Allometric options for predicting tropical tree height and crown area from stem diameter

Moses Libalah Bakonck, Adeline Fayolle, Nicolas Barbier, Nicolas Picard, Stephane Momo Takoudjou, Yannick Enock Bocko, Gislain Ii Mofack, John Mukirania Katembo, Joel Loumeto, Olga Diane Yongo, et al.

Allometric options for predicting tropical tree height and crown area from stem diameter

Moses B. Libalah (libalah_moses@yahoo.com)
University of Yaoundé I
Adeline Fayolle
University of Liège, Gembloux Agro-Bio Tech
Nicolas Barbier
UMR AMAP-IRD
Nicolas Picard
GIP Ecofor
Stéphane Momo
Plant Systematics and Ecology Laboratory (LaBosystE), Higher Teacher's Training College
Yannick Bocko
Marien Ngouabi University
Gislain II Mofack
University of Yaoundé I
John Mukirania Katembo
Institut Supérieur d'Etudes Agronomiques de Bengamisa
Joël Loumeto
Marien Ngouabi University
Olga Diane Yongo
University of Bangui
Alfred Ngomanda
Institut de Recherche en Ecologie Tropicale
Pierre Couteron
UMR AMAP-IRD
Bonaventure Sonké
Plant Systematics and Ecology Laboratory (LaBosystE), Higher Teacher's Training College
Rossi Vivien
CIRAD - UR Forêts et Sociétés

Article

Keywords:

Posted Date: November 8th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2209593/v1

Abstract

Tree height and crown area are important predictors of aboveground biomass but difficult to measure on the ground. Numerous allometric models have been established to predict tree height from diameter (H–D) and crown area from diameter (CA–D). A major challenge is to select the most precise and accurate allometric model among existing ones, depending on the species composition and forest type where the model is to be applied.

To propose a principle to select tree H–D and tree CA–D allometric models, we build a method based on *k*-fold cross-validation using a large dataset spanning six forest types from central Africa. We then compared the errors and biases using 22 previously established H–D and CA–D allometric model forms via three inter-comparable scenarios: locally derived for the forest type *vs.* regional *vs.* pantropical; regional (encompassing the forest type) *vs.* pantropical; regional (not encompassing the forest type) *vs.* pantropical model.

H–D allometries were more variable across forest types in central Africa than CA–D allometries: (i) forest type explained 6% of the variance in H–D allometry and 2% of the variance in CA–D allometry, while species explained 9% and 2% of the variance in H–D allometry and CA–D allometry, respectively; (ii) for H–D allometry, the six forest types resulted in five best-fit models whereas, for CA–D allometry, four models provided the best fit for the six forest types. We recommend using allometric models specific to the forest type, preferentially to regional ones. Regional models should in turn be preferred to pantropical allometric models.

Introduction

African tropical forests play a key role in the global carbon cycle and climate change mitigation. Emerging evidence indicates that the carbon sink in live aboveground biomass (AGB) has been stable for over three decades since 1985 [1] despite the environmental changes linked to increasing greenhouse gas concentrations [2]. Although recognized as a single biome, African tropical rainforests differ substantially within and across regions. This heterogeneity also implies variability of the response of forests to climate change. Pressure from human activities [3], deforestation [4] and degradation [5] also interplay with forest type. Therefore, efforts to implement effective conservation measures for climate change regulation and monitoring carbon credits over time must rely on accurate predictions of carbon stocks in the different forest types.

To reduce errors in the prediction of tropical tree AGB from tree diameter at breast height (D) and species wood density (ρ), it is strongly recommended to include an estimate of tree height [6,7] and tree crown dimension [8,9]. Consequently, tree heights have been predicted using African regional models that accounted for environmental factors [10], or pantropical models that accounted either for biogeographical differences between continents [11] or for climate [12]. These models encompassed different forest types. However, Africa was underrepresented in the dataset used to fit pantropical models [10,11]. Unveiling the importance of this sampling bias, Kearsley et al.[13] showed that pantropical height–diameter allometric models led to significant overestimation of AGB in central Congo Basin forest and were outperformed by site-specific models [14]. Studies focusing on few field sites in central Africa revealed contrasted performance of height–diameter models between sites [15], insinuating that tree height–diameter allometry may depend on the forest type [16] and species composition.

One historical challenge to predict tree height from stem diameter has been the choice of the allometric functional form. This challenge has been exacerbated by a plethora of proposed functional forms [17,18]. However, there is now apparent convergence towards a few functional forms which are equally supported by ecological theories. For instance, the power law function [19,20] is backed by the theory of simple allometry. The metabolic scaling theory further predicts a specific value for the coefficient of the power model: tree height scales with its stem diameter as $D^{2/3}$ [21–23]. Also, the second order polynomial on the log-log scale [12] is designed to mimic a saturation of tree height with diameter. Asymptotic models such as the Weibull [24,25] and the Michaelis-Menten [26] models are easy to interpret because maximum height is one of the coefficients of the model [10,24]. The other coefficient of the Michaelis-Menten model is also easy-to-manipulate [27]. All these functions have been used to compare height–diameter allometries between forest types in central Africa [14,28,29]. The power model often appeared to be the most pertinent among several mathematical functions in predicting tree height [30,31] and biomass [20,32].

Contrary to tree H–D allometry, the relationship between stem diameter and crown diameter has so far been overlooked in tropical rainforests (but see Loubota Panzou et al. [33]b). Yet, it is likely that large-stature species that realize faster height growth to reach crown exposure have relatively slender stems, and consequently narrow crowns to compensate for wind loads [34,35]. It is also possible that short trees develop large crowns under well illuminated and open conditions [36–39]. Another strategy can be to have a thicker stem made of lighter wood to ensure a tradeoff between the mechanical properties of the wood and the amount of matter needed to ensure a given resistance to buckling or stem damage [40], while still ensuring other functions (water transport and storage, etc.). Thus, tree height–crown allometry may change between species depending on the position of trees in the illumination gradient [41–43]. Also, variation in tree height and crown may reflect the complex multi-layered structure of tropical forests. Therefore, incorporating crown dimension in tree height predictions may improve the precision of these predictions, and consequently AGB estimation [8,42].

Consider a user with some forest inventory data that include tree diameter and wants to predict tree height or crown area (typically to predict biomass in a second step). This user has to make a choice in an ensemble of available allometric models. The goal of this study is to propose a rigorous framework to guide the choice of the height–diameter (H–D) and crown area–diameter (CA–D) allometric model depending on whether the data used to fit this model partly come or not from a forest of the same type as the user's forest. When the calibration data used to fit a model partly come from a forest of the same type as the subject forest, we shall say that the model encompasses the subject forest type. For each allometric model, D was the predictor while H (for H–D allometry) and CA (for CA–D allometry) were the dependent variables. We hypothesized that: (i) H–D and CA–D allometries differ between forest types and that these differences are due to species composition; (ii) allometric models specific to a forest type are more accurate than regional and pantropical allometric models; (iii) regional allometric models encompassing the subject forest type are more accurate than pantropical allometric models; (iv) regional

allometric models not encompassing the subject forest type are less accurate than pantropical allometric model. Regional models presumably incorporate the architectural properties of trees better than pantropical models. However, compared to regional models that do not encompass the subject forest type, pantropical allometries may have the advantage to capture a wider range of forest properties and architecture. To address these hypotheses, we compared allometric models derived for each forest type with published regional and pantropical allometric models through three scenarios: allometries derived for the forest types against regional and pantropical allometric models (scenario 1); regional allometric model encompassing the subject forest type against pantropical allometric model (scenario 2); regional allometric model not encompassing the subject forest type against pantropical allometric model (scenario 3). Finally, we discussed the choice among specific-to-the-forest-type, regional or pantropical models depending on whether the models encompass the subject forest type or not.

Results

Effect of forest type and species composition on allometries. Respectively for the H–D and CA–D allometries, the regional model with the lowest BIC was the log-normal model of the form: $\log(\mathcal{X}) = a_f + a_s + b \times \log(D) + c \times \log(D)^2$, and the linearized-power model of the form: $\log(\mathcal{X}) = a_f + a_s\, b \times \log(D)$ where $\mathcal{X}$ was tree height or crown area, a_f was a fixed effect depending on forest type f, a_s was a random effect depending on species s, and b and c were fixed coefficients. The conditional R^2 (fixed and random effects) was 80% for the H–D model and 72% for the CA–D model.

For the H–D model, the part of variance explained by tree diameter (that is $\log(D) + \log(D)^2$) was 66%, by forest type 6%, and by species composition (random-effect) 9%. When the forest type was not included in the H–D model, the proportion of variance explained by the species composition was 14%. Hence, the part of variance explained by the forest type was taken from the variance explained by the species composition. Post-hoc analysis showed three significant groups of forest types: TSE-DRC and SFM-CON; TSS-CAR and EFL-RGE; EVF-GAB and SEF-CAM (Fig. 2A).

