

Comparing basal area growth models, consistency of parameters, and accuracy of prediction

J. J. Colbert^a, Michael Schuckers^b, and Desta Fekedulegn^b

^a USDA-Forest Service, Northeastern Research Station, 180 Canfield St., Morgantown, WV 26505-3101, USA. E-Mail: jcolbert@fs.fed.us

^b Department of Statistics, West Virginia University, P.O. Box 6330, Morgantown, WV 26506-6330, USA.

Abstract: We fit alternative sigmoid growth models to sample tree basal area historical data derived from increment cores and disks taken at breast height. We examine and compare the estimated parameters for these models across a range of sample sites. Models are rated on consistency of parameters and on their ability to fit growth data from four sites that are located across a longitudinal gradient across the states of Delaware, Pennsylvania, West Virginia, and Ohio in the USA. We then examine the differences among these models in their abilities to predict short-term (decade) and longer-term growth of trees. Accuracy and potential effects of bias are discussed relative to the age and source locations of sample trees used in this study.

Keywords: Individual tree growth; Parameter estimation; Prediction; Nonlinear Modelling

1. INTRODUCTION

Analyses of four forest tree basal-area growth models were carried out using data from the states of Delaware (DE), Pennsylvania (PA), West Virginia (WV), and Ohio (OH) in the Mid-Atlantic region of the United States.

Forest conditions in the mid-Atlantic Region of the United States are quite variable, with continental climate and annual rainfall from 50 to 180+ cm but generally in the range of 100-130 cm in the areas where managed forests are most prevalent. Rainfall is generally well distributed throughout the year. Forests we studied are considered mixed mesophytic and consisted of oak-dominated, mixed oak-hickory or oak-maple forest types. These forest stands are usually quite diverse in canopy composition and contain from five to twenty or more different tree species. Elevations range from the coastal plains of DE, just 2-10 meters above sea level to the forests of the more interior central PA and WV mountains that range from 350-900 m above sea level. Our study sites range from 400 to 600 m. In OH, stands were located on the Dorr Run Management Unit of the Athens District of the Wayne National Forest, which is 210-320 m elevation.

Growth models for this region use forest inventory sampling data to estimate stocking for initial conditions. They vary in their structure but are all distance independent individual-tree or stand-table simulators. These models have been tested and compared [Schuler et al. 1993] and have been shown to perform adequately but are most accurate on particular forest types. The Stand-Damage Model [Colbert and Racine 2001, Colbert et al. 1997, Colbert and Sheehan 1995] is a more generally applicable, forest gap type model that permits calibration to forest type and growth conditions for each species. Here we examine new forms of basal area growth models that could be used to form the base model for individual tree growth.

We fit nonlinear sigmoid growth models described in detail by Fekedulegn et al. [1999]. The 4-parameter models we chose to utilize in this study have been shown to be adequate for modeling sigmoid growth with sufficient flexibility and good statistical properties [Draper and Smith 1981, Schnute 1981, Myers 1986, Vanclay 1994]. These models were fit to basal area data derived from radial increment data and estimates of inner radius and age.

2. METHODS

Source Data: A total of 190 radial growth-increment samples were used. The data were obtained from southern red oak (*Quercus falcata*) and white oak (*Quercus alba*) from the coastal plain in central and southern DE; northern red oak (*Quercus rubra*) from the ridge-and-valley area of central PA; northern red oak on the Coopers Rock State Forest in north central WV; and northern red oak from OH. Table 1 provides the numbers of sample data and type. All samples were taken at breast height (137 cm). Increment cores were taken with 4.3 mm borers and disk samples were taken from felled trees. We took samples aiming through the tree center and as close to the pith as possible.

Table 1. Sample counts by location, type and tree species.

