# Generalized additive models reveal among-stand variation in live tree biomass equations

Jacob Levine<sup>1</sup>, Perry de Valpine<sup>1</sup> and John Battles<sup>1</sup>

Jacob I. Levine, Perry de Valpine, John J. Battles. Department of Environmental Science Policy and Management, University of California Berkeley, Berkeley CA, 94720

Corresponding Author. Jacob I. Levine. (phone: (310) 754-6059; email: jacoblevine@princeton.edu).

#### Abstract

Accurate estimation of forest biomass is important for scientists and policymakers interested in carbon accounting, nutrient cycling, and forest resilience. Estimates often rely on the allometry of trees; however, limited datasets, uncertainty in model form, and unaccounted for sources of variation warrant a re-examination of allometric relationships using modern statistical techniques. We asked the following questions: (1) Is there among-stand variation in allometric relationships? (2) Is there nonlinearity in allometric relationships? (3) Can among-stand variation or nonlinearities in allometric equations be attributed to differences in stand age? (4) What are the implications for biomass estimation? To answer these questions, we synthesized a dataset of small trees from six different studies in the White Mountains of New Hampshire. We compared the performance of generalized additive models (GAMs) and linear models and found that GAMs consistently outperform linear models. The best-fitting model indicates that allometries vary among both stands and species and contain subtle nonlinearities which are themselves variable by species. Using a planned contrasts analysis, we were able to attribute some of the observed among-stand heterogeneity to differences in stand age. However, variability in these results point to additional sources of stand-level heterogeneity, which if identified could improve the accuracy of live-tree biomass estimation.

#### **Key Words**

Biomass estimation, tree allometry, generalized additive models, stand effects

## **1** Introduction

2 Accurate estimation of live tree biomass is essential for scientific, economic and policy purposes. Live trees represent a large fraction of the total forest carbon pool (Pan et al. 2011). In 3 4 North America live biomass accounts for 32% of the 103,110 Tg of carbon stored in forest 5 ecosystems (Domke et al. 2018). Estimates of live tree biomass are frequently used to quantify 6 important ecological indices such as productivity gradients, biomass recovery after disturbance, 7 and forest resilience (Whittaker and Niering 1975; Peet 1981; Reiners et al. 2012). Due to 8 concerns about global change, managers and policy-makers are interested in maximizing both the carbon storage and sequestration capacities of forests (Aalde et al. 2006; Domke et al. 2012; 9 Van Breugel et al. 2011). Additionally, market-based policy instruments, such as carbon-offset 10 trading schemes, tie estimates of forest carbon stocks to monetary transactions (Kerchner and 11 12 Keeton 2015; Newell, Pizer, and Raimi 2013).

13 It is not feasible to measure live tree biomass directly and thus estimates of forest 14 biomass rely on allometric models (Jenkins et al. 2003, Van Breugel et al. 2011). Allometric 15 models of trees relate easily-obtained dimensional measurements (e.g. diameter, height), to 16 impractical- and destructive-to-measure quantities such as tree biomass (Huxley and Teissier 17 1936). Typically, the allometric relationships used to estimate tree, stand, and forest biomass 18 depend on measurements of diameter and/or height (Jenkins et al. 2003).

A significant barrier to the development of accurate allometric models is the difficulty
and expense involved in directly measuring tree biomass (Jenkins et al. 2003; Sileshi 2014; Van
Breugel et al. 2011). In order to measure the biomass of a single tree, it must be cut-down,
measured, dissected, dried and then weighed (Fatemi et al. 2011; Whittaker et al. 1974). The
combination of destructive sampling and time-consuming analyses constrains sample size. For

example, the median sample size for tree species included in BAAD (Biomass and Allometry
Database) is 4 trees (i.e., stems ≥ 2 m tall) per location (Falster et al. 2015). Only rarely (47 out
of 497 datasets) are there more than 21 trees. Small sample sizes limit the degree of certainty in
allometric models (Weiskittel et al. 2015). This study addresses this limitation by synthesizing
six comparable, small-tree datasets collected in the White Mountains of New Hampshire over the
course of several decades.

30 Allometric equations are sometimes species-specific but are frequently developed for groups of species based on functional traits or taxonomic grouping (Jenkins et al. 2003, Picard et 31 al. 2015; Radtke et al. 2017; Dettman and MacFarlane 2018; Poudel et al. 2019). In problematic 32 cases, equations are developed for a single species and then applied to different species and/or 33 34 species groups (Weiskittel et al. 2015). Misapplication and overgeneralization of allometric 35 equations can lead to inaccuracies in live tree biomass (Weiskittel et al. 2015). Even small 36 inaccuracies at the tree level can be amplified when biomass estimates are extrapolated to the 37 stand, forest or national levels. Inaccuracies are especially problematic if they are systematic. 38 Therefore, it is important to maintain a high standard of accuracy for tree-level allometric equations. 39

The existence of spatial variability in tree allometry can also severely impact the accuracy of biomass estimates when equations are applied to different areas from where they were developed (Weiskittel et al. 2015). At the regional scale, variability in tree allometry is well explored; the consensus recommendation is that biomass equations be applied solely within the forest or region for which they were developed (Weiskittel et al. 2015, Koerper and Richardson 1980). However, little attention is paid to the existence of among-stand variation in allometric relationships, despite evidence of its existence (Fatemi et al. 2011, Wang et al. 2002). The 47 existence of among-stand variation necessitates even greater care in the application of allometric 48 equations developed from small datasets or single stands. Nevertheless, if among-stand, 49 allometric variation could be attributed to specific stand-level covariates, it could allow for more 50 accurate biomass estimation in cases where the stand attributes are known and quantified. 51 Stand age, a determinant of stand structure and composition, is one such covariate that 52 may account for some among-stand allometric variation. Even-aged stands tend to follow 53 relatively deterministic successional patterns in the absence of disturbance. For example, evenaged stands in the northern hardwood forests of the Eastern United States undergo distinct shifts 54 55 in species composition, density and vertical structure as they mature (Marks 1974). It follows that trees of the same size which exist in different stages of stand development might exhibit 56 differences in resource allocation strategies. Specifically, these strategies may differ in order to 57 58 better compete for light or other resources and vary as a function of trees' relative vertical 59 positions (Peichl and Arain 2007; Walters et al. 1993). Differences in resource allocation patterns might further vary by the interaction between species and stand age. In other words, the 60 61 way in which allometric relationships vary as a function of stand age might be different for different species. This variation in resource allocation can affect relationships between low 62 dimensional measurements such as diameter and aboveground tree biomass. Several empirical 63 studies suggest the importance of stand age in driving allometric variability (Bond 2000, Fatemi 64 et al. 2011, Wang et al. 2002, Baskerville 1983). 65

Stand or species variation may be expressed in several aspects of allometric equations.
For one, variability may be present in the intercepts and coefficients of an allometric model that
explicitly considers stand effects and/or interactions between stand effects and other predictive
variables. Second, among-stand and/or among-species variation may be expressed in the

70 functional form of the model. Most commonly, allometric scaling relationships are assumed to 71 follow the power law, owing to both metabolic scaling theory and the apparent linearity of allometric relationships on a log-log scale (Picard et al. 2015, Sileshi 2014). Other allometric 72 73 analyses consider relationships that are non-linear on a double logarithmic scale (Huxley and 74 Teissier 1936, Picard et al. 2015, MacFarlane 2015, West et al. 1999). However, the 75 identification of an appropriate nonlinear model structure is difficult, and so previous 76 considerations of nonlinearity are limited by their consideration of only a few alternative forms (e.g. geometric models, variable-density mass component models) (Picard et al. 2015, 77 MacFarlane 2015). If the form of nonlinear allometric relationships itself varies in relation to 78 species, age, stand characteristics or other environmental factors, estimation and prediction 79 would be particularly complicated. If this were the case, the limitations of choosing a few 80 81 parametric model forms could either obscure existing variability, or erroneously detect non-82 existent variability.

