JOURNAL OF MATERIALS SCIENCE 30 (1995) 1972–1976

The validity of Weibull estimators

H. PETERLIK
Institut für Festkörperphysik, Universität Wien, Boltzmanngasse 5, A-1090 Vienna, Austria

The parameters of the two-parametric Weibull distribution, the Weibull modulus and the scale parameter, were estimated by using not only analytical means but also Monte-Carlo simulations. The precision of the measurement of both parameters, i.e. their variation coefficient, has been calculated. It is shown that the variation coefficient of the scale parameter is dependent on the number of experiments, *M*, which were performed, and on the Weibull modulus itself, whereas the variation coefficient of the Weibull modulus is only dependent on *M*. Furthermore, the correctly interpreted results show that each single measurement gives the statistically correct Weibull parameters and the biasing arises only from the method of adding single measurements to obtain a mean value. Thus, in practice, when only one set of experiments for further evaluation is available, there is no need for adjustment factors.

1. Introduction

The Weibull distribution has been widely used to describe the statistical behaviour of the fracture of ceramics [1, 2]. It is based on the "weakest-link hypothesis", which means that the most serious flaw controls the strength. As in processing ceramic powders of different sizes and different compositions are used, there may arise quite complex flaw distributions and these may effect, for example, two superimposing Weibull distributions. This is one of the limits of the description of the fracture behaviour of brittle materials by the two-parametric Weibull distribution. Another limit is due to the statistical character of the Weibull distribution itself, which leads to an uncertainty in the parameters obtained by evaluation from a limited number of experiments. The purpose of the present work was to clarify the latter point.

Thus, it is assumed that the fracture stresses of a ceramic are distributed according to the two-parametric Weibull distribution, the two parameters being the Weibull modulus and the scale parameter. As already mentioned, in practice it is only possible to test a limited number of specimens. Thus the parameters of the true distribution can only be determined with a certain accuracy with respect to time and money of the material producer and the testing institution. The number of tests is the basic limit for the determination of the parameters, even if the experiments are carried out with the utmost care. The dispersion of test results is here only due to the statistical character of the fracture behaviour.

The precision of the measurement of the crack extension parameters has been investigated by Ritter *et al.* [3]. In their work, the precision of the measurement of the Weibull parameters, evaluated by linear regression, has also been investigated. Many important relations can be found in their paper. It consists of Monte-Carlo simulations of experiments as well as of analytical results. Owing to the development of computers since that time, it is now possible to carry out 10^4 times more simulated experiments, which reduces the expected error, in a rough estimation, by the square root of this number, i.e. by a factor of a hundred. In this way it is possible to control the precision of the numerical simulations with analytical results, which can be found for some special cases.

The validity of Weibull estimators has been an interesting theme of scientific research over 25 years [4–10]. Much effort was spent on the discussion of the precision and the biasing of the parameters determined by different evaluation methods, e.g. linear regression, moments and maximum likelihood method. Adjustment factors were developed to correct the biasing [4, 6, 7, 9]. The scope of the present work was to show that each single measurement of the Weibull parameters is correct and the biasing arises only from the method of adding the parameters, if one tries to obtain a mean value from a number of measurements of different laboratories from the same material. Hence, adjustment factors should not be applied, if only one set of tests is available and thus only one evaluation could be performed.

2. Theoretical background

In describing the fracture behaviour of ceramics, the two-parameter Weibull distribution is widely used

$$P_f = 1 - \exp\left[- \left(\frac{\sigma}{\sigma_0}\right)^m \right] \quad (1)$$

where P_f is the failure probability, σ the fracture strength and σ_0 and m are the Weibull parameters, the scale parameter and the Weibull modulus, respectively. In practice, it is only possible to perform a limited number of experiments to determine the true

