
Trigonometric Variable-Form Taper Equations for Australian Eucalypts

Huiquan Bi

ABSTRACT. This article introduces a new variable-form taper model that is stable in specification yet flexible in its ability to fit data for species and trees with different stem forms. The base function is constructed from trigonometric volume-ratio equations following the geometry of a tree stem. The specification for the exponent includes variables for depicting changes in stem form along a stem and variables for taking into account differences in stem form among trees of different sizes. This model is fitted to data from 25 species of Australian eucalypts and is compared with Kozak's taper model to demonstrate its characteristics: stability in specification, flexibility in fitting data for species and trees with varying stem forms and accurate predictions of taper and merchantable height. *FOR. SCI.* 46(3):397–409.

Additional Key Words: Power function, trigonometric base function, trigonometric exponent function, stem form.

TAPER EQUATIONS THAT CAN ACCURATELY PREDICT the diameter at any point on a stem from diameter at breast height and total tree height, two readily measured variables that characterize the basic dimensions of tree size, have long been the subject of research of many forest scientists. The construction of such equations has taken a number of approaches as demonstrated by the vast amount of literature in this area (e.g., Fries and Matern 1966, Goulding and Murray 1975, Max and Burkhart 1976, Clutter 1980, Amidon 1984, Biging 1984, Newnham 1988, 1992, Kozak 1988, 1997, Candy 1989, Thomas and Parresol 1991, Flewelling and Raynes 1993, Bailey 1994, Gordon et al. 1995, Fang and Bailey 1999). The weaknesses shared by many taper equations are (1) the existence of a large degree of local bias in diameter prediction over some portions of the stem, particularly the lower and/or upper stem, despite a low global bias, and (2) the failure to take into account differences in stem form between trees. A recent approach introduced by Newnham (1988, 1992) and Kozak (1988, 1997) uses a single continuous function as the base with an exponent that changes along the stem to describe the continuous change of stem form from ground to tip. Such a power function eliminates the necessity of developing segmented taper functions for different portions of the stem in order to reduce local bias. The

exponent can also be specified to change with diameter at breast height and tree height to account for the differences in stem form between trees. In comparison with other approaches such as the whole bole system of Demaerschalk and Kozak (1977) and the segmented polynomial function of Max and Burkhart (1976), this approach has the least degree of local bias and greater precision in taper predictions (Newnham 1988, 1992, Kozak 1988, Perez et al. 1990, Kozak and Smith 1993, Muhairwe 1999).

The methods of modeling the exponent have been to select a subset of variables which fits the data well enough from a larger number of candidates that are assumed to exert an influence on the exponent in multiple linear least squares regression (Newnham 1988, 1992, Kozak 1988, 1997, Perez et al. 1990, Bi and Turner 1994, Muhairwe 1999). These candidate variables, often large in number, usually include various transformations of relative height, overbark diameter at breast height, and total tree height. In the extreme cases, dozens of combinations of these transformed variables were reportedly included (Newnham 1992, Kozak 1997). These variables are usually highly intercorrelated (Kozak 1997). Such high multicollinearity leads to much inflated estimates of the standard errors of parameters and also causes the least-squares estimates of parameters to be unstable (Myers 1990,

The author is senior research scientist, Forest Research and Development Division, State Forests of NSW, P.O. Box 100, Beecroft, NSW 2119 Australia—Phone: 61 2 9872 0168; Fax: 61 2 9871 6941; E-mail: huiquanb@sf.nsw.gov.au.

Acknowledgements: I am indebted to forestry staff in both NSW and Victoria for collecting the data in the field. I thank Mr. Shimin Cai for technical assistance. Dr. Phil West, Dr. Charles Muhairwe, Dr. Peter Cornish, Mr. Vic Jurskis and Mr. Jack Simpson provided helpful comments.

Manuscript received April 19, 1999. Accepted December 16, 1999.

Copyright © 2000 by the Society of American Foresters

Fox 1991). Consequently, it is difficult to separate the individual effects of the variables for discriminating redundant ones in the model. A small change in the data can substantially alter the least-squares coefficients (e.g., Kozak 1997) and may result in a different subset of variables being selected. In addition, the variable selection itself may not be consistent in the sense that there is no guarantee that the same subset will be selected as the sample size in the data increases (Nishii 1984, Breiman 1992). The lack of consistency in variable selection compounded with high multicollinearity among the candidate predictors have made the variable-form taper models overwhelmingly data driven, and indeed very variable. As species or data sets change, different models often emerge (Newham 1988, 1992, Kozak 1988, 1997, Perez et al. 1990, Bi and Turner 1994, Muhairwe 1999). This variability has translated the variable-form taper models, particularly those of Newham (1988, 1992), into a general approach rather than a model with stable designed features apart from the specification of a general power function.

This article introduces a new variable-form taper model that is stable in specification yet flexible in its ability to fit data for species and trees with different stem forms. The base function is constructed from trigonometric volume-ratio equations. The specification for the exponent is based on Fourier transformation. It includes variables for depicting changes in stem form along the stem and variables for taking into account differences in stem form between trees of different sizes. This model is fitted, without resorting to subset selection, to data from 25 species of Australian eucalypts with a range of stem forms and is compared with Kozak's (1988) model to demonstrate its accuracy and flexibility.

Notation

The following notation will be used throughout the remainder of this article. Other notation specific to a particular equation will be listed with the equation.

- TH = total tree height in m;
 H = height above ground, $0 < H \leq TH$, in m;
 h = H/TH , relative height;
 b = $1.3/TH$, relative breast height;
 $DBHOB$ = diameter at breast height over bark;
 $DBHUB$ = diameter at breast height under bark;
 DUB = diameter under bark at relative height h ;
 d = $DUB/DBHUB$, relative diameter;
 K = a variable exponent that is a function of h and $DBHOB$ and TH .

$$B = \frac{\ln \sin\left(\frac{\pi}{2} h\right)}{\ln \sin\left(\frac{\pi}{2} b\right)}, \text{ the base function derived in this article;}$$

$$B_K = \frac{1 - \sqrt{h}}{1 - \sqrt{p}}, \text{ the base function of Kozak (1988);}$$

p = a species-specific relative height of the inflection point.

Data

The data set for this work included 5,739 trees from more than 25 species of *Eucalyptus* and *Corymbia*. The sample size among the species ranged from more than 1000 to less than 20 trees. For major commercial species, the sample size was more than 100. For a number of minor species including *E. scias*, *E. considianiana*, *E. maideni*, *E. rubida*, *E. angophoroides*, *E. bosistoana*, the sample size was less than 20, so these species were grouped together and coded as *Eucalyptus* spp. The diameter and height distributions of these sample trees were summarized for all species using Tukey's (1977) boxplots (Figure 1). The geographical areas and forest types where the samples were taken and the field measurements of stem taper were described in detail by Bi and Hamilton (1998).

The scatterplot of relative diameter against relative height was examined visually for each species to detect possible anomalies in the data. After spending much time searching and checking the original data sheets of some individual trees to verify and correct possible data errors one by one, a systematic approach of detecting abnormal data points was adopted to increase efficiency. Assuming identically distributed Gaussian errors, a nonparametric taper curve was fitted for each species using local regression, *loess* (Cleveland 1993). A locally quadratic fitting with a smoothing parameter of 0.25 was used for all species after some iterative fitting and visual examination of the smoothed taper curves and the data. The residuals of the nonparametric curve were divided into ten even intervals of relative height. The distribution of residuals within each interval was examined, and two particular values were calculated for discriminating extreme data points: the lower quartile minus twice the interquartile range and the upper quartile plus twice the interquartile range. Data points outside the range of these two values are very extreme points because this range is even greater than that used in the boxplots of Tukey (1977) to show extreme data points in an univariate distribution. The number of these extreme data points accounted for between 0.89% and 4.04% of the total number of taper measurements among the species, and about 2.05% for all species combined.

Field notes on taper measurements of these points were extracted from the database for detailed examination. Apart from a small number of data errors, most of these data points were from deformed stem sections due to the presence of fire scars, large knots or bulges, other physical damages, partial death of the stem and coppice, or epicormic growth, etc., reflecting partly the frequent fire disturbance in the regrowth forests where many of the samples were taken (Bi and Jurskis 1996). Since the taper equations were not intended for stem sections with deformities, these data points were excluded from further analysis.

