



RayExtract: A fast, scalable method for tree volume reconstruction from terrestrial laser scanning



Timothy Devereux ^{a, b,*}, Thomas Lowe ^c, Joshua Rivory ^a, Rafael Bohn Reckziegel ^{e, g}, Kim Calders ^h, Raja Ram Aryal ^a, Glen Eaton ^a, Zane Cooper ^h, Shaun Levick ^d, Stuart Phinn ^a, William Woodgate ^{a, f}

^a School of Environment, The University of Queensland, Brisbane, QLD 4072, Australia

^b CSIRO Environment, GPO Box 2583, Brisbane, QLD 4001, Australia

^c CSIRO Data61, Pullenvale, QLD 4069, Australia

^d CSIRO Environment, Locked Bag 2, Glen Osmond, SA 5064, Australia

^e CSIRO Environment, 564 Vanderlin Drive, Berrimah, NT 0828, Australia

^f CSIRO Space and Astronomy, Kensington, WA 6151, Australia

^g RIEL Charles Darwin University, Casuarina, NT 0810, Australia

^h Q-ForestLab, Department of Environment, Faculty of Bioscience Engineering, Ghent University, Ghent, East Flanders, Belgium

HIGHLIGHTS

- Introduces RayExtract, a new method for reconstructing woody volume from TLS.
- Morphological rules (Self-Similarity, Leonardo's Rule) for branch calculations.
- RayExtract automates plot reconstruction: integrates segmentation & woody modeling.
- Validated by harvest data ($n = 124$), high accuracy (CCC 0.82–0.97).
- Synthetic tree simulation ($n = 18$): high accuracy (CCC 0.97–0.98) in varied scans.
- RayExtract: predictable, efficient scaling (runtime/memory) for large applications.

ARTICLE INFO

Edited by Dr Jing M. Chen

Keywords:

Forest
Tree
Reconstruction
Segmentation
Aboveground biomass
Laser scanning

ABSTRACT

Accurate tree volume and structure are crucial for forest biomass estimation and ecosystem investigations. While terrestrial laser scanning (TLS) offers non-destructive pathways for the detailed three-dimensional tree reconstruction, current methods overestimate small branch volumes and often require tree segmentation and leaf-wood separation as a priori. This study introduces and validates RayExtract, a novel method for reconstructing woody volume from TLS data, utilising tools from the *RayCloudTools* library, to automate the extraction of tree structural metrics from point clouds. Our method incorporates two key morphological rules — Self-Similarity and Leonardo's Rule — to aid branch radius and taper calculations. Likewise, it enables rapid and automated plot-scale reconstruction by integrating tree segmentation and woody structure modelling without requiring leaf point classification. In this study, RayExtract demonstrated high accuracy across four high-quality destructive harvest reference sets with concordance correlation coefficient (CCC) values ranging from 0.82 to 0.97 ($n=124$). To explore algorithm behaviours under different leaf conditions and point densities, we implement a framework using TLS simulation of highly realistic synthetic trees. Results from the simulation framework show consistent high accuracy of total woody volume, with CCC ranging from 0.97 to 0.98 ($n=18$) across four distinct scanning configurations. Fine-scale volumetric analysis revealed that incorporating simple morphological rules can effectively inform branch taper and reduce woody volume overestimation, particularly in smaller components. Furthermore, it identifies a limitation in volumetric accuracy in trees exhibiting significant taper in the lower stem. Analysis of RayExtract's computational efficiency demonstrates that runtime and memory usage scale predictably with input data size, primarily driven by point count and the associated structural complexity within

* Corresponding author at: School of Environment, The University of Queensland, Brisbane, QLD 4072, Australia.

Email address: t.devereux@uq.edu.au (T. Devereux).

the point cloud, positioning the algorithm as well suited for large-scale applications. RayExtract represents a significant advancement in forest reconstruction, biomass estimation, and vegetation structural analysis. The method's efficiency, accuracy, and robustness across varied forest conditions mark a substantial improvement in forest structural assessment techniques using laser scanning and have broad implications for improving regional biomass estimations, and contributing to the calibration and validation of broad-scale remote sensing observations.

1. Introduction

In forest ecology and management, accurate measurements of tree woody biomass are critical for understanding ecological and economic functions (Chave et al., 2014; Picard et al., 2012). These measurements are indispensable for evaluating forest health and productivity, enabling effective conservation strategies and fostering sustainable forest management practices. Moreover, the accurate assessment of woody biomass is pivotal in global climate change mitigation efforts, as forests are a significant component of the terrestrial carbon cycle, sequestering and releasing atmospheric carbon dioxide and thus playing a key role in regulating the Earth's climate (Harris et al., 2021). A deep understanding of biomass distribution is crucial for modelling carbon fluxes and informing policy decisions aimed at reducing greenhouse gas emissions (Pan et al., 2011; Bonan, 2008). Integrating these measurements into climate strategies underscores the link between forest ecosystems and broader environmental and economic systems, further highlighting the necessity for precise and reliable biomass data.

Measuring aboveground woody biomass (AGB) involves calculating the volume occupied by a tree's aboveground components—such as the trunk, branches, and twigs—and multiplying this by the average measured woody density. Traditional AGB methods typically rely on manual measurements and allometric estimations (Chave et al., 2014; Picard et al., 2012), which are labour-intensive and may not scale accurately to larger trees (Burt et al., 2021; Demol et al., 2024). In recent years, terrestrial laser scanning (TLS) has emerged as a precise and non-destructive method for measuring aboveground woody volume at the plot scale (Calders et al., 2015; Gonzalez de Tanago et al., 2018; Momo Takoudjou et al., 2018). TLS provides a permanent, non-destructive record of three-dimensional forest structure that enables repeated measurements and analyses not feasible with traditional manual methods (Calders et al., 2020). Beyond biomass estimation, reconstructing detailed 3D tree architecture supports fine-scale applications including wood fibre analysis (Côté et al., 2021), forest fuel assessments (Arkin et al., 2025), and wildfire behavior modelling (Abdollahi and Yebra, 2025). Additionally, TLS measurements of forest structure have become integral for the validation and calibration of spaceborne biomass mapping efforts (Duncanson et al., 2019; Herold et al., 2019), such as NASA's Global Ecosystem Dynamics Investigation (GEDI) (Dubayah et al., 2020) and the European Space Agency's BIOMASS mission (Quegan et al., 2019).

While terrestrial LiDAR offers the potential for highly accurate individual tree volume measurements, challenges remain in scaling these methods to landscape levels (Calders et al., 2020). Limitations include restricted sample sizes, uncertainties in long-term monitoring effectiveness, and unresolved questions about optimal sampling strategies across diverse forest conditions, species, and plot components (Akerblom and Kaitaniemi, 2021; Disney et al., 2018). The maintenance cost, potential sampling bias, and the choice of biomass mapping frameworks (individual-tree, area-based, or hybrid) at broader scales also require further investigation. Therefore, the current strength of TLS lies in its ability to provide highly detailed, fine-scale data, rather than offering a complete solution for continuous biomass mapping at landscape scales.

Deriving metrics from TLS data is technically demanding and requires a nuanced understanding of associated biases arising from beam divergence, occlusion and scanner specific artefacts. Several processing pipelines have been developed to automate the extraction of tree structural metrics from point clouds (Martin-Ducup et al., 2021). These pipelines typically involve three main steps: individual tree

segmentation (Burt et al., 2018; Donager et al., 2021), wood-leaf classification (Wang et al., 2020; Vicari et al., 2019), and 3D Quantitative Structure Models (QSM) reconstruction (Fan et al., 2020; Raumonen et al., 2013; Hackenberg et al., 2015). QSMs represent trees as a collection of cylinders fitted to point cloud data and provide a detailed and mathematically tractable representation of tree architecture that can be used to derive tree structural metrics (Raumonen et al., 2013; Fan et al., 2020; Hackenberg et al., 2015). The accuracy and robustness of various QSMs models have been extensively validated across a range of forest types and scanning conditions (Burt et al., 2018; Calders et al., 2015; Gonzalez de Tanago et al., 2018; Momo Takoudjou et al., 2018). A comprehensive global synthesis by Demol et al. (2022a) demonstrated that TLS-derived biomass estimates achieve remarkable accuracy when compared to destructive measurements, with less than 1 % bias and a concordance correlation coefficient of 98 % across diverse forest conditions. This validation spanned an impressive range of forest types including tropical forests (Brazil, Peru, Guyana, Suriname, Cameroon, Indonesia), temperate forests (Belgium, Germany, USA), subtropical Eucalyptus forests (Australia), and urban trees (Switzerland), with measurements across various leaf conditions (leaf-on, leaf-off, needle-on, needle-off). However, this synthesis revealed important limitations: TLS performed less effectively for smaller trees (<1,000 kg) with an overestimation bias of 19.7 %, and for conifers (16 % bias compared to 0.47 % for broadleaves). While volume-based methods incorporating morphological constraints can improve radius estimation accuracy, challenges remain in applying these techniques to complex forest structures, particularly in fuel-rich environments with dense understories and clustered thin trees where occlusion becomes a significant factor. Additionally, recent studies have indicated that prevalent methods for measuring woody volume tend to overestimate the volume of smaller branch components (Morhart et al., 2024; Demol et al., 2022b). Efforts to reduce the overestimation of small branch radii have included input filtering to remove noise caused by partial LiDAR returns (Wilkes et al., 2021), allometric scaling filters (Hackenberg and Bontemps, 2023), and species-specific filters (Morales and MacFarlane, 2024) with varying levels of success. However, automatically deriving accurate tree branch volume without species-specific inputs remains an open challenge, hindering efforts for large-scale, automated forest reconstruction.

Assessment of QSM volume metrics also presents significant methodological challenges. Conventional techniques relying on destructive sampling are resource-intensive, labour-demanding, and susceptible to measurement errors. While comparing TLS-derived volume estimates against destructive harvesting provides insights into algorithm accuracy, it is prone to biases arising from errors in destructive measurements and variations in point cloud quality (Demol et al., 2022a; Morhart et al., 2024). Further complicating this, studies comparing various QSM algorithms have revealed significant performance discrepancies. For instance, a recent study benchmarking multiple QSM algorithms against a combined dataset of 160 destructively measured trees found that while some algorithms consistently estimated biomass within 20 % of destructive values, others, such as TreeQSM v2.4.1, showed substantial overestimations, averaging 3.8 times the destructive values (Cooper et al., 2025). This highlights that even with destructive data, the choice of QSM algorithm introduces considerable variability and uncertainty, underscoring the need for standardized benchmarking and careful consideration of methodology (Cooper et al., 2025). To address these limitations, the generation of TLS data from synthetic reference trees

has emerged as a promising alternative for evaluating TLS methods. This approach mitigates the aforementioned constraints associated with physical sampling. Previous studies have implemented this methodology predominantly utilising parametrically modelled trees with simplified geometries (Binney and Sukhatme, 2009; Disney et al., 2012; Raumonen et al., 2013; Abegg et al., 2023). The use of synthetic data allows for the comparison of metrics against known volumetric references, circumventing the need for in situ destructive measurements. This method not only overcomes the logistical and ethical concerns inherent in destructive sampling but also offers a scalable framework for evaluating TLS accuracy across a spectrum of tree architectures and scanning parameters.

This study introduces and validates RayExtract, a novel method for reconstructing tree volume from terrestrial laser scanning data, developed as part of the RayCloudTools (RCT) library (Lowe and Stepanas, 2021). RayExtract combines two complementary tools: *rayextract trees* for woody volume reconstruction and *rayextract terrain* for ground surface modelling. The method distinguishes itself from existing approaches through its ability to process plot level point clouds without prior leaf–wood classification, while maintaining both computational efficiency and volumetric accuracy.

To evaluate the RayExtract algorithm's capacity for estimating individual tree volume (Objective #1), we employ two distinct yet complementary approaches. Firstly, traditional destructive harvesting provides a direct real-world benchmark against established methods. However, the generalizability of these specific results is inherently limited by the relatively small sample size and the unique characteristics of the harvested trees.

Secondly, complementing the destructive sampling, a simulation framework using detailed 3D tree models allows for a focused assessment of the algorithm's behaviour under precisely controlled conditions. Crucially, this simulation approach is intended to rigorously test algorithmic behaviour and sensitivity, rather than serving as a direct substitute for real-world validation. Its primary strengths include enabling: (a) precise control over environmental and scanning parameters to examine the algorithm's response without real-world confounding factors; (b) direct comparison of algorithm performance on leaf-on versus leaf-off states, which is challenging with destructive sampling; and (c) a controlled environment ideal for internal algorithm benchmarking and sensitivity analysis. Thus, while results from simulated data yield valuable insights into the algorithm's performance characteristics and parameter sensitivities, they are interpreted separately from the real-tree benchmarks and are not presented as directly generalizable to all field scenarios.

Finally, we assess the algorithm's computational efficiency (Objective #3). This involves quantifying processing time and memory usage when applying RayExtract to real-world point cloud datasets representing a gradient of structural complexity. Although the RayExtract workflow includes individual tree extraction, the assessment presented in this paper focuses specifically on the accuracy and behaviour of the tree volume retrieval component.

2. Methods

The Methods section comprises two main parts: the algorithm description (Section 2.1) and the framework for assessing algorithmic accuracy and efficiency in extracting individual tree woody volume from point clouds (Section 2.2). While the algorithm description applies to canopy- or plot-scale point clouds containing multiple trees and includes an individual tree segmentation stage, the assessment framework focuses specifically on woody volume extraction from pre-segmented individual tree point clouds. Therefore, individual tree segmentation—a necessary prerequisite for woody volume estimation—falls outside the scope of our evaluation (Section 2.2). For clarity, when processing a single pre-segmented tree, the algorithm still executes its complete pipeline, including the segmentation step (Section 2.1.3), which

naturally identifies the single tree before proceeding to reconstruction. The algorithm architecture remains identical regardless of whether the input is plot-scale or individual tree point clouds.

2.1. Algorithm description

Self-similarity and Leonardo's rule are foundational concepts in understanding tree architecture. Self-similarity describes how parts of an object can be scaled versions of the whole, manifesting in plants either geometrically (exact scaled copies) or, more commonly, topologically through recurring connection patterns (Ferraro et al., 2005). This principle is often modelled using fractal geometry (Mandelbrot, 1983) and L-systems, which employ recursive rules to generate branching structures (Prusinkiewicz and Lindenmayer, 1990). Leonardo's rule is an empirical observation where the sum of the daughter branches' cross-sectional areas equals that of the parent branch (Eloy, 2011). RayExtract utilizes Leonardo's rule to constrain estimates at branch junctions and applies a self-similarity model where the length-to-radius ratio of branches is assumed to be approximately constant. This geometric similarity model is supported by observations in many gymnosperms (Niklas, 1994) and rainforest understory trees (Osunkoya et al., 2007). For instance, in Norway Maple, this ratio is relatively constant for branches between 1 m and 10 m in length, representing an inflection region before branches become primarily structural (Dahle and Grabosky, 2010). These rules, while not universally applicable, are a useful approximation for many tree species and are often used in allometric scaling models (West et al., 1997, 1999). By leveraging these rules, RayExtract constrains branch volume estimations and streamlines geometric calculations, enhancing both accuracy and computational efficiency. RayExtract introduces several innovations to TLS-based tree reconstruction by:

- Uniquely integrating these branch morphological rules within a graph-based segmentation approach to mitigate branch volume overestimation.
- Employing a novel quadratic distance-weighted shortest path algorithm, reducing path cross-connections between neighbouring branches.
- Incorporating a robust radius estimation less sensitive to vegetation occlusion and foliage interaction.
- Performing joint segmentation and reconstruction in a single pipeline for improved computational efficiency in large-scale applications.

