

biometrics

Additive Nonlinear Biomass Equations: A Likelihood-Based Approach

David L.R. Affleck and Ulises Diéguez-Aranda

Since Parresol's (*Can. J. For. Res.* 31:865–878, 2001) seminal article on the topic, it has become standard to develop nonlinear tree biomass equations to ensure compatibility among total and component predictions and to fit these equations using multistep generalized least-squares methods. In particular, many studies have specified equations for total tree biomass by aggregating the expectations of M component biomass models and have fit all $M + 1$ equations jointly using seemingly unrelated regression. More recently, an alternative approach has been used wherein compatibility is ensured by deriving multiplicative component expectations, with estimation carried out using instrumental variables and generalized least squares. Yet neither of these strategies considers the fundamental additivity of biomass data themselves nor the implied stochastic constraints necessary for maximum likelihood (ML) estimation. For model selection based on information criteria, stochastic simulation, Bayesian inference, or estimation with missing data, it is important to base estimation and inference on valid probabilistic models. Here, we show how aggregative and disaggregative nonlinear equations can be specified within a probabilistic framework and fit using Gaussian ML with open-source software. We use Parresol's slash pine (*Pinus elliotii* Engelm. var. *elliotii*) data to contrast model forms and predictions. We also show how the ML approach can accommodate unobserved or aggregated component biomass data and can thus be useful for integrating felled-tree data collected under different protocols.

Keywords: multivariate regression, missing data, disaggregation models, open-source statistical software

National forest inventory programs now commonly report estimates of aboveground biomass and carbon in live trees. The estimates are often derived from individual tree biomass equations developed to also furnish estimates of biomass in foliage, branches, stems, and other tree components. These component estimates may be relevant to ecophysiological growth models, timber planning efforts, bioenergy programs, or wildland fuel assessments. For many of these purposes, compatibility among the component and total tree biomass models is important. As noted by Parresol (2001), this compatibility should, at a minimum, ensure that component biomass or carbon estimates do not exceed estimates of whole-tree biomass or carbon and that component estimates sum to the estimates of the totals. Yet for more advanced analytical and estimation procedures such as stochastic simulation, mixed-effects modeling, and Bayesian inference, this level of compatibility is insufficient. These techniques require that the additive nature of tree biomass measurements themselves be recognized and that appropriate constraints be placed on stochastic model elements. In particular, they require that tree biomass models be formulated so as to jointly constitute a valid probability model.

Individual-tree equations commonly relate total and component biomass to diameter at breast height (dbh), either alone or in combination with other tree attributes. Dbh is readily measured and often accounts for the largest share of variation in tree biomass, but attributes such as tree height or live crown length can account for additional variation associated with differences in height:diameter ratios, stand density, or other factors (Brown 1978, Evert 1985). These and other predictors have been combined in myriad linear and nonlinear parametric equations to describe differences in biomass allometries across trees. The multivariate power law (i.e., $y = b_0 x_1^{b_1} x_2^{b_2} \dots x_p^{b_p}$) is often a candidate equation. This equation in particular may motivate linear regression after logarithmic transformation or the use of nonlinear regression on the original mass scale. As the calibration of multicomponent equations is intrinsically more complex in the latter case, this research focuses on the nonlinear equation setting.

Parresol (2001) developed two methods for ensuring the additivity of nonlinear component biomass equations. Both can be called aggregative in that, at least as developed in his examples, they ensured component additivity by constraining the mean function

Manuscript received September 18, 2015; accepted December 18, 2015; published online February 18, 2016.

Affiliations: David L.R. Affleck (david.affleck@umontana.edu), Department of Forest Management, University of Montana, Missoula, MT. Ulises Diéguez-Aranda (ulises.dieguez@usc.es), Sustainable Forest Management Unit, University of Santiago de Compostela.

Acknowledgments: This research was undertaken while the first author was a visiting scholar with the Sustainable Forest Management Unit of the Universidade de Santiago de Compostela. The research was supported by funding from the USDA Forest Service Forest Inventory and Analysis program, the University of Montana, and the Inland Northwest Growth and Yield Cooperative. We also thank Dr. Philip J. Radtke and two anonymous referees for insightful comments on earlier drafts.

for total biomass to be the sum of the mean functions of the biomass components. His two methods differed in the stochastic conditions attached to the models, although they are more commonly differentiated by the parameter estimation techniques employed. The first approach prescribed separate estimation of each biomass component equation, with subsequent recovery of cross-component correlations from the original measurements for estimation of precision. The second used seemingly unrelated regression methods to fit the total and component models jointly. With this latter approach, the potential cross-correlations among biomass measurements on a given tree are modeled directly. In principle, this imparts greater efficiency to parameter estimators and has helped render this method the de facto standard for nonlinear biomass equation estimation (e.g., Sanquetta et al. 2015, Zhao et al. 2015).

The inverse of the aggregative approach is to focus first on the specification of an allometric equation for total biomass and then constrain the form or parameters of component equations to ensure additivity. The latter constraints could be imposed by fitting equations characterizing the observed proportions of total biomass in each component. This could be achieved, for example, using Dirichlet regression models (Gueorguieva et al. 2008), although specialized forms (perhaps incorporating random effects) might be needed to allow for dependence among total and component biomass observations from the same tree. More commonly, nonlinear models for component biomass proportions have been fit separately, leaving one component out and estimating it subsequently via subtraction (e.g., Jenkins et al. 2003). Yet this approach does not allow for dependence among observations from the same tree and cannot ensure that subtotals of the estimated component proportions are strictly less than 1.

An alternative disaggregation strategy, apparently developed initially in the Chinese literature (Tang et al. 2000, see also Dong et al. 2015), is likewise based on the development of component biomass fraction equations but with parameter estimation carried out on the original scale (i.e., kilograms of mass). This approach can guarantee additivity of regression-based estimates of biomass components and totals. Disaggregative models of this form have been completed by specifying a stochastic structure (generally one allowing for cross-component correlations) premised on an errors-in-variables framework (see, e.g., Tang et al. 2001, Tang and Wang 2002). Like Parresol's (2001) approaches, this framework recommends a multi-step least-squares (MSLS) procedure to obtain parameter estimates.

Both aggregative and disaggregative models have been developed with due attention to the inconstant nature of component biomass variance structures and to the potential for cross-correlation between measurements taken on the same trees. However, these methods typically are fit with observations of both component and total biomass and have not explicitly recognized the fundamental additivity of these data. Biomass measurement requires destructive sampling and, except occasionally for small trees, is generally carried out component by component. That is, a tree is partitioned into components, a subsampling strategy is applied to each component, and a total biomass is obtained only afterwards by addition. This additivity of the data themselves has implications for model specification as well as the validity of probability or likelihood-based estimators and selection criteria. Estimation issues stemming from this additivity have been noted (e.g., by Bi et al. 2004) but not addressed in terms of the implied constraints on model forms. Of course, it is not necessary to develop component biomass models that are strictly valid in terms of their probabilistic structure, but such models are

useful for many stochastically oriented inference and estimation techniques. In addition, for regression models that are at least implicitly based on normal or Gaussian error structures, it is worthwhile appreciating when and why the estimation techniques fail to coincide with Gaussian maximum likelihood (ML).

A probabilistic approach to biomass equation development can also be useful when biomass data are incomplete. That is, ML estimation can be applied to make efficient use of data when measurements of certain components are missing or available only in aggregate for some trees. More generally, differences in field data collection protocols or differing definitions of biomass components across studies often present challenges in combining data sets to calibrate tree component biomass equations. For example, in a recent study of loblolly pine (*Pinus taeda* L.), Zhao et al. (2015) identified five aboveground biomass components of interest, namely stemwood, stembark, dead branches, live branches, and foliage. In another loblolly pine study, Naidu et al. (1998) identified seven biomass components including stem sapwood, stem heartwood, stembark, live branchwood, live branchbark, foliage, and roots. Each of these data sets is missing at least one component that the other includes, although they have several components in common. More interestingly, two of the components (stemwood and live branches) measured by Zhao et al. incorporate two more finely resolved subcomponents (stem sapwood and stem heartwood; live branchwood and live branchbark) available in the data of Naidu et al. By their nature, these undifferentiated components of Zhao et al. ought to contain some information about the subcomponents of Naidu et al., albeit at an aggregated level. Regardless of whether the data from these two studies would be suitable for combining to fit biomass equations, the incomplete pairing and differing resolutions of their component measurements demonstrate a challenge for analysis to be addressed here.

The primary objective of this research was to synthesize aggregative and disaggregative approaches to nonlinear biomass equation specification within a probabilistic modeling framework. A second objective was to extend the models and ML estimation algorithms to accommodate missing or aggregated observations of biomass components. Theory and analytic results relevant to these two objectives are developed in the subsequent section, both for the general case and for the special case of Gaussian ML. These are followed by empirical results obtained by estimating biomass model parameters under Gaussian ML using the slash pine (*Pinus elliotii* Engelm. var. *elliotii*) biomass data studied by Parresol (2001). The computational methodology developed to garner these results itself constitutes the product of the third objective of this research, which was to document how Gaussian ML estimation of nonlinear biomass equations can be implemented using open-source statistical software.

Biomass Models and Estimators

Notation and Additivity

Biomass studies differ in the components of biomass measured and totals sought. Some studies target total tree biomass, others only aboveground biomass. Within the aboveground portion, most studies differentiate between crown (i.e., branches and foliage) and stem, but there are variations in the precise definitions of these components and their further subdivisions. Thus, for generality, let $Y_1, Y_2, \dots, Y_M$ denote the M biomass components of a given tree and Y_t the total of interest. A fundamental identity is

$$Y_t = \sum_{m=1}^M Y_m \quad (1)$$

This identity holds by definition when the symbols represent unobserved (or true) tree biomass quantities and is often a desideratum of biomass model estimates or simulations. Yet it is important to recognize that this identity also generally holds when the symbols represent biomass measurements or data obtained from destructive sampling. This follows from the fact that total tree biomass typically goes unmeasured and is obtained instead by summing together the component biomass observations (or estimates) assembled from field and laboratory measurements.

