scientific reports

OPEN

A modification in Weibull parameters to achieve a more accurate probability distribution function in fatigue applications

H. Fakoor & J. Alizadeh Kaklar✉

Risk evaluation for fatigue failure of the engineering components is an important aspect of the engineering design. Weibull distributions are often used in preference to the log-normal distribution to analyze probability aspects of fatigue results. This study presents a probabilistic model for calculating Weibull distribution parameters to reduce the effect of percentage discretization error of experimental fatigue life and R–S–N curves for three reliability levels. By considering any result of standard fatigue test as an equivalent Weibull distribution, artificial data are generated and the accuracy of common Weibull distribution model can be improved. The results show error reduction in the Kolmogorov–Smirnov test and R-square values. Also, the Basquin model is used for different reliability levels with the same error order for risk evaluation of fatigue failure. The coefficient of variation for fatigue life increases at higher stress levels and has a linear relation with stress level for a high-cycle fatigue regime.

Fatigue failure is the formation and propagation of cracks due to a repetitive or cyclic load. It has been estimated that fatigue contributes to approximately 90% of all mechanical service failures[1]. For the first time, the French mathematician and engineer Jean-Victor Poncelet used the terminology of "fatigue" in his book in 1841[2]. For fatigue base designs, actual fatigue data should be used or, if not available, must be modeled and generated. Many models have been developed in the S–N (stress-life) approach of fatigue life estimation to depict S–N curves. The earlier models have estimated a median fatigue life, but it is necessary to calculate the risk of fatigue failure for a safe and economical design. The fatigue life shall be predetermined by the desired reliability levels referred to by the material properties, crack size, environment, and loading condition. In engineering designs, the failure of system parts must be considered with a low probability of occurrence[3].

In statistical analyses, probabilistic fatigue S–N curves are widely used to quantify scattered fatigue test data and analyze fatigue problems. Usually, for constructing the curves, fatigue tests are performed on some stress levels. Then, a distribution model (e.g., Weibull) is applied to predict fatigue life at a specified applied stress level, followed by characterizing the S–N relation. The Basquin equation is an S–N relation with a linear relationship between applied stress and fatigue life logarithmic scale. This equation can be used to express the stress-life relation in a certain survival probability confidence level. This probabilistic S–N (R-S–N) curve is an S–N relation in some probability confidence levels[4].

This paper focuses on a modification for a common probability distribution model of a physical phenomenon when the available number of experimental results is limited. The importance of this idea is that generally, engineering design costs account on average for 5% of total costs in the projects[5] and most of the cost is related to conducting experimental tests. Weibull distribution is often used in preference to analyze probability aspects of fatigue results[6] and to reduce required number of experimental tests for evaluation of fatigue life distribution. The novelty of this paper is making some artificial data by considering any result of the fatigue test as an equivalent Weibull distribution with the mean value of the same test fatigue life and cumulative probability of common Weibull method. The accuracy of common Weibull distribution model can be improved using this method.

Probability distributions

Probability theory is one of the most important aspects of statistics. A probability distribution is a mathematical function that gives the probabilities of occurrence of different possible outcomes for an experiment[7]. Table 1 presents several models that have been proposed to model probability distribution.

Urmia University, Urmia, Iran. ✉email: j.alizadeh@urmia.ac.ir

1

Model	Equation
Normal distribution or Gaussian distribution	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$
Pareto distribution	$F(x) = \begin{cases} 1 - \left(\frac{x_m}{x}\right)^\alpha, & x \geq x_m \\ 0, & x < x_m \end{cases}$
Continuous uniform distribution	$F(x) = \begin{cases} 0, & x - \mu < -\sigma\sqrt{3} \\ \frac{1}{2}\left(\frac{x-\mu}{\sigma\sqrt{3}} + 1\right), & -\sigma\sqrt{3} \leq x - \mu < \sigma\sqrt{3} \\ 1, & x - \mu \geq \sigma\sqrt{3} \end{cases}$
Weibull distribution	$F(N_f) = 1 - e^{-\left(\frac{N_f}{\theta}\right)^b}$
Gumbel distribution	$F(x; \mu, \beta) = e^{-e^{-(x-\mu)/\beta}}$
Tukey lambda distribution	$Q(p; \lambda) = \begin{cases} \frac{1}{\lambda}\left[p^\lambda - (1-p)^\lambda\right], & \lambda \neq 0 \\ \log\left(\frac{p}{1-p}\right), & \lambda = 0 \end{cases}$
Exponential distribution	$f(x, \lambda) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$

Table 1. Common models representing a probability distribution.

The probability distributions are described with statistical parameters like mean and standard deviation, as shown in Table 1. Normal distribution, named Gaussian distribution, is the most frequently used distribution function in statistical analysis. The normal distribution, which has a bell-shaped curve, has been used for independent, random variables in the survey reports, Technical Stock Market, and scientific study of many observable phenomena in nature like human height or IQ distribution. Sinclair and Dolan[8] conducted a comprehensive statistical fatigue investigation engaging 174 nominal identical, extremely polished, smooth 7075-T6 aluminum alloy specimens. They worked on 6 alternating applied stress levels at the fully reversed test conditions. A normal distribution in the logarithmic scale for the experimental results at every applied stress amplitude gives the impression of being reasonable. Derived from that and the other performed statistical fatigue test results, a logarithmic scale normal distribution of failure is usually considered in fatigue analysis. At every applied stress range, a group of S–N diagrams at different percentages of failure probabilities are established from the probability distribution functions.

The Pareto distribution is one of the power-law probability distributions in probability theory and mathematics. Pareto distribution is usually used in the explanation of data scatter in the many types of scientific studies. Particularly applied to describing the distribution of wealth in a society, fitting the trend that a large portion of wealth is held by a small fraction of the population[9].

The Gumbel distribution usually has been applied to figure the extremum (Max or Min) distribution of a number of samples of different statistical societies. Forecasting a temblor, torrent, or other types of natural disaster is of great use and value. For instance, if the database of a parameter of a phenomenon like water in the river for 10 years ago is available, Gumbel distribution can be applied to predetermine the distribution of the variation of extremum of water level at a river in the specific time. For the Gumbel distribution, the special importance is to predict the extremum distribution according to extreme value theory[10].

