Taper Models for Commercial Tree Species in the Northeastern United States

James A. Westfall, and Charles T. Scott

Abstract: A new taper model was developed based on the switching taper model of Valentine and Gregoire; the most substantial changes were reformulation to incorporate estimated join points and modification of a switching function. Random-effects parameters were included that account for within-tree correlations and allow for customized calibration to each individual tree. The new model was applied to 19 species groups from data collected on 2,448 trees across the northeastern United States. A comparison of fit statistics showed considerable improvement for the new model. The largest absolute residuals occurred near the tree base, whereas the residuals standardized in relation to tree diameter were generally constant along the length of the stem. There was substantial variability in the rate of taper at the base of the tree, particularly for trees with relatively large dbh. Similar trends were found in independent validation data. The use of random effects allows uncertainty in computed tree volumes to compare favorably with another regional taper equation that uses upper-stem diameter information for local calibration. However, comparisons with results for a locally developed model showed that the local model provided more precise estimates of volume. FOR. SCI. 56(6):515–528.

Keywords: tree form, mixed-effects model, switching function, forest inventory

FORESTERS HAVE LONG RECOGNIZED the analytical flexibility afforded by tree taper models, and the development of these models has been extensively reported in forest research literature. A number of taper model forms and parameter estimation approaches have been proposed over the last several decades. A simple parabolic model was presented by Kozak et al. (1969). Coefficients for 19 species groups were estimated using the ordinary least-squares method. Max and Burkhart (1976) used segmented polynomial models to describe loblolly pine taper characteristics. Submodels were used to predict taper in the lower, middle, and upper portions of the stem. Estimates of model parameters were obtained via nonlinear least-squares techniques. The taper model form of Max and Burkhart generally outperformed other models when fit to data from 18 Appalachian hardwood species (Martin 1981) and was also used for spruce and fir in the Northeast by Solomon et al. (1989). The segmented model approach was also taken by Flewelling and Raynes (1993) to describe the tree form of Western hemlock. Maximum likelihood parameter estimates were obtained through a multivariate nonlinear methodology, in which the estimated coefficients were used to update the error covariance structure, and this new matrix was used to re-fit the coefficients. A variable-exponent taper model was developed by Kozak (1988). The exponent was predicted by a linear function and the entire model was linearized using logarithmic transformation such that least-squares estimates of the parameters were obtained. Newnham (1992) showed that a logarithmic transformation of a variable-form taper model outperformed other taper model formulations when fitted to data from four species in Alberta. The models were fit using the stepwise procedure for least-squares linear regression. The criterion for variable selection was based on R^2 values. Valentine and Gregoire (2001) introduced a variable-form taper model in which the values of exponents were manipulated with numerical switches. In this research, the height to the base of live crown was used to define the transition point between the middle and upper segments of the bole. Model parameters were estimated using nonlinear mixed-effects methods. A mixed-effects modeling approach was also used by Yang et al. (2009b) to compare the performance of four variable-exponent taper models for describing the form of lodgepole pine.

The past work on taper models has shown that new approaches for describing tree form are continually being developed and that taper modeling is a very active area of forest research. Ongoing advancements in statistical theory and computing capabilities have allowed for more complex and robust analyses, including accounting for correlations between the height/diameter data pairs within each sample tree (Valentine and Gregoire 2001, Garber and Maguire 2003, Yang et al. 2009a). Accounting for correlated observations is necessary to avoid bias in variance estimates that result from assuming independent observations when correlations are present (Swindel 1968, Sullivan and Reynolds 1976).

Many taper modeling efforts have focused on a limited number of species (Knoebel et al. 1984, Newnham 1992, Zakrzewski and MacFarlane 2006). Similarly, with perhaps the exclusion of certain commercially important species in the southeastern United States, most taper research has been

James A. Westfall, US Forest Service, Forest Inventory and Analysis, Northeastern Research Station, 11 Campus Blvd., Suite 200, Newtown Square, PA 19073—Phone: (610) 557-4043; Fax: (610) 557-4250; jameswestfall@fs.fed.us. Charles T. Scott, US Forest Service—ctscott@fs.fed.us.

Acknowledgments: We are grateful to Harry Valentine and Tim Gregoire for reviewing the draft manuscript. We also thank the reviewers, associate editor, and editor for comments that resulted in substantial improvements to the article.

Manuscript received November 14, 2008, accepted May 11, 2010

conducted with data from relatively small areas (Garber and Maguire 2003, Ter-Mikaelian et al. 2004, Jiang et al. 2007). An exception to these paradigms is the work of Clark et al. (1991), in which models applicable to 58 species/species groups were developed from data collected in 11 states in the southeastern United States.

The purpose of this article is to present a new taper model based on the formulation of Valentine and Gregoire (2001). The new model includes estimation of submodel join points, a modified switching function, and alternative placement of a random coefficient. A mixed-effects modeling strategy was used to account for multiple measurements within each tree. Model development was carried out using data from a single species and was subsequently applied to numerous commercial tree species in the northeastern United States.

Model Specification

A number of published articles have compared different taper models (Martin 1981, Newnham 1992, Williams and Wiant 1994). It is apparent that no single model performs well for all tree species. We chose to base our work on the switching model formulated by Valentine and Gregoire (2001) because of its flexibility and biologically sensible derivation, i.e., the model was specified under the theory that the lower, middle, and upper portions of the bole have neiloidal, parabolic, and conic forms, respectively.

The original model form (in terms of diameter squared) is given by

$$d_{jk}^2 = \mathrm{dbh}_j^2\left(\frac{H_j - h_{jk}}{H_j - 1.37}\right)^{(\alpha_1+S_1)}\left(\frac{H_j - h_{jk}}{H_j - C_j}\right)^{\alpha_2 S_2} + \varepsilon_{jk} \quad (1)$$

$$S_1 = \frac{(\theta_1 + \phi_{1j})H}{1 + (h_{jk}/\theta_{2j}H_j)^{(\lambda+\phi_2)}} \quad S_2 = \frac{(h_{jk}/C_j)^{\beta_2}}{1 + (h_{jk}/C_j)^{\beta_2}}$$

$$\varepsilon_{jk} \sim N\left(0,\ \sigma_e^2\left(\mathrm{dbh}_j^{\nu_1}\exp\left(\nu_2\left[\frac{h_{jk}}{H_j}\right]^{\nu_3}\right)\right)\right) \quad (2)$$

where dbh_j is dbh (1.37 m) for tree j (cm), h_{jk} is height k along tree j (m), d_{jk} is outside bark diameter at height h_{jk} (cm), H_j is total tree height for tree j (m), C_j is height to live crown for tree j (m). ε_{jk} is residual error at height k for tree j, γ_1, γ_2, α_1, α_2, β_1, β_2, θ_1, θ_2, λ, ψ, ν_1, ν_2, ν_3, σ_e^2 are estimated fixed-effects parameters, and ϕ_{1j}, ϕ_{2j} are predicted random-effects parameters for tree j; $\phi \sim N(0, \sigma_i^2)$, $i = 1, 2$.

The lower segment is defined as the portion of the bole where $h_{jk} \leq 1.37$ m in height, the middle segment occurs between 1.37 m $\leq h_{jk} \leq C_j$, and the upper segment occurs from $C_j \leq h_{jk} \leq H_j$. Thus, the join points between the model segments occur at 1.37 m (the height at which dbh_j is measured) and C_j. Switching functions (S_1, S_2) modify their values with changing h_{jk} and were used in the model to smooth the transitions between the segments. S_1 exhibits switch-off behavior with the rate being controlled by the λ parameter. Conversely, S_2 has switch-on behavior with rate being a function of the β_2 parameter. As the rate parameter gets larger, the switch occurs over a shorter segment of the bole.

For this study, several modifications were made to the original model specification:

1. The segment join points were estimated from the data.
2. The model was reformulated to use relative height instead of actual height to facilitate estimated join points.
3. The upper bole switching function was modified.
4. The error variance function was altered slightly.
5. The placement of a random-effect parameter was changed.
6. The dependent variable was changed to diameter squared (versus cross-sectional area).

The new model form is given by

$$d_{jk}^2 = \mathrm{dbh}_j^2\left(\frac{1.37/H_j}{\gamma_1}\right)^{\psi}\left(\frac{1 - z_{jk}}{1 - \gamma_1}\right)^{(\alpha_1+\varphi_{1j})+S_1}\left(\frac{1 - z_{jk}}{1 - \gamma_2}\right)^{\alpha_2 S_2} + \varepsilon_{jk} \quad (3)$$

$$S_1 = \frac{(\theta_1 + \phi_{2j})}{1 + (z_{jk}/\theta_2)^{\lambda}} \quad S_2 = \frac{(z_{jk}/\beta_1)^{\beta_2 x_k}}{1 + (z_{jk}/\beta_1)^{\beta_2 x_k}}$$

$$\varepsilon_{jk} \sim N(0,\ \sigma_e^2(\mathrm{dbh}_j^{\nu_1}\exp(\nu_2 z_{jk}))) \quad (4)$$

where $z_{jk} = h_{jk}/H_k$, $x_j = \mathrm{dbh}_j/H_j$, and others are as defined previously.

Overall, the theoretical development of Valentine and Gregoire (2001) still provides the basis for the model formulation. Thus, it is expected that all of the estimated parameters except ν_2 would be greater than zero, with join points γ_1 and γ_2 being within the interval 0–1 and $0 < \gamma_1 < \gamma_2 < 1$. It was surmised that the incorporation of estimated join points would allow the model to better fit the data. The original model was specified such that the lower join point was at 1.37 m and the right-hand side simplified (essentially) to dbh^2 when the height along the bole was 1.37 m. The inclusion of estimated join points nullified this property and without further modification, the right-hand side would reduce to dbh^2 at the estimated join point γ_1. The term $(1.37/H)/\gamma_1)^{\psi}$ was added to account for differences between the estimated join point and the relative height at which dbh is measured. When γ_1 is greater than $1.37/H$, the value of d^2 at γ_1 is less than dbh^2, and this term multiplied against dbh^2 provides a downward adjustment that allows prediction of d^2 less than dbh^2. Similarly, when γ_1 is less than $1.37/H$, the value of d^2 at γ_1 is greater than dbh^2 and the multiplication against dbh^2 results in predicted values of d^2 that are larger than dbh^2.

