LONGLEAF PINE PLANTATIONS: GROWTH AND YIELD MODELING IN AN ECOSYSTEM RESTORATION CONTEXT

J.C.G. Goelz

Southern Research Station, USDA Forest Service,
Alexandria Forestry Center, 2500 Shreveport Hwy, Pineville, LA 71360-5500.
Phone: (318) 473-7227. Fax: (318)473-7273.
email: jcgoelz@fs.fed.us

Abstract

Restoration of longleaf pine within its historical range is actively conducted by private individuals and public agencies due to the inherent beauty of the ecosystem and the suitability as habitat for red cockaded woodpeckers and other wildlife. Managers of land restored to longleaf pine desire models that will allow long-term projections to facilitate management decisions. Managers of restored ecosystems typically desire to predict the dynamics of more than just the trees; understory vegetation, wildlife, fuels, and the effects of fire are pertinent to managers of longleaf pine. Modeling a rare ecosystem is hindered due to inadequate data covering a range of conditions and ages. Typically, plantations do not exist that approach the planned rotation age. Thus, any forest model will typically be greatly extrapolated. New diagnostics are described that suggest suitability of a model for extrapolation. These diagnostics may also be used as general goodness of fit diagnostics. If data from old plantations are lacking, older natural stands may be used to supplement the data.

Introduction

In pre-colonial times, longleaf pine (*Pinus palustris* Mill) stands were a major component of the southern coastal plain of the U.S.A. from North Carolina to Texas. As little as 1/30th of the acreage remains Franklin 1997). Restoration of longleaf pine within its former range is advocated by a private association and by governmental agencies, and an aggressive planting program has developed. Restoration of the longleaf pine forest type is desired due to inherent beauty of the mature forest as well as value for wildlife habitat, including the endangered red cockaded woodpecker (*Picoides borealis*). Although most longleaf pine plantations are established for ecosystem restoration, knowledge about the growth and development is essential for sound management by both governmental agencies and private individuals. Some land managers desire to employ lower-intensity management, particularly longer rotation ages, for which longleaf pine is well-suited. Longleaf pine is also less susceptible to most insect and disease problems than other southern pines (Boyer 1990). Longleaf pine stands subject to periodic prescribed burns are also extremely resistant to stand-replacing fires. Longleaf pine is also desirable because it produces higher-value products such as poles or pilings more frequently than similarly-sized loblolly pine (*Pinus taeda* L.), the most abundant pine species in the region.

Little growth and yield modeling has been accomplished with longleaf pine plantations. The only existing model is restricted to unthinned stands (Lohrey and Bailey 1977). Goelz and Leduc [in press, a] provide an empirical yield table for unthinned plantations, using the same dataset as Lohrey and Bailey (1977), supplemented with additional measurements for most plots, as well as additional plots arising from other silvicultural experiments. Naturally-regenerated longleaf stands have been recently (Somers and Farrar 1991).

Idiosyncrasies of Longleaf Pine

Longleaf pine has silvical characteristics that distinguish it from other pines in the U.S. (Boyer 1990) Three characteristics affect the types of diameter distributions possessed by longleaf pine, and hence influence the structure of an appropriate model. First, the "grass stage" is a period of up to several years where the seedling does not have appreciable height growth, the terminal bud remains at or near groundlevel, and the long needles resemble a bunchgrass. When recommended management practices are employed, the grass stage for most seedlings is only a year or two, although a few seedlings may linger for several years. Second, although longleaf pine is an intolerant species, saplings and poles can often persist in an intermediate or suppressed crown class for long periods. This is atypical for southern pines. Trees with live crown ratios of 30 percent or more can respond to release after years in an intermediate crown class (Boyer 1990). Third, prescribed fire is typically applied to longleaf pine stands every 2 to 5 years; prescribed fire is not standard operating procedure for other southern forests. Although mortality from prescribed fire has relatively low probability, it will occur throughout the life of the stand. Prescribed fire will also restrict ingrowth of volunteer hardwoods and loblolly pine. These characteristics provide for generally irregular, possibly multimodal, diameter distributions with long, or heavy, left-hand tails; this has motivated us to use an atypical model structure Goelz and Leduc [in press, b]. It represents an intermediate between classical model forms (Goelz [in press]), and can be viewed as a nonparametric diameter distribution model, a variant of a diameter class model (sensu, Cao and Baldwin 1999, Nepal and Somers 1992, Clutter and Jones 1980), or a variant individual tree model.

Modeling in an Ecosystem Restoration Context

The aim of this paper is to describe the development of a growth and yield model suitable in an ecosystem restoration context, using longleaf pine as an example ecosystem. Rather than identifying novel concerns in building models of forests, an ecosystem restoration context merely elevates some modeling concerns to a higher level of importance.

Availability of Data

A need for restoration presupposes a perceived deficiency in the current extent of the ecosystem. If the ecosystem is rare, a dataset that is suitable for estimating a model is unlikely. When a dataset is available, it will likely be deficient with regard to: (1) representing the range of sites typical for restoration activities across all ages from plantation establishment to final rotation, and (2) representing the current suite of silvicultural practices applied in plantation establishment and subsequent management. Our longleaf pine data largely represents plantations established with bare-root seedlings on land that was indiscriminantly burned and grazed for several years after it was clearcut. Also, the oldest plantations in our dataset were measured at age 65. This is only half of the expected rotation age of 100-150 years. Thus, the model will be greatly extrapolated with respect to age. Current restoration efforts are concentrated on converting existing stands of loblolly pine, or mixed pine-hardwood stands, or recent agricultural fields, to longleaf pine. Most contemporary longleaf pine stands are established with container stock. Thus, any model we develop will represent an extrapolation to conditions that are not represented in the dataset. We are establishing new plots that will allow us to check the comparability to our older data to current practices. This will allow future revision of the model. Our other allowance will be for the model user to specify initial survival and growth based upon their current expectations. It is the initial survival and growth that is most subject to the contemporary silvicultural practices. However, it will still be a matter of faith to expect that subsequent growth and survival will follow the trends existing in our older data.