From the above identity stems the fact that given a set of tree-level predictors $\mathbf{x}$ the joint probability law for a set of random variables $Y_1, Y_2, \dots, Y_M$, and Y_t denoting tree biomass components and their total is

$$\begin{aligned} & f_{Y_1, Y_2, \dots, Y_M, Y_t}(y_1, y_2, \dots, y_M, y_t | \mathbf{x}) \\ &= f_{Y_1, Y_2, \dots, Y_M}(y_1, y_2, \dots, y_M | \mathbf{x}) \times f_{Y_t | Y_1, Y_2, \dots, Y_M}(y_t | y_1, y_2, \dots, y_M, \mathbf{x}) \\ &= f_{Y_1, Y_2, \dots, Y_M}(y_1, y_2, \dots, y_M | \mathbf{x}) \times \begin{cases} 1 & \text{if } y_t = \sum_m y_m \\ 0 & \text{else.} \end{cases} \end{aligned}$$

That is, given $\mathbf{x}$, the joint probability model for $Y_1, Y_2, \dots, Y_M$, and Y_t is a simple multiplicative function of the model for $Y_1, Y_2, \dots, Y_M$ with the multiplier being 0 or 1 independently of $\mathbf{x}$ or any model parameters. As such, the ML estimators of the parameters governing the former can be obtained directly by maximizing only the latter. At work here is the fact that given the biomass components and their additivity, there is no additional information contained in the biomass totals.

In practice, multivariate Gaussian probability models are commonly adopted for component biomass data. Where the Gaussian assumption has not been made explicit, it has often been invoked implicitly in inferential procedures or in the calculation of information criteria such as the Akaike information criterion (AIC) (Akaike 1973). These component biomass models are often of the additive-error form

$$Y_m = g_m(\mathbf{x}_m, \beta_m) + \varepsilon_m \quad (2)$$

Here $g_m(\cdot)$ is an arbitrary and possibly nonlinear function of a set of explanatory variables ($\mathbf{x}_m$) relevant to the m th component and a q_m vector of unknown parameters β_m . Conditional on the explanatory variables, the error term ε_m is a mean-zero Gaussian disturbance with variance

$$\sigma_m^2 = v_m(\mathbf{z}_m; \theta_m) \quad (3)$$

where $v_m(\cdot)$ is a function of a vector of explanatory variables $\mathbf{z}_m$ (not necessarily identical to $\mathbf{x}_m$) and a vector of variance parameters θ_m .

Whereas observations from different trees are generally taken to be independent, cross-correlations among the error terms of biomass components on a single tree are often modeled using a set of $M(M-1)/2$ pairwise-correlation parameters. With the variance functions of Equation 3, the latter complete the definition of the variance-covariance matrix of component errors

$$V = \begin{bmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \rho_{1,2} & \sigma_1 \sigma_3 \rho_{1,3} & \dots & \sigma_1 \sigma_M \rho_{1,M} \\ \sigma_1 \sigma_2 \rho_{1,2} & \sigma_2^2 & \sigma_2 \sigma_3 \rho_{2,3} & \dots & \sigma_2 \sigma_M \rho_{2,M} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_1 \sigma_M \rho_{1,M} & \sigma_2 \sigma_M \rho_{2,M} & \sigma_3 \sigma_M \rho_{3,M} & \dots & \sigma_M^2 \end{bmatrix} \quad (4)$$

where $\rho_{m,m'}$ is the correlation between biomass components m and m' across all trees. Below, the cross-covariances in V will also be denoted $\sigma_{mm'} = \sigma_m \sigma_{m'} \rho_{m,m'}$ for simplicity.

There are two properties of this general model that complicate specification of a compatible submodel for Y_t . First, given the identity presented in Equation 1, the variance of the total biomass is not free to take any form. Instead, it is a function of the component scedastic functions and cross-correlation parameters. For example, suppose the variances of the M components are taken to scale with dbh according to component-specific parameters: i.e., $v(\text{dbh}; \theta_m) = \theta_{m1}^2 \text{dbh}^{2\theta_{m2}}$. Then the variance of $Y_t = \sum_m Y_m$ is

$$\begin{aligned} \sigma_t^2 &= \sum_m \text{var}(Y_m) + \sum_m \sum_{m'} \text{cov}(Y_m, Y_{m'}) \\ &= \sum_m \theta_{m1}^2 \text{dbh}^{2\theta_{m2}} + \sum_m \sum_{m'} \rho_{m,m'} \theta_{m1} \theta_{m'1} \text{dbh}^{\theta_{m2} + \theta_{m'2}} \end{aligned}$$

Second, if the component scedastic functions of Equation 3 have free parameters, then the cross-correlation of any given component Y_m and the total Y_t will vary across the ranges of the heteroscedasticity variables $\mathbf{z}_m$. That is, the cross-covariance between the total and, say, the first component is

$$\text{cov}[Y_t, Y_1] = E\left[\left(\sum_m \varepsilon_m\right) \varepsilon_1\right] = \theta_{11}^2 \text{dbh}^{2\theta_{12}} + \sum_{m=2}^M \rho_{m,1} \theta_{11} \theta_{m1} \text{dbh}^{\theta_{m2} + \theta_{12}}$$

This cross-covariance is a function of dbh and, given the structure of σ_t^2 , the corresponding cross-correlation cannot be reduced to a constant parameter $\rho_{t,1}$ without complex constraints on the component cross-correlations and variance parameters. In terms of allometries, the implication is that as tree dbh increases, the variance of one component (e.g., the crown) could increase faster than that of another (stemwood). Then the cross-correlation of the stemwood component with the total biomass will decrease with dbh because the variability of the total is increasing more rapidly than the variability in the stemwood component.

To our knowledge, these variance-covariance relationships have not been incorporated into biomass models that directly specify submodels for both components and totals. Yet they reveal that incorporating a submodel for total biomass by simply specifying an additional mean function and expanding V by one row and one column (incorporating one additional variance function and M additional cross-correlation parameters) will not result in a valid probability model. Instead, the implied constraints must be explicitly identified or the submodel for total biomass must be stricken. Otherwise, singularities will emerge in likelihood-based simultaneous estimation algorithms (as noted by Bi et al. 2004) and least-squares estimators will diverge from Gaussian ML estimators. From a practical standpoint, probabilistic foundations for the use of likelihood-based criteria such as the AIC will be lacking, and stochastic tree biomass simulations derived from the model will not be additive.

Finally, we note that multiplicative-error models analogous to Equation 2 have also been used in component biomass modeling. In some cases, these are transformed (e.g., to a logarithmic scale), and their stochastic specifications are completed within an additive-error framework. Alternatively, a multiplicative-error model can be reexpressed as

$$g_m(\mathbf{x}_m, \beta_m) \varepsilon_m^* = g_m(\mathbf{x}_m, \beta_m) + g_m(\mathbf{x}_m, \beta_m) (\varepsilon_m^* - 1)$$

If stochastic properties are attached directly to $\varepsilon_m = g_m(\mathbf{x}_m, \beta_m) (\varepsilon_m^* - 1)$ rather than to ε_m^* , then these models can be

subsumed into the Gaussian additive-error framework, and the variance-covariance constraints identified above still apply.

Aggregative and Disaggregative Biomass Equations

In some instances, it may be useful to specify an equation for the expected total biomass independently of any component biomass equations. Doing so would allow for maximum fidelity in characterizing totals and presumably for improved accuracy in predicting those totals. Yet in many cases, it is important to ensure that component predictions add to predictions of totals. One way to achieve this is to specify an aggregative equation for the total

$$g_t(\mathbf{x}, \beta) = \sum_m g_m(\mathbf{x}_m, \beta_m) \quad (5)$$

where $\mathbf{x} = \cup_m \mathbf{x}_m$ and $\beta = (\beta_1, \beta_2, \dots, \beta_M)^T$ collect the component predictors and parameters. Naturally, such an approach could be used to specify equations for subtotals as well (e.g., stem biomass, inclusive of both bark and wood). If the component models are linear, then Equation 5 will reduce to a congruent linear form, but in the nonlinear case, there is generally no way to simplify the expression. This aggregative approach is at the heart of the two methods described by Parresol (2001), which differ primarily in whether or not observed Y_i are used with component data to estimate the elements of β and other model parameters simultaneously. From a likelihood perspective, the system of equations that must be fit in this context is

$$Y_1 = g_1(\mathbf{x}_1, \beta_1) + \varepsilon_1 \quad (6a)$$

$$Y_2 = g_2(\mathbf{x}_2, \beta_2) + \varepsilon_2 \quad (6b)$$

⋮

$$Y_M = g_M(\mathbf{x}_M, \beta_M) + \varepsilon_M \quad (6c)$$

with the error terms conforming to the variance structure of Equation 4 when a Gaussian model is adopted. Owing to the identity defined by Equation 1, the observed totals do not contribute to the likelihood, but total biomass at given levels of $\mathbf{x}$ can subsequently be estimated as

$$\hat{Y}_t = g_t(\mathbf{x}, \hat{\beta}) = \sum_m g_m(\mathbf{x}_m, \hat{\beta}_m) = \sum_m \hat{Y}_m \quad (7)$$

ensuring additivity.

