

Biomass equations for forest regrowth in the eastern Amazon using randomized branch sampling

Mark Jonathan DUCEY¹, Daniel Jacob ZARIN², Steel Silva VASCONCELOS³, Maristela Machado ARAÚJO⁴

ABSTRACT

Forest regrowth occupies an extensive and increasing area in the Amazon basin, but accurate assessment of the impact of regrowth on carbon and nutrient cycles has been hampered by a paucity of available allometric equations. We develop pooled and species-specific equations for total aboveground biomass for a study site in the eastern Amazon that had been abandoned for 15 years. Field work was conducted using randomized branch sampling, a rapid technique that has seen little use in tropical forests. High consistency of sample paths in randomized branch sampling, as measured by the standard error of individual paths (14%), suggests the method may provide substantial efficiencies when compared to traditional procedures. The best fitting equations in this study used the traditional form $Y=a \times DBH^b$, where Y is biomass, DBH is diameter at breast height, and a and b are both species-specific parameters. Species-specific equations of the form $Y=a(BA \times H)$, where Y is biomass, BA is tree basal area, H is tree height, and a is a species-specific parameter, fit almost as well. Comparison with previously published equations indicated errors from -33% to +29% would have occurred using off-site relationships. We also present equations for stemwood, twigs, and foliage as biomass components.

KEY WORDS: Allometry, Allometric equations, Secondary growth, Tropical forest

Equações alométricas para estimativa de biomassa de floresta secundária na Amazônia Oriental usando amostragem aleatória de ramos

RESUMO

Florestas secundárias ocupam uma área extensa e crescente na bacia Amazônica, porém determinações acuradas do impacto dessas florestas nos ciclos de carbono e nutrientes têm sido dificultadas pelo número reduzido de equações alométricas. Neste estudo, nós desenvolvemos equações em nível de comunidade e espécies individuais para estimar a biomassa total da parte aérea de uma floresta secundária com 15 anos de idade na Amazônia oriental. O trabalho de campo utilizou amostragem aleatória de ramos, que é uma técnica rápida, porém pouco utilizada em florestas tropicais. Baseada no erro padrão da série de segmentos individuais (14%), a consistência da série de segmentos totais amostrados foi considerada elevada, sugerindo que o método pode ser eficiente em comparação com procedimentos tradicionais. Os melhores ajustes foram obtidos com a equação tradicional $Y=a \times DBH^b$, onde Y é a biomassa, DBH é o diâmetro à altura do peito, e a e b são parâmetros para cada espécie arbórea. Ajustes razoáveis também foram alcançados com equações da forma $Y=a(BA \times H)$, onde Y é a biomassa, BA é a área basal, H é a altura e a é um parâmetro específico para cada espécie arbórea. Comparações com equações disponíveis na literatura indicaram uma faixa de erro provável de -33% a +29% usando-se relações desenvolvidas para outros sítios. Nós também apresentamos equações para os seguintes componentes da biomassa da parte aérea: tronco, ramos e folhas.

PALAVRAS-CHAVE: Alometria, Equações alométricas, Crescimento secundário, Floresta tropical

¹ Professor da Universidade de New Hampshire, e-mail: mjdoucey@cisunix.unh.edu

² Professor da Universidade da Flórida, e-mail: zarin@ufl.edu

³ Pesquisador da Embrapa Amazônia Oriental, e-mail: steel@cpatu.embrapa.br

⁴ Professora da Universidade Federal de Santa Maria, e-mail: maristela.araujo@ufsm.br

INTRODUCTION

Approximately 650,000 km² of old-growth forest in the Brazilian Amazon has been cleared within the past 25 years (Brazil INPE, 2004), and between 30% and 50% of this area is in some stage of abandonment and regrowth (Fearnside, 1996; Houghton *et al.*, 2000). Forest regrowth occupying much of this land area are of significant ecological interest in their own right (Uhl *et al.*, 1988; Buschbacher *et al.*, 1988; Mesquita *et al.*, 2001), and contribute to large-scale biogeochemical and atmospheric cycles (Fearnside & Guimarães, 1996; Houghton *et al.*, 2000; Steininger, 2004).

Tropical forest regrowth contains complex mixtures of tree species (Uhl *et al.*, 1988; Mesquita *et al.*, 2001), and there is considerable variation in species composition of regrowth stands across this large region (Baar *et al.*, 2004). However, few species-specific allometric equations have been published (Uhl *et al.*, 1988; Saldarriaga *et al.*, 1988; Nelson *et al.*, 1999). Most studies on biomass or carbon increment in Amazonian forest regrowth rely on mixed-species equations (such as those developed by Uhl *et al.*, 1988; Saldarriaga *et al.*, 1988, and Honzák *et al.*, 1996 using data from Brown *et al.*, 1989). However, Nelson *et al.* (1999) report considerable discrepancies between the predictions of these equations, suggesting considerable error may remain when using existing mixed-species regressions in new sites.

Tropical forests present special challenges for developing accurate biomass equations. Because even a small area may contain a relatively large number of species, many trees must be measured to develop an adequate set of equations. It would be especially advantageous in this situation to be able to measure each tree as efficiently as possible, so that many trees can be measured. One approach is to reduce the portion of the tree that must be processed by using a subsampling method. Randomized branch sampling (RBS, Gregoire *et al.*, 1995) has been proposed and theoretically investigated, but there are few published reports of its use.

Here, we report the results of an aboveground biomass study in forest regrowth in the eastern Amazon. Our goals were (1) to develop species-specific allometric equations for several species not represented in previously published studies, using randomized branch sampling, a well-described methodology for tree allometry not previously applied in tropical forest regions; (2) to develop mixed-species equations appropriate for forest regrowth of similar species composition; and (3) to compare our equations with previously published equations to assess which, if any, equations might prove portable across forest regrowth regionally.

MATERIAL AND METHODS

STUDY SITE

The study was carried out at the field station of the Federal Rural University of Amazonia (UFRA: 1°17'46" S and 45°55'28" W) near the city of Castanhal, in the Bragantina Region of Pará, Brazil. The Bragantina Region is distinguished from other Amazon regions by the predominance of secondary vegetation, resulting from constant human occupation since the rubber boom at the beginning of the 20th century (Ludovino, 2001). Previously covered by humid tropical forest with a dense canopy, the study area was first cleared around 1940. Corn and manioc were the main crops of the slash-and-burn agriculture in the area. After six to eight cycles of slash and burn, the area was abandoned and allowed to regrow for the past 15 yrs. Mean annual temperature is 26°C, average relative humidity is 80%, and mean annual precipitation is 2500 mm. A dry season extends from July to December with 2 to 5 months with total rainfall less than 100 mm according to data registered from the National Agency of Electrical Energy (ANEEL) network meteorological station at Castanhal (01°17'53"S, 47°56'56"W), which was about 3 km away from our study site. The predominant soil type is dystrophic yellow latosol, stony phase I, in the Brazilian classification (Tenório *et al.*, 1999), corresponding to Sombriustox in U.S. soil taxonomy. Soils are well-drained and shallow to laterite. Predominant botanical families are Lacistemataceae, Clusiaceae, and Myrtaceae.

