Prediction of Annual Diameter Growth and Survival for Individual Trees from Periodic Measurements

Quang V. Cao

ABSTRACT. It is difficult to fit annual tree survival and diameter growth models to data that were measured, not every year, but at some interval. This study aimed to determine suitable methods to obtain parameter estimates for such a system from periodic measurements. Given a system consisting of a tree survival model and a tree diameter growth model, this paper presents an iterative method for estimating system parameters. The method involves sequentially updating of the parameters of both models within the growth period. Data from 111 plots from the Southwide Seed Source Study of loblolly pine (*Pinus taeda* L.) were then used to evaluate this iterative approach against the averaging method that assumes a constant tree survival probability and diameter growth rate during the remeasurement interval. Results indicated that the iterative method out-performed the averaging method in predicting future individual tree survival, diameter growth, and stand basal area. The iterative method was superior because it accounted for the variable rate of diameter growth and tree survival probability as functions of ever-changing stand and tree attributes. FOR. SCI. 46(1):127-131.

Additional Key Words: *Pinus taeda*; logistic regression; maximum likelihood estimators.

TWO IMPORTANT COMPONENTS OF individual tree growth models are a tree survival equation and a diameter growth function. Tree annual survival equations predict the probability that a tree survives the following year. Logistic functions have often been used to model tree mortality or survival (Hamilton 1974, Hamilton and Edwards 1976, Monserud 1976, Buchman 1979, 1983, Zhang et al. 1997), even though some exceptions exist (Glover and Hool 1979, Amateis et al. 1989, Guan and Gertner 1991a, 1991b). Annual diameter increment of a surviving tree has been predicted with tree diameter growth equations (Stage 1973, Hahn and Leary 1979, Belcher et al. 1982, Burkhart et al. 1987, Amateis et al. 1989, Zhang et al. 1997).

Problems are encountered in cases where trees are not measured every year but at some interval. McDill and Amateis (1993) evaluated various methods to fit annual growth models from periodic measurements and recommended two interpolation methods. Both methods performed better than the averaging method that assumes a constant growth for the entire period. Predicting tree annual survival from periodic measurements, on the other hand, is more complicated. Authors such as Monserud (1976) and Hamilton and Edwards (1976) have assumed that this probability is constant for the measured interval, whereas it actually varies annually with changes in stand attributes (stand age, density, etc.) and in tree attributes (diameter, height, crown ratio, etc.).

The objective of this study was to improve the methods for obtaining parameters of an individual tree growth system comprised of annual diameter growth and tree survival equations from periodic measurements.

Data

Southwide Seed Source Study

Data were from the Southwide Seed Source Study, which involved 15 loblolly pine (Pinus taeda L.) seed sources (Wells and Wakeley 1966). Only a subset of the data originating from two seed sources (southeastern North Carolina and southeastern Louisiana) was used in this study. Seedlings from the two seed sources were planted at 12 locations ranging from Texas to the Carolinas. A total of 111 plots of size 0.0164 ha each were repeatedly measured from age 10 to

The author is Associate Professor of Forestry, School of Forestry, Wildlife, and Fisheries, Louisiana Agricultural Experiment Station, Louisiana State University Agricultural Center, Baton Rouge, LA 70803 – Phone: (225) 388-4218; Fax: (225) 388-4227; E-mail: QCAO@LSU.EDU.

Manuscript received January 12, 1999. Accepted August 18, 1999.

Table 1. Distribution of 111 loblolly pine plots from the Seed Source Study, by measurement age.

No. of Plots	Measurement ages
79	10, 15, 20, 25
8	10, 15, 20, 27
8	10, 15, 22, 27
16	10, 16, 20, 27

age 25 or 27 (Table 1). Table 2 shows the distribution of observations from these plots.

Methods

After preliminary analyses, the following system of equations was selected to predict annual tree diameter growth and survival, respectively:

$$D_{i,\,t+1} - D_{i,\,t} = b_1 A_t^{b_2} B_t^{b_3} D_{i,\,t}^{b_4} + \varepsilon_{i,t} \quad (1)$$

$$Plive_{i,t+1} = \{1 + \exp[-(c_1 + c_2 A_t + c_3 B_t + c_4 D_{i,t})]\}^{-1} \quad (2)$$

where

A_t = stand age in years at time t,

B_t = basal area in m^2ha^{-1} at time t,

$D_{i,\,t}$ = diameter in cm of the i^{th} tree at time t,

$Plive_{i,\,t+1}$ = probability that the i^{th} tree survives at time (t + 1), given that it was alive at time t,

b_j and c_j = parameters to be estimated, and

$\varepsilon_{i,\,t}$ = random error.

