COMPARISON OF COMPATIBLE POLYNOMIAL TAPER EQUATIONS

A. GORDON

Forest Research Institute, New Zealand Forest Service,
Private Bag, Rotorua, New Zealand

(Received for publication 28 February 1983; revision 10 August 1983)

ABSTRACT

Polynomial taper equations in the first, second, third, fourth, and fifth powers were fitted to two data sets, conditioned to be compatible to an existing total volume equation. Accuracy and precision were compared with other models that included one higher power term, the value of the exponent for which ranged from 8 to 40. The latter models showed an improvement in prediction of merchantable and butt log volumes as well as inside bark diameter, although still with slight bias in diameter prediction that varied with relative height.

INTRODUCTION

Taper equations are functions which estimate stem diameter from known variables such as breast height diameter, tree height, and distance from the tree tip. The accepted definition of compatible taper equations is that given by Damaerschalk (1971, 1972) who introduced the term – those taper equations which yield the same total volume when integrated over total tree height as given by a volume equation. The theory developed by Demaerschalk (1971, 1972, 1973) and Munro & Demaerschalk (1974) was extended by Goulding & Murray (1976) who derived a general polynomial solution, fitting a fifth-degree model to *Pinus radiata* D. Don data from Kaingaroa Forest. Several solutions using this model are now used extensively in New Zealand Forest Service systems for preharvest planning and inventory.

However, although this model gives sufficiently accurate results for most of the current uses of taper equations, it shows clear bias trends. Some butt swell is accounted for but generally diameters are under-estimated at around 20% of height and over-estimated above 80% of height.

This study shows that, without altering the compatibility constraint or introducing new variables, an improvement in fit can be obtained by including one term of a higher order in the model. Such terms have been used previously by Fries & Matèrn (1966) and Bruce *et al.* (1968), and by Katz *et al.* (in prep.) and were suggested by Goulding & Murray (1976). Graphical comparisons clearly demonstrate the differences in bias trends between models.

The following notation is used:

H = total height (m)

DBH = breast height diameter over bark (cm)

h = distance up the stem from ground (m)

V = estimated total stem volume inside bark

K = $\pi/4\ 10^{-4}$

d = inside bark diameter at point h

z = (H – h)/H

Data

Two sets of data were used for the study. The first (Set I) contained the sectional measurements of 701 *Larix decidua* Mill. trees, covering a wide range of tree sizes and averaging 10 measurements per tree. This sample was drawn from locations throughout New Zealand. The second (Set II) was a tightly defined group of 102 *P. radiata* trees aged between 9 and 11 years, sampled in one location under one silvicultural regime. An average of eight measurements per tree were recorded. Table 1 details tree sizes.

TABLE 1—Tree sizes

		Minimum	Mean	Maximum
Data Set I	DBH (cm)	4.8	26.5	85.1
	Height (m)	3.7	21.7	37.8
	Total stem volume (m^3)	0.004	0.597	4.739
Data Set II	DBH (cm)	11.6	21.6	30.0
	Height (m)	10.0	15.4	20.7
	Total stem volume (m^3)	0.049	0.236	0.419

Over-bark diameters were measured using diameter tape at points on the stem 0.15, 0.7, 1.4, 3, 6, 9, . . . m above ground-level, to within 4.5 m of total tree height. Under-bark diameters were estimated by subtracting the sum of two Swedish-bark-gauge readings taken on opposite sides of the stem at each point. This method, adopted as standard by the New Zealand Forest Service, leads to estimates of total stem volume with little bias and a relatively high level of precision (Goulding 1979).

METHODS

Stem Volumes

Under-bark volumes were calculated by summing the volumes of each section. The sectional volume formulae used assumed the following solids: cylinder from 0.0 to 0.15 m, cone from 0.15 to 1.4 m, paraboloid from 1.4 m to the level of the last measurement, and cone to tree tip.

Function Types

Total stem volume equations of the form

$$V = a_1\ (DBH)^{a_2}\ (H^2/(H-1.4))^{a_3}$$

were fitted using logarithmic transformations and the estimate of a_1 adjusted for logarithmic bias (Finney 1941).

