Modeling Stand Density Effects on Taper for Jack Pine and Black Spruce Plantations Using Dimensional Analysis

Mahadev Sharma and John Parton

Abstract: A taper equation was developed for jack pine and black spruce trees growing at varying density using a dimensional analysis approach. Data used in this study came from stem analysis on 1,135 jack pine (*Pinus banksiana* Lamb.) and 1,189 black spruce (*Picea mariana* [Mill.] B.S.P.) trees sampled from 25 even-aged monospecific plantations in the Canadian boreal forest region of Northern Ontario. About half of the trees were randomly selected for model development, with the remainder used for model evaluation. A nonlinear mixed-effects approach was applied in fitting the taper equation. The predictive accuracy of the model was improved by including random-effects parameters for a new tree based on upper stem diameter measurements. Three scenarios of using upper stem diameter measurements to predict random effects were examined for predictive accuracy: one diameter at any height along the bole; two diameters, one each from below and above breast height; and three diameters, one from below and the other two from above breast height. The upper height at which the diameter was measured was limited to 65% of total tree height for practical reasons. For the first scenario, the model calibrated using a diameter measurement from between 34 and 38% of total height provided the best predictions of inside-bark diameters. For the second scenario, the model calibrated using one diameter from near the stump and the other from close to 65% of total height produced the least bias in predicting inside-bark diameters. For the third scenario, the model calibrated using the diameters from near the stump and at approximately 35 and 65% of total height provided the highest prediction accuracy. FOR. SCI. 55(3):268–282.

Keywords: *Picea mariana*, *Pinus banksiana*, tree form, stem profile models, stand density, thinning, nonlinear regression, fixed response, calibrated response, mixed-effects model

Development of taper equations is a basic prerequisite to estimating individual tree volumes and product yields. These equations are used to estimate diameters along the bole at any given height. Individual tree volume can then be calculated based on these diameters and corresponding heights. This is important because product recovery from different trees with the same total volume could be substantially different. Two trees with different shape, conic versus cylindrical, for example, will have different product recoveries and hence significantly different economic value.

The shape of a tree is influenced by stand density (Gray 1956, Larson 1963, Sharma and Zhang 2004). Similarly, stem form may differ among tree species growing in the same environment and the same stand conditions. For example, Sharma and Zhang (2004) reported that taper profiles for jack pine (*Pinus banksiana* Lamb.), black spruce (*Picea mariana* [Mill.] B.S.P.), and balsam fir (*Abies balsamea* [L.] Mill.) trees grown in natural stands in eastern Canada differed significantly. They further reported that the stem form also differed for black spruce trees grown in natural stands at different stand densities.

Stand density can be regulated either by planting trees at different initial spacings or by thinning stands to different densities. However, trees of a particular species grown in a plantation versus those grown in a natural stand thinned to the same density may not have the same form, especially if the thinning occurs at a later age (Sharma and Zhang 2004). As a result, tree boles cannot be completely described in simple mathematical terms. In attempts to describe tree taper, numerous models of varying complexity have been advanced. Three main approaches are applied in advancing these models. Under the first approach, tree taper is described by a simple mathematical function. Examples of such simple taper functions are those presented by Kozak et al. (1969), Ormerod (1973), Amidon (1984), Reed and Byrne (1985), Sharma and Oderwald (2001), and Sharma et al. (2002).

Under the second approach, segments of a tree stem are approximated by various geometric solids. The lower bole portion is assumed to be a neiloid frustum, the middle portion a paraboloid frustum, and the upper portion a cone (Avery and Burkhart 2002, p. 101). Examples of this approach are the segmented polynomial taper models developed by Max and Burkhart (1976), Demaerschalk and Kozak (1977), Cao et al. (1980), and Fang et al. (2000).

In the third approach, tree form is allowed to vary from one point to another along the bole, and the variable form is described by a single continuous function. Examples of this approach are the variable-exponent taper models elaborated by Kozak (1988) and the variable-form taper model presented by Newnham (1992). In these models, a single

Mahadev Sharma, Research Scientist, Ministry of Natural Resources, Ontario Forest Research Institute, 1235 Queen Street East, Sault Ste Marie, ON P6A 2E5, Canada—Phone: (705) 946-7407; Fax: (705) 946-2030; mahadev.sharma@ontario.ca. John Parton, Provincial Growth and Yield Coordinator, Ministry of Natural Resources, Timmins, ON P0N 1H0, Canada—john.parton@ontario.ca.

Acknowledgments: This study was supported by the Forestry Futures Trust Enhanced Forest Productivity Science Program and the Ontario Ministry of Natural Resources. We are grateful to Peter Newton for his assistance in using the WinDendro program for stem analysis and to Lisa Buse (Ontario Forest Research Institute) for editing a previous version of the manuscript.

Manuscript received September 12, 2008, accepted March 18, 2009 Copyright © 2009 by the Society of American Foresters

continuous function with an exponent changing from stump to top describes the neiloid, paraboloid, and several intermediate forms (Kozak 1997). Sharma and Zhang (2004) also presented a variable exponent taper equation derived using dimensional analysis. Flewelling and Raynes (1993) developed a variable-form taper model based on a system of three equations.

In most of these equations, taper is modeled in terms of dbh and total height. A few researchers have considered using crown dimensions (e.g., crown height, ratio, and diameters) as covariates (e.g., Newnham 1992), but most reported little or no improvement in resulting model performance (Burkhart and Walton 1985, Valenti and Cao 1986, Muhairwe et al. 1994). Furthermore, Leites and Robinson (2004) concluded that operational costs involved in measuring crown dimensions of standing trees might limit their use. Additional terms based on upper stem diameter measurements have been included in models in other studies (Czaplewski and McClure 1988, Rustagi and Loveless 1991, Flewelling and Raynes 1993, Kozak 1998). However, improvements in estimates were minor and dependent on the precision of the upper stem diameter measurement (Kozak 1998).

Foresters know that individual tree crown dynamics and stem form are influenced by site and stand conditions. Therefore, it makes sense that including stand density information in taper equations should improve model performance. Stand density is also more easily obtained than individual tree crown dimensions. In fact, Sharma and Zhang (2004) included stand density (trees ha^{-1}) information in modeling the taper of black spruce trees and reported improved fit statistics and predictive accuracy. However, the stand density information available to them was insufficient (three stand densities only) for making a general inference.

Data required to develop taper equations generally originate from stem analysis. These data are commonly hierarchical, with multiple measurements from individual trees resulting in a correlation among data points that is typically not removed by model fitting. As a result, residuals within a tree may be correlated in a predictable way (Leites and Robinson 2004). This correlation violates the assumption of independent observations that is the key to obtaining an unbiased estimate of the covariance matrix in regression (Valentine and Gregoire 2001).

To address this problem, recent studies have used a mixed-effects modeling technique (Fang and Bailey 2001, Leites and Robinson 2004, Trincado and Burkhart 2006). This technique has the advantage of correctly estimating the covariance matrix of correlated data and contains both fixed- and random-effects parameters in the model (Schabenberger and Pierce 2001). Fixed-effects parameters are a population average response common to all sampling units (trees) and random-effects parameters are localized responses that are specific to each sampling unit. Thus, mixed-effects models will have improved predictive accuracy when the random-effects parameters can be estimated for an unsampled location (Calama and Montero 2004). The objectives of this study were to examine the effect of stand density on taper of plantation grown jack pine and black spruce trees and to develop taper equations that incorporate stand density information for these tree species using mixed-effects modeling techniques.

Data

The data used in this study came from stem analysis of plantation-grown jack pine and black spruce trees. For both species, 25 even-aged monospecific plantations from throughout the Canadian boreal forest region of Northern Ontario were sampled. Within each plantation, three variable size circular temporary sample plots were established. The minimum plot size was set at 400 m^2 but was increased if necessary to obtain a minimum of 80 trees of the target species.

All live trees in the plot were measured, regardless of species, using Ontario's growth and yield standards (Hayden et al. 1995). Total basal area (BA/ha) and density (trees/ha) were then calculated for the plot for each species. All target species trees were sequentially numbered, and the cumulative basal area for this species was determined. Total cumulative basal area of target species trees was then divided into five BA/ha classes. Three trees that were classified as planted and did not exhibit any visible deformities, such as forks, major stem injuries, or dead or broken tops, were randomly selected from each basal area class for destructive sampling. In addition, the largest diameter tree was also selected from most of the plots. Thus, a minimum of 15 trees were sampled from each plot (temporary sample plot), which resulted in 45 to 48 trees per species from each site. In all, 1,135 jack pine and 1,189 black spruce trees were sampled from 25 sites across northern Ontario. Summary statistics for tree and stand characteristics are presented in Table 1.

Disks were cut at 0.15, 0.5, 0.9, and 1.3 m (breast height) from each sample tree. The remaining height of the tree was then divided based on one of two sampling schemes: 5 and 10% of relative height. For the 5% of relative height scheme, the remaining height was divided by 20 and the disks were cut at that interval from the breast height; similarly, for the 10% of relative height sampling scheme, the remaining height was divided by 10 and the disks were cut at the resulting interval. Twenty percent of sampled trees from each site were selected for disk sampling at 5% of relative height; the remainder were cut using the 10% of relative height sampling scheme. For each tree, this resulted in 23 disks for the 5% and 13 disks for the 10% sampling scheme.

Each sampled tree and disk were given a unique code. All disks from a tree were placed in a large breathable bag, transported, and stored at –10°C until 24 hours before preparation. Geometric mean radius was then calculated from the diameters obtained from the major (r_1) and minor (r_2) axes on each disk [i.e., $r = (r_1 \cdot r_2)^{0.5}$]. This resulted in 18,002 observations for jack pine and 18,852 observations for black spruce trees.