Pursuing an alternative approach, Tang et al. (2000) noted that it is also possible to begin with the specification of an equation for total biomass and then disaggregate this using multiplicative component-discrimination functions. For example, in a case where only stemwood (Y_w), stembark (Y_b), and crown components (Y_c) are differentiated, one might begin by constructing an equation for the expected aboveground biomass

$$E[Y_t|\mathbf{x}] = g_t(\mathbf{x}, \beta_t) \quad (8)$$

Here no constraints on the form or parameters of this function are invoked, at least not for the purposes of achieving additivity. Next, a function $g_{cs}(\mathbf{x}_{cs}, \beta_{cs})$ is specified to discriminate between the crown and stem ($Y_s = Y_w + Y_b$) fractions of the biomass such that

$$E[Y_c|\mathbf{x}_t] = g_{cs}(\mathbf{x}_{cs}, \beta_{cs}) g_t(\mathbf{x}, \beta_t) \quad (9a)$$

$$E[Y_s|\mathbf{x}_t] = [1 - g_{cs}(\mathbf{x}_{cs}, \beta_{cs})] g_t(\mathbf{x}, \beta_t) \quad (9b)$$

where $g_{cs}(\mathbf{x}_{cs}, \beta_{cs})$ maps strictly to the unit interval. For instance, from working component biomass equations [say, $h_c(\mathbf{x}_c, \beta_c)$ for Y_c and $h_s(\mathbf{x}_s, \beta_s)$ for Y_s], a discrimination function confined to [0,1] can be obtained as

$$g_{cs}(\mathbf{x}_{cs}, \beta_{cs}) = \frac{h_c(\mathbf{x}_c, \beta_c)}{h_c(\mathbf{x}_c, \beta_c) + h_s(\mathbf{x}_s, \beta_s)} \quad (10)$$

In turn, a function $g_{wb}(\mathbf{x}_{wb}, \beta_{wb})$ is then developed to discriminate between the stemwood and stembark fractions of the stem biomass

$$E[Y_w|\mathbf{x}_t] = g_{wb}(\mathbf{x}_{wb}, \beta_{wb}) [1 - g_{cs}(\mathbf{x}_{cs}, \beta_{cs})] g_t(\mathbf{x}, \beta_t)$$

$$E[Y_b|\mathbf{x}_t] = [1 - g_{wb}(\mathbf{x}_{wb}, \beta_{wb})] [1 - g_{cs}(\mathbf{x}_{cs}, \beta_{cs})] g_t(\mathbf{x}, \beta_t)$$

In this way, component biomass expectations become complex nonlinear functions of the available predictors, whereas totals (and subtotals) simplify. With the additivity identity of Equation 1 in mind, a valid Gaussian biomass model based on this system of disaggregative equations can be written in the additive-error form

$$Y_c = g_{cs}(\mathbf{x}_{cs}, \beta_{cs}) g_t(\mathbf{x}, \beta_t) + \varepsilon_c \quad (11a)$$

$$Y_w = g_{wb}(\mathbf{x}_{wb}, \beta_{wb}) [1 - g_{cs}(\mathbf{x}_{cs}, \beta_{cs})] g_t(\mathbf{x}, \beta_t) + \varepsilon_w \quad (11b)$$

$$Y_b = [1 - g_{wb}(\mathbf{x}_{wb}, \beta_{wb})] [1 - g_{cs}(\mathbf{x}_{cs}, \beta_{cs})] g_t(\mathbf{x}, \beta_t) + \varepsilon_b \quad (11c)$$

with the error terms (ε) having the general variance-covariance structure defined by Equation 4. Again, incorporation of a sub-model for observed biomass totals is not necessary.

As developed above and in Dong et al. (2015), this disaggregative biomass equation strategy dichotomizes biomass totals or subtotals into finer components. Yet the same approach could be used to split a (sub)total into 3 or more components (Tang et al. 2000). Regardless, attention must be paid to the parameterization of the discrimination functions if it is important that specific parameters be identifiable. For instance, if the crown and stem component equations on the right hand side of Equation 10 are arbitrary nonlinear functions, then it may be impossible to obtain unique estimates for each of β_c , β_s , and β_t in a simultaneous calibration of Equation set 11. On the other hand, if congruent forms are adopted for those component equations, e.g.,

$$h_c(\mathbf{x}_c, \beta_c) = h_c(\mathbf{x}, \beta_c) = \prod_k x_k^{\beta_{c,k}}$$

$$h_s(\mathbf{x}_s, \beta_s) = h_s(\mathbf{x}, \beta_s) = \prod_k x_k^{\beta_{s,k}}$$

then the discrimination function will factor

$$\begin{aligned} g_{cs}(\mathbf{x}_{cs}, \beta_{cs}) &= \frac{h_c(\mathbf{x}_c, \beta_c)}{h_c(\mathbf{x}_c, \beta_c) + h_s(\mathbf{x}_s, \beta_s)} = \frac{\prod_k x_k^{\beta_{c,k}}}{\prod_k x_k^{\beta_{c,k}} + \prod_k x_k^{\beta_{s,k}}} \\ &= \frac{1}{1 + \prod_k x_k^{\beta_{cs,k}}} \end{aligned}$$

and require a smaller number of parameters ($\beta_{cs,k} = \beta_{s,k} - \beta_{c,k}$) to be estimated. This strategy, applied by Dong et al. (2015), reduces the complexity of parameter estimation and allows for (additive) estimation of all components and totals. However, it does not permit recovery of unique values for β_c and β_s and does constrain the complexity of component equations.

Related to this issue of identifiability, Tang et al. (2001) developed an alternative stochastic formulation to replace the additive-error structure exemplified by Equation 11. Specifically, they developed an errors-in-variables structure, and Dong et al. (2015) refer to this in the context of their disaggregative biomass equation formulation. However, those studies do not detail how stochastic constraints on the error term for the biomass total (e.g., on its variance or its cross-covariances with components) can be explicitly recognized in estimation. As such, we treat below only the additive-error model structure represented by Equation 11.

In contrasting aggregative and disaggregative model structures, it is worth noting how component biomass observations contribute to the mean structure parameters in β . In the aggregative model defined by Equation 6, each component Y_m contributes directly to the estimation of its own β_m parameter vector and also indirectly (because of cross-correlations among components) to the other vectors $\beta_{m'}$. In contrast, in the disaggregative framework of Equation 11, the same kinds of indirect contributions occur, but each component Y_m contributes information directly to the estimation of β_i in Equation 8 as well as to the parameters characterizing one or more discrimination functions. This in turn has consequences for the impact of missing biomass component data, as discussed below.

Maximum Likelihood Estimation

It was noted above that under a likelihood-based approach to estimation only the likelihood of the observed component data must be maximized. That is, regarding the sample trees as independent, all parameters are estimated by maximizing

$$L(\beta, \theta, \rho | \mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_M, \mathbf{X}) \\ = \prod_{i=1}^n L_i(\beta, \theta, \rho | Y_{i1}, Y_{i2}, \dots, Y_{iM}, \mathbf{x}_{im}) \quad (12)$$

where $\theta = (\theta_1, \theta_2, \dots, \theta_M)^T$ and $\rho = (\rho_{1,2}, \rho_{1,3}, \dots, \rho_{M-1,M})^T$ contain the variance and cross-correlation parameters of the component data, $\mathbf{Y}_m = (Y_{1m}, Y_{2m}, \dots, Y_{nm})^T$ comprises the observed data on component m for all n trees, $\mathbf{X}$ collects the covariate values $\mathbf{x}_{im}$ from each tree and component, and $L_i(\cdot)$ denotes the likelihood function for the i th tree. If there is tree-to-tree dependence in one or more components, then the complete sample likelihood of Equation 12 will not factor into n tree-specific likelihoods but will nonetheless remain a function only of the component observations. Under standard conditions, the ML parameter estimators are asymptotically Gaussian distributed with variance-covariance matrix given by the inverse of the Fisher information matrix (Lehmann and Casella 1998, §6.3).

For multivariate Gaussian models, a treatment of ML estimation covering both theory and computational methods is given by Pinheiro and Bates (2000, §7.5). Most notably, the asymptotic variance-covariance matrix of the estimator of the q parameters β in a Gaussian component biomass model such as Equation 6 or 11 can be written

$$\text{var}(\hat{\beta}) = \Omega = \left[\sum_{i=1}^n F_i^T V_i^{-1} F_i \right]^{-1}$$

where V_i is the variance-covariance matrix of component errors (cf. Equation 4) for the i th tree and

$$F_i = \frac{\delta}{\delta \beta} \mathbf{g}(\mathbf{x}_i, \beta) = \frac{\delta}{\delta \beta} \begin{bmatrix} g_1(\mathbf{x}_{i1}, \beta_1) \\ g_2(\mathbf{x}_{i2}, \beta_2) \\ \vdots \\ g_M(\mathbf{x}_{iM}, \beta_M) \end{bmatrix}$$

is the $M \times q$ matrix of the tree's component mean functions differentiated with respect to their parameters (and evaluated at the true value of β). This variance matrix can itself be estimated by plugging in the ML estimates of β to allow for testing or interval estimates on the parameters themselves.

For biomass estimation purposes, point estimates of components and totals can be obtained using the same framework as Equation 7 and the ML estimates of β . Denote an estimate of the expected component, subtotal, or total biomass at a given level of $\mathbf{x}$ by $\Delta^T \mathbf{g}(\mathbf{x}, \hat{\beta})$, where Δ is a known $M \times 1$ vector (e.g., $\Delta = \mathbf{1}$ yields an estimate of total biomass). For a Gaussian biomass model, the variance of the estimate $\Delta^T \mathbf{g}(\mathbf{x}, \hat{\beta})$ can be estimated using the delta method

$$\widehat{\text{var}}[\Delta^T \mathbf{g}(\mathbf{x}, \hat{\beta})] = \Delta^T \hat{F} \hat{\Omega} \hat{F}^T \Delta \quad (13)$$

where $\hat{\Omega}$ is obtained by evaluating Ω at the ML parameter estimates and

$$\hat{F} = \left. \frac{\delta}{\delta \beta} \mathbf{g}(\mathbf{x}, \beta) \right|_{\hat{\beta}}$$

For a prediction at the individual-tree level, say $\hat{Y}_{\text{new}} = \Delta^T \mathbf{g}(\mathbf{x}, \hat{\beta})$, variance estimation would require that Equation 13 be supplemented by estimates of the intrinsic tree-to-tree variation in component, subtotal, or total biomass. Specifically, for prediction for a tree selected independently of those used to fit the biomass equations,

$$\widehat{\text{var}}[\hat{Y}_{\text{new}}] = \Delta^T \hat{F} \hat{\Omega} \hat{F}^T \Delta + \Delta^T \hat{V} \Delta$$

where $\hat{V}$ corresponds to an evaluation of Equation 4, given the ML parameter estimates and the chosen levels of the predictors $\mathbf{x}$.

