Dear Editor,

Attached are our responses to the Reviewers (that have been uploaded as well on the discussion forum) and the revised manuscript (track-changed version). We highlighted changes related to comments from Reviewer#1 in light blue, Reviewer#2 in yellow and Dr. Rutishauser (comment on the discussion forum) in green. Slight phrasing corrections, unrelated to reviewers comments, have also been done (track-changed, no highlight).

Sincerely,

The Authors

RESPONSE TO REVIEWER#1

Dear Dr. T. Baker,

Thank you for the relevant comments and suggestions you made on our manuscript. We here address the different points you raised in your review.

Reviewer comments (in italics): "*1. Practical recommendations: I was surprised given the effort the develop models for canopy mass based on direct measurements of canopy size, and their improved performance compared to models based only on diameter and height (m3 compared to m2 in Table 2), that the final recommendation is only to implement model m2 (ie just measure tree height to the base of the canopy). To me, making a few additional measurements of canopy diameter for the few largest trees in a stand would not be particularly onerous, would improve accuracy, and would be important for linking field measurements to any LiDAR studies. Why is this option rejected?*"

Our response (in plain text): We indeed put emphasis at the end of the manuscript on model m_2 which only requires trunk height as additional measurement, as we believe it much simpler to implement in a standard forest inventory protocol that already includes tree height measurements than would adding a full set of crown metrics, even on a subset of trees. We however agree that information on crown diameter, or even crown architecture, for the largest trees in a stand is highly valuable, notably for remote-sensing studies. But except for scientific studies, we are not convinced that collecting such data will become a common practice in the coming years. Unlike trunk height, measuring crown diameter is increasingly difficult as tree ages and forest canopy becomes crowded. In tropical forest canopy, crown limits are often difficult to identify, all the more when one needs to have his sights set on two opposite crown limits from a single vantage spot, as required by most Laser measurement devices. It follows that even if one only targets the largest trees in a stand, this operation is time-consuming and probably inaccurate or even biased if too quickly performed. It is not unlikely that the development of mobile terrestrial Lidar scanning technology will make it possible to extract crown data more easily in a near future. The option to include more detailed crown measurements into biomass allometric equations is therefore not rejected, but from a practical point of view, the collection of only additional trunk height data appeared to us as the most reasonable option for immediate large-scale application in operational contexts. We added a sentence at the end of the manuscript (section 4.3) to clarify our choice to the reader.

"2. Height definition: I agree that the definition of the canopy base needs to be carefully considered in the manuscript as it is an important parameter in the models. For example, it might be useful to set a minimum diameter for the lowest living branch to define these measurements (e.g. 5 cm). The authors of the manuscript involved in data collection would doubtless have valuable experience to define this carefully for tropical trees."

We indeed proposed to change our definition of crown base from "the height of the first living branch" to "the height of the first main branch". As pointed out by Dr. Rutishauser in his comment to the paper, the former definition would have included short-lived branch such as small unreiterated and/or epicormic branches, while the latter typically refers to long-lasting branches (e.g., forks). On large trees for which we advocate measuring trunk height (≥ 100 cm DBH), identifying the lowest main branch (elsewhere called "crown-forming branch", Husch et al., 2002) is often straightforward and routinely performed by foresters. However, as subjectivity increases with decreasing tree size, we understand the appeal of setting a branch diameter threshold (e.g. 5 cm). Besides setting a threshold for branch diameter, it might also be necessary to set a threshold for the vertical position of the branch along the main axis. For instance, while one may consider accounting for a 5 cm branch located e.g. 1 m bellow a growing fork (i.e. future crown base), the same 5 cm branch may be left out if it is rather located 2, 3 or 4 m bellow this point. The form of these thresholds (i.e. in cm or in % of tree DBH) might also be discussed. We believe that our field experience cannot backup all those choices that should rather be addressed using a statistical approach. Again, terrestrial LiDAR scanning technology appears particularly promising in this regard.

"3. Collinearity. The potential problems of collinearity in biomass models have been a contentious issue in the literature, and could be raised in the context of this study as well. Personally, I agree with previous work by a linked group of authors (Picard et al., 2015), that these problems (defined by considering variance inflation factors) are secondary to evaluating model performance against data, particularly now that the datasets are increasingly representative of the full range in structure of tropical trees. However, I think it

would be useful to refer briefly to this debate and the literature on this point (e.g. in section 4.3), so these points are clear to readers."

We agree and have added a paragraph in section 4.3 to refer this issue.

"4. Scaling up: I like the comments in the discussion about how the effect of these findings will depend on the size structure of the forest (section 4.2). I think it would be useful to expand this slightly to reflect on how inclusion of canopy mass will improve our understanding of broad-scale differences in biomass among forests. For example, we know that African forests tend to have more large trees than Amazonian forests (Lewis et al., 2013), and that canopy size varies with seasonality in Amazonia (Barbier et al., 2010). What implications does this work have for detecting differences in biomass among continents and along environmental gradients?"

Thanks you for this suggestion, we expanded section 4.2 accordingly.

Equation 2: H should be Ht, I think.

Correct, it has been replaced.

RESPONSE TO REVIEWER#2

Dear Reviewer #2,

Thank you for your positive review and suggestions. We here address the different points you raised in your review.

Reviewer comments (in italics): *"In the manuscript the authors suggest that incorporation of the crown mass ratio into commonly used allometric equations could improve the accuracy of forest carbon estimates. So my suggestion is to incorporate this information into the analysis by comparing calculated estimates of aboveground biomass between the proposed models and to discuss different results regarding strategies of carbon allocation between stem and crown mass and its implications for tropical carbon storage. For instance, the authors could include a table stating respective forest carbon estimates for the investigated study sites and compare reported estimates to the results derived by their novel approach accounting for a crown mass proxy. This would allow for a more direct comparison between the biomass estimates derived from the respective models and could be used to discuss the importance of incorporating crown metrics in allometric models to account for potential alterations in carbon allocation in response to projected global changes."*

Our response (in plain text): Applying the different models to plot data in order to compare the resulting aboveground biomass estimates is obviously a good idea. However, it requires having measured crown metrics, which was not systematically done for the field plots we used here. For instance, we do not have this information for the 50-ha plot at Korup NP, where the influence of forest structure on the pantropical model error is most evident. Among the 80 1-ha plots of the IRD network, we possess information on trees crown depth in 46 plots. In each plot, crown depth measurements were made on a subset of trees (N=39.2 ± 15.8) distributed over all tree diameter classes (but ≥ 10 cm of diameter at breast height, D). Although we do not have crown metric information for all large trees in those plots, we used this dataset to dig into your suggestion (Fig. 1). In Figure 1 of this response, the X axis represents plot aboveground biomass derived from m_0, the pantropical model of Chave et al. (2014). On the Y axis, plot AGB was computed with the same biomass model for trees with no information on crown depth, or with our model m_2 for trees with information on total height and crown depth (Fig 1-A). We also restricted the subset of trees on which m_2 was applied to trees with D ≥ 100 cm, as recommended in our manuscript (Fig. 1-B). For simplicity, plot AGB estimated from combined m_0 and m_2 predictions (as described above) are referred to as m_2 predictions. In Fig. 1-A, subtracting m_0 to m_2 predictions leads to an average difference of about +2 Mg. In the manuscript, we indeed showed that the averaged bias at the level of the plot network is fairly close between m0 (+6.8%, p. 19724 L. 1) and m_2 (+5%, p. 19724 L. 10). The spread of plot-level biases is also consistent with previous findings (see Fig. 5-B of the manuscript), with a tendency for m_0 to result in higher AGB estimates than m_2 (up to +20 Mg or +6% of m_0 AGB estimate), with the exception of some high-biomass plots where large trees AGB is underestimated by m_0 (up to -56 Mg, or -15.9% of m0 AGB estimate). Restricting the use of m_2 to trees with D ≥ 100 cm leads to plots AGB estimates closer to those obtained with m_0 (Fig. 1-B), notably because the overestimation of small trees AGB is not accounted for anymore. High-biomass plots still depart from the 1-1 line. Despite the limited representativity of our data on crown metrics is (even for large trees), these preliminary results seem in good agreement with the trends presented in the manuscript. Yet, we do not think that those results would bring much to the manuscript and propose to keep them here, especially since this response will be associated to the article.

Let us also note that your suggestion to discuss "*the importance of incorporating crown metrics in allometric models to account for potential alterations in carbon allocation in response to projected global changes"* echoes the comment number 4 from Reviewer#1 ("scaling up"), which led to a slight expansion of this subject in paragraph 4.2 of our manuscript.

"Page 19714; Line 4: Consider stating: "...,which play a major role in the global carbon balance (REFs)."

Agreed.

"Page 19714; Line 8: Consider changing the sentence to: "However, local forest biomass estimations commonly represent the foundation for the calibration and validation of remote sensing models."

Agreed.

“Page 19724; Line 20: Consider stating: “This threshold was mirrored by a break point in the relationship between total tree mass and the compound predictor variable used in the reference allometric model of Chave et al. (2014).”

Agreed.

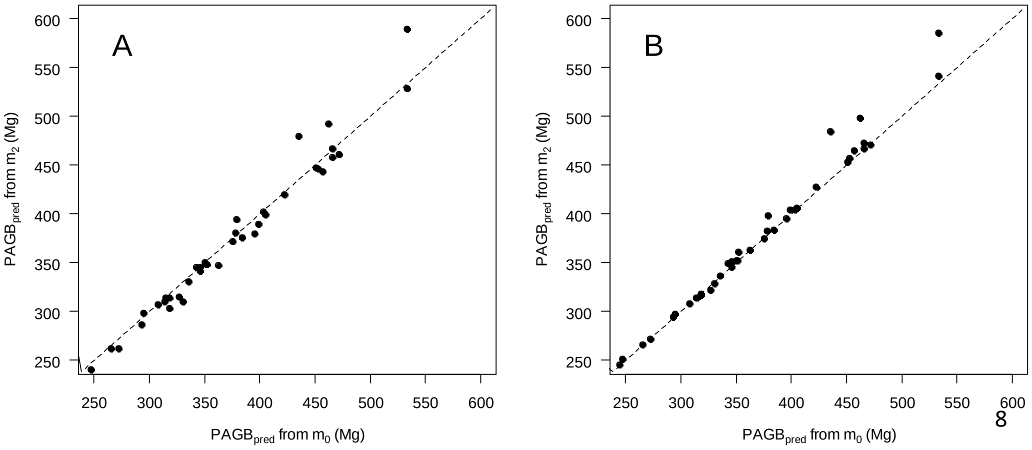

Figure 1. Above ground biomass estimation (in Mg) of 46 1-ha plots using the pantropical model m_0 (X axis) and a combination of m_0 (for trees without information on crown depth) and our model m_2 (for trees with information on crown depth) (Y axis). In plot A, we used m_2 to predict the biomass of all trees with information on crown depth, while we only used m_2 on trees with D ≥100 cm in plot B.

