



Comparing tree foliage biomass models fitted to a multispecies, felled-tree biomass dataset for the United States



Brian J. Clough^{a,*}, Matthew B. Russell^a, Grant M. Domke^b, Christopher W. Woodall^b, Philip J. Radtke^c

^a Department of Forest Resources, University of Minnesota, St. Paul, MN, United States

^b U.S. Department of Agriculture, Forest Service, Northern Research Station, St. Paul, MN, United States

^c Forest Resources and Environmental Conservation, Virginia Polytechnic Institute, Blacksburg, VA, United States

ARTICLE INFO

Article history:

Received 14 May 2015

Received in revised form 5 January 2016

Accepted 8 April 2016

Keywords:

Foliage biomass models

Component ratio models

Bayesian hierarchical models

Posterior predictive checking

Prediction uncertainty

ABSTRACT

Estimation of live tree biomass is an important task for both forest carbon accounting and studies of nutrient dynamics in forest ecosystems. In this study, we took advantage of an extensive felled-tree database (with 2885 foliage biomass observations) to compare different models and grouping schemes based on phylogenetic and geographic variation for predicting foliage biomass at the tree scale. We adopted a Bayesian hierarchical statistical framework, first to compare linear models that predict foliage biomass directly to models that separately estimate a foliage ratio as a component of total aboveground biomass, then to compare species specific models to both 'narrow' and 'broad' general biomass models using the best fitted functional form. We evaluated models by simulating new datasets from the posterior predictive distribution, using both summary statistics and visual assessments of model performance. Key findings of our study were: (1) simple linear models provided a better fit to our data than component ratio models, where total biomass and the foliar ratio are estimated separately; (2) species-specific equations provided the best predictive performance, and there was no advantage to narrow species groupings relative to broader groups; and (3) all three model schemes (i.e., species-specific models versus narrow or broad groupings proposed in national-scale biomass equations) tended to over-predict foliage biomass and resulted in predictions with very high uncertainty, particularly for large diameter trees. This analysis represents a fundamental shift in carbon accounting by employing felled-tree data to refine our understanding of uncertainty associated with component biomass estimates, and presents an ideal approach to account for tree-scale allometric model error when estimating forest carbon stocks. However, our results also highlight the need for substantial improvements to both available fitting data and models for foliage biomass before this approach is implemented within the context of greenhouse gas inventories.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Estimation of live tree biomass from forest inventory data, typically using regression models with individual tree measurements as covariates, is important for both forest carbon accounting and studies of nutrient flows across whole ecosystems (Houghton, 2003; Jenkins et al., 2003; Schroeder et al., 1997). While many models have been developed for total tree biomass, it is often desirable to have individual estimates of particular biomass components (e.g. leaves, branches, and roots) as these stocks play key roles in many ecosystem processes (Chapin et al., 2002). For example, the bulk of a tree's nitrogen is found in foliage (Wirth et al., 2004) while fine

roots are involved in the exchange of both nutrients and carbon at the plant–soil interface (Bardgett et al., 2014). However, empirical observations of tree component biomass stocks are costly and difficult to collect in the field, let alone during a systematic inventory of forests at scales ranging from an individual stand to national efforts (e.g., IPCC, 2006). Models that provide reliable estimates with realistic uncertainty bounds are necessary for forest ecosystem research, monitoring, and reporting of national greenhouse gas inventories (Baker et al., 2010).

Despite the need, a consistent methodology for estimating biomass components from forest inventories is lacking. Felled-tree data are required to estimate parameters in biomass models, but studies that have accurately sampled biomass components are rare relative to those that have sampled whole trees (Jenkins et al., 2003; Zianis et al., 2005). In addition, component stocks are often highly variable. Foliage is particularly dynamic, varying both across space

* Corresponding author.

E-mail address: bclough@umn.edu (B.J. Clough).

(i.e., individual tree sizes) and time (i.e., leaf longevity). Tree foliage biomass is influenced by a host of factors including but not limited to: genetic variation, climate, geography, stand dynamics, and disturbance events (Niinemets, 2001; Reich et al., 2014; Wirth et al., 2004). Understanding these dynamics requires models tailored to biomass components to both improve predictive accuracy and to understand the scale of uncertainty in component stocks relative to that of whole tree estimates.

Thus far, two general frameworks have been proposed for modeling foliage biomass (Bartelink, 1998). One approach involves the relationship between the area of conductive tissue located in the tree stem at the base of the crown and the amount of leaf biomass it supports (Mäkelä, 1986; Shinozaki et al., 1964; Valentine et al., 1985). This approach has been used to construct models that directly estimate leaf biomass based on stem cross section (Robichaud and Methven, 1992; Waring et al., 1982; Wirth et al., 2004; Zelliers et al., 2012):

$$\ln(BM_{fol}) = \alpha + \beta * \ln(d) \quad (1)$$

where BM_{fol} is foliage biomass, d is a proxy of sapwood area (typically diameter at breast height; dbh), and α and β are model parameters. Although the parsimonious nature of these models is attractive, their application may be limited due to poor relationships between dbh and the area of conductive tissue higher up the stem (Bartelink, 1996; Bond-Lamberty et al., 2002; Laubhann et al., 2010).

An alternative approach is to estimate total tree biomass as well as a component ratio for foliage using models dependent on tree component and region (Domke et al., 2012). An estimate of foliage biomass (or other tree components) can then be obtained for an individual tree by multiplying these two quantities together, which relies on consistent patterns of resource allocation within species functional groups. The component ratio method (CRM) is currently used by the United States (U.S.) Forest Service to estimate biomass components in the national Forest Inventory and Analysis (FIA) program (Woodall et al., 2011). This approach is appealing because in that patterns of resource allocation vary in response to predictor variables (e.g., tree size, climate, geography; Reich et al., 2014). However, most published regression equations for these foliar components are based on small datasets, while patterns of biomass allocation across species remain poorly understood (Chojnacky et al., 2013).

Beyond establishing an appropriate functional form for foliage biomass, it is also important to consider how tree species will be grouped to develop general biomass equations. An obvious approach is to use species-specific models, but this may present some disadvantages. Because felled-tree data are generally required to parameterize biomass models, the absence or underrepresentation of a particular species in such datasets can hinder the development of species-specific models. Combining multiple species into functional groups, such as those proposed by Jenkins et al. (2003) and Chojnacky et al. (2013), can address this issue by allowing analysts to obtain biomass estimates using data from species assumed to be similar to those absent in observed datasets (Weiskittel et al., 2015). The determination of such functional groups is often subjective. The method of grouping may successfully categorize the most common species into unique groups, but determining how rare and/or infrequent species are accounted for remains an open question (Picard et al., 2010). For instance, the 10 species groups proposed by Jenkins et al. (2003) were based on coarse geographic and taxonomic factors, while the 35 groups established by Chojnacky et al. (2013) considered more refined phylogenetic and allometric relationships. It is important to note that neither of these studies were based on actual biomass observations but instead used pseudodata simulated from a range of published biomass models. This highlights the need for empirical

observations of felled-tree total and foliar biomass data to inform component biomass models.

Given that model development has primarily focused on total biomass, there is a clear need for a consistent methodology for estimating component biomass stocks that has been validated using felled-tree datasets. This includes both establishing an appropriate model functional form as well as ideal groupings for nationally consistent component biomass equations. In this study, we used an extensive database of felled tree data compiled from previous published and unpublished work to compare multiple models for tree foliage biomass for the primary species occurring across the U.S. We focused on foliage since it is a highly dynamic yet poorly understood carbon pool that is an important parameter in numerous nutrient budget models (e.g., PnET, CENTURY). Our specific objectives were to: (1) develop and compare multiple functional models for foliage biomass (i.e., simple models and component ratio models) at the species level; and (2) assess the performance of species-specific equations versus the broad species grouping proposed by Jenkins et al. (2003) and the narrower grouping of Chojnacky et al. (2013) (hereafter referred to as “broad” and “narrow” groupings, respectively) when fit to observed foliage biomass data.

2. Methods

2.1. Study area

Data used in our study were collected at 130 locations in the U.S. and Canada (Fig. 1). The majority of the sampling locations were in the southeast U.S., though data were also used from the northern states, at several locations from the western U.S. and Canada, and from one location in interior Alaska. Since previously published biomass data were employed, it should be noted that the number of species and sample sizes at each location varied widely.

2.2. Legacy biomass database

For our analysis, we used an extensive database of felled-tree biomass estimates, hereafter referred to as “legacy data,” that has been assembled as part of ongoing efforts to improve the volume, biomass, and carbon models used by FIA for conducting forest resource inventories (Weiskittel et al., 2015). The database contains original records of felled tree data, from both published and unpublished work, collected between 1960 and 2013. For our modeling work we only retained observations where leaf biomass, dbh, and total height were measured. We excluded individuals that were smaller than 12.7 cm dbh, as biomass allocation patterns in seedlings and saplings differ from that of mature trees and are typically modeled separately (Jenkins et al., 2003; Nelson et al., 2014). We also removed any species with fewer than four observations to maintain as much diversity as possible within the broad and narrow species groups we considered. These filters resulted in a final sample of 2885 observations from 49 studies. A total of 65 species, 24 of the 35 Chojnacky et al. (2013) narrow species groups, and all 10 Jenkins et al. (2003) broad species groups are represented in this sample (Fig. 2; A1).

2.3. Design of Bayesian hierarchical model

Bayesian hierarchical models are useful for the development of multispecies tree component biomass models because they account for the richness in correlation structures needed in such analyses, and the analysis of posterior distributions naturally allows for a complete assessment of uncertainty in biomass predictions (Finley et al., 2008; Green et al., 1999). In our analysis, we used the same error structure for each model we considered, so

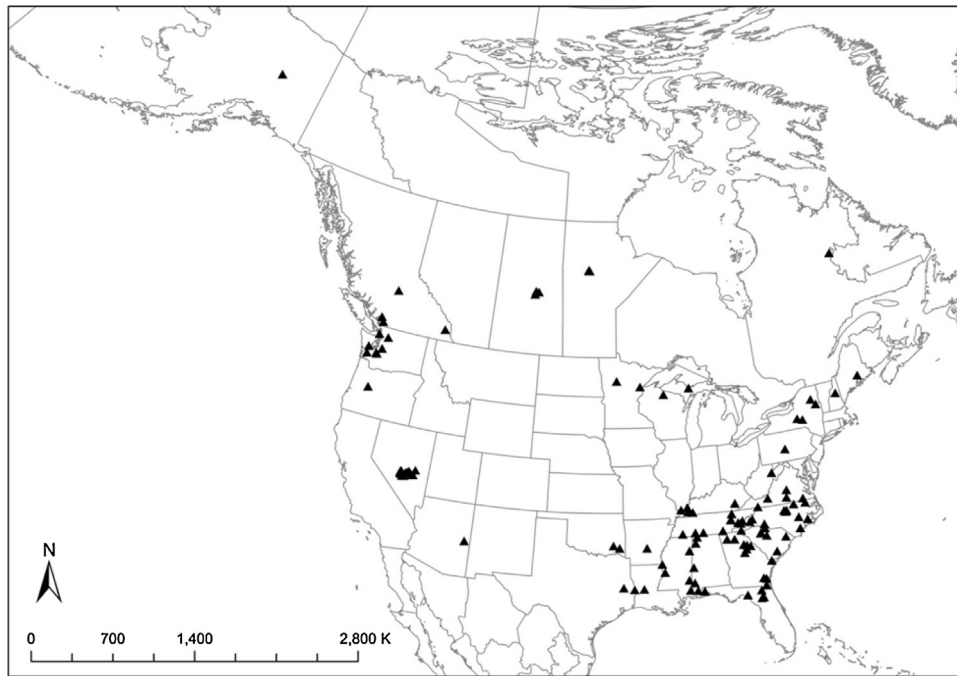


Fig. 1. Approximate locations for the 130 sampling sites in the legacy database for which measurements of foliage biomass, dbh, and height were available.