For the CA–D model, the part of variance explained by tree diameter was 65%, by forest type 2%, and by species composition (random-effect) 2%. When the forest type was not included in the CA–D model, the proportion of variance explained by the species composition was still 5%. Hence, the variance explained by the forest type was independent of the variance explained by the species composition. Post-hoc analysis showed two significant groups of forest types: EVF-GAB, SFM-CON and TSE-DRC; TSE-DRC, TSS-CAR, SEF-CAM and EFL-RGE (Fig. 2B).

Hence, the hypothesis that H–D and CA–D allometries differ between forest types is verified. Moreover, for the H–D allometry, differences between forest types are due to species composition.

Performance of H–D allometric models. *Scenario 1*. (models derived for the forest type *vs.* regional models *vs.* pantropical models): For three out of the six forest types (*viz*. EVF-GAB, SEF-CAM and TSE-DRC), the locally fitted allometric model provided the best performance according to the RB and RMSE criteria. These three forest-type specific models corresponded to two different mathematical functions, namely the linear function (for EVF-GAB and SEF-CAM) and the power function (for TSE-DRC). The regional model of Feldpausch et al. (2012) for central Africa and the regional model of Banin et al. (2012) for tropical Africa provided the best performance in two other forest types (*viz*. SFM-CON and TSS-CAR, respectively). The pantropical climate-based model of Chave et al. (2014) provided the best performance in the remaining forest type EFL-RGE (Table 3; Fig. 3A-B; Supporting Information Table S3). Except for the central Congo regional model of Kearsley et al. (2013) that had outstanding error of 5 m to 7 m and bias of − 18% to − 19% in tree height predictions (results not shown in Fig. 3A-B), the errors generated by the other allometric models for each forest type were hardly greater than 2 m. These errors ranged from a minimum of 0.64–1.23 m to a maximum of 1.14–1.96 m across forest types (Supporting Information Table S3). Despite small errors, large ranges in bias were recorded by the models. These biases ranged from a minimum of 0.002–0.43% to a maximum of 2.28–4.43% across forest types (Supporting Information Table S3).

Scenario 2. (regional models encompassing the subject forest type *vs.* pantropical models): The regional model fitted using the Meyer function provided the best performance for one out of the five validation folds, while the model of Banin et al. (2012) for tropical Africa provided the best performance for the four other folds (Table 3; Fig. 3C-D). The central Congo regional model of Kearsley et al. (2013) had again outstanding error of 31 m to 34 m and bias of − 92% to − 93% for the five folds (result not shown in Fig. 3C-D). The error in predicted tree height by the other allometric models ranged from a minimum of 5.70–6.57 m to a maximum of 7.29–8.24 m across validation folds (Supporting Information Table S4). With the exception of Kearsley et al.'s model, the bias ranged from a minimum of 0.30–1.97% to a maximum of −(8.68–11.93)% across validation folds (Supporting Information Table S4).

Scenario 3. (regional models not encompassing the subject forest type *vs.* pantropical models): Best performances were obtained with three published regional models (Table 3). The regional model of Feldpausch et al. (2011) for central Africa provided the best performance when the forest type TSE-DRC was used as validation fold. The model of Banin et al. (2012) for tropical Africa provided the best performance when TSS-CAR or SFM-CON were used as validation folds. The model of Lewis et al. (2009) for tropical Africa provided the best performance for the other three validation folds (*viz*. forest types EFL-RGE, EVF-GAB and SEF-CAM). The central Congo regional model of Kearsley et al. (2013) had again outstanding error of 27 m to 37 m and bias of − 92% to − 93% for the six folds (and was not shown in Fig. 3E-F). The errors produced by the other models ranged from a minimum of 4.50–7.08 m to a maximum of 27.24–37.38 m across validation folds (i.e. forest types) (Fig. 3E-F; Supporting Information Table S5). The bias recorded among the forest types ranged from a minimum of 0.07–8.35% to a maximum of −(92.58–94.14)% across validation folds (Supporting Information Table S5).

Performance of CA–D allometric models. *Scenario 1*. Six locally fitted models using four different mathematical functions provided the best performance to predict crown area for the six forest types (Table 3). The linear model had the lowest error and bias for SFM-CON and SEF-CAM. The linearized power model had the lowest error and bias for TSE-DRC. The parabolic-log model had the lowest error and bias for TSS-CAR and EFL-RGE. The Michailoff model had the lowest error and bias for EVF-GAB (Fig. 4A-B; Supporting Information Table S6). The errors and biases in predicted crown area for all allometric models ranged from 11 to 572 m^2 and from 0.29 to 685%, respectively. The ranges of errors and biases for each forest type are reported in Supporting Information Table S6.

Scenario 2. Five fitted regional models using four different mathematical functions provided the best performance to predict crown area for the five validation folds (Table 3). The Naslund model had the lowest error and bias for two validation folds, while the linear model, the parabolic model and the parabolic-log model had the lowest error and bias for one fold each (Fig. 4C-D; Supporting Information Table S7). The errors and biases in predicted crown area ranged from 96 to 259 m^2 and from 0.30 to − 11.93%, respectively. The ranges of errors and biases for each validation fold are reported in Supporting Information Table S7.

Scenario 3. Five fitted regional models using four different mathematical functions and one published regional model provided the best performance to predict crown area for the six validation folds corresponding to the six forest types (Table 3). The log-normal model had the lowest error and bias when the forest types TSE-DRC or SEF-CAM were used as validation folds. The parabolic model, the parabolic-log model, the Korf model and the model by Blanchard et al. (2016) for central Africa had the lowest error and bias when the forest type TSS-CAR, EFL-RGE, SFM-CON and EVF-GAB was used as validation fold, respectively (Fig. 4E-F; Supporting Information Table S8). The errors and biases in predicted crown area ranged from 67 to 282 m^2 and from 0.29 to 685%, respectively. The ranges of errors and biases for each validation fold are presented in Supporting Information Table S8.

Discussion

Analysing the accuracy of various allometric models relating tree diameter to either total tree height or crown area is a major step towards precise prediction of forest structural attributes such as biomass. To a greater extent, such analysis will facilitate the calibration of remote sensing products [58] and inform on biogeographical processes driving forest communities [10,11]. Despite a plethora of published allometric models to predict tree height (H) and crown area (CA) from stem diameter at 1.3 m breast height (D), only a few are commonly used for the H–D relationship [10,11,14,31,59] (e.g. Feldpausch et al. [10,11]; Banin et al. [14]; Ledo et al. [31]; Kearsley et al. [59]) or for the CA–D relationship [33,41,42,33,41,42b]. Sometimes, these models are calibrated on some species [28], with or without using the forest type as a predictor. One could expect that a single allometric model could have the best precision (small RMSE) and accuracy (small relative bias) for different forest types but was not the case for this study. Thus, analysing the RMSE and relative bias of predictions from multiple allometric models can help elucidate the dependence of allometry on forest type. Furthermore, our strategy of analysing both RMSE for precision and relative bias for accuracy may be preferable for assessing the performance of allometric models, because most models may be precise with very small differences in RMSE but with large differences in accuracy, i.e. relative bias.

In this study, we showed that the best model describing the allometric relationship between tree diameter and height (H–D) differed among forest types as demonstrated in Madagascar [60]. Four different mathematical functions underlined the six models provided the lowest RMSE and relative bias for the six forest types in central Africa. The linear function outperformed the other functions in two forest types. Conversely, two asymptotic, one power and one log-normal functions had the lowest RMSE and relative bias for the four other forest types. Such variations of functional forms were also reported by Loubota Panzou et al. [42a,b]. Nevertheless, a close examination of the model performances showed that very small differences in RMSE but large differences in relative bias were recorded between the best performing allometric models and the other less performing models (see Supporting Information Tables 3–8). For instance, low relative bias is crucial in AGB predictions where a bias in height prediction of the largest trees may induce significant errors in AGB plot estimates [61]. These results equally point to the plasticity of the H–D relationship to forest types, which could be controlled by abiotic and/or biotic factors. Indeed, a substantial share of the variance of the linear mixed-effect model was explained by the joint effect of species composition and forest type, thus providing support for biotic and abiotic effects (see hypothesis (i) in the Introduction). Similarly, H–D allometries have been reported to differ at various spatial scales: in tropical lowland forests across continents, allometries are modulated by precipitation seasonality and solar radiation [11]; in geographic regions within a continent, they are modulated by environment and forest structure [31]; among lowland evergreen and transition forests, they are modulated by stem density [28]; and in tropical montane forest, they are modulated by altitude [62].

