

Sample Size Reduction Based on Historical Design Information and Bayesian Statistics

Zhigang Wei

Tenneco Automotive Co., Ltd.

ABSTRACT

Numerous test data have been generated in many testing institutions over the years and the historical information from previous similar designs and operating conditions can shed light on the current and future designs since they would share some common features when the changes are not drastic. To effectively utilize the historical information for current and future designs, two steps are necessary: (1) finding an approach to consistently correlate the test data; (2) utilizing Bayesian statistics, which can provide a rigorous mathematical tool for extracting useful information from the historical data.

In this paper, a procedure for test sample size reduction is proposed based on historical fatigue S-N test data and Bayesian statistics. First, the statistical information is extracted from a large amount of fatigue test data collected over the years. Subsequently, the basic theory of the Bayesian statistics and a very effective numerical algorithm for resampling of statistical distributions are introduced. The combination of the historical data and the Bayesian statistics makes the sample size reduction and the assessment accuracy improvement possible. Finally, examples are presented and the benefit of utilizing the new analysis procedure is highlighted and discussed.

CITATION: Wei, Z., "Sample Size Reduction Based on Historical Design Information and Bayesian Statistics," *SAE Int. J. Mater. Manf.* 7(1):2014, doi:10.4271/2013-01-2440.

INTRODUCTION

Durability and reliability performance is one of the most important concerns of all engineering systems [1, 2], and welded joints are usually the weakest locations of a system under fatigue loading [3]. Durability and reliability analysis of products are essentially based on test data and the accurate interpretation of the test data. Therefore, testing is an indispensable part in a product validation process, and many repeats of tests are often required to capture the uncertainty, scatter band, and failure probability of the test data. However, the time and efforts involved in components and systems testing, especially in early development stage, can be very expensive and usually only a small number of products are available for testing. Therefore, it would be very useful to find and apply methods which can decrease the required test sample size with same or increased accuracy. Two possible methods have been considered for fatigue life assessment: (a) master fatigue S-N curve approach, and (b) historical information and Bayesian statistics approach.

First, great efforts have been made to develop the ability to accurately and quickly predict the durability and reliability

of the welded structures in the early development stage and many advanced techniques such as the structural stress/traction stress method have been developed [4]. With the method [4], fatigue data from various resources, different loading conditions and geometries have been found to collapse onto a master S-N curve [5]. If proven accurate, this method can significantly simplify design processes, and costly and time-consuming testing can be significantly reduced or even avoided. Several other versions of similar structural stress concepts and methodologies are also available and have been implemented into commercial software [6]. The linear traction stress method and the master S-N curve methods [4, 5] have been systematically studied and applied at Tenneco for many years for assessing fatigue life of welded structures and it was found that these approaches can give a generally good trend (the mean curve) but with a very big scatter band, which comes from many aspects and the direct applications of these methods in engineering designs is limited [7, 8]. Therefore, we have not reached the ideal situation, where testing can be completely replaced with pure analytical and numerical assessment, and

testing is still the essential and the most important part of a product validation process.

Second, most of the life assessment methods currently used in the industry utilize only the data observed during the current test and ignore any prior knowledge about the products or its predecessors that could be available and helpful. Historical failure data, which can be used as prior information for Bayesian statistics, can be used to obtain the statistical information such as mean curves and design curves with increased accuracy and confidence as well as reduced test sample size [9, 10, 11].

In this paper, a large amount of historical fatigue S-N data are collected and analyzed with Bayesian statistics. Examples show that utilizing the historical data in a rational way can reduce the testing burdens and increase assessment accuracy.

HISTORICAL DATA

A Master Fatigue S-N Curve for Exhaust Systems

Like other subsystems in an automobile, the exhaust system experiences cyclic loadings due to road vibrations, leading to fatigue failure over time. Welded joints are usually the weakest point of an exhaust system. Years of experience has shown that the prevailing failure modes in exhaust systems are cracked welds at joints between pipes and muffler/resonator/converter (360° welds), cracked hanger-to-pipe welds (line welds), and broken hanger rods [12]. A reliable and validated analytical method to predict fatigue life is urgently needed to assess how and where a welded part will fail in terms of fatigue life, while the iterative testing is expensive and time consuming.

Node force based linear traction stress procedure and the master S-N curve approach (Battelle fatigue master S-N curve) developed by Dong et al [4, 5] have been widely used in fatigue life assessment of welded structures. The linear traction stress approach is based on a linearization of the stress distribution across the plate thickness and along the anticipated crack path [4, 5]. The distribution of a stress component through thickness can be decomposed into membrane (average value) stress σ_m , bending stress (a linearly varying term) σ_b , and nonlinear self-equilibrating remainder, and the sum of the first two terms $\sigma_s = \sigma_m + \sigma_b$, can be used, based on fracture mechanics, to derive the equivalent traction stress ΔS_s , which will serve as a fatigue damage parameter. It has been demonstrated that a consistent characterization of stress concentration effects with mesh-insensitivity can be achieved and many fatigue test data can collapse onto a single curve [4, 5].

