Bayesian Analysis (2013) **8**, Number 1, pp. 133–168

Feedback and Modularization in a Bayesian Meta–analysis of Tree Traits Affecting Forest Dynamics

Kiona Ogle *, Jarrett Barber † and Karla Sartor ‡§

Abstract. We describe a unique application of modularization in the context of a Bayesian meta–analysis of quantitative information obtained from the literature. Incomplete reporting, resulting in large amounts of missing data, is common in many meta–analyses, and, in this study, it led to poor mixing and identifiability problems in a fully Bayesian meta–analysis model. As an alternative to the full Bayesian approach, we modularized model components (e.g., modules of covariates, sample sizes, and standard errors) to prevent missing covariate data in these modules from allowing feedback that would otherwise affect parameters in the covariate module (direct feedback control) or affect covariate effects parameters in the mean model for the response (indirect feedback control). The combination of direct and indirect feedback control greatly improves mixing and facilitates convergence of Markov chain Monte Carlo (MCMC), yielding realistic pseudo–posteriors. Thus, our modularization approach allowed us to address important limitations of existing meta–analytic methods by accommodating incomplete reporting and by considering all model quantities as stochastic, including the response variable of interest (e.g., a sample mean) and sample sizes, standard errors, and all covariates, reported or not. We illustrate our approach using data summaries extracted from literature on specific leaf area (SLA) of trees, an important functional trait linked to tree growth and forest dynamics and a key parameter in models of forest responses to climate change. A hierarchical model based on taxonomic relationships allows borrowing of strength to infer SLA for 305 tree species in the United States based on information for 158 of those species. In the context of the SLA meta–analysis, we discuss problems that arise from feedback among model components and provide ecological arguments for modularization—for "cutting feedback." We anticipate that our approach may be applied to meta–analyses of other important tree traits and to similar meta–analytical studies in general.

Keywords: cutting feedback, ecological meta–analysis, hierarchical Bayesian model, incomplete reporting, leaf traits, literature data, missing data, modularization, plant functional traits, specific leaf area, TreeTraits database, variable reporting

---

*School of Life Sciences, Arizona State University, Tempe, AZ, Kiona.Ogle@asu.edu

†School of Mathematical & Statistical Sciences, Arizona State University, Tempe, AZ, Jarrett.Barber@asu.edu

‡Department of Organismic & Evolutionary Biology, Harvard University, Cambridge, MA,

§Current affiliation: The Nature Conservancy in New Mexico, Santa Fe, NM ksartor@tnc.org

DOI:10.1214/13-BA806

1 Introduction

Recently, we have begun to see systematic exploration of and recommendations for "cutting feedback" or, synonymously, "modularization" in models wherein one or more model components or "modules" whose data or structure may be suspect, may have an unrealistic or otherwise undesirable influence on other modules (Jackson et al. 2009; Liu et al. 2009; Lunn et al. 2009a). Cutting feedback between model components combined via Bayes theorem is easily accomplished in the programming languages WinBUGS (Lunn et al. 2000) and OpenBUGS (Lunn et al. 2009b) by using their built–in "cut" functions (e.g., Spiegelhalter et al. 2003; Jackson et al. 2009; Lunn et al. 2009a; Molitor et al. 2009). It seems to us that a substantial consideration underlying a Bayesian approach to inference lies in the apparent success of combining various sources of information in the development of complex models (e.g., Molitor et al. 2009). Of course, as models become more complex, with increasing numbers of modules, modularization becomes more relevant.

In the current article, we describe a modularization case study in the form of a Bayesian meta–analysis. Our meta–analysis offers a relatively rich but easy to follow structure for illustrating modularization, and our modularized results offer an intuitively appealing interpretation. Moreover, our meta–analysis model stands on its own for its relatively simple but novel way of handling missing information associated with incomplete reporting, a pervasive aspect of meta–analytical studies that has challenged such studies to date (e.g., Lajeunesse and Forbes 2003; Wiebe et al. 2006). Significant progress has been made over the past decade with respect to statistical methods for analyzing data summaries and treatment effects reported in the literature (e.g., Hartung et al. 2008). In many cases, the primary literature is mined for quantitative information and meta–analytical tools employed to synthesize the information (Osenberg et al. 1997; Hedges et al. 1999; Osenberg et al. 1999; Gurevitch et al. 2001; Hartung et al. 2008). Still, these methods suffer several limitations.

Three important issues include heterogeneity across and within studies, simplistic treatment of reported uncertainties, and incomplete reporting (Gurevitch and Hedges 1999; Sutton and Higgins 2008). Although the first issue has been addressed by Bayesian meta–analyses that incorporate study random effects (e.g., Chung et al. 2002; Burr and Doss 2005; Higgins et al. 2009), the latter two are rarely addressed in the existing literature. Our approach explicitly accommodates these issues via study–level random effects, explicit modeling of standard errors and sample sizes, and latent modeling of missing values associated with incomplete reporting.

Our approach leads to a multi–module model wherein we might expect feedback problems associated with random effects (e.g., Lunn et al. 2009a) or with missing covariate values (Lunn et al. 2009a; Molitor et al. 2009). The latter is somewhat analogous to the situation described by Lunn et al. (2009a) wherein pharmacodynamic (PD) response data (EEG measurements) feed back to influence, perhaps undesirably, the pharmacokinetic (PK) covariate model component (plasma midazolam concentrations). Indeed, we experienced poor mixing conditions in our initial meta–analysis due to incomplete reporting and feedback issues, which initially led us to explore modularization.

In our case, we explore modularization strategies similar to those described by Lunn et al. (2009a) and Liu et al. (2009), with unique differences arising from our meta–analysis context, but that are also relevant in other hierarchical modeling settings. In particular, we introduce the notions of direct and indirect feedback to distinguish different modularization methods. Before describing our full meta–analysis approach, we first discuss a simplified example to introduce concepts used later.

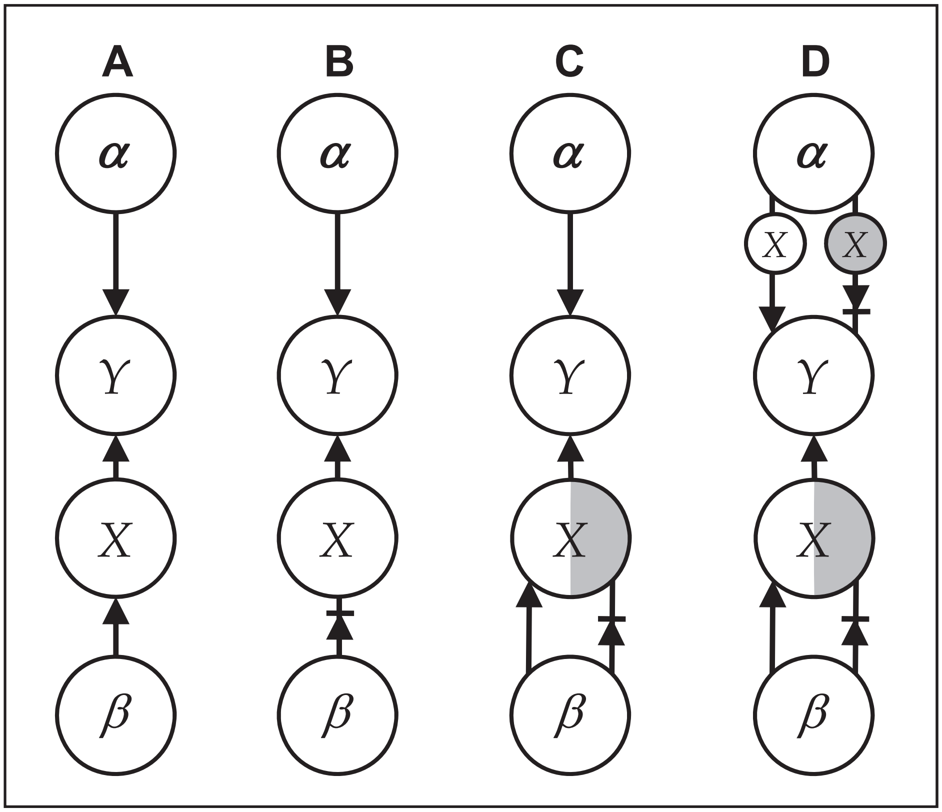

Figure 1: Definition of node and edge notation used to illustrate different approaches to controlling feedback in the presence of missing data. See text in Section 1 for discussion.

Modularization is illustrated in the very simple directed acyclic graphs (DAGs) of Figure 1. For illustrative purposes, consider a response variable (Y), a stochastic covariate (X), a covariate effect parameter (α), and a covariate parameter (β). DAG A illustrates the typical dependency between parent nodes and child nodes, where, for example, the full conditional distribution of X depends on β, and vice versa. That is, feedback occurs from X to β. DAG B indicates feedback control whereby the full conditional of β does not depend on X, but the full conditional of X still depends on β. This would correspond to a situation where β is treated as a distributional constant.

The “valve” or “diode” symbol ($\rightarrow|$) (e.g., Lunn et al. 2009a) is suggestive of the unidirectional feedback control and the modularization of X, now orphaned from its parent, β, in the sense that X no longer affects β. In DAG C, X is partitioned into observed (white fill) and missing (gray fill) values (e.g., Molitor et al. 2009), and feedback is not allowed from missing X to β and, hence, missing X are modularized; the full conditional of missing X depends on β, but the full conditional of β does not depend on missing X. DAGs B and C represent modularization associated with controlling the direct feedback between X and β (“direct” because X and β share an edge).

In DAG D (Figure 1) we may still consider edges between X and β as in DAG C, but we also depict cutting of the feedback from Y to α to be dependent on the missingness of X. A motivation for this is when we want the full conditional for α to be independent of X or a subvector of X. In this case, X and α are co–parents of Y, and the full conditional of α (the effect of X on Y) normally depends on the X covariate values. By cutting the edges between the effect α and those responses Y that are associated with missing covariate X values (gray filled circles), the full conditional of effect α no longer depends on the missing X covariate values. We refer to this as indirect feedback control because α and X do not share an edge, though, it may be argued that this is direct in the sense that the full conditional of α may depend directly on X. We further illustrate cases C and D (Figure 1) in Section 4, giving full conditionals for various cut and uncut scenarios in our specific leaf area (SLA) meta–analysis model. In particular, the indirect feedback case of Section 4.3 illustrates how missing tree age covariate values in the mean model for the response, specific leaf area, no longer allow feedback to covariate effect parameters in the mean model; the full conditional for the covariate effect parameters no longer depends on missing age covariates. See also Figure 2 in Section 3 below.

For our meta–analysis model, we focus on specific leaf area (SLA) of trees because SLA is a key parameter in models of forest dynamics (Bonan et al. 2003; Duursma et al. 2005; Tatarinov and Cienciala 2006; Ogle and Pacala 2009), which is an important consideration in the context of climate change and its potential impacts on forests and terrestrial carbon cycling (Melillo et al. 1993; Schimel 1995; Norby et al. 1999). The meta–analysis framework described here provides a potential approach for parameterizing such models with literature information. We apply our meta–analysis method to over 1800 records of SLA sample means, associated measures of uncertainty, and covariate information extracted from 182 published studies. One goal is to obtain estimates of latent SLA for 305 tree species and 85 genera in the United States, which we will later use to inform an individual–based, species–specific model of forest dynamics. In this regard, this study is the result of a larger ongoing effort aimed at quantifying tree functional traits affecting tree growth and mortality and forest dynamics. We implemented our meta–analyses in OpenBUGS (Lunn et al. 2009b) and provide the model code in the Supplementary Materials.

In Section 2, we describe the literature search that we conducted to arrive at our “data” for our meta–analysis, which we describe in Section 3. In Section 4, we discuss how we use modularization in our meta–analysis model to obtain reasonable, if inconsistent, pseudo–posteriors. We demonstrate the utility of our modularization approach with synthetic data in Section 5. We then present the results from the meta–analysis of our actual data in Section 6, and we conclude with a brief discussion of our modeling approach in Section 7.

2 Specific Leaf Area

Specific leaf area (SLA) is defined as the surface area of a leaf (or leaves) divided by the dry weight of the leaf (or leaves), with units typically reported as cm^2/g. SLA is a plant functional trait that is of broad interest because it is a key parameter in mechanistic models of plant/tree growth (Sands and Landsberg 2002; Ogle and Pacala 2009), ecosystem dynamics (Harley and Baldocchi 1995; Duursma et al. 2005; Tatarinov and Cienciala 2006), and large–scale terrestrial vegetation dynamics (Bonan et al. 2003; Picard et al. 2005), and such models often are employed to explore the potential impacts of climate change on the terrestrial biosphere (Melillo et al. 1993; Cramer et al. 2001). Thus, obtaining accurate and precise estimates of SLA, and other plant functional traits, is necessary to improve our capacity to forecast impacts of climate change on vegetation and ecosystem processes (Wramneby et al. 2008; Tang and Zhuang 2009). Such vegetation models typically are highly complex with many parameters, and often exhibit identifiability problems due to data (e.g., carbon stocks or fluxes) that are insufficient to inform all parameters (e.g., Tang and Zhuang 2008). Thus, in addition to presenting our modularization and meta–analysis approaches, two additional objectives of this study are to utilize information from the literature to obtain "representative" SLA estimates for 305 tree species that occur in forests of the United States, and to understand how SLA varies across these 305 species.

