

# Calibrating jack pine allometric relationships with simultaneous regressions

Robert Schneider, Frank Berninger, Chhun-Huor Ung, Pierre Y. Bernier, D. Edwin Swift, and S.Y. Zhang

**Abstract:** Allometric equations for estimating foliage biomass, sapwood area, and branch basal area from tree diameters and crown lengths for jack pine (*Pinus banksiana* Lamb.) in eastern Canada were calibrated using mixed models. A first model is presented that relates branch foliage biomass to branch diameter and relative position within the crown. These results show that a branch's foliage biomass is inversely proportional to its depth within the crown. At the tree level, foliage biomass was found to be proportional to crown length and to vary with stem age and slenderness. Pipe model parameters (sapwood area and branch basal area to foliage biomass) were also calculated. The sapwood area to foliage biomass parameter is proportional to stand density, whereas branch basal area to foliage biomass is constant. The tree-level allometric models were calibrated using a mixed-effects seemingly unrelated regression to account for between-model correlations.

**Résumé :** Des équations allométriques visant à estimer la biomasse foliaire, la superficie d'aubier et la surface terrière des branches à partir du diamètre des arbres et de la longueur de la cime de pins gris (*Pinus banksiana* Lamb.) de l'est du Canada ont été étalonnées à l'aide de modèles mixtes. Le premier modèle qui est présenté relie la biomasse foliaire des branches au diamètre des branches et à leur position relative dans la cime. Ces résultats indiquent que la biomasse foliaire des branches est inversement proportionnelle à sa profondeur dans la cime. À l'échelle de l'arbre, nous avons observé que la biomasse foliaire était proportionnelle à la longueur de la cime et qu'elle variait aussi en fonction de l'âge de l'arbre et de son élancement. Les paramètres du modèle tubulaire (superficie d'aubier et surface terrière des branches en fonction de la biomasse foliaire) ont aussi été calculés. Le paramètre associé à la superficie d'aubier en fonction de la biomasse foliaire est proportionnel à la densité du peuplement alors que celui de la surface terrière des branches en fonction de la biomasse foliaire est constant. Les modèles allométriques à l'échelle de l'arbre ont été étalonnés en utilisant la technique de régression sans corrélation apparente à effets mixtes pour tenir compte des corrélations entre les modèles.

[Traduit par la Rédaction]

## Introduction

Process-based models require that carbohydrates generated through photosynthesis be partitioned among a tree's constitutive components (trunk, roots, branches, and foliage). In some process-based models, the net effects of allocation and losses to litter are bounded by rules derived from allometric relationships (Landsberg and Waring 1997; Mäkelä 1997; Valentine and Mäkelä 2005). These very useful relationships maintain relative growth of the different parts of a tree within the empirically observed domain, thereby providing a certain robustness to modeled estimates and allowing for realistic model behavior over long periods. However, to better understand the weaknesses of these models, their sources of variation must be known.

Allometric methods are based on destructive measurements of tree biomass. These measurements are labor intensive and very costly to carry out. A hierarchical sampling scheme is often used to reduce fieldwork and laboratory measurements by applying the results from a small sample collected at the level of branch to the tree and stand levels. Thus, statistical methods are needed to translate branch and foliage mass measurements to tree- and stand-level estimates. Berninger et al. (2005) have shown that a large proportion of the error in the scaling up of such field measurements to produce allometric relationships is linked to stand- and tree-level effects. Such random errors within populations are usually addressed by using mixed models (Mäkinen and Isomäki 2004a, 2004b; Fortin 2006). Mixed models further allow the partitioning of this error among

Received 11 February 2008. Accepted 18 June 2008. Published on the NRC Research Press Web site at cjfr.nrc.ca on 11 September 2008.

**R. Schneider<sup>1</sup> and F. Berninger.** Département des Sciences biologiques, Université du Québec à Montréal, C.P. 8888, succursale Centre Ville, Montréal, QC H3C 3P8, Canada.

**C. Ung and P.Y. Bernier.** Natural Resources Canada, Canadian Forest Service, Canadian Wood Fibre Centre, P.O. Box 13080, Stn Ste-Foy, Québec, QC G1V 4C7, Canada.

**D.E. Swift.** Natural Resources Canada, Canadian Forest Service, Canadian Wood Fibre Centre, Atlantic Forestry Centre, P.O. Box 4000, Fredericton, NB E3B 5P7, Canada.

**S.Y. Zhang.** FPInnovations — Forintek Division, 2665 East Mall, Vancouver, BC V6T 1W5, Canada.

<sup>1</sup>Corresponding author (e-mail: schneider.robert@courrier.uqam.ca).

the different hierarchical levels, that is, the variance at the branch, tree, and stand levels.

The pipe model states that on a tree foliage biomass (or area) is directly related to the area of the sapwood downstream of this biomass (Shinozaki et al. 1964). Berninger et al. (2005) showed that parameters related to the pipe model of stem architecture and measured at tree and branch levels tend to be correlated with each other. Scaling up from the branch to the tree becomes more problematic in such situations because residuals at different levels are not independent from each other. Seemingly unrelated regression (SUR) is a statistical method that takes into account correlation between residuals of different models when estimating model parameters (Zellner 1962). Fang et al. (2001) introduced random effects in SUR using a forest growth and yield problem as an example. In a hierarchical sampling scheme, simultaneous estimation will lead to more efficient parameter estimates and unbiased parameter estimates if errors around different hierarchical levels are correlated.

The current study focuses on the estimation of total-tree foliage biomass using the simultaneous-regression mixed-model-based approach. Our objective is to determine the degree of correlation in errors among different levels of sampling as well as to quantify the importance of the random factors at different hierarchical levels (tree, plot, and site). Parameters are estimated in the context of a process-based model called Crobas (Mäkelä 1997), but similar situations arise during the calibration of many other models, as when tree measurements are used to estimate the leaf area index for remote sensing, or when scaling up sample-tree data to landscape-level estimates of biomass. Allometric relationships are used in Crobas to estimate tree productivity (via foliage biomass) and to partition growth among the different parts of the tree. The source of the errors on each parameter has to be clearly identified to ensure that the simulation results of the model are unbiased. Moreover, if some parameters show important trends with regard to random effects, a set of regional or local parameters might produce better results. Finally, a clear identification of the sources of variance of each parameter will help target research needs.

## Material

Our analysis is based on 84 jack pine (*Pinus banksiana* Lamb.) trees sampled in the summers of 2006 and 2007 within three sites spread across eastern Canada (Table 1). In the Petawawa Research Forest, 16 trees were felled in a spacing trial dating from 1970: 10 trees from the normally stocked plot (2200 stems/ha) and 6 from the low-density plot (550 stems/ha). On the Smurfit-Stone freehold in central Quebec, 50 trees were felled in several operational plantations established between 1968 and 1985 and in six naturally generated stands. The sampling in New Brunswick was undertaken in a stand that was precommercially thinned in 1966 (Zhang et al. 2006). Six stems were sampled from the low-density spacing (1320 stems/ha), six from the medium-density spacing (2200 stems/ha), and six from the high-density spacing (6720 stems/ha).

The sampled trees were selected to cover the range of tree sizes observed in the sample plots. Before felling, diameters

at breast height (DBH) and crown widths along the eight cardinal directions were measured. Once felled, total stem height and height to crown base were measured, followed by the measurement of height, diameter, insertion angle, and state (live or dead) of each branch on the stem. Of all the live branches, five sample branches per tree were systematically selected. The needles of these branches were then removed, oven dried, and weighed to yield the foliage biomass of that branch. Branch diameters with and without bark were measured, as were branch and bark biomasses. Finally, disks were taken at regular intervals along the stem (0.15 m, 0.70 m, 1.30 m, crown base, and equally spaced between crown base and stem apex – breast height such that the maximum spacing between the disks was 2 m) to measure sapwood area.

## Methods

Four relationships needed to be established. The first one is for predicting branch biomass from branch diameter. The three other relationships are at the tree level and form a system of equations that from the principles of the pipe model allows the estimation of (i) tree foliage biomass from crown length, (ii) sapwood area of the stem at crown base from tree foliage biomass, and (iii) total branch basal area of the stem from tree foliage biomass. The interrelationships among the equations are as proposed by Mäkelä (1997).

