Tree Species Traits Determine the Success of LiDAR-Based Crown Mapping in a Mixed Temperate Forest

Jack H. Hastings 1,*, Scott V. Ollinger 1,2, Andrew P. Ouimette 2, Rebecca Sanders-DeMott 2, Michael W. Palace 2,3, Mark J. Ducey 1, Franklin B. Sullivan 2, David Basler 4 and David A. Orwig 5

1 Department of Natural Resources and the Environment, University of New Hampshire, 56 College Rd, Durham, NH 03824, USA; Scott.Ollinger@unh.edu (S.V.O.); Mark.Ducey@unh.edu (M.J.D.)
2 Earth Systems Research Center, University of New Hampshire, 8 College Rd, Durham, NH 03824, USA; Andrew.Ouimette@unh.edu (A.P.O.); rebecca.sanders-demott@unh.edu (R.S.-D.); Michael.Palace@unh.edu (M.W.P.); Franklin.Sullivan@unh.edu (F.B.S.)
3 Department of Earth Sciences, University of New Hampshire, 56 College Rd, Durham, NH 03824, USA
4 Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford St, Cambridge, MA 02138, USA; davidbasler@fas.harvard.edu
5 Harvard Forest, Harvard University, 324 N. Main St, Petersham, MA 01366, USA; orwig@fas.harvard.edu
* Correspondence: jhc33@wildcats.unh.edu

Received: 27 November 2019; Accepted: 11 January 2020; Published: 17 January 2020

Abstract: The ability to automatically delineate individual tree crowns using remote sensing data opens the possibility to collect detailed tree information over large geographic regions. While individual tree crown delineation (ITCD) methods have proven successful in conifer-dominated forests using Light Detection and Ranging (LiDAR) data, it remains unclear how well these methods can be applied in deciduous broadleaf-dominated forests. We applied five automated LiDAR-based ITCD methods across fifteen plots ranging from conifer- to broadleaf-dominated forest stands at Harvard Forest in Petersham, MA, USA, and assessed accuracy against manual delineation of crowns from unmanned aerial vehicle (UAV) imagery. We then identified tree- and plot-level factors influencing the success of automated delineation techniques. There was relatively little difference in accuracy between automated crown delineation methods (51–59% aggregated plot accuracy) and, despite parameter tuning, none of the methods produced high accuracy across all plots (27—90% range in plot-level accuracy). The accuracy of all methods was significantly higher with increased plot conifer fraction, and individual conifer trees were identified with higher accuracy (mean 64%) than broadleaf trees (42%) across methods. Further, while tree-level factors (e.g., diameter at breast height, height and crown area) strongly influenced the success of crown delineations, the influence of plot-level factors varied. The most important plot-level factor was species evenness, a metric of relative species abundance that is related to both conifer fraction and the degree to which trees can fill canopy space. As species evenness decreased (e.g., high conifer fraction and less efficient filling of canopy space), the probability of successful delineation increased. Overall, our work suggests that the tested LiDAR-based ITCD methods perform equally well in a mixed temperate forest, but that delineation success is driven by forest characteristics like functional group, tree size, diversity, and crown architecture. While LiDAR-based ITCD methods are well suited for stands with distinct canopy structure, we suggest that future work explore the integration of phenology and spectral characteristics with existing LiDAR as an approach to improve crown delineation in broadleaf-dominated stands.

Keywords: LiDAR; individual tree crown delineation (ITCD); temperate forest; tree architecture
1. Introduction

Individual tree crown delineation (ITCD) via remote sensing platforms offers potential to obtain detailed tree inventory/information over large areas. ITCD has been used to map species [1], biodiversity [2], and carbon stocks [3], as well as to quantify tree structural [4] and spectral characteristics [5]. While manually delineating crowns from high resolution imagery provides accurate measurements for small scale studies [5,6], effective automated methods are necessary if efforts are to be scaled to larger geographic regions. An ideal crown delineation method would be broadly applicable across stands varying in structural and compositional complexity. Given that many forests across the globe are under increasing pressure from climate change [7], invasive pests [8], and land-use change [9], reliable methods for measuring and mapping forests takes on additional urgency. Yet, broad-scale application of automated ITCD techniques remains difficult and reliability is uncertain.

Considerable work has been done to develop and improve automated ITCD techniques [10–15]. Light Detection and Ranging (LiDAR) crown delineation methods tend to be favored over spectral methods because they are not impaired by shadow and illumination artifacts [16], and because of their ability to directly measure crown architecture [17]. However, reported accuracies of different LiDAR-based methods varies considerably [12].

The success of ITCD is directly linked to the architecture of individual tree crowns and their position relative to neighboring crowns [18]. Crown architecture controls leaf display [19], and trees must balance resource acquisition (e.g., light capture) with mechanical constraints (e.g., buckling under their own weight [20]). Tree crown shape is also plastic [21–23] and crown shape is affected by competitive interactions with neighboring trees [24,25], as well as site history and disturbance [23,26].

Despite the seemingly stochastic nature of crown and stand structural development, there are also predictable differences between needle-leaf evergreen (conifer) and deciduous broadleaf plant functional types that influence ITCD. Conifer and broadleaf species differ in their physiological traits and adaptations to acquire resources [27,28] that manifest in differences in crown shape and stand arrangement. LiDAR-based ITCD methods have been successfully applied in conifer-dominated systems [29–31], while broadleaf-dominated systems have proven to be more challenging [17,32]. Discrepancies in accuracy of ITCD methods between conifer and broadleaf systems is often attributed to the characteristic plagiotropic growth form (ellipsoidal or umbrella-shape) of broadleaf crowns that make it difficult to identify tree tops, differentiate neighboring crowns, and group split canopies of an individual crown, such as those arising from forked trees [12].

