Article

Estimating Aboveground Biomass in Tropical Forests: Field Methods and Error Analysis for the Calibration of Remote Sensing Observations

Fabio Gonçalves [1,*], Robert Treuhaft [2], Beverly Law [3], André Almeida [4], Wayne Walker [5], Alessandro Baccini [5], João Roberto dos Santos [6] and Paulo Graça [7]

[1] Canopy Remote Sensing Solutions, Florianópolis, SC 88032, Brazil
[2] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA; robert.n.treuhaft@jpl.nasa.gov
[3] Department of Forest Ecosystems & Society, Oregon State University, Corvallis, OR 97331, USA; bev.law@oregonstate.edu
[4] Departamento de Engenharia Agrícola, Universidade Federal de Sergipe, SE 49100, Brazil; andre.almeida@ufs.br
[5] Woods Hole Research Center, Falmouth, MA 02540, USA; wwalker@whrc.org (W.W.); abaccini@whrc.org (A.B.)
[6] National Institute for Space Research (INPE), São José dos Campos, SP 12227, Brazil; joao.roberto@inpe.br
[7] Department of Environmental Dynamics, National Institute for Research in Amazonia (INPA), Manaus, AM 69067, Brazil; pmlag@inpa.gov.br
* Correspondence: fabio@canopyrss.tech; Tel.: +55-48-99139-9123

Abstract: Mapping and monitoring of forest carbon stocks across large areas in the tropics will necessarily rely on remote sensing approaches, which in turn depend on field estimates of biomass for calibration and validation purposes. Here, we used field plot data collected in a tropical moist forest in the central Amazon to gain a better understanding of the uncertainty associated with plot-level biomass estimates obtained specifically for the calibration of remote sensing measurements. In addition to accounting for sources of error that would be normally expected in conventional biomass estimates (e.g., measurement and allometric errors), we examined two sources of uncertainty that are specific to the calibration process and should be taken into account in most remote sensing studies: the error resulting from spatial disagreement between field and remote sensing measurements (i.e., co-location error), and the error introduced when accounting for temporal differences in data acquisition. We found that the overall uncertainty in the field biomass was typically 25% for both secondary and primary forests, but ranged from 16 to 53%. Co-location and temporal errors accounted for a large fraction of the total variance (>65%) and were identified as important targets for reducing uncertainty in studies relating tropical forest biomass to remotely sensed data. Although measurement and allometric errors were relatively unimportant when considered alone, combined they accounted for roughly 30% of the total variance on average and should not be ignored. Our results suggest that a thorough understanding of the sources of error associated with field-measured plot-level biomass estimates in tropical forests is critical to determine confidence in remote sensing estimates of carbon stocks and fluxes, and to develop strategies for reducing the overall uncertainty of remote sensing approaches.

Keywords: forest inventory; allometry; uncertainty; error propagation; Amazon; ICESat/GLAS

1. Introduction

Our ability to estimate aboveground forest biomass from remote sensing observations has advanced substantially over the past decade, largely due to the increased availability of direct three-dimensional (3-D) measurements of vegetation structure provided by light detection and ranging (Lidar; [1]) and interferometric synthetic aperture radar (InSAR; [2]). Although approaches to forest biomass estimation based on remotely sensed structure have yet to be fully developed and validated (cf. [3]), they are already greatly expanding our knowledge of the amount and spatial distribution of carbon stored in terrestrial ecosystems, particularly in tropical forests (e.g., [4–6]), where large areas have never been inventoried on the ground [7]. Lidar remote sensing, calibrated with field measurements and combined with wall-to-wall observations from InSAR and/or passive optical systems, represents a promising alternative to more traditional approaches to biomass mapping (e.g., [8,9]) and is expected to play a key role in forest monitoring systems being developed in the context of climate change mitigation efforts such as REDD (Reducing Emissions from Deforestation and Forest Degradation), and to improve our understanding of the global carbon balance [10–13].

The typical approach to producing spatially explicit estimates of biomass from 3-D remote sensing is characterized by two primary steps. First, field estimates of aboveground biomass density are obtained from sample plot data together with published allometric equations, which allow the estimation of tree-level biomass from more easily measured quantities such as diameter, height, and wood density [14–16]. Second, the plot-level estimates of biomass are related to co-located remote sensing estimates of structure (e.g., mean canopy height) using a statistical model. The model is then applied together with remote sensing data to predict biomass in locations where ground measurements are not available [17–21]. When the 3-D measurements are spatially discontinuous, as is usually the case with Lidar, the resulting biomass predictions can be further integrated with radar and/or passive optical imagery (typically using machine learning algorithms) to produce wall-to-wall maps of biomass or carbon [5,6], although often with poorer resolution and unknown accuracy.

One of the main limitations of this scaling approach, as noted by [22], is that biomass is never measured directly (i.e., quantified by harvesting and weighing the leaves, branches, and stems of trees). Because direct measurements are laborious, time-consuming, and ultimately destructive (e.g., [23]), the remotely sensed structure is calibrated against allometrically estimated biomass (a function of diameter and sometimes height and wood density) and the final product is, in essence, "an estimate of an estimate" of biomass.

Despite significant advances in the development of allometric equations for tropical forest trees over the past decade [15,16], the allometrically-derived biomass is subject to a number of sources of error, including: (i) uncertainty in the estimation of the parameters of the allometric equation as a result of sampling error (e.g., resulting from a relatively small number of trees being harvested or bias against the harvest of trees with a "typical" form), natural variability in tree structure (i.e., trees of the same diameter, height, and wood density can display a range of biomass values), and measurement errors on the harvested trees; (ii) uncertainty associated with the choice of a particular equation or application of a given equation beyond the site(s) and/or species for which it was developed (uncertainty driven primarily by biogeographic variation in allometric relations due to soil fertility and climate); and (iii) measurement errors in the diameter, height, and wood density of the trees that the allometric equation is being applied to [22,24,25]. Combined, these sources of uncertainty have been estimated to represent approximately 50%–80% of the estimated biomass at the tree level, and over 20% at the plot scale [24,26].

Because remote sensing algorithms for prediction of forest biomass are typically calibrated with allometrically estimated biomass (see [27] for an exception), they incorporate all of the sources of uncertainty described above, in addition to those associated with the remote sensing observations. As a result, while the precision (degree of reproducibility) of remote sensing estimates of biomass can be easily assessed, their accuracy is rarely known. Although precision may be all that is needed for the relative tracking of changes in carbon stocks for REDD-like initiatives, accuracy is ultimately critical

for determining the absolute amount of carbon stored in forests, as required for global carbon budgets and climate change science [22].

While the accuracy of remote sensing-based estimates of biomass cannot be truly determined without whole-plot harvests, it can nevertheless be optimized. When calibrating remotely sensed structure to allometrically estimated biomass, it is reasonable to expect that the uncertainty in the estimated biomass will vary from plot to plot as a function of differences in, for example, tree size distribution and species composition. If these plot-level uncertainties can be estimated precisely relative to one another, the prediction accuracy of the statistical model relating biomass to remotely sensed structure can be significantly improved by giving sample plots with smaller standard deviations more weight in the parameter estimation. In practice, this can be accomplished, for example, using the method of weighted least squares (WLS), where each plot is weighted inversely by its own variance [28,29]. Although considered a standard statistical technique for dealing with nonconstant variance when responses are estimates, WLS is almost never used in the current context, in part because the uncertainty in the field biomass is rarely quantified (see [30] for an exception).

In this study, we use field plot data collected at the Tapajós National Forest, Brazil, to gain a better understanding of the uncertainty associated with plot-level biomass estimates obtained specifically for calibration of remote sensing measurements in tropical forests. In addition to accounting for sources of error that would be normally expected in conventional biomass estimates (e.g., measurement and allometric errors; [24,31–34]), we examine two sources of uncertainty that are specific to the calibration process and should be taken into account in most remote sensing studies: (1) the error resulting from spatial disagreement between field and remote sensing samples (co-location error); and (2) the error introduced when accounting for temporal differences in data acquisition.

2. Materials and Methods

2.1. Study Site

The Tapajós National Forest is located along highway BR-163, approximately 50 km south of the city of Santarém, Pará, in the central Brazilian Amazon (Figure 1). The climate is tropical monsoon (Köppen Am), with a mean annual temperature of 25.1 °C and annual precipitation of 1909 mm, with a 5-month dry season (<100 mm month^{-1}) between July and November [35]. The vegetation is dense, upland, tropical moist forest. Common genera among 193 tree species sampled in this study include *Psychotria* (Rubiaceae), *Protium* (Burseraceae), *Otoba* (Myristicaceae), *Eschweilera* (Lecythidaceae), *Pouteria* (Sapotaceae), and *Rinorea* (Violaceae) in primary forests, and *Cecropia* (Urticaceae), *Banara* (Salicaceae), and *Inga* (Fabaceae) in secondary forests. The soils are nutrient-poor oxisols and ultisols, with low pH, organic matter, and cation exchange capacity, and a high concentration of aluminum oxides [36]. Our sample sites were situated on a relatively flat plateau, with the elevation ranging from approximately 80 to 180 m.

2.2. Field Data

Field data were collected in September 2010 in 30 0.25-ha plots (50 m × 50 m) intended for calibration of Lidar data acquired by the Geoscience Laser Altimeter System (GLAS; [19]). Of the 30 plots, 8 were primary forest (PF), 8 were primary forest subject to reduced-impact selective logging (PF$_L$) between 1999 and 2003 (~3.5 trees harvested per hectare during this period; [37]), and 14 were secondary forest (SF) with different age and disturbance histories. Plots were centered on GLAS footprints selected along two sensor acquisition tracks, spanning a wide range in vertical structure and aboveground biomass (Figure 1). Individual footprint centers were located on the ground using a total station and the Differential Global Positioning System (DGPS).

