ACCURACY OF EASTERN WHITE PINE SITE INDEX MODELS DEVELOPED IN THE SOUTHERN APPALACHIAN MOUNTAINS

W. Henry McNab[1]

Abstract-Three older, anamorphic eastern white pine (*Pinus strobus* L.) site index models developed in the southern Appalachian Mountains between 1932 and 1962 were evaluated for accuracy and compared with a newer, polymorphic model developed in 1971. Accuracies of the older models were tested with data used in development of the 1971 model, in which actual site index had been determined by stem analysis. The 1971 model could not be evaluated for accuracy because independent data were unavailable. Evaluation statistics included prediction accuracy, bias, variance, mean square error, and tolerance interval. For one of the older models, prediction accuracy within 5 percent of observed site index was 100 percent, and other statistics compared favorably. Based on the premise that a polymorphic model best describes growth of eastern white pine over a range of site qualities, the site index model developed in 1932 performed surprisingly well.

INTRODUCTION

Eastern white pine (*Pinus strobus* L.) has long been recognized as one of the most valuable timber species in the southern Appalachian Mountains. This conifer is widely managed in natural and planted stands because of its desirable growth and yield characteristics, as well as the high value of its products. Site index (SI) -the average total height of the dominant and codominant trees of a stand at a specific standard age (Chapman and Myer 1949)—typically is used to measure the relative productivity of this species (Beck 1971). Site index relationships have been developed using various techniques, initially based on purely graphical methods and more recently based entirely on mathematical techniques (Chapman and Myer 1949). All types of SI relationships will be referred to as models in this paper.

Barrett first developed an SI model for eastern white pine (hereinafter white pine) in the southern Appalachian Mountains in 1932. Other models were developed as methods changed for quantifying the relationships that describe tree height increment over time. Five models based on data from the southern Appalachian Mountains are now available for white pine. The most recent model was developed by Beck (1971).

Potential problems associated with developing SI models are well known (Beck 1971, Beck and Trousdell 1973). Most problems are related to the inclusion of data from unrepresentative stands and inadequate methods of data analysis (Beck 1971). Each new SI study undoubtedly has reflected investigator intent to overcome perceived problems with earlier models. Therefore, a logical question might be: "Have white pine SI models evolved from less accuracy to greater accuracy over the past 70 years?" None of the southern Appalachian models has been tested for accuracy. This paper evaluates the accuracy of white pine SI models developed in the southern Appalachians.

METHODS

Site index Models

I examined the performance of four SI models that use a standard age of 50 years:

1. Barrett (1932) developed the first set of SI curves from "...measurements of 376 dominant and codominant trees growing in mixture with hardwoods..." He did not state his method for development of these curves, but likely based it on the guide-curve technique, where the age and height of individual trees throughout a region are measured, and one must assume that the population of site indices has been sampled adequately across all stand ages. The resulting SI model is derived from a single guide-curve that describes the average height increment relationship for the total set of sampled stands (Chapman and Meyer 1949). Site index models of this type are termed anamorphic because one curve shape describes the height-growth relationship over the entire range of site qualities sampled.

2. Doolittle and Vimmerstedt (1960) supplemented Barrett's data with additional observations from 105 plots in natural stands of pure white pine and mixed species composition in northern Georgia and western North Carolina. They, too, used the guide-curve method. However, recognizing that the rate of height growth varied with site quality, they attempted to correct for that effect using a mathematical technique based on the coefficient of variation (Chapman and Meyer 1949).

3. Vimmerstedt (1959, 1962) sampled 78 planted stands in North Carolina, Tennessee, and Georgia and established 111 plots for preparation of an SI model. Using linear regression, they developed an equation for predicting tree height at 25 years as a function of height and age, but they did not present statistics describing fit of the model. Vimmerstedt (1962) presented a conversion factor for

[1]Research Forester, Bent Creek Experimental Forest, 1577 Brevard Road, Asheville, NC 28806.

Citation for proceedings: Outcalt, Kenneth W., ed. 2002. Proceedings of the eleventh biennial southern silvicultural research conference. Gen. Tech. Rep. SRS-48. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 622 p.

