NASA Contractor Report 3697

Statistical Summaries of Fatigue Data for Design Purposes

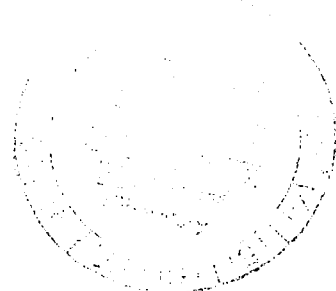Paul H. Wirsching

25

25th Anniversary
1958-1983

NASA

NASA Contractor Report 3697

Statistical Summaries of Fatigue Data for Design Purposes

Paul H. Wirsching
The University of Arizona
Tucson, Arizona

NASA

National Aeronautics
and Space Administration

**Scientific and Technical
Information Branch**

1983

ABSTRACT

Fatigue data are subject to considerable scatter and cycles to
failure, N, can have coefficients of variation typically ranging
from 25% to 75%. Presented herein are techniques for providing
statistical summaries of such data suitable for design purposes.
Special consideration is given to statistical difficulties presented
by the small sample sizes characteristic of fatigue data.

Two methods are discussed for constructing a design curve on the
safe side of the data. Both the tolerance interval and equivalent
prediction interval (EPI) concepts provide such a curve while account-
ing for both the distribution of the estimators in small samples and
the data scatter. Moreover the EPI is useful as a mechanism for pro-
viding necessary statistics on S-N data for a full reliability anal-
ysis which includes uncertainty in all fatigue design factors.

Presented are examples of statistical analyses of the general
strain life relationship. The tolerance limit and EPI techniques
for defining a design curve are demonstrated. Moreover in two ex-
amples using Waspaloy B and RQC-100 data it was demonstrated that a
reliability model could be constructed by considering the fatigue
strength and fatigue ductility coefficients as two independent ran-
dom variables.

A technique for establishing the fatigue strength for high cycle
lives was presented. This method relies on an extrapolation technique
and also accounts for "runners." Again, a reliability model or design
value can be specified.

INTRODUCTION

In general fatigue is perhaps the most important failure mode to be considered in mechanical and structural design. For some products, fatigue accounts for more than 80% of all observed service failures. Moreover, fatigue and fracture failures are sometimes catastrophic, occurring without warning and causing significant property damage and loss of life. Design for fatigue avoidance is difficult, because (a) the fatigue stresses are complicated random processes, (b) the fatigue process is influenced by many factors, and (c) many of the factors are subject to considerable uncertainty.

A major source of uncertainty is introduced by the enormous scatter in fatigue test data, with cycles to failure data having coefficients of variation typically ranging from 30 to 40% and sometimes higher than 100%. Figure 1 shows the data base for the AWS-X curve for welded tubular connections and illustrates that the scatter spans more than two orders of magnitude. The designer's decision reflecting a proper balance between risk and cost should be based on suitable consideration of such uncertainties in fatigue strength.

Presented herein is a commentary on fatigue data analysis, including a review of the literature, statistical summaries of fatigue test data, and techniques used to produce such summaries. The focus is on design application. The discussion will be restricted to the S-N approach to characterize fatigue data. This description is primarily used for small specimen data for crack initiation life estimates, but it can also apply to total life for larger structural sections.

Fig. 1

An Example of Scatter in S-N Data:
This Data is Used to Define AWS-X Curve
[See Ref. 39]
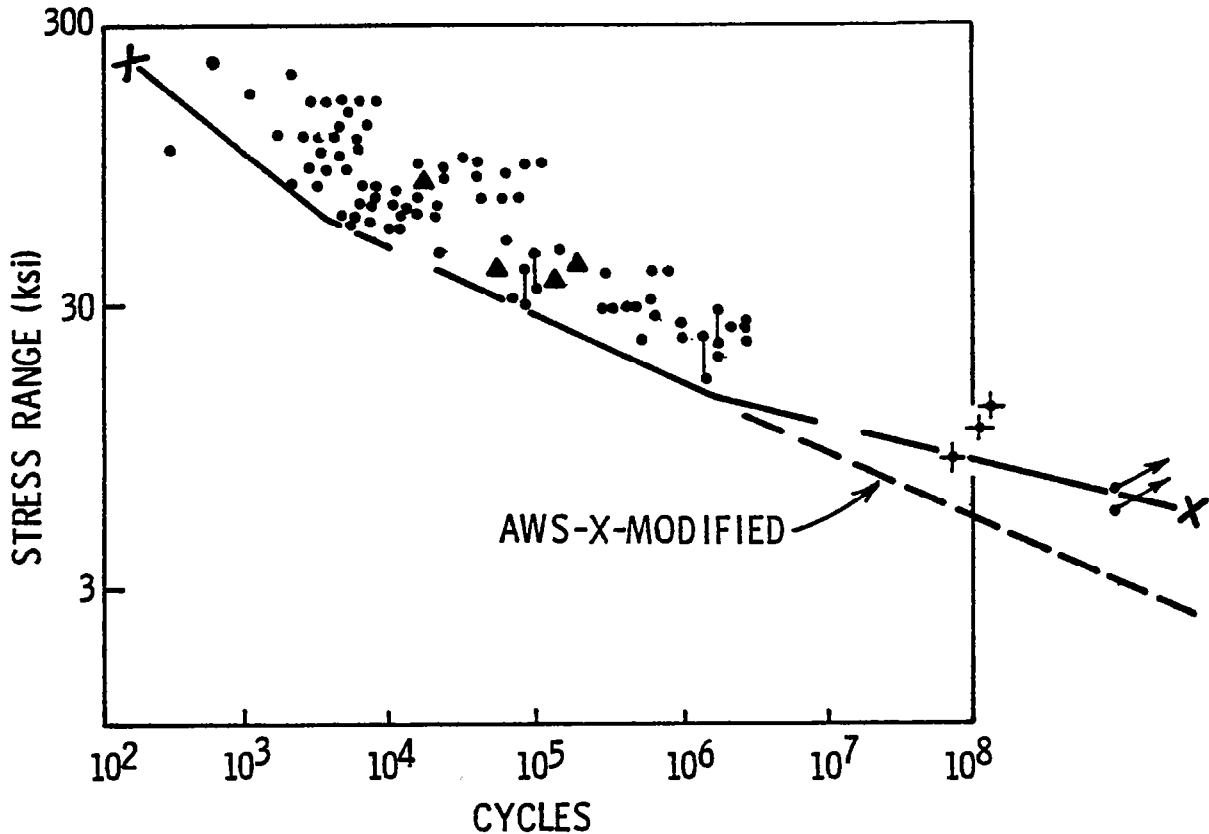
It is not the intent of this paper to make specific recommendations on how designers should model fatigue strength. Rather the purpose is to provide general information on statistical methods which will provide rational characterizations of fatigue data for design purposes. Moreover some new understanding regarding the general nature of fatigue data are presented.

2

Modern structural reliability theory has its origins in a landmark paper by A. M. Freudenthal which appeared in the 1947 Transactions of the American Society of Civil Engineers (10). Later he and E. J. Gumbel collaborated in the development of methods of fatigue reliability focusing on statistical modeling of S-N data (11), (12), (15).

The ASTM has played an active role in development of statistical methods of fatigue data analysis dating back to 1951 (33,34). Currently, subcommittee E09.06 on Statistical Aspects of Fatigue is engaged in a variety of activities relating to data analysis. In 1963 Committee E-9 on Fatigue published a guide for statistical analysis of fatigue data (1). Recently, ASTM published a work by Little dealing with statistical planning and analysis (23).

In 1973, J.T.P. Yao was instrumental in forming an ASCE Committee on Fatigue and Fracture Reliability devoted to design methods. A part of the activity of this committee deals with statistical data analysis, and it recently published a state of the art summary (8).

Numerous references are available in addition to those cited. A sample of some which provide general information on fatigue data analysis and reliability methods include the works of Lipson and Sheth (21), Little and Jebe (22), Collins (5), Yao (43), Yang (42), Kececioglu (18), Haugen (16), and Wirsching (40).

Data and information were provided by Kelly Donaldson of MTS Systems Corporation, Dennis Wolski of the Garrett Turbine Engine Company, and Bill Stamper of the Cummins Engine Company.

ENGINEERING MODELS USED TO DESCRIBE FATIGUE BEHAVIOR

The classical approach to fatigue has focused on the S-N diagram
(e.g. Fig. 1) which relates fatigue life (cycles to failure, N) to
cyclic stress amplitude S_a (or cyclic stress range S_R). Since "failure"
is usually defined specifically for a particular application, this constant-
amplitude S-N diagram can be used to relate stress to either crack initiation
period or total fatigue life. Most of the probabilistic analysis as presented
herein is equally applicable if N is either initiation period or total life
to failure. However, the physical distinction should be noted.

Baseline fatigue data are usually obtained by cycling test spec-
imens at constant-amplitude stress S (or strain) until visible crack-
ing or failure occurs. Such tests are repeated several times at different
stress levels to establish the familiar S-N curve correlating stress S
(or strain) to cycles required to initiate a fatigue crack, N. This pro-
cess is illustrated in Fig. 2.

Fatigue data typically have significant scatter which can be de-
scribed by the probability density functions (pdf) as shown in Fig. 2,
i.e., $f_{N|S}$ is the pdf of N given S and $f_{S|N}$ is the pdf of S given N.
The goal of statistical analysis is to provide summary descriptions of
the data for designers. Most commonly a design curve, as shown in Fig. 2,
on the safe side of the data is required. But for a reliability format,
the designer needs simple statistical representations of the distribution.

Fig. 2
Typical Fatigue Test
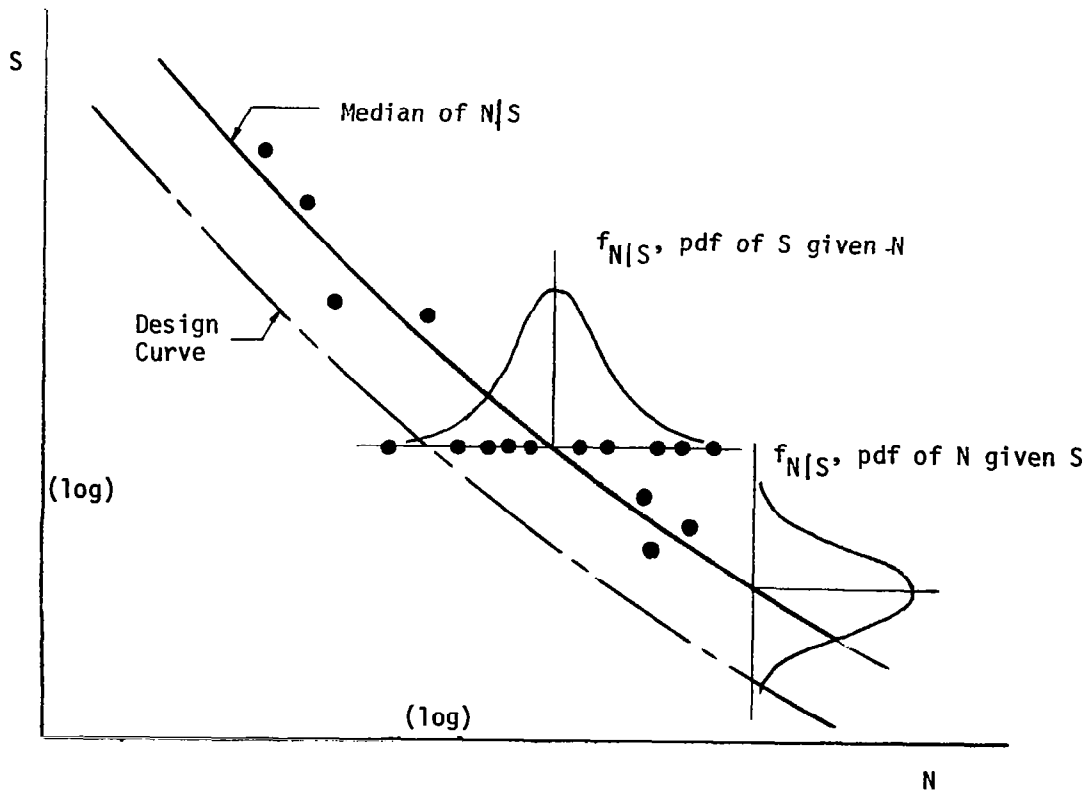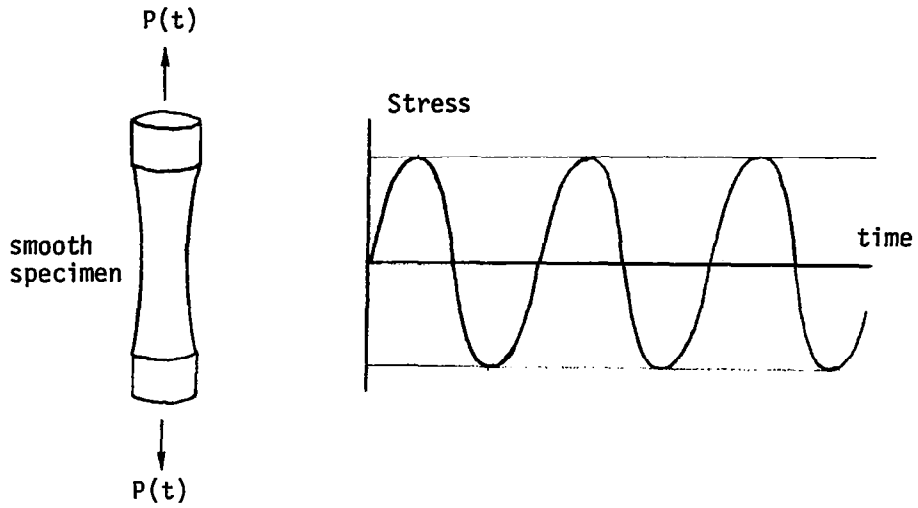
Models which characterize the S-N relationship for design purposes are described as follows.

Classical Model A commonly used S-N relationship first proposed by Basquin (3) has the form

$$NS^m = A \qquad (1)$$

where S is stress amplitude S_a, or stress range S_R, and m and A are empirical constants. Equation 1 is generally valid for the high cycle range ($N > 10^4$). In the case where a mean stress S_o is present, the term A is replaced by $A(1 - S_o/S_u)^m$ where now A corresponds to the value for zero mean tests and S_u is the ultimate strength of the material. Equation 1 plots as a straight line on log-log paper as shown in Fig. 3a.

