

STATISTICAL ANALYSIS OF FATIGUE DATA

Little/Ekvall, *editors*

 **STP 744**

AMERICAN SOCIETY FOR
TESTING AND MATERIALS

STATISTICAL ANALYSIS OF FATIGUE DATA

A symposium
sponsored by ASTM
Committee E-9 on Fatigue
AMERICAN SOCIETY FOR
TESTING AND MATERIALS
Pittsburgh, Pa., 30-31 Oct. 1979

ASTM SPECIAL TECHNICAL PUBLICATION 744
R. E. Little, University of Michigan
at Dearborn, and J. C. Ekvall,
Lockheed-California Company,
editors

ASTM Publication Code Number (PCN)
04-744000-30



AMERICAN SOCIETY FOR TESTING AND MATERIALS
1916 Race Street, Philadelphia, Pa. 19103

Copyright © by AMERICAN SOCIETY FOR TESTING AND MATERIALS 1981
Library of Congress Catalog Card Number: 81-65835

NOTE

The Society is not responsible, as a body,
for the statements and opinions
advanced in this publication.

Printed in Philadelphia, Pa.
August 1981

Foreword

The symposium on Statistical Analysis of Fatigue Data was held on 30–31 Oct. 1979 in Pittsburgh, Pa. The American Society for Testing and Materials, through its Committee E-9 on Fatigue, sponsored the event. R. E. Little of the University of Michigan at Dearborn presided as chairman, and J. C. Ekvall of the Lockheed–California Company served as cochairman. Both men served as editors of this publication.

Related ASTM Publications

Probabilistic Aspects of Fatigue, STP 511 (1972), \$19.75, 04-511000-30

Handbook of Fatigue Testing, STP 566 (1974), \$17.25, 04-566000-30

Service Fatigue Loads Monitoring, Simulation, and Analysis, STP 671 (1979), \$29.50, 04-671000-30

Fatigue Mechanisms, STP 675 (1979), \$65.00, 04-675000-30

Part-Through Crack Fatigue Life Prediction, STP 687 (1979), \$26.25, 04-687000-30

Crack Arrest Methodology and Applications, STP 711 (1980), \$44.75, 04-711000-30

Fracture Mechanics: Twelfth Conference, STP 700 (1980), \$53.25, 04-700000-30

A Note of Appreciation to Reviewers

This publication is made possible by the authors and, also, the unheralded efforts of the reviewers. This body of technical experts whose dedication, sacrifice of time and effort, and collective wisdom in reviewing the papers must be acknowledged. The quality level of ASTM publications is a direct function of their respected opinions. On behalf of ASTM we acknowledge with appreciation their contribution.

ASTM Committee on Publications

Editorial Staff

Jane B. Wheeler, *Managing Editor*
Helen M. Hoersch, *Senior Associate Editor*
Helen P. Mahy, *Senior Assistant Editor*
Allan S. Kleinberg, *Assistant Editor*

Contents

Introduction	1
Review of Statistical Analyses of Fatigue Life Data Using One-Sided Lower Statistical Tolerance Limits—R. E. LITTLE	3
Statistical Design and Analysis of an Interlaboratory Program on the Fatigue Properties of Welded Joints in Structural Steels— E. HAIBACH, R. OLIVIER, AND F. RINALDI	24
Reliability of Fatigue Testing—L. YOUNG AND J. C. EKVALL	55
Statistical Fatigue Properties of Some Heat-Treated Steels for Machine Structural Use—S. NISHIJIMA	75
Some Considerations in the Statistical Determination of the Shape of $S-N$ Curves—J. E. SPINDEL AND E. HAIBACH	89
Maximum Likelihood Estimation of a Two-Segment Weibull Distribution for Fatigue Life—P. C. CHOU AND HARRY MILLER	114
Appendix—ASTM Standard Practice for Statistical Analysis of Linear or Linearized Stress-Life ($S-N$) and Strain-Life ($\epsilon-N$) Fatigue Data (E 739-80)	129
Summary	138
Index	143

Introduction

One cannot use fatigue data competently in either design or research and development without first explaining (understanding) and assessing (measuring) variability in the test results. Maximum likelihood analysis has emerged as a major statistical tool in explaining fatigue variability—because it can be used to analyze and study even very complex mathematical fatigue models. Once an adequate statistical model has been established by appropriate study, it is vital that the associated random fatigue variability be assessed properly using test results generated by replicate experiments in a statistically planned test program. Only then may we presume to predict fatigue behavior reliably.

The two major areas considered in this Special Technical Publication are (1) maximum likelihood analysis used as a tool in the statistical analysis of fatigue data and in the study of alternative fatigue models and (2) assessment of fatigue variability using statistically planned test programs with appropriate replication. Since adequate statistical models and accurate assessment of random variability form the foundation of reliable prediction, this volume should be conceptually very useful to practitioners of fatigue analysis. In fact, it is likely that the concepts considered in this publication will become the cornerstone of statistical analyses of fatigue data in the 1980s and beyond.

The 1980s will also see routine use of elaborate digital computer software¹ for maximum likelihood analyses, as well as widespread use of the likelihood ratio test statistic, not only to study and assess the adequacy of alternative fatigue models but also to establish intervals estimates for reliable life. In this context, this publication is meant to preview what is coming in the next decade and beyond rather than to summarize what has been done recently.

The major issue to be resolved in the 1980s is how to come to grips with the discrepancies between the idealizations of test planning and mathematical analyses and the realities of practical procedures of actual test conduct so that ultimately fatigue variability may be assessed reliably. Certain aspects of this problem are presented elsewhere², but a specific example discussed here

¹Refer, for example, to Nelson, W. D., Hendrickson, R., Phillips, M. C., and Shumbart, L., "STATPAC Simplified—A Short Introduction to How to Run STATPAC, A General Statistical Package for Data Analysis," Technical Information Series Report 73 CRD 046, General Electric Co., Corporate Research and Development, Schenectady, N.Y., July 1973. (Available by writing to Technical Information Exchange, 5-237, G.E. Corp. R&D, Schenectady, N.Y. 12345.)

²Little, R. E., *ASTM Standardization News*, Vol. 8, No., 2, Feb. 1980, pp. 23-25.

will help define the issue. The current practice, as elaborated in recent textbooks and short courses, is to assume that the fatigue limit for steel is normally distributed with a standard deviation equal to (at most) 8 percent of its median value. Thus, in theory, one can estimate the alternating stress amplitude that corresponds to a probability of failure equal to 0.000001. However, several test programs have been conducted involving simple sinusoidal loading of real components (for example, high-strength bolts and forged and heat-treated valve bridges) instead of conventional laboratory specimens. The standard deviations obtained from these programs are two to three times as large as the rule-of-thumb estimate. Moreover, it has been observed that strength distributions are clearly not normal. These results indicate that the textbook estimate is generally misleading and sometimes very dangerous. The fundamental problem, of course, is that conventional laboratory tests are specifically conducted using procedures that circumvent and minimize fatigue variability. Accordingly, the results of conventional laboratory tests do not form a sound basis for predicting the fatigue variability of real components. Statistical theory indicates that we can predict fatigue behavior reliably only when the future tests of interest are nominally identical to the original tests whose data were used to compute the prediction intervals. In other words, if one wishes to predict service performance, service tests must be conducted to generate relevant data for prediction purposes. Such tests may be impractical, but, nevertheless, the discrepancy between theory and practice must be reduced. This discrepancy presents a formidable challenge to all fatigue practitioners to improve both the quality of statistical analyses and the relevance of the associated fatigue tests by appropriate planning. We hope that the reader will accept this challenge and that this publication will provide some help in that effort.

R. E. Little

University of Michigan, Dearborn, Mich.
48128; symposium chairman and editor.

J. C. Ekvall

Lockheed-California Co., Burbank, Calif.
91520; symposium cochairman and editor.

Review of Statistical Analyses of Fatigue Life Data Using One-Sided Lower Statistical Tolerance Limits

REFERENCE: Little, R. E., "Review of Statistical Analyses of Fatigue Life Data Using One-Sided Lower Statistical Tolerance Limits," *Statistical Analysis of Fatigue Data, ASTM STP 744*, R. E. Little and J. C. Ekvall, Eds., American Society for Testing and Materials, 1981, pp. 3-23.

ABSTRACT: This introductory paper explains basic probability concepts and summarizes in a fatigue context the state of the art for analyses of life data using one-sided lower statistical tolerance limits. Types I and II censoring are considered for both the two-parameter log-normal and Weibull distributions, and the corresponding approximate and exact one-sided lower tolerance limit calculations are illustrated and discussed. In addition, Antle's likelihood ratio test for discriminating between these two-parameter life distributions is summarized. The classic one-sided lower nonparametric tolerance limit analysis and a small sample modification by Little are discussed and illustrated in a fatigue context. Overall, this paper is intended to provide background and perspective for subsequent papers.

KEY WORDS: tolerance limits, one-sided lower tolerance limits, two-parameter log-normal distribution, two-parameter Weibull distribution, statistical analysis, fatigue life, fatigue

The objective of this paper is to elucidate in a fatigue context the state of the art in computation of one-sided lower statistical tolerance limits.

First, I shall provide some background and terminology for readers with little statistical training.

Background and Terminology

Consider the probability expression

$$\text{Prob} [z_{\text{lower}} < Z < z_{\text{upper}}] = \gamma \quad (1)$$

in which z_{lower} and z_{upper} are numbers (denoted by lower case letters), Z is a

¹Professor, University of Michigan-Dearborn, Mich. 48128.

4 STATISTICAL ANALYSIS OF FATIGUE DATA

random variable (denoted by a capital letter), and $0 \leq \gamma \leq 1$. Given a specific future realization of the random variable Z , say z^* , the realization will either lie within the interval from z_{lower} to z_{upper} or it will not, and we cannot tell which until we have conducted the appropriate experiment and observed its outcome. Nevertheless, we can assert that, in the long run, γ proportion of all future realizations associated with this experiment will lie within the given interval. Refer to Fig. 1.

The interval from z_{lower} to z_{upper} in probability Expression 1 is termed a two-sided probability interval. Specifically, this interval is bounded by the lower limit, z_{lower} , and the upper limit, z_{upper} . Accordingly, Expression 1 is more properly termed a two-sided probability interval expression. The associated one-sided lower probability interval expression is

$$\text{Prob} [z_{\text{lower}} < Z] = \gamma_{\text{lower}} \quad (2a)$$

and the associated one-sided upper probability interval expression is

$$\text{Prob} [Z < z_{\text{upper}}] = \gamma_{\text{upper}} \quad (2b)$$

provided that

$$(1 - \gamma_{\text{lower}}) + (1 - \gamma_{\text{upper}}) = (1 - \gamma)$$

Most statistical applications of probability expressions are based on theoretical arguments involving certain equivalent events. If, for example, we seek a probability interval to contain the mean, μ , of a normal population given the population standard deviation, σ , the appropriate equivalent events are

$$z_{\text{lower}} < Z < z_{\text{upper}} \quad (3a)$$

and

$$Z^*_{\text{lower}} < \mu < Z^*_{\text{upper}} \quad (3b)$$

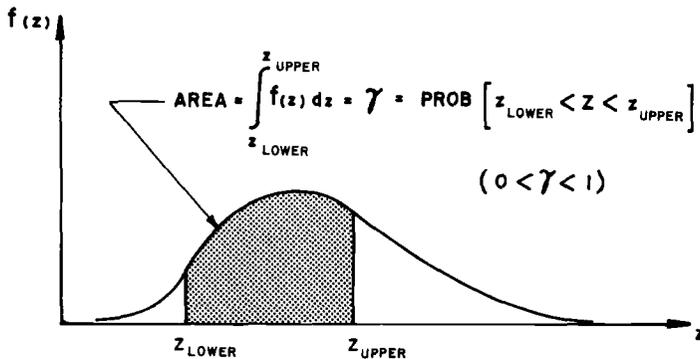


FIG. 1a—An a priori probability. The probability is γ that a single future realization of the random variable, Z , will fall within the interval $[z_{\text{lower}}, z_{\text{upper}}]$.

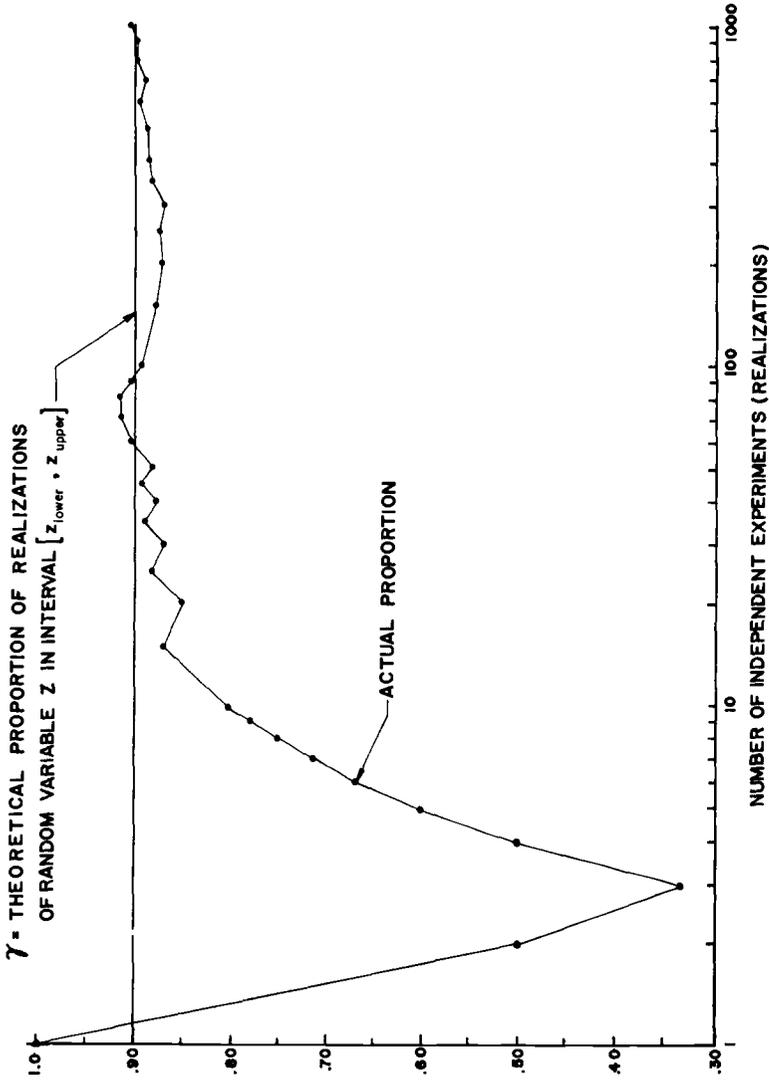


FIG. 1b—An a posteriori probability. The actual proportion of realizations that fall in the interval $[z_{lower}, z_{upper}]$ approaches γ as the number of independent experiments approaches infinity (and the deviations of the actual proportion from the theoretical proportion become smaller and smaller as n becomes larger and larger).

in which $Z = (\bar{Y} - \mu)/(\sigma/\sqrt{n})$ in Expression 3a, $Z^*_{\text{lower}} = \bar{Y} - z_{\text{upper}}\sigma/\sqrt{n}$ and $Z^*_{\text{upper}} = \bar{Y} - z_{\text{lower}}\sigma/\sqrt{n}$ in Expression 3b, and \bar{Y} (a random variable) = $\sum_{i=1}^n Y_i/n$, where Y_i is the i^{th} future (yet unknown) random observation and n is the (future) sample size.² The definition of equivalent events dictates specifically that when Expression 3a is true, then and only then is 3b true, and vice versa. The respective probabilities associated with Expressions 3a and 3b, therefore, are exactly equal, namely,

$$\text{Prob} [z_{\text{lower}} < (\bar{Y} - \mu)/(\sigma/\sqrt{n}) < z_{\text{upper}}] = \gamma \quad (4a)$$

and

$$\text{Prob} [\bar{Y} - z_{\text{upper}}\sigma/\sqrt{n} < \mu < \bar{Y} - z_{\text{lower}}\sigma/\sqrt{n}] = \gamma \quad (4b)$$

(in which z_{upper} is usually positive and z_{lower} is usually negative). The probability, γ , pertaining to application Expression 4b is established by appropriate selection of z_{lower} and z_{upper} in theory Expression 4a. Refer again to Fig. 1.

Probability Expression 4b involves a fixed (unknown) parameter and a *random interval* $[\bar{Y} - z_{\text{upper}}\sigma/\sqrt{n}, \bar{Y} - z_{\text{lower}}\sigma/\sqrt{n}]$, whereas 4a involves a fixed interval and a random variable, \bar{Y} . Given a specific future (numerical) realization of the random variable \bar{Y} , denoted \bar{y} , the quantity, $(\bar{y} - \mu)/(\sigma/\sqrt{n})$, will either lie within the interval from z_{lower} to z_{upper} or it will not, and we cannot tell which until we have conducted the appropriate experiment and observed the outcome. Nevertheless, we can assert that, in the long run, γ proportion of all possible *future* numerical values of $(\bar{y} - \mu)/(\sigma/\sqrt{n})$ will lie within the interval given in Expression 4a. In turn, using arguments based on equivalent events, we can deduce that γ proportion of all possible *future* numerical intervals $[\bar{y} - z_{\text{upper}}\sigma/\sqrt{n}, \bar{y} - z_{\text{lower}}\sigma/\sqrt{n}]$, will include the population mean, μ , even though μ is unknown. The concept of a random interval is illustrated schematically in Fig. 2. The actual proportion of the numerical intervals that indeed include the population mean, μ , may be visualized as sketched in Fig. 1b. Specifically, this proportion approaches γ in the long run (that is, as $n \rightarrow \infty$).

Probability expressions involving random intervals are usually referred to as either confidence, prediction, or tolerance expressions, depending on their use [1-3].³ Confidence expressions and their associated intervals generally pertain to the parameters of a population previously sampled, such as the mean, μ , or the standard deviation, σ , or a normal population. Prediction expressions and their associated intervals usually pertain to observations to be obtained from a specific future sample randomly drawn from a population previously sampled, whereas tolerance expressions and their associated intervals usually pertain to

²The equivalence of these events may be established in this elementary example by algebraic manipulation. However, in general, more sophisticated arguments and methodologies are needed.

³The italic numbers in brackets refer to the list of references appended to this paper.

8 STATISTICAL ANALYSIS OF FATIGUE DATA

in which Y_i and \bar{Y} are random variables and n is the sample size. The most widely used estimator for the sample standard deviation, σ , of the normal population is

$$S = \left\{ \sum_{i=1}^n (Y_i - \bar{Y})^2 / (n - 1) \right\}^{1/2} \quad (6a)$$

in which S is a random variable. For the given example data, these estimators take on the realizations \bar{y} and s , where

$$\bar{y} = \sum_{i=1}^n y_i / n = (51.4 + \dots + 51.6) / 5 = 50.10 \quad (5b)$$

and

$$s = \left\{ \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n - 1} \right\}^{1/2} = \{ [(51.4 - 50.10)^2 + \dots + (51.6 - 50.10)^2] / (5 - 1) \}^{1/2} = 1.31 \quad (6b)$$

A probability expression associated with a two-sided 100γ percent confidence interval to contain the unknown mean, μ , of a normal population may be written as

$$\text{Prob} [Z^*_{\text{lower}} < \mu < Z^*_{\text{upper}}] = \gamma \quad (7)$$

in which

$$Z^*_{\text{lower}} = \bar{Y} - t[n - 1; (1 + \gamma)/2] S / \sqrt{n}$$

$$Z^*_{\text{upper}} = \bar{Y} + t[n - 1; (1 + \gamma)/2] S / \sqrt{n}$$

and $t[n - 1; (1 + \gamma)/2]$ is the $100(1 + \gamma)/2$ percentile of the Student's t distribution, with $(n - 1)$ degrees of freedom. For any particular sample of interest, this random interval takes on the specific lower and upper limit realizations

$$z^*_{\text{lower}} = \bar{y} - t[n - 1; (1 + \gamma)/2] s / \sqrt{n}$$

and

$$z^*_{\text{upper}} = \bar{y} + t[n - 1; (1 + \gamma)/2] s / \sqrt{n}$$

Thus, for the given example data, this numerical two-sided 95 percent confidence interval for μ is bounded by

$$\begin{aligned} z^*_{\text{lower}} &= 50.10 - t[4; 0.975](1.31) / \sqrt{5} \\ &= 50.10 - 2.776(1.31) / \sqrt{5} \\ &= 50.10 - 1.63 \end{aligned}$$

and

$$z^*_{\text{upper}} = 50.10 + 1.63$$

Accordingly, the corresponding numerical two-sided 95 percent confidence interval for the unknown population mean, μ , of the normal population is [48.47, 51.73] . . . subject to the probability interpretation underlying Figs. 1 and 2. If the factor $t[n - 1; (1 + \gamma)/2]\sqrt{n}$ had been specially tabulated for the specific purposes of this calculation as $t_1(n; \gamma) = 2.776/\sqrt{5} = 1.24$, this numerical confidence interval could have been computed more conveniently as $\bar{y} \pm t_1(n; \gamma)s$.

A probability expression associated with a two-sided 100γ percent prediction interval to contain a single future observation randomly selected from a previously sampled normal population may be written as [2]

$$\text{Prob}[Z^*_{\text{lower}} < Y_{n+1} < Z^*_{\text{upper}}] = \gamma \quad (8)$$

in which

$$Z^*_{\text{lower}} = \bar{Y} - t[n - 1; (1 + \gamma)/2]S\sqrt{1 + (1/n)}$$

and

$$Z^*_{\text{upper}} = \bar{y} + t[n - 1; (1 + \gamma)/2]S\sqrt{1 + (1/n)}$$

For any particular sample of interest, this random interval takes on the specific lower and upper limit realizations

$$z^*_{\text{lower}} = \bar{y} - t[n - 1; (1 + \gamma)/2]s\sqrt{1 + (1/n)}$$

and

$$z^*_{\text{upper}} = \bar{y} + t[n - 1; (1 + \gamma)/2]s\sqrt{1 + (1/n)}$$

in which $t[n - 1; (1 + \gamma)/2]$ is the $100(1 + \gamma)/2$ percentile of the Student's t distribution with $(n - 1)$ degrees of freedom, and n is the (prior) sample size. Thus, for the given example data, this numerical two-sided 95 percent prediction interval is bounded by

$$\begin{aligned} z^*_{\text{lower}} &= 50.10 - t[4; 0.975](1.31)\sqrt{1.2} \\ &= 50.10 - 2.766(1.31)\sqrt{1.2} \\ &= 50.10 - 3.98 \end{aligned}$$

and

$$z^*_{\text{upper}} = 50.10 + 3.98$$

Accordingly, the corresponding numerical two-sided 95 percent prediction interval of a single future observation $Y(Y_{n+1})$ randomly selected from the previously sampled normal population is [46.12, 54.08] . . . subject to the probability interpretation underlying Figs. 1 and 2. If the factor

$$t[n - 1; (1 + \gamma)/2]\sqrt{1 + (1/n)}$$

had been specially tabulated for the specific purposes of this calculation as $t_2(n; \gamma) = 2.776\sqrt{1.2} = 3.04$, this numerical prediction interval could have been computed more conveniently as $\bar{y} \pm t_2(n; \gamma)s$.

The prediction interval associated with probability Expression 8 is perhaps more easily understood than the analogous confidence interval associated with Expression 7, because we can always make another observation (at least in concept) to see whether, indeed, it falls within the numerical interval—yes or no.⁴

A probability expression associated with a two-sided 100 γ percent prediction interval to contain all of k future observations, randomly selected from a previously sampled normal population, may be written as [2]

$$\text{Prob} [Z^*_{\text{lower}} < Y_{n+1} \cap Y_{n+2} \cap Y_{n+3} \cdots \cap Y_{n+k} < Z^*_{\text{upper}}] = \gamma \quad (8)$$

in which \cap (intersection) implies *all*, $z^*_{\text{lower}} = \bar{Y} - t_3(n; k; \gamma)S$, $Z^*_{\text{upper}} = \bar{Y} + t_3(n; k; \gamma)S$, and $t_3(n; k; \gamma)$ is a prediction interval factor conveniently tabulated by Hahn [1]. For example, when $n = 5$, $k = 2$, and $\gamma = 0.95$, then

$$t_3(n; k; \gamma) = t_3(5; 2; 0.95) = 3.70$$

Thus, this random prediction interval is given by $\bar{Y} \pm 3.70S$. For the given example data, the corresponding numerical two-sided 95 percent prediction interval to contain both of two future observations randomly selected from the normal population previously sampled is $\bar{y} \pm 3.70s = 50.10 \pm 4.85 = [45.25, 54.95]$. . . subject to the probability interpretation underlying Figs. 1 and 2.

A probability expression associated with a two-sided 100 γ percent tolerance interval which contains at least β proportion of all possible future observations from a previously sampled normal population may be written as

$$\text{Prob} \left[\int_{Z^*_{\text{lower}}}^{Z^*_{\text{upper}}} f_{\text{normal}}(u) du > \beta \right] = \gamma \quad (9)$$

in which $f_{\text{normal}}(u)$ is the normal probability density function, $Z^*_{\text{lower}} = \bar{Y} - t_4(n; \gamma; \beta)S$, $Z^*_{\text{upper}} = \bar{Y} + t_4(n; \gamma; \beta)S$, and $t_4(n; \gamma; \beta)$ is a tolerance limit factor widely tabulated in the statistical literature (refer, for example, to Natrella [4]). Specifically, when $n = 5$, $\gamma = 0.95$, and $\beta = 0.90$, then $t_4(5, 0.95, 0.90) = 4.28$. Thus, a random interval to contain at least 90 percent ($\beta = 0.90$) of all future observations from the previously sampled normal population with 0.95 probability ($\gamma = 0.95$) is $\bar{Y} \pm 4.28S$. For the given ex-

⁴Specifically, the replicated experiment consists of selecting a random sample of size n , computing the prediction interval, and then selecting another independent observation and observing whether it indeed falls within the computed prediction interval; this entire process is then repeated indefinitely to obtain plots similar to those in Figs. 1 and 2.

ample data, the corresponding numerical two-sided tolerance interval which contains at least 90 percent of all future observations from the previously sampled normal population with probability 0.95 is $\bar{y} \pm 4.28s = 50.10 \pm 5.61 = [44.49, 55.71]$. . . subject to the probability interpretation underlying Figs. 1 and 2.

Historically, statisticians have used the phrase “with 95 percent confidence” in place of the phrase “with 0.95 probability” when referring to a specific numerical interval (for example, the two-sided tolerance interval [44.49, 55.71]). This terminology is intended to avoid repeated use of the qualification . . . subject to the probability interpretation underlying Figs. 1 and 2. Thus, the two-sided tolerance interval expression is commonly stated verbally as “We may say with 95 percent confidence that at least 90 percent of the sampled normal population will exhibit values between 44.49 and 55.71.”

It is also relatively common to use the term “confidence” when referring to an interval containing a percentile of a distribution (rather than a parameter). For example, it might be said that “we are 95 percent confident that the tenth percentile of the sampled normal population lies within the interval $[Z^*_{\text{lower}}, Z^*_{\text{upper}}]$.” The associated probability expression may be interpreted as a tolerance limit expression, as is evident in the next section.

Figure 3 presents a plot of the example data and a sketch of the estimated normal probability density function along with diagrams for comparative purposes of the two-sided 95 percent intervals computed for the respective numerical examples. Proschan [5] provides factors to compute additional probability intervals that may be of interest to certain readers.

One-Sided Lower Tolerance Limits

I deal specifically in this paper with one-sided lower tolerance limits of the verbalized form: “We may say with γ percent confidence that (at least) β proportion of the sampled population lies above Z_{lower} .” In the section on Distribution Assumed Known I summarize exact and approximate one-sided lower tolerance limit calculations based on known distributions, namely, the two-parameter log-normal and Weibull distributions, because of their extensive use in fatigue. In the section on Life Distribution Not Assumed Known in Analysis, I discuss distribution-free one-sided lower tolerance limits, because it is indeed naive to believe that the actual fatigue life distribution is either exactly log-normal or exactly Weibull.

Test Conduct

All statistical analyses discussed herein pertain specifically to a completely randomized test program [6,7]; that is, it is implicitly assumed that all specimens are homogeneous in material and preparation and that all test con-

soring literally pertains to runouts at "long life" in a fatigue context.⁵ Type II censoring is more academic, pertaining primarily to "accelerated testing" situations where the entire test program is terminated as soon as the j^{th} failure occurs (assuming all specimens are being tested concurrently). Exact one-sided lower tolerance limit analyses are available in the statistical literature for the two-parameter log-normal and Weibull distributions given Type II censoring, but only approximate solutions are available given Type I censoring.

Regardless of the given type of censoring, the life distribution assumed, or the exactness of the analyses, the analytical procedure for the one-sided lower tolerance limits considered herein may be summarized as follows: (a) assume the distribution, (b) estimate its parameters, (c) plot the estimated distribution on probability paper (Fig. 4), (d) plot the corresponding one-sided lower 100 γ percent confidence band (Fig. 4),⁶ and (e) obtain the desired tolerance limit by finding the intersection of the relevant population proportion $(1 - \beta)$ and

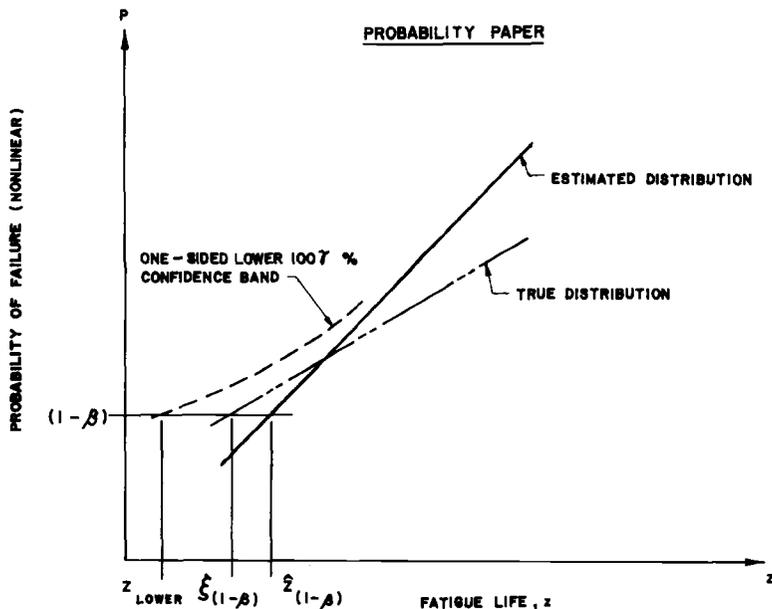


FIG. 4—A schematic drawing that defines the one-sided lower tolerance limits of interest herein, namely, one-sided lower 100 γ percent confidence limits pertaining to the $(1 - \beta)^{\text{th}}$ percentile of the assumed fatigue life distribution.

⁵Individual fatigue tests are also "suspended" after shorter durations (but prior to failure) on various occasions. Although maximum likelihood estimation techniques include suspended data also, the concept of the replicated experiment in the context of Fig. 2 is not strictly valid.

⁶The method of constructing one-sided lower confidence bands depends on whether the random interval pertains to a fixed value of z or a fixed value of $(1 - \beta)$ in the conceptually replicated experiments.

the corresponding one-sided lower 100 γ percent confidence band (Fig. 4). The associated probability expression is

$$\text{Prob}[z_{\text{lower}} < \xi_{1-\beta}] = \gamma \quad (11)$$

in which $\xi_{1-\beta}$ is the $(1 - \beta)^{\text{th}}$ percentile of the assumed distribution, and Z_{lower} is the (random) one-sided lower 100 γ percent confidence limit pertaining to the $(1 - \beta)^{\text{th}}$ percentile of the assumed distribution.

The only issues remaining pertain to the specific methods of estimating the distribution parameters and of computing the corresponding one-sided lower 100 γ percent confidence band.

Two-Parameter Weibull Distribution

Type II Censoring—I have illustrated the computation of exact one-sided lower tolerance limits for the two-parameter *Weibull* distribution given Type II censoring in a previous paper [8]. The distribution parameters are estimated using the best linear unbiased (BLU) estimation methodology, based on coefficients tabulated by White [9], and the associated one-sided lower 100 γ percent confidence limits for certain specific population percentiles are computed using special factors tabulated by Mann and Fertig [10]. These special one-sided lower confidence limit factors were established using a digital computer simulation technique in which appropriate Type II censored data were repeatedly generated and analyzed, leading ultimately to a "histogram" of observed results which closely approximates the actual sampling distribution of interest. The actual sampling distribution depends in theory upon $\xi_{1-\beta}$, but not upon the unknown parameters of the Weibull distribution. Thus, Mann and Fertig were able to satisfy probability Expression 11 by tabulating a special tolerance limit factor (which pertains to both $\xi_{1-\beta}$ and the appropriate percentile, $1 - \gamma$, of the sampling distribution of interest).

Numerical Example [8]—The following fatigue life data, randomly selected from a two-parameter Weibull population, are given:

170 000 cycles
210 000
183 000
144 000
256 000
256 000 suspended

First, for convenience, rewrite the ordered data in terms of thousands of cycles, that is, 144, 170, 183, 210, 256 (256 suspended). Next, note that if the observed fatigue life data follow the two-parameter Weibull distribution

$$F(z) = 1 - e^{-(z/\theta_1)^{\theta_2}} \quad (12)$$

then the natural logarithms of the data (denoted z^*_i in Ref 8) follow the smallest extreme value distribution

$$F(z^*) = 1 - e^{-e\left(\frac{z^* - \theta^*_1}{\theta^*_2}\right)} \tag{13}$$

We may estimate θ^*_1 and θ^*_2 in Eq 13 using the expressions [8]

$$\hat{\theta}^*_1 = \Sigma a_i z^*_i$$

and (14)

$$\hat{\theta}^*_2 = \Sigma b_i z^*_i$$

in which the a_i and b_i coefficients are given by White [9]. Refer to Table 1. Next, we may use these estimates and certain other coefficients given by White [9] in an intermediate computation to obtain best linear variant parameter estimates, \hat{a}^* and \hat{b}^* . For the given example data, the appropriate coefficients are 0.0105329 and 1.1861065, and

$$\hat{a}^* = \hat{\theta}^*_1 - \frac{0.0105329}{1.1861065} \hat{\theta}^*_2 = 5.452721$$

and (15a)

$$\hat{b}^* = \frac{\hat{\theta}^*_2}{1.1861065} = 0.220116$$

Finally, using the special tolerance limit factor tabulated by Mann and Fertig [10], we may compute the one-sided lower 95 percent confidence limit for the tenth percentile of the sampled two-parameter Weibull fatigue life distribution, that is

$$\begin{aligned} z^*_{\text{lower}} &= \hat{a}^* - \hat{b}^*(M \text{ and } F \text{ factor}) \\ &= 5.452721 - 0.220116(6.73) = 3.971 \end{aligned}$$

TABLE 1—Computation of parameter estimates for the Weibull distribution [8] ($z^*_i = \log_e z_i$).

z_i	z^*_i	a_i	b_i	$a_i z^*_i$	$b_i z^*_i$
144	4.96981	0.0057312	-0.2015427	0.028482	-1.001629
170	5.13580	0.0465760	-0.1972715	0.239205	-1.013147
183	5.20949	0.1002434	-0.1536128	0.522217	-0.800244
210	5.34711	0.1722854	-0.0645867	0.921229	-0.345352
256	5.54518	0.6751639	0.6170138	3.743905	3.421453
				5.455039	0.261081

Taking the antilog, we may say with 95 percent confidence that 90 percent of the sampled population lies above 53 000 cycles. Refer to Fig. 5.

Two-Parameter Log-Normal Distribution

Type II Censoring—I recently wrote a corresponding paper on the computation of exact one-sided lower tolerance limits for the two-parameter *log-normal* distribution with Type II censoring [11]. The methodology is identical to that in Ref 8 for the two-parameter Weibull distribution, only the coefficients change (and intermediate Calculation 15a is not required). The coefficients for best linear unbiased estimation with Type II censoring are given by Sarhan and Greenberg [12], and the associated special one-sided lower 100 γ percent confidence limit factors are tabulated by Nelson and Schmee in Ref 13. Refer to Table 2 for the estimation of θ^*_1 and θ^*_2 . The associated one-sided lower 95 percent confidence limit for the tenth percentile of the sampled two-parameter log-normal fatigue life distribution is

$$z^*_{\text{lower}} = \hat{\theta}^*_1 - \hat{\theta}^*_2(N \text{ and } S \text{ factor})$$

$$= 5.322247 - 0.303682(3.083) = 4.386 \tag{16}$$

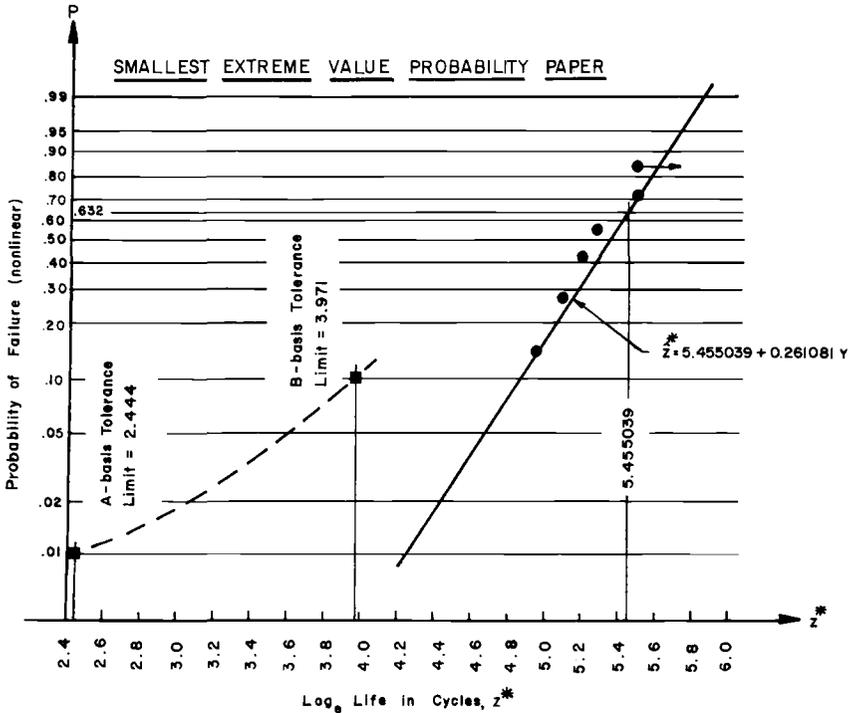


FIG. 5—Exact one-sided lower tolerance limit analysis for the text example data—assuming Type II censoring and a Weibull life distribution [8]. Refer to Table 1.

TABLE 2—Computation of parameter estimates for the log-normal distribution [11] ($z^*_i = \log_e z_i$).

z_i	z^*_i	a_i	b_i	$a_i z^*_i$	$b_i z^*_i$
144	4.96981	0.1183	-0.4097	0.587929	-2.036131
170	5.13580	0.1510	-0.1685	0.775058	-0.865382
183	5.20949	0.1680	-0.0406	0.875194	-0.211505
210	5.34711	0.1828	+0.0740	0.977452	+0.395686
256	5.54518	0.3799	+0.5448	2.106614	+3.021014
				5.322247	0.303682

Taking the antilog, we may say with 95 percent confidence that 90 percent of the sampled log-normal population lies above 80 300 cycles. Refer to Fig. 6.

Two-Parameter Weibull Distribution, Type I Censoring, and Two-Parameter Log-Normal Distribution, Type I Censoring

Suppose that the sixth specimen in these example data had actually endured 500 000 cycles before the test was terminated, that is, the sixth specimen was a runout at 500 000 cycles. Then Type I censoring obtains, and

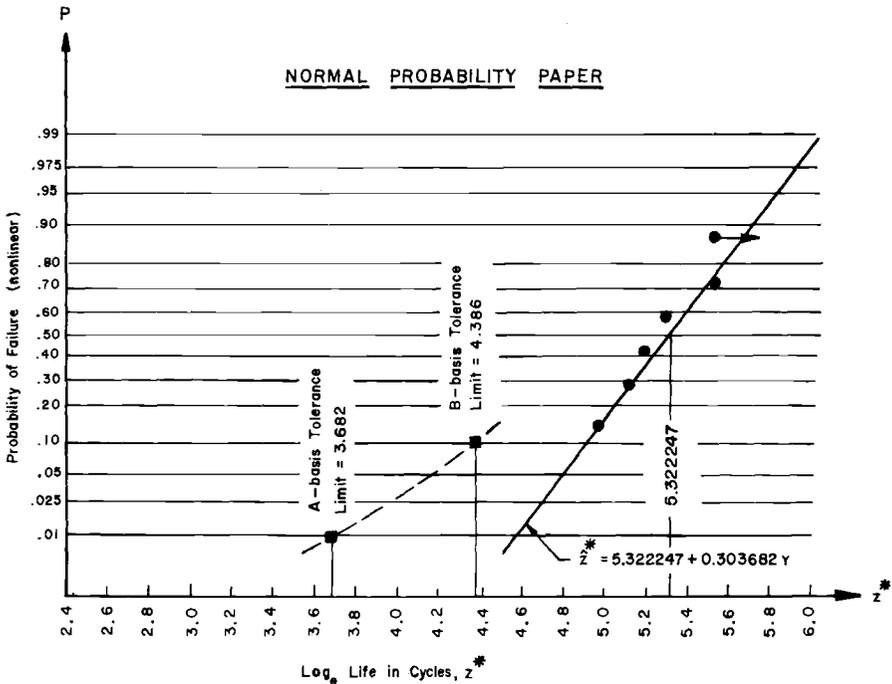


FIG. 6—Exact one-sided lower tolerance limit analysis for the text example data—assuming Type II censoring and a log-normal life distribution [11]. Refer to Table 2.

the previous analyses are not strictly valid. Maximum likelihood (ML) computer programs are available to analyze Type I censoring [14,15] but (1) the estimates of the parameters are biased, (2) the associated one-sided lower confidence limits are approximate (precise only for *large* samples), and (3) the approximate (asymptotic) confidence limits may differ depending on whether the distribution function is written using $y = (z - \theta_1)/\theta_2$, $y = \theta_2(z - \theta_1)$, $y = \theta_1 + \theta_2z$, or $y = \theta_1 + z/\theta_2$. There are several ways to correct for the bias of the estimates, and there are also different techniques to compute the associated approximate (asymptotic) one-sided lower confidence limits. Thus, there are numerous alternative analyses available—so many that a relatively comprehensive summary has not yet been attempted even in the statistical literature.

Table 3 compares one-sided lower tolerance limits computed using four different ML-based analyses for the case where the test for the sixth specimen was suspended at 256 000 cycles (Type II censoring). In general, the approximate (asymptotic) ML-based tolerance limits can differ quite markedly from the exact BLU tolerance limits for small sample sizes, depending in part on which alternative procedures are arbitrarily used in ML-based analyses. Moreover, the respective results obtained by assuming a log-normal versus a Weibull distribution can differ markedly, particularly when $(1 - \beta)$ is small, say 0.10 or less. Thus, intelligent use of such tolerance limits involves some experience and judgment regarding their sensitivity to various analytical procedures and assumptions. The more comparative analyses one generates for the given set of data, the broader perspective one has to make the necessary engineering decisions.

Discriminating Between the Two-Parameter Weibull and Log-Normal Distributions

Because the two-parameter Weibull and log-normal distributions usually differ so markedly at small percentiles ($P = 0.01$ and below), especially for small samples, a brief discussion of a statistical procedure for discerning between these two distributions may be helpful to some readers.

Dumoncaux and Antle [16] provide critical values for the ratio of maximized likelihoods to discriminate between these two distributions. First, both distributions are fitted to the data using maximum likelihood analyses,⁷ and then the respective maximum likelihood values are used to form a ratio, which is in turn compared with tabulated percentiles of the corresponding sampling distribution that were established by digital computer simulation. Refer to Table 4. Generally, it is desirable to keep the α (Type 1) error below 0.10 and while attaining a statistical power of at least 0.80. Preferably, α is at least 0.05, and the power is at least 0.90. Observe that given a complete sample (no cen-

⁷FORTTRAN listings of the appropriate computer programs may still be available by writing to Antle.

