Multisensor Data Fusion for Improved Segmentation of Individual Tree Crowns in Dense Tropical Forests
Mélaine Aubry-Kientz, Anthony Laybros, Ben Weinstein, James Ball, Toby Jackson, David Coomes, Grégoire Vincent

To cite this version:
Mélaine Aubry-Kientz, Anthony Laybros, Ben Weinstein, James Ball, Toby Jackson, et al.. Multisensor Data Fusion for Improved Segmentation of Individual Tree Crowns in Dense Tropical Forests. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, IEEE, 2021, 14, pp.3927-3936. 10.1109/jstars.2021.3069159. hal-03255607

HAL Id: hal-03255607
https://hal.inrae.fr/hal-03255607
Submitted on 9 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Multisensor Data Fusion for Improved Segmentation of Individual Tree Crowns in Dense Tropical Forests

Mélaine Aubry-Kientz, Anthony Laybros, Ben Weinstein, James G. C. Ball, Toby Jackson, David Coomes, and Grégoire Vincent

Abstract—Automatic tree crown segmentation from remote sensing data is especially challenging in dense, diverse, and multilayered tropical forest canopies, and tracking mortality by this approach is even more difficult. Here, we examine the potential for combining airborne laser scanning (ALS) with multispectral and hyperspectral data to improve the accuracy of tree crown segmentation at a study site in French Guiana. We combined an ALS point cloud clustering method with a spectral deep learning model to achieve 83% accuracy at recognizing manually segmented reference crowns (with congruence >0.5). This method outperformed a two-step process that involved clustering the ALS point cloud and then using the logistic regression of hyperspectral distances to correct oversegmentation. We used this approach to map tree mortality from repeat surveys and show that the number of crowns identified in the first that intersected with height loss clusters was a good estimator of the number of dead trees in these areas. Our results demonstrate that multisensor data fusion improves the automatic segmentation of individual tree crowns and presents a promising avenue to study forest demography with repeated remote sensing acquisitions.

Index Terms—Airborne laser scanning (ALS), data fusion, deepforest, high-resolution imagery, hyperspectral, 3-D adaptive mean-shift (AMS3D), tree crown segmentation.

I. INTRODUCTION

Airborne laser scanning (ALS) is a powerful technology for mapping forest biomass and tracking forest dynamics (e.g., [1]–[8]). Biomass maps are usually generated using “area-based” approaches, which reduce the complex information held in ALS point cloud into simple pixel-level summary information about forest structure. A standard methodology is to relate biomass estimates obtained from field plots to these simple summary statistics using regression models, and to then use these models to make predictions across ALS landscapes [9]–[18]. More complex individual tree crown (ITC) approaches seek to recognize ITC within ALS point clouds, then predict the biomass of these trees using allometric functions, and then, by summation, calculate biomass per unit area [6], [19], [20]. These ITC approaches are a little better than area-based approaches at mapping forest biomass, but their true value lies in tracking tree-level responses to environmental stressors, such as drought events [20], [21] and disease [22]. For this reason, there is strong interest in developing ITC approaches to data analysis [23].

ALS data have been extensively used to segment ITCs [24]–[26]. However, these ITC approaches are compromised by inaccuracies in the segmentation of ALS point clouds, particularly in dense, structurally diverse forests [6], [20], [23], [27]. For example, crowns can be relatively flat, preventing reliable detection of tree tops that can lead to oversegmentation of tree crowns [28], [29]. Some individuals have disjunct sections of the crown in the canopy, which result in disjoint clusters and oversegmentation.

Complementing ALS information with spectral data from multispectral or hyperspectral sensors provides opportunities to improve segmentation [30], [31], as well as identifying tree species [32] and phenological stage [33]. The immense diversity of tree species in tropical forests leads to a great diversity of spectral signatures, which can help in the crown segmentation process, because nearby trees are likely to differ [34], [35]. Assuming that neighboring segments with similar signatures are likely to belong to the same individual, spectral information can be used to merge neighboring segments and reduce oversegmentation. Therefore, the expectation is that segmentation will be better if geometrical information from ALS is complemented by spectral information.