General Strain-Life Model

The general strain-life model is now being widely used to describe strain controlled small specimen fatigue behavior over a wide range of strains (13, 19, 20). This model, summarized in Fig. 3b, considers elastic strain and plastic strain life separately. The two are added to obtain the total strain-life curve.

$$\varepsilon_a = \frac{\sigma_f'}{E} (2N)^b + \varepsilon_f' (2N)^c \qquad (2)$$

where ε_a = strain amplitude (specimens are strain cycled)

E = modulus of elasticity

σ_f' = fatigue strength coefficient

b = fatigue strength exponent

ε_f' = fatigue ductility coefficient

c = fatigue ductility exponent

Fig. 3
Fatigue Strength Models

(a) Typical High Cycle S-N Curve

$$NS^m = A$$

S

(log)

1

m

(log)

N

(b) General Strain-Life Curve

ε

Strain Amplitude

(log)

Total

$$\varepsilon_a = \frac{\sigma_f'}{E}(2N)^b + \varepsilon_f'(2N)^C$$

Stable Hysteresis Loop

σ

ε

$$\Delta\varepsilon_p \quad \Delta\varepsilon_e$$

$$\varepsilon_{pa} = \Delta\varepsilon_p/2$$

$$\varepsilon_{ea} = \Delta\varepsilon_e/2$$

$$\varepsilon_a = \varepsilon_{ea} + \varepsilon_{pa}$$

Plastic

$$\varepsilon_{pa} = \varepsilon_f'(2N)^C$$

Elastic

$$\varepsilon_{ea} = \frac{\sigma_f'}{E}(2N)^b$$

(log)

Cycles to Failure

N

7
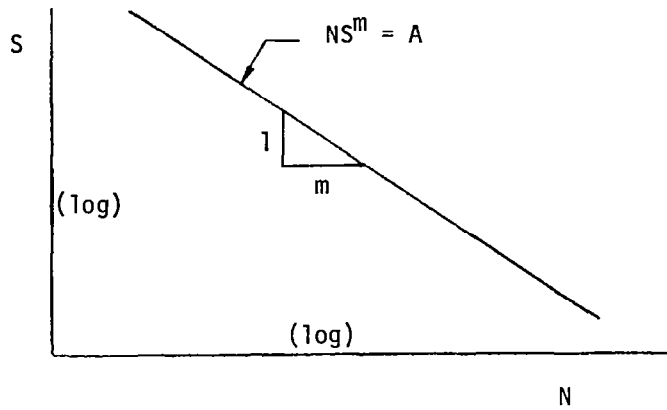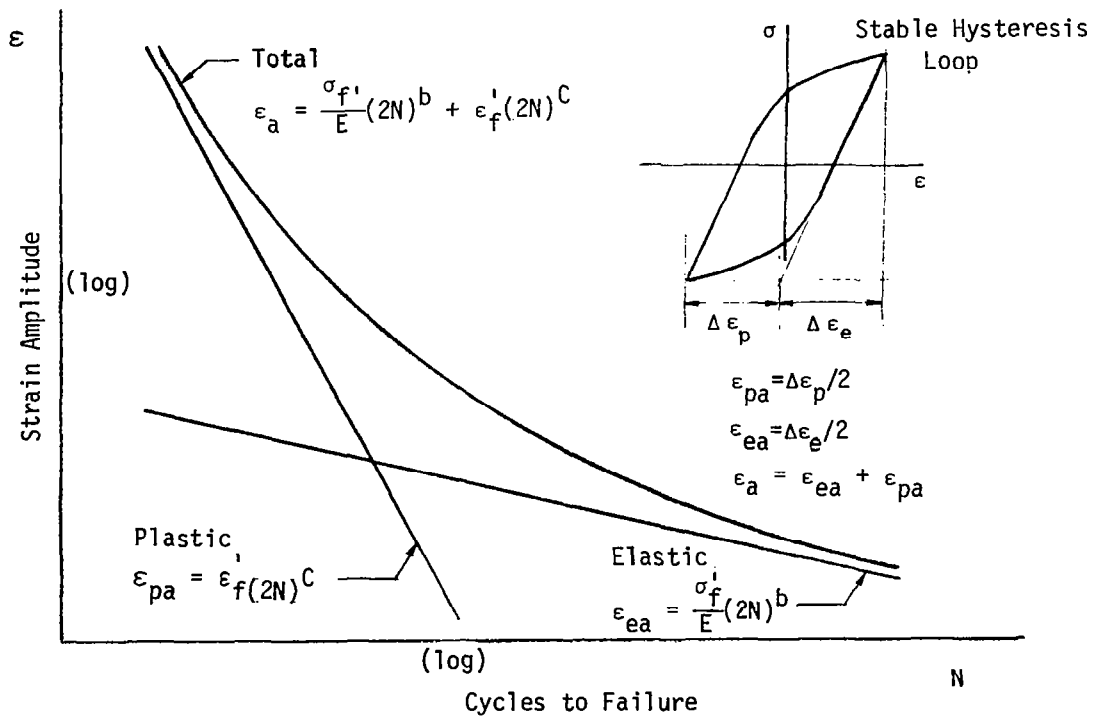
In the case where mean stress S_O is present, the term σ'_f can be replaced by $(\sigma'_f - S_O)$. A list of parameters for several materials is given in the SAE Handbook (29) and by Boardman (4).

STATISTICAL MODELS USED TO ANALYZE CYCLES TO FAILURE DATA

To make design decisions on the basis of a set of observations of a design factor, it is necessary to describe the distribution of that factor. In that regard, statistical models are usually employed. The random variable N denoting cycles to failure is usually described with a two parameter Weibull or lognormal model, but sometimes the three-parameter Weibull is used. A summary of these models as well as the normal and the three-parameter Weibull is provided in Ref. (40).

Use of the lognormal distribution has been based primarily on arguments of mathematical expediency. However, it has been pointed out by Gumbel (15) that the hazard function for the lognormal model decreases for large values of N. This does not agree with our physical understanding of progressive deterioration resulting from the fatigue process. Nevertheless the lognormal often seems to provide a "good fit" of cycles to failure data (see below).

Physical arguments favor the Weibull for most material strength variables, because it is an asymptotic distribution of minima of a sample (14). If failure of a structural element is precipitated by failure of the first of a large number of subelements, then the Weibull is likely a "good" model. Moreover, the Weibull has an increasing hazard function. However, the Weibull shape simply doesn't match the data for the large scatter data ofter observed in fatigue; the mode approaches zero as the coefficient of variation of N, C_N, approaches one.

8

A statistical test which compares various competing models on the basis of goodness of fit to a random sample has been developed by Wirsching and Carlson (36). The test, based on a modified form of the Cramer-von Mises statistic, was used to examine welded joint fatigue data to compare Weibull and lognormal distribution for consistent fitting of fatigue data.

Basically the test involves computing a "W-statistic" which is the sum of the deviation between the empirical and hypothesized cdf's (36). The summary presented in Table 1 compares the W-statistic for four distributions. The smaller the value of W, the better fit. The smallest value (circled) suggests the best fit distribution. The lognormal (LN) is clearly the winner, and in general this author has found that comparison tests on cycles to failure data favor the lognormal over the Weibull.

Comparisons of distributions should be made on the basis of fit in the tail regions, suggesting a large sample size requirement. Wirsching and Carlson (36) indicated that a sample size of at least 100 was required for good distributional resolution. Because of the expense of fatigue testing, seldom are such data sets available. In this regard it is interesting to note from Table I that the larger sample size data sets also favor the choice of the lognormal.

9

Table I

STATISTICAL TEST TO DETERMINE WHICH MODEL BEST FITS CYCLES TO FAILURE DATA
ON WELDED TUBULAR JOINTS

Investigator [**]	Sample Size n	W Statistic (Best fit is circled)			
		NOR[*]	LN	WEI	EVD[*]
Dijkstra & DeBack (168 mm chord)	10	.171	(.084)	.124	.131
Dijkstra & DeBack (457 mm chord)	19	.089	(.050)	.057	.068
Dijkstra & DeBack (914 mm chord)	11	.060	.066	(.059)	.059
AWS-X (Elastic Range)	60	.123	(.027)	.052	.083
Marshall Corrosion Fatigue	34	.144	(.026)	.042	.084
Hartt (Corrosion Fatigue)	7	.088	.079	.090	(.068)
Bouwkamp et al.	14	.067	.068	.053	(.052)
Toprac & Louis	9	(.041)	.069	.044	.073
Kurobane & Konomi	20	.174	(.049)	.092	.115
Maeda et al.	27	.115	(.035)	.042	.084
Kurobane et al.	13	.063	.048	(.045)	.048
Toprac & Louis	29	.066	(.034)	.039	.036
Gibstein	6	.114	(.059)	.093	.090
Wylde	5	.124	(.086)	.115	.106

[*] The normal (NOR) and extreme value distribution (EVD) are not serious contenders. They are included only for reference purposes.

[**] The data, analysis of that data, and the reference for each of the data sets are given in Ref. 37 and 39.

In practice, the lognormal is more commonly used by designers. Some of the reasons include: (1) the lognormal model generally has been shown to provide a reasonable description for the distribution of a wide variety of design variables, (2) statistical properties of the lognormal distribution are well defined, (3) the lognormal is easy to use in probabilistic design, (4) reliability formats using the lognormal can easily accommodate design variables having relatively large coefficients of variation, (5) the lognormal is already widely used in the design profession. For example, commonly used methods of linear model analysis for characterizing S-N fatigue data implicitly assume that cycles to failure has a lognormal distribution.

The three-parameter Weibull (TPW) is often proposed in the fatigue literature as an appropriate model for N. The cumulative distribution function (cdf) has the form

$$F_N(x) = P(N \leq x) = 1 - \exp\left[-\left(\frac{x-\xi}{\beta}\right)^\eta\right] \quad \text{for } x > \xi \quad (3)$$

where $P(\cdot)$ is "probability of." The TPW appears to be attractive because the location parameter ξ defines a non-zero lower bound on the sample space. In theory, such a model seems more realistic than the two-parameter models which permit values (albeit with small probability) down to zero.

Undesirable features of the TPW which may make its use impractical in certain cases are described in some detail in Ref. (40).

11

In summary the TPW is difficult to use. Parameter estimation requires non-trivial numerical analysis as does a full reliability analysis which includes the TPW. Moreover, credibility in the TPW wanes when the estimated location parameter $\hat{\xi}$ falls only slightly below the smallest sample point.

Probability plotting is a tool which is widely used for data analysis to (a) provide a subjective test of the hypothesis that a set of data was sampled from a given distribution family, (b) obtain estimates of the parameters, and (c) make probability calculations. The empirical distribution function, F_i, an estimate of the distribution function $F_N(n)$, can be established as described in Ref. (40). An example of a probability plot is the data of Sinclair and Dolan (32) plotted on lognormal paper as shown in Fig. 4. The empirical distribution functions illustrate the statistical scatter which is typical in fatigue data. C_N is the coefficient of variation of N (standard deviation divided by the mean). These data published in 1953 were intended to show at that time that observed scatter in cycles to fatigue is indeed an inherent characteristic of the material and not due to poor testing techniques. These plots illustrate (a) the enormous scatter typical of cycles to failure data, (b) the heteroscedastic nature of the data, i.e., scatter is a function of the stress level, and (c) the lognormal distribution seems to provide a reasonable fit.

HOW STRESS (OR STRAIN)-LIFE DATA IS ANALYZED: THE BASIC LINEAR MODEL

Fatigue data typical of the general strain life relationship, strain range partitioning, traditional S-N model, etc. are illustrated in Fig. 5. It is necessary to analyze such data for design purposes. The following two basic methods are employed: (1) define a design curve on the safe (lower) side of the data, or (2) present a statistical summary for a

12

Figure 4

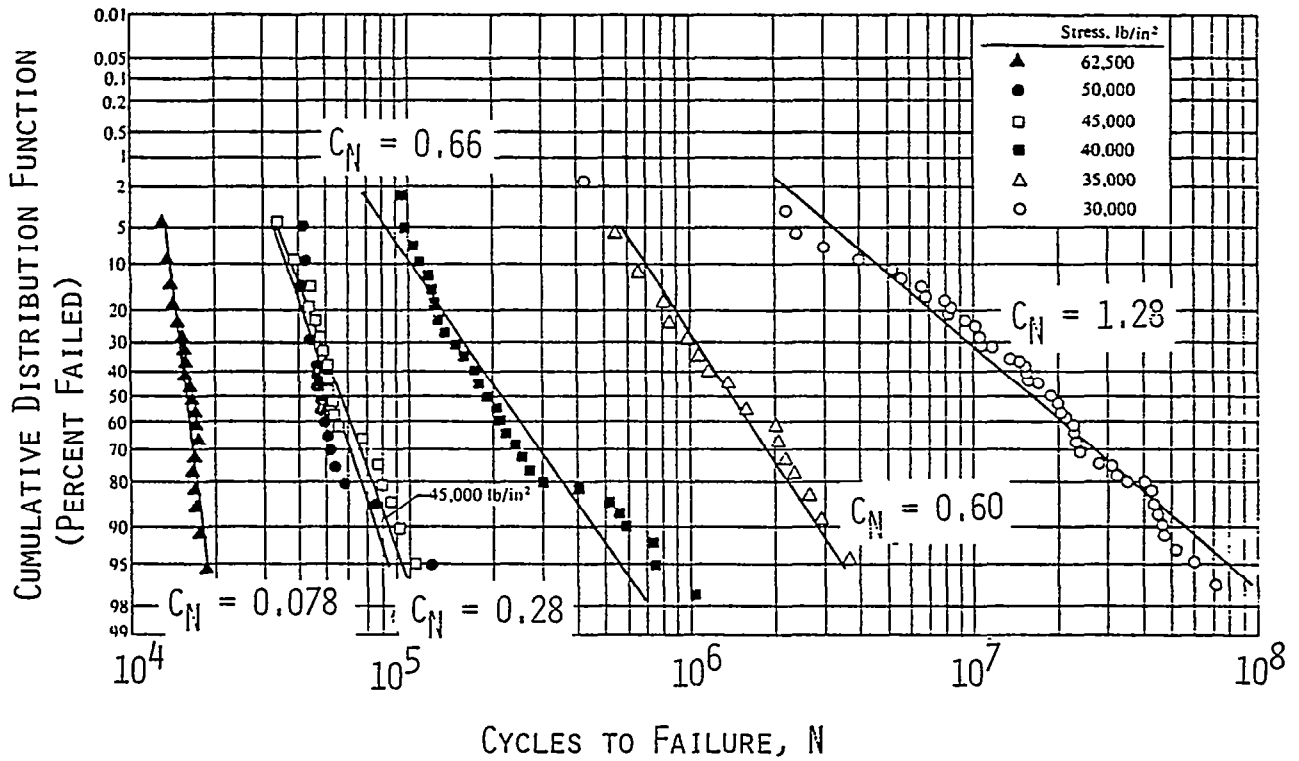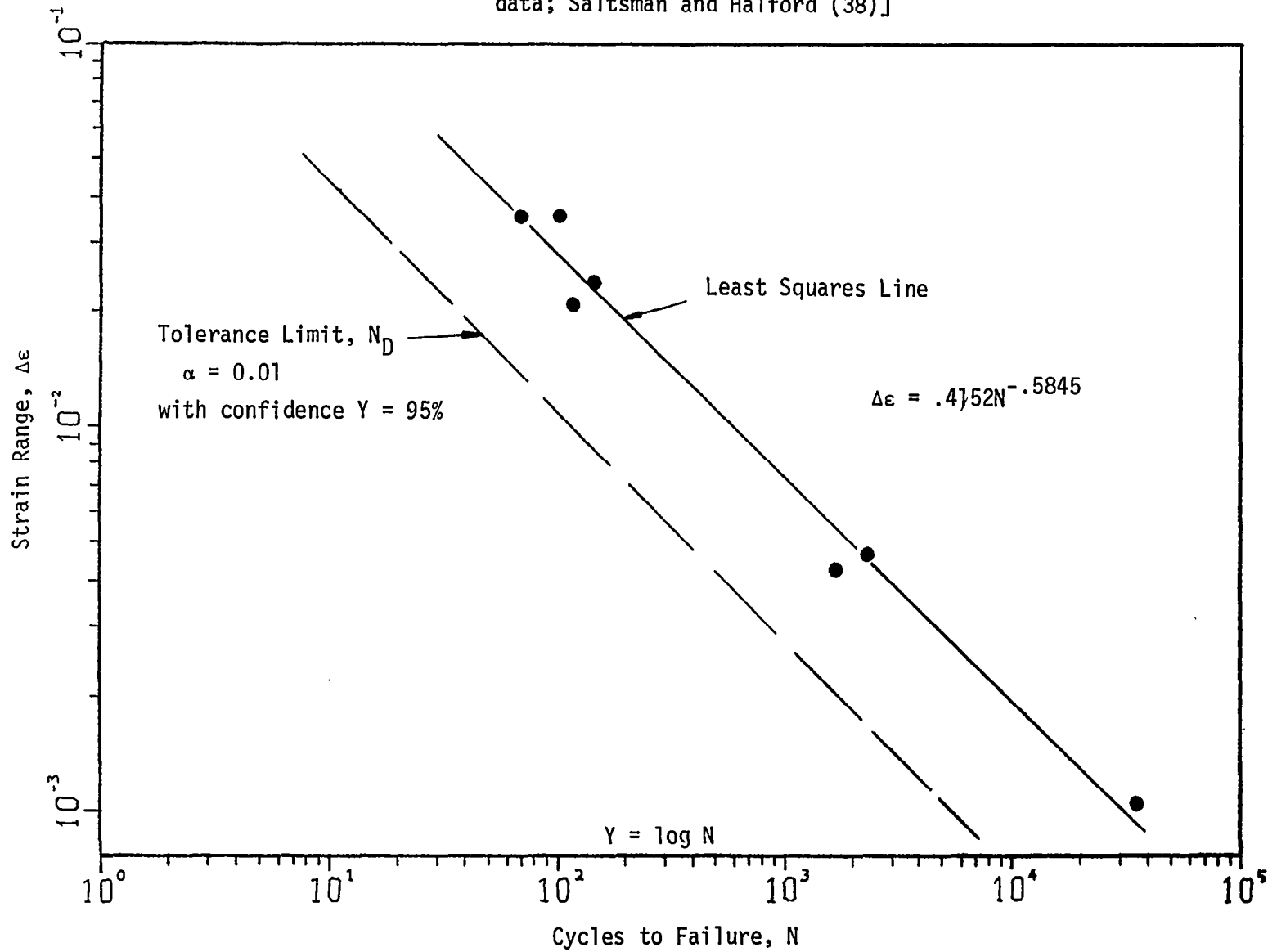EXAMPLE OF STATISTICAL VARIABILITY IN
LABORATORY FATIGUE DATA [after Sinclair
and Dolan (32)]

13

Fig. 5   Use of the Tolerance Interval to
Establish Design Curve [AISI 316
data; Saltsman and Halford (38)]

$\Delta \varepsilon = .4752N^{-.5845}$

Least Squares Line

Tolerance Limit, N_D
$\alpha = 0.01$
with confidence Y = 95%

Strain Range, $\Delta \varepsilon$

Y = log N

Cycles to Failure, N

reliability analysis or probabilistic design approach. Methods for analyzing stress-life, S-N (or strain-life, $\Delta\varepsilon$-N) data are discussed as follows.