John Tukey presented a continuous symmetric probability distribution model in which the Tukey lambda distribution function was specified in terms of its quantile function. The Tukey lambda distribution function is used to recognize a suitable distribution. Therefore, the Tukey lambda distribution usually has no direct application in statistical models. The Tukey lambda distribution has a single shape parameter that can be rearranged and defined in terms of the standard distribution.

The exponential distribution is a continuous distribution for modeling events that occur at a constant time rate. Two important applications of the exponential distribution are the modeling of radioactive decay in physics and the modeling of the posterior default probability for a set of financial assets in finance[11]. The exponential distribution can be applied to analyze the relationship between the unobservable actual values and measurement values[12]. In some cases the lifetime of a manufacturing item may fallows a mixed distribution models such as the half-normal distribution and the half-exponential distribution[13].

Among the probability distributions used to analyze fatigue problems[14,15], the Weibull distribution is one of the most common models in the logarithmical scale. Weibull progressed a new approach and used it to study fatigue actual data[16,17]. Weibull distributions have 2-parameter and 3-parameter models. The 2-parameter model is widely extended in fatigue problems and design. In this approach, the expected fatigue life range starts from zero cycles. Indeed, the 2-parameter Weibull distribution is a simplified 3-parameter Weibull distribution with a minimum expected life of zero. At the same time, the 3-parameter distribution is characterized by a finite minimum life greater than zero. For $3.3 \leq b \leq 3.5$, the Weibull distribution function is approximately normal or Gaussian, while it is exponential for $b = 1$. The coefficient of variation (standard deviation/mean) is approximately $C = 1/b$ for the two-parameter Weibull distribution. For b values between 3 and 6, (typical of fatigue), the error from this approximation is about 10 to 15%[6]. In modern life testing analysis to obtain information about fatigue life of a component, new method of experimental process is conducted, where products are tested under

higher stress than normal to get their failure information. For example, new methods such as adaptive type-I progressive hybrid censoring is planned to evaluate the failure parameters assuming that the failure causes are independent Weibull variables[18].

In a study of statistical fatigue analysis, Zhao and Liu[4] proposed a Weibull approach to the probabilistic study. They investigated stress-life for rolling contact fatigue. The study shows that the 2-parameter and 3-parameter Weibull equations have reasonable results. However, the 3-Parameter Weibull model has a lower standard deviation for fatigue life. This standard deviation decreases at higher applied stress levels. These results are consistent with those of classic fatigue studies.

Xionga et al.[19] investigated multiaxial fatigue results of magnesium alloy using the modified Smith–Watson–Topper (SWT) theories and the multiaxial Jiang criterion. The results of both theories were acceptable. Jiang et al.[20] used the Markov chain Monte Carlo method to estimate the parameters of a modified Weibull distribution. They suggested the use of Markov chain Monte Carlo estimation instead of maximum-likelihood estimation for point estimation when the sample size is less than 100. Canteli et al.[21] studied 3 types of fatigue models, namely, LCF, HCF, and VHCF, which are usually used in mechanical parts design. The study presented the actual results of the stress-based and strain-based approaches in a single methodology. Strzelecki[22] used 2-P and 3-P Weibull distribution and presented features of the S–N curve for fatigue limit investigation. The fatigue test results were used for rotary bending of S355J2 + C and C45 + C steels, and the S–N curves were specified. Acosta et al.[23] used measurement techniques based on temperature and magnetics to describe the fatigue behavior of metallic materials. Furthermore, they reduced the effort required to generate and provide S–N curves using valuable input parameters for short-time fatigue life calculation methods.

S–N Relation

In the S–N approach, many models have been developed for evaluating the S–N relation, with some shown in Table 2. Basquin[24] suggested a linear relation in the logarithmical scale between the applied stress (S) and the fatigue life (N). Basquin's equation is generally developed in the standards such as ASTM, ASME, and UNE and Guidelines such as FKM, DNV, and GL. Vidovic[25] performed an analytical study of the maximum-likelihood estimations for the parameters of a modified Weibull distribution model and indicated that their implementation in practice follows a rather simple pattern. Usabiaga et al.[26] implemented a model on the NCode2020 software to demonstrate the probable implementation in general commercial codes by major applications on fatigue design.

Crack propagation has always been a source of concern in determining inspection routines in different industries. Crack propagation at the higher applied stress amplitude can cause great uncertainty in the fatigue life estimation. Crack propagation has been studied widely to characterize different types of cracks, including edge, surface, subsurface, etc.[27–29]. Focus on these studies shows the uncertainty and wide data scatter in the fatigue life compared to other mechanical properties may be due to diversity in the crack initiation location and different crack types.

Stromeyer[30] published the empirical relation for the mathematical description of fatigue. Basquin, in his fatigue relation, had not considered the idea of the fatigue limit. Stromeyer[30] studied Wöhler's fatigue test data. To verify the existence of the fatigue limit concept, they conducted advanced rotating-beam fatigue tests on several materials to verify the existence of a definite fatigue limit. The Stromeyer law represents the Stress-Life curve by truncating the Basquin relation at the fatigue limit by plotting the load and fatigue life. In addition, Stromeyer presented a relation between fatigue samples temperature increase and fatigue limit. However, the knee point (N_{knee})[31] was not specified explicitly.

Palmgren[32] presented a new theory resembling the Basquin method and the Stromeyer method. The equation in Table 2 presents the results of the fatigue test of rolling bearings. The fundamental relation of this method involves the stressed volume of material in the rolling bearing raceway sub surfaces as the main parameter. "This volume of material is simplistically determined to have a nearly rectangular subsurface cross-sectional area bounded by the length of the maximum contact area ellipse and the depth at which the maximum failure-causing stress occurs".

In fatigue design, an adequate quantification of ISO 12107 inherent variation is one of the essential parameters for calculating the fatigue property in the various mechanical parts of systems and components. Also, it is essential to compare materials in fatigue properties, including their variation in engineering design. In this respect, statistical methods have been used widely to compare material properties. This International Standard includes a full methodology for the application of the Bastenaire model as well as other more sophisticated relationships.