0022–2461

values of the distribution, σ_0 and m. Then, in the well-known maximum likelihood evaluation method, one obtains parameters σ_{ML} and m_{ML} from the measurement of M fracture strength values σ_j, $j = 1 \ldots M$ and calculating m_{ML} from the implicit equation

$$\frac{M}{m_{ML}} + \sum_{j=1}^{M} \ln\sigma_j - M \frac{\sum_{j=1}^{M} (\sigma_j)^{m_{ML}} \ln \sigma_j}{\sum_{j=1}^{M} (\sigma_j)^{m_{ML}}} = 0 \quad (2a)$$

and σ_{ML} from

$$\mathscr{F}(\sigma_j) \equiv \sigma_{ML} = \left(\frac{\sum_{j=1}^{M} (\sigma_j)^{m_{ML}}}{M} \right)^{1/m_{ML}} \quad (2b)$$

Because the true parameters, σ_0 and m, and the parameters obtained by experiments and subsequent evaluation by the maximum likelihood method, do not necessarily coincide, the latter are denoted by the subindex ML.

The method used to obtain expectation values of measurements has been outlined elsewhere [11], where the expectation value of the subcritical crack growth parameter, n, and the standard deviation, Δn, was computed. The calculation of the expectation values of σ_{ML} and m_{ML} is quite similar: the variables σ_j obey individual Weibull distributions, see Equation 1. Hence expectation values of measurements are predicted as integrations with respect to the measure

$$dW = \prod_j d\sigma_j P'(\sigma_j) \quad (3)$$

If σ_{ML} denotes the outcome of a single measurement of $\mathscr{F}(\sigma_j)$, see Equation 2b, the expectation value is obtained by integration

$$\langle \sigma_{ML} \rangle = \int dW \, \mathscr{F}(\sigma_j) \quad (4)$$

The same procedure may be applied for computing the variance, defined as the square of the standard deviation $\Delta\sigma_{ML}$

$$(\Delta\sigma_{ML})^2 = \langle (\sigma_{ML})^2 \rangle - \langle \sigma_{ML} \rangle^2 \quad (5)$$

3. Precision of the scale parameter

The precision of the determination of the scale parameter, σ_0 was investigated. It is shown that for the maximum likelihood evaluation and for known m there exist analytical solutions for σ_{ML} and $\Delta\sigma_{ML}/\sigma_0$ which depend on the Weibull modulus, m, and the number of experiments, M, which are performed to determine σ_{ML}. The analytical solution shows the main features. The functional dependence for the general case (m unknown, see Section 3.2) is nearly identical with the special case (m known, see Section 3.1). Because of the complex way of determining m, the general case was only numerically calculated by Monte-Carlo simulations and compared with the analytical solution for known m.

3.1. Case one: the true Weibull parameter, m, is known

If the Weibull modulus, m, is known and constant, the scale parameter is obtained from Equation 2b. The expected value of σ_{ML} is then

$$\begin{aligned} \langle \sigma_{ML} \rangle &= \int dW \left(\frac{\sum_{j=1}^{M} (\sigma_j)^m}{M} \right)^{1/m} \\ &= \sigma_0 M^{-1/m} \int \prod_{j=1}^{M} dx_j e^{-x_1 - \cdots - x_M} \\ &\quad \times (x_1 + \cdots + x_M)^{1/m} \end{aligned} \quad (6)$$

by using $x_j = (\sigma_j/\sigma_0)^m$ as a new variable. Now, the number of integrals, which have to be solved, can be reduced to one by the following coordinate transformation

$$\begin{aligned} x_1 &= y_1 \\ &\vdots \\ x_{M-1} &= y_{M-1} \\ x_1 + \cdots + x_M &= y_M \end{aligned} \quad (7)$$

