Supplementary material – Tree height uncertainty biases aboveground biomass estimation more than wood density in miombo woodlands

Yambayamba *et al*., 2025

S1 Reparametrizing of existing *AGB* and *H-D* models

We used the tree harvest dataset (Table 1 in the main text) to reparametrize miombo-specific versions of eqs. 1 and 2 described in the main text. The two existing models are both based on the power law function which, in its basic form, can be expressed as:

$$y = \alpha \times x^{\beta} \quad \text{(S1)}$$

where, y is the response variable (*AGB* in our case); α is the normalization constant; x is the predictor variable (D or $\rho D^2 H$ in our case); and β is the scaling exponent. We linearized eq. S1 to obtain:

$$\ln(AGB) = \ln(\alpha) + \beta \times \ln(D), \text{ and} \quad \text{(S2)}$$

$$\ln(AGB) = \ln(\alpha) + \beta \times \ln(\rho D^2 H). \quad \text{(S3)}$$

We then used the brms package to fit log–log regression models in a Bayesian framework. We used weak priors which we assumed to follow a normal distribution for the two parameters - a truncated $N(2,5)$ for α and truncated $N(2,3)$ for β. These limits were chosen to cover realistic range of parameter estimates. We used the Metabolic Scaling Theory prediction (West et al., 1999) for the scaling exponent (β) as basis for the second prior in the power law model and relied on literature values for setting the prior for the first parameter (α) since it does not vary markedly between studies.

Since the tree harvest data came from three sites, we tested for the effect of site by fitting models with and without site as a fixed factor. Therefore, we fitted four models in total. Model parameters estimation was implemented in R using the brms package version 2.19.0 which uses Stan on the back-end and samples were drawn using the no-U-turn sampler (NUTS) (Bürkner, 2017). Inference on parameters was based on 6000 posterior samples drawn from 4 parallel chains after discarding 1000 burn-in or warmup draws per chain. The use of 4 parallel chains enabled us to estimate the Gelman-Rubin convergence statistic (Rhat), a measure of model convergence, and the Bulk_ESS and Tail_ESS statistics, which give the number of independent samples from the posterior distribution which would give the same values as those estimated from the samples used in fitting the model. If a model has successfully converged, the Rhat statistic should be less than or equal to 1.10 while both Bulk_ESS and Tail_ESS should be $\geq$ 1000 (Bürkner, 2017; Mcelreath, 2020). The four models fitted had Rhat statistic = 1.00 with both Bulk_ESS and Tail_ESS for parameter estimates being at least 1000. Further, posterior predictive checks revealed that all the models adequately fitted the observed data (Fig. S2a – d).

We then compared the four fitted models using the leave-one-out cross-validation information criterion (LOOIC). We chose LOOIC, among others such as the widely applicable information criterion (WAIC), because it is more robust even when the data has outliers or influential observations. The LOOIC statistic can be used as a measure of model predictive accuracy and to compare alternative models in the absence of validation data. Using the LOOIC approach, model fitting and prediction are done iteratively with each observation in the dataset predicted from the model fitted using the remaining data. As an example, if a dataset has 100 observations, then the model will be fitted 100 times using 99 data points in turn to fit the model and predicting the one left out at each iteration. At each iteration, a measure of prediction is calculated via pointwise log-likelihood and these are summed to obtain a single measure of

model accuracy (Vehtari et al., 2017; Bürkner et al., 2021). LOOIC was computed using the loo R package version 2.6.0 (Vehtari et al., 2023). A model with the smallest LOOIC among competing models is regarded as the best. Further, we compared the four models using the leave-one-out (LOO) adjusted R-squared (R2) - LOO-R2 (Vehtari et al., 2025). The LOO-R2 measures the predictive performance of the model for new data. Next, we compared the precision and bias/accuracy of the four models based on RMSE and PE respectively.

The analysis of the four fitted models revealed that addition of site as a fixed factor did not significantly, based on the 95% credible intervals of the LOO-R2, improve the models (Table S1). We therefore chose to use simpler models - without site as a factor – eqs. 5 and 6 in the main text. The two reparametrized models adequately fitted the data as revealed by both fit statistics (correlation coefficient (R) = 0.97 between observed and fitted values for both models) and plots of observed vs fitted values and residuals vs observed values (Fig. S3).

We used a similar modelling approach, of reparametrizing *AGB* models, to reparametrize a miombo specific version of the literature *H-D* model (eq. 8 in the main text). We also fitted *H-D* models were site was included as a fixed factor in order to asses its impact on *H-D* allometry. Specifically, we fitted four *H-D* models – the first was a simple model with only *D* as the predictor variable and the second one in which we allowed parameters to vary across species. We then included site as a fixed factor to each of the two models. We compared the performance of the four fitted *H-D* models based on the same fit statistics (LOOIC and LOO-R2) and performance metrics (RMSE and PE) as used for *AGB* models. The results of this analysis revealed that site did not significantly improve the prediction performance of the models (see Table S1). We therefore chose to use the simpler model without site as a fixed factor – eq. 7 in the main text.

SUPPLIMENTARY TABLES

Table S1: Comparison of model fit and performance for *AGB* and *HD* models fitted with and without site as a factor. For LOO-R2, the 95% credible intervals are shown in square brackets.