In this study, we develop a method for complementing ALS with RGB and hyperspectral imagery to improve ITC segmentation. We work with the mean-shift algorithm—amongst the most effective approach—currently available for segmenting ALS point clouds of tropical rainforests [23]. This algorithm draws polygons around each predicted tree crown; we then merge neighboring spectrally similar segments to reduce oversegmentation. To evaluate if segments are spectrally similar, we...
used hyperspectral data and RGB images. The resolution of the RGB images is high, allowing identifying crowns on their texture although their color is similar. We evaluate two approaches for identifying segments to merge. First, a logistic model using hyperspectral information and high-resolution RGB imagery as input, and second, a deep learning model using high-resolution RGB imagery [23], [36]. The workflows are presented in Fig. 1. Finally, we demonstrate how this approach can be used to track individual tree mortality over time.

II. MATERIALS AND METHODS

A. Data

1) Study Site: The Paracou field station is situated in a lowland tropical rain forest with gently rolling terrain near Sinnamary (5°18′N; 52°55′W) in French Guiana. Its composition is typical of Northern Guianese rain forests; more than 500 woody species attaining 10 cm diameter at breast height (DBH) have been recorded (over 118 ha), dominated by the Lecythidaceae, Fabaceae, Chrysobalanaceae, Sapotaceae, and Annonaceae families (ordered in decreasing abundance). A total of 16 plots from the experimental site were included in this study (named plot 1–16), representing more than 62 000 trees with DBH ≥ 10 cm. Different types of silvicultural treatments of increasing intensity were applied between 1986 and 1988. In this study, unlogged control plots are accounted for the majority of the area (62.5 ha).

2) Inventories: All trees of DBH ≥ 10 cm were located by Cartesian coordinates of their trunks to an estimated precision of +/−2 m, botanically identified and their DBH were measured in 2015 (all plots), 2016 (6 plots), and 2019 (15 plots). Plot corners and points along the plot border were georeferenced with centimetric accuracy using a total station when the experimental plots were set up. Trees were then positioned within subplots (20 × 20 m) using a measuring tape. Positions of newly recruited trees were later estimated from the position of older neighboring trees. Positions of the trees are taken at the bole center.

3) ALS: ALS data were acquired in October 2015, September 2016, and November 2019 by ALTOA, operating a RIEGL LMS-Q780 sensor. On all dates, the scan frequency was 400 kHz, the final point density was above 50 points m−2, the flying altitude was 800 m, and the scan angle was +/− 25°.

4) RGB Imagery: RGB images were acquired during the same flights as the ALS scans with an IXA180 phase one camera and an 8 cm ground sampling distance. Orthorectification of the imagery was performed using the canopy digital surface model (DSM) produced from the ALS data on each date. A 5-m resolution DSM was created from the point cloud by selecting the point of maximum height on a 1-m resolution grid, resampling the DSM at 5 m, and interpolating between points using a cubic spline.

5) Hyperspectral Imagery: Imaging spectroscopy was acquired only in 2016, with a Hyspex VNIR-1600 (Hyspex NEO, Skedsmokorset, Norway) sensor-mounted alongside the Riegl scanner. Its 160 bands covered a spectral range of 414–994 nm (i.e., visible to near infrared) with a spectral sampling distance of 3.64 nm. The flight took place on a cloudless day (September 19th, 2016). Images were orthorectified and georeferenced to 1 m spatial resolution with the PARGE software [37] using the canopy DSM produced from the ALS point cloud. Atmospheric
correction was applied using ATCOR [38], removing the atmospheric disturbance and obtaining apparent reflectance. Mean spatial filtering was applied with a $3 \times 3$ window.