TABLE 3—One-sided lower A-basis and B-basis tolerance limits for the text example data (Type II censoring maximum likelihood analysis).^a

	ML-Based Analysis						BLU-Based Analysis
	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	
	Two-Parameter Weibull Distribution						
B-basis	4.48	3.14	3.00	1.16	4.09	4.39	3.97
A-basis	3.60	0.79	0.68	0.26	2.80	3.40	2.44
	Two-Parameter Log-Normal Distribution						
B-basis	4.70	4.26	4.21	2.15	4.45	4.65	4.39
A-basis	4.31	3.39	3.39	1.71	3.86	4.19	3.68

^aSets *a*, *b*, *c*, and *d* pertain to elliptical *joint* asymptotic confidence regions for $y = (z - \theta_1)/\theta_2$, $y = \theta_2(z - \theta_1)$; $y = \theta_1 + \theta_2z$; and $y = \theta_1 + z/\theta_2$, respectively; Set *e* pertains to the *joint* asymptotic region defined by Bartlett's likelihood ratio procedure (which is independent of how the linear *y* versus *z* relationship is written); and Set *f* pertains to the standard asymptotic probability interval defined by Lawless' likelihood ratio procedure.

TABLE 4—Selected critical values of the ratio of maximized likelihoods (RML) [16].

<i>n</i>	$\alpha = 0.10$		$\alpha = 0.05$	
	$(RML)_c^{1/n}$	Power	$(RML)_c^{1/n}$	Power
<i>(a) Critical Values of $(RML)^{1/n}$ and Power of the Test for Log-Normal Hypothesis Versus Weibull (Alternative) Hypothesis</i>				
20	1.038	0.61	1.082	0.48
30	1.020	0.75	1.044	0.63
40	1.007	0.85	1.028	0.76
50	0.998	0.91	1.014	0.83
<i>(b) Critical Values of $(RML)^{1/n}$ and Power of the Test for Weibull Hypothesis Versus Log-Normal (Alternative) Hypothesis</i>				
20	1.041	0.57	1.067	0.43
30	1.019	0.74	1.041	0.62
40	1.005	0.85	1.026	0.75
50	0.995	0.91	1.016	0.82

soring), a “minimum” sample of about 35 is required to discriminate adequately between the two distributions. Clearly, for a sample size of 10 to 20 our ability to discriminate between the Weibull and log-normal distributions is *not* good. The discrimination problem is even more severe for censored samples.

Life Distribution Not Assumed Known in Analysis

All replicated measurements involve variability—due in part to the variability of the measurement process itself and in part to the “intrinsic” variability of

the measurand. Even in the simplest possible situation, a measurement, M , may be analytically partitioned (explained) as

$$M = X + \delta$$

in which X is a random variable with mean, μ_X , and variance, σ_X^2 , where σ_X^2 is the "intrinsic" variability of the measurand under perfect measurement or test conduct conditions and δ is a random variable with mean, μ_δ , and variance, σ_δ^2 , where σ_δ^2 is the additional (spurious) variability associated with imperfect measurement or test conduct conditions. If $\mu_\delta = 0$, the measurement is unbiased. But whether $\mu_\delta = 0$ or not, the distribution of M depends on the distributions of X and δ . Specifically, both distributions must be known to state (assert, establish) the distribution of M . In fatigue applications, the additional variability associated with material processing, manufacturing, and service loading and environments must be considered and evaluated. Thus, an analytical assumption that the log-normal or the Weibull distribution accurately describes the fatigue life of any real device always lacks credibility. I elaborate on this point in Ref 17.

Ideally, we would like to establish a lower one-sided tolerance limit which does not depend on the exact form of the underlying life distribution. The standard nonparametric lower tolerance limit meets this criterion, but it requires larger sample sizes than are practical in fatigue applications.

Standard Nonparametric One-Sided Lower Tolerance Limit

Given a life distribution with a continuous probability density function (PDF), and a randomly selected ordered sample of size n , Wilks [18] showed that

$$\text{Prob} [Z_1 \leq \xi_{1-\beta}] = \gamma = 1 - \beta^n \quad (17)$$

in which the random variable, Z_1 , is the smallest observation of an ordered life sample of size n , and $\xi_{1-\beta}$ is the $100(1 - \beta)^{\text{th}}$ percentile of the life distribution. Accordingly, the numerical realization of Z_1 , denoted z_1 , can be used to establish a nonparametric one-sided lower tolerance limit.

Sample size n should be chosen so that some appropriate value of γ is obtained in Eq 17, given some specified value of β . For example, a sample size of approximately 30 is required to establish a B-basis tolerance limit, whereas a sample of approximately 300 is required to establish an A-basis tolerance limit.⁸

Modified One-Sided Lower Nonparametric Tolerance Limit

If it appears reasonable to assume that the *slope* of PDF of the continuous life distribution is *strictly increasing* in the interval $0 < z < \xi_\alpha$, where ξ_α per-

⁸A-basis (one-sided lower) tolerance limit: $\gamma = 0.95$, $\beta = 0.99$; B-basis (one-sided lower) tolerance limit: $\gamma = 0.95$; $\beta = 0.90$.

tains to the $100\alpha^{\text{th}}$ percentile of the life distribution, then it can be shown that (mathematical details omitted)

$$\text{Prob} [Z_1/c < \xi_{1-\beta}] > 1 - [1 - (1 - \beta)c^2]^n = \gamma \quad (18)$$

where $c > 1$ and $(1 - \beta)c^2 < \alpha$. In this case, the minimum sample size required to attain a prescribed value of γ , given the desired value of β , depends on the minimum value of α that appears reasonable. For purposes of perspective, $f'(z)$ is strictly increasing up to α equal to about 0.16 for a normal distribution, 0.21 for the logistic distribution, 0.07 for the largest extreme value distribution (skewed to the right), and 0.32 for the smallest extreme value distribution (skewed to the left). Given the Weibull distribution in Eq 12, $f'(z)$ is strictly increasing only for $\theta_2 > 2$. Specifically, for $\theta_2 = 2.5$, $\alpha = 0.07$; for $\theta_2 = 3.0$, $\alpha = 0.11$; for $\theta_2 = 4.0$, $\alpha = 0.16$; for $\theta_2 = 5.0$, $\alpha = 0.20$; and for $\theta_2 = 10.0$, $\alpha = 0.26$.

Table 5 shows that, if it were reasonable to assume that $f'(z)$ is strictly increasing up to about the tenth percentile, 30 specimens could be used to obtain both an A-basis and a B-basis tolerance limit. The former would be approximately $z_1/3.13$, whereas the latter would be $z_1/1.00$. It is apparent in Table 5 that a sample size of about 20 is statistically acceptable if it appears reasonable to assume $f'(z)$ is strictly increasing up to the fourteenth percentile. But even this sample size is sufficiently large to prevent Eq 18 from finding extensive application in fatigue analyses.

Numerical Example—Suppose a sample of 22 fatigue test specimens ex-

TABLE 5—Modified one-sided nonparametric tolerance tables:
minimum sample size n versus α_{\min} for $\gamma = 0.95$ and the related C values
when $(1 - \beta)C^2 < \alpha$.

γ	α_{\min}	n_{\min}	B-Basis		A-Basis
			$\beta = 0.90$	$\beta = 0.95$	$\beta = 0.99$
0.95	0.02	149			1.41
	0.03	99			1.73
	0.04	74			1.99
	0.05	59		1.00	2.23
	0.06	49		1.09	2.44
	0.07	42		1.17	2.62
	0.08	36		1.26	2.83
	0.09	32		1.34	2.99
	0.10	29	1.00	1.40	3.13
	0.11	26	1.04	1.48	3.30
	0.12	24	1.08	1.53	3.43
	0.13	22	1.13	1.60	3.57
	0.14	20	1.18	1.67	3.73
	0.15	19	1.21	1.71	3.82
	0.20	14	1.39	1.96	4.39
	0.25	11	1.54	2.18	4.88
	0.30	9	1.68	2.38	5.32

hibited a minimum life, z_1 , equal to 121 000 cycles. Suppose further that, assuming a Weibull distribution, the best linear unbiased and maximum likelihood estimates for θ_2 are, respectively, 4.15 and 4.34. The assumption that the data follow a two-parameter Weibull distribution is roughly equivalent to the assumption that $f'(z)$ is strictly increasing up to about the seventeenth percentile. Entering Table 5, we see that for $n = 22$, $\alpha_{\min} = 0.13$, which is less than 0.17, and thus we, in turn, obtain the factor c for a B-basis tolerance limit ($c = 1.13$) and compute the desired B-basis tolerance limit as $121\ 000/1.13 = 107\ 000$ cycles. The corresponding A-basis tolerance limit is $121\ 000/3.57 = 33\ 900$ cycles.

Conclusion

Tolerance limit analyses involve the fundamental problem that spurious variability damages the credibility of quantitative (predictive) analyses (whereas this variability need not damage the credibility of comparative analyses based on appropriately planned experimental programs). Nevertheless, if predictive analyses are required (mandatory), one-sided lower tolerance limits are appropriate in situations pertaining to material specifications.

References

- [1] Hahn, G. J., *Industrial Engineering*, Vol. 2, Dec. 1970, pp. 45-48.
- [2] Hahn, G. J. and Nelson, W. B., "A Survey of Prediction Intervals and Their Applications," General Electric Co. Corporate Research and Development Technical Information Series Report No. 72CRD027, General Electric Co., Schenectady, N.Y., Jan. 1972.
- [3] Hahn, G. J., "Some Things Engineers Should Know about Statistics," General Electric Co. Corporate Research and Development Technical Information Series Report No. 73CRD291, General Electric Co., Schenectady, N.Y., Nov. 1973.
- [4] Natrella, M. G., *Experimental Statistics, Handbook 91*, U.S. Department of Commerce, National Bureau of Standards, U.S. Government Printing Office, Washington, D.C., 1963.
- [5] Proschan, F., *Journal of the American Statistical Association*, Vol. 48, 1953, pp. 550-564.
- [6] Little, R. E. and Jebe, E. H., *Manual on Statistical Planning and Analysis for Fatigue Experiments, STP 588*, American Society for Testing and Materials, Philadelphia, 1975.
- [7] Little, R. E. and Jebe, E. H., *Statistical Design of Fatigue Experiments*, Applied Science Publishers, London, England, 1975.
- [8] Little, R. E., *Journal of Testing and Evaluation*, Vol. 5, No. 4, 1977, pp. 303-308.
- [9] White, J. S., *Industrial Mathematics*, Vol. 14, Part 1, 1964, pp. 21-60.
- [10] Mann, N. R. and Fertig, K. W., *Technometrics*, Vol. 15, No. 1, Feb. 1973, pp. 87-101.
- [11] Little, R. E., *ASTM Journal of Testing and Evaluation*, Vol. 8, No. 2, 1980, pp. 80-84.
- [12] Sarhan, A. E. and Greenberg, B. G., *Contributions to Order Statistics*, Wiley, New York, 1962.
- [13] Nelson, W. and Schmee, J., *Technometrics*, Vol. 21, No. 1, 1979, pp. 43-45.
- [14] Nelson, W. and Hendrickson, R., "1972 User Manual for STATPAC—A General Purpose Program for Data Analysis and for Fitting Statistical Models to Data," General Electric Co. Corporate Research and Development Technical Information Series Report No. 72GEN009, General Electric Co., Schenectady, N.Y., May 1972.
- [15] Nelson, W. B., Hendrickson, R., Phillips, M. C., and Thumhart, L., "STATPAC Simplified—A Short Introduction to How to Run STATPAC, A General Statistical Package

for Data Analysis," General Electric Co. Corporate Research and Development Technical Information Series Report No. 73CRD046, General Electric Co., Schnectady, N.Y., July 1973.

- [16] Dumonceaux, R. and Antle, C. E., *Technometrics*, Vol. 15, No. 4, Nov. 1973, pp. 923-926.
- [17] Little, R. E., *ASTM Standardization News*, Vol. 8, No. 2, Feb. 1980, pp. 23-25.
- [18] Wilks, S. S., *Mathematical Statistics*. Wiley, New York, 1962.

Statistical Design and Analysis of an Interlaboratory Program on the Fatigue Properties of Welded Joints in Structural Steels

REFERENCE: Haibach, E., Olivier, R., and Rinaldi, F., "Statistical Design and Analysis of an Interlaboratory Program on the Fatigue Properties of Welded Joints in Structural Steels," *Statistical Analysis of Fatigue Data, ASTM STP 744*, R. E. Little and J. C. Ekvall, Eds., American Society for Testing and Materials, 1981, pp. 24-54.

ABSTRACT: The constant-amplitude fatigue behavior of welded joints in two types of normalized structural high-strength steel has been studied in an interlaboratory program within the European community. The statistical design and analysis of a part of that program is described. This part was aimed at establishing complete $S-N$ curves for three types of fillet welded joints with reference to a comprehensive statistical test plan. The test plan closely linked the activities of six laboratories involved in testing and three welding institutes fabricating the specimens under specified conditions, and it organized the repartition of the specimens to the stress levels to be applied and to the laboratories. Some restrictions, however, were imposed on the test plan due to limitations in test load capacity in some of the laboratories and to limitations in time and costs.

The results were evaluated according to the concept followed in planning, using various methods of analysis that were outlined and compared in treating the 753 test results available and in deducing characteristic figures of the fatigue strength at $2 \cdot 10^6$ cycles. The assumption of a "uniform" slope of $S-N$ curves for welded joints in structural steel proved to be reasonable. Moreover, it was possible to analyze the additional variability of the results caused by sharing the tests at each stress level among several laboratories or caused by fabricating equal portions of the specimens in three welding institutes.

KEY WORDS: fatigue tests, (complete) $S-N$ curves, welded joints, structural steel, statistical test plan, (comparative) statistical analysis, laboratory effects, welding effects, material effects, fatigue.

In order to evaluate the fatigue properties of normalized fine-grain higher strength structural steels in the welded condition, an interlaboratory program, sponsored by the European Coal and Steel Community, was carried out by

¹Director and research fellow, respectively, Fraunhofer-Institut für Betriebsfestigkeit (LBF), Darmstadt, Federal Republic of Germany.

²Head of research laboratory, Dalmine SpA, Laboratori di Ricerca, Bergamo, Italy.

seven laboratories in five countries of the European community in the period from July 1968 to December 1976. The participating laboratories were the Centre de Recherches Metallurgiques (CRM), Liege, Belgium; the Institute de Recherches de la Siderurgie Française (IRSID), St. Germain-en-Laye, France; the Fraunhofer-Institut für Betriebsfestigkeit (LBF), Darmstadt, and the Max-Planck-Institut für Eisenforschung (MPI), Düsseldorf, Germany; the Dalmine SpA, Laboratori di Ricerca, Dalmine, Bergamo, and the Acciaierie e Ferriere Lombarde Falck, Centro Ricerche e Controlli, Milano, Italy; and the Technische Hogeschool, Stevin-Laboratorium, Delft, the Netherlands. Three welding institutes were engaged in fabricating the specimens: the Centre de Recherches Metallurgiques (CRM), Liege; the Institut de Soudure (IFS), Paris; and the Instituto Italiano della Saldatura (IIS), Genova.

To ensure close cooperation among the participating laboratories and welding institutes, a working group, responsible for the detailed planning and for the realization of the test program, was constituted. Members of the working group for all or part of the contract period were E. Haibach (chairman), J. de Back, G. Bollani, J. M. Diez, H. P. Lieurade, R. Olivier, P. Rabbe, F. Rinaldi, R. V. Salkin, and P. Simon.

The results of that program and the particulars of its organization have been published in detail elsewhere [1,2].³ The present paper describes the statistical design and analysis of a main part of that interlaboratory program. It refers to $S-N$ tests for three stress ratios on three types of welded specimens in two types of steel, and these tests were shared among six laboratories. In statistical terms this is an example of an incomplete block-designed experiment. (Additional test series, not dealt with in this paper, were concerned with similar tests on notched specimens [1,2], low-cycle fatigue tests [3], tests on larger welded sections, and crack propagation tests [2]; the latter two types of test were contributed by the Stevin Laboratorium.)

Starting Point

The starting point of the program was characterized by the following situation:

1. The allowable stresses of fatigue-loaded welded joints, as given by the various codes, differed significantly even when the comparison was restricted to rather simple and well-defined types of joint [4,5].
2. Among the prevalent codes there was none that distinctly gave allowable stresses of welded joints fabricated from the newer types of fine-grain higher strength structural steels, according to Euronorm 113 [6].

This situation turned out to be unsatisfactory from a technical and economic point of view as well. In either case, a major reason was supposed

³The italic numbers in brackets refer to the list of references appended to this paper.

to be a lack of reliable fatigue data. For welded joints in higher strength materials, the available number of experimental data might have been thought to be too small to allow a new code to be set up. For welded joints in usual materials, a reanalysis of literature data revealed a wide range of scatter associated with the fatigue-strength figures reported for nominally comparable types of joints. However, it could not be determined by subsequent studies why the fatigue-strength values observed in comparable test series by different laboratories resulted in stress figures that differed by a ratio of as much as 1:3. The question remained whether the scatter was due to particular material or welding conditions, to laboratory effects, or to the method of evaluation applied to the test results. As a consequence, the differing assessment of the allowable stresses in design codes could be understood to be essentially dependent on the particular sample from the literature data that had been considered.

In order to prevent the mentioned difficulties from also being associated with the results from the intended investigation, the existing contacts among laboratories in different countries of the European community suggested setting up an interlaboratory program on a statistical basis broad enough to obtain reliable results and to allow general conclusions. Moreover, from an appropriate design of such an interlaboratory program one could expect to find some explanation of the scatter observed in the literature data.

Test Program

Two types of higher strength structural steel were selected:

- (a) a structural steel Fe E 355, in accordance with Euronorm 25, and
- (b) a vanadium-alloyed fine-grain structural steel, Fe E 460, corresponding with Euronorm 113.

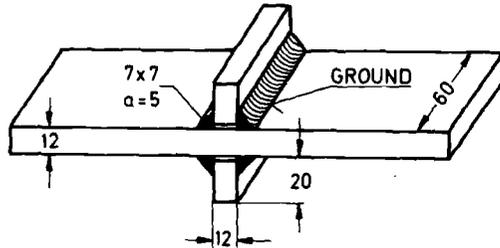
The plate materials, in 12-mm thickness, were delivered in the normalized condition, as rolled, and without any further treatment. The tolerances in plate thickness of Euronorm 29 were accepted. The chemical and mechanical properties were found to meet the standards; the actual mechanical properties are given in Table 1.

Two non-load-carrying fillet types of joint (K2 type) and a cruciform load-carrying fillet type of joint (K4 type) were tested (Fig. 1). The "K2 flat" specimens were fabricated by welding them in the flat position and by subsequently grinding the weld toes in order to provide a favorable weld profile. The fillets of the "K2 vertical" specimens were made by welding in the vertical up position, resulting in a less favorable weld profile. The K2 specimens normally fail by fatigue cracking which start at the toe of the fillet and propagates through the plate material. The K4-type specimens usually fail by cracking which starts at the root of the weld and propagates through the weld metal.

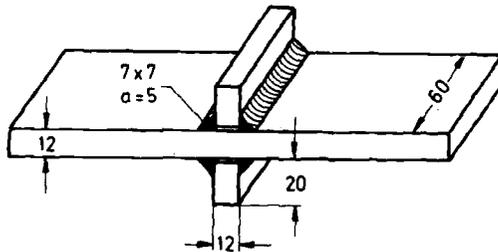
TABLE 1—Actual mechanical properties of materials.

Type of Steel	Plate Thickness, mm	Yield Strength, N/mm ²	Ultimate Strength, N/mm ²	Elongation, %	KV (−20° C), J/cm ²
Fe E 355, normalized	12	420	570	28	55
Fe E 460, normalized	12	515	660	24	54

K2 FLAT



K2 VERTICAL



K4 FLAT

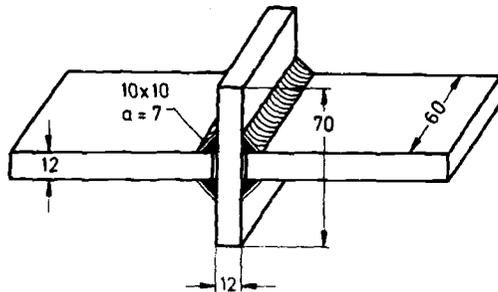


FIG. 1—Types of joint tested.

Figure 2 presents a survey of the test series provided for the part of the program considered here. In both types of material the three types of specimens were tested to establish the $S-N$ curves for completely reversed loading (stress ratio $R = -1$), for zero-tension loading ($R = 0$), and for fluctuating tension

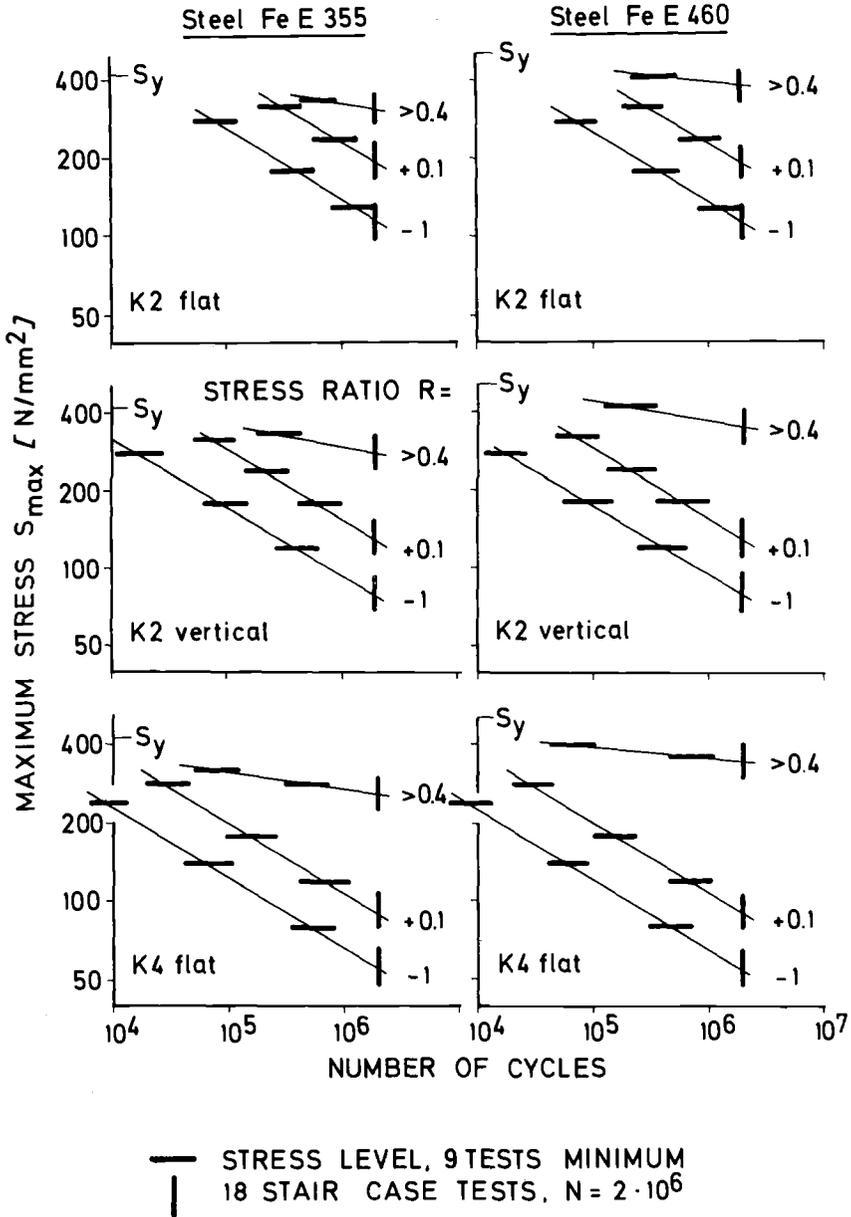


FIG. 2—Survey of the test program.

loading at a high constant mean stress (resulting in $R > 0.4$) by means of endurance tests at preselected stress levels and (separate from the statistical test plan) by staircase tests at $2 \cdot 10^6$ cycles. A total of 18 $S-N$ curves resulted from this program, and they comprised 753 individual results.

Fabrication of the Specimens

The material was ordered to be delivered in plates of 2100 by 1100 by 12 mm with special indication of the rolling direction. The plates were flame cut into assemblies of about 525 by 550 mm; each assembly contained five 60-mm wide specimens (Fig. 3). A fully randomized scheme to distribute the required number of flame-cut assemblies to the three welding institutes and to the particular types of specimens was developed in a computer program by means of random numbers assigned to each assembly. Thereafter, the assemblies were taken and grouped by following an increasing order of these random numbers. Remaining assemblies were stored as stock.

Each of the three welding institutes was ordered to fabricate one third of the estimated number of specimens of each type by following a well-defined specification. A manual welding in a special welding jig was required and the use of basic coated electrodes suitable for the parent materials and specified according to the International Standards Organization (ISO) or American Welding Society (AWS) classification. To comply with this specification, responsibility for the selection of the particular trademark of the electrodes and of the appropriate operating conditions was left to the particular institute. Further details specified were the size and shape of the fillets, the welding position, and the welding sequence, with restarting positions of new electrodes only allowed on the intermediate strips between the specimens. Finally, before the assemblies were cut each specimen was marked to identify the type of steel, the number of the plate and assembly, and the position of the specimen within the assembly.

Elaboration of the Testing Conditions

The elaboration of the testing conditions started with a forecast of the 18 $S-N$ curves to be established, under the assumption that the fatigue strength of welded joints in higher strength structural steels and in mild steel will not differ too much and that a uniform slope of $k = 3.75$ will apply to the $S-N$ curves for a constant stress ratio. This forecast, checked by some preliminary tests, allowed a detailed estimate of the test levels, of the required test loads, of the number of tests, and of the resulting testing time.

For each $S-N$ curve three approximately equidistant test levels were predetermined wherever meaningful. Of these, the upper level was definitely specified in order to observe a sufficient distance from the yield strength, for above that level the $S-N$ curve was expected to bend to the left (low cycle fatigue domain, see Fig. 4). Except for the $R > 0.4$ series, the test levels for specimens from Steel Fe E 460 were the same as for Steel Fe E 355 to ensure the possibility of directly comparing the two materials on the basis of the number of cycles to failure obtained. For the $R > 0.4$ series a different mean stress was chosen equal to two thirds of the specified minimum yield strength values, that is, $\sigma_m = 240 \text{ N/mm}^2$ or $\sigma_m = 320 \text{ N/mm}^2$, respectively, in order

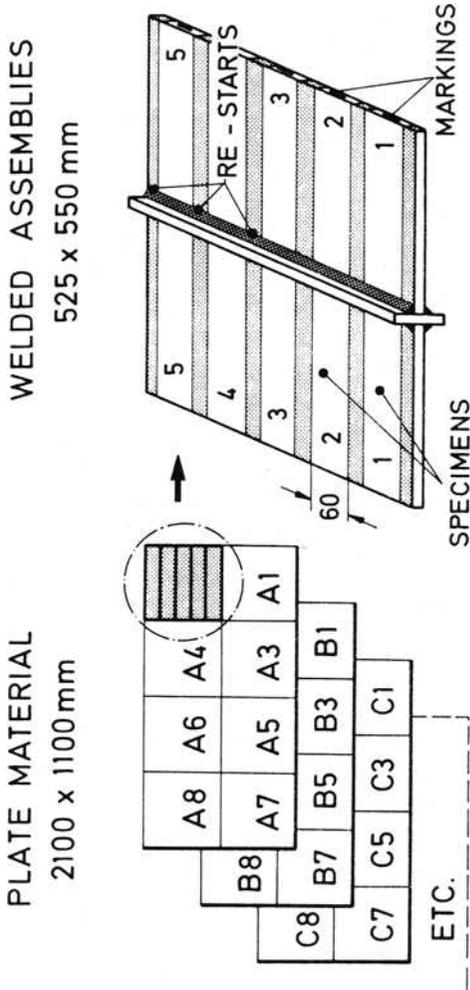


FIG. 3—Partition scheme of the delivered plate material to the welded assemblies and specimens.

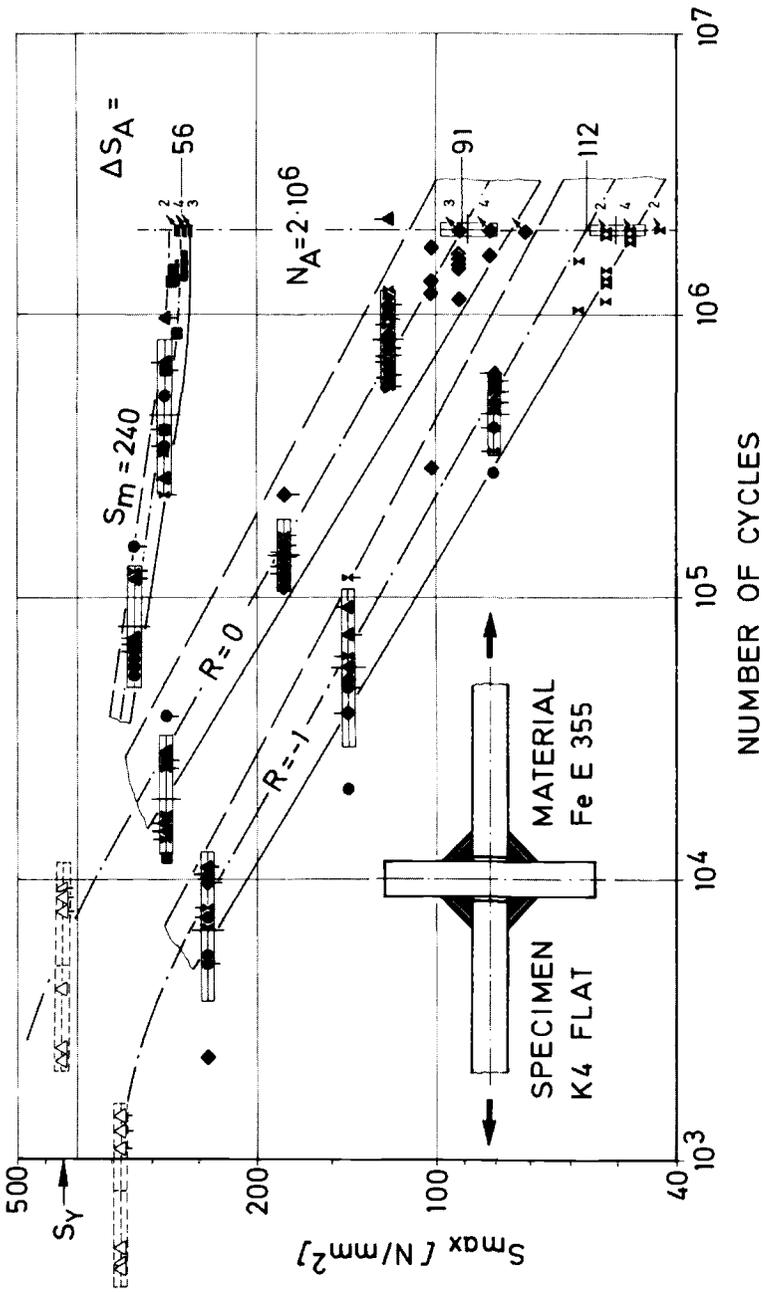


FIG. 4—Example of graphical analysis. Key: ● ▲ ▼ ■ ◆ X = six laboratories; ●●● = three welding institutes; ~~●●●~~ = low-cycle fatigue tests.

to allow for the higher strength of the Fe E 460 material. In addition to the defined test levels in the finite life region, and separate from the statistical test plan, a staircase test series at $2 \cdot 10^6$ cycles was provided for each $S-N$ curve in order to produce a particular estimate of the fatigue strength at $N = 2 \cdot 10^6$ cycles.

Later, the test plan was realized in partial stages, and each of these stages was followed by a preliminary analysis of the results so far obtained to allow the specified testing conditions of subsequent series to be adjusted, if necessary. In order to have new test results available without delay a telex code was agreed on for transmitting them to the secretariat.

Test Plan

The random number technique mentioned earlier was used again to distribute the specimens to the particular test levels in such a way that a balance of the specimens from the three welding institutes was achieved for each test level, together with a partial balance among Positions 1 to 5 of the specimens within the assemblies.

The concept of distributing the test series to the laboratories was worked out under the assumption that any laboratory effects contributing to the overall scatter of the test results could be detected with high probability. In a number of test series, however, the participation of certain laboratories was not possible because their testing machines were not capable of applying alternating loads or loads at the upper stress levels. Furthermore, the amount of work (number of tests and number of cycles) had to be balanced among the laboratories. In developing the test plan these restrictions turned out to be rather limiting; for illustration of the adopted test plan see Tables 4 and 5 in the Appendix.

A stress level was considered as an experimental unit. For the 18 $S-N$ curves (of two types of steel, three types of specimen, three stress ratios) there was a total of 42 stress levels, and theoretically these were to be combined with 18 treatments (by three welding institutes and six laboratories). Hence, $42 \times 18 = 756$ specimens would be required to provide a single replicate of each condition, but only 429 specimens were tested at the 42 levels (a minimum of 9 specimens per level), a circumstance which had direct consequences for the analyses of the results [7].

In particular, it was decided to realize a fully balanced comparison of the six laboratories by means of randomized blocks comprising those stress levels in Steel Fe E 355 that all six laboratories were able to apply (Blocks a and b in Table 4 in the Appendix), where (in Block a) 2×18 notched specimens were tested, in addition, because of their more clearly defined fatigue properties. In the remaining series for Steel Fe E 355, not all factors could be perfectly balanced. The testing laboratories were considered in a way that allowed the analysis in terms of balanced incomplete blocks, consisting of

four laboratories in testing at each stress level (Table 4). For the Fe E 460 test series the test plan was developed at a later stage, hence a similar but somewhat simpler test plan was adopted (Table 4).

Testing Procedure

For all test series the compulsory test conditions were given to the laboratories in terms of the maximum load, F_{\max} , and the minimum load, F_{\min} , as calculated from the predetermined stress levels and the test section. The test section was defined by the nominal width of the specimen and by a reference plate thickness (12.3 mm for Fe E 355 plates and 12.5 mm for Fe E 460 plates) that was determined by measuring a small sample in advance. The responsibility for a correct performance of the tests was left to the laboratories. Only some general recommendations had been given. The recommended modes of a static and a dynamic calibration of the load-measuring device were based on an ISO proposal [8]. In addition, a method was described for checking the exact alignment of the clamping devices by means of a test bar applied with strain gages. After the specimens had been delivered and the corresponding instructions had been distributed to the laboratories, the complete test program was carried out without any difficulties by following the established test plan.

After the statistically designed experiment for the finite life region was completed, the staircase test series was carried out. A complete staircase test series was performed by a single testing laboratory. It consisted of 18 randomly selected specimens in groups of 6 specimens from each welding institute. These three groups were tested sequentially in order not to disturb the staircase scheme by systematic differences that might exist between the specimens of different origin. By extrapolating straight downward on the already existing $S-N$ curve to $2 \cdot 10^6$ cycles, the mean level to start the staircase testing was determined. A step increment equal to one tenth of that mean level in terms of stress range (about 0.8 times the standard deviation as obtained in the subsequent analysis, Table 3) was expected to result in an up-and-down sequence of optimum significance in the case of welded joints.

Analysis of the Test Results

The test results available from realizing the described test plan may be analyzed under various aspects and by various statistical methods; for a complete listing of the individual results see Refs. 1 and 2. The following is a summary of the analyses aimed at elaborating some general conclusions on the validity of a uniform slope and scatter band of the $S-N$ curves, on laboratory and welding effects that may explain the reasons for the scatter of the published fatigue strength values of welded joints, and on the dependence of the fatigue strength values and their scatter on the method of

TABLE 2.—Summary of the fatigue strength values at $2 \cdot 10^6$ cycles, probability of survival 0.5, as obtained by various methods of analysis and by averaging the particular figures.^a

Type of Specimen	Stress Ratio	Steel Fe E 355										Steel Fe E 460											
		σ_{max} , N/mm ²					Average					σ_{max} , N/mm ²					Average						
		a	b	c	d	$\Delta\sigma$	σ_{max}	$\Delta\sigma$	a	b	c	d	σ_{max}	$\Delta\sigma$	Average	σ_{max}	$\Delta\sigma$	Average	σ_{max}	$\Delta\sigma$	Global Average		
K2 welded in flat position	-1	132 ^c	116	...	129 ^c	116	232	115	116	107	118	114	228	115	230	115	230	115	230	115	230	115	230
	+0.1	199	193	201	196	197	177	193	193	193	195	193	174	195	176	193	174	193	193	174	195	176	193
	> +0.4 ^b	311	310	310	140	372	390	...	372	378	116	378	116
K2 welded in vertical position	-1	78	80	73	79	77	154	76	80	76	82	78	156	78	156	78	156	78	156	78	156	78	156
	+0.1	149 ^c	133	131	152 ^c	132	119	122	133	127	129	128	115	128	117	128	115	128	115	128	115	128	117
	> +0.4 ^b	288	290	...	287	288	96	359	366	...	361	362	84	362	84
K4 welded in flat position	-1	50	56	54	52	53	106	52	56	52	61	55	110	54	108	55	110	54	108	54	108	54	108
	+0.1	89	91	97	92	92	83	86	91	92	89	89	80	91	82	89	80	91	82	89	80	91	82
	> +0.4 ^b	267	268	...	267	267	54	342	346	...	341	343	46	343	46

^aKey:

- a = from staircase tests
- b = from graphical analysis
- c = from covariance analysis
- d = from Bastenaire analysis

^bConstant mean stress, $\sigma_m = 240$ N/mm², for Steel Fe E 355 and $\sigma_m = 320$ N/mm² for Steel Fe E 460.

^cThese results from staircase tests are comparatively too high and are not considered when averaging.

evaluation applied. The results were taken as they had been obtained, but attempts were made to identify the reasons for obviously strange results; only if it was clearly justified were those results left out of the analysis.

Graphical Method

A first step in the analyzing process was a plotting of the individual results in log-log S - N diagrams for a check by inspection (Fig. 4). The different symbols used in plotting the results distinguish between the testing laboratories and the three welding institutes involved. From this it appeared that the differences between results produced by the six laboratories and between the specimens fabricated by the three welding institutes were not extremely pronounced, since the particular scatter distributions overlap. Hence the analysis was continued by a standard procedure of computing the mean and standard deviation of log N for each of the stress levels tested and by computing the mean and standard deviation of the stress (both S_{\max} and ΔS) at $2 \cdot 10^6$ cycles from the staircase test series (Tables 2 and 3). The outcome of these computations was plotted in addition to the data points to mark the span of scatter defined by a probability of survival of $P_s = 0.9, 0.5,$ and 0.1 (mean ± 1.28 standard deviations, providing a Gaussian distribution of log N). Finally, to complete the graphical analysis, S - N curves were fitted by averaging the data points. A scatter band of uniform slope and width was used to describe the S - N curves for the two constant stress ratios of $R = -1$ and $R = 0$. According to that uniform scatter band the averaging S - N curve is a straight line, which may be represented analytically in the form

$$N = N_A \cdot (\Delta S / \Delta S_A)^{-k} \quad (1)$$

for $P_s = 0.5$; $R = \text{constant}$; and $S_{\max} < S_{\text{yield}}$, where ΔS_A is the stress range for $P_s = 0.5$ at $N_A = 2 \cdot 10^6$ cycles, and where a slope $k = 3.75$, together with the indicated span of scatter ($P_s = 0.9$ to $P_s = 0.1$), is assumed to hold independent of the type of joint. Although in the planning stage this uniform scatter band of slope $k = 3.75$ was adopted only as a reasonable hypothesis derived from previous experience [9], its significance could be substantiated through the analysis of the present results. In the case of a high constant mean stress, the test results indicate some shallow type of scatter band, but when plotting the results as a function of stress range, ΔS , (instead of S_{\max}) the slope and width of the scatter band become more consistent with the uniform scatter band as well.

From comparing the results for Steels Fe E 355 and Fe E 460 at the individual stress levels, it was found that the same S - N curves could be assigned to the equivalent test series of the two plate materials in the case of $R = -1$ and $R = 0$. The S - N curves for the two types of steel differ for the high constant mean stress testing conditions because of their difference in mean stress.

TABLE 3—Scatter of the fatigue strength values in terms of coefficients of variation in percent, as obtained by various methods of analysis and by averaging the particular figures.^a

Type of Specimen	Stress Ratio	Steel Fe E 35S					Steel Fe E 460						
		a ^b	b	c	d	Average	e	a ^b	b	c	d	Average	e
K2 welded in flat position	-1	20	18	—	7	14.7	12.5	18	14	16	17	13.2	12.5
	+0.1	15	17	15	18	14.7	12.5	9	10	13	14	13.2	12.5
	> +0.4	8	14	14	—	14.7	12.5	6	11	12	12	13.2	12.5
K2 welded in vertical position	-1	9	13	12	11	9.4	12.5	19	11	13	20	12.4	12.5
	+0.1	7	6	7	12	9.4	12.5	6	9	8	17	12.4	12.5
	> +0.4	5	7	9	8	9.4	12.5	(29)	11	10	13	12.4	12.5
K4 welded in flat position	-1	8	18	11	16	13.3	12.5	9	18	14	15	13.0	12.5
	+0.1	8	11	9	15	13.3	12.5	19	9	8	16	13.0	12.5
	> +0.4	17	13	13	14	12.4	12.5	(33)	11	11	15	12.8	12.5
Average	(10.8)	13.0	11.3	12.6	12.4	12.5	(12.3)	11.5	11.7	15.4	12.8	12.5	12.5

^aKey:

- a = from staircase tests
 - b = from the stress levels tested
 - c = from covariance analysis
 - d = from Bastenaire analysis
 - e = from overall analysis of the S-N curves
- ^bStaircase test results are not considered when averaging.

As a result of the graphical analysis, the fatigue strength may be evaluated in terms of the stress range, ΔS_A , to be read from the diagrams for $P_s = 0.5$ and $N_A = 2 \cdot 10^6$ (Fig. 4 and Table 2).

Effect of Tolerance in Plate Thickness

When calculating nominal stress with some nominal or average figure of the plate thickness, the statistical variation in plate thickness is included in the scatter of the test results. In order to gain some idea about that influence, a comparative analysis of the Fe E 355 results was made by readjusting them according to the actual values of plate thickness measured for each specimen

$$N_{i, \text{readjusted}} = N_i \cdot (t_{\text{reference}} / t_{\text{measured}})^k \quad \text{with } k = 3.75 \quad (2)$$

Although this readjustment led to an average decrease of the number of cycles by 5.2 percent, which reflects a difference between the reference thickness of 12.3 mm and the average thickness measured of 1.3 percent, no significant reduction in scatter of the individual test results was achieved by this kind of evaluation.

Overall Scatter

When an analysis of the scatter adhering to the test results is intended, it is more convenient to adopt a model of the $S-N$ curves that shows a scatter band of constant width (standard deviation of $\log N$ independent of the stress level) than the uniform scatter band of variable width used in the graphical analysis. Then, with reference to the $S-N$ diagram, the scatter of the test results may be analyzed by considering the horizontal distances of the individual data points from the corresponding mean curve ($P_s = 0.5$), defined by its uniform slope, $k = 3.75$, and by the particular value of ΔS_A derived from the graphical analysis (Table 2). Due to the logarithmic scale in plotting N , these distances conform to the transformed variable $\log [N_i / N_i^*]$, where N_i is the number of cycles from Test i at the stress level ΔS_i , and N_i^* is the number of cycles to be read from the average $S-N$ curve at that level

$$N_i^* = 2 \cdot 10^6 (\Delta S_i / \Delta S_A)^{-3.75} \quad (3)$$

according to Eq 1.