83 The implementation of a non- or semi-parametric model, such as a generalized additive 84 model (GAM), can compensate for uncertainty in the form of the model by relaxing the assumption of linearity without specifying an alternative, nonlinear model form. Instead, GAMs, 85 which relate tree biomass to predictor variables (e.g. diameter, height, volume) via a combination 86 87 of smoothing splines and linear terms, automatically identify the appropriate relationships between predictor and response variables (Hastie and Tibshirani 1987; Guisan et al. 2002). 88 Thereby, these models provide freedom from the a priori assumptions about model form inherent 89 90 in the choice of a parametric linear or nonlinear function (Hastie and Tibshirani 1987; Picard et 91 al. 2015, MacFarlane 2015). Additionally, the use of GAMs allows us to examine nonlinearities 92 in a general fashion without having to compare numerous potential nonlinear forms. The

smoothing splines that constitute the non-parametric heart of GAMs are penalized fit objects
which provide balance between goodness of fit and some measure of smoothness (Wood 2011).
These splines are controlled by a smoothing parameter which adjusts the degree of penalization
placed on the "curviness" of the underlying, representative functions. This smoothing parameter
is fit by generalized cross validation and is typically constrained to be equal across splines in a
given model. By controlling the sensitivity of the splines to variation in the data, the smoothing
parameter provides the balance between fit and smoothness (Wood 2011).

We consider four potential patterns of variation in allometric relationships (Fig. 1). For 100 101 the purpose of demonstration, diameter is used as the continuous variable, and stand as the 102 categorical, though one could apply the same thinking to cases which consider either height or 103 species. In the left column, panels A and C depict relationships exhibiting among-stand variation 104 which could be modeled using linear regression. In panel A only a main effect of stand is present 105 (same slope, different intercepts), whereas panel C shows an interaction between diameter and 106 stand (different slopes and intercepts). The right column shows nonlinear relationships that 107 exhibit stand-level variability. The difference in the relationships between panels B and D 108 demonstrates the main motivation for the use of GAMs. In panel B, a nonlinear relationship is presented which is constant in the form of the relationship, but variable in the intercept. This 109 110 relationship could be modeled by a parametric nonlinear equation, given knowledge of the 111 correct model form. Panel D shows a relationship which exhibits heterogeneous nonlinearity – the relationships between diameter and biomass for each individual stand are heterogeneous not 112 113 only in the intercept or slope but also in the functional form.

This fourth form of variability – variability in functional form – highlights an additional
advantage of generalized additive models in exploring allometric relationships. Namely, they can

reveal sources of variability which are obscured by models which are rigid in functional form (i.e. linear and parametric nonlinear models). Our goal in this study is to distinguish between these forms of allometric variability across stands. Likely, the errors caused by their omission are proportionally small, but can nevertheless become important when scaled up from the tree to the national scale. Additionally, because these variations are likely caused by exogenous factors such as stand characteristics or species functional traits, they can lead to key ecological insights.

122 By examining generalized additive and linear models for allometric relationships in the 123 northern hardwood forests of New Hampshire, this study seeks to answer the following questions: (1) Is there among-stand variation in allometric relationships for linear and/or 124 nonlinear models? (2) Is there nonlinearity in allometric relationships? (3) Can among-stand 125 126 variation or nonlinearities in allometric equations be attributed to differences in stand age? (4) 127 What are implications for biomass estimation and ecological understanding? While the 128 magnitude of changes in biomass predictions is clearly important for a variety of applications, 129 detection of among-stand and among-species variation in allometric relationships is of 130 fundamental interest given these can provide insight into important ecological patterns and 131 functions.

132

# 133 Methods

*Site Description* - Data were compiled from a series of previous biomass studies
conducted in the White Mountain National Forest, New Hampshire. The full dataset includes 318
trees from 9 stands of 7 different ages. Most of the sampling was conducted in either Hubbard
Brook Experimental Forest (208 trees) or Bartlett Experimental Forest (71 trees). The remaining
trees were collected in two stands located in the vicinity of the experimental forests (Table 1).

The stands varied in age from 14 to 100+ years. Except for the oldest site, forests developed following clearcut harvests and represent even-aged stands. The oldest site is a mixed age stand that was never clearcut but did experience logging in the 1800's and early 1900's (Whittaker et al. 1974). As a result, the trees in the oldest stand ranged from a minimum of 16 years old to maximum of 260 years old. Whittaker et al. (1974) report a weighted mean age that varied from 124 years to 83 years across their sampling gradient. Thus, we identified this site as 100+ years old.

146 All stands were located in a northern hardwood forest type. At maturity, the northern 147 hardwood forest is dominated by sugar maple (Acer saccharum Marsh; ACSA), American beech 148 (Fagus grandifolia Ehrh; FAGR), and yellow birch (Betula alleghaniensis Britton; PRPE). In 149 young stands, pin cherry (Prunus pensylvanica L.f; PRPE), a pioneer species that rarely persists 150 in the forest past 30 years, is the most abundant species. As stands age, paper birch (Betula 151 papyrifera Marshall; BEPA), and to a lesser extent red maple (Acer rubrum L; ACRU) become 152 more abundant. Red spruce (*Picea rubens*, Sarg; PIRU) is the most common conifer species in 153 the northern hardwood forest and striped maple (Acer pensylvanicum L., ACPE) is a common 154 understory tree species (Bormann et al. 1970, Marks 1974, Fahey et al. 1998). Stands were grouped broadly into three age classes based on the successional patterns of 155 156 northern hardwood forests (Marks 1974). Besides reflecting our understanding of the ecology of 157 Northern Hardwood forests, grouping the stands into three distinct age classes rather than treating stand age as a continuous variable avoids the assumption that the effect of stand age is 158 159 monotonic and avoids estimating effects for stand ages we lack data on (<14, 29-100+). Stands 160 aged 14, 16 and 17 (2 stands) are labelled "YOUNG," stands aged 23, 26 (2 stands) and 29 161 labelled "MID," and the 100+ year old stand is labelled "OLD." YOUNG stands were dominated 162 by pin cherry and characterized by high stem densities. MID stands were defined by lower pin 163 cherry prevalence and higher striped maple and yellow birch components. The OLD stand was 164 dominated by yellow birch and two shade-tolerant hardwoods, sugar maple and American beech, 165 though other species like red spruce were present. This stand lacked pin cherry and had lower 166 stem density and higher basal area than other stands (Fahey et al. 1998; Fatemi et al. 2011; 167 Whittaker et al. 1974). Despite these differences in characteristics between the three age classes, 168 there was substantial enough overlap in species composition and tree size to justify allometric 169 comparison.