FIELD AND LABORATORY METHODS

All trees ≥ 1.0 cm diameter at breast height (DBH) on two 0.01 ha study plots were destructively sampled. In addition, for two taxa (*Vismia guianensis* and *Stryphnodendron* sp.), 2-3 additional trees selected from the area around the plots were also sampled to extend the size range for the allometric regressions.

Before felling, DBH and total height (to the tallest point on the tree) were measured on each tree. DBH was measured using a diameter tape on trees with DBH >5cm, and using calipers for smaller trees. However, the differences between diameter measurements using the two techniques are practically and statistically negligible; DBH as measured using a diameter tape equals the average of all possible caliper measurements (Matérn, 1956). Height was measured using a telescoping range pole. Total height was also measured immediately after felling using a fiberglass tape. Then, we employed randomized branch sampling (Jessen, 1955; Gregoire *et al.*, 1995) to subsample each tree and develop design-unbiased estimates of total biomass and its components.

Briefly, randomized branch sampling decomposes a tree into a finite collection of paths composed of individual segments. Each path connects the base of the tree to a terminating branch. A segment is the portion of a path that runs between two ramifications. Sampling begins at the base of the tree; the segment running from the base to the first ramification is sampled with probability equal to 1. All material associated with that segment is processed to determine volume and biomass components (described below). At the ramification, one must choose which segment to take. We measured the basal diameter of each segment at the ramification, and used a field computer to select a segment with probability proportional to basal diameter raised to the 2.67 power; this selection probability has been suggested as nearly optimal for total biomass (Gregoire *et al.*, 1995). All material associated with this new segment is processed, and the selection process repeats at the ramification terminating the segment. This procedure is repeated until one reaches a terminal branch. The probabilities of selection at each ramification (the conditional probabilities; Gregoire *et al.*, 1995) are recorded, and collectively they define a probability sample through the set of possible paths on the tree. Each segment is expanded by its unconditional probability, i.e. the product of the conditional probabilities of each selection at preceding ramifications, to form a design-unbiased estimator of the total of an attribute in the tree, based on the same attribute measured in each segment in the path (Gregoire *et al.*, 1995). Only one path is needed to develop an estimate for the tree. To estimate the standard error for an individual tree, multiple paths must be selected; this was done on 6 trees to evaluate the accuracy of the sampling scheme.

For each segment, we divided the plant material into five categories for processing:

1. Stemwood. We defined stemwood as all live woody material (including bark) with an outside bark diameter ≥ 1.0 cm. If a lateral branch arose from the segment with diameter ≥ 1.0 cm, that constituted a ramification and the segment terminated at the branch base. Lateral branches with diameter < 1.0 cm were defined as twigs (see below). If the diameter fell below 1.0 cm before reaching a ramification (i.e. the stem became a twig), or if no branch ≥ 1.0 cm arose from the segment, the segment was considered a terminal branch. (For very small trees, with no branches ≥ 1.0 cm, this means the initial segment was also a terminal segment, so the entire tree was processed.) For each stemwood segment, we measured the length, basal diameter, and diameter at the terminating ramification. After separating other biomass components (see below), the entire segment was cut and weighed wet in the field. For very small segments, the entire segment was returned to the lab for dry weight determination; otherwise, a "cookie" subsample was removed from the midpoint of the segment, weighed in the field for wet weight, and returned to the lab for dry weight measurement.
2. Twigs. All live woody material, including bark, < 1.0 cm in diameter was considered as twigs. All twigs were collected from each selected segment and wet weight was determined in the field. For most segments, the entire collection of twigs was returned to the lab for dry weight determination; for a few segments with very large twig volumes, an arbitrary subsample of twigs was taken, its wet weight recorded, and the subsample was returned to the lab for dry weight measurement.
3. Dead branches and twigs. Dead woody material was associated with the segment from which it arose, without respect to diameter. Dead woody material was separated from the live wood by visual inspection and by testing for flexibility, its wet weight was determined, and either the entire dead branch or a subsample taken from its midpoint (also weighed wet in the field) was returned to the lab for dry weight measurement.
4. Foliage. Foliage was separated from the branches and twigs of a segment. All foliage associated with a segment was weighed wet in the field. If the volume of foliage was small (approximately a liter or less), all the foliage was returned to the lab. Otherwise, the foliage was mixed well, and a grab sample was taken for wet weight measurement in the field and dry weight measurement in the lab.
5. Seeds and fruits. These components were rare. They were treated separately from, but similar to, foliage.

All collected material was bagged and tagged to identify the tree, path number, and segment number from which it arose. Samples were returned to the lab within 12 hours, placed in an oven at 65°C, and dried until the residual weight was stable (at least 24 hours, but generally longer for large woody samples). Nogueira *et al.* (2005) report a statistically significant but small (less than 2%) difference in wood density estimates when higher oven temperatures (e.g. 103°C) are used for large woody samples; those differences should be borne in mind when comparing these results to those from similar studies. Dry weight was determined using a calibrated electronic balance to a precision of 0.1 g.

The total dry weight for a segment was determined as the sum of the dry weights of its components. Where a component was subsampled for dry weight determination, the total dry weight of that component was estimated as the field-measured wet weight for the total material, multiplied by the dry weight to wet weight ratio of the subsample. Dry weights for the segments were then multiplied by the inverse of the unconditional selection probabilities for each segment, and summed to estimate the total dry mass of the tree (Gregoire

et al., 1995). Volume for each segment was calculated using Smalian's formula ($V=L(A+a)/2$, where V is volume, L is length, and A and a are the large and small end cross-sectional areas, respectively). The volumes were not used directly in biomass estimation, but were used to assist in error checking the dry weights. Volume of the basal segment, in association with the dry woody biomass of the basal segment, was also used to develop specific gravity for each tree as a covariate in allometric equation development.

STATISTICAL ANALYSIS

We evaluated a series of allometric models predicting total aboveground biomass as a function of commonly measured variables including DBH and total height, as well as wood density. The first set of models did not include separate coefficients for different species or taxa, but treated the entire population as a single group. The second set included separate coefficients for each species with five or more individuals in the sample, treating all other species as a pooled "other" group.

