Remote Sensing of Environment 342 (2026) 115462

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

FOR-age: Benchmarking individual tree age estimation using 3D deep learning on dense laser scanning data

Stefano Puliti [a,*], Binbin Xiang [a], Maciej Wielgosz [a], Eivind Handegard [a], Nicolas Cattaneo [a], Marta Vergarechea [a], Terje Gobakken [b], Juha Hyyppä [c], Erik Næsset [b], Mikko Vastaranta [d], Tuomas Yrttimaa [d], Rasmus Astrup [a]

[a] Norwegian Institute of Bioeconomy Research (NIBIO), Høgskoleveien 8, 1430 Ås, Norway
[b] Faculty of Environmental Sciences and Natural Resource Management, Norwegian University of Life Sciences, Ås, 1432, Norway
[c] Department of Photogrammetry and Remote Sensing, Finnish Geospatial Research Institute, National Land Survey of Finland, Vuorimiehentie 5, FI-02150 Espoo, Finland
[d] School of Forest Sciences, University of Eastern Finland, 80101 Joensuu, Finland

ARTICLE INFO

Edited by: Marie Weiss

Keywords:
Old-growth
Biodiversity
Tree growth
Forestry
Artificial intelligence (AI)
ForestFormer3D

ABSTRACT

Accurately determining the age of individual trees is important for understanding forest dynamics, tree growth, site productivity and describing ecological processes. Traditional methods, such as dendrochronological coring, are invasive, labor-intensive, and costly. This study investigates the use of deep learning (DL) to predict tree age from high-density laser scanning data as a scalable, non-invasive alternative. The dataset includes approximately 1700 tree point clouds from approx. 1 K trees across Norway, Sweden, and Finland, encompassing Norway spruce (*Picea abies)* and Scots pine (*Pinus sylvestris)* and a broad range of tree age and developmental stages, from young seedlings (1 year) to old trees (~350 years). Data were collected using terrestrial, mobile, and high-density airborne laser scanning platforms, enabling the development of sensor-agnostic models. We evaluated multiple modelling approaches, from linear regression to transformer architectures, using both training-from-scratch and fine-tuning strategies. Models fine-tuned starting from pre-trained weights from ForestFormer3D's U-Net as well as the transformer architecture (PointTransformerV3) trained from scratch, proved effective for age regression (RMSE ≤23 years). Although our analysis was limited to two tree species, we demonstrated that a single joint age-estimation model can be successfully trained for both species. We demonstrate that models trained on high-resolution data can generalize to lower-resolution, less costly inputs, provided that data augmentations that mimic reduced resolutions are included during training. This study presents a data-driven framework for estimating tree age without destructive sampling. The findings support the potential for AI-based methods to complement or replace traditional age estimation techniques in forest inventory and monitoring.

1. Introduction

Tree age is a fundamental biometric trait and ecological indicator that provides insight into a tree's physiological state, biodiversity (Wetherbee et al., 2020), forest history, growth dynamics, and recreational value of forests (Blicharska and Mikusiński, 2014; Smith et al., 2024). Despite its significance, accurately determining tree age remains labor-intensive, typically involving the extraction of increment cores or cross-sectional disks, followed by ring counting in the lab. Consequently, tree age is seldom sampled in operational forest inventories and research trials. Even large-scale monitoring efforts, such as national forest inventories, sample tree age only for a small number of trees (Breidenbach et al., 2020). This leaves a significant gap in our understanding of how population-level tree age distributions relate to the functioning and characterization of key forest processes. Further, in forest management, more accurate age information is needed to determine site productivity, predict growth, better schedule thinnings, and also to target nature conservation efforts.

Given its importance, several attempts have been made to predict tree age using tree size and growth conditions. Typically, tree size is

* Corresponding author.
E-mail address: stefano.puliti@nibio.no (S. Puliti).

https://doi.org/10.1016/j.rse.2026.115462
Received 21 July 2025; Received in revised form 21 April 2026; Accepted 26 April 2026
Available online 5 May 2026

correlated with age (Kalliovirta and Tokola, 2005) and allometric models tend to perform reasonably well in managed forests. However, outside this age range and when growing conditions vary significantly across sites, this relationship weakens (Brown et al., 2019). Allowing for nonlinearity and incorporating local growing conditions, such as topographic, geographic, and climatic features, has shown promise for large-scale applications such as national forest inventories (Lu et al., 2025; Rohner et al., 2013). However, capturing such large-scale climatic and geographic trends requires extensive, geographically distributed datasets (e.g., national forest inventories), which are not available for smaller, localized experiments aimed at within-stand age estimation for operational forest management.

Tree architectures display fine-scale structural traits beyond tree diameter at breast height (DBH) or height, such as branch thickness and droopiness, that are related to age (Handegard et al., 2021; Matthes et al., 2008; Weisberg and Ko, 2012). Tree crowns develop over decades or even centuries under the influence of tree genetics, site conditions, and competition, with their morphology reflecting life stage and growth history information (Kellomäki, 2022; Matthes et al., 2008; Piovesan and Biondi, 2021). Handegard et al. (2021) demonstrated that macro-morphological crown traits, such as branch thickness, branch angle, and stem crookedness, explain tree age for species such as Norway spruce (*Picea abies)* and Scots pine (*Pinus sylvestris)*. Traditional field-based crown trait measurements are labor-intensive, prone to substantial measurement errors, and subject to surveyor bias. Until now, gathering crown variables has been a limiting factor. The introduction of point cloud data addresses this challenge.

By taking detailed three-dimensional (3D) snapshots of tree and forest structures, laser scanning provides an objective and scalable way to characterize forest canopies, individual tree structures, and fine-grained individual tree biometrics. In recent years, laser scanning data resolution has increased, data quality improved, and there is now more versatility in data capture platforms. For example, in forest measurements, depending on the intended use, laser scanning data are often collected within the forest using either terrestrial laser scanning (TLS) or mobile laser scanning (MLS). The difference between these methods is that in TLS, the scanner remains stationary during the measurement, whereas in MLS, data are collected while the scanner is in motion. Nowadays, it is also possible to collect very high-density point cloud data from above the canopy—using aircraft, helicopters, or drones—referred to as high-density airborne laser scanning (ALSHD). TLS and MLS are well-suited for data collection on sample plots or in individual forest stands. ALSHD data can also be collected from sample plots, but increasingly also at the landscape level, and, as technology advances, over broader areas. The availability of laser scanning sensors has also increased significantly, leading to lower prices. This has enabled researchers and practitioners to collect high-point-density datasets with a relatively small investment (Výbošťok et al., 2026). As robotics and artificial intelligence continue to advance, laser scanners are poised to become ubiquitous in forestry. Such development, however, rests on the development of accurate and robust models that can provide access to fine-grained forest and tree biometric information.

Over the past couple of years, the integration of deep learning (DL) models into forest remote and proximal sensing has significantly advanced the state-of-the-art in forest and tree 3D vision. Advances in individual tree segmentation models, such as ForAInet (Xiang et al., 2024), SegmentAnyTree (Wielgosz et al., 2024), or TreeLearn (Henrich et al., 2024), are expanding the capabilities of the research community and operational actors to automate individual tree segmentation beyond what was achievable with classical methods. Simultaneously, a growing body of literature highlights how individual tree point clouds can effectively be utilized to derive fine-grained tree biometrics, including stem curve (Hyyppä et al., 2020), tree species (e.g., Allen et al., 2023; Puliti, 2025; Seidel et al., 2021), tree growth (e.g., Ahmed et al., 2025; Ahmed and Pretzsch, 2023; Puliti et al., 2024b; Puliti et al., 2023), wood quality (e.g., Cattaneo et al., 2024; Pehkonen, 2025; Pyörälä et al., 2019), traceability (Yrttimaa, 2024), and even biodiversity characteristics such as tree microhabitats (Fol et al., 2023; Rehush et al., 2018) and bird diversity (Holland et al., 2024). Collectively, these studies underscore the strong link between tree 3D architecture, fine-grained tree biometrics, and wood internal properties.

Although dense point clouds have proven suitable for measuring and estimating several key tree attributes, their use for determining tree age has received little to no research attention. Key reason for this has been the lack of costly datasets linking individual tree point clouds to corresponding tree ages. Further, there have also been no clear ideas on how metrics derived from point clouds could explain tree age beyond basic allometry. Thus, whether architectural signals of aging detectable in 3D structure can provide independent predictive value remains unexplored.

To address the data gap, we introduce a new downstream task for forest 3D computer vision and present the FOR-age dataset – the first of its kind to enable the modelling of individual tree age from 3D laser scanning data (i.e. TLS, MLS, ALSHD). This study evaluates a range of modelling approaches, from linear regression to transformer-based deep learning architectures, and compares species-specific and global models for *Picea abies* and *Pinus sylvestris*. Furthermore, we assess model robustness under reduced-resolution point cloud scenarios, simulating more efficient (i.e. cheaper) data acquisition protocols. By enabling age prediction at the level of individual trees, stands, or entire forested landscapes, this research lays the groundwork for integrating tree age distributions into forest monitoring, offering new insights into forest growth dynamics, site productivity, and biodiversity.

