

Statistical distribution for initial crack and number of loading in fatigue crack growth process



Siti Sarah Januri^{1,*}, Zulkifli Mohd Nopiah², Mohd Akramin Mohd Romlay³, Ahmad Kamal Ariffin Mohd Ihsan², Nurulkamal Masseran⁴, Shahrum Abdullah²

¹Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (Kampus Seremban), 70300 Seremban, Negeri Sembilan, Malaysia

²Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 39000 UKM Bangi, Selangor, Malaysia

³Faculty of Mechanical Engineering, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia

⁴Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 39000 UKM Bangi, Selangor, Malaysia

ARTICLE INFO

Article history:

Received 29 April 2017

Received in revised form

25 August 2017

Accepted 28 August 2017

Keywords:

Fatigue crack growth

Statistical distribution

Bootstrap analysis

ABSTRACT

A statistical distribution for crack growth technique is one of the important issues emerging from the fatigue crack propagation process. This study aims to compare three different statistical distributions for providing the best modelling of the fatigue data. The normal, the lognormal and the Weibull distribution are compared for determining a better fit for the variables. Kolmogorov-Smirnov has been chosen as the criterion of the best distribution of the variables. Ten replicate specimens of aluminium alloy A7075-T6 in constant amplitude crack tests were conducted. The number of cycles for the formation of the initial crack and initial crack length were taken as random variables. A Bootstrap approach was applied for ensuring that the chosen distribution was the best representative for this type of variables since small data was incorporated in this analysis, it was not suitable to justify the true population. Thus, the result showed that the lognormal distribution was the best distribution to represent the number of cycles and the length of the initial crack. It was found that whether the normal and lognormal types were suitable for those variables, the lognormal was more conservative for these types of variables. These two variables played the main role in life prediction. Therefore, an analysis of the statistical distribution is highly important. It is believed that these results lead to the significant prediction of fatigue lifetime.

© 2017 The Authors. Published by IASE. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

One of the important criteria to determine the mechanical effectiveness is fatigue occurrences. For example, predicting life estimation of a component in a mechanical field such as the components in an aircraft and train axle is important for minimising the cost of maintenance and inspection (Huynh et al., 2012). There are two phases that contribute to the life of fatigue; the first one is the number of loading cycles required to initiate a crack and secondly, the number of cycles it takes for that crack to propagate and result in failure. Supported by Schijve (2014) and Li et al. (2015a), the process starts with initially, a crack initiation period which is controlled by the

local stress cycles at the material surface and secondly, a crack growth period, where there is a small crack growth that is driven by the cyclic stress intensity around the crack front.

Crack initiation is a term that is used differently by scientists and engineers. For the scientists, studying the fatigue in the laboratory defined initiation as the number of cycles required to form, generate or nucleate the smallest crack that they could detect by any means. Meanwhile, for the engineers, designing and maintaining structures defined as initiation, is likely to mean the smallest crack or an engineering-sized crack that could be determined by a reliable Non-Destructive Evaluation (NDE) technique (Chan, 2010).

One of the factors that determine the durability of the material is the occurrence of crack initiation on the material performance and the advanced process provides a warning that the condition of the material has been threatened (Goszczyńska, 2014). Sangid (2013) studied the importance of crack initiation in

* Corresponding Author.

Email Address: sarahjanuri@ns.uitm.edu.my (S. S. Januri)

<https://doi.org/10.21833/ijaas.2017.010.018>

2313-626X/© 2017 The Authors. Published by IASE.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

predicting the fatigue life of components. Crack initiation dominates the process of fatigue failure, which is the most important factor, as the majority of the lifetime is spent in this stage, while the fatigue crack growth exerts a small influence on the global predictions, mainly in the high-cycle fatigue regime (Correia et al., 2013; Scharnweber et al., 2013).

Identifying the suitable statistical distribution for the initial crack length and the number of loading cycles is required to make a better prediction of the material's lifetime. An appropriate statistical approach should be used when studying the influence of the material microstructure on the fatigue behaviour of structural elements (Glodež et al., 2013). The analysis of the experimental surface crack growth data aims to determine the best statistical distribution and parameters of the relevant variables, allowing a better fit of the experimental observations (Khelif et al., 2008). Min et al. (1996) analysed the distribution of the crack initiation's life and growth life and found that all of these results would be able to provide a solid test foundation for the study of the probabilistic fatigue, probabilistic fracture mechanics, fatigue reliability and its engineering applications.

2. Related works

A statistical distribution of the relevant variables is required in fatigue problems despite the fact that it is difficult to derive a physical description of the fatigue phenomenon quantitatively. Normally, a statistical distribution is assumed and the most well-known function is the Gaussian distribution (also normal distribution). Khelif et al. (2008) found that the Weibull and lognormal distributions were proposed in the literature, regardless of the fact that the choice of appropriate distribution is a very difficult task. Thus, the statistical analysis is important to determine the best distribution of the variable rather than assuming it as a normal distribution.

Previously, Schijve (2005) analysed three distribution functions in order to describe the distribution for fatigue; (i) the normal distribution function, (ii) the 3-parameter Weibull distribution function, and (iii) the lognormal distribution function. A study by Wu and Ni (2007) also considered three statistical distributions namely the Normal, Lognormal and Weibull distributions on aluminium alloy 2025-T6 by comparing the datasets and they found that the random loading cycles and random crack sizes were best fitted by both the lognormal and Weibull distributions, respectively for most cases studied. However, there was an argument, despite the fact that the three-parameter distributions of the Weibull and lognormal types are suitable for lifetime prediction, and the two-parameter Weibull distribution is more suitable for the probabilistic fatigue design (Khelif et al., 2008). Another study by Makkonen (2009) only considered the Weibull distribution in his study, since it provides a slightly better fit with the experimental

data and therefore, the parent population of crack initiation was assumed to be Weibull-distributed.

