Supplementary Materials

Identifying Old-Growth Forests in Complex Landscapes: A New LiDAR-Based Estimation Framework and Conservation Implications

Raphaël Trouvé *, Ruizhu Jiang †, Patrick J. Baker, Sabine Kasel and Craig R. Nitschke

School of Agriculture, Food and Ecosystem Sciences, The University of Melbourne, Melbourne, VIC 3010, Australia
* Correspondence: raphael.trouve@unimelb.edu.au
† Current address: Centre for Crop Science, The University of Queensland, Brisbane, QLD 4072, Australia.

S.1 Individual tree delineation algorithm

In this section, we present and validate the individual tree detection (ITD) algorithm that was developed in Jiang (2019) and used in our manuscript. We also compare Jiang (2019) ITD results to alternative ITD algorithms that are commonly used in the field. The code for our ITD algorithm is available on GitHub. The Github repository includes an 'examples' folder that contains an R tutorial to run the ITD code in Google Colab (Jiang 2023):

https://github.com/ruizhuj/Individual_Tree_Delineation_in_Broadleaf_forest

S.1.1 Refining the watershed-ITD algorithm for broadleaved species

Individual tree delineation using Light Detection and Ranging (LiDAR) data represents a significant advancement in forest inventory and management (Popescu, Wynne et al. 2002; Wang, Gong et al. 2004; Dalponte, Orka et al. 2014; Ma, Chen et al. 2022). The process uses laser scanning data to identify and delineate individual trees within a forest stand over large scales.

One of the most common ITD algorithm is marker-controlled watershed delineation based on a Canopy Height Model (CHM) raster layer. This method works well when the tree crowns have a regular shape (Kwak, Lee et al. 2007; Madhuri Kalapala 2012; Soille 2013; Amiri 2014; Ayrey, Fraver et al. 2017). In the marker-controlled watershed approach, the CHM is usually created by taking the maximum height of height-normalized LiDAR points with cells of 1 meter or smaller resolution. This method is highly accurate in coniferous forests, where tree crowns have a conical shape and treetop is located at the center of the crown (Ke and Quackenbush 2011). However, in broadleaf-dominated forests, such as the *Eucalyptus* forests in our case study, the irregularity of tree crowns makes identifying and detecting treetops more challenging.

In this study, we develop a more robust watershed ITD method for overstorey broadleaf species by refining three key aspect of the algorithm: i) choosing a more suitable target raster, ii) improving treetop detection, and iii) identifying crown edges and forest gaps.

i) Choosing a more suitable target raster

Changing the target raster layer from a CHM to the high point densities model (DHP) has been shown to improve the accuracy of ITD algorithms for broadleaved species (Rahman, Gorte et al. 2009). The DHP method relies on the assumption that the LiDAR point density of an overstorey tree crown is greatest at the center the crown and diminishes towards its edges (Rahman and Gorte 2008; Rahman, Gorte et al. 2009). The method used by Rahman, Gorte et al. (2009) was simply to select points above a fixed reference height to distinguish the overstorey tree crowns from the understorey vegetation and ground surface. However, filtering based on a fixed reference height can be problematic if the forest structure vary in space (Rahman and Gorte 2009). To enhance the method's accuracy, a potential strategy would thus be to allow the reference height to vary with overstorey height and forest type. One promising approach is to base the reference height on the height-to-live-crown base (HCB), defined as height below which trees have no crown and above which they have a crown. We would classify LiDAR points below the HCB as belonging to the understorey and points above it as belonging to the overstorey. Height-to-live-crown

base can be predicted from total tree height (HT) (Ritchie and Hann 1987; Nutto, Spathelf et al. 2006; Rijal, Weiskittel et al. 2012; Sharma, Vacek et al. 2017) because of the strong relationship between these features (e.g., R^2 of 0.79 in conifers and 0.80 in broadleaf trees in (Rijal, Weiskittel et al. 2012)). In practice, we would use field data to calibrate a HCB = f(HT) model. As CHM is available for the entire landscape it is logical to substitute HT for CHM for the predictions of canopy surface of HCB.

ii) Improving treetop detection

The success of watershed delineation of overstorey trees relies on accurate tree-top detection (Wang, Gong et al. 2004). When tree-top detection is poor, over-segmentation can occur (i.e., an individual tree will be mistakenly divided into multiple trees), an issue that can be particularly acute for trees with large crowns. Accurate tree-top marker control prevents severe over-segmentation problems in the traditional watershed algorithm (Wang, Gong et al. 2004). A common technique for identifying treetops uses a local maximum filter to identify the peaks of tree crowns with a moving window size (commonly 3×3, 5×5, or 7×7 pixels) that can also depend on crown size (Wulder, Niemann et al. 2000; Persson, Holmgren et al. 2002; Popescu, Wynne et al. 2002; Popescu and Wynne 2004). The local maximum method makes the assumption that the peak of the tree-crown is located at or very close to the treetop (Brandtberg and Walter 1998). An image-smoothing targeted raster layer can help to reduce the noise effect (Dralle and Rudemo 1996).

iii) Identifying crown edges and forest gaps

Accurate crown edge detection reduces the identification of pseudo tree-tops (i.e., situations where we have a local height maximum on the crown edge or close to a forest gap) (Wang, Gong et al. 2004) and prevents multiple trees being merged together incorrectly (Brandtberg and Walter 1998). Canopy height models usually exhibit spatial variation in heights (Xu, Iuricich et al. 2023). Topographic analysis tools, such as slope, aspect and slope of aspect (SOA) can be used to analyze CHM (Wu, Yu et al. 2016). Slope of aspect of CHM represents the rate of change of aspect of a crown; therefore, the maximum value of SOA of CHM is located along the crown edge (“gully”) or treetop (“ridge”). Therefore, high values of SOA of CHM combined with the negative terrain surface of CHM indicate crown edges. Forest gaps can also be detected from the negative terrain surface using neighborhood analysis tools (such as the focal statistic tool in ArcGIS) to identify surfaces that are relatively lower within a larger identified neighborhood.

Incorporating all the above steps, we developed a ITD algorithm using a watershed segmentation method targeting a raster layer that combines the canopy height model (CHM) and the densities of high points model (DHP) at a fine spatial resolution of 0.5 meters (CHM×DHP). We first used the HCB = f(HT) relationship to classify LiDAR points into dominant overstorey trees and understorey trees. We then removed the LiDAR points belonging to the understorey. Canopy gaps were detected by using a focal statistic tool to identify CHM surfaces that are relatively lower within a larger identified neighborhood. The LiDAR points of vegetation located in canopy gaps were also removed. We then created DHP only based on LiDAR points of dominant overstorey canopies. Treetop detection with a local maximum filter was based on the developed raster layer CHM×DHP. Crown edge detection used the successive application of terrain surface analysis tools and neighborhood tools, including slope of CHM, aspect of CHM, SOA of CHM, and focal statistic tool. High values of SOA of CHM combined with the relatively lower

canopy height surface within a specified neighborhood were used to identify crown edges. We used a window size of 5 meters for the focal statistic tool to identify crown edges and a window size of 20 meters to identify canopy gaps. This produced a top-edge-enhanced raster layer of CHM×DHP for use in watershed segmentation. The ITD workflow is summarized below:

1. Create a 50 cm resolution canopy height model (CHM).
2. Create a height to crown base model (HCBM) based on a crown based height – top height relationship calibrated from field data (Figure S1).
3. Identify canopy gaps.
4. Filter out the understorey and midstorey points (i.e., points that are below the HCBM), retaining only overstorey LiDAR points.
5. Filter points located in canopy gaps.
6. Create a density of high points model (DHP) using overstorey crowns points.
7. Detect treetops based on a targeted raster layer of CHM×DHP.
8. Perform marker-control watershed delineation on a top-edge-enhanced CHM×DHP layer, calculated as: (CHM×DHP) × (1.2 × treetops) + (CHM×DHP) × (1 - gaps) × (1 - edges) × (1 - treetops). See Figure S2

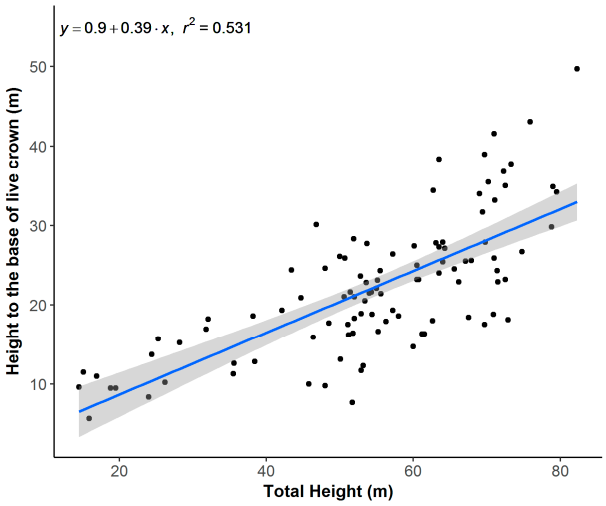

Figure S1: The linear model and R^2 of HCB-HT relationship calibrated from field data.