Consider a constant-amplitude strain-controlled fatigue test in which data pairs ($\Delta\varepsilon$, N_i) i = 1,...n are collected where N_i denotes cycles to failure associated with strain range $\Delta\varepsilon_i$, n is the sample size, $\Delta\varepsilon$ is the independent (or controlled variable, and N is the dependent variable. Data from the tests are plotted on log-log paper as shown in Fig. 5.

Because that data show a linear trend, it is reasonable to suggest that material behavior might be suitably represented by a straight line through the center of the data. Such a line would have the form

$$\Delta\varepsilon = CN^\zeta \tag{4}$$

where C and ζ are empirical constants.

Methods of basic linear model analysis are typically used to analyze such fatigue data. Consider a log transformation of variables and let

$$Y = \log N, \qquad X = \log (\Delta\varepsilon). \tag{5}$$

Thus X is the independent variable, Y is the dependent variable.

Clearly there is no functional relationship between Y and X, but there does seem to exist some kind of relation. It will be assumed that the set of data is a random sample from the following model

$$Y(x) = Y_o(x) + \delta \tag{6}$$

in which δ is a normally distributed random variable with mean equal

to zero and standard deviation equal to σ, and

$$Y_0(x) = a + bx \qquad (7)$$

where a and b are constants. Thus for specified values of X, Y is a normally distributed random variable having mean and standard deviation

$$E(Y|X) = Y_0 = a + bx \qquad (8)$$

$$\sigma(Y|X) = \sigma \qquad (9)$$

Note the assumption that σ is a constant, not a function of X. The "scatter band" of the data on log-log paper would be constant. Such data is said to be "homoscedastic."

The line $Y_0 = a + bx$, being the mean of Y, will pass through the "center" of the data. Moreover σ is a measure of the dispersion of Y for a given X. Therefore, in the linear model, a, b, and σ provide a description of the trend and dispersion of the data.

Because Y is normally distributed, N (given $\Delta\epsilon$) will be log-normal. Thus the median of N, denoted as $\tilde{N}$, is given by $Y_0 = \log \tilde{N}$. In terms of the original coordinates, the Y_0 line can be written as

$$\Delta\epsilon = C\tilde{N}^\zeta \qquad (10)$$

in which it follows from the above definitions that

$$a = -\frac{1}{\zeta}\log C, \quad b = 1/\zeta \qquad (11)$$

The parameters a and b (and thus C and ζ) and σ are not known in advance and must be estimated from the data $(\Delta\epsilon_i, N_i)$, $i = 1,\ldots n$. Equation 5 is used to translate the data into (X_i, Y_i) $i = 1,\ldots n$.

16

Using the method of least squares, a, b, and σ are estimated by $\hat{a}$, $\hat{b}$ and $\hat{s}$ respectively (25),

$$\hat{b} = \sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})/\sum_{i=1}^{n}(X_i - \bar{X})^2 \qquad (12)$$

$$\hat{a} = \bar{Y} - \hat{b}\bar{X} \qquad (13)$$

$$s^2 = \frac{1}{n-2}\sum_{i=1}^{n}[Y_i - (a + \hat{b}X_i)]^2 \qquad (14)$$

where $\bar{X}$ and $\bar{Y}$ are the sample means of X and Y respectively. Because each Y_i is a random variable, the estimates $\hat{a}$, $\hat{b}$ and s are also random variables. The "best fit" line

$$\hat{Y} = \hat{a} + \hat{b}X \qquad (15)$$

is called the <u>least squares line</u>. $\hat{Y}$ is the estimate of Y_o, the mean of Y given X.

As an example, the low cycle fatigue data for AISI 316 as presented by Saltsman and Halford (30), and shown Fig. 5 are analyzed herein. This set of data is given in Table II along with associated statistics.

Equations 12, 13, and 14 can be used to calculate $\hat{a}$, $\hat{b}$, and s, the estimates of a, b and σ, respectively.

$$a = -0.6530 \qquad b = -1.7108 \qquad s = 0.1427 \qquad (16)$$

Least squares estimators $\hat{\zeta}$ and $\hat{C}$ are obtained from Eq. 11

$$\hat{\zeta} = 1/\hat{b} = -0.5845$$

$$\hat{C} = 10^{-\hat{a}/\hat{b}} = .4152 \qquad (17)$$

17

Table II

Statistical Analysis of AISI 316 PP Data

[after Saltsman and Halford (30)]

Sample Size n=7

Strain Range $(\Delta \varepsilon)_i$	Cycles to Failure N_i	$X_i = \log(\Delta \varepsilon)_i$	$Y_i = \log N_i$
.00424	1700	-2.373	3.230
.00105	35600	-2.979	4.551
.03508	120	-1.455	2.079
.03496	68	-1.564	1.832
.00466	2333	-2.332	3.368
.02066	116	-1.685	2.064
.02360	146	-1.627	2.164

Sample Mean of X

$\overline{X} = -1.986$

Estimate of a

$\hat{a} = -.6530$

Sample Mean of Y

$\overline{Y} = 2.746$

Estimate of b

$\hat{b} = -1.711$

Other Statistics

$s = 0.1427$

$\sum x_i^2 = 29.663$

$s_X^2 = \frac{1}{n} \sum x_i^2 - (\overline{X})^2 = .2934$

$s_X = 0.5416$

The least squares line, $\Delta\varepsilon = \hat{C}\tilde{N}^{\hat{\zeta}}$ is plotted on Fig. 5 where $\hat{N}$ denotes the estimate of the median $\tilde{N}$.

Note that (a) the conditional random variable Y given X is normal and (b) the least squares line is the estimate of the mean $Y|X$. Therefore, it follows that (a) $N|\Delta\varepsilon$ is lognormal and (b) the least squares line, N, is the estimate of the median of $N|\Delta\varepsilon$.

An Alternate Form. The form of the $\Delta\varepsilon$-N relationship as given above is commonly used. However, for probabilistic design purposes, it may be more convenient to express the "Y_o line" (Eq. 10) as

$$N = A(\Delta\varepsilon)^m \tag{18}$$

Comparing Eqs. 12 and 20 the constants A and m in terms of C and ζ are given by

$$A = C^{-1/\zeta} \qquad m = 1/\zeta \tag{19}$$

For the above example, the least squares estimators are $K = .2223$ and $\hat{m} = -1.711$.

ESTABLISHING A DESIGN CURVE

A number of schemes have been used to establish a design curve for a given set of data as illustrated in Fig. 2. It is reasonable to define a design curve as a curve below which one expects the occurrence of a failure with probability α. However, the oldest (and easiest) method is just to draw the design curve to the left of all points with a "little space" between the points and the curve, e.g. see Fig. 1. (In the 1969 version of the ASME Boiler and Pressure Vessel Code, a mediam "Langer curve" is established using a least squares fit (17). Curves removed from this median, a factor of 2 on strain and 20 on life are drawn.

The design curve is a lower-bound envelope. In the linear case, it is common practice to define a design curve by drawing a line, 2 or 3 standard deviations to the left of the least squares line, and parallel to it. Such an approach does not produce an unreasonable design curve, but it does not consistently define a given reliability level, because the approach fails to recognize that a, b, and s, being estimators, are all random variables. In other words, it is incorrect (and non-conservative) to say that in general 0.135% of the failure points would be expected to lie below the median minus 2 standard deviations (on a log basis). When the distribution of the estimators are considered, that percentage is higher.

Two schemes for constructing a design curve, giving consideration to the distribution of the estimates of the least squares parameters are (1) the tolerance interval, and (2) the equivalent prediction interval.

The tolerance limit method (26) for defining the design curve relies implicitly upon the assumption of "linear failure trajectories." The concept of a failure trajectory is very useful in analyzing fatigue data for design purposes. In Fig. 6, the star represents a specimen which has failed. The dashed lines (failure trajectories) are drawn through the point according to some predefined rule. Typically the failure trajectories are parallel to the predetermined median curve. It is assumed that the failure trajectory defines the cycles to failure of that specimen if it had been tested at a different stress level. For example, a specimen which is "weak" at a high stress level

STRAIN, Δε
(or STRESS)

MEDIAN of N|Δε
(also MEDIAN of Δε|N)

PDF of N|Δε

α

Typical Failures

PDF of Δε|N

DESIGN CURVE

FAILURE TRAJECTORIES
(The star represents the failure
point, but had the specimen been
tested at another level, it would
have failed on the dashed line)

α

CYCLES TO FAILURE, N
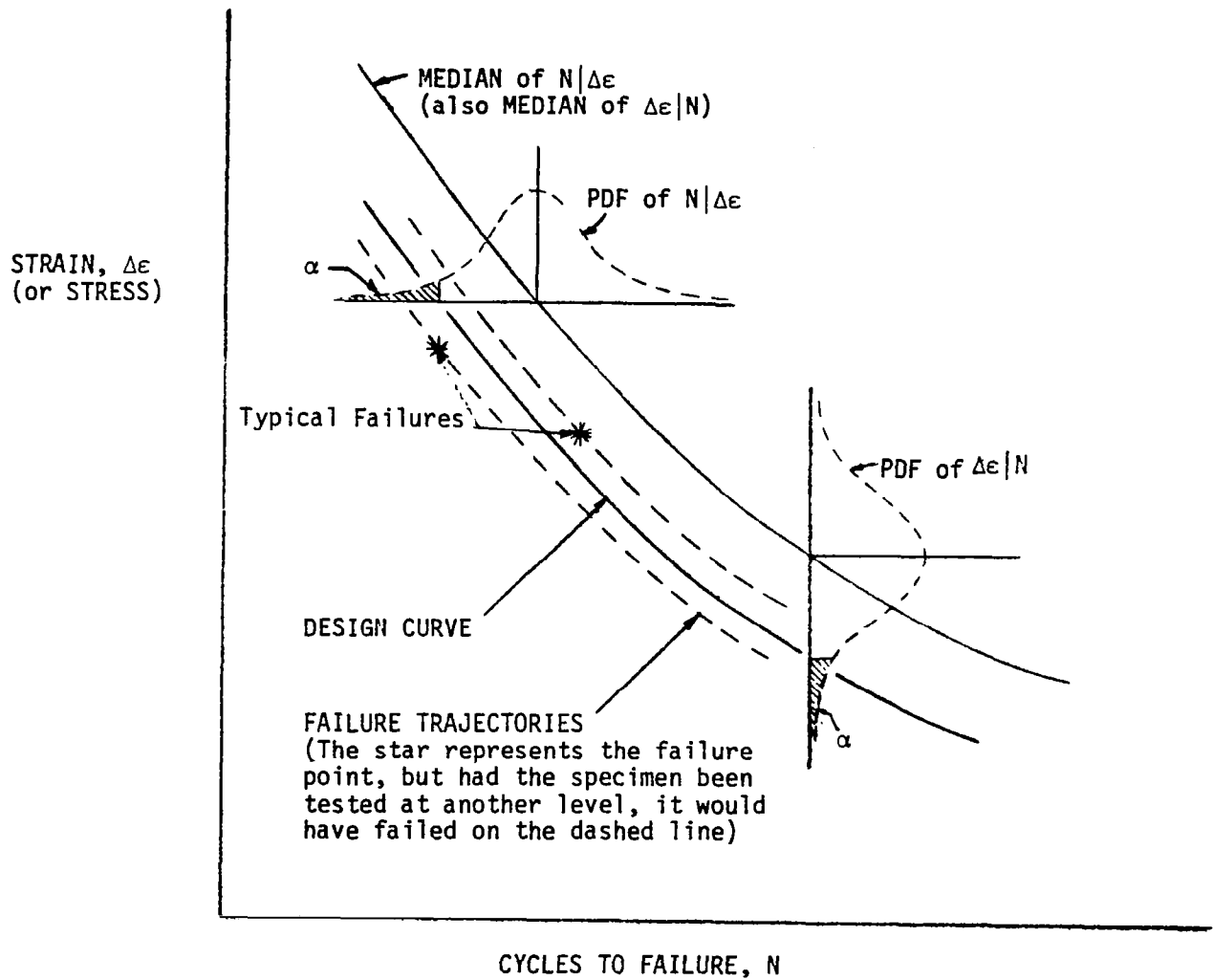
Fig. 6

Failure Trajectories and the Assumed Relationship Between Distribution of
N|Δε and Δε|N

21

would also be weak at a low stress level. Experimentalists generally recognize that a strong, low ductility specimen will have a relatively "long life" in high cycle fatigue and "short life" in low cycle fatigue. Therefore caution should be exercised in applying the failure trajectory model.

The assumption that failure trajectories describe material behavior leads to useful statistical descriptions that otherwise would not be possible. For example, we can construct the distribution of $\Delta\varepsilon$ given N as from N given $\Delta\varepsilon$ as implied in Figure 6. The design curve thus defines the lower $\alpha\%$ of failures in both the horizontal and vertical directions. This model, not valid without the failure trajectory assumption, is convenient for reliability analyses.

The tolerance limit at any stress level will lie to the left of the least squares line Y, a distance of $k_{\alpha,\gamma}s$. Where $K_{\alpha,\gamma}$ is a tolerance factor (tabulated, e.g. in Ref. 26) is the population fraction, and γ is the confidence level. Thus, the design curve is

$$\log N_D = \hat{Y} - K_{\alpha,\gamma}s \tag{20}$$

The low cycle fatigue data of Saltsman and Halford (30) are used to provide an example of a design curve based on the tolerance interval. The data are shown in Fig. 5. Assume that the decision has been made to define the design curve as the line above which no more than $\alpha = 1\%$ of the population is expected to fall with confidence $\gamma = 95\%$; n = 7. From tables (26), $K_{\alpha,\gamma} = 4.64$. The design curve is

$$\log N_D = \hat{Y} - K_{\alpha,\gamma}s = \hat{Y} - 4.64(0.1427) = \hat{Y} - 0.662 \tag{21}$$

This line is drawn in Fig. 5. Note that the values of α = .01 and γ = .95 of this example are those levels used to define the A-values

22

in MIL-HDBK-5 (24).

The equivalent prediction interval (EPI) concept is described by Wirsching and Hseih (38) and summarized in the following discussions. Define an equivalent constant standard deviation of Y as

$$\sigma_0 = s \ g(n,\alpha) \tag{22}$$

where

$$g(n,\alpha) = \exp[A(\alpha)\{\ln n\}^{-B(\alpha)}] \tag{23}$$

and

$$A(\alpha) = 1.56[\frac{1}{2}\ln(\frac{2-\alpha}{\alpha})]^{1.12} \tag{24}$$

$$B(\alpha) = 3.32 - 1.7\alpha$$

$$6 \leq n \leq 50; \quad 0.01 \leq \alpha \leq 0.15$$

$g(n,\alpha) \geq 1$ is in essence, an adjustment factor to s to account for the fact that there is uncertainty in the estimates of a and b and s. Basically the idea is to use the linear model, but with an expanded value of s (the value of σ_0 defined below) so that the N_D curve is shifted to the left to match the N_α curve.

The model, suggested by the above discussion, is as follows:

1. Let $m = \hat{b}$ be a constant

2. Assume that all of the uncertainty due to scatter in the data is accounted for in $\hat{a}$ (and therefore K) by considering the y intercept as a random variable.

3. Therefore, let the empirical relationship be

$$Y = a_0 + \hat{b}x \tag{25}$$

where a_0 has a normal distribution with mean $\hat{a}$ and standard deviation σ_0.

23

The consequences of such a model are

1. $Y|X$ has a normal distribution. (Thus N given $\Delta\epsilon$ has a lognormal distribution.)

2. The mean value of $Y|X$ is $\hat{a} + \hat{b}x$. (Thus the estimate of the median of N is $\hat{N} = \hat{A}(\Delta\epsilon)^m$.)

