Fitting Method of Small Sample \( P-S-N \) Curve Based on Weibull Distribution

T Huang\(^1\), Z W An\(^1\), Q Ma\(^1\) and M Q Han\(^1\)

\(^1\) School of Mechanical and Electronical Engineering, Lanzhou University of Technology, Lanzhou, China

E-mail: anzongwen@163.com

Abstract. In order to obtain the probabilistic stress-fatigue life curve of materials with small samples, a small sample \( P-S-N \) curve fitting method is proposed. This method is based on the premise that Weibull distribution can be applied to fatigue test data. The method first uses the sample mean fitting to obtain the median \( S-N \) curve. Secondly, constructs the random variable function based on the Weibull distribution characteristics, in addition, the sample median is used as the initial value of iteration to perform perturbation optimization to determine the shape parameter \( \alpha \) and the scale parameter \( \beta \) of Weibull distribution. The fitting of \( P-S-N \) curve is realized. Taking the fatigue data of aluminum alloy 2425-T3 as an example, the results show that the intercept relative error and slope error are less than 0.5% by comparing the fitting results of the proposed method with those of the traditional method, and the intercept error between the proposed method and the traditional method is less than 1.5% and the slope error is small when designing three groups of small sample fitting schemes (7-7-7-7, 5-5-5, 3-3-3) at 3.0%. The proposed method can obtain high-precision \( P-S-N \) curve fitting results in the case of small samples, which have a certain significance for engineering applications.

1. Introduction

The fitting method of probability stress fatigue life curve (\( P-S-N \) curve) has been the key problem of fatigue reliability of mechanical structure [1]. According to the sampling theory, the larger the number of test samples, the higher the accuracy of test evaluation, and the more truly and effectively the product quality level can be reflected [2,3]. However, due to the limitation of fatigue test cost and test time, the number of test samples cannot meet the requirements of sampling theory when the fatigue life of large-size structural parts or more expensive parts is statistically inferred, which requires an effective \( P-S-N \) curve modeling method for the existing small sample data [4-6].

Among existing \( P-S-N \) curve modeling methods for small samples, one is Bayes statistics with prior information; the other is to improve the large sample statistical method to make it suitable for small sample data analysis. Lv et al. obtained the statistical parameters of fatigue life distribution based on Bayes method, combined the prior information of life distribution parameters with small sample information, and obtained the posterior information according to Bayes formula to infer the life distribution parameters [7]. Guida et al. proposed the linear regression analysis technology of fatigue life curve under the condition of lognormal distribution [8]. Castillo et al. proposed a crack growth model based on the Buckingham theorem and then the \( P-S-N \) curves were obtained from the crack growth curves [9]. Amraoui et al. proposed a probabilistic model of \( P-S-N \) curves, which could
determine the distribution of fatigue strength using relatively small number of samples [10]. Gao et al. considered that if the P-S-N curve was extended to the region of high stress or low fatigue life, the curve will intersect at one point. From the geometric point of view, a new method for calculating P-S-N curve under the condition of small sample is proposed [11]. Xie et al and Bai et al. based on the assumption that the fatigue life at any stress level obeys the lognormal distribution and the standard deviation of fatigue life is linear with the stress level, and based on the principle of the consistency of quantiles and the principle of sample aggregation, discussed the problem of small samples of metal materials [12,13]. Under the assumption that the standard deviation of fatigue life is linear with the stress level, Bai et al. make the fatigue life of low stress level equivalent to high stress level, and realized the fitting method of P-S-N curve under the Weibull distribution of fatigue life [14]. There are two deficiencies in the above studies on the modeling of P-S-N curve with small samples: firstly, there is nonlinearity between the standard deviation of fatigue life and the stress level [4], gathering the low stress fatigue test data under the high stress according to the linear relationship will inevitably affect the accuracy of reliability analysis. Secondly, whether it is reasonable to use the failure rate curve of lognormal distribution to describe the fatigue life remains to be certified [15]. It can be seen from the damage function that Weibull distribution has a strong adaptability [16]. Although Weibull distribution cannot describe the ‘bathtub’ curve, the failure mode can be well described by properly changing its parameters, which are also one of the main reasons why Weibull distribution is widely used in the life test. However, there are few specific statistical analysis and fitting methods for the small sample P-S-N curve whose fatigue life follows Weibull distribution.