Location	Type	Species	Number
DE	IC ^a	Q. f.	24
	IC	Q. a.	32
PA	IC	Q. r.	26
WVD	Disk	Q. r.	21
Log Grade (LG)	IC	Q. r.	60
OH	IC	Q. r.	28

^a: IC = increment core sample

Increment cores were first dried and glued in place with water-soluble glue on top of wood mounts of approximately 18 mm high by 8 mm wide cross-section. The mount top is beveled so that it contains a groove 4 mm across that runs the length of the mount. Samples were oriented vertically and sanded using fine (400-1200) grit to expose cell structure. Annual radial increments were measured to the nearest 0.001 mm on a measuring stage and the radius and age at the inner edge of the innermost ring were estimated using a 1-mm scaled circular ruler, taking into consideration the curvature of the earliest growth rings and the width of those same rings. In some instances, on increment core samples and on all disk samples we were able to provide data from the pith at age 0. Basal area (inside bark) series were then produced assuming circular cross-sections at breast height.

The nonlinear models fit were the following:

1. Richards

$$\omega(t) = \beta_0 / (1 + \beta_1 \exp(-\beta_2 t))^{1/\beta_3} + \varepsilon$$

2. Weibull

$$\omega(t) = \beta_0 - \beta_1 \exp(-\beta_2 t^{\beta_3}) + \varepsilon$$

3. Chapman-Richards

$$\omega(t) = \beta_0 (1 - \beta_1 \exp(-\beta_2 t))^{1/(1-\beta_3)} + \varepsilon$$

4. von Bertalanffy

$$\omega(t) = (\beta_0^{1-\beta_3} - \beta_1 \exp(-\beta_2 t))^{1/(1-\beta_3)} + \varepsilon$$

proposed by 1. Richards [1959], 2. Ratkowsky [1983], 3. Turnbull [1963], Pienaar and Turnbull [1973], 4. von Bertalanffy [1957]. It should be noted that all parameters are assumed to be positive ($\beta_0, \beta_1, \beta_2, \beta_3 > 0$) for all models; in 2. $\beta_1 < \beta_0$; in 3. $\beta_1 < 1$; and in both 3. and 4., $\beta_3 < 1$. For the tree species used, growing in North America, biological bounds were constructed. Maximum basal area for the red oaks are less than 2 m² and less than 5 m² for *Q. alba*. We set an upper bound for β_0 to 8 m² and lower bound to 0.1 m² for all runs of each model. We use the NLIN Procedure [SAS 1999] with the Marquardt [1963] method to estimate the parameters for each model and basal area series, supplying the partial derivatives [Fekedulegn et al. 1999]. We found convergence to be quite sensitive to the starting values used. We used modal initial conditions, estimated from the data as described in Fekedulegn et al. [1999]. From the fitted values we created an initial search grid to then restart the produce (Table 2).

For each series, we fit each of the models described above. To assess the predictive power of each model, we then truncated the data set and refit the model to the truncated dataset, allowing us to examine differences and produce a test of the models to project the data forward for prediction of future basal area. Statistical summaries are provided for parameters for each model. We tested the differences among the models' ability to fit the data by examining the number of samples where the convergence criteria were met and we calculated the mean square error (MSE) for each fit and tested difference between models. We explored parameter interactions and consistency of fits among sites and between species. We also examined the interactions among parameters within models and the values of β_0 among models since this parameter is the asymptote for basal area as the tree reaches maturity.

Table 2. Initial search grid (L – lower bound; U – upper bound; S – step size).

Model		β_0	β_1	β_2	β_3
R	L	0.101	10^{-4}	10^{-3}	$5 \cdot 10^{-5}$
	U	5.0	0.1	0.3	2.0
	S	0.6	10^{-2}	0.03	0.2
W	L	0.2	10^{-2}	10^{-4}	0.1
	U	5.0	4.9	10^{-3}	4.5
	S	0.6	0.59	$2 \cdot 10^{-4}$	0.5
C-R	L	1.01	10^{-4}	10^{-3}	10^{-2}
	U	5.0	0.99998	0.1	0.99
	S	0.6	0.15	0.015	0.15
vB	L	0.2	0.36	0.002	0.36
	U	5.0	1.04	0.022	0.80
	S	0.4	0.05	0.003	0.03

3. RESULTS

3.1 Convergence

To judge the quality of models and fitting procedures, we looked at the number of samples where the convergence criteria were met. Of the 380 total series, the Chapman-Richards model consistently was fit most often, followed by the Richards, Weibull, and finally the von Bertalanffy models (Table 3).