3. The standard deviation of $Y|X$ is σ_0 (and is not a function of X).

4. $\hat{a}_0 = \log A$ is normal, and A is lognormal. The median $\tilde{A}$ and coefficient of variation C_A of A can be obtained from the lognormal (base 10) forms

$$\hat{A} = 10^{\hat{a}} \tag{26}$$

$$C_A = \sqrt{10^{(\sigma_0^2/.434)} - 1} \tag{27}$$

EXAMPLE Given the fatigue data (n = 7) as illustrated in Fig. 7, it is required to define a design S-N line which is estimated to be on the safe side of 99% of the data. This line is to be the α = .01 EPI. The basic data is summarized in Table II.

Using Eq. 25 with n = 7, α = 0.01

$$A(\alpha) = 4.64 \quad B(\alpha) = 3.30 \quad g(n,\alpha) = 1.67$$

Thus, the equivalent standard deviation is,

$$\sigma_0 = g \cdot s$$

$$= (1.67)(0.143) = 0.239 \tag{28}$$

Design Curve The 1% EPI is given as N^*, where

$$\log N^* = \hat{Y} - z_\alpha \sigma_0$$

Fig. 7 Comparison of the Lower 1% Prediction
Interval with the Equivalent
Prediction Interval [AISI 316 data,
Saltsman and Halford (38)]

1% N_D

Least Squares Line, $\hat{Y}$

$\hat{N} = .2223 \, (\Delta\varepsilon)^{-1.711}$

Lower 1%
Prediction Interval, N_α

N^*
(1% EPI)

Strain Range, $\Delta\varepsilon$

Cycles to Failure, N

25
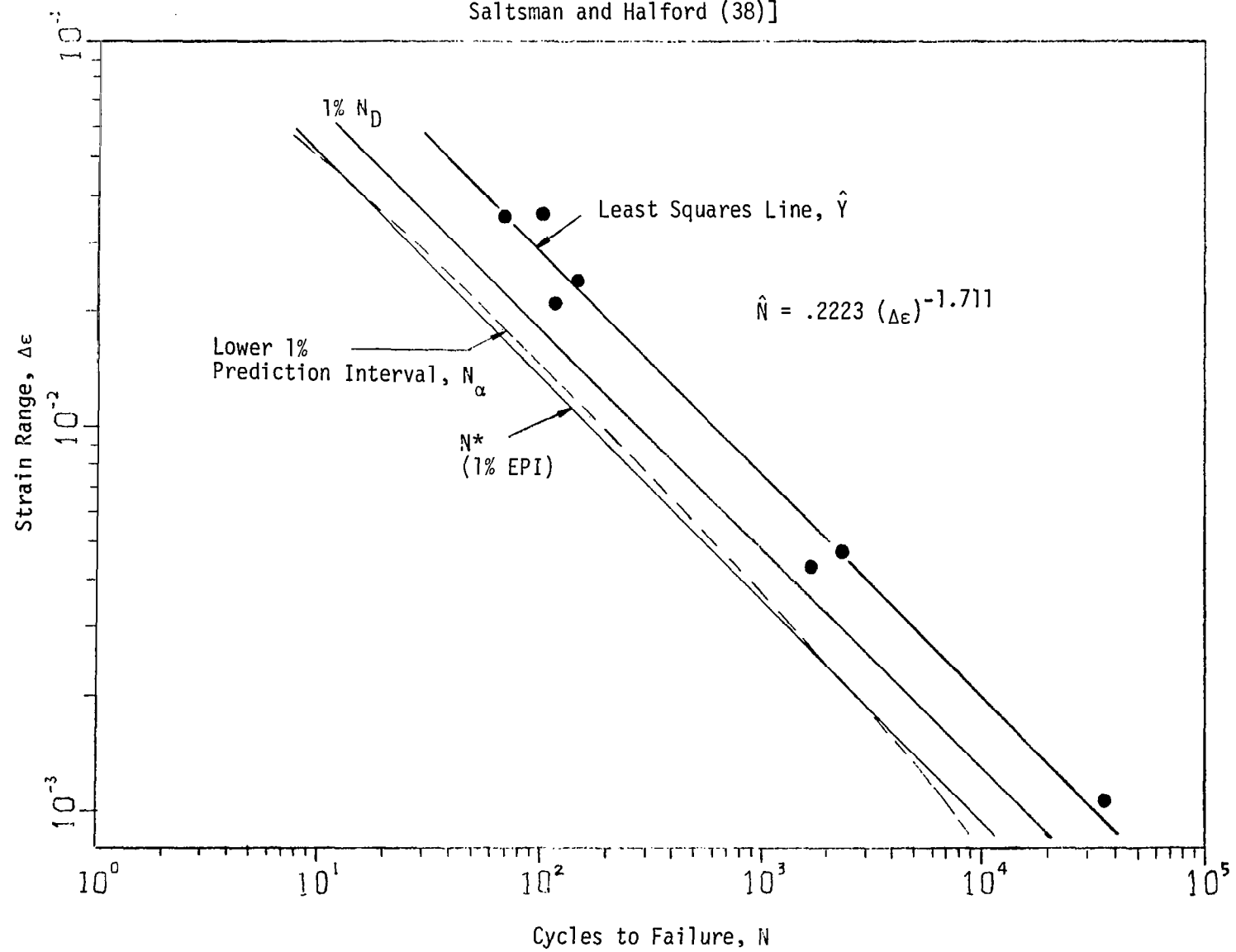
This EPI could be used as the design curve in the conventional approach. The EPI is shown in Fig. 7 along with the prediction interval N_α which the EPI approximates.

Figure 7 suggests that the EPI is a reasonable approximation to N_α the prediction interval. As the sample size becomes larger, N_α becomes flatter and the EPI becomes an even better approximation (38).

Probabilistic Format The data will be analyzed in a format which is convenient for probabilistic design procedures.

The fatigue equation will contain m and A. Using the method described above, m is a constant and equal to

$$\hat{m} = \hat{b} = -1.711$$

A will be lognormal with a median value of (Eq. 26)

$$\tilde{A} = 10^{-.6530} = .222$$

A will have a coefficient of variation of (Eq. 27 and 28) based on $\alpha = 0.01$

$$C_A = \sqrt{10^{(.239)^2/.434}} - 1 = 0.595$$

At a given strain (or stress) level, the coefficient of variation of cycle life N is equal to C_A

$$C_N = C_A \tag{29}$$

As a subjective comment, for sample sizes, $n \leq 7$, this author has found that fatigue design curves seem to be dominated by the uncertainties in $\hat{a}$, $\hat{b}$, and s. When included in the analysis, they produce what seem to be "unreasonably" conservative results. On the other hand, as

26

the sample size n becomes large, $\hat{a} \to a$, $\hat{b} \to b$ and $\hat{Y}_0 \to Y_0$, and s approaches σ, and for all practical purposes $\hat{a}$, $\hat{b}$ and s could be treated as constants for large n (typically the assumption would be reasonable for n > 50). However, because of the expense associated with fatigue testing, sample sizes will generally be small. In summary it is necessary to give full consideration to statistical distributions of these estimators when $n \leq 50$.

The issue of sample size requirements is addressed in ASTM Specification E-739 recommendations (4)

Number of Test Points Required for S-N Relationship

Type of Test	Number of Specimens
Exploratory R and D	6 to 12
Design Allowables	12 to 24
Reliability Data	12 to 24

This author agrees that these values are quite reasonable, but notes that reliability information can be provided for sample sizes, $n \geq 6$, using the EPI as described above.

SCATTER IN FATIGUE DATA: SOME RESULTS

Relatively large scatter is observed in cycles to failure data. At low cycle lives, values of C_N in the range from 0.20 to 0.40 is common for most metallic alloys. At higher lives (lower stress levels) C_N has been observed to exceed 1.00. By contrast, coefficients of variation of yield and ultimate strengths are usually less than 10%.

In an exhaustive study of fatigue data of steel and titanium alloys for aircraft applications, Whittaker has provided the summary statistics as presented in Tables III through VII (35). He used the two parameter Weibull (Eq. 3 with ξ = 0) to describe N. The shape parameter in these tables is , which is a function of the coefficient of variations;

$$C_N = n^{-0.926}.$$

Because of the large sample size studies, the significance of the Whittaker data is that some confidence can be placed in his values of COV's for different materials and conditions. Whittaker's own recommendations for representative values of coefficients of variations are

	COV (%)
Steel, TS 240 ksi	36
Steel, TS 240 ksi	48
Aluminum Alloys	22
Titanium Alloys	36

Summary fatigue data on welded joints of various structural detail collected by Ang and Munse (2) is presented in Table VIII. The average value of C_N of this data is 52%. Data summaries of welded tubular joint fatigue are provided in Table IX (41); here it is noted that the C_N is significantly higher.

Results of the RQC-100 round robin fatigue tests coordinated by ASTM are presented in Table X. The general strain-life model was used. COV's of both strain given life and life given strain are recorded. The parameters of the strain life relationship were computed by a least squares analysis using vertical deviations. The COV's of strain given life were computed (Eq. 29) using actual estimators, s; no expansion of s was made using the EPI. The COV's of life given strain were computed using Eq. 29 with s_N = s/(abs. value of slope), the s_N being the estimated log standard deviation of life given strain.

These values are very typical of those observed for other materials. What is of particular interest here is that these data summaries provide some indication of lab-to-lab variability although it is impossible to isolate this effect due to the small sample sizes.

28

Table III. —*Results of Analyses Determining the Typical Shape Parameters for Fatigue Performance of Titanium Alloys*

Data description	All data			Data below $4.(10)^5$ cycles		
	Number of groups	$C_N(\%)^*$	α	Number of groups	$C_N(\%)^*$	η
Titanium 6Al-4V	541	38	2.81	487	35	3.15
Titanium 8Al-1Mo-1V	586	35	3.07	529	32	3.44
Monolithic notched	637	37	2.93	581	33	3.33
Structures	Negligible	—	—	Negligible	—	—
Structural simulators	488	37	2.94	433	34	3.25
Room temperature	825	35	3.06	744	32	3.40
Elevated temperature	279	42	2.53	249	37	2.91
Low temperature	Negligible	—	—	Negligible	—	—
Constant amplitude	1056	38	2.83	945	34	3.17
Variable amplitude	71	16	7.19	71	16	7.19
All data	1127	37	2.93	1016	33	3.30
10^2-10^3 cycles	Negligible	—	—	Negligible	—	—
10^3-10^4 cycles	110	34	3.25	110	34	3.25
10^4-$6\cdot(10)^4$ cycles	396	28	4.02	396	28	4.02
$6\cdot(10)^4$-$4\cdot(10)^5$ cycles	429	41	2.62	429	41	2.62
$> 4\cdot(10)^5$ cycles	111	70	1.47	—	—	—

Ref. Whittaker (35)

$^*C_N = \eta^{-0.926}$

Table IV.—*Results of Analyses Determining the Typical Shape Parameters for Fatigue Performance of High-Strength Steels*

Data description	All data		
	Number of groups	$c_N(\%)$ *	η
Alloy steels	168	37	2.95
Intermediate alloys	111	46	2.31
18% Ni maraging steels	113	51	2.06
Stainless steels	314	33	3.29
Austenitic stainless steel	48	23	4.83
Air melted	44	33	3.27
Vacuum melted	94	42	2.52
0-100 ksi	43	32	3.38
101-160 ksi	43	22	5.18
161-200 ksi	131	27	4.17
201-240 ksi	285	34	3.16
241-280 ksi	132	48	2.20
281-320 ksi	83	49	2.14
321-360 ksi	—	—	—
Monolithic notched	488	40	2.72
Structures	—	—	—
Structure simulators	282	35	3.11
Room temperature	613	38	2.82
Elevated temperature	115	42	2.55
Low temperature	Negligible	—	—
Constant amplitude	770	38	2.84
Variable amplitude	Negligible	—	—
All data	770	38	2.84
10^2-10^3 cycles	143	18	6.37
10^3-10^4 cycles	127	29	3.75
10^4-$6 \cdot (10)^4$ cycles	265	42	2.58
$6 \cdot (10)^4$-$4 \cdot (10)^5$ cycles	189	48	2.21
$>4 \cdot (10)^5$ cycles	46	61	1.71

Ref. Whittaker (35)

*$c_N = \eta^{-0.926}$

Table V. — *Typical Shape Parameters for Fatigue Performance of High-Strength Steels Varying with Strength Ranges (i), (ii), and (iii)*

Data description		All data		
		Number of groups	$c_N(\%)^*$	η
(i)	**161-200 ksi**	131	27	4.17
	Alloy steels	37	33	3.33
	Intermediate alloys	Negligible	–	–
	18% Ni maraging steels	—	–	–
	Stainless steels	91	22	5.03
	Austenitic stainless steel	–	–	–
(ii)	**201-240 ksi**	285	34	3.16
	Alloy steels	Negligible	–	–
	Intermediate alloys	24	14	8.47
	18% Ni maraging steels	–	–	–
	Stainless steels	204	37	2.90
	Austenitic stainless steel	48	23	4.83
(iii)	**241-280 ksi**	132	48	2.20
	Alloy steels	47	37	2.93
	Intermediate alloys	75	54	1.96
	18% Ni maraging steels	Negligible	–	–
	Stainless steels	–	–	–
	Austenitic stainless steel	–	–	–

Ref. Whittaker (35)

$$^*c_N = \eta^{-0.926}$$

Table VI.—*Typical Shape Parameters for Fatigue Performance of Stainless Steels Varying with Strength (i) and Life (ii)*

Data description		Stainless steels		
		Number of groups	$c_N(\%)^*$	η
	All data	314	33	3.29
(i)	0-100 ksi	Negligible	–	–
	101-160 ksi	–	–	–
	161-200 ksi	91	22	5.03
	201-240 ksi	204	37	2.90
	241-280 ksi	–	–	–
	281-320 ksi	Negligible	–	–
	321-360 ksi	–	–	–
(ii)	10^2-10^3 cycles	66	18	6.49
	10^3-10^4 cycles	49	28	3.94
	10^4-$6 \cdot (10)^4$ cycles	109	39	2.77
	$6 \cdot (10)^4$-$4 \cdot (10)^5$ cycles	66	32	3.38
	$> 4 \cdot (10)^5$ cycles	24	58	1.81

Table VII.—*Typical Shape Parameters for Fatigue Performance of High-Strength Steels with Strengths Equal to or Less Than 240 ksi (i) and Greater Than 240 ksi (ii)*

| Data description		All data		
		Number of groups	$c_N(\%)^*$	η
(i)	Strength $\leqslant$ 240 ksi	502	31	3.51
	10^2-10^3 cycles	98	18	6.29
	10^3-10^4 cycles	72	26	4.24
	10^4-$6 \cdot (10)^4$ cycles	157	34	3.21
	$6 \cdot (10)^4$-$4 \cdot (10)^5$ cycles	135	32	3.37
	$> 4 \cdot (10)^5$ cycles	40	54	1.94
(ii)	Strength $>$ 240 ksi	215	49	2.17
	10^2-10^3 cycles	45	18	6.49
	10^3-10^4 cycles	43	34	3.24
	10^4-$6 \cdot (10)^4$ cycles	82	55	1.91
	$6 \cdot (10)^4$-$4 \cdot (10)^5$ cycles	39	78	1.31
	$> 4 \cdot (10)^5$ cycles	Negligible	–	–

Ref. Whittaker (35)