2.1 Literature Search and Data Extraction

Journal articles reporting SLA, specific leaf mass (SLM), leaf mass per area (LMA), or specific leaf weight (SLW) were identified by searching Web of Science (ISI Web of Knowledge, Thompson–Reuters) using different search terms, including: 'SLA', 'SLM', 'LMA', 'SLW', 'specific and (leaf or leaves or needle*) and (mass or area)', and 'leaf mass area'. We only considered sources containing SLA and SLA–related data for species occurring in the continental U.S. according to the PLANTS Database (USDA NRCS 2008), and we specifically focused on the 305 species (i.e., 285 unique species and 20 subspecies within these species) identified by the U.S. Forest Service's Forest Inventory and Analysis (FIA) program (Miles et al. 2001). These species represent 85 genera, 40 families, 26 orders, and two divisions (Magnoliophyta and Coniferophyta).

After completing the initial literature search, we conducted a second, more detailed search to locate SLA information for species that are relatively abundant in the U.S., but that were poorly represented in the initial search. This was accomplished by using individual species' scientific and common names to locate additional articles containing information on these species and by browsing the articles for SLA information.

All SLA and related information extracted was entered into the TreeTraits database using Microsoft Access; the TreeTraits database is the North American counterpart to the FET (Functional Ecology of Trees) database (Kattge et al. 2011). SLA values were reported as sample means and converted when necessary to common units (cm^2/g). When available, important metadata and covariate information were entered into the TreeTraits database, including sample location, local climate details, light environment details, soil characteristics, tree age and crown class associated with the sampled tree, and sampling methods. When available, we recorded the sample size used to calculate each SLA sample mean, the type of samples used (single leaves, multiple leaves, or whole canopy), and estimates of uncertainty (i.e., standard errors or standard deviations).

2.2 Factors Affecting Specific Leaf Area

The light environment under which a leaf develops is known to affect its SLA, with shaded leaves generally having higher SLA than sun–exposed leaves (Chen et al. 1996; Chen 1997; Reich et al. 1998b; Koike et al. 2001). Thus, we created a light environment index based on categorical information such as crown class and canopy position (i.e., lower, mid, or upper canopy), both of which affect light availability (e.g., Ellsworth and Reich 1993; Koike et al. 2001; Chmura and Tjoelker 2008), and continuous variables such as light intensity and percent sun (see Appendix A in the on–line Supplemental Materials). Each SLA record was assigned one of the following indices: shade, sun, intermediate, or unknown (i.e., missing).

Also, we expected SLA to be correlated with tree age (Steele et al. 1989; Lusk and Warton 2007). Using reported information on actual tree age, tree age class (categorical), or tree size (i.e., stem diameter at breast height, DBH), we classified SLA records as representing seedlings, saplings, or mature trees.

Methods associated with measuring SLA could also affect the reported SLA values. For example, the sidedness of the leaf area measurement (e.g., one–sided or projected, versus two–sided area) must be accounted for because one– and two–sided SLA values differ by a factor of two. The type of sample collected may also be important such that bulk (or "canopy") samples likely contain petioles or non–photosynthetic twigs and are expected to have a lower SLA compared to samples only composed of leaf material. For studies that did not report tree age, sidedness, or sample type, these variables were entered into the TreeTraits database as unknown (missing).

2.3 Summary of Literature Search Results

The literature search returned 182 publications (studies), reporting a total of 1865 SLA sample means. These records provided SLA information for 158 species (154 unique species and 4 subspecies), and 90 of the records provided information at the level of the genus. For example, we only know that the record belongs to a pine, but we do not know the particular species of pine. The records represent species from 55 of the 85 genera, 20 of the 26 orders, and both divisions.

The degree of incomplete reporting depended on the variable of interest (Table 1). The covariate tree age class shows the highest degree of reporting (available for 94.1%

of the records). SLA sample mean standard errors, sample sizes, and the covariate light environment were associated with intermediate reporting (available for about 55%–58% of the records). The unit of observation and leaf sidedness were associated with a greater degree of incomplete reporting such that they were missing for over half of the records. Sampling method was available for only one–quarter of the records. Overall, 79.2% of the records were missing some form of information such as standard errors, sample sizes, sampling methods, or important covariates (Table 1).

Variable	Reported	Missing (%)
Standard error (se)	1031	44.7
Sample size (N)	1077	42.3
Unit of observation (U)	910	51.2
	(371 leaf, 103 leaves, 436 tree)	
Sampling method (C)	1368	26.6
	(1303 leaf–level, 65 canopy–level)	
Tree age class (A)	1755	5.9
	(658 mature, 404 sapling, 693 seedling)	
Light environment (L)	1072	42.5
	(443 sunny, 293 intermediate, 336 shady)	
Sidedness (S)	910	51.2
	(539 one–sided or projected, 371 two–sided)	
All variables	388	79.2

Table 1: Summary of SLA literature information reported for 1865 records. The number of records containing information on each variable is given, and the percent of records missing information is also provided. For the categorical variables, the number of records associated with each category is shown for those cases where the information was reported. The number of records containing information on all of the variables is also provided (i.e., these records represent cases of complete reporting).

3 Bayesian Meta–analysis

We define our meta–analysis model by considering the data models for the SLA sample mean, standard error, and sample size modules. Next we describe the mean model for the latent or "true" SLA values, which explicitly links latent SLA to covariates, species identity, and model parameters. We then consider the covariate modules. We end with a description of the parameter model, which includes a hierarchical prior for species–specific latent SLA based on species' taxonomic relationships, which form a natural hierarchy. The distribution functions described below are parameterized according to Appendix A in Gelman et al. (2004).

The dependency relationships between the different modules and stochastic quantities are illustrated in the DAG in Figure 2, which is modified from the typical DAG

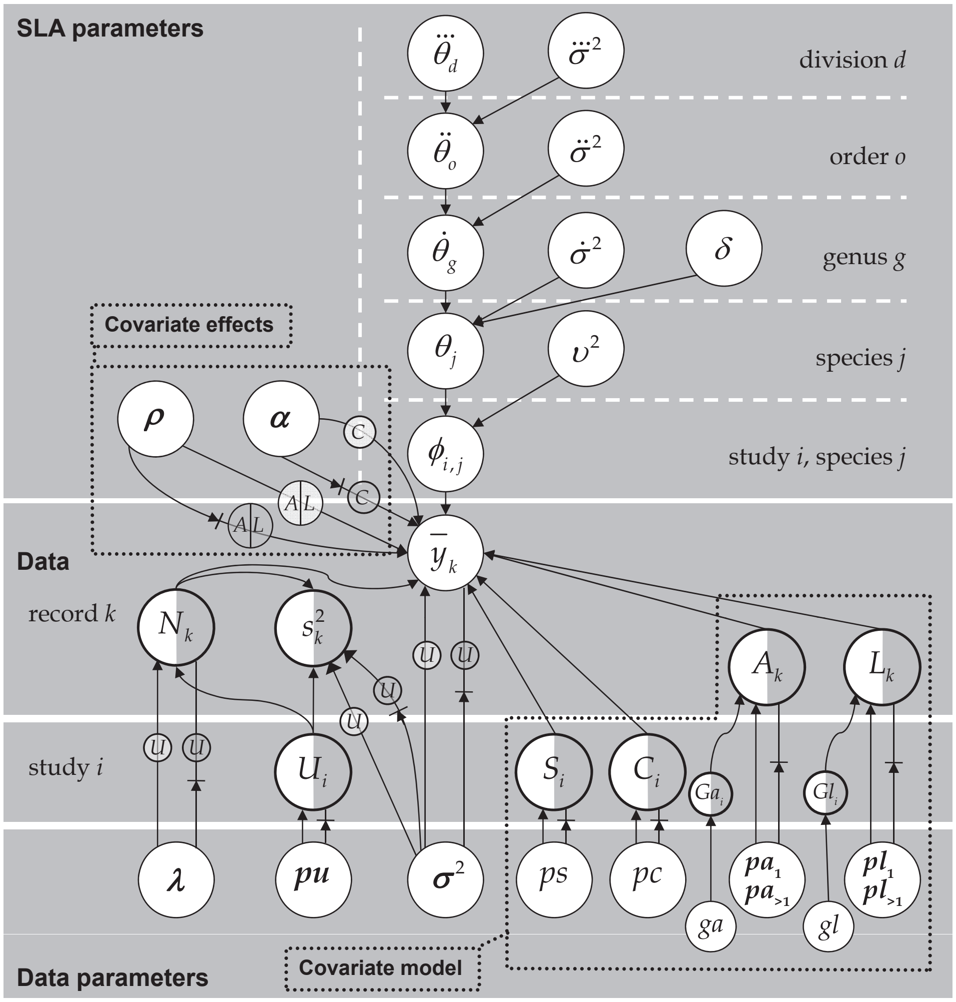

Figure 2: Graphical model illustrating the elements of the Bayesian meta–analysis model. Direct feedback control (Figure 1C) prevents missing covariates from affecting covariate model parameters (shown in this figure with dotted box labeled "Covariate model"), which are now informed only by the observed covariate data (white regions of covariate nodes), but the observed SLA mean ($\bar{y}$) is allowed to feed back to and inform the missing covariate data themselves (gray regions of covariate nodes). Indirect feedback control (Figure 1D) allows the covariate effects parameters (shown in this figure with dotted box labeled"Covariate effects" for α and ρ) to be informed only by records for which the associated covariates were reported. Circular nodes are stochastic quantities; square nodes are deterministic quantities.

notation to indicate positions of feedback control (discussed in Sections 1 and 4). We describe how we control the feedback between the missing data module and the parameter module to obtain a pseudo–posterior distribution of all unknown quantities (e.g., species– and genus–specific latent SLA, missing covariate data, light and age effects). Table 2 provides definitions of all model quantities.

3.1 SLA Data Model

Let $\bar{y}_k$ represent the SLA sample mean for record k, $k = 1, 2, \ldots, 1865$, and let $i(k)$ denote the study i, $i = 1, 2, \ldots, 182$, associated with record k. We assume $\bar{y}_k$ is normally distributed:

$$\bar{y}_k \sim Normal\left(\mu_k, \frac{\sigma^2_{U_{i(k)}}}{N_k}\right). \tag{1}$$

We assume that the $\bar{y}_k$ values are conditionally independent given the mean or latent ("true"), record–specific SLA (μ_k), the variance σ^2_U, and the sample size (N_k) used to compute $\bar{y}_k$. Note that σ^2_U depends on the unit of observation reported in each study, U_i, which indicates if $\bar{y}_k$ was computed by averaging SLA values across individual leaves ($U_i = 1$), multiple samples per tree with each containing multiple leaves ($U_i = 2$), or samples representative of an entire tree's canopy ($U_i = 3$). We define the models for N_k, μ_k, U_i, and σ^2_U in expressions (3), (4), and (5), and in Section 3.5, respectively.

3.2 Standard Error and Sample Size Models

The variance terms in expression (1) are partly informed by the record–specific standard errors (se_k) and sample sizes (N_k). (An assumption of normality at the observation level along with expression (1) compels us to consider reported standard errors by sufficiency.) Missing values caused identifiability problems between each se and N and their associated parameters. We elaborate on this in Section 4 with our discussion of feedback control involving SLA covariates, which is treated in a manner similar to that of se and N. Here, we describe the "complete" models for se and N, observed or not.

