

Article

An Open-Source Handheld Mobile LiDAR Pipeline for Tree Detection, Stem Reconstruction and Volume Estimation in *Pinus pinaster* Stands

Pedro Apóstolo ^{1,2,*}, Ivon Brandão ^{1,2}, Mateus Mendes ^{1,2,*}, Raúl Salas-González ^{1,2} and Beatriz Fidalgo ²

¹ RCM2+—Research Center in Asset Management and Systems Engineering, Coimbra Institute of Engineering, Polytechnic University of Coimbra, Rua Pedro Nunes, 3030-199 Coimbra, Portugal; ivon.brandao@esac.pt (I.B.); rsalas@esac.pt (R.S.-G.)

² Polytechnic University of Coimbra, Rua da Misericórdia, Lagar dos Cortiços, S. Martinho do Bispo, 3045-093 Coimbra, Portugal; bfidalgo@esac.pt

* Correspondence: a21280614@isec.pt (P.A.); mmendes@isec.pt (M.M.)

Abstract

Forests play a critical role in biodiversity and in the conservation of ecosystems. Reliable and efficient forest inventories are therefore essential for sustainable forest management. Handheld mobile laser scanning (HMLS) is operationally attractive for forest inventory but has rarely been validated against a destructive reference in Mediterranean *Pinus pinaster* stands. We propose an open-source five-phase HMLS pipeline comprising PDAL pre-processing with CSF ground classification, TLS2trees segmentation, CSF-anchored per-tree setup, slice-wise stem reconstruction with a sequential RANSAC prior, a one-parameter species-specific shape model, and analytical volume integration with a form-factor sanity clamp. The pipeline is calibrated against a destructive reference of 45 *Pinus pinaster* trees across four plots in central Portugal. Leave-one-plot-out cross-validation produces a volume MAPE of 9.9%, a height MAPE of 4.4% and a DBH MAPE of 3.9%. On the matched sub-sample for which the FJD Trion S1 manufacturer software and a manual CloudCompare workflow also have outputs, the proposed pipeline is competitive with the manual workflow on volume and delivers the lowest median DBH error of the three methods, while remaining fully open-source and reporting per-tree quality and uncertainty diagnostics alongside every estimate.

Keywords: handheld mobile laser scanning; HMLS; forest inventory; LiDAR; *Pinus pinaster*; stem volume; destructive validation



Academic Editor: Alicia Cordero

Received: 25 August 2026

Revised: 12 September 2026

Accepted: 15 September 2026

Published: 18 September 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

1.1. Forest Inventory and the Operational Gap of HMLS

Quantitative description of forest stands, tree positions, diameters at breast height (DBH), heights, and stem volumes is the foundation on which sustainable forest management decisions rest. Those measurements are the basis of forest inventory, which is in turn necessary for silvicultural prescription and wood-flow planning. Conventional inventories combine manual diameter tape or caliper and hypsometer measurements with allometric volume equations that propagate species- and region-dependent uncertainty into the per-tree volume estimate in European biomass models routinely reported at the tens-of-percent level at the individual-tree level [1]. Alternatives that have gained traction over the last decade include terrestrial LiDAR: static terrestrial laser scanning (TLS) and

handheld mobile laser scanning (HMLS) capture millimeter-to-centimeter-level stem geometry from the field, reducing the reliance on allometry and producing per-tree records that can be archived, audited, and compared over the years.

This paper focuses specifically on HMLS in naturally regenerated Mediterranean *Pinus pinaster* stands, a species and forest type that rank among the most important in the productive landscape of the Iberian Peninsula [2,3]. Despite its relevance, it has received limited validation using HMLS against a destructive reference.

1.2. Terrestrial LiDAR as an Alternative Measurement Platform

Terrestrial LiDAR has, over the last decade, moved from a research instrument to a credible operational alternative. Static TLS established the baseline: millimeter-level ranging, fixed-position tripod acquisitions, and multi-scan registration have been shown to extract DBH, tree height, and stem profile with accuracy that approaches destructive reference and systematically outperforms manual measurement [4–6]. HMLS, in which the sensor is carried along a walking trajectory and its pose is recovered by simultaneous localization and mapping (SLAM), trades the per-point accuracy of TLS for operational speed and full plot coverage at a single visit. Bauwens et al. [7] benchmarked a handheld SLAM scanner against TLS, reporting a DBH bias of 0.08 cm and an RMSE of 1.11 cm across 91% of the trees fully scanned at 1.3 m, while reducing plot acquisition time by roughly an order of magnitude. Earlier HMLS benchmarks reported DBH RMSE of approximately 2.9 cm on the full dataset, and approximately 1.5 cm once stems below 10 cm DBH were excluded, in managed conifer plots [8]. Subsequent work has extended HMLS to volume, height, and biomass in conifer plantations [9–11], in hardwood and mixed stands [12], in Mediterranean stands [13–15], and in multi-sensor benchmarks [16–19]. Today, the HMLS–TLS trade-off is well characterized: TLS remains the reference when millimeter precision is the objective, while HMLS is the operational tool when throughput and full-plot sampling are required.

1.3. From Point Cloud to Inventory: Pipeline Components

Converting an HMLS point cloud into an inventory table requires four tightly coupled processing steps, each of which has its own established literature: (i) denoising and ground segmentation, (ii) individual-tree segmentation, (iii) stem reconstruction and taper modeling, and (iv) volume integration and validation. The following paragraphs review the state of each step with emphasis on the implications for an end-to-end HMLS pipeline.

1.3.1. Ground Segmentation and Noise Removal in Forest Point Clouds

Ground segmentation is the first and often most decisive quality gate. The Cloth Simulation Filter (CSF) of Zhang et al. [20], reporting an average total error of 4.58% across the ISPRS filter-test benchmarks, has become the de facto standard for forested point clouds because it handles steep terrain, variable ground cover, and low-lying debris more robustly than TIN-based alternatives. Its extension to mobile-LiDAR forest clouds has reported total errors in the 0.44–1.20% range [21]. Residual noise in SLAM point clouds presents a related but distinct challenge: unlike TLS, which tends to produce uniformly distributed outliers that are rejected by standard statistical filters, HMLS output may contain spatially coherent range noise that survives statistical outlier removal and must be handled by geometry-aware post-filters, as proposed by Cabo et al. [22]. An HMLS pipeline that does not specifically guard against such coherent noise will leak residual debris into downstream segmentation.

1.3.2. Individual-Tree Segmentation

Tree instance segmentation from terrestrial point clouds has evolved from hand-crafted geometric methods toward two largely distinct families. Graph-based methods propagate tree instances from trunk seeds through the point cloud using edge-length constraints and cumulative-gap rules; TLS2trees [23] is the current open-source reference in this family and is broadly reusable across TLS and HMLS inputs, provided that semantic pre-labels (terrain/vegetation/stem) are supplied, with plot-wise segmentation RMSE reported around 60% for the challenging upper-canopy fraction in complex tropical plots. Deep learning approaches such as the Forest Structural Complexity Tool (FSCT; Krisanski et al. [24]) replace hand-crafted geometry with a PointNet++ segmentation trained on diverse forest scans, reporting a stem-class accuracy of 96.09% on the original benchmark. Both families have repeatedly been shown to perform competitively in managed stands, but neither is a complete solution: graph-based methods are sensitive to the quality of the input semantic labeling, and deep learning methods require large datasets for training.

1.3.3. Stem Reconstruction and Taper Models

Stem reconstruction by slice-wise circle fitting has a long history. The Random Sample Consensus algorithm of Fischler and Bolles [25] remains the standard geometric engine for robust circle fitting under contaminated data, and it has been combined with cylinder fitting [26] and elliptical fitting [27] to accommodate moderate eccentricity in tree cross-sections. At a coarser scale, taper models link radius to height over the whole stem [28–30].

Taper models have been incorporated into LiDAR pipelines in two distinct modes, each leaving a gap. Post hoc taper reporting fits a Kozak-type or spline form at the end of the reconstruction chain. Pitkänen et al. [31] fit a smoothing-spline taper to TLS-derived diameter sequences and reported a DBH RMSE of 4.4%. Saarinen et al. [32] applied a cubic spline as a post-fit smoother of a slice-wise diameter sequence. More recent hybrid approaches use taper equations as extrapolators above the highest reliable LiDAR slice. Lee et al. [33] reported that a hybrid LiDAR–Kozak formulation halved the stem-volume RMSE of a pure-Kozak approach. TreeQSM [6] takes a different route, fitting cylinders hierarchically without a biological taper prior.

None of the three modes uses a species-specific shape as the principal geometric parameterization of the slice-wise reconstruction itself. Boukhris et al. [34] compared parametric, semi-parametric and non-parametric taper models against TLS-derived stem profiles of 219 *Pinus nigra* trees and reported the segmented polynomial of Max and Burkhart [30], calibrated by mixed effects, as the best-performing parametric option.

1.3.4. Differences in Our Approach to the State of the Art

Existing HMLS approaches typically rely on per-tree taper models with multiple free parameters, such as the Kozak variable-exponent taper [28], whereas this work adopts a species-specific shape model with a single robustly estimated scale parameter, an approach that appears largely unexplored in the HMLS literature. Similarly, previous methods have applied post-fit smoothing of diameter profiles [32] or deep learning and geometric point-cloud completion [35], while the proposed pipeline introduces sequential RANSAC constraints and taper-driven analytical synthetic filling during stem reconstruction.

Tree volume in TLS and HMLS studies is commonly estimated using cross-section integration, mesh reconstruction, or quantitative structure models (QSMs), but validation against destructive measurements remains uncommon. Most HMLS studies use allometric or ALS-derived references [13–17], while destructive validation in *Pinus pinaster* has been limited to TLS or manual measurements [36,37]. The closest comparable HMLS study with destructive validation is that of Vandendaele et al. [38], which reported a 19% volume RMSE

on hardwoods. To the authors' knowledge, this study represents the first end-to-end HMLS pipeline for *P. pinaster* validated on a destructive reference dataset of this size (45 trees across four plots).

1.4. Research Gaps

Three specific gaps emerge from the preceding review and, together, they justify the present study.

Gap 1 is destructive validation of HMLS in maritime pine (*P. pinaster*) at a meaningful sample size. Existing HMLS work in Mediterranean species is benchmarked against allometric or ALS-derived reference, which carries an irreducible per-tree uncertainty of the order of tens of percent [1]. The few destructive studies in *P. pinaster* [36,37] use TLS or manual instruments rather than HMLS. The only published HMLS-destructive comparison (Vandendaele et al. [38], $n = 30$) is in hardwoods. A destructive validation of an end-to-end HMLS pipeline in *P. pinaster* with $n = 45$, multiple plots, and species-relevant size distribution is missing.

Gap 2 is a species-specific shape as the principal geometric parameterization of slice-wise stem reconstruction, complemented by a sequential RANSAC prior and a taper-driven synthetic-fill. Taper models have been used in LiDAR pipelines either as post hoc smoothers [31,32] or as above-the-last-slice extrapolators [33]. None constrains the slice-wise RANSAC fit within each slice with a stem-form prior, and none uses a species-specific fourth-degree polynomial as a one-parameter shape parameterization that collapses the seven free parameters of the Kozak family to a single scale parameter robustly estimated as the median of $r_i / P(x_i)$ across reliable slices.

Gap 3 is an HMLS pipeline that combines a robust DBH ensemble, a CSF-anchored ground reference, and a form-factor sanity clamp, with per-tree quality and uncertainty diagnostics surfaced alongside every reported value. The reported DBH of an HMLS pipeline is typically the model-evaluated value at $h = 1.3$ m, exposed to the specific failure mode of the chosen parameterization. A robust median across three independent estimators discards any estimator that diverges. A CSF-anchored base height eliminates a per-tree z_{base} inconsistency that propagates as a per-tree bias through every slice. A form-factor sanity clamp in the volume integration catches taper pathologies that occasionally survive the upstream quality gates.

1.5. Objectives and Contributions

The objective of this paper is to propose, calibrate against a destructive reference, and characterize an open-source HMLS processing pipeline for tree detection, stem reconstruction, and volume estimation in naturally regenerated *Pinus pinaster* stands, using a consumer-grade handheld SLAM scanner (FJD Trion S1) and a destructively measured reference of 59 *P. pinaster* trees across five plots in the Alge area of Lousã, central Portugal.

The principal contributions are five.

Destructive validation at a meaningful sample size. Per-tree comparison of pipeline volume against a destructive reference of 45 trees across four plots, with a documented exclusion criterion for the Alge2 plot ($n = 14$). To our knowledge, this is the largest destructively referenced HMLS evaluation in this species and forest type (Sections 3 and 4.1).

A species-specific one-parameter shape parameterization. A 4th-degree polynomial calibrated on 366 caliper observations across 59 destructively measured *P. pinaster* trees of the same stand ($R^2 = 0.97$) replaces the seven free parameters of the Kozak [28] variable-exponent taper by a single scale parameter, estimated robustly as a median across reliable slices, which eliminates regime-dependent model selection as a source of per-tree variance (Sections 2.4.4 and 4.2).

A CSF-anchored, plot-consistent base height. The Phase-1 CSF-classified ground points, sampled at the tree centroid by KD-tree median, replace the per-tree cloud minima as the reference for the base height, removing a per-tree offset that propagates through every downstream slice, including DBH (Sections 2.4.3 and 4.2).

A three-estimator median DBH ensemble decoupled from the volume model. The reported DBH is the median of three independent estimators. Its median absolute percentage error is 2.46% on the 45-tree sample and 2.74% on the 40-tree sub-sample matched to both baselines, below the FJD Trion Model (3.59%) and the manual CloudCompare workflow (4.92%) on the same trees (Sections 2.4.4, 3.4 and 4.2).

A sequential RANSAC prior, a taper-driven analytical synthetic-fill and a form-factor sanity clamp are integrated in a single open-source chain with per-tree quality and uncertainty diagnostics exposed alongside every reported value (Sections 2.4.4, 2.4.5 and 4.2).

The remainder of the paper is organized as follows. Section 2 describes the study area, the FJD Trion S1 acquisition, the destructive reference, the five-phase pipeline, and the calibration. Section 3 reports detection, DBH, height and volume accuracy, and the head-to-head comparison against manufacturer software and a manual workflow. Section 4 interprets the findings in the context of the HMLS literature, identifies failure modes (algorithmic and data-side), and discusses limitations and future work. Section 5 summarizes some conclusions.

2. Materials and Methods

2.1. Study Area

The area under study contains naturally regenerated maritime pine (*Pinus pinaster* Aiton) stands, located in the Alge area, which is part of the Serra da Lousã mountain range in the districts of Leiria and Coimbra, in central Portugal (Figure 1). The five plots used in this paper span an elevation range from 620 to 795 m and a topographic envelope from steep south-west-facing slopes above 30° (Alge1 and Alge2) to gentler east-, south-east- and south-facing slopes between 0° and 10° (Alge4, Alge5 and Alge6); per-plot coordinates, elevation, aspect and slope class are reported in Table 1. The stands are representative of the Mediterranean–Atlantic transition climate, with cool, wet winters and warm, dry summers typical of inland central Portugal at this elevation band. Five plots from the Alge stand block (Alge1, Alge2, Alge4, Alge5, Alge6) were destructively sampled; four are retained for the calibrated reference in this paper (Alge1, Alge4, Alge5, Alge6; $n = 45$); the fifth (Alge2; $n = 14$) was set aside because its trees were extracted by a method that damaged the stems before the destructive measurements were taken, and it is reported throughout as a sensitivity analysis (Sections 3.3.4 and 4.5). The plot identifier “Alge3” is not used in this paper because the corresponding plot was surveyed under a non-destructive measurement protocol and therefore cannot be brought into the destructively referenced evaluation reported here.

