BIOREDD+ Program

DELIVERABLE 9: LIDAR ESTIMATION OF FOREST CARBON IN BIOREDD PROJECT AREAS

Subcontract number: EEP-I-00-06-00013-00-GEM-00
Name: GeoEcoMap Inc.

August 2014

This document was produced for review by the United States Agency for International Development. It was prepared by GeoEcoMap Inc. for the BIOREDD+ Project, task order number AID-514-TO-11-00002.

GeoEcoMap, Inc
1106 Hauser Blvd
Los Angeles, CA 90019

Table of Content

1. Statement of the Task
2. Background
3. Data Sources and Processing
4. Definitions
5. Lidar Height Measurements and Accuracy
6. Field Calibration and Validation Plots
7. Lidar Biomass Estimator
8. Lidar Biomass Estimation of BioREDD Project Areas
9. Products
10. Summary and Conclusion
11. References

GeoEcoMap, Inc
1106 Hauser Blvd
Los Angeles, CA 90019

1. Statement of task 8 and 9

Final report of Lidar estimation of forest aboveground carbon estimates over Lidar transects (49 sample areas of about 83,000 ha) over the project nodes derived from integration of field inventory and airborne Lidar data. The report shall include the following components: 1. raw Lidar data at 1 m spatial resolution for surface topography and vegetation height, 2. processed Lidar waveforms at the spatial resolution of 1-ha, 3. forest aboveground biomass estimation from Lidar data in raster and vector formats at 1-ha spatial resolution, and 4. uncertainty estimates of Lidar derived forest biomass. The report includes both the meta data and the documentations outlining the processing steps, methodology, and algorithms to convert Lidar data to forest aboveground biomass carbon density.

2. Background

We used the stratified map developed during the Task 1 of the project to design the lidar data acquisition and the field sampling plots. The stratification was developed by combining four data layers: climate, soil, topography, and the land cover map. We colored the stratified image into 30 distinct colors to demonstrate the potential variations of the landscape features that may influence the forest structure, carbon stock and dynamics. The stratified map has been used to define areas for forest inventory and airborne Lidar samples. We selected 49 locations randomly in 30 stratified regions within the original project areas. We designed the 49 lidar transects to cover approximately 83,000 ha of forests randomly selected over an area of approximately 1.7 million ha along the Pacific coastal forests of Colombia where the BioREDD project has been established. The number of transects and the size of each transect was designed to reduce the variance in biomass estimation by taking into account the errors in lidar estimates of biomass, allometry, sampling errors, and natural variations of biomass. The area sampled by lidar was > 1000 ha and several lidar transects were located within each strata depending on the size of the stratum. The over-sampling compared to the ground inventory was done intentionally to meet two key requirements: 1) sampling had to be random and large enough in each stratum to reduce the standard error in estimation. We aimed at < 10% error at 95% confidence interval. However with a large number of samples acquired by the lidar data, we achieved errors with approximately one order of magnitude better with lidar data. 2) The samples were clustered along the lidar flights and therefore required to be large enough to provide the balanced sampling to estimate the mean and variance of biomass in each stratum.

3. Data Sources & Processing

All raw data processing have been completed by the SAI-FAL, the contractor from Brazil in charge of data collection, raw data processing and classification of ground and tree height from lidar data. GeoEcoMap was involved in several steps of this process:

1. Design of lidar data acquisition as random balanced sampling on forest stratification within the project area.
2. Defining the requirements for the lidar data collection. We provided in our reports detailed information about the requirements for lidar point density, horizontal and

vertical accuracy, geolocation accuracy, ground and vegetation height classification accuracy.

3. We reviewed all products and provided inputs for improvements of lidar processing, ground classification, filtering of all anomalous points.
4. We performed post-processing steps in the delivered products by further analyzing the raw data, improving the ground classification over topographically complex regions, filtering anomalously high and low points, improving digital elevation model by interpolations, and filtering and improving the missing scan line and data points in the data in order to reduce the forest structure estimation.
5. All final datasets were post-processed to include forest height metrics at 50 m spatial scale for biomass estimation. The 50 m resolution was selected to match the lidar metrics with the ground calibration and validation plots.
6. Figure 1 shows the location of sample plot data in the BioREDD project areas with location and orientation of flights randomly selected.

Fig. 3.1. Location of Airborne Lidar transects along the Pacific Coast of Colombia over stratified areas and within BIoREDD project areas.

4. Definitions

The primary focus of this report is to document the accuracy of lidar measurements of forest height, ground measurements of forest structure and estimation of forest biomass for calibration and validation of the lidar data, the calibration of the lidar measurements into forest aboveground biomass using a lidar biomass estimator, the validation of the results using cross-validation, the uncertainty estimates required for lidar estimation of the forest biomass over project areas. All calculations are performed to meet the requirements of the VM0006 methodology. The quantities calcualted and addressed in this report are:

1. Aboveground Biomass in Mg/ha
2. Aboveground Biomass Carbon in MgC/ha
3. Tree biomass allometry for converting ground measurements of forest structure to biomass
4. Tree height and diameter allometry for estimating height from diameter for trees without height measurements.
5. Wood specific gravity or wood density for converting forest volume/structure to biomass
6. Lidar biomass allometry for converting lidar measurements at the plot level to forest biomass
7. Lidar biomass estimator for estimating forest aboveground biomass in stratified forests and in BioREDD project areas.
8. Mean and variance of biomass estimation for each forest type within the BioREDD project areas.

Table 3.1. Characteristics, source and the use of remote sensing data used for carbon assessment of BioREDD region.

Data Source	Main Use of Data	Data Characteristics
Airborne small footprint lidar data collected over the BIOREDD project areas based on a stratified random sampling as described in report for the task # 2. Total of 49 flight samples were used to cover the stratified landscape of the BioREDD regions randomly. Each flight line has coverage area of at least 1000 ha, providing both tree height and ground topography over more than 83,000 ha of forests. Data were collected by Union Temporal SAI-FAL, Brazilian Company with offices in Sao Paulo and Bogota. Lidar data were collected from July or August of 2013 to January 2014. All processing and ground and canopy classification were performed by SAI-FAL with direct consultations with GeoEcoMap, Inc.	Lidar data will be used to quantify several forest height metrics at 0.25 and 1.0 ha to use the forest aboveground biomass carbon stock. Lidar data are also used to develop fraction of forest canopy cover for separating areas of degraded and secondary forests from old growth forests.	The posting of DEM and top canopy height is at 1 m. Vertical accuracy of height is about 25 cm at 95% CI. Horizontal Accuracy < 50 cm at 95% CI Geometric Accuracy < 1 m using In-Fusion Precise Point Positioning (PPP) approach. Point density of > 4 points/m^2 was requested and point density of RMS > 10 points/m^2 was achieved over 67% of the coverage area. Digital Elevation Model was corrected at GeoEcoMap by improving the ground classification macro on a TerraScan platform.
Concurrent digital photography images using a 3-band RGB camera were acquired along with the Lidar data. RGB imagery were acquired over all 49 sampling units covering more than 83,000 ha area. All RGB images were orthorectified.	RGB aerial photo imagery are used to classify intact forests, degraded, secondary forests, croplands, and grasslands, and wetlands. The imagery provided data for calibration and training of the 30 m Landsat data for land cover and land use classification over the entire pacific regions.	Spatial Resolution: 20 cm. Images are mostly cloud free but have geometric problems associated with the look angle, sun glare, and haze in the imagery.
Ground forest inventory data collected within the lidar coverage over areas with permissions. The plot data are designed to have one 1-ha and four 0.25 ha plots located	Ground data are used to calibrate and validate lidar data. By estimating forest biomass carbon stocks and developing relations between lidar height metrics and forest biomass, all lidar data will be converted to carbon stocks.	16 1-ha (100 m x 100 m) plots 128 0.25 ha (50 m x 50 m) plots 45 0.25 ha plots in one lidar transect. Total of 15 cluster plots are in Lidar transects.

5. Lidar Height Measurements and Accuracy

This section will provide the basic information about the lidar measurements of forest height and accuracy at the plot level. We used several height metrics in developing the lidar biomass estimator. These included the percentile heights (RH10, ….RH100). However, we used the mean top canopy height (TCH) as the metric for biomass estimation. The processing provided the TCH for each 50

m area over all lidar data developing (83,000 x 4) sample area for biomass estimation. The processing steps to develop 0.25 ha scale height metrics from lidar data included:

1. *Evaluation of raw data and classification from lidar (examples of figures)*

We assessed the classification of the lidar raw data into ground and canopy using the point cloud data and las files. In Microstation (using the Terrascan/Terramodel extension) and ArcGIS 10.2 (using the LP360 extension) the las files were imported and the triangulated ground model was visualized to assess the quality of the client classification. If larger interpolations in the DEM are present, they are flagged and analyzed to determine whether the DEM could be improved via additional processing (figure 7.1).