### Branch level: estimating branch foliage biomass

To determine the total foliage biomass from the sample branches, a nonlinear mixed-effects model was developed relating branch foliage biomass ( $W_f$ ) to over-bark branch diameter ( $d_{ob}$ ). The relative depth within the crown ( $x$ , distance from top of the tree divided by crown length) was also included in the model, where  $l$  represents the branch of tree  $k$ , which is nested in plot  $j$ , which, in turn, is nested within site  $i$  (eq. 1a). The random effects and error terms were also assumed to be normally distributed and independent of each other (eqs. 1b and 1c), where  $\sigma_i^2$ ,  $\sigma_{ij}^2$ ,  $\sigma_{ijk}^2$ ,  $\sigma_{res}^2$ , and  $\theta$  are estimated variance and covariance parameters, and  $\mathbf{d}_{ob,ijk}$  is a vector whose elements are the branch diameters  $d_{ob,ijkl}$ . No correlation structure for the branches within a tree was included in the model, leading to a diagonal matrix of ones symbolized by  $\Phi(0)$  in eq. 1d. An exponential variance function was used to account for the observed heteroscedasticity, where  $\Psi(\theta, \text{dia}_{ob,ijk})$  is a diagonal matrix (eq. 1d) where each diagonal element  $\psi_{m,n}$  is given by eq. 1e. The values of  $\Psi(\theta, \text{dia}_{ob,ijk})$  were used as weights during the estimation of model parameters. All models presented in this paper were calibrated using the NLME function in the R statistical software (R Development Core Team 2004; Pinheiro et al. 2005).

$$[1a] \quad W_{f,ijkl} = (\beta_1 + b_{1i} + b_{1j} + b_{1ijk}) \times d_{ob,ijkl}^{\beta_2} \times (1 + x_{ijkl})^{\beta_2} + \varepsilon_{ijkl}$$

where  $\beta_1$  and  $\beta_2$  are fixed-effect parameters;  $b_{1i}$ ,  $b_{1j}$ , and  $b_{1ijk}$  are random-effect parameters for the site, plot, and tree levels, respectively;  $W_f$  is branch foliage biomass (kilograms);  $d_{ob}$  is over-bark branch diameter (millimetres); and

**Table 1.** Stand and sample-tree characteristics (minimum and maximum values indicated in parentheses).

	Ontario	New Brunswick	Quebec
<b>Mean site characteristics*</b>			
Total precipitation (mm)	938.9	1113.9	985.6
Mean yearly temperature (°C)	4.5	5.2	1.2
Degree-days above 5 °C	1688	1175	1634
<b>Mean stand characteristics</b>			
Quadratic mean DBH (cm)	21.9 (18.5–25.2)	15.3 (13.6–17.7)	10.9 (6.0–17.7)
Basal area (m <sup>2</sup> /ha)	32.4 (26.3–39.1)	29.2 (14.1–36.2)	18.3 (8.6–31.4)
Density (stems/ha)	964 (544–1456)	1808 (575–2575)	2124 (1143–3850)
Age (years)	36	56	22.5 (20–37)
Site index <sup>†</sup>	14.0	9.0	7.5 (5.9–10.4)
No. of plots	2	3	17
<b>Mean sample-tree characteristics</b>			
DBH (mm)	208 (116–287)	163 (85–228)	159 (41–297)
Height (m)	20.04 (16.92–22.76)	15.92 (9.80–19.30)	15.06 (5.24–22.76)
Average branch diameter (mm)	19.6 (6.2–58.1)	15.8 (4.5–35.1)	12.4 (2.8–45.5)
No. of sample trees	16	18	50
No. of sample branches	71	81	190

\*From Canadian ecodistrict climate normals 1961–1990, Agriculture and Agri-Food Canada.

<sup>†</sup>Site index: dominant stand height (metres) at age 20.

$x$  is relative depth of branch within crown (distance from treetop / crown length ( $H_c$ ))

$$[1b] \quad b_i \stackrel{i.i.d.}{\sim} N(0, \sigma_i^2), \quad b_{ij} \stackrel{i.i.d.}{\sim} N(0, \sigma_{ij}^2), \quad b_{ijk} \stackrel{i.i.d.}{\sim} N(0, \sigma_{ijk}^2)$$

$$[1c] \quad \varepsilon_{ijkl} \stackrel{i.i.d.}{\sim} N(0, R_{ijk}(\sigma_{res}^2, \theta, d_{ob,ijk}))$$

$$[1d] \quad R_{ijk}(\sigma_{res}^2, \theta, d_{ob,ijk}) = \sigma_{res}^2 \Psi(\theta, d_{ob,ijk}) \Phi(0) \Psi(\theta, d_{ob,ijk})$$

$$[1e] \quad \psi_{m,n} = e^{\theta d_{ob,ijkl}} \big|_{m=n}$$

Total tree foliage biomass was estimated by using the best linear unbiased prediction, that is, by applying the fixed and random effects of each tree to its branch diameter measurements. The estimated total tree foliage biomass ( $\hat{W}_{f,ijk}$ ) corresponds to the sum of the predicted branch foliage biomass for each tree (i.e.,  $\hat{W}_{f,ijk} = \sum_l \hat{W}_{f,ijkl}$ ). This information is used in the following sections.

### Tree level: stem foliage biomass and pipe model parameters

The relationship between  $\hat{W}_{f,ijk}$  and  $H_c$  was calibrated using a nonlinear mixed-effects model, with the total tree height to DBH ratio included in the exponent (eq. 2), and random effects accounting for the plots ( $j$ ) nested within the sites ( $i$ ). As with the previous models, the different hierarchical levels of the random effects were assumed to be independent and normally distributed.

$$[2] \quad \hat{W}_{f,ijk} = (\beta_3 + b_{3,i} + b_{3,ij} + \beta_4 \text{AGE}_{ijk}) \times H_{c,ijk}^{\beta_5 + b_{5,i} + b_{5,ij} + \beta_6 \frac{H_{tot,ijk}}{DBH_{ijk}}} + \varepsilon_{2ijk}$$

where  $\beta_3$ ,  $\beta_4$ ,  $\beta_5$ , and  $\beta_6$  are fixed-effect parameters;  $b_{x,i}$  and

$b_{x,ij}$  are random-effect parameters for each site and plot, respectively;  $\hat{W}_{f,ijk}$  is tree foliage biomass (kilograms);  $\text{AGE}_{ijk}$  is stem age (years);  $H_{c,ijk}$  is crown length (metres);  $H_{tot,ijk}$  is total tree height (metres); and  $DBH_{ijk}$  is diameter at breast height (centimetres).

In its original version, the pipe model relates total tree foliage biomass to sapwood area through a single parameter (Shinozaki et al. 1964). Mäkelä (1997) inverts this relationship and predicts sapwood area from foliage biomass. It is within this framework that the pipe model parameters are estimated, and thus the traditional form is simply the inverse of the results presented here.

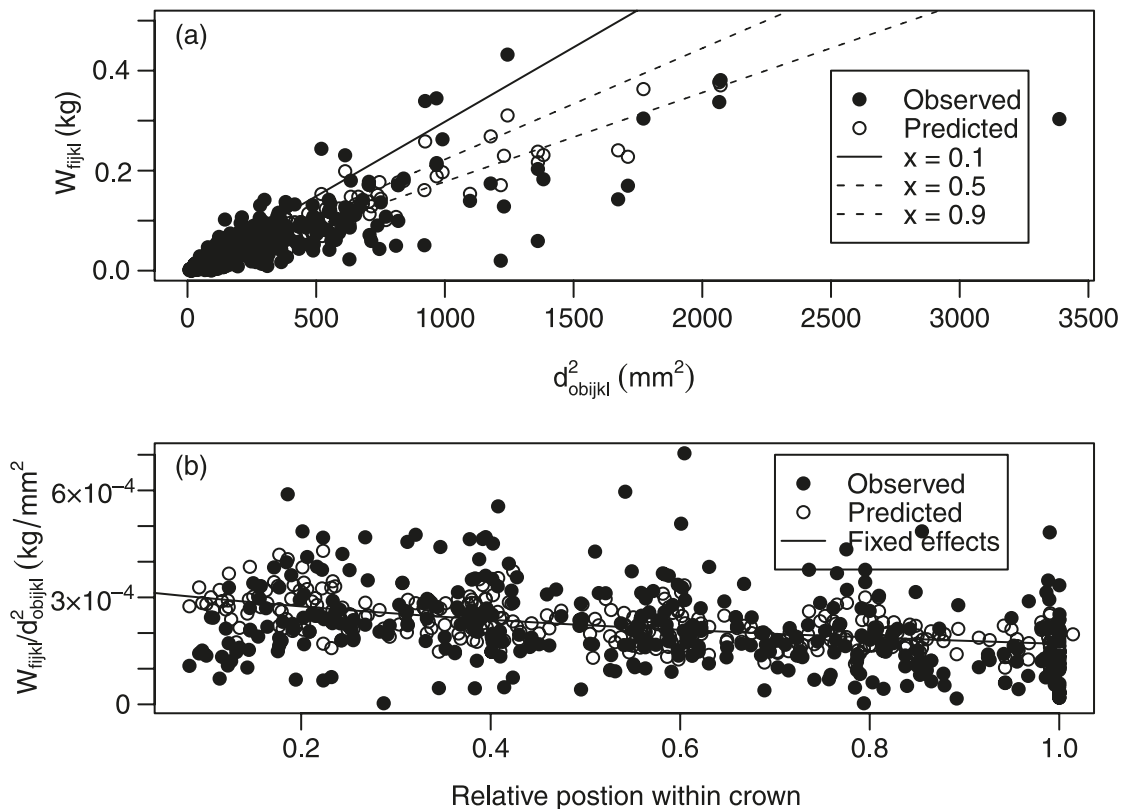
The relationships between sapwood area at the base of the crown ( $A_s$ ) or total branch basal area of the tree ( $A_b$ ) to estimated total tree foliage biomass ( $\hat{W}_{f,ijk}$ ) were calibrated using linear mixed-effect models (eq. 3). Stand density had an effect on model parameters and was thus included in the model. As with the branch foliage biomass model (eq. 1a), random effects accounting for plots ( $j$ ) nested within sites ( $i$ ) were used. The same assumptions regarding the independence and normal distribution of the random effects and error were also made.

$$[3] \quad \begin{aligned} A_{z,ijk} &= \beta_z \hat{W}_{f,ijk} + \varepsilon_{z,ijk} \\ z &= s, b \end{aligned}$$

where  $\beta_s = \beta_{s,1} + b_{s,i} + b_{s,ij} + \beta_{s,2} D_{ij}$ ;  $\beta_b = \beta_{b,1} + b_{b,i} + b_{b,ij}$ ;  $b_{z,i}$  and  $b_{z,ij}$  are site and plot random-effect parameters, respectively;  $D_{ij}$  is stand density (stems per hectare);  $s$  is sapwood area at crown base ( $A_s$ , square metres); and  $b$  is total branch basal area of the tree ( $A_b$ , square metres).

