Systematic Statistical Characterisation of Stress-Life Datasets using 3-Parameter Distributions

Joshua Hoole[1*], Pia Sartor[1], Julian Booker[2], Jonathan Cooper[1], Xenofon V. Gogouvitis[3], R. Kyle Schmidt[4]

[1] *Department of Aerospace Engineering, University of Bristol, United Kingdom, BS8 1TR*

[2] *Department of Mechanical Engineering, University of Bristol, United Kingdom, BS8 1TR*

[3] *Safran Landing Systems, Gloucester, United Kingdom, GL2 9QH*

[4] *Safran Landing Systems, Ajax, Ontario, Canada, L1S 2G8*

Abstract

The variability present in S-N datasets is typically characterised using probability distributions to enable the construction of Probability-S-N curves for design. 3-Parameter Log-Normal and Weibull distributions have been proposed as alternative distributions to the commonly assumed 2-Parameter Log-Normal distribution. This paper performs statistical characterisation of a 4340 steel S-N dataset from the Engineering Sciences Data Unit using a systematic methodology. The 3-Parameter Weibull distribution provided improved characterisation of the S-N dataset. Using a case study, it was also demonstrated that use of a 3-Parameter Weibull distribution can increase component safe-life values by 20% when compared to the 2-Parameter Log-Normal distribution.

Keywords: *Fatigue Design, Probabilistic Analysis, S-N Curves, Statistics*

1. Introduction

Within the sphere of fatigue design, many components across the aerospace, nuclear, offshore, steel structures and wind energy sectors are designed using a 'safe-life' philosophy (also known as 'life-limited') [1]. The safe-life of a component represents the number of duty cycles after which the component must be removed from service [2], regardless of whether fatigue crack propagation is observed. Safe-life components are typically designed using a 'classical' fatigue analysis approach of 'Stress-Life' (S-N) analysis based upon Miner's Rule [1].

At the core of stress-life analysis is the utility of S-N curves, which demonstrate how the number of cycles to failure (N_f) varies with the applied cyclic stress amplitude (σ_a) for a material as shown in Figure 1 [3]. S-N

* Corresponding Author. Email: jh12317@bristol.ac.uk Telephone: +447785547865 (Joshua Hoole)

curves are generated from extensive specimen testing and as a result of the inherent random nature of fatigue, significant variability is observed in the values of N_f for a given cyclic stress amplitude [4]. Whilst this paper focuses only on variability in material properties within fatigue design, it is important to note that variability is also observed within component loading and manufacturing quality (e.g. dimensional variability) [1]. The authors provide a wider discussion on the sources of variability in fatigue design in [5].

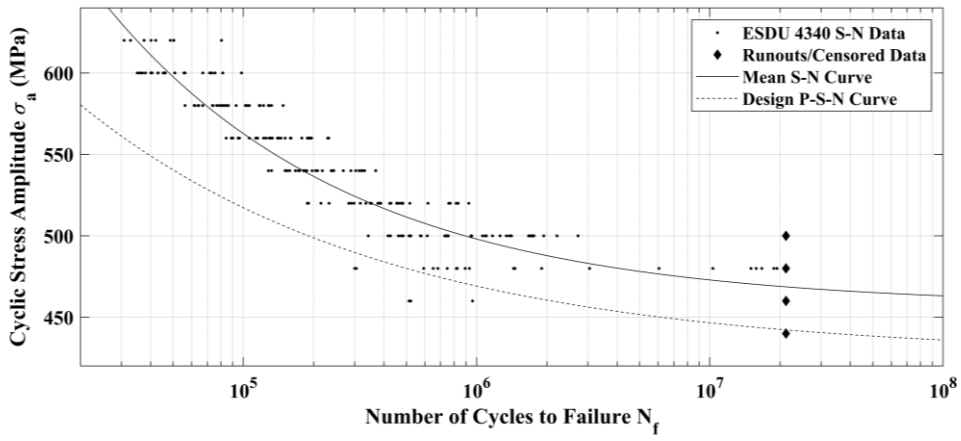

Figure 1: An example of an S-N dataset and S-N curve for 4340 steel [3]. Reproduced with permission from IHS ESDU.

To achieve reliable fatigue design for safety-critical components, the variability within N_f is often mitigated through the construction of Probability-S-N curves (P-S-N), which represent an S-N curve with a constant Probability of Survival (PoS) [6]. A recent example demonstrating the construction of P-S-N curves is presented by Zhu et al [7]. Typical PoS values are 99%, 97.7% and 95% [1] and an example of a P-S-N curve is shown in Figure 1. P-S-N curves are often referred to as 'safe', 'design', 'characteristic' or 'working' S-N curves. In order to construct P-S-N curves, a probability distribution type for N_f must be assumed and this is typically either a Log-Normal or Weibull distribution [4]. Candidate distribution types for N_f are fitted to S-N datasets using distribution fitting methods such as probability plotting or maximum likelihood estimation [8]. The selection of the final distribution type can then be performed by assessing the Goodness-of-Fit using statistical tests such as Anderson-Darling or Chi-Squared [9]. Both 2-Parameter (2P) and 3-Parameter (3P) versions of the Log-Normal and Weibull distributions exist [8, 10]. The 3P distributions differ from the 2P distributions through the introduction of a location parameter 'δ'. The δ parameter acts as a 'threshold' value, below which the probability of a value occurring is nil [8, 10]. In the context of the statistical characterisation of N_f, the δ parameter represents the minimum number of cycles to failure, thus inferring that there is a lower bound to the fatigue life of a material at a given cyclic stress amplitude [11]. Both the 2P and 3P distributions retain the shape

'λ' and scale 'σ' parameters and therefore, the 3P distributions are equivalent to the 2P distributions when $\delta = 0$. The equations for the 3P Log-Normal and 3P Weibull Probability Density Functions (PDF) are shown in Equations 1 and 2 respectively, where 'x' is the value of the random variable [10]. The impact of the threshold parameter when the shape and scale parameters are held constant is shown in Figure 2.

$$f(x;\, \lambda, \sigma, \delta) = \frac{1}{(x-\delta)\lambda\sqrt{2\pi}} exp\left\{-\frac{[ln(x-\delta)-\sigma]^2}{2\lambda^2}\right\} \quad (1)$$

$$(x;\, \lambda, \sigma, \delta) = \frac{\lambda}{\sigma}\left(\frac{x-\delta}{\sigma}\right)^{\lambda-1} exp\left\{-\left(\frac{x-\delta}{\sigma}\right)^{\lambda}\right\} \quad (2)$$

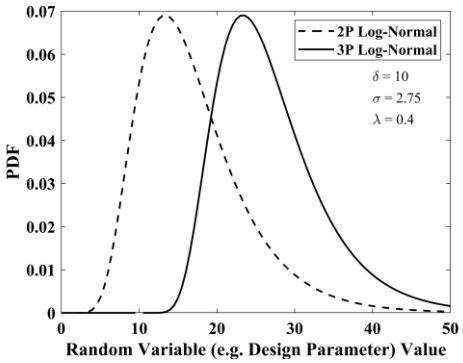

Figure 2: A comparison between a 2P Log-Normal and a 3P Log-Normal distribution, demonstrating the impact of the δ threshold (location) parameter for constant scale and shape parameters.

As the introduction of the δ parameter increases the minimum possible N_f from zero when using a 3P distribution, it is proposed that 3P distributions could result in reduced conservatism within P-S-N curves, providing they are shown to provide the best-fit to the S-N dataset. Therefore, this paper aims to identify whether 3P distributions provide improved statistical characterisation of N_f, along with quantifying the reduction in the conservatism within P-S-N curves constructed using 3P distributions. This will be achieved using a systematic methodology for the statistical characterisation of N_f at various cyclic stress amplitudes within a 4340 steel S-N dataset.

1.1. Previous Application of 3-Parameter Distributions

The application of both 3P Log-Normal and 3P Weibull distributions to characterise N_f from S-N datasets has been presented by previous work within the literature. Work performed by Zhao et al presented one of the earliest investigations into the use of 3P distributions to statistically characterise N_f within a steel S-N dataset,

using classical probability plotting methods for distribution fitting [12]. Zhao et al identified that the 2P Log-Normal distribution is more suitable for design work, due to being more conservative than the 3P Weibull distribution (i.e. predicting lower component safe-life values), despite the 3P Weibull distribution providing the best-fit to the dataset when using correlation coefficients [12]. Schijve demonstrated the applicability of 3P Log-Normal and 3P Weibull distributions to both steel and aluminium alloy S-N datasets, using probability plotting methods for distribution fitting [11]. Schijve highlighted that differences were experienced between the δ threshold values of the 3P Log-Normal and 3P Weibull distributions [11]. More recent work conducted by Wei et al, provided an investigation into comparing 2P and 3P Weibull distributions when applied to aluminium alloy S-N datasets, using probability plotting methods for fitting and the Anderson-Darling statistical test for assessing Goodness-of-Fit [13]. Wei et al concluded that the 3P Weibull distribution provided a better fit to the dataset compared to the 2P Weibull distribution, based upon the results of the Anderson-Darling test [13]. Finally, Khameneh and Azadi investigated the statistical characterisation of a cast iron material, using both probability plotting and maximum likelihood estimation methods for distribution fitting, along with the Anderson-Darling test for performing a Goodness-of-Fit test [14]. Within this work, the 3P Weibull distribution was selected as the distribution that provided the best-fit to the S-N dataset [14].

The presence of a minimum number of cycles to failure has also been supported from a physical standpoint by Schijve's work on statistical distributions of fatigue life [11]. Within this work it is suggested that the inclusion of a threshold parameter is realistic, as the zero threshold of the 2P Log-Normal and 2P Weibull distributions theoretically infers that the material could fail before loading has occurred [11]. This view is also supported by Wei et al [13]. The 3P Weibull distribution has also been proposed to represent the variability in N_f using a physical justification within the model proposed by Castillo and Fernández-Canteli [15, 16]. Within the probabilistic approach proposed by Castillo and Fernández-Canteli, it is suggested that a Weibull distribution is more appropriate to model N_f as a result of the 'weakest link' principle [16]. In addition, Castillo and Fernández-Canteli demonstrate that Weibull distributions satisfy the compatibility condition of probabilistic S-N curves [16]. The compatibility condition requires the P-S-N curves constructed based on the variability in N_f for a given stress amplitude and the variability in stress amplitude for a given N_f to be the same [16]. Therefore, there is a strong physical justification for investigating the use of 3P Weibull distributions to characterise the variability in N_f.

1.2. Paper Scope and Objectives

The review of the literature has highlighted that 3P Log-Normal and 3P Weibull distributions can be expected to provide an improved fit to S-N datasets when statistically characterising N_f [11-14], along with there being a physical basis for selecting a 3P distribution over 2P distributions [11, 15, 16]. However, the previous literature has not discussed or quantified the impact of 3P distributions on the construction of P-S-N curves for fatigue design. In addition, the distribution fitting and selection methods used within the literature often utilise only probability plotting methods, rather than the more statistically-rigorous maximum likelihood estimation approach. The common practice demonstrated in the literature also fails to perform validation of the distribution fitting and selection using multiple fitting methods. It is therefore proposed that an approach based on combining the probability plotting and maximum likelihood estimation approaches is required. Reference texts related to reliability and data analysis recommend that multiple fitting methods are to be used to achieve a robust approach to statistical characterisation [17].

Therefore, this paper aims to undertake a systematic and robust statistical characterisation of a rich S-N dataset in order to assess whether 3P distributions improve the statistical characterisation of N_f, along with evaluating the suitability of 3P distributions for constructing P-S-N curves. The reduction in conservatism when using 3P distributions will also be quantified through the application of the P-S-N curves to an S-N analysis case study based upon the Society of Automotive Engineers (SAE) keyhole geometry benchmark [18].

2. Systematic Statistical Characterisation Methodology

In order to fit the 3P distributions to an S-N dataset, a robust and systematic methodology for the statistical characterisation of S-N datasets was developed by the authors. The purpose of the novel statistical characterisation methodology is to conduct 'Fitting' and 'Selection' of the candidate distributions for N_f. 'Fitting' is the process of generating estimates of the location, shape and scale parameters of the distribution PDFs and 'selection' is the process of testing each candidate distribution for Goodness-of-Fit (GoF), in order to 'down-select' the final distribution type [8, 10]. GoF tests either accept or reject the candidate distribution. The methodology aims to maximise the amount of evidence that can be generated to support the selection of one candidate distribution over another.

