Estimating coefficients of base-age-invariant site index equations[1]

Quang V. Cao
School of Forestry, Wildlife, and Fisheries, Louisiana State University, Baton Rouge, LA 70803, U.S.A.

Received June 1, 1992
Accepted April 1, 1993

Cao, Q.V. 1993. Estimating coefficients of base-age-invariant site index equations. Can. J. For. Res. **23**: 2343–2347.

Remeasurement height data used in this study were collected from plantations of loblolly pine (*Pinus taeda* L.) established in the Southwide Pine Seed Source Study. Three different methods for estimating coefficients of five base-age-invariant site index models were evaluated. The first method involves obtaining coefficients from either a height–age equation or a differential equation. Coefficients from the remaining two methods were estimated from a height growth equation (difference equation) of which the dependent variable is either stand height or logarithm of height. Statistics used in the evaluation were mean of the differences (between the observed and predicted stand heights), mean of the absolute differences, and square root of mean squared error. Results indicated that in most cases, coefficients of the site index models evaluated should be obtained from the height growth equation.

Cao, Q.V. 1993. Estimating coefficients of base-age-invariant site index equations. Can. J. For. Res. **23** : 2343–2347.

Les données de hauteur utilisées dans cette étude proviennent du remesurage de plantations de pins à encens (*Pinus taeda* L.) établies dans le «Southwide Pine Source Study». Trois différentes méthodes d'estimation des coefficients ont été évaluées pour cinq équations d'indice de site invariables basées sur l'âge. Avec la première méthode, les coefficients sont obtenus soit à partir de l'equation hauteur–âge, soit de l'équation différentielle. Les coefficients des deux autres méthodes sont obtenus par l'équation de croissance en hauteur (équation différentielle), dans laquelle la variable dépendante est soit la hauteur du peuplement, soit le logarithme de la hauteur. Les statistiques utilisées pour l'évaluation sont la moyenne des différences (entre les hauteurs observées et prédites du peuplement), la moyenne des différences absolues et la racine carrée du carré moyen des erreurs. Les résultats montrent que, dans la plupart des cas, les coefficients des modèles d'indice de site évalués devraient être obtenus par l'équation de croissance en hauteur.

[Traduit par la rédaction]

Introduction

Forest managers have become interested in using growth and yield models to aid them in forest management planning. Height projection equations such as site index equations are an essential component in such growth and yield models. Base-age-invariant site index equations are frequently used because the base age can be changed without refitting the equations. These curves can originate from *height–age equations* of the form $H = f(A)$, or from *differential equations* of the form $\mathrm{d}H/\mathrm{d}A = g(A, H)$, where H is stand height (average height of the dominants and codominants) at age A. The resulting site index model form for either type of equation is $H = w(A, S, I)$, where S is site index at base age I.

In this study, three methods used for fitting base-age-invariant site index models were evaluated. In method *a*, coefficients of the height–age equation or differential equation are obtained through regression techniques, and these coefficients are then substituted into the site index equation. Method *b* involves rewriting the height–age equation or differential equation into a height growth equation (called difference equation) of the form of $\ln(H_2) = u(A_2, A_1, H_1)$, where H_1 and H_2 are stand heights at ages A_1 and A_2, respectively, and $\ln(x)$ denotes natural logarithm of x. This equation is then fitted to consecutive height–age pairs, and the resulting coefficients are substituted into the site index equation. Method *c* differs from method *b* only in the dependent variable of the height-growth equation, which is $H_2 = v(A_2, A_1, H_1)$.

Furnival et al. (1990) showed analytically that different methods, under some circumstances, were identical for fitting a site index equation that can be transformed into a simple linear model. These results, however, cannot be applied to a majority of current site index models, which are nonlinear in the parameters. The objective of this study is to evaluate the three alternative methods for estimating coefficients of site index equations.

Data

Remeasurement data from the Southwide Pine Seed Source Study were used. Detail was presented by Wells and Wakeley (1966) about this research project, which involved 15 loblolly pine (*Pinus taeda* L.) seed sources planted at 13 areas from Texas to Maryland. A total of 480 plots from 11 locations having height–age measurements through age 25 or 27 were selected for this study.