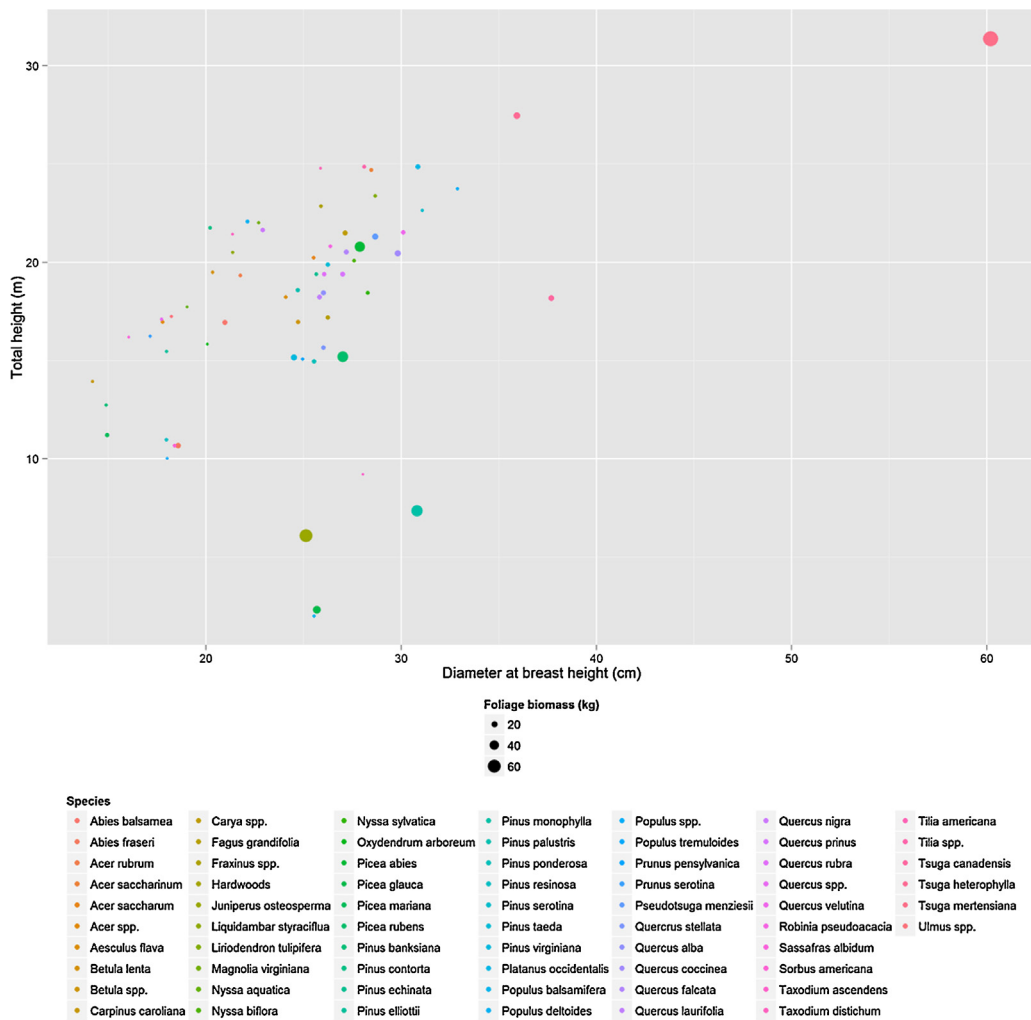


Fig. 2. Graphical summary of mean diameter-at-breast height (cm), total height (m), and foliage biomass (kg), by species, for the legacy data. A detailed tabular summary of the data may be found in the appendix.

differences will only be noted when warranted. A log-linear model for foliage biomass using the legacy data can be written as:

$$\ln(BM_{ij}) = \alpha_j + \beta_j \ln(dbh_{ij}) + \theta_k \mathbf{A} + \varepsilon, \quad (2)$$

where BM_{ij} is tree foliage biomass (kg) and dbh_{ij} is stem dbh (cm) for $i = 1, \dots, 2885$ observations, α_j and β_j are regression coefficients for j groups ($j = 65$ for species-specific models, 24 for narrow groups, or 10 for broad groups), θ_k represents a vector of k random effects to account for variation among the 49 studies in the database, \mathbf{A} is a corresponding $k \times i$ study design matrix, and $\varepsilon \sim N(0, \tau^2)$ is a zero-centered Gaussian distribution of white noise errors with variance denoted by τ^2 .

In Bayesian hierarchical modeling, our goal is to draw a reasonable posterior distribution of model parameter values (α_j , β_j , θ_k , and τ^2 in Eq. (2)) by integrating information from the data (the likelihood distribution) with a distribution of our prior belief about the parameters on which we wish to draw inference (Congdon, 2006). In a hierarchical model, we can further place hyper-prior distributions upon the parameters of these prior distributions (i.e., mean and variance for a normal prior). This property permits us to assume exchangeability among groups and to develop a nested error structure for our parameter estimates (Gelman et al., 2003). This is important in our work as we wish to estimate model parameters separately for our j species groups, but we believe that the process from which these estimates arise is similar across species. To accommodate this, we placed the following prior distributions on the regression coefficients:

$$\alpha_j \sim N(\theta_\alpha, \gamma_\alpha^2) \quad (3)$$

$$\beta_j \sim N(\theta_\beta, \gamma_\beta^2) \quad (4)$$

We assumed “weakly informative” prior distributions, as is common in applied hierarchical modeling, characterized by a flat shape and wide variance (γ^2) relative to expected estimates (Gelman, 2009, 2006). This allowed us to take advantage of the nested structure of hierarchical models even when little was known about the foliage biomass parameters a priori. In this case, we specified our hyperpriors as:

$$\theta \sim N(0, 1000) \quad (5)$$

and

$$\gamma^2 \sim Unif(0, x) \quad (6)$$

In the case of the variance, we chose a uniform range $(0, x)$ that includes the full range of expected values, but is small enough to allow for efficient convergence within a Markov chain Monte Carlo (MCMC) sampler (Gelman, 2006). Note that we allowed for exchangeability among our parameter estimates by estimating a different mean and variance for the j models, but allowed all of these to arise from the same hyper-prior distribution. We employed such a strategy because we assumed that, while differences in the relationship between foliage biomass and the independent variables among species/groups are expected, the underlying distributions are likely to overlap.

In addition to parameters related to the independent variables, we included a set of study random effects to account for the fact that the data we used was drawn from a number of different felled-tree biomass studies (Fig. 2; A1). This approach has previously been used in the context of linear mixed models for analyzing biomass data (de-Miguel et al., 2014; Moore, 2010; Pearce et al., 2010; Repola, 2008; Wirth et al., 2004; Wutzler et al., 2008). Here we again assumed exchangeability among the studies and placed the same vague priors on the mean and variance of the random effects coefficients as in (4) and (5). Finally, we completed the specification of our hierarchical model by placing a vague gamma distribution on the model variance:

$$\tau^2 \sim G(0.01, 0.01) \quad (7)$$

The hierarchical statistical framework outlined above is designed to address the challenges associated with analyzing compiled legacy biomass data where sampling methodologies, sample sizes, and distributions across species and size classes may vary substantially across locations. Two aspects of the model specifically account for these factors: (1) the inclusion of study ‘random effects’ to address variance related to study differences; and (2) the assumption of exchangeability among group-level estimates of model parameters. As discussed above the former strategy has been applied by others within a mixed modeling framework, but the latter is unique to Bayesian hierarchical model constructions since it arises as a consequence of specifying hyper-prior distributions on the priors around regression coefficients. While analyzing data compilations such as the legacy data undoubtedly introduces additional uncertainty, such resources remain important for calibrating biomass models in the absence of large-scale systematic surveys of tree allometric relationships. Fig. 3 presents a schematic diagram

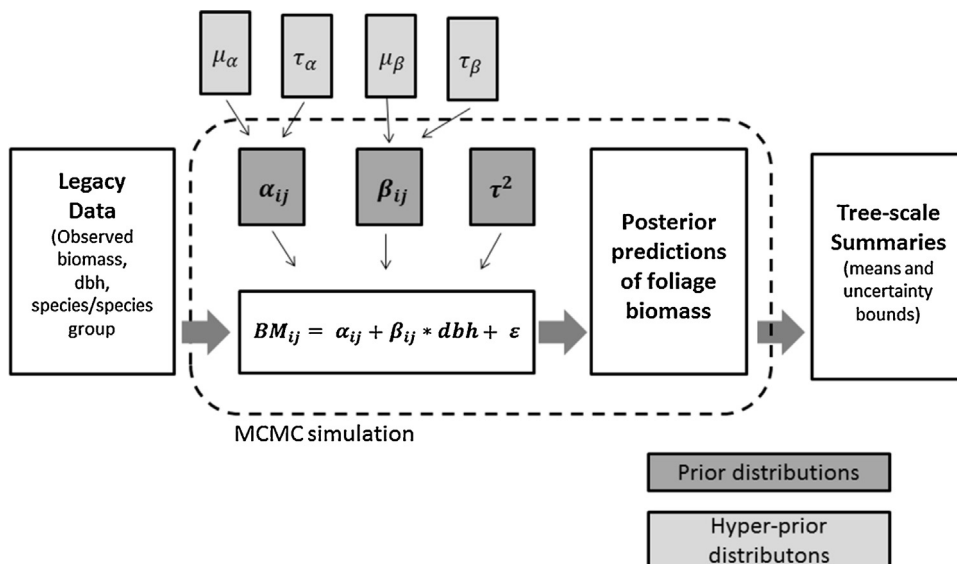


Fig. 3. Flow diagram depicting the structure of the hierarchical model and the principle steps of the model fitting and prediction process.

Table 1
Functional forms and references for the foliar biomass models compared in this study.