Other significant changes to the original model were changing the placement of a random-effect parameter in the switching function (S_1) for the lower join point and respecification of the switching function (S_2) associated with the upper join point. During the course of analysis, alternative placements of random-effects parameters were tested. The best results were obtained by moving φ_{1j} out of the S_1 switching function and associating the random effect with the α_1 parameter. It was theorized that allowing α_1 to vary among trees provided more flexibility than the original model. In addition, the S_2 switching function only depended

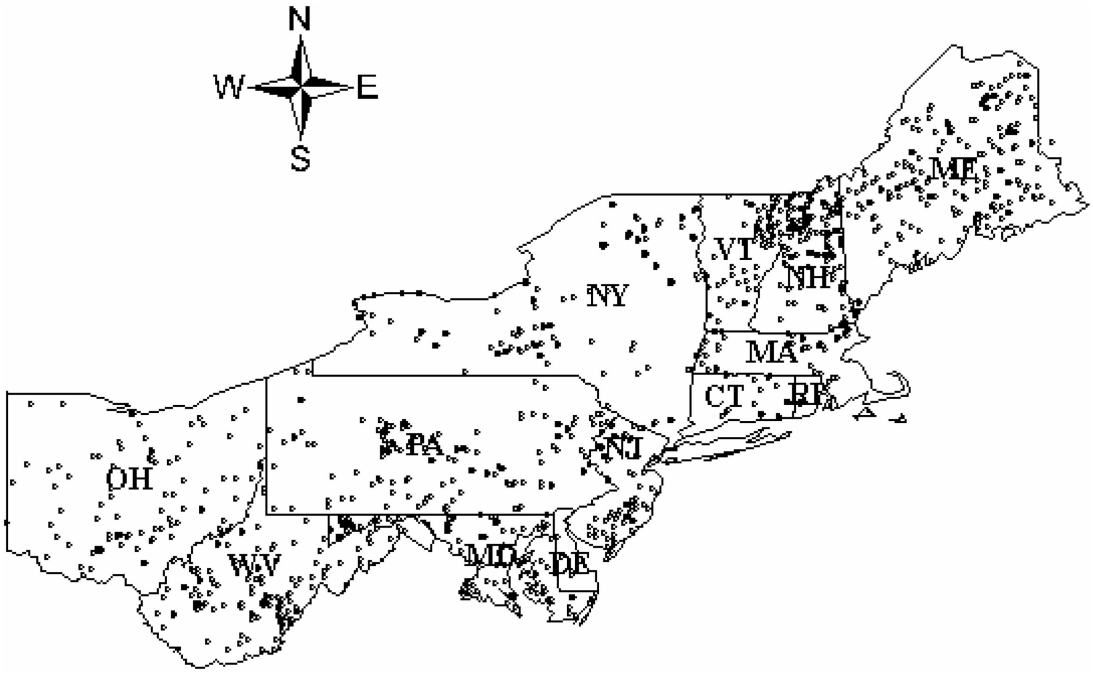

Figure 1. Spatial distribution of data collection locations across 13 states in U.S.

on the fixed parameter estimates under the new formulation. To tailor the performance to individual trees, the ratio of dbh/height was added. The function was formulated such that the length of the transition would be sensitive to tree dimensions, i.e., the transition occurs more quickly when a tree is relatively short in relation to its diameter.

The error variance Equation 4 is essentially the same as that specified by Valentine and Gregoire (2001) shown in Equation 2. However, our data did not support the additional complexity offered by the ν_3 parameter, and thus the function was simplified to two parameters.

Data

The data used to fit the model were collected by the Forest Inventory and Analysis (FIA) program of the U.S. Forest Service. The geographic range encompassed 13 states in the northeastern United States, including West Virginia, Maryland, Delaware, New Jersey, Pennsylvania, Ohio, New York, Massachusetts, Rhode Island, Connecticut, Vermont, New Hampshire, and Maine. Data were collected assuming species groupings described by Scott (1981). The sample allocation was based on species frequency and tree size information obtained from FIA inventory plots. Only one tree per species group was sampled at any given location, thus spreading the sample geographically and eliminating the potential need to account for correlations between trees at sample locations. Sample locations also were required to meet the standard FIA definition of forestland (at least 0.4 ha of area and a minimum of 36.6 m wide). Selection of individual sample trees was restricted only by characteristics that may produce inconsistent measurements, e.g., broken top or excessive lean. Figure 1 depicts the spatial distribution of sample locations.

Measurements were taken on a total of 2,448 trees over a range of tree sizes (Table 1). Height/diameter information was obtained using a Barr and Stroud dendrometer. Fixed measurement points occurred at heights of 0.3, 0.6, 0.9, 1.4, and 1.8 m. Additional measurements were taken at approximately 2.5-cm taper intervals up to total tree height. Tree heights were recorded to nearest 0.03 m, and tree diameters were recorded to the nearest 0.25 cm. There were 37,852 height/diameter data pairs obtained.

Results

To ascertain whether any improvements were gained with the reformulated model 3, comparisons with the original model 1 were performed using the data for red/silver maple (Table 2, group 19). An decrease of 16% in the Akaike information criterion (AIC) (Akaike 1974) fit statistic was gained with the new specification (smaller AIC is better). To test the efficacy of moving the random-effect parameter φ_{1j}, the new model formulation was fitted with φ_{1j} in the original location. This resulted in an increase of 8% in the AIC statistic, suggesting that an decrease of approximately 8% in the AIC value can be attributed to the change in the placement of φ_{1j}.

The model was fitted to 19 species groups (Table 2) using the SAS NLMIXED procedure. Table 3 contains the estimated fixed-effects coefficients in the taper model 3 for each species group. The estimated lower join points (γ_1) for the change from neiloidal to parabolic form generally occur near 0.1 relative height. The value of γ_1 also determines the direction and magnitude of the difference between the lower transition point and the relative height at which dbh is measured ($z_{\text{dbh}} = 1.37/H$). The relatively small differences between γ_1 and z_{dbh} result in estimated coefficients ψ being within the range of 0.13 to 0.30 across all species groups. The estimated upper join point (γ_2) for the change from parabolic to conic form exhibits high variability across the species groups. These relative height values range from 0.16

Table 1. Summary of taper data by species group for dbh, total height, number of trees, and number of data points

Group	dbh			Total height			No. trees	No. data points
	Minimum	Mean	Maximum	Minimum	Mean	Maximum		
	(cm)			(m)				
1	9.7	27.7	49.5	5.6	17.1	27.1	38	585
2	8.1	31.8	93.2	5.5	18.6	34.5	102	1,614
3	7.1	24.6	57.9	5.0	15.0	28.8	148	2,018
4	8.6	24.1	54.1	6.0	15.3	28.3	150	2,015
5	8.6	34.5	85.1	4.8	17.0	33.2	133	2,095
6	7.9	25.1	59.7	3.5	17.0	32.0	146	2,078
7	8.4	27.2	78.2	4.3	13.7	24.7	123	1,783
8	7.9	33.0	79.8	7.7	19.5	31.9	138	2,195
9	7.1	30.2	73.4	7.9	22.6	38.6	132	2,183
10	8.4	25.9	76.5	7.8	19.5	38.2	134	1,981
11	7.6	28.4	81.8	5.9	19.6	32.5	114	1,808
12	7.9	23.6	72.1	6.5	16.1	29.0	147	1,982
13	7.1	32.5	89.7	7.9	18.9	33.8	141	2,285
14	7.6	30.2	74.7	5.3	20.8	36.9	115	1,872
15	7.9	32.5	112.5	9.7	20.9	36.9	154	2,545
16	8.4	34.3	85.3	7.1	20.4	34.2	131	2,231
17	7.6	24.9	83.8	7.6	20.7	37.1	138	2,163
18	7.1	32.0	80.3	6.6	21.1	36.4	143	2,438
19	6.9	32.5	78.5	6.8	20.0	32.3	121	1,981
All	6.9	29.2	112.5	3.5	18.6	38.6	2,448	37,852