For H–D allometry, we showed that the mathematical functions describing the allometric relationship between tree diameter and crown area (CA–D) also change between forest types. Four models had lowest RMSE and relative bias across the six forest types, such variability was as also reported at pantropical scale [33]. Our results equally suggested that species composition and forest types may only partially explain the variability of CA–D allometry in conformity with [41,63]Antin et al. [41] and Shenkin et al. [63]. Ecological factors such as convergence in tree architecture, and stochastic processes such as crown packing, may have stronger influence on the relationship between diameter and crown dimension than abiotic factors. For instance, Iida et al. [64] suggested that convergence in architectural differences due to lateral crown extension for light interception may facilitate overlap in crown dimensions across species communities in Dipterocarp forests. However, some effects related to ontogeny or species regeneration guild may exist while being overshadowed by competitive convergence and stochastic processes (but see Shenkin et al. [63]). Therefore, more in depth studies are required to assess the latter hypothesis.

Another point to note from this study is that variation in allometry between the forest types in central Africa is much weaker for CA–D than for H–D. In concordance, [42]Blanchard et al. [42] highlighted that, across biogeographic areas, H–D allometry was characterized by inter-site heterogeneity and intra-site homogeneity, whereas the reverse held for CA–D allometry (i.e. inter-site homogeneity and intra-site heterogeneity). As a consequence, regional crown area–diameter models can be used to predict crown attributes (e.g. crown volume, width, depth, leaf area index, etc.) observed from remote sensing products [58], whereas a general height–diameter allometric model will be less accurate for different forest types than site-specific models. Thus, site-specific height–diameter allometric models incorporating local environmental factors are needed to accurately predict forest structural attributes (e.g. volume, biomass).

Limitations of the methods and correction. Strictly speaking, ordinary least squares (OLS) are not suitable to fit models on untransformed heteroscedastic variables such as tree heights or crown dimensions. Models may then be biased. In addition, OLS with untransformed heteroscedastic data would favour large diameter trees since they have more weight in the RMSE than trees with small diameters. To minimize such biases, we carefully built folds in the cross-validation that preserved the regular distribution of trees across diameter classes that was designed by sampling in the full dataset. As alternative,

untransformed models could have been fitted using weighted least squares (WLS), by making the variance of residuals dependent on diameter. However, WLS would have penalized fitted models because the RMSE criterion used for model comparison would then have differed from the RMSE criterion used for fitting.

Cross-validation allowed us to fairly compare models from the literature to models fitted on the study data since the validation data were independent of the calibration data. Models from the literature had fixed coefficients for all folds, whereas fitted models had different estimated coefficients for each calibration fold. As a consequence, the error induced by the uncertainty on coefficients was taken into account for fitted models, while it was not for published models. To minimize this shortcoming, we averaged RMSE over the *k* validation folds. Moreover, in allometric models, the share of the prediction error due to the uncertainty on coefficients is very small compared to that induced by the choice of the model [65–67].

Recommendation for selecting height and crown area allometries. Locally derived allometric models were preferable over regional and pantropical climate-based allometric models. Indeed, even if locally derived, regional and pantropical climate-based models had equivalent RMSE, locally derived allometric models outperformed the other models in terms of relative bias.

Although our results support overall the use of local allometric models, situations will remain where no local height or crown area data are available at local scales despite efforts to collate such data globally [68]. In the absence of local models but presence of regional models, the recommended option will depend on whether the calibration data used to fit these regional models encompassed or not the subject forest type. If regional models are encompassing the subject forest type, we recommend their use. Pantropical climate-based allometric models are also the default option in the absence of regional models.

Materials And Methods

Study sites and forest types. As a benchmark to assess the three scenarios, we used a large central African regional database collected during the PREREDD + project [44] which equally acquired the relevant institutional, national, and international guidelines and legislation for the studies on plants. Destructive biomass data were collected in logging concessions in Cameroon, Central African Republic, Republic of Congo, Democratic Republic of the Congo (DR Congo), Equatorial Guinea and Gabon with the same standardized protocol, and using appropriate sampling (following the recommendations of Chave et al. [12]). All the six sites are located at low to mid altitude (52–663 m) within the equatorial climate zone where important variation has been recorded for annual mean ranges of precipitation (1,400–2,500 mm) and temperature (20–30°C) [45]. Bedrock materials of these sites are mainly of metamorphic or sedimentary origins. Soils are ferralsols of different types (Table 1).

The six sites show contrasting floristic and functional compositions across climates, soil types and anthropogenic gradients [46], in addition to their differences in deciduousness (evergreen and semi-deciduous), hydrology, geological substrates and dominant tree species (Table 1; Supporting Information Table S1). Hence, each site corresponded to a different forest type with a semi-deciduous forest in Cameroon (SEF-CAM), a transition between semi-deciduous and evergreen forest on sandstone plateau in Central African Republic (TSS-CAR), a seasonally flooded forest including monodominant *Gilbertiodendron dewevrei* (De Wild.) J.Leonard species in the Republic of Congo (SFM-CON), a transition between semi-deciduous and evergreen forest in the DR Congo (TSE-DRC), an evergreen forest of littoral type in Equatorial Guinea (EFL-RGE) and a mainland evergreen forest in Gabon (EVF-GAB) (Table 1; Supporting Information Table S1).

Tree dataset. A total of 845 trees (52 species, 49 genera and 17 families) were sampled covering a wide range of diameter (10–208 cm), height (8–67 m) and crown area (1.43–1,495 m^2). Taxonomic identification of species, stem diameter, maximum height and crown area were measured prior to logging (Table 1; Supporting Information Table S1). Species taxonomic identification was initially made by trained botanists from respective logging concessions and confirmation in the field by our experienced botanists (ML, BS, JL, JMK, AN, ODY). Stem diameter (D, in cm) was measured at 1.3 m breast height or 30 cm away from any deformation using a diameter tape. Maximum tree height was measured using the *sine method* (*sensu* [47]). That is, a Haglöf Transponder T3® was placed on the tree at 1.3 m breast height and an observer carrying a Haglöf Vertex IV® was positioned at a relative horizontal distance from the tree that permitted viewing the topmost canopy of the tree. From this position, the maximum tree height (H in m) (average of two tallest measurements) were measured. Crown dimensions were measured as the horizontal distance from the trunk base to the projected edge of the canopy in the four cardinal directions (N, S, W and E). Whenever one of the cardinal directions could not be measured, the alternative directions of the crown (NE, SW, SE and NW, respectively) were considered. Using the values, the ellipsoidal crown area (CA in m^2) was computed as $CA = \pi \times (NS_{mean} \times EW_{mean})$, where NS_{mean} and EW_{mean} are the average crown radii for the N–S and E–W directions, respectively.

Mathematical functions for H–D and CA–D relationships. We considered 22 existing mathematical functions from the literature and fitted these separately for the H–D and for the CA–D relationships (see Supporting Information Table S2). The 22 mathematical functions possessed either 2 parameters (10 functions) or 3 parameters (12 functions). They consisted of polynomials, asymptotic and power law functions. Most of these mathematical functions are incorporated in the *Imfor R*-package [48]. To ensure consistency with the performance criteria used (see section “Evaluating performance of the allometric models” below), all models were fitted using ordinary least-squares (OLS) regression and assuming Gaussian errors. Some functions were log-transformed prior to calibration (see Supporting Information Table S2) in order to easily use the Bayesian Information Criterion (BIC). In this case, no transformation was achieved but lognormal errors were used. When log-transformation was used for calibration, a bias correction factor was used to back-transform model predictions [49,50]. For nonlinear functions, we first estimated their parameter starting values for the fitting algorithm following Huang et al.[17] and Zeide [51].

Published H–D and CA–D models. In addition to the fitted models, seven published models were taken from the literature: five regional models for H–D allometry, one pantropical model for H–D allometry, and one regional model for CA–D allometry (Table 2).

Statistical analysis. *Allometry differences between forest types and species effects*. To assess differences in the H–D and CA–D allometric models between forest types due to species composition (hypothesis (i) of the Introduction), we fitted a linear mixed-effects model [52]. Stem diameter and forest types were

used in the model as fixed effects while species as a random effect. All the 22 models were fitted using the entire tree data (Supporting Information Table S2). The best regional model was selected among these 22 models using the BIC [53]. *Post-hoc* pairwise significant differences between the forest types were assessed by Tukey HSD via least square means.

Comparing allometric models. To compare the performance of allometric models for predicting tree H and CA from D (hypotheses (ii)–(iv) of the Introduction), we adopted a cross-validation scheme [54,55]. Some models in the comparison were fitted to data while others were taken from the literature with fixed coefficients. In this latter case, the cross-validation scheme boils down to a simple validation scheme. In any case, cross-validation is appropriate to evaluate the predictive power of models as it verifies the behaviour of the models on data not used for their calibration. We performed *k*-fold cross-validation that consisted of: first splitting the dataset into *k* folds; second, using *k* – 1 folds for the calibration of the models to be fitted and one fold for the validation of all models (whether fitted or taken from the literature). The second step was repeated *k* times to ensure that all the folds were used for validation (Fig. 1).

