Evaluation of Four Methods to Estimate Parameters of an Annual Tree Survival and Diameter Growth Model

Quang Cao and Mike Strub

Abstract: An approach to simultaneously estimate parameters of an annual tree growth model was developed, in which the sum of log-likelihood functions for tree survival and diameter growth was maximized. Four methods for acquiring interim values of stand density were evaluated: (1) Updating Attributes, in which individual tree values were summarized at the end of each year within the growth period to predict interim stand-level attributes, (2) Predicting Attributes, in which stand attributes were predicted annually using a stand-level model, (3) Linear Interpolation, in which stand attributes were predicted by linear interpolation, and finally (4) Initial Values, in which stand attributes at the beginning of the growth period were used as predictors throughout the growing period, and the rate of change for tree survival and diameter was assumed to be constant for this period. Data from the Southwide Seed Source Study of loblolly pine (*Pinus taeda* L.) showed that, overall, the Updating Attributes and Predicting Attributes produced better evaluation statistics in predicting tree survival and diameter growth than did the other two methods. A simulation study confirmed that these two methods produced the least biased estimates of parameters. Compared with the Updating Attributes method, the Predicting Attributes method produced similar evaluation statistics in predicting tree survival and diameter growth and similar bias in estimating model parameters. The Predicting Attributes method therefore offers a reasonable alternative to the Updating Attributes method, because of the ease of programming in available software languages. FOR. SCI. 54(6):617–624.

Keywords: annual prediction, individual tree model, simultaneous estimation, maximum likelihood estimation, *Pinus taeda*

TREE GROWTH MODELS often describe annual changes in terms of survival and growth of individual trees. The probability that a tree survives the following year has been modeled by use of logistic regressions (Hamilton 1974, Hamilton and Edwards 1976, Monserud 1976, Buchman 1979, 1983, Zhang et al. 1997, Monserud and Sterba 1999) or other methods (Glover and Hool 1979, Amateis et al. 1989, Guan and Gertner 1991a, 1991b). Tree diameter growth equations have been routinely used to predict annual or periodic diameter growth (Stage 1973, Hahn and Leary 1979, Belcher et al. 1982, Burkhart et al. 1987, Amateis et al. 1989, Zhang et al. 1997).

The problem of estimating parameters of the annual tree growth model becomes more complicated because trees are often not measured every year but at some interval, which might vary from plot to plot even in the same study. The standard method until recently to handle this problem has been to assume a constant rate of tree survival and growth during the entire growing period and to use stand attributes at the beginning of the growth period for predicting average annual increment. This assumption is too simplistic because as time passes, both stand variables (stand height, density, and other) and tree variables (diameter and height) change and, as a result, tree survival and diameter growth should vary from year to year.

Improvement to the constant rate method started with interpolation methods developed by McDill and Amateis (1993) for modeling one variable (e.g., tree height) and later generalized by Cao et al. (2002) for many variables (tree diameter, height, and crown ratio). The interpolation methods above still assumed a constant survival rate for the growth period. Cao (2000) introduced an iterative method to account for variable rates of both tree survival and diameter growth. In this method, survival probability and diameter of each tree in the plot were predicted, and interim values of stand density (number of trees and basal area per *ha*) were updated for each year. Nord-Larsen (2006) used a similar method for modeling tree growth data with highly irregular measurement intervals.

Updating interim values of stand density annually from predicted tree survival and diameter growth is cumbersome in estimating the regression parameters and also in applying the final model. Cao (2002, 2004) alleviated this problem by using a stand-level model to predict intermediate values of stand density annually throughout the growing period.

The objective of this study was to evaluate four methods for estimating parameters of a tree survival and diameter growth system, in which interim annual stand density was updated either by summing intermediate tree attributes, from a stand-level model, from interpolated values, or from initial conditions. A maximum likelihood approach was developed to simultaneously estimate parameters of both equations in the growth system for all methods.

Data

Data used in this study were from loblolly pine (*Pinus taeda* L.) plantations in the Southwide Seed Source Study,

Quang Cao, Louisiana State University, School of Renewable Natural Resources, Baton Rouge, LA 70803—Phone: (225) 578-4218; Fax: (225) 578-4227; qcao@lsu.edu. Mike Strub, Weyerhaeuser—mike.strub@weyerhaeuser.com

Manuscript received September 28, 2007, accepted August 6, 2008

which included 15 seed sources planted at 13 locations across 10 southern states (Wells and Wakeley 1966). Each plot consisted of 49 trees planted at a 1.8 × 1.8 m spacing. Plot area was therefore approximately 0.0164 ha. Tree diameters were measured at ages 10, 15 (or 16), 20 (or 22), and 25 (or 27) years.

The fit data set, consisting of 100 plots, was randomly selected from the original data. Four methods of estimating parameters of the tree model were applied on consecutive growth periods from the fit data (total of 300 growth periods), ranging from 4 to 7 years. The validation data sets were formed by randomly selecting 100 plots from the remaining plots. All possible growth periods from the validation data (total of 600 growth periods), ranging from 4 to 17 years, were used for evaluating the two methods. Table 1 shows distribution of these plots by measurement age. Summary statistics of stand and tree variables are presented in Table 2 for the fit and validation data sets.