Equation	Ref.
$BM_{fol} = \alpha * dbh^\beta$	Bartelink (1998), Shaiek et al. (2011)
$\ln(BM_{fol}) = \alpha + \beta * \ln(dbh)$	Robichaud and Methven (1992)
$\ln(BM_{fol}) = \alpha + \beta_1 * \ln(dbh) + \beta_2 * \ln(h)$	Wirth et al. (2004), Wutzler et al. (2008)
$BM_{total} = Exp(\alpha_{BM} + \beta_{BM} * \ln(dbh))$	Jenkins et al. (2003)
$FR = Exp\left(\alpha_{FR} + \frac{\beta_{FR}}{dbh}\right)$	
$\ln(BM_{total}) = \alpha_{BM} + \beta_{BM} * \ln(dbh)$	Chojnacky et al. (2013)
$\ln(FR) = \alpha_{FR} + \beta_{FR} * \ln(dbh)$	

Variables in these models are foliage biomass (kg; BM_{fol}), diameter at breast height (cm; dbh), total tree height (m; h), total biomass (kg; BM_{total}), and foliage component ratio (FR). Note that the study ‘random effects’ and model errors are left off these equations for convenience, but the error structure described within the text was employed for all models.

of the hierarchical model structure in relationship to both the data inputs and output predicted summaries generated in our study.

2.4. Comparing multiple leaf biomass models

We considered several functional forms for predicting leaf biomass (Table 1), and examined both simple (i.e., models that estimate foliar biomass directly) and component ratio models using both linear and nonlinear functions. We generally confined our analysis to models that only included dbh as a predictor, since this is a common measurement available across national forest inventories. However, we did also include a log-linear model with height as a covariate since some studies suggest that relative dominance influences leaf biomass (Bartelink, 1998; Wirth et al., 2004). Variables related to crown dimensions, such as crown ratio or crown width, have also been used to predict foliage biomass (Laubhann et al., 2010; Nowak, 1996; Raymond and McKenzie, 2013). However, these measurements were not available for most of the database (Greater than 95% of observations lacked crown measurements).

To compare different functional forms for predicting leaf biomass, we conducted model comparisons using species-specific models. We assumed that the species-specific equations would give us the most precise estimates and that functional forms appropriate at the species level would be applicable for both the narrow and broad groupings as well. We assessed model fit to the legacy data by computing both mean absolute error (MAE) and root mean squared error (RMSE) for each function. All model fitting was accomplished via Gibbs sampling using the Just Another Gibbs Sampler (JAGS) software (Plummer, 2003), called from R via the rjags package (Plummer and Stukalov, 2014). Following a burn-in period where 10,000 iterations of sample estimates were discarded, we collected 1000 samples from each chain to compute MAE and RMSE by ‘thinning’ from a sample run period of 50,000 additional iterations to reduce autocorrelation between successive Markov chain samples. In preliminary runs of the model, this procedure resulted in posterior samples that satisfied tests for MCMC chain convergence with the Gelman and Rubin (1992) diagnostic as well as by visual inspection of trace plots and posterior distributions. The best fitted model, as determined by MAE and RMSE, was retained for the remainder of our study.

2.4.1. Comparing species grouping schemes for generalized leaf biomass models

To address the performance of different species groupings, we tested the ability of each model scheme to generate new leaf biomass datasets given the observations in the legacy database via Bayesian posterior predictive checks (Gelman et al., 1996; Meng,

1994). This approach uses the fitted model to generate many simulated datasets which can then be evaluated with exploratory graphics and/or frequentist statistical tests (Gelman et al., 2000). Since we were operating within a Bayesian framework and thus had access to posterior distributions of the model parameters, we simulated new datasets by taking draws from the posterior predictive distribution (Ntzoufras, 2009):

$$p(y_{new}^i | y, x, \alpha) = \int p(y_{new}^i | \theta) p(\theta | y, x) d\theta \tag{8}$$

where y_{new}^i represents the i th simulated datasets ($i = 1, 2, \dots, 1000$), y and x are the dependent and independent variables respectively, θ is the model parameters, and α is the hyper-parameters placed on the prior distributions of θ . These simulated data were then summarized by their mean and 95% uncertainty interval bounds to explore how well each model approximated the distribution of the observed data (Gelman, 2003). The new datasets were of the same dimension as the fitting data ($n = 2885$) and were generated following a burn-in period of 10,000 iterations to ensure the model had converged, followed by thinning from a larger sample of 50,000 to reduce autocorrelation within the MCMC chain. This procedure resulted in 1000 observations of each test statistic, which were compared to the mean and bounds of the observed data by plotting histograms, as well as by computing a Bayesian p -value to test whether the observed quantiles fell within the 95% uncertainty interval of our posterior predictive results (Meng, 1994). Lastly, to assess the relative predictive performance of each species grouping, we computed root mean squared prediction error (RMSPE) between observed and predicted values of foliar biomass for each of the 1000 simulated datasets and compared the means and uncertainty bounds of these distributions.

3. Results

3.1. Comparing functional forms for species-specific leaf biomass models

Mean absolute error ranged from 3.64 to 4.85 kg for the candidate models, while RMSE ranged from 7.69 to 14.27 kg (Table 2). The log-linear function that predicted log-transformed foliar biomass with $\ln(dbh)$ and $\ln(\text{height})$ provided the best fit according to MAE, while the nonlinear function with dbh as the only predictor was preferred according to RMSE. The log-linear component ratio model provided a better fit than the nonlinear equations proposed by Jenkins et al. (2003), though neither fit the data as well as any of the simple equations. Based on these results, we retained the log-linear model with dbh and height as covariates for the remainder of our study. While the nonlinear function had a lower RMSE, the difference between these two functions was negligible (<1%), while the improvement offered by the log-linear model according to MAE was approximately 15.5%.

Table 2
Mean absolute error (MAE) and root mean squared error (RMSE) for the five models compared in this study.

Equation	MAE (kg)	RMSE (kg)
$BM_{fol} = \alpha * dbh^\beta$	4.3	7.69 ^a
$\ln(BM_{fol}) = \alpha + \beta * \ln(dbh)$	4.56	14.27
$\ln(BM_{fol}) = \alpha + \beta_1 * \ln(dbh) + \beta_2 * \ln(h)$	3.64 ^a	7.74
$BM_{total} = Exp(\alpha_{BM} + \beta_{BM} * \ln(dbh))$	4.68	12.48
$FR = Exp\left(\alpha_{FR} + \frac{\beta_{FR}}{dbh}\right)$		
$\ln(BM_{total}) = \alpha_{BM} + \beta_{BM} * \ln(dbh)$	4.85	12.53
$\ln(FR) = \alpha_{FR} + \beta_{FR} * \ln(dbh)$		

^a Denotes the best fitted model according to MAE and RMSE. Note that for log-linear models, both observed data and model predictions were back-transformed prior to computing the test statistics.

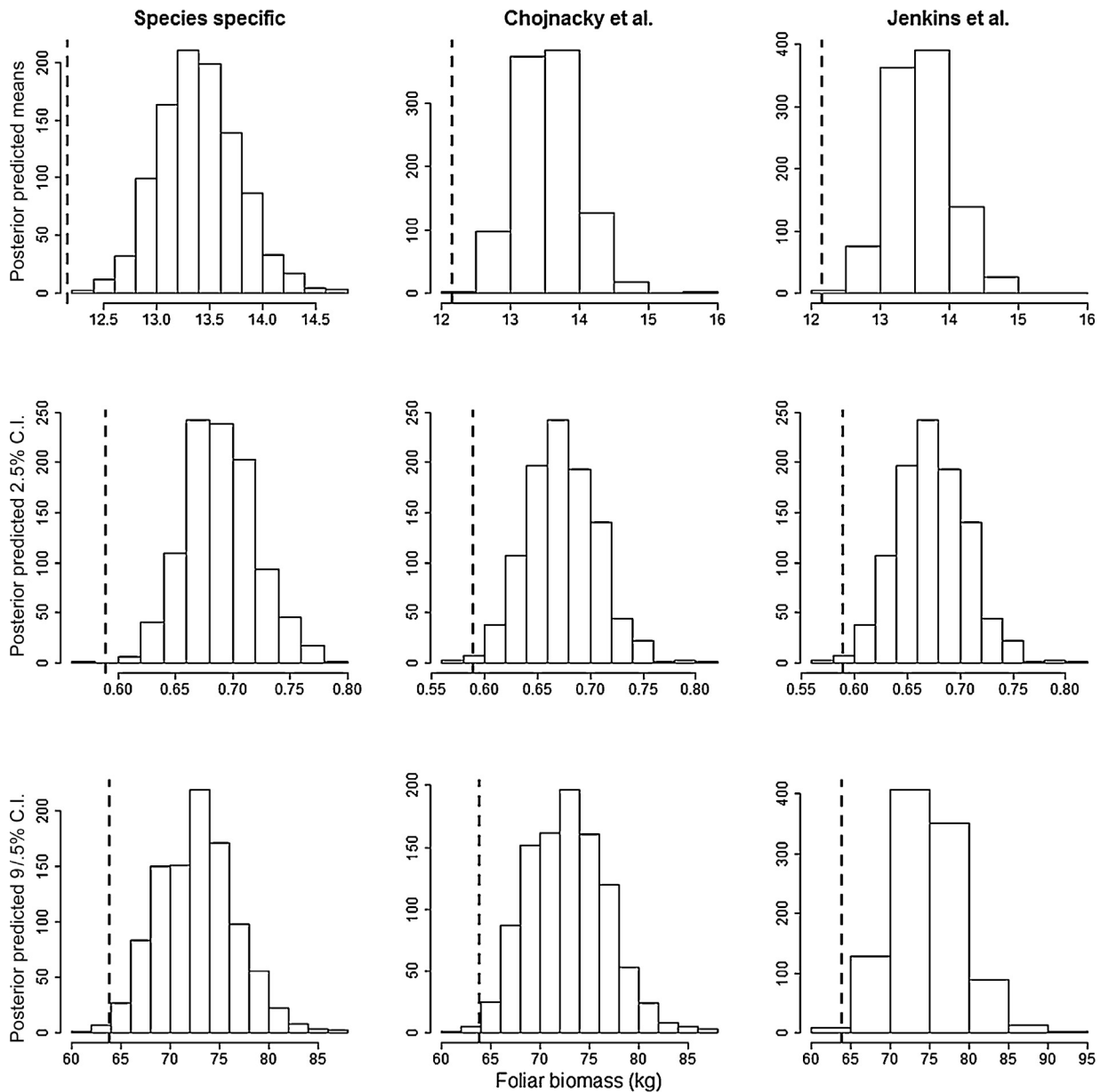


Fig. 4. Posterior predicted quantiles (mean and lower and upper uncertainty bounds) for the 1000 simulated datasets. The dashed line represents the quantiles from the observed data.

