Remote Sensing of Environment 124 (2012) 479–491

Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Estimating forest biomass and identifying low-intensity logging areas using airborne scanning lidar in Antimary State Forest, Acre State, Western Brazilian Amazon

Marcus V.N. d'Oliveira [a,1], Stephen E. Reutebuch [b,*], Robert J. McGaughey [b], Hans-Erik Andersen [b]

[a] EMBRAPA-CPAF-ACRE, Caixa Postal 392, CEP 69900-180, Rio Branco, Brazil
[b] USDA Forest Service, Pacific Northwest Research Station, PO Box 352100, Seattle, WA, USA

ARTICLE INFO

Article history:
Received 14 March 2012
Received in revised form 19 May 2012
Accepted 19 May 2012
Available online xxxx

Keywords:
Forest biomass
Airborne laser scanning
Selective logging
Tropical forest monitoring
Lidar
Amazon forest monitoring

ABSTRACT

The objectives of this study were to estimate above ground forest biomass and identify areas disturbed by selective logging in a 1000 ha Brazilian tropical forest in the Antimary State Forest (FEA) using airborne lidar data. The study area consisted of three management units, two of which were unlogged, while the third unit was selectively logged at a low intensity (approximately 10–15 m^3 ha^{-1} or 5–8% of total volume). A systematic random sample of fifty 0.25-ha ground plots were measured and used to construct lidar-based regression models for above ground biomass (AGB). A lidar model-assisted approach was used to estimate AGB for the logged and unlogged units (using both synthetic and model-assisted estimators). Two lidar explanatory variables, computed at a spatial resolution of 50 m × 50 m, were used in these predictions: 1) the first quartile height of all above ground returns (P25); and, 2) variance of the height above ground of all returns (VAR). The model-assisted AGB estimator (total 231,589 Mg ± 5,477 SE; mean 231.6 Mg ha^{-1} ± 5.5 SE; ± 2.4%) was more precise than plot-only simple random sample estimator (total 230,872 Mg ± 10,477 SE; mean 230.9 Mg ha^{-1} ± 10.5 SE; ± 4.5%). The total and mean AGB estimates obtained using the synthetic estimator (total 231,694 Mg; mean 231.7 Mg ha^{-1}) were nearly equal those obtained using the model-assisted estimator. In a second component of the analysis lidar metrics were also computed at 1 m × 1 m resolution to identify areas impacted by logging activities within the selectively harvested management unit. A high-resolution canopy relative density model (RDM) was used in GIS to identify and delineate roads, skidtrails, landings and harvested tree gaps. The area impacted by selective logging determined from the RDM was 58.4 ha or 15.4% of the total management unit. Using these two spatial resolutions of lidar analyses it was possible to identify differences in AGB in selectively logged areas that had relatively high levels of residual overstory canopy cover. The mean AGB obtained from the synthetic estimator was significantly lower in impacted areas than in undisturbed areas of the selectively logged management unit ($p = 0.01$).

Published by Elsevier Inc.

1. Introduction

Information describing forest structure and extent, along with accurate topography, are necessary for development and execution of forest management plans (inventories and zoning) and for monitoring native forests in the Brazilian Amazon. High quality information is needed to improve estimates of biomass and carbon stocks associated with reducing emissions from deforestation and forest degradation (REDD). Extensive, high quality forest structure information is very difficult and expensive to obtain in the natural forests of Acre State, Brazil. Most available databases are old, very limited in area sampled, and do not have the necessary ground resolution for robust forest operation planning and monitoring. Collection of sufficient field information is limited by the need to cover large, remote planning areas with difficult access and adverse seasonal factors (e.g., flooding during the rainy season). These conditions result in high field data collection costs, forcing compromises in the measurements collected or the number of locations sampled. For these reasons a study of airborne laser scanning or lidar (light detection and ranging) technology was undertaken in a managed natural forest in Acre State.

Various lidar systems have been used in tropical, boreal, and temperate forest studies since the 1980s. Lidar is a leading method for generating highly accurate bare earth digital elevation models (DEM) over heavily forested areas (Reutebuch et al., 2003). Previous studies undertaken in tropical (e.g., Asner et al., 2010, 2011; Drake et al., 2002a, 2002b, 2003; Dubayah et al., 2010; Kennaway et al., 2008) and temperate and boreal forests (e.g., Beets et al., 2011b; Lefsky et al., 2002; Magnussen et al., 2010; Næsset, 1997; Næsset et al., 2004; Nelson et al., 1988) have demonstrated strong relationships between lidar canopy metrics and forest structure and biomass.

Although the use of airborne lidar in forest management is not new in temperate and boreal forests (e.g., Hyyppä et al., 2008;

* Corresponding author. Tel.: +1 206 543 4710.
E-mail addresses: mvno@cpafac.embrapa.br (M.V.N. d'Oliveira), sreutebuch@fs.fed.us (S.E. Reutebuch), bmcgaughey@fs.fed.us (R.J. McGaughey), handersen@fs.fed.us (H.-E. Andersen).
[1] Tel.: +55 68 32123232

0034-4257/$ - see front matter Published by Elsevier Inc.
doi:10.1016/j.rse.2012.05.014

Reutebuch et al., 2005; Rombouts et al., 2010), there is little reported use of lidar for forest planning and monitoring in natural tropical forests. The objective of this study was to test lidar-augmented approaches for developing consistent information describing a managed natural forest in the western Amazon basin. Our goals were to study the relationship between lidar metrics and above ground biomass (AGB) in forest areas with high levels of canopy closure and to use lidar data to identify harvest areas and impacts (e.g., roads, skid trails, landings, and associated reduction in biomass) related to logging operations in selectively logged areas.

2. Methods

2.1. Study site

The FEA State Forest (FEA) is located between Rio Branco and Sena Madureira in Acre State, Western Brazilian Amazon (68° 01′ to 68° 23′ W; 9° 13′ to 9° 31′ S). FEA covers an area of 76,832 ha. The area is inhabited by approximately 380 individuals from 109 families who make their livings through extractivism (rubber tapping and Brazil-nut collection) and shift cultivation (Fig. 1).

The climate is classified as Awi (Köppen) with an annual precipitation of around 2000 mm and an average temperature of 25 °C. Wet and dry seasons can be recognized. The dry season occurs between the months of June and September. This season is used to prepare the land by slashing and burning for crops and for all operations related to forest management and logging. The rainy season lasts from October to May.

In FEA there are three types of forest: dense tropical forests with uniform canopy and emergent trees, open tropical forests with frequent occurrence of lianas and palm trees, and open forests called Tabocal that are dominated by a bamboo species locally known as Tabocas (*Guadua* sp.). The area has gentle topography with a maximum elevation range of around 300 m. The predominant soils are dystrophic yellow latosols with high clay content (Funtac, 1990).

FEA is administered by the Acre State Government through a forest management plan for sustainable timber production. Forest management in FEA started in 1985 with the establishment of a large research program coordinated by Funtac (Acre State Technological Foundation). In the designated forest management areas, regular timber extraction has been ongoing since 1999 and recently a forest concession system was adopted to regulate the execution of forest operations by logging companies. The FEA inhabitants receive social and financial benefits from the timber concessions in the forest management areas.

For this study three forest management compartments were selected. The compartments were named Modeflora-unlogged (1), Conventional-unlogged (2), and Conventional-logged (3) according to the forest operations planning methods used in each compartment and their logged or unlogged status at the time of the lidar acquisition. Selective logging in the Modeflora compartment will use a planning methodology suggested by Figueiredo et al. (2007). In this planning method, permanent protection zones in which no harvesting is allowed are established around riparian areas and for areas with slopes greater than 20%. Individual trees (DBH ≥ 35 cm) that will be harvested are identified during the forest inventory process. The locations of the protection zones and harvest trees are collected using global positioning system (GPS) receivers. The protection zone maps and GPS tree coordinates are then used in planning and execution of harvest operations. In the Conventional compartments selective logging is planned and conducted with individual harvest tree locations measured from parallel inventory lines laid out at 50 m intervals throughout the compartment; permanent protection zones are identified using existing maps. The target volume to be extracted from the Conventional management compartments is approximately 10–15 m^3 ha^{-1} (approximately 12–18 Mg ha^{-1} AGB). Logging in the Modeflora compartment had not yet commenced at the time of the lidar acquisition. Only a small portion of the southeast corner of the Conventional-unlogged compartment was harvested at the time of the lidar acquisition. For the purposes of this study, the Modeflora and Conventional-unlogged compartments were pooled into a single unlogged unit (based on their predominately unlogged condition) for comparison with the logged unit (Conventional-logged compartment).

2.2. Field measurements and plot summaries

A forest inventory was conducted in the area covered by the lidar flights in May 2010. The inventory used a systematic random sample with plots that were nominally 50 m × 50 m in size, evenly distributed along ten lines with a total of 50 sample plots and a total sampled area of 12.5 ha or 1.25% of the total study area (Fig. 2). The plot

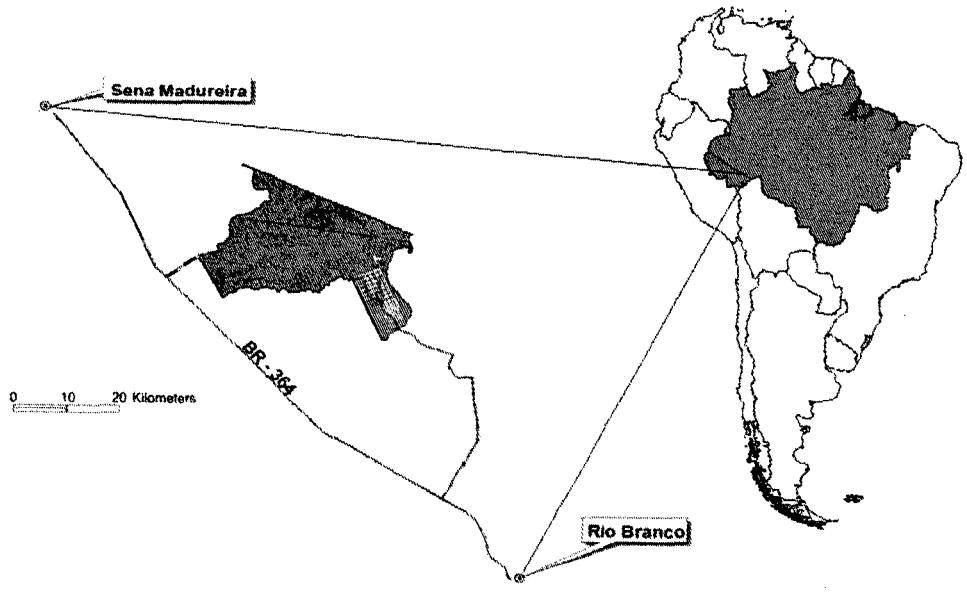

Fig. 1. Antimary State Forest location in Acre State Western Brazilian Amazon. The yellow grid areas represent the compartments of the FEA Forest Management project. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

boundaries were established using a handheld compass for azimuths and a measuring tape for length.