Table 1 presents the per-plot composition of the destructive reference and the geographic and topographic characterization of the plots. n denotes the number of destructively measured *P. pinaster* trees per plot. DBH, height and volume ranges are reported directly from the destructive reference values. Coordinates are in WGS-84 degrees–minutes–seconds (DMS); the underlying spatial reference system is EPSG:4326. Values were measured with a handheld GPS at each plot centroid during the fieldwork. Aspect classes are the dominant exposure direction; slope class is the dominant inclination band of the plot.

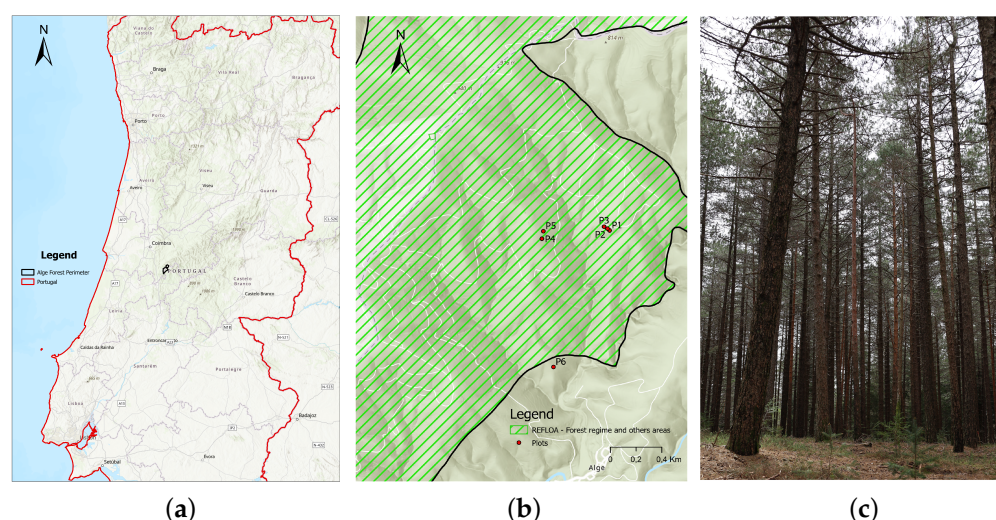


Figure 1. Study area. (a) Location of the Alge area within mainland Portugal; (b) location of the sampled plots within the Alge forest perimeter; (c) stand structure of the naturally regenerated maritime pine at one of the plots.

Table 1. Plot characteristics.

Plot	<i>n</i>	Coordinates (WGS-84)	Elevation (m)	Aspect	Slope (°)	DBH (cm)	Height (m)	Volume (m ³)
Alge1	12	40°02′19.1″ N 8°15′35.7″ W	740	SW	>30	12.3–32.6	13.6–17.6	0.08–0.62
Alge2	14	40°02′19.5″ N 8°15′36.5″ W	740	SW	>30	15.4–34.9	13.7–17.8	0.14–0.55
Alge4	15	40°02′17.1″ N 8°15′57.5″ W	795	SE	0–10	14.2–32.5	13.3–17.7	0.11–0.64
Alge5	4	40°02′18.8″ N 8°15′57.0″ W	795	E	0–10	19.8–23.3	14.4–15.9	0.20–0.33
Alge6	14	40°01′44.3″ N 8°15′54.6″ W	620	SE–S–SW	0–10	21.5–38.0	16.3–21.5	0.31–0.93
Total	59	-	620–795	-	0–30	12.3–38.0	13.3–21.5	0.08–0.93

2.2. Field Data Acquisition

Point clouds were acquired with an FJD Trion S1 handheld mobile laser scanner (HMLS), operating under simultaneous localization and mapping (SLAM). The instrument integrates 16 laser channels with a $270^\circ \times 360^\circ$ field of view, a maximum range of 120 m and a nominal acquisition rate of $320,000$ points s^{-1} . The manufacturer specifies a relative accuracy of the post-processed reconstruction of approximately 2 cm, measured under controlled conditions; no independent accuracy certification is available for forest environments, and the manufacturer explicitly notes that results may differ outside the test environment. This nominal 2 cm relative accuracy is of the same order as the caliper uncertainty of the destructive reference at small radii, and therefore sets a floor on the DBH agreement that no processing algorithm can improve upon. The operator followed a pre-defined walking trajectory within each plot, maintaining approximate visual contact with every standing stem to minimize occlusion. All five acquisitions were conducted in April 2024 under fair weather conditions, with weak-to-no wind and no precipitation during the scanning windows. The precise hour of each scan was not logged. Clouds were exported in LAS 1.4 format. The exported files contain the per-scan SLAM coordinates referenced to the trajectory start point; point coordinates are in local metric units

(typical range ± 60 m about the scan origin), the LAS header records no Variable Length Records (VLRs) carrying a Coordinate Reference System (CRS), and the file offset is zero. Each plot is therefore self-contained in its own local frame; cross-plot georeferencing is not required for the per-tree analysis reported here, which operates within each plot's frame. The geographic coordinates reported per plot in Table 1 (WGS-84, EPSG:4326) were measured separately with a handheld GPS at each plot centroid during the field campaign, independently of the LAS export. For downstream integration with GIS layers, the local LAS frame can be aligned against the corresponding GPS centroid as a registered ground control point in a separate processing step that is not part of the present pipeline.

Point-cloud completeness was assessed per tree rather than per plot, because it is the local completeness around each stem that limits the reconstruction. Two indicators are carried through the pipeline and reported with every tree: the extrapolated fraction, defined as the fraction of the tree height above the highest RANSAC-reliable slice, and a three-criterion quality flag (Section 2.4.5). Across the 45 retained trees, the extrapolated fraction has a median of 0.31 and an interquartile spread of 0.25–0.39, with three trees above 0.5; the quality flag is high for 6 trees, medium for 38 and low for 1. The spatial pattern behind these numbers is visible in Figure 2: return density is highest along the walking trajectory and decays with horizontal distance from it, and the same decay is visible vertically within the canopy. No multi-scan registration was performed, so no inter-scan misalignment can be introduced: each plot is a single continuous SLAM acquisition, and any residual drift is internal to the SLAM solution rather than a registration artifact.

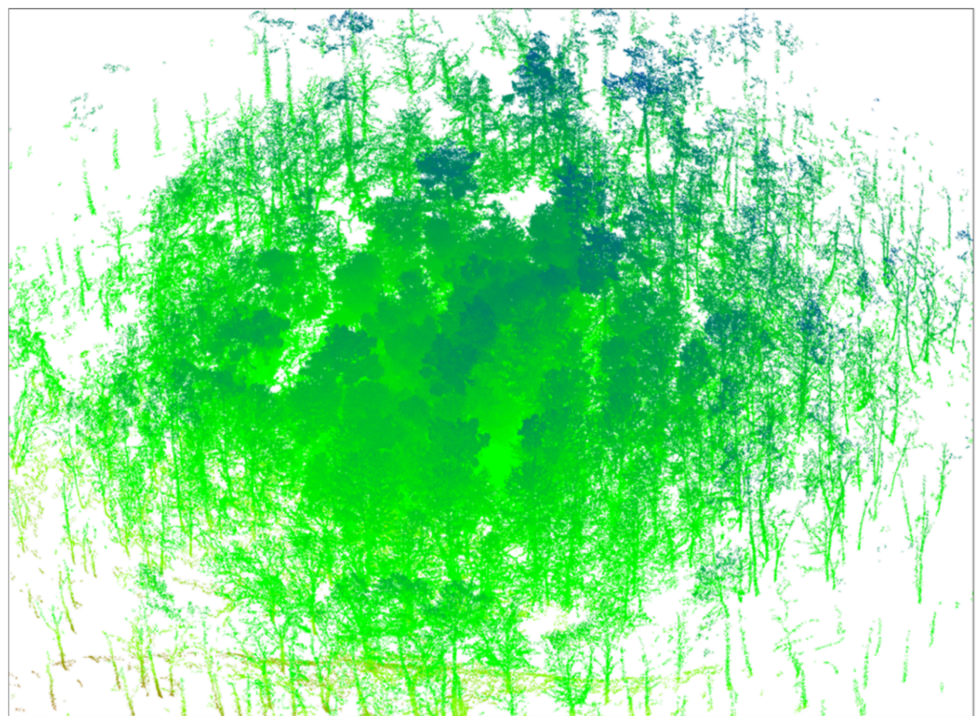


Figure 2. Raw FJD Trion S1 acquisition of one plot, seen from above and colored by height above ground (green in the lower crown grading to blue at the top). Point density is highest along the walking trajectory near the plot center and decays with horizontal distance from it, so that stems at the margin of the acquisition are only partially sampled; the same decay occurs vertically, the lower crown being densely sampled while the upper canopy, where the scanner has no line of sight from the walking trajectory, is only sparsely represented.

2.3. Destructive Ground Truth

After scanning, 59 maritime pine trees were extracted across the five Alge plots and sectioned: (i) in the field at known heights, to support stem-volume reconstruction by Smalian-type integration of caliper diameters, or (ii) measured for total stem volume by the project field protocol. The reference volumes for the 59 trees range between 0.076 and 0.934 m³, the reference DBH between 12.3 and 38.0 cm, and the reference height between 13.31 and 21.46 m. In Alge2, the extraction was performed with a tractor and grapple rather than by felling, and the stems were damaged in the process: the grapple crushes and tears the trunk at the gripping points and material is lost along the stem, so the caliper sections measured afterwards no longer correspond to the intact tree that had been scanned. Together with confirmed inconsistencies in the field records of this plot, this makes the Alge2 destructive measurements unreliable as a reference and motivates the exclusion of the plot from the calibration sample; the evidence is set out in Section 4.5 and the effect on every reported figure is in Section 3.3.4.

The destructive reference is treated as the best available reference rather than as absolute truth: caliper error at small radii reaches several millimeters, section locations can vary by 1–2 cm across measurement events, Smalian-type integration is a discrete approximation, and the mechanical handling of the stems during extraction and sectioning introduces additional uncertainty.

A further source of uncertainty is the circular cross-section assumption implicit in the caliper protocol: sectional areas are computed as circles from one or two caliper directions, whereas stems in naturally regenerated stands are frequently elliptical or irregular. This assumption is standard practice in forest inventory and is applied identically to the reference and to the pipeline output, so it does not bias the comparison; it does, however, place a floor on the accuracy attainable against this reference, and it cannot be quantified without a synchronous high-resolution scan of the same stems.

2.4. Processing Pipeline

The pipeline is implemented as a five-phase Python workflow (Phases 1, 2, 2.5, 3 and 4), illustrated in Figure 3 and orchestrated by a single entry-point script that resolves a species preset before execution. The species preset modifies the parameters that depend on stem morphology: the DBH search range, the minimum stem radius used for candidate detection, the graph edge length used by TLS2trees, the cylindrical radius used for pre-labeling, and the inlier-count threshold for stem detection. The results reported in this study use the *Pinus pinaster* preset; the values of the species-dependent parameters of the preset are listed in Table 2, and the TLS2trees instance-segmentation parameters are listed in Table 3. All processing was performed on a single workstation (HP Z4 G5; Intel Xeon w3-2423 at 2.11 GHz; 64 GB DDR5 RAM; NVIDIA RTX A2000 12 GB; Windows 11 64-bit) under Python 3.11. The pipeline is single-machine and does not require distributed execution. CUDA acceleration is used by the Phase-2 TLS2trees instance segmentation [23]. The remaining four phases, Phase 1 (PDAL), Phase 2.5, Phase 3 and Phase 4, are implemented in NumPy/SciPy/scikit-learn and run entirely on CPU. The pipeline is therefore portable across hardware tiers: it runs unchanged on workstations without a CUDA-capable GPU, at the cost of a longer processing time for the Phase-2 step, and benefits from a GPU when one is available.

Table 2. *Pinus pinaster* species preset—values of the species-dependent pipeline parameters used in this study.

Phase	Parameter	Value	Description
1	DEDUPLICATE_VOXEL	0.005 m	Voxel cell size for SLAM duplicate suppression
1	SOR_NEIGHBORS/SOR_MULTIPLIER	8/1.2	Statistical outlier removal parameters
1	CSF_RESOLUTION/CSF_RIGIDNESS	0.30 m/2	Cloth Simulation Filter ground segmentation
2.5	C_GROUND_CELL/C_GROUND_PCT	0.5 m/5%	Per-tree base height KD-tree window
2.5	C_MIN_DBH_M	0.10 m	Shrub-filter minimum DBH
2.5	H_LEAFON_CORRECTION	1.0270	Leaf-on height calibration factor
3	SLICE_STEP	0.10 m	Slice thickness for per-slice RANSAC
3	RANSAC_N_ITER	200	RANSAC iterations per slice
3	RANSAC_BAND_LO/RANSAC_BAND_HI	0.30/3.00	Sequential-prior annular admission band
3	MIN_QUALITY/MIN_ANG_COV	0.55/0.35	Slice retention thresholds
3	SKEL_DRIFT_MAX	0.06 m	Continuity-filter center-track drift bound
3	TAPER_FIT_H_MIN	0.70 m	Base-region exclusion height for taper fit
3	Species shape coefficients	(2.7342, −7.8921, 6.9028, −2.7450, 1.0)	Polynomial $P(x)$ of Equation (2), <i>P. pinaster</i>
4	E_INTEGRATION_STEP	0.01 m	Trapezoidal integration step
4	FF_LO/FF_HI	0.34/0.62	Form-factor sanity-clamp bounds
4	VOLUME_CALIB_GLOBAL	0.9518	Single-factor volume calibration (Equation (4))
4	DBH_REPORT_CALIB_GLOBAL	1.0549	Single-factor DBH calibration (Equation (5))

Table 3. TLS2trees instance-segmentation parameters for the *P. pinaster* preset.

Parameter	Value	Justification
find-stems-boundary	(1.6, 3.0) m	Raised above the TLS2trees default of (1.0, 2.0) m to avoid shrub trunks below 1.6 m being mis-identified as stems
find-stems-min-radius	0.05 m	Filters stems thinner than 5 cm radius
find-stems-min-points	200	Minimum cluster size for a candidate stem
graph-edge-length	0.7 m	Covers stem leans up to $\approx 65^\circ$ while remaining stricter than the TLS2trees default of 1.0 m
graph-maximum-cumulative-gap	35.0 m	Maximum stem height covered
min-points-per-tree	8 000	Filters small instance clusters corresponding to shrubs or scan noise
instance-tile-size	3×3 tiles	Default TLS2trees tile configuration
slice-thickness (graph)	0.3 m	Default TLS2trees graph-construction slice
add-leaves voxel/edge	0.5/1.0 m	Default TLS2trees leaf-attachment parameters

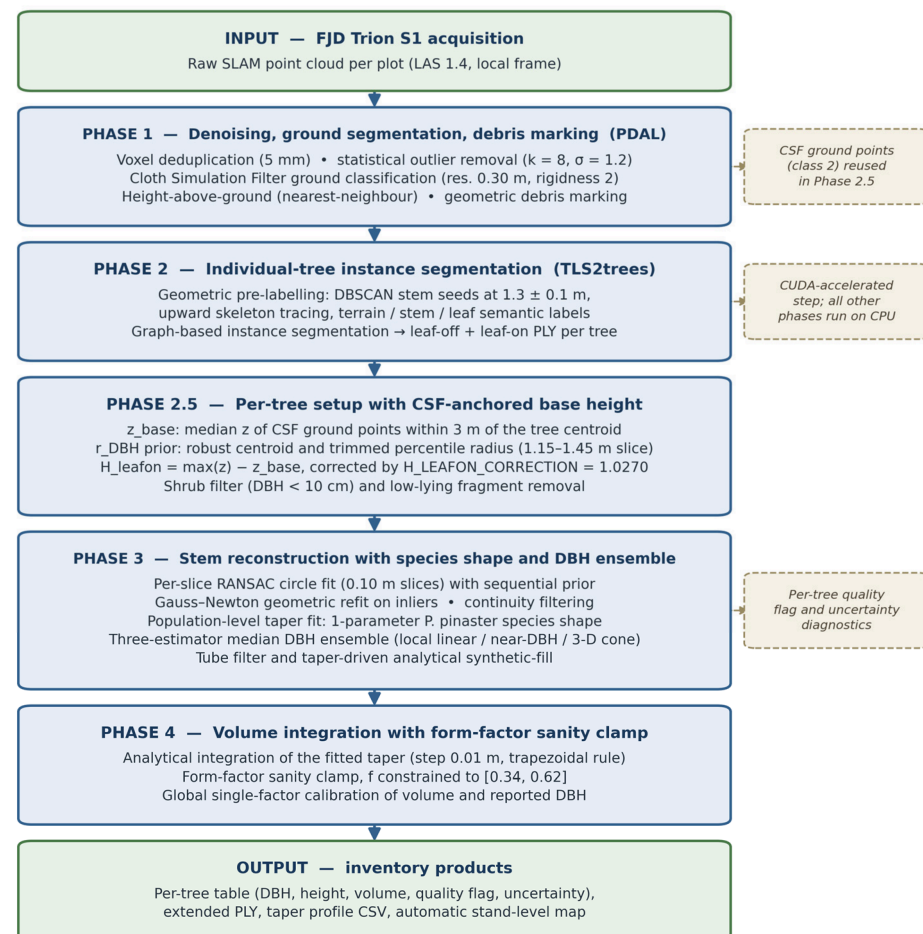


Figure 3. Five-phase architecture of the proposed open-source HMLS processing pipeline, from the raw SLAM acquisition to the per-tree inventory table and the automatic stand-level map.

