INTERNATIONAL STANDARD

ISO 12107

Second edition
2012-08-15

Metallic materials — Fatigue testing — Statistical planning and analysis of data

Matériaux métalliques — Essais de fatigue — Programmation et analyse statistique de données

Reference number
ISO 12107:2012(E)

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 12107 was prepared by Technical Committee ISO/TC 164, *Mechanical testing of metals*, Subcommittee SC 5, *Fatigue testing*.

This second edition cancels and replaces the first edition (ISO 12107:2003), which has been technically revised.

Introduction

It is known that the results of fatigue tests display significant variations even when the test is controlled very accurately. In part, these variations are attributable to non-uniformity of test specimens. Examples of such non-uniformity include slight differences in chemical composition, heat treatment, surface finish, etc. The remaining part is related to the stochastic process of fatigue failure itself that is intrinsic to metallic engineering materials.

Adequate quantification of this inherent variation is necessary to evaluate the fatigue property of a material for the design of machines and structures. It is also necessary for test laboratories to compare materials in fatigue behaviour, including its variation. Statistical methods are necessary to perform these tasks. This International Standard includes a full methodology for application of the Bastenaire model as well as other more sophisticated relationships. It also addresses the analysis of runout (censored) data.

Metallic materials — Fatigue testing — Statistical planning and analysis of data

1 Scope

1.1 Objectives

This International Standard presents methods for the experimental planning of fatigue testing and the statistical analysis of the resulting data. The purpose is to determine the fatigue properties of metallic materials with both a high degree of confidence and a practical number of specimens.

1.2 Fatigue properties to be analysed

This International Standard provides a method for the analysis of fatigue life properties at a variety of stress levels using a relationship that can linearly approximate the material's response in appropriate coordinates.

Specifically, it addresses

a) the fatigue life for a given stress, and

b) the fatigue strength for a given fatigue life.

The term "stress" in this International Standard can be replaced by "strain", as the methods described are also valid for the analysis of life properties as a function of strain. Fatigue strength in the case of strain-controlled tests is considered in terms of strain, as it is ordinarily understood in terms of stress in stress-controlled tests.

1.3 Limit of application

This International Standard is limited to the analysis of fatigue data for materials exhibiting homogeneous behaviour due to a single mechanism of fatigue failure. This refers to the statistical properties of test results that are closely related to material behaviour under the test conditions.

In fact, specimens of a given material tested under different conditions may reveal variations in failure mechanisms. For ordinary cases, the statistical property of resulting data represents one failure mechanism and may permit direct analysis. Conversely, situations are encountered where the statistical behaviour is not homogeneous. It is necessary for all such cases to be modelled by two or more individual distributions.

An example of such behaviour is often observed when failure can initiate from either a surface or internal site at the same level of stress. Under these conditions, the data will have mixed statistical characteristics corresponding to the different mechanisms of failure. These types of results are not considered in this International Standard because a much higher complexity of analysis is required.

Finally, for the S-N case (discussed in Clause 8), this International Standard addresses only complete data. Runouts of censored data are not addressed.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 3534 (all parts), *Statistics — Vocabulary and symbols*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 3534 and the following apply.

3.1 Terms related to statistics

3.1.1
confidence level
value $1 - \alpha$ of the probability associated with an interval of statistical tolerance

3.1.2
degrees of freedom
ν
number calculated by subtracting from the total number of observations the number of parameters estimated from the data

3.1.3
distribution function
function giving, for every value x, the probability that the random variable X is less than or equal to x

3.1.4
estimation
operation made for the purpose of assigning, from the values observed in a sample, numerical values to the parameters of a distribution from which this sample has been taken

3.1.5
population
totality of individual materials or items under consideration

3.1.6
random variable
variable that may take any value of a specified set of values

3.1.7
sample
one or more items taken from a population and intended to provide information on the population

3.1.8
size
n
number of items in a population, lot, sample, etc.

3.1.9
mean
μ
sum of all the data in a population divided by the number of observations

3.1.10
sample mean
$\hat{\mu}$
sum of all the data in a sample divided by the number of observations

3.1.11
standard deviation
σ
positive square root of the mean squared standard deviation from the mean from a population.

3.1.12
estimated standard deviation
$\hat{\sigma}$
positive square root of the mean squared standard deviation from the mean of a sample.

3.2 Terms related to fatigue

3.2.1
fatigue life
N
number of stress cycles applied to a specimen, at an indicated stress level, before it attains a failure criterion defined for the test

3.2.2
fatigue limit
fatigue strength at long life

NOTE Historically, this has usually been defined as the stress generating a life at 10^7 cycles.

3.2.3
fatigue strength
value of stress level S at which a specimen would fail at a given fatigue life

NOTE This is expressed in megapascals.

3.2.4
specimen
portion or piece of material to be used for a single test determination and normally prepared in a predetermined shape and in predetermined dimensions

3.2.5
stress level
S
intensity of the stress under the conditions of control in the test

EXAMPLES Amplitude, maximum, range.

3.2.6
stress step
d
difference between neighbouring stress levels when conducting the test by the staircase method

NOTE This is expressed in megapascals.

