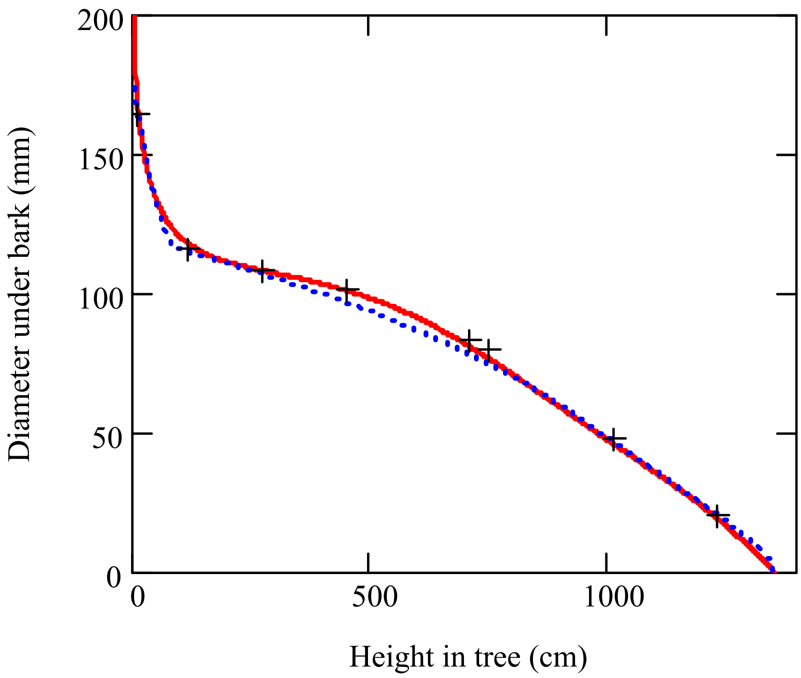
Arbetsrapport

Nr 491
2001

Taper functions for *Picea abies* (L.) Karst. *and Pinus sylvestris* L. in Sweden

Kalle Spångberg, John Arlinger, Lars Wilhelmsson, SkogForsk
Sven-Olof Lundqvist and Örjan Hedenberg, STFI



Uppsala Science Park, SE 751 83 UPPSALA • Tel: 018-18 85 00 • Fax: 018-18 86 00
skogforsk@skogforsk.se • http://www.skogforsk.se
STFI, Box 5604, SE 114 86 STOCKHOLM, SWEDEN

Key words: Taper, *Pinus sylvestris, Picea abies*, diameter, prediction, varible-form function, segmented polynomial function.

SkogForsk – Stiftelsen Skogsbrukets Forskningsinstitut

SkogForsk arbetar för ett långsiktigt, lönsamt skogsbruk på ekologisk grund. Bakom SkogForsk står skogsbolag, skogsägareföreningar, stift, gods, allmänningar, plantskolor, SkogsMaskinFöretagarna m.fl., som betalar årliga intressentbidrag. Hela skogsbruket bidrar dessutom till finansieringen genom en avgift på virke som avverkas i Sverige. Verksamheten finansieras vidare av staten enligt särskilt avtal och av fonder som ger projektbundet stöd.

SkogForsk arbetar med forskning och utveckling med fokus på tre centrala frågeställningar: Skogsodlingsmaterial, Skogsskötsel samt Råvaruutnyttjande och produktionseffektivitet. På de områden där SkogForsk har särskild kompetens utförs även i stor omfattning uppdrag åt skogsföretag, maskintillverkare och myndigheter.

Serien ***Arbetsrapport*** dokumenterar långliggande försök samt inventeringar, studier m.m. och distribueras enbart efter särskild beställning.

Forsknings- och försöksresultat från SkogForsk publiceras i följande serier:

SkogForsk-Nytt: Nyheter, sammanfattningar, översikter.
Resultat: Slutsatser och rekommendationer i lättillgänglig form.
Redogörelse: Utförlig redovisning av genomfört forskningsarbete.
Report: Vetenskapligt inriktad serie (på engelska).
Handledningar: Anvisningar för hur olika arbeten lämpligen utförs.

ISSN 1404-305X

Contents

Abstract

Taper functions were fitted to stems of Norway spruce and Scots pine from Sweden. Input variables are diameter under bark at breast height, and tree height. The taper functions were of the segmented polynomial type, or of the variable-form type. The results show that the segmented polynomial functions (Max, T. A., & Burkhart, H. E., 1976. For. Sc. 22(3):283-289) gave slightly better predictions of diameter under bark along the stems, compared to variable-form functions and established functions commonly used in Swedish forestry. Including auxiliary variables, e.g. tree age in breast height, and live crown height, did not make the variable-form function outperform the segmented polynomial function.

Introduction

This study was carried out within the SMP, or Forest-Pulp-Paper, project, whose objective is to establish predictive models for wood and fiber properties, primarily of interest for the forest industry. The objective of the present study is to establish functions that predict the diameter under bark at a given height in the stem. The stem diameter at a given height in a tree is of interest for two reasons. First, the diameter structure, or the taper, of the stem determines the volume of the merchantable logs. Secondly, the expression of growth at different heights in the tree can contribute to the explanation of the variation in wood and fiber properties (Larson, 1969; Lindström, 1996a; Lindström, 1996b; Lindström, 1997; Lundqvist et. al 2001, Wilhelmsson et. al, In press).

Once the tree is harvested, the diameters are possible to measure. But accurate predictions of volume and properties of the standing forests are also of interest for forestry planning. In most cases in forestry planning, the only diameter available is the diameter that can be practically measured in the field, i.e. the diameter within reach for a human from the ground. If the forestry planning should have a chance of describing the diameter at a higher position in the tree, then a taper function must be used.

For pine and spruce, the most commonly used functions in Sweden are segmented polynomial functions with three segments developed by Edgren & Nylinder (1949). They built a function with three segments, which is conditioned to pass through the observed diameters at 60% (d60) and 20% (d20) of the tree height. The problem is that the diameters at these heights are unknown for a standing tree, so the form quotient inside bark (d20/d60) must be predicted. The practical use of Edgren & Nylinder's function, E&N, is consequently to first predict the form factor (quota between actual volume of the stem above the butt and the volume of a cylinder with the same height as the stem and the same diameter at the base), then to predict the form quotient using diameter at breast height, tree height, and form factor, and finally chose the parameters for the taper function from tables by region, species and form quotient. This gives a function that is unique to the tree form and region, but the choice of parameters is based on predictions (unless the d20 and d60 can be measured), and the variation in parameter values is discontinuous. The diameter is expressed relative to the base diameter of the tree, so a reference diameter under bark at a known height above ground is necessary to obtain predicted diameters under bark.

