Annals of Forest Science (2020) 77: 47
https://doi.org/10.1007/s13595-020-00951-3

RESEARCH PAPER

Performances of different methods of estimating the diameter distribution based on simple stand structure variables in monospecific regular temperate European forests

Jean-Philippe Schütz[1] · Christian Rosset[2]

Received: 4 October 2019 / Accepted: 6 April 2020 / Published online: 18 May 2020

Abstract

• ***Key message*** **Using the three characteristic points of a forest stand, dg (mean quadratic diameter), d_{min} (diameter of the smallest tree) and d_{max} (diameter of the largest tree), appears informative enough to determine the parameters of the whole diameter distribution and, hence, the standing volume, with an accuracy of 2–3%. This is related principally with a particular feature of the Weibull distribution function, and the empirical dependency of the main scale parameter $\alpha + \beta$ from the mean quadratic diameter (dg): This allows the prediction of the parameter β with an unexpectedly high likelihood. This feature could be used for growth modelling as well as inventory purposes, at least for monospecific and even-aged stands and, maybe more, because this feature is proper to the function itself.**

• ***Context*** One of the most appealing applications of diameter distribution functions is to predict compliant stand diameters without needing to tally all stems, but in determining the function parameters only on the base of simple stand characteristics. This can be applied for yield model construction or inventory purposes.

• ***Aims*** The aim of this paper is to compare different methods of estimating the Weibull distribution parameters, partly based on parameter recovery method (PRM). They use a remarkable, empiric property of the Weibull function. Their performances are assessed in comparison to real distributions from a wide database of permanent Swiss yield plots repeatedly measured (time series) for Norway spruce (*Picea abies* (L.) Karst.) and European beech (*Fagus sylvatica* L.).

• ***Methods*** The Weibull distribution offers the advantage of simple but reliable estimation procedures. One of these is the main (scale) parameter β being given at a remarkable point of the function free, i.e. independent from the shape parameter γ. Because dg lies very close to this point, it correlates empirically very tightly with this parameter and thus allows for a trustful simple estimation. We compare, with appropriate statistic tests, real distribution with such obtained with the usual maximum likelihood estimation (MLE) of the Weibull parameters and those obtained with these new procedures.

• ***Results*** The results obtained from a set of 800 yield plots of regular spruce stands and 596 of beech in Switzerland illustrate the good performances of the two much simpler procedures. The accuracy of estimating the standing volume is about 1.4% for beech and 2.8% for spruce when the site index (SI) is known.

• ***Conclusion*** The three considered characteristic stand attributes (dg, d_{min} and d_{max}) appear robust enough for determining the diameter distribution with a respectable accuracy. This is enough reason for a revival of the old, but very ingenious, method of angle count sampling of Walter Bitterlich (1947).

Keywords Parameter estimation · Stand diameter · Weibull distribution · Growth modelling · Standing volume estimation

Handling Editor: Erwin Dreyer

This article is part of the topical collection on *Mensuration and modelling for forestry in a changing environment*

✉ Jean-Philippe Schütz
jph.s@bluewin.ch

Christian Rosset
Christian.Rosset@bfh.ch

[1] Department Environmental Systems Science, ETZ Zürich, Brüggliäcker 37, 8050 Zurich, Switzerland

[2] University of Applied Sciences, Zollikofen, Switzerland

1 Introduction

One of the most appealing applications of distribution functions is prediction of compliant stand diameter distributions, without needing to tally all stems but in determining the function parameters only on the base of simple stand characteristics. As predictors, we can use, for instance, the mean dbh (mean diameter at breast height), d_{max} (diameter of the largest stem), d_{min} (diameter of the smallest stem of the cohort) and dg (mean quadratic diameter), which are easily identifiable in the field. This can be applied to yield model construction (Borders et al. 1987; Schütz and Rosset 2016) or for inventory purposes. It allows for a simple and, therefore, inexpensive assessment and accurate estimations of the standing volume, for instance when G (basal area) is determined with the angle count method (Bitterlich 1947, 1984) and N (total stem numbers), with a smart phone application like MOTI (www.moti.ch).

The Weibull density function (Weibull 1939; Bailey and Dell 1973; Dubey 1967) is mainly used for such parameter estimation because of its high flexibility and some very interesting features, particularly the independency of parameters β and γ, at least at one particular position (Dubey 1967; Zanakis 1979). The reduction to the two-parameter version of the cumulative density Weibull function allows for great simplification in parameter estimation.

As a matter of principle, different methods of parameter estimation could be considered from an analytical mathematical approach—the so-called 'parameter recovery method' (PRM) after Hyink and Moser (1983)—to a more empiric-based method, by regressing the parameters on stand structure indicators from a reliable database (so-called 'parameter prediction method' (PPM)) after Clutter and Bennett (1965). Even other ways, i.e. non-parametric approximation methods, could be used, i.e. those with the Newton-Raphson algorithm. Decisive is here not the method itself, but rather the goodness of fit over the whole distribution or the accuracy in stand volume estimation.

The aim of the presented paper is to compare different methods of parameter estimation. For estimating the parameter β, we use a remarkable property of the Weibull function as well as the empiric relationship between $\alpha + \beta$ and dg. The parameter γ is derived with parameter recovery methodology. The performances are assessed in comparison to real measured distributions from a wide database of permanent Swiss yield plots repeatedly measured (time series) for Norway spruce and European beech emanating from 800 inventory for Norway spruce and 596 for European beech, and finally, as a comparison benchmark, we use the results of parameter estimation of these real stem diameter distributions fitted with usual statistical methods.

2 Material and methods

2.1 Data

The data for comparing diameter distributions are diameter lists of 800 permanent yield plots of even-aged and monospecific spruce and 597 beech stands, repeatedly measured over a long time (1882–2013), from the yield database of the Growth and Yields Research Group of the Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland. These permanent WSL yield plots for two species (at 219 sites for spruce and 138 sites for beech) are distributed all over Switzerland, with the exception of sites above 1200 m above sea level. The size of each plot is around 0.25 ha (in some case, as much as 0.5 ha). All trees are identified by a number. Measurements are recorded every 5–7 years after a thinning intervention. The dbh of all trees of the whole cohort, i.e. excluding second growth or understorey, is accurately recorded crosswise with a calliper to mm accuracy. The height is recorded on a sample of 20–40 trees, with a JAL-type hypsometer with a 7-m reference rod and an accuracy of about 0.8 m (Schmid et al. 1971) and, after 1990, with a Vertex hypsometer. Thinning interventions in WSL plots correspond generally to conventional practice in terms of thinning type and intensity, i.e. thinning from below until about 1940 and selective thinning according to Schädelin (1934) after 1940.