3.2. Comparing groupings for generalized leaf biomass equations

All three model schemes resulted in overpredictions of the mean, upper uncertainty bounds, and lower uncertainty bounds (Fig. 4). For all three species groupings, the observed mean was outside of the 95% uncertainty interval of all three quantiles. Hence, these tests demonstrated that even our best fitted model resulted in overprediction across the full range of the distribution of the observed data, regardless of the grouping used. The results of these posterior predictive checks may be more closely examined by comparing the predicted and observed results for each species grouping (Fig. 5). For most of the observations, predictive performance did not vary much between the three species grouping schemes. However, in each case, there were a handful of observations where foliage biomass was significantly overpredicted with associated uncertainty very large relative to the distribution of the data. These observations contributed to the overprediction of the mean by each dataset. The large prediction uncertainties around these

observations doubtlessly contributed to the overprediction of the upper credible interval by all three model schemes. Interestingly, these observations did not come from the same dataset, but instead arose from several different studies within the legacy database and represent several species. The common feature is that they are all among the largest trees, in terms of dbh, represented in the data we analyzed.

Even for observations with narrower credible intervals, the range of uncertainty is generally large relative to the mean estimates, suggesting considerable prediction uncertainty when generating new biomass datasets. There is a general trend of increasing uncertainty with increasing foliar biomass. This result corresponds with the composition of the legacy data, which is comprised mainly of smaller trees (Fig. 2). The species-specific equations had a lower mean prediction uncertainty than the models fit to the narrow or broad groups. When considering average performance over the 1000 simulated datasets, species specific equations offered relative improvement of about 10% over each of

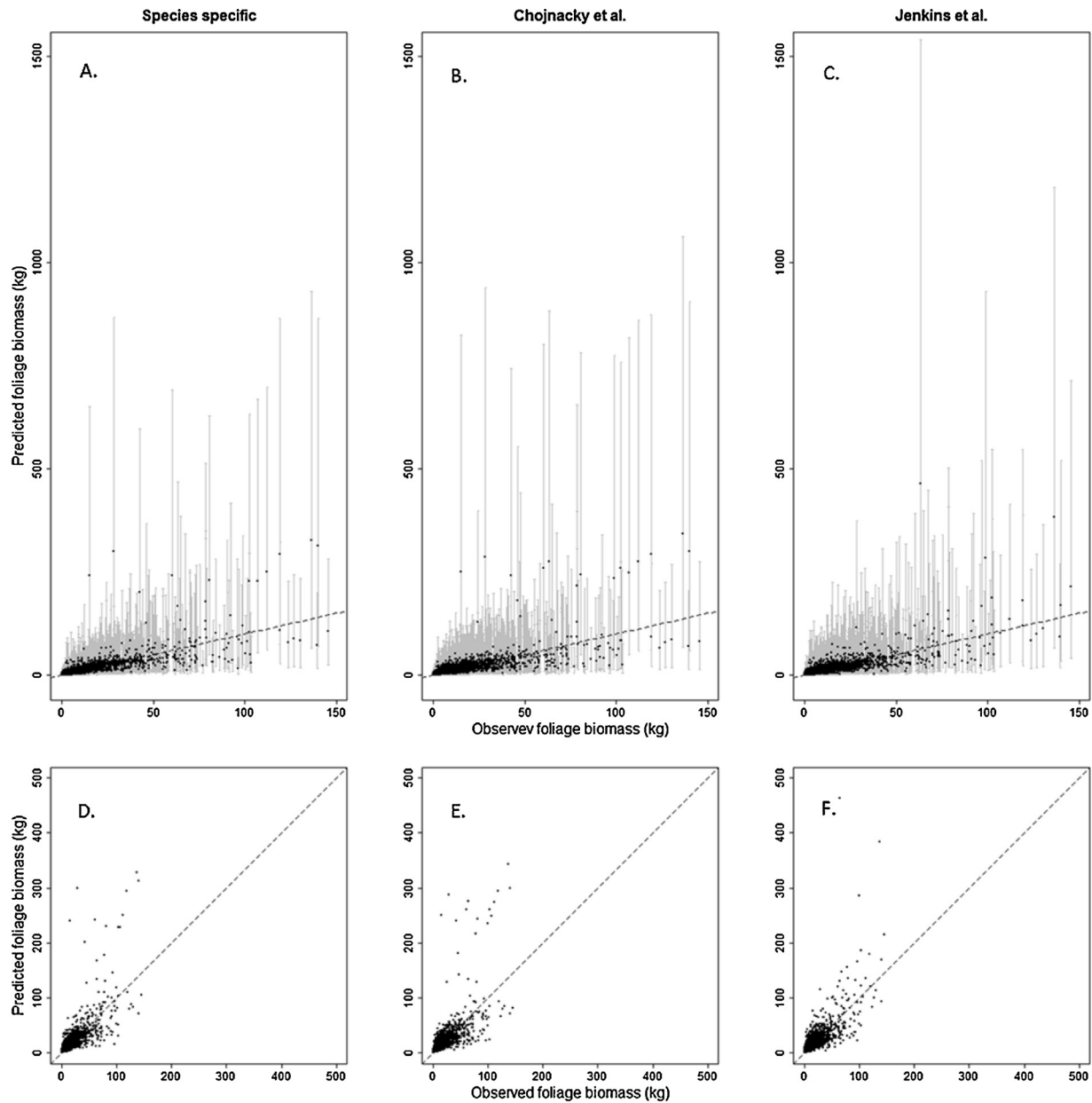


Fig. 5. Observed vs. predicted leaf biomass for the full dataset, using each species grouping. The error bars in the upper panels (Fig. 4A–C) represent the 95% uncertainty intervals of the posterior predictive samples from the 10,000 simulated datasets. The lower panels (Fig. 4D–F) show only the posterior means on a reduced set of axes, to better visualize the overall accuracy of each species grouping.

Table 3

Mean root mean squared prediction error (RMSPE) and associated lower and upper quantile bounds from 1000 simulated datasets.

	Mean RMSPE (kg)	2.5% C.I. (kg)	97.5% C.I. (kg)
Species specific	16.22	12.28	25.77
Narrow grouping (Chojnacky et al., 2013)	18.03	13.3	31.46
Broad grouping (Jenkins et al., 2003)	18.19	13.59	30.58

the other groupings (Table 3). While the range of computed RMSPE values for the 1000 trials overlap significantly for each grouping, the species specific equations also have the lowest overall 2.5 and 97.5 percent quantile bounds.

4. Discussion

The legacy data we analyzed afforded us the opportunity to compare multiple allometric models of foliage biomass for the major

tree species in North America. We found that models which directly estimate foliage biomass provided a better fit than component ratio models, and that including height as a covariate improved the fit of a log-linear function. Even among studies focused on single species or small groups of species, there has been little comparison of different allometric relationships for foliage biomass. At the individual tree scale, most work has been devoted to analyzing the effects of including additional covariates on model fit, particularly variables related to crown attributes (Bartelink, 1996; Loomis et al., 1966;

Monserud and Marshall, 1999; Tobin et al., 2006). However, no previous studies have compared these types of equations to component ratio models using felled-tree data, despite the fact that the latter are widely applied to estimate component biomass stocks at the national scale (Domke et al., 2012; Jenkins et al., 2003). The poor relative performance of component ratio models in our analysis highlights the difficulty of generalizing biomass allocation patterns, even within an individual species. The variation in foliage biomass is large and driven by a suite of landscape-level and stand-level processes (Temesgen et al., 2011) that is exacerbated when only basic tree measurements such as dbh are available (Grier and Waring, 1974; Marshall and Waring, 1986; Raymond and McKenzie, 2013; Smithwick et al., 2002). This effect was particularly problematic for large trees in our study and has been noted by other authors (Van Pelt and Sillett, 2008). In the case of large scale national forest inventories where detailed data on variables related to biomass allocation are typically not available (Weiskittel et al., 2015), simple allometric functions employing basic mensuration data (e.g., dbh and height) may provide better predictive performance.

Regardless of the functional form used, national forest carbon inventories require generalized biomass equations both to account for species that are poorly represented in felled tree datasets and to provide a consistent method for use by both government agencies and non-governmental organizations (Cháve et al., 2014; Jenkins et al., 2003). Reliable general equations, such as those of Jenkins et al. (2003) and Chojnacky et al. (2013), have been developed for total aboveground biomass for many forest ecosystems worldwide. These grouping schemes generally rely on similar patterns that emerge when analyzing multi-species datasets and appeal to allometric scaling theory as a unifying concept for applying one set of coefficients to two or more species (Chojnacky et al., 2013; Enquist and Niklas, 2001). In the case of total biomass, comparative studies have found that generalized models can perform as well as models derived from fitting to site-specific data (Case and Hall, 2008; Cháve et al., 2014; de-Miguel et al., 2014; Nelson et al., 2014). However, most studies have focused on species-specific models. It is unclear if such models are capable of accurately predicting biomass components when fit to multi-species datasets. In their pseudo-data analysis, Jenkins et al. (2003) were unable to resolve general equations for biomass components beyond separating hardwoods and softwoods, though this may be in part due to the paucity of published equations for predicting biomass components such as foliage. Since we had felled-tree data available for our analysis, we were able to apply the broad species grouping (Jenkins et al., 2003) and narrow species grouping (Chojnacky et al., 2013) used for total aboveground biomass, though our model checking procedure found that both groupings produced less accurate results than species-specific equations when simulating new datasets. It is particularly interesting to note that the models fit to broad groupings performed no worse than those that relied on narrow groups, which incorporated wood specific gravity to rigorously group species (Chojnacky et al., 2013). As these results are derived from diverse data sources, it suggests that assuming common allometric relationships across multiple taxa when developing biomass equations may be more problematic for determining component stocks than it is for total biomass. Given the dynamic nature of foliage, such patterns may be more difficult to generalize. Therefore, additional input variables (e.g., tree crown characteristics) may be required to improve prediction within the context of national forest inventories (Raymond and McKenzie, 2013; Temesgen et al., 2011).

Our results have important implications when determining forest carbon stocks and fluxes for reporting in national greenhouse gas inventories. Although foliage biomass is typically a small component of the total forest carbon stock, errors when calculating foliage biomass using individual tree models can be important when scaled across a large region with diverse forest types and

climatic conditions (e.g., U.S.). Error in foliage biomass estimates is also important for parameterizing ecosystem models such as CENTURY since it is the primary annual input to soil carbon models (Kirschbaum and Paul, 2002), and therefore, may have a disproportionate effect on uncertainty in forest soil carbon projections. We encountered difficulty in accurately estimating an allocation ratio for leaf biomass, and therefore, found simple regression equations to fit our data better than component ratio models. Further, we confirmed difficulties associated with estimating foliage biomass using only stem measurements such as dbh as inputs, which tends to result in overprediction since it fails to account for crown structure and the ratio between dbh and sapwood area (Grier and Waring, 1974; Marshall and Waring, 1986; Raymond and McKenzie, 2013). Our findings may have implications for estimating additional tree components, as pointed out by Reich et al. (2014) when observing foliage-belowground tradeoffs related to climate. Tree roots are similarly dynamic, with their morphology and total biomass influenced by belowground competition in much the same way that foliage is influenced by crown dynamics (Lin et al., 2014). While it is unknown whether the same patterns we observed for foliage will be true for other biomass components (Russell et al. Submitted for publication), we believe our results highlight a need for such analyses using felled-tree data such as found in the legacy database (Weiskittel et al., 2015).

