Structural attributes of individual trees for identifying homogeneous patches in a tropical rainforest

Cici Alexander*, Amanda H. Korstjens, Ross A. Hill

Department of Life and Environmental Sciences, Bournemouth University, Talbot Campus, Poole, Dorset BH12 5BB, UK (*Corresponding author; *Email: cici.alexander@gmail.com*)

Abstract

Mapping and monitoring tropical rainforests and quantifying their carbon stocks are important, both for devising strategies for their conservation and mitigating the effects of climate change. Airborne Laser Scanning (ALS) has advantages over other remote sensing techniques for describing the three-dimensional structure of forests. This study identifies forest patches using ALS-based structural attributes in a tropical rainforest in Sumatra, Indonesia. A method to group trees with similar attributes into forest patches based on Thiessen polygons and k-medoids clustering is developed, combining the advantages of both raster and individual tree–based methods. The structural composition of the patches could be an indicator of habitat type and quality. The patches could also be a basis for developing allometric models for more accurate estimation of carbon stock than is currently possible with generalised models.

1. Introduction

Tropical forests play a major role in regulating the Earth's climate, being a large sink for carbon dioxide, and storing much of the terrestrial carbon pool (Dixon et al. 1994). An accurate estimation of carbon components within a forest is a first step in the United Nations initiative for Reducing carbon Emissions from Deforestation and forest Degradation (REDD). However, limited knowledge about the quantity and spatial distribution of biomass at the landscape level has led to considerable uncertainties in the estimation of carbon stocks. Human activities such as logging and clearing of forests for agriculture and agro-forestry continue to alter the extent and composition of tropical rainforests. Natural causes such as death of large trees, and subsequent regrowth in the gaps, also contribute to the generation of patches in the landscape. This increases complexity in carbon estimation and causes fragmentation of habitats. Mapping and monitoring these structural changes are pre-requisites for devising strategies for conservation of many endangered species.

Airborne Laser Scanning (ALS), an active remote sensing technique based on the technique of Light Detection and Ranging (LiDAR), is now extensively used for describing the three-dimensional structure of forests to understand the habitat requirements of species and to quantify above-ground biomass (AGB), and thereby carbon stocks (Asner and Mascaro 2014). A standard approach to area-based AGB estimation with ALS data uses grid cells, which has limitations given that ALS datasets are generally obtained as point clouds. LiDAR metrics aggregated from the attributes of points within grid cells are highly scale-dependent, and in forests, a grid cell could include part of a large tree, or many small trees, depending on the cell size. Thus, Ferraz et al. (2016) noted that the predictive power of ALS-based AGB models decreased with increasing spatial resolution due to edge effects associated with tree crowns.

Patches with different canopy structure and composition can be distinguished in Canopy Height Models (CHMs) derived from ALS data, which could correspond to different habitat types and quality. These could also form the basis for carbon stock estimation which is mid-way between plot-based and individual tree–based approaches, in terms of accuracy, computational time and complexity. The aim of this study is to identify forest patches based on the structural composition of individual trees using ALS data in a tropical rainforest to facilitate estimations of habitat fragmentation and carbon stock. The objectives are: (i) to estimate the locations and attributes of single trees based on a Canopy Height Model; (ii) to group the single trees based on their structural attributes into homogeneous forest patches; and (iii) to analyse the attributes of trees within clusters of similar patches.

2. Study Area and Dataset

The study area (centre: 99.00°E; 1.89°N), with an area of 400 ha, is in Batang Toru in the province of North Sumatra, Indonesia. A history of logging and clearing of land for agro-forestry, selective logging to establish “forest gardens” and natural dynamics have created a mosaic of forest patches. The forests are home to a number of unique plant and animal species (Fredriksson et al. 2014), including the critically endangered Sumatran orang-utans (*Pongo abelii*).

ALS data were collected by PT McElhanney (Indonesia) between 23rd March and 4th April, 2015, using a Leica ALS-70 HP LiDAR system from a fixed wing aircraft. The flying height was between 900 m and 1350 m above ground level, and the scan half angle was 22.5°. This generated an ALS point cloud with an average density of 23.63 returns m^{-2}. The returns were classified into ground (0.97%) and non-ground (99.03%) using an algorithm based on adaptive TIN filtering implemented in Terrasolid software (Axelsson 2000; PT McElhanney 2015).