Most existing meta–analysis approaches treat se as a fixed quantity, discarding records without it. More recently, some have explored methods for "imputing" missing se information, but the general approach is to apply a relatively *ad hoc* method to obtain point estimates of se (or the standard deviation) that are treated as fixed values in subsequent meta–analyses (Abrams et al. 2005; Furukawa et al. 2006; Wiebe et al. 2006), thereby ignoring uncertainty in these estimates. By explicitly modeling se, we thereby use available information in observed values while accounting for uncertainties in the reported and missing estimates of se. If we assume further that individual observations of SLA are normally distributed, then basic results (Casella and Berger 2002) lead to a gamma distribution for $se^2_k = s^2_k/N_k$, conditional on N_k:

$$\frac{s^2_k}{N_k} \sim Gamma\left(\frac{N_k - 1}{2}, \frac{N_k - 1}{2} \cdot \frac{N_k}{\sigma^2_{U_{i(k)}}}\right). \tag{2}$$

Symbol	Type of Quantity
	Indices or subscripts
k	Record (“observation”) indicator
i	Study indicator
j	Species indicator
g	Genus indicator
o	Order indicator
d	Division indicator
	Data or data–related quantities
$\bar{y}$	SLA sample mean (cm^2/g)
A	Tree age class (1=mature, 2=sapling, 3=seedling)
C	Sampling method (1=leaf–level vs. 2=canopy–level)
Ga	Age class group (0=measured more than one age class; 1=measured only one age calss)
Gl	Light class group (0=measured more than one light class; 1=measured only one light class)
L	Light class (1=sunny, 2=intermediate, 3=shaded)
N	Sample size used to compute $\bar{y}$
S	Sidedness of leaf area (one–sided/projected vs. two-/all–sided)
se	Standard error of mean SLA ($\bar{y}$)
U	Unit of observation (1=leaf, 2=leaves, 3=tree)
	Data parameters
σ^2	Variance associated with SLA measurements
λ	Poisson rate parameter in sample size likelihood
pa_1	Probability vector corresponding to age classes (A) for studies measuring only one age class
$pa_{>1}$	Probability vector corresponding to age classes (A) for studies measuring more than one age class
pl_1	Probability vector corresponding to light classes (L) for studies measuring only one light class
$pl_{>1}$	Probability vector corresponding to light classes (L) for studies measuring more than one light class
ga	Probability of measuring only one age class
gl	Probability of measuring only one light class
pc	Probability of canopy–level SLA
ps	Probability of two–sided leaf area
pu	Probability vector for unit of observations
	Latent process and process parameters
μ	Latent or mean SLA (cm^2/g)
ϕ	Study by species one–sided latent SLA for mature trees in the sun (cm^2/g)
α	Age and light effects
ρ	Sampling method effect
δ	Evergreen effect for species in Magnoliophyta
θ, $\dot{\theta}$, $\ddot{\theta}$, $\dddot{\theta}$	Latent SLA (one–sided, mature trees, sun) for each species, genera, order, and division, respectively
ν^2	Study within species variance in latent SLA
$\dot{\sigma}^2$, $\ddot{\sigma}^2$, $\dddot{\sigma}^2$	Between species (within genus), between genus (within order), between order (within division) variance in latent SLA

Table 2: List of all quantities in the hierarchical Bayesian meta–analysis model for specific leaf area.

Expression (2) also applies to the small number of studies that reported a sample standard deviation (s) instead of se, in which case, the second N_k in the second argument of the gamma pdf specification in (2) is replaced with 1.

In contrast to existing meta–analysis approaches, we also model sample size, N, stochastically. We assume at least two observations were made to compute $\bar{y}_k$ and se_k (i.e., $N_k \geqslant 2$, which was true for all reported N_k), and since N_k is integer–valued with no known upper bound, we assume a Poisson distribution for $N_k - 2$:

$$N_k - 2 \sim Poisson\left(\lambda_{U_{i(k)}}\right). \tag{3}$$

The Poisson rate parameter, λ, can take on one of three values (λ_1, λ_2, λ_3) depending on the unit of observation reported for study i (i.e., U_i). We define the model for λ in Section 3.5.

3.3 SLA Mean Model

Before describing models of the covariates, we define a fairly simple, empirical model for the latent, record–specific SLA, μ_k, in expression (1), which provides context for the covariate models. The model incorporates covariates that vary at the level of record k, such as light environment L_k ($L = 1$ for sun, 2 for intermediate, 3 for shaded), tree age class A_k ($A = 1$ for mature, 2 for sapling, 3 for seedling), and species identity j, $j = 1, 2, \ldots, 305$, associated with record k. We also incorporate covariates that vary at the level of study i, such as leaf sidedness S_i ($S = 1$ for one–sided or projected, 2 for two–sided) and sampling method C_i ($C = 1$ for leaf–level, 2 for canopy–level). We implemented the following multiplicative model:

$$\mu_k = \phi_{(i,j)(k)} \cdot \alpha_{A_k,L_k} \cdot \rho_{C_{i(k)}} \cdot S_{i(k)}. \tag{4}$$

The $(i, j)(k)$ notation indexes study i and species j associated with record k. We use cell–reference coding for the matrix of age by light fixed effects, α, and for the vector of sampling method fixed effects, ρ. That is, for $A = 1$ (mature) and $L = 1$ (sunny), we set $\alpha_{1,1} = 1$, and for $C = 1$ (leaves), we set $\rho_1 = 1$. Thus, $\phi_{(i,j)(k)}$ is the latent, one–sided SLA for leaves from mature trees growing in a sunny environment for study i and species j. Multiplying by S_i converts the latent one–sided SLA to two–sided if study i measured two–sided area (i.e., S_i= 2).

We also conducted preliminary analyses with a linear model on the log scale, assigning normal priors to $\log(\phi_{(i,j)(k)})$, but this resulted in inferior model fit and unrealistic species–specific SLA (i.e., θ, see expression 11) estimates whose 97.5^{th} percentiles were an order of magnitude higher than expected. Expression (4) gives superior fit, and its multiplicative effects for light, age and canopy satisfy the ecologically reasonable expectation of comparatively larger absolute effects on SLA for larger leaves. We discuss the modeling of ϕ, α, and ρ in Section 3.5, and the models for A, L, C, and S are defined below in Section 3.4.

3.4 Covariate Data Models

We begin with the study–level covariates S, U and C. The unit of observation U is a categorical variable with support $\{1,2,3\}$, and we assume the following probability mass function (pmf) with global probability vector pu for the probability that study i measured individual leaves, multiple leaves, or trees' canopies:

$$P\left(U_i = c\right) = \begin{cases} pu_c & c = 1,2,3 \\ 0 & otherwise. \end{cases} \tag{5}$$

Sidedness S and sampling method C are binary variables, and we assume they are described by Bernoulli distributions with global probability parameters ps (probability of a two–sided measurement) and pc (probability of a canopy–level measurement), respectively. We also assume that the sampling method for "leaf" ($U_i = 1$) directly corresponds to leaf–level measurements ($C_i = 1$), thus:

$$S_i \sim Bern\left(ps\right) \tag{6a}$$

$$C_i \sim Bern\left(pc \cdot \mathrm{I}_{U_i \neq 1} + \left(1 - \mathrm{I}_{U_i \neq 1}\right)\right). \tag{6b}$$

The indicator function, $\mathrm{I}_{U_i \neq 1}$, returns 1 if $U_i = 2$ or 3 and 0 if $U_i = 1$.

The age and light covariates, A and L, both are categorical variables with support $\{1,2,3\}$. The data suggested that studies fell into one of two measurement strategy groups: those that measured SLA for only one A or L category ($Ga = Gl = 1$) and those that measured SLA for more than one category ($Ga = Gl = 0$). Similar to the models for the study–level S and C, we assume $Ga_i \sim Bern(ga)$ and $Gl_i \sim Bern(gl)$, where ga and gl are the probabilities that a study reports data for a single A or L category, respectively. If $Ga_i = 1$ ($Gl_i = 1$), then A (L) is modeled at the level of study i; conversely, if $Ga_i = 0$ ($Gl_i = 0$), then A (L) is modeled at the level of record k because records within a study can be associated with different age (light) categories. Thus, we specify the following pmfs for A and L:

$$P\left(A_i = c \,|\, Ga_i = 1\right) = \begin{cases} pa_{1c} & c = 1,2,3 \\ 0 & otherwise \end{cases} \tag{7a}$$

$$P\left(A_k = c \,|\, Ga_{i(k)} = 0\right) = \begin{cases} pa_{>1c} & c = 1,2,3 \\ 0 & otherwise \end{cases} \tag{7b}$$

$$P\left(L_i = c \,|\, Gl_i = 1\right) = \begin{cases} pl_{1c} & c = 1,2,3 \\ 0 & otherwise \end{cases} \tag{8a}$$

$$P\left(L_k = c \,|\, Gl_{i(k)} = 0\right) = \begin{cases} pl_{>1c} & c = 1,2,3 \\ 0 & otherwise. \end{cases} \tag{8b}$$

3.5 Parameter Models (Priors)

Here we give stochastic specifications for the model parameters (see Figure 2). We begin with the parameters in the SLA mean model (i.e., $\phi_{i,j}$, α, ρ), then discuss the parameters in the covariate data models (i.e., pu, pa_1, $pa_{>1}$, pl_1, $pl_{>1}$, ps, pc, ga, gl), and conclude with the priors for the remaining hyperparameters and global parameters (all other root nodes in the DAG; Figure 2).

First we define the models for the parameters in the SLA mean model (i.e., ϕ, α, ρ). Recall that $\phi_{i,j}$ is the latent, one–sided SLA value for sun–exposed leaves from mature trees of species j reported in study i. We are not interested in making inferences about SLA for every species–study combination. Rather, we wish to obtain estimates for the 305 tree species, after accounting for the study effect. For the study–within–species effects, we assume:

$$\phi_{i,j} \sim Normal\left(\theta_j, \upsilon^2\right). \tag{9}$$

Recall that SLA sample means were only reported for 158 of the 305 species. The taxonomic relationships provide a natural hierarchy for modeling the species–level parameters (θ_j). Thus, we incorporate taxonomic relationships among species (within genus), genera (within order), and orders (within division), thereby borrowing strength for poorly represented taxa (we ignore the family level since 23 of the 26 orders contained only 1 or 2 families). The use and definition of these different levels (species, genus, order, division) is based on the well–established nomenclature rules developed by Carl Linneaus in the 16^{th} century, and the hierarchical taxonomic relationships are depicted in Figure 2.

Our model assumes that each species–specific θ_j comes from a parent population defined by a normal distribution whose mean depends on genus g, $g = 1, 2, \ldots, 86$:

$$\theta_j \sim Normal([I_{j \notin (evgrn\ \&\ magnoli)} + \delta I_{j \in (evgrn\ \&\ magnoli)}]\dot{\theta}_{g(j)}, \dot{\sigma}^2). \tag{10}$$

The notation $g(j)$ indexes genus g associated with species j, and the indicator function $I_{j \in (evgrn\ \&\ magnoli)}$ returns 1 if species j is evergreen ($evgrn$) and from the Magnoliophyta division ($magnoli$), 0 otherwise. That is, $\dot{\theta}_{g(j)}$ is treated as the "base–line condition," which represents deciduous leaves in the Magnoliphyta and evergreen leaves in the Coniferophyta. (For notational convenience, we use overset dots on θ to indicate different θ's associated with the different taxonomic levels.) For practical purposes, we treat all Coniferophyta species as evergreen because there are only a few deciduous species, all of which occur in the genus Larix (larches). Thus, for evergreen species within Magnoliophyta, the mean is given by $\delta \cdot \dot{\theta}_{g(j)}$, and, for all other species, the mean is given by $\dot{\theta}_{g(j)}$.

To complete the hierarchical, taxonomic prior, we assume that $\dot{\theta}_g$ comes from a normal distribution whose mean depends on order o, $o = 1, 2, \ldots, 26$, associated with genus g; the order effects are assumed to come from a normal distribution whose mean depends on division d, $d = 1$ or 2 for Magnoliophyta and Coniferophyta, respectively, associated with order o:

$$\dot{\theta}_g \sim Normal\left(\ddot{\theta}_{o(g)}, \ddot{\sigma}^2\right). \tag{11}$$

$$\ddot{\theta}_o \sim Normal\left(\dddot{\theta}_{d(o)}, \dddot{\sigma}^2\right). \tag{12}$$

We treat the division–level mean parameters ($\dddot{\theta}_d$) as root nodes and give them diffuse, independent normal priors such that $\dddot{\theta}_1, \dddot{\theta}_2 \sim Normal(0, 1000)$.

Priors were specified for the standard deviation terms following Gelman (2006) where we assigned a $Uniform(0, 200)$ to υ in expression (9) because there are many study–by–species combinations (large "group size"), and we assigned relatively diffuse, independent folded Cauchy (i.e., $|t(0, 200^2, 1)|$) priors to $\dot{\sigma}$, $\ddot{\sigma}$, and $\dddot{\sigma}$ in expressions (10), (11), and (12).

The age–light (α), sampling method (ρ), and evergreen (δ) effects are multiplicative (see expression (12)), and we require $\alpha_{A,L} > 0$, $\rho_C > 0$, and $\delta > 0$. We specified wide, independent uniform priors for each such that $\alpha_{1,2}, \alpha_{1,3}, , \alpha_{3,3}, \rho_2, \delta \sim Uniform(0, 10)$; recall $\alpha_{1,1} = 1$ and $\rho_1 = 1$, by cell reference coding.

Next, we assign relatively noninformative, conditionally conjugate $Dirichlet([1, 1, 1])$ priors to pu, pa_1, $pa_{>1}$, pl_1, and $pl_{>1}$ associated with the U, A, and L likelihoods in expressions (5), (7), and (8). Likewise, we assign $Uniform(0, 1)$ priors to the probability parameters ps, pc, ga, and gl associated with expressions (6), (7), and (8). Finally, we assigned $Uniform(0, 200)$ priors to the standard deviations (σ_U) in expressions (1) and (2) and the Poisson rate parameters (λ_U) in expression (3) for $U = 1, 2, 3$.

4 Controlling Feedback with Modularization

4.1 Direct Feedback Control: Modularization of the Covariate Model

The specifications above lead to a full probability model for the joint posterior distribution of the missing data, latent variables, and model parameters. The full probability model allows the SLA sample means ($\bar{y}$), standard errors (se), sample sizes (N), and covariate data (A, C, L, S, U) to inform all parameters in the posterior in the usual manner, including the covariate model parameters (e.g., pu, ps, pc, etc.) (see Figure 2 for model components). However, given the large amount of missing data for se, N, and the covariates, the SLA sample means ($\bar{y}$) can potentially overwhelm the posteriors for the covariate parameters and missing data. That is, the large number of records for $\bar{y}$ could result in posteriors for the covariate parameters that are notably different than what one would obtain from the covariate data alone.

