Estimating soil organic carbon stocks in *Pinus halepensis* Mill. stands using LiDAR data and field inventories

Celia Herrero de Aza[1], David Moreno-Pérez[1], María-Belén Turrión[1], Francisco Lafuente Alvarez[1], Felipe Bravo[2], Irene Ruano[2], Frederico Tupinambá-Simões[2]

[1] Department of Soil Science and Agricultural Chemistry. University Institute for Research in Sustainable Forest Management (iuFOR), Escuela Técnica Superior de Ingenierías Agrarias (ETSIIAA). University of Valladolid, Avda. Madrid, 57. 34004. Palencia, SPAIN.

[2] Department of Plant Production and Forest Resources. University Institute for Research in Sustainable Forest Management (iuFOR), Escuela Técnica Superior de Ingenierías Agrarias (ETSIIAA). University of Valladolid, Avda. Madrid, 57. 34004. Palencia, SPAIN.

Correspondence to: Irene Ruano Benito (irene.ruano@uva.es)

Abstract. Accurate estimation of soil organic carbon (SOC) in forest ecosystems is essential for quantifying their contribution as carbon sinks and improving management strategies in the face of climate change. The objective of this study was to model the SOC content in *Pinus halepensis* forests by integrating structural variables derived from airborne LiDAR (Light Detection and Ranging) data with soil and forest inventory data. The study involved systematic soil sampling and a forest inventory on a plot covering 46.8 hectares in northern Spain. The use of LiDAR technology also provided 87 structural metrics for the plot. LiDAR metrics were then integrated with edaphic, dendrometric and stand variables to build predictive models of SOC stock by applying multivariate regression and machine learning techniques.

The results showed an average SOC of 20.33 Mg C/ha, with values ranging from 9.49 to 53.16 MgC /ha depending on bulk density, coarse fraction and total carbon concentration values, which ranged from 0.63% to 5.08%. In Outcomes of the modelling process indicated that the Random Forest algorithm performed best (R^2 = 0.81; RMSE = 7.73 Mg C/ha; MAE = 6.13 Mg C/ha), demonstrating adequate predictive capacity compared to other models. The research design made it possible to evaluate the potential of LiDAR data, acquired from airborne laser scanning (ALS) for general mapping programmes, as an effective tool for SOC estimation. This empirically validated procedure provides a useful methodological basis for advancing in SOC estimation through remote sensing and improving the quantification of soil-related ecosystem services. Our approach can assist forest managers and stakeholders in designing evidence-based strategies for soil carbon conservation in climate change scenarios such as those that affect Mediterranean *Pinus halepensis* stands, where drought and aridity may significantly influence SOC dynamics.

1. Introduction

Soil monitoring has become a central topic in the most recent European Union Environmental Strategies to ensure the sustainability of ecosystems (Panagos et al., 2020). The Soil Monitoring Law was approved in December 2025, to as a basis

for assessing, monitoring and improving soil health and resilience. It is the first common legal framework for soils in Europe, aimed at achieving healthy soils by 2050 (European Commission, 2021).

Under the new legistation, several parameters will be estimated to assess soil health status locally (enabling farmers and forest owners to analyse the impact of their management practices) and nationally (enabling countries to provide more information on the state of their soil resources (Rabot et al., 2024)). Soil organic carbon (SOC) is a basic indicator of soil health status, due to its importance in providing ecosystem services and supporting ecosystem functions (Navarrete-Poyatos et al., 2019). The role of SOC acquires even greater significance in areas that are more vulnerable to aridity (Chevallier et al., 2016). Climate change, is already altering Mediterranean forest dynamics by reducing carbon storage capacity and increasing the risk of desertification (del Río et al., 2008).

LiDAR (Light Detection and Ranging) is an active laser technology that is widely used in forest studies (Dassot et al., 2011; Borsah et al., 2023) due to its ability to accurately characterize canopy structure and estimate variables such as biomass or tree height (Oehmcke et al., 2021). By emitting and recording laser pulses, LiDAR technology allows us to characterize the three-dimensional forest canopy by generating metrics that describe canopy density, average or maximum tree height and other biomass attributes (Lefsky et al., 2002; Tupinambá-Simões et al., 2025). Advances in remote sensing techniques such as airborne laser scanning (ALS) (Navarro Cerrillo et al., 2018) have improved field inventories. Increasingly available national airborne LiDAR datasets in countries with systematic acquisition programs create new opportunities to upscale SOC estimation beyond plot-level assessments. When combined with a few basil soil observations, LiDAR-derived structural metrics can support spatially explicit SOC mapping over large areas, thereby contributing to more efficient and cost-effective soil monitoring strategies.

SOC is affected by different factors that regulate site fertility and productivity, including climate, vegetation, topography, edaphic (chemical, physical and biological) properties of the soil or bedrock, and land use and management practices (Lal, 2005; Schmidt et al., 2011; Ruiz-Peinado et al., 2017; Mayer et al., 2020). Canopy structure strongly influences the quantity and quality of carbon inputs and the soil microenvironment that regulates soil organic matter decomposition and stabilization. Forests with greater canopy height and more complex vertical structure typically produce larger quantities of litterfall and deeper, more extensive root systems (Norby et al., 2021). Aboveground litter provides a primary source of particulate organic matter, while root biomass and rhizodeposition supply carbon directly to deeper mineral soil horizons and support microbial pathways for SOC stabilization (Villarino et al., 2021). Structural complexity also modulates soil microclimate by influencing light penetration, temperature, and moisture regimes on the forest floor (Scherrer & Körner, 2011). Fine-scale LiDAR structural metrics correlate with understory microclimate conditions that affect microbial activity and decomposition rates, drawing attention to how canopy structure indirectly controls below-ground dynamics (Gril et al., 2023; Zellweger et al., 2019) that are fundamental to SOC accumulation and turnover across forest ecosystems.

Modelling has been used to estimate SOC at different scales. Recent studies have shown high accuracy in estimating SOC, total nitrogen (TN) and other soil properties using LiDAR-generated, high-resolution digital elevation models (DEMs) (Zhou et al., 2020; Zou et al., 2024). Environmental variables have been used to improve prediction accuracy (Mendes & Sommer,

2023). High-resolution multispectral sensors such as Sentinel-2 (S2), have also been used to quantify SOC content (Castaldi et al., 2019; Zhou et al., 2020; Amarnath et al., 2024). Similarly, topography-based Relief Prediction and Classification Systems (RPCS) have proven advantageous for estimating SOC dynamics (Li & McCarty, 2018). Other researchers have fitted models for forest structural variables using LiDAR data and algorithms such as k-nearest neighbors (KNN) or Random Forest (RF) (Yavari & Sohrabi, 2019; Adhikari et al., 2023; Pereira et al., 2023; Strunk & McGaughey, 2023). Rasel et al. (2017) predicted SOC by combining airborne LiDAR metrics (elevation, forest type and aboveground biomass) with the Random Forest algorithm. Navarro Cerrillo et al. (2018) designed a methodology to facilitate silvicultural decision-making in forest management by estimating SOC in *Pinus halepensis* stands in southern Spain using the KNN algorithm. Moreno Muñoz et al. (2024) developed a model using the RF algorithm to estimate SOC in mangrove ecosystems along the southern Pacific coast of Colombia, demonstrating the potential of machine learning techniques to predict edaphic variables in tropical coastal environments. Other studies have highlighted the usefulness of airborne LiDAR data in predicting soil properties. For example, Li et al. (2016) evaluated the effectiveness of LiDAR-derived variables for estimating surface soil horizon properties in a *Pinus koraiensis* forest and demonstrated their usefulness as predictors of soil properties, with coefficients of determination (R^2) ranging from 0.46 to 0.66. Hounkpatin et al. (2021) subsequently compared global and local models for predicting SOC stocks in Swedish forests using national forest inventory data and digital soil mapping approaches. Their findings suggested that local calibration has the potential for higher accuracy. Other studies applied a complex methodology that combined field data with metrics derived from airborne LiDAR to estimate changes in SOC stocks at the stand level, using the Yasso15 model to simulate edaphic carbon dynamics (Strîmbu et al., 2023).

Pinus halepensis Mill. is one of the most widely distributed Mediterranean conifers, covering approximately 3.5 million hectares from the Iberian Peninsula to the eastern Mediterranean region, thanks to its high ecological plasticity and capacity to tolerate semiarid climates and nutrient-poor soils (Derak & Cortina, 2014; Alsanousi et al., 2025). This species acts as an early colonizer of disturbed ecosystems, efficiently regenerating in bare soils with limited water availability (Serrada et al., 2008). Land-use conversion from agriculture to forest through afforestation can generate substantial carbon stocks over several decades (Charro et al., 2008). However, silvicultural practices such as systematic clear-cutting and selective thinning have been shown to influence growth, carbon storage and sequestration dynamics in *P. halepensis* plantations in central Spain (De las Heras et al., 2012; Lull et al., 2024). Altogether, its protective functions, its ability to stabilize soil, its contribution to hydrological regulation and its relevance in providing ecosystem services reinforce the role of *P. halepensis* as key species for restoration in the Mediterranean region (Derak & Cortina, 2014).

The current ecological context requires accurate methodologies for quantifying and monitoring SOC. When integrating LiDAR-derived structural data with soil and forest inventory variables, issues such as scale mismatches, spatial resolution and model generalisability must be addressed, particularly in heterogeneous Mediterranean landscapes. Accordingly, the objective of this study is to model the SOC content in *Pinus halepensis* forests by integrating structural variables derived from airborne LiDAR data with soil and forest inventory data. The aim is to generate SOC estimates that can be applied to forest management in Mediterranean environments and contribute to the development of more accurate and efficient methodologies for soil carbon

monitoring. This aligns with the EU goal of achieving reliable SOC estimates, which constitute a key indicator for assessing forest health.

