Sigma J Eng Nat Sci, Vol. 43, No. 5, pp. 1691−1700, October, 2025

Sigma Journal of Engineering and Natural Sciences
Web page info: https://sigma.yildiz.edu.tr
DOI: 10.14744/sigma.2025.00158

Research Article

Unbiased weighted least squares estimation of Weibull modulus with small samples

Burak BİRGÖREN[1,*]

[1]*Department of Industrial Engineering, TOBB University of Economics and Technology, 06510, Türkiye*

ARTICLE INFO

Article history
Received: 18 March 2024
Revised: 12 June 2024
Accepted: 05 August 2024

Keywords:
Ceramics; Composite Materials; Fracture Strength; Reliability; Weibull Distribution; Weibull Modulus; Weighted Linear Least Squares; Unbiased Estimation

ABSTRACT

Weibull distribution has been used vastly for modeling the fracture strength of ceramic and composite materials. Estimating the Weibull modulus is an important problem, particularly in cases where the sample size is small due to high experimental costs. The Ordinary Least Squares (OLS) method is the most commonly used method by materials scientists. Numerous probability estimators have been proposed using the OLS method, and most of these studies focused on unbiased estimation. Weighted Least Squares (WLS) is a promising alternative to OLS, yet, there are only a few studies on this subject. This work followed a systematic analysis to develop a new probability index for unbiased estimation of Weibull modulus using WLS. For sample sizes less than 60, the performance of the new index is shown to be significantly superior to currently used OLS methods and is better or as good as the Maximum Likelihood Estimation method.

Cite this article as: Birgören B. Unbiased weighted least squares estimation of Weibull modulus with small samples. Sigma J Eng Nat Sci 2025;43(5):1691−1700.

INTRODUCTION

The Weibull distribution is one of the most widely used probability distributions in life testing and reliability studies. The distribution was developed by Weibull [1] and has found use in many different areas of application such as wind-speed data analysis [2–5], unemployment duration analysis, survival analysis as well as reliability analysis [6]. One major application area is the modelling statistical variation in the fracture strength of many materials such as advanced ceramics, metallic matrix composites and ceramic matrix composites [7–9], and strength of micro and nano structures [10]; it is also used to describe the fracture toughness behavior of steels in ductile-brittle transition region [11], and fatigue behavior of metals [12,13].

The two-parameter Weibull distribution function is given by

$$F(s) = 1 - \exp\left[-(s/s_0)^m\right] \tag{1}$$

where m is the shape parameter or Weibull modulus, s_0 is the scale parameter or characteristic strength of the distribution, and F is the fracture probability of the material at or below uniaxial tensile stress s. Weibull modulus, m, reflects the inherent variability in strength of a material: the higher

***Corresponding author.**
*E-mail address: b.birgoren@etu.edu.tr

This paper was recommended for publication in revised form by Editor-in-Chief Ahmet Selim Dalkilic

the m the lower the dispersion of fracture stress. It is generally independent of the size of the material, assuming the material composition and microstructure remain consistent. For example, if test specimens with different cross-sectional areas used in uniaxial compression and extension tests, their Weibull modulus remains the same. The scale parameter, on the other hand, is closely related to the mean fracture stress, and it changes with the cross-sectional area.

The Weibull parameters m and s_0 are estimated from a sample of strength measurements. High experimental costs limit the number of samples to be tested. Therefore, methods achieving high estimation precision with small sample sizes are highly desirable [14]. In practice, the Ordinary Least Squares (OLS) method and the Maximum Likelihood Estimation (MLE) method have been the most commonly used methods for estimating these parameters in general reliability applications [15]. In materials science, the OLS method remains the most popular due to its simplicity and ease of use [16], while the MLE method is considered a standard method in important references such as [17] due to its well-known distributional optimality properties in large samples [15]. Both of the methods produce biased estimators, therefore there are many studies for bias correction. One approach is using empirical correction factors, [18] is an example for MLE, and [19–21] are some recent studies for OLS; and the other approach is to satisfy unbiasedness through developing a probability index to be used in OLS. Davies [16,22] provides a detailed literature review and a thorough discussion of how such indices can be developed. Despite this focus on the two-parameter model, recent research has explored the potential of the three-parameter Weibull distribution in materials science, although its practical application remains limited [23–27].

One disadvantage of OLS is that it treats each measurement equally under the assumption that the variance of the error term is a constant. However, this assumption is violated for the OLS method used for Weibull parameter estimation. Several authors have proposed the Weighted Least Squares (WLS) method as a remedy, proposing different weight factors [28–33]. Their studies showed that the WLS methods show significantly superior performance as compared to the OLS methods.

On the other hand, the WLS methods also result in biased estimators, and there has been no study in the literature for unbiased estimation of Weibull parameters by WLS to the best of our knowledge. Therefore, the main purpose of this study is to develop one such method for the unbiased estimation of the Weibull modulus. As in the OLS case, empirical correction factors and developing a new probability index are two viable approaches. This study prefers the latter because it allows decreasing estimation variance both through weight factors and fine-tuning coefficients used in the probability index.

Finally, all the Monte Carlo simulations are coded and run in the R programming language which uses Mersenne-Twister random number generator as the default generator whose cycle period is 2^{19937}-1 [34]. The following sections are organized as follows: After a concise discussion of the OLS and WLS methods, a systematic analysis is introduced to develop a new probability index for unbiased estimation of Weibull modulus using Monte Carlo simulations as the basic tool. The probability index, proposed as a result of this analysis, is then compared with the MLE, the OLS and the WLS methods using different criteria. Finally, implementation on a practical example is discussed.

