DIAMETER GROWTH MODELS USING FIA DATA FROM THE NORTHEASTERN, SOUTHERN, AND NORTH CENTRAL RESEARCH STATIONS

Veronica C. Lessard, Ronald E. McRoberts, and Margaret R. Holdaway

ABSTRACT.—Nonlinear, individual-tree, distance-independent annual diameter growth models are presented for species in two ecoregions defined by R.G. Bailey in the northern Lake States and in parts of the central and southern regions of the U.S. The models were calibrated using Forest Inventory and Analysis (FIA) data from undisturbed plots on land classified as timberland across all ownership categories. The data were generally from stands of mixed age and mixed species. The dependent variable is average annual diameter growth and the independent variables include crown ratio, crown class, stand basal area, stand basal area larger than the subject tree, physiographic class, and latitude and longitude of plot locations. The models have minimal bias and may be recalibrated easily to include new data sets.

INTRODUCTION

In response to the 1998 Farm Bill, formally known as the Agricultural Research, Extension, and Education Reform Act, the North Central Research Station (NCRS) of the USDA Forest Service, has developed an annual inventory system featuring a hexagonal grid system of FIA plots to be measured in 5-year inventory cycles, with 20 percent of the plots to be measured each year (Brand *et al.* 2000). Diameter growth models for individual trees provide a method to update the information on FIA plots not measured in the current year.

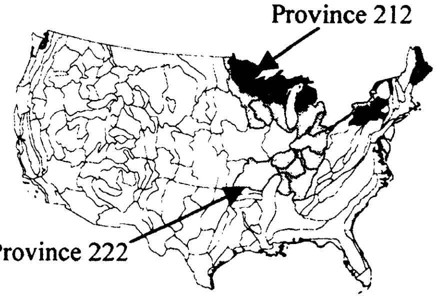

Figure 1.—*Ecoregions of North America (Bailey* et al. *1994).*

A study was conducted to develop individual-tree, distance-independent, diameter growth models using FIA data collected in previous inventories. The criteria for the models, calibrated for major species groups within two ecoregions of the central part of the United States, were that they would produce minimal bias in their estimates and be of a form that can be recalibrated easily with inventory data collected under the annual system.

ECOREGION PROVINCES

The diameter growth models were calibrated for two ecoregions: the 212-Laurentian Mixed Forest Province and the 222-Eastern Broadleaf Forest (Continental) Province defined by Bailey (1995) (fig. 1). The western contiguous portion of the Laurentian Mixed Forest Province on the U.S. side of the U.S./Canadian border falls within the northern half of the Lake States. Province 212 is a subdivision of the Warm Continental Division and is characterized by snowy, cold winters and warm summers. Most precipitation occurs in the summer but is plentiful throughout the year. Province 212 is a transition zone between the boreal forest and the broadleaf deciduous forest zones, and its habitant species include members of both zones.

The Eastern Broadleaf Forest Province is a subdivision of the Hot Continental Division. Most precipitation in Province 222 occurs during the growing season and generally decreases in quantity and adequacy as distance from the Atlantic Ocean increases. This province favors drought-resistant oak-hickory associations. Province 222 lies to the east of the prairie regions, south and west of Province 212 in the northern areas, and west of the Appalachian Mountains in the southern regions. It extends from the Minnesota/Canadian border in the north to Missouri and Tennessee in the south.

DATA

The diameter growth models were calibrated using FIA data from undisturbed plots on land classified as timberland over all ownership categories. Timberland is defined as non-reserved forestland that is producing or is capable of producing 20 cubic feet of industrial wood per year. The FIA survey design and data collection are described by Hansen *et al.* (1992). ArcView GIS was used to overlay Bailey's ecoregion map on the FIA plot locations to select plots within each of the provinces. Growth models for Province 212 were calibrated using FIA data from the following states (the numbers in parentheses refer to the year of the inventory): Michigan (1966, 1980, 1993), Wisconsin (1968, 1983, 1996), and Minnesota (1977, 1990, 1993). Models for Province 222 were calibrated using data from parts of Michigan (1966, 1980, 1993), Wisconsin (1968, 1983, 1996), Minnesota (1977, 1990, 1993), Illinois (1962, 1985, 1998), Indiana (1967, 1986, 1998), Iowa (1974, 1990), Ohio (1978, 1990), Missouri (1972, 1989), Kentucky (1974, 1987), and Tennessee (1989, 1996).