2. Materials

2.1. Study areas

The FOR-age dataset was compiled as a multi-project effort over three years. Although the dataset has been compiled from measurements across several projects, age information and high-resolution laser scanning data are available for each tree. These were conducted as parts of various efforts in Norway, Sweden, and Finland between 2023 and 2025 (see Fig. 1). These campaigns were conducted as part of six independent efforts (i.e., Handegard et al., 2021; Puliti et al., 2024a; Puliti et al., 2023). Although each study relied on a systematic or random sample plot design, they did not follow a consistent sampling approach. Instead, the selection of trees was aimed at covering a broad range of tree ages. Due to the very high costs for collecting a dataset composed of multi-modal laser scanning labelled with individual tree age, we limited our study to two of the most common forest species in Europe, *P. sylvestris* and *P. abies*, spanning from boreal forests in Fennoscandia to temperate central and eastern Europe and, in the case of *P. sylvestris*, even in Mediterranean forests and thus relevant both for economic and ecological aspects. These represent the main species in the Boreal and subalpine conifer forests, from Central Europe (in mountains) to Northern and Eastern Europe up to the Ural Mountains (de Rigo et al., 2016). The forests in the study areas included traditionally managed even-aged boreal forests (Skar, Evo, LongTerm, Valer) as well as areas that have never been managed (Handegard2021) or that have been managed using non-mechanized individual tree selective harvesting and including multi-story structure (Handegard2021, Lillomarka, PathFinder).

A summary of the main characteristics of each dataset in the FOR-age dataset is provided in Table 1.

2.2. Age measurements

Collecting individual tree age and corresponding point clouds is logistically challenging and costly, and developing a sufficiently large dataset for model training and evaluation requires integrating data from multiple sources. The FOR-age dataset includes age measurements (years at breast height) and tree point clouds obtained through various

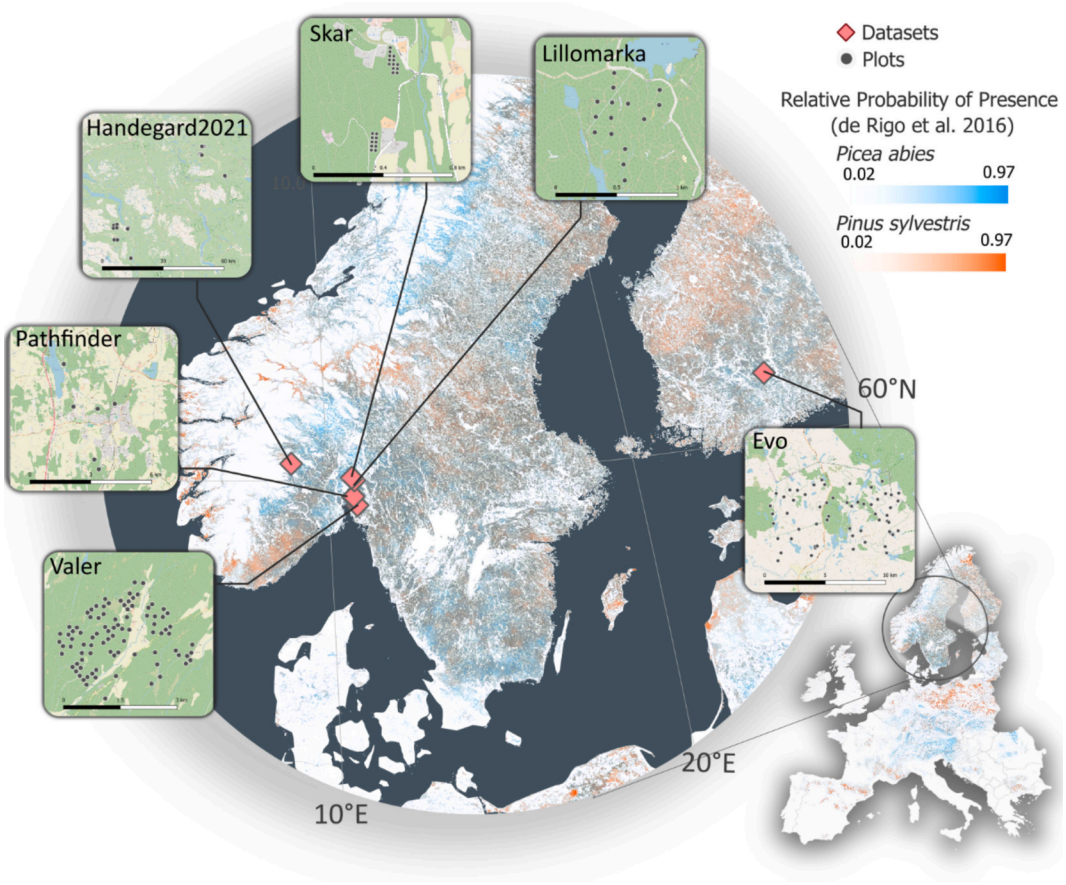

Fig. 1. Geographical overview of the different datasets composing the FOR-age dataset.

Table 1
Summary statistics for the different datasets included in the FOR-age dataset.

Dataset name	Country	n plots	n trees	n tree point clouds	Age range (years)	Age mean (years)	Laser scanning modality
Handegard2021	Norway	12	44	44	70–348	199	MLS
Skar	Norway	18	41	41	43–70	58	ALSHD
Lillomarka	Norway	15	133	266	18–223	67	MLS + ALSHD
PathFinder	Norway	4	41	57	8–209	86	MLS + ALSHD
Evo	Finland	58	335	667	28–175	71	TLS + ALSHD
Longterm	Norway	1	25	25	67	67	MLS
Valer	Norway	90	371	674	1–48	18	MLS + ALSHD
TOTAL		**199**	**992**	**1775**	**1–348**	**53**	

protocols aimed at estimating tree age:

i) Invasive sampling (Handegard2021, Skar, Lillomarka, Evo, Pathfinder) was conducted by either boring tree cores or by full tree harvest and sampling a stem disk at breast height (1.3 m); This was done for 61% of the sampled trees.
ii) Manual labelling of the number of whorls from point cloud data (Valer) in 1–48 years old trees that had visible whorls; This was done for 36% of the sampled trees.
iii) Age from planting year (LongTerm) in even-aged forest stands for which accurate information on the planting year was available; This was done for 2% of the sampled trees.

Although no dedicated reference dataset was available to validate the tree-age estimates used in this study, published work indicates that such measurements can carry uncertainties ranging from a few years to several decades (Norton and Ogden, 1990; Palik and Pregitzer, 1995). These uncertainties arise not only from general measurement error but also from limitations inherent to the specific protocols used. For example, increment coring can yield inaccurate ages when the pith is missed, and whorl counting is known to be unreliable in suppressed trees that lose branches or fail to form visible whorls—issues that may further contribute to protocol-specific errors. Such factors may influence model performance; however, in the absence of alternative age-estimation methods that are both more accurate and feasible at large spatial scales, these measurements were the best achievable given the practical constraints of this study.

High quality individual tree point clouds were manually segmented. Overall, the FOR-age dataset comprises 1775 tree scans (i.e. each single tree is sometimes scanned with both ground and airborne based scanners), a total of 581 individual *P. abies* and 411 *P. sylvestris* spanning a 350-year range and with an average of 53 years (see Table 1). The dataset is characterized by a long tail distribution (see Fig. 2) including seedling and saplings (6% of trees <10 years old), established forests (50% of the trees are between 10 and 50 years old), forest mature for harvesting (28% of the trees are between 50 and 100 years old), old trees