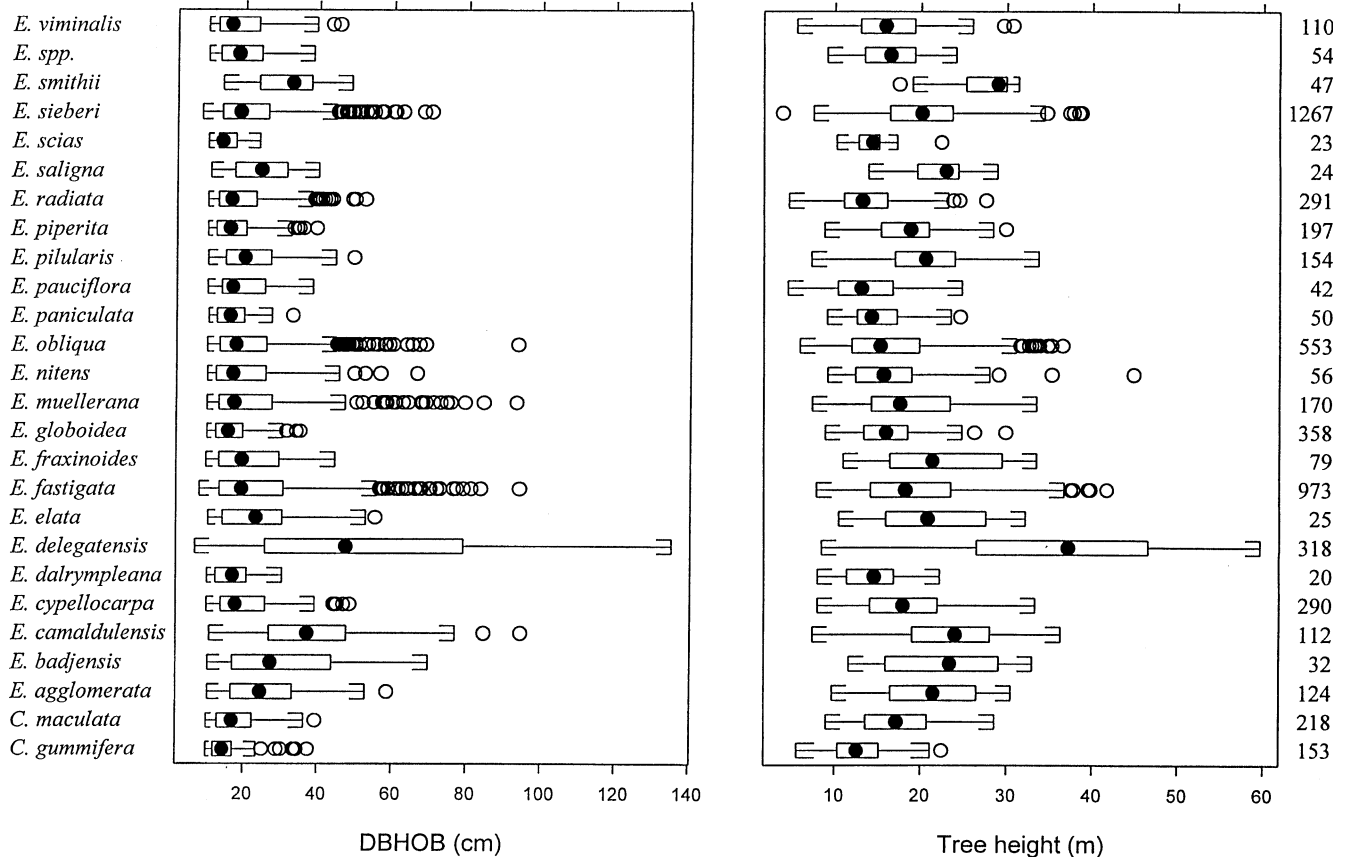


Figure 1. Boxplots of diameter overbark at breast height (DBHOB) and total height of sample trees of 25 species of *Eucalyptus* and *Corymbia*, plus a sample labelled *E. spp.* being a mixture of eucalypt species each of sample size less than 20. Numbers on the righthand side indicate the number of sample trees.

Model Derivation and Estimation

Considering a tree bole with a given relative stem profile (Figure 2), the volume of a stem section from ground to any specified top height or diameter limit can be expressed as a volume ratio (i.e., a percentage of the total stem volume). Volume ratio increases monotonically from 0 to 1 as relative height increases from 0 to 1 and as relative diameter decreases from a value greater than 1 to 0. The volume ratio to any top height limit (R_h) can be expressed as a function of relative height h , $R_h = f_1(h)$. Similarly, the volume ratio to any top end diameter limit (R_d) can be expressed as a function of relative diameter d , $R_d = f_2(d)$. At any given point on the stem profile, a volume ratio can be obtained from either R_h or R_d , and the two values must be the same such that $R_d = R_h$, and so $f_2(d) = f_1(h)$. The inverse of this relationship, $d = f_2^{-1}(f_1(h))$, provides a taper function for the stem profile.

The model developed by Bi (1999) for predicting the volume ratio to any specified top height limit for these species takes the following form:

$$R_h = \sin^{\alpha_1} \left(\frac{\pi}{2} h \right) \sin^{\alpha_2} \left(\frac{\pi}{2} h^2 \right) \sin^{\alpha_3} \left(\frac{\pi}{2} h^3 \right) \sin^{\alpha_4} \left(\frac{\pi}{2} h^4 \right) \quad (1)$$

where R_h is the ratio between stem volume from ground to relative height h and the total stem volume, α_1 to α_4 are

coefficients specific to each species. Arguments for trigonometric functions are expressed in radians. The predicted volume ratio at breast height, R_b , differs among trees because the relative breast height, b , changes with total tree height. The volume ratio to any diameter limit can be predicted from relative diameter using the following model form:

$$R_d = R_b^{d^{\frac{1}{K}}} \quad (2)$$

Equation (2) is related to Equation (1) through R_b , which is the predicted volume ratio at relative breast height from Equation (1). Another exponential function similar to Equation (2), but using the base of the natural logarithm in place of R_b , has been shown to provide accurate predictions of volume-ratio to any diameter limit for *Pinus taeda* (Van Deusen et al. 1981). For a given stem profile, the values of R_h and R_d should be the same for any given volume ratio such that

$$R_b^{d^{\frac{1}{K}}} = R_h \quad (3)$$

Rearranging Equation (3), d becomes the dependent variable of the power function

$$d = \left(\frac{\ln R_h}{\ln R_b} \right)^K \quad (4)$$

where \ln represents natural logarithm. Taking K as a variable exponent that changes with h and tree size, Equation (4)

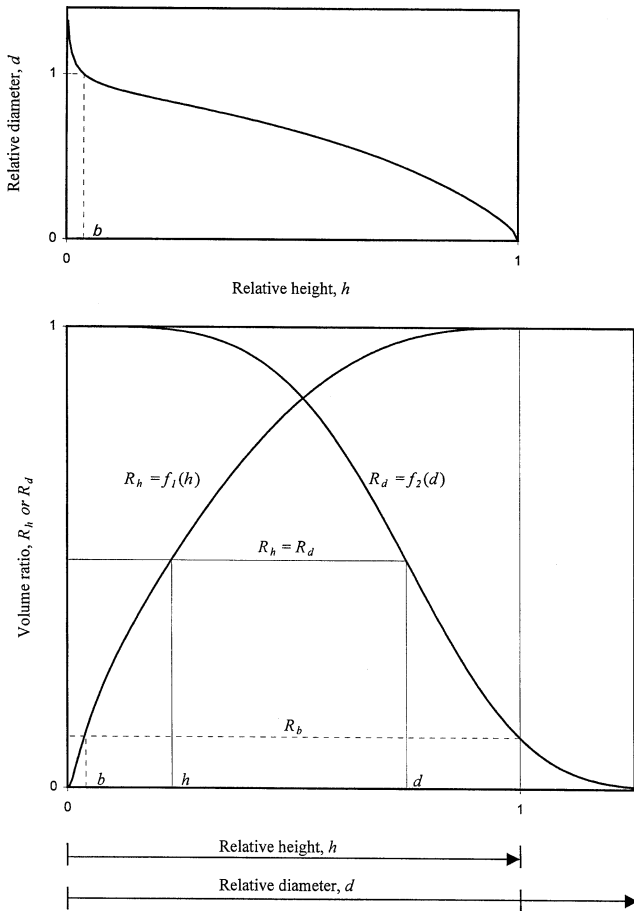


Figure 2. Diagram showing the derivation of the base function of the trigonometric variable form taper model (see Notation and text).

becomes a taper function that is consistent with Equation (1) since they are mathematically related. However, this model is not the most parsimonious because R_h and R_b are predicted already by four parameters in Equation (1). To improve model parsimony, R_h and R_b are replaced by

$$\sin\left(\frac{\pi}{2}h\right) \text{ and } \sin\left(\frac{\pi}{2}b\right),$$

respectively, such that

$$d = \left(\frac{\ln \sin\left(\frac{\pi}{2}h\right)}{\ln \sin\left(\frac{\pi}{2}b\right)} \right)^K \quad (5)$$

In comparison with the variable-form taper models of Newnham (1988) and Kozak (1988), the characteristics of Equation (5) lie in the base of the power function, which is the ratio between the logarithms of two trigonometric functions. The denominator is a negative constant for a given tree, and the numerator lies between $-\infty$ and 0 as relative height h varies from 0 to 1. The ratio tends to ∞ as h tends to 0 and equals 0 when $h = 1$. This monotonic base function also

differs among trees because the relative breast height, b , varies with total tree height. In addition, this trigonometric variable-form taper function is conditioned to pass through underbark diameter at breast height since the base function equals 1 when $h = b$ (i.e., relative height equal to relative breast height). Bi and Turner (1994) found that a sine function was a better base function than that used by Newnham (1988, 1992) for describing average stem profiles using a power function possibly because the trigonometric function has inherent inflection points and is hence more flexible.