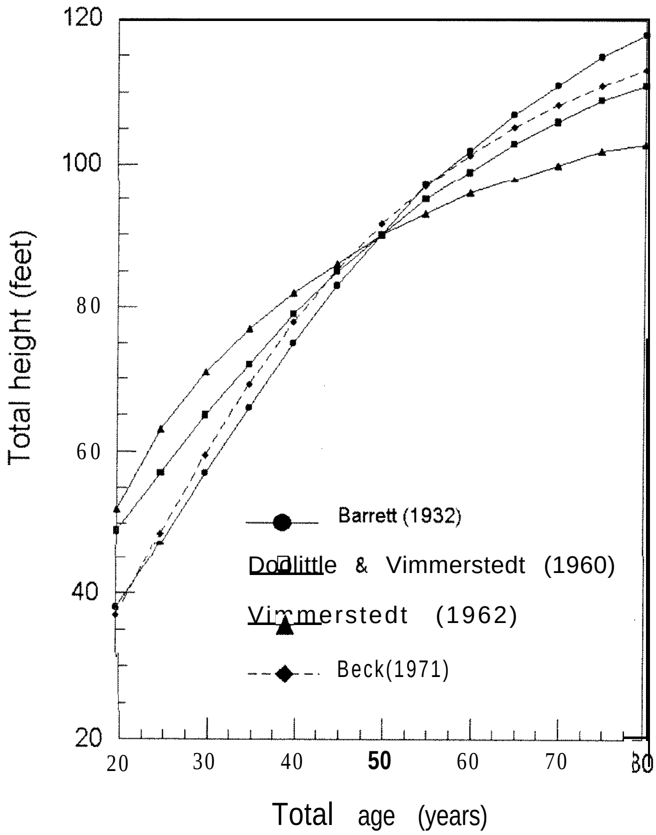

Figure l-Comparison of eastern white pine site index curves developed in the southern Appalachian Mountains for site index 90 feet.

changing SI at a standard age of 25 to a standard age of 50 years.

4. Beck (1971) sampled 43' even-aged stands of naturally established white pine in western North Carolina, northern Georgia, eastern Tennessee, and southwestern Virginia. He used stem-analysis methods to determine the total height of each sample tree at successive ages, up to and including 50 years, which provided a direct measurement of observed SI for that site. He used a non-linear sigmoid function to derive a set of polymorphic curves whose shape varied in relation to site quality.

Summarized in table 1 are characteristics and ranges of total stand ages and site indices over which each of the four SI relationships can be applied. Predicted stand height over age for each of the models is presented in figure 1 for a SI of 90.

Independent Data Set

I used field data collected by Beck (1971) as an independent data set for evaluating each of the SI models. The SI of Becks (1971) 43 stands averaged 92.7 feet (range 69 - 122), ages averaged 52.5 years (43 - 71), and total heights averaged 95.1 feet (70 - 119). About a quarter of the stands were 48-years-old or less, a quarter were 49 - 51 years, and about half of the stands were 52-years or older (table 2). Additional information on field methods is described by Beck (1971). A deficiency of this independent data set is that it is not a random sample of the population of all site indices, but Beck (1971) selected it to represent certain conditions necessary for development of his model (Beck and Trousdell 1973).

I used each of the four models to predict SI of the 43 stands. I predicted SI to the nearest foot by reading directly from published age and height graphs for the models developed by Barrett (1932), and Doolittle and Vimmerstedt (1960). I obtained predicted SI by solving equations presented by Vimmerstedt (1962) and Beck (1971). However, because independent data were used in development of Beck's SI model, this data set cannot be used to validate his model. Performance results for the model developed by Beck (1971) are presented as a standard for comparison with the other models. The most recently developed SI model (Beck 1971) is referred to as the standard model: the other three are, collectively, the old models.

Model Performance Criteria

SI model performance is associated with and implies an unspecified accuracy of prediction. Accuracy is measured in terms of: bias and precision. Bias of a model is the average difference between predicted and the observed values. Precision is a measure of the scatter of predicted SI values around their mean value. Thus, an SI model may be characterized as: (1) unbiased and precise, (2) unbiased but imprecise, (3) biased but precise, or (4) biased and imprecise. An accurate model should have attributes of being both unbiased and precise. In some instances a model could have varying degrees of bias or imprecision

