Article

Multi-Platform LiDAR Comparative Assessment for Above-Ground Biomass and Carbon Estimation in Mediterranean Woody Crops

Mateo Pastrana *, Cristina Velilla, Nelson Mattie, Alfonso Gomez, Sergio Molina

Posted Date: 4 March 2026

doi: 10.20944/preprints202603.0280.v1

Keywords: laser scanning; olive; almond; aboveground biomass; carbon accounting; machine learning

 doi:10.20944/preprints202603.0280.v1

Article

Multi-Platform LiDAR Comparative Assessment for Above-Ground Biomass and Carbon Estimation in Mediterranean Woody Crops

Mateo Pastrana [1,*], Cristina Velilla [1], Nelson Mattié [2], Alfonso Gomez [1,3] and Sergio Molina [4]

[1] School of Agricultural, Food and Biosystems Engineering, Universidad Politécnica de Madrid, Avda. Puerta de Hierro 2, 28040 Madrid, Spain

[2] Alberta Centre for Earth Observation Sciences, Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, AB, T6G 2E3, Canada

[3] Stereocarto Iberoamérica S.L., Carretera de Canillas 134, 28043. Madrid, Spain

[4] University of Jaén (UJA), Campus Las Lagunillas s/n, 23071 Jaén, Spain

* Correspondence: mateo.pastrana@alumnos.upm.es

Highlights

What are the main findings?

- Four LiDAR modalities (PNOA/ALS, Rielg ALS, ULS, and MLS) were benchmarked for plot-scale and tree-scale AGB and carbon estimations in Mediterranean woody crops.
- Species-specific allometry and fixed-area plot sampling enabled consistent biomass/carbon computations across the three farms in Córdoba (Spain).

What are the implications of the main findings?

- The National Aerial Orthophotography Plan (PNOA) can support scalable MRV workflows for orchard carbon monitoring under specific canopy/terrain conditions.
- Cross-platform harmonization and transparent validation are essential for defensible and reproducible estimates.

Abstract

Reliable aboveground biomass (AGB) estimates for woody crops are required for carbon accounting and MRV; however, it remains unclear how LiDAR modality and sampling geometry influence plot-scale and tree-scale AGB predictions in intensively managed orchards. We benchmarked four LiDAR modalities across three Mediterranean woody-crop sites in Córdoba (Spain), IFAPA, Doña María, and Villaseca using open national airborne laser scanning (PNOA/ALS), Riegl ALS, unmanned laser scanning (ULS), and mobile laser scanning (MLS). The field inventory used 58 fixed-area plots (20×50 m; 0.1 ha) collected in December 2024-January 2025 (1,867 trees) and species-specific allometries based on D2r to derive tree and plot AGB; carbon was computed using wood carbon fractions (0.445 olive; 0.457 almond) and CO2e via IPCC conversion. Plot-level LiDAR metrics (e.g., mean height, p95, maximum height, and cover proxies) were extracted from normalized point clouds and modeled with Random Forest, XGBoost, and an ensemble under an 80/20 train-test split. Mean field AGB differed among sites (33.89, 30.94 and 12.76 Mg ha−1 for Villaseca, Doña María, and IFAPA). In the provided summaries, XGBoost achieved the lowest errors at IFAPA (RMSE = 0.400 Mg ha−1; R2 = 0.994) and Villaseca (RMSE = 0.872 Mg ha−1; R2 = 0.995), whereas PNOA was competitive at Doña María (RMSE = 0.725 Mg ha−1; R2 = 0.994). The results support cross-platform LiDAR for orchard AGB mapping and identify conditions under which open national LiDAR can enable scalable MRV. In addition, we evaluated TreeQSM-based quantitative structure models (QSMs) as an independent tree-level 3D reconstruction approach and examined their site-dependent agreement with field inventory estimates.

 doi:10.20944/preprints202603.0280.v1

Keywords: laser scanning; olive; almond; aboveground biomass; carbon accounting; machine learning

1. Introduction

Woody-crop agroecosystems such as olive and almond orchards store carbon in long-lived woody biomass and are increasingly relevant for carbon accounting and measurement-reporting verification (MRV) in agriculture. Field inventories remain the benchmark for aboveground biomass (AGB) estimation, but repeated measurements are difficult to scale. LiDAR provides 3D structural measurements that enable robust plot-scale AGB modeling, often by relating plot metrics derived from height distributions and canopy cover proxies to field-derived AGB (area-based approaches) [1,2]. Orchards differ from tall forests in canopy architecture (low-to-moderate stature, row-based planting, management-driven structure), which can amplify scan-geometry effects, such as occlusion and non-uniform sampling. In Mediterranean olive systems, airborne LiDAR has demonstrated utility for structural estimation and biomass-related proxies [3,4], whereas orchard carbon assessments highlight the dependence of sequestration on planting intensity and management [5]. At broader scales, spaceborne LiDAR missions (e.g., GEDI) have reinforced the value of structural observations for biomass monitoring [6,7].

A key operational question for MRV is how much LiDAR detail is required: open national airborne LiDAR programs (e.g., Spain's PNOA-LiDAR) provide broad coverage at low densities, whereas dedicated ALS and ULS surveys and MLS provide denser sampling with distinct geometries and potential occlusion artifacts [8,9]. Cross-platform benchmarking of Mediterranean woody crops remains limited.

We benchmarked four LiDAR modalities (PNOA/ALS, Riegl ALS, ULS, and MLS) across three woody-crop sites in Córdoba (Spain) using 58 fixed-area plots (1,867 trees). We used species-specific allometries and machine-learning regression models (Random Forest and XGBoost) [10,11] to evaluate plot-scale AGB and carbon prediction and to identify conditions under which open national LiDAR can enable scalable orchard MRV.

Beyond area-based plot metrics, an increasingly important line of work leverages quantitative structural models (QSMs) to reconstruct tree geometries directly from 3D point clouds. QSM approaches, typically based on fitting cylinder primitives to woody skeletons, aim to estimate stem/branch volume and related structural attributes at the individual-tree level, enabling biomass inference with a more explicit geometric basis than summary height or density metrics. In principle, QSM-derived descriptors can reduce reliance on purely empirical relationships and improve transferability across sites when point density and occlusion conditions allow faithful reconstruction. However, QSM performance is highly sensitive to point-cloud quality and sampling geometry: noise, variable point density, occlusion, and segmentation errors can propagate into biased volume estimates, especially in complex crowns, hedgerow architectures, and intensively managed orchards where canopy structure departs from typical forest forms. Consequently, the feasibility and reliability of QSMs in woody-crop settings remain less established than those in forest terrestrial laser scanning (TLS) applications, and systematic comparisons against conventional inventory-based allometry and plot-level LiDAR models are still limited.

Therefore, in this study, we complemented cross-platform plot-scale modeling with a TreeQSM-based QSM evaluation as an independent, tree-level 3D reconstruction pathway. By comparing QSM-derived biomass estimates with field inventory across sites spanning low- to higher-biomass conditions, we assessed not only accuracy but also practical usability (e.g., proportion of trees successfully reconstructed) under real acquisition constraints. This dual perspective of geometric reconstruction (QSM) versus area-based predictors (ML models) helps clarify when 3D structural modeling offers added value for biomass and carbon estimation in Mediterranean woody-crop agroecosystems.

 doi:10.20944/preprints202603.0280.v1

2. Materials and Methods

2.1. Study Sites

The study was conducted in three farms in the province of Córdoba (southern Spain): IFAPA (Finca Experimental de la Alameda del Obispo, Córdoba, Spain), Doña María (Córdoba, Spain; Cerro del Obispo, S.L.), and Villaseca (Almodóvar del Río, Córdoba, Spain; Agrícola Villaseca S.L., Cortijo La Reina). The sites span research to commercial settings and contrasting orchard structures (Figure 1).

IFAPA: The study area covers ~6.82 ha within the Alameda del Obispo experimental farm. Research farm (IFAPA) co-located with CSIC and INIA facilities; the study focused on olive and almond plots. Olive planting layouts: 6×2.5 m, 4×1.5 m, 7×3.4 m; Almond: 7×6 m, 3.5×1.2 m, 6×3 m. Predominantly flat terrain with very gentle slopes; alluvial soils with sandy-loam texture and substantial silt; soil depth > 2 m. Mediterranean climate: mean annual precipitation of ~600 mm (October–April). For 2014: mean temperature 19°C, precipitation 572 mm, RH 64%; prevailing winds E→W; located in the Guadalquivir floodplain (flood risk noted). Olives and almonds established in 2018; drip irrigation; manual and mechanical pruning; olive harvest November-December; almond harvest August.

