Silva Fennica 1989, Vol. 23 N:o 4: 311–318

Use of the Weibull function in estimating the basal area dbh-distribution

Pekka Kilkki, Matti Maltamo, Reijo Mykkänen & Risto Päivinen

TIIVISTELMÄ: WEIBULL-FUNKTION KÄYTTÖ POHJAPINTA-ALAN LÄPIMITTAJAKAUMAN ESTIMOINNISSA

Kilkki, P., Maltamo, M., Mykkänen, R. & Päivinen, R. 1989. Use of the Weibull function in estimating the basal area dbh-distribution. Tiivistelmä: Weibull-funktion käyttö pohjapinta-alan läpimittajakauman estimoinnissa. Silva Fennica 23(4): 311–318.

The paper continues an earlier study by Kilkki and Päivinen concerning the use of the Weibull function in modelling the diameter distribution. The data consist of spruces *(Picea abies)* measured on angle count sample points of the National Forest Inventory of Finland. First, maximum likelihood estimation method was used to derive the Weibull parameters. Then, regression models to predict the values of these parameters with stand characteristics were calculated. Several methods to describe the Weibull function by a tree sample were tested. It is more efficient to sample the trees at equal frequency intervals than at equal diameter intervals. It also pays to take separate samples for pulpwood and saw timber.

Tutkimus on jatkoa Kilkin ja Päivisen männyn läpimittajakaumaa käsitelleelle tutkimukselle. Aineistona olivat valtakunnan metsien inventoinnin kuusikoepuukoealat. Ensi vaiheessa Weibull-funktion parametrit estimoitiin suurimman uskottavuuden menetelmällä. Toisessa vaiheessa laadittiin regressiomallit Weibull-parametrien ennustamiseksi metsiköstä yleisesti arvioitavien tunnusten avulla. Kuvattaessa läpimittajakaumaa otoksen avulla jokaisen puun pitäisi edustaa yhtä suurta summafrekvenssiä eikä yhtä suurta läpimitan vaihtelualuetta. Otos kannattaa poimia erikseen tukkipuille ja kuitupuille, jolloin niiden osuuksiin ei tule otannasta aiheutuvaa virhettä.

Keywords: diameter, distribution, Weibull function, *Picea abies,* maximum likelihood, estimation.
ODC 521.6

Authors' address: University of Joensuu, Faculty of Forestry, Box 111, SF-80101 Joensuu, Finland.

Accepted November 30, 1989

List of symbols

d	=diameter at breast height, cm
d_{min}	=minimum diameter of the sample trees on a sample plot, cm
T	=age, years
N	=number of trees per sample plot
G	=basal area, m^2/ha
d_{gM}	=basal area median diameter of the sample trees, cm
a,b,c	=Weibull parameters
s_e	=residual standard error of the regression model

1. Introduction

Stem-diameter distribution models are needed to replace or support empirical diameter distributions based on tree measurements. In the case of relative intensive tree tally one can make use ot the flexibility of the non-parametric methods like kernel-estimation (Droessler and Burk 1989) and percentile method (Borders et al. 1987).

In Finnish conditions, inventory by compartments is carried out without any tree tally. In large-area sampling inventories only small plots are employed and thus the number of measured trees remains low for each stand in the sample. In both cases, stem-diameter distribution models can be utilized as a means to obtain the best estimate for the tree size distribution for the growing stock.

A great number of distribution functions have been applied to the description of the diameter distributions of forest stands. Cajanus (1914) introduced the use of the Gram-Charlier series for the description of diameter distributions. He also demonstrated the potential of theoretical diameter distributions in the preparation of yield tables. One of the most popular models has been the Weibull function introduced by Bailey and Dell (1973). The popularity of the Weibull function depends largely on its simplicity and yet relatively good flexibility.

The probability density function for the Weibull random variable x is

$$f(x) = \frac{c}{b}\left(\frac{x-a}{b}\right)^{c-1} \exp\left(-\left(\frac{x-a}{b}\right)^{c}\right) \quad (a \leq x \leq \infty) \qquad (1)$$

$$= 0 \quad (x<a)$$

where
a,b,c = parameters

and the cumulative distribution function of the Weibull is

$$F(x) = 1 - \exp\left(-\left(\frac{x-a}{b}\right)^{c}\right) \quad (a \leq x \leq \infty) \qquad (2)$$

$$= 0 \quad (x<a)$$

This paper continues a study by Kilkki and Päivinen (1986), in which the trees measured from angle count points in the Finnish National Forest Inventory (NFI) were used to estimate the Weibull function parameters for Finnish pine *(Pinus sylvestris)* stands. The present data refer to spruce *(Picea abies)* stands drawn from the Finnish NFI.