Figure 1 shows the range of ages, stand densities (with the stand density index (SDI) = G_{stand} / G_{max}, where G_{max} is determined after Reineke's size density rule, after Schütz and Zingg 2010) and site indices (hdom at age 50) of the permanent yield plots considered. The reason for the difference in stand density between spruce and beech is that the database for beech contains more thinning trials with different density variants. The stem numbers per plot range from 30 to 2320, with an average of 404 for spruce, and from 20 to 3040, with an average of 287 for beech.

2.2 Methodology

The Weibull probability density function reads

$$f(x) = \gamma/\beta \left(\frac{x-\alpha}{\beta}\right)^{\gamma-1} e^{\left(-\left(\frac{x-\alpha}{\beta}\right)^{\gamma}\right)} \quad (1)$$

where α is the location parameter, β the scale parameter and γ the shape parameter.

Its cumulative form is

$$F(x) = 1-e^{-((x-\alpha)/\beta)^{\gamma}} \quad (2)$$

Usually, the best considered methodology for fitting a diameter distribution is the maximum likelihood method, as it is

Fig. 1 Frequencies of some characteristics of the yield plots used for comparison, regarding age, stand density (SDI) and site index (SI), whereas SDI is G_{stand} / G_{max}; G_{max} is assessed according to Reineke's site density rule (see Schütz and Zingg 2010); SI is the dominant height (hdom) referenced at age 50

used in most statistical standard procedures (as in, for instance, the statistical package SAS Inc.). This method, referred to as maximum likelihood estimation (MLE), is our reference benchmark for comparison, at least concerning the equality of fitting.

2.2.1 Estimating the parameter *α*

Considering the Weibull density functions (1) and its cumulative form (2), the two-parameter version is usually used for fitting the parameters by setting $\alpha = 0$ or $\alpha = d_{\min}$, the lesser measured diameter. The former simplification is not quite sufficient after Zanakis (1979) or Gorgoso et al. (2012), because $d_{\min}$ lies at the first decile interval of the cumulative distribution, and not 0.

Under such an assumption that $d_{\min}$ lies at half the first decile interval of the Weibull cumulative distribution function, we generated with Eq. (2) the 135 $d_{\min}$ values of in a range of Weibull parameters β and α, corresponding to our database, and a constant γ of 2.16, corresponding to the mean of our database. As a regressor for fitting $d_{\min} - \alpha$, we use a rough indicator of the first decile interval position, namely $((d_{\max} - d_{\min}) / N)d_{\min}$. It results in, regression analytically, a tight exponential relationship (exponent 0.4589) with an R^2 value of 0.9993 as

$$d_{\min} - \widehat{\alpha} = 1.007755\,[((d_{\max} - d_{\min})/N)d_{\min}]^{0.4589} \tag{3}$$

Hence, $\widehat{\alpha}$ could be determined from the stand characteristics and, for further parameter estimation, we can use the two-parameter Weibull version.

2.2.2 Estimating the parameter *β*

For the purpose of fitting, using only simple stand characteristics, as our aim, different methods can be used, derived mathematical analytically from function (2) as the method of moments or of the percentiles, with the former being mostly applied in PRM (Mehtätalo 2004).

The percentile methodology is quite interesting because some percentiles can be linked to simple stand characteristics easily identifiable in forests, like $d_{\max}$ or $d_{\min}$, lying at an assumed percentile value of $1/2N$ ($d_{\min}$) and $1 - (1/2N)$ (for $d_{\max}$), respectively. Moreover, in order to obtain reliable results, it is recommendable to use characteristics sufficiently distant from another on the *x*-axis, taking into account the spreading of the distribution. Thus, we should additionally include an indicator somehow in the middle (or the centre of gravity) of the distribution. Because *G* is easily assessed with the angle count method and both *N* and *G* for instance with the help of a smart phone application like MOTI (www.moti.ch), it results with dg as a very pertinent stand indicator while, in addition, being functionally linked to the basal area.

A key feature of the Weibull cumulative distribution function (2) is that the 63.2th percentile (at the point $1 - e^{-1}$) is equal to β or $\beta + \alpha$ in the case of the three-parameter Weibull function (Bailey and Dell 1973; Dubey 1967; Zanakis 1979). At that point, the influence of γ disappears. The 63.2th percentile is thus particularly predestined for parameter estimation but is evidently not identifiable easily in a forest. In contrast, dg lies very close (about 2–3%) to it, at least for monospecific and regular stands. In fact, the regression between the empiric values of $\beta + \alpha$ and dg is extremely straight and tight. (Residuals have been proven correctly distributed, not shown). This is somehow evident when looking at the scattering of the values in Fig. 2. Thus, the parameter $\alpha + \beta$ is accurately well described by a linear regression line with a remarkably high coefficient of determination ($R^2 = 0.998$ for the beech yield plots and 0.9995 for spruce). Hence, if dg (and α) is known, β can be easily obtained via Eqs. (3), (4a) and (4b). For the aforementioned Swiss plots, we obtained the following fitted regression lines (Eqs. (4a) and (4b)):

$$\widehat{\alpha} + \widehat{\beta} = 0.0703 + 1.0267\ \mathrm{d, for\ beech} \tag{4a}$$

$$\widehat{\alpha} + \widehat{\beta} = -0.0182 + 1.0311\ \mathrm{dg, for\ spruce} \tag{4b}$$

2.2.3 Estimating the parameter *γ*

This happens in the sense of the parameter recovery methodology. We can analytically derive the relations between the parameters and the empirical values of *N*, $d_{\min}$ and $d_{\max}$. This leads to the following non-linear system of equations relating these observed values with unknown parameters.