Doña María: Topographically variable farm combining new plantings with older trees; 3 parcels defined for analysis. The farm is described as 187 ha (farm description) and 261.45 ha (spatial data extent). Olive (Olea europaea) with Picual and Arbequina varieties. Square-shaped with undulating relief; mean slope 2.7%; altitude ~230 m a.s.l.; clay soil (other chemical properties not available). Mean annual temperature ~19°C; precipitation ~600 mm; RH ~65%; prevailing winds W→E; groundwater resources mentioned. Drip irrigation; mineral fertilization; manual pruning; chemical pest control; olive harvest November-December.

Villaseca: Commercial, intensively managed farm; 13 parcels derived for analysis and scalability testing. The total farm area is ~85 ha. Regular square shape and flat relief; slope 1.80%; altitude ~80 m a.s.l.; orientation NW–SE; clay-loam soil, depth ~2 m; pH 8.5; organic matter 1.20%. For 2024: mean temperature 18.5°C, precipitation 697 mm, RH 65%; prevailing winds W→E. Olive and almond planting established in 2018; drip irrigation; fertilization with liquid mineral fertilizers; phytosanitary treatments applied.

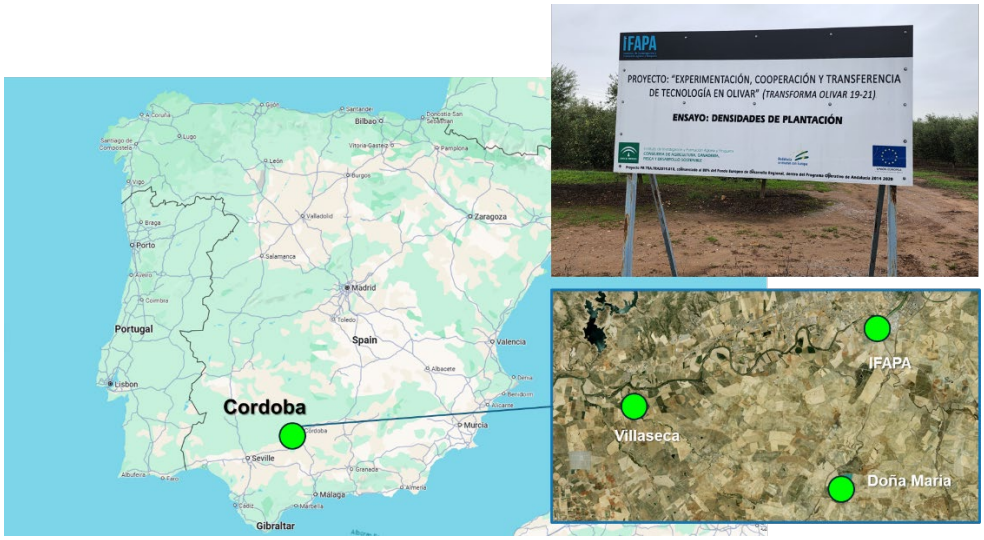

Figure 1. Location of the study sites.

 doi:10.20944/preprints202603.0280.v1

2.2. Field Inventory, Plot Design, and Biomass/Carbon Computation

Fixed-area rectangular plots of 20 m × 50 m (1000 m²; 0.1 ha) were used as the primary sampling units (Figure 2). Plots were selected using simple random sampling to minimize selection bias and ensure independence among observations. A total of 58 plots were surveyed (Doña María: 18; IFAPA: 20; Villaseca: 20). Field measurements were collected during December 2024 and January 2025 across three farms (Villaseca, Doña María, and IFAPA). Tree diameters were measured with a measuring tape, and tree height was measured (Figure 3) using two 5 m extendable poles (calibrated prior to measurements). Plot locations were handled using a mobile GNSS application (GPS Fields Area Measure) and photo documentation, and the data were recorded using standardized spreadsheets.

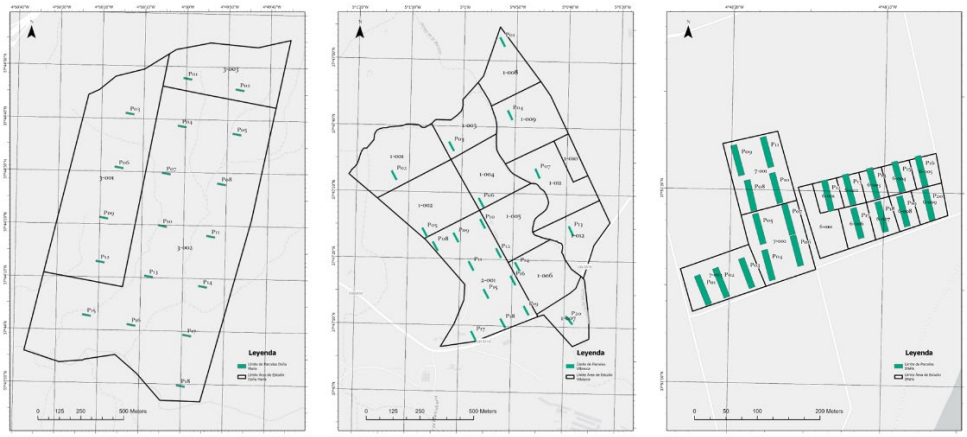

Figure 2. Study sites and field inventory plots.

Figure 3. a. Methodology and equipment used for field measurements.

 doi:10.20944/preprints202603.0280.v1

Figure 3. b. Methodology and equipment used for field measurements.

Tree-level aboveground biomass (AGB, kg dry matter) was computed using component-wise allometric equations parameterized by the trunk diameter below the first bifurcation (D2r, cm).

For almond (Prunus dulcis), component biomass was computed as:

- Branches: B_branches = 0.0688 × D2r^2.2061 × 1.1349 × ϱ;
- Stem: B_stem = 0.3346 × D2r^1.9656 × 0.9968 × ϱ;
- Foliage: B_foliage = 0.0165 × D2r^1.3212 × 0.9564 × ϱ.

For olive (Olea europaea), component biomass was computed as:

- Branches: B_branches = 0.251 × D2r^2.071 × ϱ;
- Stem: B_stem = 0.0785 × D2r^2.527 × ϱ;
- Foliage: B_foliage = 0.0446 × D2r^1.724 × ϱ.

Wood density was set to ϱ = 0.76 g·cm−3 for almond (Prunus dulcis) and ϱ = 0.56 g·cm−3 for olive (Olea europaea). D2r was measured in cm; component biomass is in kg. The equations are applicable for 5-50 cm trunk diameter. The tree AGB was computed as the sum of the components. Plot-level AGB density (Mg ha−1) was derived by summing the tree AGB within each plot, converting kg to Mg, and dividing by the plot area (0.1 ha).

Tree-level AGB (kg) was converted to carbon (C, kg C) using mean wood carbon fractions of 44.5% for olives and 45.7% for almonds (COlive = AGB×0.445; CAlmond = AGB×0.457). CO2-equivalent was computed as CO2e = 3.67 × C following IPCC molecular-weight conversion. Carbon-fraction sources include orchards and olive allometry/carbon studies [4,5,12,13].

Table 1. Field inventory summary by site (mean ± standard deviation).

Site	Statistic	N_trees_plot	D2r_cm	Height_m	AGB_Mg_ha	C_Mg_ha
Villaseca	Media	55.1	17.76	5.09	33.89	15.93
Villaseca	Desv. Est.	27.48	8.06	2.42	9.64	4.53
Doña María	Media	18.44	17.12	4.18	30.94	14.54
Doña María	Desv. Est.	0.83	2.43	0.31	9.35	4.39
IFAPA	Media	21.65	13.8	3.96	12.76	6
IFAPA	Desv. Est.	12.04	3.35	0.44	5.34	2.51

2.3. LiDAR Datasets

Four LiDAR modalities were analyzed: (i) PNOA/ALS (Spain's National Aerial Orthophotography Plan including an open airborne LiDAR program) [18], (ii) airborne laser

 doi:10.20944/preprints202603.0280.v1

scanning ALS (RIEGL LMS Q680i), (iii) unmanned laser scanning ULS (DJI Zenmuse L2), and (iv) mobile laser scanning MLS (LiDAR USA Snoopy A-Series with Velodyne HDL-32E).