Verification is required after the process of identification of distribution is performed. A Goodness of fit test is applied for the verification of the decision for choosing the suitable distribution. The Goodness of fit tests like the Kolmogorov-Smirnov or Chi-Square is applied due to the selection of the best distribution of the variables among the comparison distributions. Previously, Kolmogorov-Smirnov's goodness of fit test was used to select the best distribution among five distributions (Normal, 2-parameter lognormal, 3-parameter lognormal, 2-parameter Weibull and 3-parameter Weibull distributions) and based on the test, only the classical lognormal was seen to be suitable for the best fitting distribution for the HDPE structures (Khelif et al., 2008). Another study by Bao et al. (2009) found that the crack growth rate could be considered reasonably to follow the log-normal distribution in all the three stress levels (260MPa, 280MPa and 320MPa) and verified by the value from the Chi-Square goodness of fit test. The purpose of this study is to determine the statistical distribution of the two variables: initial crack length and a random number of loading. In the real applications, the number of loading to reach a certain crack length is estimated as a continuous variable. For experimental purposes, that number of loading is counted as a discrete variable to determine the trend of the fatigue crack growth.

2.1. Small sample size and bootstrap approach

In the engineering field, specifically in fatigue crack growth problem, it is rare to obtain a large sample size due to some limitations in the experimental works. For the statisticians, it is believed that the small data set will cause a problem in producing the results of the parameter estimates. In other words, a further analysis and results from this parameter estimate could be doubtful due to the accurateness of the predecessor analysis. Some phenomenon like the prediction analysis when involving small data set always produce misleading result (Seruga and Nagode, 2015; Bello et al., 2015). Thus, the estimation results from a small sample data set, for instance, the sample mean and sample standard deviation, are not suitable to justify the generalisation of the true population.

Bootstrap resampling is a solid and popular method to resample the original data from a small sample size. The Bootstrap method is introduced to overcome the estimation parameters' problem in generalising the true population. Estimation values from the sample are crucial for ensuring the robustness of the analysis, like a prediction of a lifetime of a material. Bootstrap method is the resampling technique from the initial sample that requires at least 1000 bootstrap resamples is sufficient to obtain accurate confidence interval estimates (Chen et al., 2015). Based on the bootstrap method, standard error and confidence interval

would be calculated to describe the uncertainty in the probabilistic models based on the limited data. The result of confidence interval and standard error from the Bootstrap technique have shown the best values compared to other methods like Monte Carlo and classical method (Khelif et al., 2008). The problem of a small sample size will contribute to a problem that leads to a misleading selection of the best-fit probability distribution. It is proven to handle the problem of large variation when modelling with a small set of samples for ensuring that the objective and reliable analysis would represent the large population (Li et al., 2015a; Bello et al., 2015; Chen et al., 2015; Suo et al., 2015).

For estimating the sampling properties of statistics from the data, bootstrap method is applicable due to simple and straightforward method (Li et al., 2015b). The characteristics of the bootstrap are; there is no need to make any assumptions about the overall distribution, and it can be inferred with the sample data by the computer technology (Suo et al., 2015). Therefore, the technique is useful for small sample size and with an unknown distribution. As a data-driven method, the bootstrap method can evaluate the measurement uncertainty from the poor information without prior information about the probability distribution of the measured data in real-time (Chen et al., 2015).

In this paper, the determination of the statistical distribution is carried out on the initial crack length and the number of cycles for the initial crack formation on aluminium alloy 7075-T6, where three different probability distribution functions have been considered to search for a better fitting: Normal, lognormal, Weibull distributions and subsequently, the Kolmogorov-Smirnov was applied to determine the best distribution for these two variables. However, the experimental results only yielded ten specimens, which is a very small dataset to generalise the actual population. Therefore, the bootstrap approach is applied to make sure that the chosen distribution was representing the variables.

3. Data measurement

In this study, the data was collected from the experiments. The number of loading cycles forming the initial crack was determined and the lengths of the crack initiation were measured. Aluminium alloy of 7075-T6 was chosen as a material in this experiment. This type of material, aluminium alloy of 7075-T6 series, is used to make aircraft, especially the wings on a naval aircraft (Newman et al., 2013).

There were only ten samples in the experiment. The dimensions of the specimens were 160.0 mm length, 60.0 mm wide and 20.0 mm thickness. They were tested at room temperature under constant amplitude loading of 45kN stress load and 0.1 fixed stress ratios in order to observe the fatigue crack growth on the surface of the material. The use of 45kN ensured that it exceeded the endurance limit load, which affirmed that the fatigue would occur for

the dimensions of this material. The length of the initial crack of the ten specimens was measured using the digital calliper and the average length of five measurements was used.

Fig. 1 shows the result of the experiment concerning the formation of semi-elliptical of initial crack length and fatigue crack growth on the material surface. The semi-elliptical shapes were formed due to the nature of the fatigue crack growth and the dimension of the material. These particles or clusters are represented as semicircular shape (Newman et al., 2013; Náhlík et al., 2017). This paper only focuses on the initial crack part, which is the first semi-elliptical flaw.

