Allometric equations for mangrove biomass estimation: a global mixed-effects analysis

Charles A. Price[1,2*], Benjamin Branoff[3*], Todd A. Schroeder[3], Skip J. Van Bloem[4], Monica Papeș[1,2], and Humfredo Marcano-Vega[3]

[1]Ecology and Evolutionary Biology
University of Tennessee
Knoxville, TN, 37996-3140, USA

[2]National Institute for Modeling Biological Systems
University of Tennessee
Knoxville, TN, 37996-3140, USA

[3]USDA Forest Service
Southern Research Station
Forest Inventory and Analysis
2730 Cherokee Farm Way
Suite 101
Knoxville, TN 37920

[4]Baruch Institute of Coastal Ecology and Forest Science
Department of Forestry and Environmental Conservation
Clemson University
Georgetown, SC, 29442, USA

* Co-Corresponding Authors

Corresponding Authors:

Charles A. Price
Ecology and Evolutionary Biology
University of Tennessee
Knoxville, TN, 37996-3140, USA
Email: cprice9@utk.edu
Phone: 865 974 9334

Benjamin Branoff
USDA Forest Service
Southern Research Station
Forest Inventory and Analysis
2730 Cherokee Farm Way
Suite 101
Knoxville, TN 37920
Email: Benjamin.Branoff@usda.gov
Phone: 865-862-2060

Abstract

Accurate estimation of mangrove biomass is a key component of the blue carbon economy, and allometric equations are a cornerstone of methodological approaches to quantify carbon from plot to landscape scales. A perennial question facing investigators is whether global allometric equations can be applied to their region of interest, or whether trees need to be harvested to provide accurate local estimates. Further, it is not always clear which combinations of tree measurements provide the most accurate estimates. We examined 16 different allometric equations for four tree measurements within a global data set of 710 trees for three common trans-Atlantic mangroves species. To determine the best performing models, we used mixed effects with site location and tree species modeled as random effects, and combinations of tree height, diameter at breast height (DBH), canopy diameter, and wood density as fixed effects resulting in 112 different models. The best performing model was a fixed effects model with three of four parameters (R^2=0.974, RMdSE=2.495) although many mixed effects models were comparable, and in general models with more parameters performed better even after penalizing model complexity. Variance partitioning and the conditional mode of univariate models indicate that site and species identity contribute relatively little to overall variance. These results provide qualified support for the use of global equations when high-quality local equations are not available. For example, a DBH only fixed effects model adds only 514 grams of median error relative to the best model (R^2=0.949, RMdSE=3.009). Our results can also help subsequent investigators prioritize data collection efforts, by highlighting data gaps, and variable combinations that yield the most accurate biomass estimation versus the time required to measure them.

Keywords: Allometry, biomass estimation, mangrove forests, mixed effects modelling

Introduction

Mangrove forests cover an area of approximately 135,000 km^2 across the globe (Leal and Spalding, 2022; Hamilton and Presotto, 2024). While a relatively small fraction of global forested land (40.6 million km^2; FAO, 2020), the outsized importance of mangroves to the blue (i.e., coastal) carbon economy has become increasingly recognized (Bertram et al., 2021; Qin et al., 2025). Due to the low oxygen levels in their soils, mangroves tend to accumulate detritus and sequester high levels of carbon below ground (Donato et al., 2011). In addition, mangroves serve as recruitment zones for a range of marine and terrestrial species (Laegdsgaard and Johnson, 2001; Mumby et al., 2004; Branoff and Campos-Cerqueira, 2021) and as storm buffers, attenuating wave impacts in areas with intense cyclonic activity (Alongi, 2008; Zhang et al., 2012).

The capacity of mangrove forests to provide these ecosystem services is strongly correlated with tree size and standing biomass (Stephenson et al., 2014; Pugh et al., 2019). In fact, biomass

estimation in forests worldwide is an area of intense interest and researchers have employed a range of approaches to estimate biomass in mangroves from traditional forest mensuration in the field, to inferring tree dimensions from remotely sensed data (typically tree height; Chave et al., 2005; Komiyama et al., 2008; Heumann, 2011; Navarro et al., 2020; Leal and Spalding, 2022).

Central to most biomass estimation efforts are empirically derived allometric equations which use harvested trees of varying size to relate standing biomass to more easily measured tree dimensions such as diameter at breast height (DBH), tree height, or canopy diameter (Komiyama et al., 2008; Rovai et al., 2021). Allometric functions are known to be influenced by the range of tree sizes examined, which for mature mangrove stands, varies considerably at geographic scales (Lugo and Snedaker, 1974; Rovai et al., 2021) and is strongly influenced by surface hydrology and soil nutrients (Cintron et al., 1978; Sherman et al., 2003); temperature, humidity and salinity (Osland et al., 2014; Devaney et al., 2021); and the intensity and frequency of disturbances (Simard et al., 2019; Krauss and Osland, 2020).

Due to the varying effects of these edaphic and climatic variables, a common practice is to harvest trees in natural areas of interest to support local biomass estimation, based on the prevailing view that locally derived allometric equations will outperform more general (global) equations developed using data collected elsewhere (Chave et al., 2005; Komiyama et al., 2008; Chave et al., 2014). In earlier work by our group, we examined the performance of local vs global equations in predicting biomass using bivariate allometric fixed effects models (Price et al., 2024), where global refers to models fit to data drawn from numerous sites throughout a given species' entire range, and local refers to data drawn from a single site or local area. We found that while some local equations had lower root mean squared errors (RMSE) than global equations, others did not, and the improvement over global equations was typically modest and strongly dependent on sample size. The application of local vs. regional/global allometry has long been a debate in forest mensuration with varying conclusions depending on site and system (Komiyama et al., 2005; Chave et al., 2014).

Univariate models (e.g. DBH or height) were used in this earlier work (Price et al., 2024) for several reasons. First, power laws of the form $ln(Y)=ln(a)+b\ ln(X)$ have a long history of theoretical and empirical use (Huxley, 1924; Schumacher and Hall, 1933), and estimating the slope (b) and intercept ($ln(a)$) of these relationships allows easy comparison to predictions from theoretical models (McMahon and Kronauer, 1976; West et al., 1999; Price et al., 2022), or values of $ln(a)$ and b derived from other empirical data fitting (Komiyama et al., 2008). Further, due to the challenges associated with measuring tree height, some researchers simply measure tree DBH, and DBH is often found to be the best single estimator of aboveground tree biomass (AGB) (Chave et al., 2005; Lima et al., 2012), with R^2 values frequently greater than 0.9 for mangroves (Komiyama et al., 2008; Price et al., 2024). Alternatively, studies based on remotely sensed data typically use tree height or canopy area to estimate AGB (Simard et al., 2006; Feliciano et al., 2017), thus the ability of height and/or canopy area to predict AGB is also of strong interest.

While our first analysis (Price et al., 2024) showed the dependence of model performance on sample size and the potential utility of single variate global equations, it did not consider the

performance of multivariate models and whether modelling site or species identity as random effects might improve model performance and identify key sources of variance. There have been many approaches to biomass estimation using regression models (Chave et al., 2005; Sileshi, 2014). While single variable models are common, if multiple predictors can improve model performance through more accurate prediction, their use is warranted. Usually one would need to account for model complexity with the increased number of parameters through the use of likelihood criterion such as Akaike information criterion (AIC) or Bayesian information criterion (BIC) (Burnham and Anderson, 1998), as well as for the potential interference of any collinear predictors via the variance inflation factor (VIF) (Menard, 1994) and whether the inclusion of random effects is warranted by examining of the Intraclass Correlation Coefficient (ICC) (Snijders and Bosker, 2012).

Here we build upon and expand the logic and approach of Chave et al. (2005) to compare the performance of a range of allometric equations in predicting aboveground biomass of the three most commonly found mangrove species throughout the Atlantic East Pacific biogeographic region (Spalding et al., 2010). We examine how field-measured variables such as tree DBH, height, and canopy diameter, with or without species wood density, perform in predicting AGB as single variables, and in a series of combinations as we describe below. We also incorporate geographic site location and species identity as random effects to determine the extent to which variance partitioning affects model choice (Feldpausch et al., 2011; Harrison et al., 2018; Cysneiros et al., 2021). We examine interspecific data groupings and use several summary statistics to inform our choice of the best models. This combined approach will allow us to address four main objectives:

- Evaluate the performance of multivariate allometric equations.
- Determine whether species and/or site effects significantly improve biomass prediction.
- Quantify the contribution of fixed and random effects to the total variance.
- Evaluate the extent to which a global DBH model is sufficient for mangrove biomass estimation.

Methods

Linking Geometry to Allometry

Superficially, a tree is a network of connected internodes approximated by a cylinder (more accurately a conical frustum). Based on this simple geometric reasoning, the volume (V) of the aboveground portion of a tree can be expressed as,

$$V = F \times \left(\frac{\pi D^2}{4}\right) \times H \quad (1)$$

where D is typically DBH and H is tree height (Cannell, 1984). The coefficient F is the form factor, which is meant to capture the taper or overall shape of a tree (Gray, 1956; Oluwajuwon et al., 2025). F is defined as the ratio of the volume of a tree's stem to that of a reference geometric shape, typically a cylinder. If a tree were an isometric cylinder, the form factor would equal 1.

The assumption that trees can be approximated as cylinders is similar to Yoda's pipe model (Shinozaki et al., 1964). Multiplying both sides of equation 1 by tree wood density (ρ) expresses the relationship in terms of mass,

$$AGB = F \times \rho \times \left(\frac{\pi D^2}{4}\right) \times H \quad (2).$$

Applying equation 2 to trees of varying size is problematic because it has long been established that the changes in *D* and *H* are allometric, meaning they change at different rates (Niklas, 1994). This is usually accounted for by expressing the proportional relationship between the two variables as $H \propto D^{\beta}$, where β is an exponent that captures the allometric change. Theoretical arguments based on elastic similarity predict that β will equal 2/3 (McMahon, 1973; McMahon and Kronauer, 1976), however both intraspecific and interspecific data compilations indicate that tree height vs DBH is frequently non-linear on logarithmic axes and often has exponents that differ from 2/3 (Niklas, 1995; Price and Schroeder, 2024).

Allowing for allometric change in equations for biomass prediction is typically represented by power law equations of the form,

$$AGB = cD^{\beta} \quad (3)$$

where *c* is a constant that effectively captures changes due to wood density or tree taper, and the exponent β reflects the change in tree shape with size. The composite function in equation 2 can similarly be represented as,

$$AGB = c(D^2H\rho)^{\beta} \quad \text{or} \quad AGB = c(C^2H\rho)^{\beta} \quad (4)$$

where *C* refers to canopy diameter, and both versions of Eq. 4 represent trees as volumes (area x length).

As tree growth is a multiplicative process (Gingerich, 2000; Kerkhoff and Enquist, 2009), data are usually examined on logarithmic axes thus equation 3 becomes,

$$log\ AGB = log\ c + \beta\ logD \quad (5)$$

These single variate equations perform reasonably well in predicting AGB of mangroves and are widely utilized. However, DBH or height alone does a poor job in capturing the overall size and shape of a tree, thus many researchers have utilized multiple variables to predict AGB (Schumacher and Hall, 1933), typically of the form,

$$log\ AGB = log\ c + \beta_1\ logD + \beta_2\ logH + \beta_3\ log\rho \quad (6)$$

Incorporating DBH, height and wood density into equation 6, may increase prediction accuracy by allowing for a multidimensional representation of the plant's form, and recognizing that large changes in wood density will impact the predicted biomass.

Two additional sources of variation are likely to impact model performance. First, site-based variability in growing conditions such as temperature, salinity, etc. are known to impact mangrove growth (Chapman et al., 2021; Ahmed et al., 2023). Second, while many species,

including the three species examined here, share distant (100+ mya) common ancestors (He et al., 2022), each species has its own unique evolutionary history and distinct adaptations, particularly with respect to surviving in unstable, saline soils (e.g., prop roots, pneumatophores, salt glands). Thus, species specific adaptations may lead to differences in allometric functions. Both site location and species identity can be incorporated into our model as random effects, which may change the scaling exponent and/or constant. For example, in the single variate case, equation 5 becomes,

$$log\ AGB\ = log\ c\ + b\ logD + \alpha_S + \beta_S\ logD + \gamma_L + \delta_L\ logD \quad (7)$$

where the subscripts α_S and β_S refer to the random effects of species on the scaling constant and exponent, respectively, and γ_L and δ_L refer to the random effects of site location on the scaling constant and exponent, respectively. Individuals of different species or under different growing conditions (sites), may have the same DBH, for example, but differ in their AGB. Random slopes/exponents and intercepts account for these differences when estimating the effect of diameter on AGB. Using our same example from above, this is representative of the "effect" or power that diameter has on AGB, where the AGB of individuals from different species or sites may respond differently in growth or biomass allocation to the same incremental change in diameter. In effect, by accounting for these sources of variability inherent within each species or site, we are acknowledging that we cannot define its source, thus we first separate the group specific variability (i.e. random effects) from the larger population before fitting for the fixed effects coefficients (slopes) on the residual variability. One goal of mixed effects modeling is to determine how important these random effects are relative to the classic and already well described "fixed" effects of DBH, height, etc. Equation 7 deals with only DBH as a predictive variable, but this same basic approach can be utilized with any number of variables.