Despite the challenges, there is a need for ITCD in many regions dominated by broadleaf and mixed stands. The temperate forests of the northeastern United States are typically dense mixed species stands with closed canopies, where crowns often overlap and have irregular shape. It remains unclear the degree to which automated ITCD techniques can be employed in the region, or what factors influence the success of these automated techniques. Here, we applied five automated LiDAR-based ITCD methods in a mixed temperate forest across a gradient of plots ranging from conifer- to broadleaf-dominated. We identified tree-level factors related to tree size and crown shape and plot-level factors related to vertical and horizontal structural and compositional complexity that influenced the outcome of our delineation analysis. We discussed how the architecture of individual trees and stand arrangement influences automated crown delineation results. Finally, we commented on how understanding the ecology of conifer and broadleaf species might best be exploited to delineate trees in temperate forests.

2. Materials and Methods

2.1. Study Site Description

This study was conducted in a Smithsonian Forest Global Earth Observatory (ForestGEO) MegaPlot [33] at the Harvard Forest, in north-central Massachusetts (42°32’ N, 72°11’ W). Located within the Prospect Hill Tract of Harvard Forest, the 35-ha MegaPlot is structurally and compositionally
representative of mature central New England forests. It encompasses a continuous forest comprised of mature eastern hemlock (Tsuga canadensis) stands, mixed-broadleaf stands, remnant red pine (Pinus resinosa) plantations, and a 3-ha swamp [34]. The age structure is dominated by 75–125 year old second growth forest [35]. Dominant species include red oak (Quercus rubra), red maple (Acer rubrum), eastern hemlock, and white pine (Pinus strobus). Other common species include Norway spruce (Picea abies), American beech (Fagus grandifolia) and birch (Betula spp.). There is little terrain variation in the MegaPlot, and elevation ranges from 340.2 to 367.8 meters above sea level. Soil are primarily moderately- to well-drained loam formed from glacial till [36].

Between 2010 and 2014 a census of the MegaPlot was conducted, where all woody stems ≥ 1 cm diameter at breast height (DBH) were mapped, measured, and identified to species [34]. The height of all stems was estimated using site-specific allometric equations [37]. For the present study, we used fifteen 20 m radius subplots located across the MegaPlot (Figure S1). Plots were selected to capture a full range of tree functional composition from conifer-dominated to broadleaf-dominated.

2.2. Remote Sensing Data

LiDAR and hyperspectral data were collected over Prospect Hill by NASA’s Goddard LiDAR, Hyperspectral and Thermal (G-LiHT) sensor package [38] between 19–21 June 2012. LiDAR point cloud, canopy height model (CHM), and hyperspectral data were downloaded on 22 October 2018 (https://glihtdata.gsfc.nasa.gov/), and clipped to a 10 m buffered extent of the MegaPlot. The LiDAR point cloud has an average density of 26.98 points per m² within the MegaPlot. The hyperspectral and LiDAR CHM data have a spatial resolution of 1 m. Although LiDAR data from NEON AOP [39] were collected in 2017, we opted to use the 2012 G-LiHT data set because of its higher point density and because changes in the canopy since 2012 were minimal.

Aerial surveys of the MegaPlot were conducted throughout the 2018 growing season by an RGB camera-equipped DJI Phantom 4 Pro unmanned aerial vehicle (UAV) equipped with a stock 20 megapixel sensor. The UAV followed a pre-programmed flight path 100 m above ground level at 20 km/h, obtaining a side overlap of 70% and a forward overlap of 80% of the images. The orthomosaic, generated using AgiSoft PhotoScan, had a spatial resolution of 0.025m but was down sampled to 0.1m for use in this analysis. Using the georeferencer plugin in QGIS (v. 2.18; [40]), UAV imagery was registered to the G-LiHT remote sensing data by identifying distinguishable features (e.g., emergent white pine crowns, MET towers, wetland boundaries) in both the UAV imagery and the hyperspectral imagery. Each UAV image was aligned with the G-LiHT imagery with 20 control points and transformed using a first-order polynomial. The resulting georeferenced UAV images were found to be in good visual agreement and tree crowns aligned with those visible in the G-LiHT hyperspectral and LiDAR imagery and field-measured stem locations. We further inspected the datasets and confirmed that there was no tree dieback or notable changes to the forest canopy between G-LiHT and UAV imagery collection.

2.3. Crown Delineation

All tree crowns visually distinguishable within the fifteen plots were manually delineated by onscreen digitizing of the UAV image collected on September 13th. This study excluded understory crowns not visible within UAV imagery. Manual delineation of individual tree crowns (M\text{ITC}) was done with a stylus pen using the FreehandEditing plugin in QGIS. While crown digitization across all plots was performed on the September 13th image for consistency, multiple dates of imagery (September 13th, October 5th, October 12th, and November 4th) were used to help distinguish crowns and identify the species of each crown based on differences in shadow and phenology (Figure 1).

We manually delineated 650 tree crowns from 14 unique species. M\text{ITC} species label and associated stem attributes (DBH and allometrically derived tree height) were manually assigned during the digitization process from the ForestGEO stem data. In rare cases where a crown could conceivably belong to one of multiple stems from either the same species or stems from different species that could not be distinguished using phenology and textural cues, the crown was assigned to the stem with the
taller allometrically derived tree height. Crown area and maximum CHM-derived crown height were calculated for each M_{ITC}. Using M_{ITC} crowns, conifer fraction of each plot was calculated as the ratio of conifer crown area to broadleaf crown area.

