On the parametric assessment of fatigue disparities

KUFOIN, Elvis N and SUSMEL, Luca <http://orcid.org/0000-0001-7753-9176>

Probabilistic Engineering Mechanics 77 (2024) 103651

Contents lists available at ScienceDirect

Probabilistic Engineering Mechanics

journal homepage: www.elsevier.com/locate/probengmech

On the parametric assessment of fatigue disparities

Elvis N. Kufoin, Luca Susmel *

Department of Civil and Structural Engineering, The University of Sheffield, Mappin Street, Sheffield, S1 3JD, UK

ARTICLE INFO

Keywords:
Statistical
Significance
Test-statistic
Collinear
Fatigue

ABSTRACT

Efficiently merging fatigue datasets from diverse sources has proven to be a strategic approach for enhancing the reliability of fatigue assessment and design within industry, while concurrently streamlining costs and time. Statistical parametric analysis is an approach that can be applied to fatigue datasets to determine whether the datasets can be deemed statistically significant (different) or statistically insignificant (similar). This paper systematically employed statistical parametric test-statistic hypotheses to assess significance. To validate this approach the paper used as a case study, fatigue data sets generated from varied notched specimens with hole diameters ranging from 0 mm to 3 mm, in addition to data from the literature. In particular, gross stresses were utilized to ensure that the only means to identify differences in the fatigue datasets was through statistical analysis. This approach was observed to work well for geometries with differences in notch geometry as small as 1 mm and was able to identify notch insensitivity in cast iron. Thus, this method can be used to differentiate fatigue datasets based on statistical parameters rather than other physical parameters.

1. Introduction

Fatigue assessment in engineering is a critical process when evaluating the structural durability and performance of structural components under dynamic loading. It predicts potential fatigue failure over time considering cyclic stresses and strains [1]. Historical failures underscore the importance of robust fatigue analysis [2]. Various methods including theoretical approaches, fracture mechanics, and non-destructive testing, to enhance fatigue behaviour understanding and improve engineering safety have been used in industry. With fatigue failure accounting for a majority of engineering failures, accurately assessing fatigue is therefore essential to prevent these failures. Engineers can proactively address weaknesses, optimize designs, and ensure long-term reliability and safety across diverse industries.

Given the critical nature of fatigue data assessment for designing resilient structures and the variability in fatigue datasets from experiments, statistical analysis serves as a potent tool for exploration. In addition, fatigue experiments are time-consuming and costly, and designing reliably requires ample fatigue data for analysis. Addressing these challenges involves utilizing datasets from different sources and applying standard approaches for amalgamation and validation, such as parametric analysis. This method evaluates differences in fatigue properties, revealing intrinsic mechanical characteristics of materials and establishing significance within diverse experimental conditions or environments. Not only does this aid in consolidating fatigue datasets for heightened reliability but it also mitigates the time and cost challenges associated with conducting numerous fatigue experiments [3–5]. A detailed methodology for parametric statistical testing constitutes a comprehensive exploration, enriching the field of fatigue data analysis with valuable insights for engineers and researchers striving to optimize designs and improve material reliability under cyclic loading conditions. Consequently, conclusions drawn from fatigue test data can be made confidently based on a predetermined level of significance.

This paper seeks to address the research question of whether statistics alone is sufficient to determine the similarity or dissimilarity between two datasets. This is not only useful for comparing fatigue data sets, but it can also be used to differentiate data sets produced using different manufacturing, specimen preparation protocols or loading conditions (tension, torsion, load ratio, etc). To answer this question, a hypothesis is formulated and tested. In this study, the hypothesis under consideration is the t-test statistics in parametric analysis. The fatigue data sets used in this study are derived from notches, which serve as a case study. By employing notches of varying geometries, different fatigue datasets are generated and utilized to validate the hypothesis.

Notches were selected as the basis for generating fatigue data for this paper because they are well-known for reducing the endurance limit of engineering materials [6,7] as well as the inverse slope [8], effectively shifting the mean S-N curve downward. They also change the level of scattering depending on the notch root radius and the material type [9].

* Corresponding author.
E-mail address: l.susmel@sheffield.ac.uk (L. Susmel).

https://doi.org/10.1016/j.probengmech.2024.103651
Received 22 March 2024; Received in revised form 4 June 2024; Accepted 10 June 2024
Available online 13 June 2024

Nomenclature

C_0	Intercept of mean S-N curve (constant)
C_1	Coefficient of independent variable (constant)
$C_{1,i}$	Coefficient of independent variable of ith data set, $i = 1,2.$.
k	Negative inverse slope
f	Degrees of freedom
F_{cal}	F-test statistic for variance calculated
F_{crit}	F-test statistic for variance from tables
$log\,N_{f,ij}$	Log of life at the replication level
$\overline{logN_{fij}}$	Mean log of life at the replication level
$\overline{log\,\sigma_i}$	Mean log of stress level, $\overline{log\,\sigma_i} = \sum_{i=1}^{n} \frac{log\,\sigma_i}{n}$
$\overline{logN_{f,i}}$	Mean log of fatigue life, $\overline{logN_{f,i}} = \sum_{i=1}^{n} \frac{log\,N_{f,i}}{n}$
$log\,N_{f,D}$	Log of estimated life for design life
m_i	Replication level at the ith stress level
N_f	Number of cycles to failure
$N_{f,i}$	Number of cycles to failure at the ith stress level
n	Number of experimental results (sample set)
N_A	Reference number of cycles to failure
N_{kp}	Number of cycles to failure at the knee point
P_s	Probability of survival
q	Index depending on the probability of survival
R	Stress ratio ($R = \sigma_{min}/\sigma_{max}$)
s	Standard deviation of log cycles to failure
$s_{x_1}{}^2$	Variance of measured quantity x
$s_e{}^2$	Equivalent variance of two homogenous data sets
$s_{C_1}{}^2$	Variance of common line when slopes of parallel lines are insignificant
x_i	Arbitrary measured quantities
t	Test statistic
t_β	Critical value corresponding to a significance level β
t_μ	Test statistic calculated for means
t'_μ	Test statistic calculated for means that are significant
t_{C_1}	Test statistic for slopes
$t(\delta)$	Test statistic for collinear lines
σ	Generic stress level (stress amplitude, maximum stress or stress range)
σ_i	ith stress level (i = 1, 2 …, n)
$\sigma_{min}, \sigma_{max}$	Minimum and maximum stress in the cycle
σ_0	Endurance limit
$\sigma_{0,P\%}$	Endurance limit at a probability of survival P
$\sigma_{0,(1-P)\%}$	Endurance limit in error at a probability of survival P
T_σ	Scatter ratio of reference stress for (1-P) % and P% probabilities of survival
β	Level of significance
δ	Random variable for collinear lines
$\mu_{Y/X}$	The expected value of $logN$ given $log\,\sigma$

By introducing a notch, it is anticipated that the fatigue data sets will be statistically significant. Altering the notch geometry allowed distinct fatigue data sets to be generated that were subsequently employed in the statistical significance testing. Moreover, notch geometry is a parameter that can be conveniently controlled and accurately measured. For this reason, three different materials – steel, cast iron, and brass were used to produce the notched specimens, with three different notch diameters: 0 mm, 1 mm, 2 mm, and 3 mm.

In the analysis of the fatigue results, only the gross stresses were taken into consideration. The gross stress is the nominal stress experienced by the specimen, assuming it is smooth and free of any geometric discontinuities that could affect the stress distribution. By using gross stresses, all specimens are therefore assumed to have the same geometry and the statistical approach is the only tool used to identify the difference in notch geometry by analysing the experimental results. By so doing, this approach focused only on utilizing statistical methods to determine whether the discrepancies in the generated fatigue data sets were statistically significant, based on a predetermined threshold.

2. Review of Wöhler curves and mean parameters from fatigue data points $(\sigma_i, N_{f,i})$

The stress-based approach to fatigue assessment relies on S–N curves, commonly known as Wöhler curves [10]. These curves are derived from subjecting identical and standardized specimens to constant amplitude cyclic loading until failure occurs. Fatigue data sets $(\sigma_i, N_{f,i})$ generated from these experiments are used to produce the S-N curve. The S-N curve in this paper will be limited to the medium cycle fatigue regime, which is crucial for understanding the endurance and fatigue life of materials, thereby optimizing design and manufacturing processes in various engineering applications. This curve in the medium cycle fatigue regime is assumed to be linear [11–15] and there are various formulations to represent this curve. One of the most popular formulation is the Basquin equation which is transformed in the log-log space as [16]:

$$LogN_f = C_o + C_1\ Log\,\sigma \tag{1}$$

where N_f is the fatigue life and σ is the stress level, which can either be the stress amplitude, range or maximum stress. C_o and C_1 are the intercept and inverse slope constant respectively which are dependent on the fatigue data. The inverse slope determines the sensitivity of the material to fatigue, and a lower value indicates that the material is more sensitive to small changes in stress or strain amplitude. Generating the S-N curve represented by equation (1) involves a few assumptions (see Refs. [11,117] and the references reported therein). The experimental data points have a standard deviation (s) which is observed in the degree of scatter and defined by the change in the residuals around the mean [18,19,20]. Therefore, the mean parameters representing the fatigue properties from the fatigue data sets can be checked for statistical significance. These mean parameters include the variance s^2, slope C_1 , and the vertical intercept C_0. These mean parameters are calculated as:

$$C_1 = \frac{\sum_{i=1}^{n}\left[log\,(\sigma_i) - \overline{log\,\sigma_i}\right]\left[log\,(N_i) - \overline{log\,N_i}\right]}{\sum_{i=1}^{n}\left[log\,(\sigma_i) - \overline{log\,\sigma_i}\right]^2} \tag{2}$$

and

$$C_0 = \overline{log\,N_i} - C_1\overline{log\,\sigma_i} \tag{3}$$

$$s^2 = \frac{\sum_{i=1}^{n}\left(logN_{fi} - log\,N_f\right)^2}{n-2} \tag{4}$$

The scatter band generated as a result of the scattering of the data is calculated using the equation [11,12,20]:

$$\text{Log}\,N_f = C_0 + C_1\,log\,\sigma_i \pm t\,s\sqrt{1 + \frac{1}{n} + \frac{\left(log\,\sigma_i - \overline{log\,\sigma_i}\right)^2}{\sum_{i=1}^{n}\left(log\,\sigma_i - \overline{log\,\sigma_i}\right)^2}} \tag{5}$$

in which C_0 and C_1 are the intercept and inverse slope constant respectively which are dependent on the fatigue data. t is the corresponding percentage point of the student's t-distribution with of degrees of freedom equal to $n-2$ and s^2 is the best guess of the variance around