Data and Statistics

Our data comes from two sources. The first is a recently published compilation of 590 individual trees from the three most common mangrove species in the Atlantic East Pacific biogeographic region (Price et al., 2024): *Avicennia germinans (L.)L., Rhizophora mangle L.,* and *Laguncularia racemosa (L.) C.F. Gaertn*. These data come from nine different sites across the region. The second dataset is from a recent mangrove harvesting effort in San Juan, Puerto Rico, which resulted in a new dataset of measurements for 120 trees across the three species (Price et al., 2025), hereafter labeled "San Juan" to distinguish it from previous mangrove studies in Puerto Rico (Golley et al., 1962). AGB and DBH were measured at all sites, height was measured at nine, and canopy diameter (hereafter just canopy) at seven of the ten total sites. Wood density data were not collected across sites, so we used data from San Juan. The citations and methods sections from all of the supporting studies are included in Appendix A.

To aid in the visualization of the overlapping data frequency distributions for the four variables (DBH, height, canopy, AGB), we employed a normal kernel smoothing function on log transformed data for each site. The smoothing function (ksdensity) was used in Matlab and

estimates density at 100 evenly spaced points based on each site's data (Bowman and Azzalini, 1997).

Using code developed in R (RCoreTeam, 2021), we examined allometric relationships for combinations of the four natural log-scaled variables: DBH, height, canopy, and wood density. We included these variables singly, as well as in groups of two, three, and all four variables as fixed effects. Site and species identity were modeled as random effects for the intercept, and combined slope and intercept of each regression equation, via the *lme4* utilities package (Bates et al., 2015). We have developed a companion GitHub repository (github.com/BBranoff/treellometry/tree/ecoinf) that includes extensive descriptions of our workflow and model implementation in R. Note, we lack site specific wood density information, so it was only incorporated as a species level model parameter in combination with the other dimensional variables (DBH, height, canopy). In previous work by our group (Price et al., 2024), our objective was to compare fitted allometric lines from multiple datasets to determine if their slopes and intercepts were different, for which standardized major axis regression (SMA) is recommended (Warton et al., 2006). Here our objective is predicting above ground biomass for which ordinary least squares (OLS) regression is the preferred choice (see Table 1, Warton et al., 2006).

In addition to standard fixed effects coefficients, we also extracted the random effects coefficients via the *ranef* function from the *lme4* package, as well as the component specific variance estimates via the *get_variance* function from the insight package (Lüdecke et al., 2019), and performed post hoc model performance comparisons via the *performance* package (Bates et al., 2015; Lüdecke et al., 2021), which extracts and compares the performance metrics discussed below. Though mixed effects models are often sensitive to disproportionately scaled predictors, our variables were all on the same scale after log transformation, thus no further scaling was performed.

With four variables (n), and without considering variable order, the number of possible combinations of variables chosen k at a time is given by the formula for the binomial coefficient, $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, which for k=1 is 4, k=2 is 6, k=3 is 4, and k=4 is 1, for a total of 15 combinations. As we lack site specific density values, we do not consider it alone as an explanatory variable which reduces our total to 14. We also include two composite models (DBH x height x density and canopy x height x density) bringing the total number to 16. We examined data with 3 effects levels: fixed effects, mixed intercepts, and mixed intercepts and slopes, at three mixed effects levels, site, species, and species and site together. This results in 7 total different effects levels which when multiplied by the number of models (16) results in 112 different model/effects combinations.

For each model we tested the underlying regression model assumptions (normality of residuals, normality of random effects, linear relationship, homogeneity of variance, multicollinearity) via the *check_model* function in the performance package (Lüdecke et al., 2021). We then used six different statistics to evaluate model performance at both fixed and mixed effects levels. The first is the coefficient of determination (R^2) which represents the amount of variance in the dependent

variable that is accounted for by the independent variables in a regression model. Traditionally, R^2 only applies to fixed effects but new methods have been developed recently to calculate such a metric for random effects as well (Nakagawa et al., 2017). We employed these methods to calculate the conditional R^2 (the variance explained by both fixed and random effects) via the *MuMIn* package in R (Bartoń, 2013).

The second and third statistics are Akaike's information criterion (AIC), and Bayesian information criterion (BIC). Both AIC and BIC inform model selection by incorporating penalties for the number of parameters. AIC penalizes the model based on the number of parameters, while BIC is based on the number of parameters and sample size (Burnham and Anderson, 1998).

Our fourth statistic is the root mean squared error (RMSE). RMSE represents the square root of the average of the squared differences between the predicted and actual values. Because our models were performed on log-scaled data, the RMSE is also on the log scale, which can be more difficult to interpret in terms of raw allometric scale. Thus, to aid error contextualization, we back transformed individual predictions using a Baskerville correction and then recalculated the RMSE on linear scaled residuals (Baskerville, 1972). While both the log and linear scale RMSE values are reported, the linear scaled values were used in performance ranking and are referred to in the text to contextualize model errors. We also used this back-transformed RMSE value divided by the mean AGB values to calculate the relative root mean squared error (RRMSE) which we multiplied by 100 to express as a percentage.

On the linear scale, the distribution of AGB values in our harvest dataset is heavily right skewed due to the relatively few large trees collected. The squared linear errors from these trees are thus disproportionately high, resulting in inflated RMSE values that don't necessarily represent the majority of observations. To account for this, we also calculated the root median squared error (RMdSE, our fifth statistic) on the back transformed residuals, which provides an alternative view of model performance on individual observations. Our sixth statistic is sigma, which represents the standard deviation of the errors or residuals, and, although similar to RMSE, sigma incorporates the model's degrees of freedom.

In addition, for mixed effects models we also calculated the intraclass correlation coefficient (ICC) which quantifies the proportion of the variance that is attributable to group level differences. ICC is calculated via the *performance* package (Lüdecke et al., 2021) as the ratio of the variance due to random effects (between group variance), to the total variance (sum of within and between group variance) (Hox et al., 2010). Previous work has suggested that the use of mixed effects models is not warranted when the ICC is less than 0.1 (Snijders and Bosker, 2012; Hox et al., 2017). Note that if there is insufficient data to calculate the ICC, our code returns NA, which also means the use of mixed effects is not warranted given the limited data. Models with low or NA ICC values were dropped from comparisons with the other models.

When parameter values are highly correlated, as they inherently are in allometric studies, one risks introducing collinearity into mixed effects models which can make it difficult to determine individual predictor effects and potentially inflate standard errors for regression coefficients. To

test for collinearity, we calculated the variance inflation factor (VIF) for the parameters in each model. The VIF for any independent variable in a model is given by $1/(1-R_v^2{}_i)$, where i represents the i^{th} independent variable, and the R_v^2 value is determined by regressing the variable in question against all other independent variables in the model. For a model with two variables the VIF values are identical for each variable. However, with three or more independent variables, each variable will have its own VIF. By convention, for a given model the maximum VIF across all independent variables is the one that is reported (Hair et al., 2022).

The validity of models with large VIF values can be a concern and some authors have suggested that models with VIF values >5 warrant further investigation and values >10 indicate collinearity that needs to be addressed (Menard, 1994). However, the use of "rules of thumb" or cutoff values to determine model validity has been questioned (O'brien, 2007). Rather than accepting or rejecting any model based on the VIF values of its parameters, we simply report the VIF values and discuss future methodological improvements that might help to reduce the influence of collinearity. One of the recommended approaches to deal with collinearity is to use functions with composite variables as in Eq. 4 (Hair et al., 2022). This is one of the reasons we also examined the two composite functions as discussed above.

In mixed effects models with multiple fixed and random effects, cross validation is frequently used to determine if model fits can be generalized to novel data. We used the *cv* function from the cv package in R (Fox and Monette, 2025) with 10 data folds and calculated the mean and standard deviation of the log scale RMSE across folds at fixed, mixed intercept, and mixed intercept and slope levels. This function creates 10 equally sized subsets (folds) of the data, repeatedly performs model fitting on the data minus each fold, and reports the resulting log scale RMSE values as their mean and standard deviation. We plotted the log RMSE from each model against the mean cross-validated log RMSE for the same model for all model types to demonstrate how stable model results were and thus how useful each model might be at predicting unseen data. For each model we also checked if the variance-covariance matrix was singular using the *isSingular* function in *lme4*. Singularity in a mixed model may indicate that the variance in some random effects is effectively zero or that the correlation is exactly ± 1. Singular models were dropped from comparison when detected.

One of our primary objectives is to compare model performance among models of differing complexity. Not all measurements were collected by all researchers, and some measures don't apply to all trees (e.g. some trees were too small to have DBH measures) leading to differences in sample size depending on the measurement in question. For example, there are 503 trees that have both DBH and AGB measures. In contrast there are 611 trees with height and AGB measures. Comparing models conditioned on different sized datasets is problematic, thus for the purposes of ranking model performance, we utilized only those trees that have all four measures (DBH, height, canopy, wood density), resulting in a subset of the full dataset containing 339 trees.

To help identify the best performing model across all variable combinations and modeling types, and the best model in each variable group using this data subset, we ranked each of our six performance statistics (R^2, AIC, BIC, RMSE, RMdSE, Sigma), from best to worst. The mean of

all six statistics' rankings was then calculated and the models assigned a global rank and sorted from best to worst (Tables S1-2).

Another of our primary objectives is to publish model coefficients for the different models so researchers from other groups can compare them to their own or use them for biomass estimation. Coefficients for models conditioned on larger sample sizes will tend to be more accurate in biomass prediction, thus we include model parameters for the full dataset (Table S3), irrespective of whether sample sizes were different for different parameter combinations.

Analyses of allometric patterns in tropical trees have focused on trees ≥5 cm DBH (Chave et al., 2005; Chave et al., 2014), as those trees constitute the majority of community level biomass in many tropical forests, and trees smaller than 5 cm DBH are not included in some site level allometric collections (Price et al., 2024). Thus, to aid in comparison with the findings (equations) from studies focused on tropical forests, we have also run our mixed effects analysis on a subset of our data containing only trees ≥5 cm DBH (204 trees with all four measurements, Table S2)). Note also that a number of trees from the Biscayne, Louisiana, and San Juan sites (206 trees in total), were too small for DBH measures and have corresponding D30 (diameter at 30 cm above ground) or basal diameter measures. Data from these trees were not included in models using DBH but were included in models using height or canopy diameter.

Results

Our combined dataset contained 710 individuals across our three focal mangrove species. Site locations are shown in Fig. 1. Sample sizes for each species and the range of measurements within the global dataset are summarized in Table 1. Table 2 contains the site level minimum, maximum, mean, and range for each of the four structural variables we examined (AGB, DBH, Height, Canopy), as well as sample sizes for each species. Sample sizes for each measurement variable were DBH=504, height=612, canopy=546 and AGB=709; The skewness for each variable was 1.67, 0.8, 1.36, and 6.81, respectively.

Probability density function (PDF) estimates of log scaled data frequency distributions for biomass, DBH, height and canopy for all 10 sites are shown in Fig. 2.

Results from the checks of model assumptions are presented as a series of pdf files, one for each model, on our companion Github page (https://github.com/BBranoff/treellometry/tree/ecoinf/Assumptions). Across all model-plot combinations, the data largely supported the regression model assumptions, although there was a modest increase in positive residuals at the small end of the size spectrum in some plots, and modest trend in negative residuals at the large end in others.

Examining all combinations of variables at fixed effects and three mixed effects levels (species & site, species, site) yields 112 different models. All non-singular fixed effects models, and mixed effects models with ICC values > 0.1 were included in the final global ranking based on R^2, AIC, BIC, RMSE, RMdSE, and Sigma(Table S1). Following removal of models that were singular, or had ICC values < 0.1 resulted in 45 models for the entire dataset (Table S1) and 48

models for the ≥5 cm subset (Table S2). For each variable group (combination of variables), all valid models are depicted in Fig. S1 and the highest ranked model from each variable group is described in Table 3.