We considered several candidate models for predicting the total aboveground biomass of trees (Y) without estimating species-specific parameters. These included two models employing tree basal area (BA , m^2), tree height (H , m ; we used height as measured before felling, not length measured after felling, to predict biomass), and a measure of wood specific gravity: either specific gravity as estimated from the basal segment of that tree (SG), or the median basal specific gravity for the species (MSG). The models in this class were

$$\text{Model 1.} \quad Y = \beta_0 (SG \times BA \times H)^{\beta_1}$$

$$\text{Model 2.} \quad Y = \beta_0 (MSG \times BA \times H)^{\beta_1}$$

In many applications, no specific gravity information is available. A model with similar form, relying only on tree dimensions, is

$$\text{Model 3.} \quad Y = \beta_0 (BA \times H)^{\beta_1}$$

However, use of this model requires measurement of tree height. Perhaps the most common allometric equation form is

$$\text{Model 4.} \quad Y = \beta_0 DBH^{\beta_1}$$

where DBH is diameter at breast height in cm. Although it is dimensionally consistent with the notion that much of the aboveground biomass of a tree comes from the stem volume, some object to forms such as Model 3 because the exponents on BA (or DBH) and H cannot vary independently. We also tested

$$\text{Model 5.} \quad Y = \beta_0 \times DBH^{\beta_1} \times H^{\beta_2}$$

Conversely, if the bulk of aboveground biomass does come from stem volume, then the additional exponent β_2 in Models 1, 2, and 3 may be considered unnecessary, and might lead to poor extrapolation even within the close vicinity of the original

data range. We also tested the following simple models with only a single estimable parameter:

$$\text{Model 6.} \quad Y = \beta_0 (SG \times BA \times H)$$

$$\text{Model 7.} \quad Y = \beta_0 (MSG \times BA \times H)$$

$$\text{Model 8.} \quad Y = \beta_0 (BA \times H)$$

It is often tempting to deal with expected heteroscedasticity (and at the same time avoid nonlinear regression) by fitting Models 1-5 after log transformation. However, this introduces a bias into resulting estimates, and complicates cross-comparison with models in the original scale. We employed nonlinear regression, estimating both the model parameters, and the parameters of a variance function, by direct maximum likelihood. For all models, the variance of individual data about the expected value was estimated as a function of the predicted value, with two additional parameters:

$$\text{Variance:} \quad s^2 = \gamma_0 Y^{\gamma_1}$$

This equation for the variance allows broad flexibility in modeling the heteroscedasticity directly, while using only one parameter more than would be needed when incorrectly assuming homoscedasticity. Each model thus required estimating 3-5 parameters: β_0 (and possibly β_1 and β_2), as well as γ_0 and γ_1 . Model selection was based on the small-sample, second-order corrected AIC_c information criterion (Akaike, 1974; Sugiura, 1978; Hurvich & Tsai, 1989). AIC_c stands for "an information criterion," or "Akaike information criterion," while the c in the subscript indicates a small-sample correction. AIC_c is calculated from the log-likelihood as

$$AIC_c = -2 \ln L + 2K + \frac{2K(K+1)}{n-K-1}$$

where $\ln(L)$ is the log-likelihood of the model, K is the number of parameters (including parameters associated with the variance), and n is the number of observations. Low values of AIC_c are considered desirable; particular values of AIC_c are not meaningful in isolation, but differences in AIC_c between models (ΔAIC_c) are (Burnham and Anderson 2002). Model selection using AIC_c is preferable to model selection using p -values, because the goal is to select an appropriate predictive model, not to test a straw null hypothesis (Burnham & Anderson, 2002). Although Burnham and Anderson (2002) warn against it, we also report pseudo- R^2 (also called model efficiency) and root mean square error here; these are reported for informational purposes only, and should not be used as model selection criteria in the presence of heteroscedasticity. We also caution that pseudo- R^2 values should not be used to compare allometric equations developed in different studies from different data sets because, just as in linear regression, the statistic depends strongly on the range and variability of the specific data set (Kvålseth, 1985).

We also fit a series of individual species or taxon models; individual species models are often considered to be more accurate in ecological work. We considered this not as a given but as a testable proposition. We fit a series of models, comparable in form to those used for the grouped equations. Specifically, we fit Models 1, 3, 4, 5, 6, and 8 as outlined above, but allowing separate parameters for each taxon. The median specific gravity term in Models 2 and 7 is uniform for a taxon, so these models are redundant. We employed the same heteroscedasticity assumption and variance model as before, allowing the two parameters of the variance model also to vary by taxon. Thus, each model required estimating 3-5 parameters per taxon. We fit separate models for seven species, and one additional group comprising all other species combined. We used an overall AIC_c to compare among equation forms, both within the taxon-specific series and with the multispecies models. The overall AIC_c was calculated from the joint log-likelihood of the models in a group, using the combined sample size and combined number of parameters (3-5 parameters for each of 8 taxa). As before, we report pseudo- R^2 and root mean squared error values, though these would be misleading in model selection due in part to the heteroscedasticity and in part to the varying number of parameters between models.

To compare the selected models with this study with those from previously published studies (Uhl *et al.*, 1988; Nelson *et al.*, 1999), we estimated the aboveground biomass for each sampled tree using the equations developed here and the previously published equations. Differences were assessed graphically, and the models were also compared for median percent error and median absolute error for each tree. Recognizing that the primary goal of biomass campaigns is

usually to estimate total biomass for a population of trees, rather than for individual trees, we also assessed overall error for the sum of the individual biomasses of the trees on the sample plots. We emphasize that we did this not to determine which set of equations is “right” in some global sense, but to test which (if any) model forms might prove portable across sites. Our working hypothesis was that models including measures other than DBH (such as SG or H) would better account for regional variability in tree form and stand composition, and would prove more portable.

The emphasis of this manuscript is on allometric equations for predicting total biomass. However, the data collected in this study are also suitable for developing equations for biomass components. We present equations for stemwood, twigs, and foliage in Appendix I.

RESULTS

Characteristics of the sampled trees are summarized in Table 1. In total, 82 trees were sampled; 77 trees represented 21 identifiable taxa while 5 trees could not be identified.

Variability of randomized branch sampling, for the six trees where two paths each were sampled, is shown in Table 2. The maximum difference between two estimates was 39%, but this was not typical. The pooled relative standard error of estimate calculated for these trees is 13.7%. Note furthermore that this would be an overestimate of the sampling variability as applied to the total population; for many trees, there were no branches ≥ 1.0 cm emanating from the first segment, so the entire tree was sampled and there is no sampling error. An increase in sampling error for larger trees does, however, add to the natural heteroscedasticity expected for tree biomass.

Table 1 - Characteristics of the sampled trees. Taxa with fewer than five trees were not analyzed individually, but placed in the “all other taxa” group. Median specific gravity of the “all other taxa” group was 0.591.

Species	n	SG	DBH (cm)		H (m)		Biomass (kg)	
			min	max	min	max	min	max
<i>Abarema jupunba</i>	5	0.606	5.0	8.8	6.0	8.2	4.754	27.568
<i>Banara guianensis</i>	1	0.591	1.9	1.9	4.7	4.7	0.816	0.816
<i>Casearia javitensis</i>	4	0.581	1.4	7.4	3.8	8.1	0.492	21.633
<i>Cupania scrobiculata</i>	1	0.674	1.0	1.0	3.0	3.0	0.181	0.181
<i>Gutteria poeppigiana</i>	1	0.304	2.8	2.8	7.3	7.3	1.123	1.123
<i>Inga flagelliformis</i>	1	0.577	4.1	4.1	5.7	5.7	4.303	4.303
<i>Lacistema aggregatum</i>	3	0.640	1.3	3.9	3.2	5.8	0.401	3.421
<i>Lacistema pubescens</i>	15	0.564	1.1	6.4	2.7	7.0	0.285	22.073
<i>Maquira guianensis</i>	2	0.533	1.2	3.1	3.1	5.8	0.310	1.824
<i>Myrcia bracteata</i>	1	0.851	1.2	1.2	3.2	3.2	0.415	0.415
<i>Myrcia sylvatica</i>	8	0.757	1.2	4.8	2.2	9.6	0.203	13.639
<i>Nectandra cuspidate</i>	3	0.475	5.5	10.4	8.0	10.5	10.081	34.306
<i>Neea glomerulifolia</i>	3	0.855	1.2	3.1	2.4	4.3	0.320	2.173
<i>Neea tamimbuca</i>	1	0.442	1.1	1.1	2.3	2.3	0.318	0.318