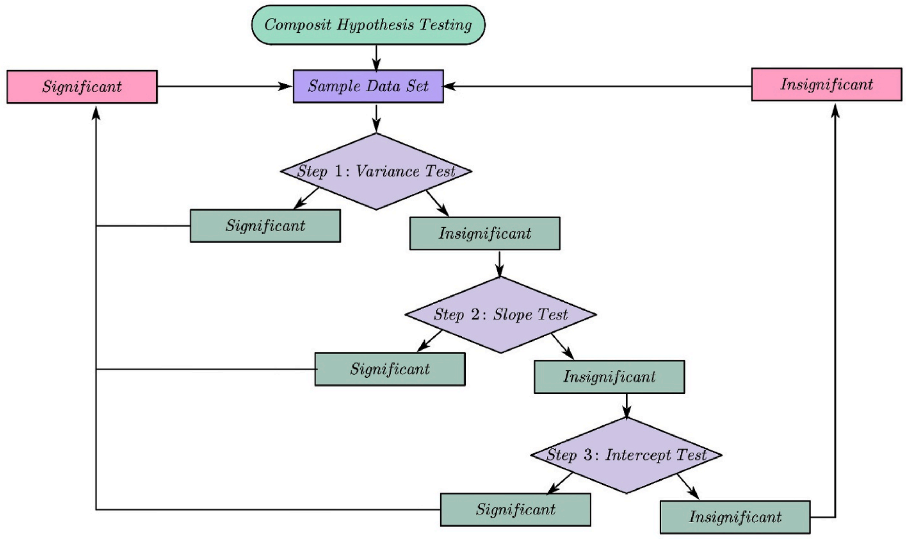


Fig. 1. The flow chart to illustrate how to test the statistical significance of two S-N curves.

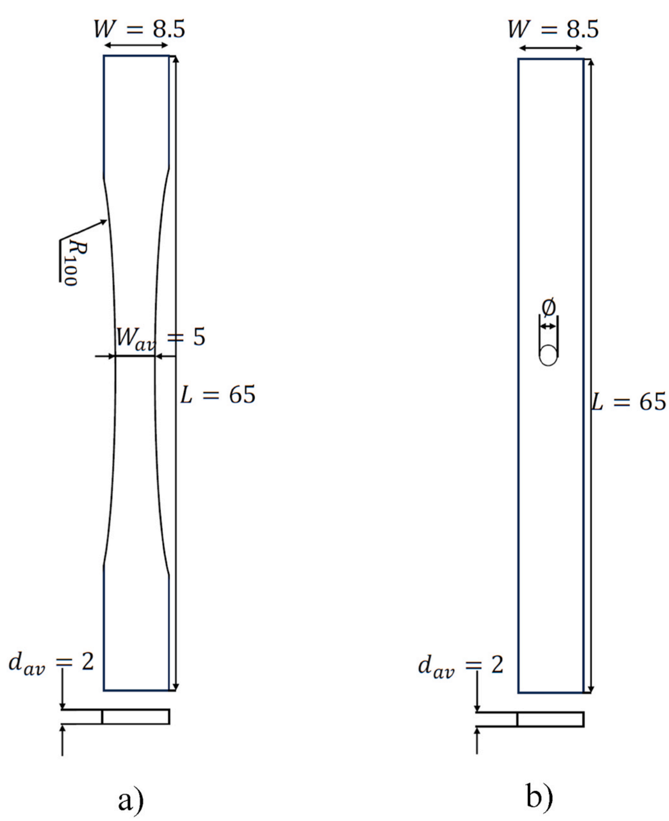


Fig. 2. Specimen geometry in mm a) plane, b) notched with a circular notch of diameter ∅

the mean curve. From these the endurance level $\sigma_{0,P\%}$, for a probability of survival P is defined as

$$\sigma_{0,P\%} = \left[\frac{10^{\left(C_0 - t\, s\sqrt{1+\frac{1}{n}}\right)}}{N_A}\right]^{\frac{1}{k}} \tag{6}$$

While the endurance limit in error at the reference cycle $\sigma_{0,(1-P)\%}$ is defined as

$$\sigma_{0,(1-P)\%} = \left[\frac{10^{\left(C_0 + t\, s\sqrt{1+\frac{1}{n}}\right)}}{N_A}\right]^{\frac{1}{k}} \tag{7}$$

From these therefore the scatter band is calculated as

$$\tau_\sigma = \frac{\sigma_{0,(1-P)\%}}{\sigma_{0,P\%}} \tag{8}$$

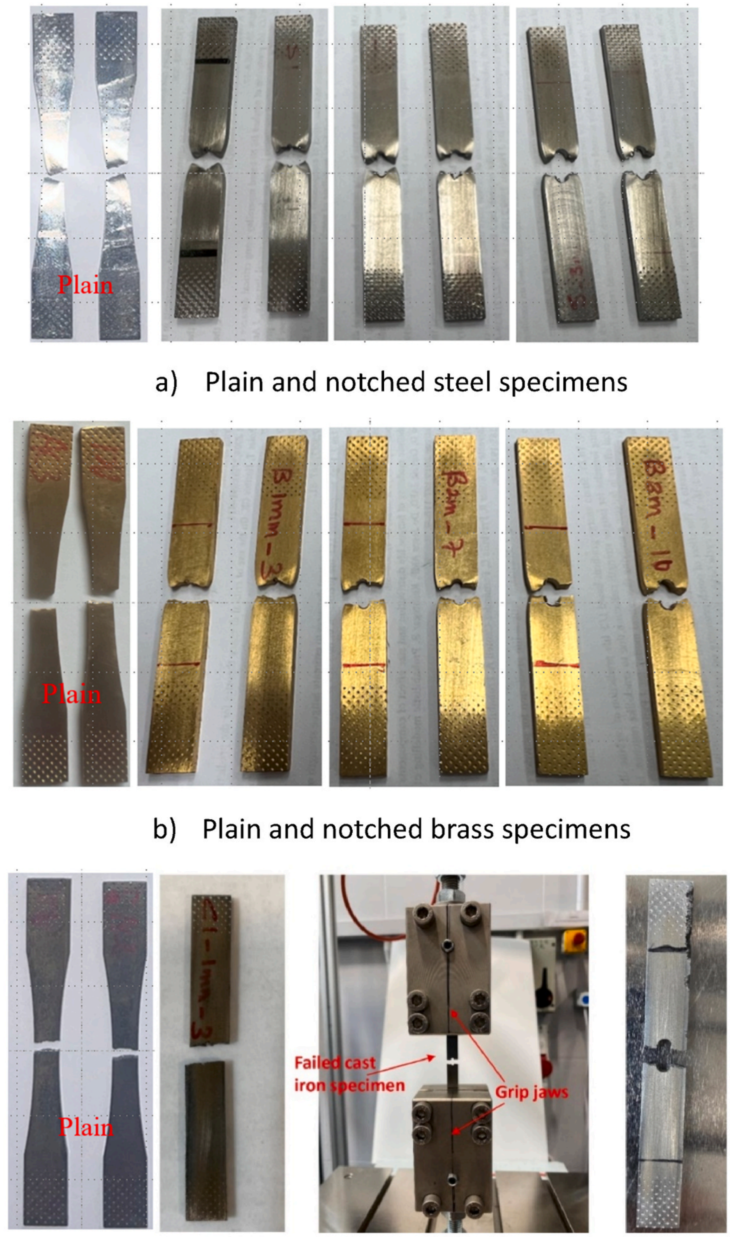
The endurance limit at a 50% probability of survival is also calculated as

$$\sigma_{0,50\%} = \left[\frac{10^{C_0}}{N_A}\right]^{\frac{1}{k}} \tag{9}$$

3. Brief overview of parametric and non-parametric analysis

Parametric and non-parametric analyses are two types of statistical procedures through which disparities can be probed between fatigue data sets [21,22]. In parametric analysis, a fixed number of parameters are considered for statistical tests [23,24]. Here, statistical tests are carried out based on some assumptions about the data sets [24]. It requires less data compared to non-parametric methods [25]. More so, parametric analysis assumed that the data has a normal distribution and this approach works best when the spread in the data set of each data set is different [26]. The parameters analysed in this approach include the variance, slope and intercepts. Meanwhile, non-parametric analysis tends to test medians. It is utilized based on fewer assumptions about data sets [27,28], and usually requires a large data set than parametric methods and has no assumed distribution to the data. Non-parametric methods can perform well in many situations but its performance is at peak when the spread of data in each group is the same. Nonparametric analysis uses the rank test for two or more groups to compare the medians [12]. Here ranks for two groups are totalled separately and the total for the smallest group should fall within the critical chi-squared lower and upper rank totals depending on the levels of significance.

In this paper, only the parametric significance test procedures will be explained. For simplicity, this procedure is used on portions of the S-N curve that are linear, the data sets also assumed to have a normal distribution and will generate S-N curves with near parallel or parallel



a) Plain and notched steel specimens

b) Plain and notched brass specimens

c) Plain cast iron specimen, and failed notched specimen attached to the grip jaws

Fig. 3. Some samples of failed specimens after fatigue tests.

characteristics. This ensures that comparisons are valid solely for stress level ranges pertinent to a given design or application. Having determined the type of analyses that will be discussed in this paper, it is therefore essential to proceed to explain the fundamental concept in statistical hypothesis testing of parametric tests, particularly the null hypothesis which plays a central role when comparing two or more data sets.