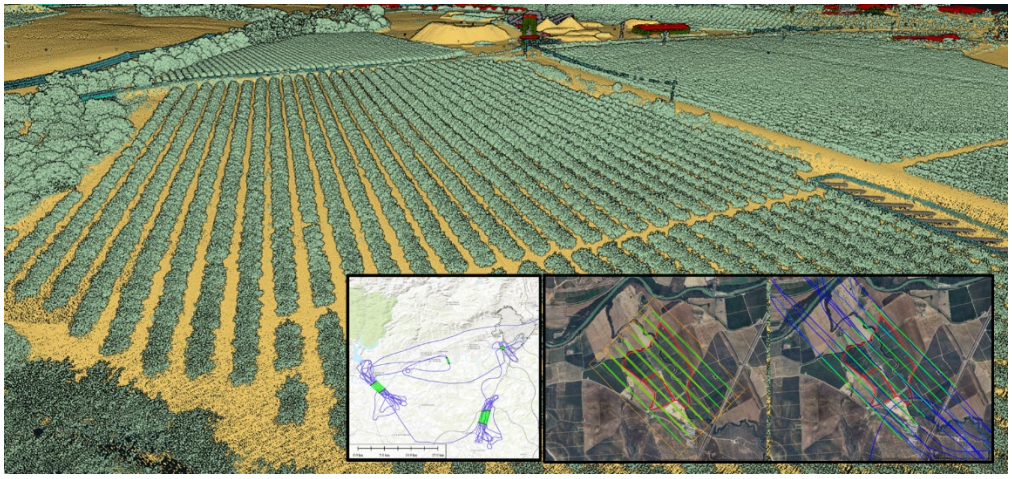

Figure 4. ALS Riegl airborne flight plans and classifies point cloud data.

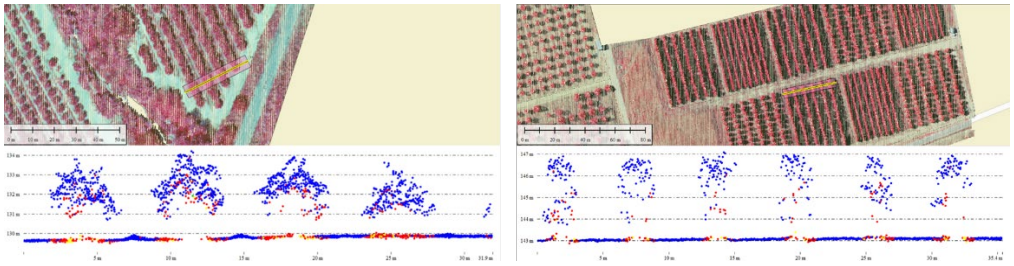

Figure 5. ALS PNOA point cloud data and cross sections.

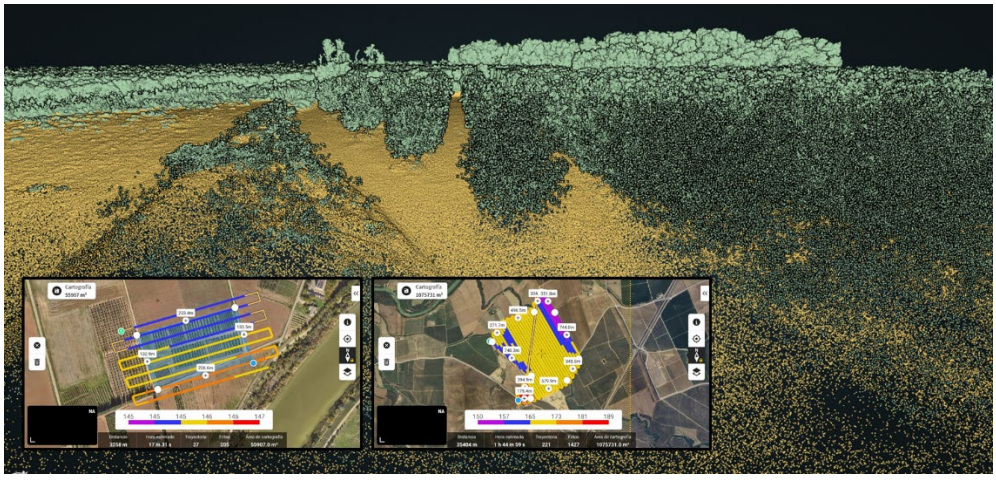

Figure 6. ULS DJI drone flight plans and classifies point cloud data.

 doi:10.20944/preprints202603.0280.v1

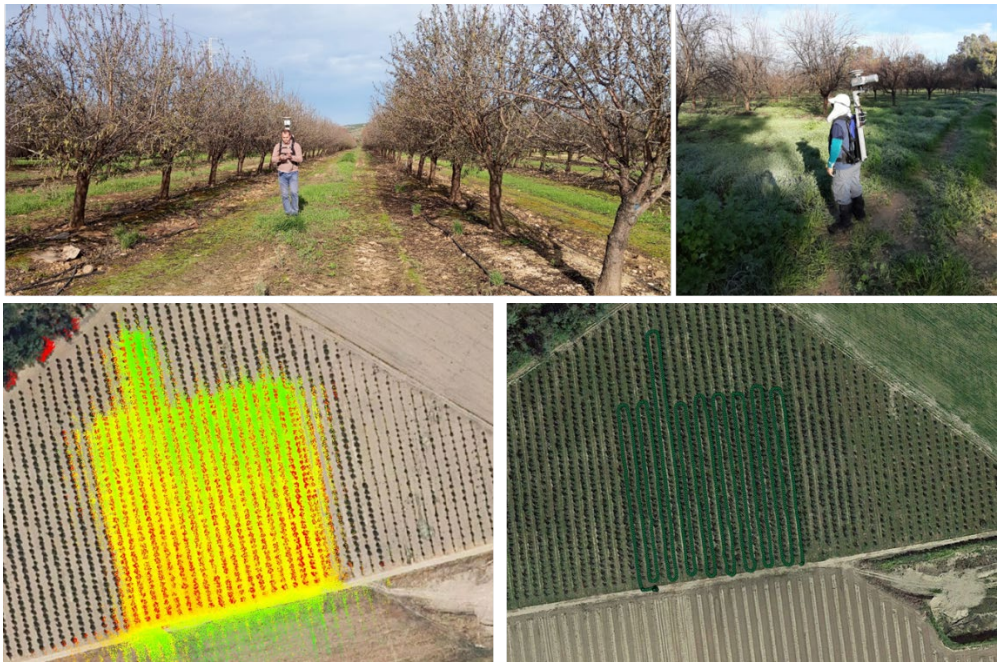

Figure 7. MLS Snoopy terrestrial trajectories and point cloud data.

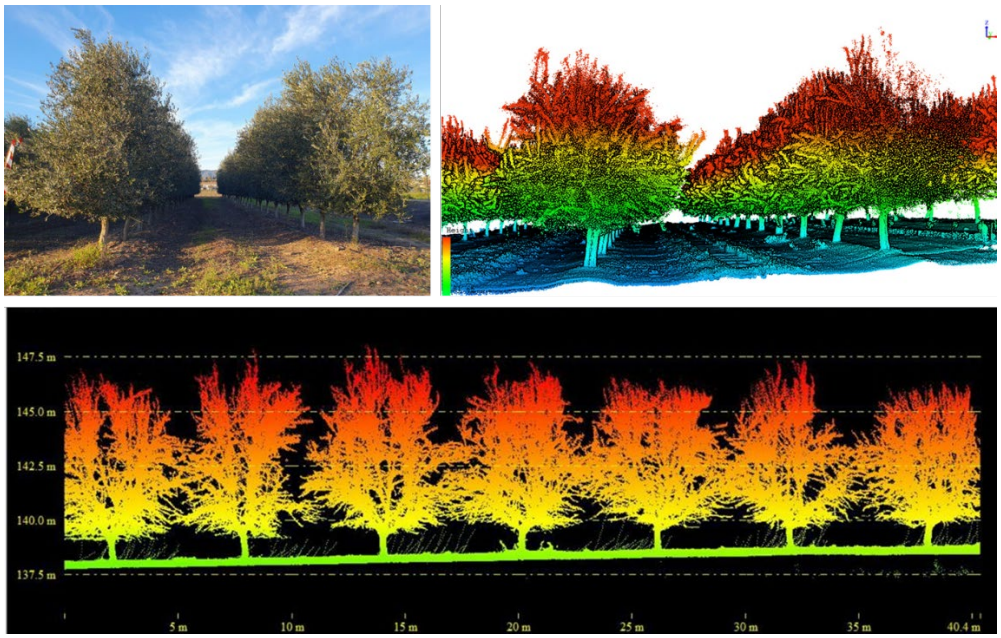

Figure 8. MLS Snoopy point cloud data.

ALS was prioritized for large extents (e.g., Doña María), ULS was used for detailed characterization in smaller areas (e.g., IFAPA), and MLS-supported close-range sampling was used (Figure 9). Dataset metadata, including point density and recommended CHM resolution, are summarized in Tables 2 and 3.

 doi:10.20944/preprints202603.0280.v1

Table 2. LiDAR modality summary (density, suggested CHM resolution, and spatial reference metadata).