Accordingly, to advance in the area of SOC estimation, this study is designed around two working hypotheses:

a). LiDAR-derived metrics can be related to SOC stocks in Mediterranean *Pinus halepensis* forests.

b). Machine learning approaches may outperform parametric models in predicting SOC.

1. Materials and methods

2.1 Description of the study area

The study area (Figure 1) is located in Ampudia, (Palencia province, Castilla y León) in northern Spain (41º 51'48'' N; 4º 46'13'' W). Situated at 860 meters above sea level (masl), the area features a sub-Mediterranean continental climate (Dsb) (AEMET, 2024) with an average temperature of 11.3 ºC. The average annual rainfall was 393 mm for the time series from 1995 to 2024. The study site comprised an afforested *Pinus halepensis* stand (60 years old); in its lower stratum, *Quercus faginea* Lam. is sporadically present and clearly subordinate in terms of coverage and dominance. Table 1 shows the main dendrometric and stand variables of the study area (de los Bueis Mellado, 2017).

Table 1. Main dendrometric and dasometric variables obtained from the forest inventory

N (trees/ha)	Ht (m)	Ho (m)	Dbh (cm)	G (m^2/ha)
845	8.5	10.1	17.7	23.20

Note: N = number of trees per hectare; Ht = average total height; Ho = dominant height; Dbh= diameter at breast height; G = basal area

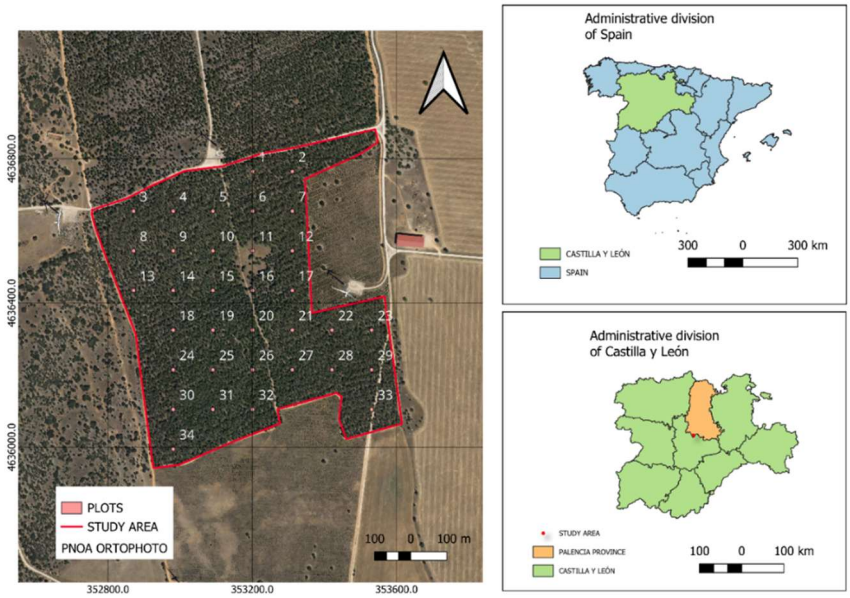

Figure 1. Location and spatial distribution of the study plots within the experimental study area. Note: the plots are overlaid on a PNOA orthophoto using UTM Zone 30N projection and the ETRS89 coordinate system.

2.2 General methodology

Data from three sources (soil sampling data, forest inventory variables and LiDAR metrics) were integrated to model SOC in this study (Figure 2). Soil sampling data parameters were determined for estimating soil carbon stock. The forest inventory provided dendrometric variables to estimate biomass and other tree and stand variables using allometric equations and tree carbon measurements. Finally, processed LiDAR datasets from the Spanish National Aerial Orthophotography Plan (PNOA in Spanish) were used to determine structural stand variables.

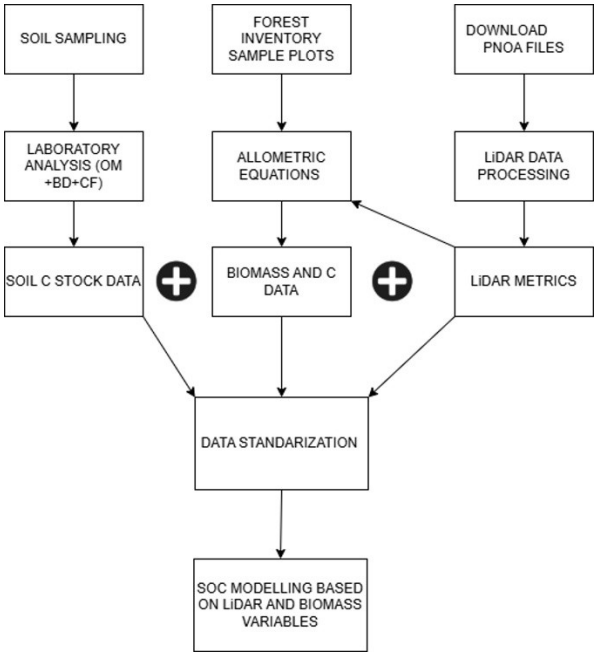

Figure 2. Methodological road map of the study

Note: OM: organic matter; BD: bulk density; CF: coarse fraction; C: carbon

2.3 **Soil sampling and laboratory analyses**

To begin the study, 34 plots were established for soil sampling. A systematic sampling design was applied to the entire study area (46.8 ha). All plots were georeferenced using high-precision, sub-meter GPS to ensure the replicability of the sampling design and facilitate accurate integration with the forest inventory and LiDAR metrics datasets.

Mineral soil samples were taken in each plot at a depth of 0-10 cm. Each plot was divided into four subplots, from each of which 6 – 8 individual randomized samples were taken and composited into one homogenized sample per plot. Samples were transported to the laboratory, where mineral soil samples were air-dried, sieved and the percentage of coarse fraction ([Ø > 2 mm]) determined. In each sample, bulk density was calculated through the core method (Blake & Hartge, 1986) with volumetric steel rings and soil dry weight.

Soil organic matter was determined by the Walkley and Black method (de Vos et al., 2007), based on the partial oxidation of organic carbon with potassium dichromate in an acidic medium. After the reaction, the excess or unconsumed dichromate was titrated with ferrous ammonium sulphate (Mohr's salt) in the presence of diphenylamine as an indicator. The percentage of easily oxidizable organic carbon was calculated from the used volume of this compound and the dry weight of the sample. To determine the total carbon in the soil, the total organic matter was divided by 1.724 (MAPA, 1994) [Eq. (1)].

$$SOC = \frac{SOM}{1.724} \quad (1)$$

147

148 Accumulated SOC_i (Mg C ha^{-1}) [Eq. (2)] was calculated from C concentration, bulk density, soil layer thickness and the
149 percentage of coarse fraction (MAPA, 1994) using the following equation [Eq. (2)] (Lee et al., 2009):

150

$$SOCi = OCi * BDi * (1 - CFi) * ti * 100 \quad (2)$$

151 Note: SOC_i is the C stock (Mg C ha^{-1}) at depth i; i: 0-10cm; OC_i is the organic C content of the soil fraction at depth i; BD_i is the bulk density
152 at depth i (Mg/m^3); CF_i is the volumetric content of the coarse fraction at depth i (%) and t is the horizon thickness of depth i (m).

153

154 All soil analyses were carried out at the ETSIIAA soil laboratory (University of Valladolid).

155 **2.4 Forest inventory and generation of equations for quantifying biomass and carbon**

156 Diameter at breast height (Dbh, with CODIMEX L manual caliper) and total height (Ht, with high-precision GEO Vertex
157 Laser) were recorded for 250 selected trees (Dissanayake, 2024) to develop a site-specific height – diameter equation. Several
158 models were tested (linear, exponential, logarithmic, polynomial and potential) using SAS software (SAS Institute Inc., 2025)
159 to identify the best fit for the Dbh – Ht relationship in the study area. The selected equation was subsequently used to estimate
160 tree biomass (W) using species-specific allometric models for *Pinus halepensis* (Ruiz-Peinado et al., 2011), separated into the
161 following biomass fractions: Ws: stem with bark (commercial volume, up to a maximum diameter of 7 cm), Wmb: medium
162 branches (diameter between 2 and 7 cm), Wthinb: thin branches (diameter smaller than 2 cm) and Wr: coarse roots.
163 Aboveground biomass (Wa) was defined as the sum of the aboveground biomass fractions of all live trees (Ws, Wmb, Wthinb).
164 Tree carbon (C) fixation was determined by multiplying each biomass value by a generic C concentration of 50.0%, based on
165 Kollmann (1959) and the Intergovernmental Panel on Climate Change (IPCC) recommendations (Penman et al. 2003).

166

2.5 ALS data from the PNOA and data processing

The LiDAR-ALS point clouds were obtained from the PNOA program, which provides pre-classified, quality-controlled datasets of validated ground classifications according to national accuracy standards. The main sensor used for collecting ALS data was the Leica ALS80. Technical specifications for the data, gathered in 2019 during flight over the study area, are presented in Table 2.

Table 2. Technical specifications for the second coverage of the PNOA – LiDAR project. Source: PNOA

Minimum point density	0.5 – 2 points/m^2
Year of flight	2019
Geodetic reference system	ETRS89 huso 30 N
RMSE Z	≤ 20 cm
Estimated planimetric accuracy	≤ 30 cm
Simultaneous image	Yes
File size	2 x 2 km
File format	LAS 1.2 format 3
DEM grid spacing	2 m x 2 m
RMSE Z (DEM)	≤ 25 cm
Estimated planimetric accuracy (DEM)	≤ 50 cm

Note: RMSE Z = Root mean square error of the vertical (Z) coordinate; DEM = Digital Elevation Model.