3.1. The t-test and null hypothesis meaning and application

A *t*-test is an inferential statistic used to determine if there is a significant difference between two measured observables. It compares the values of the measured quantities from two data sets and determines if they came from the same population. This comparison helps to determine the effect of chance on the difference, and whether the difference is outside that chance's range. T-tests are used when the data sets follow a normal distribution and have unknown variances. The *t*-test value is calculated as the ration of the difference between the measured quantities to the variation that exists in the sample data sets as shown below [12,20,29,30].

$$t = \frac{x_1 - x_2}{\sqrt{{s_{x_1}}^2 - {s_{x_2}}^2}} \quad (10)$$

Where x_1 and x_2 are any two observed quantities (such as mean, slope, etc.) with variances ${s_{x_1}}^2$ and ${s_{x_2}}^2$ respectively. Higher values of the t-score calculated using equation (10) indicate that a large difference exists between the two sample data sets. The smaller the t-value, the more similarity exists between the two sample sets. The value of the calculated test statistic is compared with a critical value t_β obtained from the critical value table known as the t-distribution table. This value

Table 1
Summary of fatigue properties at a probability of 95%, of notch specimens under uniaxial loading in tension at a frequency of 10Hz

Specimen type	Fatigue test results							
	Sample set, n	R	Inverse slope, k	Variance s^2	σ_0 [MPa]	$\sigma_{0,95\%}$ [MPa]	$\sigma_{0,5\%}$ [MPa]	τ_σ
S	17	0.1	50.9	0.09	224.5	231.5	217.6	1.1
S_1 mm	10	0.1	9.4	0.024	119.1	108.6	130.7	1.2
S_2 mm	10	0.1	8.4	0.023	96.1	86.9	106.2	1.2
S_3 mm	10	0.1	8.1	0.015	79.9	73.3	87.0	1.2
Br	5	0.1	40.1	1.17	140.6	177.7	111.2	1.6
Br_1 mm	10	0.1	8.6	0.012	72.6	67.6	77.9	1.2
Br_2 mm	10	0.1	6.9	0.011	56.3	51.7	61.4	1.2
Br_3 mm	10	0.1	6.1	0.015	46.7	41.9	52.1	1.2
CI	15	0.1	8.9	0.36	32.7	40.9	26.2	1.6
CI_1 mm	10	0.1	11.1	0.257	33.0	25.6	42.7	1.7
CI_2 mm	10	0.1	13.0	0.071	27.8	24.7	31.2	1.3
CI_3 mm	10	0.1	10.7	0.031	21.1	19.2	23.1	1.2

depends on the level of significance β, and the degrees of freedom f.

The t-distribution table can either be for a one-tail or two-tail formats. While one-tail values are used for assessing changes in data sets that have a fixed direction of change, two-tail values represent variations in more than one direction, which can either increase or decrease. For the analysis involved in this research, the two-tail distribution table will be use so that both positive and negative effects can be monitored on the measured observables.

Comparing the t-statistic with the critical value of t_β obtained from a two-tail t-distribution table, the null hypothesis is used to conclude as follows.

For:

$t \le t_\beta$, Null hypothesis accepted (no statistical significance from the two observables)
$t > t_\beta$, Alternate hypothesis or null hypothesis rejected.

What follows is a description of how parametric analysis is used in conjunction with the null hypothesis to establish statistical significance between fatigue data sets.

3.2. Statistical test on sample fatigue life data sets $(\boldsymbol{N}_{f,i})$ generated at the same stress level

Parametric analysis can be used on fatigue life data sets to ascertain whether sample sets exhibit statistical significance compared to the parent population for the fatigue lives that are generated at the same stress level. To use this approach, the fatigue lives are assumed to follow a normal distribution at this stress level. In this case, the parameters to be tested for significance are the variance and the mean. The test on variance is otherwise known as the test for homogeneity of variance and is assessed and described as below.

3.3. Statistical test on homogeneity of variances of two fatigue life data sets

Consider two normally distributed fatigue datasets at the same stress level with variances ${s_1}^2$ and ${s_2}^2$ with corresponding sample sets n_1 and n_2, respectively. The test statistic F_{cal}, for the homogeneity of the fatigue data sets is calculated as [12,29,31–33].

$$F_{cal} = \frac{{s_1}^2}{{s_2}^2} \tag{11}$$

In order to ascertain the presence of a significant difference at a designated level of significance (β), the critical value of the F_p distribution associated with $n_1 - 1$ degrees of freedom for s_1 and $n_2 - 1$ degrees of freedom for s_2, derived from statistical tables found in Refs. [17,20,30,32] is compared with the F-distribution value F_{cal}, calculated using equation (11). Should $F_{cal} \le F_p$ the sample variances are deemed not significantly different (homogeneous) in accordance with the null hypothesis. Conversely, if $F_{cal} > F_p$, the two datasets are considered significant.

3.4. Statistical test on two means when the standard deviations are not significantly different

Suppose two homogenous fatigue life data sets belong to the same population; in this scenario, the common estimate of the population variance ${s_e}^2$, is calculated as follows [29,34]:

$${s_e}^2 = \frac{(n_1 - 1){s_1}^2 + (n_2 - 1){s_2}^2}{n_1 + n_2 - 2} \tag{12}$$

in which n_1 and n_2 represent the sample sizes of the two fatigue life datasets, and s_1 and s_2 are their respective standard deviations. The test statistic for comparing their means t_μ, is calculated as [12,32]:

$$t_\mu = \frac{\overline{\log N_{f1}} - \overline{\log N_{f2}}}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \tag{13}$$

where $\overline{\log N_{f1}}$ and $\overline{\log N_{f2}}$ represent the mean values of the logarithm of fatigue life for the two fatigue life sample sets. Given a predetermined significance level β, the corresponding critical value t_β associated with degrees of freedom $f = n_1 + n_2 - 2$ is obtained from statistical tables as in Refs. [17,33,35,36]. If $/t_\mu/ > t_\beta$, it can be concluded that the populations from which the sample datasets are derived are distinct; otherwise, there is no statistically significant difference in the means of the sample sets.

3.5. Statistical test on two means when the standard deviations are significantly different

If the sample standard deviations can be verified to be significantly different, then the hypothesis that the populations' means are significant or not can be tested by calculating a test statistic t'_μ as [12,30]:

$$t'_\mu = \frac{\overline{\log N_{f1}} - \overline{\log N_{f2}}}{\sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}} \tag{14}$$

Similarly, for a predefined significance level β, the associated characteristic value t_β is determined based on a defined value of the degree of freedom defined as:

$$f = \left[\frac{c^2}{n_1 - 1} + \frac{(1 - c)^2}{n_2 - 1}\right]^{-1} \tag{15}$$

where c is a dimensionless quantity defined as [12]:

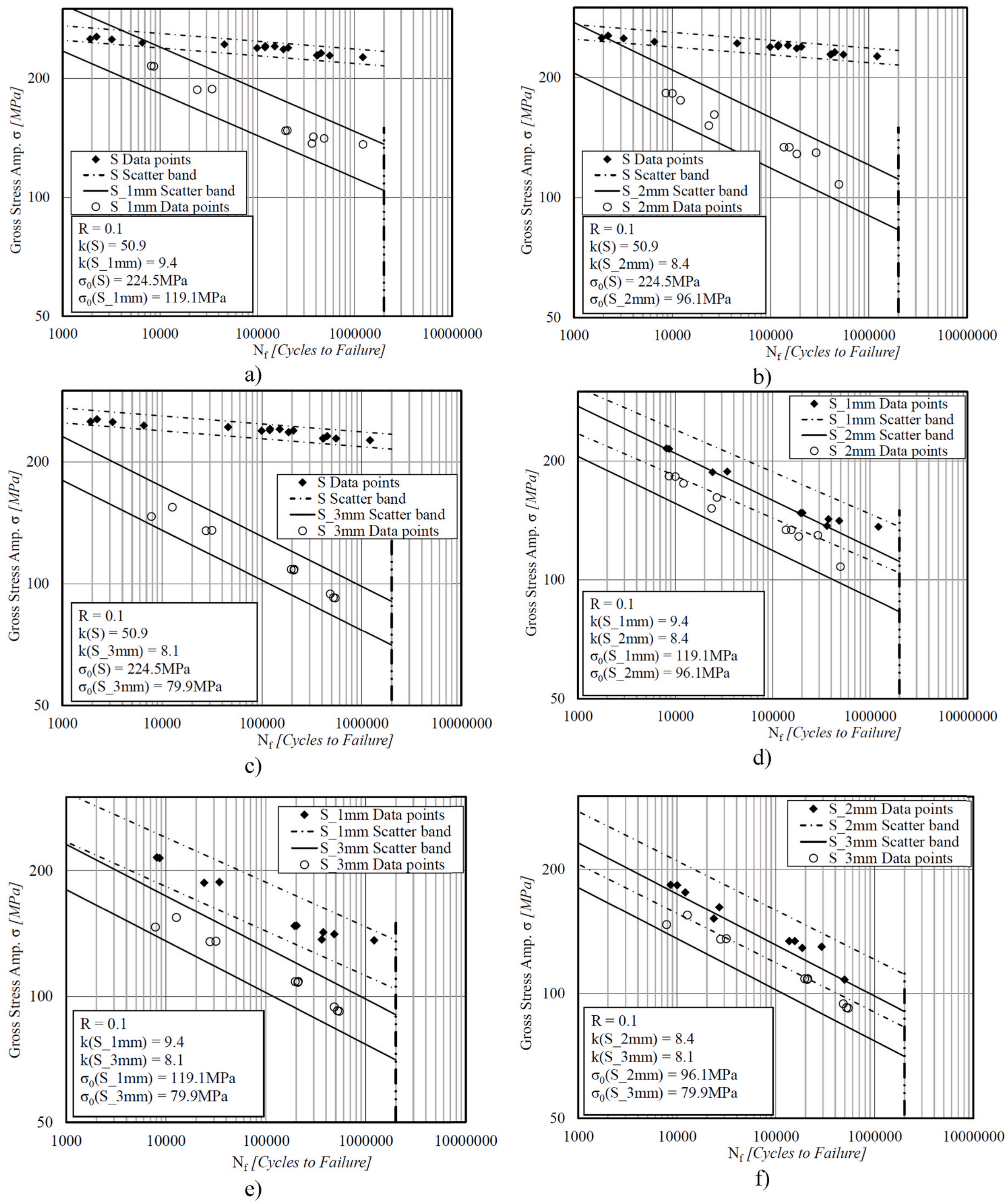


Fig. 4. Scatter bands at 95% level of confidence and 5% level of significance for: a) S and S_1 mm, b) S and S_2 mm c) S and S_3 mm d) S_1 mm and S_2 mm curves, e) S_1 mm and S_3 mm curves and f) S_2 mm and S_3 mm curves.

$$c = \frac{\frac{s_1^2}{n_1}}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \tag{16}$$

If the value for the calculated degree of freedom f is not an integer, then its value is approximated to the nearest smaller integer. The value of f is then used to extract the associated characteristic value t_β based on the predefined level of significance. In the same way, if $/ t'_\mu / < t_\beta$, then there is no significance in the means of fatigue lives of the two data sets and for $/t'_\mu/ > t_\beta$, the means are judged to be different.