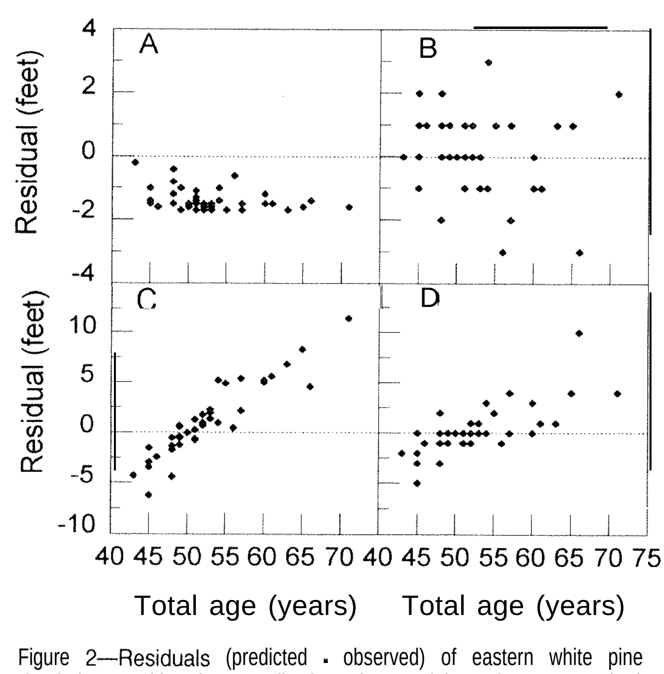

Figure 2—Residuals (predicted - observed) of eastern white pine site index resulting from application of a model used as a standard of comparison (A: Beck 1971) and three models evaluated for accuracy with an independent data set (B: Barrett, 1932; C: Vimmerstedt 1962; D: Doolittle and Vimmerstedt 1960).

'Beck (1971) used only 42 of the 43 stands sampled to develop his SI model. The identity of and reason for excluding one stand is unknown.

Table I--Site index models for eastern white pine developed in the southern Appalachian Mountains

Model source	Stand type	Model format	Standard age (yrs)	Model ranges of Age (yrs)	SI (ft)
Barrett (1932)	Natural	Graph	50	20-120	50-130
Beck (1971)	Natural	Equation	50	5-70	60-130
Doolittle and Vimmerstedt (1960)	Natural	Graph	50	20-100	50-130
Vimmerstedt (1959)	Planted	Equation	25	10-35	40-80
Vimmerstedt (1962)	Planted	Equation	50	10-59	57-115

and still have acceptable accuracy. Each condition presents a different set of implications associated with model accuracy.

I evaluated the performance of each model using a number of statistics associated with accuracy. Because I was interested in learning the difference between observed and predicted SI values, I first determined the residual of each observation (stand):

$$\text{Residual} = (\hat{Y}_i - Y_i)$$

where $\hat{Y}_i$ is the predicted SI for a stand, and Y_i is the observed SI value for the same stand. For many statistical comparisons, the standard method of calculating residuals is $(Y_i - \hat{Y}_i)$. However, I and others (Wiant 1993, Rauscher and others 2000) have used the reverse formulation because it provides results that are more easily comprehended: model overpredictions are positive errors and underpredictions are negative errors.

Bias is the mean of the residuals for all stands:

$$\text{Bias} = \Sigma\,(\text{Residual}_i)/n$$

where n is the number of sampled stands (here, 43).

The scatter of the residuals around the mean of observed SI for a model is a measure of its precision, which is quantified by the variance:

$$\text{Variance} = \Sigma\,(\bar{Y} - Y_i)^2/n-1$$

where $\bar{Y}$ is the mean of all observed SI.

The bias and variance can be combined into single statistic, the mean square error:

$$\text{MSE} = \text{bias}^2 + \text{variance}$$

which provides a measure of the model accuracy and is an indication of the model that performs best overall for estimation of SI. A disadvantage of MSE is that it cannot be used to compare relative performance of models from other studies because it is dependent on the number of observations.

Two other statistics were used to overcome the limitations of MSE and provide a more easily understood measure of future prediction errors: prediction accuracy and tolerance interval. Rauscher and others (2000) used prediction accuracy (PA) to provide a measure of the proportion of predictions that occurred within a specified distance of the observed value. I used a PA value of ±5 percent (e.g. PA-5), which is about equivalent to estimates within one SI class

Table P-Number of stands by total age and observed site index classes in the independent data set sampled by Beck (1971)

Age[a] (yrs)	Site index (ft) 70	75	80	85	90	95	100	105	110	115	120	Total
45				1	1	1	1		1		1	6
50	1	1	3	3	4	1	3	2	2		1	21
55			2	1	1	3			1	1		9
60					2			1				3
65	1	1					1					3
70		1										1
Total	2	3	5	5	8	5	5	3	4	1	2	43

“Midpoints of age and site index classes (e.g., 45 = 43 through 47, 70 = 68 through 72).