4 Statistical distributions in fatigue properties

4.1 Concept of distributions in fatigue

The fatigue properties of metallic engineering materials are determined by testing a set of specimens at various stress levels to generate a fatigue life relationship as a function of stress. The results are usually expressed as an S-N curve that fits the experimental data plotted in appropriate coordinates. These are generally either log-log or semi-log plots, with the life values always plotted on the abscissa on a logarithmic scale.

Fatigue test results usually display significant scatter even when the tests are carefully conducted to minimize experimental error. A component of this variation is due to inequalities, related to chemical composition or heat treatment, among the specimens, but another component is related to the fatigue process, an example being the initiation and growth of small cracks under test environments.

The variation in fatigue data are expressed in two ways: the distribution of fatigue life at a given stress and the distribution of strength at a given fatigue life (see References [1] to [5]).

4.2 Distribution of fatigue life

Fatigue life, N, at a given test stress, S, is considered as a random variable. It is frequently observed the distribution of fatigue life values at any stress is normal in the logarithmic metric. That is, the logarithms of the life values follow a normal distribution (See 6.4). This relationship is:

$$P(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \int_{-\infty}^{x} \exp\left[-\frac{1}{2} \left(\frac{x - \mu_x}{\sigma_x} \right)^2 \right] \mathrm{d}x \qquad (1)$$

where $x = \log N$ and μ_x and σ_x are, respectively, the mean and the standard deviation of x.

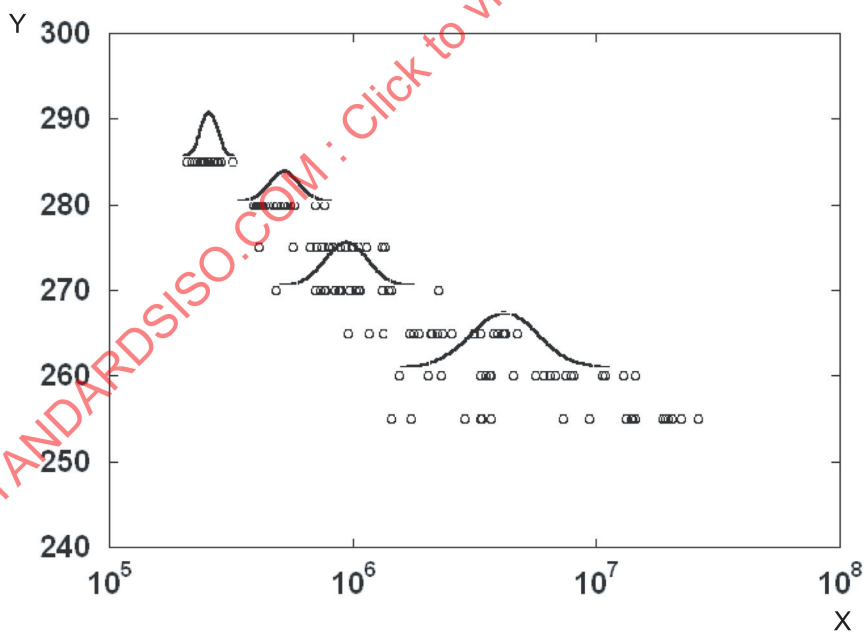
Formula (1) gives the cumulative probability of failure for x. This is the proportion of the population failing at lives less than or equal to x.

Formula (1) does not relate to the probability of failure for specimens at or near the fatigue limit. In this region, some specimens may fail, while others may not. The shape of the distribution is often skewed, displaying even greater scatter on the longer-life side. It also may be truncated to represent the longest failure life observed in the data set.

This International Standard does not address situations in which a certain number of specimens may fail, but the remaining ones do not.

Other statistical distributions can also be used to express variations in fatigue life. The Weibull [4] distribution is one of the statistical models often used to represent skewed distributions. On occasion, this distribution may apply to lives at low stresses, but this special case is not addressed in this International Standard.

Figure 1 shows an example of data from a fatigue test conducted with a statistically based experimental plan using a large number of specimens (see Reference [5]). The shape of the fatigue life distributions is demonstrated for explanatory purposes.



Key

X cycles to failure

Y stress amplitude, in MPa

Figure 1 — Concept of variation in a fatigue property — Distribution of fatigue life at given stresses for a 0,25 % C carbon steel tested in the rotating-bending mode

4.3 Distribution of fatigue strength

Fatigue strength at a given fatigue life, N, is considered as a random variable. It is expressed as the normal distribution:

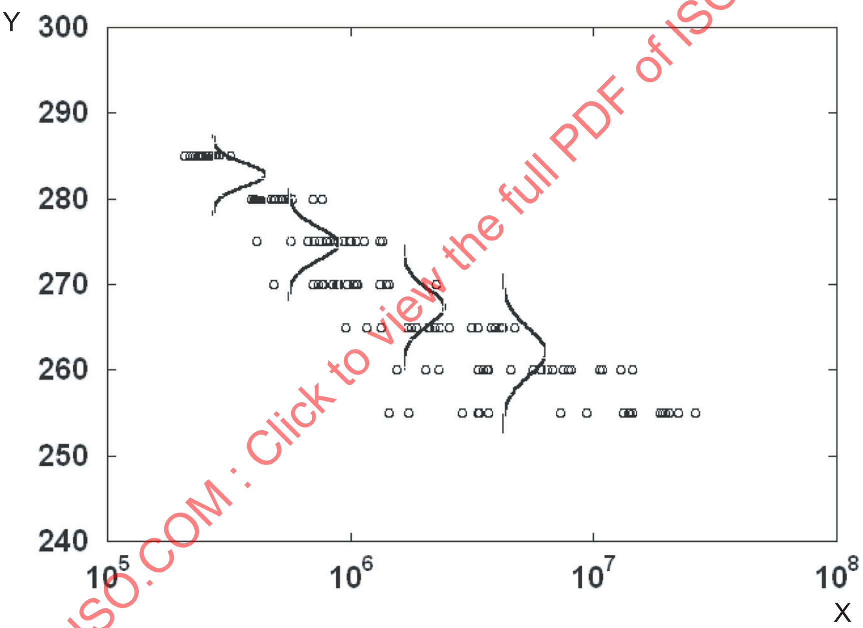
$$P(y) = \frac{1}{\sigma_y \sqrt{2\pi}} \int_{-\infty}^{y} \exp\left[-\frac{1}{2} \left(\frac{y - \mu_y}{\sigma_y} \right)^2 \right] \mathrm{d}y \tag{2}$$

where $y = S$ (the fatigue strength at N), and μ_y and σ_y are, respectively, the mean and the standard deviation of y.

Formula (2) gives the cumulative probability of failure for y. It defines the proportion of the population presenting fatigue strengths less than or equal to y.

Other statistical distributions can also be used to express variations in fatigue strength.

Figure 2 is based on the same experimental data as Figure 1. The variation in the fatigue property is expressed here in terms of strength at typical fatigue lives (see Reference [5]).



Key

X cycles to failure

Y stress amplitude, in MPa

Figure 2 — Concept of variation in a fatigue property — Distribution of fatigue strength at typical fatigue lives for a 0,25 % C carbon steel tested in the rotating-bending mode

5 Statistical planning of fatigue tests

5.1 Sampling

It is necessary to define clearly the population of the material for which the statistical distribution of fatigue properties is to be estimated. Specimen selection from the population shall be performed in a random fashion. It is also important that the specimens be selected so that they accurately represent the population they are intended to describe. A complete plan would include additional considerations.

If the population consists of several lots or batches of material, the test specimens shall be selected randomly from each group in a number proportional to the size of each lot or batch. The total number of specimens taken shall be equal to the required sample size, n.

If the population displays any serial nature, e.g. if the properties are related to the date of fabrication, the population shall be divided into groups related to time. Random samples shall be selected from each group in numbers proportional to the group size.

The specimens taken from a particular batch of material will reveal variability specific to the batch. This within-batch variation can sometimes be of the same order of importance as the between-batch variation. When the relative importance of different kinds of variation is known from experience, sampling shall be performed taking this into consideration.

Hardness measurement is recommended for some materials, when possible, to divide the population of the material into distinct groups for sampling. The groups should be of as equal size as possible. Specimens may be extracted randomly in equal numbers from each group to compose a test sample of size n. This procedure will generate samples uniformly representing the population, based upon hardness.

5.2 Allocation of specimens for testing

Specimens taken from the test materials shall be allocated to individual fatigue tests in principle in a random way, in order to minimize unexpected statistical bias. The order of testing of the specimens shall also be randomized in a series of fatigue tests.

When several test machines are used in parallel, specimens shall be tested on each machine in equal or nearly equal numbers and in a random order. The equivalence of the machines in terms of their performance shall be verified prior to testing.

When the test programme includes several independent test series, e.g. tests at different stress levels or on different materials for comparison purposes, each test series shall be carried out at equal or nearly equal rates of progress, so that all testing can be completed at approximately the same time.

6 Statistical estimation of fatigue life at a given stress

6.1 Testing to obtain fatigue life data

Conduct fatigue tests at a given stress, S, on a set of carefully prepared specimens to determine the fatigue life values for each. The number selected will be dependent upon the purpose of the test and the availability of test material. A set of seven specimens is recommended in this International Standard for exploratory tests. For reliability purposes, however, at least 28 specimens are recommended.

6.2 Plotting data on normal probability paper

Plot the fatigue lives on log-normal probability coordinates. The results should plot as a straight line. Should one or two data points (really a very low proportion of the data set) deviate from the curve, this is usually the result of invalid data. Examining test records and failed specimens is useful when there is non-conforming behaviour. The purpose is to identify a cause for such deviant behaviour to learn if these results can be discounted. Other statistical distributions e.g. Weibull may be evaluated. However, since the vast majority of unimodal fatigue results have proven to be distributed log-normally, the standard does not consider Weibull statistics. Subclause 8.3.3 gives some examples of normal probability plots constructed from data used to generate an S-N curve. Refer to these plots to understand how they will appear when the data conform well to the assumption and in other cases when there might be some issues. Please note that for the present case, the y-axis will just be the property in question as opposed to the standardized residuals given the y-axis on the presented plots in 8.3.3.

One other issue is that if the data appear to support two distinct failure distributions, the data should be segregated by the root cause. For example, results for both surface and internal initiation sites should be separated into two groups and evaluated uniquely.

6.3 Estimating distribution parameters

Calculation of the sample mean is performed as follows:

$$\hat{\mu} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{3}$$

where

$\hat{\mu}$ is the sample mean;

x_i is the ith observed value;

n is the number of data points.

