Title: Multilevel allometric growth equations improve accuracy of carbon monitoring during forest restoration.

Authors:

Brad Oberle*[a]

Piper O. Cole[a]

Garcia Frank[a]

Alexandra Gates[a]

Brittney Hall[a]

Deric Harvey[a]

Melody E. Scott[a]

Cas Setterberg[a]

Simon P. Bustetter[a]

[a]Division of Natural Sciences, New College of Florida, 5800 Bay Shore Road, Sarasota, Florida, 34243, USA.

*Corresponding author: brad.oberle@gmail.com

Highlights

- **Replacing exotic trees with native seedlings complicates forest carbon monitoring**
- **Multilevel allometric growth equations borrow strength from sampling many species**
- **Generic equations sacrifice accuracy and precision for smaller, atypical plants**
- **Underestimating C differences before and after restoration influences management**

Abstract: Managing disturbed forests for climate mitigation and biodiversity requires monitoring the carbon (C) cycle consequences of replacing established exotic vegetation with native seedlings. Standard approaches rely on allometric growth equations with unexplored limitations for measuring C changes during restoration. Most plants lack species-specific allometric growth equations, which may perform poorly for different growth forms, especially when applied to both mature trees and seedlings. To address these limitations, we generated and compared allometric growth equations for four woody species with different biogeographic origins and growth forms, including two high impact invasive species, *Cupaniopisis anacardioides* and *Schinus terebinthifolia*. By borrowing strength from sampling across species to reduce estimation error within species, Bayesian multilevel models generated more accurate and precise estimates than either independent species-level models or generic equations. Because errors increased for smaller plants and species with unusual growth forms, allometric growth equations from custom multilevel models generated higher baseline aboveground biomass estimates and lower post-restoration estimates, which has important implications for monitoring C consequences of invasive tree management.

Keywords: afforestation, Bayesian models, Brazilian Peppertree, Carrotwood, forest carbon, invasive species control, restoration

Abbreviations:

AGB	Aboveground biomass
AGE	Allometric Growth Equation
AICc	Bias Corrected Akaike Information Criterion (Hurvich and Tsai 1989)
Bdia	Stem basal diameter
BP	Brazilian Peppertree, *Schinus terebinthifolia*
CD90	Canopy diameter perpendicular to midpoint of Cdmax
CDmax	Broadest canopy diameter
CI	Confidence interval
CW	Carrotwood, *Cupaniopsis anacardioides*

D130	Stem diameter at 130 cm
D140	Stem diameter at 140 cm
Dia	Diameter covariate
DIC	Deviance Information Criterion (Spiegelhalter et al. 2002)
DRC	Stem diameter at Root Collar
Hgt	Plant height from ground to tallest observed point
Len	Length from observer to canopy high point
Sdmwmr	Sample dry mass wet mass ratio
SH	Florida native shrubs represented in this study (i.e., SS, WM)
SS	Simpson's stopper, *Myrcianthes fragrans*
WM	Southern wax mytle, *Morella cerifera*

1 Introduction

For mitigating biodiversity loss and climate change, forest management often prioritizes restoring native species and increasing carbon (C) storage (Dickie et al., 2014; Oberle et al., 2022). If target vegetation eventually stores more C than baseline conditions, these objectives may reinforce over time. However, restoration risks short-term C losses from forests dominated by woody invasives (Liao et al., 2008). Killing mature trees and replacing them with native seedlings changes biomass and stem size distributions with immediate consequences for C stocks and fluxes. As invasive necromass is removed, burned or decomposes, initial efflux may require decades before total ecosystem C storage recovers with native plant regrowth (Marchante et al., 2009). Documenting if and when restored forests recover C requires measurements and models with consistent performance across baseline and target stands that have radically different structures.

Standard methods for translating stand measurements into C estimates rely on allometric growth equations (AGEs, Eggleston et al., 2006). This class of models exploits the proportionality between different aspects of an organism's size to express biomass as a mathematical function of simpler non-destructive measurements, like stem diameter and height. In principle, these relationships reflect biophysical and anatomical constraints on development (West et al., 1999). In practice, AGEs are empirically parameterized (Chave et al., 2014; Conti et al., 2019). A set of individuals that vary in size are measured, felled, dissected, weighed and then analyzed for moisture and C content (Picard et al., 2012). The resulting biomass and size measurements generally exhibit a power-law relationship that can be parameterized using different modeling techniques and simplified under various model adequacy criteria (Clifford et al., 2013; see also Picard et al., 2015). Following best practices at every step, from species and stem selection (Roxburgh et al., 2015), to measurement and model fitting (Picard et al., 2012), typically yields relationships with high accuracy and precision, supporting their application to a range of basic and applied forest research questions.

While indispensable and ubiquitous, AGEs have at least three specific limitations when applied to C monitoring during forest restoration. First, although growing databases include hundreds of species-level AGEs (Henry et al., 2013), they constitute a small, unrepresentative fraction of the more than 150,000 global woody species (FitzJohn et al., 2014). Many impactful invasive trees and most natives from biodiversity hotspots are unrepresented. Although researchers have developed generic AGEs to apply to any unidentified species (e.g. Chave et al., 2014; Paul et al., 2016), few apply to all growth forms and many perform differently across the size classes (Conti et al., 2019; Duncanson et al., 2015). Whether or not generic AGEs provide sufficient accuracy and precision for C monitoring during forest restoration is unknown. Second, all AGEs, whether species-specific or generic, depend on one or more non-destructive size measurements. Those which predominate for trees, like diameter at breast height (DBH), are unmeasurable for seedlings shorter than the standard height, 1.3 m. Few studies have evaluated the reliability of AGEs when applied across both the mature trees that are removed and the seedlings used to replace them. Finally, standard statistical techniques for fitting power-law AGEs produce expected values of biomass that depend not only on covariate effects for non-destructive measurements, but also on error variance (Clifford et al., 2013). Error variance, in turn, depends on unstandardized features of study design, like sample size and measurement precision (Roxburgh et al., 2015). Consequently, researchers' decisions about allocating sampling effort within versus between species can have direct impacts on biomass projections from AGEs. While techniques have been proposed to leverage the strength of sampling both within and between species, this class of multilevel model has seldom been applied to AGEs (Price et al., 2009; Zapata-Cuartas et al., 2012).

We quantify some consequences of these issues and propose solutions for monitoring C during forest restoration in Florida. The state has a very high proportion of invasive plants that threaten biologically diverse forests and the ecosystem services they provide (Gordon, 1998). We present new AGEs and C estimates for two high-impact invasive species and two important natives that are widely used for revegetation. We use a range of measurements applicable across mature trees and small seedlings. We also test alternative model fitting and simplification approaches, including generalized linear models, multimodel inference and Bayesian Multilevel regression. Finally, we compare estimates from customized multilevel versus generic AGEs for aboveground biomass (AGB) and C for inventories of baseline conditions in two forest restoration projects.

2 Materials and Methods

2.1 Study system and focal species

The study occurred in Sarasota and Manatee Counties, near the Gulf of Mexico in Florida, USA. Regional climate is humid temperate-subtropical with mean annual temperature of 22.8°C and annual precipitation of approximately 1.35 m varying between a hot summer wet season and a cool winter dry season. The study included two invasive species that had spontaneously established and two native species that had been propagated in nurseries.

The first woody invasive species was *Schinus terbinthifolia* (Anacardiaceae, Brazilian Peppertree or BP). A shrub or small tree native to open habitats of eastern South America, BP has become a high impact invader on several other continents (Mukherjee et al., 2012). In south Florida, the plant was introduced in the mid-19th century for horticulture and currently occupies over 280,000 ha, or about 2% of the land area of the State (Cuda et al., 2006). Existing research includes two analyses of size and growth form. Spector and Putz (2006) found that crowding changed stem shape to be longer for a given diameter, as individuals began scrambling over adjacent vegetation, but they did not report biomass. Erickson et al. (2017) found that aboveground biomass (AGB) scaled differently with stem diameter for plants with different genetic backgrounds. However, they did not measure other aspects of plant size, like stem length or height, and used an unconventional polynomial regression without reporting model fit statistics.

The second high impact invasive species, *Cupaniopsis anacardioides* (Sapindaceae, Carrotwood or CW) is a small tree that naturally occurs across a broad range of habitats in northern Australasia (Reynolds, 1985). It was introduced to Sarasota County, Florida for horticulture in 1968 (Enloe and Langeland, 2001). Since then, it has rapidly naturalized into a range of habitats and has become a particular conservation concern along the southwest Florida coasts where it co-occurs with other invasive trees at the margins of disturbed mangroves (Lockhart et al., 1999). Despite its local importance and recognition from regional exotic plant management authorities, CW has attracted relatively little research attention, with no published AGEs.

The first native species assessed, *Myrcianthes fragrans* (Myrtaceae, Simpson's stopper or SS) is a shrub species native to Florida and several other regions including the Caribbean, Mexico, Central America, Columbia and Venezuela (Armijos et al., 2018). This plant can thrive in different environments by tolerating high salinity and drought conditions. In addition to its importance in the wild, it has also become a useful landscaping plant with low maintenance needs. It has been identified as a threatened species within Florida (Ward et al., 2003). Research on this plant has been focused on the uses of its characteristic volatile oils (Armijos et al., 2018), with no published AGE.