Missing Biomass Components

Given both the complexity and cost of obtaining tree biomass data, it is important that parameter estimation techniques are able to accommodate settings where, for some trees, some of the biomass components are not available. Little and Rubin (1987, ch. 1) discuss different mechanisms that may yield incomplete data, as well as the implications for data analysis and inference. It is assumed in the subsequent discussion that cases of incomplete data are instances of data missing at random or missing completely at random. That is, it is assumed that the missingness of any biomass components is the result of a mechanism that is independent of the values of those (or any related) components and dependent only on variables and parameters with no bearing on the relationship between tree biomass and its predictors. These are cases where missingness is uninformative, and the processes leading to incomplete data can be ignored (see Little and Rubin 1987, §5.3). To such cases ML methods are readily extended, particularly in the Gaussian setting.

With multidimensional component biomass measurements, there are two patterns of missingness that should be considered. The first and simplest is where no information for one or more components is available. For instance, if destructive sampling is carried out in leaf-off season or frozen ground prevents extraction of roots, then

no information will be available for foliage or root biomass. Furthermore, the total biomass of these trees will remain unobserved. For such trees, the observed data likelihood is obtained by integrating the complete data likelihood over the missing components. More specifically, if the first 2 biomass components are unobserved on a particular tree, then the corresponding observed data likelihood for the tree is proportional to

$$f_{Y_3, Y_4, \dots, Y_M}(y_3, y_4, \dots, y_M) \\ = \int_0^\infty \int_0^\infty f_{Y_1, Y_2, \dots, Y_M}(y_1, y_2, \dots, y_M) dy_1 dy_2 \quad (14)$$

In the Gaussian setting this integral is readily obtained. The result is itself a multivariate Gaussian density with the missing components eliminated from the complete data likelihood, i.e.,

$$f_{Y_3, Y_4, \dots, Y_M}(y_3, y_4, \dots, y_M) = (2\pi)^{-\frac{M-2}{2}} |V_o|^{-\frac{1}{2}} \cdot \exp \left\{ -\frac{1}{2} [\mathbf{y}_o - \mathbf{g}_o(\mathbf{x}; \beta)]^T V_o^{-1} [\mathbf{y}_o - \mathbf{g}_o(\mathbf{x}; \beta)] \right\} \quad (15)$$

where $\mathbf{y}_o^T = [y_3, y_4, \dots, y_M]$ is the observed data, $\mathbf{g}_o(\mathbf{x}; \beta)$ denotes the mean functions for the observed components and V_o is the matrix obtained by deleting the appropriate (here the first two) rows and columns of V

$$V_o = \begin{bmatrix} \sigma_3^2 & \sigma_{34} & \dots & \sigma_{3M} \\ \sigma_{34} & \sigma_4^2 & \dots & \sigma_{4M} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{3M} & \sigma_{4M} & \dots & \sigma_M^2 \end{bmatrix} \quad (16)$$

The likelihood function of Equation 12 covering all n trees will then be the product of complete data likelihoods for all trees without missing data and the (reduced) observed data likelihoods of the form of Equation 14 for trees with missing components. Thus, trees with different patterns of missingness will contain information on distinct subsets of model parameters and contribute differently to the overall likelihood.

The other relevant pattern of missingness for tree biomass data occurs when information is available for all components, but some components are observed only in aggregate. For example, with stem biomass components estimated from disc-based measurements, if bark and wood are not separated on some trees' discs then separate estimates of stem bark and stem wood may not be available for those trees. Or, if for some trees foliage is not separated from branches or if branches are not divided into specific size classes, then only total crown or total branch mass data will be available. These trees still contain information on all components in the sense that total tree biomass is still observed, but not all components are separately resolved. For these trees, the observed data likelihood takes a different form from Equation 14 but is again obtained by integrating over the complete data likelihood. For example, if the first two components are observed only in aggregate (i.e., $Z = Y_1 + Y_2$), then for that tree the relevant observed data likelihood is

$$f_{Z, Y_3, Y_4, \dots, Y_M}(z, y_3, y_4, \dots, y_M) \\ = \int_0^\infty \int_0^{z-y_2} f_{Y_1, Y_2, \dots, Y_M}(y_1, y_2, \dots, y_M) dy_1 dy_2$$

Although complex in the general case, this integral is also straightforward to evaluate in the Gaussian setting. This follows from the fact that the sum of Gaussian variables is itself Gaussian, with mean and variance that stem directly from the means and (co)variances of the summands. Thus,

$$f_{Y_1+Y_2, Y_3, Y_4, \dots, Y_M}(y_1 + y_2, y_3, y_4, \dots, y_M) \\ = (2\pi)^{-\frac{M-1}{2}} |V_a|^{-\frac{1}{2}} \cdot \exp \left\{ -\frac{1}{2} [\mathbf{y}_a - \mathbf{g}_a(\mathbf{x}; \beta)]^T V_a^{-1} [\mathbf{y}_a - \mathbf{g}_a(\mathbf{x}; \beta)] \right\}$$

where $\mathbf{y}_a^T = [y_1 + y_2, y_3, y_4, \dots, y_M]$ is the observed (partially aggregated) data, $\mathbf{g}_a(\mathbf{x}; \beta)$ is the vector of mean functions of observed components but with first element $g_1(\mathbf{x}_1; \beta_1) + g_2(\mathbf{x}_2; \beta_2)$, and V_a is the matrix obtained by collapsing the first two rows and columns of V

$$V_a = \begin{bmatrix} \sigma_1^2 + 2\sigma_{12} + \sigma_2^2 & \sigma_{13} + \sigma_{23} & \sigma_{14} + \sigma_{24} & \dots & \sigma_{1M} + \sigma_{2M} \\ \sigma_{13} + \sigma_{23} & \sigma_3^2 & \sigma_{34} & \dots & \sigma_{3M} \\ \sigma_{14} + \sigma_{24} & \sigma_{34} & \sigma_4^2 & \dots & \sigma_{4M} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{1M} + \sigma_{2M} & \sigma_{3M} & \sigma_{4M} & \dots & \sigma_M^2 \end{bmatrix} \quad (17)$$

Once again, these analytical operations need to be undertaken only for trees with aggregated component data. The full data likelihood of Equation 12 will consist of complete data likelihoods for trees without missing data and incomplete data likelihoods for the remaining trees.

By applying these results, ML estimation of Gaussian component biomass models with aggregative or disaggregative equations can proceed from data exhibiting either or both of these missingness patterns. As noted above, however, missing component biomass data have different consequences, depending on whether aggregative and disaggregative biomass equations are used. For instance, suppose a particular tree provides stemwood and stem bark biomass measurements but no crown biomass. With adoption of aggregative equations, the crown biomass component model must be dropped for that tree, and the tree will contribute no information for the estimation of the parameters governing expected crown mass. Yet if disaggregative equations are used, although the Y_c equation will still need to be dropped from a system such as Equation 11, the tree in question will still contribute to the estimation of all β parameters because no elements of β are unique to just one component equation. Conversely, suppose that all trees but one possess measurements of stemwood and stem bark biomass, whereas the remaining tree only has a record for undifferentiated stem biomass. In this case, the deviant tree cannot provide information useful for estimation of β_{wb} in a disaggregated model such as Equation 11; its contribution to the joint likelihood will be based on observed deviations from the expected values given by Equation 9. Yet if aggregated biomass equations are used, this tree will still contribute to the estimation of all elements of β , although its likelihood would be formed by adding together two of the equations in system 6.

Tree	D (cm)	H (m)	C (m)	Green mass (kg)			Total
				Stemwood	Stembark	Crown	
1	5.6	7.9	2.1	6.5	2.3	1.0	9.8
2	6.4	8.5	1.2	7.4	2.6	2.1	12.1
3	8.1	10.7	2.7	17.6	4.5	2.3	24.4
4	8.4	11.3	3.4	18.5	4.3	4.2	27.0
5	9.1	11.0	4.3	22.6	5.4	5.6	33.6
6	9.9	13.1	3.4	30.6	7.4	5.5	43.5
7	10.4	14.3	5.5	32.9	6.7	6.4	46.0
8	11.2	14.6	4.0	40.6	9.3	6.2	56.1
9	11.7	14.3	4.0	46.0	10.7	7.7	64.4
10	12.2	14.9	6.1	51.6	13.1	6.1	70.8
11	11.9	16.8	4.6	60.4	10.1	5.4	75.9
12	13.2	13.7	4.6	62.8	15.2	10.7	88.7
13	12.2	15.8	5.5	67.5	12.9	15.3	95.7
14	13.7	18.0	2.4	81.2	12.5	8.7	102.4
15	14.2	16.5	6.4	94.3	18.2	11.2	123.7
16	15.0	20.1	3.0	123.4	16.5	7.7	147.6
17	15.7	16.8	4.9	107.3	21.5	19.7	148.5
18	16.5	17.1	4.9	123.8	22.1	28.9	174.8
19	16.5	17.1	4.0	151.6	24.6	16.8	193.0
20	19.6	13.7	6.7	140.4	25.1	46.2	211.7
21	17.5	19.2	7.6	170.4	27.4	16.8	214.6
22	17.8	18.3	6.4	169.6	31.7	24.0	225.3
23	18.5	17.7	6.4	160.3	36.9	47.5	244.7
24	19.6	19.8	7.6	199.8	38.7	19.7	258.2
25	18.5	22.9	4.0	231.6	29.6	24.6	285.8
26	19.8	18.6	6.7	217.9	33.9	45.8	297.6
27	20.6	17.4	5.8	216.0	32.6	61.2	309.8
28	21.6	17.7	8.2	200.6	40.2	75.4	316.2
29	19.8	18.9	7.3	217.5	38.5	62.0	318.0
30	22.9	19.8	8.5	314.8	43.1	43.2	401.1
31	23.6	18.3	7.9	287.1	63.4	51.7	402.2
32	23.1	18.9	7.9	290.9	44.3	76.7	411.9
33	24.1	21.3	9.4	320.1	50.6	75.6	446.3
34	26.4	19.2	7.3	308.6	65.7	116.0	490.3
35	24.6	25.0	5.8	403.0	49.8	69.8	522.6
36	25.1	19.8	8.5	390.4	48.8	83.5	522.7
37	29.0	20.4	8.5	445.2	60.4	88.0	593.6
38	28.4	26.8	7.3	736.4	84.0	79.9	900.3
39	31.8	27.4	8.2	770.9	93.8	170.2	1034.9
40	33.0	27.7	10.4	921.3	108.0	169.2	1198.5