Field biometric measurements included diameter at breast height (*D*) measured with a diameter tape at 1.3 m and recorded to the nearest 0.1 cm; height to the base of the live crown (*H$_C$*) and total height (*H$_T$*), estimated visually by experienced members of the field crew to the nearest 0.5 m;

and crown depth (C_D), calculated as the difference between H_T and H_C. Measurements were taken for each living tree ≥5 cm in diameter in early successional stands and ≥10 cm in all other stands. For a 12.5 m × 50 m subplot extending along the major axis of the GLAS footprint, we also measured crown radius (C_R) in two orthogonal directions by projecting the edge of the crown to the ground and recording its horizontal distance to the trunk to the nearest 0.1 m using a tape measure. All trees were identified as to their species or genus (when species was uncertain) level and assigned a wood density value (ρ, oven-dry weight over green volume) derived from the literature [38,39].
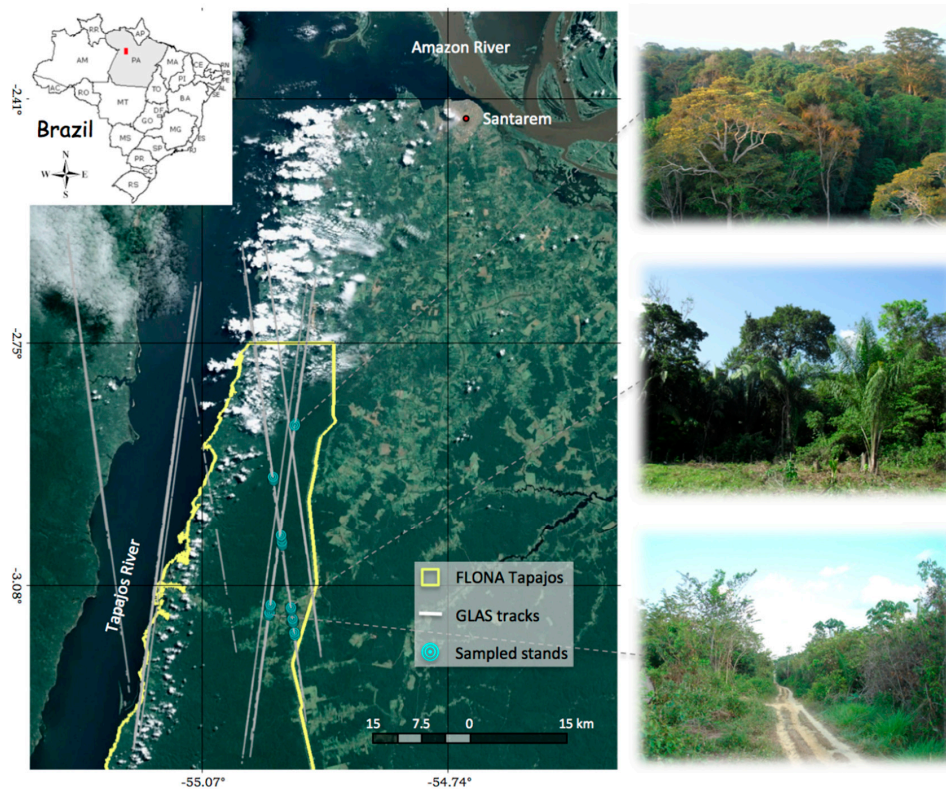

Figure 1. Geographical location of the Tapajós National Forest, PA, Brazil, outlined in yellow. The gray lines are GLAS tracks from 2003 to 2009, and the blue targets are the locations of the plots sampled in this study. The pictures on the right illustrate three of the stands where plots were located, with aboveground biomass ranging from near zero (bottom) to over 400 Mg·ha^{-1} (top).

To estimate measurement errors associated with the inventory, we conducted a blind remeasurement of 2–4 trees selected at random in each plot, resulting in a total resampling effort of 3%. For a portion of the trees that were remeasured (and additional trees selected in open areas), heights were also obtained with a laser rangefinder using the tangent method [40]. We used these observations to develop a regression model relating precise laser-measured heights to less precise, however more readily obtained, visually estimated heights (cf. [37]). This model, described in detail in [19], was applied to calibrate all visually estimated heights.

2.3. Biomass Estimation

The oven-dry aboveground mass of each live tree (M) was estimated from its diameter, total height, and wood density using an established allometric equation for tropical moist forests ([15]; Table 1). Exceptions were made for *Cecropia* spp. and palms, which differ significantly from other species in wood density and allometry [15,41] and had their biomass estimated with specific equations, as indicated in Table 1. We selected Chave's equation as the basis of our estimate because it was developed using a large number of harvested trees (1350) covering a wide range in diameter (5–156 cm), and

because it included information on tree height and wood density, which greatly improves the accuracy of biomass estimates [15,16]. In addition, approximately 43% of the trees used in the fit were harvested in the Brazilian Amazon, and 11% in the state of Pará, in sites having climatic and edaphic conditions comparable to ours. Nevertheless, we also estimated *M* in this study using two additional, widely used equations ([14,42]; Table 1), to obtain a measure of allometric uncertainty as described in Section 2.4.2.

　　The aboveground biomass density (*AGB*, Mg·ha^{-1}) at the plot level was calculated by adding the masses of all inventoried trees in the plot and dividing by the plot area (i.e., 0.25 ha). In plots where the minimum diameter was 10 cm, we corrected for the *AGB* in the 5–10 cm class by: (i) fitting a negative exponential function to the diameter distribution of the plot [33,43]

$$N_i = k \, e^{-a d_i} \tag{1}$$

where N_i was the number of trees per hectare in the *i*th diameter class with midpoint d_i, and k and a were model parameters estimated by nonlinear least squares; (ii) estimating the number of trees per hectare in the 5–10 cm class; and (iii) multiplying the resulting number by the biomass of a tree with diameter of 7.5 cm (midpoint) and wood density equal to the plot mean. To avoid errors due to the estimation of a mean height for the 5–10 cm class, we used an alternative equation based on diameter and wood density only ([15]; Table 1).

　　Because the GLAS Lidar data were acquired in 2007, three years prior to our field measurements, we applied the site-specific, stand-level growth model [44]

$$
\begin{aligned}
AGB_t &= 0.397 \, B_t \, H_{TOP_t} \,, \ with \\
B_t &= 21.057 \left(1 - e^{-0.109 \, t}\right)^{1.894} \\
H_{TOP_t} &= 23.067 \left(1 - e^{-0.074 \, t}\right)^{0.946}
\end{aligned}
\tag{2}
$$

to all SF plots to correct for the *AGB* change between observation epochs, where AGB_t is the plot biomass (Mg·ha^{-1}), estimated with the equation of [45] (Table 1); B_t is the basal area (m^2·ha^{-1}); H_{TOP_t} is the top height (m), defined as the mean total height of the tallest 20% of the trees; and *t* is the stand age (years since stand initiation). Because this growth model is based on a biomass equation that is different from the ones used in this study, we calculated the biomass change as a proportion of the plot biomass. This was done by: (i) inverting the first line of Equation (2) to estimate the plot age in years in 2010 (t_{2010}) from its measured biomass (AGB_{2010}); (ii) estimating the plot biomass at the time of the GLAS acquisition (AGB_{2007}), making $t = t_{2010} - 3$; and (iii) calculating the ratio ($AGB_{2010} - AGB_{2007}$)/AGB_{2010}. For primary forests, we assumed no biomass change in the 3-year period. This is supported by [44], who found that biomass tends to increase rapidly in early stand development at Tapajós, reaching near-asymptote as early as 40–50 years after clear-cutting (Figure 2).

Table 1. Allometric equations used to calculate individual tree biomass at Tapajós.

Category	Equation*	Source
Trees		
Cecropia spp.	$M_{T1} = exp(-2.5118 + 2.4257 \ln(D))$	[41]
All others	$M_{T2} = exp\left(-2.977 + \ln\left(\rho D^2 H_T\right)\right)$	[15]
Palms		
Attalea spp.	$M_{P1} = 63.3875 \, H_C - 112.8875$	[46]
All others	$M_{P2} = exp(-6.379 + 1.754 \ln(D) + 2.151 \ln(H_T))$	[47]
Alternative Equations		
	$M_{A1} = exp(-2.134 + 2.530 \ln(D))$	[14]
	$M_{A2} = \exp\left(-0.370 + 0.333 \ln(D) + 0.933 \, \ln(D)^2 - 0.122 \, \ln(D)^3\right)$	[42]
	$M_{A3} = \rho \, \exp\left(-1.499 + 2.148 \ln(D) + 0.207 \, \ln(D)^2 - 0.028 \, \ln(D)^3\right)$	[15]
	$M_{A4} = exp(-3.1141 + 0.9719 \ln\left(D^2 H_T\right))$	[45]

*M (kg) is the oven-dry aboveground tree biomass, D (cm) is the diameter at breast height (1.3 m), H_T (m) is the total height, H_C (m) is the height to the base of the live crown, and ρ (g·cm^{-3}) is the wood density measured as oven-dry weight over green volume.
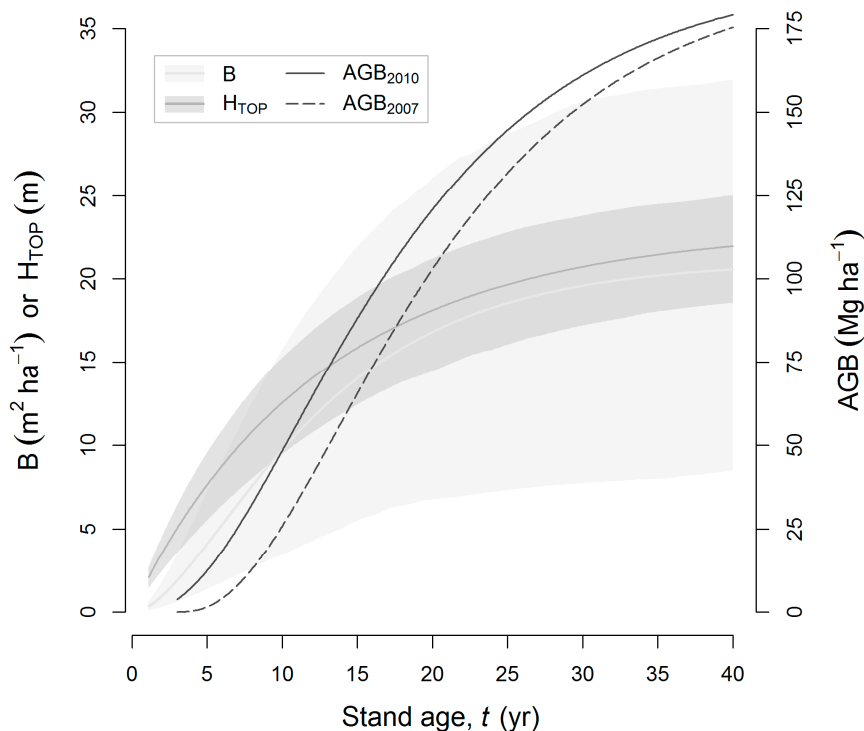
Figure 2. Variation in basal area (B), top height (H_{TOP}), and aboveground biomass (*AGB*) with stand age (*t*) at Tapajós, as described by the growth model of [44]. The solid lines represent structural values at age *t*, and the dashed line shows the expected *AGB* at the time of the GLAS acquisition (3 years prior) for a stand of age t. The gray bands for B and H_{TOP} are 95% confidence intervals from a Monte Carlo simulation [44].

2.4. Error Analysis

2.4.1. Individual Tree Measurements

Measurement errors in D, H_C, H_T, C_D, and C_R were described in terms of total error, systematic error (bias), and random error. The total error for each attribute was quantified with the root mean square deviation (RMSD)

$$\text{RMSD} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}e_i{}^2} \text{ , with } e_i = m1_i - m2_i \tag{3}$$

where n is the number of pairs of repeated measurements for the attribute, and e_i is the measurement difference for the *i*th pair, with $m1_i$ and $m2_i$ representing the original measurement and remeasurement, respectively. The systematic and random errors were quantified, respectively, as the mean and sample standard deviation (SD) of the measurement differences e_i

$$\text{Mean} = \frac{1}{n}\sum_{i=1}^{n}e_i \tag{4}$$

$$\text{SD} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(e_i - \text{Mean})^2} \tag{5}$$

We also calculated all of the above in relative terms, expressing e_i as a fraction of the average of the two measurements. For the wood density values, the standard deviation was either taken or estimated from the supplementary material provided by [39].