$^*c_N = \eta^{-0.926}$

32

Table VIII

FATIGUE DATA FOR VARIOUS DETAIL SPECIFIED IN AISC CODE

Data Provided by A.H.-S. Ang & S.H. Munse

U. of Illinois (2)

33
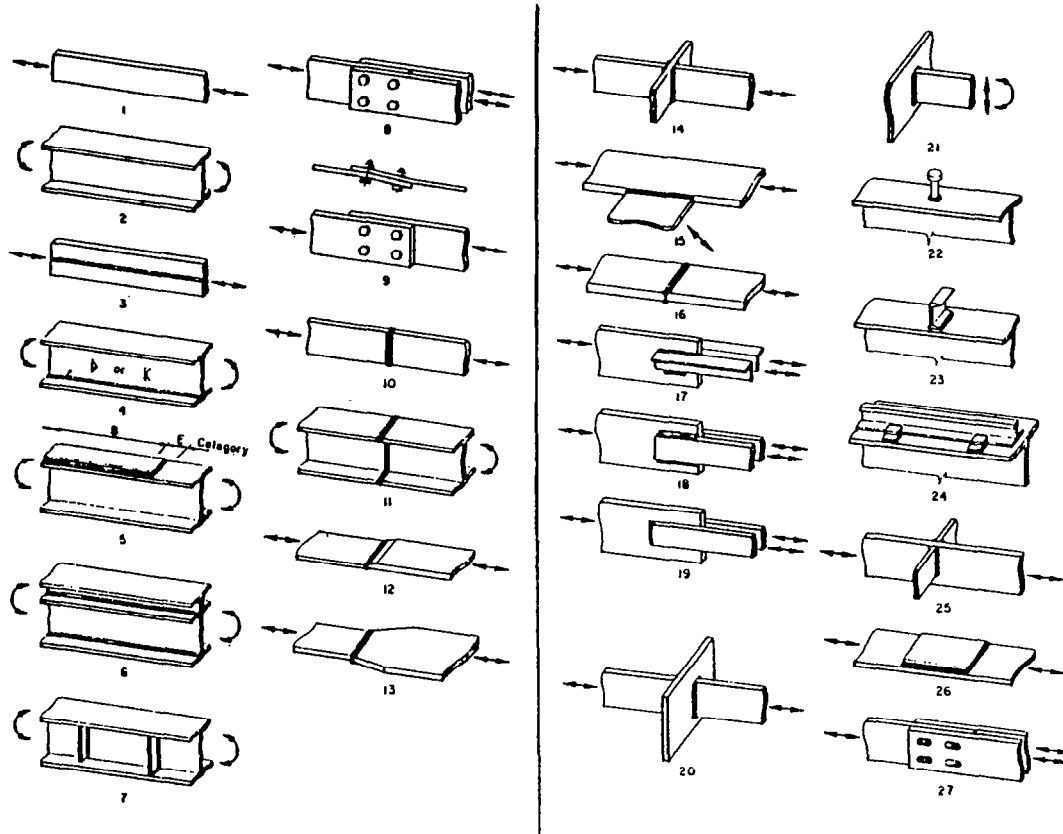
Detail No.	m	a	C_N
1	-9.7780	21.51	0.56
2	-9.7780	21.51	0.56
3	-2.7500	9.86	0.35
4	-2.7500	9.86	0.35
5	-2.7500	8.60	0.21
6	-2.7500	9.86	0.35
7	-2.7400	9.38	0.49
8	-7.6180	18.32	0.86
9	-7.4270	16.16	0.85
10	-2.8800	9.57	0.49
11	-3.8430	11.25	0.42
12	-2.8950	9.71	0.55
13	-3.2720	10.70	0.41
14	-2.7700	9.33	0.46
15	-3.0300	9.37	0.40
16	-3.7210	10.83	0.67
17	-2.6700	8.91	0.47
18	-2.6700	8.91	0.47
19	-5.3070	13.43	0.83
20	-2.7700	9.33	0.46
21	-6.6810	15.86	0.80
22	-2.7140	9.49	0.30
23	-3.2460	10.04	0.13
24	-3.2460	10.04	0.13
25	-2.7400	9.38	0.49
26	-3.7420	10.38	0.39
27	-4.6520	11.27	0.58

Fatigue Curve

$$NS^m = A$$

where, $A = 10^a$

C_N = coefficient of variation of cycles to failure

Table IX

A Statistical Summary of Some Tubular Joint Fatigue Data

(For detail on these data sets, see Ref 39 and 41)

Investigator	Comment	Slope of S-N Curve (m)	Median of K ($\bar{K}$)	COV of K (C_K)
Marshall; AWS-X Data	amalgamated data	3.41	4.75E10	1.82
Marshall; AWS-X Data (elastic range)[a]	amalgamated data	4.42	1.55E12	1.36
Marshall:	corrosion data	4.15	1.41E12	1.97
Toprac and Louis	K-joints	3.64	9.78E10	0.65
Kurobane and Konomi	K-joint	3.78	1.26E12	1.48
Maeda et al.	k-joint	2.94	3.09E10	2.14
Kurobane et al.	T-joint	3.20	1.92E12	1.70
Martin and McGregor	T-joint	4.40	1.03E15	1.03
Toprac and Louis	T-joint	3.64	7.32E11	0.82
Dijkstra and de Back	T-joint 168 mm dia chord	4.12	2.38E12	1.32
Dijkstra and de Back	T and K, 457 mm dia chord	2.79	1.11E10	0.98
Dijkstra and de Back	T and K, 914 mm dia chord	2.98	8.51E9	0.81
Unpublished T-joint data[b]	Sample size n=24	3.82	4.21E10	0.50
Unpublished K-joint data[b]	Sample size n=33	2.89	2.77E10	0.80
Unpublished T and K joint data [a,b]	Sample size n=57	3.00	5.25E10	0.73
Unpublished T and K joint data[a]	An "improved" version of previous set	3.22	1.29E11	1.25
API RP2A (Commentary on Fatigue)[a]	Large scale K-joint	4.38	4.60E12	0.73
UK-DOE Guidance Notes[a]	"T curve" for tubular joints	3.00	1.46E10	0.67

Notes: (a) Considered to be a reasonable data set for a fatigue reliability analysis
(b) This collection of data represents a screened amalgamation of points from various investigators. The collection was provided by a member of the API Technical Advisory Committee on this project. This amalgamation has not been published and has not received approval for release by the sponsoring company. The results have been made available for research purposes only.

TABLE X

THE GENERAL STRAIN-LIFE RELATIONSHIP FOR RQC-100 STEEL:
STATISTICAL SUMMARIES FROM VARIOUS LABORATORIES
(ASTM Round Robin Exercise)

LAB	N	E	σ'_f(ksi)	b	ϵ'_f	c	$COV(\epsilon_p)$ (%)	$COV(\epsilon_e)$ (%)	$COV(N_p)$ (%)	$COV(N_e)$ (%)
1	17	29500	159.85	-0.083	2.322	-.778	39.2	4.4	51.5	56.6
2	8	30325	153.0	-0.077	1.398	-.704	19.8	2.1	28.4	27.4
3	4	29964	257.09	-.134	.563	-.578	9.9	20.7	17.3	306.4
4	8	31995	118.75	-.053	.392	-.604	12.3	2.5	20.4	50.7
5	5	30560	131.04	-0.058	.643	-.649	21.7	5.3	33.9	114.2
6	8	29500	175.9	-.081	3.082	-.787	24.1	2.8	30.8	35.1
7	9	26167	158.39	-.081	1.006	-.679	16.7	3.7	24.8	48.0
8	6	29500	125.44	-.052	.633	-.620	12.3	1.6	19.9	31.8
9	8	29500	164.82	-.080	.990	-.707	8.5	2.3	12.1	29.4
10	8	30000	132.06	-.061	1.345	-.704	13.0	4.1	18.5	76.6
11	8	29350	147.66	-.066	1.533	-.693	9.9	4.4	14.4	74.3
12	7	29500	134.63	-.066	.788	-.684	16.9	4.4	25.0	74.3
13	8	31270	141.46	-.062	.796	-.659	12.3	2.1	18.7	34.4
14	6	29500	153.46	-.084	.660	-.677	4.4	2.8	6.5	33.8
15	10	29300	171.54	-.083	.964	-.651	12.3	9.2	18.9	155.8
16	8	30450	227.78	-.111	.491	-.610	23.6	11.3	39.6	134.6
17	10	30770	142.81	-.071	.517	-.639	13.9	4.8	21.9	76.9
18	10	29000	141.75	-.068	2.171	-.762	26.2	2.1	34.8	31.2
All (a)	170	29680	148.65	-.071	1.23	-.718	53.3(c)	9.2(c)	79.0(c)	209.4(c)
All (b)	148	29680	148.64	-.071	.861	-.667	29.6(c)	8.8(c)	45.7(c)	189.0(c)

(a) Amalgamation of all data (additional data not included in data sets 1 through 18)
(b) Screened amalgamation of data (data sets 1 through 18)
(c) These values may be inappropriate for design purposes; they are unreasonably high because
they reflect both uncertainties in material behavior and lab to lab variations.

EXAMPLE High temperature fatigue data (sample size, n = 44) for Waspaloy B (a nickel base superalloy) is analyzed using the strain-life relationship. The plastic, elastic, and total strain-life least squares lines are given in Figs. 8, 9, and 10 respectively. The least squares lines were fitted using vertical deviations (X and Y reversed in Eqns. 14, 15, and 16). This was for statistical convenience as described below. The fit could be performed either way, but there are significant differences as described by Boardman (4).

Also shown are the 1% EPI and the $\alpha = 0.99$, $\gamma = 0.95$ lower toler- ance intervals. Either of these curves could be used as a design curve.

These data points are very typical of strain-life data for a wide variety of metallic materials. The plastic and elastic data has a linear trend, and the data is homoscedastic (constant scatter band). Note that accurate measurements of small plastic strains are difficult and therefore some points below roughly ε_{pa} = 0.00025 which seem to violate the trend are questionable.

In a reliability format, all of the parameters of the strain life relationship should, in general, be treated as random variables. This complicated format can be simplified by assuming that E, b, and c are constants; E is known to have little variability. Basic uncertainty in material behavior is described by treating σ_f' and ε_f' as random variables, the median and coefficient of variation of which can be computed as de- scribed above in the EPI discussion.

But σ_f' and ε_f' may be dependent random variables. Three possibilities exist: (1) no, they are independent, (2) a specimen weak in high cycle

Fig. 8

An Example of Plastic Strain Life Curve

WASPALOY
(T = 1000 F)   n = 44

------ EPI (α = .01)

——— Tolerance Interval
(P = .01, γ = .95)

Least Squares Line
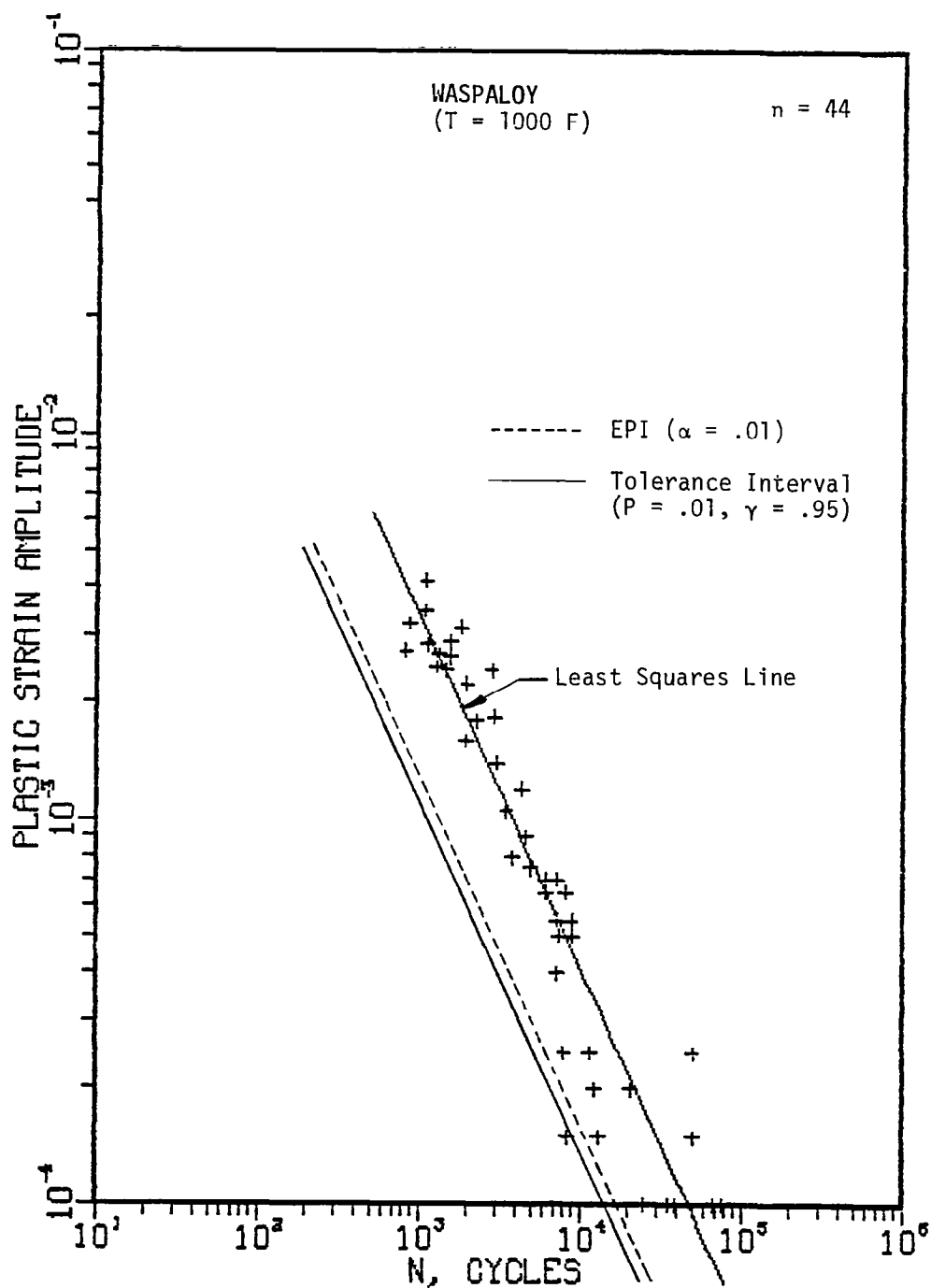
PLASTIC STRAIN AMPLITUDE