Of the 1,135 jack pine and 1,189 black spruce trees, about half (568 and 600 trees from jack pine and black spruce, respectively) were randomly selected for model

Table 1. Summary statistics for measured characteristics of plantation jack pine and black spruce trees from boreal Ontario used in this study

Variable	Frequency	Mean	SD	Minimum	Maximum
Jack pine					
BA (m^2 ha^{-1})	75	27.46	5.78	15.28	42.25
Density (trees ha^{-1}	75	1773	647	884	3302
QMD (cm)	75	14.46	2.01	10.62	19.14
dbh (cm)	1,135	17.34	4.46	6.10	34.30
Height (m)	1,135	15.47	2.54	7.93	23.17
CR	1,135	0.430	0.113	0.099	0.845
Black spruce					
BA (m^2 ha^{-1})	75	29.84	8.79	12.00	48.87
Density (trees ha^{-1})	75	2,919	896	1,471	5,579
QMD (cm)	75	11.67	2.41	6.37	16.00
dbh (cm)	1,189	13.35	3.70	2.50	24.80
Height (m)	1,189	10.85	2.47	2.98	17.85
CR	1,189	0.600	0.155	0.222	0.977

QMD, quadratic mean diameter; CR, crown ratio.

development (model data set) and the rest were reserved for model evaluation (evaluation data set).

Taper Equations

Sharma and Zhang (2004) derived a variable exponent taper equation based on a dimensionally compatible taper equation originally presented by Sharma and Oderwald (2001). They calibrated this model for jack pine, black spruce, and balsam fir trees grown in natural stands in eastern Canada and reported that this model was superior to the segmented polynomial, variable-exponent, and variable-form taper equations by Max and Burkhart (1976), Kozak (1988), and Zakrzewski (1999), respectively, in estimating tree diameters along the bole of these tree species. Their model was

$$\left(\frac{d}{D}\right)^2 = \beta_0\left(\frac{h}{h_D}\right)^{2-(\beta_1+\beta_2 x+\beta_3 x^2)}\left(\frac{H-h}{H-h_D}\right), \qquad (1)$$

where d is diameter inside bark at any given height h (m), D is dbh outside bark (cm), H is total tree height from ground to tip (m), h_D is breast height (m), x is h/H, and β_i (i = 0, 1, 2, and 3) are parameters.

Similarly, Newton and Sharma (2008) evaluated this equation for the sensitivity of different disk selection protocols using jack pine tree data collected for this study along with the equations by Max and Burkhart (1976) and Kozak (1988). They found that Equation 1 was invariant to disk selection protocols for estimating inside bark diameters and total volume of plantation jack pine trees. Therefore, this equation was fitted first to the stem analysis data collected from plantation jack pine and black spruce trees and examined for fit statistics (R^2 and mean square error) and predictive accuracy. The coefficient of determination (R^2) was 0.97 for both species and the mean square error was 0.00387 for jack pine and 0.00413 for black spruce trees. However, when these models were used to predict the diameters of the trees set aside for validation, diameters at >75% of total heights were slightly overpredicted.

The taper of trees grown in plantations was then compared with that of trees from natural stands. Results indicated that trees in plantation stands tapered more than those in natural stands. Similarly, the tree form was less parabolic in plantations than in natural stands. Therefore, the original assumption about the shape of a tree made by Sharma and Oderwald (2001) in developing a dimensionally compatible taper equation was modified and a new taper equation was developed using dimensional analysis.

Sharma and Oderwald (2001) assumed the following mathematical form to describe the overall shape of a tree:

$$d^2 = \alpha D^{\delta}\left(1-\frac{h}{H}\right)h^{\gamma},$$

where, α, δ, and γ are parameters and other variables are as defined for Equation 1. To make tree shape less parabolic, the following mathematical form was assumed:

$$\frac{d}{D} = \alpha\left(1-\frac{h}{H}\right)\left(\frac{h}{H}\right)^{\theta}, \qquad (2)$$

where $\theta = f(h)$ so that the taper equation becomes the variable exponent as the height changes from the ground to the tip of the tree and describes the various shapes along the bole. The equation is dimensionless because the dimensions of the numerator cancel those in the denominator on both sides of the equation. The constant, α, ensures that if h is breast height, h_D, then $d = D$. The term $(1 - h/H)$ ensures that if $h = H$, then $d = 0$. The constant, α, in Equation 2 can be calculated by applying the constraint; $d = D$ when $h = h_D$, where both d and D are inside or outside bark. This results in

$$\alpha = \left(\frac{H}{H-h_D}\right)\left(\frac{H}{h_D}\right)^{\theta}.$$

Substituting this value of α in Equation 2 and rearranging the terms we obtain

$$\frac{d}{D} = \left(\frac{H-h}{H-h_D}\right)\left(\frac{h}{h_D}\right)^{\theta}.$$

As mentioned earlier, both d and D in this equation are inside or outside bark. As a result, this equation cannot be used to estimate inside bark diameter using the dbh-measured outside bark. However, the equation can be modified to estimate the inside bark diameters using outside bark dbh by including a constant φ as

$$\frac{d}{D} = \varphi\left(\frac{H-h}{H-h_D}\right)\left(\frac{h}{h_D}\right)^{\theta}. \tag{3}$$

The constant φ can be calculated as

$$\varphi = \frac{D_i}{D},$$

where D_i is the inside bark dbh. Equation 3 is still dimensionless if $\theta = \mathrm{f}(h)$ is dimensionless. Sharma and Zhang (2004) expressed θ as a quadratic function of h/H in their taper equation and reported that the equation with this function accurately described the taper of natural stand jack pine, black spruce, and balsam fir trees. This function, i.e.,

$$\theta = \mathrm{f}(h) = \beta_1 + \beta_2(h/H) + \beta_3(h/H)^2,$$

where β_1, β_2, and β_3 are parameters, was also considered in this study. Because all terms in the expression are dimensionless, θ is also a dimensionless quantity. By substituting this value of θ with $x = h/H$ and replacing φ by β_0, Equation 3 becomes

$$\frac{d}{D} = \beta_0\left(\frac{H-h}{H-h_D}\right)\left(\frac{h}{h_D}\right)^{\beta_1+\beta_2 x+\beta_3 x^2}. \tag{4}$$

In this equation, the exponent $\theta = \beta_1 + \beta_2 x + \beta_3 x^2$ is the only term that determines the change in taper from one point to another as h increases along the bole. Therefore, the density effect on taper can be determined by incorporating the stand density information into the exponent as

$$\frac{d}{D} = \beta_0\left(\frac{H-h}{H-h_D}\right)\left(\frac{h}{h_D}\right)^{\beta_1+\beta_2 x+\beta_3 x^2+\beta_4 \mathrm{f(sd)}}. \tag{5}$$

where f(sd) is a function of stand density and β_4 is a parameter. To preserve the dimensionless property, the function f(sd) in Equation 5 should be dimensionless. Some options are to use (1) quadratic mean diameter (QMD)/D because QMD is determined by BA (BA/ha) and trees/ha, (2) BA/D^2, (3) $\sqrt{BA}/D$, (4) trees/ha, and functions associated with these terms. Because all of these terms are dimensionless, substituting any function of these terms in Equation 5 will ensure dimensional compatibility with the model (see Sharma and Oderwald 2002 for details on dimensional analysis applied to developing taper equations). A preliminary analysis indicated that f(sd) $= \sqrt{BA}/D$ best described the density effect on taper using Equation 5. Thus, the final taper equation that includes stand density information can be written as

$$\frac{d}{D} = \beta_0\left(\frac{H-h}{H-h_D}\right)\left(\frac{h}{h_D}\right)^{\beta_1+\beta_2 x+\beta_3 x^2+\beta_4(\sqrt{BA}/D)}. \tag{6}$$

It should be noted here that if outside bark diameter predictions are of interest β_0 should equal 1.

Nonlinear Mixed-Effects Variable Exponent Taper Equation

As Trincado and Burkhart (2006) pointed out, a nonlinear mixed-effects model can be written in the form of a two-stage model explicitly specifying within- and between-tree variations. An expression for the taper equation representing systematic and random within-tree variation associated with the jth ($j = 1, \ldots, n_i$) observation (diameter measurement) along the bole for the ith individual ($i = 1, \ldots, n$) is

$$y_{ij} = \beta_{0i}\left(\frac{H_i-h_{ij}}{H_i-h_D}\right)\left(\frac{h_{ij}}{h_D}\right)^{\beta_{1i}+\beta_{2i}x_{ij}+\beta_{3i}x_{ij}^2+\beta_{4i}(\sqrt{BA}/D_i)} + e_{ij}, \tag{7}$$

where $y_{ij} = d_{ij}/D_i$, $x_{ij} = h_{ij}/H_i$, $\mathrm{E}(e_{ij}) = 0$, and $\beta_{0i} - \beta_{4i}$ are parameters for the ith individual. Similarly, h_D is the breast height that is always constant (1.3 m) and does not change from one tree to another. This model can be generalized into a vector form as

$$\mathbf{y}_i = \beta_{0i}\left(\frac{H_i-\mathbf{h}_i}{H_i-h_D}\right)\left(\frac{\mathbf{h}_i}{h_D}\right)^{\beta_{1i}+\beta_{2i}\mathbf{x}_i+\beta_{3i}\mathbf{x}_i^2+\beta_{4i}(\sqrt{BA}/D_i)} + e_i, \tag{8}$$

where $\mathbf{y}_i = [y_{i1}, {}_{i2}, \ldots, y_{in_i}]^{\mathrm{T}}$, $\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{in_i}]^{\mathrm{T}}$, $\mathbf{h}_i = [h_{i1}, h_{i2}, \ldots, h_{in_i}]^{\mathrm{T}}$, and the conditional distribution of $\mathbf{e}_i$ given β_i is assumed to be multivariate normally distributed with $\mathrm{E}(\mathbf{e}_i) = \mathbf{0}$ and variance-covariance matrix $\mathbf{R}_i(\boldsymbol{\beta}_i, \boldsymbol{\xi})$. The vector $\boldsymbol{\xi}$ represents a vector of unknown parameters $[\sigma, \boldsymbol{\theta}', \boldsymbol{\rho}']^{\mathrm{T}}$ common for all individuals. In this case, within-tree systematic variation is characterized through the equation, and random variation is described by the distribution of the error term. If needed, the variance-covariance matrix can be expanded in a more general form to account for within-tree variance and autocorrelation (Trincado and Burkhart 2006) as

$$\mathbf{R}_i(\beta_i, \xi) = \sigma^2\mathbf{G}_i^{1/2}\Gamma_i\mathbf{G}_i^{1/2}, \tag{9}$$

where $\mathbf{G}_i$ is an $(n_i \times n_i)$ diagonal matrix that describes nonhomogeneous variance for a given tree i with n_i height-diameter measurements (SDs of the residual errors are its components), $\boldsymbol{\Gamma}_i$ is an $(n_i \times n_i)$ matrix that shows the structure of the correlation among observations for tree i, and σ^2 is a scaling factor for the error dispersion (Gregoire et al. 1995), which is the value of the residual variance of the model. If we assume that within-tree variance is homogeneous and residuals are uncorrelated, Equation 9 can be simplified as

$$\mathbf{R}_i(\boldsymbol{\beta}_i, \xi) = \sigma^2\mathbf{I}_{ni},$$

where $\mathbf{I}_{ni}$ is the identity matrix of dimensions $(n_i \times n_i)$. Errors resulting from Equation 7 will be analyzed and a more general structure, if required, will be used by incorporating the effects of within- and between-tree heterogeneous variance and correlation between residuals.