Taper functions were fitted to yield an equation in the form

$$d^2 = \frac{V}{KH} [b_1z + b_2z^2 + b_3z^3 + b_4z^4 + b_5z^5 + b_6z^P] \quad (1)$$

where $\sum_{i=1}^{n} \frac{b_i}{i+1} = 1$ (n = number of terms), and $5 < P < 41$,

as given by Goulding & Murray (1976). The theory is detailed in Appendix 1.

Fitting Procedure

Equation (1) was fitted with two, three, four, five, and six terms in the polynomial. In each attempt the regression with the least residual variance was chosen from the equations with terms as follows.

(1) The term linear in z must appear as its coefficient is not estimated directly but calculated from the other coefficients to satisfy the conditioning;

(2) All combinations of terms in $z^2 \ . \ . \ . \ z^5$, both with and without the best term in the group z^P, where $5 < P < 41$.

Table 2 shows the terms chosen in each attempt, as well as a fifth-order, five-term equation, for Data Set II. It is clear that the residual variance stabilised after the fourth term was added if one of the terms was from the group z^P.

TABLE 2—Residual variance - Data Set II

Best selection of terms	Variance accounted for (%)*	
	All data	Random subset
1, 10	78.2	82.7
1, 2, 22	80.2	85.6
1, 2, 5, 16	80.4	85.7
1, 2, 4, 5, 14	80.4	85.5
1, 2, 3, 4, 5, 12	80.4	85.3
1, 2, 3, 4, 5	79.9	85.2

* Adjusted for degrees of freedom

Because of correlation between measurements within trees the residual error is under-estimated by this fitting procedure. To check if the under-estimate is affected by the number of terms fitted, the equations were refitted to subsets of the data generated by randomly selecting one observation per tree. As shown in Table 2, the trend in the residual variance was very similar. The trends were identical for Data Set I.

Model Selection

This fitting procedure resulted in five models which included a higher power term for each data set. Those models which contained non-significant terms (99% confidence level) or which produced negative values anywhere in the domain of z $(0 \leq z \leq 1)$ were discarded.

Probably because of the large sample size, all coefficients in the six-term model fitted to Data Set I were significant at the 99.9% confidence level, but this model was not used in the comparison as there was no practical difference in fit between it and the five-term model.

Accuracy and Precision

Numerical criteria used to evaluate the models were similar to those used by Cao *et al.* (1980). The ability to predict inside-bark diameters was compared using:

(1) Bias (the mean of the differences, estimate-actual);

(2) Mean absolute difference;

(3) The standard error of estimate.

Checks were made for trends in bias related to the size of the observed values. The comparison was repeated on merchantable and butt log volume using percentage differences. Merchantable volume was defined as the inside bark volume between $h_1 =$ 0.15 m and the measurement level (h_2) with d closest to 15.0 cm. The butt log volume was taken from stump level at 0.15 m up to 6 m.

All these measures of accuracy and precision were recalculated after both data sets had been randomly divided into two equal groups, the models had been fitted to one group, and the resulting taper equations had been applied to the other. Although the resulting estimates of accuracy and precision are theoretically more accurate there were no differences in the trends between models (Table 3).

TABLE 3—Accuracy of two models estimated from all Data Set II and from a random subset of Data Set II

	Applied to estimation data		Applied to independent subset	
	Terms 1 . . . 5	Terms 1, 2, 5, 16	Terms 1 . . . 5	Terms 1, 2, 5, 16
Diameter estimation				
Mean bias (cm)	0.08	0.04	0.06	0.0
Mean absolute bias (cm)	0.75	0.70	0.71	0.67
Standard error (cm)	0.99	0.95	0.96	0.91
Butt log volume estimation				
Mean bias (dm^3)	–1.9	–1.1	–2.4	–1.5
Percentage bias in total	–1.2	–0.7	–1.5	–1.0
Standard error (dm^3)	9.5	9.3	8.4	8.4
Merchantable volume estimation				
Mean bias (dm^3)	–1.8	–1.5	–2.2	–1.8
Percentage bias in total	–1.0	–0.8	–1.3	–1.1
Standard error (dm^3)	11.3	11.4	9.7	9.9