All plants greater than 10 cm diameter at breast height (DBH) were labeled, measured and identified. The natural regeneration (DBH > 5 cm and < 10 cm) was sampled in 10 m × 10 m sub-plots located at the centre of each plot. The species were identified by Embrapa Acre parabotanists (technicians with extensive experience identifying tree species) using vernacular names.

For each tree, oven-dry AGB was estimated with Eq. (1) which was developed for a similar forest in the southern Amazon (Nogueira et al., 2008). In addition to AGB, basal area (BA) was computed using the traditional method (Eq. 2) and stem bole volume (VOL) was estimated using Eq. (3) (Funtac, 1990). Ground measurement data were summarized to the plot level for further analysis (Table 1).

$$AGB = \exp(-1.716 + 2.413^{*}\mathrm{Ln}(DBH))/1000 \quad (1)$$

$$BA = \Sigma\left(\pi^{*}(DBH/200)^{2}\right) \quad (2)$$

$$VOL = 0.000308^{*}(DBH)^{2.1988} \quad (3)$$

Plot locations (corners) were mapped using survey-grade, dual frequency (L1 and L2), dual-constellation (GPS and GLONASS) global navigation satellite system (GNSS) receivers in two separate field campaigns. One-second epoch GNSS data were collected for 10–15 minutes at each plot corner. The GNSS receivers used in this study (Javad Triumph-1) are capable of highly accurate positions in open areas, but like all GNSS receivers, have unknown and highly variable accuracy in heavily forested areas (Clarkin, 2007).

For the first GNSS field survey campaign rover receiver data were post-processed using the Rio Branco base station, located at the Acre Federal University, 90 km from the study site. Unfortunately, the Rio Branco base station did not track the Russian GLONASS satellite constellation. Therefore, the rover GLONASS data could not be included in the post-processing. To determine if inclusion of additional GLONASS satellites would improve the post-processed positions, we conducted a second GNSS field campaign in which we established a temporary base station that tracked both GPS and GLONASS satellites. We resurveyed a subset of the plot corners that had the poorest calculated positional precision from the first GPS-only campaign.

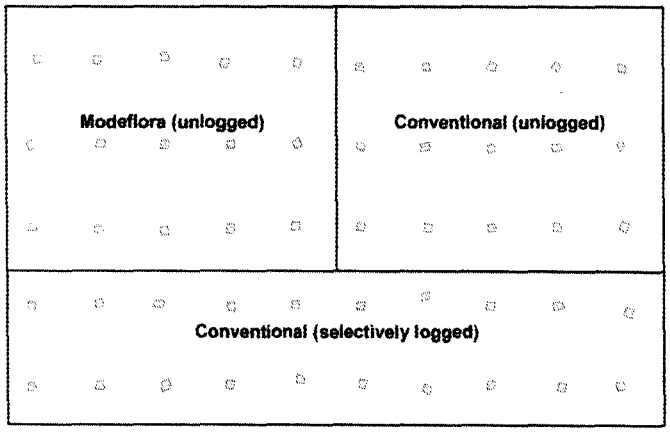

Fig. 2. Forest management compartments and 50 ground plots covered by the lidar flight in the FEA State Forest. Each ground plot is approximately 2500 sq. m. in area (nominally 50 m × 50 m). The lower compartment was selectively logged a month prior to the lidar mission.

Table 1
Number of plots (n), minimum, maximum, mean and standard deviation of basal area ($m^2\ ha^{-1}$), tree bole volume ($m^3\ ha^{-1}$), above ground biomass (Mg ha^{-1}) and means comparison among the management units and the selectively logged and unlogged units.

	Min	Max	Mean	St. dev.
Modeflora, unlogged ($n = 15$)				
BA	13.7	29.8	21.4	5.6
VOL	113.7	260.4	185.4	53.1
AGB	128.9	312.1	219.2	67.1
Conventional, unlogged ($n = 15$)				
BA	16.5	32.5	23.9	4.4
VOL	134.8	353.2	209.6	53.4
AGB	152.8	493.6	252.0	81.4
Conventional, logged ($n = 20$)				
BA	11.4	35.9	22.2	5.3
VOL	88.4	332.8	189.9	55.2
AGB	96.9	414.6	223.8	75.0
Unlogged units ($n = 30$)				
BA	13.7	32.5	22.7	5.1
VOL	113.7	353.2	197.5	53.8
AGB	128.9	493.6	235.6	75.1
All Units ($n = 50$)				
BA	11.4	35.9	22.5	5.1
VOL	88.4	353.2	194.5	53.9
AGB	96.9	493.6	230.9	74.6

ANOVA of means between management units

		Significance (p value)
BA	BA	n.s. (0.39)
VOL	VOL	n.s. (0.43)
AGB	AGB	n.s. (0.42)

ANOVA of means between logged and unlogged units

		Significance (p value)
BA	BA	n.s. (0.73)
VOL	VOL	n.s. (0.63)
AGB	AGB	n.s. (0.59)

To verify whether the position and dimensions of the disturbed areas (roads, intersections, skid trails, landings and logged tree gaps) identified by lidar analysis were consistent with disturbed areas observed in the field, 45 additional GNSS points were collected and post-processed along roads and skid trails and on landings and dimensions of landings and harvest openings were measured with a tape in the field.

Table 2
FEA lidar data acquisition and product specifications.

Specification	Description
Lidar sensor	Optech ALTM 3100EA
Flying altitude	500 m above ground
Beam divergence	0.25 mrad (1/e)
Scan angle	±5 degrees off nadir
Scan rate	70 Hz
Pulse rate	50 kHz
Swath sidelap	60%
Approximate pulse density	25 m^{-2}
Datum	SIRGAS 2000
Projection	UTM, Zone 19S
Lidar raw point cloud format	LAS format with classified ground points identified
Gridded bare earth model	1 m resolution

2.3. Lidar data

High density (25 pulses m^{-2}) discrete return lidar data were collected 29 May–3 June 2010 using an Optech ALTM3100EA scanner mounted in a Sêneca II-EMBRAER 810 C aircraft, flying at 500 m above ground. Table 2 summarizes the acquisition specifications. The total area covered by the lidar flight was 1000 ha. When the data were acquired, only the "Conventional-logged" management compartment (380 ha) and a small portion of the southeast corner of the "Conventional-unlogged" compartment had been logged (Fig. 2).

3. Data processing and analyses

3.1. Lidar processing

The FUSION lidar processing package (McGaughey, 2010) was used for processing the lidar all-returns data (first, intermediates, and last returns per pulse). Lidar returns that occurred within each of the 50 ground plot polygons were extracted from the acquisition dataset to create an all-returns point cloud file for each plot. The ground surface elevation (interpolated from the lidar bare earth digital elevation model) was then subtracted from each return to remove topographic variation within the plot. Descriptive statistics of the lidar point cloud vertical structure, using all returns above 1 m, were computed for each plot (Table 3). The 1 m minimum height above ground was used to reduce noise within the near-ground point cloud caused by low vegetation and imperfections in the ground point filtering. A canopy overstory threshold height of 2 m was used to compute lidar canopy cover metrics. Plot-level lidar metrics were merged with the summarized field plot data for regression modeling with the *R* statistical package (R Development Core Team, 2011).

A variety of raster data products were produced from the entire extent of the lidar data. Basic products were developed by the lidar data provider and provided as contract deliverables. Raster layers of forest canopy metrics were created using FUSION. The following layers were produced:

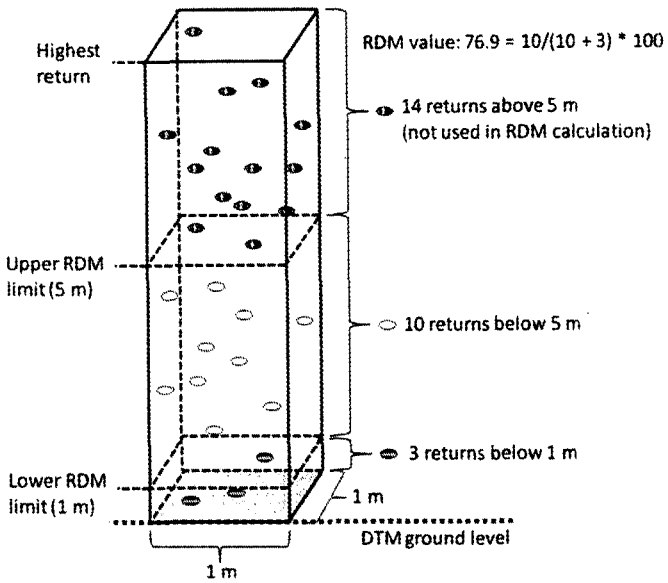

Fig. 3. Illustration of RDM calculation method. In the above example, the RDM value of 76.9 would indicate intact understory with little or no disturbance of vegetation in the 1–5 m above the ground stratum. If there were no returns in the 1–5 m stratum and 3 returns in the 0–1 m stratum, the RDM value would be 0, indicating that the understory had been crushed or removed by harvesting activities. If there were no returns in the 0–1 m stratum and the 1–5 m stratum the RDM value would be undefined (assigned a "no data" value), indicating that the cell had very dense, intact overstory not impacted by harvesting.