The three allometric relationships presented here (eqs. 2 and 3) are, however, linked, as the predicted tree foliage biomass from eq. 2 was used to determine sapwood and branch basal areas. The three tree-level equations were thus simultaneously calibrated using a seemingly unrelated re-

**Fig. 1.** (a) Branch foliage biomass ( $\hat{W}_{f,ijkl}$ ) as a function of over-bark branch diameter ( $\text{dia}_{ob,ijkl}$ ) squared, with three different relative heights ( $x$ ), and (b) branch foliage biomass to branch diameter squared as a function of relative position within the crown.



gression (SUR) to account for between-model correlations with random effects (Fang et al. 2001). Details can be found in Appendix A.

## Results

### Branch level: estimating branch foliage biomass

The branch foliage biomass model predicted that a branch near the top of the tree (low values of relative height,  $x$ ) has more foliage biomass than does a branch of the same diameter at the base of the crown (Fig. 1). Furthermore, a branch's biomass is inversely proportional to its relative height, as indicated by the sign of the  $\beta_2$  exponent (Table 2). The predicted trend of the model follows the observed data for jack pine, where a decline in the foliage biomass to branch diameter squared is observed (Fig. 1).

The importance of each component of the total predicted variance (residual, site, plot, and tree random effects) was calculated using a matrix of branch diameters ranging from 1 to 40 mm with 0.5 mm increment and relative heights ranging from 0 to 1 (0 is treetop) with 0.05 increments. Averaging the percentage of the total variance for each component yields the following picture: the random effects accounted for 98% of the total variance, with the site explaining less than 1% of the total variance, the plot random effect 59%, and the tree random effect 39%.

Moreover, a trend can be observed in the estimates of random-effect parameters for the different trees: in central Quebec, the tree random parameters were the lowest, followed by those in Ontario, whereas the trees from New Brunswick had the highest values (Fig. 2). This follows the

trend observed in the data whereby the trees sampled in central Quebec have less foliage biomass for a given branch size and crown depth than trees sampled in New Brunswick, which had the most, and Ontario.

### Tree level: stem foliage biomass and pipe model parameters

Tree foliage biomass is related to crown length, but also depends on tree age and total height to DBH ratio (Fig. 3a). The  $\beta_6$  coefficient estimate is negative, leading to lower tree biomass for stems with higher total height to DBH ratios (Table 3). Inversely, tree foliage biomass is proportional to stem age, as indicated by the positive value of the  $\beta_4$  coefficient estimate.

The pipe model parameters seem to be influenced by certain tree- and stand-level variables (Fig. 4). The stem foliage biomass to sapwood area at crown base and stem branch basal area to tree foliage biomass ratios are inversely proportional to stand density and proportional to stem age and stem size. This is reflected in the estimation of the pipe model parameters through eq. A1c, where stand density was included. Stem size (DBH) was also tested, as well as stand density, in the branch basal area to leaf biomass equation, but the fit statistics were not as good (Bayesian information criterion (BIC) = -1397 with DBH in eqs. A1c, A1d, and no stand density; BIC = -1404, as presented in eqs. A1c, A1d; BIC = -1396 with stand density in eq. A1d). The positive value of  $\beta_8$  reflects the observed trend: stands that are less dense will have more leaf biomass per sapwood area (Fig. 3c). Once the stand density was included in the model, no trends in the residuals were observed with respect to tree

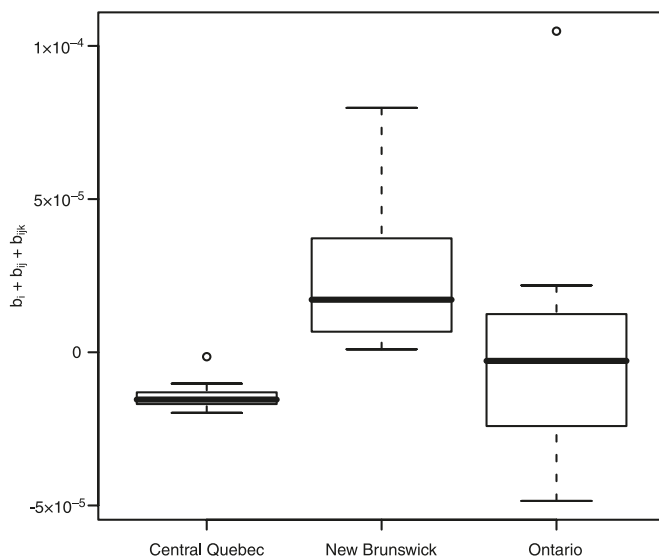
**Table 2.** Coefficients and fit statistics for the branch foliage biomass model (eq. 1a).

	Value	SE	<i>t</i>	<i>P</i>
<b>Coefficient index</b>				
$\beta_1$	0.000 325 7	0.000 022 3	13.73	<0.0001
$\beta_2$	-0.94	0.12	-7.63	<0.0001
<b>Standard errors</b>				
$\sigma_i$	$6.96 \times 10^{-10}$			
$\sigma_{ij}$	$6.08 \times 10^{-5}$			
$\sigma_{ijk}$	$4.90 \times 10^{-5}$			
$\sigma_{\text{res}}$	0.0043			
<b>Variance function parameter</b>				
$\theta$	0.0986			
<b>Fit statistics</b>				
No. of sites	3			
No. of plots	22			
No. of trees	81			
No. of observations	342			
<b>Pseudo-<math>R^{2*}</math></b>				
With random effects	0.75			
Without random effects	0.64			
<b>RMSE<sup>†</sup></b>				
With random effects	0.0315			
Without random effects	0.0376			
AIC <sup>‡</sup>	-1743			
BIC <sup>‡</sup>	-1712			
Log likelihood	880			

$$* \text{Calculated as } 1 - \left( \frac{\sum_{ijk} (y_{ijk} - \hat{y}_{ijk})^2}{\sum_{ijk} (y_{ijk} - \bar{y}_{ijk})^2} \right).$$

$$^{\dagger} \text{Calculated as } \left( \frac{\sum_{ijk} (y_{ijk} - \hat{y}_{ijk})^2}{n} \right)^{0.5}, \text{ where } n \text{ is the number of measurements.}$$

<sup>‡</sup>AIC, Akaike information criterion; BIC, Bayesian information criterion.

**Fig. 2.** Estimated random-effect parameters ( $b_{1i} + b_{1ij} + b_{1ijk}$ ) for the relationship between branch foliage biomass and branch over-bark branch diameter (eq. 1a) for each tree.

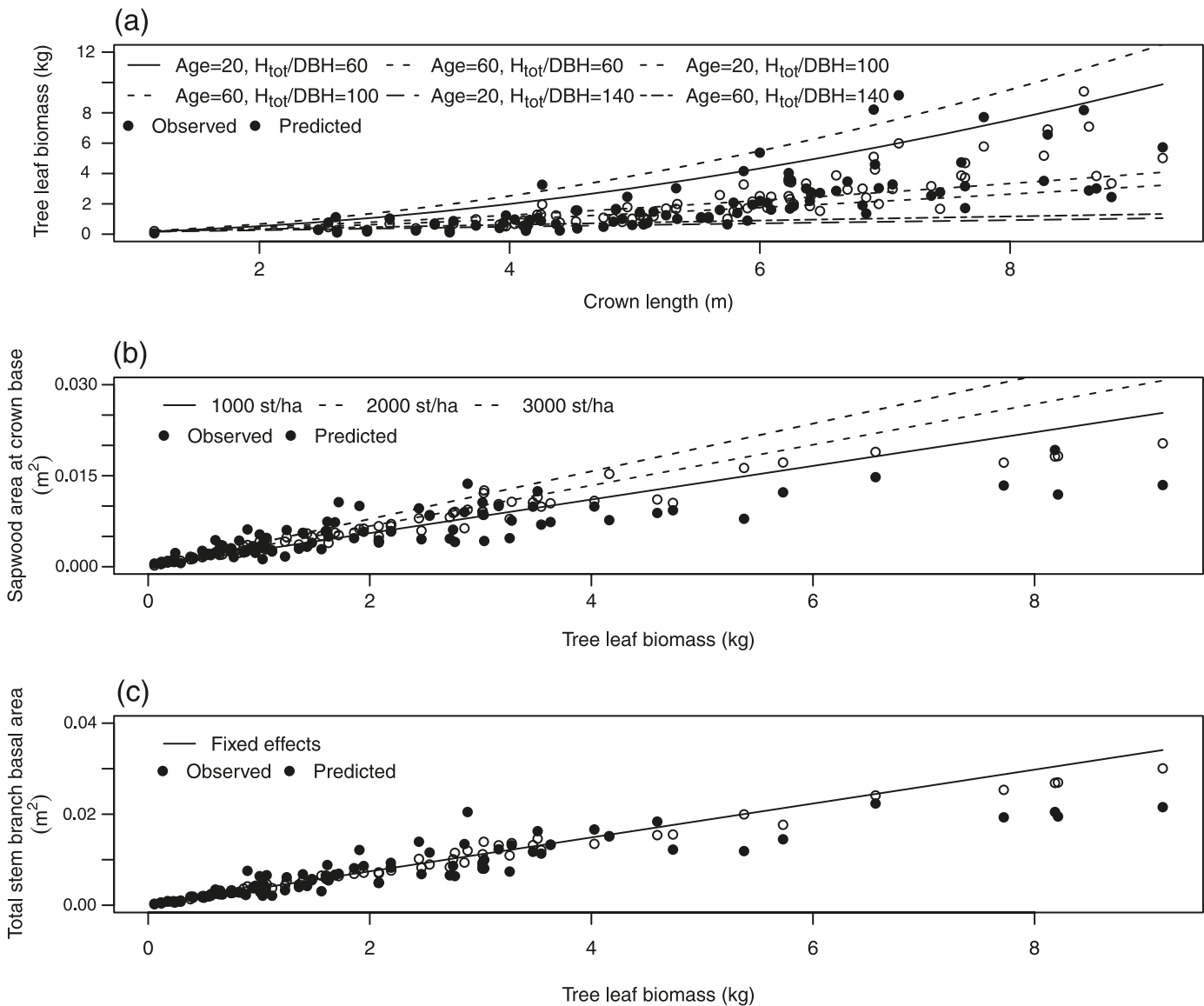
size, stand density, and basal area (Fig. 5). A slight negative trend in the residuals versus stem age is still present. Stem age was, however, statistically significant ( $P < 0.05$ ) in the pipe model parameter estimates, but the model fit statistics (Akaike information criterion (AIC) and BIC) were not improved. Moreover, the trends observed in the residuals were still present with stem age included in the model.