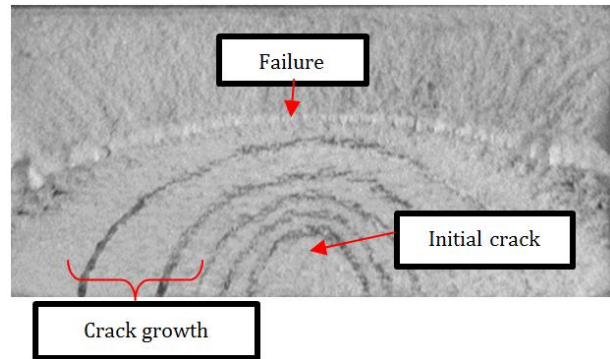


Fig. 1: Material surface of the aluminium alloy 7075-T6 from the experiment

Fig. 2 presents the data of the random loading cycles nucleating the initial crack for 10 specimens of aluminium alloy A7075-T6. The results showed that there was no relationship between the random loading cycles and initial crack since the initial crack formed does not depend on how many loading cycles can go through the material. The majority of them only required fewer random loading cycles but they produced a lengthy initial crack as compared to the others. In the meantime, some of the specimens required more loading cycles to initiate the initial crack such as specimen numbers 4, 6 and 8, even though the experiments were conducted on the same material, dimension, machine and environment. There are some factors that influence the results during the conduct of the experiment, such as machine shutdown or a human factor.

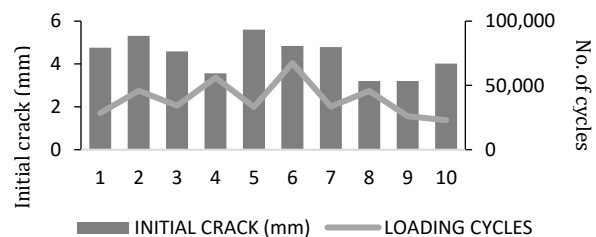


Fig. 2: Combination of the initial crack and random loading cycles

Fig. 3 proves the relationship between the two variables, whereby there is a weak relationship between the number of random loading cycles and the initial crack and is verified by the value of

correlation coefficient, which is only 3.96% correlation. Meanwhile, Fig. 4 illustrates the process of the fatigue life for the ten specimens: starting with the initial crack, crack growth and failure. It indicates that the initial crack lengths vary from 3mm to 6mm for all specimens and there is only one specimen showing random loading cycles more than 100,000 cycles to fracture.

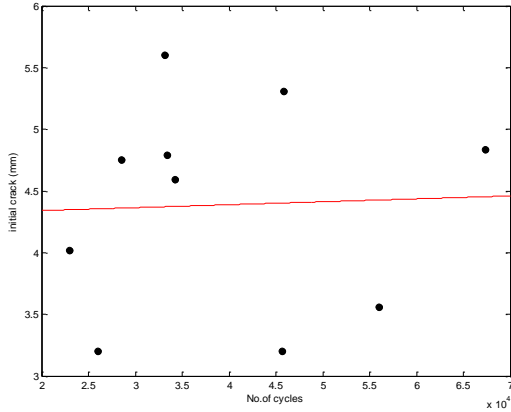


Fig. 3: Scatter chart on the crack initiation versus the number of cycles

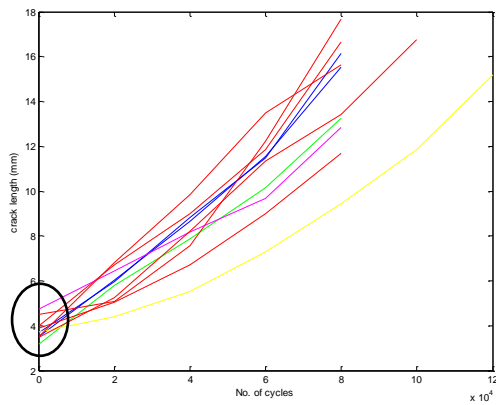


Fig. 4: Crack growth versus the number of cycles

4. Method

The distributions for the initial crack length and the number of cycles to nucleate the initial crack were selected. Based on the previous literature, the selection of the distributions was found suitable for the data lifetime particularly the fatigue crack growth data (Schijve, 2005; Wu and Ni, 2007). The probability density functions, parameters estimation and its statistical properties are given as follows:

- The probability density of a normal distribution is given by

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \mu \geq 0, \sigma \geq 0 \quad (1)$$

Where μ refers to the mean length of the initial crack length (mm) and the number of loading cycles and σ is a standard deviation parameter of the same variables. Parameter estimates can be determined using the maximum likelihood approach, which is derived from the likelihood function. Likelihood function for the Normal distribution is given as

$$\ln L(x; \mu, \sigma) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right] = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left[-\sum_{i=1}^n \frac{(x_i-\mu)^2}{2\sigma^2}\right] \quad (2)$$

Then, the estimators can be derived from the likelihood function by the differentiation of $\ln L(x; \mu, \sigma)$ with respect to μ and σ , which are given as

$$\frac{d[\ln L(\mu, \sigma^2)]}{d\mu} = \frac{\sum_{i=1}^n (x_i - \mu)}{\sigma^2}$$

and

$$\frac{d[\ln L(\mu, \sigma^2)]}{d\sigma} = \frac{-n}{2\sigma^2} + \frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^4} \quad (3)$$

Then, by solving these two partial derivatives, we can obtain the parameter estimation for both μ and σ , which are given as:

$$\hat{\mu} = \frac{\sum_{i=1}^n x_i}{n} \text{ and } \hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n} \quad (4)$$

The parameters, μ and σ^2 , provide the information about the mean and variance of the distribution. Apart from that, the coefficient of variation is σ/μ , coefficient of skewness is equal to zero and the coefficient of kurtosis is equal to three.