If $d_{\max}$ is the largest observed dbh value among the *N* ordered observations ($d_1, d_2, \ldots, d_N$), we set $F(d_{\max}) \approx 1 - \frac{1}{2N}$ which leads to $\left(\frac{d_{\max} - \alpha}{\beta}\right)^{\gamma} = -\ln(\frac{1}{2N}) = \ln(2N)$ and

$$d_{\max} = \beta(\ln(2N))^{\frac{1}{\gamma}} + \alpha \tag{5}$$

Likewise, for the smallest observation ($d_{\min}$), we have $F(d_{\min}) \approx \frac{1}{2N}$ which leads to

$$d_{\min} = \left(\beta\left(-\ln\left(1 - \frac{1}{2N}\right)\right)\right)^{1/\gamma} + \alpha \tag{6}$$

For the special case $\alpha = 0$ (with the two-parameter Weibull function), we obtain from Eqs. (5) and (6) the following solution for γ:

$$\widehat{\gamma} = \frac{\ln(\ln(2N)) - \ln\left(-\ln\left(1 - \frac{1}{2N}\right)\right)}{\ln(d_{\max}) - \ln(d_{\min})} \tag{7}$$

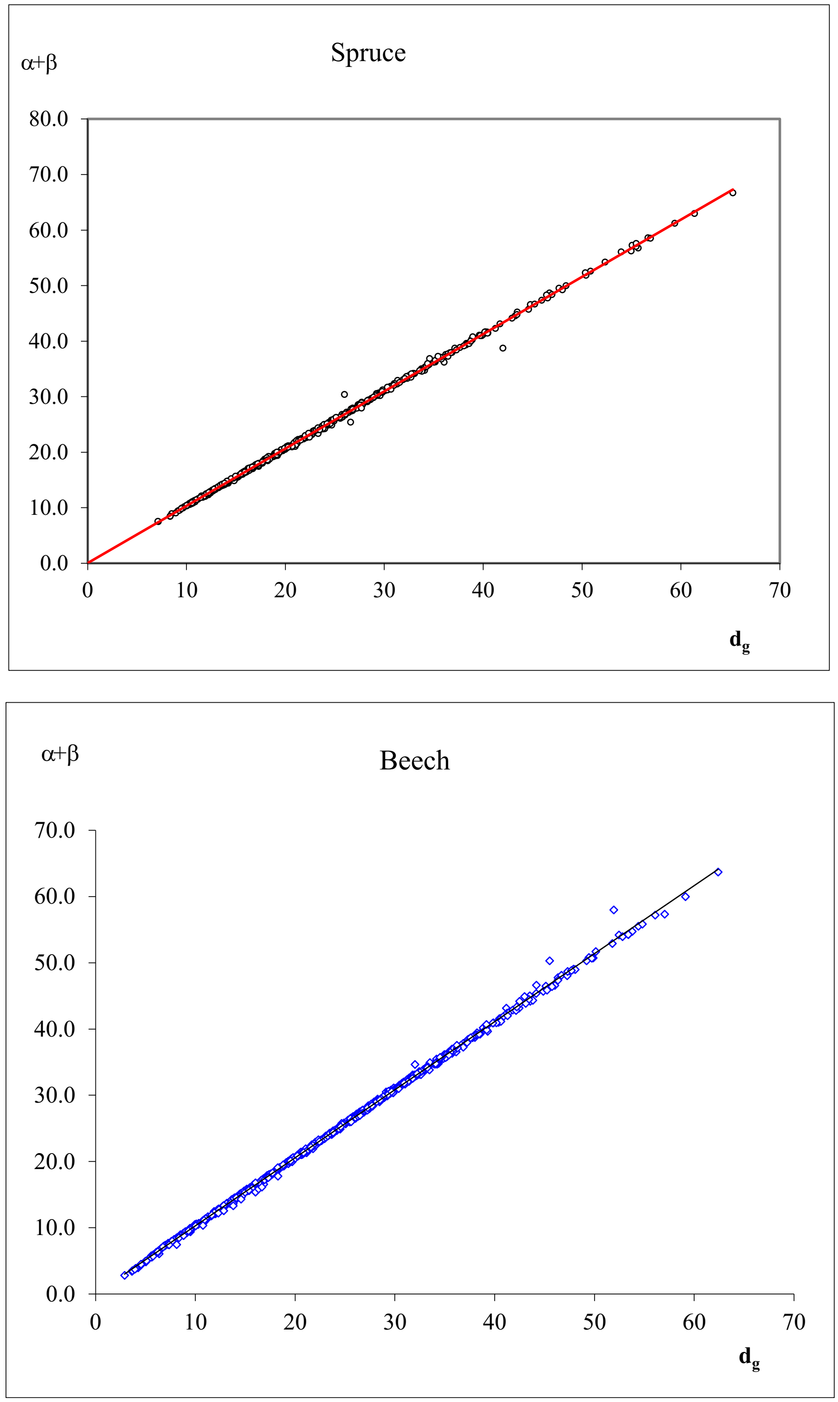


Fig. 2 Estimate of the Weibull parameter $\alpha + \beta$ obtained by a maximum likelihood estimation of the yield plot distributions, as a linear function of the quadratic main diameter (dg) for the two species spruce and beech

And for β

$$\widehat{\beta} = \frac{d_{\max} - \widehat{\alpha}}{(\ln(2N))^{1/\gamma}} \tag{8}$$

Hence, there are two ways of determining β and γ. First is in the sense of parameter prediction estimation (PRM) using Eqs. (3), (7) and (8), the method referred to as PRM1. Second is in the sense of a combination of PPM and PRM using Eqs. (3), (4a) and (4b) and then

$$\widehat{\gamma} = \frac{\ln\left(-\ln\left(\frac{1}{2N}\right)\right)}{\ln\left(\frac{d_{\max} - \widehat{\alpha}}{\widehat{\beta}}\right)} \tag{9}$$

This method is referred to as PPM1.

The difference between the PRM and PPM solutions is that only two indicators, namely d_{max} and d_{min}, are required for the PRM solution against three, adding dg, in the PPM solution. Because dg is a very relevant reference, we should expect better fitting with the PPM solution.

