J Mater Sci (2011) 46:1862–1869
DOI 10.1007/s10853-010-5014-2

Reliability of a Weibull analysis using the maximum-likelihood method

Milan Ambrožič · Lovro Gorjan

Received: 24 May 2010 / Accepted: 18 October 2010 / Published online: 30 October 2010

Abstract We have performed extensive Monte-Carlo computer simulations of the 2-parameter Weibull statistical distribution using data groups with sizes from 5 up to 100 samples. The maximum-likelihood method was used to evaluate the resulting Weibull modulus and the scale parameter, which may be different to the input values. We confirmed some trends in the evaluation of the statistical parameters for small data groups, such as a significant biasing of the Weibull modulus. We revealed the log-normal statistical distribution of the Weibull parameters obtained from repeated Monte-Carlo simulations for several groups. We also considered the influence of the measurement uncertainty on the determination of the statistical parameters. For the experimental evidence we used bend-strength data for alumina test samples from serial production in this year. The experimental data were randomly divided into several subgroups to compare the corresponding biasing of the Weibull modulus with the Monte-Carlo results.

M. Ambrožič
Jožef Stefan International Postgraduate School, Jamova 39, 1000 Ljubljana, Slovenia

M. Ambrožič
Physics Department, Faculty for Natural Sciences and Mathematics, University of Maribor, 2000 Maribor, Slovenia

L. Gorjan (✉)
Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia
e-mail: lovro.gorjan@hidria.com

L. Gorjan
Hidria AET, Tolmin, Slovenia

Introduction

The strengths measured in typical mechanical tests for brittle materials, such as ceramics, result in a Weibull statistical distribution [1]. Usually, the 2-parameter Weibull distribution is used, although in many cases the application of 3-parameter Weibull statistics or Weibull statistics corresponding to two or more different fracture modes is more appropriate [2–6]. The inherent scattering of the experimental strength values for different test samples is characterized by the Weibull modulus, and while the estimation of the scale parameter is usually accurate, the uncertainty of the Weibull modulus may be large.

Different calculation methods may be used to evaluate the Weibull parameters, e.g., the linear regression (LR) method and its variants, the maximum-likelihood (ML) method, a direct calculation from the mean value and the standard deviation of the strength of several samples (the moments method), etc. Each of the methods has its benefits and drawbacks [7, 8]. The validity of the Weibull parameters' estimations has been extensively investigated experimentally and theoretically. For instance, some authors tested the idea of dividing several measured strength values of ceramic materials into random smaller subsets to study the corresponding statistical distribution of the Weibull parameters [4, 5, 8–10]. They found a good agreement in the variation of the scale parameter with previous theoretical predictions, but the agreement for the Weibull modulus was worse and this was attributed to the deviation of the strength statistics from ideal Weibull statistics.

Monte-Carlo simulations are very useful tool for predicting the reliability of various estimation methods and their optimization [7, 8, 11–19]. The distribution functions of the estimated Weibull parameters from several

Monte-Carlo repetitions were proved to be log-normal [11–14, 20]. There is also a significant biasing, i.e., an under- or over-estimation of the Weibull parameters on average, depending on the size of the test group and the method used. Using different weight factors for different data points in the weighted LR method [7, 11, 21, 22], or improving the probability estimators (mainly adapting them to different data group sizes) [7, 17, 23] can reduce the biasing and/or the width of the probability distribution, in particular the distribution of the Weibull modulus. Among the different methods, the ML method is a standard method due to its efficiency and its ease of application when censored failure populations are encountered [24]. Since this method has proved to be particularly suitable, several variants of it have been proposed and tested, for instance the generalized maximum-likelihood method (GMLE), which uses various rank estimators [25–27].

In our previous article we treated the reliability of the Weibull parameters' estimation by using the LR method. In particular, we found that this method underestimates the Weibull modulus (in dependence of the chosen rank estimator) [28]. The results of the Monte-Carlo simulations of the data were supported qualitatively with experimental data, i.e., by almost 500 measurements of the strength of fiber-cement roofing sheets from serial production (although there was some quantitative disagreement between the theory and the experiments).

In this study we analyze a large amount of Monte-Carlo data and estimate the Weibull parameters using the ML method. In addition, we use the results of measurements of the four-point bend strength of 96% alumina samples from the serial production. We divide the strength values into random smaller subgroups. The corresponding distribution of the Weibull parameters is compared to Monte-Carlo simulations. One of the aims of this research was to confirm the expectation that the estimated Weibull parameters for several sample groups of finite size according to the ML method obey the log-normal distribution. Finally, the significance of the measurement uncertainty to the statistical results is evaluated.