Table 3-Error analysis statistics for three site index models developed for eastern white pine in the southern Appalachian Mountains compared to a model developed by Beck (1971) that was used as a standard of comparison

Category of site index model and source of model	Statistic PA-5[a] (pct)	Bias (ft)	Variance (ft)	MSE[b] (ft)	TI[c] (ft)
Site index models tested					
Barrett (1932)	100	0.17	1.56	1.59	3.03
Doolittle and Vimmerstedt (1960)	93	0.14	5.86	5.89	5.86
Vimmerstedt (1962)	77	1.14	12.22	13.53	8.47
Site index standard					
Beck (1971)	100	-1.47[d]	0.13	2.30	0.89

[a]Precision accuracy = Percent of predicted site index values within 5 pct of actual.
[b]Mean square error = $Bias^2$ + variance.
[c]Tolerance interval = Bias ± limits of SI that will include 95 pct of future errors 95 pct of the time.
[d]Significantly different from zero at the 0.05 level of probability.

of 10 feet. Reynolds (1984) suggested calculation of tolerance interval as a means of determining the limits within which most errors will occur in an SI model. The tolerance interval is equal to the mean bias plus or minus the limits of predicted SI that will include 95 percent of future errors at a 0.95 level of probability. I calculated all statistics (except PA-5) using the computer program DOSATEST, which was developed by Rauscher (1986) and refined by Wiant (1993). DOSATEST calculates a trimmed mean and jackknife standard deviation for appropriate tolerance intervals if errors are not normally distributed (Wiant 1993). Accuracy testing using these five statistics (bias, precision, MSE, prediction accuracy, and tolerance interval) and the DOSATEST software has been reported by Wiant (1993) and Rauscher and others (2000).

RESULTS AND DISCUSSION

The PA-5 statistic was highest (100 percent) for two models, Barrett (1932) and the standard (table 3), indicating that all predicted values of SI were within 5 percent of observed. Only 77 percent of stand SI values predicted by the Vimmerstedt (1962) model were within these limits. The pattern of residuals of predicted and observed SI differed for each model (figure 2).

None of the three old SI equations was significantly biased (table 3). However, the equation developed by Beck (1971) exhibited a bias of -1.47 feet (see panel A in figure 2), which was significantly different from zero at the 0.05 level of probability. For example, on a plot with tree height 90 feet at 50 years, the standard model predicts SI as about 88.5 feet. The observed bias results from the model not being constrained, or adjusted, to pass through a value of SI equal to stand height at 50 years standard age (Personal communication T.Lloyd, Research Forester, USDA Forest Service, 1577 Brevard Road, Asheville, NC 28806), as is generally customary in most SI models. Constraining the model was not addressed by Beck (1971), but likely was not done in order to provide a model of greater overall accuracy. In contrast, Trousdell and others (1974) used a similar model formulation to develop SI curves for loblloly pine (*Pinus taeda* L.) and adjusted the curves to pass through the indicated SI at age 50.

The tolerance interval was least for the standard model, which suggests a high degree of accuracy that is associated with small errors of prediction. Among the old models, tolerance interval was smallest (3.03 feet) for Barrett's (1932) and greatest (8.47 feet) for Vimmerstedt's (1962). For Barrett's model, which has a bias of zero (i.e. mean bias was 0.17, which was not significantly different from zero), the tolerance interval may be interpreted to indicate a 95 percent confidence that at least 95 percent of the population of future errors will occur within an interval of about ±3 feet of actual SI.

Mean square error, which combines the effects of bias and variance, was least for Barrett's model. The relatively large bias of the standard model (-1.47 feet) contributed to its large MSE. In many situations, however, a model with a large bias and small variance (e.g., Beck 1971) is preferable to a model with a small bias and large variance (e.g., Barrett, i932). This is because prediction errors associated with bias can be easily corrected, but accounting for error arising from imprecision is problematic.