In the case of between-tree variation, the parameter vector $\boldsymbol{\beta}_i$ varies from tree to tree and hence accounts for this

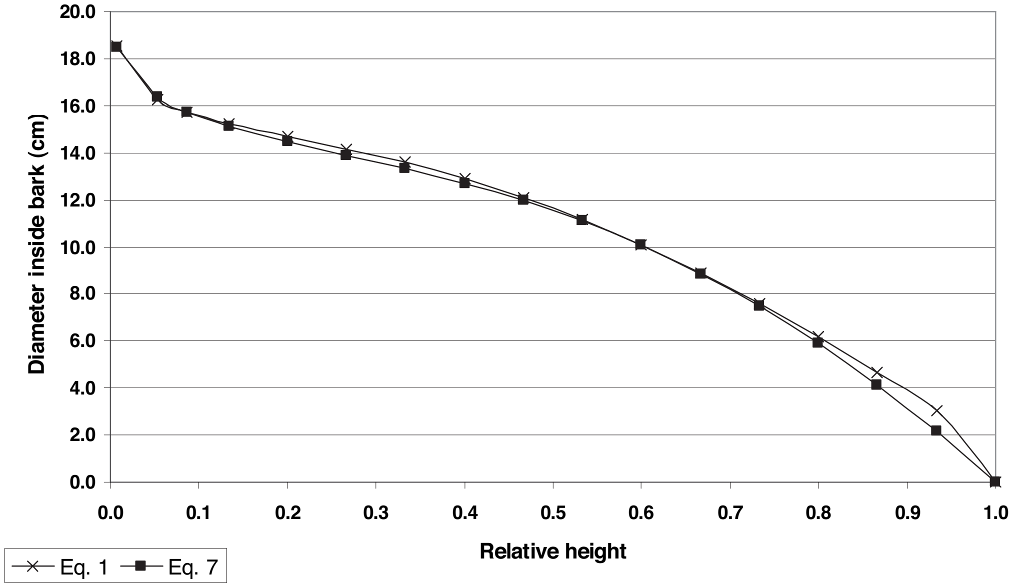


Figure 1. Tree profiles generated based on the same dbh (17.0 cm) and total height (15.0 m) values for jack pine trees using Equation 1 and Equation 7 without stand density term.

variation. Systematic and random variation in the vector canbe expressed explicitly (Pinheiro and Bates 1995, Vonesh and Chinchilli 1997, p. 319, Calama and Montero 2004, Trincado and Burkhart 2006) as

$$\boldsymbol{\beta}_i = \mathbf{A}_i\boldsymbol{\beta} + \mathbf{B}_i\mathbf{b}_i,$$

where $\boldsymbol{\beta}$ is the 5×1 vector of fixed population parameters, $\mathbf{b}_i$ is the 5×1 vector of random effects associated with the ith tree, and $\mathbf{A}_i$ and $\mathbf{B}_i$ are design matrices for the fixed and random effects specific to each tree, respectively. The vector of random effects $\mathbf{b}_i$ is assumed to be multivariate normally distributed with $E(\mathbf{b}_i) = 0$ and variance-covariance matrix $\mathbf{D}$, i.e., $\mathbf{b}_i \sim N(0, \boldsymbol{D})$. If we assume all parameters have both fixed- and random-effects components, the vector $\boldsymbol{\beta}_i$ for the ith individual can be written as

$$\boldsymbol{\beta}_i = \begin{pmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\\ \beta_4\end{pmatrix} + \begin{pmatrix}b_{0i}\\ b_{1i}\\ b_{2i}\\ b_{3i}\\ b_{4i}\end{pmatrix} = \begin{pmatrix}\beta_0 + b_{0i}\\ \beta_1 + b_{1i}\\ \beta_2 + b_{2i}\\ \beta_3 + b_{3i}\\ \beta_4 + b_{4i}\end{pmatrix} = \begin{pmatrix}\beta_{0i}\\ \beta_{1i}\\ \beta_{2i}\\ \beta_{3i}\\ \beta_{4i}\end{pmatrix}.$$

To examine the effect of stand density on taper and to determine the number of random-effects parameters needed in the model, Equation 7 was initially fitted without the density term and random-effects parameters using the NLMIXED procedure in SAS (SAS Institute, Inc. 2004). This was performed by assuming that within-tree variance is homogeneous and residuals are uncorrelated. Random-effects parameters were then added sequentially starting at β_0.

Table 2. Fit statistics for the variable exponent taper equation 7 for different combinations of random-effects parameters without and with stand density term (β_4) for jack pine and black spruce trees from boreal Ontario

Parameters	No. parameters (k)	$\hat{\sigma}^2$	$-2\ \ln(L)$[1]	AIC[1]
Jack pine				
$\beta_0, \beta_1, \beta_2, \beta_3$	5	0.001847	−31023	−31013
$\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$	6	0.001709	−31721	−31709
$\beta_{0i}, \beta_1, \beta_2, \beta_3$	6	0.001315	−32945	−32933
$\beta_{0i}, \beta_{1i}, \beta_2, \beta_3$	8	0.000866	−35288	−35272
$\beta_{0i}, \beta_{1i}, \beta_{2i}, \beta_3$	11	0.000559	−37746	−37724
$\beta_{0i}, \beta_{1i}, \beta_{2i}, \beta_{3i}$	15	0.000390	−39503	−39473
$\beta_{0i}, \beta_{1i}, \beta_{2i}, \beta_{3i}, \beta_4$	16	0.000390	−39614	−39582
Black spruce				
$\beta_0, \beta_1, \beta_2, \beta_3$	5	0.001723	−33655	−33645
$\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$	6	0.001552	−34658	−34646
$\beta_{0i}, \beta_1, \beta_2, \beta_3$	6	0.001081	−36695	−36683
$\beta_{0i}, \beta_{1i}, \beta_2, \beta_3$	8	0.000562	−40942	−40926
$\beta_{0i}, \beta_{1i}, \beta_{2i}, \beta_3$	11	0.000343	−43880	−43858
$\beta_{0i}, \beta_{1i}, \beta_{2i}, \beta_{3i}$	15	0.000255	−45184	−45154
$\beta_{0i}, \beta_{1i}, \beta_{2i}, \beta_{3i}, \beta_4$	16	0.000255	−45334	−45302

$\beta_{0i} = \beta_0 + b_{0i}$, $\beta_{1i} = \beta_1 + b_{1i}$, $\beta_{2i} = \beta_2 + b_{2i}$, and $\beta_{3i} = \beta_3 + b_{3i}$.
[1] Smaller is better.

Table 3. **Parameter estimates and fit statistics for equation 7 fitted for jack pine and black spruce trees from boreal Ontario**

	Jack pine		Black spruce	
	Estimates	SE	Estimates	SE
Parameters				
β_0	0.92230	0.00108	0.90880	0.00127
β_1	−0.05997	0.00251	−0.06670	0.00266
β_2	0.51560	0.00746	0.54100	0.00741
β_3	−0.22650	0.01026	−0.36360	0.00996
β_4	0.08383	0.00756	0.07549	0.00578
Variance components				
σ^2	0.000390	0.000006	0.000255	0.000004
var (b_0)	0.000558	0.000040	0.000900	0.000056
var (b_1)	0.000314	0.000024	0.000442	0.000029
var (b_2)	0.025240	0.001879	0.026450	0.001929
var (b_3)	0.049940	0.003568	0.048390	0.003525
cov(b_0, b_1)	−0.00007	0.000022	−0.00006	0.000030
cov(b_0, b_2)	0.000397	0.000194	0.000181	0.000244
cov(b_0, b_3)	−0.00059	0.000266	−0.00036	0.000325
cov(b_1, b_2)	−0.00138	0.000169	−0.00142	0.000182
cov(b_1, b_3)	0.001393	0.000221	0.001338	0.000235
cov(b_2, b_3)	−0.03211	0.002476	−0.03122	0.002460

Finally, the density term (β_4) was added in the presence of all four random-effects parameters (β_0–β_3). Models with different numbers of random-effects parameters and the density effects were evaluated based on goodness-of-fit criteria: i.e., twice the negative log-likelihood [$-2\ln(L)$] and Akaike's information criterion (AIC) defined as

$$\text{AIC} = -2\ln(L) + 2k,$$

where L is the likelihood function and k is the number of parameters in the model. The model with the smallest values for the goodness-of-fit criteria was considered to be the best.

Prediction of Upper Stem Diameters for a New Tree

The main purpose in developing a model is to predict the dependent variable (in this case diameter along the bole) in terms of independent variables (in this case dbh, H, and stand density) through the relationship specified in the model. In the mixed-effects modeling approach, diameter can be predicted by assuming that the random parameters are zero if no prior information is available (fixed-effects response) and predicting the random parameters for the tree for which upper stem diameter information is available for a subsample of disks (calibrated response).