Statistical model

Let our statistical variable be the four-point bend strength (called strength for brevity), σ, with the probability density function $p(\sigma)$. The cumulative probability function, also called the unreliability function, is defined as: $P(\sigma) = \int_0^\sigma p(x)\mathrm{d}x$. In the case of the 2-parameter Weibull statistics, the functions $p(\sigma)$ and $P(\sigma)$ are equal to:

$$p(\sigma) = \frac{m}{\sigma_0}\left(\frac{\sigma}{\sigma_0}\right)^{m-1}\exp\left(-\left(\frac{\sigma}{\sigma_0}\right)^m\right) \tag{1a}$$

$$P(\sigma) = 1 - \exp\left(-\left(\frac{\sigma}{\sigma_0}\right)^m\right), \tag{1b}$$

with the Weibull modulus m and the scale parameter σ_0. While various linear regression (LR) methods are operating with the function (Eq. 1b), the ML method uses directly (Eq. 1a) [7]. Here, we focus on the ML method.

After measuring (or simulating) N strength values, σ_i, $i = 1$ to N, the data are inserted into the probability density function (Eq. 1a). The ML procedure maximizes the following function with respect to the free parameters m and σ_0:

$$Y = \ln\left(\prod_{i=1}^{N} p(m, \sigma_0; \sigma_i)\right) = \sum_{i=1}^{N} \ln p(m, \sigma_0; \sigma_i). \tag{2}$$

When the derivatives of Y with respect to m and σ_0 are set to zero, the following pair of equations is obtained:

$$\frac{1}{m} = \left(\sum_{i=1}^{N} \ln\sigma_i \cdot \sigma_i^m\right) \Big/ \left(\sum_{i=1}^{N} \sigma_i^m\right) - \frac{1}{N}\sum_{i=1}^{N} \ln\sigma_i, \tag{3a}$$

$$\sigma_0 = \left(\frac{\sum_{i=1}^{N} \sigma_i^m}{N}\right)^{1/m}. \tag{3b}$$

In practice, Eq. 3a is first numerically solved for m and then σ_0 is calculated from Eq. 3b.

Knowing the Weibull parameters one can calculate various statistical parameters, for instance, the theoretical mean value $\langle\sigma\rangle$ of the strength and its standard deviation $\delta\sigma$, corresponding to the limit $N \to \infty$, and compare them to the actual values, $\langle\sigma\rangle_N$ and $\delta\sigma_N$, for a finite sample group of size N. The relevant formulae are:

$$\langle\sigma\rangle = \sigma_0 \cdot \Gamma\left(1 + \frac{1}{m}\right) \tag{4a}$$

$$\delta\sigma = \sigma_0 \cdot \sqrt{\Gamma\left(1 + \frac{2}{m}\right) - \Gamma^2\left(1 + \frac{1}{m}\right)}, \tag{4b}$$

where $\Gamma = \int_0^\infty t^{x-1}e^{-t}\mathrm{d}t$ is the gamma function, and:

$$\langle\sigma\rangle_N = \frac{1}{N}\sum_{i=1}^{N} \sigma_i \tag{4c}$$

$$\delta\sigma_N = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (\sigma_i - \langle\sigma\rangle_N)^2}. \tag{4d}$$

It is supposed that the distribution of the estimated values of the parameters is log-normal. This means that if there

are several groups of size N of independent experimental (or Monte-Carlo generated) data, and for each group different values of m and σ_0 are obtained, the resulting distribution of their natural logarithms is Gaussian:

$$p(\ln x) = \frac{1}{w\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln x - a}{w}\right)^2\right), \tag{5a}$$

where x stands for either m or for σ_0. Here, w is the half-width (standard deviation) of the Gaussian curve centered at a. The log-normal distribution (Eq. 5a) can be written directly as:

$$p(x) = p(\ln x) \cdot \frac{d(\ln x)}{dx} = \frac{1}{w\sqrt{2\pi} \cdot x} \exp\left(-\frac{1}{2}\left(\frac{\ln x - a}{w}\right)^2\right), \tag{5b}$$

which may be significantly asymmetric, depending on w. Three characteristic statistical parameters may be defined with the distribution (Eqs. 5a, 5b): the mean value (x_{mean}), the median (x_{med}) and the mode (x_{mode}). The simplest to obtain is the median: $P(x_{\text{med}}) = ½$, according to the cumulative probability function of (Eqs. 5a, 5b), that is

$$x_{\text{med}} = \exp(a). \tag{6a}$$

This value corresponds to the peak of ln x (Eq. 5a). The mode x_{mode} is defined as the value where the direct distribution function (Eq. 5b) has its maximum, while x_{mean} is the usual mean value of x according to (Eq. 5b). There are simple relations between these three parameters for the distribution (Eqs. 5a, 5b):

$$x_{\text{mode}} = x_{\text{med}} \cdot \exp(-w^2), \tag{6b}$$

$$x_{\text{mean}} = x_{\text{med}} \cdot \exp(+w^2/2), \tag{6c}$$

thus the inequality $x_{\text{mode}} < x_{\text{med}} < x_{\text{mean}}$ holds. The larger the half-width w, the larger is the difference between the three parameters. Here, we are mostly interested in the mean value and the median according the log-normal distribution of the Weibull modulus, since these parameters can be calculated simply by statistically averaging the Weibull modulus or its logarithm, respectively (see Eqs. 8a, 8b, and 8c in the discussion below).