3. Methods

3.1 Attributes of individual trees

The ground returns, with an average density of 0.23 returns m^{-2}, were used to generate a Digital Terrain Model (DTM) using FUSION v3.50 (McGaughey 2009). The ground and non-ground returns were merged, and the 95th percentile height of returns above the DTM was used to generate a CHM with a cell size of 1 m; the 95th percentile height was used instead of the maximum to exclude outliers. Individual tree locations, and their heights and crown radii were estimated from the CHM, using the *CanopyMaxima* function in FUSION. This algorithm identifies local maxima using a variable sized filtering window based on canopy height variances (Popescu et al. 2002). The number, mean height and mean canopy radius of all trees within a 25 m radius of each tree were derived, using *Generate Near Table* in ArcGIS™ (v10.1), with those summary attributes assigned to each individual tree. A 25 m buffer radius was selected because less than 1% of the trees had a crown radius larger than 12.5 m.

The tree location points (X, Y) were converted to Thiessen polygons, with the attributes of the enclosed tree assigned to the polygons. In fitting the Thiessen polygons the area of the polygon was determined by the spacing between adjacent points, with adjacency based on a Triangulated Irregular Network (TIN) generated from the points. The line connecting two points in the TIN was bisected, and these bisectors formed the edges of the Thiessen polygons.

3.2 Delineation of patches and analysis of clusters

The individual Thiessen polygons were clustered into patches using the five attributes (Height and Crown Radius of each tree, and the Count, Mean Height, and Mean Crown Radius of trees within a 25 m radius) in a k-medoids algorithm implemented in MATLAB R2015. Silhouette values, a measure of the separability of clusters, were used to determine the number of clusters; the one with the lowest number of negative Silhouette values was selected as the optimum. All adjacent polygons belonging to the same cluster were merged to generate patches in ArcGIS™. All the patches with an area less than 0.25 ha were merged based on the longest shared border. Statistical analyses were performed in MATLAB with α set to 0.001. Crown areas and Thiessen polygons were compared using a Pearson correlation. An ANOVA (one-way analysis of variance) was used to test for differences between clusters, using Scheffe's procedure for post hoc pair-wise comparisons.

4. Results

4.1 Identification of single trees

The mean height of the CHM (Figure 1) was 20.37±7.31 m. There were 34,484 trees identified with heights ≥ 5 m within the study area, with an overall tree density of 86.21 trees ha^{-1}. The mean tree height was 21.26±6.98 m, and the mean crown radius was 6.39±2.08 m. The mean number of trees within a radius of 25 m for all trees was 22.35±12.37, and their mean crown radius was 6.35±1.01 m. The mean crown area calculated from the crown radii was 141.77±93.87 m^2, whereas the mean area of Thiessen polygons was 115.99±84.13 m^2. The areas of Thiessen polygons correlated only moderately with the crown areas (r=0.4; n=34484; *p*<0.001).