Within the statistical characterisation methodology, multiple methods for fitting and GoF testing were required to provide validation. Current practice within the literature typically only uses a single fitting method and a

single GoF test. As 3P distributions are non-trivial from a fitting perspective [19], it is vital to perform validation of the distribution parameter estimates to ensure accurate values are used. The use of multiple fitting methods permits the limitations of each fitting method to be mitigated. The strengths and limitations of each fitting method will be discussed in Sections 2.3 and 2.4. The results of GoF tests also require validation as different GoF tests can often contradict one another.

A systematic statistical characterisation methodology was also developed to lead users through the same process each time, providing rules to guide decision making to 'down-select' the final distribution type. A systematic methodology is required as engineers often do not have a significant statistical background [20], and a systematic methodology can support the users' confidence in implementing statistical characterisation methods. In addition, as it has been suggested that the 2P Log-Normal distribution is often selected to characterise N_f out of a desire for mathematical simplicity and convenience [19], a systematic process supports the user by enabling a wide range of candidate distributions to be considered prior to 'down-selecting' the most appropriate distribution. Therefore, a systematic methodology supports the review and challenge of long-held statistical beliefs and assumptions [21].

The results of statistical characterisation must often be shared with other engineers, management, customers or regulatory bodies. Therefore, it is desirable to include intuitive and 'visual' methods within the statistical characterisation methodology, such that results can be shared with other stakeholders and to increase confidence in the distribution parameter estimates [17]. In addition, the methodology only utilises widely-accepted statistical methods, such that additional training resources would be available for engineers and users in the public domain.

Figure 3 presents an overview of the systematic statistical characterisation process. The flowchart in Figure 3 leads users first through dataset characterisation and selection of candidate distributions to fit. Distribution fitting is then performed using two methods in parallel for validation of the distribution parameter estimates. Following distribution fitting, two statistical GoF tests are performed in order to assess the 'fit' of each candidate distribution for each fitting method. The candidate distributions are then classified based upon whether their distribution estimates have been validated and their performance in the GoF tests. This section of the paper will briefly introduce each method used within the methodology. It should be noted that the methodology presented within this paper can be applied to all continuous datasets, both within fatigue design and the wider engineering community.

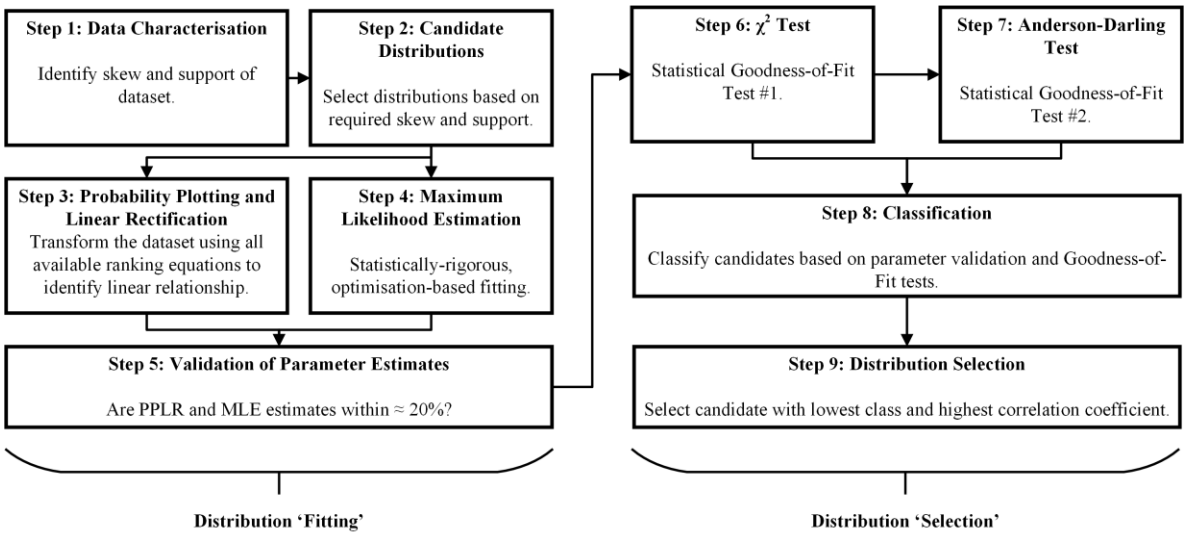

Figure 3: A visualisation of the statistical characterisation methodology developed by the authors for systematic and robust statistical characterisation.

It should be noted that the distribution fitting and goodness-of-fit methods presented within the methodology are sensitive to the sample size of the dataset. Where possible, guidance is provided about the required sample sizes for the methods within the methodology. Smaller sample sizes result in error and uncertainty in estimates of distribution parameters and lead to the requirement to construct confidence intervals on distribution parameter estimates [8]. A further discussion on confidence intervals is presented in Section 5.2. Previous work within the literature has aimed at identifying the sample sizes required to achieve a desired confidence interval [22] for S-N datasets. However, as the current paper is focused on the statistical characterisation of existing S-N datasets generated in accordance with various design standards (e.g. ASTM E739 [23]), further consideration of sample size effects is not presented.

2.1. Step 1: Dataset Characterisation

The first step of the statistical characterisation process is to characterise the dataset. The two key properties of a dataset that can guide distribution selection are 'support' and 'skew'. Support is the known possible range of values for the dataset and is typically driven by physical limitations [8]. For S-N data the support is $[0, +\infty]$ or $[a, +\infty]$ where a is the minimum N_f value introduced by the 3P distribution threshold parameter. 'Skew' is a measure of how asymmetrical a dataset is about the mean value (seen as a long 'tail' on the distribution). The sample skew 'γ' of a dataset can be computed using Equation 3, where μ_x and s_x are the mean and sample

standard deviation of the dataset respectively and N is the sample size (i.e. the number of datapoints in the dataset) [12]:

$$\gamma = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_x)^3}{s_x^3} \quad (3)$$

A positive γ value indicates that the dataset demonstrates 'positive' skew and therefore will have a right-hand distribution tail that is longer than the left, whilst a negative γ value infers the opposite [12]. S-N datasets typically demonstrate right tails and as a result, positive γ values would be expected [11, 12].

2.2. Step 2: Candidate Distributions

After identifying the required support and sample skew of the dataset, candidate distributions can be proposed. The candidate distributions should consist of distribution types that are capable of representing the dataset support and skew. In addition, statistical characterisation provides the opportunity to challenge existing assumed distribution types and therefore, any previously used distribution types for the dataset under characterisation should be included, regardless of whether or not they provide the correct support or skew characteristics [21].

2.3. Step 3: Probability Plotting and Linear Rectification

Probability plotting is a familiar tool to reliability engineers that can be used to identify whether a dataset belongs to an assumed distribution type [10, 17]. Within probability plotting, the data points of the dataset are transformed using specific empirically-based ranking equations (which approximate the relative cumulative frequency (RCF) 'F_i' or proportion of specimens that are expected to have failed by a given value) and linear rectification equations for each distribution type. If a linear relationship is observed between the ranking values and transformed dataset, the dataset can be suggested to originate from the selected distribution [10]. The coefficients from linear regression ($A0$, $A1$) can then be used to compute the distribution parameter estimates using specific linear rectification equations for each probability distribution type as found in reference texts [10]. An example of a 2P Log-Normal probability plot is shown in Figure 4.

The correlation coefficient 'r' can be computed to quantify how linear the fit is [10]. When fitting 3P distributions, the threshold parameter value that maximises r is selected as the final distribution parameter estimate [10].

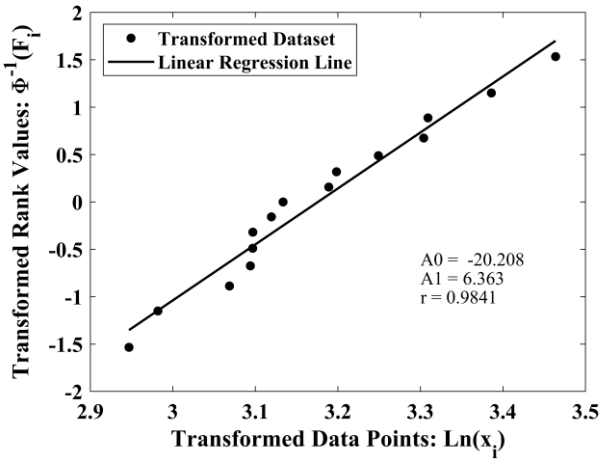

Figure 4: An example of a probability plot for a 2P Log-Normal distribution.

The correlation coefficient r is computed using Equation 4 , where $cov(X_T, F_T)$ is the covariance of the transformed data points and transformed rank values, and σ_{X_t} and σ_{F_t} are the sample standard deviations of the transformed data points and transformed values from the rank equations respectively [10].

$$r = \frac{cov(X_T, F_T)}{\sigma_{X_t}\sigma_{F_t}} \qquad (4)$$

This approach to distribution fitting will be referred to as PPLR (Probability Plotting and Linear Rectification) for the remainder of the paper. The strength of PPLR is that it is a visual and intuitive method that enables users to quickly identify the distribution parameters. The limitation of PPLR is the reliance on the selection of a ranking equation. An example of a ranking equation is the 'median rank' given in Equation 5, where i is the position of the data point in the dataset when sorted from smallest to largest value (known as 'order statistics') [10].

$$F_i = \frac{i - 0.3}{N + 0.4} \qquad (5)$$

Multiple ranking equations are available, and a selection are presented in Table A1 in the Appendix. It should be noted that each ranking equation only produces an approximation of the RCF and therefore, the selection of the ranking equation can introduce uncertainty into the PPLR fitting method. Therefore, it is recommended that all available ranking equations in Table A1 are used for the PPLR fitting and the ranking equation which maximises the correlation coefficient r should be used as the final ranking equation. As ranking equations only

provide an approximation for the RCF, an additional method for distribution fitting is required for validation of the distribution parameter estimates.

2.4. Step 4: Maximum Likelihood Estimation

Maximum Likelihood Estimation (MLE) methods are a statistically rigorous approach to estimating distribution parameter values [8]. The core philosophy behind an MLE approach is to select distribution parameter estimates that maximise the probability of the dataset being observed (this is known as the 'likelihood') [8]. Likelihood Functions (LF) are available in reference texts and are maximised using numerical solution processes and optimisation algorithms [8]. PPLR distribution parameter estimates can be used as the initial values for the numerical solution processes of the LF. Within this paper a 'brute-force' approach was used which evaluated the LF for a range of feasible distribution parameter values. The strength of the MLE approach is that for large sample sizes, the distribution parameter estimates are unbiased and the statistical uncertainty of the estimates can be assessed [8]. However, the limitation of the MLE approach lies in the non-trivial task of maximising the complex LF equations, which often require three or more maximum likelihood equations to be solved simultaneously using numerical optimisation processes [24]. This can lead to lack of convergence or convergence to a local maximum LF value rather than the global maximum, resulting in inaccurate distribution parameter estimates. Therefore, the PPLR methods (from Step 3) are required to validate the results from MLE.

2.5. Step 5: Distribution Parameter Validation

Following distribution fitting using PPLR and MLE, two sets of distribution parameter estimates will be available. In order to validate the distribution parameter estimates, the percentage difference between the PPLR and MLE estimates should be computed. From previous studies within the literature, it has been demonstrated that when using multiple fitting methods and software packages to fit a single distribution type to a dataset, the distribution parameter estimates can vary significantly [10]. Booker et al demonstrated that the estimate for standard deviation could vary across a range of approximately 20% from the minimum to maximum value, when fitting a Normal distribution to the yield strength of a carbon steel using 6 different methods [10]. Therefore, within this paper, PPLR and MLE distribution parameter estimates that have a percentage difference of less than 20% are considered to be validated. The impact of the definition of the 20% validation threshold on P-S-N curves is discussed in Section 4.1.1.