With the abbreviation $y = y_M$, it follows that

$$\langle \sigma_{ML} \rangle = \sigma_0 M^{-1/m} \int_0^\infty \left[\int dy_1 \cdots \int dy_{M-1} \right] dy \times y^{M-1} e^{-y} y^{1/m} \quad (8a)$$

$$y_1 + \cdots + y_{M-1} < 1 \quad (8b)$$

$$y_i > 0 \quad (8c)$$

The integrals in the brackets are determined by replacing $y^{1/m} \to 1$ in Equation 8, because equations 6–8 hold for arbitrary integrands. The integrals in the brackets thus evaluate to $1/\Gamma(M)$. Hence, we arrive at the solution

$$\langle \sigma_{ML} \rangle = \sigma_0 M^{-1/m} \frac{\Gamma[M + (1/m)]}{\Gamma(M)} \quad (9)$$

which gives the true parameter, σ_0, only in the limits $M \to \infty$ or $m \to \infty$. It was this result, which led to misunderstandings in the data obtained by numerical simulation in the literature [4–10]. As can be seen from this equation, the expectation value of the scale parameter, σ_{ML}, obtained by an evaluation of measured stresses, is not the true scale parameter, σ_0. This has induced discussion about the biasing of different evaluation procedures for the parameters of the Weibull distribution and the use of correction factors [4, 6, 7, 9]. But the correct interpretation from a statistical point of view is the following: N different working groups measure scale parameters from the same material, denoted by $\sigma_{ML,i}$, and each of these scale parameters was obtained by M independent fracture stress tests. Then someone collects the data and computes the arithmetic mean to obtain a more precise value for the scale parameter. Even for an infinite number of data from different working groups

he does not obtain the true scale parameter

$$\langle \sigma_{ML} \rangle = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \sigma_{ML,i}$$

$$= \sigma_0 M^{-1/m} \frac{\Gamma[M + (1/m)]}{\Gamma(M)} \neq \sigma_0 \quad (10)$$

The reason being that the distribution of $\sigma_{ML,i}$ is asymmetric and the arithmetic mean is not the correct way to calculate the true value. Though it might be a very rare case, in practice, that many different laboratories measure data from the same material, there is, however, a much simpler way to obtain the correct expectation value. One has to choose the scale in the right way to symmetrize the distribution of the scale parameters. For the scale parameter this is quite simple, taking into account the functional form of the Weibull distribution. One has to add the measurements of the different working groups analogously to Equation 2b by adding the power of m of the measured scale parameters

$$\lim_{N \to \infty} \left[\frac{1}{N} \sum_{i=1}^{N} (\sigma_{ML,i})^m \right]^{1/m} = \sigma_0 \quad (11)$$

But the usual case is that *only one* scale parameter from one material (evaluated from a certain number, M, of tests) is available. *Each* evaluation by the maximum likelihood method gives the *correct* result. Only if *more* scale parameters were measured from one material by independent sets of tests, does the arithmetic mean not give the true parameter σ_0, but has to be corrected by Equation 9 or has to be added in a different way by Equation 11. In other words, a biasing of experimentally obtained parameters arises not from a biasing of single measurements, but from the procedure to calculate the arithmetic mean of independent (correct, but asymmetrically distributed) measurements. This means that only the method of adding data arithmetically is inappropriate to obtain the true values of the parameters. A single measurement, however, is correct and one cannot draw any conclusions about a biasing. Thus, discussions about biasing and adjustment factors, should now cease, because they should not be applied if only one set of data (from a certain number of tests) was measured, which is the usual case in practice.

The essential information about the accuracy and quality of an evaluation procedure can be derived from the standard deviation and the variance, as defined in Equation 5. These are variables, which give the minimal possible dispersion. Even with the best experimental equipment one cannot obtain more precise values, because this dispersion is only due to the nature of the inherent failure mechanisms and the arising distribution of fracture stress values. It should be noted that, for asymmetric distributions, confidence intervals are better suited to describe the fracture behaviour [5, 10]. For the sake of simplicity, however, we restrict ourselves to a symmetric description, see Equation 5. Then the variation coefficient, $\Delta\sigma_{ML}/\sigma_0$ is calculated analogously to Equations 6–9. The result is

$$\left(\frac{\Delta\sigma_{ML}}{\sigma_0} \right)^2 = M^{-2/m} \left(\frac{\Gamma[M + (2/m)]}{\Gamma(M)} - \left\{ \frac{\Gamma[M + (1/m)]}{\Gamma(M)} \right\}^2 \right) \quad (12)$$

This is the minimal standard deviation, which depends only on the statistical nature of the failure mechanisms of a ceramic and can thus by no means be experimentally improved. As will be shown in Section 3.2, for the general case (m unknown), $\Delta\sigma_{ML}/\sigma_0$ shows principally the same functional behaviour on the number of measurements, M, and differs only by a small constant factor.