There is More to a Forest than the Trees

If ecosystems are the subject of interest, rather than simply a stand of trees, then at least some additional variables should be projected to be relevant to the forest manager. The focus of this paper is on "growth and yield" models, thus the scope is constrained to models where projection of stand characteristics is central, however additional variables will increase utility to a manager engaging in ecosystem restoration. The critical supplemental variables will be determined largely by the purpose for which the ecosystem is being restored.

In the example of longleaf pine, a primary purpose of restoration is on creating an understory community that is perceived to be desirable. The desirable community is dominated by a ground vegetation of grass, either wiregrass or bluestem grasses dependent on location (wiregrasses dominate in the east, while bluestem grasses dominate west of the Mississippi River). Thus, it would be helpful if a growth and yield model also described the groundcover community in some way. Groundcover biomass can be predicted well based upon tree basal area and time since last prescribed fire. Species richness of herbaceous vegetation is high in longleaf pine stands, and managers desire to maintain high species richness. Modeling the dynamics of species richness may be more difficult than simply modeling biomass.

A manager of longleaf pine will typically use prescribed fire to create and maintain the desirable understory community. Thus, a growth and yield model that incorporates fuel loadings of dead and living biomass will be much more useful than a model that does not. Existing fire behavior models may be invoked within the growth and yield model to allow simulating the effect of stand dynamics on prescribed fire behavior. It is also critical to include the effects of prescribed fire on fuel loadings and tree survival and growth.

Longleaf pine stands are desirable as breeding and foraging habitat for red cockaded woodpecker. Thus, stand characteristics that are pertinent to managing for the woodpecker are important. Large, old trees are known to be preferred for nests of the woodpecker. A restoration-relevant growth and yield model should satisfactorily predict the existence and survival of potential den trees. A valuable addition would be prediction of the survival of existing den trees. Although landscape conditions will determine habitat suitability for red cockaded woodpecker, a growth and yield model is unsatisfactory for directly considering landscape-level influences on the bird. However, the individual stand-level outputs from the growth and yield model predictions could be integrated into a landscape-level context.

Potentially, all variables of interest will be measured on the same plots as the growth and yield dataset. If this is so, appropriate methods for parameter estimation of systems of equations may be employed, such as three-stage least squares or possibly seemingly unrelated regression (Kmenta 1986). However, the data for non-tree components will likely arise from a different source than the growth and yield data set, and estimation of non-tree equations will likely be done independently of the system of equations used to describe stand growth. In some cases, preexisting models, such as wildlife habitat suitability models, may be incorporated into the system.

When the system of equations involves a subset of tree-related variables (possibly estimated using methods that consider the structure of the system (i.e. 3SLS)), with a number of accessory equations, (possibly each arising from a unique dataset), inferences about the accessory variables become difficult. If the datasets arise from experiments, rather than from a sample from the population of interest, the parameters are not necessarily consistent, and the variability of the predictions is unknown when applied to the larger population. Thus, while it may be illustrative to display long-term dynamics of one of the accessory variables, it would be less appropriate to test for differences with respect to simulated management practices on one of these variables (for example, invertebrate community species richness). Furthermore, some stand-level characteristics that are relevant to a wildlife species (such as number of trees greater than 75 cm dbh) will not be directly predicted by the growth and yield model. These contrived variables may be determined from the diameter distribution, but the variability of their prediction may not be known.

Long-term Projections

For most models designed for forest plantations, a module describing the development of natural regeneration is unnecessary. However, in an ecosystem restoration context, it is often desired to proceed to a naturally-regenerated stand following the first rotation of the plantation. There are three main ways that this can be satisfied. First, the plantation model could have a regeneration module added. At the time of the regeneration cut, the plantation model output could be used as starting values for a natural-stand growth and yield model. The second alternative would be to switch from a plantation model to a natural stand model late in the rotation, but before the regeneration cut. This would only be justified if the data failed to reject the hypothesis that the natural stand model was a satisfactory predictor for older plantations. The third alternative would be to have an integrated model that was applicable for both plantations and natural stands; some parameters could distinguish between plantations and natural stands while others would be common for both types of stands.

As previously mentioned, it is likely that rotation age for a restored forest might be much greater than the oldest plantations for that species. A land manager will likely greatly extrapolate beyond the range of the data. Any inferences from such extrapolation are suspect. However, we describe some diagnostics that can identify a model's suitability for extrapolation.

New Diagnostics Particularly Pertinent to Extrapolability With Respect to Age

Examples given in this section are purely descriptive. Although real data is used for much of the section, and models are compared, the emphasis is on describing the diagnostic, not evaluating the models. The data come from a long-term database on longleaf pine described by Goelz and Leduc [in press, a]. Specifically, we have chosen the unthinned plots with a site index of 16 to 17.5m at 25 years.

Diagnostic One, Stability of Estimates after Sequentially Deleting the Oldest-Aged Data.

The algorithm is simple: (1) estimate parameters using all data; (2) estimate parameters after eliminating the oldest 5-year age class; (3) estimate parameters after eliminating the next-oldest 5-year age class; (4) repeat (3) until data become deficient for estimating the parameters; (5) plot parameter estimates related to age of oldest data; (6) plot error criteria of choice vs. age of oldest data, including predictions of data reserved from estimation in these error criteria (i.e. error criteria are based on all data); (7) plot predicted values at the anticipated maximum age to which the model will be applied. The graphics described in (5), (6), and (7) can be used to compare the extrapolability of competing equations.

Two equations were compared for stand basal area, a Richard's function:

$$\mathrm{Y} = b_1\left(1 - e^{b_2 \mathrm{X}}\right)^{b_3} \qquad [1]$$

and a Schumacher equation:

$$\ln(\mathrm{Y}) = b_1 + b_2\left(\frac{1}{\mathrm{X}}\right) \qquad [2]$$

where Y represents basal area (m^2/ha) and X represents age, and the b_i are parameters.