Model	S–N curves relation
Basquin (1910)	$\log N = A - B \log(\Delta\sigma); \ \Delta\sigma \geq \Delta\sigma_\infty$
Stromeyer (1914)	$\log N = A - B \log(\Delta\sigma - \Delta\sigma_\infty)$
Palmgren (1924)	$\Delta\sigma = b(N + B)^{-a} + \Delta\sigma_\infty)$
Bastenaire (1972)	$N = \frac{A}{\Delta\sigma - E} \exp[-C(\Delta\sigma - E)] - B$
Ling and Pan (1997)	$F = \sum_{i=1}^{n} \left\{ \ln(\sigma(S_i)) + \frac{[\log N_i - \mu(S_i)]^2}{2\sigma^2(S_i)} \right\}$
Kohout and Vechet (2001)	$\log \left(\frac{\Delta\sigma}{\Delta\sigma_\infty} \right) = \log \left(\frac{N + N_1}{N + N_2} \right)^b$

Table 2. Common methods to estimate the S–N curves[24].

It also addresses the analysis of runout (censored) data[33]. Ling and Pan[34] presented a new method to determine R–S–N curves to minimize the cost and the number of samples needed for laboratory testing. The stress-life curves were considered in a 3-parameter form.

Kohout and Vechet[35] presented a different method to define S–N curves in the whole cyclic load domain in fatigue problems. This method incorporates all the fatigue-affected regions from ultimate strength of material to fatigue endurance limit, which is generally expressed by the Palmgren function. For every region, this method is similar to one of the previous theories, i.e., when the applied load is approximately large, the Kohout and Vechet model converts into the Basquin model. On the other hand, when the applied load is smaller than the fatigue endurance limit, this model converts to the Stromeyer function for almost infinite life and high-cycle fatigue region. Compared to the models specified above, the Kohout and Vechet model has some precedence. This method has a better curve fitting of fatigue test results, and this coefficient has unambiguous technical and geometrical meaning, which can be calculated with higher accuracy. In addition, this model is more appropriate for extrapolation and interpolation for fitted curves in the low-cycle and very-high-cycle regions.

The present work tries to develop the probabilistic S–N relationships for existing fatigue data in the following three steps: (1) collecting the fatigue test data; (2) estimating the probabilistic curves for every specified test condition (the key task in the present work is to determine the Weibull equation coefficient for scattered test data); and (3) evaluating the Basquin equation's coefficient from the previous step's data (to this end, a regression analysis will be done on the estimated fatigue life from the previous step).

Formulation of modified Weibull approach
Life distribution
Tolerance limits

Fatigue data are subjected to considerable scatter. In statistical analysis, a sample with a random data set is chosen. Obtaining data from the entire population is usually impossible or has very high undue costs. Due to sample size limitations, the sample statistical parameters, including mean median or variance values, are different from the source population. Designating a confidence level assigns a quantitative value of uncertainty or confidence. Lower and upper tolerance limits in a Weibull distributed model can be calculated using Eq. (1a) and (1b)[6]:

$$\text{Lower limit} = F\left(N_f\right) - k \tag{1a}$$

$$\text{Upper limit} = F\left(N_f\right) + k \tag{1b}$$

where k is a function of the sample size.

Replacement of sample statistical properties with source population properties involves some degree of uncertainty. This uncertainty is determined using the percent error. If the sample average is x_1, the percent error will be[4]:

$$\text{Percent error} = \pm \left[\frac{s_1 t_{\frac{\alpha}{2};\vartheta}}{\overline{X}_1 \sqrt{n}} \right] \tag{2}$$

where $\overline{X}_1$ is the sample logarithmical average life, n is the sample size, ϑ is the degree of freedom, $\frac{\alpha}{2}$ is the degree of confidence, and confidence is equal to $(1 - \alpha)$, S_1 is the logarithmical standard deviation of sample life. Note that the value of t statistics is available in standard tables.

The present study reviewed the effect of stress level on the data scatter and coefficient of variation. S_1 and $\overline{X}_1$ values were evaluated in the linear scale.

Probability distributions of samples

For the evaluation of the distribution of each sample data set, two-parameter and three-parameter Weibull distribution functions can be established as[6]:

$$F\left(N_f\right) = 1 - e^{-\left(\frac{N_f - N_{f_0}}{\varnothing - N_{f_0}}\right)^b} \tag{3}$$

where $F\left(N_f\right)$ is the failure fraction in the test data set N_f, N_{f_0} is the minimum expected fatigue life, $\theta = \varnothing - N_{f_0}$ is characteristic fatigue life (cycles when 63.2% have failed), and b is the Weibull slope or shape parameter. The terms N_{f_0}, θ, and b are 3-parameter Weibull model, and for the 2-parameter Weibull model, the parameter of N_{f_0} is zero, $N_{f_0} = 0$.

To determine the Weibull equation coefficient for scattered test data (which is the key task in the present work), a step function, i.e., $(i - 0.3)/(n + 0.4)$ is commonly used as a percent of failure. In the present work, every failure life of N_{f_i} in test results is considered as a Weibull distribution with a median value of N_{f_i} and cumulative probability of $(i - 0.3)/(n + 0.4)$. The failure fraction for median value of N_{f_i} in Eq. (3) will be 50%, so characteristic fatigue life, θ_i, the minimum expected fatigue life, γ and the Weibull slope, b for any failure life, N_i will be:

$$\text{for } i = 1 \text{ to } n : 0.5 = e^{-\left(\frac{N_i - \gamma}{\theta_i}\right)^b} \tag{4a}$$

Then:

$$\theta_i = \frac{N_i - \gamma}{\sqrt[b]{-\ln(0.5)}} \tag{4b}$$

The Weibull equation for every test result from Eq. (4b) is:

$$P_i(N@S_{ca}) = e^{-\left(\frac{N_i - \gamma}{\theta_i}\right)^b} \tag{4c}$$

where $P_i(N@S_{ca})$ is the expected failure fraction for fatigue life N at applied stress level of S_{ca}. By applying Eq. (4c) and weight factor, the modified Weibull function is derived:

$$P'(N_j@S_{ca}) = 1 - \sum_{i=1}^{n} \frac{1}{n} \int_0^{N_j} (1 - P_i(N@S_{ca}))dN \tag{5}$$

Weibull equation coefficients are determined based on the flowchart in Fig. 1. First, Weibull parameters will be calculated by the common Weibull model (Stages 1 to 3 in the flowchart). Then, for every single test data, an equivalent Weibull distribution with the mean value of the same test fatigue life and cumulative probability of common Weibull method and modified Weibull parameters will be calculated using Eqs. (4a), (4b), (4c) and (5) (Stages 4 and 5 in the flowchart). Finally, test data were compared with calculated parameters and presented probability distribution model using the Kolmogorov–Smirnov test (K–S test), as follows (Stages 6 and 7):

$$D_i = max_{i=1}^{k} \left[\left| F^*(x_i) - F(x_i) \right| \right] \tag{6}$$

This process will be repeated until the error valve becomes acceptable.