When plotting the values of N_i / N_i^* on a logarithmic scale in a probabilistic diagram (Fig. 5), an overall average value to 1 is to be expected, if the graphically derived $S-N$ curves provide an average fit of the test results. In fact, all data points in the probabilistic diagram may be fairly well described by a straight line passing through $N_i / N_i^* = 1$ at $P = 0.5$. Separately averaging the results for Steels Fe E 355 and Fe E 460 would result in figures of $N_i / N_i^* = 1.036$ or 0.952 , respectively. Moreover, all the

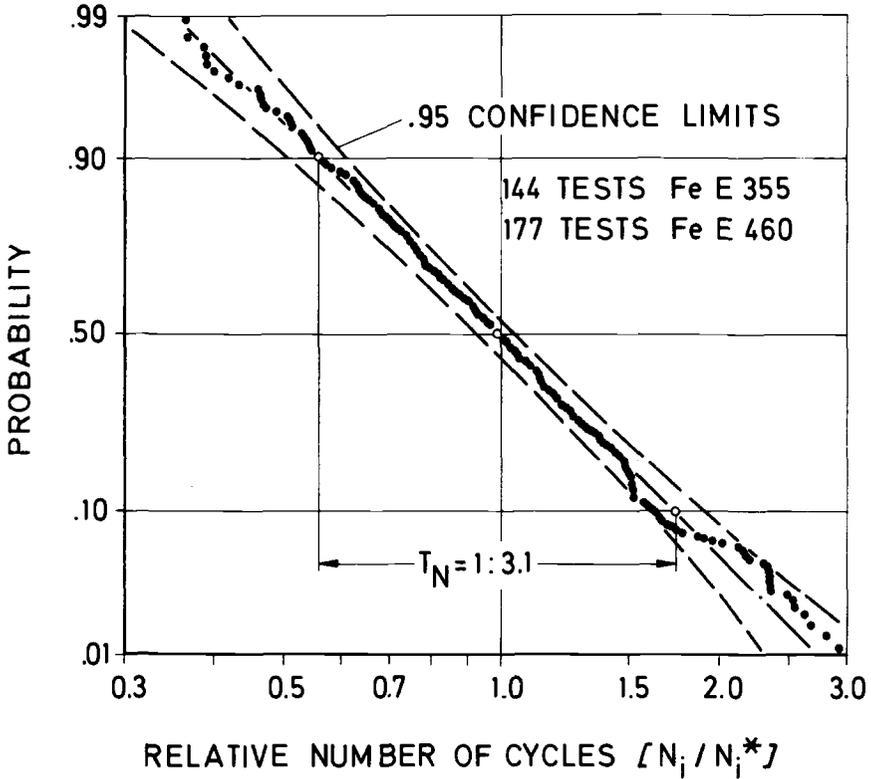


FIG. 5—Overall analysis of scatter on a relative scale: N_i = number of cycles to failure in test; N_i^* = number of cycles calculated from the average S-N curve ($P_s = 0.5$) at the stress level tested.

data points keep within the 0.95 confidence limits to be derived for the 321 test results considered [10]. The span of scatter characterized by the figure

$$T_N = \{[N_i/N_i^*]_{P=0.9}\} : \{[N_i/N_i^*]_{P=0.1}\} = 1:3.1$$

($\log [1/T_N] = 2.56 \times$ standard deviation of $\log [N_i/N_i^*]$ providing a normal distribution) hardly exceeds the values experienced with conventional test series for welded joints [9]. Hence, it may be concluded that the derived S-N curves do give a good overall average description of the test results.

Analysis of the Laboratory Effects

An analysis of the laboratory effects is presented in Fig. 6, in addition to the preliminary analysis based on only a part of the stress levels, described in the Appendix, Table 6.

Again, the transformed data, N_i/N_i^* , have been considered, and this

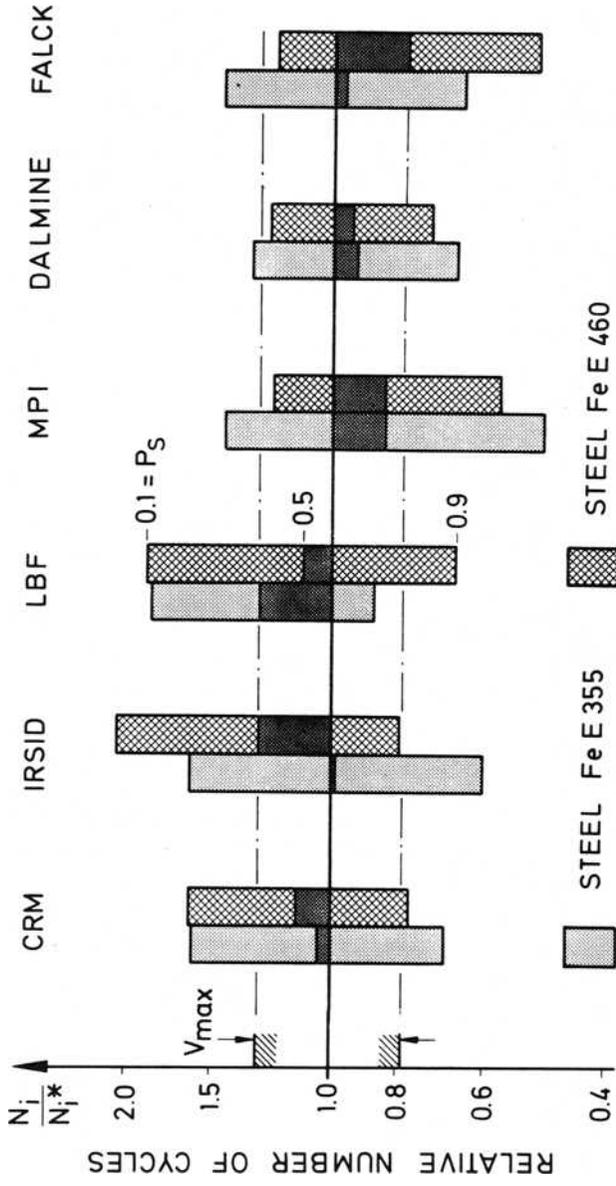


FIG. 6—Analysis of the laboratory effects.

transformation may be understood here as a means to eliminate the stress level, the stress ratio, and the type of specimen as variables from the test plan (Table 7). What remains is a separate test plan for the separately tested Fe E 355 and Fe E 460 specimens, each test plan showing six treatments (by six laboratories) with a perfect (partial) balance of the welding institutes for the Fe E 460 (Fe E 355) series, but with a different number of tests for the particular treatments (Table 7, read vertically).

In Fig. 6 the shaded beams depict the mean ($P_s = 0.5$) and the span of scatter ($P_s = 0.9$ to 0.1) of the test results, as produced by the indicated laboratories. Although some systematic trends become obvious, behind which an effect of the testing frequency may be anticipated, neither the means nor the spans of scatter have been found to differ significantly. The indicated spans of scatter include the variance due to the three welding institutes involved, however. When converted in terms of stress (according to the slope $k = 3.75$ of the $S-N$ curves), the maximum variation, V_{\max} , of the mean values around the overall mean is ± 6.5 percent, and within this level of significance the results from the six laboratories are fairly comparable.

Analysis of the Welding Effects

In the same way an analysis of the welding effects has been performed (Fig. 7) for which not only the two types of steel but also the three types of specimen have been separated into six test plans. Each of these test plans shows three treatments (three welding institutes) with an equal number of tests for the particular treatments, and with a perfect (partial) balance of the laboratories for the Fe E 460 (Fe E 355) test series. Here the laboratory effects are included in the indicated spans of scatter.

A maximum range of variation in stress, V_{\max} , of ± 9 percent was found for the K2 flat specimens with their weld toes ground, whereas the variation is considerably smaller (about ± 6.5 percent maximum) for the K2 vertical and K4 flat specimens, which are sensitive to influences of the welding procedure only. As an explanation, it may be anticipated that grinding the weld toes of the specimens introduces some additional scatter. In grinding the CRM specimens, obviously a most effective increase in fatigue life was achieved. The distinctly lower scatter associated with the IIS specimens indicates a rather careful grinding operation, particularly with those from the Fe E 460 material; but the careful grinding evidently caused a reduction of the test section, which appears as an unwanted effect not taken into account when the nominal stress was calculated.

Covariance Analysis

Considering the results of each $S-N$ curve in $\log \Delta S / \log N$ coordinates, the graphical analysis has shown that straight lines of equal slope may fit the sets

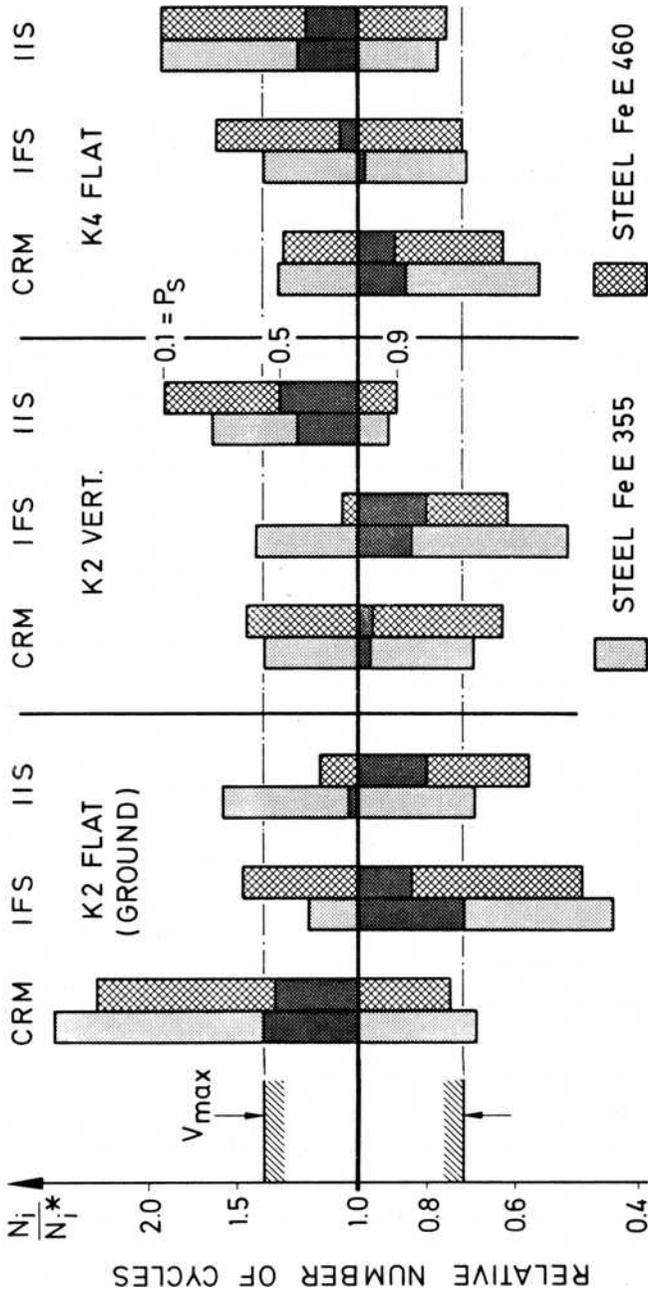


FIG. 7.—Analysis of the welding effects.

of data in a reasonable portion of the sloped part of the $S-N$ curves. Therefore, an analysis of covariance was carried out for the sloped part of the $S-N$ curves; for details see the Appendix.

Among the results of this analysis, the evaluation of eventual variations of slope is worth mentioning. In comparing the particular $S-N$ curves by taking the average slope of $k = 3.72$ as a reference, no significant differences have been found in eight out of twelve cases (two steels, three types of joint, two stress ratios, $R = -1$ and 0), but the following anomalies have been identified in the remaining four cases:

1. $S-N$ curve for K2 flat specimens in Steel Fe E 355 tested at $R = -1$. At the upper stress level failure of the specimens occurred outside of the welded test section; hence, no covariance analysis was carried out for that $S-N$ curve.

2. $S-N$ curve for K2 vertical specimens in Steel Fe E 355 at $R = -1$. A quadratic component was found in the covariance analysis, perhaps because the upper stress level was chosen too close to the yield strength of the base material; an adjusted straight line was established after eliminating the upper level, and the remaining data confirmed the average slope.

3. $S-N$ curve for K4 flat specimens in Steel Fe E 355 tested at $R = +0.1$. The analysis showed a significant difference in slope for the specimens from two welding institutes ($k = 4.80$) from those of the third ($k = 3.72$); when the cruciform-type specimens from the former two institutes were examined, their root surfaces were found to be in close contact, whereas the specimens from the third institute showed some root opening, and that difference may well explain the different fatigue behavior observed at high and low stress levels.

4. $S-N$ curve of K2 vertical specimens in Steel Fe E 460 tested at $R = -1$. A marked unbalance of the laboratory effects appeared when the averages of the data obtained in each laboratory were considered (see Fig. 12, in the Appendix).

In addition to the best fit regression lines, the scatter bands defined by $P_s = 0.9$ and 0.1 (± 1.28 standard deviations of $\log N$) have been computed and presented in diagrams; Fig. 8 gives an example. The computations followed by homogeneity tests indicate that the standard deviations of $\log N$ differ significantly when the three types of joint are compared. When treating the effect of the welding institute on the fatigue strength at $N = 5 \cdot 10^5$ cycles, there is evidence of such an effect for 7 out of the 12 $S-N$ curves, whereas in some other cases an effect may be supposed, although the data do not differ significantly. Finally, the fatigue strength values for $N = 2 \cdot 10^6$ cycles and $P_s = 0.5$ have been derived by extrapolating the best fit regression lines (Table 2). These figures deviate from the staircase test results by less than 13 per cent.

In conclusion the outcome of the covariance analysis confirms the hypothesis of a linear $\log \Delta S / \log N$ dependence and of a uniform slope ap-

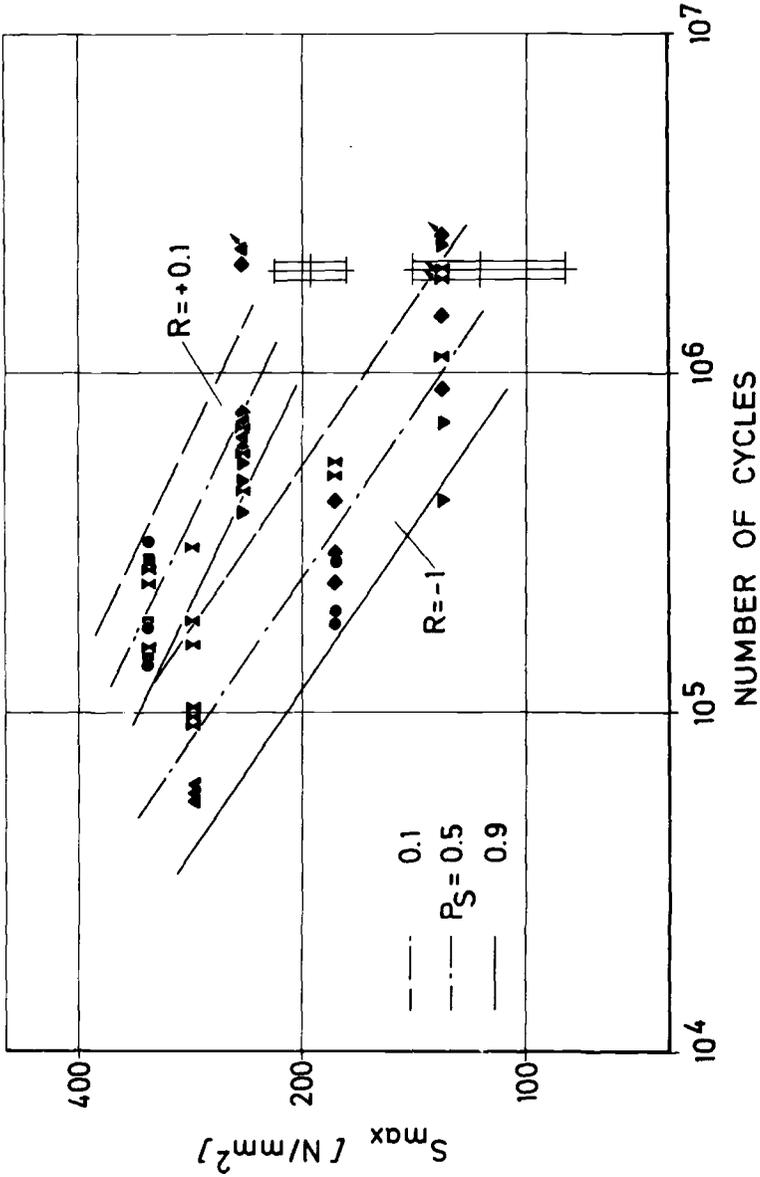


FIG. 8—Example of the covariance analysis [11] of K2 flat specimens, and Steel Fe E 460.

plying to $S-N$ diagrams of welded joints in structural steel, when established for a constant stress ratio and when plotted in terms of the stress range. For further details see Ref. [11].

Bastenaire-Type $S-N$ Curves

A computing algorithm proposed by Bastenaire [12] is aimed at fitting the following function to a set of fatigue test results

$$N = A \exp \left[- \left(\frac{S - E}{B} \right) C \right] / (S - E) \quad (5)$$

where

- N = number of cycles to fracture (or failure criterion),
- S = stress amplitude (or some other variable proportional to it),
- E = endurance limit, and
- A, B, C = parameters.

Equation 5 may take various degenerated forms; with $C = 0$, one gets the equation

$$1/N = (1/A) \cdot (S - E) \quad (6)$$

If Equation 6 is applicable to fatigue data, it means that the data can be represented by the linear regression of the random variable $1/N$ on the stress amplitude; if there are runouts at several stress levels, the statistical distribution of the fatigue life reciprocal $1/N$ must be regarded as a normal censored distribution, and it has to be processed by a special program. In general, estimation of the parameters A , B , and C is more complex because Eq 5 is nonlinear in relation to these parameters, and an iterative algorithm is used to estimate A , B , and C .

The Bastenaire type of analysis was applied to the present sets of data, too. The results, published elsewhere in detail [13], are included in Tables 2 and 3; a typical set of the so-derived $S-N$ curves is presented in Fig. 9.

Values of the Fatigue Strength at $2 \cdot 10^6$ Cycles

Table 2 gives a summary of the fatigue strength values at $2 \cdot 10^6$ cycles as obtained from

- (a) the staircase test series,
- (b) the graphical analysis by means of the uniform $S-N$ scatter band,
- (c) the covariance analysis in the $S-N$ diagram, and
- (d) fitting the Bastenaire-type $S-N$ curves.

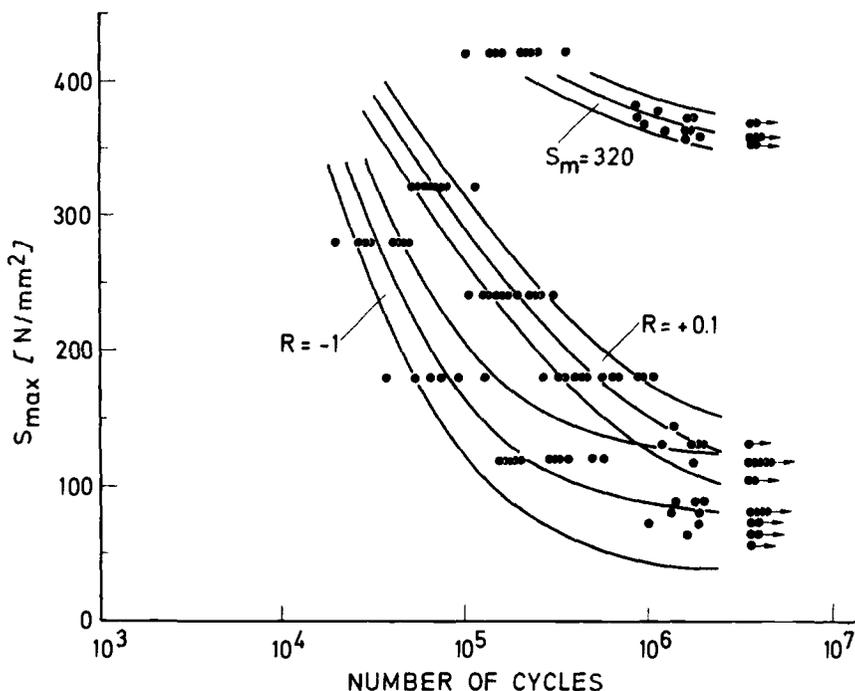


FIG. 9—Example of the Bastenaire analysis [13] of K2 vertical specimens, and Steel Fe E 460.

The probability of survival associated with these values is $P_s = 0.5$. On a general basis the results from the four types of analysis compare fairly well. The maximum differences amount to 16 percent, but a typical difference is 6 percent. In two cases the results from the staircase tests seem to be comparatively too high, but the reason for this is unknown.

In order to derive definite fatigue data applicable to design and dimensioning of welded constructions, an averaging of the equivalent figures from the various analyses appeared to be a reasonable solution. This averaging process has been made in two steps: (1) individually for each of the two types of steel considered, and (2) by a global averaging, disregarding the difference in material. The average values for the two types of material differ by less than 4 percent. The test series with a high constant mean stress have to be excluded in the second step of averaging due to the previously mentioned difference in mean stress for the two types of steel.

Scatter Associated with the Fatigue Strength Values

In combination with the fatigue strength values of Table 2, the scatter of the individual test results has to be considered (Table 3). Listed are the coefficients of variation in percent, defined as the ratio of the standard deviation,

s , to the fatigue strength value for $P_s = 0.5$, and as obtained by the following methods of analysis:

- (a) from the staircase tests,
- (b) from averaging the particular analyses of the stress levels tested,
- (c) from the covariance analysis,
- (d) from the Bastenaire-type analysis, and
- (e) from the overall analysis of the $S-N$ curves (Fig. 5).

Although the listed figures show some remarkable differences, there is little indication that these differences have any reasonable tendency. The average values prove to be much more uniform, but the values from the staircase analysis appear to be either slightly lower than the others or, in two cases, much higher, which is not surprising with this kind of test method. Therefore, the values from the staircase tests have been excluded when computing the averages for the particular types of specimen. On a global basis there is quite a good agreement between the overall analysis and the average values from the three remaining methods of analysis and also with the width of the scatter band plotted in Fig. 4. As a final result, a coefficient of variation of 0.125 may be taken as a reliable estimate of the scatter associated with the average fatigue strength values given by Table 2.

Comparison with Data from the Literature

Finally, the present test results are compared with data from the literature (Figs. 10 and 11). Of course, this comparison is lacking in the material aspect, for there are no results reported in the literature that exactly apply to the two types of materials tested here. The samples presented were selected, however, to be consistent with the present investigation in terms of specimen geometry and the stress ratio.

The comparison of fatigue data of different origin may be biased because the reported data normally were derived by rather different methods of evaluation. Therefore, a reanalysis of the individual test results was made by replotting of the results in (log-log) $S-N$ diagrams and by a graphical best fit of the uniform scatter band of slope $k = 3.75$ to the data points. The position of the scatter band defines the particular value of the characteristic stress range, $\Delta S_A = \Delta S(N = 2 \cdot 10^6, P_s = 0.5)$, and this procedure agrees with the present graphical analysis in the $S-N$ diagram. Nevertheless, the obtained fatigue strength values show a reasonable variation. Therefore, the comparison has been made in terms of a probabilistic diagram. From there it may be seen that the K2 vertical specimens obviously define a lower limit of variability, as was anticipated when specifying the vertical welding position. On the other hand the fatigue strength values of the K2 flat specimens clearly tend toward the upper limit of the distribution, and again this appears to be predetermined by the favorable weld profile specified. The fatigue strength

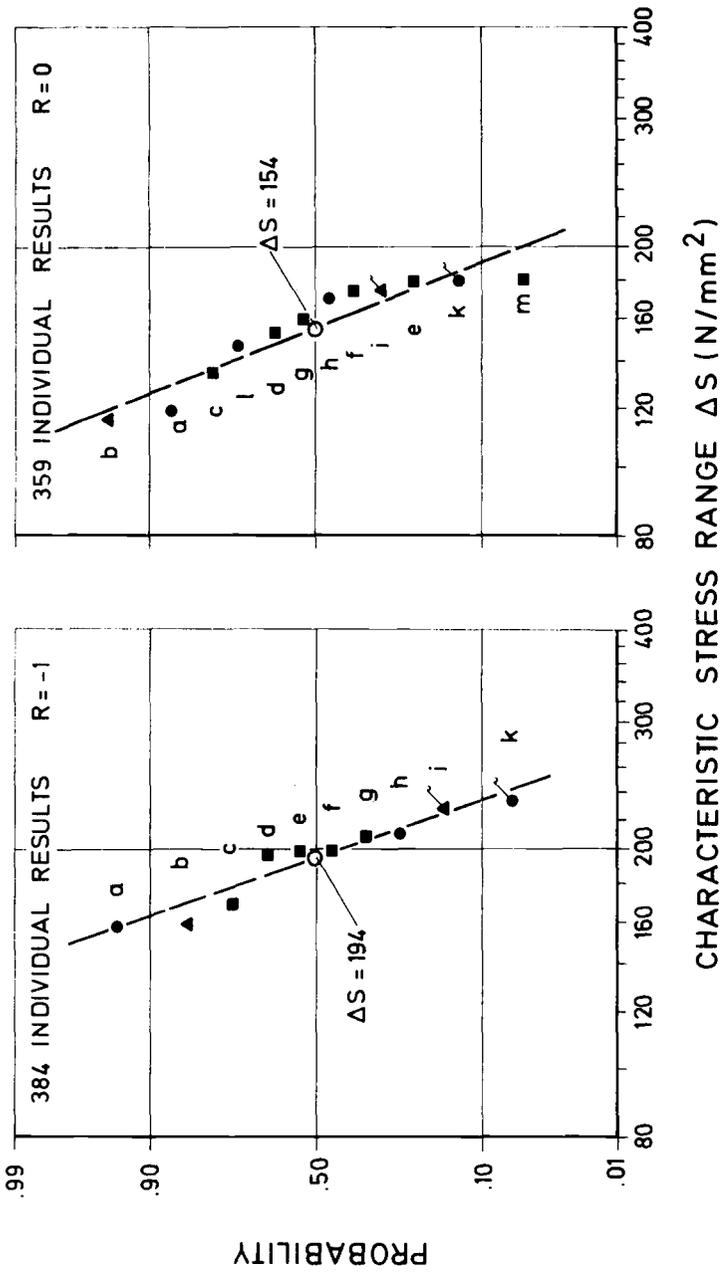


FIG. 10—Comparison of the test results with literature data for non-load-carrying fillet welds. Points a, and b: K2 vertical specimens; Points i and k: K2 flut specimens from present investigation.

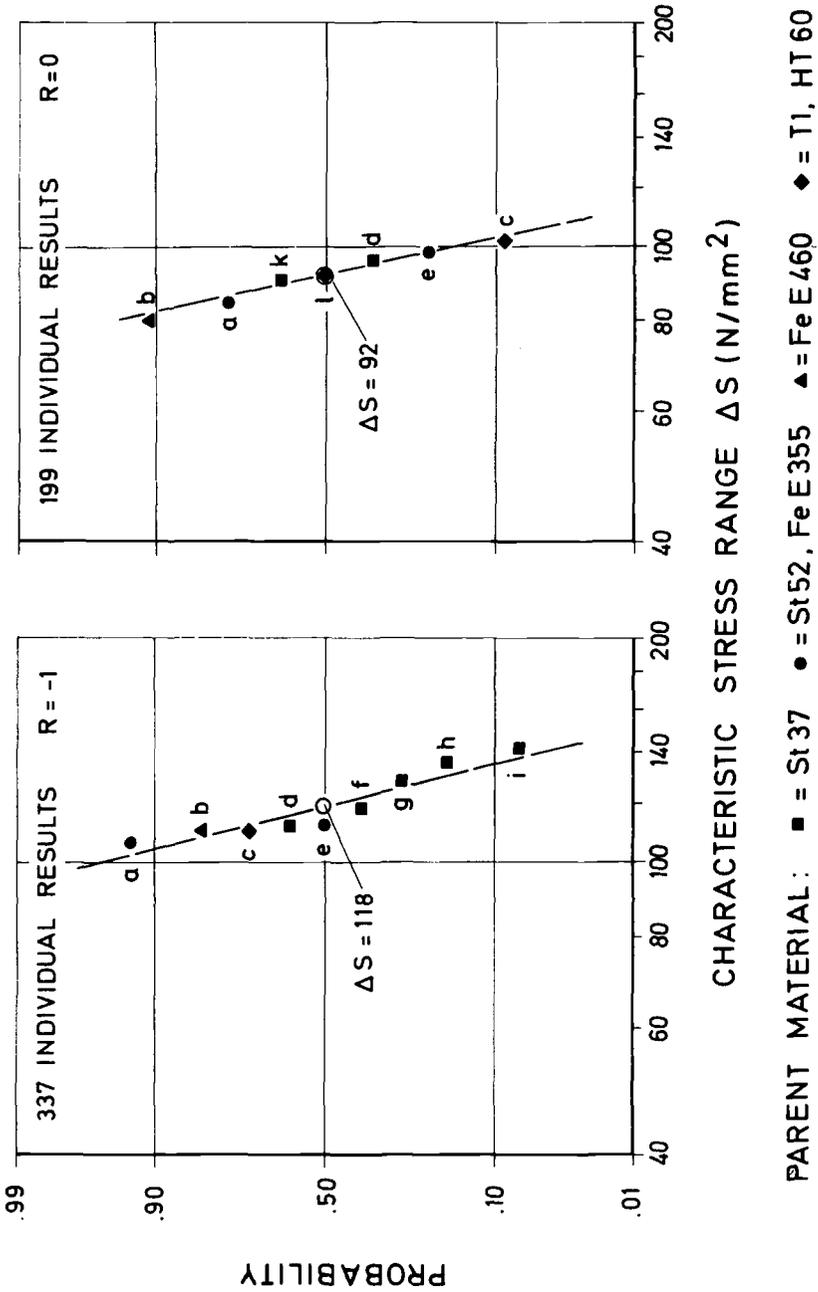


FIG. 11—Comparison of the test results with literature data for cruciform load-carrying fillet welds. Points a and b: K4 flat specimens from present investigation.

values of the K4 specimens show less variation; the present data form the lower end of the distribution. Another question of particular interest is whether any correlation exists between the observed fatigue strength values and the type of base material. The presented samples, however, clearly indicate that an effect of the base material on the fatigue strength is less significant than other effects. In particular, welding effects appear to be of major importance.

Acknowledgment

The authors acknowledge the financial support of the Interlaboratory Program by the European Coal and Steel Community and by the national contractors. They also recognize with gratitude the personal engagement and the active contributions of their colleagues in the working group in realizing the program.

APPENDIX

The distribution of the test series to the testing laboratories is outlined in Table 4. Due to the known importance of the weld profile on the fatigue life, the welding effects were given high priority in the balancing effects. Therefore, at each stress level the same number of specimens was provided from each of the three welding institutes. Additionally, it was intended for the Fe E 355 test series to involve at each stress level as great as possible a number of testing laboratories by means of a balanced incomplete block structure (Blocks b to f). For the Fe E 460 test series, a direct balance of the welding effects for each laboratory was preferred. Table 5 illustrates the repartition of the specimens to the particular test series and to the testing laboratories. In the coding of the specimens, the letters distinguish the plates, the integer number indicates the assembly, and the decimal number indicates the specimen position (see Fig. 3).

Based on the indicated blocks, an analysis of variance could be carried out on each line of Table 4 allowing for a comparison among the laboratories involved. This type

TABLE 5—*Repartition of the specimens to the particular test series and to the testing*

R	Level	CRM ^a			IRSID			LBF		
		IIS ^b	IFS	CRM	IIS	IFS	CRM	IIS	IFS	CRM
-1	3	P 2.1	F 4.5	BB 2.4	L 6.2	D 5.1		F 1.3		Z 5.2
	1		AA 8.3	X 7.2	KK 2.5	DD 7.4	T 3.3	F 1.1	LL 7.5	
	5	K 5.1		W 5.5		M 2.2	X7.1	AA 1.4	F 4.3	MM 2.2
+0.1	3	F 1.4	LL 7.3		L 6.5		DD 2.4		V 4.1	T 3.5
				L 1.2		L 2.2		Z 3.3		
	4				Z0.35	C 2.4	Z 5.1	L 6.1		BB 2.2
0.4	5				U3.3	V 4.2		B 5.4	HH 6.3	B 8.5
	3				B5.3		L 1.4		R. 3.4	BB 2.5

^aLaboratories.
^bWelding institutes.

of preliminary analysis was particularly useful for controlling, in due time, anomalous trends that might have occurred for some laboratory results and for checking for any systematic interlaboratory differences. In addition, in Block a the laboratory situation was checked by 36 tests on notched specimens for which the intrinsic scatter may be expected to be less than that for the welded specimens (three tests by each laboratory at each of two stress levels, stress ratio $R = +0.1$, Steel Fe E 355). The corresponding analysis of variance is presented in Table 6 as an example. It shows significant differences between the laboratories (approximately 10 percent in terms of stress). Also the laboratory averages differed significantly in some blocks tested and not in others. For comparison, the overall analysis in Fig. 6 shows a more satisfactory picture. Table 7 gives the reduced test plan on which Figs. 6 and 7 were based.

In spite of the laboratory effects identified through the variance analyses, the number of laboratories and the comparatively large intrinsic scatter of the welded specimens made it possible to analyze the slope of the $S-N$ curves without taking into account the differences among the laboratories in all cases but one, as mentioned in the text (Fig. 12).

For all $\log S/\log N$ diagrams, regression lines were calculated and drawn, and the 80-percent confidence limits of single values were indicated by parallel straight lines. The standard deviation used was obtained by pooling the contributions of the residual line and of the welding institutes in order to represent the actual situations in welding

TABLE 6—Example of the analysis of variance. Block a:
36 notched specimens; 2 stress levels (LEV); 6 laboratories (LABS);
variable examined; $\log N_p$

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	$F_{\text{calculated}}$
Between LABS	5	0.35730	0.07146	3.92** ^a
Between LEV	1	0.95736	0.95736	... ^b
LABS \times LEV	5	0.17480	0.03496	1.92 ^c
Residual	24	0.43727	0.01822	
Total	35	1.92673		

^a $F_{\text{tabulated}}(5,24; 0.99) = 3.90$.

^bVery large as expected; of no practical interest.

^c $F_{\text{tabulated}}(5,24; 0.90) = 2.10$.

laboratories (example referring to the K2 vertical specimens in Steel Fe E 355).

MPI			Dalmine			Falck		
IIS	IFS	CRM	IIS	IFS	CRM	IIS	IFS	CRM
	AA 8.4	FF 5.3						
AA 1.2		V 5.1						
AA 1.5	F 4.4							
CC 4.3	D 5.2	LL 6.1						
	E 7.2	BB 2.3	K 5.2	GG 7.5	V 5.3	A 6.1	E 7.4	T 3.1
		V 5.4	P 2.4	GG 7.3				
U 3.5				L 2.1	LL 6.2			
Z 3.1	L 2.5		X 6.2	HH 6.1	D 1.3			

TABLE 7—Reduced test plan applying to the transformed variable $\log(N_i/N_i^*)$.

Type of Specimen	Welding Institute	Steel Fe E 355										Steel Fe E 460										
		CRM ^a	IRSID	LBF	MPI	DAL	FA	Total	CRM	IRSID	LBF	MPI	DAL	FA	Total	CRM	IRSID	LBF	MPI	DAL	FA	Total
K2 flat	CRM	4	3	5	6	2	1	21	3	7	3	2	2	2	19	2	2	2	2	2	2	19
	IFS	4	2	6	5	3	1	21	3	7	3	2	2	2	19	2	2	2	2	2	2	19
	IIS	4	3	5	6	2	1	21	3	7	3	2	2	2	19	2	2	2	2	2	2	19
K2 vert	CRM	4	5	6	5	3	1	24	3	5	5	4	4	25	4	4	4	4	4	4	25	
	IFS	3	6	5	5	4	1	24	3	5	5	4	4	25	4	4	4	4	4	4	25	
	IIS	3	6	6	5	3	1	24	3	5	5	4	4	25	4	4	4	4	4	4	25	
K4 flat	CRM	4	6	7	6	3	1	27	4	4	4	7	5	27	5	3	3	7	5	3	27	
	IFS	4	6	6	7	3	1	27	4	4	4	7	5	27	5	3	3	7	5	3	27	
	IIS	4	7	6	6	3	1	27	4	4	4	7	5	27	5	3	3	7	5	3	27	
total		34	44	52	51	26	9	216	30	48	36	39	33	213	27	27	27	33	27	27	213	

^aSee Figs. 6 and 7.

^bLaboratories.

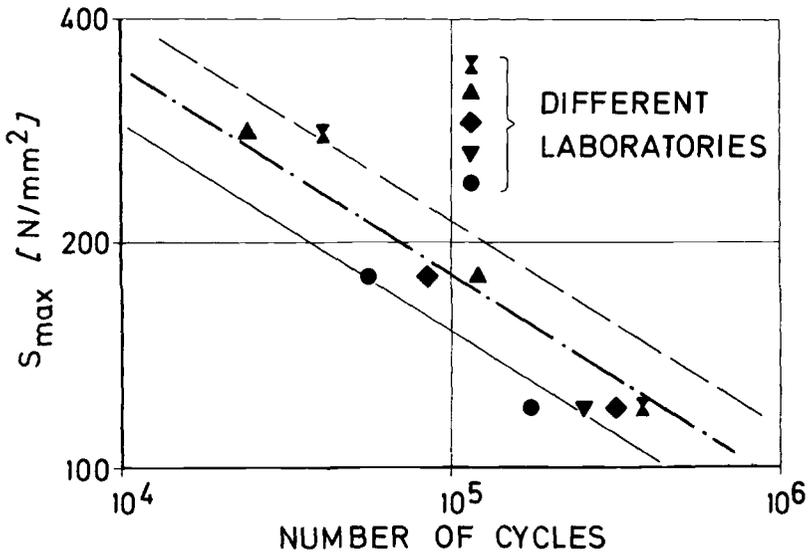


FIG. 12—Laboratory effects associated with the test series for K2 vertical specimens in Steel Fe E 460 tested at $R = -1$.

TABLE 8—Example of the analysis of covariance.^a

Source of Variation	<i>f</i>	Σx^{2b}	Σxy	Σy^2	Deviation from Regression Line			
					<i>f</i>	$\frac{\Sigma y^2 - (\Sigma xy)^2 / \Sigma x^2}{\text{Mean Square}}$		
Welding institute	CRM	5	0.0457	-0.1628	0.6827	4	0.1022	0.0256
	IFS	5	0.0461	-0.1390	0.6001	4	0.1807	0.0452
	IIS	5	0.0447	-0.1649	0.7024	4	0.0945	0.0236
Within		12				12	0.3774	0.0315
Between		2				2	0.0116	0.0056
Common		15	0.1365	-0.4667	1.9852	14	0.3891	0.0278
Adj. means ^c		2				2	0.1503	0.0754
Total		17	0.1369	-0.4737	2.1786	16	0.5398	0.0337

^aEighteen specimens: K2 vertical, Steel FeE355, stress ratio $R = -1$; 2 stress levels (upper level is disregarded because it is too close to yield strength); variables examined: $x = \log S$; $y = \log N$; null hypothesis: no significant difference of the fatigue data N_i is caused by the welding institutes; therefore, the 18 test results give a correct estimate of life expectancy for $P_s = 0.5$. Previous check: Bartlett's test on homogeneity of variance.

^bStress based on individually measured thickness of the specimens.

^cTo individuate the central point position of the regression line.

practice. For all other considerations, only the variance of the residual line was used. The covariance analysis was carried out according to the scheme of Table 8.

References

- [1] "Fatigue Investigation of Higher Strength Structural Steels in Notched and in Welded Condition," Report EUR 5357 e, Commission of the European Communities, Directorate General, Scientific and Technical Information and Information Management, Luxembourg, 1975.
- [2] "Fatigue Investigation of Typical Welded Joints in Steel FeE 460 as compared to FeE 355" Report EUR 6340 EN, Commission of the European Communities, Directorate General, Scientific and Technical Information and Information Management, Luxembourg, 1979.
- [3] Lieurade, H.-P. and Haibach, E., *Revue de Metallurgie*. Vol. 75, No. 3, 1967, pp. 177-191.
- [4] Guerrera, U., "Rules and Experimental Data in Fatigue," IIS-IIW-Document XV-228-67, Istituto Italiano della Saldatura, Genova, Italy, 1967.
- [5] Costa, G., "Welded Joints or Attachments Subjected to Fatigue Load; Ratio between the Stresses Allowed for Them by Various National Rules and the Stresses Allowed for the Unaffected Base Material by the Same Rules," IIS-IIW-Document. WG-XIII-XV-8-68, Istituto Italiano della Saldatura, Genova, Italy, 1968.
- [6] Gurney, T. R., "A Comparison of Fatigue Design Rules," *Proceedings*, Vol. 1, Conference on Fatigue of Welded Structures, 6-9 July 1970, The Welding Institute, Abington, Cambridge, England, 1971.
- [7] Natrella, M. G., *Experimental Statistics Handbook 91*, U.S. Department of Commerce, National Bureau of Standards, Washington, D.C., 1966.
- [8] "Dynamic Force Calibration of Axial Load Fatigue Testing Machines by Means of a Strain Gauge Technique." Revised draft of secretariat proposal ISO/TC 17/SC 6 (Secretariat-170) 346, Sept. 1970.
- [9] Haibach, E., *Proceedings*, Vol. 2, Conference on Fatigue of Welded Structures, 6-9 July 1970, The Welding Institute, Abington, Cambridge, England, 1971.
- [10] Henning, H. J. and Wartmann, R., *Mitteilungsblatt für mathematische Statistik*, Vol. 9, 1957, S. 168-181.
- [11] Contaretti, A. and Rinaldi, F., *La Metallurgia Italiana* No. 5, 1977.
- [12] Bastenaire, F. A., *Probabilistic Aspects of Fatigue, ASTM STP 511*, American Society for Testing and Materials, Philadelphia, 1972, pp. 3-28.
- [13] Bastenaire, F. A., Lieurade, H.-P., and Regnier, L., "Fatigue Behaviour of Cruciform Welded Joints in St E 355 and Microalloyed St E 460 High Strength Steels," *Revue de Metallurgie*, in press.

Reliability of Fatigue Testing

REFERENCE: Young, L. and Ekvall, J. C., “**Reliability of Fatigue Testing,**” *Statistical Analysis of Fatigue Data, ASTM STP 744*, R. E. Little and J. C. Ekvall, Eds., American Society for Testing and Materials, 1981, pp. 55-74.

ABSTRACT: A statistical analysis was performed on fatigue test data for aluminum, titanium, steel, and nickel materials. The data for the titanium, steel, and nickel were obtained from spectrum fatigue tests, whereas the data for aluminum were obtained from both constant amplitude and spectrum fatigue tests. The analyzed data consisted of a total of 553 *S-N* test groups with 2417 specimens and 1288 spectrum test groups with approximately 5000 specimens.

The distribution of logarithmic standard deviation of fatigue life for these test groups was analyzed with normal, logarithmic, and two-parameter Weibull probability distribution functions and with 2-deg polynomial equations. The best fit was evaluated using the coefficient of correlation and the chi-squared goodness of fit test. None of the distribution functions or polynomial equations provided the best fit for all of the distributions of the logarithmic standard deviation of fatigue life for the selected sets of test groups. A comparison is also made of three methods of calculating scatter factors.