Step (5) of the algorithm is displayed in figure 1 for equation [1] and figure 2 for equation [2]. The parameters for [1] stabilize somewhat more rapidly than do the parameters for the [2], and thus we expect[1] would be more suited for extrapolation.

Fig. 1. *Parameter estimates of equation [1], related to the maximum age of the data used in fitting basal area to age.*

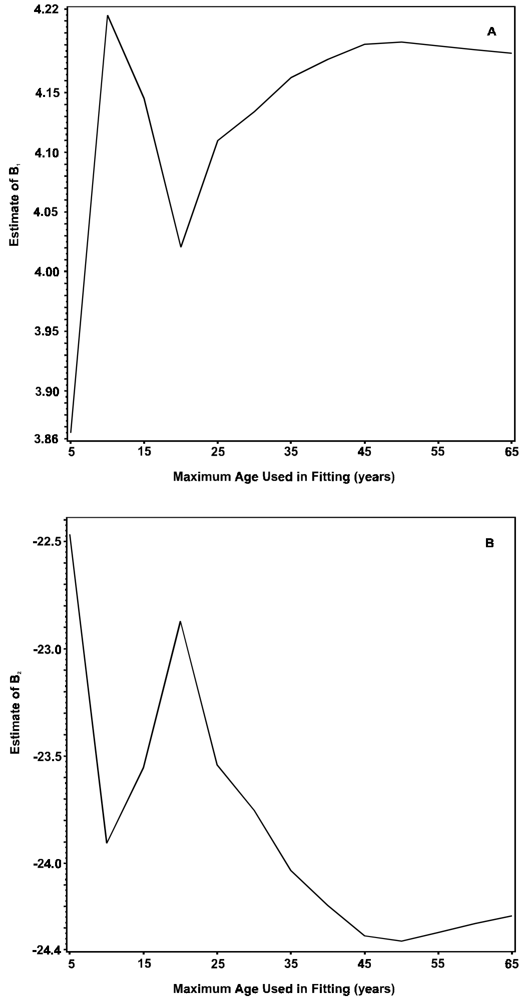

Fig. 2. *Parameter estimates of equation [2], related to the maximum age of the data used in fitting basal area to age.*

Plotting the error criteria, step (6), is done in figure 3 for bias and figure 4 for mean squared error. Equation [1] is unreliable until age 25 data are included in the estimation. However, after this point, it is the slightly better predictor. Based solely on the relative stability of the error criteria for both equations, they may both be reasonable candidates for extrapolation.

Step (7) of the algorithm is displayed in figure 5. Predicted basal area at age 120 is unreasonable for equation [1] when only data less than 15 years is available, but is reasonable when more data is used. For both equations, predicted basal area at age 120 is

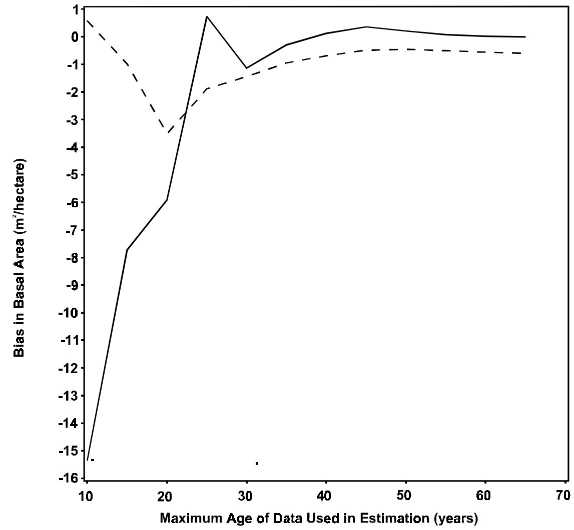

Fig. 3. *Bias in predicting basal area, related to the maximum age of the data used in fitting. Solid line represents equation [1] and dashed line represents equation [2].*

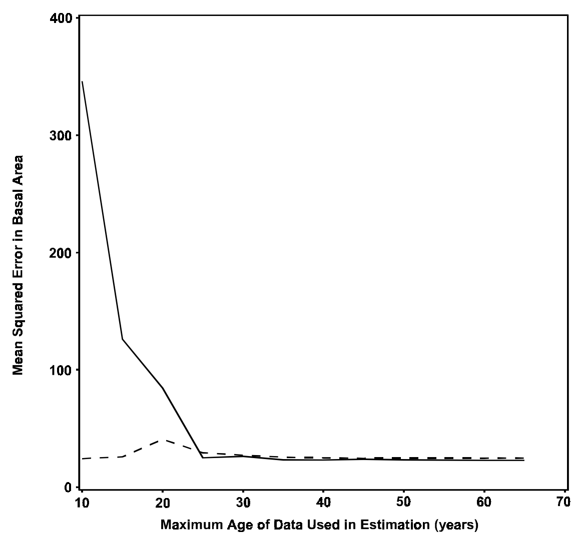

Fig. 4 *Mean squared error in predicting basal area, related to the maximum age of the data used in fitting. Solid line represents equation [1] and dashed line represents equation[2].*

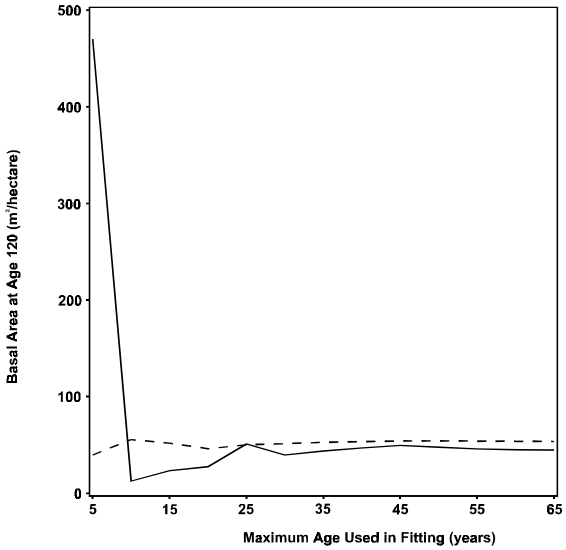

Fig. 5. *Predicted basal area at age 120, related to the maximum age of the data used in fitting. Solid line represents equation [1] and dashed line represents equation [2].*

*v*ery stable when age 30 and older data were used. Considering steps (5) through (7), equation [1] appears to be somewhat better for extrapolation, and is a slightly better predictor as well.