The large uncertainties surrounding our predictions of foliage biomass (i.e., Fig. 3), regardless of the species grouping used, is an important benchmark from this work. We used posterior predictive model checking to simulate many “new” datasets and to compare features of these distributions to the observed data. Generalized models (i.e., Jenkins et al., 2003; Chojnacky et al., 2013) are designed to be applied to predict biomass given independent data. The approach presented herein provides a better approximation to “out of sample” prediction than within-sample estimates (Hooten and Hobbs, 2015; Vehtari and Ojanen, 2012). That the ratio of our uncertainty bounds relative to the prediction means was so large is important and suggests that using methods developed for total aboveground biomass may lead to systematic underestimation of uncertainty when predicting component biomass stocks. Uncertainty in biomass estimates has received considerable attention within the literature with many studies having used MCMC approaches to account for the error in model coefficients (Lehtonen et al., 2007; Wayson et al., 2014; Wirth et al., 2004). However, in most cases such procedures require the analyst to make a priori assumptions about the dimension of the distributions underlying model parameters (Zapatas-Cuartas et al., 2012). By contrast, hierarchical models allow these quantities to be estimated by fitting the model to observed data, such as we have done here. By also drawing predictions from a posterior distribution, uncertainty associated with both model parameters and data is seamlessly incorporated into new estimates (Ntzoufras, 2009). These features make hierarchical models an attractive option for developing predictions in national forest inventories, particularly when a comprehensive felled tree biomass database is available for model development.

The converse, however, is that predictions conditioned upon felled-tree observations will reflect any biases that exist in the fitting datasets. The legacy data we used is compiled from many studies with differing sampling methodologies, diameter distributions, and sample sizes. The variable nature of these data makes it difficult to specifically account for these factors within a predictive model, beyond including a set of generic ‘study random effects’ such as we did in our analysis. Improving the geographic and taxonomic coverage of the legacy data, as well as collecting more samples of large trees for the most common species of North America, will provide a better resource for understanding contributions of different sources of uncertainty to allometric models fitted with felled-tree datasets. Efforts are already under way to make these

improvements (Weiskittel et al., 2015), but in the meantime the mean and uncertainty estimates we report must be understood within the context of the data available for our study.

Beyond the need for improved data resources, results of our work highlight several additional directions for improving component biomass models. Our work confirmed other studies where models based on simple measurements such as dbh and height led to systematic overprediction of foliage biomass (Grier and Waring, 1974; Marshall and Waring, 1986; Raymond and McKenzie, 2013). Indeed, variables related to crown structure may be important for refining leaf biomass estimates (Bartelink, 1996; Temesgen et al., 2011). As a potential application, the U.S. Forest Service collects tree crown indicator information on a subset of their standard inventory plots (Randolph, 2013), with the potential for crown measurements to occur on up to 25% of all FIA plots in the northeastern U.S. states. While these measurements represent a sample of the full FIA database, it may be possible to refine foliage biomass models for more broad application through the implementation of adjustment factors. Remote sensing technologies, including satellite-based sensors and light detection and ranging (LiDAR) systems, have been used to estimate leaf area index (LAI) and crown features at the stand and landscape level (Detto et al., 2015; Lefsky et al., 1999; Song, 2013), but little research has focused on integrating these data with estimates from individual tree models. Integrating a hierarchical statistical model with remote sensing data and landscape analyses is a promising approach for up-scaling forest carbon stocks across large areas. When only basic tree measurements such as dbh and height are available, however, our results indicate that directly estimating foliage biomass based on these predictors is preferable to estimation as a ratio of total aboveground biomass.

The hierarchical statistical framework used in our analysis is an ideal approach for monitoring and verifying forest carbon stocks, both within the context of local carbon accounting projects and national greenhouse gas inventories. The uncertainty represented in tree-scale posterior predictive distributions naturally accounts for both data-level and parameter error, and may

be easily propagated into population-level uncertainty estimates. Population-scale errors allow for robust verification of carbon stock changes by facilitating the use of basic statistical principles (i.e., testing whether changes in carbon stocks significantly differ from zero), which is not possible when fixed estimates of stocks are available. In addition a hierarchical model allows for the use of felled-tree data to constrain predictive models, eliminating the need to make prior assumptions about allometric relationships. For local accounting projects, the same general framework can be used with site or region specific felled-tree data, where available. For application to NGHGs, global analyses, or regional analyses where local data are not available, models may be fitted using general tree allometry databases. While in the past such resources have not been made public some are now available, including the global Biomass and Allometry Database for woody plants (BAAD; Falster et al., 2015). The legacy data for the United States and Canada used in our study will also soon be made public to facilitate development and calibration of allometric equations for North America. Our hope is that the availability of such data, coupled with analytical techniques such as we demonstrate here, will provide both analysts and managers with improved tools for accurately and precisely estimating forest carbon stocks.

Acknowledgements

We wish to thank David Walker, Jereme Frank, Aaron Weiskittel, and all who have contributed to the U.S. Forest Service Volume Biomass Project and the Legacy database. In addition we would like to thank John Stanovick, David Bell, Kenneth Elgersma, and three anonymous reviewers for their comments on our manuscript. This research is funded by the U.S. Department of Agriculture, Forest Service, Northern Research Station and the Minnesota Agricultural Experiment Station.

A1.

Sample sizes, summary statistics, and references for the 65 species in the legacy dataset.

Species	N	dbh (cm)			Height (m)			Foliage biomass (kg)			# Studies	Refs.
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max		
Conifers												
<i>Abies balsamea</i>	12	20.98	14.73	29.46	16.94	13.59	20.60	14.02	3.71	27.22	1	44
<i>Abies fraseri</i>	8	18.57	12.75	24.90	10.68	9.18	12.72	17.04	4.92	40.34	2	36
<i>Juniperus osteosperma</i>	21	25.14	13.00	50.50	6.09	4.10	9.40	59.26	12.40	145.50	10	31
<i>Picea abies</i>	29	27.89	13.00	43.70	20.78	13.90	25.10	44.63	4.30	139.70	1	48
<i>Picea glauca</i>	33	25.68	13.20	50.00	2.33	0.82	11.09	28.58	1.86	103.01	2	16, 49
<i>Picea mariana</i>	15	14.94	12.80	20.60	11.22	0.94	16.87	9.55	1.43	18.16	7	16, 39, 49, 50
<i>Picea rubens</i>	47	27.02	14.80	71.60	15.19	8.61	32.80	47.44	5.42	130.49	3	36, 46
<i>Pinus banksiana</i>	4	14.90	13.90	16.80	12.73	11.50	14.40	5.25	4.00	7.00	1	16
<i>Pinus contorta</i>	28	20.21	13.00	37.00	21.76	15.80	28.20	7.26	0.50	36.60	1	11
<i>Pinus echinata</i>	98	25.65	12.80	55.12	19.40	9.45	29.26	7.10	0.10	34.79	7	10, 17, 35, 42, 46, 47
<i>Pinus elliotii</i>	21	17.99	13.00	31.20	15.46	13.08	21.90	6.06	1.40	24.60	2	14, 29
<i>Pinus palustris</i>	13	24.70	18.50	31.00	18.60	16.20	22.70	9.67	3.40	22.00	2	14, 26
<i>Pinus ponderosa</i>	37	25.54	12.70	41.91	14.96	4.57	22.25	12.32	2.72	34.49	1	21
<i>Pinus resinosa</i>	25	17.97	12.70	28.70	10.97	0.10	18.96	7.17	0.00	30.16	3	23, 37, 45
<i>Pinus serotina</i>	18	31.09	14.00	46.20	22.63	13.41	26.79	4.51	0.51	10.39	1	26
<i>Pinus taeda</i>	243	26.25	12.70	59.94	19.87	5.21	35.05	11.69	0.19	63.10	11	2, 3, 10, 22, 35, 40, 46, 51
<i>Pinus virginiana</i>	6	24.51	17.27	44.20	15.16	7.21	21.34	19.15	4.76	45.17	2	10, 38
<i>Pinus monophylla</i>	58	30.82	12.70	115.60	7.35	4.00	14.00	51.01	6.40	140.20	18	31
<i>Pseudotsuga menziesii</i>	130	28.66	12.70	96.80	21.30	8.73	60.00	19.69	0.60	93.40	10	1, 12, 15, 19, 20, 24, 27
<i>Taxodium ascendens</i>	13	28.04	12.70	55.63	9.21	0.31	8.04	0.3	0.14	38.10	2	10, 32
<i>Taxodium distichum</i>	10	21.36	12.70	33.02	21.43	12.19	25.91	1.49	0.45	2.99	3	10
<i>Tsuga canadensis</i>	9	37.69	12.70	85.10	18.18	10.16	32.60	18.43	1.71	57.08	2	43, 46
<i>Tsuga heterophylla</i>	19	35.92	16.00	61.70	27.46	15.16	37.30	23.12	3.10	47.90	2	19, 25
<i>Tsuga mertensiana</i>	5	60.20	43.70	76.20	31.37	24.55	37.16	70.62	35.50	90.50	1	25

A1 (Continued)