Strength measurements

Rectangular alumina samples (sintered at 1640 °C for 3 h) with dimensions 4 × 3 × 50 mm were subjected to the four-point bending test with an inner span of 20 mm and an outer span of 40 mm. The four-point test was conducted according to the ASTM standard test [29]. The bend strength of each sample was calculated from the testing geometry and the breaking force. In this year, a little more than a thousand samples were broken for all the batches, where each batch typically contained 12 measurements. For the presentation in this article we chose chronologically the first 1000 strength values, which were subsequently statistically analyzed.

Statistical simulations

For the comparison we use two different approaches to model or evaluate the effects, such as biasing the Weibull modulus on average, when using the ML estimation, either an experimental or simulated data, as described below. In both the cases we first choose the size of the test group N (i.e., N samples in the group). To obtain an appropriate statistical evaluation we take several, i.e., N_{rep}, different groups of N samples; typically $N_{\text{rep}} = 10^4$–10^6. For each group we obtain a pair of Weibull parameters using the ML method. Since the groups differ from each other, we obtain N_{rep} different pairs of m and σ_0. For additional control, the third procedure is also used (see subsection "Hybrid method: mixing a finite number of MC data", below), which is a hybrid of both approaches.

Generating data by random generator (Monte-Carlo simulation)

We generate the Monte-Carlo (MC) data for the strengths using the prescribed input Weibull parameters m_{inp} and $\sigma_{0,\text{inp}}$. A "random" strength is calculated from the given random number $0 < r < 1$, using Eq. 1b, by:

$$\sigma = P^{-1}(r) = \sigma_{0,\text{inp}} \cdot \left[\ln\left(\frac{1}{1-r}\right)\right]^{1/m_{\text{inp}}}. \tag{7a}$$

We can also simulate the measurement uncertainty by taking the Gaussian distribution of the measurement errors: $p(\Delta\sigma) = \frac{1}{w_{\text{MU}}\sqrt{2\pi}} \exp(-\frac{1}{2}(\frac{\Delta\sigma}{w_{\text{MU}}})^2)$, where $\Delta\sigma$ means the error of the individual measurement and w_{MU} is the measurement uncertainty. We include the measurement errors in the MC simulation by:

$$\sigma = \sigma_{0,\text{inp}} \cdot \sqrt[m_{\text{inp}}]{\ln\left(\frac{1}{1-r_1}\right)} \pm w_{\text{MU}} \cdot \text{erf}^{-1}(|2r_2 - 1|), \tag{7b}$$

where r_1 and r_2 are the two independent random numbers. In the second term the negative or positive sign is chosen, whether $r_2 < 0.5$ or >0.5, respectively. For convenience, we calculated numerically in advance the error function $\text{erf}(x) = \sqrt{\frac{2}{\pi}} \int_0^x \exp(-t^2/2)\text{d}t$ for a dense grid of the independent variable x.

Mixing experimental data using a computer random generator

The procedure of dividing the experimental data into smaller random subsets (subgroups) was used before, for instance in Ref. [4]. For a chosen group size N we use a random-number generator to "shuffle" the N_{exp} experimental strength values and pick N values from them.

Since the number N_{exp} of experimental data is limited, the method of mixing experimental data itself seems at first sight to be very limited, in comparison to generating all the data with the random generator described above. But this is not actually the case. Having, for instance, only $N_{exp} = 100$ and $N = 50$, we can make 100!/(50!·50!) combinations for different groups (each with a different pair of Weibull parameters), practically an infinite number. But there is in any case one important limitation: extreme (very large and very small) strength values are usually missing in the experimental data, while in the theory they are possible. As a consequence, the range of strength values in any subset is limited by the smallest and the largest value of the entire experimental set, irrespectively of the number of subsets. Another aspect of this problem is the fact, that even for very large number of measurements extreme strength values, hypothesized in the theory, are rather unrealistic—such values have never been observed.

Hybrid method: mixing a finite number of MC data

In this control method we first generate a finite number (N_{exp}) of MC data in the same way as described in subsection "Generating data by random generator (Monte-Carlo simulation)", instead of taking real N_{exp} experimental values. An additional procedure of dividing the MC data into smaller random subsets (subgroups) is subsequently done in the same way as for the experimental data.