LiDAR data were downloaded from the PNOA portal in .LAS format along with selected map sheets that fully covered the study area (Figure 3). The datasets were spatially cropped using a previously digitized vector layer that precisely delimited the perimeter of the area under study. All LiDAR data was processed at the plot scale, treating each sampling unit independently. Data was processed in the R programming environment (Version 4.4.1; R Core Team, 2024) using a combination of geospatial and LiDAR-specific packages, including lidR, lidRmetrics, terra, and sf (Pebesma, 2016, 2019, 2023; Hijmans, 2025; Tompalski, 2025; Roussel, 2026). The lidR package was used for point cloud manipulation, terrain and canopy model generation, height normalization, individual tree detection, and extraction of forest structural metrics (Roussel et al., 2020; Peter et al., 2021).

The .LAS files were read to extract the point cloud, including spatial reference information and point-level attributes. To optimize memory usage and computational efficiency, only essential attributes (X, Y, Z coordinates, classification, and intensity) were retained. Non-essential points were filtered out and processing was restricted to relevant returns, to reduce data volume.

Ground-related points were then identified using a height-based filter, assuming that returns below 2 m corresponded predominantly to terrain surfaces. This assumption is consistent with field observations in the study area, where understory

vegetation is sparse and consists mainly of *Quercus* regeneration that rarely exceeds 2 m in height. It is important to note that this height threshold was applied exclusively for Digital Terrain Model (DTM) generation and not for vegetation structural analysis.

The DTM was generated using terrain-classified points only, at a spatial resolution of 0.5 m, and using the Inverse Distance Weighting (IDW) interpolation method. IDW was selected for its robustness and consistency when interpolating continuous surfaces from moderately dense point clouds and for its stable performance across plots with variable point distributions. Interpolation parameters were set to k = 10 nearest neighbours and power parameter p = 2, following common practice in geospatial applications (Shepard, 1968; Moussa & Abboud, 2024).

Subsequently, the LiDAR point cloud was height-normalized relative to ground level by subtracting the DTM elevation from all non-terrain returns. From the normalized point cloud, the Canopy Height Model (CHM) was generated at a spatial resolution of 0.5 m using the *pit-free* algorithm, which combines multiple height surfaces to reduce artificial depressions in canopy representation (Khosravipour et al., 2014).

Tree tops were identified from the CHM through a local maxima detection approach. Individual tree crowns were then delineated using the segmentation algorithm developed by Dalponte et al. (2016), which applies a region-growing procedure guided by CHM-derived height peaks. The selected spatial resolution and the LiDAR point density available in the study area allowed for reliable detection and segmentation of individual trees, particularly in the dominant and co-dominant canopy layers.

Once individual trees had been segmented, forest structural variables were extracted at tree level. Metrics were extracted using the lidRmetrics functions in conjunction with the lidR package, following conventional approaches to characterizing canopy structure and biophysical properties from airborne LiDAR data (Lefsky et al., 2002). Extracted variables included height-based and structural descriptors that were subsequently aggregated and used in the SOC modelling framework.

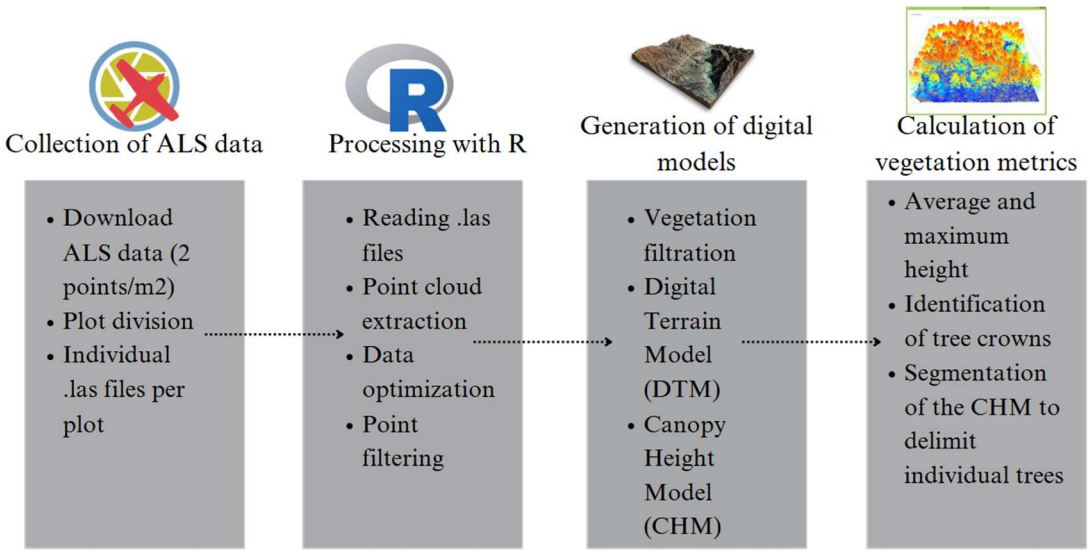

Figure 3. Conceptual map for LiDAR data processing

After individual tree detection and crown segmentation, LiDAR-derived structural metrics were computed and aggregated at the plot level. Extracted variables included descriptive height statistics such as the total number of returns, minimum, maximum, and mean canopy height, standard deviation, coefficient of variation, skewness, and kurtosis (n, zmin, zmax, zmean, zsd, zcv, zskew, zkurt). Height percentiles were calculated from the 1st to the 99th (zq1–zq99), along with return percentages above fixed and relative thresholds (e.g., 2 m, 5 m, and zmean), cumulative return metrics within predefined vertical strata, and indices derived from L-moments (L1–L4, Lskew, Lkurt). Additional metrics were computed based on the vertical leaf area density (LAD) distribution (lad_min, lad_max, lad_mean, lad_sum, lad_cv), along with statistics describing return position within the laser pulse (e.g., n_first, n_last, n_single, n_multiple, p_first, p_last).

All metrics were calculated independently for each plot and subsequently standardized for use as predictors in the SOC estimation models. A detailed description of all LiDAR-derived variables is provided in the Supplementary Material (Tables S.2–S.9).

2.6 SOC modelling and generation of the predicted SOC map

The modeling approach consisted of applying different models and machine learning algorithms to evaluate their capacity for predicting SOC based on soil data, tree biomass variables and LiDAR metrics. Four regression models with different levels of complexity were defined and implemented: a simple linear model, a second-degree polynomial model, a logarithmic model (log(x+1)) and a Random Forest (RF) model (Odebiri et al., 2021; Beisekenov et al., 2025). First, LiDAR predictor metrics with a variance of less than 0.01 were considered non-informative and removed. Collinearity between variables was examined

using the Pearson correlation matrix and those with correlation coefficients greater than 0.9 were removed to reduce redundancy and prevent multicollinearity issues. The remaining variables were normalized through standardization (mean of zero and standard deviation of one).

Second, the processed dataset was randomly divided into two subsets using a fixed random seed: a training subset comprising 75% of the observations and a test subset comprising the remaining 25%. All models were validated using a cross-validation scheme with 10 partitions (k-fold cross validation, v=10) on the training set. For hyperparameter optimization in the Random Forest model, a random search was carried out using 20 different combinations of the mtry (number of predictors considered in each tree division) and min_n (minimum number of observations in a terminal leaf) parameters, while all other parameters (e.g., number of trees, node depth, and sampling scheme) were kept at their default values to limit model complexity. The performance of each model was evaluated using the statistical parameters of mean absolute error (MAE), root mean square error (RMSE) and coefficient of determination (R^2) [Eq. (3), (4) and (5)].

$$MAE = \frac{\sum_{i=1}^{n}|y_i - x_i|}{n} \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - x_i)^2}{n}} \quad (4)$$

$$R^2 = 1 - \frac{\sum(y_i - x_i)^2}{\sum(y_i - \mu_y)} \quad (5)$$

The statistical parameters obtained for each model allowed us to identify the one with the best predictive accuracy. To generate a spatially continuous map of predicted SOC, we implemented a wall-to-wall rasterization approach using the terra package (Hijmans, 2025) in R (version 4.4.1) and. First, the complete dataset of predictor variables was structured as a spatial point object using the corresponding UTM coordinates (EPSG:25830). The area of interest was rasterized at a spatial resolution of 20 × 20 meters, consistent with the spatial density of the LiDAR-derived metrics (Trouvé et al., 2023). A multi-band SpatRaster template was constructed to cover the entire study area. The predictor variables used in the selected model were then rasterized layer by layer over this template. Once rasterized, the selected trained model was applied to the full stack of predictor layers using the predict() function, yielding a continuous surface of SOC values in Mg C/ha. The resulting raster was clipped, using a shapefile to define the limits of the forest stand. The file containing the SOC information for each pixel was exported in GeoTIFF format for visualisation and subsequent spatial analysis. Following the modelling and prediction procedure, the resulting SOC map was laid out using QGIS (version 3.40).

2. Results

3.1 Soil data results

The soil sampling results showed an average SOC of 20.33 Mg C/ha (Table 3) in a range of 9.49 Mg C/ha to 53.16 Mg C/ha. Higher SOC values (Mg C/ ha) were associated with higher concentrations of organic carbon (OC%) and total organic matter (TOM%) while Coarse fraction percentage (CF%) showed broader dispersion, from low values (<1%) in several plots to higher values in others (Table 3).

Table 3. Descriptive statistics of soil variables analysed in the laboratory (n = 34).

	Mean	Standard deviation	Minimum	Maximum	Coefficient of variation (%)
SOC_{0-10}	20.33	9.44	9.49	53.16	46.46
OC (%)	1.50	0.82	0.63	5.08	54.78
TOM (%)	2.58	1.41	1.08	8.74	54.74
BD (g/cm^3)	1.51	0.19	1.05	1.99	12.63
CF (%)	7.23	8.11	0.16	33.68	112.14

3.2 Results of tree biomass estimation

The linear equation was selected as the best Dbh-Ht model (Table 4) for estimating total height.