Ordinary Least Squares (OLS) Method

Eq. (1) becomes a straight line by a double logarithmic transformation:

$$\ln\left[-\ln\left(1-F\left(s\right)\right)\right] = m\ln s - m\ln s_0 \tag{2}$$

In order to apply the OLS method, the measurements are ranked from the smallest to the largest. F-values are assigned according to the rank i of a measurement, s_i denoting the ith smallest: $s_1 \leq s_2 \leq \ldots \leq s_n$. Then, in Eq. (2), $F(s)$ can be replaced by $F(s_i)$, and ln s can be replaced by ln s_i. The most commonly used probability estimators of $F(s_i)$, which are simply denoted as F_i, are median rank,

$$F_i = \left(i-0.3\right)/\left(n+0.4\right), \tag{3a}$$

mean rank,

$$F_i = i/\left(n+1\right), \tag{3b}$$

and hazen rank,

$$F_i = \left(i-0.5\right)/n \tag{3c}$$

There are many other estimators proposed for this purpose [16]. Considering the familiar form of a regression equation, $Y = aX + b$, the left side of Eq. (2) corresponds to Y, ln s corresponds to X, m corresponds to a, and $-m \ln s_0$ corresponds to b. Using s_i and F_i pairs in Eq. (2), a and b are obtained by the OLS procedure. Then the Weibull parameter estimates are calculated as $\hat{m} = a$ and $\hat{\sigma}_0 = \exp\left(-b/\hat{m}\right)$.

Weighted Least Squares (WLS) Method

Weighted least squares (WLS) extends ordinary least squares regression by incorporating weights into the sum of squared residuals. These weights, denoted by w_i for the ith observation, downplay the influence of data points s_i with higher variance or lower reliability. The regression model in Eq. (2) has non-constant error variance. As a result use of the weighted least squares method gives estimates with better statistical properties: Bergman [30], Faucher and Tyson [29] and Hung [31] proposed using the following weight factors in the WLS method, respectively:

$$w_i = \left[\left(1-F_i\right)\ln\left(1-F_i\right)\right]^2 \tag{4a}$$

$$w_i = 3.3F_i - 27.5\left[1-\left(1-F_i\right)^{0.025}\right] \tag{4b}$$

$$w_i = \frac{\left[\left(1-F_i\right)\ln\left(1-F_i\right)\right]^2}{\sum\left[\left(1-F_i\right)\ln\left(1-F_i\right)\right]^2} \tag{4c}$$

where hazen rank is used in Eq. (4a), and median rank is used in Eq. (4b) and (4c). Lu et al. [28] showed that $Z = -ln(1 - F(s))$ in Eq. (2) follows a standard exponential distribution. Then Eq. (2) takes the form of

$$\ln\left(Z_i\right) = m\ln\left(s_i\right) - m\ln\left(s_0\right) \tag{5}$$

where Z_i is the ith order statistic of the standard exponential variable. Using the expectation and variance of this variable, Lu et al. [28] proposed the following approximate formula for the weights:

$$w_i = \left[E\left(Z_i\right)\right]^2 / Var\left(Z_i\right) \tag{6}$$

where $E\left(Z_i\right) = \sum_{j=1}^{i}\frac{1}{n-j+1}$ and $Var\left(Z_i\right) = \sum_{j=1}^{i}\frac{1}{\left(n-j+1\right)^2}$. Note that this formula does not involve a probability index. Also, they went on to compare these four WLS methods: They pointed out that the denominator of Hung's weight factor in Eq. (4c) is constant, thus Hung's weight factor is the same as Bergman's in Eq (4a). They concluded that Bergman's weight factor (as well as that of Hung) generates a larger Mean Squared Error (MSE) than the others. Faucher and Tyson's weight factor in Eq. (4b) and their method show similar levels of performance.

As a result, we decided to use Faucher and Tyson's weight factors (Eq. 4b) in this study due to its ease of use. These weight factors were derived independently from median ranks [29] and should approximately work for any probability rank value between 0 and 1.

WLS computations are performed by various statistics packages such as MINITAB® and SPSS®. Microsoft Excel® spreadsheets can also be used for this purpose using the following closed-form formulae:

$$\hat{m} = \frac{\left(\sum w_i\right)\left(\sum w_i z_i y_i\right) - \left(\sum w_i z_i\right)\left(\sum w_i y_i\right)}{\left(\sum w_i\right)\left(w_i z_i^2\right) - \left(\sum w_i z_i\right)^2} \tag{7a}$$

and

$$\hat{s}_0 = \exp\left(-\frac{\sum w_i y_i - \hat{m}\sum w_i z_i}{\hat{m}\sum w_i}\right) \tag{7b}$$

where

$$z_i = \ln s_i \tag{7c}$$

$$y_i = \ln\left[-\ln\left(1-F_i\right)\right] \tag{7d}$$

and w_i are computed by Eq. (4b). As a side note, w_i are set to 1 for OLS computations.

In both OLS and WLS $\hat{m}/m$ is a pivotal statistic. In Monte Carlo simulations, therefore, choosing m=1 will suffice for all practical purposes. Similarly, the scale parameter will be fixed as $s_0 = 1$.

A New Probability Index Formula for Unbiased WLS

This section and following sections develop a new probability index producing an unbiased estimator of m, that is, $E(\hat{m}) = m$. Numerous probability indices have been proposed for F in addition to Eq. (3a)-(3c). Davies[16] has recently provided a comprehensive list of them. These probability indices are all developed for the ordinary least squares method and conform to the general form of

$$F_i = \left(i-a\right)/\left(n+b\right) \tag{8}$$

where $0 \le a \le 1$ and $b > -a$ (because $1 - F_i$ in Eq. (2) has to be between 0 and 1 for all $i = 1, ..., n$ to avoid negative terms in the logarithms). For example, $a = 0.3$ and $b = 0.4$ for median rank in Eq. (3a). The coefficients a and b are usually constant [16]; however, some studies aiming at minimizing the coefficient of variation of $\hat{m}$ [35,36] or making $\hat{m}$ unbiased [16,37,38] proposed a and b values as functions of n. One problem with this approach is that the values change inconsistently with changing n which leads to the lack of expressions for F which are valid for all materials engineering relevant values of n [16]. Therefore, this study aims to compute a as functions of n, and b as a constant such that WLS produces $\hat{m}$ values that are unbiased and have minimum variation.