For a fixed-effects response, the predicted diameter represents the mean behavior of the pattern of variation in diameter for given dbh, H, and stand density. Therefore, the diameters are predicted using the expression

$$\hat{d}_{ij} = \mathrm{f}(A_i\hat{\beta}, x_{ij}),$$

where $\mathbf{A}_i$ and $\hat{\boldsymbol{\beta}}$ are design matrix and the estimated vector for the fixed effects, respectively; and x_{ij} and $\hat{d}_{ij}$ are the vector of the independent variables and the predicted diameter, respectively, for the jth diameter in the ith tree.

For the calibrated response, however, model parameters are localized first by using the predicted values of the random parameters for each tree. Height-diameter pairs from a subsample of a tree are used along with stand density (i.e., independent variables in Equation 7) to predict the random parameters for that tree. The following expression can be used to predict the random parameters (Vonesh and Chinchilli 1997, p. 362):

$$\hat{b}_i = \hat{D}\hat{Z}_i^{\mathrm{T}}(\hat{R}_i + \hat{Z}_i\hat{D}\hat{Z}_i^{\mathrm{T}})^{-1}\hat{e}_i, \qquad (10)$$

where $\hat{\mathbf{D}}$ is the $q \times q$ variance–covariance matrix (q is the number of random-effects parameters included in model, 4 in this case) for the among-tree variability, $\hat{\mathbf{R}}_i$ is the $k \times k$ variance–covariance matrix for tree i, $\hat{\mathbf{e}}_i$ is the residual vector $k \times 1$ with components $\hat{e}_{ij}$ defined as

$$\hat{e}_{ij} = d_{ij} - \mathrm{f}(A_i\hat{\beta}, x_{ij}),$$

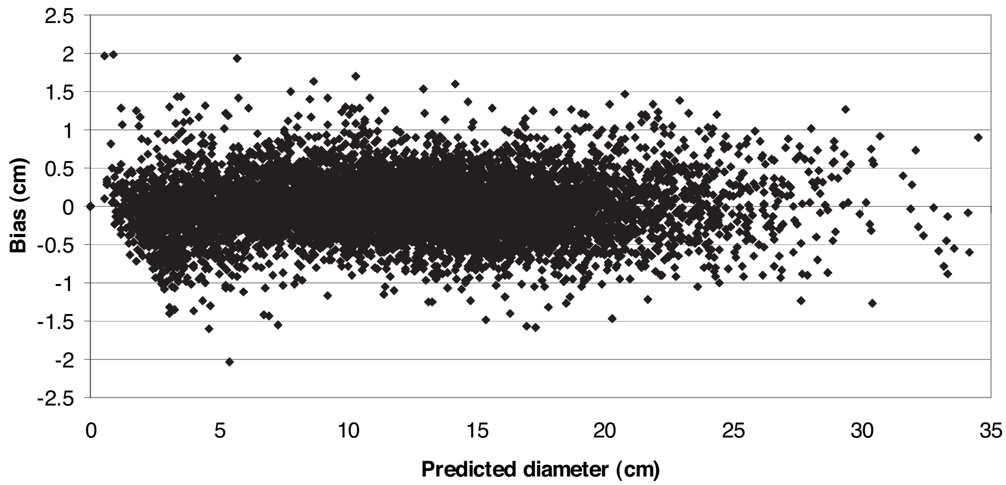
where d_{ij} = observed diameter at the jth height in the subsample from tree i, $\hat{\mathbf{Z}}_i$ is the $k \times q$ matrix evaluated at $\hat{\boldsymbol{\beta}}$ as

$$\begin{pmatrix} \dfrac{\partial \mathrm{f}(x_{i1}, \hat{\boldsymbol{\Phi}}_i)}{\partial \beta_1} \\ \vdots \\ \dfrac{\partial \mathrm{f}(x_{ik}, \hat{\boldsymbol{\Phi}}_i)}{\partial \beta_1} \end{pmatrix} \begin{matrix} \cdots \\ \\ \cdots \end{matrix} \begin{pmatrix} \dfrac{\partial \mathrm{f}(x_{i1}, \hat{\boldsymbol{\Phi}}_i)}{\partial \beta_q} \\ \vdots \\ \dfrac{\partial \mathrm{f}(x_{ik}, \hat{\boldsymbol{\Phi}}_i)}{\partial \beta_q} \end{pmatrix},$$

where $\hat{\boldsymbol{\Phi}}i = A_i\hat{\boldsymbol{\beta}}$, $\beta_1, \ldots, \beta_q$ are the fixed part of the mixed coefficients components of the vector for estimated fixed-effects $\hat{\boldsymbol{\beta}}$, and x_{ij} is the vector of independent variables corresponding to the jth diameter in the subsample of the ith tree. The predicted random effects are added to the fixed parameters to obtain localized parameters. Upper stem diameters are then predicted in terms of dbh, H, stand density, and localized parameters. Details on the prediction of random-effects parameters in a forestry context can be found in studies by Calama and Montero (2004) and Trincado and Burkhart (2006).

Prediction accuracies of the models with and without the random effect were compared by examining the bias and its SD for fixed-effects and calibrated responses along the

(a)



(b)

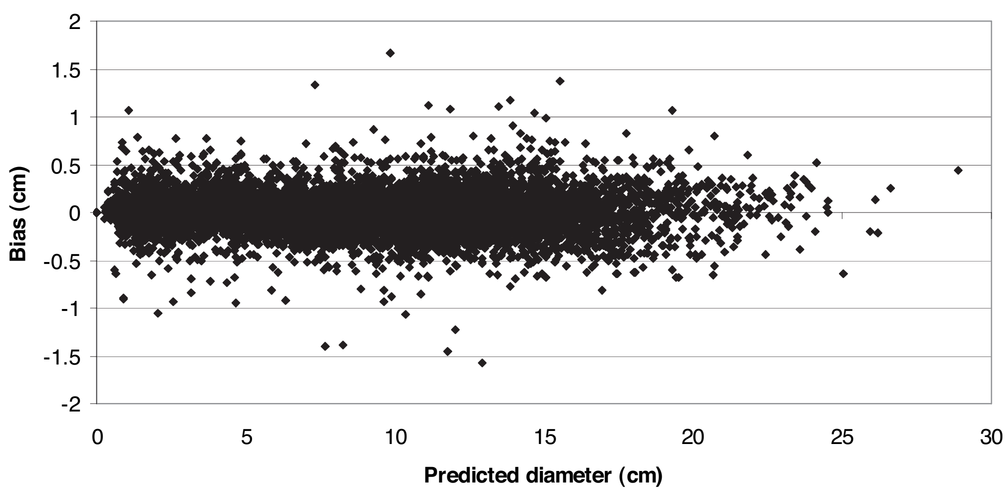


Figure 2. Bias (observed − predicted) in predicting diameters inside bark of (a) jack pine (b) black spruce trees using Equation 7.

boles of trees in the validation data set. Unless otherwise specified, the level of significance used throughout this article is 0.05 ($\alpha = 5\%$).

Results and Discussion

To examine whether the taper equation derived here describes the shape of plantation-grown jack pine and black spruce trees, Equation 7 was initially fitted without the stand density term and random-effects parameters to compare with results using Equation 1. Taper profiles were generated using these equations and identical values for dbh (17 cm) and total height (15 m) for jack pine trees (Figure 1). Diameters estimated by Equation 7 were smaller than those estimated by Equation 1 at >70% of total tree height. Similar results were obtained for black spruce trees (not shown). This finding indicates that Equation 7 adequately describes the shape of the trees used in this study. As mentioned above, Equation 1 overestimated diameters at >75% of total tree height for both species.

Equation 7 was then fitted, including random-effects parameters in the model. Initially, the model was fitted without the stand density term and was assumed to have β_0 as the only parameter associated with random effects (i.e., $\beta_{0i} = \beta_0 + b_{0i}$). In the second step, β_1 was also assumed to be associated with the random-effects parameter (i.e., $\beta_{1i} = \beta_1 + b_{1i}$) in addition to β_{0i}. Next, β_2 was associated with the random effects (i.e., $\beta_{2i} = \beta_2 + b_{2i}$) in addition to β_{0i} and β_{1i} and so on. Finally, the stand density term was added with its coefficient (β_4) as a fixed-effects parameter. Table 2 displays the goodness-of-fit statistics for Equation 7 for all these scenarios including the model with all the parameters (β_0–β_4) as fixed effects for both species, illustrating the extent to which inclusion of random-effects parameters improved

