

Fatigue Life Distribution Estimation



D. Gary Harlow

Abstract Modeling fatigue life is complex whether it is applied to structures or experimental programs. Through the years several empirical approaches have been utilized. Each approach has positive aspects; however, none have been acceptable for every circumstance. On many occasions the primary shortcoming for an empirical method is the lack of a sufficiently robust database for statistical modeling. The modeling is exacerbated for loading near typical operating conditions because the scatter in the fatigue lives is quite large. The scatter may be attributed to microstructure, manufacturing, or experimental inconsistencies, or a combination thereof. Empirical modeling is more challenging for extreme life estimation because those events are rare. The primary purpose herein is to propose an empirically based methodology for estimating the cumulative distribution functions for fatigue life, given the applied load. The methodology incorporates available fatigue life data for various stresses or strains using a statistical transformation to merge all the life data so that distribution estimation is more accurate than traditional approaches. Subsequently, the distribution for the transformed and merged data is converted using change-of-variables to estimate the distribution for each applied load. To assess the validity of the proposed methodology percentile bounds are estimated for the life data. The development of the methodology and its subsequent validation is illustrated using three sets of fatigue life data which are readily available in the open literature.

Keywords Cycle dependent parameters · Data fusion · Fatigue life transformation · Weibull distribution

Nomenclature

α, β, γ Weibull cdf parameters
 α_s significance

D. Gary Harlow (✉)
Mechanical Engineering and Mechanics, Lehigh University, 19 Memorial Drive West, Bethlehem,
PA 18015–3085, USA
e-mail: dgh0@lehigh.edu

AD	Anderson–Darling goodness of fit test
cdf	cumulative distribution function
cv	coefficient of variation
$\Delta\varepsilon$	strain range
$\Delta\sigma$	stress range
FLT	Fatigue Life Transformation
$F(\bullet)$	cdf
KS	Kolmogorov–Smirnov goodness of fit test
MLE	maximum likelihood estimation
N_A	arbitrary normalization constant
N_f	cycles to failure
n	sample size
m	number of different values of applied stress
p	percentile
s	sample standard deviation
s_A	arbitrary normalization constant
S–N	stress–number of cycles
t	time, cycles
$W(\alpha, \beta, \gamma)$	three parameter Weibull cdf
$y_{1/2}$	median
\bar{y}	sample average

1 Introduction

A difficulty with fatigue life data is the characterization of its variability, which can be several orders of magnitude [1], especially for loading near operating conditions. The variability is attributable to experimental error, as well as material microstructure or processing. Thus, estimation and prediction of fatigue life is challenging. A key concern is the characterization of the cumulative distribution function (cdf) for fatigue life, given an applied load, which may be either the stress range $\Delta\sigma$ or the strain range $\Delta\varepsilon$. The lower tail portion of the cdf which depicts high reliability is especially critical; however, that is where scatter is more pronounced, and sample sizes are smaller. Empirically modeling fatigue life has been considered numerous times. A simple internet search for *statistical fatigue life modeling* yields in excess of 10 million citations. A relatively recent work is [2], in which the authors incorporate statistical analysis with traditional stress-life, strain-life, or crack propagation models. While empiricism is used, the thrust is to incorporate as much physically motivated modeling as possible. More frequently, investigators attempt to fit a stress-life (S-N) curve through the data, especially the medians for given $\Delta\sigma$ or $\Delta\varepsilon$. A nice review of such practices is [3]. Another example of statistical modeling of fatigue data is contained in [4] in which over two chapters are devoted statistical methodologies. Other examples of statistical stress-life analysis may be found in [5, 6]. While there

are many more papers, books, and conferences publications on statistical fatigue modeling, these are representative. In spite of all these references, statistical fatigue analysis is still an open area of investigation.

Herein, an empirically based approach for accurately estimating the fatigue life cdf, given $\Delta\sigma$ or $\Delta\varepsilon$ is proposed. The methodology merges fatigue life data using a statistical transformation for the estimation. The statistical technique increases the sample size by merging fatigue data for more precise assessment. This is necessitated because there is often large variability in S-N data, and the sample sizes are small. Validation of the modeling is essential, especially for prediction of life outside of the range of experimental observations. The validity of the methodology is evaluated by considering percentile bounds estimated for the S-N data. The development of the methodology and its subsequent validation is illustrated using three different fatigue life datasets.

A fundamental issue in fatigue life estimation is the choice of an underlying cdf. The cdf used in the ensuing analyses is a three-parameter Weibull cdf. A fairly recent example of a traditional statistical S-N methodology using the three-parameter Weibull cdf is [7] where fatigue of structural and rolling contact problems are considered. There seems to be a need for more experimental data to enhance modeling in almost all fatigue analyses. This is addressed in [8] by using normalization for the fatigue life data so that all the data are merged. The normalized data are then modeled with a three-parameter Weibull cdf. Even though the intention in [8] is similar to the emphasis herein, the methodology is somewhat different.

2 Fatigue Life Data

Fatigue life data are most often presented on an S-N plot which shows the fatigue data for a given load. The load is typically stress or strain. Thus S-N can represent stress-life or strain-life. An additional way in which the fatigue data are presented is on a probability plot. Both of these representations will be used subsequently. Three different sets of fatigue life data are considered for the proposed method.