- Digital elevation model (DEM) with 1 m resolution: the lidar vendor created the DEM model using the Terrasolid commercial software package (www.terrasolid.fi).
- Canopy surface model (CSM) with 1 m resolution: the CSM was created using the highest return in each cell.
- Canopy height model (CHM) with 1 m resolution: the CHM was created by subtracting the DEM from the CSM.
- Relative density models (RDM) with 1 m resolution: The FUSION "Cover" algorithm was used to create raster layers of a relative percentage of lidar returns within a user-specified above ground height stratum. For each raster, the percentage was computed by dividing the number of all returns in a user-specified height stratum by the sum of returns within and below the height stratum (Fig. 3). The RDMs were used as an indicator of vegetation density in the stratum.
- Reference image with 1 m resolution: an image color was created by color coding lidar first return above ground heights to serve as background image for data exploration and display. The image resolution is sufficient to clearly recognize overstory tree crowns, facilitating visual interpretation of the height of the canopy.
- Lidar raster layers of forest canopy structure with 50 m resolution: the same forest structure metrics (Table 3) that were computed from each ground plot point cloud were also computed at a raster cell resolution equal to the nominal ground plot size over the entire study area.

Table 3
Summary of the lidar forest structure variables derived from the lidar point cloud for each ground plot and for each 50 m by 50 m grid cell in the study area. The height above ground of each return was computed by subtracting the ground surface elevation below the lidar return from the lidar return elevation prior to computing these metrics.

Minimum height above ground
Maximum height above ground
Mean height above ground
Quadratic mean height above ground
Median height above ground
Mode height above ground
Standard deviation of height above ground
Variance of height above ground
Coefficient of variation of height above ground
Interquartile distance of height above ground
Skewness of height above ground
Height kurtosis of height above ground
AAD (average absolute deviation from the mean height) of height above ground
Height L-moments (L1, L2, L3, L4)
Height L-moment skewness
Height L-moment kurtosis
Percentile height values (1st, 5th, 10th, 20th, 25th, 30th, 40th, 50th, 60th, 70th, 75th, 80th, 90th, 95th, 99th percentiles) of height above ground
Percentage of first returns above a specified height (canopy cover estimate)
Percentage of first returns above the mean height
Percentage of first returns above the mode height
Percentage of all returns above a specified height (alternate canopy cover estimate)
Percentage of all returns above the mean height
Percentage of all returns above the mode height

3.2. Regression modeling of above ground biomass

Multiple linear regression techniques were used to develop relationships between plot-level lidar metrics (Table 3) and field-measured AGB, VOL, and BA. Lidar predictor variables were selected using the best subsets approach (R Package 'leaps', Lumley, 2009) in the *R* statistical package (R Development Core Team, 2011). The variance inflation factor (VIF) statistic was used to eliminate highly collinear predictor variables (Fox & Monette, 1992). If VIF exceeded 5.0 for a candidate

predictor variable, it was dropped from the regression model. The Box-Cox method in the R MASS package (Venables & Ripley, 2002) was used to explore possible power transformations of AGB and VOL to reduce non-constant variance and satisfy the assumption of constant variance implicit in linear regression techniques. The square-root transformation of AGB and VOL was used to reduce non-constant variance. The method outlined by Miller (1984) was used to correct for bias introduced during the back-transformation of the final lidar regression models into their original units.

3.3. GIS processing

3.3.1. Mapping of predicted above ground biomass, volume, and basal area

The lidar regression models (with back-transformed bias correction) and the FUSION raster layers of the lidar predictor variables were used in ArcGIS Spatial Analyst Raster Calculator (ESRI, 2011) to map forest structure (AGB, VOL, and BA) across the study area at 50 m resolution (Fig. 4).

3.3.2. Mapping of areas impacted by harvesting

As shown in Fig. 5, areas impacted by harvesting were not readily evident in the 1-m resolution lidar CHM; therefore, the RDMs were displayed in GIS to identify and quantify the areas impacted by logging operations in the managed areas. Several RDMs were examined to determine which height stratum limits would best identify impacted areas. By visual inspection, the RDM that included lidar returns above 1 m and below 5 m was chosen to identify the areas with impacts associated with the network of roads, skid trails and landings, and to locate and characterize canopy gaps created when trees were removed. Each cell in the RDM raster layer contains the percentage of returns below the upper limit of the height stratum that fall between the stratum's upper and lower limit (Fig. 3). In GIS a gray scale representation of the RDM was visually interpreted (Fig. 5). Dark cells correspond to areas with a low percentage of returns in the RDM height stratum. Light cells correspond to areas with a high percentage of returns in the stratum. Cells where no returns occurred below the upper height stratum limit (no lidar data) were assigned a light color. These "no data" cells occurred where the canopy cover above the RDM upper height limit was very dense, effectively blocking sufficient lidar energy from reflecting off lower canopy elements or the ground. By visual inspection of the three-dimensional structure of the lidar point cloud, it was determined that these "no data" areas were intact forest canopy that had not been thinned by harvest activities.

By examining the RDM for the 1 m to 5 m height stratum, the areas impacted by forest operations (roads, landings, skidder trails and canopy gaps due to tree removal) were visually distinguished from the undisturbed areas. The impacted areas were manually digitized in GIS. A 6-m buffer was added to the digitized centerlines of main roads and a 4-m buffer to skidtrails. A 20-m buffer was added to the digitized center point of landings and a 25-m buffer to harvested tree gaps to account for typical widths of these features (Fig. 6). The polygons of buffered impacted areas were converted to a 5-m resolution raster. This impact raster was intersected with the lidar-predicted AGB raster (50 m resolution). Those cells in the AGB raster that contained impacted cells were classified as impacted by harvesting (Fig. 6). A similar visual assessment of the RDM was used to assign a classification of either "impacted" or "non-impacted" to each ground plot. GNSS points collected in the field in tree gaps and along roads and skidtrails were overlaid on the RDM to determine if these field verification points were within the impacted area.

3.4. Lidar-based AGB estimation procedure

Both lidar model-assisted and synthetic estimation approaches were used to estimate total AGB and mean AGB within the management units. These approaches utilize correlated information collected on a larger number of sampling units (N), and at a lower cost, (e.g. lidar metrics for each 50-m grid cell) to improve the estimation of a parameter of interest (e.g., AGB). From a statistical standpoint, the model-assisted estimator for total AGB can be expressed in the following form (Särndal et al., 1992, p. 231):

$$\widehat{TotalAGB_{ma}} = \sum_U \hat{y}_k + \frac{N}{n_s}\sum_s (y_k - \hat{y}_k) \quad (4)$$

The *model-assisted* estimator has two components: the sum of the lidar-predicted values ($\sum_U \hat{y}_k$) for the complete sample (U) of lidar grid cells, and a correction term based on the mean residual of the plot sample (s): ($\frac{N}{n_s}\sum_s (y_k - \hat{y}_k)$. The *synthetic* estimator of total AGB is obtained by using only the predicted values ($\sum_U \hat{y}_k$) without the residual correction term (Särndal et al., 1992, p. 399). The effect of the residual correction term is to remove bias from the model and, as a result, the model-assisted estimator is approximately design-unbiased (i.e. bias goes to zero at large sample sizes), while the synthetic estimator is not design-unbiased. In the case where the estimation is within subsets of the population (i.e. management units in our case, or impacted vs. non-impacted forest classes), the model-assisted estimator is modified such that the first term represents a sum over the subset and only residuals within that subset are used in the calculation of the second (residual correction) term (Särndal et al., 1992, p. 399). Model-assisted estimators for smaller subsets are also approximately design-unbiased. In model-assisted

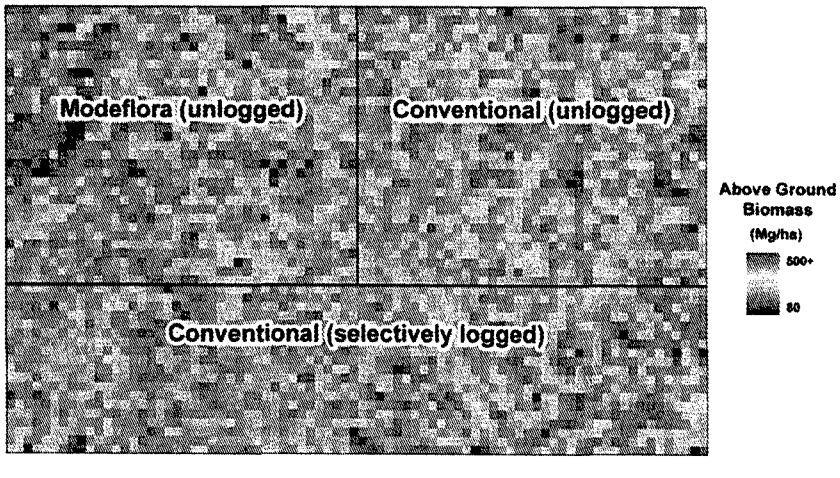

Fig. 4. Lidar model predictions of above ground biomass for the FEA study site by management area (50 m cells).

estimation, the reduction in the variance due to use of the model is directly related to the strength of the linear relationship between the predictor variables and inventory parameter. Not surprisingly, the variance of model-assisted estimators is always higher than that of the synthetic estimator, since variability due to model bias is accounted for in the model-assisted estimator. The variance of the field plot-based estimates with simple random sampling (SRS) is given by Särndal et al. (1992, p. 68):

$$\hat{V}_{srs} = N^2 \frac{1-f}{n_s} \frac{\sum_s (y_k - \bar{y})^2}{n_s - 1} = N^2 \frac{1-f}{n_s} S_y^2 \tag{5}$$

The variance of the model-assisted estimator is given by Särndal et al. (1992, p. 402):

$$\hat{V}_{ma_d} = N_d^2 \left(\frac{1}{n_{s_d}} - \frac{1}{N_d}\right) \frac{\sum_{s_d} \left(e_{k_s} - \bar{e}_{s_d}\right)^2}{n_{s_d} - 1} \tag{6}$$

It should be noted that in the case where the domain of interest is the entire area, this variance estimator can also be expressed as (Särndal et al., 1992, p. 276):

$$N^2 \left(\frac{1}{n_s} - \frac{1}{N}\right) S_y^2 \left(1 - R^2\right) = \hat{V}_{srs} \left(1 - R^2\right) \tag{7}$$

where n_s is the number of 0.25 ha sample plots, N is the population size (e.g. 4000 in the 1000 ha study area), $\hat{V}_{srs}$ is the variance of the plot SRS, and R^2 is the coefficient of determination of the multiple regression model. It is clear from these formulae that the reduction in variance gained by using the model-assisted approach compared to the SRS approach is a direct function of the strength of the regression relationship (R^2); and, in the case of estimation within domains, the variance is a direct function of the variability of the residuals about their mean within the domains (rightmost term in Eq. 6).