The purpose of the paper is to provide methods for the derivation of the theoretical set of tree diameters and their frequencies in the case when only stand characteristics like basal area, median basal area diameter and age are known. It is assumed that the derived distributions are used separately for pine and spruce in mixed stands. The efficiencies of the different methods to select the sample from the estimated diameter distribution are also tested.

2. Material and methods

Data were drawn from the seventh Finnish National Forest Inventory (Kuusela and Salminen 1980). The data consisted of all sample tree angle count points with a least five spruces belonging to the dominant crown storey. Only spruces were included in the data. Less than ten sample plots were rejected during the data processing, because their trees evidently did not represent the same crown storey.

In the final data there were 1746 angle count points with 14416 trees. The data are mainly from spruce stands in southern and central Finland; only two angle count points were from northern Finland. The number of spruces per angle count point varied from 5 to 23 and the average number was 8.3. The average basal area was 16.6 m^2/ha. The arithmetic mean diameter of the trees was 22.3 cm, the minimum diameter 2.2 cm, and the maximum diameter 63.0 cm.

Fig. 1 presents an example of the empirical cumulative basal area distribution.

The Weibull parameters for each set of diameters were estimated using the maximum likelihood method. The natural logarithm of the likelihood function were solved using the ZXMWD-subroutine of the IMSL-library (IMSL... 1984). The initial values for the parameters were generated by the library program. The allowed ranges for the parameters were

$$0.5 \cdot 0.3^{(1/N)} d_{min} \leq a \leq 0.3^{(1/N)} d_{min}$$

$$1 \leq b \leq 50$$
$$1.1 \leq c \leq 15.$$

The basal area median diameter estimates were calculated using formula

$$d_{gM} = a + b(-\ln(0.5))^{1/c} \qquad (3)$$

If the basal area median diameter is known, only two of the Weibull parameters need to be predicted in the application phase of the Weibull function. Kilkki and Päivinen (1986) chose parameters a and c for estimation and parameter b to be solved from the formula

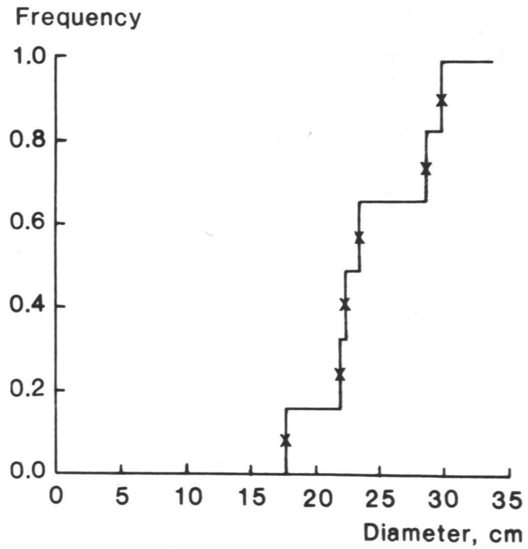

Fig. 1. Example of an empirical cumulative dbh basal area distribution (x = cumulative frequencies used in the nonlinear regression analysis as the values of the dependent variables).

$$b = \frac{d_{gM}-a}{(-\ln(0.5))^{1/c}} \qquad (4)$$

In this study, the regression models were derived for parameter b instead of c. Then, parameter c can be solved from the formula

$$c = \frac{\ln(-\ln(0.5))}{\ln((d_{gM}-a)/b)} \qquad (5)$$

Weighted regression analysis was used. The weights corresponded to the inverse of the residual variances of the approximate Weibull cumulative distribution function with parameter estimates from the nonlinear regression analysis. The parameters were estimated by stepwise regression analysis. The predicting variable was rejected on the 0.01 risk level.

3. Results

The regression models for predicting the Weibull parameters for basal area diameter distribution are:

$$a = 0.001389 \quad (t = 0.0)$$
$$0.517444\, d_{gM} \quad (t = 32.0) \qquad (6)$$
$$(s_e = 3.78)$$

$$\ln(b) = -0.346223 \quad (t = 4.4)$$
$$0.934993\, d_{gM} \quad (t = 34.5)$$
$$-0.000925\, G \quad (t = 3.3) \qquad (7)$$
$$(s_e = 0.298)$$

If the minimum diameter of a stand is measured, regression model only for parameter b needs to be estimated. Then the minimum diameter can be an independent variable and the regression model for parameter b is the following:

$$b = 0.629537 \quad (t = 7.3) \qquad (8)$$
$$1.050618\, d_{gM} \quad (t = 305.2)$$
$$-1.020776\, a \quad (t = 247.5)$$
$$0.014405\, G \quad (t = 5.6)$$
$$-0.001986\, T \quad (t = 3.5)$$
$$(s_e = 0.608)$$

Due to the logarithmic transformation, Taylor's series correction has to be made when applying model (7). Similarly, Taylor's series correction has to be made when parameter c (formula 5) is derived.