Results and discussion

Monte-Carlo data generation without measurement uncertainty

First, we took various pairs of the input parameters m_{inp} and $\sigma_{0,inp}$ (the latter in arbitrary units) as well as different size groups N to obtain some general findings. In this step the measurement uncertainty w_{MU} is set to zero. The repetition number N_{rep} between 10^4 and 10^5 already gives good representative statistics, but to be more certain about the results we sometimes used a few millions of repetitions. Some general expectations were confirmed:

1. The distribution of the values of the Weibull modulus m is independent of the input scale parameter $\sigma_{0,inp}$, but, on the other hand, the input Weibull modulus m_{inp} influenced the statistics of the scale parameter. Both distributions, for m and σ_0, fit the log-normal distribution (Eqs. 5a, 5b) very well.
2. The ML method, on average, overestimates the Weibull modulus. Both the shift of the peak with respect to the input Weibull modulus (i.e., the difference $a - \ln m_{inp}$) and the half-width w in the distribution (Eqs. 5a, 5b) depend only on the group size N but not on m_{inp} and other parameters; the shift and the half-width decrease rapidly with increasing N.
3. On the other hand, the scale parameter is slightly underestimated, but this difference can be neglected, except for the smallest input Weibull modulus. The corresponding half-width of the Gaussian distribution of $\ln \sigma_0$ depends not only on N, but also on m_{inp}. The relative width in the distribution of the scale parameter can be neglected in comparison with that of the Weibull modulus.

Since the distribution function of the scale parameter is relatively narrow and the underestimation of its most probable value is small, we will focus mainly on the discussion of the distribution of the Weibull modulus. We can conveniently estimate the "best value" of m after several repetitions N_{rep} in two characteristic ways. The first method corresponds to the most probable value of $\ln m$ corresponding to the peak of the distribution function (Eq. 5a), i.e., the median (Eq. 6a):

$$\langle m_{LN} \rangle \equiv \exp(\langle \ln m \rangle). \tag{8a}$$

In (Eq. 7a) the brackets $\langle \ldots \rangle$ denote statistical averaging. The index "*LN*" means averaging of the logarithms of m, in contrast to the usual (direct) averaging of m according to the mean value:

$$\langle m_{DIR} \rangle \equiv \langle m \rangle. \tag{8b}$$

In accordance with this notation, we rewrite Eq. 6c:

$$\langle m_{DIR} \rangle = \exp(w^2/2) \cdot \langle m_{LN} \rangle \tag{8c}$$

and this was confirmed by our simulations.

Because of equal shifts of the logarithms for all the input values of m_{inp}, we can state that the logarithmically averaged value of m in Eq. 8a differs from input m by the common "bias" factor $k_1(N)$:

$$\langle m_{LN} \rangle = k_1(N) \cdot m_{inp}. \tag{9a}$$

We can define a similar bias factor for the direct averaging of m:

$$\langle m_{DIR} \rangle = k_2(N) \cdot m_{inp}. \tag{9b}$$

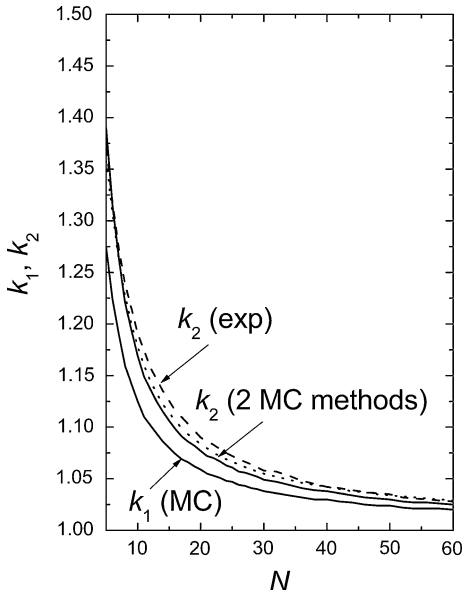

Fig. 1 *Solid lines*: the dependence of MC bias factors k_1 and k_2 for m on N (solid lines). Other testing parameters: $N_{rep} = 10^5$, $m_{inp} = 10$, $\sigma_{0,inp}$ is arbitrary. The coefficients k_2 from mixing 1000 experimental strengths (*dashed line*) and from hybrid method of mixing 1000 MC strengths (*dotted line*) are added for comparison

The factor $k_2(N)$ is larger than $k_1(N)$ by the factor $\exp(w(N)^2/2)$ according to Eq. 8c. They differ significantly only for the small group size N, where the half-width w in Eqs. 5a, 5b and 8a, 8b, 8c is considerable (Fig. 1).

Our finding that the statistical distribution of the ratio m/m_{inp} is independent of the choice of particular values of the input Weibull parameters is in agreement with the results of other authors [11, 12], but this does not hold for the moments method [7]. As regards the biasing of the averaged Weibull modulus, other authors, for instance Peterlik, Orlovskaja et al. [4, 5], also found that the ML method overestimates the average Weibull modulus, while the scale parameter is slightly underestimated. On the other hand, the LR method with some frequently used, simple formulas for the median ranks tends to underestimate the Weibull modulus [7, 8, 25]. Much effort has been invested in optimizing the different estimation methods, in particular minimizing the biasing of the Weibull modulus [7, 8, 11–19].