It should be noted that the implicit assumption for Equation [2] is that $Plive_{i,\,t+1}$ is independent of the tree's history and is a function of current stand variables (age and stand basal area) and tree diameter.

The Tree Survival Model

Let y_i = 1 if tree i survived the q-year growth period (q varied from 4 to 7 yr for the Southwide Seed Source Study data), and 0 otherwise. The probability (p_i) of tree i surviving the period is the product of annual survival probabilities:

$$p_i = \prod_{j=1}^{q} \text{probability that tree } i \text{ survives at time } (t+j) \mid \text{tree } i \text{ was alive at time } (t+j-1)], \text{ or}$$

$$p_i = \prod_{j=1}^{q} Plive_{i,\,t+j}$$

$$= \prod_{j=1}^{q} \{1 + \exp[-(c_1 + c_2 A_{t+j-1} + c_3 B_{t+j-1} + c_4 D_{i,\,t+j-1})]\}^{-1} \quad (3)$$

Steps for computing p_i from Equation (3) are outlined as follows. At year (t+1), the survival probability ($Plive_{i,\,t+1}$) of tree i was predicted from Equation (2), and its diameter ($D_{i,\,t+1}$) was predicted from Equation (1)]. Stand basal area was then estimated from:

$$B_{t+1} = \left(K \sum_{i=1}^{m} Plive_{i,\,t+1}\, D_{i,\,t+1}^2 \right) / a \quad (4)$$

where m = number of trees in a plot, a = plot size in hectares, and $K = \pi/40{,}000$ = a factor to convert diameter in cm into basal area in m^2.

Then the new values of tree diameter, stand age, and basal area were used to predict tree survival and diameter growth for the following year. The process was repeated until the end of the period was reached.

Equation (3) can be considered as a modified form of logistic regression. Maximum likelihood estimators were obtained using the weighted least squares approach introduced by Walker and Duncan (1967).

The Diameter Growth Model

Since tree diameters were not measured annually, the parameters of equation [1] could not be obtained directly. Equation (1) can be rewritten as a series of annual diameter growth equations that cover a period of q years:

$$D_{i,\,t+1} - D_{i,\,t} = b_1 A_t^{b_2} B_t^{b_3} D_{i,\,t}^{b_4} + \varepsilon_{i,t} \quad (5.1)$$

$$\vdots$$

$$D_{i,\,t+q} - D_{i,\,t+q-1} = b_1 A_{t+q-1}^{b_2} B_{t+q-1}^{b_3} D_{i,\,t+q-1}^{b_4} + \varepsilon_{i,t+q-1} \quad (5.q)$$

where $\varepsilon_{i,\,t}$'s are random errors.

Summing equations (5.1) through (5.q) results in

$$D_{i,\,t+q} = D_{i,\,t} + \sum_{j=1}^{q} b_1 A_{t+j-1}^{b_2} B_{t+j-1}^{b_3} D_{i,\,t+j-1}^{b_4} + \eta_i \quad (6)$$

where η_i is random error.

Table 2. Distribution of observations from 111 plots in loblolly pine plantations used from the Seed Source Study.

		Trees/ha		Basal area (m^2/ha)			Diameter (cm)	
Age	No. of Plots	Mean	SD[1]	Mean	SD	No. of Trees	Mean	SD
10	111	1806	677	21.08	8.24	3284	11.7	3.5
15	95	1485	634	28.24	8.20	2312	14.9	4.3
16	16	1289	115	31.31	3.60	338	17.1	4.1
20	103	1034	373	30.94	9.27	1746	18.9	4.9
22	8	961	187	21.05	4.87	126	16.3	3.7
25	79	870	421	33.67	12.20	1127	21.3	5.7
27	32	736	208	30.12	8.67	386	22.1	5.5
All	444	1281	642	28.10	10.21	9319	15.68	5.73

[1] SD = standard deviation.

In this approach, called the summation method, nonlinear regression techniques can be employed to obtain the parameter estimates for Equation (6).

Equation (6) is very similar to the diameter growth model from the "one-step" interpolation method developed by McDill and Amateis (1993) except for the random component. In their formulation, the random error is divided by the proportion of periodic growth attributable to the first year. In a preliminary Monte Carlo analysis, simulated data that mimicked the Southwide Seed Source Study data were used to evaluate these two methods. Estimates of parameters in the annual diameter growth model were found to be closer to the true values for the summation method (equation 6) than for the one-step interpolation method (McDill and Amateis 1993), in terms of mean squared error. Therefore the summation method was used to obtain parameter estimates for the annual diameter growth model.