The second native species included in this study, *Morella cerifera* (L.) (Myricaceae, syn. *Myrica cerifera*, Wax Myrtle, or WM) is a small evergreen tree or large shrub. It is native to the southeastern United States, including Florida (Van Deelen, 1991). WM can grow 5-10 m tall and are considered to have a 'fast' growth rate. They can survive well in different levels of sunlight, and are highly salt tolerant and require little care after establishment. Experimental evidence suggests that WM is specifically resistant to allelopathic properties of BP (Dunevitz and Ewel, 1981). There are no published AGEs of this species.

2.2 Non-destructive measurements

For each species, we selected stems along a logarithmic size distribution and collected several non-destructive measurements before harvesting stems and then separating and weighing tissues. However, differences among species in their sizes, growth forms and published research required different field protocols.

For CW, all plants had spontaneously recruited at the Tidy Island Preserve, a disturbed mangrove swamp, in sandy soils created from dredge spoil at the northeast corner of the property located at approximately 27° 27' 5" N, 82° 39' 14"W. We selected 19 individual stems along the diameter sequence exp({4:32}/10), with the smallest individual measuring 1.5 cm and the largest 25.4 cm DBH. We avoided individuals with deformities at the point of measurement. For each individual, we used diameter tapes to measure diameter at three locations: above the root collar (*DRC*) as the lowest point above root flare, at 130 cm (*D130*), and at 140 cm (*D140*). We also measured three features of the tree canopy. After walking the canopy perimeter, we measured the longest canopy diameter (*CDMAX*) in meters using a 50 meter tape held level. At the midpoint of *CDMAX*, we measured the perpendicular canopy diameter (*CD90*). From these two variables, we calculated the elliptical area of the canopy (*Can*) in m^2 as $CDMAX*CD90*\boldsymbol{\pi}$. Then, from a location from which the highest point of the canopy and a point on the ground directly below it were both visible, we measured the distance from the high point projection onto the ground to the observer in m using a 50 meter tape held level (*LEN*) and the inclinations in degrees from the observer both to the top of the canopy (*UP*) and to the ground point beneath it (*DOWN*) using a clinometer. We calculated plant height (*Hgt*) in m as $(\tan(UP/180*\boldsymbol{\pi})+\tan(DOWN/180*\boldsymbol{\pi}))*\mathrm{LEN}$, except for the smallest plant, for which we directly measured height from the ground with a 50 m tape.

Upon completing non-destructive measurements, we felled individuals at a point below *D130* and as close to the ground as possible. We separated aboveground biomass into 2 to 4 tissue categories depending on plant size and reproductive status. For each individual plant, we separated (1) leaves, including petioles and flowers when present from (2) branches, representing stems less than 7.5 cm basal diameter branching above the main stem at 140 cm. We also separated (3) trunks, representing the main stem below 140 cm and all stem segments with basal diameter greater than 7.5 cm. Finally, for individuals that were fruiting we separated (4) fruits, including peduncles. We weighed the total wet mass (*TWM*) of each tissue to the nearest 100 g in the field using a Pesola scale. To determine dry mass, we collected a subsample of every tissue present from each individual. For the leaf subsample, we avoided leaves that were directly exposed to the soil or open air during weighing. For the branch subsample, we included subsamples roughly corresponding to every quartile of the branch size distribution. For the trunk subsample, when applicable, we used a chainsaw to cut a segment including the *D130* and *D140* measurement locations. Fruit subsamples, when present, were also collected so as to avoid fruits that were exposed to the ground or the open air during weighing. We enclosed all tissue subsamples in sealed, labeled 4L polyethylene bags, which we transported to the lab for temporary storage at 4°C for a maximum of 28 days until further processing.

Because BP had some published biomass scaling information, we collected a smaller number of individuals and fewer non-destructive measurements. All individuals were growing on edges between unmanaged habitats and mowed lawns on the New College of Florida campus, near 27° 22' 50”N, 82° 33' 45”W. We identified 11 stems approximately following the diameter sequence exp({-1:10}/*2.25), with the smallest individual measuring 0.8 cm and the largest, 8.3 cm, *D130*, avoiding individuals with deformities or branches below 130 cm. We measured *DRC, D130* and *Hgt* using the same methods described for CW. Because the vining growth habit complicated the delineation of individual canopies, we did not measure canopy dimensions. We separated BP individuals into only two tissue types: (1) leaves, including petioles and (2) stems. We weighed each tissue type separately in the field to the nearest 100 g and gathered tissue subsamples following the same rules to minimize contamination and increase representativeness. We immediately processed BP tissue samples in the laboratory.

For the two native shrub seedlings, SS and WM, we measured nursery-grown individuals in 3.78l pots that spanned the size range used for revegetation. Because these species’ seedlings tended to grow with multiple stems, we collected different sets of measurements for individual plants, defined as all of the tissues in a single pot, and every stem defined as a single shoot emerging from the surface of the potting soil. At the level of the individual, we measured *Hgt* in cm as the vertical distance from the soil to the top of the canopy. We also measured one canopy diameter in an arbitrary direction, followed by a second perpendicular diameter at the midpoint of the first and calculated elliptical canopy area using the same formula that we applied to CW. At the level of each stem, we also measured *Hgt* in cm using a 5 m tape and *DRC* in cm to the nearest mm immediately above the soil.

For the shrubs, we separated individuals into two tissue types. For (1) stem wood, we cut stems at the point closest to the root flare or above the topmost root. We recorded the mass of the stems and branches then dried them at 65° C for three days and remeasured the dry mass. For (2) leaves, we collected all of the leaves for every individual plant and dried them at 65° C for three days and recorded the dry mass.

2.3 Biomass estimation

We estimated the moisture content of every tissue type for every individual of each species by weighing, drying and reweighing the entire tissue subsample. To control for the effects of transpiration and condensation during CW transportation and storage, we measured the wet mass of the entire bagged subsample, then dried both the poly bags and contained subsamples. We then subtracted the dry mass of the bag from the total mass of the bagged subsample for the original subsample wet mass. We dried the leaf and fruit subsamples in paper bags at 65°C for two days. We dried branch and trunk subsamples at 103°C for three days. Once dry, we reweighed subsamples and calculated the ratio of dry mass to wet mass (*sdmwmr*). We calculated the AGB of each individual by summing the products of the field measured wet mass for each tissue type present (i in $1\ldots n$) and the estimated *sdmwm* for the corresponding tissue subsample.

$$AGB = \sum_{i-1}^{n} TWM_i * sdmwmr_i \qquad \text{(Eq. 1)}$$

2.4 Wood density and Carbon Content

We measured wood density and C content of both wood and leaves for each species. At the Tidy Island site, we collected wood and leaf tissue samples from two mature individuals (i.e. *D130* > 7.5 cm) of both CW and BP. We collected trunk wood samples using an increment borer driven to approximately the midpoint of the trunk from a location below *D130.* We collected leaf tissue samples from accessible branches on the same individuals and placed them in polyethylene bags. Samples were transported to the lab and stored temporarily at 4°C.

For measuring wood density from trees, we split cores in half and measured the fresh volume of the inner and outer segments separately using the displacement method (Nicotra et al., 2010). We then dried the core segments at 103° C for three days and remeasured their dry mass. We calculated the wood density of each segment as the ratio of fresh volume to dry mass. For each individual we calculated the wood density as the weighted average of the inner and outer segments. Finally, for each species, we calculated wood density as the raw average for each individual.

We homogenized leaf and wood tissue sample using a Wiley Mill and physically aggregated homogenized subsamples for total carbon / nitrogen content analysis using a Thermo Electron Flash EA1112. We calculated mean C concentrations for each tissue and species as the simple average of collected samples.

To estimate the overall C content of aboveground biomass for each species, we first calculated the ratio of wood dry mass to leaf dry mass for every individual, excluding reproductive tissues when present. Then, for each species, we tested whether the wood:leaf ratio varied with stem diameter (*D130* for BP and CW and *DRC* of the largest stem for the pooled shrub individuals), using linear regression. No species exhibited significant size-based variation in aboveground tissue ratios (BP p = 0.06, CW p = 0.86, combined shrubs p = 0.10), so we calculated each species' aboveground C content as the mean of the tissue-mass weighted average C content of wood and leaves. We used a simple average of SS and WM for the combined category of shrubs. Trait measurements for all species are available in Supplementary Table S1.

2.5 Data analysis

2.5.1 Custom AGE fitting methods

We investigated three different statistical methods for identifying the most adequate form for AGEs. We conducted all analyses in R version 4.0.2 (R Core Team, 2020). The first method, which focuses on the single best linear model, referred to hereafter simply as "Top", involves identifying a linear model for the natural logarithm of the response measurement (i.e. AGB or tissue mass) over a set of candidate models consisting of combinations of non-destructive candidate measurements (e.g. *D130*, *Hgt*, *Can*, etc.). While this approach is widely applied for estimating AGEs (Picard et al., 2012), it can eliminate information from alternative models that may have very similar adequacies.

The second method, which involves inference across multiple models, referred to hereafter as "Multiple," addresses the major limitation of the first by weighting coefficients for all candidate predictors based on the cumulative information criterion scores. While multimodel inference has been widely applied in ecology (Grueber et al., 2011), it has been seldom used for estimating AGEs, potentially due to unclear methods for estimating weighted residual error variance, a necessary step for correcting for bias in lognormal regression.