The analysis outlined above is carried out on fatigue life datasets under the assumption that they have been collected at uniform stress levels and follow a normal distribution. When dealing with datasets acquired from a variety of stress levels, the parametric analysis is extended by considering the parameters within the mean curves generated from these datasets, as elaborated below.

4. Parametric analysis of the parameters of two mean curves

To statistically compare the significance of two mean S-N curves using the null hypothesis, three steps are employed [29,31]. The initial step entails testing whether the variance or standard deviations (s) around the distinct lines can be assumed to be drawn from the same

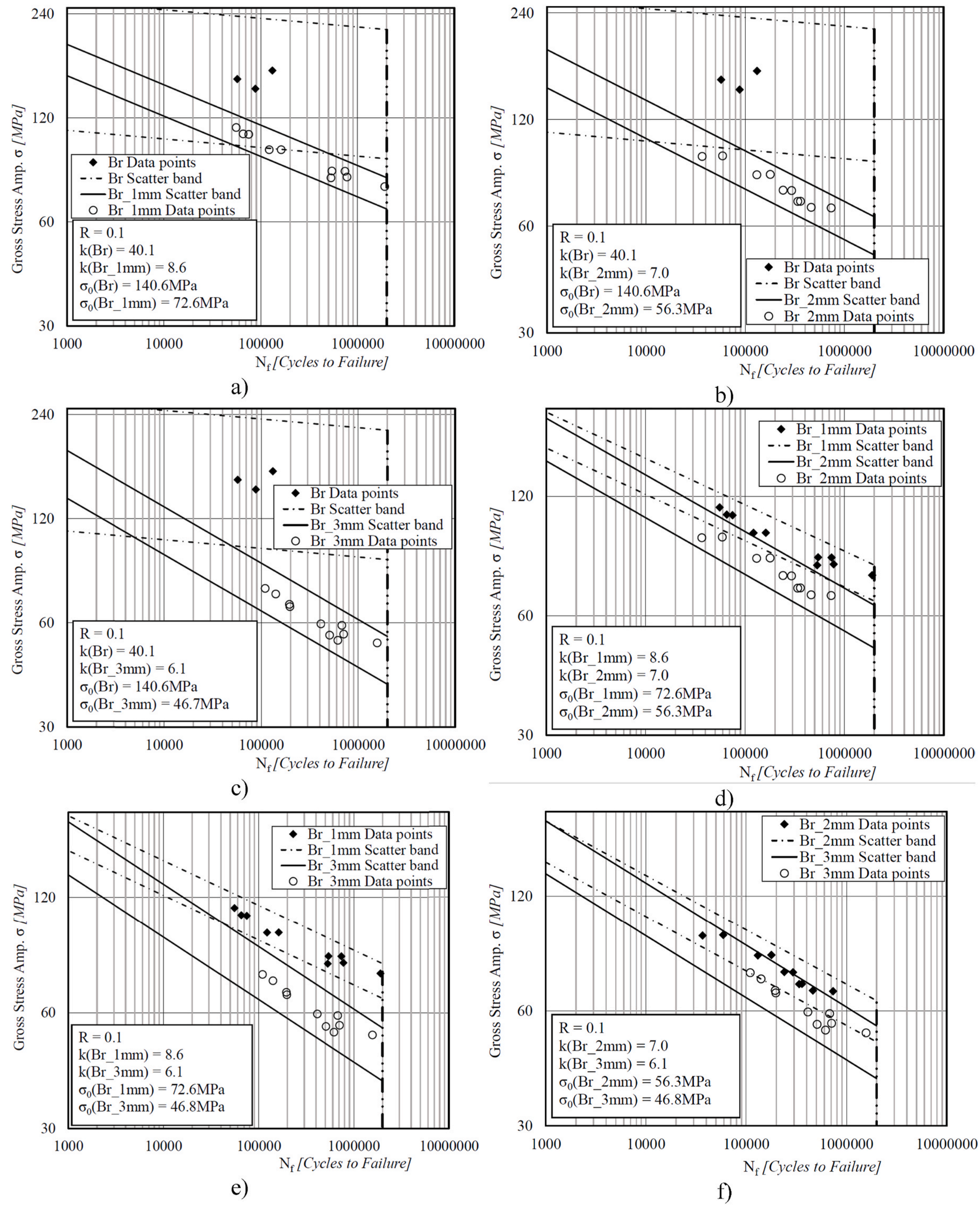


Fig. 5. Scatter bands at 95% level of confidence and 5% level of significance for: a) Br and Br_1 mm, b) Br and B_2 mm, c) Br and Br_3 mm, d) Br_1 mm and Br_2 mm, e) Br_1 mm and Br_3 mm and f) Br_2 mm and Br_3 mm curves.

population, thus demonstrating homogeneity. Secondly, the examination involves assessing whether the two mean curves can be viewed as parallel (approximated by the same inverse slope C_1). And lastly, determining if the two parallel regression lines can be considered collinear, meaning they lie on the same line [36].

4.1. Test that the variance of the data set is homogenous

Homogeneity of variance is used to describe a data set that has the same variance as another [20,30]. This equivalence can be visually identified through consistent scatter on a scatter plot or by observing equivalent standard deviations in the derived parameters like sample size, mean, slope and variance [29,37]. If, upon inspection, the data exhibits heteroscedasticity, a statistical hypothesis test is carried out by constructing a test for the homogeneity of variances between the fatigue data sets.

Consider for example two data sets 1 & 2 having sample sizes n_1 and n_2 respectively, with corresponding degrees of freedom $(n_1 - 2)$, and

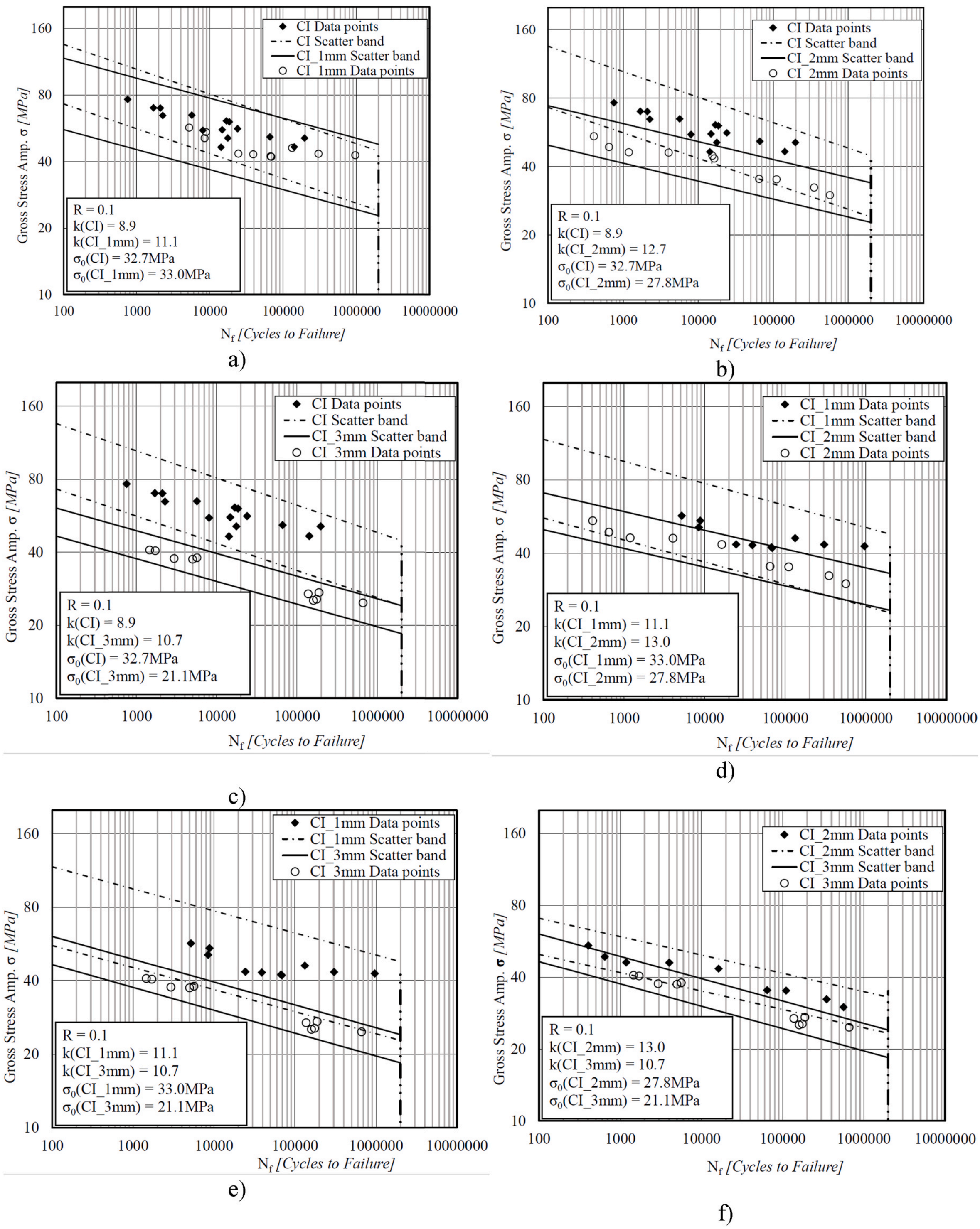


Fig. 6. Scatter bands at 95% level of confidence and 5% level of significance for: a) CI and CI_1 mm, b) CI and CI_2 mm, c) CI and CI_3 mm, d) CI_1 mm and CI_2 mm, e) CI_1 mm and CI_3 mm and f) CI-2mm and CI_3 mm.

