Journal of Applied Sciences

ISSN 1812-5654

science alert

ANSInet
an open access publisher
http://ansinet.com

Journal of Applied Sciences 13 (9): 1607-1611, 2013
ISSN 1812-5654 / DOI: 10.3923/jas.2013.1607.1611

Integration of Historical Data and a Small Sample of Data Reliability Assessment Methods

Xie Anqi, Wang Xiaohong and Yang Ming
School of Reliability and System Engineering, Beihang University,
Beijing 100191, China

Abstract: To extract valuable information from a small sample of fatigue failure data completely and accurately, the author proposes a theory of historical data fusion for a certain kind of products under different stress levels. It is presumed that their fatigue life distribution obeys three-parameter Weibull distribution. The shape parameter is estimated with historical shape parameter under other stress levels under a proper assumption. The relationship between the scale parameter and stress is the main study content. The scale parameter is estimated with data under different stress levels using the Inverse Power Law Model and a critical approximation. At last using the small sample the positional parameter can be computed. The existing historical data fusion methods do not take the differences among three parameters into account. For the reason that different parameters have different relationship with historical data, so it is more proper to use corresponding method to each parameter.

Key words: Three-parameter weibull distribution, historical data fusion, fatigue life

INTRODUCTION

With advances in technology, the quality of products is being continuously improved and therefore the service life is being extended, which create a new challenge to Reliability and life testing. The traditional reliability test and life test requires a relative high number of samples or test time. Applied into test of high reliability products, the traditional methods cause high cost and low benefit-cost ratio. Thus, the study on alternative methods is popular. This field of study has two directions. First direction is to improve the test stress, such as accelerated environmental testing. Second direction is the study on data processing, such as the Bayesian method and the bootstrap method (Hongshuang, 2006). At the same time, it is a new studying direction to use the existing historical data and information of the same or similar products to improve the accuracy of data fitting of small samples.

Currently there do not exist a lot of research on the data fusion of historical fatigue life, however, the booting of the quality effects the cost of money and time in life testing, for which it is urgent to study on this field. And fatigue life is always fit with three-parameter Weibull distribution and thus the key point is put on the estimation of the three parameters. The existing historical data fusion methods cope with the different parameters with the same or similar means, however in fact, the three parameters have different relationship with experimental stress, for which those methods have some limitation and shortages. In this paper, based on the correlation coefficient optimization method, the weighted linear regression method and the inverse power law model this method solves the problem of the differences among the three parameters in estimating.

MATH TOOLS

This method of historical data fusion is based on the premise that the mechanical fatigue life distribution obeys the Weibull distribution. And the second, to find the relationship between stress and life we must use the inverse power law model. The third, the weighted linear regression method is used in estimating the scale parameter. At last the correlation coefficient optimization method is used in estimating the positional parameter.

Three-parameter Weibull Distribution:Weibull distribution functions are given by:

$$F(t) = 1 - e^{-(t-g)^m/t_0} \quad (1)$$

$$F(t) = \frac{m}{t_0}(t - ?)^{m-1} e^{-(t-?)^m/t_0} \quad (2)$$

where, m is the shape parameter, t_0 is the scale parameter, γ is the positional parameter.

Assume $t_0 = \eta^m$, function (1), (2) switch into:

$$F(t) = 1 - e^{-((t-?)/?)^m} \quad (3)$$

Corresponding Author: Xie Anqi, School of Reliability and System Engineering, Beihang University, Beijing 100191, China

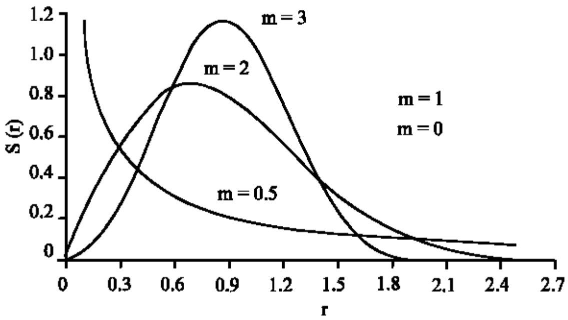

Fig. 1: Weibull probability density function with different parameters

$$F(t) = \frac{m}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{m-1} e^{-\left(\frac{t-\gamma}{\eta}\right)^m} \quad (4)$$

where, η is called true scale parameter.