For simplicity, the taper model specified in Equation (5) can be written as a general power function:

$$d = B^K \quad (6)$$

where B represents the base, a monotonic function of relative height h . Because no biological or statistical theory can be relied on to identify the “true” model for K , an empirical function purely for the purpose of prediction will have to come from exploratory model building. A preliminary analysis regressing $\ln d$ against $\ln B$ without the intercept term (i.e., effectively taking K as a constant) resulted in residuals showing cyclic patterns for every species. The common approach of overcoming such patterned residuals has been the use of polynomial functions of h in variable-form taper functions (Newnham 1988, 1992, Kozak 1988, 1997, Perez et al. 1990, Muhairwe 1999), and more so in other nonsegmented taper functions which often involve many terms and in some cases a very high power term (e.g., Bruce et al. 1968, Gordon 1983). However, the theory of Fourier analysis suggests that any cyclic function can be decomposed into a weighted sum of mutually orthogonal sine and cosine pairs (Mathews 1987). Often a few large Fourier weights are able to provide a compact structural summary of the observed periodicity in empirical data. Analyses using two and three pairs of sine and cosine with different frequencies showed almost no cyclic patterns in the residuals. After further exploratory analysis of various equation forms and consideration of model parsimony, the variable exponent K was specified as a function of relative height h , $DBHOB$, and tree height TH for all species as follows:

$$K = a_1 + a_2 \sin\left(\frac{\pi}{2}h\right) + a_3 \cos\left(\frac{3\pi}{2}h\right) + a_4 \sin\left(\frac{\pi}{2}h\right) / h + a_5 DBHOB + a_6 h \sqrt{DBHOB} + a_7 h \sqrt{TH} \quad (7)$$

where a_1 to a_7 are parameters. The first three trigonometric variables were used for depicting changes in stem form along the stem, and the last three variables were used for taking into account differences in stem form between trees of different sizes.

Substituting Equation (7) into Equation (6) and taking logarithm on both sides, $\ln d$ becomes a linear function of variables that are multiplications of the logarithm of base B and all variables in Equation (7) as shown in Equation (8) below.

$$\ln d = \left(a_1 + a_2 \sin\left(\frac{\pi}{2}h\right) + a_3 \cos\left(\frac{3\pi}{2}h\right) + a_4 \sin\left(\frac{\pi}{2}h\right) / h + a_5 DBHOB + a_6 h \sqrt{DBHOB} + a_7 h \sqrt{TH} \right) \ln B \quad (8)$$

Equation (8) was fitted to data using least squares regression to obtain parameter estimates for each species. Positive autocorrelations and heteroskedasticity present in the data were not taken into account in the least squares regression. When both are present, the least squares estimates of regression coefficients remain unbiased and consistent, although no longer efficient (Myers 1990). Prediction accuracy is little affected by autocorrelation in the error term and multicollinearity among predictor variables in the equation (Kozak 1997). Even when the correlated error structure is accounted for in fitting taper models, the improvement in prediction accuracy was too small to be of practical importance (Williams and Reich 1997).

Predicting Underbark Diameter at Breast Height

The dependent variable of the trigonometric taper functions is relative diameter d (i.e., the ratio of diameter underbark at any height above ground to that at breast height). To convert any predicted relative diameter to absolute underbark stem diameter, $DBHUB$ must first be estimated. To facilitate the application of the taper equations, a set of equations was developed for these species using the following equation form

$$DBHUB = DBHOBe^{\frac{1}{b_0 + b_1 DBHOB + b_2 \ln DBHOB + b_3 \ln TH}} \quad (9)$$

where $b_0 - b_3$ are parameters. This equation form was selected from several linear and nonlinear model forms by comparing their prediction accuracy within the range of observed data and consistency in extrapolation beyond this range. The equation was linearized and parameters were estimated using least squares regression.

Validation

Prediction Accuracy

To assess the prediction accuracy of the estimated taper equations, the resampling approach taken by Bi (1999) was adopted. For each species with N sample trees, Equation (8) was fitted N times. Each time, all data points of one tree were removed from the fitting process, and predicted values of relative diameter were obtained for them using the coefficients estimated from the remaining data. The difference between the observed and predicted values was taken as the error of prediction, ϵ . The mean squared error of prediction (MSEP) was taken as a measure of prediction accuracy:

$$\begin{aligned} \text{MSEP} &= E(\epsilon^2) = E(\epsilon - E(\epsilon) + E(\epsilon))^2 \\ &= \text{Var}(\epsilon) + (E(\epsilon))^2 \end{aligned} \quad (10)$$

where $\text{Var}(\epsilon)$ and $E(\epsilon)$ were the variance and expectation of prediction error, indicating the precision and the bias of prediction respectively (Wackerly et al. 1996). The average size of prediction error is indicated by $E(|\epsilon|)$. For a normally distributed ϵ with zero mean and variance σ^2 , $E(|\epsilon|)$ equals $\sqrt{2/\pi}\sigma$ (Flury 1990, Greene 1993). The accuracy of predic-

tion was evaluated over ten relative height intervals for each species. Within each interval, $E(\epsilon)$, $E(|\epsilon|)$ and the 2.5th and the 97.5th percentiles of prediction error were calculated and plotted to show the bias and precision of prediction.

For eight species with more than 200 sample trees (Figure 1), the taper equations were further evaluated to see if they adequately reflected the differences in stem shape associated with changes in tree size that was found with some species (Forslund 1991, Allen 1993, Bi and Turner 1994, Muhairwe 1994). For each of these species, the data were divided into eight size classes according to $DBHOB$ and tree height in a manner similar to that of Flewelling and Raynes (1993). A nonparametric height-diameter curve was fitted at first using local regression, *loess* (Cleveland 1993), with a locally quadratic fitting and a smoothing parameter of 0.75. The height-diameter curve divided the data points of each species into two parts (Figure 3). Points above the curve had greater than average height at a given diameter, while points below the curve showed the contrary. Then the data were further divided into four diameter classes by using the median, the upper and lower quartiles shown in Figure 1. Thus the data of each species was divided into eight size classes in the height-diameter space (Figure 3). For each group, the accuracy of prediction was evaluated over ten relative height intervals by calculating $E(\epsilon)$, $E(|\epsilon|)$ and the 2.5th and the 97.5th percentiles of prediction error.

Model Comparison

Prediction accuracy of the trigonometric taper function for each species was compared with that of the taper model of Kozak (1988):

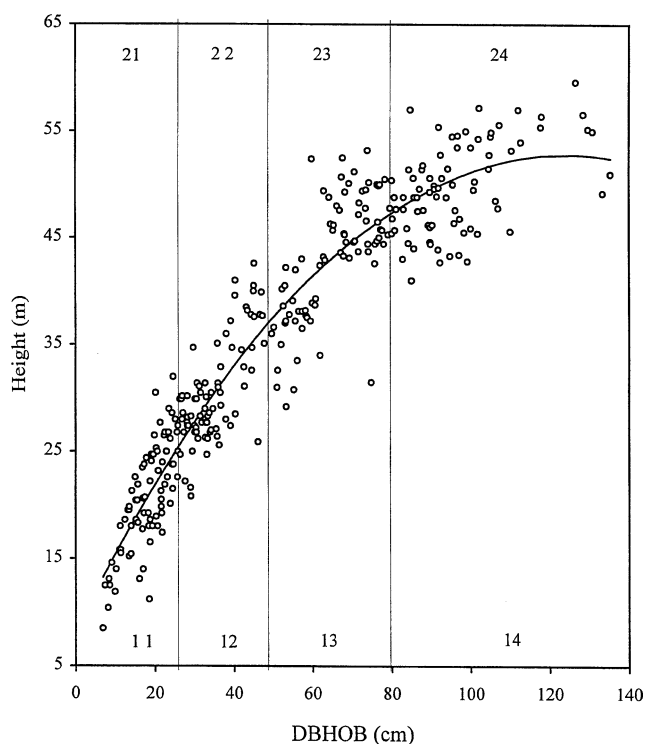


Figure 3. Height plotted against $DBHOB$ for *E. delegatensis* as an example of the division of data into 8 size groups in the height-diameter space for the eight species, each with more than 200 sample trees.

$$\ln DUB = a_0 + a_1 \ln DBHOB + a_2 DBHOB + \ln B_K [b_1 h^2 + b_2 \ln(h + 0.001) + b_3 h^{0.5} + b_4 e^h + b_5 (DBHOB / TH)] \quad (11)$$

This model was fitted using least squares regression for each species. Errors of prediction were obtained using the same resampling approach as described above. However, instead of taking p as the species-specific relative height of inflection point in the base function B_K as defined by Kozak (1988), the relative height of centroid was used, which ranged from 0.206 to 0.258 among the 25 species (Bi 1999). Variations in the value of p between 0.15 and 0.35 were shown to have little effect on the prediction accuracy of the variable-exponent model (Perez et al. 1990, LeMay et al. 1993).