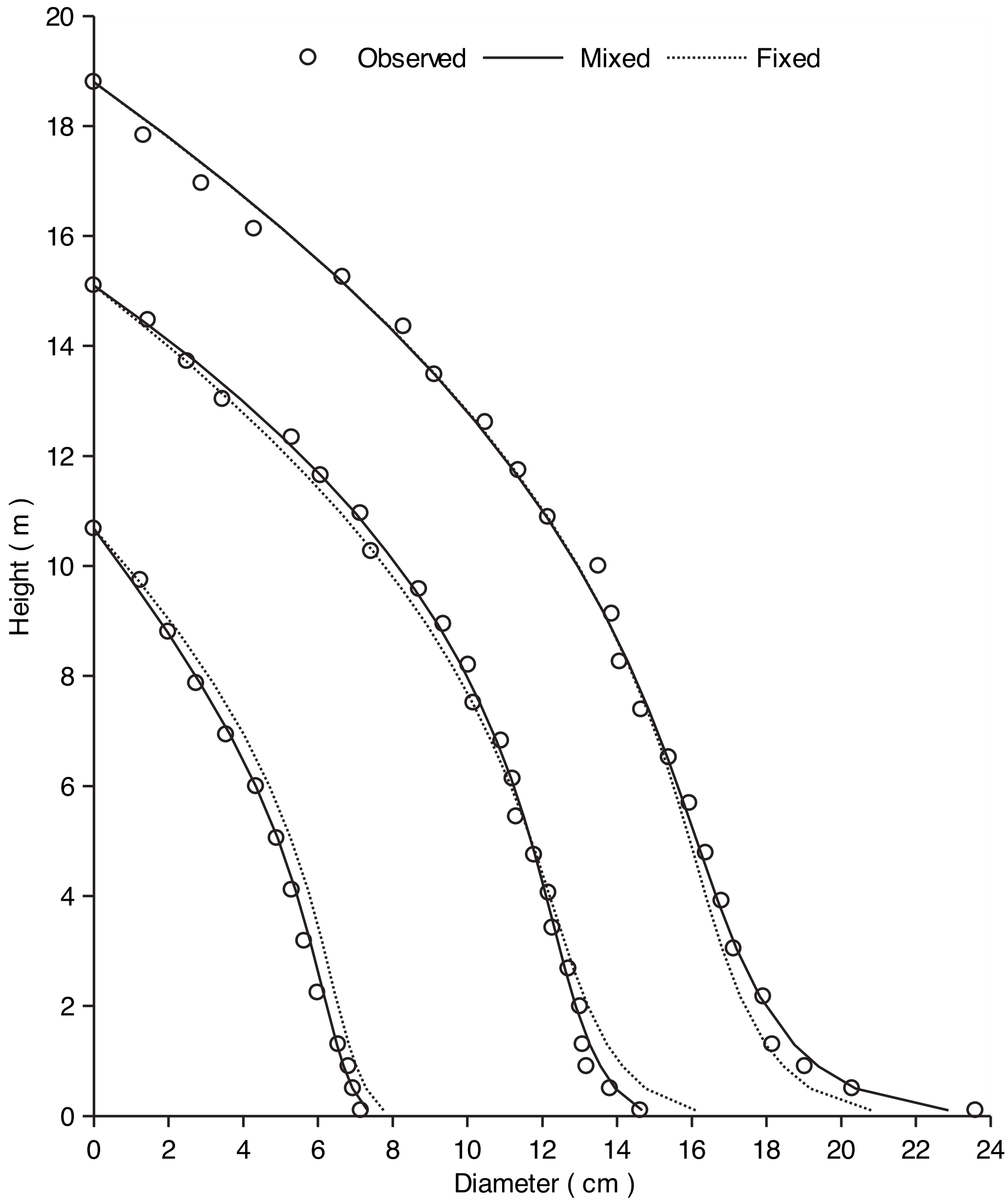


Figure 3. Tree profiles for three randomly selected jack pine trees, one from each of three classes (dominant, intermediate, and suppressed) generated using fixed-effects and mixed-effects models with β_0, β_1, β_2, and β_3 having random effects in Equation 7.

the fit statics. An attempt to associate β_4 with a random-effects parameter was not successful as the model could not be fitted with five random-effects parameters.

Other models with four random-effects parameters obtained from different combinations of β_i values ($i = 0, 1, 2, 3$, and 4) associated with random effects were also fitted. However, the model with the random effects associated with β_0–β_3 resulted in the best model in terms of fit statistics for both tree species. Estimates for parameters for this model along with fit statistics are presented in Table 3. Parameter estimates including variance components for black spruce were consistent with their counterparts for jack pine. The estimates were consistent in the sense that negative values for one species were also negative for the other species. In addition, the magnitudes of these estimates for one species were very close to those for the other species. All the variance components were significant except cov(b_0, b_3) and cov(b_1, b_2) for black spruce trees. ‘

To determine whether the error structure resulting from Equation 7 with four random-effects parameters was homogeneous and uncorrelated, inside bark values were predicted for the model data set. Bias (observed − predicted) in predicting these diameters was calculated for all diameters from each tree for both species. The bias was then plotted against the predicted diameters (Figure 2). Trends in the error structure did not suggest heterogeneous variance or correlation. This implies that the within-tree variance-covariance matrix (Equation 9) can be written as

$$\mathbf{R}_i(\boldsymbol{\beta}_i, \boldsymbol{\xi}) = \sigma^2\mathbf{I}_{ni},$$

where $\mathbf{I}_{ni}$ is the identity matrix of dimensions ($n_i \times n_i$).

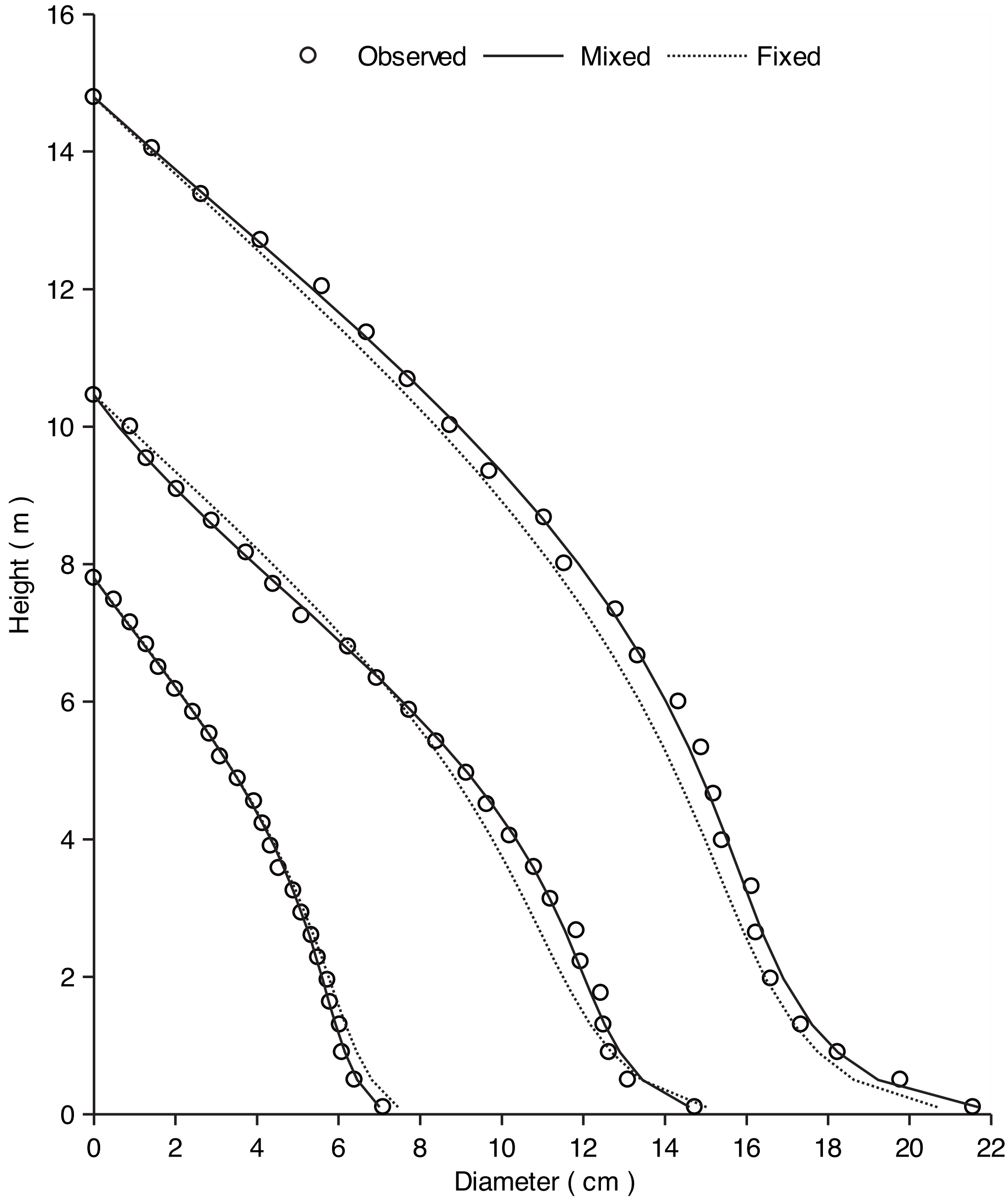


Figure 4. Tree profiles for three randomly selected black spruce trees, one from each of three classes (dominant, intermediate, and suppressed) generated using fixed-effects and mixed-effects models with β_0, β_1, β_2, and β_3 having random effects in Equation 7.

Estimates for σ^2 were 0.00039 and 0.000255 for jack pine and black spruce, respectively (Table 3).

For between-tree variation, the parameter vector $\boldsymbol{\beta}_i$ for the ith individual tree can be expressed as

$$\boldsymbol{\beta}_i = \mathbf{A}_i\boldsymbol{\beta} + \mathbf{B}_i\mathbf{b}_i = \begin{pmatrix} \beta_0 + b_{0i} \\ \beta_1 + b_{1i} \\ \beta_2 + b_{2i} \\ \beta_3 + b_{3i} \\ \beta_4 \end{pmatrix} = \begin{pmatrix} \beta_{0i} \\ \beta_{1i} \\ \beta_{2i} \\ \beta_{3i} \\ \beta_4 \end{pmatrix}$$

$$\mathbf{b}_i \sim \mathrm{N}(\mathbf{0}, \mathbf{D}),$$

where $\boldsymbol{\beta}$ is a $[\beta_0, \beta_1, \beta_2, \beta_3, \beta_4]^{\mathrm{T}}$ vector of fixed effects, $\mathbf{b}_i$ is a $[b_{0i}, b_{1i}, b_{2i}, b_{3i}]^{\mathrm{T}}$ vector of random effects, $\mathbf{A}_i = \mathbf{I}_5$ is a (5 $\times$ 5) identity matrix (design matrix) for the fixed effects, and $\mathbf{B}_i$ is a $[1000, 0100, 0010, 0001, 0000]^{\mathrm{T}}$ design matrix for the random effects. As mentioned above, the vector of random effects, $\mathbf{b}_i$, is assumed to be multivariate normally distributed with $\mathrm{E}[\mathbf{b}_i] = \mathbf{0}$ and variance covariance matrix **D.** These matrices are:

for jack pine

$$\mathbf{D} = \begin{pmatrix} 0.000558 & -0.000070 & 0.000397 & -0.000590 \\ -0.000070 & 0.000314 & -0.001380 & 0.001393 \\ 0.000397 & -0.001380 & 0.025240 & -0.032110 \\ -0.000590 & 0.001393 & -0.032110 & 0.049940 \end{pmatrix}$$

for black spruce

$$\mathbf{D} = \begin{pmatrix} 0.000900 & -0.000060 & 0.000181 & -0.000360 \\ -0.000060 & 0.000442 & -0.001420 & 0.001338 \\ 0.000181 & -0.001420 & 0.026450 & -0.031220 \\ -0.000360 & 0.001338 & -0.031220 & 0.048390 \end{pmatrix}.$$