N, CYCLES
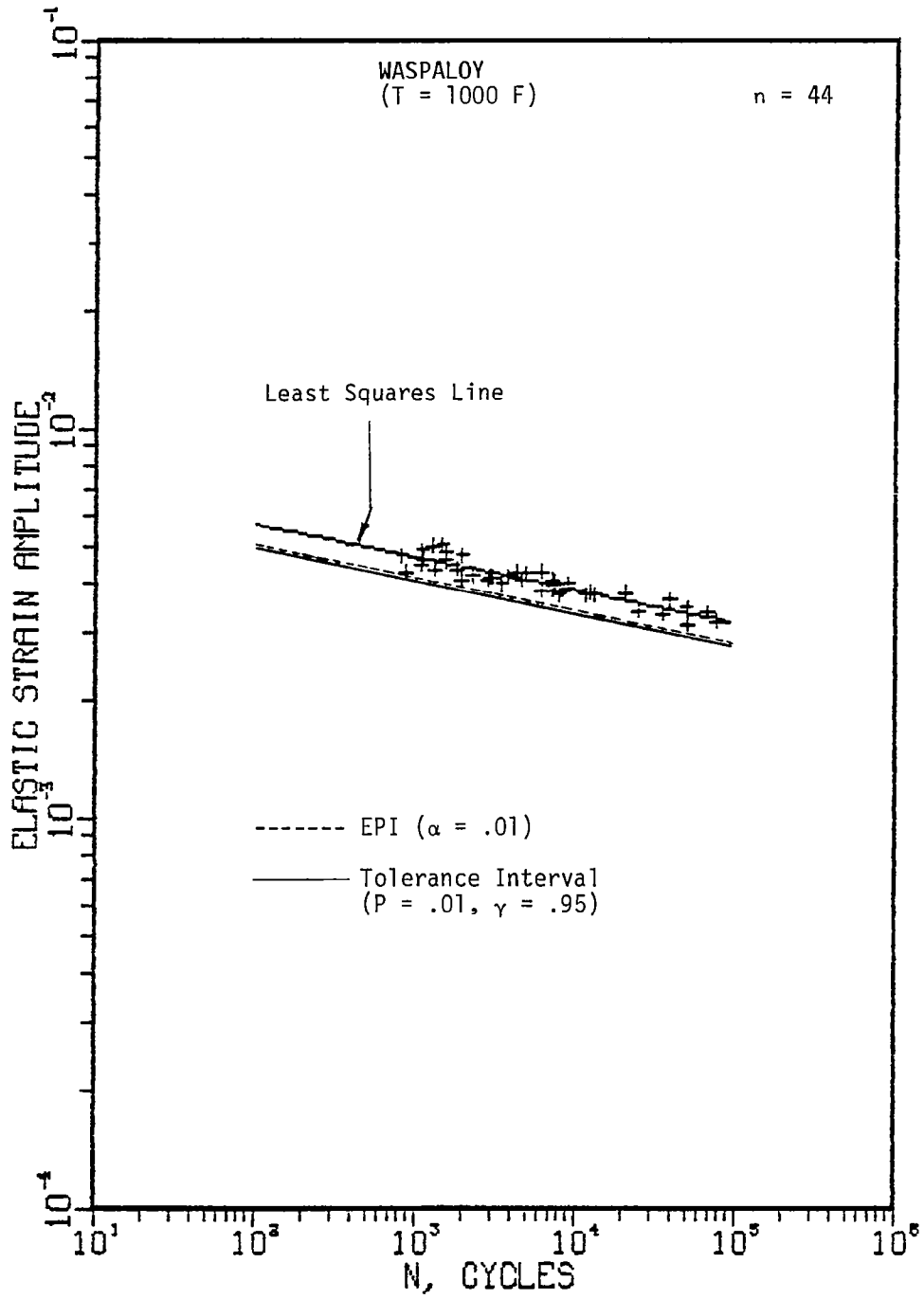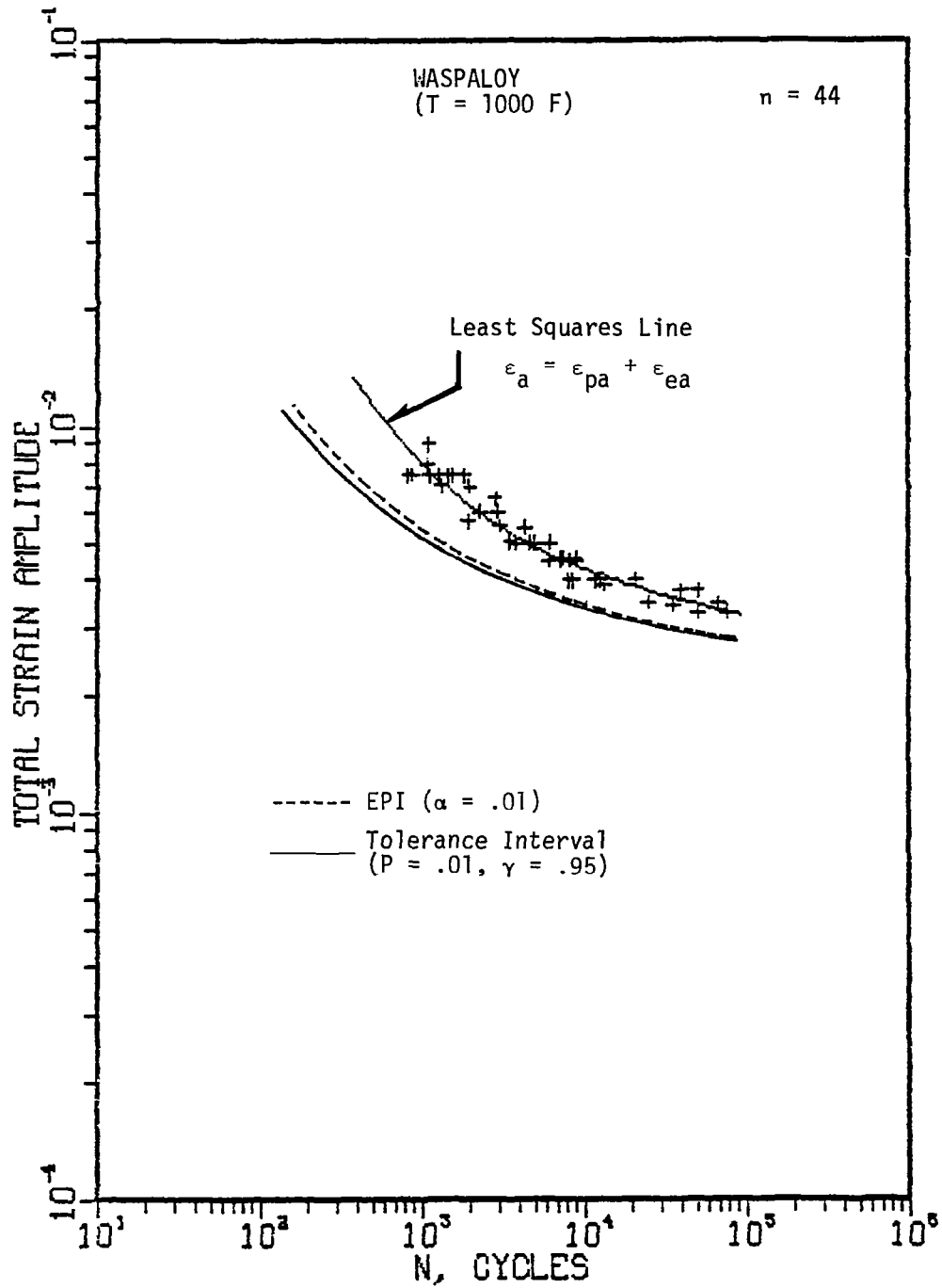
Fig. 9

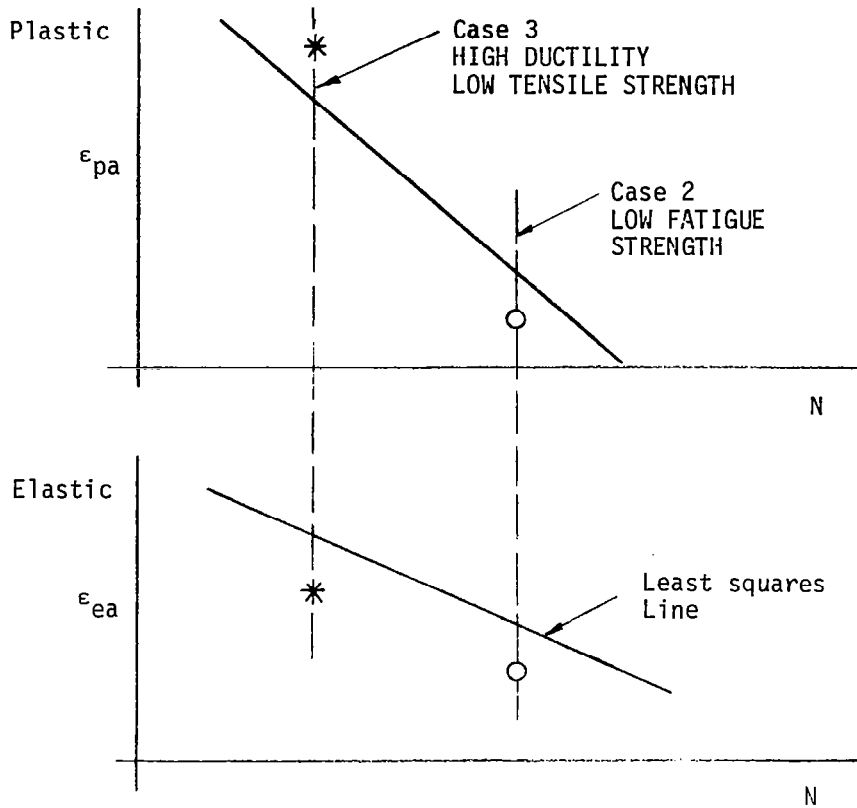An Example of an Elastic Strain Life Curve

Fig. 10

An Example of the Total Strain Life Curve

fatigue may also be weak in low cycle fatigue because of an inherent flaw or weakness, (3) a specimen may have relatively high ductility and low tensile strength, weak in high cycle fatigue and strong in low cycle fatigue (or vice versa). Failure points for the latter two cases in which σ_f' and ε_f' would be dependent are illustrated in Fig. 11a. This figure suggests that vertical deviations could be used to examine dependency.

Vertical deviations from the least squares line are defined in Fig. 11b. The plastic and elastic deviation for each specimen in the Waspaloy B data set is plotted in Fig. 12. For Case 2 behavior, the data would be concentrated in the 1st and 3rd quadrant, Case 3 in the 2nd and 4th quadrant. The relative uniform spread suggests that it is not unreasonable to assume that σ_f' and ε_f' are independent. In summary the fatigue strength (strain level given cycle life) is given by the general strain life relationship of Eq. 2 where E=207 GPa(30×10^3ksi), b= -0.0843, c= -0.9126, σ_f' and ε_f' are lognormally distributed random variables. Using a 5% EPI, σ_f' has a median = 1839 MPa (266.7 ksi) and COV = 0.049; ε_f' has a median = 3.47 and COV = 0.425.

The fortuitous result with the Waspaloy allows us to avoid having to deal with dependent variables in a reliability format. But is it general? Clearly, the deviations should be examined for each material and condition. In another study, a sample of n=148 points of anal-gamated data on RQC-100 (see Table X) also produced a reasonably uni-form scatter of deviations as shown in Fig. 13.

40
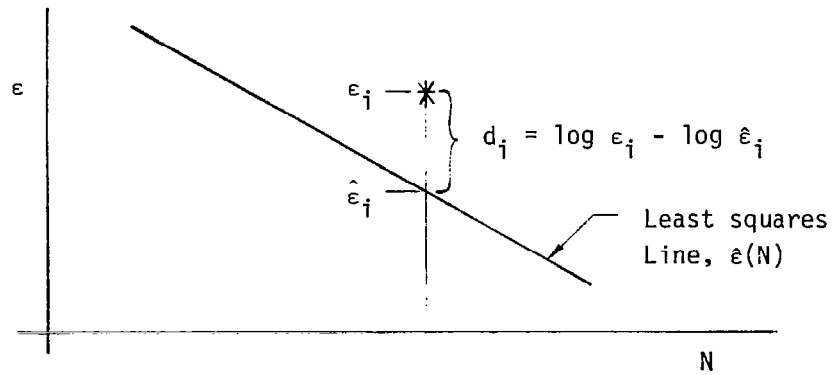
Fig. 11

(a) Deviations from least squares line

Plastic

ε_{pa}

Case 3
HIGH DUCTILITY
LOW TENSILE STRENGTH

Case 2
LOW FATIGUE
STRENGTH

N

Elastic

ε_{ea}

Least squares
Line

N

(b) Definition of Deviations

ε

ε_i —

$\hat{\varepsilon}_i$ —

$d_i = \log \varepsilon_i - \log \hat{\varepsilon}_i$

Least squares
Line, $\hat{\varepsilon}(N)$

N

41

Fig. 12

WASPALOY
(T=1000F, n=44)

Vertical Deviations (log basis)

$$\frac{d_i}{\frac{\Sigma |\overline{d_i}|}{n}}$$
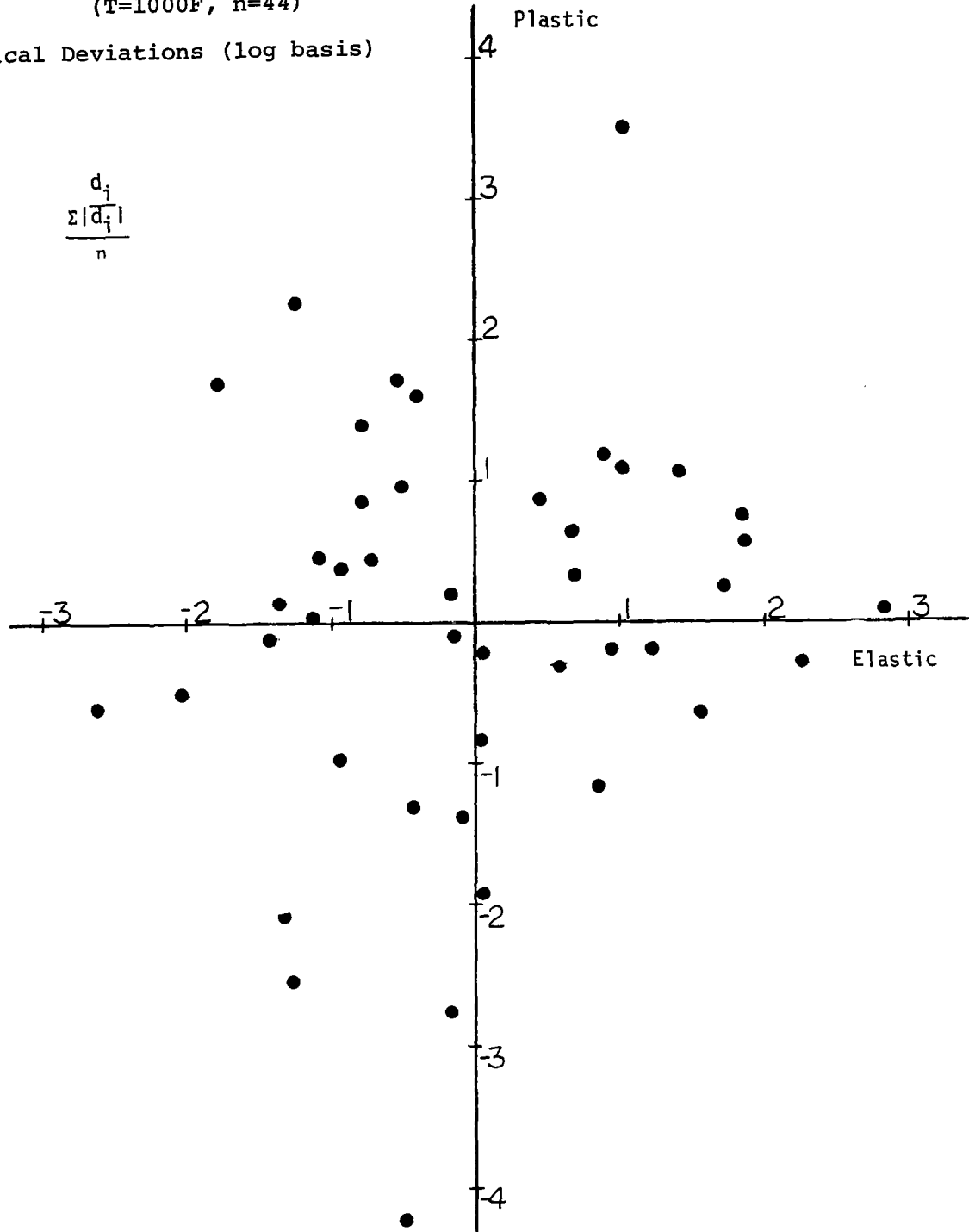
Fig 13

RQC-100 (n=148)

Vertical Deviations (log basis)

$$\frac{\frac{d_i}{\Sigma|d_i|}}{n}$$

Plastic

Elastic
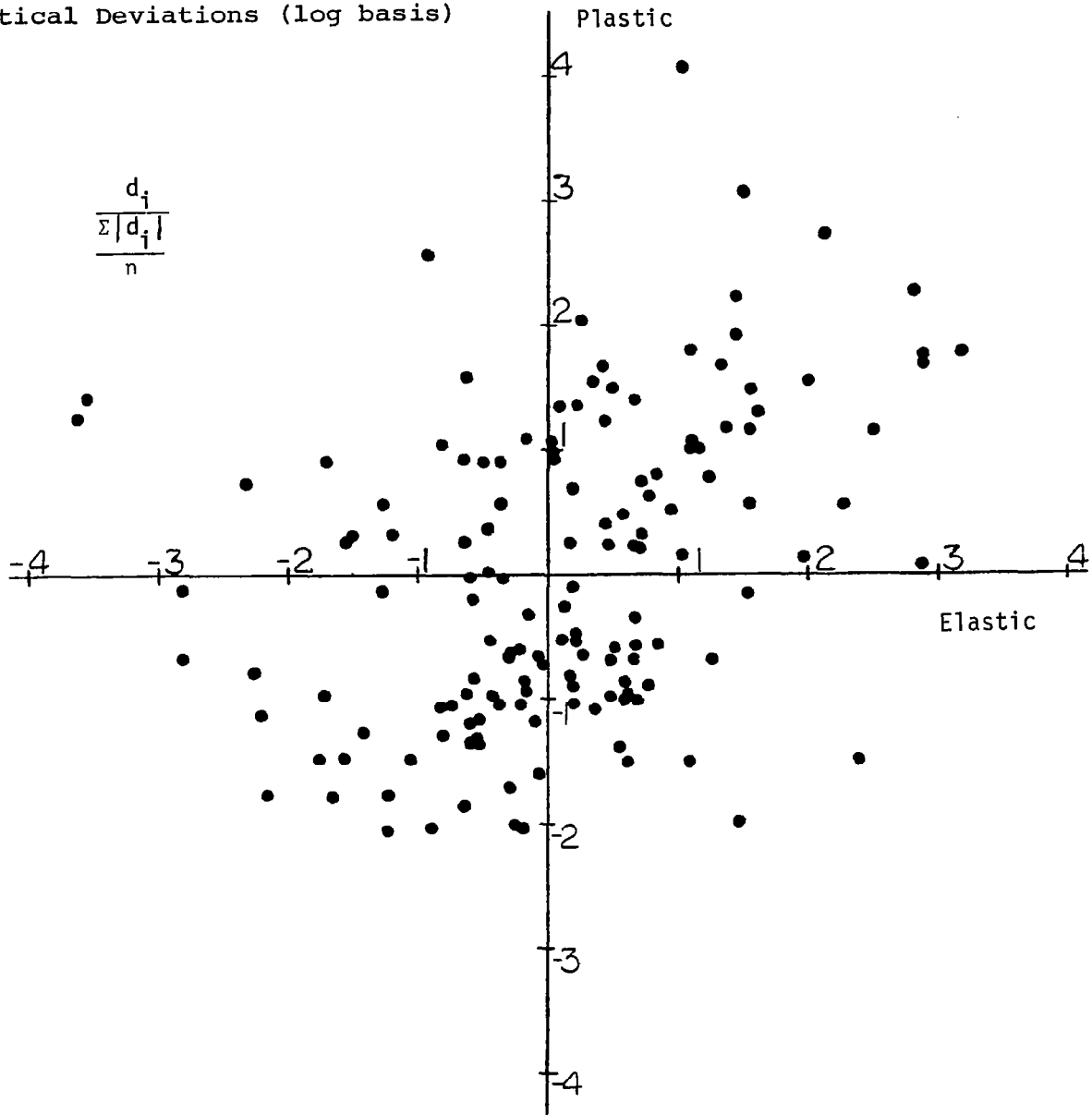
A MODEL FOR THE DISTRIBUTION OF FATIGUE STRENGTHS AT HIGH CYCLE LIVES

The number of load cycles experienced during the lifetime of a wide variety of structural and mechanical components may be in the range from 10^7 to 10^{10} cycles. Often it is necessary for the designer to know the distribution of fatigue strength (the stress endurance limit) for a given life N out in the high cycle region where the S-N curve is assumed to be flat.