Prediction accuracy of the two models was compared in terms of predictions of both relative diameter and underbark stem diameter to see how each performed when used for describing stem form and predicting underbark stem diameter. The dependent variable of Kozak's (1988) model is underbark diameter, DUB , not relative diameter, d , as in the trigonometric taper function. Therefore, the predicted values of DUB from Kozak's (1988) model were divided by $DBHUB$ to obtain predicted relative diameter, and the predicted values of relative diameter from the trigonometric taper functions were multiplied by $DBHUB$ predicted from Equation (9) to obtain predicted DUB . Values of the mean squared error of prediction ($MSEP$) and bias in the prediction of both relative and absolute underbark diameter were compared between the two models across species, species by relative height classes for all species, and species by size groups and by relative height classes for the eight species with large samples. Since it was too lengthy to report, this detailed comparison was summarized by two ratios for each species, $R_{MSEP} = MSEP_K / MSEP_B$ and $R_{bias} = |E(\epsilon)_K| / |E(\epsilon)_B|$, where $MSEP_K$ is the mean squared error of prediction of Kozak's (1988) model and $MSEP_B$ is that of the trigonometric taper function, $|E(\epsilon)_K|$ is the absolute value of local bias over 10 relative height intervals for Kozak's (1988) model and $|E(\epsilon)_B|$ is that of the trigonometric taper function. The two ratios indicated the overall predictive performance and the average magnitude of local bias of the trigonometric taper function relative to Kozak's (1988) model. For each of the 8 species with large samples, an additional value of R_{bias} was obtained from a total of 80 values of local bias because the data were divided by 8 size groups and by 10 relative height intervals.

Accuracy of Merchantable Height Estimation

Apart from taper prediction, taper functions are also used for estimating the merchantable height of a tree given a specified top end diameter (Kozak and Smith 1993). Since the trigonometric taper functions cannot be transposed to give explicit mathematical solutions of relative height for a given relative diameter, numerical solutions have to be obtained through iterations. Using the trigonometric taper functions and $DBHUB$ predicted from Equation (9), the predicted merchantable height for a range of top limit underbark diameter from 6 cm to 30 cm with an even interval of 2 cm was obtained through numerical iterations for all trees with predicted $DBHUB$ greater than the specified top diameter limit. For many species, a top limit underbark diameter of 8 cm is the current specification in management. The observed merchantable height was obtained through linear interpolation for each tree using taper measurements immediately below and above the point of interpolation. Linear interpolation was used because taper measurements above breast height were taken at 1.5 m intervals for most trees and at 3 m intervals only for a small number of samples of mainly *E. delegatensis*. Quadratic interpolation would improve the accuracy of merchantable height estimation very little. The difference between the observed and predicted values of merchantable height was taken as the error of prediction. For each species, the mean, the lower and upper quartile, the 2.5th and the 97.5th percentiles of prediction error for each specified top end diameter were calculated and plotted to show the bias and precision of prediction.

Results

The complete expression of the trigonometric taper model is the same for all species [see Equation (12) below].

The estimated parameters are shown in Table 1. Three stem profiles were made for a small, average, and large tree using the taper function for *E. delegatensis* as an example (Figure 4). These stem profiles illustrated the changes in stem form along the stem and also differences in stem form among trees of different sizes as depicted by the trigonometric taper function. Predictions of underbark diameter from $DBHOB$ and tree height can be obtained by multiplying the predicted relative diameter with underbark diameter at breast height predicted from Equation (9) and parameters in Table 2.

There was little local bias across relative height classes in the predictions of relative diameter for all species [Figure (5)]. For *C. maculata*, *E. badjensis*, *E. delegatensis*, *E. fastigata*, *E. fraxinoidies*, *E. pilularis*, and *E. piperita*,

$$d = \left(\frac{\ln \sin\left(\frac{\pi}{2} h\right)}{\ln \sin\left(\frac{\pi}{2} b\right)} \right)^{a_1 + a_2 \sin\left(\frac{\pi}{2} h\right) + a_3 \cos\left(\frac{3\pi}{2} h\right) + a_4 \sin\left(\frac{\pi}{2} h\right) / h + a_5 DBHOB + a_6 h \sqrt{DBHOB} + a_7 h \sqrt{TH}} \quad (12)$$

Table 1. Parameters of the trigonometric variable-form taper model [Equation (12)] for the 25 species. The fitting statistics for the regression involving log transformed relative diameter are R^2 and root mean squared error (RMSE). I^2 and IRMSE represent the corresponding fit statistics for the regression between observed and predicted relative diameter obtained through resampling.

Species	a_1	a_2	a_3	a_4	a_5	a_6	a_7	R^2	RMSE	I^2	IRMSE
<i>C. gummifera</i>	1.5745	-0.4691	-0.0995	-0.8095	0.0028	0.0468	-0.0489	0.92	0.132	0.90	0.100
<i>C. maculata</i>	0.9477	-0.1125	-0.0050	-0.4595	-0.0018	0.0731	-0.0632	0.96	0.103	0.97	0.053
<i>E. agglomerata</i>	1.9268	-0.4932	-0.0437	-0.9866	0.0020	0.0348	-0.0617	0.95	0.131	0.94	0.080
<i>E. badjensis</i>	1.6678	-0.1985	-0.0214	-0.9098	-0.0003	0.0603	-0.0872	0.99	0.079	0.98	0.043
<i>E. camaldulensis</i>	0.8364	0.2216	-0.0031	-0.4110	0.0015	0.0547	-0.1332	0.92	0.159	0.94	0.077
<i>E. cypellocarpa</i>	0.4361	0.1186	0.0552	-0.1528	-0.0011	0.0641	-0.0632	0.95	0.113	0.94	0.070
<i>E. dalrympleana</i>	0.8266	-0.3792	-0.0470	-0.3323	-0.0022	0.0342	0.0367	0.95	0.128	0.95	0.069
<i>E. delegatensis</i>	1.2992	-0.0564	0.0096	-0.7248	0.0017	-0.0099	-0.0254	0.96	0.107	0.96	0.054
<i>E. elata</i>	1.2717	0.2267	0.0515	-0.6965	-0.0003	0.0955	-0.1777	0.97	0.103	0.97	0.053
<i>E. fastigata</i>	0.7965	0.0513	0.0312	-0.3703	-0.0011	0.0737	-0.0874	0.97	0.099	0.97	0.052
<i>E. fraxinoides</i>	1.1467	-0.0718	0.0048	-0.6167	0.0015	0.0201	-0.0424	0.97	0.089	0.97	0.046
<i>E. globoidea</i>	1.0495	-0.2678	-0.0062	-0.4681	-0.0025	0.1120	-0.0905	0.95	0.108	0.93	0.076
<i>E. muellerana</i>	1.9618	-0.5850	-0.0783	-0.9549	-0.0004	0.0826	-0.0953	0.95	0.131	0.93	0.088
<i>E. nitens</i>	0.9654	0.0889	0.0329	-0.5134	0.0004	0.0496	-0.0833	0.95	0.130	0.97	0.055
<i>E. obliqua</i>	0.5744	0.2210	0.0748	-0.2838	0.0001	0.0554	-0.0863	0.96	0.108	0.95	0.064
<i>E. paniculata</i>	1.3308	-0.2888	-0.0334	-0.6855	-0.0008	0.1121	-0.1022	0.95	0.107	0.94	0.075
<i>E. pauciflora</i>	1.1689	0.3397	0.0372	-0.6628	0.0043	0.0675	-0.2003	0.95	0.134	0.94	0.083
<i>E. pilularis</i>	1.0546	-0.2531	-0.0168	-0.4707	-0.0020	0.0880	-0.0711	0.97	0.095	0.97	0.054
<i>E. piperita</i>	0.5564	-0.1492	0.0162	-0.1696	-0.0015	0.0696	-0.0349	0.94	0.120	0.97	0.053
<i>E. radiata</i>	0.7775	0.1525	0.0025	-0.3706	0.0004	0.0683	-0.1247	0.92	0.158	0.94	0.079
<i>E. saligna</i>	1.0391	-0.5304	-0.0298	-0.4386	-0.0017	0.0565	0.0128	0.97	0.094	0.96	0.059
<i>E. scias</i>	0.2577	0.2832	0.0692	-0.0243	-0.0090	0.1519	-0.1540	0.93	0.131	0.94	0.074
<i>E. sieberi</i>	0.8948	-0.0415	0.0349	-0.3895	-0.0004	0.0744	-0.0951	0.96	0.106	0.95	0.071
<i>E. smithii</i>	1.9891	-0.5870	-0.0256	-0.9106	-0.0020	0.0863	-0.0927	0.96	0.109	0.93	0.085
<i>E. spp.</i>	0.9834	0.1636	0.0406	-0.4887	0.0006	0.1178	-0.1874	0.94	0.136	0.91	0.097
<i>E. viminalis</i>	1.3184	-0.0296	0.0158	-0.7196	0.0012	0.0420	-0.0847	0.96	0.105	0.96	0.060