$(n_2 - 2)$. Additionally, assume that ${s_1}^2$ and ${s_2}^2$ represent the variances of the respective data sets 1 and 2, with $s_1 > s_2$. The test statistic F_{cal} is calculated according to equation (11) [12,17,30,32]. By referring to standard F-distribution tables, the critical value entry (F_{crit}) corresponding to the degrees of freedom $f_1 = (n_1 - 2)$ and $f_2 = (n_2 - 2)$ is extracted and subsequently compared with the calculated value.

If it turns out that $F_{cal} \le F_{crit}$, it can be concluded that both variance estimates are homogeneous and can be considered as independent estimators of the population. Consequently, the null hypothesis is validated, signifying that the difference in variance of the two mean curves is not significant. In this situation, the variance estimate ${s_e}^2$ defining both data sets is defined as:

$$ {s_e}^2 = \frac{(n_1 - 1){s_1}^2 + (n_2 - 1){s_2}^2}{(n_1 - 2) + (n_2 - 2)} \tag{17} $$

To summarize, the two sample variances are considered significant if the equation below is validated.

Table 2
Summary of statistical analysis using the null hypothesis for the fatigue data sets generated by testing brass, cast iron and steel specimens with 1 mm, 2 mm and 3 mm notches.

Source of data sets compared		Variance test ($\beta = 1.7\%$)	Slope test ($\beta = 1.7\%$)	Intercept test ($\beta = 1.7\%$)	Conclusion ($\sum\beta \cong 5\%$)
Steel	S and S_1 mm	−0.76	**31.94**	**1.40**	Significant
	S and S_2 mm	−0.53	**32.89**	**2.01**	Significant
	S and S_3 mm	**1.44**	**33.44**	**2.58**	Significant
	S_1 mm and S_2 mm	−4.00	−1.47	**0.48**	Significant
	S_1 mm and S_3 mm	−3.50	−0.88	**1.09**	Significant
	S_2 mm and S_3 mm	−3.58	−1.87	**0.41**	Significant
Brass	Br and Br_1 mm	**106.83**	**10.99**	−0.52	Significant
	Br and Br_2 mm	**112.92**	**12.05**	−0.103	Significant
	Br and Br_3 mm	**88.24**	**12.97**	−0.06	Significant
	B_1 mm and B_2 mm	−4.01	−0.77	**0.52**	Significant
	B_1 mm and B_3 mm	−3.87	−0.03	**0.90**	Significant
	B_2 mm and B_3 mm	−3.80	−1.69	**0.23**	Significant
Cast Iron	CI and CI_1 mm	−3.00	−6.78	−0.27	Insignificant
	CI and CI_2 mm	−3.22	−1.49	**1.01**	Significant
	CI and CI_3 mm	**0.10**	−2.72	**1.61**	Significant
	CI_1 mm and CI_2 mm	−2.04	−6.75	**0.69**	Significant
	CI_1 mm and CI_3 mm	**3.30**	−7.36	**1.36**	Significant
	CI_2 mm and CI_3 mm	−2.87	−0.86	**0.69**	Significant

Table 3
Summary of statistical analysis using the null hypothesis for the fatigue data sets generated by testing brass, cast iron and steel specimens with 1 mm, 2 mm and 3 mm notches around their mean stresses.

Source of data sets compared		Variance test ($\beta = 1.7\%$)	Slope test ($\beta = 1.7\%$)	Intercept test ($\beta = 1.7\%$)	Conclusion ($\sum\beta \cong 5\%$)
Steel	S_1 mm and S_2 mm	−4.00	−1.47	**−2.55**	Insignificant
	S_1 mm and S_3 mm	−3.50	−0.88	**−0.56**	Insignificant
	S_2 mm and S_3 mm	−3.58	−1.87	**−3.35**	Insignificant
Brass	B_1 mm and B_2 mm	−4.01	−0.77	**−0.82**	Insignificant
	B_1 mm and B_3 mm	−3.87	−0.03	**1.09**	Significant
	B_2 mm and B_3 mm	−3.80	−1.69	**−2.69**	Insignificant
Cast Iron	CI_1 mm and CI_2 mm	−2.04	−6.75	**−12.40**	Insignificant
	CI_1 mm and CI_3 mm	**3.30**	−7.36	**−9.97**	Significant
	CI_2 mm and CI_3 mm	−2.87	−0.86	**−0.34**	Insignificant

Table 4
Fatigue data from Ref. [41].

As-forged cantilever		As-forged bending	
σ (MPa)	N(cycles)	σ (MPa)	N(cycles)
892	5951	546	43586
607	23826	398	94495
596	39661	396	110938
551	43176	273	698654
396	144489	274	820228
396	195427	227	851781
324	328393	227	981303
324	367768		
274	753441		
248	623852		
249	760585		
229	1327244		

$$\frac{s_n^2}{s_d^2} - F_{crit}(1-\beta, f_n, f_d) < 0 \quad (18)$$

In this equation, s_n^2 and s_d^2 represent the variances of the numerator and denominator of equation (11) respectively. Additionally, f_n and f_d correspond to the degrees of freedom of the numerator and the denominator, respectively. And $F_{crit}(1-\beta, f_n, f_d)$ denotes the critical value obtained from entries corresponding to the significance level β, and the respective degrees of freedom.

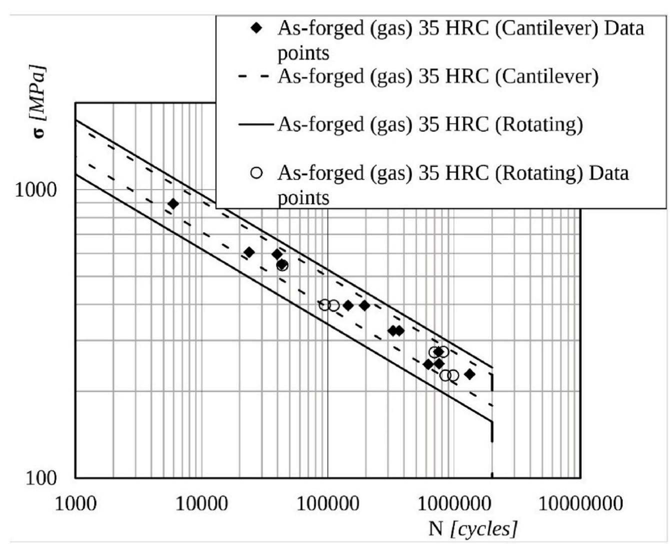


Fig. 7. Scatter band at 95% level of confidence and 5% level of significance for significant data set of as-forged (cantilever) and as-forged (bending) in Ref. [41].

Table 5
Summary of parametric statistical analysis of fatigue data set from Ref. [41].

Source of data sets compared	Variance test ($\beta = 1.7\%$)	Slope test ($\beta = 1.7\%$)	Intercept test ($\beta = 1.7\%$)	Conclusion ($\sum\beta \cong 5\%$)
Considering the entire S-N curve with extrapolationd				
Forged bending-cantilever	−5.64	−0.89	−0.07	Insignificant
Stress levels in the vicinity of the mean stress				
Forged bending-cantilever	−5.64	−0.89	−2.29	Insignificant

4.2. Test that the lines are parallel

Let $C_{1,1}$ and $C_{1,2}$ represent the slopes of two data sets 1 & 2 respectively. The t-test statistic for the significance of the slopes t_{C_1} is calculated as [30,32]:

$$t_{C_1} = \frac{\left|C_{1,1} - C_{1,2}\right|}{\sqrt{\left(\frac{1}{\sum_{i=1}^{n_1}\left(\log\sigma_{i,1} - \overline{\log\sigma_{i,1}}\right)^2} + \frac{1}{\sum_{i=1}^{n_2}\left(\log\sigma_{i,2} - \overline{\log\sigma_{i,2}}\right)^2}\right)s_e}} \quad (19)$$

This value of t_{C_1} is then compared with the student's distribution value t_β obtained from statistical tables aligning with degrees of freedom $f_1 + f_2 = (n_1 - 2) + (n_2 - 2)$ and the chosen significance level. Should the value of t_{C_1} be lower than this critical value, the slopes are deemed not statistically significant. Conversely, if t_{C_1} is more than the critical value, there is no bases to accept the null hypothesis.