System of Tree survival and Diameter Growth Models

Two methods for fitting the system of tree annual survival and diameter growth equations to periodic measurement data were evaluated in this study. The averaging method assumed for each tree a constant annual survival probability and annual diameter growth during the entire growing period. The averaging method has been used by most tree-level growth and yield models. The iterative method allowed the survival probability and diameter growth rate of each tree to vary from year to year, depending on stand variables (stand age and basal area) and its current diameter. This method combined the above formulations for tree survival and diameter growth models.

Averaging method

The annual tree survival probability [Equation (2)] was assumed unchanged during the growing period. As a result, the probability of tree *i* surviving the period was

$$p_i = \{1+\exp[-(c_1 + c_2 A_t + c_3 B_t + c_4 D_{i,t})]\}^{-q} \qquad (7)$$

where q = length of growing period (4 to 7 yr for the Southwide Seed Source Study).

Parameters of Equation (7) were then obtained using the maximum likelihood procedures.

In the averaging method, the annual diameter growth [Equation (1)] was computed as the average growth over the period. Equation (1) can be rewritten in terms of average annual diameter growth for the q-year growing period:

$$(D_{i,\,t+q} - D_{i,\,t})\,/\,q \;=\; b_1 A_t^{b_2} B_t^{b_3} D_{i,t}^{b_4} + \varepsilon_{i,t} \qquad (8)$$

Nonlinear regression techniques were used to obtain parameters for Equation (8).

Iterative method

In this method, survival probability and diameter growth for each tree were assumed to vary every year, as part of the environmental changes that influence mortality and growth rate of all trees in a stand. Iterative procedures were therefore necessary. In this approach, the maximum likelihood method was used for the tree annual survival model and the summation method for the diameter growth model. The following algorithm dealt with each equation in a sequential manner:

Step 1: Parameters for Equation (8) obtained earlier from the averaging method were used as starting parameter estimates for the annual diameter growth equation (1).

Step 2: Parameter estimates for the tree survival equation (2) were determined, using the current set of parameters for the diameter growth equation (1).

Step 3: Parameter estimates for the diameter growth equation (1) were determined, using the current set of parameters for the tree survival equation (2).

Step 4: Go back to step 2. Quit when all parameter estimates changed very little (less than 10^{-6}) from one iteration to the next.

Evaluation Criteria

The following criteria were used to evaluate the above two methods in predicting future tree survival, diameter growth, and stand basal area of the 111 remeasured plots from the Southwide Seed Source Study.

Tree survival

1. Mean deviation $= \Sigma(y_i - p_i)/n$
2. Mean absolute deviation $= \Sigma|y_i - p_i|/n$
3. Log-likelihood value

where y_i = 1 (survived) or 0 (dead), p_i = predicted survival probability for tree *i* after a *q*-year period, n = 7806 = total number of trees in the data, and the sum includes values of *i* from 1 to *n*.

Tree diameter growth

1. Mean deviation $= \Sigma(\Delta D_i - \Delta\hat{D}_i)\,/\,n_s$
2. Mean absolute deviation $= \Sigma\,|\,\Delta D_i - \Delta\hat{D}_i\,|\,/\,n_s$
3. Fit index $= 1 - \dfrac{\Sigma(\Delta D_i - \Delta\hat{D}_i)^2}{\Sigma(\Delta D_i - \overline{\Delta D})^2}$

where ΔD_i, $\Delta\hat{D}_i$ = observed and predicted q-year diameter growth (cm) for tree *i*, respectively, $\overline{\Delta D}$ = mean *q*-year diameter growth, n_s = 6033 = total number of surviving trees in the data, and the sum includes values of *i* from 1 to n_s. Fit index is analogous to R^2 in linear regression, with a maximum value of 1 for perfect prediction.

Stand basal area

1. Mean deviation $= \Sigma(B_{j,\,t+q} - \hat{B}_{j,\,t+q})\,/\,m$
2. Mean absolute deviation $= \Sigma\,|\,B_{j,\,t+q} - \hat{B}_{j,\,t+q}\,|\,/\,m$
3. Fit index $= 1 - \dfrac{\Sigma(B_{j,\,t+q} - \hat{B}_{j,\,t+q})^2}{\Sigma(B_{j,\,t+q} - \overline{B}_{t+q})^2}$

where $B_{j,t+q}$ and $\hat{B}_{j,\,t+q}$ = observed and predicted basal area (m^2/ha), respectively, of plot *j* at the end of a *q*-year period,

Table 3. Parameter estimates for the diameter growth and tree survival models[1] based on data from loblolly pine plantations, by method.