A number of statistical methods of estimating the fatigue strength at a given cycle life are available. Such techniques, summarized by Reemsnyder (28), Lipson and Sheth (21), Little (23), and Collins (5) include the survival method, and its more refined form, the Probit method, the staircase or up-and-down method, the step-test method, and the Prot method. These procedures typically require long cycle lives and relatively large sample sizes; testing tends to become expensive.
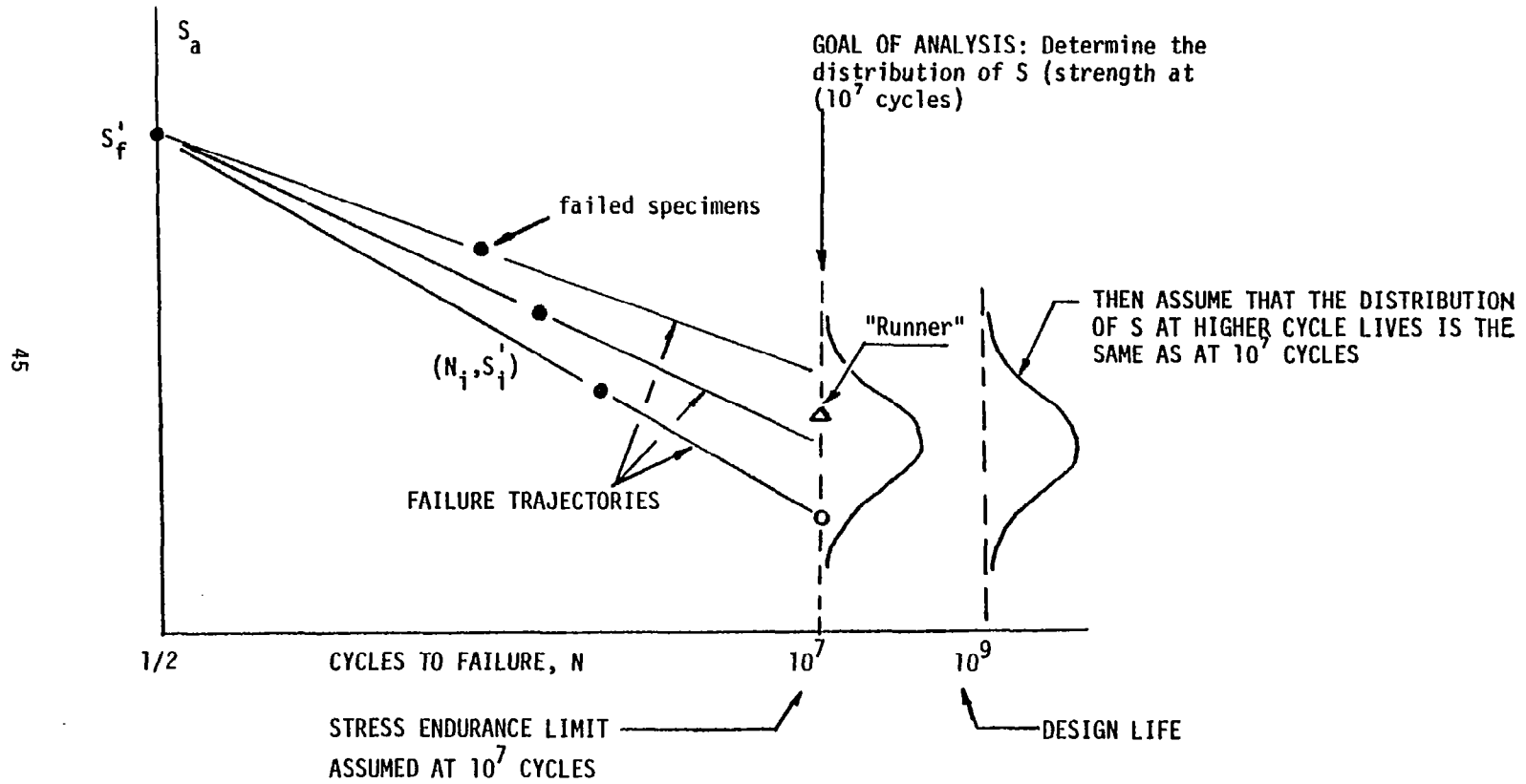
Fatigue testing is expensive, partly because of the length of time that it takes to apply these long cycle lives. Described as follows is an accelerated test method which has been used successfully. Stress levels are chosen so that specimens fail at relatively short lives (typically 10^5 to 10^6 cycles). The data are extrapolated to higher cycle lives.

Consider a constant amplitude high cycle fatigue test in which pairs of data (S_i', N_i), i = 1, n are collected. Data from a hypothetical test are shown in Fig. 14. An endurance limit is assumed as illustrated by the horizontal segment of this S-N curve. It is assumed that the endurance limit occurs at 10^7 cycles. Furthermore, it is assumed that the distribution of S at a given N at 10^7 cycles will be the same for any higher life cycle as shown in Fig. 14.

Figure 14

A MODEL FOR HIGH CYCLE FATIGUE

(Typical Use; Machine Components)

Stress Amplitude

GOAL OF ANALYSIS: Determine the distribution of S (strength at 10^7 cycles)

failed specimens

"Runner"

THEN ASSUME THAT THE DISTRIBUTION OF S AT HIGHER CYCLE LIVES IS THE SAME AS AT 10^7 CYCLES

(N_i, S_i')

FAILURE TRAJECTORIES

1/2   CYCLES TO FAILURE, N 10^7 10^9

STRESS ENDURANCE LIMIT ASSUMED AT 10^7 CYCLES

DESIGN LIFE

A fatigue test plan specifies a termination of the loading on a component which does not fail before a given life. This specimen is called a "runner" and the data points are shown by the triangles of Fig. 14 in which the test was suspended at 10^7 cycles. This data must be included in the analysis.

To obtain a random sample of S, now defined as a random variable denoting the fatigue strength at 10^7 cycles, the following procedure is used. It is assumed that the fatigue strength for a given specimen is defined by a straight line (failure trajectory) connecting the fatigue strength coefficient at σ_f' at 1/2 cycle and (N_i, S_i) as shown in Fig. 14. The data point (solid point in Fig. 14) is projected to the life cycle where the endurance limit occurs, which is assumed here to be 10^7 cycles. The sample point (circle in Fig. 14) is denoted as S_i. By such a scheme one can obtain a random sample of S.

Incomplete failure data consisting of levels on failed components and unfailed components are called "multiply censored." This suspended data can be analyzed utilizing the "median rank concept" and suspended items approach (21). The median ranks extracted from this approach will be used to establish the empirical distribution function of fatigue strengths.

Lipson and Sheth have described a method for obtaining the empirical distribution function from a random sample which includes suspended items (21). This method is a combination of a modified sudden death approach and Johnson's concept of median ranking . The method involves

an adjustment of the order number based upon the position of the suspended items. Step-by-step instructions are illustrated by the following:

1. Suppose that the failure data on n specimens consist of the failure stresses (at 10^7 cycles for the failed units and the stress levels for the unfailed units (see Table XI and Fig. 15). Order the failed units in the sample from the smallest to largest failure stresses as shown in Column 3.

2. Obtain the number of suspended items which precede each failed unit (Column 2).

3. Determining the "new increment" of each failed unit by using the formula

$$\text{New Increment} = \frac{(n + 1) - \text{previous failure order number}}{1 + \text{number of items following present suspended set}}$$

The new increment is recalculated each time a suspension is encountered in the ordered stress table (see Column 4).

4. Calculate the order number for each failed unit. This calculation is done by simple addition of the last order number and the new increment (Column 5).

5. Obtain the median rank (empirical distribution function) of the order number for each failed unit by the formula (Column 6).

The empirical distribution function $F(S_{(j)})$ of Column 6 is used as a basis for selecting a suitable statistical model.

47

Table XI

Analysis of Suspended Fatigue Data

EXAMPLE: Given a sample of fatigue failure stress at 10^7 cycles (in ksi); n = 13; the data are ordered.

| 11.7 | 12.2 | 12.5 | 12.7 | 12.8 | 13.3 | 13.8 | 14.0* | 14.0* | 14.0* | 14.4 | 14.7 | 15.0* |

*Specimen did not fail at 10^7 cycles.

Organize data as shown below:

(1) Failure Number	(2) Suspended Items Preceeding Failure	(3) Stress $S_{(j)}$	(4) New Increment**	(5) Order Number** (j)	(6) Median Rank $F(S_{(j)}) = \dfrac{j - 0.3}{n + 0.4}$
1	0	11.7	1.00	1.00	0.052
2	0	12.2	1.00	2.00	0.127
3	0	12.5	1.00	3.00	0.201
4	0	12.7	1.00	4.00	0.276
5	0	12.8	1.00	5.00	0.351
6	0	13.3	1.00	6.00	0.425
7	0	13.8	1.00	7.00	0.500
8	3	14.4	1.75	8.75	0.631
9	0	14.7	1.75	10.50	0.761

**An example of how the new increment and order number are calculated:

The eight failure is preceded by three suspended items. Therefore, to find the increment, as shown in Step 3 above,

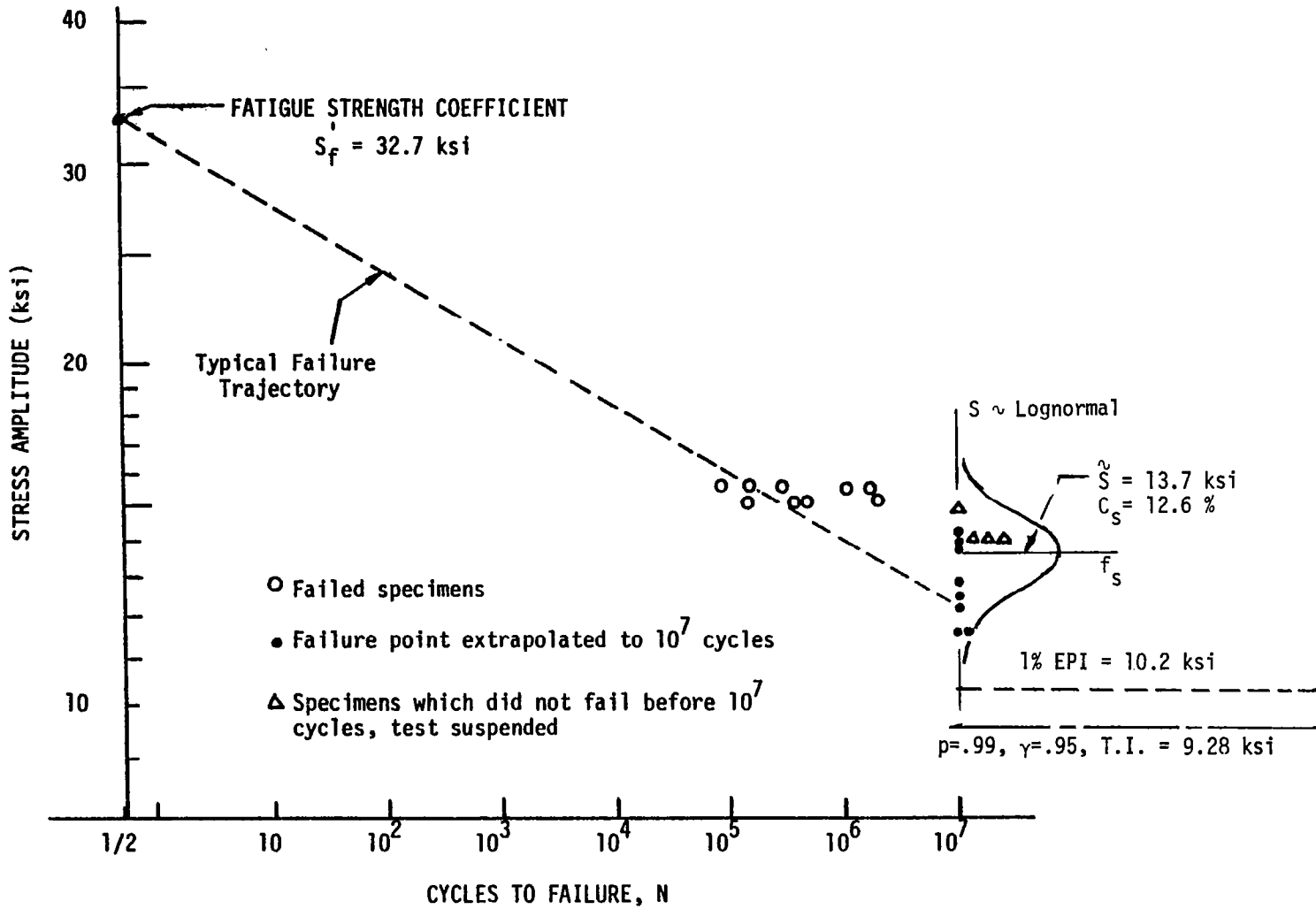$$I = \frac{(13 + 1) - 7}{1 + 3} = 1.75$$

Thus, the new order number is j = 7.00 + 1.75 = 8.75

The remaining order numbers are determined by using the same procedure and

$$F = \frac{8.75 - 0.3}{13 + .4} = 0.631$$

Figure 15

S-N DIAGRAM FOR EXAMPLE OF DATA EXTRAPOLATED INTO HIGH CYCLE RANGE

The empirical distribution function (Column 6) produces a linear plot on lognormal paper as shown in Fig. 16. Assume that fatigue strength S is lognormally distributed. Estimates of the lognormal parameters for S are computed as suggested in Fig. 16.

Unfortunately, the distributions of the estimators are not known so that the tolerance interval and EPI concepts cannot be used to provide accurate statistical summaries for design purposes. However, as a first approximation they can provide design values which do at least recognize uncertainties in estimators. The estimated lower tolerance limit for $P = .99$ and $\gamma = .95$ and for $\bar{X} = 2.62$, $\sigma_X = 0.107$ and $n = 13$ is $X_L = 2.228$ and $S_L = \exp (X_L) = 9.28$ ksi (26). This value is plotted on Fig. 15.

The equivalent prediction interval for a single variable is discussed in Ref. 40. Again the assumption is made that the distribution of the estimators is the same as for a complete sample. It follows $\alpha = 0.01$ and $n = 13$ that $g (.01, 13) = 1.18$, and the equivalent standard deviation is $\sigma_0 = 0.126$. The lower 1% EPI is $X_L = 2.325$, and $S_L = \exp (X_L) = 10.23$ ksi (38,40). This value is also shown in Fig. 15. For a reliability analysis, S can be assumed to be lognormal with median $\tilde{S} = 13.7$ ksi and $C_S = 12.6\%$ (40).

Fatigue analysis using this extrapolation approach depends on where the endurance limit is chosen. For steels, it is generally considered that the knee of the S-N curve falls below 10^7 cycles so the choice of 10^7 may tend to produce conservative statistics.