KEY WORDS: statistical analysis, fatigue, standard deviation, test life reduction factors, normal distribution function, logarithmic distribution function, Weibull distribution function, 2-deg polynomial equations, chi-squared goodness of fit test, spectrum fatigue test data, *S-N* fatigue test data, aluminum, titanium, steel, nickel

Nomenclature

- c Joint probability confidence level of (c_1) (c_2)
- c_1 Mean life confidence level
- c_2 Logarithmic standard deviation confidence level
- e Cell size
- F_{tu} Ultimate tensile strength
- g Factor of one half
- i i^{th} variance in a set of variances or count number
- j j^{th} specimen in a test group
- k Number of test groups
- \bar{l} Mean log of fatigue life of specimens tested in a group

¹Design specialist and staff scientist, respectively, Lockheed-California Co., Burbank, Calif. 91520.

l_j	Log of fatigue life of j^{th} specimen in a test group
m	Total number of logarithmic standard deviations or total number of data points
MN/m ²	Meganewtons per square metre
n	Total number of specimens
n_i	Number of specimens in i^{th} test group
P	Probability
R	Reliability level or probability of no failure
r	Number of coefficients in equation or correlation coefficient
S	Unbiased sample standard deviation
S_i	Sample standard deviation of the i^{th} test group
$S_{p,c}$	Scatter factor for probability of failure, P , and joint confidence level, c
W	Weibull variate equal to $\ln \ln [1/1 - P]$
X	Normal deviate
X_i	The i^{th} fatigue failure in a test group composed of n specimens
Y	Standard deviation of log test life or logarithm of standard deviation of log test life, as defined when used
z	Normal deviate
z_{c1}	Normal deviate for mean life confidence level, $c1$
z_p	Normal deviate for probability of failure, P
β	Confidence level
ν	Degrees of freedom
σ	Population standard deviation
σ_{c2}	Standard deviation for confidence level, $c2$
σ_i	Standard deviation for i^{th} test group
σ_n	Standard deviation for n^{th} test group
σ_s	Unbiased estimate of population standard deviation
σ_β	Standard deviation for confidence level, β
χ^2	Chi-squared variate

The fatigue life of an aircraft structure is generally based on the results of representative full-scale component fatigue tests. Due to budget and schedule restraints, only one or two components are usually tested for critical areas of the structure. To obtain confidence in achieving a specified fatigue life, the test results are factored by a life-reduction or scatter factor. Since only a few structural details are replicated in any component, the confidence level based only on the component test results is not very high. Therefore, additional test data are used to define the confidence level in achieving a specified reliability corresponding to a probability of no fatigue failures for untested structures.

The variability in fatigue test results is generally expressed in terms of a probability distribution function. Two of these functions commonly used to represent fatigue data are the two-parameter logarithmic and two-parameter Weibull distribution functions. The choice of one function or the other is a

matter of preference since most studies have indicated that either function can be used to represent fatigue scatter of identically tested specimens over the probability, P , range of $0.05 < P < 0.95$. No one probability distribution function provides the best fit to available large sets of data for probability values below 0.05 and above 0.95. Stagg [1-3]² summarizes many investigations that have been made to determine the type of distribution function for fatigue test data and concludes that the logarithmic distribution provides a reasonably good approximation for most data sets. In this paper, the scatter characteristics of fatigue data are investigated in terms of a logarithmic distribution of fatigue lives.

The pooling of data sets for evaluating fatigue scatter has also been studied by various investigators [1-10]. The effect some variables have on a fatigue scatter is summarized in the following sections.

Type of Material

The scatter for both Type 7075 and Type 2024 aluminum alloys is similar [2,7,10] and did not vary with product form [10] except for hand forgings [7], in which it was greater. For titanium and low-strength steel ($F_{tu} < 1655$ MN/m², that is, 240 ksi) the scatter in $S-N$ data is the same or greater than for aluminum alloys [9]. The combined scatter for spectrum loading tests conducted on low-strength steel and nickel-base alloys was less than that for aluminum alloys [5].

Type of Test History

Test histories can be classified generally as constant amplitude and variable amplitude. Three types of variable amplitude fatigue tests that have been conducted for aircraft structures are illustrated in Fig. 1. For scatter evaluations, previous investigations have generally combined data for the three types of spectrum tests. Tests have also been conducted with the sequence varied within a block or flight. Jost and Verinder [8] found that the loading sequence did not significantly influence scatter in the fatigue life of full-scale structure tests.

Tension-tension and tension-compression $S-N$ data [6] exhibit similar scatter, which is less than that for compression-compression fatigue data [7]. Five studies [6-10] indicate that the scatter for $S-N$ tests is greater than for variable amplitude loading spectrum tests. Other studies [4,7,8] indicate that the scatter characteristics are similar for $S-N$ testing and spectrum testing over the life range of 10^2 to 10^6 cycles for aluminum alloys or if the stress range experienced by the material is similar for the two types of testing [3,5]. Jost [8] also found

²The italic numbers in brackets refer to the list of references appended to this paper.

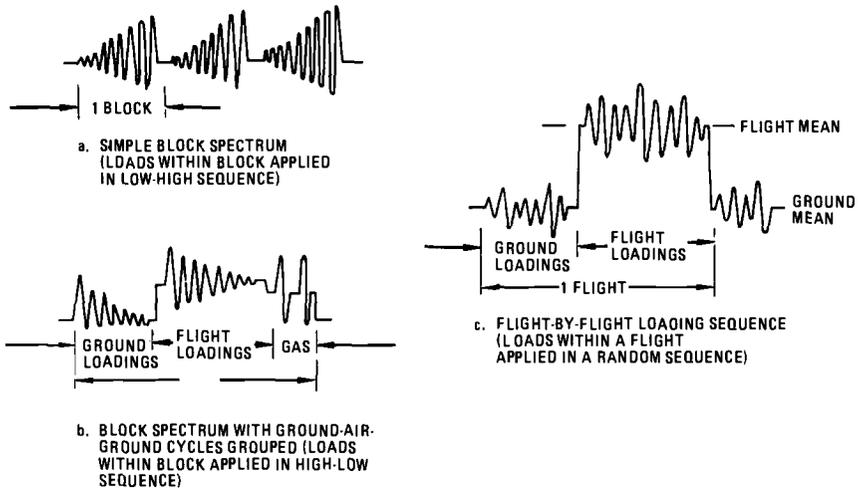


FIG. 1—Types of fatigue loading histories.

that maneuver-loading spectra of structures exhibited less scatter than $S-N$ tests and gust-loading spectra exhibited more scatter than $S-N$ tests.

Number of Fatigue Loading Cycles

The scatter in spectrum fatigue test data was found to be independent of fatigue life [4, 5, 10]. For constant amplitude fatigue tests, the scatter was constant for aluminum alloys over the life range of 70 to 10^6 cycles [4, 7] and 10^2 to 10^7 cycles [8] and for titanium alloys up to 4×10^5 cycles [9]. Schijve and Jacobs [4] reported that the scatter did not vary with the number of cycles applied in spectrum fatigue tests. However, Abelkis [6] indicates that scatter varied with fatigue life and generally increased with increasing life for both $S-N$ and spectrum testing.

Type of Specimen

Unnotched fatigue specimens have been found to exhibit less scatter than notched specimens [1, 6, 7]. Scatter does not vary with the geometry of the notch [7, 10]. The scatter characteristics are similar for aluminum notched, joint, and structural specimens [5, 6, 10] and for titanium notched, lap, and butt-joint specimens [9]. The scatter is greater for bonded joints [7].

Type of Loading (Testing)

Axial-loaded and flexure-loaded specimens exhibit similar scatter characteristics which are less than for notched rotating bending tests [7]. Tests con-

ducted using mechanical shakers had more scatter than servo control test data [10].

Specimen Fabrication

Ultracarefully prepared specimens [7] or specimens made with a very fine finish [10] exhibit less scatter than specimens made by using conventional fabrication methods and finishes.

In this study, scatter characteristics were collected and evaluated for test data applicable to aircraft structures. The two primary variables studied were (1) material: aluminum, titanium, and combined steel and nickel-base alloys; and (2) type of loading history: simple block spectra loading, flight-by-flight spectra loading, and constant amplitude loading.

Test Data

The variable amplitude-loading spectrum test data analyzed for scatter were obtained from literature searches [11, 12]. The data selected for analysis consisted of only test groups with three or more variable amplitude loading levels and with all of the test specimens failed, that is, no runouts. In addition, test groups that contained unnotched or bonded specimens were not included. Also omitted were test groups that contained specimens either tested under narrow-band random loading or subjected to any service experience or preload before testing.

The selected and statistically analyzed spectrum test data are summarized in Table 1. Test groups containing only two specimens with simultaneous failures were not included in the statistical analysis because the logarithmic standard deviation of these groups was equal to zero.

The spectrum test data summarized in Table 1 were obtained on specimens fabricated from aluminum, titanium, steel, and nickel alloys. The aluminum alloys were Types 2024, 2218, 7075, 7079, and 7178. The titanium alloys were alpha-Ti-8Al-1Mo-1V, alpha-beta-Ti-6Al-4V, and beta-Ti-13V-11Cr-3Al. Finally, the steel and nickel alloys that were statistically analyzed together consisted of low-alloy steel, carbon steel, Types 4340 and 300M high-strength steels, Types 301, Ph14-8Mo, and 17-7PH stainless steels, and Inco 718 nickel-base alloy.

The scatter characteristics of constant amplitude fatigue test results were statistically analyzed by using selected data [7, 10]. The data were selected on the following basis:

1. Only aluminum alloys were considered; these alloys were Types 2020, 2024, 2124, 6061, 7075, 7076, 7079, 7175, and 7178.
2. Test groups containing only two specimens were not included in this statistical analysis. These specimens were not considered to be part of the same population since two-specimen groups were sometimes tested in series. There

TABLE 1—Summary of spectrum-loaded test data.

Material	Loading Sequence	Initial No. of Test Groups	Initial No. of Specimens	No. of Test Groups with Zero Log Normal Standard Deviation	Statistically Analyzed Groups	No. of Analyzed Test Groups with		
						Hole-Notched Specimens	Built-up Structural Specimens	Lug Specimens
Aluminum	flight-by-flight	653	2191	34	619	343	257	19
	block	328	1454	3	325	163	147	15
Titanium	flight-by-flight	232	827	13	219	194	25	...
	flight-by-flight	42	141	0	42	38	4	...
Steel and Nickel	block	87	458	4	83	33	50	...

is generally less scatter associated with specimens tested in series than there is with two independently tested specimens.

3. Fatigue test data were not included for unnotched specimens. Fatigue failures generally do not occur in unnotched areas of aircraft structures, and, therefore, the scatter characteristics of unnotched data are not pertinent to these types of structures. Also the scatter for unnotched specimens has been shown to be different from that for notched or joint-type specimens.

4. The data for spot-welded and bonded specimens were not used because the fatigue characteristics of these specimens are significantly different from the corresponding characteristics for notched or mechanically fastened specimens.

5. Only axially loaded specimens were considered since this is the most representative loading for primary aircraft structures.

6. The only test groups considered were those containing all failures, that is, no runouts.

7. The fatigue data for test groups containing a specimen with a minimum life of less than 100 cycles were not included because such test lives are well below the operational life of aircraft structures.

8. The data obtained in mechanical shaker test machines were excluded because the fatigue test results from this type of equipment are generally not used to evaluate the fatigue characteristics of aircraft structures.

The final amount of $S-N$ data available for statistical analysis consisted of 604 test groups and 2683 specimens.

To see if there was any trend for scatter of constant-amplitude test data to increase with the mean fatigue life, the $S-N$ data were subdivided into 24 data sets of approximately 100 specimens each (91 minimum and 150 maximum). The average logarithmic standard deviation, that is, standard deviation of the log of life, and the geometric mean fatigue life was calculated for each data set consisting of 25 or 29 test groups. The results of these calculations are plotted in Fig. 2.

The average logarithmic standard deviation for the 24 data sets was 0.127. The last two data sets indicated a considerably higher average value of the logarithmic standard deviation than the first 22 sets. Therefore, the decision was made to exclude all $S-N$ test groups with mean test lives above a million cycles. The remaining 553 test groups with 2417 test specimens were statistically analyzed assuming that the logarithmic standard deviation is independent of fatigue life.

Regression Analysis of the Distribution of Logarithmic Standard Deviations

A series of linear regression analyses was performed by using an SPSS computer program [13] to determine if the standard deviation of the logarithm of fatigue life for a set of fatigue test groups would fit a normal, logarithmic, Weibull, or quadratic distribution function. These groups were tested under

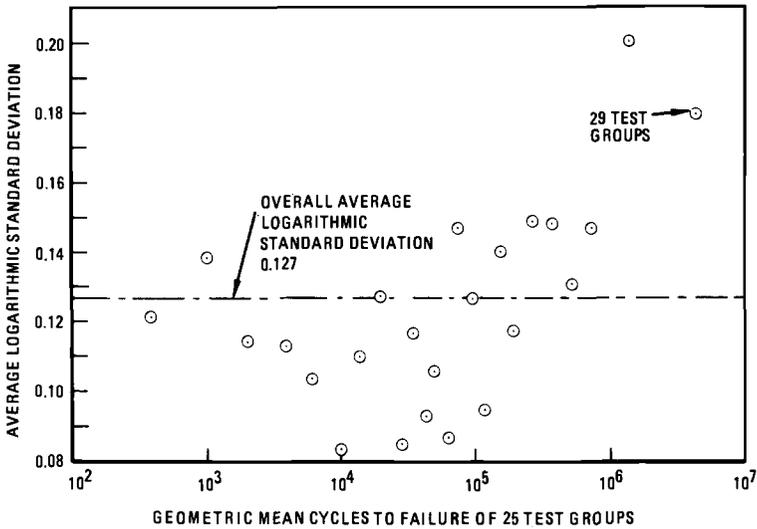


FIG. 2—Correlation of the average logarithmic standard deviation versus cycles to failure for 24 data sets consisting of 25 or 29 test groups.

either variable amplitude spectrum type or constant amplitude S - N type loadings. Each test group consists of a compatible set of essentially identical specimens subjected to virtually the same fatigue test loads. The standard deviation, based on the logarithms of all the fatigue failure lives in a test group, was calculated from the following equation

$$\text{standard deviation} = \left[\frac{1}{n-1} \sum_{i=1}^{i=n} \log X_i^2 - \left(\sum_{i=1}^{i=n} \log X_i \right)^2 \right]^{1/2} \quad (1)$$

with X_i being the i^{th} fatigue life failure in a test group composed of n specimens. To distinguish the standard deviation of log life from the standard deviation for the distribution of standard deviation, the standard deviation in Eq 1 will be referred to as logarithmic standard deviation.

The first analysis was made with the normal distribution function for logarithmic standard deviation and was based on the linear regression of the uniformly spaced normal deviate, z , with respect to compatible sets of logarithmic standard deviations. The value of z corresponds to the number of multiples of standard deviation in the normal distribution of logarithmic standard deviation by which an observed value is different from the median value of logarithmic standard deviation. Each value of z also represents an area under the curve of the normal distribution function, which is related to a specified probability of occurrence, P , of an observed value of logarithmic standard deviation.

The probability, P , was evaluated by ordering the values of logarithmic stan-

standard deviation sequentially from the lowest to the highest value. Each of the ordered values was assigned a count number which varied continuously from 1 to m , where m is defined as in Eq 2. The probability corresponding to each value of logarithmic standard deviation was then computed by using the following equation [14]

$$P = \frac{i - g}{m - (2g + 1)} \tag{2}$$

where

- i = count number,
- $g = 1/2$, and
- m = total number of logarithmic standard deviations being analyzed.

For this type of sequentially ordered data, the median value of logarithmic standard deviation has a count number of $[(m/2) - g]/m$ if m is odd. If m is even, the median is the average value of $[(m/2) - g]/m$ th and the $[(m/2) - g + 1]/m$ th ranked values of logarithmic standard deviation.

The linear regression analysis of the normal distribution of logarithmic standard deviations led to an equation of the form expressed in Eq 3

$$Y = A + BX \tag{3}$$

with

$$Y = \text{logarithmic standard deviation}$$

and

$$X = z$$

Equation 3 was also used for the linear regression of a logarithmic distribution with

$$Y = \log_{10} \text{ of the logarithmic standard deviation}$$

and

$$X = z$$

The linear equations and correlation coefficients, r , derived by this method are shown in Table 2. The correlation coefficient, r , for the logarithmic distribution was significantly larger than that for the normal distribution (which is not shown). The improvement in correlation occurs because multiplying by 2 the right-hand side of the linear logarithmic equations on this table would provide regression of the logarithmic variance, that is, the square of the logarithmic standard deviation, with Y being a logarithm to the base 10. In most instances, the variance is more normally distributed than the standard deviation, and the \log_{10} of the standard deviation follows the statistical trends of the variance.

The linear regression was conducted of the \log_{10} of the logarithmic standard deviation with respect to the Weibull variate, W , with the variables in Eq 3 given by

$$Y = \log_{10} \text{ of the logarithmic standard deviation}$$

TABLE 2—Summary of equations representing the probability

Item	Material	Type of Testing	No. of Test Groups	Linear Logarithmic Equation with X as the Variable	
				r	Equation
1	aluminum	flight-by-flight	290, n = 2	0.978	Y = -1.289 + 0.479X
2		flight-by-flight	329, n > 2	0.993	Y = -1.131 + 0.307X
3		block	325, n ≥ 2	0.981	Y = -1.187 + 0.363X
4		flight-by-flight and block	654, n ≥ 2 lines 2 + 3	0.986	Y = -1.158 + 0.337X
5	titanium	flight-by-flight	219, n ≥ 2	0.985	Y = -1.121 + 0.373X
6	steel and nickel	flight-by-flight	42, n ≥ 2	0.978	Y = -1.128 + 0.455X
7		block	83, n ≥ 2	0.978	Y = -0.946 + 0.329X
8		flight-by-flight and block	125, n ≥ 2 lines 6 + 7	0.982	Y = -1.007 + 0.384X
9	aluminum	S-N	553, n > 2	0.993	Y = -1.039 + 0.313X

^aKey:

n = number of specimens in a test group

Y = log₁₀ of the logarithmic standard deviation of fatigue test life

X = number of multiples of the standard deviation of the normal distribution of Y by which an observed value is different from the median value of Y (Note: the median value of Y corresponds to the geometric median value of the distribution of logarithmic standard deviation of fatigue test life.)

W = ln ln [1/(1 - P)]

P = probability of equalling or exceeding a specified value of Y

and

$$X = W = \ln \ln \left[\frac{1}{1 - P} \right]$$

where,

$$P = \text{value expressed in Eq 2}$$

Some improvements occurred in the coefficients of correlation, r, when the linear regression of the Weibull method was employed, as shown in Table 2. The resulting linear equations are also shown in Table 2.

Finally, polynomial regression analyses were performed by using an analytical expression of the type given in Eq 4 for both the logarithmic and Weibull distributions of the logarithm standard deviation.

$$Y = A + BX + CX^2 \tag{4}$$

The resulting quadratic equations are shown in Table 2.

The chi-squared test was conducted to determine the goodness of fit of the equations given in Table 2 to the data for the logarithmic standard deviations. Each of the nine data sets was subdivided into a number of cells, e. If the total

distribution of logarithmic standard deviation of fatigue test life.^a

2-deg Quadratic Logarithmic Equations with X as the Variable	Linear Weibull Equations with W as the Variable		2-deg Quadratic Weibull Equations with W as the Variable
	r	Equation	
$Y = -1.218 + 0.479X - 0.0710X^2$	0.998	$Y = -1.068 + 0.383W$	$Y = -1.072 + 0.401W + 0.00711W^2$
$Y = -1.109 + 0.307X - 0.0219X^2$	0.988	$Y = -0.993 + 0.239W$	$Y = -1.001 + 0.279W + 0.0158W^2$
$Y = -1.142 + 0.363X - 0.0453X^2$	0.996	$Y = -1.021 + 0.288W$	$Y = -1.025 + 0.306W + 0.00730W^2$
$Y = -1.122 + 0.337X - 0.0361X^2$	0.995	$Y = -1.005 + 0.265W$	$Y = -1.010 + 0.289W + 0.00930W^2$
$Y = -1.076 + 0.373X - 0.0449X^2$	0.994	$Y = -0.951 + 0.295W$	$Y = -0.958 + 0.327W + 0.0131W^2$
$Y = -1.102 + 0.455X - 0.0264X^2$	0.964	$Y = -0.926 + 0.355W$	$Y = -0.950 + 0.447W + 0.0409W^2$
$Y = -0.906 + 0.329X - 0.0403X^2$	0.993	$Y = -0.795 + 0.262W$	$Y = -0.798 + 0.272W + 0.00393W^2$
$Y = -0.960 + 0.384X - 0.0474X^2$	0.992	$Y = -0.832 + 0.305W$	$Y = -0.839 + 0.338W + 0.0136W^2$
$Y = -1.017 + 0.313X - 0.0213X^2$	0.989	$Y = -0.898 + 0.244W$	$Y = -0.905 + 0.280W + 0.0144W^2$

number of data points was less than 200 (Items 6, 7, and 8), the cell size was determined by dividing the total number of points by 5. For more than 200 points the cell size was determined from

$$e = 4 [0.75(m - 1)^2]^{1/5} \tag{5}$$

where m = total number of data points. The results of the chi-squared goodness of fit test are summarized in Table 3. The degrees of freedom for each data set were calculated from

$$\nu = e - 1 - r \tag{6}$$

where r = number of coefficients used to define the equations given in Table 2.

An equation is considered a good fit to the test data at the 5-percent significance level, that is, when there is a 5 percent probability of obtaining a chi-squared variate greater than the tabulated value. Only data Sets 7 and 9 are acceptable for the linear logarithmic distribution. The linear Weibull distribution would be accepted for data Sets 1, 3, 5, 7, and 8. For data Set 7 the linear Weibull distribution provides a better fit than the linear logarithmic distribution. Except for data Set 7, the quadratic equations in X or W provide better fits to the data than the linear equations. The degree of fit for data Set 7 is about the same for both the linear and quadratic equations.

None of the equations provided a good fit to data Set 6. However, when data Set 6 was combined with data Set 7 (data Set 8), a good fit was obtained for both the linear and quadratic equations in W .

TABLE 3—Summary of chi-squared goodness of fit test for equations used to represent probability distribution of $\log_{10} \sigma_s$.

Item	Material	No. of Test Groups	Type of Testing	No. of Cells	Degrees of ν^a	Linear Logarithmic Equations		Quadratic Logarithmic Equations		Linear Weibull Equations		Quadratic Weibull Equations	
						Chi-Square Variate %	Probability of Obtaining Chi-Squared Variate %	Chi-Square Variate %	Probability of Obtaining Chi-Squared Variate %	Chi-Square Variate %	Probability of Obtaining Chi-Squared Variate %	Chi-Square Variate %	Probability of Obtaining Chi-Squared Variate %
1	aluminum	290	flight-by-flight	36	33, 32	81.42	<0.05	39.21	10 to 20	36.73	20 to 30	30.03	50 to 60
2	aluminum	329	flight-by-flight	38	35, 34	50.42	2.5 to 5.0	43.26	10 to 20	65.21	0.1 to 0.5	52.5	1.0 to 2.5
3	aluminum	325	block	38	35, 34	54.18	1.0 to 2.5	39.45	20 to 30	42.72	10 to 20	29.86	60 to 70
4	aluminum	654	flight-by-flight and block	50	47, 46	97.99	<0.05	60.68	5 to 10	99.98	<0.05	63.43	2.5 to 5.0
5	titanium	219	flight-by-flight	33	30, 29	46.96	2.5 to 5.0	24.96	60 to 70	33.40	30 to 40	24.66	60 to 70
6	steel and nickel	42	flight-by-flight	8	5, 4	14.38	1.0 to 2.5	12.86	1.0 to 2.5	39.14	<0.05	13.24	1.0 to 2.5
7	steel and nickel	83	block	17	14, 13	20.43	10 to 20	15.52	20 to 30	11.83	60 to 70	11.42	50 to 60
8	steel and nickel	125	flight-by-flight and block	25	22, 21	38.80	1.0 to 2.5	34.40	2.5 to 5.0	20.40	50 to 60	16.40	70 to 80
9	aluminum	553	S-N	47	44, 43	51.88	20 to 30	42.19	50 to 60	69.05	1.0 to 2.5	52.39	10 to 20

^a ν = linear, quadratic.

To get an idea of the meaning of goodness of fit, the test data are compared with the distribution equations in Figs. 3, 4, and 5 for data Sets 2, 5 and 9. It is obvious from Fig. 3 that the quadratic equation in X provides the best fit of those shown and that the other two equations are not acceptable by the chi-squared test. The logarithmic distribution plots as a straight line in Fig. 3 where the X scale is linear for z . In Fig. 4, the Weibull distribution is acceptable, whereas the logarithmic distribution is not. The Weibull distribution plots as a straight line in Fig. 4, where the W scale is linear for $\ln \ln [1/(1 - P)]$. Finally, the quadratic distribution in X provides a good fit to the data in Fig. 5 where the logarithmic and Weibull distributions are unacceptable. Also the X scale is linear for z in Fig. 5.

The values obtained for the logarithmic standard deviations from the four equations for each data set are summarized in Table 4 for the 95, 90, and 50 percent probability values. There is not much difference in the 50 percent probability values obtained for each distribution; however, there is a significant difference near the tails of the distribution. Except for data Set 6, the Weibull distribution yields the lowest calculated values for σ_β and the logarithmic distribution the highest values, with the X and W quadratic distributions falling in between for probability values greater than 90 percent. When equations are used for calculating the value of σ_β , the equation providing the best fit, as indicated by the chi-squared test in Table 3, was used.

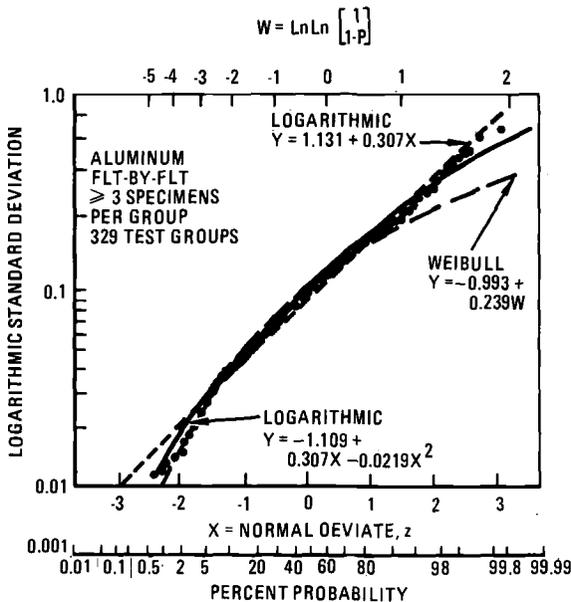


FIG. 3—Comparison of probability distribution equations with data for aluminum flight-by-flight test of a group with three or more specimens.

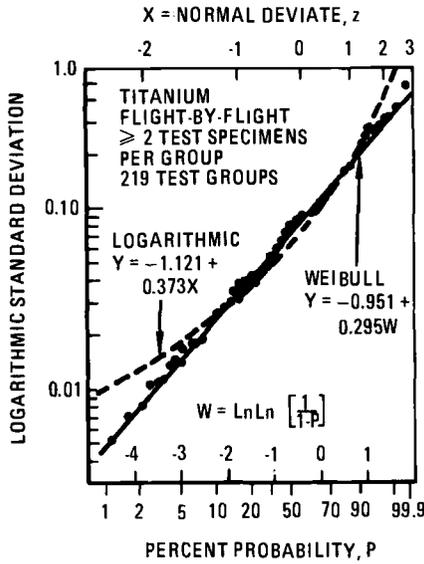


FIG. 4—Comparison of probability distribution equations with data for titanium flight-by-flight tests.

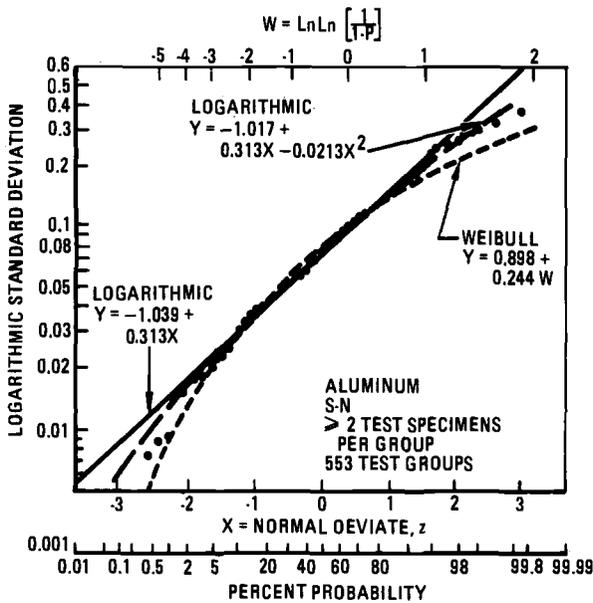


FIG. 5—Comparison of probability distribution equations for aluminum S-N tests of a group size of three or more.

TABLE 4—Summary of calculated value of logarithmic standard deviation for various levels of probability.

Item	Type-of-Test Data	Prob-ability, %	Logarithmic Standard Deviation			
			Linear Equation in X	Quadratic Equation in X and X^2	Linear Equation in W	Quadratic Equation in W and W^2
			1	flight-by-flight aluminum, 2 specimens/group	50 90 95	0.05 0.21 0.32
2	flight-by-flight aluminum, > 2 specimens/group	50 90 95	0.07 0.18 0.24	0.08 0.18 0.22	0.08 0.16 0.19	0.08 0.17 0.21
3	block spectrum aluminum, ≥ 2 specimens/group	50 90 95	0.07 0.19 0.26	0.07 0.18 0.22	0.07 0.17 0.20	0.07 0.17 0.21
4	Items 2 and 3 spectrum loading aluminum, ≥ 2 specimens/group	50 90 95	0.07 0.19 0.25	0.08 0.18 0.22	0.08 0.16 0.19	0.08 0.17 0.21
5	flight-by-flight titanium, ≥ 2 specimens/group	50 90 95	0.08 0.23 0.31	0.08 0.21 0.26	0.09 0.20 0.24	0.08 0.21 0.26
6	flight-by-flight steel and nickel ≥ 2 specimens/group	50 90 95	0.07 0.28 ① ^a	0.08 0.33 ①	0.09 0.23 ①	0.08 0.28 ①
7	block spectrum steel and nickel ≥ 2 specimens/group	50 90 95	0.11 0.30 0.39	0.12 0.28 0.34	0.13 0.27 0.31	0.13 0.27 0.32
8	Items 6 and 7 spectrum loading steel and nickel > 2 specimens/group	50 90 95	0.10 0.31 0.42	0.11 0.28 0.35	0.11 0.26 0.32	0.11 0.28 0.35
9	$S-N$ aluminum, > 2 specimens/group	50 90 95	0.09 0.23 0.30	0.10 0.22 0.28	0.10 0.20 0.23	0.10 0.22 0.26

^a① is beyond the range of data.

The results summarized in Table 4 indicate that aluminum alloys have the least scatter, titanium somewhat higher, and the combined data for steel and nickel-base alloys the largest. The scatter characteristics for the various aluminum and titanium alloys were similar. However, the scatter for the various steel and nickel materials may be a function of material strength, as indicated by results reported by Whittaker [9]. Therefore, the scatter for the steel materials summarized in Table 4 should not be considered typical of all steel and nickel alloys. More spectrum fatigue data are needed for steel alloys to determine the relation between fatigue scatter and material strength level.

Comparison of Scatter Factor Equations

The scatter factors were calculated by using Eq 7 [6], 8 [15,16], and 9 [17,18], as expressed here for the various sets of data, and are summarized in Table 5 for reliability values of 0.90 and 0.95 and a confidence level of 0.95. The appropriate values of the logarithmic standard deviation used in the calculations are also given in Table 5.

$$S_{p,c} = \text{antilog } \sigma \left[\frac{z_{c1}}{\sqrt{n}} + z_p \right] \tag{7}$$

σ is determined from evaluating scatter of available appropriate test data and assuming that sample variances, S_i^2 , are from populations with equal variances, that is, $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2$. Then an unbiased estimate of the population standard deviation is given by

$$\sigma = \sigma_s = \left[\frac{\sum_{i=1}^k S_i^2 (n_i - 1)}{\sum_{i=1}^k n_i - k} \right]^{1/2}$$

where

$$s^2 = \sum_{j=1}^n \frac{(l_j - \bar{l})^2}{n - k} \text{ for each sample}$$

$$S_{p,c} = \text{antilog } \sigma_{c2} \left[\frac{z_{c1}}{\sqrt{n}} + z_p \right] \tag{8}$$

where

$$c = (c1)(c2)$$

σ_{c2} is determined from evaluating scatter of available appropriate test data assuming that sample variances, S_i^2 , are from populations with equal variances, that is

$$\sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2.$$

Then S_i^2/σ^2 is a chi-squared distribution with $n - k = \nu$ degrees of freedom, therefore, σ_{c2} is given by

$$\sigma_{c2} = \frac{\sigma_s}{(\chi^2/\nu)_{c2}^{1/2}}$$

TABLE 5—Calculated values of scatter factor for aluminum with $R = 0.90$ and 0.95 , $c = \beta = 0.95$, and $n = 1$.

Item	Type of Test	No. of Test Groups	Logarithmic Standard Deviation				Scatter Factors for $R = 0.90$ and $c = \beta = 0.95$			Scatter Factors for $R = 0.95$ and $c = \beta = 0.95$		
			σ_s	σ_c , $c2 = 0.975$	σ_β , $\beta = 0.95$	σ_μ	Eq 7, $c1 = 0.95$	Eq 8, $c1 = 0.975$	Eq 9	Eq 7, $c1 = 0.95$	Eq 8, $c1 = 0.975$	Eq 9
1	flight-by-flight	290	0.12	0.130	0.24	0.24	2.25	2.64	2.72	2.48	2.94	3.62
2	flight-by-flight	329	0.11	0.116	0.22	0.22	2.10	2.38	2.51	2.30	2.62	3.25
3	block	325	0.12	0.125	0.21	0.21	2.25	2.54	2.40	2.48	2.82	3.08
4	flight-by-flight and block	654	0.11	0.113	0.22	0.22	2.10	2.32	2.51	2.30	2.55	3.25
5	flight-by-flight	219	0.13	0.138	0.26	0.26	2.40	2.80	2.96	2.68	3.14	4.03
8	flight-by-flight and block	125	0.17	0.181	0.35	0.35	3.14	3.86	4.31	3.62	4.49	6.52
9	S-N	553	0.15	0.155	0.28	0.28	2.75	3.18	3.22	3.12	3.62	4.48

for very large sample sizes the confidence limits are small and $\sigma_{c2} \approx \sigma$. If sample variances, σ_i^2 , do not fit a chi-squared distribution, then it cannot be assumed that σ_s came from populations with equal variances [6]. Then σ_s is replaced by σ_β and

$$S_{p.c} = \text{antilog } z_p \sigma_\beta \left[1 + \frac{1}{n} \right]^{1/2} \quad (9)$$

The scatter factors determined from Eq 7 are lower than those determined by either Eq 8 or 9. This is to be expected since the unbiased estimate (50 percent confidence) of the logarithmic standard deviation is used in Eq 7. Except for the block loading tests at $R = 0.90$, the scatter factors from Eq 9 are the largest.

The values of S_i^2/σ^2 should form a χ^2/ν distribution if the variances of a set of fatigue test groups of the same sample size all have the same value of population variance, σ^2 . Such a distribution would form a linear plot on chi-squared paper with a degree of freedom equal to the sample size minus 1. In Fig. 6 two sets of S_i^2/σ^2 values for test groups of sample size 4 are plotted on chi-squared paper with 3 degrees of freedom for aluminum $S-N$ and spectrum fatigue test data. Since neither set of the plotted values is linear, it cannot be concluded that either set of test groups is from populations with equal variance. Similar plots for sets of test groups with other sample sizes of 5, 6, and 7 specimens were also examined, and none of the sets formed linear plots on the appropriate chi-squared paper. Therefore, Eq 9 is recommended for general use to determine the reliability and confidence levels associated with a specified life based on a limited number of tests conducted on full-scale structures.

Conclusions

No single distribution function provided the best fit to the standard deviation of the log of test life for all of the analyzed sets of test data. The two-parameter Weibull distribution provided the best fit to the logarithmic standard deviations for the majority of the spectrum test data between probabilities of 10 and 90 percent. The logarithmic distribution of these standard deviations provided the best fit for $S-N$ test data in the same range of probabilities. Below and above probabilities of 10 and 90 percent, respectively, the corresponding 2-deg polynomial equations provide the best correlation.

This investigation indicates that aluminum alloys have the least scatter, with the scatter for titanium alloys a little higher, and that for the combined data for steel and nickel base alloys the largest. Because the set of data is small and the scatter for the various steel and nickel alloys may be a function of material strength, these findings should not be considered typical of all steel and nickel alloys.

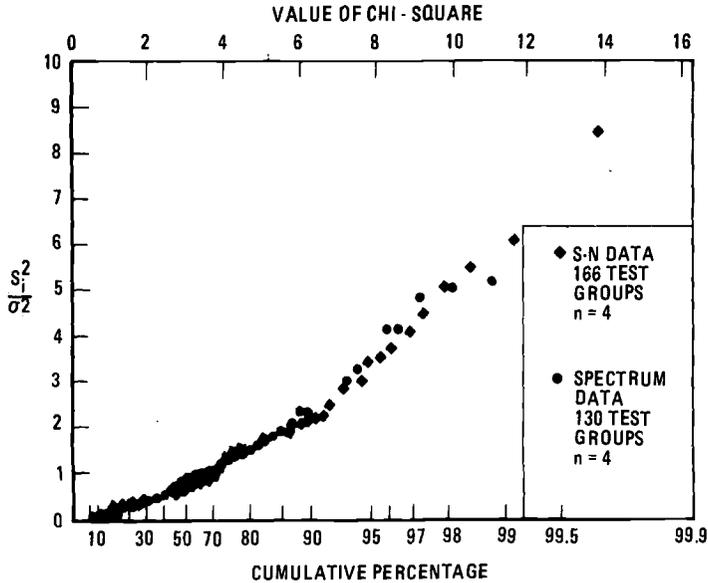


FIG. 6—Chi-squared probability plot for 3 degrees of freedom of aluminum S-N and spectrum data of sample size 4.

Equations 7 and 8 assume that the test sample variances are from populations with equal variances. The variance of logarithmic fatigue life for fatigue test groups does not seem to fit a chi-squared distribution as it would if the test groups were from populations with the same variance. Therefore, the use of Eq 9 appears to be the logical choice for determining test life reduction factors.

References

- [1] Stagg, A. M., "An Investigation of the Scatter in Constant Amplitude Fatigue Test Results of Aluminium Alloys 2024 and 7075," C.P. No. 1093, Aeronautical Research Council, London, England, April 1969.
- [2] Stagg, A. M., "An Investigation of the Scatter in Variable Amplitude Fatigue Test Results of 2024 and 7075 Materials," C.P. No. 1123, Aeronautical Research Council, London, England, May 1969.
- [3] Stagg, A. M., "Scatter in Fatigue-Elements and Sections from Aircraft Structures," RAE Technical Report 69155, Royal Aeronautical Establishment, Farnborough, England, 1969.
- [4] Schijve, J. and Jacobs, F. A., "Program Fatigue Tests on Notched Light Alloy Specimens of 2024 and 7075 Material," NRL-TR M. 2070, National Aeronautical Research Institute, Amsterdam, the Netherlands, 1960.
- [5] Davis, C. S. and Young, L., "A Comparison of Fatigue Life and Reliability from Constant and Variable Amplitude Loading Tests," *Proceedings, Twenty-third Annual National Forum, American Helicopter Society, Inc.*, Washington, D.C., 10-12 May 1967.
- [6] Abelkis, P. R., "Fatigue Strength Design and Analysis of Aircraft Structures, Part I. Scatter Factors and Design Charts," AFFDL-TR-66-197, Part I, Air Force Flight Dynamics Laboratory, Wright-Patterson Air Force Base, Dayton, Ohio, June 1967.
- [7] Whittaker, I. C. and Besuner, P. M., "A Reliability Analysis Approach to Fatigue Life

- Variability of Aircraft Structures," AFML-TR-69-65, Air Force Material Laboratory, Wright-Patterson Air Force Base, Dayton, Ohio, April 1969.
- [8] Jost, G. S. and Verinder, F. E. "A Survey of Fatigue Life Variability in Aluminum Alloy Aircraft Structures," Report SM. 329, Aeronautical Research Laboratories, Department of Supply, Melbourne, Australia, Feb. 1971.
- [9] Whittaker, I. C. "Development of Titanium and Steel Fatigue Variability Model for Application of Reliability Analysis Approach to Aircraft Structures," AFML-TR-72-236, Air Force Material Laboratory, Wright-Patterson Air Force Base, Dayton, Ohio, Oct. 1972.
- [10] Impellizzeri, L. F., Siegel, A. E., and McGinnis, R. A., "Evaluation of Structural Reliability Analysis Procedures as Applied to a Fighter Aircraft," AFML-TR-73-150, Air Force Material Laboratory, Wright-Patterson Air Force Base, Dayton, Ohio, Sept. 1973.
- [11] Young, L., Davis, C. S., and McCulloch, A. J., "An Evaluation of Fatigue Test Life Reduction Factors for Spectrum Loading Tests—Part II—Fatigue Test Data," Lockheed Report LR 19024, Lockheed Co., 11 Aug. 1965.
- [12] Young, L. and Ekvall, J. C., "Reliability of Fatigue Testing," Lockheed Report LR 29270, Lockheed Co., Dec. 1979.
- [13] Nie, N. H., Hull, C. H., Venkins, J. G., Steinbenner, K., and Bent, D. H., *Statistical Package for the Social Sciences*. 2nd ed., McGraw-Hill, New York, 1975.
- [14] Cochran, W. G., *Annals of Mathematical Statistics*, Vol. 23, 1952, pp. 315-345.
- [15] Kaechele, L., "Probability and Scatter in Cumulative Fatigue Damage," Memo RM-3688-PR, Rand Corp., Santa Monica, Calif., Dec. 1963.
- [16] Butler, J. P., *International Conference on Structural Safety and Reliability*. A. M. Freudenthal, Ed., Pergamon Press, New York, 1972, pp. 181-211.
- [17] Impellizzeri, L. F., *Structural Fatigue in Aircraft*. ASTM STP 404, American Society for Testing and Materials, Philadelphia, 1966.
- [18] McCulloch, A. J. and Walsh, J. E., *Journal of the American Statistical Association*, Vol. 62, March 1967, pp. 45-47.

Statistical Fatigue Properties of Some Heat-Treated Steels for Machine Structural Use

REFERENCE: Nishijima, S., "Statistical Fatigue Properties of Some Heat-Treated Steels for Machine Structural Use," *Statistical Analysis of Fatigue Data, ASTM STP 744*, R. E. Little and J. C. Ekvall, Eds., American Society for Testing and Materials, 1981, pp. 75-88.

ABSTRACT: A practical method for establishing the P - S - N diagram is proposed. The method is essentially based on consideration of the distribution of strength deviation values for individual test data determined against the mean S - N curve of the population. Thus an S - N curve for 1 percent failure probability can be derived from the fatigue test results using 100 specimens.

A series of statistically planned fatigue tests was conducted according to the proposed method, in order to obtain basic fatigue data about different materials most commonly used in mechanical industries. This paper deals with the results for some typical carbon and low-alloy steels with different heat treatments, the tests being on smooth specimens in rotating bending, and discusses the variation in statistical properties between materials.

KEY WORDS: fatigue, statistical analysis, statistical properties, P - S - N diagrams, fatigue strength, fatigue strength variation, rotating bending fatigue tests

Statistical fatigue properties of metallic materials have been recognized as one of the important classes of information needed for reliable design of machines and structures that experience fluctuating loads during use. Many investigations have been conducted that reveal fundamental statistical characteristics of fatigue failure, which have resulted in the development of various statistical approaches to laboratory fatigue testing [1-4].²

In fatigue tests, the applied stress (or strain), S , is an independent variable, and the number of cycles to failure, N , is a dependent variable described by a certain distribution function whose shape changes according to S . The scatter in fatigue life at a prescribed stress level is generally studied as the relation

¹Head, First Laboratory, Fatigue Testing Division, National Research Institute for Metals, Tokyo, 153 Japan.

