LIDAR-based estimation of bole biomass for precision management of an Amazonian forest: Comparisons of ground-based and remotely sensed estimates

Evandro Orfanó Figueiredo a, Marcus Vinício Neves d'Oliveira b, Evaldo Muñoz Braz c, Daniel de Almeida Papa d, Philip Martin Fearnside e,f,*

a Embrapa Acre, Rodovia BR-364, km 14, CEP 69590-056 Rio Branco, Acre, Brazil
b Embrapa Acre, Rodovia BR-364, km 14, CEP 69590-056 Rio Branco, Acre, Brazil
c Embrapa Florestas, Estrada da Ribeira Km 111, CP 319, CEP 83411-000 Colombo, Paraná, Brazil
d Embrapa Acre, Rodovia BR-364, km 14, CEP 69590-056 Rio Branco, Acre, Brazil
e Instituto Nacional de Pesquisas da Amazônia (INPA), Av. André Araújo, 2936, Petrópolis, CEP 69067-375 Manaus, Amazonas, Brazil
f Brazilian Research Network on Climate Change (RedeClima), Brazil

ARTICLE INFO

Article history:
Received 4 October 2015
Accepted 12 October 2016
Available online xxxx

Keywords:
Airborne LIDAR
Linear regression models
Forest profiling
Logging
Rainforest
Tropical forest

ABSTRACT

Based on airborne LIDAR data on canopy morphology and height of Amazon forest trees, we developed allometric models to estimate dry biomass stored in the boles of dominant and co-dominant individuals and compared these results with those from equations based on traditional variables such as diameter at breast height (DBH). The database consisted of 142 trees of interest for logging in a forest under management for timber in Brazil's state of Acre. The trees chosen for study were selected through proportional sampling by diameter class (ranging from 45 to 165 cm DBH) in order to properly represent the dominant and co-dominant tree populations with diameters appropriate for harvest. Subsequent to LIDAR profiling of these trees, they were felled, subjected to a battery of dimensional measurements and sampled for wood-density determination. A set of models was generated, followed by model selection and identity testing in order to compare groups of basic wood density (low, medium and high). The morphometric variables of the crown had high explanatory power for bole biomass independent of whether the allometric equations included DBH. When calculating bole biomass with equations that include basic wood density, the best estimate is obtained using variables for both DBH and crown morphology. To obtain an allometric equation that encompasses species in all three classes of basic density, one should either use only independent variables representing crown dimensions or complement these with variables for basic density (BD) and total height (H). The study demonstrates the feasibility of using ground-based measurements to calibrate biomass models that include only LIDAR-based variables, allowing much larger areas to be surveyed with reasonable accuracy. The present study is designed to produce data needed for forest management, but the methods developed here can be adapted to studies aimed at reducing the uncertainty in biomass estimates of whole forests (not just harvestable trees) for use in quantifying carbon emissions from forest loss and degradation.

© 2016 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

In recent years great advances have been made in the planning and implementation of forest-management operations in Brazil’s Amazon region using precision management techniques (Figueiredo et al., 2007). Precision management integrates the use of geographical positioning system (GPS) and geographical information system (GIS) technology. Airborne LIDAR (Light Detection And Ranging) technology has recently been shown to have wide application in precision management of tropical forests, allowing information on relief and hydrographical structure to be obtained with sub-meter accuracy over large tracts of forest (d’Oliveira et al., 2012). Dubayah et al. (2010), Stark et al. (2012), Sullivan et al. (2014) and Palace et al. (2015) have described LIDAR’s potential in modeling forest carbon stocks, while Hunter et al. (2013) proposed corrective measures to improve estimates of forest biomass parameters using LIDAR. Use of laser profiling improves the quality of planning for infrastructure (such as the network of roads, storage yards and skid trails in the monitoring of forest operations) and in estimating the volume and biomass of managed forests.

http://dx.doi.org/10.1016/j.rse.2016.10.026
0034-4257/© 2016 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Despite technological advances in optical physics, remote sensing, GIS and computing (Hudak et al., 2012; Lim and Treitz, 2004; Næset and Gobakken, 2008; Simbaña et al., 2016), it is still necessary to further develop basic knowledge of forest components, such as the understanding of plant biomass. Biomass estimates are considered to be empirical since the models used to describe a response variable do not identify the causes or explain the phenomena that affect the behavior of this variable (Clark and Clark, 2000; Scolforo et al., 2004, 2008; Vanclay, 1994; Whittaker and Woodwell, 1971). Allometric equations used to estimate volume, biomass and carbon stocks in forests have usually been prepared based on destructive plots, correlating measurements of whole trees with two variables that are possible to measure in the field: height and diameter at breast height measured 1.3 m above the ground or above any buttresses (DBH) (da Silva, 2007; Higuchi et al., 1998).

Morphometric variables of the crown have high correlations with dendrometric parameters of the bole such as DBH and height (Durlo and Denardi, 1998; Orellana and Koehler, 2008; Wink et al., 2012). However, estimating the values of these variables by measuring individual trees in tropical forests is a major challenge. Even for measuring diameter and height, the dense understory, crooked trunks and the presence of roots in strut or tabular form are obstacles to making precise measurements. Especially for variables that cannot be measured directly in the field (i.e., height) the gain produced by inclusion of the variable in the model should not be smaller than the error associated with its measurement in the field (d’Oliveira et al., 2012). In this environment, it is difficult, from a practical standpoint, to obtain morphometric variables for the canopy in conventional forest inventories (Ferraz et al., 2015; Wulder et al., 2012). However, LIDAR data allow measurements of total height and morphological variables for the canopies of co-dominant and dominant trees to be obtained with great precision.

The aim of the present study is to develop allometric equations for estimating stem biomass of dominant and co-dominant trees under precision forest management. A combination of field-based measurements and LIDAR-derived estimates of canopy geometry was used to estimate bole biomass. The equations are based on morphometric variables for the canopy obtained from LIDAR, together with traditionally employed variables such as DBH, total height (Ht), and the apparent density (AD) and basic density (BD) of the wood. Identity was also assessed for groups of models of wood density.