2.6. Step 6: Chi-squared Goodness-of-Fit Test

The first statistical test to be used within the methodology is the Chi-Squared (χ^2) GoF test. The χ^2 test compares the frequency of observations from the dataset with the expected frequency from the fitted probability distribution, by dividing the dataset into n_{bin} bins using rules provided by D'Agostino and Stephens [9]. The χ^2 statistic 'χ_s^2' is computed using Equation 6, where O_j is the observed frequency in each bin and E_j is the expected frequency, computed using the fitted Cumulative Distribution Function (CDF) [9]:

$$\chi_s^2 = \sum_{j=1}^{n_{bin}} \frac{\left(O_j - E_j\right)^2}{E_j} \qquad (6)$$

The χ_s^2 value is then compared to a critical value 'χ_c^2' at a specified significance level, 'α' (conventionally α = 5%), which is available in statistical tables. The distribution is then accepted if $\chi_s^2 < \chi_c^2$ or rejected if $\chi_s^2 \geq \chi_c^2$. The strength of a χ^2 GoF test is its ability to generate critical values for any candidate distribution type. The limitation of the χ^2 GoF test is that it is accurate only for larger sample sizes and should not be used for sample sizes of N < 15 [10] and should be used with caution below sample sizes of N < 50 [21]. Therefore, an additional GoF test is required to provide validation for the typical sample sizes of S-N datasets.

2.7. Step 7: Anderson-Darling Goodness-of-Fit Test

The second statistical GoF test used within the methodology is the Anderson-Darling (A-D) test. Similar to the χ^2 test, the A-D test quantifies the difference between the frequency as computed by the fitted CDF and the expected cumulative frequency from the dataset [9]. The test statistic for the A-D GoF test is A^2, which is then compared to a critical value A_{crit}^2 at a given significance level α. Tabulated A_{crit}^2 values are available for Normal, 2P Log-Normal and 2P Weibull distributions, but are limited or non-existent for 3P distributions [9, 25]. In order to produce A_{crit}^2 values for 3P distributions, an approach known as a 'parametric bootstrap' must be used [26]. The parametric bootstrap generates a random sample from the fitted distribution and re-fits the candidate distribution to the random sample [26]. The A^2 statistic is then computed, and this process is repeated for many iterations (typically 10,000 times) [25]. The estimate of the A_{crit}^2 value is then found from the $1 - \alpha$ percentile of the bootstrap results [25, 26]. As with the χ^2 GoF test, the distribution is accepted if $A^2 < A_{crit}^2$ and rejected if $A^2 \geq A_{crit}^2$.

A limitation of the A-D test is the requirement to generate A^2_{crit} values using a parametric bootstrap. When coupled with a complex MLE numerical solution approach, 10,000 iterations can typically require a few hours of computational run-time.

2.8. Step 8: Distribution Classification

One of the most significant challenges during statistical characterisation is the synthesis of all of the information generated during distribution fitting and GoF tests, in order to select the distribution that most accurately characterises the dataset. Section 3.3 provides a template for a table that can be used to record the results of fitting and GoF tests in one central location. In order to further assist users in selecting the final distribution type and to increase the systematic nature of the methodology, a series of classifications that consolidate the results of distribution fitting and GoF tests have been defined by the authors to simplify the distribution selection process.

Table 1 shows the classifications to be used and are based on the amount of evidence available that supports the candidate distribution being the best-fit to the dataset. Therefore, distributions that have validated parameter estimates and are accepted by all GoF tests will have the lowest classification. All other classifications represent a 'loss' of supporting evidence. Classifications for sample sizes of $N < 15$, where only the A-D GoF test can be performed, are shown in Table 2. The classifications synthesise all of the results from fitting and GoF tests (up to 25 elements of information) into a single value for each candidate distribution.

Table 1: The candidate distribution classification definitions, based on validated distribution parameters and the outcomes of the GoF tests for sample sizes greater than or equal to $N = 15$.

Class	Validated Distribution Parameter Estimates?	PPLR – GoF Results		MLE – GoF Results	
		χ^2	A-D	χ^2	A-D
1	Y	Accept	Accept	Accept	Accept
2	N	Accept	Accept	Accept	Accept
2	Y	3 Accept, 1 Reject – Reduced Validation across GoF Tests.			
3	N				
4	All other cases, including: • Validated or Unvalidated Parameters, rejected by 2 or more out of 4 GoF Tests. • Validated or Unvalidated Parameters, rejected by All GoF tests. • Incorrect skew or support.				

Following classification, the lowest possible class should be identified. If only Class 4 distributions are available, a wider search for other potential distribution types should be conducted. Class 3 distributions may be used as an initial approximation to the dataset but should warrant further investigation. Both Class 1 and Class 2 distributions can be considered as suitable distributions, with a Class 1 distribution being preferable to

a Class 2. It should be noted that current statistical characterisation practice within the literature uses only one fitting method and one GoF test for distribution selection [11-13]. In this case, no validation of distribution parameter estimates or GoF tests is performed and this is equivalent to a Class 4 in Table 1. Therefore, it can be seen that using Class 2 as the minimum acceptable class represents increased statistical rigour when characterising S-N datasets compared to existing practice.

Table 2: The candidate distribution classification definitions, based on validated distribution parameters and the outcomes of the GoF tests for sample sizes less than $N = 15$.

Class	Validated Distribution Parameter Estimates?	A-D GoF for PPLR	A-D GoF for MLE
1	Y	Accept	Accept
2	N	Accept	Accept
3	Y	Only 1 Test Accepts.	
4	All other cases, including: • Unvalidated Parameters, rejected by 1 or more A-D test. • Incorrect skew or support.		

2.9. Step 9: Final Distribution Selection

If only one distribution type is available in the lowest of Class 1 or 2, it should be selected as the final distribution type. If multiple candidate distributions are present within the lowest class, the correlation coefficient value 'r' from PPLR can be used to rank the remaining distributions. An r value closer to 1 represents an improved distribution fit to the dataset. Therefore, the candidate distribution in the lowest class with the highest r value should be selected as the final distribution type. Providing the MLE and PPLR distribution parameters are consistent (i.e. pass the validation check in Step 5), the r value can also be used to represent the GoF of the MLE parameter estimates.

2.10. Alternative Approaches to Distribution Fitting and Selection

The methodology presented within this paper is based solely on 'classical' or 'frequentist' methods of distribution fitting and GoF testing. The use of such methods within the methodology is required to ensure familiarity with engineers and existing practitioners.

An alternative approach to distribution fitting and selection is the Bayesian approach [27]. The Bayesian approach is based upon more advanced statistical concepts and therefore has not been used within the presented methodology, as engineers typically do not have a significant statistical background [20]. The Bayesian approach however, has been successfully applied to S-N datasets within the literature, with the view to exploiting the ability of the Bayesian approach to 'update' distribution parameter estimates based on new data [27]. However, as the purpose of the methodology presented within this paper is aimed at characterising existing

S-N datasets of specific materials, it is not expected that there will be additional data available to update the probability distributions.

It should also be noted that other GoF tests are available, such as the Kolmogorov-Smirnov (KS) [9], Cramer-von-Mises (CvM) [9], Akaike Information Criterion (AIC) [28] and Bayesian Information Criterion (BIC) [29]. The KS and CvM GoF tests were not included in the methodology as the A-D test provides greater sensitivity in distribution tails compared to the KS test [9] and the CvM test is not as widely used within the fatigue design literature. AIC and BIC, which are based on the MLE methods, only provide a 'ranking' of distribution types, rather than the 'accept-reject' approach of the A-D and χ^2 GoF tests. As the correlation coefficient from PPLR already provides a 'ranking' of the distribution types, the AIC and BIC GoF tests have not been included within the methodology. However, in the event that further evidence is required to select one Class 1 distribution over another, the AIC and BIC could be considered.

3. Statistical Characterisation of 4340 Steel S-N Dataset

This section of the paper will present the application of the statistical characterisation methodology to a real S-N dataset, with the aim of investigating the application of 3P Log-Normal and 3P Weibull distributions to statistically characterise the variability in the value of N_f across the S-N dataset. The application of the methodology will be demonstrated on a single stress amplitude from the S-N dataset in Section 3.2. The results for the complete S-N dataset will then be presented and discussed.

3.1. Engineering Sciences Data Unit 4340 Steel S-N Dataset

The S-N dataset used within this study is for 4340 steel, which is a high strength steel and typical of the materials used within aircraft landing gear components [6, 30]. The Engineering Sciences Data Unit (ESDU) provide a rich S-N dataset for 4340 steel [3]. The dataset is generated from fully-reversed (i.e. zero mean stress) rotating-bending testing of coupons [3].

The dataset and mean S-N curve for 4340 steel are shown previously in Figure 1 [3]. Within the S-N dataset, sample sizes vary from N = 10 to N = 42, with the majority of stress levels having sample sizes between N = 25 and N = 30 [3]. It should be noted that the ESDU 4340 steel dataset [3] contains both 'failure' results (i.e. separation of the specimens) and 'run-out' results, whereby the test was terminated at a pre-determined number of load cycles. 'Run-outs' are Type I censored data points [8, 24] and require additional statistical characterisation methods beyond the scope of this paper (see Section 5.2). Stress levels containing only failure

results are known as 'complete' samples and these were statistically characterised using the methodology presented in Section 2.

3.2. Demonstration of Statistical Characterisation Methodology on a Single Stress Level

Considering the dataset at σ_a = 520 MPa, the support and skew of the dataset were identified within Step 1 of the statistical characterisation methodology. The sample skew value of $\gamma = 0.948$ suggested that the dataset shows positive skew. Based upon the discussion in Section 2.1, the support of the distribution was expected to be only positive values (i.e. $[0, +\infty]$) or above a given threshold (i.e. $[a, +\infty]$). Step 2 required the selection of the candidate distributions. As the scope of this paper is limited to investigating the application of 3P distributions, only the 2P Log-Normal, 3P Log-Normal, 2P Weibull and 3P Weibull distributions were considered (all of these distributions provided the required support and skew characteristics).

The probability plots from the PPLR distribution fitting process in Step 3 are shown in Figure 5. It can be seen from Figure 5 that all candidate distributions presented linear relationships between the transformed data points and rank values. It was found that the Mean ranking equation (see Table A.1. in Appendix) maximised the correlation coefficient for each candidate distribution. It can also be seen from Figure 5 that the 3P Weibull distribution minimised the deviation in the linear relationship at the lower tail of the 2P Weibull probability plot. This suggests an improved fit to the dataset when using a 3P distribution, supporting the presence of a threshold parameter. The PPLR distribution parameter estimates are shown in the results table in Table 3.

Step 4 of the statistical characterisation methodology performed MLE fitting of the distributions and the resulting parameter estimates are shown in Table 3. Table 3 also shows that the maximum percentage difference between PPLR and MLE parameter estimates varied significantly across the different candidate distributions, from 4.853% to 48.612%. Using 20% difference as the threshold for validation (see Section 2.5), the results inferred that the 2P Log-Normal and 2P Weibull distributions are the only candidate distributions with validated parameter estimates (Step 5). It can also be seen from Table 3 that the threshold parameter estimates for the 3P Log-Normal and 3P Weibull distributions varied significantly to one another. This is consistent with the work conducted by Schijve who also observed inconsistency between the threshold values for 3P Log-Normal and 3P Weibull distributions [11].