The reliability of the following methods in the derivation of the Weibull parameters was tested:

Medhod 1: Maximum likelihood estimation

Method 2: Parameter a with maximum likelihood estimation, b with a regression model

Method 3: Parameters a and b with regression models

The estimates for the first, third, and fourth powers of the dbh's were compared with their values derived from the original diameter measurements. The mean errors ($\bar{x}$) and root mean square errors (RMSE) are:

	Σ d	Σ d %	Σ d
Method 1			
$\bar{x}$	0.69	-0.11	-0.08
RMSE	1.39	0.20	0.48
Method 2			
$\bar{x}$	0.54	.20	1.03
RMSE	2.49	2.67	8.57
Method 3			
$\bar{x}$	0.53	.38	1.49
RMSE	3.80	2.97	9.10

Some tests were made with Weibull functions using diameters with opposite signs. The idea of this approach refers to the possibility to measure the maximum diameter and use it directly as parameter a. This method gave in some cases slightly better results in the estimation of the higher moments.

After the theoretical diameter distribution has been estimated by predicting the Weibull parameters with stand characteristics, a set of tree diameters and their frequencies has to be sampled from the distribution. These data are used in the derivation of the stand characteristics, such as volume, saw timber volume, etc.

The efficiency of three different sampling methods was tested. The Weibull distributions were defined by the maximum likelihood Weibull parameter estimates of the data. Altogether 1746 sets of Weibull parameters were available.

In the first method (A) each tree represents an equal diameter interval (Fig. 2 a). In the second method (B) each tree represents an equal frequency (Fig. 2 b). The third method (C) corresponds to the method A except for the fact that the samples are taken separately for saw timber (d > 18 cm) and pulpwood (0<d≤18 cm).

The standard errors of the estimates of the diameter powers 1, 3, and 4 were used as measures of precision. The "true" values were obtained by taking a sample at 1 mm

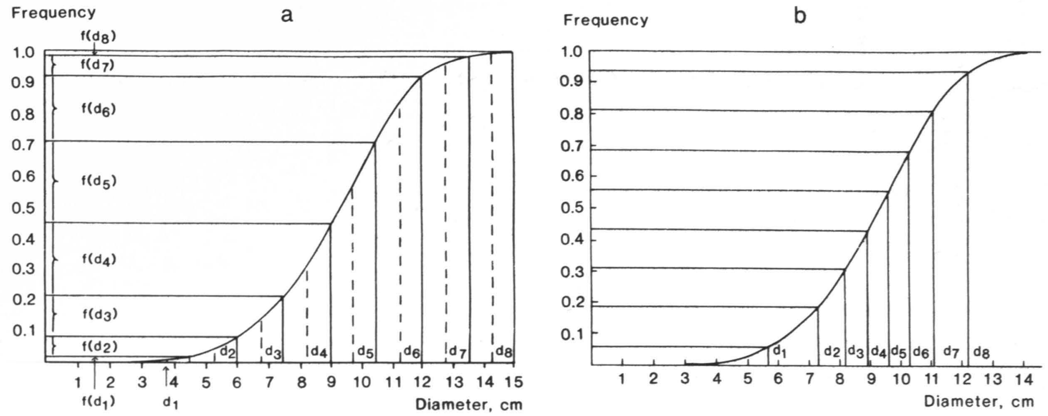

Fig. 2. a) Sampling a Weibull distribution using equal diameter intervals (method A). b) Sampling a Weibull distribution using equal cumulative frequency intervals (method B).

intervals from the minimum diameter up to the diameter of 1000 mm. The number of sample diameters varied from 5 to 30. Only the samples with both saw timber and pulpwood were accepted. Consequently the number of diameter distributions was reduced to 1022.

The efficiency of method A depeds on the minimum and maximum diameters at which the cumulative distribution function is cut. The minimum diameters corresponding to the sum frequencies 0.01, 0.001, and 0.0001, and the maximum diameters corresponding to the sum frequencies 0.99, 0.999, and 0.9999 were respectively tested. The results are given in Fig. 3.

When the diameter limits are set to the sum frequencies 0.01 and 0.99 method B is superior except with the fourth power and with a sample size of less than 7. When the diameter limits corresponding to the sum frequencies 0.001 and 0.999 or 0.0001 and 0.9999 are used, method A performs a little better with large samples. Separate sampling of saw timber and pulpwood (method C) generally impairs the results.