A design principle that recurs across the pipeline is the asymmetric use of the same intermediate representation for different downstream purposes. The slice-wise per-tree profile, in particular, is read in four distinct modes within Phase 3: (i) as a sequential prior for the next RANSAC search; (ii) as input to the population-level taper fit (with an explicit exclusion of base-region slices that are otherwise retained); (iii) as a skeleton for tube filtering and synthetic completion; and (iv) as one of three independent DBH estimators. Each mode applies its own acceptance criteria to the same underlying data.

2.4.1. Phase 1, Denoising, Ground Segmentation, and Debris Marking

Phase 1 is a deterministic PDAL [39] pipeline that performs voxel deduplication, statistical outlier removal, ground classification, and height normalization, followed by a pipeline-specific geometric debris-marking step. Voxel deduplication uses a 5 mm cell (`filters.voxelcenternearestneighbor`), retaining one point per cell to suppress the SLAM duplicates introduced by loop closures. Statistical outlier removal (`filters.outlier`, $k = 8$ neighbors, σ -multiplier = 1.2) marks outliers as classification 7 and filters range removes them. Ground segmentation uses the Cloth Simulation Filter (CSF; Zhang et al. [20]) with resolution 0.3 m, rigidity 2 (medium), 500 iterations, and smoothing enabled. Height above ground (HAG) is computed with the nearest-neighbor estimator (`filters.hag_nn`).

A geometric debris filter suppresses low-lying branches, forest-floor residues and scanning artifacts. The output LAS carries the standard classification (terrain = 2 for CSF-identified ground) and a debris label in an extra dimension. Figure 4 illustrates the two principal Phase-1 products on one of the Alge plots: the height-above-ground normalization

of the cloud and the CSF ground classification. A specific output of Phase 1 used downstream is the set of CSF-classified ground points, which Phase 2.5 samples at each tree centroid to compute a per-tree base height that is consistent across the plot (Section 2.4.3).

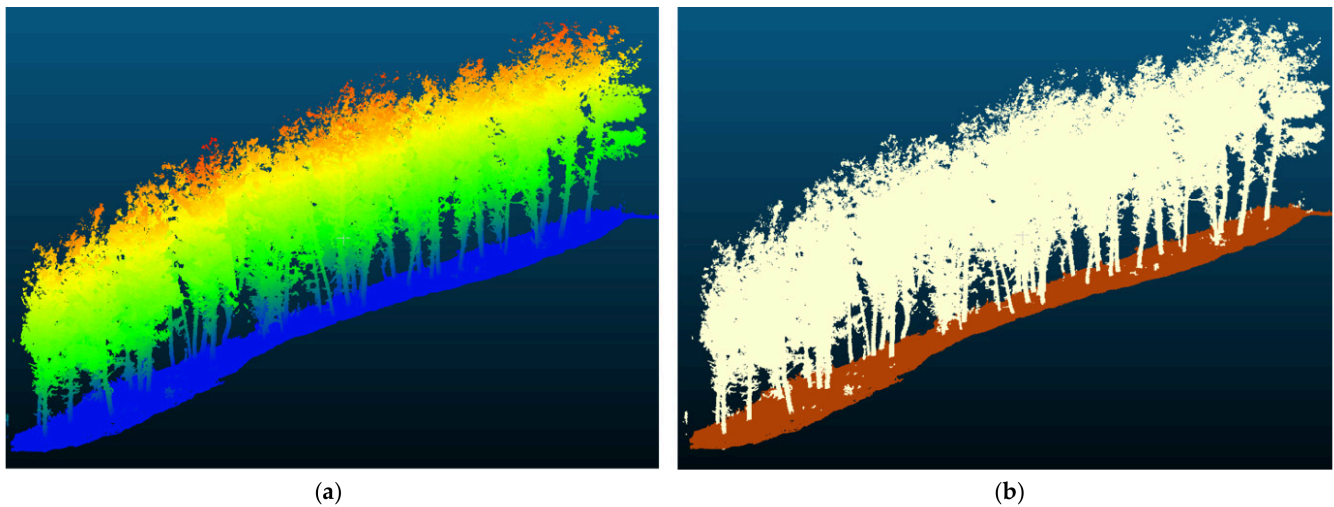


Figure 4. Principal outputs of Phase 1 on one of the Alge plots. (a) Point cloud after denoising, colored by height above ground; (b) result of the Cloth Simulation Filter, with the classified ground points highlighted against the remaining returns.

2.4.2. Phase 2, Tree Instance Segmentation

Instance segmentation couples a geometry-based pre-labeling step with the graph-based tree instance segmentation of TLS2trees (Wilkes et al. [23]; tls-tools-ucl/TLS2trees). Pre-labeling assigns terrain/stem/leaf labels by skeleton tracing from the DBH slice: candidate stems are identified at 1.3 ± 0.1 m by DBSCAN ($\epsilon = 0.1$ m, MinPts = 20) and each candidate centroid is traced upward through 0.25 m vertical slices toward the canopy apex with a 0.5 m horizontal search radius. Points below 0.3 m are labeled terrain; points within 0.20 m of the DBH centroid or of the traced skeleton are labeled stem; the remaining above-ground points are labeled leaf. The labeled PLY is passed to TLS2trees instance.py with the *P. pinaster* preset parameters listed in Table 3, which returns a leaf-off and a leaf-on PLY per detected tree. Figure 5 shows both steps on one plot.

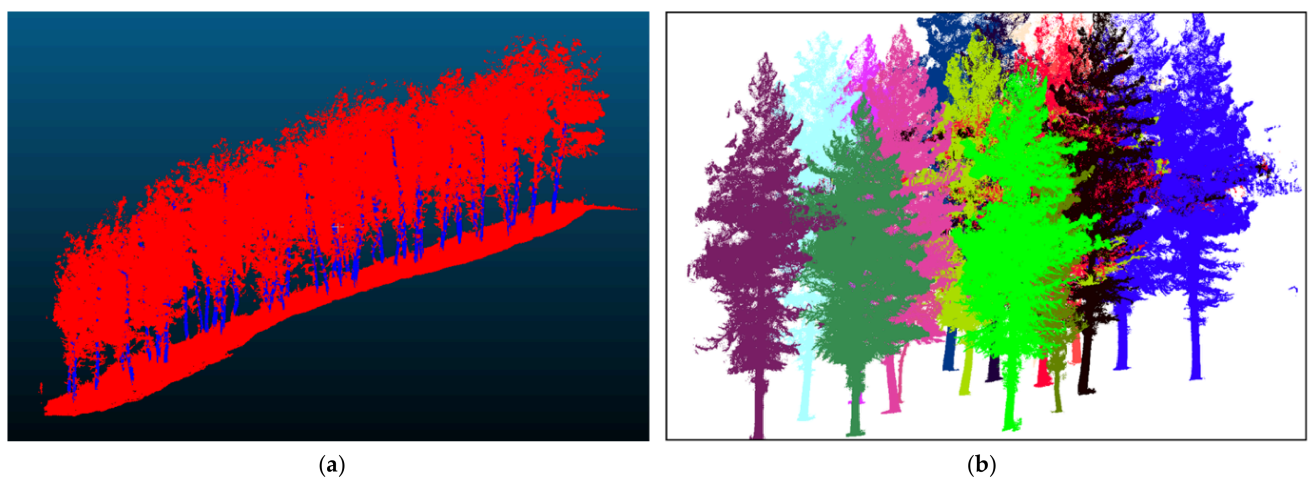


Figure 5. Phase 2 on one of the Alge plots. (a) Geometric pre-labeling, with the points labeled as stem separated from ground, branches and foliage; (b) output of the TLS2trees graph-based instance segmentation, with one color per individual tree.

2.4.3. Phase 2.5, Per-Tree Setup with CSF-Anchored Base Height

Phase 2.5 prepares each segmented tree for Phase 3 by computing three per-tree scalars, summarized in Table 4. The principal change with respect to earlier versions of this pipeline is the source of the per-tree base height.

Table 4. The three per-tree scalars computed by Phase 2.5.

Scalar	How It Is Computed	Role Downstream
Base height z_{base}	Median z of the Phase-1 CSF ground points within a 3 m horizontal radius of the tree centroid (or the 25 nearest points if fewer than five fall inside the radius), retrieved from a per-plot 2-D KD-tree on the (x, y) of the points classified as ground.	Origin of the height axis for every slice of Phase 3, and therefore for the DBH slice and for the total tree height.
DBH prior radius r_{DBH}	On the slice with $\text{HAG} \in [1.15, 1.45]$ m: robust centroid by trimming to the 80th percentile of the distances to the median centroid, then the median of percentiles 60–85 of the point-to-centroid distances.	Prior for the RANSAC search at the first slice of Phase 3. It is not the reported DBH, which comes from the three-estimator ensemble of Section 2.4.4.
Leaf-on tree height H_{leafon}	$H_{\text{raw}} = \max(z_{\text{leafon}}) - z_{\text{base}}$, scaled by $H_{\text{LEAFON_CORRECTION}} = 1.0270$ (Section 2.5). Percentile 99.5, percentile 99 and conical extrapolation of the apex were tested and rejected as worse than $\max(z)$ for these canopies.	Reported tree height and upper integration limit of the volume of Section 2.4.5.

The base height is the single most consequential of the three. Because every slice height is measured as $z - z_{\text{base}}$, an inconsistent z_{base} propagates as a per-tree offset through the whole stem profile, including the DBH slice. Anchoring it on the Phase-1 CSF ground instead of on the per-tree cloud minima removes a per-tree variance of order 1–2 dm. Measured in isolation, by re-running the pipeline with the CSF anchor disabled, the change moves volume MAPE from 10.7 to 9.9% and DBH MAPE from 5.3 to 4.0% under LOPO evaluation, and the uncalibrated DBH bias from -11% to -5% . A large fraction of the bias that the earlier calibration was correcting was therefore a height-misalignment artifact rather than a geometric error.

Phase 2.5 also rejects candidate trees with DBH below 10 cm (shrub filter) and filters low-lying shrub fragments within $2.5 \times r_{\text{DBH}}$ of each retained trunk. Output is `*_clean.leafoff.ply` plus `*_meta.json` per tree.

2.4.4. Phase 3, Stem Reconstruction with Species Shape and DBH Ensemble

Phase 3 reconstructs the external stem profile in six steps: per-slice RANSAC circle fitting with a sequential prior and a geometric Gauss–Newton refit on inliers; continuity filtering of the slice sequence; population-level taper fit using the species shape; a three-estimator DBH ensemble; a tube filter; and a taper-driven synthetic-fill regularization.

Per-slice RANSAC with sequential prior and geometric refit. The cleaned leaf-off cloud is sliced at $\text{SLICE_STEP} = 0.10$ m. At each slice, a RANSAC circle fit is performed with $\text{RANSAC_N_ITER} = 200$ iterations. The fit is constrained by the sequential prior, the center (seq_{cx}, seq_{cy}) and radius seq_r , accepted at the immediately preceding slice, through three mechanisms: an inlier-search annulus of $[0.30 \times seq_r, 3.00 \times seq_r]$ (widened from $[0.40, 2.50]$ in a parameter sweep validated by leave-one-plot-out), radius bounds $[seq_r \times 0.45, seq_r / 0.45]$, and a center-drift bound of $seq_r \times 1.5$. The prior is initialized at

the first slice with the median in-slice centroid and r_DBH from Phase 2.5 and propagates upward as each new slice is accepted.

After the RANSAC core has identified inliers, the radius and center are refined by a **Gauss–Newton geometric refit on the inliers**, minimizing $\sum_i$ with R as the mean inlier distance. This refit replaces the “best 3-point circle” of the standard RANSAC output, which on partial-arc slices is geometrically biased; the Gauss–Newton refit uses the full inlier set and yields a substantially less noisy radius sequence. A slice is *retained* if its fit quality (inlier fraction) exceeds 0.55 and its angular coverage exceeds 0.35.

Continuity filtering. Retained slices are subject to a second filter on the spatial coherence of the center track: the sequence of (cx, cy) is smoothed by a five-slice rolling median; slices whose center deviates from the smoothed track by more than $SKEL_DRIFT_MAX = 0.06$ m are rejected as inconsistent with a single coherent stem.

Population-level taper fit, species shape. Slices that survive the geometric and continuity filters and lie above $TAPER_FIT_H_MIN = 0.70$ m are passed to the population-level taper fit. The base region ($h < 0.70$ m) is excluded from the fit (to prevent buttressing from inflating the base radius) but is retained for the skeleton trace, the tube filter, and the synthetic-fill.

The principal taper parameterization is the *P. pinaster* species shape:

$$r(h) = k \cdot P\left(\frac{h}{H}\right) \quad (1)$$

where P is a 4th-degree polynomial that captures the normalized stem-form $d(h)/d_0$ of *P. pinaster*. The polynomial is calibrated on 366 caliper observations across 59 destructively measured trees from the same stand:

$$P(x) = 2.7342x^4 - 7.8921x^3 + 6.9028x^2 - 2.7450x + 1, R^2 = 0.9705 \quad (2)$$

with $x = h/H \in [0, 1]$. The single free parameter k (the base half-diameter, r at $h = 0$) is estimated as the median of $r_i/P(x_i)$ across all retained slices. The use of the median rather than a least-squares fit is deliberate: it gives the species shape robustness against contaminated slices that would defeat an unconstrained free fit, by ensuring that a small subset of high-residual slices cannot pull the entire reconstructed profile. Two alternative parameterizations are retained as configuration options for species or regions without a calibrated polynomial: a 2-parameter power-law $r(h) = r_0$ and the 7-parameter Kozak [28] variable-exponent taper.

DBH ensemble (median of three estimators). The DBH reported by the pipeline is the median of three independent estimators, each strong in a different regime and each failing differently (Table 5). The median is itself the regularizer: it discards any one estimator that diverges from the other two, which is what decouples the reported DBH from the failure modes of the volume taper model.