We chose not to smooth CHM data (e.g., Gaussian filtering) prior to crown delineation analyses because our preliminary results showed smoothing either made no marked improvement on delineation success, or, in certain cases, might otherwise be clumped during manual interpretation.

We tested five automated individual tree crown (A_{ITC}) delineation techniques (Table 1) available in the R (v. 3.5.1; [41]) lidR package [42]. Four routines are surface-based methods applied to a rasterized CHM, and the fifth is a 3D method applied to a LiDAR point cloud. Dalponte2016 (DALPONTE) is a surface-based seed and region growing method [43]. Silva2016 (SILVA) is a surface-based seed and voronoi tessellation method [31]. Simple Watershed (SW)† is a surface-based watershed segmentation [44]. Marker-controlled Watershed (MCWS) is a watershed segmentation that relies on a priority seed map. Li2012 (LI)‡ is a 3D region growing method applied to a point cloud [29]. All techniques were run using the lasttrees function. Treetop priority seed points used in DALPONTE, SILVA, and MCWS were created with the tree_detection function using the lmf (local maximum filtering) algorithm [45]. SW did not rely on a priority seed map, and LI has a tree top detection built into the function. Final A_{ITC} polygons were generated using the tree_hulls function, by creating a 2D concave hull around the segmented point cloud. We chose not to smooth CHM data (e.g., Gaussian filtering) prior to crown delineation analyses because our preliminary results showed smoothing either made no marked improvement on delineation success, or, in certain cases, decreased overall accuracy of the methods.

**Figure 1.** Manual crown delineation was performed using 2018 high resolution UAV imagery. All delineations were done on the September 13th image (left panel), but other dates of imagery were used to help differentiate crowns growing in close proximity. The right panel (October 12th) gives an example of phenologic differences between species that can be leveraged to help separate crowns that might otherwise be clumped during manual interpretation.

**Table 1.** Five automated LiDAR-based individual tree crown delineation routines were evaluated in this study. † Four routines are surface-based methods applied to rasterized canopy height models. ‡ The fifth routine is a 3D method applied to a point cloud. All routines were implemented in the R package lidR, developed by Roussel and Auty [42].

<table>
<thead>
<tr>
<th>Crown Delineation Routine</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marker-controlled Watershed (MCWS) †</td>
<td>[43]</td>
</tr>
<tr>
<td>Simple Watershed (SW) †</td>
<td>[43]</td>
</tr>
<tr>
<td>Dalponte2016 (DALPONTE) †</td>
<td>[42]</td>
</tr>
<tr>
<td>Silva2016 (SILVA) †</td>
<td>[30]</td>
</tr>
<tr>
<td>Li2012 (LI) ‡</td>
<td>[29]</td>
</tr>
</tbody>
</table>
2.4. Parameter Tuning and Accuracy Assessment

To apply each crown delineation method, we optimized parameters against manually delineated crowns. Each automated delineation technique has different input parameters controlling how the algorithm searches and delineates the CHM or point cloud, and methods vary in input parameter complexity. Parameters include search window sizes, maximum height or radius values and drop-off thresholds. We first applied each automated delineation technique with default parameters. We specified a $3 \times 3$ moving window for the $l_{mf}$ tree top detection during default parameterization because a default parameter was not given.

We then tuned each technique’s input parameters to find (1) the best plot-tuned parameters—potentially unique parameters that maximized accuracy for each individual plot and (2) the best generalized parameters—a single set of parameters that achieved the highest accuracy when evaluated across all 15 plots. Parameter tuning was done using a bootstrapping approach, where input parameters were randomly selected within a predefined range (Tables S1 and S2). Following each delineation iteration, accuracy was assessed by comparing the generated AITC polygons to the reference MITC polygons. Automated delineations were paired to manual delineations (details provided in supplemental material) so that any given MITC was labeled as either correctly or incorrectly delineated. A detection accuracy score (DA) was assigned to each iteration:

$$DA = \frac{n_{TP}}{N},$$

where, $n_{TP}$ is the number of correctly delineated AITC and $N$ is the number of MITC [46]. A given AITC was considered correctly delineated (true positive) if $\geq 50\%$ of the area of both AITC and MITC overlap (Figure 2; e.g., [47,48]). Each routine, except LI, was iterated 500 times. LI was only iterated 200 times because it was substantially slower than the surface-based methods and because maximum accuracy achieved did not improve beyond the first 100 iterations. We retained tuning iterations for the highest plot-tuned parameter accuracy and the highest generalized parameter accuracy (aggregated across the 15 plots) for each automated crown delineation technique.

To further understand how each method performed at the crown-level, we characterized the incorrect MITC delineations by type of error. Therefore, each crown was ultimately assigned to one of four categories based on their overlap with AITC (Figure 2):

1. Over-segmentation: The intersecting area between AITC and MITC is greater than or equal to 50% of the area of only AITC, indicating that the automated crown is smaller than manual crown.
2. True Positive: The intersecting area between AITC and MITC is greater than or equal to 50% of the area of both AITC and MITC (as defined above), indicating that the automated crown is well matched to the manual crown.
3. Under-segmentation: The intersecting area between AITC and MITC is greater than or equal to 50% of the area of only MITC, indicating that the automated crown is larger than the manual crown.
4. False Positive: The intersecting area between AITC and MITC is greater than or equal to 50% of the area of neither AITC and MITC, indicating poor matching of the automated and manual crowns.

Given that any MITC can only be linked to one AITC, in the case where multiple AITC fell within a single MITC (as can occur with over-segmentation), the MITC was assigned to the AITC that had the highest proportion of overlapping area with the particular MITC.