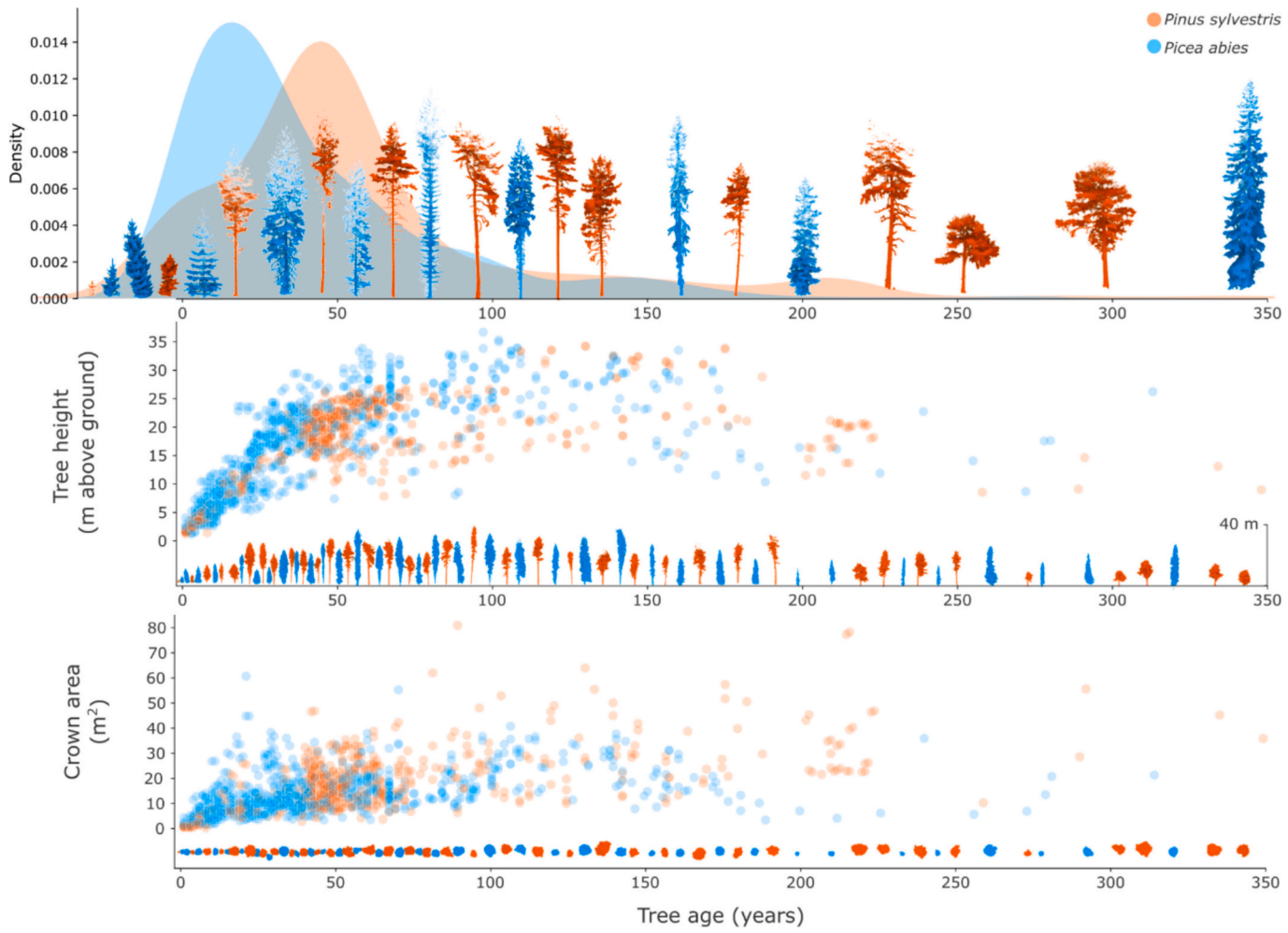

Fig. 2. Age distributions within the FOR-age dataset are divided by species and age class, and their relationships with tree height and crown area.

(9% of the trees are between 100 and 200 years old) and with only a small portion of very old trees (2% of trees older than 200 years old). The observed distribution partly results from our sampling strategy aimed at capturing a broad range of tree ages and also resembles patterns found in many European forests, where historical overexploitation and later management have reduced the occurrence of old trees, now mostly scattered across the landscape (Gold et al., 2006). It also reflects the natural disturbance regimes that shape age distributions differently across species such as *P. abies* and *P. sylvestris*. Shade-tolerant spruce often exhibits a reverse-J distribution due to gap dynamics, while pine typically regenerates in cohorts following fire events and tends to live longer. (See Fig. 3.)

2.3. Point cloud data capture and segmentation

The laser scanning data were captured during several field campaigns and varying laser scanning sensors, scanning protocols falling within the following categories:

- Ground based:
  - o **MLS data**: captured using Geoslam ZEB-Horizon (FARO, 2025b) in 250 m^2 circular (Handegard2021, Lillomarka, Valer) or 2500 m^2 square plots (PathFinder, Longterm) based on an irregular pattern aimed at covering the full area of the plot while also ensuring loop closure. Further details on the specific data acquisition protocols can be found in Puliti et al. (2023) and Puliti et al. (2024a). The raw GeoSLAM ZEB-Horizon data were processed into point clouds using the simultaneous FARO Connect2 (FARO, 2025a)
  - o **TLS data (Evo)**: Terrestrial laser scanning data were captured in May 2023 using a Leica RTC360 time-of-flight laser scanner. The scan design consisted of nine scan positions evenly distributed across 1024 m^2 square plots. Artificial reference targets were used

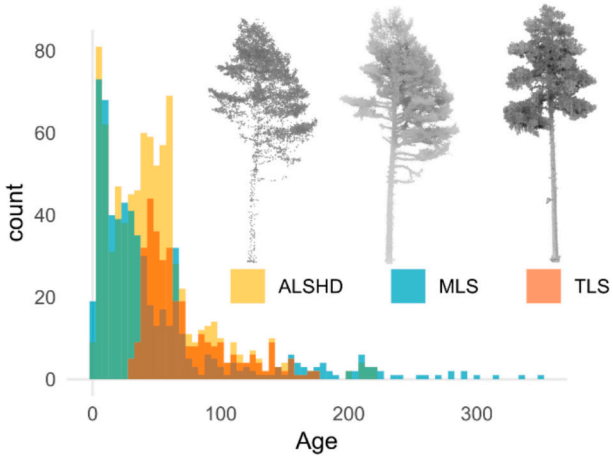

Fig. 3. Distribution of the FOR-age individual tree point clouds by laser scanning modality.

to co-register individual scans at the sample plot level in the Register360 software provided by the scanner manufacturer. The resulting scanner-local point clouds were georeferenced against RTK GNSS-measured tree maps by identifying corresponding local-global coordinate pairs for 5–8 trees per sample plot and computing a rigid transformation (XY-translation and *Z*-axis rotation) accordingly.

- High density airborne laser scanning (ALSHD)
  - o **ULS data** (Skar, PathFinder): captured using a multirotor unmanned aerial vehicle (UAV) equipped with a Riegl VUX1-UAV and MiniVUX-1UAV (Riegl, 2025a). The ULS data were collected at 60–120 m above ground level (AGL), following a double crosshatch pattern, with large lateral overlap (90%) and a wide field of view (FOV = 140°). Further details on the ULS data in this study can be found in Puliti et al. (2020); Puliti et al. (2023).
  - o **Heli-ALS** (Lillomarka, Evo): captured using helicopters with various sensors onboard. The Lillomarka dataset was captured using a Riegl VUX-240 (Riegl, 2025c) at a flying height of 200 m AGL, with a flight speed of 25 m sec^{-1}, a FOV of 75°, at a pulse rate of 1800 kHz, a scan rate of 350 lines per second (LPS), and with a 80% lateral overlap resulting in an average density of 1500 returns m^{-2}. The Evo dataset was captured in June 2023 using a three-sensor system consisting of a Riegl VUX-1HA (wavelength of 1550 nm, a scan rate of 119 LPS, and a FOV of 360° FOV), a Riegl MiniVUX-1DL (905 nm, 120 LPS, 46°), and a Riegl VQ-840-G (532 nm, 92 LPS, 40°) laser scanners. For accurate positioning and orientation, the system was equipped with a NovAtel Pwrpak7 GNSS receiver and a NovAtel ISA-100C IMU. The system was carried by a helicopter that flew over the sample plots four times—north-to-south, east-to-west, south-to-north, and west-to-east—at an altitude of 100 m and a flight speed of 14 m sec^{-1}. This resulted in an average surface density of approximately 1200 returns m^{-2}. The trajectory data was post-processed using a Waypoint Inertial Explorer provided by the sensor manufacturer with base station measurements for differential correction. The point cloud data was processed in RiProcess software, including a boresight adjustment step to adjust sensor alignments for accurate point cloud alignment between sensors.
  - o **Fixed-wing ALS (Valer)**: Airborne laser scanning data were captured for parallel flight strips using a manned fixed wing aircraft mounted with a dual-channel Riegl VQ-1560 II S (Riegl, 2025b), flown at 940 m AGL, with a flight speed of 62 m sec^{-1}, a FOV of 58.52°, a pulse rate of 2000 kHz, a scan rate of 267 LPS, and with an 81% lateral overlap resulting in an average density of 700 returns m^{-2}. Such acquisition, due to the higher altitude and higher flight speed, represents an interesting alternative for improving scalability compared to UAV or even helicopter laser scanning data collection.

Because for most of the trees (i.e. 84%) MLS or TLS and ALSHD data were fully co-registered, the FOR-age dataset is sensor- and platform-agnostic, allowing for the development of methods that address variations in point cloud density, occlusions and quality.

2.4. Data split

In this study, the dataset was partitioned into training, validation, and test sets using stratified sampling to ensure balanced representation across different age classes and dominant tree species. To ensure spatial separation among the training, validation, and test sets, the split was performed at the plot level, with the dominant tree species defined as the one with the most trees within each plot. The average tree age per plot was computed and binned into 10-year-wide bins. The dominant species and age class were merged to generate a new stratification column, which was then used to split the data into training, validation, and test sets in the proportions of 70%, 15%, and 15%, respectively. This stratified approach ensured that each subset maintained the same distribution of age classes and dominant species, thereby preserving the heterogeneity of the original dataset.

2.5. GPS-time augmentation

To boost model transferability to sparser point clouds, we adopted a similar augmentation strategy to SegmentAnyTree (Wielgosz et al., 2024), in which point clouds are sparsified and used to improve the model's transferability to sparser and thus more efficiently captured datasets. Unlike Wielgosz et al. (2024), who achieved this by randomly downsampling the original point clouds, in this study, we focused on a more realistic augmentation strategy that leverages GPS-time values available for most point clouds. GPS-time represents the time (seconds) at which a laser pulse was emitted and can therefore be used to separate different passes over the same tree during the acquisition time. Using this information, we divided trees with available GPS-time data to simulate the following more efficient data capture strategies (see Fig. 4): i) Individual pass, ii) Every other pass, and iii) Every third pass.