With the linear traction stress approach, the fatigue data generated over several years at Tenneco Inc., have been systematically collected, analyzed to assess fatigue life of welded structures [7, 8]. It was found that the linear traction approach can give a generally good trend and the mean curve of the master fatigue S-N curve matches that of Tenneco 360°

welds data well [7, 8]. However, the Tenneco data show a bigger scatter band, which is caused by many factors and the direct application of the methods in engineering design is limited [8]. 360° welds include bushing, flange-to-pipe, and pipe-to-pipe etc. [12]. The previous data, termed of Phase-I data, have been reported [8]. And a large amount of new test data of similar types, termed as Phase-II data, has been recently collected, analyzed and added to the Tenneco database. All individual test data of both Phase-I and Phase-II data are plotted together and shown in Fig.1 of the Appendix.

At a stress level, six tests are repeated. It is clear that both Phase-I and Phase-II data demonstrate similar data patterns. The linear fit parameters obtained with several fit methods [13] are provided in Table 1 and Table 2 of the Appendix, respectively, for Phase-I and Phase-II data. It is found that the differences in fit parameters are not significant, therefore, these two data sets can then be combined to generate a new master S-N curve for exhaust weld structures, and the values of the fit parameters are provided in Table 3 of the Appendix. In curve fitting, the linearized relationship $Y = A + BX$ is used, where $Y = \log_{10}(\Delta S)$ and $X = \log_{10}(N)$. ΔS and N are equivalent structural stress (SS) range and the cycles to failure, respectively.

It is clear from Fig.1 that the data pattern of Tenneco 360° welds data does not belong to the standard vertical pattern nor the horizontal pattern [13]. In fact, the data cannot be easily trimmed into standard vertical and horizontal patterns because of limited data range [13]. Therefore, the perpendicular offset method should be used to obtain a fit curve [13]. The fit parameters for the whole data with the three methods are listed Table 3 and the fit mean curves obtained from both perpendicular (solid line) [13] and standard [14] horizontal (dash-dot line) methods are also plotted in Fig.2 of the Appendix. The new Battelle master S-N mean curve [8] obtained from the perpendicular method for the Battelle database is also plotted in Fig.2 for comparison purpose. It is clear that the difference between these two methods is significant for Tenneco 360° welds data, and the fit curve obtained with the horizontal method is not accurate at all and misleading. The rotated fit curve with horizontal method (the standard method) can be easily described by the equilibrium argument [13]. It is interesting to note that both the new Battelle master S-N curve and Tenneco master S-N curve for 360° welds give a very similar mean curve.

The fit parameters based on the perpendicular method for the new Battelle master S-N curve, the current Tenneco S-N curve (Phase-I) and the new Tenneco master S-N curve (Phase-I & Phase-II) are listed in Table 4 of the Appendix. σ_N is standard deviation in terms of fatigue life (after log transformation). The detailed data analysis procedure can be found elsewhere [8, 13]. The mean curves of both the new Battelle S-N curve and the new Tenneco master S-N curve are plotted in Fig.3 of the Appendix. C and h in Table 4 can be used to express the life-stress range relationship with $N =$

$C(\Delta S)^{-h}$. h represents the inverse slope of the mean S-N curve of the data in log-log scale.

It is found from Fig.3 that the life spans four orders of magnitude within the range of $\pm 3\sigma_N$. It should be noted that the master S-N curve was obtained by analyzing all of the data obtained under various conditions. Variation and uncertainty of materials, loading control and the analysis methods all make contributions to the big scatter band. Therefore, the scatter band of the Tenneco fatigue master S-N curve is too wide for direct applications. In fact, the test data in Fig.1 and Fig.2 are obtained using many test sets, which consist of six individual test data points at each stress level, and the life span for each test set is usually less than one order of magnitude and the scatter band for a specific test is usually much narrower. An example is shown in Fig.4 of the Appendix. Therefore, alternative methods need to be explored to find useful information for life assessment. The distribution of mean and standard deviation of test data will be presented in the following subsection for this purpose.

Distribution of the Mean and the Standard Deviation

The mean distribution of all test sets for data shown in Fig.1 and Fig.2 are plotted in Fig.5 in a log-log plot. To protect the confidential information the detailed values in Fig. 5 and in some figures of the following sections are not shown. The mean curves obtained using the three fit methods [13] are plotted in Fig.5(a) and the design curves based on the perpendicular method is shown in Fig.5(b). Implicitly, the lognormal distribution is assumed in the data fit process, and it provides a reasonable good data fit. Again, the Phase-I and Phase-II data give a similar data pattern. The two-parameter lognormal (with mean μ and standard deviation σ) probability density function (pdf) and the corresponding cumulative distribution function (cdf) are expressed in Eq.(1a) and Eq. (1b), respectively.