The basic simulation procedure employed in this study involves generating a sample of n values from a Weibull distribution with parameters m=1 and $s_0 = 1$; estimating $\hat{m}$ using WLS with certain a and b values in Eq. (8) (or any index function with two constants such as in Eq. (10) as will be discussed below) and repeating this R times to compute the sample mean and the standard deviation of $\hat{m}$.

In this procedure the sample mean of $\hat{m}$, $\hat{m}_{ave}$, can be considered as a stochastic function $g(a, b, n, R)$. For a given n, finding the values of a and b producing an unbiased estimator ($\hat{m}_{ave} = 1$) is equivalent to finding the root of $g(a, b, n, R)$ -1. *The root-finding procedure* employed to compute the a-b combination in Eq. (8) is (or in Eq. (10) as will be discussed shortly) as follows: While keeping b at a certain value b', a is increased incrementally by δ, and at any step going from a' to $a' + \delta$, if $\hat{m}_{ave}$, estimated by the basic simulation procedure, goes from a value less than 1 to a value greater than 1, or the vice versa, this indicates that there exists an unbiased estimator for $b = b'$ and the values of $a' + \delta / 2$ and b' can be considered as the approximate values producing an unbiased estimator. The precision of the procedure depends on the choice of δ. This is repeated for all b values in a given search interval in increments of δ: ....$b' - \delta$, b', $b' + \delta$, ...

As an initial attempt to produce approximate a and b values in Eq. (8) a small group sample sizes n = 6, 15, 30, 75, 100 and 150, are selected. In Eq. (8) b is usually between 0 and 1 [16], hence $0 \le a \le 1$ and $0 \le b \le 1$. The root-finding procedure was repeated for each sample size with $\delta = 0.01$ and the simulation run number is $R = 10^5$. However, in order to avoid numerical errors, the values of 0 and 1 for a were avoided throughout all the simulations in this study; 0 is always replaced by 0.001 and 1 by 0.999 (a = 0.001, 0.01, 0.02, ... 0.99, 0.999 and b = 0, 0.01, 0.02, ..., 0.99, 1). This simulation, unfortunately, was not able to produce a constant b value producing an unbiased estimator for all n values. Further analysis with b values larger than 1 also showed that as the b value increased the amount of bias became larger.

Therefore, the probability index, used for OLS, cannot be used for WLS. This necessitates another index form suitable for WLS. Another such form used in probability plots, in general, has the following form [28,39]:

$$F_i = (i-a)/(n-2a+1) \tag{9}$$

When this formula is tried for the same group of samples sizes, the root finding procedure was able to find an unbiased estimator for each sample size with $\delta = 0.01$ and $R = 10^5$ (consider $b = 1$). Next, we decided to try a more general formula with two constants as in Eq. (7):

$$F_i = (i-a)/(n-2a+b) \tag{10}$$

This allows choosing among a range of b values producing an unbiased estimator with the smallest variance. Following the discussion on bounds for Eq. (8), $0 \le a \le 1$ and $a < b$.

Search for a Value Of *b* with the Smallest Variance

The root-finding procedure is repeated for Eq. (10) with the same group of sample sizes, changing b from 1 to 3, a from 0 to 1 (strictly speaking from 0.001 to 0.999) with $\delta = 0.01$ and $R = 10^5$. This procedure produced an unbiased estimator except for $b > 2.6$. Among all b values between 1 and 2.6, the standard deviations, multiplied by $\sqrt{n}$ for normalization, are averaged over all sample sizes. The smallest average value is produced by $b = 1.57$. Figure 1 shows the normalized standard deviations for selected b values; note that the minimums are achieved at different values of a for each n, which are not presented here. It indicates that change in b may cause significant differences in the standard deviation for $n \le 30$; also, the value $b = 1.5$, which is close to value of $b = 1.57$, consistently produces small standard deviation values for each sample size. This can be seen in Figure 1 for n = 6 and 15. As a result of this analysis b = 1.57 was fixed as a constant in Eq. (10).

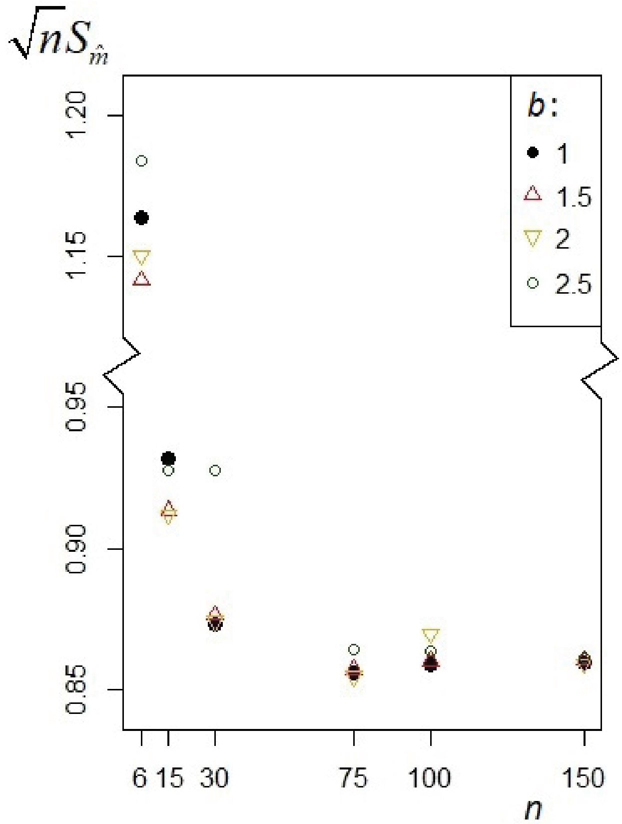

Figure 1. Samples size vs. normalized standard deviations of $\hat{m}$ for selected b values.