3. Methods

3.1. Model development

In this study, we benchmarked several model architectures - from linear regression to 3D state-of-the-art deep learning models- and modelling strategies -, training from scratch or fine-tuning - with increasing degree of complexity and computational needs, including the following strategies:

- ***Scratch-LR*: Training from scratch a linear regression model** to provide a baseline of the simplest approach, as suggested by Kalliovirta and Tokola (2005), we fitted a linear regression model using tree height (m) and crown projection area (m^2) as predictor variables.
- ***Scratch-PT3***: Point Transformer v3 (PT3) (PT3; Wu et al., 2024) model was trained from scratch (i.e. randomly initialized weights) for direct age regression using only 3D coordinates (X, Y, Z) as input. The model employed a five-stage encoder–decoder backbone (PT-v3m1) with progressively increasing channels (32 to 512) and attention heads (2 to 32), using randomized point orderings (*Z*-order and Hilbert curves) and a 512-dimensional embedding. Training was conducted over 300 epochs using AdamW (lr = 1e-4, weight_decay = 0.05) with a linear PolyLR schedule and a weighted mean squared error loss to mitigate age-class imbalance. To improve generalization, extensive point cloud augmentations were applied: normalization, random rotation around the vertical axis (±1 rad), anisotropic scaling (0.7–1.5×), dropout of up to 20% of points, spatial shifts (±0.2 m), voxel-based sampling (1 cm grid), and point shuffling. Hyperparameter tuning was done using the validation set.
- ***FineTune-FF3D*:** To get an understanding of the possibility of using an existing state-of-the-art forest panoptic segmentation model as a backbone to fine-tune a model head targeted at the prediction of tree age rather than tree masks. Assuming that the encoded information from the panoptic forest model regarding semantics (e.g. wood/leaf separation) and individuals is also relevant for estimating tree age, this approach can leverage existing knowledge of forest 3D structures to boost performance and reduce the need for labelled data. In this study, we relied on ForestFormer3D (Xiang et al., 2025), the latest state-of-the-art forest panoptic 3D segmentation model. Such a model relies on a transformer architecture and has proven to address advanced tasks such as the instance segmentation of trees into very complex canopy structures and to deal with the multi-scale nature of trees (i.e. transferable to trees of all sizes). We implemented age regression with ForestFormer3D by keeping its 3D sparse U-Net backbone, removing the transformer decoder, and attaching a

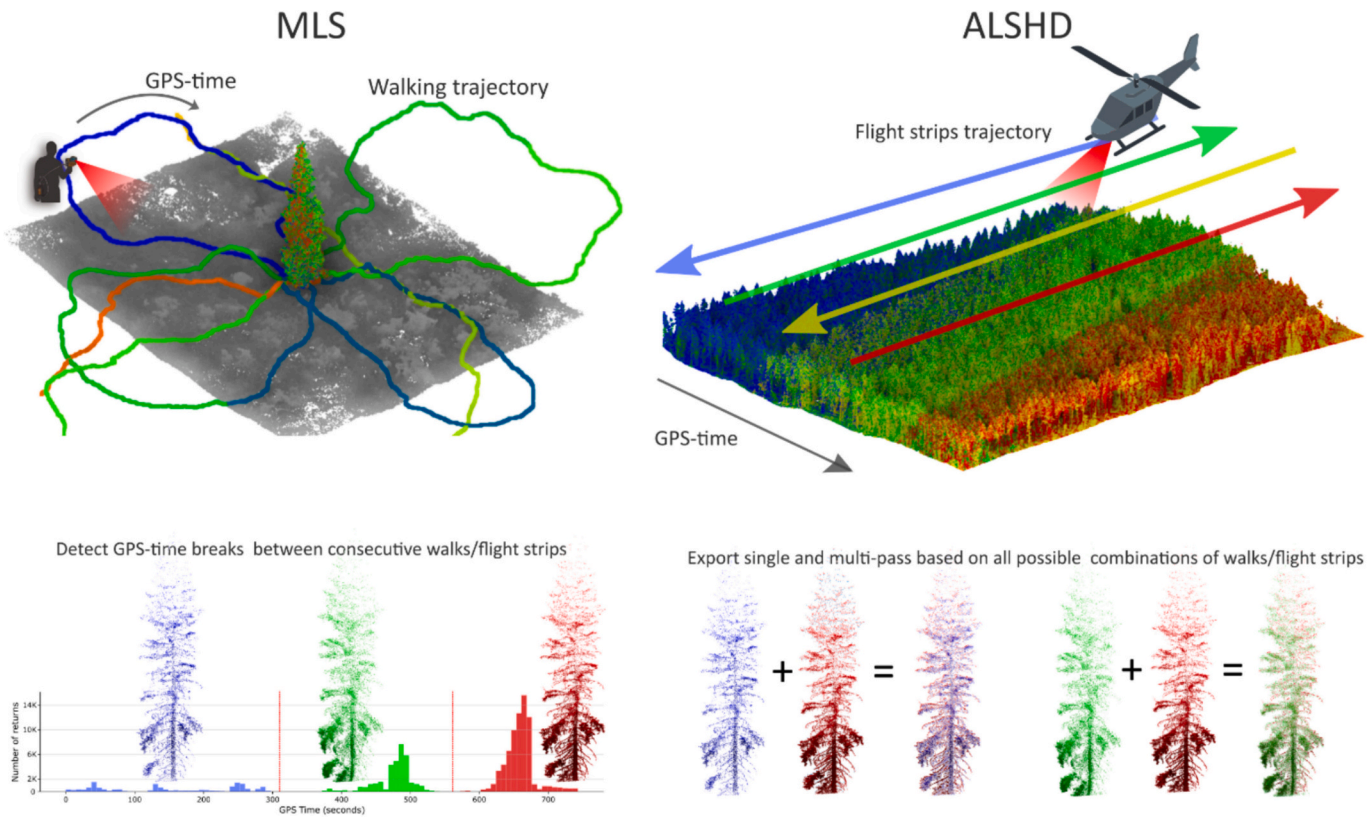

Fig. 4. Schematic illustration of the GPS-time augmentation for both MLS and ALSHD point clouds. Trajectories are color-coded by GPS-time, with cooler colors indicating the start of data collection and warmer colors indicating the end.

regression head that receives the 3D sparse U-Net backbone's output features and predicts a single tree age value. The backbone was initialized with weights from pre-training on a forest panoptic segmentation task and is fine-tuned jointly with the regression head. Training is performed on 0.05 m voxels for 1200 epochs using AdamW, a cosine learning rate schedule (initial learning rate = 0.0001), and weight decay of 0.05. Random data augmentation includes horizontal flips, rotations around the vertical axis, isotropic scaling between 0.8 and 1.2, and Gaussian point jitter. As for *scratch_PT3*, the hyperparameters were tuned on the validation set. A weighted mean squared error (MSE) loss is used. Each age is weighted by the inverse square root of its sample frequency to mitigate class imbalance. The model with the smallest RMSE on the validation split is used for the final prediction.

All experiments were conducted on a single NVIDIA A100-SXM4-80GB GPU. The training stage used a batch size of 64 with 6 data-loading workers and required approximately 31 h in total.

Fig. 5 illustrates a schematic workflow of this study from data collection to method evaluation.

3.2. Model evaluation

The model evaluation aimed to benchmark the above-described strategies, providing insight into the potential for modelling tree age across species and laser scanning modalities. The model evaluation was performed by applying the different models to the test data and computing the root mean squared error (RMSE) as a measure of accuracy.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (\hat{y}_i - y_i)^2}{N}} \quad (1)$$

The mean difference (MD) was used as a measure of systematic errors in the models' predictions and computed as

$$MD = \frac{\sum_{i=1}^{N} \hat{y}_i - y_i}{N} \quad (2)$$

where $\hat{y}_i$ and y_i are the predicted and reference age for the i^{th} tree, and N is the total number of trees in the test data. The model residuals and predictions were further visually assessed.

In addition, as a measure of the model's ability to explain the variation of the measured age, we reported the coefficient of determination (R^2) computed as:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2} \quad (2a)$$

where $\bar{y}$ is the average of the reference age values.

For the best performing method, we further analyzed the performance by age class. Further, we assessed the possibility of developing a joint-species model and compared it with species-specific models. Additional evaluation focused on the model's ability to generalize to sparser point clouds, obtained through more efficient data capture trajectories. This has important implications for scalability (i.e. the sparser the data, the larger the scale). The GPS-time sliced test data were used to evaluate performance across various scenarios, including single walk/flight-line, every-other, every-third, and so forth.