2. Materials and methods

2.1. Study site

The studies were conducted in a 315-ha area of forest management in the Antimary State Forest (68° 01’ to 68° 23’ W; 9° 13’ to 9° 31’ S) under SmartWood Certification No. SW-FM/COC-1670 and Environment Institute of Acre (IMAC) Operating License No. 530/2008 (renewal). This protected area is located in the municipalities of Bujari and Sena Madureira, Acre state, Brazil (Fig. 1).

The area encompassing the Antimary State Forest has an average annual rainfall of 2000 mm and average temperature of 25 °C (Acre, 2000). A dry season from June to September is the period when the logging is performed. The forest consists of three main types: dense, open, and open with bamboo. These three forest types occur intermittently in the study area. The predominant soils are dystrophic and yellow latosols.
(Oxisols) with gently rolling topography and a maximum elevation of around 300 m (Acre, 2000).

2.2. Forest inventory

In 2009 an inventory census was conducted in order to plan the logging in the 4000-ha 2010 annual-production unit (APU). The survey was conducted considering all commercial timber species with DBH above 30 cm. The location of the trees and the planning techniques were performed according to the procedures recommended for precision management (Figueiredo et al., 2007). Thus, for each tree the geographical coordinates and a barometric point were determined using a high-sensitivity GPS. The botanical specimens of the species in the sample were deposited for identification in the herbarium of the Federal University of Acre Zoobotanical Park (UFAC/PZ), Rio Branco, Acre, Brazil.

2.3. LIDAR data

High density (25 pulses m$^{-2}$) discrete-return LIDAR data were collected before logging between 29 May and 3 June 2010, covering a total of 1000 ha inside the 2010 APU. The Optech ALTM 3100 (Aerial Laser Terrain Mapper) system was used, carried on a twin-engine Piper Seneca II aircraft, model Neiva/Embraer 810C. The flight was conducted at an average speed of 210 km h$^{-1}$ at a height of 300 m; the LIDAR system had a beam diameter of 20 cm, beam divergence of 0.3 mrad, scanning angle of 15°, and scanning frequency of 58.7 Hz. (D'Oliveira et al., 2012).

As a fixed benchmark on the ground the RBMC INBO 93911 reference was adopted from the Brazilian Network for Continuous Monitoring of Systems (GNSS). The data were processed in the Universal Transverse Mercator (UTM) coordinate system (Zone 19 South) and the SIRGAS 2000 reference system. Two pairs of Trimble 5700 GPSs with L1/L2 carriers were used. The average intensity of profiling was 43.03 dots m$^{-2}$.

The LIDAR reflection data were initially placed in a single structured file forming a 315-ha mosaic with approximately 130 million pulses. Files were processed on a three-dimensional basis using Quick Terrain Modeler software, which is specific for this purpose.

2.4. Quantification of bole volume

The sample structure of the dominant and co-dominant trees of interest for logging under forestry-management conditions was measured considering a proportional sampling by diameter class (Table 1). The sampled trees were cut and cubic scaling was performed using the Smalian method. After cutting the trees, the stumps were geo-referenced in the herbarium of the Federal University of Acre Zoobotanical Park (UFAC/PZ), Rio Branco, Acre, Brazil. The sampled trees were cut and cubic scaling was performed using the Smalian method. After cutting the trees, the stumps were geo-referenced considering a proportional sampling by diameter class (Table 1).

2.5. Quantification of biomass

After cutting each tree in the sample, wedges were removed with an average length of 30 cm and thickness of 10 cm. The wood samples (including bark and sapwood) were taken from the first and last log of each tree, using a chainsaw.

To obtain the apparent density (green weight/green volume in g cm$^{-3}$), the samples were immersed in water to achieve constant weight, as recommended for the hydrostatic balance method (ABNT, 2003). An electronic balance was used with 25 kg capacity and a sensitivity of 2 g. Subsequently, the dry weight of the sample for calculating basic density (dry weight/green volume) was obtained after the samples reached constant weight at 105 °C (± 2 °C). Using the data on the density of the samples from each tree of interest, we calculated the dry biomasses from the volume obtained by rigorous cubic scaling.

2.6. Processing the LIDAR point cloud to obtain canopy morphometric variables

Each sampled tree was georeferenced in the field and the canopy of each tree was subsequently identified in the LIDAR point cloud. The advantage of working with dominant and co-dominant trees is that individual tree crowns can be pulled out of the LIDAR data even in a structurally complex forest because these tree crowns are at the same level or above the main canopy height (Detto et al., 2015; Hunter et al., 2015; Tochon et al., 2015). Physiognomic characteristics of the environment we studied (open forest and open forest with bamboo) were also of fundamental importance for isolating dominant and co-dominant trees in the point cloud.

We interpreted the LIDAR data manually to calculate morphometric variables with Quick Terrain Modeler ×64 software using the following procedure:

a) Point clouds are processed with a hue filter, which allows processing with all of the data on height and on the texture of the target surface;
b) The reflection of the point cloud is controlled using the Voxel Autosize method, in which the points projected on the surface are processed to assign different sizes to each point based on its position in relation to the plane of visualization; the points that are close to the plane appear larger, while the points that are away from the plane appear smaller (Applied Imagery, 2010). Thus, points on the boundary of the profile structure are highlighted, providing a contrast with the second and third points present within the structure. This process facilitates drawing the outline of the dominant and co-dominant trees and, hence, the process of isolating the trees (Fig. 2), and

c) Sample trees are separated from the LIDAR point cloud by making a three-dimensional polygon surrounding the canopy. Initially, a region that includes the area of the canopy projection of the tree of interest and the understory beneath this canopy was cut out of the point cloud. Next, the tree crown and the projection of the stem were isolated by editing the polygons and by successive cuttings.