Computing Values of *a* as a Function of *n* When *b* = 1.57

The next task is to compute values of a as a function of n. Our initial trials with the newly computed index function

$$F_i = (i-a)/(n-2a+1.57) \tag{11}$$

revealed the stochastic function $\hat{m}_{ave} = g(a, b, n, R)$ is a considerably flat function with respect to a. When $\delta = 0.01$ and $R = 10^5$ were used, this caused the function value to fluctuate above and below 1 for neighboring values of a, making it difficult to specify a precise a value producing an unbiased estimator. Increasing δ and R simultaneously and running the simulations for all n are expected to result in total simulation times that will be prohibitively long.

A two-step solution is proposed for this problem: In the first step, $R = 10^5$ was increased to $R = 10^6$ while keeping $\delta = 0.01$ for a larger set of selected sample sizes n = 6, 8, 10, 12, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140 and 150. Then, using the root-finding procedure, approximate values of a were computed. Using them, values

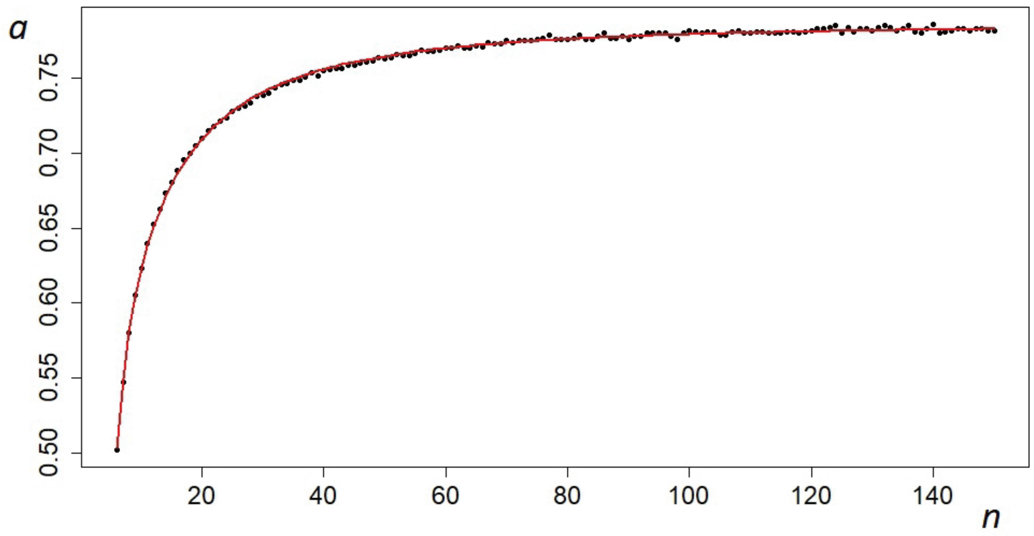

Figure 2. Sample size vs. *a* values, provided in the appendix, and the fitted curve in Eq. (12).

for all *n* were approximated: for example, if $a = 0.63$ for $n = 10$ and $a = 0.65$ for $n = 12$, then as $a \approx 0.64$ for $n = 11$.

In the second step, $\delta = 0.01$ was reduced to $\delta = 0.001$ along with $R = 10^6$. The root finding procedure was applied to an interval of length 0.024 centering the previously computed values of *a* for all $n = 6,\ldots, 150$. The length 0.024 was determined by some trial-and-error; it was to be kept to a minimum because newly selected values of δ and *R* considerably increased the simulation time. This particular interval length managed to produce values of *a* for all *n*; values of a are provided in the appendix.

In order to fit the values to the sample size, the following function is proposed to be used in Eq. (11):

$$a = -7.19247 + 7.97925\exp(-(n^{-1.5})) + 4.52723\exp(-n^{0.6}) \quad (12)$$

The fit is considerably good with an R^2 value of 0.9989, residual standard error of 0.0015 and a maximum percent error of 0.45. The values and fitted function are shown in Figure 2, which visually shows a very good fit.

Consequently, this study proposes using the WLS methods with the probability index in Eq. (11) together with coefficients in Eq. (12) using the weights in Eq. (4b).

COMPARISON WITH OTHER METHODS

It is tricky to select methods for comparison purposes, because there is no WLS method developed for unbiased estimation of the Weibull modulus *m*. We selected Davies' [16] probability index formula as the first method of comparison; it is designed for unbiased estimation of the mean of *m* (he proposed others for median and mode) to be used with OLS. While OLS can be seen as a special case of WLS, where all the weight are one, OLS and WLS are treated as two separate methods in this study. Davies [16] showed that the formula results in unbiased estimates with low variation. As the second and third methods of comparison, the MLE method and Faucher and Tyson's WLS method [29] are selected. While their estimates are biased, they may still produce estimates close to the true value, which can be measured by the Mean Squared Error (*MSE*): $MSE = E(\hat{m} - m)^2$. *MSE* is also related with bias: $MSE = Var(\hat{m}) + (E(\hat{m}) - m)^2$, where the second part of the summation is the square of the bias of $\hat{m}$. Note that this part is zero for the proposed method and Davies' method because of their unbiasedness.

The basic simulation procedure described previously is used with the selected methods for comparison and selected sample sizes; that is, parameters are set as m=1 and $s_0 = 1$ for Weibull random number generation; the simulation run number is set as R=10^6. The *MSE* values as well as the sample means of $\hat{m}$ are shown in Table 1.

First, let's look at the precision of the unbiasedness: the first column of the proposed method in Table 1 should be ideally 1, but there are slight deviations which are due to the simulation error and the errors incurred during the root finding procedure. Davies' method is designed for $n \geq 10$, this explains the relatively high deviations for $n = 6$ and 8.