- The probability density of the lognormal distribution is given by:

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right], \mu \geq 0, \sigma \geq 0 \quad (5)$$

Where, μ is a shape parameter and σ is a scale parameter of the length of the initial crack and number of loadings. The likelihood function of the Lognormal distribution is given as

$$\ln L(x; \mu, \sigma) = \frac{n}{2} \ln(\sigma^2) - \frac{n}{2} \ln(2\pi) - \ln(\prod_{i=1}^n x_i) - \sum_{i=1}^n \frac{(\ln(x_i) - \mu)^2}{2\sigma^2} \quad (6)$$

Then, by using the same approach, the maximum likelihood estimator for the Lognormal parameters is determined as

$$\hat{\mu} = \frac{\sum_{i=1}^n \ln x_i}{n} \text{ and } \hat{\sigma} = \frac{\sum_{i=1}^n (\ln x_i - \frac{\sum_{i=1}^n \ln x_i}{n})^2}{n} \quad (7)$$

The statistical properties of the Lognormal distribution are provided by the mean, $\exp[\mu + \sigma^2/2]$, the variance, $\exp(\sigma^2)[\exp(\sigma^2) - 1]\exp(2\mu)$, the coefficient of variation, $\sqrt{[\exp(\sigma^2) - 1]}$, the coefficient of skewness, $[\exp(\sigma^2) + 2]\sqrt{[\exp(\sigma^2) - 1]}$ and the coefficient of kurtosis $\exp(4\sigma^2) + 2\exp(3\sigma^2) + 3\exp(2\sigma^2) - 3$ (Krishnamoorthy, 2006).

- The probability density of the Weibull distribution is given by

$$f(x; \alpha, \beta) = \frac{\beta}{\alpha} \left(\frac{x}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{x}{\alpha}\right)^\beta\right], \beta \geq 0, \alpha \geq 0 \quad (8)$$

where, β is a shape parameter and α is a scale parameter for the initial crack length and the

number of loading cycles. The likelihood function of the Weibull distribution can be written as

$$L(x; \alpha, \beta) = \prod_{i=1}^n \frac{\beta}{\alpha^\beta} (x_i)^{\beta-1} \exp \left[-\left(\frac{x_i}{\alpha}\right)^\beta \right]. \quad (9)$$

Then, using the same approach, the maximum likelihood estimator for the Weibull distribution can be determined as

$$\hat{\beta} = \left[\left(\sum_{i=1}^n x_i^\beta \ln x_i \right) \left(\sum_{i=1}^n x_i^\beta \right)^{-1} - n^{-1} \sum_{i=1}^n \ln x_i \right]^{-1}$$

and

$$\hat{\alpha} = \left[\left(\frac{1}{n} \sum_{i=1}^n x_i^\beta \right)^{\frac{1}{\beta}} \right] \quad (10)$$

The maximum likelihood estimator for the parameters of Weibull distributions can be determined numerically using methods such as the Newton-Rapson, scoring, EM algorithm, quasi-Newton and the Nelder-Mead method (Masseran et al., 2013). The statistical properties of the Weibull distribution are provided by the mean, $\Gamma(1 + 1/\beta)$, the variance, (Krishnamoorthy, 2006)

$$\alpha^2 \Gamma(1 + 2/\beta) - [\Gamma(1 + 1/\beta)]^2,$$

the coefficient of variation,

$$\sqrt{[\Gamma(1 + 2/\beta) - [\Gamma(1 + 1/\beta)]^2] / \Gamma(1 + 1/\beta)},$$

and the coefficient of skewness

$$\frac{\Gamma(1+3/\beta) - 3\Gamma(1+1/\beta)\Gamma(1+2/\beta) + 2[\Gamma(1+1/\beta)]^3}{[\Gamma(1+1/\beta) - (\Gamma(1+1/\beta))^2]^{1/2}}$$

Since the data is considered as a small sample size, the bootstrapping analysis is considered to generalise the population. Bootstrap samples are constructed based on a random sampling with replacement from the original dataset. With that, each observation in the bootstrap sample set may appear once, more than once or not at all. With the constructed bootstrap sample set, the statistics of concern (e.g., sample mean, sample SD, and K-S scores) are obtained.

Based on the Li et al. (2015b), the procedure is as follows: Let, original data set, $X = \{x_i, i = 1, 2, \dots, N\}$, then, a bootstrap sample set $B_j = \{B_{1j}, B_{2j}, \dots, B_{Nj}\}$ is constructed by random sampling with replacement from X . From this bootstrap sample set, the sample mean and SD are given as

$$\bar{B}_{Nj} = \frac{1}{N} \sum_{i=1}^N B_{ij} \text{ and } S_{Nj} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (B_{ij} - \bar{B}_{Nj})^2}$$

The above procedure is repeated and the N_s bootstrap sample sets are obtained. For each set of bootstrap samples, the sample mean, standard deviation and the KS values are computed.