Tube filter and synthetic-fill regularization. The fitted taper defines an expected radius $r_{\text{fitted}}(h)$ along the full tree height. Two post-fit cleaning operations exploit this: a tube filter that removes points beyond $r_{\text{fitted}}(h) \times 1.25$ from the smoothed skeleton, and a taper-driven synthetic-fill that completes empty angular sectors at each slice using synthetic points generated on $r_{\text{fitted}}(h)$ with small radial and vertical jitter ($\sigma_r = 3$ mm, $\sigma_z = 8$ mm). The synthetic points are sampled from the fitted analytical taper, not from a learned distribution as in Bornand et al. [35] and not by angular interpolation between neighboring sectors. This produces a topologically complete representation of the stem surface for downstream mesh construction without introducing volume information beyond what the fitted taper already carries.

Table 5. The three DBH estimators combined by the median ensemble of Phase 3.

Estimator	Definition	Strength	Failure Mode
Robust local linear	Robust linear regression of the radii of the retained slices in $h \in [0.9, 2.5]$ m, evaluated at $h = 1.30$ m and refined by two passes of MAD outlier removal.	Stable on monopodial stems with a continuous slice sequence.	Degenerates when the slice sequence has gaps.
Near-DBH median	Median radius of the retained slices within $h \in [1.15, 1.45]$ m, widened to $[1.0, 1.6]$ m if too few slices survive.	Simplest and least model-dependent of the three.	Exposed to contamination of a single slice.
3-D cone fit	Huber-loss least-squares fit of a 6-parameter 3-D cone to all in-band points in $h \in [0.9, 2.5]$ m, returning the radius at $h = 1.30$ m.	Most resistant to partial-arc bias, since it uses one axis for the whole band.	Fails ungracefully on multi-stem or forked trees.

Figure 6 follows one tree through the chain, from the segmented leaf-on cloud to the reconstructed stem surface. The output of Phase 3 is, per tree: an extended PLY containing the cleaned stem points plus the synthetic completion vertices; a JSON metadata file carrying the fitted parameters, residual sigma, retained-slice count, extrapolated fraction, and per-tree quality flag; and a discrete profile CSV sampling $r_{\text{fitted}}(h)$ at 0.10 m intervals.

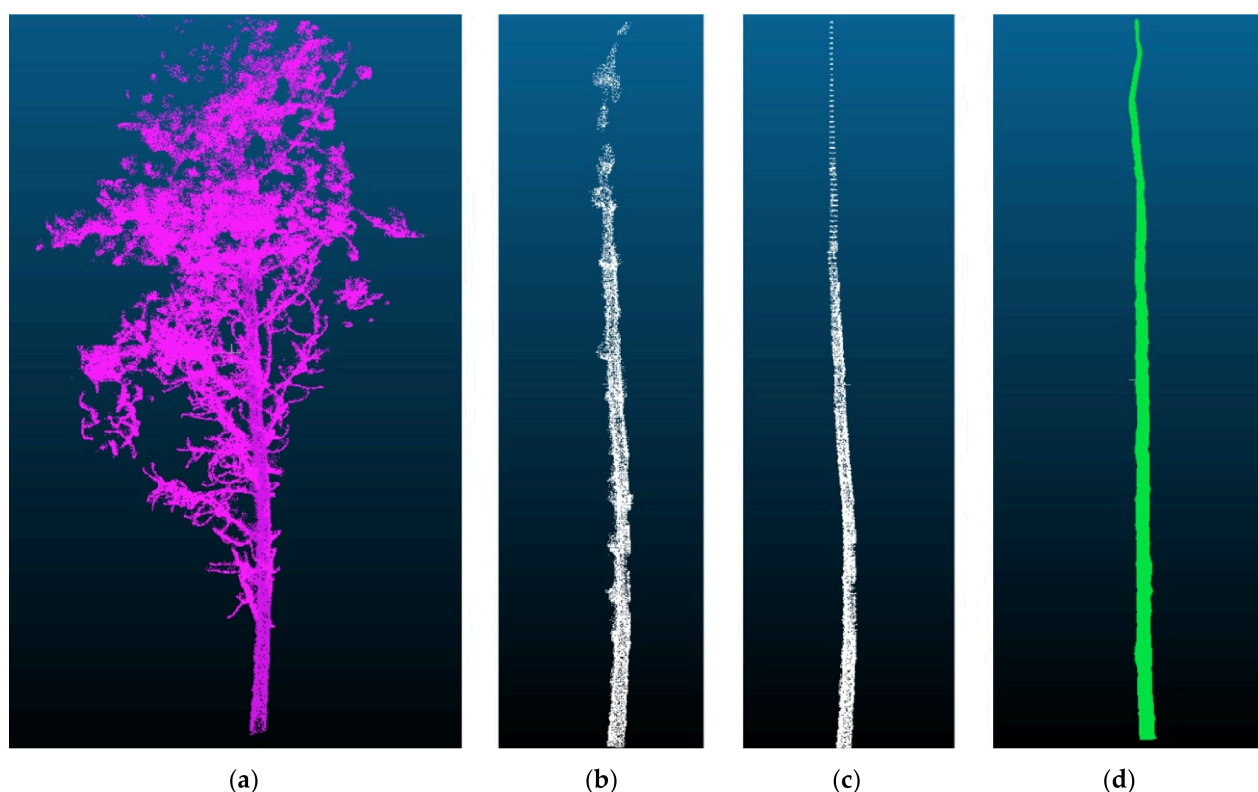


Figure 6. Stem reconstruction of a single tree. (a) Tree as delivered by the Phase-2 instance segmentation, leaf-on; (b) cleaned stem points after the Phase-2.5 setup; (c) same stem after the taper-driven synthetic-fill, with the empty angular sectors completed from the fitted species shape; (d) reconstructed stem surface used for the volume integration of Section 2.4.5.

2.4.5. Phase 4, Volume Integration with Form-Factor Sanity Clamp

Phase 4 computes per-tree volume by analytical integration of the fitted taper profile:

$$V_{\text{raw}} = \int_0^H \pi \cdot r_{\text{fitted}}^2 \, dh \quad (3)$$

where V_{raw} is the uncalibrated stem volume, $r_{\text{fitted}}(h)$ is the radius of the fitted taper at height h above the tree base, and H is the total tree height. The integral is evaluated numerically with an integration step of 0.01 m and the trapezoidal rule.

Form-factor sanity clamp. A sanity clamp on the implied form factor $f = V / (\pi/4 \cdot \text{DBH}^2 \cdot H)$ is applied, where V is the stem volume, DBH the reported diameter at breast height, and H the total tree height. If f lies outside the *P. pinaster* range [0.34, 0.62], V_{raw} is scaled to bring f to the nearer bound.

The bounds were derived from the destructive reference of the calibration sample: the implied form factor of the 45 retained trees ranges from 0.384 to 0.553, with a mean of 0.450 and a standard deviation of 0.033, so the lower bound lies slightly more than three standard deviations below the mean and the upper bound well above the largest observed value. The clamp is therefore deliberately permissive, and it catches taper-model pathologies that occasionally survive the upstream quality gates. It does not bind on any tree of the destructive set, including the Alge2 plot: the implied form factor of the pipeline estimates ranges from 0.353 to 0.532, so the clamp cannot have altered any statistic reported in Section 3. Its role is as a guard for deployment outside the calibration sample; in an earlier version of the pipeline, in which the taper fit had seven free parameters, it recovered a tree that the free fit had placed at +180% deviation.

Global single-factor calibrations. Two near-unity global factors are applied to the output of the integration:

$$V_{\text{final}} = V_{\text{raw}} \cdot 0.9518 \quad (4)$$

$$\text{DBH}_{\text{reported}} = \text{DBH}_{\text{ensemble median}} \cdot 1.0549 \quad (5)$$

In Equations (4) and (5), V_{final} and $\text{DBH}_{\text{reported}}$ are the calibrated values delivered by the pipeline, V_{raw} is the volume obtained from the integration of Equation (3), and $\text{DBH}_{\text{ensemble median}}$ is the median of the three DBH estimators of Section 2.4.4. The two multiplicative factors are the log-mean of the ratio between the destructive reference and the raw pipeline output across the 45-tree calibration sample (cf. Section 2.5). In earlier versions of this pipeline, the correction had to be a step function ($\times 0.84$ below and $\times 1.16$ above a DBH of 16 cm) combined with a per-model split ($\times 1.05 / \times 1.10$), because the raw output was biased in different directions for different tree sizes and for different taper models. In the present design, a single scalar close to unity is sufficient for each variable. Physically, this means that the geometry the pipeline reconstructs is already close to the real stem: the upstream design changes (CSF-anchored z_{base} , species-shape taper, DBH ensemble, Gauss–Newton geometric refit, form-factor clamp) removed the size-dependent and model-dependent distortions that the earlier step calibration had been compensating for, and what remains is a small, uniform residual scale.

Per-tree quality and uncertainty. Each tree carries a quality flag (high, medium, low) derived from a count of three criteria: DBH angular coverage $\geq 50\%$, extrapolated fraction below 20%, residual sigma below 20 mm, and an uncertainty value that grows linearly with the extrapolated fraction (base 5% + $30 \times \text{frac}_{\text{extrapolated}}$).

2.5. Calibration of Empirical Constants

Four empirical constants absorb the residual systematic bias of the pipeline: a multiplicative correction of the estimated total height ($H_{\text{LEAFON_CORRECTION}}$, Phase 2.5),

a multiplicative correction of the integrated stem volume (VOLUME_CALIB_GLOBAL, Phase 4), a multiplicative correction of the reported DBH (DBH_REPORT_CALIB_GLOBAL, Phase 4), and the lower and upper bounds of the form-factor sanity clamp. All four were calibrated against the 45-tree destructive reference. We adopt a single-factor global calibration (a multiplicative scalar per variable) rather than the per-class step function used in earlier versions of this pipeline family, because the upstream design changes described above produced an output that the global scalar absorbs the residual bias on at the LOPO-evaluated level without per-class structure. The calibration constants are: H_LEAFON_CORRECTION = 1.0270 (a 2.7% aggregate height correction); VOLUME_CALIB_GLOBAL = 0.9518 (a 3.4% aggregate volume correction); and DBH_REPORT_CALIB_GLOBAL = 1.0549 (a 5.5% aggregate DBH correction). The form-factor clamp bounds are [0.34, 0.62], derived from the destructive set.

Under the leave-one-plot-out protocol, all four constants are re-estimated from scratch inside each training fold, using only the three plots that form that fold: H_LEAFON_CORRECTION, VOLUME_CALIB_GLOBAL and DBH_REPORT_CALIB_GLOBAL are recomputed as the log-mean of the reference-to-raw ratio over the training trees, and the form-factor bounds are re-derived from the training trees alone. The held-out plot contributes to none of these estimates. The species-shape polynomial of Equation (2) is the one exception and is not re-fitted per fold; it is calibrated once on the caliper observations of the destructive set, so the LOPO figures are not fully independent of the destructive data in that single respect. This residual dependency is stated as a limitation in Section 4.5.

The four constants are near unity. This indicates that the raw bias of the geometric pipeline on the calibration sample is small in magnitude, and that the per-class step calibration used in earlier versions of this pipeline family is no longer required to absorb it. We mention this as an indication of the upstream design's coherence on the present dataset, not as evidence of out-of-sample unbiasedness, which is what the LOPO statistics of Section 3 are intended to characterize.

2.6. Validation Metrics

Detection performance was evaluated on the 45 destructively measured reference trees by matching reconstructed tree positions to the destructive reference positions using a horizontal proximity criterion of 2 m on the DBH-centroid coordinates (cx_{DBH}, cy_{DBH}) from the per-tree taper metadata. Because the destructive reference is a planned subsample of each plot rather than an exhaustive census, the unmatched pipeline instances cannot be labeled as true or false positives: an instance without a destructive counterpart may be a correctly detected tree that was simply not felled, or a genuine commission error, and the two cases are indistinguishable without a full stem map of each plot. Precision and F1-score are therefore not estimable on this dataset, and reporting them would require assuming that every unfelled instance is a false positive, which is demonstrably wrong. What is estimable is Recall over the reference subsample, which we report as Detection: the fraction of destructively measured trees recovered by the pipeline.

DBH and height accuracy were assessed against field measurements using mean absolute error (MAE), root-mean-square error (RMSE), and bias (mean signed error), and the percentage equivalents (MAPE, MedAPE, relative RMSE, signed bias as percentage). Volume accuracy was reported against (i) the 45 destructively measured retained trees and (ii) the full $n = 59$ sample as a sensitivity analysis, in terms of MAPE, MedAPE, signed bias, RMSE, Lin's concordance correlation coefficient (CCC), and Pearson's r . Uncertainty in the aggregated error metrics is quantified by a non-parametric bootstrap over trees (10,000 resamples), reported as 95% percentile confidence intervals. Differences between the proposed pipeline and each baseline are tested with the two-sided Wilcoxon signed-

rank test on the per-tree absolute percentage errors of the matched sample, which is paired by construction and makes no distributional assumption.

2.7. Baseline Comparisons

Per-tree DBH, total height, and stem volume produced by the proposed pipeline are compared on the same 45 trees against two operational alternatives: (i) the FJD Trion S1 manufacturer software (FJD Trion Model, release of 24 October 2024), processing the same raw scan files in its default configuration; and (ii) a manual CloudCompare workflow on the same raw scan files, performed by a single forestry engineer with formal training on LiDAR processing. In the manual workflow, the per-tree point clouds previously segmented by the FJD Trion Model are imported into the 3DFIN plugin for CloudCompare, which computes the per-tree dendrometric values.

The 3DFIN parameters were left at the plugin defaults except for the following: cloth resolution set to 0.4; stripe upper limit 3.5 m; stripe lower limit 0.7 m; pruning intensity 2; lowest section 0.3 m; distance between sections 0.2 m; and section width 0.05 m. Per-tree volume was computed from the 3DFIN section diameters by combining Smalian's formula for the stem segments below the canopy top with the cone-volume formula for the apical extension; the same convention was applied to the destructive caliper sections to allow the comparison to be performed on a like-for-like volume estimator. For each method, we report on the same 45-tree sample the same statistics as for the pipeline.

Two caveats apply. First, both baselines were run on the project's distribution (a clip of the original SLAM acquisition), and whether the manufacturer software internally uses the full acquisition is an open question we return to in Section 4.3. Second, the baseline output files provided by the manufacturer software cover 40 of the 45 retained destructively measured trees: five Alge6 trees with identifiers 55–59 in the project numbering are absent from the manufacturer export. The manufacturer software does not log the candidates it rejects, so the reason for their absence could not be established from the exported files; the head-to-head agreement statistics in Section 3.4 are therefore reported on the 40 common trees, with the five absent trees indicated where relevant.

Use of generative artificial intelligence. A large-language-model assistant (Anthropic Claude 4.5 Haiku) was used to support the drafting and editing of the manuscript. All scientific content, data analysis, results, and interpretation are the authors' own. The assistant was used as a writing aid; the authors take full responsibility for the final content of the manuscript.

3. Results

3.1. Tree Detection

Across the four plots retained for quantitative analysis, the pipeline produced 243 individual-tree instances by TLS2trees segmentation (Table 6). Every one of the 45 destructively measured reference trees was detected and successfully matched to a pipeline instance by coordinate proximity (2 m threshold). The remaining 198 instances correspond to field trees that were not felled and therefore have no measured destructive reference. Most are likely to correspond to genuine *P. pinaster* stems standing in the plots at the time of acquisition, with a smaller number of cases where TLS2trees split a stem into two instances or labeled a shrub of marginal size as a candidate tree. Recall over the destructive subsample is therefore $45/45 = 100\%$. The 198 unmatched instances cannot be partitioned into correct detections and commission errors without an exhaustive stem map, as explained in Section 2.6; their count is reported for transparency, but no precision or F1 figure is derived from it.