2.5. Statistical Analysis

To understand the factors that influenced automated crown delineation success, we evaluated tree-level variables (field-measured DBH, CHM-derived MITC crown height, and MITC crown area) and metrics of plot-level vertical and horizontal structural and compositional complexity (canopy complexity, uniformity of crown spacing, relative density, trees per plot, and species diversity). Plot-level metrics only included stem attributes associated with MITC data.
Plot canopy complexity was estimated using the Rumple Index [49], a ratio of canopy surface area to projected ground area, describing the roughness of the canopy terrain surface. The horizontal spatial distribution of trees in each plot was calculated from ITC centroids using an aggregation index (AGI) [50], which estimates the uniformity of crown spacing. Plot stand density was expressed using a mixed species relative density equation [51], as well as by summing the number of trees per plot. We estimated the diversity of species on each plot using three common diversity indices: Shannon’s Diversity Index (H), Pielou’s Evenness Index (J), and species richness [52]. All predictor variables were standardized to have a mean of zero and a standard deviation of one by subtracting the mean and dividing by one standard deviation [53].

To identify important plot-level variables we performed univariate linear regressions between plot-level metrics and plot-tuned accuracy (n = 15) for all five crown delineation routines. We then built global multiple linear regression models including all significant (α < 5%) variables from the univariate analyses. Multicollinearity of the global model was evaluated using variance inflation factor (VIF), and we sequentially removed highly inter-correlated variables until VIF of all variables was <10 [54]. The best model for plot-level performance was chosen using a corrected Akaike Information Criterion (AIC) to account for small sample size [55].

Finally, we built mixed-effect logistic regressions to help understand which of the tree- and plot-level factors influenced agreement of AITC with MITC as a linear function of covariates in a logistic regression [56]. Logit models were built in the R package lme4 [57]. Each global model included...
tree-level variables and the plot-level variables found to be significant during the univariate analyses described above. We controlled for plot-level variability by including plot as a random effect in each model. Model selection was performed by backward elimination from the global model, and the final model was chosen by minimum AIC [55]. We took the number of times a variable was included across the five models as an indication of the importance of that variable on crown delineation.

Model accuracies were evaluated using a 10-fold cross validation, where the developed logistic relationships were each trained on 90% of the data and tested on the remaining 10%. Training and testing were performed on all 10 folds of data and the results were averaged to give an estimate of each model’s accuracy.

3. Results


Of the 650 tree crowns we manually delineated, 379 were conifer crowns, and 271 were broadleaf crowns. The range in height, DBH, and crown area was comparable between conifer and broadleaf species. On average conifers were taller and had larger DBH (Figure 3), while median conifer crown area was 27% smaller than broadleaf crowns.

![Figure 3. Density distribution of tree-level variables showing differences between conifer and broadleaf functional groups.](image)

Trees per plot ranged from 25 to 66 across the 15 plots (mean 43), and structural complexity and composition varied substantially. Average plot canopy height ranged from 15.5 m to 28.9 m, and
average crown area ranged from 13.75 m$^2$ to 47 m$^2$. Conifer fraction, as estimated by crown area, ranged from 14% to 96%. Six of the 15 plots were characterized as conifer-dominated with $\geq 50\%$ conifer fraction.

Plot-level characteristics showed varying relationships with conifer fraction (Figure S2). For example, there was a strong positive linear relationship between conifer fraction and rumple ($r: 0.89; p < 0.001$). Rumple ranged from 1.34 to 2.16, with conifer-dominated plots occupying the upper end of this range (1.69–2.16). Species evenness also had a strong linear relationship with conifer fraction ($r: −0.84; p < 0.001$), as did Shannon’s Index ($r: −0.76; p < 0.001$). Conifer fraction showed a weak, but non-significant relationship with trees per plot ($r: 0.41; p = 0.12$) and AGI ($r: 0.41; p = 0.13$).

3.2. Automated Crown Delineation Accuracy

3.2.1. Influence of Parameter Tuning and Differences in Model Accuracies

The influence of generalized parameter tuning compared to default parameters varied by method (1–41%) (Table 2). While further plot-tuning of method parameters only marginally improved overall accuracy scores (+2–5%), we chose to continue the analyses using plot-tuned results because plot-level accuracy (Table S3) improved by as much as 36% (LI) and because we were interested in understanding the factors that influenced the highest quality delineations.

Following plot-tuning overall accuracy and plot-level accuracy did not vary substantially across delineation methods. Overall accuracy ranged from 51% by SWS to 59% by LI. Though LI was marginally more accurate (+4%) than the second highest overall accuracy (MCWS: 55%), it came at a substantial increase in processing time and complexity of input parameters (and necessarily required parameter tuning to achieve high accuracy).

Plot-level accuracy ranged from a low of 27% (DALPONTE and SWS; plot 11) to a high of 90% (MCWS; plot 14). The difference between the most- and least-accurately delineated plot was >40% for all methods. Accuracy of all methods varied similarly across all plots (Figure S3), and significantly related to conifer fraction ($p < 0.05$).

Table 2. Overall site accuracy (0-1 range) of five different automated crown delineation routines. The table included default, generalized, and plot-tuned parameters. † Conifer and broadleaf accuracies are from plot-tuned model runs.