2.7. Independent variables

A total of 18 independent variables were used, considering stem data, apparent density (AD), basic density (BD) of the wood, the altitude of the site (As) (elevation above mean sea level of the ground at the location of the tree) and morphometric variables of the crown. The variables for the bole were diameter at breast height measured at 1.3 m above the ground or above any buttresses (DBH, in cm) and height of insertion of the crown (Hic in m), also known as the “commercial height.”

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportional-sampling data base stratified by diameter class of the dominant and co-dominant trees with volumes measured for use in fitting allometric models.</td>
</tr>
<tr>
<td>Diameter Class</td>
</tr>
<tr>
<td>I</td>
</tr>
<tr>
<td>II</td>
</tr>
<tr>
<td>III</td>
</tr>
<tr>
<td>IV</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>
The following variables adapted from Burger (1939) were obtained from the LIDAR point cloud to describe the morphometry of the canopy (Fig. 3): length of the crown (Lc in m), length of the branches (Lb in m), mean crown diameter (CD in m), total height of the tree (Ht in m), percentage of canopy (PC = LC/Ht in %), degree of slenderness (Ds = Ht/DBH), index of protuberance (IP = CD/DBH), index of enclosure (IE), form of the crown (Fe = CD/Lc), volume of the crown (Vc in m³) from the solid of rotation that best models the crown, index of living space (ILS = (CD/DBH)²), crown projection area (CPA in m²) representing the area under the canopy, and mantle of the crown (MC in m²) representing the area of the surface enclosing the crown. The DBH of each tree was measured in the field during the forest inventory. The one-year time difference between the forest inventory and the LIDAR flight do not significantly affect the use of LIDAR-derived variables and DBH to develop the biomass estimation models.

Of all the variables evaluated, only the data for Vc, CPA and MC were obtained during the processing of the LIDAR point cloud, verification in the field not being possible. This was due to the difficulty of measuring this information in the forest, as already reported by Chambers et al. (2007) and Ferraz et al. (2016).

2.8. Allometric equations for bole biomass

Initially we evaluated the intensity of the linear relationship between the dependent variables – dry biomass (Bd) and green biomass (Bg) – and the independent variables. This step assists in the initial indication of the most significant variables for model building (Statgraphics, 2006).

Only independent variables with correlation coefficient values less than −0.65 or >0.65 were selected for the routine that checks all possible models (Ryan, 2011). The inclusion of independent variables in allometric models was limited to a maximum of three, which allows models with up to four parameters (βj), including the intercept (β0).

In order to reduce the number of models initially indicated by the selection routine, additional criteria were incorporated into the screening process: the equations could have no multicolinearity and were required to have normally distributed residuals that are independent and homoscedastic. The following statistics were calculated for this additional screening: multicolinearity array, standardized error distribution, Durbin-Watson (DW) test and Hartley F-maximum (SAS Institute, 1990; Statgraphics, 2006).

Equations for estimating biomass were obtained for the models proposed by the selection routine. The statistical significance of each independent variable was examined using the Fisher test (F test). Variables with significance levels <0.15 were removed to simplify the polynomial.

For each allometric equation we performed an analysis of the influence of the independent observations. An observation was only considered to be influential when it produced substantial changes in the calculated statistical values with and without the observation in accord with the following measures of atypical status: elements in the principal diagonal of the H matrix, DFFITS and Cook’s distance (Chatterjee and Hadi, 1986; Figueiredo, 2005; Souza, 1998; Statgraphics, 2006). Selection of the best equation was based on a graphical analysis of the residuals expressed as percentages, standard error expressed in absolute and in percentage terms (Syx and Syx%), the Pressp criterion and the adjusted coefficient of determination (R²aj%) (SAS Institute, 1990; Souza, 1998).

2.9. Identity of models by density group

After selecting the best allometric equation for dry biomass, we performed an identity test of the models. Because only linear models were involved, we used the procedure described by Graybill (1976) to assess the need for fits made either individually or by basic-density group for three classes of basic density (BD): Low (BDlo: BD < 0.5), medium (BDmd: BD ≥ 0.5 < 0.7) and high (BDhi: DB ≥ 0.7) (do Vale et al., 2005; Nogueira et al., 2005). The procedure consists of minimizing the sum of squares. For linear models, the identity test allows the F test to be used to assess the significance of the difference between the total sum of squares of the regressions fitted for each basic-density group considered by itself (full model: Ω) and the sum of squares of the regressions fitted for all three basic-density groups (reduced model: w) (Figueiredo, 2005; SAS Institute, 1990; Statgraphics, 2006). Table 2 presents the analysis of variance for testing the identity of the linear regression models.

Three hypotheses were tested for the wood-density groups. The first group (a) is for the fit of the original database composed of species in all three basic-density classes (low, medium and high); the second group (b) is composed of species with low and medium basic density, and the third and last group (c) is composed of species with medium and high basic density.

3. Results

The sample structure studied was classified as having 17.6% species with low basic density (BD < 0.5), 40.8% moderately hard (0.5 ≤
BD 0.7), and 41.6% hard (BD 0.7). When basic densities of wood samples collected from the base of the trunk (first log) were compared with those from the upper end (last log) using the t-test, a statistical difference was observed (t = 2.15, α = 0.05, p = 0.03). The basic density of the first log (mean ± standard deviation = 0.69 ± 0.03, n = 142) was 7.3% higher than the density of the last log (0.65 ± 0.03, n = 142).

Table 2

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>df</th>
<th>Sum of squares</th>
<th>Mean square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full model</td>
<td>(A × p)</td>
<td>SS Reg.(Ω)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduced model</td>
<td>p</td>
<td>SS Reg.(w)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference for hypothesis testing</td>
<td>(A − 1)p</td>
<td>SS Reg.(Ω)−SS Reg.(w)</td>
<td>SS(difference)/(A − 1)p</td>
<td>MS(difference)/MS(residual)</td>
</tr>
<tr>
<td>Residual</td>
<td>N−(A × p)</td>
<td>SS Total(Ω)−SS Reg.(Ω)</td>
<td>SS(residual)/N−(A × p)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>N</td>
<td>SS Total(Ω)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

df = degrees of freedom; SS = sum of squares; MS = mean square; F = F-test statistic; A = number of density classes; p = number of parameters in the reduced model (w); N = number of observations in the full model (Ω).
The overall mean basic density based on the samples from the first and last logs was 0.67 (n = 142) with a coefficient of variation (CV) of 24.4%.