R–S–N relation
In Basquin power law, a log–log straight linear relationship is considered between the applied stress cycles and the number of cycles to failure. R–S–N curves are depicted by determining the fatigue life using the Weibull
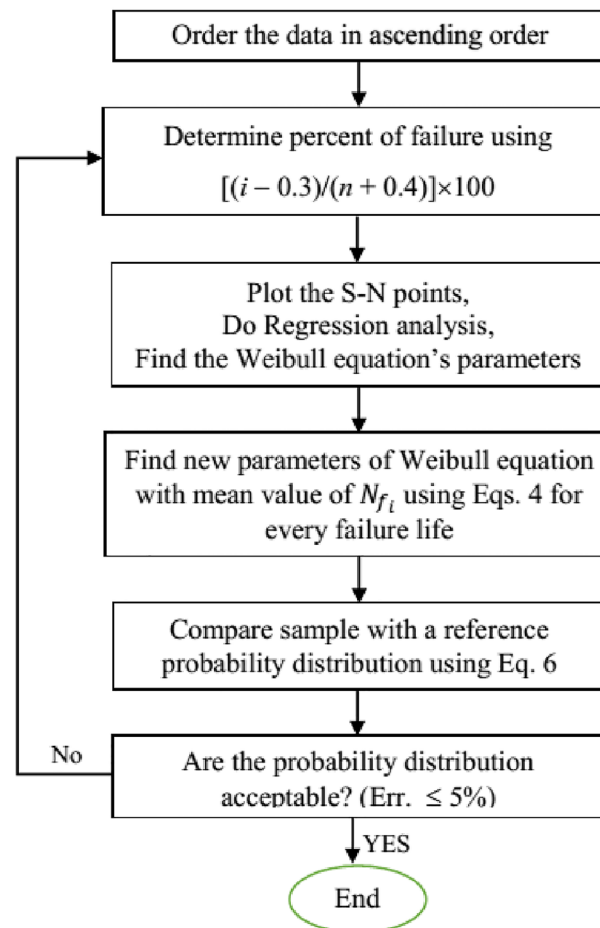

Figure 1. Flowchart of parameters calculation for Modified Weibull distribution.

equation of every data set for each reliability at every specified stress level. Then, regression analysis is done to fit a linear curve to the S–N data.

$$\log(N) = A - B \log(\Delta\sigma); \ \ \Delta\sigma \geq \Delta\sigma_\infty \tag{7}$$

In statistical theory, regression analysis is a process of identifying data trends. Regression analysis commonly uses regression analysis to evaluate relationships between some factors, including a dependent factor and other independent factors (variables). This process shows which variable is essential and which f variable can be ignored. It also shows how these variables affect each other. Linear regression is usually applied to estimate data trends. This method depends on the problem and determines one or more lines to fit the data with minimum error according to a specific mathematical calculation like ordinary least squares. This technique calculates a unique line with a minimum difference between the true data and that line compared to the sum of squared differences. In the statistical analysis of fatigue problem, selecting the best curve fitting method sometimes become very complicated. To avoid uncertainty about a divergent solution, we applied the K–S test to the presented sequences method at the final step.

Results and discussion

The test data used to evaluate the modified Weibull probability distribution parameters are extracted from the literature[4,36–39] and presented as six grouped sets of fatigue life data in Table 3. Every data set is a set of fatigue life experimentally measured at same condition and every data group is some data set at different stress levels. C and F data groups have been accomplished only in a single stress level and thus in this study only are applied to modify Weibull parameters. The other data groups in Table 3 are fatigue life scatter at different applied stress levels and are used to modify Weibull parameters and also evaluate stress-life relation for various reliability levels. For numerical solution of the mentioned procedure, a simple M-File code was written in Matlab and Matlab curve fitting tool was applied for regression analysis.

Survival probabilities of test data

The statistical parameters can be determined by repeating the steps of flowchart in Fig. 1. Also, the P-N_f (survival probability-fatigue life) relation and Weibull parameters can be evaluated using Eqs. (3), (4c), and (5). Applying regression analysis and curve fitting, Weibull parameters of b, θ, and N_{f_0} can be determined in terms of the least square method (Fig. 2). The difference between test results and the distribution model shows good compliance in the modified Weibull model. The effect of sample size and deviation from Weibull distribution was reduced in the test points.

Further, by applying Eqs. (3), (4c), and (5), curve fitting, and determining Weibull parameters, the error between curves and test data sets can be obtained using Eq. (6) for the K–S test and R-square value, respectively. The results of the present modeling are given in Table 4. Here, the value of θ in Weibull parameters is almost constant. All the K–S test and R-square values show less error in the modified Weibull model, which complies with curves and test data in Fig. 2.

As shown in Table 4, using the modified method, the error values of the K–S test declined, and the R-square value approached 1. In all cases, the Weibull slope or shape parameter (b) decreased. The characteristic life θ has changed from 1.5 to 14%, but the minimum time or cycles to failure N_{f_0} has changed considerably up to 10