Diagnostic One, for Quadratic Mean Diameter.

This example differs from the previous by considering quadratic mean diameter, rather than basal area. For equation [1], two of the parameters stabilized after age 40 data was included (figure 6b,c), however the asymptote parameter did not stabilize at all (figure 6a). For equation [2], neither parameter stabilized (figure 7). These results suggest that neither is appropriate for extrapolation. Given the strong trends, probably neither equation well-represents the true function.

When only the youngest ages were used, equation [1] was inferior to equation [2] for both bias (figure 8) and mean squared error

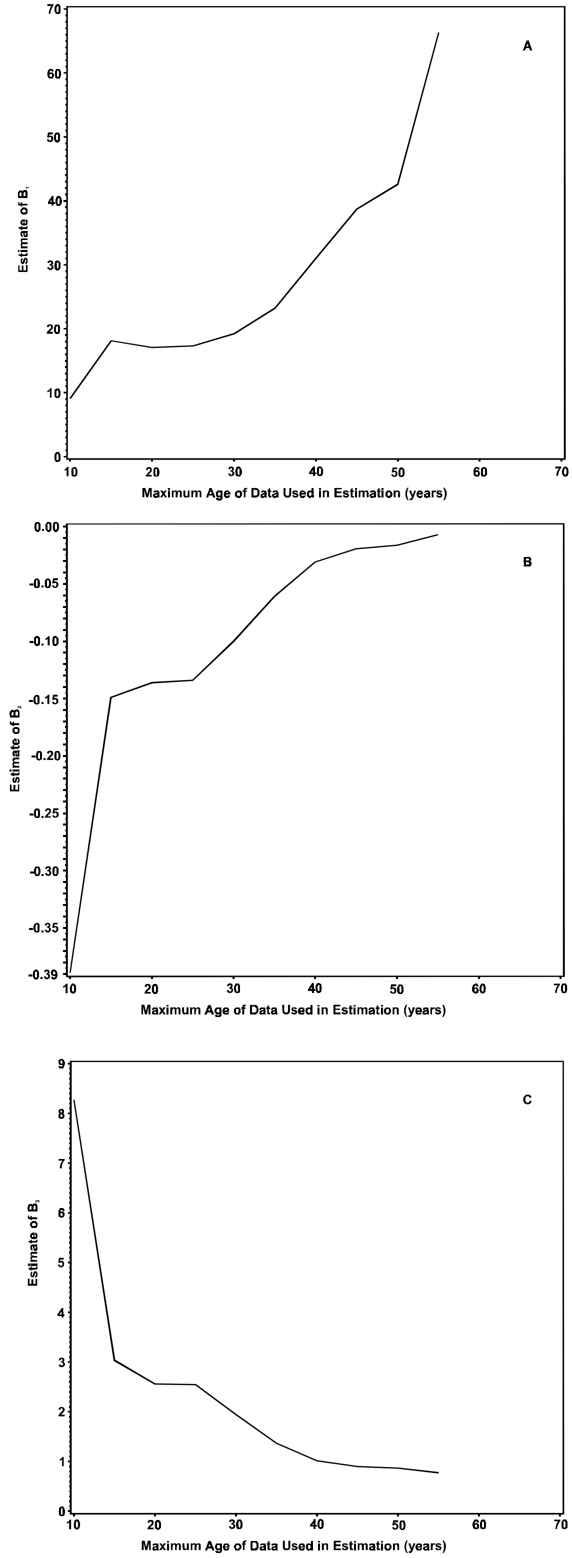

Fig. 6. *Parameter estimates of equation [1], related to the maximum age of the data used in fitting quadratic mean diameter to age.*

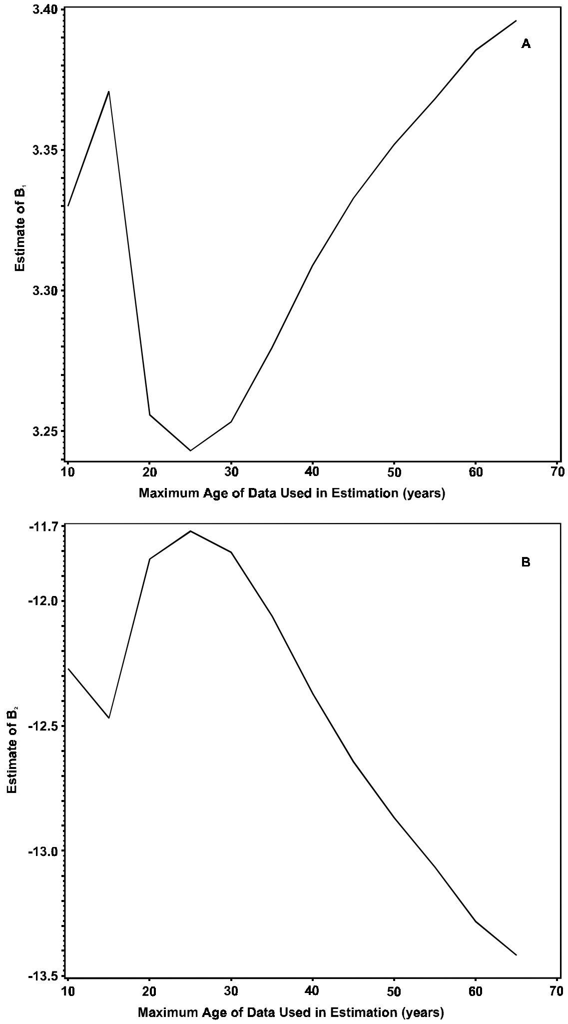

Fig. 7. *Parameter estimates of equation [2], related to the maximum age of the data used in fitting quadratic mean diameter to age.*

(Figure 9). After age 35 data were used, equation [1] had become superior for both error criteria. The lines for equation [1] did not flatten out until the older data were used. The lines for equation [2] did not fluctuate as greatly as equation [1]. The graphs suggest that neither equation represents a compelling candidate model, although the evidence is not as strong as from the parameter values. Although equation [1] was superior to equation [2], it did not level-off until after age 40 data were used.