The Weibull Life Expectancy and Variance are given by:

$$E(T) = \gamma + \eta\Gamma(1+1/m) \quad (5)$$

$$Var(T) = \eta^2[\Gamma\left(1+\frac{2}{m}\right) - (\Gamma(1+1/m))^2] \quad (6)$$

where, Γ (1+1/m) is a funtion of Γ.

Inverse power law model: Life characteristics decrease by negative power function while the mechanical or electrical stresses rise. This model is mainly used in machinery and electrical products and has been confirmed by many experimental data. As this paper only discusses mechanical fatigue life obeying Weibull distribution , it is reasonable to adopt this model in the following estimation.

Inverse power law model is given by:

$$\frac{dM}{dt} = K = \Lambda S^{\alpha} \quad (7)$$

where, M is a life characteristics, such as the median life expectancy or the mean life expectancy, Λ is a constant, α is a constant decided by activation energy, S is the stress, As the applied stress is constant, points on both sides of the Eq,:

$$M\text{-}M_0 = \Lambda F^{\alpha} t \quad (8)$$

Then:

$$t = \frac{M - M_0}{\Lambda} \bullet F^{-\alpha} \quad (9)$$

The relationship between life characteristics and stress levels is nonlinear. After some proper transformation , including logarithmic transformation, the relationship above becomes linear. This linearization process can make the function more convenient to use.

Logarithmic transformation:

$$\ln t = \ln(M - M_0) - \ln? + (-a)\bullet \ln S = \ln(M - M_0) - \ln? - a\ln S \quad (10)$$

When the value of M reaches a specified value M_{SF}, it can be judged that the product fails. At this time the value of t is the life of the product:

$$L = \frac{M_{SF} - M_0}{\Lambda S^{\alpha}} \quad (11)$$

$$\ln L = \ln(M_{SF} - M_0) - \ln\Lambda - \alpha \ln S \quad (12)$$

Assume that:

$$A = In\ (M_{SF} = M_0)\text{-}In\Lambda \quad (13)$$

$$B = -\alpha \quad (14)$$

Then:

$$In\ L = A+B\ In\ S \quad (15)$$

The function above show the linear relationship between the life and the stress,which is used in the next chapter.

Correlation coefficient optimization method: Correlation coefficient optimization method is applicable in estimating the positional parameter. The following is the specific process.

Switch the original function into:

$$\ln(-\ln(1 - F(t))) = m\bullet\ln(t - \gamma) - \ln(\eta^m) \quad (16)$$

Assume that:

$$Y = \ln\left[-\ln(1 - F(x))\right] \quad (17)$$

$$X = \ln (x-\gamma) \quad (18)$$

$$B = \ln \eta^m \quad (19)$$

Then:

$$Y = mX-B \quad (20)$$

Then, converse the sample data (x_i, $F(x_i)$) into (X_i, Y_i) and compute the correlation coefficient R (X, Y):

$$R(X,Y) = \frac{(\sum_{i=1}^{n} X_i Y_i - n\bar{X}\bar{Y})}{\sqrt{\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right)(\sum_{i=1}^{n} Y_i^2 - n\bar{Y}^2)}} \quad (21)$$

Obviously, R (X, Y) is a function of γ. When R (X, Y) obtains the maximum, the value of γ is the estimates of it. It can be proved by derivation that when the following formula establishes correlation coefficient R (X, Y) obtains the maximum:

$$(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2)\sum_{i=1}^{n}\frac{\bar{Y}-Y_i}{x_i - ?} - (\sum_{i=1}^{n} X_i Y_i - n\bar{X}\cdot\bar{Y})\sum_{i=1}^{n}\frac{\bar{Y}-Y_i}{x_i - ?} = 0 \quad (22)$$

The solution of the function above is the positional parameter's estimate value.

Weighted linear regression method: The assumption of least square method is:

$$Y_i = \beta_0+\beta_1 X_i+\varepsilon \quad (23)$$

Where:

$$\varepsilon \sim (0, \sigma_i^2)$$

Which means the residuals ε doesn's change with the dependent variable. But in fact the residuals ε does change with the dependent variable, which is called heteroscedasticity. This is given by:

$$Y_i = \beta_0+\beta_1 X_i+\varepsilon_i \quad (24)$$

Where:

$$\varepsilon \sim (0, \sigma_i^2)$$

Assum that:

$$\sigma_i^2 = k_i \sigma^2 \quad (25)$$

Regression model changes into:

$$Y_i/k_i = \beta_0/k_i+\beta_1 X_i/k_i+\varepsilon \quad (26)$$

After substitution:

$$Y_i' = \beta_0'+\beta_1' X_i'+e' \quad (27)$$

Where:

$$\varepsilon \sim (0, \sigma_i^2)$$

Obviously the substitution eliminate the heteroscedasticity.