Closing a gap in tropical forest biomass estimation: Accounting for crown mass variation in pantropical allometries

P. Ploton[1,2], N. Barbier[1], S.T. Momo[1,3], M. Réjou-Méchain[1,4,5], F. Boyemba Bosela[6], G. Chuyong[7], G. Dauby[8,9], V. Droissart[1,10], A. Fayolle[11], R.C. Goodman[12], M. Henry[13], N.G. Kamdem[3], J. Katembo Mukirania[6], D. Kenfack[14], M Libalah[3], A. Ngomanda[15], V. Rossi[4,16], B. Sonké[3], N. Texier[1,3], D. Thomas[17], D. Zebaze[3], P. Couteron[1], U. Berger[18] and R. Pélissier[1]

[1] Institut de Recherche pour le Développement, UMR-AMAP, Montpellier, France

[2] Institut des sciences et industries du vivant et de l'environnement, Montpellier, France

[3] Laboratoire de Botanique systématique et d'Ecologie, Département des Sciences Biologiques, Ecole Normale Supérieure, Université de Yaoundé I, Yaoundé, Cameroon

[4] Centre de coopération internationale en recherche agronomique pour le développement, Montpellier, France

[5] Geomatics and Applied Informatics Laboratory (LIAG), French Institute of Pondicherry, Puducherry, India

[6] Faculté des Sciences, Université de Kisangani, Kisangani, Democratic Republic of Congo

[7] Department of Botany and Plant Physiology, University of Buea, Buea, Cameroon

[8] Institut de Recherche pour le Développement, UMR-DIADE, Montpellier, France

[9] Evolutionary Biology and Ecology, Faculté des Sciences, Université Libre de Bruxelles, Brussels, Belgium

[10] Herbarium et Bibliothèque de Botanique africaine, Université Libre de Bruxelles, Brussels, Belgium

[11] Research axis on Forest Resource Management of the Biosystem engineering (BIOSE), Gembloux Agro-Bio Tech, Université de Liège, Gembloux, Belgium

[12] Yale School of Forestry and Environmental Studies, New Haven, USA

[13] Food and Agriculture Organization of the United Nations, UN-REDD Programme, Rome, Italy

[14] Center for Tropical Forest Science, Harvard University, Cambridge, USA

[15] Institut de Recherche en Ecologie Tropicale, Libreville, Gabon

[16] Université de Yaoundé I, UMMISCO, Yaoundé, Cameroon

[17] Department of Botany and Plant Pathology, Oregon State University, Corvallis, USA

[18] Technische Universität Dresden, Faculty of Environmental Sciences, Institute of Forest Growth and Forest Computer Sciences, Tharandt, Germany

Correspondence to: P. Ploton (pierre.ploton@ird.fr)

Abstract

Accurately monitoring tropical forest carbon stocks is an outstanding challenge. Allometric models that consider tree diameter, height and wood density as predictors are currently used in most tropical forest carbon studies. In particular, a pantropical biomass model has been widely used for approximately a decade, and its most recent version will certainly constitute a reference in the coming years. However, this reference model shows a systematic bias for the largest trees. Because large trees are key drivers of forest carbon stocks and dynamics, understanding the origin and the consequences of this bias is of utmost concern. In this study, we compiled a unique tree mass dataset on 673 trees destructively sampled in five tropical countries (101 trees > 100 cm in diameter) and an original dataset of 130 forest plots (1 ha) from central Africa to quantify the prediction error of biomass allometric models at the individual and plot levels when explicitly accounting or not ~~accounting~~ for crown mass variations. We first showed that the proportion of crown to total tree aboveground biomass is highly variable among trees, ranging from 3 to 88 %. This proportion was constant on average for trees < 10 Mg (mean of 34 %) but, above this threshold, increased sharply with tree mass and exceeded 50 % on average for trees ≥ 45 Mg. This increase coincided with a progressive deviation between the pantropical biomass model estimations and actual tree mass. Accounting for a crown mass proxy in a newly developed model consistently removed the bias observed for large trees (> 1 Mg) and reduced the range of plot-level error from -23–16 % to 0–10 %. The disproportionally higher allocation of large trees to crown mass may thus explain the bias observed recently in the reference pantropical model. This bias leads to far-from-negligible, but often overlooked, systematic errors at the plot level and may be easily corrected by accounting for a crown mass proxy for the largest trees in a

stand, thus suggesting that the accuracy of forest carbon estimates can be significantly improved at a minimal cost.

1 Introduction

Monitoring forest carbon variation in space and time is both a sociopolitical challenge for climate change mitigation and a scientific challenge, especially in tropical forests, which play a major role in the ~~world ~~global carbon balance (Hansen et al., 2013; Harris et al., 2012; Saatchi et al., 2011). Significant milestones have been reached in the last decade thanks to the development of broad-scale remote sensing approaches (Baccini et al., 2012; Malhi et al., 2006; Mitchard et al., 2013; Saatchi et al., 2011). However, local forest biomass estimations commonly represent the foundation for the calibration and validation of remote sensing models.~~However, local forest biomass estimations are still the bedrock of most (if not all) of these approaches for the calibration and validation of remote sensing models.~~ As a consequence, uncertainties and errors in local biomass estimations may propagate dramatically to broad-scale forest carbon stock assessment (Avitabile et al., 2011; Pelletier et al., 2011; Réjou-Méchain et al., 2014). Aboveground biomass (*AGB*) is the major pool of biomass in tropical forests (Eggleston et al., 2006). The *AGB* of a tree (or *TAGB*) is generally predicted by empirically derived allometric equations that use measurements of the size of an individual tree as predictors of its mass (Clark and Kellner, 2012). Among these predictors, diameter at breast height (D) and total tree height (H) are often used to capture volume variations between trees, whereas wood density (ρ) is used to convert volume to dry mass (Brown et al., 1989). The most ~~currently ~~frequently used allometric equations for tropical forests (Chave et al., 2005, 2014) have the following form: $TAGB = \alpha * (D^2 * H * \rho)^\beta$, where diameter, height and wood density are combined into a single compound variable related to dry mass through a power law of parameters α and β. This model form, referred to hereafter as our reference allometric model form, performs well when $\beta = 1$ or close to 1 (Chave et al., 2005, 2014), meaning that trees can roughly be viewed as a standard geometric solid for which the parameter α determines the shape (or form factor) of the geometric approximation. However, the uncertainty associated with this model is still very high, with an average error of 50 % at the tree level, illustrating the high natural variability of mass between trees with similar D, H and ρ values. More importantly, this reference allometric model shows a systematic underestimation of *TAGB* of approximately 20 % in average for the heaviest trees (> 30 Mg) (Fig. 2 in Chave et al. 2014), which may contribute strongly to uncertainty in

Formatted: Highlight

Formatted: Highlight

biomass estimates at the plot level. It is often argued that, by definition, the least-squares regression model implies that tree-level errors are globally centered on 0, thus limiting the plot-level prediction error to approximately 5-10 % for a standard 1-ha forest plot (Chave et al., 2014; Moundounga Mavouroulou et al., 2014). However, systematic errors associated with large trees are expected to disproportionally propagate to plot-level predictions because of their prominent contribution to plot *AGB* (Bastin et al., 2015; Clark and Clark, 1996; Sist et al., 2014; Slik et al., 2013; Stephenson et al., 2014). Thus, identifying the origin of systematic errors in such biomass allometric models is a prerequisite for improving local biomass estimations and thus limiting the risk of uncontrolled error propagation to broad-scale extrapolations.

As foresters have known for decades, it is reasonable to approximate stem volume using a geometric shape. Such an approximation, however, is questionable for assessing the total tree volume, including the crown. Because β is generally close to 1 in the reference allometric model, the relative proportion of crown to total tree mass (or crown mass ratio) directly affects the adjustment of the tree form factor α (e.g., Cannell 1984). Moreover, the crown mass ratio is known to vary greatly between species, reflecting different strategies of carbon allocation. For instance, Cannell (1984) observed that coniferous species have a lower proportion of crown mass (10-20 %) than tropical broadleaved species (over 35 %), whereas temperate softwood species were found to have a lower and less variable crown mass ratio (20-30 %) than temperate hardwood species (20-70 %; Freedman et al., 1982; Jenkins et al., 2003). In the tropics, distinct crown size allometries have been documented among species functional groups (Poorter *et.al.* 2003; Poorter, Bongers, et Bongers 2006; Van Gelder, Poorter, et Sterck 2006). For instance, at comparable stem diameters, pioneer species tend to be taller and to have shorter and narrower crowns than understory species (Poorter et al., 2006). These differences reflect strategies of energy investment (tree height *vs.* crown development) ~~that~~ are likely to result in different crown mass ratios among trees with similar D^2, H and ρ values. Indeed, Goodman et al. (2014) obtained a substantially improved biomass allometric model when crown diameter was incorporated into the equation to account for individual variation in crown size.

Destructive data on tropical trees featuring information on both crown mass and classical biometric measurements (D, H, ρ) are scarce and theoretical work on crown properties largely remains to be validated with field data. In most empirical studies published to date, crown mass models use trunk diameter as a single predictor (e.g., Nogueira et al.

2008; Chambers et al. 2001). Such models often provide good results ($R^2 \geq 0.9$), which reflect the strong biophysical constraints exerted by the diameter of the first pipe (the trunk) on the volume of the branching network (Shinozaki et al., 1964). However, theoretical results suggest that several crown metrics would scale with crown mass. For instance, Mäkelä et Valentine (2006) modified the allometric scaling theory (Enquist, 2002; West et al., 1999) by incorporating self-pruning processes into the crown. The authors showed that crown mass is expected to be a power function of the total length of the branching network, which they approximated by crown depth (i.e., total tree height minus trunk height). The construction of the crown and its structural properties have also largely been studied in the light of the mechanical stresses faced by trees (such as gravity and wind; e.g., McMahon et Kronauer 1976; Eloy 2011). Within this theoretical frame, crown mass can also be expressed as a power function of crown diameter (King and Loucks, 1978).

In the present study, we used a unique tree mass dataset containing crown mass information on 673 trees from five tropical countries and a network of forest plots covering 130 ha in central Africa to (i) quantify the variation in crown mass ratio in tropical trees; (ii) assess the contribution of crown mass variation to the reference pantropical model error, either at the tree level or when propagated at the plot level; and (iii) propose a new operational strategy to explicitly account for crown mass variation in biomass allometric equations. We hypothesize that the variation in crown mass ratio in tropical trees is a major source of error in current biomass allometric models and that accounting for this variation would significantly reduce uncertainty associated with plot-level biomass predictions.

2 Materials and Methods

2.1 Biomass data

We compiled tree *AGB* data from published and unpublished sources providing information on crown mass for 673 tropical trees belonging to 132 genera (144 identified species), with a wide tree size range (i.e., diameter at breast height, *D*: 10-212 cm) and aboveground tree masses of up to 76 Mg. An unpublished dataset for 77 large trees (with $D \geq 67$ cm) was obtained from the fieldwork of PP, NB and SM in semi-deciduous forests of Eastern Cameroon (site characteristics and field protocol in Supplement S1.1 and S1.2.1). The remaining datasets were gathered from relevant published studies: 29 trees from Ghana (Henry et al., 2010), 285 trees from Madagascar (Vieilledent et al., 2011), 51 trees from Peru

(Goodman et al., 2014, 2013), 132 trees from Cameroon (Fayolle *et al.*, 2013) and 99 trees from Gabon (Ngomanda *et al.*, 2014). The whole dataset is available from the Dryad Data Repository (http://dx.doi.org/10.5061/dryad.f2b52), with details about the protocol used to integrate data from published studies presented in the Supplementary Information (S1.2.2). For the purpose of some analyses, we extracted from this crown mass database (hereafter referred to as Data_{CM1}) a subset of 541 trees for which total tree height was available (Data_{CM2}; all but Fayolle *et al.* 2013) and another subset of 119 trees for which crown diameter was also available (Data_{CD}; all but Vieilledent *et.al.* 2011, Fayolle *et.al.* 2013, Ngomanda *et.al.* 2014 and 38 trees from our unpublished dataset). Finally, we used as a reference the data from Chave et al. (2014) on the total mass (but not crown mass) of 4,004 destructively sampled trees of many different species from all around the tropical world (Data_{REF}).

2.2 Forest inventory data

We used a set of 81 large forest plots (> 1 ha), covering a total area of 130 ha, to propagate *TAGB* estimation errors to plot-level predictions. The forest inventory data contained the taxonomic identification of all trees with a diameter at breast height (D) ≥ 10 cm, as well as total tree height measurements (H) for a subset of trees, from which we established plot-level *H vs. D* relationships to predict the tree height of the remaining trees. Details about the inventory protocol along with statistical procedures used to compute plot *AGB* (or *PAGB*) from field measurements are provided in the Supplementary Information (S1.3). Among these plots, 80 were from a network of 1-ha plots established in humid evergreen to semi-deciduous forests belonging to 13 sites in Cameroon, Gabon and the Democratic Republic of Congo (unpublished data[1]). In addition, we included a 50-ha permanent plot from Korup National Park, in the evergreen Atlantic forest of western Cameroon (Chuyong et al., 2004), which we subdivided into 1-ha subplots. Overall, the inventory data encompassed a high diversity of stand structural profiles ranging from open-canopy *Marantaceae* forests to old-growth monodominant *Gilbertiodendron dewevrei* stands and including mixed *terra firme* forests with various levels of degradation.

---

[1] metadata available at http://vmamapgn-test.mpl.ird.fr:8080/geonetwork/srv/eng/search#|7dd46c7d-db2f-4bb0-920a-8afe4832f1b3

2.3 Allometric model fitting

We fitted the pantropical allometric model of Chave et al. (2014) to log-transformed data using ordinary least-squares regression:

$$\ln(TAGB) = \alpha + \beta * \ln(D^2 * H * \rho) + \varepsilon \qquad (1)$$

with *TAGB* (in kg) representing the aboveground tree mass, *D* (in cm) the tree stem diameter, *H* (in m) the total tree height, ρ (in g.cm^{-3}) the wood density and ε the error term, which is assumed to follow a normal distribution N ~ (0, RSE²), where RSE is the residual standard error of the model. This model, denoted m_0, was considered as the reference model.