The initial step is the extraction of a terrain elevation model from point cloud data. The *rayextract terrain* tool is utilised for this purpose. This process constructs a three-dimensional mesh that accurately delineates the terrain's surface contours. The terrain model is used to identify tree bases, after which the point cloud is segmented into individual trees and reconstructed as QSMs simultaneously using the *rayextract trees* command. Although the methods described apply to 3D reconstruction for both the stand and individual tree levels, this study focuses on comparison against pre-segmented individual trees. The RayExtract method is similar in its general approach to the distance segmentation methods (Tagliasacchi et al., 2016; Hu et al., 2017), as originally described in Xu et al. (2007). These methods find the shortest path from a tree base to all points in the tree, then segment these paths by distance to generate the nodes N_s . However, there are significant differences compared to these other approaches. The method is summarised and split into subsections in Fig. 1. Throughout the description, we use the k_x format for user-definable parameters, while the default values are found in Table 1.

2.1.1. Reconstruction input

There are two input files required for the tree/forest reconstruction. The first is a point cloud containing 3D surface points $p_i \in R^3, i \in 0..n$. These points should sample the stems, branches and leaves of all the trees present in the scene, in addition to the ground and other vegetation types or landscape surfaces. The second input is the ground surface, which is reconstructed using the sand model method

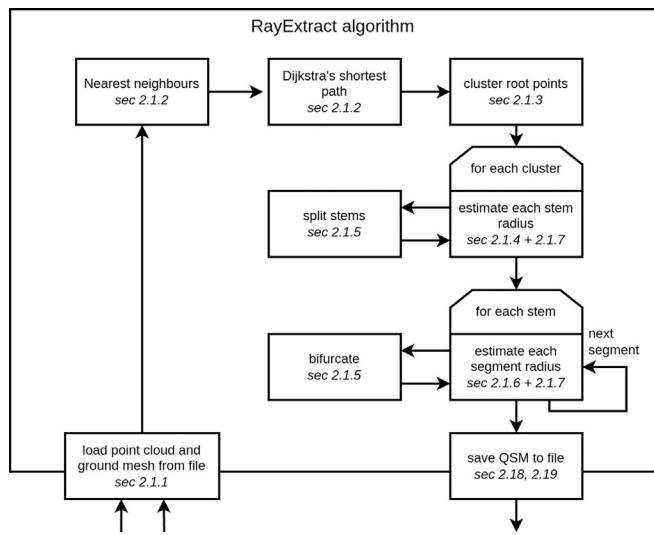


Fig. 1. Overview of the RayExtract method pipeline. Boxes indicate major steps and refer to the sections where each computational process is described.

(Lowe and Pinski, 2022) implemented in the RCT library (Lowe and Stepanas, 2021). This generates a triangular mesh of the highest lower bound of the point cloud based on a conical kernel of gradient 1. These two input files are sufficient for running the forest reconstruction based on the following shortest path from the ground method.

2.1.2. Shortest path graph

The first step of distance segmentation methods is to find the shortest path from a tree base to all points in the tree. The underlying algorithm is Dijkstra's shortest path algorithm (Dijkstra, 1959), which operates on any undirected graph with non-negative edge weights. The nodes are derived from the points \mathbf{p}_i and the edges connect to the nearest neighbour points up to some distance limit. The maximum distance k_d is configurable, and the number of neighbours is limited to a maximum of 20 to prevent excessive computation without compromising the accuracy of

Table 1
RayExtract user-definable parameters and default values used in the methods description.

User Parameter	Command Line Name	Description	Value Used
k_c	-c	Cylinder length to width ratio	4
k_p	-crop_length	Crops small branches to this distance from end	1 m
k_d	-distance_limit	Maximum edge length in shortest path forest	1 m
k_m	-max_diameter	Maximum trunk diameter in segmenting trees	0.9 m
k_o	-global_taper_factor (k_o^3)	0–1 for per-tree to per-scan taper estimation	0.67
k_f	-gravity_factor	Larger values preference vertical trees	0.3
k_h	-height_min	Minimum height counted as a tree	2 m
k_j	-girth_height_ratio	The amount up each tree's height to estimate trunk girth	0.12
k_g	-g	Maximum point gap relative to branch length	0.016
k_s	-s	Maximum point span relative to segment radius	4.5

the pathfinding. For fast nearest neighbour estimation we use a k_d tree algorithm (Elseberg et al., 2012). The multi-root variant of Dijkstra's algorithm is initialised with multiple root points and so it generates a shortest path forest rather than shortest path tree. More formally, it generates a disjoint acyclic undirected graph. These root points are the vertices of the ground mesh, thus the tree bases do not have to be calculated in advance.

A square distance weight is used for the graph edges to encourage the paths to follow the chains of points up the branches in the point clouds, rather than cross-connecting to neighbouring branches. To understand the propensity to erroneously cross-connect to a neighboring branch, we can model a case where two stems are sampled by points at Δx metres apart that create a cross-connection over a gap of length g_1 or less using the linear edge weight Fig. 2(a). With the quadratic edge weight, it can be shown that the equivalent gap size is:

$$g_2 = \sqrt{g_1 \Delta x} \quad (1)$$

Therefore the gap size (and the propensity to cross-connect) is less than for the Euclidean case whenever $\Delta x < g_1$. The more densely the branches are sampled the greater this advantage.

We also minimise cross-connections by preferring smoother paths over sharp direction changes. For this reason the edge weight is given a directional weighting prior to squaring such that right angle direction changes and greater are excluded Fig. 2(b). The combined weighting method is as follows:

$$w_{ij} = \frac{1}{h_i} \left(\frac{|\mathbf{v}_{ij}|}{\max(0, \hat{\mathbf{v}}_{ij} \cdot \hat{\mathbf{v}})} \right)^2 \quad (2)$$

$$\mathbf{v}_{ij} = \mathbf{p}_j - \mathbf{p}_i$$

where $\hat{\mathbf{v}}$ is the direction vector from \mathbf{p}_i to its parent's parent in the shortest path forest. Here h_i is the tree height above point \mathbf{p}_i , which is approximated by collecting all points \mathbf{p}_l within the same $k_m \times k_m$ metre grid cell as \mathbf{p}_i and taking the vertical span: $h_i = \max_l \mathbf{p}_{lz} - \min_l \mathbf{p}_{lz}$. The $1/h_i$ factor is to prevent shortest paths from flowing up small trees into big trees where a lidar point may be a long way from the trunk (Fig. 3).

To discourage paths that stray too far laterally—which are unrealistic for a tree supporting its weight under gravity—the edge weight w_{ij} is multiplied by the gravity multiplier g_{ij} :

$$g_{ij} = 1 + k_f \left(\mathbf{v}_{rj_x}^2 + \mathbf{v}_{rj_y}^2 \right) \quad (3)$$

$$\mathbf{v}_{rj} = \mathbf{p}_j - \mathbf{p}_{\text{root}} \quad (4)$$

where k_f is the gravity factor constant. The shortest path forest is encoded as a parent $i_{\text{par}(i)}$ point index for each point i .

2.1.3. Initial tree segmentation

The initial tree segmentation takes the shortest path forest and generates a list of root segments representing the base of each tree. Each segment of index s contains a set of root points $R_s = i_0, i_1, \dots, i_m$, which are the vertices of the input ground mesh. These form the roots of the shortest path forest contained in that segment. To generate these root segments, the Euclidean path distance $d(i)$ from each point i to its most distant leaf point is first calculated. This is done by iterating from leaf points through the parents to update the largest path distance from point to leaf. Then, the root points are clustered into individual tree base candidates on which to start building each tree. This clustering is done in a grid fashion using the same $k_m \times k_m$ metre horizontal gridding of the point cloud and calculating the maximum distance $d_{\max}(s)$ to a leaf point within each 2×2 grid cell block, indexed by s , as seen in Fig. 4. A segment is generated for each 2×2 block, and every root point is inserted into R_s for the s corresponding to the largest $d_{\max}(s)$. Using 2×2 blocks avoids the generation of multiple tree segments when a single tree crosses the grid

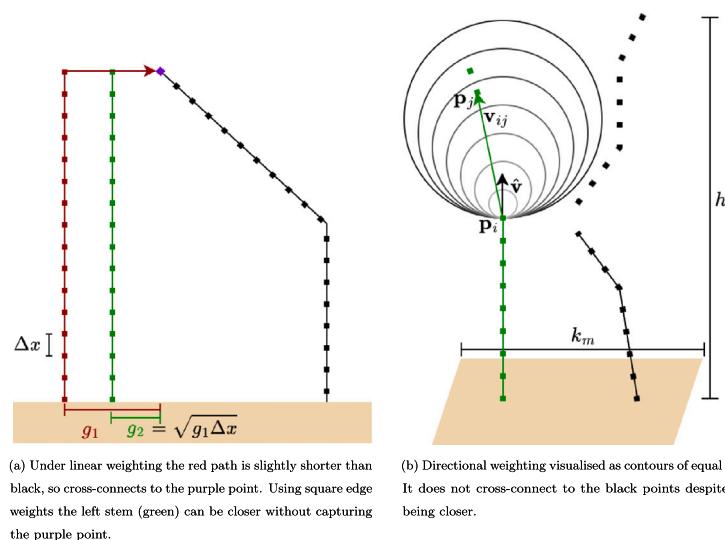


Fig. 2. Benefit of squared edge weighting (a) and directional weighting (b). The weighting function in Eq. (2) applies both modifiers, first directional weighting then squaring of the weight.

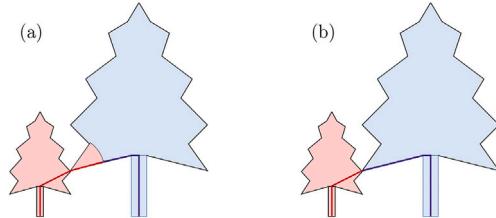


Fig. 3. Demonstration of need for $1/h_i$ factor in the shortest path edge weighting Eq. (2). (a) With equal edge weights the paths from the shorter tree can occupy nearby regions of the larger tree. (b) With edge weights scaled inversely to tree height, point ownership is more accurate for proximate trees.

boundary. For this reason, the maximum tree diameter in the point cloud should be less than k_m . All initial segments s where $d_{\max}(s)$ is less than a minimum tree height k_h are then removed. This is useful for ignoring undergrowth and ground mesh vertices that are not at tree trunks.

2.1.4. Trunk radius estimation

For each tree candidate, the next step is to estimate its trunk radius. This is more critical than the radius of other branches because the trunk is the thickest part of the tree and thus provides the best opportunity to accurately determine the overall width scale of the tree. Accordingly, the radius is estimated at three heights around an average height that is a given ratio k_I of the tree's height. Each radius is estimated using the method in Section 2.1.7, and the smallest of these three is used as the radius estimate. This is because noise and vegetation in the vicinity of a trunk almost always lead to an overestimation of the trunk radius. The set of points from the base to k_I of the tree's height together with its radius then becomes the base segment, indexed s .

To avoid nearby stems being considered as a single stem, we perform the branch segmentation described in the following section, and repeat the trunk radius estimation whenever the points have been split into multiple stems. After this, the method performs the segment reconstruction phase of Section 2.1.6.

2.1.5. Branch segmentation

To determine whether the segment has branched, its set of endpoints is examined, projected by the segment's direction into 2D. A branching event is considered if the set of points spans more than some multiple k_s of the segment's radius, or if there is a gap between neighboring points

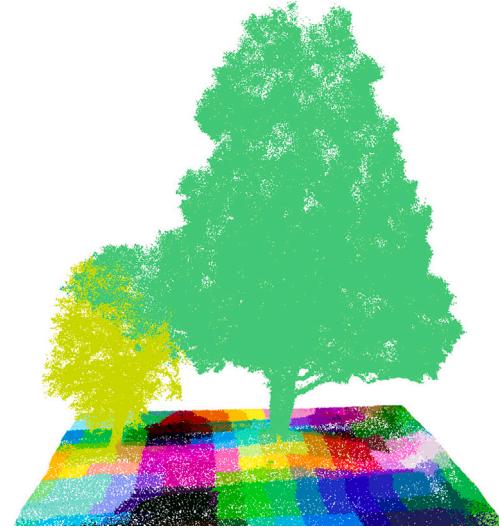


Fig. 4. Input cloud coloured by initial segmentation. The $k_m \times k_m$ metre tiles are visible, but coloured according to the cell with maximum distance to leaf (d_{\max}) in a 2x2 cell neighbourhood. Consequently the central tree (with trunk intersecting four cells) has a single colour and so the ground points of that colour form the root points of a single base segment. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

that exceeds some chosen ratio k_g of the branch length $d_{\max}(s)$. If this occurs, it is necessary to know which points belong to which child branch, requiring a point clustering algorithm.

An agglomerative clustering method is used (Fig. 4), whereby all points are ordered by the distance to their closest neighbour and these edges are iterated through in turn, up until edge length $k_g \times d_{\max}(s)$. With each newly considered edge, the graphs are concatenated if the diagonal of the bounding box of the combined graph is less than $k_s \times$ radius. The result of this set of concatenations is a set of n connected graphs.

2.1.6. Segment reconstruction

In segment reconstruction, the process starts with the set of root points for the segment and a parent radius. Using a length-to-width

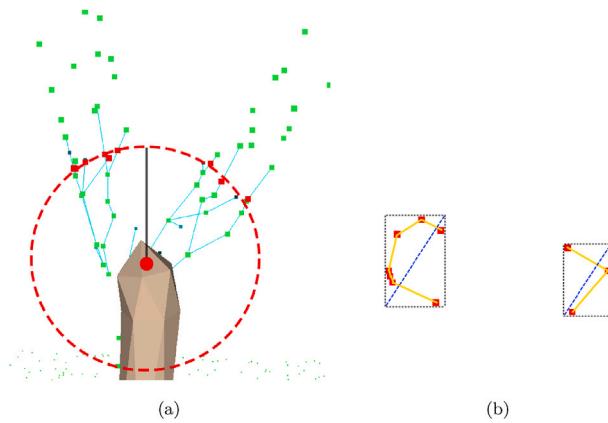


Fig. 5. Branch segmentation and reconstruction process. (a) Side view: segment length $k_c \times \text{parent_radius}$ determines end points shown in red. (b) Plan view: agglomerative clustering groups end points up to edge length $k_g \times d_{\max}(s)$ and while bounding box diagonal (dashed blue line) $\leq k_s \times \text{radius}$. Clustered end points become the root points R_{s+l} for child segments. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ratio k_c , the length of the segment is calculated from this parent radius Fig. 5(a). The shortest path edges are then followed to obtain the first points beyond the segment. A linear interpolation with their parents provides the endpoints at the top face of the cylindrical segment. These points are used for radius estimation, as described in the next section.

Before iterating to the next segment, it is necessary to estimate this segment's radius according to Section 2.1.7 and then decide whether this segment is a branching segment according to Section 2.1.5.

Each of the segmented n sets of endpoints then becomes the root points $R_{s+l}, l \in (1..n)$ for the next branch segment at index $s+l$ Fig. 5(b). This segment reconstruction is repeated for each new segment in turn until the path distance to the end $d_{\max}(s)$ is less than a crop length k_p . This parameter allows the user to trade off visual fidelity with accuracy and speed; low values reconstruct tiny branches, but these are typically uncertain due to lidar coverage limitations and are more likely to be leaf points than larger branches.

By setting the appropriate segment crop length parameter, our algorithm naturally avoids fitting to leaf points without requiring explicit leaf point removal as a preprocessing step. This approach maintains reconstruction quality while simplifying the workflow, as the algorithm intrinsically distinguishes between the more structured branching patterns and the noisier, less organised leaf point distributions during the reconstruction process.

2.1.7. Radius estimation

To estimate the radius of a cylindrical segment, we employ a circle-fitting approach in the plane perpendicular to the branch axis. Let the branch direction be defined from the parent position \mathbf{p}_{par} to the segment centroid Fig. 6(a).