The apparent density (AD) resulting from the wood sample composition of the first and last log was 0.96 and the CV was 11.1%. Comparing the apparent densities of the first and last logs, there was no statistical difference between the two groups (t = 1.54, α = 0.05, p = 0.12).

The mean AD of the base of the tree was 0.97 ± 0.02 (n = 142) and the mean AD of the last log was 0.95 ± 0.02 (n = 142). The mean moisture content of the first log was 0.34 ± 0.09 (n = 142), and for the last log it was 0.34 ± 0.10 (n = 142).

In the tested equations where the independent variables were restricted to those for canopy morphometry, 14.1% of sampled trees were identified as atypical points. When DBH was included as an independent variable in the model this percentage dropped to 6.3%, and when both DBH and Ht were included it dropped to 4.9%.

The descriptive results for the main independent variables are shown in Table 3.

Evaluation of the linear correlation between the independent and dependent variables provides an important tool for the initial selection of independent variables, especially when considering a large number of possible variables of interest. Table 4 describes the results of linear correlation of eighteen variables of interest, including crown morphometric variables, environmental variables and bole variables, with the dependent variable being either green or dry bole biomass.

Correlation analysis between crown morphometric variables and the main dendrometric characters of the bole (Table 5) helps in understanding the importance that the crown data have for representing bole variables and, consequently, the volume and biomass present in the trunks of the trees.

The results of the selection procedure for all possible models were divided into two categories. The first group (Equations 1 to 5 in Table 6) is for estimation of dry biomass of the bole, while the second group (Equations of 6 to 11 in Table 6) is for green biomass.

Figs. 4 and 5 present the standard error percentage graphs for the best equations considering the inclusion or the exclusion DBH, and Figs. S1 and S2 (Supplementary material) present, real and estimated values. The plots of residuals in Fig. 4 are for estimates of dry biomass (Bd) and in Fig. 5 for green biomass (Bg).

All of the selected equations and the equations originating from the combination of subsets for identity testing of the models had Durbin-Watson values between 1.897 and 2.262 (p = 0.409 and 0.926); since the p-value is > 0.05, there is no indication of serial autocorrelation in the residuals for the selected models. The assumption of homoscedasticity (Hartley F-max) was also confirmed in the different diameter classes of the sample. Kolmogorov-Smirnov tests of the frequency of the residuals for all selected equations do not allow rejection of the hypothesis of a normal distribution with 95% confidence.

The main purpose of the identity test of model is to evaluate whether the allometric equations fit using the datasets for the three wood-density classes (low, medium and high) have similar characteristics, or whether fits should be made individually for each density class or for groupings of two density classes. Table 7 presents the results for the identity test of models for estimating the dry biomass of the bole and Table 8 the results for models that estimate green biomass.

4. Discussion

Allometric equations should seek to reduce possible sources of variation, avoiding generalizations based on models derived from specific data from case studies (Baker et al., 2004; Bohlman and O’Brien, 2006; Crow and Schlaegel, 1988; Ferraz et al., 2016; Scolforo et al., 2008). Because of this, allometric equations generated from a given forest structure are not sufficient for general use in the field of forest management. Models must be fitted to reflect the characteristics of each region and interest.

Because the present study is focused on the bole biomass of trees intended for the timber industry that are cut as part of forest management, sampling is for trees with DBH ≥ 45 cm. The sample composition therefore differs from that in studies intended to estimate total tree biomass and the composition of the whole forest, such as Higuchi et al. (1998), da Silva (2007) and Nogueira et al. (2007). In most studies for estimating the biomass of native forests, a sample is taken that begins with the smaller diameter classes (e.g., DBH ≥ 5 cm), and generally has < 10% of the sampled individuals with DBH ≥ 50 cm. This change in sample composition directly impacts biomass estimates, mainly due to changes in diameter structure, species and density. These results are therefore not necessarily applicable in other situations.

The statistically significant difference in the density of samples from the first and last logs corroborates the results obtained by Scolforo et al. (2004), Nogueira et al. (2005, 2007) and Silveira et al. (2013). Thus, if basic density is included in estimating the biomass of a species group, a compound sample should be collected along the bole because using samples taken from a single position on the bole was found to misrepresent basic density by an average of 7.28%. The same behavior was not observed when using apparent density, which does not differ significantly between the base and the top of the bole. This variable was not chosen by the routine checking all possible models (Ryan, 2011).

Table 3

<table>
<thead>
<tr>
<th>Description</th>
<th>Units</th>
<th>n</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Coefficient of variation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPA</td>
<td>(m²)</td>
<td>142</td>
<td>354.1</td>
<td>235.2</td>
<td>66.4%</td>
<td>106.8</td>
<td>1,713.8</td>
</tr>
<tr>
<td>Lc</td>
<td>(m)</td>
<td>142</td>
<td>16.3</td>
<td>3.9</td>
<td>37.9%</td>
<td>4.1</td>
<td>24.9</td>
</tr>
<tr>
<td>Ub</td>
<td>(m)</td>
<td>142</td>
<td>15.4</td>
<td>4.2</td>
<td>27.5%</td>
<td>6.1</td>
<td>25.0</td>
</tr>
<tr>
<td>DBH</td>
<td>(cm)</td>
<td>142</td>
<td>77.9</td>
<td>23.1</td>
<td>29.7%</td>
<td>112.6</td>
<td>2,001.7</td>
</tr>
<tr>
<td>DC</td>
<td>(m)</td>
<td>142</td>
<td>21.8</td>
<td>6.5</td>
<td>29.8%</td>
<td>12.0</td>
<td>49.1</td>
</tr>
<tr>
<td>Hic</td>
<td>(m)</td>
<td>142</td>
<td>23.7</td>
<td>4.5</td>
<td>19.2%</td>
<td>12.2</td>
<td>36.1</td>
</tr>
<tr>
<td>Ht</td>
<td>(m)</td>
<td>142</td>
<td>39.1</td>
<td>4.9</td>
<td>12.6%</td>
<td>26.4</td>
<td>53.7</td>
</tr>
<tr>
<td>MC</td>
<td>(m³)</td>
<td>142</td>
<td>400.7</td>
<td>290.7</td>
<td>69.1%</td>
<td>112.6</td>
<td>2,001.7</td>
</tr>
<tr>
<td>Vc</td>
<td>(m³)</td>
<td>142</td>
<td>2,495.9</td>
<td>2,449.1</td>
<td>98.1%</td>
<td>235.2</td>
<td>14,998.1</td>
</tr>
<tr>
<td>Vol</td>
<td>(m³)</td>
<td>142</td>
<td>8.3</td>
<td>6.9</td>
<td>82.7%</td>
<td>1.5</td>
<td>55.9</td>
</tr>
</tbody>
</table>