3.4.1. Hypothesis testing via bootstrapping

A bootstrap algorithm was used to determine if mean AGB was significantly different between identified impacted and non-impacted areas. The bootstrap approach to hypothesis testing was based on the algorithm presented in Efron and Tibshirani (1994), and in the case of the model-assisted estimator, incorporated the bootstrap variance estimation technique for the regression estimator in two-phase sampling developed by Sitter (1997). In this approach, at each iteration, a sample of size n_s was randomly drawn with replacement from the field plots; then, a regression model was developed using this sample and applied to the lidar grid cells within a given domain (i.e., area of the logged unit that were either impacted or non-impacted). These lidar predictions are used to calculate a model-assisted estimate of biomass within a given domain. The correction term is developed using the residuals for a random selection of plots equal in number to the actual number of plots located in the domain. This procedure is carried out many times ($n = 5000$), resulting in a bootstrap distribution of the difference between means under the null hypothesis. The percentage of bootstrap samples that are more extreme than the observed value (i.e., mean AGB) from the domain of interest

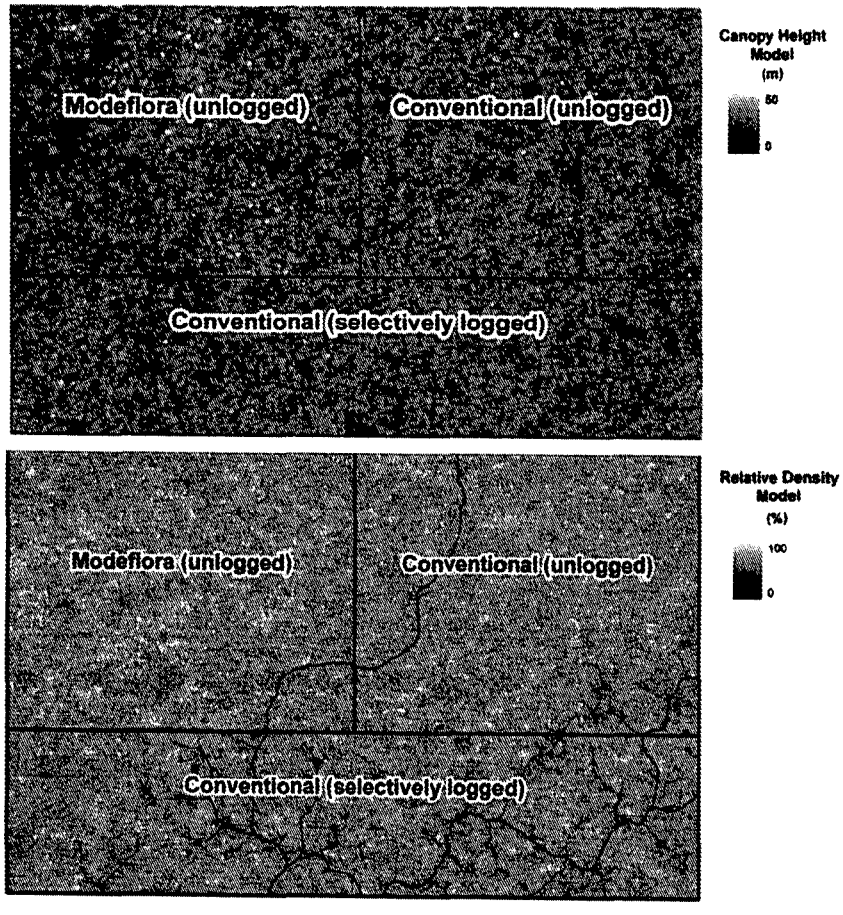

Fig. 5. High-resolution (1 m) canopy height model (top). Note that roads and skidtrails are not evident. Relative density model (RDM) for FEA study area using all lidar points above 1 m and below 5 m above ground (bottom). The gray scale represents the lidar return density (black low, light high) in the RDM height stratum. Main and secondary roads, landings, skid trails and logged tree gaps are black.

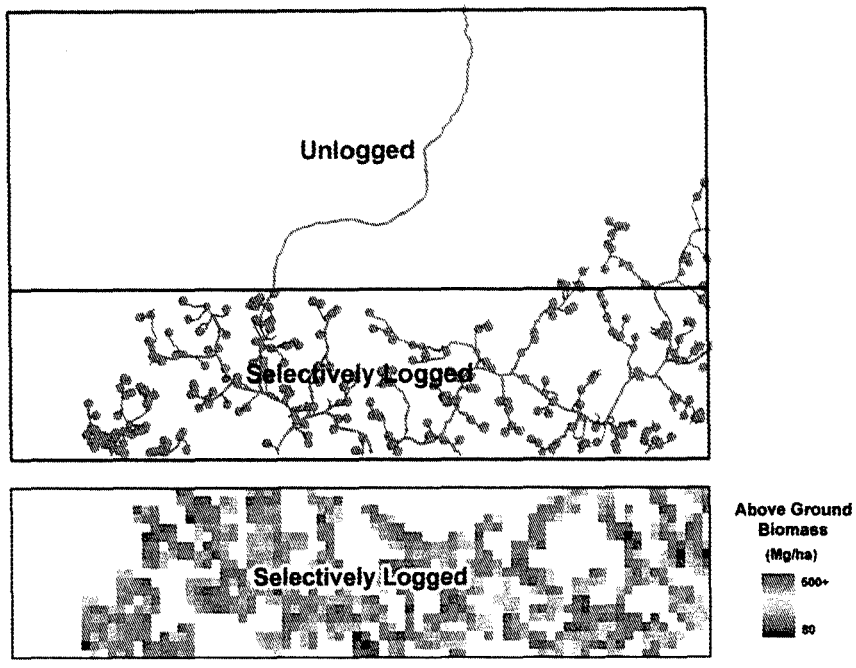

Fig. 6. Areas impacted by harvesting (top) were digitized from the RDM and buffered in GIS. Bottom panel illustrates the lidar model predictions of AGB for 50 m cells identified as impacted by harvesting.

(e.g., impacted plots) represents the Achieved Significance Level (ASL), a bootstrap analogue to the *p*-value for a traditional hypothesis test. It should be noted that for the synthetic and plot-based SRS estimates, the ASL obtained from this bootstrapping approach could be directly compared to the *p*-values obtained from traditional ANOVA testing of the equality of means for AGB in impacted vs. non-impacted areas.

4. Results

4.1. Regression modeling and biomass estimation

Table 4 summarizes the models evaluated for AGB, VOL, and BA. Lidar return quadratic mean height above ground (QElevMean) and

Table 4
Lidar-based above ground biomass, volume, and basal area regression models developed using the FEA ground plots ($n = 50$). Square root transformation of response variables and back transformed equation with bias correction factors (CF) are included for biomass and volume regressions.

Forest structure variable	Regression model	CF	Multiple R^2	Root-mean-square error
Sqrt(Biomass) (Mg ha^{-1})	$\sqrt{AGB} = 3.119^{*} + 0.564^{***} \times P25 + 0.062^{***} \times VAR$	-	0.70	1.28
Biomass (Mg ha^{-1})	$AGB = (3.119 + 0.564 \times P25 + 0.062 \times VAR)^2 + CF$	1.74	0.72	40.20
Sqrt(Biomass) (Mg ha^{-1})	$\sqrt{AGB} = -1.583 + 0.796^{***} \times QElevMean$	-	0.71	1.37
Biomass (Mg ha^{-1})	$AGB = (-1.583 + 0.796 \times QElevMean)^2 + CF$	1.95	0.67	43.25
Sqrt(Biomass) (Mg ha^{-1})	$\sqrt{AGB} = -0.834 + 0.837^{***} \times ElevMean$	-	0.63	1.43
Biomass (Mg ha^{-1})	$AGB = (-0.834 + 0.837 \times ElevMean)^2 + CF$	2.12	0.63	46.12
Sqrt(Volume) (m^3 ha^{-1})	$\sqrt{VOL} = 4.062^{***} + 0.496^{***} \times P25 + 0.046^{***} \times VAR$		0.68	1.10
Volume (m^3 ha^{-1})	$VOL = (4.062 + 0.496 \times P25 + 0.046 \times VAR)^2 + CF$	1.21	0.69	30.48
Sqrt(Volume) (m^3 ha^{-1})	$\sqrt{VOL} = 0.562 + 0.636^{***} \times QElevMean$	-	0.65	1.15
Volume (m^3 ha^{-1})	$VOL = (0.562 + 0.636 \times QElevMean)^2 + CF$	1.32	0.66	31.96
Sqrt(Volume) (m^3 ha^{-1})	$\sqrt{VOL} = 1.000 + 0.677^{***} \times ElevMean$	-	0.63	1.17
Volume (m^3 ha^{-1})	$VOL = (1.000 + 0.677 \times ElevMean)^2 + CF$	1.38	0.63	33.04
BA (m^2 ha^{-1})	$BA = -3.503 + 1.447^{***} \times P25 + 0.101^{***} \times VAR$	-	0.63	3.18
BA (m^2 ha^{-1})	$BA = -11.512^{**} + 1.631^{***} \times QElevMean$	-	0.59	3.32
BA (m^2 ha^{-1})	$BA = -11.011^{**} + 1.770^{***} \times ElevMean$	-	0.60	3.29

* denotes $p < 0.05$ significance; *** denotes $p < 0.001$ significance.