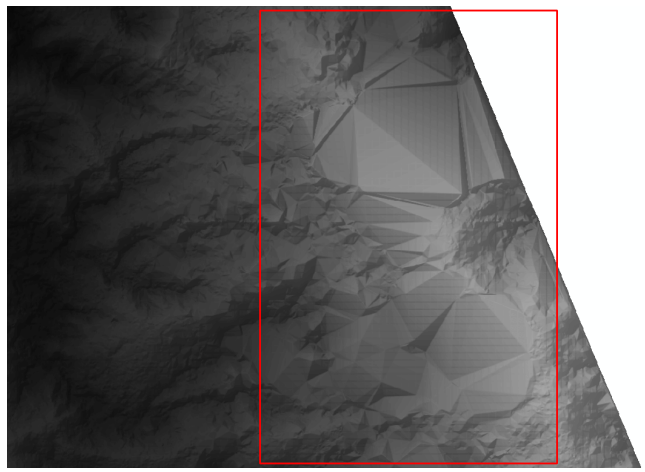
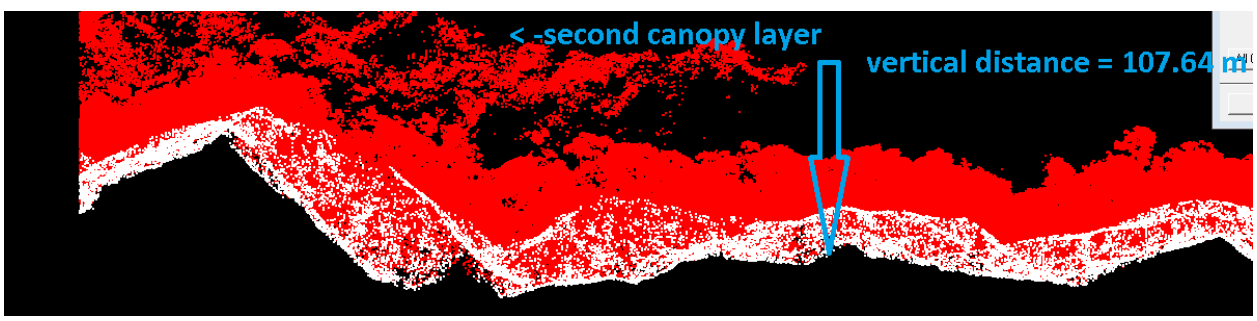


Fig. 5.1 Large interpolation errors in the DEM derived from lidar data.

2. *Filtering the high and low points*

In forested ecosystems the outlier values are generally high points caused by low clouds or birds or any objects on the path of the lidar and top canopy height that appears well above the top of the forest canopy. If the client classification does not re-classify these points as error points, we removed them manually so they don't provide large errors canopy height values. We re-classified all error points by first flagging them in ArcGIS and reclassifying them in Terrascan. Majority of lidar sample areas do not have these high and low points< However, when they are present, they are often sparsely scattered throughtout the lidar transects but are sparse and have very low impact on the overall accuracy of lidar top canopy estimates at the plot scale.



Fig, 5.2 High points in a second 'ghost' canopy layer.

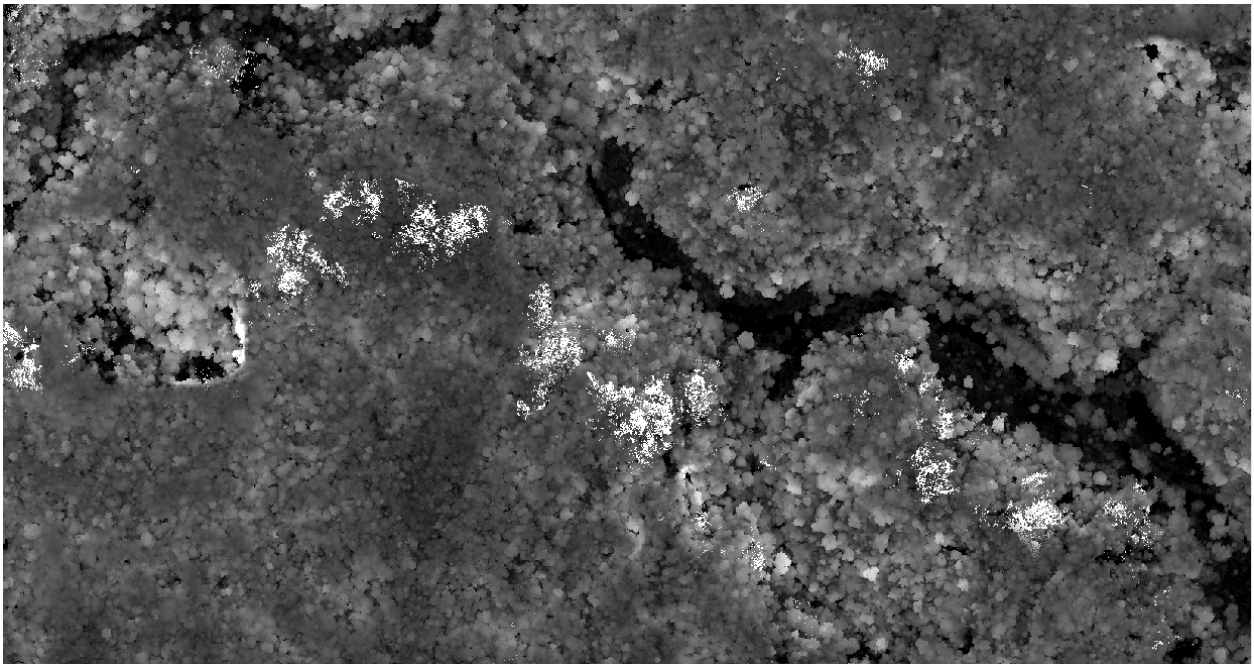


Fig. 5.3. High points caused by cloud strikes nd birds (white speckling points in throughout the middle of the image)

3. Processing the data in height/Aggregation of the data into 0.25 ha grids

The final two steps are executed in one program. After the classification has been finalized we output a 1 m resolution ground DEM and a canopy CSM from Terrascan in 1km^2 blocks from Terrascan in Arc grid format. Once all tiles have been output from Terrascan we process them in ArcGIS using python scripts. The python script executes the following commands:

Mosaic all tiles for the ground into a continuous DEM geotiff file

Mosaic all tiles for the canopy into a continuous CSM geotiff file

Outline the extent of the mosaicked rasters in a vector shapefile

Subtract the DEM from the CSM on a cell by cell basis to create a canopy height model (CHM) which represents the maximum canopy return in a 1-m resolution raster

Aggregation of the CHM to 50 m grid cells which calculate the mean of all the CHM cells

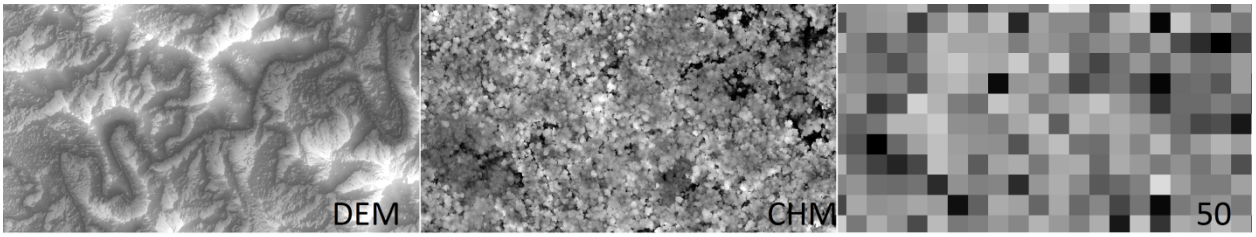


Fig. 5.4. Digital Elevation Model (left), Canopy Height Model (center) and CHM resampled at 50 m (right).

6. Field Calibration and Validation Plots

This section provides a basic description of biomass plot measurements, which are further described in the reports provided by CONIF and Alvaro Duque, for the Forest Carbon Monitoring System and calibration and validation of the Lidar data. For the BioREDD project areas, we established both permanent and temporary sampling plots to be used to estimate carbon stocks across the landscape and for implementation of future monitoring for detecting changes. The size and shape of the sample plots for tree biomass measurement is a trade-off between accuracy, precision, time, and cost for measurement, and finally the compatibility with the remote sensing data. The most appropriate size and shape may also be dependent on the vegetation type found in the sampling area. In addition, the plots must be large enough to reduce errors associated with geolocation and edge effects when used to calibrate the RS data. The minimum size of plots selected for this project is 0.25 ha (Asner and Mascaro, 2013; Meyer and Saatchi et al. 2013).