Note that the symbol "^" means an estimation based upon a sample.

The sample standard deviation is calculated using the following relationship:

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n} \left(x_i - \hat{\mu}\right)^2}{n-1}} \tag{4}$$

6.4 Quantitative evaluation of the assumption of normality

A number of statistical tests have been developed attempting to quantitatively consider the assumption of normality. These tests can sometimes generate conflicting results. However, one that seems quite useful is the Anderson-Darling Test. The details for performing this evaluation as well as others can be found in Reference [9]. Also, there are commercially available statistical software packages that perform quantitative evaluations of normality.

6.5 Estimating the lower limit of the fatigue life

Estimate the lower limit of the fatigue life at a given probability of failure, assuming a normal distribution, at the confidence level $1 - \alpha$ from the equation:

$$\hat{x}_{(P,1-\alpha)} = \hat{\mu}_x - k_{(P,1-\alpha,\nu)}\hat{\sigma}_x \tag{5}$$

The coefficient $k_{(P,\,1-\alpha,\,\nu)}$ is the one-sided tolerance limit for a normal distribution, as given in Table B.1. P corresponds to the reliability of the prediction (say 99 % probability) and $1 - \alpha$ is the confidence of the reliability statement. These values are generated by integration of the non-central t distribution with non-centrality parameter:

$$\delta = \sqrt{n} \tag{6}$$

The number of degrees of freedom, ν, is the same number used in estimating the standard deviation. For the present case, this is $n - 1$.

A worked example is given in A.1.

7 Statistical estimation of fatigue strength at a given fatigue life

7.1 Testing to obtain fatigue strength data

Conduct fatigue tests to generate strength data for a set of specimens in a sequential way using the method known as the staircase method (see Reference [7]).

It is necessary to have rough estimates of the mean and the standard deviation of the fatigue strength for the materials to be tested. Start the test at a first stress level preferably close to the estimated mean strength. Also select a stress step, preferably close to the standard deviation, by which to vary the stress level during the test.

If no information is available about the standard deviation, a step of about 5 % of the estimated mean fatigue strength may be used as the stress step.

Test a first specimen, randomly chosen, at the first stress level to find if it fails before the given number of cycles. For the next specimen, also randomly chosen, increase the stress level by a step if the preceding specimen did not fail, and decrease the stress by the same amount if it failed. Continue testing until all the specimens have been tested in this way.

Exploratory research requires a minimum of 15 specimens to estimate the mean and the standard deviation of the fatigue strength. Reliability data requires at least 30 specimens.

A worked example of the staircase method is given in A.2.1, together with worked examples of the analyses described in 7.2 and 7.3.

7.2 Statistical analysis of test data

Count the frequencies of failure and non-failure of the specimens tested at different stress levels. Use the analysis for the group with the least number of observations.

Denote the stress levels arranged in ascending order by $S_0 \leq S1 \leq \ldots \leq S_l$, where l is the number of stress levels, denote the number of events by f_i, and denote the stress step by d. Estimate the parameters for the statistical distribution of the fatigue strength, Formula (2), from:

$$\hat{\mu}_y = S_0 + d\left(\frac{A}{C} \pm \frac{1}{2}\right) \tag{7}$$

$$\hat{\sigma}_y = 1{,}62d\left(D + 0{,}029\right) \tag{8}$$

where:

$$A = \sum_{i=1}^{l} i f_i$$

$$B = \sum_{i=1}^{l} i^2 f_i$$

$$C = \sum_{i=1}^{l} f_i$$

$$D = \frac{BC - A^2}{C^2}$$

In Formula (7), take the value of ± 1/2 as:

− 1/2 when the event analysed is failure;

+ 1/2 when the event analysed is non-failure.

In Reference [7], it is stated that Formula (8) is valid only when $D > 0{,}3$. This condition is generally satisfied when $d / \hat{\sigma}_y$ is selected properly within the range 0,5 to 2.

7.3 Estimating the lower limit of the fatigue strength

Estimate the lower limit of the fatigue strength at a probability of failure P for the population at a confidence level of $1 - \alpha$, if the assumption of a normal distribution of the fatigue strength is correct, from the equation:

$$\hat{y}_{(P,\,1-\alpha)} = \hat{\mu}_y - k_{(P,\,1-\alpha,\,\nu)}\hat{\sigma}_y \qquad (9)$$

where the coefficient $k_{(P,\,1-\alpha,\,\nu)}$ is the one-sided tolerance limit for a normal distribution, as given in Table B.1.

Take as the number of degrees of freedom, ν, the number that was used in estimating the standard deviation. For the present case, this is $n - 1$.

7.4 Modified method when standard deviation is known

A modified staircase method, with fewer specimens, is possible if the standard deviation is known and only the mean of the fatigue strength needs to be estimated (see Reference [8]).

Conduct tests as in the staircase method described in 7.1, by decreasing or increasing the stress level by a fixed step depending whether the preceding event was a failure or non-failure, respectively. Choose the initial stress level close to the roughly estimated mean and the stress step approximately equal to the known standard deviation.

A minimum of six specimens is required for exploratory tests and at least 15 for reliability data.