Table 1 - Continuation

Species	n	SG	DBH (cm)		H (m)		Biomass (kg)	
			min	max	min	max	min	max
<i>Ocotea guianensis</i>	8	0.516	5.0	18.4	7.7	18.0	5.692	197.659
<i>Poecilanthe effuse</i>	10	0.835	1.1	6.9	3.2	7.3	0.339	26.570
<i>Stryphnodendron pulcherrimum</i>	1	0.611	21.4	21.4	11.2	11.2	262.346	262.346
<i>Tapirira guianensis</i>	1	0.362	5.1	5.1	8.2	8.2	4.533	4.533
<i>Talisia longifolia</i>	1	0.933	3.0	3.0	4.0	4.0	2.869	2.869
<i>Vismia guianensis</i>	6	0.585	1.5	5.0	4.6	6.4	0.671	7.832
<i>Vochysia inundata</i>	1	0.439	5.7	5.7	7.0	7.0	11.581	11.581
Unidentifiable taxa	5	0.591	1.0	3.6	2.4	7.3	0.249	3.550

Table 2 - Comparison of biomass estimates for trees on which two randomized branch sampling paths were measured. Estimated biomass for a tree is the average of the estimates calculated for the two paths. Percent difference is calculated from the two estimates as $100(\max-\min)/\text{average}$.

Species	DBH, cm	H, m	Estimated Biomass, kg	Percent Difference
<i>Myrcia sylvatica</i>	4.0	9.6	12.7	14.2
<i>Nectandra cuspidate</i>	10.4	10.5	34.3	39.0
<i>Ocotea guianensis</i>	5.0	7.7	5.7	19.8
<i>Ocotea guianensis</i>	6.6	8.7	13.8	1.5
<i>Stryphnodendron</i> sp.	21.4	11.2	262.3	12.4
<i>Vismia guianensis</i>	5.0	5.5	7.8	0.4

A comparison of the mixed-species models, as fit to all taxa combined, is shown in Table 3. The best model was Model 2; its AIC_c of 314.3 constitutes the benchmark against which other models in this group should be judged. Somewhat surprisingly, Model 2 uses the median specific gravity for a taxon, rather than the specific gravity as measured from the basal segment of the subject tree. The individual-tree specific gravity is, however, subject to measurement error which can propagate to estimates of biomass. The second-best model, Model 7, also employs median specific gravity but is a ratio model, shedding the potentially destabilizing exponent. Its AIC_c of 316.1 differs from the best AIC_c by $\Delta AIC_c=1.8$. This value falls within the range of 0-2 interpretable as providing

“substantial” support for the model (Burnham & Anderson, 2002, p. 70).

Specific gravity information is not always available in field studies. The three best-performing mixed species models that do not include specific gravity are models 3, 5, and 6, in that order. The AIC_c of Model 3, 322.6, yields a $\Delta AIC_c=8.3$, somewhere between “considerably less” support and “essentially none” in the terminology of Burnham & Anderson (2002). Nonetheless, this is a simple model that might be useful when height measurements are feasible but little other information is available. The most common form of allometric model, Model 4, has a dismal $\Delta AIC_c=24.3$ and is essentially unsupported as a model form for use with mixed species in this context.

An overview of the species- or taxon-specific models, including their pooled AIC_c values, is given in Table 4. The construction of these pooled AIC_c models allows direct comparison with the mixed-species models reported in Table 3. The best taxon-specific model form is the traditional power function of DBH, Model 4. Parameters for the eight taxonomic groups, along with model performance and the shape of the variance function, are shown in Figure 1. This model substantially outperforms the best mixed-species models, including those that require density information. This equation has the advantage and disadvantage of using the

Table 3 - Mixed-species allometric equations. Y is estimated biomass (kg), SG is individual-tree specific gravity, MSG is median specific gravity for the species or group, BA is basal area (m^2), H is tree height (m), and DBH is diameter at breast height (cm). Variance reflects the spread of individual trees about the predicted curve. r^2 and RMSE are shown for informational purposes only; neither is an apt goodness-of-fit criterion, because both assume homoscedastic data. Low values of AIC_c indicate better model performance.

Model	Equation	Variance	r^2	RMSE	AIC_c
Model 1	$Y=736.5(SG \times BA \times H)^{0.9331}$	$s^2=0.5698Y^{1.316}$	0.95	0.87	319.3
Model 2	$Y=770.8(MSG \times BA \times H)^{0.9460}$	$s^2=0.5110Y^{1.358}$	0.96	0.82	314.3
Model 3	$Y=387.8(BA \times H)^{0.8916}$	$s^2=0.4032Y^{1.583}$	0.92	1.11	322.6
Model 4	$Y=0.2237 \times DBH^{2.260}$	$s^2=0.4154Y^{1.549}$	0.97	0.68	338.6
Model 5	$Y=0.0985 \times DBH^{1.879} \times H^{0.7355}$	$s^2=0.4154Y^{1.552}$	0.94	0.98	323.3
Model 6	$Y=929.3(SG \times BA \times H)$	$s^2=0.8520Y^{1.194}$	0.96	0.77	323.8
Model 7	$Y=938.8(MSG \times BA \times H)$	$s^2=0.7024Y^{1.241}$	0.97	0.72	316.1
Model 8	$Y=569.4(BA \times H)$	$s^2=0.7912Y^{1.356}$	0.92	1.12	333.1

least information (other than taxonomic identity and DBH) of any model. A concern with models of this form is that extrapolation outside the measured range of DBH can be poor. Moreover, shifts in the relationship between DBH, height, and biomass can occur between secondary forest sites.

The second best model, Model 8, is a very simple ratio model that requires tree height in addition to tree basal area (or DBH). Parameters for the eight taxonomic groups, along with model performance and the shape of the variance function, are shown in Figure 2. This model also outperforms the mixed-species models, including those with density information. For this model, relative to the taxon-specific Model 4, $\Delta AIC_c = 9.9$, or somewhere between “considerably less” support and “essentially none” in the terminology of Burnham & Anderson (2002). For sites similar to that of this study, the Model 4 formulation would clearly be preferred.

A graphic comparison between the best mixed-species equations in this study, and the comparable equations of Uhl *et al.* (1988) and Nelson *et al.* (1999), is shown in Figure 3. Model 6 of Nelson *et al.*, which requires specific gravity, tended to overpredict in relation to Model 2 of this study. All the other equations, including the equation of Uhl *et al.* employing specific gravity, and those of Nelson *et al.* using DBH with height or DBH alone, tended to underpredict for this data set. The results are summarized numerically in Table 5. The equations generated in this study often missed

Table 4 - Summary of the species- or group-specific allometric equations. Y is estimated biomass (kg), SG is individual-tree specific gravity, MSG is median specific gravity for the species or group, BA is basal area (m²), H is tree height (m), and DBH is diameter at breast height (cm). Parameters a, b, and c are fit individually for each species or group. Low values of AIC_c indicate better model performance.