170 Study sites were all located at similar elevations (300-600m) with the "OLD" stand being 171 the highest. Sites were measured during different years in the period from 1973 to 2004, but 172 measurement protocols followed those described in Whittaker et al. 1972. Given their close 173 physical proximity, we assume the stands developed in relatively similar climates. However, 174 temporal differences could be present as a result of documented trends in climate and 175 atmospheric pollution (Fatemi et al. 2011).

176 Sampling - In each stand, a sample of trees was chosen to include a full distribution of present sizes without consideration for each species' contribution to stand density or size 177 distributions. Trees were selected to have good health and vigor, and diseased, damaged and 178 179 multi-stem trees were excluded. Given the large presence of beech bark disease in the younger 180 stands, this constrained the selection of beech individuals. Diameter at breast height (DBH, breast height = 1.37 m), and total height of each tree was measured on site. Following this, trees 181 182 were felled and then separated into stem, branch and foliage components. These components 183 were weighed in the field. Tissue samples were taken from each component of each tree, oven-184 dried and weighed in order to determine moisture content. These relationships were then applied to determine total dry aboveground biomass of each tree. Parabolic volume was calculated fromheight and diameter according to eq. 1 (Whittaker and Woodwell 1968).

187 (1)  $Volume = 0.5 \pi Radius^2 Height$ 

*Analysis* - First, diameter, height and aboveground biomass were log transformed for
consistency with previous allometric studies, to ensure comparability between linear and
generalized additive models, and to meet the assumptions of linear modeling. Additionally, trees
from the OLD stand which were larger than the largest trees from the MID and YOUNG stands
were not considered in the analysis. This filtering was done to remove any tree size bias from the
comparison between stands of different ages. This reduced our total sample size to 273 trees.

To determine the best predictive allometric model, we performed model selection on 182 194 195 potential model formulations. Models were ranked separately for both the linear and generalized 196 additive model forms using the Akaike Information Criterion corrected for small sample size 197 ,AICc, in the statistical analysis software R (R Core Team 2018 version 3.5.1; Burnham and 198 Anderson 2004). We started with full models for the linear (Equations 1a and 1b) and for the 199 GAM (Equations 2a and 2b), and then compared all simpler parameter and interaction 200 combinations. For the linear models, possible parameter and interaction combinations were fit 201 and then ranked using the package MuMIn (Barton, K. 2018; version 1.42.1). Models including 202 parabolic volume as a predictor were constructed and fit separately from models including 203 diameter and height. We reasoned that since parabolic volume is calculated from diameter and 204 height, including all three quantities risks fitting redundant parameters in the models.

205

206 (1)(a)  $\log(biomass) \sim \log(DBH) * \log(height) * species * stand ID$ 

207 (b)  $\log(biomass) \sim \log(parabolic volume) * species * stand ID$ 

208	(2)(a)	$\log(biomass) \sim s(\log(DBH) : species: stand ID)$	

209  $+ s(\log(height) : species: stand ID) + species * stand ID$ log(biomass) ~ s(log(parabolic volume): species: stand ID) 210 (*b*) 211 + species \* stand ID

212 Formulae written in R syntax. 's()' denotes smoothing splines on the parenthetical variables.

213

230

214 An analogous set of generalized additive models was constructed, fit and ranked. 215 However, specifications differed slightly due to the nature of GAMs. Models including parabolic 216 volume were again considered separately. Interactions between continuous variables were 217 specified as multivariate smoothing splines. In models where different splines were fit for each 218 species and/or stand ID, splines were specified to have the same smoothness penalty (fit by 219 generalized cross-validation). For both the linear and generalized additive models, we adopted a 220 fixed effect approach to controlling for the nested structure of the data. There are three reasons 221 why we opted for fixed effect as opposed to mixed effect models: 1) There is no random effects 222 analogue for among-stand variation in splines. 2) The assumption that random effects are 223 normally distributed is constraining and without a priori justification and 3) a random effects 224 structure precludes the post-hoc comparisons we hope to use to assess the effect of stand age. 225 Full details regarding methods of model construction can be found in the supplementary 226 material. GAMs were fit using the package mgcv (Wood 2011; version 1.8.24). 227 After performing model selection independently for both model types, we ranked all of 228 the models by AICc in order to compare performance between model forms. AICc imposes a 229 stronger penalty on model complexity than AIC and was chosen in order to avoid fitting models which were overly complex given the size of the dataset. The 50 best-fitting models of each type

are shown in Tables S1 and S2 (Supplementary Material). Then, we generated graphs for three
different measures of model fit (AICc, Log-Likelihood, and adjusted R<sup>2</sup>) as a function of the
(estimated) degrees of freedom in order to compare model performance at various levels of
model complexity (Figure 2).

235 The flexibility of Generalized Additive Models results in three potential limitations: They 236 cannot be used to estimate values outside the range on which they were fit, can result in 237 biologically inconsistent fits and are prone to bias towards over-complex models in the case of 238 AIC based selection (Burnham and Anderson 2004; Wood et al. 2016; Greven and Kneib 2010). The first concern is mollified by the fact that we are primarily interested in identifying variation 239 240 rather than developing predictive equations. The second is mitigated by fitting the smoothing 241 parameters using generalized cross-validation. The third concern is the most relevant for this 242 study: a generalized additive model might appear to outperform a linear model on a given dataset 243 even if there are no important nonlinearities, or unaccounted-for sources of variation (i.e. panels 244 B and D of Figure 1), simply due to AIC over-fitting. To check for this sort of spurious result, we 245 simulated a null distribution of AICc differences between best fitting GAMs and linear models 246 and used this to assess the difference found in our analysis. Similar to parametric bootstrapping 247 where the best-fit model is used to approximate the sampling distribution of parameter estimates 248 via simulation, we used the best-fit linear model for simulations. Specifically, we maintained the 249 same values for the explanatory variables found in our dataset but drew new log(biomass) values 250 from the best-fit linear model, including its mean predictions and residual variance. Then, we ran 251 the same model selection protocol used on the real data to determine the AICc difference 252 between the best fitting GAM and the best fitting linear model. We repeated this procedure 5,000 253 times and used the distribution of simulated differences as the null distribution for testing

whether the difference from the real data was significant. If the observed value is in the upper tailof the null distribution, it provides evidence against spurious overfitting due to use of AICc.

Following determination of the best-fitting model, post-hoc tests were employed to 256 257 determine whether some of the among-stand variation could be attributed to the effect of stand 258 age. Specifically, we compared tree allometries among the YOUNG, MID and OLD age groups. 259 We performed tests separately for each species that was measured in the OLD stand as well as at 260 least one YOUNG stand and one MID stand. Four species met this requirement: striped maple, 261 sugar maple, yellow birch, and American beech. We first subset the data by species, then fit both 262 a GAM and linear model to the subsetted data using the model terms in the overall best-fitting 263 model (Table 3). Then, using the multcomp package in R (Hothorn et al. 2008; version 1.4.8) we 264 specified general linear hypothesis tests in order to test for significant differences in the 265 parameters fit for YOUNG, MID, and OLD stand effects. These tests were performed separately 266 for parameters fit by linear and generalized additive models. The parameters that we considered 267 varied depending on where stand effects were included in the best-fitting model (i.e. as an 268 additive effect vs. interaction effect with diameter or height).

To assess the effects of model choice on biomass estimation we generated tree-level log(biomass) predictions for each in tree in the study. Then, we transformed the estimates into real terms, correcting for log-transform bias using the methodology outlined in Sprugel (1983). We summed estimates by stand age and species in order to better understand where differences in estimation exist between the two model forms. While good for relative comparisons, this method likely underestimates prediction errors given the predictions are generated on the same data for which the models are fit.