Species	N	dbh (cm)			Height (m)			Foliage biomass (kg)			# Studies	Refs.
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max		
Hardwoods												
<i>Acer</i> spp.	4	17.77	13.20	25.90	16.97	12.65	24.96	6.77	1.53	15.58	1	26
<i>Acer rubrum</i>	142	21.76	12.70	52.40	19.34	12.19	33.55	6.35	0.41	44.09	16	8, 10, 13, 28, 33
<i>Acer saccharinum</i>	14	28.47	13.72	49.02	24.69	16.15	29.57	8.50	1.91	22.04	1	10
<i>Acer saccharum</i>	54	25.51	14.22	60.70	20.24	14.00	32.40	6.58	0.36	22.04	3	4, 33, 46
<i>Aesculus flava</i>	10	24.09	12.70	48.00	18.23	9.07	29.40	7.56	0.18	33.67	3	36, 46
<i>Betula</i> spp.	44	24.71	12.70	46.99	16.97	1.27	29.87	9.35	0.18	47.67	8	10, 49
<i>Betula lenta</i>	20	20.35	12.95	39.60	19.49	14.94	28.20	5.51	0.65	18.59	3	13, 28, 46
<i>Carpinus caroliniana</i>	8	14.19	12.70	16.00	13.94	9.75	17.68	4.08	2.27	6.35	2	10
<i>Carya</i> spp.	102	27.14	12.70	58.67	21.48	9.75	36.58	15.12	0.36	82.92	17	5, 10, 28, 46
<i>Fagus grandifolia</i>	21	26.25	14.48	51.60	17.19	9.39	23.40	11.67	1.41	96.33	4	8, 10, 36, 46
<i>Fraxinus</i> spp.	88	25.89	12.95	59.69	22.86	8.23	32.61	6.83	0.14	45.45	12	10, 46
<i>Liquidambar styraciflua</i>	380	21.36	12.70	48.01	20.50	11.58	33.22	3.91	0.09	29.57	32	10, 26, 46
<i>Liriodendron tulipifera</i>	107	28.67	13.21	55.80	23.38	12.80	39.69	6.75	0.05	24.99	13	10, 13, 26, 28, 46
<i>Magnolia virginiana</i>	5	22.71	16.76	32.51	22.01	17.68	26.82	4.54	1.41	11.97	2	10
<i>Nyssa aquatica</i>	56	19.04	12.95	50.55	17.73	12.19	21.34	1.96	0.50	7.98	5	10, 26
<i>Nyssa biflora</i>	71	27.58	13.21	46.74	20.08	10.36	28.04	6.31	0.45	33.70	1	10
<i>Nyssa sylvatica</i>	20	28.29	13.97	45.21	18.44	10.36	30.48	8.57	1.32	24.36	6	10
<i>Oxydendrum arboreum</i>	11	20.08	13.21	34.60	15.83	10.67	20.50	2.40	0.64	4.37	4	10, 26, 28
<i>Platanus occidentalis</i>	31	30.86	12.95	49.78	24.85	14.63	35.05	14.37	0.23	43.41	8	10
<i>Populus</i> spp.	9	32.89	13.65	50.80	23.73	1.44	32.92	5.03	0.45	10.57	4	10, 18, 36
<i>Populus balsamifera</i>	11	25.54	14.50	46.50	2.00	1.51	2.55	4.78	1.54	8.23	1	49
<i>Populus deltoides</i>	43	22.14	12.70	36.83	22.06	12.47	34.60	8.22	1.13	22.44	2	9, 34
<i>Populus tremuloides</i>	27	18.01	13.06	30.50	10.02	0.78	23.60	2.52	0.76	6.28	3	7, 41, 49
<i>Prunus pensylvanica</i>	6	24.95	13.80	41.90	15.07	1.41	27.00	6.13	1.05	19.70	3	18, 46
<i>Prunus serotina</i>	17	17.14	13.21	23.37	16.24	12.16	20.73	5.62	0.39	15.00	6	10, 13, 36
<i>Quercus</i> spp.	12	17.73	13.46	28.70	17.10	14.17	22.80	4.45	0.17	14.92	2	5, 26
<i>Quercus alba</i>	202	26.02	12.70	64.00	18.45	9.45	33.70	15.18	1.22	89.04	19	5, 10, 26, 28, 46
<i>Quercus coccinea</i>	78	29.82	12.95	56.39	20.45	13.41	32.00	19.12	1.63	82.33	10	10, 28, 46
<i>Quercus falcata</i>	50	27.19	12.95	47.75	20.51	12.50	26.82	14.99	0.95	57.56	8	10
<i>Quercus laurifolia</i>	33	25.81	12.95	43.94	18.24	11.28	27.43	12.72	0.86	51.44	4	10
<i>Quercus nigra</i>	75	22.92	12.70	42.67	21.64	13.11	30.48	11.67	0.77	55.61	9	10
<i>Quercus prinus</i>	71	27.00	12.70	57.50	19.40	8.84	29.50	12.47	1.00	48.40	6	10, 28
<i>Quercus rubra</i>	21	26.06	12.95	72.90	19.40	12.30	30.60	9.86	2.13	40.65	6	10, 28, 46
<i>Quercus stellata</i>	28	26.02	12.95	53.09	15.65	10.06	20.12	11.06	1.59	36.20	3	10, 46
<i>Quercus velutina</i>	39	30.11	12.70	55.37	21.52	12.50	29.57	12.09	0.09	44.59	5	10
<i>Robinia pseudoacacia</i>	24	26.37	14.22	41.66	20.82	10.97	30.18	7.61	1.81	20.09	4	6, 10, 46
<i>Sassafras albidum</i>	4	16.05	13.46	21.84	16.20	13.40	22.25	2.43	0.68	4.04	2	10, 46
<i>Sorbus americana</i>	4	18.39	12.80	25.70	10.68	8.78	12.59	6.97	2.73	12.34	2	36, 46
<i>Tilia</i> spp.	18	28.11	14.22	41.91	24.86	19.20	32.31	9.08	0.14	22.82	1	10
<i>Tilia americana</i>	4	25.88	17.80	32.80	24.78	23.00	26.30	2.84	0.57	4.95	1	46
<i>Ulmus</i> spp.	19	18.23	13.46	28.45	17.25	12.80	23.47	4.08	0.82	18.37	10	10

¹(Archibald, 1983), ²(Baker, 1962; Baker et al., 1974), ³(Baldwin, 1987), ⁴(Bickelhaupt et al., 1973), ⁵(Blackmon et al., 1977), ⁶(Boring and Swank, 1984), ⁷(Bray and Dudkiewicz, 1963), ⁸(Briggs et al., 1989), ⁹(Carter and Smith, 1971), ¹⁰(Clark et al., 1986a, 1986b, 1985), ¹¹(Comeau, 1976), ¹²(Devine et al., 2013), ¹³(Adams et al., 2012), ¹⁴(Garbett, 1977), ¹⁵(Gholz, 1979), ¹⁶(Gower and Vogel, 1999), ¹⁷(Gyawali, 2008), ¹⁸(Harrison, 1981), ¹⁹(Harrison et al., 2009), ²⁰(Heilman, 1961), ²²(Houser, 1980), ²³(Hutnik, 1964), ²⁴(Kline et al., 1973), ²⁵(Krumlik and Kimmins, 1973), ²⁷(Marshall and Waring, 1986), ²⁸(Martin et al., 1998), ²⁹(Mead, 1971), ³⁰(Metz and Wells, 1965), ³¹(Miller et al., 1981), ³²(Mitsch and Ewel, 1979), ³³(Mroz et al., 1984), ³⁴(Mueller, 1976), ³⁵(Neisch, 1980), ³⁶(Smith, 1977), ³⁷(Nowak, 1996), ³⁸(Olah, 1972), ³⁹(Rencz and Auclair, 1980), ⁴⁰(Roth, 2010), ⁴¹(Ruark et al., 1987), ⁴²(Sabatia, 2007), ⁴³(Santee, 1970), ⁴⁴(Shannon, 1976), ⁴⁵(Singer, 1964), ⁴⁶(Sollins and Anderson, 1971), ⁴⁷(Vaidya, 1961), ⁴⁸(Jokela et al., 1986), ⁴⁹(Yarie et al., 2007), ⁵⁰(Rencz and Auclair, 1980), ⁵¹(Metz and Wells, 1965).