Trees included in the calculation of total plot basal area per acre (BA) and total plot basal area per acre greater than the subject tree (BAL) for time 1 were restricted to all living trees recorded on the plot in the first inventory. Trees included in the calculation of BA and BAL at time 2 were restricted to the remeasured trees that were alive and included in the plot at time 1 and still alive at time 2. Plot variables associated with individual trees were maintained with each tree record. The data were sorted by species and the diameter growth models were fit to the data. Only remeasured trees that were alive during both inventories were used to calibrate the individual-tree diameter growth models. For states with data from three inventories, data were extracted separately for growth periods from the first to the second inventory and from the second to the third inventory, and the two growth periods were treated as separate observations.

The data were split into two databases used to calibrate and validate the models. Every fourth plot was systematically assigned to the validation database. The remaining 75 percent of the plots were used for calibration of the models. Associated plot information was retained with each tree record, and tree records were sorted into species groups.

Diameter at breast height (DBH) is used as the major predictor variable in the growth model for predicting change in DBH. The average observed change in DBH, calculated as the ratio of the difference in DBH measurements at the two inventories and the number of years between inventories, is the dependent variable. Both individual tree and plot variables are considered for use as growth predictors. Individual tree variables include initital crown ratio (CR) and initial crown class (CC). Crown ratio is the percentage of total tree height that is crown and is assigned in FIA data to one of nine categories, where each of the category values 1-8 represents 10 percent interval widths and the final category represents 81-100 percent. Crown class is recorded in FIA data in five categories ranging from nonsuppressed to suppressed.

Variables related to the plot information include BA, BAL, physiographic class (PC), latitude (LAT), and longitude (LNG). Information about competition within the stand is given by BAL and BA, while physiographic class gives information related to site soil and water conditions that affect site productivity. Latitude and longitude are surrogates for climatic conditions in the regional models.

MODELING METHODOLOGY

Mathematical Form of the Diameter Growth Model

The form of the diameter growth models is the product of two components, an average DBH growth model and a modifier. The average model is based on a two-parameter gamma probability density function using DBH as the independent variable to predict diameter growth rates. The modifier is a product of exponentials, each of which incorporates a single additional independent variable. The modifier gives greater accuracy and precision to the growth model by allowing the predicted growth values to increase or decrease from those given by the average model. The form of the diameter growth model is

$$\Delta DBH = AVE(DBH) * MOD(X_1,..., X_7) \quad (1a)$$

where

$$AVG(DBH) = \beta_1 \exp(-\beta_2 DBH) DBH^{\beta_3}, \quad (1b)$$

and

$$MOD(X_1, ..., X_7) = \prod_{i=1}^{7} e^{f_i(X_i)} \quad (1c)$$

The X_is and the f_is are defined in table 1. The functions incorporate the difference of the observed and average ecosystem values of the variables. The greater the deviation of the observed variable's value from the average ecosystem value, the greater the impact that variable has on the model prediction. The values, -90 and 46, in the functions 6 and 7 are the average values of longitude and latitude in Province 212 data. The average longitude and latitude values are -88 and 40, respectively, for Province 222.

The parameterization of the gamma probability density function used in the average component (1b) of the model is a simplified form of that given by Johnson and Kotz (1970). In the formulation, β_2 serves as the scale parameter that defines the spread of the distribution, and the shape parameter, β_3, establishes the peakedness of the curve. The parameter, β_1, is a multiplier to better adjust the model fit to the data either upward or downward in conjunction with the rates of average annual growth. The parameter estimation routine failed to converge for some species. For those species, either $\hat{\beta}_2$ or $\hat{\beta}_3$ was set to 0 in separate calibrations of the model, thus changing the gamma function to a power function or an exponential function, respectively. Of the two functions, the power function generally provided the better fit, based on the mean square error (MSE) of the resulting fits of the models to the data.

In addition to the independent variables used in the diameter growth models (DBH, CR, CC, BA, BAL, PC, LAT, and LNG), a number of other independent variables were explored for inclusion in the diameter growth models by Holdaway (2000). These included the number of trees per acre, average stand diameter, and the ratio of DBH to average stand diameter. Our goal was to build diameter growth models that would explain the most variability while using the smallest number of variables (and associated parameters) to achieve a parsimonious model. No variable was included in the model if the asymptotic 95 percent confidence interval for the parameter estimate associated with the variable included zero. Additionally, a comparison criterion, given by Linhart and Zucchini (1986), accounting for both an estimate of the MSE and the number of estimated parameters was used to determine the variables for inclusion in the models.