The influence of the number of predictors on our six main summary statistics is shown in Fig. 3 for both fixed effects, and mixed effects models. Models with more predictors tend to have higher R^2 values and lower RMSE, RMdSE, AIC, BIC, and Sigma values. Also, some models with fewer predictors (1 or 2) have summary statistics that are close in value to models with more predictors.

The conditional mode of random effects for all sites and species and for each predictive variable in univariate models (minus that for wood density, given the lack of site-specific values) are depicted in Fig. 4. In general, site level effects were larger than species, especially for models utilizing height and canopy diameter, while the DBH model reflected relatively little random effects.

Partitioning the variance in aboveground biomass to fixed effects, random effects, and residuals yielded means and standard errors for each component among all models (Fig. 5), as well as those for each variable in each model family in a more detailed representation (Fig S2-3). Total estimated explained variance was at least 80% (mean = 93%), and fixed effects dominate at about 91% of explained variance on average, accounting for the vast majority (Fig 5A). With fixed effects removed (Fig. 5B), it is easier to compare the random effects, which account for 2.4% of the total variance on average and as in Fig 4, show site level variance components generally account for almost all of the random variance, especially in models utilizing height.

Plots of RMSE with and without k-fold cross validation indicate that fixed and mixed-effectsmodels with ICC values >0.1 (Fig. 6) have good predictive power and can be generalized to unseen data from the same population (e.g. the same sites and species).

Model performance varied according to the performance metric (Figs. 3, 4, S1), but more complex models (i.e. more predictors) often performed best, even in metrics that penalize additional parameters. Many of the top performing models included site-level or site- and species-level random effects on both intercept and slope, while species-only random effects typically performed similarly to many of the fixed-effects only models.

Of the 504 trees with DBH measures, 265 of those were from trees with DBH ≥ 5 cm. Of those 87 (33%) were *A. germinans*, 101 (38%) were *L. racemosa*, and 77 (29%) were *R. mangle*, thus a relatively even sampling across species. Table S3 contains the ranked fixed effects models only for the entire dataset, and for models fit to trees with DBH ≥ 5 cm. Not surprisingly, the model coefficients from the two data groupings (all and ≥ 5 cm) differ, however the model rankings across the two data groupings are largely preserved (Table S3). Comparing the rank of models retained for the entire dataset to the same models in the ≥ 5 cm subset yielded a Pearson rank correlation of 0.954.

Discussion

The development of accurate equations to estimate mangrove biomass can support numerous areas of scientific inquiry and management. Both field-based and remote measurements (height, canopy) using drones, planes or satellites rely on allometric equations to estimate biomass, from individual trees to landscape level (Simard et al., 2006; Heumann, 2011; Hu et al., 2020). Such efforts also underpin carbon stock assessments to comply with national or international guidelines (Arnold et al., 2023; Bernardino et al., 2024).

A central question within the field has been whether locally derived equations are necessary, or whether mass estimates derived from global equations are sufficient (Chave et al., 2005; Komiyama et al., 2005; Price et al., 2024). Most studies that have harvested these species, and of which our global data compilation is comprised, have operated under the assumption that local equations are necessary (Fromard et al., 1998; Soares and Schaeffer-Novelli, 2005; Smith and Whelan, 2006; Yepes et al., 2016; Zanvo et al., 2023; Price et al., 2025). Clearly, one cannot even address the question of whether local or global equations are required without a dataset comprised of numerous sites, and species-specific allometries. Thus, at least initially, local data and equations are necessary. However, as the collective sample size increases, one can begin to ask if there are conditions under which the use of global equations is warranted, and whether more data are required to inform the issue.

The most commonly cited paper for equations to estimate mangrove biomass within the Atlantic East Pacific biogeographic region is Chave *et al*. (2005), who combined mangrove data from two papers (Imbert and Rollet, 1989; Fromard et al., 1998) totaling 149 trees across three species (*L. racemosa, A. germinans* and *R. mangle*). At the time of its publication, the number of trees and sites represented in Chave et al. (2005) was larger than previous analyses. Combining more recent data from (Price et al., 2024) and (Price et al., 2025) yields 710 trees for *L. racemosa* (239), *A. germinans* (255) and *R. mangle* (216), over a four-fold increase in sample size compared to Chave et al. (2005) and increases the number of sites represented from two to ten. Note, the data from both studies analyzed by Chave et al. (2005) are included in our global compilation.

The summary statistics contained within Tables 3 and S1-2 indicate that in general, mixed models that allow for site and species differences, often outperform global, fixed effects equations, particularly for the ≥5 cm subset. However, the best overall model was the one that incorporated all four measured variables as fixed effects (Tables 3, S1). In addition, the two composite models we examined also tended to perform well (Tables 3, S1-2). However, like the results from previous work (Price et al., 2024; Price et al., 2025), the increase in prediction accuracy over some simpler models is rather modest. For example, looking at Table 3 and focusing solely on RMSE, the difference between the RMSE of the best performing model and the RMSE of several other highly ranked models was not large. The difference between models in RMdSE is even more modest, where the difference between the best model (2.495 kg) and a fixed effects model using DBH only (3.009 kg) was 514 grams (Table 3).

Despite their strong performance, the ICC values were frequently low for the mixed effects models with more variables (Tables 3, S1-2). For example, the best overall mixed intercept model (Table 3) had an ICC value of 0.16, meaning that 16% of the variance could be attributed

to between site or species differences. Moreover, some models, including a large proportion of mixed intercept and slope models, returned a value of NA which means that the model in question is singular, or that the random effect variance for that variable grouping is effectively zero. This can be due to a lack of variance, overfitting with a random effects structure that is too complex, or sample sizes that are insufficient to detect between group variances relative to the total variance. Subsequent work with larger datasets could help to determine which of these factors most influence ICC, and the extent to which mixed effects models are justified, although the cost-benefit of cutting down protected mangrove trees would need to be considered if the cutting is purely for model refinement and not concurrent with other societal needs (e.g. Price et al., 2025).

Additionally, the VIF values of some of our top performing models indicate significant collinearity. This was not true for all models, but many of the models with lower VIF values had ICC values that indicated poor model performance. As mentioned previously, one way to deal with high levels of collinearity is using composite variables, and our two composite models performed well overall, particularly the product of DBH, height, and wood density (Table 3). However, composite models still require more field data collection, and the improvement over simple models (DBH or height, fixed effects) is not substantial.

While we have utilized the mean rank of six individually ranked performance metrics, it is somewhat arbitrary to assign equal weight to each. We have included a global rank largely for illustrative purposes, and the rankings of our six performance metrics are strongly correlated with one another (Table S4), so examining their mean does give a sense of their performance across these metrics. However, we caution against overreliance on this single value in judging models. Performance metrics such as AIC or BIC are widely used in model selection and weigh model fit against model complexity, so rankings based on those metrics may be sufficient for many researchers (Table S1). However, if accuracy with minimal error is more highly valued, model complexity may not be a concern, and thus rankings based on explained variance (R^2), prediction accuracy (RMSE or RMdSE), or prediction error (sigma) may be more valued.

Variance decomposition (Fig. 5) demonstrates that the overwhelming majority of variance (>90%) is due to fixed effects (DBH, height, canopy, density). This is not surprising in an analysis of allometric trends as the data and models, by design, are evaluating change due to variability in tree size. Fig. 5B also shows that the percent of total variance attributed to site differences (mean = 4.13%) was an order of magnitude greater than that at the species level (mean = 0.27%). Quantifying this proportional variability reinforces our main result that the field measured linear dimensions drive most of the variability in AGB, and while random effects can improve model performance, their overall contribution is small, around 1.9% from this study and mostly only greater than this for site level differences in models utilizing height.

The mean RMSE from k-fold cross validation for each model is almost identical to that from the entire data set in all cases (Fig. 6, Table S1). In fixed effects, mixed intercept, and mixed intercept and slope models the points all fall along the 1:1 line in both the entire data set (6A) and for the ≥5 cm subset (6B). This indicates that these models are robust and could be used for biomass prediction for this collection of sites and species. However, for researchers wanting to

predict AGB of trees in new sites, it is advisable to use coefficients from the fixed effects models (Table S3).

The two composite models we examined both performed relatively well in comparison to other models. Some of the improved performance is likely attributable to the fact that a composite variable that is a product of three variables functions as a single variable, which will tend to lower its AIC/BIC scores. As described in the introduction, being the product of an area times a length times density, the relationship between the composite variables and the mass they are attempting to predict is conceptually straightforward. However, the assumption that trees can be approximated by a series of cylinders, or a single large cylinder, is not entirely consistent with actual tree form. Exhaustive descriptions of the dimensions and architecture of entire trees are rare. With the advent of terrestrial laser scanning, quantitative structural models fit to point clouds obtained from real trees are getting closer, but methodological concerns such as over fitting small branches, or fitting cylinders to branches that are not truly cylindrical still leave a gap in our understanding of tree form (Calders et al., 2015).

Mangroves often grow in regions subject to periodic disturbance, and many of the sites in our dataset experience disturbance in the form of seasonal cyclones (hurricanes) which can lead to tree mortality or significant loss of branches and foliage. This allows light to penetrate further into the canopy; thus, plants can invest more in growing outwards rather than upwards. As canopies close, competition for light requires trees to invest more in growing upwards. These general trends may explain why tree height and canopy diameter exhibit more plasticity in random effects (Fig. 4). In contrast, DBH represents the amount of conductive, storage, and biomechanical tissue required to support the mass of the tree above, and to link it with below ground nutrient acquisition. As an inherent outcome of these physical interdependencies, DBH appears to exhibit much lower site or species level variability (Fig. 4), consistent with what other groups have found (Ter-Mikaelian and Korzukhin, 1997; Jenkins et al., 2003; Chave et al., 2014).

The variability in allometry across sites, however minimal, does seem to be significant for one or more predictor variables. Trees at the Benin, Biscayne, Guadeloupe, Louisiana and Puerto Rico sites exhibited the greatest random effects on the intercept for models involving height (Fig. 4). This may be due in part to the fact that data collected at these sites are found at the small (Louisiana) or large (Guadeloupe, Benin) end of the size spectrum (Fig. 2, Table 2). For Biscayne and Louisiana, the datasets are comprised of smaller mangroves and thus this deviation is unavoidable. For Guadeloupe and Benin, this highlights the value in collecting trees over the entire size range when harvesting samples to develop allometric equations (Price et al., 2024; Price et al., 2025). Data from Louisiana and Puerto Rico also segregated in Figs. 2 and 4, but their influence on the conditional means of random effects was weaker because only a single species was collected at each site. These relatively large random effects for height, and to a lesser extent for canopy, suggest that the biomass relationship with these predictors is more dependent on site. In contrast, random effects associated with DBH were minimal to zero, suggesting tree biomass responds equally to this metric regardless of site. The same can be said for all predictors in regard to species level random effects, which seem to have little influence on the models.

Thus, species or site level variability in allometry and predictive error would be relatively low when using a global diameter-based model but would increase when using height or canopy across multiple sites.

Only one of the studies we compiled reported locally measured wood density values for one of our species (*A. germinans*, Zanvo et al., 2023), thus we used values from a recent data collection effort by member of our group (Price et al., 2025). The values obtained in this work (Table 1) are similar to the values reported by other researchers for these species (Quadros and Zimmer, 2017; Virgulino-Júnior et al., 2020). Mangrove wood density has been shown to vary modestly, but systematically with tree size and soil salinity (Schmitz et al., 2006; Virgulino-Júnior et al., 2020), thus it is possible that with detailed site and size specific information, the influence of wood density on model predictions would change. Future efforts may uncover the extent to which it varies within and across environmental and tree size gradients.

Several groups have examined what constitutes a sufficient sample size in tree allometry, with data informed, recommended minimum sample sizes ranging from 17 to 166 individuals depending on the tree size range and underlying variability (Chave et al., 2004; Roxburgh et al., 2015). Trees in locations with large size ranges will tend to require larger sample sizes and Chave et al. (2004) recommend harvesting at least 100 trees for species in tropical forests. Over much of their range, the mangrove species we have examined are much shorter in stature than tropical forest trees, so smaller sample sizes may be adequate in some areas, but some of our site-specific sample sizes are likely too small (Table 2). Moreover, mangroves can obtain very large sizes, with DBH measures reaching 96.7 cm and height 57 m in the tallest forests (Castellanos-Galindo et al., 2021). We are unaware of datasets that contain harvested trees near such sizes.