An explanation for the relatively poor performance of the Vimmerstedt (1962) model is likely due to several causes. First, unlike the other SI models evaluated, this one was developed in planted stands of white pine but tested using data from natural stands. Effects of stand establishment-method and species composition on SI relationships for white pine are not well known, although planted seedlings typically exhibit greater height growth than natural seedlings until about 5 years (Personal communication, Brian flitter, Forestry Supervisor, Biltmore Estate, One North Pack Square, Asheville, NC 28801). Second, over 80 percent of sample trees used in development of the Vimmerstedt model were less than 25 years of age, which tended to weight the curves away from height patterns at a standard age 50 years. Last, Vimmerstedt (1962) presented without explanation a single factor for converting SI at base age 25 to base age 50. Application of the single factor suggests that total height at age 50 would be 1.4335 times that measured at age 25 on all sites. It seems likely that use of a single conversion factor would reduce accuracy of SI models at higher and lower site qualities. In comparison, Trousdell and others (1974) found that height of loblolly pine at 50 years ranged from about 1.4 to 1.7 times that at age 25, depending on site quality. The combination of these and other unknown factors likely contributed to reduced performance of the Vimmerstedt (1962) model.

The tests I conducted were restricted to stand ages 43 - 71 years, which covered only about half the age ranges applicable for most of the models. Tests of the models at younger ages were not possible due to lack of independent data. However, performance of the SI models for younger stand ages may be implied by their performance at the older ages. Assuming that the standard model offers the best representation of height for white pine at all ages, the model developed by Barrett (1932) probably would perform well in younger stands.

CONCLUSIONS

This study has shown that accuracy of eastern white pine SI models varies in the southern Appalachian Mountains. None of the three tested SI models exhibited performance superior to the most recently developed polymorphic model (Beck 1971), which, however, could not be evaluated because a satisfactory data set was not available. One of the anamorphic models (Barrett 1932) compared favorably to the standard model, and several components of its accuracy (bias and MSE) were slightly superior to the standard. The data presented in table 3 are statistics of fit for Beck's (1971) model, rather than independent tests of accuracy.

Results of this study should be useful to researchers for designing new studies and in helping managers decide which SI model to use. One reason I made this study was recognition of how little information is in the literature on the topic of SI validation testing. Site index models are one of the most commonly used forms of prediction equations in forestry; they typically are developed, presented, and used with no accompanying evaluation of performance. The DOSATEST program provides an easy-to-use tool for making tests of accuracy. The primary conclusions are that plantation SI curves seem to differ from natural stands, and that curves developed at two different times for the same region using very different model developmental techniques produced very similar results.

ACKNOWLEDGMENTS

I thank H. Michael Rauscher for providing the DOSATEST software. I also thank F. Thomas Lloyd and Mike Rauscher for many discussions on the topic of accuracy tests in forest measurements that led me to make this study. Tom Lloyd and Brian Ritter reviewed an early draft of this paper. Their suggestions for improvements are greatly appreciated.

LITERATURE CITED

Barrett, L.I. 1932. Second growth white pine height growth classification. Tech. Notes 5. U.S. Department of Agriculture, Forest Service, Appalachian Forest Experiment Station. 1 p.

Beck, D.E. 1971. Polymorphic site index curves for white pine in the Southern Appalachians. Res. Pap. SE-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 8 p.

Beck, D.E.; Trousdell, K.B. 1973. Site index: accuracy of prediction. Res. Pap. SE-l 08. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 7 p.

Chapman, H.H.; Meyer, W.H. 1949. Forest mensuration. New York: McGraw-Hill Book Co. 522 p.

Doolittle, W.T.; Vimmerstedt, J.P. 1960. Site index curves for natural stands of white pine in the Southern Appalachians. Res. Note 141. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 2 p.

Rauscher, H.M. 1986. The microcomputer scientific software series 4: testing prediction accuracy. Gen. Tech. Rep. NC-l 07. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station. 19 p.

Rauscher, H.M.; Young, M.J.; Webb, C.D.; Robison, D.J. 2000. Testing the accuracy of growth and yield models for southern hardwood forests. Southern Journal of Applied Forestry. 24: 176-185.

Reynolds, M.R. 1984. Estimating the error in model predictions. Forest Science. 30: 454-469.

Trousdell, K.B.; Beck, D.E.; Lloyd, F.T. 1974. Site index for loblolly pine in the Atlantic Coastal Plain of the Carolinas and Virginia. Res. Pap. SE-115 Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 11 p.

Vimmerstedt, J.P. 1959. Site index curves for Southern Appalachian white pine plantations. Res. Note 131. Asheville, NC: US. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 2 p.

Vimmerstedt, J.P. 1962. Southern Appalachian white pine plantations: site, volume, and yield. Stn. Pap. 149. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 13 p.

Wiant, H.V., Jr. 1993. DOSATEST for testing accuracy. Compiler. 11: 28-29.