Because the identification of d_{max} and particularly d_{min} is somehow uncertain, it could be useful to analyse whether both indicators could be derived empirically. In fact, scatter plots d_{max} (Fig. 3) as well as d_{min} (Fig. 4) in relation to dg of the Swiss data exhibit a non-linear, relatively tightly correlated empirical relationship. We found a similar empirical non-linear relationship between d_{max} and d_{min} with the explicative variable G / dg (identical to dg N), a rough indicator for the yield level, with an R^2 value of 0.611 for beech and 0.603 for spruce plots. Both regressors can be combined in a bilinear regression of the following form. For beech, the second explanatory variable (G / dg) is not significant in the case of d_{min}, so this member is not considered. The exponents have been derived from the single relationship fitted with a power function.

$$d_{min} = 0.244648\ \mathrm{dg}^{1.2384}, \text{ for beech with an } R^2 \text{ value of } 0.90 \tag{10a}$$

$$d_{min} = 0.292814804 + 0.216214381\left(\mathrm{dg}^{1.24}\right) + 1.095854229\,(G/\mathrm{dg})^{-1.23}, \tag{10b}$$

for spruce with an R^2 value 0.899

$$d_{max} = 1.401206877 + 3.689977935\ \mathrm{dg}^{0.77} - 5.325663339\,(G/\mathrm{dg})^{-0.79}, \tag{11a}$$

for beech with an R^2 value of 0.948

$$d_{max} = 0.902011561 + 3.2457725592\left(\mathrm{dg}^{0.8}\right) - 4.936440585\,(G/\mathrm{dg})^{-0.75}, \tag{11b}$$

for spruce with an R^2 value of 0.925

The contribution of both explanatory variables is highly significant.

In the frame of developing growth models on smart phones for practitioners (see Schütz and Rosset 2016), it appears interesting to test how good the performances of volume estimation would be if renouncing the measurements of d_{max} and d_{min} based instead on Eqs. (10a), (10b), (11a) and (11b) and using solely dg. In this sense, we tested two alternative estimation methods, referred to as PRM2 and PPM2, with empiric estimates of d_{max} and d_{min} introduced in otherwise the same resolution equation systems as for PRM1 and PPM1.

Summarising, in the following, we tested 4 methods of estimating the Weibull parameters (methods 2–5 below) to determine the distribution of diameters and compare it to method 1 obtained from fitting the tallied diameters with a statistic classical method (MLE):

1. MLE, the maximum likelihood resulting from computation with a tallied diameter list and the statistic software package SAS Inc.
2. PRM1, parameter recovery based on d_{max} and d_{min}
3. PPM1, a combination of parameter recovery and empiric parameter prediction based on d_{max}, d_{min} and dg
4. PRM2, the same based on empiric estimate of d_{max} and d_{min} and assessment of dg
5. PPM2, the same based on empiric estimate of d_{max} and d_{min} and assessment of dg

2.3 Assessment of the goodness of fit

Criteria for the quality of fitting a distribution depend on the modelling aims. In our case, the equality on the whole distribution is more determinant than the adequation in every diameter class. The so-called 'Kolmogorov-Smirnov test' (Smirnov 1948; Reynolds et al. 1988) is often applied. It measures the maximal deviation between a real cumulative distribution and an estimated distribution. In our case, this test does not seem particularly adequate. So, we applied the following method for assessing the goodness of fit.

We compare the empirical quantiles of the dbhs in the stands with the corresponding quantiles obtained from the fitted Weibull distribution (so-called Q-Q plots) (Fig. 5). Ideally, one should obtain the identity regression line $y = x$ and a coefficient of determination (R^2) close to 1. We defined an R^2 value of 0.95 as an acceptable limit for such goodness of fitting.

Because one of the main purposes of such modelling is to determine the standing volume, the impact of the largest trees is more determinant than that of the small ones. Adjustment in the upper part is therefore more pertinent. Thus, as a fitting criterion on the stand level, we can consider the accuracy of estimating the standing volume (V_s stem without branches) to be the relative absolute difference between $V_{inventory}$ and $V_{estimated}$. In the same token accuracy of estimation, the ddom (diameter of 100 trees/ha in %) could be considered too.

The stem volume (V_s) is estimated with a two-variable volume function (with two independent variables d_i and h_i) of the form

$$V_s = \alpha\, e^{\left(\beta \ln(d_i) + \gamma\left(\ln(d_i)^2\right) + \delta \ln(h_i) + \varepsilon\left(\ln(h_i)^2\right)\right)} \tag{12}$$

Equation (12) was established especially for spruce and beech on the basis of numerous lying tree measurements in 2-m sections (15,285 for spruce and 8295 for beech) from the WSL database (unpublished). Because our database of yield plots does not contain all h_i values, only a sample, we need to find an estimation for the missing height. Therefore, single empiric height values of yield plots with h_i measurement (226 different h_i measures