The n segment points are projected onto the plane perpendicular to the branch direction, yielding 2D points $\mathbf{q}_i \in \mathbb{R}^2$, where $i = 1, \dots, n$ Fig. 6(b).

To estimate the circle center, we transform the projected points into a paraboloid representation:

$$\mathbf{u}_i = (q_{i,x}, q_{i,y}, \frac{1}{2}|\mathbf{q}_i|^2) \quad (5)$$

where $q_{i,x}$ and $q_{i,y}$ are the x and y components of \mathbf{q}_i . The key insight is that if the original points \mathbf{q}_i lie on a circle, the transformed points \mathbf{u}_i will lie on a plane in 3D space (Fig. 6(d), Appendix A).

We centre the transformed points by subtracting their mean:

$$\mathbf{v}_i = \mathbf{u}_i - \bar{\mathbf{u}} \quad (6)$$

$$\bar{\mathbf{u}} = \frac{1}{n} \sum_{i=1}^n \mathbf{u}_i \quad (7)$$

We fit a plane to the centered points \mathbf{v}_i using least squares Fig. 6(d). The plane gradients (g_x, g_y) are computed as:

$$g_x = \frac{S_{xz}S_{yy} - S_{yz}S_{xy}}{S_{xx}S_{yy} - S_{xy}^2} \quad (8)$$

$$g_y = \frac{S_{yz} - g_x S_{xy}}{S_{yy}} \quad (9)$$

where the summation terms are:

$$S_{xx} = \sum_{i=1}^n v_{i,x}^2, \quad S_{yy} = \sum_{i=1}^n v_{i,y}^2, \quad S_{xy} = \sum_{i=1}^n v_{i,x}v_{i,y} \quad (10)$$

$$S_{xz} = \sum_{i=1}^n v_{i,x}v_{i,z}, \quad S_{yz} = \sum_{i=1}^n v_{i,y}v_{i,z} \quad (11)$$

The estimated centre of the circle in the original 2D projection is:

$$\mathbf{c} = (g_x, g_y) \quad (12)$$

For each projected point \mathbf{q}_i , we compute its distance to the estimated center:

$$r_i = |\mathbf{q}_i - \mathbf{c}| \quad (13)$$

The segment radius r is then calculated using a robust averaging method that reduces sensitivity to outliers:

$$r = \left(\frac{1}{n} \sum_{i=1}^n r_i^{1/4} \right)^4 \quad (14)$$

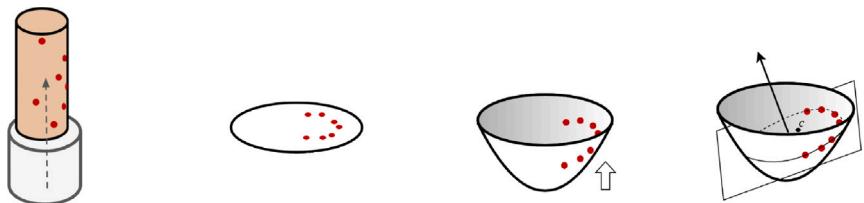


Fig. 6. Radius estimation process using paraboloid transformation for robust circle fitting. The method projects 3D points onto a perpendicular plane, transforms them to a paraboloid representation where circular points become coplanar, then extracts the circle center directly from the fitted plane gradients.

This quartic mean is more robust to large-radius outliers (e.g., vegetation surrounding a branch) compared to the arithmetic mean.

The fitting accuracy for this segment is quantified as:

$$a = \frac{r}{e} \quad (15)$$

where e is the mean absolute deviation of the radii:

$$e = \frac{1}{n} \sum_{i=1}^n |r_i - r| \quad (16)$$

A higher accuracy value a indicates more consistent radii and thus a better cylindrical fit.

2.1.8. Taper estimation

In forest point clouds, dense vegetation often causes occlusion and noise, making branch radius estimation challenging, especially for smaller branches. In order to produce reliable tree structures, it is necessary to rely on more than just branch segment points to estimate radius. Accordingly, prior information about trees, such as Leonardo's rule and self-similarity, is exploited. These two priors fix the ratio of segment radius to trunk radius for every segment, denoted as q_s for each segment s . The trunk radius relative to the tree's length L_t is called the tree taper and is the free parameter that is estimated. While the two priors heavily constrain the tree's branch radii, they also provide robustness to complex or noisy point clouds and allow all of the segment radius estimations to be accumulated into a singular taper estimation per tree. It's important to note that tapering within a tree can be variable (Burkhart and Tomé, 2012). Despite this variability, the method of accumulating estimations into a singular taper estimation is currently one of the best available approaches for dealing with complex forest point clouds. However, the assumption that accumulated estimations consistently average out noise to give a more reliable estimate requires further investigation. The limitations and potential implications of this approach are discussed in detail in the Discussion.

The accumulated taper estimation for a tree t is a weighted average of the taper at each segment s :

$$T(t) = \frac{\sum_s w(s, t) \frac{r_s}{q_s L_t}}{W(t)} \quad (17)$$

where $W(t) = \sum_s w(s, t)$. The segment weight $w(s, t)$ is in proportion to the radius estimation accuracy a_s and to the branch volume, since larger volume branches would be expected to have a better signal to noise.

$$w(s, t) = w_s^{\text{junc}} a_s (q_s L_t)^3 \quad (18)$$

where w_s^{junc} is a junction weight, which is 1 by default and 1/4 on segments that are at a branching point. This reflects the uncertainty in estimating the radius on branching segments, which is expected to be less cylindrical. Even with these two constraining priors, in dense forests the smaller trees may not have a sufficiently clear lidar signal to estimate the taper gradient. Therefore an optional third prior is included: that all trees in the scan share the same taper gradient, thus utilising information from neighbouring trees.

This prior is enforced in a soft manner depending on the size of the tree, through the parameter k_o . Firstly, the weighted average taper \bar{T} is calculated for the scan and its weighted average weight \bar{W} :

$$\bar{T} = \frac{\sum_t T(t) W(t)}{\sum_t W(t)} \quad (19)$$

$$\bar{W} = \frac{\sum_t \sum_s w^2(s, t)}{\sum_t W(t)} \quad (20)$$

The taper $T^*(t)$ is then chosen for a particular tree t as a weighted blend of the tree's estimated taper $T(t)$ and the whole-scan estimated taper \bar{T} :

controlled through the blend parameter k_o :

$$T^*(t) = \frac{k_o \bar{T} \bar{W} + (1 - k_o) T(t) W(t)}{k_o \bar{W} + (1 - k_o) W(t)} \quad (21)$$

When $k_o = 0$ each tree's taper is individually estimated, and when $k_o = 1$ all trees use the scan-wide taper estimation. For intermediate values of k_o the scan-wide taper estimation is applied preferentially to the trees with the least weight $W(t)$, these are the shorter trees, with the least taper certainty.

From this taper the final radius \hat{r} is calculated at any segment of a tree:

$$\hat{r}(s, t) = T^*(t) q_s L_t \quad (22)$$

2.1.9. Reconstruction output

The output is a text file describing the forest as a disjoint Directed Acyclic Graph, stored as a list G of nodes containing a position \mathbf{p} , radius r and parent node s_{par} .

$$G = (N_0, N_1, \dots, N_m) \quad (23)$$

$$N_s = (\mathbf{p}, r, s_{\text{par}}) \quad (24)$$

Each node together with its parent node describes a cylindrical segment of a tree branch. To index both the node and the segment that connects that node to its parent s is used, whereas i, j index the lidar points, and t will be used to index each tree. The output is therefore a piecewise-cylindrical representation of the forest. It is a disjoint graph as there is one acyclic graph per tree in the forest. This output G is referred to as the forest's Quantitative Structure Model (QSM). Ultimately, the main processing stages of the RayExtract algorithm are depicted in Fig. 7, in a synthetic point cloud, complementary to Fig. 1, where the method's workflow was presented.

2.2. Assessment of individual tree volume

We employ two approaches to assess the woody volume estimations produced by RayExtract (Objective #1): (1) a comparison of total volume outputs with real-world destructive harvest measurements, and (2) an evaluation using synthetic reference trees. This dual-method approach leverages the empirical precision of actual measurements while also taking advantage of the controlled variability and detailed analysis offered by simulated reference data. The synthetic approach is used to demonstrate RayExtract's performance in higher branching orders in both leaf-on and leaf-off point clouds in two survey configurations at different scanner distances that produce contrasting sampling densities (Objective #2). RayExtract was run using default parameters, each tree's taper was set to be individually estimated as per Section 2.1.8.

2.2.1. Destructive harvest data

A high-quality comprehensive representation of tree morphology was compiled for validating RayExtract from pre-existing datasets based on (Demol et al., 2022a). Specifically, a total of 124 destructively harvested trees were obtained with coincident TLS from openly available sources (Table 2). This dataset contained a diverse range of maximum tree heights and trunk diameters (Fig. 8). The main criteria for inclusion were: (i) the data was collected using a scanner capable of providing multiple returns and, (ii) multiple scans were collected around each harvested tree to minimise occlusion. Importantly, these data originated from various studies, each conducted with different survey configurations, including variations in scanners and settings. RayExtract was run on each tree estimating the taper from the point data using default parameters as shown in Table 1.

Destructive harvest statistical analysis. To assess the accuracy of RayExtract volume outputs, reconstructed volumes were compared to

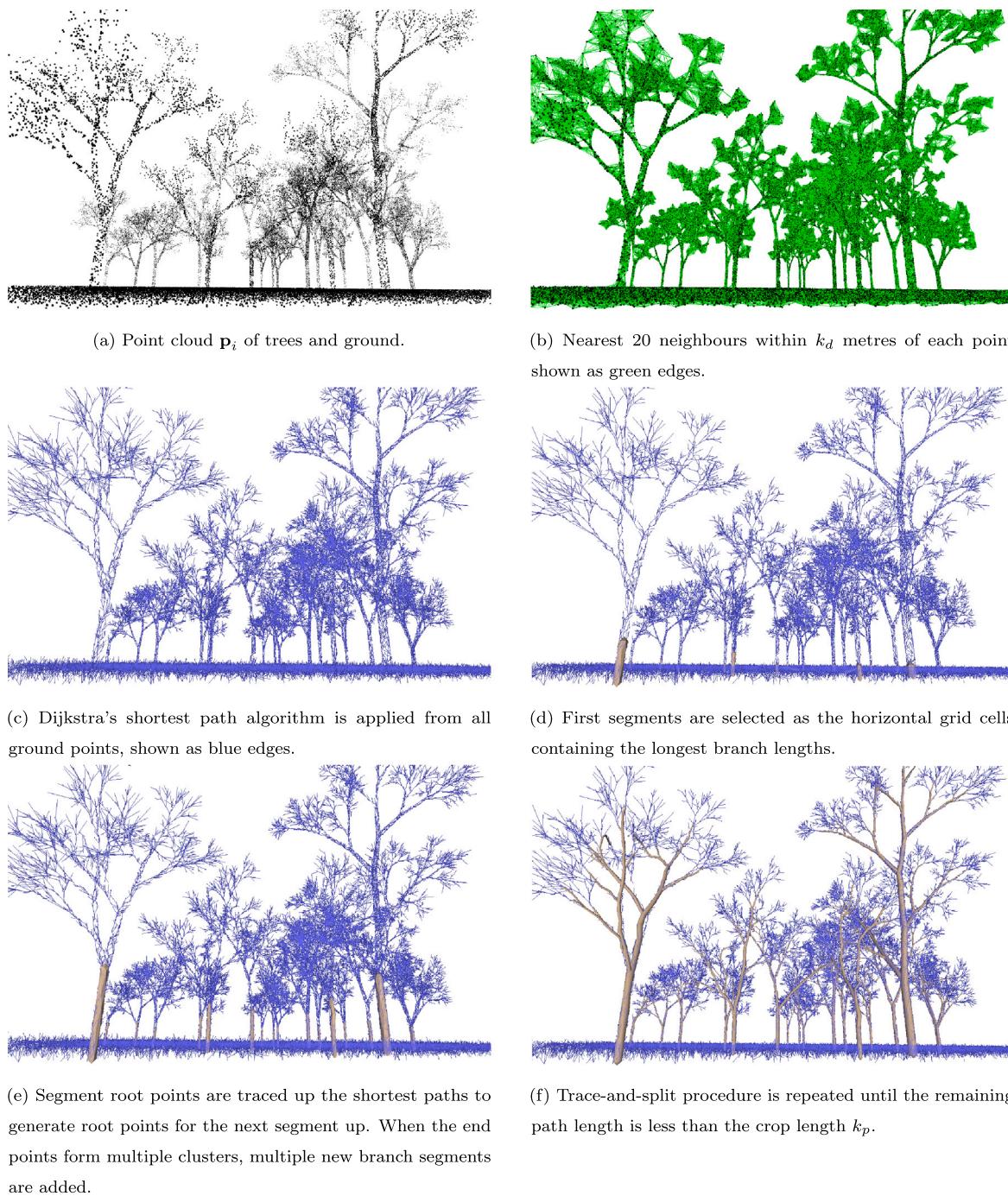


Fig. 7. Example of the RayExtract algorithm pipeline on a simulated forest point cloud, complementary to method workflow in Fig. 1.

Table 2

Summary of the destructive harvest reference data used in TLS studies dealing with biomass validation and in the present study.

Study	Country	Forest type	Tree type	Leaf conditions	TLS device	Sampling	Sampled trees	TreeQSM version
Gonzalez de Tanago et al. (2018)	Peru, Guyana, Indonesia	Tropical rain forest	Broadleaves	On	RIEGL VZ-400	MS per tree	29	v2.0
Burt et al. (2021)	Brazil	Tropical rain forest	Broadleaves	On	RIEGL VZ-400	MS per tree	4	v2.3.2
Demol et al. (2021)	Belgium	Temperate decidual	Broadleaves, Conifers	Off, On	RIEGL VZ-400, VZ-1000	20 m regular grid	65	v2.3.0
Lau et al. (2019)	Guyana	Tropical rain forest	Broadleaves	On	RIEGL VZ-400	MS per tree	26	v2.2.3

*MS = Multiple scans.

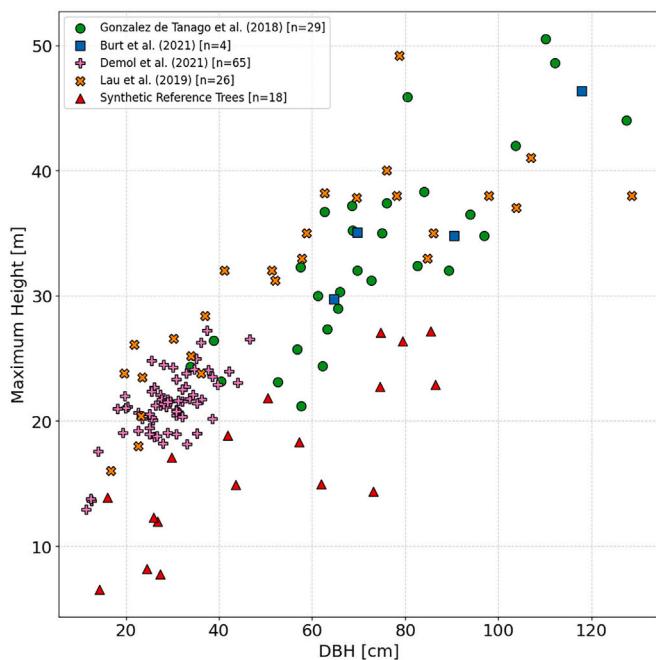


Fig. 8. Distribution of Diameter at Breast Height (DBH) and Maximum Height for 142 trees, comprising 124 destructively harvested specimens from four studies and 18 synthetic reference trees. Destructively harvested trees range from 11.5 to 128.7 cm in DBH and 12.92 to 50.50 m in height, while synthetic trees span 14.28 to 86.50 cm in DBH and 6.56 to 27.14 m in height.

the reference volumes derived from destructive harvest. The strength of the linear relationship between the reconstructed and reference volumes was quantified using the coefficient of determination (R^2) and the concordance correlation coefficient (CCC). The comparison of total volume from synthetic data used the same approach.