The third method, Bayesian Multilevel modeling, referred to hereafter simply as “Multilevel,” combines strengths of both the first and second approaches in that it explicitly specifies the error variance for the lognormal distribution while simultaneously estimating parameters with their relative weights (Ogle and Barber, 2008). Furthermore, the Multilevel approach enables flexible models for the parameters themselves, including by pooling the error variance estimates from several small studies for a potentially more robust estimate of the global error variance (Price et al., 2009).

We fit AGEs using all three methods to two different datasets (Supplemental Table S2). The first dataset consisted of shoot wood mass measurements for every individual stem (56 total stems) as the response variable and corresponding non-destructive stem-level size measurements as covariates. The second dataset consisted of aboveground biomass measurements for every individual (40 total individuals) with applicable individual-level size measurements as covariates. The individual biomass dataset included only three individuals of WM so we pooled all individual level-measurements together for the aboveground biomass analysis of the two shrub species. Preliminary analyses indicated that different diameter measurements for the same stem were highly correlated (Pearson’s $r > 0.95$) so we included only one diameter measurement per stem in all analyses: *D130* for the tree species and *DRC* for the shrub species. We natural-log transformed all continuous covariates prior to estimating parameter values. For the purposes of comparing modeling approaches, we mean centered all log-transformed covariates.

To implement the first two AGE methods, Top and Multiple, we used an exhaustive search of candidate linear models defined by all possible combinations of covariates and ranked by the sample-size corrected Akaike Information Criterion (AICc, Hurvich and Tsai, 1989). We implemented the model search and weighting using the “glmulti” function in R Package “glmulti” (Calcagno and Mazancourt, 2010), specifying a natural log transformation of the response variable, all first order model terms for the candidate set, a linear model fit function and AICc as the information criterion. We identified the Top model by the lowest AICc. For all coefficient effects in the Top model as well as the error variance, we calculated the mean and 95% CI on the log-transformed scale using the “confint” function in R Package “stats”. We directly calculated the variance of the residuals and estimated 95% Cl as the ratio of the sum of the squared residuals to the 2.5% and 97.5% quantiles of a Chi-squared distribution with degrees of freedom equal to the degrees of freedom of the Top model. For the Multiple model, we characterized the AICc-weighted coefficients and corresponding 95% CI using the method proposed by Johnson and Omland (2004) as implemented by the function “coef” in R Package “glmulti”. We note that this method does not provide a straightforward estimate for the variance of the residuals, so it cannot predict unbiased values on the untransformed scale. Instead of using the Weighted model for prediction, we include these estimates and confidence intervals for the purpose of comparing coefficient estimates from all three methods.

We estimated the Multilevel model in a Bayesian context using Gibbs sampling as implemented using JAGS software through the R Package “rjags” (Plummer et al., 2022). We directly specified a log-normal likelihood for the response variable (i 1…n) for all species (*k* 1..m) with the log mean as a linear function of candidate predictors:

$$Mass_i \sim logNormal(\beta_0 + \beta_{dk} * CDia_i + \beta_{hk} * CHgt_i + \beta_{ck} * CCan_i + \beta_{sk} * CStem_i, \tau.) \quad \text{(Eq. 2)}$$

Where the β’s represent coefficients for the mean-centered log-transformed covariates for each species *k*, and τ. represents the precision, which is the inverse of the variance, σ^2. We considered two different models for the error variance, σ^2. The first gave every species an independent error variance (σ^2_k), which is equivalent to independent regressions, while the second pools all of the *i* observations for a single error variance.

For any covariate that was not applicable for a species or response variable, (e.g. *D130* for shrubs), we fixed the corresponding coefficient (i.e. β_{Dsh}) to zero. For all other coefficients, we estimated their value and weight simultaneously using a latent indicator variable approach (O’Hara and Sillanpää, 2009):

$$\beta_{.k} = \beta_{t_{.k}} * p_{.k} \quad \text{(Eq. 3)}$$

Where $\beta_{t_{.k}}$ is the latent unweighted coefficient value and $p_{.k}$ is a weight expressed as a proportion. We estimated each coefficient’s weight under a Bernoulli likelihood with a model-wide proportion of important coefficients (π):

$$p_{.k} \sim Bernoulli(\pi) \quad \text{(Eq. 4)}$$

We set vague priors for all regression and weighting coefficients. For the latent coefficient effects, $\beta_{t_{.k}}$, we set broad normal priors (mean 0, standard deviation 100). For the global proportion of important coefficients, π, we use a vague Beta (0.5,0.5) hyperprior. For the error standard deviations σ^2, we used broad (minimum 0, maximum 10) uniform priors.

We characterized the posterior using three independent Markov Chain Monte Carlo simulations with an adaptive burn-in of 2,000 iterations, followed by 200,000 samples, discarding every 20th sample to reduce memory storage requirements. We checked to ensure that the effective sample sizes for all parameters of interest were greater than 1000 using the “effectiveSize” function in R package “coda” (Plummer et al., 2020) and characterized the posterior mean and 95% CI using the R function “summary.” We then drew 20,000 additional samples to estimate the Deviance Information Criterion (DIC, Supplementary Table S3), which is an analog of AIC that is appropriate for Multilevel models (Spiegelhalter et al., 2002).

2.5.2 Model Comparisons

We quantified the accuracy and precision of two different custom and generic models for AGE in our dataset. The first custom model was the set of Top GLMs, which we used to generate predicted values on the original untransformed scale of the data:

$$\widehat{Mass}_{ik} = exp\left(\beta_0 + \beta_{.k} * \log(X_{.ik} - \overline{X_{.k}}) \ldots + \frac{\sigma_k^2}{2}\right) \qquad \text{(Eq. 5)}$$

Where the β's represent coefficient estimates from the Top GLM model for each species *k*, including every covariate *X* and the final term is the standard bias correction for lognormal regression as a function of the error variance σ^2. For projecting the Multilevel model, we included any covariate that occurred in a GLM with a weight greater than 20%, had an estimated importance value over 0.2 or had a posterior 95% CI that excluded 0. We refit the simplified Multilevel models using the same procedure described in 2.5.1 for models with either independent error variance or pooled error variance approaches and recalculated the DICs (Supplementary Table S3). We then applied Eq. 5 using the posterior mean estimates for the indicator-weighted simplified coefficients and the pooled error variance.

The first generic AGE required the only variable that was approximately measured for every individual in the dataset: Basal Diameter (*BDia*). The specific equation was proposed by Paul et al. (2016) for multistemmed shrubs from Australia and has the form:

$$\widehat{Mass_{\iota}} = 1.128 * exp\big((2.474 * \log(BDia_i)) - 2.757)\big) \qquad \text{(Eq. 6)}$$

For BP and CW we used the *DRC* for *BDia*. For the shrubs we used the cumulative diameter at the root collar for all stems. We note that the measurement location was slightly different from the measurements used to parameterize the Paul et al. (2016) model, diameter at 10 cm.

The second generic AGE approach applied two different models to the different growth forms in the dataset. For the small tree, CW, and the large shrub, BP, we used the global model proposed by Chave et al. (2014):

$$\widehat{Mass_{\iota}} = 0.0673 * \left(\rho + D130_i^2 * Hgt_i\right)^{0.976} \qquad \text{(Eq. 7)}$$

Where ρ represents the stem wood density estimate for the species (Supplementary Table S1).

For the shrub seedlings, we used the top global model proposed by Conti et al. (2019):

$$\widehat{Mass_{\iota}} = exp((-2.281 + 1.525 * \log(BDia_i) + 0.831 * \log(CD_i) + 0.523 * \log(Hgt_i)) \qquad \text{(Eq. 8)}$$

Where CD is the average canopy diameter.

For both custom and generic AGEs, we compared predicted to observed AGB values to quantify model accuracy and precision. To quantify accuracy, we calculated the Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\sum_{i=1}^{n}\left(Mass_i - \widehat{Mass_{\iota}}\right)^2 \frac{1}{n}} \qquad \text{(Eq. 9)}$$

To quantify bias, we calculated the average proportional difference between projected and observed values, relative to the observed values (Chave et al. 2016):

$$Bias = \sqrt{\sum_{i=1}^{n}\left(\frac{\widehat{Mass_{\iota}} - Mass_i}{Mass_i}\right)\frac{1}{n}} * 100\% \qquad \text{(Eq. 10)}$$

We calculated both RMSE and Bias for species-level datasets and for the overall dataset.

2.5.3 Model Effects

We compared AGB and C projections for two forest restoration projects. The first project occurred in the 2 ha dredge spoil mound at the Tidy Island Preserve. Focusing on how invasive removal and native revegetation influences C storage along the shore, the inventory consisted of 24 12.5 x 8 m monitoring plots arrayed every 25 m around the perimeter of the spoil mound. The center of each plot was located at mean high tide with the long axis perpendicular to the wrack line. Within the entire 100 m^2 plot, we identified every free standing tree greater than 7.5 cm at 130 cm height that was 50% rooted within the plot. Within a central 12.5 x 1 m transect we further identified every sapling greater than 2 cm at 130 cm height. We measured CW individuals using the same methodology used for parameterizing the AGEs. For BP, we measured stem diameter at 130 and 140 cm only.