**B. ALS Crown Segmentation**

A 3-D adaptive mean-shift (AMS3D, [39]) algorithm was used to segment tree crowns because previous work has shown that it performs better than other methods at our study site [27]. The AMS3D algorithm assumes the point cloud to be a multimodal distribution where each mode is defined as local maxima in density and height and corresponds to a single location within an ITC [23]. The bandwidth was applied as a Pollock kernel as it allowed a wider variety of crown shapes to be segmented [40]. The shape of the bandwidth could vary from a cone to an ellipsoid with parameter $m$ [41], [42]. The size of the bandwidth adapted to the point height to allow higher crowns to be larger [23]. The segmentation was performed on the Computree platform [43].

Based on previous analysis [27], the main drawback of the ALS approach is the tendency to oversegment large crowns. These large crowns are visible from the top, and therefore, their segmentation might be improved by using additional spectral information. In this study, we focus on the segmentation of these upper canopy tree crowns. To do so, we first need to convert the 3-D point clusters into 2-D polygons for which spectral information is available. Once the point cloud had been segmented, polygons were created by first rasterizing the point cloud and keeping the cluster value of the highest point in a pixel. The pixel size was chosen to ensure that all pixels were assigned a cluster value. In our retrievals, the point density was high (>50 points m$^{-2}$) and we rasterized at a resolution of 0.5 m. We applied a majority filter ($3 \times 3$ pixels) to reduce noise. Then, polygons were created based on the clustered cells in the raster.

**C. Manual Correction**

Manual segmentation was conducted on a subset of trees in 2016 and the map of segmented trees was considered as our ground truth. We first drew polygons where we thought they should be crowns using the canopy height model (CHM) and the high-resolution RGB images from 2015 and 2016, and the hyperspectral information from 2016. These segments and their species label were then validated in the field. Hereafter, these segments are referred to as the “manually segmented crowns”; they account for 706 of the crowns in plots 3, 8, 9, 10, 12, and 16 from 155 species.

Manual correction of the automatic AMS3D segmentation was conducted in three plots of the study site (plots 4, 7, and 15) with RS data from 2016. This correction was guided by the multispectral and hyperspectral data. The correction consisted of merging, splitting, or redrawing segments. These segments are hereafter referred to as the “reference segments.”

**D. Texture in the RGB Images**

As the resolution of the RGB images was high, we expected different crowns to exhibit different textures although their color was similar (see Fig. 2). We used the high-resolution pictures to compute texture indices based on the gray-level co-occurrence matrix (GLCM). We derived seven statistics (mean, homogeneity, variance, dissimilarity, entropy, correlation, and contrast) from the GLCM for each band (red, green, and blue) using the GLCM package in R [44].

**E. Principal Component Analysis (PCA) of Hyperspectral Data**

While the spatial resolution of the hyperspectral imagery was lower than that of the RGB images, the spectral resolution was much higher, and different crowns often had differentiable spectrums. To reduce the number of dimensions, we used a PCA and discarded the first PCA component that corresponded to illumination [45], [46]. Before applying the PCA, the hyperspectral data were centered and scaled. The scikit-learn library was used to run the PCA and for centering and scaling the data [47].

**F. Segment Merging Logistic Model**

For any two adjacent AMS3D segments, we computed the probability that fusion was necessary (i.e., that the segments...
were a part of the same crown and incorrectly segmented by AMS3D. This probability depends on the following rules.

1) Allometry of height to the crown area of the two adjacent segments to prevent the creation of a small tree with too large a crown or a tall tree with too small a crown.
2) Minimum change in height inside a crown (Δheight).
3) Minimum change in hyperspectral values (ΔHSI).
4) Consistency in RGB texture (Δtexture).

The model used was a linear model, including the previous terms with a logit link of the following form:

\[ p_f = \logit^{-1} \left( \text{allometry, } \Delta \text{height, } \Delta \text{HSI, } \Delta \text{texture} \right) \]  

where \( p_f \) is the probability of fusion. To select the number of components of the PCA of hyperspectral data to keep, we gradually increased the number of components in the model and computed the Akaike information criteria (AIC). Six components were eventually selected as this model had the lowest AIC (see Fig. 2).