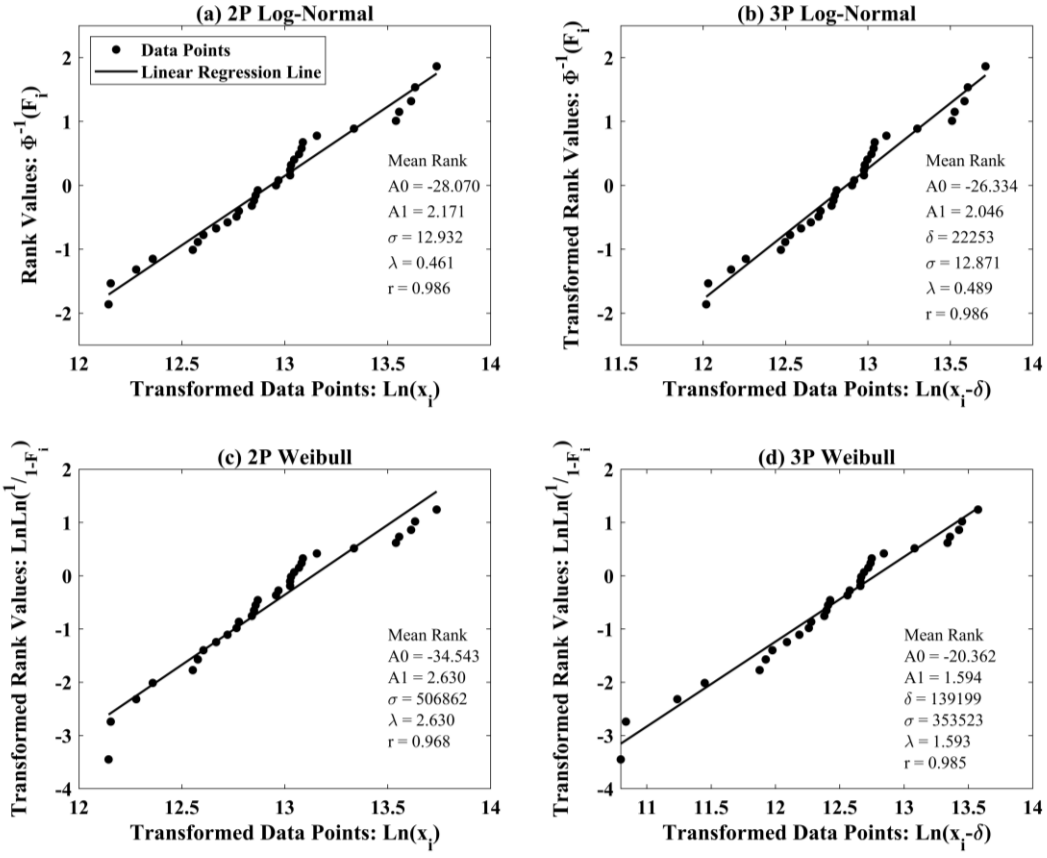

Figure 5: The probability plots for (a) 2P Log-Normal, (b) 3P Log-Normal, (c) 2P Weibull and (d) 3P Weibull distributions, showing the linear regression coefficients and distribution parameter estimates for the S-N dataset at σ_a = 520 MPa. The threshold parameter can be seen to improve the linear fit from the 2P to the 3P Weibull distribution.

Table 3: The probability plotting and maximum likelihood distribution parameter estimates for the candidate distributions for the number of cycles to failure at σ_a = 520 MPa. The maximum difference is the maximum percentage difference between the PPLR and MLE estimates. The PPLR ranking equation is also shown as (*italics*).

Distribution	Fitting Method (*Ranking Equation*)	Distribution Parameter Estimate			Maximum Difference
		Location/Threshold δ	Scale σ	Shape λ	
2P Log-Normal	PPLR (*Mean*)	-	12.932	0.461	10.333%
	MLE	-	12.932	0.418	
3P Log-Normal	PPLR (*Mean*)	22253	12.861	0.472	48.612%
	MLE	43304	12.809	0.464	
2P Weibull	PPLR (*Mean*)	-	504751	2.767	4.853%
	MLE	-	508269	2.508	
3P Weibull	PPLR (*Mean*)	146676	342490	1.603	21.187%
	MLE	176620	298760	1.408	

Table 4 shows the results of the χ^2 GoF test from Step 7. The 'A' prefix demonstrates that the test accepted the candidate distribution, whilst an 'R' prefix with *italic* values shows that the test rejected the candidate distribution. It can be seen from Table 4, that the 3P Weibull distribution when fitted using a PPLR approach

was the only candidate distribution to be rejected by the χ^2 GoF test at a significance level of α = 5%. The 3P Weibull distribution fitted using MLE was accepted at the α = 5% significance level by the χ^2 GoF test.

Table 4: The results of the goodness-of-fit tests for the candidate distributions for number of cycles to failure at σ_a = 520 MPa. The 'A' and 'R' prefixes represent that the distribution was accepted or rejected by the goodness-of-fit test respectively. The final distribution classifications are also shown.

Distribution	Fitting Method (*Ranking Equation*)	χ_s^2	χ_c^2 $\alpha = 5\%$	A^2	A_c^2 $\alpha = 5\%$	r	Class
2P Log-Normal	PPLR (*Mean)*	4.055	A 5.992	0.445	A 0.752	0.986	1
	MLE	2.810	A 5.992	0.386	A 0.752	-	
3P Log-Normal	PPLR (*Mean)*	3.012	A 5.992	0.429	A 0.601	0.986	2
	MLE	1.831	A 5.992	0.383	A 0.649	-	
2P Weibull	PPLR (*Mean)*	3.535	A 5.992	0.990	R *0.757*	0.968	4
	MLE	3.648	A 7.815	0.904	R *0.757*	-	
3P Weibull	PPLR (*Mean)*	9.516	R *3.842*	0.521	A 0.885	0.985	3
	MLE	2.686	A 5.992	0.509	A 0.891	-	

The results from the A-D GoF test are shown next in Table 4 from Step 8. It can be seen that the 2P Weibull distribution, when fitted with PPLR and MLE was rejected at the α = 5% significance level. This also highlights the importance of validating GoF tests, as the rejection of the 2P Weibull distribution by A-D contradicts its acceptance by the χ^2 test. In a similar fashion, the 3P Weibull distribution fitted using PPLR was accepted by the A-D GoF test despite being rejected by χ^2 test.

The final column of Table 4 shows the classification of each candidate distribution for Step 10, based upon the Class definitions shown previously in Table 1. It can be seen from Table 4, that the 2P Log-Normal distribution was Class 1, the 3P Log-Normal distribution was Class 2 (due to unvalidated distribution parameter estimates), the 2P Weibull distribution was Class 4 (due to A-D GoF test rejecting both PPLR and MLE estimates) and the 3P Weibull distribution was Class 3 (due to unvalidated parameter estimates and χ^2 rejecting the PPLR estimates). Therefore, in this isolated case, the 2P Log-Normal distribution should be selected, due to having the lowest class number. Despite having a similar correlation coefficient value to the 2P Log-Normal distribution, the 3P Log-Normal distribution is rejected due to the large difference of 48.612% between the location parameter from PPLR and MLE and therefore resulting in unvalidated distribution parameter estimates.

3.3. Statistical Characterisation of Complete 4340 Steel S-N Dataset

The systematic statistical characterisation methodology described in Section 2.2 and demonstrated in Section 3.1 was then applied to the remaining complete 4340 S-N datasets at σ_a = 620, 600, 580 and 540 MPa (identified as Stress Level 1, 2, 3, 4 and 5 respectively) [3]. The results table template is shown in Table 5a and 5b and has been populated for all complete datasets within the ESDU 4340 steel S-N dataset. Stress Level 6 at σ_a = 520

MPa as characterised in Section 3.2 is included in Table 5b for completeness. This section will aim to summarise the results of the statistical characterisation. The clearest approach to reviewing the vast amount of information contained within Tables 5a and 5b is to focus on each candidate distribution type individually. Only the A-D GoF test could be performed at Stress Level 1 at σ_a = 620 MPa due to the sample size of N = 10.

Table 5a: The distribution fitting and goodness-of-fit results for Stress Levels 1, 2 and 3 (σ_a = 620 MPa, 600 MPa and 580 MPa respectively). The 'A' and 'R' prefixes demonstrate that a goodness-of-fit test either accepts or rejects a candidate distribution. The final candidate distribution classification is shown on the right-hand side of the table.

Stress Level 1	σ = 620 MPa	Sample Size N		10	Sample Skew γ	1.528		Best Fit Distribution			**2P Log-Normal**
Distribution	Fitting Method (*Rank*)	Distribution Parameter Estimate			Maximum Difference	χ_s^2	χ_c^2 $\alpha = 5\%$	A^2	A_c^2 $\alpha = 5\%$	r	Class
		Location δ	Scale σ	Shape λ							
2P Log-Normal	PPLR (*EV*)	-	10.645	0.300	4.590%	-	-	0.404	A 0.752	0.959	1
	MLE	-	10.665	0.286		-	-	0.384	A 0.752	-	
3P Log-Normal	PPLR (*EV*)	25795	9.552	0.743	25.584%	-	-	0.252	A 0.635	0.982	2
	MLE	29337	9.195	1.011		-	-	0.308	A 1.001	-	
2P Weibull	PPLR (*EV*)	-	48562	3.717	15.990%	Incorrect Skew Behaviour. Reject 2P Weibull Distribution.					4
	MLE	-	49602	3.205							
3P Weibull	PPLR (*Hazen*)	30167	45010	0.935	59.542%	Produces Exponential Distribution. Reject 3P Weibull Distribution.					4
	MLE	30800	111251	0.630							
Stress Level 2	σ = 600 MPa	Sample Size N		17	Sample Skew γ	0.761		Best Fit Distribution			**3P Log-Normal**
Distribution	Fitting Method (*Rank*)	Distribution Parameter Estimate			Maximum Difference	χ_s^2	χ_c^2 $\alpha = 5\%$	A^2	A_c^2 $\alpha = 5\%$	r	Class
		Location δ	Scale σ	Shape λ							
2P Log-Normal	PPLR (*Mean*)	-	10.903	0.351	16.012%	6.253	R *5.992*	0.390	A 0.752	0.981	2
	MLE	-	10.903	0.303		3.448	A 5.992	0.392	A 0.752	-	
3P Log-Normal	PPLR (*Mean*)	26745	10.136	0.713	15.335%	3.750	A 5.992	0.253	A 0.626	0.990	2
	MLE	31590	9.857	0.796		4.719	R *3.842*	0.277	A 0.630	-	
2P Weibull	PPLR (Mean)	-	63416	3.348	2.350%	Incorrect Skew Behaviour. Reject 2P Weibull Distribution.					4
	MLE	-	63202	3.428							
3P Weibull	PPLR (*Hazen*)	33711	25148	1.207	41.284%	3.280	A 5.992	0.205	A 1.002	0.992	4
	MLE	35096	17800	0.860		MLE Produces Exponential Distribution				-	
Stress Level 3	σ = 580 MPa	Sample Size N		26	Sample Skew γ	0.503		Best Fit Distribution			**3P Weibull**
Distribution	Fitting Method (*Rank*)	Distribution Parameter Estimate			Maximum Difference	χ_s^2	χ_c^2 $\alpha = 5\%$	A^2	A_c^2 $\alpha = 5\%$	r	Class
		Location δ	Scale σ	Shape λ							
2P Log-Normal	PPLR (*Mean*)	-	11.394	0.309	11.506%	2.247	A 5.992	0.349	A 0.752	0.987	1
	MLE	-	11.394	0.277		2.563	A 5.992	0.398	A 0.752	-	
3P Log-Normal	PPLR (*Mean*)	32771	10.897	0.489	18.280%	3.067	A 3.842	0.285	A 0.616	0.989	2
	MLE	40102	10.732	0.517		4.318	R *3.842*	0.353	A 0.616	-	
2P Weibull	PPLR (*Mean*)	-	101850	3.871	0.821%	Incorrect Skew Behaviour. Reject 2P Weibull Distribution.					4
	MLE	-	101837	3.903							
3P Weibull	PPLR (*Mean*)	51936	45682	1.400	12.577%	1.722	A 3.815	0.253	A 0.940	0.992	1
	MLE	55015	40578	1.330		1.884	A 3.815	0.264	A 0.911	-	

3.3.1. 2P Log-Normal Distribution

From Table 5a and 5b, it can be seen that the 2P Log-Normal distribution was predominately a Class 1 distribution across the six stress levels as it was accepted by both GoF tests at the α = 5% significance level with validated distribution parameter estimates (i.e. PPLR and MLE parameter estimates are within 20%). The only instance whereby 2P Log-Normal distribution was higher than Class 1 was at Stress Level 2, due to rejection by the χ^2 GoF test for the PPLR distribution parameter estimates. The 2P Log-Normal distribution

also had high correlation coefficient values across the stress levels. The consistency of the 2P Log-Normal distribution to be a Class 1 distribution, along with its high correlation coefficient values suggests that the 2P Log-Normal distribution could provide accurate statistical characterisation of the value of N_f across the S-N dataset. The demonstration of the statistical validity of the 2P Log-Normal distribution is in agreement with the existing and commonly held assumption that N_f is 2P Log-Normally distributed [12, 19].