We quantified the proportion of total stems and basal area represented by the two focal species, BP and CW, in the upland 50 m^2 portion of all 24 plots. For the best estimates from a customized model, we projected AGB and C using Eq 5. using posterior mean parameter estimates from the simplified Multilevel model with pooled errors (e.g Eq. 2-4, Supplementary Table S4). For the best estimates from a generic model, we projected AGB and C using Eq. 7 as originally estimated by Chave et al (2014). This equation requires individual height, which we did not measure for BP in the inventory because stems from many individuals overlapped in thickets where they grew. Instead, we substituted the 75th percentile of individual heights from our AGE dataset (4.4 m), which was approximately the same height as the BP thicket. Otherwise, we excluded two individuals of CW that had survived a blow down during a storm and had no canopy measurements. We converted AGB estimates from both AGE approaches to C concentration using species-level aboveground C concentrations (i.e. Section 2.4, Supplementary Tables S1, S4) and then scaled plot level concentrations to Mg C ha^{-1}. Finally, we tested whether different AGEs would generate significantly different C concentration estimates under realistic survey conditions using paired *t*-tests for each focal species by size class and their combined C concentrations across all 24 inventory plots.

The second project involved native afforestation over turf in a dredge spoil site fringed by the same invasive plants. The goal for the inventory was establishing a baseline C concentration for calculating sequestration during native regrowth. We recorded data for a total of 53 multiscale inventory plots on 5 parallel transects, with the three longest measuring 161 m and including 11 plots and two shorter transects with 10 plots each in between. We recorded aboveground biomass measurements for plots with more than 50% planted coverage for a total of 33 3 m radius plots surveyed for seedling biomass. For each woody seedling, we measured the *DRC* of the largest stem using calipers, the height using a 5 m tape, the length of the canopy going parallel to the transect using a 5 m tape, and the length of the canopy going perpendicular to the transect using a 5 m tape. We also recorded the number of stems. *DRC* was recorded immediately above the soil. To determine the height we recorded the highest point of the individual starting from the root collar and moving along the main stem.

For the best projections from a customized model, we again used posterior mean estimates from the simplified Multilevel model as estimated by Eq. 2-5 for both shrub species combined (Supplementary Table S4). For the best estimates from a customized generic model, we used Eq. 8 as originally proposed by Conti et al. (2019). Because this equation requires basal diameter, whereas we had only measured the diameter at the root collar of the largest stem and the total number of stems, we estimated basal area for shrubs in the inventory using a regression fitted to the AGE dataset. Specifically, we use the R package “glmulti” to estimate weighted coefficients for total basal diameter from every first order model including both predictors, largest stem diameter at the root collar and total stem number. We then use the R function “predict” to estimate weighted values using the method proposed by Johnson and Omland (2004) for all individuals in the inventory dataset (Supplementary Table S4). We calculated the mean canopy diameter from both diameter measurements we collected in the field. We tested for significant differences in C concentration estimates between AGE approaches using paired *t*-tests for all 33 inventory plots.

3 Results

3.1 Different methods generate similar covariate effects

Aboveground biomass and stem wood mass varied with different sets of non-destructive covariates (Supplementary Table S4). All of the Top models for AGB included one or two covariates. For BP, the top model for AGB included only ln(*D130*) and contributed more than 84% of the cumulative model weight. The Top model for the small invasive tree, CW, included both ln(*D130*) and ln(*Can*) but contributed a smaller proportion of the model weight (53%), with an additional 30% of model weight contributed by a simpler model with only ln(*D130*). For the two native shrubs measured as seedlings, ln(*Hgt*) was the only covariate included in the Top model, but a model that added an effect for ln(*DRC*) also contributed more than 20% of the model weight. Models for wood mass by stem included similar sets of covariates and relative weights, with subtle differences, including the Top model for CW stem wood mass having only ln(*D130*) as a covariate and the disaggregated data for the native shrub SS including ln(*DRC*) in addition to ln(*Hgt*).

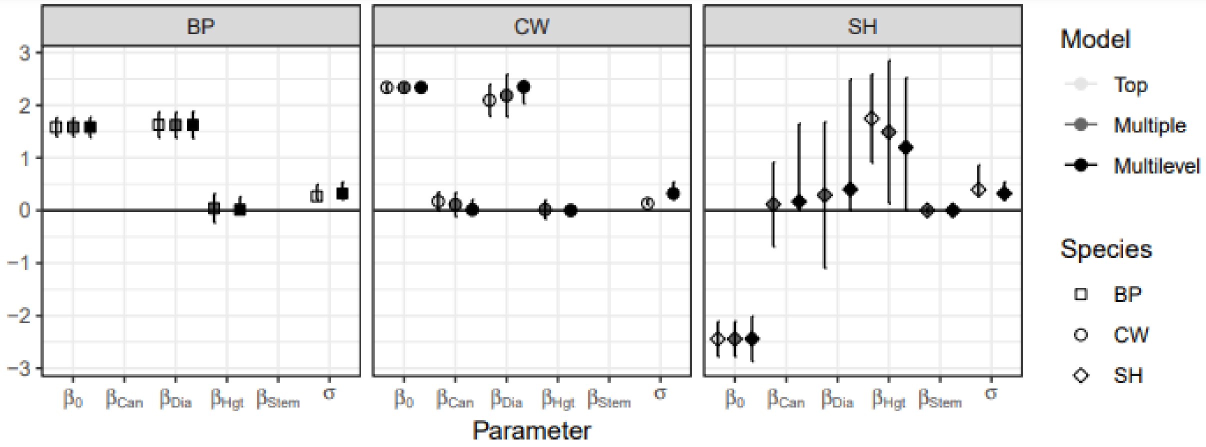


Figure 1: Aboveground biomass allometric growth parameter estimates for *Schinus terebinthifolia* (BP), *Cupaniopsis anacardioides* (CW) and two native shrubs (SH) generated either by the single GLM with the lowest AICc (Top), multimodel inference across all possible GLMs weighted by their AICc (Multiple)

or a Bayesian Multilevel model with latent indicator weighting and pooled measurement error (Multilevel). Covariates were log transformed and mean centered prior to estimation. Error bars represent 95% Confidence Intervals. Parameters for covariates that were excluded or unestimable are not depicted.

Parameter estimates from the Top, Multiple and Multilevel methods were very similar (Figure 1, Supplementary Figure S1). All three methods generated intercept estimates that were within 1% of one another, although the CIs for the Multilevel method tended to be narrower than either GLM-based method. Estimated covariate effects were similar, but with noticeable differences between methods. Covariate effects that did not occur in the Top GLM generally had weighted estimates with 95% CIs that included 0. Otherwise, estimates varied with the concentration of model weight. For BP, which had very concentrated model weight in a simple model, the parameter estimates for ln(*D130*) were very similar (<1% absolute difference) between all methods with nearly identical 95% CIs. For CW, which had a larger set of models contributing weight, the parameter estimates for the effect of ln(*D130*) varied by more than 10% between methods and estimates for the effect of ln(*Can*) varied by almost 100%. However, the 95% confidence intervals broadly overlapped and all three methods generated 95% CIs that included 0. Variation in estimates for the native shrubs were more variable still, with up to 20% absolute differences for the effect of ln(*Hgt*) and much broader 95% CIs. Finally, the pooled estimate for the error standard deviation from the Multilevel model had a much smaller 95% CI than the corresponding estimates from the Top GLMs, and was lower in magnitude for both BP and SH, which had fewer samples. Variation in parameter estimates and uncertainty were similar for models of stem wood dry mass (Supplementary Figure S1). Disaggregated stem wood data for the two shrub species had much narrower and non-overlapping covariate 95% CIs than the corresponding covariates for individual-level shrub AGB.

3.2 Custom models have higher accuracy and lower bias than generic methods

Custom AGEs generated AGB predictions that were more accurate and precise than generic AGEs (Table 1, Table 2, Figure 2). Model accuracy, as measured using RMSE, also decreased with increasing plant size, with the average errors differing by orders of magnitude between very small SH, intermediate BP and relatively large CW. The simplified Multilevel model performed best overall, with slightly lower RMSE and Bias than the set of Top GLM models. The Multilevel model made the most accurate predictions for CW AGB, which included the largest individuals in the study, and the least biased predictions for SH AGB, which was the group with the fewest individuals sampled. Otherwise, both custom models performed similarly and much better than either generic model. The AGE proposed by Paul et al. (2016), which was the only generic model suitable for all individuals in our dataset based on the measured covariates, performed most poorly, deviating from observed CW AGB by over 40 kg on average and producing estimates of SH AGB that were more than five times larger than actually observed. Compared to using a single AGE for all growth forms, using different generic models for trees (i.e. BP and CW using Chave et al. 2014) versus shrubs (i.e. SS and WM using Conti et al. 2019) produced more accurate biomass predictions, except for BP.

Table 1: Root Mean Squared Error (kg) of aboveground biomass predictions from four allometric growth equations for Brazilian Peppertree (BP), Carrotwood (CW) and two native shrubs (SH). Custom Top and Multilevel models are described in 2.5.1. The Generic Multistem equation was proposed by Paul et al. 2016 and the Generic Tree and Shrub equations were proposed by Conti et al. 2019 and Chave et al. 2014 respectively.

Model	**BP**	**CW**	**SH**	**Overall**
Custom Top	3.73	16.16	0.05	11.33
Custom Multilevel	3.75	14.05	0.04	10.57
Generic Multistem	4.23	43.16	0.61	29.84
Generic Tree + Shrub	8.07	35.07	0.07	24.57

Table 2: Bias (%) of aboveground biomass predictions from four allometric growth equations for Brazilian pepper (BP), Carrotwood (CW) and two native shrubs (SH). Custom Top and Multilevel models

are described in 2.5.1. The Generic Multistem equation was proposed by Paul et al. 2016 and the Generic Tree and Shrub equations were proposed by Conti et al. 2019 and Chave et al. 2014 respectively.