Methods

Forest modelers have generally described individual tree survival and growth as functions of stand attributes (stand age, site quality, and stand density), tree attributes (tree diameter and height), and some measure of competition (e.g., total basal area of trees having larger diameters). The following system of equations was adopted after preliminary analyses to predict tree annual survival and diameter growth, respectively:

$$p_{i,t+1} = \left[1 + \exp\left(\alpha_1 + \frac{\alpha_2}{A_t} + \alpha_3\hat{H}_t + \alpha_4\mathrm{RS}_t + \frac{\alpha_5\hat{d}_{i,t}}{\hat{D}q_t}\right)\right]^{-1}, \tag{1a}$$

$$\hat{d}_{i,t+1} = \hat{d}_{i,t} + \exp\left[\beta_1 + \frac{\beta_2}{A_t} + \beta_3\mathrm{RS}_t + \beta_4\mathrm{bal}_{i,t} + \beta_5\ln(\hat{d}_{i,t})\right], \tag{1b}$$

where $p_{i,t+1}$ is the probability that the ith tree survives at time $(t + 1)$, given that it was alive at time t; A_t is stand age in years at time t; $\hat{H}_t$ is average height of the dominants and codominants in m at time t, predicted from $\hat{H}_t = \exp\{\lambda_1 + [\ln(\hat{H}_{t-1}) - \lambda_1](A_{t-1}/A_t)\}$; $\mathrm{RS}_t = (10{,}000/\hat{N}_t)^{0.5}/\hat{H}_t$ is the relative spacing at time t; $\hat{N}_t$ is stand density at time t in terms of number of trees per ha; $\hat{B}_t$ is stand density at time t in terms of basal area (m^2/ha); $\hat{d}_{i,t}$ is dbh in cm of the ith tree at time t; $\hat{D}q_t$ is the quadratic mean diameter in cm at time t; $\mathrm{bal}_{i,t}$ is basal area in m^2/ha at time t of the trees having diameters larger than $\hat{d}_{i,t}$; $\ln(\cdot)$ is the natural logarithm; and α, β, and λ are regression coefficients.

Four methods for deriving interim values when estimating parameters of the tree growth and survival models were evaluated: (1) Updating Attributes, in which individual tree values were summarized at the end of each year within the growth period to predict interim stand-level attributes; (2) Predicting Attributes, in which stand attributes were predicted annually using a stand-level model; (3) Linear Interpolation, in which stand attributes were predicted by linear interpolation; and finally (4) Initial Values, in which stand attributes at the beginning of the growth period were used as predictors throughout the growing period, and the rate of change for tree survival and diameter was assumed to be constant for this period.

Method 1: Updating Attributes

In this method, survival probability and diameter of each tree in the plot were predicted annually. Interim values of stand density were updated for each year from

$$\hat{N}_t = (1/s)\sum p_{i,t}, \quad \text{and} \tag{2a}$$

$$\hat{B}_t = (K/s)\sum p_{i,t}\hat{d}_{i,t}^2, \tag{2b}$$

where $K = \pi/40{,}000$ is a factor to convert diameter from cm into basal area in m^2, s is plot size in ha, and the summation sign includes all trees in the plot.

Annual values of $\mathrm{bal}_{i,t}$ for tree i were updated from Equation 2b, except that the summation sign in this case included only trees having diameters larger than $\hat{d}_{i,t}$. $\hat{D}q_t$ was calculated as the square root of the average of the $\hat{d}_{i,t}$ squared.

Method 2: Predicting Attributes

Instead of updating interim values of stand density by annually summarizing plot data, the following stand-level model was used to predict intermediate values of stand density throughout the growing period:

$$\hat{N}_{t+1} = \exp\left[\frac{A_t}{A_{t+1}}\ln(\hat{N}_t) + \left(1 - \frac{A_t}{A_{t+1}}\right)\left(\gamma_1 + \frac{\gamma_2}{A_t} + \gamma_3\hat{N}_t\right)\right], \tag{3a}$$

$$\hat{D}q_{t+1} = \exp\left[\frac{A_t}{A_{t+1}}\ln(\hat{D}q_t) + \left(1 - \frac{A_t}{A_{t+1}}\right)\left(\delta_1 + \frac{\delta_2}{A_t} + \delta_3\mathrm{RS}_t + \delta_4\hat{D}q_t\right)\right], \tag{3b}$$

Basal area per ha of trees larger than the ith tree in terms of diameter was projected annually as

$$\mathrm{bal}_{i,t+1} = \mathrm{bal}_{i,t} + \exp\left[\kappa_1 + \frac{\kappa_2}{A_t} + \frac{\kappa_3\hat{H}_t}{A_t} + \kappa_4\mathrm{RS}_t + \kappa_5\hat{N}_t + \kappa_6\mathrm{bal}_{i,t}\right]. \tag{4}$$

Table 1. Distribution of plots in the fit and validation data set, by measurement age

Measurement ages (yr)	No. of plots: Fit	No. of plots: Validation
10, 15, 20, 25	61	64
10, 15, 20, 27	12	8
10, 15, 22, 27	12	8
10, 16, 20, 27	15	20
Total	100	100

Table 2.   Summary statistics of stand and tree variables for the fit and validation data sets