The data were randomly divided into a fit data set and a validation data set. The fit data (240 plots) were used to estimate the regression coefficients. The validation data set (remaining 240 plots) was used solely to validate the ability of the models to accurately predict height in the population. The height–age distributions for the fit and test data sets are presented in Table 1. There are a total of 1680 consecutive height–age pairs for the fit data set and 1677 pairs for the validation data set.

Models considered

Models developed by Beck (1971), Graney and Burkhart (1973), Trousdell et al. (1974), Burkhart and Tennent (1977*a*, 1977*b*), and many others were not considered in this study because these models required site index as an independent variable and thus were not base-age invariant. Complicated segmented site index equations such as those by Devan and Burkhart (1982) and Borders et al. (1984) were also not included in the evaluation.

The five published models selected for evaluation are listed in Table 2. Models 1–3 belong to the first group of site index models that originated from height–age equations. The second

[1]Louisiana Agricultural Experiment Station manuscript 93-22-7031.

Printed in Canada / Imprimé au Canada

TABLE 1. Summary statistics for stand height of loblolly pine plantations for the fit and validation data sets

Stand age (years)	Fit data set				Validation data set			
	n	Min. (m)	Mean (m)	Max. (m)	n	Min. (m)	Mean (m)	Max. (m)
1	240	0.18	0.36	0.67	238	0.21	0.35	0.73
3	240	0.64	1.58	2.65	240	0.64	1.56	2.65
5	240	1.64	3.56	5.88	239	1.58	3.53	5.67
10	240	4.24	9.02	12.04	240	4.69	9.02	12.07
15	205	8.81	13.03	16.76	208	8.02	13.06	16.43
16	35	14.51	16.34	17.65	32	14.57	15.92	17.59
20	217	12.47	17.24	21.18	229	10.30	17.00	21.34
22	23	11.25	14.96	17.62	11	10.61	14.88	17.28
25	161	14.87	19.92	24.32	182	11.80	19.73	24.05
27	79	12.13	19.79	24.54	58	11.43	20.26	24.87
All	1680	0.18	9.28	24.54	1677	0.21	9.25	24.87

TABLE 2. List of site index models included in the analysis

Model	No.	Equation
Group 1: site index models from height–age equations		
Schumacher (1939)*	1*a*	$\ln(H) = b_1 + b_2/A$
	1*b*	$\ln(H_2) = \ln(H_1) + b_2(1/A_2 - 1/A_1)$
	1*c*	$H_2 = H_1 \exp[b_2(1/A_2 - 1/A_1)]$
	1	$\ln(H) = \ln(S) + b_2(1/A - 1/I)$
Chapman–Richards	2*a*	$H = b_1[1 - \exp(b_2 A)]^{b_3}$
	2*c*	$H_2 = H_1\left[\frac{1 - \exp(b_2 A_2)}{1 - \exp(b_2 A_1)}\right]^{b_3}$
	2	$H = S\left[\frac{1 - \exp(b_2 A)}{1 - \exp(b_2 I)}\right]^{b_3}$
Bailey and Clutter (1974)	3*a*	$\ln(H) = b_1 + b_2 A^{b_3}$
	3*b*	$\ln(H_2) = b_1 + [\ln(H_1) - b_1]\,(A_2/A_1)^{b_3}$
	3*c*	$H_2 = \exp[b_1 + [\ln(H_1) - b_1]\,(A_2/A_1)^{b_3}]$
	3	$\ln(H) = b_1 + [\ln(S) - b_1]\,(A/I)^{b_3}$
Group 2: site index models from differential equations		
Clutter and Lenhart (1968)†	4*a*	$dy/dx = c_1 + c_2 y + c_3 x$
	4*b*	$\ln(H_2) = b_1 + b_2/A_2 + [\ln(H_1) - b_1 - b_2/A_1]\exp[b_3(1/A_2 - 1/A_1)]$
	4*c*	$H_2 = \exp\{b_1 + b_2/A_2 + [\ln(H_1) - b_1 - b_2/A_1]\exp[b_3(1/A_2 - 1/A_1)]\}$
	4	$\ln(H) = b_1 + b_2/A + [\ln(S) - b_1 - b_2/I]\exp[b_3(1/A - 1/I)]$
Amateis and Burkhart (1985)	5*a*	$dy/dx = b_1 y + b_2 y/x$
	5*b*	$\ln(H_2) = \ln(H_1)\,(A_1/A_2)^{b_2}\exp[b_1(1/A_2 - 1/A_1)]$
	5*c*	$H_2 = \exp\{\ln(H_1)\,(A_1/A_2)^{b_2}\exp[b_1(1/A_2 - 1/A_1)]\}$
	5	$\ln(H) = \ln(S)\,(I/A)^{b_2}\exp[b_1(1/A - 1/I)]$