276 Stand and species differences were visualized using prediction plots. To create these 277 plots, predictions for log(biomass) were first generated using the best-fitting model. For ease of interpretation, plots only show predictions over either diameter or height even though predictions 278 279 were generated over both. In order to avoid predictions over unrealistic combinations of height 280 and diameter (for example a tree with 1cm height and 1,000cm DBH), a linear regression of 281 diameter and height was used to choose values for the non-displayed variable. Four plots were 282 generated: two showing distinct relationships for each stand age and two showing distinct 283 relationships for each species. Within each group (stand age or species), one plot shows diameter and the other shows height (Figures 4-7). All plots were generated using the package ggplot2 in 284 285 R (Wickham 2016).

286

### 287 Results

288 The best-fitting GAM (AICc = 34.3) outperformed the best-fitting linear model (AICc = 289 55.0). Notably, the 15 best-fitting GAMs exhibit lower AICc scores than the best-fitting linear 290 model (Tables 2, 3 and Tables S1 and S2). Furthermore, when measures of model fit were 291 plotted as a function of model complexity (degrees of freedom), the generalized additive models 292 consistently outperformed linear models of similar complexity (Figure 2). Moreover, the 293 observed AICc difference between the best fitting GAM and the best fitting linear model, 20.7, 294 falls in the far upper range of the values simulated from the null distribution (P = 0.00042, mean 295 = -8.77, sd = 8.82; Figure 3).

The overall best model (a GAM) includes terms for an additive effect of species and stand ID, different splines on log(DBH) for each species, and a single spline on log(height) (Table 3). The best-fitting linear model includes terms for an additive effect of both species and

299 stand ID as well as different slopes on log(DBH) for each species and stand ID combination, and 300 different slopes on log(HT) for each stand ID. Interaction terms between continuous predictors 301 and stand ID or species indicate variation similar to the relationships shown in panels C and D of 302 Fig. 1. These terms are included in both the best fitting linear model and generalized additive 303 model for species, but only in the best fitting linear model for stand ID. Additive stand ID and 304 species terms indicate relationships similar to those shown in panels A and B of Fig. 1. These 305 terms are included for both species and stand ID for both the best fitting linear and generalized 306 additive model. Parabolic volume is included in one of the six best-fitting generalized additive models, but none of the six best fitting linear models (Tables 2 and 3). The importance of 307 308 species, stand ID, log(DBH) and log(height) in predicting log(biomass) is robust to the choice of 309 model form with at least one parameter for each variable included in the six best fitting models 310 of both types.

311 General linear hypothesis tests detected statistically significant differences in allometric 312 relationships between YOUNG and MID stands, and MID and OLD stands for each of the four 313 species considered, and between YOUNG and OLD stands for three of the four species considered. Table 4 shows the full set of results from the planned contrasts analysis. However, 314 these results are not consistent across or even within model forms for most combinations of ages 315 316 and species. Thus, while we are able to establish some correlation between stand age and among-317 stand variation in allometry, the association is highly variable. This variability suggests the existence of unexplained stand-level drivers of variability. 318

Model choice resulted in small discrepancies in total biomass estimation. The linear model predicted higher total biomass (5,362 kg) than the generalized additive model (5,318 kg) when all trees in the study were summed. The actual total biomass of all measured trees in the study was 5,292 kg. Thereby, the best fitting linear model overestimated biomass by 1.3%,

323 whereas the best fitting generalized additive model overestimated biomass by 0.5%. The root

mean squared error (RMSE) was lower for the GAM than the LM (5.22 kg vs. 5.82 kg). The

325 effects of omitting stand and species parameters were slightly greater. The best fitting model that

included no stand ID terms overestimated biomass by 1.5%, and the best fitting model that

327 included no species terms overestimated biomass by 2.9%. Both of these models were GAMs.

328

# 329 Discussion

330 Generalized additive models reveal subtle (on a log scale) deviations from linearity in aboveground tree allometry. This result is consistent at all levels of model complexity (Figure 2). 331 332 The most parsimonious model for log(biomass) shows that stands differ in intercept, while 333 species differ in both intercept and model form. Thereby, this model (model 1 in Table 2) 334 exhibits a common model form for all stands, with additive variation in log(biomass) among 335 stands (see Panel B, Fig. 1). Among species, the model includes variation in model form as well 336 as additive variation in log(biomass) (see Panel D, Fig. 1). Thereby, the best fitting model 337 exhibits variable nonlinearities with species. Each species has a different spline shape, and analogously, a different model form. 338

Stand ID terms are included in 22 of the 25 best fitting models, including all six of the best fitting linear models and all six of the best fitting generalized additive models. The consistent inclusion of these terms indicates the existence of among-stand variation in the relationships between height, diameter and biomass. The best fitting model indicates that this variation occurs in the intercept, meaning that for every combination of height and diameter, stands vary in the logarithm of aboveground biomass by an additive factor.

345 The documentation of this variation raises two issues. First, stand level variation is rarely 346 included in the application of allometric models. Second, and perhaps more importantly, it 347 points to sources of variation resulting from unmeasured stand-level covariates. For example, 348 heterogeneity in species composition, tree size distributions, soil fertility, and/or microclimate 349 could result in allometric variability among stands. These drivers of variation, if present, are of 350 great interest because they may lead to a better understanding of ecological dynamics. If these 351 drivers are discovered and modeled, they may be used to develop improved predictive equations. 352 One potential driver of allometric variation is stand age. When stands were grouped into three age classes, significant differences were found for all but one combination of species and 353 354 age group pair (sugar maple, YOUNG vs OLD). This indicates that at least some of the among-355 stand variation in allometry can be attributed to stand age. The structure of even-aged northern 356 hardwood stands changes deterministically with age, resulting in differences in resource 357 availability for trees at different stages (Marks 1974). Differences in resource availability can 358 spur allometric variation if growth form is plastic. For example, leaf area and branch density 359 both respond to changes in light availability, resulting in the sort of additive variation in 360 log(biomass) among trees of a given size exhibited in the best-fitting model (Williams et al 1999; 361 Duchesneau et al. 2001; Delagrange et al. 2004). The process of self-thinning can stimulate 362 heterogeneity in leaf and branch density among trees of a single size by altering the structure of a 363 stand (Mohler et al. 1978; Bi and Turvey 1996; Rio et al. 2016). As stands age, suppressed individuals die, freeing up growing space for surviving individuals. In order to take advantage of 364 365 additional light resources, individuals then increase their leaf area and branch density (Mohler et 366 al. 1978; Marks 1974; Bi and Turvey 1996). Through this mechanism, individuals of a given size 367 can vary in biomass as a function of stand development, and therefore stand age.