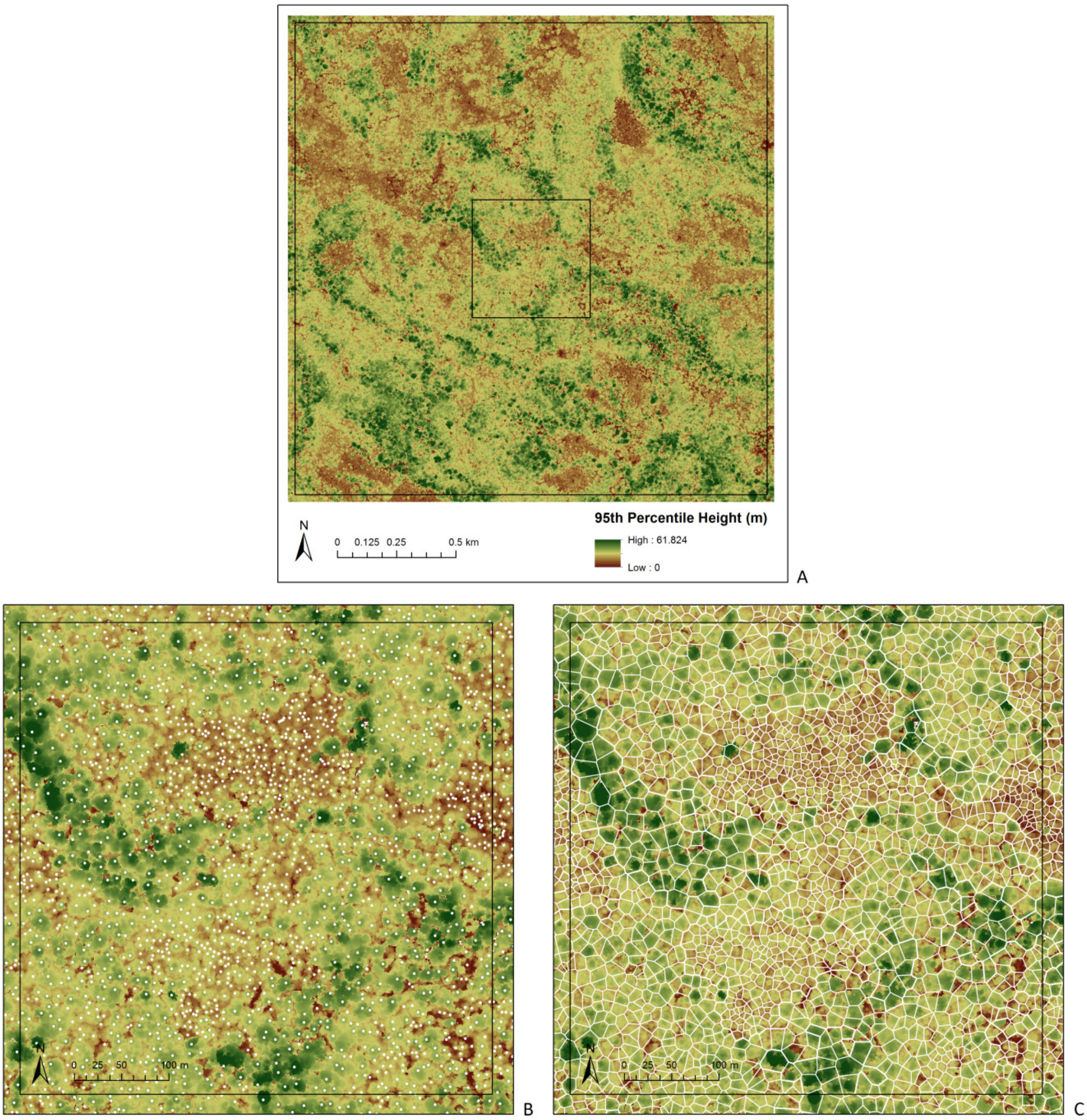

Figure 1: Canopy Height Model generated from the ALS dataset (A); locations of individual trees in a subset of the study area (B); and the Thiessen polygons generated from the locations of individual trees (C)

4.2 Delineation of forest patches

The tree clustering process identified an optimum number of five cluster types based on the five input structural variables. The shortest trees (Cluster 2) occupied only 4.58% of the area, while accounting for 13.38% of the tree count, while the tallest trees (Cluster 3) occupied 21.97% of the area, with only 8.86% of the tree count. Cluster 4 (mean tree height: 25.58 m), covered the largest area (37.26%), based on the clustered Thiessen polygons (Table 1). There were 1082 patches when the Thiessen polygons were merged based on clusters, with a mean area of 0.37±2.94 ha. These were merged into 189 patches with a mean area of 2.11±8.71 ha, by iterative merging of patches with an area less than 0.25 ha (Figure 2).