The first set considered is one of the special cases taken from [9]. Fatigue testing was conducted on 2024-T4 aluminum alloy specimens. The fatigue tests were performed on rectangular specimens with dimensions of 110 mm long, 52 mm wide, and 1 mm thick, and with a center cut circular hole of radius 5 mm. Holes were cut using standard procedures with a lathe, and burrs were removed by polishing techniques. Testing was conducted in laboratory environment where temperatures of 295–297 K (approximately 22–24 °C) and relative humidities of 50–56% were observed. Constant amplitude tests were performed at a frequency of 30 Hz on a single machine with a single operator in order to minimize experimental error. Because of the extensiveness of this data, they have been used in a variety of analyses; e.g., [10–12]. These data are summarized in Table 1. A total of 222 specimens were tested using eight different values for $\Delta\sigma$. The specimens were tested to fracture. The sample coefficients of variation (c_v) are nearly the same, approximately 9%, for the larger

Table 1 Statistical summary of fatigue life data for 2024-T4 specimens [9]

Load, $\Delta\sigma$ (MPa)	Size (m)	Average (\bar{x})	Standard deviation (s)	cv (%)
255	21	18,200	1760	9.6
235	30	28,700	2500	8.7
206	30	59,400	4230	7.1
177	30	146,000	12,600	8.6
157	30	264,000	22,600	8.6
137	30	519,000	96,200	18.5
127	30	1,710,000	1,090,000	63.8
123	21	4,530,000	2,660,000	58.7

values of $\Delta\sigma$. When $\Delta\sigma$ is 137 MPa, the cv is about double, and for the two smaller values of $\Delta\sigma$ the scatter increases significantly. The fatigue life data are plotted on an S-N graph in Fig. 1 in the traditional linear versus logarithm S-N format. As $\Delta\sigma$ is reduced, the increase in the scatter in life is apparent. Modeling the increasing variability for decreasing $\Delta\sigma$ is the challenge for accurate fatigue life prediction.

The second set of fatigue data to be considered is data collected at room temperature for ASTM A969 hot dipped galvanized sheet steel with a gauge thickness of 1.78 mm [13]. ASTM A969 is a cold-rolled, low carbon, extra deep drawing steel. This steel is very ductile and soft, and it is age resistant. The automotive industry uses it in applications where severe forming is required, e.g., inner door components, dash panels, body side components, and floor pans with spare tire tubs. Fatigue tests for the ASTM A969 specimens were conducted using a triangular waveform at 25 Hz. The fatigue tests were terminated, i.e., designated as a failure, when the tensile load dropped by 50% of the maximum load. A total of 69 specimens were tested to failure. The data are summarized in Table 2. The cv given $\Delta\varepsilon$ is more scattered than those for the 2024-T4 data. The smallest value of $\Delta\varepsilon$ has the largest scatter, but the second

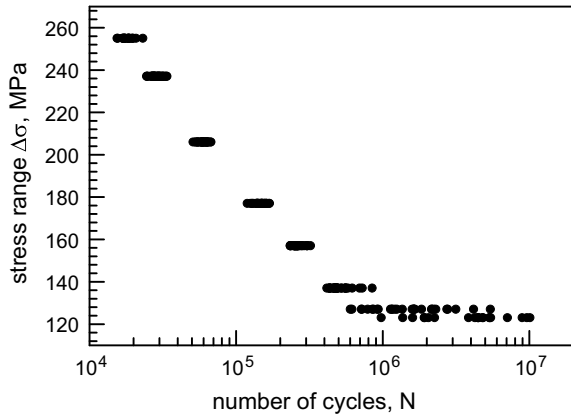
Fig. 1 Fatigue life data for 2024-T4 specimens [9]

Table 2 Statistical summary of fatigue life data for ASTM A969 specimens [13]

Load, $\Delta\epsilon$ (mm/mm)	Size (m)	Average (\bar{x})	Standard deviation (s)	cv (%)
0.0080	3	4,100	815	19.8
0.0060	12	14,000	3,800	27.2
0.0050	6	34,600	7,030	20.3
0.0040	12	56,500	8,860	15.7
0.0030	12	107,000	12,900	12.1
0.0024	6	199,000	29,800	15.0
0.0020	11	499,000	68,300	13.7
0.0018	7	1,030,000	391,000	37.8

largest $\Delta\epsilon$ is rather large as well. This behavior is seen graphically on the S-N plot in Fig. 2. Consequently, these data are not as statistically well behaved as the 2024-T4 data. These data have been used to investigate other types of fatigue modeling [14, 15].

The third set to assess is 9Cr-1Mo steel which were collated from a round-robin test program and were reported in [16]. This steel is creep strengthened, and it is frequently used in thermal power plants to improve the energy efficiency of the power plant by increasing operating temperatures and pressures. Specifically, 9Cr-1Mo is often used for steam generator components of both fossil fired and nuclear power plants. The material from which the data were generated was a single cast, rolled plate with a nominal tensile strength of 623 MPa [16]. A total of 130 specimens were tested to failure. The data are summarized in Table 3. The cv given $\Delta\epsilon$ is even more scattered than the above datasets. In fact, there does not seem to be any discernible pattern. The data are shown on Fig. 3. The reason for the unusual statistical behavior would require more in depth analysis than is provided in [16]. Usually round-robin testing has considerably more variability in results because testing conditions and

Fig. 2 Fatigue life data for ASTM A969 specimens [13]