Even though $\mathrm{E}[b_i] = \mathbf{0}$, the vector of random-effects parameters ($\mathbf{b}_i = [b_{0i}, b_{1i}, b_{2i}, b_{3i}]^{\mathrm{T}}$) for an individual tree

Table 4. Smallest values of mean biases (cm) (observed − predicted) and their SDs in predicting diameters along the bole of the trees from among the models that used one, two, and three diameters as prior information along with the corresponding values from the model without any prior information for jack pine and black spruce trees from validation data sets

	Diameter 7		Diameters 1 and 10		Diameters 1, 7, and 10		No prior	
Relative height	Bias	SD	Bias	SD	Bias	SD	Bias	SD
Jack pine								
$0.0 \le h/H \le 0.1$	0.0104	0.7394	−0.1222	0.4092	−0.1060	0.3945	0.0073	0.7493
$0.1 < h/H \le 0.2$	0.1733	0.4654	0.1277	0.5112	0.1742	0.4700	0.1776	0.5534
$0.2 < h/H \le 0.3$	0.1155	0.3923	0.0785	0.5448	0.1211	0.4112	0.1340	0.6372
$0.3 < h/H \le 0.4$	−0.0187	0.0381	−0.0943	0.5780	−0.0177	0.0418	−0.0211	0.7396
$0.4 < h/H \le 0.5$	−0.1015	0.4447	−0.1911	0.5592	−0.1254	0.4024	−0.0981	0.7801
$0.5 < h/H \le 0.6$	−0.0194	0.5669	−0.1178	0.5481	−0.0760	0.4644	−0.0198	0.8434
$0.6 < h/H \le 0.7$	0.0717	0.6764	−0.0348	0.0595	−0.0346	0.0531	0.0743	0.8615
$0.7 < h/H \le 0.8$	0.1649	0.8111	0.0645	0.5871	0.0101	0.6321	0.1657	0.9074
$0.8 < h/H \le 0.9$	0.0987	0.9101	0.0191	0.7360	−0.0805	0.7829	0.1014	0.9430
$0.9 < h/H \le 1.0$	−0.0206	0.7744	−0.0650	0.7083	−0.1701	0.7660	−0.0189	0.7824
Black spruce								
$0.0 \le h/H \le 0.1$	−0.0737	0.6869	−0.0729	0.3133	−0.0685	0.3093	0.0781	0.7059
$0.1 < h/H \le 0.2$	0.0786	0.3094	0.0520	0.3241	0.0607	0.3002	0.0769	0.3714
$0.2 < h/H \le 0.3$	0.1174	0.2896	0.1044	0.3741	0.1121	0.2960	0.1328	0.4466
$0.3 < h/H \le 0.4$	0.0082	0.1472	0.0057	0.3641	0.0060	0.1502	0.0279	0.4443
$0.4 < h/H \le 0.5$	−0.0504	0.2623	−0.0528	0.3571	−0.0484	0.2358	−0.0343	0.4707
$0.5 < h/H \le 0.6$	−0.0513	0.3619	−0.0664	0.3352	−0.0554	0.2841	−0.0350	0.5134
$0.6 < h/H \le 0.7$	−0.0089	0.4277	−0.0212	0.0561	−0.0201	0.0519	0.0062	0.5512
$0.7 < h/H \le 0.8$	0.0418	0.4541	0.0247	0.2923	0.0156	0.3349	0.0602	0.5436
$0.8 < h/H \le 0.9$	0.0606	0.4401	0.0442	0.3381	0.0246	0.3954	0.0715	0.4814
$0.9 < h/H \le 1.0$	0.0568	0.3281	0.0447	0.3036	0.0221	0.3311	0.0612	0.3371

Diameter 1 was measured at heights 0.15 m from the ground. Diameters 7 and 10 corresponded to heights at 30 and 60% of the section between 1.3 m and the tip of the tree, respectively.

could be different from zero. It can be predicted for each tree using stem diameter information if available for a subsample of disks along the bole (see Appendix). These predicted values, in combination with the fixed effects parameters displayed in Table 3, will result in a unique number for each coefficient (except for β_4) for each tree.

To visually examine the model fit, three trees from each species were randomly selected, one from each of three classes (dominant, intermediate, and suppressed). Predicted diameters from the regression (random-effects parameters b_0, b_1, b_2, and b_3 included in the model) were plotted against height along with their corresponding observed values (Figures 3 and 4). Diameters predicted without the random-effects parameters in the models were also plotted for comparison. This shows that the mixed-effects model better fits the taper data than the model using only fixed effects.

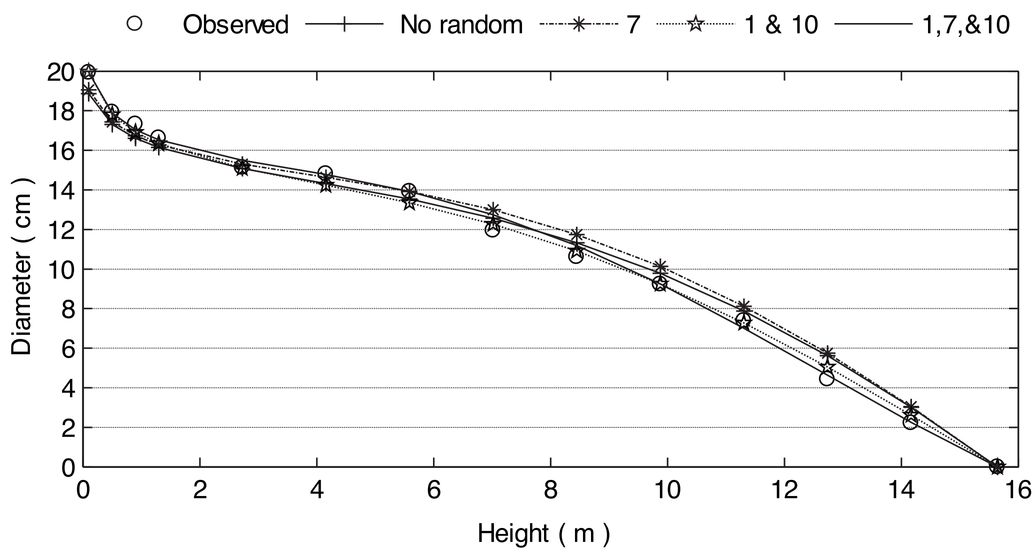
Equation 7 was further evaluated using the validation data sets. Four random-effects parameters (b_{0i}, b_{1i}, b_{2i}, and b_{3i}) were predicted for each tree from evaluation data sets using stem diameters. Disks from the stump to 65% of total height were used. Because disk 4 was sampled at the breast height (1.3 m) for all trees, diameters measured from disks 1, 2, and 3 (at heights 0.15, 0.5, and 0.9 m from the ground, respectively) and from 5, 6, 7, 8, 9, and 10 (at 10, 20, 30, 40, 50, and 60% of the section between 1.3 m and the tip of the tree, respectively) were used to predict the random-effects parameters. For the rest of the article, these diameters will be referred to as diameters 1, 2, 3, 5, 6, 7, 8, 9, and 10, respectively.

Using all nine diameter measurements from a tree as prior information for predicting random-effects parameters was not practical. Therefore, to determine the minimum number of diameters and their optimum location along the bole to accurately predict the desired parameters, three scenarios were evaluated. The first scenario used a single diameter measurement. This resulted in nine sets of random parameters predicted using one diameter measured at each height along the stem. The second scenario used two diameters, one from below and one from above breast height. This resulted in 18 sets of random parameters predicted using diameter combinations (1, 5), (1, 6), (1, 7), (1, 8), (1, 9), (1, 10), (2, 5), (2, 6), (2, 7), (2, 8), (2, 9), (2, 10), (3, 5), (3, 6), (3, 7), (3, 8), (3, 9), and (3, 10).

Finally, three diameters were used by combining one from below and two from above breast height in the third scenario. This resulted in 42 sets of random parameters predicted using diameter combinations (1, 5, 7), (1, 5, 8), (1, 5, 9), (1, 5, 10), (1, 6, 7), (1, 6, 8), (1, 6, 9), (1, 6, 10), (1, 7, 8), (1, 7, 9), (1, 7, 10), (1, 8, 9), (1, 8, 10), (1, 9, 10), (2, 5, 7), (2, 5, 8), (2, 5, 9), (2, 5, 10), (2, 6, 7), (2, 6, 8), (2, 6, 9), (2, 6, 10), (2, 7, 8), (2, 7, 9), (2, 7, 10), (2, 8, 9), (2, 8, 10), (2, 9, 10), (3, 5, 7), (3, 5, 8), (3, 5, 9), (3, 5, 10), (3, 6, 7), (3, 6, 8), (3, 6, 9), (3, 6, 10), (3, 7, 8), (3, 7, 9), (3, 7, 10), (3, 8, 9), (3, 8, 10), and (3, 9, 10). Proc IML in SAS was implemented to predict the random-effects parameters for each tree for each prior information combination scenario using Equation 7 for both species.