Table 4. Height diameter models fitted with forest inventory data

Model	Equation	R^2
Linear	$Ht = 0.2731\ Dbh + 4.9865$	0.546
Exponential	$Ht = 6.4408\ e^{0.0233Dbh}$	0.532
Logarithmic	$Ht = 6.6774 \ln(Dbh) - 9.4012$	0.519
Polynomial	$Ht = -0.001\ Dbh^2 + 0.3299\ Dbh + 4.2507$	0.535
Potential	$Ht = 1.837\ Dbh^{0.5778}$	0.542

Meanwhile, results for tree biomass (Table S.1.) showed stem biomass (Ws) to be the dominant aboveground biomass fraction across plots, accounting for an average of 57.2% of total biomass. This was followed by biomass averages of 27.4% for medium branches (Wb2–7) and 15.4% to thin branches (Wb0.5–2). The belowground root biomass (Wr) compartment is then combined with the aboveground fractions to calculate total tree biomass.

3.3 SOC modelling based on LiDAR and biomass metrics

Ten-fold cross-validation on the training set showed higher statistical parameter values for the Random Forest (RF) model (Table 5). The R^2 value (0.811) reflected the explained variance while the RMSE (7.73 Mg C/ha) and MAE (6.13 Mg C/ha) values quantified the prediction error magnitude. The other models explained a good amount of variance but with larger errors (Table 5). In the end, Random Forest showed the best statistical parameters for SOC estimation and was selected as the optimal candidate (Figure 4).

Table 5. Average performance metrics obtained through cross-validation (10-fold) for soil organic carbon (SOC) estimation models

Model	RMSE (Mg C/ha)	MAE (Mg C/ha)	R^2
Random Forest	7.73	6.13	0.811
Logarithmic	331	264	0.871
Polynomial	745	455	0.794
Linear	661	498	0.721

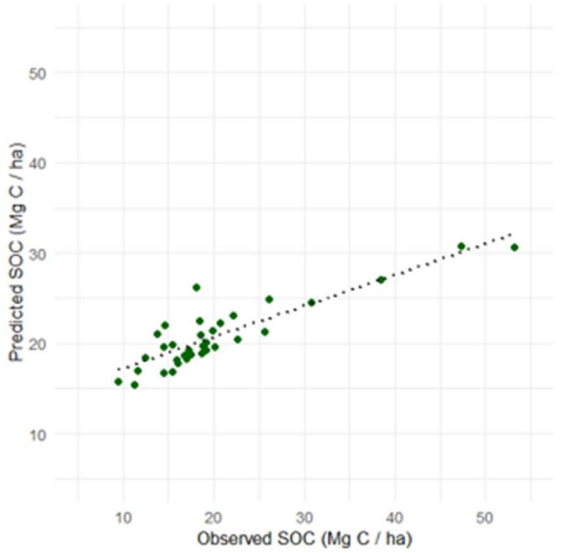

Figure 4. Random Forest model performance: Predicted vs Observed Soil Organic Carbon (SOC) after cross-validation

The RF model made it possible to estimate SOC content for the entire study area. Spatial distribution showed values ranging from 16.2 to 23.5 Mg C/ha, with remarkable variability throughout the forest stand (Figure 5).

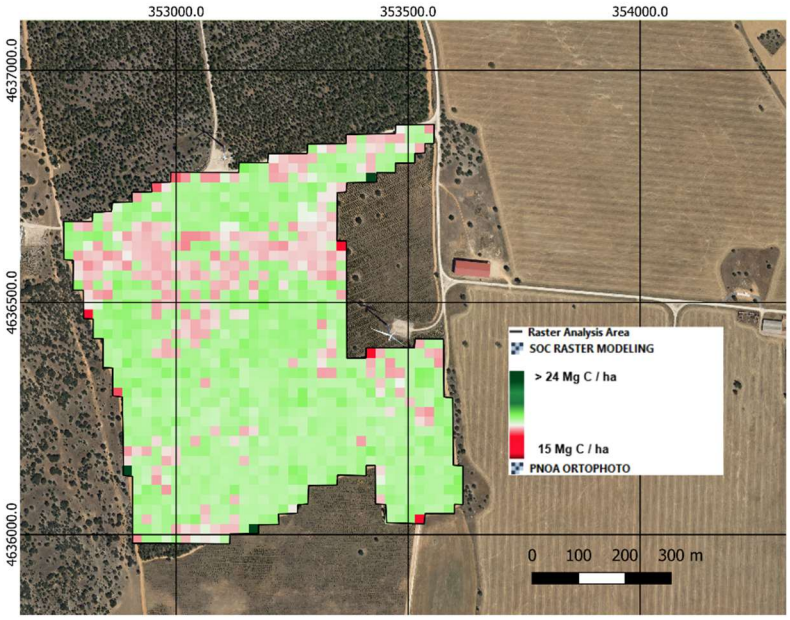

Figure 5. Predicted spatial distribution of soil organic carbon (Mg C/ha) using the Random Forest model

3. Discussion

In this study, a SOC estimation approach was tested in a Mediterranean *Pinus halepensis* ecosystem by applying machine learning techniques to a combination of (a) soil carbon stock data, (b) forest inventory variables and (c) structural metrics derived from LiDAR point clouds processed at plot level. The results provided relevant SOC predictions with potential for scalability to other ecosystems where field sampling is limited but airborne LiDAR data might be available. Our approach could even be applied across broader spatial domains, thus aligning with current efforts to develop harmonized SOC monitoring strategies at the international level.

In this study, the airborne LiDAR data proved effective in capturing structural metrics of the forest canopy by providing valuable proxies for estimating SOC in Mediterranean ecosystems. Models based exclusively on soil properties and conventional forest inventory variables (e.g., bulk density, organic carbon concentration, tree diameter or basal area) rely on point-based measurements that often fail to capture the spatial heterogeneity of forest structure and its influence on soil processes. While these approaches provide valuable local estimates, they have only limited capacity to represent vertical canopy complexity, stand structural variability and biomass spatial distribution patterns. These features indirectly regulate

litter inputs, root dynamics and microclimatic conditions, which affect SOC accumulation (Yang et al., 2024). Such limitations can be partially overcome through the integration of LiDAR-derived structural metrics that provide spatially continuous, three-dimensional information about forest canopy architecture, including height distribution, vertical complexity and canopy density (Olivera et al., 2021; Pascual et al., 2023). Our results indicate that LiDAR metrics, as integrative proxies of vegetation processes, provide helpful information about forest structure that can improve the predictive performance of SOC models and enhance SOC estimation in heterogeneous Mediterranean ecosystems.

We evaluated several SOC estimation approaches in this study and found that the Random Forest (RF) model performed best, achieving a coefficient of determination of 0.811. Beyond its predictive accuracy, this result emphasizes the capacity of non-parametric methods for capturing nonlinear relationships and interactions between structural LiDAR metrics and soil carbon content (Rasel et al., 2017). The associated errors (RMSE = 7.73 Mg C/ha and MAE = 6.13 Mg C/ha) were lower than those obtained in linear, polynomial and logarithmic models. These findings are consistent with those of other studies that have shown more accurate estimates with machine learning techniques for estimating soil properties using LiDAR data (Hengl et al., 2018; Navarro Cerrillo et al., 2018; Hu et al., 2023; Misebo et al., 2024; Alonso-Sarria et al., 2025). Rasel et al. (2017) fitted a model that estimated 69% of SOC based on biomass variables obtained through ALS sensors, while Stumpf et al. (2024), estimated 64% of SOC at a depth of 1.2 m. Navarro-Cerrillo et al. (2018) combined low-density ALS data with nearest neighbor (kNN) models to estimate 82% of SOC at depth of 10 cm in *Pinus halepensis* plantations in southeastern Spain. Navarrete-Poyatos et al. (2019) reported a significant improvement in SOC prediction by combining ALS data with Random Forest. Finally, Pascual et al. (2023) quantified stored carbon by integrating LiDAR metrics with aboveground biomass and litter estimations, highlighting the relevance of incorporating multiple ecosystem data pools in the search for more complete assessment of carbon content.

Our findings aligned with previous research regarding the observed tendency of the Random Forest model to underestimate SOC values, particularly in plots with higher stock values. Agaba (2024) reported that RF underestimated SOC in heterogeneous mountainous landscapes, especially at higher altitudes and in areas with steep slopes. They attributed this pattern to insufficient sampling density in complex surfaces and the limited representation of extreme SOC values in the training data. Similarly, Ou et al. (2024) demonstrated that including site attributes and climatic variables significantly improved SOC prediction in cropland soils and noted how model performance could decline without them.

Model validation in this study was based on cross-validation, considering the maximum available observations for model training. This is a common and accepted strategy in exploratory and proof-of-concept studies. While cross-validation provides a robust internal assessment of model performance, future studies could incorporate independent datasets or multi-site designs. Nonetheless, our results showed that LiDAR-derived structural metrics capture a substantial proportion of SOC variability. Forest variables (tree density, tree basal area, tree biomass, species diversity) are related to atmospheric carbon fluxes and the contribution of litter, roots and exudates (Muñoz-Rojas et al., 2016; Goberna et al., 2007) to the carbon cycle, all of which reflect SOC dynamics. The absence of certain edaphic variables known to influence SOC stabilization processes (e.g., soil texture, pH, biological indicators) in this study reflects a deliberate focus on assessing the explanatory power of forest structural

information derived from LiDAR. Including these in future studies, along with edaphic properties such as microbial dynamics or physical edaphic parameters that may influence SOC stabilization (Muñoz-Rojas et al., 2015; Doetterl et al., 2025), could further refine model predictions.

Integrating remote sensing techniques, such as ALS, with machine learning-based prediction models has proven itself an effective strategy for obtaining a scalable, operational approach to assessing SOC in contexts with extensive national LiDAR coverage. Stevens et al. (2013) integrated site variables derived from digital terrain models with field data to generate SOC predictions in agricultural soils. Soon after, de Brogniez et al. (2015) also used generalized additive models to map SOC based on the Land Use/Cover Area statistical Survey (LUCAS) database. These studies highlighted the usefulness of multivariate approaches that combine environmental variables derived from remote sensors with edaphic information. Our work contributes to this line of research by demonstrating that, even in scenarios with a small number of plots, the structural information obtained by ALS can suffice to generate robust SOC maps at an operational scale. Ballabio et al. (2016) demonstrated use of non-linear modelling techniques, such as adaptive regression splines (MARS), to map physical soil properties. Similarly, our findings could be used to generate future carbon mapping schemes on a larger scale and provide replicable methodologies based on open data and models. This possibility is directly relevant and applicable to European climate neutrality objectives, which prioritize accurate carbon quantification in forest soils (Panagos et al., 2020).