Figure 1. Slash pine biomass data from Parresol (2001) with tree dbh (D), total height (H), and live crown length (C). Shaded cells identify data masked from some analyses; within these cells component values shown in white are assumed missing (i.e., unknown along with the tree total) while those shown in black are assumed known only in aggregate (i.e., stem mass known but not stemwood and stembark separately).

Application Materials

ML estimation of biomass model parameters is illustrated using the green mass data previously presented by Parresol (2001). The data are from 40 slash pine trees grown in unthinned plantations in Louisiana, USA. Biomass totals are exclusive of 15-cm stumps, and the crown component represents the branch and foliage masses. Figure 1 reproduces the complete data and also identifies components taken as missing in subsequent analyses.

Biomass Models

An aggregative component biomass model for the slash pine data was developed using the biomass equations selected by Parresol (2001). Expressed as exponential functions these are

$$g_c(D, H; \beta_c) = \exp[\beta_{c1} + \beta_{c2} \ln D + \beta_{c3} \ln H] \quad (18a)$$

$$g_b(D; \beta_b) = \exp[\beta_{b1} + \beta_{b2} \ln D] \quad (18b)$$

$$g_w(D, H; \beta_w) = \exp[\beta_{w1} + \beta_{w2} \ln(D^2 H)] \quad (18c)$$

where D and H represent tree dbh and total height and the subscripts c , b , and w identify the crown, stembark, and stemwood

equations. For error variance functions, Parresol (2001) recommended

$$v_c(D, H; \theta_c) = \theta_{c1} D^{\theta_{c2}} e^{\theta_{c3} H^2}$$

$$v_b(D; \theta_b) = \theta_{b1} D^{\theta_{b2}}$$

$$v_w(D, H; \theta_w) = \theta_{w1} (D^2 H)^{\theta_{w2}}$$

but a simpler set of variance functions

$$v_m(D; \theta_m) = \theta_{1m}^2 D^{2\theta_{2m}} \quad (19)$$

were evaluated here. To complete the variance-covariance matrix, an unstructured set of pairwise cross-correlations ($\rho_{m,m'}$) was used to describe the associations among component error terms for the same tree, and error terms for distinct trees were presumed independent.

Against this, a disaggregative model of the form of Equation system 11 was then specified. This model was based on an equation for expected total biomass, a function to discriminate between crown and stem mass, and a function to discriminate between stembark and stemwood mass. The forms adopted for these equations were

$$g_t(D, H; \beta_t) = \exp[\beta_{t1} + \beta_{t2} \ln D + \beta_{t3} \ln H] \quad (20a)$$

$$g_{cs}(D; \beta_{cs}) = \frac{1}{1 + \exp[\beta_{cs1} + \beta_{cs2} \ln D + \beta_{cs3} \ln H]} \quad (20b)$$

$$g_{wb}(D; \beta_{wb}) = \frac{1}{1 + \exp[\beta_{wb1} + \beta_{wb2} \ln D + \beta_{wb3} \ln H]} \quad (20c)$$

Substituting these equations into Equation 11 provides the overall model structure. To complete the model, variances of component errors were characterized using Equation 19 and cross-covariances using a set of unstructured pairwise correlation parameters (as above).

Maximum Likelihood Estimation

The aggregative and disaggregative models were initially fit to the complete set of component data in Figure 1. The nlme package (Pinheiro et al. 2014) for the open-source program R (R Core Team 2014) was used for Gaussian ML. Specifically, as detailed in Appendix A, the gnls function of this package was used to obtain ML estimates. This function operates by alternating between optimization of the likelihood for β conditional on estimates of variance parameters and vice versa (see Pinheiro and Bates 2000, §7.5.2). To allow for simultaneous estimation of all three component biomass equations, the three mean functions (formulated on an exponential scale) were collapsed into a single function using logical flags, and the variance structure was specified using nlme's varComb() function.

For the aggregative model, the starting values of the β parameters were obtained from separate ordinary least-squares calibrations undertaken after logarithmic transformation of both the observed biomass and the component expectation Equations 18. This approach was also used to obtain starting values for β_t in the disaggregative model. The starting values for β_{cs} and β_{wb} in that model's discrimination functions were obtained from separate log-linear regressions of $Y_c/(Y_w + Y_b)$ and Y_b/Y_w , respectively, on D and H . With starting values for β in hand, model parameters were then estimated by ML using gnls.

After calibrations of the models against the complete data set, the patterns of missingness shown in Figure 1 were simulated, both separately (i.e., only crown component missing or only total stem biomass available) and in combination. The gnls algorithms can readily accommodate missing component data. Specifically, provided that components are consistently (across trees) ordered and identified by integer values (Appendix A), then missing values can be

included for unobserved components and gnls will appropriately factor the likelihoods of the affected trees to the form of Equation 15. On the other hand, missingness in the form of aggregated components (e.g., overall stem mass is observed, rather than separate stemwood and stembark masses) cannot be accommodated by standard inputs to this function. In particular, gnls presumes that component scedastic functions and component cross-correlation functions can be specified separately, using nlme's built-in variance and correlation class structures. This imparts computational efficiencies because the variance and correlation parameters can be separated in optimization. However, with this form of missingness the resulting covariance structure given by Equation 17 cannot be factored into variance and correlation matrices in the same way for all trees. One way to address this is to develop a nlme covariance class structure that incorporates both variance and cross-correlation parameters as well as the structure of the data for each particular tree. This is illustrated by an R script in the Supplemental Material[§] that redefines nlme's basic symmetric correlation class to provide a symmetric covariance class taking an additional argument describing the data structure for each tree. This last argument allows the general variance-covariance matrix (Equation 4) to be collapsed to a form (such as Equation 17) according to how the component biomass data themselves have been collapsed. With this covariance structure used as a correlation class in gnls, ML estimates of model parameters can be obtained when both patterns of missingness occur (Appendix B).

Model Comparison and Evaluation

As a matter of standard practice, circumspect evaluation of model-data agreement via graphical and statistical measures is recommended. Here, however, our emphasis is on demonstration of alternative estimation techniques for complete/incomplete data sets with respect to a fixed set of model forms. Therefore, models fit by ML to the complete and simulated-incomplete data are compared in terms of their estimated parameters and SEs. In addition, the resulting biomass estimation surfaces are evaluated and model root mean square errors (RMSEs) are compared. The latter are computed on the original, unweighted scale

$$\text{RMSE} = \sqrt{\frac{1}{40} \sum_i e_{mi}^2} \quad (21)$$

where the $e_{mi} = Y_{mi} - \hat{Y}_{mi}$ are the raw residuals for the m th component. Owing to the generalized variance structure of Equation 19, the raw residuals have unequal variability and are not well suited for residual diagnostics. Better for diagnostic purposes are the normalized residuals $r_i = \hat{V}_i^{-1/2} e_i$, where $e_i = (e_{1i}, e_{2i}, \dots, e_{Mi})^T$ is the vector of raw residuals for the i th tree, $\hat{V}_i$ is the corresponding estimated variance-covariance matrix, and $\hat{V}_i^{-1/2}$ is a Choleski factorization of that matrix. Normalized residuals are approximately Gaussian distributed with mean 0 and variance 1 if the model is correctly specified (Pinheiro and Bates 2000, p. 239).

Inasmuch as the RMSE statistic as calculated above makes no adjustment for the numbers of parameters estimated by the various models, we also compute the finite-sample corrected form of AIC (AICc) for each model

$$\text{AICc} = -2L(\beta, \theta, \rho | \mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_M, \mathbf{X}) + \frac{2nM d}{nM - d - 1} \quad (22)$$

Table 1. Estimated component biomass equation parameters and SEs for the aggregative model obtained by ML from the complete (Agg-ML) and incomplete (Agg-ML_{miss}) data as well as from the MSLS fits (Agg-2SLS and Agg-3SLS) reported in Parresol (2001).

Parameter	Agg-ML	Agg-ML _{miss}	Agg-2SLS	Agg-3SLS
$\hat{\beta}_{b1}$	-3.1465 0.1149	-3.1709 0.1136	-3.0731 0.1530	-3.0782 0.1151
$\hat{\beta}_{b2}$	2.2355 0.0426	2.2469 0.0424	2.2093 0.0522	2.2112 0.0397
$\hat{\beta}_{c1}$	-3.5593 0.5183	-3.6433 0.5635	-3.5980 0.4675	-4.2676 0.3111
$\hat{\beta}_{c2}$	3.5219 0.2376	3.5055 0.2617	3.6804 0.2570	3.5098 0.1470
$\hat{\beta}_{c3}$	-1.1222 0.3733	-1.0681 0.4094	-1.2624 0.3750	-0.8719 0.1930
$\hat{\beta}_{w1}$	-4.1560 0.1123	-4.1503 0.1109	-4.1127 0.1112	-4.1200 0.1059
$\hat{\beta}_{w2}$	1.0637 0.0133	1.0630 0.0132	1.0585 0.0131	1.0590 0.0124

where d counts the number of parameters in the mean and covariance functions. For comparability, RMSE and AICc are computed using the complete set of observations even for models fit to simulated-incomplete data.