If the test is conducted on n specimens at stress levels $S_1, S_2, \ldots, S_n$ in a sequential way, then the mean fatigue strength is determined by averaging the test stresses, S_2 to S_{n+1}, beyond the first, without regard to whether each event was a failure or a non-failure:

$$\hat{\mu}_y = \frac{\sum_{i=2}^{n+1} S_i}{n} \qquad (10)$$

The test at S_{n-1} is not carried out, but the stress level itself is determined from the result of nth test.

Estimate the lower limit of the fatigue strength for the population from Formula (9). Take as the number of degrees of freedom that corresponding to the standard deviation used for the test or, if this number is unknown, take it as $n - 1$.

In the modified staircase method, it is necessary to know the standard deviation of the fatigue strength. It may be estimated from the S-N curve as described in Clause 8.

A worked example is given in A.2.

8 Statistical estimation of the S-N curve

8.1 Introduction

Analysis of S-N fatigue is performed for the purpose of fitting an appropriate mathematical relationship to test data to generate a curve which yields approximately 50 % probability of failure. Typically, the data exist at a number of stress or strains and represent a continuous single distribution that is log-normally distributed with constant variance as a function of stress or strain.

The basic relationships employed to describe behaviour are used to reflect either linear or curvilinear response. Figures 3 and 4 demonstrate the behaviour in question. Figure 5 presents a case which occurs only on occasion.

This more complicated behaviour can be managed by use of the Bastenaire equation. This relationship is useful when the data demonstrates an asymptotic flattening of the curve in the very high life regime while simultaneously displaying a convex downward shape in the high stress or strain region.

Mathematically, the Bastenaire relationship has the following form:

$$N = \frac{A}{S-E} \exp\left[-\left(\frac{S-E}{B} \right)^C \right] \qquad (11)$$

where

N	is the fatigue life;
S	is the stress or strain;
A, B, C, E	are curve fit parameters.

The Stromeyer relationship, useful when there is no high stress downward concavity is:

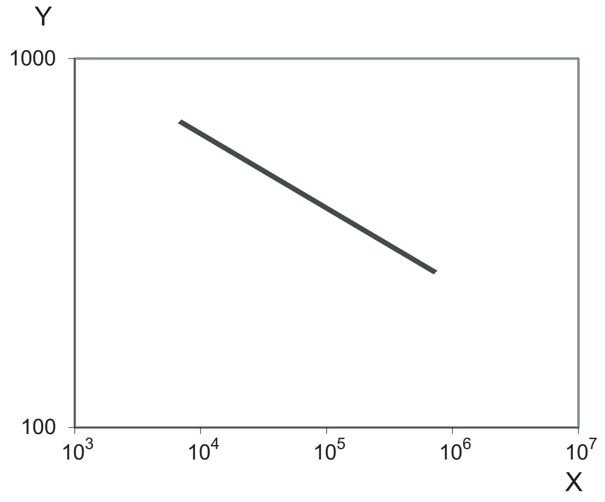
$$\log_{10}(N) = A + B \log_{10}(S - E) \qquad (12)$$

where

N	is the fatigue life;
S	is the stress or strain;
E	is stress at very long life and must be less than all the stress or strain values in the data;
A, B	are curve fit parameters.

Application of the Bastenaire equation or its simplification, Formula (12), cannot be performed using the method of linear least squares. More advanced concepts are required and are beyond the scope of the present document. They can be found in References [10] and [11]. Additionally, a future standard is planned to address a method suitable for this analysis as well as other advanced topics that remain to be defined.

Note that high stress behaviour can result from performing stress-controlled tests at maximum stress levels exceeding the yield strength. In general, this is an improper test technique because cyclic ratcheting may result. Conversely, this high stress behaviour has been observed in strain-controlled tests as well. Strain-controlled testing is sometimes purposely conducted at strain levels producing stresses exceeding the yield strength. These are usually valid in the absence of specific testing issues, etc., invalidating the results.