Furthermore, certain differences in the tree foliage biomass to sapwood area and tree foliage biomass to branch basal area ratios are observed in relation to the different sites, where the trees from the central Quebec site seem to be slightly different from those of the two other sites (Fig. 6). This is translated into differences in the random-effect estimates (Table 4): (1) trees in Ontario have more leaf biomass for a given crown length, age, and height to DBH ratio, followed by the trees in New Brunswick and then in Quebec; (2) inversely, the trees in New Brunswick have less sapwood area at crown base and branch basal area for a given tree foliage biomass when compared with the other two sites, in the order of Ontario followed by Quebec.

The relative importance of the variance components was averaged using a matrix of stem DBH ranging from 10 to 25 cm with an increment of 1 cm, total height between 5 and 20 m with 0.5 m increments, crown lengths ranging from 30% to 70% of the total height with 10% increments,



**Fig. 3.** Observed and predicted values for (a) tree foliage biomass ( $\hat{W}_{f,ijk}$ ) as a function of crown length ( $H_{c,ijk}$ ), with three different total height to DBH ratios and two different ages, (b) tree sapwood area at crown base ( $A_{s,ijk}$ ) as a function of estimated tree foliage biomass ( $\hat{W}_{f,ijk}$ ) for three different stand densities, and (c) total branch basal area of the stem ( $A_{b,ijk}$ ) as a function of estimated tree foliage biomass ( $\hat{W}_{f,ijk}$ ) for three different stand densities.



age ranging between 20 and 70 years with a 10 year increment, and stand density varying between 1000 and 3000 stems/ha with a 500 stems/ha increment. For determining the foliage biomass of the tree from the crown length, the site random effect explained 72% of the total variance, leaving 28% of the total variance due to the fixed effects (residual variance). No plot effects were significant in the first equation. The plot random effect accounted for less than 1% of the total variance for both pipe model parameter estimates, with no site random effect, leaving the residual variance explaining 99% of the total variance.

The correlations between the error terms of each equation were moderate (Table 3), and all of the error terms were negatively correlated among each other ( $\rho_{1,2} = -0.31$ ,  $\rho_{1,3} = -0.32$ ,  $\rho_{2,3} = -0.43$ ). Inversely, the variance of the plot-level random effect of the pipe model parameters was highly positively correlated ( $\sigma_4 = 0.997$ ).

Finally, the use of the SUR calibration method led to a more efficient estimation of the fixed effects (Fig. 7). The standard errors of the estimates are generally smaller with the simultaneous regression technique than with the independent calibration of each equation.

## Discussion

Foliage biomass of individual branches was influenced by the position of the branch along the stem. For a given diameter, a branch will have more foliage at the top of the tree than at the bottom because top branches receive less shading (Berninger and Nikinmaa 1994). When using the model to predict branch foliage biomass, information on the random effects, especially those concerning the plot and the tree, are important for reducing the error associated with the prediction. Of the two, the plot effect seems to be the most im-

**Table 3.** Coefficients and fit statistics for determining the tree foliage biomass, sapwood area at crown base, and stem branch basal area (eqs. A1a–A1j).

	Value	SE	<i>t</i>	<i>P</i>
<b>Coefficient index</b>				
$\beta_3$	0.1211	0.0356	3.39	0.0008
$\beta_4$	0.0009	0.0003	2.92	0.0038
$\beta_5$	2.6732	0.1186	22.53	<0.0001
$\beta_6$	−0.0126	0.0012	−9.93	<0.0001
$\beta_7$	0.0022	0.0004	5.21	<0.0001
$\beta_8$	$5.81 \times 10^{-7}$	$2.30 \times 10^{-7}$	2.50	0.0133
$\beta_9$	0.0037	0.0002	21.13	<0.0001
<b>Standard errors and their correlations (eq. A1e)</b>				
$\sigma_1$	0.0406			
$\sigma_2$	0.000 33			
$\sigma_3$	0.000 50			
$\sigma_4$	0.997			
$\sigma$	0.3755			
<b>Variance function parameters (eq. A1h)</b>				
$\theta_1$	0.8841			
$\theta_2$	0.9532			
$\theta_3$	1.0224			
<b>Correlation parameters (eq. A1j)</b>				
$\rho_{1,2}$	−0.305			
$\rho_{1,3}$	−0.315			
$\rho_{2,3}$	−0.429			
<b>Fit statistics</b>				
No. of sites	3			
No. of plots	22			
No. of trees	84			
No. of observations	252			
<b>Pseudo-<math>R^2</math>* (with random effects, without random effects)</b>				
$W_f$	0.72, 0.53			
$A_s$	0.53, 0.39			
$A_b$	0.78, 0.65			
<b>RMSE<sup>†</sup> (with random effects, without random effects)</b>				
$W_f$	1.03, 1.34			
$A_s$	0.0027, 0.0030			
$A_b$	0.0027, 0.0034			
AIC <sup>‡</sup>	−1467			
BIC <sup>‡</sup>	−1404			
Log likelihood	751			

$$^*\text{Calculated as } 1 - \left( \frac{\sum_{ijk} (y_{ijk} - \hat{y}_{ijk})^2}{\sum_{ijk} (y_{ijk} - \bar{y}_{ijk})^2} \right).$$

$$^{\dagger}\text{Calculated as } \left( \frac{\sum_{ijk} (y_{ijk} - \hat{y}_{ijk})^2}{n} \right)^{0.5}, \text{ where } n \text{ is the number of measurements.}$$

<sup>‡</sup>AIC, Akaike information criterion; BIC, Bayesian information criterion.

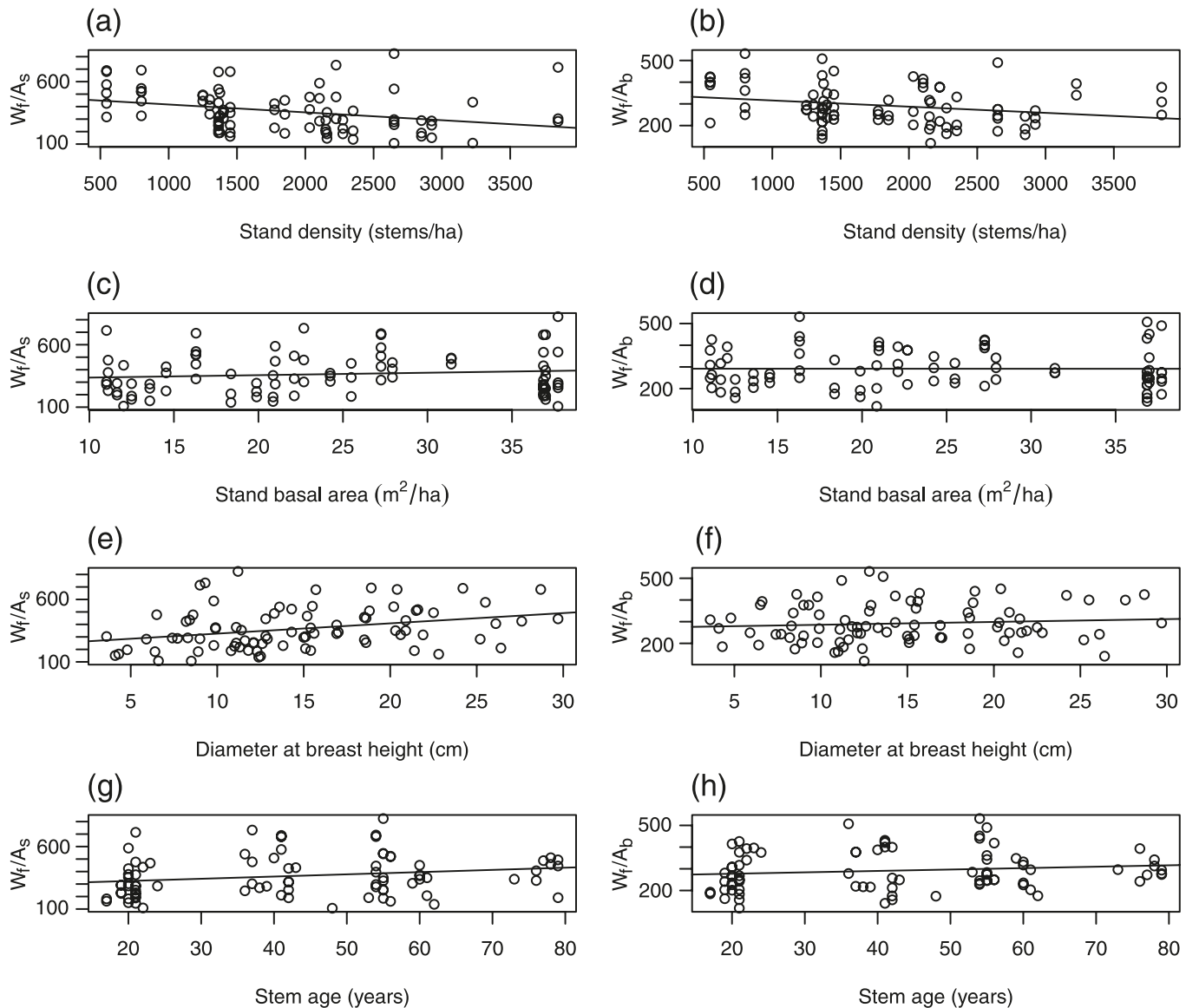
portant, and further data would be needed to verify whether stand-level information would increase the predictive capacity of the fixed effects.