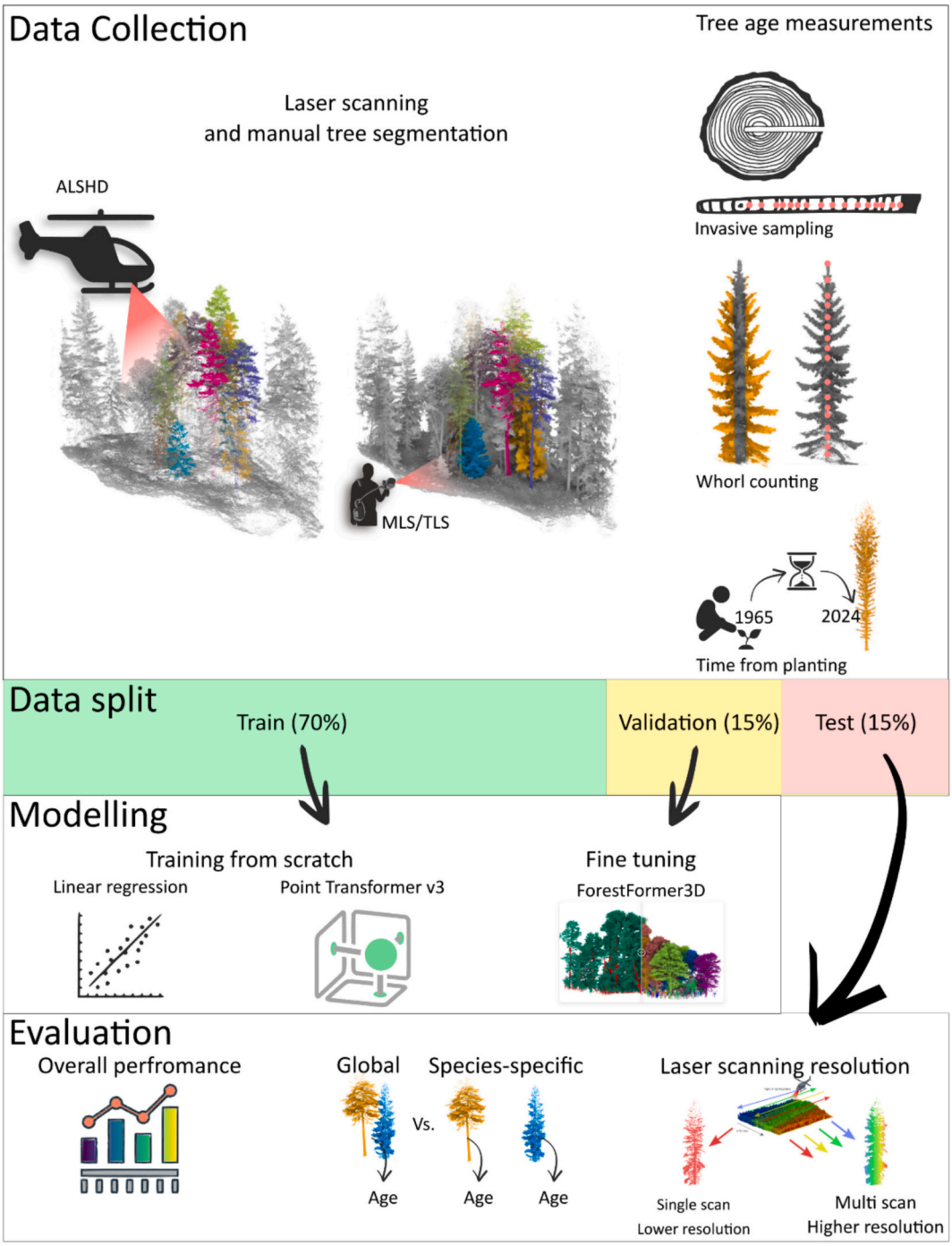

Fig. 5. Schematic workflow of this study from data collection to method evaluation.

4. Results and discussion

4.1. Comparison of deep learning architectures and training strategies

From the overall performance comparison of the different modelling scenarios against the withheld test data, the *fineTune_FF3D* model emerged as the best model (RMSE = 21 years; MD = 2 years; $R^2 = 0.74$), closely followed by *scratch_PT3* (RMSE = 23 years; MD = -4 years; $R^2 = 0.68$). *fineTune_FF3D* showed particularly improved accuracy in young and old forests (see Fig. 6) and reduced systematic errors (i.e. *MD*) by about half (see Fig. 6). This highlights that state-of-the-art point cloud transformer architectures (i.e. PointTransformer v3) can effectively capture fine-grained tree biometric traits even when trained from scratch. Further, this study shows the possibility to effectively build upon FF3D as forest-specific pretrained 3D backbone for tasks extending beyond its original design, underscoring its versatility.

Unlike the low RMSE (6 years) reported by Kalliovirta and Tokola (2005), our linear regression model based on tree height and crown area performed poorly (RMSE = 34 years) and saturated at approximately 100 years, limiting its use for both forest management and conservation. These findings align with the widespread pattern that tree age and tree size are poorly or even negatively correlated for old trees (Brown et al., 2019; Castagneri et al., 2013; Piovesan and Biondi, 2021).

Interestingly, the ability of our best model (i.e. *fineTune_FF3D*) to describe age variability ($R^2 = 0.74$) was greater than previously reported when using non-invasive human interpretation and measurement

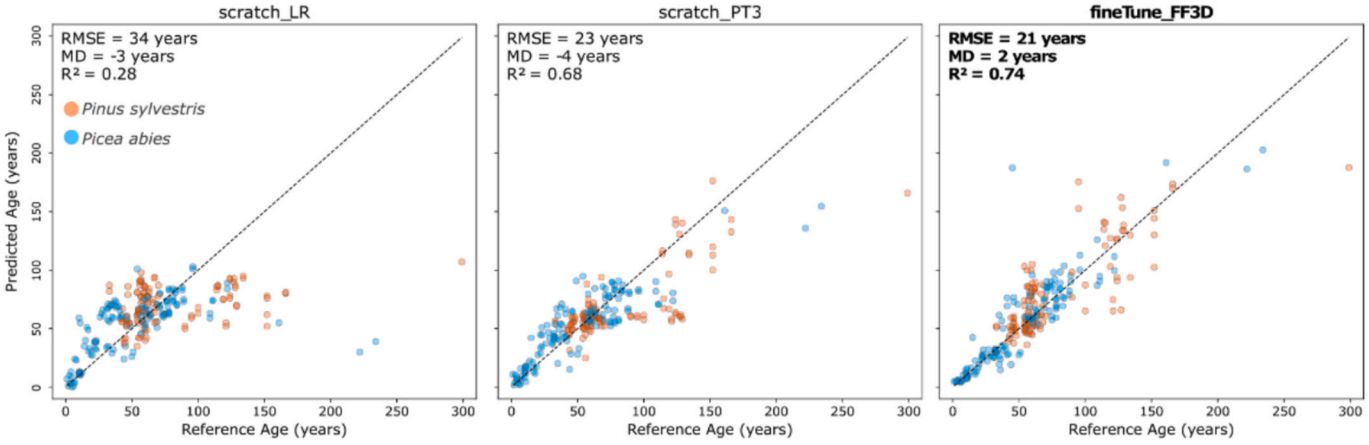

Fig. 6. Scatterplots of the predicted vs reference age values for all the benchmarked methods, including the dashed 1:1 line.

methods. Previous models to predict tree age for European and North American conifers without tree cores have reported maximum R^2 of 0.5–0,7 (Brown et al., 2019; Handegard et al., 2021; Matthes et al., 2008; Weisberg and Ko, 2012). Although the trees in their studies had greater variation in age and shape than in ours, our findings demonstrate that 3D deep learning has evolved into a powerful modelling technique, capable of capturing fine-grained tree biometrics such as tree age, as well as or even better than human non- invasive methods. This is particularly interesting, especially because laser scanning data lacks the richness of semantic information captured by human vision that allows us to capture key morphological traits important for tree age determination, such as bark structure, spiral grain, and the presence of lichens indicative of old trees (Handegard et al., 2021; Marmor et al., 2011).

For *fineTune_FF3D*, a closer examination of performance across age classes (Fig. 7) showed that both random errors (RMSE) and systematic errors (MD) generally increased with tree age. While some increase is expected due to the larger magnitude of reference ages, the errors became notably higher for very old trees, likely because training samples in these age classes were sparse. It is also worth noting that larger errors may still be acceptable for older trees, since in many applications it is sufficient to know that a tree exceeds a certain age threshold, whereas more precise age estimates are typically more important for younger trees in management contexts.

Prediction errors for *P. abies* showed less variability and fewer systematic biases than those for *P. sylvestris* (see Fig. 7a), although this may partly reflect their different age distributions, with *P. abies* spanning a narrower range. The comparison of residuals between ground-based (MLS and TLS) and airborne (ALSHD) laser scanning data (Fig. 7b)

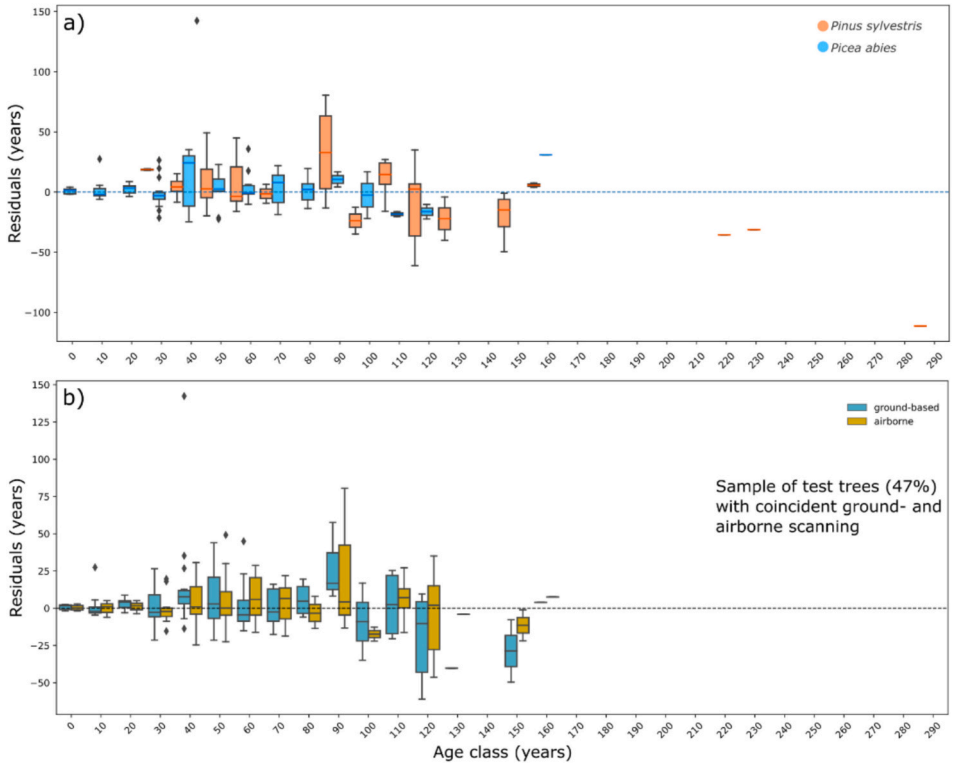

Fig. 7. Boxplots representing the residuals (i.e. reference age minus predicted age) distribution grouped by (a) tree species by (b) laser scanning data for 10 year-wide tree age classes. The latter group was defined using a subsample of 47% of the test trees where coincident ground-based (TLS and MLS) and airborne (ALSHD) was availabe. The boxes represent the interquartile range with the median line, whiskers extend to 1.5 times the interquartile range, points denote outliers, and the dashed horizontal line marks zero.