2.3. Evaluation of algorithm behavior using a simulation framework

This section details the methods used to simulate TLS point clouds, derive reference volumes, and evaluate QSM volumetric accuracy using synthetic trees. The key steps covered are: (1) description of the 3D digital reference tree models; (2) outline of TLS simulation methods, including scanner type, survey approach, and derivation of leaf-on/leaf-off point clouds; (3) introduction of a voxel-based method for precise calculation of reference woody volume and vertical profiles; and (4) description of the statistical analysis used to compare estimated volumes from simulated point clouds with the reference model volumes.

3D digital reference tree models. The synthetic reference dataset consisted of 18 tree models exhibiting diversity in height and diameter at breast height (DBH) (Fig. 9), representing the genera *Platanus*, *Eucalyptus*, and *Acacia*. These high-fidelity models were sourced commercially from the Evermotion Archmodels Vol. 238 collection (Evermotion, 2024). This collection features models originally generated by Evermotion utilising the GrowFX plugin for 3ds Max. Their process involved an initial generation using spline curves and bevel properties, which was then followed by manual editing guided by reference photographs to enhance realism and avoid reliance on purely predefined parametric constraints. This resulted in complex and realistic geometries suitable for mitigating the circularity that can arise when validating a parametric reconstruction method with purely parametric reference models.

TLS simulation. Synthetic TLS data were generated using the Helios++ lidar simulation framework (Winiwarter et al., 2022). The

scanner settings were configured to replicate those of a RIEGL VZ-400 terrestrial scanner operating at 300 KHz, as shown in Appendix Table A.4. This scanner was chosen as it is commonly used in TLS vegetation capture due to its high accuracy at long range and its ability to provide multiple returns per pulse (Wilkes et al., 2017). Additionally, all the destructive harvest TLS datasets used for comparison in this study were collected with this scanner or the similar RIEGL VZ-1000 (Table 2). To quantify the effects of scanner distance and sampling density on the volumetric reconstructions, the simulations were classified into close and far survey configurations (Fig. 10). The close simulations include four scan positions arranged in a square configuration with 20 m sides, with each synthetic tree positioned centrally. This resulted in each position being approximately 10 m from the target tree. Each scan position involved both a vertical and a 90° tilted scan. Tilted scans are a requirement for taller trees at this distance due to the 100° horizontal (-40°–60°) field of view of the RIEGL VZ-400. The far simulations involved 4 positions with both upright and tilted scans in a square configuration with 40 m sides, meaning the positions were approximately 20 m from the target tree. Ten simulation iterations were completed for each tree, with the tree model rotated on the z-axis by a random value between 0° and 360° on each iteration. The simulations were configured to preserve full waveform noise resulting from beam divergence. An example of a synthetic tree mesh, simulated TLS and RayExtract reconstruction is shown in Fig. 11.

For each simulation type, close and far, RayExtract was applied to simulated point clouds to quantify the effects of leaf-on and leaf-off conditions on the final volume estimates: the first used all points as input, including leaf points, while the second used only points derived from woody components. Leaf and wood returns were separated in each simulation output using a known classification generated from the material ID within the tree mesh file Fig. 11(b).

Voxel derived volume measurement. To achieve precise volumetric measurements and enable direct comparison of volume distribution along the vertical tree profile, a voxel-based approach using a 1 cm resolution grid was implemented. Smaller voxels lead to reduced volume underestimation at mesh boundaries; therefore, a voxel size of 1 cm³ was selected as an optimal balance between accuracy and computational efficiency. This method superimposes a regular fine-scale voxel grid over each tree mesh and employs raycasting in multiple directions from the centre of each voxel. The total enclosed volume within each mesh is determined by aggregating voxels where all raycasting hits are recorded, effectively capturing the internal structure of the tree. Fig. 12 illustrates this concept, showing the centres of interior voxels (grey points) within a reference tree mesh (green edges). While this approach may slightly underestimate volume due to edge effects, it was chosen over direct measurements from structural models for its simplicity and universal applicability to any 3D mesh. This technique allows for consistent volume comparisons between reference tree meshes and RayExtract QSM reconstructions. When focusing solely on volume analysis, this approach is preferable to directly comparing branch segments, as it effectively manages errors arising from unreliable branch segment matches between reconstruction and reference.

Volume distribution by height. Volume profiles were calculated for all synthetic reference trees and RayExtract output meshes derived from simulated point clouds. Voxel grids within mesh interior space were computed for each reference mesh and RayExtract output meshes as per Section 2.3. Total volume by height was calculated by summing the voxel volumes within each 0.2 m height bin. The error was calculated as the difference between the reconstructed volume and the reference tree volume in each height bin (in cubic metres, m³).

Simulation total volume statistical analysis. For the comparison of simulated outputs to the synthetic reference, the statistical analysis methods used for total tree volume, as described in Section 2.2.1, were applied.

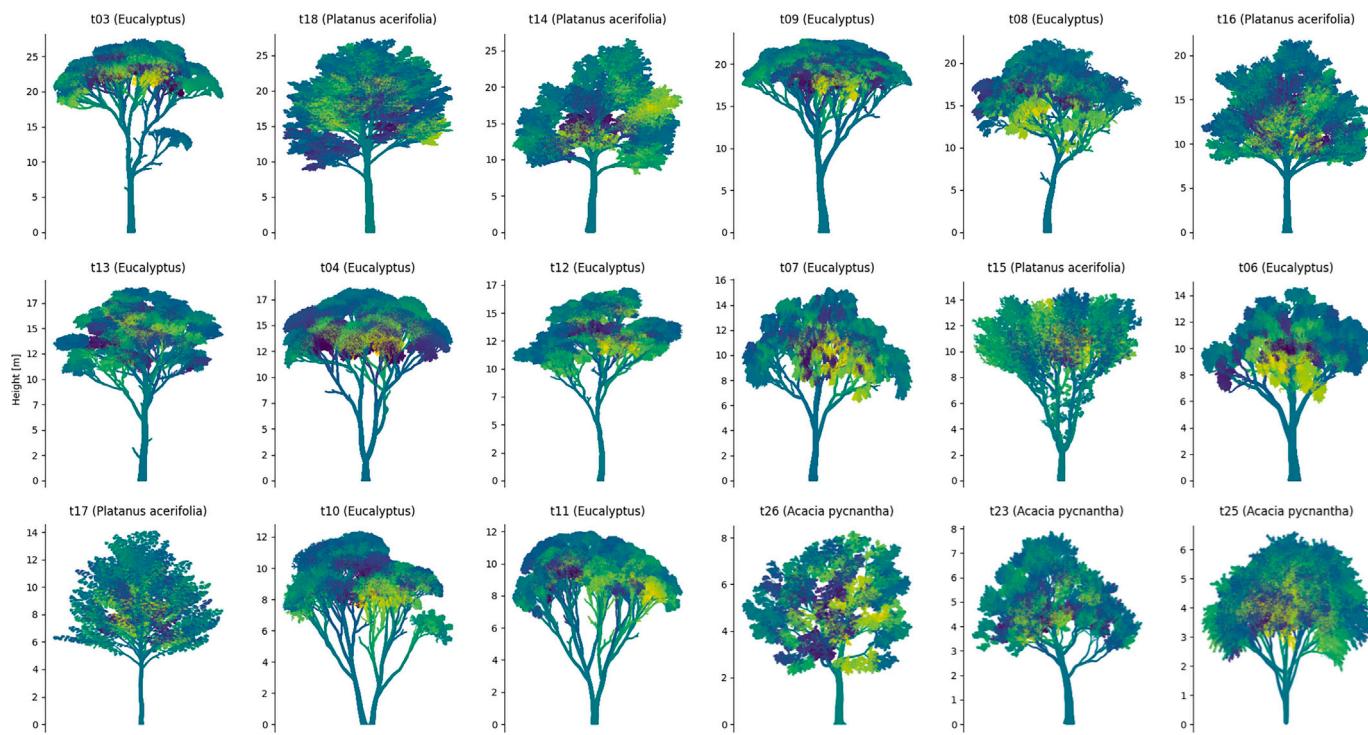


Fig. 9. The 18 synthetic trees used in TLS simulations sorted by height (bottom right to top left). The mesh vertices are coloured by distance along the point cloud Y axis to aid visual contrast. Each model's species is given in the subtitle.

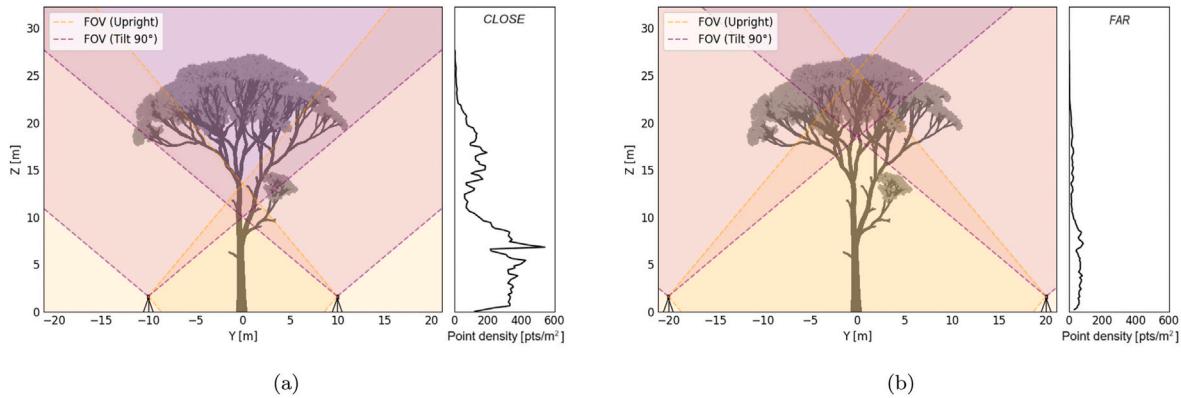


Fig. 10. Impact on point cloud density for a synthetic tree (t03) in close and far range scanner positioning, at 10 m (a) and 20 m (b) from the tree trunk (tripods in graphs). Coloured regions represent scanner field of view (FOV) of upright (orange) and 90° (purple) tilt angles. Side plots show the point density per unit of woody surface area along the tree height (points/m²), demonstrating how scanner distance and configuration (angle) affect the digitisation of the tree structure. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Although, total volume was derived using the previously described interior voxel method, rather than from the QSM cylinders. This analysis was conducted for simulations under both leaf-on and leaf-off conditions, as well as for close and far survey configurations.

2.4. Algorithm efficiency

To evaluate the performance characteristics of RayExtract (Objective #3), we conducted a comprehensive analysis of memory consumption and runtime efficiency. Input size was quantified using two metrics: the number of trees derived from inventory reference data and the total number of points in the point cloud. Point clouds were acquired for one-hectare reference plots of the Terrestrial Ecosystem Research Network (TERN), in Australia: Calperum, a Mallee system (Levick et al., 2024a); Litchfield, a tropical savannah (Levick et al., 2024b), and; Tumbarumba,

a Wet Eucalypt system (Levick et al., 2023). Scanning was performed using a RIEGL VZ-2000i terrestrial laser scanner operating at a 600 kHz laser repetition rate with a 0.03 mrad angular spacing. Scan positions were regularly spaced at 10 metres at the Tumbarumba site, and 20 metres at the Calperum and Litchfield sites. Following acquisition, scans were registered and downsampled using RIEGL RiSCAN PRO software. A 3 cm regular voxel grid resolution was applied to the point cloud for standard density. Point coordinates were encoded as 32-bit floating-point values. All processing and performance benchmarking were executed on a desktop computer running Fedora Linux 42 (Kernel 6.14.0), equipped with an Intel Core i7-14700 K CPU (28 logical cores up to 5.60 GHz), 128 GB of RAM, and utilising an NVMe SSD (Gen4 PCIe, rated up to 5,150 MB s⁻¹ read/write speed) for primary storage and processing. Computational resource usage (memory, CPU time) for processing each site was tracked and summarized (Table 3).

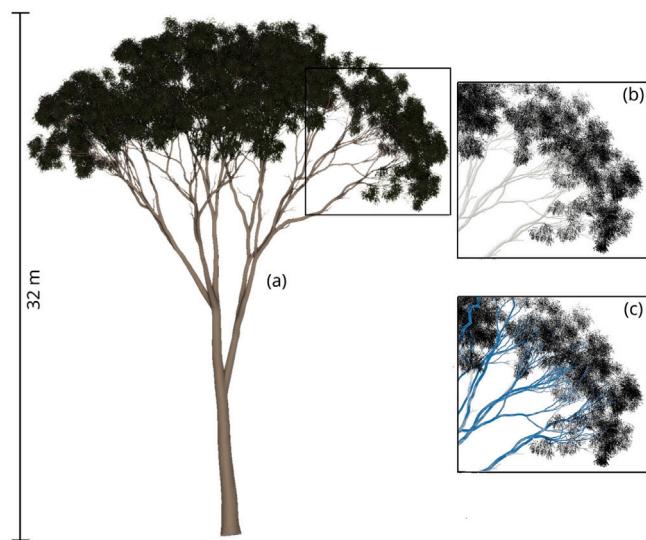


Fig. 11. Visualisation of a synthetic Eucalyptus tree model and its reconstructions. (a) Full view of the synthetic reference Eucalyptus tree (t05), showing detailed branch structure and foliage. (b) Close-up of a simulated TLS point cloud, depicting both wood (grey) and leaf points (black) with high detail. (c) Corresponding reconstruction using RayExtract run on the entire point cloud, highlighting the method's ability to identify and reconstruct woody structures (blue) from the combined wood and leaf point input. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

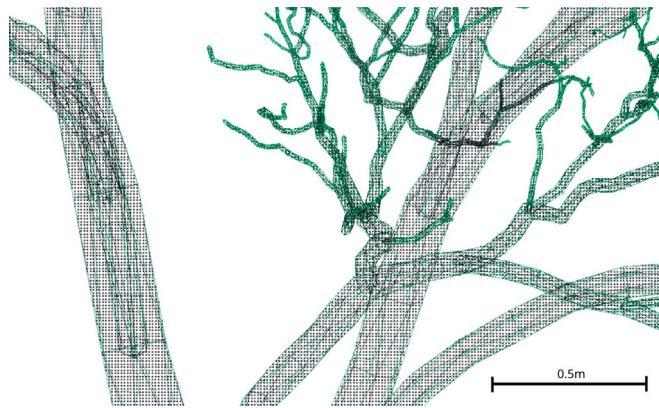


Fig. 12. Visualisation of the voxel-based volume measurement technique. The figure shows a reference tree mesh (green edges) overlaid with the centres of interior voxels (grey points). Each grey point represents the centre of a 1 cm³ voxel determined to be inside the tree volume through raycasting. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Results and discussion

3.1. Destructive harvest

Comparison of RayExtract's total tree volume output against destructive harvest reference datasets demonstrates strong agreement across multiple studies (Fig. 14). The analysis reveals high coefficients of determination (R^2) ranging from 0.74 to 0.93 and concordance correlation coefficient (CCC) values from 0.89 to 0.97. This indicates robust linear relationships between harvested and reconstructed volumes for the four reference datasets tested. These results are particularly noteworthy given that, with the exception of Demol et al. (2021), RayExtract was run on all point data including leaf points. A comparison of RayExtract

and TreeQSM woody volume published by each study is also shown in Fig. 14, demonstrating RayExtract's superior agreement with reference data in all studies except Burt et al. (2021). It should be noted that each reference study employed different versions of TreeQSM (Table 2) and that TreeQSM has been shown to produce different results between versions (Morales and MacFarlane, 2024). Additionally, each study used different methods for leaf point removal, a prerequisite for processing leaf on point clouds with TreeQSM.