A promising taper function that can be fitted to Swedish data is the segmented polynomial function of Max & Burkhart (1976). They have built a function, SPF, which describes the area under bark, relative to the area over bark at breast height, as a function of relative height in the tree. The function has shown good precision in predicting the diameter along the stems of pine species in particular. The input is diameter at breast height and tree height. An assumption with the SPF is that the taper of all trees can be described with one function, so there is no adaptation to the individual tree. However, the dependent variable y has a reference point where y=1 at breast height, but the relative height of the breast height differs with tree height. This means that the relative height at which the relative area is 1 will change with tree height (figure 1). Thus, the assumption that one function can describe all trees is a simplify-cation.

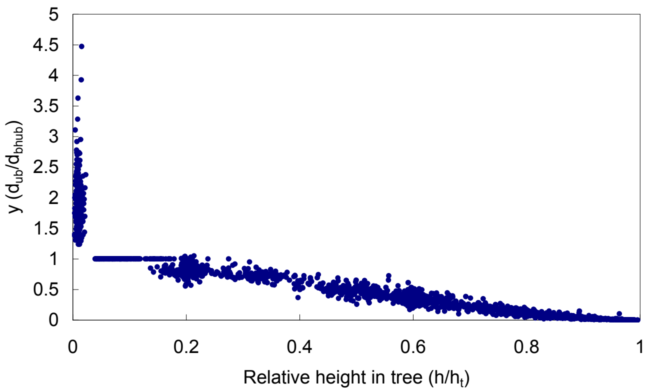


Figure 1.
y in the adaptation of Max & Burkhart's (1976) segmented polynomial function, SPF, for spruce data.

In the original definition, y is the squared diameter under bark relative to the squared diameter at breast height over bark (Max and Burkhart, 1976), but here the diameter at breast height under bark is used as the reference value according to

$$y = b_1(x-1) + b_2(x^2-1) + b_3(a_1-x)^2 I_1 + b_4(a_2-x)^2 I_2 + \varepsilon \qquad \text{(SPF)}$$

or,

$$d_{ub} = \sqrt{d_{bhub}^{\ 2}\left(b_1(x-1) + b_2(x^2-1) + b_3(a_1-x)^2 I_1 + b_4(a_2-x)^2 I_2\right)}$$

where $y=d_{ub}^2/d_{bhub}^2$, $x=h/h_t$, d_{ub} is diameter under bark, d_{bhub} is diameter under bark at breast height 1.3 m above ground level, h is height in stem, h_t is tree height, b_1, b_2, b_3 and b_4 are constants, and

$I_1=1$ if $x \le a_1$, else $I_1=0$
$I_2=1$ if $x \le a_2$, else $I_2=0$

A more recent methodology for describing stem taper was introduced by Newnham (1992) who describes the taper using an exponential relationship between relative diameter and relative height in the tree, in a so-called variable-form function, VFF:

$$\left(\frac{d_{ub}}{d_{bhub}}\right)^k = \frac{h_t - h}{h_t - 130} \qquad \text{(VFF)}$$

or

$$d_{ub} = d_{bhub}\left(\frac{h_t - h}{h_t - 130}\right)^{\frac{1}{k}}$$

where d_{ub} is diameter under bark, d_{bhub} is diameter under bark at breast height 1.3 m above ground level, h is height in stem, h_t is tree height, and k is a variable controlling the curve. The exponent k is variable along the stem and between different trees (figure 2). Newnham also allows a constant g that can be used to convert the units used for relative diameter and relative height in tree, but that constant is not used here, nor by him. Newnham estimated ln(k) using variables derived from the diameter at breast height and the tree height. Although crown related variables, such as live crown ratio, would be logical variables to include in the estimate of ln(k), Newnham chose not to include them since they are seldom available in the data. However, in the present study, k was allowed to be described by such variables also due to their influence on the allocation of growth to the stem (Larson, 1969). The VFF has been successful in predicting diameter along the stems, particularly of other species than *Pinus*, in North America. It has recently been fitted to Scots pine in Sweden by Petersson (1999), P7.

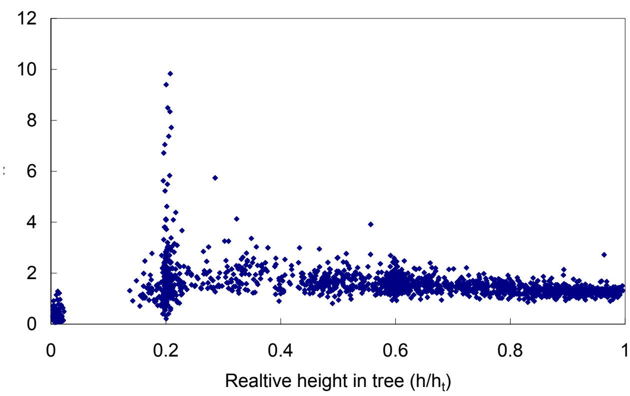


Figure 2.
k in Newnham's (1992) variable-form function, VFF, for spruce data.

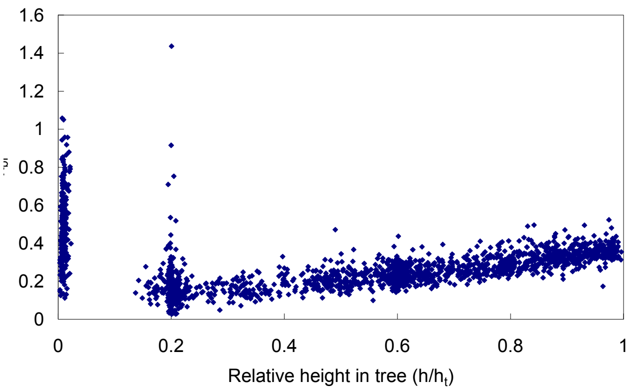


Figure 3.
k_{bi} in Bi's (2000) trigonometric variable-form function, VFFBi, for spruce data.