368 Fatemi et al. (2011) also explored the effect of stand age on tree allometry in northern 369 hardwood forests by comparing predictions from equations developed from a new dataset of young-stand trees to those from equations developed from old-stand trees in Whittaker et al. 370 371 (1974). Fatemi et al. (2011) reported large differences in predicted biomass between the young 372 and old stand equations for yellow birch, though these were less pronounced for sugar maple and 373 American beech. As in Fatemi et al. (2011), we found significant differences for American 374 beech, sugar maple and yellow birch, though we also detected significant differences for striped 375 maple. These dissimilarities in results may be related to the alternative analytic approaches used by the two studies. In Fatemi et al. (2011), differences in allometric relationships between the 376 377 young and old stands were determined by comparing biomass predictions from allometric 378 equations on diameter developed in the old stand to those developed in the young stand. We 379 instead compared the fit of parameter estimates for stands of various age classes. It is therefore 380 possible that the results between the two studies could vary, as dissimilarities in parameter 381 estimates do not necessarily translate to similar discrepancies in prediction. Regardless, both 382 studies support the finding that allometric heterogeneity is present between stands.

383 The best-fitting generalized additive model includes nonlinear splines on both diameter and height, with individual splines on diameter for each species. This indicates among-species 384 385 heterogeneity in the form of allometric nonlinearities (on a log-log scale). One potential 386 explanation for this among-species variation in model form is that it reflects differences in the way that species respond to the varying competitive circumstances posed by being a small vs. 387 388 large tree in a given stand. To illustrate, one might predict that a shade-intolerant, pioneer species 389 such as pin cherry would be invariable (i.e. nearly linear) in allometry given that it can only 390 survive in full sun. Two trees living in full sun, regardless of their size, are unlikely to vary in

growth form (Brisson 2001; Muth and Bazzaz 2002). On the other hand, a shade-tolerant climax species such as American beech might be more nonlinear in allometry given that it can survive in either shade or sun. In even-aged stands like the ones in this study, trees on the large end of the size distribution are likely canopy trees (except in the OLD stand) and are therefore not lightlimited in the way that small trees in the same stand would be. For a given species, growth form is highly dependent on light availability, and so the allometries for these two trees are likely to differ (Williams et al 1999; Duchesneau et al. 2001, Delagrange et al. 2004).

398 Some evidence for this interpretation can be seen in Figures 4 and 5, where there appears 399 to be a higher degree of nonlinearity for shade tolerant species (e.g. American beech and striped 400 maple) than shade intolerant species (e.g. pin cherry and yellow birch). However, it is 401 inappropriate to draw firm conclusions from visual inspection of the prediction plots. This 402 method is purely heuristic (only one continuous explanatory variable is shown at a time), and 403 nonlinearities are sensitive to the existence of outliers, especially in relatively small datasets like 404 this one. Moreover, it is plausible that nonlinearities are really the result of unmeasured sources 405 of variation. Therefore, we suggest that future research efforts continue to explore the existence 406 of and explanations for allometric nonlinearities in both larger datasets and more diverse 407 geographies. Generalized additive models provide a valuable tool to this end.

Among-species variability also appeared in the planned contrasts analysis. There, the presence of significant stand age related differences in allometry varied by species for a given model form and age group comparison. This result points to an additional aspect of complexity in allometric relationships: variability in allometry due to stand conditions, including stand age, may also be affected by species differences. The reasoning behind this explanation is very similar to the explanation for species-specific nonlinearities discussed above. A tree of a given 414 size is likely to experience a different competitive circumstance in a young even-aged stand than 415 in an old even-aged stand, primarily because of differing size and species distributions. The 416 degree to which this competitive circumstance varies among different stand ages is also 417 influenced by the growth strategy and functional traits of the species (Williams et al. 1999; 418 Delagrange et al. 2004). For example, intermediate and late successional species can be expected 419 to exhibit more complex, dynamic relationships with stand conditions than early successional 420 species (Küppers 1989). Again, while the existence of among-species variation in allometry is 421 clear in this study, further research is required in order to attribute this variation to characteristics 422 of the species. If allometric variation is attributable to species traits, this could aid the 423 development of better allometric equations for species which lack sufficient data to develop 424 equations directly.

425 The differences in biomass estimates produced by the generalized additive and linear 426 models are small. For the 273 trees in this study, the total biomass predicted by the linear model 427 was approximately 1.3%, higher than the actual biomass of the sample trees. While the 428 generalized additive model also overpredicted the total biomass, it was by only 0.5%. 429 Overpredictions were worsened when stand ID and species terms were omitted from the models. 430 These prediction errors are likely underestimates given that they are generated using the same 431 data as the models. However, we cannot extrapolate these results to other datasets due to the 432 inclusion of Stand ID terms in the model. Additionally, we are limited in our ability to extrapolate our results to larger scales given that the largest tree in this study was 17.6 cm in 433 diameter. That being said, accurate estimates of live tree biomass in young stands are important 434 435 for ecological monitoring and management. Aboveground, young-stand biomass is an important 436 indicator of ecosystem function, being directly related to carbon cycling, primary productivity

and tree demography. Many ecological studies of forest systems seek to quantify aboveground
biomass in order to assess processes such as post-disturbance biomass recovery, and forest
resilience (Reiners et al. 2012; Whittaker and Niering 1975). Thereby, systematic overestimation
of aboveground, young-stand biomass has important ramifications for our understanding of
forest resilience and recovery.

Furthermore, while the difference in total biomass estimates is proportionally small, it 442 443 points to an opportunity for important refinements. That generalized additive models provide 444 consistently better fits than linear models and result in even slightly more accurate biomass estimates points to the presence of additional sources of variability that is not captured by linear 445 models. As discussed previously, this heterogeneity shows up as variable nonlinearities in the 446 447 generalized additive models (Figure 1, Panel D) as well as in the additive effects of stand ID and 448 species (Figure 1, Panel B). The omission of these variables leads to larger prediction errors than 449 the ones resulting from choosing a linear model over a GAM.

450 Due to limitations in the available data we only examine stand age as a possible driver of 451 allometric variation in this study. However, future studies should prioritize the identification of 452 the sources of both among-stand and among-species variation. For example, among-stand 453 variation may result from differences in the physical environment (i.e. soil type, aspect, 454 hydrology etc.), or ecological characteristics (i.e. tree size distributions and density, species 455 composition) while among-species variation could be the result of variation in functional traits 456 such as shade tolerance, or wood structure. Attributing variation to these characteristics could lead to both improvements in the predictive accuracy of biomass equations and in a better 457 458 biological understanding of allometric relationships.

459

460 Conclusion

In this study we found that allometric relationships for aboveground biomass not only vary by species but also by stand. We documented substantial evidence for the importance of stand age in driving among-stand variation in allometry, but heterogeneity in these results points to additional drivers of allometric variation. The best-fitting model contains subtle nonlinearities which are variable among species. These nonlinearities may indicate unaccounted for sources of variation, or even interesting biological mechanisms.

The nonlinearities uncovered by the GAMs lead to reduced prediction error when compared to the more standard linear models. Overestimation of biomass by the linear models compared to the GAMs is proportionally small, but nevertheless indicates the opportunity for refinements in the accuracy of biomass estimation equations, especially if the sources of variation can be identified and modeled in the future.

472

#### 473 Acknowledgements

Allometric data archived as part of the National Science Foundation's Long-Term
Ecological Research network. Hubbard Brook Experimental Forest and Bartlett Experimental
Forest are operated and maintained by the USDA Forest Service, Northern Research Station,
Newton Square, PA, US. We appreciate the cooperation and logistical support of the White
Mountain National Forest. Funding for this research was provided by National Science
Foundation's Long-Term Ecological Research project (DEB 1114804).