Model	**BP**	**CW**	**SH**	**Overall**
Custom Top	6.87	1.68	15.47	6.34
Custom Multilevel	6.39	4.07	7.43	5.52
Generic Multistem	-28.73	24.80	529.00	122.19
Generic Tree + Shrub	-66.22	-20.49	63.64	-15.27

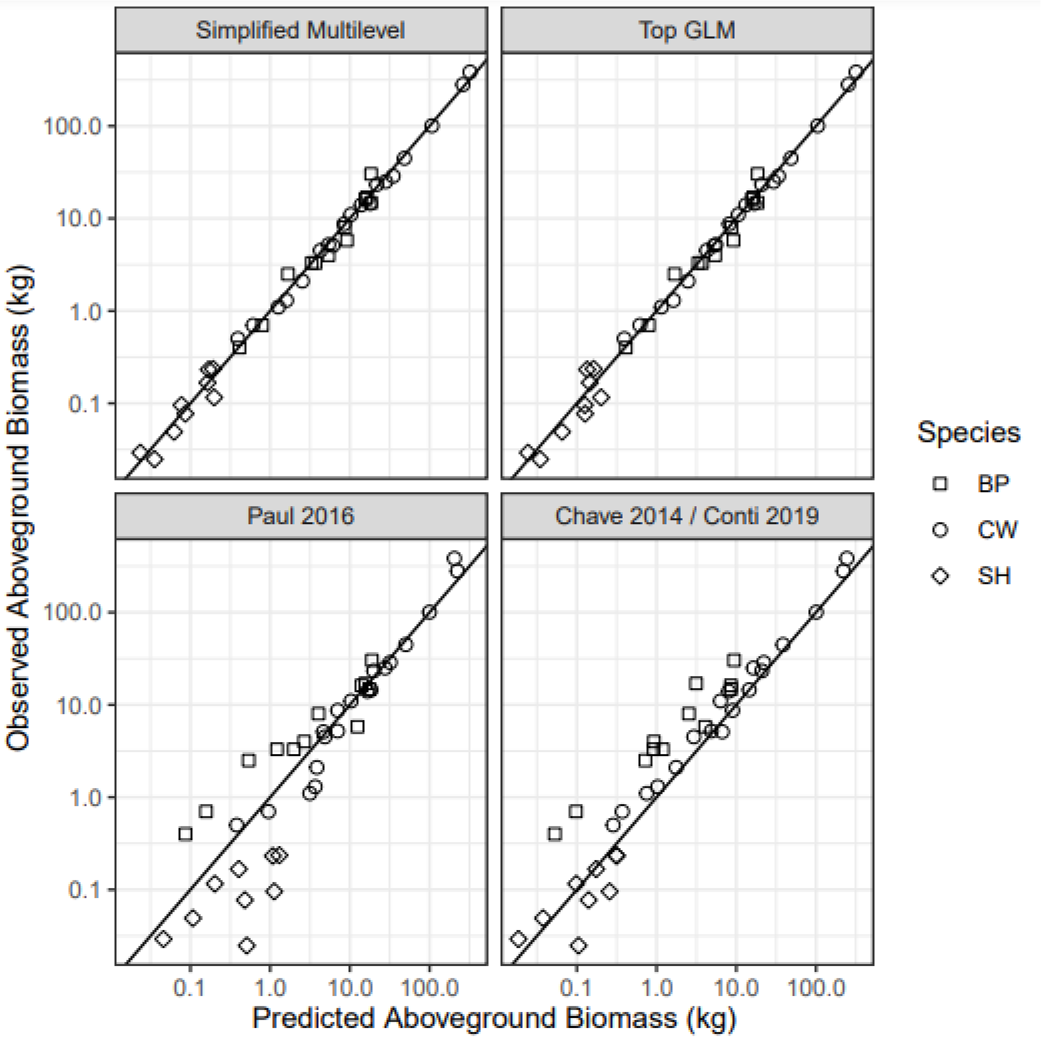


Figure 2: Two methods for generating custom aboveground biomass allometric growth equations, estimated as GLMs with the lowest AICc (Top) or a simplified Bayesian Multilevel model (Multilevel), outperform two generic approaches for plants of different size and growth form, either Paul et al. 2016 for all individuals or Chave et al. 2014 for trees combined with Conti et al. 2019 for shrubs. Axes are log 10

transformed to facilitate comparison across invasive small trees (CW), large shrubs (BP) and native shrub seedlings (SH).

3.3 Custom models generate significantly different estimates of baseline C before invasive removal and after replanting

Differences in accuracy and precision between custom and generic AGEs produced divergent estimates of aboveground C concentrations before and after forest restoration. At the pre-restoration location, either BP or CW was present in 19/24 0.05 ha inventory plots. Both focal species collectively accounted for 11.9% of tree basal area (>7.5 cm *D130*) and 44% of sapling basal area (2cm < *D130* < 7.5 cm). The remaining basal area was represented by species with existing AGEs that we did not include in our comparison. The Multilevel AGEs generated consistently higher estimates of aboveground C concentrations before removal (Table 3). Across both species and size classes combined, the Custom Multilevel AGE estimated 32% more C in invasive aboveground biomass than the best-performing Generic AGE for trees (i.e. Chave et al. 2014, difference = 2.81 Mg ha^{-1} of C, paired *t*-test, n = 24, t = 2.64, p = 0.014). Divergence between estimates increased for smaller individuals, with the Custom AGEs estimating significantly higher C concentrations for both CW saplings (difference = 0.38 Mg ha^{-1} of C, paired *t*-test, n = 24, t = 2.63, p = 0.015) and BP adults (difference = 0.88 Mg ha^{-1} of C, paired *t*-test, n = 24, t = 2.34, p = 0.029). Custom Multilevel AGEs also generated higher estimates for CW trees, which represented the majority of aboveground C for these two species, and BP saplings, which had the most divergent estimates, but estimates were not significantly different (p = 0.21 and p = 0.18 respectively).

Table 3: Aboveground C Density (average Mg ha^{-1} [s.e.]) estimated using either a custom Multilevel allometric growth equation or a generic model for trees (Chave et al. 2014) for Brazilian Peppertree (BP) and Carrotwood (CW) individuals in two size classes, saplings (2 - 7.5 cm DBH) and trees (>7.5 cm DBH), as encountered in 24 0.05 ha inventory plots of pre-restoration conditions at a preserve in southwest Florida. Asterices indicate categories for which the average C concentration differed significantly between estimation methods.

Method	**BP Sapling**	**BP Tree***	**CW Sapling***	**CW Tree**	**Combined***
Custom Multilevel	1.59 [1.14]	1.91 [0.84]	1.71 [0.64]	6.33 [2.43]	11.53 [3.32]
Generic Tree	0.69 [0.49]	1.03 [0.46]	1.33 [0.51]	5.67 [2.04]	8.72 [2.46]

At the second forest restoration site, where native seedlings had just been planted, Custom and Generic AGE methods also generated different C concentration estimates, but in the opposite direction. Based on 876 woody seedlings measured in 33 3 m radius inventory plots, the Custom Multilevel method estimated 12.5% less C in aboveground biomass compared to the best performing Generic AGE for shrubs (i.e. Conti et al. 2018, difference = -0.008 Mg ha^{-1} of C, paired *t*-test, n = 33, $t = -2.586$, p = 0.015). When scaled up to the entire 0.25 ha project area, the Custom Multilevel AGE estimated 0.070 Mg C (± 0.004 S.E.) in aboveground biomass compared to 0.080 Mg C (± 0.005 S.E.) estimated using the Generic AGE.

4 Discussion

For monitoring C in forest restoration projects that replace established invasives with native seedlings, Multilevel allometric growth equations can address important limitations to conventional approaches. Compared to generic AGEs, which perform differently across size classes and growth forms (Conti 2019), custom AGEs provided more accurate AGB estimates, especially for small plants and species with unusual growth forms. In two forest restoration projects that controlled invasive trees by replacing them with native seedlings, the choice of AGE significantly influenced estimates of baseline conditions. Compared to traditional statistical methods that identify the single top model for each species, multilevel methods, which borrow strength from sampling among species to reduce project-wide error and bias, can facilitate sampling more species with improved performance.

4.1 Custom allometric growth equations outperform generic methods for monitoring biomass during invasive tree control

New AGEs for two impactful invasive species, Brazilian Peppertree and Carrotwood, not only fill gaps in growing species-level AGE databases (Henry et al., 2013; McPherson et al., 2016), but also reinforce known limitations of generic AGEs for projecting aboveground biomass (Conti et al., 2019). Existing global models, while invaluable in the absence of species-specific information, focus on mature trees of intermediate size (Chave et al., 2014). Consequently, they may deviate more for both very large individuals, which contribute the most biomass (Gonzalez de Tanago et al., 2018), as well as the smallest individuals, including seedlings and shrubs. In an analysis of worldwide patterns of shrub biomass, Conti et al. (2019) found that several global models underestimated the biomass of the largest individuals by 50% while overestimating the biomass of the smallest individuals by a similar magnitude. We observed similar patterns in both our training dataset and baseline inventories. Two generic models underestimated the biomass of species with larger individuals (BP and CW) while greatly overestimating the biomass of shrub seedlings. When compared to more accurate and precise custom AGEs, aboveground biomass projections from generic AGEs underestimated established invasive biomass and overestimated aboveground biomass of native shrub seedlings. Consequently, generic AGEs would project much shorter times for ecosystem C recovery following invasive removal and native replanting than custom AGEs. In situations where both the quantity and timing of C storage are important, such as restoration funded by C finance, custom AGEs may provide needed accuracy.

