9 m, which is comparable to the values obtained for the best models in the benchmark of Kaartinen *et al.* [24].

Method #2 shows comparable matching rates to Method #1, but with a commission rate twice as high. In the lower height layers up to 10 m tree height, only up to 5% of the extracted trees could be correctly matched. In contrast to Method #1, trees down to a height of 2.5 m could be detected, which might lead to the slightly higher percentage value. However, the method is based on rasterized ALS data and therefore the rather low matching rate in the lower height layers was expected. The spatial accuracy of the method is comparable to the results of Method #1.

High commission rates in the results of Methods #3, #5, and #6 indicate that these methods tend to over-perform, which means they show high commission rates. Methods #3 and #5 are based on 3D operations in multiple canopy height models or directly in the 3D point cloud, while Method #6 is based on local maximum detection in a canopy height model, which uses a small local maxima kernel (3 × 3 pixels) and no preliminary smoothing. The small kernel tends to find local irregularities in the canopy height model and since these irregularities can be located even inside a single tree crown, the small kernel tends to detect too many potential trees. The result is the highest commission rate within this benchmark. The alternative Method #7 shows better results in terms of commission rate as the rate of the 5 × 5 kernel is seven times lower than the one from the 3 × 3 kernel. Methods #3 and #5 seem to be too sensitive in the detection process and the 3D clustering especially tends to detect multiple trees within a given single tree crown. Beside the fact of high commission rates for these methods, Method #5 shows up to 17% of correctly matched trees in the lower layers up to 10 m tree height. Compared to other methods, this is clearly the best result. Method #3 shows the best height accuracy with a RMS value of 0.7 m. Both Method #7 and #3 show the lowest matching rates in the uppermost height layer with trees taller than 20 m. In total, Method #7 shows comparable results to results of Method #1 and is counted as one of the best results within this benchmark.

Method #4 shows a relatively high matching rate of 49% (RMS) but in contrast the commission rate is high. This indicates that the method found many trees that could not be linked to the reference data. In the lower layers below 10 m tree height, up to 7% (RMS) of the available reference trees were correctly matched, while up to 40% (RMS) of the detected trees are sorted to commission errors. The rather low matching value can be explained by the methodology. It uses a smoothed rasterized canopy model, which follows the upper canopy and therefore the detection rate of smaller trees in subdominant layers is believed to be low. In the highest height layer with trees taller than 20 m, a matching rate of 76% (RMS) could be obtained, which is one of the highest values in this benchmark for this height class. The spatial accuracy of the method is comparable to that from Method #1. In general, the spatial accuracy of all methods does not differ very much.

Method #8 shows a high matching rate of 44% paired with a low commission rate of 13%. Based on these values, the results of Method #8 are close to the results of Method #1 and among the best within this benchmark. In the lower levels with tree heights up to 10 m, the method obtained a matching rate of up to 9%, which counts, together with Method #4 and #5, as the best obtained results. In general, a 9%

matching rate in lower height intervals is, compared to the other methods, a good result, but from an overall perspective such a low detection rate is unsatisfying. Like other methods that rely on maximum search in a rasterized canopy height model of the uppermost canopy, the low rate can be explained by the methodology.

5.4. Matching Results per Forest Type

The class of single-layered coniferous forests shows the best results of all tested classes as a high matching rate of 60% combined with a low commission rate of 37% is given. This result seems feasible as coniferous trees have, in most cases, a clearly defined tree crown shape. This means that the tree top appears as a clear peak in the canopy height model. Since most of the tested methods within this benchmark rely on local maximum detection on the canopy height model, the good result for single-layered coniferous forests was expected. The best performing methods for this forest type were Methods #1, #3, and #4.

The class of multi-layered coniferous forest as well as the class of multi-layered mixed forest show the lowest matching rates in this benchmark. Only a matching rate of up to 38% (RMS) could be obtained. The commission rate of the multi-layered mixed forest is twice as high as the rate found for the multi-layered coniferous forest, which shows a value of 22% RMS. The low matching rate can be explained by the methodology of the tested methods. Trees in lower layers are challenging for all tested methods. The higher commission rate for the multi-layered mixed forest can be linked to more complex crowns for deciduous trees, which results in over performing detection results. The best results for the multi-layered coniferous forest were obtained by Methods #2, #4, and #10. For the multi-layered deciduous forest, the best results were obtained by Methods #1, #4, and #8.