*Where H_i is the average height of the dominants and codominants at age A_i and S is site index at base age I.
†Where $y = \ln(H)$, $x = 1/A$, $b_1 = -(c_1 + c_3/c_2)/c_2$, $b_2 = -c_3/c_2$, and $b_3 = c_2$.

group consists of the remaining two site index models that were derived from differential equations. Coefficients in a site index equation (e.g., eq. 1) can be obtained from any of the three related equations (e.g., 1*a*, 1*b*, and 1*c*). The letters *a*, *b*, and *c* included in the equation numbers corresponded to the three methods evaluated in this study.

Group 1: site index models from height–age equations

The Schumacher (1939) equation (model 1) is a well-known function that was used by Langdon (1959) for slash pine (*Pinus elliottii* Engelm.), Chaiken and Nelson (1959) for Virginia pine (*Pinus virginiana* Mill.), Schlaegel et al. (1969) for yellow-poplar (*Liriodendron tulipifera* L.), and Lenhart (1971) for loblolly pine, to name a few. Furnival et al. (1990) showed that methods 1*a* and 1*b* provided identical results when data from all possible height–age pairs were used.

The Chapman–Richards equation [2*a*] was suggested by Richards (1959) and Chapman (1961) as a generalization of the original growth model derived by von Bertalanffy (1957). The resulting site index model [2] was used by Pienaar and

TABLE 3. Parameter estimates for the height equations for loblolly pine plantations

Model	No.	b_1	b_2	b_3
Schumacher (1939)	1*a*	2.628 14	−4.027 00	
	1*b*		−2.527 87	
	1*c*		−3.374 90	
Chapman–Richards	2*a*	24.601 16	−0.084 53	1.810 44
	2*c*		−0.076 46	1.668 65
Bailey and Clutter (1974)	3*a*	9.633 63	−10.719 97	−0.151 31
	3*b*	12.144 77		−0.113 82
	3*c*	4.981 14		−0.363 99
Clutter and Lenhart (1968)	4*a*	−2.391 06	1.481 63	−2.992 77
	4*b*	−56.407 76	−34.751 65	0.413 56
	4*c*	3.313 53	−13.166 04	7.595 83
Amateis and Burkhart (1985)	5*a*	−5.146 32	0.000 082 166	
	5*b*	3.554 58	−0.814 08	
	5*c*	−4.839 66	0.002 867	

Shiver (1980) for loblolly pine, Lenhart et al. (1986) for loblolly and slash pine, and others. Since it does not make sense to use $\ln(H_2)$ as the dependent variable in the height growth equation for this model, eq. [2*b*] was not considered in this study.

The height–age model used by Bailey and Clutter (1974) was based on a modified form of the Schumacher equation. The assumption that the slope b_2 in eq. [3*a*] varies depending on the site resulted in a polymorphic site index model.

Group 2: site index models from differential equations

Clutter and Lenhart (1968) and Amateis and Burkhart (1985) used a differential equation approach for predicting site index of loblolly pine. These two models are labeled 4 and 5, respectively. The dependent variable (d*y*/d*x*) in each differential equation listed will be computed from the difference of two observations:

$$\frac{dy}{dx} = \frac{y_2 - y_1}{x_2 - x_1}$$

where $y_i = \ln(H_i)$ and $x_i = 1/A_i$; $i = 1$ or 2. The x and y independent variables were calculated as the average of the two adjacent observations.

All models presented above are base-age invariant, and all except the Schumacher and Chapman–Richards models produce polymorphic site index curves.