Age (yr)	No. of plots	Number of trees/ha		Basal area (m^2/ha)		No. of trees	Tree diameter (cm)	
		Mean	SD	Mean	SD		Mean	SD
Fit data								
10	100	2,220	442	22.41	6.53	3,638	10.9	3.2
15	85	1,939	495	30.38	6.39	2,701	13.6	3.8
16	15	1,184	195	28.27	3.59	291	17.0	3.8
20	88	1,266	352	32.45	6.96	1,825	17.6	4.0
22	12	951	126	19.23	6.47	187	15.5	4.3
25	61	1,171	362	36.00	9.29	1,170	19.3	4.4
27	39	803	198	29.65	7.29	513	21.0	5.3
Validation data								
10	100	2,160	400	23.49	6.69	3,540	11.3	3.3
15	80	1,817	501	30.83	6.73	2,382	14.2	3.9
16	20	1,297	207	30.15	3.83	425	16.7	4.0
20	92	1,219	331	32.84	8.05	1,838	18.0	4.3
22	8	938	230	19.15	5.84	123	15.6	3.9
25	64	1,091	383	35.73	11.04	1,144	19.8	4.8
27	36	786	201	30.93	7.16	464	22.2	5.3

Method 3: Linear Interpolation

Intermediate values of independent variables in Equations 1a and 1b were obtained through linear interpolation in this method. If x_t and x_{t+q} are a stand variable measured at the beginning (t) and end ($t + q$) of a q-year growing period, respectively, then the value of that variable at time $t + j$ is determined as

$$x_{t+j} = x_t + (x_{i+q} - x_t)(j/q), \quad i = 0, 1, \ldots, q. \tag{5}$$

Method 4: Initial Values

In the three methods outlined above, tree survival probability and diameter growth are allowed to vary from year to year during the growth period. The Initial Values method, added to the evaluation for completeness, assumes that the survival probability and diameter growth of every tree in the stand remain unchanged during the growing period. Equations 1a and 1b are simplified as

$$plive_i = \left[1 + \exp\left(\alpha_1 + \frac{\alpha_2}{A_t} + \alpha_3 H_t + \alpha_4 \mathrm{RS}_t + \frac{\alpha_5 d_{i,t}}{\hat{D}q_t}\right)\right]^{-q}, \tag{6a}$$

$$\Delta d = \exp\left[\beta_1 + \frac{\beta_2}{A_t} + \beta_3 \mathrm{RS}_t + \beta_4 \mathrm{bal}_{i,t} + \beta_5 \ln(d_{i,t})\right], \tag{6b}$$

where $plive_i$ is probability that the ith tree survives the q-year growth period, and $\Delta d = (d_{i,t+q} - d_{i,t})/q$.

Simultaneous Estimation of Parameters

Various methods for simultaneous estimation of parameters of different equations in a regression system have been used (Burkhart and Sprinz 1984, Reed and Green 1984, Van Deusen 1988, Zhang et al. 1997, Hasenauer et al. 1998, Cao 2006). A maximum likelihood procedure was developed in this study to estimate parameters of both tree survival and diameter growth equations simultaneously.

The probability of the ith tree surviving the q-year growing period, where q varied from 4 to 7 years in this study, is the product of annual probabilities:

$$plive_i = \prod_{j=t+1}^{t+q} P_{i,j}. \tag{7}$$

Equation 7 can be considered as a modified form of logistic regression. Maximum likelihood estimators can be obtained either indirectly (Cao 2000, 2002, 2004) by using the weighted least-squares approach or directly by maximizing the log-likelihood function (Nord-Larsen 2006). Flewelling and Monserud (2002) recommended either a weighted nonlinear least-squares method or a maximum likelihood method for unequal period lengths. The log-likelihood function in this case is

$$\ln(L_1) = \sum_{i=1}^{n_1} z_i, \tag{8}$$

where n_1 is the total number of trees in the fit data set and $z_i = \ln(plive_i)$ if tree i was alive at the end of the period and $\ln(1 - plive_i)$ otherwise.

For tree diameter, Equation 1b was used repeatedly to produce $\hat{d}_{i,t+q}$, which predicted diameter of tree i at the end of the growth period ($d_{i,t+q}$):

$$d_{i,t+q} = \hat{d}_{i,t+q} + \varepsilon_i, \tag{9}$$

where ε_i = error, assumed to be normally distributed with mean 0 and variance σ^2. The log-likelihood for Equation 9 is as follows:

$$\ln(L_2) = -n_2 \ln(\sigma) - \frac{n_2}{2}\ln(2\pi) - \frac{1}{2\sigma^2}\sum_{i=1}^{n_2}(d_{i,t+q} - \hat{d}_{i,t+q})^2, \tag{10}$$

where n_2 is the total number of trees in the fit data set that survived the entire growing period.

We propose to simultaneously estimate the parameters of both Equations 7 and 9 using the maximum likelihood technique. The combined likelihood function is motivated by considering the following bivariate distribution of tree diameter growth and survival:

$$P(X_i, Y_i) = \begin{cases} P(X_i|Y_i = 1) \cdot P(Y_i = 1), & \text{if } Y_i = 1, \\ P(X_i|Y_i = 0) \cdot P(Y_i = 0), & \text{if } Y_i = 0, \end{cases} \tag{11}$$

where X_i is the diameter growth of tree i during the period, and $Y_i = 1$ if tree i survives the period and 0 otherwise.

In the case of mortality, we assume that the tree died at the beginning of the growth period and there was no diameter growth. Therefore $X_i = 0$, resulting in $P(X_i|Y_i = 0) = 1$. Equation 11 can be rewritten as

$$\ln[P(X_i, Y_i)] = \begin{cases} \ln[P(X_i|Y_i = 1)] + \ln[P(Y_i = 1)], & \text{if } Y_i = 1, \\ \ln[P(Y_i = 0)], & \text{if } Y_i = 0, \end{cases} \tag{12}$$

which leads to

$$\ln(L) = \ln(L_2) + \ln(L_1). \tag{13}$$

An alternate justification is if errors in mortality prediction and diameter growth estimation are independent of each other, then the log-likelihood equation for both events is the sum of the two log-likelihood equations.