Therefore, this paper presents a small sample P-S-N curve modeling method for Weibull distribution of fatigue life. According to the characteristics of Weibull distribution, the equivalent data of Weibull distribution with shape parameter $\alpha$ and scale parameter $\beta = 1$ are obtained by using the experimental data at all stress levels. Combined with the median S-N curve and the perturbation optimization method, the shape parameters and scale parameters of fatigue life distribution at all stress levels are calculated to obtain the S-N curve at any survival probability, and the fatigue of aluminum alloy 2425-T3 material is used the data of labor test is taken as an example to verify the method proposed in this paper.

2. P-S-N curve fitting principle

When the same specimen is tested at different stress levels, and the obtained fatigue life data had the same probability points at each stress level. In other words, the fatigue life of the same specimen under different stress levels must be located at the same probability quantile in the parent distribution of each life. This principle is life probability quantile consistency principle [12] expressed as (figure 1).

\[ P(n_{ij}) = P(n_{kj}) \]  

where $n_{ij}$ represents the fatigue life of the sample labeled $j$ at the level of stress $i$; $n_{kj}$ represents the fatigue life of the sample labeled $j$ at the level of stress $k$; $P$ represents the failure probability.

The Weibull distribution is one of the most widely used life distribution in reliability engineering [17]. It is assumed that the fatigue life of materials obeys the two-parameter Weibull distribution, and its cumulative probability distribution function is

\[ F(n) = 1 - \exp \left( -\left( \frac{n}{\beta} \right)^{\alpha} \right) \]  

where $n$ is fatigue life, $\alpha$ is the shape parameter and $\beta$ is the scale parameter.
Based on the principle of consistency of life probability quantiles, the relationship between the two-parameter Weibull distribution and the same fatigue life probability quantiles with different stress levels is as follows:

\[
1 - \exp\left[-\left(\frac{n_{ij}}{\beta_i}\right)^{\alpha_i}\right] = 1 - \exp\left[-\left(\frac{n_{kj}}{\beta_k}\right)^{\alpha_k}\right]
\]

(3)

\[
\left(\frac{n_{ij}}{\beta_i}\right)^{\alpha_i} = \left(\frac{n_{kj}}{\beta_k}\right)^{\alpha_k}
\]

(4)

where \(\alpha_i, \beta_i\) are shape parameters and scale parameters under the \(i\)-th stress level respectively; \(\alpha_k, \beta_k\) are shape parameters and scale parameters under the \(k\)-th stress level respectively.

Generally, it is assumed that the shape parameter \(\alpha\) of the fatigue life is determined by the characteristics of the material, and the scale parameter \(\beta\) depends on the applied load [17]. Therefore, for the same material, the shape parameter of the fatigue life distribution under any stress level is equal, that is, \(\alpha_i = \alpha_2 = \cdots = \alpha_i = \cdots\), according to equation (4), it can be obtained

\[
\frac{n_{ij}}{\beta_i} = \frac{n_{2j}}{\beta_2} = \cdots = \frac{n_{ij}}{\beta_i} = \cdots
\]

(5)

According to equation (2), we find that the failure probability is 0.6321 when the fatigue life is equal to the scale parameter (i.e. \(n = \beta\)). When the fatigue life is equal to the mean (median) \(\bar{n}\), the failure probability is 0.5.

Figure 2 indicates that if we can find the relationship between the mean value of life (median) \(\bar{n}\) and the scale parameter \(\beta\), we can get the parameter value of Weibull distribution, and then fit the \(P-S-N\) curve.
Figure 2. Probability density characteristics of Weibull distribution.

According to equation (5), we construct \( y = \begin{bmatrix} N_i / \beta_i, N_i / \beta, \ldots \end{bmatrix} \), and take it into equation (1) to obtain that \( y \) obeys Weibull distribution has a shape parameter of \( \alpha \) and a scale parameter of 1 (hereinafter referred to as the Weibull \((\alpha, 1)\)).

\[
F(y) = 1 - \exp(-y^\alpha)
\]

(6)

where: \( \beta_i \) is the scale parameter under the \( i \)-th level of stress, and \( N_i \) is the fatigue life value of all specimens under the \( i \)-th level of stress.