	Equation	AIC_c
Model 1	$Y = a(SG \times BA \times H)^b$	311.2
Model 3	$Y = a(BA \times H)^b$	313.0
Model 4	$Y = a \times DBH^b$	279.1
Model 5	$Y = a \times DBH^b \times H^c$	385.4
Model 6	$Y = a(SG \times BA \times H)$	290.6
Model 8	$Y = a(BA \times H)$	289.0

the aboveground biomass of individual trees by 19-25%. However, these errors were more or less evenly distributed, so that estimates of the total biomass of the sample (i.e., the plot-level biomass estimates) were within 5%. Errors for individual trees using the off-site allometric equations were on the order of 30-40%, but much of this error was comprised of bias; errors for the total aboveground biomass were as great as 33%. Contrary to expectations, the best performer among the off-site allometric equations was Model 2 of Nelson *et al.* (1999), which uses DBH alone. Its error of 17% for plot-level aboveground biomass estimates might still be judged unacceptable in some situations.

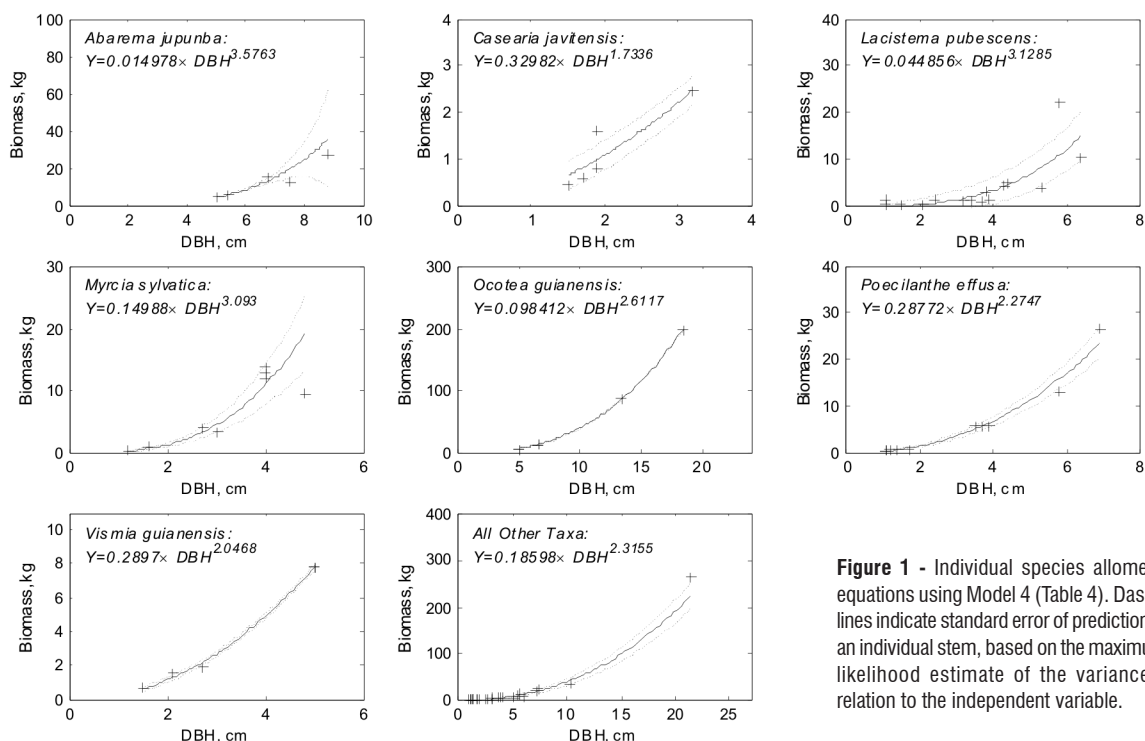


Figure 1 - Individual species allometric equations using Model 4 (Table 4). Dashed lines indicate standard error of prediction for an individual stem, based on the maximum-likelihood estimate of the variance in relation to the independent variable.

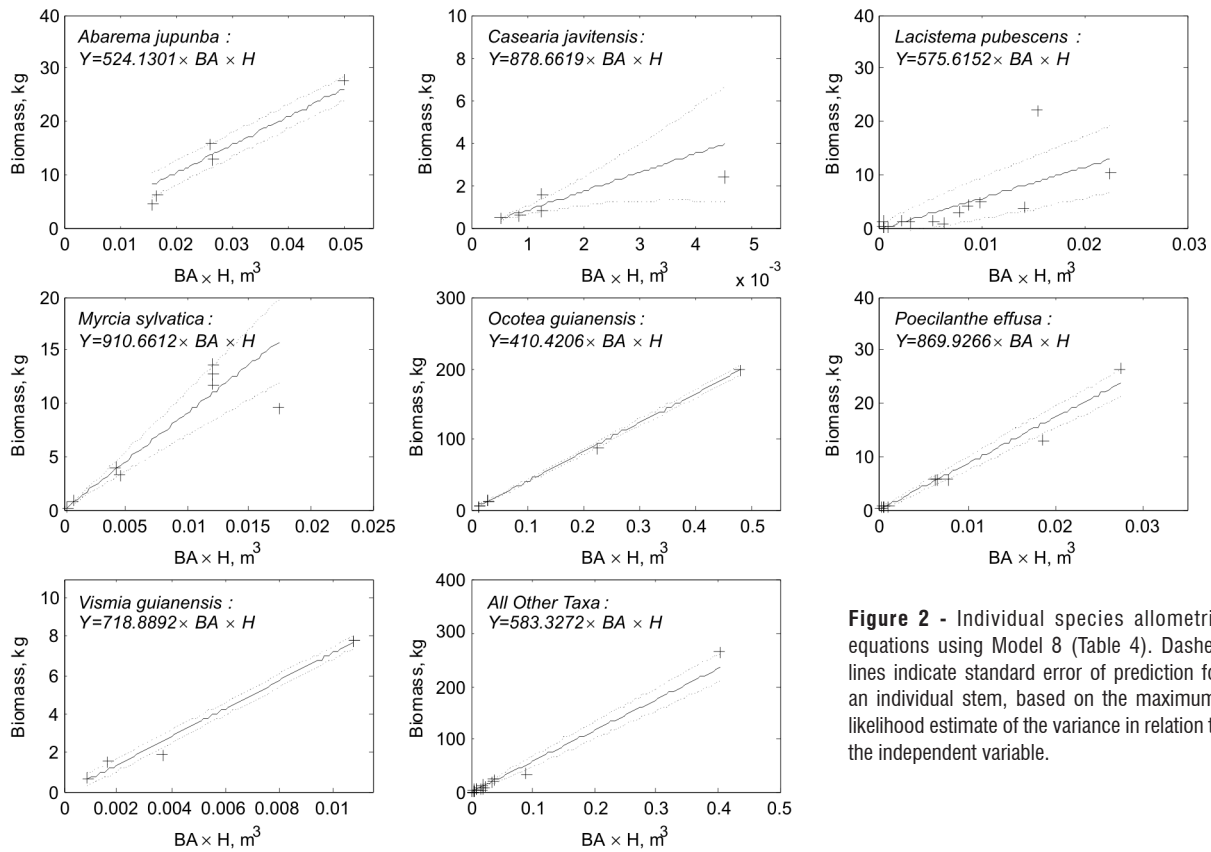


Figure 2 - Individual species allometric equations using Model 8 (Table 4). Dashed lines indicate standard error of prediction for an individual stem, based on the maximum-likelihood estimate of the variance in relation to the independent variable.