To assess the sensitivity of m_0 to crown mass variations, we built a model (m_1) that restricted the volume approximation to the trunk compartment and included actual crown mass as an additional covariate:

$$ln(TAGB) = \alpha + \beta * \ln(D^2 * Ht * \rho) + \gamma * \ln(Cm) + \varepsilon \qquad (2)$$

Formatted: Highlight

with *Cm* representing the crown mass (in kg) and *Ht* the trunk height (i.e., height of the first ~~living~~ main branch, in m). Note that model m_1 cannot be operationally implemented (which would require destructive measurements of crowns) but quantifies the maximal improvement that can be made through the inclusion of crown mass proxies in a biomass allometric model.

Formatted: Highlight

2.4 Development of crown mass proxies

We further developed crown mass proxies to be incorporated in place of the real crown mass (*Cm*) in the allometric model m_1. From preliminary tests of various model forms (see Appendix A), we selected a crown mass sub-model based on a volume approximation similar to that made for the trunk component (sm_1):

$$\ln(Cm) = \alpha + \beta * \ln(D^2 * Hc * \rho) + \varepsilon \qquad (3)$$

where *D* is the trunk diameter at breast height (in cm) and *Hc* the crown depth (that is *H* – *Ht*, in m), available in our dataset $Data_{CM2}$ (n=541).

In this sub-model, tree crowns of short stature but large width are assigned a small *Hc*, thus a small mass, whereas the volume they occupy is more horizontal than vertical. We thus tested in sub-model sm_2 (eq. 4) whether using the mean crown size (eq. 5), which accounts for both *Hc* and *Cd* (the crown diameter in m available in our dataset $Data_{CD}$ (n=119)) reduces the error associated with sm_1:

$$\ln(Cm) = \alpha + \beta * \ln(D^2 * Cs * \rho) + \varepsilon \qquad (4)$$

$$Cs = \frac{(Hc+Cd)}{2} \qquad (5)$$

Finally, Sillett et al. (2010) showed that for large, old trees, a temporal increment of *D* and *H* poorly reflects the high rate of mass accumulation within crowns. We thus hypothesized that the relationship between *Cm* and $D^2*Hc*\rho$ (or $D^2*Cs*\rho$) depends on tree size and fitted a quadratic (second-order) polynomial model to account for this phenomenon (Niklas, 1995), if any:

$$\ln(Cm) = \alpha + \beta * \ln(D^2 * Hc * \rho) + \gamma * \ln(D^2 * Hc * \rho)^2 + \varepsilon \qquad (6)$$

$$\ln(Cm) = \alpha + \beta * \ln(D^2 * Cs * \rho) + \gamma * \ln(D^2 * Cs * \rho)^2 + \varepsilon \qquad (7)$$

where eqs. 6 and 7 are referred to as sub-models 3 and 4, respectively.

2.5 Model error evaluation

2.5.1 Tree-level

From biomass allometric equations, we estimated crown mass (denoted Cm_{est}) or total tree aboveground mass (denoted $TAGB_{est}$) including (Baskerville, 1972) bias correction during back-transformation from the logarithmic scale to the original mass unit (i.e., kg). In addition to classical criteria of model fit assessment (adjusted R^2, Residual Standard Error, Akaike Information Criterion), we quantified model uncertainty based on the distribution of individual relative residuals (in %), which is defined as follows:

$$s_i = \left(\frac{Y_{est,i} - Y_{obs,i}}{Y_{obs,i}}\right) * 100 \qquad (8)$$

where $Y_{obs,i}$ and $Y_{est,i}$ are the crown or tree biomass values in the calibration dataset (i.e., measured in the field) and those allometrically estimated for tree *i*, respectively. We reported the median of $|s_i|$ values, hereafter referred to as "*S*", as an indicator of model precision. For a tree biomass allometric model to be unbiased, we expect s_i to be locally centered on zero for any given small range of the tree mass gradient. We thus investigated the distribution of s_i values with respect to tree mass using local regression (loess method; Cleveland, Grosse & Shyu 1992).

2.5.2 Plot level

Allometric models are mostly used to make plot-level AGB predictions from non-destructive forest inventory data. Such plot-level predictions are obtained by simply summing individual predictions over all trees in a plot ($PAGB_{pred} = \sum_i TAGB_{pred}$). Prediction errors at the tree level are thus expected to yield an error at the plot level, which may depend on the actual tree mass distribution in the sample plot when the model is locally biased. To account for this effect, we developed a simulation procedure, implemented in two steps, ~~that~~ which propagated $TAGB_{pred}$ errors to $PAGB_{pred}$. The first step consists in attributing to each tree i in a given plot a value of $TAGB_{sim}$ corresponding to the actual AGB of a similar felled tree selected in Data_{REF} based on its nearest neighbor in the space of the centered-reduced variables D, H and ρ (here taken as species average from Dryad Global Wood Density Database, Chave et al., 2009; Zanne et al., 2009). In a second step, the simulation propagates individual errors of a given allometric model using the local distribution of s_i values as predicted by the loess regression: For each $TAGB_{sim}$, we drew a s_{sim} value from a local normal distribution fitted with the loess parameters (i.e., local mean and standard deviation) predicted for that particular $TAGB_{sim}$. Thus, we generated for each 1-ha plot a realistic $PAGB_{sim}$ (i.e., based on observed felled trees) with repeated realizations of a plot-level prediction error (in %) computed for n trees as follows:

$$\mathrm{S_{plot}} = \frac{\sum_{i=1}^{n}(\mathrm{s_{sim}(i)*TAGB_{sim}(i)})}{\sum_{i=1}^{n}\mathrm{TAGB_{sim}(i)}}. \quad (9)$$

For each of the simulated plots, we provided the mean and standard deviation of 1000 realizations of the plot-level prediction error.

All analyses were performed with R statistical software 2.15.2 (R Core Team, 2012), using packages lmodel2 (Legendre, 2011), segmented (Muggeo, 2003), FNN (Beygelzimer et al., 2013) and msir (Scrucca, 2011).

3 Results

3.1 Contribution of crown to tree mass

Our crown mass database (Data_{CM1}; 673 trees, including 128 trees > 10 Mg) revealed a huge variation in the contribution of crown to total tree mass, ranging from 2.5 to 87.5 % of total aboveground biomass, with a mean of 35.6 % (± 16.2 %). Despite this variation, a linear regression (model II) revealed a significant increase in the crown mass ratio with tree mass of approximately 3.7 % per 10 Mg (Fig. 1-A). A similar trend was observed at every site, except for the Ghana dataset (Henry et al. 2010), for which the largest sampled tree (72 Mg) had a rather low crown mass ratio (46 %). Overall, this trend appeared to have been driven by the largest trees in the database (Fig. 1-B). Indeed, the crown mass ratio appeared to be nearly constant for trees ≤ 10 Mg with an average of 34.0 % (± 16.9 %), and then to increase progressively with tree mass, exceeding 50 % on average for trees ≥ 45 Mg.

3.2 Crown mass sub-models

All crown mass sub-models provided good fits to our data ($R^2 \geq 0.9$, see Table 1). However, when information on crown diameter was available (Data_{CD}), models that included mean crown size in the compound variable (i.e., *Cs*, a combination of crown depth and diameter, in sm_2 and sm_4) gave lower AICs and errors (RSE and *S*) than models that included the simpler crown depth metric (i.e., *Hc* in sm_1 and sm_3). The quadratic model form provided a better fit than the linear model form (e.g., sm_3 *vs.* sm_1 fitted on Data_{CM2}), which can be explained by the non-linear increase in crown mass with either of the two proxy variables ($D^2*Hc*\rho$ or $D^2*Cs*\rho$). The slope of the relationship between crown mass and, for example, $D^2*Hc*\rho$ presented a breaking point at approximately 7.5 Mg (Davies' test $P < 0.001$) that was not captured by sub-model sm_1 (Fig. 2-A, full line), leading to a substantial bias in back-transformed crown mass estimations (approximately 50 % of observed crown mass for $Cm_{obs} \geq 10$ Mg, Fig. 2-B). The quadratic sub-model sm_3 provided fairly unbiased crown mass estimations (Fig. 2-C). Because the first-order term was never significant in the quadratic sub-models, we retained only the second-order term as a crown mass proxy in the biomass allometric models (i.e., $(D^2 * Hc * \rho)^2$ for model m_2 and $(D^2 * Cs * \rho)^2$ for model m_3).

3.3 Accounting for crown mass in biomass allometric models

The reference model (m_0) proposed by Chave et al. (2014) presented, when fitted to DATA_{REF}, a bias that was a function of tree mass, with a systematic *AGB* over-estimation for trees < approximately 10 Mg and an under-estimation for larger trees, reaching approximately 25 % for trees greater than 30 Mg (Fig. 3-A). This bias pattern reflected a breaking point in the relationship between $D^2*H*\rho$ and $TAGB_{obs}$ (Davies' test $P < 0.001$) located at approximately 10 Mg (Fig. 3-B). Accounting for actual crown mass (*Cm*) in the biomass allometric model (i.e., model m_1) corrected for a similar bias pattern observed when m_0 was fitted to $DATA_{CM2}$ (Fig. 4-A). This result shows that variation in crown mass among trees is a major source of bias in the reference biomass allometric model, m_0.

Using our simulation procedure, we propagated individual prediction errors of m_0 and m_1 to the 130 1-ha field plots from central Africa (Fig. 4-B). This process revealed that the reference pantropical model (m_0) led to an average plot-level relative prediction error (S_{plot}) ranging from -23 % to +16 % (mean = +6.8 %) on $PAGB_{pred}$, which dropped to +1 to +4 % (mean = +2.6 %) when the model accounted for crown mass (m_1).

Because in practice crown mass cannot be routinely measured in the field, we tested the potential of crown mass proxies to improve biomass allometric models. Model m_2, which used a compound variable integrating crown depth i.e., $(D^2 * Hc * \rho)^2$ as a proxy of crown mass outperformed m_0 (Table 2). Although the gain in precision (RSE and S) over m_0 was rather low, the model provided the ~~striking~~ major advantage of being free of significant local bias on large trees (> 1 Mg; Fig. 5-A). At the plot level, this model provided a much higher precision (0 to 10 % on $PAGB_{pred}$) and a lower bias (average error of 5 %) than the reference pantropical model m_0 (Fig. 5-B). Using a compound variable integrating crown size i.e., $(D^2 * Cs * \rho)^2$ as a crown mass proxy (model m_3), thus requiring both crown depth and diameter measurements, significantly improved model precision (m_3 *vs.* m_2, Table 2) while preserving the relatively unbiased distribution of relative residuals (results not shown).

4 Discussion

Using a dataset of 673 individuals including most of the largest trees that have been destructively sampled to date, we discovered tremendous variation in the crown mass ratio among tropical trees, ranging from 3 to 88 %, with an average of 36 %. This variation was not independent of tree size, as indicated by a marked increase in the crown mass ratio with tree mass for trees ≥ 10 Mg. This threshold was mirrored by~~echoed~~ a breaking point in the relationship between total tree mass and the compound predictor variable used in the reference allometric model of Chave et al. (2014). When the compound variable is limited to trunk mass prediction, and a crown mass predictor is added to the model, the bias towards large trees is significantly reduced. As a consequence, error propagation to plot-level *AGB* estimations is largely reduced. In the following section, we discuss the significance and implication of these results from both an ecological and a practical point of view with respect

Formatted: Highlight

to resource allocation to the tree compartments and to carbon storage in forest aboveground biomass.