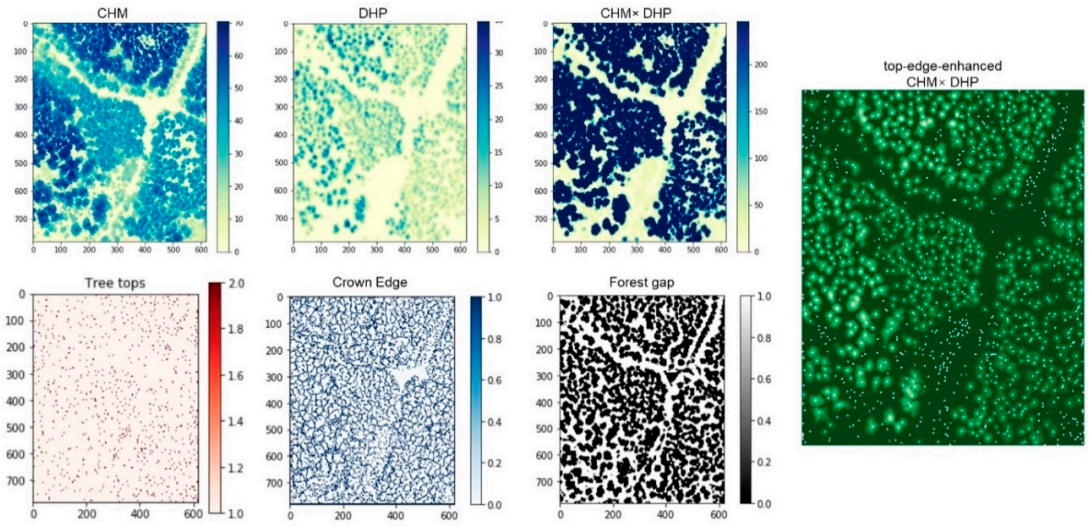

Figure S2: workflow for generating the top-edge-enhanced raster layer of CHM×DHP.

Table S1: The confusion matrix of ITD outputs and equations for each validation metrics. True condition is based on directly observed data from stem-mapped plots and predicted condition is from LiDAR-derived ITD analyses.

		True condition		Equations for individual metrics	
		Condition positive (CP)	Condition negative (CN)	ACC = (TP + TN)/(TP+FP+TN+FN)	
Predicted condition	Predicted condition positive (PP)	True positive (TP) = ITD correctly detected	False positive (FP) = ITD merged crown + ITD over-segmented	True positive rate (TPR) or Sensitivity = TP/CP = TP/ (TP + FP)	Commission rate (CR) or false positive rate (FPR) = FP/PP = FP/(TP+FP)
	Predicted condition negative	False negative (FN) = missed treetops + merged treetops	True negative (TN)=0	Omission rate (OR) = FN/CP = FN/ (TP + FN)	True negative rate (TPR) or Specificity =TN/CN=TN/(TN+FP)

S1.2 Individual tree detection validation

We validated our method by comparing our ITD predictions with field measurement data. We first established three 1-ha plots in mountain ash forest to validate our ITD algorithm. In each plot, all dominant overstorey trees were mapped, identified to species, and measured for DBH, total height (HT), height to the base of the live crown (HCB), and mean crown width (CW). We estimated mean crown width as the mean distance from the tree bole to the northern, southern, western, and eastern edge of the tree crown. Within each plot in an area of approximately 2000 m^2 (20 × 100m), mid- and understorey trees with DBH>10cm were mapped and we measured their stem and crown attributes. These three plots were located in forest with distinct growth stages (1939 regrowth, 1977 regrowth, and multi-cohort stand). In total, 873 trees were measured, including 309 dominant overstorey trees, 422 mid- and understorey trees,

and 142 suppressed canopy or dead trees. The dominant overstorey species was *Eucalyptus regnans* and the most common mid- and understorey species were *Acacia dealbata* and *A. frigescens,* respectively.

The ground truth validation was carried out using the coordinates of live overstory trees within three 1-ha stem mapped plots. Four validation indicators were calculated to evaluate the ITD method compared to the ground truthing: overall accuracy (ACC), sensitivity (true positive rate, TPR), commission rate (CR) and omission rate (OR) (Sačkov, Hlásny et al. 2017). True Positives (TP) represents the proportion of crowns accurately identified by the ITD method; False Negatives (FN) is the proportion of treetops missed by the ITD method, or treetops of crowns that were erroneously merged with adjacent crowns; False Positives (FP) is the proportion of crowns that were over-segmented by the ITD method or crowns containing two or more true crowns. The equations for these validation metrics are shown in Table S1. For crowns that were correctly segmented (True positive), we also estimated the R^2 between LiDAR-derived crown width and the crown widths measured in the field to evaluate the ability of our ITD approach to accurately delineate individual crowns and their sizes. The LiDAR-derived crown width was estimated based on the projection area of each delineated polygon for each overstorey tree crown by ITD.

The ITD approach that we developed for this study was able to correctly detect individual trees in different growth stages of the mountain ash forest (Table S3, Figure S3), although the accuracy varied among growth stages. Our algorithm was most accurate in the 1939 regrowth and multi-cohort stands that were dominated by 1939 regrowth mountain ash trees (ACC = 0.8 in both plots). The dominant overstorey crowns identified by ITD in these plots had an 86% chance of being the true crowns and an 8% chance to being over-segmented (i.e., a single tree crown being classified as two or more individual crowns by ITD). The ITD algorithm had a 13% chance of missing a true crown. In 1977 regrowth forest (i.e., 38-years old), the ITD algorithm was much less effective (ACC = 45%). Only 53% crowns were correctly detected and 47% of crowns were merged with their neighbors or being missed. Figure S3 shows the outcomes of our ITD method for stem mapping, indicating whether each detection was correct or incorrect. Missed field crowns in the 1939 regrowth and multi-cohort stands were most likely suppressed overstorey trees, while missed field crowns in the 1977 regrowth forest were located in “canopy gaps” identified by our ITD method. For detected trees, we found a linear relationship between the LiDAR-derived crown width and the crown widths measured in the field (R^2 = 0.64; relative error of 18%; see Figure S4). The relationship followed the 1:1 line (intercept not significantly different from zero, slope not significantly different from one), indicating no detected bias in LiDAR reconstructed crown widths.

Table S2: Accuracy of individual tree delineation results for each plot.

		Plot1	Plot2	Plot3
		1939 regrowth	**1977 regrowth**	**multi-cohort**
Number of trees				
	Field measurement	65	207	37
TRUE Positive (TP)	ITD correctly detect	56	109	32
FALSE Negative (FN)	Missed field treetops	3	7	5
	Merged field treetops	6	91	0
FALSE Positive (FP)	ITD >2 crowns	3	33	0
	ITD over-segmented	2	0	3
Model Validation				
	Accuracy (ACC)	0.800	0.454	0.800
	Sensitivity (TPR)	0.862	0.527	0.865
	Omission Rate (OR)	0.138	0.473	0.135
	Commission Rate (CR)	0.082	0.232	0.086

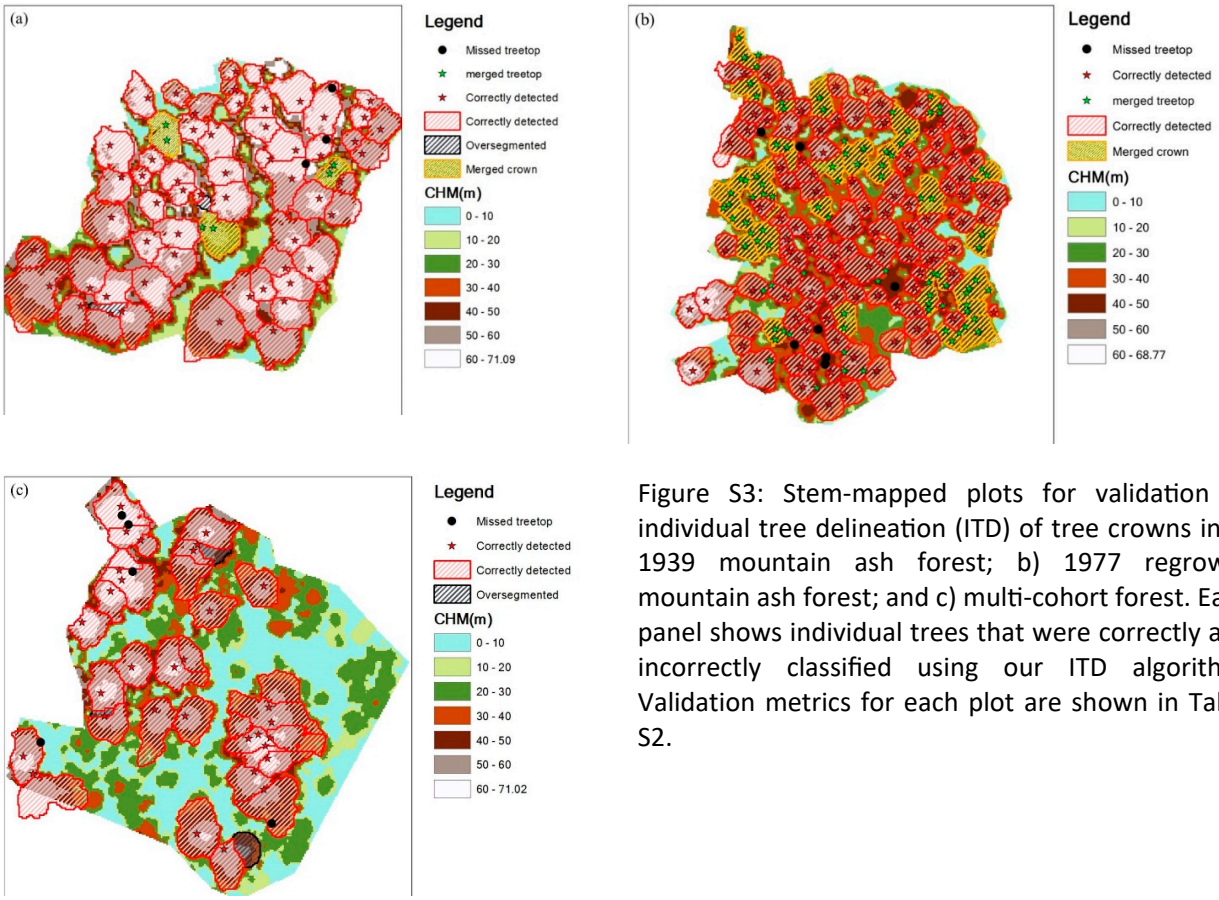

Figure S3: Stem-mapped plots for validation of individual tree delineation (ITD) of tree crowns in a) 1939 mountain ash forest; b) 1977 regrowth mountain ash forest; and c) multi-cohort forest. Each panel shows individual trees that were correctly and incorrectly classified using our ITD algorithm. Validation metrics for each plot are shown in Table S2.