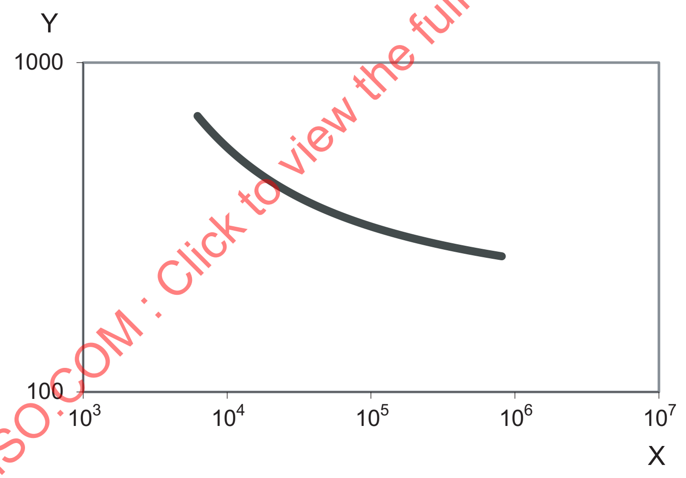


Key

X cycles to failure

Y stress or strain, stress units (MPa) presented

Figure 3 —Typical linear fatigue response

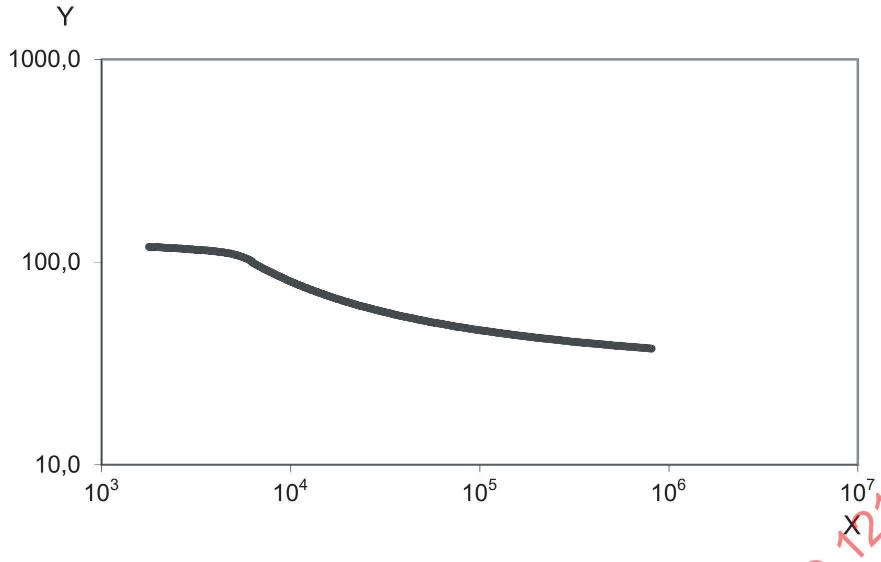


Key

X cycles to failure

Y stress or strain, stress units (MPa) presented

Figure 4 —Typical curvilinear fatigue response



Key

X cycles to failure

Y stress or strain, stress units (MPa) presented

Figure 5 — S-N response occasionally observed

The mathematical models appropriate for the majority of the cases the fatigue practitioner will encounter are given below. However, there will be cases, noted above, that occur that our outside the domain of the methodologies presented and require more sophisticated approaches. However, in general these cases occur rather infrequently and the majority of S-N curves can be evaluated using the techniques presented for linear or curvilinear response (Figures 3 and 4, respectively). The statistical techniques presented below are based on the method of linear least squares [12],[13]. The more advanced methods require either nonlinear regression methods or maximum likelihood estimation (MLE).

These relationships are:

— Linear fatigue response model

$$\mathrm{Log}_{10}(N) = b_0 + b_1 \log_{10}(S) \tag{13}$$

where b_0, b_1 , ...b_n are linear regression coefficients and S can be either stress or strain.

— Curvilinear fatigue response

$$\mathrm{Log}_{10}(N) = b_0 + b_1 \log_{10}(S) + b_2 \log_{10}^2(S) \tag{14}$$

8.2 Estimation of regression parameters

8.2.1 Estimation of the parameters for the linear model[1)]

For all the data, take the logarithms of the stress (or strain) values and the corresponding observed lives. Base 10 logarithms are encouraged. Logarithms to another base, for example base e, are acceptable. However, base 10 is recommended, as most practitioners seem to use these values.

Specifically, the model then has the form:

$$\hat{Y}_i = b_0 + b_1 X_i \tag{15}$$

where $\hat{Y}_i$ is the predicted value of the dependent variable, $X_i = \log_{10}(S_i)$ and $Y_i = \log_{10}(N_i)$.

$$b_1 = \frac{\sum_{i=1}^{n} X_i Y_i - \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n}}{\sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}} \tag{16}$$

$$b_0 = \frac{\left(\sum_{i=1}^{n} Y_i - b_1 \sum_{i=1}^{n} X_i\right)}{n} \tag{17}$$

Standard deviation calculated from the results of regression analysis is defined as:

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2}{n - p}} \tag{18}$$

where p is the number of parameters estimated in the model, in this case p = 2.

Correlation coefficient: A useful parameter to assist in the evaluating the quality of the fit is the correlation coefficient, R^2. This parameter presents the proportion of the variation of the data explained by the model to the total variation. The following relationship is normally used for the linear case. It is generalized later in the discussion of the quadratic model.

$$R^2 = \frac{\left(\sum_{i=1}^{n} X_i Y_i - \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n}\right)^2}{\left[\left(\sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}\right)\left(\sum_{i=1}^{n} Y_i^2 - \frac{\left(\sum_{i=1}^{n} Y_i\right)^2}{n}\right)\right]} \tag{19}$$

Usually, a value of 0,9 or better is indicative of a good fit.

1) Classically, linear regression refers to any linear combination of explanatory variables. For the purposes of this International Standard, however, a linear model simply means a relationship of the form $y = mx + b$.

8.2.2 Estimation of the regression parameters for the quadratic model

Linear regression models that is, all models in which the parameters are linear or can be linearized by a suitable transformation (for example logarithmic) can be calculated using the methods of linear algebra. References [12] and [13] provide the details.

Basically, the solution to the general linear problem is given by:

$$\boldsymbol{b} = (\boldsymbol{X'X})^{-1} \boldsymbol{X'Y} \tag{20}$$

where

- $\boldsymbol{b}$ is the matrix of the calculated regression parameters;
- $\boldsymbol{X'}$ is the transpose of the matrix of x values; the independent variables;
- $\boldsymbol{X}$ is the matrix of X values; the independent variables;
- $\boldsymbol{Y}$ is the matrix of Y values; the dependent variable.