<table>
<thead>
<tr>
<th>Routine</th>
<th>Default</th>
<th>Generalized</th>
<th>Plot-tuned</th>
<th>Conifer †</th>
<th>Broadleaf †</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCWS</td>
<td>0.49</td>
<td>0.53</td>
<td>0.55</td>
<td>0.65</td>
<td>0.40</td>
</tr>
<tr>
<td>SWS</td>
<td>0.08</td>
<td>0.49</td>
<td>0.51</td>
<td>0.54</td>
<td>0.46</td>
</tr>
<tr>
<td>DALPONTE</td>
<td>0.46</td>
<td>0.48</td>
<td>0.52</td>
<td>0.63</td>
<td>0.38</td>
</tr>
<tr>
<td>SILVA</td>
<td>0.48</td>
<td>0.49</td>
<td>0.54</td>
<td>0.67</td>
<td>0.39</td>
</tr>
<tr>
<td>LI</td>
<td>0.38</td>
<td>0.55</td>
<td>0.59</td>
<td>0.69</td>
<td>0.46</td>
</tr>
</tbody>
</table>

3.2.2. Differences in Accuracy across Species

All methods more accurately delineated conifer crowns (mean 64%) than broadleaf crowns (mean 42%). Each method had trade-offs in accuracy at the species level, and no single method stood out as having the highest accuracy across all species (Figure 4). For example, SILVA delineated red pine especially well (81%), but had consistently low broadleaf accuracy scores. Alternatively, SWS had the lowest red pine accuracy (53%) but excelled at delineating red oak in comparison to other methods (+9%).

3.3. Variables Influencing Accurate Automated Crown Delineation

3.3.1. Linear Regressions

Five plot-level variables (J, trees per plot, rumple, H, and AGI) were found to be significant ($p < 0.05$) in at least one univariate regression (Figure 5). All five variables were initially included in
the multivariate global model. However, trees per plot and AGI were highly correlated ($r = 0.97$), and H and J were highly correlated ($r = 0.94$). To constrain VIF values to below 10, H and trees per plot were removed from the analysis. For each of the five multivariate regression analyses, the model including only J as an independent variable was selected as the best model. This relationship between accuracy and J was significant ($p < 0.05$) for all models (Table S4).

**Figure 4.** Automated crown delineation routines generally showed similar patterns of accuracy across the dominant species.

**Figure 5.** Prediction accuracy in relation to plot-level (a) evenness, (b) rumple index, and (c) aggregation index for one crown delineation method (DALPONTE). Points are colored to show fraction of conifer crown area per plot (conifer fraction). The relationship between accuracy and evenness was significant ($p < 0.05$) across all methods. However, the relationship between accuracy and aggregation index was only significant for DALPONTE and SILVA, and the relationship between accuracy and rumple index was only significant for DALPONTE, SILVA and LI.
3.3.2. Logistic Regressions

Global logit models consisted of tree-level variables (DBH, height, and crown area) and the plot-level variables (rumple, J and AGI) identified in the linear regression analyses. Results of the final logit models are shown in Table 3. Cross validation model accuracy ranged from 61% (MCWS) to 70% (SWS), suggesting that while we captured the most impactful variables in predicting crown delineation, there are additional factors unaccounted for. There was no single variable in the logit models that was included across all five ITCD approaches. However, all but one model (SWS) consisted of at least one tree-level variable related to tree size and one plot-level variable related to tree arrangement.

Table 3. Results of the logit models assessing the important tree- and plot-level variables influencing the odds of successful individual tree crown delineation. All variables were standardized prior to analyses. The table includes the 10-fold cross-validation (CV %) model accuracy estimates, coefficients of variables in the models and the corresponding standard error (SE) (*p < 0.05, **p < 0.01, ***p < 0.001).

<table>
<thead>
<tr>
<th>Routine</th>
<th>Variable</th>
<th>CV %</th>
<th>Coefficient</th>
<th>SE (Coef)</th>
<th>Z Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCWS</td>
<td>Intercept</td>
<td>60.77</td>
<td>0.19</td>
<td>0.16</td>
<td>1.16</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>Crown Area</td>
<td></td>
<td>−0.38</td>
<td>0.11</td>
<td>−3.41</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>DBH</td>
<td></td>
<td>0.42</td>
<td>0.14</td>
<td>3.06</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Height</td>
<td></td>
<td>0.39</td>
<td>0.28</td>
<td>1.39</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>J</td>
<td></td>
<td>−0.68</td>
<td>0.26</td>
<td>−2.64</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Rumple</td>
<td></td>
<td>−0.49</td>
<td>0.26</td>
<td>−1.89</td>
<td>0.06</td>
</tr>
<tr>
<td>SWS</td>
<td>Intercept</td>
<td>70</td>
<td>−0.1</td>
<td>0.33</td>
<td>−0.3</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>Crown Area</td>
<td></td>
<td>0.76</td>
<td>0.12</td>
<td>6.39</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Height</td>
<td></td>
<td>1.25</td>
<td>0.25</td>
<td>4.91</td>
<td>0</td>
</tr>
<tr>
<td>DALPONTE</td>
<td>Intercept</td>
<td>62</td>
<td>0.12</td>
<td>0.09</td>
<td>1.35</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>AGI</td>
<td></td>
<td>0.36</td>
<td>0.12</td>
<td>2.88</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>DBH</td>
<td></td>
<td>0.6</td>
<td>0.11</td>
<td>5.51</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>J</td>
<td></td>
<td>−0.25</td>
<td>0.12</td>
<td>−2.11</td>
<td>0.03</td>
</tr>
<tr>
<td>SILVA</td>
<td>Intercept</td>
<td>65.32</td>
<td>0.27</td>
<td>0.2</td>
<td>1.33</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>AGI</td>
<td></td>
<td>0.47</td>
<td>0.21</td>
<td>2.21</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Height</td>
<td></td>
<td>1.04</td>
<td>0.24</td>
<td>4.37</td>
<td>0</td>
</tr>
<tr>
<td>LI</td>
<td>Intercept</td>
<td>61.54</td>
<td>0.41</td>
<td>0.11</td>
<td>3.78</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Crown Area</td>
<td></td>
<td>−0.2</td>
<td>0.1</td>
<td>−1.97</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>DBH</td>
<td></td>
<td>0.5</td>
<td>0.11</td>
<td>4.39</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>J</td>
<td></td>
<td>−0.33</td>
<td>0.12</td>
<td>−2.75</td>
<td>0.01</td>
</tr>
</tbody>
</table>

All tree-level variables were important, each showing up in three of the five logit models, though not always together. Height and DBH always had a positive effect on successful crown delineation. Crown area had a negative effect in two models, but it had a positive effect on delineation for the SWS method.