Table 5b: The distribution fitting and goodness-of-fit test results for Stress Levels 4, 5 and 6 (σ_a = 560 MPa, 540 MPa and 520 MPa respectively). The 'A' and 'R' prefixes demonstrate that a goodness-of-fit test either accepts or rejects a candidate distribution. The final candidate distribution classification is shown on the right-hand side of the table.

Stress Level 4	σ = 560 MPa	Sample Size N	28	Sample Skew γ	0.413	Best Fit Distribution				**3P Weibull**
Distribution	Fitting Method (*Rank*)	Distribution Parameter Estimate: Location δ	Scale σ	Shape λ	Maximum Difference	χ_s^2	χ_c^2 $\alpha = 5\%$	A^2	A_c^2 $\alpha = 5\%$	r	Class
2P Log-Normal	PPLR (*Mean*)	-	11.848	0.320	10.405%	2.077	A 5.992	0.234	A 0.752	0.992	1
	MLE	-	11.849	0.290		1.498	A 5.992	0.248	A 0.752	-	
3P Log-Normal	PPLR (*Mean*)	0	11.848	0.320	Inconsistent Threshold Parameter. Reject 3P Log-Normal Distribution						4
	MLE	13798	11.740	0.317							
2P Weibull	PPLR (*Mean*)	-	161041	3.791	0.903%	Incorrect Skew Behaviour. Reject 2P Weibull Distribution.					4
	MLE	-	161065	3.826							
3P Weibull	PPLR (*Mean*)	72364	83901	1.589	16.223%	3.230	A 5.992	0.243	A 1.013	0.993	1
	MLE	80325	72189	1.471		1.602	A 3.842	0.294	A 0.898	-	
Stress Level 5	σ = 540 MPa	Sample Size N	28	Sample Skew γ	0.518	Best Fit Distribution				**3P Weibull**	
Distribution	Fitting Method (*Rank*)	Distribution Parameter Estimate: Location δ	Scale σ	Shape λ	Maximum Difference	χ_s^2	χ_c^2 $\alpha = 5\%$	A^2	A_c^2 $\alpha = 5\%$	r	Class
2P Log-Normal	PPLR (*Mean*)	-	12.280	0.325	10.533%	1.566	A 5.992	0.286	A 0.752	0.991	1
	MLE	-	12.280	0.294		1.368	A 5.992	0.315	A 0.752	-	
3P Log-Normal	PPLR (*Mean*)	54037	11.972	0.439	20.553%	1.266	A 3.842	0.237	A 0.673	0.992	2
	MLE	68017	11.871	0.432		1.042	A 3.842	0.270	A 0.607	-	
2P Weibull	PPLR (Mean)	-	248920	3.697	3.235%	Incorrect Skew Behaviour. Reject 2P Weibull Distribution.					4
	MLE	-	249202	3.580							
3P Weibull	PPLR (*Mean*)	110601	130134	1.617	15.184%	1.657	A 3.842	0.234	A 0.916	0.993	1
	MLE	122587	112979	1.458		3.574	A 5.992	0.228	A 0.901	-	
Stress Level 6	σ = 520 MPa	Sample Size N	31	Sample Skew γ	0.948	Best Fit Distribution				**2P Log-Normal**	
Distribution	Fitting Method (*Rank*)	Distribution Parameter Estimate: Location δ	Scale σ	Shape λ	Maximum Difference	χ_s^2	χ_c^2 $\alpha = 5\%$	A^2	A_c^2 $\alpha = 5\%$	r	Class
2P Log-Normal	PPLR (*Mean*)	-	12.932	0.461	10.333%	4.055	A 5.992	0.445	A 0.752	0.986	1
	MLE	-	12.932	0.418		2.810	A 5.992	0.386	A 0.752	-	
3P Log-Normal	PPLR (*Mean*)	22253	12.871	0.489	48.612%	3.012	A 5.992	0.429	A 0.601	0.986	2
	MLE	43304	12.809	0.464		1.831	A 5.992	0.383	A 0.649	-	
2P Weibull	PPLR (*Mean*)	-	506862	2.630	4.853%	3.535	A 7.815	0.990	R *0.757*	0.968	4
	MLE	-	508269	2.508		3.648	A 7.815	0.904	R *0.757*	-	
3P Weibull	PPLR (*Mean*)	139199	353523	1.593	21.187%	9.516	R *3.842*	0.521	A 0.885	0.985	3
	MLE	176620	298760	1.408		2.686	A 5.992	0.509	A 0.891	-	

3.3.2. 2P Weibull Distribution

The 2P Weibull distribution was classified at all stress levels as Class 4 as shown in Table 5a and 5b. This was as a result of both the PPLR and MLE fitting process producing λ shape estimates that are in excess of $\lambda = 3$, producing a distribution that will demonstrate either symmetrical or negative skew [31] at Stress Level 1 to 5. This contradicts the positive sample skew values computed for each dataset. At Stress Level 6, the 2P Weibull distribution was rejected by the A-D test at 5% for both the PPLR and MLE distribution parameter estimates. Therefore, the 2P Weibull distribution does not provide accurate statistical characterisation of the N_f values across the 4340 S-N dataset.

3.3.3. 3P Log-Normal Distribution

From Tables 5a and 5b, it can be seen that the 3P Log-Normal distribution was predominately a Class 2 distribution across the 6 stress levels, although it was Class 4 for Stress Level 4 due to inconsistent PPLR and MLE location parameter estimates. Despite having larger correlation coefficient values compared to the 2P Log-Normal distribution and therefore suggesting a better fit, the percentage difference between the PPLR and MLE distribution parameter estimates typically exceeded the 20% threshold, resulting in unvalidated parameter estimates at Stress Levels 2 to 6. The MLE distribution parameter estimates for the 3P Log-Normal distribution were also rejected by the χ^2 test at $\alpha = 5\%$ for Stress Level 2 and 3. The most significant observation regarding the 3P Log-Normal distribution from Table 5a and 5b was the instability and inconsistency in the δ threshold parameter estimates. Firstly, it can be seen for Stress Level 4 that the PPLR threshold parameter estimate was equal to zero and therefore the 3P Log-Normal distribution was equivalent to the 2P Log-Normal distribution. In addition, it can be seen that the threshold parameter value did not continually increase for a reducing stress, as shown in Figure 6. It would be expected that the minimum number of cycles to failure defined by the threshold parameter would increase with reducing stress amplitude, in a similar manner to the mean number of cycles to failure increasing with reducing cyclic stress amplitude (i.e. the classical S-N curve response). As can be seen in Figure 6, the threshold parameter increases from Stress Level 1 to 3, but then rapidly decreased and fluctuated across Stress Level 4, 5 and 6. The instability in the threshold parameter and high class numbers means that the 3P Log-Normal distribution should not be used characterise the N_f values within then 4340 S-N data set.

3.3.4. 3P Weibull Distribution

Regarding the 3P Weibull distribution, Table 5a and 5b shows that for half of the characterised Stress Levels (3, 4 and 5), the 3P Weibull was a Class 1 distribution, with correlation coefficients at Stress Levels 3, 4 and 5 that were higher than the other Class 1 distribution, 2P Log-Normal. Table 5a and 5b also show that the 3P Weibull distribution provided the best-fit for half of the stress levels. The 2P Log-Normal distribution only provided the best-fit for 2 stress levels. In addition, Figure 6 shows that the δ threshold parameter of the 3P Weibull distribution was significantly more stable, showing a continual increase in threshold value for reducing stress, compared to the unstable behaviour of the 3P Log-Normal distribution threshold parameter.

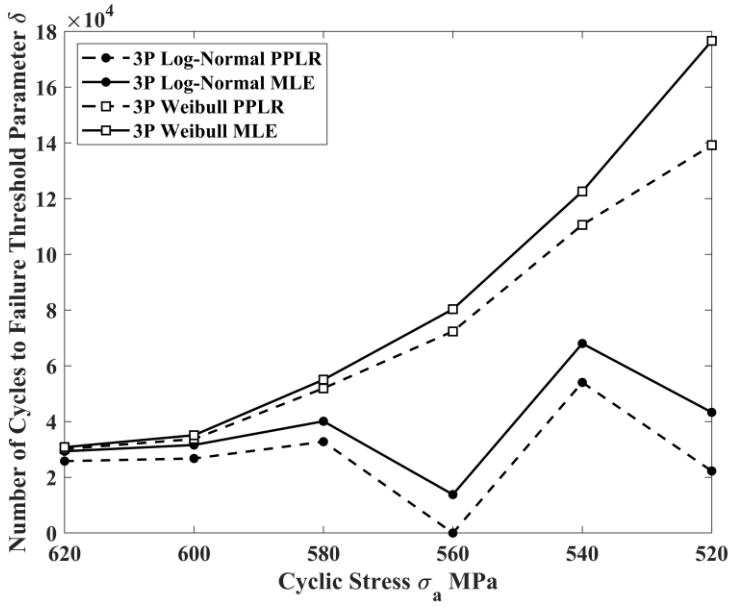

Figure 6: A visualisation of how the 3P threshold parameter varies with reducing stress amplitude. The 3P Weibull distribution shows a constantly increasing threshold parameter, whilst the 3P Log-Normal distribution shows an unstable threshold parameter value.

However, at Stress Levels 1, 2 and 6, the 3P Weibull distribution was classified as either Class 4 or Class 3 and this was expected to be as a result of the challenges of fitting the 3P Weibull distribution. For Stress Level 1, an Exponential distribution was produced by both PPLR and MLE methods (as shown by the shape parameter values of $\lambda < 1$) and the 3P Weibull distribution was therefore rejected due to the incorrect skew behaviour. The incorrect skew behaviour is suggested to be as a result of the small sample size of $N = 10$. At Stress Level 2, PPLR produced a 3P Weibull distribution, with a high correlation coefficient and was accepted by both GoF tests, whilst MLE produced an Exponential distribution. This suggests that the optimisation process used to maximise the 3P Weibull LF converged to a local maximum, rather than the global maximum. Had the global maximum been identified, it would be expected that the 3P Weibull distribution would have also been Class 1

and provided the best-fit at Stress Level 2. For Stress Level 6, the PPLR-fitted distribution was rejected by the χ^2 test, whilst the MLE-fitted distribution was accepted by both GoF tests. This contradiction between the two fitting methods could be as a result of PPLR ranking only providing an approximation of the RCF, therefore introducing inaccuracy into the distribution fitting process.

Therefore, Table 5a and 5b show that when the 3P Weibull distribution had been successfully fitted to the S-N dataset with validated distribution parameter estimates, it was seen to have the highest correlation coefficient from PPLR and should therefore could be considered as the final distribution along with the 2P Log-Normal distribution for the S-N dataset. Possible approaches to overcome the challenges of fitting the 3P Weibull distribution at Stress Levels 1, 2 and 6 are discussed further in Section 5.1. The identification of the 3P Weibull distribution providing the best-fit to S-N datasets supports the findings of Zhao et al [12], Wei et al [13] and Khameneh and Azadi [14].

Whilst the 3P Weibull was shown to provide the best-fit to the S-N dataset at a number of stress levels, the 2P Log-Normal distribution was also shown to be predominately a Class 1 distribution. Therefore, the common assumption that N_f values are 2P Log-Normal distributed [12, 19] has also been shown to be statistically valid.

4. Case Study: Evaluating the Reduction in Conservatism from 3-Parameter Distributions

As the 3P Weibull distribution was shown to provide the best-fit at a number of the stress levels within the 4340 steel S-N dataset, the impact of characterising the variability in N_f values using 3P Weibull distributions on P-S-N curves and a component's safe-life was investigated and quantified using an S-N analysis case study. It should be noted that as both the 2P Weibull and 3P Log-Normal distributions were found to provide unacceptable fits to the 4340 steel S-N dataset in Section 3, they have not been considered for the construction of P-S-N curves in the following sections.