Table 2. Species groups and within-group composition

Group	Species composition
1	Red pine (*Pinus resinosa* [Ait.])
2	Eastern white pine (*Pinus strobus* [L.])
3	White spruce (*Picea glauca* [(Moench) Voss]), black spruce (*Picea mariana* [(Mill.) B.S.P.]), red spruce (*Picea rubens* [Sang.])
4	Balsam fir (*Abies balsamea* [(L.) Mill.])
5	Eastern hemlock (*Tsuga canadensis* [(L.) Carr.])
6	Larch (introduced) (*Larix* spp. [Mill.]), tamarack (native) (*Larix laricina* [(Du Roi) K. Koch]), Norway spruce (*Picea abies* [(L.) Karst.]), jack pine (*Pinus banksiana* [Lamb.]), shortleaf pine (*Pinus echinata* [Mill.]), Table Mountain pine (*Pinus pungens* [Lamb.]), pitch pine (*Pinus rigida* [Mull.]). Scotch pine (*Pinus sylvestris* [L.]), loblolly pine (*Pinus taeda* [L.]), Virginia pine (*Pinus virginiana* [Mill.])
7	Atlantic white cedar (*Pinus rigida* [Mill.]), Eastern redcedar (*Pinus sylvestris* [L.]). northern white cedar (*Pinus taeda* [L.])
8	Sugar maple (*Acer saccharum* [Marsh.])
9	Yellow poplar (*Liriodendron tulipifera* [L.])
10	White ash (*Fraxinus americana* [L.]), black ash (*Fraxinus nigra* [Marsh.]), green ash (*Fraxinus pennsylvanica* [Marsh.]), balsam poplar (*Populus balsamifera* [L.]), eastern cottonwood (*Populus deltoides* [Bartr. ex Marsh.]), bigtooth aspen (*Populus grandidentata* [Michx.]), quaking aspen (*Populus tremuloides* [Michx.])
11	Black cherry (*Prunus serotina* [Ehrh.])
12	Yellow birch (*Betula alleghaniensis* [Britton]), sweet birch (*Betula lenta* [L.]), river birch (*Betula nigra* [L.]), paper birch (*Betula papyrifera* [Marsh.]), gray birch (*Betula populifolia* [Marsh.])
13	American beech (*Fagus grandifolia* [Ehrh.])
14	Basswood (*Tilia* spp. [L.]). American basswood (*Tilia americana* [L.]), white basswood (*Tilia heterophylla* [Vent.])
15	Sweetgum (*Liquidambar styraciflua* [L.]), blackgum (*Nyssa svlvatica* [Marsh.]), scarlet oak (*Quercus coccinea* [Muenchh.]). southern red oak (*Quercus falcata* var. falcata [Michx.]), shingle oak (*Quercus imbricaria* [Michx.]). pin oak (*Quercus palustris* [Muenchh.]), willow oak (*Quercus phellos* [L.]), northern red oak (*Quercus rubra* [L.]), black oak (*Quercus velutina* [Lam.])
16	Chestnut oak (*Quercus prinus* [L.])
17	Hickory (*Carya* spp. [Nutt.]), bitternut hickory (*Carya cordifonnis* [(Wangenh.) K. Koch]), pignut hickory (*Carya glabra* [(Mill.) Sweet]), shellbark hickory (*Carya laciniosa* [(Michx. f.) Loud.]), shagbark hickory (Carya ovata [(Mill.) K. Koch]), mockernut hickory (Carya tomentosa [(Poir.) Nutt.])
18	Buckeye (*Aesculus* spp. [L.]), Ohio buckeye (*Aesculus glabra* [Willd.]), yellow buckeye (*Aesculus octandra* [Marsh.]), hackberry (*Celtis occidentalis* [L.]), honeylocust (*Gleditsia triacanthos* [L.]), American holly (*Ilex opaca* [Ait.]), butternut (*Juglans cinerea* [L.]), black walnut (*Juglans nigra* [L.]), magnolia (*Magnolia* spp. [L.]), cucumbertree (Magnolia acuminata [L.]), sycamore (*Platanus occidentalis* [L.]), white oak (*Quercus alba* [L.]), swamp white oak (*Quercus bicolor* [Willd.]), bur oak (*Quercus macrocarpa* [Michx.]), chinkapin oak (*Quercus muehlenbergii* [Engelm.]), black locust (*Robinia pseudoacacia* [L.]), black willow (*Salix nigra* [Marsh.]), elm (*Ulmus* spp. [L.]), American elm (*Limus americana* [L.]), slippery elm (*Ulmus rubra* [Muhl.])
19	Red maple (*Acer rubrum* [L.]), silver maple (*Acer saccharinum* [L.])

Table 3. Estimated fixed-effects coefficients (SE) for taper model 3

Group	θ_1	θ_2	α_1	α_2	γ_1	γ_2	ψ	λ	β_1	β_2
1	7.6044	0.0148	1.2379	0.3304	0.0759	0.6611	0.3008	1.1569	0.5462	3.0627
	(2.10)	(0.01)	(0.17)	(0.15)	(0.00)	(0.16)	(0.01)	(0.22)	(0.09)	(1.07)
2	7.1438	0.0123	0.8978	0.7872	0.0989	0.4985	0.2049	0.9247	0.5715	2.0482
	(1.47)	(0.01)	(0.12)	(0.10)	(0.01)	(0.04)	(0.01)	(0.14)	(0.05)	(0.30)
3	6.8745	0.0110	1.1241	0.4107	0.1376	0.4842	0.2038	1.2598	0.4986	2.7865
	(1.94)	(0.00)	(0.06)	(0.05)	(0.01)	(0.06)	(0.01)	(0.15)	(0.04)	(0.68)
4	5.3693	0.0171	1.4212	0.3003	0.0890	0.6485	0.1916	1.8873	0.4764	2.6383
	(1.10)	(0.00)	(0.06)	(0.05)	(0.01)	(0.07)	(0.01)	(0.22)	(0.04)	(0.54)
5	7.2442	0.0152	1.4008	0.8306	0.0856	0.4724	0.2011	1.5716	0.6994	1.9524
	(2.89)	(0.01)	(0.04)	(0.13)	(0.01)	(0.05)	(0.01)	(0.23)	(0.06)	(0.43)
6	5.2913	0.0411	1.1291	0.6831	0.0745	0.5798	0.1896	1.5776	0.6616	6.0645
	(1.03)	(0.01)	(0.05)	(0.05)	(0.01)	(0.03)	(0.01)	(0.21)	(0.02)	(1.25)
7	5.4000	0.0256	1.9295	0.8142	0.0943	0.9642	0.2761	1.8605	1.3432	1.3438
	(0.64)	(0.00)	(0.04)	(0.28)	(0.01)	(0.02)	(0.01)	(0.18)	(0.18)	(0.10)
8	6.5790	0.0111	1.0682	1.1833	0.1031	0.2624	0.1516	1.1482	0.6637	3.0996
	(1.40)	(0.00)	(0.07)	(0.13)	(0.01)	(0.08)	(0.01)	(0.15)	(0.03)	(0.45)
9	6.8715	0.0058	1.0542	0.9805	0.1141	0.3607	0.1412	0.9822	0.6614	3.6589
	(1.14)	(0.00)	(0.06)	(0.08)	(0.01)	(0.04)	(0.01)	(0.10)	(0.02)	(0.48)
10	3.3085	0.0276	1.2634	0.9088	0.1098	0.5198	0.1840	1.7842	0.6719	5.1178
	(0.37)	(0.00)	(0.04)	(0.07)	(0.01)	(0.04)	(0.01)	(0.21)	(0.02)	(0.73)
11	3.2042	0.0479	1.2507	0.8075	0.0800	0.4170	0.2227	2.7226	0.7065	4.6476
	(0.30)	(0.01)	(0.05)	(0.12)	(0.01)	(0.12)	(0.01)	(0.42)	(0.04)	(1.07)
12	7.5437	0.0103	0.9961	1.1042	0.1313	0.3539	0.2091	0.9478	0.5995	3.4205
	(0.82)	(0.00)	(0.09)	(0.10)	(0.01)	(0.04)	(0.01)	(0.08)	(0.02)	(0.61)
13	8.9843	0.0107	0.7621	1.3734	0.0956	0.1650	0.1924	1.2237	0.4626	1.0954
	(1.66)	(0.00)	(0.27)	(0.39)	(0.01)	(0.04)	(0.01)	(0.15)	(0.03)	(0.33)
14	5.3708	0.0072	0.8737	0.9876	0.1123	0.3053	0.1303	0.8358	0.6718	3.4707
	(0.59)	(0.00)	(0.09)	(0.09)	(0.01)	(0.07)	(0.01)	(0.11)	(0.02)	(0.41)
15	12.8336	0.0125	0.9038	1.0950	0.0935	0.3971	0.2038	1.0457	0.5508	3.4681
	(1.87)	(0.00)	(0.10)	(0.08)	(0.01)	(0.03)	(0.01)	(0.08)	(0.02)	(0.46)
16	9.5073	0.0066	0.7679	1.2678	0.1379	0.3926	0.2039	0.8442	0.5208	2.6000
	(1.45)	(0.00)	(0.11)	(0.09)	(0.01)	(0.03)	(0.01)	(0.08)	(0.02)	(0.35)
17	9.4326	0.0085	0.9135	0.8853	0.1284	0.3969	0.1887	0.8447	0.5547	4.0183
	(0.94)	(0.00)	(0.11)	(0.08)	(0.01)	(0.05)	(0.01)	(0.07)	(0.03)	(0.64)
18	9.0505	0.0241	1.2980	0.7684	0.0684	0.4555	0.1769	1.6684	0.5408	4.1821
	(1.10)	(0.00)	(0.06)	(0.05)	(0.01)	(0.03)	(0.01)	(0.20)	(0.02)	(0.77)
19	7.5707	0.0105	1.5273	0.7684	0.0931	0.4223	0.1441	1.3910	0.6453	4.0737
	(1.72)	(0.00)	(0.05)	(0.08)	(0.01)	(0.06)	(0.02)	(0.17)	(0.04)	(1.03)

for American beech (species group 13) to 0.96 for cedars (species group 7). Unlike the estimated values for γ_1, a notable difference between softwood and hardwood species groups is exhibited in estimates for γ_2. For softwood species (species groups 1–7) the average relative height for the upper join point was 0.62, whereas for hardwood species (species groups 8–19) the average was 0.37. We surmise that this result is due to the general stem form differences between softwood and hardwood species (excurrent versus deliquescent, respectively).

The value of the S_1 switching function decreases with increasing relative height, which facilitates the transfer between the lower and middle segments of the bole. The behavior of S_1 was consistent among all species groups, for which the estimated value of λ produced a response curve similar to an inverse exponential function. The initial magnitude defines the amount of swell at the base of the tree and is primarily driven by the estimate of θ_1. For most species groups, the shape value ($\alpha_1 + S_1$) decreases to a value between 1.0 and 2.0 with increasing relative height, indicating the transition from a neiloidal lower section to a parabolic middle section.