In assessing the influence of size range and sample size of tropical tree allometry, Chave et al. (2004) recommended the use of “regional or pan-tropical compilations” over site-specific models based on small sample sizes. We agree with this general conclusion which would highlight the need to expand harvesting of mangroves, so a greater range of sites and size ranges are represented in allometric data compilations. Due to logistical or conservation constraints, mangroves are particularly challenging to harvest. However, trees from the larger end of the size spectrum are underrepresented in our compiled dataset and empirical data from the forests that contain such trees would likely improve model estimates. Few would likely relish cutting down such majestic trees, and alternatives such as quantitative structural models derived from terrestrial laser scanning (TLS) might obviate the need for harvesting (Calders et al., 2015; Demol et al., 2022).

With sufficient time and resources, and minimal concerns for the impact of harvesting trees, researchers interested in the best biomass estimate should collect all three dimensional variables (DBH, height, canopy diameter), wood density, and harvest a sufficient number of trees per species (Chave et al., 2004; >50, Paul et al., 2018) to generate local allometric equations. While our sample sizes and the number of sites are larger than any previous analyses, the question of local vs global equations could still be informed by more data and we cannot rule out the possibility that increases in sample size will increase the variance attributable to species or site

level differences. This is especially true for sites that exhibit conditions far outside of the "normal" growing conditions, such as those at the latitudinal limits of the species (Louisiana data, Osland et al., 2014) or where nutrient or salinity concentrations result in anomalous growth (e.g. dwarfism) (Biscayne data, Ross et al., 2001).

Our global dataset is comprised of just three species because, for much of their range, they are the only three species found (Spalding et al., 2010). Moreover, throughout their entire range they are often (but not always) the most common species and comprise a substantial portion of the mangrove community composition both in individual numbers and mass. Additional data from species that are somewhat common within the Atlantic East Pacific biogeographic region (*Avicennia schaueriana, Rhizophora racemosa, Rhizophora harrisonii, Conocarpus erectus*) could further inform this issue.

Overall, our work offers support for using global interspecific fixed effect equations. Among commonly collected field-based measurements, the easiest variable to measure rapidly is DBH. Given the considerable additional time it takes to measure tree height and/or canopy diameter in the field, and the error associated with these measures, we suspect that many researchers would be willing to accept the additional level of error from simply measuring DBH. Measuring tree height using height poles or hypsometers is notoriously difficult in some mangrove forests because the high leaf and stem density and canopy overlap between trees makes seeing and identifying the highest tree branch quite challenging. This is in addition to the fact that navigating and even standing within some mangrove forests can be a challenge due to the presence of large overlapping prop roots, unstable soils/mud, or the presence of dangerous fauna.

For researchers wanting to use equations from our analyses to predict biomass in their system of interest, in the absence of site specific or regional equations, we recommend using the fixed effects models as the parameter values (slope(s), intercept) for the mixed effects models in Tables 3, S1-2 are specific to this particular collection of species and sites and may not be applicable to other sites or species aggregations. Moreover, the variance attributable to random effects is not large. To that end we have compiled all the fixed effects models into a single Table S3 for models fit to the entire dataset, and models fit only to trees ≥ 5 cm DBH.

For land managers interested in using forest mensuration to estimate stand level biomass for blue carbon accounting (i.e. REDD+), evaluating the time required to measure trees using one measurement (DBH) vs multiple measurements (DBH, height, canopy, density) should be considered. There may be instances where the speed with which DBH can be measured on a larger plot area might take precedence. Moreover, the relatively good performance of height as a single predictor may support the use of remote sensing approaches (lidar, photogrammetry) that yield height estimates from individual tree detection, or canopy height models. Our data indicates that the development of site specific allometric models will typically, but modestly, outperform global equations. In instances where developing site-specific equations is not possible or desirable, global or regional equations should be sufficient, particularly if tree sizes are within the range of the data presented here.

Summary

Investigators interested in predicting the biomass of mangroves from local to global scales seek to develop allometric equations that provide the most accurate estimates. While equations based on locally harvested data with adequate sample sizes will likely yield accurate estimates, the improvement over global models may be modest. If harvesting trees is not preferred, our work suggests that measuring the DBH, height, canopy diameter, and wood density of individual species will yield the most accurate mangrove biomass estimates. If time or resources further limit field work, or if investigators want to prioritize the number of trees measured over the number of different measurements for each tree, then our work indicates that simply measuring the DBH (or height) of each tree will modestly increase the median error of biomass estimates. While our dataset is the largest yet used to address such questions in mangroves, with over 700 trees across the three species from 10 different sites, confirmation of these patterns with data from more sites and species, and particularly more large trees, will help to determine the extent to which global or local equations provide the best information.

Figure captions

Figure 1. Plot locations for the 10 studies considered in our analysis. Location symbol color and the color bar correspond the maximum single tree ABG (kg) at each site.

Figure 2. Probability density function (PDF) estimates from kernel smoothing. Note that the data from the sites with small individuals (Biscayne, Louisiana) segregate toward the left side of each panel and data from sites that collected large individuals (Benin, Guadeloupe), the right.

Figure 3. Bivariate plots of our seven summary statistics (Panels A-F, R^2, AIC, BIC, Sigma, linear RMSE, and linear RMdSE), as a function of the number of predictors. The different modelling types; fixed effects, mixed intercept and mixed intercept and slope, are denoted with blue, red and green circles, respectively. The symbols for each number of predictors are offset slightly on the x-axis to increase visibility, e.g. all of the symbols on the far left of Panel A have just a single predictor. Note that while models with more predictors tend to perform better (higher R^2, lower everything else), many of the models with 1 or 2 predictors also perform reasonably well. Please consult Tables 3, S1-2 for specific model details.

Figure 4. Conditional means of random effects for single variable models with full random effects (random slopes and intercepts on both species and site). Significant random effects are inferred from means that deviate from zero more than their corresponding error bars. Relatively large site level intercept effects are most apparent for height, while random effects on DBH and species are relatively low. Species present at each site are noted parenthetically using first letter of each binomial next to the site names.

Figure 5. Variance decomposition. Partitioning the variance in aboveground biomass to fixed effects, random effects, and residuals; means and standard errors for each component across all models. Total estimated explained variance is at least 90% on average, and fixed effects dominate, accounting for the vast majority of variance (Fig 5A). Fixed effects removed (Fig. 5B) to compare the random effects, which account for no more than roughly 7% of the variance. This corroborates the view that species and location level information add relatively little to the variability in allometric relationships we are examining. The same plots for each model family are provided in Figs S2-3.

Figure 6. Bivariate plots of the mean standardized RMSE from k-fold cross validation vs. the standardized RMSE for the same model without cross validation (Table S1) for fixed (red circles), mixed intercept (green circles) and mixed intercept and slope models (red circles). Panel A contains values for the full data set and Panel B those trees ≥ 5 cm DBH. For both datasets there is strong agreement between the log RMSE from each model and that companion mean log RMSE from k-fold cross validation as all points fall roughly along the black 1:1 line.

Supplementary Material

Figure S1. Radar charts of model performance metrics for each group of parameters show which combination of mixed-effects resulted in the best model performance (Table 3). The seven axes

in each figure correspond to the seven summary statistics presented in Table 2 (R^2, Sigma, AIC, BIC, RMSE, RMdSE and RMSE.CVsd). These metrics are in the same position in each panel as in the top left. Scores are scaled such that higher values, and thus larger polygons, represent the best models.

Figure S2. Variance decomposition for the 16 different model groups. As with the across model comparison (Fig. 5), fixed effects dominate the variance in a model types. The variance attributable to fixed effects exceeds 80% in all models.

Figure S3. Variance decomposition for the 16 different model groups with fixed effects removed. Removing the variance attributable makes it easier to see that contributed by random effects, which is very small. The largest variance component observed was the model with height as a single predictor variable, within which, the intercept term accounted for ~7.5% of the total variance.

Appendix A. List of the articles from which our data compilation is comprised along with the relevant Methods section for each.

Appendix B. Glossary of column header terms for Tables S1-2.

Table S1. Summary statistics for the 112 different allometric models. Detailed descriptions of the column headers are given in Appendix B.

Table S2. Summary statistics for the mixed effects models grouped by model type. Detailed descriptions of the column headers are given in Appendix B.

Table S3. Fixed effects models for the entire dataset (size = all) and for trees ≥ 5 cm DBH (Data size = ≥ 5cm). The models within each data grouping (All or ≥ 5cm) are sorted by their rank among the fixed effects models.

Table S4. Pearson correlation coefficients for summary statistics. We calculated the Pearson correlation coefficient pairwise between each of our summary statistics. The top six rows correspond to the entire dataset (all) and the bottom six for the ≥ 5 cm DBH subset. Note the correlations are generally high.

Acknowledgements

We would like to thank Aerostar Holdings LLC for granting access to the San Juan site and for assisting with security clearances and other logistics. Jaime Pabón (Aerostar) provided logistical support and Marisleissis Orona (Aerostar) provided daily escort, supervision of onsite activities, and helped with data acquisition and recording. Walter Soler (Ambienta) assisted with logistical support including the hiring of our chainsaw operator, Alejandro Rodríguez-Rojas (Econet). Ixia Avilés-Vázquez and Luis Ortiz-López from U.S. Forest Service, Forest Inventory and Analysis and Nicole Pillot (Boricorps) and Morgan Chaudry (Clemson University) also helped with data collection. Boricorps and the Jobos Bay National Estuarine Research Reserve in Puerto Rico provided field and lab support. This is Technical Contribution 7552 of the Clemson University Experimental Station. Funding for SJVB provided by USDA NIFA McIntire-Stennis project

award no. SC-1700647. This manuscript was prepared in part by U.S. government employees, therefore it is considered in the public domain and not subject to copyright.

Data Availability Statement

All data and code in support of this manuscript are available at:
https://github.com/BBranoff/treellometry

References

Ahmed, S., Sarker, S.K., Friess, D.A., Kamruzzaman, M., Jacobs, M., Sillanpää, M., Naabeh, C.S.S., Pretzsch, H., 2023. Mangrove tree growth is size-dependent across a large-scale salinity gradient. Forest Ecology and Management 537, 120954.
Alongi, D.M., 2008. Mangrove forests: Resilience, protection from tsunamis, and responses to global climate change. Estuarine Coastal and Shelf Science 76, 1-13.
Arnold, F.E., Piazza, M., Wynn, K.Z., Htut, P., Tun, S.S., 2023. Mangrove biomass and carbon estimates for REDD plus from national forest inventory in two regions of Myanmar. International Forestry Review 25, 283-299.
Bartoń, K., 2013. MuMIn: Multi-model inference.
Baskerville, G.L., 1972. Use of Logarithmic Regression in the Estimation of Plant Biomass. Canadian Journal of Forest Research 2, 149-149.
Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67, 1 - 48.
Bernardino, A.F., Mazzuco, A.C.A., Costa, R.F., Souza, F., Owuor, M.A., Nobrega, G.N., Sanders, C.J., Ferreira, T.O., Kauffman, J.B., 2024. The inclusion of Amazon mangroves in Brazil's REDD+ program. Nature Communications 15, 1549.
Bertram, C., Quaas, M., Reusch, T.B.H., Vafeidis, A.T., Wolff, C., Rickels, W., 2021. The blue carbon wealth of nations. Nature Climate Change 11, 704-709.
Bowman, A.W., Azzalini, A., 1997. Applied smoothing techniques for data analysis. Oxford University Press Inc, New York.
Branoff, B.L., Campos-Cerqueira, M., 2021. The Role of Urbanness, Vegetation Structure, and Scale in Shaping Puerto Rico's Acoustically Active Mangrove Fauna Communities. Frontiers in Marine Science Volume 8 - 2021.
Burnham, K.P., Anderson, D.R., 1998. Model Selection and Inference: A Practical Information-Theoretic Approach. Springer New York, New York, NY.
Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P., Herold, M., Culvenor, D., Avitabile, V., Disney, M., Armston, J., Kaasalainen, M., 2015. Nondestructive estimates of above-ground biomass using terrestrial laser scanning. Methods in Ecology and Evolution 6, 198-208.
Cannell, M.G.R., 1984. Woody biomass of forest stands. Forest Ecology and Management 8, 299-312.
Castellanos-Galindo, G.A., Casella, E., Tavera, H., Zapata Padilla, L.A., Simard, M., 2021. Structural Characteristics of the Tallest Mangrove Forests of the American Continent: A Comparison of Ground-Based, Drone and Radar Measurements. Frontiers in Forests and Global Change Volume 4 - 2021.
Chapman, S.K., Feller, I.C., Canas, G., Hayes, M.A., Dix, N., Hester, M., Morris, J., Langley, J.A., 2021. Mangrove growth response to experimental warming is greatest near the range limit in northeast Florida. Ecology 102, e03320.