Our results contrast with those for loblolly pine (*Pinus taeda* L.), where the slope of the branch foliage biomass to branch size was relatively constant throughout the top two-thirds of the crown and dropped in the bottom third (Valentine et al. 1994). On the other hand, Baldwin et al.

(1997) reported a slightly nonlinear decrease in branch foliage biomass with increasing crown depth, such as the results presented in this paper for jack pine. Our data show, however, a slightly stronger decrease in branch foliage biomass in the upper parts of the crown.

Moreover, the time of the sampling seems to affect the loblolly pine branch foliage biomass to branch size relationship, through differences either among or within years

**Fig. 4.** Estimated tree foliage biomass ( $\hat{W}_{f,ijk}$ ) to sapwood area at crown base ( $A_{s,ijk}$ ) (a, c, e, g) and to total branch basal area ( $A_{b,ijk}$ ) (b, d, f, h) illustrated as a function of stand density, stand basal area, diameter at breast height, and stem age.

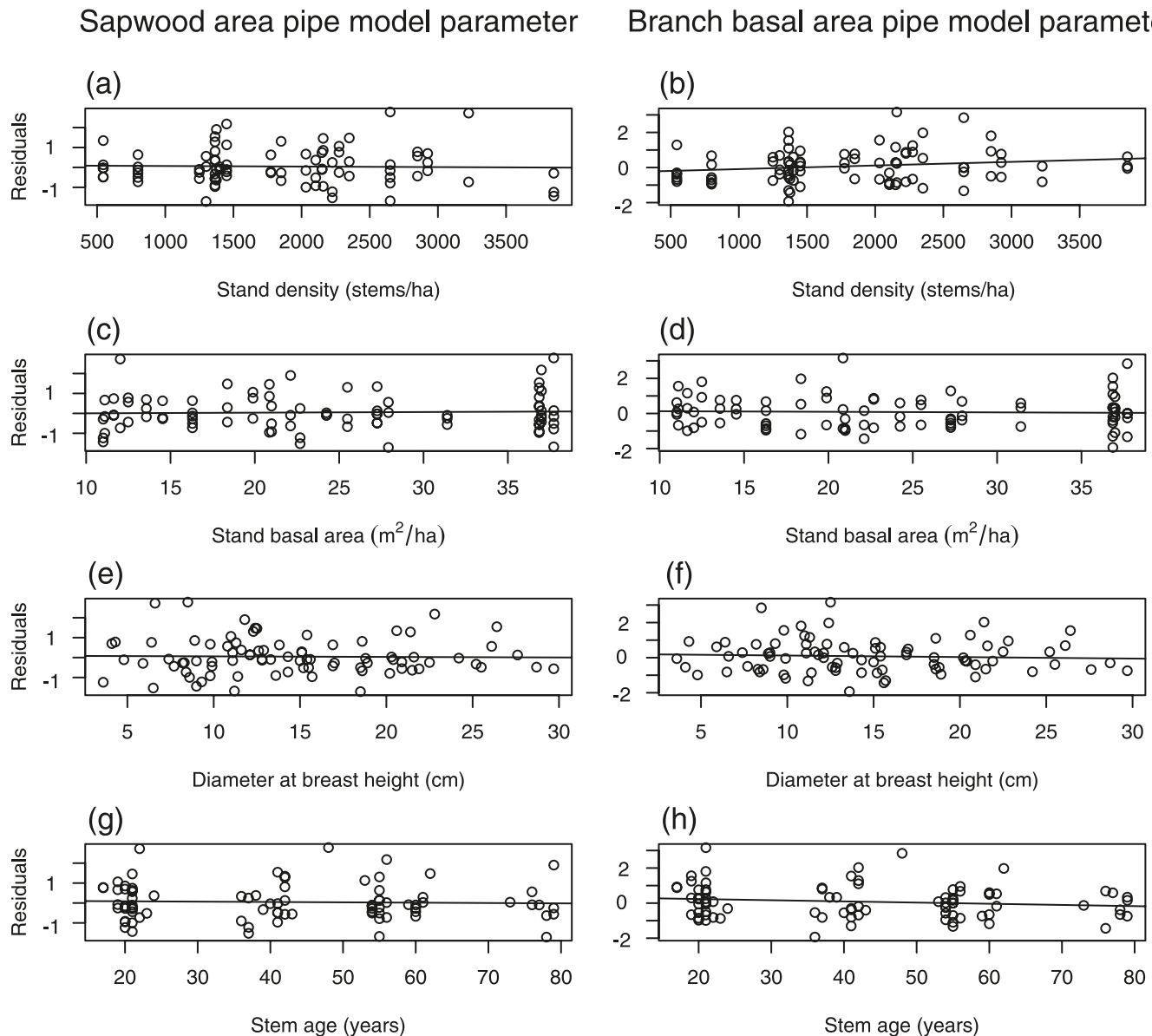


(Valentine et al. 1994; Baldwin et al. 1997). The seasonal and yearly differences were not observed within our data at either the branch or the tree level. If this had been the case, the central Quebec random effects would have had a greater spread because sampling was carried out in spring, summer, and fall, whereas the Ontario site was sampled in spring, and the New Brunswick site in summer. However, the main objective of this model was to estimate whole-tree foliage biomass. The use of random effects and the best unbiased prediction when applying the model to scale from branch to tree estimation accounts for any seasonal effects in sampling by applying a model that is specific to each individual (Littell et al. 1996). However, random site and plot effects were significant (with site effects being dominant). This indicates that there is a variation source that is not accounted for in the model, be it biological, genetic, or environmental. Moreover, the observed among-site variance could also be explained by interannual temporal variations, or might be due to some unexplainable random effect.

At the tree level, foliage biomass is inversely related to the height to DBH ratio, with trees that have a greater height/DBH ratio having less foliage biomass than smaller trees. The height/diameter ratio has frequently been used as a measure of suppression of trees. Jack pine seems to behave, in this respect, similarly to other pines (Shelburne et al. 1993; Vanninen and Mäkelä 1999; Berninger et al. 2005). Stem age was also found to affect the amount of foliage biomass a tree has, with foliage biomass being proportional to stem age. Mäkelä (1997) assumed that crown shape is constant throughout the life-span of the tree. However, this is not the case for jack pine. First, according to the preliminary analysis of our data, the vertical distribution of the foliage of jack pine seems to be affected by stem age, with the height of maximum foliage density being proportional to stem age. Second, height increment follows the same trend, with older trees having smaller increments than younger ones. This is also demonstrated by older trees producing a greater number of whorls per unit length of crown;



**Fig. 5.** Residuals of the pipe model equations (eqs. A1c and A1d) with respect to stand density, stand basal area, diameter at breast height, and stem age.

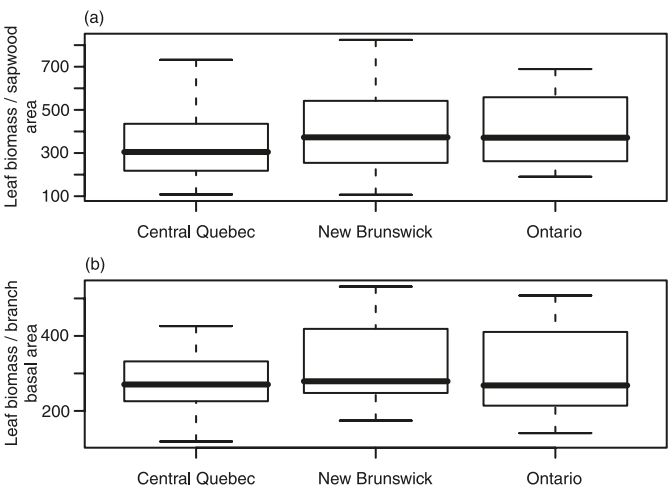


hence the observation that older trees have more foliage biomass than younger ones for a given crown length.