The parameters of both Equations 7 and 9 were simultaneously estimated by maximizing $\ln(L)$, or minimizing

$$-\ln(L) = n_2 \ln(\sigma) + \frac{n_2}{2}\ln(2\pi) + \frac{1}{2\sigma^2}\sum_{i=1}^{n_2}(d_{i,t+q} - \hat{d}_{i,t+q})^2 - \sum_{i=1}^{n_1} z_i. \tag{14}$$

In each iteration, the variance, σ^2, was estimated from

$$\left(\frac{1}{n_2 - k_2}\right)\sum_{i=1}^{n_2}(d_{i,t+q} - \hat{d}^*_{i,t+q})^2,$$

where $k_2 = 5$ is the number of parameters in Equation 1b, and $\hat{d}^*_{i,t+q}$ is the predicted value for $d_{i,t+q}$ from the previous iteration.

SAS procedure NLIN (SAS Institute, Inc., 2004) was used to solve the minimization problem of Equation 14. Let us define $y_i = 0$ and

$$\hat{y}_i = \sqrt{-z_i + I_{i,t}\left(\ln(\sigma) + \frac{(d_{i,t+q} - \hat{d}_{i,t+q})^2}{2\sigma^2}\right)},$$

where $I_{i,t} = 1$ if tree i survived at age A_{t+q} and 0 otherwise. Minimizing the sum of squared errors or $\sum(y_i - \hat{y}_i)^2$ was equivalent to minimizing $-\ln(L)$ and yielded maximum likelihood estimates of the parameters of the growth system. The algorithm used to compute $\hat{y}_i$ for the ith tree is shown in the Appendix. The Initial Values approach (method 4) directly predicts survival and growth for a measurement interval and therefore does not requite the logic described in this algorithm.

Evaluation

Short growth periods (4–7 years) based on consecutive growing periods in the fit data set were used for estimating the parameters. On the other hand, evaluation from the validation data involved all possible growing periods, including short (4–7 years), medium (10–12 years), and long (15–17 years) projection periods. The intention was to determine how well the model extrapolated beyond the data range. Criteria used in evaluating predictions of tree survival probability were mean difference (between observed and predicted values) and mean absolute difference. For tree diameters, the evaluation statistics were mean difference, mean absolute difference, and fit index (which is computationally identical to R^2 in linear regression).

Results and Discussion

Parameter estimates and their standard errors for Equations 3 and 4 are presented in Table 3. These equations were used to predict interim values of $\hat{N}_{t+i}$, $\hat{D}q_{t+i}$, and $\text{bal}_{i,t+1}$ to be used in the Predicting Attributes method.

Table 4 shows parameter estimates and their standard errors for the tree survival and diameter growth system from all four methods. All parameters presented in Table 4 were significant at the 5% level. The Akaike information criterion values from all methods were similar, with methods 3 and 4 being slightly better than methods 1 and 2.

Evaluation statistics for predicting tree survival (Figure 1) included mean difference between observed and predicted survival probabilities and mean absolute difference. An unexpected result was that the bias, or mean difference, values were smaller for long projection periods compared with those for medium projection lengths. On average, all methods overpredicted tree survival, except for the Initial Values method at short projections. Method 3 (Linear Interpolation) produced the smallest mean absolute difference values. Evaluation statistics for methods 1 and 2 were

Table 3. Parameter estimates for Equations 3 and 4

Equation	Parameter	Estimate	SE
3a	γ_1	3.9467	0.2096
	γ_2	32.4557	3.3179
	γ_3	0.0002	0.0001
3b	δ_1	3.1544	0.1366
	δ_2	−8.1406	1.0025
	δ_3	1.5231	0.2615
	δ_4	0.0291	0.0045
4	κ_1	15.5784	1.3973
	κ_2	78.8367	4.4454
	κ_3	−5.3739	0.3801
	κ_4	−11.6949	1.2830
	κ_5	−2.0285	0.1557
	κ_6	0.0344	0.0015

Equation form:
(3a) $\hat{N}_{t+1} = \exp[A_t/A_{t+1}\ln(\hat{N}_{t+i}) + (1 - A_t/A_{t+1})(\gamma_1 + \gamma_2/A_t + \gamma_3\hat{N}_t]$.
(3b) $\hat{D}q_{t+i} = \exp[(A_t/A_{t+1})\ln(\hat{D}q_t = (1 - A_t/A_{t+1})(\delta_1 + \delta_2/A_t + \delta_3\text{RS}_t + \delta_4\hat{D}q_t]$.
(4) $\text{bal}_{i,t+i} = \text{bal}_{i,t} + \exp[\kappa_1 + \kappa_2/A_t = \kappa_3\hat{H}_t/A_t + \kappa_4\text{RS}_t + \kappa_5\hat{N}_t \kappa_6\text{bal}_{i,t}]$.