To compare the models in the three scenarios (i.e. scenario 1: models derived for the forest type *vs.* regional models *vs.* pantropical models, scenario 2: regional models encompassing the subject forest type *vs.* pantropical models, and scenario 3: regional models not encompassing the subject forest type *vs.* pantropical models), we used two folding schemes based on the dataset of 845 trees from all forest types. We ensured similarity in diameter structure for each of the schemes and within each fold by assigning trees into the following nine diameter classes [10–20[, [20–30[, [30–40[, [40–50[, [50–60[, [60–70[, [70–100[, [100–150[, > 150 cm. We then built the *k*-fold by randomly splitting each diameter class into *k* folds (Figs. 1 and Supporting Information Figure S1).

For H–D allometries, we compared six models from the literature with the subject forest type models based on 22 mathematical functions and the regional models based on the same 22 mathematical functions. Results were presented only for eight H–D allometric models; i.e. the best model for the subject forest type, the best regional model, and the six models from the literature (Table 2).

For CA–D allometries, we compared one model from the literature with subject forest type models based on 22 mathematical functions and regional models based on the same 22 mathematical functions. Results were presented only for three CA–D allometric models; i.e. the best model for the subject forest type, the best regional model, and the model from the literature (Table 2).

For both H–D and CA–D allometries, literature models were not re-calibrated on the training dataset during the *k* fold comparison.

5-fold cross-validation for scenarios 1 and 2. We randomly split each forest type dataset into 5 folds (Fig. 1). We repeated the following steps five times by: (i) combining 4 folds for calibration (80% of the dataset) and reserving 1 fold for validation (20% of the dataset); (ii) calibrating a local equation for each forest type, for scenario 1; (iii) calibrating a regional equation using all 4 × 6 calibration folds from the 6 forest types combined, for scenario 2. Steps (i)-(iii) were repeated 5 times so that each fold was used for validation (Fig. 1).

6-fold cross-validation for scenario 3. We split the whole tree dataset into 6 folds according to the six forest type (Supporting Information Figure S1). We repeated the following steps six times by: (i) combining 5 folds for calibration and reserving 1 fold for validation; (ii) calibrating a regional model using the 5 calibration folds. Steps (i)-(ii) were repeated 6 times so that each fold (*i.e.* each forest type) was used for the validation (Supporting Information Figure S1). Thus, the forest type of the validation fold is the one excluded from calibration.

Evaluating the performance of the allometric models. For each scenario, we evaluated the performance of the models from the literature, those of the subject forest type, and those of the regional models by averaging, over the *k*-fold iterations, two criteria computed on the validation fold:

$$\text{Relative bias (\%)}: \text{RB} = 100 \times \frac{1}{n}\sum_{i=1}^{n}\frac{\widehat{X}_i - X_i}{X_i}$$

$$\text{Root mean square error (m or m}^2\text{)}: \text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\widehat{X}_i - X_i)^2}$$

where $\widehat{X}_i$ and X_i are respectively, the predicted and observed height (or crown area) and n represents the total number of trees.

For each *k* fold iterations in scenario 1, RB and RMSE of forest type, regional and pantropical models were computed using the validation fold of each forest type separately. RB and RMSE were then averaged over the 5 iterations. For each *k* fold iterations in scenario 2, RB and RMSE of regional and pantropical models were computed on each the validation fold. For each *k* fold iterations in scenario 3, RB and RMSE of regional and pantropical models were computed on each the validation fold which corresponded to a forest type. Note that RMSE is always positive while the RB indicates whether the error was positive or negative. To determine the most suitable allometric model among the locally-derived forest type, regional and pantropical models for H–D and CA–D relationships, we compared their relative error (RMSE) and then relative bias (RB). We first selected the model with the lowest RMSE. Second, we selected other models whose difference in RMSE with the lowest RMSE was less than 1 unit (error of 1 m tall tree or 1 m^2 crown area seems meaningful for forest biometry). Lastly among this selection, we retained the one with the lowest RB as the best model.

All statistical analyses were performed using version 4.1.2 of the R statistical software [56].

Declarations

Data Availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Acknowledgements

This study is a contribution from the regional project on REDD+ (PreREDD+) in the Congo basin that was funded by the Global Environment Fund (TF010038) and was administered by the World Bank to the COMIFAC. The second component of the project aimed at “Building allometric equations for the forests of the Congo basin” was led by the ONFi/TEREA/Nature+ consortium and data collection was carried out by national scientific institutions including INDEFOR-AP in Equatorial Guinea, IRET in Gabon, University of Yaoundé 1 in Cameroon sup- ported by researchers from IRD, University of Bangui in CAR, University of Marien NGouabi in Congo and the University of Kisangani in DRC. Logistic support necessary for the field and laboratory measurements was provided by the six timber companies including COMALI in Equatorial Guinea, Rougier Gabon, GRUMCAM in Cameroon, SEFCA in CAR, CIB-OLAM in Congo, and CFT in DRC. The authors also thank the large number of technicians and students that were involved in the data collection and Sébastien BAUWENS for the data collection protocol, Adrien PEROCHES for data quality controls and Samuel QUEVAUVILLERS for data management.

Author contributions

M.B.L, A.F., V.R. and N.B. designed the research; A.F., N.P., N.B., A.N., S.T.M., P.C.,Y.B., G.M., J.M.K.,J.L., B.S., O.D.Y., V.R. and participated in data collection; M.B.L, and V.R. analysed the data and wrote the first draft, A.F., N.P., N.B., A.N., S.T.M., P.C., and O.D.Y. provided useful feedback.

Competing interests

The authors declare no competing interests.

References

1. Hubau, W. *et al.* Asynchronous carbon sink saturation in African and Amazonian tropical forests. *Nature* **579**, 80–87 (2020).
2. IPCC, I. P. O. C. C. *Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Intergovernmental Panel on Climate Change* (2014).
3. Abernethy, K., Maisels, F. & White, L. J. T. Environmental Issues in Central Africa. *Annu. Rev. Environ. Resour.* **41**, 1–33 (2016).
4. Hansen, M. C. *et al.* High-resolution global maps of 21st-century forest cover change. *Science (80-.).* **342**, 850–853 (2013).
5. Pearson, T. R. H., Brown, S., Murray, L. & Sidman, G. Greenhouse gas emissions from tropical forest degradation: An underestimated source. *Carbon Balance Manag.* **12**, (2017).
6. Chave, J. *et al.* Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia* **145**, 87–99 (2005).
7. Nogueira, E. M., Nelson, B. W., Fearnside, P. M., França, M. B. & Oliveira, Á. C. A. de. Tree height in Brazil’s ‘arc of deforestation’: Shorter trees in south and southwest Amazonia imply lower biomass. *For. Ecol. Manage.* **255**, 2963–2972 (2008).
8. Goodman, R. C., Phillips, O. L. & Baker, T. R. The importance of crown dimensions to improve tropical tree biomass estimates. *Ecol. Appl.* **24**, 680–698 (2014).
9. Ploton, P. *et al.* Closing a gap in tropical forest biomass estimation: Taking crown mass variation into account in pantropical allometries. *Biogeosciences* **13**, 1571–1585 (2016).
10. Feldpausch, T. R. *et al.* Height-diameter allometry of tropical forest trees. *Biogeosciences* **8**, 1081–1106 (2011).
11. Banin, L. *et al.* What controls tropical forest architecture? Testing environmental, structural and floristic drivers. *Glob. Ecol. Biogeogr.* **21**, 1179–1190 (2012).
12. Chave, J. *et al.* Improved allometric models to estimate the aboveground biomass of tropical trees. *Glob. Chang. Biol.* **20**, 3177–3190 (2014).
13. Kearsley, E. *et al.* Conventional tree height-diameter relationships significantly overestimate aboveground carbon stocks in the Central Congo Basin. *Nat. Commun.* **4**, 86–94 (2013).
14. Kearsley, E. *et al.* Model performance of tree height-diameter relationships in central Congo Basin. *Ann. For. Sci.* **74**, 1–13 (2017).
15. Fayolle, A. *et al.* Taller trees, denser stands and greater biomass in semi-deciduous than in evergreen lowland central African forests. *For. Ecol. Manage.* **374**, 42–50 (2016).
16. Ketterings, Q. M., Coe, R., Van Noordwijk, M., Ambagau’, Y. & Palm, C. A. Reducing uncertainty in the use of allometric biomass equations for predicting above-ground tree biomass in mixed secondary forests. *For. Ecol. Manage.* **146**, 199–209 (2001).
17. Huang, S., Titus, S. J. & Wiens, D. P. Comparison of nonlinear height–diameter functions for major Alberta tree species. *Can. J. For. Res.* **22**, 1297–1304 (1992).
18. Mehtätalo, L., Gregoire, T. G. & de Miguel Magaña, S. Modeling height-diameter curves for prediction. *Can. J. For. Res.* **837**, 150409143753006 (2015).
19. King, D. A. Allometry and Life History of Tropical Trees. *J. Trop. Ecol.* **12**, 25–44 (1996).
20. Fang, Z. & Bailey, R. L. Height–diameter models for tropical forests on Hainan Island in southern China. *For. Ecol. Manage.* **110**, 315–327 (1998).
21. West, G. B., Brown, J. H. & Enquist, B. J. A general model for the origin of allometric scaling laws in biology. *Science (80-.).* **276**, 122–126 (1997).
22. West, G. B., Brown, J. H. & Enquist, B. J. A general model for the structure and allometry of plant vascular systems. *Nature* **400**, 664–667 (1999).
23. Muller-Landau, H. C. *et al.* Testing metabolic ecology theory for allometric scaling of tree size, growth and mortality in tropical forests. *Ecol. Lett.* **9**, 575–588 (2006).
24. Bailey, R. The potential of Weibull-type functions as flexible growth curves. *Can. J. For. Res.* **10**, 117–118 (1980).
25. Weibull, W. A Statistical Distribution Function of Wide Applicability. *J. Appl. Mech.* **18**, 293–297 (1951).