Model	**LOOIC**	**LOO-R2**	**RMSE**	**PE (%)**
Fitted *AGB* models				
AGB~D	148.5	0.93 [0.91, 0.95]	200.9	-4.9
AGB~D + site	159.6	0.93 [0.90, 0.95]	181.0	-5.0
$AGB \sim D^2H\rho$	138.7	0.94 [0.91, 0.95]	220.2	-8.1
$AGB \sim D^2H\rho$+site	122.8	0.94 [0.92, 0.96]	169.3	-7.6
Fitted *H-D* models				
H~D	968.9	0.71 [0.69, 0.72]	2.9	-3.6
H~D + site	828.4	0.72 [0.70, 0.73]	2.9	-2.7
H~D/species	-90.4	0.77 [0.76, 0.78]	2.6	-2.7
H~D/species+site	-268.6	0.78 [0.76, 0.79]	2.5	-2.7

Table S2: Fitted aboveground biomass (*AGB*) models parameter estimates, including standard errors and 95% credible intervals.

Model	**Functional form**	**Parameter**		
		α Estimate (Est. error) [95% CI]	β Estimate (Est. error) [95% CI]	σ Estimate (Est. error) [95% CI]
Fitted $\boldsymbol{AGB - D}$ model	$\ln(AGB) = \ln(\alpha) + \beta * \ln(D)$	-2.483 (0.172) [-2.814, -2.157]	2.590 (0.056) [2.482, 2.699]	-0.957 (0.058) [-1.067, -0.838]
Fitted $\boldsymbol{AGB - \rho D^2 H}$	$\ln(AGB) = \ln(\alpha) + \beta * \ln(\rho D^2 H)$	-2.125 (0.160) [-2.432, -1.808]	0.923 (0.020) [0.884, 0.961]	-0.985 (0.059) [-1.097, -0.867]

In order to use fitted models and obtain *AGB* in kg per tree, the equations have to be back transformed and with the addition of the Baskerville correction factor ($\hat{\sigma}^2 * 0.5$) - where $\hat{\sigma}$ is the estimated residual standard error of the model (Baskerville, 1972). The equations thus took the form:

Fitted $\boldsymbol{AGB - D}$ model: $AGB = \exp(-2.483 + 2.590 \times \ln(D) + \exp(-0.957)^2 \times 0.5)$ (S4)

Fitted $\boldsymbol{AGB - \rho D^2 H}$ model: $AGB = \exp(-2.125 + 0.923 \times \ln(\rho D^2 H) + \exp(-0.985)^2 \times 0.5)$ (S5)

Table S3: Fitted height-diameter model parameter estimates, including standard errors and 95% credible intervals.

Model	**Functional form**	**Parameter**		
		α Estimate (Est. error) [95% CI]	β Estimate (Est. error) [95% CI]	σ Estimate (Est. error) [95% CI]
Fitted $\boldsymbol{H-D}$ model	$\ln(H) = \ln(\alpha) + \beta \times \ln(D)$	0.556 (0.036) [0.486, 0.628]	0.613 (0.012) [0.590, 0.637]	-1.427 (0.020) [-1.466, -1.387]

Table S4: Estimated group (random) level parameters for 79 uniquely identified species + mixed species (unidentified – 21 trees or 0.5% of the total trees (4187) inventoried) (provided as an attached Excel file).

Table S5: Metrics for literature and fitted height-diameter models in different tree stem diameter (*D*) size classes of the *H-D* modelling and tree harvest (validation) datasets. RMSE is the root mean square error; and PE is the percentage error. The better model in each category is shown in bold.

Model	**RMSE (m)**				**PE (%)**			
	All trees	Small trees	Medium trees	Large trees	All trees	Small trees	Medium trees	Large trees
Height-diameter modelling data (for fitting the H-D model)								
Eq. (11) (LocalHD model)	**2.6**	**2.2**	**3.1**	**3.9**	**-2.7**	**-3.5**	**-1.8**	**6.0**
Eq. (12) (RegionalHD model)	4.4	3.0	6.3	7.8	-22.4	-14.1	-31.4	-32.1
Tree harvest data (for validation of fitted model)								
Eq. (11) (LocalHD model)	**3.0**	**2.5**	**2.8**	**5.4**	**-0.7**	**-9.5**	**-0.6**	**23.4**
Eq. (12) (RegionalHD model)	4.9	3.5	5.8	5.0	-25.5	-18.7	-30.2	-22.9