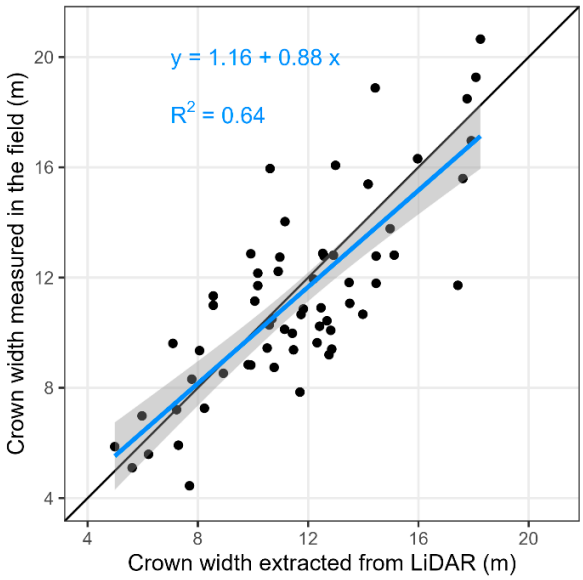

Figure S4: The relationship between field-measured crown width and crown width extracted from LiDAR. The relative error of the relationship (coefficient of variation) was around 18%.

S1.3 Comparison with existing ITD methods

In addition to the watershed delineation method, a few studies using raster-based ITD methods, such as layer stacking and localized contour, have shown a high level of accuracy for individual tree delineation. Layer stacking is a method which slices the entire forest point cloud at 1-m height intervals and isolates trees in each layer. Merging the results from all layers produces representative tree profiles (Ayrey, Fraver et al. 2017; Ma, Chen et al. 2022). Layer stacking has high accuracy for detecting the location of individual trees. Ayrey et al. (2017) was able to identify 75% of trees correctly detected in pure deciduous stands; however, layer stacking tends to underestimate crown size which could be problematic in our study. Localized contour seeks to capture the topological structure of the canopy and quantifies the topological relationships of tree crowns using a graph theory-based localized contour tree method which can be used to segment individual tree crowns in a manner analogous to detecting hills on a topographic map (Wu, Yu et al. 2016). This approach essentially creates crown contours based on the CHM. In coniferous forests, the overall accuracy of the localized contour ITD method reached 75 to 94% (Wu, Yu et al. 2016). The lidR package in R (https://r-lidar.github.io/lidRbook/itd-its.html#itd) also incorporates innovative methods such as Local Maximum Filter (LMF) and Dalponte and Coomes' region-growing algorithm (Dalponte and Coomes 2016; Roussel, Auty et al. 2020). The LMF recognizes local maxima in CHM as treetops, providing a basis for initial tree detection. The LMF supports the use of both fixed and variable window sizes, allowing for flexible and adaptive tree detection. Subsequently, the Dalponte et al. (2016) algorithm utilizes a region-growing strategy that progressively delineates individual tree crowns from a given 'seed' point, which is typically the local maximum determined by the LMF.

To test the performance of our approach, we compared it to traditional watershed-ITD (Madhuri Kalapala 2012; Amiri 2014), layer stacking (Ayrey et al., 2017), localized contour (Wu, Yu et al. 2016) and LidR

methods. We used these approaches to delineate the crowns in one-hectare stem-mapped plot located in 1939 regrowth mountain ash stand. Four validation indicators (overall accuracy (ACC), sensitivity (true positive rate, TPR), commission rate (CR) and omission rate (OR)) were calculated and used to compare the model performance with our refined watershed-ITD approach. We also compared the crown shape and size delineated from these four methods.

The comparison of the developed algorithms with layer-stacking, localized contour, and marker-control watershed segmentation are summarized in Table S3 and Figure S5. The highest accuracy of treetop detection was layer-stacking method (ACC=90%). Trees detected by layer-stacking had a 92% chance of being the true crowns. No crown was over-segmented by layer-stacking. When compared to our ITD algorithm (top-edge enhanced watershed delineation), layer stacking correctly identifed 9.6% more trees. Traditional marker-control watershed segmentation, localized contour and LidR method both performed badly in our stem mapping plots. Traditional marker-control watershed segmentation method trends to have highest rate to over-segment trees among four methods. The overall accuracy of localized contour was the lowest among all four methods (ACC=46%). Crowns delineated by lidR were merged highest number of crowns with their neighbors (19/65 = 29.2%). This suggests that a direct application of the region-growing method on the CHM of broadleaf forests may not yield optimal segmentation outcomes.

In terms of characterizing crown shape, layer stacking tended to underestimate crown size, because it excluded layers that were farthest from the tree center. Crown shapes delineated by localized contour were also smaller than the actual tree crown shapes, due to the fact that only closed contour lines were considered in this method. Traditional marker-control watershed segmentation algorithms often mistakenly merged smaller trees within or around the crowns of primary trees, and so it tended to overestimate crown size. Although the overall accuracy of our ITD algorithm was slightly less than for layer-stacking, it delineated tree crowns precisely along the crown edge, which provided more accurate information of the size off tree crowns.

Table S3: Accuracy of individual tree delineation results for each method.

		Traditional Watershed	Layer stacking	local contour	LidR	ITD in this study
	Number of trees					
TRUE Positive (TP)	Field measurement	65	65	65	65	65
	ITD correctly detect	52	60	38	46	56
FALSE Negative (FN)	Missed field treetops	0	1	9	0	3
	Merged field treetops	13	4	18	19	6
FALSE Positive (FP)	ITD >2 crowns	6	2	9	12	3
	ITD over-segmented	25	0	9	6	2
	Model Validation					
	Accuracy (ACC)	0.54	0.90	0.46	0.55	0.80
	Sensitivity (TPR)	0.80	0.93	0.59	0.71	0.86
	Omission Rate (OR)	0.20	0.01	0.42	0.29	0.14
	Commission Rate (CR)	0.37	0.03	0.32	0.28	0.08

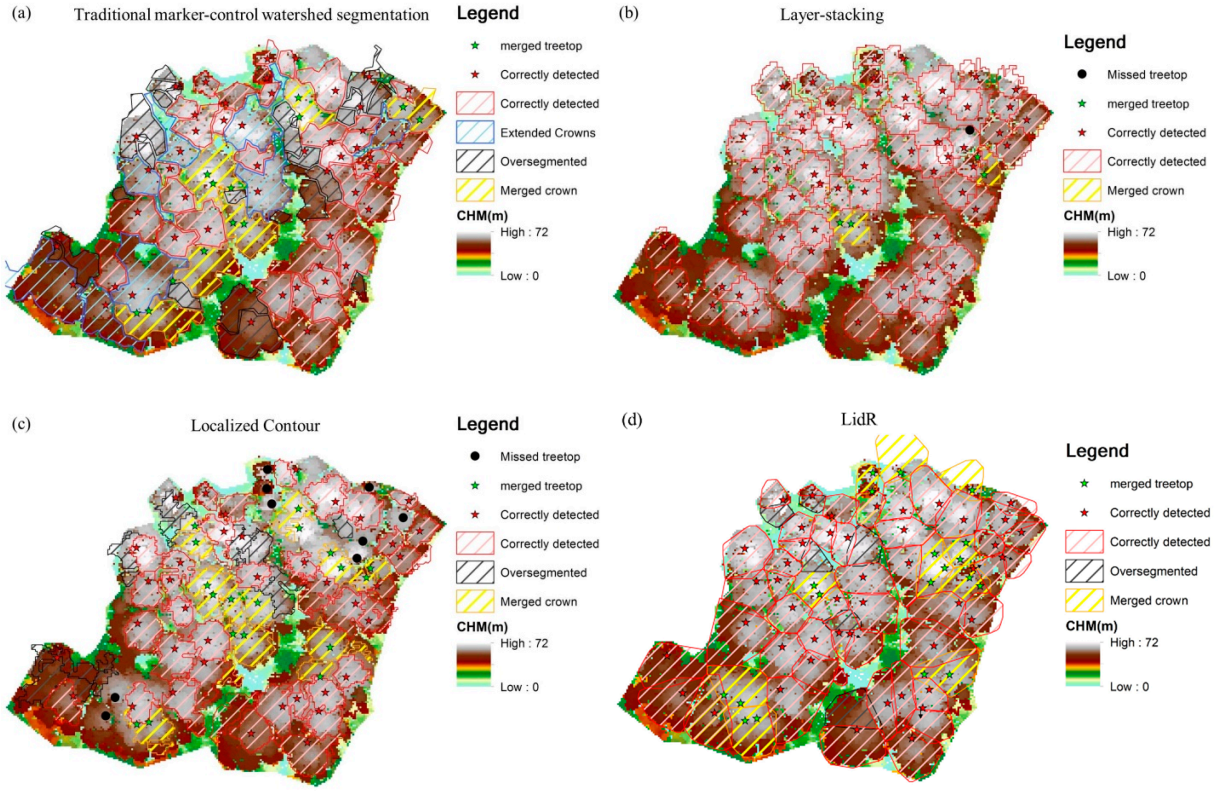

Figure S5: Stem-mapped plots for plot1-1939 regrowth forest of delineated tree crowns using: a) traditional marker-control watershed segmentation method; b) layer-stacking method; c) localized contour method; and d) lidR method. Each panel shows individual trees that were correctly and incorrectly classified using our ITD algorithm.