Then, the bootstrap mean and SD estimates of the sample mean value, M_N can be calculated by

$$M_{N,mean} \approx \frac{1}{N_s} \sum_{j=1}^{N_s} \bar{B}_{Nj}$$

and

$$\sigma_{M_N} \approx \frac{1}{N_s-1} \sum_{j=1}^{N_s} (\bar{B}_{Nj} - M_{N,mean})^2$$

Similarly, the bootstrap mean and SD estimates of the sample SD, S_N , are derived as below:

$$S_{N,mean} \approx \frac{1}{N_s} \sum_{j=1}^{N_s} S_{Nj}$$

and

$$\sigma_{S_N} \approx \sqrt{\frac{1}{N_s-1} \sum_{j=1}^{N_s} (S_{Nj} - S_{N,mean})^2}.$$

In this paper, $N=5000$ is adopted. A 5000 resampling was selected to ensure that the results were reliable.

5. Results and discussion

The descriptive statistic was analysed to describe the experimental data obtained. The Kolmogorov-Smirnov goodness of fit test (K-S test) was applied to investigate the statistical distribution of the number of the loading cycles and the length of the crack initiation. Tables 1 and 2 show the result of the parameters estimation, the goodness of fit test and the Coefficient of Variation (COV) values for each distribution of the length of initial crack and number of loading cycles for the original data set, respectively.

5.1. Determining probability distribution

As discussed above, the three distributions were also compared in the bootstrapping analysis. Fig. 5 indicates the PDF of the mean of number of cycles on the three distributions: normal, lognormal and Weibull. It clearly shows that the Weibull curve is shifted out from the histogram while, the normal and lognormal curves fit perfectly on the histogram. As the number of specimens tested is small for the experiment, the classical parameter estimates are not suitable. Hence, to verify the value of parameter estimates, two methods were compared to take into account the small sample estimates: the resampling technique and the Bootstrap method. Table 3 shows the obtained mean and the standard deviation, and the confidence intervals for the two variables. The Bootstrap technique was seen to give the narrowest confidence intervals for the given data. It could be interpreted as 95% confident that the average number of cycles of all populations was between 31212 and 47426 cycles, whereas, the mean of initial crack length was between 3.5mm and 4.097mm. Moreover, the standard error and standard deviation, in this technique, were seen to be lower than the other methods. Standard deviation explained the shape of the distribution, which was the distance of the individual data from the mean value, while, the standard error explained the distance of the sample mean to the true mean of the true population.

Figs. 6-9 show the probability density functions of the sample mean and the standard deviation of the initial crack and the number of cycles that nucleate

the initial crack, respectively. Their sampling properties are summarised in Table 4-6.

Table 1: Descriptive statistic for the initial crack length

Distribution	Parameter estimates	K-S test	p-value	Mean	Variance	COV
Normal	$\mu = 3.825$	0.171	0.885	3.825	0.225	0.124
	$\sigma = 0.474$					
Log-normal	$\mu = 1.335$	0.153	0.948	3.827	0.212	0.120
	$\sigma = 0.120$					
Weibull	$\alpha = 8.410$	0.194	0.779	3.808	0.291	0.142
	$\beta = 4.034$					

Table 2: Descriptive statistic for the number of loading cycles

Distribution	Parameter estimates	K-S test	p-value	Mean	Variance	COV
Normal	$\mu = 39328.4$	0.240	0.536	39328.4	2.01×10^8	0.360
	$\sigma = 14165.7$					
Log-normal	$\mu = 10.525$	0.196	0.771	39507.3	1.96×10^8	0.355
	$\sigma = 0.344258$					
Weibull	$\alpha = 3.1091$	0.233	0.575	39386.1	1.92×10^8	0.352
	$\beta = 44035$					

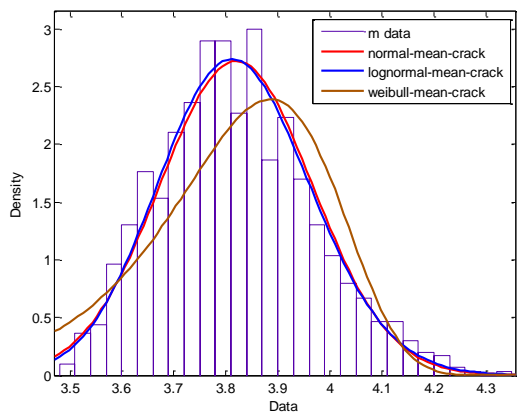


Fig. 5: PDF mean of the initial crack for the three distributions

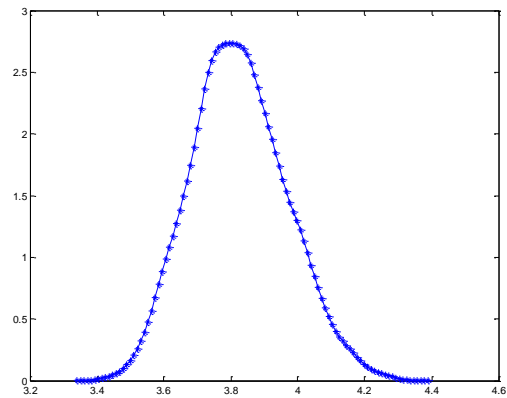


Fig. 6: Mean of the bootstrap initial crack

Table 3: Parameter estimates and confidence intervals by three different methods

	Initial crack length	Number of loading cycles
Classical	3.8243	39328.4
Standard Deviation	0.476441217	14165.6681
Standard Error	0.150663942	4479.577577
Confidence Interval	3.528998674	30548.42795
	4.119601326	48108.37205
Resampling	4.1254	39048.3
Standard Deviation	0.518029214	13229.06913
Standard Error	0.16381522	4183.398977
Confidence Interval	3.804322167	30848.838
	4.446477833	47247.762
Bootstrap	3.825605	39319.65
Standard Deviation	0.440736	13085.3
Standard Error	0.13835	4136.172
Confidence Interval	3.554429	31211.97
	4.09676	47425.77