Table 3. Adequacy of Marquardt fit to full and reduced series (counts in parentheses).

Model	Full	Reduced	Total
	Percent		
R	83 (158)	72 (136)	77 (294)
W	71 (135)	75 (142)	73 (277)
C-R	96 (182)	94 (178)	95 (360)
vB	63 (120)	70 (133)	67 (253)

For all models, the fitting procedure converged for the majority of these data and there was no significant difference among the numbers of series fit for each model. It can also be seen that the length of the data series did not have a consistent effect on convergence. There was no data set that was significantly different in the number of series converging (Table 4). Among datasets, the Chapman-Richards model consistently converged most often. It should be noted that we first attempted to use less dense initial grids and found that the procedure did not converge for a very large number of samples, across all models.

Table 4. Adequacy of Marquardt fit among data sets.

Model	802	812	LG	OH	PA	WV
	Percent					
R	61	85	82	85	73	74
W	87	29	85	57	81	76
CR	97	98	95	100	88	88
vB	62	81	64	85	62	45

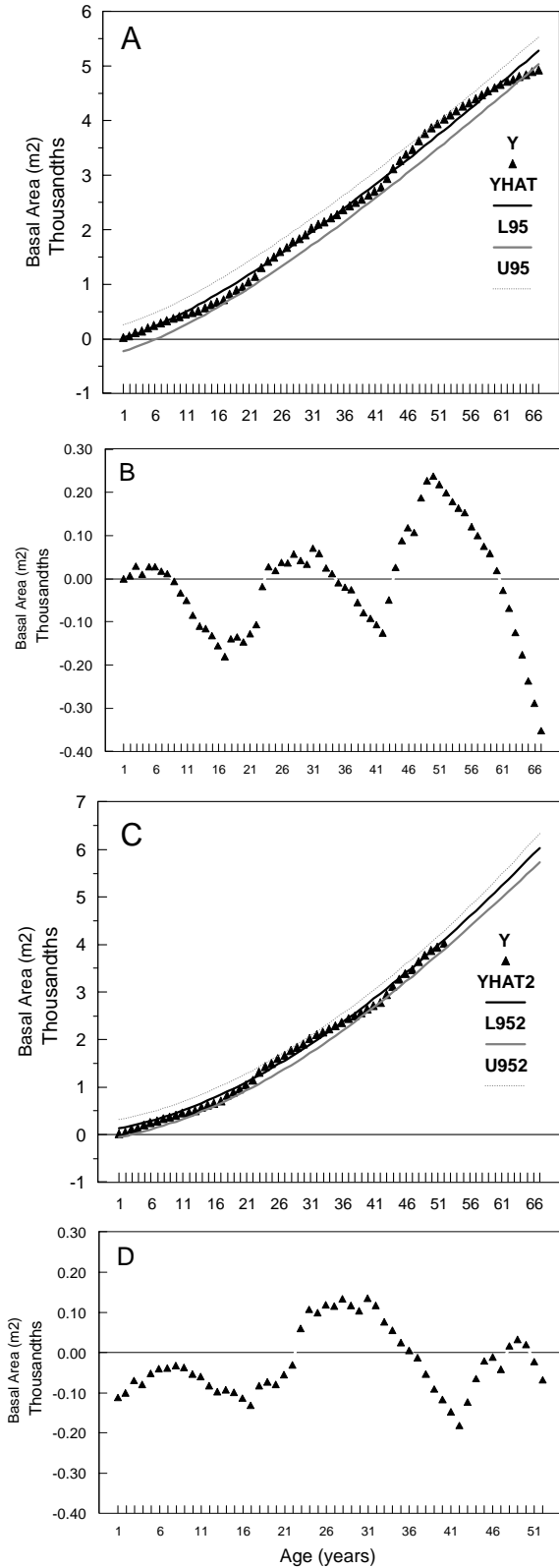
Sites: 802 = DE, *Q. alba*; 812 = DE, *Q. falcata*; LG = WV-LG; WV = WV Disks.