Both RMSE and AICc are also computed for the aggregative models fit (on the original scale) by Parresol (2001) using 2-stage (2SLS) and 3-stage (3SLS) least squares. For these models, a Gaussian likelihood was used in AICc calculations along with the published variance functions. AICc for the 2SLS fit used the cross-correlation parameter estimates presented by Parresol (p. 870). Estimates of these parameters for the 3SLS fit were not published but were provided in computer code supplied by B.R. Parresol (USDA Forest Service, pers. comm., May 2001).

Results

ML estimates of component biomass equation parameters for the aggregative (Agg-ML) system are shown in Table 1 alongside the 2SLS (Agg-2SLS) and 3SLS (Agg-3SLS) estimates given in Parresol (2001). Both the estimates themselves and their estimated SEs changed only slightly when the missing data patterns of Figure 1 were imposed (Agg-ML_{miss}). The estimated SEs of the crown equation parameter estimates all increased, whereas those of the stembark and stemwood equation estimates fell, albeit to a lesser degree. The increased SEs for $\hat{\beta}_c$ presumably result from the complete loss of crown biomass data for 7 trees. On the other hand, the stemwood and stembark equation estimates are still informed by data from all 40 trees, although to a lesser extent by the 7 trees with only aggregated stem biomass data. The estimated SEs reported for Parresol's MSLS fits are similar to those from the ML fits, although those for his Agg-3SLS fit are generally the smallest.

For the disaggregative system, estimates of the parameters characterizing total tree biomass varied less between the complete (Dis-ML) and incomplete (Dis-ML_{miss}) data fits than did estimates of the stem-crown or stemwood-stembark discrimination parameters (Table 2). The same was true of the estimated SEs, although these increased across the board when missingness was simulated. Under the Dis-ML_{miss} model, all 40 trees contribute to the estimation of β_t even when the missing data patterns of Figure 1 are

[§] Supplementary data are available with this article at <http://dx.doi.org/10.5849/forsci.15-126>.

Table 2. Estimated component biomass equation parameters and standard errors for the disaggregative model obtained by ML from the complete (Dis-ML) and incomplete (Dis-ML_{miss}) data.

Parameter	Dis-ML	Dis-ML _{miss}
$\hat{\beta}_{c1}$	-3.1772	-3.1536
	0.1586	0.1638
$\hat{\beta}_{c2}$	2.3157	2.3319
	0.0742	0.0780
$\hat{\beta}_{c3}$	0.6449	0.6217
	0.1133	0.1188
$\hat{\beta}_{sc1}$	0.1101	-0.0006
	0.5703	0.6109
$\hat{\beta}_{sc2}$	1.5058	1.4795
	0.2643	0.2872
$\hat{\beta}_{sc3}$	-2.1843	-2.1122
	0.4144	0.4492
$\hat{\beta}_{wb1}$	0.9623	1.1852
	0.2507	0.2671
$\hat{\beta}_{wb2}$	0.0831	0.2304
	0.1195	0.1300
$\hat{\beta}_{wb3}$	-1.0221	-1.2433
	0.1830	0.2034

Table 3. ML estimates of covariance parameters for aggregative and disaggregative models using complete (ML) and simulated-incomplete (ML_{miss}) data.

Parameter	Agg-ML	Agg-ML	Dis-ML	Dis-ML _{miss}
$\hat{\theta}_{c1}$	0.0019	0.0025	0.0019	0.0024
$\hat{\theta}_{c2}$	2.9058	2.8228	2.9122	2.8287
$\hat{\theta}_{b1}$	0.0022	0.0019	0.0028	0.0018
$\hat{\theta}_{b2}$	2.5660	2.5943	2.4649	2.6128
$\hat{\theta}_{w1}$	0.0043	0.0037	0.0053	0.0048
$\hat{\theta}_{w2}$	2.8195	2.8710	2.7470	2.7813
$\hat{\rho}_{cb}$	-0.0164	0.1153	-0.0225	0.0898
$\hat{\rho}_{cw}$	-0.0721	-0.0473	-0.0829	-0.0543
$\hat{\rho}_{wb}$	0.3559	0.4119	0.3303	0.4055

imposed. However, the seven trees missing crown biomass provide less information for estimating β_{sc} , and the seven trees lacking separate stembark and stemwood biomass measurements provide no information for estimating β_{wb} .

Estimates of covariance parameters by model and data set are given in Table 3. The variance parameter estimates reinforce the heteroscedastic nature of biomass observations, both across components ($\hat{\theta}_{w1}$ are approximately double $\hat{\theta}_{b1}$ or $\hat{\theta}_{c1}$) and across tree sizes ($\hat{\theta}_2 > 0$). Table 3 does not show estimated SEs, but the ML estimates of the cross-correlation between stemwood and stembark biomass were consistently more than 2 estimated SEs from 0, whereas the other cross-correlation estimates were consistently within 2 estimated SEs of 0. After Equations 18 were fit separately by 2SLS, Parresol (2001, p. 870) recovered cross-correlations of 0.29 (crown-stembark), 0.01 (crown-stemwood), and 0.19 (stemwood-stembark). The first and last of these are different from the ML estimates in Table 3. Estimated cross-correlations under his Agg-3SLS fit were not reported but were communicated as 0.013 (crown-stembark), -0.115 (crown-stemwood), and 0.283 (stemwood-stembark). These agree more closely with the ML correlations in Table 3.

Despite variations in functional forms and parameter estimates, component and total tree biomass estimates were similar across models. The most pronounced differences between the Agg-ML and Dis-ML estimates were for trees with low heights given their dbh. For instance,

Table 4. Component RMSE for alternative models, parameter estimation routines, and calibration data sets.

Model	Crown	Bark	Wood	Total
	(kg).....			
Agg-ML	13.3	5.1	26.9	30.4
Agg-ML _{miss}	13.2	5.2	26.9	30.7
Agg-2SLS	13.9	5.0	26.7	31.4
Agg-3SLS	12.8	5.0	26.7	29.8
Dis-ML	13.4	4.8	26.2	29.6
Dis-ML _{miss}	13.4	5.4	26.2	30.3

applying Agg-ML coefficients, trees with dbh 20 cm and height 15 m are estimated to have, on average

$$g_c(20, 15; \hat{\beta}_c) = \exp[-3.559 + 3.522 \ln 20 - 1.122 \ln 15] = 52.1 \text{ kg}$$

$$g_b(20; \hat{\beta}_b) = \exp[-3.146 + 2.236 \ln 20] = 34.8 \text{ kg}$$

$$g_w(20, 15; \hat{\beta}_w) = \exp[-4.156 + 1.064 \ln(20^2 15)] = 163.7 \text{ kg}$$

of biomass in their crowns, stembark, and stemwood components, for an estimated aboveground biomass of 250.6 kg. Applying Equation 13, the estimated SEs for these components are 4.8, 0.8, and 2.5 kg, respectively, and 5.6 kg for the total. With the Dis-ML model, the total biomass is estimated as

$$g_c(20, 15; \hat{\beta}_t) = \exp[-3.177 + 2.316 \ln 20 + 0.645 \ln 15] = 246.2 \text{ kg}$$

with estimated SE 8.1 kg. The crown, stembark, and stemwood components are, respectively,

$$246.2 \text{ kg} \left[1 - \frac{1}{1 + e^{(0.11 + 1.506 \ln 20 - 2.184 \ln 15)}} \right] = 53 \text{ kg}$$

$$246.2 \text{ kg} \left[\frac{1}{1 + e^{(0.11 + 1.506 \ln 20 - 2.184 \ln 15)}} \right] \left[1 - \frac{1}{1 + e^{(0.962 + 0.083 \ln 20 - 1.022 \ln 15)}} \right] = 33.7 \text{ kg}$$

$$246.2 \text{ kg} \left[\frac{1}{1 + e^{(0.11 + 1.506 \ln 20 - 2.184 \ln 15)}} \right] \left[\frac{1}{1 + e^{(0.962 + 0.083 \ln 20 - 1.022 \ln 15)}} \right] = 159.6 \text{ kg}$$

with respective estimated SEs of 4.7, 1.7, and 6.8 kg. For a given dbh, the respective estimates from the Agg-ML and Dis-ML models converge as height is increased.

Looking across the range of the predictors, RMSE differed little across models and estimation routines (Table 4). As expected, the RMSE for total biomass was lower for models fit to the complete data set. However, this was not the case for all components: for some components the RMSE was the same as or lower for Agg-ML_{miss} or Dis-ML_{miss} relative to those for Agg-ML or Dis-ML. This can occur because, as calculated here, RMSE is an unweighted measure of data-model agreement, and each model fit is associated with a unique set of weight (i.e., variance) functions (cf. Table 3).

Figures 2 and 3 show Gaussian probability plots of the normalized residuals from the Agg-ML and Dis-ML models. The distributions of residuals from these and other models had light left tails. Assessed against 100 parametric bootstrap resampling results, however, the normalized residual plots exhibited no pronounced outliers or deviations from Gaussian densities.