System	Point spacing (m)	Density (pts/m²)	Intensity (bits)	Color	Multiple returns	Scan angle (deg)	Flight/ Survey date	EPSG	Orthometric heights	Stated Accuracy XY/Z
PNOA ALS	0.447	5.0	16	RGBI	Yes (4)	-19.62 – 3.17°	27/09/2024	25830	EGM2008-REDNAP	0.30 / 0.10 m
Riegl ALS	0.146	46.8	8	RGB	Yes (7)	-30 – 30°	21/11/2024	32630	EVRS	0.06 / 0.06 m
ULS	0.02	2470.0	8	RGB	No	-35 – 35°	18/12/2024	32630	EVRS	0.10 / 0.05 m
MLS	0.005	36545.0	8	N/A	No	-90 – 65°	10/12/2024	32630	EVRS	0.02 / 0.02 m

Table 3. Sensor/platform and acquisition specifications (as provided).

	PNOA/ALS (national)	Riegl ALS	ULS	MLS
Sistema LiDAR	Optech T2000	Riegl LMS-Q680i	DJI Zenmuse L2	LiDAR USA Velodyne hdl-32e
Sistema GNSS/IMU	POS AV™ AP60 (OEM)	IGI IMU LIE	DJI 200 Hz	Snoopy INS (OEM)
Sw Plan/Naveg./Post	Optech FMS y LMS (LiDAR Mapping Suite)	IGI Plan v1.5.5 /Aerocontrol/ Inertial Explorer 8.9 y AeroOffice 5.58	DJI Terra	Scan look PC LiDAR 360 Inertial explorer
Hv (m)	2.200	420	80	2
FOV	50	60	70	40ºV – 360ºHZ
Frec.- Escaneo (Hz)		400		
Frec. Pulso (Kz)	700-1.100	266	240	1.200
Velocidad	160 Knot	85 Knot	6 m/s	1 m/s
Densidad nominal (ps/m2)	5	11	300	33.000
FWF	Sí, no disponible	Sí	No	No
Retornos	4	7	3	1
Intensidad	16	8	8	8
Color	RGBI	ND	RGB	ND
Rec. Transversal	>15%	>90%	>90%	N/A
EERR	IGN/RAP	IGN/RAP (CRBD)	IGN/RAP	IGN/RAP
Precisión (XY/Z m)	0,30/0,10	0,06/0,06	0,10/0,05	0,02/0,02

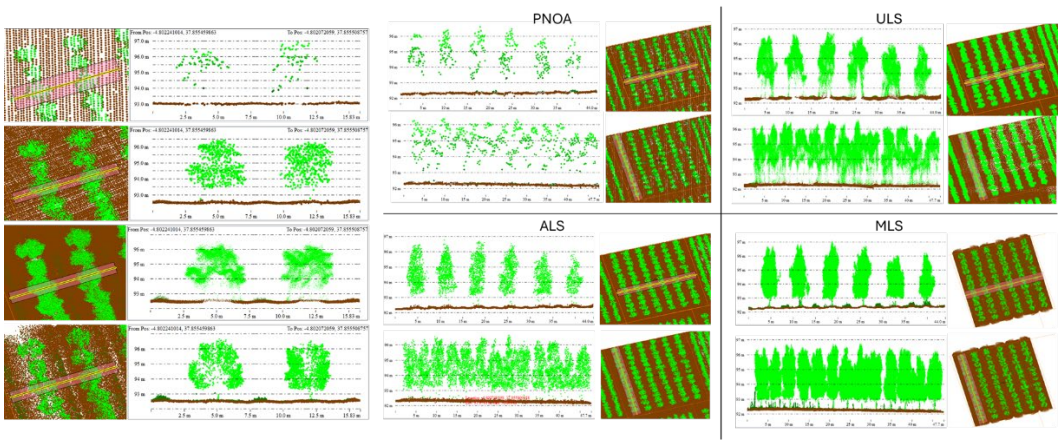

Figure 9. LiDAR point cloud data; cross-sections from all sensors.

 doi:10.20944/preprints202603.0280.v1

2.4. Pre-Processing and Metric Extraction

LiDAR point clouds were processed to support cross-platform comparison. Pre-processing included noise filtering and outlier removal, ground/vegetation separation, point cloud classification and segmentation using deep learning algorithms, terrain interpolation to produce a digital terrain model (DTM), and height normalization to express point elevations relative to the ground. Canopy height models (CHMs) were generated by rasterizing normalized heights; pit-free CHM approaches can reduce canopy surface artifacts in discrete-return LiDAR [8].

For metric extraction, LAS/LAZ files were handled programmatically (e.g., using Laspy in Python and LasCanopy from LAStools), and plot-level structural metrics were computed from normalized point clouds. The project workflow describes plot metrics stored in CSV files per site and system, including mean height, 95th percentile of height (p95), maximum height, and canopy-cover proxies. Additional geometric descriptors (e.g., point density and slope/curvature metrics) were considered during preprocessing to characterize terrain and sampling effects.

Because sources differ in coordinate reference systems and orthometric height references, all datasets should be harmonized to a common horizontal CRS and vertical datum prior to comparison. The final submission should document the transformation parameters and quantify the residual co-registration errors.

2.5. Modeling and Validation

LiDAR-derived metrics (e.g., mean height, p95, maximum height, and canopy-cover proxies) were extracted for each plot and each LiDAR system (ALS, ULS, MLS, and PNOA) and stored in per-system CSV tables. The field AGB (response) and LiDAR metrics (predictors) were merged using a plot identifier. Non-predictive identifiers were removed, and predictors were standardized using a z-score scaler (StandardScaler). Random Forest (RandomForestRegressor) and XGBoost (XGBRegressor) models were trained, and an ensemble estimate was derived as the arithmetic mean of the two model predictions.

Data were split into training and test subsets using an 80/20 partition to evaluate generalization on unseen plots. The performance was reported using R^2, RMSE, MAE, and bias (absolute and relative). The project design indicates that Villaseca and Doña María were used for model development/testing, while IFAPA served as a control/validation site; however, the exact external-validation configuration should be explicitly stated in the final analysis if models were trained without IFAPA data. To compare structural variables and derived biomass/carbon among sites and/or systems, both parametric (ANOVA) and non-parametric tests (Kruskal-Wallis) were used depending on distributional assumptions; Dunn post-hoc tests were used for pairwise comparisons when non-parametric significance was detected. Normality was assessed using Shapiro-type tests in the analysis workflow.

In addition to the plot-level area-based models, we evaluated a tree-level 3D reconstruction approach using TreeQSM (Figures 10 and 11). In our workflow, TreeQSM was applied to point clouds generated only by MLS, in which individual-tree subsets could be isolated with sufficient quality (i.e., adequate point density, limited noise, and reduced occlusion). Prior to the QSM reconstruction, point clouds were filtered to remove outliers and non-woody points, and tree-level subsets were extracted using the available segmentation outputs and field plot information. QSM outputs were inspected for geometric plausibility (e.g., continuous branching structure and realistic cylinder radii) and flagged when reconstruction artifacts indicated segmentation failure or insufficient sampling.

 doi:10.20944/preprints202603.0280.v1

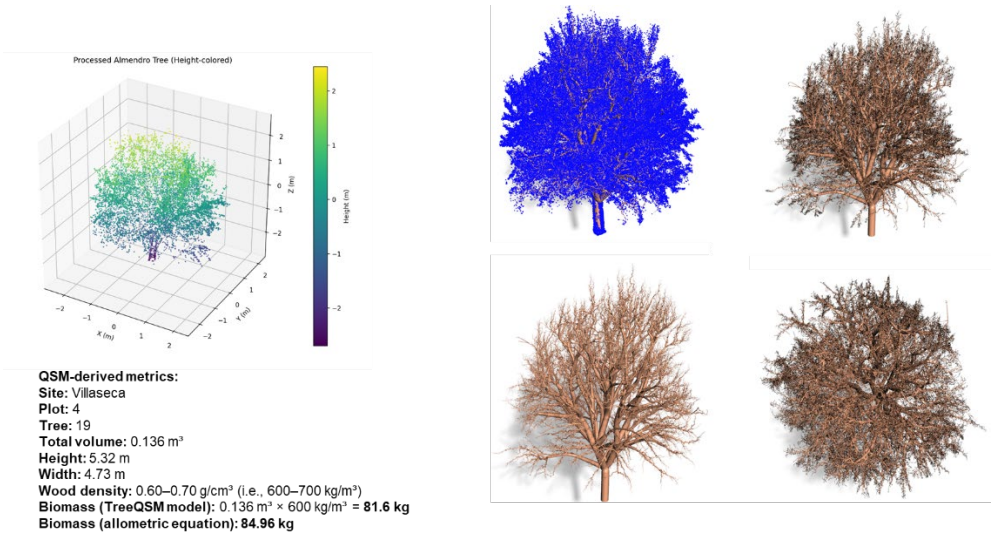

Figure 10. QSM Reconstructed 3D model of the almond tree.