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., Yamakura, T., 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. Oecologia 145, 87-99.
Chave, J., Condit, R., Aguilar, S., Hernandez, A., Lao, S., Perez, R., 2004. Error propagation and scaling for tropical forest biomass estimates. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 359, 409-420.
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., Vieilledent, G., 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. Global Change Biology 20, 3177-3190.
Cintron, G., Lugo, A.E., Pool, D.J., Morris, G., 1978. Mangroves of Arid Environments in Puerto Rico and Adjacent Islands. Biotropica 10, 110-121.
Cysneiros, V.C., Coelho de Souza, F., Gaui, T.D., Pelissari, A.L., Orso, G.A., Machado, S.d.A., de Carvalho, D.C., Silveira-Filho, T.B., 2021. Integrating climate, soil and stand structure into allometric models: An approach of site-effects on tree allometry in Atlantic Forest. Ecological Indicators 127, 107794.
Demol, M., Verbeeck, H., Gielen, B., Armston, J., Burt, A., Disney, M., Duncanson, L., Hackenberg, J., Kükenbrink, D., Lau, A., Ploton, P., Sewdien, A., Stovall, A., Takoudjou, S.M., Volkova, L., Weston, C., Wortel, V., Calders, K., 2022. Estimating forest above-ground biomass with terrestrial laser scanning: Current status and future directions. Methods in Ecology and Evolution 13, 1628-1639.
Devaney, J.L., Pullen, J., Feller, I.C., Parker, J.D., 2021. Low humidity and hypersalinity reduce cold tolerance in mangroves. Estuarine, Coastal and Shelf Science 248, 107015.
Donato, D.C., Kauffman, J.B., Murdiyarso, D., Kurnianto, S., Stidham, M., Kanninen, M., 2011. Mangroves among the most carbon-rich forests in the tropics. Nature Geoscience 4, 293-297.
FAO, 2020. Global Forest Resources Assessment 2020, Main report, Rome.
Feldpausch, T.R., Banin, L., Phillips, O.L., Baker, T.R., Lewis, S.L., Quesada, C.A., Affum-Baffoe, K., Arets, E.J.M.M., Berry, N.J., Bird, M., Brondizio, E.S., de Camargo, P., Chave, J., Djagbletey, G., Domingues, T.F., Drescher, M., Fearnside, P.M., França, M.B., Fyllas, N.M., Lopez-Gonzalez, G., Hladik, A., Higuchi, N., Hunter, M.O., Iida, Y., Salim, K.A., Kassim, A.R., Keller, M., Kemp, J., King, D.A., Lovett, J.C., Marimon, B.S., Marimon-Junior, B.H., Lenza, E., Marshall, A.R., Metcalfe, D.J., Mitchard, E.T.A., Moran, E.F., Nelson, B.W., Nilus, R., Nogueira, E.M., Palace, M., Patiño, S., Peh, K.S.H., Raventos, M.T., Reitsma, J.M., Saiz, G., Schrodt, F., Sonké, B., Taedoumg, H.E., Tan, S., White, L., Wöll, H., Lloyd, J., 2011. Height-diameter allometry of tropical forest trees. Biogeosciences 8, 1081-1106.
Feliciano, E.A., Wdowinski, S., Potts, M.D., Lee, S.-K., Fatoyinbo, T.E., 2017. Estimating Mangrove Canopy Height and Above-Ground Biomass in the Everglades National Park with Airborne LiDAR and TanDEM-X Data. Remote Sensing 9, 702.
Fox, J., Monette, G., 2025. cv: Cross-Validating Regression Models, In: 2.0.4, R.p.v. (2.0.4, R.p.v.)2.0.4, R.p.v.s).
Fromard, F., Puig, H., Mougin, E., Marty, G., Betoulle, J.L., Cadamuro, L., 1998. Structure, above-ground biomass and dynamics of mangrove ecosystems: new data from French Guiana. Oecologia 115, 39-53.
Gingerich, P.D., 2000. Arithmetic or Geometric Normality of Biological Variation: an Empirical Test of Theory. Journal of Theoretical Biology 204, 201-221.
Golley, F., Howard, T.O., Ronald, F.W., 1962. The Structure and Metabolism of a Puerto Rican Red Mangrove Forest in May. Ecology 43, 9-19.
Gray, H.R., 1956. The form and taper of forest-tree stems. University of Oxford, Imperial Forestry Institute.
Hair, J.F., Babin, B.J., Anderson, R.E., Black, W.C., 2022. Multivariate Data Analysis. Cengage Learning.

Hamilton, S.E., Presotto, A., 2024. A Global Database to Monitor Annual Mangrove Forest Change, 2000-2020. International Journal of Applied Geospatial Research 15.
Harrison, X.A., Donaldson, L., Correa-Cano, M.E., Evans, J., Fisher, D.N., Goodwin, C.E.D., Robinson, B.S., Hodgson, D.J., Inger, R., 2018. A brief introduction to mixed effects modelling and multi-model inference in ecology. PeerJ 6, e4794.
He, Z., Feng, X., Chen, Q., Li, L., Li, S., Han, K., Guo, Z., Wang, J., Liu, M., Shi, C., Xu, S., Shao, S., Liu, X., Mao, X., Xie, W., Wang, X., Zhang, R., Li, G., Wu, W., Zheng, Z., Zhong, C., Duke, N.C., Boufford, D.E., Fan, G., Wu, C.-I., Ricklefs, R.E., Shi, S., 2022. Evolution of coastal forests based on a full set of mangrove genomes. Nature Ecology & Evolution 6, 738-749.
Heumann, B.W., 2011. Satellite remote sensing of mangrove forests: Recent advances and future opportunities. Progress in Physical Geography-Earth and Environment 35, 87-108.
Hox, J., Moerbeek, M., van de Schoot, R., 2017. Multilevel Analysis: Techniques and Applications, Third Edition (3rd ed.) ed. Routledge.
Hox, J.J., Maas, C.J.M., Brinkhuis, M.J.S., 2010. The effect of estimation method and sample size in multilevel structural equation modeling. Statistica Neerlandica 64, 157-170.
Hu, T., Zhang, Y., Su, Y., Zheng, Y., Lin, G., Guo, Q., 2020. Mapping the Global Mangrove Forest Aboveground Biomass Using Multisource Remote Sensing Data. Remote Sensing 12, 1690.
Huxley, J.S., 1924. Constant differential growth ratios and their significance. Nature 114, 895-896.
Imbert, D., Rollet, B., 1989. Phytomasse aerienne et production primaire dans la mangrove du Grand Cul-De-Sac Marin (Guadeloupe, Antillas Francaises). Bull. Ecol. 20, 27-39.
Jenkins, J., Chojnacky, D., Heath, L., Birdsey, R., 2003. National Scale Biomass Estimators for United States Tree Species. Forest Science 49, 12-35.
Kerkhoff, A., Enquist, B., 2009. Multiplicative by nature: Why logarithmic transformation is necessary in allometry. Journal of theoretical biology 257.
Komiyama, A., Ong, J.E., Poungparn, S., 2008. Allometry, biomass, and productivity of mangrove forests: A review. Aquatic Botany 89, 128-137.
Komiyama, A., Poungparn, S., Kato, S., 2005. Common allometric equations for estimating the tree weight of mangroves. Journal of Tropical Ecology 21.
Krauss, K.W., Osland, M.J., 2020. Tropical cyclones and the organization of mangrove forests: a review. Annals of Botany 125, 213-234.
Laegdsgaard, P., Johnson, C., 2001. Why do juvenile fish utilise mangrove habitats? Journal of Experimental Marine Biology and Ecology 257, 229-253.
Leal, M., Spalding, M.D. (Leal, M., Spalding, M.D.(Leal, M., Spalding, M.D.s), 2022. The State of the World's Mangroves 2022. Global Mangrove Alliance.
Lima, A., Suwa, R., Ribeiro, G.H., Kajimoto, T., Santos, J., Silva, R., Souza, C., Barros, P., Noguchi, H., Ishizuka, M., Higuchi, N., 2012. Allometric models for estimating above-and below-ground biomass in Amazonian forests at São Gabriel da Cachoeira in the upper Rio Negro, Brazil.
Lüdecke, D., Ben Shachar, M., Patil, I., Waggoner, P., Makowski, D., 2021. performance: An R Package for Assessment, Comparison and Testing of Statistical Models. The Journal of Open Source Software 6, 3139.
Lüdecke, D., Waggoner, P., Makowski, D., 2019. insight: A Unified Interface to Access Information from Model Objects in R. Journal of Open Source Software 4, 1412.
Lugo, A.E., Snedaker, S.C., 1974. The Ecology of Mangroves. Annual Review of Ecology and Systematics 5, 39-64.
McMahon, T., 1973. Size and shape in biology. Science 179, 1201-1204.
McMahon, T.A., Kronauer, R.E., 1976. Tree structures: Deducing the principle of mechanical design. Journal of Theoretical Biology 59, 443-466.
Menard, S., 1994. Applied Logistic Regression Analysis. Sage, Thousand Oaks, CA.

Mumby, P.J., Edwards, A.J., Arias-Gonzalez, J.E., Lindeman, K.C., Blackwell, P.G., Gall, A., Gorczynska, M.I., Harborne, A.R., Pescod, C.L., Renken, H., Wabnitz, C.C.C., Llewellyn, G., 2004. Mangroves enhance the biomass of coral reef fish communities in the Caribbean. Nature 427, 533-536.
Nakagawa, S., Johnson, P.C.D., Schielzeth, H., 2017. The coefficient of determination R(2) and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. J R Soc Interface 14.
Navarro, A., Young, M., Allan, B., Carnell, P., Macreadie, P., Ierodiaconou, D., 2020. The application of Unmanned Aerial Vehicles (UAVs) to estimate above-ground biomass of mangrove ecosystems. Remote Sensing of Environment 242.
Niklas, K.J., 1994. Plant Allometry: The Scaling of Form and Process. Univ. Chicago Press, Chicago.
Niklas, K.J., 1995. Size-dependent Allometry of Tree Height, Diameter and Trunk-taper. Annals of Botany 75, 217-227.
O'brien, R.M., 2007. A Caution Regarding Rules of Thumb for Variance Inflation Factors. Quality & Quantity 41, 673-690.
Oluwajuwon, T.V., Ogbuka, C.E., Ogana, F.N., Hossain, M.S., Israel, R., Lee, D.J., 2025. Describing and Modelling Stem Form of Tropical Tree Species with Form Factor: A Comprehensive Review. Forests 16, 29.
Osland, M.J., Day, R.H., Larriviere, J.C., From, A.S., 2014. Aboveground Allometric Models for Freeze-Affected Black Mangroves (Avicennia germinans): Equations for a Climate Sensitive Mangrove-Marsh Ecotone. PLOS ONE 9, e99604.
Paul, K.I., Radtke, P.J., Roxburgh, S.H., Larmour, J., Waterworth, R., Butler, D., Brooksbank, K., Ximenes, F., 2018. Validation of allometric biomass models: How to have confidence in the application of existing models. Forest Ecology and Management 412, 70-79.
Price, C.A., Branoff, B., Cummins, K., Kakaï, R.G., Ogurcak, D., Papeş, M., Ross, M., Whelan, K.R.T., Schroeder, T.A., 2024. Global Data Compilation Across Climate Gradients Supports the Use of Common Allometric Equations for Three Transatlantic Mangrove Species. Ecology and Evolution 14, e70577.
Price, C.A., Drake, P., Veneklaas, E.J., Renton, M., 2022. Flow similarity, stochastic branching, and quarter-power scaling in plants. Plant Physiology 190, 1854-1865.
Price, C.A., Schroeder, T.A., 2024. Flow similarity model predicts allometric size dependence, curvature, and covariation of 285 North American tree species. Ecosphere 15, e4835.
Price, C.A., Schroeder, T.A., Branoff, B., Van Bloem, S.J., Pillot-Torres, N., Chaudry, M.H., Papeş, M., Marcano-Vega, H., 2025. Allometric relationships for the linear dimensions and biomass of Puerto Rican mangroves. Forest Ecology and Management 593, 122908.
Pugh, T.A.M., Arneth, A., Kautz, M., Poulter, B., Smith, B., 2019. Important role of forest disturbances in the global biomass turnover and carbon sinks. Nature Geoscience 12, 730-735.
Qin, Y., Zhu, G., Liang, M., Feng, B., 2025. Exploring Global Research Trends in Mangrove Ecosystems: A Bibliometric Analysis of Key Issues and Future Directions (1996–2022). Land Degradation & Development 36, 3163-3177.
Quadros, A.F., Zimmer, M., 2017. Dataset of "true mangroves" plant species traits. Biodiversity Data Journal 5.
RCoreTeam, 2021. R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.
Ross, M., Ruiz, P., Telesnicki, G., Meeder, J., 2001. Estimating above-ground biomass and production in mangrove communities of Biscayne National Park, Florida (U.S.A.). Wetlands Ecology and Management 9, 27-37.
Rovai, A., Twilley, R., Castañeda-Moya, E., Midway, S., Friess, D., Trettin, C., Bukoski, J., Stovall, A., Pagliosa, P., Fonseca, A., Mackenzie, R., Aslan, A., Sasmito, S., Sillanpää, M., Cole, T., Purbopuspito, J.,