The single-layered mixed forest shows a matching rate of 47%, which is the second best matching rate for the classified results. In contrast, a very high commission rate of 104% is given. The high rate can be explained by the fact that deciduous tree crowns tend to be more complex than coniferous ones. Single tree crowns may consist of multiple local peaks in the canopy height model, which may be correctly detected as local maximum but do not represent the tree stem. The best performing methods for this forest type are Methods #1 and #8.

In general, it can be seen that the single-layered forest types show better results than the multi-layered ones. This was expected as forest structure has a significant influence on the results. Between the single-layered coniferous and mixed class, a considerable difference in the matching rates as well as commission rates is noticeable. This confirms the findings of Vauhkonen *et al.* [26], who tested the performance within coniferous and deciduous plots in Germany.

5.5. Overall Performance

The overall performance brings together all matching results from all tested methods. An overall matching rate of 47% (RMS) was found. This value aligns with the benchmark results presented in Kaartinen *et al.* [24], as well as with the results for the study areas in Germany and Norway in the benchmark published by Vauhkonen *et al.* [26]. In contrast, novel methods such as, for example, a purely point cloud-driven method presented in Vega *et al.* [47] enable a higher overall performance, *i.e.*, 75% for mixed multi-layered mountainous forest in the French Alps.

The overall best performing methods are #1, #2, #7, and #8. The other four tested methods show commission errors that are too high. For the spatial accuracy, a horizontal accuracy of 1.7 m (RMS) and a vertical accuracy of 1.0 m (RMS) could be obtained. These values are comparable to other previously carried out benchmarks. The performance of the different methods differs more for the tree detection than for the extracted tree heights. This was also reported by Vauhkonen *et al.* [26].

5.6. Perspectives

The trade-off between omission and commission errors turns out to be a critical point regarding tree detection. Some detection methods are probably intrinsically more efficient because they are able to extract more relevant information from the point cloud, as is expected from point cloud-based methods. However, as exemplified by Methods #6 and #7, which differ only by the LM kernel size, it turns out that the choice of algorithm parameters such as raster resolution, kernel size, and horizontal or vertical exclusion thresholds may have a major impact on detection results. From the image processing point of view, extracting trees is basically separating the signal from the noise. Depending on the forest structure (and on the caliper threshold, which defines the tree object) and on the acquisition parameters, the filters required for tree extraction have to be chosen or at least tuned specifically. In order to improve the detection algorithm available for forest practitioners, it seems important to A) have datasets that allow us to test the robustness of algorithms on a wide range of forest structures, and B) design algorithms able to optimize their setting, either based on internal (Lidar itself) or external (tree allometry) data. For the comparison of results, an automated matching procedure like that presented in this paper is of high relevance. Moreover, the choice of a trade-off criterion between the omission and commission errors would make comparisons easier, but it has to be application-oriented.

6. Conclusions

This study demonstrated that Forest Inventory data can be automatically matched to single tree detection results obtained from airborne laser scanning data. Furthermore, eight single tree detection methods were tested based on a unique dataset of different forest types originating from eight areas within the Alpine Space. The proposed method for automatically matching forest inventory data and remotely sensed data worked efficiently. In general, all tested methods achieve comparable results for the matching rates, but do differ for the extraction rates and omission/commission rates. The tree detection rates show a higher variation than the estimated tree heights. A method based on local maxima detection within a canopy height model using variable-sized moving windows is rated as the best performing algorithm. Complex multi-layered forests were challenging for all tested methods. A point cloud clustering-based method gained the best results for trees in subdominant layers, which is rated as an advantage over raster-based methods. The best detection results were obtained for single-layered coniferous forests.