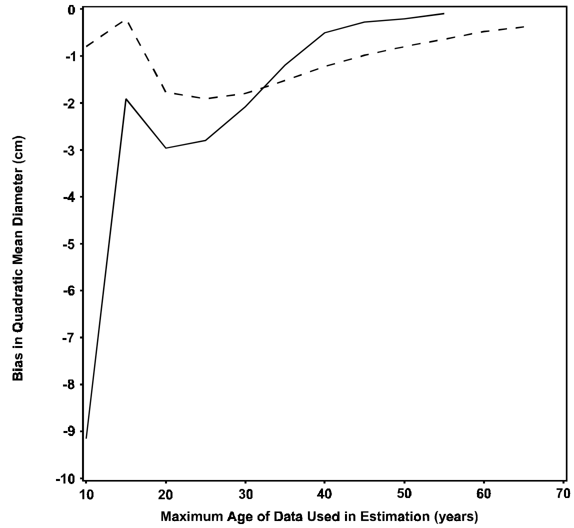

Fig. 8. *Bias in quadratic mean diameter, related to the maximum age of the data used in fitting. Solid line represents equation [1] and dashed line represents equation [2].*

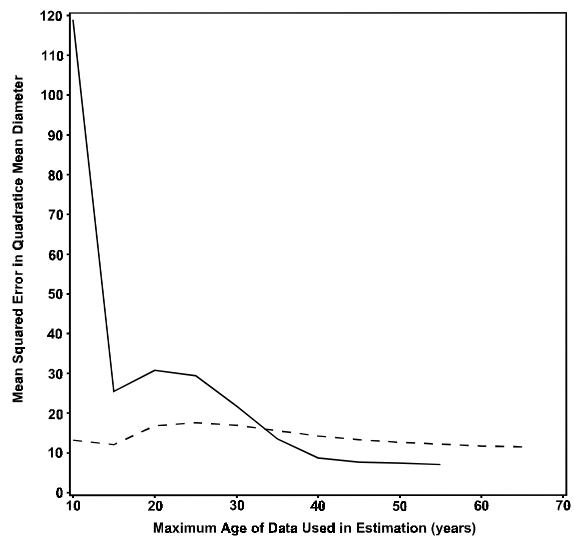

Fig. 9 *Mean squared error in predicting quadratic mean diameter, related to the maximum age of the data used in fitting. Solid line represents equation [1] and dashed line represents equation [2].*

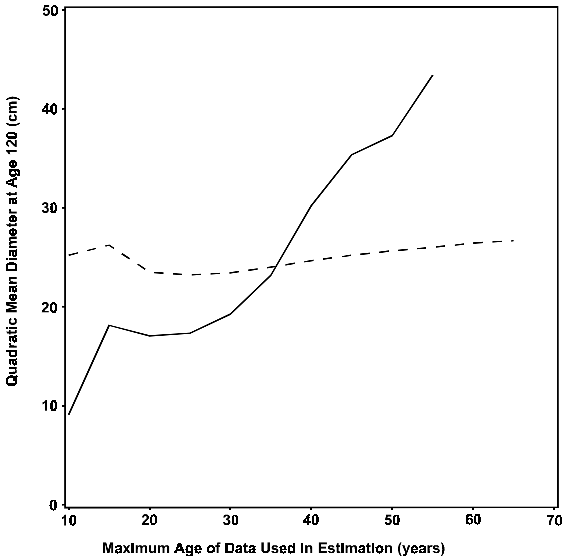

Fig. 10. *Predicted quadratic mean diameter at age 120, related to the maximum age of the data used in fitting. Solid line represents equation [1] and dashed line represents equation [2].*

The predicted quadratic mean diameter at age 120 is presented in figure 10. For equation [1], this value does not stabilize at all as older data is added. On the other hand, equation [2] does stabilize, but at about 25 cm dbh, which is too low. This provides more evidence that neither function is appropriate for this data.

***Diagnostic One, Simulated Data*.**

In this example, data were generated by equation [1] with parameter values the same as the fit to "all data", and with a normal random error added with a variance of 0.059 m^2/ha. There were 100 plots generated, each with values at ages 15 to 60 years. In this example, fits were made to individual plots. As equation [1] is the function used to generate the data, we expect ideal behavior for equation [1] in this diagnostic.

Figure 11 represents box-plots for the parameters of equation [1] from the 100 plots. Average parameter values varied little as