The switching function S_2 facilitates transition between the middle and upper segments of the bole with S_2 increasing toward 1.0 with increasing h. As a multiplier of α_2, the potential range of the shape parameter $\alpha_2 \times S_2$ is $(0.0, \alpha_2)$. The upper limit defined by α_2 exhibits a wide range of values across species groups, from 0.30 for balsam fir (species group 4) to 1.37 for American beech (species group 13). Similar to that for γ_2, there was a notable difference in estimated α_2 between hardwood and softwood species groups. The mean value of α_2 for softwood groups was 0.59, whereas the mean for hardwood groups was 1.01.

Figure 2 depicts the behavior of the S_2 switching function in species group 19 (red/silver maple) for trees having x (= dbh/H) = 1.5 and 3.0 cm/m. For trees having relatively little taper, the transition between segments begins earlier and progresses over a longer range of relative height. Conversely, trees having relatively high rates of taper exhibit transitions that begin later and occur over a shorter relative height range. The more rapid transitions associated with increasing x result in the fundamental shape for the upper stem segment (e.g., conic) being realized at lower relative heights.

Residuals from 3 were examined for each species group and no systematic problems were revealed. Figure 3 displays plots of raw residuals against predicted values, relative height (z), dbh, and total tree height (H) for sugar maple

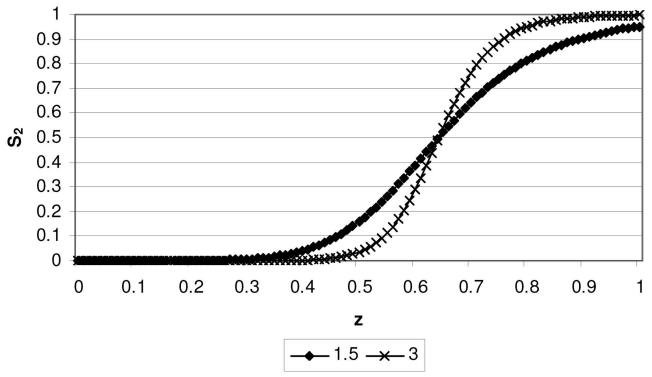

Figure 2. Behavior of S_2 for x (dbh/H) = 1.5 and 3.0 cm/m for species group 19 (red/silver maple).

(species group 8). Residual error for this species group is representative of the average amount of residual error across all groups. The residuals are approximately evenly distributed, regardless of the comparative metric. It is apparent that the variability of the residuals is higher for larger trees. In addition, most of this variability is contained in the lower half of the stem ($z \leq 0.5$), with a slightly prominent concentration near the base of the tree ($z \leq 0.05$).

The random-effects parameters φ_{1j} and φ_{2j} were associated with fixed-effects coefficients α_1 and θ_1, respectively. The variances of the random-effects components (σ_1^2, σ_2^2) are shown in Table 4. The values of σ_1^2 were very similar across all species groups, indicating similar intertree variation about α_1. The value of α_1 largely controls the shape in the middle segment of the model. The values of σ_2^2 exhibit a fairly wide range across species groups. As the random-effect φ_{2j} provided modification to the fixed-effect θ_1 for each tree, the value of σ_2^2 reflects the relative variability of butt swell within a species group. Red/silver maple (species group 19) and balsam fir (species group 4) exhibited the highest variability, whereas the least variability was found in pines (species group 6).

In some applications, the use of random-effects parameters eliminates the heteroscedasticity in the error variance (Jones 1990). However, examination of residuals from Equation 3 indicated that patterned variance still existed. This was probably due to specific tree characteristics that random effects were unable to resolve. For instance, larger trees have a more wavering profile in the lower bole than smaller trees. A model that assumes decreasing diameter with increasing height cannot accommodate this variability well, regardless of inclusion of random-effects parameters, thus resulting in larger absolute residuals than for smaller trees that tend to follow the assumed shapes. Similarly, trees generally have more form variation in the lower stem, e.g., due to out-of-roundness, and thus length along the bole affects the amount of error variance.

Estimated coefficients for the function that describe error variance Equation 4 along with estimated variances of the random effects are shown in Table 4. The negative-valued ν_2 indicates that the most variability is encountered near the base of the tree, and variance decreases with increasing h. It is also shown that variability increases with increasing dbh, as evidenced by positive-valued ν_1. If an unbiased model is

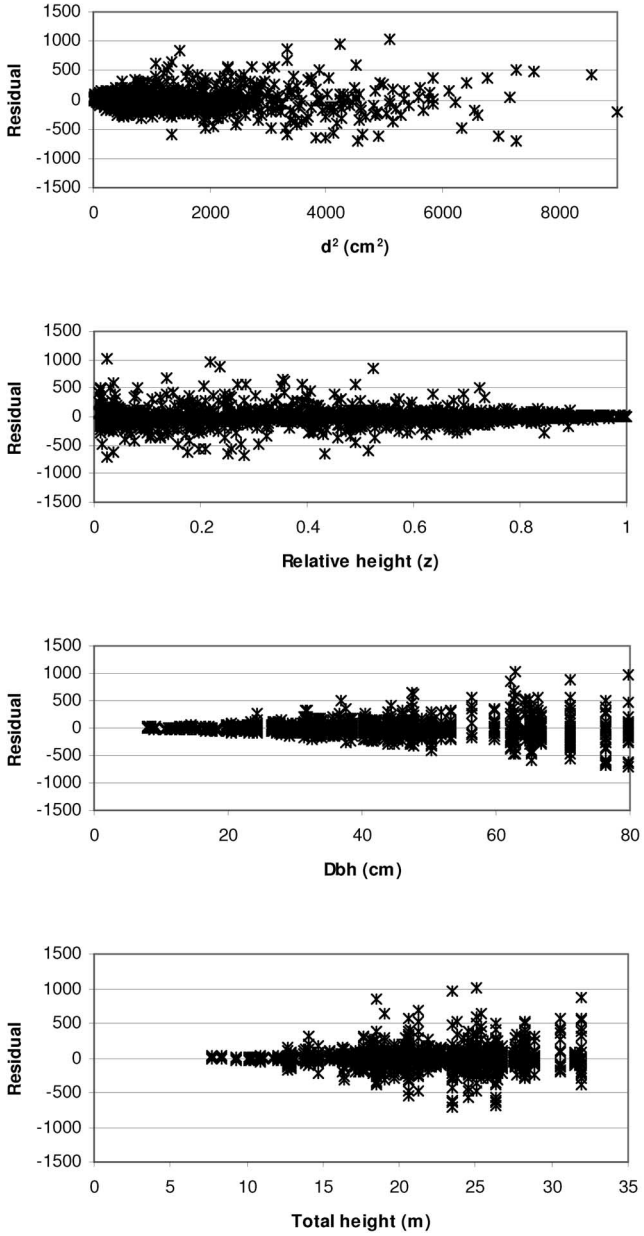

Figure 3. Plots of residual values from model [3] vs. predicted values (d^2), relative height (z), dbh, and total height for sugar maple (species group 8).

assumed, the predicted variance of ε is akin to a mean squared-error statistic and the root mean squared error (RMSE) is obtained by taking the square root of the variance. The magnitude of the RMSE in relation to tree dbh and relative height is depicted in Figure 4.

Table 5 summarizes the error from Equation 3 by species group across relative height classes. The magnitude of error is characterized by two metrics: mean absolute deviation (MAD) (cm^2) and root mean squared relative residual (RMSRR) (%). The trends in MAD generally reflect what was shown in the residual plots (Figure 2), in which the largest values are at the lower relative height classes and MAD decreases with increasing height. In the 0.00–0.04 relative height class (in which the largest MADs were generally encountered), the deviations ranged from 27.7 to 175.3 cm^2 for balsam fir (species group 4) and Eastern

Table 4. Estimated fixed-effects coefficients (SE) for variance function 4 and variances of random-effects from taper model 3

Group	σ_e^2	ν_1	ν_2	σ_1^2	σ_2^2
1	0.0734 (0.03)	3.2805 (0.14)	−3.0847 (0.20)	0.0982	1.4934
2	0.0146 (0.00)	3.9015 (0.07)	−3.2048 (0.14)	0.1094	2.3184
3	0.0491 (0.01)	3.6232 (0.07)	−3.5589 (0.11)	0.1104	5.4370
4	0.0429 (0.01)	3.4143 (0.08)	−2.7934 (0.13)	0.0724	18.1725
5	0.0411 (0.01)	3.7729 (0.07)	−4.5450 (0.12)	0.1317	3.7505
6	0.2819 (0.09)	3.1675 (0.10)	−4.0688 (0.09)	0.0972	0.4393
7	0.1133 (0.03)	3.3872 (0.07)	−4.1175 (0.13)	0.1031	13.2516
8	0.0330 (0.01)	3.7631 (0.06)	−3.5830 (0.13)	0.2275	10.4670
9	0.0114 (0.00)	4.0362 (0.05)	−4.3288 (0.12)	0.0802	1.8686
10	0.0052 (0.00)	4.1963 (0.06)	−2.8835 (0.13)	0.1129	4.4449
11	0.0275 (0.01)	3.7305 (0.08)	−2.9463 (0.14)	0.1037	3.0032
12	0.0282 (0.01)	3.8727 (0.07)	−3.7804 (0.15)	0.1625	9.6901
13	0.0357 (0.01)	3.8073 (0.06)	−3.9789 (0.14)	0.1572	5.2652
14	0.0261 (0.01)	3.9352 (0.06)	−3.9959 (0.13)	0.0897	5.4129
15	0.0264 (0.01)	3.8581 (0.06)	−4.1602 (0.13)	0.0940	13.9132
16	0.0418 (0.01)	3.7182 (0.06)	−3.5527 (0.14)	0.1736	8.7560
17	0.0219 (0.00)	4.0124 (0.07)	−4.1099 (0.15)	0.0825	8.0381
18	0.0255 (0.01)	3.9670 (0.05)	−4.5241 (0.14)	0.1228	8.3896
19	0.0658 (0.02)	3.7386 (0.06)	−4.9284 (0.14)	0.2412	20.9797