Random parameters predicted using each diameter combination under each scenario were combined with their fixed effects counterparts in Equation 7, and diameters along the bole for each tree for each species were estimated for the validation data sets. To evaluate the predictive ability of the models over the entire length of the stem, the relative height of each tree was divided into 10 sections. Within

(a)



(b)

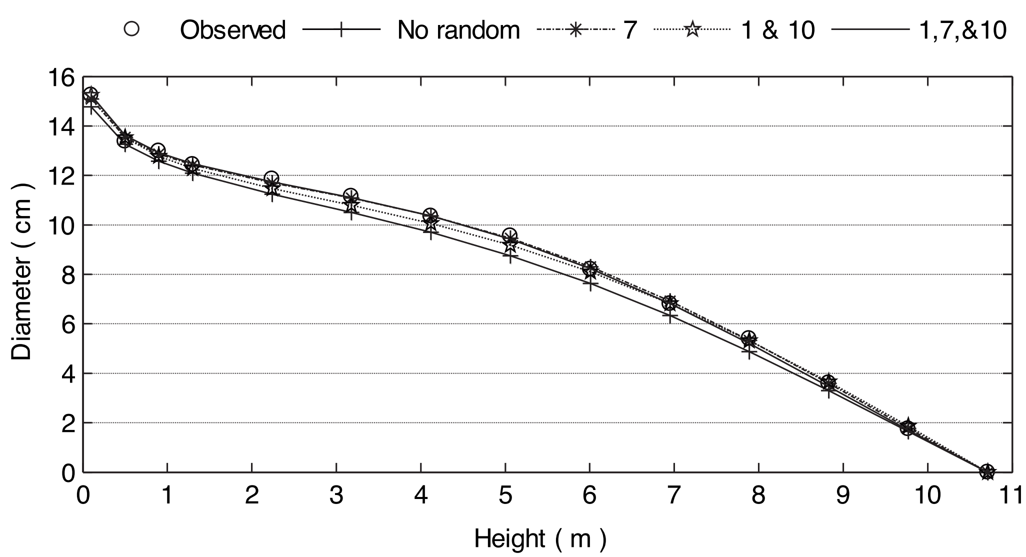


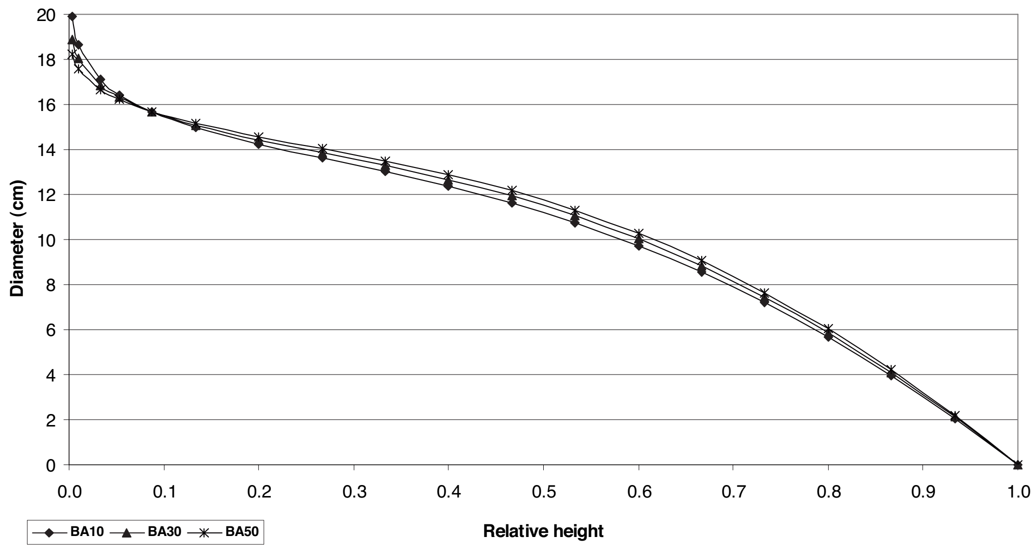
Figure 5. Calibrated responses for the trees that were closest to the average dbh and total heights for (a) jack pine (dbh = 17.5 cm and height = 15.65 m) and (b) black spruce (dbh = 13.3 cm and height = 10.71 m) using models with one, two, and three diameters (i.e., diameters corresponding to disks 7, 1 and 10, and 1, 7, and 10, respectively) as prior information, with the mean responses and observed diameters.

each section, the diameters measured were compared with the diameters predicted from these models by calculating bias (observed − predicted) in predicting diameters along the bole. Mean bias and its SD were calculated for each section using all trees for each prior information combination under each scenario for each species. The model resulting in the smallest values of mean bias and its SD across the length of a tree was considered to be the best.

Mean bias and its SD within each section were also computed for the model without any prior information (mean response model). To compare models resulting from different prior information, the model with the smallest values of bias and SD across the length of a tree was selected for each scenario for each species. These values were also calculated for the mean response model (Table 4). Interestingly, the same prior information resulted in the best model for both species under the same scenario.

The magnitudes of average bias that resulted from the best models under all prior information scenarios and from the model with no prior information are very similar along the length of the trees (Table 4). However, SDs of the bias for the model with no prior information are consistently

(a)



(b)

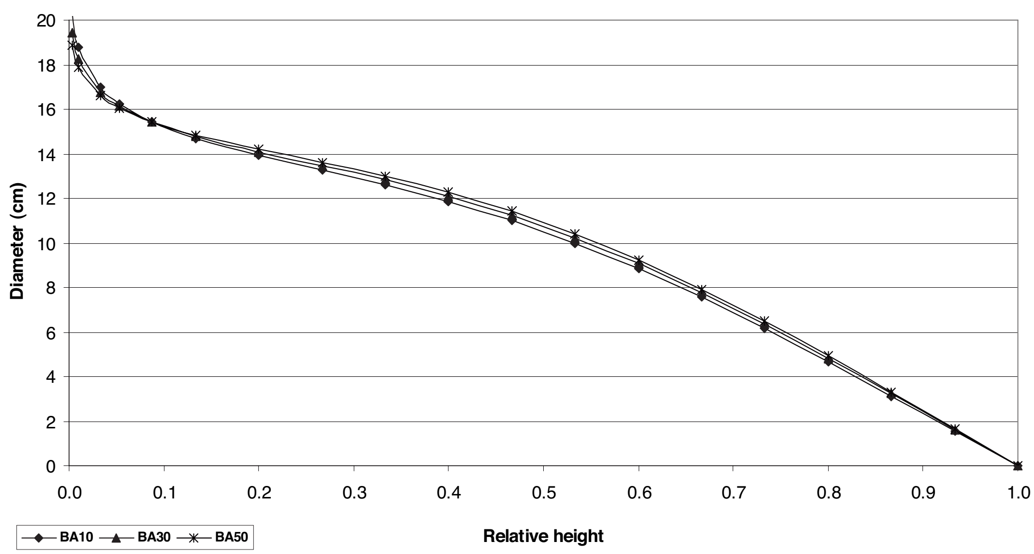


Figure 6. Tree profiles (mean responses) generated from Equation 7 using dbh = 17.0 cm and total height = 15.0 m at different stand densities (BA = 10, 30, and 50 m^2 ha^{-1}) for (a) jack pine and (b) black spruce.

larger than their corresponding values for the models with prior information for all scenarios. Thus, increasing the number of diameters used in predicting random parameters results in smaller SDs for up to 70% of total tree height for both species. Above this point, the SD of the bias is larger for the third scenario than for the second.

Because the model that uses two diameters resulted in the smaller SD than the one using three diameters above 70% of total tree height, a model that uses three diameters in predicting its random parameters and results in smaller SDs than the one that uses two diameters across the length of the trees was also investigated. In this model, one diameter below breast height was combined with two diameters at different points above breast height. The model with diameters 1, 8, and 12 as prior information resulted in consistently smaller SDs across tree height than the best model using two diameters for both species. The magnitudes of average bias from this model, however, were very similar to their corresponding values using the best model with two diameters.

The best models that used one, two, and three diameters [i.e., models with diameters 7, (1, 10), and (1, 7, 10) as priors] in predicting random-effects parameters were further evaluated by producing the associated taper profiles. These profiles (calibrated responses) were generated for the trees that were closest to the average trees in terms of dbh and total heights for both species (i.e., dbh = 17.5 cm and height = 15.65 m for jack pine and dbh = 13.3 cm and height = 10.71 m for black spruce); mean responses and observed diameters are also shown (Figure 5). Even though the model calibrated using diameters 1, 8, and 12 resulted in smaller SDs of bias in predicting diameters, taper profiles for this model were not produced here as this model would not be practically realistic.

Relying on two diameters (1 and 10) to predict random-effects parameters is sufficient to capture the taper profile for this particular jack pine tree (Figure 5). On the other hand, a calibrated response using only one diameter is even better than the one using two diameters for black spruce trees. These results indicate that two diameters (near the stump and between 60 and 65% of total tree height) are required to accurately calibrate the taper equation for jack pine but one diameter near stump height could be enough to calibrate the model for black spruce trees. The diameters used to predict random-effects parameters should be measured accurately. Otherwise, the resulting bias using mixed-effects models (calibrated response) could be worse than just using fixed-effects models (mean response).

Finally, the effect of stand density was analyzed visually by producing tree profiles (mean responses) using Equation 7 for dbh = 17.0 cm and H = 15.0 m at different stand densities (BA = 10, 30, and 50 m^2/ha) (Figure 6). These results show that the trees have larger butt diameters and more taper at lower than at higher stand density. However, the difference in bole diameter between trees at lower and higher stand densities diminishes as stand density increases. In addition, density affected the taper of jack pine more than that of black spruce.

The taper equation presented here has a unique feature. It has the flexibility that a diameter at any point along the bole can be used in place of dbh as assumed by other models, allowing the reference diameter to be measured at any convenient point on the stem. This feature is especially useful when the diameter of a tree cannot be measured at breast height because of an irregular stem (swelling, bump, depression, branch, and so on). In this case, breast height (h_D) in the equation must be replaced by the height at which the reference diameter was measured. This can be easily applied if the outside bark diameter predictions are of interest. In the case of inside bark diameters, however, the bark thickness at the point where the reference diameter was measured should be closed to the one at the breast height because the models were fitted using the ratio of inside bark diameters to outside bark dbh.

As mentioned earlier, the data used in this study were from trees that did not exhibit any visible deformities. Therefore, caution should be applied when these models are used. For example, if a stand has many trees with visible deformities, such as forks, major stem injuries, or dead or broken tops, stand volume calculated based on these equations could be overestimated.