4.1 Crown mass ratio and the reference biomass model error

We observed an overall systematic increase in the crown mass ratio with tree mass. This ontogenetic trend has already been reported for some tropical canopy species (O'Brien et al., 1995) and likely reflects changes in the pattern of resource allocation underlying crown edification in most forest canopy trees (Barthélémy and Caraglio, 2007; Hasenauer and Monserud, 1996; Holdaway, 1986; Moorby and Wareing, 1963; Perry, 1985). The overall increase in the carbon accumulation rate with tree size is a well-established trend (Stephenson et al., 2014), but the relative contribution of the trunk and the crown to that pattern has rarely been investigated, particularly on large trees for which branch growth monitoring involves a tremendous amount of work. Sillett et al. (2010) collected a unique dataset in this regard, with detailed growth measurements on very old (up to 1850 years) and large (up to 648 cm *D*) individuals of *Eucalyptus regnans* and *Sequoia sempervirens* species. For these two species, the contribution of crown to *AGB* growth increased linearly with tree size and thus the crown mass ratio. We observed the same tendency in our data for trees ≥ 10 Mg (typically with *D* > 100 cm). This result thus suggests that biomass allometric relationships may differ among small and large trees, thus explaining the systematic underestimation of *AGB* for large trees observed by Chave et al. (2014). The latter authors suggested that underestimations ~~that~~ induced by ~~this~~ their model ~~underestimation~~ was due to a potential "majestic tree" sampling bias, in which scientists would have more systematically sampled trees with well-formed boles and healthy crowns. We agree that such an effect cannot be completely ruled out, and it is probably all the more significant that trees ≥ 10 Mg represent only 3 % of the reference dataset of Chave et al. (2014). Collecting more field data on the largest tree size classes should therefore constitute a priority if we are to improve multi-specific, broad-scale allometric models, and the recent development of non-destructive *AGB* estimation methods based on terrestrial LiDAR data should help in this regard (e.g., Calders et al., 2015). However, regardless of whether the non-linear increase in crown mass ratio with tree mass held to a sampling artifact, we have shown that it was the source of systematic error in the reference model that used a unique geometric approximation with an average form factor for all trees. This finding agrees with the results of Goodman *et al.* (2014) in Peru, who found significant improvements in biomass estimates of large trees when biomass models included tree crown radius, thus partially accounting for crown ratio variations. Identifying predictable patterns of crown mass ratio variation, as performed for crown size allometries specific to some functional groups (Poorter *et.al.* 2003; Poorter, Bongers, et Bongers 2006; Van Gelder, Poorter, et Sterck 2006), therefore appears to be a potential way to improve allometric models performance.

4.2 Model error propagation depends on targeted plot structure

The reference pantropical model provided by Chave *et al.* (2014) presents a bias pattern that is a function of tree size (i.e., average over-estimation of small tree *AGB* and average underestimation of large tree *AGB*). Propagation of individual errors to the plot level therefore depends on tree size distribution in the sample plot, with over- or under-estimations depending on the relative importance of small or large trees in the stand (e.g., young secondary forests *vs.* old-growth forests; see Appendix B for more information on the interaction between model error, plot structure and plot size). This effect is not consistent with the general assumption that individual errors should compensate at the plot level. Although the dependence of error propagation on tree size distribution has already been raised (Magnabosco Marra et al., 2015; Mascaro et al., 2011), it is generally omitted from error propagation procedures (e.g., Picard, Bosela, et Rossi 2014; Moundounga Mavouroulou et al. 2014; Chen, Vaglio Laurin, et Valentini 2015). When propagating local bias to ~~our~~ 130 1-ha plots in central Africa, the reference pantropical model led to plot-level errors ranging from -15% to +8%. The presence of large trees, in particular their relative contribution to stand total *AGB*, explained most ~~of~~ between-plots error variation (Appendix B). We can therefore hypothesize that in the Neotropics where large trees are less common in forests than in the Paleotropics (Lewis et al., 2013; Slik et al., 2013), the model would more systematically over-estimate plots *AGB*. Interestingly, most of the plots undergoing a systematic *AGB* under-estimation (i.e. high number of large trees) were located in the Atlantic forests of Western Cameroon (Korup NP), where large individuals of *Lecomtedoxa klaineana* (Pierre ex Engl) – a so-called “biomass hyperdominant” species (*sensu* Bastin et al. 2015) – are particularly abundant. Interactions between model error and forest structure may thus also hinder the detection of spatial variations in forests *AGB* between forest types as well as at local scales e.g., between patches dominated or not by *Lecomtedoxa klaineana* trees. At the landscape or regional scale, plot-level errors may average out if the study area is a mosaic of forests with varying tree size distributions. However, if plot estimations are used to calibrate remote

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

sensing products, individual plot errors may propagate as a systematic bias in the final extrapolation (Réjou-Méchain et al. 2014).

4.3 Accounting for crown mass variation in allometric models

We propose a modeling strategy that decomposes total tree mass into trunk and crown masses. A direct benefit of addressing these two components separately is that it should reduce the error in trunk mass estimation because the trunk form factor is less variable across species than the whole-tree form factor (Cannell, 1984). We modeled tree crown using a geometric solid whose basal diameter and height were the trunk diameter and crown depth, respectively. Crown volume was thus considered as the volume occupied by branches if they were squeezed onto the main stem ("as if a ring were passed up the stem"; Cannell 1984). Using a simple linear model to relate crown mass to the geometric approximation (sm_1, sm_2) led to an under-estimation bias that gradually increased with crown mass (Fig. 2-B). A similar pattern was observed on all crown mass models based on trunk diameter (Appendix A) and reflected a significant change in the relationship between the two variables with crown size. Consistently, a second-order polynomial model better captured such a non-linear increase in crown mass with trunk diameter-based proxies and thus provided unbiased crown mass estimates (Fig. 2-C). Our results agree with those of Sillett et al. (2010), who showed that ground-based measurements such as trunk diameter do not properly render the high rate of mass accumulation in large trees, notably in tree crowns, and may also explain why the dynamics of forest biomass are inferred differently from top-down (e.g., airborne LiDAR) or bottom-up views (e.g., field measurement; Réjou-Méchain et al., 2015).

Changes in trunk and crown mass along tree ontogeny are not independent and indeed, both variables appeared tightly correlated in our dataset. Including crown mass (or a proxy for this variable) as an additive covariate to the trunk mass proxy may thus raise the debate on collinearity between predictors in biomass allometry models (see Picard et al., 2015; Sileshi, 2014). For instance, models m_1 and m_2 calibrated on Data$_{CM2}$ led to a variance inflation factor (VIF) of 5.4 and 8.8, respectively, which is higher than the range of values commonly considered as critical (2-5, Sileshi, 2014). Yet, we have shown than the inclusion of a separate crown component to the models reduced model residuals (greater precision) and improved their distribution over the *AGB* gradient (greater accuracy), because it allowed us to capture a general trend in our dataset of a relative increase of crown mass proportion with tree mass. Assuming that this phenomenon holds in new sets of tropical trees and that we adequately

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

sampled the correlation structure between crown and trunk masses, ~~then~~ the issue of predictors collinearity should therefore not dramatically inflate model~~s~~ prediction errors (Picard et al., 2015).

Formatted: Highlight

Formatted: Highlight

From a practical point of view, our tree biomass model m_2, which requires only extra information on trunk height (if total height is already measured) provides a better fit than the reference pantropical model and removes estimation bias on large trees. In scientific forest inventories, total tree height is often measured on a sub-sample of trees, including most of the largest trees in each plot, to calibrate local allometries between H and D. We believe that measuring trunk height on those trees does not represent a cumbersome amount of additional effort because trunk height is much more easily measured than total tree height. We thus recommend using model m_2 —at least for the largest trees, i.e., those with $D \geq 100$ cm — and encourage future studies to assess its performance from independent datasets. Including more detailed crown measurements into biomass allometric equations could also become a reasonable option in a near future, provided the development of new technologies, like (mobile) terrestrial Lidar scanning, will make it possible to easily extract crown data and gather large-scale datasets.

Formatted: Highlight

Appendix A: Crown mass sub-models

A.1 Method

Several tree metrics are expected to scale with crown mass, particularly crown height (Mäkelä and Valentine, 2006), crown diameter (King and Loucks, 1978) or trunk diameter (e.g., Nogueira et al. 2008; Chambers et al. 2001). In this study, we tested whether any of these variables (i.e., trunk diameter, crown height and crown diameter) prevailed over the others in explaining crown mass variations. Power functions were fitted in log-transformed form using ordinary least-squares regression techniques (models sm_{1-X}):

$$\ln(Cm) = \alpha + \beta * \ln(X) + \varepsilon \qquad \text{(A1)}$$

where Cm is the crown mass (in kg); X is the structural variable of interest, namely D for trunk diameter at breast height (in cm), Hc for crown depth (in m), or Cd for crown diameter (in m); α and β are the model coefficients and is ε the error term assumed to follow a normal distribution.

We also assessed the predictive power of the three structural variables on crown mass while controlling for variations in wood density (ρ, in g.cm^{-3}), leading to models sm$_{2\text{-X}}$:

$$\ln(Cm) = \alpha + \beta * \ln(X) + \gamma * \ln(\rho) + \varepsilon \qquad \text{(A2)}$$

where γ is the model coefficient of ρ.

Similarly to the cylindrical approximation of a tree trunk, we further established a compound variable for tree crown based on *D* and *Hc*, leading to model sm$_3$:

$$\ln(Cm) = \alpha + \beta * \ln(D^2 * Hc * \rho) + \varepsilon \qquad \text{(A3)}$$

where crown height is a proxy for the length of the branching network. Results obtained using sm$_3$ are presented in the manuscript as well as in this appendix for comparison with those obtained using sm$_{1\text{-x}}$ and sm$_{2\text{-x}}$.

A.2 Results & Discussion

Among the three structural variables tested as proxies for crown mass, trunk diameter provided the best results. Model 1-D presented a high R² (0.88), but its precision was low, with an *S* (i.e., the median of unsigned s_i values) of 43 % (Table A1). Moreover, model error increased appreciably with crown mass (Fig. A1, caption A). For instance, model estimations for an observed crown mass of approximately 20 Mg ranged between 5 and 55 Mg. Nevertheless, sm$_{1\text{-D}}$ outperformed sm$_{1\text{-Hc}}$ (Data$_{\text{CM2}}$, AIC of 1182 *vs.* 1603, respectively) and was slightly better than sm$_{1\text{-Cd}}$ (Data$_{\text{CD}}$, AIC of 257 *vs.* 263, respectively), suggesting that the width of the first branching network pipe is a stronger constraint on branches' mass than the external dimensions of the network (i.e., *Hc*, *Cd*).

The model based on crown depth (sm$_{1\text{-Hc}}$) was subjected to a large error (*S* of c. 80 %; Table A1) and clearly saturated for a crown mass ≥ 10 Mg (Fig. A1, caption B). Because crown depth does not account for branch angle, it does not properly render the length of the branching network. The saturation threshold observed on large crowns supports the observations of Sillett et al. (2010): Tree height, from which crown depth directly derives, levels off in large/adult trees, but mass accumulation—notably within the crowns—continues far beyond this point. It follows that crown depth alone does not allow for the detection of the highest mass levels in large/old tree crowns.

The model based on crown diameter presented a weaker fit than sm$_{1\text{-D}}$, with a higher AIC (Data$_{\text{CD}}$, 263 *vs.* 257) and an individual relative error approximately 10 % higher (*S* of approximately 50 % and 40 %, respectively; Table A1). However, crown diameter appeared more informative regarding the mass of the largest crowns than trunk diameter (Fig. A2, captions A and B). In fact, the individual relative error of $sm_{1\text{-}Cd}$ on crowns ≥ 10 Mg was only 26 %, versus 47 % for $sm_{1\text{-}D}$.

Accounting for variations in wood density improved the model based on trunk diameter. As shown in Fig. A1, using a color code for wood density highlighted a predictable error pattern in model estimations: Trunk diameter tends to over- or under-estimate the crown mass of trees with high or low wood density, respectively. This pattern is corrected for in $sm_{2\text{-}D}$, which presents a lower AIC than $sm_{1\text{-}D}$ (i.e., 1079) and an individual relative error approximately 15 % lower (i.e., 37 %; Table A1). Interestingly, whereas $sm_{2\text{-}D}$ appeared to be more accurate than $sm_{1\text{-}D}$ in its estimations of large crown mass (Fig. A1, caption C), it also presented an under-estimation bias that gradually increased with crown mass. Including ρ in the model based on *Cd* improved the model fit (AIC of 251 *vs.* 262 for $sm_{2\text{-}Cd}$ and $sm_{1\text{-}Cd}$, respectively) and decreased the individual relative error by approximately 15 %. Similarly to $sm_{1\text{-}Cd}$, $sm_{2\text{-}Cd}$ was outperformed by its counterpart based on *D* (AIC of 185). Moreover, the gain in precision in $sm_{2\text{-}Cd}$ was localized on small crowns, whereas estimations on large crowns were fairly equivalent (Fig. A2, caption C-D). Model 2-D was more precise on crowns ≥ 10 Mg, with an individual relative error of 23 % versus 32 % for $sm_{2\text{-}Cd}$.