Data group	Data set	Fatigue life (10^6 cycles)
A[4]	Data set 1 (S = 4900 MPa)	22.5, 38.8, 64.2, 141, 141, 170, 185, 210, 361, 416, 421, 486, 829, 1390
	Data set 2 (S = 5500 MPa)	8.77, 15, 24.8, 54, 64.9, 65.2, 70.7, 80.2, 137, 158, 160, 184, 312, 520
	Data set 3 (S = 6100 MPa)	5.88, 9.43, 14.6, 29, 34.1, 34.2, 36.7, 41, 65.7, 74.2, 75.1, 85, 135, 211
	Data set 4 (S = 6500 MPa)	1.63, 2.88, 3.72, 4.38, 5.98, 7.99, 13.9, 18.3, 50.7, 57.3, 66.1, 70.7, 103, 179
B[36]	Data set 5 (S = 126 MPa)	0.766126, 0.775688, 1.084235, 1.236622, 1.272766, 1.313096, 1.333748, 1.334498, 1.401002, 1.514865, 1.689579, 1.882968, 1.904087, 1.964597
	Data set 6 (S = 144 MPa)	0.28268, 0.30957, 0.3551, 0.398319, 0.45602, 0.45648, 0.496652, 0.5232, 0.554726, 0.647726, 0.66155, 0.683264, 0.739415, 0.827473, 0.8856, 0.897183, 0.989, 1.20064
	Data set 7 (S = 180 MPa)	0.112302, 0.1278, 0.1335, 0.13371, 0.13384, 0.138322, 0.142491, 0.235782, 0.2898, 0.302732, 0.3303
C[37]	Data set 8 (S = 67.62 MPa)	0.128657, 0.398586, 0.656841, 0.799445
D[38]	Data set 9 (L = 140 mm, S = 46.14 MPa)	8.704, 10.7, 13.17, 13.23, 22.13
	Data set 10 (L = 140 mm, S = 76.96 MPa)	2.33, 2.156, 2.091, 1.25, 1.695
	Data set 11 (L = 140 mm, S = 114.66 MPa)	0.467, 1.16, 0.693, 0.759, 5.67
	Data set 12 (L = 140 mm, S = 160.53 MPa)	0.28, 0.202, 0.214, 0.769
E[39]	Data set 13 (L = 19 mm, S = 53 MPa)	49.2, 49.6, 45, 44.1
	Data set 14 (L = 19 mm, S = 62.5 MPa)	9.3, 13.2, 14, 15.7
	Data set 15 (L = 19 mm, S = 89.5 MPa)	3.87, 3.950, 2.42, 2.82
	Data set 16 (L = 19 mm, S = 132 MPa)	1.22, 1.66, 2.39, 2.66
F[39]	Data set 17 (concrete with steel fiber, V = 1%, S = 0.8 MPa)	0.040135, 0.050139, 0.058141, 0.082345, 0.0984, 0.12862, 0.146544, 0.23916, 0.292488, 0.460482

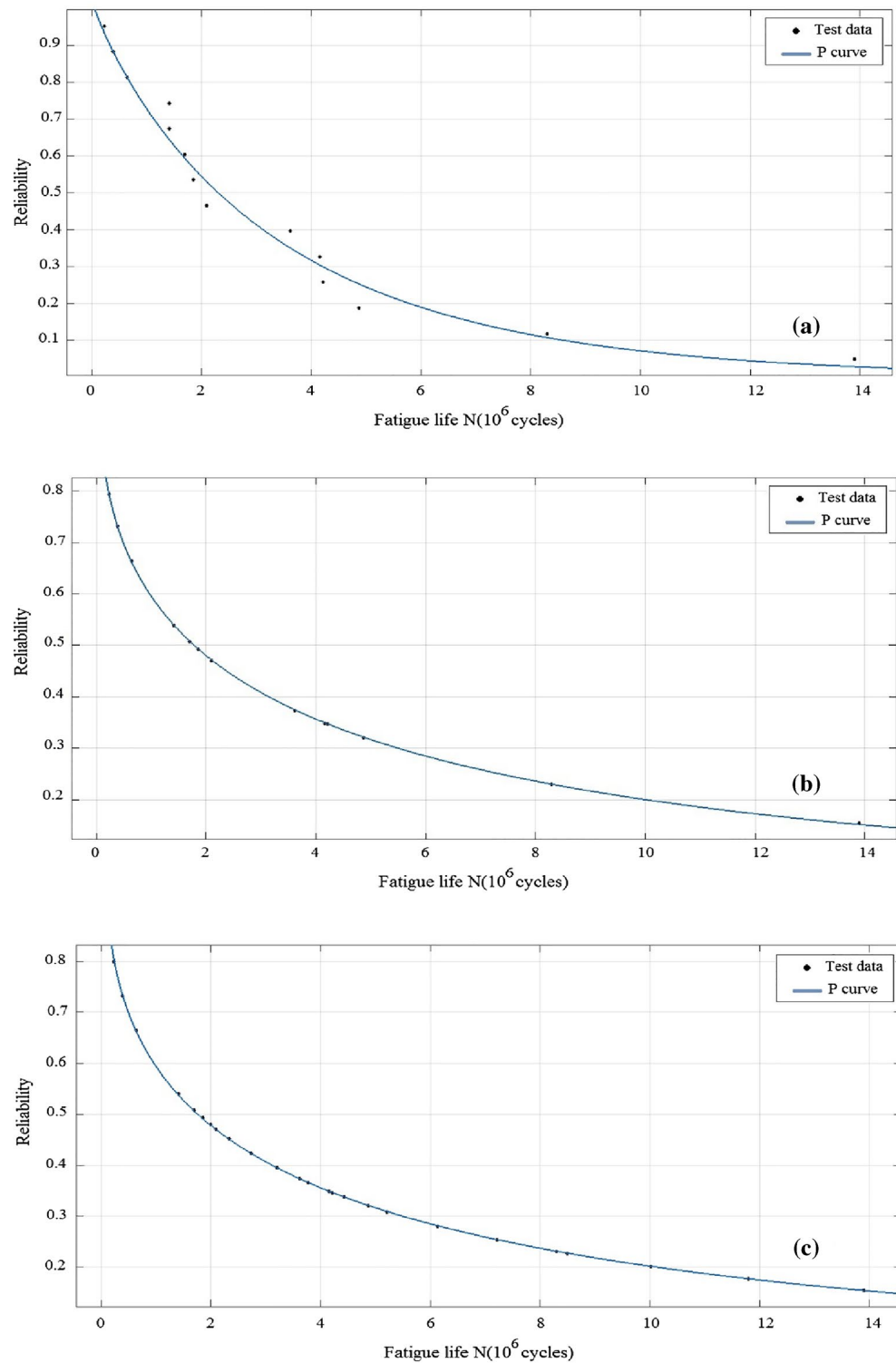**Table 3.** Test data used to evaluate the modified Weibull parameters.

Figure 2. Calculated reliability for (**a**) Common Weibull distribution, (**b**) Modified Weibull distribution and (**c**) Modified Weibull distribution with artificially increase of test data.

times. The values of minimum expected cycles to failure obtained from Weibull Distribution in this approach should not be used directly for design.