SUPPLIMENTARY FIGURES

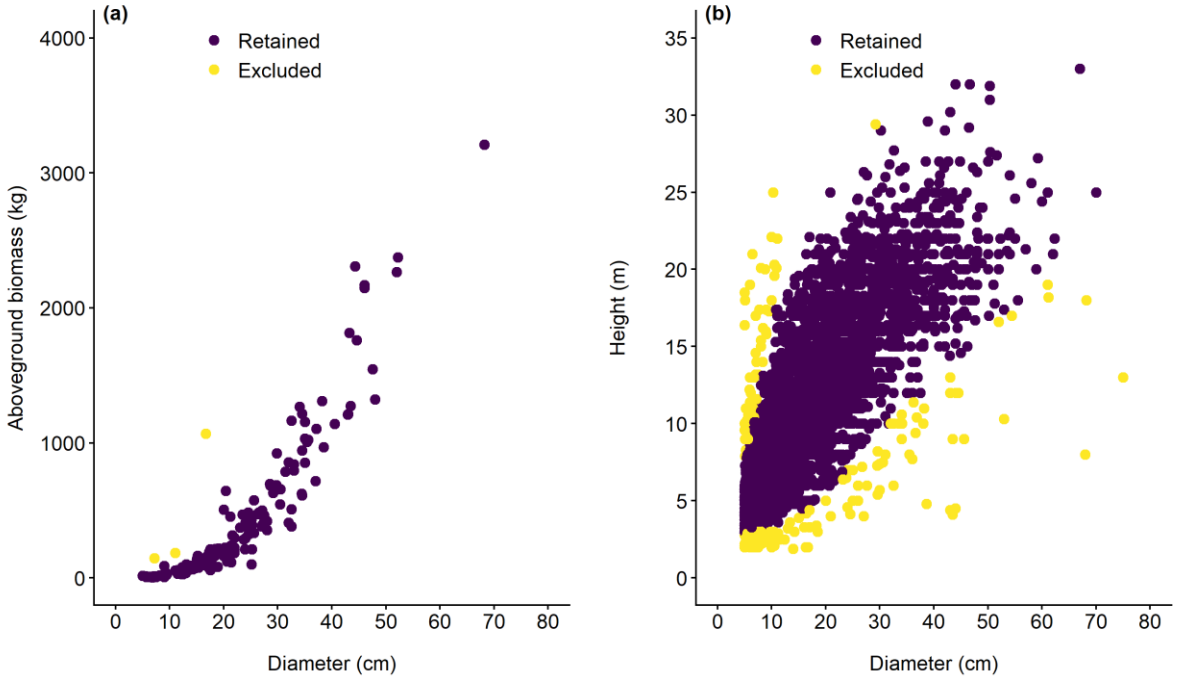


Figure S1. Relationship between aboveground biomass and diameter **(a)** and height and diameter **(b)** showing the data that was excluded (yellow filled circles) and retained (purple filled circles) after using the Cook's distance and studentized residuals to check for outliers in the bivariate relationship between *AGB* and *D* and *H* and *D* (on a log-log scale).

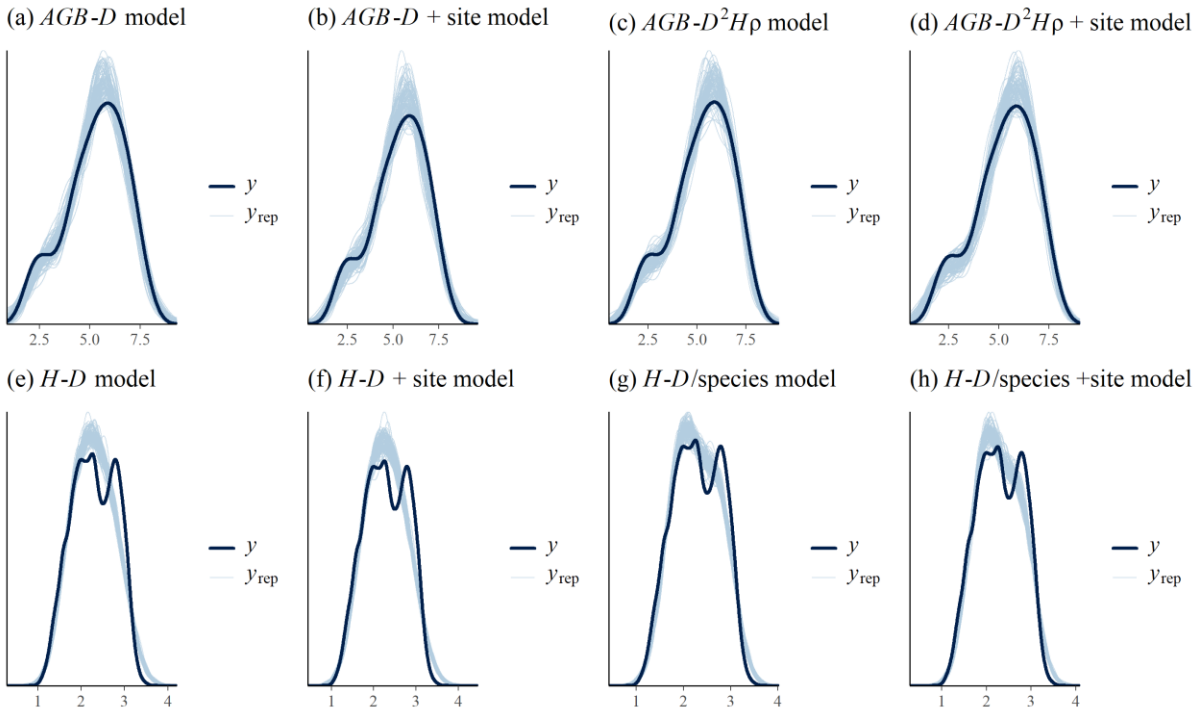


Figure S2. Posterior predictive checks for the fitted *AGB* **(a)** – **(d)** and *H-D* **(e) – (h)** models. The dark blue curve (y) represents observed tree *AGB* **(a)** – **(d)** and *H* **(e) – (h)** while the light blue curve/line represents 100 new samples (y_{rep}) of predicted tree *AGB* **(a)** – **(d)** and *H* **(e) – (h)** based on posterior distribution of parameters for each model. Model fits are considered adequate if the two distributions, i.e. y and y_{rep}, look similar or overlap. Note that the values for **(a)** – **(h)** are based on log transformed values.