Tables 4-5 exhibit the descriptive analysis of the three selected distributions of the initial crack length and number of cycles based on the mean and the standard deviation values from the bootstrap data, respectively. This includes all the parameter estimation values, including the COV values, which were slightly lower than the original data in Table 1 and Table 2. Based on these results, the Weibull distribution was seen to be unsuitable for both the variables even for a small or a large-scale data, and

the COVs of the mean values for the three candidate distributions of both variables were comparable.

Table 4: Descriptive statistic for the initial crack length based on the bootstrap dataset

Distribution	Parameter estimation	Mean	Var	COV
Normal	$\mu = 3.820$	3.820	0.021	0.039
	$\sigma = 0.147$			
Lognormal	$\mu = 1.340$	3.820	0.021	0.038
	$\sigma = 0.0382$			
Weibull	$\alpha = 3.892$	3.810	0.035	0.049
	$\beta = 25.486$			

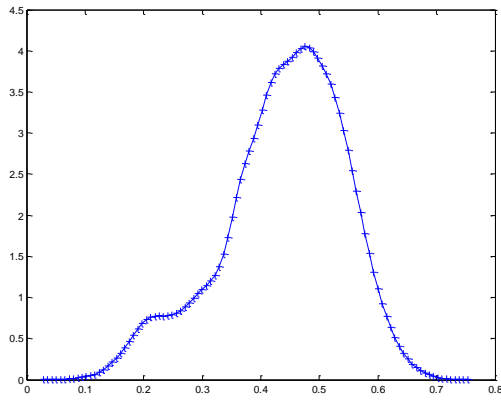


Fig. 7: Standard deviation of the bootstrap initial crack

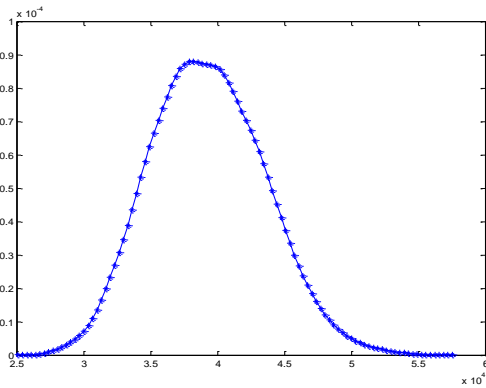


Fig. 8: Mean of the bootstrap number of cycles

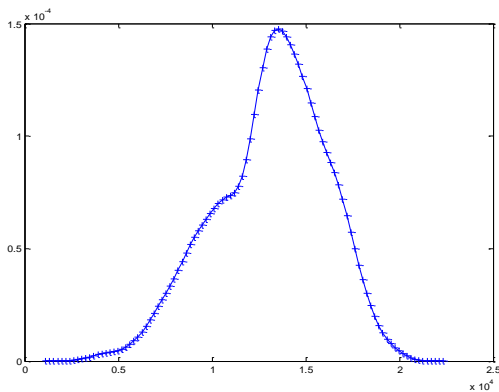


Fig. 9: Standard deviation of the bootstrap number of loading cycles

Figs. 10-11 show the probability density functions of the KS values associated with the three candidate distributions for the initial crack and number of loading cycles, respectively. The sampling properties of the KS values for the original data are also listed in Table 1 and Table 2. The COVs of the KS values for the normal and lognormal candidate distributions of initial crack and number of cycles are comparable. However, the COV of the KS value for the Weibull distribution was seen to be significantly different from the data. The best-fit distribution can be identified from the KS values associated with the three candidate distributions for each bootstrap sample.

Table 5: Descriptive statistic for the random loading cycles based on the bootstrap dataset

Distribution	Parameter estimation	Mean	Var	COV
Normal	$\mu = 39291.9$	39291.9	1.80	0.11
	$\sigma = 4237.68$		$\times 10^7$	
Lognormal	$\mu = 10.573$	39292.6	1.81	0.11
	$\sigma = 0.108$		$\times 10^7$	
Weibull	$\alpha = 41216.5$	39136.1	2.41	0.13
	$\beta = 9.563$		$\times 10^7$	

Table 6: Sampling properties of the mean, standard deviation and COV values for the KS values for two variables

	Initial crack			Number of cycles		
	Mean	SD	COV	Mean	SD	COV
Mean value	3.828	0.143	0.037	39386	4258.7	0.108
Std. dev	0.442	0.104	0.236	13104	3001.1	0.229
Normal	0.167	0.031	0.184	0.182	0.031	0.170
Log-normal	1.806	0.181	0.101	1.719	0.164	0.095
Weibull	0.180	5.518	30.647	0.195	5.718	29.305