Model evaluation

Parameter estimates of the site index equations were obtained from consecutive height–age pairs from the fit data set using the three methods described above. Evaluation of the models involved only the validation data set. Evaluation data consisted of all possible combinations of projection lengths (A_1, A_2), separated into three groups: (*i*) 2–5 years, (*ii*) 6–10 years, and (*iii*) 11–26 years. The site index equations were employed to predict stand height H_2 at age A_2 based on height H_1 at age A_1. The performance of the three alternative methods will be evaluated for each of the five models based on the following three criteria:

(1) Mean difference:

$$\overline{\text{Diff}} = \sum_{i=1}^{n} \frac{H_i - \hat{H}_i}{n}$$

where H_i and $\hat{H}_i$ are the ith observed and predicted values, respectively, for stand height in meters and n is number of observations.

(2) Mean absolute difference:

$$\overline{|\text{Diff}|} = \sum_{i=1}^{n} \frac{|H_i - \hat{H}_i|}{n}$$

(3) Root mean squared error:

$$\text{RMSE} = \left(\sum_{i=1}^{n} \frac{(H_i - \hat{H}_i)^2}{n} \right)^{0.5}$$

The statistics computed from the three alternative methods were then ranked for each site index model.

Results and discussion

Coefficients of five models estimated from the fit data set using three methods are presented in Table 3. The three methods were evaluated for each site index model based on the evaluation statistics (Table 4). Rankings of the three methods for each site index model are shown in Table 5. Results were similar for most site index models with a few exceptions.

For group 1, which included models derived from height–age equations (models 1 through 3), method *c* of obtaining coefficients from the height growth equations performed very well. It ranked best among these three methods. The only exception was the Schumacher model, which ranked behind method *a* for projection length over 10 years.

Site index models from differential equations (group 2) also responded well to method *c*, which performed best for projection length over 5 years. Method *c* ranked a close second when the length of projection was between 2 and 5 years.

TABLE 4. Evaluation statistics for the site index models for loblolly pine plantations

Model	No.	Projection length 2–5 years (n = 1584)				Projection length 6–10 years (n = 1186)				Projection length 11–26 years (n = 2253)			
		$\overline{\text{Diff}}$	$\overline{\lvert\text{Diff}\rvert}$	RMSE	Sum of ranks	$\overline{\text{Diff}}$	$\overline{\lvert\text{Diff}\rvert}$	RMSE	Sum of ranks	$\overline{\text{Diff}}$	$\overline{\lvert\text{Diff}\rvert}$	RMSE	Sum of ranks
Schumacher (1939)	1*a*	0.38 (1)	3.06 (3)	3.49 (3)	7	3.55 (1)	5.26 (2)	5.64 (2)	5	7.66 (1)	8.53 (1)	9.69 (1)	3
	1*b*	2.13 (3)	2.27 (1)	2.73 (2)	6	6.22 (3)	6.22 (3)	6.46 (3)	9	12.49 (3)	12.49 (3)	12.92 (3)	9
	1*c*	1.35 (2)	2.40 (2)	2.70 (1)	5	5.05 (2)	5.14 (1)	5.61 (1)	4	10.47 (2)	10.47 (2)	11.09 (2)	6
Chapman–Richards	2*a*	−0.21 (2)	0.79 (2)	1.05 (2)	6	−0.47 (2)	1.75 (2)	2.40 (2)	6	−1.64 (2)	3.96 (2)	5.42 (2)	6
	2*c*	0.04 (1)	0.69 (1)	0.90 (1)	4	2.30 (1)	1.50 (1)	1.92 (1)	3	0.55 (1)	3.05 (1)	3.94 (1)	3
Bailey and Clutter (1974)	3*a*	−0.23 (3)	0.85 (1)	1.09 (2)	6	−0.28 (3)	1.99 (2)	2.36 (2)	7	−0.59 (3)	2.76 (2)	3.48 (2)	7
	3*b*	−0.19 (1)	0.91 (3)	1.16 (3)	7	−0.20 (2)	2.26 (3)	2.63 (3)	8	−0.22 (2)	2.98 (3)	3.74 (3)	8
	3*c*	−0.21 (2)	0.86 (2)	1.07 (1)	5	0.13 (1)	1.36 (1)	1.72 (1)	3	0.14 (1)	2.20 (1)	2.77 (1)	3
Clutter and Lenhart (1968)	4*a*	0.41 (1)	1.50 (3)	2.86 (3)	7	−0.10 (1)	4.71 (3)	22.03 (3)	7	−3.29 (3)	15.53 (3)	107.39 (3)	9
	4*b*	0.44 (2)	0.87 (1)	1.12 (1)	4	1.20 (3)	1.93 (2)	2.35 (2)	7	2.67 (2)	3.31 (2)	4.09 (2)	6
	4*c*	0.52 (3)	0.89 (2)	1.13 (2)	7	0.44 (2)	1.27 (1)	1.62 (1)	4	1.02 (1)	1.93 (1)	2.46 (1)	3
Amateis and Burkhart (1985)	5*a*	0.99 (2)	1.53 (1)	1.93 (1)	4	2.56 (2)	4.05 (2)	5.18 (2)	6	7.06 (1)	10.64 (2)	12.64 (2)	5
	5*b*	−0.24 (1)	3.23 (3)	3.83 (3)	7	−0.57 (1)	9.16 (3)	10.06 (3)	7	7.60 (2)	14.53 (3)	16.33 (3)	8
	5*c*	1.21 (3)	1.53 (1)	1.93 (1)	5	3.22 (3)	3.98 (1)	5.15 (1)	5	8.60 (3)	10.20 (1)	12.07 (1)	5