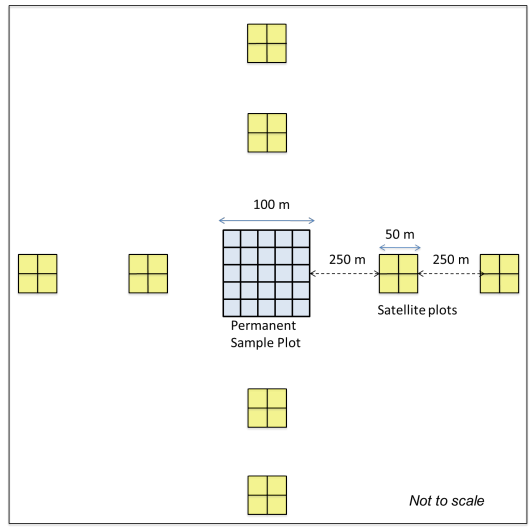


Fig. 6.1 Shape and size of permanent and satellite plots within the lidar transect.

The 0.25 ha size plot is a compromise between small plot sizes that are efficient in the field inventory approaches (e.g. < 0.1 ha) but provide models with large uncertainty between RS and ground estimates of AGB, and large plots (e.g. 1.0 ha) that are costly and difficult to establish in the field but provide models with the low unceratinty (~ 10% relative error and no bias) (Meyer et al., 2013; Asner and Mascaro, 2014).

In the BioREDD project, the plots are only used for calibration and validation of lidar to AGB model. The minimum number of plots required for developing and validating the remote sensing to AGB predictive model was estimated to be 45, with 30 plots for calibration and 15 plots for validation. This requirements were based on a similar methodology as in CDM A/R tools for Calculation of the number of sample plots for measurements within A/R CDM project activities, and based on minimum number for developing remote sensing and AGB predictive model (Asner and Mascaro, 2014). In general one predictive model is enough to estimate forest biomass from lidar in BioREDD regions if the ground allometric model used in estimating AGB at the plot level is also one. If the number of ground allometry changes in the project area because of the presence of various forest types, as in Chave et al., 2005 for wet, moist, and dry forests, then the number of remote sensing predictive models will also change.

In the BioREDD project we increased the number of calibration and validation plots in order to make sure that the numbers are enough for developing more than one lidar to AGB predictive model if necessary. A total 15 cluster plots were used in 15 lidar transects covering different strata in the BioREDD regions. Each cluster plot met the following requirements:

1. Each cluster plot had one permanent plot of 1-ha in size (100 m x 100 m) and 8 satellite or auxiliary plots at 0.25 ha in size (50 m x 50 m).
2. The center location of permanent plot was selected randomly in the lidar transect in the forest area to allow sampling all forest conditions as in fragmented, degraded or secondary forests. The satellite plots were located at 250 m and 500 m away from

the center permanent plots in four cardinal directions (Fig. 6.1).

3. The measurements in the permanent plots followed the RAINFOR protocols (REF?) All trees with DBH ≥ 10 cm in diameter were measured in each plot, with trees being tagged and recorded, x,y coordinates were recorded for permanent plots but not the satellite plots, and trees were identified in all plots by species for wood density values.
4. A minimum number of 50 trees in different DBH classes were selected for height measurements and development of DBH-H allometry for the plots or forest type. The number of trees with height measurement represented between 7 and 16% of the total number of trees registered in each site. All trees in San Pablo already had height measurements.
5. A tree-diameter height (H-D) relationship or predictive model was developed for each cluster plot and also combined for the region. Fig.6.2 shows the allometry used to convert tree diameter to height for all the trees in the plots. We used a Weibull model to estimate the height of trees that were not measured directly:

$$H = a[1 - \exp(-bD^c)]$$

where H is the tree height of individual trees measured in the field, and D is diameter at breast height of these trees. The coefficients a, b, and c were estimated from the measured diameter and height and provided for the regional model as:

$$H = 40.51\left[1 - \exp(-0.0599D^{0.7096})\right]: \text{ For all 15 random plots}$$

$$H = 39.79\left[1 - \exp(-0.0675D^{0.7458})\right]: \text{ For 45 systematic plots}$$

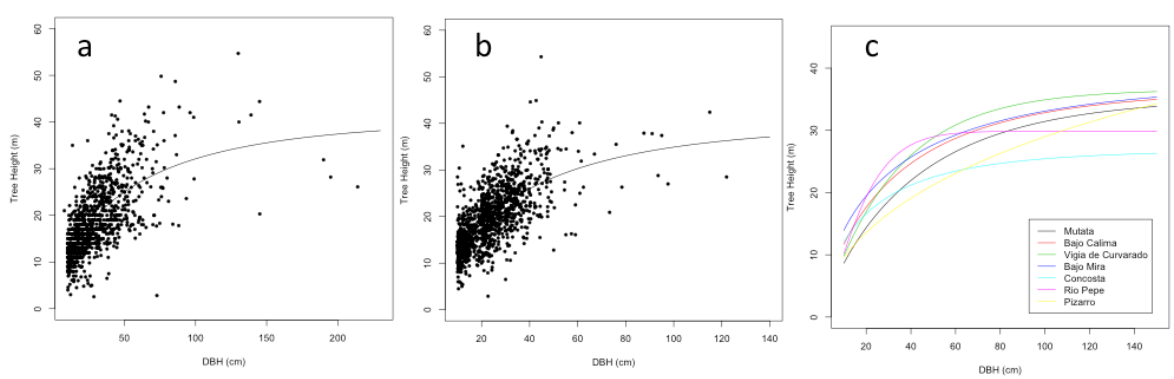


Fig. 6.2. Diameter-height models developed from the tree height measurements in the sample plots with (a) showing all data used for a regional model using the random cluster plots, (b) showing the model for 45 systematic plots, and (c) showing individual models used for different project areas. Models used in figure a and b are approximately the same and do not produce significantly differnet biomass results.

6. Each tree identified in the plot was given a wood density based on the Global Wood Density Database for South America. For trees with not identification, we used the average wood density of the plot for biomass calculation. During the BioREDD project, we also measurement wood density of a selected number of species at the branch and stem levels for improving the wood density values for the region. The average plot level wood density varied in the BioREDD region suggesting the heterogeneity of wood density at the landscape scale in the Pacific coastal forests of Colombia.

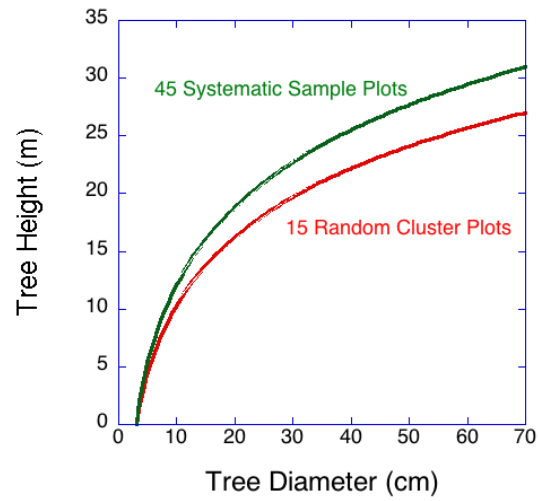


Fig. 6.3. Comparison of two H-D models for the Choco forest types derived from 15 1-ha random plots and 45 0.25 ha systematic plots.

Table 6.1. Average wood density of forests in BioREDD regions from trees collected in random cluster plots.

BioREDD Project Region	Average Wood Density
Acapa	0.658
Bahia_Malaga	0.551
Bajo_Calima_I	0.562
Bajo_Calima_II	0.593
Bajo_Mira	0.569
Buenavista	0.675
Canton_S_Pablo	0.601
Chigorodo	0.507
Chontadural	0.503
Concosta	0.636
Curvarado	0.533
Pizarro	0.518
Rio_Cajambre	0.494
Rio_Pepe_I	0.487
Rio_Pepe_II	0.524

For AGB estimations of sample plots, we used the allometric model developed for the study region by the BioREDD team. First we used the Chave et al. 2005 model for moist forests to estimate the biomass for each plots. However, the BioREDD project provided funding to develop local allometric equations for the trees of the Pacific coastal forests. The models developed for the region

included three types of forests in the regions: Terra firme (Colinas), Inundated forests (Guandal), and mangrove forests (Manglar). Trees were harvested in the region and combined by harvested trees from past research activities to develop the models. The models included uncertainty analysis and comparison with Chave et al. (2005) and (2014) allometric equations and differences were reported (see BioREDD project report BR-SUBK-FP-022). The equations are summarized below:

Colinas $\ln(AGB) = -2.130+2.015\times\ln(D)+0.724\times\ln(H)+1.002\times\ln(WD)$

Guandal $\ln(AGB) = -2.328+1.833\times\ln(D)+0.724\times\ln(H)+0.151\times\ln(WD)$

Manglar $\ln(AGB) = -2.818+2.185\times\ln(D)+0.724\times\ln(H)+0.650\times\ln(WD)$

For estimation of the biomass of palm trees in all types of forests, we used the allometric equation developed by Saldarriaga (2014).

$$\ln(AGB) = -0.173+0.700\times\ln(D^2\times H\times WD)$$

For biomass values estimated in individual permanent and satellite plots see table 6.3. For each plot, we also have an indentifier and discription by the field group that summarizes the type an condition of the forests such as degraded, secondary forests, Guandal, Colinas, Manglar, etc. Most plots fall in mixed forests, therefore the biomass values estimated from allometric models are applied on the tree level to provide estimates of the biomass at the plot level.