Warren, M., Murdiyarso, D., Mofu, W., Riul, P., 2021. Macroecological patterns of forest structure and allometric scaling in mangrove forests. Global Ecology and Biogeography 30.
Roxburgh, S.H., Paul, K.I., Clifford, D., England, J.R., Raison, R.J., 2015. Guidelines for constructing allometric models for the prediction of woody biomass: How many individuals to harvest? Ecosphere 6, art38.
Schmitz, N., Verheyden, A., Beeckman, H., Kairo, J.G., Koedam, N., 2006. Influence of a Salinity Gradient on the Vessel Characters of the Mangrove Species Rhizophora mucronata. Annals of Botany 98, 1321-1330.
Schumacher, F.X., Hall, F.D.S., 1933. Logarithmic expression of timber-tree volume. Journal of Agricultural Research 47, 0719-0734.
Sherman, R.E., Fahey, T.J., Martinez, P., 2003. Spatial patterns of Biomass and aboveground net primary productivity in a mangrove ecosystem in the Dominican Republic. Ecosystems 6, 384-398.
Shinozaki, K., Yoda, K., Hozumi, K., Kira, T., 1964. A QUANTITATIVE ANALYSIS OF PLANT FORM-THE PIPE MODEL THEORY : I.BASIC ANALYSES. 日本生態学会誌 14, 97-105.
Sileshi, G.W., 2014. A critical review of forest biomass estimation models, common mistakes and corrective measures. Forest Ecology and Management 329, 237-254.
Simard, M., Fatoyinbo, L., Smetanka, C., Rivera-Monroy, V.H., Castañeda-Moya, E., Thomas, N., Van der Stocken, T., 2019. Mangrove canopy height globally related to precipitation, temperature and cyclone frequency. Nature Geoscience 12, 40-45.
Simard, M., Zhang, K.Q., Rivera-Monroy, V.H., Ross, M.S., Ruiz, P.L., Castaneda-Moya, E., Twilley, R.R., Rodriguez, E., 2006. Mapping height and biomass of mangrove forests in Everglades National Park with SRTM elevation data. Photogrammetric Engineering and Remote Sensing 72, 299-311.
Smith, T.J., Whelan, K.R.T., 2006. Development of allometric relations for three mangrove species in South Florida for use in the Greater Everglades Ecosystem restoration. Wetlands Ecology and Management 14, 409-419.
Snijders, T., Bosker, R., 2012. Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. Sage Publishers, London.
Soares, M.L.G., Schaeffer-Novelli, Y., 2005. Above-ground biomass of mangrove species. I. Analysis of models. Estuarine, Coastal and Shelf Science 65, 1-18.
Spalding, M., Kainuma, M., Collins, L., 2010. World Atlas of Mangroves. Taylor and Francis, New York.
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., Zavala, M.A., 2014. Rate of tree carbon accumulation increases continuously with tree size. Nature 507, 90-93.
Ter-Mikaelian, M.T., Korzukhin, M.D., 1997. Biomass equations for sixty-five North American tree species. Forest Ecology and Management 97, 1-24.
Virgulino-Júnior, P.C.C., Gardunho, D.C.L., Silva, D.N.C., Fernandes, M.E.B., 2020. Wood density in mangrove forests on the Brazilian Amazon coast. Trees 34, 51-60.
Warton, D.I., Wright, I.J., Falster, D.S., Westoby, M., 2006. Bivariate line-fitting methods for allometry. Biol Rev Camb Philos Soc 81, 259-291.
West, G.B., Brown, J.H., Enquist, B.J., 1999. A general model for the structure and allometry of plant vascular systems. Nature 400, 664-667.
Yepes, A., Zapata, M., Bolivar, J., Monsalve, A., Espinosa, S.M., Sierra-Correa, P.C., Sierra, A., 2016. Tree above-ground biomass allometries for carbon stocks estimation in the Caribbean mangroves in Colombia. Revista De Biologia Tropical 64, 913-926.

Zanvo, S.M.G., Mensah, S., Salako, K.V., Glèlè Kakaï, R., 2023. Tree height-diameter, aboveground and belowground biomass allometries for two West African mangrove species. Biomass and Bioenergy 176, 106917.
Zhang, K.Q., Liu, H.Q., Li, Y.P., Xu, H.Z., Shen, J., Rhome, J., Smith, T.J., 2012. The role of mangroves in attenuating storm surges. Estuarine Coastal and Shelf Science 102, 11-23.

	A. germinans	*L. racemosa*	*R. mangle*
Sample Size	255	239	216
Mean DBH (cm)	9.785	7.375	7.177
Min DBH (cm)	0.257	0.500	0.500
Max DBH (cm)	42.430	38.700	32.389
Std Dev DBH (cm)	7.768	7.869	7.951
Mean Height (m)	4.682	6.068	5.360
Min Height (m)	0.094	0.210	0.255
Max Height (m)	18.100	17.820	16.600
Std Dev Height (cm)	5.121	5.221	5.348
Mean Canopy Diameter (m)	2.032	1.653	1.710
Min Canopy Diameter (m)	0.050	0.050	0.050
Max Canopy Diameter (m)	9.700	6.380	7.050
Mean Mass Dry (kg)	53.804	28.623	40.784
Min Mass Dry (kg)	0.001	0.004	0.005
Max Mass Dry (kg)	1552.159	698.280	1038.911
Std Dev Mass Dry (kg)	115.508	118.557	123.342
Mean Wood Density (g/cm^3)	0.720	0.661	0.754
Min Wood Density (g/cm^3)	0.605	0.57	0.641
Max Wood Density (g/cm^3)	0.863	0.783	0.848
Std Dev Wood Density (g/cm^3)	0.039	0.038	0.039

Table 1. Sample size and summary statistics of tree structural parameters by species for the mean values, ranges and standard deviations of the five field-measured parameters underpinning our analysis. Note, wood density values are from the San Juan dataset.

	Everglades	Guadeloupe	Puerto Rico	Biscayne	French Guiana	Bertioga	Guaratiba	Benin	Louisiana	San Juan
AGB min (kg)	0.140	10.000	0.775	0.004	0.128	0.097	0.280	4.988	0.001	0.001
AGB max (kg)	186.850	1145.800	16.290	9.315	1552.159	310.233	279.711	480.073	2.452	698.280
AGB mean (kg)	33.583	151.216	3.989	0.857	56.046	42.410	43.416	107.346	0.298	32.981
AGB range (kg)	186.710	1135.800	15.515	9.311	1552.031	310.136	279.431	475.086	2.451	698.279
DBH min (cm)	0.500	6.600	1.200	0.648	0.257	0.600	1.100	3.183	NA	0.525
DBH max (cm)	21.500	40.700	5.300	5.860	42.430	20.000	22.200	41.730	NA	38.700
DBH mean (cm)	6.469	16.140	2.480	2.163	5.247	7.373	7.546	17.187	NA	7.286
DBH range (cm)	21.000	34.100	4.100	5.212	42.174	19.400	21.100	38.547	NA	38.175
Height min (m)	1.500	8.300	3.600	0.430	NA	1.200	1.900	2.850	0.310	0.094
Height max (m)	17.820	18.100	8.600	5.570	NA	12.100	15.100	13.200	1.570	13.800
Height mean (m)	7.202	13.540	4.960	2.367	NA	6.743	7.229	6.697	0.835	4.907
Height range (m)	16.320	9.800	5.000	5.140	NA	10.900	13.200	10.350	1.260	13.706
Canopy min (m)	NA	1.800	NA	0.020	NA	0.395	0.625	1.950	0.050	0.062
Canopy max (m)	NA	9.700	NA	2.700	NA	7.000	6.380	7.750	1.735	8.600
Canopy mean (m)	NA	3.820	NA	0.539	NA	2.375	2.352	4.732	0.463	1.540
Canopy range (m)	NA	7.900	NA	2.680	NA	6.605	5.755	5.800	1.685	8.539
A. germinans (n)	8	21	-	34	57	-	-	39	56	40
L. racemosa (n)	10	17	-	58	37	45	33	-	-	39
R. mangle (n)	14	17	10	65	4	33	32	-	-	41

Table 2. Structural data for mangrove trees pooled across species for the 10 sites. Species totals are given in the last three rows. Note not all sites collected data on each structural parameter (indicated by NA). Note, we lack site specific wood density estimates.

Table 3. Highest ranked model from each model group (group of variables). The number of each model group is in the first column. Each model in column 2 has a corresponding group of variables from which it is composed and we selected the best performing model for each group of variables for inclusion here. Extended description of the column headers can be found in Appendix B. All models in all model groups can be found in Tables S1-2. Diff Best Model (kg) is the difference in kg between each model and the best performing model. The three single variable fixed effects models are included in the last three rows for comparison.

ModelN	Model	MixedEffects Effects	Effects	Model Rank	R^2	AIC	BIC	RMdSE linear (kg)	RRMdSE %	Difference Best Model (kg)	RMSE linear (kg)	RRMSE %	Difference Best Model (kg)	Sigma	ICC	Maximum VIF
16	DBH, Height, Canopy, Density	NA	Fixed	1	0.974	252.770	275.726	2.495	15.456	0.000	34.609	57.452	0.000	0.348	NA	11.555
4	Comp.DBH.H.Den	Site	MixedIntSlope	2	0.976	283.623	306.579	2.594	16.070	0.099	33.266	55.223	-1.343	0.342	0.631	NA
12	DBH, Height, Canopy	Species&Site	MixedInt	3	0.976	249.039	275.821	2.607	16.153	0.113	37.673	62.538	3.064	0.334	0.601	9.290
1	DBH	Site	MixedIntSlope	5	0.971	323.714	346.670	2.517	15.593	0.022	28.574	47.427	-6.035	0.366	0.502	NA
14	DBH, Canopy, Density	Site	MixedInt	6	0.969	340.396	363.352	2.334	14.462	-0.160	33.298	55.269	-1.211	0.384	0.191	3.375
7	DBH, Canopy	Species&Site	MixedInt	8	0.969	344.266	367.222	2.412	14.941	-0.085	34.509	57.301	-0.066	0.383	0.244	3.320
13	DBH, Height, Density	Site	MixedInt	10	0.966	352.908	375.864	2.727	16.896	0.232	44.532	73.932	9.923	0.393	0.100	5.829
6	DBH, Height	Species&Site	MixedInt	11	0.966	361.754	384.710	2.722	16.864	0.227	45.521	75.621	10.945	0.394	0.249	6.037
8	DBH, Density	Site	MixedInt	18	0.963	404.265	423.395	2.653	16.436	0.159	47.034	78.078	12.425	0.425	0.106	1.003
15	Height, Canopy, Density	Site	MixedInt	26	0.957	451.552	474.508	4.254	26.356	1.764	59.433	98.661	24.824	0.452	0.206	3.727
9	Height, Canopy	Site	MixedInt	27	0.957	452.622	471.752	3.884	24.059	1.389	59.397	98.601	24.788	0.454	0.196	3.506
5	Comp.Can.H.Den	NA	Fixed	32	0.925	603.523	615.001	5.167	32.012	2.672	52.907	87.827	18.298	0.586	NA	NA
2	Height	Species&Site	MixedIntSlope	33	0.923	714.713	749.147	7.386	45.764	4.892	72.875	120.975	38.266	0.641	0.683	NA
10	Height, Density	Site	MixedInt	35	0.903	723.035	742.168	8.099	50.168	5.604	83.536	138.673	48.927	0.668	0.538	1.005
11	Canopy, Density	Site	MixedInt	37	0.841	872.416	891.507	7.106	44.026	4.611	78.723	130.684	44.114	0.848	0.108	1.005