An improvement of the variable-form function is reported by Bi who has developed a trigonometric variable-form function , VFFBi, with a trigonometric base function, which by the exponent k_{bi}, (figure 3) explains the relative diameter (Bi, 2000). This base is itself variable in that the difference in relative height of breast height between trees of different sizes varies. Three trigonometric variables, and three stem variables, are used to describe the variation in stem shape within the stem. The choice of trigonometric variables was made according to Fourier analysis in order to overcome a cyclic pattern of variation in the residuals from regressing ln(d) against ln(b). Again, the exponential relation between diameter and relative height changes along the stem as well as between stems of different size (figure 3). According to Bi, the trigonometric variable-form function is more stable than Newnham's, and Kozak's similar functions, for which the choice of variables to include in the function changes with different data sets (Kozak 1988; Newnham 1988; Newnham 1992; Kozak 1997; Petersson, 1999).

$$\frac{d_{ub}}{d_{bhub}} = \left(\frac{\ln\left(\sin\left(\frac{\pi h}{2h_t}\right)\right)}{\ln\left(\sin\left(\frac{\pi}{2}b\right)\right)} \right)^{k_{Bi}} \qquad \text{VFFBi}$$

where

$b = 130/h_t$

$$k_{Bi} = \left(a_1 + a_2 \sin\left(\frac{\pi h}{2h_t}\right) + a_3 \cos\left(\frac{3\pi h}{2h_t}\right) + a_4 \frac{\sin\left(\frac{\pi h}{2h_t}\right)}{\frac{h}{h_t}} + a_5 d_{bhob} + a_6 \frac{h\sqrt{d_{bhob}}}{h_t} + a_7 \frac{h\sqrt{h_t}}{h_t} \right)$$

where d_{ub} is diameter under bark, d_{bhub} is diameter (mm) under bark at breast height 1.3 m above ground level, h is height in stem (cm), h_t is total tree height (cm), and k_{bi} is a variable controlling the curve.

Material and Methods

Discs were taken from specified heights along the stems of 252 Norway spruce trees , *Picea abies* (L.) Karst., originating from 42 spruce dominated stands and 120 Scots pine trees, *Pinus sylvestris* (L.), originating from 20 pine dominated stands (Arlinger et al 2001). The stands were systematically sampled along gradients of latitude and temperature sum (Morén & Perttu, 1994) in order to cover the predominant climatic conditions in most parts of Sweden, (figure 4).

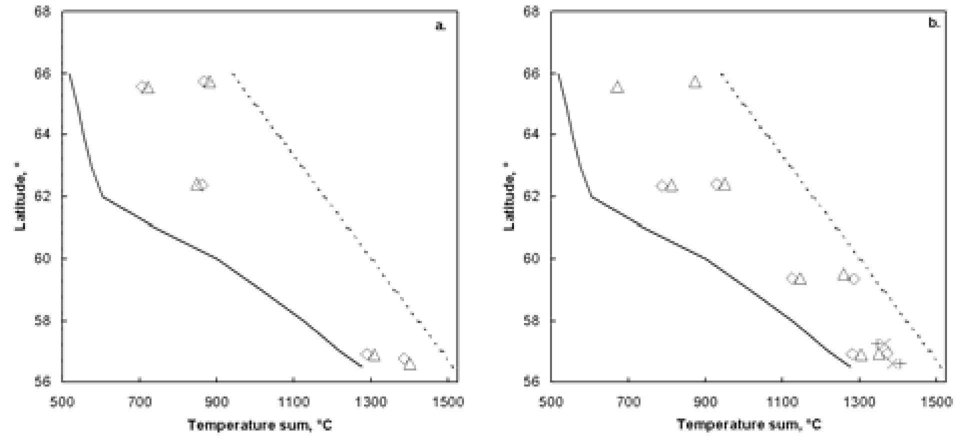


Figure 4.
Distribution of Scots pine (left) and Norway spruce (right) stands according to latitude, temperature sum, maturity class, and pairs of high and low site fertility (Δ = pair of old stands, ◊ = pair of young stands). Additional sampling around latitude 57° included Norway spruce stands in regions with different average levels of precipitation, × = additional old spruce stand and + = additional young spruce stand).

In order to span the variation in the number of annual rings and the diameter of logs, both younger and older stands on high and low fertility sites were sampled. Two relatively evenly stocked circular plots consisting of at least 25 healthy trees were selected within each stand. All trees were calipered and classified into four breast height diameter classes (thick, average, thin and very thin). A site classification was made for each plot. One sample tree per diameter class and plot was randomly selected among each of the three largest diameter classes. The sample trees were felled and 3-9 sample heights determined depending on tree height and age. A crosscut disc was removed five cm above the nearest node below the sample height. The discs were frozen at the end of each day in the field and transported to the lab at STFI, the Swedish Pulp and Paper Research Institute. The diameter, bark thickness, and growth ring pattern was measured on the thawed disc in the lab using image analyses (Lundqvist, 1998). The diameter under bark was measured in the field at four different positions along the stem: butt height, 20% and 60% of the tree height, and at the third node of branches below the top. Just as the crosscuts, these field measurements were made five cm above the nearest node below the intended height. The sampling resulted in a data set that spanned a large variation in tree characteristics (table 1).

Table 1.
Descriptive statistics for trees in data set.

	Norway spruce N=251 trees			Scots pine N=117 trees		
Variable	Mean	Min.	Max.	Mean	Min.	Max.
Tree age in breast height (1.3 m above ground)	56	17	208	69	19	130
Diameter under bark in breast height (cm)	18	4.4	47	16	6.0	35
Tree height above ground (m)	17	5.0	33	15	7.0	26

The image analysis measured the radii from the pith to the bark in two opposite directions, north to south, which were marked on the disc in the field. Since the fieldwork was carried out under the growing season, the latest annual growth ring was more or less developed. Therefore, the outermost annual growth ring was excluded in the data. But the field measurements of the diameter under bark included all annual growth rings, so the diameter from the image analysis was increased by the average width of the three outermost annual growth rings in the disc.

The data was checked for outliers by plotting the diameter and height of the observation tree by tree. Abnormal observations that were the results of obvious measurement or disc numbering errors, such as observations much larger or smaller than both observations below and above in the tree, were removed. One tree and three discs were removed as outliers from the spruce data set, and five observations from the pine data set. Several observations that deviate from the general patterns of k, k_{bi} and y were, however, left for a lack of arguments why their values were results of unusual errors rather than results of uneven growth, stem ovality etc.

In order to obtain accurate predictions of stem diameter at a given height above ground level, the SPF, VFF and VFFBi were fitted to Norway spruce and Scots pine data respectively. The precision was evaluated as root mean square error, RMSE,

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(d_i - \hat{d}_i)^2}{n-p}}$$

or relative to the observed value

$$RMSE(\%) = 100\sqrt{\frac{\sum_{i=1}^{n}(\frac{d_i - \hat{d}_i}{d_i})^2}{n-p}}$$

where d is the observed diameter under bark, $\hat{d}$ is the predicted diameter under bark, n is the number of observations and p is the number of parameters in the function. The precision is also presented by classes of commercial dimension, i.e. residue where $d < 50$ mm, thin pulpwood where $50 \leq d < 100$ mm, thick pulpwood where $100 \leq d < 150$ mm, and timber where $d \geq 150$ mm. These diameter boundaries reflect the current commercial dimensions in Sweden where 5 cm under bark is the lowest small end diameter for pulpwood, and 15 cm is a common minimum timber small end diameter. The distinction between thin and thick pulpwood is, however, particular to this study.