Another possible means of evaluating the Weibull parameters is the linear regression. The parameters obtained by this evaluation procedure are denoted by the subindex LR, i.e. m_{LR} and σ_{LR}. In this case, these Weibull parameters are obtained from a diagram of the logarithmic strength values versus appropriate values of the fracture probability $\ln \ln (1 - P_f)^{-1}$, where for $j = 1, \ldots, M$ experiments

$$P_f = \frac{j - 0.5}{M} \quad (13)$$

has been recommended [10]. Because here the fracture probability is chosen by appointment, unfortunately one cannot use the above approach for the maximum likelihood method, where the fracture probabilities, P_f, are abitrary (statistically distributed). Contrary to the linear regression method, the maximum likelihood method offers a number of analytical possibilities to calculate expectation values, e.g. the distribution of the logarithms of the experimentally obtained scale parameters in terms of the Gamma-function and its derivatives

$$\langle \ln \sigma_{ML} \rangle = \ln \sigma_0 \left[\frac{\Gamma'(M)}{\Gamma(M)} - \ln M \right] \quad (14)$$

Another argument for the use of the maximum likelihood procedure is the slightly better accuracy, described by the smaller variation coefficient; see the results in Section 3.2.

3.2. Case two: the true Weibull parameter m is unknown

This is the general case, in which the Weibull modulus, m_{ML}, is obtained by a set of experiments, evaluated from Equation 2a, and the scale parameter, σ_{ML} is evaluated from Equation 2b. An analytical solution could not be found, because m_{ML} is only given by the implicit and transcendent Equation 2a. By performing numerical calculations and comparing them with the analytical solution for known m, see Equation 12, it fortunately turns out that the functional dependence of σ_{ML} on the number of experiments, M, is of the same behaviour as for the case treated in Section 3.1 and differs only by a factor, α, close to 1. The

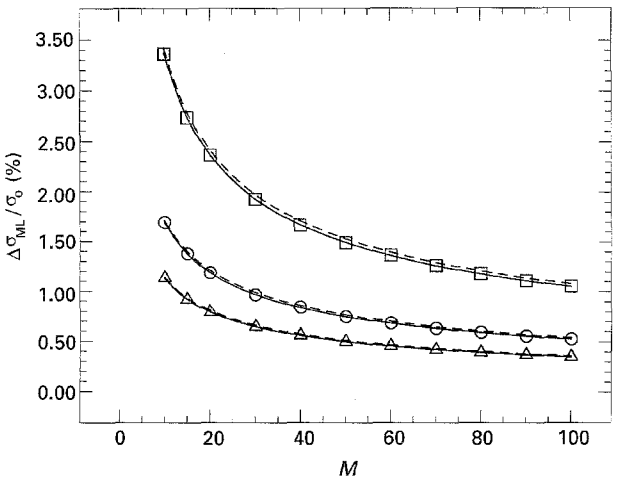

Figure 1 The variation coefficient $\Delta\sigma_{ML}/\sigma_0$ obtained by the maximum likelihood evaluation of numerically simulated experiments for $m = (\square)$ 10, $(\bigcirc)$ 20 and $(\triangle)$ 30, depending on the number, M, of experiments performed. (——) Obtained from equation 15a, (– – –) obtained by evaluation from linear regression, with fracture probabilities chosen according to equation 13.