Looking into each reference dataset, the highest R^2 of 0.93 (CCC: 0.97) was achieved in Gonzalez de Tanago et al. (2018) with 29 trees. The best-fit line ($y = 0.99x$) closely matched a perfect linear trend, indicating minimal bias. A R^2 of 0.93 (CCC: 0.96) for a larger sample of 65 trees was achieved for Demol et al. (2021), demonstrating consistent accuracy across a more extensive dataset. Results for Lau et al. (2019) showed strong performance with an R^2 of 0.87 (CCC: 0.94) for 26 trees, though with a slight overestimation trend ($y = 1.11x$), particularly for larger trees. Results for Burt et al. (2021), with the smallest sample of 4 trees, achieved an R^2 of 0.74 (CCC: 0.89), suggesting more variability but still maintaining good agreement. The data from Burt et al. (2021), although limited, suggests a tendency to underestimate volumes ($y = 0.82x$) in larger trees. Notably, some individual trees deviate significantly from the trend lines such as MDD04_012 (Gonzalez de Tanago et al. 2018), CAXH_T3 (Burt et al. 2021), and GUY_100_17 (Lau et al. 2019), and are illustrated in Fig. 15. These outliers represent challenging cases and are discussed in Sections 4.1 and 4.2.

3.2. Simulation framework analysis

RayExtract reported very similar concordance correlation coefficient (CCC) values of 0.99 across varying leaf conditions and point densities (Fig. 16). This suggests that RayExtract is robust to differences in leaf conditions and point cloud resolution when estimating tree volume. Such features are considerable advantages of RayExtract compared to other QSM algorithms, as leaf point classification is time consuming and can introduce bias in QSMs outputs (Ali et al., 2024).

The synthetic tree volume profiles, derived from a voxel-based volume measurement method (Section 2.3), offer a detailed view of reconstruction accuracy along the vertical profile of each tree (Fig. 18). The comparison between leaf-on and leaf-off reconstructions provides crucial information about RayExtract's performance under different foliage conditions. In most cases, leaf-off reconstructions showed smaller errors compared to leaf-on scenarios, particularly in the upper canopy regions. This difference was more pronounced in trees with denser canopies, such as *Platanus acerifolia*, suggesting that leaf occlusion affects reconstruction accuracy. However, the relatively small difference between leaf-on and leaf-off results for most trees indicates that RayExtract maintains robust performance even in the presence of foliage. Our analysis revealed consistent patterns in error distribution along tree heights. Many trees showed larger errors in the mid-canopy region, where branching complexity is typically highest. Lower trunk areas generally exhibited smaller errors, likely due to their simpler geometry and better visibility in scans. Upper canopy regions often showed increased errors, possibly due to occlusion and lower point density in these areas.

RayExtract demonstrates a tendency to underestimate branch volume in higher branching orders, a phenomenon most prominently observed in trees t16 and t09 (Fig. 17). This underestimation is likely due to three factors. One factor is the over-identification of branching segments, which influences the radius estimation of each segment. The influence of Leonardo's Rule, which states that the sum of the cross-sectional areas of a tree's child branches is equal to the cross-sectional area of the parent branch, can lead to reduced radius estimations for all child branches when branches are incorrectly segmented and the total number of branches is not accurately measured. The second factor is that the reference tree may have more than a realistic number of child branches per segment. For these reasons, the branching architecture calculated (i.e., the number of segments, and their



Fig. 13. Representative images of the TERN reference sites where TLS data was collected for computational performance analysis: Robson Creek Tropical Rainforest, Tumbarumba Wet Eucalypt, Litchfield Tropical Savannah, and Calperum Mallee Woodlands.

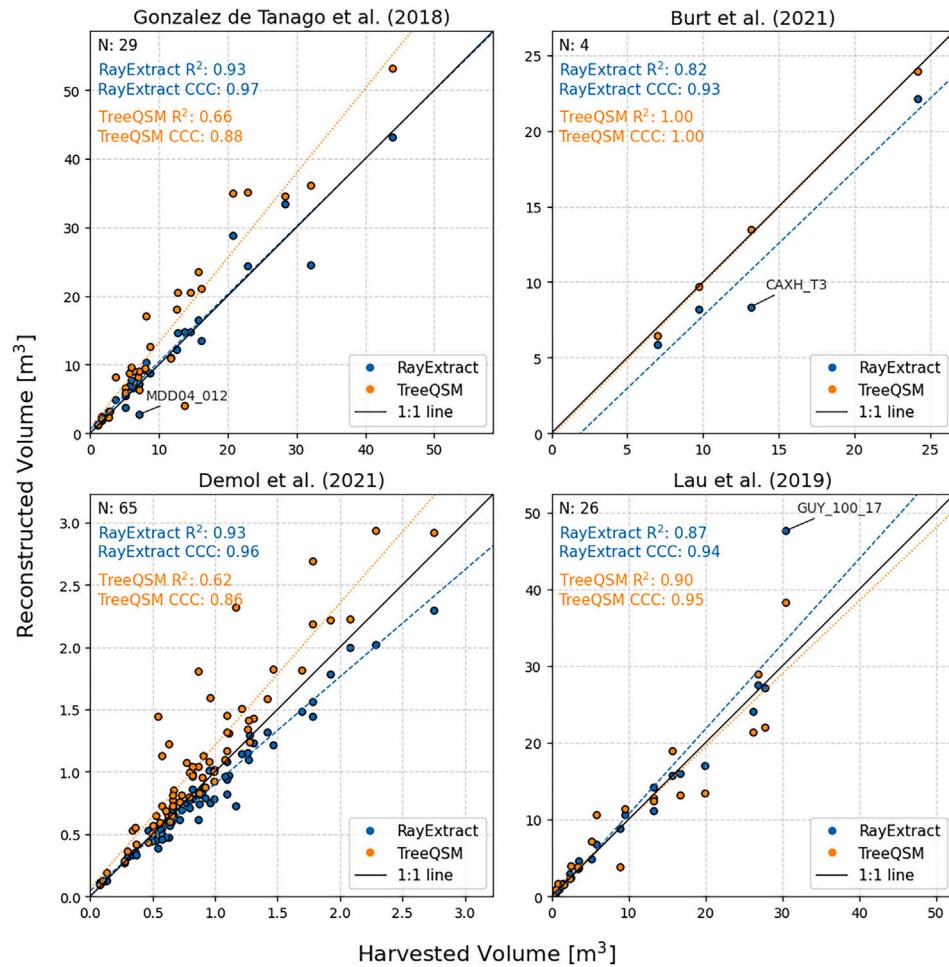


Fig. 14. Comparison of reconstructed vs. harvested tree volume for RayExtract and TreeQSM methods across four studies. Scatter plots show reconstructed volume vs. harvested volume for individual trees, with performance metrics (R^2 and CCC) provided for each method and study. The black diagonal line represents perfect agreement between reconstructed and harvested volumes. Annotated points indicate notable outliers discussed in the text. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

branching order) would likely be incorrect in these cases; the validation of the branching architecture is challenging and thus suggested as future work. The third factor contributing to volume underestimation is that the taper factor estimated from the point cloud data is too large. This overestimation of branch taper results in a more rapid decrease in branch diameter along its length, further reducing the

calculated volume of higher-order branches. Overall, these results collectively demonstrate RayExtract's robustness in estimating tree volumes across a range of tree species and structural complexities. However, they also highlight specific challenges, particularly in accurately reconstructing complex branching structures and accounting for significant stem taper.

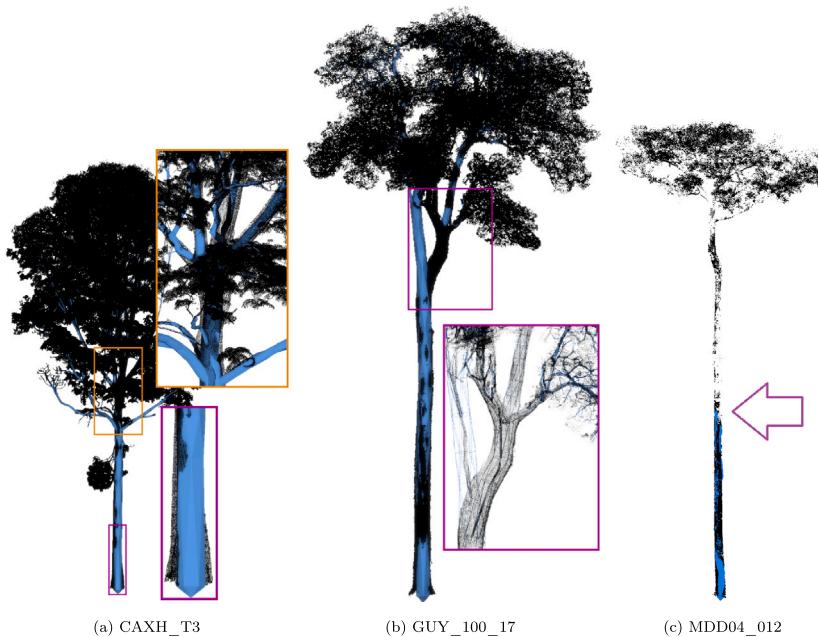


Fig. 15. Reference tree point clouds with identified issues in the QSMs modelled with RayExtract: (a) low volume from underestimation in buttress and larger branches (Burt et al., 2021, CAXH_T3); (b) overestimation of small branches coming from the main stem (Lau et al., 2019, GUY_100_17), and; (c) underestimation of volume due to large occlusion on main stem preventing entire reconstruction (Gonzalez de Tanago et al., 2018, MDD04_012). Tree reconstruction issues highlighted with auxiliary coloured panels.

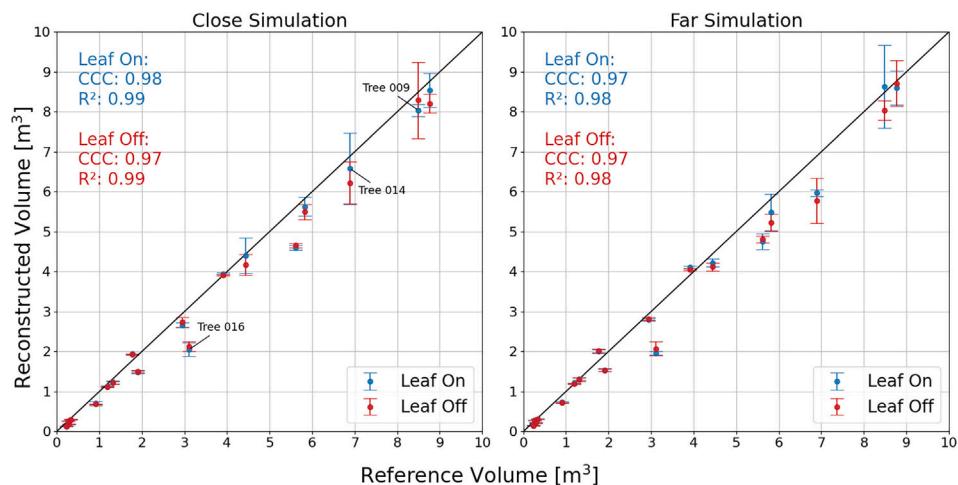


Fig. 16. Comparison of reconstructed volume versus reference volume for close and far simulations under leaf-on and leaf-off conditions. The plots show the correlation between reconstructed and reference volumes for each reference tree, with CCC and R^2 values provided for each scenario. The dashed line represents a 1:1 relationship. Error bars indicate measurement uncertainty. Both close and far simulations demonstrate high accuracy in volume reconstruction, with slight variations between leaf-on and leaf-off conditions. Annotated points indicate notable outliers which vertical profiles are visualised in Fig. 17.

3.3. Algorithm efficiency

We evaluated the performance characteristics (runtime efficiency and memory consumption) of the RayExtract algorithm (Objective #3) using TLS datasets acquired from three structurally distinct plots: a mallee woodland, tropical savannah, and wet eucalypt forest (Fig. 13). These sites provided a valuable gradient in input data complexity, measured both by the number of inventoried trees and the total point count within the processed point clouds (Table 3).

The results clearly indicate that the computational demands of RayExtract scale with the complexity and size of the input point cloud data. As the point count and corresponding file size increased across the sites – from Calperum (68 trees; 20.5 million points; 0.91 GB) to

Litchfield (492 trees; 85.1 million points; 3.81 GB), and peaking at Tumbarumba (383 trees; 132 million points; 5.92 GB) – both the required runtime (\approx 5, 6, and 11 min, respectively), and peak random access memory usage (15.46, 37.94, and 56.95 GB, respectively) showed a corresponding increase.

Comparing the Litchfield and Tumbarumba sites reveals that point count and data size are directly affecting RayExtract's performance, although the tree count discrepancy between datasets (492 vs 383 trees, respectively). The Tumbarumba dataset, possessing roughly 55 % more points, required substantially more processing time (\approx 90 % longer runtime) and memory (\approx 50 % more RAM). This higher point count reflects the greater overall structural complexity within the wet eucalypt forest (e.g., more intricate branching patterns leading to a higher number

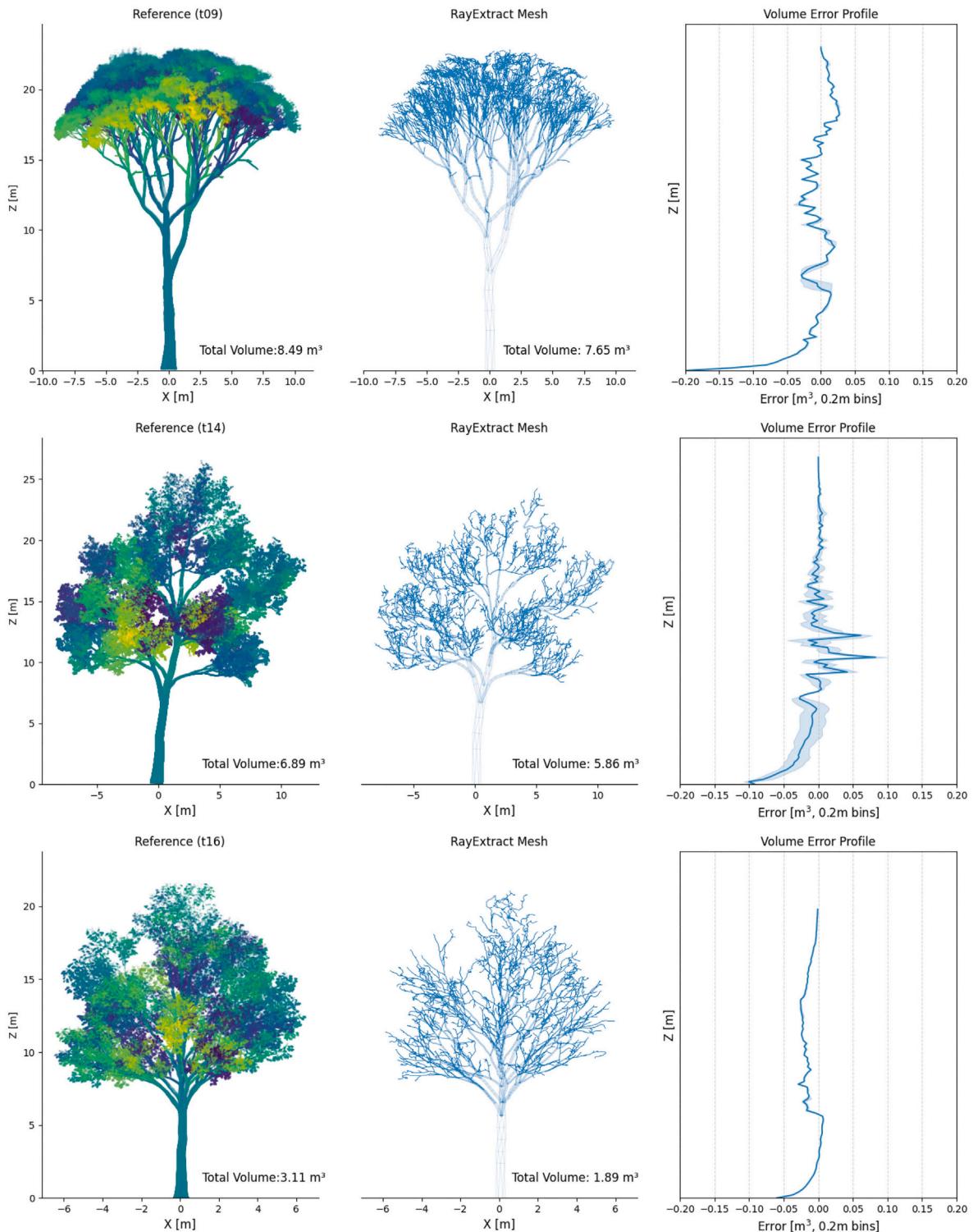


Fig. 17. Three outlying trees: t09, t14, and t16. For each tree, three visualisations are provided: the synthetic reference with total volume (mesh vertices are coloured by distance along the point cloud Y axis to aid visual contrast), the RayExtract-generated reconstruction with its corresponding total volume and a volume error profile showing reconstruction accuracy along the tree's vertical axis.

of detectable stem and branch segments in the data) compared to the tropical savanna.