Future studies should investigate the effect of different point densities on the detection results. Multiple datasets of the same area acquired with different flight parameters (*i.e.*, viewing angles, heights above ground, footprint size) would be necessary to perform this analysis robustly. Such datasets are rarely available. Automated absolute georeferencing between FI data and ALS data (co-registration) as well as an automated classification of FI plots in different forest types (*i.e.*, single-/multi-layered forests)

based on the ALS data would help us to overcome the manual steps performed in the presented study. Finally, the performance of novel, point cloud-driven single tree detection methods [47] should be tested on the unique dataset from the Alpine Space presented herein.

Author Contributions

Lothar Eysn: paper writing; data quality check; tree matching procedure, preparation of results; organization of the benchmark; paper revision. Jean-Matthieu Monnet: paper revision; application of detection method. Markus Hollaus, Norbert Pfeifer, and Frédéric Berger: supervision. Eva Lindberg, Michele Dalponte, Milan Kobal, Marco Pellegrini, Emanuele Lingua, and Domen Mongus: application of their detection method and writing the corresponding method description.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Hyyppä, J.; Holopainen, M.; Olsson, H. Laser Scanning in Forests. *Remote Sens.* **2012**, *4*, 2919–2922.
2. Holopainen, M.; Vastaranta, M.; Hyyppä, J. Outlook for the next generation's precision forestry in Finland. *Forests* **2014**, *5*, 1682–1694.
3. Franklin, S.E. *Remote Sensing for Sustainable Forest Management*; CRC Press; Taylor & Francis Group: Boca Raton, FL, USA, 2001.
4. Wulder, M.A.; Bater, C.W.; Coops, N.C.; Hilker, T.; White, J.C. The role of lidar in sustainable forest management. *For. Chron.* **2008**, *84*, 807–826.
5. Akay, A.E.; Oğuz, H.; Karas, I.R.; Aruga, K. Using lidar technology in forestry activities. *Environ. Monitor. Assess.* **2009**, *151*, 117–125.
6. LidarComm. Second Largest Forest Products Company Realizes Multiple Benefits from LIDAR and Ortho Imagery. Available online: http://www10.giscafe.com/nbc/articles/view_article.php?articleid=432054 (accessed on 9 April 2015).
7. NEWFOR Alpine Space Programme. European Territorial Cooperation 2007–2013. Available online: http://www.alpine-space.eu/projects/projects/detail/NEWFOR/show/ (accessed on 15 January 2015).
8. Hollaus, M.; Wagner, W.; Schadauer, K.; Maier, B.; Gabler, K. Growing stock estimation for alpine forests in Austria: A robust lidar-based approach. *Can. J. For. Res.* **2009**, *39*, 1387–1400.

9. Holmgren, J. Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. *Scand. J. For. Res.* **2004**, *19*, 543–553.

10. Hyyppa, J.; Kelle, O.; Lehikoinen, M.; Inkinen, M. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 969–975.

11. Yao, W.; Krzystek, P.; Heurich, M. Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform lidar data. *Remote Sens. Environ.* **2012**, *123*, 368–380.

12. Korpela, I.; Dahlin, B.; Schäfer, H.; Bruun, E.; Haapaniemi, F.; Honkasalo, J.; Ilvesniemi, S.; Kuutti, V.; Linkosalmi, M.; Mustonen, J. Single-Tree Forest Inventory Using Lidar and Aerial Images for 3D Treetop Positioning, Species Recognition, Height and Crown Width Estimation. In Proceedings of ISPRS workshop on Laser Scanning and SilviLaser 2007, Espoo, Finland, 12–14 September 2007; pp. 227–233.

13. Dalponte, M.; Bruzzone, L.; Gianelle, D. A system for the estimation of single-tree stem diameter and volume using multireturn lidar data. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 2479–2490.

14. Ancelin, P.; Barthelon, C.; Berger, F.; Cardew, M.; Chauvin, C.; Courbaud, B.; Descroix, L.; Dorren, L.; Fay, J.; Gaudry, P.; *et al. Guide des Sylvicultures de Montagne Alpes du Nord Françaises*; ONF/CEMAGREF: Grenoble, France, 2006; p. 154.

15. Lafond, V.; Lagarrigues, G.; Cordonnier, T.; Courbaud, B. Uneven-aged management options to promote forest resilience for climate change adaptation: Effects of group selection and harvesting intensity. *Ann. For. Sci.* **2014**, *71*, 173–186.