Weighted Regression

Heterogeneity in the variation of the residuals was adequately addressed using linear regression to find a functional form relating standard deviation of the residuals to predicted growth as (McRoberts *et al.* 2000):

$$E[\ln(\hat{\sigma})] = \alpha_1 + \alpha_2 \ln(\Delta D\hat{B}H) \quad , (2)$$

where E(.) represents the statistical expectation, $\Delta D\hat{B}H$ is the average predicted annual diameter growth for predicted diameter growth classes, $\hat{\sigma}$ is the standard deviation of the

Table 1.—*The functional forms for variables used in the modifier (the exponential of the sum of these functions) are presented*

Variables	i	Functional form $f_i(X_i)$
Crown ratio	1	$\beta_4 * (CR - 4)$
Plot basal area larger than the subject tree	2	$\exp(\beta_5 * (BAL - 50)) - 1$
Plot basal area	3	$\beta_6 * (BA - 100)$
Crown class	4	$\beta_7 * (\exp(CC/3) - 2.718)$
Physiographic class	5	$\beta_8 * (PC - 5) + \beta_9 * (PC - 5)^2$
Longitude	6	$\beta_{10} * (LNG + 90)$
Latitude	7	$\beta_{11} * (LAT - 46)$

residuals for predicted diameter growth classes, and the αs are parameters to be estimated.

Bias Assessments

The fit of the models was verified using the calibration data set, while the data from the 25 percent of all plots initially set aside were used to validate the models. Verification and validation were carried out by applying the models to the data. Predicted average annnual growth was compared with actual average annual growth over the interval between measurements. The resulting residuals (observed minus predicted average annual growth for the interval between measurements) from the calibration and validation data were analyzed separately.

Annualized residual medians, median ratios, standard deviations, and r^2 values were calculated to examine the prediction bias of the models. Relative bias presents the median bias as a percent of the median observed growth and gives perspective to the importance of bias from a biological viewpoint. Relative bias was calculated as 100 percent times the ratio of the median annualized growth residual to the median annualized observed growth. The r^2 values are calculated as the square of the coefficient of correlation between the annualized predicted and observed values of diameter growth.

RESULTS

Parameter Estimates

Model parameter estimates for 10 of the most frequently occurring tree species in the calibration data sets for Province 212 and Province 222 are given in tables 2 and 3, respectively. Variables in the modifier component of the model not found to be important for a particular species were assigned parameter values of 0, which is analogous to multiplying the model by 1.

Model Verification and Validation

The results of the analysis of the residuals, calculated as the difference of the observed and predicted annual diameter growth values for the observations in the calibration and validation data sets, are given in tables 4 and 5, respectively. Negative values of the median residuals and relative bias indicate overestimation of the models.

CONCLUSIONS

For most species, observed average annual change in DBH ranges from 0.05 to 0.15 inches per year, and DBH is measured to the nearest 0.1 inch. The values of the median residuals for the models applied to the calibration and validation data for Province 212 and Province 222 show the models are relatively unbiased, with median overestimation generally near or less than 0.01 inch per year. The median bias

Table 2.—*Parameter estimates for the annual diameter growth models (1a-c) are given for the 10 most frequently occurring species groups in the 212-Laurentian Mixed Forest Province. Variables to which the parameters, $\hat{\beta}_1$-$\hat{\beta}_{11}$, correspond are given in the column headings.*

Species	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$	$\hat{\beta}_8$	$\hat{\beta}_9$	$\hat{\beta}_{10}$	$\hat{\beta}_{11}$
	DBH	*DBH*	*DBH*	*CR*	*BAL*	*BA*	*CC*	*PC*	*PC²*	*LNG*	*LAT*
Softwoods											
Black spruce	0.0432	0.0321	0.2194	0.1952	0	0	-0.0447	-0.1514	0	0	0
Balsam fir	0.0497	0.0829	0.6517	0.1402	0	-0.0012	-0.0589	-0.1125	0	-0.0354	-0.1074
Tamarack	0.0333	0	0.4205	0.1570	0	0	-0.1069	-0.1262	0	0	0
N. white-cedar	0.0326	0.0202	0.4353	0.1575	0	-0.0005	-0.0667	-0.0895	0	-0.0113	-0.0586
Hardwoods											
N. red oak	0.0724	0	0.2237	0.0349	-0.0029	0	-0.1447	0	0	-0.0246	-0.1090
Hard maple	0.0293	0.0566	0.8377	0.0959	0	-0.0010	-0.1955	0	0	-0.0215	-0.1595
Soft maple	0.0509	0.0167	0.4651	0.1281	0	-0.0014	-0.1274	0	-0.0307	-0.0544	-0.1363
Black ash	0.0535	0	0.2278	0.1323	0	-0.0006	-0.1042	-0.0727	0	-0.0295	-0.1598
Quaking aspen	0.0902	0.0333	0.3930	0.1144	-0.0023	0	-0.1084	0	-0.0280	-0.0261	-0.0940
Paper birch	0.0639	0	0.1284	0.1330	-0.0021	0	0	0	-0.0616	-0.0565	-0.2244