The result of the prediction was compared to that when using Edgren & Nylinder´s (1949) functions for Scots pine and Norway spruce, E&N, and Petersson's (1999) fitted variable-form function for Scots pine, P7. For this purpose, the data was divided into a validation and a calibration data set. The division into calibration and validation data sets was aimed at reaching the same variability in each data set, while making the data sets unique from each other. With this in mind, the data was divided at stand level, so that plots within the same stand and trees within the same plot would end up in the same data set. This way, any eventual similarities within or between plots within stand would not affect the precision of the predictions. One stand from each combination of classes of stand age, growth rate and temperature sum, was randomly chosen to the validation data. The SPF and VFF functions were refitted to the calibration data, and the parameters from the calibration were used to predict the validation data. Mean and standard deviation describes the distributions of the residuals.

Results

The variable k in the VFF functions (figure 2) is defined by relative height and relative diameter as follows:

$$k = \frac{\ln\left(\frac{h_t - h}{h_t - 130}\right)}{\ln\left(\frac{d_{ub}}{d_{bhub}}\right)}$$

An improvement in the resulting taper functions has been noted when the variation in the natural logarithm of k, ln(k), rather than k, has been explained (Newnham, 1992; Kozak and Smith, 1993). The ln(k) was, consequently, described by least squares regression using SAS, PROC REG, SELECTION=MAXR (Anon. 1989). With this procedure, several functions are tested in order to find the function with the largest R^2 while the number of independent variables is low. Three sub-functions were used to explain the variation in ln(k), to be used in different situations depending on which input data is available. In the first case, the independent variables available were transformations and combinations of diameter at breast height under bark and tree height. The resulting function is called VFFdh. In the other cases, variables describing the crown length, tree age, and temperature sum, which can be described as a function of latitude and longitude (Morén & Perttu, 1994) were also allowed. These functions were named VFF4, VFF5 or VFF6 according to the number of variables. The sub-functions are described in (table 2). The VFFBi was fitted by explaining $\ln(k_{bi})$ as a function of the variables given by Bi (table 3) using SAS, PROC REG (Anon. 1989).

Table 2.
Parameters for each function that describes the dependent variable ln(k) in the variable-form function. Fitted to entire data for each species.

Spruce: VFFdh (n=1929)

$$\ln(k) = 0.316 + 1.16x - 6.97x^3 \frac{d_{bhob}}{h_t} - 0.00448 \frac{h_t}{\sqrt{h}} \qquad \text{MSE}=0.10133$$

Spruce: VFF4 (n=1929)

$$\ln(k) = 0.632 + 1.11x - 6.29x^3 \frac{d_{bhob}}{h_t} - 0.00460 \frac{h_t}{\sqrt{h}} - 0.429xh_c \qquad \text{MSE}=0.09756$$

Spruce: VFF6 (n=1929)

$$\ln(k) = 1.24 + 0.872x - 3.98 \frac{d_{bhob}}{h_t} + 7.39 \times 10^{-6} \frac{h_t^2}{h} - 0.0102 \frac{h_t}{\sqrt{h}} - 3.92 \frac{h_t - h_c}{h_t - 130} + 3.99 \frac{h_t - h_c}{h_t}$$

MSE=0.08991

Pine: VFFdh (n=844)

$$\ln(k) = 0.369 + 1.27x - 5.92x^3 \frac{d_{bhob}}{h_t} - 0.00439 \frac{h_t}{\sqrt{h}} \qquad \text{MSE}=0.10908$$

Pine: VFF4 (n=844)

$$\ln(k) = 0.189 + 1.31x - 5.77x^3 \frac{d_{bhob}}{h_t} - 0.00453 \frac{h_t}{\sqrt{h}} - 0.00227a_{bh} \qquad \text{MSE}=0.10306$$

Pine: VFF5 (n=844)

$$\ln(k) = 0.592 + 31.6x^2 \frac{d_{bhob}}{h_t} - 22.5x^3 \frac{d_{bhob}}{h_t} - 47.6x^2 \left(\frac{d_{bhob}}{h_t}\right)^2 - 0.00388 \frac{h_t}{\sqrt{h}} - 153 \times 10^{-6} t_{sum} \frac{d_{bhob}}{d_{baw}}$$

MSE=0.10539

where $x=(h_t-h)/(h_t-130)$ where h is height in tree (cm), h_t is tree height (cm), d_{bhob} is diameter (mm) over bark at breast height 1.3 m above ground level, a_{bh} is number of annual growth rings at breast height, h_c is height (cm) to the first node with living branches, t_{sum} is the temperature sum (Morén & Perttu, 1994), and d_{baw} is basal area weighted mean diameter (mm) over bark at breast height within the plot.

Table 3.
The natural logarithm of k_{bi} as a function of tree variables.

Spruce:

$$\ln(k_{Bi}) = \left(13.385 - 5.0930\sin\left(\frac{\pi h}{2h_t}\right) - 1.1175\cos\left(\frac{3\pi h}{2h_t}\right) - 8.45305\frac{\sin\left(\frac{\pi h}{2h_t}\right)}{\frac{h}{h_t}} + 0.00125 d_{bhob} + 0.03286\frac{h\sqrt{d_{bhob}}}{h_t} - 0.0363\frac{h\sqrt{h_t}}{h_t} \right)$$

n=1928, MSE=0.08633

Pine:

$$\ln(k_{Bi}) = \left(11.662 - 4.2600\sin\left(\frac{\pi h}{2h_t}\right) - 0.90291\cos\left(\frac{3\pi h}{2h_t}\right) - 7.6278\frac{\sin\left(\frac{\pi h}{2h_t}\right)}{\frac{h}{h_t}} + 578.8\cdot 10^{-6} d_{bhob} + 0.03484\frac{h\sqrt{d_{bhob}}}{h_t} - 0.03396\frac{h\sqrt{h_t}}{h_t} \right)$$

n=844, MSE=0.09768

All parameters are significant at α=0.05 or better, except the second to last, with a value 0.03484, for pine.

For all cases of VFF, the natural logarithm of k and k_{bi}, is explained by multiple linear regression. This introduces a bias when the predicted ln(k) and ln(k_{bi}) are retransformed to k and k_{bi}. This is a phenomena associated with the different distributions of a variable and the logarithm of the variable. When a predicted variable is retransformed from the logarithmic scale, the distribution will be skewed. A correction for this skewness can be made when the variable is retransformed:

$$k_{pcorr} = e^{\left(\ln(k)_p + MSE/2\right)} \quad (1)$$

where k_{pcorr} is the predicted and retransformed variable k in VFF functions, $\ln(k)_p$ is the predicted natural logarithm of the variable k, and MSE is the Mean Square Error of the function predicting the natural logarithm of the variable k (Brownlee, 1965). This correction for the logarithmic bias is included in the functions so that

$$d_{ub} = d_{bhub}\left(\frac{h_t - h}{h_t - 130}\right)^{\frac{1}{k_{pcorr}}} \qquad \text{VFF (2)}$$

$$d_{ub} = d_{bhub}\left(\frac{\ln\left(\sin\left(\frac{\pi h}{2h_t}\right)\right)}{\ln\left(\sin\left(\frac{\pi}{2}b\right)\right)}\right)^{k_{Bipcorr}} \qquad \text{VFFBi (3)}$$

where $b = 130/h_t$

Such a correction has not been discussed in earlier work on the VFF (Kozak, 1988; Newnham, 1992), but it leads to an improvement of the distribution for k and k_{bi}. The arithmetic mean of $k_{pdhuncorr}$ is 1.354 but the mean of $k_{pdhcorr}$ is 1.424, which is closer to the mean of the observed k 1.423 (table 4).