Table 6.2. Mean and standard deviation of forest biomass of 15 cluster plots within the BioREDD project areas.

Name of Culster Plots	Number of WD samples	Average WD of Trees	Average WD of Trees and Palms	Average number of trees per ha	AGB (Mg ha^{-1})
Acapa	30	0.658	0.630	357	144.5 ± 50.6
Bahia_Malaga	81	0.551	0.531	636	217.6 ± 65.9
Bajo_Calima_I	100	0.562	0.537	694	154.9 ± 56.6
Bajo_Calima_II	70	0.593	0.563	634	170.5 ± 37.4
Bajo_Mira	30	0.569	0.543	460	105.9 ± 24.6
Buenavista	73	0.675	0.657	505	304.7 ± 40.0
Canton_S_Pablo	38	0.601	0.588	569	214.0 ± 51.7
Chigorodo	98	0.507	0.505	485	260.8 ± 95.4
Chontadural	168	0.503	0.480	538	207.4 ± 29.6
Concosta	84	0.636	0.606	612	364.5 ± 472.0
Curvarado	73	0.533	0.522	442	342.7 ± 157.0
Pizarro	81	0.518	0.495	396	162.1 ± 51.6
Rio_Cajambre	31	0.494	0.457	1300	125.6 ± 39.0
Rio_Pepe_I	160	0.487	0.476	538	167.1 ± 31.8
Rio_Pepe_II	160	0.524	0.502	769	215.9 ± 60.6

In addition to above plots, we also developed a set of 45 plots, each 0.25 ha in size for spatial uncertainty analysis of the lidar estimation of biomass. We selected one lidar transect with

relatively easy access and with no previous plots and developed 45 systematically located plots (Fig. 6.3). Number of plots to estimate mean biomass density at 10% error with 95% confidence was calculated to be 45 0.25 ha plots. We calcualted the mean and the standard error by using a lidar biomass allometry model developed for the region, predicting the mean and variance and using the variance to estimate the number of plots required for the ground measurements.

The sampling design followed the standard methodology as in CDM AR tools for the number of sample plots as in:

$$n = \left(\frac{1.96\sigma}{\varepsilon}\right)^2$$

n : Number of plots to estimate mean biomass

1.96: Coefficient for 95%confidence interval

σ: Standard Deviation of Distribution

ε: Error

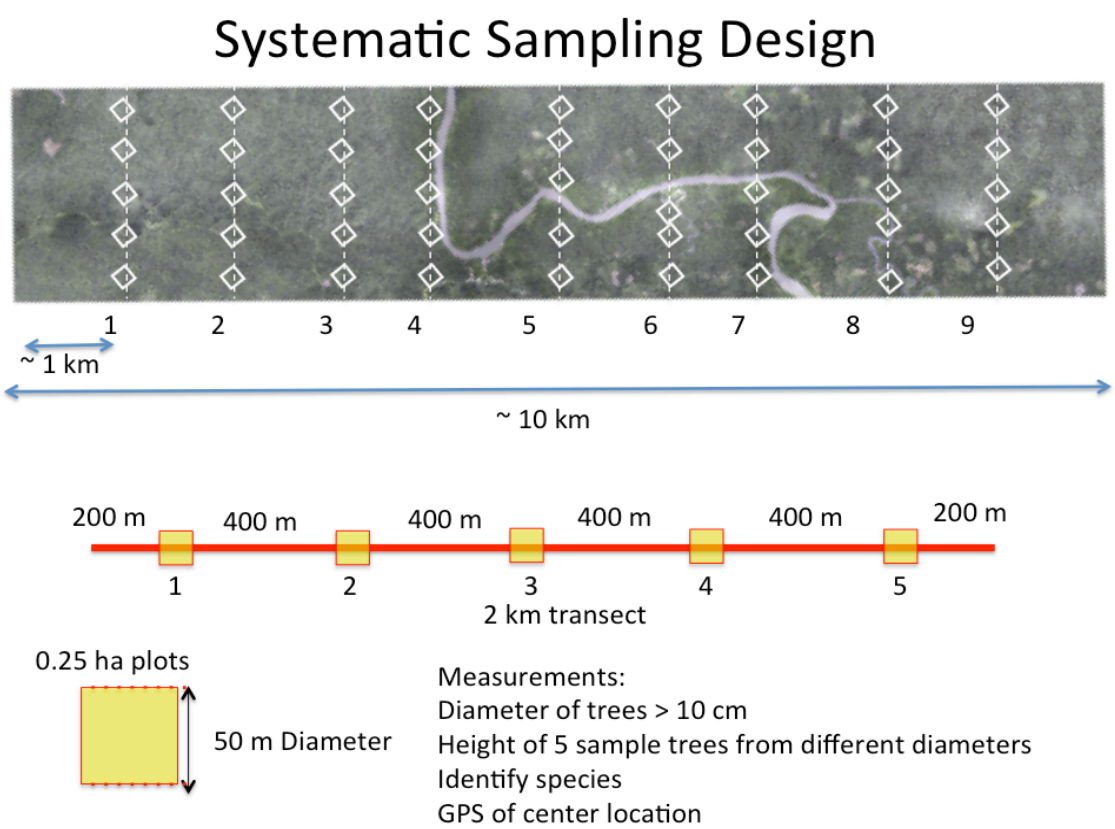


Fig. 6.3. Schematic of the systematic sampling design of a lidar transect for quantifying the spatial uncertainty of lidar estimation of bimass using lidar to AGB predictive model.

Table 6.2. Forest biomass estimated for all permanent and satellite plots in the BioREDD project regions.