Vc = volume of the crown from the solid of rotation that best models the crown (m³); MC = mantle of the crown (surface area of the solid of rotation in m²); CPA = crown projection area (m²); DC = diameter of the crown (mean in m); DBH = diameter at breast height measured in the field 1.3 m above the ground or above any butteneses (m); Lc = length of the crown (m); Ht = total height of the tree (m); Ub = length of the branches (m); Hic = height of insertion of the crown [height above the ground of the first living branch]; Vol = volume of the bole from rigorous cubic scaling (m³).
Vc = volume of the crown from the solid of rotation that best models the crown (m³); MC = mantle of the crown (surface area of the solid of rotation in m²); CPA = crown projection area (m²); DC = diameter of the crown (mean in m); DBH = diameter at breast height measured in the field 1.3 m above the ground or above any buttresses (m); IE = index of enclosure = DC/HT; Lc = length of the crown (m); Ht = total height of the tree (m); PC = percentage of crown = (Lc/HT) × 100 (%); Lb = length of the branches (m); IP = index of proteuberance = DC/DBH; ILS = Index of living space = (DC/DBH)^2; Hic = height of insertion of the crown [height above the ground of the first living branch]; Fc = form of the crown = DC/Lc; Ds = degree of slenderness = HT/DBH; As = altitude of the site = elevation of the ground above mean sea level as measured with LIDAR; DB = basic density of the wood; AD = apparent density of the wood.

Results of the correlation between the independent variables for crown morphometry and traditionally used bole variables (DBH, Ht and BD) were incorporated into a single model. Even so, only approximately 5% of the observations were influential.

Although trees with broken crowns are easily identified when the LIDAR point cloud is interpreted, the models cannot represent them. These trees were therefore excluded as atypical observations. In a forest management strategy the trees with broken crowns should be prioritized for harvest due to their lower reproductive capacities, their lower yields and because felling them causes the least impact due to their reduced canopies. However, they should be measured in the field, since the models are unable to generate reliable estimates. The CPA, MC and Vc variables were the independent variables that were most influenced by the existence of broken crowns. The largest variation was seen in the volume of the crown (Vc) because the consequences of a partial loss of the crown are strongly reflected in the three-dimensional volume results.

The results of a linear correlation between all independent variables and the modeled independent variable indicate that DBH, MC, CPA, Vc and CD have strong linear correlations with dry biomass (Bd) (Table 4). This information is important for fitting linear models or for assessing the explanatory power of a given type of information. This is especially evident for the four morphometric variables for the crown (MC, Vc, CPA and CD), which have strong explanatory power with correlation-coefficient values above 0.7. The MC and Vc variables obtained exclusively by the processing the LIDAR point cloud explain >50% of the variation in data for estimating biomass stored in the stems of dominant and co-dominant trees. This corroborates the findings of Schneider and Schneider (2008), who emphasize the importance of understanding the horizontal space of the forest, i.e., the individuals with the largest crowns have the greatest capacity for storing biomass and represent the potential of the forest site. A number of authors have developed and studied methods for estimating tree biomass from canopy morphology, eliminating field measurements (Balzotti et al., 2016; Bouvier et al., 2015; Ferraz et al., 2016; Shugart et al., 2015; Zolkos et al., 2013).

Analyzing the correlation between the data on the bole (DBH, Dmax, Dmin, Hc and Ht) and morphometric information on the crown can help better understand the importance of each independent variable derived from the crown (Table 5). The mantle of the crown (MC) and the crown projection area (CPA), in addition to providing good predications of DBH and Dmax, also exhibit strong correlations with the Dmin. The diameter of the thin end of the bole, or Dmin, provides important information relevant to the yield obtained by the forest industry and is also important for the estimating biomass stored in the trunk. Therefore, trees with large crown structures have greater DBH and Dmin values, and, consequently, greater stocks of biomass, corroborating the results obtained by Duncanson et al. (2015) and Balzotti et al. (2016).