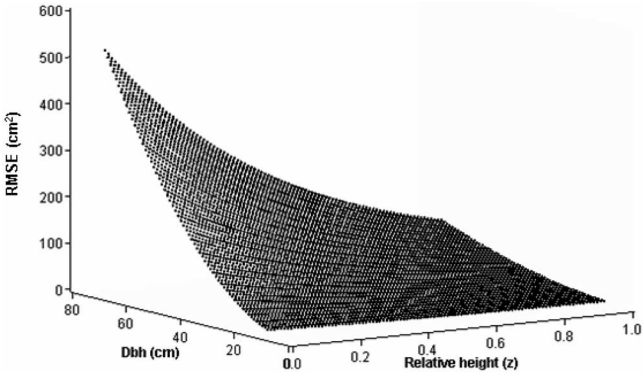

Figure 4. Magnitude of the root mean-squared error (RMSE) in relation to tree dbh (cm) and relative height for species group 8 (sugar maple).

hemlock (species group 5), respectively. Trends with height classes were much less apparent in RMSRR values, in which error in relation to bole diameter was relatively stable. For the most part, the mean relative error ranged between 1.0 and 5.0% across all height classes and all species groups.

The improved fit to individual trees due to the inclusion of random-effects parameters is shown in Figure 5. The better fit is apparent via visual inspection as the prediction line is more aligned with the observed data under a mixed-effects model (Figure 5b). The enhanced predictive capability in the lower bole is most obvious on the largest tree, for which the substantial diameter change between the first two measurements is more accurately described. Note also that better predictions are obtained throughout the entire tree length with the mixed-effects model-fitting technique.

No closed-form solution exists for 3; thus, computation of outside-bark cubic volume must be accomplished through numerical integration. The taper-based volumes were compared with those derived from application of Smalian's formula (Avery and Burkhart 2002, p. 101), where volumes of each section of the tree were computed and section volumes were summed to the heights of 10.1 cm top diameter (pulpwood) and 17.8 cm (softwood)/22.9 cm (hardwood) top diameter (sawlog). Table 6 shows the results for both mixed-effects and fixed-effects (φ_1, $\varphi_2 = 0$) models. The mean differences in volume estimates are not statistically different from zero for either the fixed-effects or mixed-effects model. However, the inclusion of random-effects parameters in the mixed-effects model results in smaller SDs of volume differences. Both the mean percent absolute difference and the SD of the percent absolute differences are substantially reduced by the mixed-effects model. The results for the mixed-effects model are similar to those obtained by Clark et al. (1991), who used upper stem diameter information to calibrate the predictions to individual tree characteristics. When the fixed-effects model is implemented, the percent absolute errors were roughly twice as large as those reported by Clark et al. (1991).

One would expect that taper models developed for relatively small areas would perform better than those applicable to large geographic areas. For example, Jiang et al. (2005) presented compatible taper and volume equations for yellow-poplar in two regions of West Virginia. These local equations produced SEs of volume estimates that were approximately 0.005–0.01 m^3. In contrast, the model presented here had SEs ranging from roughly 0.05 to 0.10 m^3.

Validation

An evaluation of model predictive accuracy was undertaken using independent data. No other data are known for the area that covers the spatial extent and species diversity found in the study data. As such, data from wood utilization studies conducted by the U.S. Forest Service North Central Research Station (NCRS) (now part of the Northern Research Station) and tree taper data collected by U.S. Forest Service Eastern Region (R9) were combined into a validation data set (Table 7). The NCRS data have no geographic overlap with the data used to fit the models; however, many of the same species are present. The R9 data were collected entirely on national forestland, some of which was within the study area (about 20% of the R9 data).

Table 5. Summary of mean absolute deviation (MAD) and root mean square relative residual (RMSRR) by relative height class for 19 species groups

	Relative height (z)									
Species group	0.00–0.04	0.05–0.14	0.15–0.24	0.25–0.34	0.35–0.44	0.45–0.54	0.55–0.64	0.65–0.74	0.75–0.84	0.85–0.94
1										
MAD (cm^2)	49.9	34.9	42.5	31.5	30.6	30.9	27.6	20.7	23.2	14.8
RMSRR (%)	1.8	1.5	2.1	1.5	1.5	1.7	2.0	1.7	2.8	3.1
2										
MAD (cm^2)	85.9	57.2	95.3	100.3	113.6	59.4	75.2	44.5	29.0	38.6
RMSRR (%)	2.5	2.2	3.4	3.4	4.1	2.7	5.0	3.4	4.7	8.1
3										
MAD (cm^2)	72.5	27.5	47.8	56.7	48.1	36.7	27.7	22.8	22.4	14.5
RMSRR (%)	2.7	1.6	2.2	2.8	2.5	2.1	2.1	2.2	3.6	3.9
4										
MAD (cm^2)	27.7	24.3	38.4	39.7	28.9	25.2	22.3	21.6	15.4	13.5
RMSRR (%)	1.3	1.5	2.1	2.1	1.9	1.8	1.9	2.5	2.9	3.8
5										
MAD (cm^2)	175.3	75.4	112.6	94.9	71.9	59.0	45.0	41.8	28.1	18.2
RMSRR (%)	5.0	3.0	4.0	3.6	3.1	3.0	2.9	4.1	3.5	3.9
6										
MAD (cm^2)	69.2	27.1	42.8	39.2	39.1	34.8	24.6	17.7	17.5	11.4
RMSRR (%)	3.1	1.6	2.3	2.1	2.2	2.1	1.9	1.6	2.3	3.0
7										
MAD (cm^2)	71.3	42.9	66.4	69.4	44.0	42.7	31.7	26.3	16.0	11.4
RMSRR (%)	2.4	3.1	3.7	3.3	3.0	3.0	2.6	2.4	2.6	3.7
8										
MAD (cm^2)	80.9	60.9	126.2	112.3	101.5	88.1	60.4	53.2	30.1	13.8
RMSRR (%)	2.4	2.4	4.2	4.1	4.4	4.6	4.2	5.0	5.8	8.1
9										
MAD (cm^2)	80.5	49.0	81.0	92.7	79.6	61.7	40.7	42.9	24.7	11.2
RMSRR (%)	2.6	2.2	3.2	3.8	3.3	3.1	2.5	4.5	3.7	3.6
10										
MAD (cm^2)	48.3	41.2	65.2	71.0	94.4	88.3	36.7	32.6	25.1	10.9
RMSRR (%)	1.8	2.1	3.0	3.3	4.9	6.0	3.3	3.7	4.8	3.6
11										
MAD (cm^2)	61.9	37.2	82.4	82.5	63.8	65.0	44.7	39.6	27.5	15.0
RMSRR (%)	2.2	1.8	3.5	3.8	3.3	3.5	3.0	3.5	3.7	5.2
12										
MAD (cm^2)	61.1	38.6	75.8	85.4	61.5	44.6	43.5	32.0	19.9	10.4
RMSRR (%)	2.3	2.0	3.6	4.2	3.4	3.5	3.4	4.3	4.2	3.2
13										
MAD (cm^2)	103.2	68.9	137.4	129.2	128.7	85.7	61.4	50.7	35.1	13.4
RMSRR (%)	3.0	2.7	5.0	4.9	5.5	4.1	7.3	5.1	5.2	4.0
14										
MAD (cm^2)	120.4	69.7	95.0	85.0	92.4	92.9	57.4	47.1	39.2	15.2
RMSRR (%)	3.6	2.8	4.0	3.5	3.9	4.2	3.5	3.7	5.8	3.4
15										
MAD (cm^2)	102.2	55.7	96.1	106.3	93.4	78.3	64.8	44.4	25.4	14.2
RMSRR (%)	2.9	2.1	3.1	3.9	3.9	4.0	4.5	3.9	3.5	3.1
16										
MAD (cm^2)	100.8	74.4	125.6	130.0	138.5	97.0	67.9	52.0	32.9	15.3
RMSRR (%)	2.8	2.8	4.0	4.6	5.8	6.0	4.0	4.8	4.1	3.7
17										
MAD (cm^2)	55.2	50.7	82.9	55.9	73.4	60.1	36.3	27.7	23.6	10.5
RMSRR (%)	2.0	2.7	4.2	3.0	5.4	4.1	2.9	3.3	3.7	3.5
18										
MAD (cm^2)	103.2	66.3	119.8	131.0	109.8	73.1	62.5	39.5	29.7	14.5
RMSRR (%)	3.2	2.9	5.1	5.8	5.0	3.9	4.5	3.6	3.8	3.5
19										
MAD (cm^2)	102.6	79.0	134.4	139.4	125.5	91.3	46.6	38.3	27.0	12.8
RMSRR (%)	2.9	3.2	5.1	5.2	5.6	5.0	3.3	3.6	4.2	3.4

Model predictions of outside bark diameter squared (d^2) and outside bark cubic volumes (via integration) were compared with the observed data for each species group. The predictions were calculated with random-effects parameters set equal to zero (the expected value). Table 8 provides the same error statistics that were used above for d^2 predictions; the MAD and percent RMSRR. Results were similar to those obtained with the model-fitting data (Table 5) in that the largest MAD tended to occur in the lower relative height classes and MAD generally decreased as height increased.