Estimating SOC using data derived from soil sampling, forest inventory variables and LiDAR metrics offers an innovative and non-destructive approach to assessing carbon reservoirs in forest ecosystems. However, this study has certain limitations that should be considered when interpreting the results and planning future research. One limitation is related to the resolution and coverage of the LiDAR data used in the study, which were obtained from the PNOA. We relied on this data and did not undertake independent, ground-based DTM validation. PNOA LiDAR products are generated under standardized national protocols and include validated ground classifications. While these data are available and cover large areas, their point density and temporal resolution may be limited. This affects the accuracy of detail in forest structure characterization and could influence subsequent SOC estimation (Johnson et al., 2022).

The role of *Pinus halepensis* in afforestation and restoration programmes is particularly relevant in semiarid Mediterranean ecosystems, where soil degradation heavily constrains ecological recovery. Although monospecific plantations may initially be associated with reduced understory diversity or slower successional dynamics, this species has shown a remarkable ability to colonise poor soils, stabilise eroded surfaces and initiate soil-improvement processes through organic matter inputs and enhanced structure (Maestre & Cortina, 2004; Chirino et al., 2006). Appropriate management through silvicultural treatments such as thinning, pruning or enrichment plantings can mitigate early structural homogenisation and promote both floristic and functional diversification in restored stands (Navarro et al., 2010; Gordo et al., 2020). Afforestation efforts on eroded marly and gypsiferous slopes since the mid-twentieth century have demonstrated the effectiveness of *P. halepensis* in reducing erosion, facilitating soil restoration and establishing stable tree cover in sites where few species can thrive (Ruano et al., 2022). Current observations indicate low vulnerability to increasing aridity for existing stands in the species' distribution range. In projected climate change scenarios, the expansion potential of *P. halepensis* and its capacity of to sustain key ecosystem

services make it a viable alternative to less drought-tolerant species (Ruano et al., 2022). Beyond its role in recovering degraded areas, the species stands out as an effective carbon sink, with numerous studies demonstrating its capacity to sequester carbon in both biomass and soils (Ruíz-Navarro et al., 2009; López-Senespleda et al., 2021; Santonja et al., 2022). As a species that enhances soil properties, fosters the accumulation of soil organic carbon and facilitates successional processes, *P. halepensis* has great strategic importance in forest planning and the development of resilient afforestation strategies under conditions of increasing aridity.

For estimating carbon concentrations, automated dry combustion (elemental analysis) has been widely regarded as the reference approach in many recent SOC studies and has progressively been replaced by wet oxidation methods. The Walkley–Black method (and its variants) continues to be widely used in large-scale soil surveys, legacy monitoring programmes and soil inventory datasets. It is supported by institutional standard operating procedures and harmonization frameworks (Lettens et al., 2007). Importantly, recent methodological studies have shown that although Walkley–Black is based on incomplete oxidation of soil organic carbon, its results can be made comparable to dry combustion data through appropriate correction or calibration factors and its performance is sufficiently reproducible for routine and inventory-oriented applications (Shamrikova et al., 2022). The ongoing application of the Walkley–Black method is well documented in Mediterranean soil studies, including national-scale assessments and regional databases from Spain, where oxidizable organic carbon measurements have been systematically employed and, as required, harmonized with total carbon data obtained through dry combustion (Calvo de Anta, 2020). When applied consistently across samples and interpreted in a comparative framework, the Walkley–Black method provides internally coherent SOC estimates that are suitable for spatial analyses, modelling approaches and comparative assessments (Lettens et al., 2007; Shamrikova et al., 2022). Moreover, the method relies on standard laboratory procedures and equipment commonly available in soil laboratories, which has contributed to its widespread and sustained use in soil monitoring and inventory contexts, particularly in large-scale or resource-limited studies (El Mdrssa, 2025).

The integration of multiple data sources is recommended to improve future methodological approaches. Combining LiDAR metrics with information obtained from hyperspectral images or radar data makes it possible to capture a wider range of variables related to forest and soil properties, thus strengthening the robustness of the models (Tafur et al., 2022). Integration of multi-temporal or permanent data would also improve the accuracy of the models while facilitating assessment of edaphic carbon stability when disturbances occur (Guillaume et al., 2021).

Conclusions

This study confirms the potential of airborne LiDAR data, combined with estimates of forest inventory data and soil data, for predicting SOC content in *Pinus halepensis* forest ecosystems. The Random Forest model clearly outperformed the parametric models used and showed the best predictive capacity, presenting a coefficient of determination (R^2) of 0.811 and moderate mean errors (RMSE = 7.73 Mg C/ha; MAE = 6.13 Mg C/ha) during cross-validation. The integration of structural metrics derived from available LiDAR data made it possible to efficiently consider the spatial variability of SOC without the need for

intensive field sampling. These results are consistent with previous studies that support the use of machine-learning models for predicting complex soil properties, especially in systems with high structural heterogeneity. However, higher LiDAR data resolution and the inclusion of certain key soil variables could improve the results. Our findings reinforce the usefulness of remote sensing and machine learning tools for estimating carbon stocks in forest soils, providing a methodological approach to support sustainable forest management and environmental monitoring strategies in Mediterranean contexts.

CRediT authorship contribution statement

CHA and DMP: writing —original draft preparation, review and editing, conceptualization, methodology, data curation and formal analysis, supervision. **CHA and DMP** wrote the manuscript. Herrero and Moreno have contributed equally to the paper. **MBT**: writing—review and editing, methodology, supervision. **FLA**: writing—review and editing, methodology, supervision **FBO**: writing—review and editing, methodology, supervision. **IRB**: methodology, supervision. **FTS**: writing—review and editing, conceptualization, data curation and formal analysis, supervision.

Declaration of Competing Interest

We have no competing interests to declare.

Acknowledgements

This study was made possible through funding from the IMFLEX project (PID2021-1262750B-C22) of the State Research Agency; the European Regional Development Fund (ERDF); the Ministry of Science and Innovation (MICINN); and the European Union plan for recovery, transformation and resilience. Thanks to Andrea Blanch for the English revision.

References

Adhikari, A., Montes, C. R., and Peduzzi, A.: A comparison of modeling methods for predicting forest attributes using LiDAR metrics, Remote Sens., 15, 1284, https://doi.org/10.3390/rs15051284, 2023.

AEMET: Agencia Estatal de Meteorología: Mapas climáticos de España (1981–2010) y ETo (1996–2016), Gobierno de España, https://www.aemet.es/es/serviciosclimaticos/datosclimatologicos/atlas_climatico (last access: 23 September 2025), 2024.

Agaba, S.: Mapping soil organic carbon using different machine learning models as an application of digital soil mapping, PhD thesis, Università degli Studi di Milano-Bicocca, Milan, https://boa.unimib.it/handle/10281/512359, 2024.

Alonso-Sarria, F., Blanco-Bernardeau, A., Gomariz-Castillo, F., Jiménez-Bastida, H., and Romero-Díaz, A.: Estimation of soil properties using machine learning techniques to improve hydrological modeling in a semiarid environment: Campo de Cartagena (Spain), Earth Sci. Inform., 18, 323, https://doi.org/10.1007/s12145-025-01833-w, 2025.

Alsanousi, A. A., Abdul-Hamid, H., Mohamed, J., and Masoud, M.: *Pinus halepensis* Mill. in the Mediterranean region: a review of ecological significance, growth patterns, and soil interactions, iForest, 18, 30–37, https://doi.org/10.3832/ifor4566-017, 2025.

Amarnath, A., Dasar, G. V., and Gowda, G. B.: Geospatial mapping of soil organic carbon in Sirsi forest division using remote sensing techniques, Int. J. Res. Agron., 7, 878–884, https://doi.org/10.33545/2618060X.2024.v7.i9sl.1625, 2024.

Aroca-Fernandez, J. M., Diez-Pastor, J. F., Latorre-Carmona, P., Elvira, V., Camps-Valls, G., Pascual, R., and Garcia-Osorio, C.: A collaborative platform for soil organic carbon inference based on spatiotemporal remote sensing data, https://doi.org/10.48550/arXiv.2504.13962, 2025.

Ballabio, C., Panagos, P., and Montanarella, L.: Mapping topsoil physical properties at European scale using the LUCAS database, Geoderma, 261, 110–123, https://doi.org/10.1016/j.geoderma.2015.07.006, 2016.

Beisekenov, N., Banakinaou, W., Ajayi, A. D., Hasegawa, H., and Tadao, A.: Remote sensing-based soil organic carbon monitoring using advanced machine learning techniques under conservation agriculture systems, Smart Agricultural Technology, 11, 101036, https://doi.org/10.1016/j.atech.2025.101036, 2025.

Blake, G. R., and Hartge, K. H.: Bulk density, in Klute, A. (ed.), Methods of Soil Analysis: Part 1. Physical and Mineralogical Methods, 2nd edn., Soil Science Society of America and American Society of Agronomy, Madison, Wisconsin, 363–375, https://doi.org/10.2136/sssabookser5.1.2ed.c13, 1986.

Borsah, A. A., Nazeer, M., and Wong, M. S.: LiDAR-based forest biomass remote sensing: A review of metrics, methods, and assessment criteria for the selection of allometric equations, Forests, 14, 2095, https://doi.org/10.3390/f14102095, 2023.

Breidenbach, J., Ivanovs, J., Kangas, A., Nord-Larsen, T., Nilsson, M., and Astrup, R.: Improving living biomass C-stock loss estimates by combining optical satellite, airborne laser scanning, and NFI data, Can. J. For. Res., 51, 1–14, https://doi.org/10.1139/cjfr-2020-0518, 2021.