The influence of our "response" model of $\bar{y}$ on the covariate model component is analogous to the situation described by Lunn et al. (2009a) wherein pharmacodynamic (PD) response data (EEG measurements) "feed back" to influence, perhaps undesirably, the pharmacokinetic (PK) covariate model component (plasma midazolam concentrations). The example here is different from that of Lunn et al. (2009a) because it involves missing covariates rather than latent truth underlying values observed with error, but the effect of feedback can be similarly undesirable. In fact, this was the case here, and we encountered other problems upon applying Bayes theorem to the full probability (FP) model, without feedback control, in WinBUGS and OpenBUGS. For example, the MCMC chains exhibited poor mixing, they did not converge after running the model for many hundreds of thousands of iterations, and the chains for some variance components were exploring unrealistic regions of the parameter space.

The poor mixing problems appear to reflect problems of nonidentifiability between the covariate and study effects parameters (α, ρ, ϕ) and missing covariate values. This poor mixing is similar to that reported by Liu et al. (2009) in their fully Bayesian implementation of a Gaussian process response surface emulation of complex computer models (e.g., Sacks et al. 1989). As in Liu et al. (2009), the mixing is so poor for our FP implementation that we are unable to declare convergence with any certainty. However, we speculate that confounding between the covariate effects, study effects, variance parameters, and imputed covariates caused weak identifiability among these components to contribute to the poor mixing. Without convergence, this is difficult to verify or to otherwise assess a problem. It is possible that the model components are approximately correct and that we may fix the mixing problem by using additional data or informative priors. We have neither. And, if the models are bad, we have little information to suggest how to remodel.

To proceed with the analysis, we choose to cut feedback between the SLA mean model for $\bar{y}$ and the covariate model (see “Covariate model” box in Figure 2). In doing so, consistency between local specifications and a joint posterior is no longer guaranteed and, strictly speaking, we step beyond the Bayesian paradigm. But, as we have argued, and as Liu et al. (2009) say, “computation is otherwise not possible.” The modularization of the covariate model eliminated feedback from the $\bar{y}$ data to the covariate model parameters via the missing covariates. These results provide insight into additional justification for cutting feedback beyond simply getting the model to run, as we discuss below. Now, the $\bar{y}$ data do not influence the covariate model parameters, and $P(pa, pl, pc, ps, pu, \lambda, \sigma^2, \ldots)$ is a “distributional constant” (Lunn et al. 2009a) with respect to the $\bar{y}$ model.

We illustrate the differences between the posteriors associated with the Bayesian FP model and the pseudo–posteriors associated with the model that implements feedback control (FC) by modularization of the covariate model, thereby eliminating feedback from the $\bar{y}$ data to the covariate model parameters (see Figure 2). We show these differences for a subset of parameters in the model by giving their full conditionals under the FP model and under the FC model. For example, let the covariate parameter p represent ps, pc, or pu, and let the missing and observed covariates (S, C, or U) be represented by $Zmiss$ and $Zobs$, respectively. The full conditional distribution for p under the FP model is given by:

$$P(p|...) \propto P(Zobs|p) \cdot P(Zmiss|p) \cdot P(p). \tag{13}$$

Severing the feedback through $Zmiss$ to p results in the following full conditional for p under the FC model:

$$P(p|...) \propto P(Zobs|p) \cdot P(p). \tag{14}$$

Expressions (13) and (14) demonstrate the effect of eliminating the feedback of the $\bar{y}$ data to the covariate parameters via the missing covariate data.

Note that the pseudo–posteriors for the covariate parameters (e.g., p as above) should be interpreted as providing estimates of the parameters corresponding to the observed data. If the data are missing at random, then the parameter estimates should represent the true covariate parameters. If the data are not missing at random, however, then the parameter estimates will likely be biased for the true parameters. If we wish to obtain estimates of the true parameters, then we would need to consider the (pseudo) posteriors that are informed by both the missing and observed data components. Although not the focus of this paper, we briefly describe how we obtained these estimates in Section 4.3.

A more extreme approach would cut feedback from the $\bar{y}$ model to the imputed missing values ($Zmiss$) as in Lunn et al. (2009a) and Molitor et al. (2009), in which case $P(Zmiss)$ is a distributional constant with respect to the $\bar{y}$ model. Our choice described above may be seen as a compromise between the FP approach and the latter approach since the $\bar{y}$ data are allowed to influence imputed covariate values ($Zmiss$) but not the parameters of the covariate model (p) because these parameters are not affected by the missing covariates. (We tried the latter, more extreme case, but it, too, suffered from notable mixing problems.) Incidentally, we note a similarity to traditional data–assimilation (e.g., Wang et al. 2000) wherein data are allowed to optimally adjust states—analogous to missing covariates here—but are not allowed to influence the parameters of the state equation, which are often fixed—actual constants—in the assimilation procedure. Thus, as data–assimilation procedures move beyond the tradition of fixing parameters in the state model, modularization methodology seems to become more relevant for data assimilation.

In the above models, covariates and covariate parameters are *directly* connected (see the discussion of Figure 1 in Section 1) such that the covariates are child nodes of parent covariate parameters. Thus, we refer to the feedback control described by expressions (13) and (14) as direct feedback control. Eliminating the direct feedback between the $\bar{y}$ data (or imputed covariates) and the covariate parameters did not, however, solve the poor mixing problem. We suspect that this is due to a persistent identifiability problem between the missing covariates and the covariate effects (e.g., α, ρ). For example, since this model still allows the missing covariates to affect α and ρ, one can imagine the situation where we are missing, for example, sidedness (S) information for many studies. Thus, if most of the imputed S values are equal to 2, then α (or ρ) would decrease by a factor of 2 (see expression (4)) compared to the situation where the imputed S values are equal to 1. Thus, α could vary by at least a factor of 2 depending on the current state of the imputed S values, and many of the MCMC chains appeared to get stuck in a particular state.

4.2 Indirect Feedback Control: Modularization of the Covariate Effects

To avoid unwanted influence of the imputed covariates on the covariate effects (α, ρ), we also severed the feedback between the imputed covariate data and α and ρ (see "Covariate effects" box in Figure 2). We refer to this as *indirect* feedback control because the covariates (e.g., A, L, C) and the covariate effects (α, ρ) do not share an edge in the DAG (Figure 2), but they both appear in the SLA mean model (both share edges with $\bar{y}$), thereby indirectly influencing each other. For example, even though C and ρ do not share an edge, we do not allow the missing values for C ($Cmiss$) to affect ρ.

To illustrate the effect of the indirect feedback control, we again compare the full conditionals for the Bayesian FP model with those from the indirect FC model. First, we consider the full conditional for α. We partition $\bar{y}$ into two vectors, $\bar{y}'$ and $\bar{y}''$, where $\bar{y}'$ contains the $\bar{y}_k$ where A and L are observed, and $\bar{y}''$ contains the remaining $\bar{y}_k$ associated with records missing A or L. Under the FP model, the full conditional for α is:

$$\begin{aligned} P(\alpha|\cdots) \propto\ & P(\bar{y}'|\alpha,\rho,Aobs,Lobs,N,S,C,\sigma^2,\phi) \\ & \times P(\bar{y}''|\alpha,\rho,Amiss,Lmiss,N,S,C,\sigma^2,\phi)P(\alpha). \end{aligned} \tag{15}$$

Under indirect FC, however, the full conditional for α does not depend on the likelihood of $\bar{y}''$ because the records associated with $\bar{y}''$are missing the covariates A or L such that:

$$P(\alpha|\cdots) \propto P(\bar{y}'|\alpha,\rho,Aobs,Lobs,N,S,C,\sigma^2,\phi)P(\alpha). \tag{16}$$

Note, however, that the full conditionals in expression (16) are conditional on both observed and missing N, S, and C associated with the records defining $\bar{y}'$.

Now consider the full conditional for ρ, which describes the effect of C, and partition $\bar{y}$ into $\bar{y}^*$ and $\bar{y}^{**}$. Here, $\bar{y}^*$ contains the $\bar{y}_k$ where C is observed ($Cobs$) and $\bar{y}^{**}$ contains the remaining $\bar{y}_k$ associated with missing C values ($Cmiss$). Under the FP model, the full conditional for ρ is:

$$\begin{aligned} P(\rho|...) \propto\ & P(\bar{y}^*|\alpha,\rho,A,L,N,S,Cobs,\sigma^2,\phi) \\ & \times P(\bar{y}^{**}|\alpha,\rho,A,L,N,S,Cmiss,\sigma^2,\phi)P(\rho). \end{aligned} \tag{17}$$

Similar to expression (16), the full conditional for ρ under indirect FC is:

$$P(\rho|...) \propto P(\bar{y}^*|\alpha,\rho,A,L,N,S,Cobs,\sigma^2,\phi)P(\rho). \tag{18}$$

Again, the full conditional in expression (18) is conditional on both observed and missing A, L, N, and S associated with the $\bar{y}^*$ records.

It should be clear that all "observations" of $\bar{y}$ are used to inform the parameters of most interest, that is, the latent SLA values (e.g., θ, $\dot{\theta}$,$\ddot{\theta}$ and $\dddot{\theta}$). That is, the full conditionals for these parameters in the FC model are analogous to the standard full conditionals that we would derive for the Bayesian FP model.

The above FC model, described by expressions (14), (16), and (18), defines our final model of choice (as in Figure 2). The combination of direct and indirect feedback control solved the identifiability and mixing problems associated with the other models, and resulted in realistic parameter estimates. We note that the direct feedback control is described in Lunn et al. (2009a) and Molitor et al. (2009), but we are not aware of any studies that implement such indirect feedback control. In summary, in our final FC model, the $\bar{y}$ data do not feed back to affect the covariate parameters (pa_1, $pa_{>1}$, pu, ps, λ, etc.) but they do affect the missing covariate data ($Amiss$, $Lmiss$, $Cmiss$, etc.) (direct FC); and, the missing covariate data do not influence the covariate effect parameters (α and ρ) in the SLA mean model (indirect FC) (see Figure 2).

4.3 Effect of Modularization on Covariate Parameters

As a simple model diagnostic to evaluate the impact of the feedback control on the parameter estimates, we evaluated potential discrepancies between the observed and imputed covariate data with respect to covariate parameter estimates. That is, we implemented the covariate–only model (see "Covariate model" box in Figure 2, but without cutting feedback), with the observed covariate data and with the missing covariate data filled–in with the maximum (pseudo–) a posteriori (MAP) estimates obtained from the FC model described above. That is, we used a completed set of covariate data. We then implemented the covariate–only model to obtain its Bayesian posterior distributions for the covariate parameters associated with the observed data and for those associated with the imputed (MAP) data. That is, for each covariate Z, we created an indicator variable x^Z such that $x^Z = 1$ if Z was originally reported and $x^Z = 2$ if Z was originally missing but now replaced with the MAP estimate from the above FC model.

Omitting details of the complete covariate–only model, we describe the module for S as an example, where we modified expression (6) and the prior for ps such that:

$$\begin{aligned} S_i &\sim Bern\left(ps_{x_i^S}\right) \\ ps_x &\sim Uniform(0,1) \quad x = 1, 2. \end{aligned} \tag{19}$$

Similar modifications were made to the se, N, L, U, A, and C models whereby all covariate parameters and hyperparameters were modeled separately for the observed ($x^Z = 1$) and the imputed ($x^Z = 2$) data. Again, this model did not involve feedback control, there was no missing data since all original missing values were replaced with the MAP estimates, and it did not involve the $\bar{y}$ data. The results from this analysis are summarized in Section 6.5.

As noted in Section 4.1, if the covariate data are not missing at random, then neither ps_1 (equivalently, ps in expression 8) or ps_2 reflect the "true" ps, and the same applies to the other covariate parameters. If a goal of such a meta–analysis is to learn about the true covariate parameters, then one would want to use both the imputed and observed data simultaneously to obtain the posterior distributions of the parameters, independent of the observation status (observed vs. missing) of the covariate data. We did this within the FC model by essentially creating a second likelihood for each "complete" covariate dataset whereby at each posterior draw, the missing data were "replaced" with the imputed data, creating the complete dataset Z^{comp}. Returning to the sidedness example, let S^{comp} represent the complete dataset, and we assigned a likelihood to S^{comp} analogous to expression (8):

$$S_i^{comp} \sim Bern\left(ps^{true}\right). \tag{20}$$

Here, ps^{true} is assigned a $Uniform(0,1)$ prior, and it represents the true covariate parameter, which is informed by both the observed and imputed (missing) covariate data. Thus, if a goal is to learn about the probability of measuring one– vs two–sided leaf area, then one would obtain the posterior distribution for ps^{true}.