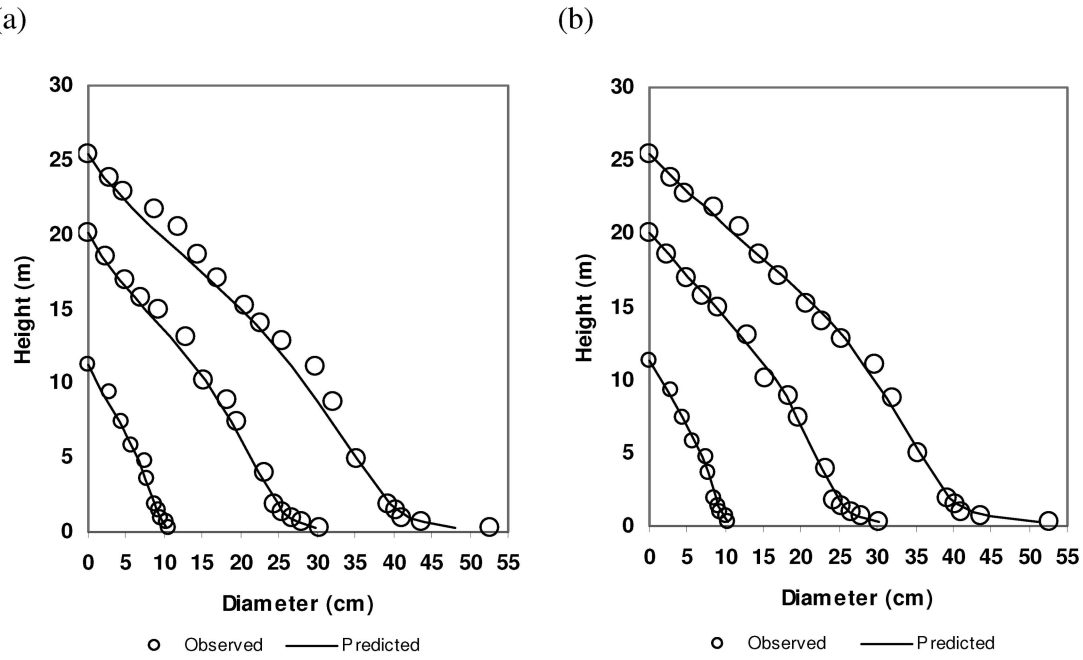

Figure 5. Comparison between observed data and predictions for (a) fixed-effects and (b) mixed-effects models.

Table 6. Mean difference and mean percent absolute difference (SD) of volume estimates by species group from numerical integration of taper equation 3 with (mixed) and without (fixed) random-effects parameters

Species group	Pulpwood top diameter (10.1 cm dob)				Sawlog top diameter (dob)[1]			
	Fixed		Mixed		Fixed		Mixed	
	Bias (m^3)	Percent |D|	Bias (m^3)	Percent |D|	Bias (m^3)	Percent |D|	Bias (m^3)	Percent |D|
1	0.045 (0.09)	8.4 (6.50)	0.012 (0.02)	3.0 (3.70)	0.058 (0.09)	8.7 (5.80)	0.025 (0.03)	3.9 (3.80)
2	0.014 (0.19)	7.7 (6.60)	0.009 (0.09)	4.2 (3.50)	0.021 (0.20)	7.6 (5.80)	0.015 (0.11)	4.6 (4.10)
3	0.018 (0.10)	9.5 (6.50)	0.009 (0.04)	4.8 (3.60)	0.036 (0.11)	8.1 (5.90)	0.024 (0.06)	5.1 (3.90)
4	0.002 (0.06)	8.4 (6.00)	0.006 (0.02)	4.5 (4.20)	0.011 (0.06)	7.9 (5.10)	0.017 (0.03)	5.7 (3.80)
5	0.014 (0.15)	9.1 (6.60)	0.016 (0.08)	4.7 (3.70)	0.027 (0.16)	8.1 (6.30)	0.031 (0.09)	5.6 (4.60)
6	0.032 (0.11)	9.6 (6.70)	0.009 (0.03)	4.7 (4.20)	0.055 (0.12)	8.6 (5.90)	0.025 (0.04)	5.5 (3.40)
7	0.012 (0.09)	9.6 (10.80)	0.007 (0.04)	5.7 (7.30)	0.029 (0.09)	8.8 (6.30)	0.022 (0.04)	5.8 (4.00)
8	−0.024 (0.20)	9.0 (8.10)	0.004 (0.07)	4.6 (3.40)	0.006 (0.20)	8.1 (6.10)	0.029 (0.10)	5.4 (4.50)
9	−0.005 (0.17)	7.2 (6.40)	0.021 (0.11)	4.5 (3.80)	0.039 (0.20)	6.2 (4.90)	0.063 (0.16)	4.8 (3.60)
10	0.018 (0.11)	9.5 (7.20)	0.010 (0.05)	4.2 (3.90)	0.043 (0.11)	8.4 (5.00)	0.043 (0.07)	5.9 (3.70)
11	0.003 (0.16)	9.5 (9.20)	0.012 (0.05)	4.4 (3.50)	0.031 (0.17)	8.9 (6.40)	0.042 (0.07)	5.5 (4.30)
12	0.004 (0.13)	9.6 (6.90)	0.009 (0.05)	6.1 (4.50)	0.027 (0.18)	9.3 (6.40)	0.037 (0.08)	7.0 (4.60)
13	−0.038 (0.21)	9.0 (7.00)	0.004 (0.09)	4.8 (3.30)	−0.019 (0.24)	7.9 (6.10)	0.035 (0.14)	6.3 (3.80)
14	−0.029 (0.25)	7.1 (6.30)	0.004 (0.12)	4.4 (3.70)	−0.012 (0.30)	7.0 (6.30)	0.034 (0.17)	5.4 (4.30)
15	−0.008 (0.20)	7.9 (5.30)	0.009 (0.07)	4.3 (3.60)	0.024 (0.20)	7.7 (5.00)	0.041 (0.11)	5.3 (4.20)
16	−0.007 (0.21)	9.1 (7.70)	0.012 (0.09)	4.4 (3.40)	0.013 (0.21)	8.4 (5.90)	0.039 (0.11)	6.0 (4.30)
17	0.014 (0.16)	9.6 (8.30)	0.006 (0.08)	6.0 (4.10)	0.039 (0.23)	10.9 (10.00)	0.040 (0.13)	8.0 (4.90)
18	−0.020 (0.29)	10.8 (11.80)	0.014 (0.12)	4.7 (4.70)	0.005 (0.31)	9.9 (11.50)	0.045 (0.16)	6.1 (4.80)
19	0.000 (0.22)	10.7 (8.60)	0.016 (0.11)	5.2 (4.20)	0.040 (0.23)	8.2 (5.80)	0.054 (0.14)	6.6 (5.20)

dob, diameter outside bark.
[1] 17.8 cm (softwood)/22.9 cm (hardwood) top diameter (dob).

Compared with the model-fitting data, the MAD values in the lower bole ($z \leq 0.04$) were approximately three times larger, which to some extent is attributable to the lack of random effects in this evaluation. These differences decreased with increasing height. The RMSRR values were roughly 2 times larger than those observed in analysis of the model-fitting data. There were no notable trends in RMSRR with relative height. Overall, the hardwood species prediction error was larger than that for softwood species.

Table 9 shows the comparisons between taper-based pulpwood volumes and those from Smalian's formula. The mean differences were approximately twice as large as those found in the model-fitting data for the fixed-effects model (Table 6). However, the dispersion in the distributions of these differences remained about the same. For the mean percent absolute differences, the means and SDs were somewhat larger and more variable. The differences between taper-based sawtimber volumes and those from Smalian's formula exhibited characteristics very similar to those of pulpwood volume differences (Table 9).

Table 7. Summary of combined North Central Research Station and U.S. Forest Service Eastern Region R9 validation data by species group for dbh, total height, number of trees, and number of data points

	dbh (cm)			Total height (m)				
Group	Minimum	Mean	Maximum	Minimum	Mean	Maximum	No. trees	No. data points
1	6.6	25.1	58.9	7.3	18.3	26.5	1,279	6,140
2	6.6	35.1	83.8	7.1	19.7	29.8	172	856
3	8.9	24.6	66.0	7.9	16.9	32.3	405	1,669
4	9.7	22.4	46.0	8.0	16.3	24.0	453	1,745
5	14.7	39.1	69.1	8.9	19.1	28.2	73	347
6	5.6	23.6	49.3	2.3	17.4	28.2	1,304	5,672
7	8.9	21.3	43.7	5.5	11.7	25.9	332	864
8	5.1	31.2	95.3	6.7	21.8	30.4	757	4,713
9	22.4	44.5	73.4	19.0	30.6	37.9	36	302
10	7.4	29.5	96.3	7.1	22.4	38.9	2,000	10,815
11	14.2	42.4	94.0	10.2	27.2	34.4	67	580
12	6.6	26.2	71.4	7.8	19.2	27.2	707	2,735
13	11.4	35.1	90.7	12.4	22.6	32.3	57	299
14	13.7	37.1	89.4	14.1	24.1	35.1	207	1,395
15	6.4	39.4	111.8	5.0	21.9	34.0	964	5,173
16	23.6	42.9	66.0	21.3	27.7	32.0	11	117
17	14.7	39.6	98.3	11.5	25.2	32.2	76	321
18	5.8	36.8	94.0	5.5	20.2	32.9	608	2,917
19	4.3	27.4	80.8	7.2	20.2	31.7	562	3,013
All	4.3	29.8	111.8	2.3	20.3	38.9	10,070	49,673

Discussion

The results indicate the modifications made to the original model allowed for significant improvement in describing the observed data. The incorporation of estimated join points and the alteration of the S_2 switching function was accompanied by a concomitant decrease (8%) in the AIC (Akaike 1974) fit statistic. The addition of estimated join points required three additional fixed-effects parameters in the model: one for each of the two join points (γ_1, γ_2) and the third (ψ) for the additional scaling term that accounted for the difference between the estimated join point height and 1.37 m (1.37 m was the fixed join point in the derivation of Valentine and Gregoire (2001)). The estimated parameter ψ results in diameters that do not scale exactly to dbh when height equals 1.37 m. This is not considered a deficiency, however, as neither the new nor original model formulations would otherwise return dbh exactly because the $\alpha_2 S_2$ exponent $\neq 0$ at 1.37 m, such that the last term in the model is not exactly unity.