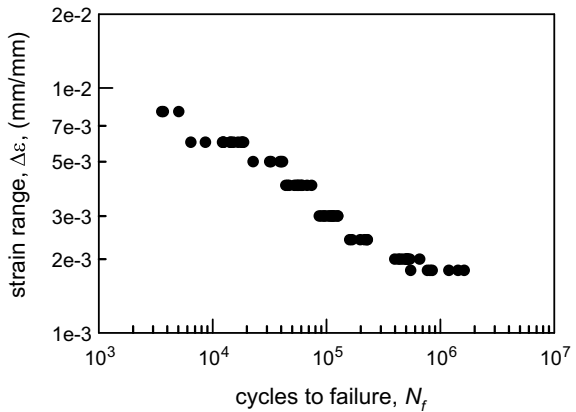
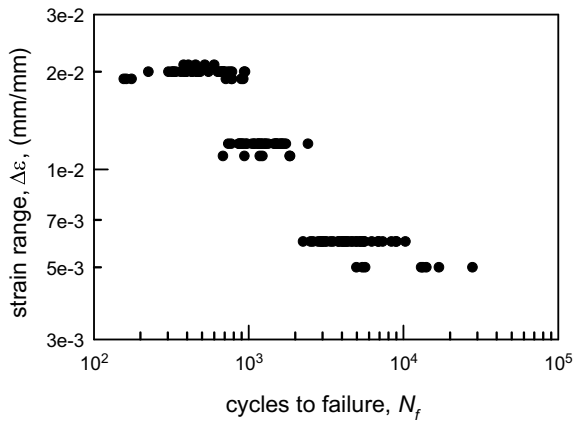


Table 3 Statistical summary of fatigue life data for 9Cr-1Mo specimens [16]

Load, $\Delta\epsilon$ (mm/mm)	Size (m)	Average (\bar{x})	Standard deviation (s)	cv (%)
0.021	5	472	88	18.8
0.020	32	542	189	34.9
0.019	8	572	345	60.2
0.012	34	1,260	357	28.4
0.011	6	1,290	475	36.9
0.006	37	4,820	2,000	41.6
0.005	8	12,700	7,700	60.8

Fig. 3 Fatigue life data for 9Cr-1Mo specimens [16]

methodologies are not consistent. Nevertheless, the data will serve as an excellent case for the proposed modeling approach.

3 Data Fusion for Fatigue Life Analysis

The following is a purely empirical method to improve fatigue life modeling. Because fatigue data are usually limited in number for relatively few different loading conditions, modeling is crude. A methodology that has been developed to account for uncertainty for static properties [17] is adapted for fatigue life data. The basis of the approach is a linear transformation of a collection of experimental observations $\{y_j : 1 \leq j \leq n\}$ into another set of values $\{z_j : 1 \leq j \leq n\}$ so that both sets have the same average and sample standard deviation. Let

$$z_j = ay_j + b. \quad (1)$$

The choices of a and b in Eq. (1) are easily determined by simple algebra to be the following:

$$a = \frac{s_A}{s_y} \text{ and } b = N_A - \frac{s_A}{s_y} \bar{y}, \quad (2)$$

where \bar{y} is the average and s_y is the sample standard deviation for $\{y_j : 1 \leq j \leq n\}$, and N_A and s_A are arbitrary values chosen for normalization.

For fatigue data the life times are distributed over several orders of magnitude that the procedure in Eqs. (1) and (2) is applied to the natural logarithm of the life times. Let m be the number of different values of applied stress or strain, i.e., $\{\Delta\sigma_k : 1 \leq k \leq m\}$ or $\{\Delta\varepsilon_k : 1 \leq k \leq m\}$. Given $\Delta\sigma_k$ or $\Delta\varepsilon_k$ the associated life times are $\{N_{k,j} : 1 \leq j \leq n_k\}$ where n_k is its sample size. Let

$$y_{k,j} = \ln(N_{k,j}) \quad (3)$$

be the transformed life times. Substituting Eq. (2) into Eq. (1) leads to the following:

$$z_{k,j} = \frac{s_A}{s_{y,k}}(y_{k,j} - \bar{y}_k) + N_A. \quad (4)$$

Thus, the averages and sample standard deviations of $\{y_{k,j} : 1 \leq j \leq n_k\}$ and $\{z_{k,j} : 1 \leq j \leq n_k\}$ are equal to each other. The next step is to merge all the transformed $z_{k,j}$ values from Eq. (4) for $1 \leq j \leq n_k$ and $1 \leq k \leq m$. The purpose in using the merged values is to have a more extensive dataset for estimation of the cdf. This is especially critical for estimating the extremes of the cdf more accurately. Subsequently, an appropriate cdf $F_Z(z)$ is found that characterizes the merged data. It is assumed that this cdf also characterizes the subsets $\{z_{k,j} : 1 \leq j \leq n_k\}$ of the merged set. With this assumption and the linear transformation in Eq. (4), the cdfs for $\{y_{k,j} : 1 \leq j \leq n_k\}$ $F_{y,k}(y)$ can be derived from $F_Z(z)$ as follows:

$$F_{y,k}(y) = F_Z\left(\frac{s_A}{s_{y,k}}(y - \bar{y}_k) + N_A\right). \quad (5)$$

The approach is designated as the Fatigue Life Transformation (FLT). Recall that the above methodology is applied to natural logarithms. In order to make observations on the actual fatigue lives, the values must be changed back to actual cycles.