²The italic numbers in brackets refer to the list of references appended to this paper.

between N and the proportion of failed specimens, P , prior to the N cycles. The scatter in N is believed to be due mainly to the variation in the properties of each specimen, differences in the test conditions, and uncertainty in the fracture process itself.

Another concept for scatter in fatigue is the P - S relation, in which the percentage of failed specimens, P , at different levels of S is considered by prescribing the number of cycles, N . This is the same as examining the distribution of specimens whose life is coincident with N , and in such a case the value of S can be regarded as the strength of the specimens. The scatter in S is believed to have the same cause as that in N .

It is known that these two concepts of scatter in fatigue, life scatter (P - N) and strength scatter (P - S), can be adequately expressed on the P - S - N diagram, a family of S - N curves each corresponding to a particular value of failure probability, P . An ordinary way of establishing such a P - S - N diagram for a given material is to determine experimentally the P - N relations at different S values and draw S - N curves using N values expected for the P prescribed. Obviously a hundred tests are needed to get N values at $P = 1$ percent for only a single value of S , and thus nearly a thousand tests are needed to determine a complete P - S - N set up to $P = 1$ percent.

The aim of the present paper is to give a more practical method of producing P - S - N diagrams with an accessible number of specimens and to show some typical results for several of the heat-treated steels most popularly employed in mechanical industries. The data presented here are a part of the series of studies undertaken at the National Research Institute for Metals (NRIM), Tokyo, Japan, with the view of providing engineers with the basic statistical fatigue data for current structural materials. Detailed and more comprehensive results will be found in the references: the statistical fatigue properties for carbon and chromium-molybdenum steels, including their heat-to-heat variations [5]; for differently heat-treated carbon, chromium-molybdenum, nickel-chromium-molybdenum, and Type 403 steels, showing the effect of a notch [6]; for Types 5083 and 7075 aluminium alloys, also showing the effect of a notch [7]; and for butt-welded joints of high-strength steels, demonstrating the effect of stress concentrations due to reinforcement [8].

Theory

Figure 1 demonstrates a typical data scatter in the fatigue of smooth specimens of a ferrite-pearlite carbon steel, S25C (see later paragraphs for experimental details). One can easily recognize the two-way distributions schematized in the figure—those in life and those in strength. The shape of the life distribution is strongly dependent on the stress level, but the strength distribution appears to show similarities at different life levels, as was shown experimentally in some previous works [9, 10].

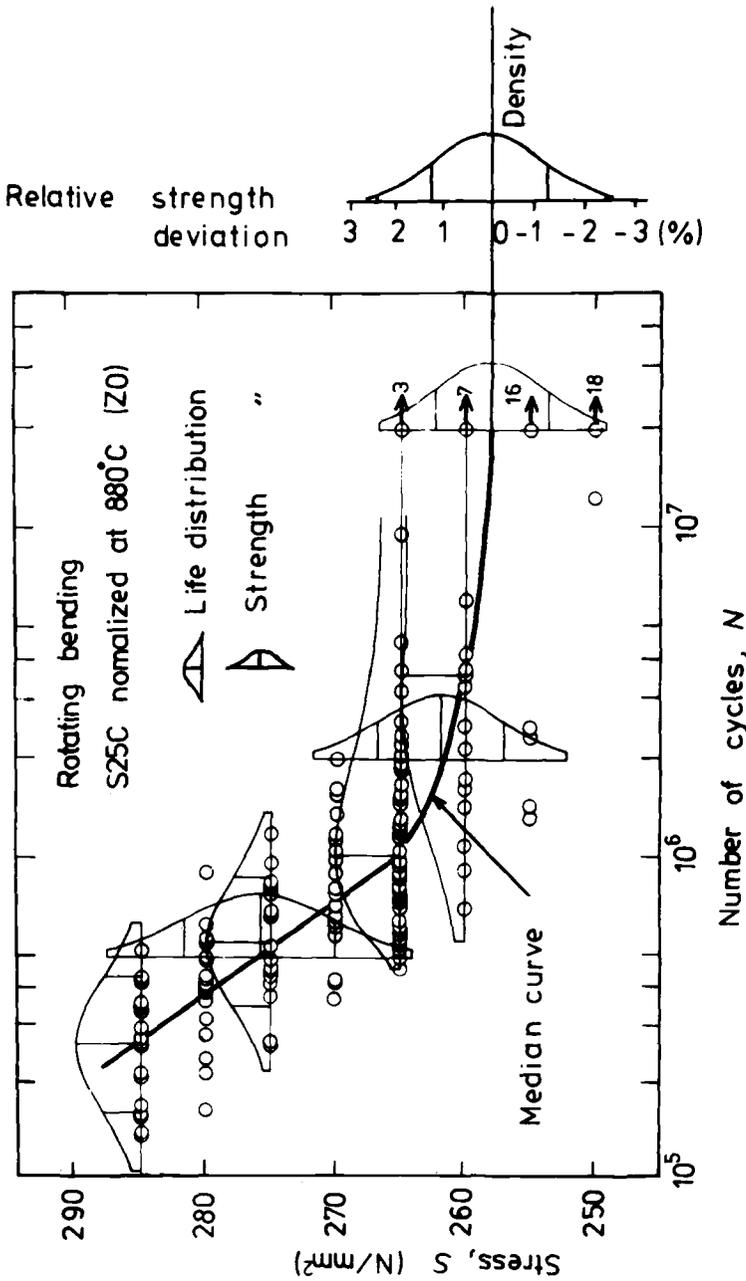


FIG. 1—Typical data scatter in fatigue test results.

Let a specimen have a life N_j at a stress S_i . On the basis of the strength distribution at N_j , the strength inherent in the specimen is reflected in the deviation from the mean, $S_i - S_m(N_j)$, where $S_m(N_j)$ designates the mean strength of the material at N_j . If we assume that the deviation in strength for different N_j values varies in proportion to S_m , or, in other words, that the coefficient of variation in strength is constant, we can consider for many specimens the set of relative strength deviation values

$$r_{ij} = \frac{S_i}{S_m(N_j)} - 1 \quad (1)$$

as a statistical variable reflecting the strength distribution of the material. For the example in Fig. 1, the distribution of r_{ij} is illustrated on the right-hand side of the figure.

Two questions now arise: how to determine $S_m(N_j)$ and how to examine the distribution of r_{ij} .

One of the means of determining S_m at a prescribed N_j is probit analysis using weighting, based on the normal distribution law [11]. Introducing a normalized variable, x_i , using the standard deviation, s , where

$$x_i = \frac{S_i - S_m}{s} \quad (2)$$

the density and the probability of failure at N_j are, respectively

$$\phi(x_i) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x_i^2}{2}\right) \quad (3)$$

and

$$P_i = \psi(x_i) = \int_{-\infty}^{x_i} \phi(x) dx \quad (4)$$

The values of x_i in Eq 2 can be related to observed values of P_i using the inverse function of ψ , as

$$x_i = \psi^{-1}(P_i) = \frac{1}{s} S_i - \frac{S_m}{s} \quad (5)$$

Since Eq 5 represents a linear relation between the test stress, S_i , and the observed value, $x_i = \psi^{-1}(P_i)$, we can estimate S_m and s from experimental data by an ordinary least squares method. The computation is done taking into account the weighting for x_i ,

$$w_i = \frac{n_i \{ \phi(x_i) \}^2}{\psi(x_i) \{ 1 - \psi(x_i) \}} \quad (6)$$

where n_i is the number of tests performed at S_i . An iterative process is necessary because x_i in Eqs 5 and 6 is also influenced by S_m and s . The practice for the calculations is found in Ref 11. Finally, the solutions are of the form

$$S_m = \frac{\sum w_i S_i}{\sum w_i} - \frac{\sum w_i x_i}{\sum w_i} s \quad (7)$$

$$s = \frac{\sum w_i \cdot \sum w_i S_i^2 - (\sum w_i S_i)^2}{\sum w_i \cdot \sum w_i x_i S_i - \sum w_i x_i \cdot \sum w_i S_i} \quad (8)$$

and the standard error in S_m is

$$e = s \sqrt{\frac{1}{\sum w_i} + \frac{(S_m - \sum w_i S_i / \sum w_i)^2}{\sum w_i S_i^2 - (\sum w_i S_i)^2 / \sum w_i}} \quad (9)$$

With these calculations being done for various N_j values, the best-fit mean S - N curve can be computed numerically, using $1/e^2$ as the weighting for S_m . Figure 1 shows the best-fit polynomial curve, avoiding any prejudice as to the shape of the S - N curve, which agrees very well with the observed values of median life at different stress levels.

It now becomes possible to examine the distribution of the relative strength deviation values (Eq 1) by regarding all of them as a set of order statistics. The cumulative probability corresponding to a specimen of rank number J is evaluated in Eq 10, the approximated equation for median rank plotting

$$P = \frac{J - 0.3}{M + 0.4} \quad (10)$$

where M is the total number of tests. If there are some data from unfailed specimens, the rank value should be calculated according to the theory of multiply censored order statistics [12], as

$$J = J_0 + \frac{M + 1 - J_0}{M + 2 - K} \quad (11)$$

In this equation, K is the rank number of data counted regardless of the failure or nonfailure of specimens, and J_0 is the rank value attributed to the failed datum immediately before the K^{th} under consideration. The density distribution of the relative strength deviation demonstrated in Fig. 1 was obtained in this way.

The relative strength distribution having been determined, a complete P - S - N diagram can be composed easily by shifting the mean S - N curve according to the deviation values for the prescribed failure probabilities pre-

dicted from the distribution. Thus, only a hundred specimens will be enough to obtain a $P-S-N$ set up to $P = 1$ percent.

Experimental Work

The materials studied were sampled from those commercially available as hot-rolled bars of 19 to 25 mm in diameter, which conformed to the Japanese Industrial Standard (JIS). Here the results for 0.25C steel, 0.45C steel, 0.35C-1Cr-0.2Mo steel, 0.39C-1.7Ni-1Cr-0.2Mo steel, and 0.15C-0.5Si-12Cr steel will be discussed.

Table 1 gives the chemical composition from the check analysis. Different heat treatments were performed at NRIM, using salt baths, to investigate the effect of heat treatment on the scatter in fatigue: these treatments were normalizing and high and low-temperature tempering after quenching.

Fatigue tests were conducted at NRIM using 20 rotating bending machines of the uniform bending moment type, at 3000 rpm, conforming to the JIS testing standard (Z 2274). Atmospheric conditions in the testing room were not severely regulated; some informal data are as follows: temperature, 13 to 27°C; humidity, 35 to 85 percent; sulfur dioxide (SO₂), 0.03 ppm, and oxides of nitrogen (NO_x), 0.08 ppm. The specimens were of a smooth cylindrical type with 8-mm diameters, finished by longitudinal polishing with 600-grade waterproof silicon carbide papers. The number of specimens used in a series of tests was variable but more than 100.

All the processes of sampling, heat treatments, and the execution of tests were carefully controlled, because of statistical considerations in order to avoid any intrusion of extraneous effects on the resulting variation in fatigue properties. In this regard, Vickers hardness was measured on the shank of each specimen before testing, verifying that the hardness distribution obeys a normal distribution law. An example is displayed in Fig. 2 for S25C steel.

Results

Table 2 lists the mean and coefficient of variation (percentage of estimated standard deviation to mean) for each material, obtained from 10 to 20 ten-

TABLE 1—Chemical composition (weight percent).

JIS Designation	Heat Code	C	Si	Mn	P	S	Cu	Cr	Ni	Mo
S25C	Z	0.28	0.30	0.46	0.026	0.015	0.043	0.039	0.018	...
S45C	W	0.46	0.28	0.81	0.014	0.016	0.058	0.058	0.022	...
SCM435	D	0.35	0.24	0.72	0.016	0.017	0.14	0.96	0.070	0.17
SNCM439	Q	0.40	0.25	0.68	0.023	0.020	0.11	0.71	1.71	0.16
SUS403	O	0.14	0.27	0.44	0.022	0.018	...	11.76	0.16	...

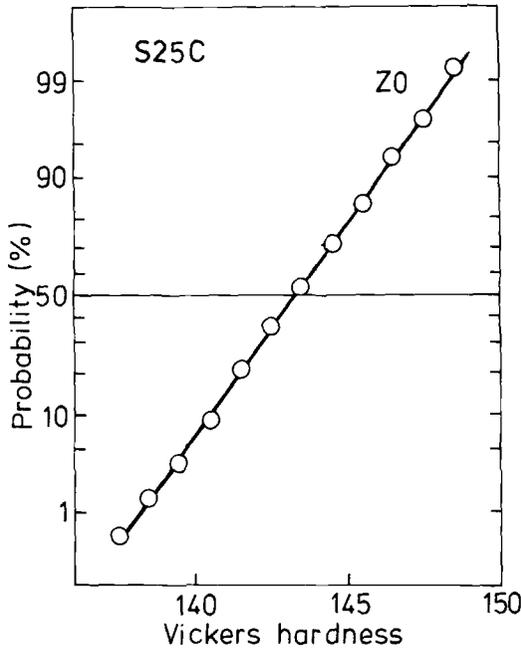


FIG. 2—Example of the Vickers hardness distribution of specimens.

sion tests. The results of the Vickers hardness measurements, as well as the fatigue strength at 10^7 cycles, are also tabulated in the same manner for comparison. The difference in heat treatment and, therefore, in microstructure of the materials is identified by a numeric code, as explained in the footnote to the table.

For the results of the fatigue tests, the values of S_m were estimated by probit analysis at 0.1 intervals in $\log N$. This resulted normally in 20 to 30 plots of S_m-N data, which could then be used to compute the mean $S-N$ curve, as the best-fit polynomial equation. Figure 3 compares, on normal probability paper, the distribution shapes for the relative strength deviation values (Eq 1) for different materials. Obviously, the distributions appear almost normal in these cases.

In Fig. 4 are displayed $P-S-N$ diagrams, each as a set of $S-N$ curves for prescribed failure probabilities. The curves labeled 1, 10, 90, or 99 percent were obtained by shifting laterally the mean $S-N$ curve by the amount, S , which was equivalent to the corresponding relative deviation value in Fig. 3. As can be seen in each diagram in Fig. 4, the distribution of data points is satisfactorily expressed by the curves, which confirms that the method of analysis was applied successfully. It should be particularly noted that in the case of SCM435 steel tempered at 450°C , the $S-N$ curves fall again for cycle

TABLE 2—Mechanical properties given as mean (top) and coefficient of variation in percent (bottom).

JIS Designation	Test Code ^a	Yield Stress, N/mm ²	Tensile Strength, N/mm ²	Elongation, %	Reduction of Area, %	Vickers Hardness	Fatigue Strength at 10 ⁷ Cycles N/mm ²
S25C	Z0	332	506	38.5	60.9	143	259
		1.5	1.7	3.2	1.9	1.6	1.7
S45C	W0	450	689	30.5	54.3	198	314
		2.1	1.3	1.9	1.5	1.0	1.7
		667	811	25.5	63.1	270	421
	W1	3.3	1.0	3.1	1.1	1.4	2.1
		923	1041	16.6	61.3	331	598
SCM435	D1	3.7	1.8	5.4	1.5	0.9	2.4
		1243	1373	11.4	51.1	434	698
	D2	6.8	4.6	9.6	3.9	1.1	2.6
		882	974	19.7	63.9	314	539
SNM439	Q1	0.7	0.4	8.6	0.9	1.1	2.7
		461	653	26.9	73.3	219	385
SUS403	O1	1.3	0.9	4.8	0.5	0.7	0.9

^aNumeric code represents the types of heat treatment and microstructure:

- 0 = normalized, presenting ferrite-pearlite structure
- 1 = quench-tempered, high-temperature-tempered martensite
- 2 = quench-tempered, low-temperature-tempered martensite

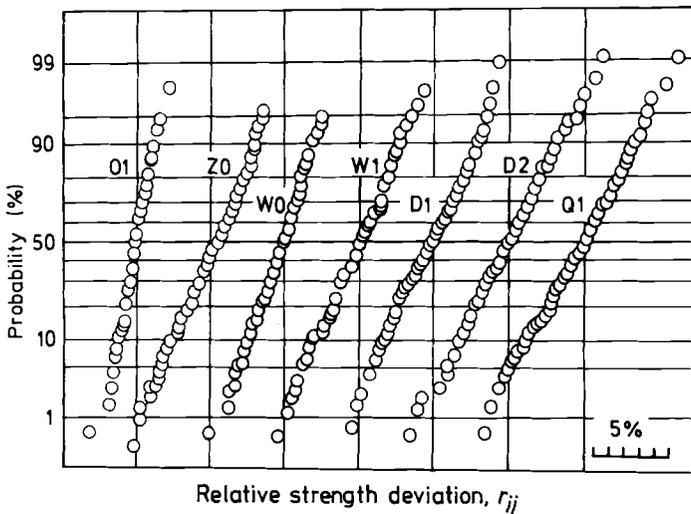


FIG. 3—Distribution of relative fatigue strength deviation (normal probability paper).

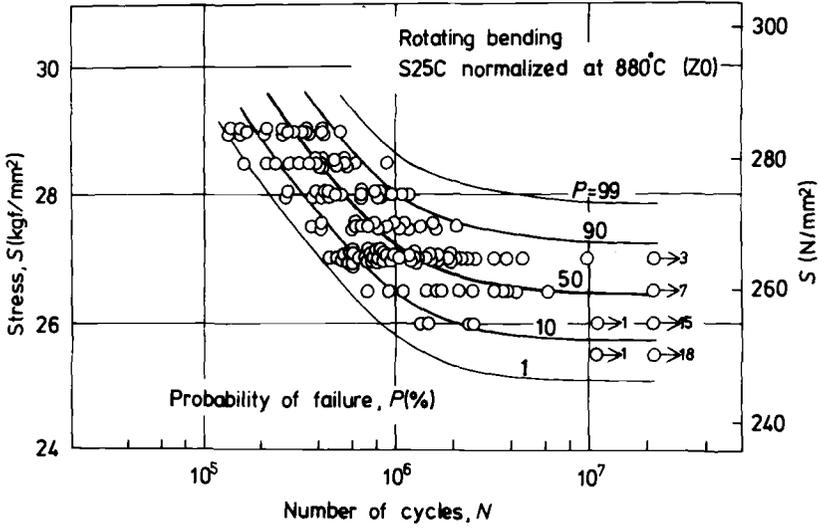


FIG. 4a—P-S-N diagram for S25C steel normalized at 880°C.

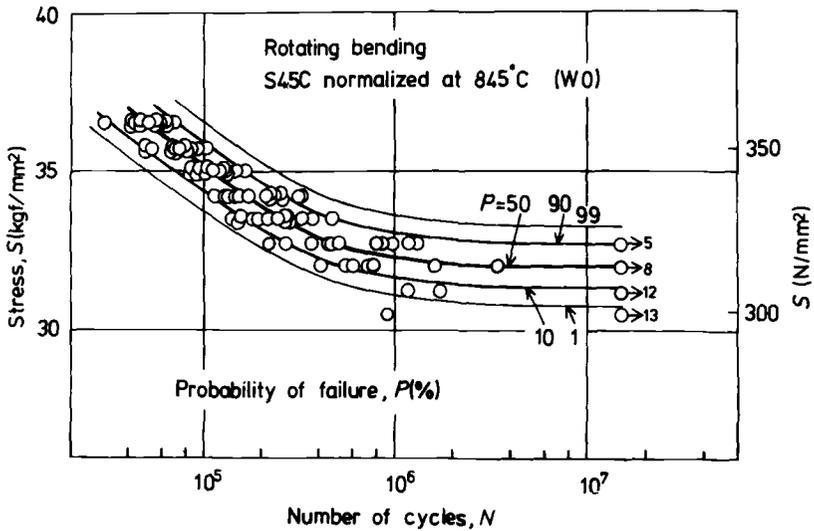


FIG. 4b—P-S-N diagram for S45C steel normalized at 845°C.

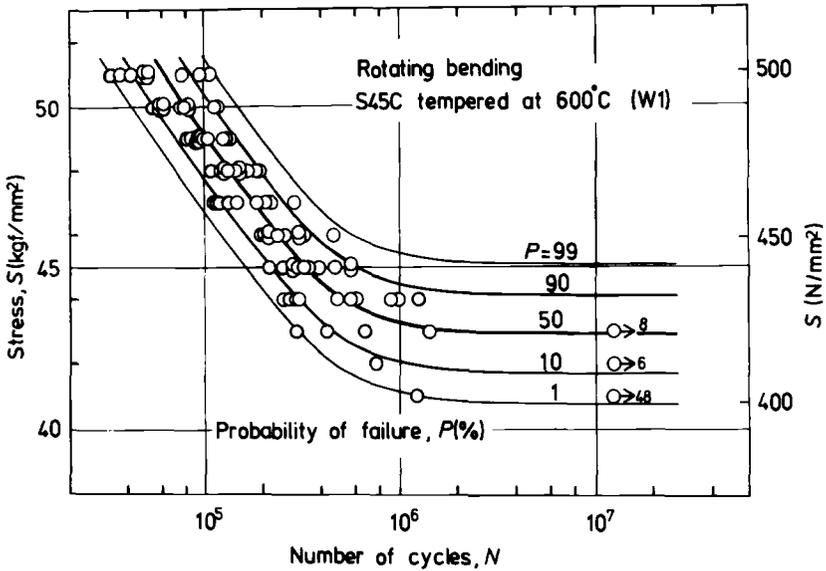


FIG. 4c—P-S-N diagram for S45C steel tempered at 600°C.

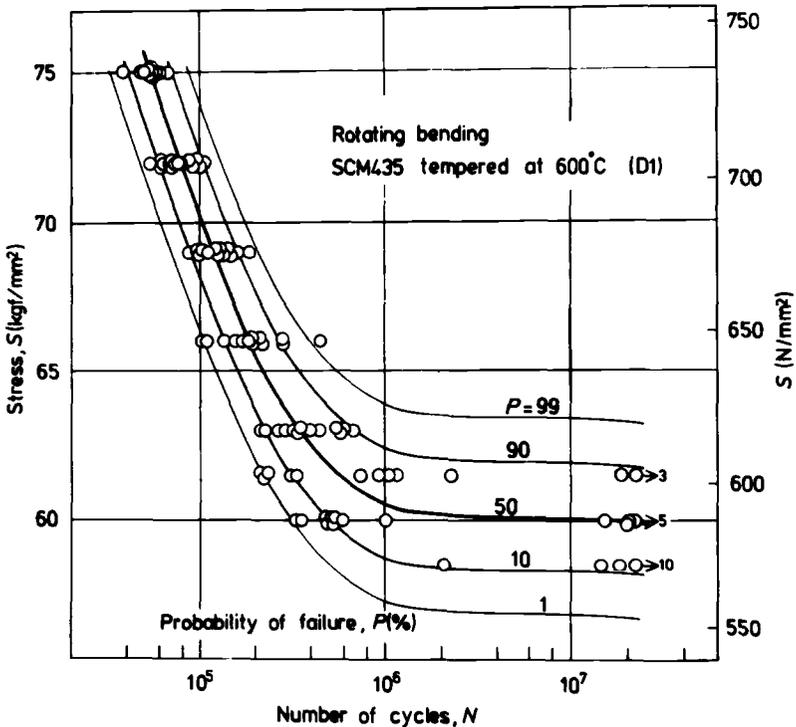


FIG. 4d—P-S-N diagram for SCM435 steel tempered at 600°C.

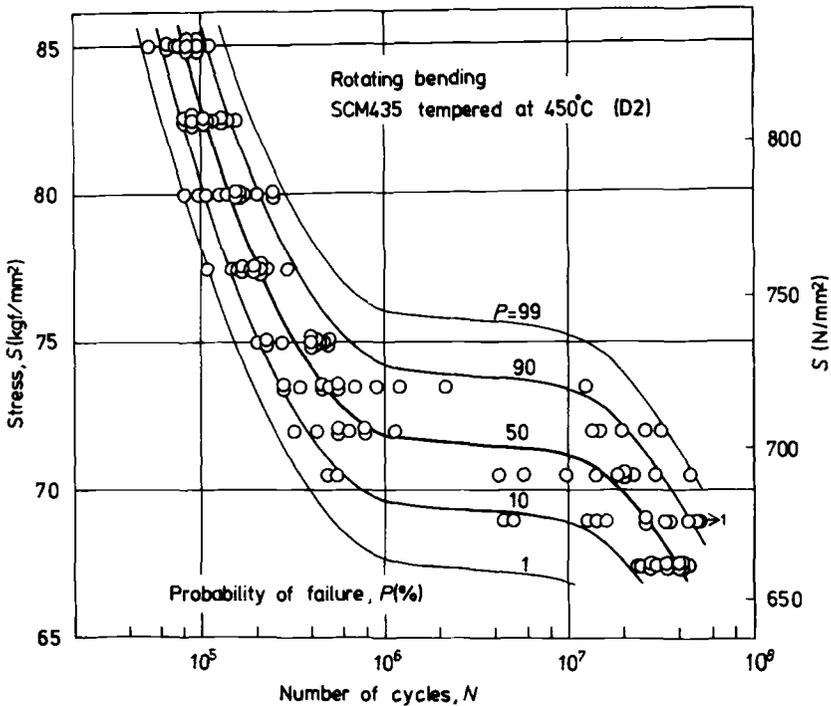


FIG. 4e—P-S-N diagram for SCM435 steel tempered at 450°C.

ranges higher than 10^7 cycles. A similar anomaly can be observed for SNCM439 steel tempered at 630°C, but the degree of change is less pronounced.

Discussion

In the present work, the normality of the fatigue strength distribution and the constancy of the coefficient of variation have been assumed as the basis of the statistical analysis. In reality, both of these are not mandatory. As seen in Fig. 3, the relative strengths in fatigue are distributed normally in this work, as is the Vickers hardness of the materials, but it could be different for other experimental conditions [6, 7]. If other distribution functions were chosen for analysis, it would be enough to modify Eqs 2 to 6 as required, according to the distribution. The assumption of a constant coefficient of variation is also an empirical hypothesis observed for a wide range of materials tested, with both smooth and notched specimens, but it is not general [7]. A more generalized transformation than that in Eq 1 could be used to homogenize the variation, as required. The use of the coefficient of variation has, however,

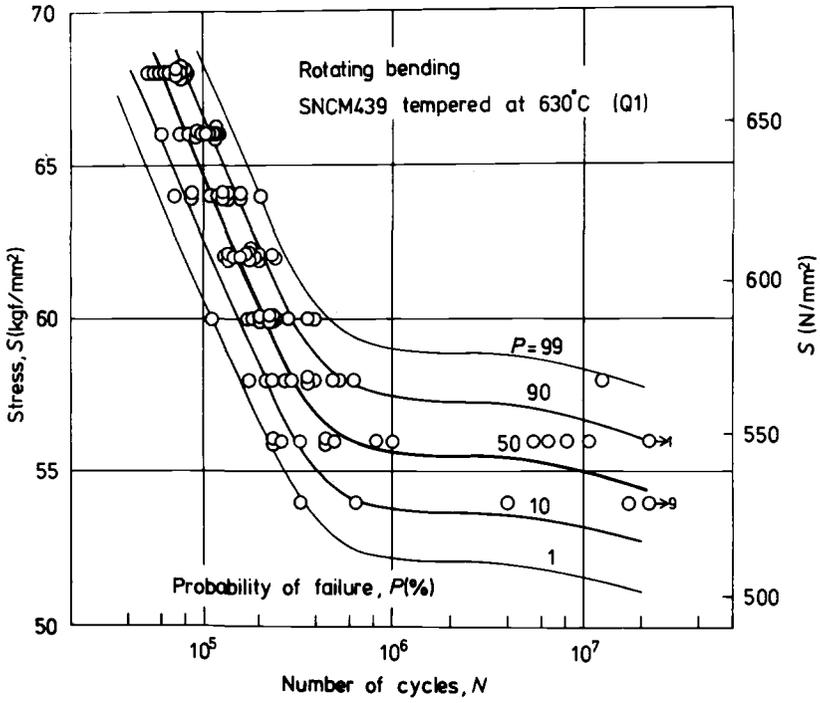


FIG. 4f—P-S-N diagram for SNCM439 steel tempered at 630°C.

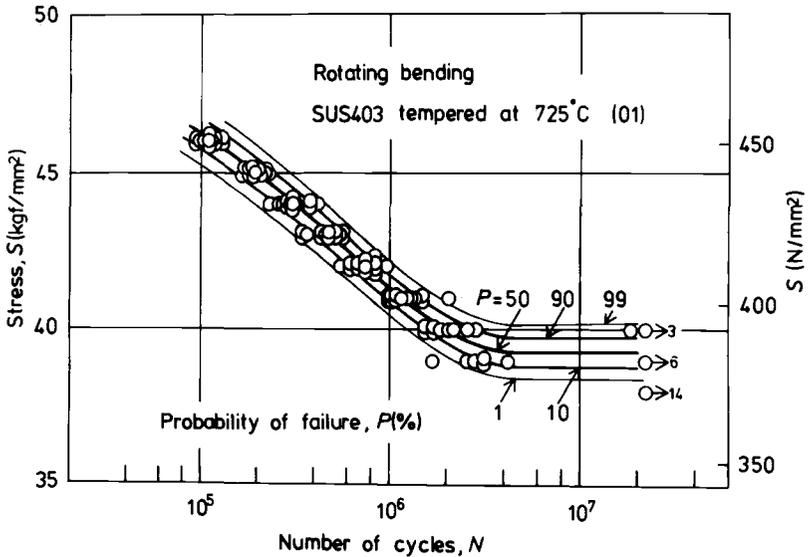


FIG. 4g—P-S-N diagram for SUS403 steel tempered at 725°C.

the great advantage of allowing the possibility of direct comparison between materials and between different mechanical properties.

The second fall in the $S-N$ curve in the very high cycle region raises grave problems in practice. In general, this phenomenon becomes more evident when the steel is less tempered and consequently harder, whereas it is scarcely observed in sharply notched specimens, even for very hard materials [6]. These observations are for fatigue in laboratory air, and similar effects are familiar in corrosion fatigue [13]. This seems to suggest a harmful effect of atmosphere, which should not be neglected in long-term fatigue life evaluation. For simplicity, it is assumed in Fig. 4 that the coefficient of variation in fatigue strength is constant even for very high cycle ranges. It would seem, however, that further extensive studies are needed to know whether this assumption is pertinent.

As stated earlier, some variation in fatigue test results is believed to be attributable to the nonuniformity of specimens in strength, which can be evaluated from the variation in the Vickers hardness of specimens. The coefficient of variation of the hardness, ζ_H , is in general smaller than that of the resultant fatigue strength variation coefficient, ζ_w , as shown in Table 2. This difference may be considered as being due to a variation introduced during the fatigue process itself, or what may be called net fatigue variation. This is evaluated from

$$\zeta_0 = \sqrt{\zeta_w^2 - \zeta_H^2} \quad (13)$$

and appears to increase with increasing material hardness, as seen in Fig. 5. This tendency is consistent with our knowledge of materials properties in that hard materials are generally sensitive to the presence of defects, which affect

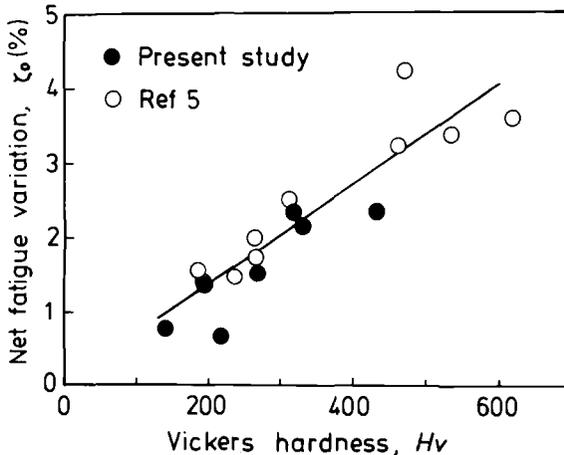


FIG. 5—Change of net fatigue variation with Vickers hardness of materials.

directly the scatter in the initiation of fatigue cracks, on one hand, and to the suggested harmful effect of the environment, on the other.

Conclusion

A new method of analysis has been proposed to investigate the statistical properties of materials in fatigue. More than 100 specimens were used in the present series of fatigue tests to obtain an $S-N$ curve for 1 percent failure probability, although the number of specimens could be reduced according to the required level of failure probability. The success of the method was demonstrated by the analysis of test data for smooth specimens of several heat-treated carbon and low-alloy steels. It was shown that the fatigue strength of these materials follows a normal distribution, as did their Vickers hardness. The net fatigue variation, or the variation introduced during the fatigue process itself, appeared to increase with increasing materials hardness. It is believed that the proposed method of analysis offers an accessible approach to the evaluation of the statistical nature of fatigue for a variety of materials, and thus contributes to the enrichment of our knowledge in materials science.

References

- [1] Weibull, W., *Fatigue Testing and Analysis of Results*, Pergamon Press, New York, 1961.
- [2] *Probabilistic Aspects of Fatigue, ASTM STP 511*, American Society for Testing and Materials, Philadelphia, 1972.
- [3] Little, R. E., *Manual on Statistical Planning and Analysis of Fatigue Experiments, ASTM STP 588*, American Society for Testing and Materials, Philadelphia, 1975.
- [4] Little, R. E. and Jebe, E. H., *Statistical Design of Fatigue Experiments*, Applied Science Publishers Ltd., London, 1975.
- [5] Nishijima, S., Masuda, C., Abe, T., Ohta, Y., Takeuchi, E., Komatsu, A., Ishii, A., Matsuyama, T., and Sumiyoshi, H., *Transactions of the National Research Institute for Metals, Tokyo*, Vol. 19, 1977, pp. 119-132.
- [6] Nishijima, S., Masuda, C., Abe, T., Takeuchi, E., Komatsu, A., Ishii, A., Matsuyama, T., Sumiyoshi, H., Tanaka, Y., and Ohtsubo, S., *Transactions of the National Research Institute for Metals, Tokyo*, Vol. 19, 1977, pp. 327-343.
- [7] Nishijima, S., Masuda, C., Abe, T., Takeuchi, E., Ishii, A., Sumiyoshi, H., and Tanaka, Y., *Transactions of the National Research Institute for Metals, Tokyo*, Vol. 20, 1978, pp. 314-320.
- [8] Nishijima, S. and Takeuchi, E., *Transactions of the National Research Institute for Metals, Tokyo*, Vol. 21, 1979, pp. 74-84.
- [9] Bastenaire, F., Bastien, M., and Pomey, G., *Acta Technica Academiae Scientiarum Hungaricae*, Tom. 35-36, 1961, pp. 7-26.
- [10] Nishijima, S., *Proceedings, Fifteenth Japan Congress on Materials Research, Society of Materials Science, Kyoto, Japan, 1972*, pp. 7-12.
- [11] Pearson, E. S. and Hartley, H. O., *Biometrika Tables for Statisticians*, Vol. 1, Cambridge University Press, New York, 1962, p. 4.
- [12] Johnson, L. G., *The Statistical Treatment of Fatigue Experiments*, Elsevier, New York, 1964, p. 37.
- [13] Iwamoto, K., *Transactions of the Japan Society of Mechanical Engineers*, Vol. 30, No. 212, 1964, pp. 500-503.

Some Considerations in the Statistical Determination of the Shape of S-N Curves

REFERENCE: Spindel, J. E. and Haibach, E., "Some Considerations in the Statistical Determination of the Shape of S-N Curves," *Statistical Analysis of Fatigue Data, ASTM STP 744*, R. E. Little and J. C. Ekvall, Eds., American Society for Testing and Materials, 1981, pp. 89-113.

ABSTRACT: This paper discusses the need for defining the shape of S-N curves, the various kinds of test data available for this purpose, and the problems in their statistical evaluation, including the assumptions that must be made.

A method is proposed for finding the "cutoff" point, that is the endurance at which conventionally shaped S-N curves change slope to the horizontal. It is based on maximum likelihood principles and deals with runouts in a statistically acceptable way. A sharp or a continuous transition from the horizontal to the sloped straight line $\log S/\log N$ curve may be considered.

The method can be used for analysis and comparison of fatigue test results with a computer program, described and listed elsewhere by the authors, but it is subject to certain amendments, which are described in the paper.

KEY WORDS: S-N curve, $\log S/\log N$ curve slope, endurance limit, statistical analysis, maximum likelihood method, computer program, fatigue

The determination of the shape of S-N curves and, in particular, of the endurance limit or the "cutoff point" at which they change to a nearly horizontal line, is not a purely academic exercise. For quite a number of components the endurance limit is a design criterion. Moreover, it is found in cumulative damage calculations that the position of the cutoff point has a considerable effect on the stress that is calculated as tolerable for a given load spectrum when such calculations are based on Miner's rule as modified by fracture mechanics considerations [1].³ The main interest here lies in the design of relatively large welded components for which test data are scarce at high values of endurance.

¹ Bridge development engineer, British Railways Board, London NW1 6JU, Great Britain.

² Director, Fraunhofer-Institut für Betriebsfestigkeit (LBF), Darmstadt, Federal Republic of Germany.

³ The italic numbers in brackets refer to the list of references appended to this paper.

Figure 1 shows S - N curves with differing slopes ($k = -3.00$ to -3.75) and different cutoff points ($N_E = 2 \cdot 10^6$, $10 \cdot 10^6$ or ∞); these differences appear between S - N curves for welded joints given in existing standards, although these standards are based on more or less the same series of published test results. From the differing S - N curves, the endurance under spectrum loading (Gaussian random process) have been calculated by the modified version of Miner's rule and plotted in terms of the maximum stress amplitude, \bar{S}_a , in the spectrum ($\bar{S}_a = 5.25S_{rms}$). Considerable differences are found not only in the stress at the endurance limit but also in the fatigue strength for spectrum loading calculated for high-endurance values, that is, values above 10^7 cycles where experimental verification is not practicable because of the testing time and cost. Obviously the effect of the different cutoff points is more pronounced than that of the differing slopes.

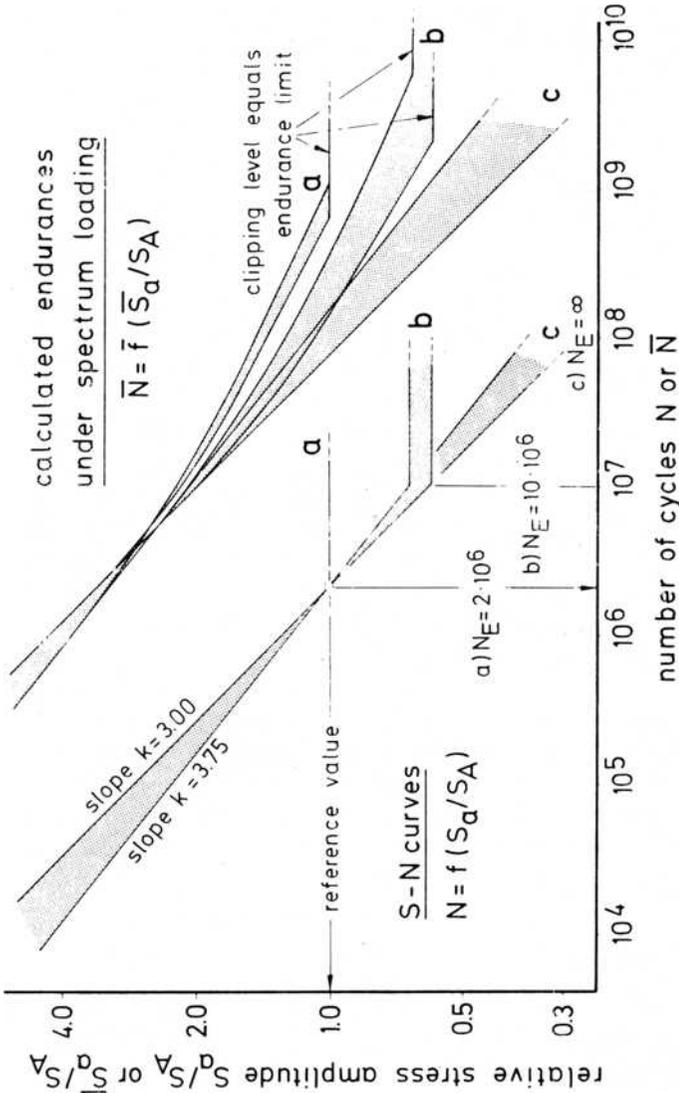
The authors had to deal with this problem, concerning the slopes of straight-line approximations to the S - N curve, and developed a method that would not only determine the "best" common slope to fit a number of sets of individual test data (meaning data obtained on similar types of specimens in different laboratories) but would also deal with runouts in a statistically acceptable way [2]. This technique was based on maximum likelihood and used the support, as Edwards terms the logarithm of the likelihood [3], to select best common lines. The aim of this analysis was not so much to determine a good summary description of an individual set of data as to determine the parameters of the most likely parent population common to a number of sets of data for apparently similar types of specimens and, further, by way of finding confidence limits, to determine the range of such populations which must reasonably be considered to have produced the data.

It had been hoped that the same technique would allow the position of the cutoff point to be determined when the shape of the S - N curve is approximated by a pair of straight lines, one of which is horizontal; but certain difficulties, discussed in this paper, were encountered in this procedure, which led to a need for modifying the procedure and introducing an additional criterion for defining the location of this point.

Possible Definitions of the S - N Curve

Any statistical analysis must be based on some assumptions about the shape of the S - N curve, however large or small the number of parameters used to define it. Various formulas for continuous curves to represent the stress-endurance relationship have been suggested since Wöhler (Table 1). These functions represent simple straight-line approximations in either log-linear or log-log coordinates that change to the horizontal at the endurance limit, or hyperbolic or S-shaped functions that approach the endurance limit, E , and the ultimate strength, R , as well.

The straight-line approximation may be found to give a poor fit in the case



- N = Number of cycles under constant amplitude loading.
- S_d = Stress amplitude under constant amplitude loading.
- \bar{N} = Number of cycles under spectrum loading.
- \bar{S}_d = Maximum stress amplitude under spectrum loading = $5.25 \times S_{rms}$.
- S_A = Reference stress amplitude at $2 \cdot 10^6$ cycles, equal for Curves a, b, and c.

FIG. 1—S-N Curves, a, b, and c differing in Slope k and in the cutoff point, N_E , and the resulting differences of the calculated endurance curves under spectrum loading (Gaussian process, clipped at $5.25 \times rms$).

TABLE 1—Proposed equations for S-N curves.^a

1870	Wöhler	$\log N$	$= a - b \cdot S$	for $S \geq E$
1910	Basquin	$\log N$	$= a - b \cdot \log S$	for $S \geq E$
1914	Strohmeyer	$\log N$	$= a - b \cdot \log [S - E]$	
1924	Palmgren	$\log (N + d)$	$= a - b \cdot \log [S - E]$	
1949	Weibull	$\log (N + d)$	$= a - b \cdot \log [(S - E)/(R - E)]$	
1955	Stüssi	$\log N$	$= a - b \cdot \log [(S - E)/(R - S)]$	
1963	Bastenaire	$\log N$	$= a - 1 \cdot \log [S - E] + [(S - E)/b]^c$	

^a Key:

E = endurance limit

R = ultimate strength

$a, b, c,$ and d = parameters

of well-documented data, as shown in Fig. 2. Equally, the assumption that the logarithms of the endurance are normally distributed with the same standard deviation at all stress levels is evidently untrue. The S-shaped types of S - N curves, like those suggested by Weibull, Stüssi, or Bastenaire, may be considered to provide a more appropriate representation of the data if the analysis is aimed at the best possible description of an individual set of data (Fig. 3). It is evident that the more usual sets of fatigue test results (Fig. 4, for example), which contain relatively few points and show considerable scatter, cannot be used to determine the shape of an S - N curve on their own. In those cases it is hardly possible to find values for even a single parameter, such as the slope, with reasonable narrow confidence limits.

Regarding the various suggestions for the statistical distribution of test results at individual stress levels, it has been shown, for example, that there is little to choose between a log-normal and an extreme value (Weibull) distribution in the range in which data exist [4]. That one of these distributions leads to ridiculous results when extrapolated beyond that range is, in the authors' view, of little interest in the analysis of data, though it may well be important in attempts to calculate the reliability of large numbers of structural components by extrapolating from endurances reached in tests to those required in service.