4.4 Implementation

The Bayesian meta–analysis model with feedback control (FC model), as described above, was coded in the OpenBUGS program version 3.0.3 (Lunn et al. 2009b), and the code is provided in Appendix B in the on–line Supplemental Materials. We used the "cut" function to implement feedback control described above (e.g., Spiegelhalter et al. 2003; Jackson et al. 2009; Lunn et al. 2009a). We employed the "ones trick" (see WinBUGS manual, Spiegelhalter et al. 2003) to assign the likelihood to S^{comp} in expression (20). Starting values were obtained by conducting preliminary analyses that separately implemented individual model components (e.g., covariate module, standard error and sample size module, etc.). Relatively dispersed starting values were used for running the modules, and once each module approximately converged, we used the final MCMC samples to initialize the full FC model. We ran three parallel MCMC chains for 110,000 iterations per chain. Chains were monitored and evaluated for convergence to the pseudo–posterior using the Brooks–Gelman–Rubin (BGR) diagnostic tool in OpenBUGS (Brooks and Roberts 1998; Gelman et al. 2004). We discarded iterations associated with the "burn–in" period (first 40,000), and post burn–in samples were thinned every 5 iterations to reduce storage requirements, which also had the effect of reducing autocorrelation within each chain, resulting in a total sample of about 42,000 from the posterior distribution.

5 Evaluation of Modularization with Synthetic Data

Before we present the results for the SLA data and model described in Sections 2–4, we highlight a simulation experiment that demonstrates the utility of our proposed modularization approach. We created a synthetic SLA dataset by assuming a simplified version of the above model (Section 3). We used the reported age (A), light environment (L), unit of observation (U), and sidedness (S) covariate data. We filled in missing study–level S data with $S = 1$ (one–sided), and we filled in each missing record–level A and L data and study–level U data by generating a categorical value (1, 2, or 3) according to the proportions of each category that occurred in the observed data. Likewise, we generated values for the missing sample size (N) and standard error (se) data according to the distribution of observed N and se data. Given these complete data, we then generated record–level SLA values (i.e., $\bar{y}$) according to expression (1) and a simplified mean model, $\mu_k = \theta_{j(k)} \cdot \alpha_{A_k,L_k} \cdot S_{i(k)}$. Thus, for the synthetic data, we do not incorporate a sampling method (C) effect (ρ) or study by species random effects (ϕ). To generate the synthetic $\bar{y}$ data, we specified "true" values for θ and α that were similar to the posterior means we obtained from the analysis of the real SLA data. Given this complete dataset, we created an incomplete dataset by "removing" A, L, U, S, N, and se data according to the missingness pattern in the real SLA dataset.

Next, we implemented four modeling approaches with the synthetic data. The first two specified the full probability (FP) model (i.e., without feedback control [FC]); one was applied to the complete synthetic dataset (no missing data), and the other to the incomplete synthetic dataset. Fitting the FP model to the complete dataset allowed us to estimate the "true" values for parameters that were not directly specified during the data generation process, such as means and variances in the taxonomic hierarchy (e.g., see expressions (10)–(12)). The other two approaches applied different modularization specifications to the incomplete synthetic dataset. One modularization approach only involved direct FC within the covariate models; the other approach implemented both direct and indirect FC (as in Figure 2). All approaches involved the model components defined in expressions (1)–(3), (5), (6a), (10), (11), and (12). Expression (4) was modified as described above (i.e., $\mu_k = \theta_{j(k)} \cdot \alpha_{A_k,L_k} \cdot S_{i(k)}$), and the models for A and L in expressions (7) and (8) were simplified such that the likelihood of each data component involved a global probability vector (e.g., pa and pl), analogous to the likelihood for U in expression (5). Expressions (6b) and (9) were not employed because they define the models for sampling method and study random effects. The four models were implemented in OpenBUGS following a procedure similar to that described in Section 4.4.

Despite trying a variety of different values to initiate the MCMC algorithm, chains for some of the model parameters, including the deviance, did not converge for the FP model when applied to the incomplete dataset. Similar lack of convergence was encountered with the direct FC model. The most problematic parameters were those associated with the se and N likelihoods (i.e., σ^2_U and λ_U), which have a direct effect on the deviance via the likelihood for $\bar{y}$. Conversely, the chains converged for all parameters when the FP model was applied to the complete synthetic dataset and when both indirect and direct FC were used with the incomplete dataset.

To compare the posterior parameter estimates obtained from the four different modeling approaches, we computed posterior statistics based on the chain that achieved the lowest deviance for the FP and direct FC models (incomplete dataset). All three chains were used to compute posterior statistics for the other two models. The FP approach and the two modularization approaches yielded very similar posterior parameter estimates that agreed with the true values for most parameters. However, there was a subset of parameters for which one or more modeling approaches yielded posterior estimates that did not agree with the true values (see Table 3). In general, the number

Parameter	Truth	No Control (FP)	Direct FC	Direct & Indirect FC
$\dot{\sigma}$	2.5	**6.3**	**3.9**	2.5
		(5.4, 7.3)	**(3.3, 4.6)**	(2.1, 3.0)
$\ddot{\sigma}$	15.5	17.5	16.8	15.8
		(13.1, 22.8)	(12.6, 21.8)	(12.3, 20.3)
$\dddot{\sigma}$	6.1	8.4	10.6	7.3
		(1.4, 19.8)	(1.0, 24.8)	(0.2, 18.9)
$\sigma_{U=1}$	47.5	**39.9**	**39.6**	47.7
		(39.4, 40.5)	**(39.1, 40.2)**	(46.6, 48.2)
$\sigma_{U=2}$	55.4	**133.8**	56.6	55.1
		(132.5, 135.1)	(55.3, 57.7)	(54.0, 56.2)
$\sigma_{U=3}$	71.1	**110.4**	**108.0**	**106.6**
		(109.1, 111.7)	**(106.7, 109.2)**	**(105.2, 108.0)**
$\lambda_{U=1}$	19.2	19.5	**15.4**	19.1
		(19.1, 19.9)	**(15.0, 15.8)**	(18.7, 19.6)
$\lambda_{U=2}$	48.1	**200.0**	**51.8**	48.5
		(200.0, 200.0)	**(50.7, 53.0)**	(47.1, 49.9)
$\lambda_{U=3}$	79.4	**43.2**	**32.0**	**32.0**
		(42.6, 43.9)	**(31.4, 32.5)**	**(31.4, 32.3)**
$\theta_{j=214}$	98.2	**51.9**	99.1	99.2
		(50.1, 53.8)	(96.0, 102.3)	(95.9, 102.5)
$\theta_{j=277}$	97.3	**70.2**	**93.1**	96.7
		(68.0, 72.4)	**(90.0, 96.1)**	(92.5, 101.3)
$\theta_{j=278}$	99.0	97.7	**58.9**	97.3
		(93.2, 102.1)	**(56.2, 61.5)**	(92.7, 101.8)
$\theta_{j=298}$	115.9	116.0	**74.0**	100.7
		(112.1, 120.1)	**(59.0, 78.7)**	(71.1, 118.6)
$pu_{c=1}$	0.15	**0.29**	**0.52**	**0.51**
		(0.23, 0.36)	**(0.45, 0.59)**	**(0.44, 0.59)**
$pu_{c=2}$	0.10	**0.38**	**0.15**	0.14
		(0.31, 0.45)	**(0.10, 0.20)**	(0.09, 0.20)
$pu_{c=3}$	0.76	**0.33**	**0.33**	**0.34**
		(0.26, 0.40)	**(0.27, 0.40)**	**(0.28, 0.42)**
$pl_{c=1}$	0.66	**0.63**	0.65	0.65
		(0.60, 0.65)	(0.63, 0.67)	(0.63, 0.68)
$pl_{c=2}$	0.16	0.18	0.17	0.17
		(0.16, 0.20)	(0.15, 0.18)	(0.15, 0.18)
$pl_{c=3}$	0.18	0.19	0.18	0.18
		(0.17, 0.21)	(0.17, 0.20)	(0.16, 0.20)

Table 3: Posterior means and central 95% credible intervals (CI) for a synthetic SLA dataset. Bolded results do not agree with the true values (i.e., 95% CI does not contain the truth). Results are shown for parameter estimates that differed from the truth for at least one modeling approach. The MCMC chains did not converge to the same posterior space for the No (feedback) Control (FP) and the Direct FC models, and thus, posterior results are based on the chain with the lowest deviance; posterior results are based on all 3 chains for the Direct & Indirect FC model.

of disagreements was notably higher for the FP approach, intermediate for the direct FC model, and lowest for the direct and indirect FC model. Importantly, the FP model yielded MCMC chains for $\lambda_U = 2$ that quickly moved towards the upper limit (200) of the uniform prior, for all chains. When this limit was increased to an unrealistically high value (e.g., 1000), the FP chains still bumped against the upper limit. In addition to a greater number of disagreements, the FP model also yielded posterior means for seven species–level θ's that greatly differed from the true value (i.e., by 25% to 47%); the direct FC model resulted in four θ's that differed by 25–41%, and the direct and indirect FC model only yielded two θ's that differed from the true values (by 25%). Thus, the results from the synthetic data analysis support our decision to modularize the SLA meta–analysis model.

6 Model Results and Implications for SLA

Here we focus on the results related to the SLA mean model and key parameters of interest associated with the FC model, involving both direct and indirect feedback control, as described for the real data in Sections 2–4 and Figure 2. For brevity, we do not discuss the results associated with the covariate models as we saw no problems with these results, but we do discuss some of the results in Section 6.5 as related to the model diagnostics. Here we discuss model goodness–of–fit with respect to how well the meta–analysis model can replicate observed $\bar{y}$ (Section 6.1), posterior estimates of covariate effects and the importance of accounting for different sources of variability despite incomplete reporting (6.2), posterior estimates of latent SLA (θ, $\dot{\theta}$) for the 305 species and 185 genera considered here (6.3), sources of variability affecting SLA (6.4), and implications of modularization and feedback control for covariate parameters (6.5).

6.1 Model Goodness–of–fit Associated with SLA–related Data

Plots of replicated ($\bar{y}_{rep}$) versus observed $\bar{y}$ (e.g., Gelman et al. 2004) indicate good agreement between the model and data (Figure 3). In addition to results for all records (Figure 3A), we partitioned $\bar{y}_{rep}$ and $\bar{y}$ into two groups according to record completeness. We defined the complete reporting group as all of the records for which all of covariate information was reported, and the incomplete reporting group included records that were missing data for at least one covariate. The goal of evaluating model goodness of fit separately for the two groups was to qualitatively explore if incomplete reporting affected model fit. The model was equally successful at describing the observed $\bar{y}$ for both the complete and incomplete reporting groups (compare Figure 3B vs. 3C, respectively).

Although pairs of ($\bar{y}$, $E(\bar{y}_{rep})$) generally fall close to the 1:1 line (Figure 3), the actual containment percentages for the nominal 95% central posterior credible intervals (CIs) for $\bar{y}_{rep}$ are 88.9% for complete reporting, and 91.7% for incomplete reporting. Thus, predictions are accurate, but coverage is lower than expected, which may be due to the fact that we did not allow for study–specific parameters associated with the sample size (N) and standard error (se) data (i.e., λ and σ are only allowed to vary by U, the

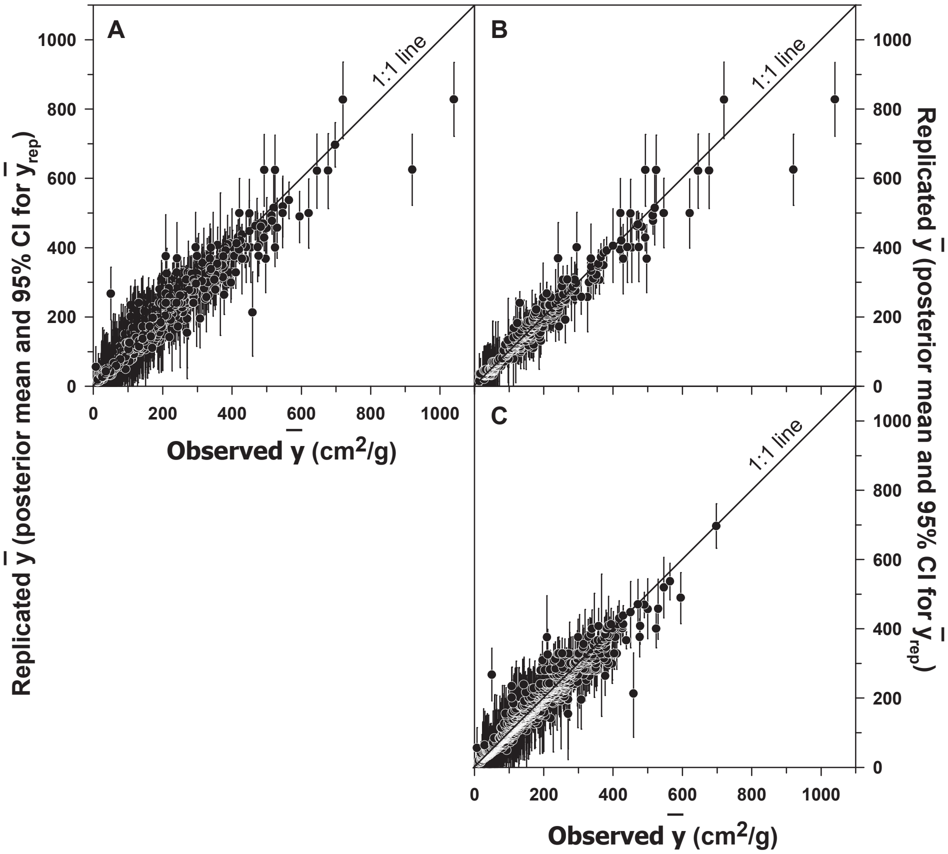

Figure 3: Predicted versus observed ($\bar{y}$) SLA values; predicted values are the posterior means ($E(\bar{y}_{rep})$, dots) and 95% posterior credible intervals (CIs, whiskers) for $\bar{y}_{rep}$. (A) all 1865 records ($E(\bar{y}_{rep}) = 11.57 + 0.93\bar{y}$, $R^2 = 0.94$); (B) the 388 records associated with complete reporting ($E(\bar{y}_{rep}) = 12.86 + 0.91\bar{y}$, $R^2 = 0.94$); and (C) the 1477 records associated with incomplete reporting ($E(\bar{y}_{rep}) = 10.84 + 0.93\bar{y}$, $R^2 = 0.94$).

unit of observation). That is, although incorporation of the N and se information, as per expressions (2) and (3), allowed us to explicitly account for uncertainty associated with each reported $\bar{y}$, we assumed that the parameters associated with the N and se distributions (λ and σ) were the same for all records and studies.