RESULTS

The problems of the five-term, fifth-order model are shown graphically in Fig. 1 (a). This figure shows approximate confidence intervals on the mean bias in estimating diameter, calculated over 5% height classes, for each model. The distinct trends in diameter estimation error with proportion of tree height result from an inappropriate model. This curve is followed closely in shape, but with less precision, when the bias is examined from the same model fitted to the small sample of Data Set II (Fig. 1 (c)). However, the best four- and five-term models which include a higher power term show a definite improvement (Fig. 1 (b) and (d)). This is reflected in Table 3 as nearly all measures of accuracy and precision improve when the model contains a high power term.

Figures 2 and 3 show predicted stem profiles overlaid by the basic data of Set II. The profiles are scaled as proportion of height over proportion of breast-height-diameter-over-bark and drawn for the average volume of the data set. The model with Terms 1, 2, 5, and 16 clearly estimates stem taper in a more descriptive and less arbitrary fashion than the five-term, fifth-order equation.

The coefficients of the best equations are as follows:

Data Set I	Volume	$a_1 = -10.434$	$a_2 = 1.835$	$a_3 = 1.147$		
	Taper	$b_1 = 0.122$	$b_2 = 6.933$	$b_3 = -8.468$		
		$b_4 = 3.429$	$b_5 = 0.0$	$b_6 = 1.244$	$P = 20$	
Data Set II	Volume	$a_1 = -10.322$	$a_2 = 1.767$	$a_3 = 1.200$		
	Taper	$b_1 = 0.888$	$b_2 = 1.917$	$b_3 = 0.0$		
		$b_4 = 0.0$	$b_5 = -0.869$	$b_6 = 1.050$	$P = 16$	

DISCUSSION

If the sole purpose of a taper function were to predict diameters there exist a number of models (Max & Burkhart 1976; Ormerod 1973) which have been shown to be more accurate than the compatible five-term, fifth-order polynomial. Modifying the model, by including a high power term, makes a considerable improvement in both diameter and stem volume prediction.

However, this modified polynomial still shows a slight tendency to produce greater bias in diameter estimation with increasing proportion of tree height (Fig. 1 (b), (d)). As Cao *et al.* (1980) noted "some precision on estimating diameters is apparently sacrificed to ensure the taper equation is compatible" for compatibility results in the diameter estimates having an underlying dependence on the sectional volume method and formulae used to compute the actual volume of sectionally measured trees. Although the New Zealand Forest Service method for estimating actual volume has been shown to be quite accurate (Goulding 1979), the error in the "true" volume will always have an effect on diameter prediction by a compatible taper function.

The bias shown by the modified model is small, often negligible, and must be weighed against the simplicity and flexibility of this solution. Fitting and parameter estimation are accomplished using standard linear regression procedures, and the com-