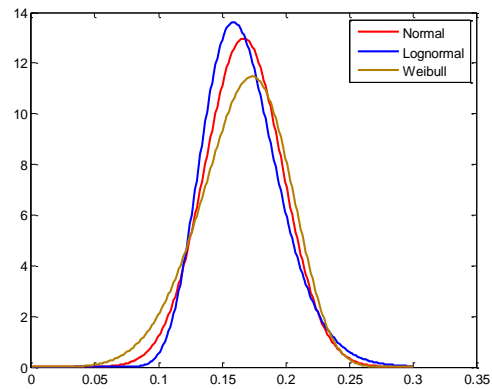


Fig. 10: KS-values for initial crack length

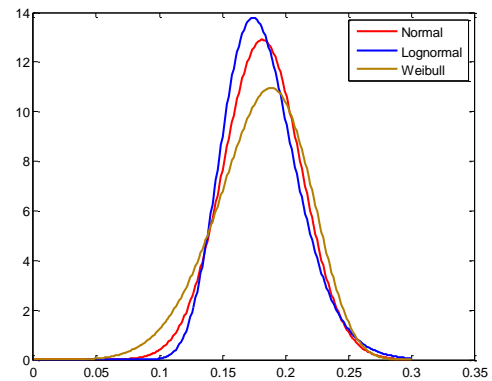


Fig. 11: KS values for number of loading cycles

Note that the distribution cannot be determined as the best-fit distribution due to the small sample size. In this respect, the bootstrap method shows a great advantage over the traditional methods.

Fig. 12 illustrates a comparison of the probability plot for all three candidate's distributions of the initial crack length. The behaviour of all distributions is quite similar which is between 3.65mm and 4.1 mm. There are two intersections illustrated in the graph, which explained the likelihood of the initial crack formation happening at 25% and 90%, respectively, which are around 3.8mm and 4.1mm.

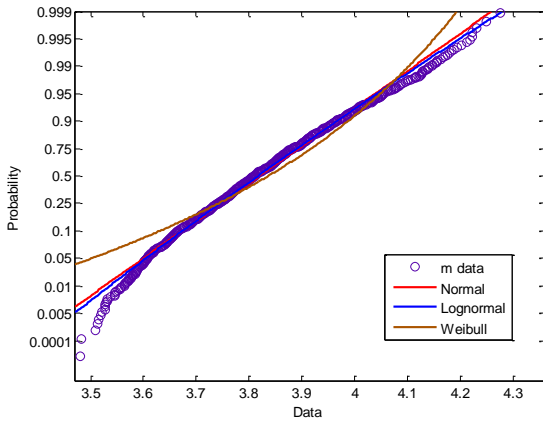


Fig. 12: Probability plot of the initial crack length for the three distributions

Fig. 13 and Fig. 14 exhibit a comparison for all the distributions for the number of loading cycles. Fig. 13 presents the fact that all distributions have agreed to predict the 25% likelihood of initial crack length being formed during 35,000 loading cycles, and 90% likelihood when there are 45,000 loading cycles.

From the observations, the Normal and Weibull distributions demonstrate a linear relation or almost a similar behaviour in explaining the distribution of the data. Conversely, the Lognormal demonstrates the curve relation. Naturally, the curve relation is better in presenting the real situation problems. Fig. 14 supports the lognormal distribution, showing that it is the best distribution to represent the variable.

6. Conclusion

A constant amplitude-loading test was conducted on an aluminium alloy A7075-T6 in order to identify the statistical distribution. Based on the experimental findings, there is no relationship between random loading cycles and the length of the initial crack. The Normal, lognormal and the Weibull distributions were selected and compared and the selection of the best distribution was verified by the Kolmogorov-Smirnov test.

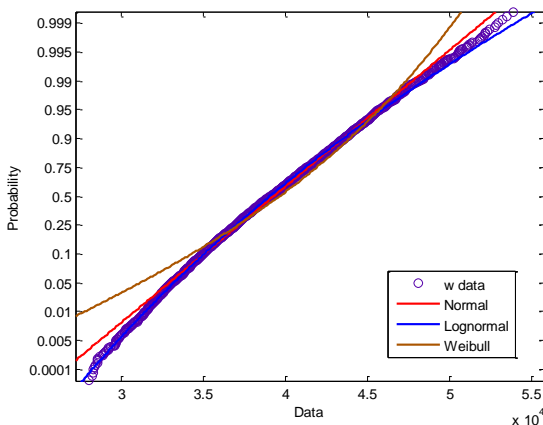


Fig. 13: Comparison of the probability plot of the three distributions of random loading cycles

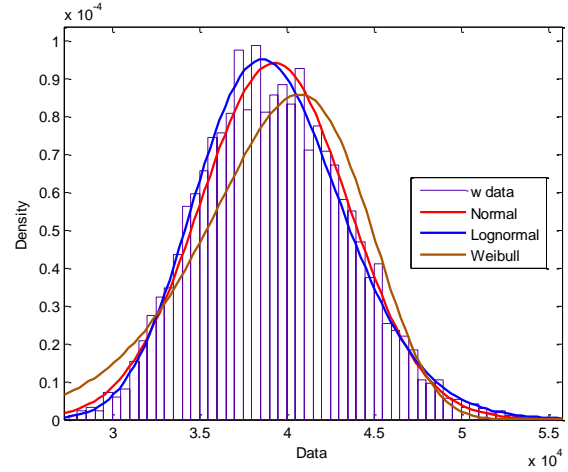


Fig. 14: Comparison of the probability density function of the three distributions of number of loading cycles

The major conclusions of this investigation have been summarised below. First, two variables were selected as the main factors in Firstly, the two variables were selected as the main factors in explaining the fatigue crack growth problems: initial crack length and random loading cycles. Secondly, the numerical results showed that the bootstrap method could effectively model the variations of the sample statistics and the KS scores, even though the sample size of the original dataset was as small as N=10. Thirdly, by the bootstrap method, it was significantly seen that the best distribution of the two variables in the crack problem, was the lognormal distribution, which was conducted graphically and quantitatively. It is believed that the outcomes from this study would be able to model the fatigue time accurately. This study considered the random loading and the initial crack length. To predict the fatigue life, there are three important processes. Therefore, the statistical distribution for the fatigue life should be considered for future research to ensure the modelling of the fatigue life distribution more accurately.