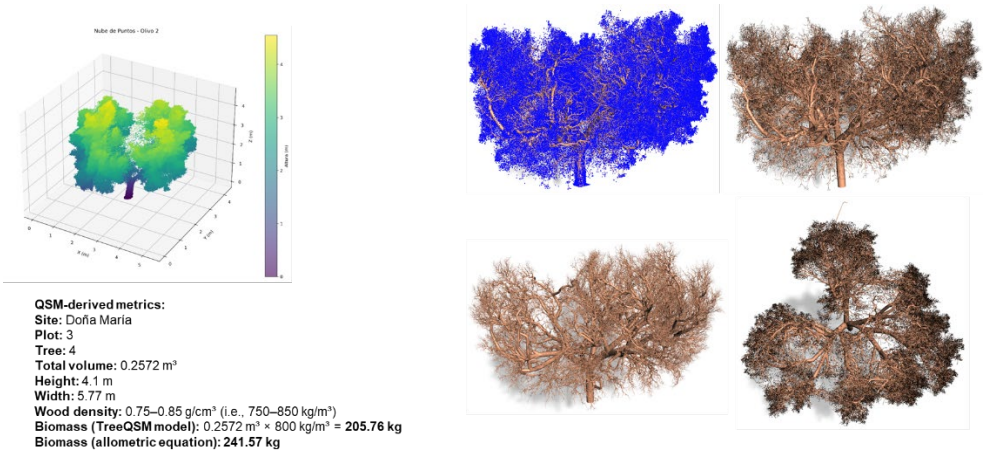

Figure 11. QSM Reconstructed 3D model of the olive tree.

TreeQSM-derived woody volume estimates were converted to biomass using species-appropriate wood density values and, where required, converted to carbon and CO_2-equivalent using the same carbon fractions and CO_2 conversion factor applied to inventory-derived estimates to ensure comparability. The QSM pathway is conceptually distinct from the area-based approach: rather than predicting plot AGB from summary LiDAR metrics, QSM estimates are grounded in the explicit reconstruction of 3D tree geometry. This provides a complementary line of evidence to assess whether geometric reconstruction can reduce bias relative to purely empirical predictors, particularly across sites with contrasting orchard architectures.

Because QSM performance is strongly dependent on point cloud completeness and density, we quantified practical usability by reporting the proportion of trees for which TreeQSM produced usable reconstructions. Reconstructions were considered usable only when (i) the fitted model represented a coherent woody structure and (ii) key structural attributes were within plausible ranges for the species and management system. Trees were excluded when the reconstruction was compromised by insufficient point density, excessive noise, severe occlusion, or segmentation errors. This usability metric is critical for operational deployment because QSM methods can deliver high-fidelity outputs only for a subset of trees under typical orchard acquisition conditions.

TreeQSM-derived biomass estimates were evaluated against field inventory at the tree level using standard regression diagnostics and error metrics, including the coefficient of determination (R^2), bias, and dispersion of residuals. When appropriate, the results were summarized by site to test

 doi:10.20944/preprints202603.0280.v1

for site-dependent agreement and systematic over- or underestimation patterns. In parallel with the machine-learning models, we emphasize that QSM–inventory agreement should be interpreted in light of sampling geometry and data quality constraints; therefore, we report both accuracy metrics and the fraction of trees successfully reconstructed as complementary indicators of performance and applicability.

3. Results

3.1. Field Biomass and Carbon Stocks

The field inventory indicated strong site differences in structure and biomass (Table 1). Mean AGB was 33.89 Mg ha−1 in Villaseca, 30.94 Mg ha−1 in Doña María, and 12.76 Mg ha−1 in IFAPA (Figures 12a, 12b and 12c). These contrasts are consistent with the differences in planting layout, management, and site conditions described in Section 2.1.

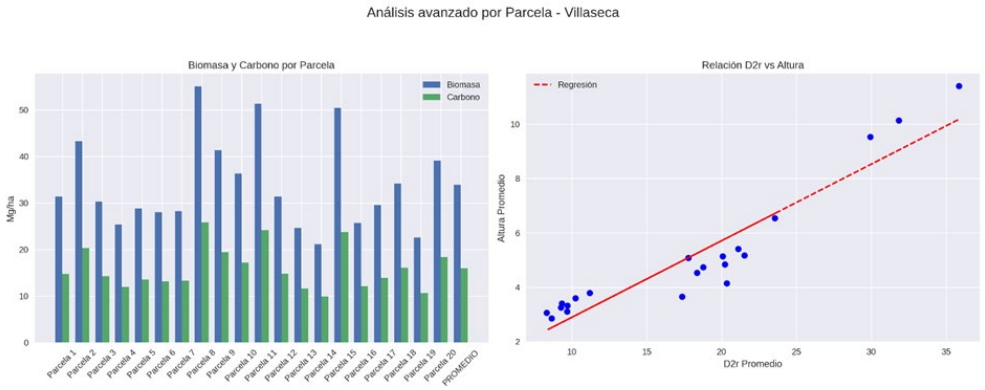

Figure 12. a. Villaseca site, biomass estimations per site per plot.

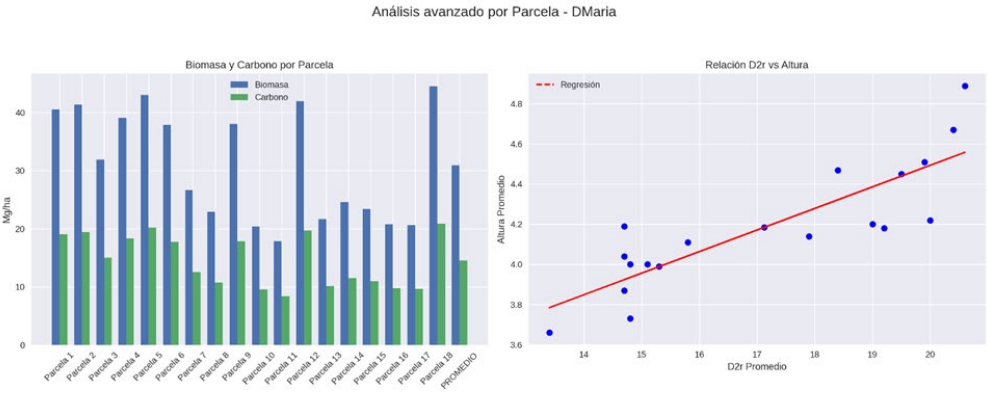

Figure 12. b. Doña Maria site, biomass estimations per site per plot.

 doi:10.20944/preprints202603.0280.v1

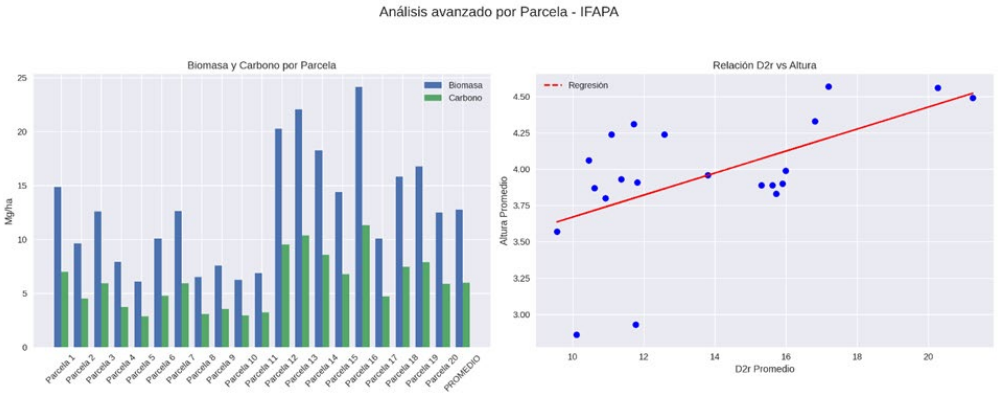

Figure 12. c. IFAPA site, biomass estimations per site per plot.

3.2. Model Performance

The model performance summaries are reported in Table 5 and Figures 13–15. In the reported results, XGBoost achieved low errors at IFAPA and Villaseca, whereas PNOA was competitive at Doña María. Performance interpretation should explicitly consider train–test partitioning and any site-based external validation if applied.