To test the hypothesis of no systematic difference between the first and second measurements of a given attribute, we used either a paired t-test or the alternative Wilcoxon signed-rank test, depending on the assessment of distributional assumptions and the presence of outliers [29]. To determine whether measurement variation increased with the magnitude of the measurement, we: (i) divided the measurements for a given attribute into four classes with an equal number of samples; (ii) regressed the SD calculated for each size class on the average value of the measurement for that class; and (iii) tested whether the slope was significantly different from zero. Finally, we tested if measurement differences varied with forest type by including forest type as a factor in the regression of the absolute measurement difference on the magnitude of the measurement—i.e., incorporating different intercepts and slopes for SF, PF_L and PF—and testing for the equality of the coefficients using the extra-sum-of-squares F-test [29].

2.4.2. Biomass

Measurement errors in diameter (σ_D), height (σ_H), and wood density (σ_ρ) were propagated to the biomass estimate by expanding the allometric equations in Table 1 to a Taylor series and retaining only first-order terms. For a model like M_{T2} (Table 1), of the form $M = aD^kH\rho$, with ρ uncorrelated with both D and H, we expressed the uncertainty in the mass of a tree (σ_M) in terms of measurement errors as [24]

$$
\begin{aligned}
\sigma_M &= \left[\sigma_D^2 \left(\frac{\partial M}{\partial D} \right)^2 + \sigma_H^2 \left(\frac{\partial M}{\partial H} \right)^2 + \sigma_\rho^2 \left(\frac{\partial M}{\partial \rho} \right)^2 + 2\sigma_{DH}^2 \left(\frac{\partial M}{\partial D} \right) \left(\frac{\partial M}{\partial H} \right) \right]^{1/2} \\
&= M \left(k^2 \frac{\sigma_D^2}{D^2} + \frac{\sigma_H^2}{H^2} + \frac{\sigma_\rho^2}{\rho^2} + 2k \frac{\sigma_{DH}^2}{DH} \right)^{1/2}
\end{aligned}
\tag{6}
$$

where the terms in parentheses in the upper equation are the partial derivatives of M with respect to each of the dendrometric quantities, D, H, and ρ; and σ_{DH} is the covariance between D and H.

We accounted for two sources of allometric uncertainty: (i) the uncertainty related to the model residuals, σ_A, estimated as [24,48]

$$
\sigma_A = \left[e^{(2\hat{\sigma}^2 + 2\ln M)} - e^{(\hat{\sigma}^2 + 2\ln M)} \right]^{1/2} = \left(e^{\hat{\sigma}^2} - 1 \right)^{1/2} \langle M \rangle
\tag{7}
$$

where $\hat{\sigma}$ is the standard error of the regression on log-transformed data (see Table 1), and $M = M \times \exp(\hat{\sigma}^2/2)$ is an unbiased estimate of the back-transformed biomass prediction M; and (ii) the uncertainty involved in the selection of the allometric equation, σ_S, estimated by calculating the mass of each tree with three independent equations (M_{T2}, M_{A1}, and M_{A2} in Table 1) and obtaining the standard deviation of the resulting values. The quantification of a third source of allometric uncertainty, the uncertainty in the determination of the model parameters as a result of sampling error (e.g., [26]), would require access to the destructive harvest data used in the development of the allometric equations and was not considered in this study.

Spatial disagreement between the field plots and the GLAS footprints introduced additional uncertainty in the biomass estimates. This co-location error, σ_C, was introduced because of positional errors associated with both data sets [49,50], and because the size, shape, and orientation of the field and the GLAS samples did not exactly coincide (Figure 3). We took a Monte Carlo approach, based on binomial statistics [20], to estimate the difference in biomass between what was measured in each 50 m × 50 m field plot and what was actually present in the area covered by the GLAS footprint (Figure 3). The binomial-statistical approach constructs an ensemble of possible tree masses (M) based on those measured in the field. The method of ensemble-member construction is to assume that the set of Ms actually measured for each tree are the only values allowable for all ensemble members. For a single area, in just the FUI (field) area of Figure 3 for example, say there were 100 trees. There are 100 tree mass "bins"—bins labeled by the mass of the actually-measured tree—and the only way that another statistical ensemble member can be realized is by changing the population number of trees in

each bin according to the binomial distribution, which means the probability of populating a single mass bin of mass *B* with *x* trees is

$$P_B(x|n_p) = \begin{pmatrix} n_p \\ x \end{pmatrix} p^x (1-p)^{n_p-x} \qquad (8)$$

where p is the probability of finding a single tree in the mass bin, p = $1/n_p$, and n_p is the total number of trees in the plot, 100 in this example. On average, each mass bin *B* will be populated with 1 tree, ($n_p \times p = 1$), a result of binomial statistics, and the total average mass of the plot will be exactly what was measured in the field (the mass of the sum of the 100 bin labels). The fundamental assumption of binomial statistics is that each outcome—the number of trees in each bin—is independent of all other outcomes. That is, for example, if we measured one tree at 500 kg and one at 1000 kg, the probability of finding one tree in a tree mass bin of 500 kg is the same as finding one in a mass bin of 1000 kg.
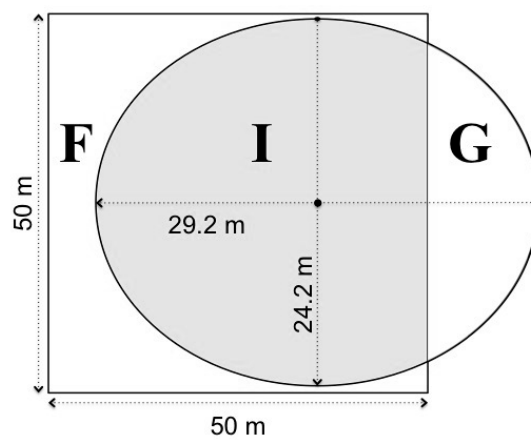

Figure 3. Schematic diagram illustrating an arbitrary intersection of a GLAS footprint (ellipse) and a field plot (square), portrayed in their correct relative nominal sizes and shapes. Co-location error was introduced due to differences in the areas sampled by each technique (G ∪ I for GLAS vs. F ∪ I for field). The gray zone "I" represents the area covered effectively by both techniques and averaged 75% of the footprint area for the 30 plots used in this study.

Figure 3 shows 3 areas relevant to considering area mismatch. The field measurements were taken with the rectangular boundaries shown, and the GLAS measurements were taken within the ellipse. "F" signifies the part of the field measurement area not in common with GLAS, and, similarly, "G" refers to the part of the GLAS measurement area not in common with the field. "I" refers to the area in common between field and GLAS, and for that area, the field-GLAS mismatch is zero. It is the difference in biomasses of areas F∪I and G∪I in Figure 3 that is of interest in assessing the area "mismatch", where F∪I is the zone measured in the field and G∪I is the zone observed by GLAS. Probabilities as in Equation (8) for the same mass bins as in the field are constructed. That is, it is assumed that the spectrum of tree masses is the same for all areas of Figure 3. In order to construct a probability for one of the zones, n_p in Equation (8) must be replaced by n_z, the number of total trees in the zone. This total number is assumed to be in proportion to area. We therefore took the probability of *x* trees in bin *B* of zone *z* to be

$$P_B(x|n_z, p) = \begin{pmatrix} n_z \\ x \end{pmatrix} p^x (1-p)^{n_z-x} \qquad (9)$$

Monte Carlo values of x_i were generated for the *i*th bin of zone *z* with the probability distribution of Equation (9). Total zone *z* biomasses (B_z) were generated by summing over all N bins (the number of trees measured in the field) for each throw:

$$B_z = \sum_{i=0}^{N} x_i B_i \tag{10}$$

The difference $(B_{G \cup I}/\text{Area}_{G \cup I}) - (B_{F \cup I}/\text{Area}_{F \cup I})$ was calculated for each Monte Carlo throw. This procedure was repeated automatically 10^4 times and the standard deviation of the biomass differences was taken as the co-location error.

In addition to the error sources discussed above, we included in our error budget the uncertainty associated with the corrections based on Equations (1) and (2), described in Section 2.3. The error of estimating biomass for the 5–10 cm diameter class, σ_{5-10}, was assumed to be the same as for the 10–15 cm class, where all trees were actually measured and the error could be determined. This assumption was tested and verified on plots where the minimum diameter was 5 cm. The uncertainty associated with the application of the growth model, σ_G, was estimated by propagating the uncertainties in the parameters of Equation (2) to the determination of the biomass change, in a framework similar to Equation (6).

Errors σ_M, σ_A, and σ_S were calculated at the tree level and added in quadrature to obtain plot-level estimates on a per-hectare basis. These errors were in turn combined in quadrature with σ_C, σ_{5-10}, and σ_G, calculated directly at the plot level, to obtain an estimate of the overall uncertainty under the assumption of additivity and statistical independence.

3. Results

3.1. Tree Measurement Errors

Uncertainties resulting from differences in repeated measurements of D, H_C, H_T, C_D, and C_R are summarized in Table 2. D was the most precisely measured quantity, with a RMSD of less than 2%. Repeated measurements of height were typically within 1 m of each other (RMSD = 15%–18%), with approximately half of the H_C, and a quarter of the H_T observations showing identical repeated measurements. C_D and C_R measurements showed considerably less agreement (RMSD of 31 and 26%, respectively). However, with the exception of D, there was no evidence of a systematic difference between first and second measurements (Table 2 and Figure 4). For D, the data suggested that the second measurement produced values that were lower to a statistically significant degree when compared to the first measurement, although the estimated median difference of less than 0.1 cm has no practical significance.

Table 2. Summary statistics of differences between repeated measurements of diameter (D), height to the base of the live crown (H_C), total height (H_T), crown depth (C_D), and crown radius (C_R) for trees sampled at Tapajós. Statistics include total error (RMSD), systematic error (mean), and random error (SD), in both absolute and relative terms, as described in Section 2.4.1. The number of observations was 104, except for C_R (n = 144).

Attribute	Range	Differences					
		RMSD	Mean	SD	% That Is:		
					0	≤10%	≤25%
D (cm)	5.5–110.5	0.8 (1.8%)	0.1 (0.5%)	0.8 (1.8%)	23.1	100	100
H_C (m)	1.5–31.0	1.8 (17.7%)	0.1 (0.6%)	1.8 (17.8%)	47.1	54.8	83.7
H_T (m)	5.0–40.0	2.3 (15.2%)	−0.2 (−1.7%)	2.3 (15.2%)	24.0	53.8	93.3
C_D (m)	1.0–20.0	1.8 (30.7%)	−0.3 (−4.8%)	1.8 (30.5%)	32.7	33.7	71.2
C_R (m)	0.7–8.0	0.8 (25.7%)	0.0 (0.0%)	0.8 (25.8%)	11.8	37.5	68.8

Measurement variation increased significantly with the magnitude of the measurement across all attributes (Figure 4). The estimated rates of increase in the standard deviation of the measurement differences were 4, 10, 7, 18, and 26% for D, H_C, H_T, C_D, and C_R, respectively. When differences

were expressed as a percentage of the measurement, H_C and C_R showed no significant trend. The relative differences in D also increased with D (at a rate of 0.04% cm^{-1}), although the evidence was only suggestive, and the differences in H_T and C_D actually decreased with the magnitude of the measurements, at rates of 0.4% m^{-1} and 1.9% m^{-1}, respectively. There was no evidence that absolute differences between repeated measurements (both the mean and the rate of change) varied with forest type, after accounting for differences in the magnitude of the measurements.