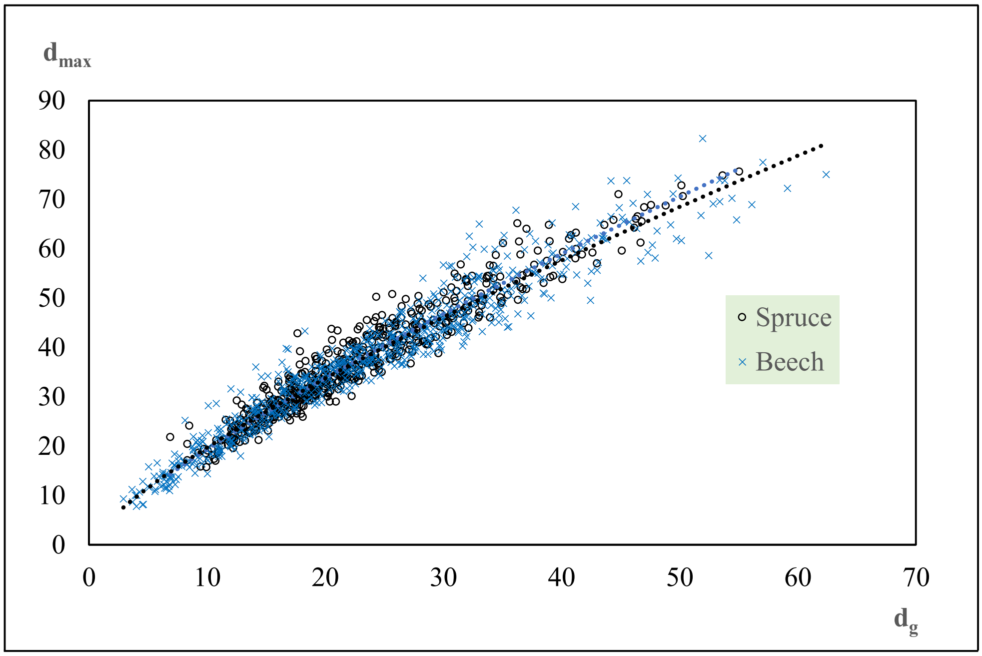


Fig. 3 Scatter plot of d_{max} versus dg based on 813 permanent yield plots of spruce and 667 of beech in Switzerland

from 62 yield plots) are fitted apart with the semi-logarithmic model function (data not shown)

$$h_i = \alpha h_i + \beta h_i \ln(d_i) \tag{13}$$

Thereafter, only the regressor βh_i is required because it gives the curvature of the actual height functions. For modelling purposes, the height curves should be adjusted to the development stage, which depends on the SDI (ddom at age 50). So for each development stage, we determine αh_i to pass through the corresponding point (hdom; ddom) of the actual stand height development.

Summarising, when regressed on the variables dg and ln(dg) or dg^2, we obtain the βh_i of the Swiss plots

$$\beta h_i = 5.40133 + 0.24498 \text{ dg} - 0.00394(\text{dg})^2 \text{ for spruce} \tag{14a}$$

$$\beta h_i = 1.87183 + 1.72503 \ln(\text{dg}) \text{ for beech} \tag{14b}$$

The missing coefficient αh_i is obtained when fitting the height curves to pass through the point hdom:ddom,

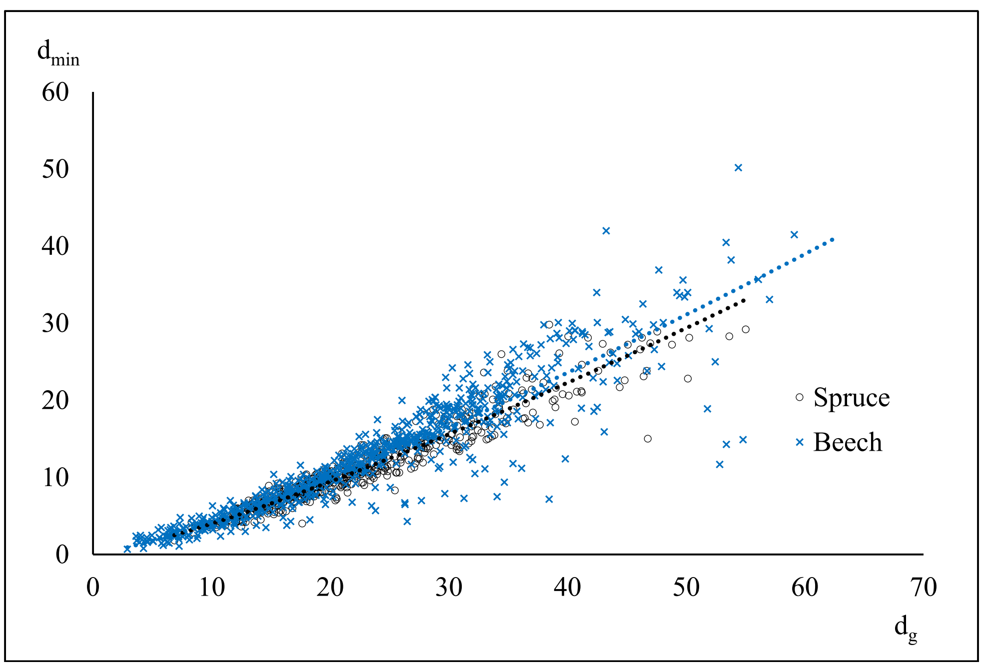


Fig. 4 Scatter plot of d_{min} versus dg based 813 permanent yield plots of spruce and 667 of beech in Switzerland (so-called 'Q-Q plots')