26. Stage, A. R. *Prediction of height increment for models of forest growth. Intermountain Forest and Range Experiment Station* **164**, (Forest Service, US Department of Agriculture, 1975).
27. Molto, Q. *et al.* Predicting tree heights for biomass estimates in tropical forests -A test from French Guiana. *Biogeosciences* **11**, 3121–3130 (2014).
28. Fayolle, A. *et al.* Taller trees, denser stands and greater biomass in semi-deciduous than in evergreen lowland central African forests. *For. Ecol. Manage.* **374**, 42–50 (2016).
29. Loubota Panzou, G. J., Bocko, Y. E., Mavoungou, A. Y. & Loumeto, J. J. Height-diameter allometry in African monodominant forest close to mixed forest. *J. Trop. Ecol.* **37**, 98–107 (2021).
30. Djomo, A. N., Ibrahima, A., Saborowski, J. & Gravenhorst, G. Allometric equations for biomass estimations in Cameroon and pan moist tropical equations including biomass data from Africa. *For. Ecol. Manage.* **260**, 1873–1885 (2010).
31. Feldpausch, T. R. *et al.* Tree height integrated into pantropical forest biomass estimates. *Biogeosciences* **9**, 3381–3403 (2012).
32. Fonton, N. H. *et al.* Analyzing Accuracy of the Power Functions for Modeling Aboveground Biomass Prediction in Congo Basin Tropical Forests. *Open J. For.* **07**, 388–402 (2017).
33. Loubota Panzou, G. J. *et al.* Pantropical variability in tree crown allometry. *Glob. Ecol. Biogeogr.* **30**, 459–475 (2021).
34. Kohyama, T., Suzuki, E., Partomihardjo, T., Yamada, T. & Kubo, T. Tree species differentiation in growth, recruitment and allometry in relation to maximum height in a Bornean mixed dipterocarp forest. *J. Ecol.* **91**, 797–806 (2003).
35. Poorter, L., Bongers, F., Sterck, F. J. & Woll, H. Beyond the regeneration phase: differentiation of height-light trajectories among tropical tree species. *J. Ecol.* **93**, 256–267 (2005).
36. O'Brien, S. T., Hubbell, S. P., Spiro, P., Condit, R. & Foster, R. B. Diameter, height, crown, and age relationships in eight neotropical tree species. *Ecology* **76**, 1926–1939 (1995).
37. Sterck, F. J. & Bongers, F. Ontogenetic Changes in Size, Allometry, and Mechanical Design of Tropical Rain Forest Trees. *Am. J. Bot.* **85**, 266–272 (1998).
38. Poorter, L. Growth responses of 15 rain-forest tree species to a light gradient: The relative importance of morphological and physiological traits. *Funct. Ecol.* **13**, 396–410 (1999).
39. Loubota Panzou, G. J. *et al.* Architectural differences associated with functional traits among 45 coexisting tree species in Central Africa. *Funct. Ecol.* **32**, 2583–2593 (2018).
40. King, D. A., Davies, S. J., Tan, S. & Nur Supardi, M. N. Trees approach gravitational limits to height in tall lowland forests of Malaysia. *Funct. Ecol.* **23**, 284–291 (2009).
41. Antin, C., Pélissier, R., Vincent, G. & Couteron, P. Crown allometries are less responsive than stem allometry to tree size and habitat variations in an Indian monsoon forest. *Trees - Struct. Funct.* **27**, 1485–1495 (2013).
42. Blanchard, E. *et al.* Contrasted allometries between stem diameter, crown area, and tree height in five tropical biogeographic areas. *Trees - Struct. Funct.* **30**, 1953–1968 (2016).
43. Martin-Ducup, O. *et al.* Terrestrial laser scanning reveals convergence of tree architecture with increasingly dominant crown canopy position. *Funct. Ecol.* **34**, 2442–2452 (2020).
44. Fayolle, A. *et al.* A regional allometry for the Congo basin forests based on the largest ever destructive sampling. *For. Ecol. Manage.* **430**, 228–240 (2018).
45. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
46. Réjou-Méchain, M. *et al.* Unveiling African rainforest composition and vulnerability to global change. *Nature* **593**, 90–94 (2021).
47. Larjavaara, M. & Muller-Landau, H. C. Measuring tree height: A quantitative comparison of two common field methods in a moist tropical forest. *Methods Ecol. Evol.* **4**, 793–801 (2013).
48. Mehtatalo, L. Functions for forest biometrics - R package 'lmfor'. *CRAN - An R software for statistical analysis* 43 (2015). Available at: https://cran.r-project.org/web/packages/lmfor/index.html. (Accessed: 22nd July 2017)
49. Baskerville, G. Use of Logarithmic Regression in the Estimation of Plant Biomass. *Can. J. For.* **2**, 49–53 (1972).
50. Mascaro, J., Litton, C. M., Hughes, R. F., Uowolo, A. & Schnitzer, S. A. Minimizing Bias in Biomass Allometry: Model Selection and Log-Transformation of Data. *Biotropica* **43**, 649–653 (2011).
51. Zeide, B. Analysis of growth equations. *For. Sci.* **39**, 594–616 (1993).
52. Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-effects models using lme4. *arXiv:1406.5823v1[stat.CO]23* 1–51 (2014). doi:10.1177/009286150103500418
53. Neath, A. A. & Cavanaugh, J. E. The Bayesian information criterion: Background, derivation, and applications. *Wiley Interdiscip. Rev. Comput. Stat.* **4**, 199–203 (2012).
54. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proc. 14th Int. Jt. Conf. Artif. Intell. - Vol. 2* **2**, 1137–1143 (1995).
55. Efron, B. & Tibshirani, R. Improvements on Cross-Validation: The .632+ Bootstrap Method. *J. Am. Stat. Assoc.* **92**, 548–560 (1997).
56. R Development Core Team. A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria* (2021).
57. Lewis, S. L. *et al.* Increasing carbon storage in intact African tropical forests. *Nature* **457**, 1003–1006 (2009).
58. Jucker, T. *et al.* Allometric equations for integrating remote sensing imagery into forest monitoring programmes. *Glob. Chang. Biol.* **23**, 177–190 (2017).
59. Ledo, A. *et al.* Re-evaluation of individual diameter: Height allometric models to improve biomass estimation of tropical trees. *Ecol. Appl.* **26**, 2374–2380 (2016).

60. Vieilledent, G. *et al.* A universal approach to estimate biomass and carbon stock in tropical forests using generic allometric models. *Ecol. Appl.* **22**, 572–583 (2012).
61. Sullivan, M. J. P. *et al.* Field methods for sampling tree height for tropical forest biomass estimation. *Methods Ecol. Evol.* **2018**, 1–11 (2018).
62. Imani, G. *et al.* Height-diameter allometry and above ground biomass in tropical montane forests: Insights from the Albertine Rift in Africa. *PLoS One* **12**, (2017).
63. Shenkin, A. *et al.* The Influence of Ecosystem and Phylogeny on Tropical Tree Crown Size and Shape. *Front. For. Glob. Chang.* **3**, (2020).
64. Iida, Y. *et al.* Tree architecture and life-history strategies across 200 co-occurring tropical tree species. *Functional Ecology* **25**, 1260–1268 (2011).
65. Chave, J. *et al.* Error propagation and scaling for tropical forest biomass estimates. *Philos. Trans. R. Soc. B Biol. Sci.* **359**, 409–420 (2004).
66. Réjou-Méchain, M., Tanguy, A., Piponiot, C., Chave, J. & Hérault, B. Biomass: an R Package for Estimating Above-Ground Biomass and Its Uncertainty in Tropical Forests. *Methods Ecol. Evol.* **8**, 1163–1167 (2017).
67. Picard, N., Boyemba Bosela, F. & Rossi, V. Reducing the error in biomass estimates strongly depends on model selection. *Ann. For. Sci.* **72**, 811–823 (2015).
68. Jucker, T. *et al.* Tallo–a global tree allometry and crown architecture database. *Glob. Chang. Biol.* 5254–5268 (2022). doi:10.1111/gcb.16302

Tables

Table 1 Summary characteristics of the forest types in six different sites. Countries and forest types are abbreviated as follows: semi-deciduous forest in Cameroon (SEF-CAM), transition between semi-deciduous and evergreen forest on sandstone plateau in Central Africa Republic (TSS-CAR), seasonally flooded forest including monodominant *Gilbertiodendron dewevrei* (De Wild.) J.Leonard species in the Republic of Congo (SFM-CON), transition between semi-deciduous and evergreen forest in the DR Congo (TSE-DRC), evergreen forest of Littoral type in Equatorial Guinea (EFL-RGE), and mainland evergreen forest in Gabon (EVF-GAB). The geographical coordinates in latitude (Lat.) and longitude (Lon.), mean annual temperature range (MAT in °C), mean annual rainfall range (MAR in mm), sampling intensity (number of trees, number of species and diameter range in cm) and logging companies are equally provided.