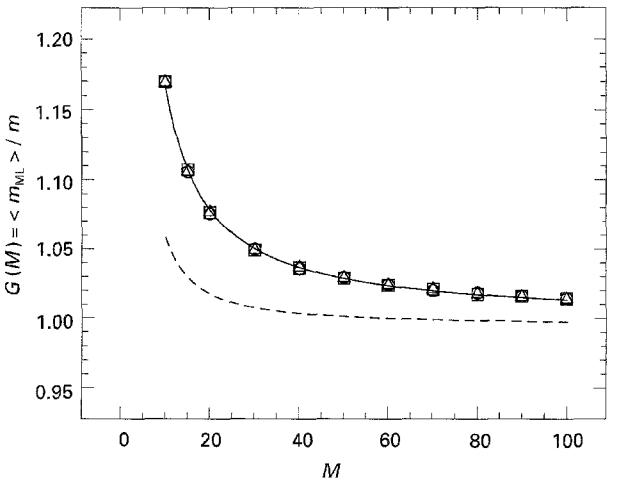

Figure 2 Dependence of the correction function $G(M)$ on the number, M, of experiments for $m = (\square)$ 10, $(\bigcirc)$ 20 and $(\triangle)$ 30. (——) Calculated from equation 16; (– – –) obtained by evaluation from linear regression.

expectation value is then

$$\left(\frac{\Delta\sigma_{ML}}{\sigma_0}\right)^2 = \alpha M^{-2/m}\left(\frac{\Gamma[M + (2/m)]}{\Gamma(M)} - \left\{\frac{\Gamma[M + (1/m)]}{\Gamma(M)}\right\}^2\right) \quad (15a)$$

with

$$\alpha = 1.05 \pm 0.003 \quad (15b)$$

where the error in Equation 15b is obtained from the numerical simulations in the range of the number of experiments from $M = 10$–100. In Fig. 1 the numerically simulated results for the variation coefficient $\Delta\sigma_{ML}/\sigma_0$ can be seen for three different Weibull moduli, $m = 10$, 20 and 30. The solid lines are obtained from Equation 15a and fit the numerical data very well. For comparison, the dashed lines represent the results for $\Delta\sigma_{LR}/\sigma_0$ from an evaluation by linear regression using Equation 13 as definition for the probability of fracture. The variation coefficient of the linear regression method is slightly higher than that of the maximum likelihood evaluation procedure.

Using Equation 15a the lowest bound for the determination of the scale parameter may be calculated for an arbitrary number of tests performed, or the number of tests necessary to obtain a certain accuracy may be calculated in advance.

4. Experimentally obtained Weibull-modulus, m_{ML}

As has been already mentioned, the dependence of the experimentally determined Weibull modulus, m_{ML}, on the number of tests performed cannot be computed by the method outlined above, because it is only given by an implicit and transcendent equation. Thus, the dependence of m_{ML} and $\Delta m_{ML}/m$ is given by a polynomial fit of the numerically simulated tests. If a number of measured m_{ML} is available, and the respective values are added arithmetically, one has to correct the obtained m_{ML} by a function $G(M)$ to obtain the correct value. Usually there is only one case in practice: if one calculates the crack extension parameter, n, from experiments with constant loading rate, the proposal for the CEN-standard [12] requires ten tests at five loading rates each differing by one order of magnitude. Thus, five different Weibull moduli are obtained. The corrected mean value gives the highest precision (the lowest standard deviation), if the evaluation of the crack extension parameter is performed by the maximum likelihood procedure [13]. Fig. 2 shows the dependence of the correction function $G(M)$ on M. It is independent of the Weibull modulus (see numerically computed values for $m = 10$, 20 and 30. This observed independence of the actual value of m is in agreement with previous analytical considerations of Bain and Antle [4]. The solid line in Fig. 2 is calculated by the fit

$$G(M) \equiv \frac{\langle m_{ML}\rangle}{m} = 1 + 2.1049\, M^{-1.1} \quad (16)$$

The dashed line, obtained by the linear regression method with the fracture probabilities chosen according to equation 13, is shown for comparison. An arithmetic mean from m_{LR}-values (calculated by linear regression) is less biased than an arithmetic mean from m_{ML} values (calculated by maximum likelihood), which means that the result of the arithmetic mean is not as bad for the linear regression procedure as for the maximum likelihood evaluation.