Acknowledgment

Special thanks to the research grant FRGS/2/2013/KT06/UKM/1 for the financial support.

References

Bao R, Zhang JY, Yang HQ, and Fei BJ (2009). Stochastic characteristics of fatigue crack propagation for TA15 titanium alloy. In: Aliabadi MH, Abela S, Baragetti S, Guagliano M, and Lee HS (Eds.), *Advances in fracture and damage mechanics VIII*: 873–876. Trans Tech Publications, Stafa-Zurich, Switzerland.

Bello AO, Bamiduro AT, Chuwkwu AU, and Osowole IO (2015). Bootstrap nonlinear regression application in a design of an experiment data for fewer sample size. *International Journal of Research (IJR)*, 2(2): 428–441.

Chan KS (2010). Roles of microstructure in fatigue crack initiation. *International Journal of Fatigue*, 32(9): 1428–1447.

Chen Y, Jiang S, Yang J, Song K, and Wang Q (2015). Grey bootstrap method for data validation and dynamic uncertainty

- estimation of self-validating multifunctional sensors. *Chemometrics and Intelligent Laboratory Systems*, 146: 63–76.
- Correia JaFO, De Jesus AMP, and Fernández-Canteli A (2013). Local unified probabilistic model for fatigue crack initiation and propagation: Application to a notched geometry. *Engineering Structures*, 52: 394–407.
- Glodež S, Šori M, and Kramberger J (2013). Prediction of micro-crack initiation in high strength steels using Weibull distribution. *Engineering Fracture Mechanics*, 108: 263–274.
- Goszczyńska B (2014). Analysis of the process of crack initiation and evolution in concrete with acoustic emission testing. *Archives of Civil and Mechanical Engineering*, 14(1): 134–143.
- Huynh KT, Castro IT, Barros A and Bérenguer C (2012). Modeling age-based maintenance strategies with minimal repairs for systems subject to competing failure modes due to degradation and shocks. *European Journal of Operational Research*, 218(1): 140–151.
- Khelif R, Chateaufneuf A, and Chaoui K (2008). Statistical analysis of HDPE fatigue lifetime. *Meccanica*, 43: 567–576.
- Krishnamoorthy, K (2006). *Handbook of statistical distributions with applications*. Chapman and Hall/CRC, Boca Raton, USA.
- Li DQ, Tang XS, and Phoon KK (2015a). Bootstrap method for characterizing the effect of uncertainty in shear strength parameters on slope reliability. *Reliability Engineering & System Safety*, 140: 99–106.
- Li L, Shen L, and Proust G (2015b). Fatigue crack initiation life prediction for aluminium alloy 7075 using crystal plasticity finite element simulations. *Mechanics of Materials*, 81: 84–93.
- Makkonen M (2009). Predicting the total fatigue life in metals. *International Journal of Fatigue*, 31(7): 1163–1175.
- Masseran N, Razali AM, Ibrahim K, Zaharim A, and Sopian K (2013). The probability distribution model of wind speed over east Malaysia. *Research Journal of Applied Sciences, Engineering and Technology*, 6(10): 1774–1779.
- Min L, Qing S, and Qing-Xiong Y (1996). Large sample size experimental investigation on the statistical nature of fatigue crack initiation and growth. *International Journal of Fatigue*, 18(2): 87–94.
- Náhlík L, Pokorný P, Ševčík M, Fajkoš R, Matušek P, and Hutař P (2017). Fatigue lifetime estimation of railway axles. *Engineering Failure Analysis*, 73: 139–157.
- Newman JC, Anagnostou EL, and Rusk D (2013). Fatigue and crack-growth analyses on 7075-T651 aluminum alloy coupons under constant- and variable-amplitude loading. *International Journal of Fatigue*, 62: 133–143.
- Sangid MD (2013). The physics of fatigue crack initiation. *International Journal of Fatigue*, 57: 58–72.
- Scharnweber M, Tirschler W, Oertel CG, and Skrotzki W (2013). Initiation and propagation of short cracks in austenitic-ferritic duplex steel. *Materials Science and Engineering A*, 583: 52–60.
- Schijve J (2005). Statistical distribution functions and fatigue of structures. *International Journal of Fatigue*, 27(9): 1031–1039.
- Schijve J (2014). The significance of fatigue crack initiation for predictions of the fatigue limit of specimens and structures. *International Journal of Fatigue*, 61: 39–45.
- Šeruga D and Nagode M (2015). A method for long-term creep-rupture strength prediction based on a small sample of experimental results, smoothed bootstrapping and time-temperature parameters. *Materials and Design*, 67(0): 180–187.
- Suo H, Gao J, Gao Z, Jiang H, and Wang R (2015). Bootstrap and maximum entropy based small-sample product lifetime probability distribution. *International Federation of Automatic Control*, 48(3): 219–224.
- Wu WF and Ni CC (2007). Statistical aspects of some fatigue crack growth data. *Engineering Fracture Mechanics*, 74(18): 2952–2963.