The obtained values of the biasing factors, such as k_1 or k_2, with the help of Monte-Carlo simulations are useful, since by knowing them the experimentalist can correct (unbias) the ML estimations of the Weibull modulus from the experimental data. The simple idea is to divide the calculated Weibull modulus by the biasing factor to obtain a "more correct" (unbiased—UB) value. To follow the discussion from the experimentalist's point of view, let us denote the true (unknown) value of m for large collection of data just by m_{inp} (in the spirit of the input parameter for Monte-Carlo simulations). The estimated value for one test group is just m and the corresponding unbiased value is $m_{UB} = m/k(N)$. Of course, the correction with k does not ensure exact matching of m_{UB} with m_{inp} just for one group, but lowers this deviation on average (for several test groups chosen from the same large collection of data). The question arises as to which bias factor (≡unbias quotient) is more appropriate to use: k_1, k_2 or some other. This depends on the unbiasing criterion. If one wants to eliminate the sum of errors $m_{UB}-m_{inp}$ (but not the sum of their absolute values, see Eq. 10a below) for several test groups, the unbias quotient k_2 is the right choice. In fact, we have verified that the tabulated unbias factors in the standard ASTM C 1293—95 [24] for different N agree very well with our k_2; in the standard the corresponding multiplication with $1/k_2$ is used, so the unbias factors are smaller than one. However, there are objections against the use of unbias factors in the literature (see the comments below).

But what if one uses different criteria for the error in the evaluation of the Weibull moduli? Let us define three criteria of the relative error for a comparison:

$$e_1 = \frac{\left|\langle m_{UB} - m_{inp}\rangle\right|}{m_{inp}} \tag{10a}$$

$$e_2 = \frac{\langle\left|m_{UB} - m_{inp}\right|\rangle}{m_{inp}} \tag{10b}$$

$$e_3 = \frac{\langle(m_{UB} - m_{inp})^2\rangle^{1/2}}{m_{inp}}. \tag{10c}$$

Only the error e_1 can become zero for $m_{UB} = m/k_2(N)$ and for several test groups (or large N_{rep} in the Monte-Carlo simulation) since individual errors can have both signs and cancel. The other two error estimations, which are more common, are always positive.

In the next step of the Monte-Carlo simulations we relaxed the unbias quotient k, which can take continuous values other than the "optimal" value $k_2(N)$. We calculated the three errors (ten) for three different values of N. We did many repetitions N_{rep} to obtain a very accurate statistical averaging $\langle\ldots\rangle$ of the individual errors. The results are shown in Fig. 2. The minima of the errors e_2 and e_3 are shifted to higher k factors with respect to the factor k_2 from the standard that annihilates the error e_1. Thus, when different criteria for the error in the determination of the Weibull modulus are preferred, other unbias factors should be taken instead of the standard ones. Even more, Song et al. [17] warn against uncritically using different correction factors to the Weibull parameters from experimental data: when only one set of strength data is available, the correction factors should not even be applied. Although some standards, like ASTM C 1239—95 [24], demand the unbias factors in every case, this is questionable. This dilemma can be appreciated by not using the unbias factor to "correct"

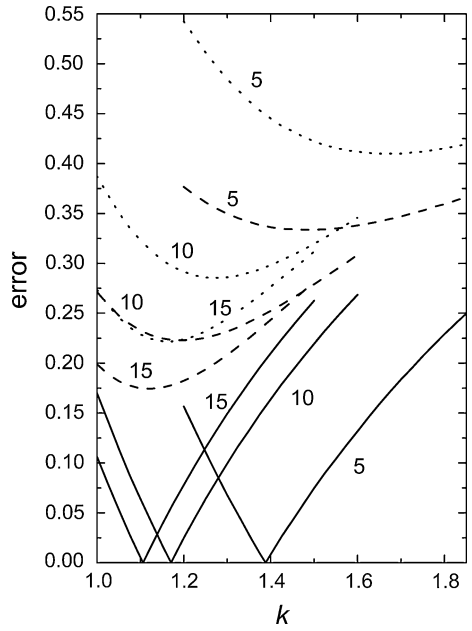

Fig. 2 The three relative statistical estimation errors as functions of unbias quotient k (so that $m_{UB} = m/k$ for each group); other parameters: $N_{rep} = 2{\cdot}10^6$, $m_{inp} = 10$, $\sigma_{0,inp}$ is arbitrary, $N = 5$, 10, and 15 (indicated on graphs). Errors: e_1 (Eq. 10a, *solid lines*), e_2 (Eq. 10b, *dashed lines*), e_3 (Eq. 10c, *dotted lines*)

the Weibull parameters for one set of experimental data; instead of this, the estimated values of these parameters should be accompanied by 90% confidence intervals.