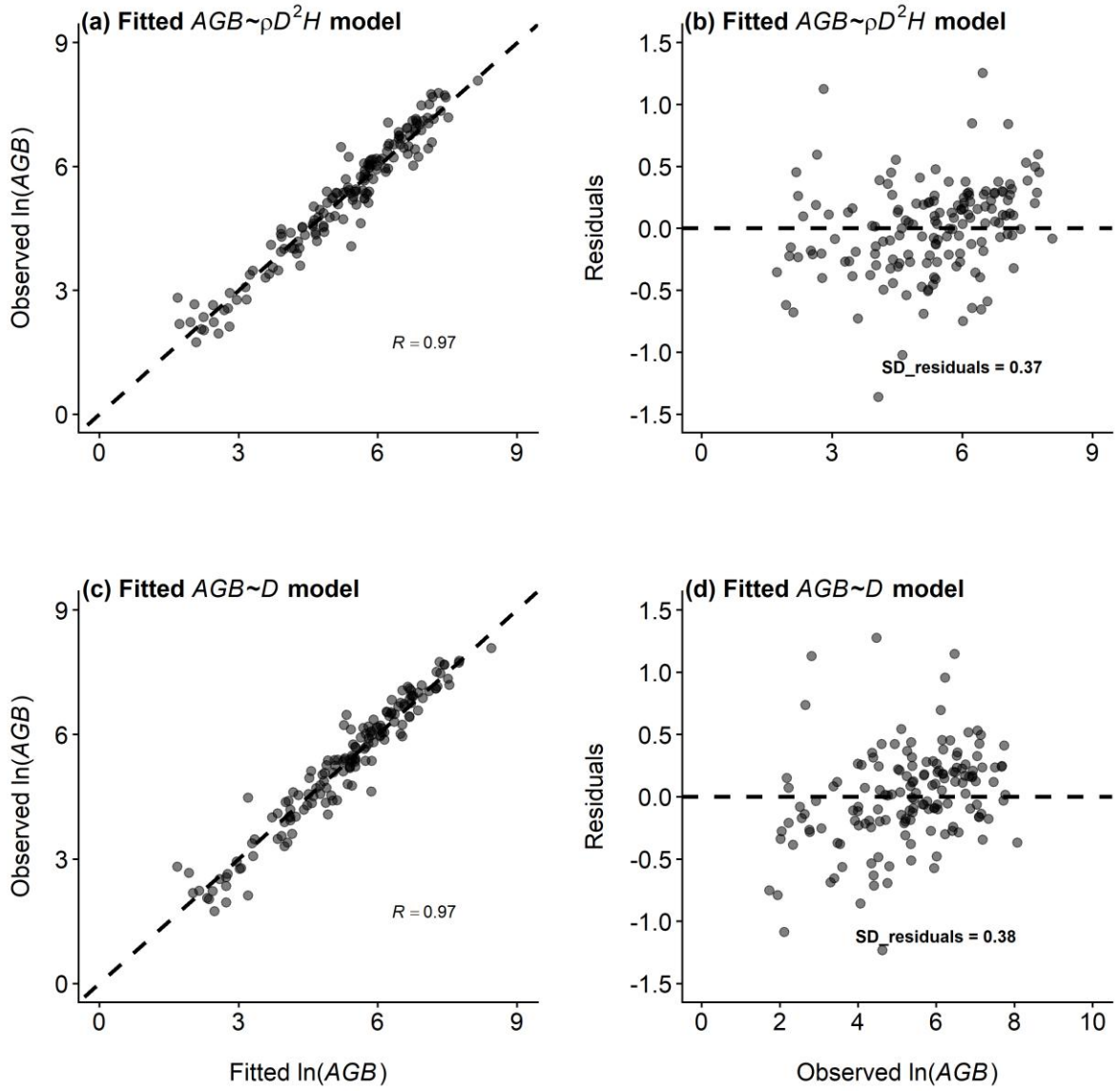


Figure S3. Comparison of goodness of fit for the two fitted *AGB* models, including plots of observed versus fitted values **(a)** and **(c)**; residuals versus observed values **(b)** and **(d)**. In **(b)** and **(d)**, SD_residuals is the standard deviation of the residuals for each model. The black filled circles/dots represent observed data.