the precision of prediction was relatively high among the species as shown by the narrower confidence intervals across all relative height classes for these species. For *C. gummifera*, *E. camaldulensis*, *E. pauciflora*, and *E. radiata*, the precision was relatively low. As expected, the prediction in the section closest to the ground was generally less precise than that in other stem sections. The average size of error in relative diameter prediction was below 0.1 for all relative height classes and for all species (Figure 5).

For the eight species with sample size larger than 200 trees (Figure 1), the division of data of each species by 8 size groups and by 10 relative height classes provided more than 600 observations of local bias and precision in relative diameter prediction. Some relative height classes close to the tip of smaller size groups had less than 10 observations, and they were merged with the adjacent relative height class. Almost all values of local bias in relative diameter prediction were within ± 0.05 , with 95% of the observations within ± 0.035 , and 90% within ± 0.025 . There was not a pattern of local bias that varied systematically with tree size among the 8 species. Two species, *E. delegatensis* and *E. sieberi*, were shown as examples (Figure 6). The former had the largest size range and the latter the largest number of sample trees.

The trigonometric taper functions were generally less biased and more precise than the taper model of Kozak (1988) for predicting both relative diameter and absolute underbark diameter. On average, the mean squared error of prediction of Kozak's model was 7.4%

greater than that of the trigonometric taper function for predicting relative diameter and 14% greater for predicting absolute underbark diameter for the 26 species (Table 3). In either case, the largest value of the ratio between $MSEP_K$ and $MSEP_B$ was 50% or more. The magnitude of local bias of Kozak's model was on average 28% greater than that of the trigonometric taper equation for predicting relative diameter and 25% greater for predicting absolute underbark diameter. For the eight species with data divided into eight size groups in height-diameter space, the magnitude of local bias of Kozak's model was 44% greater than that of the trigonometric taper equation for predicting relative diameter and 19% greater for predicting absolute underbark diameter when averaged over all size groups and relative height intervals.

The bias in the estimation of merchantable height using the trigonometric taper functions was small, not more than 1 m for all top end diameters and all species (Figure 7). Precision of the estimation was the best for *E. pilularis* and *E. piperita*, and the worst for *E. camaldulensis* and *E. agglomerata* among the 25 species. As expected, the 90% confidence intervals of the error of estimation generally increased in width as top end diameter became larger. They were not symmetric for certain top end diameters of some species such as *E. cypellocarpa*. In comparison with the confidence intervals, the interquartile range where 50% of the prediction error fell was generally much narrower, mostly within 1.5 m or less for all species over the range of top limit underbark diameter from 6 cm to 30 cm.

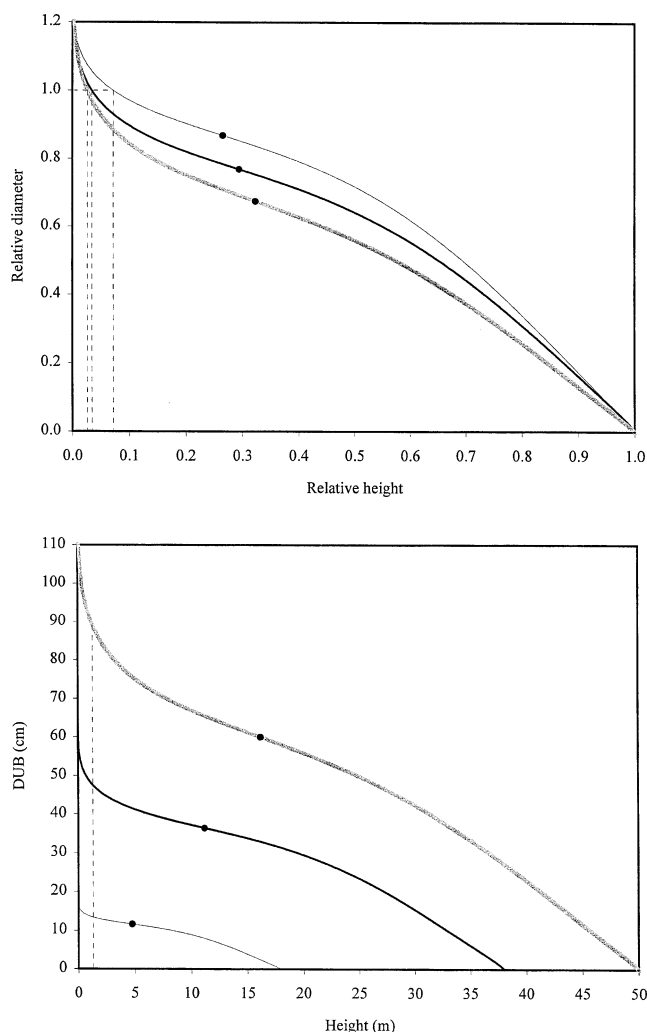


Figure 4. Relative and absolute stem profiles for a small, an average and a large tree derived from the trigonometric variable-form taper function for *E. delegatensis*. The points on the stem profiles are points of inflection obtained numerically from the taper function. The broken lines show breast height and relative breast height for the absolute and relative stem profiles of the three trees.

Discussion

The trigonometric variable-form taper model has overcome the weakness of unstable specification in the variable-form taper models introduced by Newnham (1988, 1992) and Kozak (1988, 1997). The characteristics of this model lie in its base function and the specification for the exponent K . The construction of the base function from trigonometric volume-ratio equations follows the geometry of a tree stem and is also constrained to pass through diameter at breast height. The specification for the exponent K is based on Fourier transformation. It has three trigonometric variables for depicting changes in stem form along the stem and three other variables for taking into account differences in stem form between trees of different sizes. These characteristics have provided the flexibility observed in the model in fitting data without resorting to variable selection for species and trees from a range of growth conditions and with varying stem forms. This flexibility is particularly useful for minimizing local bias and improving global prediction accuracy.

Among the 25 species, *E. delegatensis* had the largest size range and changes in stem form associated with tree size were more evident (Bi and Hamilton 1998). Such changes were adequately described by the trigonometric taper function (Figure 4). The relative stem profile predicted from the trigonometric taper equation was more paraboloidal for small trees, consistent with the findings of Forslund (1991) and Allen (1993). The basal swell and the neiloidal proportion at lower stem increased with tree size. Consequently, the relative height of the point of inflection (where the taper curve changes from neiloidal to paraboloidal) on the predicted stem profile also increased with tree size (Figure 4). Also, local bias was small and practically negligible along the stem in all size groups in the height and diameter space (Figure 3, Figure 6).

Unlike Kozak's (1988) model, which is constrained to pass through a fixed inflection point for a given species regardless of tree size, the trigonometric taper function is more flexible. It allows both the base function and the exponent to vary with tree size. So the point of inflection derived from the trigonometric taper function can vary with tree size. For the three examples in Figure 4, the relative height of the point of inflection, derived numerically from the trigonometric taper equation, ranged from 0.266 to 0.324, with a mean of 0.295. From the smallest diameter tree to the largest within the data space (Figure 3), the point of inflection on the predicted stem profile increased from a relative height of 0.268 to 0.347, almost 10% of the total tree height. This predicted increase differs from the findings of Demaerschalk and Kozak (1977) that relative height of the inflection point is relatively constant within a species regardless of tree size. Such size-related changes in relative height of the inflection point may explain why variations in the relative height of the inflection point in the base function of Kozak's (1988) model had little effect on its prediction accuracy (Perez et al. 1990, Lemay et al. 1993). The flexibility of the trigonometric taper model to depict such changes in stem form could be a major contributor to its comparatively superior predictive performance.