The strongest crown mass predictor, *D*, was used as the basis of a geometric solid approximating crown volume (D^2*Hc) and, in turn, mass ($D^2*Hc*\rho$) in model sm_3. With one less parameter than $sm_{2\text{-}D}$, sm_3 presented a lower AIC than the former model (i.e., 1012), but the two models provided a fairly similar fit to the observations (RSE of 0.65 *vs.* 0.61 and S of 37 % *vs.* 36 % for $sm_{2\text{-}D}$ and sm_3, respectively). This result indicates that when *D* and ρ are known, information on crown depth is of minor importance for predicting crown mass. However, this conclusion applies to our dataset only because *Hc* might be more informative regarding crown mass variations when considering sites/species with more highly contrasting *D-H* or *D-Hc* relationships.

Similarly to $sm_{2\text{-}D}$, sm_3 presented an under-estimation bias that increased gradually with crown mass (illustrated in Fig. A1 caption D).

Appendix B: Plot-level error propagation

We used the error propagation procedure described in the Methods section of the manuscript to estimate the mean plot-level *AGB* prediction error that could be expected from m_0 calibrated on $DATA_{REF}$ (i.e., the pantropical model proposed in Chave et al. 2014). Model error was propagated on 130 1-ha sample plots of tropical forest in central Africa, a network of 80 1-ha plots (field inventory protocol in Supplement Information S1.3) to which we added 50 1-ha plots from Korup 50-ha permanent plot (Chuyong et al., 2004). We further sub-sampled Korup 50-ha permanent plot in sub-plots of varying sizes (from 25 ha to 0.1 ha) to evaluate the effect of plot size on plot-level *AGB* prediction error.

From the simulated $PAGB_{sim}$ for the 130 1-ha plots, we estimated that the reference pantropical model, m_0, propagated to $PAGB_{pred}$ a mean prediction error (over 1000 realizations of S_{plot}) that ranged between -15 % and +7.7 % (Fig. B1-A), mostly caused by trees with mass ≥ 20 Mg (Fig. B1-B). This trend was particularly evident in the undisturbed evergreen stands of Korup (triangles in Fig. B1-A-B), where patches of *Lecomtedoxa klaineana* (Pierre ex Engl) individuals largely drove the *PAGB* predictions (R^2= 0.87, model II OLS method). This species generates high-statured individuals of high wood density, which frequently exceed 20 Mg and result in underestimates of plot-level biomass. Interestingly, some high-biomass plots could still be over-estimated when $PAGB_{pred}$ was concentrated in trees weighting less than 20 Mg.

As a consequence of m_0 bias concentration in large trees, plot-level prediction errors for the 50 ha in Korup tended to stabilize near 0 for subplots ≥ 5 ha only. Below this threshold (i.e., for subplots ≤ 1 ha), the median error is positive but negative outliers are more frequent (Fig. B2). Indeed, on the one hand, small plots are less likely to encompass large trees and have a positive prediction error of up to approximately +7.5 %. On the other hand, a single large tree can strongly affect $PAGB_{pred}$, occasionally leading to a large underestimation of small plots *AGB* that can exceed -15 % for a 0.25-ha and -20 % for a 0.1-ha subplot.

Author contributions. Conceived and designed the experiments: PP, NB and RP. Collected data (unpublished destructive data and field inventories): SM, BS, NGK, ML, DZ, NT, FBB, JKM, GD, VD. Shared data: GC, DK, DT, AF, AN, MH, RCG. Analyzed the data: PP. Analysis feedback: RP, NB, VR, MRM, UB. Wrote the paper: PP, RP and MRM. Writing feedback: NB, AF, VR, PC, MH, RCG.

Acknowledgments. Destructive data from Cameroon were collected with the financial support from the IRD project PPR FTH-AC ‘Changements globaux, biodiversité et santé en zone forestière d’Afrique Centrale’ and the support and involvement of Alpicam Company. A portion of the plot data were collected with the support of the CoForTips project as part of the ERA-Net BiodivERsA 2011-2012 European joint call (ANR-12-EBID-0002). PP was supported by an Erasmus Mundus PhD grant from the 2013-2016 Forest, Nature and Society (FONASO) doctoral program.

Data Accessibility

Destructive dataset available at http://dx.doi.org/10.5061/dryad.f2b52

REFERENCES

Avitabile, V., Herold, M., Henry, M. and Schmullius, C.: Mapping biomass with remote sensing: a comparison of methods for the case study of Uganda, Carbon Balance Manag., 6(7), 1–14, 2011.

Baccini, A., Goetz, S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P. S. A., Dubayah, R. and Friedl, M. A.: Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps, Nat. Clim. Change, 2(3), 182–185, 2012.

Barthélémy, D. and Caraglio, Y.: Plant Architecture: A Dynamic, Multilevel and Comprehensive Approach to Plant Form, Structure and Ontogeny, Ann. Bot., 99(3), 375–407, doi:10.1093/aob/mcl260, 2007.

Baskerville, G. L.: Use of Logarithmic Regression in the Estimation of Plant Biomass, Can. J. For. Res., 2(1), 49–53, doi:10.1139/x72-009, 1972.

Bastin, J.-F., Barbier, N., Réjou-Méchain, M., Fayolle, A., Gourlet-Fleury, S., Maniatis, D., de Haulleville, T., Baya, F., Beeckman, H. and Beina, D.: Seeing Central African forests through their largest trees, Sci. Rep., 5(13156), doi:doi:10.1038/srep13156, 2015.

Beygelzimer, A., Kakadet, S., Langford, J., Arya, S., Mount, D. and Li, S.: FNN: fast nearest neighbor search algorithms and applications. R package version 1.1., 2013.

Brown, S., Gillespie, A. J. and Lugo, A. E.: Biomass estimation methods for tropical forests with applications to forest inventory data, For. Sci., 35(4), 881–902, 1989.

Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P., Herold, M., Culvenor, D., Avitabile, V., Disney, M., Armston, J. and Kaasalainen, M.: Nondestructive estimates of above-ground biomass using terrestrial laser scanning, edited by S. McMahon, Methods Ecol. Evol., 6(2), 198–208, doi:10.1111/2041-210X.12301, 2015.

Cannell, M. G. R.: Woody biomass of forest stands, For. Ecol. Manag., 8(3–4), 299–312, doi:10.1016/0378-1127(84)90062-8, 1984.

Chambers, J. Q., dos Santos, J., Ribeiro, R. J. and Higuchi, N.: Tree damage, allometric relationships, and above-ground net primary production in central Amazon forest, For. Ecol. Manag., 152(1), 73–84, 2001.

Chave, J., Andalo, C., Brown, S., Cairns, M. A., Chambers, J. Q., Eamus, D., Fölster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J.-P., Nelson, B. W., Ogawa, H., Puig, H., Riéra,

B. and Yamakura, T.: Tree allometry and improved estimation of carbon stocks and balance in tropical forests, Oecologia, 145(1), 87–99, doi:10.1007/s00442-005-0100-x, 2005.

Chave, J., Coomes, D., Jansen, S., Lewis, S. L., Swenson, N. G. and Zanne, A. E.: Towards a worldwide wood economics spectrum, Ecol. Lett., 12(4), 351–366, doi:10.1111/j.1461-0248.2009.01285.x, 2009.

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., Henry, M., Martínez-Yrízar, A., Mugasha, W. A., Muller-Landau, H. C., Mencuccini, M., Nelson, B. W., Ngomanda, A., Nogueira, E. M., Ortiz-Malavassi, E., Pélissier, R., Ploton, P., Ryan, C. M., Saldarriaga, J. G. and Vieilledent, G.: Improved allometric models to estimate the aboveground biomass of tropical trees, Glob. Change Biol., 20(10), 3177–3190, doi:10.1111/gcb.12629, 2014.

Chen, Q., Vaglio Laurin, G. and Valentini, R.: Uncertainty of remotely sensed aboveground biomass over an African tropical forest: Propagating errors from trees to plots to pixels, Remote Sens. Environ., 160, 134–143, doi:10.1016/j.rse.2015.01.009, 2015.

Chuyong GB, Condit R, Kenfack D, Losos E, Sainge M, Songwe NC and Thomas DW: Korup forest dynamics plot, Cameroon. In: Losos EC, Leigh EG Jr (eds) Forest diversity and dynamism: findings from a large-scale plot network. University of Chicago Press, Chicago, 506-516, 2004

Clark, D. B. and Clark, D. A.: Abundance, growth and mortality of very large trees in neotropical lowland rain forest, For. Ecol. Manag., 80(1–3), 235–244, doi:10.1016/0378-1127(95)03607-5, 1996.

Clark, D. B. and Kellner, J. R.: Tropical forest biomass estimation and the fallacy of misplaced concreteness, J. Veg. Sci., 23(6), 1191–1196, doi:10.1111/j.1654-1103.2012.01471.x, 2012.

Cleveland, W. S., Grosse, E. and Shyu, W. M.: Local regression models, Stat. Models S, chapter 8, 309–376, 1992.

Eggleston, H. S., Buendia, L., Miwa, K., Ngara, T. and Tanabe, K.: IPCC guidelines for national greenhouse gas inventories, Inst. Glob. Environ. Strateg. Hayama Jpn., 2006.

Eloy, C.: Leonardo's rule, self-similarity and wind-induced stresses in trees, Phys. Rev. Lett., 107(25), 258101, doi:10.1103/PhysRevLett.107.258101, 2011.

Enquist, B. J.: Universal scaling in tree and vascular plant allometry: toward a general quantitative theory linking plant form and function from cells to ecosystems, Tree Physiol., 22(15-16), 1045–1064, doi:10.1093/treephys/22.15-16.1045, 2002.

Fayolle, A., Doucet, J.-L., Gillet, J.-F., Bourland, N. and Lejeune, P.: Tree allometry in Central Africa: Testing the validity of pantropical multi-species allometric equations for estimating biomass and carbon stocks, For. Ecol. Manag., 305, 29–37, doi:10.1016/j.foreco.2013.05.036, 2013.

Freedman, B., Duinker, P. N., Barclay, H., Morash, R. and Prager, U.: Forest biomass and nutrient studies in central Nova Scotia., Inf. Rep. Marit. For. Res. Cent. Can., (M-X-134), 126 pp., 1982.

Goodman, R. C., Phillips, O. L. and Baker, T. R.: Data from: The importance of crown dimensions to improve tropical tree biomass estimates, [online] Available from: http://dx.doi.org/10.5061/dryad.p281g (Accessed 17 May 2015), 2013.

Goodman, R. C., Phillips, O. L. and Baker, T. R.: The importance of crown dimensions to improve tropical tree biomass estimates, Ecol. Appl., 24(4), 680–698, 2014.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J. and Loveland, T. R.: High-resolution global maps of 21st-century forest cover change, science, 342(6160), 850–853, 2013.

Harris, N. L., Brown, S., Hagen, S. C., Saatchi, S. S., Petrova, S., Salas, W., Hansen, M. C., Potapov, P. V. and Lotsch, A.: Baseline map of carbon emissions from deforestation in tropical regions, Science, 336(6088), 1573–1576, 2012.

Hasenauer, H. and Monserud, R. A.: A crown ratio model for Austrian forests, For. Ecol. Manag., 84(1–3), 49–60, doi:10.1016/0378-1127(96)03768-1, 1996.

Henry, M., Besnard, A., Asante, W. A., Eshun, J., Adu-Bredu, S., Valentini, R., Bernoux, M. and Saint-André, L.: Wood density, phytomass variations within and among trees, and allometric equations in a tropical rainforest of Africa, For. Ecol. Manag., 260(8), 1375–1388, doi:10.1016/j.foreco.2010.07.040, 2010.

Holdaway, M. R.: Modeling Tree Crown Ratio, For. Chron., 62(5), 451–455, doi:10.5558/tfc62451-5, 1986.

Jenkins, J. C., Chojnacky, D. C., Heath, L. S. and Birdsey, R. A.: National-Scale Biomass Estimators for United States Tree Species, For. Sci., 49(1), 12–35, 2003.

King, D. and Loucks, O. L.: The theory of tree bole and branch form, Radiat. Environ. Biophys., 15(2), 141–165, doi:10.1007/BF01323263, 1978.

Legendre, P.: lmodel2: Model II Regression. R package version 1.7-0, See Httpcran R-Proj. Orgwebpackageslmodel2, 2011.