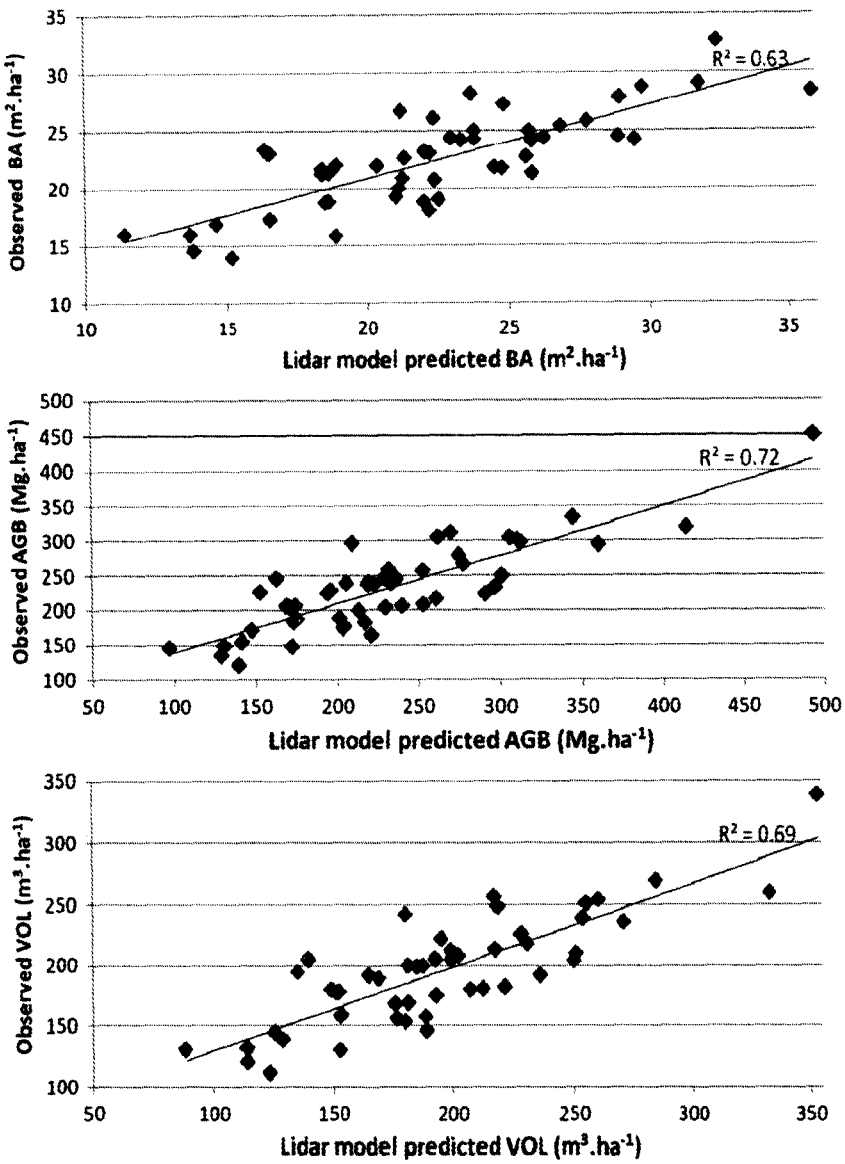

Fig. 7. Predicted versus observed ground plot values for basal area (top), above ground biomass (middle), and tree bole volume (bottom) for models that use P25 and VAR lidar explanatory variables.

mean height (ElevMean) provided univariate regression models with the best overall fit based on proportion of explained variance and root-mean-square error. The 25th percentile height above ground of all lidar returns (P25) and the variance of all lidar return heights above ground (VAR) provided robust, parsimonious multivariate models. Inclusion of more lidar variables produced models with marginally higher R^2 statistics; however, such models tended to over fit the observed data (often resulting in high VIF values between similar lidar explanatory metrics, e.g., 80th percentile height and 90th percentile height). Including highly collinear lidar metrics would limit their usefulness as predictive models over the wide range of forest structure conditions in the FEA study site. Fig. 7 displays the predicted and observed BA, VOL, AGB values for each plot for the models that use P25 and VAR lidar explanatory variables. The multivariate model was used for further comparisons of AGB estimates.

Table 5 summarizes AGB inventory estimates (unit total, mean and standard error of the mean) from the ground plots for the logged and unlogged management units, as well as impacted and non-impacted areas. Table 6 summarizes the AGB estimates (unit total, mean and standard error of the mean) for the units developed from the lidar synthetic and model-assisted estimators. The AGB means from the plots alone are similar to those obtained using the lidar synthetic- and model-assisted estimators (Table 6); however the model-assisted standard errors are improved (standard errors were not calculated for the synthetic estimator because there was no accounting for bias).

Table 5
Summary of SRS plot-based inventory estimates (total and mean above ground biomass and standard error), FEA State Forest.

		Total biomass (Mg)		Mean biomass (Mg ha^{-1})		
	n	Estimate	SE	Estimate	SE	RSE
Unlogged compartments	30	146081.8	9237.2	235.6	13.6	6.3%
Modeflora	15	67958.8	6796.4	219.2	17.2	7.9%
Conventional unlogged	15	78123.0	6431.2	252.0	20.9	8.3%
Logged compartment						
Conventional logged	20	85027.0	5361.6	223.8	16.7	7.4%
Impacted	10	38961.6	3743.4	219.5	16.8	7.7%
Non-impacted	10	46171.7	4027.3	228.0	29.9	13.1%
Total	50	230871.6	10477.4	230.9	10.5	4.5%

Table 6
Summary of lidar model-assisted inventory estimates (total and mean above ground biomass and model-assisted standard error; SE), FEA State Forest.

			Total Biomass (Mg)			Mean Biomass (Mg ha^{-1})			
	N	*n*	Synthetic	Model-assisted	SE	Synthetic	Model-assisted	SE	RSE
Unlogged compartments	2480	30	144520.7	142806.3	4611.7	233.1	230.3	7.4	3.2%
Modeflora	1240	15	71475.7	68824.1	2813.1	230.6	222.0	9.1	4.1%
Conventional unlogged	1240	15	73045.0	73982.2	3756.8	235.6	238.7	12.0	5.1%
Logged compartment									
Conventional logged	1520	20	87172.8	88649.6	3104.6	229.4	233.3	8.2	3.5%
Impacted	710	10	39993.8	40899.2	1880.4	225.3	230.0	11.0	4.6%
Non-impacted	810	10	47179.0	47720.0	2636.3	233.0	236.0	12.0	5.5%
Total	4000	50	231693.5	231589.0	5477.1	231.7	231.6	5.5	2.4%

4.2. Logging impact assessment

Despite the relatively high resolution (1 m) of the lidar CHM, it was not sufficient to visually identify harvest impacted and non-impacted areas in the study site (Fig. 5). Manipulation of the RDM in GIS produced a raster image where the areas with low numbers of lidar returns above 1 m and below 5 m could be visually distinguished from areas with high numbers of returns in this height stratum. Ground data points collected to validate the impacted areas identified in the RDM align very well with areas with low RDM values. All of the 45 GNSS points collected along roads, skidtrails, on landings, and in tree gaps associated with selective logging were within the lidar-predicted impacted raster cells. Thus, we infer that dark areas connected by linear features in Fig. 5 represent areas where vegetation was removed or crushed during the construction of roads, skidtrails, landings or removal of trees, resulting in understory canopy gaps. The impacted area was estimated to be 58.4 ha or 15.4% of the logged unit. Half of the ground plots (10/20) within the unit were classified as impacted.

The bootstrap distribution corresponding to the null hypothesis of no difference between means in impacted and non-impacted areas, and the observed difference, is shown in Fig. 8. Using the bootstrap hypothesis testing approach, neither the mean AGB within impacted forest areas calculated from the ground plots (ASL = 0.80) nor the mean AGB within impacted areas calculated using the lidar model-assisted approach (ASL = 0.75) were found to be significantly different from the mean AGB within non-impacted areas. However, the AGB means computed using the lidar synthetic estimator were significantly different between impacted and non-impacted areas (ASL = 0.01). The lidar estimate of mean AGB obtained from synthetic estimator (225.3 Mg ha^{-1}) of the 50-m cells impacted by logging (Fig. 6) was 7.7 Mg ha^{-1} lower than the mean AGB (233.0 Mg ha^{-1}) for the non-impacted cells. The difference of means for the P25 lidar explanatory variable between the impacted and non-impacted areas was highly significant ($p<0.0016$); whereas, there was no significant difference for VAR ($p<0.59$), indicating that changes in lower canopy structure as measured by the P25 variable (and identified in the RDM) caused the majority of the difference in the lidar AGB regression model. The corresponding difference in the lidar model-assisted estimate of mean AGB between impacted (230.0 Mg ha^{-1}) and non-impacted (236.0 Mg ha^{-1}) areas was 6.0 Mg ha^{-1}. However, the relatively high variances associated with the model-assisted estimators (Table 6) did not provide adequate statistical power to detect a significant difference between these means.

5. Discussion

5.1. Regression model selection for above ground biomass

In this study, our R^2 values (0.63–0.72) of the AGB regression models were lower than those reported in similar studies in predominately coniferous temperate and boreal forests (e.g., Li et al., 2008; Lim & Treitz, 2004; Means et al., 1999; Næsset et al., 2004) but similar to studies conducted in mixed conifer and hardwood forest types (e.g., Lefsky et al., 1999; Popescu et al., 2003) and tropical forests (e.g., Asner et al., 2009; Drake et al., 2002a, 2002b; Kennaway et al., 2008). This is not surprising given that we were limited to a single AGB allometric equation based solely on diameter for all species.

Selection of P25 and VAR as lidar explanatory variables provided a predictive model for AGB that is logically related to canopy height, variation and transparency expected in highly diverse native tropical canopies. Steinhilb and Erickson (1972) found that center of mass (CM) of the above ground tree material (bole, limbs, and foliage) of three conifer and two deciduous species (Steinhilb & Erickson, 1970; Steinhilb & Winsauer, 1976) occurred between 36% and 41% of total tree height. Fridley and Tufts (1989) reported that the CM of Loblolly pine (Pinus taeda) occurred between 36% and 44% of tree height. In all these studies of tree CM, the height above ground of the CM is directly correlated with AGB. For the FEA field plots, the mean P25 height above ground occurred at 35% of the mean plot-level 99th percentile height (P99). P99 corresponds to plot canopy near-maximum height. (Lidar return maximum height (P100) is not always a reliable measure of maximum canopy height because the lidar point cloud can include returns from birds or other anomalous returns above the canopy that are not filtered out of the raw lidar data). The P25 lidar metric is likely to be well correlated with the CM of the AGB of other tree species and, by summation of individual plot trees, correlated to the CM of plot-level AGB. Interestingly, in a study in which trees were destructively sampled by Beets et al. (2011b), the 30th percentile height (P30) of all returns was highly predictive of AGB in young radiata pine (*Pinus radiata*) plantations. P30 and a lidar percent cover metric were the lidar explanatory variables used to predict carbon in a nation-wide forest inventory of New Zealand forest plantations established after 1989 as required under the United Nations Kyoto Protocol (Beets et al., 2011a).