Possible Statistical Methods

The statistical methods that can be used to analyze fatigue test results range from the crude blunderbuss approach, in which all results from similar types of specimens are treated as one sample and a straight line is put through the lot by regression analysis, to the ultrarefined approach, where multiple regressions are calculated with transformed data to produce various shapes of curves. These methods can be further extended to deal with runouts and various assumptions about the statistical distribution of the individual test results.

The method of analysis appropriate in any given case should be selected on

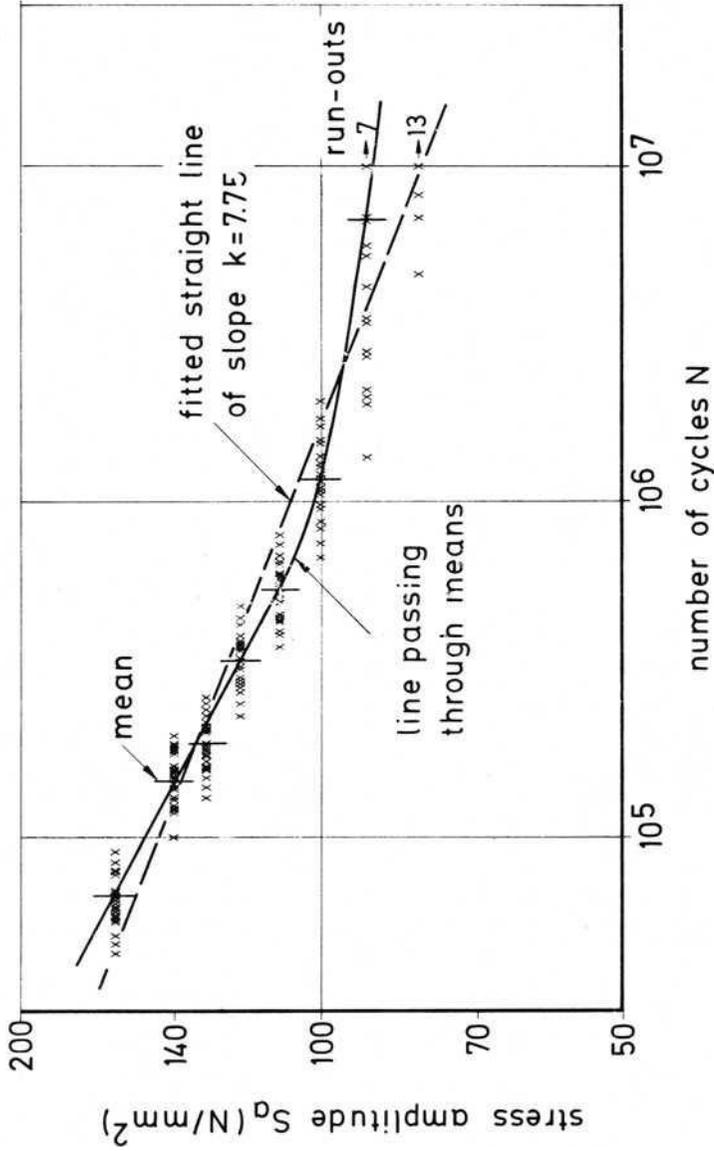


FIG. 2.—S-N curve established by 20 tests per stress level (data from Serensen, Troufiakov, Dvoretzky, and Kramarenko, 1972).

the principle that the simplest statistical model that will fit the available data with acceptable precision is the right one to use. In this context, precision refers to the inferences drawn about the parent population that might have produced the data rather than to a description of the individual set of data, which is, after all, given by a table of results without any trouble whatsoever.

Multiple regression may be appropriate for well-documented data, such as those shown in Fig. 2, but it must be pointed out that the data will have to be transformed to give dependent variables with a constant standard deviation if available statistical packages are used in the analysis. Otherwise the regression and the confidence limits calculated will be distorted. In any case, if the data contain runouts, a method based on maximum likelihood is needed to take account of these (Fig. 2).

The same methods, however, cannot be applied to data such as those in Fig. 4 because it is not possible to determine whether the results show constant or variable standard deviation at different stress levels. In such cases it is the authors' contention that any information concerning the shape of the $S-N$ curve must be derived from a number of comparable sets of such data. A graphical analysis developed on the assumption of a uniform shape of the $S-N$ curves for comparable test series but allowing for variation in the parameter S_A that defines the fatigue strength at $2 \cdot 10^6$ cycles has been described elsewhere in detail [5] (Fig. 4). A similar analysis in terms of conventional regression analysis would require the determination of the individual lines, a test to check whether they could all have come from a common population of such lines, and then the determination of the best fit line and its confidence limits [6].

This whole process is far simpler if the method of maximum likelihood is accepted. The support, defined as the logarithm of the likelihood, can be calculated for any number of parameters separately for all the sets of data, and the total support for each parameter can be obtained by summing the appropriate support values. In this way, the best supported values for various parameters of the $S-N$ curve can be found on the assumption that all the sets of data have these parameters in common. It is prudent, of course, to check that none of the individual sets is so far removed from the "best fit" parent population as to make the hypothesis untenable.

Combining Sets of Data

The technique of combining individual sets of data can be illustrated by applying it to samples drawn from a normally distributed population: If the population mean is μ and its variance σ^2 , the probability of drawing a sample value, y_i , is

$$\frac{1}{\sigma \sqrt{2\pi}} \exp \left[-\frac{1}{2} \cdot \left(\frac{y_i - \mu}{\sigma} \right)^2 \right] dy \quad (1)$$

the logarithm of which is

$$-\frac{1}{2} \cdot \left(\frac{y_i - \mu}{\sigma} \right)^2 - \ln \sigma + \ln \frac{dy}{\sqrt{2\pi}} \tag{2}$$

The last term in this expression is constant for all values of μ and σ and, therefore, of no interest in likelihood ratio or support calculations. The support from a sample of n values of y_i is given by

$$\text{SUP} = -\frac{1}{2} \left\{ \frac{1}{\sigma^2} \sum_1^n (y_i - \mu)^2 + n \cdot \ln \sigma^2 \right\} \tag{3}$$

$$= -\frac{n}{2} \left\{ \frac{1}{\sigma^2} \cdot \left[(\bar{y} - \mu)^2 + s^2 \right] + \ln \sigma^2 \right\} \tag{4}$$

if \bar{y} is the sample mean and s^2 is the sample variance. [Here s^2 is the square of the maximum likelihood estimator, $\Sigma(y_i - \bar{y})^2/n$, and not that of the unbiased estimator having $n - 1$ as its divisor].

For runouts in fatigue tests, the probability is given by the condition that no failure has occurred up to a certain value of y_i , where y_i would be the logarithm of the endurance. The probability is

$$\frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{(y_i - \mu)/\sigma} \exp \left[-\frac{1}{2} \xi^2 \right] d\xi \tag{5}$$

and the support is

$$\text{SUP} = \ln \left\{ \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{(y_i - \mu)/\sigma} \exp \left[-\frac{1}{2} \xi^2 \right] d\xi \right\} \tag{6}$$

Inferences about the parent population in terms of μ and σ are obtained by maximizing support in terms of these variables, algebraically where possible and numerically otherwise.

The numerical method is illustrated in Table 2, which shows matrices of support values for various combinations of μ and σ for two such samples. The first two matrices allow the most likely joint values of μ and σ to be found and produce the same results as conventional statistical methods. If it is assumed, or known, that the two samples have been drawn from the same population, then Matrix 3, whose values of support are the sums of the values in Matrices 1 and 2, must be used. A mean of $\mu = 9$ and a standard deviation of $\sigma = 4$ are found as the best estimates.

If it is believed that the means are independent of the standard deviations,

TABLE 2—*Matrices of support values for various combinations of the mean, μ , and standard deviation, σ , for two samples drawn from a normally distributed population.*

Standard Deviation σ of Population	Mean, μ , of Population							
	4.5	6.0	7.5	9.0	10.5	12.0	13.5	
Matrix 1: Sample 1 ($\bar{x} = 6; s = 2; n = 3$)								Matrix 1b
1	-9.4	-6.0	-9.4	-19.5	-36.4	-60.0	-90.4	-6.0
2	-4.4	-3.6	-4.4	-7.0	-11.2	-17.1	-24.7	-3.6
3	-4.3	-4.0	-4.3	-5.5	-7.3	-10.0	-13.3	-4.0
4	-4.7	-4.5	-4.7	-5.4	-6.4	-7.9	-9.8	-4.5
5	-5.2	-5.1	-5.2	-5.6	-6.3	-7.2	-8.4	-5.1
6	-5.6	-5.5	-5.6	-5.9	-6.4	-7.0	-7.9	-5.5
7	-6.0	-6.0	-6.0	-6.2	-6.6	-7.1	-7.7	-6.0
8	-6.0	-6.3	-6.4	-6.5	-6.8	-7.2	-7.7	-6.3
Matrix 1a								
	-4.3	-3.6	-4.3	-5.4	-6.3	-7.0	-7.7	
Matrix 2: Sample 2 ($\bar{x} = 12; s = 3; n = 3$)								Matrix 2b
1	-97.9	-67.5	-43.9	-27.0	-16.9	-13.5	-16.9	-13.5
2	-26.6	-19.0	-13.1	-8.8	-6.3	-5.5	-6.3	-5.5
3	-14.2	-10.8	-8.2	-6.3	-5.2	-4.8	-5.2	-4.8
4	-10.3	-8.4	-6.9	-5.9	-5.2	-5.0	-5.2	-5.0
5	-8.7	-7.5	-6.6	-5.9	-5.5	-5.4	-5.5	-5.4
6	-8.1	-7.3	-6.6	-6.1	-5.8	-5.8	-5.8	-5.8
7	-7.8	-7.2	-6.7	-6.4	-6.2	-6.1	-6.2	-6.1
8	-7.8	-7.3	-6.9	-6.7	-6.5	-6.5	-6.5	-6.5
Matrix 2a								
	-7.8	-7.2	-6.6	-5.9	-5.2	-4.8	-5.2	

TABLE 2—Continued

Standard Deviation σ of Population	Mean, μ , of Population							
	4.5	6.0	7.5	9.0	10.5	12.0	13.5	
	Matrix 3: Samples 1 and 2 Combined							Matrix 3b
1		-73.5	-53.3	-46.5	-53.3	-73.5		-19.5
2	-31.0	-22.5	-17.5	-15.8	-17.5	-22.5	-31.0	-9.1
3	-18.5	-14.8	-12.5	-11.8	-12.5	-14.8	-18.5	-8.8
4	-15.0	-12.9	-11.7	-11.2	-11.7	-12.9	-15.0	-9.5
5	-14.0	-12.6	-11.8	-11.5	-11.8	-12.6	-14.0	-10.5
6	-13.7	-12.8	-12.2	-12.0	-12.2	-12.8	-13.7	-11.3
7	-13.9	-13.2	-12.8	-12.6	-12.8	-13.2	-13.9	-12.1
8	-14.2	-13.6	-13.3	-13.2	-13.3	-13.6	-14.2	-12.8
9	-14.5	-14.1	-13.8	-13.8	-13.8	-14.1	-15.5	
	Matrix 3a							
	-12.1	-10.8	-10.9	-11.3	-11.5	-11.8	12.9	

the best supported value of the mean to fit both samples is obtained by summing the best supported values for mean from each column independently of the corresponding values of standard deviation (Matrices 1a, 2a, and 3a). According to this hypothesis, the best supported mean value would be 6.9 (found by interpolation of Matrix 3a) instead of 9. If, however, it is believed that the standard deviations are independent of the means (Matrices 1b, 2b, and 3b), the best supported value of the standard deviation common to the two samples would be a value of about 2.5 (found by interpolation of Matrix 3b), which is quite different from that which would be obtained if both samples were taken to be from one population (3.93, or approximately 4 from Matrix 3).

Confidence Limits

It is not enough to determine the most likely value of a parameter. Since it is known that samples drawn from a given population will show variations in their properties, it must also be accepted that a given sample could have

come from a range of populations which do not differ "overmuch" from the best supported population. What is accepted as "overmuch" will depend on the particular circumstances and the known scatter in experimental results to be expected from a particular type of test. In conventional statistical analyses this range of a parameter is given by confidence limits, which are defined by stating the probability of being right, in the long run, in asserting that the true value of the parameter lies within these limits [7].

In the method of support, Edwards replaces this criterion with one of not exceeding a certain likelihood ratio, which, in terms of the logarithm of that ratio, amounts to a loss of 2 units of support to correspond roughly to the conventional 95 percent confidence limit and a loss of 4.5 units of support to correspond to the conventional 99.8 percent confidence limit [3].

Which of these approaches one uses may be a matter for argument, though Edwards gives strong reasons for preferring his method. In terms of practical application to the type of analysis described in the following sections, Edwards's method is very much simpler.

Applying this idea to the figures shown in Matrix 3 of Table 2 would, for example, suggest that the value of μ could, with a loss of 2 units of support, lie anywhere between 6 and 12, a result which is hardly surprising in view of the large sample variance. It should be noted that this range is obtained by choosing the populations with the widest range of μ , irrespective of the value of σ . It corresponds to that which would be found from Student's t distribution and thus takes full account of the sample size.

Support calculations can also be used to test the hypothesis that the variances in Samples 1 and 2 might be different values for samples taken from populations with a common variance. If the best support values for standard deviations in Matrices 1*b* and 2*b* are added, it will be found by interpolation that the best supported value is about 2.5, and its support is -8.90 (Matrix 3*b*). This is the support for the hypothesis that the two standard deviations are different values obtained from populations with the same variance. It may be compared with the support for the hypothesis that the two standard deviations come from separate populations, for which the total support is given from the Matrices 1*b* and 2*b* by $(-3.6) + (-4.8) = -8.4$. This value is only a little greater than that for the previous hypothesis, and therefore the hypothesis that the two variances come from sources with the same variance could be accepted. A similar computation for the values of the mean, μ , (Matrices 1*a*, 2*a*, and 3*a*) would show a loss of support of 2.4 for the hypothesis that the two means are values drawn from populations with the same mean.

The Support Method Applied to S-N Curves

The only difference between this calculation and a form of regression analysis is that both μ and σ become functions of the independent variable, x . These functions are defined by a number of parameters, for example, two

parameters, m and c , to define a straight line of the mean line, $\mu(x)$, as $mx + c$ and either a constant or similar simple function to define $\sigma(x)$. The best supported (most likely) values of these parameters are then found by evaluating the support for all possible combinations and picking the best supported joint combination. It is evident from Eq 3 that this process will produce a least squares solution if a normally distributed population is assumed and is thus equivalent to normal regression analysis. Confidence limits for regressions are more complex, but those for the mean values of the slope and other parameters of the mean line can be taken from matrices, as illustrated for sample values in Table 3.

When the support method is applied to the analysis of $S-N$ data, the transformed variables $\log N$ and $\log S_a$ are most commonly treated. To specify the shape of the $S-N$ curve, (a) the fatigue strength S_A at $2 \cdot 10^6$ cycles, (b) the slope, k , (c) the standard deviation, s , of $\log N$ or $\log S_a$ (see the following section), and (d) the position of change of slope N_E at stress S_E are to be considered as parameters of a simple model (Fig. 5 Curve a). The support matrices may be established by a computer program [8] as a function of two of these parameters, such as joint values of S_A and k , with the standard deviation and the value of N_E kept constant. (For a description and listings of the developed computer programs, see Refs 8 and 9).

Matrices 1 and 2 in Table 3 apply to comparable sets of results. When the two sets are combined by summing the matrices in Matrix 3, the best supported joint values of slope and fatigue strength are $k = -3.37$ and $S_A = 68$ N/mm². If we were to allow for differing values of S_A , however, the best supported value of slope independent from S_A would be found by individually summing the highest support values in each column; a best supported value of slope of $k = -3.75$ results (Matrix 1 + Matrix 2).

The sample for Matrix 2 has been plotted in Fig. 4. As the cutoff point was chosen outside the range of the present data to simplify the example ($N_E > 10^7$), direct comparison may be made of the best slope of the regression line and its confidence limits, as shown in Fig. 4, with the best supported value of k and the framed areas of the matrix in which the loss of support is less than 2 units.

The Support Method for Defining the Shape of $S-N$ Curves

If a more detailed analysis of the shape of the $S-N$ curve (Fig. 5) is intended by means of the maximum likelihood method, a slightly modified concept for the support computations is to be preferred [9]. The matrix of the support values is determined as a function of the Slope k and the standard deviation, s , for a fixed value of the cutoff point, N_E , whereas the endurance limit, S_E , is determined by maximizing the support for each combination of the other parameters. This computation is repeated for various values of the cutoff

TABLE 3—Matrices of support values for various combinations of Slope k and fatigue strength S_A for two comparable S-N curves.

		Slope k										Matrix		
		-2.25	-2.62	-3.00	-3.37	-3.75	-4.12	-4.50	-4.87	-5.25	Matrix 1	Matrix 2 (1 + 2)	Matrix 3	
Matrix 1	$S_A = 84 \text{ N/mm}^2$	-5	-4	-3	-3	-2	-3	-3	-3	-4	-2			
	76	1	3	4	5	5	5	5	4	3	5			
	68	4	6	6	6	5	3	1	-3	-7	6			
	61	5	4	3	0	-4	-18	-16	-24	-34	5			
	55	-2	-2	-7	-14	-23	-34	-47	-63	-88	2			
Matrix 2	$S_A = 84 \text{ N/mm}^2$	-23	-16	-11	-6	-2	1	2	3	3	3			
	76	-13	-6	-1	4	7	8	8	7	5	8			
	68	-6	0	5	8	8	7	5	-8	-7	8			
	61	-1	4	6	6	4	-2	-9	-19	-32	6			
	55	1	4	3	-1	-8	-19	-33	-58	-70	4			
Matrix 3	$S_A = 84 \text{ N/mm}^2$	-28	-20	-14	-9	-4	-2	-1	0	-1	1			
	76	-12	-3	-3	9	12	13	13	11	8	13			
	68	-2	6	11	14	13	10	6	-3	-14	14			
	61	4	8	9	6	0	-12	-25	-43	-66	11			
	55	3	2	-4	-15	-31	53	-80	-113	-150	6			
Analysis slope	Matrix 1	4.7	5.6	6.0	5.7	5.1	5.1	4.6	3.8	2.5				
	Matrix 2	1.0	3.6	6.1	7.6	8.4	8.2	8.4	7.3	4.9				
	Matrix (1 + 2)	5.7	9.2	12.1	13.3	13.5	13.3	13.0	11.1	7.4				
	Matrix 3	4	8	11	14	13	13	13	11	8				

Analysis S_A

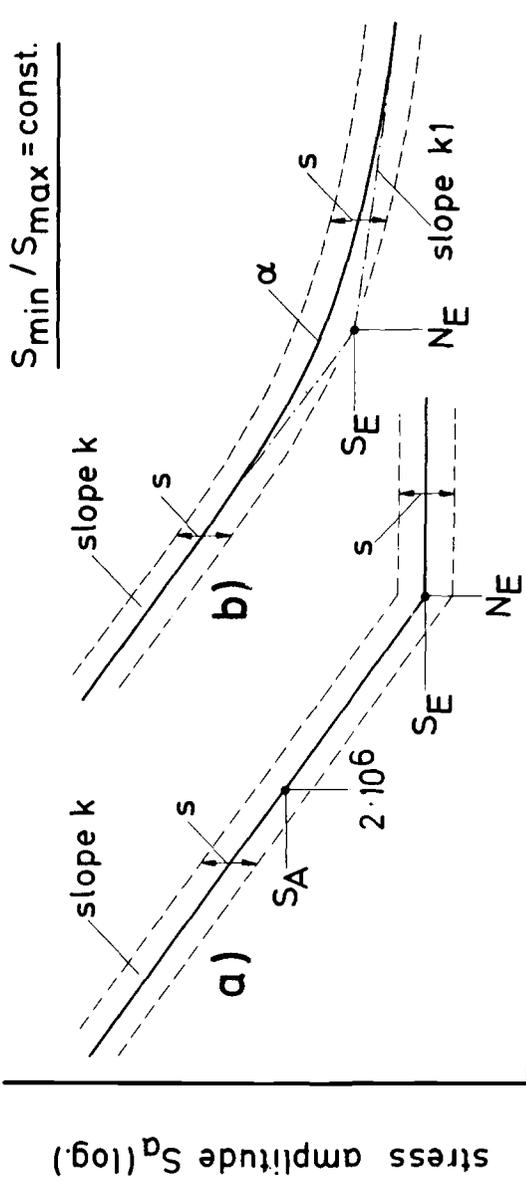


FIG. 5—Shapes of the S-N curve considered in this paper: Curve a, simple model; Curve b, extended model.

point to find the most appropriate value from the corresponding support values.

Another modification studied for comparison with the simple model of the $S-N$ curve (Fig. 5 Curve a) was an extended model showing a continuous change of slope from the steep line of Slope k to a flat (not necessarily horizontal) line of Slope k_1 , in order to allow a better fit to the realistic shape of $S-N$ curves [9] (Fig. 5 Curve b). An additional parameter, α , was introduced to specify the mode of transition. A first estimate of α among typical values $\alpha = 10$ to 100 may be found by inspection of the set of data in question; for $\alpha = \infty$, the extended model degenerates to the simple one changing slope abruptly. Analytically, the $S-N$ curve for the extended model is given by the equation

$$Y = \frac{1}{2}(k + k_1) \cdot x + \frac{1}{2}(k - k_1) \left\{ |x| + \frac{1}{\alpha} \cdot \ln \left[1 + e^{-2\alpha|x|} \right] \right\} \quad (7)$$

where $Y = \log N/N_E$, and $x = \log S/S_E$.

In examining sets of data, it appears as a marked characteristic of the point where these curves change from a relatively steep straight line to the nearly horizontal that there is a considerable increase in the standard deviation of the endurance at that level (Fig. 2). Among several possibilities to account for this phenomenon, the adopted solution was based on the assumption that each individual test result represents a similar shape $S-N$ curve but shifted vertically in relation to the mean line (Fig. 5). Physically, this assumption is based on the concept that the fatigue strength of a specimen is a function of the size of a defect at the origin of the failure, that the sizes of the defects are random and log-normally distributed, and that the random values of endurance found in tests are a measure of these defect sizes, and hence of the fatigue strength expressed as a stress at a certain endurance. In other words, for analysis of the scatter the data points have to be converted to some reference endurance value according to the $S-N$ relationship defined by the shape of the $S-N$ curve. For a linear $\log S/\log N$ relationship of Slope k , for example, one would obtain the standard deviation of $\log S$ equal to the standard deviation of $\log N$ divided by $|k|$.

This concept leads to parallel scatter bands and implies the assumption of a standard deviation of $\log S$ independent of $\log N$. Hence, if $\log S$ is considered to be normally distributed, the distribution for the logarithm of the endurance will be grossly unsymmetrical in regions of changing slope. It does, however, show the increased scatter of the endurance in that region, which can be seen in experimental data (Fig. 2).

In combination with the extended model of the $S-N$ curve (Fig. 5 Curve b), this change in standard deviation also offers an additional criterion for the fit of the curve, namely, that standard deviation of $\log S$, determined separately for the lower and upper stress levels using the same curve, should not differ significantly. In terms of support calculations, the condition is checked by finding the loss of support when the support for the line which best satisfies the other conditions calculated for the line as a whole (SUP 1) is compared with that for the same line, but calculated separately and summed for the steep and flat portions by allowing for different standard deviations while maintaining all other parameters (SUP 2) [9]. The total support for any given line with its change of slope point is then that for the best fit line which also best satisfies the additional criterion. This leads to

$$(SUP 1) + (SUP 1 - SUP 2) \quad \text{or} \quad (2 \cdot SUP 1 - SUP 2) \quad (8)$$

Illustration of the Described Method

The described method of determining the shape of $S-N$ curves will be illustrated with reference to the well-established set of data presented in Fig. 2 first (Table 4). In case of an $S-N$ curve with the slope changing abruptly to the horizontal, the best supported cutoff point is found by interpolation between computer runs 3 and 4. It yields a value of $N_E = 1.8 \cdot 10^6$ with a support of about 237.5, combined with a slope of $k = -5.8$ and an endurance limit of $S_E = 90 \text{ N/mm}^2$. An obviously better fit is obtained with an $S-N$ curve of continuously changing slope (Run 8), for which the maximum support is SUP 1 = 251.7, applying to a cutoff point $N_E = 3.0 \cdot 10^6$, a slope $k = -5.0$, and an endurance limit $S_E = 75.2 \text{ N/mm}^2$. But it has also to be

TABLE 4—Jointly best supported parameters of the $S-N$ curve as a function of the cutoff point derived for the data from Fig. 2.

Run No.	Cutoff Point, N_E	Slope k	SD of $\log S$, s	Endurance Limit, S_E , N/mm^2	Support SUP 1	Additional Criterion (2SUP1 - SUP2)
Slope Changing Abruptly to the Horizontal						
1	$0.6 \cdot 10^6$	4.00	0.0326	98.1	213.0	
2	$1.0 \cdot 10^6$	5.00	0.0228	94.8	230.9	
3	$1.5 \cdot 10^6$	5.50	0.0176	91.2	236.5	
4	$2.0 \cdot 10^6$	6.00	0.0162	89.5	236.5	
5	$3.0 \cdot 10^6$	6.50	0.0175	86.7	221.6	
6	$10.0 \cdot 10^6$	7.75	0.0246	79.5	130.7	
Slope Changing Continuously to the Horizontal ($\alpha = 20$)						
7	$2.0 \cdot 10^6$	4.50	0.0215	76.8	249.5	249.0
8	$3.0 \cdot 10^6$	5.00	0.0194	75.2	251.7	250.9
9	$5.0 \cdot 10^6$	5.50	0.0176	72.6	250.8	248.2
10 ($\alpha = 25$)	$3.0 \cdot 10^6$	5.25	0.0185	77.8	251.1	

recognized that the confidence limits for N_E and k are much wider in the latter case. From the additional criterion, a maximum support (2·SUP 1 – SUP 2) of 250.9 is obtained. By fitting a parabola to the calculated points (Runs 7 and 9), it may be deduced that by considering SUP 1, a maximum of support is to be expected for a value $N_E = 3.46 \cdot 10^6$ with confidence limits of 2.09 and $5.74 \cdot 10^6$. Using the additional criterion (2·SUP 1 – SUP 2), the result is a maximum supported value $N_E = 3.11 \cdot 10^6$ with confidence limits of 1.98 and $4.89 \cdot 10^6$. Hence, the value of N_E is found to be slightly lower, and the confidence limits become narrower when the additional criterion is used. Figure 6 allows a comparison of the derived curves with the data points.

Considering a cutoff point at $N_E = 10^7$, the abruptly changing S - N curve gives, with a rather low support of 130.7, the simple straight-line approximation of slope $k = -7.75$, as shown in Fig. 2. If a more realistic estimate of a straight line to fit the sloped part of the S - N curve is wanted, the data at the lower three or four stress levels in Fig. 2 should be neglected (Table 5). Although a decision of that kind seems reasonable for the present set of data, it should be noted that, practically, it is quite unusual to have such well-established sets of data available, so that neglecting data would result not only in a loss of information but also in a guess rather than an inference based on fact.

Therefore, in order to check the method with some more realistic sets of data, a smaller sample was selected from the data in Fig. 2, comprising the tests on the levels $S_a = 90, 110,$ and 140 N/mm^2 only. The results are given in Table 6, and they may be compared with those in Table 4. Although the best supported parameters of the abruptly changing S - N curve (Run 18) compare fairly well with those derived from the complete set, the continuously changing S - N curve gives a somewhat strange result (Fig. 7) typical of data such as those shown in Fig. 8. It is clear from this figure that for all data of this type the "best fit" line will be such that the horizontal portion passes through the row of results at the lowest stress level and, in the extreme case shown, leaves the value of N_E undefined. The previous sample illustrates that this would also be true if some form of continuous curve were used instead of two straight lines. The least squares method of fitting S - N curves with a cutoff point is, therefore, unsatisfactory for data for which the sloped part of the S - N curve is not significantly determined. The special reason that the method did work quite satisfactorily in the case of the abruptly changing S - N curve in the previous example is the exceptional situation that the lowest level is rather close to the endurance limit expected. The obvious difference in the standard deviation of the logarithm of the endurance, however, suggests that the flat portion of the curve should pass somewhere near the results at the lower stress level.

In case of this simplified example, however, most of the considered curves in the bottom of Table 6 are similar in shape near the three stress levels, and

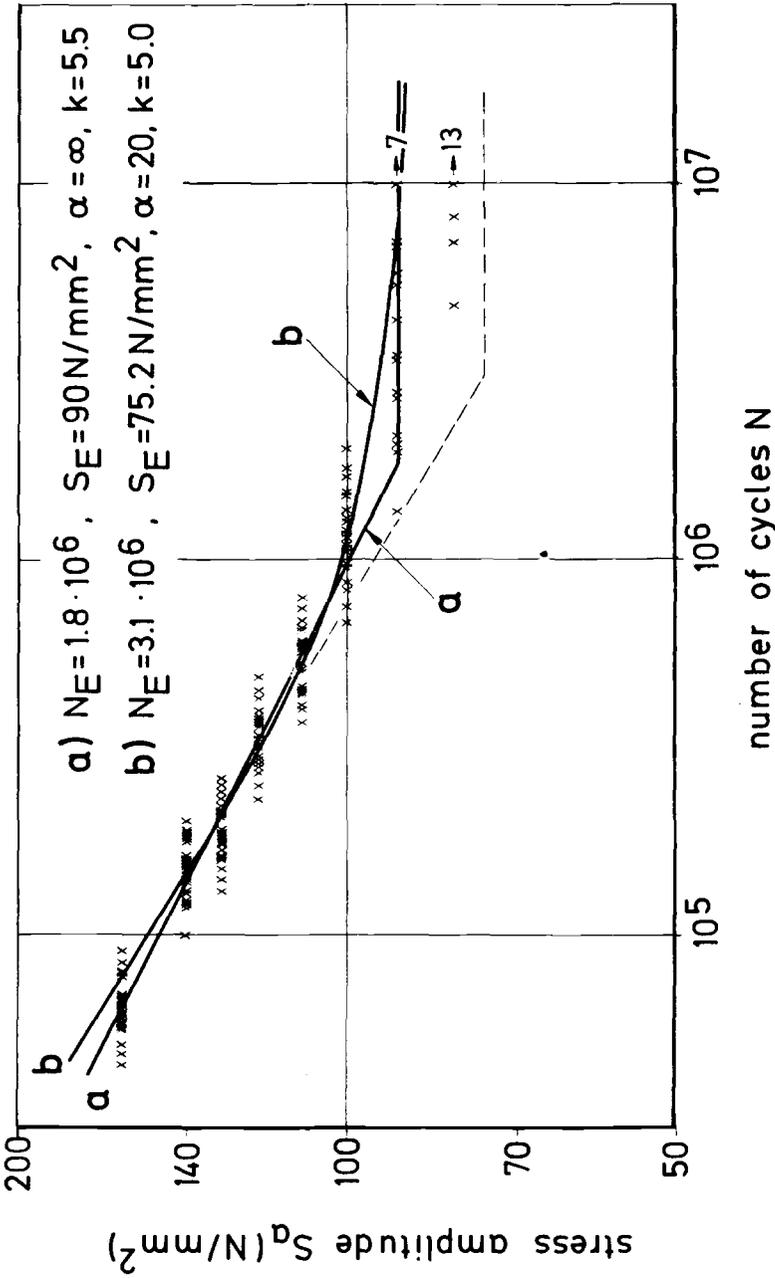


FIG. 6—Best supported S-N curves: Curve a, changing slope abruptly to the horizontal; Curve b, changing slope continuously to the horizontal as fitted to the set of data from Fig. 2.

TABLE 5—Jointly best supported parameters of a straight line S-N curve ($N_E = 10^7$) when neglecting the data at the lower stress levels of Fig. 2.

Run No.	Slope k	SD of Log S_s	Endurance Limit, S_E , N/mm ² .				Support, SUP 1			
			All Data	Without Data at the Lower			All Data	Without Data at the Lower		
				2 Levels	3 Levels	4 Levels		2 Levels	3 Levels	4 Levels
11	7.75	0.0246	<u>79.5^a</u>	78.4	79.0	79.9	<u>130.7</u>	161.2	136.7	114.1
12	6.00	0.0162		<u>68.5</u>	<u>68.2</u>	<u>68.3</u>		<u>220.3</u>	<u>192.1</u>	<u>154.9</u>
13	5.75	0.0138		<u>66.7</u>	<u>66.3</u>	<u>66.3</u>		<u>211.8</u>	<u>194.6</u>	<u>157.2</u>
14	5.50	0.0144		<u>64.8</u>	<u>64.3</u>	<u>64.2</u>		<u>203.3</u>	<u>194.3</u>	<u>157.2</u>
15	5.25	0.0185			62.1	61.9			190.9	155.6

^a Best supported values (underlined) to be compared within columns only, as support changes with the number of tests considered.

 TABLE 6—Jointly best supported parameters of the S-N curve as a function of the cutoff point derived for a selected sample of data from Fig. 2 (tests on levels $S_a = 90, 110, \text{ and } 140 \text{ N/mm}^2$ only).

Run No.	Cutoff Point, N_E	Slope k	SD of Log S_s	Endurance Limit, S_E , N/mm ²	Support, SUP 1	Additional Criterion (2 SUP 1 — SUP 2)
Slope Changing Abruptly to the Horizontal						
16	$0.6 \cdot 10^6$	3.50	0.0326	97.3	84.5	
17	$1.0 \cdot 10^6$	4.50	0.0215	93.0	98.4	
18	$1.5 \cdot 10^6$	5.50	0.0144	91.4	106.8	
19	$2.0 \cdot 10^6$	6.00	0.0132	89.8	100.3	
20	$3.0 \cdot 10^6$	6.75	0.0169	88.1	81.0	
Slope Changing Continuously to the Horizontal ($\alpha = 20$)						
21	$1.0 \cdot 10^6$	3.25	0.0244	78.4	107.0	107.1
22	$1.5 \cdot 10^6$	3.75	0.0211	76.4	105.0	105.0
23	$2.0 \cdot 10^6$	4.25	0.0186	76.0	104.4	104.5
24	$3.0 \cdot 10^6$	4.75	0.0167	74.2	103.3	103.3
25	$4.5 \cdot 10^6$	5.25	0.0151	72.7	99.6	
26	$6.0 \cdot 10^6$	5.50	0.0176	71.2	94.4	

therefore the additional criterion does not furnish any better discrimination. This also illustrates the fact that over a narrow range of stress values various curves may be equally satisfactory, though they may be quite different when extrapolated.

As pointed out in the foregoing, the situation in treating sets of data similar to those in Fig. 7 may be markedly improved if the parameters of the S-N curve are derived by combining a number of comparable sets of such data. An example is given in Table 7. In this analysis, nine comparable sets

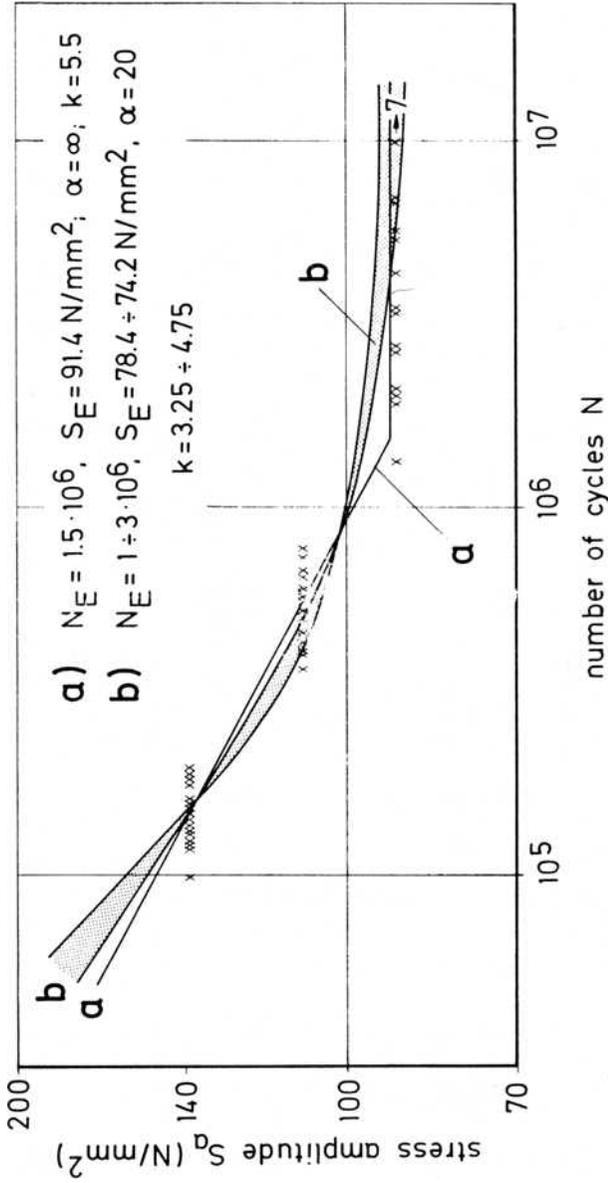


FIG. 7—Best supported S-N curves: Curve a. changing slope abruptly to the horizontal; Curve b. changing slope continuously to the horizontal as fitted to a sample of data from Fig. 2. (Stress levels, $S_u = 90, 110,$ and 140 N/mm^2 only.)

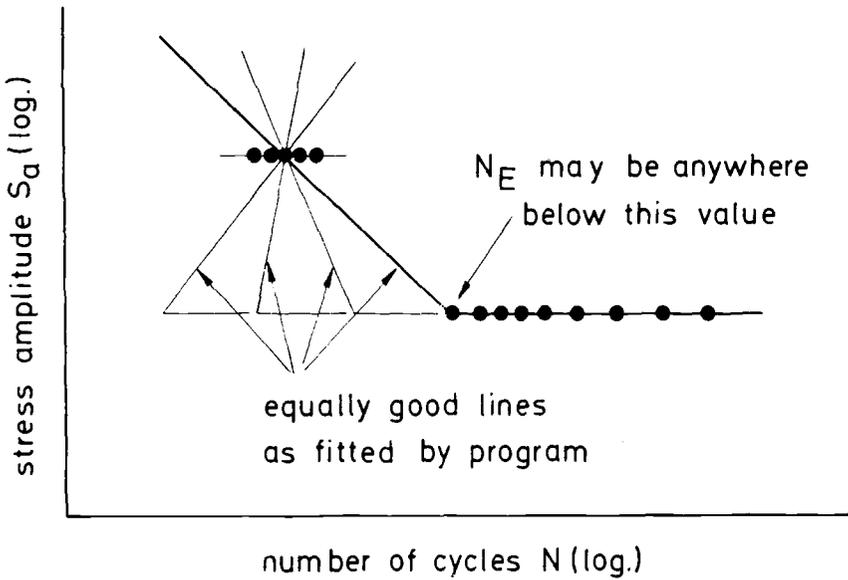


FIG. 8—The strange shape of lines fitted by program is typical of the extreme case of data restricted to two stress levels. (Neither maximum likelihood nor least square fits can distinguish between the lines shown.)

TABLE 7—Jointly best supported parameters as a function of the cutoff point to define a common shape of the S-N curve for nine comparable test series on notched specimens in quenched and tempered steel 42CrMo4 (stress concentration factor $K_t = 2.5, 3.6, \text{ or } 5.2$; stress ratio $R = -1 \text{ or } 0$).

Run No.	Cutoff Point, N_E	Slope k	SD of Log S , s	Sum of Support, SUP 1
Slope Changing Abruptly to the Horizontal				
27	$0.20 \cdot 10^6$	4.75	0.0204	339.7
28	$0.25 \cdot 10^6$	5.00	0.0194	347.1
29	$0.30 \cdot 10^6$	5.25	0.0185	339.5
30	$0.45 \cdot 10^6$	5.50	0.0207	281.2
Slope Changing Continuously to the Horizontal ($\alpha = 30$)				
31	$0.30 \cdot 10^6$	4.50	0.0215	332.9
32	$0.325 \cdot 10^6$	4.50	0.0215	334.1
33	$0.35 \cdot 10^6$	4.75	0.0204	332.4
34	$0.45 \cdot 10^6$	5.00	0.0194	330.4

of data have been combined by summing the appropriate support values to define the "shape parameters" N_E , k , and s , while the position of the thus-defined $S-N$ curves was fitted individually for each set through the value determined for the endurance limit, S_E . The jointly best supported shape parameters are $N_E = 0.25 \cdot 10^6$, $k = -5.0$, and $s = 0.0194$, in case of the abruptly changing $S-N$ curve, or $N_E = 0.325 \cdot 10^6$, $k = -4.5$, and $s = 0.0215$, in case of the continuously changing $S-N$ curve. In both cases the confidence limits are reasonably narrow, but the support of the abruptly changing $S-N$ curve is significantly higher than for the continuously changing curve. This is because the abruptly changing curve provides a better supported approximation of most of the present sets.

A similar analysis by the graphical method is presented in Fig. 9 [10]. Here the test results were plotted on a relative scale for which the individual value of the endurance limit, S_E , was the reference stress. As a consequence, all data points coincide in one scatter band. The shape parameters to be derived from the scatter band, $N_E = 0.3 \cdot 10^6$, $k = -5.0$, and $s < 0.04$, are found in good agreement with the parameters determined by the maximum likelihood analysis. In Table 8 the common shape parameters derived by the three methods mentioned and the corresponding values of the endurance limits are compared with those obtained by an individual best fit to each of the nine sets of results. While the values of the endurance limit derived individually or by the common shape $S-N$ curve show only slight differences, the individual values of Slope k differ considerably. Further, for all sets a highly significant loss of support has to be stated for the common shape $S-N$ curve when compared with the individual best fitting shape of the curve.

This example shows that inferences as to the shape parameters of the $S-N$ curve are more closely defined when based on a large number of results even if these have been taken from several independent sets of data (Table 7). This is only achieved, however, by sacrificing accuracy, since common shape parameters do not fit individual results as well as individual parameters (Table 8). Whether, and to what extent, this can be accepted is a matter of professional judgment and not mathematics.

The example also illustrates the differences in results that different methods and assumptions can give.

Conclusions

1. The most appropriate statistical method for the analysis of fatigue test results depends on the data to be analyzed. If the results contain runouts, the method used must allow their proper consideration. The inferences drawn from a statistical analysis depend on the method used. Its choice must be a matter of careful professional judgment.

2. $S-N$ curves need several parameters for their definition, at least two for a straight line and more for the complex shapes required for certain data.