The changes to the S_2 switching function also contributed to better model fit. It was noted that the S_1 switching function in the original model was not forced to be centered at 1.37 m (the lower join point) and also had random-effects parameters that facilitated adaptation to individual trees. However, S_2 was centered on the upper join point and had no among-tree flexibility in behavior. Two changes were made to S_2 to provide characteristics similar to those of S_1: the center of the switch between model segments was estimated from the data, and the ratio of dbh/H was added to the exponent to make the length of the transition zone sensitive to tree taper. In many cases, the center of the switch (β_1) differed substantially from the upper join point (γ_2), indicating that removal of the forced equality was advantageous. The effect of adding dbh/H was to lengthen the transition between model segments for trees that tapered relatively slowly or shorten the transition for trees exhibiting relatively high rates of taper.

It was also found that moving the placement of a random-effect parameter resulted in better fit statistics for the data used in this study. Ideally, placement of random coefficients is determined by fitting the model to each subject (tree) individually and evaluating which parameters exhibit considerable among-tree variation. For our data, this resulted in satisfactory convergence for only a subset of the trees, thus making it difficult to ascertain the range of variability in parameters across all trees. Given this outcome, the alternative was to evaluate the model fits with random effects at various placements (this was the approach taken in development of the original model). The efficacy of the various placements was evaluated primarily via the AIC statistic and the statistical significance of the variance of the random effect. The best results were obtained by moving the random parameter from the denominator of S_1 and associating it with the α_1 parameter instead. This outcome was not surprising given that the prediction error was largest near the base of the tree and predictions in this lower section are primarily controlled by the α_1 parameter (and to a lesser extent S_1). Thus, modifications that directly affect lower bole predictions among trees should result in improvement.

Valentine and Gregoire (2001) specified a three-parameter variance function 2 in their original model. The new model formulation 4 used only two parameters (the exponent ν_3 was removed) because inclusion of the additional parameter caused difficulties with convergence of the optimization process in model fitting, or the estimated parameter was not significantly different from 1 and the improvement in fit statistics was slight. Conceivably, the extra parameter was unnecessary because the modifications to the model altered the complexity of the variance structure.

Although the MADs tended to decrease with increasing

Table 8. Summary of mean absolute deviation (MAD) and root mean square relative residual (RMSRR) by relative height class for 19 species groups in US Forest Service North Central Research Station/Eastern Region R9 validation data

	Relative height (z)									
Species group	0.00–0.04	0.05–0.14	0.15–0.24	0.25–0.34	0.35–0.44	0.45–0.54	0.55–0.64	0.65–0.74	0.75–0.84	0.85–0.94
1										
MAD (cm^2)	67.5	19.5	41.1	49.6	54.0	48.0	53.2	50.9	43.9	20.6
RMSRR (%)	3.5	1.1	2.3	2.6	3.0	3.2	4.0	4.2	4.7	3.6
2										
MAD (cm^2)	171.4	18.3	92.6	99.3	69.8	75.7	97.8	102.5	77.7	74.5
RMSRR (%)	5.3	1.0	4.0	4.3	3.7	3.9	6.1	10.5	9.2	7.8
3										
MAD (cm^2)	144.8	4.5	37.2	38.2	45.5	51.7	51.7	45.3	41.0	4.7
RMSRR (%)	6.1	0.4	2.0	2.6	3.9	5.1	5.1	5.7	5.6	1.5
4										
MAD (cm^2)	137.3	4.2	16.7	41.8	45.9	48.6	35.7	35.1	29.5	6.5
RMSRR (%)	6.1	0.5	1.2	3.1	4.0	4.9	4.6	4.9	3.8	1.9
5										
MAD (cm^2)	211.3	23.8	111.6	233.2	198.0	168.5	126.9	77.8	35.5	22.0
RMSRR (%)	5.7	1.1	5.2	7.0	6.9	7.9	7.2	6.2	4.0	2.6
6										
MAD (cm^2)	100.1	6.4	39.9	50.6	48.6	54.5	53.5	48.0	41.1	32.4
RMSRR (%)	4.7	0.7	2.8	3.4	4.4	4.8	5.0	5.1	4.5	5.7
7										
MAD (cm^2)	166.5	8.5	17.1	40.7	35.7	35.5	39.2	30.9	31.1	14.7
RMSRR (%)	7.3	1.2	1.8	3.8	3.6	4.2	5.0	3.5	3.7	2.6
8										
MAD (cm^2)	200.5	33.8	106.8	117.1	127.4	147.6	106.3	94.5	46.7	16.5
RMSRR (%)	7.2	3.0	6.3	6.2	7.1	12.8	9.0	11.2	5.7	4.5
9										
MAD (cm^2)	119.3	54.9	101.9	99.4	134.6	164.9	136.7	141.0	104.0	117.3
RMSRR (%)	4.2	1.7	2.8	3.2	5.0	6.9	6.7	11.6	10.0	9.6
10										
MAD (cm^2)	158.7	15.0	65.8	67.6	87.2	83.5	78.7	49.5	45.2	22.1
RMSRR (%)	5.9	1.6	3.1	3.6	5.0	4.9	5.7	5.1	5.5	4.2
11										
MAD (cm^2)	162.8	97.0	212.2	186.1	231.4	180.7	174.0	119.1	54.5	109.0
RMSRR (%)	5.1	4.4	7.8	6.7	9.6	7.7	10.0	7.0	5.0	7.9
12										
MAD (cm^2)	148.7	13.4	57.9	76.2	79.1	70.3	57.0	46.4	50.3	24.7
RMSRR (%)	5.8	2.4	3.7	4.1	5.0	6.0	7.0	5.9	6.9	5.6
13										
MAD (cm^2)	197.5	28.7	111.4	116.1	206.3	103.4	141.8	90.9	47.9	27.5
RMSRR (%)	6.6	1.0	3.9	5.3	8.0	5.8	11.0	8.3	6.8	4.0
14										
MAD (cm^2)	319.9	31.9	123.1	110.3	115.0	107.4	110.8	89.6	60.4	21.2
RMSRR (%)	9.9	1.7	4.8	4.8	5.9	5.5	7.2	8.1	7.7	3.6
15										
MAD (cm^2)	397.2	40.4	161.0	189.8	165.9	179.8	135.6	87.8	53.0	33.0
RMSRR (%)	11.4	2.0	7.6	7.5	7.3	11.3	9.6	8.8	6.6	4.8
16										
MAD (cm^2)	144.3	66.2	95.6	86.5	126.0	167.7	160.2	104.9	88.2	41.8
RMSRR (%)	3.8	2.2	2.6	2.7	4.3	6.8	8.2	9.2	8.7	4.1
17										
MAD (cm^2)	154.8	81.1	152.1	157.7	204.8	301.6	127.3	88.7	82.1	78.2
RMSRR (%)	4.5	2.2	7.9	5.9	8.0	21.6	6.8	10.4	15.5	7.3
18										
MAD (cm^2)	256.5	29.0	132.3	163.0	191.3	176.4	140.1	78.4	56.2	37.8
RMSRR (%)	7.6	2.7	5.2	8.5	10.8	9.9	12.5	9.7	6.8	5.6
19										
MAD (cm^2)	193.0	22.7	76.3	91.9	114.3	76.6	78.0	69.9	39.4	8.8
RMSRR (%)	7.4	1.6	4.5	6.8	8.7	7.1	8.9	8.9	8.1	4.8

relative height, there was considerable variability below relative heights of approximately 0.5 (Table 5). This variability was more prominent in the hardwood groups, which is probably attributable to forking of the stem in this area of the tree. The consequent abrupt decreases in tree diameter associated with a small change in relative height often produce substantial deviations from the general trend predicted by the model. In contrast, RMSRR values were fairly similar across height classes. The RMSRR statistic provides an alternative way to view the magnitude of errors: the

Table 9. Mean difference and mean percent absolute difference (SD) of volume estimates by species group from numerical integration of taper Equation 3 (φ_{1j}, $\varphi_{2j} = 0$) for US Forest Service North Central Research Station/Eastern Region R9 validation data

	Pulpwood top diameter (10.1 cm dob)		Sawlog top diameter (dob)[1]	
Species group	Bias (m^3)	Percent |D|	Bias (m^3)	Percent |D|
1	0.031 (0.06)	12.1 (7.60)	0.031 (0.07)	12.0 (6.90)
2	−0.069 (0.13)	12.2 (8.30)	−0.092 (0.14)	12.3 (8.30)
3	−0.028 (0.06)	14.0 (10.00)	−0.045 (0.07)	16.5 (10.60)
4	−0.024 (0.05)	16.2 (9.00)	−0.040 (0.05)	19.2 (8.40)
5	0.052 (0.15)	16.1 (11.80)	0.051 (0.15)	16.8 (11.60)
6	−0.004 (0.07)	12.7 (9.80)	−0.018 (0.08)	14.1 (9.30)
7	−0.013 (0.05)	23.4 (16.10)	−0.025 (0.05)	26.6 (15.40)
8	−0.015 (0.22)	10.9 (8.80)	−0.047 (0.27)	13.2 (9.30)
9	0.038 (0.20)	8.0 (5.20)	0.024 (0.22)	8.5 (5.10)
10	0.023 (0.12)	10.1 (6.60)	0.001 (0.13)	10.1 (6.40)
11	0.019 (0.28)	10.2 (10.10)	0.022 (0.27)	10.4 (10.00)
12	0.003 (0.07)	11.5 (7.50)	−0.027 (0.11)	12.0 (7.40)
13	−0.052 (0.19)	10.9 (8.00)	−0.091 (0.21)	11.0 (7.10)
14	−0.033 (0.18)	9.9 (8.50)	−0.040 (0.24)	11.3 (8.00)
15	−0.010 (0.22)	12.1 (9.10)	−0.019 (0.21)	12.9 (8.00)
16	0.073 (0.11)	5.1 (4.40)	0.109 (0.11)	5.5 (4.40)
17	−0.070 (0.26)	13.2 (12.50)	−0.141 (0.28)	12.7 (8.00)
18	−0.040 (0.18)	13.2 (9.80)	−0.063 (0.19)	14.8 (8.80)
19	−0.011 (0.15)	11.8 (8.50)	−0.038 (0.18)	12.0 (8.30)

dob, diameter outside bark.
[1] 17.8 cm (softwood)/22.9 cm (hardwood) top diameter (diameter outside bark).

errors are evaluated in relation to tree bole diameter. The nearly invariant percent error is indicative of the heterogeneous variability along the length of the stem.