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

$$F(x) = \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right) \right]; \sigma > 0, -\infty < x < \infty \quad (1-b)$$

where μ is the mean and σ^2 is the variance; $\operatorname{erf}()$ is the error function

Fig.6 shows the distribution of the standard deviation of the test data. It is found that the Phase-I and Phase-II data also give a similar data pattern. Based on the characteristics of the data pattern, three-parameter Weibull distribution is used to fit the data and the fit curve obtained from MinTab is plotted in Fig.7 (a). The two-parameter Weibull fit is also provided in Fig.7(b) for comparison. From the data plot shown in Fig.7(b), it can be clearly seen that the data do not

fall on a straight line, and therefore, the two-parameter Weibull distribution does not give a good fit at the lower tail, which shows a concave downturn trend, indicating a threshold parameter needs to be introduced. Therefore, the three-parameter Weibull distribution can capture the major trend and provide a better data correlation.

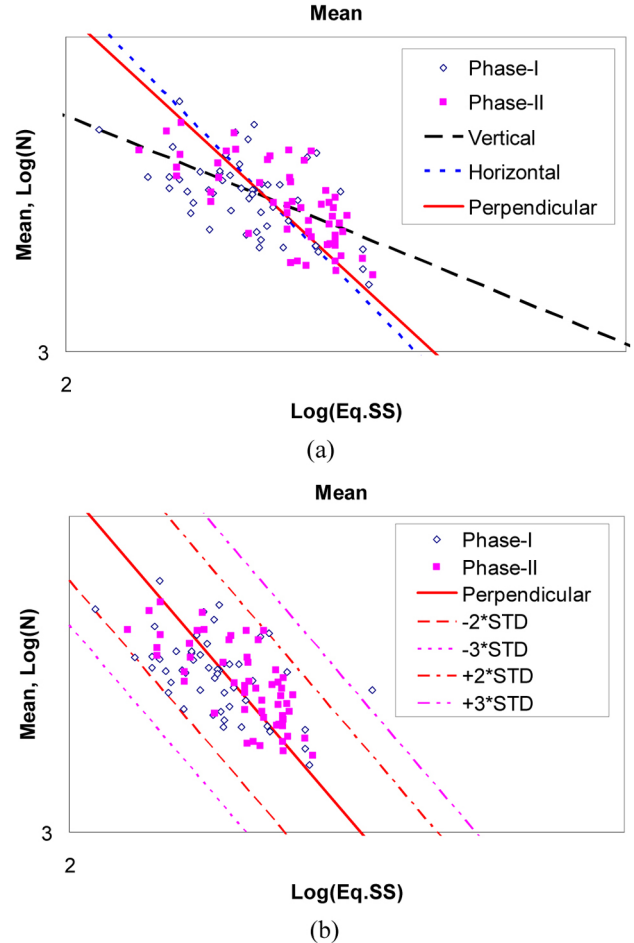


Fig.5. The construction of (a) mean curves and, (b) the design curves

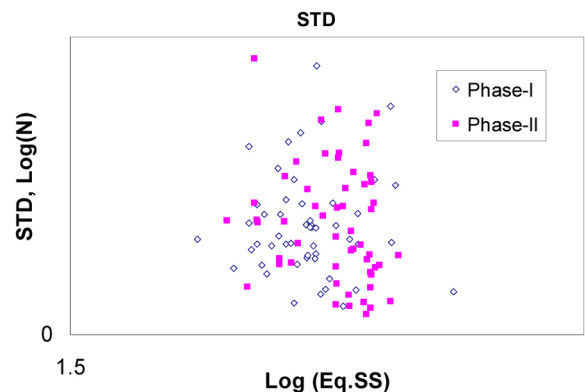


Fig.6. The distribution of the standard deviation (STD)