In terms of wood density, about 90% of the inventoried trees showed a coefficient of variation (CV) of less than 20%. The CV was 15% on average (median of 14%), but ranged from 0 to as high as 64%, depending on the method used to assign the wood density value (e.g., species-vs. genus-level database estimates).
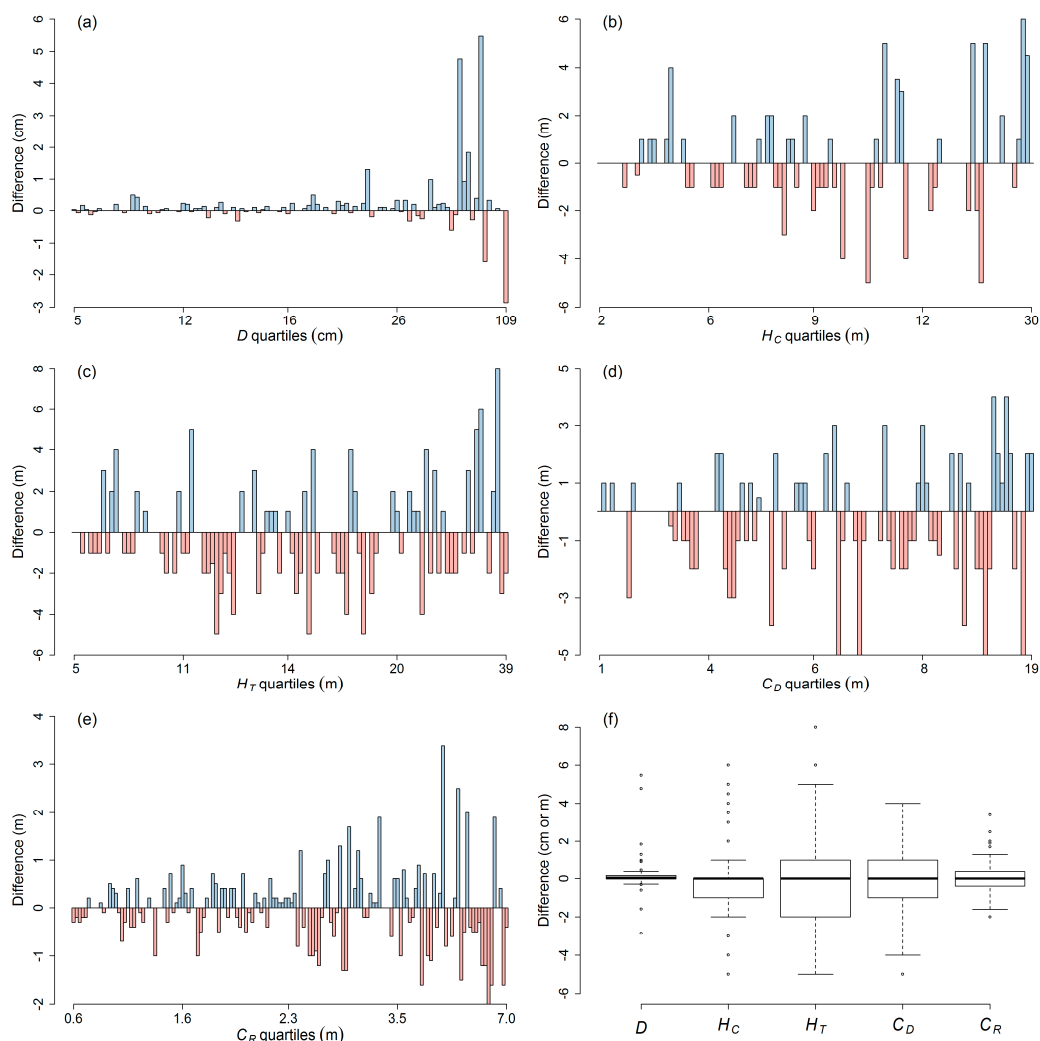
Figure 4. Differences between repeated measurements of: (**a**) diameter; (**b**) height to the base of the live crown; (**c**) total height; (**d**) crown depth; and (**e**) crown radius, ordered by the magnitude of the measurement (estimated as the average of the two measurements). (**f**) Side-by-side boxes represent the middle 50% of the distributions, with medians marked by a thick black line. The whiskers extend to the smallest and largest differences not more than 1.5 box-lengths away from the box, and the dots represent extreme values.

3.2. Field Biomass

Plot-level aboveground biomass ranged from 1.9 to 130.1 Mg·ha^{-1} in secondary forests, and from 162.6 to 423.6 Mg·ha^{-1} in primary forests, with no apparent difference between PF and PF$_L$ plots.

The overall mean was 174.8 ± 134 (SD) Mg·ha^{-1}, and the median was 172.8 Mg·ha^{-1}. These results are summarized by forest type in Table 3, along with a number of other stand characteristics.

Table 3. Characteristics of secondary (SF), selectively-logged (PF_L), and primary (PF) forest stands used in this study. Field plots were 0.25 ha in size (50 m × 50 m). The minimum diameter was 10 cm, except for seven early successional stands where the minimum was 5 cm. Values are median and range (in parentheses).

Attribute	Forest Type		
	SF (14 Plots)	PF_L (8 Plots)	PF (8 Plots)
Number of species (0.25 ha^{-1})	28 (7–45)	41 (36–49)	42 (31–45)
Stem density (trees ha^{-1})	488 (132–1052)	380 (340–456)	354 (304–424)
Basal area (m^2·ha^{-1})	8.7 (1.5–16.9)	22.8 (18.9–31.3)	24.8 (15.9–30.4)
Mean height (m)	12.9 (7.1–19.9)	19.7 (17.2–21.8)	18.0 (13.5–22.1)
Mean wood density (g·cm^{-3})	0.50 (0.37–0.54)	0.65 (0.54–0.70)	0.62 (0.59–0.63)
Biomass (Mg·ha^{-1})	37.3 (1.9–130.1)	285.8 (183.0–423.6)	293.0 (162.6–417.5)
Mean crown depth (m)	5.2 (2.9–8.3)	7.0 (6.4–8.2)	7.0 (5.5–9.5)
Mean crown radius (m)*	2.3 (0.8–3.0)	3.0 (2.2–3.5)	3.1 (2.2–4.1)

* Estimated from trees located in a central 12.5 m × 50 m subplot.

Figure 5 shows the average stem density (bars) and tree height (circles) per diameter class for all secondary (dark gray) and primary (light gray) forest plots. The solid lines show the fit of Equation (1) to the average diameter distributions, and the dashed lines show the fit of a similar exponential decay model to the tree height data. The diameter distributions followed an inverted J-shaped curve typical of tropical forests, with the ratio of the number of trees in successive diameter classes roughly constant (~1.9 for SF and 1.7 for PF/PF_L). Secondary forests showed considerably fewer (and generally shorter) trees than primary forests at any given diameter class, except for the smallest classes (<20 cm). Primary forests showed fairly balanced diameter distributions (both PF and PF_L stands), while secondary forests contained virtually no trees above 60 cm diameter.
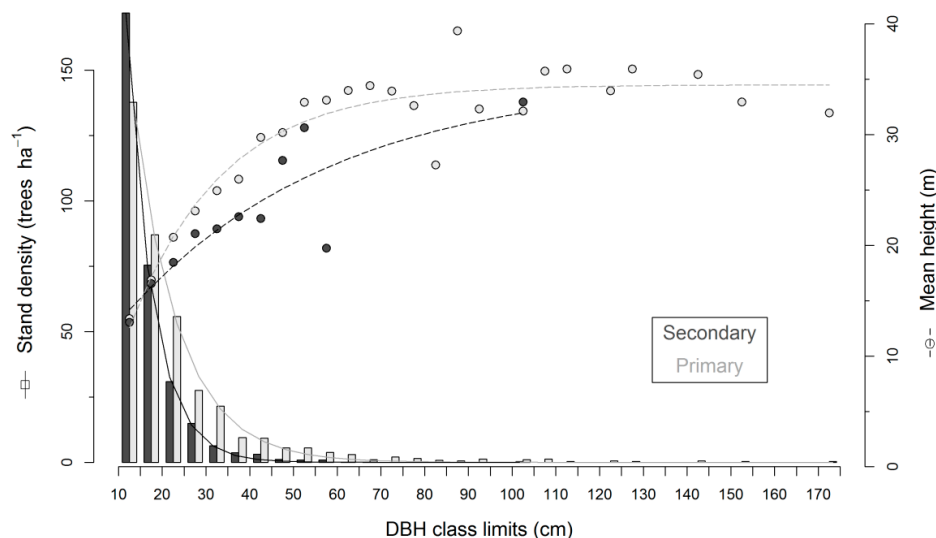

Figure 5. Average stem density (bars) and tree height (circles) per diameter class (trees $\geq$ 10 cm) for secondary (dark gray) and primary (light gray) forests. The solid lines show the fit of Equation (1) to the average diameter distributions, and the dashed lines show the fit of a similar exponential decay model to the tree height data.

The estimated frequency of trees 5–10 cm diameter with Equation (1) averaged 475 ha^{-1} for mid-successional forests and 234 ha^{-1} for primary forests, representing 0.8 to 20% of the total plot

biomass. As shown in Figure 6, these results are consistent with those of stands where trees 5–10 cm diameter were actually measured in the field. For young secondary forests, Figure 6 suggests that the contribution of trees in this class is a linear function of the stand biomass, decreasing rapidly from about 70 to 15% as biomass increases from near 0 to 50 Mg·ha^{-1}. For stands with biomass greater than 50 Mg·ha^{-1}, the contribution of trees 5–10 cm diameter declines exponentially from an initial value of 12%, leveling off at about 1.4% after ~280 Mg·ha^{-1}.

The estimated biomass change for the three-year period between field and GLAS measurements ranged from 7% in the oldest SF stand (~27 years) to 97% in the youngest stand (~4 years). As suggested by Figure 2, biomass accumulation rates varied nonlinearly with stand age (from a low of 0.6 to a maximum of 6.6 Mg·ha^{-1}·yr^{-1}), with the highest rates observed for stands 10 to 14 years old.
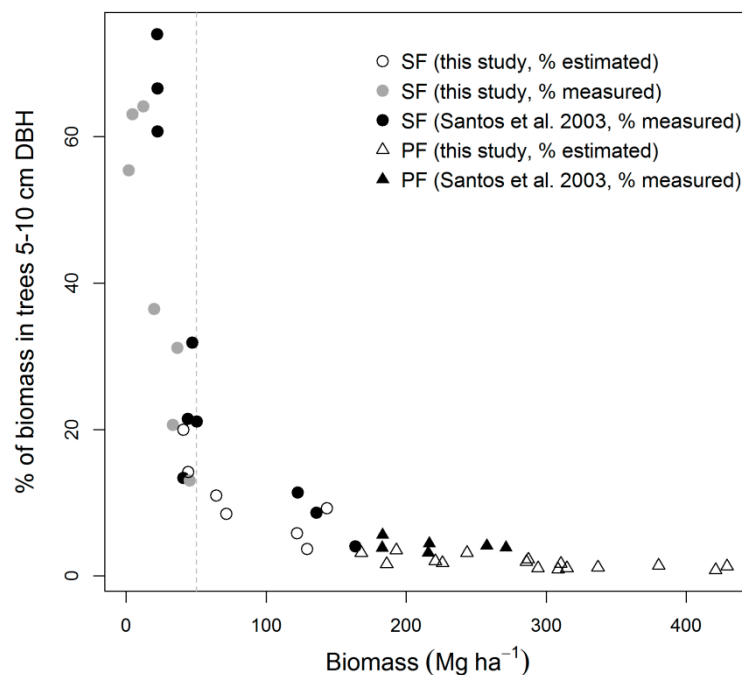

Figure 6. Relationship between stand biomass at Tapajós (trees ≥ 5 cm diameter) and the fraction of that biomass found in trees 5–10 cm diameter. The symbols in black represent plots measured in 2000, as described by [51].