Before comparing the methods let's note that for two estimators, $\hat{m}_1$ and $\hat{m}_2$, the ratio of their MSEs are called relative efficiency, defined as $RE[\hat{m}_1, \hat{m}_2] = MSE(\hat{m}_1)/MSE(\hat{m}_2)$. A *RE* value of 0.80 implies that the necessary sample size for the method using $\hat{m}_1$ is 80% of that needed for $\hat{m}_2$ to achieve approximately equal MSE value [15]. This may translate as the approximate percent saving in testing costs when the method of $\hat{m}_1$ is preferred to that of $\hat{m}_2$.

When the *MSE*'s of the two unbiased methods are compared, Table 1 shows that Davies' method has significantly larger values than the proposed method for all sample sizes. It is to the degree that the levels achieved by the proposed method can be achieved with an additional five test specimens (e.g. $n = 15$ vs $n = 20$) for $n \geq 20$, and with an additional 10 specimen for $n \geq 35$; this result can be seen by comparing *MSE* values for different n values, and also agrees with the *RE* column (note that *RE* produces an approximate result). Note that, since these methods are unbiased, *MSE* of $\hat{m}$ is equivalent to the sample variance of $\hat{m}$, $S^2_{\hat{m}}$(strictly speaking, they are approximately equal due to small deviations from unbiasedness).

While the MLE method has been a popular alternative to OLS and WLS, it suffers from severe overestimation bias in small samples. For instance, in Table 1, the MLE estimate for $n = 10$ (1.1698) overestimates the unknown Weibull modulus by 16.98%. Table 2 decomposes *MSE* values in Table 1 into bias and variance components, revealing that the bias becomes negligible (below 2%) for sample sizes exceeding $n = 70$; note that $MSE = (Bias)^2 + Variance$ (here, for $n = 10$, $0.1517 = 0.1698^2 + 0.1229$). Table 2 highlights the substantial bias in MLE estimates for small samples significantly impacts the overall *MSE*.

The *MSE* values of the MLE method are at least 10% larger than those of the proposed method for $n \leq 30$. The difference is so severe that the levels achieved by the proposed method for $n \leq 12$ can be achieved by the MLE method with an addition of at least three test specimens. However, as the sample size increases, the MLE methods starts to show better performance as a result of its optimality properties in large samples. For $n \geq 60$, the MLE method starts to perform better, and the *MSE* for the proposed method becomes 14% larger than that of the MLE method at $n = 150$.

When compared with the proposed method, Faucher and Tyson's method has smaller *MSE* values for all sample sizes, and the bias is not severe as can be seen in Table 2. The proposed method's *MSE* is about 6% above that of Faucher and Tyson's method for $n \leq 15$, but as the sample size increases, this percentage goes below 2% after $n \geq 35$ and below 1% after $n \geq 60$. These results do not indicate that Faucher and Tyson's method is better than the proposed method, because it has significant bias. It underestimates $\hat{m}$ and approximately 5% underestimation should be expected for n = 6-20. The percentage of underestimation

Table 1. Sample mean and MSE values of $\hat{m}$ for the compared methods

	Proposed (P)		**Davies (D)**			**MLE (M)**			**Fau&Tys (F)**		
n	**Mean**	**MSE**	**Mean**	**MSE**	**RE(P,D)**	**Mean**	**MSE**	**RE(P,M)**	**Mean**	**MSE**	**RE(P,F)**
6	1.0008	0.2130	1.0368	0.2420	0.8802	1.3341	0.4848	0.4394	0.9752	0.2031	1.0487
8	1.0000	0.1289	1.0111	0.1453	0.8871	1.2249	0.2399	0.5373	0.9511	0.1193	1.0805
10	0.9996	0.0924	1.0030	0.1067	0.8660	1.1698	0.1517	0.6091	0.9459	0.0859	1.0757
12	0.9988	0.0719	0.9997	0.0854	0.8419	1.1358	0.1081	0.6651	0.9454	0.0676	1.0636
15	0.9998	0.0545	0.9998	0.0670	0.8134	1.1055	0.0749	0.7276	0.9495	0.0519	1.0501
20	1.0000	0.0389	0.9998	0.0498	0.7811	1.0762	0.0482	0.8071	0.9558	0.0376	1.0346
25	1.0003	0.0303	1.0005	0.0399	0.7594	1.0598	0.0352	0.8608	0.9617	0.0295	1.0271
30	1.0001	0.0249	1.0004	0.0334	0.7455	1.0489	0.0276	0.9022	0.9661	0.0244	1.0205
35	1.0004	0.0212	1.0006	0.0288	0.7361	1.0418	0.0227	0.9339	0.9701	0.0208	1.0192
40	1.0002	0.0184	1.0002	0.0254	0.7244	1.0362	0.0191	0.9634	0.9730	0.0182	1.0110
45	1.0000	0.0163	0.9999	0.0226	0.7212	1.0318	0.0166	0.9819	0.9753	0.0161	1.0124
50	1.0001	0.0147	0.9998	0.0205	0.7171	1.0286	0.0146	1.0068	0.9776	0.0145	1.0138
55	1.0001	0.0133	0.9999	0.0187	0.7112	1.0261	0.0131	1.0153	0.9795	0.0132	1.0076
60	0.9999	0.0122	0.9996	0.0172	0.7093	1.0236	0.0118	1.0339	0.9808	0.0121	1.0083
65	0.9999	0.0112	0.9999	0.0159	0.7044	1.0217	0.0108	1.0370	0.9821	0.0112	1.0000
70	0.9999	0.0105	0.9999	0.0148	0.7095	1.0202	0.0099	1.0606	0.9833	0.0104	1.0096
75	1.0000	0.0098	1.0001	0.0139	0.7050	1.0189	0.0092	1.0652	0.9844	0.0097	1.0103
80	0.9998	0.0092	1.0000	0.0131	0.7023	1.0176	0.0085	1.0824	0.9852	0.0091	1.0110
85	1.0000	0.0086	1.0001	0.0123	0.6992	1.0166	0.0080	1.0750	0.9862	0.0086	1.0000
90	0.9999	0.0082	1.0001	0.0117	0.7009	1.0156	0.0075	1.0933	0.9868	0.0081	1.0123
95	1.0000	0.0077	1.0001	0.0111	0.6937	1.0148	0.0070	1.1000	0.9876	0.0077	1.0000
100	0.9999	0.0073	1.0001	0.0105	0.6952	1.0140	0.0067	1.0896	0.9881	0.0073	1.0000
125	1.0000	0.0059	0.9999	0.0085	0.6941	1.0113	0.0052	1.1346	0.9905	0.0059	1.0000
150	0.9999	0.0049	0.9996	0.0071	0.6901	1.0093	0.0043	1.1395	0.9920	0.0049	1.0000

becomes less severe as the sample size increases and goes below 2% for $n \geq 60$.