4.1. Impact of 3-Parameter Weibull Distributions on P-S-N Curves

The first investigation considered the N_f value required to achieve a given PoS at each stress level for the fitted 2P Log-Normal and 3P Weibull distributions ($N_{f_{PoS}}$). $N_{f_{PoS}}$ is computed using the inverse of the CDF of the fitted distribution [6]. Figure 7 shows the 2P Log-Normal distribution and 3P Weibull distribution fitted to the dataset at Stress Level 6 and the presence of the threshold parameter for the 3P Weibull distribution can be clearly seen. Figure 8 shows the original 4340 S-N dataset [3], with the $N_{f_{PoS}}$ values at 99% PoS for both the 2P Log-Normal distribution and 3P Weibull distribution for Stress Levels 2 to 6 (Stress Level 1 has been omitted

due to the small sample size). The 99% $PoS\ N_{f_{PoS}}$ values are also shown in Table 6. A PoS of 99% was selected as this is a typical PoS used for safety-critical components in the aerospace sector [1].

It can be seen from Table 6 and Figures 7 and 8 that the 3P Weibull distribution 99% $PoS\ N_{f_{PoS}}$ values were approximately 20% larger than those of the 2P Log-Normal distribution. This suggests that the 3P Weibull distribution, when successfully fitted to the dataset, provides a reduction in the conservatism of the P-S-N curves of approximately 20%. This would suggest, neglecting all other sources of variability in fatigue design, that if components were designed using the 3P Weibull distribution 99% PoS P-S-N curve, they would have a theoretical design safe-life 20% larger than if designed using the 2P Log-Normal distribution 99% PoS P-S-N curve. Alternatively, the increase in available fatigue life could enable higher cyclic stresses to be applied to the component, permitting structural elements with reduced sectional areas and thicknesses. This would result in lighter-weight structural components when designing with the 3P Weibull P-S-N curve. Therefore, the improved statistical characterisation brought about by the systematic methodology selecting the 3P Weibull distribution could enable engineers to reduce conservatism within fatigue design, leading to longer component lives and increased component efficiency [2].

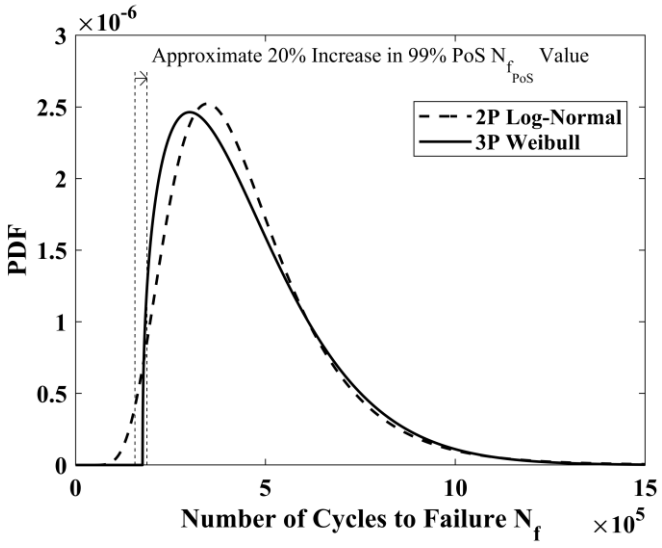

Figure 7: A comparison of the 2P Log-Normal and 3P Weibull distributions at Stress Level 6 (520 MPa), clearly highlighting the presence of the 3P Weibull distribution δ threshold parameter.

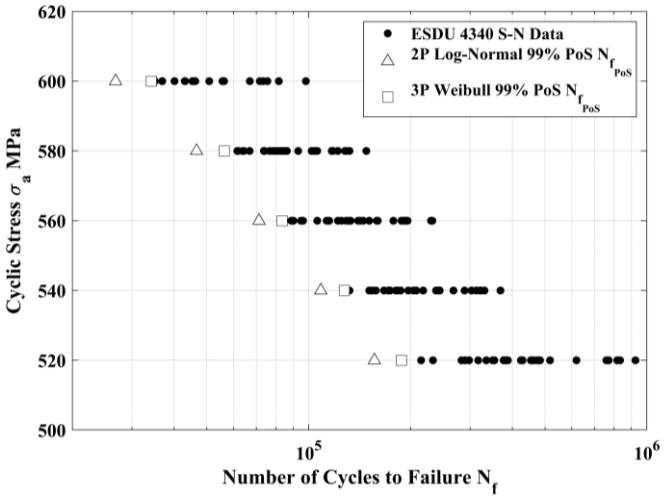

Figure 8: A comparison of 99% PoS number of cycles to failure values generated at each stress level by the 2P Log-Normal and 3P Weibull distribution. 4340 S-N dataset reproduced with permission from IHS ESDU [3].

Table 6: The 99% PoS number of cycles to failure values $N_{f_{PoS}}$ generated at each stress-level from the 2P Log-Normal and 3P Weibull distributions. The 3P Weibull distribution can be seen to provide an increase in $N_{f_{PoS}}$ of approximately 20%. The successful fitting method for the 3P Weibull distribution is also shown.

Stress Level	σ_a (MPa)	99% PoS $N_{f_{PoS}}$ Value		Percentage Increase in $N_{f_{PoS}}$ for 3P Weibull
		2P Log-Normal	3P Weibull	
2	600	26854 (MLE)	34267 (PPLR)	27.605%
3	580	46610 (MLE)	56291 (MLE)	20.770%
4	560	71315 (MLE)	83490 (MLE)	17.072%
5	540	108560 (MLE)	127408 (MLE)	17.362%
6	520	156426 (MLE)	187999 (MLE)	20.184%

Whilst the focus of this paper is the construction of P-S-N curves at a specified PoS for deterministic design, the probabilistic scatter band when using the 2P Log-Normal and 3P Weibull distributions can also be compared. Figure 9 shows the $N_{f_{PoS}}$ at 99%, 50% (i.e. 'mean') and 1% PoS values. It can be observed from Figure 9 that the 50% PoS $N_{f_{PoS}}$ values are similar for both the 2P Log-Normal and 3P Weibull distributions. In addition, Figure 9 shows that the 50% and 1% PoS $N_{f_{PoS}}$ have a smaller difference between the 2P Log-Normal and 3P Weibull distributions, typically differing by only 10%. It should also be noted that at 50% PoS $N_{f_{PoS}}$ values, the 2P Log-Normal distribution predicts higher $N_{f_{PoS}}$ values than the 3P Weibull. These observations from Figure 9 suggest that as the design P-S-N PoS value increases, that the potential life increase offered by the 3P Weibull reduces.

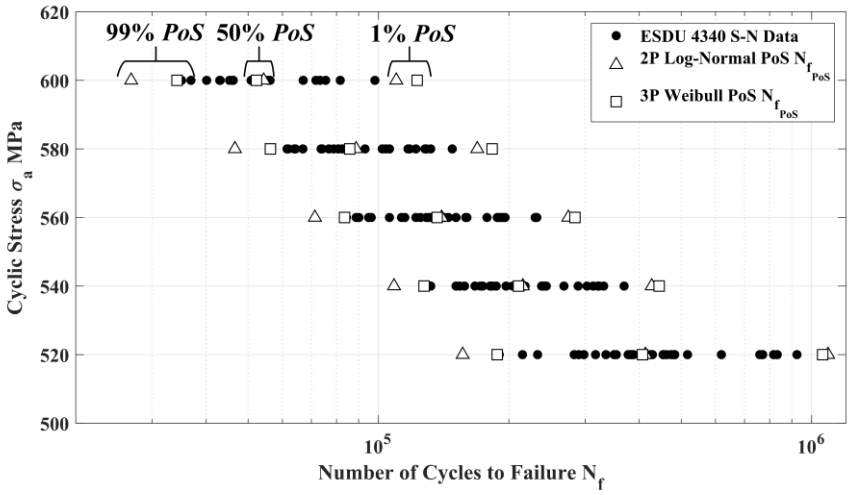

Figure 9: A comparison of the 99%, 50% and 1% PoS scatter bands constructed using the 2P Log-Normal and 3P Weibull distributions. 4340 S-N dataset reproduced with permission from IHS ESDU [3].

4.1.1. Impact on P-S-N Curves of Validation Threshold

The results in Table 7 were used to investigate the impact of the 20% validation threshold of percentage difference between the PPLR and MLE distribution parameter estimates. Table 7 shows the maximum percentage difference between the distribution parameter estimates and the corresponding difference in the 99% PoS $N_{f_{PoS}}$ values for the Class 1 2P Log-Normal distribution and 3P Weibull distributions. It should be noted that Table 7 omits values from Stress Level 2 as Table 5a shows that no Class 1 distributions were identified. It can be observed from Table 7 that all distribution parameter estimates with a maximum difference of less than 20% resulted in no more than a 10% difference in the 99% PoS $N_{f_{PoS}}$ values. Within the literature there is limited guidance on acceptable percentage errors and differences between N_f values and whilst work presented in the literature has stated the use of a 5% acceptable error, there is little justification for the selection of such a value [22]. Therefore, a conscious effort is required within the fatigue design community to define acceptable error values for N_f from S-N curves, accounting for both error in the statistical characterisation and selection of S-N curve shape. This could be achieved in future work by investigating the impact on predicted component lives of the error in N_f values.

Table 7: The relation between the maximum difference between PPLR and MLE distribution parameter estimates and the resulting difference in the 99% PoS number of cycles to failure values.

Stress Level	Distribution	Max Difference in Parameter Estimates from PPLR and MLE	Difference in 99% PoS $N_{f_{PoS}}$ between PPLR and MLE Distributions
3	2P Log-Normal	11.506%	7.142%
	3P Weibull	12.577%	4.705%
4	2P Log-Normal	10.405%	6.769%
	3P Weibull	16.223%	7.768%
5	2P Log-Normal	10.533%	6.958%
	3P Weibull	15.184%	7.258%
6	2P Log-Normal	10.333%	9.546%
	3P Weibull	21.187%	15.484%

4.2. Impact of 3-Parameter Weibull Distribution on Component Safe-Life

In order to further investigate the impact of assuming a 3P Weibull distribution for the construction of P-S-N curves, a case study was defined in order to quantify the difference in a component safe-life when using P-S-N curves constructed using 2P Log-Normal and 3P Weibull distributions. The SAE keyhole benchmark shown in Figure 10 was used as a representative component geometry [18]. The original benchmark geometry was retained, although it was assumed that the component was manufactured from 4340 steel, in order to utilise the statistical characterisation of the ESDU dataset [3]. The SAE 'transmission' spectrum shown in Figure 10 was used as the case study loading spectrum [18]. Rainflow counting was applied to the loading spectrum to identify individual load cycle ranges [4]. These were then converted to stress ranges, based on the nominal component geometry and a stress concentration factor of K_T = 3, to account for the keyhole notch [1]. The Goodman mean stress correction was employed to convert the stress ranges into fully-reversed cyclic stress amplitudes σ_a [4].

Assuming a Basquin S-N curve shape (as is commonly performed for S-N datasets) [2, 4], 99% PoS P-S-N curves were constructed for both the 2P Log-Normal $N_{f_{PoS}}$ and 3P Weibull $N_{f_{PoS}}$ values using 'minimum' fitting. The resulting curves are shown on log-log scales in Figure 11. Minimum fitting ensures that all datapoints used to fit the curve, either lie on, or on the 'conservative' side of the curve, such that the P-S-N curve represents a lower N_f than the $N_{f_{PoS}}$ value. This can be seen in Figure 11 for both the 2P Log-Normal and 3P Weibull P-S-N curve, whereby the $N_{f_{PoS}}$ at σ_a = 600 MPa and σ_a = 520 MPa lie on the P-S-N curves, whilst at all other stress levels the P-S-N curves under predict the N_f value. The fatigue limit 'σ_{FL}' (i.e. the cyclic stress amplitude below which damage is assumed to not be accumulated) was assumed to be Normally distributed [4]. The 99% PoS fatigue limit was computed based upon a mean fatigue limit value of σ_{FL} = 457

MPa and a sample standard deviation of σ = 13. The value for σ_{FL} and standard deviation was computed using the 'Probit' method [4].