Table 4.
Comparing distributions of individual tree values of k and k_{bi} and parameter predicted k_p values with and without correction for logarithmic bias.

Variable	Spruce n=1928 Mean	Spruce Std Dev	Pine n=844 Mean	Pine Std Dev
ln(k)	0.1749	0.7069	0.3461	0.6478
$\ln(k)_{pdh}$	0.1749	0.6313	0.3461	0.5577
$\ln(k)_{p4}$	0.1749	0.6343	0.3461	0.5631
$\ln(k)_{p6}$	0.1749	0.6404	0.3461	0.5612
k	1.423	0.7735	1.669	0.9262
$k_{pdhcorr}$	1.424	0.4876	1.665	0.5682
$k_{pdhuncorr}$	1.354	0.4635	1.576	0.5381
k_{p4corr}	1.424	0.4954	1.664	0.5754
$k_{p4uncorr}$	1.356	0.4718	1.580	0.5465
k_{p5corr}			1.666	0.5782
$k_{p5uncorr}$			1.580	0.5486
k_{p6corr}	1.425	0.5074		
$k_{p6uncorr}$	1.363	0.4851		
$\ln(k_{bi})$	-1.388	0.4658	-1.530	0.4754
$\ln(k_{bi})_p$	-1.388	0.3617	-1.530	0.3591
k_{bi}	0.2768	0.1335	0.2396	0.1047
$k_{bipcorr}$	0.2779	0.09996	0.2420	0.08358
$k_{bipuncorr}$	0.2661	0.09574	0.2304	0.07959

index p for predicted value; indices dh, 4, 5, 6 and bi for the different functions; index corr for corrected for logarithmic bias; index uncorr for not corrected for logarithmic bias.

The SPF was also fitted to the data (table 5) using SAS, PROC NLIN, METHOD=DUD (Anon. 1989). The minimum determined with the non-linear technique was found within a wide range of starting values for the parameters. XXXVFF since the exponent k is undefined when $\ln(d_{ub}/d_{bhub})=0$, i.e. ($d_{ub}=d_{bhub}$). Three observations above breast height had undefined ln(k) since their d_{ub} was larger than d_{bhub}. In order to compare SPF and VFF when fitted to the same dataset, observations not useful for the VFF were excluded also when fitting the SPF (table 6).

Table 5.
Parameters in SPF fitted to entire data.

Parameter	Estimate for spruce	Standard error	Estimate for pine	Standard error
b1	-3.8158	0.7046	-5.6198	1.7545
b2	1.8128	0.4082	2.7310	0.9785
b3	-2.0333	0.3729	-3.1194	0.9282
b4	254.4	26.3346	141.9	30.7873
a1	0.6778	0.0557	0.7611	0.0537
a2	0.0677	0.00328	0.0757	0.00757
n	2193		972	

All parameters are significant at α=0.05.

Table 6.
The precision of the fitted functions in explaining the variation in d_{ub}.

	n	Std.dev. d (mm)	RMSE (mm)	RMSE (%)
Spruce: SPF	1928	94.2	14.2	21.9
Spruce: VFFBi	1928	94.2	13.9	14.9
Spruce: VFFdh	1928	94.2	19.3	19.0
Spruce: VFF4	1928	94.2	18.9	17.7
Spruce: VFF6	1928	94.2	15.7	17.4
Pine: SPF	844	74.7	9.92	29.1
Pine: VFFBi	844	74.7	9.89	18.0
Pine: VFFdh	844	74.7	11.1	21.5
Pine: VFF4	844	74.7	11.0	24.0
Pine: VFF5	844	74.7	11.0	32.5

The precision of the functions was evaluated on the same observations (table 6). The VFFBi showed the lowest RMSE for spruce when all observations are considered (table 6) as well as when the pulpwood dimensions of the logs are considered (figure 5). The SPF showed comparable precision to VFFBi in nominal values, RMSE (mm), but worse in relative, RMSE (%). The VFF gave, generally, lower precision than SPF in the lower section of the trees. The ratio between the residual variation around the function, $RMSE^2$ and the variance of d_{ub}, std.dev^2, was 0.022 for VFFBi and 0.023 for SPF, which means that a large portion of the variation in d_{ub} can be explained with the functions.

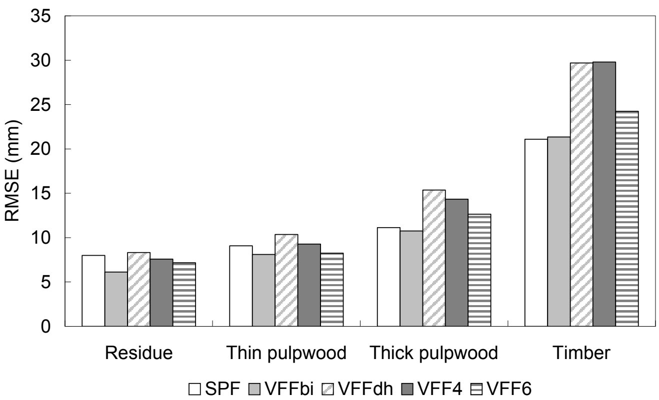


Figure 5.
RMSE of the relative residual ((observation-prediction)/observation) for spruce diameter under bark, by commercial dimension class. Residue where d < 50 mm, thin pulpwood where 50 ≤ d < 100 mm, thick pulpwood where 100 ≤ d < 150 mm, and timber where d ≥ 150 mm.

In the validation test, where the parameters were estimated on the calibration set and used to predict the validation set, the SPF function had the smallest bias, i.e. mean of the residuals, in the spruce pulpwood dimensions (table 7). Edgren & Nylinder's (1949) function gave slightly larger bias, but the standard deviations were comparable to those of the fitted functions. The VFFBi had a positive bias for all dimension classes. The residuals from VFFBi compared well to the other VFF functions, in that VFFBi gave lower bias than the other VFF functions for two commercial dimensions. The predicted diameters for two spruce trees are plotted to visualize the differences between VFFBi and SPF (figure 6). A cyclic fluctuation in VFFBi compared to SPF is apparent, which is not the case when Bi fitted the function to *Eucalyptus sp.* (Bi, 2000).