Practical Example

An application of the proposed method was carried out on the results of an experimental study [40]: 19 identical composite specimens were prepared from quasi-isotropic carbon-epoxy sheets with (0°)3 configuration, 0.89 mm thickness, and 295 gr/m2 weight and the tension experiments were carried out using an Instron 8516+ universal testing machine. The measured fracture strength values are presented in Table 3.

To enhance the application study in terms of sample size, in addition to the original data set given in Table 3, two random subsets of sizes 6, 10, and 15 were drawn from the same data set. These subsets consist of {522, 532.7, 476.5, 521.6, 439, 507.3} for n = 6, {513.6, 552,519, 521.6, 439,450.9, 463.5, 497.5, 476.5, 477} for n =10, and {532.7, 502.5, 442, 519, 502.7, 477, 552, 522, 439, 513.6, 521.6, 450.9, 476.5, 507.3, 463.5} for n = 15.

To implement the proposed method, Eq. (7) is used where F_i is computed by Eq. (11) and Eq. (12). The following Table 4 gives the estimated Weibull modulus values. As n increases, the *MSE* decreases as can be seen from Table 1,

Table 2. Decomposition of *MSE* values of $\hat{m}$ for biased estimation methods

	MLE (M)				Fau&Tys (F)			
n	**Mean**	**MSE**	**Bias**	**Variance**	**Mean**	**MSE**	**Bias**	**Variance**
6	1.3341	0.4848	33.41%	0.3732	0.9752	0.2031	-2.48%	0.2025
8	1.2249	0.2399	22.49%	0.1893	0.9511	0.1193	-4.89%	0.1169
10	1.1698	0.1517	16.98%	0.1229	0.9459	0.0859	-5.41%	0.0830
12	1.1358	0.1081	13.58%	0.0897	0.9454	0.0676	-5.46%	0.0646
15	1.1055	0.0749	10.55%	0.0638	0.9495	0.0519	-5.05%	0.0493
20	1.0762	0.0482	7.62%	0.0424	0.9558	0.0376	-4.42%	0.0356
25	1.0598	0.0352	5.98%	0.0316	0.9617	0.0295	-3.83%	0.0280
30	1.0489	0.0276	4.89%	0.0252	0.9661	0.0244	-3.39%	0.0233
35	1.0418	0.0227	4.18%	0.0210	0.9701	0.0208	-2.99%	0.0199
40	1.0362	0.0191	3.62%	0.0178	0.9730	0.0182	-2.70%	0.0175
45	1.0318	0.0166	3.18%	0.0156	0.9753	0.0161	-2.47%	0.0155
50	1.0286	0.0146	2.86%	0.0138	0.9776	0.0145	-2.24%	0.0140
55	1.0261	0.0131	2.61%	0.0124	0.9795	0.0132	-2.05%	0.0128
60	1.0236	0.0118	2.36%	0.0112	0.9808	0.0121	-1.92%	0.0117
65	1.0217	0.0108	2.17%	0.0103	0.9821	0.0112	-1.79%	0.0109
70	1.0202	0.0099	2.02%	0.0095	0.9833	0.0104	-1.67%	0.0101

Table 3. Fracture strength of carbon-epoxy composite material specimens (megapascals)

Test No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Fracture strength [MPa]	532.7	502.5	442	473	519	502.7	477	510	522	552
Test No.	11	12	13	14	15	16	17	18	19	
Fracture strength [MPa]	522	439	513.6	497.5	521.6	450.9	476.5	507.3	463.5	

Table 4. Weibull modulus estimates ($\hat{m}$) for the fracture strength data

n	**Proposed (P)**	**Davies (D)**	**MLE (M)**	**Fau&Tys (F)**
6	16.0764	14.5886	21.9916	15.3496
10	13.8474	15.2852	15.6205	13.2601
15	15.6385	16.2970	17.3370	14.9150
19	17.6756	18.1770	18.8625	16.9160

hence estimates with larger sample sizes have higher precision. Therefore, the estimates for $n = 19$ should be closer to the true, unknown Weibull modulus value.

The estimates for $n = 19$ are close to each other due to their small *MSE* values (check the values for $n = 20$ in Table 1). However, the smallest *MSE*'s belong to the proposed WLS method and Faucher and Tyson's WLS method, thus their estimates, 17.6756 and 16.916 can be argued to be closer to the true m value.

For smaller n values the estimates are farther away from each other, due to increased *MSE* and increased bias. Consider $n = 6$: *MSE* is a measure of expected distance of an estimate from the true value, and they are 5-10 times as large as those for $n = 20$. Also, due to the substantial overestimation problem in the MLE method, which can be seen in Table 2, MLE estimates are larger than the others in Table 4.

CONCLUSION

This research introduced a novel probability index that, for the first time, offers unbiased WLS estimators of the Weibull modulus with a satisfactory degree of precision. To this end, a systematic study was carried out to compute parameters of a probability index to be used with the WLS method for the unbiased estimation of the Weibull modulus. First Faucher and Tyson's weights were selected among several weighting schemes. Then two parameters of a generic probability index were computed by using a step-by-step approach. The result is a simple index formula presented in Eq. (10) wherein the single parameter a, produced from simulations, is fitted as a function of the sample size in Eq. (11). Use of two parameters in this study allowed variance reduction of the estimator as well as satisfying unbiasedness.