Conclusions

A taper equation was developed for jack pine and black spruce plantations growing at varying densities. The equation was derived using a dimensional analysis approach by incorporating stand density. Different aspects of stand density (e.g., BA/ha, trees/ha, QMD, and their derivatives) were examined, but BA/ha proved most significant in describing taper. The density effect on taper was more pronounced for jack pine than for black spruce trees. A nonlinear mixed-effects approach was applied in fitting the taper equations for both species. Assuming random effects for four of the five parameters significantly improved the fit statistics (AIC and mean square error).

Inclusion of random-effects parameters for a new tree based on upper stem diameter measurements improved the predictive accuracy of the model. Three scenarios were examined: one diameter at any height along the bole; two diameters, one from below and another from above breast height; and three diameters, one from below and the other two from above breast height. The upper height at which the diameter to be measured for calibration was limited to 65% of total tree height because it is generally not practical to measure diameters above this point.

Under the first scenario, the model calibrated using a diameter measurement from between 34 and 38% of total height resulted in the best model for predicting inside bark diameters. Under the second scenario, the model calibrated using one diameter measured near the stump and the other near 65% of total height produced the least bias in predicting inside bark diameters. In the third scenario, the model calibrated using diameters measured near the stump and at approximately 35 and 65% of total height provided the best predictive accuracy. In our opinion, the improvement obtained by including a third diameter is insufficient to justify the additional costs of acquiring these measurements.

Literature Cited

AMIDON, E.L. 1984. A general taper functional form to predict bole volume for five mixed-conifer species in California. *For. Sci.* 30:166–171.

AVERY, T.E., AND H.E. BURKHART. 2002. *Forest Measurements*. 5th ed. McGraw-Hill, New York, NY. 456 p.

BURKHART, H.E., AND S.B. WALTON. 1985. Incorporating crown ratio into taper equations for loblolly pine trees. *For. Sci.* 31:478–484.

CALAMA, R., AND G. MONTERO. 2004. Interregional nonlinear height-diameter model with random coefficients for stone pine in Spain. *Can. J. For. Res.* 34:150–163.

CAO, Q.V., H.E. BURKHART, AND T.A. MAX. 1980. Evaluation of two methods for cubic volume prediction of loblolly pine to any merchantable limit. *For. Sci.* 26:71–80.

CZAPLEWSKI, R.L., AND J.P. MC.CLURE. 1988. Conditioning a segmented stem profile model for two diameter measurements. *For. Sci.* 34:512–522.

DEMAERSCHALK, J.P., AND A. KOZAK. 1977. The whole bole system: A conditioned dual equation system for precise prediction of tree profiles. *Can. J. For. Res.* 7:488–497.

FANG, Z., AND R.L. BAILEY. 2001. Nonlinear mixed-effects modeling for slash pine dominant height growth following intensive silviculture treatments. *For. Sci.* 47:287–300.

FANG, Z., BORDERS, B.E., AND BAILEY, R.L. 2000. Compatible volume-taper models for loblolly and slash pine based on system with segmented-stem form factors. *For. Sci.* 46:1–12.

FLEWELLING, J.W., AND L.M. RAYNES. 1993. Variable-shape stem-profile predictions for western hemlock. Part I. Predictions from dbh and total height. *Can. J. For. Res.* 23:520–536.

GRAY, H.R. 1956. *The form and taper of forest-tree stems.* Imperial Forest Institute. Oxford, UK. Paper 32.

GREGOIRE, T.G., O. SCHABENBERGER, AND J.P. BARRETT. 1995.

Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent plot measurements. *Can. J. For. Res.* 25:137–156.

HAYDEN, J. KERLEY, D. CARR, T. KENEDI, AND J. HALLARN. 1995. *Field Manual for Establishing and Measuring Permanent Sample Plots.* Ontario Ministry of Natural Resoures, Sault Ste. Marie, ON.

KOZAK, A. 1988. A variable-exponent taper equation. *Can. J. For. Res.* 18:1363–1368.

KOZAK, A. 1997. Effects of multicollinearity and autocorrelation on the variable-exponent taper functions. *Can. J. For. Res.* 27:619–629.

KOZAK, A. 1998. Effects of upper stem measurements on the predictive ability of a variable-exponent taper equation. *Can. J. For. Res.* 28:1078–1083.

KOZAK A, D.D. MUNRO, AND J.H.G. SMITH. 1969. Taper functions and their application in forest inventory. *For. Chron.* 45:278–283.

LARSON, P.R. 1963. *Stem form development of forest trees.* For. Sci. Monogr. 5. Society of American Foresters, Bethesda, MD.

LEITES, L.R., AND A.P. ROBINSON. 2004. Improving taper equations of loblolly pine with crown dimensions in a mixed-effects modelling framework. *For. Sci.* 50:204–212.

MAX, T.A., AND H.E. BURKHART. 1976. Segmented polynomial regression applied to taper equations. *For. Sci.* 22:283–289.

MUHAIRWE, C.K., V.M. LEMAY, AND A. KOZAK. 1994. Effects of adding tree, and site variables to Kozak's variable-exponent taper equation. *Can. J. For. Res*. 24: 252–259.

NEWNHAM, R.M. 1992. Variable-form taper functions for four Alberta tree species. *Can. J. For. Res.* 22:210–223.

NEWTON, P.F., AND M. SHARMA. 2008. Effect of disk selection protocol on taper equation performance. *Scand. J. For. Res.* 23:358–370.

ORMEROD, D.W. 1973. A simple bole model. *For. Chron.* 49: 136–138.

PINHEIRO, J.C., AND D.M. BATES. 1995. *Model building for nonlinear mixed effects models.* Dept. of Biostatistics, Univ. of Wisconsin, Madison, WI. Tech. Rep. 91. 11 p.

REED, D.D., AND J.C. BYRNE. 1985. A simple, variable form volume estimation system. *For. Chron.* 61:87–90.

RUSTAGI, K.P., AND R.S. LOVELESS, JR. 1991. Compatible variable-form volume and stem profile equations for Douglas-fir. *Can. J. For. Res*. 21: 143–151.

SAS INSTITUTE, INC. 2004. SAS 9.1.3. SAS Institute Inc., Cary, NC.

SCHABENBERGER, O., AND F.J. PIERCE. 2001. *Contemporary statistical models for the plant and soil sciences.* CRC Press LLC, Boca Raton, FL. 738 p.

SHARMA, M., AND R.G. ODERWALD. 2001. Dimensionally compatible volume and taper equations. *Can. J. For. Res.* 31:797–803.

SHARMA, M., R.G. ODERWALD, AND R.L. AMATEIS. 2002. A consistent system of equations for tree and stand volume. *For. Ecol. Manag.* 165:179–187.

SHARMA, M., AND S.Y. ZHANG. 2004. Variable-exponent taper equations for jack pine, black spruce, and balsam fir in eastern Canada. *For. Ecol. Manag.* 198:39–53.

TRINCADO, G., AND H.E. BURKHART. 2006. A generalized approach for modeling and localizing stem profile curves. *For. Sci.* 52:670–682.

VALENTINE, H.T., AND T.G. GREGOIRE. 2001. A switching model of bole taper. *Can. J. For. Res.* 31:1400–1409.

VALENTI, M.A., AND Q.V. CAO. 1986. Use of crown ratio to improve loblolly pine taper equations. *Can. J. For. Res.* 16:1141–1145.

VONESH, E.F., AND V.M. CHINCHILLI. 1997. linear and nonlinear models for the analysis of repeated measurements. Marcel Dekker, Inc., New York, NY.

ZAKRZEWSKI, W.T. 1999. A mathematically tractable stem profile model for jack pine in Ontario. *North. J. Appl. For.* 16:138–143.

Appendix

The following SAS codes can be used to calculate four random-effects parameters (b_{0i}, b_{1i}, b_{2i}, and b_{3i}) in Equation 7. Values of parameter estimates used in this example are from jack pine trees.

```
DATA one;
input Tree dbh THT Disk ht DIB Density;
/* tht = total height, Disk = disk number, ht = height from ground, dib = inside bark
   diameter at height ht, density = sqrt(baph)/dbh; */
data one; set one;
%macro predict(nd); *nd = number of disks to be used from a tree to calculate random
effect parameters;
z1 = ht/tht; z2 = z1**2;
a0 = 0.9,223; a1 = -.05,997; a2 = .5,156; a3 = -.2,265; a4 = .08,383;
*for jack pine from table 3;
pred = dbh*a0*((ht/1.3)**(a1+a2*z1+a3*z2+a4*density))*(tht-ht)/(tht-1.3);
da0 = pred/a0; da1 = pred*log(ht/1.3); da2 = pred*z1*log(ht/1.3);
da3 = pred*z2*log(ht/1.3);
resid = dib-pred; run;
proc iml;
create ranvars var {b0 b1 b2 b3};
vcov = {0.000558 -0.00007 0.00307  -0.00059,
        -0.00007 0.000314 -0.00138 0.001393,
        0.00397  -0.00138 0.025420 -0.03211,
        -0.00059 0.001393 -0.03211 0.049490};
```

```
*vcov = variance covariance matrix (D) for jack pine trees presented in the text;
mse = .00,039; G = mse*I(&nd); *mse = sigma**2 from table 3;
do i = 1 to k; *k = number of trees;
p = i;
use one var {Tree dbh tht ht dib da0 da1 da2 da3 resid} where (Tree = p);
read next &sub var {da0 da1 da2 da3} into Z;
use one var {Tree dbh tht ht diameter inside bark da0 da1 da2 da3 resid} where (Tree = p);
read next &sub var {resid} into E;
ZT = T(Z);
R = G;
M = Z*vcov*ZT + R; N = inv(M);
RANDOM = vcov*ZT*N*E;
RANDOMT = T(RANDOM);
edit ranvars;
append from randomt;
close;
end;
abort;
run;
%end;
mend predict;
%predict(nd);
proc print data = ranvars; run;
quit;
```