The limitations of generic AGEs for both small individuals as well as species with unusual growth compounded in the case of Brazilian Peppertree. Typically classified as a shrub due to its relatively short stature and tendency to produce multiple stems below 130 cm, BP alters its growth when crowded, producing longer stems that scramble over adjacent vegetation, somewhat like a liana (Spector and Putz, 2006). Growing this way intertwines the canopies of adjacent individuals, complicating height and canopy measurements that are required by several generic AGEs. Even with all necessary measurements, unusual growth may explain why generic methods underperformed in the training dataset. The Chave et al. (2014) model, which was parameterized for trees with a single upright stem, underestimated AGB of BP by 65%. Some of this "missing" biomass may occur in BP's sprawling horizontal growth. To the extent that other invasive trees succeed because of unusual growth forms (Callaway and Ridenour, 2004) or idiosyncratic genetic histories (Erickson et al., 2017; Mukherjee et al., 2012), site- and species-specific AGEs may improve accuracy for these widespread and impactful plants.

4.2 Strengths, weakness and possible extensions of Multilevel methods for allometric growth equations

In demonstrating improved performance for custom AGEs, our results reinforce previous analyses. However, generating quality AGEs requires considerable effort, expense and expertise (Picard et al., 2012). Best practices modeled on standard statistical approaches recommend harvesting more than 30 individuals across the full size range for every species, which may be prohibitively expensive where uncharacterized species are prevalent and individuals can attain great size (Roxburgh et al., 2015). We present an alternative statistical approach that can facilitate a more efficient sampling design by simultaneously estimating AGEs for multiple species. Multilevel Bayesian modeling borrows strength from sampling across species to constrain project-wide error and generate more accurate and precise projections, especially for species represented by relatively few samples (Price et al., 2009). No single species in our study met the individual sampling threshold recommended for conventional analysis. The corresponding Top GLM method generated error estimates that tracked each species group's sampling intensity. However, by pooling error variation across all species in the analysis, the Multilevel method generated error estimates that were lower and more precise for species with relatively fewer samples.

While pooling error in a multilevel model increased overall performance, detailed differences in accuracy and precision between responses, species and methods indicated some limitations. The strongest pooling that we employed was for estimating aboveground biomass of two shrub species. With six and three individuals sampled, neither SS nor WM had more individuals than candidate covariates, precluding statistical analysis of individual AGB by species (Roxburgh et al., 2015). However, when analyzing wood mass per stem, which involved more samples and fewer covariates, the estimated effects of stem diameter and height differed significantly between species, indicating different patterns of stem wood allocation. Pooling both species combined variation within and between species, which likely contributed to the larger 95% CIs for covariate effects in the individual-level AGB analysis. Even so, the pooled Multilevel method generated more accurate and precise estimates of shrub AGB than either generic model, which necessarily pool variation from even more species.

A more subtle example of the effects of pooling comes from contrasting the bias estimates between custom methods for shrubs versus Carrotwoods. The overall reduction in bias for the Multilevel model compared to the Top model reflected much more precise estimates for 9 individual shrubs compensating for a slight increase in bias for 19 individual Carrotwoods. More consistent precision across groups may reflect the role of sampling error in the bias correction for log-normal models. The standard correction adds a fraction of the estimated error variance to the log linear equation for the mean (Clifford et al., 2013). High error variance associated with small sample sizes would tend to increase the magnitude of this correction factor and therefore the magnitude of upward bias correction in groups with fewer individuals. Pooling error estimates across species in the multilevel model moderated the differences in bias correction, possibly contributing to the slight overestimation of CW biomass in the training dataset. Alternative bias corrections may have different sensitivity to sampling error and could provide a valuable avenue for future research (Clifford et al., 2013), as would estimation methods that are less sensitive to bias, such as weighted regression (Parresol, 2001).

Even though Multilevel methods cannot overcome fundamental sampling limitations, the flexibility of the approach provides promising avenues for future development. One technique applied here, latent variable weighting, simultaneously estimates covariate effects and their weights across models with every potential combination of terms (O’Hara and Sillanpää, 2009). Similar to multimodel inference for GLMs, latent variable weighting may have boosted the accuracy of carrotwood AGB relative to the single Top model, which ignored information from other models with similar adequacy scores. Another potential extension for studies involving even more species is hierarchical parameter models. Analogous to random effects in GLMs, hierarchical models interpolate between complete pooling, which we employed for both shrub species, and no pooling, which corresponds to the conventional approach. Doing so may increase the power to characterize subtle differences between species with relatively few samples (Price et al., 2009). Furthermore, hierarchical models for both parameters and error can involve phylogenetic structure, further boosting accuracy (Oberle et al., 2016). Phylogenetic effect structures may also enable model-based projections for related species based on evolutionary rates, rather than the taxonomic aggregation and substitution advocated by some databasing efforts (McPherson et al., 2016). Finally, implementing multilevel models in a Bayesian context facilitates incorporating previously estimated equations through informative priors even when the original data are unavailable (Zapata-Cuartas et al., 2012). For all of these reasons, we believe that Bayesian Multilevel methods provide a promising area for developing powerful, flexible AGEs.

5 Conclusions

Ultimately, the value of multilevel models and custom allometric growth equations depends on the research questions and restoration objectives at hand. For many purposes, including describing the stand structure and C dynamics of well-characterized boreal and temperate forests, existing models provide outstanding performance. However, in regions with less research and more tree species, but equally urgent goals to control invasive species while maximizing C storage, more precise estimation methods are necessary. Remotely sensed data may one day provide the geographic, temporal and taxonomic resolution necessary for accurate forest restoration monitoring (Gonzalez de Tanago et al., 2018). In the meantime, stem-based surveys combined with custom multilevel allometric growth equations can leverage the strength from sampling across species to improve projections across size classes and growth forms.

Author Contributions

Brad Oberle: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing-Original draft preparation, Writing-review and editing. **Piper Cole:** Methodology, Investigation, Writing- Reviewing and Editing. **Garcia Frank:** Investigation, Data curation, Writing- Reviewing and Editing. **Alexandra Gates:** Conceptualization, Investigation, Data curation, Writing- Reviewing and Editing. **Brittney Hall:** Conceptualization, Investigation, Data curation, Writing-Original draft preparation, Writing- Reviewing and Editing. **Deric Harvey:** Methodology, Investigation, Writing- Reviewing and Editing. **Melody Scott:** Conceptualization, Investigation, Data curation, Writing-Original draft preparation, Writing- Reviewing and Editing. **Cas Setterberg:** Methodology, Investigation, Writing- Reviewing and Editing. **Simon Bustetter:** Methodology, Investigation, Data curation, Writing- Reviewing and Editing.

Acknowledgements

This project has been funded wholly or in part by the United States Environmental Protection Agency under assistance agreement [MX 00D680-18/6207] to New College of Florida. The contents of this document do not necessarily reflect the views and policies of the Environmental Protection Agency, nor does the EPA endorse trade names or recommend the use of commercial products mentioned in this document. Funding for shrub biomass and the baseline inventory of the Colony Cove afforestation project was provided by Equity Lifestyles Inc. We thank Charles Reith and Zach Zildjian for facilitating the Colony Cove project. We also thank many New College of Florida students who conducted destructive sampling and laboratory analysis of Brazilian pepper as part of the 2018 and 2020 Forest Ecology Laboratory. Dorn Martell and Tom Smith assisted with data collection.

Data statement

All data necessary to reproduce the fitted allometric growth equations will be made available in a public repository (globallometry.org) upon acceptance of the manuscript.

References

Armijos, C., Valarezo, E., Cartuche, L., Zaragoza, T., Finzi, P.V., Mellerio, G.G., Vidari, G., 2018. Chemical composition and antimicrobial activity of Myrcianthes fragrans essential oil, a natural aromatizer of the traditional Ecuadorian beverage colada morada. J. Ethnopharmacol. 225, 319–326. https://doi.org/10.1016/j.jep.2018.07.018

Calcagno, V., Mazancourt, C. de, 2010. glmulti: An R Package for Easy Automated Model Selection with (Generalized) Linear Models. J. Stat. Softw. 34, 1–29. https://doi.org/10.18637/jss.v034.i12

Callaway, R.M., Ridenour, W.M., 2004. Novel Weapons: Invasive Success and the Evolution of Increased Competitive Ability. Front. Ecol. Environ. 2, 436–443. https://doi.org/10.2307/3868432

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. Glob. Change Biol. 20, 3177–3190. https://doi.org/10.1111/gcb.12629

Clifford, D., Cressie, N., England, J.R., Roxburgh, S.H., Paul, K.I., 2013. Correction factors for unbiased, efficient estimation and prediction of biomass from log–log allometric models. For. Ecol. Manag. 310, 375–381. https://doi.org/10.1016/j.foreco.2013.08.041

Conti, G., Gorné, L.D., Zeballos, S.R., Lipoma, M.L., Gatica, G., Kowaljow, E., Whitworth-Hulse, J.I., Cuchietti, A., Poca, M., Pestoni, S., Fernandes, P.M., 2019. Developing allometric models to predict the individual aboveground biomass of shrubs worldwide. Glob. Ecol. Biogeogr. 28, 961–975. https://doi.org/10.1111/geb.12907

Cuda, J.P., Ferriter, A.P., Manrique, V., Medal, J.C., n.d. Recommendations from the Brazilian Peppertree Task Force Florida Exotic Pest Plant Council.