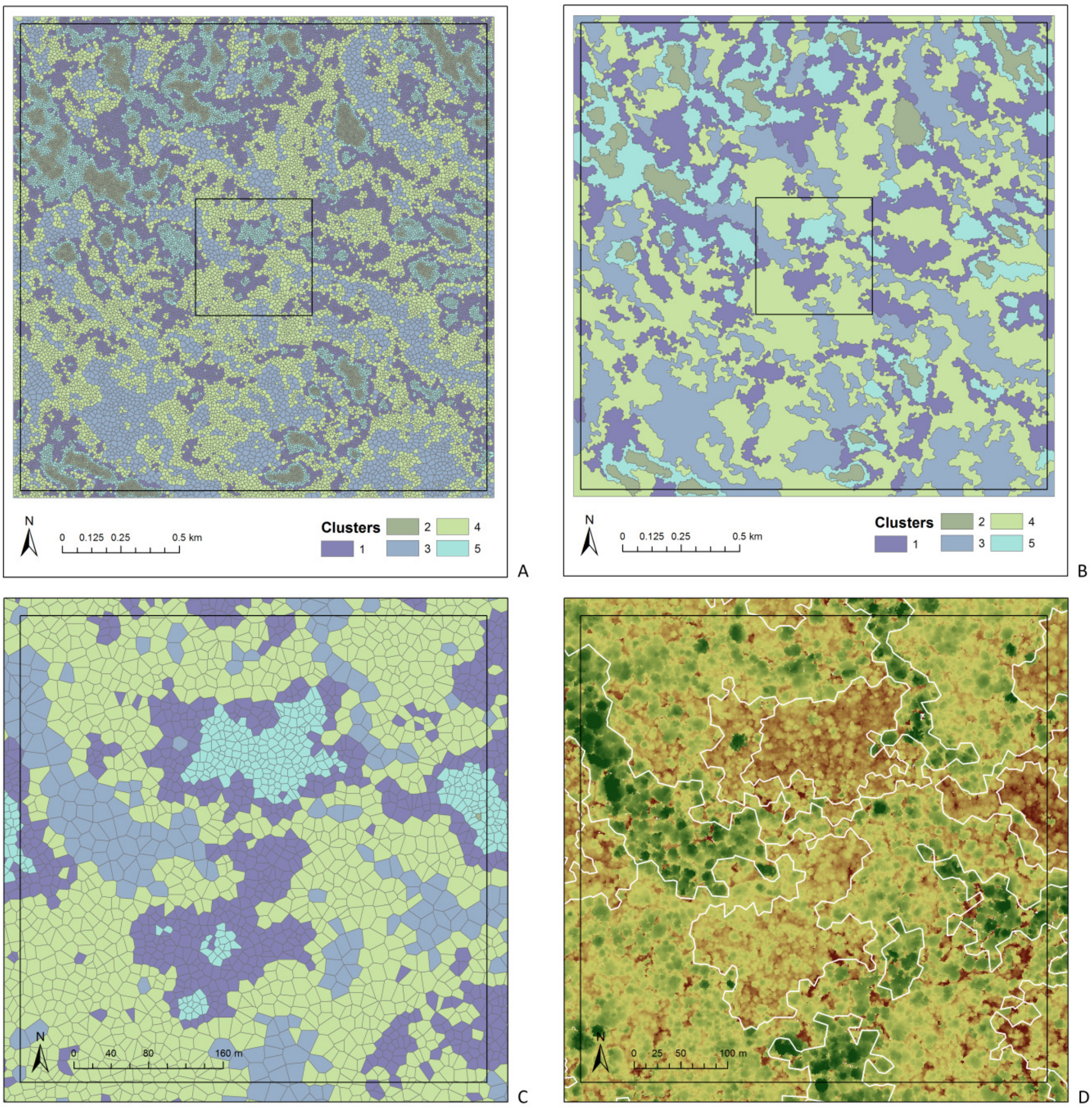

Figure 2: Thiessen polygons grouped into clusters based on the attributes of individual trees for the whole study area (A) and for a subset (C); the merged patches after dissolving the patches with area less than 0.25 ha for the whole study area (B), and boundaries of patches overlaid on the Canopy Height Model for a subset (D)

4.3 Analysis of clusters

The mean height, mean crown radius and density of trees in each patch (Table 1; Figure 3) were significantly different between the clusters (all $p<0.001$; $F_{4,183}=1032.41$; $F_{4,183}=132.9$; $F_{4,183}=679.3$ respectively). When the clusters were compared pairwise, all differences were significant except for the crown radii for clusters 2 and 5 ($p=0.002$), and the density of trees for clusters 3 and 4 ($p=0.042$).