6.2 Covariate Effects

Incorporation of sampling–related covariates into our meta–analysis allows us to explicitly account for important sampling artifacts. For example, the incorporation of extra, non–leaf material (e.g., twigs, stems) for a canopy ($C = 2$) versus leaf–level ($C = 1$)

sample results in a larger denominator in the division of leaf area by sample dry weight, underestimating the actual SLA (area/weight). Thus, we expected canopy–level samples to yield lower $\bar{y}$ values than leaf–level, but $\bar{y}$ values alone cannot tell us how much lower. In fact, the posterior mean for the multiplicative effect of a canopy sample (ρ_2, see expression (4)) was less than 1.0 (Figure 4); the leaf–level effect was treated as the reference value such that $\rho_1 = 1$. That is, leaf–level samples, which more accurately represent the real SLA, produce $\bar{y}$ values that are about 1.65 times ($1/\rho_2$) higher than those associated with canopy–level samples. Evidently, our meta–analysis leads to plausible results beyond the information contained in the $\bar{y}$ data alone. This result supports our treatment of missing covariates; omission of records with missing C information may have compromised our ability to obtain a precise estimate for ρ given that about 27% of the records were missing C, and canopy samples ($C = 2$) accounted for less than 5% of the reported values.

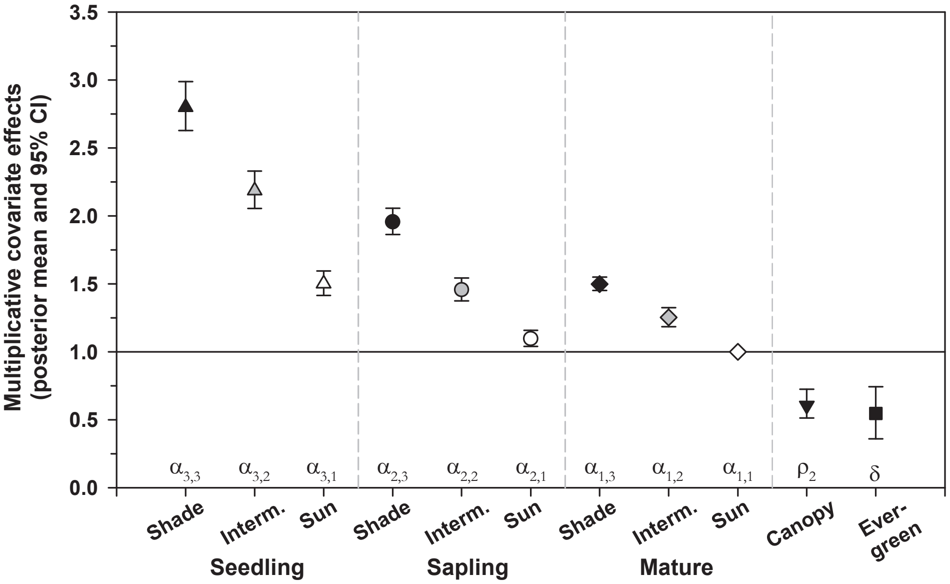

Figure 4: Posterior means and 95% credible intervals (CI) for the covariate effects in the SLA mean model, including the multiplicative light and age effects (α, see expression (4)), the sampling method effect (ρ, expression (4)), and the evergreen effect (δ, expression (10)).

Moreover, incorporation of leaf habit (evergreen vs. deciduous), tree age (A), and light environment (L) effects allowed us to obtain realistic estimates of species–specific, latent, one–sided SLAs that account for these important covariates, even if these covariates are associated with incomplete reporting. Within the division Magnoliophyta, species with evergreen leaves have significantly lower SLA values than deciduous species within the same genus, with deciduous leaves having mean SLAs about 1.8 times ($1/\delta$) greater than those of evergreen leaves (see δ in Figure 4). Likewise, across all species,

A and L are critical determinants of latent SLA. Seedlings generally have SLA values about 1.5 to 1.9 times higher than those of adult (mature) trees (see α terms in Figure 4). Seedlings also had higher SLA values than saplings. Leaves produced under low light (shade) have SLAs that are about 1.5 to 1.9 times higher than those of leaves produced under high light (sun) (Figure 4). A and L interacted to affect SLA such that differences in light environment had the greatest multiplicative effect on latent SLA within seedlings, and differences in tree age had the greatest multiplicative effect under shaded conditions (Figure 4). The effect of age class and light environment is comparable to the multiplicative difference between evergreen and deciduous leaves and between canopy– and leaf–level samples. All of these results are consistent with our ecological understanding of factors affecting SLA (Lusk and Warton 2007) such that SLA tends to be higher for young versus old trees (Steele et al. 1989), deciduous versus evergreen species (Reich et al. 1998a; Ackerly et al. 2002), and shaded versus sun–exposed leaves (Popma and Bongers 1988; Dong 1993).

6.3 Latent SLA

We illustrate leaf–level, 1–sided, species–specific latent SLA estimates for the two most extreme age–light categories: for seedlings in the shade (highest SLA) and mature trees in the sun (lowest SLA). Posterior estimates of $\theta_j \cdot \alpha_{3,3}$ (seedlings in the shade) and θ_j (mature trees in the sun) for Coniferophyta and deciduous Magnoliphyta, and associated $\theta_j \cdot \delta \cdot \alpha_{3,3}$ and $\theta_j \cdot \delta$ for evergreen Magnoliophya are shown in Figure 5. For the mature–sun classification, the posterior means and 95% CIs for θ_j ranged from 33.8 [5.1, 62.6] cm^2/g for *Juniperus monosperma* (one–seed juniper; $n = 4$ records for this species) to 179.4 [135.3, 225.3] cm^2/g for *Juglans nigra* (black walnut; $n = 2$ records). These species also defined the range predicted for the seedling–shade classification such that $\theta_j \cdot \alpha_{3,3}$ spanned 94.5 [14.7, 175.7] to 502.3 [376.8, 637.8] cm^2/g. Thus, one–sided, leaf–level SLA is expected to differ by a factor of approximately 15 across the range of species, age classes, and light environments considered here.

In general, latent SLA varies across taxonomic groups according to our expectations (Figure 5). For example, angiosperms (division Magnoliophyta) had higher SLA than conifers (division Coniferophyta) (e.g., Figure 5A), such that the posterior means and 95% CIs for $\dddot{\theta}_1$ and $\dddot{\theta}_2$ were 124.1 [106.4, 138.3] and 45.3 [7.3, 72.1] cm^2/g, respectively. Within the Magnoliophyta, oaks have the lowest SLA, which reflects their leaf morphology (tend to have thick, tough leaves), while maples have relatively high SLA (tend to have large, thin leaves) (Figure 5A).

The precision of the species– and genus–level latent SLA estimates varied across taxonomic groups, but was essentially independent of taxonomic identity. Rather, this precision was related to the number of records associated with each genus (results not shown). The precision of the θ_j estimates was also similar for well–represented species (Figure 5C) versus poorly represented species (Figure 5B). In this case study, the ability to provide "tight" estimates for poorly represented species can be explained by the borrowing of strength among species within each genus, which is facilitated by the

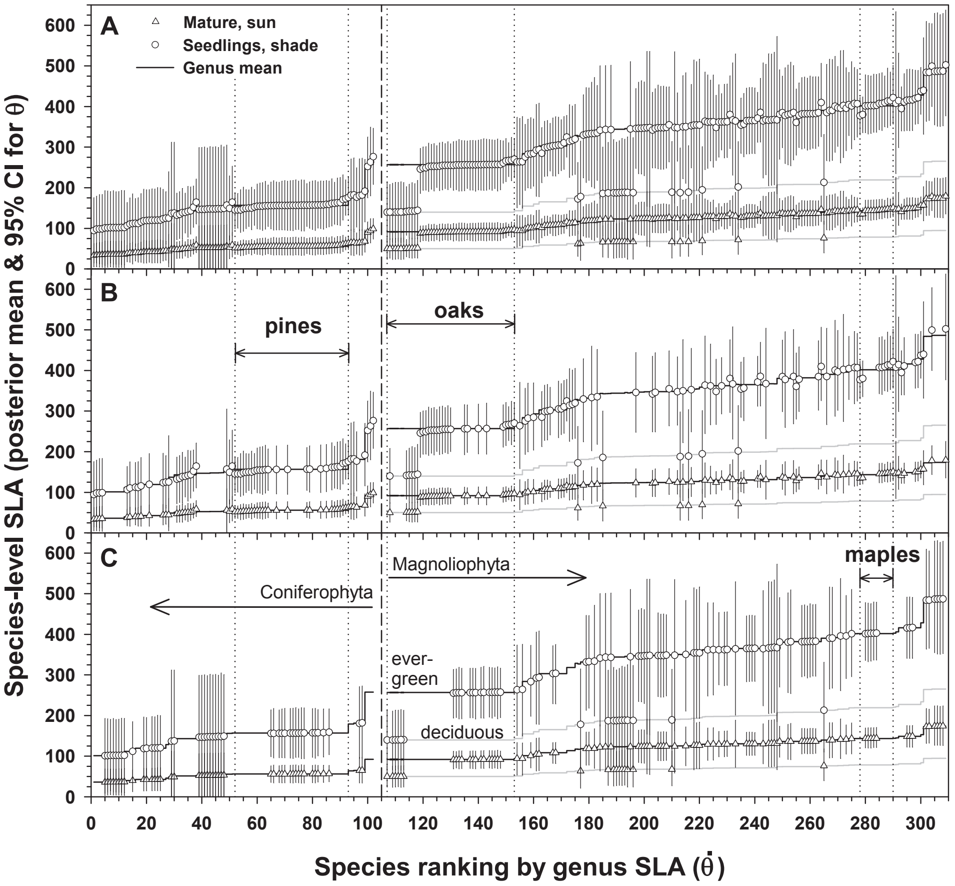

Figure 5: Latent, one–sided SLA estimates for the two "extreme" age by light categories. For each species j, symbols are the posterior means for θ_j (mature, sun; triangles) and $\theta_j \cdot \alpha_{3,3}$ (seedling, shade; circles); whiskers represent the central 95% credible interval (CI). Within the Magnoliophyta, these effects correspond to deciduous species, and evergreen species are given by $\theta_j \cdot \delta$ and $\theta_j \cdot \alpha_{3,3} \cdot \delta$. Estimates are shown for: (A) all 305 species, (B) the 158 species represented in the TreeTraits database, and (C) the 147 species not represented in TreeTraits. The solid lines are the posterior means for the genus–level values (based on $\dot{\theta}$); within the Magnoliophyta, the black lines correspond to deciduous and the gray lines to evergreen species. The dashed vertical line separates the two divisions; the dotted vertical lines delineate three genera (pines, oaks, maples) with many species.

taxonomic–based hierarchical model for θ_j. This borrowing of strength is fairly pronounced because the between species variability (within a genus) is small, and it is noticeably smaller than the variability due to study and genus (e.g., Figure 6). Thus, posterior means of θ_j for poorly represented species are nearly identical to the posterior mean for the associated genus, $\dot{\theta}_g$.