But really relevant and interesting is the accuracy of the evaluation procedure, described by the variation coefficient $\Delta m_{ML}/m$. As is shown in Fig. 3, it is independent of the actual value of m. The solid line is calculated from a fit given by

$$\frac{\Delta m_{ML}}{m} = 0.042\,22 + 2.3375\, M^{-0.8836} \quad (17)$$

Again, for comparison, the dashed line is computed by the linear regression method and has a slightly higher standard deviation than the maximum likelihood

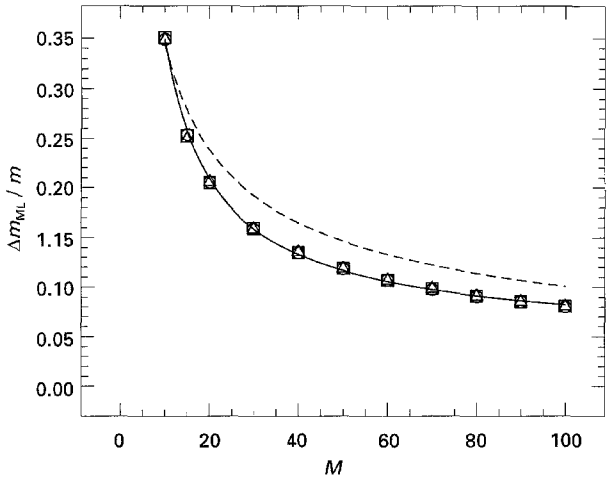

Figure 3 The variation coefficient of the Weibull parameter m, $\Delta m_{ML}/m$, for $m = (\square)$ 10, $(\bigcirc)$ 20, and $(\triangle)$ 30, depending on the number of experiments, M. (——) Obtained from equation 17; (– – –) linear regression is shown for comparison.

evaluation procedure. From the two fits, which are valid to a precision of less than 2% in the range $M = 10$–100, one can calculate the lower bound of the standard deviation for the experimentally obtained Weibull parameter m_{ML}.

5. Conclusion

By analytical solutions and numerical calculations it has been shown that the maximum likelihood evaluation procedure gives correct results for the Weibull parameters determined from a set of experiments. Only if more than one working group measured data from the same material, must the data be corrected, if arithmetically added. The same relation holds for other evaluation procedures, such as the linear regression method. The variation coefficient, i.e. the standard deviation divided by the mean of the value itself, is given for both Weibull parameters, as the dependence on the number of measurements, M. It is slightly lower for the maximum likelihood than for the linear regression evaluation procedure.

Acknowledgement

The support of the Austrian Fonds zur Förderung der wissenschaftlichen Forschung, project P8990-TEC, is gratefully acknowledged.

References

1. W. WEIBULL, *J. Appl. Mech.* **18** (1951) 293.
2. D. MUNZ and T. FETT, "Mechanisches Verhalten keramischer Werkstoffe" (Springer, Berlin, 1989).
3. J. RITTER, N. BANDYOPADHYAY and K. JAKUS, *Am. Ceram. Soc. Bull.* **60** (1981) 798.
4. L. J. BAIN and C. E. ANTLE, *Technometrics* **9** (1967) 621.
5. D. R. THOMAN, L. J. BAIN and C. E. ANTLE, *ibid.* **11** (1969) 445.
6. K. TRUSTRUM and A. DE S. JAYATILAKA, *J. Mater. Sci.* **14** (1979) 1080.
7. B. BERGMAN, *J. Mater. Sci. Lett.* **3** (1984) 689.
8. J. D. SULLIVAN and P. H. LAUZON, *ibid.* **5** (1986) 1245.
9. B. FAUCHER and W. R. TYSON, *ibid.* **7** (1988) 199.
10. A. KHALILI and K. KROMP, *J. Mater. Sci.* **26** (1991) 6741.
11. H. PETERLIK, *J. Eur. Ceram. Soc.*, **13** (1994) 509.
12. CEN standard draft EN 843-3: "Determination of subcritical crack growth parameters from constant stressing rate flexural strength tests, 1992".
13. H. PETERLIK and K. KROMP, *Fortschrittsber. Deutsch. Keram. Ges.* **9** [1] (1994) 78.

Received 3 June
and accepted 24 August 1994