The VAR lidar metric (variance of lidar return heights above ground, including all returns) is related to both the variability of the canopy height and its transparency with regards to passage of lidar pulses through gaps in foliage and branches. In areas with tall canopy, if the canopy is dense, there are proportionally fewer second and third returns reducing VAR; whereas, if the canopy is sparse, a lidar pulse has a greater probability of generating a second or third return from below the canopy surface, increasing VAR. Additionally, in areas with dense canopy the distance from the first return at the canopy surface to second or third returns would likely be less than in areas with sparse canopy, again resulting in lower values for VAR in dense canopy. At the same time, the value of P25 would be higher for dense canopy areas compared to the same canopy height with sparse cover. The all-returns lidar VAR explanatory variable is strongly correlated with the P99 canopy height ($r = 0.81$) and has a positive coefficient in the AGB model (Table 4). Thus, in this type of forest with highly variable canopy height, VAR not only characterizes canopy surface variability and transparency, but to a large extent, canopy dominant height.

5.2. Improvement of AGB area estimates by use of lidar-based methods

The bootstrap variance estimation technique for the regression estimator in two-phase analysis produced AGB inventory estimates that

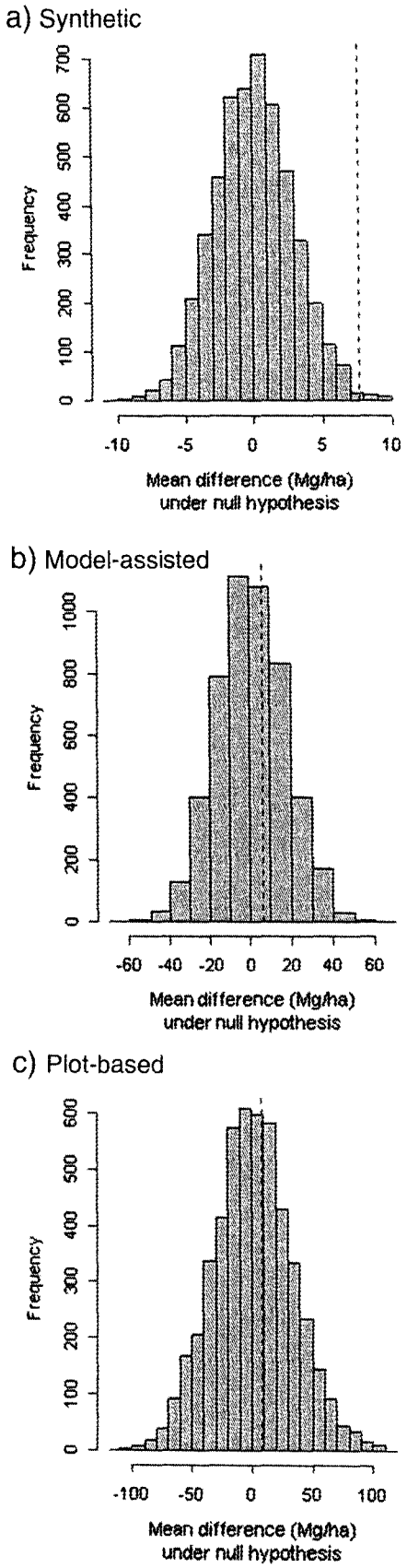

Fig. 8. Bootstrap distributions under null hypothesis (no difference between means) for impacted vs. non-impacted areas (vertical dashed line indicates observed value) using: a) lidar synthetic estimation (ASL = 0.01); b) lidar model-assisted estimation (ASL = 0.75); and c) plot-based estimation (ASL = 0.80).

were more precise (i.e., lower standard errors) than estimates based solely on the field plots. Our AGB standard error (5.5 Mg ha^{-1}, 2.4% of mean) was similar in magnitude to AGB errors reported by Gonzalez et al. (2010) for temperate conifer forests using a Monte Carlo model-assisted approach. Because lidar data were collected over the entire site, it was possible to map lidar estimates of AGB at 50 m resolution, which, along with the CHM, provides forest planners with more spatially accurate and detailed planning information than is possible via ground data collection methods.

5.3. Selective logging impacts

Assessment and monitoring of small scale forest disturbance (e.g., single tree and small group selection logging) using passive remote sensing techniques in natural tropical forests is particularly difficult. Passive optical remote sensing techniques (e.g., aerial photography and multispectral satellite imagery) are limited in their capacity to distinguish structural changes occurring below the top of the canopy (Coops et al., 2007). When viewed from above, the post-disturbance canopy surface is within the range of variability observed in undisturbed areas. Canopy cover of residual dominant and codominant tree crowns can be very high, obscuring roads, skidtrails, landings, and removal of individual trees. Thus, selective logging impacts are very difficult to assess from remotely-sensed data (e.g., Asner et al., 2004).

We found in this study that information from field plots combined with conventional lidar-derived products such as a CHM (Fig. 5) and predicted AGB (Fig. 4) at relatively fine scales (1 m and 50 m, respectively) did not provide sufficient information to locate and characterize selective logging disturbances. The lidar model-assisted prediction of the mean AGB did not show a significant difference between undisturbed areas and those portions of the selectively logged management unit where roads, skidtrails, landings and harvested tree gaps occurred. The fact that there was a highly significant difference between mean AGB estimates obtained using the lidar synthetic estimator indicates that increased efforts to improve the lidar regression models (i.e. reducing residual error due to plot location error, imprecision of allometric models, etc.—see Section 5.4) will allow for improved estimates of reduction in biomass (and carbon) due to selective logging activities.

Because lidar is an active remote sensing technology that provides a measure of the three-dimensional structure of the canopy, different portions of the lidar point cloud can be used to measure different portions of the canopy structure (Andersen et al., 2006). The RDM provided a characterization of the canopy structure for the height stratum between 1 and 5 m above ground where many of the effects related to recent logging activities were likely to be evident. In areas where no selective logging had occurred, dense overstory canopy often completely blocked the passage of lidar pulses in this height layer, resulting in no logical mathematical value for the RDM cell (division by zero). In areas with sparser overstory canopy a sufficient portion of laser energy passed into or completely through this RDM height layer. If numerous lidar returns were generated in this layer, it was assumed that foliage and branches occurred in the layer, indicating that the canopy midstory layer was not heavily impacted. However, if most of the laser energy passed through this layer (i.e., there were no returns or only a few returns in this layer) and reflected from vegetation below 1 m or the ground, the midstory might have been heavily impacted by logging or the area might have experienced a natural disturbance.

The RDM was used to identify specific features associated with selective logging (truck roads, skidtrails, landings, and tree gaps) where there were low numbers of midstory lidar returns. These features were not evident in the 1-m resolution CHM (Fig. 5) because they were concealed by the high level of residual overstory canopy that remained after selectively logging. These features have distinctive characteristics in the RDM that assist in their identification: roads and skidtrails are narrow, linear features; landings and tree gaps are

typically 20–25 m roughly circular shapes connected by roads and skidtrails.

The additional criterion for separating features in the RDM due to natural disturbances from recent logging features is imposed by the systems used to extract logs. In FEA all logging was carried out with ground-based systems: log skidders and trucks. Therefore, all harvested tree locations were connected by roads, landings, or skidtrails that appear as linear features in the RDM. Only those areas that had low RDM values and were near these linear features were identified as logged areas. Other scattered dark areas that were not connected by dark linear features were likely natural openings in the canopy midstory. Obviously, this approach would not be effective in areas that were selectively logged with aerial systems such as helicopter or full-suspension cableways.

Once these logging features were identified, the synthetic estimator was used to test for differences in lidar-predicted AGB (at 50 m resolution) between undisturbed areas and impacted areas within the selectively logged unit. Although a low percentage of the total volume (less than 10%) was harvested, the difference in the AGB means (7.7 Mg ha^{-1}) was highly significant ($p<0.01$) and of the same magnitude as the planned harvest volume (12–18 Mg ha^{-1} planned removals). A pre-logging, spatially-explicit map of AGB would be needed to definitively attribute this difference to logging. However, because the impacted areas are dispersed throughout the entire unit (Fig. 5) and spatially associated with logging features, it is likely that the difference in AGB is associated with selective logging and not due simply to natural variation in the unit.

The area impacted by selective logging as determined from the RDM layer was 15.4% of the management unit. This is similar to what Asner et al. (2004) calculated through intensive field mapping of roads, skidtrails, landings, and tree gaps for conventional selective logging in eastern Amazonia. The lidar RDM approach to identifying logging impact areas may be a viable alternative to such intensive field surveys. Further research is needed to test whether object-based image analysis methods can be used to automate the delineation of roads, skidtrails, landings, and tree gaps within the RDM. Additionally, although not the focus of our study, a robust estimate of canopy gap fraction associated with the selective logging is easily computed from the lidar point cloud. Canopy gap fraction is a key data element in the regional selective logging satellite mapping approach presented by Asner et al. (2004).

More research is needed to determine what density of lidar is necessary to reliably identify skidtrails and small logging roads under heavy residual canopy. The density of our lidar acquisition (25 pulses m^{-2}) was high, allowing generation of the RDM at 1 m resolution with extremely high levels of detail. For example, several manually cleared walking trails used by rubber tappers and forestry inventory crews were evident in the RDM. This level of detail is beyond that needed for monitoring selective logging.

Logging in our study area had started about a month before lidar data were acquired. Therefore, regenerating vegetation in tree gaps, skidtrails, and landings had not grown over 1 m tall, the lower limit of the RDM layer. If the lidar acquisition had been delayed, the lower and upper limits of the RDM would likely need to be adjusted. Further research is needed to evaluate if the RDM approach can be used when logging has occurred one or more years prior to lidar acquisition.

By using two scales of analysis and the lidar synthetic estimator we were able to identify a statistically significant difference in estimates of AGB between selectively logged and unlogged areas that was not evident from ground plots or from model-assisted lidar predictions. In our study, we had lidar coverage over the entire study area; however, a similar approach for estimating area recently selectively logged versus undisturbed could be incorporated into a multi-level inventory framework (Andersen et al., 2011).

5.4. Factors affecting the accuracy and precision of lidar-based AGB estimation

There are two likely major sources of unexplained variance of the AGB regression model (Mascaro et al., 2011). The first is mis-registration of the ground plot locations to the lidar point cloud. The second is poor estimation of AGB due to limited allometric equations and ground measurements.