References

- Adams, M.B., Edwards, P.J., Ford, W.M., Schuler, T.M., Thomas-Van Gundy, M., Wood, F., 2012. Fernow experimental forest: research history and opportunities. In: US Forest Service Gen. Tech. Re EFR-2.
- Archibald, R.M., 1983. Effects of Nitrogen Fertilization on Allometric Relationships in Douglas-Fir in Western Washington. University of Washington, Seattle, Washington.
- Baker, D.J., Richards, G., Grainger, A., Gonzalez, P., Brown, S., DeFries, R., 2010. Achieving forest carbon information with higher certainty: a five-part plan. *Environ. Sci. Policy* 13, 249–260.
- Baker, J.A., 1962. Dry Matter Production in Plantations of Loblolly Pine. Duke University.
- Baker, J.B., Switzer, G.L., Nelson, L.E., 1974. Biomass production and nitrogen recovery after fertilization of young loblolly pines. *Soil Sci. Soc. Am. J.* 38, 958–961.
- Baldwin, V.C., 1987. Green and dry-weight equations for above-ground components of planted Loblolly Pine trees in the west Gulf Region. *South. J. Appl. For.* 11, 212–218.
- Bardgett, R.D., Mommer, L., De Vries, F.T., 2014. Going underground: root traits as drivers of ecosystem processes. *Trends Ecol. Evol.* 29, 692–699.
- Bartelink, H.H., 1998. A Model of Dry Matter Partitioning in Trees, pp. 91–101.
- Bartelink, H.H., 1996. Allometric relationships on biomass and needle area of Douglas-fir. *For. Ecol. Manag.* 86, 193–203, [http://dx.doi.org/10.1016/S0378-1127\(96\)03783-8](http://dx.doi.org/10.1016/S0378-1127(96)03783-8).
- Bickelhaupt, D.H., Leaf, A.L., Richard, N.A., 1973. Effect of branching habit on above-ground dry weight estimates of *Acer saccharum* stands. In: Young, H.E. (Ed.), IUFRO Biomass Studies. University of Maine, Nancy, France and Vancouver, BC, pp. 221–230.
- Blackmon, B.G., Baker, J.B., Cooper, D.T., 1977. Nutrient use by three geographic sources of eastern cottonwood. *Can. J. For. Res.* 9, 532–534.
- Bond-Lamberty, B., Wang, C., Gower, S.-T., 2002. Aboveground and belowground biomass and sapwood area allometric equations for six boreal tree species of northern Manitoba. *Can. J. For. Res.* 32, 1441–1450.
- Boring, L.R., Swank, W.T., 1984. The role of black locust (*Robinia pseudoacacia*) in forest succession. *J. Ecol.* 72, 749–766.
- Bray, J.R., Dudkiewicz, L.A., 1963. The composition, biomass and productivity of two *Populus* forests. *Bull. Torrey Bot. Club* 90, 298–308.
- Briggs, R.D., Porter, J.H., White, E.H., 1989. Component Biomass Equations for *Acer rubrum* and *Fagus grandifolia*.
- Carter, M., Smith, E., 1971. Dry Weight and Nutrient Accumulation in Young Stands of Cottonwood (*Populus deltoides* Bartr.).
- Case, B.S., Hall, R.J., 2008. Assessing prediction errors of generalized tree biomass and volume equations for the boreal region of west-central Canada. *Can. J. For. Res.* 38, 878–889.
- Chapin, F.S., Matson, P.A., Mooney, H.A., 2002. Principles of Terrestrial Ecosystem Ecology. Springer Science & Business Media Inc., New York, NY.
- 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-Yrizar, A., Mugasha, W.A., Müller-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.* 3177–3190, <http://dx.doi.org/10.1111/gcb.12629>.
- Chojnacky, D.C., Heath, L.S., Jenkins, J.C., 2013. Updated generalized biomass equations for North American tree species. *Forestry* 87, 129–151, <http://dx.doi.org/10.1093/forestry/cpt053>.
- Clark, A., Phillips, D.R., Frederick, D.J., 1986a. Weight, Volume and Physical Properties of Major Hardwood Species in the Upland-South. USDA Forest Service.
- Clark, A., Phillips, D.R., Frederick, D.J., 1986b. Weight, Volume and Physical Properties of Major Hardwood Species in the Piedmont. USDA Forest Service.
- Clark, A., Phillips, D.R., Frederick, D.J., 1985. Weight, Volume and Physical Properties of Major Hardwood Species in the Gulf and Atlantic Coastal Plains. USDA Forest Service.
- Comeau, P.G., 1976. Net Primary Production of Lodgepole Pine (*Pinus contorta*) in Some Ecosystems in Southern British Columbia. University of Victoria, Canada.
- Congdon, P., 2006. Bayesian Statistical Modeling, 2nd ed. Wiley.
- de-Miguel, S., Mehtatalo, L., Durkaya, A., 2014. Developing generalized, calibratable, mixed effects meta-models for large-scale biomass prediction. *Can. J. For. Res.* 44, 648–656.
- Detto, M., Asner, G.P., Müller-Landau, H.C., Sonnentag, O., 2015. Spatial variability in tropical forest leaf area density from multireturn LiDAR and modelling. *J. Geophys. Res. Biogeosci.* 120, 294–309.
- Devine, W.D., Footen, P.W., Harrison, R.B., Terry, T.A., Harrington, C.A., Holub, S.M., Gould, P.J., 2013. Estimating Tree Biomass, Carbon and Nitrogen in Two Vegetation Control Treatments in an 11-Year-Old Douglas-Fir Plantation on a Highly Productive Site. USDA Forest Service, Portland, Oregon.
- Domke, G.M., Woodall, C.W., Smith, J.E., Westfall, J.A., McRoberts, R.E., 2012. Consequences of alternative tree-level biomass estimation procedures on U.S. forest carbon stock estimates. *For. Ecol. Manag.* 270, 108–116, <http://dx.doi.org/10.1016/j.foreco.2012.01.022>.
- Enquist, B.J., Niklas, K.J., 2001. Invariant scaling relations across tree-dominated communities. *Nature* 410, 655–660.
- Falster, D.S., Duursma, R.A., Ishihara, M.I., Barneche, D.R., FitzJohn, R.G., Vårhammar, A., Aiba, M., Ando, M., Anten, N., Aspinwall, M.J., Baltzer, J.L., Baraloto, C., Battaglia, M., Battles, J.J., Bond-Lamberty, B., van Breugel, M., Camac, J., Claveau, Y., Coll, L., Dannoura, M., Delagrangé, S., Domec, J.-C., Fatemi, F., Feng, W., Gargaglione, V., Goto, Y., Hagihara, A., Hall, J.S., Hamilton, S., Harja, D., Hiura, T., Holdaway, R., Hutley, L.S., Ichie, T., Jokela, E.J., Kantola, A., Kelly, J.W.G., Kenzo, T., King, D., Kloeppel, B.D., Kohyama, T., Komiyama, A., Laclau, J.-P., Lusk, C.H., Maguire, D.A., le Maire, G., Mäkelä, A., Markesteijn, L., Marshall, J., McCulloh, K., Miyata, I., Mokany, K., Mori, S., Myster, R.W., Nagano, M., Naidu, S.L., Nouvellon, Y., O'Grady, A.P., O'Hara, K.L., Ohtsuka, T., Osada, N., Osunkoya, O.O., Peri, P.L., Petritan, A.M., Poorter, L., Portsmouth, A., Potvin, C., Ransijn, J., Reid, D., Ribeiro, S.C., Roberts, S.D., Rodríguez, R., Saldaña-Acosta, A., Santa-Regina, I., Sasa, K., Selaya, N.G., Sillett, S.C., Sterck, F., Takagi, K., Tange, T., Tanouchi, H., Tissue, D., Umehara, T., Utsugi, H., Vadeboncoeur, M.A., Valladares, F., Vanninen, P., Wang, J.R., Wenk, E., Williams, R., de Aquino Ximenes, F., Yamaba, A., Yamada, T., Yamakura, T., Yanai, R.D., York, R.A., 2015. BAAD: a biomass and allometry database for woody plants. *Ecology* 96, 1445, <http://dx.doi.org/10.1890/14-1889.1>.
- Finley, A.O., Banerjee, S., McRoberts, R.E., 2008. A Bayesian approach to quantifying uncertainty in multi-source forest area estimates. *Environ. Ecol. Stat.* 15, 241–258.
- Garbett, W.S., 1977. Aboveground Biomass and Nutrient Content of a Mixed Slash-Longleaf Pine Stand in Florida. University of Florida.
- Gelman, A.E., 2009. Bayes, Jeffreys, prior distributions and the philosophy of statistics. *Stat. Sci.* 24, 176–178.
- Gelman, A.E., 2006. Prior distributions for variance parameters in hierarchical models. *Bayesian Anal.* 71, 369–382.
- Gelman, A.E., 2003. A Bayesian formulation of exploratory data analysis: a goodness-of-fit testing. *Int. Stat. Rev.* 71, 369–382.
- Gelman, A.E., Carlin, J.B., Stern, H.S., Rubin, D.B., 2003. Bayesian Data Analysis, 2nd ed. CRC Press, London.
- Gelman, A.E., Goegebeur, Y., Tuerlinckx, F., Van Mechelen, I., 2000. Diagnostic checks for discrete data regression models using posterior predictive simulations. *J. R. Stat. Soc. Ser. B (Statistical Methodol.)* 49, 247–268.
- Gelman, A.E., Meng, X.-L., Stern, H., 1996. Posterior predictive assessment of model fitness via realized discrepancies. *Stat. Sin.* 6, 733–807.
- Gelman, A.E., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7, 457–511.
- Gholz, H., 1979. Equations for Estimating Biomass and Leaf Area of Plants in the Pacific Northwest. Corvallis, OR.
- Gower, S.-T., Vogel, J.G., 1999. BOREAS TE-06 Allometry Data. Oak Ridge, Tennessee, USA.
- Green, E.J., MacFarlane, D.W., Valentine, H.T., Strawderman, W.E., 1999. Assessing uncertainty in a stand growth model by Bayesian synthesis. *For. Sci.* 45.
- Grier, C.C., Waring, R.H., 1974. Conifer foliage mass related to sapwood area. *For. Sci.* 20, 205–206.
- Gyawali, N., 2008. Aboveground Biomass Partitioning Due to Thinning in Naturally Regenerated Even-Aged Shortleaf Line Stands in Southeast Oklahoma. Oklahoma State University.
- Harrison, R.B., 1981. Biomass, Calcium, Magnesium and Potassium Dynamics of Young Even-Aged White Mountain, New Hampshire Forest Stands Regenerating from Clearcutting. University of New Hampshire.
- Harrison, R.B., Terry, T.A., Licata, C.W., Flaming, B.L., Meade, R., Guerrini, I.A., Strahm, B.D., Xue, D., Lolley, M.R., Sidell, A.R., Wagoner, G.L., Briggs, D., Turnblom, E.C., 2009. Biomass and stand characteristics of a highly productive mixed Douglas-Fir and Western Hemlock plantation in coastal Washington. *West. J. Appl. For.* 24, 180–186.
- Heilman, P.E., 1961. Effects of Nitrogen Fertilization on the Growth and Nitrogen Nutrition of Low-Site Douglas-Fir Stands. University of Washington.
- Hooten, M., Hobbs, N., 2015. A guide to Bayesian model selection for ecologists. *Ecol. Monogr.* 85, 3–28.
- Houghton, R.A., 2003. Why are estimates of the terrestrial carbon balance so different? *Glob. Change Biol.* 9, 500–509, <http://dx.doi.org/10.1046/j.1365-2486.2003.00620.x>.
- Houser, J.N., 1980. Estimation of Aboveground Biomass and Inorganic Nutrient Content of a 25-Year-Old Loblolly Pine (*Pinus taeda* L.) Plantation. Texas A&M University, College Station, Texas.
- Hutnik, R.J., 1964. Accumulation and Production of Dry Matter in Red Pine (*Pinus resinosa* Ait.) Plantations at Various Spacings. Sch. For. Duke University, Durham, North Carolina.
- IPCC, 2006. 2006 IPCC Guidelines for National Greenhouse Gas Inventories: Volume 4 Agriculture, Forestry, and Other Land Use, Prepared by the National Greenhouse Gas Inventories Programme. IGES, Japan.
- Jenkins, J.C., Chojnacky, D.C., Heath, L.S., Birdsey, R.A., 2003. National-scale biomass estimators for United States tree species. *For. Sci.* 49.
- Jokela, E.J., van Gurp, K.P., Briggs, R.D., White, E.H., 1986. Biomass estimation equations for Norway spruce in New York. *Can. J. For. Res.* 16, 413–415.
- Kirschbaum, M., Paul, K., 2002. Modelling C and N dynamics in forest soils with a modified version of the CENTURY model. *Soil Biol. Biochem.* 34, 341–354.
- Kline, J.R., Reed, K.L., Waring, R.H., Stewart, M.L., 1973. Internal Report 123: Direct Measurement of Transpiration and Biomass in Coniferous Trees.
- Krumlik, G.J., Kimmins, J.P., 1973. Studies of biomass distribution and tree form in old virgin forests in the mountains of south coastal British Columbia, Canada. In: Young, H.E. (Ed.), IUFRO Biomass Studies. University of Maine, Nancy, France and Vancouver, BC, pp. 363–374.
- Laubhann, D., Eckmüller, O., Sterba, H., 2010. Applicability of non-destructive substitutes for leaf area in different stands of Norway spruce (*Picea abies* L.