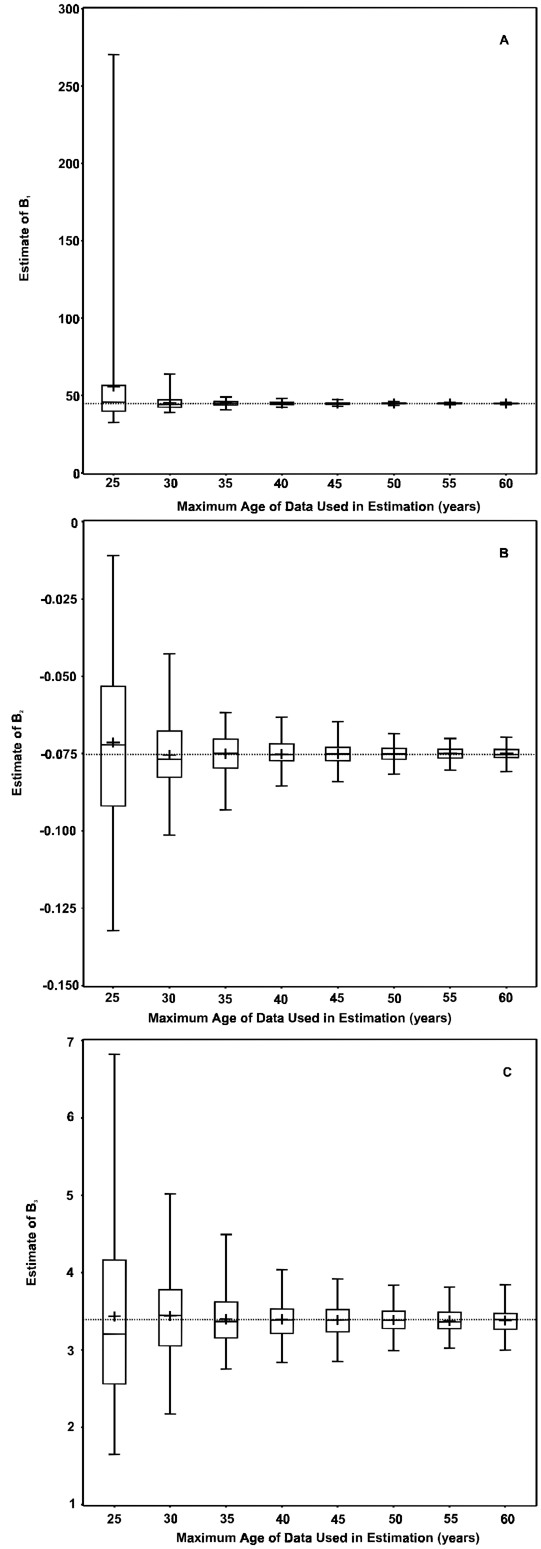

Fig. 11. *Parameter estimates of equation [1], related to the maximum age of the data used in fitting basal area to age, using simulated data.*

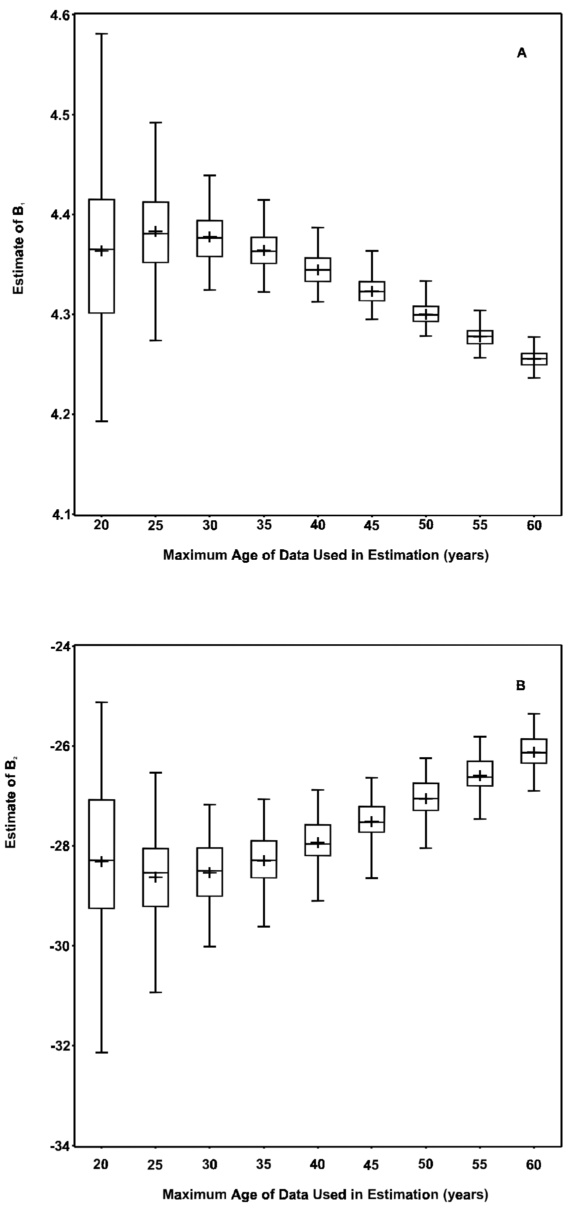

Fig. 12. *Parameter estimates of equation [2], related to the maximum age of the data used in fitting basal area to age, using simulated data.*

more data was used in estimation, however, the variability of those parameters decreased. Figure 12 represents the results from equation [2]. Neither parameter has stabilized as older data were added. Thus, we can infer that equation [2] is not well-suited for extrapolation. More generally, we can infer that equation [2] does not well-represent the true function.

The error criteria confirm this finding. For bias (figure 13), and mean squared error (figure 14), equation [1] is superior to equation [2] in terms of magnitude, once the number of ages used in estimation exceeded the number of parameters (age 30), and also levels off much more quickly. Equation [1] is much more suitable for extrapolation, as would be expected since equation [1] is the true function. Thus, this diagnostic can also be considered to be a general goodness of fit diagnostic.

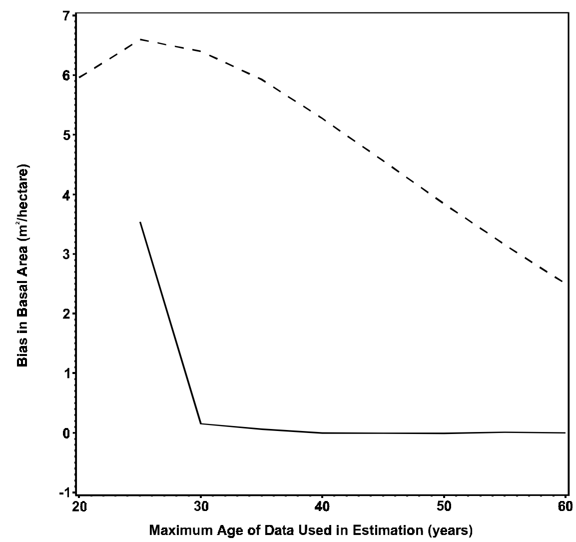

Fig. 13. *Bias in basal area, related to the maximum age of the simulated data used in fitting. Solid line represents equation [1] and dashed line represents equation [2].*

When an algebraic difference equation is employed ($Y_i= f(Y_j, A_i, A_j)$, where A_j signifies age at time j, which is generally less than A_i), it may be estimated by using the immediately-previous observation as a predictor (j=i-1). Borders et al. (1988) termed this “nonoverlapping growth intervals”.