3.3. Biomass Error

The contribution of the different error sources to the overall uncertainty in the field biomass is summarized in Table 4 and detailed below. Figure 7 explores the calculated sensitivity of our binomial approach in Equation (8), showing the dependence of the co-location error on the spatial overlap between field and GLAS samples. The gray and black lines represent the average co-location error for secondary and primary forests, respectively, when the overlap is artificially changed from 0% to 100%. When the overlap is zero, the binomial model yields an average co-location error of 29% of the estimated biomass for SF plots and an error of 42% for PF plots. These errors decrease slowly (and almost linearly) as the overlap increases from 0 to about 60% overlap, and then converge rapidly to zero as the overlap approaches 100%. On average, overlaps ≥ 75% are needed in primary forests to attain co-location errors not exceeding 20%. In secondary forests, this same level of co-location error can be achieved with overlaps ≥ 50%. The estimated overlap between GLAS and our field plots ranged between 50 and 91%, except for one secondary stand where the overlap was zero—the plot missed the GLAS footprint by about 26 m. The resulting co-location errors (σ_C) were typically 13%–26% and dominated the overall uncertainty in both mid-successional and primary stands (Table 4).

Table 4. Uncertainties in field-based estimates of plot biomass at Tapajós. The error is presented in terms of the median value and the interquartile range (in parentheses) of the relative errors of all applicable plots. The last three columns give the percentage of variance in the biomass estimate which is due to each error source. Values are means for early-successional (SF_{early}), mid-successional (SF_{mid}), and primary (PF/PF_L) forest plots.

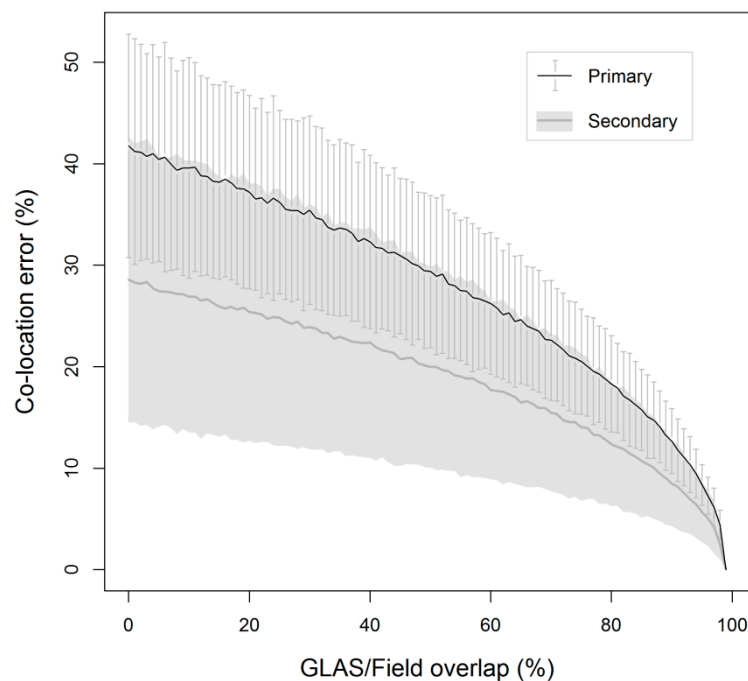| Error Source	Error (%)	% of Total Variance		
		SF_{early}	SF_{mid}	PF/PF_L
Measurement (σ_M)	6.4 (4.3–9.0)	6.4	4.8	7.6
Allometry (model residuals, σ_A)	7.5 (5.0–10.2)	3.7	8.2	12.9
Allometry (model selection, σ_S)	7.1 (4.8–10.6)	9.4	10.6	14.1
Co-location (σ_C)	19.1 (13.0–25.6)	28.3	45.8	65.2
Trees 5–10 cm diameter ($\sigma_{5–10}$)	1.1 (0.7–3.5)	NA	11.6	0.2
Growth model (σ_G)	12.0 (7.0–18.6)	52.2	19.0	NA
Total	25.4 (20.2–33.9)	100	100	100

Figure 7. Dependence of the co-location error estimated with the binomial approach on the spatial overlap between GLAS and field measurements. The lines represent the mean error for secondary (gray) and primary (black) stands when the overlap is artificially changed from 0% to 100%. The gray band and the error bars represent ±1 SD.

Uncertainties in diameter (~2%), height (~15%), and wood density (~14%) resulted in a median error of 25% in the mass of individual trees. Nonetheless, this error dropped to only about 6% when scaled to the plot level (σ_M, Table 4). In secondary forests, the alternative allometric equations of Brown [14] and Chambers [42] (M_{A1} and M_{A2}, Table 1) overestimated the Chave-based plot biomass by an average of 29 and 54%, respectively. In primary forests, the Brown equation showed no systematic bias, whereas the Chambers equation resulted in slight underestimates in high-biomass stands (~11%). The errors associated with the choice of the allometry (σ_S) were typically 5%–11%, similar to the errors related to the model residuals (σ_A). The individual contributions of measurement and allometric errors to the final uncertainty were generally below 15%, and slightly lower in secondary forests than in primary forests (Table 4).

To gain a better understanding of the uncertainty associated with allometric methods, we also compared our reference biomass estimates obtained with the Chave equation M_{T2} (Table 1) using field-measured D and H_T, and taxon-specific ρ derived from the literature, with four alternative estimates produced with the Chave equations, but making small changes in the input data to illustrate the variability that can be expected when common, suboptimal field data sets are used: Use of the Chave equation M_{T2} with a regional average wood density of 0.667 g·cm^{-3} [39], as opposed to taxon-specific densities, resulted in overestimation of biomass values by about 23% in secondary forests and no bias in primary forests. When M_{T2} was applied using taxon-specific wood densities, but heights derived from a regional height-diameter relationship [52], the plot-level biomass was 11% lower on average due to a negative bias in height of 2.6 m. Use of the Chave equation without the height term (M_{A3}, Table 1) resulted in plot-level biomass ~20% higher, regardless of the successional status. When this equation was applied using the regional average wood density of 0.667 g·cm^{-3}, the discrepancies in secondary forests were higher still (48%). We should note that the biomass of cecropias and palms, estimated by the specific equations M_{T1}, M_{P1}, and M_{P2} (Table 1), was held constant across all comparisons. Although this introduced some dependence across biomass estimates, these species typically accounted for only about 3% of the total plot biomass.

In primary forests, where the minimum measured diameter was 10 cm, the error of estimating biomass for the 5–10 cm diameter class (σ_{5-10}) contributed less than 1% to the total variance and could safely be neglected. However, this error was about seven times larger in mid-successional forests, being comparable to other sources in magnitude (Table 4). In secondary forests, the projection of biomass values backward in time three years induced errors (σ_G) of the order of 7%–19%. This term dominated the uncertainties in early successional stands, accounting for about half of the total variance on average, and represented the second largest component in mid-successional stands (Table 4). The dependence of σ_G on the temporal difference between field and remote sensing acquisitions is illustrated in Figure 8 for SF plots of different ages. As expected, σ_G increases significantly with the time gap in data acquisition. The increase is faster for younger forests, which display higher values of σ_G than older forests at any given temporal interval (the greater the relative change in biomass, the greater the uncertainty in the model estimate).
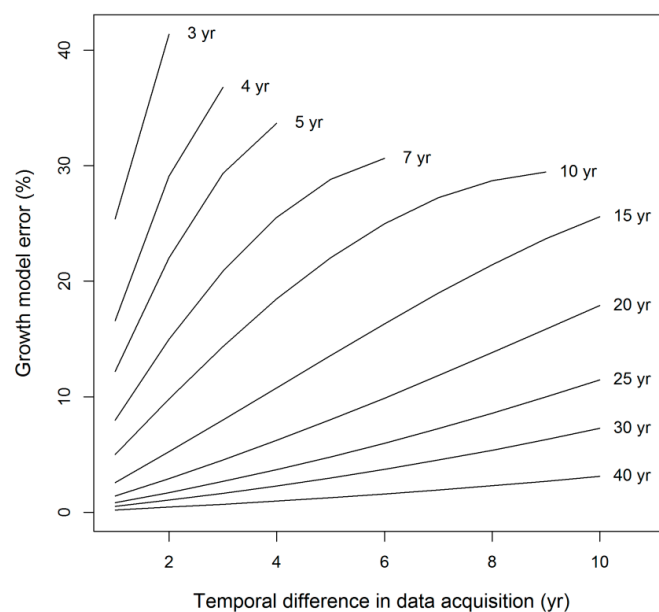

Figure 8. Dependence of the biomass error introduced by the growth model of [44] on the temporal difference between field and remote sensing acquisitions (i.e., remote sensing data acquired 1, 2, 3, ..., 10 years prior to the field measurements). The dependence is illustrated for secondary forests of different ages, as indicated by the labels to the right of the lines.

The overall uncertainty in the field biomass was typically 25% (for both secondary and primary forests), but ranged from 16% to 53%. Measurement (σ_M), allometric (both σ_A and σ_S), and co-location (σ_C) errors increased significantly with plot biomass, at rates of 8, 10, 11, and 22%, respectively (Figure 9). The error of estimating biomass for the 5–10 cm diameter class (not shown in the figure) increased significantly with biomass in secondary forests (at a rate of 10%), but showed no trend in old-growth forests. The growth model error showed no linear trend with biomass, but was a straight-line function of the biomass accumulation rate, increasing by about 0.7 Mg·ha^{-1} for each unit increase in the growth rate. As a result of the above trends, the overall uncertainty also increased with biomass, at a combined rate of 28%. However, there was no evidence that the mean errors or rates of error increase differed among forest types, after accounting for differences in biomass.
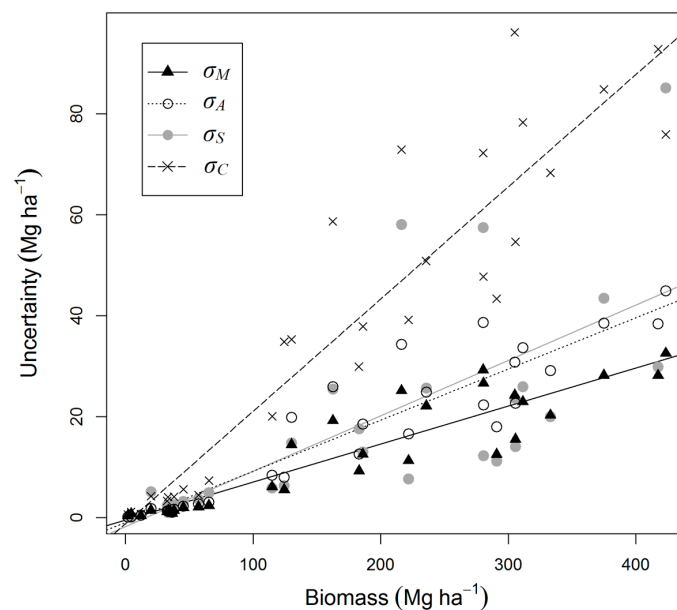

Figure 9. Stand biomass versus measurement (σ_M), allometric (σ_A and σ_S), and co-location (σ_C) errors, with the estimated regression lines. Sources of error that are not common to all plots (i.e., σ_{5-10} and σ_G) were omitted from the figure for clarity. By looking at points intersected by an imaginary vertical line at any level of biomass in the x-axis, one can see the relative contribution of the different error sources for the plot represented by that biomass.