Monte-Carlo simulations considering measurement uncertainty

Only when w_{MU} becomes comparable to $\delta\sigma$ does the measurement uncertainty influence the statistical results significantly, as has already been reported in our previous article where the LR method was used. The (virtual) decrease of m with increasing w_{MU} was to be expected since the measurement uncertainty widens the strength distribution function, in a similar way as a lower m would do. Here, we must emphasize that although w_{MU} has been incorporated into the simulation by (Eq. 7b), the inverse procedure of getting the Weibull parameters from the Monte-Carlo force data is still based on the function (Eq. 1a) for an exact measurement. The ratio between the virtual and the input Weibull modulus as a function of the ratio of the measurement uncertainty and the virtual standard deviation of the strength has been calculated in the following way. The obtained averaged modulus $\langle m_{DIR} \rangle$ is first corrected, i.e., divided by $k_2(N)$, to obtain the unbiased m_{MU}, which is still burdened with the measurement uncertainty error. Next, the standard deviation $\delta\sigma$ of the strength (Eq. 4b) is calculated for this m_{UM} and for the input scale parameter $\sigma_{0,inp}$ (since the output value of the scale parameter differs negligibly from the input value both for ML biasing and the measurement uncertainty contribution). The m_{MU}/m_{inp} dependence on $w_{MU}/\sigma_{0,inp}$ or on $w_{MU}/\delta\sigma$ is shown in Fig. 3 for $m_{inp} = 5$ and 15 and for $N = 5$ and 30. The curves for the different m_{inp} and N differ from each other and it would be impractical to use them in experiments for "unbiasing" m with respect to the measurement uncertainty. We would obtain similar curves if we took the normalization $w_{MU}/\delta\sigma_N$ (according to Eq. 4d and averaging such standard deviations over N_{rep}) instead of the theoretical $w_{MU}/\delta\sigma$ on the horizontal axis. However, we have checked that the following relation holds quite well, even for a large measurement uncertainty:

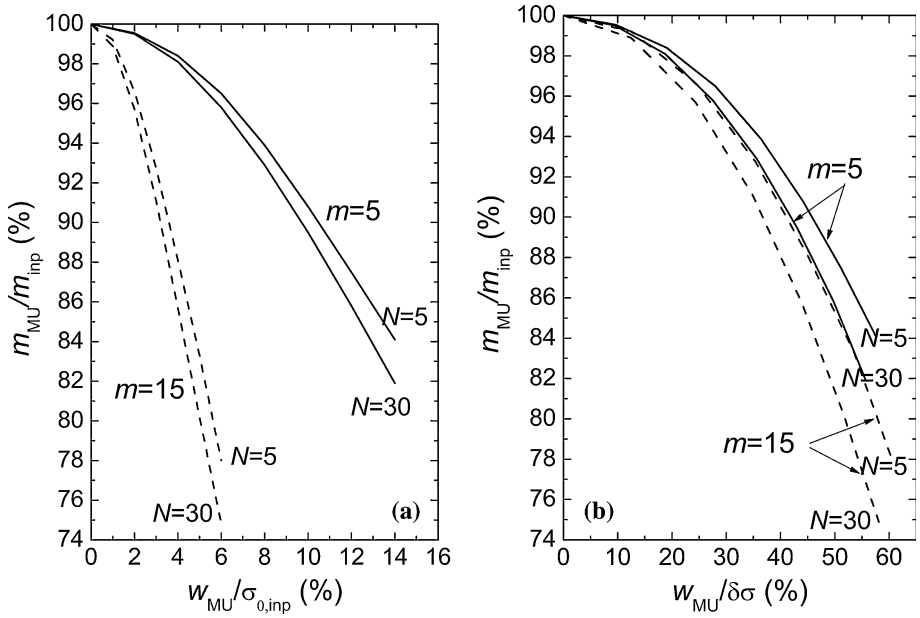

Fig. 3 Dependence of m_{MU}/m_{inp} on $w_{MU}/\sigma_{0,inp}$ **a** or $w_{MU}/\delta\sigma$ **b** for $m_{inp} = 5$ (*solid lines*) and 15 (*dashed lines*) and for $N = 5$ and 30. Other parameters: $N_{rep} = 10^6$, $\sigma_{0,inp}$ is arbitrary. Approximate matching of the curves for $m_{inp} = 5$, $N = 30$ and $m_{inp} = 15$, $N = 5$ in **b** is a coincidence

$$\Gamma\left(1+\frac{2}{m_{\mathrm{MU}}}\right)-\Gamma^2\left(1+\frac{1}{m_{\mathrm{MU}}}\right) = \Gamma\left(1+\frac{2}{m_{\mathrm{inp}}}\right)-\Gamma^2\left(1+\frac{1}{m_{\mathrm{inp}}}\right)+\left(\frac{w_{\mathrm{MU}}}{\sigma_{0,\mathrm{inp}}}\right)^2, \qquad (11)$$

i.e., the squared standard deviation of the strength according to the virtual m_{UM} is the sum of squared standard deviations for the input (real) Weibull distribution and for the Gaussian measurement error distribution. Equation 11 can thus be useful for experimentalists. First, the parameters m and σ_0 are obtained for real experimental data with the group size N. Next, m should be unbiased by dividing it with $k_2(N)$ to obtain m_{MU}. This value is inserted into Eq. 11, which is subsequently solved for the real parameter m (in the experiment m and σ_0 sit on the right-hand side of the equation instead of m_{inp} and $\sigma_{0,\mathrm{inp}}$).