480

481 References

482	Aalde, H., P. Gonzalez, M. Gytarsky, T. Krug, W.A. Kurz, S. Ogle, J. Raison, D. Schoene, N.H.
483	Ravindranath, N.G. Elhassan, L.S. Heath, N. Higuchi, S. Kainja, M. Matsumoto, M.J.S.
484	Sánchez, and Z. Somogyi. 2006. Forest Land. In Intergovernmental Panel on Climate
485	Change. [Eggleston H.S., L. Buendia, K. Miwa, T. Ngara, and K. Tanabe (eds).] National
486	Greenhouse Gas Inventory Guidelines. Institute for Global Environmental Strategies,
487	Hayama, Japan.
488	Baskerville, G. L. 1983. Balsam fir foliage regressions are age dependent. Canadian Journal of
489	Forest Research, 13(6), 1248-1251. https://doi.org/10.1139/x83-164
490	Barton, K. 2013. Package 'MuMIn'. Version 1.18.
491	Bi, H. and Turvey, N.D., 1996. Competition in mixed stands of Pinus radiata and Eucalyptus
492	obliqua. Journal of applied ecology, pp.87-99.
493	Bond, B.J. 2000. Age-related changes in photosynthesis of woody plants. Trends in Plant
494	Science, 5(8), 349-353. https://doi.org/10.1016/S1360-1385(00)01691-5
495	Burnham, K.P., and Anderson, D.R. 2004. Multimodel Inference. Sociological Methods and
496	Research, 33(2), 261-304. https://doi.org/10.1177/0049124104268644
497	Del Río, M., Pretzsch, H., Alberdi, I., Bielak, K., Bravo, F., Brunner, A., Condés, S., Ducey,
498	M.J., Fonseca, T., von Lüpke, N. and Pach, M. 2016. Characterization of the structure,
499	dynamics, and productivity of mixed-species stands: review and perspectives. European
500	Journal of Forest Research, 135(1), 23-49.
501	Delagrange, S., Messier, C., Lechowicz, M.J. and Dizengremel, P., 2004. Physiological,
502	morphological and allocational plasticity in understory deciduous trees: importance of
503	plant size and light availability. Tree physiology, 24(7), pp.775-784.
504	Delucia, E.H., Sipe, T.W., Herrick, J., and Maherali, H. 1998. Sapling biomass allocation and

- 505 growth in the understory of a deciduous hardwood forest. *American Journal of Botany*,
- 506 *85*(7), 955–963. https://doi.org/10.2307/2446362
- 507 Domke, G.M., Becker, D.R., D'Amato, A.W., Ek, A.R., and Woodall, C.W. 2012. Carbon
- 508 emissions associated with the procurement and utilization of forest harvest residues for
- 509 energy, northern Minnesota, USA. *Biomass and Bioenergy*, *36*, 141–150.
- 510 https://doi.org/10.1016/J.BIOMBIOE.2011.10.035
- 511 Domke, G.M., Williams, C.A., Birdsey, R., Coulston, J., Finzi, A., Gough, C., Haight, B.,
- 512 Hicke, J., Janowiak, M., de Jong, B., Kurz, W.A., Lucash, M., Ogle, S., Olguín-
- 513 Alvarez, M., Pan, Y., Skutsch, M., Smyth, C., Swanston, C., Templer, P., Wear, D.,
- and Woodall, C.W. 2018. Chapter 9: Forests. *In Second State of the Carbon Cycle*
- 515 *Report (SOCCR2): A Sustained Assessment Report* [Cavallaro, N., Shrestha, G.,
- 516 Birdsey, R., Mayes, M.A., Najjar, R.G., Reed, S.C., Romero-Lankao, P., and Zhu, Z.
- 517 (eds.)]. U.S. Global Change Research Program, Washington, DC, USA, pp. 365-398,
- 518 https://doi.org/10.7930/SOCCR2.2018.Ch9.
- Duchesneau, R., Lesage, I., Messier, C. and Morin, H., 2001. Effects of light and intraspecific
  competition on growth and crown morphology of two size classes of understory balsam
  fir saplings. *Forest ecology and management*, *140*(2-3), pp.215-225.
- 522 Fahey, T.J., Battles, J.J., and Wilson, G.F. 1998. Responses of early successional northern
- hardwood forests to changes in nutrient availability. *Ecological Monographs*, 68(2), 183–
- 524 212. https://doi.org/10.1890/0012-9615(1998)068[0183:ROESNH]2.0.CO;2
- 525 Fatemi, F.R., Yanai, R.D., Hamburg, S.P., Vadeboncoeur, M.A., Arthur, M.A., Briggs, R.D., and
- 526 Levine, C.R. 2011. Allometric equations for young northern hardwoods: the importance of
- 527 age-specific equations for estimating aboveground biomass. *Canadian Journal of Forest*

- 528 *Research*, *41*(4), 881–891. https://doi.org/10.1139/x10-248
- 529 Greven, S. Kneib, T. On the behaviour of marginal and conditional AIC in linear mixed models,

530 *Biometrika*, 97(4), 773–789, <u>https://doi.org/10.1093/biomet/asq042</u>

- 531 Guisan, Antoine, Thomas C. Edwards Jr, and Trevor Hastie. "Generalized linear and generalized
- additive models in studies of species distributions: setting the scene." *Ecological*
- 533 *modelling* 157.2-3 (2002): 89-100. https://doi.org/10.1016/S0304-3800(02)00204-1
- Hastie, T., and Tibshirani, R. 1987. Generalized Additive Models: Some Applications. *Journal*
- 535 *of the American Statistical Association*, 82(398), 371–386.
- 536 https://doi.org/10.1080/01621459.1987.10478440
- 537 Hothorn, T., Bretz, F., Westfall, P., Heiberger, R. M., Schuetzenmeister, A., and Scheibe, S.
- 538 (2016). Package 'multcomp'. Simultaneous inference in general parametric models.
  539 Project for Statistical Computing, Vienna, Austria.
- 540 Huxley, J.S., and Teissier, G. 1936. Terminology of Relative Growth. Nature, 137(3471), 780-
- 541 781. https://doi.org/10.1038/137780b0
- 542 Intergovernmental Panel on Climate Change, National Greenhouse Gas Inventory Guidelines,
- 543 vol. 4, Institute for Global Environmental Strategies, Hayama, Japan, pp. 4.1-4.83
- 544 Jenkins, J.C., Chojnacky, D.C., Heath, L.S., and Birdsey, R.A. 2003. National-scale biomass
- 545 estimators for United States tree species. *Forest Science*, 49(1), 12–35.
- 546 https://doi.org/10.2737/NE-GTR-310
- 547 Kerchner, C.D., and Keeton, W.S. 2015. California's regulatory forest carbon market: Viability
- for northeast landowners. *Forest Policy and Economics*, *50*, 70–81.
- 549 https://doi.org/10.1016/J.FORPOL.2014.09.005