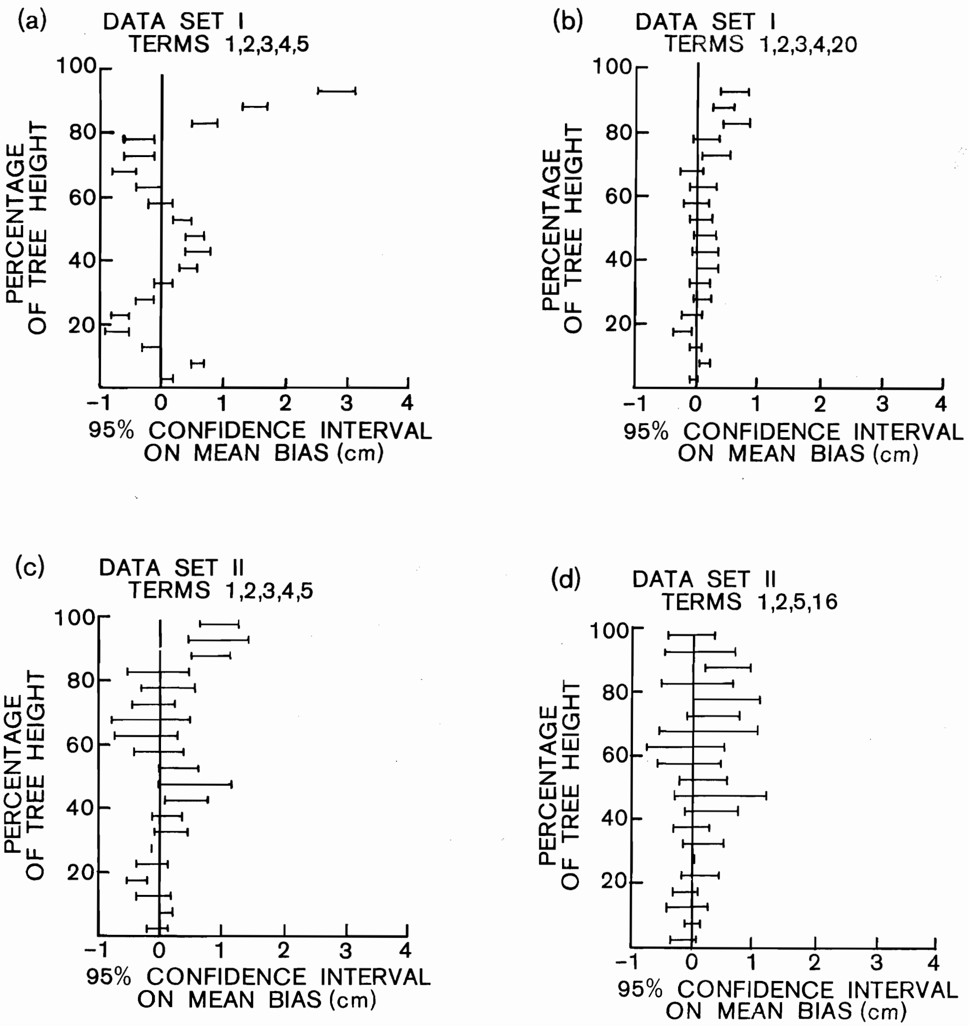

FIG. 1—Trends in diameter estimation bias.

plexities of using grafted submodels and non-linear methods (Demaerschalk & Kozak 1977; Max & Burkhart 1976) are avoided. No problems have been encountered in implementation and use of the compatible polynomial taper functions as they can be derived around existing stem volume functions where their compatibility ensures ready acceptance. Inclusion in systems of measurement is relatively efficient as the function can be integrated analytically for volume calculation and easily solved for stem length.

Thus, in balance, the modified polynomial taper function appears a viable and practical method for estimating tree diameter and volume and is a definite improvement on the five-term, fifth-order polynomial.