Calvo de Anta, R., Luís, E., Febrero-Bande, M., Galiñanes, J., Macías, F., Ortíz, R., and Casás, F.: Soil organic carbon in peninsular Spain: Influence of environmental factors and spatial distribution, *Geoderma*, 370, 114365, https://doi.org/10.1016/j.geoderma.2020.114365, 2020.

Castaldi, F., Hueni, A., Chabrillat, S., Ward, K., Buttafuoco, G., Bomans, B., Vreys, K., Brell, M., and van Wesemael, B.: Evaluating the capability of Sentinel-2 data for soil organic carbon prediction in croplands, ISPRS J. Photogramm. Remote Sens., 147, 267–282, https://doi.org/10.1016/j.isprsjprs.2018.11.026, 2019.

Centro Nacional de Información Geográfica (CNIG): Datos LiDAR del Plan Nacional de Ortofotografía Aérea (PNOA), Ministerio de Transportes, Movilidad y Agenda Urbana, https://centrodedescargas.cnig.es , 2020.

Charro, E., Hernández Navarro, S., Martín Gil, J., Moyano Gardini, A., and Ruiz Potosme, N.: Estimación del secuestro de carbono en suelos bajo masas forestales de *Pinus halepensis* en Castilla y León (España), Cuadernos de la Sociedad Española de Ciencias Forestales, 25, 1–11, https://doi.org/10.31167/csef.v0i25.9671, 2008.

Chevallier, T., Hamdi, S., Gallali, T., Brahim, N., Cardinael, R., Bounouara, Z., Cournac, L., Chenu, C., and Bernoux, M.: Soil carbon as an indicator of Mediterranean soil quality, in Soils and Climate Change: Challenges and Perspectives, IRD Éditions, 627–636, https://books.openedition.org/irdeditions/24018?lang=en, 2016.

Chirino, E., Bonet, A., Bellot, J., & Sánchez, J. R.: Effects of 30-year-old Aleppo pine plantations on runoff, soil erosion, and plant diversity in a semi-arid landscape in south eastern Spain, Catena, 65, 19–29, https://doi.org/10.1016/j.catena.2005.09.003, 2006.

Dalponte, M. and Coomes, D.A.: Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data, Methods in Ecology and Evolution, 7, 1236–1245, 2016.

Dassot, M., Constant, T., and Fournier, M.: The use of terrestrial LiDAR technology in forest science: Application fields, benefits and challenges, Ann. For. Sci., 68, 959–974, https://doi.org/10.1007/s13595-011-0102-2, 2011.

de Brogniez, D., Ballabio, C., Stevens, A., Jones, R. J. A., Montanarella, L., & van Wesemael, B.: A map of the topsoil organic carbon content of Europe generated by a generalized additive model. European Journal of Soil Science, *66*(1), 121–134, https://doi.org/10.1111/ejss.12193, 2015.

de las Heras, J., Moya, D., López-Serrano, F. R., and Rubio, E.: Carbon sequestration of naturally regenerated *Aleppo pine* stands in response to early thinning, New For., 44, 457–470, https://doi.org/10.1007/s11056-012-9356-2, 2012.

de los Bueis Mellado, T.: Relationships between the dynamics of *Pinus halepensis* Mill. and *Pinus sylvestris* L. plantations and environmental parameters: A basis for sustainable management of stands, Tesis doctoral, Universidad de Valladolid, Valladolid, Spain, 2017.

del Río, M., Barbeito, I., Bravo-Oviedo, A., Calama, R., Cañellas, I., Herrero, C., and Bravo, F.: Carbon sequestration in Mediterranean pine forests, in *Managing Forest Ecosystems: The Challenge of Climate Change*, Montero, G., Cañellas, I., and Ruiz-Peinado, R. (eds.), Springer, Dordrecht, 221–245, https://doi.org/10.1007/978-1-4020-8343-3_13, 2008.

Derak, M., and Cortina, J.: Multi-criteria participative evaluation of *Pinus halepensis* plantations in a semiarid area of southeast Spain, Ecol. Indic., 43, 56–68, https://doi.org/10.1016/j.ecolind.2014.02.017, 2014.

de Vos, B., Lettens, S., Muys, B., and Deckers, J. A.: Walkley-Black analysis of forest soil organic carbon: Recovery, limitations and uncertainty, Soil Use Manage., 23, 221–229, https://doi.org/10.1111/j.1475-2743.2007.00084.x, 2007.

Dissanayake, C. T. M.: LiDAR and photogrammetry for automated stem volume estimation at the single-tree level, Master's thesis, Erasmus Mundus Master in Mediterranean Forestry and Natural Resources – MEDFOR, Universidad de Valladolid & Università degli Studi di Padova, 2024.

Doetterl, S., Berhe, A. A., Heckman, K., Lawrence, C., Schnecker, J., Vargas, R., Vogel, C., and Wagai, R.: A landscape-scale view of soil organic matter dynamics, Nat. Rev. Earth Environ., 6, 67–81, https://doi.org/10.1038/s43017-024-00621-2, 2025.

El Mderssa, M., Atlas, S., Ettaqy, A., El Mouridi, Z., Abbas, Y., & Ibijbijen, J.: Comparative evaluation of three methods for soil organic carbon quantification in Morocco's Middle Atlas forests, *Discover Soil,* 2, 98, https://doi.org/10.1007/s44378-025-00127-3, 2025.

495

496 European Commission: EU soil strategy for 2030: Reaping the benefits of healthy soils for people, food, nature and climate
497 (COM(2021) 699 final), European Commission, Brussels, https://eur-lex.europa.eu/legal-
498 content/EN/TXT/?uri=CELEX:52021DC0699, 2021.

499 Goberna, M., Sánchez, J., Pascual, J. A., and García, C.: *Pinus halepensis* Mill. plantations did not restore organic carbon,
500 microbial biomass and activity levels in a semi-arid Mediterranean soil, Appl. Soil Ecol., 36, 107–115,
501 https://doi.org/10.1016/j.apsoil.2006.12.003, 2007.

502 Gordo, J., González, A., Cubero, D., Rojo, L. I., Martínez, C., Finat, L., Hernández, J., & Reque, J.: Actuaciones para aumentar
503 la diversidad en repoblaciones forestales en la meseta Norte, Montes, 139, 26–31, 2020.

504 Guillaume, T., Bragazza, L., Levasseur, C., Libohova, Z., and Sinaj, S.: Long-term soil organic carbon dynamics in temperate
505 cropland-grassland systems, Agric. Ecosyst. Environ., 305, 107184, https://doi.org/10.1016/j.agee.2020.107184, 2021.

506 Gril, E., Laslier, M., Gallet-Moron, É., Durrieu, S., Spicher, F., Le Roux, V., Brasseur, B., Haesen, S., Van Meerbeek, K.,
507 Decocq, G., Marrec, R., & Lenoir, J.: Using airborne LiDAR to map forest microclimate temperature buffering or
508 amplification, Remote Sensing of Environment, 298, 113820, https://doi.org/10.1016/j.rse.2023.113820, 2023.

509 Hengl, T., Nussbaum, M., Wright, M. N., Heuvelink, G. B. M., and Gräler, B.: Random forest as a generic framework for
510 predictive modeling of spatial and spatio-temporal variables, PeerJ, 6, e5518, https://doi.org/10.7717/peerj.5518, 2018.

511 Hernández-Alonso, H., Madrigal-González, J., Tornos-Estupiña, L., Santiago-Rodríguez, A., Alonso-Rojo, P., Morera-Beita,
512 A., and Silla, F.: Tree-size heterogeneity modulates the forest age-dependent carbon density in biomass and top soil stocks on
513 Mediterranean woodlands, Plant Soil, 486, 361–373, https://doi.org/10.1007/s11104-023-05874-2, 2023.

514 Hijmans, R. J.: Spatial data analysis, R package *terra* version 1.8-60, CRAN, https://doi.org/10.32614/CRAN.package.terra,
515 2025.

516 Hounkpatin, K. O. L., Stendahl, J., Lundblad, M., and Karltun, E.: Predicting the spatial distribution of soil organic carbon
517 stock in Swedish forests using a group of covariates and site-specific data, SOIL, 7, 377–398, https://doi.org/10.5194/soil-7-
518 377-2021, 2021.

519 Hu, T., Yu, C., Dou, X., Zhang, Y., Li, G., and Sun, L.: Simulation of soil organic carbon dynamics in postfire boreal forests
520 of China by incorporating high-resolution remote sensing data and field measurement, Fire, 6, 414,
521 https://doi.org/10.3390/fire6110414, 2023.

522 Johnson, L. K., Mahoney, M. J., Bevilacqua, E., Stehman, S. V., Domke, G. M., and Beier, C. M.: Fine-resolution landscape-
523 scale biomass mapping using a spatiotemporal patchwork of LiDAR coverages, Int. J. Appl. Earth Obs. Geoinf., 114, 103059,
524 https://doi.org/10.1016/j.jag.2022.103059, 2022.

525 Khosravipour, A., Skidmore, A. K., Isenburg, M., Wang, T., and Hussin, Y. A.: Generating pit-free canopy height models
526 from airborne LiDAR, Photogramm. Eng. Remote Sens., 80, 863–872, https://doi.org/10.14358/PERS.80.9.863, 2014.

527 Kollmann, F.: Tecnología de la madera y sus aplicaciones, IFIE, Madrid, 1959.

528 Lal, R.: Forest soils and carbon sequestration, For. Ecol. Manage., 220, 242–258, https://doi.org/10.1016/j.foreco.2005.08.015,
529 2005.

Lee, J., Hopmans, J. W., Rolston, D. E., Baer, S. G., and Six, J.: Determining soil carbon stock changes: simple bulk density corrections fail, Agric. Ecosyst. Environ., 134, 251–256, https://doi.org/10.1016/j.agee.2009.07.006, 2009.