Table 4. Parameter estimates for the tree survival and diameter growth system, by method

Equation	Parameter	Estimate	SE
Method 1: Updating Attributes (AIC = 50,130.47)			
1a	α_1	8.3311	1.0619
	α_2	−16.0255	5.3075
	α_3	−0.1656	0.0307
	α_4	−20.6922	2.5972
	α_5	−4.6547	0.2296
1b	β_1	−3.9862	0.2070
	β_2	12.0646	0.7114
	β_3	3.6259	0.2689
	β_4	−0.0179	0.0015
	β_5	0.7137	0.0492
Method 2: Predicting Attributes (AIC = 50,129.92)			
1a	α_1	7.9605	1.0448
	α_2	−15.6037	5.2580
	α_3	−0.1597	0.0303
	α_4	−19.6350	2.4601
	α_5	−4.5685	0.2596
1b	β_1	−4.0130	0.2554
	β_2	12.8627	0.8418
	β_3	3.5465	0.3290
	β_4	−0.0190	0.0019
	β_5	0.7158	0.0604
Method 3: Linear Interpolation (AIC = 49,461.31)			
1a	α_1	5.1964	0.9547
	α_1	−17.1025	5.1495
	α_3	−0.0577	0.0283
	α_4	−7.5066	1.9364
	α_5	−5.6513	0.2787
1b	β_1	−4.4337	0.2632
	β_2	13.1355	0.8673
	β_3	3.6156	0.3445
	β_4	−0.0209	0.0021
	β_5	0.8527	0.0632
Method 4: Initial Values (AIC = 50,118.71)			
6a	α_1	7.9471	0.8535
	α_2	−8.016	4.0435
	α_3	−0.2036	0.0296
	α_4	−18.0461	1.8575
	α_5	−4.5836	0.2578
6b	β_1	−3.5770	0.2546
	β_2	8.1286	0.7538
	β_3	2.9625	0.3123
	β_4	−0.0216	0.0021
	β_5	0.6825	0.0623

AIC, Akaike information criterion.
Equation form:
(1a) $p_{i,t+i} = [1 + \exp(\alpha_1 + \alpha_2/A_t + \alpha_3\hat{H}_t + \alpha_4 \mathrm{RS}_t + \alpha_5 \hat{d}_{i,t}/\hat{D}q_t]^{-1}$.
(1b) $\hat{d}_{i,t+i} = \hat{d}_{i,t} + \exp[\beta_1 + \beta_2/A_t + \beta_3 \mathrm{RS}_t + \beta_4 \mathrm{bal}_{i,t} + \beta_5 \ln \hat{d}_{i,t}]$.
(6a) $plive_i = [1 + \exp(\alpha_1 + \alpha_2/A_t + \alpha_3\hat{H}_t + \alpha_4 \mathrm{RS}_t + \alpha_5 d_{i,t}/Dq_t]^{-q}$.
(6b) $\Delta d = \exp[\beta_1 + \beta_2/A_t + \beta_3 \mathrm{RS}_t + \beta_4 \mathrm{bal}_{i,t} + \beta_5 \ln(d_{i,t})]$.

similar, with method 1 (Updating Attributes) being slightly better.

The last two methods, based on constant rate (Initial Values) and interpolation (Linear Interpolation), estimate parameters by using observed independent variables. Barreto and Howland (2006) showed that when these methods are applied to simultaneous equations (in this case survival and growth), biased parameter estimates result. They label this simultaneity bias. Simultaneity bias occurs when a system of two or more equations contains the dependent variable from one equation as an independent variable in another equation. They show that methods that use esti-

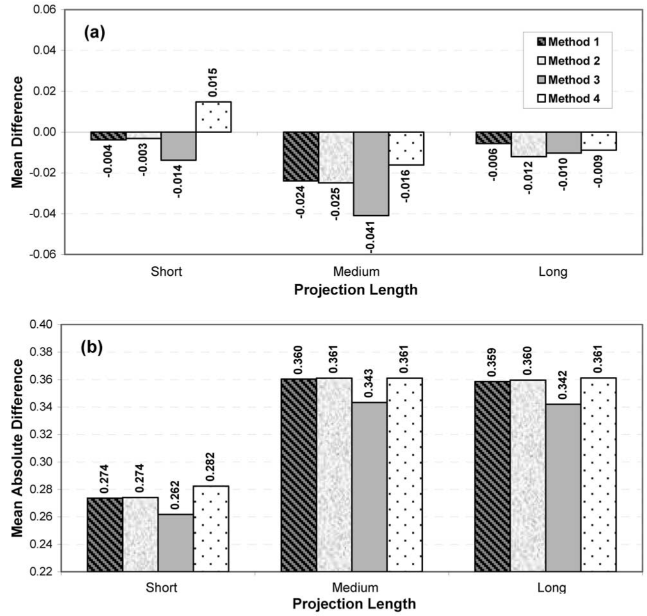

Figure 1. Evaluation statistics based on tree survival probability computed from the validation data set for different projection lengths. These statistics are (a) mean difference and (b) mean absolute difference. The four methods evaluated are (1) Updating Attributes, (2) Predicting Attributes, (3) Linear Interpolation, and (4) Initial Values.

mated (from other equations in the system of simultaneous equations) independent variables result in unbiased parameter estimates. These results are so well known in the econometrics literature that accounts of the history of the discovery of this bias have been reported in Morgan (1990).

The Updating Attributes and Predicting Attributes methods rely on estimated independent variables and therefore should not have simultaneity bias. Results from this study confirmed that bias was indeed less for these two methods.

Figure 2 presents evaluation statistics for predicting tree diameters computed from the validation data set for different projection lengths. These statistics are mean difference between observed and predicted diameters, mean absolute difference, and fit index. Methods 1 and 2 were close, outscoring the other two methods in all cases. The Initial Values method was consistently the worst among all methods. Again as expected, bias was less for the first two methods that rely on estimated independent variables and should not have simultaneity bias.