To estimate the parameters of the model, the manually segmented crowns were compared with the AMS3D segments. If two segments were a part of the same crown, the value of fusion for this pair was 1, while if two segments were not a part of the same crown, it was set to 0. We first introduced the different distance measures independently (either an HSI distance measure or a texture distance measure) and selected the distance for each data type that had the lowest AIC. As the number of variables did not change, this meant that we selected the variables that gave the model of maximum likelihood.

G. Variable Selection

Different distance measures could be used to compare the hyperspectral and texture differences between segments [48]. If the vectors being compared are considered orthogonal, distance functions, such as Euclidean or weighted Euclidean distance, or a difference based on the angular information, such as spectral angle, can be used. We may also consider the distance between the distributions with distance measures, such as Bhattacharyya or cumulative spectrum. We compared six commonly used distance or dissimilarity measures applied to the six components of the PCA run on the hyperspectral images, or on the texture indices (seven indices for each of the three colors): the spectral angle mapper (SAM, [49]), Canberra distance [50], Euclidean distance of cumulative spectrum [48], root mean square, and Bhattacharyya [51] and Euclidean distance.

H. DeepForest

The DeepForest package\(^1\) was recently developed to segment trees from aerial RGB images [52]–[54]. The model uses a deep learning convolutional neural network (CNN) to predict individual crowns [53]. The deep learning model was pretrained with annotations from sites of the U.S. National Ecological Observation Network. The model was tested with data from our study site, confirming that the method is robust and flexible enough to be applied in tropical forests [53]. We fine-tuned the pretrained model with the bounding boxes drawn around the reference crowns that had been segmented by the AMS3D algorithm and corrected by hand (on 18.75 ha, see Section II-C). Then, the model was applied to other plots for which manually segmented crowns were available (plots 3, 8, 9, 10, 11, 12, and 16, summing to 706 manually segmented crowns; see Fig. 1). The resulting bounding boxes were then used to refine the AMS3D segmentation. If two segments from the AMS3D segmentation intersected by more than 50% with the same bounding box generated by DeepForest, they were merged.

I. Segmentation Validation

It is difficult to validate crown segmentation in dense tropical forests as not all trees are visible in airborne imagery. Therefore, one cannot expect to tally all the trees from an aerial view and directly use stem counts as validation data. This limitation prevents the computation of detection rate, omission, and commission errors, usually used to validate ITC segmentation.

Instead, we used the crowns that were manually segmented for validation and computed congruence, and over- and under-segmentation for these crowns. Following the procedure, as described in [27], for each machine-segmented crown intersecting with a manually delineated crown, the Jaccard index was calculated. Its value measures the ratio of the area of the intersection of the two polygons over the area of their union. A crown was considered correctly segmented if the Jaccard index was above 0.5. To test the tendency of algorithms to oversegment an additional test was conducted. For each reference crown, we detected every automatic crown with more than 50% of its area inside the reference crown. They were then merged and the new Jaccard index was calculated. If it was above 0.5, the crown was considered oversegmented. A similar test was applied to detect undersegmentation in which the roles of reference and automatic segments were inverted. To detect how well the method performed with small or large crowns, we then realized this validation only with small (DBH < 30 cm), medium (30 cm < DBH < 50 cm), and large trees (DBH > 50 cm).

Then, following again the procedure of the article presented in [27], we paired trees from the inventories with segmented crowns from AMS3D only and AMS3D corrected with DeepForest, and we computed the RMSE of a species-specific allometric model estimating DBH from crown height and size. The lower the RMSE, the better the fit of the model, which meant that the crowns were better segmented. Computing the RMSE for different tree size classes allowed us to identify which size class was better segmented. We expected the correction with DeepForest to reduce the RMSE for larger trees.

J. Tracking Tree Mortality Over Time

To evaluate how ITC segmentation could improve the estimation of tropical forests’ dynamics, we compared the estimates of tree mortality based on the ITC segmentations and the area of loss of canopy height between two dates. Inventories from 2015 and 2019 were compared to retrieve the number and positions of dead trees in 15 plots of the Paracou field station totaling 94 ha to provide a means of validating the approach.