Table 6. Per-plot detection summary. *npipe* = total pipeline instances; *nGT* = destructively measured reference trees; *nmatched* = trees successfully matched.

Plot	<i>npipe</i>	<i>nGT</i>	<i>nmatched</i>	Notes
Alge1	90	12	12	retained
Alge2	49	14	14	excluded, Section 4.5
Alge4	49	15	15	retained
Alge5	18	4	4	retained
Alge6	86	14	14	retained
Total retained	243	45	45	

3.2. DBH and Height Accuracy

DBH and height were assessed against destructively measured reference values on the 45 retained trees. Summary statistics are reported in Table 7.

Table 7. DBH and height error statistics for the proposed pipeline on the 45 destructively measured *P. pinaster* trees (Alge2 excluded).

Metric	<i>n</i>	MAPE	MedAPE	Bias	RMSE (rel.)
DBH (cm)	45	3.77%	2.46%	+0.13%	5.59%
Height (m)	45	4.04%	3.27%	+0.13%	5.11%

DBH. Aggregated MAPE is 3.77%, with median absolute percentage error (MedAPE) of 2.46% and an effectively zero signed bias (+0.13%). The MedAPE of 2.46%, the central statistic for the typical tree, is below the corresponding figures of both baseline methods (FJD Trion Model 3.59%; manual CloudCompare 4.92%; cf. Section 3.4); on the median tree, the pipeline outperforms both baselines on DBH.

Height. MAPE is 4.04% with bias +0.13%, dominated by the apex-occlusion outliers identified per tree in Sections 3.3 and 4.3.

3.3. Volume Accuracy

3.3.1. Aggregated Agreement

Pipeline volume estimates were compared to the destructive reference values on the 45 retained trees. Aggregated agreement is reported in Table 8 and visualized as 1:1 scatter plots in Figure 7.

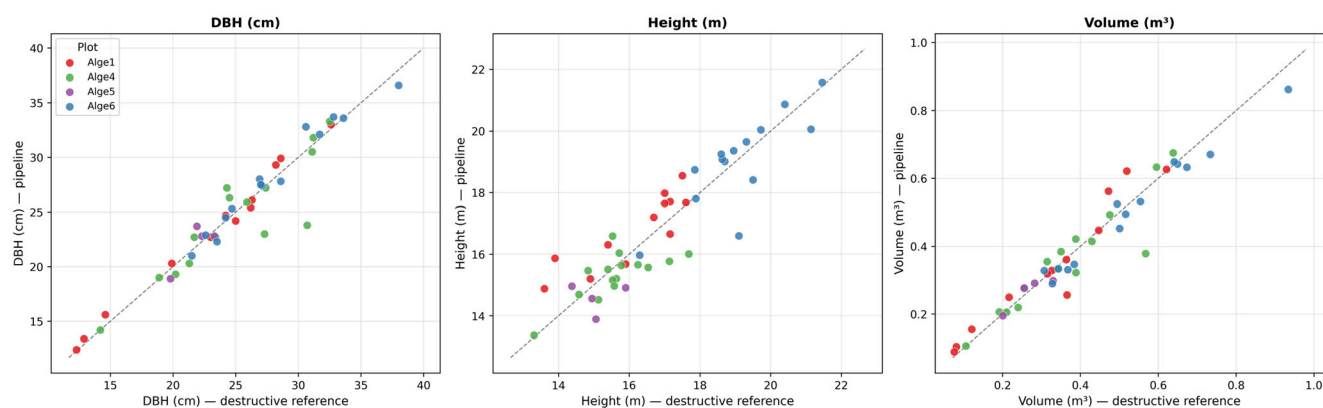


Figure 7. Pipeline-estimated versus destructively measured per-tree DBH, height and volume for the 45 retained *P. pinaster* trees (Alge1, Alge4, Alge5, Alge6). The dashed line represents the 1:1 reference. The pipeline preserves the rank ordering of trees by volume with small error (CCC = 0.965 on the in-sample calibration).

Table 8. Aggregated volume agreement against destructive reference ($n = 45$ *P. pinaster*, Alge2 excluded), under in-sample calibration and under leave-one-plot-out (LOPO) cross-validation. The LOPO statistics are the headline figures of this paper.

Regime	n	Vol MAPE	Vol MedAPE	Vol Bias	DBH MAPE	H MAPE
In-sample	45	9.00%	7.65%	+0.77%	3.8%	4.0%
LOPO	45	9.9%	—	+0.7%	3.9%	4.4%

Under in-sample calibration, volume MAPE is 9.00% (MedAPE 7.65%), with a signed bias of +0.77% and relative RMSE of 12.22%. Under LOPO cross-validation (constants re-fitted on three plots, evaluated on the held-out fourth), volume MAPE is 9.9% with bias +0.7%, DBH MAPE 3.9% and height MAPE 4.4% (Table 8); the LOPO regime is the headline statistic of this paper because it is the only out-of-sample evaluation of the four calibration constants on the final destructive reference. The corresponding R^2 between pipeline and reference volume under the in-sample fit is 0.9315, Lin's CCC is 0.9645, and Pearson's r is 0.9655. Of the 45 trees, 38% lie within $\pm 5\%$ of the reference volume, 76% within $\pm 10\%$, 91% within $\pm 20\%$, and 98% within $\pm 30\%$. The bootstrap 95% confidence intervals of the in-sample aggregated MAPE are [6.7, 11.6] % for volume, [3.2, 5.0] % for height and [2.7, 5.1] % for DBH (10,000 resamples over the 45 trees). The width of the volume interval reflects the influence of the four outlier trees of Section 3.3.3 on a sample of this size.

The LOPO per-plot decomposition is reported in Section 3.3.2 alongside the in-sample per-plot values.

For transparency, the full-sample result including Alge2 ($n = 59$) is reported as a sensitivity analysis at the end of Section 3.3 (after the per-plot decomposition). The defense of the Alge2 exclusion is given in Section 4.5.

3.3.2. Per-Plot Agreement

Per-plot volume agreement is reported in Table 9 and visualized in Figures 8 and 9.

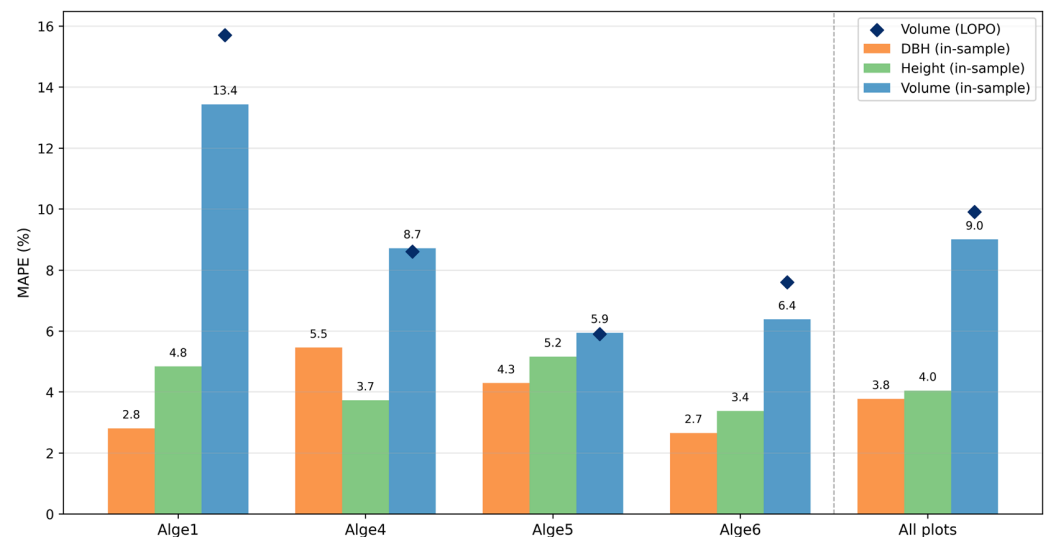


Figure 8. Per-plot MAPE for DBH, height and volume across the four retained plots under in-sample calibration (bars), with the per-plot volume MAPE under leave-one-plot-out cross-validation superimposed as diamonds. The aggregate over the 45 trees is shown on the right.

Alge5 ($n = 4$, with a small-sample caveat) reports the lowest absolute residual MAPE (5.94%). Alge6, with the largest mean DBH (≈ 29 cm) and the most uniform stem distribution, follows closely at 6.39% MAPE with a modest negative bias. Alge4 sits in the middle

(8.71%). Alge1, which contains a long tail of small trees and the trees with the highest per-tree extrapolation, produces the highest per-plot MAPE (13.44%), with most of the residual concentrated in the outliers itemized in Section 3.3.3.

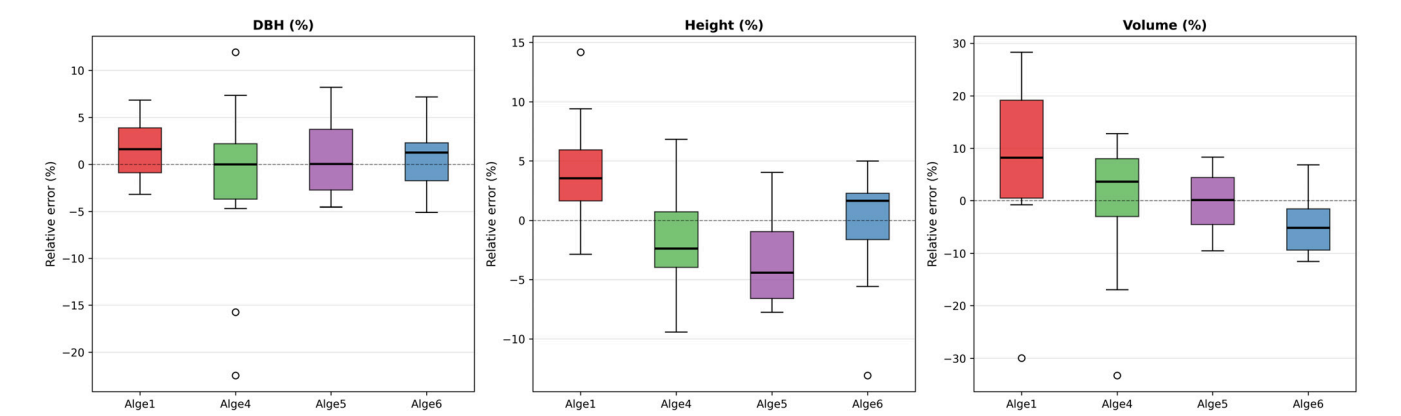


Figure 9. Distribution of the per-tree relative error by plot, for DBH, height and volume, on the 45 retained *P. pinaster* trees. Box edges show the 25th and 75th percentiles; whiskers extend to $1.5 \times \text{IQR}$; the central line is the median. Plot-level volume MAPE under LOPO is reported in Table 9.

Table 9. Per-plot volume error statistics on the four retained plots.

Plot	<i>n</i>	In-Sample MAPE	In-Sample Bias	LOPO MAPE	LOPO Bias
Alge1	12	13.44%	+8.32%	15.7%	+11.0%
Alge4	15	8.71%	−0.20%	8.6%	−0.7%
Alge5	4	5.94%	−0.23%	5.9%	−0.3%
Alge6	14	6.39%	−4.37%	7.6%	−6.4%
Aggregate	45	9.00%	+0.77%	9.9%	+0.7%

3.3.3. Trees with Absolute Volume Deviation Above 25%

Four trees in the retained sample exhibit an absolute volume deviation above 25% after the full pipeline and the global calibration. All four are small or structurally atypical stems, and all four are flagged as medium quality by the per-tree diagnostic; their individual figures and dominant error mechanisms are reported in Table 10. The per-tree distribution of signed volume deviation is plotted in Figure 10, where the four outliers are visible at the extremes of the distribution.

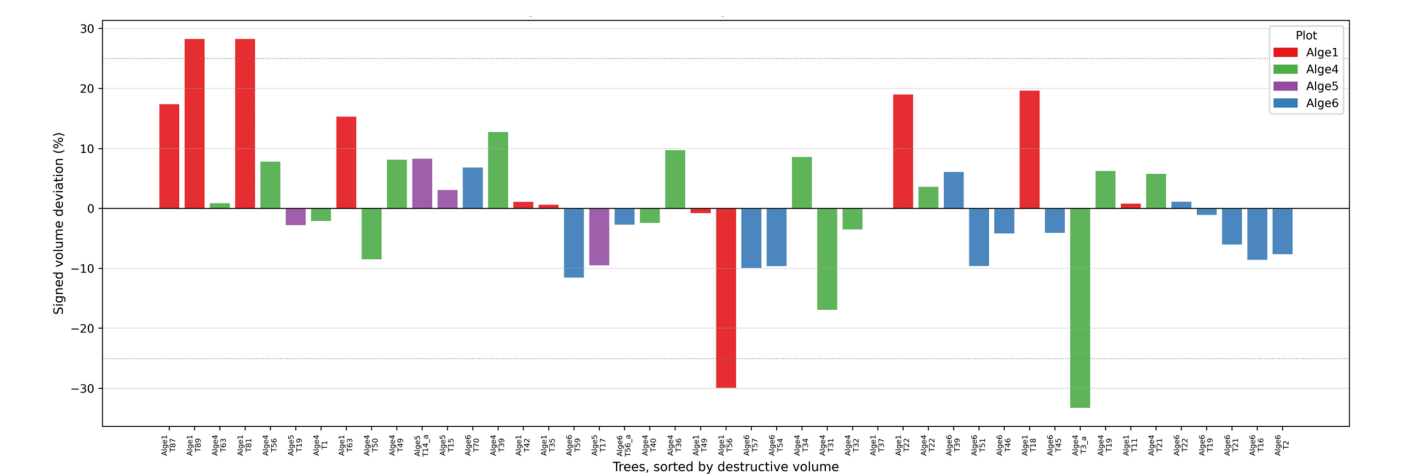


Figure 10. Per-tree signed volume deviation (%) against destructive reference, for the 45 retained *P. pinaster* trees, sorted along the *x*-axis by destructive volume. The four outliers identified in Section 3.3.3 (Alge1 T56, Alge1 T81, Alge1 T89, Alge4 T3_a) are visible at the extremes of the distribution.

Table 10. Trees with an absolute volume deviation above 25% in the 45-tree retained sample. All four were reconstructed with the species shape and flagged as medium quality by the per-tree diagnostic.

Tree	Destructive Volume (m ³)	Pipeline Volume (m ³)	Deviation	Extrapolated Fraction	Dominant Error Mechanism
Alge1 T56 (GT#1)	0.365	0.256	−29.9%	0.786	Second-highest extrapolated fraction in the sample: 79% of the stem is reconstructed from the species shape rather than measured directly (sample median $\approx$ 30%)
Alge1 T81 (GT#2)	0.121	0.155	+28.3%	0.433	Small-stem regime (DBH 14.6 cm); the small absolute volume amplifies the relative deviation
Alge1 T89 (GT#4)	0.081	0.104	+28.3%	0.234	Smallest reference DBH of the outlier set (12.9 cm); height over-estimated by 14.2%, which propagates to the volume
Alge4 T3_a (GT#35)	0.567	0.379	−33.3%	0.416	DBH under-estimated (reference 30.7 cm, pipeline 23.8 cm, −22.5%); the DBH ensemble median converged on an under-estimate

3.3.4. Sensitivity Analysis: The Full 59-Tree Sample

The Alge2 plot was set aside from the calibration sample for the reasons given in Section 4.5. For transparency, the same aggregated statistics were recomputed on the full destructive set of 59 trees. Volume MAPE rises from 9.00% to 12.55%, median absolute percentage error from 7.65% to 8.18%, signed bias from +0.77% to +6.22% and relative RMSE from 12.22% to 19.96%; Lin’s concordance correlation coefficient falls from 0.965 to 0.946 and DBH MAPE rises from 3.77% to 5.07%. Height MAPE is essentially unchanged, at 4.13% against 4.04%. Taken alone, the 14 trees of Alge2 have a volume MAPE of 23.94% with a mean bias of +23.72%, a DBH MAPE of 9.25% with a bias of +5.85%, and a height MAPE of 4.42% with a bias of +0.34%. The degradation is therefore confined to the two channels that derive from stem diameter and leaves the length channel untouched, which is the signature of the reference-side problem described in Section 4.5 rather than of a pipeline failure on this plot.