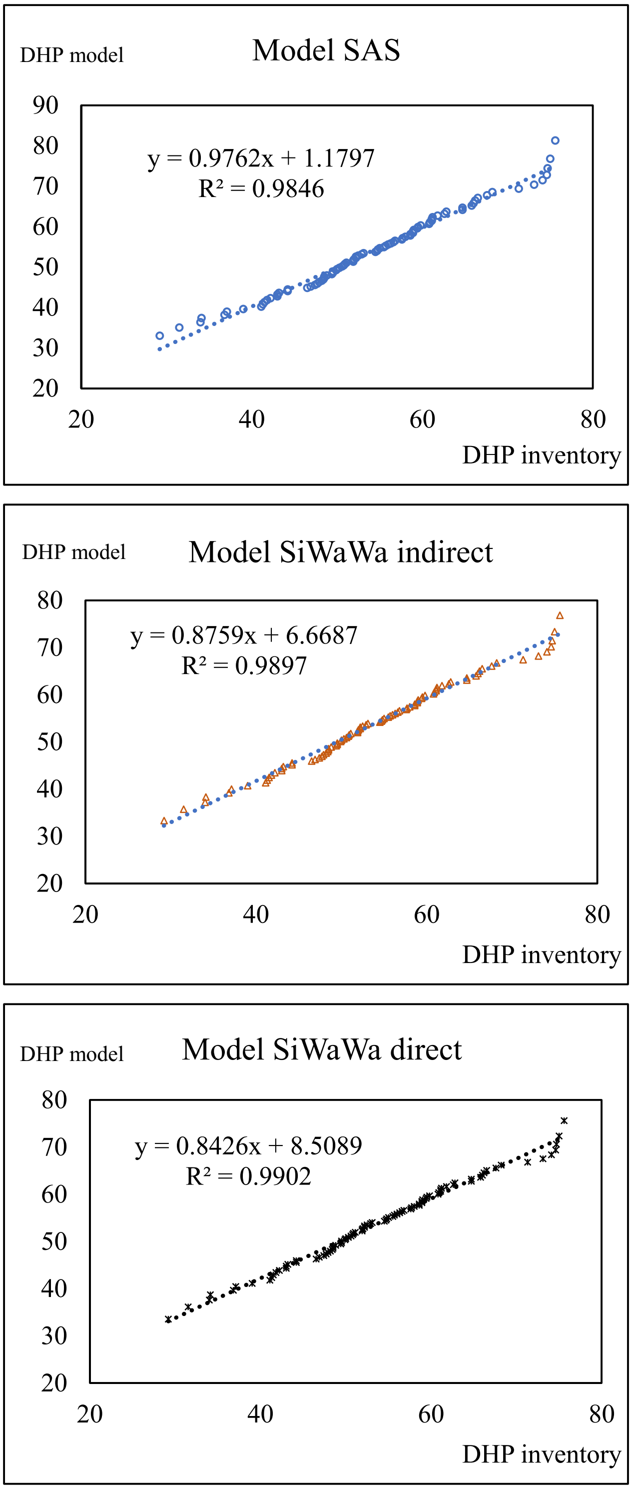


Fig. 5 Linear regression of empiric diameter quantiles against the fitted quantiles of distribution estimated with the Weibull parameters obtained by the three different estimation procedures for spruce, considered as a method of goodness of fitting

determined with a SDI function (not shown, see Schütz and Rosset 2016).

Thus

$$\alpha h_i = \text{hdom} - \beta h_i\,(\text{dg})\,\ln(\text{ddom}) \tag{15}$$

3 Results

Table 1 summarises the results for goodness of fit. As significant for the main test (Q-Q plots) is considered a R^2 value higher than 0.95.

Figure 6 illustrates two examples of diameter distribution—one with well-balanced (unimodal) distribution over the diameter classes and one with unequal distribution. Because of compensation effects, the former leads to acceptable fitting over the whole distribution.

The method of the maximum likelihood as benchmark appears compliant, based on the R^2 value of the Q-Q plots in 90.1% (beech) and 99% (spruce), respectively. This method ranks first in 81.6% (beech) and 93.5% (spruce), respectively. In comparison, the method based on the empiric relationship between parameters β and dg (PPM1 and PPM2) presents compliant results in 65.1% (68.7) for PPM1 and 72.3% (55.9) for PPM2, respectively. The result of such fitting without tallying all diameters surpasses even the results of the classical fitting (with the maximum likelihood estimate) based on a full inventory of the diameters in 24.6 (18.65)% of the case for beech and in 3.2 (1.4)% for spruce.

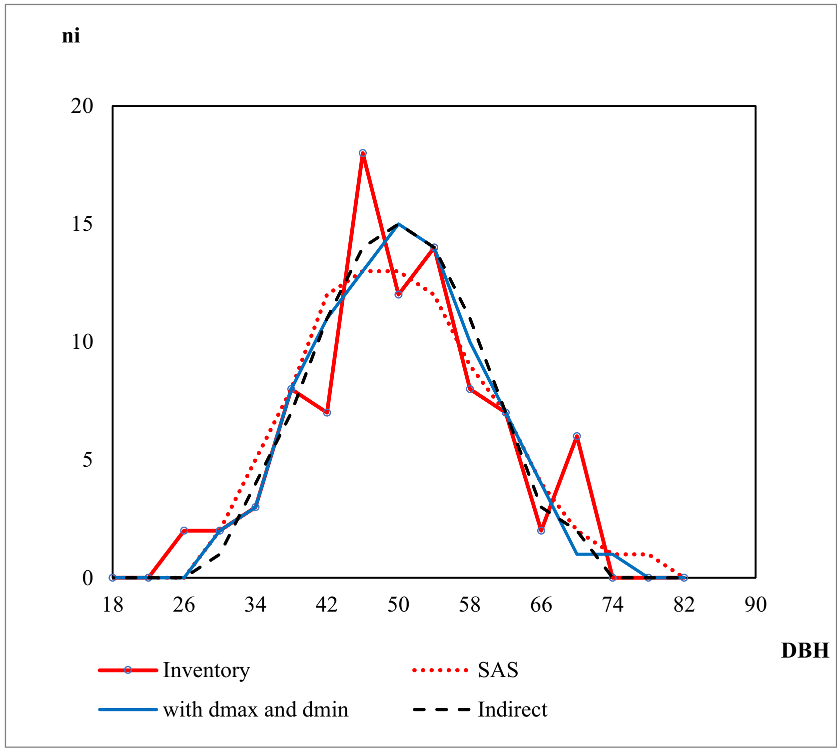
As expected, the PRM based only on $d_{\min}$ and $d_{\max}$ performs less well with 21.3% (32.5) and 53.7% (29.9) for PRM2.

Table 1 Goodness of fit for comparing the five methods of adjustment (see Section 2.2)

Comp. inventory with	MLE	PRM1	PPM1	PRM2	PPM2
Q-Q plots[1]					
Beech (%)	90.1	21.3	65.1	53.7	72.3
Spruce (%)	99.0	32.5	68.7	29.7	55.7
Ranking % 1st[2]					
Beech	81.6	2.9	24.6	4.2	18.6
Spruce	93.5	0.9	3.2	1.0	1.4
Accuracy for V_s					
Beech (%)	1.2	7.1	0.2	− 2.9	− 0.9
Spruce	0.4	− 1.3	− 1.2	− 8.2	− 1.9
Accuracy for ddom					
Beech	0.5	− 1.4	− 3.2	− 3.2	− 2.3
Spruce	0.6	− 5.0	− 4.8	− 7.2	− 5.7

[1] % number of plots with $R^2 > 0.95$

[2] % plots ranking first over the 5 methods on the base of the R^2 value of Q-Q plots