was conducted on a subsample comprising 47% of the test trees (119 trees) with coincident observations from both data sources (mean age: 61; standard deviation: 35 years). The model showed comparable accuracy on airborne (RMSE = 16 years; MD = 2 years) and terrestrial (RMSE = 22 years; MD = 3 years) data, demonstrating the platform-agnostic capability of *fineTune_FF3D*. For younger trees (<20 years), both data sources exhibit a small spread of residuals, which increases with age. Overall, the dispersion of residuals tends to be larger for ground-based data, particularly in intermediate age classes. For age classes ≥100 years, airborne data appear to show a somewhat smaller spread compared to ground-based estimates; however, the limited number of observations in these older age classes prevents drawing robust conclusions. Residuals for both data sources remain generally centered around zero, indicating only a small bias. In addition, residual variability increases with age and includes more frequent outliers.

From the assessment of classification accuracy across discrete age classes, we observed that the most advanced strategy achieved overall accuracies of up to 76–77% (see Fig. 8). Among them, *fineTune_FF3D* consistently outperformed others across most age groups and was the only scenario capable of correctly classifying some trees older than 200 years. As expected, *scratch_LR* performed poorly, particularly in identifying trees older than 100 years.

From hereon, further results will focus exclusively on the *fineTune_FF3D* scenario, as this was found to be the best-performing method.

It is also important to acknowledge that we did not estimate confidence intervals for the deep learning predictions, although they are an important metric for quantifying uncertainty at the individual-tree or stand level. Methods for estimating variance and uncertainty in deep learning models (e.g. bootstrapping) rely on repeated model training on resampled data and are therefore computationally intensive, which precluded their use in this study.

4.2. Joint-species or species-specific models

To investigate whether a single joint model can effectively replace species-specific age estimation models, we compared the performance of the *fineTune_FF3D* model trained on both pine and spruce trees with that of two species-specific models trained separately on each species. As shown in Fig. 9, the joint model outperformed species-specific models, with improvements of up to 6 years in RMSE. This result suggests that pooling data across the two studied tree species can be beneficial, likely because splitting the dataset into species-specific subsets substantially reduces the amount of training data available to each model. In addition, the joint model may implicitly encode species information while simultaneously learning age-related structural patterns that are shared across species. Despite interspecies differences, crown structural changes with age (e.g. increased branch mortality and weakened apical growth) may follow a similar pattern across species. Rather than focusing on cross-species transfer, our objective is to assess whether a unified model can provide a more accurate and streamlined alternative to maintaining separate species-specific models, which is particularly attractive for large-scale biophysical trait prediction.

4.3. GPS-time augmentation

Based on Fig. 10a and b we can see that training the model with the GPS-time augmentation did not improve overall performance at full resolution compared with the model trained on the original data.

Despite the lack of improvement on the data at original resolution, we also found that when comparing the available trees with single-pass data (i.e. GPS time was not available for all trees), the model trained using the GPS-time augmentation showed a better ability to deal with sparser tree 3D representations (see Fig. 10) and thus highlighting its increased capacity to transfer toward more efficient laser scanning data capture (i.e. reduced passes mean less flightlines or walks) or scan edges. The improvement in transferability to sparser data was substantially larger than the loss in performance on full-resolution data, indicating that GPS-time augmentation might be a useful strategy for realistic simulation of sparser point clouds.

4.4. Future outlook

Our study introduces a novel forest 3D computer vision task and provides a first proof-of-concept demonstrating the value of fine-scale 3D structural information (beyond diameter and height) for tree age prediction. The results show that deep features extracted from state-of-the-art 3D deep learning models can capture rich structural detail, enabling accurate characterization of individual tree biometrics. At the same time, several aspects warrant further investigation to support the development of operational applications. Key areas for future research include:

- **Combining structural with growing conditions**: A potential strategy to improve performance would be to embed topographic-, geographic-, and climate-variables (Lu et al., 2025) as well as tree social status proxies (e.g. shading indices) into the proposed deep learning models, enabling them to ground their predictions in specific growing conditions and potentially increase accuracy. Incorporating information on local growing conditions is highly sensible; however, the spatial granularity of some variables—particularly climatic ones—may be suitable for large-scale applications (e.g. National Forest Inventories) but too coarse for forest management that requires within-stand age distributions.
- **Performance in complex forests**: Although our study focuses on Fennoscandian forests, the FOR-age dataset includes stands representing a gradient of management types, from even-aged, single-

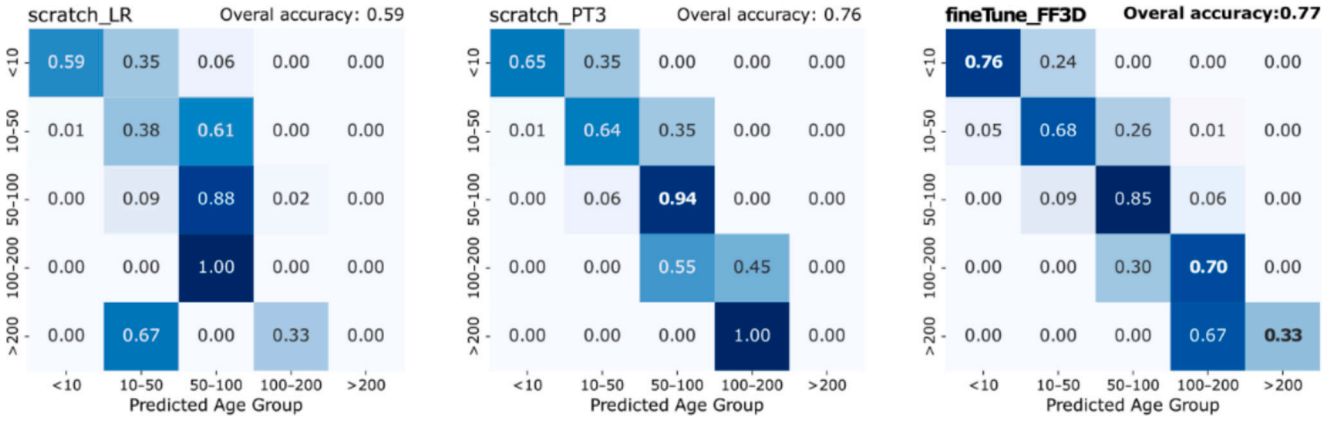

Fig. 8. Confusion matrices for discrete age class classification across all tested scenarios. Highlighted values represent the best performance.