plot name	plot type	AGB (Mg ha)	plot name	plot type	AGB (Mg ha)	plot name	plot type	AGB (Mg ha)
Acapa	Permanent	175.88	Bajo Mira	Satellite1	144.56	Curvarado	Satellite1	344.22
Bahia Malaga	Permanent	246.24	Bajo Mira	Satellite2	100.40	Curvarado	Satellite2	510.91
Bajo Calima I	Permanent	155.32	Bajo Mira	Satellite3	121.37	Curvarado	Satellite3	178.86
Bajo Calima II	Permanent	190.77	Bajo Mira	Satellite4	211.96	Curvarado	Satellite4	179.75
Bajo Mira	Permanent	126.67	Bajo Mira	Satellite5	133.43	Curvarado	Satellite5	245.48
Buenavista	Permanent	274.50	Bajo Mira	Satellite6	148.13	Curvarado	Satellite6	493.72
Canton S Pablo	Permanent	226.65	Bajo Mira	Satellite7	154.80	Curvarado	Satellite7	158.76
Chigorodo	Permanent	216.43	Bajo Mira	Satellite8	164.30	Curvarado	Satellite8	510.61
Chontadural	Permanent	206.41	Buenavista	Satellite1	346.75	Pizarro	Satellite1	189.55
Concosta	Permanent	259.92	Buenavista	Satellite2	268.30	Pizarro	Satellite2	161.38
Curvarado	Permanent	411.27	Buenavista	Satellite3	229.32	Pizarro	Satellite3	141.34
Pizarro	Permanent	239.12	Buenavista	Satellite4	296.95	Pizarro	Satellite4	215.58
Rio Cajambre	Permanent	195.28	Buenavista	Satellite5	341.26	Pizarro	Satellite5	95.71
Rio Pepe I	Permanent	126.72	Buenavista	Satellite6	321.43	Pizarro	Satellite6	92.23
Rio Pepe II	Permanent	180.27	Buenavista	Satellite7	301.12	Pizarro	Satellite7	135.64
Acapa	Satellite1	137.40	Buenavista	Satellite8	336.38	Pizarro	Satellite8	195.64
Acapa	Satellite2	173.81	Canton S Pablo	Satellite1	211.72	Rio Cajambre	Satellite1	73.20
Acapa	Satellite3	124.65	Canton S Pablo	Satellite2	265.05	Rio Cajambre	Satellite2	93.26
Acapa	Satellite4	192.74	Canton S Pablo	Satellite3	272.93	Rio Cajambre	Satellite3	204.89
Acapa	Satellite5	51.21	Canton S Pablo	Satellite4	209.51	Rio Cajambre	Satellite4	124.43
Acapa	Satellite7	97.14	Canton S Pablo	Satellite5	244.62	Rio Cajambre	Satellite5	167.41
Acapa	Satellite8	135.03	Canton S Pablo	Satellite6	194.91	Rio Cajambre	Satellite6	187.39
Bahia Malaga	Satellite1	253.38	Canton S Pablo	Satellite7	97.56	Rio Cajambre	Satellite7	175.45
Bahia Malaga	Satellite2	330.53	Canton S Pablo	Satellite8	201.09	Rio Cajambre	Satellite8	173.85
Bahia Malaga	Satellite3	187.42	Chigorodo	Satellite1	337.02	Rio Pepe I	Satellite1	172.20
Bahia Malaga	Satellite4	141.16	Chigorodo	Satellite3	259.22	Rio Pepe I	Satellite2	136.07
Bahia Malaga	Satellite5	188.53	Chigorodo	Satellite4	351.37	Rio Pepe I	Satellite3	187.62
Bahia Malaga	Satellite6	283.75	Chigorodo	Satellite5	202.34	Rio Pepe I	Satellite4	221.66
Bahia Malaga	Satellite7	156.72	Chigorodo	Satellite6	404.31	Rio Pepe I	Satellite5	199.23
Bahia Malaga	Satellite8	167.32	Chigorodo	Satellite7	156.81	Rio Pepe I	Satellite6	182.15
Bajo Calima I	Satellite1	283.62	Chigorodo	Satellite8	157.21	Rio Pepe I	Satellite7	149.62
Bajo Calima I	Satellite2	171.65	Chontadural	Satellite1	218.40	Rio Pepe I	Satellite8	143.69
Bajo Calima I	Satellite3	156.04	Chontadural	Satellite2	224.92	Rio Pepe II	Satellite1	291.16
Bajo Calima I	Satellite4	168.97	Chontadural	Satellite3	199.49	Rio Pepe II	Satellite2	294.98
Bajo Calima I	Satellite5	138.97	Chontadural	Satellite4	250.25	Rio Pepe II	Satellite3	149.82
Bajo Calima I	Satellite6	105.63	Chontadural	Satellite5	146.51	Rio Pepe II	Satellite4	240.97
Bajo Calima I	Satellite7	129.09	Chontadural	Satellite6	206.46	Rio Pepe II	Satellite5	112.91
Bajo Calima I	Satellite8	87.09	Chontadural	Satellite7	217.88	Rio Pepe II	Satellite6	223.10
Bajo Calima II	Satellite1	230.76	Concosta	Satellite1	177.43	Rio Pepe II	Satellite7	240.50
Bajo Calima II	Satellite2	167.13	Concosta	Satellite2	314.75	Rio Pepe II	Satellite8	218.70
Bajo Calima II	Satellite3	162.05	Concosta	Satellite3	211.06			
Bajo Calima II	Satellite4	160.19	Concosta	Satellite4	232.79			
Bajo Calima II	Satellite5	135.05	Concosta	Satellite5	182.53			
Bajo Calima II	Satellite6	160.18	Concosta	Satellite6	118.76			
Bajo Calima II	Satellite7	214.66	Concosta	Satellite7	162.10			
Bajo Calima II	Satellite8	113.48	Concosta	Satellite8	1574.12			

Table 6.3. Estimated biomass for 45 sample plots used for lidar spatial uncertainty analysis.

parcela	AGB (Mg ha)	parcela	AGB (Mg ha)	parcela	AGB (Mg ha)
1	281.87253	16	162.7849	31	262.17065
2	322.51324	17	171.31248	32	215.96968
3	170.48509	18	138.61664	33	227.55025
4	193.61127	19	196.20612	34	156.77078
5	240.94267	20	133.91098	35	150.92538
6	281.30362	21	261.47452	36	322.06229
7	224.39587	22	225.94531	37	276.3736
8	165.81203	23	158.71066	38	235.0863
9	195.56256	24	218.24144	39	143.73756
10	204.08147	25	171.15575	40	367.83783
11	253.41548	26	269.78772	41	374.5163
12	272.69293	27	193.27368	42	350.38623
13	376.7012	28	228.84946	43	222.2756
14	187.66628	29	226.4919	44	286.888
15	214.20047	30	210.44151	45	162.85114

7. Lidar Biomass Estimator

In this section, we sumamrize: 1) the Lidar estimator or predictive model, 2) estimation of biomass from all 83,000 ha lidar data, and 3) estimation of mean and variance of lidar derived AGB for each stratum and land cover type in the BioREDD project areas. In the next section, we will provide the steps taken to insure the uncertainty in estimates are rigourous and valid for the entire project area. The sampling approach of the lidar data is summarized in Fig. 7.1

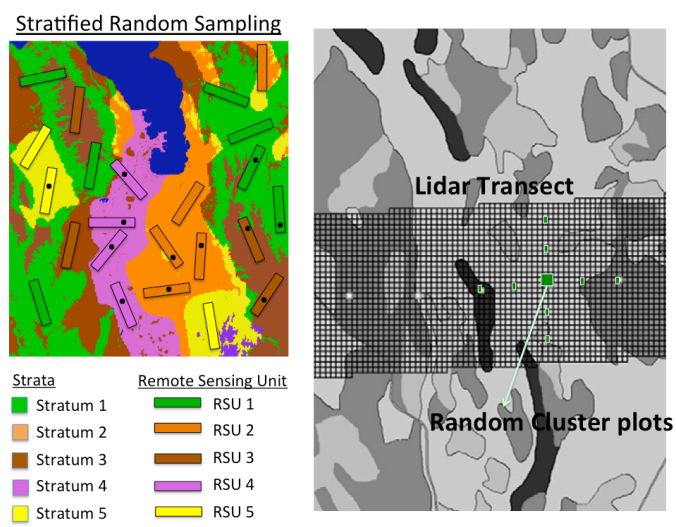


Fig. 7.1 Schematic of lidar random sampling of stratified landscapes over the BioREDD region on the left panel. The remote sensing units (RSU) are the lidar transect flight lines. On the right panel the lidar transect with pixels representing 0.25 ha (50 m x 50 m) aggregated from 1-m lidar data with location of sample plots for calibration and validation of lidar datais shown.

Lidar Biomass Predictive model

From lidar data we derived various height metrics to examine the accuracy of the best relationship for aboveground biomass estimation. These metrics included the percentile of energy from lidar data using the point cloud information or only the top canopy height for each lidar pixel (1 m). As

shown in various earlier work, the top canopy height at the native pixel level of lidar is the most reliable metric to use for biomass estimation (Meyer et al. 2013; Asner and Mascaro, 2014). The quality of point cloud data is highly variable over individual lidar transects or between transects if data were taken at different dates with different lidar flight configuration or environmental conditions such as canopy moisture because of the impacts on NIR (Near Infrared Lidar) signal penetration into the canopy.

Using the top canopy height at each pixel as the basic measurement of lidar data, we developed several metrics such as mean top canopy height (TCH) and 25, 50, 75, and 100 (maximum height) percentiles of 2500 points at 50 m x 50 m scale. We tested three types of equations for evaluating the contribution of these metrics for biomass estimation:

$$AGB = ah_{TCH}^{b} + \varepsilon$$

$$AGB = a_0 + a_1 h_{25p}^{b_1} + a_2 h_{50p}^{b_2} + a_3 h_{75p}^{b_3} + a_4 h_{100p}^{b_4} + \varepsilon$$

$$AGB = a_0 + a_1 h_{25p} + a_2 h_{50p} + a_3 h_{75p} + a_4 h_{100p} + \varepsilon$$

Where h_{TCH} represents the mean top canopy height, and the metrics are 25, 50, 75, and 100 percentiles. In all equations $\epsilon \sim N(0, \sigma^2)$ represents the uncertainty in measurements or when observations deviate from a power-law by accounting for the heteroskedasticity. After testing all three equations, we arrived to the conclusion that the first equation using the simle mean top canopy height (TCH) is the most reliable model as the other models, although slightly improve the coefficient of derminiation with AGB and the RMSE error of estimation, the are more subject to errors and sometimes require a higher number of plot data for robust estimations of model coefficients. The TCH model has only two coefficients (a,b) and can be readily developed using a minimum number of 30 points or ground plots.