\(^{1}\)Online. [Available]: https://github.com/weecology/DeepForest
TABLE I
AIC OF LOGISTIC MODELS INCLUDING VARIABLES INDEPENDENTLY (LOWEST AIC IN RED)

<table>
<thead>
<tr>
<th>HSI distances</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectral angle mapper</td>
<td>8636</td>
</tr>
<tr>
<td>Canberra</td>
<td>8858</td>
</tr>
<tr>
<td>Euclidean distance of cumulative spectrum</td>
<td>8822</td>
</tr>
<tr>
<td>root mean square</td>
<td>8723</td>
</tr>
<tr>
<td>Bhattacharyya distance</td>
<td>8689</td>
</tr>
<tr>
<td>Euclidean</td>
<td>8775</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Texture (RGB) distances</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectral angle mapper</td>
<td>9734</td>
</tr>
<tr>
<td>Canberra</td>
<td>9697</td>
</tr>
<tr>
<td>Euclidean distance of cumulative spectrum</td>
<td>9721</td>
</tr>
<tr>
<td>root mean square</td>
<td>9688</td>
</tr>
<tr>
<td>Bhattacharyya distance</td>
<td>9560</td>
</tr>
<tr>
<td>Euclidean</td>
<td>9867</td>
</tr>
</tbody>
</table>

TABLE II
CONGRUENCE OF THE REFERENCE CROWNS WITH DIFFERENT SEGMENTATION METHODS

<table>
<thead>
<tr>
<th>Segmentation method</th>
<th>Congruence</th>
<th>Over-segmentation</th>
<th>Under-segmentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeepForest</td>
<td>73.8%</td>
<td>4.5%</td>
<td>1.7%</td>
</tr>
<tr>
<td>AMS3D</td>
<td>78.2%</td>
<td>12.7%</td>
<td>1.8%</td>
</tr>
<tr>
<td>AMS3D then</td>
<td>82.9%</td>
<td>5.2%</td>
<td>2.5%</td>
</tr>
<tr>
<td>DeepForest</td>
<td>82.2%</td>
<td>8.9%</td>
<td>1.6%</td>
</tr>
<tr>
<td>AMS3D then logistic model</td>
<td>78.2%</td>
<td>8.9%</td>
<td>1.6%</td>
</tr>
</tbody>
</table>

We used the CHM (resolution 0.5 m) from 2015 to 2019 to identify height loss in the canopy. Following the criteria used in [55], contiguous clusters larger than 4 m² and with >3 m height loss between CHM were classified as canopy subsidence areas. Trees, which were recorded as alive in 2015, intersected with these clusters, and were missing in the 2019 inventory, were classified as dead.

Automatically segmented crowns that had more than 50% of their area intersecting with a canopy subsidence area were classified as lost crowns. We compared the number of lost crowns with the number of dead trees. The number of dead trees (and lost crowns) increases with canopy subsidence number and area, so we also compared the number of dead trees to the number of lost crowns divided by the sum of the canopy subsidence area.

III. RESULTS

A. AMS3D

In total, 132,678 clusters were segmented by AMS3D (2016 point cloud). Most of these clusters are too small to correspond to canopy trees, and once rasterized and drawn as polygons, 49,168 crowns remained. A total of 78.2% of the manually segmented crowns had a congruent segment. However, 12.7% of the crowns were oversegmented (see Table II). Tree crown segments were manually corrected in plots 4, 7, and 15 (see Fig. 3). Out of the 8,097 crowns segmented by AMS3D in these plots, 1,222 were fused with another and 234 were split, creating a reference dataset comprising 6,875 segmented tree crowns.

B. Logistic Regression Model

The AICs for logistic models containing hyperspectral information were consistently lower than for models containing textural information, indicating the value of hyperspectral imaging in tree recognition. The best-supported logistic model included the spectral angle mapper from the hyperspectral data, and the Bhattacharyya distance from the texture (see Table I).