NOTE: $\overline{\text{Diff}}$, mean difference; $\overline{\lvert\text{Diff}\rvert}$, mean absolute difference; and RMSE, root mean squared error. For each model, number in parentheses following each statistic denotes the rank (1 being best) of the method based on that statistic.

TABLE 5. Rankings (1 being best) of three methods for estimating coefficients by site index model

Model	Projection length 2–5 years			Projection length 6–10 years			Projection length 11–26 years		
	Method *a*	Method *b*	Method *c*	Method *a*	Method *b*	Method *c*	Method *a*	Method *b*	Method *c*
Group 1: site index models from height–age equations									
Schumacher (1939)	3	2	1	2	3	1	1	3	2
Chapman–Richards	2		1	2		1	2		1
Bailey and Clutter (1974)	2	3	1	2	3	1	2	3	1
Group 2: site index models from differential equations									
Clutter and Lenhart (1968)	2	1	2	2	2	1	3	2	1
Amateis and Burkhart (1985)	1	3	2	2	3	1	1	3	1

Method a performed as well as method c for the Amateis–Burkhart model and was in second place for other models. In the case of the Clutter–Lenhart model, however, method a resulted in very high values of mean absolute difference and root mean square error.

Method b differed from method c in using $\ln(H_2)$, not H_2, as the dependent variable of the height growth equation. Method c outperformed method b in most cases, proving that the variable of interest, H_2, should be selected as the dependent variable. An additional method developed by Bailey and Clutter (1974) specifically for fitting their site index equation was also investigated in this study and found to be a close second to method c for model 3.

Evaluation statistics for short-, medium-, and long-ranged height projection showed that the Chapman–Richards, Bailey–Clutter, and Clutter–Lenhart models (2, 3, and 4, respectively) performed very well compared with the other two site index models evaluated in this study.

Data from consecutive differences versus all possible differences

Borders et al. (1988) raised the issue of using all possible differences in modeling basal area and volume growth. They evaluated the use of growth series from consecutive differences, all possible differences, and longest growth interval in two basal area projection models and found that projection capabilities from these different data were dependent on model form.

The evaluation in the present study was based on coefficients of site index models estimated from consecutive differences of height–age pairs. In addition, a similar evaluation based on coefficients from all possible differences was also performed. The rankings of the three methods remained essentially unchanged.

Conclusions

These results showed that method c of estimating coefficients from height growth equations was consistently better than the other two methods in projecting height, with a few exceptions. The performance of a model could suffer if an inappropriate method was selected, as in the case of method a for the Clutter–Lenhart model in long-term projection. Similar results were obtained when data were derived either from consecutive differences or from all possible differences. In general, the Chapman–Richards, Bailey–Clutter, and Clutter–Lenhart models performed well for all projection lengths evaluated in this study.

Amateis, R.L., and Burkhart, H.E. 1985. Site index curves for loblolly pine plantations on cutover site-prepared lands. South. J. Appl. For. **9**: 166–169.