4 Flt Analysis for 2024-T4 Fatigue Life Data

To evaluate the effectiveness of the proposed FLT methodology, the fatigue life data summarized in Table 1 and shown on Fig. 1 is considered. Recall that the FLT is

applied to the natural logarithm of the fatigue data; see Eq. (3). The arbitrarily chosen values for N_A and s_A are 26 and 1, respectively. The rather large value for N_A was chosen to assure that $z_{k,j}$ in Eq. (4) is positive. The 222 FLT data are shown on Fig. 4, where the axes are labeled to be easily read. Each set of data for a given $\Delta\sigma_k$ are transformed using FLT. These transformed data are well grouped so that it is reasonable to merge them. Figure 5 shows the entire 222 FLT values merged into a common sample space. The FLT merged data contain approximately 7–10 times more data than those for each given $\Delta\sigma_k$. Thus, estimation for the cdf is necessarily more accurate which results in a better characterization of its lower tail. Also, notice that the cycles are transformed using the FLT procedure; they are not actual cycles to failure, i.e., they are not equivalent to the data shown on Fig. 1. The solid line is the maximum likelihood estimation (MLE) for a three-parameter Weibull cdf $W(\alpha, \beta, \gamma)$, where α is the shape parameter, β is the scale parameter, and γ is the location parameter. The form of $W(\alpha, \beta, \gamma)$ is

$$F(x) = 1 - \exp\{-(x - \gamma)/\beta\}^\alpha, \quad x \geq \gamma. \quad (6)$$

Fig. 4 FLT fatigue life data for 2024-T4 specimens given $\Delta\sigma$ [9]

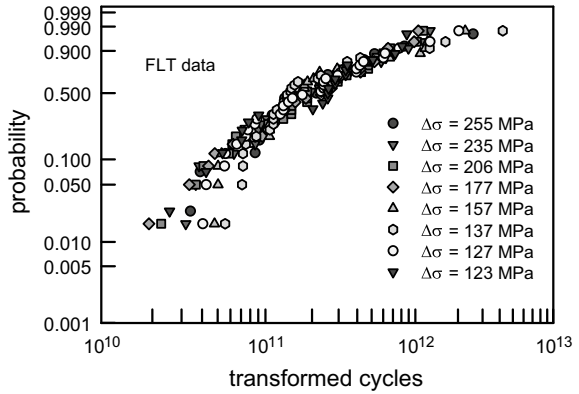
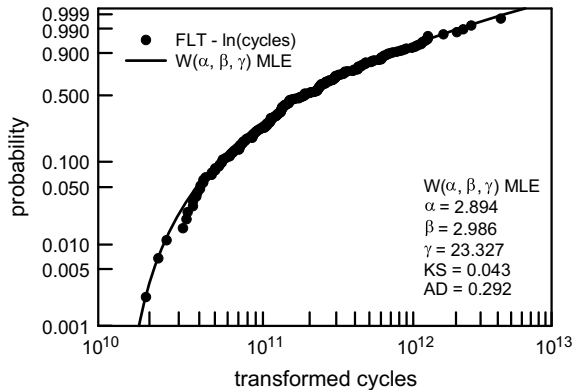


Fig. 5 Merged FLT fatigue life data for 2024-T4 specimens [9]; Weibull MLE



Graphically, the fit is excellent. The Kolmogorov-Smirnov (KS) and Anderson-Darling (AD) goodness of fit test statistics are 0.043 and 0.292, respectively. Both of which indicate that the MLE is acceptable for any significance level α_s less than 0.3. Consequently, the MLE $W(\alpha, \beta, \gamma)$ cdf is an excellent representation of the merged FLT data. The MLE estimated parameters are $\hat{\alpha} = 2.894$; $\hat{\beta} = 2.986$; and $\hat{\gamma} = 23.327$.

The three-parameter Weibull cdf $W(\alpha, \beta, \gamma)$ shown in Eq. 6 was selected for consideration because it has become a very popular cdf to represent fatigue data since its namesake used it for that purpose [18]. Two popular resources for the Weibull cdf, which advocate its use and contain examples of its applications, are [19, 20]. The primary reason for its choice, however, is because there is a minimum value represented by γ . Typically, any collection of fatigue life data is spread over two or three orders of magnitude. Consequently, a nonzero minimum value is required to appropriately represent the fatigue data.

Now, it is assumed that the MLE estimated parameters for the FLT merged data are acceptable for each of its subsets $\{z_{k,j} : 1 \leq j \leq n_k\}$. The cdf $F_{y,k}(y)$ for each given $\Delta\sigma_k$ can be determined using Eq. (5) as follows:

$$F_{y,k}(y) = 1 - \exp\left\{-\left[\left(\frac{s_A}{s_{y,k}}\right)(y - \bar{y}_k) + N_A - \hat{\gamma}\right]/\hat{\beta}\right\}^{\hat{\alpha}}. \quad (7)$$

Recall that the arbitrary constants N_A and s_A are 26 and 1, respectively. Equation (7) can be rewritten to put it into the standard Weibull cdf form $W(\alpha, \beta, \gamma)$;

$$F_{y,k}(y) = 1 - \exp\left\{-\left[(y - [\bar{y}_k + (\hat{\gamma} - N_A)\left(\frac{s_{y,k}}{s_A}\right)])/\hat{\beta}\left(\frac{s_{y,k}}{s_A}\right)\right]^{\hat{\alpha}}\right\}, \quad (8)$$

where the shape parameter $\hat{\alpha}$ is the same for each individual cdf, but the scale parameter $\hat{\beta}_k$ and location parameter $\hat{\gamma}_k$ are