The exponent of eq. 3 is referred to as the fractal dimension of the crown, whereas the slope is known as the surface area density (Mäkelä 1997). Using the mean slenderness ( $H_{\text{tot}}/\text{DBH} = 96$ ) and age (40 years) from our data set, the predicted values for jack pine are 1.46 and 0.15, respectively. These are quite different from those used by Mäkelä (1997) in the Scots pine (*Pinus sylvestris* L.) simulations, where the fractal dimension was 2.5 (Mäkelä and Sievänen 1992) and the surface density was 0.05 (Mäkelä 1997). The problem, however, when empirically calibrating such a power function used to relate foliage biomass to crown length is that there is a high correlation between the slope and the exponent. The interpretation of each coefficient thus becomes impossible. Nevertheless, foliage biomass for a given crown length is lower for jack pine than for Scots pine.

The results indicate that the pipe model relationships hold quite well, but that other dendrometric variables seem to affect the parameters to a certain extent. With the data available, it was difficult to separate the influence of stem size and age from that of stand density. However, the model showed that when stand density is included, most of the trends observed in the residuals disappeared. These results agree with those presented for other species, for which shifts in the ratio of leaf biomass to sapwood area have been observed for stems of different sizes (Mäkelä and Albrektson 1992; Shelburne et al. 1993). On the other hand, Monserud and Marshall (1999) and Berninger et al. (2005) did not find any decrease in pipe model ratios with tree height. Age, however, is difficult to separate from tree size. Based on grafting experiments, Vanderklein et al. (2007) claimed that age has no effect on tree hydraulics and allometric relations, but that tree size could be a determining factor in the rela-

**Fig. 6.** Estimated tree foliage biomass ( $\hat{W}_{f,ijk}$ ) to sapwood area at crown base ratio ( $A_{s,ijk}$ ) (a) and to total branch basal area ( $A_{b,ijk}$ ) (b) ratio for each sampled tree.



**Table 4.** Site random-effect estimates for the tree foliage biomass and site average plot random effects for sapwood area at crown base models and branch basal area model (eqs. A1b–A1d).

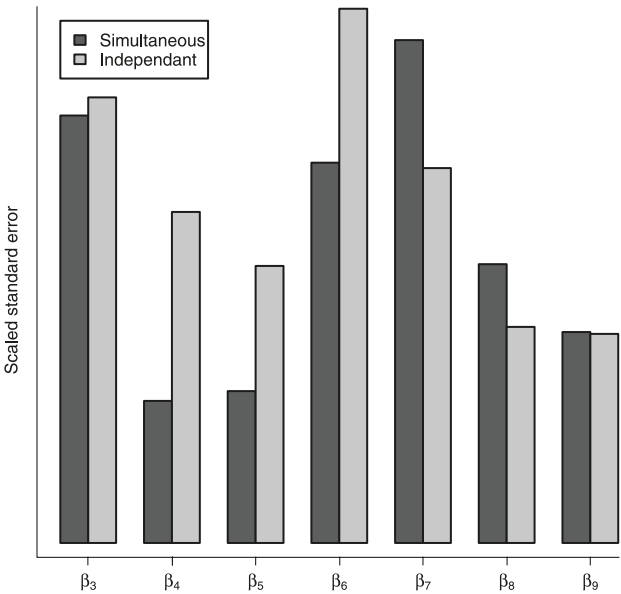
Site	$b_{3i}$	$\bar{b}_{7i}$	$\bar{b}_{9i}$
New Brunswick	0.0053	$-9.92 \times 10^{-5}$	$-1.54 \times 10^{-4}$
Ontario	0.0448	$-9.19 \times 10^{-8}$	$-1.29 \times 10^{-7}$
Quebec	-0.0501	$1.75 \times 10^{-5}$	$2.71 \times 10^{-5}$

tionship between foliage mass and sapwood area. If hydraulic theory is used, tree height (or more precisely the length of the pathway from the roots to the foliage) should influence the pipe model parameters (e.g., Mencuccini and Grace 1996). However, in our study, foliage mass per sapwood area tended to increase rather than decrease with increasing tree size.

The foliage biomass to sapwood basal area at crown base ratio in Scots pine is inversely proportional to site fertility (Vanninen et al. 1996; Berninger et al. 2005) and average site temperature (Berninger and Nikinmaa 1994; Palmroth et al. 1999). Large variations in the tree foliage biomass to stem basal area at crown base ratio between neighboring plots have also been reported for Scots pine and Norway spruce (*Picea abies* (L.) Karst.; Lehtonen 2005). However, the data used to calibrate the models presented in this paper showed no trends in site fertility (dominant stand height at age 20) with respect to pipe model parameters. Moreover, only a slight trend is observed between average annual temperature and the pipe model parameters (as illustrated by the foliage biomass to sapwood area and branch basal area): the warmer site (New Brunswick) seems to have higher ratios than the cooler one (central Quebec). Such a generalization must, however, be done while keeping in mind that our sampling was restricted to three geographic regions.

Growth rate differences could be used to explain differences in pipe model parameter shifts, as suppressed trees in a given stand, or stems in dense stands, have lower growth rates than dominant trees, or trees in sparser stands. Stems with smaller rings will have higher-density wood (e.g.,

**Fig. 7.** Comparison of the scaled standard errors of each parameter estimated in eqs. A1b–A1d. Scaling of the standard error was done by dividing the standard error by a power of 10 to be on the same scale, the same divider was used for the simultaneous and independent standard error of a given parameter (e.g., standard error of  $\beta_6$  divided by 1000 for both the simultaneous and independent calibration of the models).



Zobel and van Buijtenen 1989). This means that trees will have lower hydraulic conductance, leading to a need for more sapwood per leaf biomass. These trends were also observed in balsam fir (*Abies balsamea* L.), for which radial growth was found to be positively correlated with the ratio of projected leaf area to sapwood basal area (Coyea and Margolis 1992).

The use of the seemingly unrelated regression (SUR) method to calibrate the tree-level models was justified by the strong correlation between the errors (Gallant 1987). Although the parameter estimates in our case were not very different from the independent calibration, the SUR method ensures that they are statistically sound and unbiased. The lack of statistical differences between the parameters estimated by the different fitting techniques (independent vs. SUR) could be in part due to the fact that two models use the same form ( $A_x = \beta_x W_f$ ,  $x = s, b$ ). Nevertheless, the parameter estimates using the SUR method were generally more efficient.

Simultaneous regressions (SUR, 3SLS, etc.), although well known in the forestry literature (Hasenauer et al. 1998; Fang et al. 2001), have not been widely used for allometric equations (Carvalho and Parresol 2003), and our study is, to our knowledge, the first to use mixed models with SUR for allometric relationships. Carvalho and Parresol (2003) used SUR methods for tree biomass estimations, but did not use discrete site and tree effects in their equations. Partial knowledge of site parameters will increase the reliability of our parameter values, as much of the variation in the random effects occurs among sites. Variation among plots in the same regions was small, although this may be a partial artifact. Trees from the

Ontario and New Brunswick sites came from spacing experiments. This implies that within-site variation of site fertility and stand age was negligible among the plots and trees of these sites, and that site fertility and stand density were not correlated. The Quebec values are, however, from different stands and had a wider range for stand age and site fertility, thus representing better regional parameter estimates, as well as a likely greater correlation between site quality and stand densities. The variance explained for the pipe model by the random effects was moderate but not negligible. For most applications, regional allometric relationships might not be necessary (in accordance with Lambert et al. 2005) but the mixed regression allows a flexible calibration of these relationships if one can find that a given site variable will improve the relationships across all sites.

Process-based growth models such as Crobas (Mäkelä 1997) rely on the robustness of the underlying allometric relationships. Pipe model parameters have been found to vary geographically and with stand characteristics (Nikinmaa 1992; Berninger and Nikinmaa 1994; Palmroth et al. 1999; Mencuccini and Bonosi 2001). Numerous studies seem to indicate through empirical evidence that variations in pipe model parameters exist, with no clear theory about the cause of these fluctuations being put forth. With the model results presented in this paper, the robustness of the pipe model parameters for jack pine seems to be a plausible hypothesis given the low importance of the random effects. The picture is, however, different when it comes to estimating tree foliage biomass from crown length. The site random effect was very important, accounting for most of the variance. Furthermore, the plot random effects were found to be nonsignificant. Thus regional relationships might be important in predicting stem foliage biomass relations unless a site-level covariable derived from forest cover maps can be used to absorb some of the site-level random effect. The regional size over which local parameters hold cannot, however, be determined with the available data. Further work would be needed to determine the causes for the observed shifts and the size of the area where the parameters are valid.

The correlation between the error terms can be used to gain information when using the models in a predictive manner (Fang et al. 2001). Moreover, the negative correlation of the error terms indicates that trees with greater than average foliage biomass for a given crown length will have lower than average sapwood area and total branch basal area for a given foliage biomass. In the same manner, the error associated with the ratio of sapwood area to foliage biomass is negatively correlated with the error of branch basal area to foliage biomass ratio. This last correlation is counterbalanced by the highly positive correlation between the plot random effects of both pipe model equations.

In conclusion, the pipe model allometric equations seem to hold well for jack pine, although there were shifts due to stand density. In light of the model variance estimations, local calibration of the pipe model parameters is not required for most applications, whereas local estimation of stem foliage biomass using crown length as predictor might need local efforts. Moreover, the correlation between the different equations should be considered when estimating the different parameters, especially if the forms of the equations dif-

fer. These equations can nevertheless be used to estimate foliage biomass, sapwood area, and branch basal area of a stem using crown length, total stem height, DBH, stem age, and stand density with little error on the prediction.