3.4. Comparison Against Baseline Methods

The proposed pipeline was compared head-to-head against the FJD Trion Model and a manual CloudCompare workflow on the 40 retained destructively measured trees for which both baseline outputs are available (5 Alge6 trees absent from the manufacturer export; cf. Section 2.7). Aggregated agreement statistics are reported in Table 11, and per-plot volume statistics in Table 12. Where the pipeline statistic in Table 11 differs slightly from the 45-tree pipeline figures of Section 3.3 (e.g., 9.26% vs. 9.00% on volume MAPE), the difference is the sample composition; both versions are reported for transparency.

Three observations follow directly from Table 11. First, the FJD Trion Model leads on aggregated MAPE across the three variables, with its largest lead in the height channel (4.25% against 2.21%), which Section 4.3 traces to a data-side rather than algorithmic cause. Second, the proposed pipeline is essentially tied with the manual CloudCompare workflow on aggregated volume MAPE (9.26% against 9.12%, a 0.14 pp gap well within the noise of the destructive reference itself) and is within 0.2 pp of the proprietary software on

aggregated DBH MAPE. Third, the proposed pipeline delivers the lowest DBH MedAPE of the three methods on the median tree (2.74% against 3.59% and 4.92% on the matched 40-tree sample), which is the central statistic for the typical tree and indicates that the three-estimator DBH ensemble of Section 2.4.4 is operationally robust. We note that the comparison on signed bias is not directly informative: the pipeline applies an empirical bias-correction step (Equations (4) and (5)) calibrated on the destructive reference, while the baselines do not, so the pipeline's near-zero bias is in part a property of its calibration design.

Table 11. Head-to-head aggregated agreement against destructive reference on the 40 trees for which both baseline outputs are available (5 Alge6 trees of project identifiers 55–59 absent from the manufacturer export; cf. Section 2.7). Pipeline statistics in this table are recomputed on the same $n = 40$ sub-sample for comparability; the standalone $n = 45$ figures of Sections 3.2 and 3.3 are reproduced in the rightmost column for reference.

Variable	Method	MAPE	MedAPE	Bias	(Pipeline $n = 45$)
Volume	Proposed pipeline	9.26%	7.86%	+1.40%	9.00%/7.65%/+0.77%
	FJD Trion Model	6.73%	6.23%	−3.97%	-
	CloudCompare manual	9.12%	8.26%	−6.60%	-
Height	Proposed pipeline	4.25%	3.63%	+0.11%	4.04%/3.27%/+0.13%
	FJD Trion Model	2.21%	1.70%	+0.66%	-
	CloudCompare manual	2.94%	2.12%	−0.40%	-
DBH	Proposed pipeline	3.88%	2.74%	+0.15%	3.77%/2.46%/+0.13%
	FJD Trion Model	3.69%	3.59%	−3.40%	-
	CloudCompare manual	4.51%	4.92%	−4.20%	-

Table 12. Per-plot volume MAPE for the three methods on the four retained plots.

Plot	n	Proposed Pipeline	FJD Trion	CloudCompare	Best
Alge1	12	13.44%	8.10%	11.90%	Trion
Alge4	15	8.71%	6.80%	9.30%	Trion (pipeline > CC)
Alge5	4	5.94%	6.70%	9.60%	Proposed pipeline
Alge6	14	6.39%	5.50%	7.20%	Trion (pipeline > CC)

A paired Wilcoxon signed-rank test on the per-tree absolute volume errors of the matched sample finds no statistically significant difference between the proposed pipeline and either baseline (pipeline vs. FJD Trion Model, $p = 0.44$; pipeline vs. manual CloudCompare workflow, $p = 0.56$). The apparent ordering of the aggregated MAPE values is therefore not resolved by the present sample size, and the three methods should be read as statistically indistinguishable on volume rather than as ranked.

Per-plot volume agreement is reported in Table 12.

The proposed pipeline leads on Alge5 (5.94% against 6.70% and 9.60%) and outperforms the manual CloudCompare workflow on three of the four retained plots (Alge4, Alge5, Alge6); the proprietary software leads on the remaining three plots (Alge1, Alge4, Alge6) with the largest lead in Alge1 (8.10% against 13.44%), the plot with the highest concentration of small stems and high-extrapolation outliers.

3.5. Pipeline Output Stratification

Three stratifications of the 45-tree sample reveal the structure of the residual error distribution: by taper parameterization, by DBH class and by extrapolated fraction.

By taper parameterization. All 45 retained trees in the calibrated run use the species shape. The Kozak and power-law alternatives, retained as configuration options for species or regions without a calibrated polynomial, were not selected in any tree of the *P. pinaster* sample.

By DBH class and by extrapolated fraction. Per-tree residuals against DBH and against extrapolated fraction are visualized in Figures 11 and 12. Two patterns are apparent: residuals are dispersed approximately symmetrically about zero across the DBH range, with a slightly higher dispersion for DBH below 16 cm (the small-stem regime of the earlier per-class calibration), driven principally by Alge1 T81 and T89; the absolute volume error increases monotonically with the extrapolated fraction: 4.9% for the six trees below 0.2, 7.7% for the 27 trees between 0.2 and 0.4, 12.5% for the 10 trees between 0.4 and 0.6 and 21.4% for the two trees above 0.6 (Spearman $\rho = +0.33$, Pearson $r = +0.37$, $n = 45$). The relationship is real but weak at the individual-tree level: the tree with the highest extrapolated fraction in the sample (Alge4 T39, 0.918) is estimated to within +12.8%, while the largest under-estimation (Alge1 T56, 0.786) is −29.9%. Extrapolated fraction therefore behaves as a risk indicator rather than as a predictor of error, which is how it is used in the per-tree uncertainty value of Section 2.4.5.

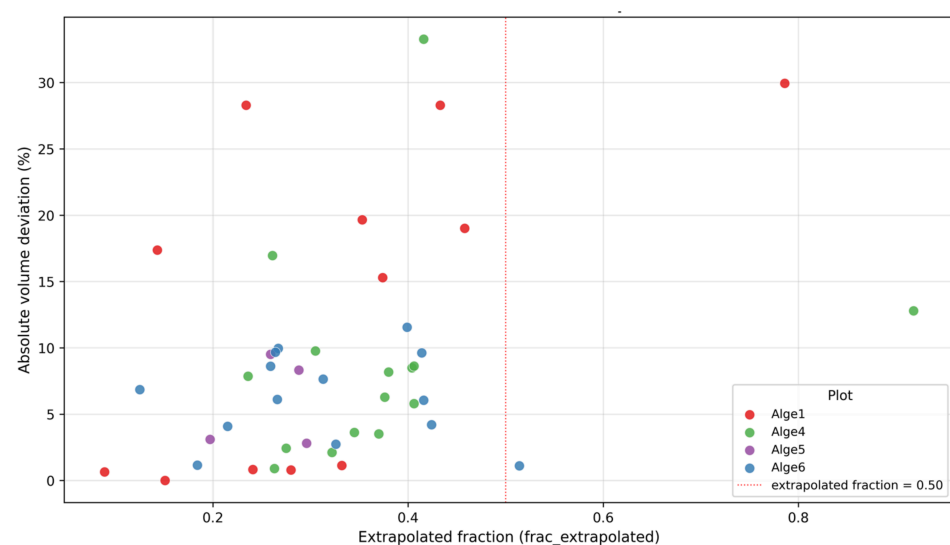


Figure 11. Per-tree absolute volume deviation (%) against the per-tree extrapolated fraction (frac_extrapolated), defined as the fraction of the tree height above the highest RANSAC-reliable slice. Each point is one tree of the 45-tree retained sample; color indicates plot.

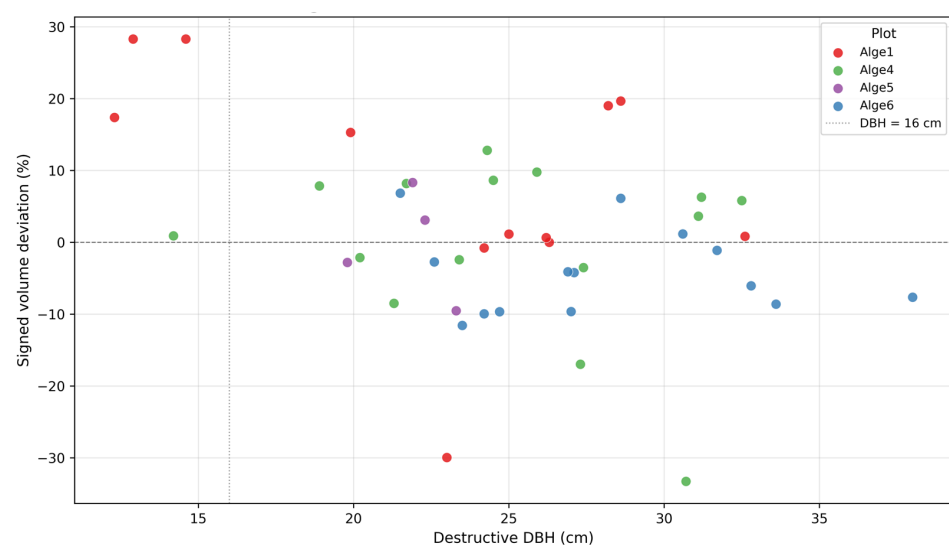


Figure 12. Per-tree signed volume deviation (%) against destructive DBH (cm), for the 45 retained *P. pinaster* trees. Color indicates plot. The dispersion of small-DBH residuals (below 16 cm) is visibly larger than that of large-DBH residuals, reflecting the dynamic-range effect discussed in Section 4.3.

3.6. Processing Times

Phase-by-phase wall time for the run reported here is summarized in Table 13. Tree IDs are reassigned at each run by TLS2trees and matching across runs is performed by horizontal coordinate proximity, not by ID.

Table 13. Per-plot processing time and detected-instance count for the run reported here.

Plot	Cloud Size	Detected	nGT	Wall Time
Alge1	1.5 GB	90	12	139.1 min
Alge2	0.3 GB	49	14	16.7 min
Alge4	1.1 GB	49	15	90.5 min
Alge5	0.3 GB	18	4	16.6 min
Alge6	1.4 GB	86	14	121.9 min
Total	4.6 GB	292	59	384.8 min $\approx$ 6.4 h

4. Discussion

4.1. Volume Agreement in Context

The proposed pipeline achieved a mean absolute percentage error (MAPE) of 9.00% for tree volume estimation on 45 destructively measured *Pinus pinaster* trees, with low bias (+0.77%) and strong agreement with the reference measurements (Lin's CCC = 0.965). Compared with the literature, these results are competitive with or better than previously reported HMLS studies: they outperform the only published HMLS benchmark validated against destructive measurements [38], which reported a merchantable-wood volume RMSE of 19% on 30 hardwoods, and they fall at the lower end of the 10–20% error range typically reported for TLS in dense forest. The relative RMSE of 12.22% obtained here is also consistent with the multi-system SLAM benchmark of Muhojoki et al. [40], which reports handheld-SLAM stem-volume RMSE of the order of 10% under easy conditions and of the order of 20% under difficult conditions. Importantly, the achieved accuracy is below the uncertainty commonly associated with allometric reference models [1], supporting the potential of HMLS as a practical alternative for operational forest inventories.

Comparison with the operational baselines showed that the proprietary FJD Trion S1 software achieved the lowest overall volume error, while the proposed pipeline performed similarly to the manual CloudCompare workflow and delivered the best median DBH accuracy. The purpose of this comparison is not to position the pipeline as a direct competitor to commercial software, which is optimized for a different set of constraints and whose processing chain is not open to inspection, but to establish that an auditable open-source chain reaches the same accuracy band on DBH, matches the manual workflow on volume, and does so while reporting a per-tree quality flag and uncertainty value that neither baseline provides. High concordance and correlation coefficients demonstrate that the pipeline preserves the relative ranking of tree volumes, an important property for stand-level decision making. Error analysis further indicated that the residual inaccuracies were primarily associated with a small number of challenging trees, particularly small stems, rather than with a systematic plot-level bias, and that the built-in quality diagnostic successfully identified the cases requiring manual review.

4.2. Design Choices and Methodological Contributions

The proposed pipeline incorporates five methodological choices that collectively improve the robustness and accuracy of tree reconstruction and volume estimation from HMLS data. For clarity about what is adopted and what is new: the Cloth Simulation Filter [20], the TLS2trees instance segmentation [23], the RANSAC circle-fitting engine [25] and the Kozak taper family [28,29] are existing methods used here without modification. What

this work proposes is the species-specific one-parameter shape, the sequential RANSAC prior between consecutive slices, the taper-driven synthetic-fill, the three-estimator median DBH ensemble, the CSF-anchored base height and the form-factor clamp, together with their integration into a single reproducible chain with per-tree diagnostics.

These are: (1) a CSF-based, plot-consistent ground reference for estimating tree base height, which reduces the error propagated to DBH and volume estimation (isolated effect, measured by ablation: volume MAPE $10.7 \rightarrow 9.9\%$ and DBH MAPE $5.3 \rightarrow 4.0\%$ under LOPO); (2) a species-specific taper model, which replaces per-tree taper fitting with a single calibrated shape and a robust median-based scaling, reducing model instability; (3) a three-estimator DBH ensemble that combines independent diameter estimates and reports their median, making DBH estimation resilient to the failure of any individual estimator; (4) a sequential RANSAC framework with taper-guided synthetic point completion, which improves circle fitting and compensates for incomplete scans; and (5) an operational quality layer comprising geometric circle refitting, a biological plausibility check through the form-factor constraint, and per-tree quality and uncertainty diagnostics.

Rather than relying on a single algorithmic improvement, the pipeline follows a layered design in which each component addresses a different source of error, from ground reference estimation through stem reconstruction to quality control. Together, these choices reduce bias, improve the stability of the DBH and volume estimates, and provide the diagnostic information that supports practical deployment in operational forest inventories. The species shape is, however, empirical: it requires a destructive calibration sample for each new species or region, and the polynomial reported here is the *P. pinaster*/Lousã instance.

Contribution of the Individual Components

The pipeline was developed as a sequence of revisions, each motivated by a failure mode observed on the destructive reference, so the contribution of its main components can be read from the comparison between the present design and earlier versions on the same trees. The CSF-anchored base height removes a per-tree offset that propagates to every slice: on Alge6 T73, the height and volume deviations move from -14.7% and -18.1% to $+2.3\%$ and $+8.4\%$. The one-parameter species shape removes regime-dependent model selection: Alge4 T49, at -37.0% under the free Kozak fit, falls within range under the species shape, as does Alge4 T27 at -22.7% . The cumulative effect is visible in the calibration itself, which earlier required a step function with a per-model split and now requires a single scalar per variable (Section 2.5). These are observational comparisons across successive versions of the pipeline rather than a controlled ablation; a systematic leave-one-component-out experiment is identified as future work, and the tagged release is publicly archived so that it can be carried out independently.