S.2 Workflow on binary classification of old-growth forests

In this section, we describe the workflow used to reconstruct the presence of old-growth forests from aerial LiDAR. We also discuss the uncertainties associated with each step of our workflow. We recognize that the uncertainties associated with each step can interact in complex ways, which is why we also validated the modelled old growth map using an independent dataset, as explained in the manuscript. The classification of old-growth forests from aerial LiDAR data followed six steps:

Step 1) Light Detection And Ranging (LiDAR) acquisition. The aerial LiDAR data was collected in 2015-2016 by RPS Group and provided to the Victorian Government Department of Environment, Energy, and Climate Action, (DEECA). The average LiDAR point density within the 337,548 ha surveyed was around 28 points / m^2 (Trouvé, Jiang et al. 2023). Figure S6 shows a map of the LiDAR surveyed area in the Central Highlands of Victoria. The root mean squared error in raw LiDAR point height was around 10 cm. The digital elevation and canopy height model production method can also result in uncertainty in the predicted canopy height and shape reliability (degree of fidelity in the spatial pattern uncertainty) that can affect later steps in the workflow. This step has a small impact on the total uncertainty of the workflow.

Step 2) We delineated individual tree crown by using an individual tree detection (ITD) algorithm on the LiDAR data (Figure S7) (Jiang 2019). Details for our ITD algorithm are available in section S1 of this Supplementary Materials. Following this step, we have a list with the estimated location, tree height, and crown width of the 22 million trees detected in the LiDAR footprint (Trouvé, Jiang et al. 2023). In the Central Highlands region, we can summarize the uncertainties associated with the ITD algorithm as follows:

- In 1939 regrowth and multi-cohort stands dominated by 1939 regrowth trees (76 years old trees), the overall accuracy of the algorithm was 80% (Jiang 2019). The dominant overstorey crowns identified by ITD in these plots had an 86% chance of being the true crowns and an 8% chance to being over-segmented (i.e., a single tree crown being classified as two or more individual crowns by the ITD). The ITD algorithm had a 13% chance of missing a true crown. Missed crowns in the 1939 regrowth and multi-cohort stands were most often suppressed overstorey trees. Missing suppressed trees is unlikely to impact the classification of old-growth forests as these forests are primarily classified based on the presence of large trees; however, it could potentially affect other aspects of the data (e.g., estimating basal area or biomass per hectare based on the identified trees, but these objectives are out of scope for this study).
- The ITD algorithm was much less effective in younger stands due to overlapping crowns. For example, in 1977 regrowth forest (i.e., 38 years old trees), the overall accuracy was only 45% (Jiang 2019). Only 53% crowns were correctly detected and 47% of crowns were merged with their neighbors or missed. Low accuracy in identifying individual trees in young regrowth is also unlikely to impact old-growth forests classification as old-growth forests are primarily classified based on the presence of large trees.
- For detected trees, there was a good linear relationship between the LiDAR-derived crown width and the crown widths measured in the field (R^2 = 0.64; with no detected bias; see Figure S4). The

highlights the relative accuracy of the crown dimensions generated by the ITD algorithm, which is important for developing estimates of the DBH associated with each individual tree.

Step 3) We used a diameter at breast height (DBH, in cm) vs. crown width (in m) allometry to reconstruct the DBH of each detected tree. The allometry was calibrated from LiDAR detected crown width and DBH field measurements. Our dataset included DBH field data from 158 individual trees (89 from stem mapping plots and 69 from large old tree plots) (Jiang 2019). We matched the DBH of each tree with its LiDAR crown width as reconstructed by our ITD algorithm. The allometric relationship was calibrated using a generalized linear model with a log-link and Gamma distribution to account for the heteroscedasticity in the data. The DBH-crown allometry follows:

$$\text{DBH} = e^{1.899}\,\text{crown-width}^{1.188}$$

The equation is illustrated in Figure S8. As can be seen in Figure S8, there is some residual uncertainty in the individual DBH predicted by the DBH-crown allometry. The relative error (i.e., coefficient of variation) for the residuals of the DBH-crown width model was 30%.

Step 4) We grouped detected trees by cohorts of similar DBH using finite mixture models (Fedrigo, Stewart et al. 2018). For each hectare, we fitted four models containing a mixture of one to four normal distributions. The best model, in terms of leave-one-out cross-validation, determines the number of cohorts present in each hectare. At the end of this step, we had an estimate of the number of LiDAR detected trees per ha, the number of cohorts per ha, the mean tree size and standard deviation of each cohort as well as the proportion of trees and crown cover in each cohort.

There are three potential sources of uncertainty in the finite mixture modelling step: 1) the spatial aggregation on a per ha basis; 2) model selection (i.e., how many cohorts are present in the pixel); and 3) parameter uncertainty in the selected model. There is a bias-variance trade-off concerning the spatial aggregation: larger pixels contain more delineated trees than smaller pixels which facilitate the finite mixture modelling (distributions); however, larger pixels also make it more likely to meet heterogeneous forest types and confound the spatial and vertical components of forest structure diversity. For example, in our workflow, boundaries of previous logging coupes are often classified as multi-cohorts (which should indicate vertical structural diversity) due to the presence of both large and small trees in the same pixel. In reality, boundaries of logging coupes represent the juxtaposition of two separate mono-cohort stands (spatial structural diversity). Concerning model selection, while the multi-cohort nature of some stands is evident and can be easily picked up at this step (e.g., in Figure S9), if cohorts overlap in term of tree size it can be difficult to determine the number of cohorts present in the stand. Following a principle of parsimony, when two models have similar goodness-of-fits, we selected the simplest model (e.g., if the models with one and two cohorts have similar fits, we select the model with one cohort). The selected model also reports uncertainties in terms of model parameters: mean DBH of each cohort (represented by the colored vertical bands in Figure S9) and proportion of trees contained in each cohort.

Additionally, since the finite mixture modelling aggregates the tree data by cohort, it can potentially smooth out some of the errors that arose in steps 1 to 3.

Step 5) We estimated the age of each cohort based on cohort size and cohort growth rate calibrated for the area. We used the HWPLOT database (Trouvé, Nitschke et al. 2017), which is a network of silvicultural experiments spread through Victoria, to estimate DBH growth from 213 control (i.e., unthinned) plots. The growth of trees was well correlated with the annual heat moisture index (AHMI = (T + 10) / (P / 1000), where T is mean annual temperature in °C and P is annual precipitation in mm / year), which is a measure of aridity (Nitschke, Nichols et al. 2017). We used a double logistic regression to represent the asymmetric niche response (Huisman, Olff et al. 1993) of cohort growth rate (ΔDBH, in cm / year) to AHMI. The cohort growth equation, which is illustrated in the top row of Figure S10, follows:

ΔDBH = 0.697 + 1.138 + (0.853 - 1.138) / (1 + (AHMI / 11.954)^30) - (1.138 + (0.697 - 1.138) / (1 + (AHMI / 20)^10))

The model had a relative error of 15%. This growth rate error can propagate to the estimate of stand age of the pixel. While age estimates based on the size of an individual tree are notoriously inaccurate, the cohort modelling approach done in step 4 mitigates part of the issue because the mean growth of a cohort is much less variable than the growth of individual trees within the same cohort and increases the signal-to-noise ratio of the data.

Step 6) We used a rule-based classification key to estimate the presence of old-growth forests in each pixel. The rule-based classification key follows the current definition of old-growth forest in Victoria which is based on the Woodgate definition (Woodgate 1994; Woodgate, Peel et al. 1996). Under the current regulation, a stand is classified as old growth if it established prior to 1900, has less than 10% crown cover of regrowth forest (i.e., trees that established after 1900), and the influence of past disturbances (fire and logging) is no longer discernible. We estimated regrowth cover based on the mixture modelling and age reconstruction outputs. Wildfire and logging disturbances were identified from the FIRE_HISTORY (medium to high severity fires), FIRE_SEV09_POLY (fire severity classes 1 and 2 for crown burn and crown scorch), and LASTLOG25 GIS layers downloaded from https://data.vic.gov.au. We then used the combined spatial layers of the cohort mapping, regrowth cover, and disturbance to classify the entire landscape into one of four simplified growth stage classes (see Table 1 in the manuscript and Figure S11).

The rule-based classification can potentially amplify some of the errors coming from the previous steps. For example, in single cohort stands, stands estimated to have recruited in 1899 are considered old growth, while stands estimated to have recruited in 1901 are not. For multicohort stands that have at least one pre-1900 cohort, having a regrowth crown cover of 11% (rather than say 9%) disqualifies the stand from being considered old growth. When stands are close to the recruitment date and regrowth percentage thresholds, small errors can make the difference between being classified as old growth or not. Filtering for disturbances can also add some noise to the definition of old-growth forests as the mapping of fire footprint and fire severity is not always accurate.