Despite the usefulness of taper equations in forest management, their development has been very limited for native tree species in Australia (Goodwin 1992, Muhairwe 1999). As the management of native forests becomes increasingly intensive, accurate estimates of stem volume and taper will be needed for native tree species. Recently, equations have been developed to predict total stem volume from ground to tip for 25 native tree species in southern New South Wales and Victoria (Bi and Hamilton 1998). In addition, trigonometric volume-ratio equations, which predict stem volume to any specified top height limit as a percentage of the total stem volume, were developed for these species (Bi 1999). When used together with the companion total stem volume equations, they allow direct predictions of volume of any stem section. The trigonometric variable-form taper equations form a part of a system of equations for stem volume and taper predictions for these species. They will be used primarily

Table 2. Parameters and fitting statistics for Equation (9) for predicting underbark diameter at breast height for the 25 species.

Species	b_0	b_1	b_2	b_3	RMSE	R^2
<i>C. gummifera</i>	-5.4249	-0.2918	4.4868	-1.6979	1.096	0.942
<i>C. maculata</i>	1.5514	-0.0784	-0.9343	-0.7221	0.699	0.987
<i>E. agglomerata</i>	-3.1890	-0.2951	7.7669	-6.2804	1.491	0.979
<i>E. badjensis</i>	2.6711	0.0224	2.0702	-5.8250	0.827	0.997
<i>E. camaldulensis</i>	7.3801	-0.0257	-3.4093	-0.9238	1.503	0.991
<i>E. cypellocarpa</i>	-6.3526	-0.2435	4.2712	-2.2543	0.705	0.992
<i>E. dalrympleana</i>	-1.9227	0.0486	-1.9439	-0.5452	0.832	0.978
<i>E. delegatensis</i>	-19.2493	-0.1124	6.2012	-1.7207	1.681	0.996
<i>E. elata</i>	-2.5573	0.1193	-2.9895	0.0000	0.606	0.998
<i>E. fastigata</i>	-4.7471	-0.0962	1.7585	-0.8877	0.940	0.993
<i>E. fraxinoides</i>	-19.4705	-0.2660	10.7957	-5.3715	0.444	0.997
<i>E. globoidea</i>	-4.8415	-0.1452	2.1128	-0.6441	0.881	0.956
<i>E. muellerana</i>	-3.4680	-0.0988	1.1051	-0.5494	0.945	0.997
<i>E. nitens</i>	-0.8269	-0.0715	0.3794	-1.4807	0.602	0.997
<i>E. obliqua</i>	-0.9364	-0.0628	0.5001	-0.9724	1.064	0.990
<i>E. paniculata</i>	-3.6234	-0.1485	1.9373	-0.5739	0.913	0.960
<i>E. pauciflora</i>	12.2289	0.5486	-11.8786	0.0000	0.643	0.992
<i>E. pilularis</i>	2.4828	-0.0790	1.3682	-3.0222	0.831	0.988
<i>E. piperita</i>	-1.7287	-0.1347	1.3574	-1.3270	0.553	0.989
<i>E. radiata</i>	-6.8957	-0.3296	-1.2053	3.6160	0.917	0.987
<i>E. saligna</i>	-11.3675	-0.2593	4.2534	0.0000	1.481	0.965
<i>E. scias</i>	3.1396	0.2192	-3.5289	0.0000	0.787	0.934
<i>E. sieberi</i>	-5.7986	-0.0434	0.1301	0.5839	0.850	0.990
<i>E. smithii</i>	1.8002	0.0728	-2.7889	0.0000	0.955	0.985
<i>E. spp.</i>	-14.7972	-0.5416	8.1128	-0.9621	0.931	0.981
<i>E. viminalis</i>	-5.6868	-0.1732	2.4357	-1.2902	0.696	0.990

for predicting underbark diameter and estimating merchantable height at specified top end diameters. For several minor species, the size range of the sample trees was relatively small; any extrapolation much beyond the size range should be made with caution.

Literature Cited

- ALLEN, P.J. 1993. Average relative stem profile comparisons for tree size classes of Caribbean pine. *Can. J. For. Res.* 23:2594–2598.
- AMIDON, E.L. 1984. A general taper functional form to predict bole volume for five mixed-conifer species in California. *For. Sci.* 30:166–171.
- BAILEY, R.L. 1994. A compatible volume-taper model based on the Schumacher and Hall generalised constant form factor volume equation. *For. Sci.* 40:303–313.
- BI, H. 1999. Predicting stem volume to any height limit for native tree species in Southern New South Wales and Victoria. *N.Z. J. For. Sci.* 29:318–331.
- BI, H., AND F. HAMILTON. 1998. Stem volume equations for native tree species in southern New South Wales and Victoria. *Aus. For.* 61:275–286.
- BI, H., AND V. JURSKIS. 1996. Yield equations for irregular regrowth forests of *Eucalyptus fastigata* on the south-east tablelands of New South Wales. *Aus. For.* 59:151–160.
- BI, H., AND J. TURNER. 1994. Long term effects of superphosphate fertilization on stem form, taper and stem volume estimation of *Pinus radiata*. *For. Ecol. Manage.* 70:285–297.
- BIGING, G.S. 1984. Taper equations for second-growth mixed conifers of northern California. *For. Sci.* 30:1103–1117.
- BREIMAN, L. 1992. The little bootstrap and other methods for multidimensionality selection in regression: X-fixed prediction error. *J. Am. Stat. Assoc.* 87:738–754.
- BRUCE, D.R., R.O. CURTIS, AND C. VANCOEVERING. 1968. Development of a system of taper and volume tables for red alder. *For. Sci.* 14:339–350.
- CANDY, S.G. 1989. Compatible tree volume and variable-form stem taper models for *Pinus radiata* in Tasmania. *N.Z. J. For. Sci.* 19:97–111.
- CLEVELAND, W.S. 1993. Visualizing data. AT&T Bell Laboratories, Murray Hill, NJ. 360 p.
- CLUTTER, J.L. 1980. Development of taper functions from variable-top merchantable volume equations. *For. Sci.* 26:117–120.
- DEMAERSCHALK, J.P., AND A. KOZAK. 1977. The whole-bole system: A conditioned dual-equation system for precise prediction of tree profiles. *Can. J. For. Res.* 7:488–497.
- FANG, Z., AND R.L. BAILEY. 1999. Compatible volume and taper models with coefficients for tropical species on Hainan Island in Southern China. *For. Sci.* 45:85–100.
- FLEWELLING, J.W., AND L.M. RAYNES. 1993. Variable-shape stem-profile predictions for western hemlock. Part I. Predictions from DBH and total height. *Can. J. For. Res.* 23:520–536.
- FLURY, B. 1990. Principal points. *Biometrika* 77:33–41.
- FORSLUND, R.R. 1991. The power function as a simple stem profile examination tool. *Can. J. For. Res.* 21:193–198.
- FOX, J. 1991. Regression diagnostics. Sage Univ. Pap. Ser. on Quantitative applications in the social sciences, Series No:07-079. Sage, Newbury Park, CA.
- FRIES, J., AND B. MATERN. 1966. On the use of multivariate methods for the construction of tree taper curves. P. 85–117 in *Res. Notes*, No. 9. Internat. Advisory Group of Forest Statisticians. Dep. of For. Biom., Roy. Coll. of For., Stockholm, Sweden. 370 p.
- GOODWIN, A. 1992. A taper equation for *Eucalyptus obliqua*. P. 454–467 in *Proc. of IUFRO conf. on Integrating forest information over space and time*. ANUTECH Pty Ltd., Canberra, Australia. 499 p.
- GORDON, A.D. 1983. Comparison of compatible polynomial taper equations. *N.Z. J. For. Sci.* 13:146–155.