The performance of methods A and B in the estimation of the saw timber (d > 18 cm) proportion was also studied (see Fig. 4). In this test, 1635 distributions with both saw log size and pulpwood size trees were accepted. Method B was superior to method A for all diameter limits.

In conclusion, it can be said that if the number of systematically sampled diameters is relatively small method B should be used. The choice between methods A and C depends on the emphasis laid on the precision requirements of the saw timber estimate vs. the moment estimates. If the number of sampled diameters is large (>50), the results from all methods are equally good. In methods A and C, however, the percentiles corresponding to the minimum and maximum diameters should be small.

It appears that the error of the results increases markedly when the value of the parameter c is small. Thus, the size of the sample might be adjusted with the value of parameter c.

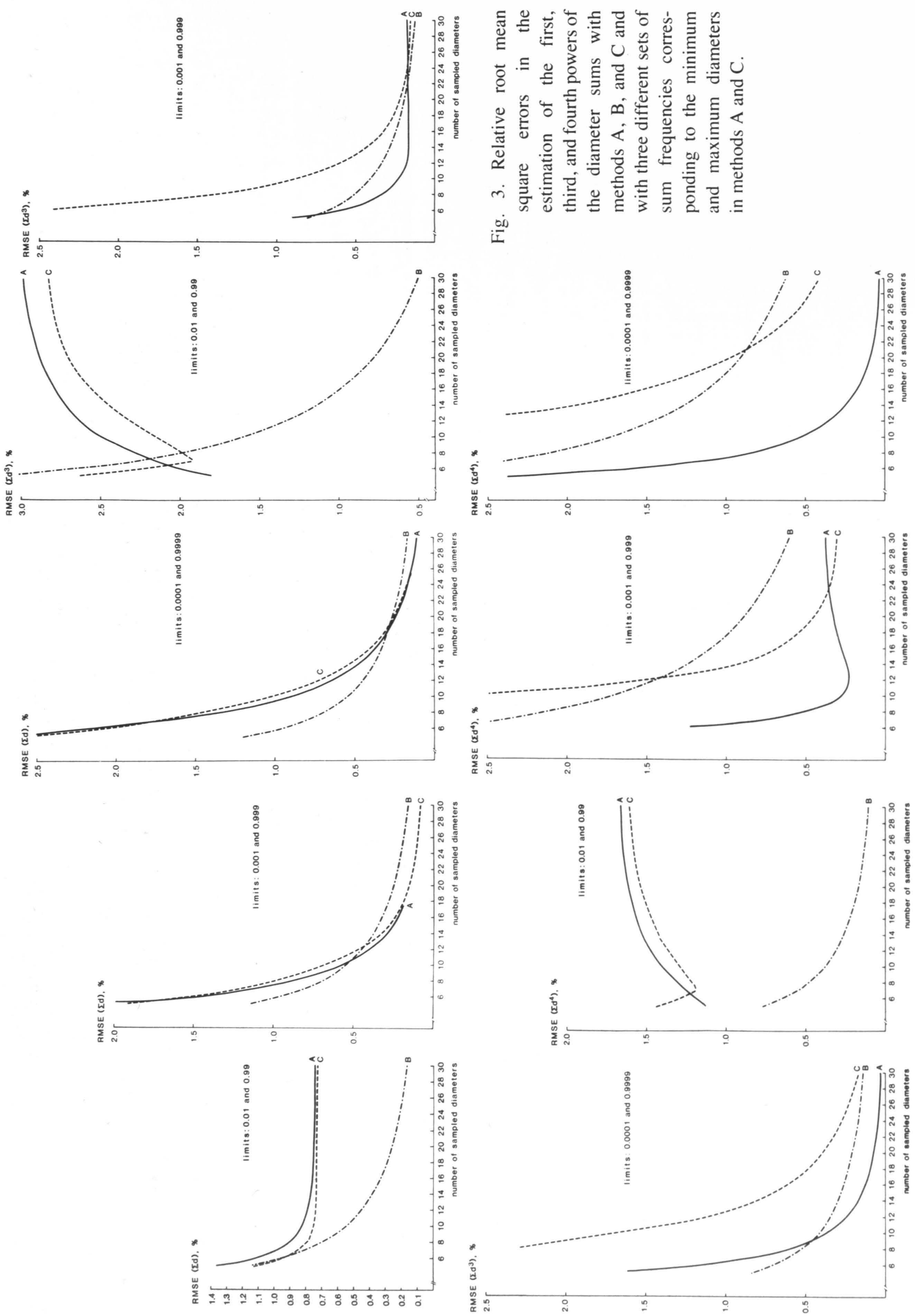

Fig. 3. Relative root mean square errors in the estimation of the first, third, and fourth powers of the diameter sums with methods A, B, and C and with three different sets of sum frequencies corresponding to the minimum and maximum diameters in methods A and C.