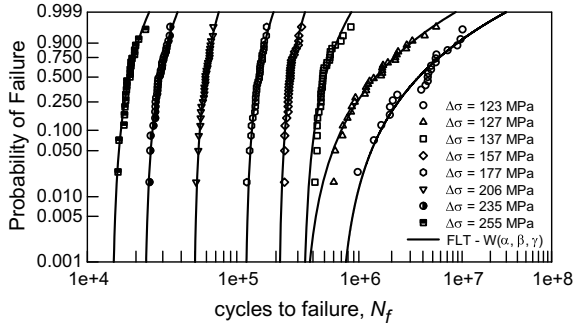
$$\hat{\beta}_k = \hat{\beta}\left(\frac{s_{y,k}}{s_A}\right); \text{ and } \hat{\gamma}_k = \bar{y}_k + (\hat{\gamma} - N_A)\left(\frac{s_{y,k}}{s_A}\right), \quad (9)$$

which are explicitly dependent on the sample parameters for the fatigue life data given $\Delta\sigma_k$ and the arbitrary constants N_A and s_A . Recall that the FLT is for $\ln(N_f)$; consequently, the range of values is considerably smaller than that for N_f . Table 4 contains the FLT cdf $W(\alpha, \beta_k, \gamma_k)$ parameters for each given $\Delta\sigma_k$. Figure 6 shows the fatigue life data plotted on two parameter Weibull probability paper with the corresponding FLT cdfs $W(\alpha, \beta_k, \gamma_k)$. Graphically, these cdfs appear to fit the data well. Indeed, the KS goodness of fit test indicates that all these cdfs are acceptable for any α_s less than 0.25. The AD test, which focuses on the quality of the fit in the tails, is more discriminating. The cdfs when $\Delta\sigma_k$ equals 127, 177, 206, 235, or 255 MPa are acceptable for any α_s less than 0.25. When $\Delta\sigma_k$ equals 123 or 157 MPa, however, the cdfs are acceptable for any α_s less than 0.05. When $\Delta\sigma_k$ equals 137, the AD test implies that the cdf is not acceptable. Although it is not obvious on Fig. 6

Table 4 FLT Weibull parameters for natural logarithm of fatigue life data for 2024-T4 specimens [9]

Load, $\Delta\sigma$ (MPa)	$\hat{\alpha}$	$\hat{\beta}_k$	$\hat{\gamma}_k$
255	2.894	0.280	9.556
235	2.894	0.258	10.032
206	2.894	0.214	10.799
177	2.894	0.261	11.657
157	2.894	0.247	12.259
137	2.894	0.491	12.705
127	2.894	1.678	12.690
123	2.894	1.972	13.375

Fig. 6 FLT Weibull cdfs for fatigue life data for 2024-T4 specimens [9]



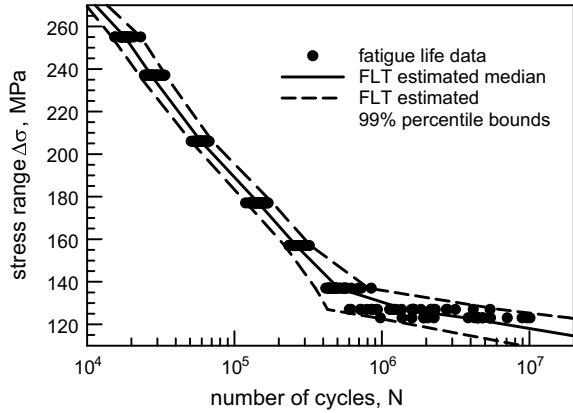
because the cycles scale is so large, both the upper and lower tail of the FLT cdf are sufficiently different from the life data. Even so, the overall deduction is that the FLT transformation is acceptable for these fatigue life data. Again, the KS test supports that conclusion, and there is only one value for $\Delta\sigma_k$, 137 MPa, for which the AD test indicates otherwise.

Another way to assess the quality of the proposed FLT methodology is to consider percentiles p of the estimated cdfs. The percentiles are given by

$$y_p = \hat{\gamma}_k + \hat{\beta}_k [-\ln(1 - p)]^{1/\hat{\alpha}}, \quad (10)$$

which are computed from Eqs. (8) and (9). One of the most common percentiles that is considered is the median $y_{1/2}$, i.e., p is 0.5. Figure 7 is an S-N graph, identical to Fig. 1, where the solid line is the FLT estimated median, and the dashed lines are the FLT estimated 99% percentile bounds. The 99% bounds are very tight while encapsulating the entire set of data for each $\Delta\sigma_k$. They also follow the trend in the S-N data in that they are very narrow when the life data have very little variability, but they are broader when the life data have larger variability. This lends credence to the FLT approach for the 2024-T4 fatigue data.

Fig. 7 Fatigue life data for 2024-T4 specimens [9] with FLT percentile bounds



5 FLT Analysis for ASTM A969 Fatigue Life Data

The second applications of the FLT method is for the ASTM A969 fatigue data. Again, the arbitrarily chosen values for N_A and s_A are 26 and 1, respectively. Figure 8 shows the entire 69 FLT values merged into a common sample space. The solid line is the MLE $W(\alpha, \beta, \gamma)$, Eq. (6). The KS and AD goodness of fit test statistics are 0.047 and 0.524, respectively. The MLE is acceptable according to the KS test for any significance level α_s less than 0.3. For the AD test, however, it is acceptable only for α_s less than 0.2 because there is some deviation between the data and the MLE in the lower tail. Even so, the FLT data are well represented by the MLE $W(\alpha, \beta, \gamma)$ cdf. The MLE estimated parameters are $\hat{\alpha} = 3.057$; $\hat{\beta} = 2.964$; and $\hat{\gamma} = 23.289$.