Parameter	Averaging method	Iterative method
b_1	2.14064	3.33693
b_2	–1.03790	–1.12541
b_3	–0.55226	–0.58777
b_4	1.11740	1.12378
c_1	3.54019	3.41156
c_2	–0.13331	–0.09397
c_3	–0.09974	–0.07388
c_4	0.31516	0.22882

1 $$D_{i,t+1} - D_{i,t} = b_1 A_t^{b_2} B_t^{b_3} D_{i,t}^{b_4} + \varepsilon_{i,t}$$

$$Plive_{i,t+1} = \{1 + \exp[-(c_1 + c_2 A_t + c_3 B_t + c_4 D_{i,t})]\}^{-1}$$

(variables are defined in text)

$\overline{B}_{t+q}$ = mean basal area at the end of a q-year period, $m = 333$ = total number of plot remeasurements (or plot–age combinations), and the sum includes values of j from 1 to m.

A good method to estimate the parameters of the system should produce low values of mean deviation and mean absolute deviation, and high values of log–likelihood and fit index.

Results and Discussion

The parameter estimates for the diameter growth and tree survival models from the averaging and iterative methods are shown in Table 3. The iterative method converged quickly after 4 iterations.

Table 4 presents evaluation statistics for tree survival, diameter growth, and stand basal area. It is obvious that the iterative method consistently out-performed the averaging (constant growth rate) approach in producing significantly lower values of mean deviation and mean absolute deviation ($P < 0.0001$), while generating higher values of log-likelihood and fit index.

For the tree survival model, the deviation ($y_i - p_i$) measured how close the predicted survival probability was to the observed survival status of tree i. The absolute deviation, $| y_i - p_i |$, can be considered as the probability of wrong prediction; it is equal to $(1 - p_i)$ if the tree survived and p_i otherwise. The mean deviation and the mean absolute deviation were reduced from 0.017 to 0.0006, and from 0.27 to 0.26, respectively, when the averaging method was replaced by the iterative method. The iterative approach also increased log-likelihood from –3513 to –3429.

Table 4. Evaluation statistics for two methods based on tree survival, diameter growth, and stand basal area.

Evaluation Statistic	Averaging method	Iterative method
Tree Survival, trees/ha (n = 7806)		
Mean deviation	0.0166	0.0006
Mean absolute deviation	0.2721	0.2621
Log-likelihood	–3,513	–3,429
Tree Diameter Growth, cm (n = 6033)		
Mean deviation	0.2696	0.0025
Mean absolute deviation	1.1067	1.0769
Fit index	0.1904	0.2135
Stand Basal Area, m2/ha (n = 333)		
Mean deviation	1.0058	0.2379
Mean absolute deviation	4.6840	4.4892
Fit index	0.5530	0.5666

For the diameter growth model, the iterative approach decreased bias from 0.27 cm to 0.002 cm, and mean absolute deviation from 1.11 cm to 1.08 cm, while increasing fit index from 0.1904 to 0.2135.

Stand basal area is the sum of individual basal areas of surviving trees, and thus should reflect the interrelated performance of both tree survival and diameter growth prediction equations. Bias in predicting stand basal area was reduced from 1.01 m2/ha (averaging method) to 0.24 m2/ha (iterative method). This was a reduction in bias from 3.30% to 0.78%. Substituting the iterative method for the averaging method reduced mean absolute deviation from 4.68 to 4.49 m2/ha, while increasing fit index from 0.5530 to 0.5666.

The iterative method provided better results because it accounted for the variable rate of diameter growth and tree survival probability as functions of ever-changing stand and tree attributes. This new approach should work well with other data sets of loblolly pine as well as other species.

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.

BELCHER, D.M., M.R. HOLDAWAY, AND G.J. BRAND. 1982. A description of STEMS – the Stand and Tree Evaluation and Modeling System. USDA For. Serv. Gen. Tech. Rep. NC-79, 18 p.

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

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

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. Sch. For. Wildl. Res., Va. Polytech. Inst. & State Univ., Publ. No. FWS-l-87, 47 p.

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.

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

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

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

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.

STAGE, A.R. 1973. Prognosis model for stand development. USDA For. Serv. Res. Pap. INT-137, 32 p.

WALKER, S.H., AND D.B. DUNCAN. 1967. Estimation of the probability of an event as a function of several independent variables. Biometrika 54:167-179.

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. 40 p.

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.