## Acknowledgements

Financial support was given by the Fonds québécois de recherche sur la nature et les technologies. We also acknowledge the support of Robert Allard, Steve D'eon, Smurfit-Stone Container Canada Inc., and the Petawawa Research Forest. Finally, the reviews, editorial comments, and suggestions of Harry Valentine, Phil Comeau, and Caroline Simpson as well as those of the reviewers were much appreciated.

## References

- Baldwin, V., Jr., Peterson, K., Burkhart, H., Amateis, R., and Dougherty, P. 1997. Equations for estimating loblolly pine branch and foliage weight and surface area distributions. *Can. J. For. Res.* **27**: 918–927. doi:10.1139/cjfr-27-6-918.
- Berninger, F., and Nikinmaa, E. 1994. Foliage area – sapwood area relationships of Scots pine (*Pinus sylvestris*) trees in different climates. *Can. J. For. Res.* **24**: 2263–2268. doi:10.1139/x94-290.
- Berninger, F., Coll, L., Vanninen, P., Mäkelä, A., Palmroth, S., and Nikinmaa, E. 2005. Effects of tree size and position on pipe model ratios in Scots pine. *Can. J. For. Res.* **35**: 1294–1304. doi:10.1139/x05-055.
- Carvalho, J.R., and Parresol, B.R. 2003. Additivity in tree biomass components of Pyrenean oak (*Quercus pyrenaica* Willd.). *For. Ecol. Manage.* **179**: 269–276. doi:10.1016/S0378-1127(02)00549-2.
- Coyea, M., and Margolis, H. 1992. Factors affecting the relationship between sapwood area and leaf area of balsam fir. *Can. J. For. Res.* **22**: 1684–1693. doi:10.1139/x92-222.
- Fang, Z., Bailey, R., and Shiver, B.D. 2001. A multivariate simultaneous prediction system for stand growth and yield with fixed and random effects. *For. Sci.* **47**: 550–562.
- Fortin, M. 2006. Étude rétrospective de la croissance en diamètre du sapin baumier (*Abies balsamea* (L.) Mill.) et de l'épinette rouge (*Picea rubens* Sarg.) en peuplements mixtes après une coupe à diamètre limite. *For. Chron.* **81**: 791–800.
- Gallant, A.R. 1987. Nonlinear statistical models. Wiley series in probability and mathematical statistics. John Wiley and Sons, New York.
- Hall, D.B., and Bailey, R.L. 2001. Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. *For. Sci.* **47**: 311–321.
- Hasenauer, H., Monserud, R.A., and Gregoire, T.G. 1998. Using simultaneous regression techniques with individual-tree growth models. *For. Sci.* **44**: 87–95.
- Lambert, M., Ung, C., and Raulier, F. 2005. Canadian national tree aboveground biomass equations. *Can. J. For. Res.* **35**: 1996–2018. doi:10.1139/x05-112.
- Landsberg, J.J., and Waring, R.H. 1997. A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning. *For. Ecol. Manage.* **95**: 209–228. doi:10.1016/S0378-1127(97)00026-1.
- Lehtonen, A. 2005. Estimating foliage biomass in Scots pine (*Pinus sylvestris*) and Norway spruce (*Picea abies*) plots. *Tree Physiol.* **25**: 803–811. PMID:15870050.
- Littell, R.C., Milliken, G.A., Stroup, W.W., and Wolfinger, R.D. 1996. SAS system for mixed models. SAS Institute Inc., Cary, N.C.

- Mäkelä, A. 1997. A carbon balance model of growth and self-pruning in trees based on structural relationships. *For. Sci.* **43**: 7–24.
- Mäkelä, A., and Albrektson, A. 1992. An analysis of the relationship between foliage biomass and crown surface area in *Pinus sylvestris* in Sweden. *Scand. J. For. Res.* **7**: 297–307.
- Mäkinen, H., and Isomäki, A. 2004a. Thinning intensity and long-term changes in increment and stem form of Scots pine trees. *For. Ecol. Manage.* **203**: 21–34. doi:10.1016/j.foreco.2004.07.028.
- Mäkinen, H., and Isomäki, A. 2004b. Thinning intensity and growth of Scots pine stands in Finland. *For. Ecol. Manage.* **201**: 311–325. doi:10.1016/j.foreco.2004.07.016.
- Mäkelä, A., and Sievänen, R. 1992. Height growth strategies in open-grown trees. *J. Theor. Biol.* **159**: 443–467. doi:10.1016/S0022-5193(05)80690-3.
- Mencuccini, M., and Bonosi, L. 2001. Leaf/sapwood area ratios in Scots pine show acclimation across Europe. *Can. J. For. Res.* **31**: 442–456. doi:10.1139/cjfr-31-3-442.
- Mencuccini, M., and Grace, J. 1996. Developmental patterns of above-ground hydraulic conductance in a Scots pine (*Pinus sylvestris* L.) age sequence. *Plant Cell Environ.* **19**: 939–948. doi:10.1111/j.1365-3040.1996.tb00458.x.
- Monserud, R.A., and Marshall, J.D. 1999. Allometric crown relations in three northern Idaho conifer species. *Can. J. For. Res.* **29**: 521–535. doi:10.1139/cjfr-29-5-521.
- Nikinmaa, E. 1992. Analysis of the growth of Scots pine: matching patterns with function. *Acta For. Fenn.* **235**: 1–68.
- Palmroth, S., Berninger, F., Lloyd, J., Nikinmaa, E., Pulkkinen, P., and Hari, P. 1999. No water conserving behaviour is observed in Scots pine from wet to dry climates. *Oecologia (Berlin)*, **121**: 302–309. doi:10.1007/s004420050932.
- Pinheiro, J.C., and Bates, D.M. 2000. Mixed-effects models in S and S-PLUS. Statistics and Computing Series, Springer-Verlag, New York.
- Pinheiro, J., Bates, D., DebRoy, S., and Sarkar, D. 2005. NLME: linear and nonlinear mixed effects models. R package Version 3.1–56. R Foundation for Statistical Computing, Vienna, Austria. Available from <http://www.R-project.org> [accessed 15 February 2008].
- R Development Core Team. 2004. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from <http://www.R-project.org> [accessed 15 February 2008].
- Shelburne, V.B., Hedden, R.L., and Allen, R.M. 1993. The effects of site, stand density, and sapwood permeability on the relationship between leaf area and sapwood area in loblolly pine (*Pinus taeda* L.). *For. Ecol. Manage.* **58**: 193–209. doi:10.1016/0378-1127(93)90145-D.
- Shinozaki, K.K., Yoda, K., Hozumi, K., and Kira, T. 1964. A quantitative analysis of plant form — the pipe model theory. 1. Basic analyses. *Jpn. J. Ecol.* **14**: 97–105.
- Valentine, H.T., and Mäkelä, A. 2005. Bridging process-based and empirical approaches to modeling tree growth. *Tree Physiol.* **25**: 769–779. PMID:15870047.
- Valentine, H.T., Baldwin, V.C., Gregoire, T.G., and Burkhart, H.E. 1994. Surrogates for foliar dry matter in loblolly pine. *For. Sci.* **40**: 576–585.
- Vanninen, P., and Mäkelä, A. 1999. Fine root biomass of Scots pine stands differing in age and soil fertility in southern Finland. *Tree Physiol.* **19**: 823–830. PMID:10562399.
- Vanninen, P., Ylitalo, H., Sievänen, R., and Mäkelä, A. 1996. Effects of age and site quality on the distribution of biomass in Scots pine (*Pinus sylvestris* L.). *Trees (Berlin)*, **10**: 231–238.
- Vanderklein, D., Martinez-Vilalta, J., Lee, S., and Mencuccini, M. 2007. Plant size, not age, regulates growth and gas exchange in

grafted Scots pine trees. *Tree Physiol.* **27**: 71–79. PMID:17169908.

Zellner, A. 1962. An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias. *J. Am. Stat. Assoc.* **57**: 348–368. doi:10.2307/2281644.

Zhang, S.Y., Chauret, G., Swift, E., and Duchesne, I. 2006. Effects of precommercial thinning on tree growth and lumber quality in jack pine (*Pinus banksiana* Lamb.) stands in New Brunswick, Canada. *Can. J. For. Res.* **36**: 945–952. doi:10.1139/X05-307.

Zobel, B.J., and van Buijtenen, J.P. 1989. Wood variation: its causes and control. Springer-Verlag.