In the manuscript, we took three additional steps to address the uncertainties that can accumulate in our workflow: 1) we validated the binary old growth map using 49 independent field plots; 2) We used Monte Carlo simulations to propagate uncertainties in steps 2 to 5 (ITD, crown-DBH allometry, finite mixture modelling, and growth rate equation for age reconstruction) and quantify the impact of the technical aspect of the work on the extent of old-growth forests the landscape; 3) We ran a sensitivity analysis to determine the impact of each rule-based classification filter on the extent of old growth in the region.

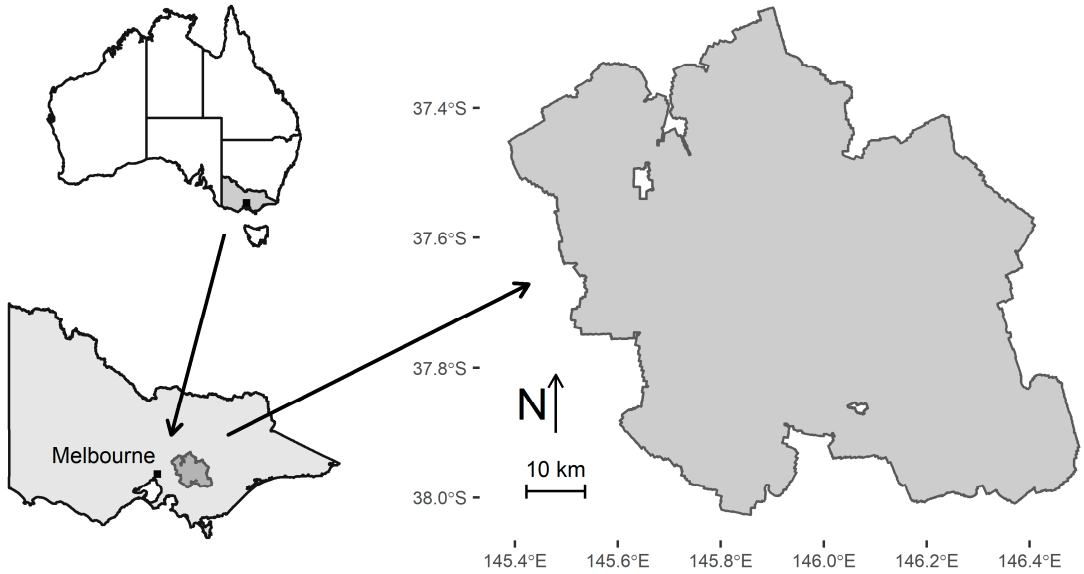

Figure S6: Location of the Central Highlands of Victoria in Australia (left panels) with a close up of the surveyed LiDAR area (right panel).

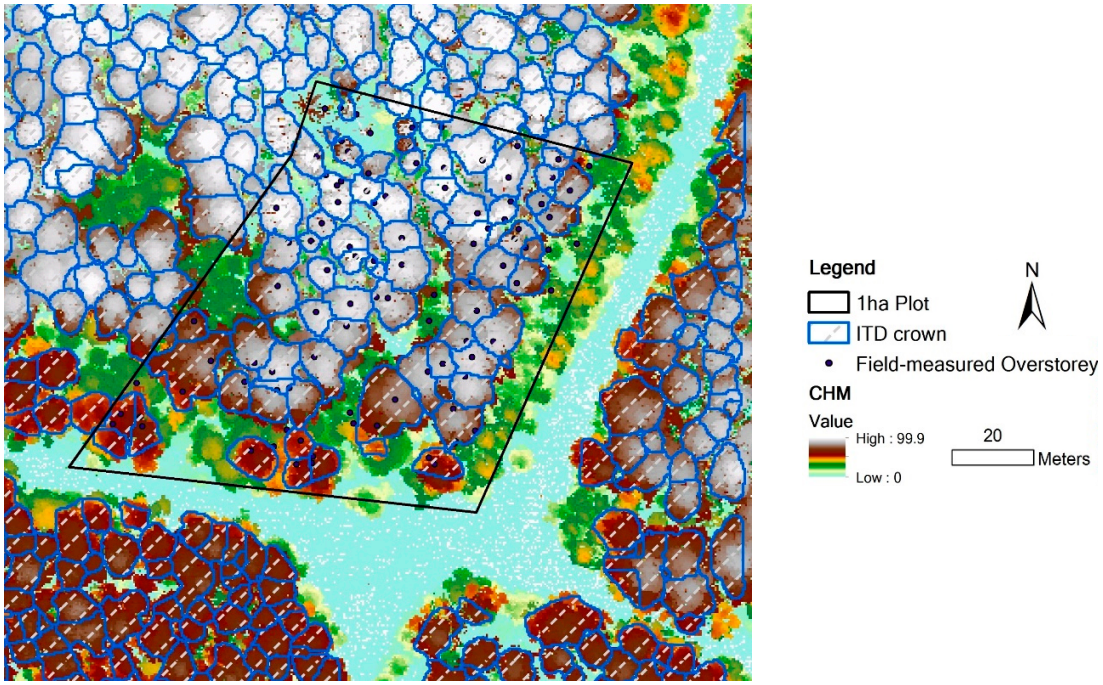

Figure S7: Example of individual tree delineation in a 1939 regrowth plots Central Highlands of Victoria overlayed on the canopy height model (CHM, in meters). See Table S2 and Figure S3 for a more comprehensive analysis of the algorithm's accuracy.

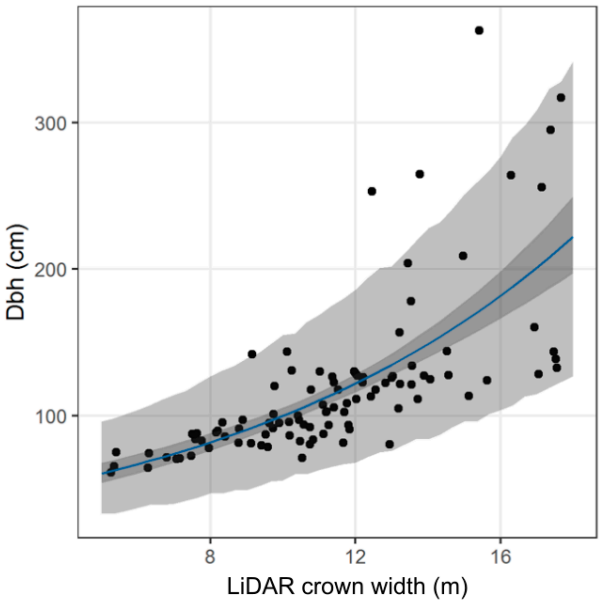

Figure S8: Relationship between DBH and LiDAR crown width in the Central Highlands area. Dots are raw data points, the solid blue line represents mean predictions, dark shaded area represents the 95% credible intervals (i.e., uncertainty in model parameters), and the light shaded area represents the 95% prediction intervals (i.e., includes parameter and sampling uncertainties).

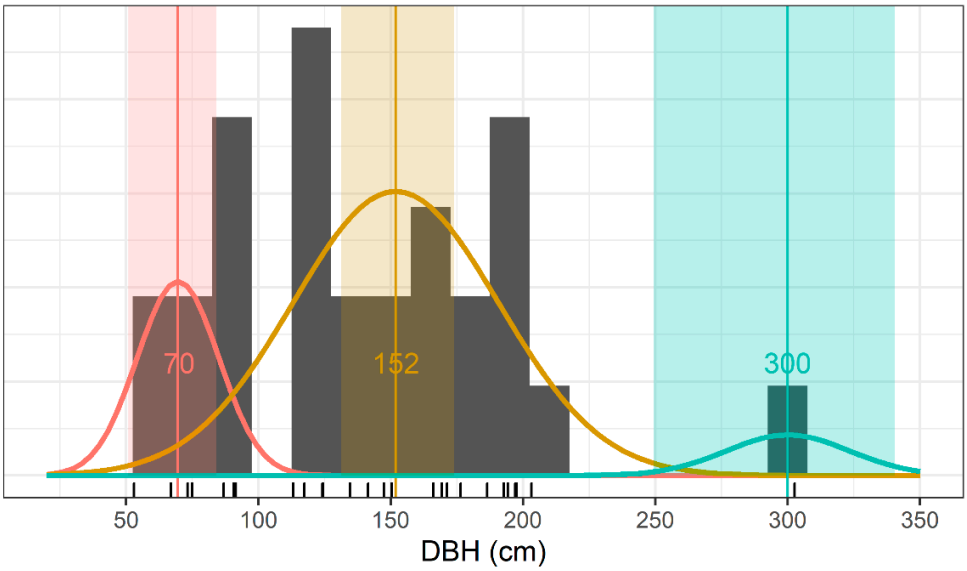

Figure S9: Finite mixture modelling of cohort size. In this specific 1-hectare plot, we identified three distinct cohorts (70 ± 8, 152 ± 11, and 300 ± 21 cm).