Table 3.—*Parameter estimates for the annual diameter growth models (1a-c) are given for the 10 most frequently occurring species groups in the 222-Eastern Broadleaf Forest (Continental) Province. Variables to which the parameters, $\hat{\beta}_1$-$\hat{\beta}_{11}$, correspond are given in the column headings.*

Species	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_7$	$\hat{\beta}_8$	$\hat{\beta}_{10}$	$\hat{\beta}_{11}$
	DBH	*DBH*	*DBH*	*CR*	*BAL*	*CC*	*PC*	*LNG*	*LAT*
Hardwoods									
Select white oak	0.0658	0.0105	0.3150	0.0372	-0.0031	-0.1959	0.1441	0.0264	-0.0200
Other white oak	0.0813	0	0.1103	0.0575	-0.0025	-0.2107	0.1509	0.0823	0
Select red oak	0.0910	0.0163	0.3374	0.0422	-0.0022	-0.1773	0.1458	0.0242	-0.0087
Other red oak	0.0871	0.0290	0.3832	0.0685	-0.0036	-0.1378	0.1145	0.0485	-0.0082
Select hickory	0.0630	0	0.1656	0.0523	-0.0018	-0.2359	0	0.0197	0
Other hickory	0.0522	0.0253	0.4692	0.0809	-0.0016	-0.2380	0.3244	0.0246	0.0683
Hard maple	0.0643	0.0547	0.5735	0.1021	-0.0035	-0.1963	0.1204	0.0157	0
Soft maple	0.0841	0.0313	0.5007	0.0793	-0.0026	-0.1842	0.1006	0	-0.0365
White and green ash	0.0621	0.0340	0.5834	0.0972	-0.0013	-0.1673	0	0.0282	0
Elm	0.0428	0.0535	0.8250	0.1166	-0.0015	-0.0541	0.1033	0.0522	0.0829

Table 4.—*Residual analysis of the diameter growth models fit to FIA calibration (Cal) and validation (Val) data from the 212-Laurentian Mixed Forest Province*

Species	Observations		Median residual		Standard deviation		Relative bias		r^2	
	Cal	Val	Cal	Val	Cal	Val	Cal	Val	Cal	Val
	Number		*In/yr*		*In/yr*		*Percent*			
Softwoods										
Black spruce	5,921	2,184	-0.006	-0.005	0.036	0.034	-13.2	-13.3	0.297	0.266
Balsam fir	5,916	1,974	-0.008	-0.005	0.051	0.051	-11.0	-6.8	0.383	0.413
Tamarack	2,572	986	-0.011	-0.009	0.050	0.045	-20.0	-15.5	0.294	0.287
Northern white-cedar	7,496	2,497	-0.006	-0.005	0.040	0.041	-10.0	-8.9	0.280	0.288
Hardwoods										
Northern red oak	2,900	991	-0.009	-0.010	0.060	0.055	-7.0	-7.8	0.299	0.388
Hard maple	7,214	2,551	-0.006	-0.005	0.054	0.056	-7.2	-6.0	0.354	0.339
Soft maple	6,487	2,075	-0.008	-0.006	0.061	0.065	-7.3	-6.3	0.328	0.303
Black ash	3,591	1,311	-0.006	-0.003	0.041	0.040	-10.5	-5.7	0.207	0.202
Quaking aspen	12,264	4,033	-0.008	-0.007	0.066	0.066	-5.7	-5.0	0.239	0.228
Paper birch	5,981	1,763	-0.007	-0.006	0.046	0.043	-9.8	-9.3	0.173	0.185

is generally within 15 percent of the median observed growth rate, as given by the relative bias (tables 4 and 5). Some larger percentages are due to the small median observed growth rates in the denominators of the relative bias ratios for slower growing species.