In the case of statistical insignificance, the mean curves are considered to be parallel. Hence, a common estimate for the slopes C_1 is determined. This estimation takes the form of the weighted average of both slopes $C_{1,1}$ and $C_{1,2}$, and can be calculated as [29]:

$$C_1 = \frac{\sum_{i=1}^{n_1}\left(\log\sigma_{i,1} - \overline{\log\sigma_{i,1}}\right)^2 \times C_{1,1} + \sum_{i=1}^{n_2}\left(\log\sigma_{i,2} - \overline{\log\sigma_{i,2}}\right)^2 \times C_{1,2}}{\sum_{i=1}^{n_1}\left(\log\sigma_{i,1} - \overline{\log\sigma_{i,1}}\right)^2 + \sum_{i=1}^{n_2}\left(\log\sigma_{i,2} - \overline{\log\sigma_{i,2}}\right)^2} \quad (20)$$

The resulting estimate of the variance of C_1 denoted as ${s_{C_1}}^2$ is such

Table 6
Summary of fatigue test results under uniaxial loading in tension (R = 0.1) at a frequency of 10Hz

Material type	Fatigue test results							
	1 mm notch		2 mm notch		3 mm notch		Plain	
	σ_0 [MPa]	N_f [Cycles]	σ_0 [MPa]	N_f [Cycles]	σ_0 [MPa]	N_f [Cycles]	σ_0 [MPa]	N_f [Cycles]
Steel	215.2	8103	183.0	8606	154.7	12684	75.5	109362
	214.6	8643	182.7	10000	146.7	7826	72.7	140561
	188.0	34155	175.5	12105	135.6	31457	67.8	194403
	187.2	24110	161.6	26734	135.3	27529	66.9	196858
	147.8	203115	151.7	23547	108.6	210278	59.6	410938
	147.5	193751	133.8	155536	108.6	196032	59.0	676739
	142.4	374787	133.8	136806	108.1	208275	55.7	706821
	140.9	483252	129.7	291168	94.4	481826	55.3	507491
	136.9	362177	128.8	185460	92.5	519988	53.4	616481
	136.2	1215396	107.9	496775	92.2	540895	52.5	1568334
							75.5	109362
							72.7	140561
							67.8	194403
							66.9	196858
							59.6	410938
							59.0	676739
							55.7	706821
Cast iron	56.9	5176	54.4	406	40.8	1455	76.4	751
	54.4	8798	48.8	645	40.5	1708	70.0	1690
	51.1	8442	46.2	1183	37.9	5641	69.8	2101
	46.1	132604	46.1	4025	37.6	2933	64.8	5651
	43.4	24487	44.7	15354	37.4	4988	64.6	2264
	43.4	303117	43.5	16266	27.2	187383	61.0	16839
	43.2	39080	35.3	64586	26.9	137409	60.5	18501
	42.8	963511	35.2	110298	25.6	174677	56.3	23930
	42.3	67485	32.3	347842	25.3	159794	55.7	14745
	42.1	68918	30.0	563403	24.7	660480	55.4	7987
							51.7	66103
							51.1	17621
							51.0	197025
							46.6	142029
							46.4	14235
Brass	112.6	55514	94.8	59018	75.5	109362	192.2	8
	107.9	65518	94.4	36739	72.7	140561	178.1	13
	107.5	74827	84.0	178161	67.8	194403	164.4	131262
	97.1	161869	83.8	130672	66.9	196858	155.4	57180
	97.0	121909	75.8	240122	59.6	410938	145.7	87888
	84.2	541945	75.7	292589	59.0	676739		
	84.1	739815	70.5	359750	55.7	706821		
	80.9	776713	70.5	337359	55.3	507491		
	80.5	530996	67.8	460968	53.4	616481		
	75.9	1903387	67.5	732233	52.5	1568334		

that the combined estimate of both variances and the error in the estimation of C_1 has one degree of freedom and is calculated as:

$$s_{C_1}{}^2 = \frac{\left(C_{1,1} - C_{1,2}\right)^2}{\frac{1}{\sum_{i=1}^{n_1}\left(\log \sigma_{i,1} - \overline{\log \sigma_{i,1}}\right)^2} + \frac{1}{\sum_{i=1}^{n_2}\left(\log \sigma_{i,2} - \overline{\log \sigma_{i,2}}\right)^2}} \quad (21)$$

However, $s_{C_1}{}^2$ is always small [29] when the two lines are considered parallel. Alternatively, the estimated common variance $s_e{}^2$ is defined as in equation (17) and according to the null hypothesis, the individual slopes are significant if $t_{C_1} > t_\beta$ in equation (19).

4.3. Test that the parallel lines are collinear

The two mean curves from the fatigue data sets are collinear if both have similar intercept and slope. Suppose the expected values for both curves are defined as $\mu_{N_{f,1}/\sigma_1} = C_0 + C_1 \log \sigma_1$ and $\mu_{N_{f,2}/\sigma_2} = C_0 + C_1 \log \sigma_2$ for any stress level σ, with both intercept and slope constant. Then at the mean point, there is a random variable δ defined as:

$$\left(\overline{\log N_{1,i}} - \overline{\log N_{2,i}}\right) - C_1\left(\overline{\log \sigma_{1,i}} - \overline{\log \sigma_{2,i}}\right) = \delta \quad (22)$$

δ is normally distributed with mean equal to zero and variance $Var(\delta)$ defined by Ref. [28].

$$Var(\delta) = s^2\left[\frac{1}{n_1} + \frac{1}{n_2} + \frac{\left(\overline{\log \sigma_{1,i}} - \overline{\log \sigma_{2,i}}\right)^2}{\sum_{i=1}^{n_1}\left(\log \sigma_{i,1} - \overline{\log \sigma_{i,1}}\right)^2 + \sum_{i=1}^{n_2}\left(\log \sigma_{i,2} - \overline{\log \sigma_{i,2}}\right)^2}\right] \quad (23)$$

Then a test statistic $t(\delta)$ for the random variable δ can be calculated as shown in equation (15) where s_e is the new-pooled estimate of the combined variance corresponding to $[(n_1 - 2) + (n_2 - 2) + 1]$ degrees of freedom.

$$t(\delta) = \frac{C_{0,1} - C_{0,2}}{\sqrt{\left[\frac{1}{n_1} + \frac{1}{n_2} + \frac{\left(\overline{\log \sigma_{1,i}} - \overline{\log \sigma_{2,i}}\right)^2}{\sum_{i=1}^{n_1}\left(\log \sigma_{i,1} - \overline{\log \sigma_{i,1}}\right)^2 + \sum_{i=1}^{n_2}\left(\log \sigma_{i,2} - \overline{\log \sigma_{i,2}}\right)^2}\right] s_e}} \quad (24)$$

Hence, the two mean curves are considered not to be collinear if the test statistic $t(\delta)$ is significant. i.e. $/t(\delta)/ > t_\beta$ where t_β is the characteristic value associated with the predefined significance level and degrees of freedom. If $/t(\delta)/ \le t_\beta$, then the null hypothesis confirms that the two curves are collinear [29].

In conclusion, when considering the null hypothesis, two mean curves from two fatigue data sets are considered insignificant if the variances, gradients, and intercepts are found to be statistically insignificant. If there is no substantial evidence to support the acceptance of significance for any of these parameters, then, in accordance with the criteria of the null hypothesis, there is no foundation to assert that the data sets are insignificant.

4.4. Composite hypotheses and its significance level

It is important to highlight that in composite hypothesis testing, the null hypothesis is evaluated separately for each parameter of variance, intercept, and slope, all at a specified significance level. The cumulative level of significance is obtained by summing all the individual significance levels when each parameter is tested as illustrated in the flow chart in Fig. 1. This can potentially lead to a high chance of rejecting these hypotheses even if all three parameters are accurate. As a result, it is advisable to employ a lower significance level for each individual parameter test when conducting composite hypothesis testing [12,20,29]. For instance, a significance level of 1.7% could be used, resulting in an approximate combined level of significance of about 5%. Under a similar condition, when testing for the consistency in sample variance and means, a significance level of 2.5% would be appropriate. This approach helps reduce the increased risk of erroneous rejections associated with composite hypothesis testing. In general, the level of significance of each of the parameters tested is determined by dividing the significance level used to test the null hypothesis by the number of parameters considered [29–33].

The composite hypothesis is deemed acceptable only when all the test statistics pertaining to the parameters of variance, intercept, and slope are insignificant. If any of these parameters lack sufficient evidence to support the acceptance of the null hypothesis, it follows that there is no valid ground to accept the significance of two S-N curves from the data sets.

5. Experimental procedure and results

To evaluate the accuracy and reliability of the mentioned test statistics, experimental fatigue data were generated. These datasets were derived from both plane specimens and specimens altered by notches of different geometries with varied sharpness. The geometries ranged from a 0 mm diameter hole notch (plain specimens) to a 3 mm diameter hole notch on the chosen materials.

5.1. Materials, specimen geometry and experimental testing

The materials chosen for this analyses were ex-service pipeline materials of cast iron ($348HV_{0.2}$), brass ($111HV_{0.2}$), and X52 carbon steel ($184HV_{0.2}$), where $HV_{0.2}$ denotes the Vickers Hardness using an applied load of 2 N. The specimens underwent surface preparation to make them smooth by manual machining. For plane specimens, the dog-bone design was used as shown in Fig. 2a. The dimensions consisted of a length of 65 mm, a thickness of 2 mm, and a width of 5 mm. The thickness measurement was taken at different points along the length of each specimen and the results averaged. For notched specimens, the notches were created from rectangular strips using drills with varying diameters (1–3 mm) to produce circular holes at the centre of the specimens. The dimensions of the specimens measured on average a length of 65 mm, 8.5mmm in width and had a cross sectional thickness of 2 mm as shown in Fig. 2b.

The fatigue machine used to generate the data was the Multipurpose Servo Hydraulic Testing machine, the LFV-L series, with a static load test capacity of up to 25 kN, and a recommended fatigue testing capacity of 20 kN. Each specimen was loaded into the test machine in tension ($R = 0.1$) at a chosen stress amplitude and a frequency of 10Hz until the specimen failed by complete breakage or survived 2×10^6 cycles (runout). For each of the runout results, the specimen was retested at a higher stress amplitude.

The failure criterion was defined as the complete breakage of specimens occurring at the critical region with a small net area, in which the crack propagation part of the total life was considered to be negligible. Some of the failed specimens are shown in Fig. 3.

5.2. Experimental results and application of parametric test analysis

Table 1 summarises the fatigue results generated from testing the notched specimens of steel, cast iron and brass with different notch radii. The experimental data used to generate these fatigue results are also summarized in Table 6. Considering any two data sets for each material type, the scatter bands and data points are illustrated on respective graphs to visually depict the significance of the data. The chosen approach uses gross stress amplitudes with fatigue life. The utilization of gross stress amplitudes is intended to minimize the impact of notch geometry variations within the compared data sets. The primary objective is to assess the effectiveness of using statistical approaches in identifying the presence of different notch geometries in the datasets.