The probability of two segments (A and B) being a part of the same crowns was expressed as

\[
p_f(A, B) = \logit^{-1}\left(\beta_0 + \beta_1 \frac{D_A}{H_A} + \beta_2 \frac{D_B}{H_B} + \beta_3 |H_A - H_B| + \beta_4 \text{SAM}_{H_{SI}, A, B} + \beta_5 \text{Bhat}_{A, B}\right) \tag{2}
\]

where \(p_f\) is the probability of fusion for two segments A and B, \(D_A\) and \(D_B\) are their diameters, \(H_A\) and \(H_B\) are their heights, \(\text{SAM}_{H_{SI}, A, B}\) is the spectral angle mapper from the hyperspectral data, and \(\text{Bhat}_{A, B}\) is the Bhattacharyya distance between the segments A and B.
data, $\text{Bhat}_{A,B}$ is the Bhattacharyya distance from the texture, and $\beta_0,\ldots,5$ are the parameters.

C. DeepForest

Once applied to plots 3, 8, 9, 10, 12, and 16, DeepForest drew 10,297 boxes around crowns, while AMS3D segmented 23,173 polygons in these plots. Both algorithms had good congruence but tended to oversegment, AMS3D being worse at oversegmenting than DeepForest (see Table II).

D. Merging ALS-Derived Crowns Based on HIS and RGB Data

Both correction methods (DeepForest and the logistic model) correct over-segmentation by merging segments (see Fig. 4), but the congruence rate is improved only with the DeepForest correction method (see Table II). While the logistic model reduced the number of oversegmented crowns from 12.7% to 8.9%, the DeepForest correction reduced it to 5.2%. Moreover, 82.9% of the reference crowns were well segmented (congruence >0.5) with the DeepForest correction.

We expected that the segmentation approaches that successfully applied AMS3D and DeepForest would benefit from the advantages of both algorithms: the ability of AMS3D to segment small crowns and the ability of DeepForest to bound larger ones. Using the methodology, as described in [24], we paired the automatic segments with the stems from the field inventories. Then, using an allometric relation between DBH, $H$, and crown diameter, we computed the RMSE between the observed DBH and the DBH predicted by the model. The better the segmentation method, the lower the RMSE should be. For large trees (DBH > 50 cm), the correction with DeepForest decreased the RMSE, which showed that large crowns were better segmented once DeepForest correction had been applied. However, for very small trees (DBH < 15 cm), the RMSE was higher after the DeepForest correction (see Fig. 5). This was due to the fusion of small segments with larger ones, removing these small segments from the dataset and leaving fewer small segments available for the pairing algorithm. The segmentation validation realized only on small trees (DBH < 30 cm) showed that the crowns of these trees are well segmented (82.4% of the 68 manually segmented crowns), but the correction is not useful for these trees. Medium trees (30 cm < DBH < 50 cm) are almost as well segmented (81.8% of the 351 manually segmented crowns), the correction allows to reduce oversegmentation and to reach 83.5% of well-segmented crowns. Big trees (DBH > 50 cm) benefit the most from the correction and 82.9 of the 287 manually segmented crowns are well segmented after the correction (against 72.8% without correction).

E. Tracking Tree Mortality Over Time

In the 15 sampled plots, 3,361 trees died between 2015 and 2019, which represents 5.6% of the living trees of 2015. Out of the 2,706 trees that were situated in canopy subsidence areas (contiguous clusters larger than 4 m² and with more than 3 m height loss between the CHM of 2015 and 2019), 503 (18.6%) were classified as dead because they were missing in the 2019 inventory. The number of trees that died between 2015 and 2019 per 6.25 ha plot was related to the newly created canopy.
Fig. 6. (a) Canopy subsidence area and (b) number of crowns intersecting with a canopy subsidence area in relation with the number of dead trees in these areas.

subsidence area (see Fig. 6). However, the number of lost crowns (crowns present in 2015 intersecting by more than 50% with canopy subsidence detected in 2019) was found to be a much stronger predictor of the number of trees that were observed to have died in the field plots ($R^2$ increased from 0.595 to 0.831, see Fig. 6). Using the segmentation of AMS3D only, the correlation was not as strong ($R^2 = 0.729$), highlighting the benefit of combining ALS and spectral information to track tree mortality.