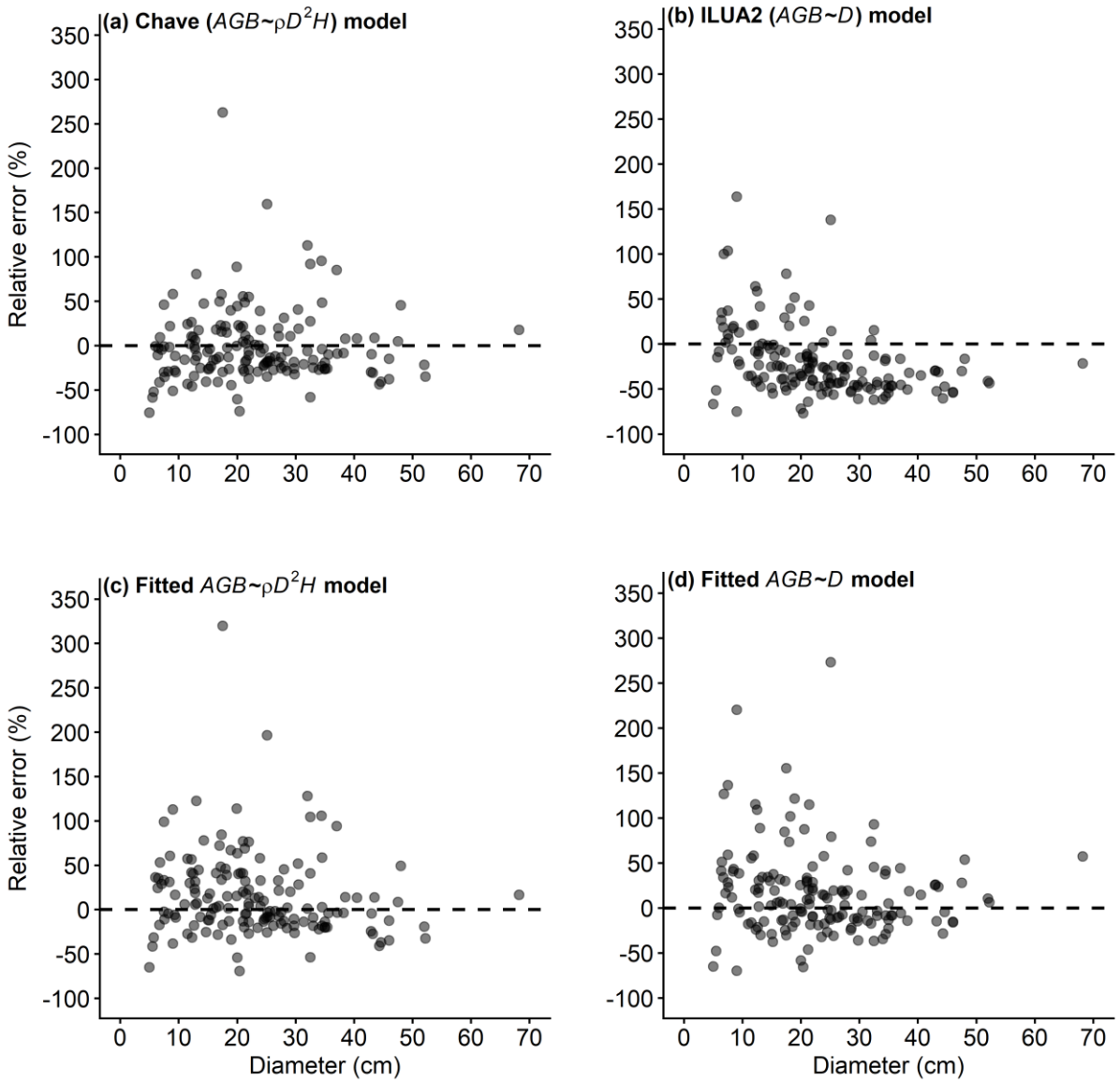
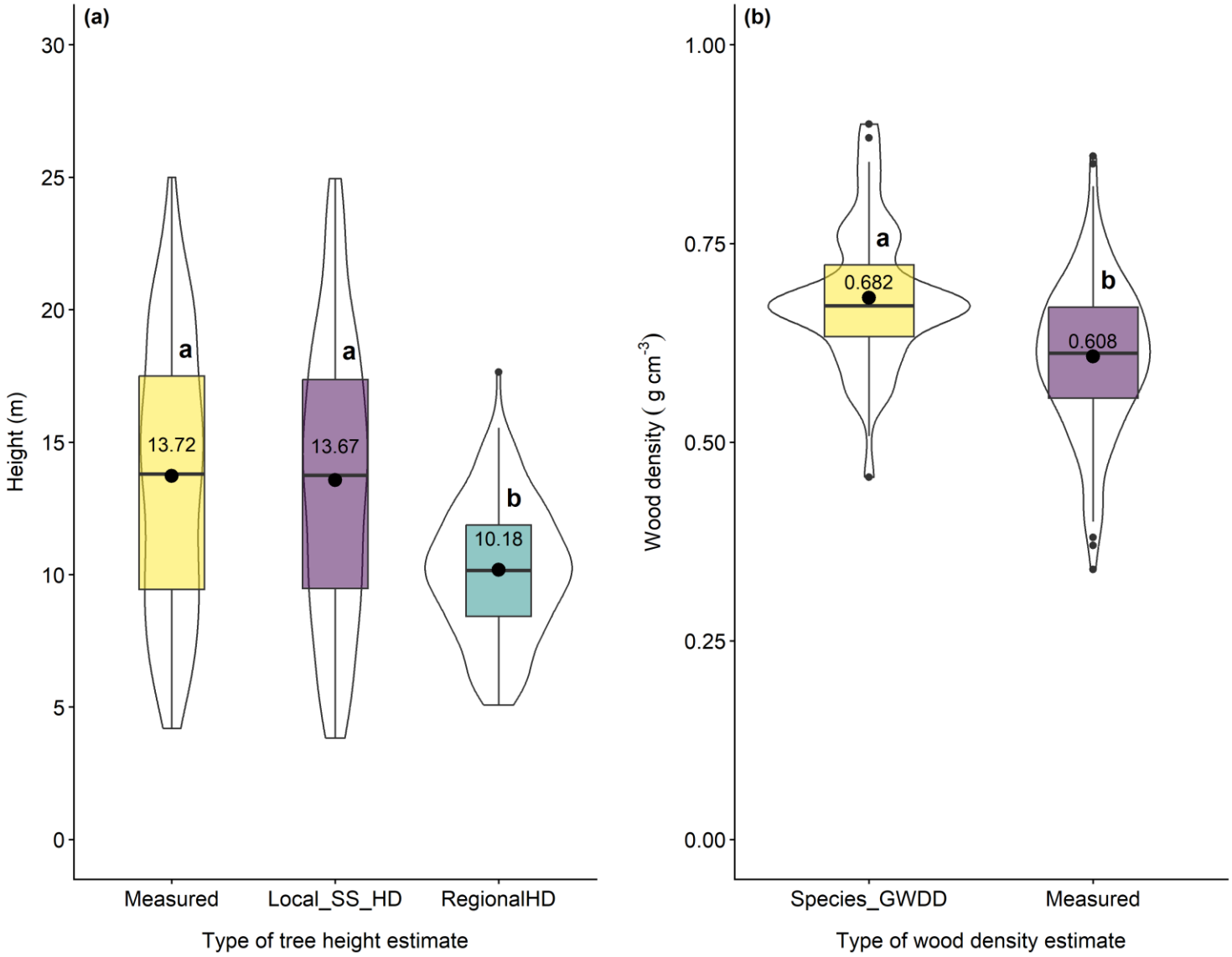