Additional relationships of interest are:

$$R_{\text{SST}} = \sum_{i=1}^{n} \left(Y_i - \bar{Y}_i \right)^2 \tag{21}$$

$$R_{\text{SSR}} = \sum_{i=1}^{n} \left(\hat{Y}_i - \bar{Y}_i \right)^2 \tag{22}$$

$$R_{\text{SSE}} = \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i \right)^2 \tag{23}$$

where

- R_{SST} is the sum of squares total;
- R_{SSR} is the sum of squares regression;
- R_{SSE} is the sum of squares error;
- $\bar{Y}$ is the average of all the $\boldsymbol{Y'}$ in the model.

The standard deviation for the regression model, given in Formula (18) is alternatively expressed as:

$$\hat{\sigma} = \sqrt{\frac{R_{\text{SSE}}}{n - p}} \tag{24}$$

where

- $\hat{\sigma}$ is the estimated standard deviation;
- p is the number of parameters (b's) estimated in the model

  p = 2 for a model of the form $y = b_0 + b_1x$ and

  p = 3 for a quadratic model, $y = b_0 + b_1x + b_1x^2$;
- $n - p$ is the degrees of freedom.

The correlation coefficient presented in Formula (15) is more generally given by:

$$R^2 = \frac{R_{\mathrm{SSR}}}{R_{\mathrm{SST}}} \tag{25}$$

8.3 Analysis approach

In general, the simplest model that captures behaviour should be used for the analysis. Usually, a more complex relationship will give a better fit and certainly will increase the R^2 value, but the final model should only be more complex than the linear relationship if it can be shown to significantly reduce the scatter. Analysis of the data using this strategy is encouraged and the following details will help in determining the significance of increasing the complexity.

8.3.1 Plot the curve on the S-N diagram

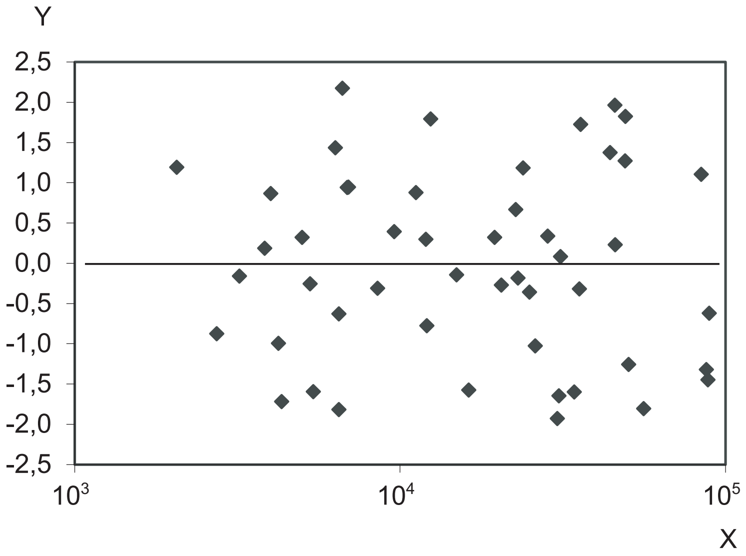
The first analysis, the one assuming linear response in S-N coordinates, should be plotted with the data used to generate the curve to obtain an assessment of the overall fit.

8.3.2 Residuals plots

Evaluation of the quality of fit is undertaken by evaluating plots of residuals and plots of residual versus cumulative normal probability. A residual is defined as:

$$e_i = \left(Y_i - \hat{Y}_i\right) \tag{26}$$

One property of the residuals is that they should sum to zero. A plot of the residual versus the corresponding predicted $(\hat{Y}_i)$ $\left[\text{which is, } \log_{10}(\text{Life})\right]$ values should more or less uniformly populate the plot. An example of an acceptable plot and plots with issues that require resolution are given below. Ideally, the residuals should more or less uniformly populate plot curve without biases or trends. As always, the sum of the residuals should be zero, so the residuals should be centred about the zero value on the y-axis. Note that the residuals on the y-axis have been scaled by the standard deviation. These are referred to as the standardized residuals. Figure 6, as noted, demonstrates the case when a model is adequately capturing response. Figure 7 is the classic case where data was evaluated using a linear response model, but requires a quadratic expression. When the quadratic model is applied, the residuals plot will then appear similar to that shown in Figure 6.

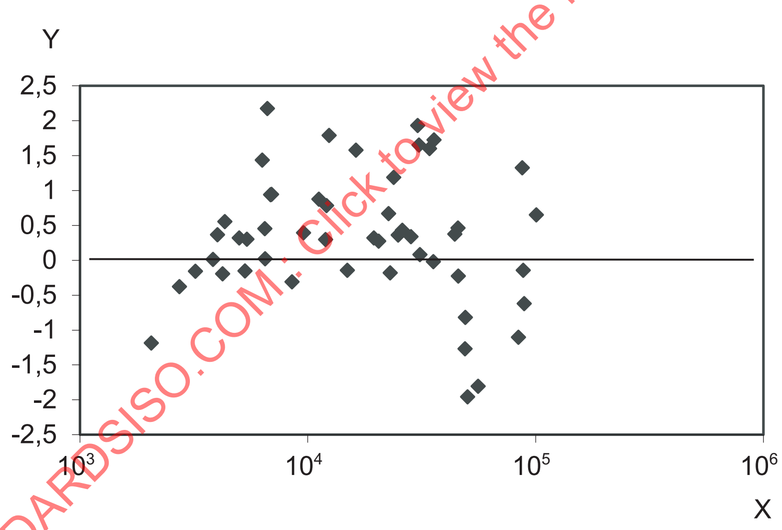


Key

X predicted life

Y standardized residuals

Figure 6 — Residuals plot demonstrating results for a model adequately capturing behaviour