Mixing experimental data

In the next step we "shuffled" well the 1000 experimental strength values using a random generator. The "true" values of the Weibull parameters were taken from the ML estimation for all 1000 values, since for $N = 1000$ both the bias and the width w of the distribution (Eqs. 5a, 5b) are very small. These ("true") values correspond to the input values for the Monte-Carlo data generation, and they are $m_{\mathrm{inp}} = 9.08$, $\sigma_{0,\mathrm{inp}} = 316.48$ MPa. For comparison, we calculated the Weibull parameters for all 1000 experimental strength values with the LR method using the simple rank-estimator function:

$$P_i = \frac{i-0.3}{N+0.4}, \qquad (12a)$$

which is often used in the literature [13, 17]. Several other rank estimators are mentioned in the literature [13, 17–19, 30], but for very large data sets, such as ours, any of them is applicable. We did not bother about choosing the most optimal of them, because in this investigation we focus on ML method, which does not use such estimators. We obtained the values $m = 9.11$ and $\sigma_0 = 316.48$ MPa, with the correlation coefficient of 99.84% (we use the regression on y-coordinate, i.e., minimizing the sum of the squares of vertical distances of data points from the straight line, see Ref. [28]). This is in a good agreement with the ML estimation. Figure 4 shows a plot of 1000 experimental measurements (diamond points), where quantities on x and y axis are $\ln(\sigma)$ and $\ln(\ln(1/1-P))$, respectively, so that weibull distribution follows a straight line. The fitting lines of ML and LR are so close together, that they are indistinguishable from each other, therefore only one line is drawn. A significant deviation of a few experimental data points for the lowest values of the strength from the straight line in the linear diagram in Fig. 4 are only artefacts of the

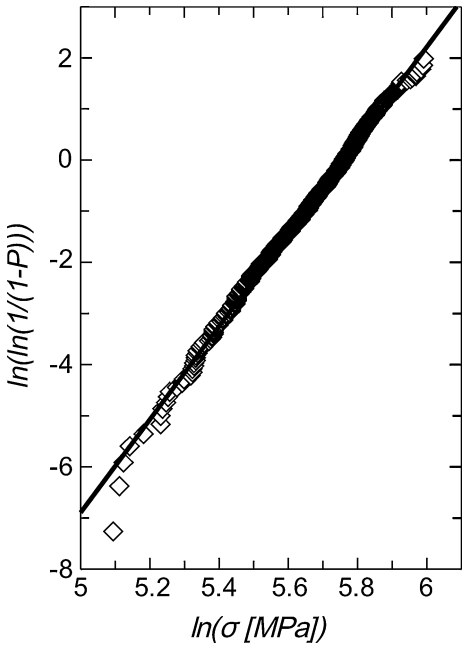

Fig. 4 Strength distribution of 1000 experimental measurements (*diamond*) and fitted Weibull distribution—*solid line*, which represents fits from both ML and linear regression technique

style of data plotting—the deviations are largely stretched for low values of P.

Different group sizes N were considered systematically again and the N values were randomly selected from $N_{\mathrm{exp}} = 1000$ values in each repetition. The bias factors k_1 and k_2 were calculated again and we found that they are slightly higher than those from the pure Monte-Carlo simulations above. The distribution of the values of both Weibull parameters fit the log-normal function very well, as for the Monte-Carlo generated data. The relation $k_2(N) = k_1(N)\cdot\exp(w(N)^2/2)$ still holds well. Thus, working with experimental data confirms the theoretical expectations for most points, except that the factors k_1 and k_2 are not exactly the same as theoretically predicted. For evidence and a comparison only the factor $k_{2,\mathrm{exp}}(N)$ corresponding to experimental version of Eq. 9b is shown in Fig. 1 (dashed line).

Finally, we performed the hybrid procedure described in subsection "Hybrid method: mixing a finite number of MC data". We generated the 1000 pieces of data for the material strength using some input Weibull modulus, e.g., 10 (there is only one repetition now). We stored these values in the file, as if they were real experimental data, and we subsequently used the mixing method described above. We found that the bias factor k_2 is similar to that obtained in the usual MC simulations described in subsection "Generating data by random generator (Monte-Carlo simulation)" (dotted line in Fig. 1). When we repeated the whole procedure on another set of 1000 MC

data, the result was almost the same. We concluded that mixing a finite number of data (either experimental or MC data) is not problematic if the data correspond well to the Weibull statistics.

There is a possible plausible explanation for the slight difference between the factors k_2 from the three methods. First of all, biasing of k_2 for $N = 1000$ is not the right cause because it is too small: $k_2(1000) = 1.003$ according to the method "Generating data by random generator (Monte-Carlo simulation)". One reason for the deviations may be possibly non-ideal Weibull distribution. Another reason is a small but still non-negligible standard deviation in the distribution of the logarithms of m: Monte-Carlo simulations give $w = 0.25$ in the distribution (Eq. 5a, 5b) for $N = 1000$. This means that the estimated m from the 1000 samples is in the range from 97.6 to 102.5% of its "true" value (corresponding to "Hybrid method: mixing a finite number of MC data" or to hypothetical experimental distribution with infinitely many samples) with the probability of about 2/3. These differences in the estimated m are directly manifested in the differences in k_1 and k_2. But the most probable reason seems to be the fact that particularly experimental data lack the theoretically possible very low and very high strength values, as mentioned above.