Figure S4. Relationship between relative errors (in % and calculated as: $((\widehat{AGB_i} - AGB_i)/AGB_i) * 100$; where AGB_i is the measured AGB for tree i; and $\widehat{AGB_i}$ is the estimated AGB for tree i) of aboveground biomass estimation and stem diameter based on the two literature models (eqs. 1 and 2) and the two fitted models (eqs. 5 and 6) in the main text.


(a)
Height (m)
30
25
20
15
10
5
0
a
a
b
13.72
13.67
10.18
Measured
Local_SS_HD
RegionalHD
Type of tree height estimate
(b)
Wood density (g cm-3)
1.00
0.75
0.50
0.25
0.00
a
b
0.682
0.608
Species_GWDD
Measured
Type of wood density estimate

158

Figure S5. Comparison of methods for tree **(a)** *H* estimates and **(b)** ρ estimates typically used as inputs in *AGB* models. In **(a)** and **(b)**, using one-way ANOVA on the tree harvest dataset, we found a significant difference ($F_{2,462}$, = 31.28, P<0.001) among the three methods of *H* estimates and a significant difference ($F_{1,308}$ = 51.95, P<0.001) between the two methods of ρ estimates. Different lowercase letters, for the box plots, indicate significantly different estimates based on Tukey HSD pairwise comparison. The number in each box plot is the calculated mean for the *H* **(a)** and ρ **(b)** estimates. Local_SS_HD is the locally fitted species specific *HD* model (eq. 7 in the main text) while RegionalHD is the literature regional *HD* model (eq. 8) in the main text) used in this study. Species_GWDD is the dataset of wood density values extracted from the Global Wood Density Database (GWDD).