Using Eqs. (7)–(9), Fig. 9 has the ASTM A969 data with the FLT cdfs $W(\alpha, \beta_k, \gamma_k)$. Graphically, the FLT cdfs appear to characterize the data well. In fact, the KS test indicates that all these cdfs are acceptable for any α_s less than 0.3. The AD test, however, implies that the FLT cdfs are marginal, at best. Clearly, the FLT cdfs are

Fig. 8 Merged FLT fatigue life data for ASTM A969 specimens [13]; Weibull MLE

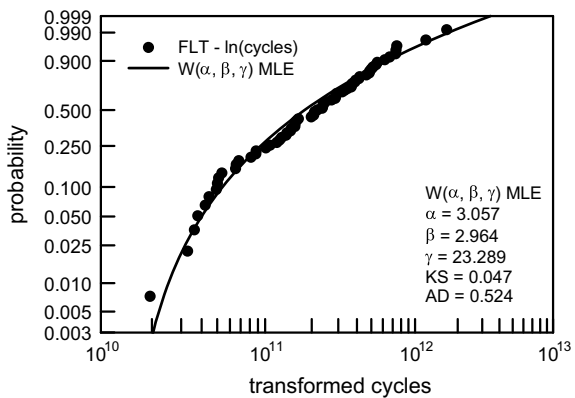


Fig. 9 FLT Weibull cdfs for fatigue life data for ASTM A969 specimens [13]

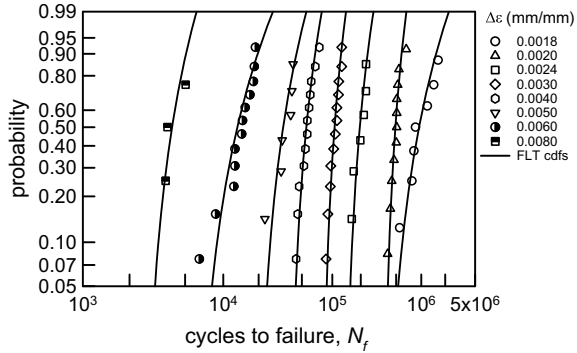
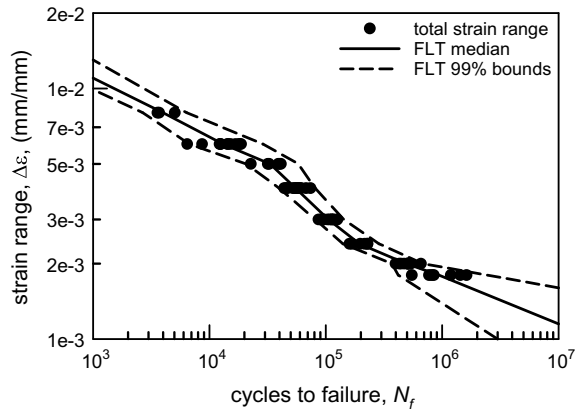


Fig. 10 Fatigue life data for ASTM A969 specimens [13] with FLT percentile bounds

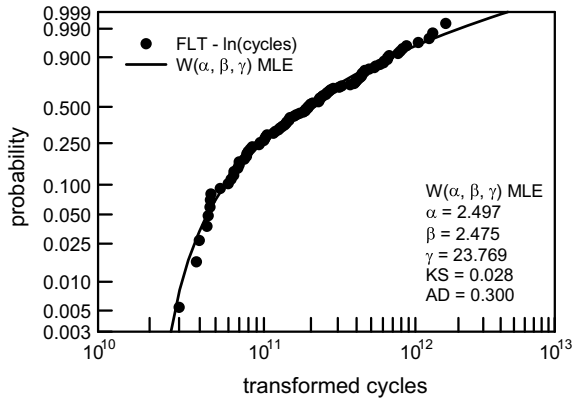


not as accurate in the tails. No doubt, larger samples for each $\Delta\epsilon_k$ would help with characterization of the extremes. Using Eq. (10), Fig. 10 is an S-N graph, identical to Fig. 2, with the addition of the FLT estimated median, and the FLT estimated 99% percentile bounds. The 99% bounds are very tight, and the all the data are within the bounds for each $\Delta\epsilon_k$. They are somewhat jagged because they follow the pattern of the S-N data. The analysis is not as crisp as that for the 2024-T4 data; however, there seems to be merit in using the FLT approach for the ASTM A969 fatigue data.

6 FLT Analysis for 9Cr-1Mo Fatigue Life Data

The third application of the FLT method for the 9Cr-1Mo fatigue data does not perform very well. Figure 3 shows unusually large scatter for the higher values of $\Delta\epsilon_k$ which is a good test for the FLT methodology. The arbitrary scaling factors are the same; N_A and s_A are 26 and 1, respectively. Figure 11 shows the 130 FLT merged values with the MLE $W(\alpha, \beta, \gamma)$. The KS and AD goodness of fit test statistics

Fig. 11 Merged FLT fatigue life data for 9Cr-1Mo specimens [16]; Weibull MLE



are 0.028 and 0.300, respectively. The MLE is acceptable according to the KS and AD tests for any significance level α_s less than 0.3. The merged data are very well characterized by this cdf. The corresponding MLE estimated parameters are $\hat{\alpha} = 2.497$; $\hat{\beta} = 2.475$; and $\hat{\gamma} = 23.769$.