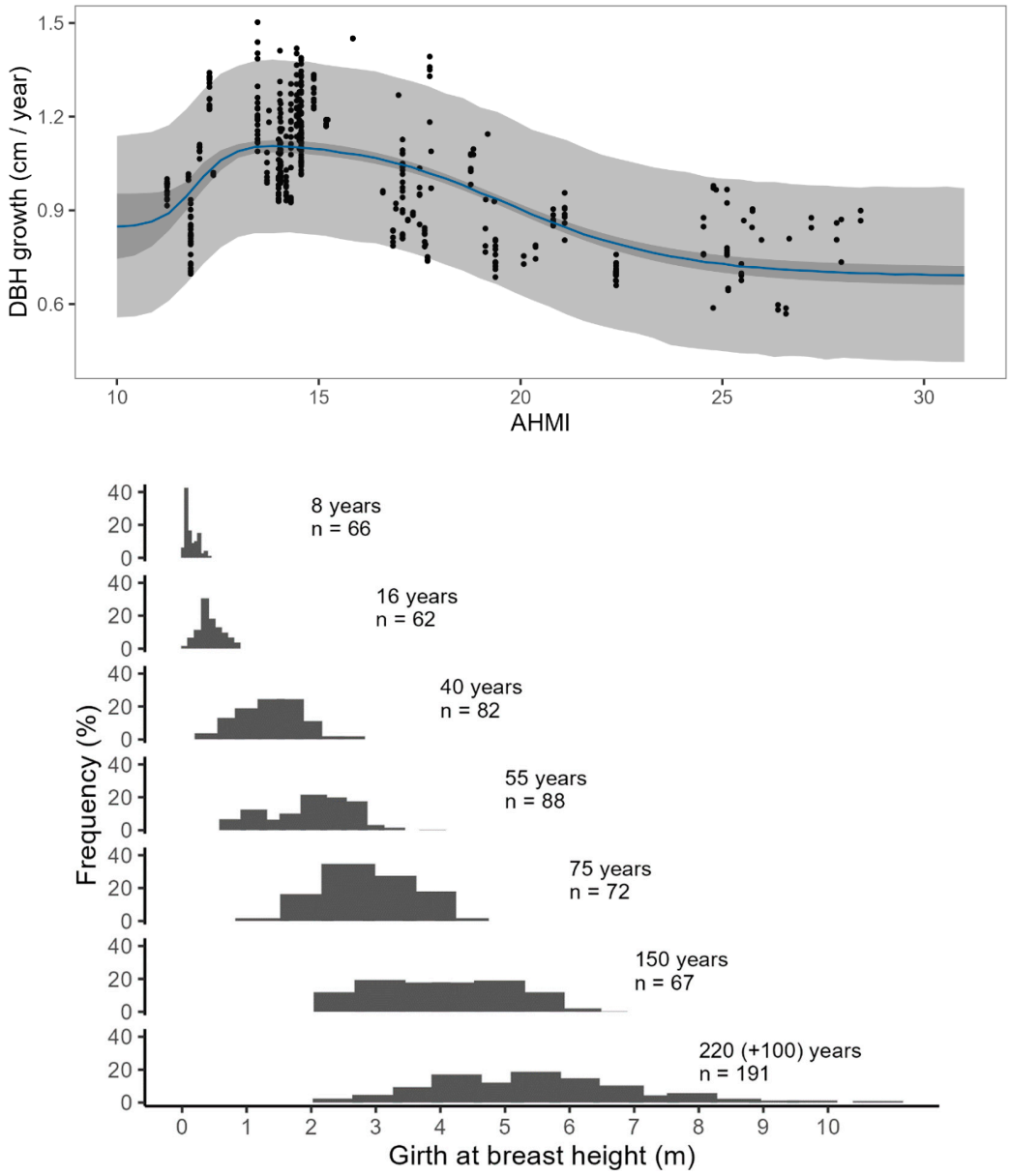

Figure S10: Top. Cohort DBH growth as a function of annual heat moisture index. Dots are observed data points, the solid blue line represents mean predictions, dark shaded area represents the 95% credible intervals (i.e., uncertainty in model parameters), and the light shaded area represents the 95% prediction intervals (i.e., includes parameter and sampling uncertainties from the log-normal likelihood). The dataset used for calibrating this function is described in (Trouvé, Nitschke et al. 2017). Bottom. Frequency histograms of girth at breast height for even-aged stands of *Eucalyptus regnans* of different ages. Redrawn from Ashton (1976). In Ashton's paper, DBH = Age^1.02, which can be approximated as a mean DBH growth of 1 to 1.1 cm per year. Ashton's growth rates are consistent with the growth rates predicted in the top row (in our calibration dataset, *Eucalyptus regnans* covered a range of AHMI between 12 and 18).

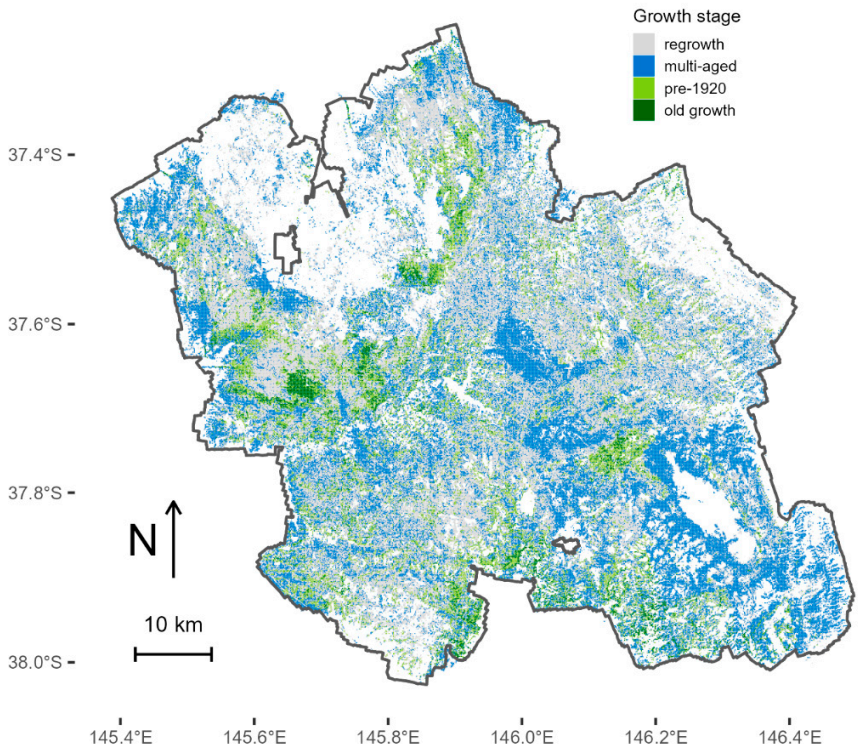

Figure S11: Map of simplified growth stage for the Central Highlands area (this corresponds to Figure 3 in the manuscript). 45.9% of the forest was classified as regrowth, 37.3% as multi-aged, 14.0% as pre-1920, and 2.7% as old growth.

S.3 Assessing uncertainty in old growth extent associated with the technical aspect of our workflow

This section details our approach for evaluating uncertainty in our estimate of old growth extent. We used Monte Carlo simulations to propagate uncertainty through the different technical steps of our workflow (Figure S12).

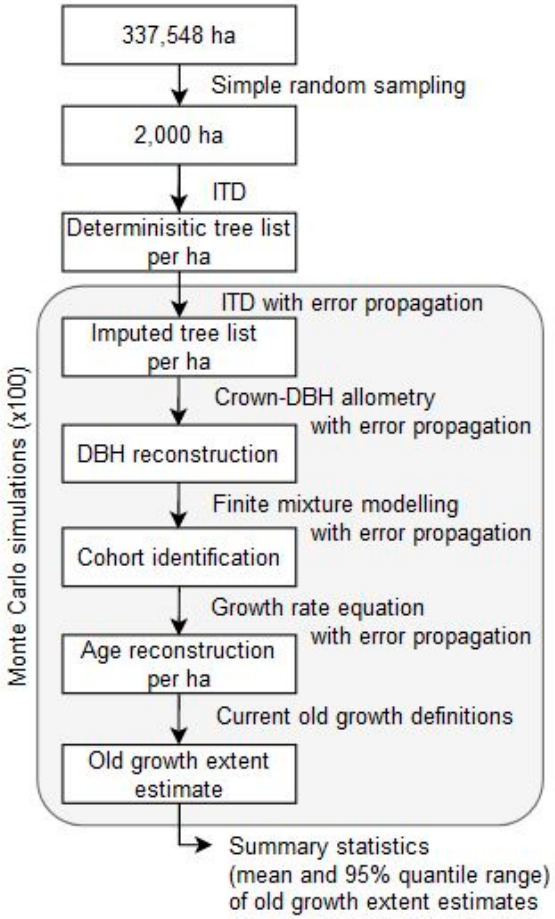

Figure S12: Conceptual diagram of uncertainty propagation using Monte Carlo simulations.

Simple random sampling. Because of computational constraints, we could not run the Monte Carlo analysis on the entire map. Instead, we ran the uncertainty analysis on a simple random sample of 2,000 ha (i.e., pixels) in the landscape. The sample size of 2,000 ha was chosen as a trade-off between Binomial sampling uncertainty[1], the number of Monte Carlo replicates that we would be able to run, and computational feasibility.

[1] The standard error formula for Binomial sampling uncertainty is SE = sqrt(p * (1 - p) / n)). With p = 2.7% and n = 2,000 samples, SE = 0.36%, which gives a 95% confidence interval range (i.e., ± 2 SE range) of around 1.5% on the proportion of old growth extent due to sampling.

The Monte Carlo analysis focused on propagating uncertainty in the main steps of our workflow: the ITD step, the diameter reconstruction step, the cohort identification step, and the age reconstruction step (see Figure 1 in the manuscript and Figure S12). The procedure was repeated 100 times. At the end of each run, we applied the current rule-based classification of old growth and computed old growth extent. This gave us a potential distribution of old growth extent, which represents the uncertainty associated with the technical aspects of our workflow and which we summarized by computing the mean and 95% quantile range of the distribution.