5.3. Application of parametric analyses on generated data sets

It is assumed that the data sets from each group of specimens/material are normally distributed at each stress level and are statistically independent of each other. Additionally, the variance at each stress level is constant, and therefore, the data sets follow the same form of the S-N curve and have the same residual standard deviations. Linear regression is applied to generate these mean S-N curves. Figs. 4–6 provide a summary of the S-N curves and scatter bands for steel, brass, and cast-iron specimens generated in this investigation, with a 95% confidence level and a significance level of 5%. These curves also present the estimated endurance limits after 2×10^6 cycles, along with the values of their inverse slopes. The parameters of these mean curves are then utilized to assess statistical significance of the data sets. The significance test-statistics for variance, slope, and intercept, calculated using gross stress amplitudes for all materials, are summarized in Table 2 by employing equations (18), (19) and (24). In this table, a negative entry indicates that the calculated test statistic is smaller than the critical value or the p-value corresponding to the chosen level of significance.

As depicted in Table 1, the anticipated notch effect was evident in the fatigue data sets [6,7], with the endurance limit decreasing upon the introduction of the notch. Statistical significance is observed in the characteristic test value retention when comparing each notch with the plain fatigue data, notably due to a change in slope. Upon examination of the variances of the S_1 mm and S_2 mm data sets, along with their inverse slopes, it becomes apparent that the two data sets are statistically insignificant when considered together. However, the mean curves representing these data sets do not lie on the same plane and are not collinear. Therefore, it is concluded, with a 95% confidence level, that these data sets cannot be represented by the same line, indicating statistical significance. Similarly, S_1 mm and S_3 mm data sets exhibit statistically insignificance in variance and inverse slope. However, their mean curves are not collinear, leading to the conclusion that these two data sets are also statistically significant. The same pattern is observed with the S_2 mm and S_3 mm data sets, confirming their statistical significance. Given that these fatigue data sets were generated by testing specimens with different notch radii, the *t*-test statistical analysis proves capable of detecting changes in the geometry of these notches for steel specimens.

Considering the brass specimens, Fig. 5 provides a summary of the data sets at a 95% confidence level and a 95% probability of survival. Similar to the steel specimens, the data sets in Table 2 show that the fatigue properties change with the introduction of notches and are deemed statistically significant with changes in the slope being more statistically significant. Comparing the notched specimens amongst themselves also shows statistical significance due to the absence of collinearity between each pair of data. This lack of collinearity is evident from the positive characteristic test values, indicating that the difference in intercepts for each curve exceeds the characteristic value at the chosen probability. Therefore, the statistical test analysis effectively detects changes in geometry within the brass fatigue data sets.

For cast iron specimens, the statistical significance test does not indicate any changes upon the introduction of a 1 mm notch geometry. This finding aligns with the explanation provided in Refs. [38,39] regarding the low sensitivity of cast iron to notches. Additionally, statistical significance is only observed in the notched specimens due to collinearity, as will be further elucidated. Fig. 6 illustrates the scatter bands for any two data sets for cast iron at a 95% confidence level and a 95% probability of survival, utilizing gross stresses. The results of the statistical test analyses for significance are summarized in Table 2.

As observed from the test statistic values for cast iron, the data sets from cast iron specimens with 1 mm and 2 mm notches, and using gross stresses, the spread in these data sets suggest that they are insignificant and drawn from the same population. The inverse slopes of these sets are also statistically insignificant. However, the mean curves representing these datasets are not collinear, indicating statistical significance. Similarly, for data sets from specimens with 1 mm and 3 mm notches the spread shows that they do not belong to the same population, even though the inverse slopes from these datasets are not statistically significant. In addition, the mean curves representing the data sets are not collinear. Therefore, it is reasonable to conclude that these datasets are indeed significant. In addition, specimens with 2 mm and 3 mm notches are also statistically significant because the mean curves representing both sets are not collinear. Thus, it is concluded that, by using gross stresses and statistical approaches, the data sets from cast iron are all significant. This approach has been able to verify that notches with varied geometry will have an impact on the fatigue property of cast iron even when some types of cast iron show less sensitivity to small notch radii as stated in Ref. [40].

The analysis was conducted by examining the mean curves derived from the compared data sets. If the stress levels, considered in generating fatigue data during testing, are statistically insignificant, their mean stress levels should also be statistically insignificant. Ref. [30] recommends performing a check for statistical significance only for stress levels around the mean stress levels during testing. In this scenario, the variance and slope test statistics are calculated using equations (18) and (19) respectively, while equation (24) is adjusted for checking collinearity, as follows:

$$t(\delta) = \frac{\left|C_{0,1} - C_{0,2}\right|}{\sqrt{\left[\frac{1}{n_1} + \frac{1}{n_2} + \frac{\left(\overline{\log \sigma_{1,i}}\right)^2}{\sum_{i=1}^{n_1}\left(\log \sigma_{i,1} - \overline{\log \sigma_{i,1}}\right)^2 +} + \frac{\left(\overline{\log \sigma_{2,i}}\right)^2}{\sum_{i=1}^{n_2}\left(\log \sigma_{i,2} - \overline{\log \sigma_{i,2}}\right)^2}\right] s_e}} \quad (25)$$

In which $C_{0,1}$ and $C_{0,2}$ are the estimated intercepts of the regression lines through the two data sets and s_e is the estimate of the common variance of the two data sets defined by equation (17). By applying this to the experimental data, Table 3 summarises the significance test conclusions arrived at. The results in this table show that restricting the stress range to a narrow region around the mean stress field for the statistical significance tests, the endurance limit must only change by approximately 36% (see Table 1) to indicate a potential statistical significance within the fatigue data sets. Therefore, restricting the statistical significance analysis to only sections of the data sets will result in conclusions about significance that are not consistent with the fatigue properties revealed by the entire data set. One explanation for this is the K_t which is around 3 in this case. In the medium/low cycle fatigue regime, the elastic peak stress is larger than the yield stress.

5.4. Application of parametric test analysis of data sets from literature (insignificance)

Notch parameters have a significant impact on the fatigue strength of materials. However, what if parameters that have a lesser influence on the material's fatigue strength are considered? In such cases there might be minimal change in the fatigue strength when these parameters are utilized. To investigate this scenario, we analysed data from the literature to highlight situations where fatigue data sets become insignificant. Consider the data in Ref. [41], in which the data sets were generated by testing plain specimens with two types of surface finishes: a smooth-polished surface finish and a hot-forged surface finish with different levels of hardness. These data sets were used to evaluate and quantify forged surface finish effect at several hardness levels. It should be noted that this project is only interested in using the data sets produced in this reference rather than reviewing the research reported therein.

The data set considered in this case is shown in Ref. [41] of this reference for as-forged in bending and as-forged in cantilever testing. This data is summarized in Table 4, while the overlapping scatter bands, from the S-N curves plotted at confidence level of 95% with a probability of survival of 95%, derived from this data set is as shown in Fig. 7.

The analysis for this pair of data sets is summarized in Table 5 and it clearly shows that the difference in the data set is statistically insignificant at a level of 5%. These two data sets can be seen as drawn from the same population, despite the difference in the endurance limits extrapolated at 2 million cycles. In this example, the endurance limits differ by about 3%. Because the data sets are insignificant, it can be concluded that there is no difference in the fatigue behaviour of the forged surface finish in rotating bending and cantilever bending.

It can be observed that if there is a case of statistical insignificance observed when evaluating both the entire S-N curves, the same holds true when examining stress levels around the mean stress levels considered during testing for both data sets.

6. Discussion

In this investigation it has been established that statistical tests can be used to compare different fatigue data sets by assessing the significance of differences and test hypotheses under varying conditions. By using gross stresses, it has been inferred that no geometrical information is attributed a priori to each data set. The results have shown that the fatigue data sets are statistically significant and thus have been generated from specimens with different geometries. Indeed, the data sets have been generated from specimens with different notch dimensions, root radius.

Thus, parametric analysis employs a well-established statistical method that provide robust conclusions based on the level of significance. This greatly helps in making informed decisions between about fatigue data sets for design purposes. It offers an objective way to assess the significance of differences in fatigue properties by reducing subjective biases in interpretation of results. It is data driven, consistent, and efficiently handles all the parameters that define the mean curve and further offers a more comprehensive assessment of the fatigue behaviour as opposed to visual assessments. When utilized accurately, parametric analysis helps minimize errors that might arise from misinterpreting data or drawing conclusions solely based on visual observations. Furthermore, this analysis has demonstrated a comprehensive approach that can be used to determine the feasibility of merging fatigue datasets from various sources for enhanced reliability. This approach aims to minimize the time and cost associated with conducting fatigue experiments, as well as comparing the impact of specific characteristics on the mechanical properties of a material. Some of these characteristics may include surface finish, environmental impacts or conditions, curing, machining, etc. This is important because more data increases the level of reliability and fatigue experiments are costly and time consuming. Using data sets from a variety of sources will reduce the cost in time and money. More so, in the case of fatigue data sets used in the very high cycle fatigue regime, where a substantial amount of data are needed to establish distributions at each stress level, this process becomes especially valuable. In addition, this statistical parametric analysis approach of fatigue data sets will assist in determining whether the differences observed in fatigue behaviour of are due to chance, material properties, or the conditions in which the data were collected. This approach aids in understanding what causes variations in fatigue behaviour and whether they have meaningful reasons behind them.

However, when dealing with ferrous metals lacking a fatigue limit, the comparison of data sets is influenced by the stress range considered in generating the data sets. Stress levels significantly distant from the mean stress are more susceptible to the collinearity condition. Restricting the analysis to the vicinity of the mean stress can result in different conclusions regarding significance levels compared to situations where stress levels far from the mean stress are involved.

7. Conclusion

This paper has clarified and substantiated the application of parametric statistical analysis for identifying disparities in specimen geometry. The method proves effective in detecting even small geometric alterations through the analysis of data set behaviour. However, a notable challenge lies in determining the significance of changes in geometry, particularly when monitored using significance levels. Nevertheless, this approach remains valuable as long as a clearly defined significance level is employed.