3	Canopy	Species&Site	MixedInt	38	0.842	875.599	894.729	7.176	44.461	4.681	76.916	127.683	42.307	0.850	0.116	NA
1	DBH	NA	Fixed	29	0.949	468.350	479.828	3.009	18.640	0.514	61.780	102.557	27.171	0.480	NA	NA
3	Canopy	NA	Fixed	41.5	0.822	894.665	906.143	8.159	50.554	5.665	76.510	127.010	41.901	0.900	NA	NA
2	Height	NA	Fixed	45	0.811	915.262	926.740	13.876	85.971	11.381	86.632	143.813	52.023	0.928	NA	NA

Declaration of interests

☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Charles A Price reports financial support was provided by USDA Forest Service Southern Research Station. Reports a relationship with that includes:. Has patent pending to. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Graphical abstract

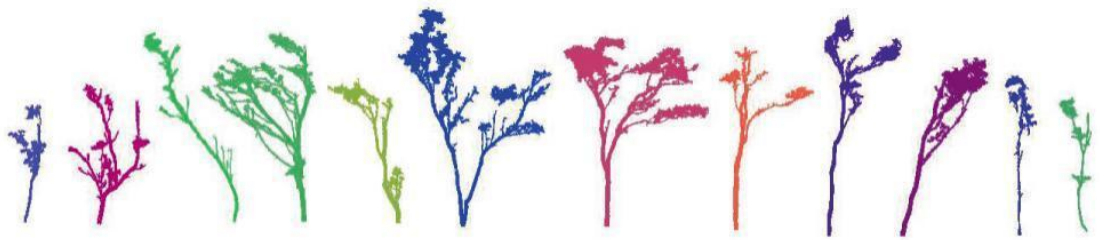
Highlights

- The use of local vs global allometric equations is a long-standing issue in mangrove ecology
- We analyzed a global data compilation for Atlantic East Pacific mangroves from 10 sites
- Mixed effects models outperformed fixed effects models, but the improvement was modest
- Diameter at breast height (DBH) exhibits the least variation as a predictive variable across species and sites
- Variance partitioning indicates that fixed effects dominate

How does location and species identity affect mangrove biomass allometry?

Fixed Effects
DBH
Height
Canopy
Wood density

Mixed Effects
Species ID
Location



16 equations
112 different models
Seven Statistics
R^2
AIC
BIC
RMSE
RMdSE
Sigma
ICC

10 sites, 710 trees, 3 species

Mangrove seedling

Best model: R^2=0.974, RMdSE=2.495 (kg)

$$\log AGB = \log c + \beta_1 \, logD + \beta_2 \, logH + \beta_3 \, logC$$

The best model predicts above ground biomass (AGB) using DBH (*D*), height (*H*) and canopy diameter (*C*), species ID and location

However, just measuring DBH (irrespective of species or location) only adds 161 grams of error to the average tree

Global model: R^2=0.931, RMdSE=2.656 (kg)

$$\log AGB = \log c + \beta_1 \, logD$$

Graphics Abstract

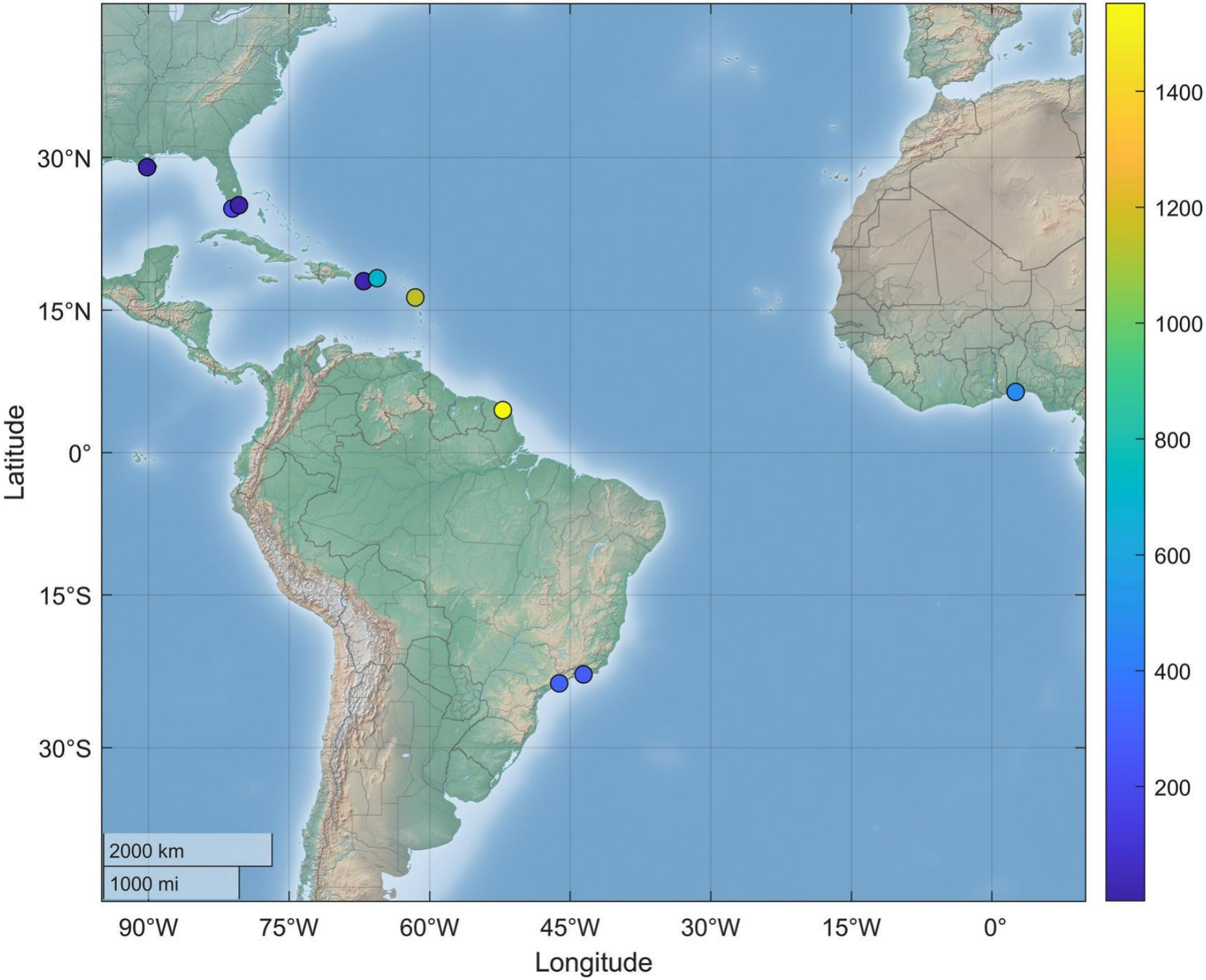

30°N
15°N
0°
15°S
30°S
Latitude
90°W
75°W
60°W
45°W
30°W
15°W
0°
Longitude
2000 km
1000 mi
1400
1200
1000
800
600
400
200

Figure 1

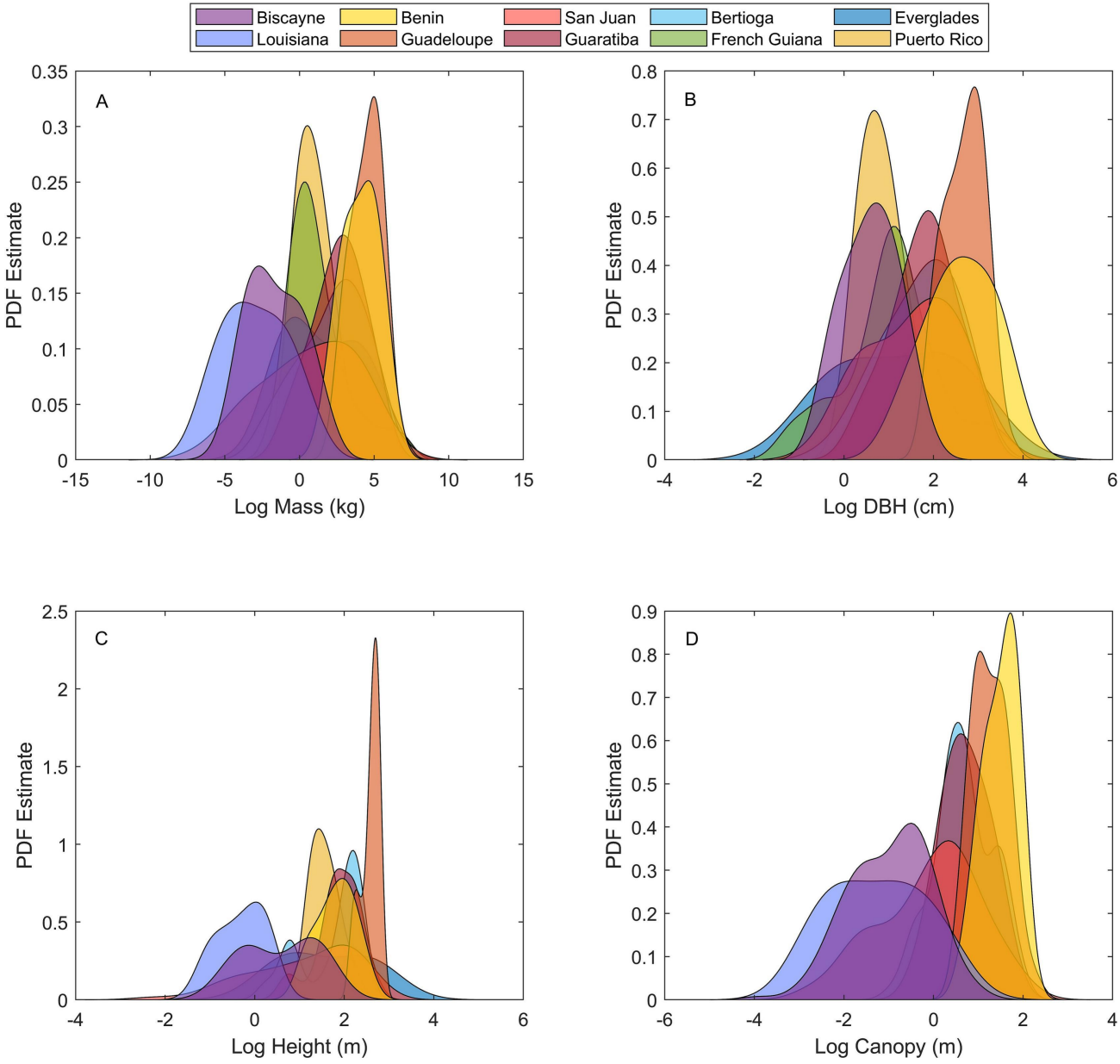

Biscayne
Louisiana
Benin
Guadeloupe
San Juan
Guaratiba
Bertioga
French Guiana
Everglades
Puerto Rico
A
PDF Estimate
Log Mass (kg)
B
PDF Estimate
Log DBH (cm)
C
PDF Estimate
Log Height (m)
D
PDF Estimate
Log Canopy (m)