Lettens, S., De Vos, B., Quataert, P., Van Wesemael, B., Muys, B., & Van Orshoven, J.: Variable carbon recovery of Walkley–Black analysis and implications for national soil organic carbon accounting. European Journal of Soil Science, 58(6), 1244–1253, https://doi.org/10.1111/j.1365-2389.2007.00928.x, 2007.

Lefsky, M. A., Cohen, W. B., Parker, G. G., and Harding, D. J.: LiDAR remote sensing for ecosystem studies: lidar, an emerging remote sensing technology that directly measures the three-dimensional distribution of plant canopies, can accurately estimate vegetation structural attributes and should be of particular interest to forest, landscape, and global ecologists, BioScience, 52, 19–30, https://doi.org/10.1641/0006-3568(2002)052[0019:LRSFES]2.0.CO;2, 2002.

Li, W., Hu, X., Su, Y., Tao, S., Ma, Q., and Guo, Q.: A new method for voxel-based modelling of three-dimensional forest scenes with integration of terrestrial and airborne LiDAR data, Methods Ecol. Evol., 15, 569–582, https://doi.org/10.1111/2041-210X.14290, 2024.

Li, C., Xu, Y., Liu, Z., Tao, S., Li, F., and Fang, J.: Estimation of forest topsoil properties using airborne LiDAR-derived intensity and topographic factors, Remote Sens., 8, 561, https://doi.org/10.3390/rs8070561, 2016.

Li, X., & McCarty, G. W.: Topographic metric predictions of soil redistribution and organic carbon in cropland fields. Catena, 160, 222–232, https://doi.org/10.1016/j.catena.2017.09.026, 2018.

López-Senespleda, E., Calama, R., and Ruiz-Peinado, R.: Estimating forest floor carbon stocks in woodland formations in Spain, Sci. Total Environ., 788, 147734, https://doi.org/10.1016/j.scitotenv.2021.147734, 2021.

Lull, C., Gil-Ortiz, R., Bautista, I., del Campo, A., and Lidón, A.: The short-term effects of heavy thinning on selected soil carbon pools and microbial activity in a young *Aleppo pine* forest, Forests, 15, 658, https://doi.org/10.3390/f15040658, 2024.

Maestre, F. T., & Cortina, J.: Are *Pinus halepensis* plantations useful as a restoration tool in semiarid Mediterranean areas?, Forest Ecology and Management, 198, 303–317, 2004.

MAPA: Métodos oficiales de análisis, Ministerio de Agricultura, Pesca y Alimentación, Madrid, 662 pp., 1994.

Mayer, M., Prescott, C. E., Abaker, W. E. A., Augusto, L., Cécillon, L., Ferreira, G. W. D., James, J., Jandl, R., Katzensteiner, K., Laclau, J. P., Laganière, J., Nouvellon, Y., Paré, D., Stanturf, J. A., Vanguelova, E. I., and Vesterdal, L.: Tamm review: Influence of forest management activities on soil organic carbon stocks: A knowledge synthesis, For. Ecol. Manage., 466, 118127, https://doi.org/10.1016/j.foreco.2020.118127, 2020.

Mendes, W. de S., & Sommer, M.: The power of integrating proximal and high-resolution remote sensing for mapping SOC stocks in agricultural peatlands, Plant and Soil, 492(1–2), 501–517, https://doi.org/10.1007/s11104-023-06198-x, 2023.

Misebo, A. M., Hawryło, P., Szostak, M., and Pietrzykowski, M.: Spatial estimation of soil organic carbon, total nitrogen, and soil water storage in reclaimed post-mining site based on remote sensing data, Ecol. Indic., 166, 112228, https://doi.org/10.1016/j.ecolind.2024.112228, 2024.

Moreno Muñoz, A. S., Guzmán Alvis, Á. I., and Benavides Martínez, I. F.: A random forest model to predict soil organic carbon storage in mangroves from southern Colombian Pacific coast, Estuar. Coast. Shelf Sci., 299, 108674, https://doi.org/10.1016/j.ecss.2024.108674, 2024.

Moussa, H., and Abboud, M.: The methodology of applying inverse distance weighting interpolation method in determining normal heights, Resourceedings, 4, 1–6, https://doi.org/10.21625/resourceedings.v4i1.1068, 2024.

Muñoz-Rojas, M., Erickson, T. E., Martini, D. C., Dixon, K. W., and Merritt, D. J.: Climate and soil factors influencing seedling recruitment of plant species used for dryland restoration, SOIL, 2, 287–298, https://doi.org/10.5194/soil-2-287-2016, 2016.

Muñoz-Rojas, M., Jordán, A., Zavala, L. M., de la Rosa, D., Abd-Elmabod, S. K., and Anaya-Romero, M.: Impact of land use and land cover changes on organic carbon stocks in Mediterranean soils (1956–2007), Land Degrad. Dev., 26, 168–179, https://doi.org/10.1002/ldr.2194, 2015.

Navarrete-Poyatos, M. A., Navarro-Cerrillo, R. M., Lara-Gómez, M. A., Duque-Lazo, J., Varo, M. de los A., and Rodríguez, G. P.: Assessment of the carbon stock in pine plantations in southern Spain through ALS data and k-nearest neighbor algorithm based models, Geosciences, 9, 442, https://doi.org/10.3390/geosciences9100442, 2019.

Navarro, F. B., Jiménez, M. N., Cañadas, E. M., Gallego, E., Terrón, L., & Ripoll, M. A.: Effects of different intensities of overstory thinning on tree growth and understory plant-species productivity in a semi-arid Pinus halepensis Mill. afforestation, Forest Systems, 19, 410–417, https://revistas.inia.es/index.php/fs/article/download/1463/1373/, 2010.

Navarro-Cerrillo, R. M., Duque-Lazo, J., Rodríguez-Vallejo, C., Varo-Martínez, M. Á., and Palacios-Rodríguez, G.: Airborne laser scanning cartography of on-site carbon stocks as a basis for the silviculture of *Pinus halepensis* plantations, Remote Sens., 10, 1660, https://doi.org/10.3390/rs10101660, 2018.

Norby, R. J., Warren, J. M., Iversen, C. M., Childs, J., Jawdy, S. S., & Walker, A. P.: Forest stand and canopy development unaltered by 12-year elevated CO_2 exposure, but canopy structure strongly influences growth and carbon allocation, Tree Physiology, 42(3), 428–440, https://doi.org/10.1093/treephys/tpab107

Odebiri, O., Odindi, J., & Mutanga, O.: Basic and deep learning models in remote sensing of soil organic carbon estimation: A brief review, International Journal of Applied Earth Observation and Geoinformation, 102, 102389. https://doi.org/10.1016/j.jag.2021.102389, 2021.

Oehmcke, S., Li, L., Revenga, J. C., Nord-Larsen, T., Trepekli, K., Gieseke, F., and Igel, C.: Deep learning based 3D point cloud regression for estimating forest biomass, in *Proceedings of the 29th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (SIGSPATIAL '21)*, Beijing, China, 2–5 November 2021, 284–293, https://doi.org/10.1145/3557915.3561471, 2021.

Oliveira, C. P. D., Ferreira, R. L. C., Da Silva, J. A. A., Lima, R. B. D., Silva, E. A., Silva, A. F. D., Lucena, J. D. S. D., Dos Santos, N. A. T., Lopes, I. J. C., Pessoa, M. M. D. L., and Melo, C. L. S.-M. S. D.: Modeling and Spatialization of Biomass and Carbon Stock Using LiDAR Metrics in Tropical Dry Forest, Brazil, Forests, 12(4), 473. https://doi.org/10.3390/f12040473, 2021.

Ou, J., Wu, Z., Yan, Q., Feng, X., and Zhao, Z.: Improving soil organic carbon mapping in farmlands using machine learning models and complex cropping system information, Environ. Sci. Eur., 36, 9, https://doi.org/10.1186/s12302-024-00912-x, 2024.

Panagos, P., Ballabio, C., Scarpa, S., Borrelli, P., Lugato, E., and Montanarella, L.: Soil related indicators to support agro-environmental policies, European Commission, Luxembourg, https://doi.org/10.2760/011194, 2020.

Panagos, P., van Liedekerke, M., Borrelli, P., Köninger, J., Ballabio, C., Orgiazzi, A., Lugato, E., Liakos, L., Hervas, J., Jones, A., and Montanarella, L.: European Soil Data Centre 2.0: Soil data and knowledge in support of the EU policies, Eur. J. Soil Sci., 73, e13315, https://doi.org/10.1111/ejss.13315, 2022.

Pascual, A., Godinho, S., and Guerra-Hernández, J.: Integrated LiDAR-supported valuation of biomass and litter in forest ecosystems: A showcase in Spain, Sci. Total Environ., 897, 165364, https://doi.org/10.1016/j.scitotenv.2023.165364, 2023.

Pebesma, E.: sf: Simple features for R, R package version 0.9-8, https://doi.org/10.32614/CRAN.package.sf, 2016.

Pebesma, E., Bivand, R.: Spatial Data Science: With applications in R. Chapman and Hall/CRC, doi:10.1201/9780429459016, https://r-spatial.org/book/, 2023

Pebesma, E.: Simple Features for R: Standardized Support for Spatial Vector Data, The R Journal, 10(1), 439–446, doi:10.32614/RJ-2018-009, https://doi.org/10.32614/RJ-2018-009, 2018.

Penman, J., Gytarsky, M., Hiraishi, T., Krug, T., Kruger, D., Pipatti, R., Buendia, L., Miwa, K., Ngara, T., Tanabe, K., and Wagner, F.: Good practice guidance for land use, land-use change and forestry, IPCC/IGES, Hayama, Japan, http://www.ipcc-nggip.iges.or.jp/public/gpglulucf/gpglulucf_contents (last access: 25 October 2024), 2003.