Table 5 shows results of Duncan's multiple range test performed on the overall mean of the evaluation statistics from the four methods (with projection periods and trees considered as blocks). On the basis of all evaluation statistics, the four methods were classified into three distinct groups that were significantly different from one another at the 5% level. The overall means from methods 1 and 2 are not significantly different; whereas they were significantly different from those produced by methods 3 and 4.

Separate Estimation of Parameters

Alternative methods were explored for estimating parameters of the tree survival and diameter growth equations separately. Least-squares techniques were used to estimate

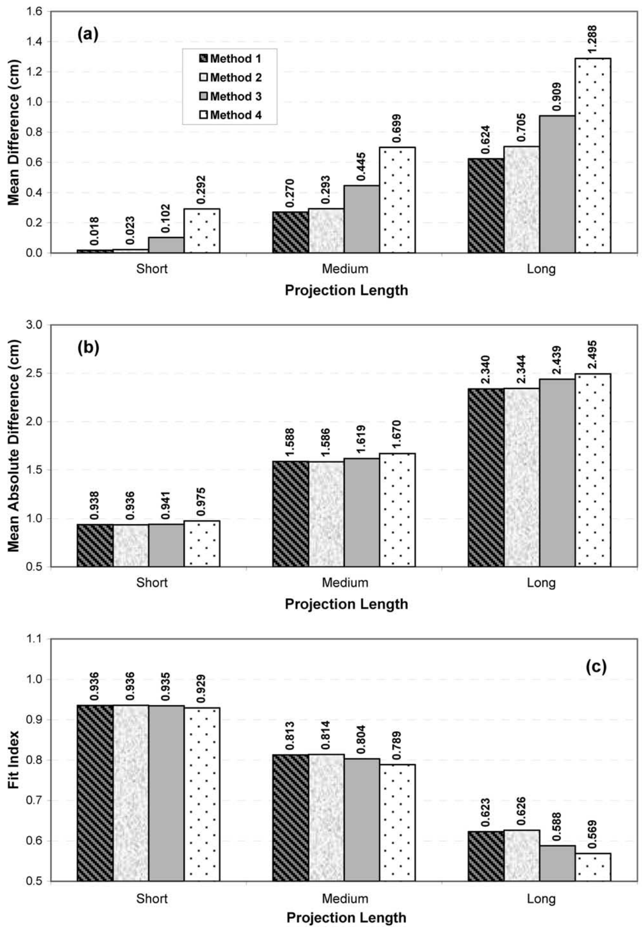

Figure 2. Evaluation statistics based on tree diameter computed from the validation data set for different projection lengths. These statistics are (a) mean difference, (b) mean absolute difference, and (c) fit index (similar to R^2). The four methods evaluated are (1) Updating Attributes, (2) Predicting Attributes, (3) Linear Interpolation, and (4) Initial Values.

parameters of the diameter growth Equation 1b, whereas maximum likelihood estimation of parameters of the survival Equation 1a was obtained through weighted nonlinear regression.

The Updating Attributes method required an iterative procedure outlined by Cao (2000), in which new parameter estimates for one equation were determined by use of the current set of parameters for the other equation. The iterative procedure converged quickly after four iterations. In the Predicting Attributes method, parameters from the diameter growth equation were estimated first, followed by those from the survival equation. This was similar to the two-step procedure developed by Cao (2002).

These alternative methods of separate estimation yielded parameter estimates that were similar to those obtained from simultaneous estimation and produced evaluation results that were almost identical to those from simultaneous estimation.

Simulation Study

Whether or not a method is appropriate for estimating parameters of a model depends on whether the model is sufficiently specified and whether eccentricities were present (e.g., outliers) in the data or both. Monte Carlo analysis was used in this study to ensure that the model was correctly specified and that the four methods could be reliably and accurately evaluated. The method that provided estimates closest to the true (known) parameters was considered the best method.

Simulation growth data were generated from the fit data set, using Equations 1a and 1b for predicting annual changes in tree survival and diameter. Coefficients from the Updating Attributes method (Table 4) were used for this procedure. A total of 500 simulated data sets were generated. Each data set contained initial measurements from the 100 plots in the fit data as the initial data. At the end of each year in the growing period (varying between 4 and 7 years), a tree survived if a uniform (0, 1) random number was less than the survival probability computed for that tree from Equation 1a. Diameter growth of each surviving tree was then computed every year by adding a random term to the predicted annual growth from Equation 1b. The random error term was from a normal distribution with mean zero and variance equal to the mean squared error from Equation 1b. The random error was truncated on both sides at plus or minus predicted annual growth to ensure that positive diameter growth was attained and also that the error distribution remained symmetrical.

Figure 3 shows that methods 1 (Updating Attributes) and 2 (Predicting Attributes) produced estimators that were the least biased (<7%), with method 1 slightly edging method 2. As expected, large bias up to 30 and 60% resulted from method 3 (Linear Interpolation) and method 4 (Initial Values) method, respectively, owing to simultaneity bias.

Updating versus Predicting Attributes

Programming Consideration

The Predicting Attributes method is relatively straightforward and can be easily applied using available statistical software such as SAS proc NLIN. The Updating Attributes is much more difficult to program using NLIN, requiring the use of an array of all tree diameters in a plot for each observation. A similar programming procedure was shown by Strub and Cieszewski (2002) in dealing with estimating parameters of a site index equation.

Furthermore, the Updating Attributes method consumed significant time in summarizing plot information at the end of each year. This method required more than eight times the CPU time needed for the Predicting Attributes method (42.96 versus 5.01 seconds).