The low values of the squared coefficients of correlation, r^2, indicate that much unexplained variation remains. This is not surprising when one considers that the data, used in both the calibration and validation of the models, were generally collected on uneven-aged, mixed-species stands that range over an entire ecosystem. Future work on these models will investigate incorporation of climatic data such as temperature and precipitation. This information should contribute to better fitting models by explaining more of the uncertainty.

ACKNOWLEDGMENTS

We want to thank Doug Griffith of the Northeastern Research Station and Joseph Glover of the Southern Research Station for kindly providing FIA data from Kentucky and Ohio and from Tennessee, respectively. We would also like to thank Mark Nelson and Dan Wendt for their suggestions and assistance in applying Arc-Info to this research.

Table 5.—*Residual analysis of the diameter growth models fit to FIA calibration (Cal) and validation (Val) data from the 222-Eastern Broadleaf Forest (Continental) Province*

Species	Observations		Median residual		Standard deviation		Relative bias		r^2	
	Cal	Val	Cal	Val	Cal	Val	Cal	Val	Cal	Val
	Number		*In/yr*		*In/yr*		*Percent*			
Hardwoods										
Select white oak	5,726	1,784	-0.007	-0.010	0.061	0.061	-5.7	-8.4	0.326	0.349
Other white oak	1,708	475	-0.005	-0.003	0.052	0.051	-6.3	-3.6	0.291	0.315
Select red oak	2,185	717	-0.013	-0.010	0.082	0.086	-8.7	-6.0	0.215	0.223
Other red oak	1,474	1,628	-0.009	-0.015	0.073	0.074	-6.4	-10.7	0.239	0.258
Select hickory	1,400	396	-0.008	-0.004	0.049	0.058	-10.0	-4.9	0.238	0.266
Other hickory	1,616	535	-0.006	-0.008	0.053	0.054	-7.8	-9.4	0.358	0.362
Hard maple	2,317	764	-0.007	-0.007	0.069	0.067	-8.8	-7.6	0.365	0.333
Soft maple	1,602	466	-0.013	-0.018	0.103	0.110	-8.8	-13.8	0.304	0.260
White and green ash	1,685	451	-0.010	-0.008	0.087	0.082	-6.8	-5.6	0.270	0.289
Elm	1,256	401	-0.011	-0.010	0.088	0.098	-12.2	-10.0	0.372	0.360

LITERATURE CITED

Bailey, R.G. 1995. **Description of the ecoregions of the United States**. 2d ed. Revised and expanded. Misc. Publ. 1391. Washington, DC: U.S. Department of Agriculture, Forest Service. With separate map at 1:7,500,000.

Bailey, R.G.; Avers, P.E.; King, T.; McNabb, W.H., eds. 1994. **Ecoregions and subregions of the United States (map)**. Washington, DC: U.S. Geological Survey. Scale 1:7,500,000, colored. Prepared for the USDA Forest Service.

Brand, G.J.; Nelson, M.D.; Wendt, D.G.; Nimerfro, K.K. 2000. **Technical aspects of the hexagon/panel system**. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C., eds. Proceedings of the 1st annual FIA symposium; 1999 November 2-3; San Antonio, TX. Gen. Tech. Rep. NC-213. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station: 8-13.

Hansen, M.H.; Frieswuk, T.; Glover, J.F.; Kelly, J.F. 1992. **The Eastwide forest inventory data base: users manual**. Gen. Tech. Rep. NC-151. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station.

Holdaway, M.R. 2000. **The AFIS tree growth model for updating annual forest inventories in Minnesota**. In: Hansen, M.H.; Burk, T.E., eds. Integrated tools for natural resources inventories in the 21st century: proceedings of the IUFRO conference; 1998 August 16-20; Boise, ID. Gen. Tech. Rep. NC-212. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station: 507-514.

Johnson, N.L.; Kotz, S. 1970. **Continuous univariate distributions**. New York: Houghton Mifflin.

Linhart, H.; Zucchini, W. 1986. **Model selection**. New York: John Wiley.

McRoberts, R.E.; Holdaway, M.R.; Lessard, V.C. 2000. **Comparing the STEMS and AFIS growth models with respect to the uncertainty of predictions**. In: Hansen, M.H.; Burk, T.E., eds. Integrated tools for natural resources inventories in the 21st century: proceedings of the IUFRO conference; 1998 August 16-20; Boise, ID. Gen. Tech. Rep. NC-212. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station: 539-548.

ABOUT THE AUTHORS

Veronica C. Lessard is a Research Forester and Ronald E. McRoberts and Margaret R. Holdaway are Mathematical Statisticians at the North Central Research Station, USDA Forest Service, St. Paul, MN.