6.4 Sources of Variability

In addition to the aforementioned covariate effects, other factors appear critical to understanding sources of variation affecting SLA. Clearly, variability attributed to studies is large compared to the variability within the different taxonomic levels (Figure 6). That is, the sources of variation between species within genus ($\dot{\sigma}$), between genera within orders ($\ddot{\sigma}$), and between orders within division ($\dddot{\sigma}$) are small compared to the variability between studies (υ) (Figure 6). The overall variation attributed to different taxonomic levels within the complete taxonomy is comparable to, but slightly less than υ (Figure 6). The between study standard deviation (υ) also exceeds the expected within study standard deviations for $\bar{y}$ (i.e., $\frac{\sigma}{\sqrt{E(N)}}$: $\frac{\sigma_1}{\sqrt{\lambda_1+2}}$, $\frac{\sigma_2}{\sqrt{\lambda_2+2}}$,$\frac{\sigma_3}{\sqrt{\lambda_3+2}}$; Figure 6), which are comparable to the taxonomy standard deviations ($\dot{\sigma}$, $\ddot{\sigma}$, $\dddot{\sigma}$). Although we accounted for the effect of light (L), this indicates that other site–level environmental factors—such as, but not limited to, water and/or nutrient availability (Knops and Reinhart 2000; Niinemets 2001; Wright et al. 2004; Hoffmann et al. 2005; Meier and Leuschner 2008)—appear to be stronger determinants of SLA than evolutionary forces (as encapsulated by the taxonomy), which is consistent with other studies evaluating the role of taxonomy versus environment on SLA (Wilson et al. 1999). Although environmental influences are strong, taxonomy is still important, and it also provides a method for estimating latent SLA for species not in the TreeTraits database.

6.5 The Effect of Modularization on Covariate Parameters

As a model diagnostic, we implemented the "covariate only" model with the complete covariate dataset whereby missing values were replaced with the MAP estimates from the FC model (see Section 4.3). Although the posteriors obtained for covariate parameters based on the observed data often differed from those obtained from the MAP estimates, the differences make sense given our knowledge about the missingness mechanisms. For example, the probability ps of a study reporting 2–sided leaf area ($S = 2$) was relatively high under the observed data, but was notably less under the imputed data (posterior means and 95% CIs for ps_1 and ps_2 were 0.229 [0.152, 0.318] and 0.093 [0.042, 0.161], respectively). Note, however, if we are interested in learning about the true covariate parameter, then we would use the posterior for ps^{true} obtained from the FC model. In this case, the posterior mean for ps^{true} was 0.163 and the 95% CI was [0.115, 0.221]. Note that the posterior mean for ps^{true} is essentially a weighted average of the posterior means for ps_1 and ps_2, where the weights are approximately equal to the proportion of reported S (for ps_1) and missing S (for ps_2).

The higher values for ps under the observed compared to the imputed data are expected because the standard approach is to measure 1–sided leaf area ($S = 1$) (Cornelissen et al. 2003). We suspect that authors are more inclined to explicitly indicate 2–sided leaf area because this deviates from the standard approach, while those em-

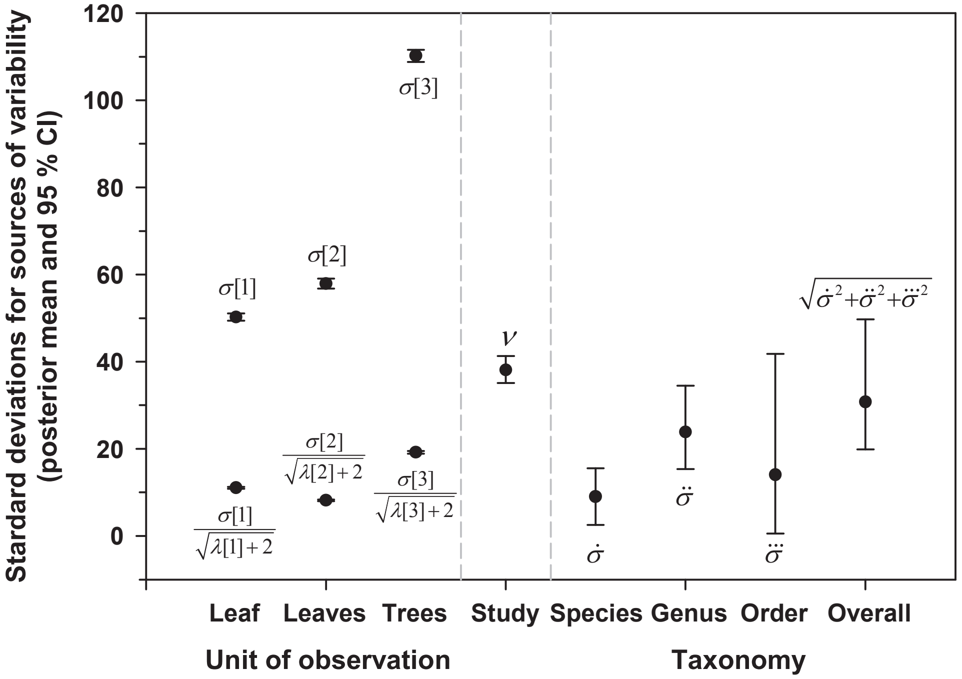

Figure 6: Posterior means and central 95% credible intervals (CI) for different standard deviation terms. Results are shown for the sample–to–sample SLA variability (σ_1, σ_2, and σ_3, for leaf, leaves, and trees, respectively), variability associated with the SLA sample mean, $\bar{y}$ ($\frac{\sigma}{\sqrt{E(N)}} = \frac{\sigma}{\sqrt{\lambda+2}}$ terms associated with each σ), study within species variability (υ), and taxonomic variability ($\dot{\sigma}$, $\ddot{\sigma}$, and $\dddot{\sigma}$ for between species within genus, between genera within order, and between orders within division, respectively, including the combined (overall) variation explained by the taxonomy ($\sqrt{\dot{\sigma}^2 + \ddot{\sigma}^2 + \dddot{\sigma}^2}$)).

ploying the standard approach may be less likely to indicate this explicitly. Similar arguments can be constructed for the differences in the posteriors for *pu*, *pl*, and *pa* associated with their respective observed versus imputed covariate data. Thus, such “nonrandom” missingness mechanisms provide further justification for modularization of the covariate model and severing feedback between the imputed covariates and the covariate effects parameters because we do not necessarily expect the missing data to have comparable covariate parameters as the observed data (e.g., Molitor et al. 2009).

7 Discussion

Feedback control and modularization are proving to be useful tools in relatively complex models wherein one or more modules may have an unrealistic or undesirable influence on other modules (Jackson et al. 2009; Liu et al. 2009; Lunn et al. 2009a; Molitor et al. 2009). We illustrate a unique application of modularization within a Bayesian meta–analysis of information extracted from the literature. Modularization within the meta–analysis was necessary to accommodate a large amount of missing information due to incomplete reporting, solving potential identifiability issues associated with estimating missing covariate data, associated covariate effects parameters, and data covariance parameters. Modularization of the covariate module and covariate effects module was easily accomplished in OpenBUGS (Lunn et al. 2009b) using the built–in cut function (e.g., Spiegelhalter et al. 2003; Jackson et al. 2009; Lunn et al. 2009a; Molitor et al. 2009).

Our approach advances upon existing meta–analytic methodologies because it is capable of accommodating all sources of information—whether complete or incomplete—and dealing with multiple sources of uncertainty within a transparent and statistically rigorous manner. This is not the first study, however, to describe a hierarchical Bayesian approach to meta–analysis (e.g., Sutton and Abrams 2001; Chung et al. 2002; Schmid et al. 2004; Burr and Doss 2005; Conlon et al. 2007; Dunson et al. 2008), and within the broad field of ecology, several examples exist in the fisheries literature (e.g., Liermann and Hilborn 1997; Harley and Myers 2001; Millar and Methot 2002; Barrowman et al. 2003; Helser and Lai 2004). Existing Bayesian approaches, however, still suffer from problems associated with incomplete reporting and ignoring the stochastic nature of the standard error, sample size, and covariate data. This is the first study that we are aware of that explicitly tackles all of these issues, and our modularization approach in particular allows us to address the commonplace issue of incomplete reporting.

Within our meta–analysis model, we use basic probability theory and sufficiency arguments to develop a stochastic model for the SLA standard error data, and we also treat the sample size and all covariate data as stochastic. These stochastic models provide a framework for addressing incomplete reporting, whereby we modularize the missing information such that missing data do not feed back to affect parameters in their associated stochastic models (direct feedback control), and they do not affect their associated effects parameters in the SLA mean model (indirect feedback control). By implementing both direct and indirect feedback control, we are able to avoid identifiability issues between missing covariate data, covariate parameters, covariate effect parameters, and SLA covariance parameters. Direct feedback control has been presented elsewhere (e.g., Jackson et al. 2009; Lunn et al. 2009a; Molitor et al. 2009), but we are not aware of any studies that discuss the potential utility of indirect feedback control. The performance and utility of indirect feedback control remain to be explored in more general settings, and may be most relevant in complex models such as the SLA meta–analysis described here. In this study, implementation of both direct and indirect feedback control were necessary to improve mixing, achieve convergence, obtain realistic pseudo–posteriors for the parameters of interest (e.g., Figures 4 and 5), and produce a model that successfully reproduced the observed SLA sample mean ($\bar{y}$) data (Figure 3).

Although we illustrate and apply our modularized meta–analysis model to literature information obtained for specific leaf area (SLA), we suggest that many aspects of our approach are amenable to the more general problem of meta–analysis for parameter estimation, across an array of scientific fields. Our modularization approach may also help direct future improvements to data assimilation techniques. More generally, implementation of direct and indirect feedback control may be necessary for successful implementation of complex models that are characterized by the coupling of diverse modules and data sources.

References

Abrams, K. R., Gillies, C. L., and Lambert, P. C. (2005). "Meta-analysis of heterogeneously reported trials assessing change from baseline." *Statistics in Medicine*, 24: 3823–3844. 141

Ackerly, D. D., Knight, C. A., Weiss, S. B., Barton, K., and Starmer, K. P. (2002). "Leaf size, specific leaf area and microhabitat distribution of chaparral woody plants: contrasting patterns in species level and community level analyses." *Oecologia*, 130(3): 449–457. 157

Barrowman, N. J., Myers, R. A., Hilborn, R., Kehler, D. G., and Field, C. A. (2003). "The variability among populations of coho salmon in the maximum reproductive rate and depensation." *Ecological Applications*, 13(3): 784–793. 161

Bonan, G. B., Levis, S., Sitch, S., Vertenstein, M., and Oleson, K. W. (2003). "A dynamic global vegetation model for use with climate models: concepts and description of simulated vegetation dynamics." *Global Change Biology*, 9(11): 1543–1566. 136, 137

Brooks, S. P. and Roberts, G. O. (1998). "Convergence assessment techniques for Markov chain Monte Carlo." *Statistics and Computing*, 8(4): 319–335. 151

Burr, D. and Doss, H. (2005). "A Bayesian semiparametric model for random-effects meta-analysis." *Journal of the American Statistical Association*, 100(469): 242–251. 161

Casella, G. and Berger, R. L. (2002). *Statistical Inference*. Pacific Grove, CA: Duxbury. 141

Chen, H. Y. H. (1997). "Interspecific responses of planted seedlings to light availability in interior British Columbia: survival, growth, allometric patterns, and specific leaf area." *Canadian Journal of Forest Research*, 27(9): 1383–1393. 138

Chen, H. Y. H., Klinka, K., and Kayahara, G. J. (1996). "Effects of light on growth, crown architecture, and specific leaf area for naturally established *Pinus contorta* var *latifolia* and *Pseudotsuga menziesii* var *glauca* saplings." *Canadian Journal of Forest Research*, 26(7): 1149–1157. 138

Chmura, D. J. and Tjoelker, M. G. (2008). "Leaf traits in relation to crown development, light interception and growth of elite families of loblolly and slash pine." *Tree Physiology*, 28(5): 729–742. 138

Chung, Y. S., Dey, D. K., and Jang, J. H. (2002). "Semiparametric hierarchical selection models for Bayesian meta analysis." *Journal of Statistical Computation and Simulation*, 72(10): 825–839. 134, 161

Conlon, E. M., Song, J. J., and Liu, A. (2007). "Bayesian meta-analysis models for microarray data: a comparative study." *BMC Bioinformatics*, 8: doi:10.1186/1471–2105–8–80. 161

Cornelissen, J. H. C., Lavorel, S., Garnier, E., Diaz, S., Buchmann, N., Gurvich, D. E., Reich, P. B., ter Steege, H., Morgan, H. D., van der Heijden, M. G. A., Pausas, J. G., and Poorter, H. (2003). "A handbook of protocols for standardised and easy measurement of plant functional traits worldwide." *Australian Journal of Botany*, 51(4): 335–380. 159

Cramer, W., Bondeau, A., Woodward, F. I., Prentice, I. C., Betts, R. A., Brovkin, V., Cox, P. M., Fisher, V., Foley, J. A., Friend, A. D., Kucharik, C., Lomas, M. R., Ramankutty, N., Sitch, S., Smith, B., White, A., and Young-Molling, C. (2001). "Global response of terrestrial ecosystem structure and function to CO_2 and climate change: results from six dynamic global vegetation models." *Global Change Biology*, 7(4): 357–373. 137

Dong, M. (1993). "Morphological plasticity of the clonal herb *Lamiastrum galeobdolon* (L.) Ehrend. & Polatschek in response to partial shading." *New Phytologist*, 124(2): 291–300. 157

Dunson, D. B., Xue, Y., and Carin, L. (2008). "The matrix stick-breaking process: Flexible Bayes meta-analysis." *Journal of the American Statistical Association*, 103(481): 317–327. 161