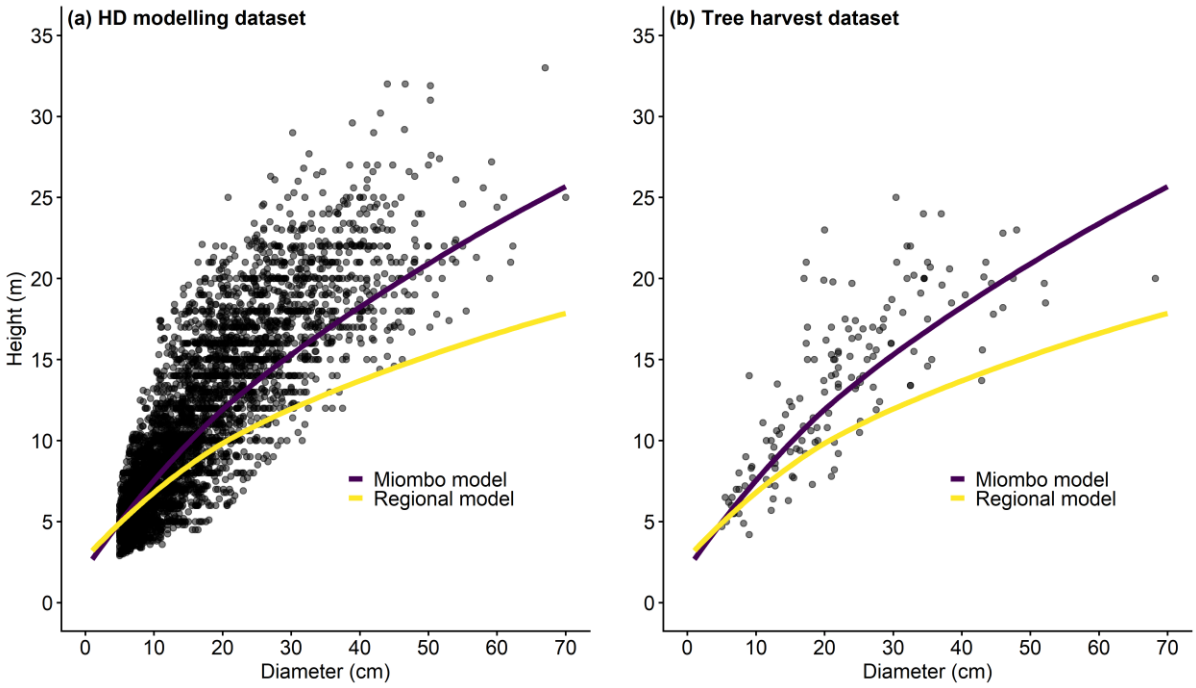


Figure S6. Relationship between tree height (*H*) and stem diameter (*D*) for **(a)** the 4187 trees used to develop the species-specific, miombo *H–D* allometry and **(b)** the 155 trees in the tree harvest dataset used for *AGB* modelling. The line of best-fit of the miombo power-law model fit to these data (eq. 7 in main text) is shown in purple, while the yellow curve shows the fit of the regional *H–D* allometry for tropical dry forests derived from the literature (eq. 8 in main text). The black filled circles/points show the observed data for each dataset. To visualise the models, we generated a predicted *H* for *D* values in the range 1-70 cm (with 1cm interval) and converted point predictions into lines via a locally estimated scatterplot smoothing (LOESS).

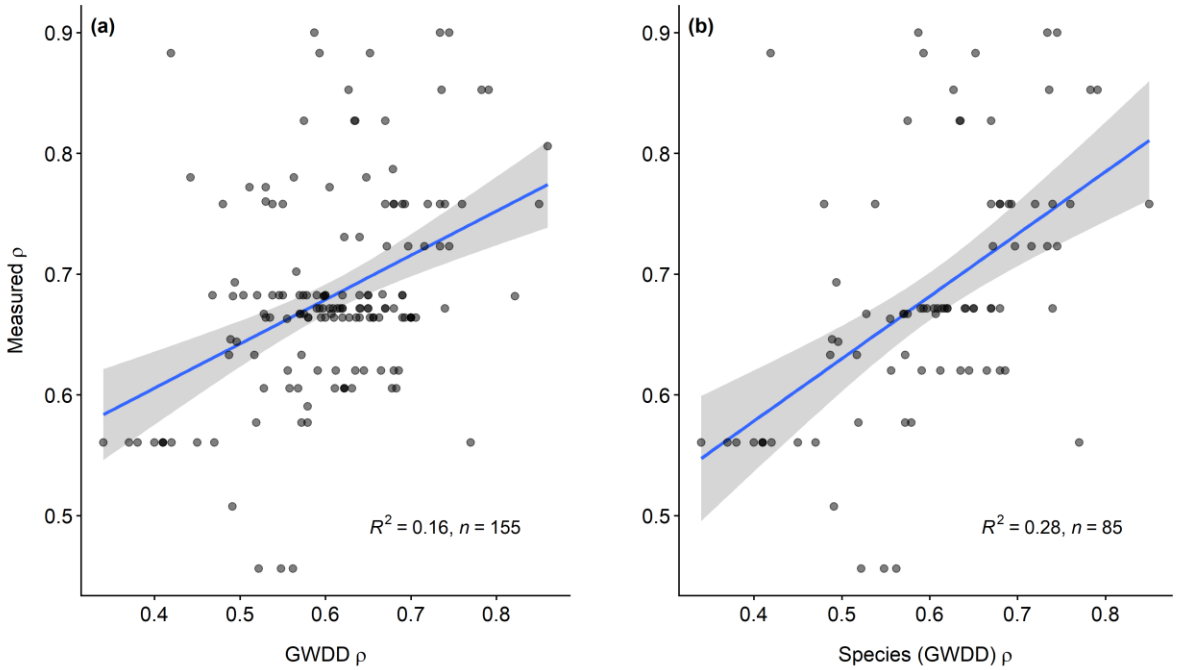


Figure S7. Scatterplot of wood density (ρ) values extracted from the global wood density database (GWDD) versus measured wood density values. In **(a)**, the relationship is shown for all the 155 harvested trees whose ρ was measured **(b)** shows the relationship only for trees whose measured ρ values were matched to those in the GWDD at species level. A total of 19 out of 37 species were matched to those in the GWDD. Thus, the $n = 85$ in **(b)** is the number of trees for the 19 species that are in the GWDD. The coefficient of determination (R^2) between GWDD and measured ρ is shown along with a line of best fit (blue colour) and its 95% confidence bound (grey colour). The black filled circles represent the ρ values.