4. Discussion

The aim of the study was to provide methods for estimation of the spruce stand diameter distribution, when only stand characteristics are known. Weibull function was selected to describe the basal area diameter distribution. It was found that very few stand variables in addition to those directly describing the diameter distribution are of value in the estimation of the Weibull parameters. When applying the method, it must be kept in mind that measurement or estimation errors on the stand variables used as predictors increase the error in the Weibull parameter estimates.

Attention should be paid to the method used when the diameters are sampled from the theoretical diameter distribution function. When the number of sampled diameters is relatively small, systematic samples with equal probabilities for each observation seem the most efficient. In addition to the separate samples for saw timber and pulpwood size trees it might be appropriate to take separate samples for waste wood size trees, too.

One possibility to apply the Weibull distribution is to use both the basal area and the number of trees as input variables. If the basal area median diameter (d_{gM}) and the minimum diameter (a) or the saw timber percentage are also known, all Weibull parameters can be derived analytically. This method could be applied in the field in order to avoid unrealistic combinations of the basal area, number of trees, basal area median diameter, minimum diameter, and saw timber percentage in field estimates. A field computer could tell us whether a certain combination of values is consistent with any Weibull function.

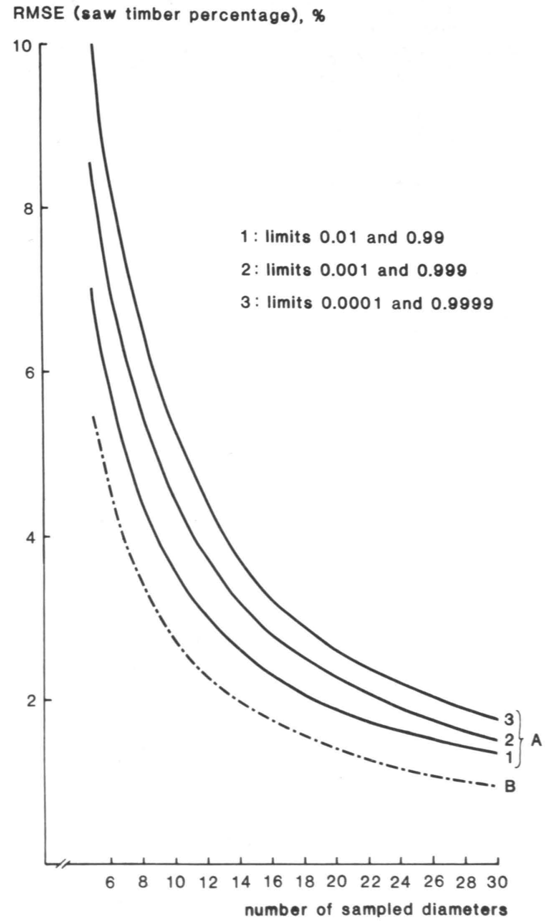

Fig. 4. Mean root square errors of the saw timber percentage with method A (with three sets of minimum and maximum diameters) and method B.

References

Bailey, R.L. & Dell, T. R. 1973. Quantifying diameter distribution with the Weibull-function. Forest Science 19(2): 97–104.

Borders, B.E., Souter, R.A., Bailey, R.L. & Ware, K.D. 1987. Percentile-based distributions characterize forest stand tables. Forest Science 33(2): 570–576.

Cajanus, W. 1914. Über die Entwicklung gleichaltriger Waldbestände. Eine statistische Studie. I. Acta Forestalia Fennica 3. 142 p.

Droessler, T.D. & Burk, T.E. 1989. A test of nonparametric smoothing of diameter distributions. Scandinavian Journal of Forest Research 4: 407–415.

IMSL Library, user's Manual. 1984. International Mathematic and Statistical Libraries INC., Houston, Texas.
Kilkki, P. & Päivinen, R. 1986. Weibull-function in the estimation of the basal area DBH-distribution. Silva Fennica 20(2): 149–156.
Kuusela, K. & Salminen, S. 1980. Ahvenanmaan maakunnan ja maan yhdeksän eteläisimmän piirimetsälautakunnan alueen metsävarat 1977–1979. Summary: Forest resources in the province of Ahvenanmaa and the nine southernmost Forestry Board Districts in Finland 1977–1979. Folia Forestalia 446. 90 p.

Total of 7 references