Interim Stand Attributes

Resulting stand attributes (number of trees per ha and quadratic mean diameter) at the end of the growth period were similar for both methods, clustering about the 1:1 line. For basal area per ha of trees larger than a subject tree (bal), the Predicting Attributes method matched the Updating Attributes method for small to medium bal, but tended to underpredict for larger bal values. This may be because Equation 4 did not include tree diameter as an independent variable.

Table 5. **Overall mean of evaluations statistics for the four methods**

Evaluation statistic*	Parameter estimating method			
	Updating attributes	Predicting attributes	Linear interpolation	Initial values
Tree survival				
Mean difference	-0.0112^{b}	-0.0124^{b}	-0.0226^{c}	-0.0006^{a}
Mean absolute difference	0.3204^{b}	0.3211^{b}	0.3058^{c}	0.3252^{a}
Tree diameter				
Mean difference	0.1804^{c}	0.2011^{c}	0.3202^{b}	0.5566^{a}
Mean absolute difference	1.3340^{c}	1.3328^{c}	1.3588^{b}	1.4016^{a}

* For each evaluation statistic, means with the same letter in the same row are not significantly different at the 5% level from Duncan's multiple-range test (with projection periods and trees as blocks).

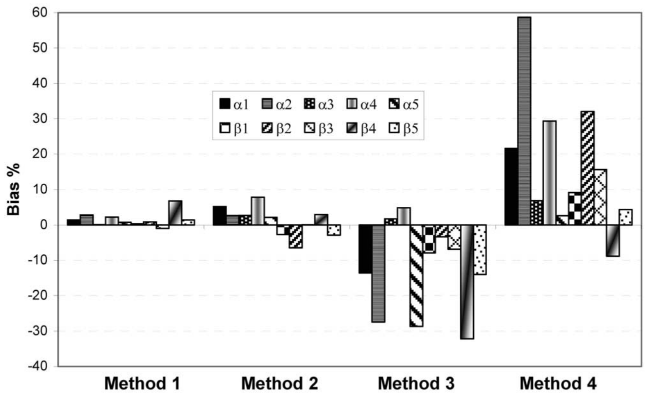

Figure 3. Percent bias in estimating 10 regression parameters using four methods: (1) Updating Attributes; (2) Predicting Attributes; (3) Linear Interpolation; and (4) Initial Values. Average bias was computed from a simulation study involving 500 replications.

Applying the Model

To be consistent, the model should be applied using the same method from which its parameters were obtained. For example, if the Updating Attributes method was used for parameter estimation, then stand attributes in the intermediate years within the growing period should also be updated from the plot summary when it is used to make predictions. However, results showed that using the wrong method when applying the model did not hurt the performance of the model. Similar evaluation statistics were obtained regardless of which method was used in applying the model.

Conclusions

Simultaneous estimation of parameters of the annual tree growth model was obtained by maximizing the sum of log-likelihood functions for tree survival and diameter growth. Overall, the Updating Attributes and Predicting Attributes methods produced better evaluation statistics in predicting tree survival and diameter growth than did the other two methods. Results from an additional simulation study showed that these two methods produced the least biased estimates of parameters. Compared with the Updating Attributes method, the Predicting Attributes method produced similar evaluation statistics in predicting tree survival and diameter growth and similar bias in estimating model parameters. The Predicting Attributes method runs faster and is easier to program in available software languages and therefore offers a reasonable alternative to the Updating Attributes method.

Literature Cited

AMATEIS, R.L., H.E. BURKHART, AND T.A. WALSH. 1989. Diameter increment and survival equations for loblolly pine trees growing in thinned and unthinned plantations on cutover, site-prepared lands. *South. J. Appl. For.* 13:170–174.

BARRETO, H. AND F.M. HOWLAND. 2006. *Introductory econometrics using Monte Carlo simulation with Microsoft Excel.* Cambridge University Press, New York, NY.

BELCHER, D.M., M.R. HOLDAWAY, AND G.J. BRAND. 1982. *A description of STEMS—The stand and tree evaluation and modeling system.* US For. Serv. Gen. Tech. Rep. NC-79. 18 p.

BUCHMAN, R.G. 1979. Mortality functions. P. 47–55 in *A generalized forest growth projection system applied to the Lake States region*. US For. Serv. Gen. Tech. Rep. NC-49.

BUCHMAN, R.G. 1983. *Survival predictions for major Lake States tree species*. US For. Serv. Res. Paper NC-233, 7 p.

BURKHART, H.E., AND P.T. SPRINZ. 1984. Compatible cubic volume and basal area projection equations for thinned old-field pine plantations. *For. Sci.* 30:86–93.

BURKHART, H.E., K.D. FARRAR, R.L. AMATEIS, AND R.F. DANIELS. 1987. *Simulation of individual tree growth and stand development in loblolly pine plantations on cutover, site-prepared areas.* School of Forestry and Wildlife Research, Virginia Polytechnic Inst. and State Univ. Publ. No. FWS-l-87. 47 p.

CAO, Q.V. 2000. Prediction of annual diameter growth and survival for individual trees from periodic measurements. *For. Sci.* 46:127–131.

CAO, Q.V. 2002. Annual tree growth predictions based on periodic measurements. P. 7–13 in *Proc. of the IUFRO symposium on statistics and information technology in forestry*, September 8–12, 2002. Virginia Polytechnic Inst. and State Univ., Blacksburg, VA.