It is likely this increased structural complexity, represented by more tree segments and branches within the point cloud, rather than point count in isolation, that places higher computational demands

on the RayExtract algorithm, leading to increased runtime and memory usage. The structurally simpler Calperum mallee site, with the lowest tree and point counts, exhibited the lowest computational demand, and illustrated the need for at least five minutes per hectare.

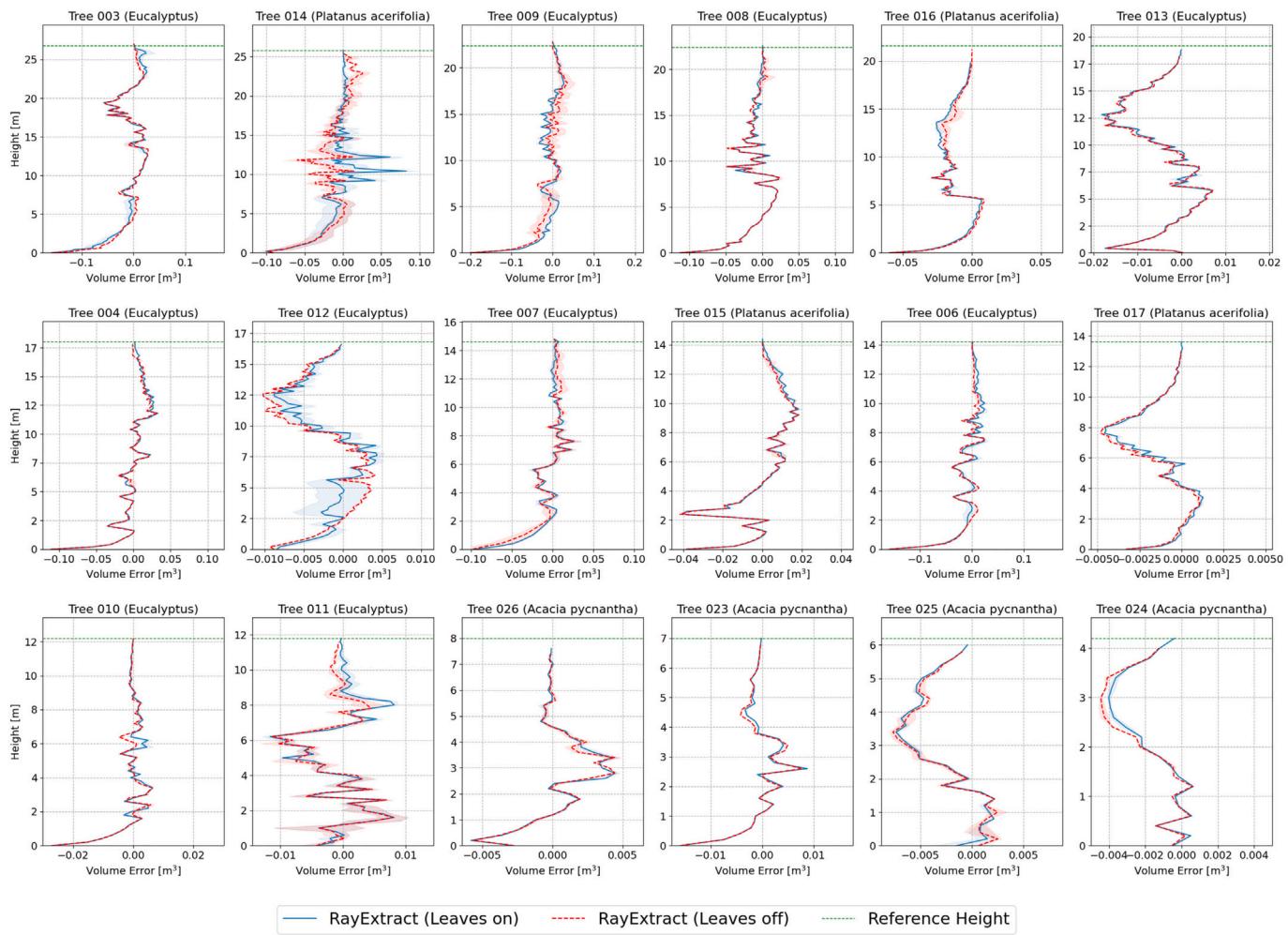


Fig. 18. Volume error profiles of synthetic trees reconstructed using both leaf-on and leaf-off point clouds in the close scanning configuration. The error is calculated as the difference between the reconstructed volume and the reference tree volume (in m^3) in each height bin (0.2 m). Axis limits are variable for each subplot.

Table 3

RayExtract processing performance metrics for TLS data of one-hectare plots (TERN sites, Australia) from input point clouds to output QSMs of the tree populations.

Site	Vegetation	Tree Count	Point Count (million)*	File size (GB)	Runtime (min:s)	RAM (GB)
Calperum	Mallee	68	20.5	0.91	5:21	15.46
Litchfield	Tropical savanna	492	85.1	3.81	5:56	37.94
Tumbarumba	Wet eucalypt	383	132.0	5.92	11:19	56.95
Robson Creek	Tropical Rainforest	1058	260.0	12.58	23.55	110.28

*Point clouds were projected and downsampled using regular voxels with 3 cm side lengths. A regular grid-sampling strategy was employed with 10 m steps at Tumbarumba and 20 m steps at Litchfield and Calperum sites. Robson Creek used 10 m steps with an additional tilted scan configuration.

4. Limitations and future work

4.1. Volumetric errors

The volume underestimation particularly evident in Fig. 15(a) (tree CAXH_T3), can be partially attributed to RayExtract's current limitations in accurately modelling buttress roots. These buttress roots, characteristic of numerous tropical tree species, present a complex structural challenge that is not easily captured by the cylinder primitives currently employed in the volume estimation approach. To address this, a specialised meshing method, similar to the approach in TreeQSM v2.4.1 (Raumonen, 2022), is proposed. This would be an optional step, enhancing the representation of buttressed trees without affecting performance

on others. To complement this, an automated filtering mechanism using machine learning (e.g., CNNs trained on diverse tree forms) could identify trees requiring this specialised meshing. Combining improved buttress modelling with automated detection could significantly enhance volume estimation accuracy across challenging tropical morphologies, aiding large-scale inventories in species-rich rainforests. Validating these proposed improvements will be a focus of future work.

As illustrated in Fig. 15(b), using tree GUY_100_17 as an example, the algorithm's misinterpretation of certain vertical structures can lead to significant volumetric errors. In this case, a small auxiliary limb extends vertically alongside the main stem, maintaining partial connectivity with the upper canopy. The reconstruction algorithm erroneously

identifies this vertical limb as a continuation of the main stem, while incorrectly classifying the actual stem as a child branch. This misclassification triggers an inappropriate application of taper rules, particularly Leonardo's rule, to the true main stem. Consequently, the algorithm reduces the radii of all subsequent child branches emanating from the actual main stem, propagating the error throughout the upper portions of the tree structure. This cascading effect results in a substantial underestimation of branch volumes in the upper canopy, highlighting the algorithm's current limitations in distinguishing between vertical main stems and parallel ascending branches. Future work should prioritize establishing quantitative relationships between morphological complexity and model accuracy to better predict and mitigate reconstruction errors in complex tropical tree forms.

4.2. Point cloud quality

Due to variations in the conditions under which each reference data set was obtained, pinpointing the exact sources of bias presents a challenge. Factors such as occlusion (both within the canopy and of the ground), noise, and point density significantly influence the accuracy of the output, as established in previous studies (Abegg et al., 2023; Morhart et al., 2024). These variables can affect the quality of the point cloud data, leading to potential inaccuracies in volume estimation. These effects are most prevalent within the data from Gonzalez de Tanago et al. (2018, tree MDD04_012, Fig. 15(c)).

A practical concern arises regarding ground coverage, particularly in common field scenarios employing fewer scan positions, which can lead to occlusion in the ground layer, or when the ground is obscured by undergrowth. While RayExtract does not have explicitly defined minimum requirements for ground point density, achieving adequate ground characterization is crucial for optimal performance. Following established acquisition protocols, like those recommended by (Wilkes et al., 2017), helps mitigate these issues by promoting sufficient ground sampling even with practical limitations on scan setups, thereby ensuring more reliable results. Furthermore, employing scanners capable of recording multiple returns per pulse (i.e. ≥ 4) remains recommended to enhance data capture, potentially penetrating undergrowth and characterizing complex structures more effectively. Issues such as inconsistent point cloud quality give further relevance to the use of synthetic data for benchmarking TLS methods. Despite these challenges, the accuracy of RayExtract across diverse real-world data sources (Fig. 14), often captured under such variable conditions, underscores its effectiveness and robustness.

Bohn Reckziegel et al. (2025) tested RayExtract in a one-hectare tropical savanna digitised by two distinct platform-sensor combinations providing high-density point clouds: a modern drone-LiDAR system and the standard TLS (similarly to the reference studies). Under facilitated understory conditions (after fire), data qualities differed in total vegetation points ($>$ in TLS) while ground points were fairly proportionate between datasets. The overall tree count was 15 % higher with TLS data, and occlusion at the stem and ground interface has aided the segmentation errors within the drone data. Though the different nature of the datasets (i.e. viewpoints, static versus mobile), trees with higher dimensions (>12 cm) were equally segmented and reconstructed, while the stand volumetric mismatch deviated only by 1 %, and mean stand metrics such as DBH and tree height agreed within 3 %. The study of (Bohn Reckziegel et al., 2025) indicated that RayExtract may also work with high-resolution drone-LiDAR data for the characterization of the overstorey, without major structural losses, in open forests.

4.3. Synthetic reference data

While synthetic trees offer controlled evaluation advantages, the models used here have limitations. Firstly, the specific generation methodology, including photo-guided refinement, remains unpublished. Secondly, although chosen to reduce circularity compared to purely parametric references, using any simulated data risks methodological

circularity if the generation software (GrowFX) and the tested algorithm (e.g., RayExtract) share underlying principles. This potential overlap, though mitigated by manual refinement, means simulation-based performance might be slightly overestimated compared to real-world tests. Consequently, these simulations, while useful for specific tests (like leaf-on/off), are less suited for comprehensive sensitivity analysis of QSM responses to diverse branching characteristics (counts, lengths, curvatures, taper). More sophisticated generative models offering explicit control over architectural parameters are needed. Such advanced synthetic datasets would enable more rigorous testing of QSM performance and sensitivity across a controlled range of tree forms, complementing destructive and field validation, and ultimately advancing QSM reliability.

4.4. Improving branch radius constraints

Earlier research into tree morphology revealed the complex interplay between physical and environmental factors that influence branch radius patterns. Edder (1946) provided a foundational theoretical framework that focused on the tree trunk as an optimal mechanical support, postulating that diameter distributions follow predictable patterns based on maintaining constant maximum compressive stress across all cross-sections under combined gravity and wind loads. King and Loucks (1978) examined both bole and branch morphology through the lens of optimal design, hypothesizing that branch radius allocation minimizes wood use while providing necessary structural support. While RayExtract currently employs simpler constraints such as self-similarity and Leonardo's rule for radius estimation from point clouds, these earlier works suggest more sophisticated morphological models incorporating mechanical support, wind resistance, and resource allocation could further refine our approach. This is particularly relevant as our current radius constraints, while effective at preventing unrealistic branch expansion, may be sensitive to initial stem selections and could potentially underperform in dense forest scenarios. Future iterations of RayExtract could benefit from less parametric, more adaptive models, potentially addressing current limitations in distinguishing between closely positioned branches while reducing parameter tuning requirements.

4.5. Advancing branch-level accuracy assessment

While our destructive harvest comparison provides strong evidence for total volume accuracy, we acknowledge several fundamental limitations. First, destructive harvest data typically provides only whole-tree measurements without detailed branch-specific references—a constraint common across TLS-based reconstruction studies that restricts our ability to validate individual structural components. Our complementary use of synthetic tree simulations partially addresses this limitation by enabling analyses such as vertical error profiling, yet we recognize that the synthetic dataset used here has limited scope.

More broadly, our harvest reference data may not adequately capture the full range of structural complexities found in global forest ecosystems. Additionally, while Section 2.1 emphasizes our novel integration of morphological rules, we have not quantified the independent contribution of these components to reconstruction accuracy.

Future work should prioritize developing more comprehensive algorithm assessment frameworks through several approaches: (1) exploring methods for obtaining detailed field measurements of branch geometry where operationally feasible, (2) conducting ablation studies that systematically remove specific algorithmic components (e.g., Leonardo's rule, species-specific taper corrections) to isolate their individual contributions to accuracy, and (3) expanding datasets to encompass greater structural and taxonomic diversity representative of global forest conditions.

4.6. Improving input data quality and preprocessing

The reliability of any tree reconstruction algorithm, including RayExtract, is fundamentally linked to the quality of the input point

cloud data, which is attributed to the completeness of the samples. Future work should focus on developing techniques to better quantify point cloud quality, using occlusion mapping or sampling completeness assessments (Zhang et al., 2024), and systematically exploring the effects of various filtering methods and noise reduction algorithms on the precision and reliability of QSM-derived woody volume. This is crucial for understanding and mitigating sources of error originating from the initial data capture and processing steps.

4.7. Enhancing segmentation accuracy and error propagation

A critical aspect of real-world application is the accurate segmentation of individual trees from complex plot-level point clouds. Future research will investigate the sensitivity of RayExtract's volume (and consequently, biomass) estimations to segmentation errors, particularly in diverse and dense forest stands. This involves quantifying the impact of under- and over-segmentation on final volume estimates, as illustrated in Bohn Reckziegel et al. (2025) for TLS and high-density drone-LiDAR data. Furthermore, we will explore methods for improving segmentation accuracy itself, potentially through advanced techniques such as machine learning-derived point weighting schemes (e.g., leveraging semantic segmentation probabilities) or incorporating multi-spectral information. The goal is to develop strategies that minimize the propagation of segmentation errors into volumetric and biomass estimates, especially at the plot scale. Additionally, terrain complexity represents another important consideration for plot-level applications, particularly on steep slopes where topographic effects can influence both data acquisition geometry and point cloud quality. While the current study focuses on individual tree reconstruction where these effects are minimized, future work extending RayExtract to full plot-scale assessments should address terrain-related challenges systematically.