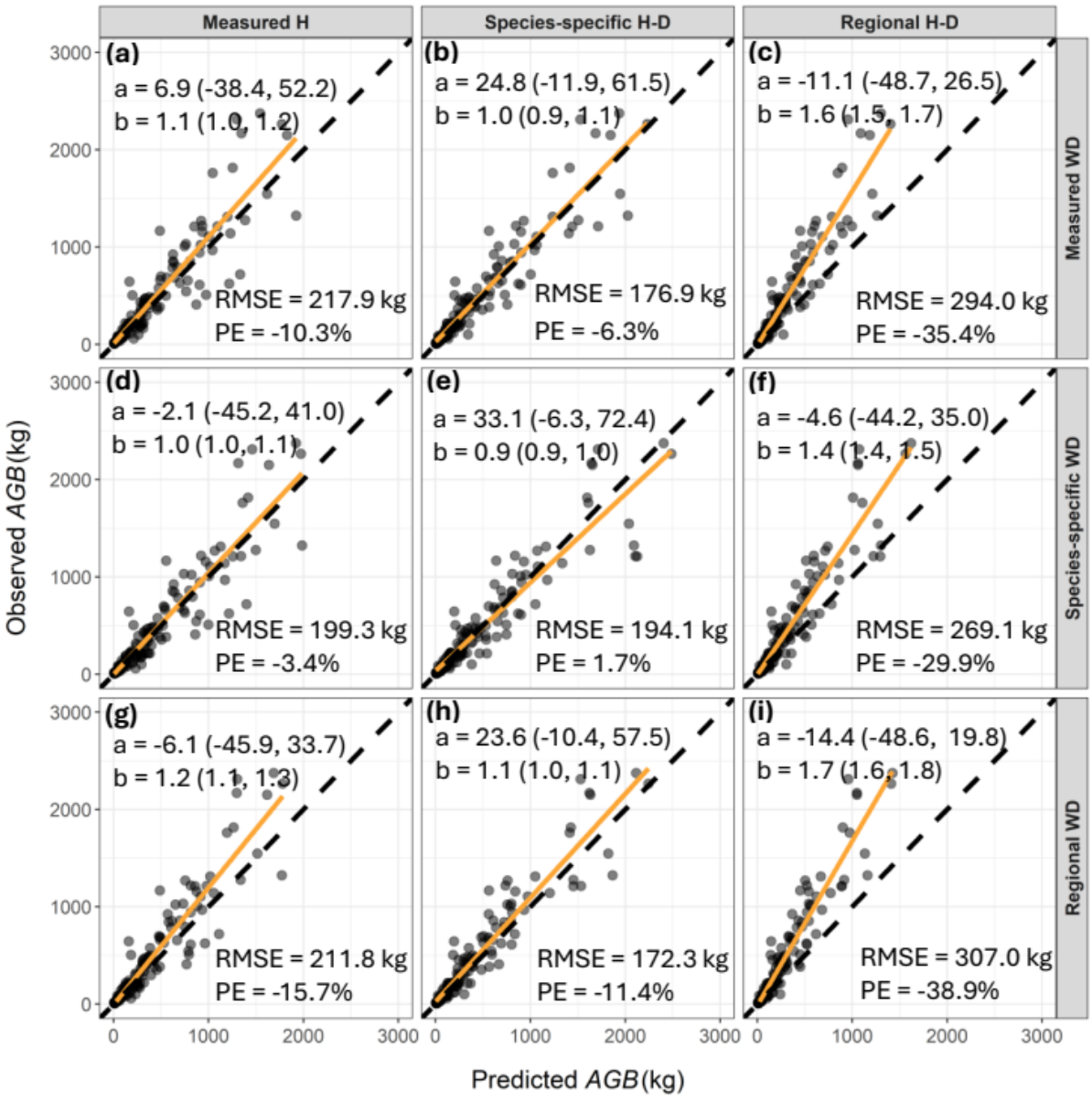


Figure S8. Errors in aboveground biomass (*AGB*) estimation arising from uncertainty in tree height (*H*) and wood density (ρ) for 154 harvested trees – excluding the largest observed/measured *AGB* which was seemingly an outlier when compared to other data values shown in Figure 4 of the main text. Each panel corresponds to one of the nine scenarios described in Table 3 (in the main text) in which we altered inputs of *H* and ρ in the *AGB* model developed by Chave et al. (2014). *H* inputs vary in uncertainty from left to right, from field measured, to those estimated using a species-specific *H-D* model (eq. 7), and finally those

derived from a regional *H-D* model (eq. 8). ρ inputs vary in uncertainty from top to bottom, from field measured ρ values, to species-specific (or closest taxonomic unit) mean values derived from the Global Wood Density Database, and finally those based on a single regional mean value for African tropical forests derived from the literature. The orange coloured solid lines correspond to the fit of a linear model between observed and predicted *AGB* values, while dashed lines illustrate a 1:1 relationship. Fit statistics for each model are reported in the panels, including the intercept, a, and slope ,b, of the regression line (with 95% confidence intervals in brackets), as well as the RMSE and PE.

References

Baskerville, G.: Use of logarithmic regression in the estimation of plant biomass, Canadian Journal of Forest Research, 2, 49-53, 1972.

Bürkner, P.-C.: brms: An R Package for Bayesian Multilevel Models Using Stan, Journal of Statistical Software, 80, 1 - 28, 10.18637/jss.v080.i01, 2017.

Bürkner, P.-C., Gabry, J., and Vehtari, A.: Efficient leave-one-out cross-validation for Bayesian non-factorized normal and Student-t models, Computational Statistics, 36, 1243-1261, 10.1007/s00180-020-01045-4, 2021.

McElreath, R.: Statistical rethinking: A Bayesian course with examples in R and Stan, CRC press, 2020.

Vehtari, A., Gelman, A., and Gabry, J.: Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC, Statistics and Computing, 27, 1413-1432, 10.1007/s11222-016-9696-4, 2017.

Bayesian R2 and LOO-R2: https://avehtari.github.io/bayes_R2/bayes_R2.html, last access: 28/03/2026.

Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Bürkner, P.-C., Paananen, T., and Gelman, A.: "loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models." R package version 2.6.0, https://mc-stan.org/loo/, 2023.

West, G. B., Brown, J. H., and Enquist, B. J.: A general model for the structure and allometry of plant vascular systems, Nature, 400, 664-667, 1999.