- Karst.) focusing on traditional forest crown measures. *For. Ecol. Manag.* 260, 1498–1506, <http://dx.doi.org/10.1016/j.foreco.2010.07.048>.
- Lefsky, M., Cohen, A., Acker, W.B., Parker, S.A., Spies, G.G., Harding, T.D.A., 1999. Lidar remote sensing of the canopy structure and biophysical properties of Douglas-Fir western Hemlock forests. *Remote Sens. Environ.* 70, 339–361, [http://dx.doi.org/10.1016/S0034-4257\(99\)00052-8](http://dx.doi.org/10.1016/S0034-4257(99)00052-8).
- Lehtonen, A., Cienciala, E., Tatarinov, F., Makipaa, R., 2007. Uncertainty estimation of biomass expansion factors for Norway spruce in the Czech Republic. *Ann. For. Sci.* 64, 133–140.
- Lin, Y., Huth, F., Berger, U., Grimm, V., 2014. The role of belowground competition and plastic biomass allocation in altering plant mass-density relationships. *Oikos* 123, 248–256.
- Loomis, R.M., Phares, R.E., Crosby, J.S., 1966. Estimating foliage and branchwood quantities in shortleaf pine. *For. Sci.* 12, 30–39.
- Mäkelä, A., 1986. Implications of the pipe model theory on dry matter partitioning and height growth in trees I. Introduction, 103–120.
- Marshall, J.D., Waring, R.H., 1986. Comparison of methods of estimated leaf-area index in old-growth Douglas-fir. *Ecology* 67, 975–979.
- Martin, J.G., Kloepfel, B.D., Schaefer, T.L., Kimbler, D.L., McNulty, S.G., 1998. Aboveground biomass and nitrogen allocation of ten deciduous southern Appalachian tree species. *Can. J. For. Res.* 28, 1648–1659.
- Mead, D.J., 1971. Movement of Added Nitrogen and Phosphorus in a Pine Forest Ecosystem. University of Florida.
- Meng, X.-L., 1994. Posterior predictive *p*-values. *Ann. Stat.* 22, 1142–1160.
- Metz, L.J., Wells, C.G., 1965. Weight and Nutrient Content of the Aboveground Parts of Some Loblolly Pines. US Forest Service Research Paper SE-17, Asheville, NC, USA.
- Miller, E.L., Meeuwig, R.O., Budy, J.D., 1981. Biomass of Singleleaf Pinyon and Utah Juniper. Research Paper INT-273.
- Mitsch, W.J., Ewel, K.C., 1979. Comparative biomass and growth of cypress in Florida wetlands. *Am. Midl. Nat.* 101, 417–426.
- Monserud, R.A., Marshall, J.D., 1999. Allometric crown relations in three northern Idaho conifer species. *Can. J. For. Res.* 29, 521–535.
- Moore, J.R., 2010. Allometric equations to predict the total above-ground biomass of radiata pine trees. *Ann. For. Sci.* 67, 806.
- Mroz, G.D., Gale, M.R., Jurgensen, M.F., Frederick, D.F., Clark, A., 1984. Composition, structure, and aboveground biomass of two old-growth northern hardwood stands in Upper Michigan. *Can. J. For. Res.*
- Mueller, C.W., 1976. The accumulation of dry matter in plantations of eastern cottonwood on alluvial sites of the Mississippi River valley. Mississippi State University.
- Neisch, D.A., 1980. Estimating Aboveground Biomass for Shortleaf and Loblolly Pines in east Texas. Stephen F. Austin State University.
- Nelson, A.S., Weiskittel, A.R., Wagner, R.G., Saunders, M.R., 2014. Development and evaluation of aboveground small tree biomass models for naturally regenerated and planted species in eastern Maine, U.S.A. *Biomass Bioenergy* 68, 215–227, <http://dx.doi.org/10.1016/j.biombioe.2014.06.015>.
- Niinemets, U., 2001. Global-scale climatic controls of leaf dry mass per area, density, and thickness in trees and shrubs. *Ecology* 82 (2), 453–469, <http://dx.doi.org/10.2307/2679872>.
- Nowak, D.J., 1996. Notes: estimating leaf area and leaf biomass of open-grown deciduous urban trees. *For. Sci.* 42, 504–507.
- Ntzoufras, I., 2009. *Bayesian Modeling Using WinBUGS*. Wiley.
- Olah, F.D., 1972. Growth Characteristics of Open-Grown Virginia pine (*Pinus virginiana*, Mill.). Virginia Polytechnic Institute and State University, Blacksburg, VA.
- Pearce, H.G., Anderson, W.R., Fogarty, I.G., Todoroki, C.L., Anderson, S.A.J., 2010. Linear mixed-effects models for estimating biomass and fuel loads in shrublands. *Can. J. For. Res.* 4, 2015–2026.
- Picard, N., Mortier, F., Rossi, V., Gourlet-Fleury, S., 2010. Clustering species using a model of population dynamics and aggregation theory. *Ecol. Model.* 221, 152–160.
- Plummer, M., 2003. JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling.
- Plummer, M., Stukalov, A., 2014. Package “rjags”: Bayesian Graphical Models Using MCMC.
- Randolph, K.C., 2013. Development history and bibliography of the US Forest Service crown-condition indicator for forest health monitoring. *Environ. Monit. Assess.* 185, 4977–4993.
- Raymond, C.L., McKenzie, D., 2013. Comparing Algorithms for Estimating Foliar Biomass of Conifers in the Pacific Northwest.
- Reich, P.B., Luo, Y., Bradford, J.B., Poorter, H., Perry, C.H., Oleksyn, J., 2014. Temperature drives global patterns in forest biomass distribution in leaves, stems, and roots. *Proc. Natl. Acad. Sci. U. S. A.* 111, 13721–13726, <http://dx.doi.org/10.1073/pnas.1216053111>.
- Rencz, A.N., Auclair, A.N., 1980. Dimension analysis of various components of black spruce in subarctic lichen woodland. *Can. J. For. Res.* 10, 491–497.
- Repola, J., 2008. Biomass equations for birch in Finland. *Silva Fenn.* 42, 605–624.
- Robichaud, E., Methven, I.R., 1992. The applicability of the pipe model theory for the prediction of foliage biomass in trees from natural, untreated black spruce stands. *Can. J. For. Res.* 22, 1118–1123.
- Roth, B.E., 2010. *Genotype × Environment Interactions in Selected Loblolly and Slash Pine Plantations in the Southeastern United States*. University of Florida.
- Ruark, G.A., Martin, G.L., Bockheim, J.G., 1987. Comparison of constant and variable allometric ratios for estimating *Populus tremuloides* biomass. *For. Sci.* 33, 294–300.
- Sabatia, C.O., 2007. Effect of Thinning on Partitioning of Aboveground Biomass in Naturally Regenerated Shortleaf Pine. Oklahoma State University.
- Santee, W.R., 1970. A Dimensional Analysis of Eastern Hemlock (*Tsuga canadensis*). University of Georgia.
- Schroeder, P., Brown, S., Jiangming, M., Birdsey, R., Cieszewski, C., 1997. Biomass estimation for temperate broadleaf forests of the United States using inventory data. *For. Sci.* 43, 424–434.
- Shaiek, O., Loustau, D., Trichet, P., Meredieu, C., Bachtobji, B., Garchi, S., Hedi, E.L., Aouni, M., 2011. Generalized biomass equations for the main aboveground biomass components of maritime pine across contrasting environments. *Ann. For. Sci.* 68, 443–452.
- Shannon, C.A., 1976. Uptake and Distribution of Biomass and Nutrient Elements in Balsam Fir. University of Minnesota.
- Shinozaki, K., Yoda, K., Hozumi, K., Kira, T., 1964. A quantitative analysis of plant form. The pipe model theory. II. Further evidence of the theory and its application in forest ecology. *Jpn. J. Ecol.* 14, 133–139.
- Singer, F.P., 1964. Accumulation of Organic Matter in Red Pine and Norway Spruce Plantations of Various Spacings. Pennsylvania State University.
- Smith, N.J., 1977. Estimates of Above-Ground Biomass, Net Primary Production and Energy Flows in 8 to 10 Year Old Red Alder (*Alnus rubra* Bong.) Ecosystems. University of British Columbia.
- Smithwick, E.A.H., Harmon, M.E., Remillard, S.M., Acker, S.A., Franklin, J.F., 2002. Potential upper bounds of carbon stores in forests of the Pacific Northwest. *Ecol. Appl.* 12, 1303–1317.
- Sollins, P., Anderson, R.M., 1971. Dry-Weight and Other Data for Trees and Woody Shrubs of the Southeastern United States. Oak Ridge, Tennessee, USA.
- Song, C., 2013. Optical remote sensing of forest leaf area index and biomass. *Prog. Phys. Geogr.* 37, 98–113.
- Temesgen, H., Vicente, M., Weiskittel, A., Wilson, D., 2011. Sampling strategies for efficient estimation of tree foliage biomass. *For. Sci.* 57, 153–163.
- Tobin, B., Black, K., Osborne, B., Reidy, B., Bolger, T., Nieuwenhuis, M., 2006. Assessment of allometric algorithms for estimating leaf biomass, leaf area index and litter fall in different-aged Sitka spruce forests. *Forestry* 79, 453–465.
- Vaidya, M.S.L., 1961. Dry Matter Production and Nutrient Accumulation in Plantations of Shortleaf Pine. Duke University.
- Valentine, H.T., Forest, N., Station, E., Service, U.F., 1985. *Tree-Growth Models: Derivations Employing the Pipe-model Theory*, pp. 579–585.
- Van Pelt, R., Sillett, S.C., 2008. Crown development of coastal *Pseudotsuga menziesii*, including a conceptual model for tall conifers. *Ecol. Monogr.* 78, 283–311.
- Vehtari, A., Ojanen, J., 2012. A survey of Bayesian predictive methods for model assessment, selection and comparison. *Stat. Surv.* 6, 142–228.
- Waring, R.H., Schroeder, P.E., Oren, O., 1982. Application of the pipe model theory to predict canopy leaf area. *Can. J. For. Res.* 12, 556–560.
- Wayson, C.A., Johnson, K.D., Cole, J.A., Olguin, M.I., Carrillo, O.I., Birdsey, R.A., 2014. Estimating uncertainty of allometric biomass equations with incomplete fit error information using a pseudo-data approach: methods. *Ann. For. Sci.*, <http://dx.doi.org/10.1007/s13595-014-0436-7>.
- Weiskittel, A.R., Macfarlane, D.W., Radtke, P.J., Affleck, D.L.R., Temesgen, H., Westfall, J.A., Woodall, C.W., Coulston, J.W., 2015. A call to improve methods for estimating tree biomass for regional and national assessments. *J. For.* (in press).
- Wirth, C., Schumacher, J., Schulze, E.-D., 2004. Generic biomass functions for Norway spruce in Central Europe—a meta-analysis approach toward prediction and uncertainty estimation. *Tree Physiol.* 24, 121–139, <http://dx.doi.org/10.1093/treephys/24.2.121>.
- Woodall, C.W., Heath, L.S., Domke, G.M., Nichols, M.C., 2011. Methods and Equations for Estimating Aboveground Volume, Biomass, and Carbon for Trees in the U.S. Forest Inventory, 2010.
- Wutzler, T., Wirth, C., Schumacher, J., 2008. Generic biomass functions for Common beech (*Fagus sylvatica*) in Central Europe: predictions and components of uncertainty. *Can. J. For. Res.* 38, 1661–1675, <http://dx.doi.org/10.1139/X07-194>.
- Yarie, J., Kane, E., Mack, M., 2007. Aboveground Biomass Equations for the Trees of Interior Alaska. Fairbanks, AK, USA.
- Zapatas-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.
- Zellers, C.E., Saunders, M.R., Morrissey, R.C., Shields, J.M., Bailey, B.G., Dyer, J., Cook, J., 2012. Development of allometric leaf area models for intensively managed black walnut (*Juglans nigra* L.). *Ann. For. Sci.* 69, 907–913, <http://dx.doi.org/10.1007/s13595-012-0215-2>.
- Zianis, D., Muukkonen, P., Makipaa, R., Mencuccini, M., 2005. Biomass and stem volume equations for tree species in Europe. *Silva Fenn. Monogr.* 4, 63.