Lewis, S. L., Sonke, B., Sunderland, T., Begne, S. K., Lopez-Gonzalez, G., van der Heijden, G. M. F., Phillips, O. L., Affum-Baffoe, K., Baker, T. R., Banin, L., Bastin, J.-F., Beeckman, H., Boeckx, P., Bogaert, J., De Canniere, C., Chezeaux, E., Clark, C. J., Collins, M., Djagbletey, G., Djuikouo, M. N. K., Droissart, V., Doucet, J.-L., Ewango, C. E. N., Fauset, S., Feldpausch, T. R., Foli, E. G., Gillet, J.-F., Hamilton, A. C., Harris, D. J., Hart, T. B., de Haulleville, T., Hladik, A., Hufkens, K., Huygens, D., Jeanmart, P., Jeffery, K. J., Kearsley, E., Leal, M. E., Lloyd, J., Lovett, J. C., Makana, J.-R., Malhi, Y., Marshall, A. R., Ojo, L., Peh, K. S.-H., Pickavance, G., Poulsen, J. R., Reitsma, J. M., Sheil, D., Simo, M., Steppe, K., Taedoumg, H. E., Talbot, J., Taplin, J. R. D., Taylor, D., Thomas, S. C., Toirambe, B., Verbeeck, H., Vleminckx, J., White, L. J. T., Willcock, S., Woell, H. and Zemagho, L.: Above-ground biomass and structure of 260 African tropical forests, Philos. Trans. R. Soc. B Biol. Sci., 368(1625), 20120295–20120295, doi:10.1098/rstb.2012.0295, 2013.

Magnabosco Marra, D., Higuchi, N., Trumbore, S. E., Ribeiro, G., dos Santos, J., Carneiro, V. M. C., Lima, A. J. N., Chambers, J. Q., Negrón-Juárez, R. I. and Holzwarth, F.: Predicting biomass of hyperdiverse and structurally complex Central Amazon forests–a virtual approach using extensive field data, 2015.

Mäkelä, A. and Harry T.: Crown ratio influences allometric scaling of trees, Ecology, 87(12), 2967–2972, doi:10.1890/0012-9658(2006)87[2967:CRIASI]2.0.CO;2, 2006.

Malhi, Y., Wood, D., Baker, T. R., Wright, J., Phillips, O. L., Cochrane, T., Meir, P., Chave, J., Almeida, S. and Arroyo, L.: The regional variation of aboveground live biomass in old-growth Amazonian forests, Glob. Change Biol., 12(7), 1107–1138, 2006.

Mascaro, J., Litton, C. M., Hughes, R. F., Uowolo, A. and Schnitzer, S. A.: Minimizing Bias in Biomass Allometry: Model Selection and Log-Transformation of Data, Biotropica, 43(6), 649–653, doi:10.1111/j.1744-7429.2011.00798.x, 2011.

McMahon, T. A. and Kronauer, R. E.: Tree structures: deducing the principle of mechanical design, J. Theor. Biol., 59(2), 443–466, 1976.

Mitchard, E. T., Saatchi, S. S., Baccini, A., Asner, G. P., Goetz, S. J., Harris, N. L. and Brown, S.: Uncertainty in the spatial distribution of tropical forest biomass: a comparison of pan-tropical maps, Carbon Balance Manag., 8(1), 10, doi:10.1186/1750-0680-8-10, 2013.

Moorby, J. and Wareing, P. F.: Ageing in Woody Plants, Ann. Bot., 27(2), 291–308, 1963.

Moundounga Mavouroulou, Q., Ngomanda, A., Engone Obiang, N. L., Lebamba, J., Gomat, H., Mankou, G. S., Loumeto, J., Midoko Iponga, D., Kossi Ditsouga, F., Zinga Koumba, R., Botsika Bobé, K. H., Lépengué, N., Mbatchi, B. and Picard, N.: How to improve allometric equations to estimate forest biomass stocks? Some hints from a central African forest, Can. J. For. Res., 44(7), 685–691, doi:10.1139/cjfr-2013-0520, 2014.

Muggeo, V. M. R.: Estimating regression models with unknown break-points, Stat. Med., 22(19), 3055–3071, doi:10.1002/sim.1545, 2003.

Ngomanda, A., Engone Obiang, N. L., Lebamba, J., Moundounga Mavouroulou, Q., Gomat, H., Mankou, G. S., Loumeto, J., Midoko Iponga, D., Kossi Ditsouga, F., Zinga Koumba, R., Botsika Bobé, K. H., Mikala Okouyi, C., Nyangadouma, R., Lépengué, N., Mbatchi, B. and Picard, N.: Site-specific versus pantropical allometric equations: Which option to estimate the biomass of a moist central African forest?, For. Ecol. Manag., 312, 1–9, doi:10.1016/j.foreco.2013.10.029, 2014.

Niklas, K. J.: Size-dependent Allometry of Tree Height, Diameter and Trunk-taper, Ann. Bot., 75(3), 217–227, doi:10.1006/anbo.1995.1015, 1995.

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

O’Brien, S. T., Hubbell, S. P., Spiro, P., Condit, R. and Foster, R. B.: Diameter, Height, Crown, and Age Relationship in Eight Neotropical Tree Species, Ecology, 76(6), 1926–1939, doi:10.2307/1940724, 1995.

Pelletier, J., Ramankutty, N. and Potvin, C.: Diagnosing the uncertainty and detectability of emission reductions for REDD + under current capabilities: an example for Panama, Environ. Res. Lett., 6(2), 024005, doi:10.1088/1748-9326/6/2/024005, 2011.

Perry, D. A.: The competition process in forest stands, Attrib. Trees Crop Plants, 481–506, 1985.

Picard, N., Bosela, F. B. and Rossi, V.: Reducing the error in biomass estimates strongly depends on model selection, Ann. For. Sci., 72(6), 811–923, doi:10.1007/s13595-014-0434-9, 2014.

Picard, N., Rutishauser, E., Ploton, P., Ngomanda, A. and Henry, M.: Should tree biomass allometry be restricted to power models?, For. Ecol. Manag., 353, 156–163, doi:10.1016/j.foreco.2015.05.035, 2015.

Poorter, L., Bongers, F., Sterck, F. J. and Wöll, H.: Architecture of 53 rain forest tree species differing in adult stature and shade tolerance, Ecology, 84(3), 602–608, doi:10.1890/0012-9658(2003)084[0602:AORFTS]2.0.CO;2, 2003.

Poorter, L., Bongers, L. and Bongers, F.: Architecture of 54 moist-forest tree species: traits, trade-offs, and functional groups, Ecology, 87(5), 1289–1301, doi:10.1890/0012-9658(2006)87[1289:AOMTST]2.0.CO;2, 2006.

R Core Team: R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, available at: http://www.R-project.org/ (last access: January 2013), 2012.

Réjou-Méchain, M., Muller-Landau, H. C., Detto, M., Thomas, S. C., Le Toan, T., Saatchi, S. S., Barreto-Silva, J. S., Bourg, N. A., Bunyavejchewin, S., Butt, N., Brockelman, W. Y., Cao, M., Cárdenas, D., Chiang, J.-M., Chuyong, G. B., Clay, K., Condit, R., Dattaraja, H. S., Davies, S. J., Duque, A., Esufali, S., Ewango, C., Fernando, R. H. S., Fletcher, C. D., Gunatilleke, I. A. U. N., Hao, Z., Harms, K. E., Hart, T. B., Hérault, B., Howe, R. W., Hubbell, S. P., Johnson, D. J., Kenfack, D., Larson, A. J., Lin, L., Lin, Y., Lutz, J. A., Makana, J.-R., Malhi, Y., Marthews, T. R., McEwan, R. W., McMahon, S. M., McShea, W. J., Muscarella, R., Nathalang, A., Noor, N. S. M., Nytch, C. J., Oliveira, A. A., Phillips, R. P., Pongpattananurak, N., Punchi-Manage, R., Salim, R., Schurman, J., Sukumar, R., Suresh, H. S., Suwanvecho, U., Thomas, D. W., Thompson, J., Uríarte, M., Valencia, R., Vicentini, A., Wolf, A. T., Yap, S., Yuan, Z., Zartman, C. E., Zimmerman, J. K. and Chave, J.: Local spatial structure of forest biomass and its consequences for remote sensing of carbon stocks, Biogeosciences, 11(23), 6827–6840, doi:10.5194/bg-11-6827-2014, 2014.

Réjou-Méchain, M., Tymen, B., Blanc, L., Fauset, S., Feldpausch, T. R., Monteagudo, A., Phillips, O. L., Richard, H. and Chave, J.: Using repeated small-footprint LiDAR acquisitions to infer spatial and temporal variations of a high-biomass Neotropical forest, Remote Sens. Environ., 169, 93–101, 2015.

Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L. and Hagen, S.: Benchmark map of forest carbon stocks in tropical regions across three continents, Proc. Natl. Acad. Sci., 108(24), 9899–9904, 2011.

Scrucca, L.: Model-based SIR for dimension reduction, Comput. Stat. Data Anal., 55(11), 3010–3026, 2011.

Shinozaki, K., Yoda, K., Hozumi, K. and Kira, T.: A quantitative analysis of plant form-the pipe model theory: I. Basic analyses, 日本生態学会誌, 14(3), 97–105, 1964.

Sileshi, G. W.: A critical review of forest biomass estimation models, common mistakes and corrective measures, For. Ecol. Manag., 329, 237–254, doi:10.1016/j.foreco.2014.06.026, 2014.

Sillett, S. C., Van Pelt, R., Koch, G. W., Ambrose, A. R., Carroll, A. L., Antoine, M. E. and Mifsud, B. M.: Increasing wood production through old age in tall trees, For. Ecol. Manag., 259(5), 976–994, doi:10.1016/j.foreco.2009.12.003, 2010.

Sist, P., Mazzei, L., Blanc, L. and Rutishauser, E.: Large trees as key elements of carbon storage and dynamics after selective logging in the Eastern Amazon, For. Ecol. Manag., 318, 103–109, doi:10.1016/j.foreco.2014.01.005, 2014.

Slik, J. W., Paoli, G., McGuire, K., Amaral, I., Barroso, J., Bastian, M., Blanc, L., Bongers, F., Boundja, P. and Clark, C.: Large trees drive forest aboveground biomass variation in moist lowland forests across the tropics, Glob. Ecol. Biogeogr., 22(12), 1261–1271, 2013.

Stephenson, N. L., Das, A. J., Condit, R., Russo, S. E., Baker, P. J., Beckman, N. G., Coomes, D. A., Lines, E. R., Morris, W. K., Rüger, N., Álvarez, E., Blundo, C., Bunyavejchewin, S., Chuyong, G., Davies, S. J., Duque, Á., Ewango, C. N., Flores, O., Franklin, J. F., Grau, H. R., Hao, Z., Harmon, M. E., Hubbell, S. P., Kenfack, D., Lin, Y., Makana, J.-R., Malizia, A., Malizia, L. R., Pabst, R. J., Pongpattananurak, N., Su, S.-H., Sun, I.-F., Tan, S., Thomas, D., van Mantgem, P. J., Wang, X., Wiser, S. K. and Zavala, M. A.: Rate of tree carbon accumulation increases continuously with tree size, Nature, advance online publication, doi:10.1038/nature12914, 2014.

Van Gelder, H. A., Poorter, L. and Sterck, F. J.: Wood mechanics, allometry, and life-history variation in a tropical rain forest tree community, New Phytol., 171(2), 367–378, doi:10.1111/j.1469-8137.2006.01757.x, 2006.

Vieilledent, G., Vaudry, R., Andriamanohisoa, S. F. D., Rakotonarivo, O. S., Randrianasolo, H. Z., Razafindrabe, H. N., Rakotoarivony, C. B., Ebeling, J. and Rasamoelina, M.: A universal approach to estimate biomass and carbon stock in tropical forests using generic allometric models, Ecol. Appl., 22(2), 572–583, doi:10.1890/11-0039.1, 2011.

West, G. B., Brown, J. H. and Enquist, B. J.: A general model for the structure and allometry of plant vascular systems, Nature, 400(6745), 664–667, doi:10.1038/23251, 1999.

Zanne, A. E., Lopez-Gonzalez, G., Coomes, D. A., Ilic, J., Jansen, S., Lewis, S. L., Miller, R. B., Swenson, N. G., Wiemann, M. C. and Chave, J.: Data from: towards a worldwide wood economics spectrum. Dryad Digital Reposit., 2009.