We examined MSE for each model, series length, and source data set and found that there was no significant difference among these classifications. We found the choice of limits and mesh size to have a large effect on the number of converging series and the quality of the fit (MSE). We found good quality fits ($MSE < 10^{-4}$) in 87.5% of the non-converging samples. It should be noted that since the procedure used did not permit the inclusion of non-constant constraints (boundary conditions) among parameters, there was some data truncation for early growth for some samples when fitting the von Bertalanffy model. It is sufficient that $\beta_0^{1-\beta_3} > \beta_1$ for $\hat{\omega}$ to be positive for all t . When this condition is not met, $\hat{\omega}$ will remain undefined for small t , and the associated data are ignored during the fitting procedure.

3.2 Model Fits: Graphs and Residual Plots

While improvement can be made in obtaining convergence, results obtained on samples that did not converge appear often to be adequate and represent the data well throughout the range of those data. The errors about the non-converged fitted curves are often no worse or even better than another sample taken at another radius from the same tree where convergence criteria were met. The ability to find an adequate fit does not appear to be associated with either the length or starting point of the data series. Figure 1 shows the fitted summaries for such a sample. The truncated series did not converge. Convergence was obtained for both series from a second increment core sample taken from the same tree. It was found that problems with convergence did not appear to be associated consistently with either the full or truncated series. Influences like individual tree release, weather, or insect defoliation can cause fluctuations in the growth pattern that are not well represented by these models.

Figure 1. Graph of basal area data and Chapman-Richards model with 95 percent confidence intervals: (A) converged to the full dataset; (B) residuals about \hat{Y} from A.; (C) non-convergence to the truncated dataset; (D) residuals for C. Note that the residuals are clustered closer to the curve in the truncated fit.

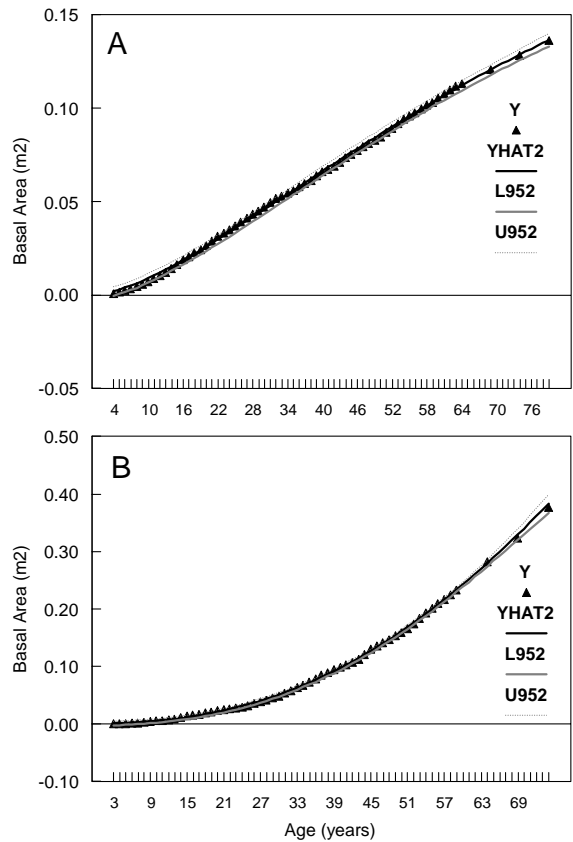


3.3 Predictions

We examined 5-, 10- and 15-year predictions. Figure 2 shows two trees where the truncated series demonstrates that the prediction from the model can adequately be used for predicting the trend of diameter growth. Table 5 shows the values for each of the predictions as well as the actual data.