Table 1: Attributes of patches within the five clusters before (above) and after (below) merging

Cluster ID	1	2	3	4	5
Number of Patches	414	62	181	268	157
	62	20	39	33	35
Total Number of trees	9832	4615	3054	9219	7764
	9573	4256	3057	9693	7905
Total Number of trees (%)	28.51	13.38	8.86	26.73	22.51
	27.76	12.34	8.86	28.11	22.92
Total Surface Area (%)	24.20	4.58	21.97	37.26	11.99
	23.31	4.28	21.52	38.55	12.34
Overall Density (Trees ha^{-1})	101.57	251.99	34.75	61.85	161.90
	102.68	248.60	35.51	62.87	160.13
Mean Height of Trees (m)	19.89±3.64	13.93±2.57	35.68±5.05	25.58±3.55	16.54±3.18
	19.99±3.90	13.99±2.61	34.49±6.46	25.32±4.37	16.60±3.29
Mean Crown Radius of Trees (m)	6.33±1.87	4.91±1.40	7.82±2.37	7.33±2.04	5.65±1.69
	6.39±1.89	4.93±1.42	7.62±2.37	7.24±2.08	5.65±1.70

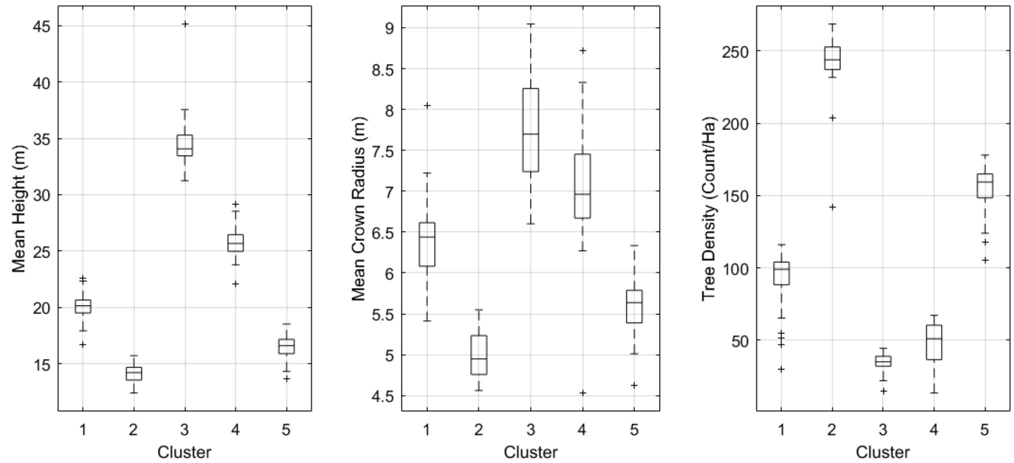

Figure 3: Box Plots of mean height, mean crown radius and overall density of trees in the five clusters (after merging)

5. Discussion and Conclusion

Identification of homogeneous patches in tropical forests based on tree heights, crown radii and density could have relevance for estimating habitat fragmentation and biomass. The method developed in this study, based on Thiessen polygons and k-medoids clustering, groups trees of similar structural attributes combining the advantages of raster and individual tree–based methods. The structural composition of the patches could be an indicator of habitat type and quality for species which are increasingly under threat from anthropogenic and natural disturbances. Distances between suitable habitats, in the case of fragmentation, could potentially be more accurately estimated using these tree crown–following tessellations rather than grid cells, especially if they are at a low resolution.

Natural and anthropogenic factors have contributed to the generation of a mosaic of forest patches in the study area, which were clearly visible in the CHM. The tallest trees with the largest crown radii (Cluster 3) occupied a large percentage of the area but had relatively low tree density. Mapping the extent of these tall patches is important, even if the accuracy of estimated tree density is low, since the large trees account for most of the biomass in tropical forests. They also serve as a focal point for biological activity and create large gaps at death, altering the forest structure dynamics in addition to releasing the sequestered carbon (Chambers et al. 2007; Ferraz et al. 2016). Note that the average estimated tree density in the study area (86.21 trees ha^{-1}) could be lower than the actual density since the algorithm identifies only the dominant and co-dominant trees as represented in the CHM (McGaughey 2009).