Table 7
Spruce validation results in terms of residuals by classes of commercial dimensions when predicting diameter under bark on validation data using parameters fitted to calibration data (the distribution with the lowest bias in bold, the best of the VFF distributions underlined).

	N	E&N bias (mm)	E&N std.dev (mm)	SPF bias (mm)	SPF std.dev (mm)	VFFBi bias (mm)	VFFBi std.dev (mm)	VFFdh bias (mm)	VFFdh std.dev (mm)	VFF4 bias (mm)	VFF4 std.dev (mm)	VFF6 bias (mm)	VFF6 std.dev (mm)
Residue (d<50 mm)	150	-3.59	6.12	-3.93	7.93	**0.52**	**6.15**	-2.50	8.04	-3.07	6.19	-3.20	6.18
Thin pulpwood (50 ≤ d < 100 mm)	164	-3.23	6.96	**0.84**	**9.66**	3.61	7.63	-3.68	9.95	-4.12	8.10	-2.82	7.69
Thick pulpwood (100 ≤ d < 150 mm)	154	-3.26	9.43	**-1.14**	**10.99**	2.09	9.43	-3.47	12.2	-3.88	10.5	-2.57	8.54
Timber (150 mm ≤ d)	275	10.8	27.8	-5.54	19.9	5.60	19.2	2.72	23.9	**2.45**	**23.6**	4.29	22.2
Total	743												

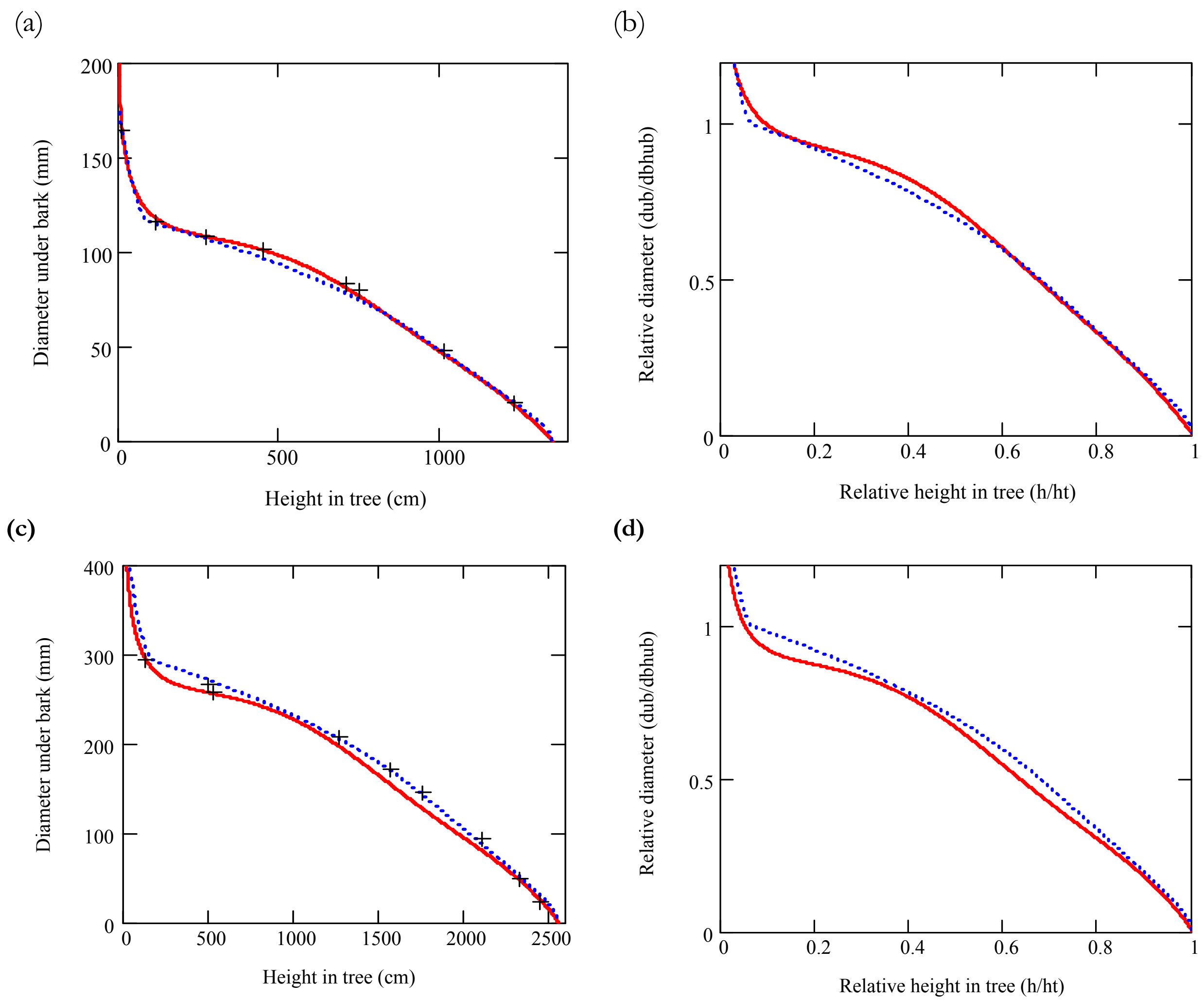


Figure 6.
Spruce, predicted taper with Bi's function (thicker line) and Max & Burkhart's function (dotted line) and observed values (plus sign) for diameter under bark for a small tree (ht=1350 cm, dbhpb=125 mm) in nominal values (a), in relative diameter and height (b) and for a large tree (ht=2552 cm, dbhpb=305 mm) (c and d).

For pine, VFFBi and SPF had similar RMSE, slightly lower than the other VFF (table 6 and figure 7). Just as for spruce, the SPF was inferior to VFFBi in the residue dimensions, but similar or better in the commercial dimensions. The ratio between the residual variation around the function, $RMSE^2$ and the variance of d_{ub}, $Std.dev^2$, was 0.018 for both SPF and VFFBi, which means that a large portion of the variation in d_{ub} can be explained with the functions. The ratio was slightly smaller than for spruce, which means that the functions were more successful in explaining the variation in pine taper than for spruce trees, but on the other hand the variation in the data was larger for spruce than for pine (table 1).