CAO, Q.V. 2004. Annual tree growth predictions from periodic measurements. P. 212–215 in *Proc. of the Southern silviculture research conference.* US For. Serv. Gen. Tech. Rep. SRS-71.

CAO, Q.V. 2006. Predictions of individual-tree and whole-stand attributes for loblolly pine plantations. *For. Ecol. Manag.* 236:342–347.

CAO, Q.V., S. LI, AND M.E. MCDILL. 2002. Developing a system of annual tree growth equations for the loblolly pine-shortleaf pine type in Louisiana. *Can. J. For. Res.* 32:2051–2059.

FLEWELLING, J., AND MONSERUD, R.A. 2002. A comparison of methods for estimating mortality models. P. 168–177 in *Proc. of the Second forest vegetation simulator (FVS) conference*, Crookston, N.L., and R.N. Havis (comps.). February 12–14, 2002, Fort Collins, CO, Proc. RMRS-P-25. US Department of

Agriculture, Forest Service, Rocky Mountain Research Station, Ogden, UT.

GLOVER, G.R., AND J.N. HOOL. 1979. A basal area ratio predictor of loblolly pine plantation mortality. *For. Sci.* 25:275–282.

GUAN, B.T., AND G. GERTNER. 1991a. Using a parallel distributed processing system to model individual tree mortality. *For. Sci.* 37:871–885.

GUAN, B.T., AND G. GERTNER. 1991b. Modeling red pine tree survival with an artificial neural network. *For. Sci.* 37: 1429–1440.

HAHN, J.T., AND R.A. LEARY. 1979. Potential diameter growth functions. P. 22–26 in *A generalized forest growth projection system applied to the Lake States region.* US For. Serv. Gen. Tech. Rep. NC-49.

HAMILTON, D.A., JR. 1974. *Event probabilities estimated by regression*. US For. Serv. Res. Paper INT-152, 18 p.

HAMILTON, D.A., JR., AND B.M. EDWARDS. 1976. *Modeling the probability of individual tree mortality*. US For. Serv. Res. Paper INT-185, 22 p.

HASENAUER, H., R.A. MONSERUD, AND T.G. GREGOIRE. 1998. Using simultaneous regression techniques with individual-tree growth models. *For. Sci.* 44:87–95.

MCDILL, M.E., AND R.L. AMATEIS. 1993. Fitting discrete-time dynamic models having any time interval. *For. Sci.* 39:499–519.

MONSERUD, R.A. 1976. Simulation of forest tree mortality. *For. Sci.* 22:438–444.

MONSERUD, R.A., AND H. STERBA. 1999. Modeling individual tree mortality for Austrian forest species. *For. Ecol. Manag.* 113:109–123.

MORGAN, M. 1990. *A history of econometric ideas*. Cambridge University Press, New York, NY. 296 p.

NORD-LARSEN, T. 2006. Modeling individual-tree growth from data with highly irregular measurement intervals. *For. Sci.* 52:198–208.

REED, D.D., AND E.J. GREEN. 1984. Compatible stem taper and volume ratio equations. *For. Sci.* 30:977–990.

SAS INSTITUTE, INC. 2004. *SAS/STAT 9.1 user's guide.* SAS Institute, Inc., Cary, NC. 5121 p.

STAGE, A.R. 1973. *Prognosis model for stand development.* US For. Serv. Res. Paper INT-137, 32 p.

STRUB, M., AND C.J. CIESZEWSKI. 2002. Fitting global site index parameters when plot or tree site index is treated as a local nuisance parameter. P. 97–107 in *Proc. of the IUFRO symposium on statistics and information technology in forestry.* September 8–12, 2002. Virginia Polytechnic Inst. and State Univ., Blacksburg, VA.

VAN DEUSEN, P.C. 1988. Simultaneous estimation with a squared error loss function. *Can. J. For. Res.* 18:1093–1096.

WELLS, O.O., AND P.C. WAKELEY. 1966. Geographic variation in survival, growth, and fusiform-rust infection of planted loblolly pine. *For. Sci. Monogr.* 11:1–40.

ZHANG, S., R.L. AMATEIS, AND H.E. BURKHART. 1997. Constraining individual tree diameter increment and survival models for loblolly pine plantations. *For. Sci.* 43:414–423.

Appendix

*Algorithm for computing $\hat{\mathbf{y}}_i$.**

Loop for t from A_{t+1} to A_{t+q}
  Loop for tree j from 1 to n_1
    Compute $\hat{d}_{j,t}$ and $p_{j,t}$
    End {loop for tree j}
  Compute $\hat{N}_{t+q}$ and $\hat{B}_{t+q}$:
    Method 1: Updating attributes from plot summary
    Method 2: Predicting attributes from Equation 3
    Method 3: Interpolating intermediate attributes
  End {loop for t}
Compute $\hat{y}_i$.

*

$$\hat{y}_i = \sqrt{-z_i + I_{i,t}\left(\ln(\sigma) + \frac{(d_{i,t+q} - \hat{d}_{i,t+q})^2}{2\sigma^2}\right)},$$

where $I_{i,t}$ is 1 if tree i survived at the end of the growth period and 0 otherwise; σ^2 is the variance of ε_i from $d_{i,t+q} = \hat{d}_{i,t+q} + \varepsilon_i$; $d_{i,t+q}$ and $\hat{d}_{i,t+q}$ are observed and predicted tree diameter at the end of the growth period; $zi = \ln(plive_i)$ if tree i was alive at the end of the period and $\ln(1 - plive_i)$ otherwise; and $plive_i$ is the probability that tree i survived at the end of the period.