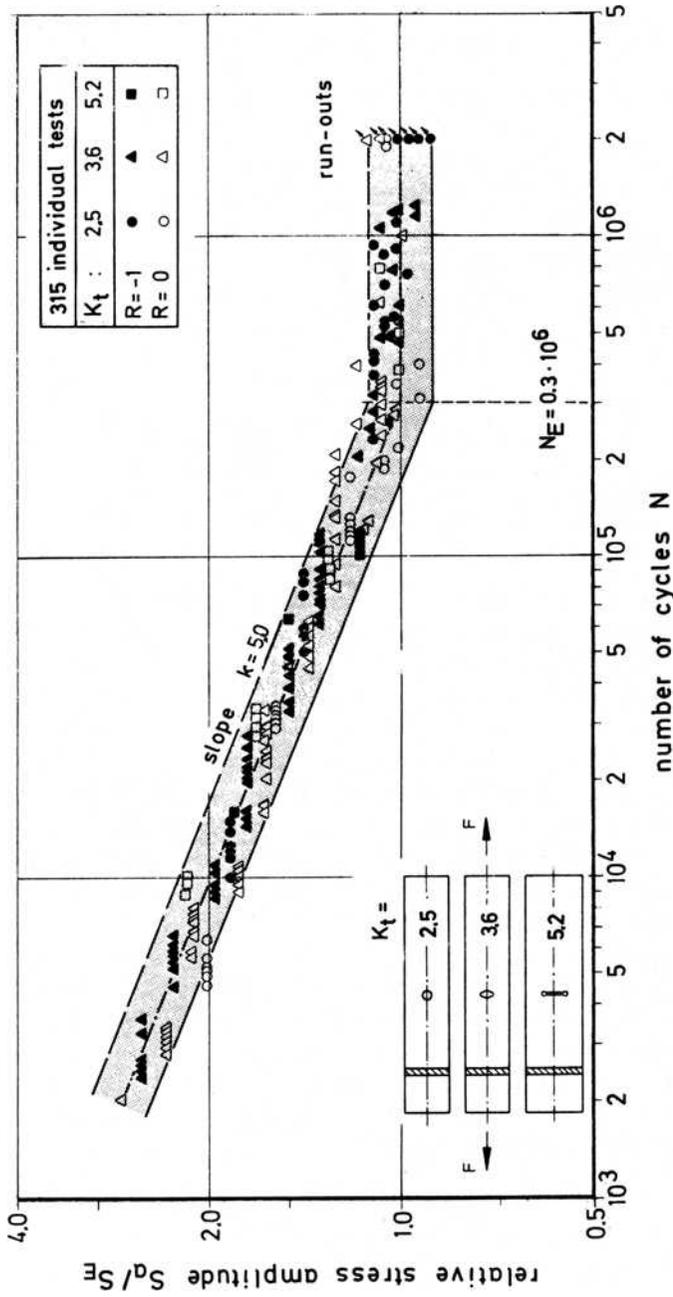


FIG. 9—Results from nine sets combined for determining common shape parameters of the S-N curve and plotted on a relative scale for graphical analysis (notched specimens in quenched and tempered steel 42CrMo4).

TABLE 8—Parameters of the graphically derived, of the individually best fitting, and of the commonly best supported $S-N$ curves for nine comparable test series on notched specimens in quenched and tempered steel 42CrMo4.^a

Set No.	Graphical Fit, $N_E = 0.30 \cdot 10^6$, $k = 5.0$, S_E	$S-N$ Curve Changing Slope Abruptly						$S-N$ Curve Changing Slope Continuously			
		$N_E = 0.25 \cdot 10^6$, Individual Best Fit			$N_E = 0.25 \cdot 10^6$, $k = 5.0$			$N_E = 0.35 \cdot 10^6$, Individual Best Fit		$N_E = 0.35 \cdot 10^6$, $k = 4.5$	
		k	S_E	SUP 1	S_E	SUP 1	k	S_E	SUP 1	S_E	SUP 1
1	215	4.50	214	41.1	216	38.5	4.25	190	38.5	192	35.3
2	180	5.00	190	43.1	191	40.7	4.25	168	37.3	170	36.4
3	165	4.75	171	58.4	173	54.7	4.25	150	54.1	153	51.6
4	135	5.25	146	52.0	144	50.8	4.50	127	49.9	129	50.4
5	165	5.25	176	35.8	172	35.0	5.25	167	37.0	153	32.0
6	135	5.50	147	28.7	141	21.5	5.50	138	35.6	125	22.8
7	190	5.25	202	65.1	199	58.0	4.75	178	60.4	175	53.8
8	160	5.50	175	19.3	173	14.6	5.25	156	25.1	154	21.6
9	105	4.50	106	45.0	109	33.7	4.25	95	36.6	97	30.8

^a Values of endurance limit S_E in N/mm^2 . The individual best fit is computed with the individual standard deviation, common shape fitted with the global value of standard deviation.

Unless the set of data to be analyzed contains a large number of results, these parameters can only be determined statistically within wide confidence limits. They can be determined within closer limits using a number of different sets of data from similar tests on the assumption that the shapes of the $S-N$ curves are similar for all these data and that the curves differ only in one parameter. (A method for doing this is outlined in the paper).

3. For other than straight-line $S-N$ curves, the method of least squares does not necessarily lead to the physically most acceptable line, particularly concerning the position where the curve changes from a relatively steep slope to a flat slope. It is shown that the position of this point is indeterminate in the case of certain sets of data and is not well determined by the least squares fit even when good data are used.

4. In view of this difficulty, an additional criterion is required for the best fit curve, namely, that the curve should also be a best fit on the assumption that the standard deviation of the logarithm of the stress remains constant along the curve. This further criterion produces a more reasonable result for the location of the point at which the change of slope occurs. In practice, it is better to use a shape of $S-N$ curve with a gradual change of slope.

Acknowledgment

Part of the work described in this paper was developed from investigations undertaken for the Office de Recherches et d'Essais de l'Union Internationale des Chemins de Fer and carried out at British Railways Board, Lon-

don, and at Fraunhofer-Institut für Betriebsfestigkeit (LBF), Darmstadt. The authors wish to acknowledge the help they have had from these organizations.

References

- [1] Haibach, E., *Proceedings*, Conference on Fatigue of Welded Structures, Brighton, England, The Welding Institute, 1970, pp. xx-xxii.
- [2] Spindel, J. E. and Haibach, E., *International Journal of Fatigue*, Vol. 1, No. 2, 1979, pp. 81-88.
- [3] Edwards, A. W. F., *Likelihood, an Account of the Statistical Concept of Likelihood and Its Application to Scientific Inference*. Cambridge University Press, Cambridge, England, 1972.
- [4] Jacobi, G. H. and Nowack, H., *Probabilistic Aspects of Fatigue*. ASTM STP 511, American Society for Testing and Materials, Philadelphia, 1975, pp. 61-74.
- [5] Haibach, E. and Atzori, B., *Aluminium*. Vol. 51, No. 4, 1975, pp. 267-272.
- [6] Hald, A., *Statistical Theory with Engineering Applications*, Wiley, New York, 1952.
- [7] Kendal, M. G. and Stuart, A., *The Advanced Theory of Statistics*. Griffin & Co., London, 1972.
- [8] Spindel, J. E. and Haibach, E., "The Statistical Analysis of Fatigue Test Results," Technical Document DT 71 (D 130) E, Office de Recherches et d'Essais de l'Union Internationale des Chemins de Fer (ORE), Utrecht, The Netherlands, 1979.
- [9] "Fatigue Phenomena in Welded Connections of Bridges and Cranes," Report D 130/RP 10, Office de Recherches et d'Essais de l'Union Internationale des Chemins de Fer (ORE), Utrecht, The Netherlands, 1979.
- [10] Haibach, E. and Matschke, C., "Schwingfestigkeit von Stahl 42 CrMo 4 bei verschiedenen Formzahlen und Spannungsverhältnissen," Report No. FB-153, Fraunhofer-Institut für Betriebsfestigkeit (LBF), Darmstadt, Germany, 1980.

Maximum Likelihood Estimation of a Two-Segment Weibull Distribution for Fatigue Life

REFERENCE: Chou, P. C. and Miller, Harry, "Maximum Likelihood Estimation of a Two-Segment Weibull Distribution for Fatigue Life," *Statistical Analysis of Fatigue Data, ASTM STP 744*, R. E. Little and J. C. Ekvall, Eds., American Society for Testing and Materials, 1981, pp. 114-128.

ABSTRACT: A two-segment distribution is proposed for representation of the fatigue life of modern high-performance composite materials. Each segment is a two-parameter Weibull distribution. Equations of the maximum likelihood method in estimating parameters are derived, and an iterative solution scheme is presented. Several example problems are included.

KEY WORDS: maximum likelihood estimation, Weibull distribution, censored samples, composite materials, fatigue, strength, statistical analysis

Nomenclature

- k Subscript which denotes a segment ($k = 1, 2$)
- $F(x)$ Cumulative distribution function
- $f(x)$ Probability density function
- α_k Weibull shape parameter of the k^{th} segment
- β_k Weibull scale parameter of the k^{th} segment
- θ_k Alternate form of the scale parameter ($\theta_k = \beta_k^{\alpha_k}$)
- δ Value of x separating the two segments (that is, the intersection point)
- N Total sample size ($N = n_1 + n_2 + n_3 + n_4$)
- n_1 Number of failed sample points with a value $\leq \delta$
- n_2 Number of failed sample points with a value $> \delta$
- n_3 Number of suspended sample points with a value $\leq \delta$
- n_4 Number of suspended sample points with a value $> \delta$

¹Professor and visiting assistant professor, respectively, Department of Mechanical Engineering, Drexel University, Philadelphia, Pa. 19104.

- x_i Value of the i^{th} failed specimen ordered so that $x_i \leq x_{i+1}$
 y_i Value of the i^{th} suspended specimen ordered so that $y_i \leq y_{i+1}$

Due to their light weight and high strength, modern high-performance composite materials, such as graphite fiber embedded in epoxy matrix, have been used as structural parts in military and commercial aircraft. They are also being used in sporting goods (skis, tennis rackets, and golf clubs) and are being considered as structural material for automobiles in the 1980s. At the present, the composites have one disadvantage; that is, their strength and fatigue life have larger scatter than those of metals. Extensive testing is currently being carried out to characterize these materials in order to understand their behavior better and to facilitate design applications.

In studying the test data of fatigue life distribution of composites we have observed that in certain cases the distribution is best represented by two distribution functions, one in the low life region and one in the high life region. In this paper, we shall present the maximum likelihood method of estimating parameters of a two-segment distribution where each segment is a two-parameter Weibull distribution.

Two-segmented distribution first appeared in Weibull's paper [1].² In introducing the distribution now bearing his name, Weibull considered two types of this distribution—a simple type and a complex type. His simple type is a standard three-parameter Weibull; the complex type is the sum of two subpopulations. The distribution function of the complex type appears as two straight-line segments in the Weibull coordinate. Weibull showed a few examples, including one on the length of cyrtoideae (a kind of sea shell), and one on the fatigue life of steel; the latter example will also be used for our present approach. Weibull used three-parameter Weibulls for each of his subpopulations and used the trial-and-error method and curve fitting by eye in estimating parameters.

In 1959, Kao [2] also discussed the two-segmented Weibull distribution in connection with failure of electronic tubes. He proposed that the failure can be classified into two types; one is sudden or catastrophic failure (infant mortality) and the other is wear-out or delayed failure. The distribution function of the life of the tube is the sum of two distributions, each of a subpopulation. Kao called this a "mixed distribution," which is similar to the "multi-risk" model discussed by Herman and Patell [3]. Kao further demonstrated that when the characteristic life of the wear-out distribution is large, the mixed distribution can be approximated by a "composite distribution," which is essentially that of the two-segmented Weibull discussed in this paper. He used a two-parameter Weibull for each segment, placed some restrictions on the values of the two shape parameters, and used graphical curve fitting in estimating the parameters.

²The italic numbers in brackets refer to the list of references appended to this paper.

Srivastava [4] studied the problem of life distribution of a specimen subjected to two alternating stress levels. He assumed that the Weibull shape parameters for these two stress levels are the same but the scale parameters are different. The combined distribution for many alternating periods at each of the stress levels is derived.

In reliability engineering, the concept of multisegment distribution is also being used. One example is the piecewise-linear failure rate (hazard function) model, which is one version of the well-known "bathtub curve" [5]. Mann, Schafer, and Singpurwalla [6] also discuss a "two component composite distribution," which is similar to that described by Kao [2].

It seems that the maximum likelihood estimation (MLE) method has not been applied to the two-segment Weibull distribution with unknown parameters. In this paper, we shall first derive the equations of the MLE method, with progressive censoring capability. The general approach is similar to that used by Cohen [7], who has applied the MLE to a single Weibull function, with progressive censoring. A few illustrative examples are given. Two of these examples are for the fatigue life of composite materials.

Two-Segment Weibull Distribution

Let us consider the two-segment Weibull distribution with a probability density function defined as

$$f(x) = \frac{\alpha_k}{\theta_k} x^{\alpha_k - 1} \exp\left[-\frac{x^{\alpha_k}}{\theta_k}\right] \quad (1)$$

and the corresponding cumulative distribution function

$$F(x) = 1 - \exp\left[-\frac{x^{\alpha_k}}{\theta_k}\right] \quad (2)$$

where

$$k = 1 \quad \text{for } x \leq \delta$$

and

$$k = 2 \quad \text{for } x > \delta$$

Each of the two segments is a two-parameter Weibull distribution. In general, at $x = \delta$, $f(x)$ is discontinuous, but $F(x)$ is continuous. When the domain of x extends from zero to infinity, we have the condition

$$\int_0^{\infty} f(x) dx = 1 \quad (3)$$

which is equivalent to $F(\delta)_k = 1 = F(\delta)_k = 2$.

Combining Eqs 2 and 3, we obtain

$$\theta_1 = \theta_2 \delta^{\alpha_1 - \alpha_2} \tag{4}$$

This equation reduces the number of independent parameters from five to four. We shall consider α_1 , α_2 , θ_2 , and δ as our independent parameters.

With a given value of δ , we shall use the maximum likelihood method in estimating the values of the parameters α_k and θ_k for a random sample of N specimens containing $(n_1 + n_2)$ failed specimens and $(n_3 + n_4)$ suspended, or censored, specimens. The censoring can be progressive, that is, any number of specimens can be censored at any time.

The likelihood function for this distribution may be written as

$$L = \text{const} \prod_{i=1}^{n_1} \frac{\alpha_1}{\theta_1} x_i^{(\alpha_1 - 1)} \exp\left[-\frac{x_i^{\alpha_1}}{\theta_1}\right] \cdot \prod_{i=n_1+1}^{n_1+n_2} \frac{\alpha_2}{\theta_2} x_i^{(\alpha_2 - 1)} \exp\left[-\frac{x_i^{\alpha_2}}{\theta_2}\right] \cdot \prod_{i=1}^{n_3} \exp\left[-\frac{y_i^{\alpha_1}}{\theta_1}\right] \cdot \prod_{i=n_3+1}^{n_3+n_4} \exp\left[-\frac{y_i^{\alpha_2}}{\theta_2}\right] \tag{5}$$

Eliminating θ_1 by Eq 4 and taking logarithms of Eq 5, we have

$$\ln L = \ln C + n_1[\ln \alpha_1 + (\alpha_2 - \alpha_1) \ln \delta - \ln \theta_2] + (\alpha_1 - 1) \sum_{i=1}^{n_2} \ln x_i - \delta^{\alpha_2 - \alpha_1} \theta_2^{-1} \left[\sum_{i=1}^{n_1} x_i^{\alpha_1} + \sum_{i=1}^{n_3} y_i^{\alpha_1} \right] + n_2 \ln \frac{\alpha_2}{\theta_2} + (\alpha_2 - 1) \sum_{i=n_1+1}^{n_1+n_2} \ln x_i - \theta_2^{-1} \left[\sum_{i=n_1+1}^{n_1+n_2} x_i^{\alpha_2} + \sum_{i=n_3+1}^{n_3+n_4} y_i^{\alpha_2} \right] \tag{6}$$

Our task now is to find values of the four parameters that maximize the likelihood function, L , or its logarithm, $\ln L$. From Eq 6, it can be seen that $\ln L$ is a continuous function of the parameters α_1 , α_2 , and θ_2 , but a discontinuous function of δ . In maximizing $L' = \ln L$, we cannot use directly the equation $\partial L' / \partial \delta = 0$, because L' has a discontinuity at every data point $\delta = x_i$.

To solve this problem of maximization of a discontinuous function, we shall adopt the following procedure. First, we shall find the values of α_1 , α_2 , and θ_2 that maximize L' for a given value of δ . Then, we shall find the value of δ , when combined with its corresponding values of α_1 , α_2 , and θ_2 , that will

maximize L' . For the first step, we derive the equations of vanishing of the derivatives of L' with respect to α_1 , α_2 , and θ_2 , or

$$\begin{aligned} \frac{\partial L'}{\partial \alpha_1} = 0 &= n_1 \alpha_1^{-1} - n_1 \ln \delta + \sum_{i=1}^{n_1} \ln x_i \\ &\quad - \theta_2^{-1} \delta^{\alpha_2 - \alpha_1} \left[\sum_{i=1}^{n_1} x_i^{\alpha_1} \ln x_i + \sum_{i=1}^{n_3} y_i^{\alpha_1} \ln y_i \right] \\ &\quad + \theta_2^{-1} \delta^{\alpha_2 - \alpha_1} \ln \delta \left[\sum_{i=1}^{n_1} x_i^{\alpha_1} + \sum_{i=1}^{n_3} y_i^{\alpha_1} \right] \quad (7) \end{aligned}$$

$$\begin{aligned} \frac{\partial L'}{\partial \alpha_2} = 0 &= n_1 \ln \delta - \theta_2^{-1} \delta^{\alpha_2 - \alpha_1} \ln \delta \left[\sum_{i=1}^{n_1} x_i^{\alpha_1} + \sum_{i=1}^{n_3} y_i^{\alpha_1} \right] + n_2 \alpha_2^{-1} \\ &\quad + \sum_{i=1+n_1}^{n_1+n_2} \ln x_i - \theta_2^{-1} \left[\sum_{i=n_1+1}^{n_1+n_2} x_i^{\alpha_2} \ln x_i + \sum_{i=n_3+1}^{n_3+n_4} y_i^{\alpha_2} \ln y_i \right] \quad (8) \end{aligned}$$

$$\begin{aligned} \frac{\partial L'}{\partial \theta_2} = 0 &= -n_1 \theta_2^{-1} + \theta_2^{-2} \delta^{\alpha_2 - \alpha_1} \left[\sum_{i=1}^{n_1} x_i^{\alpha_1} + \sum_{i=1}^{n_3} y_i^{\alpha_1} \right] \\ &\quad - n_2 \theta_2^{-1} + \theta_2^{-2} \left[\sum_{i=n_1+1}^{n_1+n_3} x_i^{\alpha_2} + \sum_{i=n_3+1}^{n_3+n_4} y_i^{\alpha_2} \right] \quad (9) \end{aligned}$$

For a given value of δ , Eqs 7, 8, and 9 are solved by a numerical iterative scheme for values of α_1 , α_2 , and θ_2 . A first estimation of the values of α_1 and α_2 is made. These values are substituted into Eq 9 and the value of θ_2 is solved. These values are then used in Eqs 7 and 8 to obtain new estimates of α_1 and α_2 . This process is repeated until the values of all the parameters have converged.

This solution scheme has been programmed for use on an IBM 370 computer. The convergence criterion used in this program compares the value of each parameter to its respective value in the previous iteration. If

$$|\alpha_k - \alpha_{k(\text{previous})}| < 0.0001 \quad (10)$$

and if

$$\left| \frac{[\theta_k - \theta_{k(\text{previous})}]}{\theta_k} \right| < 0.0001 \quad (11)$$

the values of α_k and θ_k are considered satisfactory and the iteration process is stopped. For all the data sets we have studied, convergence has always occurred within 40 iterations, even when the initial estimates were an order of magnitude higher than their final value.

This set of values of $\alpha_1, \alpha_2,$ and $\theta_2,$ which maximizes L' for a given $\delta,$ will be denoted by $\bar{\alpha}_1, \bar{\alpha}_2,$ and $\bar{\theta}_2.$ The likelihood function is then $L'(\bar{\alpha}_1, \bar{\alpha}_2, \bar{\theta}_2, \delta).$

To find the value of δ that maximizes $L'(\bar{\alpha}_1, \bar{\alpha}_2, \bar{\theta}_2, \delta)$ let us consider first the set of discrete values $\delta = x_{n_1}.$ Let

$$L'(x_{n_1}) = L'(\bar{\alpha}_1, \bar{\alpha}_2, \bar{\theta}_2, x_{n_1}) \tag{12}$$

Then the increment of L' between $\delta = x_{n_1+1}$ and $\delta = x_{n_1}$ is

$$\begin{aligned} \Delta L' &= L'(x_{n_1+1}) - L'(x_{n_1}) \\ &= \ln\left(\frac{\alpha_1}{\alpha_2}\right) + (n_1 - 1)(\alpha_1 - \alpha_2)[\ln x_{n_1} - \ln x_{n_1+1}] \\ &\quad - \frac{1}{\theta_2} (\alpha_2 - \alpha_1) (x_{n_1})^{\alpha_2 - \alpha_1 - 1} (x_{n_1+1} - x_{n_1}) \left[\sum_1^{n_1} x_i^{\alpha_1} \right] \\ &\quad - \frac{1}{\theta_2} (x_{n_1})^{\alpha_2 - \alpha_1} (x_{n_1+1})^{\alpha_1} + \frac{1}{\theta_2} (x_{n_1+1})^{\alpha_2} \tag{13} \end{aligned}$$

If among the x_i values $x_{n_1} = x_m$ gives a maximum value of $L',$ then

$$\begin{aligned} \Delta L' &> 0 \quad \text{for } n_1 < m \\ \Delta L' &< 0 \quad \text{for } n_1 \geq m \end{aligned} \tag{14}$$

From Eqs 13 and 14, the value of $\delta = x_m$ that maximizes L' among the discrete points $\delta = x_{n_1}$ can be obtained. Next, let us investigate the value of L' for values of δ in between the data points x_{n_1} and $x_{n_1+1},$ or $x_{n_1} < \delta < x_{n_1+1}.$ In varying δ within this range, the values of n_1 and n_2 do not change, and L' is a continuous function of $\delta.$ Taking the derivative, we obtain, for $x_{n_1} < \delta < x_{n_1+1}$

$$\frac{\partial L'}{\partial \delta} = \frac{n_1(\alpha_1 - \alpha_2)}{\theta_1 \delta} \left[-\beta_1^{\alpha_1} + \frac{1}{n_1} \sum_{i=1}^{n_1} x_i^{\alpha_1} \right] \tag{15}$$

For the range of values of δ between x_{n_1} and $x_{n_1+1},$ the sign of $\partial L'/\partial \delta$ is governed by the terms within the brackets of Eq 15. For the usual case of $\beta_1 > x_{n_1},$ it is negative. If $\beta_1 < x_{n_1},$ it could be positive. In either case, its sign

does not change when δ is varied within the range. Therefore, the maximum value of L' occurs at the data point $\delta = x_{n_1} = x_m$.

Illustrative Examples

We shall present four examples. The first one involves a set of data points taken from a known two-segment distribution. The second one is the fatigue life of Bofors ST-37 steel, which was originally studied by Weibull. The last two examples involve the fatigue life of graphite-epoxy composite materials.

Idealized Data Set

In the first example, we shall start with a hypothetical two-segment Weibull with known values of parameters, select a few points from it, and then apply the MLE to determine the parameters corresponding to these points. These are then compared with the original distribution. The hypothetical distribution selected has the following values

$$\begin{aligned}\alpha_1 &= 2.0 & \alpha_2 &= 0.5 \\ \beta_1 &= 2.3 \times 10^5 & \beta_2 &= 4.0 \times 10^5 \\ \delta &= 1.913 \times 10^5\end{aligned}$$

Twenty points were selected from this distribution with equal ΔF between points. These points, together with the curve of MLE of the distribution, are shown in Fig. 1. The agreement is satisfactory. It can be shown that as the number of points increases, the estimated values of the parameters approach the original value.

In addition, for this data, the value of L' was calculated at several values of δ . The results are shown in Fig. 2. As can be seen, L' attains its maximum at the data point x_{10} , and has continuously decreasing value between data points.

Bofors Steel

The data for this example are taken from Weibull's paper [1]. Fatigue life data of 235 specimens of Bofors St-37 steel under rotating beam tests were recorded. The lives of individual specimens were not given; only the numbers of specimens that failed within certain life intervals were tabulated. These data are reproduced in Table 1. In applying our MLE equations, we have assumed that all specimens with life within a given interval have life at the upper limit of the interval. The results are shown in Fig. 3. The data points are shown by vertical lines bounded by circles, the location of which are calculated according to the median rank formula [8].

Weibull's original fitted curve is also shown in Fig. 3. He used two three-

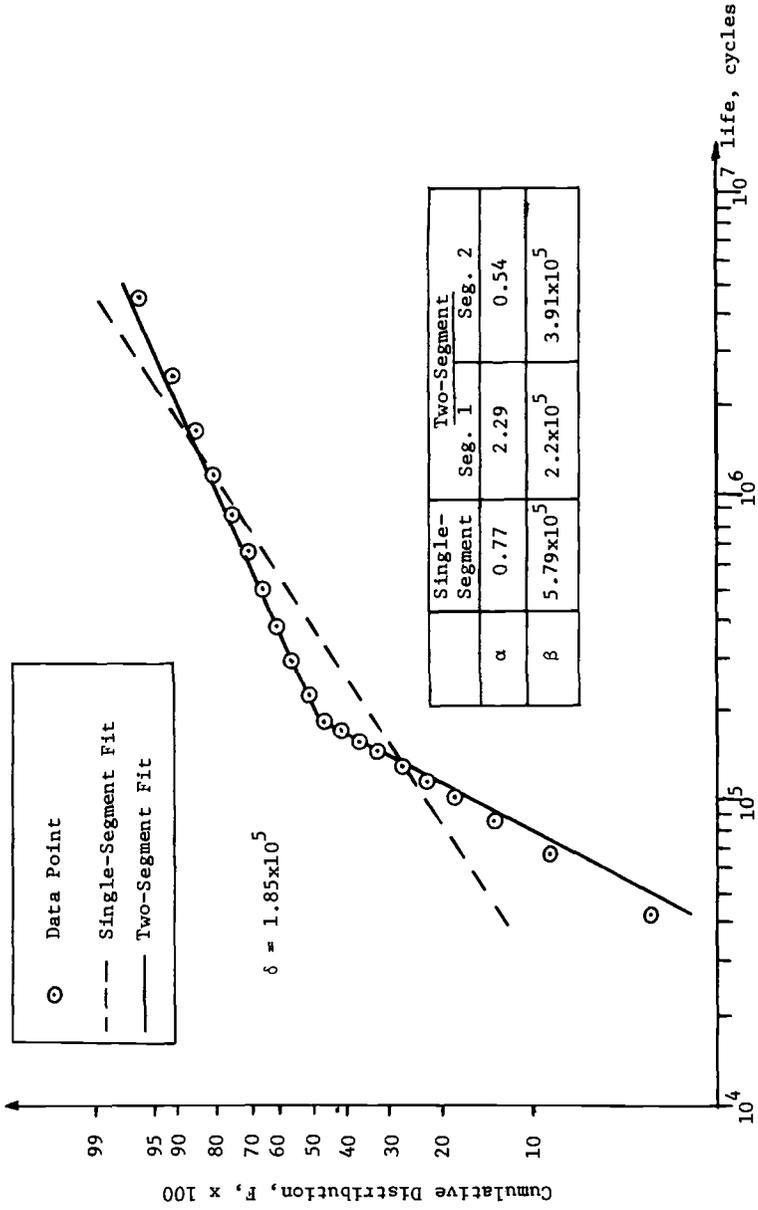


FIG. 1—Comparison between two-segment and simple Weibull distributions for simulated data.

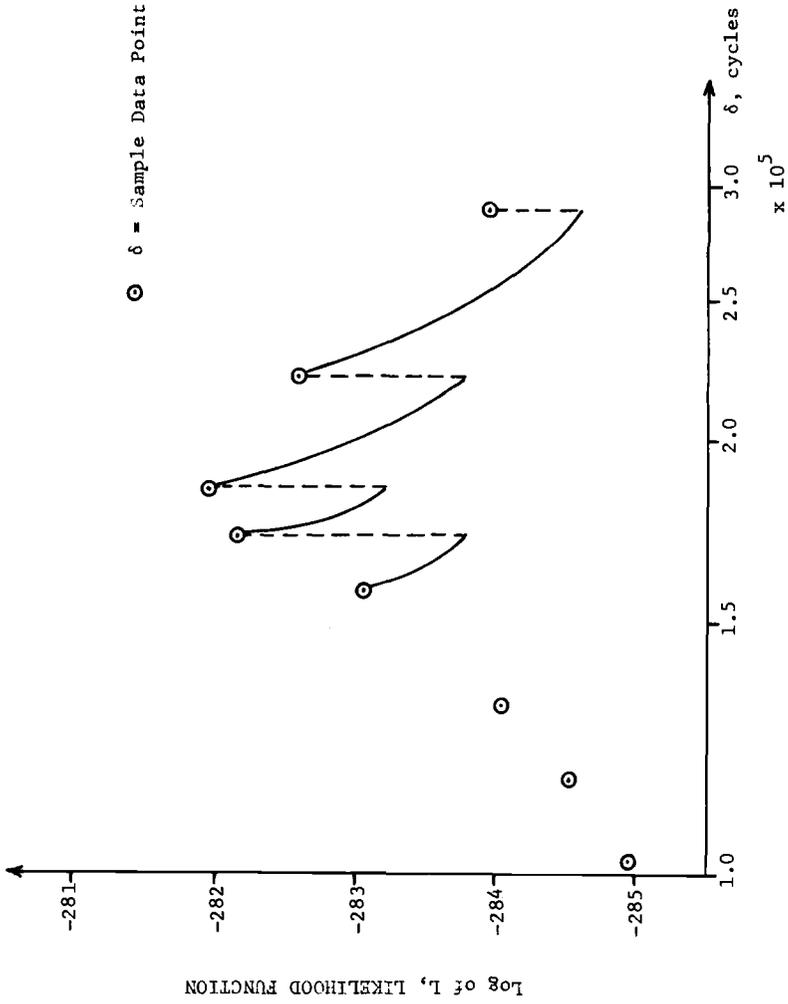


FIG. 2—Value of logarithm of the likelihood function for several values of the partition parameter δ using the simulated data set.

TABLE 1—*Fatigue of Bofors steel [1].^a*

Life, Cycles	Number of Specimens	Life, Cycles	Number of Specimens
12 500 to 17 500	5	47 501 to 52 500	6
17 501 to 22 500	43	52 501 to 57 500	4
22 501 to 27 500	78	57 501 to 62 500	3
27 501 to 32 500	44	62 501 to 67 500	2
32 501 to 37 500	23	67 501 to 72 500	1
37 501 to 42 500	14	72 501 to 77 500	1
42 501 to 47 500	8	77 501 to 82 500	1
		82 501 to 87 500	1
		87 501 to 92 500	1

^aRotating beam test at $\pm 32 \text{ kg/mm}^2$.

parameter Weibull distributions and fitted the curve to the data points visually.

Figure 3 is plotted in the "Weibull coordinates"; two-parameter Weibull functions appear as straight lines, while three-parameter Weibull distributions do not.

Fatigue Life of Composite Material, Complete Samples

In Ref 9, Ryder and Walker tested graphite-epoxy composite laminates which were typical of those used for aircraft structures. We shall study his data for fatigue life under tension-tension fatigue of the Laminate II composites. Details of the specimen lay-up, testing condition, and fatigue life are given in Table 2. Twenty failed data points are available, which represents a complete sample without censoring.

The results are shown in Fig. 4. The solid curve is the estimated two-segment Weibull, and the dotted line is an MLE of a single-function Weibull. The two-segment Weibull shows a good fit to the data.

Fatigue Life of Composite Material, Censored Samples

In Ref 10, Wang, Chou, and Alper have studied the fatigue life of unidirectional graphite-epoxy composites. They used 24 specimens, 20 fatigued to failure and 4 suspended (censored) at 10^6 cycles. Their data are reproduced in Table 3. The estimated distribution is shown in Fig. 5. Again, the fit is satisfactory.

In this case it was found that L' attained maximum value locally at two places. The first maximum occurred at the value shown in the figure. The second occurred at the third from the last failed data point (x_{17}). In order to obtain the best fit to the data in the low life region, the first maxima was chosen for the representation of this sample.

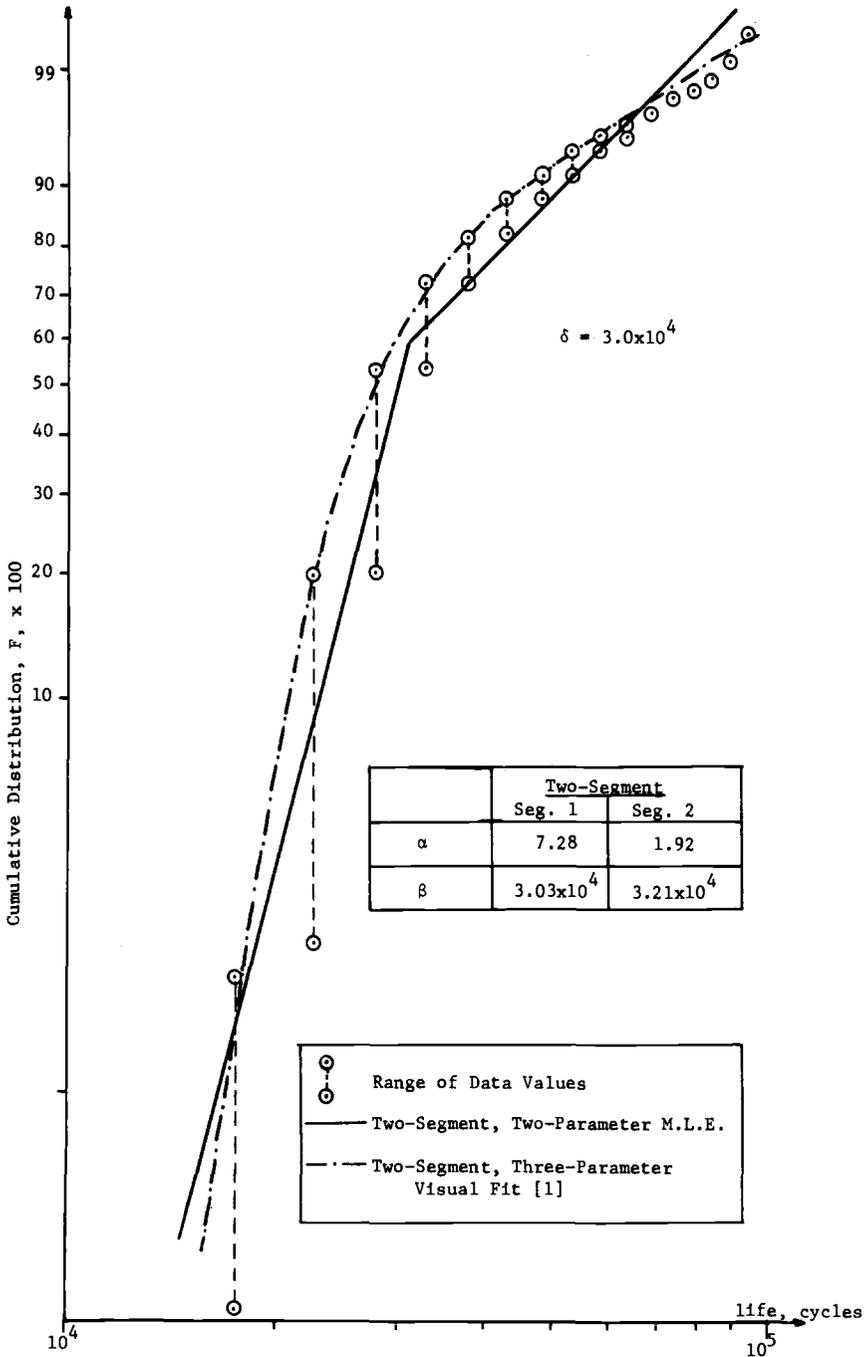


FIG. 3—Comparison between the experimental data and the two-segment distribution for fatigue life of ST-37 steel (Ref 1).

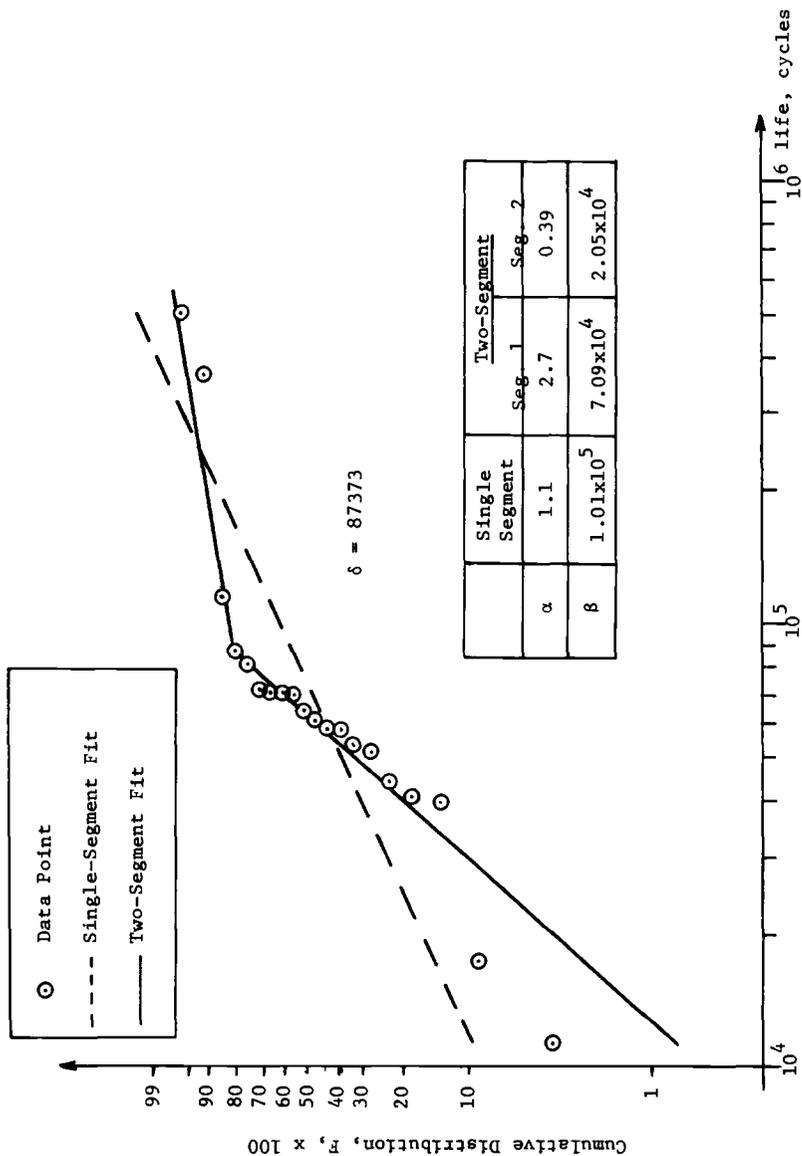


FIG. 4—Comparison between the two-segment and the single-segment Weibull distributions for fatigue life of graphite-epoxy laminates (data from Ref 9).

TABLE 2—*Ryder-Walker tension-tension fatigue tests* [9].^a

11 491	51 848	64 070	81 571
17 578	54 187	69 711	87 373
40 270	58 530	70 049	116 667
41 200	59 320	70 497	367 644
44 830	60 912	71 400	513 600

^aFatigue life, cycles: maximum stress = 50 ksi, $F = 10$ Hz, Gr/E (0/+45/90/-45₂/90/+45/0)_s.

TABLE 3—*Wang-Chou-Alper tension-tension fatigue tests* [10].^a

30	288	5 984	15 754	1 000 000 ^b
69	380	8 609	18 995	1 000 000 ^b
90	1 570	11 362	22 515	1 000 000 ^b
260	3 269	12 119	97 009	1 000 000 ^b
286	5 653	15 529	149 356	

^aFatigue life, cycles: 6 ply Gr/E unidirectional; maximum stress = 171 ksi, $F = 9.5$ Hz.

^bSuspended (censored).

Concluding Remarks

We have shown that the fatigue life of certain materials can be represented by a two-segment distribution, each segment a two-parameter Weibull. The maximum likelihood method is applied for parameter estimation with satisfactory results. We have not studied the failure mechanism or the cause of the failure. It is very likely that two separate failure mechanisms are present. The identification of the fatigue life distribution with two segments of distribution will facilitate the search for the failure mechanisms.

The three-parameter Weibull distribution is often used to represent data that do not agree with a single two-parameter Weibull. With the present method of conveniently fitting a two-segment, two-parameter Weibull, it seems that there is no need to use the three-parameter Weibull. If the variable involved should have a domain from zero to infinity, like fatigue life, there is no physical reason to impose a finite minimum value, as the location parameter does in the three-parameter Weibull. Also, the two-parameter Weibull has the convenience that its shape parameter gives an indication of the degree of scatter in terms of the central value, just like the coefficient of variation. For the three-parameter Weibull, the shape parameter gives the degree of scatter in terms of the central value minus the location parameter, which is more difficult in making comparisons. For instance, in terms of the two-parameter Weibull, the population which has the larger shape parameter has smaller scatter. This type of statement cannot be made for the shape parameter of the three-parameter Weibull. Further discussion on this point will be made in a separate paper.

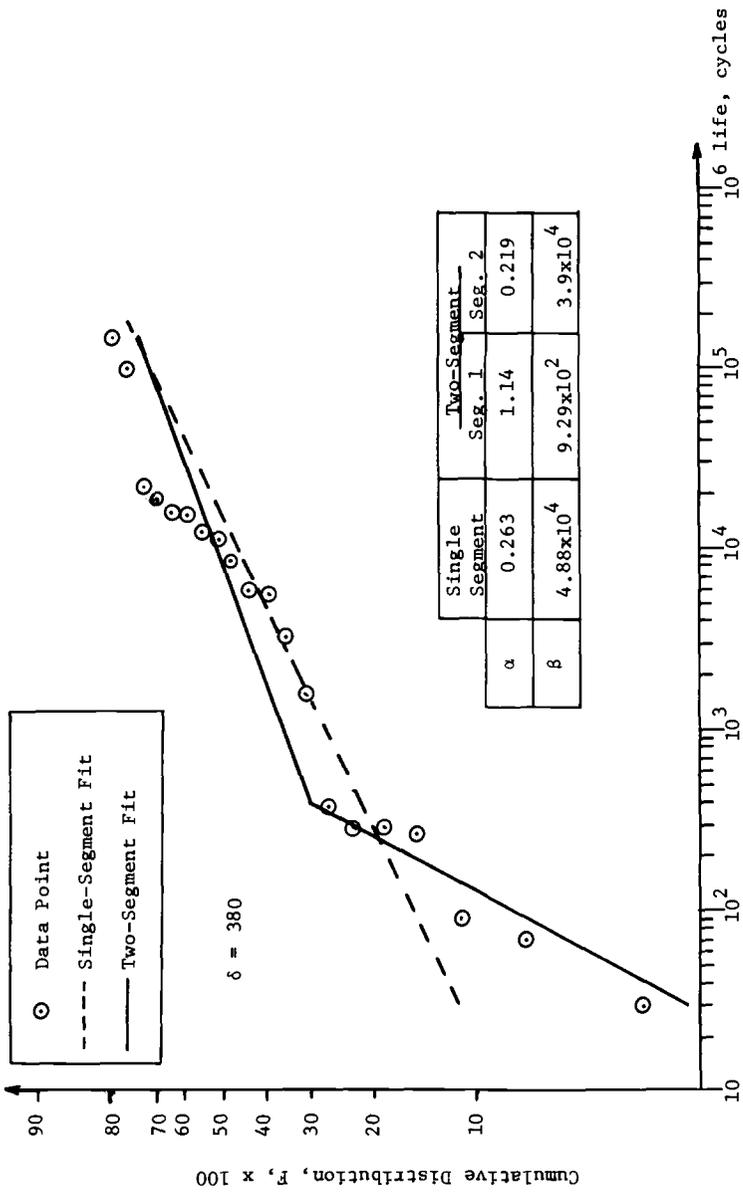


FIG. 5—Comparison between the two-segment and the single-segment Weibull distributions for fatigue of unidirectional composites (data from Ref 10).

Acknowledgment

This work was supported partially by U.S. Air Force Materials Laboratory through a contract to Dyna East Corp., and partially by Ford Motor Co. through a grant to Drexel University.