Table 5 provides measures of model-data agreement in terms of

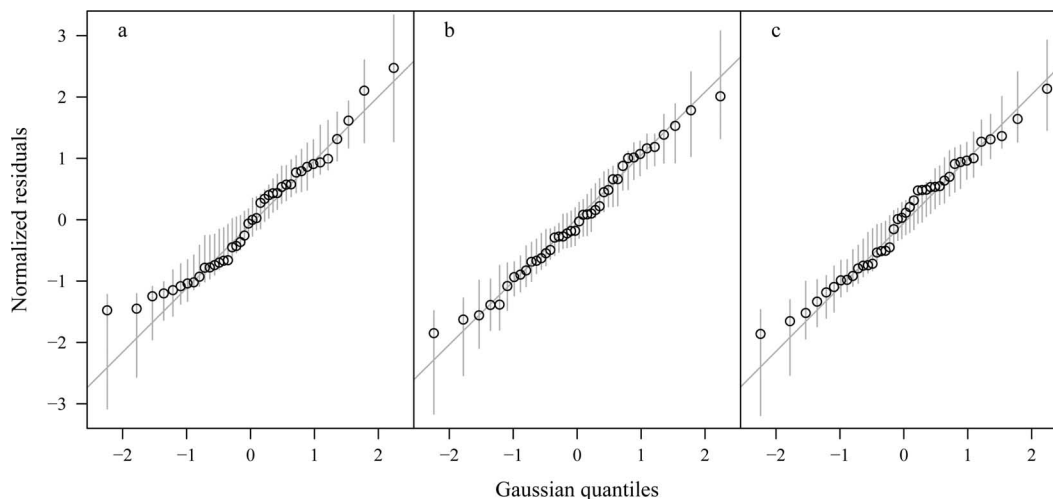


Figure 2. Gaussian probability plots of normalized residuals from the Agg-ML model for crown (a), stembark (b), and stemwood (c) biomass. Gray line segments denote ranges of normalized residuals from 100 parametric bootstrap sample fits.

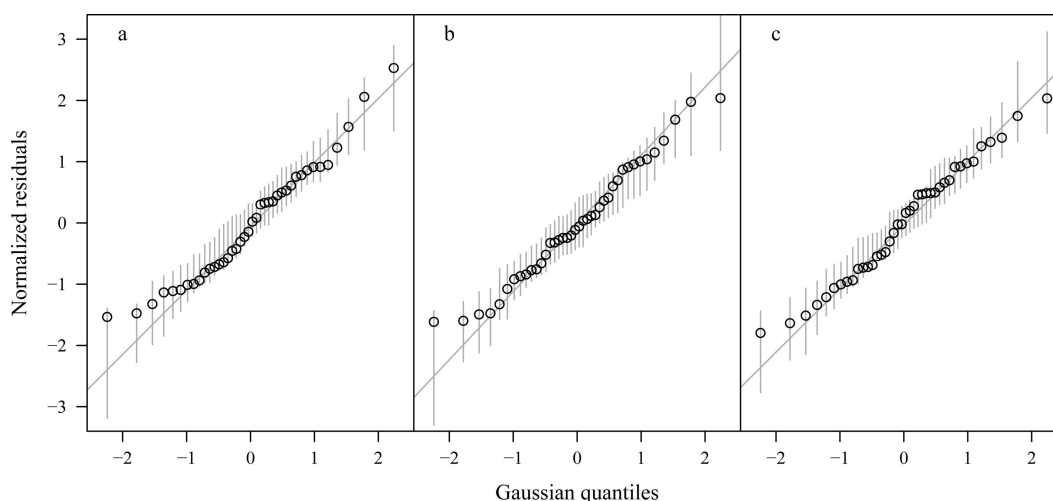


Figure 3. Gaussian probability plots of normalized residuals from the Dis-ML model for crown (a), stembark (b), and stemwood (c) biomass. Gray line segments denote ranges of normalized residuals from 100 parametric bootstrap sample fits.

Table 5. Number of parameters, complete data log-likelihood, and AICc by model, estimation technique, and calibration data set.

Model	d	lnL	AICc
Agg-ML	16	-373.7	784.6
Agg-ML _{miss}	16	-374.5	786.4
Agg-2SLS	17	-377.8	795.6
Agg-3SLS	22	-377.1	808.5
Dis-ML	18	-373.2	789.1
Dis-ML _{miss}	18	-375.0	792.8

d, number of parameters; lnL, complete data log-likelihood.

Gaussian probabilities. By these measures, the ML-estimated models provided better descriptions of the slash pine data set, with Agg-ML providing the best characterization after accounting for the dimension of the parameter vector. As expected, the models obtained from the incomplete data were associated with smaller log-likelihoods as these were evaluated using all the data. The models obtained by Parresol (2001) had smaller likelihoods and larger AICc than any ML fits. This is not surprising given that the MSLS fits were not obtained by optimizing a Gaussian probability model. Indeed, the structure implied by the model Parresol used to obtain the Agg-3SLS parameter estimates

does not correspond with a probability model of any kind. We emphasize, however, that this does not imply that his MSLS models are inferior to the ML models obtained here in terms of predictive accuracy. The RMSEs of the models are very similar (Table 4) such that the differences from a likelihood or AIC perspective may be due in large part to differences in the fitted covariance structures.

Discussion

Aggregative and disaggregative biomass equations can be developed and fit within statistical frameworks built on valid probability models. Doing so requires that the additivity of the biomass data themselves be recognized. In turn, this implies that the dependence structure of biomass data collected from the same tree be described only for the observable components and that only the component biomass observations (not observed totals or subtotals) be used in parameter estimation. This departs from previous presentations of aggregative and disaggregative systems (Parresol 2001, Dong et al. 2015) where component data as well as their (conditionally redundant) totals are used in MSLS estimation procedures. In principle, valid probability models could also be calibrated using observations of $M - 1$ biomass components plus the total, but doing so would

require that additional constraints be imposed to preclude negative estimates of the M th component. Notably, these considerations also carry over to the linear modeling context.

If a system of component biomass models has a valid probabilistic structure, its parameters can be estimated by ML. This is advantageous when a credible probability law can be advanced because ML then allows for efficient use of the available data. It also facilitates the use of information-based criteria for model selection and is crucial for Bayesian model development and inference. In addition, it allows for consistent treatment of partially observed component biomass data. Specifically, if component data are missing (or aggregated) at random (cf. Little and Rubin 1987, ch. 1), then ML remains fully efficient. For a given data set, its use obviates the need to drop trees or impute missing component data. For instance, the proposed approach can be readily applied where high excavation costs preclude measurement of belowground biomass on all sample trees. More broadly, its use can facilitate assimilation of data sets collected by researchers measuring different biomass components. For example, one destructive sampling effort may yield only estimates of total branch biomass, whereas another may identify the distribution of branch biomass across several size-classes. By recognizing the aggregated nature of the former data set, the two could be pooled to jointly estimate branch biomass within each size-class. Of course, in evaluating partially observed component biomass data, one must carefully consider whether the missingness mechanism is informative or not. Trees lacking observations of foliage biomass owing to defoliation by insects may, for example, also have had biomass distributions that predisposed them to attack. In such cases, the missingness mechanism itself must be modeled if partially observed data are to be pooled with complete records (Little and Rubin 1987, ch. 11).

Joint calibration of systems of biomass component equations allows for the use of cross-component dependencies in parameter estimation. In principle, this improves estimator accuracy at the system level. At the same time, by minimizing a system-level loss function, one potentially sacrifices a greater precision that might be obtained for a particular component or subtotal if it were fit separately. This could be disadvantageous if a particular component or subtotal is of overriding importance. For instance, in the application studied here, a model for total slash pine biomass might yield more accurate estimates if fit exclusively to the observed totals (see e.g., Parresol 1999). In addition, more accurate estimates of totals may be obtained by fitting an additive system of equations but using both the component data and the observed totals. For example, Parresol's (2001) 3SLS procedure reduced the RMSE for total biomass relative both to his 2SLS procedure and the ML fit obtained here (Table 4). It achieves this by reweighting the component residuals in the objective function. For ML the latter is a quadratic function of the component residuals weighted by the variance functions of those components and their cross-correlations; for 3SLS the component residuals are also weighted by a specified variance function for the total and an additional set of cross-correlations (between components and totals). In this indirect reweighting of component residuals, as well as in the fact that it implies an invalid probability model, it is a curious (albeit effective) way to seek improved precision for totals. An alternative strategy using ML is to fix or premultiply elements of the component variance functions (e.g., the θ_{1m} in Equation 19) to reflect greater importance of some components vis-à-vis others. This allows for importance-based weighting to overcome uncertainty-based weighting, although it is unclear how this could be used to improve estimates of (sub)totals as opposed to individual components.

Joint ML estimation of nonlinear component biomass models is

readily achieved using open-source software, as documented above, at least in the Gaussian setting. Moreover, in this same setting, missing or aggregated component observations can be accommodated by exploiting the extendable structure of the nlme package for R. This is important given the common invocation, implicit or explicit, of Gaussian models in biomass modeling. Nonetheless, generalization of these methods beyond the standard Gaussian model would serve to facilitate nonlinear mixed-effects modeling to account for dependence across trees as well as the characterization of fundamentally non-Gaussian data. This is the subject of continuing research.