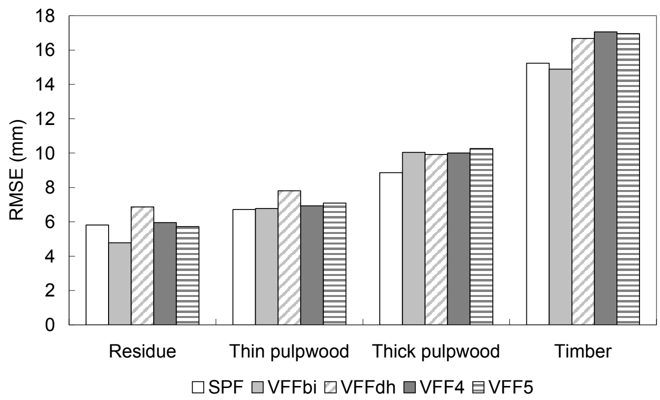


Figure 7.
RMSE (mm) measured on the common residuals from the fitted functions for pine diameter under bark, by commercial dimension class. Residue where d < 50 mm, thin pulpwood where 50 ≤ d < 100 mm, thick pulpwood where 100 ≤ d < 150 mm, and timber where d ≥ 150 mm.

RMSE (mm) measured on the common residuals from the fitted functions for pine diameter under bark, by commercial dimension class. Residue where d < 50 mm, thin pulpwood where 50 ≤ d < 100 mm, thick pulpwood where 100 ≤ d < 150 mm, and timber where d ≥ 150 mm.In the validation, the P7 and SPF functions gave the lowest bias for the residuals by commercial dimensions (table 8). The fitted VFF in this study could not predict the validation data better than the VFF fitted by Petersson, i.e. P7. Again, the E&N function gave slightly larger bias, but the standard deviations were comparable to those of the fitted functions. The residuals from VFFBi were not convincingly better distributed than other VFF functions.

Table 8.
Pine validation results in terms of residuals by classes of commercial dimensions when predicting diameter under bark on validation data using parameters fitted to calibration data (the distribution with the lowest bias in bold, the best of the VFF distributions underlined).

	n	E&N bias (mm)	E&N std.dev (mm)	P7 bias (mm)	P7 std.dev (mm)	SPF bias (mm)	SPF std.dev (mm)	VFFBi bias (mm)	VFFBi std.dev (mm)	VFFdh bias (mm)	VFFdh std.dev (mm)	VFF4 bias (mm)	VFF4 std.dev (mm)	VFF5 bias (mm)	VFF5 std.dev (mm)
Residue (d<50 mm)	47	-5.18	6.35	**0.51**	**6.59**	-3.60	5.50	<u>-1.84</u>	<u>5.03</u>	-2.28	5.40	-2.83	5.08	-2.49	4.56
Thin pulpwood (50 ≤ d < 100 mm)	57	-3.90	6.83	-2.26	8.88	**0.17**	**7.22**	1.11	7.45	<u>-0.33</u>	<u>7.34</u>	-1.46	6.31	0.90	7.58
Thick pulpwood (100 ≤ d < 150 mm)	65	1.53	7.47	**-1.07**	**8.51**	1.60	7.79	3.06	8.57	2.64	7.77	<u>1.35</u>	<u>6.68</u>	3.08	8.44
Timber (150 mm ≤ d)	87	7.76	13.36	2.82	13.1	**0.35**	**12.1**	5.84	12.4	5.84	14.7	<u>4.86</u>	<u>15.0</u>	5.59	14.8
Total	256														

Discussion

The objective of the present study is to establish functions that predict the diameter under bark at a given height in the stem. This objective has been reached by the fitted SPF functions (eq. SPF, table 5), which are well fitted to the data, i.e. has low RMSE, and have good predictive capacity as shown by the small bias in the validation. The VFFBi had slightly lower RMSE but showed an inferior predictive capacity in the validation. This suggests that the VFFBi is too flexible in its fit to the data, giving over-fitted parameters that are less capable of predicting validation data. The VFFBi also predicts a fluctuating shape of the stem compared to the SPF. Bi's own use of the function on *Eucalyptus sp.* (Bi, 2000) doesn't show such a fluctuation, but rather a smooth trend as for SPF. The RMSE for diameter under bark in Bi's study ranges from 4.4 to 17 mm, with a mean of 9.2 mm. This agrees with the fit on pine in this study, but the fit on spruce gives a higher RMSE than Bi's. Until the variable-form functions, which conceptually are preferable with their adaptation to the individual tree, are fitted more successfully to Swedish data, the fitted SPF gives better predictions of diameter under bark than the previously used functions of Edgren & Nylinder (1949).

The VFF has been fitted to Swedish pine (Petersson, 1999). The SPF in this study gave predictions similar to those of Petersson's varible-form function, here called P7. Peterson's P7 with seven variables in the sub-function for ln(k) had a RMSE of 7.8 mm , which is lower than the RMSE≈10 of SPF and VFFBi, and RMSE≈11 of the other variable-form functions in this study. That many variables in the sub-function were not deemed justifiable in this study, since the improvement in R^2 of the sub-functions dropped as the number of variables increased. Petersson also developed a segmented function for pine, but since its predictive properties were similar to his variable-form function, the segmented function was not tested here.

The negative aspect of the SPF is its generalization that the relation between relative height and relative diameter is the same for all trees (see also discussion about parameter a_2 below) . This is apparent as the predicted d_{ub} at breast height differs from d_{bhub}, (figure 8). In spite of this increase in bias in breast height as the tree height increases, the SPF shows smaller bias in the timber dimensions than the other functions tested, so the SPF is yet to be challenged by a more flexible function.

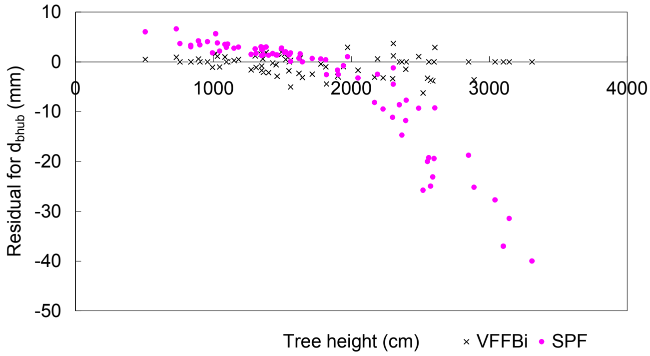


Figure 8.
Residuals for d_{bhub} from predictions on spruce validation data.

The VFF was expected to outperform the SPF, since the variable-form function has been shown to give better predictions than SPF in studies in North America (Newnham, 1992). This was not the case in the present study, where the SPF gave lower RMSE than VFFdh, VFF4 , VFF5 and VFF6, for both Norway spruce and Scots pine. In the VFF alternatives with 4 or more variables in the sub-function, crown variables were allowed as input. The choice of variables turned out to include variables related to crown length in the best VFF models for spruce, but not for pine. This is not surprising as pine generally has a smaller variation in crown height than spruce. The parameter a_2 in SPF, describing the join point between the neiloid butt section, and the paraboloid middle section of the stem, was estimated to 7% of the stem height for both spruce and pine. This is lower than in southern USA (Max and Burkhart, 1976), where naturally regenerated stands of *Pinus taeda* had a_2 of roughly 9%, while plantation stands of *P taeda* had a somewhat higher a_2 of 11%. The upper join point, a_1, of Scots pines in the present study lies between the values for naturally regenerated and plantation stands of *P taeda*. For Norway spruce, the upper join point was lower than for the pine species, which seems reasonable considering the slower natural branching of spruce.