Country (Country-forest type code)	Forest type	Logging company	GPS Coordinates (Lat.; Lon.)	Geological substrate	Seasons (duration)	MAT (°C) min-max.	MAR (mm)	Soil (colour)	Sampling intensity
Cameroon (SEF-CAM)	Semi-deciduous	*Alpicam-Grumcam*	3°42' – 4°3' N ;14°14' – 14°34' E	Granites & Schistes	Dry (Jun–Aug & Nov–Mar); Rain (Apr–May & Sep–Dec)	20–25	1500–2000	Ferralsols (Reddish-brown)	Trees = 132; Species = 15; Diameter range = [11.5-180 cm]
Central African Republic (TSS-CAR)	Transition between Evergreen & Semi-deciduous on sandstone plateau	*SEFCA*	3.919;16.896	Sedimentary, Metamorphic, igneous	Dry (Jun–Aug & Nov–Mar); Rain (Apr–May & Sep–Dec)	23–27	1400–1600	Ferralsols (Reddish-brown)	Trees = 143; Species = 16; Diameter range = [10.5-173 cm]
Republic of Congo (SFM-CON)	Seasonally flooded including monodominant species	*CIB-OLAM*	2°13'–2°14'N; 17°00'–17°07'E	Quartzite	Dry (Jun–Aug & Dec–Feb); Rain (Mar–May & Sep–Nov)	24–28	1400–1700	Xanthic Ferralsols	Trees = 141; Species = 16; Diameter range = [10.3-208 cm]
Democratic Republic of the Congo (TSE-DRC)	Transition between Semi-deciduous & Evergreen	*CFT*	0°1'–0°8' N;25°20'–25°31'E	Sedimentary	Dry (Jun–Sep & Dec–Feb); Rain (Mar–May & Oct–Nov)	20–26	1500–2000	Ferralsols (Yellowish-red)	Trees = 142; Species = 16; Diameter range = [11.7-160.5cm]
Gabon (EVF-GAB)	Evergreen	*Rougier Haut-Abanga*	0°15'–0°50' N ;10°30'–11°30' E	Metamorphic or granites	Dry (Jun–Sep & Dec–Jan); Rain (Feb–May & Oct–Nov)	24–26	1800–2000	Ferralsols	Trees = 178; Species = 16; Diameter range = [12.3-169.3 cm]
Equatorial Guinea (EFL-RGE)	Littoral Evergreen	*COMALI*	1.316;9.537	Sedimentary	Dry (Jul–Sep & Dec–Feb); Rain (Mar–Jun & Oct–Nov)	23–30	2000–2500	Ferralsols (yellow-Sandy)	Trees = 109; Species = 14; Diameter range = [11-172 cm]

Table 2 H–D and CA–D allometric models compared in this study.

Allometric models	References
H–D allometry	
Best fitted model for the subject forest type	this study
Best fitted regional model for each fold	this study
Pantropical climate-based model	Chave et al., 2014
Central Congo regional equation for mixed forest type	Kearsley et al., 2013
Tropical Africa regional model	Banin et al., 2012
Tropical Africa regional model	Lewis et al., 2009
Central Africa regional model	Feldpausch et al., 2011
Central Africa regional model	Feldpausch et al., 2012
CA–D allometry	
Best fitted model for the subject forest type	this study
Best fitted regional model for each fold	this study
Central Africa regional model	Blanchard et al., 2016

Table 3 Best performance height-diameter (H–D) and crown area-diameter (CA–D) allometric models for central African forest types. Forest types are described in Table 1. The name of fitted models is given by the mathematical function used (as given in Table S2) followed by "_FT" if the model was fitted at the forest type level or "_RG" if fitted at regional level. The values of coefficients *a*, *b*, *c* then are the estimates resulting from the model fit. The name of the models taken from the literature is after their author and date of publication. Coefficient values are then the published values.

				Determination coefficients		Model coefficients			Starting values		
Scenario	**Validation fold**	**Model name**	**Mathematical function**	**RMSE (m)**	**RB (%)**	**a**	**b**	**c**	**A**	**B**	**C**
H–D Allometric models											
Local forest type	SEF	Linear_FT	$\mathcal{X} = a + b \times D$	1.526	-0.003	18.657	0.298				
	EVF	Linear_FT	$\mathcal{X} = a + b \times D$	1.326	-0.439	21.070	0.242				
	TSE	Power_FT	$\mathcal{X} = a \times D^b$	1.165	-0.065	9.418	0.271		6.161	0.360	
	SFM	Feldpausch et al. (2012)	$\mathcal{X} = a \times (1-\exp(-b \times D^c))$	0.650	0.111	50.453	-0.047	0.812			
	TSS	Banin et al. (2012)	$\mathcal{X} = a - b \times \exp(-c \times D)$	1.124	0.262	45.100	-42.800	0.025			
	EFL	Chave et al. (2014)	$\mathcal{X}=\exp(a-E + b \times D - c \times D^2))$	1.290	0.085	0.893	0.760	0.034			
Regional encompassing the subject forest type	Region-Fold1 to 4	Banin et al. (2012)	$\mathcal{X} = a - b \times \exp(-c \times D)$	6.140	-1.972	45.100	-42.800	0.025			
	Region-Fold5	Meyer_RG	$\mathcal{X} = a \times (1-\exp(-b \times D))$	6.150	0.305	46.125	0.028		85.475	0.007	
Regional not encompassing the subject forest type	SFM	Banin et al. (2012)	$\mathcal{X} = a - b \times \exp(-c \times D)$	4.760	-0.076	45.100	-42.800	0.025			
	TSS	Banin et al. (2012)	$\mathcal{X} = a - b \times \exp(-c \times D)$	4.606	-1.477	45.100	-42.800	0.025			
	EFL	Lewis et al. (2009)	$\mathcal{X} = a \times (1-\exp(-b D^c))$	5.064	-0.595	54.010	-0.053	0.759			
	SEF	Lewis et al. (2009)	$\mathcal{X} = a \times (1-\exp(-b D^c))$	6.950	-1.053	54.010	-0.053	0.759			
	EVF	Lewis et al. (2009)	$\mathcal{X} = a \times (1-\exp(-b D^c))$	6.045	0.331	54.010	-0.053	0.759			
	TSE	Feldpausch et al. (2011)	$\mathcal{X} = a + b \times \log(D)$	7.868	8.387	1.153	0.554				
CA–D allometric models											
Local forest type	SEF	Linear_FT	$\mathcal{X} = a + b \times D$	18.636	3.045	-94.516	4.807				
	SFM	Linear_FT	$\mathcal{X} = a + b \times D$	15.960	7.403	-78.431	3.778				
	EVF	Michailoff_FT	$\mathcal{X} = a \times \exp(-b \times D^{-1})$	13.631	5.982	609.076	92.745		247.054	49.395	
	TSS	Parabolic-log_FT	$\mathcal{X} = a + b \times \log(D) + c \times \log(D)^2$	11.851	0.299	1151.943	-757.198	125.665			
	TSE	Linearized-power_FT	$log(\mathcal{X}) = a + b \times \log(D)$	29.544	8.916	-1.681	1.609				
	EFL	Parabolic-log_FT	$\mathcal{X} = a + b \times \log(D) + c \times \log(D)^2$	12.429	3.844	653.583	-475.373	89.104			
Regional encompassing the subject forest type	Region-Fold1	Linear_RG	$\mathcal{X} = a + b \times D$	97.337	-1.662	-67.517	3.874				
	Region-Fold2	Parabolic-log_RG	$\mathcal{X} = a + b \times \log(D) + c \times \log(D)^2$	100.412	50.662	1128.264	-738.262	122.809			
	Region-Fold3	Naslund_RG	$\mathcal{X} = D^2/(a + b \times D)^2$	101.337	45.072	3.654	0.019		4.609	0.017	
	Region-Fold4	Naslund_RG	$\mathcal{X} = D^2/(a + b \times D)^2$	131.422	64.542	3.585	0.020		4.517	0.018	
	Region-Fold5	Parabolic_RG	$\mathcal{X} = a + b \times D + c \times D^2$	104.798	47.833	-46.120	3.162	0.004			