While we tried to obtain the most accurate positions possible for the plot corners, the dense, multi-layered canopy and tall, large-diameter stems in our study site represent difficult conditions for acquiring accurate GNSS positions. Næsset (2001), Clarkin (2007), Andersen et al. (2009) and Valbuena et al. (2010) have all completed studies of GNSS accuracy in conifer dominated forests; however, we have not found similar studies in dense tropical rainforests where GNSS signal tracking may be considerably more difficult. In heavy canopy, the direct path from a satellite to the GNSS receiver can be completely blocked and only the reflected (multipath) signal is received. Multipathing occurs when the signal reflects off of a nearby reflective surface (e.g., large tree bole) before being detected by the receiver. Because the multipath signal has traveled farther than the straight-line distance to the satellite, it can cause large positional errors in heavily forested areas, particularly when only a low number of satellites can be tracked, making it impossible to identify and eliminate the multipath signal from the position solution.

In our first GNSS field campaign, only GPS satellite signals were available from the Rio Branco base station. Examination of the individual epoch positions post-processed using only GPS data revealed widely scattered clouds of positions that varied by as much as 50 m horizontally. This was particularly pronounced at plot corners where a low number of GPS satellites were tracked due to heavy canopy. Clarkin (2007) found that increasing the number of satellites tracked and used in the calculation of a position increases the accuracy of the position when surveying locations in heavy forest. As the number of tracked satellites increases, the likelihood of having sufficient non-multipath satellite signals increases, allowing better detection and elimination of any multipath signals, resulting in more reliable post-processed positions.

In our second GNSS field campaign, we established a temporary base station that tracked both GPS and GLONASS satellites and then resurveyed a subset of the plot corners that had the poorest calculated positional precision from the first GPS-only campaign. For 98 field locations for which both GPS and GLONASS base station data were collected, the number of satellites tracked increased from 6.6 with GPS alone to 11.5 with GPS and GLONASS. The mean post-processed precision of the field locations improved from 0.45 m (SD = 0.44 m) for GPS alone to 0.31 m (SD = 0.16 m) when both GPS and GLONASS data were used. The mean horizontal difference in field locations computed from GPS alone versus GPS and GLONASS was 1.71 m (SD = 1.99 m). However, 7 field locations had horizontal differences greater than 5 m. The mean number of GPS signals available for post-processing these positions was only 4.3, compared to 8.5 signals with GPS and GLONASS. We believe the reason for these large differences in GPS-only versus GPS and GLONASS positions is due to multipath signals that the post-processing software was unable to eliminate due to the minimal number of available GPS satellite signals. For this reason, we recommend that GNSS systems capable of collecting and post-processing GPS and GLONASS data be used when surveying plot locations under heavy, non-deciduous forest canopy. For the final set of plot corner locations, we used the GNSS rover file that had the greatest number of satellites available for post-processing.

The precision of plot corner locations reported by the GNSS post-processing software was generally less than 0.5 m which is similar to location precision reported by Valbuena et al. (2010) using similar equipment, processing, and field protocols. However, the accuracy of the locations cannot be determined without a rigorous ground survey

using conventional closed traverse methods. Such a ground survey was not possible for this study due to time and cost limitations. The Valbuena et al. (2010) study, which was conducted in mature coniferous forests in Spain, reported actual horizontal error of plot positions approximately 2–3 times the precision reported by the post-processing software. Therefore, we would expect our horizontal error to likely be greater, in the range of 2–5 m, due to the dense, tropical canopy.

Frazer et al. (2011) examined the influence of sample plot size and location errors on accuracy and uncertainty of lidar biomass estimates using simulated stand and lidar datasets. Their results indicated that with large plot sizes (>0.125 ha), plot positional errors of up to 5 m had little effect on lidar-derived estimates. With a plot size of 0.25 ha and estimated plot location error of less than 5 m, the results of Frazer et al. (2011) would suggest that our lidar explanatory metrics for each plot should not have been strongly affected by misregistration of ground plots to the lidar point cloud. However, their simulated stand was composed of a single conifer species (Douglas-fir) that has a well behaved conical or cylindrical crown form, unlike many of the dominant tropical tree species found in our study site that have widely spreading, irregular crown shapes. Asner et al. (2009) also reported that AGB estimation errors due to misalignment of lidar and field plot data were in the range of 0 to 10 Mg ha^{-1} in a study of tropical forests conducted in Hawaii. More research is needed to better estimate GNSS positional errors in dense tropical forests and to establish guidelines for minimum data collection time at each plot location, the most robust GNSS post-processing methods, and whether post-processing using multiple GNSS constellations is required to insure plot location accuracy of 5 m or better.

In this study, stem diameters were measured using the generally applied rules for tropical forests (diameter measurement at breast height (1.3 m) or above basal buttress). DBH measurements by field personnel can vary depending on the field crew's experience, tree characteristics (e.g. bark, trunk irregularity, bifurcated trunks, lianas), and topography (slope). In addition, variations on the height of the DBH measurement can produce significant differences on the estimated parameters (e.g., Brokaw & Thompson, 2000). The equation used to compute above ground biomass had only diameter as the independent variable. Diameter alone is used, especially in tropical forests, because it is easy to obtain with a relatively high accuracy. The use of height as a second independent variable would likely improve individual tree AGB estimates; however, it is extremely difficult to measure accurate tree heights in tropical forests (e.g., Asner et al., 2011; d'Oliveira et al., 2011). Thus, biomass equations based only on DBH are generally used and accepted in tropical forests studies (Brown, 1997; Chave et al., 2008; Higuchi et al., 1998; Nogueira et al., 2008; Sierra et al., 2007). It is very likely that diameter will continue to be the main variable measured for biomass studies in tropical forests. We also recognize that using a single biomass equation in an ecosystem with hundreds of tree species (with different growth patterns, crown architecture, and crown expansion as a result of canopy position) may introduce high levels of measurement error for individual tree AGB and likely contributes significantly to the remaining unexplained variance in our lidar AGB model. Further tests are needed to determine if use of allometric equations for individual species and adjustments for regional wood density differences would reduce ground plot AGB measurement error and improve model fits, as reported by Asner et al. (2011).

As several previous studies have found (Drake et al., 2003; Gobakken & Naesset, 2008; Hopkinson, 2007; Rombouts et al., 2010), we also would expect the AGB model would be affected by lidar sensor parameters, acquisition mission specifications, and seasonal conditions. In our study, lidar data were flown at the end of the rainy season before any significant leaf fall. We would expect that results would be different if lidar data were reacquired with the same sensor and mission specifications in the dry season with significant leaf-off conditions. The lidar sensor used in this study was a discrete-return system that did not provide calibrated return intensity data; therefore, we did not use any lidar metrics related to portion of energy returned from the canopy (intensity of discrete returns). Instead, we were limited to metrics which describe the vertical structure of the lidar point cloud.

5.5. Direct lidar models and applications for forest management

There are several products that can be derived directly from the LIDAR point cloud that are not dependent on ground plot measurements. Specifically, these are: bare earth, canopy surface, canopy height, canopy cover, and canopy relative density models and high-resolution orthographic intensity images created from return intensity values. Lidar bare earth models describe topography and drainages at unprecedented resolution, allowing reliable road and operations planning with greatly reduced field work. Canopy height and cover models provide a direct method for monitoring change over time of the canopy surface. Lidar-based high-resolution orthographic images are useful as base maps where recent photographic imagery of comparable scale is not available. As shown in the FEA site, the RDMs provide a method to detect fine-scale disturbance in the under- and midstory canopy. For all these products, further tests are needed to determine if lidar acquisitions with lower pulse densities and greater flying heights can provide sufficiently accurate products as lower costs.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.rse.2012.05.014.

References

Andersen, H. -E., Clarkin, T., Winterberger, K., & Strunk, J. (2009). An accuracy assessment of positions obtained using survey- and recreational-grade global positioning system receivers across a range of forest conditions within the Tanana Valley of interior Alaska. *Western Journal of Applied Forestry, 24*(3), 128–136.

Andersen, H. -E., Reutebuch, S. E., & Mcgaughey, R. J. (2006). Active remote sensing. In G. Shao, & K. Reynolds (Eds.), *Computer applications in sustainable forest management* (pp. 281–304). Dordrecht, Netherlands: Springer-Verlag.

Andersen, H. -E., Strunk, J., Temesgen, H., Atwood, D., & Winterberger, K. (2011). Using multi-level remote sensing and ground data to estimate forest biomass resources in remote regions: A case study in the boreal forests of interior Alaska. *Canadian Journal of Remote Sensing, 37*(6), 596–611.

Asner, G. P., Hughes, R. F., Varga, T. A., Knapp, D. E., & Kennedy-Bowdoin, T. (2009). Environmental and biotic controls over aboveground biomass throughout a tropical rain forest. *Ecosystems, 12*, 261–278.

Asner, G. P., Keller, M., & Silva, J. N. M. (2004). Spatial and temporal dynamics of forest canopy gaps following selective logging in the eastern Amazon. *Global Change Biology, 10*, 765–783.

Asner, G. P., Mascaro, J., Muller-Landau, H. C., Vieilledent, G., Vaudry, R., Rasamoelina, M., et al. (2011). A universal airborne LiDAR approach for tropical forest carbon mapping. *Oecologia*. http://dx.doi.org/10.1007/s00442-011-2165-z.

Asner, G. P., Powell, G. V. N., Mascaro, J., Knapp, D. E., Clark, J. K., Jacobson, J., et al. (2010). High-resolution forest carbon stocks and emissions in the Amazon. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 16738–16742.

Beets, P. N., Brandon, A. M., Goulding, C. J., Kimberley, M. O., Paul, T. S. H., & Searles, N. (2011a). The inventory of carbon stock in New Zealand's post-1989 planted forest for reporting under the Kyoto protocol. *Forest Ecology and Management, 262*(6), 1119–1130.

Beets, P. N., Reutebuch, S., Kimberley, M., Oliver, G., Pearce, S., & McGaughey, R. (2011b). Leaf area index, biomass carbon and growth rate of radiata pine genetic types and relationships with LiDAR. *Forests, 2*(3), 637–659.

Brokaw, N., & Thompson, J. (2000). The H for DBH. *Forest Ecology and Management, 129*, 89–91.

Brown, S. (1997). Estimating biomass and biomass change of tropical forests. *FAO Forestry Paper, 134*, 55p.