A variable-form function has been fitted to Norway spruce in France (Saint-André, 1999) with a RMSE=15 mm (where RMSE is the square root of the quota between sum of squared residuals and n, rather than (n-p) as in the present study). This fit is in northern France on a small sample of 24 trees, but the result agrees with those presented in this paper, and does not indicate that the fit of a variable-form function could be substantially improved when it comes to predicting the diameter of Norway spruce.

Two instruments were used to measure diameter under bark, the caliper in the field and the scanner in the lab. This is a source of error in the data, since there might be a systematic difference between the instruments and different measurement errors associated with each instrument. These cons were, as we believed in the project group, balanced by the pros that the data could be increased at a relatively low cost. The number of discs that were collected was limited by the cost of measuring the wood and fiber properties of each disc, which is a part of the other activities in the SMP project.

Conclusion

The SPF were the most suitable functions to predict the stem taper of trees whose d_{bhub} and tree height, h_t, were known. These functions, one for Norway spruce and one for Scots pine, were fitted to, and validated with, data covering both species in Swedish stands within the range of latitude 57° to 68° and elevations up to slightly over 400 m above see level. The VFF or VFFbi for spruce and pine may be useful alternatives to the SPF as they resulted in almost comparable precision. The SPF seems to give slightly better predictions of d_{ub} than the frequently used functions developed by Edgren & Nylinder (1949) when these functions were used with estimated, not measured diameter ratios D60/D20. The steps in predicting a diameter using the VFF functions are the following: predict the ln(k) at the given height using the functions in tables 2 or 3, then transform the ln(k) to k while compensating for logarithmic bias using the equation for k_{pcorr}, (1), and use the k_{pcorr} , or $k_{Bipcorr}$, to predict the d_{ub} using the equation (2) or (3) respectively.

Acknowledgements

This study is part of the *Skog-Massa-Papper* (or Forest-Pulp-Paper) project, which is a joint project between SkogForsk and STFI, and funded by NUTEK's OptiTrä-program. Several persons at the two institutes and elsewhere have been involved in the collection, preparation and measurements of the samples, and the structuring of the database. We thank particularly Hans Fryk, Johan Hägerdahl, Lars Olsson, Hilding Ekman, Thomas Grahn, and the staff at Hemglass where we keep our frozen samples.

References

Anon. 1989. SAS/STAT(r) User's Guide, Version 6, Fourth Edition. 846. SAS Institute Inc., Cary. ISBN 1-55544-376-1.

Arlinger, J., Wilhelmsson, L., Lundqvist, S.-O. and Hedenberg, Ö. 2001. Vedprovtagning - stragi, dokumentation och utfall. Delrapport projektet Skog-Massa-Papper . Arbetsrapport 487, SkogForsk (In Swedish).

Bi, H. 2000. Trigonometric Variable-form Taper Equations for Australian Eucalypts. For. Sci. 46(3): 397-409.

Brownlee, K. A. 1965. Statistical theory and methodology in science and engineering. 590 pp. John Wiley and sons, New York.. ISBN 0-471-11355-7.

Edgren, V. & Nylinder, P. 1949. Funktioner och tabeller för bestämning av avsmalning och formkvot under bark för tall och gran i norra respektive södra Sverige. Meddel. från Statens skogsforskningsinstitut. 38(7): 25. (In Swedish with English summary)

Kozak, A. 1988. A variable-exponent taper equation. Can. J. For. Res. 18(11): 1363-1368.

Kozak, A. 1997. Effects of multicollinearity and autocorrelation on the variable-exponent taper functions. Can. J. For. Res. 27(5): 619-629.

Kozak, A. & Smith, J. H. G. 1993. Standards for evaluating taper estimating systems. Forestry Chronicle 69(4): 438-444.

Larson, P. R. 1969. Wood formation and the concept of wood quality. Yale University, School of Forestry. Bulletin 74, 54 pp.

Lindström, H. 1996a. Basic Density in Norway spruce, Part II. Predicted by stem taper, mean growth ring width and factors correlated to crown development. Wood and Fiber Science 28(2): 240-251.

Lindström, H. 1996b. Basic density in Norway spruce. Part III. Development from pith outwards. Wood and Fiber Science 28(4): 391-405.

Lindström, H. 1997. Fiber length, tracheid diameter, and latewood percentage in Norway spruce: development from pith outwards. Wood and Fiber Science 29(1): 21-34.

Lundqvist, S.-O., Grahn, T., Hedenberg, Ö. & Olsson, L. 1998. Fibertyper i barrved. Mätmetoder och presentationssätt. STFI, Stockholm. STFI-rapport P19: 19 pp. (In Swedish.)

Lundqvist, S O. Ekenstedt, F. Grahn, T. Hedenberg, Ö. Olsson, L. Arlinger, J. Wilhelmsson, L.. Evans, R. 2001. Variability in Fiber Properties of Norway Spruce of Importance for Mechanical Pulping. In: Proceedings International mechanical pulping conference, Helsinki, Finland, 4-8 June 2001, vol. 2 B, pp 9-18. ISBN:951-97513-7-8

Max, T. A. & Burkhart, H. E. 1976. Segmented polynomial regression applied to taper equations. Forest Science 22(3): 283-289.

Morén, A.-S. & Perttu, K. L. 1994. Regional temperature and radiation indices and their adjustment to horizontal and inclined forest land. Stud. For. Suec. 194.

Newnham, R. M. 1992. Variable-form taper functions for four Alberta tree species. Can. J. For. Res. 22(2): 210-223.

Petersson, H. 1999. A segmented stem profile model for Pinus sylvestris. Forest Ecology and Management 124(1): 13-26.

Saint-André, L., Leban, J.-M., Houllier, F. & Daquitaine, R. 1999. Comparaison de deux modèles de profil de tige et validation sur un échantillon indépendant. Application à l'épicéa commun dans le nord-est de la France. Ann. For. Sci. 56: 121-132. (In French with English summary.)

Wilhelmsson, L., Arlinger, J., Spångberg, K., Lundqvist, S.-O., Hedenberg, Ö., Olsson, L. & Grahn, T. Models for predicting wood properties in stems of Picea abies and Pinus sylvestris in Sweden. Scand. J. For. Res. (In press). 22 pp.