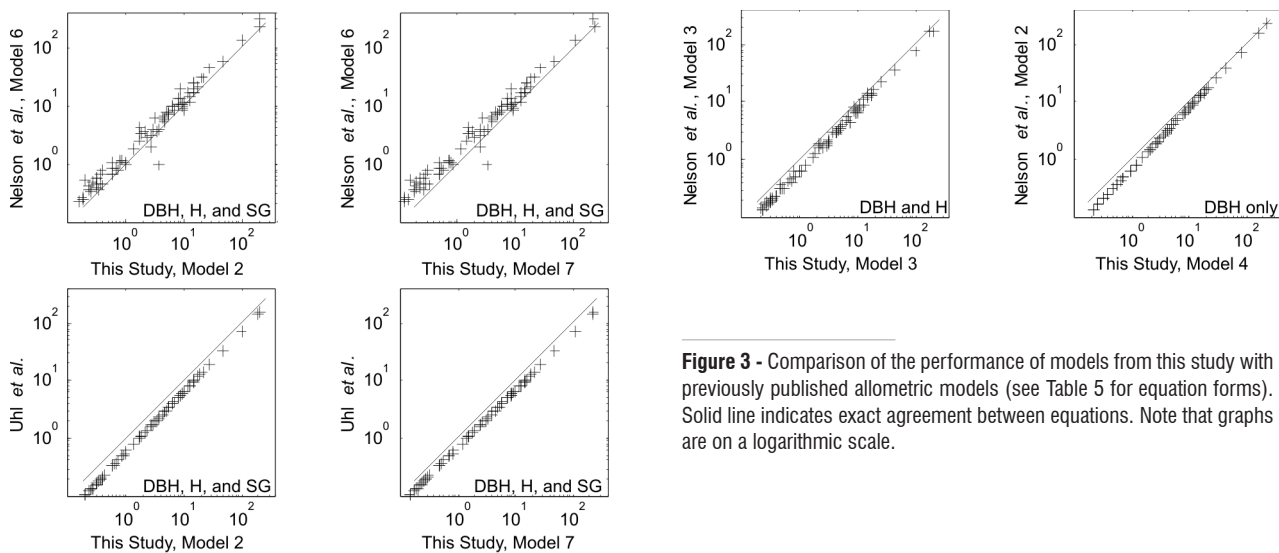


Figure 3 - Comparison of the performance of models from this study with previously published allometric models (see Table 5 for equation forms). Solid line indicates exact agreement between equations. Note that graphs are on a logarithmic scale.

Table 5 - Comparison with previously published allometric equations. Models of Nelson *et al.* (1999) and Uhl *et al.* (1982) are shown with biomass, rather than ln(biomass), as the dependent variable to facilitate comparison. Median error and median absolute error are calculated for individual trees; relative error of total describes the difference from the plot-level estimate calculated using the directly measured tree biomasses.

Source	Equation	Median Error (%)	Median Absolute Error (%)	Relative Error of Total (%)
Models Requiring DBH (or BA), H, and MSG or SG				
This Study, Model 2	$Y = 770.8(\text{MSG} \times \text{BA} \times \text{H})^{0.9460}$	2.0	18.9	-1.8
This Study, Model 7	$Y = 938.8(\text{MSG} \times \text{BA} \times \text{H})$	-11.2	23.9	2.0
Nelson <i>et al.</i> , Model 6	$Y = 0.2299 \times \text{DBH}^{2.1569} \times \text{H}^{0.3888} \times \text{SG}^{0.7218}$	36.9	39.7	28.6
Uhl <i>et al.</i>	$Y = 0.0514 \times (\text{DBH}^2 \times \text{H} \times \text{SG})^{0.991}$	-40.4	41.2	-33.0
Models Requiring DBH and H				
This Study, Model 3	$Y = 387.8(\text{BA} \times \text{H})^{0.8916}$	9.4	23.9	-4.8
Nelson <i>et al.</i> , Model 3	$Y = 0.0804 \times \text{DBH}^{2.1400} \times \text{H}^{0.4644}$	-26.1	30.9	-23.9
Models Requiring DBH Only				
This Study, Model 4	$Y = 0.2237 \times \text{DBH}^{2.260}$	8.7	25.7	-4.4
Nelson <i>et al.</i> , Model 2	$Y = 0.1358 \times \text{DBH}^{2.4128}$	-21.0	32.1	-17.0

DISCUSSION

The choice of an allometric model depends on the biological context and the sampling context, as well as the statistical properties of the model. Unless one is prepared to conduct a site-specific allometric study, the stability of predictions across sites is a key question. This is particularly important for forest regrowth; the height-diameter relationship, which is undoubtedly important for total aboveground biomass as well as biomass allocation, has long been known to vary with stand density and stand development (Oliver & Larson, 1996).

Viewed from this perspective, the second-best set of models in this study (taxon-specific models using Model 8, a simple ratio model employing tree basal area and height) seems to have much to offer. The models incorporate species composition and differences in form using a single parameter, while the lack of an exponent means the estimates will scale linearly and safely for trees slightly outside the range of the data. Gross extrapolation still remains dangerous, of course. In particular, very small trees (those shorter than breast height) may have zero basal area, but still have finite biomass. Also, changes in tree form with increasing size are not accounted for if the single parameter remains constant.

A major objection to models such as Model 8 has been the use of tree height. For example, Nelson *et al.* (1999, p. 166) note that telescoping height poles often provide underestimates of tree height in tropical forests. They argue that tree height should only be used in allometric equations when it provides clear improvement in predictions, and when it can be measured accurately. In this study, tree height measured before felling with a height pole averaged 5% less than length measured after felling with a tape (standard error of 1%). Measurement variability added an additional 9.7% random error (in the root mean square sense) to the bias. Part of this “bias” is undoubtedly because height, rather than length, is measured in the field, and some trees lean;

part is undoubtedly due to parallax and imperfect lines of sight. However, both sources of error (bias and variance) are accounted for in the model fitting process. If a small bias is relatively consistent between field crews, and a similar bias is present in the original fitting data, the impact is negligible. If height measurements were so variable as to be inherently unreliable, then adding height would not improve predictions even in the fitting data. Still, the ability to measure heights accurately and consistently cannot be taken for granted, and it must be weighed against the possibility of error introduced by shifts in the height-diameter relationship.