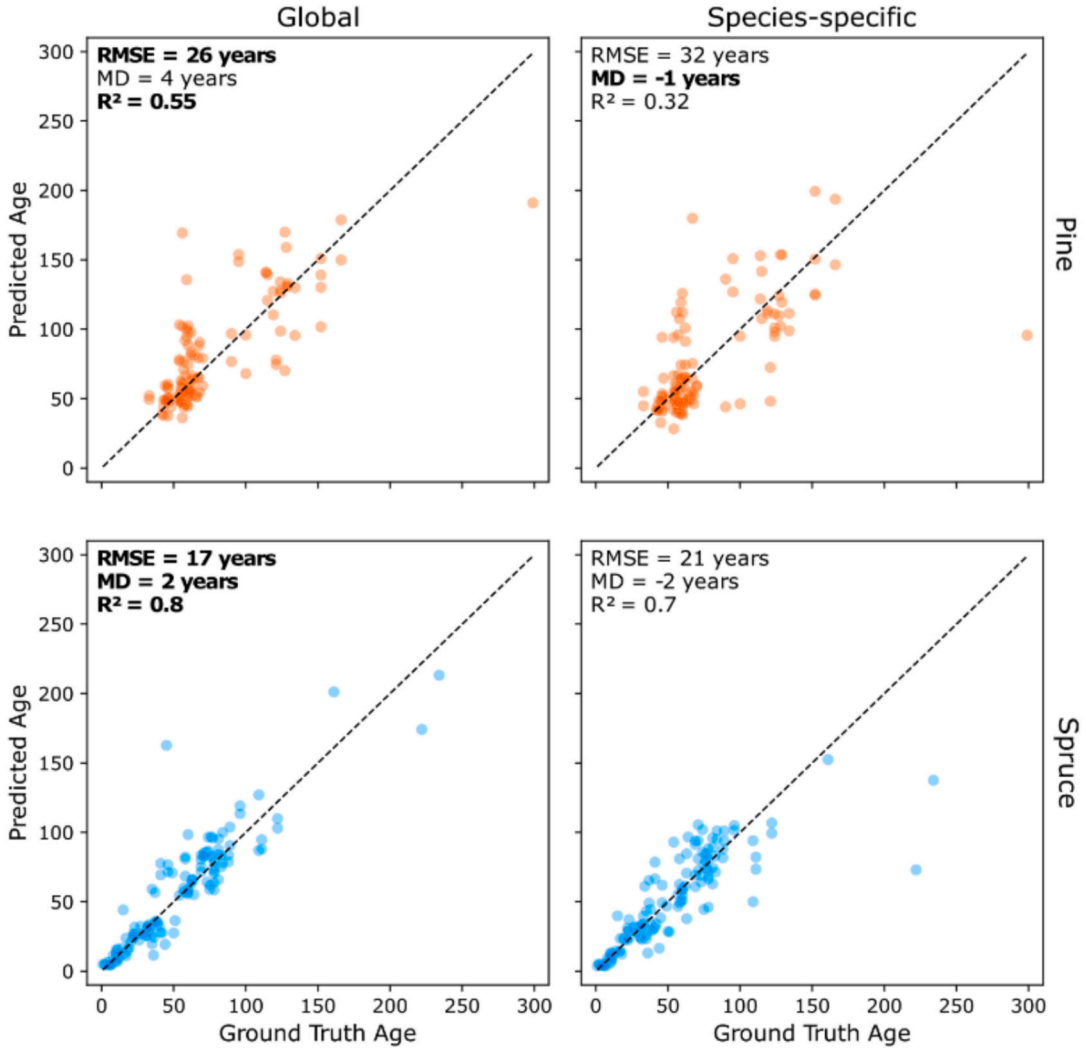

Fig. 9. Scatterplot matrix representing the performance comparison between the global model and species-specific models for Scots pine and Norway spruce.

layered stands to multi-story stands managed through single-tree selective removals, as well as unmanaged forests. Nevertheless, further work is needed to assess how well these models generalize to more structurally complex, multi-layered forest conditions and to denser forests, where tree-crown-size/age relationships differ from those in mature forests. This is particularly relevant in the context of increasing interest in continuous cover forestry (Mason et al., 2021), where the presence of suppressed but old trees may pose additional challenges for accurate age prediction.

- **Explainability**: The markedly higher performance of deep learning models compared to simple linear regression highlights an important advancement in modelling capabilities, but it also introduces challenges related to interpretability. Understanding which structural attributes drive predictions is valuable for many ecological applications, where explaining underlying mechanisms or providing practical guidance is essential. Future work should therefore explore the use of explainable AI approaches to identify the structural features most influential for age prediction, while acknowledging that interpreting such methods can remain difficult.

5. Conclusion

This study introduced a novel downstream task and dataset for estimating individual tree age from 3D point cloud data – an objective previously unexplored in forest remote sensing. By evaluating a range of classical and advanced deep learning techniques, we found that the latest 3D deep learning architectures achieve state-of-the-art accuracy in predicting fine-grained tree biometrics such as tree age, underscoring their relevance for ecological applications.

The most effective strategy emerged from fine-tuning a pre-trained 3D panoptic segmentation model, which leverages embedded forest structural and semantic information. This result highlights the benefits of such models, which can be used beyond segmentation tasks and serve as versatile backbones for a wide range of forest-related downstream tasks.

Our findings demonstrate that it is possible to predict individual tree age with high accuracy, enabling the characterization of forest developmental stages and distinguishing between mature and old-growth trees—information critical for implementing both forest management and conservation strategies. Expanding access to tree age data would allow forest communities to be characterized not only by which species are present and how large trees are, but also by age demographics, opening new directions for research in forest ecology. In addition, improved age information can strengthen forest yield and growth models, many of which rely on accurate estimates of tree or stand age as a key predictor of growth dynamics and productivity. However, the superior performance of the deep learning models comes at the expense of explainability, as it remains unclear which features they use to infer tree age. This limitation may pose challenges for ecological applications that depend on understanding causal relationships.

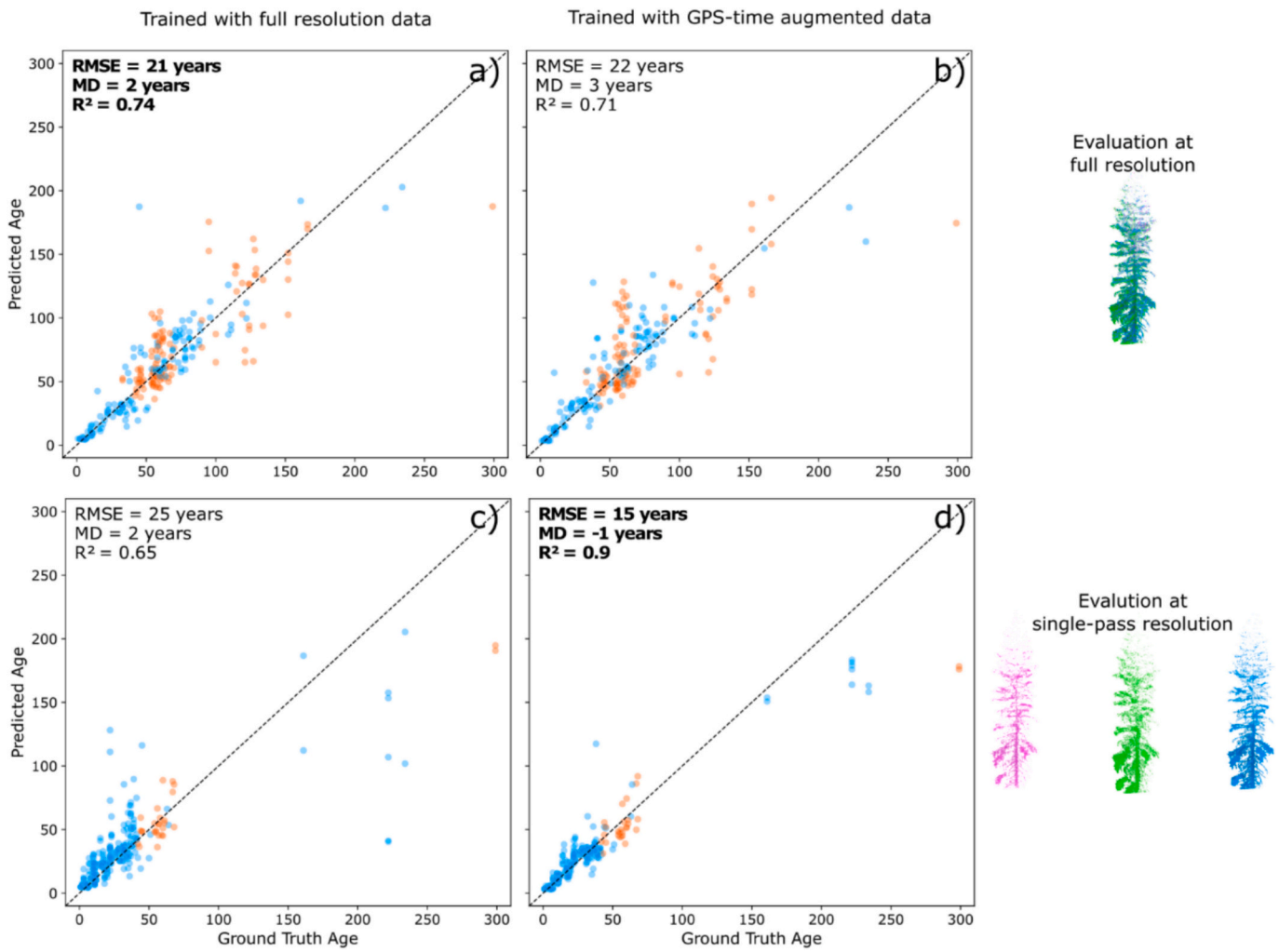

Fig. 10. Scatterplot matrix representing the performance comparison at the original point cloud resolution (a and b) and at single-pass resolution (c and d) between the *fineTune_FF3D* trained using the original resolution data (a and c) only and the *fineTune_FF3D* model trained using GPS-time augmentation (b and d).

Furthermore, while the study was limited to two species, the model's performance suggests that advanced architectures can generalize across species. This challenges the prevailing assumption that structure–biometric tree relationships are species-specific and opens the door to the development of global, scalable forest monitoring tools. The inclusion of data from multiple sites (i.e. varying environmental conditions and management histories) and from multiple platforms and sensors further demonstrates the model's robustness with respect to its transferability to new laser scanning data and forest environments.

Additionally, we showed that model transferability to sparser datasets can be significantly enhanced through realistic augmentations—specifically by slicing tree point clouds based on GPS-time. This approach preserved performance on high-resolution data while enabling effective deployment on cheaper datasets.

Overall, our study marks a significant step forward in the application of 3D deep learning to forest science, offering scalable, accurate, and species-agnostic tools to better understand and manage forest ecosystems.