4.8. Applications in digital twin forests and radiative transfer

The study demonstrated RayExtract's potential to streamline the generation of comprehensive digital twin forests, tasks previously requiring considerable manual effort (e.g., Wytham Woods; Calders et al. (2018)). Moreover, the applications of RayExtract and RayCloudTools go beyond the universe of TLS, as they have been used for LiDAR data from varied platform-sensor combinations (Bohn Reckziegel et al., 2025; Lowe and Stepanas, 2021). Integrating such high-resolution 3D reconstructions into radiative transfer models offers powerful opportunities. It allows for more detailed exploration and analysis of canopy structure, light interception, and energy flux predictions derived from various remote sensing data sources (Malenovský et al., 2023; Cimoli et al., 2024).

5. Conclusion

This study introduced RayExtract, a novel method for reconstructing woody volume from TLS data. The findings demonstrated that RayExtract is a fast and accurate method for individual tree volume measurement. By incorporating morphological rules, specifically Leonardo's Rule and the principle of self-similarity, RayExtract effectively reduced the common issue of overestimating volume in smaller branches found in current tree reconstruction methods. Comparison against destructively harvested data showed strong agreement, underscoring RayExtract's robustness across diverse forest types and data collection conditions. Furthermore, the synthetic assessment framework provided valuable insights into the algorithm's behaviour and performance across varying point cloud qualities and tree morphologies. RayExtract exhibited high degrees of accuracy, particularly in estimating volume distribution along the vertical profile, with notable performance in higher branching orders. RayExtract also demonstrated a high level of efficiency and scalability when applied to plot-scale datasets. This efficiency, coupled with its ability to operate on entire forest point cloud datasets without prior leaf-wood classification, positions RayExtract as a valuable tool for forest inventory and biomass estimation.

CRediT authorship contribution statement

Timothy Devereux: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Thomas Lowe:** Writing – review & editing, Software, Methodology, Conceptualization. **Joshua Rivory:** Writing – review & editing. **Rafael Bohn Reckziegel:** Writing – review & editing. **Kim Calders:** Writing – review & editing. **Raja Ram Aryal:** Writing – review & editing. **Glen Eaton:** Writing – review & editing. **Zane Cooper:** Data curation. **Shaun Levick:** Writing – review & editing, Supervision. **Stuart Phinn:** Writing – review & editing, Supervision. **William Woodgate:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

William Woodgate reports financial support was provided by the Australian Research Council. Timothy Devereux reports equipment, drugs, or supplies were provided by the National Collaborative Research Infrastructure Strategy. Kim Calders reports financial support was provided by the European Union. Timothy Devereux reports financial support was provided by the Commonwealth Scientific and Industrial Research Organisation. 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

Funding. TD received funding through a CSIRO postgraduate scholarship. SL and TL received funding through the CSIRO's Digital Water and Landscapes initiative (3D-AGB project). WW was supported by an Australian Research Council DECRA Fellowship (DE190101182). We acknowledge the support of the Terrestrial Ecosystem Research Network (TERN), supported by the National Collaborative Research Infrastructure Strategy (NCRIS). KC and ZC were funded by the European Union (ERC-2021-STG Grant agreement No. 101039795). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Source Code Availability. RayExtract is a component of the RayCloudTools library developed by the Robotic Perception and Autonomy Group within CSIRO Data61. Source code is available at <https://github.com/csiro-robotics/raycloudtools>. RayCloudTools is an open-source collection of C++ command line tools for processing ray cloud data, built upon the associated *raylib* library.

Appendix A. Appendix

An arbitrary set of points on a circle can be parameterised as $(c + r \sin \alpha, r \cos \alpha)$ for some c, r and for a set of angle values α . Here we show that raising these points to the paraboloid $z(x, y) = x^2 + y^2$ gives a set of points in a 3D plane.

The raised circle is therefore the set of points

$$p(\alpha) = (c + r \sin \alpha, r \cos \alpha, (c + r \sin \alpha)^2 + (r \cos \alpha)^2) \quad (\text{A.1})$$

We show that the points sit on the plane:

$$z = 2cx - c^2 + r^2 \quad (\text{A.2})$$

Proof. Substituting $p(\alpha)$ into this plane equation gives:

$$(c + r \sin \alpha)^2 + (r \cos \alpha)^2 = 2c(c + r \sin \alpha) - c^2 + r^2 \quad (\text{A.3})$$

Expanding out the parentheses:

$$c^2 + 2cr \sin \alpha + r^2(\sin^2 \alpha + \cos^2 \alpha) = 2c^2 + 2cr \sin \alpha - c^2 + r^2 \quad (\text{A.4})$$

Table A.4
Lidar scanner specifications set within Helios ++ .

Parameter	Value
Accuracy (m)	0.005
Beam Divergence (rad)	0.0003
Pulse Frequency (kHz)	300
Pulse Length (ns)	5
Minimum Range (m)	1.5
Max Vertical Scan Angle (deg)	60
Min Vertical Scan Angle (deg)	-40
Scan Frequency (Hz)	120
Beam Quality	3

Simplifying:

$$c^2 + 2cr \sin \alpha + r^2 = c^2 + 2cr \sin \alpha + r^2$$

(A.5)



Data availability

Data will be made available on request.

References

- Abdollahi, A., Yebra, M., 2025. Challenges and opportunities in remote sensing-based fuel load estimation for wildfire behavior and management: a comprehensive review. *Remote Sens.* 17, 415. <https://doi.org/10.3390/rs17030415>. <https://www.mdpi.com/2072-4292/17/3/415>.
- Aebegg, M., Bösch, R., Kükenbrink, D., Morsdorf, F., 2023. Tree volume estimation with terrestrial laser scanning — testing for bias in a 3d virtual environment. *Agric. For. Meteorol.* 331, 109348. <https://doi.org/10.1016/j.agrmet.2023.109348>. <https://www.sciencedirect.com/science/article/pii/S0168192323000424>.
- Ali, M., Lohani, B., Hollaus, M., Pfeifer, N., 2024. Benchmarking geometry-based leaf-filtering algorithms for tree volume estimation using terrestrial lidar scanners. *Remote Sens.* 16, 1021. <https://doi.org/10.3390/rs16061021>. <https://www.mdpi.com/2072-4292/16/6/1021>.
- Arkin, J., Coops, N.C., Daniels, L.D., Plowright, A., 2025. Canopy and surface fuel estimations using RPAS and ground-based point clouds. *For.: Int. J. For. Res.* 98, 15–28. <https://doi.org/10.1093/forestry/cpa020>. <https://academic.oup.com/forestry/article/98/1/15/7146018>.
- Binney, J., Sukhatme, G.S., 2009. 3d tree reconstruction from laser range data. In: 2009 IEEE International Conference on Robotics and Automation, pp. 1321–1326, <https://doi.org/10.1109/ROBOT.2009.5152684>. <https://ieeexplore.ieee.org/document/5152684>. ISSN: 1050-4729.
- Bohn Reckziegel, R., Lowe, T., Devereux, T., Johnson, S.M., Rochelmeyer, E., Hutley, L.B., Doody, T., Levick, S.R., 2025. Assessing the reliability of woody vegetation structural characterisation from UAV-LS in a tropical savanna. *Sci. Remote Sens.* 11, 100178. <https://doi.org/10.1016/j.srs.2024.100178>. publisher: Elsevier.
- Bonan, G.B., 2008. Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science* 320, 1444–1449. <https://doi.org/10.1126/science.1155121>. <https://www.science.org/doi/10.1126/science.1155121>.
- Burkhart, H.E., Tomé, M., 2012. Tree form and STEM taper. In: Burkhart, H.E., Tomé, M. (Eds.), *Modeling Forest Trees and Stands*. Springer Netherlands, Dordrecht, pp. 9–41. https://doi.org/10.1007/978-90-481-3170-9_2.
- Burt, A., Boni Vicari, M., Da Costa, A.C.L., Coughlin, I., Meir, P., Rowland, L., Disney, M., 2021. New insights into large tropical tree MASS and structure from direct harvest and terrestrial lidar. *R. Soc. Open Sci.* 8, rsos.201458, 201458. <https://doi.org/10.1098/rsos.201458>. <https://royalsocietypublishing.org/doi/10.1098/rsos.201458>.
- Burt, A., Disney, M., Calders, K., 2018. Extracting individual trees from lidar point clouds using treeeg. *Methods Ecol. Evol.* 2041–210X.13121. <https://doi.org/10.1111/2041-210X.13121>.
- Calders, K., Adams, J., Armston, J., Bartholomeus, H., Bauwens, S., Bentley, L.P., Chave, J., Danson, F.M., Demol, M., Disney, M., Gaulton, R., Krishna Moorthy, S.M., Levick, S.R., Saarinen, N., Schaaf, C., Stovall, A., Terryn, L., Wilkes, P., Verbeeck, H., 2020. Terrestrial laser scanning in forest ecology: expanding the horizon. *Remote Sens. Environ.* 251, 112102. <https://doi.org/10.1016/j.rse.2020.112102>. <https://linkinghub.elsevier.com/retrieve/pii/S0034425720304752>.
- Calders, K., Newham, G., Burt, A., Murphy, S., Raumonen, P., Herold, M., Culvenor, D., Avitabile, V., Disney, M., Armston, J., Kaasalainen, M., 2015. Nondestructive estimates of above-ground biomass using terrestrial laser scanning. *Methods Ecol. Evol.* 6, 198–208. <https://doi.org/10.1111/2041-210X.12301>. <http://doi.wiley.com/10.1111/2041-210X.12301>.
- Calders, K., Wilkes, P., Disney, M., Armston, J., Schaefer, M., Woodgate, W., 2018. Chapter 19. Terrestrial lidar for measuring above-ground biomass and forest structure. Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M.S., Delitti, W.B., Duque, A., Eid, T., Fearnside, P.M., Goodman, R.C., Henry, M., Martínez-Yrízar, A., Mugasha, W.A., Muller-Landau, H.C., Mencuccini, M., Nelson, B.W., Ngomanda, A., Nogueira, E.M., Ortiz-Malavassi, E., Pélissier, R., Ploton, P., Ryan, C.M., Saldarriaga, J.G., Vieilledent, G., 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. *Glob. Change Biol.* 20, 3177–3190. <https://doi.org/10.1111/gcb.12629>. <https://onlinelibrary.wiley.com/doi/10.1111/gcb.12629>.
- Cimoli, E., Lucieer, A., Malenovský, Z., Woodgate, W., Janoutová, R., Turner, D., Haynes, R.S., Phinn, S., 2024. Mapping functional diversity of canopy physiological traits using UAS imaging spectroscopy. *Remote Sens. Environ.* 302, 113958. <https://doi.org/10.1016/j.rse.2023.113958>. <https://www.sciencedirect.com/science/article/pii/S0034425723005102>.
- Cooper, Z.T., Terryn, L., Burt, A., Cherlet, W., Demol, M., Devereux, T., Lau, A., Raumonen, P., den Broeck, W.A.V., Verbeeck, H., Calders, K., 2025. Benchmarking quantitative structure modelling algorithms against destructively sampled aboveground biomass estimates. Manuscript submitted for publication.
- Côté, J.F., Luther, J.E., Lenz, P., Fournier, R.A., Van Lier, O.R., 2021. Assessing the impact of fine-scale structure on predicting wood fibre attributes of boreal conifer trees and forest plots. *For. Ecol. Manag.* 479, 118624. <https://doi.org/10.1016/j.foreco.2020.118624>. <https://linkinghub.elsevier.com/retrieve/pii/S0378112720313931>.
- Dahle, G.A., Grabosky, J.C., 2010. Allometric patterns in Acer Platanoides (Aceraceae) branches. *Trees* 24, 321–326. <https://doi.org/10.1007/s00468-009-0401-5>. <http://link.springer.com/10.1007/s00468-009-0401-5>.
- Demol, M., Aguirre-Amuchastegui, N., Bernoñaite, G., Disney, M., Duncanson, L., Elmendorp, E., Espejo, A., Furey, A., Hancock, S., Hansen, J., Horsley, H., Langa, S., Liang, M., Locke, A., Manjate, V., Mapanga, F., Omidvar, H., Parsons, A., Peneva-Reed, E., Perry, T., Puma Vilca, B.L., Rodríguez-Veiga, P., Sutcliffe, C., Upham, R., de Walque, B., Burt, A., 2024. Multi-scale lidar measurements suggest miombo woodlands contain substantially more carbon than thought. *Commun. Earth Environ.* 5, 1–11. <https://doi.org/10.1038/s43247-024-01448-x>. <https://www.nature.com/articles/s43247-024-01448-x>. publisher: Nature Publishing Group.
- Demol, M., Calders, K., Verbeeck, H., Gielen, B., 2021. Forest above-ground volume assessments with terrestrial laser scanning: a ground-truth validation experiment in temperate, managed forests. *Ann. Bot.* 128, 805–819. <https://doi.org/10.1093/aob/mab110>. <https://academic.oup.com/aob/article/128/6/805/6362525>.
- Demol, M., Verbeeck, H., Gielen, B., Armstrong, J., Burt, A., Disney, M., Duncanson, L., Hackenberg, J., Kükenbrink, D., Lau, A., Ploton, P., Sewdien, A., Stovall, A., Takoudjou, S.M., Volkova, L., Weston, C., Wortel, V., Calders, K., 2022a. Estimating forest above-ground biomass with terrestrial laser scanning: current status and future directions. *Methods Ecol. Evol.* 13, 1628–1639. <https://doi.org/10.1111/2041-210X.13906>.
- Demol, M., Wilkes, P., Raumonen, P., Krishna Moorthy, S., Calders, K., Gielen, B., Verbeeck, H., 2022b. Volumetric overestimation of small branches in 3d reconstructions of *Fraxinus excelsior*. *Silva Fennica* 56, <https://doi.org/10.14214/sf.10550>. <https://www.silvafennica.fi/article/10550>.
- Dijkstra, E.W., 1959. A note on two problems in connexion with graphs. *Numer. Math.* 1, 269–271. <https://doi.org/10.1007/BF01386390>. <http://link.springer.com/10.1007/BF01386390>.
- Disney, M., Lewis, P., Raumonen, P., 2012. Testing a new vegetation structure retrieval algorithm from terrestrial lidar scanner data using 3d models. In: *Silvilaser 2012*, 12th International Conference on LiDAR Applications for Assessing Forest Ecosystems, 16–19 September 2012, Vancouver, Canada. Silvilaser Vancouver 2012, pp. 1–9. <https://researchport.tuni.fi/en/publications/testing-a-new-vegetation-structure-retrieval-algorithm-from-tree>.
- Disney, M., Boni Vicari, M., Burt, A., Calders, K., Lewis, S.L., Raumonen, P., Wilkes, P., 2018. Weighing trees with lasers: advances, challenges and opportunities. *Interface Focus* 8, 20170048. <https://doi.org/10.1098/rsfs.2017.0048>. <https://royalsocietypublishing.org/doi/full/10.1098/rsfs.2017.0048>.
- Donager, J.J., Sánchez Meador, A.J., Blackburn, R.C., 2021. Adjudicating perspectives on forest structure: how do airborne, terrestrial, and mobile lidar-derived estimates compare? *Remote Sens.* 13, 2297. <https://doi.org/10.3390/rs1312297>. <https://www.mdpi.com/2072-4292/13/12/2297>. number: 12 Publisher: Multidisciplinary Digital Publishing Institute.
- Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armstrong, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P.L., Qi, W., Silva, C., 2020. The global ecosystem dynamics investigation: high-resolution laser ranging of the earth's forests and topography. *Sci. Remote Sens.* 1, 100002. <https://doi.org/10.1016/j.srs.2020.100002>. <http://www.sciencedirect.com/science/article/pii/S2666017220300018>.
- Duncanson, L., Armstrong, J., Disney, M., Avitabile, V., Barbier, N., Calders, K., Carter, S., Chave, J., Herold, M., Crowther, T., Falkowski, M., Kellner, J., Labrière, N., Lucas, R., MacBean, N., McRoberts, R., Meyer, V., Næsset, E., Nickeson, J., Williams, M., 2019. The importance of consistent validation of global forest aboveground biomass products. *Surv. Geophys.* 40, <https://doi.org/10.1007/s10712-019-09538-8>.
- Eloy, C., 2011. Leonardo's rule, self-similarity, and wind-induced stresses in trees. *Phys. Rev. Lett.* 107, 258101. <https://doi.org/10.1103/PhysRevLett.107.258101>. <http://link.aps.org/doi/10.1103/PhysRevLett.107.258101>. publisher: American Physical Society.
- Elseberg, J., Magnenat, S., Siegwart, R., Rueckert, A., 2012. Comparison of Nearest-Neighbor-Search Strategies and Implementations for Efficient Shape Registration. *Università Degli Studi Di Bergamo*, https://doi.org/10.6092/JOSER.2012.03_01_P2. <https://aisberg.unibg.it//handle/10446/86202>.
- Esser, M.H.M., 1946. Tree trunks and branches as optimum mechanical supports of the crown: i. The trunk. *Bull. Math. Biophys.* 8, 65–74. <https://doi.org/10.1007/BF02478463>.
- Evermotion, 2024. Archmodels vol. 238. https://evermotion.org/shop/show_product/archmodels-vol-238/17446. place: Bialystok, Poland.
- Fan, G., Nan, L., Dong, Y., Su, X., Chen, F., 2020. Adqsm: a new method for estimating above-ground biomass from TLS point clouds. *Remote Sens.* 12, 3089. <https://doi.org/10.3390/rs12143089>.