For the set of rolling contact fatigue life (i.e., data set 1), the smaller fatigue life will have a larger survival probability. Survival probability data can be evaluated using Eqs. (4a) and (5). Figure 2 shows the results of regression analysis and curve fitting. For the test data, reliability of fatigue life is calculated based on the modified and common Weibull distributions. As indicated in Fig. 2, modified Weibull parameters have better compliance

Test data[4,37-39]	Distribution function	Weibull parameters			Error	
		b	θ	N_{f_0}	K–S test (%)	R-square
Data set 1	Step function	0.89	349,600,000	5,497,000	8	0.986
	Modified Weibull	0.62	344,000,000	9,050,000	6	0.999
Data set 2	Step function	0.90	132,806,849	4,384,680	10.2	0.967
	Modified Weibull	0.62	137,984,686	4,019,219	5.4	0.999
Data set 3	Step function	1	63,647,140	2,938,570	8.7	0.969
	Modified Weibull	0.7	65,883,507	2,160,990	5.2	0.999
Data set 4	Step function	0.58	32,506,615	1,479,600	11.6	0.963
	Modified Weibull	0.4	37,336,316	2,574,857	9.2	0.997
Data set 6	Step function	1.59	602,100	117,100	9.3	0.93
	Modified Weibull	0.86	592,300	79,200	5	1
Data set 8	Step function	0.85	169,100	61,130	16	0.976
	Modified Weibull	0.69	176,000	140,000	2	1
Data set 16	Step function	1.1	179,400	1647	10.2	0.980
	Modified Weibull	0.7	168,000	15,938	5.8	0.993

Table 4. Results of modified Weibull parameters.

between the Weibull model and test results, in the both cases of equal and artificially increased sample size. Here, the effect of the step function for percent of failure and deviation from the Weibull probability distribution function has been eliminated.

Figure 3 shows the R-Square values of common Weibull and presented model with respect to sample size. It can be seen that as the sample size increases, the error of the common Weibull model becomes negligible. Small sample size cases had a greater jump in the percent failure function, leading to higher error values.

The effect of normalized stress level (with respect to the maximum stress in each data set) on the fatigue life scatter is presented in Fig. 4. As can be seen from Fig. 4, the lower applied stress amplitudes have less data scattering.

R–S–N relation

By applying Eq. (7), Weibull parameters of datasets 5–7 from Table 3, and using regression analysis, Basquin parameters A and B can be determined in terms of the least square method. The S–N curves in Fig. 5 show that for the higher survival probability, the expected fatigue life would decrease at every stress level. Engineering designs usually are in the 0.01 percent probability of failure range[6]. Therefore, extrapolation is required. By extrapolating these curves to the helpful percent probability of failure range, the curves would intersect, which is unreasonable. Also, Table 4 shows that the value of least fatigue life has not reasonable behavior with an increase in the stress level.

The deviation from the linear relation in the logarithmic scale is calculated considering the Basquin S–N relation and using regression analysis. Figure 6 shows the curve fitting error in the 3 reliability levels. As can be seen, the error increased at a higher reliability level.
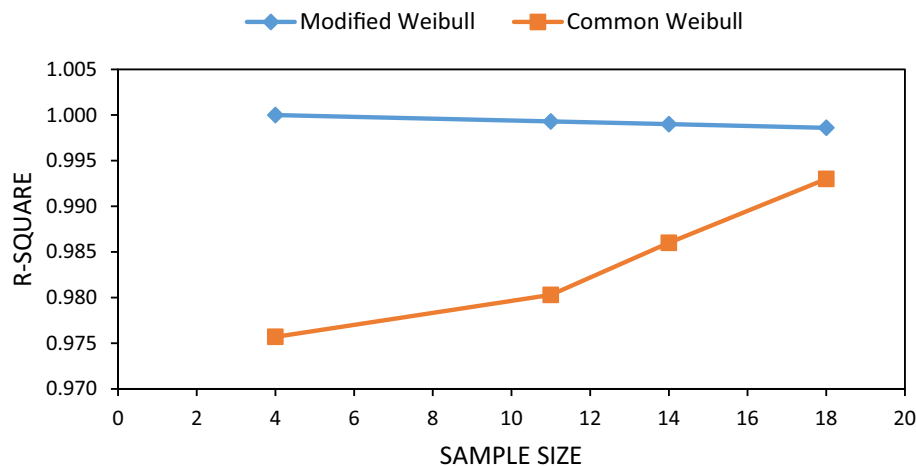

Figure 3. Effect of sample size on the R-Square value for Weibull distribution (data sets 5, 6, 7 and 8).
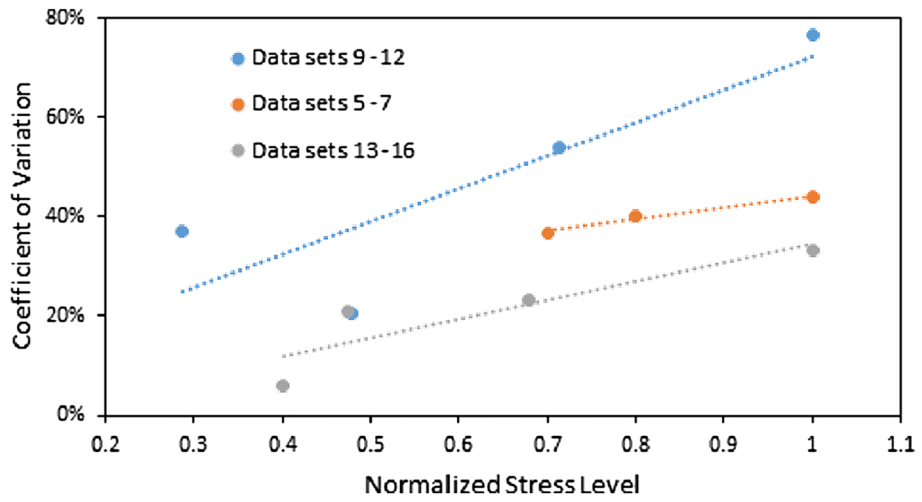
Figure 4. Effect of stress level on the coefficient of variation of fatigue life.
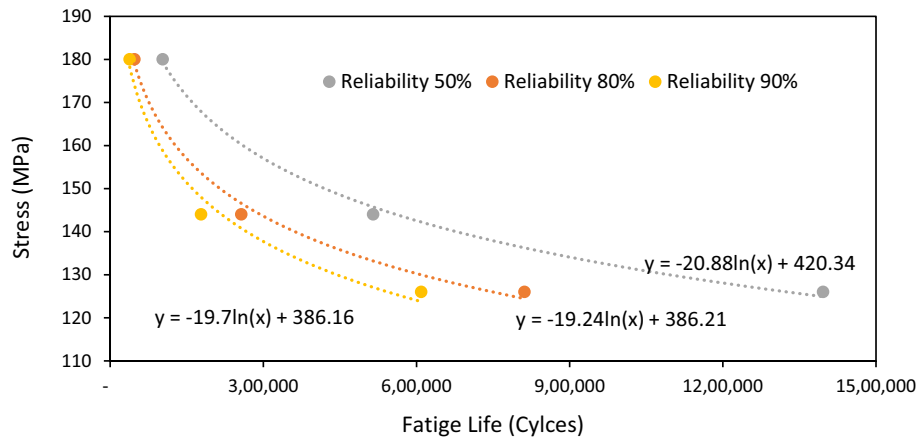