4. Discussion

4.1. Precision of Individual Tree Measurements

Tree diameter measurements are not difficult to obtain, involve limited subjectivity, and can usually be made with a high degree of precision (e.g., [53–56]). The small variation in diameter measurements observed in this study (RMSD = 0.8 cm or 1.8%) is consistent with previous findings and likely resulted from divergences in tape placement, with repeated measurements taken at slightly different tree heights or angles. Other potential sources of variation include mistakes reading the tape, recording error, and data entry error, all of which are difficult to detect if the resulting values are not particularly unusual.

Despite the obvious subjectivity associated with ocular height estimates (both H_C and H_T), they were surprisingly precise, with a combined RMSD of only 2 m [19]. For comparison, Kitahara et al. [56] reported nearly the same level of precision (1.8 m) for repeated height measurements made with a modern ultrasonic hypsometer (Haglöf Vertex) in less dense temperate forests with relatively lower structural complexity. In a recent study also conducted at Tapajós, Hunter et al. [57] obtained a precision of 4.7 m for heights obtained with a clinometer and a measuring tape, keeping angles below 50°

and correcting for slope to minimize measurement error. Factors contributing to variation in our height measurements include difficulty in determining the location of the treetop due to occlusion by surrounding vegetation, and disparity in the perception of where the base of the crown is located.

Variation was considerably higher for crown measurements. Crown depth, calculated as the difference between H_T and H_C, was very similar in absolute precision to the ocular height estimates. However, while a RMSD of 2 m typically represents a relatively small percentage of a tree height, it corresponds to a large fraction of a typical crown depth measurement (6 m on average for the trees sampled in this study). The precision of crown radius measurements was better than 1 m, but represented ~26% in relative terms. These measurements required some level of personal judgment and were affected by visibility restrictions in ways similar to the height estimates. In addition, crown spread was typically 25%–45% of the tree height and the resulting high levels of crown overlap among trees made it often challenging to identify the correct branches for the measurement. We note that the horizontal position of the crown edge is somewhat difficult to determine from directly below and suggest that the precision of crown radius measurements would likely be improved by sighting the edges along a clinometer held at a 90-degree angle. In terms of height measurements (and the derived crown depth), uncertainties may be reduced with the aid of a telescoping height measuring pole. Although not necessarily practical, the pole could be used to obtain direct height measurements for small trees (up to 10–15 m), and serve as a height reference for the ocular estimation of taller trees.

Not surprisingly, measurements of the attributes in Table 2 were more precise for small trees than for large trees (most sources of measurement variation become more pronounced as tree size increases). Although measurement variation generally increased with the dimension of the measurement, the magnitude of this effect differed substantially among attributes, with stem diameter showing the lowest rate of increase, followed by height, and crown dimensions. For tree height, Hunter et al. [57] observed an eightfold increase in measurement variation (from 1.1 to 8.2 m) after dividing the data into four diameter classes with an equal number of trees. This contrasts sharply with the less than twofold increase observed in this study (from 1.8 to 2.9 m), indicating that the precision of the ocular height estimates was not only high, but also displayed a relatively low, yet statistically significant, dependence on tree height.

Because precision is not constant across the range of diameters and heights, it is important to account for this variation when propagating measurement errors to determine the uncertainty in biomass. The standard deviation of the differences between repeated measurements, calculated by quartiles of the ranked set of measurements, is provided in Table A1 for reference.

While measurement uncertainty was generally not negligible, with precision clearly declining with increasing tree size, we found no systematic errors. In addition, we found no differences in precision (or rates of decline in precision with increasing tree size) between primary and secondary forests, after accounting for tree size. This suggests that measurement precision was fairly robust to changes in measurement conditions (induced by changes in stem density, species composition, leaf area index, etc.), with divergences in overall precision being largely attributable to differences in tree size distribution (see Figure 5). We stress that the results presented here refer strictly to reproducibility of measurements, and that no reference is made to the agreement of those measurements with the true, unknown values (i.e., accuracy).

4.2. Biomass Estimation and Its Error

Our results show that co-location error, defined in this study as the uncertainty in the biomass estimate resulting from the spatial disagreement between field and Lidar samples (i.e., field plots including trees not captured by GLAS and/or excluding trees that were actually captured), accounts for a substantial portion of the total error. In agreement with our findings for stands at La Selva Biological Station, Costa Rica [20], co-location error dominated the overall uncertainty in the field biomass, except in early-successional forests where the application of the growth model resulted in larger errors on average (Table 4). The results illustrated in Figure 7 are consistent with the expectation

of lower co-location error with increasing Lidar/field overlap, as well as lower errors for secondary forests compared to primary forests, given their lower species diversity and more homogeneous canopy structure (cf. Figure 5). We should note that the binomial approach in (8) assumes that the tree size-frequency distribution of each individual zone depicted in Figure 3 (i.e., F, I, and G) is similar to that observed for the full 50 m × 50 m field plot (F ∪ I). This explains the relatively low maximum values of co-location error in Figure 7 when the overlap is zero. We should also note that although the binomial approach is presented here using GLAS as an example, it depends exclusively on the field data and on the amount of overlap between the field and the remote sensing samples, and thus could be applied regardless of the remote sensing data type.

In a previous study in a tropical rainforest in Hawaii, Asner et al. [49] found that misalignment of Lidar and field data introduced errors in biomass estimates of only 0–10 Mg·ha^{-1} (0%–3.5% of the median biomass). While differences in floristic composition, vegetation structure, and plot size make direct comparisons between studies difficult, differences in methodology most likely account for much of the observed discrepancy. In their study, Asner et al. estimated the co-location error by varying the location and size of the Lidar "plots" by small amounts (10%); regressing the Lidar metrics obtained for each new location/size against the (fixed) field-measured biomass; and determining the variation in the biomass predictions resulting from variation in the Lidar metrics. This is conceptually different from the approach used in this study, where the field-based estimate of biomass was varied instead. This distinction is important because inclusion/exclusion of trees in the calibration plots (particularly big trees) can cause large changes in the field estimate of biomass that may not be proportionally reflected in the vertical structure captured by the Lidar (and in turn, in the Lidar estimate of biomass). As shown by [58], even relatively small changes in field-based estimates of biomass in tropical forests (as a result of accounting for portions of trees that fall outside the plot boundary) can have a significant impact on the relationship with Lidar metrics, accounting for as much as 55% of the error associated with Lidar-biomass models.

In comparison to the co-location error, measurement and allometric errors were relatively small. Uncertainties in height and wood density values were large relative to the uncertainty in diameter and contributed the bulk of the uncertainty in biomass due to measurement variation. This measurement error was fairly large at the tree level, but decreased significantly at the plot scale because measurement variation was unbiased (Table 2) and tree-level errors were added in quadrature to produce realistic plot-level estimates.

The overestimation of the reference, Chave-based biomass in secondary forests by the equations of Brown [14] and Chambers [42] (M_{A1} and M_{A2}, Table 1) was largely explained by the omission of wood density information in the models. These alternative mixed-species equations were derived from primary forest trees, which tend to have much denser wood than the secondary forest trees to which they were applied (cf. Table 3 and [41]). When we corrected M_{A1} and M_{A2} by including a dependence on wood density as in [24], the overestimation of the reference biomass in secondary forests decreased by a factor of 3 and 2, respectively. From our tests with the Chave equations with and without height, we would expect these differences to decrease by an additional ~20% if M_{A1} and M_{A2} also included a dependence on tree height, and if tree allometry is somewhat conserved across moist tropical sites as indicated by [16]. Thus, most of the variation captured by σ_S was apparently due to the use of allometric equations, which differed with respect to the inclusion of height and wood density information.

As with the measurement error, allometric errors were assumed to be uncorrelated and decreased significantly (by a factor of 3–4) when scaled to the plot level. While this assumption seems reasonable for σ_A (cf. [24–26,32]) given the random nature of the regression errors (assumed to be normally distributed with mean zero), one could argue that the error due to the choice of the allometric equation (σ_S) is systematic and unlikely to be independent (trees with similar diameter, for example, can have nearly the same σ_S). One way of testing if the sum in quadrature is appropriate is to calculate σ_S directly at the plot level by taking the standard deviation of the plot-level biomass estimates obtained

with the alternative equations. This resulted in a median error of 13% for our 30 plots, which is only slightly higher than that obtained by adding tree-level errors in quadrature (Table 4). The error for primary forests alone was virtually the same (9%), confirming the tendency of individual tree errors to offset each other when combined to generate plot-level estimates. We note that [33] observed the same level of error (10%) when estimating the biomass density (trees ≥ 15 cm diameter) of an area of 392 ha at Tapajós using four alternative equations (one of which is M_{A1}). A similar error (13%) was also reported by [24] for estimates obtained with eight different equations in a 50 ha plot in Panama, after correcting for variation in wood density.

The estimation of biomass for the modeled diameter class of 5–10 cm made a negligible contribution to the final uncertainty in primary forests, but represented a significant source of error in mid-successional forests. This is not surprising when one considers the rapid decrease in the contribution of trees 5–10 cm diameter to biomass with increasing biomass, as observed in Figure 6. Because in young secondary forests (<50 Mg·ha^{-1}) trees 5–10 cm diameter can account for as much as 70% of the aboveground biomass, modeling the stem frequency of this class with Equation (1) and estimating biomass for the modeled class has the potential to introduce large errors. For these young forests, we recommend that trees 5–10 cm diameter be directly measured (or tallied) in the field.

The biomass accumulation rates of secondary forests estimated with the site-specific growth model of Neeff & Santos [44] agreed well with rates from a long-term study of SF regrowth in the central Amazon [59], as well as with a more general estimate for tropical moist forests based on data from a number of sites [60]. However, the error associated with this model dominated the overall uncertainty in the biomass of the youngest stands and represented the second largest contribution in older secondary forests. In (2), stand biomass is defined as the product of basal area and top height (the two state variables that are explicitly modeled), and most of its uncertainty results from the relatively poor fit of the basal area model (Figure 2). Although site-specific, the model is based on growth rates inferred from a chronosequence, which may differ substantially from actual growth rates due to stand differences in land-use history (e.g., old agricultural fields vs. abandoned pastures) [59]. This probably explains most of the variability observed by Neeff & Santos [44] in basal area for a given age, and the relatively large uncertainty associated with the model parameters.