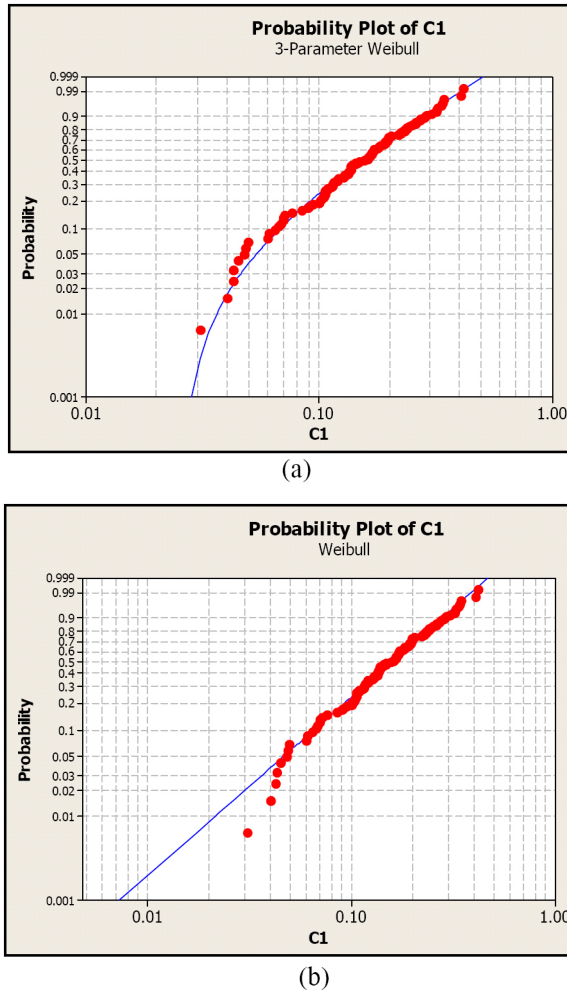


Fig. 7. Data fit of the standard deviation using (a) three-parameter and, (b) two-parameter Weibull distributions

The three-parameter Weibull pdf and cdf are as follows:

$$f(x) = \frac{\beta}{\eta} \left(\frac{x-\gamma}{\eta} \right)^{\beta-1} \exp \left[- \left(\frac{x-\gamma}{\eta} \right)^{\beta} \right]; \quad x \geq 0, \eta > 0, \beta > 0, x > \gamma \quad (2-a)$$

$$F(x) = 1 - \exp \left[- \left(\frac{x-\gamma}{\eta} \right)^{\beta} \right]; \quad x \geq 0, \eta > 0, \beta > 0, x > \gamma \quad (2-b)$$

where η , β and γ are the scale, shape, and location or shift or threshold parameters, respectively. When $\gamma = 0$, the three-parameter Weibull functions are the two-parameter Weibull distributions. The threshold parameter γ give the lower bound of the Weibull cdf. As the name implies, the threshold parameter locates the cdf along the abscissa. Changing the value of γ has the effect of “sliding” the cdf and its associated function to the right because value of γ must be positive. The existence of a positive threshold parameter γ reflects the intrinsic scatter nature of fatigue data and the scatter characteristics cannot be eliminated and should be considered in curve fit of the standard deviation.

DESIGN CURVE CONSTRUCTION

Based on the current test data, the Frequentist method can be used to calculate the statistical information such as mean, standard deviation and design limit/curve. For a given limited test data, in addition to the mean curve and standard deviation, the design limit/curve is also an important parameter in durability and reliability designs [15]. Generally, a design limit/curve is constructed to ensure that a majority of the fatigue or other component relevant life data falls above the minimum or lower bound value with a certain confidence interval [1, 16]. Depending on the test specimen size and the required reliability, a confidence level is specified in order to quantify these uncertainties in design. The construction of a design limit/curve is generally dependent upon material cost, safety policy, and industry standards. Several methods have been proposed and some of these methods have been coded by engineering standards bodies, including ASME code [16]. For a homoscedastic distribution assumption, the design limit Y_0 , is generally expressed in Eq.(3) with Y as a function of the independent variable X .

$$Y_0 = \bar{Y} - Ks \quad (3)$$

where $\bar{Y}$ is the mean value or curve and can be easily calculated from a given set of test data; The sample standard deviation S for fatigue data is calculated with Eq.(4).

$$s = \sqrt{\frac{1}{f} \sum_{i=1}^n [\log(N_i) - \overline{\log(N_i)}]^2} \quad (4)$$

n is the total sample size and $f = n - 2$ is chosen for problems with two undetermined parameters (two or more than two stress levels) and $f = n - 1$ should be used for problems with only one undetermined parameters (one stress level) [15]; K is a factor generally related to confidence interval, sample size, and failure probability, and is given in Eq.(5) for a specific proportion $(100(1 - p)\%)$. i.e., $R(1 - p)$, of the population be above or below a prescribed confidence level $(100(1 - \gamma)\%)$, i.e., $C(1 - \gamma)$, or overall, R90C90 [15].

$$K = \frac{z_p(1 - \phi) + \left\{ z_p^2(1 - \phi)^2 - \left[(1 - \phi)^2 - z_r^2/(2f) \right] \left(z_p^2 - az_r^2 \right) \right\}^{1/2}}{(1 - \phi)^2 - z_r^2/(2f)} \quad (5)$$

where $\phi = 1/(4f)$. More details can be found in [15].

BAYESIAN STATISTICS AND SAMPLING-IMPORTANCE-RESAMPLING PROCEDURE

In contrast to the Frequentist approach, Bayesian approach considers previous knowledge obtained from