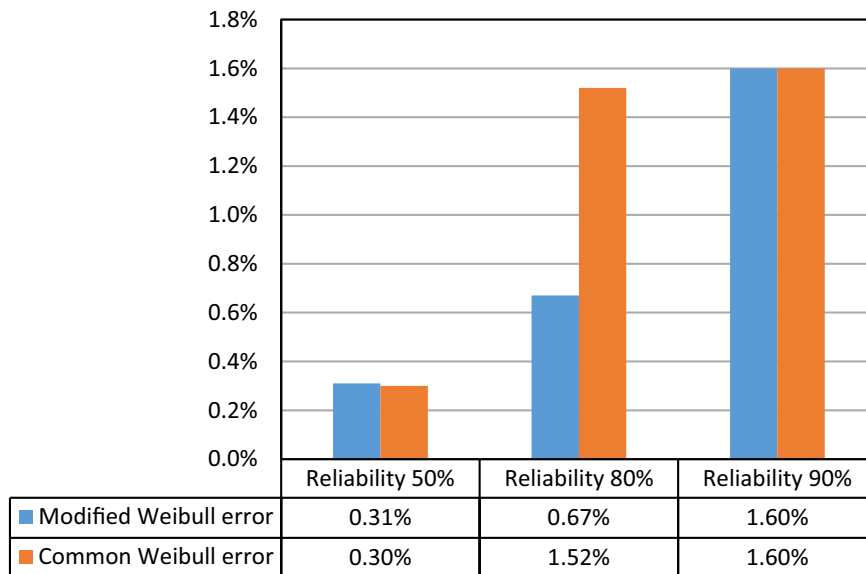**Figure 5.** R–S–N curves for 50%, 80%, and 90% reliability.

	Reliability 50%	Reliability 80%	Reliability 90%
Modified Weibull error	0.31%	0.67%	1.60%
Common Weibull error	0.30%	1.52%	1.60%

Figure 6. Curve fitting error (1 minus R-Square) in R–S–N curves.

Conclusion

A general method for estimating the parameters of Weibull distribution for modeling general fatigue life data scatter for every applied load and stress region was developed in this study. By considering any result of fatigue test as an equivalent Weibull distribution, artificial data are generated and the accuracy of common Weibull distribution model can be improved. Next, the corresponding fatigue life was evaluated at any reliability using the determined distribution model for any specified stress level. Using Basquin Stress-Life relation and fitting the S–N curves, R–S–N curves were obtained at any reliability level. Overall, the major results of this study can be outlined as follows:

1. The presented method causes the failure percent of test data to increase smoothly and be close to the Weibull distribution curve.
2. Modified coefficients of the distribution function have an acceptable error in the K–S test and R-square value. Also, the difference between test results and probability distribution function was decreased.
3. The effect of discretizing of percent of failure in the sample fatigue life results was decreased, and this approach can be used in all the other small sample size test data. Modifying Weibull parameters has a greater effect on the decreasing error.
4. The fatigue life scatters increase in the higher stress levels.
5. R–S–N curves were obtained using the Basquin Stress-Life relation.
6. The results show the curves cannot be extrapolated to minimum expected life and very high levels of reliability.

Future research(es) may involve using neutrosophic statistics to extend this study. Neutrosophic statistics is the extension of classical statistics and is applied when the data is coming from an uncertain environment like new pandemic or from a complex process like fatigue problems in engineering[40–43].

Data availability

All data generated or analyzed during this study are included in this published article.