ITD error propagation: the refined watershed ITD algorithm that we used is deterministic. This means that rerunning the algorithm multiple times gives the same output, which can be problematic for Monte Carlo error propagation. However, we can use post-processing and data imputation methods, calibrated on the ITD accuracy in the validation sample, to inform error propagation in the ITD step. The ITD outputs has three categories of errors: trees can be missing, undersegmented, and oversegmented. The validation sample can inform us about how frequent these errors happen, and the relative size of trees affected by each of these errors. Each replication of this step creates a virtual tree list for each ha by randomly changing the deterministic tree list delineated by the ITD.

> **Missing trees** are trees that were present in the field but that were not detected by the ITD. The correction for missing trees is to add trees to the virtual tree list. Typically, these trees are smaller than average (see Figure S13a). According to our pooled validation samples (1939 and multi-cohort, which is our main target for old growth), 88 trees were detected by the ITD and 8 were missing. We simulated the number of missing trees per ha by randomly drawing from a Poisson distribution with mean set at 8/88 times the number of ITD-detected trees in each hectare. We then added these missing trees to our virtual tree list for each ha. For each ha, we predicted a count of missing trees by sampling from a Poisson distribution with mean 8/88 times the number of ITD detected trees in the ha to add to our virtual tree list per ha. We used resampling methods to impute the crown width of these missing trees. The crown width was randomly selected from the trees already present in the plot, with the probability for each tree to be selected taken from a softmax transformation of relative crown width (RCW) with a coefficient of -3: exp(-3 * RCW) / sum(exp(-3 * RCW)), as calibrated from the validation sample (see Figure S13a).
>
> **Oversegmented trees.** Oversegmented trees are large trees that have been wrongly split into smaller trees by the ITD. The correction for oversegmented trees is to merge the ITD crown area for the oversegmented tree with the crown area of a neighbor ITD detected tree. Next, we estimate the increased crown width of the neighbor tree as if it were circular. The potential number of oversegmented trees per ha was sampled from a Poisson distribution with mean 5 / 88 times the number of ITD detected tree in each ha. We then sampled which ITD detected trees were oversegmented (typically these were smaller trees; which was modelled using a softmax of RCW with a coefficient of -10; see Figure S13b). We then randomly attributed the extra crown area to one of the other ITD detected tree (with a probability of being sampled a softmax function of RCW with a coefficient of 1; see Figure S13c).

Undersegmented trees. Undersegmented trees (also called merged trees) happen when the ITD algorithm merges several trees into one larger tree. The correction for undersegmented trees is to replace it with two smaller trees while preserving the total crown area. The potential number of undersegmented trees per ha was sampled from a Poisson distribution with mean 3 / 88 times the number of ITD detected tree in each ha. We then randomly selected which trees might have been undersegmented. In our validation sample, we find no association between RCW and the probability of being undersegmented (Figure S13d). Therefore, we used a uniform distribution to select which tree (if any) was undersegmented. Undersegmented trees are split into two individuals. We assume that each individual has half the crown area of the initial tree. We then calculate the crown width for these new trees by assuming a circular crown shape.

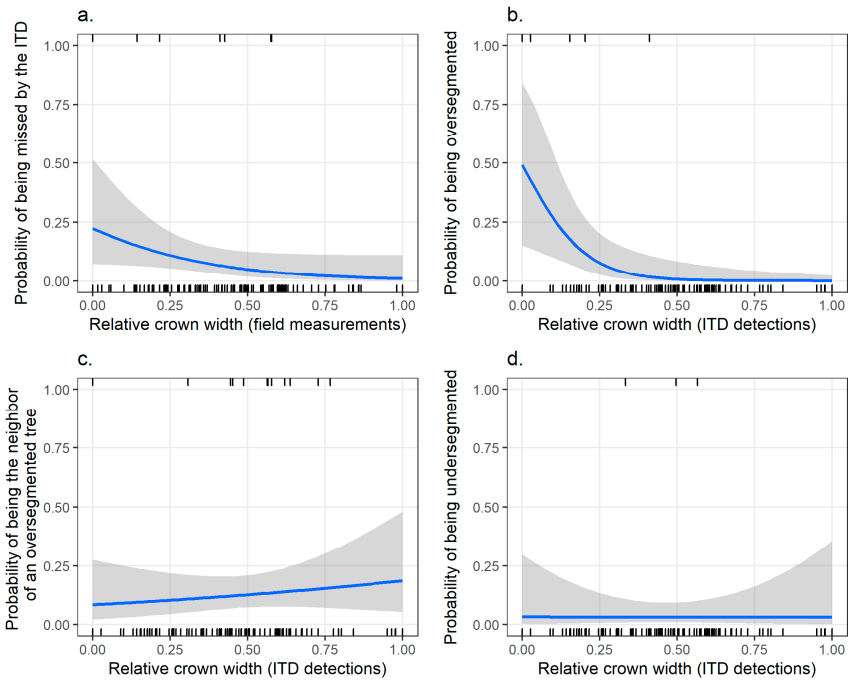

Figure S13: Different ITD classification errors as a function of relative crown width. a. Small trees (as measured in the field) were more likely to have been missed by the ITD algorithm (slope of -3). b. Small trees (as detected by the ITD) were also more likely to have been oversegmented (slope of -10). c. The neighbor of oversegmented trees were slightly larger than average (slope of 1). d. The probability of being undersegmented was independent of tree size (slope not significantly different from zero). Ticks show the validation data. The solid blue line is a logistic regression fit to the data. The shaded area is the 95% confidence intervals associated with the logistic regression fit. These modelled probabilities can be used to inform the relative size of trees affected by each error type.

At the end of the ITD error propagation step, each 1-ha pixels has been duplicated 100 times with 100 slightly different number of trees per ha and crown width distribution. While we imputed missed, oversegmented, and undersegmented trees, we did not propagate uncertainty in the ITD detected crown width, even though they varied slightly from the field measurements (see Figure S4). This is because our crown-DBH allometry was based on LiDAR-detected crown width, not field measurements.

Crown-DBH allometry error propagation. The crown DBH allometry is used to reconstruct the DBH of trees from the virtual list of reconstructed ITD crown width. In the crown-DBH allometry, there are two

types of uncertainty: parameter uncertainty (shown as dark shading in Figure S8) and residual uncertainty (represented by light shading in Figure S8, equal to 30% relative uncertainty). To address parameter uncertainty, each Monte Carlo run randomly selects a pair of potential parameter values from the variance-covariance of parameter estimates in the crown-DBH allometry model. To address residual uncertainty, we added IID tree-level residuals (on the multiplicative scale and sampled from a lognormal distribution with sigma = 0.26, i.e, 30% relative uncertainty) to reflect the fact that the individual DBH varies from the mean prediction.

Finite mixture modelling error propagation. We fitted finite mixture models to each potential DBH distribution per pixel, determining the number of cohorts per ha and a mean DBH per cohort, along with error estimates of mean DBH per cohort. To propagate uncertainty, we sampled the DBH of each cohort from a normal distribution with a mean centered on the estimated mean DBH and a standard deviation based on the estimated standard error. At the end of this stage, each pixel has 100 unique combinations of cohort structures (i.e., number of cohorts per ha, mean tree size of each cohort).

Growth rate error propagation. We use a growth rate equation to reconstruct the age of each cohort. Similar to the crown allometry equation, the growth rate equation has two types of uncertainty: parameter uncertainty (shown as dark shading in Figure S10, top row) and residual uncertainty (represented by light shading in Figure S10). To address parameter uncertainty, each Monte Carlo run randomly selects a joint distribution of potential parameter values from the variance-covariance of parameter estimates in the growth rate equation. To address residual uncertainty, we added IID tree-level residuals (on the multiplicative scale and sampled from a lognormal distribution with sigma = 0.14, i.e, 15% relative uncertainty) to reflect the fact that growth rates per pixel can vary from the mean prediction.

Compute summary statistics of old growth extent. The Monte Carlo procedure detailed above was repeated 100 times. At the end of each run, we computed the proportion of old growth extent in our sample according to the current old growth definition. This gave us a distribution of old growth extent, which we then summarized by computing the mean and 95% quantile range of the distribution. In this uncertainty propagation and under the current old growth definition, an average of 3.13% of the forest was considered old growth. The standard deviation of the estimate was 0.78% with a 95% quantile range of 1.7 to 4.8%. However, this only represents our modelling process uncertainty. We also need to account for the sampling error in our analysis (i.e., the fact that we didn't run the error propagation for all the pixels in the landscape to reduce computational complexity). Based on 2,000 sampled pixels and a mean of 3.13%, the Binomial sampling standard error is 0.39%. We used the root sum square method to add the sampling error to the modelling error, resulting in a total standard error of 0.88%. This gives us an uncertainty range (mean ±2 SE) of 1.4% to 4.9%, which is what we report in the manuscript.

S.4 Partitioning of number of cohorts per hectares by Ecological Vegetation classes (EVC)

A breakdown of number of cohorts per ha per EVC in the Central Highlands of Victoria is shown in Table S4.

Table S4: Partitioning of number of cohorts per hectares for the eight most common Ecological Vegetation Classes (EVC[1]) in the Central Highlands region of Victoria, Australia evaluated in this study.