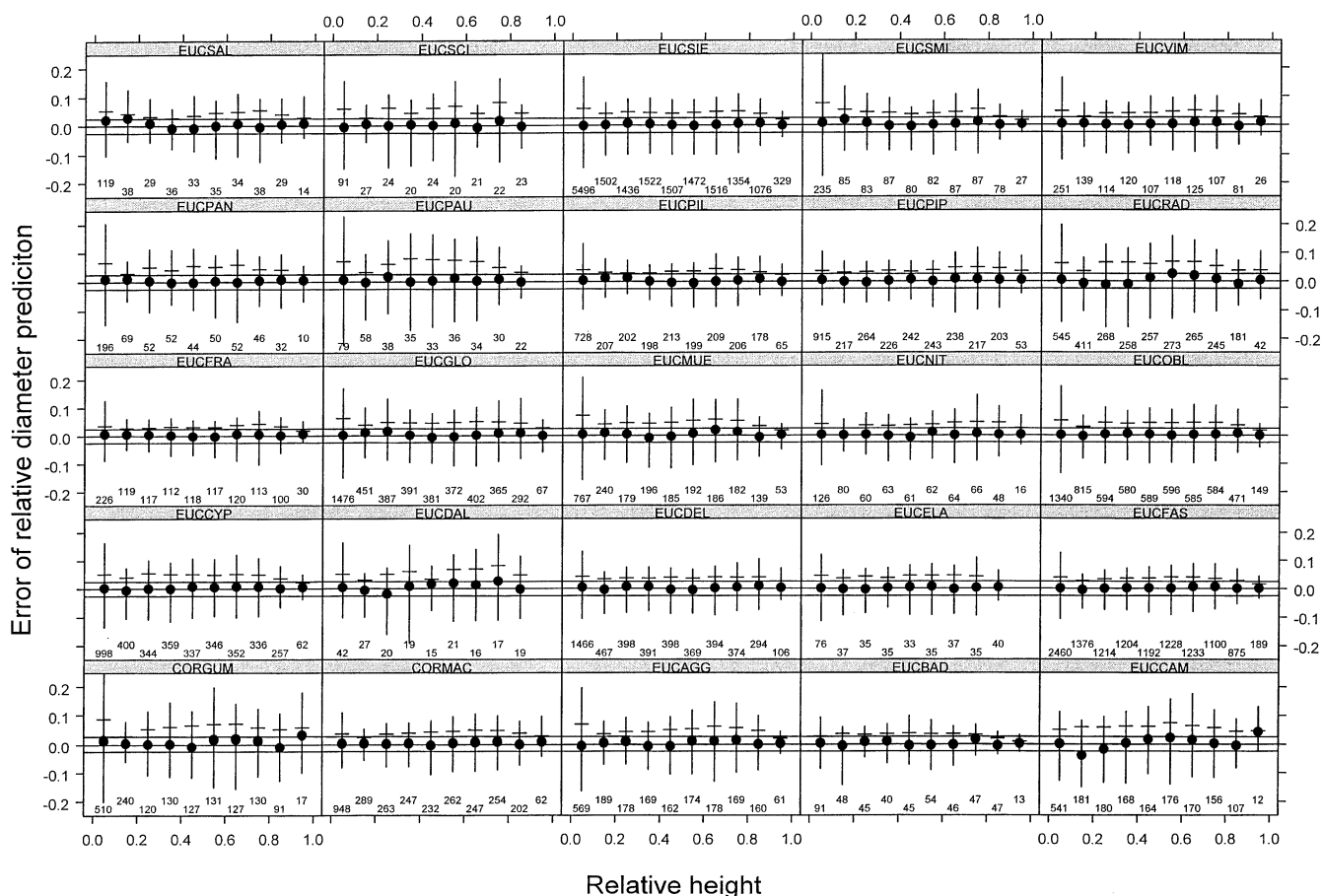


Figure 5. Bias and precision of taper prediction along relative height classes for the 25 species as indicated by the six-letter code for the trigonometric variable-form taper functions. The filled circles represent the bias for the corresponding relative height classes. Within each relative height class, 90% of the prediction errors fall within the interval shown by the vertical bar. The mean absolute values of prediction error are indicated by the small horizontal line segment crossing the vertical bar. The horizontal lines immediately above and below the x-axis mark a range of ± 0.025 . The number below each bar indicates the number of data points in the relative height class.

GORDON, A.D., C. LUNDGREN, AND E. HAY. 1995. Development of a composite taper equation to predict over- and under-bark diameter and volume of *Eucalyptus saligna* in New Zealand. *N.Z. J. For. Sci.* 25:318–327.

GOULDING, C.J., AND J.C. MURRAY. 1975. Polynomial taper equations that are compatible with tree volume equations. *N.Z. J. For. Sci.* 5:313–322.

GREENE, W.H. 1993. *Econometric analysis*. Ed 2. MacMillan, New York. 791 p.

KOZAK, A. 1988. A variable-exponent taper equation. *Can. J. For. Res.* 18:1363–1368.

KOZAK, A. 1997. Effects of multicollinearity and autocorrelation on the variable-exponent taper functions. *Can. J. For. Res.* 27:619–629.

KOZAK, A., AND J.H.G. SMITH. 1993. Standards for evaluating taper estimating systems. *For. Chron.* 69:438–444.

LEMAY, V.M., A. KOZAK, C.K. MUHAREWE, AND R.A. KOZAK. 1993. Factors affecting the performance of Kozak's (1988) variable-exponent taper functions. *Proc. of IUFRO conf. on Modern methods of estimating tree and log volume*. West Virginia Univ. Publ. Serv., Morgantown, WV, P. 168.

MATHEWS, J.H. 1987. *Numerical methods for computer science, engineering and mathematics*. Prentice-Hall International, Inc., London. 507 p. (AU: Place of Publisher?)

MAX, T.A., AND H.E. BURKHART. 1976. Segmented polynomial regression applied to taper equations. *For. Sci.* 22:283–289.

MUHAIKWE, C.K. 1994. Tree form and taper variation over time for interior lodgepole pine. *Can. J. For. Res.* 24:1904–1913.

MUHAIKWE, C.K. 1999. Taper equations for *Eucalyptus pilularis* and *E. grandis* for the north coast in New South Wales, Australia. *For. Ecol. Manage.* 113:251–269.

MYERS, R.H. 1990. *Classical and modern regression with applications*. PWS-KENT, Boston. 488 p.

NEWHAM, R.M. 1988. A variable form taper function. *For. Can. Petawawa Nat. For. Inst. Inf. Rep.*, PI-X-83.

NEWHAM, R.M. 1992. Variable-form taper functions for four Alberta tree species. *Can. J. For. Res.* 22:210–223.

NISHII, R. 1984. Asymptotic properties of criteria for selection of variables in multiple regression. *Ann. Stat.* 12:758–765.

PEREZ, D.N., H.E. BURKHART, AND C.T. STIFF. 1990. A variable-form taper function for *Pinus oocarpa* Schiede in central Honduras. *For. Sci.* 36:186–191.

THOMAS, C.E., AND B.R. PARRESOL. 1991. Simple, flexible, trigonometric taper equations. *Can. J. For. Res.* 21:1132–1137.

TUKEY, J.W. 1977. *Exploratory data analysis*. Addison-Wesley, Reading, MA. 688 p.

VAN DEUSEN, P.C., A.D. SULLIVAN, AND T.G. MATNEY. 1981. A prediction system for cubic foot volume of Loblolly pine applicable through much of its range. *South. J. Appl. For.* 5:186–189.

WACKERLY, D.D., W. MENDENHALL, AND R.L. SCHEFFER. 1996. *Mathematical statistics with applications*. Duxbury Press, Belmont. 798 p.

WILLIAMS, M.S., AND R.M. REICH. 1997. Exploring the error structure of taper equations. *For. Sci.* 43:378–386.