The tallest trees were not necessarily those with the largest diameters (DBH, Dmax and Dmin). This shows the low hypsometric relation

<table>
<thead>
<tr>
<th>Order</th>
<th>Independent variable</th>
<th>Sample size</th>
<th>Correlation (r)</th>
<th>Order</th>
<th>Independent variable</th>
<th>Sample size</th>
<th>Correlation (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DBH</td>
<td>142</td>
<td>0.52</td>
<td>1</td>
<td>DBH</td>
<td>142</td>
<td>0.63</td>
</tr>
<tr>
<td>2</td>
<td>MC</td>
<td>142</td>
<td>0.79</td>
<td>2</td>
<td>Vc</td>
<td>142</td>
<td>0.72</td>
</tr>
<tr>
<td>3</td>
<td>CPA</td>
<td>142</td>
<td>0.77</td>
<td>3</td>
<td>MC</td>
<td>142</td>
<td>0.70</td>
</tr>
<tr>
<td>4</td>
<td>Vc</td>
<td>142</td>
<td>0.77</td>
<td>4</td>
<td>CPA</td>
<td>142</td>
<td>0.70</td>
</tr>
<tr>
<td>5</td>
<td>DC</td>
<td>142</td>
<td>0.72</td>
<td>5</td>
<td>DC</td>
<td>142</td>
<td>0.66</td>
</tr>
<tr>
<td>6</td>
<td>IE</td>
<td>142</td>
<td>0.56</td>
<td>6</td>
<td>Ht</td>
<td>142</td>
<td>0.60</td>
</tr>
<tr>
<td>7</td>
<td>Ht</td>
<td>142</td>
<td>0.54</td>
<td>7</td>
<td>Lc</td>
<td>142</td>
<td>0.51</td>
</tr>
<tr>
<td>8</td>
<td>Lc</td>
<td>142</td>
<td>0.51</td>
<td>8</td>
<td>IE</td>
<td>142</td>
<td>0.47</td>
</tr>
<tr>
<td>9</td>
<td>PC</td>
<td>142</td>
<td>0.38</td>
<td>9</td>
<td>Lb</td>
<td>142</td>
<td>0.42</td>
</tr>
<tr>
<td>10</td>
<td>Lb</td>
<td>142</td>
<td>0.36</td>
<td>10</td>
<td>PC</td>
<td>142</td>
<td>0.36</td>
</tr>
<tr>
<td>11</td>
<td>Hc</td>
<td>142</td>
<td>0.26</td>
<td>11</td>
<td>Hc</td>
<td>142</td>
<td>0.28</td>
</tr>
<tr>
<td>12</td>
<td>Fc</td>
<td>142</td>
<td>0.12</td>
<td>12</td>
<td>BD</td>
<td>142</td>
<td>0.14</td>
</tr>
<tr>
<td>13</td>
<td>As</td>
<td>142</td>
<td>0.01</td>
<td>13</td>
<td>Fc</td>
<td>142</td>
<td>0.11</td>
</tr>
<tr>
<td>14</td>
<td>AD</td>
<td>142</td>
<td>-0.08</td>
<td>14</td>
<td>AD</td>
<td>142</td>
<td>0.08</td>
</tr>
<tr>
<td>15</td>
<td>BD</td>
<td>142</td>
<td>-0.12</td>
<td>15</td>
<td>As</td>
<td>142</td>
<td>0.03</td>
</tr>
<tr>
<td>16</td>
<td>IP</td>
<td>142</td>
<td>-0.13</td>
<td>16</td>
<td>IP</td>
<td>142</td>
<td>-0.12</td>
</tr>
<tr>
<td>17</td>
<td>ILS</td>
<td>142</td>
<td>-0.15</td>
<td>17</td>
<td>ILS</td>
<td>142</td>
<td>-0.13</td>
</tr>
<tr>
<td>18</td>
<td>Ds</td>
<td>142</td>
<td>-0.65</td>
<td>18</td>
<td>DA</td>
<td>142</td>
<td>-0.57</td>
</tr>
</tbody>
</table>

Statistics in the analysis; this atypical feature is explained by the existence of trees with broken crowns. The data were best modeled when canopy morphometric variables and traditionally used bole variables (DBH, Ht and BD) were incorporated into a single model. Even so, only approximately 5% of the observations were influential.

Table 5: Results of the correlation between the independent variables used for crown morphometry and the independently employed in dendrometric models, Antimary State Forest, Acre, Brazil.

<table>
<thead>
<tr>
<th>Independent variables traditionally used in dendrometric models</th>
<th>DBH</th>
<th>Dmax</th>
<th>Dmin</th>
<th>Hc</th>
<th>Ht</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ds</td>
<td>-0.6301†</td>
<td>-0.8204†</td>
<td>-0.8011†</td>
<td>0.0798</td>
<td>-0.0167</td>
</tr>
<tr>
<td>ILS</td>
<td>-0.2621</td>
<td>-0.2604</td>
<td>-0.2150</td>
<td>-0.2665</td>
<td>0.0902</td>
</tr>
<tr>
<td>IP</td>
<td>-0.2476</td>
<td>-0.2457</td>
<td>-0.1988</td>
<td>-0.2429</td>
<td>0.1268</td>
</tr>
<tr>
<td>Fc</td>
<td>0.1313</td>
<td>0.1346</td>
<td>0.1956</td>
<td>0.1158</td>
<td>-0.1695</td>
</tr>
<tr>
<td>Lb</td>
<td>0.3299</td>
<td>0.3358</td>
<td>0.2699</td>
<td>-0.3465</td>
<td>0.5634</td>
</tr>
<tr>
<td>PC</td>
<td>0.3797</td>
<td>0.3741</td>
<td>0.3490</td>
<td>-0.2317</td>
<td>0.3797</td>
</tr>
<tr>
<td>Lc</td>
<td>0.4800</td>
<td>0.4717</td>
<td>0.4257</td>
<td>0.0886</td>
<td>0.6457</td>
</tr>
<tr>
<td>IE</td>
<td>0.5826</td>
<td>0.5843</td>
<td>0.6046</td>
<td>-0.2262</td>
<td>0.1622</td>
</tr>
<tr>
<td>DC</td>
<td>0.7053*</td>
<td>0.7027*</td>
<td>0.6899</td>
<td>0.0410</td>
<td>0.5266</td>
</tr>
<tr>
<td>Vc</td>
<td>0.7238*</td>
<td>0.7143*</td>
<td>0.6874</td>
<td>0.0540</td>
<td>0.5258</td>
</tr>
<tr>
<td>CPA</td>
<td>0.7307*</td>
<td>0.7265*</td>
<td>0.7170*</td>
<td>0.0323</td>
<td>0.4681</td>
</tr>
<tr>
<td>MC</td>
<td>0.7552*</td>
<td>0.7531*</td>
<td>0.7434*</td>
<td>0.0461</td>
<td>0.4398</td>
</tr>
</tbody>
</table>