The proposed method requires only the use of the new probability index along with Faucher and Tyson's weights. It is the only WLS method with unbiased Weibull modulus estimator in the literature. When compared with a recent OLS method proposing a new probability index producing unbiased Weibull modulus estimators, it exhibited significantly smaller variance. *MSE* is used for comparison among methods producing biased estimators. Results showed that the proposed method performs better than the MLE method for sample sizes less than 60, and its performance is close to that of Faucher and Tyson's method, but with slightly larger *MSE* values. The advantage of the proposed method is most prominent in the range of sample sizes smaller than 125 where the percent bias in the aforementioned biased methods is greater than 1% and they should be used with caution.

In materials science, high experimental costs limit the number of samples to be tested; and in general, smaller the sample size are preferred. In general, sample sizes are less than 60 in such experiments, which makes the proposed method advantageous in terms of its smaller MSE and unbiasedness property. While this paper focuses on materials science applications, it should be pointed out that the proposed method can be used in other application areas such as wind speed analysis and reliability analysis.

In conclusion, the implementation of the proposed WLS method is almost as simple as that of the OLS method and we hope that this helps materials scientists and engineers who generally prefer OLS and MLE in practice.

DATA AVAILABILITY STATEMENT

The authors confirm that the data that supports the findings of this study are available within the article. Raw data that support the finding of this study are available from the corresponding author, upon reasonable request.

CONFLICT OF INTEREST

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ETHICS

There are no ethical issues with the publication of this manuscript.

STATEMENT ON THE USE OF ARTIFICIAL INTELLIGENCE

Artificial intelligence was not used in the preparation of the article.

REFERENCES

[1] Weibull W. A statistical distribution function of wide applicability. J Appl Mech 1951:1-5.
[2] Kaplan YA. Comparison of the performance of the methods used to find the weibull parameters at different heights. Arab J Sci Eng 2021;46:12145–53. [CrossRef]
[3] Usta I, Arik I, Yenilmez I, Kantar YM. A new estimation approach based on moments for estimating Weibull parameters in wind power applications. Energy Convers Manag 2018;164:570–578. [CrossRef]
[4] Kantar YM, Usta I, Yenilmez I, Arık I. A study on estimation of wind speed distribution by using the Modified Weibull distribution. Bilişim Teknol Derg 2016;9:63.
[5] Kantar YM, Arık I, Usta I, Yenilmez I. Comparison of some estimation methods of the two parameter Weibull distribution for unusual wind speed data cases. Bilişim Teknol Derg 2016;9:81.
[6] Lai C-D, Murthy D, Xie M. Weibull distributions and their applications. Springer Handb., Springer; 2006, p. 63–78. [CrossRef]
[7] Barbero E, Fernández-Sáez J, Navarro C. Statistical analysis of the mechanical properties of composite materials. Compos Part B Eng 2000;31:375–381. [CrossRef]

[8] Lei W-S, Yu Z, Zhang P, Qian G. Standardized Weibull statistics of ceramic strength. Ceram Int 2021;47:4972–4993. [CrossRef]

[9] Lei W-S, Zhang P, Yu Z, Qian G. Statistics of ceramic strength: Use ordinary Weibull distribution function or Weibull statistical fracture theory? Ceram Int 2020;46:20751–20768. [CrossRef]

[10] Bernal RA. On the application of Weibull statistics for describing strength of micro and nanostructures. Mech Mater 2021;162:104057. [CrossRef]

[11] Barbero E, Fernandez-Saez J, Navarro C. Statistical distribution of the estimator of Weibull modulus. J Mater Sci Lett 2001;20:847–849. [CrossRef]

[12] Yang X, Xie L, Yang Y, Zhao B, Li Y. A comparative study for parameter estimation of the Weibull distribution in a small sample size: An application to spring fatigue failure data. Qual Eng 2022:1–13. [CrossRef]

[13] Liu J, Zhang Z, Zhu L, Wang J, He Y. Research on fatigue life evaluation method of shaft parts based on small sample. Int J Struct Integr 2023;14:709–732. [CrossRef]

[14] Yalçınkaya M, Birgören B. Shortest confidence intervals for weibull modulus for small samples in materials reliability analysis. Gazi Univ J Sci 2023:1. [CrossRef]

[15] Genschel U, Meeker WQ. A comparison of maximum likelihood and median-rank regression for Weibull estimation. Qual Eng 2010;22:236–255. [CrossRef]

[16] Davies IJ. Unbiased estimation of Weibull modulus using linear least squares analysis—A systematic approach. J Eur Ceram Soc 2017;37:369–380. [CrossRef]

[17] Composite Materials Handbook. Vol. 1. Polymer matrix composites guidelines for characterization of structural materials, SAE International on behalf of CMH-17, a division of Wichita State University. SAE International; 2012.

[18] Hirose H. Bias correction for the maximum likelihood estimates in the two-parameter Weibull distribution. IEEE Trans Dielectr Electr Insul 1999;6:66–68. [CrossRef]

[19] Davies I. Empirical correction factor for the best estimate of Weibull modulus obtained using linear least squares analysis. J Mater Sci Lett 2001;20:997–999. [CrossRef]

[20] Davies I. Best estimate of Weibull modulus obtained using linear least squares analysis: an improved empirical correction factor. J Mater Sci 2004;39:1441–1444. [CrossRef]

[21] Jia X, Xi G, Nadarajah S. Correction Factor for Unbiased Estimation of Weibull Modulus by the Linear Least Squares Method. Metall Mater Trans A 2019;50:2991–3001. [CrossRef]

[22] Davies IJ. Unbiased estimation of the Weibull scale parameter using linear least squares analysis. J Eur Ceram Soc 2017;37:2973–2981. [CrossRef]