CRediT authorship contribution statement

Stefano Puliti: Writing – original draft, Visualization, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Binbin Xiang:** Writing – review & editing, Validation, Software, Methodology, Formal analysis. **Maciej Wielgosz:** Validation, Software, Methodology, Investigation, Formal analysis. **Eivind Handegard:** Writing – review & editing, Investigation, Data curation. **Nicolas Cattaneo:** Writing – review & editing, Data curation. **Marta Vergarechea:** Writing – review & editing, Data curation. **Terje Gobakken:** Writing – review & editing, Data curation. **Juha Hyyppä:** Writing – review & editing, Data curation. **Erik Næsset:** Writing – review & editing, Data curation. **Mikko Vastaranta:** Writing – review & editing, Data curation. **Tuomas Yrttimaa:** Writing – review & editing, Data curation. **Rasmus Astrup:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Data availability statement

The FOR-age data used in this study is publicly available at the following link: https://zenodo.org/uploads/19853987 The code for training and validation of the model is available in the following github repository: https://github.com/SingleTree-EU/FORage For benchmarking against the withheld test set, please refer to the instuctrions in the FOR-age benchmark - Codabench competition (https://www.codabench.org/competitions/16014/).

Declaration of competing interest

The authors declare the following financial interests/personal

relationships which may be considered as potential competing interests:

Stefano Puliti reports financial support was provided by Bio-based Industries Joint Undertaking. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is part of the Center for Research-based Innovation SmartForest: Bringing Industry 4.0 to the Norwegian forest sector (NsFR SFI project no. 309671, smartforest.no).

The project is supported by the Circular Bio-based Europe Joint Undertaking and its members under grant agreement No 101157488 (SingleTree).

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or CBE JU. Neither the European Union nor the CBE JU can be held responsible for them.

References

Ahmed, S., Pretzsch, H., 2023. TLidar-based crown shape indicates tree ring pattern in Norway spruce (*Picea abies* (L.) H. Karst) trees across competition gradients. A modeling and methodological approach. Ecol. Indic. 148, 110116.

Ahmed, S., et al., 2025. Crown structure indicates tree secondary growth, competition legacy, and growth potential of dominant species in europe. Ecol. Indic. 170, 113074.

Allen, M.J., et al., 2023. Tree species classification from complex laser scanning data in mediterranean forests using deep learning. Methods Ecol. Evol. 14, 1657–1667.

Blicharska, M., Mikusiński, G., 2014. Incorporating social and cultural significance of large old trees in conservation policy. Conserv. Biol. 28, 1558–1567.

Breidenbach, J., et al., 2020. A century of national forest inventory in Norway – informing past, present, and future decisions. Forest Ecosyst. 7, 46.

Brown, P.M., Gannon, B., Battaglia, M.A., Fornwalt, P.J., Huckaby, L.S., Cheng, A.S., Baggett, L.S., 2019. Identifying old trees to inform ecological restoration in montane forests of the central Rocky Mountains, USA. Tree-Ring Res. 75 (1), 34–48.

Castagneri, D., et al., 2013. Age and growth patterns of old Norway spruce trees in trillemarka forest, Norway. Scand. J. For. Res. 28, 232–240.

Cattaneo, N., et al., 2024. Estimating wood quality attributes from dense airborne LiDAR point clouds. Forest Ecosyst. 11, 100184.

FARO, 2025a. FARO Connect Software. Available at: https://www.faro.com/en/Products/Software/FARO-Connect-Software.

FARO, 2025b. GeoSLAM ZEB Horizon RT Mobile Scanner. Available at: https://www.faro.com/en/Products/Hardware/GeoSLAM-ZEB-Horizon-RT.

Fol, C.R., et al., 2023. Evaluating state-of-the-art 3D scanning methods for stem-level biodiversity inventories in forests. Int. J. Appl. Earth Obs. Geoinf. 122, 103396.

Gold, S., et al., 2006. The development of european forest resources, 1950 to 2000. Forest Policy Econ. 8, 183–192.

Handegard, E., et al., 2021. Identifying old Norway spruce and scots pine trees by morphological traits and site characteristics. Scand. J. For. Res. 36, 550–562.

Henrich, J., et al., 2024. TreeLearn: a deep learning method for segmenting individual trees from ground-based LiDAR forest point clouds. Eco. Inform. 84, 102888.

Holland, A., et al., 2024. Terrestrial lidar reveals new information about habitats provided by large old trees. Biol. Conserv. 292, 110507.

Hyyppä, E., et al., 2020. Accurate derivation of stem curve and volume using backpack mobile laser scanning. ISPRS J. Photogramm. Remote Sens. 161, 246–262.

Kalliovirta, J., Tokola, T., 2005. Functions for estimating stem diameter and tree age using tree height, crown width and existing stand database information. Silva Fenn. 39, 227–248.

Kellomäki, S., 2022. *Management of Boreal Forests: Theories and Applications for Ecosystem Services*. Springer Nature.

Lu, J., et al., 2025. Tree age estimation across the U.S. using forest inventory and analysis database. For. Ecol. Manag. 584, 122603.

Marmor, L., et al., 2011. Effects of forest continuity and tree age on epiphytic lichen biota in coniferous forests in Estonia. Ecol. Indic. 11, 1270–1276.

Mason, W.L., et al., 2021. Continuous cover forestry in Europe: usage and the knowledge gaps and challenges to wider adoption. Forestry: Intern. J. Forest Res. 95, 1–12.

Matthes, U., et al., 2008. Predicting the age of ancient thuja occidentalis on cliffs. Can. J. For. Res. 38, 2923–2931.

Norton, D.A., Ogden, J., 1990. Problems with the use of treerings in the study of forest population dynamics. In: Kairiukstis, E.R.C.A.L.A. (Ed.), Methods Ofdendrochronology: Applications in the Environmental Sciences. Kluwer AcademicPublishers, pp. 284–288.

Palik, B.J., Pregitzer, K.S., 1995. Variability in early height growth rate of forest trees: implications for retrospective studies of stand dynamics. Can. J. For. Res. 25, 767–776.

Pehkonen, M., et al., 2025. Identification and segmentation of branch whorls and sawlogs in standing timber using terrestrial laser scanning and deep learning. Forestry 98 (5), 712–725.

Piovesan, G., Biondi, F., 2021. On tree longevity. New Phytol. 231, 1318–1337.

Puliti, S., et al., 2020. Estimation of forest growing stock volume with UAV laser scanning data: can it be done without field data? Remote Sens. 12, 1245.

Puliti, S., et al., 2023. Tree height-growth trajectory estimation using uni-temporal UAV laser scanning data and deep learning. Forestry: Intern. J. Forest Res. 96, 37–48.

Puliti, S., et al., 2025. Benchmarking tree species classification from proximally sensed laser scanning data: introducing the FOR-species20K dataset. Methods Ecol. Evol. 16 (4), 801–818.

Puliti, S., et al., 2024a. Deliverable 1.2 field Protocol Established –assessing PathFinders Superplots with Advanced Methods. Available at: https://pathfinder-heu.eu/wp-content/uploads/2025/01/PathFinder_D1.2.pdf.

Puliti, S., et al., 2024b. BranchPoseNet: characterizing tree branching with a deep learning-based pose estimation approach. arXiv preprint arXiv:2409.14755.

Pyörälä, J., et al., 2019. Variability of wood properties using airborne and terrestrial laser scanning. Remote Sens. Environ. 235, 111474.

Rehush, N., et al., 2018. Identifying tree-related microhabitats in TLS point clouds using machine learning. Remote Sens. 10, 1735.

Riegl, 2025a. Riegl-miniVUX-1UAV. Available at: http://www.riegl.com/products/unmanned-scanning/riegl-minivux-1uav/.

Riegl, 2025b. Riegl VQ-1560 II S. Available at: http://www.riegl.com/nc/products/airborne-scanning/produktdetail/product/scanner/70/.

Riegl, 2025c. Riegl VUX-240. Available at: http://www.riegl.com/products/unmanned-scanning/riegl-vux-240/.

de Rigo, D., et al., 2016. The European Atlas of Forest Tree Species: Modelling, Data and Information on Forest Tree Species. European Atlas of Forest Tree Species: Publ. Off. EU, Luxembourg.

Rohner, B., et al., 2013. Towards non-destructive estimation of tree age. For. Ecol. Manag. 304, 286–295.

Seidel, D., et al., 2021. Predicting tree species from 3D laser scanning point clouds using deep learning. Front. Plant Sci. 12.

Smith, A., et al., 2024. An investigation into the age structure of Norway spruce and scots pine stands in Norway. Ecol. Indic. 167, 112627.

Výbošťok, J., et al., 2026. An open and novel low-cost terrestrial laser scanner prototype for forest monitoring. Sensors 26, 63.

Weisberg, P.J., Ko, D.W., 2012. Old tree morphology in singleleaf pinyon pine (pinus monophylla). For. Ecol. Manag. 263, 67–73.

Wetherbee, R., et al., 2020. Veteran trees are a source of natural enemies. Sci. Rep. 10, 18485.

Wielgosz, M., et al., 2024. SegmentAnyTree: a sensor and platform agnostic deep learning model for tree segmentation using laser scanning data. Remote Sens. Environ. 313, 114367.

Wu, X., et al., 2024. Point transformer v3: simpler faster stronger. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4840–4851. Available at:

Xiang, B., et al., 2024. Automated forest inventory: analysis of high-density airborne LiDAR point clouds with 3D deep learning. Remote Sens. Environ. 305, 114078.

Xiang, B., Wielgosz, M., Puliti, S., Král, K., Krůček, M., Missarov, A., Astrup, R., 2025. Forestformer3d: a unified framework for end-to-end segmentation of forest lidar 3d point clouds. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 24717–24727.

Yrttimaa, T., et al., 2024. Quantifying architectural uniqueness of Scots pine trees using terrestrial laser scanning: toward individual tree fingerprinting. Forestry 98 (4), 465–477.