- 10.3390/rs12183089. <https://www.mdpi.com/2072-4292/12/18/3089>. number: 18 Publisher: Multidisciplinary Digital Publishing Institute.
- Ferraro, P., Godin, C., Prusinkiewicz, P., 2005. TOWARD a QUANTIFICATION OF SELF-SIMILARITY IN PLANTS. *Fractals* 13, 91–109. <https://doi.org/10.1142/S0218348X05002805>. <https://www.worldscientific.com/doi/abs/10.1142/S0218348X05002805>.
- Hackenberg, J., Bontemps, J.D., 2023. Improving quantitative structure models with filters based on allometric scaling theory. *Appl. Geom.* <https://doi.org/10.1007/s12518-023-00537-4>. <https://link.springer.com/pdf/10.1007/s12518-023-00537-4>.
- Hackenberg, J., Specker, H., Calders, K., Disney, M., Raunonen, P., 2015. Simpletree—an efficient open source tool to build tree models from TLS clouds. *Forests* 6, 4245–4294. <https://doi.org/10.3390/f6114245>. <https://www.mdpi.com/1999-4907/6/11/4245>. number: 11 Publisher: Multidisciplinary Digital Publishing Institute.
- Harris, N.L., Gibbs, D.A., Baccini, A., Birdsey, R.A., Bruun, S.D., Farina, M., Fatoyinbo, L., Hansen, M.C., Herold, M., Houghton, R.A., Potapov, P.V., Suarez, D.R., Roman-Cuesta, R.M., Saatchi, S.S., Slay, C.M., Turubanova, S.A., Tyukavina, A., 2021. Global maps of twenty-first century forest carbon fluxes. *Nat. Clim. Change* 1–7. <https://doi.org/10.1038/s41558-020-00976-6>. publisher: Nature Publishing Group.
- Herold, M., Carter, S., Avitabile, V., Espejo, A.B., Jonckheere, I., Lucas, R., McRoberts, R.E., Næsset, E., Nightingale, J., Petersen, R., Reiche, J., Romijn, E., Rosenqvist, A., Rozendaal, D.M.A., Seifert, F.M., Sanz, M.J., De Sy, V., 2019. The role and need for space-based forest biomass-related measurements in environmental management and policy. *Surv. Geophys.* 40, 757–778. <https://doi.org/10.1007/s10712-019-09510-6>
- Hu, S., Li, Z., Zhang, Z., He, D., Wimmer, M., 2017. Efficient tree modeling from airborne lidar point clouds. *Comput. Graph.* 67, 1–13. <https://doi.org/10.1016/j.cag.2017.04.004>. <https://www.sciencedirect.com/science/article/pii/S0097849317300420>.
- King, D., Loucks, O.L., 1978. The theory of treebole and branch form. *Radiat. Environ. Biophys.* 15, 141–165. <https://doi.org/10.1007/BF01323263>. <http://link.springer.com/10.1007/BF01323263>.
- Lau, A., Martius, C., Bartholomeus, H., Shenkin, A., Jackson, T., Malhi, Y., Herold, M., Bentley, L.P., 2019. Estimating architecture-based metabolic scaling exponents of tropical trees using terrestrial lidar and 3d modelling. *For. Ecol. Manag.* 439, 132–145. <https://doi.org/10.1016/j.foreco.2019.02.019>. <https://www.sciencedirect.com/science/article/pii/S0378112718318036>.
- Levick, S., Devereux, T., Johnson, S., Woodgate, W., 2023. Tumbarumba wet eucalypt terrestrial lidar, 2022. <https://doi.org/10.25901/sd66-pq95>. <https://portal.tern.org.au/metadata/TERN/9ddb0c50-1310-4288-b56f-a396d4e7d94d>.
- Levick, S., Devereux, T., Johnson, S., Woodgate, W., 2024a. Calperum mallee terrestrial lidar, 2021. <https://doi.org/10.25901/px6c-wj02>. <https://portal.tern.org.au/metadata/TERN/6fd837d7-11c0-4301-bfd-353b8462e019>.
- Levick, S., Devereux, T., Woodgate, W., Johnson, S., 2024b. Litchfield Savanna terrestrial lidar, 2021. <https://doi.org/10.25901/160a-ak75>. <https://portal.tern.org.au/metadata/TERN/41f65b2b-bb74-46ce-83aa-50643afc3246>.
- Lowe, T., Pinskier, J., 2022. Tree reconstruction using topology optimisation. *Remote Sens.* 15, 172. <https://doi.org/10.3390/rs15010172>. <https://www.mdpi.com/2072-4292/15/1/172>.
- Lowe, T.D., Stepanas, K., 2021. Raycloudtools: a concise interface for analysis and manipulation of ray clouds. *IEEE Access* 9, 79712–79724. <https://doi.org/10.1109/ACCESS.2021.3084954>. <https://ieeexplore.ieee.org/document/9444433/>.
- Malenovský, Z., Lamsal, K., Janoutová, R., Devereux, T., Woodgate, W., Hambrecht, L., Cimoli, E., Lucieer, A., Homolová, L., Reagieg, O., Wang, Y., Gastellu-Etchegorry, J., 2023. 3d radiative transfer modelling of forest canopies reconstructed from terrestrial laser scanning: a case of tall Australian eucalypts. In: IGARSS 2023 – 2023 IEEE International Geoscience and Remote Sensing Symposium. IEEE, Pasadena, CA, USA, pp. 2683–2686. <https://doi.org/10.1109/IGARSS52108.2023.10281746>. <https://ieeexplore.ieee.org/document/10281746/>.
- Mandelbrot, B.B., 1983. The Fractal Geometry of Nature. New York, <https://ui.adsabs.harvard.edu/abs/1983whf.book.M>.
- Martin-Ducup, O., Mofack, G., Wang, D., Raunonen, P., Ploton, P., Sonké, B., Barbier, N., Couteron, P., Pélissier, R., 2021. Evaluation of automated pipelines for tree and plot metric estimation from TLS data in tropical forest areas. *Ann. Bot.* 128, 753–766. <https://doi.org/10.1093/aob/mcab051>. <https://academic.oup.com/aob/article/128/6/753/6238622>.
- Momo Takoudjou, S., Ploton, P., Sonké, B., Hackenberg, J., Griffon, S., de Coligny, F., Kamdem, N.G., Libalah, M., Mofack, G.I., Le MoguéDéember, G., Pélissier, R., Barbier, N., 2018. Using terrestrial laser scanning data to estimate large tropical trees biomass and calibrate allometric models: a comparison with traditional destructive approach. *Methods Ecol. Evol.* 9, 905–916. <https://doi.org/10.1111/2041-210X.12933>. <http://doi.wiley.com/10.1111/2041-210X.12933>.
- Morales, A., MacFarlane, D.W., 2024. Reducing tree volume overestimation in quantitative structure models using modeled branch topology and direct twig measurements. *For.: Int. J. For. Res.* cpae046. <https://doi.org/10.1093/forestry/cpae046>. <https://academic.oup.com/forestry/advance-article/doi/10.1093/forestry/cpae046/7801610>.
- Morhart, C., Schindler, Z., Frey, J., Sheppard, J.P., Calders, K., Disney, M., Morsdorf, F., Raunonen, P., Seifert, T., 2024. Limitations of estimating branch volume from terrestrial laser scanning. *Eur. J. For. Res.* <https://doi.org/10.1007/s10342-023-01651-z>
- Niklas, K.J., 1994. Plant Allometry: the Scaling of form and Process. University of Chicago Press, Chicago.
- Osunkoya, O.O., Omar-Ali, K., Amit, N., Dayan, J., Daud, D.S., Sheng, T.K., 2007. Comparative height-crown allometry and mechanical design in 22 tree species of Kuala Belalong rainforest, Brunei, Borneo. *Am. J. Bot.* 94, 1951–1962. <https://doi.org/10.3732/ajb.94.12.1951>. <https://bsapubs.onlinelibrary.wiley.com/doi/10.3732/ajb.94.12.1951>.
- Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L., Canadell, J.G., Ciais, P., Jackson, R.B., Pacala, S.W., McGuire, A.D., Piao, S., Rautiainen, A., Sitch, S., Hayes, D., 2011. A large and persistent carbon sink in the world's forests. *Science* 333, 988–993. <https://doi.org/10.1126/science.1201609>. <https://science.scienmag.org/content/333/6045/988>. publisher: American Association for the Advancement of Science Section Research Article.
- Picard, N., Saint-André, L., Henry, M., 2012. Manual for Building Tree Volume and Biomass Allometric Equations: from Filed Measurement to Prediction. Food and Agriculture Organization of the United Nations (FAO) [u.a.], Rome.
- Prusinkiewicz, P., Lindenmayer, A., 1990. Graphical modeling using l-systems. In: Prusinkiewicz, P. (Ed.), The Algorithmic Beauty of Plants. Springer New York, New York, NY, pp. 1–50. https://doi.org/10.1007/978-1-4613-8476-2_1. series Title: The Virtual Laboratory.
- Quegan, S., Le Toan, T., Chave, J., Dall, J., Extrayat, J.F., Minh, D.H.T., Lomas, M., D'Alessandro, M.M., Paillou, P., Papathanassiou, K., Rocca, F., Saatchi, S., Scipal, K., Shugart, H., Smallman, T.L., Soja, M.J., Tebaldini, S., Ulander, L., Villard, L., Williams, M., 2019. The European Space Agency BIOMASS mission: measuring forest above-ground biomass from space. *Remote Sens. Environ.* 227, 44–60. <https://doi.org/10.1016/j.rse.2019.03.032>. <https://linkinghub.elsevier.com/retrieve/pii/S0034425719301233>.
- Raunonen, P., 2022. TreeQSM user manual. https://github.com/InverseTampere/TreeQSM/blob/master/Manual/TreeQSM_documentation.pdf.
- Raunonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., Holopainen, M., Disney, M., Lewis, P., 2013. Fast automatic precision tree models from terrestrial laser scanner data. *Remote Sens.* 5, 491–520. <https://doi.org/10.3390/rs5020491>. <https://www.mdpi.com/2072-4292/5/2/491>.
- Tagliasacchi, A., Delame, T., Spagnuolo, M., Amenta, N., Telea, A., 2016. 3D Skeletons: a state-of-the-art report. *Comput. Graph. Forum* 35, 573–597. <https://doi.org/10.1111/cgf.12865>. <https://onlinelibrary.wiley.com/doi/10.1111/cgf.12865>.
- Gonzalez de Tanago, J., Lau, A., Bartholomeus, H., Herold, M., Avitabile, V., Raunonen, P., Martius, C., Goodman, R.C., Disney, M., Manuri, S., Burt, A., Calders, K., 2018. Estimation of above-ground biomass of large tropical trees with terrestrial lidar. *Methods Ecol. Evol.* 9, 223–234. <https://doi.org/10.1111/2041-210X.12904>. <http://doi.wiley.com/10.1111/2041-210X.12904>.
- Vicari, M.B., Disney, M., Wilkes, P., Burt, A., Calders, K., Woodgate, W., 2019. Leaf and wood classification framework for terrestrial lidar point clouds. *Methods Ecol. Evol.* 10, 680–694. <https://doi.org/10.1111/2041-210X.13144>. <https://onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.13144>.
- Wang, D., Takoudjou, S.M., Casella, E., 2020. Lewos: a universal leaf-wood classification method to facilitate the 3d modelling of large tropical trees using terrestrial lidar. *Methods Ecol. Evol.* 11, 376–389. <https://doi.org/10.1111/2041-210X.13342>. <https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.13342>.
- West, G.B., Brown, J.H., Enquist, B.J., 1997. A general model for the origin of allometric scaling laws in Biology. *Science* 276, 122–126. <https://doi.org/10.1126/science.276.5309.122>.
- West, G.B., Brown, J.H., Enquist, B.J., 1999. A general model for the structure and allometry of plant vascular systems. *Nature* 400, 664–667. <https://doi.org/10.1038/23251>.
- Wilkes, P., Lau, A., Disney, M., Calders, K., Burt, A., Gonzalez de Tanago, J., Bartholomeus, H., Brede, B., Herold, M., 2017. Data acquisition considerations for terrestrial laser scanning of forest plots. *Remote Sens. Environ.* 196, 140–153. <https://doi.org/10.1016/j.rse.2017.04.030>. <https://www.sciencedirect.com/science/article/pii/S003442571730189X>.
- Wilkes, P., Shenkin, A., Disney, M., Malhi, Y., Bentley, L.P., Vicari, M.B., 2021. Terrestrial laser scanning to reconstruct branch architecture from harvested branches. *Methods Ecol. Evol.* 12, 2487–2500. <https://doi.org/10.1111/2041-210X.13709>. <https://onlinelibrary.wiley.com/doi/10.1111/2041-210X.13709>.
- Winiwarter, L., Esmoris Pena, A.M., Weiser, H., Anders, K., Martínez Sánchez, J., Searle, M., Höfle, B., 2022. Virtual laser scanning with HELIOS + : a novel take on ray tracing-based simulation of topographic full-waveform 3d laser scanning. *Remote Sens. Environ.* 269, 112772. <https://doi.org/10.1016/j.rse.2021.112772>. <https://linkinghub.elsevier.com/retrieve/pii/S0034425721004922>.
- Xu, H., Gossett, N., Chen, B., 2007. Knowledge and heuristic-based modeling of laser-scanned trees. *ACM Trans. Graph.* 26, 19. <https://doi.org/10.1145/1289603.1289610>.
- Zhang, D., Král, K., Kruček, M., Cushman, K., Kellner, J.R., 2024. Near-complete sampling of forest structure from high-density drone lidar demonstrated by ray tracing. *Remote Sens.* 16, <https://doi.org/10.3390/rs16152774>.
- Akerblom, M., Kaitaniemi, P., 2021. Terrestrial laser scanning: a new standard of forest measuring and modelling? *Ann. Bot.* 128, 653–662. <https://doi.org/10.1093/aob/mcab111>. <https://academic.oup.com/aob/article/128/6/653/6364917>.