Regional not encompassing the subject forest type	SEF	Log-normal_RG	$log(\mathcal{X}) = a + b \times \log(D) + c \times \log(D)^2$	122.152	60.397	-0.285	0.981	0.062			
	SFM	Korf_RG	$\mathcal{X} = a \times \exp(-b \times D^c)$	89.763	81.147	10833.950	18.270	0.359	1941.531	22.983	0.540
	EVF	Blanchard et al. (2016)	$\mathcal{X} = a + b \times \log(D)$	107.375	15.656	1.240	0.718				
	TSS	Parabolic_RG	$\mathcal{X} = a + b \times D + c \times D^2$	67.439	26.898	-41.010	3.056	0.005			
	TSE	Log-normal_RG	$log(\mathcal{X}) = a + b \times \log(D) + c \times \log(D)^2$	140.169	46.068	0.208	0.729	0.094			
	EFL	Parabolic-log_RG	$\mathcal{X} = a + b \times \log(D) + c \times \log(D)^2$	104.152	1.650	1219.950	-785.654	128.403			

Figures

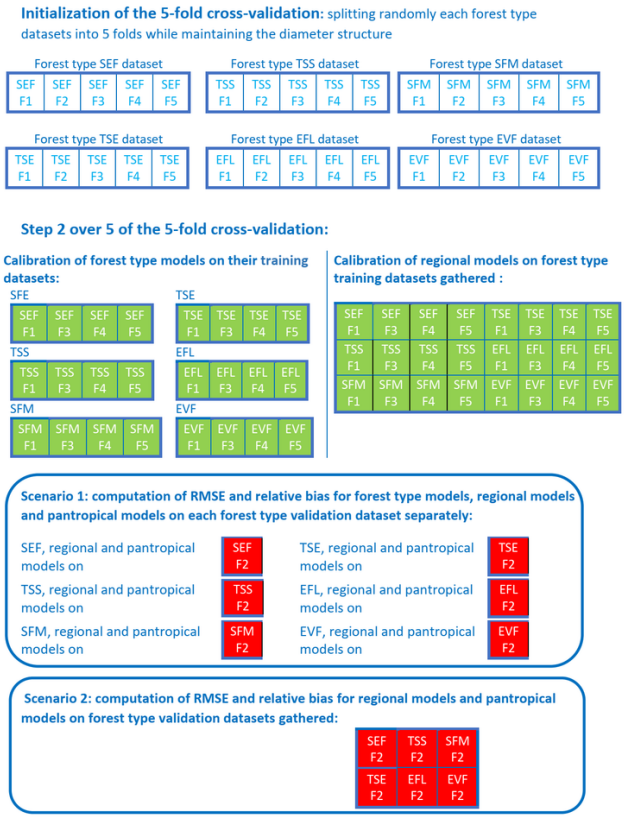

Figure 1

k-fold cross-validation schemes for the first and second scenarios. Green blocks represent training dataset and red blocks represent validation datasets. Green arrows represent fitting and red arrows represent performance evaluation. SEF-CAM: semi-deciduous forest type in Cameroon; TSS-CAR: transition between semi-deciduous and evergreen forest on sandstone plateau in Central Africa Republic; SFM-CON: seasonally flooded forest including monodominant *Gilbertiodendron dewevrei* (De Wild.) J.Leonard species in Republic of Congo; TSE-DRC: transition between semi-deciduous and evergreen forest in the Democratic Republic of Congo; EFL-RGE: evergreen forest of Littoral type in Equatorial Guinea; and EVF-GAB: mainland evergreen forest in Gabon.

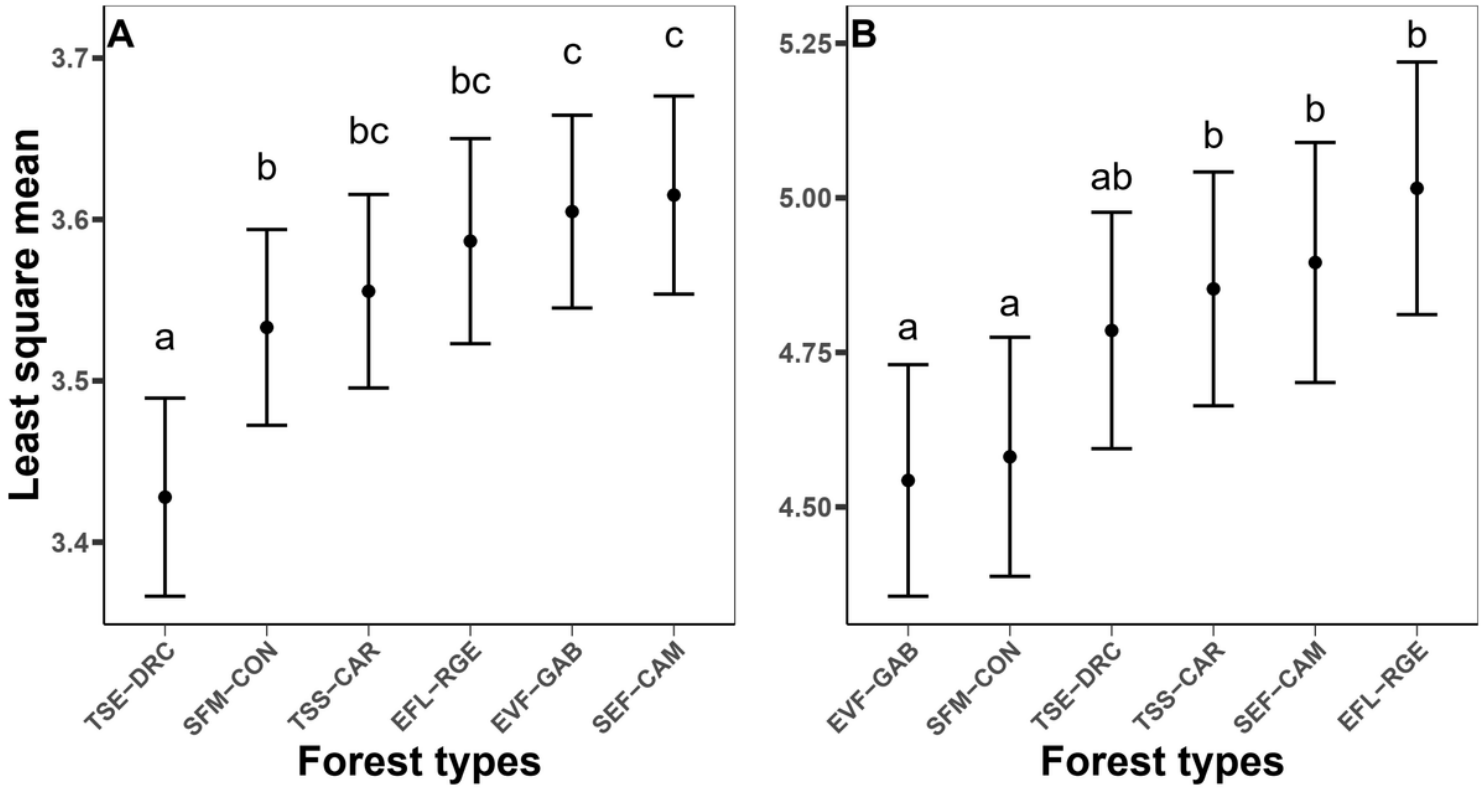

Figure 2

Pairwise comparison of forest type effect on H–D allometric models (A) and CA–D allometric models (B). Forest types are defined in Figure 1. Error bars with the same letters are not significantly different.