The areas of Thiessen polygons did not have a high correlation with crown areas since the polygons were constructed based only on the distances between adjacent points. The associated Thiessen polygon would be larger than the estimated crown area if the distance from a tree to adjacent trees is greater than its crown radius. However, this would not pose a problem in the delineation of patches, and the estimation of patch areas could in fact, be more accurate than one based on crown areas if tree crowns overlap. The distance for locating neighbouring trees could be modified based on the study area, especially if there are a high proportion of very large or very small trees.

Although generalised allometric models have been developed for tropical forests (Chave et al. 2005), species-specific and site-specific models are considered to be more accurate. Even within a single species, the biomass may vary depending on environmental factors such as rainfall, soil or topography (Litton and Kauffman 2008). The species diversity in tropical forests is extremely high, and AGB estimations based on individual tree species may be difficult with the current knowledge of taxonomy and tree species distribution. Allometric models based on identified clusters within a landscape could be developed based on systematic field surveys, rather than random sampling of stands or limited species-specific models, leading to more accurate estimation of carbon stock in tropical forests.

Acknowledgement

This research has received funding through EU Marie Skłodowska-Curie Actions (H2020-MSCA-IF-2014) under grant agreement n° [657607], and is part of the LEAP (Landscape Ecology and Primatology) project. We are grateful to Graham Usher and Gabriella Fredriksson (Sumatran Orangutan Conservation Program) and Johannes Sagala (Sarulla Operations Ltd.) for the Airborne Laser Scanner data, to PT McElhanney for airborne data collection, and initial processing, and to Robert J. McGaughey (USDA) for help with FUSION.

References

Asner, G.P., & Mascaro, J. (2014). Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric. *Remote Sensing of Environment, 140*, 614-624

Axelsson, P. (2000). DEM generation from laser scanner data using adaptive tin models. *International Archives of Photogrammetry and Remote Sensing, XXXIII, Part B4*, 110-117

Chambers, J.Q., Asner, G.P., Morton, D.C., Anderson, L.O., Saatchi, S.S., Espírito-Santo, F.D.B., Palace, M., & Souza Jr, C. (2007). Regional ecosystem structure and function: ecological insights from remote sensing of tropical forests. *Trends in Ecology & Evolution, 22*, 414-423

Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chambers, J.Q., Eamus, D., Fölster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J.-P., Nelson, B.W., Ogawa, H., Puig, H., Riéra, B., & Yamakura, T. (2005). Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia, 145*, 87-99

Dixon, R.K., Solomon, A.M., Brown, S., Houghton, R.A., Trexier, M.C., & Wisniewski, J. (1994). Carbon pools and flux of global forest ecosystems. *Science, 263*, 185-190

Ferraz, A., Saatchi, S., Mallet, C., & Meyer, V. (2016). Lidar detection of individual tree size in tropical forests. *Remote Sensing of Environment, 183*, 318-333

Fredriksson, G., Nowak, M., & Usher, G. (2014). Biodiversity Monitoring Stage 1: Camera trapping and Wildlife Surveys. In PanEco / Yayasan Ekosistem Lestari (Ed.), *Biodiversity Monitoring for Sarulla Operations Ltd.* Indonesia

Litton, C.M., & Kauffman, J.B. (2008). Allometric Models for Predicting Aboveground Biomass in Two Widespread Woody Plants in Hawaii. *Biotropica, 40*, 313-320

McGaughey, R.J. (2009). FUSION/LDV: Software for LIDAR data analysis and visualization. *US Department of Agriculture, Forest Service, Pacific Northwest Research Station: Seattle, WA, USA, 123*

Popescu, S.C., Wynne, R.H., & Nelson, R.F. (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

PT McElhanney (2015). LiDAR & DAP Survey. In PT McElhanney (Ed.), *Project Report for BUT Sarulla Operations Ltd*. Indonesia