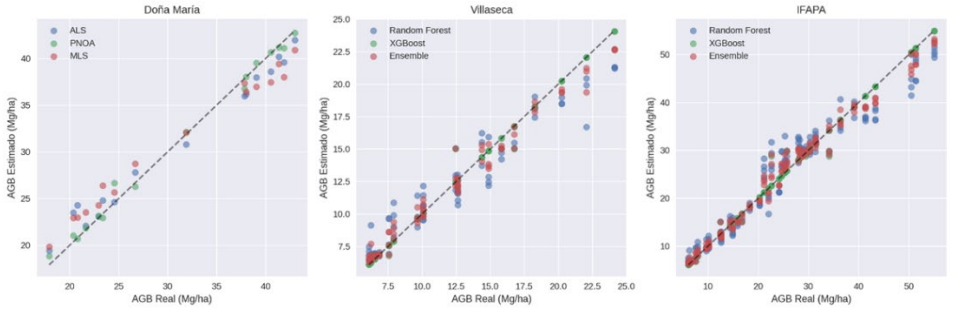

Figure 13. Scatter Plots.

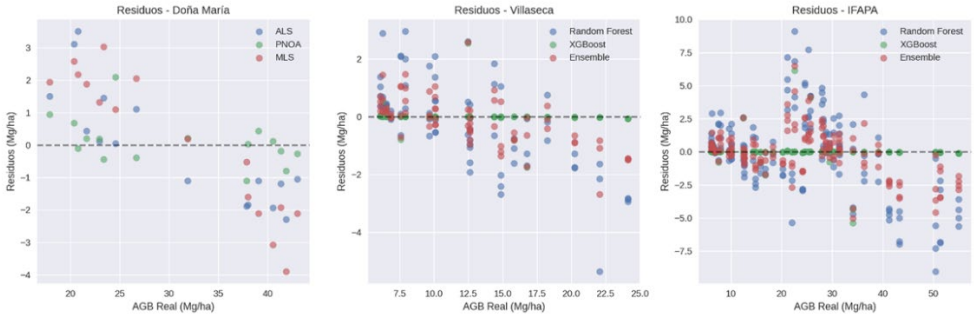

Figure 14. Analysis of Residuals.

 doi:10.20944/preprints202603.0280.v1

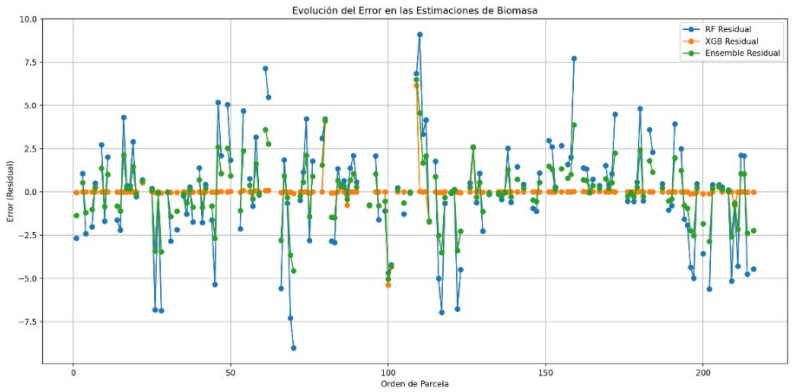

Figure 15. Biomass estimation error evolution.

3.3. Cross-Site Comparison of Results

Statistical testing (ANOVA and Kruskal-Wallis) indicated that Villaseca, Doña María, and IFAPA differed significantly in D2r, tree height, trees per plot, biomass, carbon, and CO_2e, with consistently very low p-values. Doña María shows higher and more stable biomass and carbon stocks (small standard errors), suggesting relatively homogeneous growth conditions and management. In contrast, IFAPA exhibited greater dispersion, and the observed inverse D2r-biomass pattern indicates stronger environmental and/or management-driven variability within the site. Villaseca presents an intermediate and, in some metrics, a contrasting response relative to the other sites, making it particularly useful as a benchmark under differing structural and management conditions.

3.4. Statistical Comparison Across Sites (ANOVA and Kruskal–Wallis)

To assess whether tree structure and biomass-related variables differed among the three study sites (Villaseca, Doña María, and IFAPA), we applied both a parametric one-way analysis of variance (ANOVA) and a non-parametric Kruskal–Wallis test. The results are summarized in Table 4 and Figures 13, 14, and 15.

ANOVA results. One-way ANOVA indicated statistically significant differences among the sites for most variables. Specifically, the mean D2r differed across sites ($F \approx 3.35$, $p = 0.0421$), as did the mean tree height ($F \approx 3.51$, $p = 0.0365$). The number of trees per plot showed a strong site effect ($F = 27.28$, $p < 0.0001$). Similarly, plot-level biomass density differed markedly among sites ($F = 39.48$, $p < 0.0001$), and the same strong site effect was observed for carbon density ($F = 39.48$, $p < 0.0001$). Collectively, these results suggest that differences in site conditions and management practices are associated with significant variations in orchard structure, productivity, and carbon stocks.

Kruskal–Wallis results. Because normality cannot always be assumed, we complemented the ANOVA with the Kruskal–Wallis test. The Kruskal–Wallis statistic for D2r suggested a trend toward site differences ($H \approx 5.34$, $p = 0.0694$), although it was marginally non-significant at the conventional $\alpha = 0.05$ level. For height, no significant differences were detected ($H \approx 2.34$, $p = 0.3104$). In contrast, the number of trees per plot remained highly site-dependent ($H \approx 28.86$, $p < 0.0001$), and both biomass density ($H \approx 37.97$, $p < 0.0001$) and carbon density ($H \approx 37.97$, $p < 0.0001$) exhibited strong consistent differences among sites.

Overall, the agreement between ANOVA and Kruskal–Wallis for tree density (N_trees), biomass, and carbon provides robust evidence that site-specific conditions and management exert a strong influence on these key variables, whereas differences in D2r and height appear to be weaker and more sensitive to distributional assumptions.

Table 4. Summary integrating the results of both tests: ANOVA and Kruskal-Wallis.

Variable	ANOVA: F-statistic (p-value)	Kruskal-Wallis: H-statistic (p-value)

 doi:10.20944/preprints202603.0280.v1

D2r	3.35 (0.0421)	5.34 (0.0694)
Tree Height	3.51 (0.0365)	2.34 (0.3104)
N_Tree	27.28 (<0.0001)	28.86 (<0.0001)
Biomass	39.48 (<0.0001)	37.97 (<0.0001)
Carbon	39.48 (<0.0001)	37.97 (<0.0001)

Table 5. Model performance summaries extracted from project documentation.

Site	Group	RMSE_Mg_ha	Rel_RMSE_%	Bias_Mg_ha	Rel_Bias_%	R^2	Mean_Rel_Error_%	SD_Rel_Error_%
Doña María (platform comparison)	ALS	1.743	5.671	-0.0675	-0.219	0.985	1.382	6.865
Doña María (platform comparison)	PNOA	0.725	2.359	0.1025	0.333	0.994	0.713	2.809
Doña María (platform comparison)	MLS	2.168	7.053	0.065	0.211	0.988	2.298	7.496
IFAPA (model comparison)	Random Forest	1.542	12.559	-0.182	-1.484	0.932	1.859	13.150
IFAPA (model comparison)	XGBoost	0.400	3.257	-0.003	-0.025	0.994	0.105	3.275
IFAPA (model comparison)	Ensemble	0.849	6.918	-0.092	-0.755	0.981	0.982	7.185
Villaseca (model comparison)	Random Forest	3.060	12.853	-0.010	-0.042	0.950	2.547	12.797
Villaseca (model comparison)	XGBoost	0.872	3.666	-0.004	-0.017	0.995	0.137	3.796
Villaseca (model comparison)	Ensemble	1.714	7.200	-0.007	-0.029	0.984	1.342	7.223

4. Discussion

At the low-biomass IFAPA site, TreeQSM closely matched the field inventory (51.14 vs. 51.11 kg tree^{-1}; Δ = 0.03 kg tree^{-1}; R^2 = 0.61). In contrast, TreeQSM tended to overestimate biomass at Doña María and Villaseca (+115.78 and +43.03 kg tree^{-1}, respectively; R^2 = 0.42 and 0.59), indicating site-dependent bias. Among the machine-learning approaches, XGBoost consistently outperformed Random Forest (residual SD 0.876 vs. 3.07), and the ensemble improved Random Forest performance in ~60% of cases. Comparisons across LiDAR sources also revealed systematic differences (e.g., elevation offsets of ~1.23 m vs. 0.68 m between ULS and ALS), underscoring the need for cross-sensor calibration. TreeQSM applicability was further constrained by data quality: only 28.28% of the reconstructed 3D models were usable owing to point density limitations, noise, occlusion, and segmentation failures (of the 1,867 scanned trees, only 522 were successfully reconstructed into models, corresponding to approximately 28% of the total). Where feasible, 3D structural reconstructions provide stronger inputs than raw point-cloud metrics for allometric modeling and regression, improving geometric fidelity and reducing bias. Overall, the proposed workflow is transferable and can be replicated across a wide range of agroforestry and woody-crop settings.