SUMMARY

Metallic materials exhibit very significant scatter in fatigue data. It is necessary to provide a statistical summary of such data from which designers can make decisions which balance risk and cost. This paper

Figure 16

Lognormal Probability Plot for Fatigue Strength S beyond 10^7 Cycles from the Data of Table XI
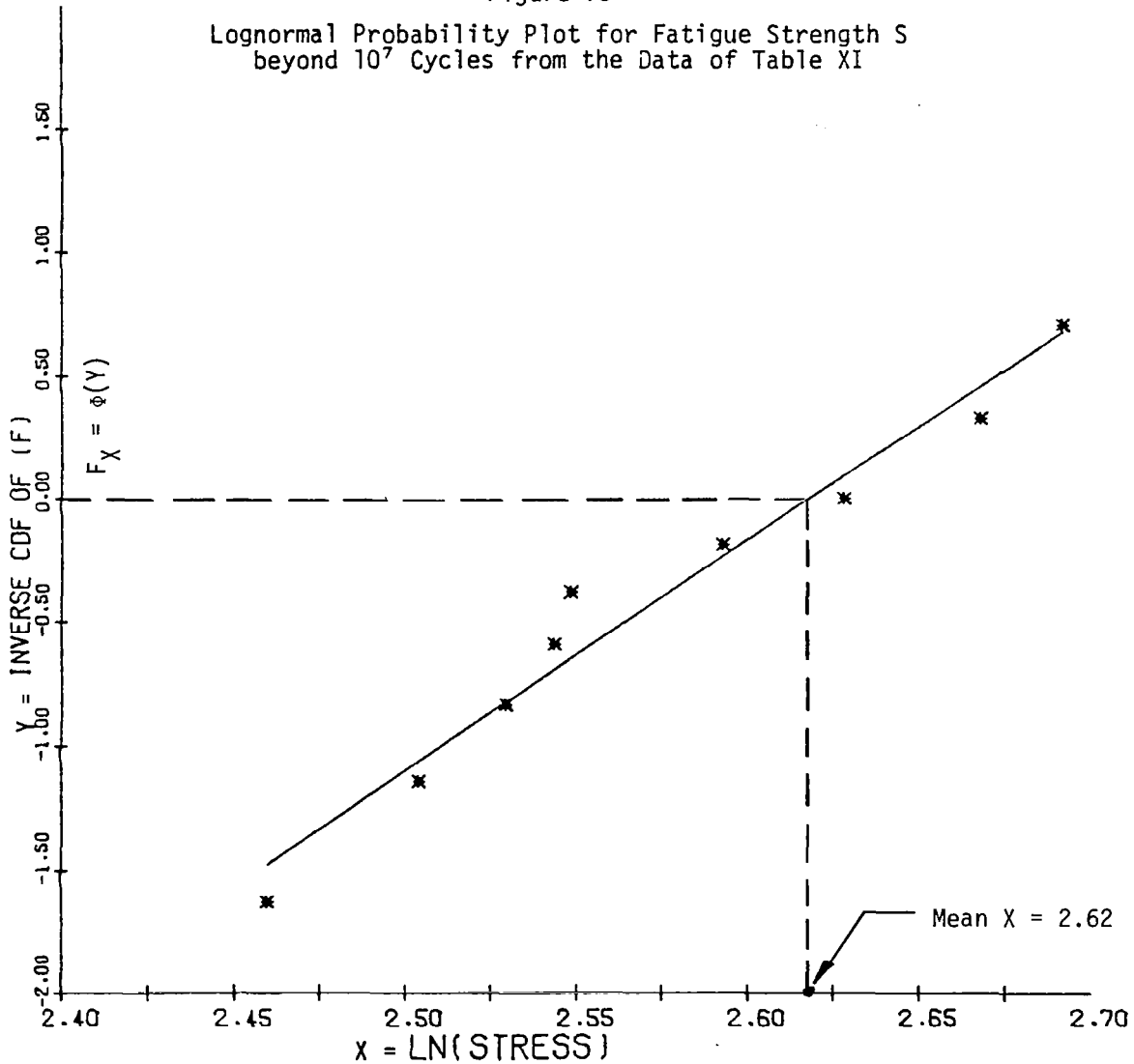
Statistics on X (computed from the least squares analysis of the probability plot)

μ_X = 2.62

σ_X = 0.107

Statistics on S,

Median (S) = $\tilde{S}$ = exp (μ_X) = 13.7 ksi

COV of S = C_S = $\sqrt{\exp{(\sigma_X^2)-1}}$ = 0.107

C_S = 0.126 using σ_0, the EPI concept

summarized some methods of analysis of fatigue data and presented examples which illustrate the general character of fatigue data for materials. No specific recommendations are made regarding how data should be treated, but general information is presented which should provide guidance in constructing strategies for fatigue avoidance.

General conclusions and observations are as follows.

1. The method of analysis of fatigue data depends in part how the fatigue failure condition is defined as well as the consequences of failure and the cost of overdesign.

2. The simple linear model (on a log basis) can be used with accuracy for the general strain-life relationships, strainrange partitioning data, high cycle welded joint data, and other applications.

3. Two methods of analyzing S-N data are presented; (a) the design curve which is a lower bound on the safe side of the data, established from a tolerance interval or equivalent prediction interval (EPI), and (b) parameters of a characteristic equation presented as random variables so that fatigue strength can be treated as a random variable in a reliability format.

4. The concept of failure trajectories was described as a model for defining the distribution of N given S for the whole range of S, and for converting to the distribution of S given N. This convenient model can be useful in other applications involving strength data.

5. Typical values of the COV for fatigue data for a wide variety of metallic materials

N, Cycles to failure (life given stress)	30 to 75%
S, Fatigue Strength (stress given life)	5 to 15%

For many smooth specimen data, the COV of N is about 40%. The COV of S beyond the endurance limit is about 10%.

6. The statistical distribution of the least squares estimators is an important consideration in constructing a design model from those small sample sizes which are common in fatigue. As a rule, if $n \geq 7$, the influence of estimator variability is not excessive although it should be considered for n up to 50.

7. Based on a fairly limited number of experiences, it appears that basic similarities exist in the general character of fatigue data. Aluminum and titanium alloys, steels, and nickel base superalloys (at high temperatures) seem to exhibit the same linear and homoscedastic trends in strain life data, and also have about the same amount of scatter.

NOTATION

a   y intercept of $Y_o(x)$

$\hat{a}$ least squares estimate of a

a_o normal variate with mean $\hat{a}$ and standard deviation σ_o

A   coefficient in S-N curve; Eq 1 and 20

$A(\alpha)$ Eq 26

b   fatigue strength exponent; also slope of $Y_o(x)$

$\hat{b}$ least squares estimate of b

B   coefficient in S-N curve; Eq 3

$B(\alpha)$ Eq 26

c   fatigue ductility exponent

C   coefficient in strain-life equation; Eq 6

C_A coefficient of variation of A

C_N coefficient of variation of N

COV   coefficient of variation

cdf   cumulative distribution function

E   modulus of elasticity

$E(Y|X)$ expected value of Y given X

EPI   equivalent prediction interval

$f_{N/S}$ pdf of N given S

$f_{S/N}$ pdf of S given N

$F_N(x)$ cdf of N

NOTATION (continued)

$g(n,\alpha)$	reduction function
$K_{\alpha,\gamma}$	tolerance factor
m	exponent of S-N curve; Eq 1 and 20
n	sample size
N	cycles to failure
N_i	ith observed value of N
$\dot{N}^*$	EPI S-N curve
N_D	design S-N curve
pdf	probability density function
s	estimate of σ
S	amplitude (or range) of fatigue stresses
S_i	ith observed value of S
S_e	stress endurance limit
S_o	mean stress
S_n	ultimate strength
TPW	three parameter Weibull
X	log ($\Delta\varepsilon$)
Y	log N
Y_o	expected value of Y given X
$\hat{Y}$	least squares estimate of Y_o

NOTATION (continued)

z_a standard normal variate

α population fraction; 100α = % less than

β scale parameter in Weibull cdf

γ confidence level

$\Delta\varepsilon$ strain range

δ normal random variable with mean zero and standard deviation of one

ε_a strain amplitude

ε_f' fatigue ductility coefficient

ζ exponent in strain-life equation; Eq 6

η shape parameter in Weibull cdf

ξ location parameter in Weibull cdf

σ standard deviation of Y given X

σ_0 equivalent standard deviation

σ_f' fatigue strength coefficient

REFERENCES

1. A Guide for Fatigue Testing and Statistical Analysis of Fatigue Data, ASTM STP 91-A, American Society for Testing and Materials, 1963.

2. Ang, A. H-S. and Munse, W. H., "Practical Reliability Basis for Structural Fatigue," presented at ASCE Structural Engineering Conference, New Orleans, LA, April 1975, Preprint 2494.

3. Basquin, H. O., "The Exponential Law of Endurance Tests," Proceedings of the ASTM, Vol. 10, Part II, 1910.

4. Boardman, B. E., "Crack Initiation Fatigue Data, Analysis, Trends, and Estimation," Proceedings of the SAE Fatigue Conference, P-109, SAE, 1982

5. Collins, J. A., Failure of Materials in Mechanical Design, McGraw-Hill, 1981.

6. Criteria for Design of Elevated Temperature Class I Components in Section III, Division 1 of the ASME Boiler and Pressure Vessel Code, May 1976

7. Dowling, N. E., Brose, W. R. and Wilson, W. K., "Notched Member Fatigue Life Predictions by the Local Strain Approach," Fatigue Under Complex Loading; Analyses and Experiments, AE-7, SAE, 1977.

8. "Fatigue Reliability," a four-part series, Journal of the Structural Division, ASCE, Vol. 108, No. ST1, Jan. 1982.

9. Fatigue Under Complex Loading: Analyses and Experiments, AE-6, SAE, 1977.

10. Freudenthal, A. M., "Safety of Structures," Transactions, ASCE, Vol. 112, 1947, pp. 125-180.

11. Freudenthal, A. M. and Gumbel, E. J., "On the Statistical Interpretation of Fatigue Tests," Proceedings of the Royal Society of London, Series A, Vol. 216, 1953, pp. 309-322.

12. Freudenthal, A. M. and Gumbel, E. J., "Physical and Statistical Aspects of Fatigue," Advances in Applied Mechanics, Vol. 4, 1956.

13. Fuchs, H. O. and Stephens, R. K., Metal Fatigue in Engineering, Wiley, 1980.

14. Gumbel, E. J., Statistics of Extremes, Columbian Univ., Press 1958.

15. Gumbel, E. J., "Parameters in the Distribution of Fatigue Life," Journal of the Engineering Mechanics Division, ASCE, Vol. 89, No. EM5, Oct. 1963.

16. Haugen, E. B., Probabilistic Mechanical Design, Wiley, 1980.

17. Jaske, C. E., and O'Donnell, W. J., "Fatigue Design Criteria for Pressure Vessel Alloys," Journal of Pressure Vessel Technology, ASME, Nov. 1977.

18. Kececioglu, D., et al., "Combined Bending Torsion Fatigue Reliability of AISI 4340 Steel Shafting with K_t = 2.34," ASME Paper No. 74-WA/DE-12, 1974.

19. Morrow, J., "Fatigue Properties of Metals", Sec. 3.2 of Fatigue Design Handbook, SAE, 1968.

20. Lawrence, F. V., et al., "Estimating the Fatigue Crack Initiation Life of Welds," Fatigue Testing of Weldments, ASTM STP-648, 1978.

21. Lipson, C., and Sheth, N. J., Statistical Design and Analysis of Engineering Experiments, McGraw-Hill, 1973.

22. Little, R. E., and Jebe, E. H., Statistical Design of Fatigue Experiments, Applied Science, 1975.

23. Little, R. E., Manual on Statistical Planning and Analysis, ASTM, STP 588, 1975.

24. MIL-HDBK-5B, Metallic Materials and Elements for Aerospace Vehicle Structures, U.S. Department of Defense.

25. Mood, A. M. and Graybill, F. A., Introduction to the Theory of Statistics, McGraw-Hill, 1963.

26. Natrella, M. G., Experimental Statistics, NBS Handbook 91, 1963

27. Proceedings of the SAE Fatigue Conference, P-109, Dearborn, MI, April 14-16, 1982

28. Reemsnyder, H. S., "Procurement and Analysis of Structural Fatigue Data," Journal of the Structural Division, ASCE, Vol. 95, No. ST 7, July 1969.

29. SAE Handbook, Society of Automotive Engineers, 1981.

30. Saltsman, J. F. and Halford, G. R., "Application of Strainrange Partitioning to the Prediction of Creep Fatigue Lives of AISI Types 304 and 316 Stainless Steel,: ASME Paper No. 76-PVP-26, 1976.

31. "Section VIII, Division II - Alternative Rules," ASME Boiler and Pressure Vessel Code, ASME, New York, 1974.

32. Sinclair, G. M. and T. J. Dolan, "Effect of Stress Amplitude on the Variability in Fatigue Life of 75S-T6 Aluminum Alloy," Transactions of the ASME, July 1953.

33. Symposium on Statistical Aspects of Fatigue, ASTM, STP 121, June 1951.

34. Symposium on Fatigue With Emphasis on Statistical Approach, ASTM STP 137, June 1952.

35. Whittaker, I. C., "Development of Titanium and Steel Fatigue Variability Model for Application of Reliability Approach to Aircraft Structures," AFML-TR-72-236, Wright-Patterson AFB, Ohio, October 1972.

36. Wirsching, P. H., and Carlson, J. R., "Model Identification for Engineering Variables," Journal of the Engineering Mechanics Division, Vol. 103, No. EM1, Feb. 1977.

37. Wirsching, P. H., and Light, M. C., "Probability Based Fatigue Design Criteria for Offshore Structures," API-PRAC 15, Final Report, First-Year Effort, American Petroleum Institute, Nov. 1979.

38. Wirsching, P. H. and S. Hsieh, "Linear Model in Probabilistic Fatigue Design," The Journal of the Engineering Mechanics Division, ASCE, Vol. 106, No. EM6, Dec. 1980.

39. Wirsching, P. H., "Probability Based Fatigue Design Criteria for Offshore Structures," Final Report API PRAC 80-15, American Petroleum Institute, Feb. 1981.

40. Wirsching, P. H., "The Application of Probabilistic Design Theory to High Temperature Low Cycle Fatigue, NASA, CR-165488, NASA/Lewis Research Center, Nov. 1981.

41. Wirsching, P. H., "Probability Based Fatigue Design Criteria for Offshore Structures," Final Report (in preparation), American Petroleum Institute, 1982.

42. Yang, J. N., "Fatigue Reliability Analysis of Aircraft Structures," Probabilistic Methods in Structural Engineering, (ed. M. Shinozuka, J. T. P. Yao), ASCE, 1981.

43. Yao, J. T. P., "Fatigue Reliability and Design," Journal of the Structural Division, ASCE, Vol. 100, No. ST9, September 1974.

1. Report No. NASA CR-3697	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle STATISTICAL SUMMARIES OF FATIGUE DATA FOR DESIGN PURPOSES		5. Report Date July 1983
		6. Performing Organization Code
7. Author(s) Paul H. Wirsching		8. Performing Organization Report No. None
		10. Work Unit No.
9. Performing Organization Name and Address The University of Arizona Aerospace and Mechanical Engineering Dept. Tucson, Arizona 85721		11. Contract or Grant No. NAG-3-41
		13. Type of Report and Period Covered Contractor Report
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		14. Sponsoring Agency Code 505-33-72 (E-1602)

15. Supplementary Notes

16. Abstract

Fatigue data are subject to considerable scatter and cycles to failure, N, can have coefficients of variation typically ranging from 25% to 75%. Presented herein are techniques for providing statistical summaries for such data suitable for design purposes. Special consideration is given to statistical difficulties presented by the small sample sizes characteristic of fatigue data. Two methods are discussed for constructing a design curve on the safe side of the data. Both the tolerance interval and equivalent prediction interval (EPI) concepts provide such a curve while accounting for both the distribution of the estimators in small samples and the data scatter. Moreover the EPI is useful as a mechanism for providing necessary statistics on S-N data for a full reliability analysis which includes uncertainty in all fatigue design factors. Presented are examples of statistical analyses of the general strain life relationship. The tolerance limit and EPI techniques for defining a design curve are demonstrated. Moreover in two examples using Waspaloy B and RQC-100 data it was demonstrated that a reliability model could be constructed by considering the fatigue strength and fatigue ductility coefficients as two independent random variables. A technique for establishing the fatigue strength for high cycle lives was presented. This method relies on an extrapolation technique and also accounts for "runners." Again, a reliability model or design value can be specified.

17. Key Words (Suggested by Author(s)) Reliability; Probabilistic design; Fatigue; Statistics; Tolerance interval	18. Distribution Statement Unclassified - unlimited STAR Category 39

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of pages 62	22. Price* A04