Using the following formulae:

$$\begin{cases} \hat{\beta}_1 = \dfrac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \\ \hat{\beta}_0 = \bar{y}-\hat{\beta}_1\bar{x} \end{cases} \quad (28)$$

Compute the value of:

$$\hat{\beta}_1 - \hat{\beta}_0$$

DETERMINE THE LIFETIME DISTRIBUTION FUSING THE HISTORICAL DATA

Outline of the method: Historical data can increase the amount of information and improve the accuracy in the estimate with small samples. But in the study of historical data fusion, previous methods did not take the differences among the three parameters into account. This study discusses that the three parameters have different relationship with stress ,and use corresponding means to estimate the three parameter.

Determination of parameters

Determination of shape parameters: In the premise of constant failure mechanism, lifetime distributions of products in different stresses share the same shape parameter (Sun, 2008). Thus, the shape parameter can be estimated with historical shape parameters using the weighted average method:

$$\hat{m} = p_1\hat{m}_1 + p_2\hat{m}_2 + \ldots + p_k\hat{m}_k \quad (29)$$

Where:

$$p_i = \frac{d_i}{\sum_{i=1}^{k} d_i} \quad (30)$$

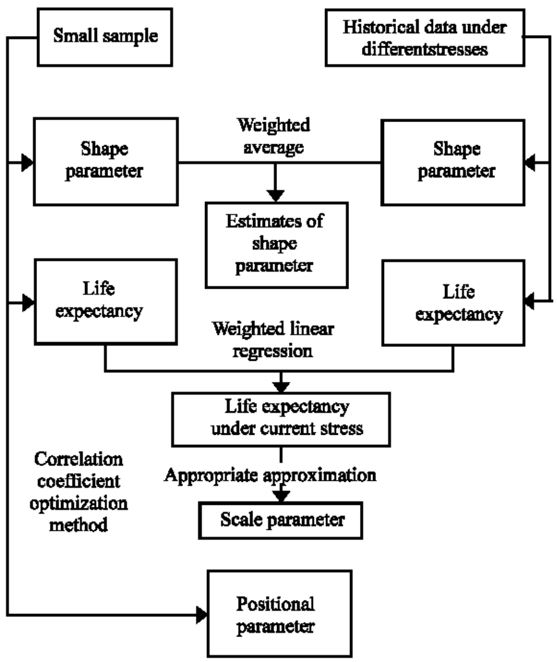

Fig. 2: Flow chart of the method

$$d_i = \frac{\sigma_i}{\sqrt{n_i}} \frac{|S_i - S_o|}{S_o} \quad (31)$$

where, $\widehat{m_i}$ is the historical shape parameter estimates, p_i is the weights of each historical shape parameter estimates, n_i is number of samples in group i, σ_i standard deviation of logarithmic fatigue life, S_0 is the current stress.

Determination of scale parameter: The main innovation of this paper is the means to estimate scale parameter. On one hand,scale parameter have close relationship with stress, however, it is necesary to use life expectancy as a bridge to find out the relationship between scale parameter and stress level. And in this process an approximation is also needed. On the other hand, with the difference between the current stress and historical stresses increasing, the variance is also increasing and the value of the data under the specific stress for estimation of the current scale parameter is shrinking. So, weighted linear regression should be taken into estimation in this situation.

Weibull life expectation is $\gamma+\eta\Gamma$ (1+1/m). As in the most of the situations there are 8-9 orders of magnitude difference between γ and $\eta\Gamma$ (1+1/m) (Hongshuang, 2006). Thus, it is reasonable to ignore γ in computing the life expectation and the life expectation is simplified as $\eta\Gamma$ (1+1/m).

As mentioned in the last chapter, the logarithm of the low cycle fatigue life and the logarithm of the lifetime have a linear relationship in a certain interval (Li and Lu, 2007). With the historical data in different stress levels the linear relationship can be determined by linear regression. But with the difference between the current stress and historical stress increasing, the variance is increasing. For this reason weighted linear regression is chosen in this situation.