Table 1. Crown mass sub-models. Model variables are *Cm* (crown mass, kg), *D* (diameter at breast height, cm), *Hc* (crown depth, m), *Cs* (average of *Hc* and crown diameter, m) and ρ (wood density, g.cm^{-3}). The general form of the models is ln(Y) = a+ b* ln(X) + c*ln(X)². Model coefficient estimates are provided along with the associated standard error denoted SE$_i$, with i as the coefficient. Coefficients' probability value (pv) is coded as follows: pv $\leq 10^{-4}$: '***', pv $\leq 10^{-3}$: '**', pv $\leq 10^{-2}$: '*', pv $\leq$ 0.05 : '.' and pv $\geq$ 0.05 : 'ns'. Models' performance parameters are R² (adjusted R square), RSE (residual standard error), *S* (median of unsigned relative individual errors, in %), AIC (Akaike Information Criterion), dF (degree of freedom).

model	**Dataset**	**Model input**			**Model parameters**						**Model performance**				
		Y	**X**	**X²**	**a**	**b**	**c**	**SE$_a$**	**SE$_b$**	**SE$_c$**	**R²**	**RSE**	***S***	**AIC**	**dF**
sm$_1$	Data$_{CM2}$ (n=541)	*Cm*	$D^2*Hc*\rho$	-	-2.6345***	0.9368***		0.1145	0.0125		0.91	0.615	36.0	1012.6	539
sm$_3$			$D^2*Hc*\rho$	$(D^2*Hc*\rho)^2$	0.9017.	0.1143ns	0.0452***	0.5049	0.1153	0.0063	0.92	0.588	35.2	965.2	538
			-	$(D^2*Hc*\rho)^2$	1.3990***		0.0514***	0.0605		0.0007	0.92	0.588	35.5	964.2	539
sm$_1$	Data$_{CD}$ (n=119)	*Cm*	$D^2*Hc*\rho$	-	-2.9115***	0.9843***		0.3139	0.0289		0.91	0.516	31.8	184.1	117
sm$_2$			$D^2*Cs*\rho$	-	-3.0716***	0.9958***		0.2514	0.0231		0.94	0.414	21.8	131.9	117
sm$_3$			$D^2*Hc*\rho$	$(D^2*Hc*\rho)^2$	-0.2682ns	0.4272 ns	0.0283.	1.4077	0.2908	0.0147	0.91	0.510	29.7	182.3	116
			-	$(D^2*Hc*\rho)^2$	1.7830***		0.0498***	0.1774		0.0015	0.91	0.512	32.2	182.5	117
sm$_4$			$D^2*Cs*\rho$	$(D^2*Cs*\rho)^2$	-0.5265ns	0.4617.	0.0270*	1.1443	0.2356	0.0119	0.94	0.407	128.7	25.9	116
			-	$(D^2*Cs*\rho)^2$	1.6994***		0.0502***	0.1421		0.0012	0.94	0.412	130.5	25.8	117

Table 2. Models used to estimate tree *AGB*. Model parameters are D (diameter at breast height, cm), H (total height, m), Ht (trunk height, m), Hc (crown depth, m), Cm (crown mass, kg), Cs (average of Hc and crown diameter, m) and ρ (wood density, g.cm^{-3}). The general form of the models is $\ln(Y) = a + b^{*} \ln(X_1) + c^{*}\ln(X_2)$. Model coefficient estimates are provided along with the associated standard error denoted SE_i, with i as the coefficient. Coefficients' probability value (pv) is coded as follows: $pv \leq 10^{-4}$: '***', $pv \leq 10^{-3}$: '**', $pv \leq 10^{-2}$: '*', $pv \leq 0.05$: '.' and $pv \geq 0.05$: 'ns'. Models' performance parameters are R² (adjusted R square), RSE (residual standard error), *S* (median of unsigned relative individual errors, in %), AIC (Akaike Information Criterion), dF (degree of freedom).

model	**Dataset**	**Model input**			**Model parameters**						**Model performance**				
		Y	**X_1**	**X_2**	**a**	**b**	**c**	**SE_a**	**SE_b**	**SE_c**	**R^2**	**RSE**	***S***	**AIC**	**dF**
m_0	$Data_{REF}$ (n=4004)	*AGB*	$D^2{*}H{*}\rho$		-2.7628***	0.9759***		0.0211	0.0026		0.97	0.358	22.1	3130.7	4002
m_0	$Data_{CM2}$ (n=541)	*AGB*	$D^2{*}H{*}\rho$		-2.5860***	0.9603***		0.0659	0.0066		0.98	0.314	18.9	284.8	539
m_1			$D^2{*}Ht{*}\rho$	*Cm*	-0.5619***	0.5049***	0.4816***	0.0469	0.0098	0.0096	0.99	0.199	9.8	-205.7	538
m_2			$D^2{*}Ht{*}\rho$	$(D^2{*}Hc{*}\rho)^2$	0.3757***	0.4451***	0.0281***	0.0974	0.0186	0.0010	0.98	0.298	17.8	231.5	538
m_0	$Data_{CD}$ (n=119)	*AGB*	$D^2{*}H{*}\rho$		-3.1105***	1.0119***		0.1866	0.0160		0.97	0.268	15.0	28.1	117
m_1			$D^2{*}Ht{*}\rho$	*Cm*	-0.5851***	0.4784***	0.5172***	0.1117	0.0203	0.0185	0.99	0.142	7.0	-121.2	116
m_2			$D^2{*}Ht{*}\rho$	$(D^2{*}Hc{*}\rho)^2$	-0.2853ns	0.5804***	0.0216***	0.2499	0.0397	0.0019	0.97	0.272	14.5	32.5	116
m_3			$D^2{*}Ht{*}\rho$	$(D^2{*}Cs{*}\rho)^2$	0.5800*	0.4263***	0.0283***	0.2662	0.0444	0.0021	0.98	0.246	12.3	9.3	116

Table A1. Preliminary crown mass sub-models. Model parameters are *D* (diameter at breast height, cm), *Hc* (crown depth, m), *Cm* (crown mass, kg), *Cd* (crown diameter, in m), *Cs* (average of *Hc* and *Cd*, m) and ρ (wood density, g.cm^{-3}). The general form of the models is $\ln(Y) = a + b*\ln(X_1) + c*\ln(X_2)$. Model coefficients' estimates are provided along with the associated standard error denoted SE_i, with i as the coefficient. Coefficients' probability value (pv) is coded as follows: $pv \leq 10^{-4}$: '***', $pv \leq 10^{-3}$: '**', $pv \leq 10^{-2}$: '*', $pv \leq 0.05$: '.' and $pv \geq 0.05$: 'ns'. Models' performance parameters are R² (adjusted R square), RSE (residual standard error), *S* (median of unsigned relative individual errors, in %), AIC (Akaike Information Criterion), dF (degree of freedom).

model	**Dataset**	**Model input**			**Model parameters**						**Model performance**				
		Y	**X_1**	**X_2**	**a**	**b**	**C**	**SE_a**	**SE_b**	**SE_c**	**R²**	**RSE**	**S**	**AIC**	**dF**
1-D	Data$_{CM2}$ (n=541)	*Cm*	*D*		-3.6163***	2,5786***		0.1514	0.0409		0.88	0.719	42.8	1181.6	539
1-Hc			*Hc*		-0.1711ns	2.6387***		0.1574	0.0673		0.74	1.060	82.2	1602.8	539
2-D			*D*	*ρ*	-3.0876***	2.6048***	1.1202***	0.1462	0.0372	0.1048	0.90	0.653	36.7	1079.4	538
2-Hc			*Hc*	*ρ*	-0.3952*	2.6574***	-0.3274.	0.1959	0.0679	0.1712	0.74	1.058	80.6	1601.1	538
3			*D²*Hc*ρ*		-2.6345***	0.9368***		0.1145	0.0125		0.91	0.615	36.0	1012.6	539
1-D	Data$_{CD}$ (n=119)	*Cm*	*D*		-3.4603***	2.5684***		0.4692	0.1075		0.83	0.702	39.8	257.4	117
1-Hc			*Hc*		1.3923*	2.2907***		0.5392	0.1938		0.54	1.149	77.4	374.7	117
1-Cd			*Cd*		-0.1181ns	2.8298***		0.3403	0.1218		0.82	0.718	52.7	262.8	117
2-D			*D*	*ρ*	-2.7296***	2.6293***	1.5243***	0.3528	0.0793	0.1523	0.91	0.516	30.5	185.3	116
2-Hc			*Hc*	*ρ*	1.1181ns	2.3356***	-0.2326ns	0.6869	0.2063	0.3596	0.54	1.152	82.9	376.3	116
2-Cd			*Cd*	*ρ*	0.4677ns	2.7954***	0.7538***	0.3585	0.1158	0.2009	0.84	0.681	44.5	251.2	116

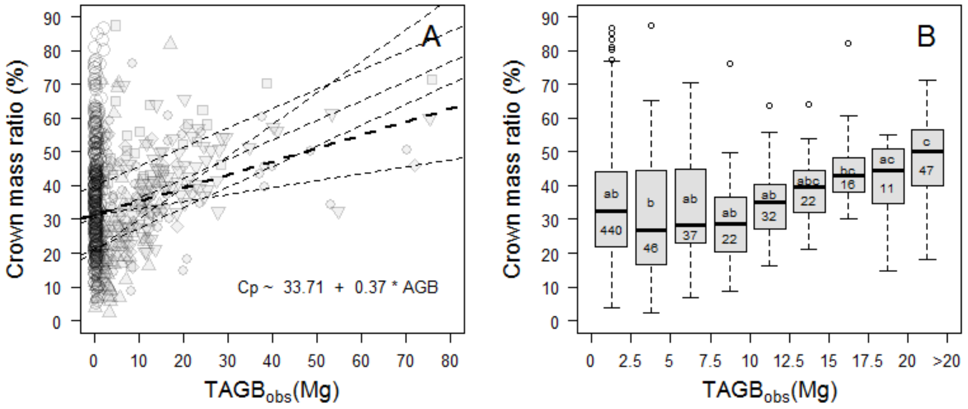

Figure 1. (A) Distribution of crown mass ratio (in %) along the range of tree mass ($TAGB_{obs}$, in Mg) for 673 trees. Dashed lines represent the fit of robust regressions (model II linear regression fitted using ordinary least square) performed on the full crown mass dataset (thick line; one-tailed permutation test on slope: p-value < 0.001) and on each separate source (thin lines), with symbols indicating the source: empty circles from Vieilledent et al. (2011; regression line not represented since the largest tree is 3.7 Mg only); solid circles from Fayolle et al. (2013); squares from Goodman et al. (2013, 2014); diamonds from Henry et al. (2010); head-up triangles from Ngomanda et al. (2014); and head-down triangles from the unpublished data set from Cameroon. (B) Boxplot representing the variation in crown mass ratio (in %) across tree mass bins of equal width (2.5 Mg). The last bin contains all trees ≥ 20 Mg. The number of individuals per bin and the results of non-parametric pairwise comparisons are represented below and above the median lines, respectively.

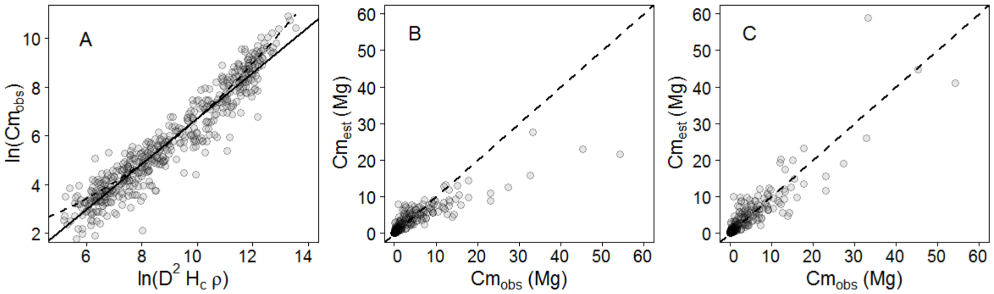

Figure 2. (A) Observed crown mass versus the compound variable $D^2*Hc*\rho$ (in log scale), displaying a slightly concave relationship. The crown mass sub-model 1 does not capture this effect (model fit represented with a full line in caption A), resulting in biased model predictions (caption B), whereas sub-model 3 does not present this error pattern (model fit represented as a dashed line in caption A, observed crown mass against model predictions in caption C). Models were fitted on Data_{CM2}.