Duursma, R. A., Marshall, J. D., Nippert, J. B., Chambers, C. C., and Robinson, A. P. (2005). "Estimating leaf-level parameters for ecosystem process models: a study in mixed conifer canopies on complex terrain." *Tree Physiology*, 25(11): 1347–1359. 136, 137

Ellsworth, D. S. and Reich, P. B. (1993). "Canopy structure and vertical patterns of photosynthesis and related leaf traits in a deciduous forest." *Oecologia*, 96(2): 169–178. 138

Furukawa, T. A., Barbui, C., Cipriani, A., Brambilla, P., and Watanabe, N. (2006). "Imputing missing standard deviations in meta-analyses can provide accurate results." *Journal of Clinical Epidemiology*, 59(1): 7–10. 141

Gelman, A. (2006). "Prior distributions for variance parameters in hierarchical models." *Bayesian Analysis*, 1(3): 515–533. 146

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004). *Bayesian Data Analysis*. Boca Raton: Chapman and Hall/CRC Press. 139, 151, 154

Gurevitch, J., Curtis, P. S., and Jones, M. H. (2001). "Meta-analysis in ecology." *Advances in Ecological Research*, 32: 199–247. 134

Gurevitch, J. and Hedges, L. V. (1999). “Statistical issues in ecological meta-analyses.” *Ecology*, 80(4): 1142–1149. 134

Harley, P. C. and Baldocchi, D. D. (1995). “Scaling carbon dioxide and water vapour exchange from leaf to canopy in a deciduous forest. I. Leaf model parametrization.” *Plant Cell and Environment*, 18(10): 1146–1156. 137

Harley, S. J. and Myers, R. A. (2001). “Hierarchical Bayesian models of length-specific catchability of research trawl surveys.” *Canadian Journal of Fisheries and Aquatic Sciences*, 58(8): 1569–1584. 161

Hartung, J., Knapp, G., and Sinha, B. K. (2008). *Statistical Meta-Analysis with Applications*. Hoboken, NJ: John Wiley & Sons, Inc. 134

Hedges, L. V., Gurevitch, J., and Curtis, P. S. (1999). “The meta-analysis of response ratios in experimental ecology.” *Ecology*, 80(4): 1150–1156. 134

Helser, T. E. and Lai, H. L. (2004). “A Bayesian hierarchical meta-analysis of fish growth: with an example for North American largemouth bass, *Micropterus salmoides*.” *Ecological Modelling*, 178(3-4): 399–416. 161

Higgins, J. P. T., Thompson, S. G., and Spiegelhalter, D. J. (2009). “A re-evaluation of random-effects meta-analysis.” *Journal of the Royal Statistical Society Series A-Statistics in Society*, 172: 137–159. 134

Hoffmann, W. A., Franco, A. C., Moreira, M. Z., and Haridasan, M. (2005). “Specific leaf area explains differences in leaf traits between congeneric savanna and forest trees.” *Functional Ecology*, 19(6): 932–940. 159

Jackson, C. H., Best, N. G., and Richardson, S. (2009). “Bayesian graphical models for regression on multiple data sets with different variables.” *Biostatistics*, 10(2): 335–351. 134, 151, 161

Kattge, J., Ogle, K., Boenisch, G., Diaz, S., Lavorel, S., Madin, J., Nadrowski, K., Noellert, S., Sartor, K., and Wirth, C. (2011). “A generic structure for plant trait databases.” *Methods in Ecology and Evolution*, 2: 202–213. 138

Knops, J. M. H. and Reinhart, K. (2000). “Specific leaf area along a nitrogen fertilization gradient.” *American Midland Naturalist*, 144(2): 265–272. 159

Koike, T., Kitao, M., Maruyama, Y., Mori, S., and Lei, T. T. (2001). “Leaf morphology and photosynthetic adjustments among deciduous broad-leaved trees within the vertical canopy profile.” *Tree Physiology*, 21(12-13): 951–958. 138

Lajeunesse, M. J. and Forbes, M. R. (2003). “Variable reporting and quantitative reviews: a comparison of three meta-analytical techniques.” *Ecology Letters*, 6(5): 448–454. 134

Liermann, M. and Hilborn, R. (1997). “Depensation in fish stocks: a hierarchic Bayesian meta-analysis.” *Canadian Journal of Fisheries and Aquatic Sciences*, 54(9): 1976–1984. 161

Liu, F., Bayarri, M., and Berger, J. (2009). "Modularization in Bayesian analysis, with emphasis on analysis of computer models." *Bayesian Analysis*, 4(1): 119–150. 134, 135, 147, 161

Lunn, D., Best, N., Spiegelhalter, D., Graham, G., and Neuenschwander, B. (2009a). "Combining MCMC with 'sequential' PKPD modelling." *Journal of Pharmacokinetics and Pharmacodynamics*, 36(1): 19–38. 134, 135, 136, 146, 147, 148, 150, 151, 161

Lunn, D., Spiegelhalter, D., Thomas, A., and Best, N. (2009b). "The BUGS project: Evolution, critique and future directions (with discussion)." *Statistics in Medicine*, 28: 3049–3082. 134, 136, 151, 161

Lunn, D. J., Thomas, A., Best, N., and Spiegelhalter, D. (2000). "WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility." *Statistics and Computing*, 10: 325–337. 134

Lusk, C. H. and Warton, D. I. (2007). "Global meta-analysis shows that relationships of leaf mass per area with species shade tolerance depend on leaf habit and ontogeny." *New Phytologist*, 176(4): 764–774. 157

Meier, I. C. and Leuschner, C. (2008). "Leaf size and leaf area index in *Fagus sylvatica* forests: Competing effects of precipitation, temperature, and nitrogen availability." *Ecosystems*, 11(5): 655–669. 159

Melillo, J. M., McGuire, A. D., Kicklighter, D. W., Moore, B., Vorosmarty, C. J., and Schloss, A. L. (1993). "Global climate change and terrestrial net primary production." *Nature*, 363(6426): 234–240. 136, 137

Miles, P., Brand, G., Alerich, C., Bednar, L., Woudenberg, S., Glover, J., and Ezzell, E. (2001). "The Forest Inventory and Analysis Database: Database Description and Users Manual." USDA Forest Service, Report NC-218. 137

Millar, R. B. and Methot, R. D. (2002). "Age-structured meta-analysis of US West Coast rockfish (Scorpaenidae) populations and hierarchical modeling of trawl survey catchabilities." *Canadian Journal of Fisheries and Aquatic Sciences*, 59(2): 383–392. 161

Molitor, N. T., Best, N., Jackson, C., and Richardson, S. (2009). "Using Bayesian graphical models to model biases in observational studies and to combine multiple sources of data: application to low birth weight and water disinfection by-products." *Journal of the Royal Statistical Society Series A-Statistics in Society*, 172: 615–637. 134, 136, 148, 150, 160, 161

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. 159

Norby, R. J., Wullschleger, S. D., Gunderson, C. A., Johnson, D. W., and Ceulemans, R. (1999). "Tree responses to rising CO_2 in field experiments: implications for the future forest." *Plant Cell and Environment*, 22(6): 683–714. 136

Ogle, K. and Pacala, S. W. (2009). "A modeling framework for inferring tree growth and allocation from physiological, morphological and allometric traits." *Tree Physiology*, 29: 587–605. 136, 137

Osenberg, C. W., Sarnelle, O., and Cooper, S. D. (1997). "Effect size in ecological experiments: The application of biological models in meta-analysis." *American Naturalist*, 150(6): 798–812. 134

Osenberg, C. W., Sarnelle, O., and Goldberg, D. E. (1999). "Meta-analysis in ecology: Concepts, statistics, and applications." *Ecology*, 80(4): 1103–1104. 134

Picard, G., Woodward, F. I., Lomas, M. R., Pellenq, J., Quegan, S., and Kennedy, M. (2005). "Constraining the Sheffield dynamic global vegetation model using streamflow measurements in the United Kingdom." *Global Change Biology*, 11(12): 2196–2210. 137

Popma, J. and Bongers, F. (1988). "The effect of canopy gaps on growth and morphology of seedlings of rain forest species." *Oecologia*, 75(4): 625–632. 157

Reich, P. B., Ellsworth, D. S., and Walters, M. B. (1998a). "Leaf structure (specific leaf area) modulates photosynthesis-nitrogen relations: evidence from within and across species and functional groups." *Functional Ecology*, 12(6): 948–958. 157

Reich, P. B., Tjoelker, M. G., Walters, M. B., Vanderklein, D. W., and Bushena, C. (1998b). "Close association of RGR, leaf and root morphology, seed mass and shade tolerance in seedlings of nine boreal tree species grown in high and low light." *Functional Ecology*, 12(3): 327–338. 138

Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989). "Design and analysis of computer experiments." *Statistical Science*, 4: 409–423. 147

Sands, P. J. and Landsberg, J. J. (2002). "Parameterisation of 3-PG for plantation grown *Eucalyptus globulus*." *Forest Ecology and Management*, 163(1-3): 273–292. 137

Schimel, D. S. (1995). "Terrestrial ecosystems and the carbon cycle." *Global Change Biology*, 1(1): 77–91. 136

Schmid, C. H., Stark, P. C., Berlin, J. A., Landais, P., and Lau, J. (2004). "Meta-regression detected associations between heterogeneous treatment effects and study-level, but not patient-level, factors." *Journal of Clinical Epidemiology*, 57(7): 683–697. 161

Spiegelhalter, D., Thomas, A., Best, N., and Lunn, D. (2003). "WinBUGS Version 1.4 User Manual." Technical report, Medical Research Council Biostatistics Unit. 134, 151, 161

Steele, M. J., Coutts, M. P., and Yeoman, M. M. (1989). "Developmental changes in Sitka spruce as indices of physiological age. I. Changes in needle morphology." *New Phytologist*, 113(3): 367–375. 138, 157

Sutton, A. J. and Abrams, K. R. (2001). "Bayesian methods in meta-analysis and evidence synthesis." *Statistical Methods in Medical Research*, 10(4): 277–303. 161

Sutton, A. J. and Higgins, J. P. T. (2008). "Recent developments in meta-analysis." *Statistics in Medicine*, 27: 625–650. 134

Tang, J. Y. and Zhuang, Q. L. (2008). "Equifinality in parameterization of process-based biogeochemistry models: A significant uncertainty source to the estimation of regional carbon dynamics." *Journal of Geophysical Research-Biogeosciences*, 113(G4): G04010. 137

— (2009). "A global sensitivity analysis and Bayesian inference framework for improving the parameter estimation and prediction of a process-based Terrestrial Ecosystem Model." *Journal of Geophysical Research-Atmospheres*, 114: D15303. 137

Tatarinov, F. A. and Cienciala, E. (2006). "Application of BIOME-BGC model to managed forests 1. Sensitivity analysis." *Forest Ecology and Management*, 237(1-3): 267–279. 136

USDA NRCS (2008). *The PLANTS Database*. Baton Rouge, LA: USDA Natural Resources Conservation Service National Plant Data Center http://www.plants.usda.gov. 137

Wang, B., Zou, X., and Zhu, J. (2000). "Data assimilation and its applications." *Proceedings of the National Academy of Sciences*, 97(21): 11143–11144. 148

Wiebe, N., Vandermeer, B., Platt, R. W., Klassen, T. P., Moher, D., and Barrowman, N. J. (2006). "A systematic review identifies a lack of standardization in methods for handling missing variance data." *Journal of Clinical Epidemiology*, 59(4): 342–353. 134, 141

Wilson, P. J., Thompson, K., and Hodgson, J. G. (1999). "Specific leaf area and leaf dry matter content as alternative predictors of plant strategies." *New Phytologist*, 143(1): 155–162. 159

Wramneby, A., Smith, B., Zaehle, S., and Sykes, M. T. (2008). "Parameter uncertainties in the modelling of vegetation dynamics - Effects on tree community structure and ecosystem functioning in European forest biomes." *Ecological Modelling*, 216(3-4): 277–290. 137

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J. H. C., Diemer, M., Flexas, J., Garnier, E., Groom, P. K., Gulias, J., Hikosaka, K., Lamont, B. B., Lee, T., Lee, W., Lusk, C., Midgley, J. J., Navas, M. L., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V. I., Roumet, C., Thomas, S. C., Tjoelker, M. G., Veneklaas, E. J., and Villar, R. (2004). "The worldwide leaf economics spectrum." *Nature*, 428(6985): 821–827. 159

Acknowledgments

This work was supported by a National Science Foundation (NSF) grant awarded to K.O. and J.J.B. in 2009 (DBI–0850361), NSF grants awarded to K.O. in 2003 and 2006 (#0630119), and Wyoming NASA Space Grant Consortium grants awarded to K.O. and J.J.B. in 2008 (NASA grant #NNG05G165H). The authors thank Brenda Thompson for implementing earlier versions of the model in WinBUGS; Christian Wirth, Jeremy Lichstein, and Jens Kattge for stimulating discussions that contributed to this work; and Katya Belykh and Kimberly Garvie for contributing to earlier versions of the TreeTraits database.