Figure 2

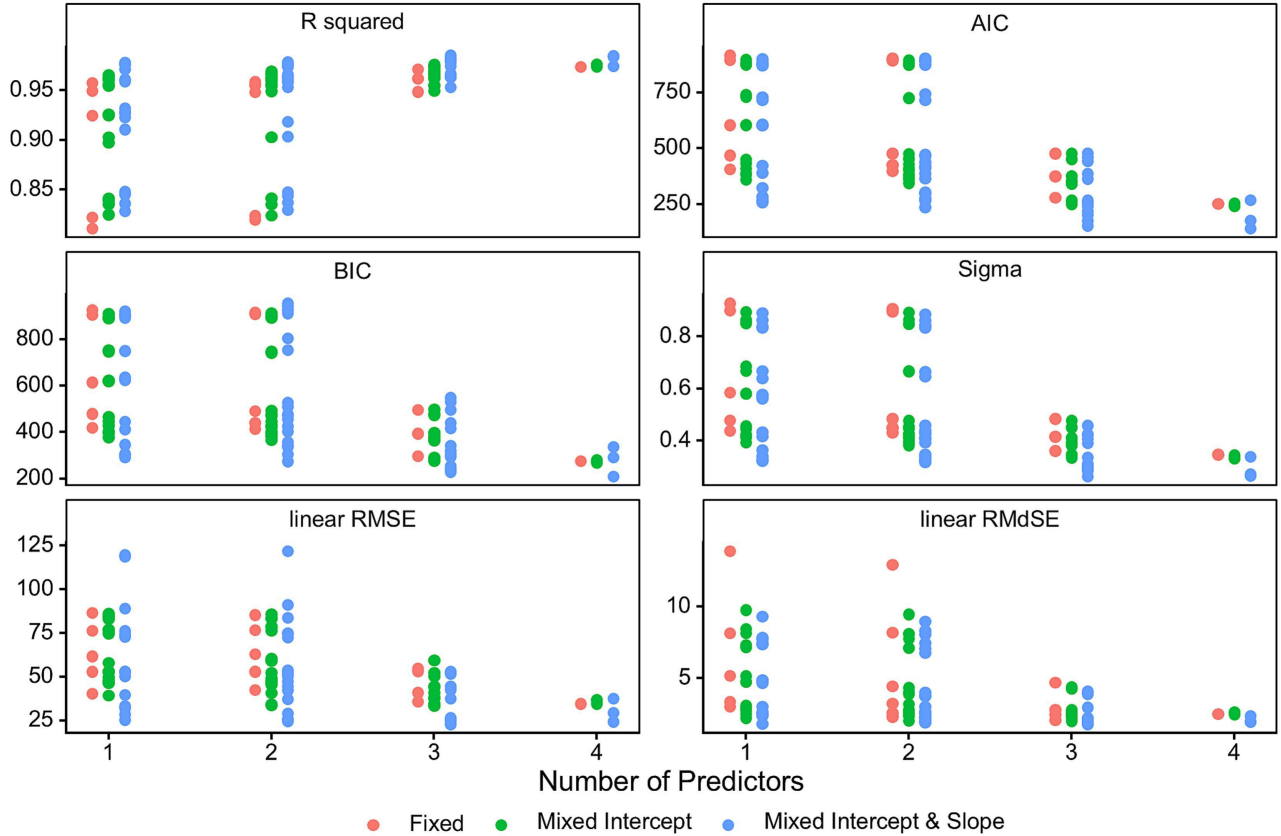

R squared
0.95
0.90
0.85
AIC
750
500
250
BIC
800
600
400
200
Sigma
0.8
0.6
0.4
linear RMSE
125
100
75
50
25
linear RMdSE
10
5
1
2
3
4
Number of Predictors
Fixed
Mixed Intercept
Mixed Intercept & Slope

Figure 3

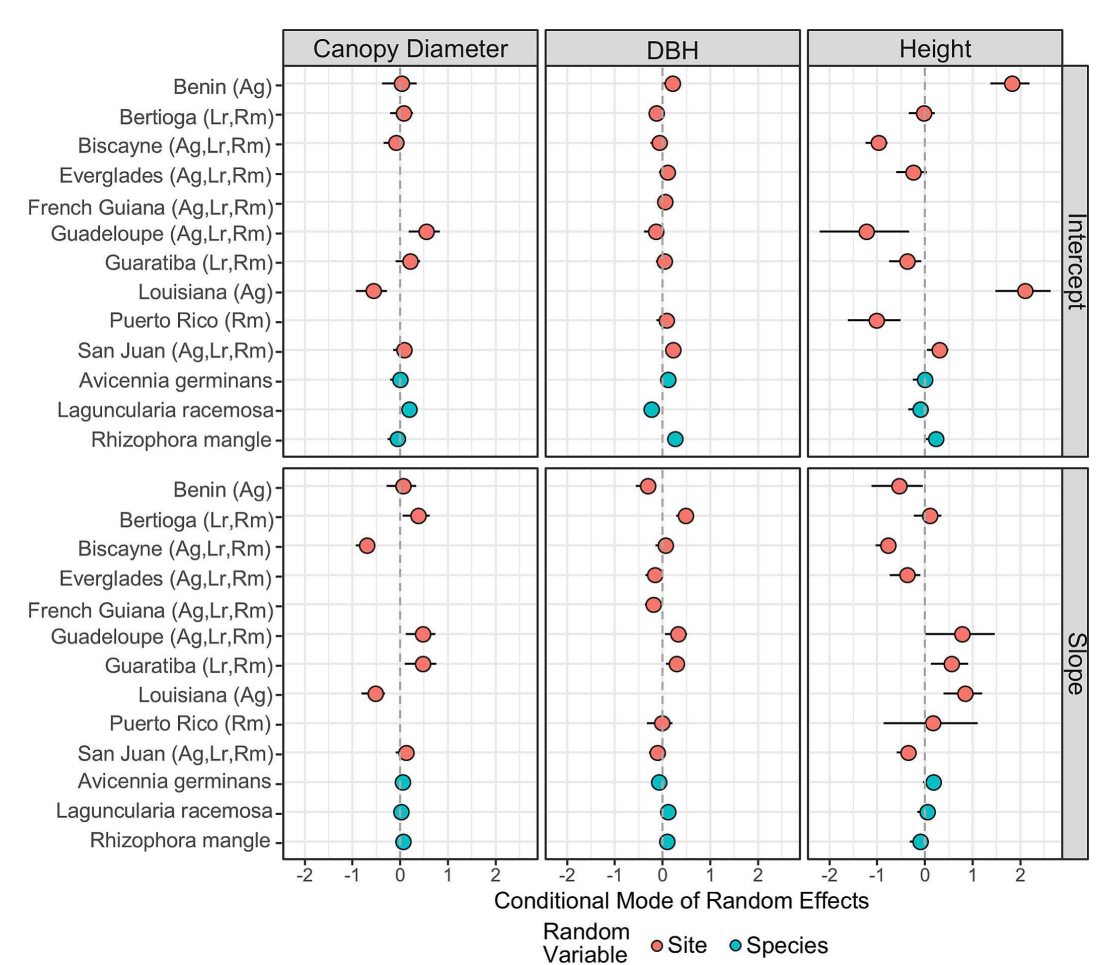

Canopy Diameter
DBH
Height
Intercept
Slope
Benin (Ag)
Bertioga (Lr,Rm)
Biscayne (Ag,Lr,Rm)
Everglades (Ag,Lr,Rm)
French Guiana (Ag,Lr,Rm)
Guadeloupe (Ag,Lr,Rm)
Guaratiba (Lr,Rm)
Louisiana (Ag)
Puerto Rico (Rm)
San Juan (Ag,Lr,Rm)
Avicennia germinans
Laguncularia racemosa
Rhizophora mangle
-2
-1
0
1
2
Conditional Mode of Random Effects
Random Variable
Site
Species

Figure 4

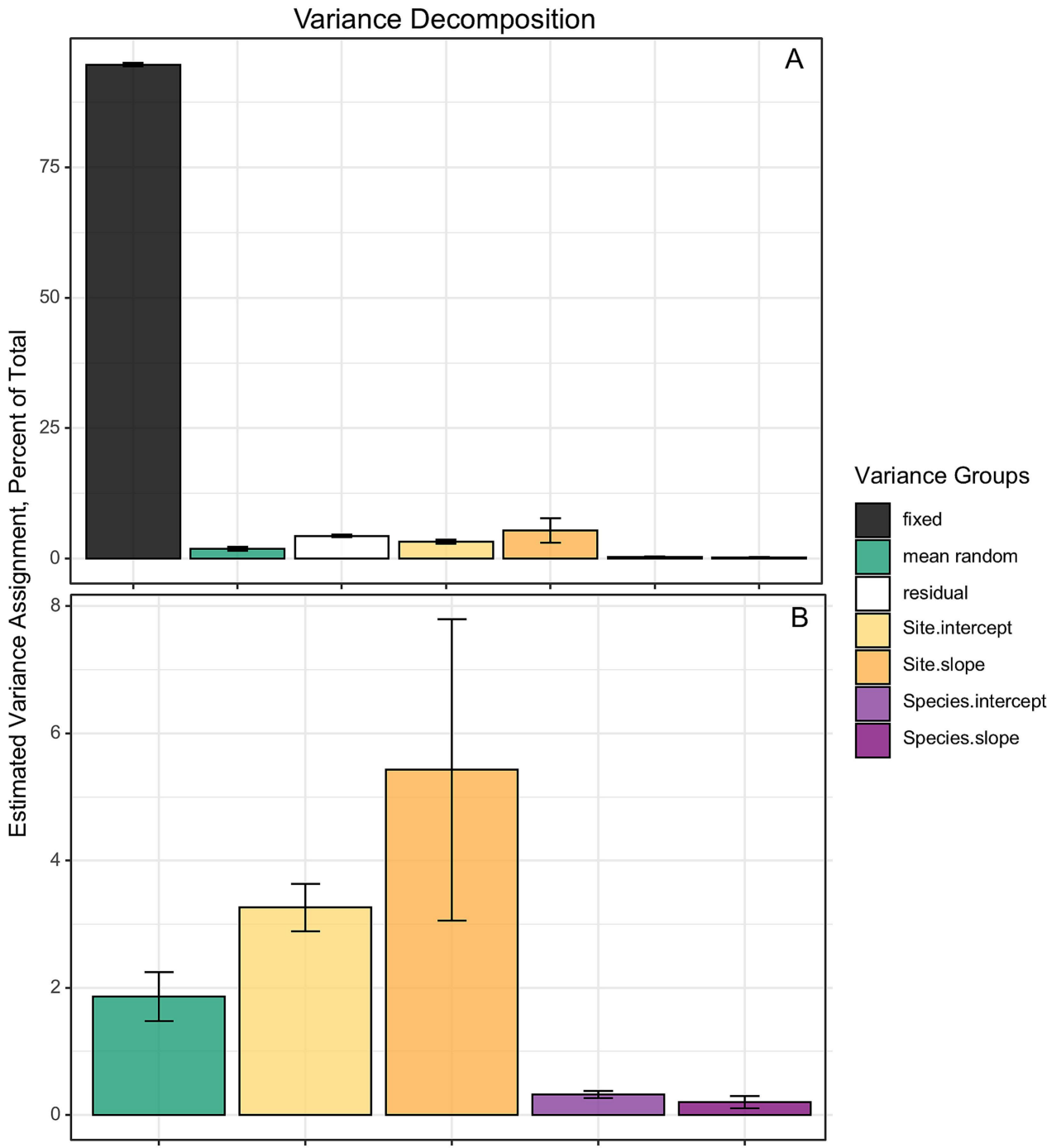

Variance Decomposition
A
B
Estimated Variance Assignment, Percent of Total
75
50
25
0
8
6
4
2
0
Variance Groups
fixed
mean random
residual
Site.intercept
Site.slope
Species.intercept
Species.slope

Figure 5

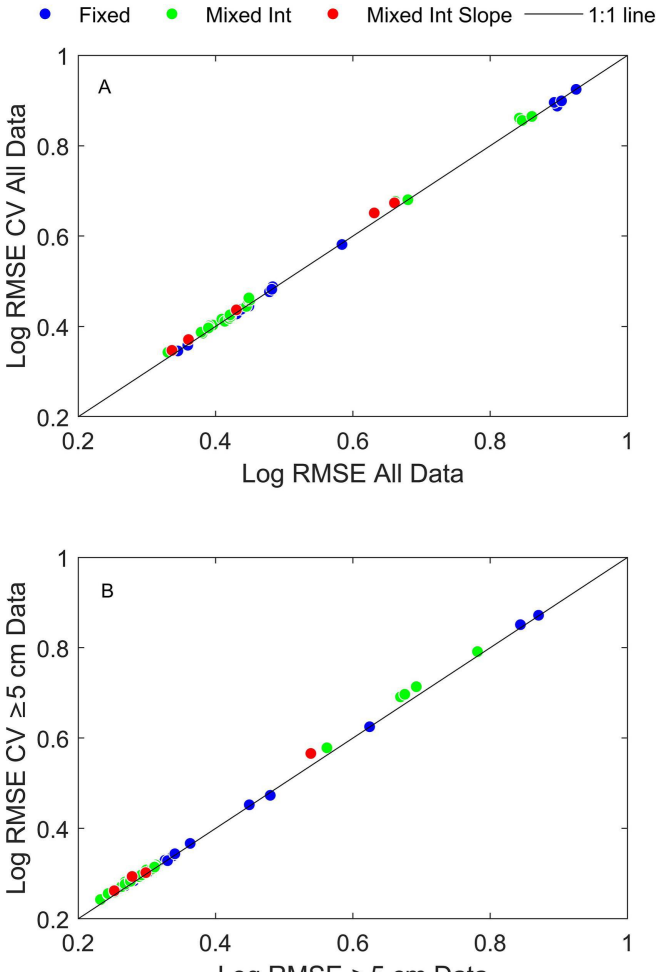

Fixed
Mixed Int
Mixed Int Slope
1:1 line
A
Log RMSE CV All Data
Log RMSE All Data
B
Log RMSE CV ≥5 cm Data
Log RMSE ≥5 cm Data
0.2
0.4
0.6
0.8
1

Figure 6