As mentioned in the last chapter,the relationship between logarithmic life InL and logarithmic stress InS is given by:

$$\text{InL} = \text{A}+\text{B InS}$$

Using weighted linear regression parameters A and B can be determined:

$$\hat{\eta} = \frac{\hat{L}}{\Gamma(1+1/\hat{m})} \quad (32)$$

Using the formula above the estimates of scale parameter can be determined.

Though the estimates from small samples have direct meaning, its accuracy is so poor that corretion by historical data is necesary

$$\hat{\eta} = \frac{\widehat{\eta_1}R_1 + \widehat{\eta_2}R_2}{R_1 + R_2} \quad (33)$$

where, $\widehat{\eta_1}$ is the estimates from historical data, $\widehat{\eta_2}$ is the estimates from the small sample, R_1 and R_2 is the correlation coefficient of corresponding lifetime expectation and average lifetime under other stress levels. The previous historical data fusion method has noticed the effect of different stress levels on the three parameters, but they did not find the proper way to connect them. In terms of the mechanical sample, the method proposed in this paper takes adventage of Inverse Power Law Model and introducts a approximation based on engineering practice, which improve the accuracy of historical data fusion.

Determination of positional parameters: The poitional parameter can be determine by correlation coefficient optimization method with the data of the small sample.

EXAMPLE ANALYSIS

There are 600 samples, respectively under 666, 583 and 478 MPa by experiments. Ten fatigue life samples under 583MPa are randomly chosen as a small sample of experimental data. Tanaka *et al*. (1984) The samples of other two stress levels, 666 and 478 MPa are used as historical data. The small sample is shown in the Table 1.

Table 1: Small sample fatigue life data

275029	328296	372329	398485	4228890
44624	47096	496056	538543	588051

Table 2: Historical data

Stess (MPa)	Logarithmic	Shape	Scale
666	4.54	2.70	2.274×10^{10}
478	5.30	3.10	2.540×10^{17}

Table 3: Historical data fusion results

Stess (MPa)	Logarithmic	Shape	Scale
583	4.40	2.90	3.546×10^{14}

Table 4 large sample result

Stress (MPa)	Logarithmic	Shape	Scale
583	4.74	2.75	4.863×10^{14}

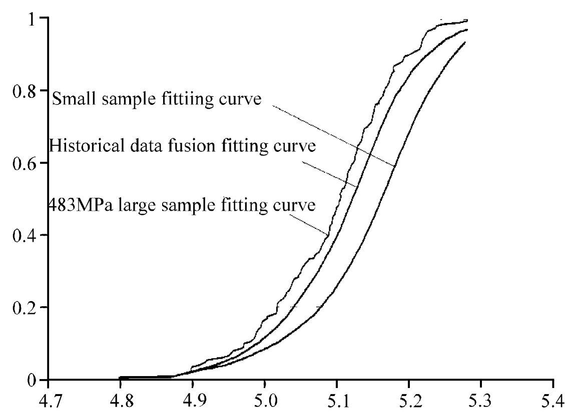

Fig. 3: Comparation of fitting curves

The three parameters of the historical data is shown in the Table 2.

Based on the data above, the historical data fusion results is shown in the Table 3.

With the 200 samples under 583 MPa the actual three parameters can be determined and shown in the Table 4.

Based on the small sample result and the historical data fusion result two curves are fit. It is obvious that the accuracy is improved.

CONCLUSION

In this study the three parameters is estimated with three different means according to their different relationship with stress level. The main content in this study is to improve estimation accuracy of scale parameter with the relationship among scale parameter, life expectation and stress, which is proved to be valid. Shape parameter estimation is based on the assumption of constant failure mechanism, so it has some of limitation. And because there is not enough study on the relationship between the positional parameter and the historical data, the positional parameter estimation accuracy is not improved by historical data. But judged from the example, this method has a relative high level of accuracy.

REFERRENCES

Hongshuang, L., 2006. Small-Sample Analysing and Research Method. Shaanxi Publication, China.

Li, H.S. and Z.Z. Lu, 2007. New method for estimating fatigue life three-parameter P-S-N curves. J. Mechanical Strength, 29: 300-304.

Sun, Z.L., 2008. A condition for constant failure mechanism. Electron. Product Reliabil. Environ. Testing, 4: 6-8.

Tanaka, S., M. Ichikawa and S. Akita, 1984. A probabilistic investigation of fatigue life and cumulative cycle ratio. Eng. Fract. Mechan., 20: 501-513.