IV. DISCUSSION

Although the combination of AMS3D and DeepForest is complex and may not seem to be the simplest ITC segmentation to use, it gives better results, is easier to implement, and also faster than the logistic model developed. Other methods, mostly based on the CHM, may be simpler, although they often need a fair amount of parameter tuning. However, these methods do not perform well in tropical dense forests. By using the entire point cloud and not the CHM, and allowing the higher trees to have larger crowns, AMS3D has proven to be the most effective method to segment ITC on ALS data in our study site [27].

The methodology shows that combining 3-D ALS data with 2-D spectral data improves the segmentation of ITC in a dense tropical forest. Our approach captures the complementary aspect of what made individual crowns appear as coherent objects—i.e., change in the spectral/textural characteristics of regions or the presence of crown edges. Hyperspectral spectral data, even reduced to six components, proved more effective than the texture in merging neighboring segments from the same crown. However, hyperspectral data are uncommon and tend to be more expensive to acquire. The use of other texture features may be more cost-effective if associated with the appropriate similarity measure, as proposed, for instance, in [56] with the color contrast occurrence matrix and the Kullback–Leibler divergence.

DeepForest outperformed the logistic model approach despite using only RGB images. The correction realized with DeepForest reduced the oversegmentation by 12% without increasing undersegmentation. This is promising as RGB images are cheaper and more commonly available than hyperspectral images. The application of DeepForest to hyperspectral imagery has not yet been attempted and remains an area of active research. Combining multiple sensors for object detection in remote sensing imagery requires overcoming two main obstacles: first, how to balance data with differing spatial resolution and data input types (e.g., point cloud versus rasters), and second, accounting for potential errors in georeferencing among data products [57]. In addition, it remains unclear whether it is better to a priori select bands and hyperspectral features, as we did in our logistic model, or learn directly from large hyperspectral data cubes. Other deep learning methods have recently been developed to segment tree crowns, often using the CNNs. The mask R-CNN for instance does not simply draw bounding boxes but delineates crowns exactly [58], while the faster-CNN has been applied to ALS data [59], [60] to extract points corresponding to individual trees. These deep learning methods present additional promising opportunities for fusing data from different sensors. Deep learning methods that combine spatial and spectral learning are rapidly evolving [61] and have the potential for simultaneously learning on multiple data inputs.

Multidate remote sensing data have recently been used to study trees’ dynamics, for instance, to track treefalls in African savannas [62]–[64] or harvest trees in boreal forests [65]; to estimate fire severity, drought-induced mortality, and track tree growth and loss in the Sierra Nevada, California [66]–[69]; or to detect tree crowns’ shapes changes in Amsterdam, the Netherlands [41]. However, using multidate ALS to track tree growth in Scotland showed density-dependent biases [70]. In tropical forests, mortality has been studied through tree gaps’ dynamics [55] and branch fall [71]. We showed here that adding ITC information to tree gaps’ dynamics improves our ability to track individual tree mortality.
Tracking individual tree fate (recruitment, growth, and mortality) from repeat ALS requires a high level of segmentation accuracy. The sources of discrepancy between the repeated ALS datasets are expected to occur over time, including geometric issues from the imperfect alignment of point clouds or from vegetation sway by the wind. Similarly, even the best quality imagery will suffer from its own limitations and a slight change in sun angle can reveal or mask, otherwise invisible or visible crown. Spectral data can suffer from illumination problems and parts, or even entire crowns can be hidden by shadows of neighboring trees. So, multiple images combined with ALS may indeed be required to achieve robust tracking of individual trees. This combination can be done directly during the segmentation of the point cloud if consistent spectral information can be assigned to all points, as in the case of multispectral ALS [29]. However, geometrical and spectral characteristics are often acquired with separate sensors and the association of both kinds of data can be hampered by misalignment. A geometrically consistent fusion of both data can be achieved by backprojection of ALS point clouds on the image plane so that each point has its own spectral information (see for instance [72] for hyperspectral imagery and [73] for multispectral imagery). This colored point cloud can be then used to produce an image by interpolating the color of the points projected on a plane. However, this amounts to resampling and interpolating the spectral information and degrades the resolution of the image leading to the loss of fine texture. Moreover, hyperspectral imagery often has a lower spatial resolution than that of the RGB imagery but is spectrally more discriminating.