- Koerper, G. J., and Richardson, C. J. 1980. Biomass and net annual primary production
  regressions for Populus grandidentata on three sites in northern lower Michigan.
- 552 *Canadian Journal of Forest Research*, *10*(1), 92-101. https://doi.org/10.1139/x80-015
- 553 Küppers, M. 1989. Ecological significance of above-ground architectural patterns in woody
- plants: A question of cost-benefit relationships. *Trends in Ecology and Evolution*, 4(12),
- 555 375–379. https://doi.org/10.1016/0169-5347(89)90103-1
- 556 MacFarlane, D.W. 2015. A generalized tree component biomass model derived from principles
- of variable allometry. *Forest Ecology and Management*, *354*, 43–55.
- 558 https://doi.org/10.1016/j.foreco.2015.06.038
- 559 Marks, P.L. 1974. The Role of Pin Cherry (Prunus pensylvanica L.) in the Maintenance of
- 560 Stability in Northern Hardwood Ecosystems. *Ecological Monographs*, 44(1), 73–88.
  561 https://doi.org/10.2307/1942319
- Mohler, C.L., Marks, P.L., and Sprugel, D.G. 1978. Stand structure and allometry of trees during
  self-thinning of pure stands. *The Journal of Ecology*: 599-614.
- 564 Newell, R.G., Pizer, W.A., and Raimi, D. 2013. Carbon Markets 15 Years after Kyoto: Lessons
- 565 Learned, New Challenges. *Journal of Economic Perspectives*, 27(1), 123–146.
- 566 https://doi.org/10.1257/jep.27.1.123
- 567 Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L.,
- 568 Shvidenko, A., Lewis, S.L., Canadell, J.G. and Ciais, P., 2011. A large and persistent
- carbon sink in the world's forests. *Science*, *333*(6045), pp.988-993
- 570 Peet, R.K., 1981. Changes in biomass and production during secondary forest succession. *Forest succession*, (pp. 324-338). Springer, New York, NY.
- 572 Peichl, M., and Arain, M.A. 2007. Allometry and partitioning of above- and belowground tree

- 573 biomass in an age-sequence of white pine forests. *Forest Ecology and Management*, 253(1–
- 574 3), 68–80. https://doi.org/10.1016/J.FORECO.2007.07.003
- 575 Picard, N., Rutishauser, E., Ploton, P., Ngomanda, A., and Henry, M. 2015. Should tree biomass
- allometry be restricted to power models? *Forest Ecology and Management*, *353*, 156–163.
- 577 https://doi.org/10.1016/j.foreco.2015.05.035
- 578 Reiners, W.A., Driese, K.L., Fahey, T.J., & Gerow, K.G. 2012. Effects of three years of
- regrowth inhibition on the resilience of a clear-cut northern hardwood forest. *Ecosystems*, *15*(8), 1351-1362. https://doi.org/10.1007/s10021-012-9589-0
- 581 Shaw, C.H., Hilger, A.B., Metsaranta, J., Kurz, W.A., Russo, G., Eichel, F., Stinson, G., Smyth,
- 582 C. and Filiatrault, M., 2014. Evaluation of simulated estimates of forest ecosystem carbon
- stocks using ground plot data from Canada's National Forest Inventory. *Ecological*
- 584 *Modelling*, *272*, pp.323-347.
- 585 Sileshi, G.W. 2014. A critical review of forest biomass estimation models, common mistakes and
- 586 corrective measures. *Forest Ecology and Management*, *329*, 237–254.
- 587 https://doi.org/10.1016/j.foreco.2014.06.026
- Sprugel, D. G. 1983. Correcting for bias in log-transformed allometric equations. *Ecology*, 64(1),
  209-210. https://doi.org/10.2307/1937343
- 590 Van Breugel, M., Ransijn, J., Craven, D., Bongers, F., and Hall, J. S. 2011. Estimating carbon
- 591 stock in secondary forests: Decisions and uncertainties associated with allometric biomass
- 592 models. *Forest Ecology and Management*, *262*(8), 1648–1657.
- 593 https://doi.org/10.1016/j.foreco.2011.07.018
- 594 Walters, M.B., Kruger, E.L., and Reich, P.B. 1999. Growth, biomass distribution and CO2
- exchange of northern hardwood seedlings in high and low light: relationships with

- 596 successional status and shade tolerance. *Oecologia*, 94(1), 7–16.
- 597 https://doi.org/10.1007/BF00317294
- Wang, J.R., Zhong, A.L., and Kimmins, J.P. 2002. Biomass estimation errors associated with the
  use of published regression equations of paper birch and trembling aspen. *Northern*

600 *Journal of Applied Forestry*, 19(3), 128-136. https://doi.org/10.1093/njaf/19.3.128

- 601 Weiskittel, A.R., MacFarlane, D.W., Radtke, P.J., Affleck, D.L.R., Temesgen, H., Woodall, C. et
- al. 2015. A Call to Improve Methods for Estimating Tree Biomass for Regional and
- 603 National Assessments. Journal of Forestry, 113(4), 414–424. https://doi.org/10.5849/jof.14-
- 604 091
- West, G.B., Brown, J.H., and Enquist, B.J. 1997. A general model for the origin of allometric
  scaling laws in biology. *Science (New York, N.Y.)*, 276(5309), 122–126.
- 607 https://doi.org/10.1126/science.276.5309.122
- Wickham, H. 2011. ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics*, 3(2),
  180-185.
- 610 Williams, H., Messier, C. and Kneeshaw, D.D., 1999. Effects of light availability and sapling
- 611 size on the growth and crown morphology of understory Douglas-fir and lodgepole pine.
- 612 *Canadian Journal of Forest Research*, *29*(2), pp.222-231.
- White, J. and Harper, J.L., 1970. Correlated changes in plant size and number in plant
  populations. *The Journal of Ecology*, pp.467-485.
- 615 Whittaker, R.H., Bormann, F.H., Likens, G.E., and Siccama, T.G. 1974. The Hubbard Brook
- 616 ecosystem study: forest biomass and production. Ecological Monographs, 44(2): 233-254.
- 617 https://doi.org/10.2307/1942313

- 618 Whittaker, R.H., and Niering, W.A. (1975). Vegetation of the Santa Catalina Mountains,
- Arizona. V. Biomass, production, and diversity along the elevation gradient. *Ecology*,
  56(4), 771-790. https://doi.org/10.2307/1936291
- 621 Whittaker, R.H., and Woodwell, G.M. 1968. Dimension and Production Relations of Trees and
- 622 Shrubs in the Brookhaven Forest, New York. *Journal of Ecology*, 56(1), 1–25.
- 623 https://doi.org/10.2307/2258063
- 624 Wood, S. N. (2001). mgcv: GAMs and generalized ridge regression for R. *R news*, *1*(2), 20-25.
- 625 Wood, S. N., Pya, N. and Säfken, B. 2016. Smoothing Parameter and Model Selection for
- 626 General Smooth Models. *Journal of the American Statistical Association*, 111(516),
- 627 1548-1563, DOI: <u>10.1080/01621459.2016.1180986</u>
- 628 Woodall, C.W., Coulston, J.W., Domke, G.M., Walters, B.F., Wear, D.N., Smith, J.E., Andersen,
- 629 H.E., Clough, B.J., Cohen, W.B., Griffith, D.M. and Hagen, S.C., 2015. The US forest
- 630 carbon accounting framework: stocks and stock change, 1990-2016. Gen. Tech. Rep. NRS-
- 631 154. Newtown Square, PA: US Department of Agriculture, Forest Service, Northern
- 632 *Research Station.* 49 p., 154, pp.1-49.
- 633 Yoda, K. 1963. Self-thinning in overcrowded pure stands under cultivated and natural conditions
- 634 (Intraspecific competition among higher plants. XI). J. Inst. Polytech. Osaka City Univ.
- 635 Ser. D., 14, 107-129.
- 636

## Tables:

Table 1: Summary of datasets, sources, locations, stand ages, sample sizes (n), and species.