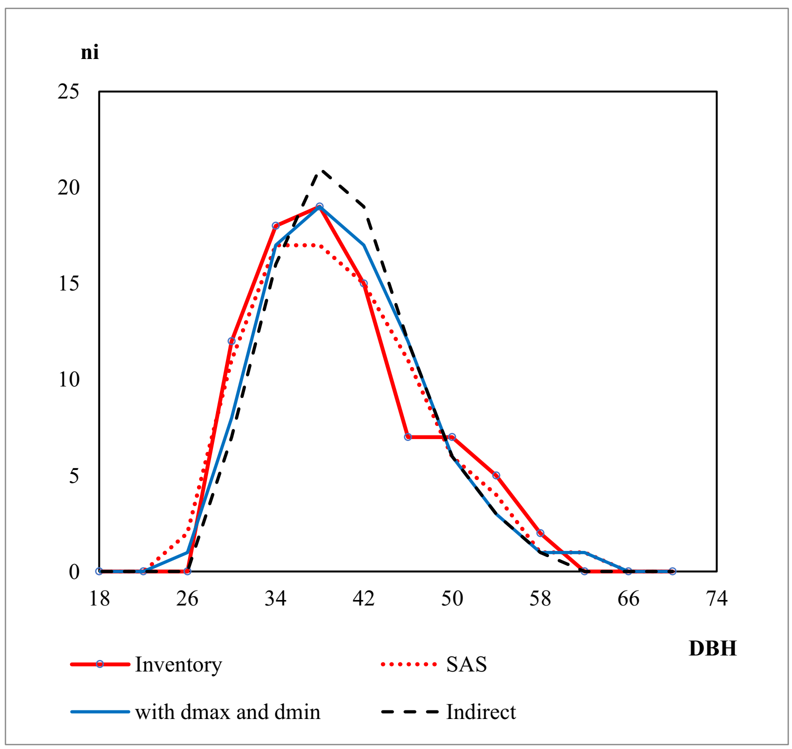


Fig. 6 Stem number distribution of spruce plot 21-066, inventory year 1937, at age 139 (above). Below, beech plot 42-024, inventory year 1978, at age 125. Compared to plot inventory (bold red line), the distribution with parameter estimation using a statistical program (SAS Institute Inc.) is presented in red dotted, the simple SiWaWa indirect method in bold blue line (this study) and the same with the real d_{max} and d_{min} in black dotted line

The methods without assessing d_{max} and d_{min}, but based on the indirect empiric way (PRM2, PPM2), perform better, at least in the case of beech (53.7% over 21.3 for PRM and 72.3% over 21.2 for PPM).

Estimation accuracy of the standing volume is 1.2% (resp. 0.4% for spruce) for the MLE reference method. More astonishing is the remarkable performance of the PPM1 method with practically similar (even in average better) results (0.2–1.2%). Here too, the method with using only dg (PPM2) performed quite respectably (2.9, 1.4%). In contrast, the PRM did not perform very well.

4 Discussion

It may be astonishing that a simple solution, with only dg as information, allows for the prediction of diameter distribution with the same significance (or even better) than the usual fitting of a full diameter inventory and, consequently, allows us to assess the standing volume with an accuracy of 0.9–1.9%. If generalisable, this would open bright prospects for growth modelling and for forest inventory.

These empiric results emanate from yield plots selected for even aged and monospecific composition for two main tree species in Switzerland, albeit with enough representation in numbers and sites. This already demonstrates the adequacy for growth modelling, at least for the above-mentioned site conditions and, let us say, for temperate Central Europe, in so far as a growth model usually simplifies the growth prediction to monospecific species. Similar results had been found for four other main tree species—namely Scot pine (*Pinus sylvestris* L.), Douglas fir (*Pseudotsuga menziesii* Mirb.), European white oaks (*Quercus robur* L. and *Quercus petraea* Liebl.) and European common ash (*Fraxinus excelsior* L.) (Schütz and Rosset 2016), with similar straightness and high R^2 for the $\alpha + \beta$ dependency to dg, confirming the method validity that dg lies in the vicinity of the 63.2th percentile and, therefore, where the influence of γ disappears. There are good reasons to think that this remarkable feature could be generalised, with the eventual exception of too irregular stands. Such a possible generalisation should be addressed in further research.

From the three applied independent variables for predicting the diameter distribution, dg is the most pertinent because it lies at the centre of gravity of the distribution and is directly related to the basal area and, thus, with the standing volume. This fact explains why the parameter recovery method using only $d_{\min}$ and $d_{\max}$ does not match so well. As demonstrated, dg is very tightly correlated with the main parameter β responsible for the spreading of the distribution. This is in relation to the properties of the Weibull function itself and not somehow a serendipitous empiric result.

Once β is assessed, γ can be derived so the function passes through to the point $d_{\max}$ (see Eq. (9)). In this regard, the question arises whether only one large tree is sufficiently representative and consistent to characterise the distribution. Figure 3 and Eqs. (10a) and (10b) show that $d_{\max}$, in relation to dg, is quite consistent, with additional G / dg (equal to dg N) as an independent regressing variable. The statistical multiple correlation lies with an R^2 value of 0.948 and 0.925, respectively. In previous studies, we showed in a thinning experiment that the position of the largest trees is quite stable on the long run and relatively uninfluenced by the thinning intensity (Schütz et al. 2015). This could be explained by the fact that in every population, the largest trees have sufficient growth momentum to maintain their socially dominant position without suffering from competition from the neighbours. One condition is, of course, that the largest tree should stem from the same cohort, excluding, for instance, standard remnants. So, we can conclude that $d_{\max}$ appears to be sufficiently robust and appropriate to characterise the right part (i.e. the most important part for standing volume assessment) of the distribution. Kangas and Maltamo (2000), using $d_{\max}$ $d_{\min}$ and a basal area–weighted mean diameter, found a similar standing volume accuracy between 1 and 2%, which is more than the value 3.6–5.7% found by Siipilehto and Mehtätalo (2013).