V. CONCLUSION

Efficiently and accurately segmenting tree crowns is challenging in dense tropical forests. We have shown that a new approach combining the state-of-the-art methods for analyzing ALS and RGB data (i.e., AMS3D and DeepForest, respectively) performs better than either method on its own. We have also demonstrated that our approach can be used to track the mortality of individuals, which could radically improve dynamics’ models used to predict forest responses to anthropogenic change.

ACKNOWLEDGMENT

The authors would like to thank the reviewers and the editor for their constructive comments on their article.

REFERENCES


Mélaine Aubry-Kientz received the M.Sc. degree in bioinformatics and modeling from INSA Lyon, France, in 2011, and the Ph.D. degree in ecology from Université de Guyane, French Guiana, in 2014. She was a Postdoctoral fellow at AMAPLab, Montpellier, France, and is now a Lecturer in AgroParisTech, UMR EcoFoG, Kourou, French Guiana. Her research is about different aspects of forest dynamics modeling, with a focus on the tropical forests of French Guiana and the use of remote sensing technologies.

Anthony Laybros received the M.Sc. degree in remote sensing from Université de Rennes 2, France, in 2016, and the Ph.D. degree in ecology and remote sensing from the Université de Montpellier, France, in 2021. His research interests are tropical forest, hyperspectral imagery, LiDAR point clouds, and machine learning methods, including deep learning methods. The development of operational methods in these different fields and their routine applications is also one of its aspirations.

Ben Weinstein received the Ph.D. degree in ecology and evolution from Stony Brook University, New York NY, USA, in 2016. He is a Postdoctoral Fellow with the University of Florida, Gainesville, FL, USA, studying computer vision for ecological monitoring.

James G. C. Ball received the M.Phys. degree in physics from the University of Oxford, Oxford, U.K., in 2014, and the M.Sc. degree in environmental technology from Imperial College London, London, U.K., in 2017. He is currently working toward the Ph.D. degree with Forest Ecology and Conservation Group, Department of Plant Sciences, University of Cambridge, Cambridge, U.K., specializing in the application of remote sensing to forest dynamics.

His research contributes to Labex CEEA’s strategic project for decrypting tropical forest phenology (PhenObs). He is interested in the application of machine learning and computer vision to understand how tropical forests are responding to global change.

Mr. Ball is a member of Magdalene College, Cambridge, U.K., and the Cambridge Centre for Earth Observation, University of Cambridge Conservation Research Institute, Cambridge, U.K.


He is a Postdoctor with the Forest Ecology and Conservation Group, University of Cambridge (https://coomeslab.org/dr-toby-jackson/), Cambridge, U.K. His work focuses on changes in forest structure over time and the effects of wind as a driver of these changes. He uses terrestrial and airborne LiDAR data to map forest structure and to assess these changes, and a range of sensors to collect data on wind speeds and tree response to wind loading.

David Coomes received the B.A. degree in natural sciences and the Ph.D. degree in botany from the University of Cambridge, Cambridge, U.K., in 1989 and 1994, respectively.

He leads the Forest Ecology and Conservation Group, University of Cambridge, Cambridge, U.K., and is the Director of the Centre for Earth Observation.

Grégoire Vincent received the M.Sc. in agronomy from the Institut National Polytechnique de Lorraine, France, in 1989, and the Ph.D. degree in systems modeling from Université de Lyon, France, in 1993.

He is a Forest Ecologist. His current work focuses on leveraging new remote sensing technologies to improve our understanding of tropical forest response to on-going climate change.