We then tested the following two models in terms of their accuracy in estimating biomass in different forest types.

$$AGB = a(WDh_{TCH})^{b} + \varepsilon$$

$$AGB = ah_{TCH}^{b} + \varepsilon$$

where WD represents the average WD density at the plot level added as linear term to improve the estimation of AGB over areas with different WD. We have already calculated the average WD of each plot within the cluster plots (permanent and satellites) and also calculated estimates of mean WD for forest type identified during the ground measurements and comparison of coordinates with the land cover map of the region. The advantage of model 1 withiout WD component is its straightforward application with the lidar data. Model 2 with WD requires that the regional or landscape variations of

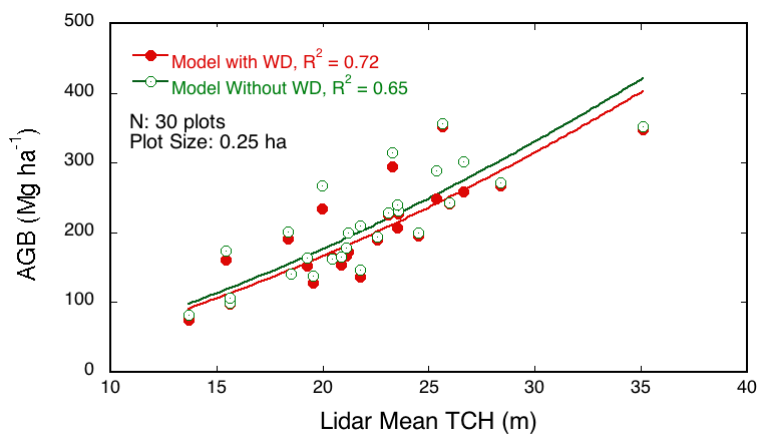


Fig. 7.2 Lidar biomass model derived from mean top canopy height (TCH) and aboveground biomass estimated from ground data. The

WD is known. Using the above two equations, we tested the goodness to fit of both models to the plot data evaluated by AIC (Akaike Information Criterion) . The AIC is a measure of the goodness-of-fit that penalizes parameter-rich models, as required by the principle of parsimony (Burnham & Anderson,2002). The results as shown below:

$$AGB = 1.484(WDh_{TCH})^{1.575} \quad (AIC = 2780, \sigma = 0.32)$$
$$AGB = 1.707h_{TCH}^{1.548} \quad (AIC = 2850, \sigma = 0.34)$$

The model with WD performed slightly better than the model without WD (higher AIC and σ). We removed the plots with significant numbe of mangrove trees and Guandal, we noticed that the models were not significantly different. We concluded that the model with WD performs better than the model without WD only in forests with mangrove trees and in inundated forests (guandal) by improving the RMSE of estimation by approximately 10%. Figure 7.2 shows the difference between the models and the improvement.
The uncertainty of the above models have been evaluated using cross validation approach and using plots not used in developing the model. In the above equation we only used 30 plots to develop the model in order to demonstrate how the methodology works with the minimum number of plots. Overall, we have the following number of plots to evaluate the lidar model: 1) 15 1-ha permanent plots within the lidar transects. We divided each plot to four 0.25 ha plots to increase the number of plots for calibration and validation. This resulted to 60 plots at 0.25ha, 2) we have 15 sets of satellite plots falling in the lidar transects with each set having 8 satellite plots at 0.25 ha for a total number of 120 with 109 plots accuractely located in lidar images, and 3) we collected 45 plots in 0.25 ha systematically within one lidar transect. Overall, we had 214 plots at 0.25 ha in size to develop and test the accuracy of lidar predictive model.

Table 7.1. Number of plots established in the BioREDD project and used in developing and testing the accuracy of Lidar biomass model.

Forest inventory plots	**Size and shape**	**Number of plots in Lidar transects**	**Number of plots at 0.25 ha size**	**Number of 0.25 ha in lidar transects**
16 permanent plots	1-ha, 100 m x 100 m	15	64	60
16 cluster of satellite plots	0.25 ha (50 m x 50 m)	15	128	109
1 set of systematic sampling plots	0.25 ha (50 m x 50 m)	45	45	45
Total	**-**	**75**	**237**	**214**

We use the following steps to evaluate the performance of the model or the lidar estimator on the above sample points. We use general statistical terms in representing the errors: 1) the mean squared error (MSE) of an estimator measures the average of the squares of the "errors", that is, the difference between the estimator and what is estimated. The MSE is the second moment (about the origin) of the error, and thus incorporates both the variance of the estimator and its bias. For an

unbiased estimator, the MSE is the variance of the estimator. Like the variance, MSE has the same units of measurement as the square of the quantity being estimated.

Statistical Analysis

The statistical analysis includes evaluating the performance of the model selected above based on regressing a dependent variable (AGB) against one or several independent variables (in our case, WD and H). We follow the general form of the model as shown above by assuming ε as an error term as a normally distributed with zero mean and SD of σ. If the model as the one selected above as p parameters (p=2 for WD and h), then the σ is defined as:

$$AGB = a(WDh_{TCH})^b + \varepsilon$$

$$\ln(AGB) = a + b\ln(WDh_{TCH}) + \varepsilon$$

$$RSE = \sigma = \sqrt{\frac{1}{N-p}\sum_{i=1}^{N}\varepsilon_i^2}$$

$$N(\varepsilon) = N(0,\sigma^2): \quad \text{Distirubtion of errors}$$

The model as shown above can be linearized in the ln form to simplify the model as a linear regression model. The model be used to estimate AGB from the parameters developed at the plot level, i.e. average wood density *WD (g cm^{-3}) and h_{TCH} (m)*. The estimated value of AGB can be written as:

$$AGB_{est} = \overline{\exp\left[a + b\ln(WDh_{TCH}) + \varepsilon\right]} = \overline{\exp(\varepsilon)} \times \exp\left[a + b\ln(WDh_{TCH})\right]$$

where

$$\overline{\exp(\varepsilon)} = \int \exp(\varepsilon)N(\varepsilon)\,d\varepsilon \quad \text{with } N(\varepsilon) = N(0,\sigma^2)$$

$$\overline{\exp(\varepsilon)} = \exp(\sigma^2/2)$$

then

$$AGB_{est} = \exp\left[\sigma^2/2 + a + b\ln(WDh_{TCH})\right]$$

The last equation provides the unbiased estimator for AGB using the height and wood density. To examine the model, we used the plot data and calculated the average systematic error (bias) and the coefficient of variation (CV) as follo

$$bias = \frac{1}{N}\sum_{i=1}^{N}\frac{(AGB_{est}(i) - AGB_{obs}(i))}{AGB_{obs}(i)}$$

$$RSE = \sqrt{\frac{1}{N-P}\sum_{i=1}^{N}\left[AGB_{est}(i) - AGB_{obs}(i)\right]^2}$$

$$MAGB = \frac{1}{N}\sum_{i=1}^{N}AGB_{obs}$$

$$CV = \frac{RSE}{MAGB}$$

where RSE is the residual standard error representing the random errors. The standard deviation of estimation error can be computed as: $SD = \sqrt{RSE^2 - bias^2}$.

The table below provides the model fit and the accuracy of the model in estimating the AGB. The bias can be readily excluded in the estimator. From the table, it appears that having a minimum of 30 plots for calibration and 15 plots for validation will be enough to acheive at the required lidar biomass predictor model and uncertainty assessment. More plots provides only slighty improved results. Note that here, we used one allometric model for all forests. In the case, more allometric models, for each model, a minimum of 30 plots are required for calibration.

Table 7.2. Lidar biomass predictive model Model goodness of fit and uncertainty assessment using the plot data available for the BioREDD project.