According to the probability density characteristic diagram of Weibull distribution in figure 2, using the principle of perturbation optimization, we make the following assumptions: the initial value is \( \beta_i = \bar{n}_i \), the step size is \( \eta = 0.001 \), and the number of iterations is 4 (the value starts from 0), then

\[
\beta_i = \bar{n}_i + \eta \Delta
\]

(7)

Take the value from equation (7) into equation (5) for the calculation, get \( \beta_i = \bar{n}_i \); \( \beta_i = \bar{n}_i \), and then take it into equation (6) for the calculation to see whether equation (6) obeys Weibull \((\alpha, 1)\) distribution. If it is not obeyed, bring \( \beta_i \) into equation (7) again, get a new value of \( \beta_i \), repeat the instructions in the previous step and bring it into equation (5), get a new \( \beta_i = \bar{n}_i \); \( \beta_i = \bar{n}_i \), until equation (6) is obeying Weibull \((\alpha, 1)\) distribution, and the iteration stops. But we know that for \( \beta_0 \), it can only approach to 1 infinitely, and it cannot be equal to 1 (unless the number of samples is very large), so we use the perturbation optimization principle, when the error between the \( \beta_0 \) value of equation (6) and 1 is less than the predefined error (such as 0.001 [12]), it is the termination condition of our iteration, at this time \( \beta_0, \beta_0, \ldots \) is our solution, and the formula is

\[
\left| \frac{\beta_0 - 1}{\beta_0} \right| < 0.001
\]

(8)

where: \( \beta_0 \) is the value of \( \beta_i \) after the first iteration, and \( \beta_0 \) is the value of \( \beta \) fitted by equation (6).
3. Specific method of P-S-N curve fitting

When the fatigue life follows two-parameter Weibull distribution, several different test schemes are designed to adapt to different situations. The basic scheme is to determine the four constant cyclic stress levels stress. In the fatigue test, because the dispersion of the fatigue life of the material under the high stress level is small, and the fatigue test process time is short, X test pieces are selected at the high stress level, and the stress of each level is other the number of subsamples was x (X ≥ x).

Taking 30 fatigue specimens in the fatigue experiment scheme 15-5-5-5 as examples, determine a reasonable value of η, shape parameter α and scale parameter β under each stress level are determined, and the method of fitting the P-S-N curve of fatigue life is as follows:

(1) Select four stress levels roughly covering the stress range of high cycle fatigue, i.e. the fatigue live ranging from $10^4 \sim 10^6$ times. Test 15 specimens at the highest stress level $S_1$, denote the fatigue lives at $n_{1,j}(j=1 \sim 15)$; test 5 specimens at each of the other three stress levels and denote the fatigue lives as $n_{i,j}(i=2 \sim 4, j=1 \sim 5)$.

(2) Estimate the mean value of logarithmic life at each stress level (mathematical expectation)

$$\bar{n}_j = \frac{1}{m_j} \sum_{i=1}^{m_j} n_{j,i} \quad j = 1, 2, 3, 4$$

where, $m_j$ is the number of specimens at the level of the first level of j stress.

(3) According to the logarithm life mean value $\bar{n}_j (1 \sim 4)$ under each stress level estimated by equation (9), the relationship between stress and life is power-law. The median $S-N$ curve equation is fitted by the least square method, and the logarithm life mean value $\bar{n}_1$, $\bar{n}_2$, $\bar{n}_3$, $\bar{n}_4$ under four stress levels are solved by the regression equation.

(4) Taking $\bar{n}$ values obtained from into the equation (7). Since $\beta_1$, $\beta_2$, $\beta_3$, $\beta_4$ is not the correct act $\beta_i$ values, but iterates from $\bar{n}$ to $\beta_i$ steps by step, it follow the Weibull distribution of equation (6). Only when equation (6) is the subject to Weibull ($\alpha$, 1) can we show that $\beta_i$ has been found. Finally, $\beta_i$ values can be discovered by perturbation optimization.

Because the fitting of $\beta_i$ is obtained from all the above data, the accuracy is high. The lower the stress, the higher the dispersion of the material fatigue life, and the larger the error of the fitted life mean, the corrected life mean can be obtained from equation (5) according to known $\beta_i$, so that the fitted median $S-N$ curve accuracy will be higher than that of the median $S-N$ curve directly fitted by the original small sample.