Literature Cited

- AKAIKE, H. 1973. Information theory and an extension of the maximum likelihood principle. P. 267–281 in *Second International Symposium on Information Theory*, Petrov, B.N., and F. Csaki (eds.). Akademiai Kiado, Budapest.
- BI, H., J. TURNER, AND M.J. LAMBERT. 2004. Additive biomass equations for native eucalypt forest trees of temperate Australia. *Trees* 18:467–479.
- BROWN, J.K. 1978. *Weight and density of crowns of Rocky Mountain conifers*. USDA For. Serv., Res. Pap. INT-197, Intermountain Forest and Range Experiment Station, Ogden, UT. 56 p.
- DONG, L., L. ZHANG, AND F. LI. 2015. A three-step proportional weighting system of nonlinear biomass equation. *For. Sci.* 61:35–45.
- EVERT, F. 1985. *Systems of equations for estimating oven-dry mass of 18 Canadian tree species*. Canadian Forestry Service, Infor. Rep. PI-X-59, Petawawa National Forestry Institute, Chalk River, ON, Canada. 58 p.
- GUEORGIEVA, R., R. ROSENHECK, AND D. ZELTERMANA. 2008. Dirichlet component regression and its applications to psychiatric data. *Comput. Stat. Data Anal.* 58:5344–5355.
- JENKINS, J.C., D.C. CHOJNACKY, L.S. HEATH, AND R.A. BIRDSEY. 2003. National-scale biomass estimators for United States tree species. *For. Sci.* 49:12–35.
- LEHMANN, E.L., AND G. CASELLA. 1998. *Theory of point estimation*, 2nd ed. Springer, New York. 590 p.
- LITTLE, R.J.A., AND D.B. RUBIN. 1987. *Statistical analysis with missing data*. John Wiley and Sons, New York. 408 p.
- NAIDU, S.L., E.H. DELUCIA, AND R.B. THOMAS. 1998. Contrasting patterns of biomass allocation in dominant and suppressed loblolly pine. *Can. J. For. Res.* 28:1116–1124.
- PARRESOL, B.R. 1999. Assessing tree and stand biomass: A review with examples and critical comparisons. *For. Sci.* 45:573–593.
- PARRESOL, B.R. 2001. Additivity of nonlinear biomass equations. *Can. J. For. Res.* 31:865–878.
- PINHEIRO, J., D. BATES, S. DEBROY, D. SARKAR, AND R. CORE TEAM. 2014. *nlme: Linear and nonlinear mixed effects models*. R package version 3.1-118, R Foundation for Statistical Computing, Vienna, Austria.
- PINHEIRO, J.C., AND D.W. BATES. 2000. *Mixed-effects models in S and S-PLUS*. Springer, New York. 530 p.
- R CORE TEAM. 2014. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- SANQUETTA, C.R., A. BEHLING, A.P.D. CORTE, S.P. NETTO, A.B. SCHIKOWSKI, AND M. DO AMARAL. 2015. Simultaneous estimation as alternative to independent modeling of tree biomass. *Ann. For. Sci.* 72(8):1099–1112.
- TANG, S., Y. LI, AND Y. WANG. 2001. Simultaneous equations, error-in-variable models, and model integration in systems ecology. *Ecol. Model.* 142:285–294.
- TANG, S., AND Y. WANG. 2002. A parameter estimation program for the error-in-variable model. *Ecol. Model.* 156:225–236.
- TANG, S., H. ZHANG, AND H. XU. 2000. Study on establish and estimate method of compatible biomass model. *Sci. Silv. Sin.* 36(Suppl. 1):19–27 (in Chinese with English abstract).
- ZHAO, D.H., M. KANE, D. MARKEWITZ, R. TESKEY, AND M. CLUTTER. 2015. Additive tree biomass equations for midrotation loblolly pine plantations. *For. Sci.* 61:613–623.

Appendix A

The following R code obtains ML parameter estimates for the aggregated and disaggregated component biomass models from the complete slash pine data set. The code presumes that the data presented in Figure 1 (and in the Supplemental Materials file slash.csv) are held in a data frame called slash. These data must first be transposed (one row per tree $\times$ component) and appropriately coded for subsequent functions:

```
# transpose biomass data to long form
slashlong <- data.frame(biomass=with(slash,c(crown,bark,wood)),
  dbh=with(slash,rep(dbh,3)),
  ht=with(slash,rep(ht,3)),
  comp=factor(rep(c("crown","bark","wood"),each=nrow(slash)),
    levels=c("crown","bark","wood"),ordered=T),
  tree=rep(1:nrow(slash),3) )
# order data by tree and component
slashlong <- slashlong[order(slashlong$tree,slashlong$comp),]
# associate integer codes with components
slashlong$comp.index <- as.integer(slashlong$comp)
```

The slashlong data frame can then be used to fit the aggregated (steps A1–A3) and disaggregated (D1–D3) models by Gaussian maximum likelihood:

```
# A1. Fit log-linear versions of component models for starting values
woodlm <- lm(log(wood) ~ log(dbh*dbh*ht), data=slash)
barklm <- lm(log(bark) ~ log(dbh), data=slash)
crownlm <- lm(log(crown) ~ log(dbh) + log(ht), data=slash)

loglinstart <- c(coef(woodlm),coef(barklm),coef(crownlm))
names(loglinstart) <- c("w.b1","w.b2","b.b1","b.b2","c.b1","c.b2","c.b3")

# A2. Specify a joint mean function in exponential form
agg.biom <- function(dbh,ht,comp, w.b1,w.b2, b.b1,b.b2, c.b1,c.b2,c.b3){
  exp(w.b1 + w.b2*log(dbh*dbh*ht))*(comp=="wood") +
  exp(b.b1 + b.b2*log(dbh))*(comp=="bark") +
  exp(c.b1 + c.b2*log(dbh) + c.b3*log(ht))*(comp=="crown") }

# A3. Fit aggregative component biomass model
agg.fit <- gnls(biomass ~ agg.biom(dbh,ht,comp,w.b1,w.b2,b.b1,b.b2,c.b1,c.b2,c.b3),
  data=slashlong,na.action=na.pass,
  start=loglinstart,
  weights=varComb(varIdent(form = ~ 1 | comp),
    varPower(form = ~ I(dbh) | comp)),
  correlation=corSymm(form = ~ comp.index | tree),
  control=gnlsControl(nlsTol=.00001,msVerbose=T),verbose=T)

# D1. Fit log-linear versions of total/discrimination equations for starting values
totalbiom <- lm(log(total) ~ log(dbh) + log(ht), data=slash)
stemfrac <- lm(log(crown/(wood+bark)) ~ log(dbh) + log(ht), data=slash)
woodfrac <- lm(log(bark/wood) ~ log(dbh) + log(ht), data=slash)

startval <- c(coef(totalbiom),coef(stemfrac),coef(woodfrac))
names(startval) <- c("t.1","t.2","t.3", "sc.1","sc.2","sc.3", "wb.1","wb.2","wb.3")

# D2. Specify a joint mean function in exponential form
disagg.biom <- function(dbh,ht,comp,t.1,t.2,t.3,sc.1,sc.2,sc.3,wb.1,wb.2,wb.3){
  exp(t.1 + t.2*log(dbh) + t.3*log(ht))/
  (1 + exp(-sc.1 - sc.2*log(dbh) - sc.3*log(ht))*(comp=="crown"))/
  (1 + exp(sc.1 + sc.2*log(dbh) + sc.3*log(ht))*(comp!="crown"))/
  (1 + exp(-wb.1 - wb.2*log(dbh) - wb.3*log(ht))*(comp=="bark"))/
  (1 + exp(wb.1 + wb.2*log(dbh) + wb.3*log(ht))*(comp=="wood")) }

# D3. Fit disaggregative component biomass model
disagg.fit <- gnls(biomass ~ disagg.biom(dbh,ht,comp,t.1,t.2,t.3,sc.1,sc.2,sc.3,
  wb.1,wb.2,wb.3),
  data=slashlong,na.action=na.pass,
  start=startval,
  weights=varComb(varIdent(form = ~ 1 | comp),
    varPower(form = ~ I(dbh) | comp)),
  correlation=corSymm(form = ~ comp.index | tree),
  control=gnlsControl(nlsTol=.00001,msVerbose=T),verbose=T)
```

Appendix B

The procedures below perform Gaussian ML of the aggregative and disaggregative component biomass systems, given the patterns of missing and aggregated data displayed in Figure 1. To do so, the “observed” data first need to be stored in a new data frame with appropriate deletions and aggregations. Below, this new data frame is called slashlong2 and includes a new comp factor level “stem” (comp.index=4) with values only for the seven trees having combined wood and bark observations. In addition, the structure of the available data for each tree must be coded using matrices S_i :

$$S_i = \begin{cases} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \text{for complete tree records;} \\ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} & \text{for trees with missing crown records;} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} & \text{for trees with aggregated stem records;} \\ \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} & \text{for trees with aggregated stem records} \\ & \text{and missing crown records.} \end{cases}$$

Combining these matrices into a list called comps and sourcing covSymmFunction.R (see Supplemental Material), ML estimates can be obtained using the following code:

```
# AM1. Specify a joint mean function that allows for aggregated data
agg.biom2 <- function(dbh,ht,comp, w.b1,w.b2, b.b1,b.b2, c.b1,c.b2,c.b3){
  exp(w.b1 + w.b2*log(dbh*dbh*ht))*(comp=="wood" | comp=="stem") +
  exp(b.b1 + b.b2*log(dbh))*(comp=="bark" | comp=="stem") +
  exp(c.b1 + c.b2*log(dbh) + c.b3*log(ht))*(comp=="crown") }

# AM2. Fit aggregative system to missing data
agg.fit2 <- gnls(biomass ~ agg.biom2(dbh,ht,comp,w.b1,w.b2,b.b1,b.b2,c.b1,c.b2,c.b3),
  data=slashlong2,
  start=loglinstart,
  correlation=covSymm(value=double(3),
    varvalue=c(1,1,0,0,0),
    form = ~ comp.index | tree,
    varform= ~ dbh|comp,
    comps=comps),
  control=gnlsControl(nlsTol=.00001,msVerbose=T),verbose=T)

# DM1. Fit disaggregative component biomass model to missing data
disagg.fit2 <- gnls(biomass ~ disagg.biom(dbh,ht,comp,t.1,t.2,t.3,sc.1,sc.2,sc.3,
  wb.1,wb.2,wb.3),
  data=slashlong2,
  start=startval,
  correlation=covSymm(value=double(3),
    varvalue=c(1,1,0,0,0),
    form = ~ comp.index | tree,
    varform= ~ dbh|comp,
    comps=comps),
  control=gnlsControl(nlsTol=.00001,msVerbose=T),verbose=T)
```

The covSymm() function used in the above gnls steps defines a symmetric but collapsible covariance class.