The FLT cdfs $W(\alpha, \beta_k, \gamma_k)$, Eqs. (7)–(9), for 9Cr-1Mo are shown on Fig. 12. Graphically, the FLT cdfs appear to be acceptable, at least for the cases with more data. Clearly, when $\Delta\epsilon_k$ is 0.019, the fit is borderline. The KS test indicates that all these cdfs are acceptable for any α_s less than 0.3. On the other hand, the AD test indicates that none of the FLT cdfs are acceptable. The compressed graphical scale for the cycles to failure masks the poor fit of the FLT cdfs to the tails of the data. Figure 13 is the S-N graph, Fig. 3, with the FLT estimated median, and the FLT estimated 99% percentile bounds. Because the FLT cdfs are not very representative of the fatigue data, the 99% bounds are erratic. All the data are within the bounds for each $\Delta\epsilon_k$. That is somewhat positive. They are quite jagged because they follow the S-N data pattern. One of the difficulties in modeling these data are that there are three values for $\Delta\epsilon_k$ that have several replicates, but the other values have only a few. Also the scatter in the data when $\Delta\epsilon_k$ is 0.020 appears to be larger than expected.

Fig. 12 FLT Weibull cdfs for fatigue life data for 9Cr-1Mo specimens [16]

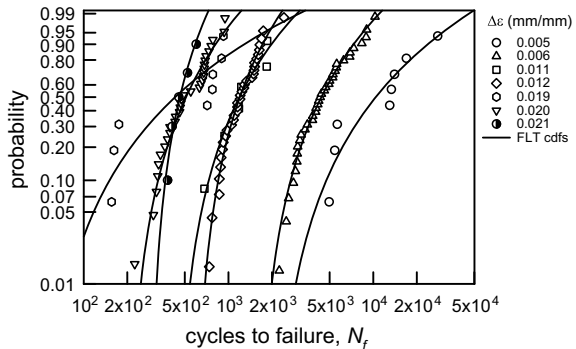
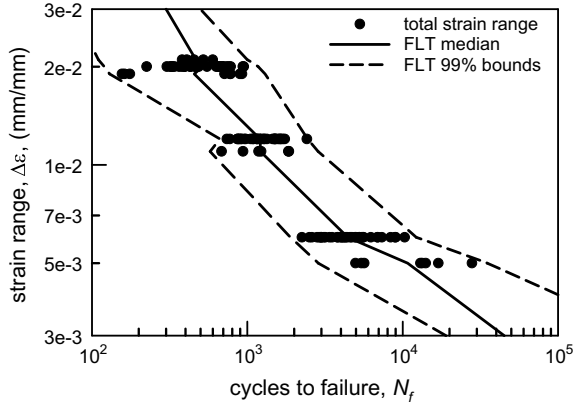


Fig. 13 Fatigue life data for 9Cr-1Mo specimens [16] with FLT percentile bounds



Since there is no explanation for this in [16], the reason would purely speculation. The analysis is marginal for the 9Cr-1Mo data. There may be merit in using the FLT approach for some insight, but conclusions would need to be made very cautiously.

7 Observations and Conclusions

Three sets of fatigue life data were considered for the proposed FLT method. These datasets were selected because of they have replicate data for several applied stress or strain ranges, $\Delta\sigma_k$ or $\Delta\epsilon_k$, respectively. The primary reason for the proposed FLT method is to accurately model the statistical nature of fatigue life data. Specifically, the estimation of underlying cdfs is crucial. It is well known that fatigue life data have rather large amounts of variability particularly for applied loads similar to typical operating conditions. Modeling these data is very challenging. Associated with this is that many fatigue life data sets have relatively few choices for $\Delta\sigma_k$ or $\Delta\epsilon_k$, and each choice has limited observations.

The FLT introduced in this paper attempts to help improve fatigue life modeling by using a statistically based transformation to merge data, thereby increasing the effective sample size. The FLT approach transforms the fatigue life data for each given $\Delta\sigma_k$ or $\Delta\epsilon_k$ so that the averages and standard deviations are the same. Subsequently, the data are merged, and a suitable cdf is statistically estimated for the entire collection. The cdf for each given $\Delta\sigma_k$ or $\Delta\epsilon_k$ is obtained by standard change-of-variables methods using the cdf that characterizes the entire transformed and merged data.

Using the 2024-T4 fatigue life in [9], the FLT is very promising partly because there are eight different values for $\Delta\sigma_k$ and a total of 222 data being considered. A three-parameter Weibull cdf $W(\alpha, \beta, \gamma)$ is an excellent representation of the merged FLT data, and a $W(\alpha, \beta, \gamma)$ is also appropriate characterization for the underlying cdfs given $\Delta\sigma_k$. Consequently, there is assurance that the lower tail behavior is adequately

modeled because of the methodology. Additionally, the validation for the approach is the computed FLT percentiles for the S-N data. The computed FLT 99% percentiles fully encompass all the S-N data, but more importantly, the bounds are quite tight. The conclusion from this analysis is that the FLT methodology is warranted.

To corroborate this conclusion, two other sets of fatigue life data were considered; ASTM A969 [13] and 9Cr-1Mo [16]. In both cases, the merged FLT data are well characterized by a $W(\alpha, \beta, \gamma)$. The KS test indicates that the transformed $W(\alpha, \beta, \gamma)$ is also suitable for the underlying cdfs given $\Delta\varepsilon_k$, but the AD test is more discriminating. The tail behavior of the underlying cdfs given $\Delta\varepsilon_k$ are marginally acceptable at best, if at all. The 99% percentiles for the ASTM A969 S-N data are quite good. They encompass the data, and they are tight. For the 9Cr-1Mo S-N data, however, the percentile lines are not very regular. They do encompass the data, but the data have so much variability that little is gained by the analysis.