[23] Garrido A, Caro-Carretero R, Jimenez-Octavio JR, Carnicero A, Such M. A new approach to fitting the three-parameter Weibull distribution: An application to glass ceramics. Commun Stat-Theory Methods 2021;50:3403–3420. [CrossRef]

[24] Deng B, Jiang D, Gong J. Is a three-parameter Weibull function really necessary for the characterization of the statistical variation of the strength of brittle ceramics? J Eur Ceram Soc 2018;38:2234–2242. [CrossRef]

[25] Yang F, Ren H, Hu Z. Maximum likelihood estimation for three-parameter Weibull distribution using evolutionary strategy. Math Probl Eng 2019;2019:6281781. [CrossRef]

[26] Nohut S. Three-parameter (3P) weibull distribution for characterization of strength of ceramics showing R-Curve behavior. Ceram Int 2021;47:2270–2279. [CrossRef]

[27] Guo J, Kong X, Wu N, Xie L. Weibull parameter estimation and reliability analysis with small samples based on successive approximation method. J Mech Sci Technol 2023;37:5797–5811. [CrossRef]

[28] Lu H, Chen C, Wu J. A note on weighted least-squares estimation of the shape parameter of the Weibull distribution. Qual Reliab Eng Int 2004;20:579–586. [CrossRef]

[29] Faucher B, Tyson W. On the determination of Weibull parameters. J Mater Sci Lett 1988;7:1199–2203. [CrossRef]

[30] Bergman B. Estimation of Weibull parameters using a weight function. J Mater Sci Lett 1986;5:611–614. [CrossRef]

[31] Hung W. Weighted least-squares estimation of the shape parameter of the Weibull distribution. Qual Reliab Eng Int 2001;17:467–469. [CrossRef]

[32] Mert Kantar Y. Estimating variances in weighted least-squares estimation of distributional parameters. Math Comput Appl 2016;21:7. [CrossRef]

[33] Kantar YM. Generalized least squares and weighted least squares estimation methods for distributional parameters. REVSTAT-Stat J 2015;13:263–282.

[34] Dutang C, Wuertz D. A note on random number generation 2009. https://cran.r-project.org/web/packages/randtoolbox/vignettes/fullpres.pdf.

[35] Song L, Wu D, Li Y. Optimal probability estimators for determining Weibull parameters. J Mater Sci Lett 2003;22:1651–1653. [CrossRef]

[36] Wu D, Lu G, Jiang H, Li Y. Improved estimation of Weibull parameters with the linear regression method. J Am Ceram Soc 2004;87:1799–1802. [CrossRef]

[37] Tiryakioğlu M. An unbiased probability estimator to determine Weibull modulus by the linear regression method. J Mater Sci 2006;41:5011–5013. [CrossRef]

[38] Tiryakioğlu M, Hudak D. On estimating Weibull modulus by the linear regression method. J Mater Sci 2007;42:10173–10179. [CrossRef]

[39] D'Agostino R. Goodness-of-fit-techniques. Routledge; 2017. [CrossRef]

[40] Dirikolu MH, Aktas A, Birgoren B. Statistical analysis of fracture strength of composite materials using Weibull distribution. Turk J Eng Environ Sci 2002;26:45–48.

APPENDIX

Table A. values of *a* to be used in Eq. (11) for WLS

n	*a*	*n*	*a*	*n*	*a*	*n*	*a*	*n*	*a*
6	0.5025	35	0.7485	64	0.7695	93	0.7795	122	0.7825
7	0.5475	36	0.7485	65	0.7715	94	0.7795	123	0.7835
8	0.5805	37	0.7505	66	0.7705	95	0.7795	124	0.7845
9	0.6055	38	0.7535	67	0.7735	96	0.7795	125	0.7795
10	0.6235	39	0.7515	68	0.7725	97	0.7775	126	0.7835
11	0.6395	40	0.7545	69	0.7725	98	0.7755	127	0.7805
12	0.6525	41	0.7555	70	0.7745	99	0.7795	128	0.7825
13	0.6625	42	0.7565	71	0.7735	100	0.7815	129	0.7825
14	0.6735	43	0.7565	72	0.7745	101	0.7805	130	0.7815
15	0.6805	44	0.7585	73	0.7745	102	0.7795	131	0.7835
16	0.6885	45	0.7585	74	0.7745	103	0.7805	132	0.7845
17	0.6955	46	0.7595	75	0.7755	104	0.7805	133	0.7835
18	0.6995	47	0.7605	76	0.7765	105	0.7785	134	0.7815
19	0.7045	48	0.7615	77	0.7785	106	0.7785	135	0.7825
20	0.7095	49	0.7635	78	0.7755	107	0.7805	136	0.7845
21	0.7145	50	0.7625	79	0.7755	108	0.7815	137	0.7805
22	0.7175	51	0.7635	80	0.7755	109	0.7795	138	0.7795
23	0.7215	52	0.7655	81	0.7765	110	0.7795	139	0.7825
24	0.7235	53	0.7645	82	0.7785	111	0.7805	140	0.7855
25	0.7275	54	0.7645	83	0.7755	112	0.7805	141	0.7795
26	0.7295	55	0.7665	84	0.7755	113	0.7795	142	0.7805
27	0.7315	56	0.7685	85	0.7775	114	0.7795	143	0.7815
28	0.7335	57	0.7675	86	0.7795	115	0.7795	144	0.7825
29	0.7375	58	0.7675	87	0.7765	116	0.7805	145	0.7825
30	0.7385	59	0.7685	88	0.7765	117	0.7805	146	0.7815
31	0.7395	60	0.7695	89	0.7775	118	0.7795	147	0.7825
32	0.7435	61	0.7695	90	0.7755	119	0.7805	148	0.7825
33	0.7455	62	0.7715	91	0.7775	120	0.7815	149	0.7815
34	0.7465	63	0.7695	92	0.7775	121	0.7825	150	0.7815