16. Dorren, L. *Rockyfor3D (v5.2) Revealed—Transparent Description of the Complete 3D Rockfall Model*; Association ecorisQ: Geneva, Switzerland, 2015; p. 32.

17. Kania, A.; Lindberg, E.; Schroiff, A.; Mücke, W.; Holmgren, J.; Pfeifer, N. Individual Tree Detection as Input Information for Natura 2000 Habitat Quality Mapping. In Proceedings of the Remote Sensing and GIS for Monitoring Habitat Quality—RSGIS4HQ, Vienna, Austria, 24–25 September 2014; p. 3.

18. Brosofske, K.D.; Froese, R.E.; Falkowski, M.J.; Banskota, A. A review of methods for mapping and prediction of inventory attributes for operational forest management. *For. Sci.* **2014**, *60*, 733–756.

19. Lindberg, E.; Hollaus, M. Comparison of methods for estimation of stem volume, stem number and basal area from airborne laser scanning data in a hemi-boreal forest. *Remote Sens.* **2012**, *4*, 1004–1023.

20. Maltamo, M.; Gobakken, T. Predicting Tree Diameter Distributions. In *Forestry Applications of Airborne Laser Scanning*; Maltamo, M., Næsset, E., Vauhkonen, J., Eds.; Springer: Dordrecht, The Netherlands, 2014; Volume 27, pp. 177–191.

21. Saad, R.; Wallerman, J.; Lämås, T. Estimating stem diameter distributions from airborne laser scanning data and their effects on long term forest management planning. *Scand. J. For. Res.* **2014**, *11*, doi:10.1080/02827581.2014.978888.

22. Xu, Q.; Hou, Z.; Maltamo, M.; Tokola, T. Calibration of area based diameter distribution with individual tree based diameter estimates using airborne laser scanning. *ISPRS J. Photogramm. Remote Sens.* **2014**, *93*, 65–75.

23. Næsset, E. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sen. Environ.* **2002**, *80*, 88–99.

24. Kaartinen, H.; Hyyppä, J.; Yu, X.; Vastaranta, M.; Hyyppä, H.; Kukko, A.; Holopainen, M.; Heipke, C.; Hirschmugl, M.; Morsdorf, F.; *et al.* An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sens.* **2012**, *4*, 950–974.

25. Kaartinen, H.; Hyyppä, J. *EuroSDR/ISPRS Project, Commission II "Tree Extraction"; Final Report*; EuroSDR (European Spatial Data Research): Dublin, Ireland, 2008.

26. Vauhkonen, J.; Ene, L.; Gupta, S.; Heinzel, J.; Holmgren, J.; Pitkänen, J.; Solberg, S.; Wang, Y.; Weinacker, H.; Hauglin, K.M.; *et al.* Comparative testing of single-tree detection algorithms under different types of forest. *Forestry* **2012**, *85*, 27–40.

27. Eysn, L.; Hollaus, M.; Monnet, J.-M.; Dalponte, M.; Kobal, M.; Pellegrini, M.; Lindberg, E.; Mongus, D.; Berger, F. *NEWFOR—Single Tree Detection Benchmark—Report*; NEWFOR: Wien, Austria, 2014; p. 87.

28. NEWFOR. The NEWFOR Single Tree Detection Benchmark Dataset. Available online: http://www.newfor.net/download-newfor-single-tree-detection-benchmark-dataset/ (accessed on 1 Feburary 2015).

29. QGIS Development Team QGIS Geographic Information System. Open Source Geospatial Foundation Project. Available online: http://www.qgis.org (accessed on 11 March 2015).

30. Monnet, J.-M.; Mermin, E.; Chanussot, J.; Berger, F. Tree Top Detection Using Local Maxima Filtering: A Parameter Sensitivity Analysis. In Proceedings of Silvilaser 2010, the 10th International Conference on LiDAR Applications for Assessing Forest Ecosystems, Freiburg, Germany, 14–17 September 2010; p. 9.

31. Monnet, J.-M. Using Airborne Laser Scanning for Mountain Forests Mapping: Support Vector Regression for Stand Parameters Estimation and Unsupervised Training for Treetop Detection. Ph.D. Thesis, University of Grenoble, Grenoble, France, 25 October 2011.