Based on these three examples, the FLT approach should be employed for fatigue life data analysis when an empirical method is desired. The FLT method excellent for one of the cases, 2024-T4, acceptable for another case, ASTM A969, and marginal for the other case, 9Cr-1Mo. As with all empirical analyses, caution must be exercised when it is implemented. Limited applied loads with limited replicate data for each load hinders accurate modeling for any method including the FLT. As with all empirical methods, the more data there is, the better the accuracy will be. The example which was the worst, 9Cr-1Mo, seems to be poor because there is overly large scatter in the data for the higher loading conditions coupled with applied loads with only a few replicates. Again, the FLT methodology should be implemented with care.

Many sets of experimental fatigue life data contain censoring. This will be investigated in the future. In this case the cdf estimation is more advanced, especially for a three-parameter cdf. In principle the FLT methodology should be similar except for the adjustment for censoring. All things considered, the proposed FLT approach has sufficient promise that further investigation and analysis is certainly warranted. The overarching observation is that the FLT approach is useful if the fatigue data are reasonably well behaved.

References

1. Harlow DG, Wei RP, Sakai T, Oguma N (2006) Crack growth based probability modeling of S-N response for high strength steel. *Int J Fatigue* 28:1479–1485
2. Castillo E, Fernández-Canteli A (2009) A unified statistical methodology for modeling fatigue damage. Springer Science and Business Media B.V., Berlin
3. Schneider CRA, Maddox SJ (2003) Best practice guide on statistical analysis of fatigue data. International Institute of Welding, IIW-XIII-WG1-114-03, United Kingdom
4. Collins JA (1993) Failure of materials in mechanical design: analysis, prediction, prevention, 2nd edn. Wiley-Interscience Publications, New York
5. Little RE, Ekvall JC (eds) (1981) Statistical analysis of fatigue data. ASTM STP744, American Society for Testing and Materials, Philadelphia

6. Zheng X-L, Jiang B, Lü H (1995) Determination of probability distribution of fatigue strength and expressions of P-S-N curves. *Eng Fracture Mech* 50:483–491
7. Shimizu S (2005) P-S-N/P-F-L curve approach using three-parameter Weibull distribution for life and fatigue analysis of structural and rolling contact components. *Tribol Trans* 48:576–582
8. Fernández-Canteli A, Blasón S, Correia JAFO, de Jesus AMP (2014) A probabilistic interpretation of the miner number for fatigue life prediction. *FratturaedIntegritàStrutturale* 30:327–339
9. Shimokawa T, Hamaguchi Y (1985) Relationship between fatigue life distribution, notch configuration, and S-N curve of a 2024-T4 Aluminum Alloy. *Trans ASME J Eng Mater Technol* 107:214–220
10. Harlow DG (2005) Probability versus statistical modeling: examples from fatigue life prediction. *Int J Reliab Qual Saf Eng* 12:1–16
11. Harlow DG (2011) Generalized probability distributions for accelerated life modeling. *SAE Int J Mater Manuf* 4:980–991
12. Sakai T, Nakajima M, Tokaji K, Hasegawa N (1997) Statistical distribution patterns in mechanical and fatigue properties of metallic materials. *Mater Sci Res Int* 3:63–74
13. IF_DDQ_HDG70G_Strain_Life_Fatigue.xls www.autosteel.org (2004). Automotive Applications Council, Steel Market Development Institute, Auto/Steel-Partnership Fatigue Data
14. Harlow DG (2014) Low cycle fatigue: probability and statistical modeling of fatigue life. In: *Proceedings of the ASME 2014 pressure vessels and piping conference, PVP2014–28114*, Anaheim, CA, 20–24 July 2014, V06BT06A045. ISBN 978–0–7918–4604–9
15. Harlow DG (2017) Confidence bounds for fatigue distribution functions. In: *Proceedings of the ASME 2017 pressure vessels and piping conference, PVP2017–65416*, Waikoloa, HI, 16–20 July 2017. V01AT01A011. ISBN 978–0–7918–5790–8
16. Thomas GB, Varma RK (1993) Evaluation of low cycle fatigue test data in the BCR/VAMAS intercomparison programme. Commission of the European Communities, BCR Information Applied Metrology, Brussels, Luxembourg
17. Harlow DG (2007) Probabilistic property prediction. *Eng Fract Mech* 74:2943–2951
18. Weibull EHW (1951) A statistical distribution function of wide applicability. *ASME J Appl Mech* 18:293–297
19. Abernethy RB (2006) In: Abernethy RB (ed) *The new Weibull handbook*, 5th edn. 536 Oyster Road, North Palm Beach, Florida, pp 33408–34328
20. Rinne H (2009) *The Weibull distribution*. CRC Press, Chapman-Hall, Taylor and Francis Group, Boca Raton

D. Gary Harlow is a Professor in Mechanical Engineering and Mechanics at Lehigh University in Bethlehem, PA. Since 2008 he has been the department chair. He received his BA in Mathematics (1973) from Western Kentucky University and his MS (1976) and Ph.D. in Applied Probability and Stochastic Processes (1977) from Cornell University. His research has primarily centered on scientifically and mechanistically based probability modeling and statistical analyses. Specifically, his research has included modeling of failure processes (creep, creep-fatigue, very high cycle fatigue, corrosion pitting, and corrosion fatigue) in materials (aluminum alloys, steels, and composites); stochastic fracture mechanics; mechanical and system reliability; applications of stochastic processes; and methods in applied statistics for materials characterization. His research has been sponsored throughout his career by the NSF, FAA, DoD, and DARPA. He is a fellow of ASME.