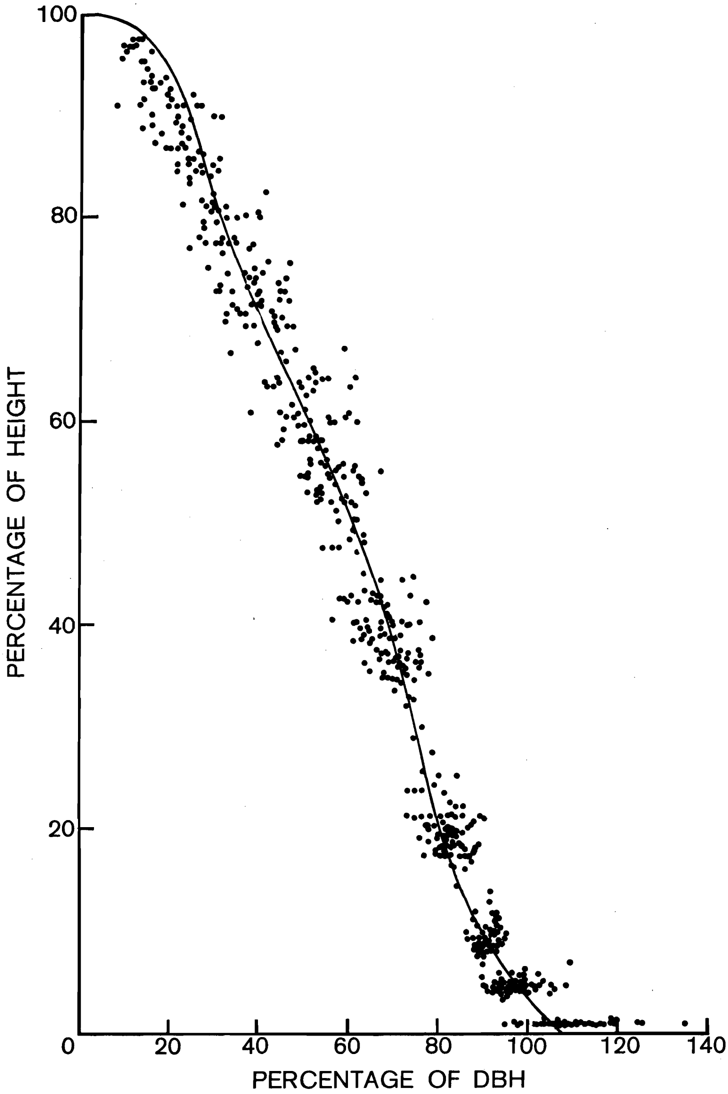

FIG. 2—Data set II. Terms 1, 2, 3, 4, and 5.

REFERENCES

BRUCE, D.; CURTIS, R. O.; VANCOEVERING, C. 1968: Development of a system of taper and volume tables for red alder. **Forest Science 14(3):** 339–50.

CAO, Q. V.; BURKHART, H. E.; MAX, T. A. 1980: Evaluation of two methods for cubic-volume prediction of loblolly pine to any merchantable limit. **Forest Science 26(1):** 71–80.

DEMAERSCHALK, J. P. 1971: Taper equations can be converted to volume equations and point sampling factors. **Forestry Chronicle 47(6):** 352–4.

——— 1972: Converting volume equations to compatible taper equations. **Forest Science 18(3):** 241–5.