Vc = volume of the crown from the solid of rotation that best models the crown (m³); MC = mantle of the crown (surface area of the solid of rotation in m²); CPA = crown projection area (m²); DC = diameter of the crown (mean in m); DBH = diameter at breast height measured in the field 1.3 m above the ground or above any buttresses (m); IE = index of enclosure = DC/HT; Lc = length of the crown (m); Ht = total height of the tree (m); PC = percentage of crown = (Lc/HT) × 100 (%); Lb = length of the branches (m); IP = index of proteuberance = DC/DBH; ILS = Index of living space = (DC/DBH)^2; Hic = height of insertion of the crown [height above the ground of the first living branch]; Fc = form of the crown = DC/Lc; Ds = degree of slenderness = HT/DBH; Dmax = diameter at the thick end of the bole (m); Dmin = diameter at the thin end of the bole (m); * = correlation that explains information with values equal to or >50%.
of the population we studied. Allometric equations based on simple input variables such as DBH may not adequately represent the variations in Amazonian forest. This is reinforced by the findings of Crow and Schlaegel (1988), Santos (1996), Fearnside (1997, 2007), Clark and Clark (2006) and Duncanson et al. (2015). This increases the importance of including the independent variable Ht (total height); this variable was present in 66% of the equations selected for estimating green and dry biomass. The total height of the tree (Ht) is discarded as a variable in regression models due to the difficulty of measurement in the field. According to Overman et al. (1994), Brown et al. (1995), Higuchi et al. (1998) and da Silva (2007), inclusion of tree height in allometric models contributes little to improving results and is subject to accuracy due to the difficulty of obtaining measurements. However, in large pan-tropical data sets, height data can significantly improve allometric estimates (Asner and MASCARO, 2014; BANIT et al., 2012; CHAVE et al., 2014; DUNCANSON et al., 2015; FLEDPAAUS et al., 2011, 2012).

When airborne LIDAR is employed, the total height (Ht) variable gains importance and is more practical to collect because the data of tree height and of precision for total height (Ht) are technologically overcome. The height variable should be incorporated into allometric models whenever possible because it is both logically related to biomass and is the variable that is most easily obtained by processing the LIDAR point cloud. However, before obtaining the heights of the trees from the LIDAR point cloud, one should normalize the heights and delete the set of floating points above the forest canopy. These points represent flocks of birds that are detected during profiling with airborne LIDAR. Neglecting this step in the processing can result in overestimation of the total height (Ht).

Another important aspect that shows the importance of including the height variable was described by Brown (2002). This is that equations that only consider diameter and that represent large areas of forest with variations in forest type tend not to adequately reflect the true biomass of trees at any given location. Moreover, the height of the dominant trees has always been seen as reflecting the quality of a forest ecosystem. The height variable is linked to the ecosystem’s capacity to stock biomass (Baltzetti et al., 2016; CASSOL, 1982; DUNCANSON et al., 2015; FERRAZ et al., 2016; FIGUEIREDO, 2005; SCOLFORO, 1998; STANSFIELD et al., 1992).

Eight equations were selected to estimate the biomass of the bole (Bd) (Table 6); the best results were achieved by Equations 1, 3, 4, and 2. In the four best models the variables DBH and basic density were always included. Equation 1 resulted in the best statistical precision and the lowest validation value (Press ν). In Equations 2 and 4, crown morphometric variables were incorporated: crown projection area (CPA), volume of the crown (Vc) and mantle of the crown (MC), i.e., the three independent variables of the crown with the highest correlations with bole biomass. Equation 3 represents the models traditionally used to estimate dry biomass. These models included the variables DBH, BD and Ht, and the statistical results are similar to those of equations that include canopy variables. Equations fit with better statistical results are possible; however, these apply to simpler forest structures, as in temperate forests (Cao et al., 2016; Gregoire et al., 2016).

Equations with simple inputs for Bd were not selected by the statistical routine. This was because they failed to adequately represent the largest variations in the dendrometric characters of the sampled trees. Equations 5, 6, 7 and 8 can be used when one chooses only variables obtained with LIDAR, whether or not they are associated with wood-density variables. However, the precision of the results is lower than that of equations that include DBH. In the equations using only LIDAR variables (Equations 5, 6, 7 and 8; Fig. 4), there was a tendency to underestimate the bole biomass for trees with diameters between 45 and 60 cm. The models’ difficulty in representing this diameter range is due to the fact that it contains the greatest number of species of commercial interest in forest management. For diameter intervals representing larger trees, the number of species decreases and the sources of variation therefore also decrease.

The distributions of the residuals (expressed as percentages) for the first four equations are all similar (Fig. 4). All of these equations include DBH and BD as variables. When these four equations are graphed we observe substantial heterogeneity in the residuals in the smaller diameter classes, but this did not represent an impediment to meeting the assumptions for regression because the distribution of errors showed no trends and converged to zero.

Equations 9 and 13 (for green biomass) had the highest accuracy and the lowest validation values (Press ν). The distributions of percentage residuals for the two equations (Fig. 4) show similarities, with little scattering of percentage errors and a lack of tendencies to over- or underestimate biomass. For estimates of green biomass, the traditionally used linear models with directly measured explanatory variables (DBH, Ht and AD) have good predictive capacity, and incorporation of morphometric variables of the crown contributes little to improving accuracy. It was even possible to fit models with a single input (DBH – Equation 14), corroborating the results of Higuchi et al. (1998).
However, when opting to use only DBH as an explanatory variable, it is clear that the dispersion of residuals is greater in all diameter ranges. The green biomass Equations 15, 16, 17 and 18, which only use variables obtained from processing the LIDAR point cloud, had lower statistical accuracy than models that include DBH. This was also the case for the equations for dry biomass, where the inclusion of DBH as an explanatory variable decreased the dispersion of residuals for the smaller diameter classes.