The estimated biomass change of over 30% per year for the youngest secondary forests emphasizes the importance of accounting for temporal differences between field and remote sensing data. Here, we assumed no biomass change for primary forests, but these changes have been determined to be small at Tapajós (~1% per year) relative to the average biomass [35,61,62]. We would expect the resulting errors to be negligible, unless some disturbance occurred between observation epochs (e.g., logging, fire, windstorms, etc.). This is generally detectable in the field and/or in remote sensing data [63], and neither data source revealed the occurrence of a significant disturbance event during the period of this study.

The overall uncertainty in the plot-level biomass observed in this study (typically 20%–34%) was similar in relative size to that observed by Treuhaft et al. [20] for a tropical wet forest in Costa Rica (26%–31%) where measurement, allometric, and co-location errors were also considered. It is also in line with the total plot-level error observed by Chave et al. [24] (~23%–27%, as calculated from their Table 3) and Chen et al. [26] (20%), although different methods were used for the assessment of uncertainty. Recent estimates of biomass from Lidar [19] and InSAR [21] observations for the plots used in this study showed RMS errors about the field-estimated biomass of 20%–35%. The uncertainty in the field biomass can therefore represent a dominant term in the overall error budget for biomass maps derived from remote sensing, and should be taken into account in calibration and validation efforts.

Because co-location and temporal errors have the potential to account for a large fraction of the total variance (~70% on average for the plots in this study, but as much as 94% in individual plots), they emerge as obvious targets for reducing uncertainty in studies relating tropical forest biomass to remotely sensed data. Temporal errors can be minimized by conducting field campaigns as close as possible in time to the remote sensing data acquisition. In cases where this is not possible (e.g., when

using historical data), our results underscore the importance of selecting an appropriate growth model to account for biomass change in secondary forests, and of quantifying the uncertainty associated with this model. Reducing co-location errors requires not only the acquisition of high-precision differential GPS measurements for plot location (often complemented with topographic surveying in closed-canopy forests), but also that field and remote sensing samples agree as much as possible in size, shape, and orientation. A simple statistical approach such as the one presented in Section 2.4.2 can be used to account for errors due to partial overlap.

Finally, we note that although measurement and allometric errors were relatively unimportant when considered alone, combined they accounted for roughly 30% of the total variance on average (as much as 64% in individual plots) and should not be ignored. Steps can be taken to reduce uncertainties in height and wood density measurements, as well as in allometric equations. However, reducing co-location and temporal errors may be a more cost-effective solution for reducing the overall uncertainty when resources are limited. For instance, the total error in Table 4 would drop by nearly half if co-location and temporal errors were zero, but only by about 20% if we disregard measurement and allometric errors instead.

5. Conclusions

The objective of this study was to use field plot data collected in the central Amazon to gain a better understanding of the uncertainty associated with plot-level biomass estimates obtained specifically for calibration of remote sensing measurements in tropical forests (see [19] for details on the calibration performed using the plots of this study). We found that the overall uncertainty in the field biomass was typically 25% for both secondary and primary forests, but ranged from 16% to 53%. Co-location and temporal errors accounted for a large fraction of the total variance (>65%) when compared to sources of error that are commonly assessed in conventional biomass estimates, emerging as important targets for reducing uncertainty in studies relating tropical forest biomass to remotely sensed data. Although measurement and allometric errors were relatively unimportant when considered alone, combined they accounted for roughly 30% of the total variance on average and should not be ignored. Our results suggest that a thorough understanding of the sources of error associated with plot-level biomass estimates in tropical forests is critical to determine confidence in remote sensing estimates of carbon stocks and fluxes, and to develop strategies for reducing the overall uncertainty of remote sensing approaches.

Author Contributions: F.G. conceived, designed, and performed the experiment; analyzed the data; and wrote the paper. R.T. assisted with study design, data collection, data analysis, and writing of the paper. B.L. assisted with study design, interpretation of results, and writing of the paper. A.A. assisted with data collection and contributed to parts of the analysis. W.W. and A.B. assisted with data analysis and writing of the paper. J.R.S. and P.G. assisted with data collection and writing of the paper.

Appendix A

Table A1. Standard deviations (SD) of differences in repeated measurements of diameter (*D*), height to the base of the live crown (*H_C*), total height (*H_T*), crown depth (*C_D*), and crown radius (*C_R*) for trees sampled at Tapajós. Values are shown by quartile of the ranked set of measurements, in both absolute and relative terms (see Section 2.4.1 for details). The total number of observations was 104, except for *C_R* (n = 144).

Attribute	Probabilities			
	0%–25%	**25%–50%**	**50%–75%**	**75%–100%**
D (cm)				
Quantiles	5.5–12.3	12.3–16.1	16.1–26.1	26.1–110
SD	0.1 (1.4%)	0.1 (0.9%)	0.3 (1.3%)	1.6 (2.8%)
H_C (m)				
Quantiles	1.5–6	6–9.3	9.3–12.3	12.3–31
SD	1 (22.9%)	1.2 (15.1%)	2.1 (19.6%)	2.4 (13.1%)
H_T (m)				
Quantiles	5–11.4	11.4–14.5	14.5–19.5	19.5–40
SD	1.5 (17.9%)	2.2 (17.4%)	2.3 (14.3%)	2.9 (10.2%)
C_D (m)				
Quantiles	1–4	4–6	6–8.5	8.5–20
SD	0.9 (36.8%)	1.6 (34.3%)	2 (29.6%)	2.4 (21.9%)
C_R (m)				
Quantiles	0.7–1.6	1.6–2.3	2.3–3.5	3.5–8
SD	0.3 (26.5%)	0.4 (21.8%)	0.8 (30.1%)	1.2 (25.2%)

References

1. Lefsky, M.A.; Cohen, W.B.; Parker, G.G.; Harding, D.J. Lidar remote sensing for ecosystem studies. *Bioscience* **2002**, *52*, 19–30. [CrossRef]
2. Treuhaft, R.N.; Law, B.E.; Asner, G.P. Forest attributes from radar interferometric structure and its fusion with optical remote sensing. *Bioscience* **2004**, *54*, 561–571. [CrossRef]
3. Zolkos, S.G.; Goetz, S.J.; Dubayah, R. A meta-analysis of terrestrial aboveground biomass estimation using Lidar remote sensing. *Remote Sens. Environ.* **2013**, *128*, 289–298. [CrossRef]
4. Asner, G.P.; Powell, G.V.N.; Mascaro, J.; Knapp, D.E.; Clark, J.K.; Jacobson, J.; Kennedy-Bowdoin, T.; Balaji, A.; Paez-Acosta, G.; Victoria, E.; et al. High-resolution forest carbon stocks and emissions in the Amazon. *Proc. Natl. Acad. Sci. USA.* **2010**, *107*, 16738–16742. [CrossRef] [PubMed]
5. Baccini, A.; Goetz, S.J.; Walker, W.S.; Laporte, N.T.; Sun, M.; Sulla-Menashe, D.; Hackler, J.; Beck, P.S.A.; Dubayah, R.; Friedl, M.A.; et al. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nat. Clim. Chang.* **2012**, *2*, 182–185. [CrossRef]
6. Saatchi, S.S.; Harris, N.L.; Brown, S.; Lefsky, M.; Mitchard, E.T.A.; Salas, W.; Zutta, B.R.; Buermann, W.; Lewis, S.L.; Hagen, S.; et al. Benchmark map of forest carbon stocks in tropical regions across three continents. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 9899–9904. [CrossRef] [PubMed]
7. Houghton, R.A. Aboveground forest biomass and the global carbon balance. *Glob. Chang. Biol.* **2005**, *11*, 945–958. [CrossRef]
8. Houghton, R.; Lawrence, K.; Hackler, J.; Brown, S. The spatial distribution of forest biomass in the Brazilian Amazon: A comparison of estimates. *Glob. Chang. Biol.* **2001**, *7*, 731–746. [CrossRef]
9. Malhi, Y.; Wood, D.; Baker, T.R.; Wright, J.; Phillips, O.L.; Cochrane, T.; Meir, P.; Chave, J.; Almeida, S.; Arroyo, L. The regional variation of aboveground live biomass in old-growth Amazonian forests. *Glob. Chang. Biol.* **2006**, *12*, 1107–1138. [CrossRef]
10. Asner, G.P. Tropical forest carbon assessment: Integrating satellite and airborne mapping approaches. *Environ. Res. Lett.* **2009**, *4*, 34009. [CrossRef]

11. Houghton, R.A.; Hall, F.; Goetz, S.J. Importance of biomass in the global carbon cycle. *J. Geophys. Res.* **2009**, *114*. [CrossRef]

12. Goetz, S.; Dubayah, R. Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. *Carbon Manag.* **2011**, *2*, 231–244. [CrossRef]

13. Hall, F.G.; Bergen, K.; Blair, J.B.; Dubayah, R.; Houghton, R.; Hurtt, G.; Kellndorfer, J.; Lefsky, M.; Ranson, J.; Saatchi, S.; et al. Characterizing 3D vegetation structure from space: Mission requirements. *Remote Sens. Environ.* **2011**, *115*, 2753–2775. [CrossRef]

14. Brown, S. *Estimating Biomass and Biomass Change of Tropical Forests: A Primer*; Food and Agriculture Organization of the United Nations (FAO): Rome, Italy, 1997.

15. Chave, J.; Andalo, C.; Brown, S.; Cairns, M.A.; Chambers, J.Q.; Eamus, D.; Fölster, H.; Fromard, F.; Higuchi, N.; Kira, T.; et al. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia* **2005**, *145*, 87–99. [CrossRef] [PubMed]

16. Chave, J.; Réjou-Méchain, M.; Búrquez, A.; Chidumayo, E.; Colgan, M.S.; Delitti, W.B.C.; Duque, A.; Eid, T.; Fearnside, P.M.; Goodman, R.C.; et al. Improved allometric models to estimate the aboveground biomass of tropical trees. *Glob. Chang. Biol.* **2014**, *20*, 3177–3190. [CrossRef] [PubMed]

17. Asner, G.P.; Mascaro, J. Mapping tropical forest carbon: Calibrating plot estimates to a simple Lidar metric. *Remote Sens. Environ.* **2014**, *140*, 614–624. [CrossRef]

18. Drake, J.B.; Knox, R.G.; Dubayah, R.O.; Clark, D.B.; Condit, R.; Blair, J.B.; Hofton, M. Above-ground biomass estimation in closed canopy neotropical forests using Lidar remote sensing: Factors affecting the generality of relationships. *Glob. Ecol. Biogeogr.* **2003**, *12*, 147–159. [CrossRef]

19. Gonçalves, F.G. Vertical Structure and Aboveground Biomass of Tropical Forests from Lidar Remote Sensing. Ph.D. Thesis, Oregon State University, Corvallis, OR, USA, November 2014.