Bailey, R.L., and Clutter, J.L. 1974. Base-age invariant polymorphic site curves. For. Sci. **20**: 155–159.

Beck, D.E. 1971. Polymorphic site index curves for white pine in the Southern Appalachians. USDA For. Serv. Southeast. For. Exp. Stn. Res. Pap. SE-80.

Borders, B.E., Bailey, R.L., and Ware, K.D. 1984. Slash pine site index from a polymorphic model by joining (splining) non-polynomial segments with an algebraic difference method. For. Sci. **30**: 411–423.

Borders, B.E., Bailey, R.L., and Clutter, M.L. 1988. Forest growth models: parameter estimation using real growth series. *In* Forest growth modelling and prediction. Vol. 2. Proceedings of the IUFRO Conference. *Edited by* A.R. Ek, S.R. Shifley, and T.E. Burk. USDA For. Serv. Gen. Tech. Rep. NC-120. pp. 660–667.

Burkhart, H.E., and Tennent, R.B. 1977*a*. Site index equations for radiata pine in New Zealand. N.Z. J. For. Sci. **7**: 408–416.

Burkhart, H.E., and Tennent, R.B. 1977*b*. Site index equations for Douglas fir in Kaingaroa Forest. N.Z. J. For. Sci. **7**: 417–419.

Chaiken, L.E., and Nelson, T.C. 1959. Site index curves for Piedmont Virginia pine. USDA For. Serv. Southeast. For. Exp. Stn. Res. Note 135.

Chapman, D.G. 1961. Statistical problems in dynamics of exploited fisheries populations. *In* Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, 20 June – 30 July 1960, Berkeley, Calif. Vol. 4. *Edited by* Jerzy Neyman. University of California Press, Berkeley. pp. 153–168.

Clutter, J.L., and Lenhart, J.D. 1968. Site index curves for old-field loblolly pine plantations in the Georgia Piedmont. Ga. For. Res. Counc. Rep. 22. Ser. 1.

Devan, J.S., and Burkhart, H.E. 1982. Polymorphic site index equations for loblolly pine based on a segmented polynomial differential model. For. Sci. **28**: 544–555.

Furnival, G.M., Gregoire, T.G., and Valentine, H.T. 1990. An analysis of three methods for fitting site-index curves. For. Sci. **36**: 464–469.

Graney, D.L., and Burkhart, H.E. 1973. Polymorphic site index curves for shortleaf pine in the Ouachita Mountains. U.S. For. Serv. South. For. Exp. Stn. Res. Pap. SO-85.

Langdon, O.G. 1959. Site index curves for South Florida slash pine. USDA For. Serv. Southeast. For. Exp. Stn. Res. Note 133.

Lenhart, J.D. 1971. Site index curves for old field loblolly pine plantations in the interior West Gulf Coastal Plain. Texas For. Pap. 8.

Lenhart, J.D., Hunt, E.V., and Blackard, J.A. 1986. Site index equations for loblolly and slash pine plantations on non-old-fields in east Texas. South. J. Appl. For. **10**: 109–112.

Pienaar, L.V., and Shiver, B.D. 1980. Dominant height growth and site index curves for loblolly pine plantations in the Carolina Flatwoods. South. J. Appl. For. **4**: 54–59.

Richards, F.J. 1959. A flexible growth function for empirical use. J. Exp. Bot. **10**: 290–300.

Schlaegel, B.E., Kulow, D.L., and Baughman, R.N. 1969. Empirical yield tables for West Virginia yellow-poplar. West Virginia University, Morgantown. West Va. Agric. Exp. Stn. Bull. 574T.

Schumacher, F.X. 1939. A new growth curve and its application to timber-yield studies. J. For. **37**: 819–820.

Trousdell, K.B., Beck, D.E., and Lloyd, F.T. 1974. Site index for loblolly pine in the Atlantic Coastal Plain of the Carolinas and Virginia. USDA For. Serv. Res. Pap. SE-115.

von Bertalanffy, L. 1957. Quantitative laws in metabolism and growth. Q. Rev. Biol. **32**: 217–231.

Wells, O.O., and Wakeley, P.L. 1966. Geographic variation in survival, growth, and fusiform-rust infection of planted loblolly pine. For. Sci. Monogr. 11.