	Number of cohorts per ha					
EVC (ha)	1	2	3	4	Total (ha)	Total (%)
EVC 29: damp forest	56,845	34,617	6,899	272	98,633	29.2
EVC 30: wet forest	68,590	24,288	4,759	242	97,879	29.0
EVC 39: montane Wet Forest	24,412	11,956	3,313	126	39,807	11.8
EVC 23: herb-rich foothill forest	13,873	8,766	1,082	38	23,759	7.0
EVC 45: Shrubby Foothill Forest	7,686	8,699	1,911	77	18,373	5.4
EVC 18: riparian forest	11,161	4,375	688	23	16,247	4.8
EVC 38: montane damp forest	10,923	4,233	778	23	15,957	4.7
EVC 31: cool temperate rainforest	8,166	2,144	366	5	10,681	3.2
Others EVCs	8,808	6,263	1,115	26	16,212	4.8
Total (ha)	210,464	105,341	20,911	832	337,548	
Total (%)	62.4	31.2	6.2	0.2		100.0

[1] EVC 29 is Damp Forest (open forest dominated by a mixture of *E. obliqua*, *E. cypellocarpa*, and *E. radiata*), EVC 30 is Wet Forest (tall open forest dominated by *E. regnans*), EVC 39 is Montane Wet Forest (tall open forest above 1000 m asl. dominated by *E. delegatensis*, *E. regnans*, or *E. nitens*), EVC 23 is Herb-Rich Foothill Forest (medium open forest dominated by a mixture of *E. radiata, E. obliqua*, *and E. cypellocarpa* with a high cover of grass on the ground), EVC 45 is Shrubby Foothill Forest (medium eucalypt forest dominated by a mixture of *E. obliqua*, *E. sieberi*, and *E. radiata*), EVC 18 is Riparian Forest (tall forest along river banks dominated by *E. viminalis* and *E. obliqua*), EVC 38 is Montane Damp Forest (open forest above 800 m asl. dominated by a mixture of *E. obliqua*, *E. cypellocarpa*, and *E. radiata*) and EVC 31 is Cool Temperate Rainforest (closed forest dominated by *Nothofagus cunninghamii*, *Atherosperma moschatum*, and *Acacia* spp., with occasional emergent eucalypts (cool temperate mixed forest)).

References:

Amiri N (2014) Assessment of Marker-Controlled Watershed segmentation algorithm for individual tree top detection and crown delineation.

Ashton D (1976) The development of even-aged stands of *Eucalyptus regnans* F. Muell. in central Victoria. Australian Journal of Botany 24:397-414. doi: http://dx.doi.org/10.1071/BT9760397

Ayrey E, S Fraver, JA Kershaw, LS Kenefic, D Hayes, AR Weiskittel, BE Roth (2017) Layer Stacking: A Novel Algorithm for Individual Forest Tree Segmentation from LiDAR Point Clouds. Canadian Journal of Remote Sensing 43:16-27. doi: 10.1080/07038992.2017.1252907

Brandtberg T, F Walter (1998) Automated delineation of individual tree crowns in high spatial resolution aerial images by multiple-scale analysis. Machine Vision and Applications 11:64-73

Dalponte M, DA Coomes (2016) Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. Methods in ecology and evolution 7:1236-1245

Dalponte M, HO Orka, LT Ene, T Gobakken, E Naesset (2014) Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data. Remote Sensing of Environment 140:306-317. doi: 10.1016/j.rse.2013.09.006

Dralle K, M Rudemo (1996) Stem number estimation by kernel smoothing of aerial photos. Canadian Journal of Forest Research 26:1228-1236

Fedrigo M, SB Stewart, S Kasel, V Levchenko, R Trouvé, CR Nitschke (2018) Radiocarbon dating informs tree fern population dynamics and disturbance history of temperate forests in southeast Australia. Radiocarbon 61:445-460. doi: 10.1017/RDC.2018.119

Huisman J, H Olff, LFM Fresco (1993) A hierarchical set of models for species response analysis. Journal of Vegetation Science 4:37-46. doi: 10.2307/3235732

Jiang R (2019) Using LiDAR for landscape-scale mapping of potential habitat for the critically endangered Leadbeater's possum, The University of Melbourne, Victoria, Australia

Jiang R (2023) Individual Tree Delineation in Broadleaf forest. https://github.com/ruizhuj/Individual_Tree_Delineation_in_Broadleaf_forest/blob/main/README.md. https://zenodo.org/record/8267560. Accessed 21/08/2023 https://zenodo.org/record/8267560

Ke YH, LJ Quackenbush (2011) A review of methods for automatic individual tree-crown detection and delineation from passive remote sensing. International Journal of Remote Sensing 32:4725-4747. doi: 10.1080/01431161.2010.494184

Kwak DA, WK Lee, JH Lee, GS Biging, P Gong (2007) Detection of individual trees and estimation of tree height using LiDAR data. J For Res 12:425-434. doi: 10.1007/s10310-007-0041-9

Ma K, Z Chen, L Fu, W Tian, F Jiang, J Yi, Z Du, H Sun (2022) Performance and sensitivity of individual tree segmentation methods for UAV-LiDAR in multiple forest types. Remote Sensing 14:298

Madhuri Kalapala DVSR, Dr. K. Srinivas (2012) Robust Tree Crown Delineation using Novel Marker Controlled Watershed Segmentation Algorithm. International Journal of Engineering Research & Technology Vol.1

Nitschke CR, S Nichols, K Allen, C Dobbs, SJ Livesley, PJ Baker, Y Lynch (2017) The influence of climate and drought on urban tree growth in southeast Australia and the implications for future growth under climate change. Landscape and Urban Planning 167:275-287. doi: https://doi.org/10.1016/j.landurbplan.2017.06.012

Nutto L, P Spathelf, I Seling (2006) Management of individual tree diameter growth and implications for pruning for Brazilian Eucalyptus grandis Hill ex Maiden. Floresta 36

Persson A, J Holmgren, U Söderman (2002) Detecting and measuring individual trees using an airborne laser scanner. Photogramm Eng Rem S 68:925-932

Popescu SC, RH Wynne (2004) Seeing the trees in the forest. Photogrammetric Engineering & Remote Sensing 70:589-604
Popescu SC, RH Wynne, RF Nelson (2002) Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size. Computers and electronics in agriculture 37:71-95
Rahman M, B Gorte (2009) Tree crown delineation from high resolution airborne lidar based on densities of high points. ISPRS
Rahman M, B Gorte, A Bucksch (2009) A new method for individual tree measurement from airborne LiDAR. Proc SilviLaser:14-16
Rahman MA, B Gorte (2008) Individual tree detection based on densities of high points of high resolution airborne LiDAR. GEOBIA:350-355
Rijal B, AR Weiskittel, JA Kershaw (2012) Development of regional height to diameter equations for 15 tree species in the North American Acadian Region. Forestry 85:379-390. doi: 10.1093/forestry/cps036
Ritchie MW, DW Hann (1987) Equations for predicting height to crown base for fourteen tree species in southwest Oregon.
Roussel J-R, D Auty, NC Coops, P Tompalski, TR Goodbody, AS Meador, J-F Bourdon, F De Boissieu, A Achim (2020) lidR: An R package for analysis of Airborne Laser Scanning (ALS) data. Remote Sensing of Environment 251:112061
Sačkov I, T Hlásny, T Bucha, M Juriš (2017) Integration of tree allometry rules to treetops detection and tree crowns delineation using airborne lidar data. iForest-Biogeosciences and Forestry 10:459
Sharma RP, Z Vacek, S Vacek, V Podrázský, V Jansa (2017) Modelling individual tree height to crown base of Norway spruce (Picea abies (L.) Karst.) and European beech (Fagus sylvatica L.). PloS one 12:e0186394
Soille P (2013) Morphological image analysis: principles and applications, Springer Science & Business Media
Trouvé R, R Jiang, M Fedrigo, MD White, S Kasel, PJ Baker, CR Nitschke (2023) Combining Environmental, Multispectral, and LiDAR Data Improves Forest Type Classification: A Case Study on Mapping Cool Temperate Rainforests and Mixed Forests. Remote Sensing of Environment 15. doi: https://doi.org/10.3390/rs15010060
Trouvé R, CR Nitschke, AP Robinson, PJ Baker (2017) Estimating the self-thinning line from mortality data. Forest Ecology and Management 402:122-134. doi: https://doi.org/10.1016/j.foreco.2017.07.027
Wang L, P Gong, GS Biging (2004) Individual tree-crown delineation and treetop detection in high-spatial-resolution aerial imagery. Photogrammetric Engineering & Remote Sensing 70:351-357
Woodgate PW (1994) A study of the old-growth forests of East Gippsland. Department of Conservation and Natural Resources. Melbourne, Victoria, Australia., Melbourne
Woodgate PW, BD Peel, JE Coram, SJ Farrell, KT Ritman, A Lewis (1996) Old-growth forest studies in Victoria, Australia Concepts and principles. Forest Ecology and Management 85:79-94. doi: https://doi.org/10.1016/S0378-1127(96)03752-8
Wu B, BL Yu, QS Wu, Y Huang, ZQ Chen, JP Wu (2016) Individual tree crown delineation using localized contour tree method and airborne LiDAR data in coniferous forests. International Journal of Applied Earth Observation and Geoinformation 52:82-94. doi: 10.1016/j.jag.2016.06.003
Wulder M, KO Niemann, DG Goodenough (2000) Local maximum filtering for the extraction of tree locations and basal area from high spatial resolution imagery. Remote Sensing of Environment 73:103-114. doi: Doi 10.1016/S0034-4257(00)00101-2
Xu X, F Iuricich, L De Floriani (2023) A topology-based approach to individual tree segmentation from airborne LiDAR data. GeoInformatica:1-30