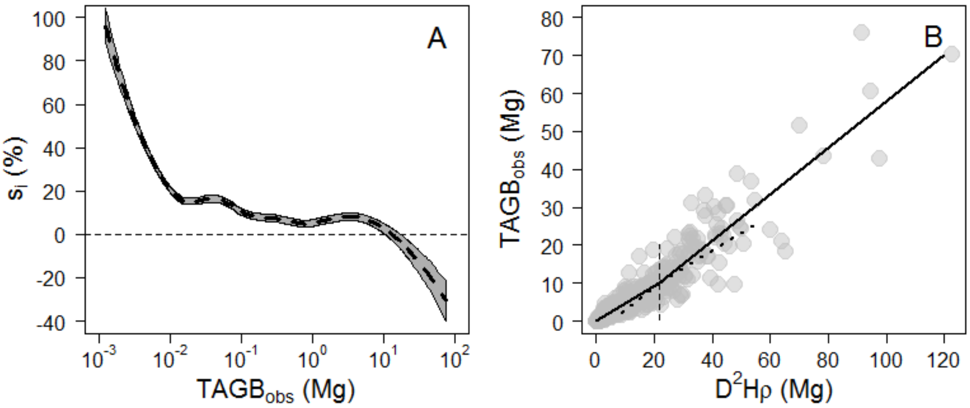

Figure 3. (A) Relative individual residuals (s_i in %) of the reference pantropical model of Chave et.al. (2014) against the tree *AGB* gradient. The thick dashed line represents the fit of a local regression (loess function, span = 0.5) bounded by standard errors. (B) Observed tree *AGB* ($TAGB_{obs}$) versus the compound variable $D^2*H*\rho$ with *D* and *H* being the tree stem diameter and height, respectively, and ρ the wood density. A segmented regression revealed a significant break point (thin vertical dashed line) at approximately 10 Mg of $TAGB_{obs}$ (Davies test p-value < 2.2e-16).

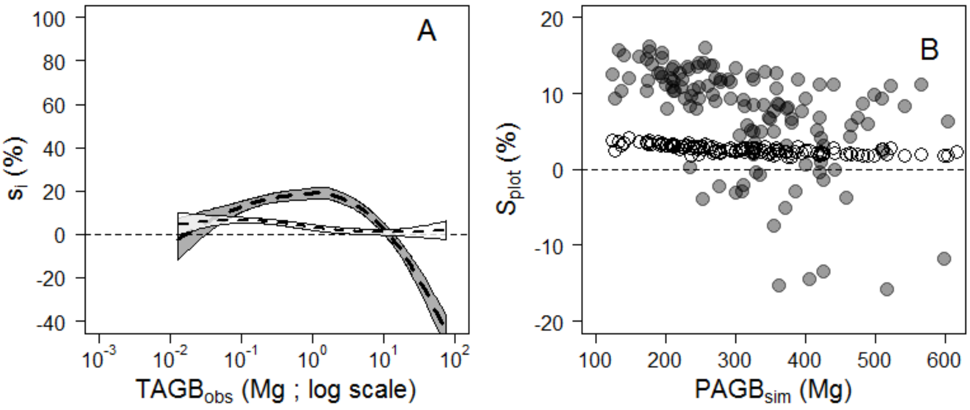

Figure 4. (A) Relative residuals (s_i, in %) of the reference pantropical model m_0 (grey background) and our model m_1 including crown mass (white background). Thick dashed lines represent fits of local regressions (loess function, span = 1) bounded by standard errors. (B) Propagation of individual estimation errors of m_0 (solid grey circles) and m_1 (empty circles) to the plot level.

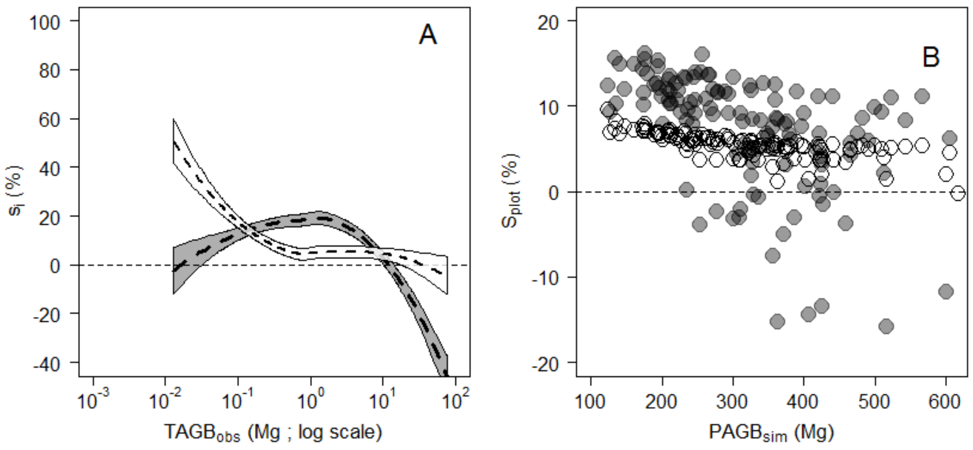

Figure 5. (A) Relative individual residuals (s_i, in %) obtained with the reference pantropical model m_0 (grey background) and with our model including a crown mass proxy, m_2 (white background). Thick dashed lines represent fits of local regressions (loess function, span = 1) bounded by standard errors. (B) Propagation of individual residual errors of m_0 (solid grey circles) and m_2 (empty circles) to the plot level.

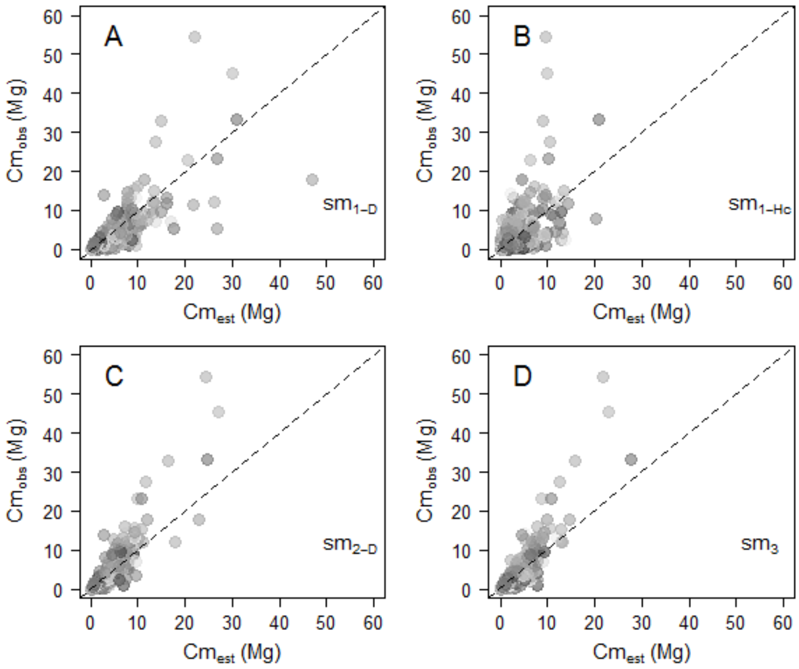

Figure A1. Observed against estimated crown mass (in Mg) for models 1-D (caption A), 1-Hc (caption B), 2-D (caption C), 3 (caption D). Models were calibrated on $Data_{CM2}$. Tree wood density was standardized to range between 0 and 1 and represented as a grayscale (with black the lowest values and white the highest values).

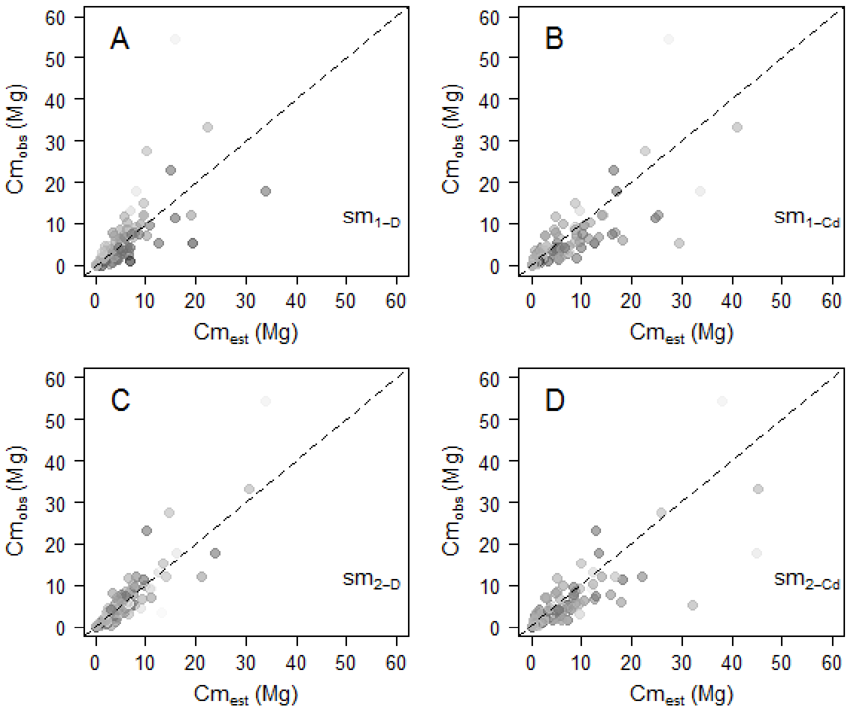

Figure A2. Observed versus estimated crown mass (in Mg) for models 1-D (caption A), 1-Cd (caption B), 2-D (caption C), 2-Cd (caption D). Models were calibrated on $Data_{CD}$. Tree wood density was standardized to range between 0 and 1 and is represented as a grayscale (with black the lowest values and white the highest values).

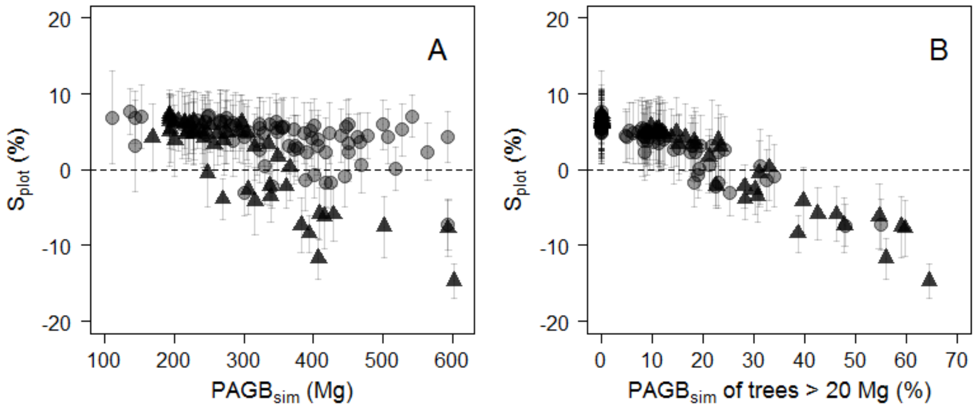

Figure B1. Plot-level propagation of individual-level model error. (A) Mean relative error (S_{plot}, in %) and standard deviation of 1000 random error sampling against simulated plot *AGB* and (B) against the fraction (%) of simulated plot *AGB* accounted for by trees > 20 Mg. Plots from Korup permanent plot are represented by triangles.

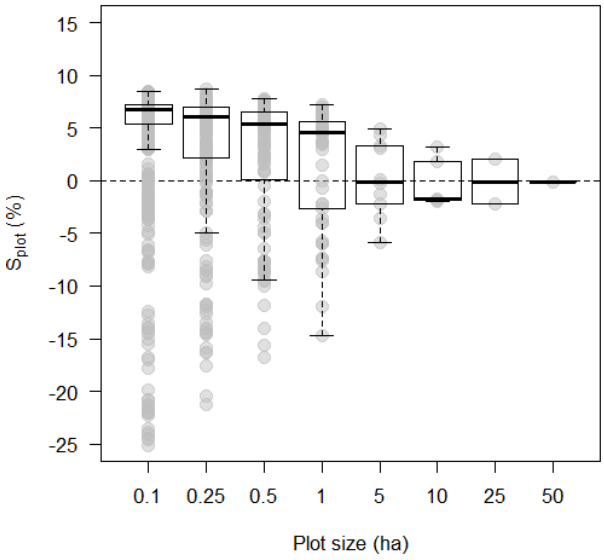

Figure B2. Plot-level relative error (S_{plot}, in %) as a function of plot size (in ha) in Korup permanent plot. Individual plot values are represented by grey dots.