4.3. Failure Modes and the Data-Side Ceiling on Accuracy

The residual error of the pipeline against the 45-tree destructive reference is concentrated in a small number of trees whose mechanisms can be traced individually to either algorithmic or data-side causes.

Algorithmic failure modes resolved by the present design. Three trees that were outliers in earlier versions of this pipeline family, Alge4 T49 (-37.0% volume), Alge4 T27 (-22.7%) and Alge6 T73 (-18.1% volume, -14.7% height), fall below the 25% threshold in the present evaluation. The first two were previously reconstructed with the Kozak parameterization under low-coverage extrapolation and are recovered by the species shape with median-of-ratios scaling; the third is recovered by the CSF-anchored base height, its earlier height under-estimation having been a z_{base} offset artifact rather than a limit on apex recovery. The mechanisms are itemized in Table 14.

Table 14. Summary of the failure modes identified in the 45-tree retained sample, with their probable cause, supporting evidence and operational implication.

Failure Mode	Probable Cause	Evidence	Implication
Regime-dependent taper failure (resolved in the present design)	Free-parameter Kozak fit diverging under low slice coverage	Alge4 T49 (−37.0% in an earlier version) and Alge4 T27 (−22.7%) fall within range under the species shape	The one-parameter species shape removes model selection as a source of per-tree variance
Base-height offset (resolved in the present design)	z_base estimated from the per-tree cloud minima, biased upward by undergrowth and root collars	Alge6 T73 moves from −14.7% height and −18.1% volume to +2.3% and +8.4% under the CSF anchor	A plot-consistent ground reference removes a per-tree offset that propagates to every slice
High extrapolated fraction (contained, not eliminated)	Stem geometry above the highest reliable slice was never measured	Alge1 T56: 78.6% extrapolated, −29.9% volume (the highest extrapolated fraction, Alge4 T39 at 91.8%, is estimated to within +12.8%)	Flagged by the per-tree uncertainty value; only better scan coverage can remove it
DBH ensemble under-estimate (contained, not eliminated)	All three estimators biased low on a structurally atypical stem	Alge4 T3_a: DBH 30.7 → 23.8 cm, volume −33.3%	Flagged as medium quality; requires manual review in operational deployment
Small-stem dynamic range (contained, not eliminated)	A small absolute residual maps to a large relative deviation	Alge1 T81 (+28.3%) and Alge1 T89 (+28.3%), DBH 14.6 and 12.9 cm	Relative error is not the appropriate metric below approximately 15 cm DBH
Canopy apex absent from the input cloud (data-side ceiling)	No line of sight to the canopy top from any waypoint of the walking trajectory	Alge5 T17, Alge5 T15 and Alge6 T22 have no apex returns in the raw SLAM cloud	Not addressable by any algorithm; requires multi-position scanning or a complementary UAV pass

Algorithmic failure modes are contained but not eliminated by the present design. Four trees in the retained sample exhibit an absolute volume deviation above 25%: Alge1 T56 (−29.9%), Alge1 T81 (+28.3%), Alge1 T89 (+28.3%) and Alge4 T3_a (−33.3%); their individual figures are reported in Table 10. T56 and T3_a are the two cases where these elements contained the residual but did not eliminate it: T56 because 78.6% of its stem is reconstructed by extrapolation, and no algorithmic correction can recover geometry that was never measured; T3_a because the median of the three DBH estimators converged on an under-estimate for this specific tree. T81 and T89 are small-stem dynamic-range effects, where a small absolute residual maps to a large relative deviation. All four trees are flagged by the per-tree quality diagnostic as medium quality, which surfaces them for inspection in any operational deployment.

Data-side ceiling: three trees whose apex does not exist in the input cloud. Part of the residual height MAPE of 4.04% across the 45-tree sample is not algorithmic but data-side: excluding the three trees whose apex is absent from the input cloud lowers it to 3.68%. Three trees in the sample exhibit persistent height under-estimation that no element of the pipeline corrects: Alge5 T17 (GT#42, dH −6.23%), Alge5 T15 (GT#44, dH −7.77%), and Alge6 T22 (GT#51, dH −13.09%). We inspected the original SLAM cloud (the pre-Phase-1 file, in the project's Baldios_Alge_clip_Pedro distribution) at the coordinates of each of these trees using a separate diagnostic script. The result is conclusive: the canopy apex of

each of these three trees is not present in the input cloud, neither before nor after Phase 1 processing. Phase 1 removes approximately 10% of points across the cloud and cannot be the cause; the apex was never captured by the SLAM acquisition at the present plot. The three trees stood in particularly tall and dense neighborhoods within their plots, consistent with the operator's walking trajectory not having a line of sight to the canopy top from any waypoint within useful range. Figure 2 shows the mechanism directly: seen from above, the return density over the plot decays sharply with height, so that the upper canopy is only sparsely sampled.

The implication is direct: a pipeline of any sophistication, processing the present clouds, would be unable to recover the apex of these three trees because the apex returns are not in the data. The FJD Trion S1 manufacturer software reports a height MAPE of 2.21% on the same 45 trees; that figure is, in our reading, attainable only if the manufacturer software is processing the full SLAM acquisition, rather than the clip we received. We cannot directly verify this hypothesis without access to the manufacturer software's input pipeline; we present it as a defensible interpretation of the empirical pattern and mark it as such.

The data-side ceiling on volume MAPE is connected to, but not strictly equal to, the data-side ceiling on height MAPE. With $V \propto r^2 \cdot H$, the apex absence imposes a downward bias on height on the three affected trees; whether the volume residual follows depends on whether the DBH residual on the same trees is also near zero. In the present sample, the three apex-absent trees split: Alge5 T17 (height -6.2% , volume -9.5%) shows the expected co-directional propagation; Alge5 T15 (height -7.8% , volume $+3.1\%$) and Alge6 T22 (height -13.1% , volume $+1.2\%$) show the height under-estimation partially or wholly compensated by a positive DBH residual on the same tree, so that the volume residual is small or reversed. The data-side ceiling therefore acts firmly on the height channel; its propagation to the volume channel is mediated by the DBH and is not uniform across the three trees. Multi-position scanning or a complementary drone-based above-canopy pass would close the gap and is the natural extension of the protocol for users who require sub-3% height MAPE in similar stand conditions.

4.4. The Operational Value of the Automatic Stand-Level Map

Beyond the accuracy discussion, the pipeline produces a per-plot map that translates the point cloud into a form directly usable by forest managers: tree positions, heights encoded by color, volumes encoded by marker size, and plot-level totals in a side panel (Figure 13). This artifact converts a processing pipeline into an inventory tool: the user does not need to open a point-cloud viewer or interact with a CSV to understand the stand. The map is particularly aligned with the sustainability orientation of operational forest inventory in Mediterranean stands: a walking operator with an entry-level HMLS can document a stand in the time of a conventional inventory protocol, with a complete per-tree volume record and a spatial representation that can be archived, audited, and compared across years.

The per-tree quality classification (high/medium/low) and the per-tree uncertainty value (linear in the extrapolated fraction) are written into the per-tree CSV and would normally be exposed to a forest manager using the pipeline output. The framework is not a statistical confidence interval and we do not present it as such; it is a per-tree diagnostic that captures the conditions under which the pipeline produced the estimate. The four outlier trees in Section 3.3.3 are all flagged medium or low by the quality classifier, i.e., the diagnostic and the destructive evidence agree on which trees are problematic.

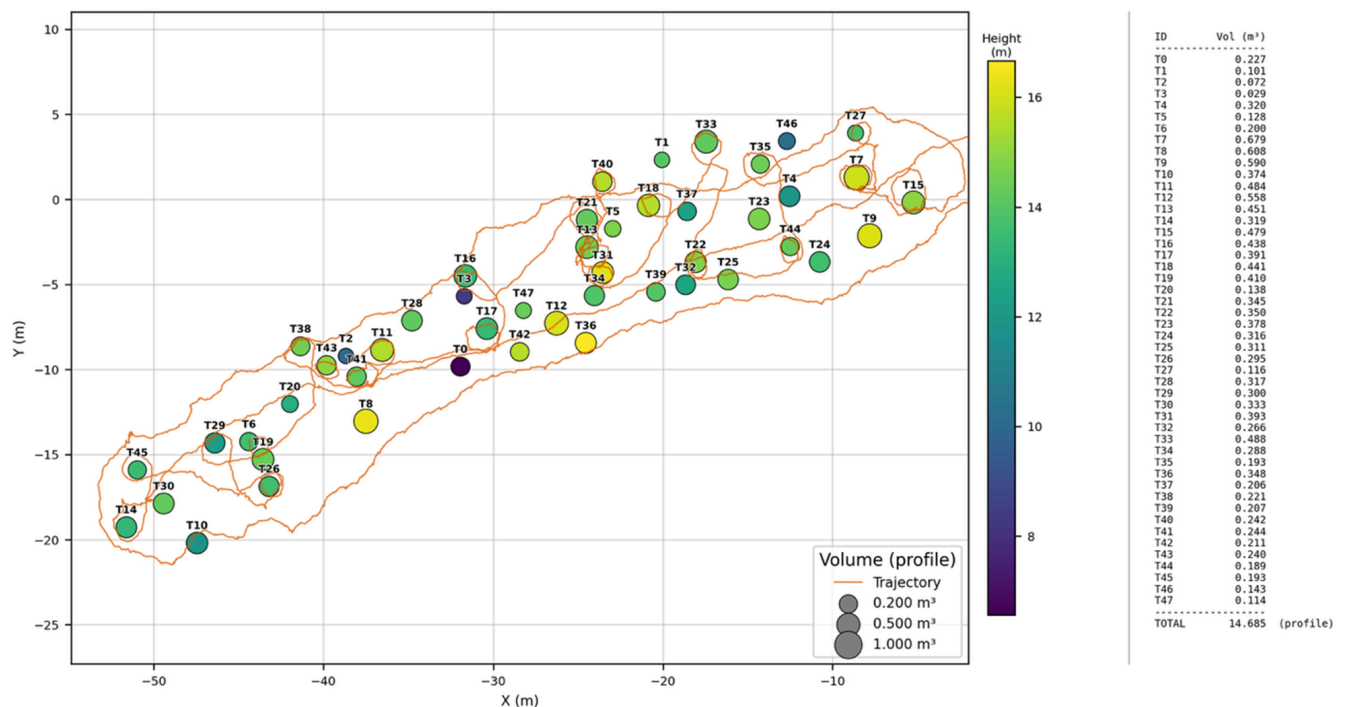


Figure 13. Automatic stand-level map produced by the pipeline for one of the Alge plots: tree positions in the plot frame, total height encoded by color, stem volume encoded by marker size, and plot-level totals in the side panel.

4.5. Limitations

The main limitations of the proposed model are as follows.

Limited generalizability. The study was conducted on a single species (*Pinus pinaster*) in one region of Portugal. Although the pipeline could in principle be adapted to other species, doing so will require new destructive calibration and validation before its performance can be generalized.

Risk of overfitting to the calibration set. Several components were calibrated on the same destructive dataset used for evaluation, and the leave-one-plot-out protocol mitigates but does not remove this dependency. The four scalar constants and the form-factor bounds are re-estimated inside each training fold and are therefore genuinely out-of-sample; the species-shape polynomial of Equation (2) is fitted once on the caliper observations of the full destructive set, including the Alge2 plot, and is not re-fitted per fold. The reported LOPO figures are therefore optimistic with respect to a fully independent dataset. The four plots also lie within a single stand block and share site, silvicultural history and size distribution, so inter-fold variability understates what a new region would introduce. Independent validation on stands outside this block is the necessary next step.

Exclusion of the Alge2 plot. The 14 trees of Alge2 were not felled but pulled out with a tractor and grapple. The grapple crushes and tears the trunk at the gripping points and material is lost along the stem, so the caliper sections measured after extraction no longer describe the intact tree that the scanner had recorded. The reference for this plot is therefore systematically smaller than the tree the pipeline measured, and the error pattern is what that mechanism predicts. The two channels derived from stem diameter are strongly and one-directionally affected: volume is over-estimated on 13 of the 14 trees, with a mean bias of +23.7%, and DBH on 11 of 14, with a mean bias of +5.8%. The channel that does not depend on diameter is unaffected: height bias is +0.3%, in line with the retained plots, where the volume and DBH biases are +0.8% and +0.1%. A pipeline defect would not produce a bias confined to the diameter-derived channels of one plot while leaving its

height estimates intact; a degraded reference does. The field records of this plot also contain confirmed inconsistencies, including a coordinate swap between the trees of identifiers 20 and 21. The full-sample results are reported in Section 3.3.4 so that the reader can judge the effect directly.

Baseline comparison limitations. The comparison against the FJD Trion Model must be read with caution for three reasons we cannot resolve from outside the software: its internal processing chain is undocumented; it is not known whether it operates on the same clipped acquisition supplied to the proposed pipeline or on the full SLAM record, and the height results of Section 4.3 are more consistent with the latter; and its export omitted five of the retained trees without stating a reason. The manual CloudCompare workflow was executed by a single operator, so operator variability is confounded with method. The figures of Table 11 should therefore be read as an indication of the accuracy band each method occupies under these conditions, rather than as a definitive ranking.

Uncertainty in the reference measurements. Although destructive measurements represent the best available ground truth, they are not error-free. Measurement inaccuracies, the discretization implicit in Smalian-type integration, and the handling of the stems during extraction and sectioning introduce uncertainty into the reference volumes used for validation. The Alge2 case discussed above is an extreme instance of this: when the extraction method itself damages the stem, the destructive measurement stops being the best available reference and becomes the dominant source of apparent error.

Absence of a TLS reference tier. The design compares HMLS output directly against destructive measurements, which confounds the intrinsic accuracy of the extraction algorithms with the error introduced by the quality of the HMLS point cloud. A synchronous TLS acquisition of the same stems before felling would have separated the two, and would also have allowed the circular cross-section assumption of the caliper protocol to be quantified. Such an acquisition was outside the envelope of this campaign and is recommended for future destructive studies.

Restricted acquisition conditions. Data were collected under favorable weather and forest conditions with relatively low understory density. The performance of the pipeline under more challenging conditions, such as dense vegetation, heavy rain, or strong wind, remains untested, and incomplete canopy data limited the height estimation accuracy for some trees.

4.6. Future Work

Future work will extend the destructive validation to additional tree species, including *Quercus suber*, *Pinus sylvestris* and *Pinus nigra*, enabling species-specific taper calibration and error analysis. The pipeline will also be tested using full SLAM acquisitions rather than clipped point clouds, in order to improve height estimation. Integration with UAV-LiDAR, for instance, through a SLAM-plus-UAV registration such as the RANSAC-ICP approach of Zhang et al. [41], is also proposed to recover the missing canopy information and further increase height accuracy. The current per-tree uncertainty indicator will be calibrated against the observed errors, so as to provide statistically meaningful confidence intervals instead of operational guidance values. Finally, the species-specific taper model will be extended by incorporating bark-thickness corrections, calibrated individually for each species against destructive reference measurements.

5. Conclusions

This paper presents an open-source handheld mobile laser scanning (HMLS) pipeline for tree detection, stem reconstruction, and volume estimation in *Pinus pinaster* stands, validated against the largest destructively measured dataset currently reported for this

species. The pipeline achieved 9.9% volume MAPE, 4.4% height MAPE, and 3.9% DBH MAPE under leave-one-plot-out cross-validation, demonstrating performance competitive with a manual CloudCompare workflow and approaching the accuracy of the proprietary FJD Trion software, while providing the best median DBH accuracy of the three methods. Its main methodological contributions are a CSF-based ground reference, a species-specific taper model, a multi-estimator DBH ensemble, sequential RANSAC-based stem reconstruction, and integrated quality and uncertainty diagnostics.