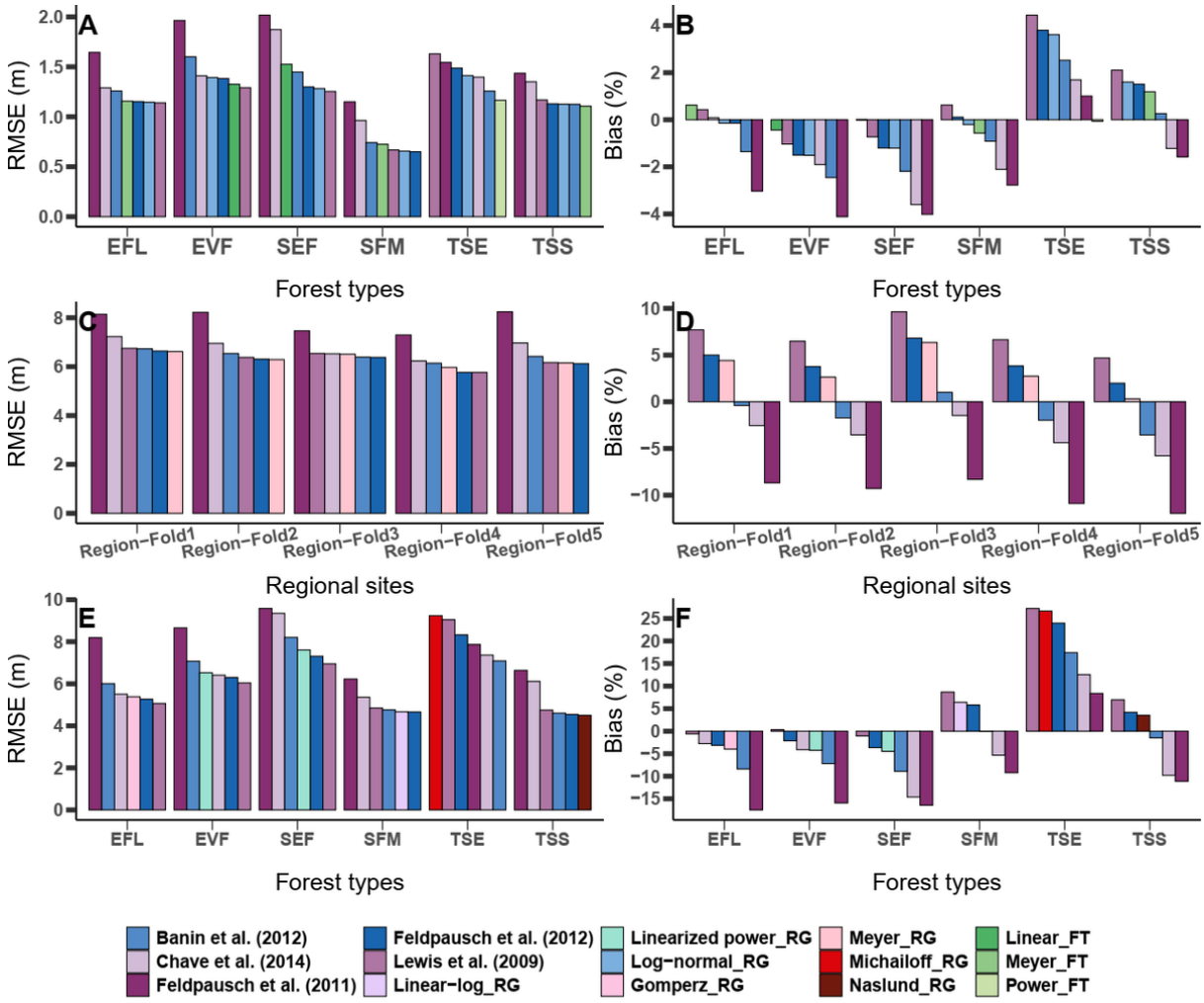

Figure 3

Performances in terms of RMSE (left column) and bias (right column) of the height-diameter allometric models compared for three scenarios. Each bar corresponds to a model as identified by its colour. The different group of bars on the x-axis correspond to the different validation folds used to assess model performance. Some models were fitted: their name is given by the mathematical function used (as given in Table S2) followed by "_FT" if the model was fitted at the forest type level or "_RG" if fitted at regional level. Some models were taken from the literature: their name is after their author and date of publication. The model by Kearsley et al. (2013), although included in the model comparison, was not shown here due to its poor performance. Panels A and B: model comparison for scenario 1 (models derived for the forest type *vs.* regional models *vs.* pantropical models); the different forest types on the x-axis also correspond to those used for model calibration. Panels C and D: model comparison for scenario 2 (regional model encompassing the subject forest type *vs.* pantropical models). Panels E and F: model comparison for scenario 3 (regional model not encompassing the subject forest type *vs.* pantropical models); the different forest types on the x-axis also correspond to those excluded from model calibration. RMSE is relative mean squared error (A, C and D) and bias is relative bias (B, D and F). Abbreviations for forest types are defined in Figure 1.

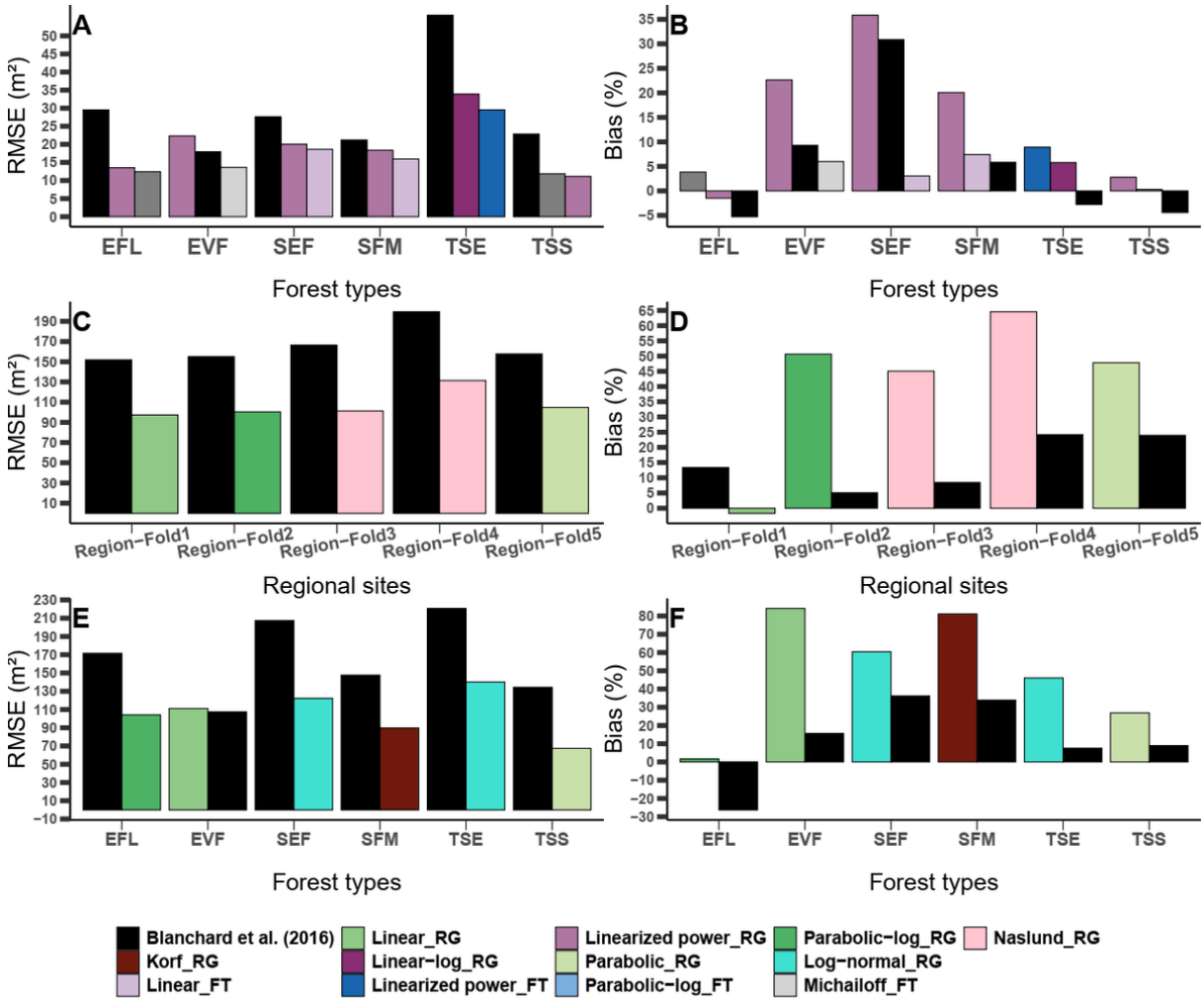

Figure 4

Performances in terms of RMSE (left column) and bias (right column) of the crown area-diameter allometric models compared for three scenarios. Each bar corresponds to a model as identified by its colour. The different blocks of bars on the x-axis correspond to the different validation folds used to assess model performance. Some models were fitted: their name is given by the mathematical function used (as given in Table S2) followed by "_FT" if the model was fitted at the forest type level or "_RG" if fitted at regional level. One model, named Blanchard et al. (2016), was taken from the literature. Panels A and B: model comparison for scenario 1 (models derived for the forest type *vs.* regional models *vs.* pantropical models); the different forest types on the x-axis also correspond to those used for model calibration. Panels C and D: model comparison for scenario 2 (regional model encompassing the subject forest type *vs.* pantropical models). Panels E and F: model comparison for scenario 3 (regional model not encompassing the subject forest type *vs.* pantropical models); the different forest types on the x-axis also correspond to those excluded from model calibration. RMSE is relative mean squared error (A, C and D) and bias is relative bias (B, D and F). Abbreviations for forest types are defined in Figure 1.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- LibalahetalSI12TreeHeightCrownAreaDiameterAllometriessrep.docx
- LibalahetalSI38TreeHeightCrownAreaDiameterAllometriessrep.xlsx