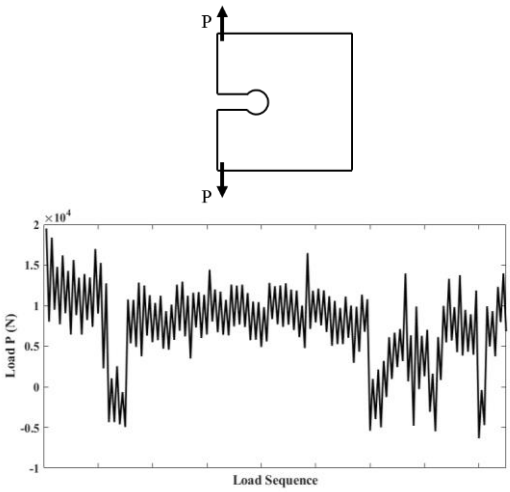

Figure 10: The SAE Keyhole benchmark geometry and a section of the 'transmission' loading spectrum [18].

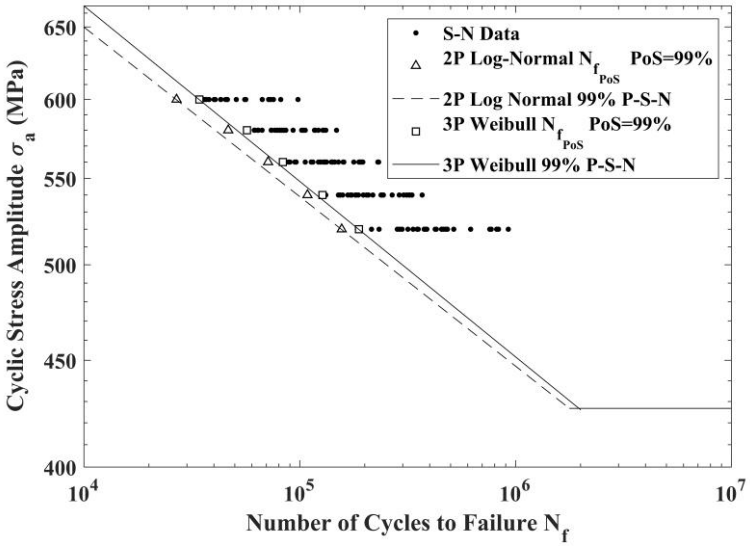

Figure 11: The 99% PoS P-S-N curves generated from minimum fitting using the 2P Log-Normal and 3P Weibull distributions. 4340 S-N dataset reproduced with permission from IHS ESDU [3].

Fatigue damage accumulation was then computed using Miner's Rule based upon the fully-reversed stress amplitudes and the total damage accumulated (D_T) when using each P-S-N curve is shown in Table 8. Assuming that component failure occurs when D_T = 1 (i.e. all available fatigue life has been consumed), the component safe-life could be computed as the inverse of the accumulated fatigue damage (i.e. the number of times that the

'transmission' loading spectrum can be applied before component retirement) [4]. The resulting safe-life values are shown in Table 8 and it can be seen that the use of the 3P Weibull distribution results in a 22.98% increase in the component safe-life value. Therefore, the 3P Weibull distribution has also been shown to reduce the conservatism in component safe-life values, compared to the 2P Log-Normal distribution. The increase in safe-life of 22.98% is also consistent with the increase in $N_{f_{PoS}}$ for the 3P Weibull distribution shown previously in Table 6.

Table 8: Accumulated Fatigue Damage and resulting Safe-Life values for the SAE Keyhole benchmark case study when using 2P Log-Normal and 3P Weibull distributions to construct P-S-N curves.

Fatigue Analysis Results	99% PoS P-S-N Curve Distribution	
	2P Log-Normal	3P Weibull
Accumulated Damage D_T	8.95×10^{-5}	7.28×10^{-5}
Safe-Life	11172	13739

5. Discussion

Firstly, the application of the systematic and robust statistical characterisation methodology defined in Section 2 demonstrated that the 3P Weibull distribution did provide an improved fit to the S-N dataset at a number of stress levels when compared to the 2P Log-Normal, 3P Log-Normal and 2P Weibull distributions. The 2P Log-Normal distribution was also found to provide an acceptable fit to the S-N dataset, validating a long-held assumption about the variability observed in N_f. The statistical methodology demonstrated the desired systematic process and enabled the selection of the 3P Weibull distribution as the final distribution type. The methodology, through the use of the distribution classifications in Table 1 and 2 permitted the 25 individual elements for each candidate distribution to be synthesised into a single class number, simplifying the distribution selection process. The methodology demonstrated that 3P distributions (specifically the 3P Weibull distribution) can provide good fits to S-N datasets and therefore, 3P distributions should be considered as candidate distributions in future statistical characterisation of S-N datasets.

As has been demonstrated in Table 6 and Table 8, the use of the 3P Weibull distribution to statistically characterise the variability in N_f values within datasets can reduce the conservatism within both P-S-N curves and component safe-life values, when compared to the commonly-used and assumed 2P Log-Normal distribution. Within the case study, the 3P Weibull distribution was shown to increase 99% PoS $N_{f_{PoS}}$ values by approximately 20%, whilst the SAE keyhole benchmark life was increased by 22.98%. This therefore demonstrates that the utilisation of the 3P Weibull distribution, providing it has been shown to be the best fit to

the S-N dataset, could reduce conservatism within fatigue design, potentially resulting in more lightweight components with longer design safe-life values. The 2P Log-Normal distribution was found to also be selected as a Class 1 distribution by the statistical characterisation methodology. Therefore, safety-critical industrial sectors may wish to retain the 2P Log-Normal as a result of the increased conservatism it introduces into fatigue design and component safe-life values. In addition, the 3P Weibull distribution presents a number of challenges regarding distribution fitting and P-S-N curve construction, coupled with increased complexity when compared to the 2P Log-Normal distribution, which will be explored in the remainder of this section.

5.1. Challenges of Fitting 3-Parameter Weibull Distribution

Whilst the 3P Weibull distribution was selected as the final distribution type to statistically characterise N_f values, there are currently limitations surrounding its adoption, including poor fits at Stress Levels 1, 5 and 6 and challenges regarding the construction of P-S-N curves for fatigue design.

Firstly, the comprehensive results recorded in Tables 5a and 5b enabled the interrogation of the results from distribution fitting and testing, in order to highlight potential distribution fitting errors and to identify why the 3P Weibull distribution was not consistently a Class 1 distribution at all stress levels. At Stress Level 1, the rejected 3P Weibull distribution was expected to be as a result of the small sample size of N = 10. Work by Wei et al [13], also experienced challenges when fitting 3P Weibull distributions to datasets with small sample sizes. In order to confirm the statistical validity of the 3P Weibull distribution at Stress Level 1, a larger sample size is required. Increased sample sizes would also improve the accuracy of statistical characterisation across the complete S-N dataset. Based on the methods used within the statistical characterisation methodology, the minimum required sample size would be N = 15, as the χ^2 test is inaccurate for sample sizes N < 15 [10]. It should be noted that such a sample size is consistent with the sample sizes suggested for S-N curve generation in ASTM E739 [23].

At Stress Level 2, the 3P Weibull distribution was classified as Class 4 as a result of the MLE distribution parameters producing an Exponential distribution leading to incorrect skew characteristics. This is expected to be as a result of the MLE numerical solution process converging to a local maximum. An improved MLE solution process (e.g. sophisticated optimisation methods) could be employed to assess whether more accurate MLE distribution parameters could be generated. In addition, improved MLE solution processes would reduce the computational expense of generating bootstrap A-D critical values and could also remove the need for the user to interrupt or 'tune' the MLE fitting process to ensure convergence.

At Stress Level 6, it is expected that the reliance on the mean ranking equation for PPLR resulted in inaccurate distribution parameter estimates. A proposed solution to this is to conduct a literature review to identify if there are additional ranking equations beyond those presented in Table A.1 in the Appendix [32]. It can also be observed that all of the ranking equations are in the form shown in Equation 7, where f_1 and f_2 are the constant terms for each ranking equation (e.g. for the median rank equation, f_1 = 0.3 and f_2 = 0.4). f_1 and f_2 can vary between [0, 0.5] and [0, 1] respectively [32].

$$F_i = \frac{i - f_1}{N + f_2} \qquad (7)$$

Therefore, an additional proposed solution is that a 'sweep' through various combinations of f_1 and f_2 values could be performed to identify the pair of values that maximises the correlation coefficient r. As the computational expense of PPLR fitting is negligible, there is little additional resource burden in identifying the optimum f_1 and f_2 values that maximise the correlation coefficient. This is to be explored by the authors during future work.

Finally, it was observed from Table 5a and 5b that the maximum difference values between PPLR and MLE distribution parameter estimates were greatest for the 3P Log-Normal and 3P Weibull distributions. This is expected to be as a result of the greater complexity of the PPLR and MLE fitting methods compared to the 2P distributions, resulting in distribution parameter estimates that exhibit a greater percentage difference for the 3P distributions. The increased maximum difference values for the 3P distributions further highlights the need for using both PPLR and MLE methods simultaneously to validate distribution parameter estimates.

5.2. Challenges of Constructing P-S-N Curves using 3-Parameter Weibull Distribution

Beyond the challenge of fitting 3P Weibull distributions to datasets, there are further challenges regarding the construction of P-S-N curves for fatigue design based upon 3P Weibull distributions. These challenges must be overcome to widen the utility of the 3P Weibull distribution within fatigue design.

Firstly, as discussed in Section 3.1, S-N datasets are typically comprised of both complete failure and run-out or 'censored' data. In order to generate P-S-N curves based upon the entire S-N dataset, censored distribution fitting and selection methods would be required. Whilst censored 'run-out' datasets have been outside the scope of this paper, the authors have investigated both fitting and testing of censored S-N datasets. The authors have observed that the MLE equations are significantly more complex [24], although approximate distribution

parameter estimates can be sourced from modified PPLR approaches [17]. The greatest challenge when working with censored datasets is GoF testing. As the χ^2 GoF test cannot be used for censored datasets, only the A-D test is able to be performed, and critical values are limited (e.g. Normal) and are not always directly applicable to Type I censoring [9]. Therefore, future work should focus on developing GoF tests for Type I censored datasets. The authors have also trialled a parametric bootstrap approach for censored samples [33], with the view to generating A^2_{crit} values using bootstrap samples with the same number of censored observations as the original sample. However, the statistical rigour of this approach is still being assessed. Recent work presented by Toasa Caiza and Ummenhofer details an approach that accounts for run-outs in the 3P Weibull distribution [34]. The implementation of statistical characterisation methods for Type I censored data would also enable more sophisticated P-S-N curve shapes to be fitted to the 99% PoS $N_{f_{PoS}}$. An example of an improved P-S-N curve shape compared to the Basquin P-S-N curves used within the case study (shown in Figure 11), is the power law assumed for the S-N curve as shown previously in Figure 1 [3].

The other challenge regarding the application of the 3P Weibull distribution to construct P-S-N curves is that design P-S-N curves are typically expressed with both a PoS and a Confidence Level (CL) value [1]. CLs are used to produce Confidence Intervals (CIs), which represent the plausible range of the distribution parameter estimates that bracket the population distribution parameter [8]. A CL of 95% represents that for a series of samples, 95% of the confidence intervals would bracket the population distribution parameter [8]. As the CLs impact the width of the CI of a distribution parameter estimate, any values computed using the PDF or CDF (such as PoS $N_{f_{PoS}}$ values) are also affected. Whilst methods and tables for constructing CIs are available for the 2P Log-Normal distribution [35], methods are not currently available for the 3P Weibull distribution [36]. This means that PoS/CL design P-S-N curves cannot be currently constructed for the 3P Weibull distribution. In addition, the width of CIs can give engineers an indication of how well a candidate distribution represents the dataset. Large CIs represent uncertainty in the distribution parameter estimates and therefore, can be used to suggest a poor fit to the dataset. Whilst parametric bootstrap methods can be used to estimate CIs [37], an investigation by the authors showed that the bootstrap samples for the δ threshold parameter had a bi-modal distribution and it is therefore currently unclear whether such an approach would produce accurate CIs for 3P Weibull distributions.