With that said, the best set of models in this study did not employ tree height, but relied on DBH alone in concert with species or taxon identification. In this, our results are similar to those of Nelson *et al.* (1999), who reported good fits with traditional power functions of DBH. This study adds to the limited menu of taxon-specific allometric equations for Amazonian forest regrowth.

The final result of this study, the general lack of agreement between locally-developed mixed species equations and other published equations, can only be viewed as discouraging. The previously published equations of Uhl *et al.* (1988) and Nelson *et al.* (1999) are not wrong; they are simply not right for this site. Likewise, the mixed-species equations developed here may be found wanting when applied to other forest regrowth stands. This remains true, apparently, even when specific gravity and tree height are accounted for in the models. The instability of allometric predictions across these sites indicates that additional work is needed to better understand the spatial variability in tree size – biomass relations. The demonstrated efficiency of using randomized branch sampling for biomass allometry may permit more frequent development of site-specific equations than has previously occurred using traditional procedures.

ACKNOWLEDGMENTS

We thank the students who participated in this study during biomass allometry field course held at the UFRA field station in 2002, Professor Izildinha Miranda for organizational support prior to and during the course, and Wilson J. Dias de Oliveira for assistance with sample processing. We also thank Moisés Mourão Jr. for text revision. This research was conducted under cooperative agreements between the University of Florida, UFRA, and the Empresa Brasileira de Pesquisa Agropecuária, and was supported by a grant from the Andrew Mellon Foundation.

LITERATURE CITED

- Akaike, H. 1974. A new look at the statistical model identification. *IEEE Trans. Automatic Control*, AC-19: 716-723.
- Baar, R.; Cordeiro, M.R.; Denich, M.; Fölster, H. 2004. Floristic inventory of secondary vegetation in agricultural systems of East-Amazonia. *Biodivers. Conserv.*, 13(7): 501-528.
- Brazil, INPE. 2004. Monitoramento da floresta Amazônica brasileira por satélite – Projeto Prodes. (www.obt.inpe.br/prodes). Access: 23/June/05.
- Brown, S.; Gillespie, A.J.R.; Lugo, A.E. 1989. Biomass estimation methods for tropical forests with applications to forestry inventory data. *For. Sci.*, 35(4): 881-902.
- Burnham, K.P.; Anderson, D.R. 2002. *Model Selection and Multimodel Inference: a Practical Information-Theoretic Approach*, 2nd ed. Springer, New York, USA. 488pp.
- Buschbacher, R.; Uhl, C.; Serrão, E.A.S. 1988. Abandoned pastures in eastern Amazonia. II. Nutrient stocks in the soil and vegetation. *J. Ecol.*, 76(3): 682-699.
- Diniz, T.D. de A.S. 1986. Caracterização climática da Amazônia Oriental, 3-13. EMBRAPA/CPATU-GTZ, Belém, Pará, Brazil.
- Fearnside, P.M. 1996. Amazonian deforestation and global warming: carbon stocks in vegetation replacing Brazil's Amazon forests. *For. Ecol. Manage.*, 80(1): 21-34.
- Fearnside, P.M.; Guimarães, W.M. 1996. Carbon uptake by secondary forests in Brazilian Amazonia. *For. Ecol. Manage.*, 80(1): 35-46.
- Gregoire, T.G.; Valentine, H.T.; Furnival, G.M. 1995. Sampling methods to estimate foliage and other characteristics of individual trees. *Ecology*, 76(4): 1181-1194.
- Honzák, M.; Lucas, R.M.; Amaral, I.; Curran, P.J.; Foody, G.M.; Amaral, S. 1996. Estimation of the leaf area index and total biomass of tropical regenerating forests: comparison of methodologies. pp. 365-381 In: Gash, J.H.; Nobre, C.A.; Roberts, J.M.; Victoria, R.L. (Eds.) *Amazonian Deforestation and Climate*. Institute of Hydrology, UK.
- Houghton, R.A.; Skole, D.L.; Nobre, C.A.; Hackler, J.L.; Lawrence, K.T.; Chomentowski, W.H. 2000. Annual fluxes of carbon from deforestation and regrowth in the Brazilian Amazon. *Nature*, 403: 301-304.
- Hurvich, C.M.; Tsai, C.-L. 1989. Regression and time series model selection in small samples. *Biometrika*, 76(2): 297-307.
- Jessen, R.J. 1955. Determining the fruit count on a tree by randomized branch sampling. *Biometrics*, 11(1): 99-109.
- Kvålseth, T.O. 1985. Cautionary note about R^2 . *Am. Stat.*, 39(4): 279-285.
- Ludovino, R.M.R. 2001. Análise da diversidade e da dinâmica da agricultura familiar na Amazônia Oriental: O caso da Zona Bragantina. Ph.D. dissertation, Universidade Técnica de Lisboa, Instituto Superior de Agronomia, Lisboa, Portugal.
- Matérn, B. 1956. On the geometry of the cross-section of a stem. *Medd. Stat. Skogsforskn. Inst.*, 46(11). 28 pp.
- Mesquita, R.C.G.; Ickes, K.; Ganade, G.; Williamson, B.B. 2001. Alternative successional pathways in the Amazon Basin. *J. Ecol.*, 89(4): 528-537.
- Nelson, B.W.; Mesquita, R.; Pereira, J.L.G.; Souza, S.G.A. de; Batista, G.T.; Couto, L.B. 1999. Allometric regressions for improved estimate of secondary forest biomass in the central Amazon. *For. Ecol. Manage.*, 117(3): 149-167.
- Nogueira, E.M.; Nelson, B.W.; Fearnside, P.M. 2005. Wood density in dense forest in central Amazonia, Brazil. *For. Ecol. Manage.*, 208(1-3): 261-286.
- Oliver, C.D.; Larson, B.C. 1996. *Forest Stand Dynamics*, update ed. Wiley, New York, NY, USA. 544pp.
- Saldarriaga, J.G.; West, D.C.; Tharp, M.L.; Uhl, C. 1988. Long-term chronosequence of forest succession in the upper Rio Negro of Colombia and Venezuela. *J. Ecol.*, 76: 938-958.
- Steininger, M.K. 2004. Net carbon fluxes from forest clearance and regrowth in the Amazon. *Ecol. Appl.*, 14(4): S313-S322.
- Sugiura, N. 1978. Further analysis of the data by Akaike's information criterion and the finite corrections. *Comm. Stat. Theor. Methods*, A7: 13-26.
- Tenório, A.R.D.M.; Graça, J.J.D.C.; Góes, J.E.M.; Mendez, J.G.R.; Gama, J.R.M.F.; Silva, P.R.O.D.; Chagas, P.S.M.D.; Silva, R.N.P.D.; Américo, R.R.; Pereira, W.L.M. 1999. Mapeamento dos solos da estação de piscicultura de Castanhal, PA. *FCAP Informe Técnico* 25: 5-26.
- Uhl, C.; Buschbacher, R.; Serrão, E.A.S. 1988. Abandoned pastures in eastern Amazonia. I. Patterns of plant succession. *J. Ecol.*, 76(3): 663-681.

Recebido em 03/08/2008

Aceito em 03/03/2009

APPENDIX I. EQUATIONS FOR ESTIMATING BIOMASS OF STEMS, TWIGS, AND FOLIAGE.