Calibration Sample Size	Model fit Assessments			Validation Sample Size	Cross Validation Results		
	σ	AIC	DF		RSE	Bias	CV
30	0.329	47	3989	15	37. 96	2.27%	18.8%
50	0.326	45	3990	25	35. 24	2.01%	16.2%
100	0.319	45	3980	50	33. 96	2.20%	15.8%
150	0.311	43	3980	75	31. 96	1.98%	14.8%

8. Lidar Biomass Estimation of BioREDD Project Areas

The lidar data acquired over the BioREDD region followed a stratified random sampling approach where the remote sensing sampling units (RSU) are the flight lines of lidar with a minimum area coverage. Each flight line has a given swath with regular grid cells of 0.25 ha with biomass estimations that are considered the population elements. Flight lines or sampling units are designed to provide a balanced random sample of stratified classes and therefore for each strata j there are η_j sampling units and within each sampling units there are η_{ij} lidar grid cells. The lidar grid cells within each flight line are considered clustered and the estimator for the mean and variance of

biomass density for each stratum are given as follows (Naesset et al., 2013; Neigh et al., 2013; Stahl et al., 2011).

$$\mu_j = \frac{\frac{1}{n_j}\sum_{i=1}^{n_j} F_{ij}(\alpha_j)}{\frac{1}{n_j}\sum_{i=1}^{n_j} \eta_{ij}}$$

$$\mathrm{var}(\mu_j) = \frac{1}{\bar{\eta}_j^2}\frac{\sum_{i=1}^{n}\left(F_{ij}(\alpha_j) - \mu_j \eta_{ij}\right)^2}{n(n-1)} + \frac{1}{\bar{\eta}_j^2}\sum_{k_1}^{p_j}\sum_{k_2}^{p_j} Cov(\alpha_{k_1 j}, \alpha_{k2j})\bar{F}'_{k_1 j}\bar{F}'_{k2j}$$

where

$$\bar{F}'_{k_1 j} = \sum_{i=1}^{n}\sum_{t=1}^{T}\frac{\partial f(x_{it}, \alpha_{k_1})}{\partial \alpha_{k_1}}$$

where

$$AGB = f(x, \alpha, \varepsilon) = a(WDh_{TCH})^b + \varepsilon$$

where μ_j is the mean carbon density of stratum j, F_{ij} is the carbon density estimates for sample i in stratum j, n_j is the number of RSU intersecting stratum j, α_j represents the vector of parameters used in the biomass model (e.g. a and b in lidar model in above AGB equation, and η_{ij} is the number of lidar biomass units (cluster size) in each sample unit i intersecting stratum j. Note that if lidar biomass estimates are at 1-ha units, η_{ij} represents the size of the lidar transect or RSU in ha. In above equation, $cov(\alpha_{k_1 j}, \alpha_{k_2 j})$ is the covariance of k_1 and k_2 coefficients of the lidar biomass predictive model, represented by function *f* for stratum j. The first term in equation in variance estimate represents the sampling error and the second term describes the model error for each stratum. The above relations will be simplified if only one lidar model is used for all strata.

The lidar model-assisted estimator of biomass (carbon density) is approximately a design-unbiased estimator irrespective of the model of choice when the number of samples collected in each RSU is large (which is the case in airborne lidar collection). The design is unbiased because within each stratum, ground plots are used to make the lidar predictor model unbiased (as discussed above) (Naesset et al., 2011). The variance estimator also propagates the error from the lidar predictive model for estimation at the statra and land cover class scales.

We collected 45 ground samples in a lidar sampling unit or transect to also estimate if there has been any bias in estimation of biomass at scale of sampling unit (lidar flight transect). The mean for each sample unit can be readily rewritten as:

$$\mu = \frac{\sum_j \hat{\mu}_j}{N} + \frac{\sum_j \hat{e}_j}{n}$$

$$\hat{e}_j = \mu_j - \hat{\mu}_j$$

where $\hat{e}_j$ is the difference between ground estimated biomass μ_j and lidar estimated biomass $\hat{\mu}_j$ using the lidar model. If the number of plots n used for estimating $\hat{e}_j$ is large (> 10) (Särndal, 1984).

Having more than 30 ground plots will allow us to remove the bias in the system and have an unbiased estimate of mean and standard error.

Using the above relations, we followed the following steps to estimate the mean and variance of biomass density in each strata and also based on the land cover types used in the VM0006 methodology.

We used the following steps to estimate the aboveground biomass for the BioREDD regions at the scale of strata, LULC, and project areas:

1. The unbiased lidar predictive model was used to estimate the biomass for each grid cell within the lidar sampling units. Since the grid cells are 0.25 ha each, we developed approximately 332,000 (i.e. 4 x 83,000) estimates at grid cells. Each grid cell has a lattitude and longitude identifier that can be traced to a location within each strata in the BioREDD project region. Each estimate has an error associated with it according to relations provided in the previous section.
2. The grid point estimates and the form of the mean and variance estimator shown in this section are used to estimate the mean biomass in each strata and the BioREDD project area by using the proportion area of each stratum in the project area as given by:

   $$\hat{\mu}_{PA} = \frac{1}{PA}\sum_{j=1}^{m} w_j \mu_j$$

   where w_j is the proportion area of each stratum in PA (Project Area)

   μ_j is the biomass density for stratum j

   m is the number of strata in the project area
3. For the LULC classes, we use the same methodology as the strata and replace the stratification by LULC map and use the lidar grid cells as samples for estimating biomass mean and variance of LULC class types.

9. Products

The products are provided as biomass carbon estimates and uncertainty for strata, selected BioREDD project areas, and LULC class types in each BioREDD project regions. We also provide the forest carbon density for all 28 initial BioREDD region.

Table 9.1. Mean and standard error of forest above ground biomass carbon density in each stratum using lidar derived biomass values at the pixel scale (0.25 ha) over the areas intersected by each stratrum. In this analysis only stratified areas associated with the forest areas are considered.

Stratum #	# of pixels	# of intersected RSU	Mean AGB (Mg ha^{-1})	Standard Error (Mg ha^{-1})	Min AGB (Mg ha^{-1})	Max AGB (Mg ha^{-1})
1	1148	3	220.6	1.97	31.36	458.21
4	2191	3	47.379	0.422	29.75	201.1
5	139	1	113.26	5.98	30.31	336.9
6	341	1	66.047	1.87	29.8	208.88
8	4298	4	112.42	0.884	29.77	386.24
11	18989	7	245.02	0.658	29.82	634
12	10563	8	160.44	0.584	29.93	507.13
13	4589	5	123.71	0.657	29.78	324.07
14	14320	6	179.54	0.838	29.78	657.16
15	7117	5	129.54	0.783	29.74	494.08
16	399	1	187.71	2.66	37.51	389.67
17	49115	7	208.48	0.303	29.74	516.06
18	18431	5	179.99	0.473	29.77	548.24
19	25937	8	198.84	0.436	29.76	541.89
20	14882	5	161.1	0.546	29.8	493.17
21	1464	2	140.47	1.9	29.79	396.26
22	12088	4	132.33	0.587	29.75	466.73
24	3384	2	246.15	1.22	29.75	495.32
27	3354	1	246.71	1.27	30.34	573.5
30	286	1	192.87	2.59	92.99	472.91
31	10350	4	245.2	0.731	30.18	625.28
32	5121	3	271.52	0.889	30.03	504.83
34	3013	2	242.08	1.68	29.76	541.2
38	14489	7	250.76	0.44	30.61	560.31
39	1336	1	248.83	1.34	35.79	443.54
41	3818	3	216.53	1.06	30	712.54
42	358	1	250.87	2.66	122.4	457.4
44	4503	3	215.68	0.786	29.99	444.32
48	223	1	260.39	3.63	130.1	398.84
51	60	1	203.68	5.76	131.6	315.07

Table 9.2. Mean and standard error of biomass in LULC classes within the BioREDD regions. Under each class the top number is the mean and the bottom number is the standard error calculated from the relations shown above from lidar derived biomass at pixel scale (0.25 ha).

BioREDD Project Areas	VM0006 Land Use and Land Cover Types						
	Intact Forest	Degraded Forest	Pasture	Crops	Wetlands	Settlement	Other
Chicao	275.29 0.80	79.81 1.48	22.9 1.58	48.62 1.31		9.76 0.51	
Apartado Buenavista	273.37 0.76	70.28 1.06	18.61 1.46	34.01 1.40			17.12 1.39
La Madre	250.16 1.52	95.91 0.96	18.92 1.36	30.71 1.50	32.83 0.99		
Rio Montaño	304.01 2.77	111.31 0.78	22.72 1.21	39.46 1.55	27.65 0.87		33.71 0.86
Vigia de Curvarado y Santa Ros	128.92 0.98	62.75 1.44	17.59 0.81	25.71 1.34	29.72 1.41	4.79 0.22	20.01 1.32
Chontadural Cañero	296.63 0.84	135.57 1.02	14.3 0.84	28.53 1.13			
Bellavista y Union Pitalito	248.27 0.76	131.02 1.23	14.34 0.98	28.94 1.50	12.70 0.81	14.26 0.94	
Polines	313.12 1.22	111.35 2.08	17.32 0.97	23.30 1.11			
Yaberarado (Abibe Chigorodo)	240.03 1.28	141.77 1.96	37.78 2.43	46.77 2.11		12.1 0.63	
Bajo Grande	234.64 0.98	88.75 0.96	22.34 0.51	24.08 0.88			
Santa Rosa de Ijua	254.74 0.64	122.92 0.86	17.60 0.67	31.60 0.73		8.66 0.54	
Ordo Siviru Aguaclara	263.44 0.66	118.19 0.97	10.01 0.59	22.02 0.77			
Jaikerazavi (Abibe Mutata)	218.41 1.22	80.07 2.18	22.15 0.93	41.40 0.73		15.78 0.34	
Rio Pepe	238.29 0.85	76.70 1.11	21.37 0.58	31.08 0.65		22.73 0.38	44.63 0.41
Mayor del Canton de San Pablo	249.16 0.64	123.38 1.38	15.77 0.83	38.21 1.38		25.07 0.53	23.41 1.77
Pizarro	204.42 1.41	89.28 1.28	18.22 0.77	47.12 1.57		13.24 0.38	31.97 0.34
Bajo Mira y Frontera	172.63 0.76	58.60 0.74	15.12 0.51	27.16 0.54	17.56 0.71	15.30 0.45	24.48 0.47
Acapa	255.88 1.66	60.09 2.13	16.30 0.76	36.48 1.1	11.49 0.89	22.34 1.21	
San Andres Usaraga	227.14 1.18	89.16 1.04	16.32 0.54	35.24 0.63		20.30 1.29	
Siviru	188.42 1.21	65.15 0.86	23.65 0.44	31.58 0.86	14.70 0.22	21.56 0.38	
Pedeguita y Mancilla	264.98 2.10	81.59 1.62	21.21 0.65	31.32 0.98	12.14 0.20	21.81 0.66	12.81 0.19