20. Treuhaft, R.N.; Gonçalves, F.G.; Drake, J.B.; Chapman, B.D.; dos Santos, J.R.; Dutra, L.V.; Graça, P.M.L.A.; Purcell, G.H. Biomass estimation in a tropical wet forest using Fourier transforms of profiles from Lidar or interferometric SAR. *Geophys. Res. Lett.* **2010**, *37*. [CrossRef]

21. Treuhaft, R.; Gonçalves, F.; Santos, J.R.; Keller, M.; Palace, M.; Madsen, S.N.; Sullivan, F.; Graca, P.M.L.A. Tropical-forest biomass estimation at X-band from the spaceborne TanDEM-X interferometer. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 239–243. [CrossRef]

22. Clark, D.B.; Kellner, J.R. Tropical forest biomass estimation and the fallacy of misplaced concreteness. *J. Veg. Sci.* **2012**, *23*, 1191–1196. [CrossRef]

23. Araújo, T.M.; Higuchi, N.; Carvalho, J.A. Comparison of formulae for biomass content determination in a tropical rain forest site in the state of Pará, Brazil. *For. Ecol. Manag.* **1999**, *117*, 43–52. [CrossRef]

24. Chave, J.; Condit, R.; Aguilar, S.; Hernandez, A.; Lao, S.; Perez, R. Error propagation and scaling for tropical forest biomass estimates. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2004**, *359*, 409–420. [CrossRef] [PubMed]

25. Ahmed, R.; Siqueira, P.; Hensley, S.; Bergen, K. Uncertainty of forest biomass estimates in north temperate forests due to allometry: Implications for remote sensing. *Remote Sens.* **2013**, *5*, 3007–3036. [CrossRef]

26. Chen, Q.; Laurin, G.V.; Valentini, R. Uncertainty of remotely sensed aboveground biomass over an African tropical forest: Propagating errors from trees to plots to pixels. *Remote Sens. Environ.* **2015**, 134–143. [CrossRef]

27. Colgan, M.S.; Asner, G.P.; Swemmer, T. Harvesting tree biomass at the stand level to assess the accuracy of field and airborne biomass estimation in savannas. *Ecol. Appl.* **2013**, *23*, 1170–1184. [CrossRef] [PubMed]

28. Bevington, P.R. *Data Reduction and Error Analysis for the Physical Sciences*; Bruflodt, D., Ed.; McGraw-Hill: New York, NY, USA, 1969; Volume 336.

29. Ramsey, F.; Schafer, D. *The Statistical Sleuth: A Course in Methods of Data Analysis*, 2nd ed.; Duxbury/Thomson Learning: Pacific Grove, CA, USA, 2002.

30. Chen, Q.; McRoberts, R.E.; Wang, C.; Radtke, P.J. Forest aboveground biomass mapping and estimation across multiple spatial scales using model-based inference. *Remote Sens. Environ.* **2016**, *184*, 350–360. [CrossRef]

31. Brown, I.F.; Martinelli, L.A.; Thomas, W.W.; Moreira, M.Z.; Cid Ferreira, C.A.; Victoria, R.A. Uncertainty in the biomass of Amazonian forests: An example from Rondônia, Brazil. *For. Ecol. Manag.* **1995**, *75*, 175–189. [CrossRef]

32. Phillips, D.L.; Brown, S.L.; Schroeder, P.E.; Birdsey, R.A. Toward error analysis of large-scale forest carbon budgets. *Glob. Ecol. Biogeogr.* **2000**, *9*, 305–313. [CrossRef]

33. Keller, M.; Palace, M.; Hurtt, G. Biomass estimation in the Tapajos National Forest, Brazil examination of sampling and allometric uncertainties. *For. Ecol. Manag.* **2001**, *154*, 371–382. [CrossRef]

34. Molto, Q.; Rossi, V.; Blanc, L. Error propagation in biomass estimation in tropical forests. *Methods Ecol. Evol.* **2013**, *4*, 175–183. [CrossRef]

35. Vieira, S.; de Camargo, P.B.; Selhorst, D.; da Silva, R.; Hutyra, L.; Chambers, J.Q.; Brown, I.F.; Higuchi, N.; dos Santos, J.; Wofsy, S.C.; et al. Forest structure and carbon dynamics in Amazonian tropical rain forests. *Oecologia* **2004**, *140*, 468–479. [CrossRef] [PubMed]

36. Silver, W.L.; Neff, J.; McGroddy, M.; Veldkamp, E.; Keller, M.; Cosme, R. Effects of soil texture on belowground carbon and nutrient storage in a lowland Amazonian forest ecosystem. *Ecosystems* **2000**, *3*, 193–209. [CrossRef]

37. Gonçalves, F.; Santos, J. Composição florística e estrutura de uma unidade de manejo florestal sustentável na Floresta Nacional do Tapajós, Pará. *Acta Amaz.* **2008**, *38*, 229–244. [CrossRef]

38. Reyes, G.; Brown, S.; Chapman, J.; Lugo, A. *Wood Densities of Tropical Tree Species*; General Technical Report (GTR) SO-88; USDA Forest Service, Southern Forest Experiment Station, Institute of Tropical Forestry: New Orleans, LA, USA, February 1992.

39. Chave, J.; Muller-Landau, H.C.; Baker, T.R.; Easdale, T.A.; ter Steege, H.; Webb, C.O. Regional and phylogenetic variation of wood density across 2456 neotropical tree species. *Ecol. Appl.* **2006**, *16*, 2356–2367. [CrossRef]

40. Larjavaara, M.; Muller-Landau, H.C. Measuring tree height: a quantitative comparison of two common field methods in a moist tropical forest. *Methods Ecol. Evol.* **2013**, *4*, 793–801. [CrossRef]

41. Nelson, B.W.; Mesquita, R.; Pereira, J.L.G.; de Souza, S.G.A.; Batista, G.T.; Couta, L.B. Allometric regressions for improved estimate of secondary forest biomass in the Central Amazon. *For. Ecol. Manag.* **1999**, *177*, 149–167. [CrossRef]

42. Chambers, J.Q.; Dos Santos, J.; Ribeiro, R.J.; Higuchi, N. Tree damage, allometric relationships, and above-ground net primary production in central Amazon forest. *For. Ecol. Manag.* **2001**, *152*, 73–84. [CrossRef]

43. Van Laar, A.; Akça, A. *Forest Mensuration*; Springer: Dordrecht, The Netherlands, 2007.

44. Neeff, T.; dos Santos, J.R. A growth model for secondary forest in Central Amazonia. *For. Ecol. Manag.* **2005**, *216*, 270–282. [CrossRef]

45. Brown, S.; Gillespie, A.; Lugo, A. Biomass estimation methods for tropical forests with applications to forest inventory data. *For. Sci.* **1989**, *35*, 881–902.

46. Anderson, A.B. *The Biology of Orbignya Martiana (Palmae), A Tropical Dry Forest Dominant in Brazil*; UMI Dissertation Information Service, University Microfilms International: Gainesville, FL, USA, 1983.

47. Saldarriaga, J.; West, D.; Tharp, M.; Uhl, C. Long-term chronosequence of forest succession in the upper Rio Negro of Colombia and Venezuela. *J. Ecol.* **1988**, *76*, 938–958. [CrossRef]

48. Baskerville, G. Use of logarithmic regression in the estimation of plant biomass. *Can. J. For. Res.* **1972**, *2*, 49–53. [CrossRef]

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

50. Wang, G.; Zhang, M.; Gertner, G.Z.; Oyana, T.; McRoberts, R.E.; Ge, H. Uncertainties of mapping aboveground forest carbon due to plot locations using national forest inventory plot and remotely sensed data. *Scand. J. For. Res.* **2011**, *26*, 360–373. [CrossRef]

51. Santos, J.R.; Freitas, C.C.; Araujo, L.S.; Dutra, L.V.; Mura, J.C.; Gama, F.F.; Soler, L.S.; Sant'Anna, S.J.S. Airborne P-band SAR applied to the aboveground biomass studies in the Brazilian tropical rainforest. *Remote Sens. Environ.* **2003**, *87*, 482–493. [CrossRef]

52. Feldpausch, T.R.; Banin, L.; Phillips, O.L.; Baker, T.R.; Lewis, S.L.; Quesada, C.A.; Affum-Baffoe, K.; Arets, E.J.M.M.; Berry, N.J.; Bird, M.; et al. Height-diameter allometry of tropical forest trees. *Biogeosciences* **2011**, *8*, 1081–1106. [CrossRef]

53. McRoberts, R.E.; Hahn, J.T.; Hefty, G.J.; Cleve, J.R. Variation in forest inventory field measurements. *Can. J. For. Res.* **1994**, *24*, 1766–1770. [CrossRef]

54. Barker, J.R.; Bollman, M.; Ringold, P.L.; Sackinger, J.; Cline, S.P. Evaluation of metric precision for a riparian forest survey. *Environ. Monit. Assess.* **2002**, *75*, 51–72. [CrossRef] [PubMed]

55. Elzinga, C.; Shearer, R.C.; Elzinga, G. Observer variation in tree diameter measurements. *West. J. Appl. For.* **2005**, *20*, 134–137.

56. Kitahara, F.; Mizoue, N.; Yoshida, S. Evaluation of data quality in Japanese national forest inventory. *Environ. Monit. Assess.* **2009**, *159*, 331–340. [CrossRef] [PubMed]

57. Hunter, M.O.; Keller, M.; Victoria, D.; Morton, D.C. Tree height and tropical forest biomass estimation. *Biogeosciences* **2013**, *10*, 8385–8399. [CrossRef]

58. Mascaro, J.; Detto, M.; Asner, G.P.; Muller-Landau, H.C. Evaluating uncertainty in mapping forest carbon with airborne Lidar. *Remote Sens. Environ.* **2011**, *115*, 3770–3774. [CrossRef]

59. Feldpausch, T.R.; Prates-Clark, C.D.C.; Fernandes, E.C.M.; Riha, S.J. Secondary forest growth deviation from chronosequence predictions in central Amazonia. *Glob. Chang. Biol.* **2007**, *13*, 967–979. [CrossRef]

60. Silver, W.L.; Ostertag, R.; Lugo, A.E. The potential for carbon sequestration through reforestation of abandoned tropical agricultural and pasture lands. *Restor. Ecol.* **2000**, *8*, 394–407. [CrossRef]

61. Rice, A.H.; Pyle, E.H.; Saleska, S.R.; Hutyra, L.; Palace, M.; Keller, M.; de Camargo, P.B.; Portilho, K.; Marques, D.F.; Wofsy, S.C. Carbon balance and vegetation dynamics in an old-growth Amazonian forest. *Ecol. Appl.* **2004**, *14*, 55–71. [CrossRef]

62. Pyle, E.H.; Santoni, G.W.; Nascimento, H.E.M.; Hutyra, L.R.; Vieira, S.; Curran, D.J.; van Haren, J.; Saleska, S.R.; Chow, V.Y.; Carmago, P.B.; et al. Dynamics of carbon, biomass, and structure in two Amazonian forests. *J. Geophys. Res. Biogeosci.* **2008**, *113*. [CrossRef]

63. Frolking, S.; Palace, M.W.; Clark, D.B.; Chambers, J.Q.; Shugart, H.H.; Hurtt, G.C. Forest disturbance and recovery: A general review in the context of spaceborne remote sensing of impacts on aboveground biomass and canopy structure. *J. Geophys. Res. Biogeosci.* **2009**, *114*. [CrossRef]