It is concluded that:

- Parametric analysis can detect changes of 1 mm in the geometry of specimens used to generate data sets.
- The mean points of fatigue data sets influence statistical parametric analysis. Parametric analysis for two or more mean S-N curves is very effective for stress ranges near the mean stress. Thus, using the analysis for points extrapolated far off from the mean stress considered during testing can lead to misleading or erroneous conclusions.
- Collinearity stands out as the most influential factor in establishing statistical independence. When S-N curves exhibit collinearity, it indicates that the precision of the estimated coefficients in the common S-N curve diminishes due to inflation in the variance and standard error of the coefficient estimates. This reduction in precision weakens the statistical power of the final S-N curve, thereby decreasing its reliability. Although this issue can be addressed by considering variance inflation factors, it is not applicable in this context since the various parameters determining changes in the data sets cannot be determined a priori. Therefore, it falls beyond the scope of this paper.

CRediT authorship contribution statement

Elvis N. Kufoin: Writing – original draft, Visualization, Methodology, Formal analysis, Data curation. **Luca Susmel:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] J. Jia, Fatigue assessment, in: Essentials Appl. Dyn. Anal. RIsk Eng., Springer-Verlag Berlin Heidelberg, Berlin, 2014, pp. 321–372, https://doi.org/10.1007/978-3-642-37003-8_17.

[2] L. Zhang, B. Jiang, P. Zhang, H. Yan, X. Xu, R. Liu, J. Tang, C. Ren, Methods for fatigue-life estimation: a review of the current status and future trends, Nanotechnol. Precis. Eng. 6 (2023) 025001, https://doi.org/10.1063/10.0017255.

[3] Y. Feng, C. Gao, Y. He, T. An, C. Fan, H. Zhang, Investigation on tension-tension fatigue performances and reliability fatigue life of T700/MTM46 composite laminates, Compos. Struct. 136 (2016) 64–74, https://doi.org/10.1016/j.compstruct.2015.09.057.

[4] F. Campbell, Fatigue and Fracture, Understanding the Basics, ASM International, Materials Park, Ohio, 2012.

[5] W. Feng, F. Xu, W. Xie, M. Zhang, Fatigue performances and reliability analysis of fatigue life for scarf-repaired composite laminates, J. Adhes. 98 (2022) 367–389, https://doi.org/10.1080/00218464.2020.1834389.

[6] Y. Murakami, T. Takagi, K. Wada, H. Matsunaga, Essential structure of S-N curve: prediction of fatigue life and fatigue limit of defective materials and nature of scatter, Int. J. Fatig. 146 (2021) 106138, https://doi.org/10.1016/j.ijfatigue.2020.106138.

[7] J. Mei, S. Xing, A. Vasu, J. Chung, R. Desai, P. Dong, The fatigue limit prediction of notched components – a critical review and modified stress gradient based approach, Int. J. Fatig. 135 (2020) 105531, https://doi.org/10.1016/j.ijfatigue.2020.105531.

[8] J.R. Duran, On fatigue life predictions for notched members by the nominal stress-based and the local strain-based methods, Appl. Comput. Mech. 16 (2022), https://doi.org/10.24132/acm.2022.678.

[9] B.J. Tian, J.J. Xiong, J.Z. Liu, A new approach for evaluating fatigue lives of multi-fastener mechanical joints based on a nominal stress concept and minimal datasets, Int. J. Fatig. 80 (2015) 257–265, https://doi.org/10.1016/j.ijfatigue.2015.06.013.
[10] F.B. Cartiaux, A. Ehrlacher, F. Legoll, A. Libal, J. Reygner, Probabilistic formulation of Miner's rule and application to structural fatigue, Probabilist. Eng. Mech. 74 (2023), https://doi.org/10.1016/j.probengmech.2023.103500.
[11] E. Kufoin, L. Susmel, Quantitative review of probabilistic approaches to fatigue design in the medium cycle fatigue regime, Probabilist. Eng. Mech. 75 (2024) 103–589, https://doi.org/10.1016/j.probengmech.2024.103589.
[12] R.C. Rice, B.C. Laboratories, Fatigue Data Analysis, ASM Handb. Comm, 2015.
[13] R.I. Stephens, A. Fatemi, R.R. Stephens, H.O. Fuchs, Metal Fatigue in Engineering, Second, John Wiley & Sons, INC, New York, 2001.
[14] Y.L. Lee, J. Pan, R.B. Hathaway, M.E. Barkey, Fatigue Testing and Analysis : Theory and Practice, Elsevier Butterworth-Heinemann., Amsterdam ; Boston, 2005.
[15] Q. Zou, J. Wen, Bayesian model averaging for probabilistic S-N curves with probability distribution model form uncertainty, Int. J. Fatig. 177 (2023) 107955, https://doi.org/10.1016/j.ijfatigue.2023.107955.
[16] A. Bengtsson, I. Rychlik, Uncertainty in fatigue life prediction of structures subject to Gaussian loads, Probabilist. Eng. Mech. 24 (2009) 224–235, https://doi.org/10.1016/j.probengmech.2008.06.004.
[17] ASTM E739-10, Standard Practise for Statistical Analysis of Linear or Linearized Stress-Life (S-N) and Strain-Life (ε-N) Fatigue Data, ASTM, 2010, pp. 1–7, https://doi.org/10.1520/E0739-23.2.
[18] Y. Peng, J. Chen, J. Dong, Experimental Data Assessment and Fatigue Design Recommendation for Stainless-Steel Welded Joints, Metals (Basel) 9 (2019) 723. https://doi.org/10.3390/met9070723.
[19] O. Kaleva, H. Orelma, Statistical properties of the model parameters in the continuum approach to high-cycle fatigue, Probabilist. Eng. Mech. 63 (2021), https://doi.org/10.1016/j.probengmech.2021.103117.
[20] C.R. Schneider, S.J. Maddox, BEST PRACTICE GUIDE ON STATISTICAL ANALYSIS OF FATIGUE DATA, 2003.
[21] G. Rugg, Using Statistics [electronic Resource] : A Gentle Introduction, McGraw-Hill Education, Maidenhead, 2007.
[22] F. Castanié. Digital Spectra Analysis: Parametric, Non-Parametric and Advanced Methods, John Wiley & Sons, Incorporated, Newark, 2011.
[23] S.F. Masri, A Hybrid Parametric/nonparametric Approach for the Identification of Nonlinear Systems, 1994.
[24] A.J. Vickers, Parametric versus non-parametric statistics in the analysis of randomized trials with non-normally distributed data, BMC Med. Res. Methodol. 5 (2005) 35, https://doi.org/10.1186/1471-2288-5-35.
[25] S. Pajevic, P.J. Basser, Parametric and non-parametric statistical analysis of DT-MRI data, J. Magn. Reson. 161 (2003) 1–14, https://doi.org/10.1016/S1090-7807(02)00178-7.
[26] V. Bagdonavicius, Non-parametric Tests for Cendored Data (Electronic Resource), ISTE, Hoboken, N.J. : Wiley, London, 2011.
[27] L. Shi, L. Khalij, H. Bai, C. Gautrelet, A non-parametric evaluation method in staircase test for improving fatigue limit assessment, Probabilist. Eng. Mech. 70 (2022), https://doi.org/10.1016/j.probengmech.2022.103373.
[28] C. Soize, A nonparametric model of random uncertainties for reduced matrix models in structural dynamics, France. www.elsevier.com/locate/probengmech, 2000.
[29] R.E. Little, E.H. Jebe, Statistical Design of Fatigue Experiments, Applied Science Publishers, London, 1975.
[30] F. Lefebvre, I. Huther, G. Parmentier, M. Huther, Document XIII-2807-19 BEST PRACTICE GUIDELINE for STATISTICAL ANALYSES of FATIGUE RESULTS, 2019.
[31] R.E. Little, J.C. Ekvall, Statistical analysis of fatigue data, ASTM STP 744, American Society for Testing and Materials (1981).
[32] R.C. Rice, Statistical aspects of design, Mater. Sel. Des. 20 (1997) 72–86, https://doi.org/10.31399/asm.hb.v20.a0002431.
[33] S.M. Marco, H.E. Frankel, M.N. Torrey, C.A. Moyer, A guide for fatigue testing and the statistical analysis of fatigue data, Am. J. Orthod. 51 (1963) 318, https://doi.org/10.1016/0002-9416(65)90116-8.
[34] C.C. Engler-Pinto, J. Lasecki, S.R. Frisch, A.D. Michael, E.A. John, Statistical approaches applied to fatigue test data analysis, SAE Int. 114 (2005) 422–431.
[35] L. Jegou, Y. Marco, V. Le Saux, S. Calloch, Fast prediction of the Wöhler curve from heat build-up measurements on Short Fiber Reinforced Plastic, Int. J. Fatig. 47 (2013) 259–267, https://doi.org/10.1016/j.ijfatigue.2012.09.007.
[36] L. Tanner, S. Collier, H.F. Dodge, R.J. Hader, G.J. Lieberman, W. Youden, ASTM MANUAL ON FITTING STRAIGHT LINES, Prentice Hall Inc., Englewood Cliffs, N.J, 1959.
[37] STP46393S, ASTM manual on fitting straight lines, ASTM man, Fitting Straight Lines (2009) 1–28, https://doi.org/10.1520/stp46393s, 1.
[38] D. Taylor, M. Hughest, D. Allent, Notch fatigue behaviour in cast irons, Int. J. Fatig. 18 (1996) 439–445.
[39] P. Livieri, E. Maggiolini, R. Tovo, On the notch sensitivity of cast iron under multi-axial fatigue loading, Frat, Integrita Strutt 30 (2014) 558–568, https://doi.org/10.3221/IGF-ESIS.30.67.
[40] A. Durmuş, H. Aydın, M. Tutar, A. Bayram, K. Yiğit, Effect of the microstructure on the notched tensile strength of as-cast and austempered ductile cast irons, Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci. 226 (2012) 2214–2229, https://doi.org/10.1177/0954406211433248.
[41] R. Louks, L. Susmel, The linear-elastic Theory of Critical Distances to estimate high-cycle fatigue strength of notched metallic materials at elevated temperatures, Fatig. Fract. Eng. Mater. Struct. 38 (2015) 629–640, https://doi.org/10.1111/ffe.12273.