## Appendix A. Seemingly unrelated regression

### The system of equations

The exogenous variables of the system are crown length, DBH, and total tree height in eq. 2, whereas tree foliage biomass ( $\hat{W}_{f,ijk}$ ) is endogenous to the system. Exogenous variables do not depend on equations within the system, whereas endogenous variables are established by one of equations in the system. To account for the correlation between the different equations, the methodology proposed by Fang et al. (2001) was used to carry out a mixed-effects nonlinear seemingly unrelated regression (SUR) estimation of the parameters (cf. Gallant 1987). This was done by writing the equations as a vector (eq. A1a) with the predicted variables identified as groups (eqs. A1b–A1d;  $y_{ijk,1} = \hat{W}_{f,ijk}$ ;  $y_{ijk,2} = A_{s,ijk}$ ;  $y_{ijk,3} = A_{b,ijk}$ ). During the simultaneous calibration, certain site and plot random effects (eq. 2: slope for plot random effect ( $b_{3,ij}$ ), exponent for site and plot random effects ( $b_{4,i}$ ,  $b_{4,ij}$ ); eq. 3: slopes for site random effect in the pipe model parameters  $b_{s,i}$ ,  $b_{b,i}$ ) were found to be statistically nonsignificant and were dropped from the final model (eqs. A1b–A1d).

$$[A1a] \quad \mathbf{Y}_{ijk} = \begin{bmatrix} y_{ijk,1} \\ y_{ijk,2} \\ y_{ijk,3} \end{bmatrix} = \begin{bmatrix} \hat{y}_{ijk,1} \\ \hat{y}_{ijk,2} \\ \hat{y}_{ijk,3} \end{bmatrix} + \varepsilon_{ijk}$$

$$[A1b] \quad \hat{y}_{ijk,1} = (\beta_3 + b_{3,i} + \beta_4 \text{AGE}_{ijk}) H_{c,ijk}^{\beta_5 + \beta_6 \frac{H_{tot}}{DBH}}$$

$$[A1c] \quad \hat{y}_{ijk,2} = (\beta_7 + b_{7,ij} + \beta_8 D_{ij}) y_{ijk,1}$$

$$[A1d] \quad \hat{y}_{ijk,3} = (\beta_9 + b_{9,ij}) y_{ijk,1}$$

$$b_{3,i} \sim N(0, \sigma_1^2)$$

$$[A1e] \quad \begin{bmatrix} b_{7,ij} \\ b_{9,ij} \end{bmatrix} \sim N(0, \Psi_1) \quad \text{where } \Psi_1 = \begin{bmatrix} \sigma_2^2 & \sigma_4 \\ \sigma_4 & \sigma_3^2 \end{bmatrix}$$

$$[A1f] \quad [\varepsilon_{2ijk} \quad \varepsilon_{sijk} \quad \varepsilon_{bijk}]^T | b_i, b_{ij} \sim N(0, \mathbf{R}_{ijk})$$

where:

$$[A1g] \quad \mathbf{R}_{ijk} = \sigma^2 \mathbf{G}_{ijk}^{1/2} (\beta_{ij}, \boldsymbol{\theta}) \boldsymbol{\Gamma}_{ijk}(\rho) \mathbf{G}_{ijk}^{1/2} (\beta_{ij}, \boldsymbol{\theta})$$



$$[A1h] \quad \mathbf{G}_{ijk}(\beta_{ij}, \boldsymbol{\theta}) = \text{diag}\left(f_{\hat{W}_{f,ijk}}^{2\theta_1}, f_{\hat{A}_{s,ijk}}^{2\theta_2}, f_{\hat{A}_{b,ijk}}^{2\theta_3}\right)$$

$$[A1i] \quad (f_{\hat{W}_{f,ijk}}, f_{\hat{A}_{s,ijk}}, f_{\hat{A}_{b,ijk}}) = (E[y_{ijk,1}], E[y_{ijk,2}], E[y_{ijk,3}])$$

$$[A1j] \quad \boldsymbol{\Gamma}_{ijk}(\boldsymbol{\rho}) = \begin{bmatrix} 1 & \rho_{1,2} & \rho_{1,3} \\ & 1 & \rho_{2,3} \\ & & 1 \end{bmatrix}$$

The  $\mathbf{R}$  matrix (eq. A1g) was used to estimate the variance–covariance matrix between the different equations. The correlation between the error terms was modeled by the general correlation matrix  $\boldsymbol{\Gamma}$  (eq. A1j). The variance of each group was estimated by a power of the mean function in the  $\mathbf{G}$  matrix (eqs. A1h, A1i): the matrix is diagonal with the variance of each equation being one element, given by a power of the mean variance function. An underlying assumption of the method is that the errors within the plots for the same equation are independent, that is, that the error for a prediction of a given variable for a tree does not depend on the error of the prediction of another tree.

### Prediction errors

It is well established in the literature how to make predictions with mixed-effects models (Pinheiro and Bates 2000; Fang et al. 2001; Hall and Bailey 2001), and therefore this will not be detailed further here. However, the variance of the predicted variables will be explored to examine the relative importance of the random effects on the total variance of the prediction. The total variance of a prediction using the SUR equations ( $\sigma_{\text{tot}}^2$ ) is the sum of the independent variances (eq. A2).

$$[A2] \quad \sigma_{\text{tot}}^2 = \sigma_{\text{resid}}^2 + \sigma_i^2 + \sigma_j^2$$

where  $\sigma_{\text{resid}}^2$  is the residual variance,  $\sigma_i^2$  is the variance due to the site random effect, and  $\sigma_j^2$  is the variance due to the plot random effect.

The residual variance is estimated for each point by eqs. A1h and A1i using the power of the mean function, and is explicitly given in eq. A3.

$$[A3] \quad \left. \begin{aligned} \sigma_{\text{resid}, W_f}^2 &= \sigma^2 \hat{W}_f^{2\theta_1} \\ \sigma_{\text{resid}, A_s}^2 &= \sigma^2 \hat{A}_s^{2\theta_2} \\ \sigma_{\text{resid}, A_b}^2 &= \sigma^2 \hat{A}_b^{2\theta_3} \end{aligned} \right\}$$

where  $\sigma_{\text{resid}, W_f}^2$ ,  $\sigma_{\text{resid}, A_s}^2$ ,  $\sigma_{\text{resid}, A_b}^2$  are residual variances for the tree foliage biomass, sapwood area at crown base, and total branch basal area of the stem, respectively;  $\hat{W}_f$ ,  $\hat{A}_s$ ,  $\hat{A}_b$  are predicted tree foliage biomass, sapwood area at crown base, and total branch basal area of the stem, respectively;  $\sigma^2$  is the estimated residual variance of the model; and  $\theta_1$ ,  $\theta_2$ , and  $\theta_3$  are estimated variance parameters.

The variances due to the site and plot random effects were calculated using the random-effect design matrix and the variance–covariance matrix (eq. A4). The  $\mathbf{Z}$  matrix is an  $m \times n$  matrix, with  $m$  corresponding to the number of

predictions and  $n$  to the number of random effects in the equation. As all three equations have a single random effect (eqs. A1b–A1d), the  $\mathbf{Z}$  matrix only has one column ( $n = 1$ ) for all the predicted variables. This further implies that the variance–covariance matrix  $\mathbf{G}$  is reduced to a  $1 \times 1$  matrix with no covariance component.

$$[A4] \quad \sigma_{\text{random}}^2 = \tilde{\mathbf{Z}} \mathbf{G} \tilde{\mathbf{Z}}^T$$

where  $\tilde{\mathbf{Z}}$  is the design matrix for the random effects, and  $\mathbf{G}$  is the variance–covariance matrix of the random effects.

The values of the columns in the  $\mathbf{Z}$  matrix are given by numerically evaluating the first derivative of the equation with respect to the random effect (eqs. A5a–A5c).

$$[A5a] \quad \frac{\partial \hat{W}_f}{\partial b_{3,i}} = H_{c,ijk}^{\beta_s + \beta_b \frac{H_{\text{tot}}}{DBH}}$$

$$[A5b] \quad \frac{\partial \hat{A}_s}{\partial b_{7,ij}} = \hat{W}_f$$

$$[A5c] \quad \frac{\partial \hat{A}_b}{\partial b_{9,ij}} = \hat{W}_f$$

The value of the only cell in the  $\mathbf{G}$  matrix is given by the random-effect variance ( $\sigma_1^2$ ,  $\sigma_2^2$ , and  $\sigma_3^2$  in eq. A1e). Equation A2 can thus be rewritten using eqs. A3 to A5 (eq. A6). For the tree foliage biomass, no plot random effects were kept in the final model, as were no site random effects in the sapwood area at crown base and branch basal area models. The same approach can be used to calculate the variance of the branch foliage biomass given in eq. 1.

$$[A6] \quad \left. \begin{aligned} \sigma_{W_f}^2 &= \sigma^2 \hat{W}_f^{2\theta_1} + \sigma_1^2 H_{c,ijk}^{\beta_s + \beta_b \frac{H_{\text{tot}}}{DBH}} + 0 \\ \sigma_{A_s}^2 &= \sigma^2 \hat{A}_s^{2\theta_2} + 0 + \sigma_2^2 \hat{W}_f \\ \sigma_{A_b}^2 &= \sigma^2 \hat{A}_b^{2\theta_3} + 0 + \sigma_3^2 \hat{W}_f \end{aligned} \right\}$$

This development assumes that all the information on the right-hand side of the equations is known. If the total stem foliage biomass is estimated through eq. A1b and inserted into eqs. A1c and A1d, correlation of the errors and error propagation must also be considered. This will not be undertaken here, as it is clearly detailed in Fang et al. (2001).

### References

- Fang, Z., Bailey, R., and Shiver, B.D. 2001. A multivariate simultaneous prediction system for stand growth and yield with fixed and random effects. *For. Sci.* **47**: 550–562.
- Gallant, A.R. 1987. *Nonlinear statistical models*. Wiley series in probability and mathematical statistics. John Wiley and Sons, New York.
- Hall, D.B., and Bailey, R.L. 2001. Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. *For. Sci.* **47**: 311–321.
- Pinheiro, J.C., and Bates, D.M. 2000. *Mixed-effects models in S and S-PLUS*. Statistics and Computing Series, Springer-Verlag, New York.