By combining fixed-area plot inventory, species-specific allometry, and cross-platform LiDAR metrics, this study supports accurate plot-scale AGB and carbon estimation in Mediterranean woody

 doi:10.20944/preprints202603.0280.v1

crops. The project results indicate that open national LiDAR (PNOA) can be competitive under certain conditions, supporting scalable MRV. Key methodological priorities for final submission are (i) complete documentation of CRS/vertical harmonization and co-registration error, and (ii) transparent reporting of validation design, including potential spatial autocorrelation controls.

 doi:10.20944/preprints202603.0280.v1

4.1. Comparative Performance of LiDAR Modalities

Our results confirm that differences among LiDAR modalities influence not only point density and geometric precision, but also critically, the ability to characterize canopy three-dimensional structure, which ultimately governs aboveground biomass (AGB) estimation. MLS provided the finest spatial detail and most faithful representation of individual-tree structure, capturing crown architecture, intra-canopy gaps, and even trunk bases with high clarity. Its close-range sampling and very high point density (≈36,500 pts m^{-2}) largely compensated for the absence of multiple returns, making MLS a valuable "3D ground-truth" reference for calibrating structural metrics and validating other modalities.

ALS has emerged as the most balanced modality for operational biomass modeling, combining broad airborne coverage with multi-return capability (up to seven returns) and high vertical precision (≈0.06 m). This configuration enables robust retrieval of vertical canopy structures, including lower strata, which supports the modeling of height percentiles (e.g., RH metrics), canopy cover proxies (FCC), and height heterogeneity—features that are essential for plot-scale AGB estimation and regional upscaling.

Despite its exceptional point density (>2,000 pts m^{-2}), ULS exhibited limitations linked to inter-strip misalignments and the lack of multiple returns in some configurations, which can increase uncertainty in lower-canopy characterization. Nevertheless, ULS remains highly attractive for detailed sampling over limited extents owing to its operational agility and comparatively low cost, provided that geometric strip adjustment and radiometric/intensity harmonization are implemented.

The PNOA LiDAR (≈5–8 pts m^{-2}; up to four returns) provides a strong foundation for systematic mapping and long-term monitoring at large scales. Although its moderate density does not allow the canopy internal structure to be resolved with the same fidelity as ALS or MLS, its nationwide coverage and temporal consistency make it a key resource for repeated carbon monitoring and scalable inventory updates.

4.2. Implications for AGB Modeling

AGB modeling is strongly conditioned by the quality and characteristics of the input point clouds. In this study, metrics derived from ALS and PNOA, particularly height percentiles (e.g., RH50, RH95) and canopy cover fraction (FCC), showed stable relationships with field observations and supported allometric/regression models with relative errors typically below ~10–15% (site-dependent). For ULS and MLS, high-density and close-range sampling enable precise characterization of crown volume proxies and, in principle, tree-level biomass estimation; however, these modalities require modality-specific calibration and normalization to account for scan-angle effects, occlusion, and radiometric discrepancies among systems.

A key methodological issue is that the absence of multiple returns (or an inconsistent return structure) in the ULS and MLS can introduce positive bias in height-based percentiles, particularly when lower-canopy penetration is limited. This bias can be reduced through cross-modal fusion with ALS (multi-return reference), pass-level classification strategies, or harmonized point-cloud normalization workflows that explicitly model sampling completeness. Overall, our findings reinforce the value of LiDAR-derived structural metrics, especially RH metrics, height variance, and FCC, as robust predictors for biomass estimation in Mediterranean woody crops, consistent with prior evidence.

4.3. Scalability and Hierarchical Integration Strategy

A hierarchical integration of LiDAR modalities MLS, ALS, and PNOA has emerged as an effective and scalable framework for biomass and carbon estimation across spatial scales. Under this strategy, the MLS functions as a high-resolution 3D reference dataset to calibrate structural metrics, develop geometry-driven relationships, and validate local predictions. ALS then enables regional upscaling of these relationships, leveraging its intermediate density, strong vertical accuracy, and

 doi:10.20944/preprints202603.0280.v1

multi-return structure to capture vertical canopy variability over larger extents. Finally, PNOA LiDAR provides a pathway to extend AGB models to national-scale carbon inventories and long-term monitoring, benefiting from consistent coverage and repeat acquisitions.

This tiered approach optimizes acquisition costs by reserving ultra-dense surveys for calibration/benchmarking while using airborne sources for regional mapping and national programs for periodic updates. Importantly, it also strengthens the coherence between local observations and regional products, supporting reproducible methodologies for biomass and carbon quantification in woody-crop systems.

4.4. Limitations and Future Perspectives

The main limitations identified were (i) the limited number of vertical control points for rigorous cross-system vertical harmonization, (ii) the absence or inconsistency of multiple returns for some sensors/configurations, and (iii) radiometric/intensity discrepancies among systems. Future work should prioritize intensity normalization and radiometric harmonization algorithms, AI-based segmentation methods tailored to orchard architectures (including hedgerow systems), and multitemporal analyses using successive PNOA LiDAR acquisitions to quantify canopy growth dynamics and carbon accumulation over time.

In addition, integrating LiDAR with high-resolution spectral imagery (multispectral or hyperspectral) is expected to improve biomass and productivity estimates by combining structural information with canopy conditions and vigor signals. Such multi-sensor fusion is particularly promising for woody agricultural systems, where management practices can decouple the structure from physiological status across seasons.

5. Conclusions

This study demonstrates that multi-platform LiDAR, combined with a rigorous fixed-area field inventory and species-specific allometry, can support robust aboveground biomass (AGB), carbon, and CO_2e estimations in Mediterranean woody-crop systems. The field results revealed clear, statistically significant differences among Villaseca, Doña María, and IFAPA in key structural and carbon-related variables (e.g., tree density, biomass, and carbon), highlighting the importance of site conditions and management in shaping productivity and carbon storage. Across sites, the trunk diameter measured below the first bifurcation (D2r) behaved consistently and was strongly informative for overall tree development, supporting its use as a practical structural indicator. However, the orchard architecture, particularly hedgerow systems, suggests that LiDAR validation and inference are more defensible at the hedge-segment or plot level than at the individual-tree level.

In modeling, machine-learning approaches provide accurate plot-scale predictions, with XGBoost generally outperforming Random Forest and ensemble strategies, improving stability in several cases. Cross-sensor comparisons indicate that differences among LiDAR sources (e.g., ULS vs. ALS) can introduce systematic offsets, reinforcing the need for harmonization and calibration prior to pooling metrics or transferring models. TreeQSM-based quantitative structure modeling provided an informative, geometry-driven complement to area-based metrics but showed strong sensitivity to point-cloud quality and usability constraints, with performance varying by site and acquisition conditions. These limitations should be explicitly considered before operational deployment in orchard MRV. Overall, the workflow is transferable and can be replicated across agroforestry and woody-crop settings, provided that site-specific calibration, transparent validation design, and uncertainty reporting are implemented.

Author Contributions: Mateo Pastrana: Conceptualization, Methodology, Investigation, Formal analysis, Data curation, Visualization, Writing-original draft preparation, Writing-review and editing, Project administration. Mateo Pastrana led the overall research design, multi-platform LiDAR integration strategy, biomass and carbon modeling workflow, statistical analysis, and manuscript preparation. Cristina Velilla: Supervision, Conceptualization, Methodology, Writing-review and editing, Funding acquisition. Cristina Velilla provided

 doi:10.20944/preprints202603.0280.v1

doctoral supervision, high-level methodological guidance, and critical scientific revision of the manuscript. Sergio Molina: Methodology, Validation, Resources, Writing-review, and editing. Sergio Molina contributed expert input on ALS (PNOA) data processing standards, LiDAR validation strategies, and technical reviews of the derived products. Nelson Mattie: Conceptualization, Methodology, Investigation, Data curation, LiDAR data processing workflows, python programming, analytical development for biomass estimation, Formal analysis, Writing-review and editing. Alfonso Gomez: Conceptualization, Validation, Investigation. Alfonso Gomez contributed to the experimental design refinement, validation framework, and support in data interpretation. All authors have read and agreed to the published version of the manuscript.

Funding: [NEEDS INPUT: Specify funding sources, including APC funding if applicable, or state "This research received no external funding."].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The processed point clouds, derived plot metrics, and analysis scripts were stored in a FAIR-aligned repository for the project. Data are available from the corresponding author upon reasonable request. [NEEDS INPUT: Provide repository link/DOI if public.].