References

1. Campbell, F. C. *Elements of Metallurgy and Engineering Alloys* (ASM International, 2008).
2. Murakami, Y., Takagi, T., Wada, K. & Matsunaga, H. Essential structure of S-N curve: Prediction of fatigue life and fatigue limit of defective materials and nature of scatter. *Int. J. Fatigue.* **146**, 106138 (2021).
3. Castilloa, E. & Canteli, F. A compatible regression Weibull model for the description of the three-dimensional fatigue σ$_M$-N-R field as a basis for elative damage approach. *Int. J. Fatigue.* **155**, 106596 (2022).
4. Zhao, Y. & Liu, H. Weibull modeling of the probabilistic S–N curves for rolling contact fatigue. *Int. J. Fatigue.* **66**, 47–54 (2014).
5. Kaiser, M. *The Offshore Pipeline Construction Industry* (Gulf Professional Publishing, 2020).
6. Stephens, R. I., Fatemi, A., Stephens, R. O. & Fonchs, H. O. *Metal Fatigue in Engineering* (Wiley, 2000).
7. Ash, R. B. *Basic Probability Theory* 66–69 (Dover Publications, 2008).
8. Sinclair, G. M. & Dolan, T. J. Effect of stress amplitude on statistical variability in fatigue life of 75S–T6 aluminum alloy. *Trans. ASME.* **75**, 867–872 (1953).
9. Vilfredo, P. Cours d'economie politique. *J. Polit. Econ.* **6**, 549–552 (1898).
10. Gumbel, E. J. Les valeurs extrêmes des distributions statistiques. *Ann. l'Inst. Henri Poincaré.* **5**, 115–158 (1935).
11. Kissell, R. & Poserina, J. *Optimal Sports Math, Statistics, and Fantasy* 103–135 (Academic Press, 2017).
12. Li, Y., Chiang, J. Y., Bai, Y. & Chai, K. C. Estimation of process performance index for the two-parameter exponential distribution with measurement error. *Sci. Rep.* **13**, 2327 (2023).
13. Naveed, M. *et al.* Control chart for half normal and half exponential power distributed process. *Sci. Rep.* **13**, 8663 (2023).
14. Störzel, K. & Baumgartner, J. Statistical evaluation of fatigue tests using maximum likelihood. *Mater. Test.* **63**, 714–720 (2021).
15. Wang, B., Islam, F. & Mair, G. W. Evaluation methods for estimation of Weibull parameters used in Monte Carlo simulations for safety analysis of pressure vessels. *Mater. Test.* **63**, 379–385 (2021).
16. Weibull, W. A statistical distribution function of wide applicability. *J. Appl. Mech.* **73**, 293–297 (1951).
17. Weibull, W. *Fatigue Testing and Analysis of Results* (Pergamon Press, 1961).
18. Lone, ASh., Rahman, A. & Islam, A. Step-stress partially accelerated life testing plan for competing risk using adaptive type-I progressive hybrid censoring. *Pak. J. Stat.* **33**(4), 237–248 (2017).
19. Xiong, Y., Yu, Q. & Jiang, Y. Multiaxial fatigue of extruded AZ31B magnesium alloy. *Mater. Sci. Eng. A.* **546**, 119–128 (2012).
20. Jiang, H., Xie, M. & Tang, L. C. Markov chain Monte Carlo methods for parameter estimation of the modified Weibull distribution. *J. Appl. Stat.* **35**, 647–658 (2008).
21. Fernández Canteli, A., Castillo, E., Blason, S., Correia, J. A. F. O. & de Jesus, A. M. P. Generalization of the Weibull probabilistic compatible model to assess fatigue data into three domains: LCF, HCF and VHCF. *Int. J. Fatigue* **159**, 106771 (2022).
22. Strzelecki, P. Determination of fatigue life for low probability of failure for different stress levels using 3-parameter Weibull distribution. *Int. J. Fatigue* **145**, 106080 (2021).
23. Acosta, R. *et al.* Evaluation of S-N curves including failure probabilities using short-time procedures. *Mater. Test.* **63**, 705–713 (2021).
24. Dhillon, B. S. *Design Reliability: Fundamentals and Applications* (CRC Press, 1999).
25. Vidović, Z. On MLEs of the parameters of a modified Weibull distribution based on record values. *J. Appl. Stat.* **46**, 715–724 (2019).
26. Usabiaga, H., Muniz-Calvente, M., Ramalle, M., Urresti, I. & Fernández, C. A. Improving with probabilistic and scale features the Basquin linear and bi-linear fatigue models. *Eng. Fail. Anal.* **116**, 104728 (2020).
27. Jafari, A. & Alizadeh Kaklar, J. Determination of the critical length of a subsurface crack in a monobloc R7T railway wheel using FEM analysis. in *Proceedings of the ASME 2010 International Mechanical Engineering Congress and Exposition. 2012; Volume 11: New Developments in Simulation Methods and Software for Engineering Applications; Safety Engineering, Risk Analysis and Reliability Methods; Transportation Systems*, 863–868 (2010).

28. Abdoli, A., Khezri, J. & Alizadeh, K. J. A new weight function for one-dimensional subsurface cracks under general loading. *Fatigue Fract. Eng. Mater. Struct.* **43**, 433–443 (2020).
29. Samadlou, F. & Alizadeh, K. J. Propagation pattern for a two-dimensional subsurface crack under a moving contact pressure. *Eng. Fract. Mech.* https://doi.org/10.1016/j.engfracmech.2020.107002 (2020).
30. Stromeyer, C. E. The determination of fatigue limits under alternating stress conditions. *Proc. R. Soc. Lond.* **90**, 411–425 (1914).
31. Battelle Memorial Institute, National Research Council. *Prevention of the Failure of Metals Under Repeated Stress* (Wiley, 1941).
32. Palmgren, A. G., Lebensgauer, D. & Kugellagern, V. Life length of roller bearings or durability of ball bearings. *Z. Vereinesd Dtsch. Ingenieure* **14**, 339–341 (1924).
33. Metallic Materials: Fatigue Testing: Statistical Planning and Analysis of Data. ISO 12107 (2012).
34. Ling, J. & Pan, J. A maximum likelihood method for estimating P-S-N curves. *Int. J. Fatigue.* **19**, 415–419 (1997).
35. Kohout, J. & Veˇchet, S. A new function for fatigue curves characterization and its multiple merits. *Int. J. Fatigue.* **23**, 175–183 (2001).
36. Caiza, P. D. T. & Ummenhofer, T. Consideration of the runouts and their subsequent retests into S-N curves modelling based on a three-parameter Weibull Distribution. *Int. J. Fatigue.* **106**, 70–80 (2018).
37. Giancaspro, J., Taam, W. & Wong, R. Modified joint Weibull approach to determine Load enhancement Factors. *Int. J. Fatigue.* **31**, 782–790 (2009).
38. Fatigue design curves for welded joints in air and seawater under variable amplitude loading. Offshore Technology Report: OTO 1999 058 (Failure Control Engineering & Materials Consultants, 2000).
39. Singh, S. P. & Kaushik, S. K. Flexural fatigue life distributions and failure probability of steel fibrous concrete. *ACI Mater. J.* **97**, 658–667 (2000).
40. Asalam, M. A new sampling plan using Neutrosophic process loss consideration. *Symmetry.* **10**, 32–37 (2018).
41. Hameed, M. S., Ahmad, Z., Shahbaz Ali, S. H., Muhammad Kamran, M. & Babole, A. R. L. An approach to (μ, ν, ω)-single-valued neutrosophic submodules. *Sci. Rep.* **13**, 751 (2023).
42. Duran, V., Topal, S., Smarandache, F. & Aslam, M. Using the four-valued Rasch model in the preparation of neutrosophic form of risk maps for the spread of COVID-19 in Turkey. *Cogn. Data Sci. Sustain. Comput.* **1**, 43–69 (2023).
43. Sundareswaran, R. *et al.* Assessment of structural cracks in buildings using single-valued neutrosophic DEMATEL model. *Mater. Today Proc.* **65**, 1078–1085 (2022).

Author contributions

H.F. wrote the main manuscript text and J.A.K. reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.