The importance of $d_{\min}$ is more subject to concerns. In more irregular stands, in mixed stands and in unevenly aged ones, $d_{\min}$ is probably not so convenient and not so clearly identifiable in opposition to $d_{\max}$, which represents the biological best-fitted self-dominating sustainable element of the cohort. In such cases, the empiric relationship between $d_{\min}$ and dg (and eventually G / dg) can be useful. This explains why the PPM2 method performs better than PPM1. In older stands, a secondary crop installs itself spontaneously, so $d_{\min}$ should belong to the cohort being considered, excluding, for instance, second growth or under-storey. Then again, the smallest trees are more unstable and, so, could be damaged after logging operations and could otherwise die from overcrowding. Nevertheless, Fig. 4 shows that its distribution is far from unacceptable. Because $d_{\min}$ influences the origin of the distribution, it contributes less from the points of view of productivity or value. The use of the two smallest trees, as Zanakis (1979) proposes, or even several more, to somehow smooth this point of reference, does not change the problem.

Strictly considered, $d_{\max}$ (as well as $d_{\min}$) depends on the reference area on which they have been recorded. It is quite evident that, for small areas, the largest recorded tree does not have the same significance as for very large areas. Thus, it seems better to assess $d_{\max}$ over a sufficiently large area.

5 Conclusions

Using the three characteristic points of a forest stand, dg, $d_{\min}$ and $d_{\max}$, appears informative enough to determine the parameters of the whole diameter distribution and, hence, the standing volume with quite a valuable accuracy. This is connected principally with a particular feature of the Weibull distribution function, which allows the prediction of the main parameter β from dg with an unexpectedly high likelihood. This feature could be used for growth modelling as well as inventory purposes, at least for monospecific and even-aged stands but, maybe more, because this feature is proper to the function itself. Once the scale parameter β is estimated, the two other parameters can be recovered so that the distribution passes through the values $d_{\max}$.

This leads to a revival of an old but very ingenious method of the angle count sampling of Walter Bitterlich (1947).

Furthermore, it is quite inexpensive and easy in comparison to plot sampling or full inventory but, nevertheless, performs similarly or even better (Piqué et al. 2011). The remarkable evolution of information technologies (ITs) allows for the application of such growth models for smart phones or tablets as useful and accurate devices based on these findings (Schütz and Rosset 2016) allowing interactive participation too.

Acknowledgements We are deeply grateful to Dr. Daniel Mandallaz, a lecturer in statistics and forest inventory at ETH-Zurich, for his highly appreciated advices in mathematic statistics, particularly for deriving the recover solution. We are very grateful too to the Swiss Federal Institute for Forest Snow and Landscape Research (WSL), Birmensdorf, Switzerland, and particularly to Andreas Zingg for making the data from the yield database available. Without the high proficiency of technicians responsible for yield plot surveys, the accuracy of the data would never have been of the required high quality needed for such modelling. They worked unselfishly, partly during their whole tenure but always in a responsible manner and mostly under difficult conditions. Many thanks therefore to Jakob Werner, Emil Müller, Fritz Pfäffli, Paul Imhof, Enrico Cereghetti, Christian Matter, Ernst Baumann and many others unknown.

Contribution of co-authors Long-time collaboration on the development of application of the Weibull function for modelling purposes; collaboration to text structuration

References

Bailey RL, Dell TR (1973) Quantifying diameter distributions with the Weibull function. For Sci 19:97–104

Bitterlich W (1947) Die Winkelzählmessung. Allg Forst und Holzwirtschaftliche Z 58:94–96

Bitterlich W (1984) The relascope idea; relative measurements in forestry. Commonwealth Agricultural Bureaux, Slough

Borders BE, Souter RA, Bailey RL, Ware KD (1987) Percentile-based distributions characterize forest stand tables. For Sci 33:570–576

Clutter JL, Bennett FA (1965) Diameter distribution in old-field slash pine plantation. Ga For Res Counc Rep:13

Dubey SD (1967) Some percentile estimators for Weibull parameters. Technometrics 9:119–129

Gorgoso JJ, Rojo A, Camara-Obregon A, Dieguez-Aranda U (2012) A comparison of estimation methods for fitting Weibull, Johnson's and beta functions to Pinus pinaster, Pinus radiata, and Pinus sylvestris stands in northwest Spain. For Syst 21:446–459

Hyink DM, Moser JW (1983) A generalized framework for projecting forest yield and stand structure using diameter distribution with additional information. For Sci 29:85–95

Kangas A, Maltamo M (2000) Performance of percentile based diameter distribution prediction and Weibull method in independent data sets. Silva Fenn 34:381–398

Mehtätalo L (2004) Predicting stand characteristics using limited measurements. FFRI research report 929, Finnisch Forest Research Institute (dissertation)

Piqué M, Obon B, Condés S (2011) Comparison of relascope and fixed-radius plots for the estimation of forest stand variables in northeast Spain: an inventory simulation approach. Eur J For Res 130:851–859

Reynolds MRJ, Burk TE, Huang WC (1988) Goodness-of-fit tests and model selection procedures for diameter distribution models. For Sci 34:373–399

Schädelin W (1934) Die Durchforstung als Auslese- und Veredlungsbetrieb höchster Wertleistung. P Haupt, Bern-Leipzig

Schmid P, Roiko-Jokela P, Mingard P, Zobeiry M (1971) The optimal determination of the volume of standing trees. Mitt Forst Bundesversuchsanstalt Wien 91:33–54

Schütz JP, Rosset C (2016) Des modèles de production et d'aide à la décision sur smartphones. Rev For Fr 68:427–439

Schütz JP, Zingg A (2010) Improving estimations of maximum stand density by combining Reineke's size-density rule and the yield level, using the example of spruce (Picea abies (L.) Karst.) and European beech (Fagus sylvatica L.). Ann For Sci 67:507–518

Schütz JP, Ammann PL, Zingg A (2015) Optimizing the yield of Douglas-fir with an appropriate thinning regime. Eur J For Res 134:469–480

Siipilehto J, Mehtätalo L (2013) Parameter recovery vs. parameter prediction for the Weibull distribution validated for scots pine stands in Finland. Silva Fennica 47:1–22

Smirnov N (1948) Table for estimating the goodness of fit of empirical distributions. Ann Math Stat 19:279–281

Weibull W (1939) A statistical theory of the strength of material. Ing Vetensk Akad Handl Nr. 153. Generalstabens Litografiska Anstalts Förlag, Strockholm

Zanakis SH (1979) A simulation study of some simple estimators for the three-parameter Weibull distribution. J Stat Comput Simul 9:101–116