Acknowledgments: The ALS, ULS and MLS LiDAR datasets were acquired by Stereocarto Iberoamerica S.L. The authors acknowledge their technical support for the data acquisition. The company had no role in the study design, data analysis, or interpretation of results. The authors acknowledge IFAPA and the farm owners/managers (Doña María and Villaseca) for their access and logistical support. The authors have reviewed and edited the manuscript and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

- AGB: Aboveground biomass
- ALS: Airborne laser scanning
- CHM: Canopy height model
- CO2e: Carbon dioxide equivalent
- DTM: Digital terrain model
- GEDI: Global Ecosystem Dynamics Investigation
- GNSS: Global Navigation Satellite System
- IFAPA: Instituto de Investigación y Formación Agraria y Pesquera
- IMU: Inertial measurement unit
- MLS: Mobile laser scanning
- MRV: Measurement–reporting–verification
- PNOA: National Aerial Orthophotography Plan (Spain)
- ULS: Unmanned laser scanning

Appendix A

Appendix A.1. Plot Centroid Coordinates

Table A1. UTM coordinates of plot centroids used for field inventory and LiDAR metric extraction.

Site	Plot	Easting	Northing
Doña María	1	338442.891	4177491.422
Doña María	2	338479.943	4177785.198
Doña María	3	338572.069	4178069.313
Doña María	4	338619.368	4178361.078

 doi:10.20944/preprints202603.0280.v1

Doña María	5	338682.799	4178665.661
Doña María	6	338773.056	4178955.806
Doña María	7	338791.065	4179213.691
Doña María	8	338360.591	4178729.133
Doña María	9	338453.508	4179003.026
Doña María	10	338486.634	4179280.926
Doña María	11	338154.990	4177847.548
Doña María	12	338257.349	4178129.910
Doña María	13	338338.505	4178423.586
Doña María	14	337894.781	4177905.352
Doña María	15	337974.339	4178216.396
Doña María	16	338084.834	4178765.124
Doña María	17	338149.355	4179078.420
Doña María	18	337995.796	4178472.399
Villaseca	1	322488.703	4183949.082
Villaseca	2	322625.775	4184007.307
Villaseca	3	322734.451	4184066.151
Villaseca	4	322547.632	4184144.189
Villaseca	5	322673.036	4184206.089
Villaseca	6	322477.923	4184274.370
Villaseca	7	322407.454	4184406.609
Villaseca	8	322311.322	4184366.350
Villaseca	9	322605.783	4184333.879
Villaseca	10	322531.865	4184470.533
Villaseca	11	322525.107	4184568.767
Villaseca	12	322119.568	4184696.117
Villaseca	13	322787.640	4184703.639
Villaseca	14	322660.101	4184975.439
Villaseca	15	322943.758	4184435.829
Villaseca	16	322934.346	4184020.871
Villaseca	17	322693.976	4184269.419
Villaseca	18	322262.459	4184430.594
Villaseca	19	322625.344	4185315.724
Villaseca	20	322387.154	4184830.815
IFAPA	1	341170.695	4191371.022
IFAPA	2	341232.111	4191326.678
IFAPA	3	341205.203	4191261.724
IFAPA	4	341217.546	4191383.992
IFAPA	5	341191.305	4191315.048
IFAPA	6	341251.737	4191276.441
IFAPA	7	341314.457	4191314.919
IFAPA	8	341348.624	4191324.491
IFAPA	9	341387.266	4191334.331
IFAPA	10	341427.348	4191344.763
IFAPA	11	341464.099	4191353.589
IFAPA	12	341475.656	4191300.190
IFAPA	13	341435.214	4191289.675
IFAPA	14	341400.434	4191280.931
IFAPA	15	341361.690	4191270.790
IFAPA	16	341219.843	4191204.017
IFAPA	17	341265.128	4191226.961
IFAPA	18	341185.157	4191191.251
IFAPA	19	341144.323	4191175.878

 doi:10.20944/preprints202603.0280.v1

IFAPA	20	341115.363	4191164.295

References

1. 1. Asner, G.P.; Mascaro, J. Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric. Remote Sens. Environ. 2014, 140, 614–624. https://doi.org/10.1016/j.rse.2013.09.023
2. 2. Silva, C.A.; Hudak, A.T.; Vierling, L.A.; Klauberg, C.; Garcia, M.; Ferraz, A.; Keller, M.; Eitel, J.U.H.Saatchi, S. Impacts of airborne lidar pulse density on estimating biomass stocks and changes in a selectively logged tropical forest. Remote Sens. 2017, 9, 1068. https://doi.org/10.3390/rs9101068
3. 3. Estornell, J.; Velázquez-Martí, B.; López-Cortés, I.; Salazar, D.; Fernández-Sarría, A. Estimation of wood volume and height of olive tree plantations using airborne discrete-return LiDAR data. GISci. Remote Sens. 2014, 51, 17–29. https://doi.org/10.1080/15481603.2014.883209
4. 4. Velázquez-Martí, B.; López-Cortés, I.; Salazar-Hernández, D.M. Dendrometric analysis of olive trees for wood biomass quantification in Mediterranean orchards. Agrofor. Syst. 2014, 88, 755–765. https://doi.org/10.1007/s10457-014-9718-1
5. 5. López-Bellido, P.J.; López-Bellido, R.J.; García-Morillo, J.; Fernández-García, P. Assessment of carbon sequestration and the carbon footprint in olive groves in Southern Spain. Int. J. Sustain. Dev. World Ecol. 2016, 23, 1–13. https://doi.org/10.1080/17583004.2016.1213126
6. 6. Dubayah, R.; et al. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. Remote Sens. Environ. 2020, 240, 111779. https://doi.org/10.1016/j.rse.2020.111779
7. 7. Duncanson, L.; et al. Aboveground biomass density models for NASA's GEDI mission. Remote Sens. Environ. 2022, 270, 112845. https://doi.org/10.1016/j.rse.2021.112845
8. 8. Khosravipour, A.; Skidmore, A.K.; Isenburg, M.; Wang, T.; Hussin, Y.A. Generating pit-free canopy height models from airborne lidar. Photogramm. Eng. Remote Sens. 2014, 80, 863–872. https://doi.org/10.14358/PERS.80.9.863
9. 9. Kukko, A.; Kaartinen, H.; Hyyppä, J.; Chen, Y. Multiplatform mobile laser scanning: Usability and performance. Sensors 2012, 12, 11712–11733. https://doi.org/10.3390/s120911712
10. 10. Breiman, L. Random forests. Mach. Learn. 2001, 45, 5–32. https://doi.org/10.1023/A:1010933404324
11. 11. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. https://doi.org/10.1145/2939672.2939785
12. 12. IPCC. 2006 IPCC Guidelines for National Greenhouse Gas Inventories; Intergovernmental Panel on Climate Change: Geneva, Switzerland, 2006.
13. 13. Kebede, B.; Soromessa, T. Allometric equations for aboveground biomass estimation of Olea europaea L. subsp. cuspidata in Mana Angetu Forest. Ecosyst. Health Sustain. 2018, 4, 1–12. https://doi.org/10.1080/20964129.2018.1433951
14. 14. Henry, M.; Bombelli, A.; Trotta, C.; Alessandrini, A.; Birigazzi, L.; Sola, G.; Vieilledent, G.; Santenoise, P.; Longuetaud, F.; Valentini, R.; Picard, N.; Saint-André, L. GlobAllomeTree: International platform for tree allometric equations to support volume, biomass and carbon assessment. iForest 2013, 6, 326–330. https://doi.org/10.3832/ifor0901-006
15. 15. Frank, B.; Mauro, F.; Allensworth, E. allometric: Structured Allometric Models for Trees. R package version 3.0.0, 2024. Available online: https://allometric.org/ (accessed on 26 February 2026).
16. 16. Proietti, S.; et al. Carbon footprint of an olive tree grove. Appl. Energy 2014, 127, 115–124. https://doi.org/10.1016/j.apenergy.2014.04.019
17. 17. Roussel, J.-R.; et al. lidR: An R package for analysis of airborne laser scanning (ALS) data. Remote Sens. Environ. 2020, 251, 112061. https://doi.org/10.1016/j.rse.2020.112061
18. 18. LiDAR, Especificaciones técnicas. Plan Nacional de Ortofotografía Aérea (PNOA), Instituto Geográfico Nacional (IGN). Available online: https://pnoa.ign.es/pnoa-lidar/especificaciones-tecnicas (accessed on 26 February 2026).

 doi:10.20944/preprints202603.0280.v1