Figure 2. The data, fitted Chapman-Richards model, and 95% confidence intervals (A) using the first 59 years of a PA sample and (B) using the first 64 years of an OH sample (see Table 5).

3.4 Model Asymptotes



We examined the parameters for each model. β_0 was the only parameter that was comparable among models. We found that the truncated series ($p = 0.0104$) and the von Bertalanffy model ($p < 0.001$) had significantly higher predicted asymptotes. The asymptotes were the lowest and did not differ significantly between the Richards and Weibull models while for these two, β_0 did differ from the other two models. There were a number of samples for which the predicted asymptote was well above the biological maximum for the tree species. Parameter median, Q1, and Q3 values are provided for each model in Section 3.5.

Table 5. Actual and predicted basal area for two example tree: Y = measured basal area; \hat{Y} = predicted basal area, full data series; \hat{Y}_t = predicted basal area, built from truncated data series.

PI ^a	0	5	10	15
E118234A				
Y	0.1133	0.1207	0.1283	0.1364
\hat{Y}	0.1120	0.1204	0.1284	0.1360
\hat{Y}_t	0.1121	0.1206	0.1287	0.1364
T03B				
Y	0.233	0.2819	0.323	0.3764
\hat{Y}	0.2344	0.2789	0.3258	0.3747
\hat{Y}_t	0.234	0.2796	0.3295	0.3832

^a: Prediction interval length in years.

3.5 Other Parameters

Parameters for each model were examined both graphically and statistically. Here we present what we think are the most relevant observations from that review. The parameter statistics are for the samples where the convergence criteria were met. Q1 and Q3 are the 25% and 75% quantiles.

	Richards		
	Q3	Median	Q1
β_0	0.270	0.143	0.100
β_1	1.61 –3 ^a	0.21 –3	0.11 –3
β_2	0.0428	0.0280	0.0196
β_3	0.22 –3	0.04 –3	0.02 –3

^a: Here as elsewhere in this paper, this notation indicates a decimal shift (1.61×10^{-3}).

For this model, β_0 was the lowest across all fits; the 99th percentile estimate was 3.89 m²; β_1 showed the widest range of any parameter across all models; β_2 had the narrowest range when compared to the other models' t coefficients; β_3 had the second widest range across all models (over 5 orders of magnitude) but for most of the fits, estimates converged to values less than 10^{-3} .

	Weibull		
	Q3	Median	Q1
β_0	0.360	0.152	0.100

β_1	0.360	0.152	0.100
β_2	3.49 –4	0.61 –4	0.10 –4
β_3	2.33	1.98	1.66

As we improved the convergence across models and samples, we found just one striking relationship among parameters. An almost perfect linear relationship existed between β_0 and β_1 in this model; β_0 was above the estimates for the Richards model but it still tended to under-predict expected maximum diameter for these data. The time coefficient, β_2 , showed the widest variation across models but because t in this model form carries an exponential factor, β_3 , the variation in these two parameters compensated one another. These two parameters showed a strong log-linear relationship. Another interesting note is that while the convergence criteria were met, one sample showed no fit; \hat{Y} was essentially constant.

Chapman-Richards

	Q3	Median	Q1
β_0	2.79	0.255	0.118
β_1	0.999	0.987	0.228
β_2	0.0245	0.0143	0.0055
β_3	0.966	0.689	0.509

Under the Chapman-Richards formulation, β_0 showed the most biologically reasonable range of values, but it did considerably over-estimate on more than 10% of the samples. β_1 ranged over four orders of magnitude but did not show any noticeable interaction with other parameters. Except for the Weibull form, β_3 showed the narrowest range of variability, with the exception of one sample.

No consistent pattern emerged between species or among sites in the parameterization of these models except that β_3 in the Weibull Model shows some slight location dependence. These differences are mitigated by the fact that the range in differences are similar to what is obtained within West Virginia data between those samples taken from whole tree dissections and those taken from increment cores.