Plot-level variables were not consistently important across methods. Species evenness was the most important, showing up in three models, and each time having a negative effect on crown delineation accuracy (*p < 0.05). AGI was included as positively affecting delineation in two models (*p < 0.05), while rumple was only included in one model and negatively affected delineation odds (*p = 0.06).
4. Discussion

4.1. Differences between Segmentation Methods

Automated crown delineation shows promise but remains difficult to apply in closed canopy mixed species forests. Despite methodological differences between segmentation routines, we found similar patterns in accuracy and in which crowns were best delineated. Even with parameter tuning, none of the methods produced high overall accuracy (51–59%), and the variation in accuracy across plots was similar for each method (i.e., the most and least accurate plots identified by each method were generally the same). Furthermore, all methods better delineated conifer crowns than deciduous crowns.

We found that the point cloud segmentation method (LI) only marginally outperformed the CHM-based segmentation methods. This may be because our validation method (manual delineation of UAV imagery) allowed for accurate surveying of canopy trees but did not attempt to address understory trees. Point cloud-based methods outperform CHM-based methods in detection and delineation of understory and suppressed crowns [58], but there is less of an advantage for delineation of canopy trees.

Emerging point cloud segmentation methods have taken novel approaches (e.g., Multiclass Graph Cut [15]; Mean Shift [59]) that may or may not be better suited to delineate closed-canopy temperate forests with difficult deciduous crown architecture. However, superior performance of point cloud segmentation methods is likely to be largely limited by availability of LiDAR data with high measurement density [60]. An appeal of exploring CHM-based segmentation is the wide and growing availability of high-quality CHM data, which can make these techniques more broadly applicable than point cloud segmentation approaches.

One reason that methods that relied on an external tree detection algorithm (DALPONTE, SILVA, and MCWS) all performed similarly may be because we used the same tree detection algorithm (local maximum filtering) for each method. Given that we found such little variation between these CHM-based segmentation methods, this may suggest that the choice of tree detection method (e.g., [61]) is of equal or greater importance than the choice of segmentation method (e.g., region growing vs. watershed).

Similarities in crown delineation patterns indicate that the ability to delineate crowns is driven by physical canopy traits across plots rather than by methodological differences in crown delineation algorithms. Conifer and broadleaf species have evolved traits that distinguish their ability to compete for resources and respond to disturbance and competition. In turn these traits influence tree height, crown architecture (crown spreading and leaf-display), and how crowns interact with neighboring crowns.

4.2. Tree Architecture

4.2.1. Tree Size

Taller trees and larger DBH trees were more likely to be correctly delineated using automated methods. This is in part because large trees often hold dominant positions in the canopy and tend to have more symmetrical crown shape [21]. Yet, this is also because conifers in the canopy tended to be taller and have larger DBH (Figure 3). Conifer species identified in the plots have lower average wood density than the broadleaf species [51,62]. Lower wood density is energetically efficient for height growth [63,64], while resulting in lower stand density contributions of individual trees [51]. In our study region, in higher diameter size classes, conifer species diverge from broadleaf species, growing taller. This divergence is associated with both lower wood density and evergreen foliage [65].

Conifers, especially white pine, are larger (DBH and height) at Harvard Forest because of site history and growth strategy. Much of the northeastern United States landscape has been shaped by historical land use [66]. White pine are successful colonizers on disturbed sites, and many of the large white pines are old-field pines that invaded agricultural and pastoral fields following abandonment in the mid-1800s [67]. Low density wood, and relatively high photosynthetic rates [27,63] allow white
pine to achieve rapid vertical growth, and they continue to avoid direct competition by occupying a higher canopy stratum than broadleaf species. On a canopy height model, emergent white pines appear as hotspots (Figure S4) because they often stand five or more meters above the continuous canopy; thus, they are easily detected and delineated by automated crown delineation methods.

4.2.2. Crown Spread

Smaller diameter crowns were more likely to be successfully delineated which, as was the case with height, is likely related to differences between conifer and broadleaf traits. Conifer crowns tend to spread less than broadleaf crowns, though it is possible to find white pine or hemlock that are comparable in spread to broadleaf crowns. However, conifers maintain a more rigid, apically controlled, growth form and are less likely to exhibit crown plasticity [68]. This results in a singularly defined orthotropic bole and the characteristically conical crown shape, and it is far rarer to find conifers with forked trunks and split crowns. This may be because many conifers have evolved in resource poor conditions, and invest in long-lasting low-nitrogen (N) foliage [69]. Convergent leaf and canopy structural properties (conical crown shape, clumped foliage) promote light scattering and more even/diffuse light conditions throughout the canopy, which in turn increases radiation use efficiency. Ollinger [70] pointed out that plants grown (or adapted to grow) in resource poor conditions allocate fewer resources to wood compared to foliage, constraining crown spread.