Future improvements are expected to come primarily from enhanced data acquisition rather than from algorithmic changes. The processing code is released under an open-source license and the pipeline produces operational inventory outputs with per-tree uncertainty estimates. The results reported here are specific to naturally regenerated *Pinus pinaster* in a single stand block of central Portugal: the processing chain is species-independent, but the species-shape polynomial and the four calibration constants are not, and transferring the pipeline to another species or region requires a new destructive calibration before comparable accuracy can be claimed.

Author Contributions: Conceptualization, P.A., I.B., R.S.-G., B.F. and M.M.; methodology, P.A., R.S.-G. and I.B.; software, P.A.; validation, P.A. and I.B.; formal analysis, P.A., I.B., R.S.-G., B.F. and M.M.; investigation, P.A., I.B., R.S.-G., B.F. and M.M.; resources, I.B., R.S.-G., M.M. and B.F.; data curation, P.A. and I.B.; writing (original draft preparation), P.A.; writing (review and editing), P.A., I.B., M.M., R.S.-G. and B.F.; visualization, P.A. and I.B.; supervision, M.M., R.S.-G. and B.F.; project administration, R.S.-G. and B.F.; funding acquisition, R.S.-G. and B.F. All authors have read and agreed to the published version of the manuscript.

Funding: This work is financed by the European Union (NextGenerationEU) through the project no. 34TRANSFORM, Digital Transformation of the Forestry Sector for a More Resilient and Low-Carbon Economy. The project is co-financed by Centro 2020, Portugal 2020, and the European Union through the European Social Fund (ESF).

Data Availability Statement: The pipeline source code is publicly available on GitHub at <https://github.com/Apostolo6/ApostoloLDS> (last accessed on 14 September 2026), and the tagged release used in this study is archived at Zenodo under DOI 10.5281/zenodo.20717660 (<https://doi.org/10.5281/zenodo.20717660>)—(last accessed on 14 September 2026); the specific GitHub tag is <https://github.com/Apostolo6/ApostoloLDS/tree/v1.0.0> (last accessed on 14 September 2026). The destructive reference data and the per-tree pipeline outputs used in this study are available from the corresponding author upon reasonable request, subject to the project’s data-sharing agreement.

Acknowledgments: The authors would like to acknowledge the Government of Portugal and the European Union–NextGenerationEU for supporting this research through the PRR—Portugal’s Recovery and Resilience Plan, under project No. 34 TRANSFORM—Digital Transformation of the Forestry Sector for a More Resilient and Low-Carbon Economy. The authors also express their gratitude to the Polytechnic University of Coimbra (UPC) and the Escola Superior Agrária de Coimbra (ESAC) for providing the FJD Trion S1 scanner and the field datasets used in this study, and the RCM2+ (Research Center in Asset Management and Systems Engineering) for institutional support.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

HMLS	handheld mobile laser scanning
TLS	terrestrial laser scanning
ALS	airborne laser scanning

SLAM	simultaneous localization and mapping
PDAL	point data abstraction library
CSF	cloth simulation filter
DBH	diameter at breast height
HAG	height above ground
RANSAC	random sample consensus
MAPE	mean absolute percentage error
MedAPE	median absolute percentage error
RMSE	root-mean-square error
CCC	Lin's concordance correlation coefficient
LOPO	leave-one-plot-out (cross-validation)
MAD	median absolute deviation
GT	ground truth
GIS	geographic information system
QSM	quantitative structure model
FSCT	Forest Structural Complexity Tool
IPC	Polytechnic University of Coimbra
ESAC	Escola Superior Agrária de Coimbra
ESF	European Social Fund

References

1. Zianis, D.; Muukkonen, P.; Mäkipää, R.; Mencuccini, M. Biomass and Stem Volume Equations for Tree Species in Europe. In *Silva Fennica Monographs 4*; Finnish Society of Forest Science and Finnish Forest Research Institute: Helsinki, Finland, 2005; p. 63. Available online: <https://silvafennica.fi/show/mono/mono4> (accessed on 14 September 2026).
2. Instituto da Conservação da Natureza e das Florestas. 6.^o Inventário Florestal Nacional: Relatório Final; ICNF: Lisboa, Portugal, 2019.
3. Ministerio para la Transición Ecológica y el Reto Demográfico. *Anuario de Estadística Forestal 2019*; Gobierno de España: Madrid, Spain, 2021. Available online: <https://www.miteco.gob.es> (accessed on 14 September 2026).
4. Liang, X.; Kankare, V.; Hyypä, J.; Wang, Y.; Kukko, A.; Haggrén, H.; Yu, X.; Kaartinen, H.; Jaakkola, A.; Guan, F.; et al. Terrestrial laser scanning in forest inventories. *ISPRS J. Photogramm. Remote Sens.* **2016**, *115*, 63–77. [\[CrossRef\]](#)
5. Calders, K.; Newnham, G.; Burt, A.; Murphy, S.; Raunonen, P.; Herold, M.; Culvenor, D.; Avitabile, V.; Disney, M.; Armston, J.; et al. Nondestructive estimates of above-ground biomass using terrestrial laser scanning. *Methods Ecol. Evol.* **2015**, *6*, 198–208. [\[CrossRef\]](#)
6. Raunonen, P.; Kaasalainen, M.; Åkerblom, M.; Kaasalainen, S.; Kaartinen, H.; Vastaranta, M.; Holopainen, M.; Disney, M.; Lewis, P. Fast Automatic Precision Tree Models from Terrestrial Laser Scanner Data. *Remote Sens.* **2013**, *5*, 491–520. [\[CrossRef\]](#)
7. Bauwens, S.; Bartholomeus, H.; Calders, K.; Lejeune, P. Forest Inventory with Terrestrial LiDAR: A Comparison of Static and Hand-Held Mobile Laser Scanning. *Forests* **2016**, *7*, 127. [\[CrossRef\]](#)
8. Ryding, J.; Williams, E.; Smith, M.J.; Eichhorn, M.P. Assessing Handheld Mobile Laser Scanners for Forest Surveys. *Remote Sens.* **2015**, *7*, 1095–1111. [\[CrossRef\]](#)
9. Oveland, I.; Hauglin, M.; Gobakken, T.; Næsset, E.; Maalen-Johansen, I. Automatic Estimation of Tree Position and Stem Diameter Using a Moving Terrestrial Laser Scanner. *Remote Sens.* **2017**, *9*, 350. [\[CrossRef\]](#)
10. Chen, S.; Liu, H.; Feng, Z.; Shen, C.; Chen, P. Applicability of personal laser scanning in forestry inventory. *PLoS ONE* **2019**, *14*, e0211392. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Hyypä, E.; Yu, X.; Kaartinen, H.; Hakala, T.; Kukko, A.; Vastaranta, M.; Hyypä, J. Comparison of Backpack, Handheld, Under-Canopy UAV, and Above-Canopy UAV Laser Scanning for Field Reference Data Collection in Boreal Forests. *Remote Sens.* **2020**, *12*, 3327. [\[CrossRef\]](#)
12. Vandendaele, B.; Martin-Ducup, O.; Fournier, R.A.; Pelletier, G.; Lejeune, P. Mobile Laser Scanning for Estimating Tree Structural Attributes in a Temperate Hardwood Forest. *Remote Sens.* **2022**, *14*, 4522. [\[CrossRef\]](#)
13. Tupinambá-Simões, F.; Pascual, A.; Guerra-Hernández, J.; Ordóñez, C.; de Conto, T.; Bravo, F. Assessing the Performance of a Handheld Laser Scanning System for Individual Tree Mapping, A Mixed Forests Showcase in Spain. *Remote Sens.* **2023**, *15*, 1169. [\[CrossRef\]](#)
14. Tupinambá-Simões, F.; Pascual, A.; Guerra-Hernández, J.; Ordóñez, C.; de Conto, T.; Bravo, F. Accuracy of tree mapping based on hand-held laser scanning comparing leaf-on and leaf-off conditions in mixed forests. *J. For. Res.* **2024**, *35*, 93. [\[CrossRef\]](#)

15. Giannetti, F.; Passarino, L.; Aleandri, G.; Borghi, C.; Vangi, E.; Anzilotti, S.; Raddi, S.; Chirici, G.; Travaglini, D.; Maltoni, A.; et al. Efficiency of Mobile Laser Scanning for Digital Martelosopes for Conifer Forests in the Mediterranean Region. *Forests* **2024**, *15*, 2202. [\[CrossRef\]](#)
16. Gollob, C.; Ritter, T.; Nothdurft, A. Forest Inventory with Long Range and High-Speed Personal Laser Scanning (PLS) and Simultaneous Localization and Mapping (SLAM) Technology. *Remote Sens.* **2020**, *12*, 1509. [\[CrossRef\]](#)
17. Gollob, C.; Ritter, T.; Kraßnitzer, R.; Tockner, A.; Nothdurft, A. Measurement of Forest Inventory Parameters with Apple iPad Pro and Integrated LiDAR Technology. *Remote Sens.* **2021**, *13*, 3129. [\[CrossRef\]](#)
18. Mokroš, M.; Mikita, T.; Singh, A.; Tomašík, J.; Chudá, J.; Wężyk, P.; Kuželka, K.; Surový, P.; Klimánek, M.; Zięba-Kulawik, K.; et al. Novel low-cost mobile mapping systems for forest inventories as terrestrial laser scanning alternatives. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *104*, 102512. [\[CrossRef\]](#)
19. Vatandaşlar, C.; Zeybek, M. Extraction of forest inventory parameters using handheld mobile laser scanning: A case study from Trabzon, Turkey. *Measurement* **2021**, *177*, 109328. [\[CrossRef\]](#)
20. Zhang, W.; Qi, J.; Wan, P.; Wang, H.; Xie, D.; Wang, X.; Yan, G. An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation. *Remote Sens.* **2016**, *8*, 501. [\[CrossRef\]](#)
21. Cai, S.; Zhang, W.; Qi, J.; Wan, P.; Shao, J.; Shen, A. Applicability Analysis of Cloth Simulation Filtering Algorithm for Mobile LiDAR Point Cloud. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII-3*, 107–111. [\[CrossRef\]](#)
22. Cabo, C.; Del Pozo, S.; Rodríguez-Gonzálvez, P.; Ordóñez, C.; González-Aguilera, D. Comparing Terrestrial Laser Scanning (TLS) and Wearable Laser Scanning (WLS) for Individual Tree Modeling at Plot Level. *Remote Sens.* **2018**, *10*, 540. [\[CrossRef\]](#)
23. Wilkes, P.; Disney, M.; Armston, J.; Bartholomeus, H.; Bentley, L.; Brede, B.; Burt, A.; Calders, K.; Chavana-Bryant, C.; Clewley, D.; et al. TLS2trees: A scalable tree segmentation pipeline for TLS data. *Methods Ecol. Evol.* **2023**, *14*, 3083–3099. [\[CrossRef\]](#)
24. Krisanski, S.; Taskhiri, M.S.; Gonzalez Aracil, S.; Herries, D.; Muneri, A.; Gurung, M.B.; Montgomery, J.; Turner, P. Forest Structural Complexity Tool, An Open Source, Fully-Automated Tool for Measuring Forest Point Clouds. *Remote Sens.* **2021**, *13*, 4677. [\[CrossRef\]](#)
25. Fischler, M.A.; Bolles, R.C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **1981**, *24*, 381–395. [\[CrossRef\]](#)
26. Liang, X.; Litkey, P.; Hyypä, J.; Kaartinen, H.; Vastaranta, M.; Holopainen, M. Automatic Stem Mapping Using Single-Scan Terrestrial Laser Scanning. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 661–670. [\[CrossRef\]](#)
27. Ye, W.; Qian, C.; Tang, J.; Liu, H.; Fan, X.; Liang, X.; Zhang, H. Improved 3D Stem Mapping Method and Elliptic Hypothesis-Based DBH Estimation from Terrestrial Laser Scanning Data. *Remote Sens.* **2020**, *12*, 352. [\[CrossRef\]](#)
28. Kozak, A. A variable-exponent taper equation. *Can. J. For. Res.* **1988**, *18*, 1363–1368. [\[CrossRef\]](#)
29. Kozak, A. My last words on taper equations. *For. Chron.* **2004**, *80*, 507–515. [\[CrossRef\]](#)
30. Max, T.A.; Burkhart, H.E. Segmented polynomial regression applied to taper equations. *For. Sci.* **1976**, *22*, 283–289. [\[CrossRef\]](#)
31. Pitkänen, T.P.; Raunonen, P.; Kangas, A. Measuring stem diameters with TLS in boreal forests by complementary fitting procedure. *ISPRS J. Photogramm. Remote Sens.* **2019**, *147*, 294–306. [\[CrossRef\]](#)
32. Saarinen, N.; Kankare, V.; Yrttimaa, T.; Viljanen, N.; Honkavaara, E.; Holopainen, M.; Hyypä, J.; Huuskonen, S.; Hynynen, J.; Vastaranta, M. Assessing the effects of thinning on stem growth allocation of individual Scots pine trees. *For. Ecol. Manag.* **2020**, *474*, 118344. [\[CrossRef\]](#)
33. Lee, D.; Kwak, D.-A.; Park, K.; Yun, J. Advancing Stem Volume Estimation Using Multi-Platform LiDAR and Taper Model Integration for Precision Forestry. *Remote Sens.* **2025**, *17*, 785. [\[CrossRef\]](#)
34. Boukhris, I.; Puletti, N.; Vonderach, C.; Guasti, M.; Lahssini, S.; Santini, M.; Valentini, R. Comparative analysis of taper models for *Pinus nigra* Arn. using terrestrial laser scanner acquired data. *iFor. Biogeosci. For.* **2024**, *17*, 203–212. [\[CrossRef\]](#)
35. Bornand, A.; Abegg, M.; Morsdorf, F.; Rehush, N. Completing 3D point clouds of individual trees using deep learning. *Methods Ecol. Evol.* **2024**, *15*, 1623–1638. [\[CrossRef\]](#)
36. Prendes, C.; Cabo, C.; Ordóñez, C.; Majada, J.; Canga, E. An algorithm for the automatic parametrization of wood volume equations from Terrestrial Laser Scanning point clouds: Application in *Pinus pinaster*. *GISci. Remote Sens.* **2021**, *58*, 1130–1150. [\[CrossRef\]](#)
37. Mura, M.; Mura, M.; Scotti, R. Evaluation of methods to improve the direct estimation of standing trees volume. *iFor. Biogeosci. For.* **2025**, *18*, 87–92. [\[CrossRef\]](#)
38. Vandendaele, B.; Martin-Ducup, O.; Fournier, R.A.; Pelletier, G. Evaluation of mobile laser scanning acquisition scenarios for automated wood volume estimation in a temperate hardwood forest using quantitative structural models. *Can. J. For. Res.* **2024**, *54*, 774–792. [\[CrossRef\]](#)
39. Butler, H.; Chambers, B.; Hartzell, P.; Glennie, C. PDAL: An open source library for the processing and analysis of point clouds. *Comput. Geosci.* **2021**, *148*, 104680. [\[CrossRef\]](#)

40. Muhojoki, J.; Tavi, D.; Hyyppä, E.; Lehtomäki, M.; Faitli, T.; Kaartinen, H.; Kukko, A.; Hakala, T.; Hyyppä, J. Benchmarking Under- and Above-Canopy Laser Scanning Solutions for Deriving Stem Curve and Volume in Easy and Difficult Boreal Forest Conditions. *Remote Sens.* **2024**, *16*, 1721. [[CrossRef](#)]
41. Zhang, S.; Wang, H.; Wang, C.; Wang, Y.; Wang, S.; Yang, Z. An Improved RANSAC-ICP Method for Registration of SLAM and UAV-LiDAR Point Cloud at Plot Scale. *Forests* **2024**, *15*, 893. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.