References

- [1] Weibull, W., *Journal of Applied Mechanics*, Vol. 18, 1951, pp. 293-297.
- [2] Kao, J. H. K., *Technometrics*, Vol. 1, No. 4, Nov. 1959, pp. 389-407.
- [3] Herman, R. J. and Patell, K. N., *Technometrics*, Vol. 13, No. 2, May 1971, pp. 385-396.
- [4] Srivastava, T. N., *I.E.E.E. Transactions on Reliability and Quality Control*, Vol. R-23, No. 2, June 1974, pp. 115-118.
- [5] Shooman, M. L., *Probabilistic Reliability: An Engineering Approach*, McGraw-Hill, New York, 1968, p. 194.
- [6] Mann, N. R., Schafer, R. E., and Singpurwalla, N. D., *Methods for Statistical Analysis of Reliability and Life Data*, Wiley, New York, 1974, pp. 140-141.
- [7] Cohen, A. C., *Technometrics*, Vol. 7, No. 4, Nov. 1965, pp. 579-588.
- [8] Johnson, L. G., *Industrial Mathematics*, Vol. II, 1951, pp. 1-9.
- [9] Ryder, J. T. and Walker, E. K., "Ascertainment of the Effect of Compressive Loading on the Fatigue Lifetime of Graphite Epoxy Laminates for Structural Application," AFML-TR-76-241, Wright-Patterson Air Force Base, Dayton, Ohio, 1976.
- [10] Wang, A. S. D., Chou, P. C., and Alper, J., "Effects of Proof-Test on the Strength and the Fatigue Life of a Unidirectional Composite," *Fatigue of Fibrous Composite Materials*, ASTM STP 723, American Society for Testing and Materials, Philadelphia, 1981.

APPENDIX



Designation: E 739 – 80

Standard Practice for STATISTICAL ANALYSIS OF LINEAR OR LINEARIZED STRESS-LIFE ($S-N$) AND STRAIN-LIFE ($\epsilon-N$) FATIGUE DATA¹

This standard is issued under the fixed designation E 739; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reappraisal.

1. Scope

1.1 This practice pertains only to $S-N$ and $\epsilon-N$ relationships that may be reasonably approximated by a straight line (on appropriate coordinates) for a specific interval of stress or strain. It presents elementary procedures that presently reflect good practice in modeling and analysis. However, because the actual $S-N$ or $\epsilon-N$ relationship is approximated by a straight line only within a specific interval of stress or strain, and because the actual fatigue life distribution is unknown, it is *not recommended* that (a) the $S-N$ or $\epsilon-N$ curve be extrapolated outside the interval of testing, or (b) the fatigue life at a specific stress or strain amplitude be estimated below approximately the fifth percentile ($P \approx 0.05$). As alternative fatigue models and statistical analyses are continually being developed, later revisions of this practice may subsequently present analyses that permit more complete interpretation of $S-N$ and $\epsilon-N$ data.

2. Applicable Documents

2.1 ASTM Standards:

- E 206 Definitions of Terms Relating to Fatigue Testing and the Statistical Analysis of Fatigue Data²
- E 467 Recommended Practice for Verification of Constant Amplitude Dynamic Loads in an Axial Load Fatigue Testing Machine²
- E 468 Recommended Practice for Presentation of Constant Amplitude Fatigue Test Results for Metallic Materials²
- E 513 Definitions of Terms Relating to Con-

stant-Amplitude, Low-Cycle Fatigue Testing²

2.2 Special Technical Publications:³

STP 313 *ASTM Manual on Fitting Straight Lines*

STP 588 *Manual on Statistical Planning and Analysis for Fatigue Experiments*

3. Significance and Use

3.1 Materials scientists and engineers are making increased use of statistical analyses in interpreting $S-N$ and $\epsilon-N$ fatigue data. Statistical analysis applies when the given data can be reasonably assumed to be a random sample of (or representation of) some specific defined population or universe of material of interest (under specific test conditions), and it is desired either to characterize the material or to predict the performance of future random samples of the material (under similar test conditions), or both.

4. Terminology

4.1 The terms used in this practice shall be used as defined in Definitions E 206 and E 513. In addition, the following terminology is used:

4.1.1 *independent variable*—the selected and controlled variable (namely, stress or strain). It

¹ This practice is under the jurisdiction of ASTM Committee E-9 on Fatigue and is the direct responsibility of Subcommittee E09.06 on Statistical Aspects of Fatigue.

Current edition approved June 12, 1980. Published August 1980.

² *Annual Book of ASTM Standards*, Part 10.

³ Available from ASTM, 1916 Race St., Philadelphia, Pa. 19103.



is denoted X herein when plotted on appropriate coordinates.

4.1.2 *replicate (repeat) tests*—nominally identical tests on different randomly selected test specimens conducted at the same nominal value of the independent variable X . Such replicate or repeat tests should be conducted independently; for example, each replicate test should involve a separate set of the test machine and its settings.

4.1.3 *dependent variable*—the fatigue life N (or the logarithm of the fatigue life).

NOTE 1—Log (N) is denoted Y herein.

4.1.4 *log-normal distribution*—the distribution of N when log (N) is normally distributed. (Accordingly, it is convenient to analyze log (N) using methods based on the normal distribution.)

4.1.5 *run out*—no failure at a specified number of load cycles (Recommended Practice E 468).

NOTE 2—The analyses illustrated herein do not apply when the data include either run-outs (or suspended tests). Moreover, the straight-line approximation of the S - N or ϵ - N relationship may not be appropriate at long lives when run-outs are likely.

NOTE 3—For purposes of statistical analysis, a run-out may be viewed as a test specimen that has either been removed from the test or is still running at the time of the data analysis.

5. Types of S - N and ϵ - N Curves Considered

5.1 It is well known that the shape of S - N and ϵ - N curves can depend markedly on the material and test conditions. This practice is restricted to linear or linearized S - N and ϵ - N relationships, for example,

$$\log N = A + B(S) \text{ or} \quad (1a)$$

$$\log N = A + B(\epsilon), \text{ or} \quad (1b)$$

$$\log N = A + B(\log S) \text{ or} \quad (2a)$$

$$\log N = A + B(\log \epsilon) \quad (2b)$$

in which S and ϵ may refer to (a) the maximum value of constant-amplitude cyclic stress or strain, given a specific value of the stress or strain ratio, or of the minimum cyclic stress or strain, (b) the amplitude or the range of the constant-amplitude cyclic stress or strain, given a specific value of the mean stress or strain or (c) analogous information stated in terms of some appropriate independent (controlled) variable.

NOTE 4—In certain cases the amplitude of the

stress or strain is not constant during the entire test for a given specimen. In such cases some effective (equivalent) value of S or ϵ must be established for use in analysis.

5.1.1 The fatigue life N is the dependent (random) variable in S - N and ϵ - N tests, whereas S or ϵ is the independent (controlled) variable.

NOTE 5—In certain cases the independent variable used in analysis is not literally the variable controlled during testing. For example, it is common practice to analyze low-cycle fatigue data treating the range of plastic strain as the controlled variable, when in fact the range of total strain was actually controlled during testing. Although there may be some question regarding the exact nature of the controlled variable in certain S - N and ϵ - N tests, there is never any doubt that the fatigue life is the dependent variable.

NOTE 6—In plotting S - N and ϵ - N curves, the independent variables S and ϵ are plotted along the ordinate, with life (the dependent variable) plotted along the abscissa. Refer, for example, to Fig. 1.

5.1.2 The distribution of fatigue life (in any test) is unknown (and indeed may be quite complex in certain situations). For the purposes of simplifying the analysis (while maintaining sound statistical procedures), it is assumed herein that the logarithms of the fatigue lives are normally distributed, that is, the fatigue life is log-normally distributed, and that the variance of log life is constant over the entire range of the independent variable used in testing (that is, the scatter in log N is assumed to be the same at low S and ϵ levels as at high levels of S or ϵ). Accordingly, log N is used as the dependent (random) variable in analysis. It is denoted Y . The independent variable is denoted X . It may be either S or ϵ , or log S or log ϵ , respectively, depending on which appears to produce a straight line plot for the interval of S or ϵ of interest. Thus Eqs 1 and 2 may be re-expressed as

$$Y = A + BX \quad (3)$$

Equation 3 is used in subsequent analysis. It may be stated more precisely as $\mu_{Y|X} = A + BX$, where $\mu_{Y|X}$ is the expected value of Y given X .

NOTE 7—For testing the adequacy of the linear model see 8.2.

NOTE 8—The expected value is the mean of the conceptual population of all Y 's given a specific level of X . (The median and mean are identical for the symmetrical normal distribution assumed herein for Y .)



6. Test Planning

6.1 Test planning for *S-N* and *ε-N* test programs is discussed in Chapter 3 of Ref (2). Planned grouping (blocking) and randomization are essential features of a well-planned test program. In particular, good test methodology involves use of planned grouping to (a) balance potentially spurious effects of nuisance variables (for example, laboratory humidity) and (b) allow for possible test equipment malfunction during the test program.

7. Sampling

7.1 It is vital that sampling procedures be adopted which assure a random sample of the material being tested. A random sample is required to state that the test specimens are representative of the conceptual universe about which both statistical and engineering inference will be made.

NOTE 9—A random sampling procedure provides each specimen that conceivably could be selected (tested) an equal (or known) opportunity of actually being selected at each stage of the sampling process. Thus, it is poor practice to use specimens from a single source (plate, heat, supplier) when seeking a random sample of the material being tested unless that particular source is of specific interest.

NOTE 10—Procedures for using random numbers to obtain random samples and to assign stress or strain amplitudes to specimens (and to establish the time order of testing) are given in Chapter 4 of Ref (3).

7.1.1 *Sample Size*—The minimum number of specimens required in *S-N* (and *ε-N*) testing depends on the type of test program conducted. The following guidelines given in Chapter 3 of Ref (2) appear reasonable.

Type of Test	Minimum Number of Specimens ^a
Preliminary and exploratory (exploratory research and development tests)	6 to 12
Research and development testing of components and specimens	6 to 12
Design allowables data	12 to 24
Reliability data	12 to 24

^a If the variability is large, a wide confidence band will be obtained unless a large number of specimens are tested (See 8.1.1).

7.1.2 *Replication*—The replication guidelines given in Chapter 3 of Ref (2) are based on the following definition:

% replication = 100 [(total number of different stress or strain levels used in testing/total number of specimens tested)]

Type of Test	Percent Replication ^a
Preliminary and exploratory (research and development tests)	17 to 33 min
Research and development testing of components and specimens	33 to 50 min
Design allowables data	50 to 75 min
Reliability data	75 to 88 min

^a Note that percent replication indicates the portion of the total number of specimens tested that may be used for obtaining an estimate of the variability of replicate tests.

7.1.2.1 *Replication Examples*—Good replication: Suppose that 10 specimens are used in research and development for the testing of a component. If two specimens are tested at each of five stress or strain amplitudes, the test program involves 50 % replications. This percent replication is considered adequate for most research and development applications. Poor replication: Suppose eight different stress or strain amplitudes are used in testing, with two replicates at each of two stress or strain amplitudes (and no replication at the other six stress or strain amplitudes). This test program involves only 20 % replication, which is not generally considered adequate.

8. Statistical Analysis (Linear Model $Y = A + BX$, Log-Normal Fatigue Life Distribution with Constant Variance Along the Entire Interval of X Used in Testing, No Run-outs or Suspended Tests or Both, Completely Randomized Design Test Program)

8.1 For the case where (a) the fatigue life data pertain to a random sample (all Y_i are independent), (b) there are neither run-outs nor suspended tests and where, for the entire interval of X used in testing, (c) the *S-N* or *ε-N* relationship is described by the linear model $Y = A + BX$ (more precisely by $\mu_{Y|X} = A + BX$), (d) the (two parameter) log-normal distribution describes the fatigue life N , and (e) the variance of the log-normal distribution is constant, the maximum likelihood estimators of A and B are as follows:

$$\hat{A} = \bar{Y} - \hat{B}\bar{X} \tag{4}$$

$$\hat{B} = \frac{\sum_{i=1}^k (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^k (X_i - \bar{X})^2} \tag{5}$$

where the symbol "caret" ($\hat{}$) denotes estimate (estimator), the symbol "overbar" ($\bar{}$) denotes average (for example, $\bar{Y} = \sum_{i=1}^k Y_i/k$ and $\bar{X} =$

$\sum_{i=1}^k X_i/k$, $Y_i = \log N_i$, $X_i = S_i$ or ϵ_i , or $\log S_i$ or $\log \epsilon_i$ (refer to Eqs 1 and 2), and k is the total number of test specimens (the total sample size). The recommended expression for estimating the variance of the normal distribution for $\log N$ is

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^k (Y_i - \hat{Y}_i)^2}{k - 2} \tag{6}$$

in which $\hat{Y}_i = \hat{A} + \hat{B}X_i$ and the $(k - 2)$ term in the denominator is used instead of k to make $\hat{\sigma}^2$ an unbiased estimator of the normal population variance σ^2 .

NOTE 11—An assumption of constant variance is usually reasonable for notched and joint specimens up to about 10^8 cycles to failure. The variance of unnotched specimens generally increases with decreasing stress (strain) level (see Section 9). If the assumption of constant variance appears to be dubious the reader is referred to Ref (5) for the appropriate statistical test.

8.1.1 *Confidence Intervals for Parameters A and B*—The estimators \hat{A} and \hat{B} are normally distributed with expected values A and B , respectively, (regardless of total sample size k) when conditions (a) through (e) in 8.1 are met. Accordingly, confidence intervals for parameters A and B can be established using the t distribution, Table 1. The confidence interval for A is given by $\hat{A} \pm t_p \hat{\sigma}_A$, or

$$\hat{A} \pm t_p \hat{\sigma} \left[\frac{1}{k} + \frac{\bar{X}^2}{\sum_{i=1}^k (X_i - \bar{X})^2} \right]^{1/2} \tag{7}$$

and for B is given by $\hat{B} \pm t_p \hat{\sigma}_B$, or

$$\hat{B} \pm t_p \hat{\sigma} \left[\sum_{i=1}^k (X_i - \bar{X})^2 \right]^{-1/2} \tag{8}$$

in which the value of t_p is read from Table 1 for the desired value of P , the confidence level associated with the confidence interval. This table has one entry parameter (the statistical degrees of freedom, n , for t). For Eqs 7 and 8, $n = k - 2$.

NOTE 12—The confidence intervals for A and B are exact if conditions (a) through (e) in 8.1 are met exactly. However, these intervals are still reasonably accurate when the actual life distribution differs slightly from the (two-parameter) log-normal distribution, that is, when only condition (d) is not met exactly, due to the robustness of the t statistic.

NOTE 13—Because the actual median $S-N$ or $\epsilon-N$ relationship is only approximated by a straight line within a specific interval of stress or strain, confidence intervals for A and B that pertain to confidence levels

greater than approximately 0.95 are not recommended.

The meaning of the confidence interval associated with, say, Eq 8 is as follows (NOTE 14). If the values of t_p given in Table 1 for, say, $P = 95\%$ are used in a series of analyses involving the estimation of B from independent data sets, then in the long run we may expect 95% of the computed intervals to include the value B . If in each instance we were to assert that B lies within the interval computed, we should expect to be correct 95 times in 100 and in error 5 times in 100: that is, the statement “ B lies within the computed interval” has a 95% probability of being correct. But there would be no operational meaning in the following statement made in any one instance: “The probability is 95% that B falls within the computed interval in this case” since B either does or does not fall within the interval. It should also be emphasized that even in independent samples from the same universe, the intervals given by Eq 8 will vary both in width and position from sample to sample. (This variation will be particularly noticeable for small samples.) It is this series of (random) intervals “fluctuating” in size and position that will include, ideally, the value B 95 times out of 100 for $P = 95\%$. Similar interpretations hold for confidence intervals associated with other confidence levels. For a given total sample size k , it is evident that the width of the confidence interval for B will be a minimum whenever

$$\sum_{i=1}^k (X_i - \bar{X})^2$$

is a maximum. Since the X_i levels are selected by the investigator, the width of confidence interval for B may be reduced by appropriate test planning. For example, the width of the interval will be minimized when, for a fixed number of available test specimens, k , half are tested at each of the extreme levels X_{min} and X_{max} . However, this allocation should be used only when there is strong *a priori* knowledge that the $S-N$ or $\epsilon-N$ curve is indeed linear—because this allocation precludes a statistical test for linearity (8.2). See Chapter 3 of Ref (2) for a further discussion of efficient selection of stress (or strain) levels and the related specimen allocations to these stress (or strain) levels.

NOTE 14—This explanation is similar to that of STP 313.

E 739

8.1.2 *Confidence Band for the Entire Median S-N or e-N Curve (that is, for the Median S-N or e-N Curve as a Whole)*—If conditions (a) through (e) in 8.1 are met, an exact confidence band for the entire median S-N or e-N curve (that is, all points on the linear or linearized median S-N or e-N curve considered simultaneously) may be computed using the following equation:

$$\hat{A} + \hat{B}X \pm \sqrt{2F_p} \hat{\sigma} \left[\frac{1}{k} + \frac{(X - \bar{X})^2}{\sum_{i=1}^k (\bar{X}_i - \bar{X})^2} \right]^{1/2} \quad (9)$$

in which F_p is given in Table 2. This table involves two entry parameters (the statistical degrees of freedom n_1 and n_2 for F). For Eq 9, $n_1 = 2$ and $n_2 = (k - 2)$. For example, when $k = 7$, $F_{0.95} = 5.7861$.

A 95 % confidence band computed using Eq 9 is plotted in Fig. 1 for the example data of 8.2.1. The interpretation of this band is similar to that for a confidence interval (8.1.1). Namely, if conditions (a) through (e) are met, and if the values of F_p given in Table 2 for, say, $P = 95\%$ are used in a series of analyses involving the construction of confidence bands using Eq 9 for the entire range of X used in testing; then in the long run we may expect 95 % of the computed hyperbolic bands to include the straight line $\mu_{YX} = A + BX$ everywhere along the entire range of X used in testing.

NOTE 15—Because the actual median S-N or e-N relationship is only approximated by a straight line within a specific interval of stress of strain, confidence bands which pertain to confidence levels greater than approximately 0.95 are not recommended.

While the hyperbolic confidence bands generated by Eq. 9 and plotted in Fig. 1 are statistically correct, straight-line confidence and tolerance bands parallel to the fitted line $\hat{\mu}_{YX} = \hat{A} + \hat{B}X$ are sometimes used. These bands are described in Chapter 5 of Ref (3).

8.2 *Testing the Adequacy of the Linear Model*—In 8.1 it was assumed that a linear model is valid, namely that $\mu_{YX} = A + BX$. If the test program is planned such that there is more than one observed value of Y at some of the X_i levels where $i \geq 3$, then a statistical test for linearity can be made based on the F distribution, Table 2. The log life of the j th replicate specimen tested in the i th level of X is subsequently denoted Y_{ij} .

Suppose that fatigue tests are conducted at l different levels of X and that m_i replicate values of Y are observed at each X_i . Then the hypothesis of linearity (that $\mu_{YX} = A + BX$) is rejected when the computed value of

$$\frac{\sum_{i=1}^l m_i (\bar{Y}_i - \hat{Y}_i)^2 / (l - 2)}{\sum_{i=1}^l \sum_{j=1}^{m_i} (Y_{ij} - \bar{Y}_i)^2 / (k - l)} \quad (10)$$

exceeds F_p , where the value of F_p is read from Table 2 for the desired significance level. (The significance level is defined as the probability in percent of incorrectly rejecting the hypothesis of linearity when there is indeed a linear relationship between X and μ_{YX} .) The total number of specimens tested, k , is computed using

$$k = \sum_{i=1}^l m_i \quad (11)$$

Table 2 involves two entry parameters (the statistical degrees of freedom n_1 and n_2 for F). For Eq 10, $n_1 = (l - 2)$, and $n_2 = (k - l)$. For example, $F_{0.95} = 6.9443$ when $k = 8$ and $l = 4$.

The F test (Eq 10) compares the variability of average value about the fitted straight line, as measured by their mean square (NOTE 15) (the numerator in Eq 10) to the variability among replicates, as measured by their mean square (the denominator in Eq 10). The latter mean square is independent of the form of the model assumed for the S-N or e-N relationship. If the relationship between μ_{YX} and X is indeed linear, Eq 10 follows the F distribution with degrees of freedom, $(l - 2)$ and $(k - l)$. Otherwise Eq 10 is larger on the average than would be expected by random sampling from this F distribution. Thus the hypothesis of a linear model is rejected if the observed value of F (Eq 10) exceeds the tabulated value F_p . If the linear model is rejected, it is recommended that a nonlinear model be considered, for example:

$$\mu_{YX} = A + BX + CX^2 \quad (12)$$

NOTE 16—Some readers may be tempted to use existing digital computer software which calculates a value of r , the so-called correlation coefficient, or r^2 , the coefficient of determination, to ascertain the suitability of the linear model. This approach is not recommended. (For example, $r = 0.993$ with $F = 3.62$ for the example of 8.3.1, whereas $r = 0.988$ and $F = 21.5$ for similar data set generated during the 1976 E09.08 low-cycle fatigue round robin).

NOTE 17—A mean square value is a specific sum of squares divided by its statistical degrees of freedom.

E 739

8.3 Numerical Examples:

8.3.1 Example 1: Consider the following low-cycle fatigue data (taken from a 1976 E09.08 round-robin test program (laboratory 43)):

$\Delta\epsilon_p/2$ Plastic Strain Amplitude—Unitless	N Fatigue Life Cycles
0.01636	168
0.01609	200
0.00675	1000
0.00682	1180
0.00179	4730
0.00160	8035
0.00165	5254
0.00053	28617
0.00054	32650

(a) Estimate parameters A and B and the respective 95 % confidence intervals. First, restate (transform) the data in terms of logarithms (base 10 used herein due to its wide use in practice).

$X_i = \log(\Delta\epsilon_{pi}/2)$ (Independent Variable)	$Y_i = \log N_i$ (Dependent Variable)
-1.78622	2.22531
-1.79344	2.30103
-2.17070	3.00000
-2.16622	3.07188
-2.74715	3.67486
-2.79588	3.90499
-2.78252	3.72049
-3.27572	4.45662
-3.26761	4.51388

Then, from Eqs 4 and 5:

$$\hat{A} = -0.24474 \quad \hat{B} = -1.45144$$

Or, as expressed in the form of Eq 2b:

$$\widehat{\log N} = -0.24474 - 1.45144 \log(\Delta\epsilon_p/2)$$

Also, from Eq 6:

$$\hat{\sigma}^2 = 0.07837/7 = 0.011195$$

Or,

$$\hat{\sigma} = 0.1058$$

Accordingly, using Eq 7, the 95 % confidence interval for A is ($t_p = 2.3646$) $[-0.6435, 0.1540]$, and, using Eq 8, the 95 % confidence interval for B is $[-1.6056, -1.2974]$.

The fitted line $\hat{Y} = \widehat{\log N} = -0.24474 - 1.45144 \log(\Delta\epsilon_p/2) = -0.24474 - 1.45144X$ is displayed in Fig. 1, where the 95 % confidence band computed using Eq 9 is also plotted. (For example, when $\Delta\epsilon_p/2 = 0.01$, $X = -2.000$, $\hat{Y} = 2.65814$, $\hat{Y}_{\text{lower band}} = 2.65814 - 0.15215 = 2.50599$ and $\hat{Y}_{\text{upper band}} = 2.65814 +$

$$0.15215 = 2.81029.)$$

The fitted line can be transformed to the form given in Appendix XI of Recommended Practice E 606 as follows:

$$\widehat{\log N} = -0.24474 - 1.45144 \log(\Delta\epsilon_p/2)$$

$$\log(\Delta\epsilon_p/2) = -0.16862 - 0.68897 \widehat{\log N}$$

$$\Delta\epsilon_p/2 = 0.67823 (\hat{N})^{-0.68897}$$

Substituting cycles (\hat{N}) to reversals ($2\hat{N}_f$) gives

$$\Delta\epsilon_p/2 = 0.67823 \left(\frac{2\hat{N}_f}{2}\right)^{-0.68897}$$

$$\Delta\epsilon_p/2 = 0.67823 (1/2)^{-0.68897} (2\hat{N}_f)^{-0.68897}$$

$$\Delta\epsilon_p/2 = 1.09340 (2\hat{N}_f)^{-0.68897}$$

The above alternative equation is shown on Fig. 1.

Ancillary Calculations:

$$\bar{X} = -2.53172 \quad \bar{Y} = 3.42990$$

$$\sum_{i=1}^9 (X_i - \bar{X})^2 = 2.63892$$

$$\sum_{i=1}^9 (X_i - \bar{X})(Y_i - \bar{Y}) = -3.83023$$

$$\hat{\sigma}_A = \hat{\sigma} \left[\frac{1}{9} + \frac{(-2.53172)^2}{2.63892} \right]^{1/2} = 0.1686$$

$$\hat{\sigma}_B = \hat{\sigma} [2.63892]^{-1/2} = 0.06513$$

(b) Test for linearity at the 5 % significance level.

We shall ignore the slight differences among the amplitudes of plastic strain and assume that $l = 4$ and $\kappa = 9$. Then, at each of the four X_i levels, we shall compute \hat{Y}_i using $\hat{Y}_i = -0.24414 - 1.45144\bar{X}_i$ and \hat{Y}_i using $\hat{Y}_i = \Sigma Y_{ij}/m_i$. Accordingly, $F_{0.95} = 5.79$, whereas F computed (using Eq 10) = 3.62. Hence, we do not reject the linear model in this example.

Ancillary Calculations:

$$\text{Numerator } (F) = 0.0532/2$$

$$\text{Denominator } (F) = 0.0368/5$$

8.3.2 Example 2: Consider the following low-cycle fatigue data (also taken from a 1976 E09.08 round-robin test program (laboratory 34)):

$\Delta\epsilon_p/2$ Plastic Strain Amplitude—Unitless	N Fatigue Life Cycles
0.0164	153
0.0164	153
0.0069	563
0.0069	694

E 739

$\Delta\epsilon_p/2$ Plastic Strain Amplitude—Unitless	N Fatigue Life Cycles
0.00185	3515
0.00175	3860
0.00054	17500
0.00058	20330
0.000006	60350
0.000006	121500

The F test (Eq 10) in this case indicates that the linear model should be rejected at the 5% significance level (that is, F calculated = 9.08, where $F_{3,5,0.05} = 5.41$). Hence estimation of A and B for the linear model is not recommended. Rather, a nonlinear model should be considered in analysis.

9. Other Statistical Analyses

9.1 When the Weibull distribution is assumed to describe the distribution of fatigue life at a given stress or strain amplitude, or when the fatigue data include either run-outs or suspended tests (or when the variance of log life increases noticeably as life increases), the appropriate statistical analyses are more complicated than illustrated herein. The reader is referred to Ref (4) for an example of relevant digital computer software.

NOTE 18—It is not good practice either to ignore run-outs or to treat them as if they were failures. Rather, maximum likelihood analyses of the type illustrated in Ref (4) are recommended.

REFERENCES

- (1) *ASTM Manual on Fitting Straight Lines*, STP 313, Am. Soc. Testing Mats., 1962.
- (2) *Manual on Statistical Planning and Analysis for Fatigue Experiments*, STP 588, Am. Soc. Testing Mat., 1975.
- (3) Little, R. E., and Jebe, E. H., *Statistical Design of Fatigue Experiments*, Applied Science Publishers, London, 1975.
- (4) Nelson, W. B., et al, "STATPAC Simplified—A Short Introduction To How To Run STATPAC, A General Statistical Package for Data Analysis," *Technical Information Series Report 73CRD 046*, July, 1973, General Electric Co., Corporate Research and Development, Schenectady, N. Y.
- (5) Brownlee, K. A., *Statistical Theory and Methodology in Science and Engineering*, John Wiley and Sons, New York, 2nd Ed. 1965.

TABLE 1 Values of t_p (Abstracted from STP 313)

n^a	$P, \%^A$	
	90	95
4	2.1318	2.7764
5	2.0150	2.5706
6	1.9432	2.4469
7	1.8946	2.3646
8	1.8595	2.3060
9	1.8331	2.2622
10	1.8125	2.2281
11	1.7959	2.2010
12	1.7823	2.1788
13	1.7709	2.1604
14	1.7613	2.1448
15	1.7530	2.1315
16	1.7459	2.1199
17	1.7396	2.1098
18	1.7341	2.1009
19	1.7291	2.0930
20	1.7247	2.0860
21	1.7207	2.0796
22	1.7171	2.0739

^A P is the probability in percent that the random variable t lies in the interval from $-t_p$ to $+t_p$.

^a n is not sample size, but the degrees of freedom of t , that is, $n = k - 2$.



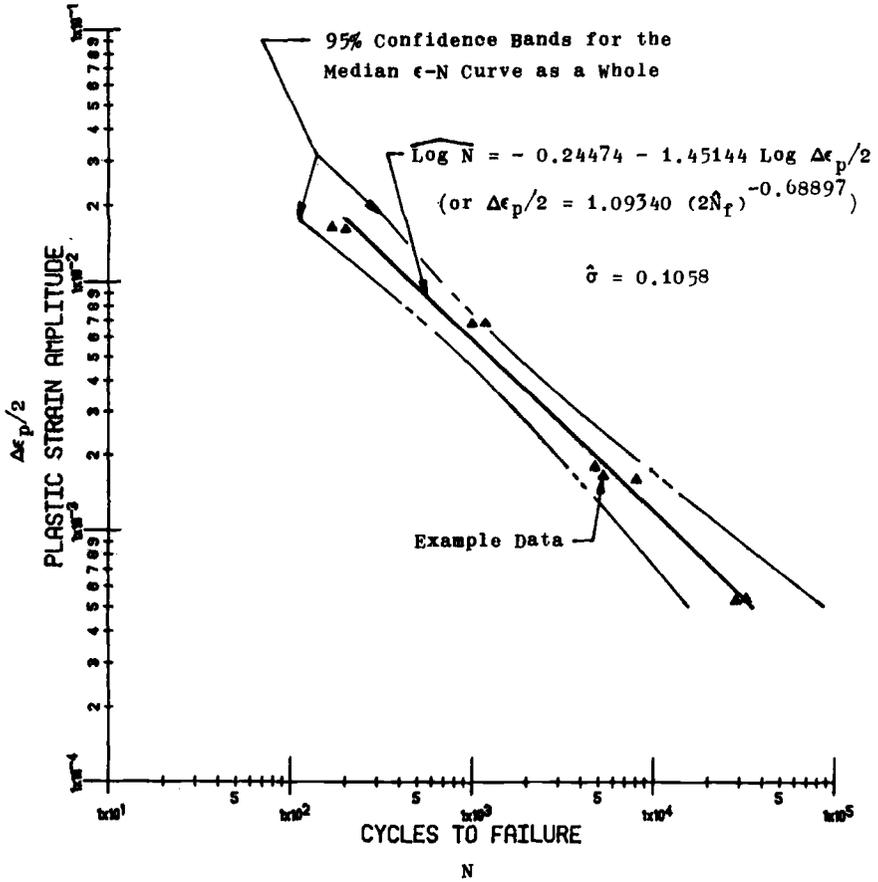
E 739

TABLE 2 Values of F_{α} ^A (Abstracted from STP 313)

		Degrees of Freedom, n_1			
		1	2	3	4
Degrees of Freedom, n_2	1	161.45 4052.2	199.50 4999.5	215.71 5403.3	224.58 5624.6
	2	18.513 98.503	19.000 99.000	19.164 99.166	19.247 99.249
	3	10.128 34.116	9.5521 30.817	9.2766 29.457	9.1172 28.710
	4	7.7086 21.198	6.9443 18.000	6.5914 16.694	6.3883 15.977
	5	6.6079 16.258	5.7861 13.274	5.4095 12.060	5.1922 11.392
	6	5.9874 13.745	5.1433 10.925	4.7571 9.7795	4.5337 9.1483
	7	5.5914 12.246	4.7374 9.5466	4.3468 8.4513	4.1203 7.8467
	8	5.3177 11.259	4.4590 8.6491	4.0662 7.5910	3.8378 7.0060
	9	5.1174 10.561	4.2565 8.0215	3.8626 6.9919	3.6331 6.4221
	10	4.9646 10.044	4.1028 7.5594	3.7083 6.5523	3.4780 5.9943
	11	4.8443 9.6460	3.9823 7.2057	3.5874 6.2167	3.3567 5.6683
	12	4.7472 9.3302	3.8853 6.9266	3.4903 5.9526	3.2592 5.4119
	13	4.6672 9.0738	3.8056 6.7010	3.4105 5.7394	3.1791 5.2053
	14	4.6001 8.8616	3.7389 6.5149	3.3439 5.5639	3.1122 5.0354
	15	4.5431 8.6831	3.6823 6.3589	3.2874 5.4170	3.0556 4.8932

^A In each row, the top figures are values of F corresponding to $P = 95\%$, the bottom figures correspond to $P = 99\%$. Thus, the top figures pertain to the 5% significance level, whereas the bottom figures pertain to the 1% significance level. (The bottom figures are not recommended for use in Eq 9).

 E 739



NOTE—The 95% confidence band for the ϵ - N curve as a whole is based on Eq 10. (Note that the dependent variable, fatigue life, is plotted here along the abscissa to conform to engineering convention.)

FIG. 1 Fitted Relationship Between the Fatigue Life N (Y) and the Plastic Strain Amplitude $\Delta\epsilon_p/2$ (X) for the Example Data Given

Summary

The paper by Little is an appropriate introductory paper for this Special Technical Publication. There is much confusion among practicing engineers regarding the use and meaning of probability statements associated with fatigue test data. Little explains the difference between confidence, prediction, and tolerance expressions and includes illustrative examples. He also covers one-sided lower statistical tolerance limits for log normal and Weibull distributions for Type I and Type II censoring. He then discusses the use of maximum likelihood analysis to generate estimates that may be compared with analogous estimates generated by alternative statistical procedures. Fatigue life data are generally not either log normal or Weibull distributed; therefore, a nonparametric approach for establishing tolerance limits, as discussed by Little, is more appropriate. This, however, requires larger sample sizes than are generally practical in fatigue applications. For most future applications the maximum likelihood estimation technique, which is referenced and referred to in Little's article, will be the preferred statistical analysis procedure for analyzing fatigue life data.

The paper by Haibach, Olivier, and Rinaldi discusses practical aspects of planning and conducting round-robin fatigue test programs, which in this case consisted of 753 $S-N$ tests carried out by six test laboratories with specimens fabricated by three welding institutes. Anyone contemplating a round-robin or interlaboratory test program should review this paper thoroughly. With any program of this type it is necessary to keep test scatter among laboratories to a minimum so the effects of significant variables can be accounted for in the statistical analysis with a minimum number of specimens. In this regard preliminary estimates of the results were made and confirmed by preliminary testing so the proper load levels could be selected for testing. Also, recommended modes of static and dynamic calibration of the load-measuring device were made, and a method was described for checking the exact alignment of the clamping devices. The variables considered in the study included two materials, three types of welded specimens, three stress ratios, and two to four stress levels. The program gives very complete results for the constant amplitude fatigue properties for the materials and specimens tested. However, at some additional cost, programmed tests should have been conducted to evaluate the effect of representative service loadings on the fatigue life of these materials and joints.

Young and Ekvall have made a very extensive evaluation of variability in spectrum (5000 specimens) and $S-N$ (2417 specimens) fatigue test data. The

results provide an estimate of the variability one can expect in fatigue tests of different materials (aluminum, titanium, steel, and nickel) and with different types of testing (flight-by-flight, block loading, and constant amplitude). A review is given of different statistically based equations used for determining scatter factors which have been applied to test results for safe-life designed parts. The paper illustrates how probability distribution functions can be used to fit the variability in fatigue test data by using regression analysis. The methods can be applied to evaluate the scatter in other types of test data, such as those from fracture toughness tests, crack growth tests, and others.

Nishijima considers the distribution of strength deviation values of individual test results about the median $S-N$ curve. With this approach all the samples tested can be combined to develop the median $S-N$ curve and a $P-S-N$ diagram. Nishijima uses probit analysis to estimate the medium fatigue strengths at small intervals in the life scale. The values obtained from the probit analysis were used to compute the best-fit polynomial equation of the $S-N$ curve. The deviations about the median curve are pooled to determine the distribution of relative strength values and to develop $P-S-N$ diagrams. Although the fatigue strength for the seven materials considered appears to be normally distributed, Nishijima points out that not all fatigue limit data support an assumption of normality. The advantages of the approach are that the data can be pooled to estimate variability in fatigue strength, runout data can be included in the analysis, and a more general shape of the median $S-N$ curve can be obtained. This last point was demonstrated by the falloff in the curve for two materials above 10^7 cycles, which was attributed to the effect of environmental exposure. Nevertheless, the problem remains that a fairly large number of specimens is still required to develop a reliable $P-S-N$ diagram.

Spindel and Haibach studied some of the problems associated with fitting $S-N$ curves to fatigue test data. They used the "support approach," based on maximum likelihood principles, to investigate fitting $S-N$ curves with up to five parameters: (1) the fatigue strength at 2×10^6 cycles, (2) the slope of the linear portion of the $S-N$ curve, (3) the standard deviation of the logarithm of applied stress amplitude, and (4) and (5) the position at which the slope changes, in terms of cycles and stress amplitude. A computer program, discussed elsewhere, was used to define the optimum combination of parameters by a trial and error procedure—that is, by varying some parameters while other parameters are held constant. A method of combining sets of similar data and defining confidence limits is also discussed. The results of this method indicate that $S-N$ curves can be determined within closer limits by using several sets of data from similar tests, assuming that the curves are similar and differ only in one parameter. Also, a better fit curve can be obtained if it is assumed that the standard deviation of the logarithm of stress remains constant along the curve and the slope is permitted to change gradually from the high-stress to the high-cycle range of data. This last assumption differs from that used in Nishijima's paper, where the coefficient of variation is

assumed constant, rather than the standard deviation. Also both of these assumptions differ from that used in regression analysis, where the variance of the logarithm life is assumed constant over the entire range of the independent variable, stress. The methods discussed by Nishijima and by Spindel and Haibach are more appropriate for large samples of test data.

Chow and Miller illustrate the use of maximum likelihood analysis in estimating the parameters of a relatively sophisticated mathematical model. The model assumed in their analysis is a two-segment distribution where each segment is a two-parameter Weibull distribution. The maximum likelihood method includes data for any number of specimens censored at any time as well as for failed specimens. An iterative procedure for obtaining three parameters while holding the intersection point constant was programmed on the IBM 370 computer. Convergence to a solution was obtained within 40 iterations for all the data sets these authors studied. The optimum value of the intersection points is then determined for the three parameters obtained from the maximum likelihood analysis. It appears that the values obtained for the three parameters will depend on the initial choice of the intersection point, and one should perform the analysis for several values of the intersection point to obtain the best fit. Also the examples of test data given in the paper do not clearly demonstrate that the two-segment Weibull distribution is better than a single three-parameter Weibull distribution.

This volume also includes a reprint of the ASTM Standard Practice for Statistical Analysis of Linear or Linearized Stress-Life ($S-N$) and Strain-Life ($\epsilon-N$) Fatigue Data (E 739-80). The practice introduces maximum likelihood estimation and a statistical test for assessing the adequacy of the mathematical model, but it is rather a modest document, leaving room for numerous extensions. In particular, one hopes that in the 1980s the practice will be extended to cover maximum likelihood analyses which permit (a) employing a nonlinear $S-N$ curve, (b) including suspended tests (runouts) in the data, (c) assuming that the variance of fatigue life increases along the $S-N$ curve, (d) presuming that a Weibull or some other distribution describes the random variability of fatigue life along the $S-N$ curve, (e) considering independent multiple modes (causes) of failure, and (f) comparing two or more $S-N$ curves using the likelihood ratio test statistic.

R. E. Little

University of Michigan, Dearborn, Mich.
48128; symposium chairman and editor.

J. C. Ekvall

Lockheed-California Co., Burbank, Calif.
91520; symposium cochairman and editor.

Index

A

a posteriori probability, 5
a priori probability, 4
 A-basis tolerance limit, 21
 Analysis of covariance, 40, 41, 53
 Analysis of variance, 50, 51

B

B-basis tolerance limit, 21
 Bias, 21

C

Censoring
 Type I, 12
 Type II, 12
 Censored samples, 14, 123
 Chi-square distribution, 70, 72, 73
 Chi-square paper, 72, 73
 Chi-square test (*see* Goodness of fit test)
 Coefficient of variation, 78
 Mechanical properties, 78, 85-87
 Confidence intervals, 6
 Confidence limits, 6, 38, 94, 98, 99, 132
 Correlation coefficient, r , 63-65, 133
 Covariance, analysis of, 40, 41, 53

D

Degrees of freedom, 8, 65
 Distribution
 Log-normal (*see* Log-normal distribution)
 Normal (*see* Normal distribution)
 Weibull (*see* Weibull distribution)

E

Endurance limit, 89, 91, 92, 104-112
 Experiment design
 Incomplete block, 25, 50

F

Fatigue strength, 36, 76, 100, 101
 Coefficient of variation, 78
 Fatigue tests
 Types of loading, 57, 58
 Runouts, 96, 130

G

Goodness of fit test
 Chi-square test, 64-66

H

Hardness distribution (*see* Vickers hardness)

I

Interlaboratory test program, 25-32

L

Least squares analysis, 100, 109, 112
 Weighted, 78, 79

Linear regression analysis, 61

Normal distribution, 63

Quadratic (*see* Polynomial regression analysis)

Weibull distribution, 64

Logarithmic standard deviation, 61-63

Probability distribution equations, 64, 65

Probability distribution plots, 67, 68

Probability and confidence level values, 71

Log-normal distribution, 16, 57, 64-68, 92, 95, 131

M

Material properties

Mechanical, 27, 82

Maximum likelihood analysis, 18, 90, 95, 109, 110

Estimator, 96

Likelihood ratio, 18, 19, 99, 140

Support, 90, 95

Median rank (*see* Probability plotting position)

Median value, 63

Miner's rule, 89, 90

Multiple regression analysis, 64, 65

N

Nonparametric tolerance limit, 20

Normal distribution, 62, 63, 78, 95

Normal population

Mean estimator, 7

Standard deviation estimator, 7

O

One-sided probability intervals, 4

One-sided lower tolerance limits, 11

Ordered data, 14, 63, 79

P

P-S-N diagram, 76

Polynomial regression analysis (*see* Multiple regression analysis)

Population standard deviation, 70

Precision, 95

Probabilistic diagrams, 37, 38, 47, 67, 68, 73

Probability

a posteriori, 5

a priori, 4

Expression, 3, 6, 8, 10

Interval, 5

Limits, 4

Normal, 96

Plotting positions, 63, 79

Probit analysis, 78

R

Random interval, 6

Random numbers, 29, 32

Random variable, realizations of, 4

Replication recommendations, for *S-N* testing, 131

S

Sample size, 131

Sampling, 131

Scatter

In fatigue life, 59, 76

In fatigue strength, 45, 76, 103

Laboratory effects, 38
 Variables affecting, 57, 76
 Welding effects, 40
 Scatter factors, 70
S-N curve
 Analytical form, 35, 44, 90, 91,
 94, 100, 103, 104-112, 130
 Confidence band, 137
 Graphical plots, 31, 33, 45, 77,
 82-86, 91, 93, 94, 100, 102,
 106, 108, 111, 137
 Slope, 29, 35, 42, 90, 91, 94, 100,
 103, 104-112, 130-132
 Support method, 99, 100, 104-110
 Staircase tests 28, 32, 33
 Standard deviation, 62, 78, 95, 96,
 99, 100, 103, 104, 107, 109,
 112
 Standard error of the median fatigue
 limit estimate, 79
 Support values, 97, 98, 100, 101
 Suspended tests (*see* Censoring)

T

Test plan, 32, 49
 Randomized Blocks, 32
 Replicates, 32

Treatments, 32, 40
 Test program (conduct), 11, 131
 Tolerance Interval, 6, 10
 Tolerance Limit
 A-basis, 20
 B-basis, 20
 One-sided lower, 11
 Nonparametric, 20

U

Unbiased estimator of population
 variance, 132

V

Variability (*see* Scatter)
 Variance, 63, 95, 99
 Analysis of, 50, 51
 Vickers hardness distribution, 81
 Coefficient of Variation, 87

W

Weibull distribution, 14, 18, 56,
 61, 64-68, 92, 135
 Two-segment, 116
 Welded joints, 26, 27, 76, 90