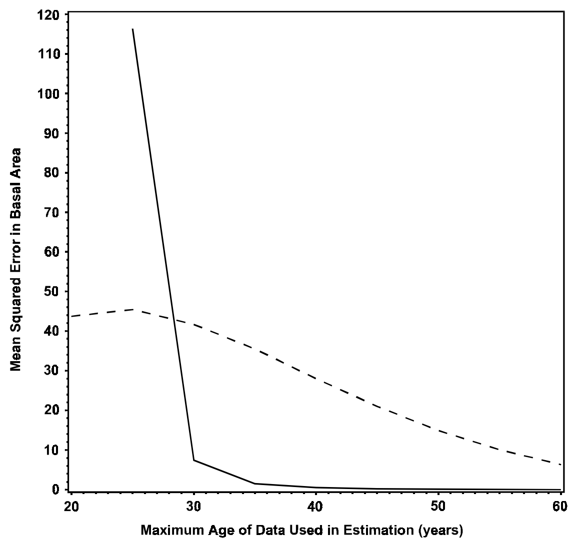

Fig. 14. *Mean squared error in predicting basal area, related to the maximum age of the simulated data used in fitting. Solid line represents equation [1] and dashed line represents equation [2].*

I suggest a diagnostic for “extrapolability” and general goodness of fit is to apply the model to longer-term predictions. Thus, the algorithm for this diagnostic is: (1) fit a difference function using a nonoverlapping growth intervals dataset; (2) calculate error criteria for predictions of all possible multiples of growth intervals, and plot the error criteria against the number of growth intervals. If the data do not generally follow the same interval between measurement, then length of the growth interval will be more suitable than would number of intervals.

The two difference equations that were used were a Richards function:

$$Y_2 = Y_1\left(\frac{1-e^{b_1 X_2}}{1-e^{b_1 X_1}}\right)^{b_2} \qquad [3]$$

and a Schumacher-Coile type model:

$$\ln(Y_2) = \frac{X_1}{X_2}\ln(Y_1) + b_1\left(1-\frac{X_1}{X_2}\right) \qquad [4]$$

Bias is plotted against number of growth intervals in figure 15, and mean squared error is plotted against number of growth intervals in figure 16. Although the two models perform similarly when the equation was used to project one growth interval, the Richards function becomes progressively worse than the Schumacher-Coile model as the projection length is increased. The inference from this diagnostic is that the Schumacher-Coile model is superior to the Richards function with regard to long-term projection. Note that it would also be relevant to conduct diagnostic one with these models; they are estimated sufficiently differently that the results for model [1] do not relate directly to model [3] nor do the results for model [2] relate to model [4].

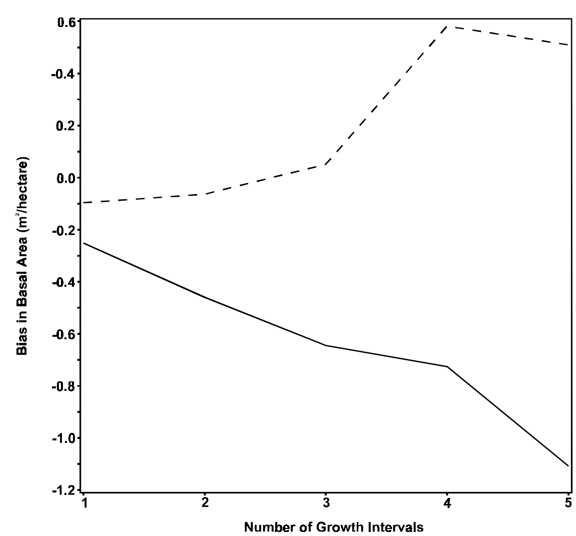

***Fig. 15**. Bias in basal area, related to the number of growth intervals of the prediction. Solid line represents equation [3] and dashed line represents equation [4].*

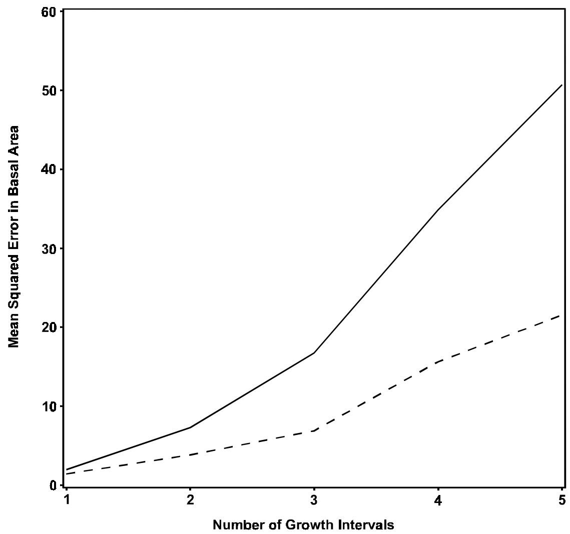

Fig. 16. *Mean squared error in basal area, related to the number of growth intervals of the prediction. Solid line represents equation [3] and dashed line represents equation [4]*.

Generality of these Diagnostics

These diagnostics have utility beyond assessing the extrapolability with respect to age. First, they may be used in a similar fashion to assess extrapolability with respect to other variables. Second, they may be used as general goodness of fit diagnostics. A model that predicts the oldest data poorly, when that data is excluded from the estimation (diagnostic one), probably does not well-represent the true underlying function. A model that predicts long term projections poorly, when estimation is based upon short-term projections (diagnostic two), probably does not well-represent the true underlying function. However, care must be exercised in the use of these diagnostics or short-term accuracy may be sacrificed. Furthermore, standard model selection procedures should also be applied.