Chave, J., Olivier, J., Bongers, F., Chatelet, P., Forget, P. M., Meer, P., et al. (2008). Above-ground biomass and productivity in a rain forest of eastern South America. *Journal of Tropical Ecology, 24*, 355–366.

Clarkin, T. 2007. Modeling global navigation satellite system positional error under forest canopy based on LIDAR-derived canopy densities. M.S. Thesis. College of Forest Resources, University of Washington, Seattle, WA. 99pp.

Coops, N. C., Hilker, T., Wulder, M. A., St-Onge, B., Newnham, G., Siggins, A., et al. (2007). Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. *Trees - Structure and Function, 21*(3), 295–310.

d'Oliveira, M. V. N., Alvarado, E. C., Santos, J. C., & Carvalho, J. A., Jr. (2011). Forest natural regeneration and biomass production after slash and burn in a seasonally dry forest in the Southern Brazilian Amazon. *Forest Ecology and Management, 261*(9), 1490–1498.

Drake, J. B., Dubayah, R. O., Clark, D. B., Knox, R. G., Blair, J. B., Hofton, M. A., et al. (2002a). Estimation of tropical forest structural characteristics using large-footprint lidar. *Remote Sensing of Environment, 79*, 305–319.

Drake, J. B., Dubayah, R. O., Knox, R. G., Clark, D. B., & Blair, J. B. (2002b). Sensitivity of large-footprint lidar to canopy structure and biomass in a neotropical rainforest. *Remote Sensing of Environment, 81*, 378–392.

Drake, J. B., Knox, R. G., Dubayah, R. O., Clark, D. B., Condits, R., Blair, J. B., et al. (2003). Above-ground biomass estimation in closed canopy neotropical forest using lidar remote sensing: Factors affecting the generality of relationships. *Global Ecology and Biogeography, 12*, 147–159.

Dubayah, R. O., Sheldon, S. L., Clark, D. B., Hofton, M. A., Blair, J. B., Hurtt, G. C., et al. (2010). Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica. *Journal of Geophysical Research, 115*, 1–17 G00E09. http://dx.doi.org/10.1029/2009JG000933.

Efron, B., & Tibshirani, R. (1994). *An introduction to the bootstrap.* Boca Raton, FL: Chapman and Hall/CRC.

ESRI (2011). *ArcGIS Desktop: Release 9.3.* Redlands, CA: Environmental Systems Research Institute.

Figueiredo, E. O., Braz, E. M., & Oliveira, M. V. N. (2007). *Manejo de precisão em florestas tropicais: Modelo digital de exploração florestal.* Rio Branco, Acre: Embrapa Acre 183pp.

Fox, J., & Monette, G. (1992). Generalized collinearity diagnostics. *JASA, 87*, 178–183.

Frazer, G. W., Magnussen, S., Wulder, M. A., & Niemann, K. O. (2011). Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. *Remote Sensing of Environment, 115*(2), 636–649.

Fridley, J. L., & Tufts, R. A. (1989). Analytical estimates of Loblolly pine tree center of mass and mass moment of inertia. *Forest Science, 35*(1), 126–136.

FUNTAC (1990). *Estrutura do plano de manejo de uso múltiplo da floresta Estadual do Antimari.* Rio Branco, Acre: Acre State Technological Foundation 103 pp.

Gobakken, T., & Naesset, E. (2008). Assessing effects of laser point density, ground sampling intensity and field sample plot size on biophysical stand properties derived from airborne laser scanner data. *Canadian Journal of Forest Research, 38*, 1095–1109.

Gonzalez, P., Asner, G. P., Battles, J. J., Lefsky, M. A., Waring, K. M., & Palace, M. (2010). Forest carbon densities and uncertainties from lidar, QuickBird, and field measurements in California. *Remote Sensing of Environment, 114*, 1561–1575.

Higuchi, N., Santos, J., Ribeiro, R. J., Minettte, L., & Biot, Y. (1998). Biomassa da parte aerea da da vegetacao da floresta tropical umida de terra-firme da Amazonia Brasileira. *Acta Amazonica, 28*, 153–166.

Hopkinson, C. (2007). The influence of flying altitude, beam divergence and pulse repetition frequency on canopy penetration and laser pulse return distribution characteristics. *Canadian Journal of Remote Sensing, 33*, 312–324.

Hyyppä, J., Hyyppä, H., Leckie, D., Gougeon, F., Yu, X., & Maltamo, M. (2008). Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. *International Journal of Remote Sensing, 29*(5), 1339–1366.

Kennaway, T. A., Helmer, E. H., Lefsky, M. A., Brandeis, T. A., & Sherrillc, K. R. (2008). Mapping land cover and estimating forest structure using satellite imagery and coarse resolution lidar in the Virgin Islands. *Journal of Applied Remote Sensing, 2*, 023551.

Lefsky, M. A., Cohen, W. B., Harding, D. J., Parker, G. G., Acker, S. A., & Gower, S. T. (2002). Lidar remote sensing of above-ground biomass in three biomes. *Global Ecology and Biogeography, 11*, 393–399.

Lefsky, M. A., Harding, D., Cohen, W. B., Parker, G., & Shugart, H. H. (1999). Surface lidar remote sensing of basal area and biomass in deciduous forests of eastern Maryland, USA. *Remote Sensing of Environment, 67*, 83–98.

Li, Y., Andersen, H. -E., & McGaughey, R. (2008). A Comparison of statistical methods for estimating forest biomass from light detection and ranging. *Western Journal of Applied Forestry, 23*(4), 223–231.

Lim, K. S., & Treitz, P. M. (2004). Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators. *Scandinavian Journal of Forest Research, 16*, 558–570.

Lumley, T. (2009). Leaps: Regression subset selection (using Fortran code by Alan Miller). http://CRAN.R-project.org/package=leaps Accessed 20 December 2011.

Magnussen, S., Næsset, E., & Gobakken, T. (2010). Reliability of LiDAR derived predictors of forest inventory attributes: A case study with Norway spruce. *Remote Sensing of Environment, 114*, 700–712.

Mascaro, J., Detto, M., Asner, G. P., & Muller-Landau, H. C. (2011). Evaluating uncertainty in mapping forest carbon with airborne LiDAR. *Remote Sensing of Environment, 115*, 3770–3774.

McGaughey, R. J. (2010). *FUSION/LDV: Software for LIDAR data analysis and visualization.* United States Department of Agriculture, Forest Service, Pacific Northwest Research Station 154 pp. (http://forsys.cfr.washington.edu/fusionlatest.html last accessed March 2012).

Means, J. E., Acker, S. A., Harding, D. J., Blair, J. B., Lefsky, M. A., Cohen, W. B., et al. (1999). Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the Western Cascades of Oregon. *Remote Sensing of Environment, 67*, 298–308.

Miller, D. M. (1984). Reducing transformation bias in curve fitting. *The American Statistician, 38*(2), 124–126.

Næsset, E. (1997). Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sensing of Environment, 61*(2), 246–253.

Næsset, E. (2001). Effects of differential single- and dual-frequency GPS and GLONASS observations on point accuracy under forest canopies. *Photogrammetric Engineering and Remote Sensing, 67*, 1021–1026.

Næsset, E., Gobakken, T., Holmgren, J., Hyyppä, H., Hyyppä, J., Maltamo, M., et al. (2004). Laser scanning of forest resources: The Nordic experience. *Scandinavian Journal of Forest Research, 19*(6), 482–499.

Nelson, R., Krabill, W., & Tonelli, J. (1988). Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment, 24*(2), 247–267.

Nogueira, E. M., Fearnside, P. M., Nelson, B. W., Barbosa, R. I., & Keizer, E. W. H. (2008). Estimates of forest biomass in the Brazilian Amazon: New allometric equations and adjustments to biomass from wood-volume inventories. *Forest Ecology and Management, 256*(11), 1853–1867.

Popescu, S. C., Wynne, R. H., & Nelson, R. F. (2003). Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Canadian Journal of Remote Sensing, 29*(5), 564–577.

R Development Core Team (2011). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Reutebuch, S. E., Andersen, H. -E., & McGaughey, R. J. (2005). Light detection and ranging (LIDAR): An emerging tool for multiple resource inventory. *Journal of Forestry, 103*(6), 286–292.

Reutebuch, S. E., McGaughey, R. J., Andersen, H. -E., & Carson, W. W. (2003). Accuracy of a high-resolution lidar terrain model under a conifer forest canopy. *Canadian Journal of Remote Sensing, 29*(5), 527–535.

Rombouts, J. H., Ferguson, I. S., & Leech, J. W. (2010). Campaign and Site effects in LiDAR prediction models for site quality assessment of radiata pine plantations in South Australia. *International Journal of Remote Sensing, 31*(5), 1155–1173.

Särndal, C. -E., Swensson, B., & Wretman, J. (1992). *Model-assisted survey sampling.* NY, NY: Springer 694 pp.

Sierra, C. A., Del Valle, J. I., Orrego, S. A., Moreno, F. H., Harmon, M. E., Zapata, M., et al. (2007). Total carbon stocks in a tropical forest landscape of the Porce region, Colombia. *Forest Ecology and Management, 243*, 299–309.

Sitter, R. (1997). Variance estimation for the regression estimator in two-phase sampling. *Journal of the American Statistical Association, 92*(438), 780–787.

Steinhilb, H. M., & Erickson, J. R. (1970). *Weights and centers of gravity for quaking aspen trees and boles.* USDA For. Serv. Res. Note NC-91*N.C. For. Exp. Stn., St. Paul, Minn.* 4 pp.

Steinhilb, H. M., & Erickson, J. R. (1972). *Weights and centers of gravity for red pine, white spruce, and balsam fir.* USDA For. Serv. Res. Paper NC-75*N.C. For. Exp. Stn., St. Paul, Minn.* 7 pp.

Steinhilb, H. M., & Winsauer, S. A. (1976). *Sugar Maple: Tree and bole weights, volumes, centers of gravity, and logging residue.* USDA For. Serv. Res. Paper NC-132*N.C. For. Exp. Stn., St. Paul, Minn.* 7 pp.

Valbuena, R., Mauro, F., Rodriguez-Solano, R., & Manzanera, J. A. (2010). Accuracy and precision of GPS receivers under forest canopies in a mountainous environment. *Spanish Journal of Agricultural Research, 8*(4), 1047–1057.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). NY, NY: Springer.