Data summary				
Stand ID Data Set	Source	Location	tand Age n Species	S
1 FatemiBEF	Fatemi et al. 2011	Bartlett Experimental Forest	14 18 red ma	ple, sugar maple, yellow birch, American beech, pin cherry
2			16 17 red ma	ple, sugar maple, yellow birch, paper birch, American beech, pin cherry
3			26 19 red ma	ple, sugar maple, yellow birch, paper birch, American beech, pin cherry
4			29 17 red ma	ple, sugar maple, yellow birch, paper birch, American beech, pin cherry
5 BattlesRock	Fahey et al. 1998	White Mountain National Forest	17 27 yellow	birch, paper birch, American beech
6 SiccamaW5	Unpublished data	Hubbard Brook Experimental Forest	17 64 red ma	ple, sugar maple, yellow birch, paper birch, American beech, pin cherry
7 BattlesGB	Fahey et al. 1998	White Mountain National Forest	23 12 yellow	birch, paper birch, American beech
8 FaheyW4	Unpublished data	Hubbard Brook Experimental Forest	26 49 striped	I maple, sugar maple, yellow birch, paper birch, American beech, pin cherry
9 WhittakerW6	Whittaker et al. 1974	Hubbard Brook Experimental Forest	100+ 50 red ma	ple, sugar maple, yellow birch, paper birch, American beech, rubens

Table 2: Summarized model selection results for the generalized additive model form. Only the six best-fitting models are shown. The estimated degrees of freedom (edf), AICc, change in AICc (delta), and relative model support (weight) are reported for each model.

Model Selection Results (GAM)									
Rank	Model	edf	AICc	delt	ta	weight			
	1 log(biomass) ~ species + stand ID + s(log(DBH), by = species) + s(log(height))		32	34.3	0.00	8.05E-01			
	2 log(biomass) ~ species + stand ID + s(log(height), by = species) + s(log(DBH))		33	38.8	4.49	8.51E-02			
	3 log(biomass) ~ species + stand ID + s(log(volume), by = species)		32	39.7	5.40	5.39E-02			
	4 log(biomass) ~ species + s(log(DBH), by = species) + s(log(height), by = stand ID)		36	41.5	7.22	2.17E-02			
	5 log(biomass) ~ species:stand ID + s(log(DBH), by = species) + s(log(height))		61	41.9	7.62	1.79E-02			
	6 log(biomass) ~ species + stand ID + s(log(DBH), log(height), by = species)		37	43.3	8.99	8.98E-03			

Table 3: Summarized model selection results for the linear model form. Only the six best-fitting models are shown.

Model Selection Results	(Linear Model)
-------------------------	----------------

Rank	Model	df		AIC	delta	weight
1	log(biomass) ~ log(DBH) + log(height) + species + stand ID		42	55.0	20.68	2.61E-05
	+ log(DBH):species + log(DBH):stand ID + log(height):stand ID					
2	log(biomass)~log(DBH) + log(height) + species + stand ID + log(DBH):species		73	55.7	21.44	1.77E-05
	+ log(DBH):stand ID + log(neight):stand ID + species:stand ID					
3	log(biomass) ~ log(DBH) + log(height) + species + stand ID		65	57.9	23.56	6.16E-06
	+ log(DBH):species + log(DBH):stand ID + species:stand ID					
4	log(biomass) ~ log(DBH) + log(height) + species + stand ID		42	58.1	23.80	5.47E-06
	+ log(DBH):stand ID + log(height):species + log(height):stand ID					
5	log(biomass) ~ log(DBH) + log(height) + species + stand ID + log(DBH):stand ID		73	59.6	25.30	2.58E-06
	+ log(height):species + log(height):stand ID + species:stand ID					
6	log(biomass) ~ log(DBH) + log(beight) + species + stand ID		65	60.7	26.40	1.49E-06
Ũ	+ log(DBH):stand ID + log(height):species + species:stand ID			50.7	20.40	252.00

Table 4: Results of planned contrast analysis. Estimates and p-values are shown for each comparison and model. Light blue highlighting denotes significance at the p = 0.1 level, Dark highlighting indicates significance at the p = 0.05 level. For each individual species the results of two tests are shown, one for the additive effect of stand in best-fitting linear model (labelled "Linear"), and one for the additive effect of stand in the best-fitting generalized additive model.

	striped maple		sugar maple		yellow	birch	American beech		
YOUNG vs OLD	Est.	p-value	Est.	p-value	Est.	p-value	Est.	p-value	
Linear	-7.69	0.03	-11.78	0.08	-14.88	6.64E-05	-0.48	0.93	
GAM	-0.137	0.15	-0.11	0.43	2.58	< 1E-16	2.62	< 1E-16	
YOUNG vs MID									
Linear	7.84	0.02	6.85	0.25	0.33	0.91	11.84	0.03	
GAM	-0.12	0.18	4.28	< 1E-16	2.78	< 1E-16	10.96	< 1E-16	
MID vs OLD									
Linear	-15.52	3.11E-15	-18.634	1.49E-10	-15.21	1.94E-10	-12.32	6.26E-06	
GAM	-0.009	0.88	-4.39	< 1E-16	-0.19	2.97E-02	-8.33	< 1E-16	

**Comparison Tests** 

## **Figure Captions**



Figure 1: Four potential expressions of among-stand heterogeneity in allometric relationships as a function of log(diameter). Relationships in the left column are linear while relationships in the right column are nonlinear. The top row contains relationships with no interaction effect between stand and allometry, while the bottom row contains relationships with an interaction effect. The relationships displayed in panel D demonstrate the case of nonlinearity that does not follow a simple equation possibly due to stand history or characteristics, which motivates the consideration of generalized additive models for allometric equations.



Figure 2: Three measures of model fit as a function of model complexity. From left to right the panels show the Akaike Information Criterion with correction for small sample sizes (AICc), Log-Likelihood, and adjusted R<sup>2</sup> as a function of the (estimated) degrees of freedom. Each point represents one of the fitted models. Only the best-performing models for each model type are displayed.



Figure 3: Histogram of simulated differences (blue) in AICc value between best fitting generalized additive model and the best fitting linear model for 5,000 iterations (AICc difference = AICc Linear – AICc GAM). Negative values indicate cases in which the linear model performed better (had lower AICc values) than the GAM. The red line indicates the difference in AICc between these models for the observed data.



Figure 4: Predicted log(biomass) values for the generalized additive model, shown here as a function of log(DBH). log(height) values were chosen using predictions from a linear model of height on diameter. Relationships are separated by species, and stand age is shown as a gradient from light to dark. Panels display results for each species.



Figure 5: Predicted log(biomass) values for the generalized additive model, shown here as a function of log(height). Relationships are separated by species, and stand age is shown as a gradient from light to dark. log(diameter) values were chosen using predictions from a linear model of diameter on height. Panels display results for each species.



Figure 6: Predicted log(biomass) values for the generalized additive model, shown here as a function of log(DBH). log(height) values were chosen using predictions from a linear model of height on diameter. Relationships are separated by stand age, and species are delineated by color. Panels display results for each stand age.



Figure 7: Predicted log(biomass) values for the generalized additive model, shown here as a function of log(height). log(diameter) values were chosen using predictions from a linear model of diameter on height. Relationships are separated by stand age, and species are delineated by color. Panels display results for each stand age