Perea-Ardila, M. A., Andrade-Castañeda, H. J., and Segura-Madrigal, M. A.: Estimación de biomasa aérea y carbono con teledetección en bosques alto-andinos de Boyacá, Colombia, Rev. Cartográfica, 102, 99–123, https://doi.org/10.35424/rcarto.i102.821, 2021.

Pereira, F. F., Sussel, T., Mendes, G., Coelho, J., Magalhães de Andrade, R., Luiz, M., Reiss, L., Fortes, J., Renk, C., Correia, T., Santos, S., Jorge, S., and Simões, C.: Comparison of LiDAR- and UAV-derived data for landslide susceptibility mapping using random forest algorithm, Landslides, 20, 579–600, https://doi.org/10.1007/s10346-022-02001-7, 2023.

Peter, J. S., Drake, J., Medley, P., and Ibeanusi, V.: Forest structural estimates derived using a practical, open-source LiDAR-processing workflow, Remote Sens., 13, 4763, https://doi.org/10.3390/rs13234763, 2021.

QGIS Development Team: QGIS Geographic Information System, version 3.40, Open Source Geospatial Foundation, https://qgis.org, 2024.

R Core Team: R: A language and environment for statistical computing, version 4.4.1, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/, 2024.

Rabot, E., Saby, N. P. A., Martin, M. P., Barré, P., Chenu, C., Cousin, I., Arrouays, D., Angers, D., and Bispo, A.: Relevance of the organic carbon to clay ratio as a national soil health indicator, Geoderma, 443, 116829, https://doi.org/10.1016/j.geoderma.2024.116829, 2024.

Rasel, S. M. M., Groen, T. A., Hussin, Y. A., and Diti, I. J.: Proxies for soil organic carbon derived from remote sensing, Int. J. Appl. Earth Obs. Geoinf., 59, 157–166, https://doi.org/10.1016/j.jag.2017.03.004, 2017.

Roussel, J. R., Auty, D., Coops, N. C., Tompalski, P., Goodbody, T. R. H., Meador, A. S., Bourdon, J. F., de Boissieu, F., and Achim, A.: lidR: An R package for analysis of airborne laser scanning (ALS) data, Remote Sens. Environ., 251, 112061, https://doi.org/10.1016/j.rse.2020.112061, 2020.

Roussel, J., Auty, D.: Airborne LiDAR Data Manipulation and Visualization for Forestry Applications. R package version 4.2.3, https://cran.r-project.org/package=lidR, 2026.

Ruano, I., Calama, R., & Mutke, S.: Fichas de impactos, vulnerabilidad y oportunidades de adaptación al cambio climático para ecosistemas arbolados: Pinares de Pinus halepensis. En: Bravo, F. (ed.), Adaptación al cambio climático: directrices para la adaptación de la gestión del patrimonio natural y la política forestal al cambio climático en Castilla y León, iuFOR – Universidad de Valladolid, 507 pp., 2022.

Ruiz-Navarro, A., Barberá, G. G., Navarro-Cano, J. A., Albaladejo, J., and Castillo, V. M.: Soil dynamics in *Pinus halepensis* reforestation: Effect of microenvironments and previous land use, Geoderma, 153, 353–361, https://doi.org/10.1016/j.geoderma.2009.08.024, 2009.

Ruiz-Peinado, R., Bravo-Oviedo, A., López-Senespleda, E., Bravo, F., and del Río, M.: Forest management and carbon sequestration in the Mediterranean region: A review, For. Syst., 26, eR04S, https://doi.org/10.5424/fs/2017262-11205, 2017.

Ruiz-Peinado, R., del Río, M., and Montero, G.: New models for estimating the carbon sink capacity of Spanish softwood species, For. Syst., 20, 176–188, https://doi.org/10.5424/fs/2011201-11643, 2011.

Santonja, M., Pereira, S., Gauquelin, T., Quer, E., Simioni, G., Limousin, J. M., Ourcival, J. M., Reiter, I. M., Fernandez, C., and Baldy, V.: Experimental precipitation reduction slows down litter decomposition but exhibits weak to no effect on soil organic carbon and nitrogen stocks in three Mediterranean forests of southern France, Forests, 13, 1485, https://doi.org/10.3390/f13091485, 2022.

SAS Institute Inc., : SAS software, version 9.4M8, SAS Institute Inc., Cary, NC, 2025.

Schmidt, M. W. I., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D. A. C., Nannipieri, P., Rasse, D. P., Weiner, S., and Trumbore, S. E.: Persistence of soil organic matter as an ecosystem property, Nature, 478, 49–56, https://doi.org/10.1038/nature10386, 2011.

Shamrikova, E. V., Zonova, V., Kondratenok, B. M., Lu-Lyan-Min, E. I., Tumanova, E. A., Davydova, A. P., Vanchikova, E. V., Libohova, Z., Lapteva, E. M., & Suvannang, N.: Transferability between soil organic matter measurement methods for database harmonization, Geoderma, 412, 115547, https://doi.org/10.1016/j.geoderma.2022.115547, 2022.

Serrada Hierro, R., Montero González, G., and Reque Kilchenmann, J. A. (eds.): Compendio de selvicultura aplicada en España, 2 vols., Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Madrid, 2008.

Shepard, D.: A two-dimensional interpolation function for irregularly spaced data, in *Proceedings of the 23rd ACM National Conference*, ACM Press, New York, 517–524, https://doi.org/10.1145/800186.810616, 1968.

Stevens, A., Nocita, M., Tóth, G., Montanarella, L., and van Wesemael, B.: Prediction of soil organic carbon at the European scale by visible and near infrared reflectance spectroscopy, PLoS ONE, 8, e66409, https://doi.org/10.1371/journal.pone.0066409, 2013.

Scherrer, D., & Körner, C.: Topographically controlled thermal-habitat differentiation buffers alpine plant diversity against climate warming. Journal of Biogeography, 38(2), 406–416, doi: 10.1111/j.1365-2699.2010.02407.x

2011.

669 Strîmbu, V. F., Næsset, E., Ørka, H. O., Liski, J., Petersson, H., and Gobakken, T.: Estimating biomass and soil carbon change
670 at the level of forest stands using repeated forest surveys assisted by airborne laser scanner data, Carbon Balance Manag., 18,
671 22, https://doi.org/10.1186/s13021-023-00222-4, 2023.

672 Strunk, J. L., and McGaughey, R. J.: Stand validation of LiDAR forest inventory modeling for a managed southern pine forest,
673 Can. J. For. Res., 53, 71–89, https://doi.org/10.1139/cjfr-2022-0032, 2023.

674 Stumpf, F., Behrens, T., Schmidt, K., and Keller, A.: Exploiting soil and remote sensing data archives for 3D mapping of
675 multiple soil properties at the Swiss national scale, Remote Sens., 16, 2712, https://doi.org/10.3390/rs16152712, 2024.

676 Tafur, E., Veneros, J., García, L., Gamarra, Ó., Farje, J., and Santistevan, M.: Técnicas no destructivas para la estimación de
677 la biomasa forestal aérea, Idesia, 40, 7–17, https://doi.org/10.4067/s0718-34292022000300007, 2022.

678 Trouvé, R., Jiang, R., Fedrigo, M., White, M. D., Kasel, S., Baker, P. J., and Nitschke, C. R.: Combining environmental,
679 multispectral, and LiDAR data improves forest type classification: A case study on mapping cool temperate rainforests and
680 mixed forests, Remote Sens., 15, 60, https://doi.org/10.3390/rs15010060, 2023.

681 Tompalski, P.: lidRmetrics: Derived LiDAR metrics for airborne laser scanning data (Versión 0.x.x) [Paquete de software R].
682 https://ptompalski.github.io/lidRmetrics/, 2025.

683 Tupinambá-Simões, F., Pascual, A., Guerra-Hernández, J., Ordóñez, C., Barreiro, S., and Bravo, F.: Combining hand-held and
684 drone-based LiDAR for forest carbon monitoring: Insights from a Mediterranean mixed forest in central Portugal, Eur. J. For.
685 Res., in press, https://doi.org/10.1007/s10342-025-01772-7, 2025.

686 Villarino, S. H., Pinto, P., Jackson, R. B., & Piñeiro, G.**:** Plant rhizodeposition: A key factor for soil organic matter formation
687 in stable fractions. Science Advances, 7(16), eabd3176. https://doi.org/10.1126/sciadv.abd3176, 2021.

688 Yang, B., Zhang, X., & Li, Y.: Stand spatial structure is more important than species diversity in enhancing the carbon sink of
689 fragile natural secondary forest, Ecological Indicators, 158, 111449, https://doi.org/10.1016/j.ecolind.2023.111449, 2024.

690 Yavari, F., and Sohrabi, H.: Estimation of available canopy fuel of coppice oak stands using low-density airborne laser
691 scanning (LiDAR) data, in *Advances in Science, Technology and Innovation*, Springer, Cham, 171–173,
692 https://doi.org/10.1007/978-3-030-01440-7_40, 2019.

693 Zellweger, F., Braunisch, V., Baltensweiler, A., & Bollmann, K.: Microclimate buffering by forest structure: implications for
694 climate change vulnerability and adaptation, Journal of Ecology, *107*(5), 2072–2085, 2019.

695 Zhou, T., Geng, Y., Chen, J., Pan, J., Haase, D., and Lausch, A.: High-resolution digital mapping of soil organic carbon and
696 soil total nitrogen using DEM derivatives, Sentinel-1 and Sentinel-2 data based on machine learning algorithms, Sci. Total
697 Environ., 729, 138244, https://doi.org/10.1016/j.scitotenv.2020.138244, 2020.

698 Zou, J., Wei, Y., Zhang, Y., Liu, Z., Gai, Y., Chen, H., Liu, P., and Song, Q.: Remote sensing inversion of soil organic matter
699 in cropland combining topographic factors with spectral parameters, Front. Environ. Sci., 12, 1420557,
700 https://doi.org/10.3389/fenvs.2024.1420557, 2024

701