Los Rios La Larga y Tumarado	278.72 1.15	123.33 56.23	26.11 0.62	32.77 1.76	13.48 1.12	27.96 0.60	27.70 0.45
La Costa - Concosta	273.26 2.34	84.62 1.31	33.69 1.26	38.42 2.01		16.83 0.48	32.30 1.34
Rio Piliza	268.85 1.28	119.54 1.28	21.91 1.56	29.57 2.02		35.94 2.01	62.35 1.66
Rio Baudo - Acaba	256.59 1.20	105.10 1.25	18.35 1.74	27.76 1.99	47.02 1.20	21.12 1.19	36.47 1.72
Rio Calima (Bajo Calima)	270.82 1.35	124.92 1.44	30.38 1.36	33.05 2.08	52.71 2.90	11.61 0.54	14.69 0.8
Rio Cajambre	267.41 1.38	117.35 2.35	31.50 1.32	53.70 2.35	69.68 1.36		
Bahia Malaga La Plata	276.93 1.53	125.16 3.15	33.21 0.93	33.74 2.70	13.6 1.38		

Next, we estimate the aboveground carbon density and standard error in priority project areas selected for establishing REDD+ projects. The area of the project areas are smaller than the above regions and cover a combination of forest areas. For estimating carbon from biomass, we have used the fraction of 0.485 used in most tropical studies. This factor is less than 0.5 recommended by the IPCC guidlines.

Table 9.3. Mean and standard error of the carbon density for priority project areas. The class numbers are: 1. Forest, 2. Degraded forest, 3. Grasslands, 4. Croplands, 5. Wetlands, 6. Settlements, 7. Other lands.

BioREDD Project Areas	Land Cover Types	AGB_Mean (MgC/ha)	AGB_SE (MgC/ha)
BMBC	1	137.8481	1.4812
	2	53.98446	1.7984
	3	15.77896	2.1676
	4	20.5056	1.2012
	5	15.9641	2.0988
	6	6.560251	0.43949
	7	17.9625	0.63584
CDD	1	161.02431	1.6804
	2	36.97894	1.1710
	3	16.97222	1.2806
	4	8.501385	2.1625
	5	10.82348	0.91725
	6	0.2939836	0.043650
	7	34.03397	1.6881
SUPP	1	141.4029	2.1329
	2	68.05023	2.2782
	3	10.40879	1.8794
	4	5.847085	1.7809
	5	0	0.0000
	6	8.3129	0.32215
	7	26.69632	2.1279
BMF_ACAPA	1	153.49675	4.5374
	2	78.30404	2.0869
	3	12.93648	1.5032
	4	17.64887	1.2884
	5	15.29623	1.0375
	6	7.283083	0.26140
	7	22.66439	0.94841
CAJAMBRE	1	136.079	1.7262
	2	69.33057	2.2327

	3	17.17667	1.7318
	4	28.52794	1.4162
	5	24.08764	1.7097
	6	5.75	0.33953
	7	5.8	0.056657
CONCOSTA	1	148.83324	2.0843
	2	81.77203	2.1042
	3	15.7625	1.5443
	4	22.15574	1.5248
	5	23.50919	1.3750
	6	3.2725	0.15316
	7	16.48239	0.83079
MUTATA	1	129.2673	2.1823
	2	71.78804	2.7106
	3	16.34767	2.1922
	4	29.79207	2.2564
	5	0	0.0000
	6	14.37875	2.1568
	7	0	0.0000
RIO PEPE ACABA	1	130.7909	1.9814
	2	52.50507	1.8920
	3	15.88363	1.8362
	4	27.70095	1.5082
	5	0	0.0000
	6	16.96421	0.44335

10. Summary and Conclusion

This report provided the methodology and the estimation of forest biomass from lidar inventory sampling of the BioREDD region. The relations provided in the report are also used in the methodology tool document currently under review for approval. The sampling approach used in this study mimics the statistical field inventory approach. By collecting a large number of remote sensing data, we were able to not only estimate the forest biomass but reduce the standard error significantly compared to often limited ground sampling approaches. The results from our analysis can be summarized as:

1. An unbiased estimator for the mean and variance of the biomass based on lidar stratified random sampling was developed.
2. The predictor model is based on two parameters: mean top canopy height derived from lidar at the plot scale (here 0.25 ha), and the average wood density at the landscape, regional, or forest type scales.
3. A minimum of 30 sample plots was required to develop the lidar biomass model and about 15 plots for validation and uncertainty assessment for each key forest type with different ground allometry. Adding more plots for calibration and validation only slightly improved the model fit and the uncertainty in biomass estimation.
4. The estimate of variance also includes the error propagation from lidar biomass modeling and sampling uncertainty. Since the number of biomass samples from lidar measurements is large, the standard error including the spatial autocorrelation remain small.

5. Here, we focused only on the aboveground biomass. In the next report, we will provide the carbon numbers for all pools necessary to estimate the emissions from deforestation and degradation.

10. References

Asner, G. P., Mascaro, J., Muller-Landau, H. C., Vieilledent, G., Vaudry, R., Rasamoelina, M., ... & van Breugel, M. (2012). A universal airborne LiDAR approach for tropical forest carbon mapping. *Oecologia*, *168*(4), 1147-1160.

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

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M. S., Delitti, W. B., ... & Vieilledent, G. (2014). Improved allometric models to estimate the aboveground biomass of tropical trees. *Global change biology*.

Chave J*,* Andalo C*,* Brown S *et al. (*2005*)* Tree allometry and improved estimation of carbon stocks and balance in tropical forests. Oecologia*,* 145*,* 87–99.

Gregoire, T. G., Ståhl, G., Næsset, E., Gobakken, T., Nelson, R., & Holm, S. (2011). Model-assisted estimation of biomass in a LiDAR sample survey in Hedmark County, Norway. Canadian Journal of Forest Research-Revue CanadienneDe Recherche Forestiere, 41, 83–95.
Meyer, V., Saatchi, S. S., Chave, J., Dalling, J. W., Bohlman, S., Fricker, G. A., ... & Hubbell, S. (2013). Detecting tropical forest biomass dynamics from repeated airborne Lidar measurements. *Biogeosciences*, *10*(8), 5421-5438.

Næsset, E., Gobakken, T., Solberg, S., Gregoire, T. G., Nelson, R., Ståhl, G., et al. (2011). Model-assisted regional forest biomass estimation using LiDAR and InSAR as auxiliary data: A case study from a boreal forest area. Remote Sensing of Environment, 115, 3599–3614.

Neigh, C. S., Nelson, R. F., Ranson, K. J., Margolis, H. A., Montesano, P. M., Sun, G., ... & Andersen, H. E. (2013). Taking stock of circumboreal forest carbon with ground measurements, airborne and spaceborne LiDAR. *Remote Sensing of Environment*, *137*, 274-287.

Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. A., Salas, W., et al. (2011). Benchmark map of forest carbon stocks in tropical regions across three continents.Proceedings of the National Academy of Sciences of the United States of America, 108,9899–9904.

Ståhl, G., Holm, S., Gregoire, T. G., Gobakken, T., Næsset, E., & Nelson, R. (2011). Model-based inference for biomass estimation in a LiDAR sample survey in Hedmark County, Norway. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere, 41, 96–107.