Dickie, I.A., Bennett, B.M., Burrows, L.E., Nuñez, M.A., Peltzer, D.A., Porté, A., Richardson, D.M., Rejmánek, M., Rundel, P.W., van Wilgen, B.W., 2014. Conflicting values: ecosystem services and invasive tree management. Biol. Invasions 16, 705–719. https://doi.org/10.1007/s10530-013-0609-6

Duncanson, L., Rourke, O., Dubayah, R., 2015. Small Sample Sizes Yield Biased Allometric Equations in Temperate Forests. Sci. Rep. 5, 17153. https://doi.org/10.1038/srep17153

Dunevitz, V., Ewel, J., 1981. Allelopathy of wax myrtle (*Myrica cerifera*) on *Schinus terebinthifolius*. Fla. Sci. 44.

Eggleston, H.S., Buendia, L., Miwa, K., Ngara, T., Tanabe, K., 2006. 2006 IPCC Guidelines for National Greenhouse Gas Inventories.

Enloe, S.F., Langeland, K.A., n.d. SS-AGR-165/AG111: Invasive Plants in Natural Area Weeds: Carrotwood (Cupaniopsis anacardioides) [WWW Document]. URL https://edis.ifas.ufl.edu/publication/AG111 (accessed 3.7.23).

Erickson, K.D., Pratt, P.D., Rayamajhi, M.B., Horvitz, C.C., 2017. Introduction History Influences Aboveground Biomass Allocation in Brazilian Peppertree (Schinus terebinthifolius). Invasive Plant Sci. Manag. 10, 247–253. https://doi.org/10.1017/inp.2017.23

FitzJohn, R.G., Pennell, M.W., Zanne, A.E., Stevens, P.F., Tank, D.C., Cornwell, W.K., 2014. How much of the world is woody? J. Ecol. 102, 1266–1272. https://doi.org/10.1111/1365-2745.12260

Gonzalez de Tanago, J., Lau, A., Bartholomeus, H., Herold, M., Avitabile, V., Raumonen, P., Martius, C., Goodman, R.C., Disney, M., Manuri, S., Burt, A., Calders, K., 2018. Estimation of above-ground biomass of large tropical trees with terrestrial LiDAR. Methods Ecol. Evol. 9, 223–234.

https://doi.org/10.1111/2041-210X.12904

Gordon, D.R., 1998. Effects of Invasive, Non-Indigenous Plant Species on Ecosystem Processes: Lessons from Florida. Ecol. Appl. 8, 975–989. https://doi.org/10.1890/1051-0761(1998)008[0975:EOINIP]2.0.CO;2

Grueber, C.E., Nakagawa, S., Laws, R.J., Jamieson, I.G., 2011. Multimodel inference in ecology and evolution: challenges and solutions. J. Evol. Biol. 24, 699–711. https://doi.org/10.1111/j.1420-9101.2010.02210.x

Henry, M., Bombelli, A., Trotta, C., Alessandrini, A., Birigazzi, L., Sola, G., Vieilledent, G., Santenoise, P., Longuetaud, F., Valentini, R., Picard, N., Saint-André, L., 2013. GlobAllomeTree: international platform for tree allometric equations to support volume, biomass and carbon assessment. IForest Biogeosciences For. 6, E1. https://doi.org/10.3832/ifor0901-006

Hurvich, C.M., Tsai, C.-L., 1989. Regression and time series model selection in small samples. Biometrika 76, 297–307. https://doi.org/10.1093/biomet/76.2.297

Johnson, J.B., Omland, K.S., 2004. Model selection in ecology and evolution. Trends Ecol. Evol. 19, 101–108. https://doi.org/10.1016/j.tree.2003.10.013

Liao, C., Peng, R., Luo, Y., Zhou, X., Wu, X., Fang, C., Chen, J., Li, B., 2008. Altered ecosystem carbon and nitrogen cycles by plant invasion: a meta-analysis. New Phytol. 177, 706–714. https://doi.org/10.1111/j.1469-8137.2007.02290.x

Lockhart, C.S., Austin, D.F., Jones, W.E., Downey, L.A., 1999. Invasion of Carrotwood (Cupaniopsis anacardioides) in Florida Natural Areas (USA). Nat. Areas J. 19, 254–262.

Marchante, E., Kjøller, A., Struwe, S., Freitas, H., 2009. Soil recovery after removal of the N2-fixing invasive Acacia longifolia: consequences for ecosystem restoration. Biol. Invasions 11, 813–823. https://doi.org/10.1007/s10530-008-9295-1

McPherson, E.G., Doorn, N.S. van, Peper, P.J., 2016. Urban tree database and allometric equations. Gen Tech Rep PSW-GTR-253 Albany CA US Dep. Agric. For. Serv. Pac. Southwest Res. Stn. 86 P

253. https://doi.org/10.2737/PSW-GTR-253

Mukherjee, A., Williams, D.A., Wheeler, G.S., Cuda, J.P., Pal, S., Overholt, W.A., 2012. Brazilian peppertree (Schinus terebinthifolius) in Florida and South America: evidence of a possible niche shift driven by hybridization. Biol. Invasions 14, 1415–1430. https://doi.org/10.1007/s10530-011-0168-7

Nicotra, A., Sack, L., Santiago, L., Cornwell, W., 2010. PrometheusWiki: Protocols, methods, explanations and updated standards for ecological and environmental plant physiology.

Oberle, B., Breithaupt, J., McTigue, A.M., Stryker, R., Cladas, M., Raulerson, G., Young, D.F., 2022. Restoration objectives create surface carbon cycle trade-offs in coastal habitats. Restor. Ecol. 30, e13563. https://doi.org/10.1111/rec.13563

Oberle, B., Ogle, K., Zuluaga, J.C.P., Sweeney, J., Zanne, A.E., 2016. A Bayesian model for xylem vessel length accommodates subsampling and reveals skewed distributions in species that dominate seasonal habitats. J. Plant Hydraul. 3, e003–e003. https://doi.org/10.20870/jph.2016.e003

Ogle, K., Barber, J.J., 2008. Bayesian Data—Model Integration in Plant Physiological and Ecosystem Ecology, in: Lüttge, U., Beyschlag, W., Murata, J. (Eds.), Progress in Botany, Progress in Botany. Springer, Berlin, Heidelberg, pp. 281–311. https://doi.org/10.1007/978-3-540-72954-9_12

O’Hara, R.B., Sillanpää, M.J., 2009. A review of Bayesian variable selection methods: what, how and which. Bayesian Anal. 4, 85–117. https://doi.org/10.1214/09-BA403

Parresol, B.R., 2001. Additivity of nonlinear biomass equations. Can. J. For. Res. 31, 865–878. https://doi.org/10.1139/x00-202

Paul, K.I., Roxburgh, S.H., Chave, J., England, J.R., Zerihun, A., Specht, A., Lewis, T., Bennett, L.T., Baker, T.G., Adams, M.A., Huxtable, D., Montagu, K.D., Falster, D.S., Feller, M., Sochacki, S., Ritson, P., Bastin, G., Bartle, J., Wildy, D., Hobbs, T., Larmour, J., Waterworth, R., Stewart, H.T.L., Jonson, J., Forrester, D.I., Applegate, G., Mendham, D., Bradford, M., O’Grady, A., Green, D., Sudmeyer, R., Rance, S.J., Turner, J., Barton, C., Wenk, E.H., Grove, T., Attiwill, P.M., Pinkard, E., Butler, D., Brooksbank, K., Spencer, B., Snowdon, P., O’Brien, N., Battaglia, M., Cameron, D.M., Hamilton, S., McAuthur, G., Sinclair, J., 2016. Testing the generality of above-ground biomass allometry across plant functional types at the continent scale. Glob. Change Biol. 22, 2106–2124. https://doi.org/10.1111/gcb.13201

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

Picard, N., Saint-André, L., Henry, M., 2012. Manual for building tree volume and biomass allometric equations: from field measurement to prediction. Man. Build. Tree Vol. Biomass Allometric Equ. Field Meas. Predict. FAOFood Agric. Organ. United Nations2012.

Plummer, M., Best, N., Cowles, K., Vines, K., Sarkar, D., Bates, D., Almond, R., details, A.M. coda author, 2020. coda: Output Analysis and Diagnostics for MCMC.

Plummer, M., Stukalov, A., Denwood, M., 2022. rjags: Bayesian Graphical Models using MCMC.

Price, C.A., Ogle, K., White, E.P., Weitz, J.S., 2009. Evaluating scaling models in biology using hierarchical Bayesian approaches. Ecol. Lett. 12, 641–651. https://doi.org/10.1111/j.1461-0248.2009.01316.x

R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Reynolds, S.T., 1985. Sapindaceae, in: Flora of Australia, Volume 25, Melianthaceae to Simaroubaceae. Australian Government Publishing Service, Canberra, pp. 4–163.