In summary, the statistical characterisation of N_f within S-N datasets can be seen to be a trade-off between reduced conservatism with fatigue design (in the form of P-S-N curves and component safe-life values) and increased challenges within the distribution fitting, GoF testing and construction of P-S-N curves. These

challenges must be overcome to confirm that the 3P Weibull distribution provides the best-fit across the 4340 S-N dataset.

6. Future Work

Regarding future work, the authors intend to consider a wider range of 3P distributions, along with demonstrating the utilisation of the 3P Weibull distribution within a probabilistic approach to fatigue design.

Firstly, it has been identified that there are other 3P distributions within the literature, such as the Generalised Extreme Value (GEV) distribution, which encompasses the Gumbel, Fréchet and Weibull distributions [6]. The GEV distribution has been recently applied to a 300M steel S-N dataset (a modification of 4340 steel) [6] and therefore future work should investigate the application of the GEV distribution to the 4340 S-N dataset.

Finally, the authors' wider research focuses on the development of a probabilistic approach to fatigue design, in order to compute the probability of fatigue failure for a safe-life component [5]. Probabilistic approaches statistically characterise the variability in each design parameter (e.g. materials data, loads data, etc.) and propagate the variability through to the component safe-life [38, 39]. Therefore, the development of a systematic statistical characterisation methodology has supported the need for accurate statistical characterisation within probabilistic design. Regarding the statistical characterisation of S-N datasets within probabilistic fatigue design, due to its better fit to the 4340 S-N dataset at some stress levels, the 3P Weibull distribution should be used to characterise the variability in N_f values within a probabilistic approach. The authors' future work aims to quantify the impact on the estimated probability of failure when using the 2P Log-Normal and 3P Weibull distributions within a probabilistic approach.

7. Conclusions

This paper has presented the application of a novel and systematic statistical characterisation methodology with a view to identifying whether 3-Paramater distributions can improve the statistical characterisation of the number of cycles to failure within stress-life (S-N) datasets. The reduction of conservatism within Probability-S-N (P-S-N) curves and component safe-life values was quantified using a case study based upon the SAE keyhole benchmark and a rich S-N dataset for 4340 steel from the Engineering Sciences Data Unit. The following conclusions can be drawn from the statistical characterisation of the S-N dataset and the case study:

1) The statistical characterisation methodology provides a systematic and traceable process for 'down-selecting' distribution types, based upon validation of distribution fitting and testing using multiple methods.
2) The 3-Parameter Weibull distribution showed a superior fit to the number of cycles to failure at a number of stress levels within the S-N dataset.
3) The common practice of using the 2-Parameter Log-Normal distribution to characterise the variability in the number of cycles to failure has been demonstrated as statistically valid for the S-N dataset.
4) The use of the 3-Parameter Weibull distribution resulted in a 22.98% increase in the component safe-life when used to construct P-S-N curves compared to the 2-Parameter Log-Normal distribution.
5) Further work is required to confirm the statistical validity of the 3-Parameter Weibull distribution, along with the identification of methods to construct confidence intervals for the distribution parameters of the 3-Parameter Weibull distribution.

It should be noted that the results shown within this paper will be specific to the 4340 S-N dataset and should only be used to provide an indication of the potential benefits of improved and systematic statistical characterisation of S-N datasets and the reduced conservatism introduced by the 3P Weibull distribution. It is hoped that the systematic statistical characterisation methodology presented within this paper will assist other engineers when characterising their own S-N datasets.

Acknowledgements

This paper represents work performed as part of the Aerospace Technology Institute funded "Large Landing Gear of the Future" Project (Grant Number:113077) in collaboration with Safran Landing Systems. The authors extend their personal thanks to IHS ESDU for providing permission to reproduce the 4340 S-N dataset [3] within this paper.

References

[1] Hoole, J., Sartor, P., Booker, J. D., Cooper, J. E., Gogouvitis, X. V. and Schmidt, R. K., *Evaluating the Impact of Conservatism in Industrial Fatigue Analysis of Life-Limited Components*, 12th International Fatigue Congress (Fatigue 2018), MATEC Web of Conferences, **165**, 2018.
[2] Suresh S., *Fatigue of Materials*, 2nd Edition, Cambridge University Press, 1998.
[3] Engineering Sciences Data Unit, *Endurance of high-strength steels*, Data Item 04019a, 2006.
[4] Schijve, J., *Fatigue of Structures and Materials*, Springer, 2009.

[5] Hoole, J., Sartor, P., Booker, J. D., Cooper, J. E., Gogouvitis, X. V., Ghouali, A. and Schmidt, R. K., *A Framework to Implement Probabilistic Fatigue Design of Safe-Life Components*, ICAF 2019 – Structural Integrity in the Age of Additive Manufacturing, edited by A. Niepokolczycki and J. Komorowski, Lecture Notes in Mechanical Engineering, Springer International Publishing, Switzerland, 2020, pp. 321-335.

[6] Bag, A., Delbergue, D., Bocher, P., Lévesque, M. and Brochu, M., *Statistical analysis of high cycle fatigue life and inclusion size distribution in shot peened 300M steel*, International Journal of Fatigue, **118**, 126-138, 2019.

[7] Zhu, S.-P., Liu, Q., Lei, Q. and Wang, Q., *Probabilistic fatigue life prediction and reliability assessment of a high pressure turbine disc considering load variations*, International Journal of Damage Mechanics, **27** (10), 1569-1588, 2018.

[8] Bury, K. *Statistical Distributions in Engineering*, Cambridge University Press, 1999.

[9] D'Agostino, R. B., Stephens, M. A., *Goodness-of-Fit Techniques*, STATISTICS: textbooks and monographs, Volume 68, Marcel Dekker Inc., 1986.

[10] Booker, J. D., Raines, M. and Swift, K. G., *Designing Capable and Reliable Products*, Butterworth Heinemann, 2001.

[11] Schijve, J., *Statistical distribution functions and fatigue of structures*, International Journal of Fatigue, **27**, 1031-1039, 2005.

[12] Zhao, Y. X., Gao, Q. and Wang, J.-N., *An approach for determining appropriate assumed distribution of fatigue life under limited data*, Reliability Engineering and System Safety, **67**, 1-7, 2000.

[13] Wei, Z., Luo, L., Gao, L. and Nikbin, K., *Statistical and Probabilistic Analysis of Fatigue Life Data with Two- and Three-Parameter Weibull Distribution Functions*, ASME 2016 Pressure Vessels and Piping Conference, Vancouver, Canada, 2016

[14] Khameneh, M. J., Azadi, M., *Reliability prediction, scatter-band analysis and fatigue limit assessment of high-cycle fatigue properties in EN-GJS700-2 ductile cast iron*, 12th International Fatigue Congress (Fatigue 2018), MATEC Web of Conferences, **165**, 2018.

[15] Castillo, E., Fernández-Canteli, A., Hadi, A. S. and López-Aenlle, M., *A fatigue model with local sensitivity analysis*, Fatigue & Fracture of Engineering Materials & Structures, **30**, 2, 149-168, 2007.

[16] Castillo, E. and Fernández-Canteli, *A Unified Statistical Methodology for Modelling Fatigue Damage*, Springer, 2009.

[17] Nelson, W. N., *Applied Life Data Analysis*, John Wiley & Sons, 1982.

[18] Society of Automotive Engineers, *Fatigue Under Complex Loading: Analyses and Experiments*, 1977.

[19] Wirsching, P. H., *Statistical Summaries of Fatigue Data for Design Purposes*, NASA Contractor Report 3697, 1983.

[20] Goh, Y. M., McMahon, C. A. and Booker, J. D., *Improved utility and application of probabilistic methods for reliable mechanical design*, Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, **223**, 3, 199-214, 2009.

[21] Siddall, J. N., *Probabilistic Engineering Design: Principles and Applications*, Marcel Dekker Inc., 1983.

[22] Gope, P. C., *Determination of Minimum Number of Specimens in S-N Testing*, Journal of Engineering Materials and Technology, **124**, 421-427, 2002.

[23] ASTM International, *E739-10 Standard Practice for Statistical Analysis of Linear or Linearized Stress-Life (S-N) and Strain-Life (ε-N) Fatigue Data*, 2015.

[24] Mann, N. R., Schafer, R. E. and Singpurwalla, N. D., *Methods for Statistical Analysis of Reliability and Life Data*, John Wiley & Sons, 1974.

[25] Evans, J. W., Johnson, R. A. and Green, D. W., *Two- and Three- Parameter Weibull Goodness-of-Fit Tests*, United States Department of Agriculture, FPL-RP-493, 1989.

[26] Stute, W., Manteiga, W. G. and Quindimil, M. P., *Bootstrap Based Goodness-Of-Fit Tests*, Metrika: International Journal for Theoretical and Applied Statistics, **40**, 243-256, 1993.

[27] Guida, M. and Penta, F., *A Bayesian analysis of fatigue data*, Structural Safety, **32**, 64-76, 2010.

[28] D' Anna, G., Giorgio, M. and Riccio, A., *Estimating fatigue life of structural components from accelerated data via a Birnbaum-Saunders model with shape and scale stress dependent parameters*, Procedia Engineering, **167**, 10-17, 2016.

[29] Liu, X.-W., Lu, D.-G. and Hoogenboom, P. C. J., *Hierarchical Bayesian fatigue data analysis*, International Journal of Fatigue, **100**, 418-428, 2017.

[30] Braga, D. F. O., Tavares, S. M. O., da Silva, L. F. M., Moreira, P. M. G. P. and de Castro, P. M. S. T., *Advanced design for lightweight structures: Review and prospects*, Progress in Aerospace Sciences, **69**, 29-39, 2014.

[31] Lawless, J. F., *Statistical Models and Methods for Lifetime Data*, John Wiley & Sons, 1982.

[32] Cunnane, C., *Unbiased Plotting Positions – A Review*, Journal of Hydrology, **37**, 205-222, 1978.

[33] Efron, B., *Censored Data and the Bootstrap*, Journal of the American Statistical Association, **76**, 374, 312-319, 1981.

[34] Caiza Toasa, P. D., Ummenhofer, T., *Consideration of the runouts and their subsequent retests into S-N curves modelling based on a three-parameter Weibull distribution*, International Journal of Fatigue, **106**, 70-80, 2018.

[35] Owen, D. B., *Factors for One-Sided Tolerance limits and for Variables Sampling Plans*, Sandia Corporation SCR-607, 1963.

[36] Wirsching, P. H. *The Application of Probabilistic Design Theory to High Temperature Low Cycle Fatigue*, NASA, CR-165488, 1981.

[37] Sutherland, H. J., Veers, P. S., *The Development of Confidence Limits for Fatigue Strength Data*, 2000 ASME Wind Energy Symposium, Reno, U.S.A., 2000.

[38] Haugen, E. B., *Probabilistic Mechanical Design*, John Wiley & Sons, 1980.

[39] Ocampo, J., Millwater, H., Singh, G., Smith, H., Abali, F., Nuss, M., Reyer, M. and Shiao, M., *Development of a Probabilistic Linear Damage Methodology for Small Aircraft*, Journal of Aircraft, **48**, 6, 2090-2106 (2011).

Appendix

This appendix provides the ranking equations employed within the probability plotting and linear rectification method utilised within the systematic statistical characterisation framework [27]. Table A.1 shows the f_1 and f_2 constants of the ranking equation, in the form of Equation A.1. F_i is the relative cumulative frequency for

the order statistic i. The name of each ranking equation as used in Tables 3, 4, 5a and 5b is also given in Table A.1.

$$F_i = \frac{i - f_1}{N + f_2} \qquad \text{(A.1.)}$$

Table A.1: *The ranking equations and f_1 and f_2 constants to be used for PPLR distribution fitting.*

Ranking Equation (*Abbreviation*)	f_1	f_2
Large Samples (*LS*)	0	0
Hazen Formula (*Hazen*)	0.5	0
Mean Rank (*Mean*)	0	1
Gumbel (*Gumbel*)	0.4	0.2
Extreme Value (*EV*)	0.35	0
Median Rank (*Median*)	0.3	0.4
Normal (*Normal*)	0.3175	0.365