(5) The scale and shape parameters of fatigue life at different stress levels are obtained, the points of fatigue life with different failure probabilities at each stress level can be calculated, and then fit the P-S-N curve.
4. Calculation example and analysis discussion

In order to verify the P-S-N curve fitting method for small samples of two parameter Weibull distribution mentioned in this paper, the fatigue data of aluminum alloy 2425-T3 material in reference [12] are taken as an example, and the P-S-N curve fitting method (hereinafter referred to as the ‘traditional method’) for the same probability life points of each stress level is obtained by fitting the fatigue life data with all the fatigue life data of each stress level Through comparative analysis, the fatigue life data are shown in table 1.
Table 1. Fatigue life test data of aluminium alloy 2425-T3 by group method.

| Stress (MPa) | Number of samples | Logarithm fatigue life                                      |
|-------------|------------------|------------------------------------------------------------|
| 400         | 14               | 4.477, 4.400, 4.426, 4.462, 4.592, 4.411, 4.447, 4.402,    |
|             |                  | 4.665, 4.475, 4.458, 4.551, 4.525, 4.641                  |
| 350         | 15               | 4.784, 4.842, 4.776, 4.813, 4.813, 4.860, 4.798, 4.776,  |
|             |                  | 4.758, 4.770, 4.755, 4.837, 4.736, 4.842, 4.796          |
| 300         | 15               | 5.028, 5.074, 5.016, 4.894, 4.993, 5.071, 5.024, 5.035,  |
|             |                  | 4.954, 5.039, 5.098, 5.057, 5.092, 5.082, 5.005          |
| 200         | 15               | 5.603, 5.544, 5.528, 5.630, 5.594, 5.540, 5.581, 5.548,  |
|             |                  | 5.426, 5.567, 5.554, 5.627, 5.630, 5.596, 5.626          |

If life data of each stress level is limited and there are only five fatigue life test data, then the fatigue life test data of a small sample (5-5-5-5) scheme is formed. The method in this paper is used to fit the data, and the final shape parameter $\alpha$ and scale parameter $\beta$ are determined by perturbation optimization principle.

Figure 4. Fit the 2425-T3 median S-N curve according to different sample data

Figure 4 shows the median S-N curve of aluminum alloy 2425-T3 material fitted by the original test data according to the traditional group method, the median S-N curve fitted by the proposed method with the original test data, and the median S-N curve fitted by this method with the life sample data of test scheme (5-5-5-5). Among them, according to all 59 test samples, the median S-N curve fitted by the method proposed in this paper basically coincides with the median S-N curve fitted by the traditional method, with a relative error of 0.41% for slope and 0.23% for intercept. For the test scheme (5-5-5-5), with the decrease of the number of samples, the accuracy is reduced, but the difference between the fitting curves is not significant. Compared with the median S-N curve fitted by the traditional method, the relative error of the slope is 0.63%, and the relative error of the intercept is 0.44%. It can be observed that the method proposed in this paper is effective in fitting $P$-$S$-$N$ curve.
Figure 5. $P-S-N$ curves with 2425-T3 confidence of 1% and 99% based on different samples

Figure 5 shows the $P-S-N$ curve of aluminum alloy 2425-T3 material with confidence of 95%, survival probability of 99% and confidence of 95%, survival probability of 1% based on the traditional group method, and the $P-S-N$ curve of 2425-T3 material with confidence of 95%, survival probability of 99% and confidence of 95% and survival probability of 1% based on the life sample data of test scheme (5-5-5-5). According to figure 5, there is a very small amount of data outside the range of 1% reliability and 99% reliability fitted by the traditional group method, indicating that the traditional group method is dangerous. For all the data of this method are within the range of 1% and 99% of the reliability of the fitting scheme (5-5-5-5), it shows that the method proposed in this paper has good applicability.

Figure 6. 2425-T3 median $S-N$ curves fitted according to different small sample data

In checking to see the applicability of the method set out in the present paper, two kinds of small sample test schemes (7-7-7-7) and (3-3-3-3) are designed for comparative study. The result is shown in figure 6. The accuracy of scheme (7-7-7-7) is slightly lower than that of scheme (5-5-5-5), which is mainly due to the fact that a small number of samples far away from the life of the parent are also calculated in the random sampling scheme (7-7-7-7), The accuracy of fitting is reduced. The low
accuracy of the scheme (3-3-3-3) is mainly due to the lack of test data, but the slope error and intercept error of either scheme are less than 3.0% and 1.0%, respectively.