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. https://doi.org/10.1890/ES14-00251.1

Spector, T., Putz, F.E., 2006. Biomechanical Plasticity Facilitates Invasion of Maritime Forests in the southern USA by Brazilian pepper (Schinus terebinthifolius). Biol. Invasions 8, 255–260. https://doi.org/10.1007/s10530-004-5571-x

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A., 2002. Bayesian measures of model complexity and fit. J. R. Stat. Soc. Ser. B Stat. Methodol. 64, 583–639. https://doi.org/10.1111/1467-9868.00353

Van Deelen, T.R., 1991. Morella cerifera [WWW Document]. Morella Cerifera. URL https://www.fs.usda.gov/database/feis/plants/shrub/morcer/all.html (accessed 3.7.23).

Ward, D.B., Austin, D.F., Coile, N.C., 2003. Endangered and Threatened Plants of Florida, Ranked in Order of Rarity. Castanea 68, 160–174.

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. https://doi.org/10.1038/23251

Zapata-Cuartas, M., Sierra, C.A., Alleman, L., 2012. Probability distribution of allometric coefficients and Bayesian estimation of aboveground tree biomass. For. Ecol. Manag. 277, 173–179. https://doi.org/10.1016/j.foreco.2012.04.030

Supplementary Material

Supplementary Table S1: Trait data for two Florida invasive species *Schinus terebinthifolia* (BP), *Cupaniopsis anacardioides* (CW), and two natives widely used in revegetation, *Myrcianthes fragrans* (Simpson's stopper) and *Morella cerifera* (WM), including one value for both native shrub species pooled (SH).

Taxon	**Trait (unit)**	**Value**
BP	Wood Density (g/cm^3)	0.5087
BP	Wood C (%)	46.71
BP	Leaf C (%)	46.1
BP	Wood N (%)	0.308
BP	Leaf N (%)	1.562
BP	Wood : Leaf Dry Mass	4.182
CW	Wood Density (g/cm^3)	0.5925
CW	Wood C (%)	48.16
CW	Leaf C (%)	46.58
CW	Wood N (%)	0.389
CW	Leaf N (%)	1.544
CW	Wood : Leaf Dry Mass	9.617
SS	Wood C (%)	45.69

SS	Leaf C (%)	43.77
SS	Wood N (%)	0.48
SS	Leaf N (%)	0.86
WM	Wood C (%)	47.15
WM	Leaf C (%)	49.9
WM	Wood N (%)	0.89
WM	Leaf N (%)	1.84
SH	Wood : Leaf Dry Mass	6.528

Supplementary Table S2: covariate sampling by species code (Stems:Individuals)

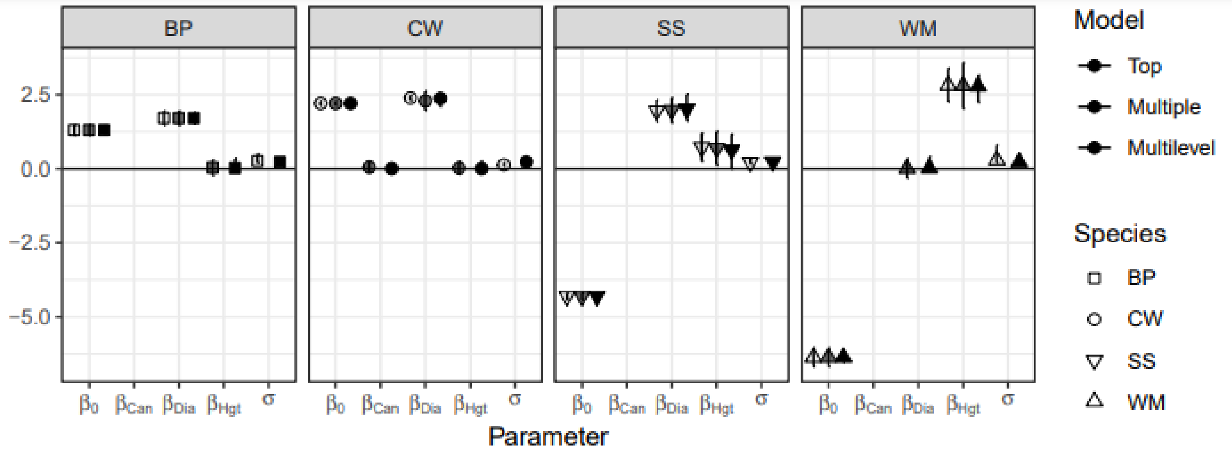
Covariate	**Units**	**BP**	**CW**	**SS**	**WM**
DRC	cm	11:11	19:19	18:6	7:3
D130	cm	11:11	19:19	NA	NA
D140	cm	NA	19:19	NA	NA
Canopy Area	m^2	NA	19:19	NA:6	NA:3
Height	m	11:11	19:19	18:6	7:3
Stem count	n	NA	NA	NA:6	NA:3

Supplementary Table S3: Deviance information criterion (DIC) scores for Multilevel models by response variable variance model and simplification. For details on taxon abbreviations, model fitting and simplification criteria, see Materials and Methods. For simplified equations, see Supplementary Table S4.

Response	Taxon (samples)	Variance Model	Covariates	DIC
Shoot Wood Dry Mass (kg)	BP (11), CW (19), SS (18), WM (7)	independent	all	-114.2
Shoot Wood Dry Mass (kg)	BP (11), CW (19), SS (18), WM (7)	independent	simplified	-113.6
Shoot Wood Dry Mass (kg)	BP (11), CW (19), SS (18), WM (7)	pooled	all	-122.2
Shoot Wood Dry Mass (kg)	BP (11), CW (19), SS (18), WM (7)	pooled	simplified	-116.1
Whole Plant Aboveground Biomass (kg)	BP (11), CW (19), SH (9)	independent	all	102.3
Whole Plant Aboveground Biomass (kg)	BP (11), CW (19), SH (9)	independent	simplified	97.7
Whole Plant Aboveground Biomass (kg)	BP (11), CW (19), SH (9)	pooled	all	96.6
Whole Plant Aboveground Biomass (kg)	BP (11), CW (19), SH (9)	pooled	simplified	93.7

Supplementary Table S4: Simplified expressions for estimating Shoot Wood, Aboveground biomass or Basal Diameter for different taxa using different covariates. All equations except Basal Diameter were fitted using simplified, latent variable weighted Multilevel models with pooled lognormal errors. Basal Diameter was fitted using Johnson and Omland (2004) AICc weighting for all first order linear models.

Species	**Response**	**Covariates**	**Function**
BP	Whole Plant Shoot Wood Dry Mass (kg)	Dia_130_cm	exp(1.314+1.712*[ln(Dia_130_cm)-1.315]+0.2530^2/2)
BP	Whole Plant Aboveground Biomass (kg)	Dia_130_cm	exp(1.581+1.626*[ln(Dia_130_cm)-1.315]+0.2501^2/2)
BP	Whole Plant Aboveground C (kg)	Dia_130_cm	exp(1.581+1.626*[ln(Dia_130_cm)-1.315]+0.2501^2/2)*0.4659
CW	Whole Plant Shoot Wood Dry Mass (kg)	Dia_130_cm	exp(2.208+2.392*[ln(Dia_130_cm)-1.801]+0.2530^2/2)
CW	Whole Plant Aboveground Biomass (kg)	Dia_130_cm, Can_msq	exp(2.338+2.285*[ln(Dia_130_cm)-1.801]+ 0.053*[ln(Can_msq)-2.771]+0.2501^2/2)
CW	Whole Plant Aboveground C (kg)	Dia_130_cm, Can_msq	exp(2.338+2.285*[ln(Dia_130_cm)-1.801]+ 0.053*[ln(Can_msq)-2.771]+0.2501^2/2)*0.4801
CW	Whole Plant Aboveground Biomass (kg)	Dia_130_cm	exp(2.338+2.370*[ln(Dia_130_cm)-1.801]+0.2492^2/2)
CW	Whole Plant Aboveground C (kg)	Dia_130_cm	exp(2.338+2.370*[ln(Dia_130_cm)-1.801]+0.2492^2/2)*0.4801
SS	Single Shoot Wood Dry Mass (kg)	Dia_RC_Lg_cm	exp(-4.311+2.436*[ln(Dia_RC_cm)+0.7481]+0.2530^2/2)
WM	Single Shoot Wood Dry Mass (kg)	Hgt_m	exp(-6.365+2.820*[ln(Hgt_m+0.1638]+0.2530^2/2)
SH	Whole Plant Aboveground Biomass (kg)	Dia_RC_Lg_cm, Hgt_m	exp(-2.445+0.886*[ln(Dia_RC_Lg_cm)-0.1928]+1.324*[ln(Hgt_m+0.0281]+0.2501^2/2)
SH	Whole Plant Aboveground C (kg)	Dia_RC_Lg_cm, Hgt_m	exp(-2.445+0.886*[ln(Dia_RC_Lg_cm)-0.1928]+1.324*[ln(Hgt_m+0.0281]+0.2501^2/2)*0.4648
SH	Basal Diameter (cm)	Dia_RC_Lg_cm, Stems_n	-0.2157+0.7473*Dia_RC_Lg_cm+0.5162*Stems_n



Supplementary Figure S1: Stem wood dry mass covariate effects for *Schinus terebinthifolia* (BP), *Cupaniopsis anacardioides* (CW), *Myrcianthes fragrans* (SS) and *Morella cerifera* (WM) generated either by the single GLM with the lowest AICc (Top), multimodel inference across all possible GLMs weighted by their AICc (Multiple) or a Bayesian Multilevel model with latent indicator weighting and pooled measurement error (Multilevel). Covariates were log transformed and mean centered prior to estimation and error bars represent 95% Confidence Intervals. Parameters for effects that were excluded or unestimable are not depicted.