Finally, we mention the comparison of fitting experimental data to Weibull and other generally known distributions, such as Gaussian. This can be done by comparing the goodness of resulting fit parameters [31]. We have used the LR method to fit the 1000 experimental strength values to Gaussian distribution and obtained the correlation coefficient 98.89% (compared to 99.84% for Weibull distribution, see above). This difference is too small to disregard distributions other than Weibull. There are some reasonable objections against the non-critical use of the Weibull distribution. For instance, this distribution is theoretically based on the power-law distribution of the flaw sizes in the specimens which validity may be very limited, particularly in the limit of very small flaws [30].

Conclusions

We found that the distribution of both Weibull parameters using the ML estimation on either Monte-Carlo or mixed experimental data agrees well with the log-normal distribution. The biasing of the Weibull parameter is significant for smaller sizes of the data group, but the Monte-Carlo and experimental data result in slightly different bias factors. We have pointed out that the evaluation of the unbiasing factors for the Weibull modulus must be done very carefully, according to the preferential criterion of the estimation accuracy. We also showed that the measurement uncertainty can be included in the estimation procedure well by the rule of adding the squares of errors of different origin.

Acknowledgements This research was supported by the Ministry of Education and Sport of Republic of Slovenia and the European Social Fund. We thank the company Hidria AET for providing the experimental data.

References

1. Weibull W (1951) J Appl Mech 18:293
2. ReliaSoft's Weibull (1992) ++, Life Data Analysis Reference. ReliaSoft Publishing
3. Curtis RV, Juszczyk AS (1998) J Mater Sci 33:1151. doi:10.1023/A:1004361222711
4. Orlovskaja N, Peterlik H, Marczewski M, Kromp K (1997) J Mater Sci 32:1903. doi:10.1023/A:1018521310570
5. Peterlik H, Orlovskaja N, Steinkellner W, Kromp K (2000) J Mater Sci 35:707. doi:10.1023/A:1004757317724
6. Li QS, Fang JQ, Liu DK, Tang J (2003) Cem Concr Res 33:1631
7. Wu D, Zhou J, Li Y (2006) J Mater Sci 41:5630. doi:10.1007/s10853-006-0344-9
8. Wu D, Zhou J, Li Y (2006) J Eur Ceram Soc 26:1099
9. Peterlik H (1995) J Mater Sci 30:1972. doi:10.1007/BF00353020
10. Danzer R, Lube T, Supancic P (2001) Z Metallkunde 92:773
11. Bergman B (1984) J Mater Sci Lett 3:689
12. Khalili A, Kromp K (1991) J Mater Sci 26:6741. doi:10.1007/BF02402669
13. Gong J (2000) J Mater Sci Lett 19:827
14. Barbero E, Fernandez-Saez J, Navarro C (2001) J Mater Sci Lett 20:847
15. Davies IJ (2001) J Mater Sci Lett 20:997
16. Wu DF, Li YD, Zhang JP, Chang L, Wu DH, Fang ZP, Shi YH (2001) Chem Eng Sci 56:7035
17. Song L, Wu D, Li Y (2003) J Mater Sci Lett 22:1651
18. Griggs JA, Zhang Y (2003) J Mater Sci Lett 22:1771
19. Davies IJ (2004) J Mater Sci 39:1441. doi:10.1023/B:JMSC.0000013913.84004.cd
20. Tanaka T, Nakayama H, Sakaida A, Imamichi T (1995) Mater Sci Res Int 1:51
21. Faucher B, Tyson WR (1988) J Mater Sci Lett 7:1199
22. Langlois R (1991) J Mater Sci Lett 10:1049
23. Cacciari M, Mazzanti G, Montanari GC (1996) IEEE Trans On El Ins 3:18
24. ASTM C 1239—95 (1995) Standard practice for reporting uniaxial strength data and estimating Weibull distribution parameters for advanced ceramics. American Society for Testing and Materials, Philadelphia
25. Jacquelin J (1993) IEEE Trans On El Ins 28:65
26. Hirose H (1999) IEEE Trans On Dielec El Ins 6:66
27. Kantar YM, Senoglu B (2008) Comput Geosci 34:1900
28. Ambrožič M, Vidovič K (2007) J Mater Sci 42:9645. doi:10.1007/s10853-007-1967-1
29. ASTM C 1161—94 (1994) Standard test method for flexural strength of advanced ceramics at ambient temperature. American Society for Testing and Materials, Philadelphia
30. Danzer R (2006) J Eur Ceram Soc 26(15):3034
31. Lu C, Danzer R, Fiescher FD (2002) Phys Rev E 65:067102