In contrast to conifers, broadleaf species invest in high-N deciduous foliage, with higher photosynthetic capacities and shorter leaf lifespans. To benefit from this strategy, broadleaf species must maximize direct light interception. Intermediate- and shade-tolerant broadleaf species (e.g., red oak and red maple) achieve this by spreading their crowns to maximize foliage display on a more even plane. Thus, many broadleaf species have weak apical control that results in plagiotropic growth forms [71]. Weak apical control allows multiple stems to compete for a dominant terminal position, the result of which can be a broad and flat crown, often with forked trunks and multiple differentiated sections within a single crown (i.e., crown splitting).

The ability to spread branches laterally is also positively associated with wood density. Higher wood density is correlated with structural properties, including resistance to splitting, rupture stress, dynamic breakage, and increased elasticity [20]. While low density wood is a lower carbon-cost approach to attain vertical expansion, broadleaf species with denser wood can expand lateral branching without compromising structural integrity [63,64]. This is in agreement with crown radius - DBH allometric equations developed by Sullivan et al. [37] at the Harvard Forest. They found the slope of the crown radius - DBH relationship to be steeper for broadleaf than conifer species, and that this relationship was related to wood specific gravity.

Red oak, in particular, often have substantial crown spread and split crowns. This type of architecture presents two major challenges for automated tree crown delineation: (1) It is difficult to define a singular local maximum (tree detection) and (2) crowns either interdigitate with neighboring crowns—resulting in under-segmentation, or crowns split—resulting in over-segmentation. We found all methods most often over-segmented red oak (Figure 6). These results agree with other studies that found broadleaf canopies are often over-segmented [17].
Figure 6. Two-dimensional density plot showing different patterns of crown delineation accuracy between conifer and broadleaf functional groups. The circles provide crown size comparison for two end member species: red pine (shown in red circles) and red oak (shown in blue circles). This figure corresponds with delineation categories described in Figure 3: the top right quadrant signifies true delineations, the top left signifies over-segmentation, the bottom right signifies under-segmentation, and the bottom left signifies false positive. This figure shows data generated using the DALPONTE method, although all methods produce similar patterns.

4.2.3. Mechanical Interactions

Mechanical interactions between neighboring crowns is another major dynamic controlling lateral branch expansion, perhaps even more than resource competition [72]. Crown shyness—gaps that form between adjacent crowns, often of the same species—can result from mechanical bud abrasion and branch damage during crown collisions [73]. While mechanical interactions occur between all adjacent crowns in closed-canopy stands, crown shyness is controlled by branch fragility and rates of regrowth following lateral branch damage [72].

Crown shyness is a common occurrence in even-aged conifer-dominated stands [74]. Crown shyness is especially visible in red pine dominated plots (Figure S5) located in an even-aged remnant pine plantation [75]. Shyness likely contributed not only to the high accuracy in these plots (as high as 80%), but also the fidelity of the delineations because gaps between adjacent crowns create defined borders for delineation (Figure 6). In comparison to broadleaf species with strong, dense branches (e.g., red oak), red pine is more susceptible to collision damage. High height:diameter ratios coupled with low wood density make the crowns of red pine susceptible to wind damage [76] through increased crown mobility and resulting high-impact crown collisions [77] that maintain crown shyness.

4.3. Species Evenness

We found that species evenness was the most important plot-level variable associated with crown delineation success. As species evenness decreased, the likelihood of successful delineation increased. Evenness was likely important because of (1) its negative relation to conifer fraction, and (2) a relationship between evenness and canopy space filling efficiency.
There was a strong relationship between species evenness and conifer fraction (Figure S2); the least even plots had the highest conifer fraction, while the most even plots tended to have the lowest conifer fraction. It is important to note that two of the low evenness conifer plots were artificial in the sense that they are remnant red pine plantation [75], though red pine can grow naturally in monocultures. However, the other low evenness conifer plot was in a natural mature hemlock stand, a common occurrence in northeastern temperate forests. Hemlock stands often have low species evenness because of deep shade cast that precludes the establishment of less shade tolerant species [78]. Further, conifer needles have high carbon to nitrogen ratios, resulting in lower soil fertility which can deter broadleaf growth [27,79].

In addition to the influence of conifer fraction, the evenness—accuracy relationship may also be reflective of increased efficiency of canopy space filling (i.e., crown packing) in higher diversity plots [22,80], which makes successful delineation more challenging. In low diversity stands, trees from the same species compete similarly for growing space [81], while in higher diversity stands niche partitioning and complementarity of crown architecture promote partitioning of resources [82,83] allowing for more efficient use of available canopy space [84,85]. As plot diversity increases crown packing increases, and it becomes increasingly difficult to differentiate neighboring crowns (Figure 7).

Figure 7. G-LiHT LiDAR point cloud comparison highlighting the differences in structure between a high-evenness broadleaf-dominated stand (A) and a low-evenness conifer-dominated stand (B). Warmer colors represent higher points in the canopy. The conifer-dominated stand exhibits higher canopy rumple, and uniformity of crown shape. The conical, less-plastic shape of conifer crowns may also reduce canopy space filling efficiency.

To further investigate the potential relationship between species evenness and crown packing, we calculated average plot NDVI [86] from the G-LiHT hyperspectral data as a proxy estimate of leaf area index and foliar density [87], assuming increased crown packing would be related to higher LAI. We found evenness is strongly related to NDVI \( p < 0.001; r: 0.9 \). However, because NDVI is also related to conifer fraction [88], we performed a partial correlation test. After accounting for conifer fraction,
NDVI was still positively correlated (r: 0.58) with species evenness, lending support to the idea that the evenness - accuracy relationship is both a result of conifer fraction and increased crown packing in higher diversity plots. Future work would benefit from a more detailed analysis exploring how crown delineation results are expressly affected by local neighborhood tree species diversity (e.g., [89]), and the interaction of neighboring crowns.