Supplementary Data

When models do not seem to have high extrapolability, and as a test even when they do seem to be extrapolable, the data may be supplemented with information from older stands that may be indicative of old plantations. There are several potential sources for this supplemental data. A very rare, old, small plantation may exist. Even if a detailed history of the plantation is unavailable, it may provide information regarding limits of size of individual trees or limits of stand density. Old, monospecific natural stands may also be helpful for setting limits. In figure 17, we supplemented our quadratic mean diameter data with some estimates for old-growth stands in Texas provided by Wahlenberg (1946, p236).

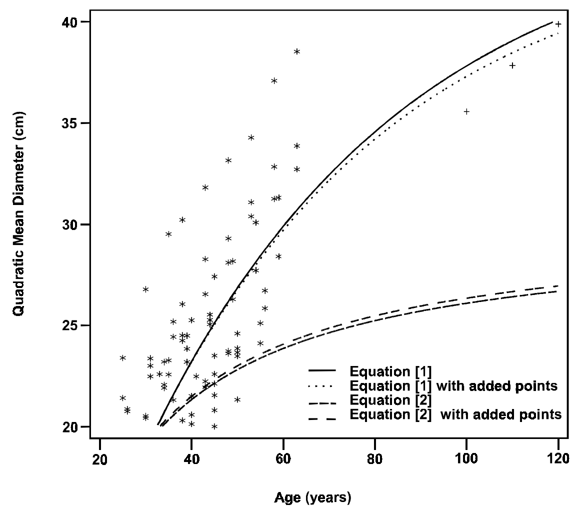

Fig. 17. *Influence of the addition of three data points from old-growth stands on estimation of quadratic mean diameter from a data set of longleaf pine plantations. The graph is truncated to allow distinction between lines.*

Although the few data points did not appreciably alter these equations, that is not necessarily the case. When such supplementary data are used, and their effect is large, careful thought should ensure that they are representative of older plantations. If the resulting equation is little-changed within the range of the data, but complies with the supplemental data outside that range, the use of the supplemental data may be reasonable. If the addition of the supplemental data greatly affects the behavior of the equation within the range of the actual data, then use of the supplemental data may be ill-advised. Supplemental data, experience, and anecdotal descriptions may also be used as informal "reality checks" on model behavior.

Conclusion

Model building in an ecosystem restoration context typically has difficulties that are not parcel of standard growth and yield modeling. Data is likely deficient in representing the range of ages to final rotation age, and may be deficient as far as the scope of sites and range of management practices that will be applied. Typically interest will be not solely in the trees, but also in one or more other component of the ecosystem.

Acknowledgements

This paper benefited from discussions I had over ten years ago with my former Ph.D. advisor, T.E. Burk, University of Minnesota. The paper was reviewed by D.J. Leduc of the Southern Research Station, U.S. Forest Service, Pineville, LA. Mr. Leduc also made the figures.

References

Borders, B.E., R.L. Bailey, and M.L. Clutter. 1988. Forest growth models: parameter estimation using real growth series. *in* Proceedings, IUFRO conference: Forest growth modeling and prediction, 23-27 Aug. 1987, Minneapolis, MN. *Edited by* A.R. Ek, S.R. Shifley, and T.E. Burk. Vol. 2. USDA Forest Service, North Central For. Exp. Stn. St. Paul, MN. Gen. Tech. Rep. NC-120. pp. 660-667

Boyer, W.D. 1990. *Pinus palustris* Mill. Longleaf Pine. *In* Silvics of North America. Vol. 1, Conifers. *Coordinated by* R.M. Burns and B.H. Honkala. USDA Forest Service, Washington, DC. Agric. Handb. 654. pp. 405-412.

Cao, Q.V. and V.C. Baldwin, Jr. 1999. A new algorithm for stand table projection models. Forest Science 45: 506-511.

Clutter, J.L. and E.P. Jones, Jr. 1980. Prediction of growth after thinning of old-field slash pine plantations. USDA Forest Service,Southeast. For. Exp. Stn. Ashville, NC. Res. Pap. SE-217. 19p.

Franklin, R.M. 1997. Stewardship of Longleaf Pine Forests: A Guide for Landowners. Longleaf Alliance Rept. No. 2. The Longleaf Alliance, Solon Dixon Forestry Education Center, Andalusia, AL. 44p.

Goelz, J.C.G. [in press] Growth and yield models viewed from a diameter distribution perspective. J. Math. Model. Scient. Comput.

Goelz, J.C.G. and D.J. Leduc. [in press, a]. Long-term studies on development of longleaf pine plantations. *In* Proceedings of the Third Longleaf Alliance Regional Conference. Forest for Our Future. 16-18 October 2000, Alexandria, LA. *Edited by* J.L. Kush. The Longleaf Alliance and Auburn University, Auburn University, AL.

Goelz, J.C.G. and D.J. Leduc [in press,b]. A Model describing growth and development of longleaf pine plantations: Consequences of observed stand structures on structure of the model. *In* Proceedings of the 11th Biennial Southern Silvicultural Research Conference, 19-21 March 2001, Knoxville, TN. Edited by K. Outcalt. USDA For. Serv. Gen. Tech. Rep. SRS-XXX.

Kmenta, J. 1986. Elements of Econometrics, 2nd ed. Macmillan Publ. Co., New York, NY. 786p.

Lohrey, R.E. and R.L. Bailey. 1977. Yield tables and stand structure for unthinned longleaf pine plantations in Louisiana and Texas. USDA Forest Service, Southern Forest and Range Experiment Station. Res. Pap. SO-133. 53p.

Nepal, S.K. and G.L. Somers. 1992. A generalized approach to stand table projection. Forest Science 38: 120-133.

Somers, G.L. and R.M. Farrar, Jr. 1991. Biomathematical growth equations for natural longleaf pine stands. Forest Science 37: 227-244.

Wahlenberg, W.G. 1946. Longleaf Pine. Its Use, Ecology, Regeneration, Protection, Growth, and Management. Charles Lathrop Pack Forestry Foundation, Washington, DC. 429p.