32. Dalponte, M.; Frizzera, L.; Gianelle, D. Estimation of Forest Attributes at Single Tree Level Using Hyperspectral and ALS data. In Proceedings of the ForestSAT 2014, Riva del Garda, Italy, 4–7 November 2014.

33. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cyb.* **1979**, *9*, 62–66.

34. Sambugaro, M.; Colpi, C.; Marzano, R.; Pellegrini, M.; Pirotti, F.; Lingua, E. Utilizzo Del Telerilevamento Per l'Analisi Della Biodiversità Strutturale: Il Caso Studio Della Riserva Forestale di Clöise (Asiago, VI). In Proceedings of the 17th Conferenza Nazionale ASITA, Riva del Garda, Italy, 5–7 November 2013; pp. 1171–1178.

35. Koch, B.; Heyder, U.; Weinacker, H. Detection of individual tree crowns in airborne lidar data. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 357–363.

36. ESRI ArcGIS Desktop: Release 10.2. Available online: http://www.esri.com/software/arcgis/arcgis-for-desktop (accessed on 30 November 2014).

37. Lindberg, E.; Eysn, L.; Hollaus, M.; Holmgren, J.; Pfeifer, N. Delineation of tree crowns and tree species classification from full-waveform airborne laser scanning data using 3-d ellipsoidal clustering. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 3174–3181.

38. Eysn, L.; Hollaus, M.; Schadauer, K.; Pfeifer, N. Forest delineation based on airborne lidar data. *Remote Sens.* **2012**, *4*, 762–783.

39. Hollaus, M.; Mandlburger, G.; Pfeifer, N.; Mücke, W. Land Cover Dependent Derivation of Digital Surface Models from Airborne Laser Scanning Data. In Proceedings of the ISPRS Commission III Symposium PCV2010, Saint-Mandré, France, 1–3 September 2010; p. 6.

40. OPALS Orientation and Processing of Airborne Laser Scanning Data. Available online: http://geo.tuwien.ac.at/opals (accessed on 1 Feburary 2015).

41. Mongus, D.; Lukač, N.; Žalik, B. Ground and building extraction from lidar data based on differential morphological profiles and locally fitted surfaces. *ISPRS J. Photogramm. Remote Sens.* **2014**, *93*, 145–156.

42. Foundation, P.S. Python Language Reference, Version 2.7. Available online: http://www.python.org (accessed on 15 July 2014).

43. Wallace, L.; Lucieer, A.; Watson, C.S. Evaluating tree detection and segmentation routines on very high resolution UAV lidar data. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 7619–7628.

44. Dorigo, W.; Hollaus, M.; Wagner, W.; Schadauer, K. An application-oriented automated approach for co-registration of forest inventory and airborne laser scanning data. *Int. J. Remote Sens.* **2010**, *31*, 1133–1153.

45. Monnet, J.-M.; Mermin, É. Cross-correlation of diameter measures for the co-registration of forest inventory plots with airborne laser scanning data. *Forests* **2014**, *5*, 2307–2326.

46. Hauglin, M.; Lien, V.; Næsset, E.; Gobakken, T. Geo-referencing forest field plots by co-registration of terrestrial and airborne laser scanning data. *Int. J. Remote Sens.* **2014**, *35*, 3135–3149.

47. Vega, C.; Hamrouni, A.; El Mokhtari, S.; Morel, J.; Bock, J.; Renaud, J.-P.; Bouvier, M.; Durrieu, S. PTrees: A point-based approach to forest tree extraction from lidar data. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *33*, 98–108.

48. Kandare, K.; Dalponte, M.; Gianelle, D.; Chan, J.C.W. A New Procedure For Identifying Single Trees in Understory Layer Using Discrete LiDAR Data. In Proceedings of 2014 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Quebec City, QC, Canada, 13–18 July 2014; pp. 1357–1360.

49. Reitberger, J.; Schnörr, C.; Krzystek, P.; Stilla, U. 3D segmentation of single trees exploiting full waveform lidar data. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 561–574.