It is often useful to estimate not only the total aboveground biomass of a tree, but also the biomass of major components. Here, we present equations for estimating the biomass of stemwood, twigs, and foliage based on the data for this study. All components are as defined in the Methods section. Equations were fit using the same procedure as was used for total biomass. For simplicity, we provide here the species-specific equations that use a power function of DBH (i.e. following Model 4), because DBH is commonly measured and may be the only variable (other than species) available in some investigations. We also present the best-fitting equation (as judged using AIC_c) or set of equations that uses median specific gravity or height, as these may be available in some situations and can often improve model performance.

Results for stemwood are shown in Table A.1. In addition, it is notable that the best fitting equation for predicting stemwood biomass, out of all the sets evaluated, is that based on Model 1:

$$Y=679.1(SG \times BA \times H)^{0.9767}$$

Almost as good ($\Delta AIC_c=0.36$) is that based on Model 6:

$$Y=757.1(SG \times BA \times H)$$

Both models had an r^2 of 0.97, and the fit (as judged using AIC_c) was dramatically better than the next best model ($\Delta AIC_c > 35$). However, both require density from a wood sample taken from the first stem segment. This is wildly impractical in most investigations, but it might be possible to use these equations advantageously in a subsampling approach.

Results for twigs and foliage are shown in Table A.2 and Table A.3, respectively. The best fitting model in both cases follows the same form, Model 2, which uses median specific gravity of the stemwood but otherwise is not species-specific. We note that particular caution should be taken with the foliar biomass equations, as foliar biomass may be influenced by seasonal phenology. Also, we caution that estimates calculated from the equations in Tables A1 through A3 may not necessarily add exactly to estimates of the total aboveground biomass calculated using the equations presented in the main paper.

Table A2. Allometric equations for predicting twig biomass. Y is estimated biomass (kg), MSG is median specific gravity for the species or group, BA is basal area (m²), H is tree height (m), and DBH is diameter at breast height (cm). r^2 and RMSE are shown for informational purposes only; neither is an apt goodness-of-fit criterion, because both assume homoscedastic data. The performance of the Model 2 equation is substantially better than that of the Model 4 equations ($\Delta AIC_c=29.34$).

Species	Model 2		
	Model	r^2	RMSE
All species	$Y=10.095(MSG \times BA \times H)^{0.579}$	0.73	0.06
Species	Model 4		
	Model	r^2	RMSE
<i>Abarema jupunba</i>	$Y=0.000147 \times DBH^{4.485}$	0.98	0.08
<i>Casearia javitensis</i>	$Y=0.0359 \times DBH^{1.951}$	0.81	0.03
<i>Lacistema pubescens</i>	$Y=0.0665 \times DBH^{1.036}$	0.51	0.04
<i>Myrcia sylvatica</i>	$Y=0.0591 \times DBH^{1.847}$	0.40	0.12
<i>Ocotea guianensis</i>	$Y=0.0506 \times DBH^{1.498}$	0.66	0.47
<i>Poecilanthe effusa</i>	$Y=0.0549 \times DBH^{1.749}$	0.75	0.13
<i>Vismia guianensis</i>	$Y=0.0433 \times DBH^{1.514}$	0.76	0.07
Other species	$Y=0.0749 \times DBH^{1.350}$	0.78	0.13

Table A1. Allometric equations for predicting stemwood biomass. Y is estimated biomass (kg), BA is basal area (m²), H is tree height (m), and DBH is diameter at breast height (cm). r^2 and RMSE are shown for informational purposes only; neither is an apt goodness-of-fit criterion, because both assume homoscedastic data. The performance of the species-specific Model 3 equations is substantially better than that of the Model 4 equations ($\Delta AIC_c=9.69$).

Species	Model 3			Model 4		
	Model	r^2	RMSE	Model	r^2	RMSE
<i>Abarema jupunba</i>	$Y=3260.2(BA \times H)^{1.5813}$	0.86	1.75	$Y=0.0238 \times DBH^{3.233}$	0.87	1.71
<i>Casearia javitensis</i>	$Y=61.8(BA \times H)^{0.6712}$	0.78	0.15	$Y=0.2151 \times DBH^{1.750}$	0.76	0.16
<i>Lacistema pubescens</i>	$Y=137.1(BA \times H)^{0.8227}$	0.83	0.28	$Y=0.1329 \times DBH^{2.027}$	0.80	0.30
<i>Myrcia sylvatica</i>	$Y=376.6(BA \times H)^{0.9169}$	0.77	0.96	$Y=0.1022 \times DBH^{3.341}$	0.38	1.58
<i>Ocotea guianensis</i>	$Y=543.0(BA \times H)^{1.1278}$	0.85	9.59	$Y=0.0549 \times DBH^{2.777}$	0.99	1.99
<i>Poecilanthe effusa</i>	$Y=800.5(BA \times H)^{1.0429}$	0.99	0.19	$Y=0.1557 \times DBH^{2.508}$	0.99	0.16
<i>Vismia guianensis</i>	$Y=1543.7(BA \times H)^{1.2222}$	0.99	0.12	$Y=0.1281 \times DBH^{2.402}$	0.99	0.12
Other species	$Y=386.0(BA \times H)^{0.9704}$	0.93	2.05	$Y=0.1073 \times DBH^{2.461}$	0.99	0.63

Table A3. Allometric equations for predicting foliar biomass. Y is estimated biomass (kg), MSG is median specific gravity for the species or group, BA is basal area (m²), H is tree height (m), and DBH is diameter at breast height (cm). r² and RMSE are shown for informational purposes only; neither is an apt goodness-of-fit criterion, because both assume homoscedastic data. The performance of the Model 2 equation is only slightly better than that of the Model 4 equations ($\Delta AIC_c = 1.79$). The Model 4 equation has an exponent of 0 for *Abarema jupunba* (i.e., DBH was not a useful predictor of foliar biomass for the sample data).

Model 2			
Species	Model	r ²	RMSE
All species	$Y = 15.151(\text{MSG} \times \text{BA} \times \text{H})^{0.658}$	0.51	0.19
Model 4			
Species	Model	R ²	RMSE
<i>Abarema jupunba</i>	$Y = 0.1196$	0.00	0.13
<i>Casearia javitensis</i>	$Y = 0.0369 \times \text{DBH}^{2.191}$	0.82	0.04
<i>Lacistema pubescens</i>	$Y = 0.0488 \times \text{DBH}^{0.877}$	0.21	0.04
<i>Myrcia sylvatica</i>	$Y = 0.0535 \times \text{DBH}^{0.947}$	0.18	0.05
<i>Ocotea guianensis</i>	$Y = 0.0060 \times \text{DBH}^{1.700}$	0.82	1.12
<i>Poecilanthe effusa</i>	$Y = 0.0844 \times \text{DBH}^{1.749}$	0.71	0.19
<i>Vismia guianensis</i>	$Y = 0.0468 \times \text{DBH}^{1.791}$	0.96	0.04
Other species	$Y = 0.0414 \times \text{DBH}^{1.785}$	0.81	0.20