In this paper, the $P-S-N$ curve test data of gear bending fatigue in reference [7] are also fitted, and three test schemes (8-7-8), (8-5-5) and (8-3-3) in the original text are respectively carried out. The results are shown in figure 7. The median $S-N$ curve obtained by the two-parameter Weibull distribution fitting method proposed in this paper is fitted with the median $S-N$ curve of 95% confidence in slope and intercept by the traditional method the errors are 1.4334%, 0.1273%, 4.8985%, 0.4781% and 5.8790%, 0.5999% respectively. This shows that this method is feasible.

Figure 7. Fit the median $S-N$ curve of gear bending fatigue test according to different small sample data.

Based on the analysis of the above test data, the fitting accuracy of the $P-S-N$ curve fitting method of the two-parameter small sample Weibull distribution proposed in this paper is relatively high. When the number of samples is more and closer to the life of the parent, the fitting effect of the $P-S-N$ curve is more obvious, but the error will be too large due to the deviation of individual samples from the parent itself, which are inevitable errors. In order to fit the $P-S-N$ curve with high precision as much as possible, the only way is to increase the number of samples, but for the limited number of samples, this method has smaller error, which is a better choice.

5. Conclusions
In this paper, a new fitting method of $P-S-N$ curve for small samples is proposed, which accords with the two parameter Weibull distribution, by using the principle of life quantile consistency and an iterative method.

(1) In this paper, the method is applied to design a variety of stress levels of small sample test program, and the test data and literature data are used to verify, compared with the traditional method, the proposed method has a higher accuracy in the number of small samples.

(2) The core of this method is to use the fatigue characteristic of Weibull distribution, calculate the eigenvalue by using the mathematical function, and then get the relatively accurate parameter value by using the iterative principle. It has high recognition, simple and intuitive method, and is easy to be used in practical engineering.

Acknowledgements
This research was supported by the National Natural Science Foundation of China under contract number 51665029, and the Industrial Support Plan for Colleges and Universities in Gansu Province of China grant number 2020C-12.
References
[1] Li C H, Wu S C and Liu Y X 2019 *Journal of Mechanical Engineering* 55 42-53
[2] Li Y H, Zhang Y and Shi S S 2020 *Journal of Mechanical Engineering* 18 165-170
[3] Zhai J M and Li X Y 2012 *International Journal of Fatigue* 44 107-15
[4] Liu W F, Liu Z M and Xue H 2018 *Journal of the China Railway Society* 40 52-58
[5] Shimizu S, Toshia K and Tsuchiya K 2010 *International Journal of Fatigue* 32 565-75
[6] Fouchereau R, Celeux G and Pamphile P 2014 *International Journal of Fatigue* 68 217-23
[7] Lv Z and Yao W X 2008 *Mechanics in Engineering* 5 9-15
[8] Guida M and Penta F 2010 *Structural Safety* 32 64-76
[9] Castillo E, Fernandez-Canteli A and Siegele D 2014 *International Journal of Fatigue* 187 159-72
[10] Amraoui AE, Gharad AE and Bensalah MO 2014 *Applied Mathematical Sciences* 8 1323-31
[11] Gao J X, An Z W and Liu B 2017 *Proceedings of the Institution of Mechanical Engineers* 231 130-7
[12] Xie L Y, Liu J Z and Wu N 2014 *International Journal of Fatigue* 63 62-7
[13] Bai X, Xie L Y and Qian W X 2016 *Journal of Mechanical Engineering* 52 206-12
[14] Bai En J, Huang S T and Lie L Y 2019 *Journal of Xi'an Jiaotong University* 53 96-101
[15] Xie L Y 2009 *Basic theory and method of mechanical reliability* (Beijing: Science Press) p 28
[16] Xie L Y 2014 *Journal of Mechanical Engineering* 50 27-35
[17] Sun Q, Dui H N and Fan X L 2014 *International Journal of Fatigue* 69 16-21