4. DISCUSSION

Each of these models will provide reasonable fit to radial increment data and permit estimates of

future basal area under nominal conditions. There is consistency between models in terms of MSE. Research to date suggests that MSE would be a better screening criterion than meeting SAS's Proc NLIN convergence criteria for retaining results. We attempted to use other fitting methods but found that the Marquardt method performed best.

To strengthen our understanding of the power and consistency of these models to perform across this region, we plan to expand the data to include balance among species and to classify our analyses to account for canopy strata and site factor effects within species. We will fit data from older trees to ascertain how tree age affects parameters, particularly the asymptote. We will explore the ranges for parameters of these models that give rise to realistic trends for mature and over-mature trees.

As mentioned earlier, convergence is highly dependent on starting values. We found that the use of a starting grid will usually provide good results. When boundary conditions are considered, convergence and the quality of final values can be further assured. Some care must be taken to deal with nonlinear boundary conditions that may not be used under some procedures.

We plan to compare these models to growth models used in forest management in this region and to the diameter growth model used as the basis for predicting diameter increment in forest gap simulators.

5. ACKNOWLEDGEMENTS

We thank Michael A. Valenti, Delaware State Forest Service, and Phil Perry, Wayne National Forest, USDA Forest Service for contributing data. We especially thank Sandra Fosbroke for assistance in editing and graphic development. And finally, we thank our field and lab crews who laboured to collect and prepare the samples.

6. REFERENCES

Bertalanffy, L. von. Quantitative laws in metabolism and growth. *Quantitative Rev. Biology* 32: 218-231, 1957.

Colbert, J.J., and G. Racin. How to use the Stand-Damage Model: Version 2.0 (Computer program). Gen. Tech. Rep. NE-281. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 79 p., 2001.

Colbert, J.J., P. Perry, B. Onken. Preparing for the gypsy moth – design and analysis of stand management: Dorr Run, Wayne National

Forest. In: Communicating the role of silviculture in managing the national forests: proceedings of the national silviculture workshop; 1997 May 19-22; Warren, PA. Gen. Tech. Rep. NE-238. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station: 76-84, 1997.

Colbert, J.J., K.A. Sheehan. Description of the Stand-Damage Model: part of the gypsy moth life system model. Gen. Tech. Rep. NE-208. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 111 p., 1995.

Draper, N.R. and H. Smith. *Applied regression analysis*. 2nd edition. John Wiley & Sons Inc. New York. 709 p., 1981.

Fekedulegn, D., M.P. Mac Siurtain, and J.J. Colbert. Parameter estimation of nonlinear growth models in forestry. *Silva Fennica* 33(4): 327-336, 1999.

Marquardt, D.W. An algorithm for least squares estimation of nonlinear parameters. *Journal of Society Industrial and Applied Mathematics* 11: 431-441, 1963.

Myers, R.H. *Classical and modern regression with applications*. Duxbury Press, Boston. 359 p.

Pienaar, L.V. and K.J. Turnbull. The Chapman-Richards generalization of von Bertalanffy's growth model for basal area growth and yield in even-aged stands. *Forest Science* 19(1): 2-22, 1973.

Ratkowsky, D.A. *Nonlinear regression modelling*. Marcel Dekker, New York, 276 p., 1983.

Richards, F.J. A flexible growth function for empirical use. *Journal of Experimental Botany* 10:290-300, 1959.

SAS. 2000. *SAS Online Doc Version 8*, February 2000, SAS Institute Inc., Cary, NC, USA, 1986.

Schnute, J. A versital growth model with statistically stable parameters. *Canadian Journal of Fishery and Aquatic Sciences* 38: 1128-1140, 1981.

Schuler, T.M., D.A. Marquis, R.L. Ernst, B.T. Simpson. Test of four stand growth simulators for the northeastern United States. Res. Pap. NE-676. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 14 p., 1993.

Turnbull, K.J. Population dynamics in mixed forest stands: a system of mathematical models of mixed growth and structure. Ph.D. Dissertation, University of Washington, 1963.

Vancly, J. K. *Modelling forest growth and yield*. CAB International, Wallingford, UK. 380 p., 1994.