Our findings show that automated LiDAR-based ITCD methods have substantial promise for delineation of large trees. Despite lower accuracy for smaller size trees, these results are encouraging given the important role large trees play in terrestrial ecosystems [90], especially in terms of carbon accumulation [91,92]. We were able to successfully delineate 62-70% of all trees ≥ 40 cm DBH, which is promising for the prospect of tree-centric carbon mapping [3,43].

We also found these methods to perform especially well in conifer-dominated stands. High accuracy mapping of conifers has important implications for forest management and conservation. In particular, the ability to delineate mature eastern hemlock could provide important data for monitoring and mapping hemlock woolly adelgid (HWA) infestations [93]. Given the impact HWA has on the structure and composition of infested forests [94], existing crown delineation methods could aid in measuring and mapping HWA impacts.

Much of the northeast United States is comprised of aggrading second growth forest [66]. However, while our plots cover a range of structure and composition, they are undoubtedly still just a sample of the different forest types found across the northeast. LiDAR-crown delineation methods are likely to show varying degrees of accuracy based on additional factors influencing structure, such as stage of forest succession [95]. Relatively young stands in stem-exclusion stage (sensu [81]) are likely to be especially difficult to delineate because of high-stem density and intense competition. Given that tree size [96], stand structural complexity [97], and canopy surface complexity [98] often increase with stand age our results suggest that mature- and old-growth stands could be delineated with high accuracy.

4.5. Moving forward

While structurally distinct conifers stands are ideally suited for LiDAR-based crown delineation, these methods may not be ideal for delineating broadleaf canopies due the tendency of broadleaf crowns to be irregularly shaped, have forked trunks, and overlap with neighboring crowns. Broadleaf species often lack the distinctive crown architecture of conifers that allows for higher accuracy LiDAR-based crown delineation.

An alternate path for improving crown mapping of mixed- and deciduous-dominated stands may lie within the species-specific phenological and spectral traits of broadleaf species. Indeed, much of the information we relied upon to manually delineate tree crowns—subtle differences in hue and texture from RGB images—is lost in a LiDAR dataset. Even more information may be available from hyperspectral or multi-temporal RGB imagery (e.g., [99]). Many studies have shown great success for spectrally distinguishing canopy species using hyperspectral (e.g., [1]) and multi-temporal imagery (e.g., [6]), while fewer studies have made use of this wealth of information available to delineate mixed- and broadleaf-dominated forests [100–102].

Of course, the use of high resolution imagery in crown delineation has been explored (e.g., [47,48]). However, many of the studies have relied on panchromatic [4] or single band imagery [103]. Future work could benefit from development of spectral or integrated LiDAR—spectral delineation methods for distinguishing broadleaf tree crowns. The increasing wide-spread availability of spectral platforms—including airborne sensor packages that acquire co-aligned hyperspectral and LiDAR data (e.g., G-LiHT [38] and NEON AOP [39])—adds incentive to develop effective methods because of the potential to apply methods broadly.
5. Conclusions

The ability to automatically delineate individual tree crowns in all types of forests would be a major step forward for applications of remote sensing in ecology and land management. We found that LiDAR-based crown delineation methods worked well in conifer-dominated stands but were less reliable in broadleaf-dominated stands. Overall, discrepancies in accuracy appear to be driven by differences in underlying traits controlling tree architecture and how trees interact with one another rather than differences between different LiDAR-based delineation methods. While broadleaf crowns often lack the same level of structural distinction as conifers, they have unique phenology and spectral characteristics that may be exploited to improve delineation techniques. Our work demonstrates a need to develop crown delineation techniques that integrate both structural and spectral characteristics to effectively delineate mixed species stands.

Supplementary Materials: The following are available online at http://www.mdpi.com/2072-4292/12/2/309/s1.


Funding: This research was funded by the National Sciences Foundation (NSF Grants #1638688, 1920908), USDA NHAES (Hatch NH00634), NASA (Grant #12AK56G), the New Hampshire Space Grant Consortium (NASA Grant #NNX15AH79H). This work is a publication of the Harvard Forest LTER program (NSF Grant #832210).

Acknowledgments: The authors would like to acknowledge the numerous field technicians at Harvard Forest who helped collected the ForestGEO stem data.

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

References


36. Orwig, D.A.; Boucher, P.; Paynter, I.; Saenz, E.; Li, Z.; Schaaf, C. The potential to characterize ecological data with terrestrial laser scanning in Harvard Forest, MA. *Interface Focus* 2018, 8, 20170044. [CrossRef]


46. Yin, D.; Wang, L. How to assess the accuracy of the individual tree-based forest inventory derived from remotely sensed data: A review. *Int. J. Remote Sens.* 2016, 37, 4521–4553. [CrossRef]


50. Pomerening, A. Approaches to quantifying forest structures. *Forestry* 2002, 75, 305–324. [CrossRef]


53. McCune, B.; Grace, J.B. *Analysis of Ecological Communities*; MJM Sotware Design: Gleneden Beach, OR, USA, 2002.


62. Miles, P.D.; Smith, W.B. Specific Gravity and Other Properties of Wood and Bark for 156 Tree Species Found in North America. In *Research Note Nrs-38*; United States Forest Service: Newtown Square, PA, USA, 2009; Volume 35.


66. Thompson, J.R.; Carpenter, D.N.; Cogbill, C.V.; Foster, D.R. Four Centuries of Change in Northeastern United States Forests. *PLoS ONE* 2013, 8, e72540. [CrossRef]


70. Ollinger, S.V. Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytol.* 2011, 189, 375–394. [CrossRef]


© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).