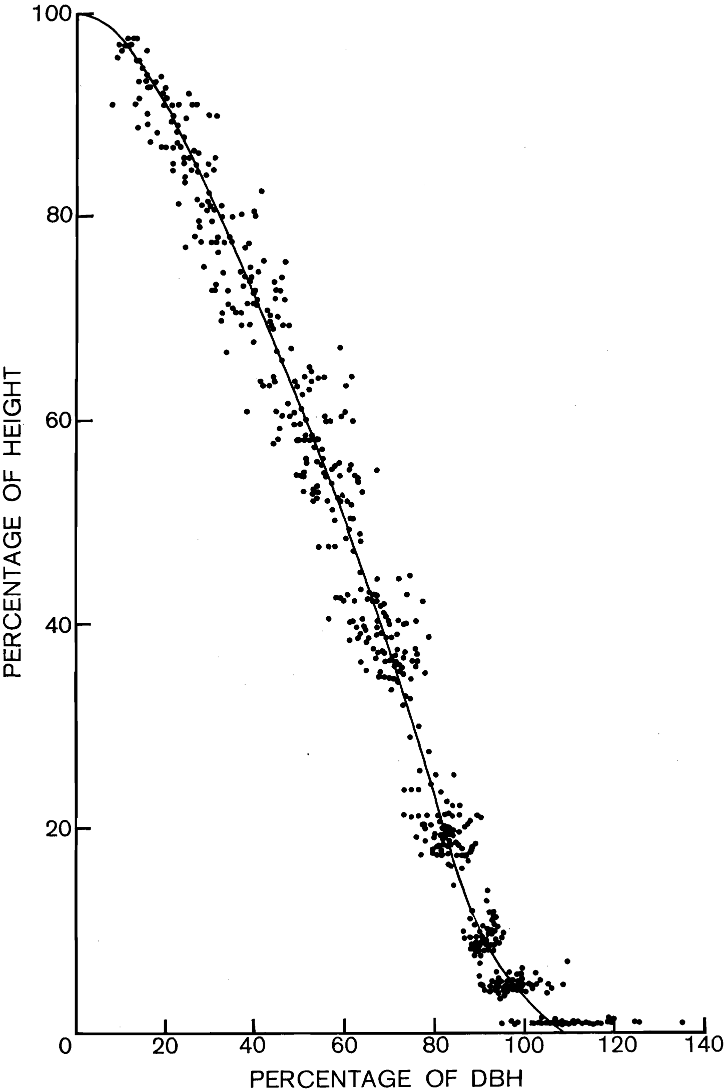

FIG. 3—Data set II. Terms 1, 2, 5, and 16.

——— 1973: Integrated systems for the estimation of tree taper and volume. **Canadian Journal of Forest Research 3(1):** 90-4.

DEMAERSCHALK, J. P.; KOZAK, A. 1977: The whole-bole system: a conditional dual-equation system for precise prediction of tree profiles. **Canadian Journal of Forest Research 7:** 488–97.

FINNEY, D. J. 1941: On the distribution of a variate whose logarithm is normally distributed. **Journal of the Royal Statistical Society (supplement) 7(2):** 155–61.

FRIES, J.; MATERN, B. 1966: On the use of multivariate methods for the construction of tree taper curves. 2nd Conference of IUFRO Advisory Group of Forest Statisticians, Stockholm. **Royal College of Forestry, Department of Forest Biometry, Research Notes No. 9.**

GOULDING, C. J. 1979: Cubic spline curves and calculation of volume of sectionally measured trees. **New Zealand Journal of Forestry Science 9:** 89–99.

GOULDING, C. J.; MURRAY, J. C. 1976: Polynomial taper equations that are compatible with tree volume equations. **New Zealand Journal of Forestry Science 5:** 313–22.

KATZ, A.; DUNNINGHAM, A. G.; GORDON, A.: A compatible volume and taper equation for New Zealand **Pinus radiata** D. Don grown under the direct sawlog regime (in prep.).

MAX, T. A.; BURKHART, H. E. 1976: Segmented polynomial regression applied to taper equations. **Forest Science 22(3):** 283–9.

MUNRO, D. D.; DEMAERSCHALK, J. P. 1974: Taper-based versus volume-based compatible estimating systems. **Forestry Chronicle 50(5):** 197–9.

ORMEROD, D. W. 1973: A simple bole model. **Forestry Chronicle 49:** 136–8.

APPENDIX I

A BRIEF THEORY OF THE METHOD

Total stem volume is given by

$$V_s = K \int_0^H d^2 \, dh \qquad [1]$$

As
$$d^2 = \frac{V}{KH} (b_1 z + b_2 z^2 + \ldots b_n z^n) \qquad [2]$$

where
$$z = \frac{H - h}{H}$$

then
$$V_s = \frac{V}{H} \left(\left(\frac{b_1}{2} + \frac{b_2}{3} + \ldots \frac{b_n}{n+1}\right) H\right)$$

and *if*
$$\sum_{i=1}^{n} \frac{b_i}{(i+1)} = 1 \qquad [3]$$

then $V_s = V$ and [2] is compatible. Rearranging [2] using [3] in the form

$$b_1 = 2\left(1 - \sum_{i=2}^{n} \frac{b_i}{(i+1)}\right) \qquad [4]$$

leads to the linear equation

$$\frac{d^2 KH}{V} - 2z = \sum_{i=2}^{n} b'_i \left((i+1) z^i - 2z\right)$$

where
$$b'_i = \frac{b_i}{i+1} \qquad [5]$$

the b' coefficients can be estimated by regression methods and the coefficients of [2] calculated using [4] and [5].

From [1] and [2] volume from the top of the tree to point h is given by

$$V_t = V \sum_{i=1}^{n} \left(\frac{b_i}{i+1} z^{i+1}\right)$$

Note: The constraint [3] is a joint conditioning of all the coefficients and so [4] can be rewritten to eliminate any term in the fitting with no effect on the solution.