Identity testing of models seeks to assess whether it is better to fit models that consider a single data group with species in all three categories of basic wood density (low, medium and high), or whether the data should be divided into subgroups (Graybill, 1976). Equations 1, 2 and 4, which estimate dry biomass using DBH as the independent variable, show significant differences as indicated by the F test (Table 7). Better results are obtained when models are fitted separately for each density group. The only equation that considers DBH as an independent variable and that allows the formation of a single database with all densities is Equation 3. This equation is traditionally used, and considers as independent variables DBH, AD and Ht.

Equations 6, 7 and 8 only use variables that are obtained with airborne LIDAR, especially CPA, Ht, FC and As. These variables can be fitted to the data for datasets for the different wood basic density groups. This
Fig. 5. Distribution of residuals (expressed as percentages) for the best equations for estimating green biomass, with DBH and without DBH.
allows for greater ease in fitting and in using the equations for estimating the biomass of dominant and co-dominant trees in forest-management areas.

The equations for dry biomass that have been developed with variables representing the canopy (CPA and FC), together with HT or altitude of the site (As), were better able to represent the set of information on the sampled forest population, regardless of the density of the wood. If one chooses not to use models that consider the canopy variables, models should either be fitted separately for each group of basic wood density or they should include the basic density (BD) as an independent variable so that species of all densities can be fitted in a single model.

The great advantage of using a model without independent variables that are measured in the field (DBH and density), such as Equations 5, 7, and 8, is that the equation can make estimates for dominant and co-dominant trees as soon as one has a model that has been fitted and validated for a given location. This optimizes the work in annual operational planning in forest management.

## 5. Conclusion

The study shows the viability of measuring crown morphological variables and total height in Amazonian forest based on airborne LiDAR data alone. These variables are sufficient to obtain useful estimates of forest biomass, and especially the biomasses of large commercially valuable trees that are of interest for forest management. On-the-ground measurements of traditional variables such as diameter at breast height and wood density are needed to calibrate the remotely sensed variables to biomass in each general location, but once this is done, remotely sensed data can be gathered and the estimates extended.
to much wider areas than would be practical to assess with traditional forest inventories.

The variables with the highest linear correlations with the dry biomass of the bole are: diameter at breast height (DBH), basic wood density (BD), mantle of the crown (MC), crown projection area (CPA), volume of the crown (VC) and crown diameter (CD). In estimating dry biomass, the best results for accuracy and distribution of residuals are always achieved when both crown morphometric variables and traditional variables for the bole (DBH, total height and basic density) are used. When opting to include wood density as an explanatory variable, one must determine basic density from samples collected along the length of the bole. Trees with broken crowns (about 14%) can be identified as outliers by analysis of influence and should not be modeled by equations that use crown morphology variables. The mantle of the crown (MC) and the crown projection area (CPA) are strongly correlated with the length of the bole (DBm) and are important independent variables for forest yield and for the biomass stored in the trunk.

The equations for estimating dry biomass (Bd) that only include crown morphometry variables and/or the altitude of the site (As) provide lower accuracy as compared to models with DBH. However, they have good performance in estimating the biomass of dominant and co-dominant trees, regardless of the basic wood density group, and can be used based only on data obtained by processing the point cloud from airborne LiDAR.

The application of LiDAR on a wide scale in Amazonian forest can potentially make significant contributions to improving estimates of forest biomass (thus reducing uncertainty in estimates of greenhouse-gas emissions from deforestation). These estimates can also contribute to detecting and monitoring forest degradation, and to assuring the sustainability of forest management.

Acknowledgments

The authors’ research is supported by: Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), Conselho Nacional do Desenvolvimento Científico e Tecnológico (CNPq: Proc. 305880/2007-1: 304020/2010-9; 573810/2008-7; 755853/2008-5), Brazilian Research Network on Climate Change (RedeClima) and Instituto Nacional de Pesquisas da Amazônia (INPA: PRJ1303). Reviewer comments were very helpful.

Appendix A. Supplementary data

Supplementary data for this article can be found online at doi:10.1016/j.rse.2016.10.026.

References


Schneider, P.R., Schneider, P.R.P., 2008. Introdução aos Modelos de Regressão Linear e Não-Linear. Embrapa-SPI/Embrapa-SEA, Brasília, DF, Brazil (505 pp).


SUPPLEMENTARY MATERIAL

LIDAR-based estimation of bole biomass for precision management of an Amazonian forest: Comparisons of ground-based and remotely sensed estimates

Evandro Orfanó Figueiredo¹, Marcus Vinicio Neves d’Oliveira², Evaldo Muñoz Braz³, Daniel de Almeida Papa⁴, Philip Martin Fearnside⁵*

¹ Embrapa Acre, Rodovia BR-364, km 14, CEP 69900-056, Rio Branco, Acre, Brazil, evandro.figueiredo@embrapa.br
² Embrapa Acre, Rodovia BR-364, km 14, CEP 69900-056, Rio Branco, Acre, Brazil, marcus.oliveira@embrapa.br
³ Embrapa Florestas, Estrada da Ribeira Km 111, CP 319, CEP 83411-000, Colombo, Paraná, Brazil, evaldo.braz@embrapa.br
⁴ Embrapa Acre, Rodovia BR-364, km 14, CEP 69900-056, Rio Branco, Acre, Brazil, daniel.papa@embrapa.br
⁵ Instituto Nacional de Pesquisas da Amazônia (INPA), Av. André Araújo, 2936 - Petrópolis, CEP: 69067-375, Manaus, Amazonas, Brazil, pmfearn@inpa.gov.br

*Corresponding author.

Figure S1 - Distribution of real and estimated values for the best equations for estimating dry biomass, with DBH and without DBH.

Figure S2 - Distribution of real and estimated values for the best equations for estimating green biomass, with DBH and without DBH.
Figure S1 - Distribution of real and estimated values for the best equations for estimating dry biomass, with DBH and without DBH.
Figure S2 - Distribution of real and estimated values for the best equations for estimating green biomass, with DBH and without DBH.