The intertree and intratree form variation is quantified by the error function 4, and the general trend is illustrated in Figure 4. Across all species groups, deviations of model predictions from observed values were greatest near the base of the tree. A review of the residuals indicated that no trends were evident in relation to latitude, longitude, elevation, or dbh; however, it was noted that trees having the most variability also had relatively large dbh. One reason for this phenomenon is the wavering profile of the lower bole (≤3 m) that often develops with increasing tree size. Essentially, the bole diameter does not always decrease with increasing h, such that sometimes the diameter increases (temporarily) as h increases. This type of pattern makes it difficult to accurately define the form in this area of the tree. Another source of lower bole variation for this study is the lack of circularity of the lower bole. The use of a dendrometer requires that the tree be sighted from a fixed point. The view from any given point may make the bole diameter appear larger or smaller than the actual diameter due to out-of-roundness. As h increases, the stem tends to become circular and the diameter measurement is essentially the same from all view points. One may speculate that other random sources of variation (e.g., micro-site conditions) may also be contributing to the model error near the base of the tree.

Valentine and Gregoire (2001) note that most taper data sets have only a single observation below dbh and recommend additional measurements in the lower bole. Our results indicate there is a tradeoff when one is obtaining more measurements in the lower bole, because the additional data are collected in a highly variable portion of the tree. This finding may suggest that the error in the lower portion of the tree is underestimated in studies in which only one measurement is taken below dbh.

The fitted models were applied to independent data to evaluate predictions of diameter squared (d^2) at relative height z and to compare cubic volumes obtained via integration with those calculated via Smalian's formula. For d^2, the values of MAD and RMSRR were generally larger but exhibited similar patterns related to relative height z as were found in analyses of the model residuals. For volumes, both negative and positive mean differences occurred across species groups. This result suggests that the trees in the validation data may have more taper (negative difference) or less taper (positive difference) than those in the study data. These results are consistent with the results of Jiang et al. (2005), who found that taper of yellow poplar differed between two ecoregions in West Virginia and those of Brooks and Wiant (2008), which indicated form differences across five ecoregions for six hardwood species in the central Appalachian region. Over all groups, the largest differences between observed data points and predicted values occurred near the base of the tree, which was consistent with the model residuals. The similar results for lower-bole deviations between the fit and validation data suggest that a method that better accounts for intertree lower-bole variability is needed.

Comparison of computed cubic volumes with those derived from Smalian's formula indicated that the taper model provides unbiased volume estimates. The improved fit to individual trees afforded by the mixed-effects model was verified by the smaller deviations and SEs compared with those for the fixed-effects model. The similarities in performance with the regional models of Clark et al. (1991) suggest that the use of random effects essentially provides

the same benefit as using upper-stem information for calibration purposes (see below for additional discussion). However, the use of random effects in the regional models did not attain the predictive accuracy of the local models developed for yellow poplar by Jiang et al. (2005). The use of a locally developed model (if available) should provide more accurate estimates than use of the regional model.

The results suggest that prediction of random-effects parameters for new observations will provide better performance compared with implementation of only the fixed-effects model. Prediction of random effects requires some new information by which a better representation of the subject of interest is obtained (Lappi 1991, Fang and Bailey 2001). The efficacy of the random effects is largely dependent on the amount of new information acquired, i.e., location and number of height/diameter pair measurements. The mixed-effects model results are suggestive of the upper limit for prediction accuracy as numerous measurements from each tree were used in model fitting. An in-depth description of use of additional stem measurements to calibrate a mixed-effects taper model is given by Trincado and Burkhart (2006). Implementation of these methods is recommended. Often, useful new information is obtained from height/diameter data in the upper stem that indicates a rate of taper for the tree of interest. Although these methods are appealing, a note of caution is needed. The improvements in model predictions for individual trees rely on measurements of new information that is at least as accurate as that of the modeling data. Substantial inaccuracies or biases could result in poorer predictions than those obtained from the mean-response (fixed-effects) model.

Conclusions

The enhancements to the original model provided increased flexibility to better describe the varying forms of individual trees. Specifically, the replacement of prespecified join points between the model sections with estimated join points allowed the model to better conform to the observed data. In addition, reformulation of the S_2 switching function to relax the restriction of having the function centered on the upper join point γ_2 (to be consistent with the S_1 switching function) and to make the length of the transition between segments sensitive to tree dimensions provided additional functionality. Conjointly, the model fit was substantially improved by moving the random-effect parameter φ_{1j} out of the S_1 switching function and associating φ_{1j} with the fixed-effects parameter α_1. With the use of data for the red/silver maple species group (19), the modifications resulted in a considerable decrease (16%) in the AIC statistic compared with the original model.

A limitation of these models is the inability to account for local variations in tree form. The expense of obtaining a sample of sufficient spatial density to detect localized profile differences at a regional scale was beyond the scope of this study. Users should evaluate the predictive properties of the model described herein before using it in lieu of a locally calibrated taper model.

For species not represented in the study data, coefficients from the species group having the most similar stem form characteristics could be used (with caution). In addition, the validation results suggest that the regional scope of these models may allow for use outside the geographic boundaries of the study area. However, an unknown amount of bias may be introduced when the models are used beyond the range of intended application.

Literature Cited

AKAIKE, H. 1974. A new look at the statistical model identification. *IEEE Trans. Autom. Cont*. 19:716–723.

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

BROOKS, J.R., AND H.V. WIANT, JR. 2008. Ecoregion-based local volume equations for Appalachian hardwoods. *North. J. Appl. For.* 25:87–92.

CLARK, A., III, R.A. SOUTER, AND B.E. SCHLAEGEL. 1991. *Stem profile models for Southern tree species.* US For. Serv. Res. Paper. SE-282. 124 p.

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

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.

GARBER, S.M., AND D.A. MAGUIRE. 2003. Modeling stem taper of three central Oregon species using nonlinear mixed effects models and autoregressive error structures. *For. Ecol. Manag.* 179:507–522.

JIANG, L., J.R. BROOKS, AND G.R. HOBBS. 2007. Using crown ratio in yellow-poplar compatible taper and volume models. *North. J. Appl. For.* 24:271–275.

JONES, R.H. 1990. Serial correlation or random subject effects? *Commun. Stat. Simul. Comp.* 19:1105–1123.

KNOEBEL, B.R., H.E. BURKHART, AND D.E. BECK. 1984. Stem volume and taper functions for yellow-poplar in the southern Appalachians. *South. J. Appl. For.* 8:185–188.

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

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

LAPPI, J. 1991. Calibration of height and volume equations with random parameters. *For. Sci.* 37:781–801.

MARTIN, A.J. 1981. *Taper and volume models for selected Appalachian hardwood species.* US For. Ser. Res. Paper NE-490. 22 p.

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

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

SCOTT, C.T. 1981. *Northeastern forest survey revised cubic-foot volume models.* US For. Serv. Res. Note NE-304. 3 p.

SOLOMON, D.S., T.D. DROESSLER, AND R.C. LEMIN, JR. 1989. Segmented quadratic taper models for spruce and fir in the Northeast. *North. J. Appl. For.* 6:123–126.

SULLIVAN, A.D., AND M.R. REYNOLDS, JR. 1976. Regression problems from repeated measurements. *For. Sci.* 22:382–385.

SWINDEL, B.F. 1968. On the bias of some least-squares estimators of variance in a general linear model. *Biometrika* 55:313–316.

TER-MIKAELIAN, M.T., W.T. ZAKRZEWSKI, G.B. MACDONALD, AND D.H. WEINGARTNER. 2004. Stem profile models for young trembling aspen in northern Ontario. *Ann. For. Sci.* 61: 109–115.

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.

WILLIAMS, T.B., AND H.V. WIANT, JR. 1994. Evaluation of nine taper systems for four Appalachian hardwoods. *North. J. Appl. For.* 11:24–26.

YANG, Y., S. HUANG, AND S.X. MENGR. 2009a. Development of a tree-specific stem profile model for white spruce: A nonlinear mixed model approach with a generalized covariance structure. *Forestry* 82:541–555.

YANG, Y., S. HUANG, G. TRINCADO, AND S.X. MENGR. 2009b. Nonlinear mixed-effects modeling of variable-exponent taper equations for lodgepole pine in Alberta, Canada. *Eur. J. For. Res*. 128:415–429.

ZAKRZEWSKI, W.T., AND D.W. MACFARLANE. 2006. Regional stem profile model for cross-border comparisons of harvested red pine (*Pinus resinosa* Ait.) in Ontario and Michigan. *For. Sci*. 52:468–475.