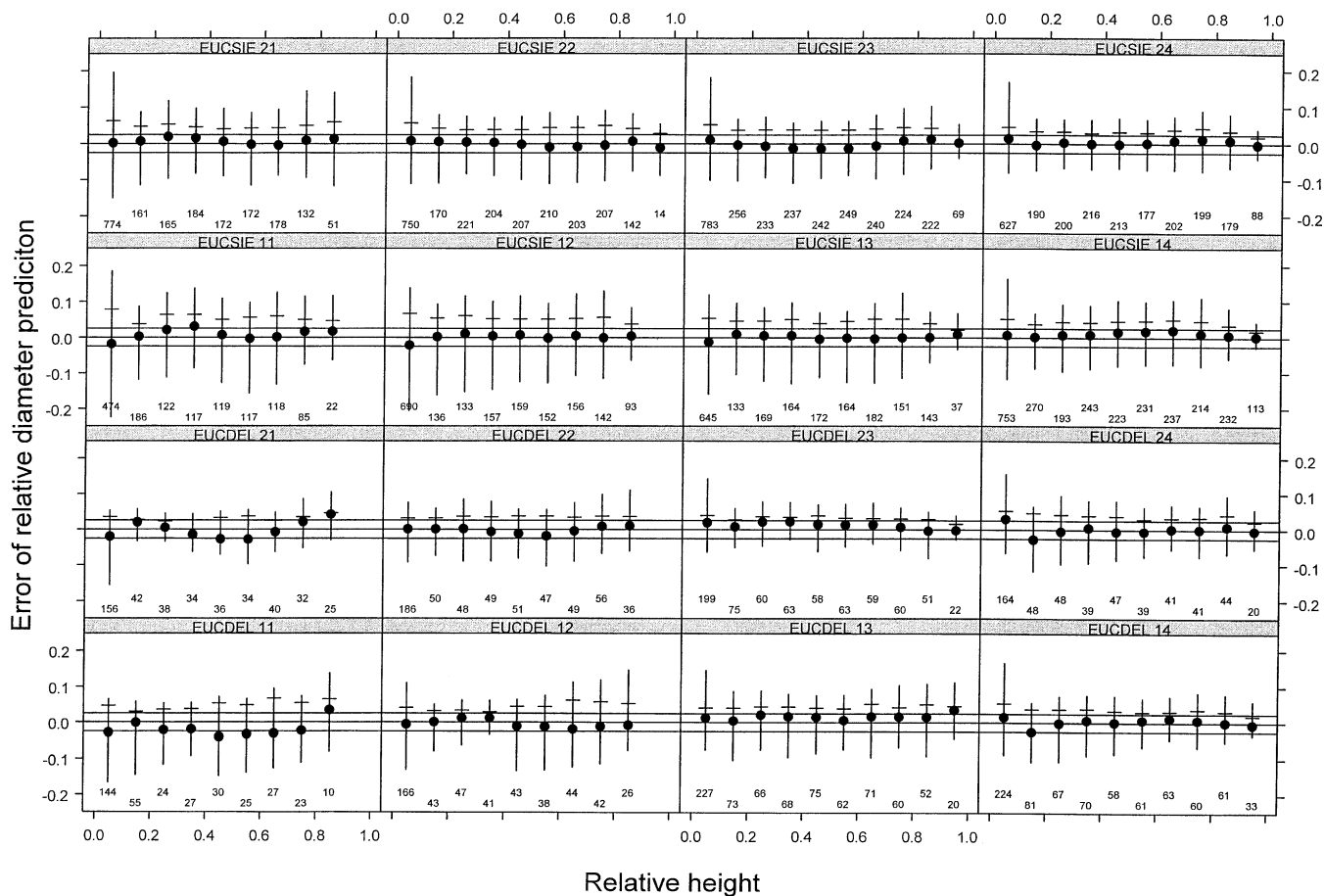


Figure 6. Bias and precision of taper prediction along relative height classes in each size group of *E. delegatensis* and *E. sieberi* (indicated by the code in the strip of each panel) for the trigonometric variable-form taper functions. The filled circles represent the bias for the corresponding relative height classes. Within each relative height class, 90% of the prediction errors fall within the interval shown by the vertical bar. The mean absolute values of prediction error are indicated by the small horizontal line segment crossing the vertical bar. The horizontal lines immediately above and below the x-axis mark a range of ± 0.025 . The number below each bar indicates the number of data points in the relative height class.

Table 3. Comparative prediction accuracy of the trigonometric variable-form taper function and Kozak's (1988) model for predicting relative diameter and absolute underbark diameter. R_{MSEP} and R_{bias} are the ratio of mean squared error of prediction and the ratio of the mean absolute local bias of Kozak's (1988) taper model over that of the trigonometric taper function. The extra value of R_{bias} for each of the 8 species with larger sample sizes was based on a total of 80 relative height classes in 8 size groups.

Species	Relative diameter		Underbark diameter		
	R_{MSEP}	R_{bias}	R_{MSEP}	R_{bias}	
<i>C. gummifera</i>	1.07	1.38	1.19	1.74	
<i>C. maculata</i>	1.02	1.01	1.00	1.15	0.95
<i>E. agglomerata</i>	1.02	1.42	1.08	0.68	
<i>E. badjensis</i>	1.07	1.14	1.26	1.22	
<i>E. camaldulensis</i>	1.01	0.94	0.98	0.99	
<i>E. cypelloarpa</i>	1.02	0.89	1.04	0.93	1.13
<i>E. dalrympleana</i>	0.92	1.12	0.87	1.13	
<i>E. delegatensis</i>	1.23	1.28	1.24	1.17	1.18
<i>E. elata</i>	1.22	1.73	1.47	2.40	
<i>E. fastigata</i>	1.13	2.16	1.06	1.61	1.49
<i>E. fraxinoides</i>	1.00	0.66	1.02	1.00	
<i>E. globoidea</i>	1.01	1.11	1.06	1.01	1.11
<i>E. muellerana</i>	1.10	0.83	1.50	2.05	
<i>E. nitens</i>	1.06	1.79	1.33	1.92	
<i>E. obliqua</i>	1.08	0.98	1.16	1.05	1.32
<i>E. paniculata</i>	0.97	1.70	1.02	1.05	
<i>E. pauciflora</i>	1.52	1.25	1.25	1.24	
<i>E. pilularis</i>	1.04	1.04	1.01	1.04	
<i>E. piperita</i>	1.05	0.92	1.11	1.30	
<i>E. radiata</i>	1.09	1.46	1.00	0.85	0.86
<i>E. saligna</i>	1.09	1.93	1.09	0.87	
<i>E. scias</i>	1.04	1.79	1.47	1.90	
<i>E. sieberi</i>	1.05	1.52	1.08	1.27	1.45
<i>E. smithii</i>	1.01	1.03	0.94	0.83	
<i>E. spp.</i>	1.06	1.03	1.31	0.89	
<i>E. viminalis</i>	1.04	1.14	1.08	1.10	
<i>Mean</i>	1.07	1.28	1.14	1.25	1.19

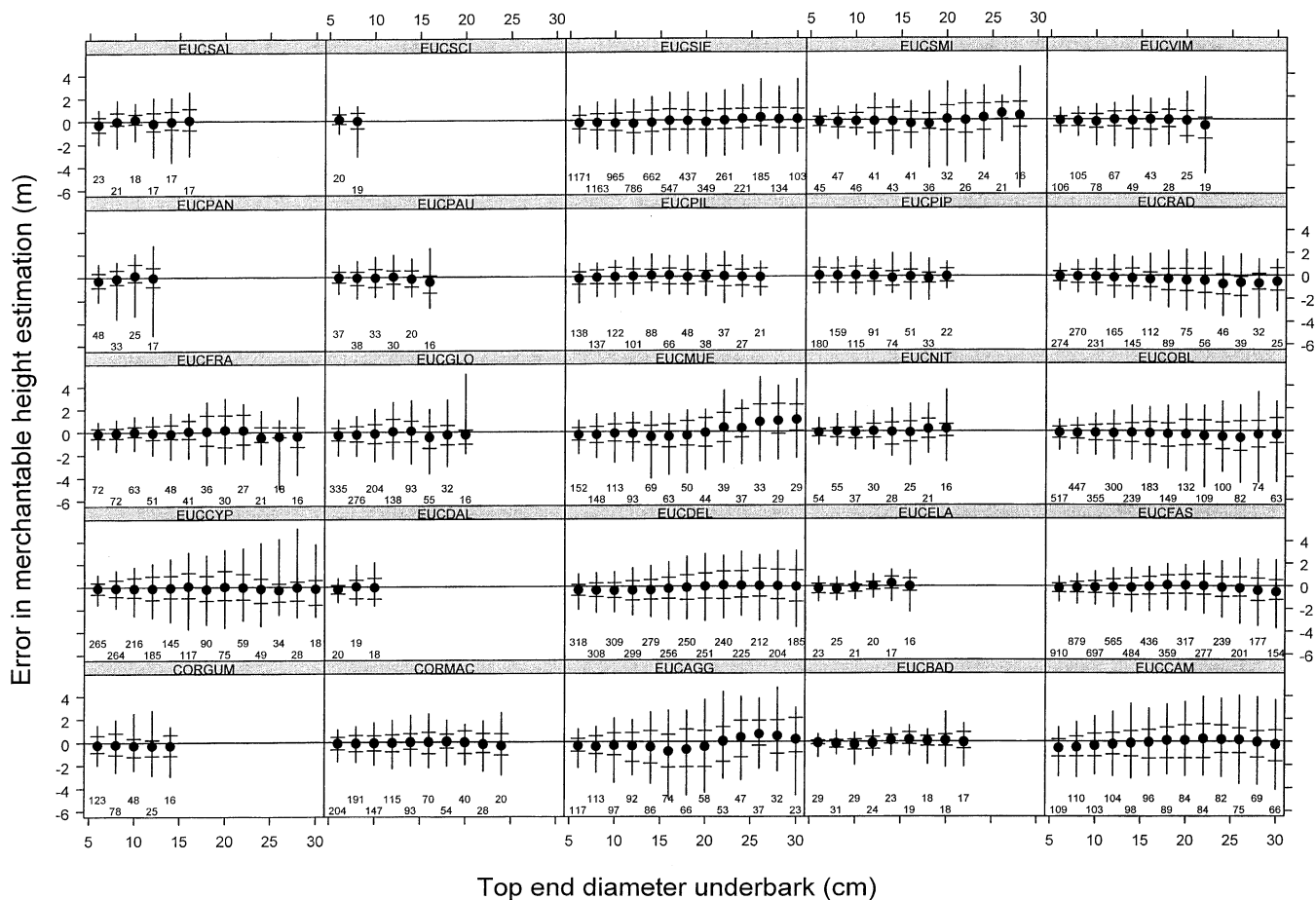


Figure 7. Bias and precision of merchantable height estimation for a number of top end underbark diameters for the 25 species. The filled circles represent the mean error of merchantable height estimation. For each top end diameter, 90% of the prediction errors fall within the interval shown by the vertical bar, and 50% fall within the horizontal line segments cross the vertical bar. The number below each bar indicates the number of data points.