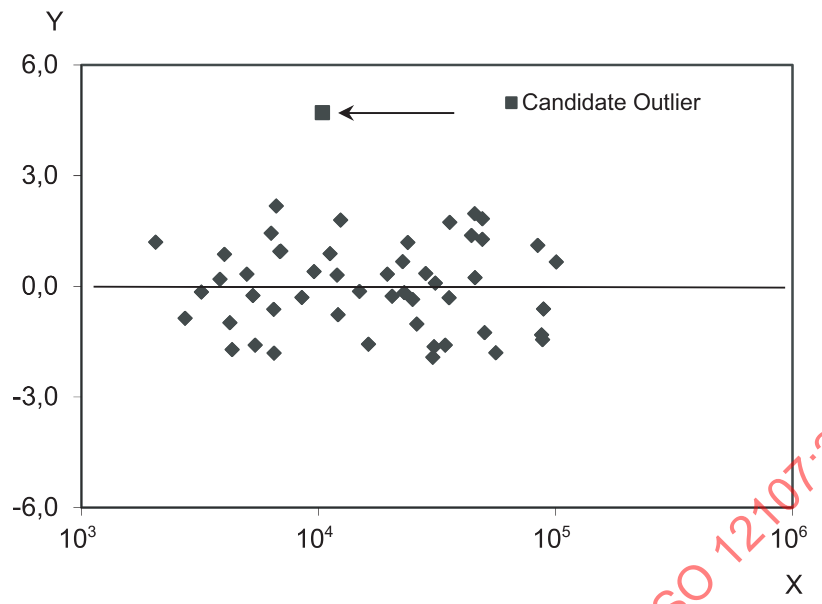


Key

X predicted life

Y standardized residuals

Figure 7 — Residuals plot demonstrating classic behaviour when a linear response model is applied to results requiring a quadratic relationship



Key

X predicted life

Y standardized residuals

Figure 8 — An acceptable residuals plot for the fit, but also demonstrating a candidate outlying result

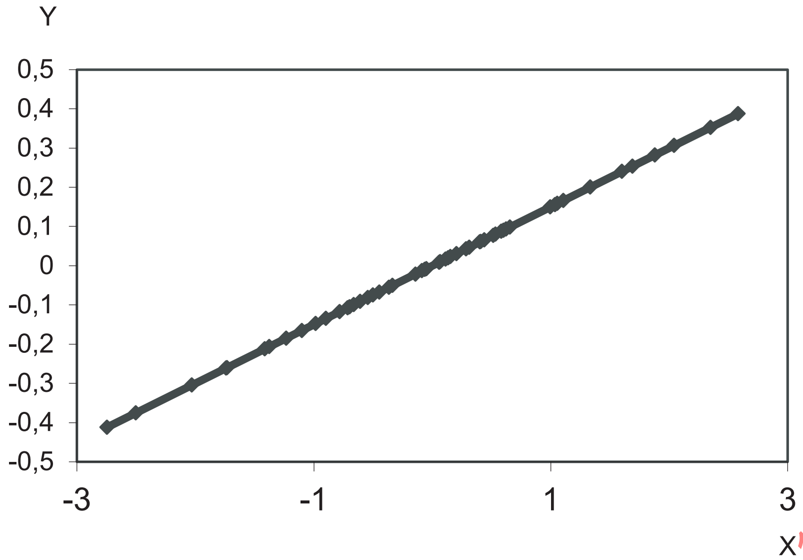
Finally, when a possible outlier(s) is observed, it is appropriate to conduct a careful review of the test records that accompany the results to see if there were any machining and/or testing discrepancies. In addition, it is recommended that a metallurgical and fractographic evaluation of results be made in order to determine if the specimen is a valid observation or if it may have been damaged. Note that the disposition of outlying data can be problematic and subjective. Quantitative statistical tests to evaluate outlying data are available, but these can often generate conflicting results. This is particularly true for cases where the points(s) are suspiciously aberrant, but no so far deviant as to be clearly erroneous.

Unduly high (long-life) results, as well as short-life values, can be problematic; both can inappropriately inflate the scatter and shift the curves. The likely result is higher than necessary estimates of the standard deviation. This can lead to lower tolerance limits than appropriate.

In cases with outliers, the correct perspective is to be judicious, neither automatically rejecting tests that might be problematic, nor retaining such results because no physical reason can be identified. The guidance is to err conservatively in such cases by retaining observations where there is insufficient cause to reject them.

8.3.3 Normal probability plot

Plots of the residuals, or standardized residuals, versus cumulative normal probability are also useful in evaluation of the fit of the model to the data. The residuals are assumed to be normally distributed and this assumption can be evaluated by determining if the residual from the analysis plot reasonably as a straight line.

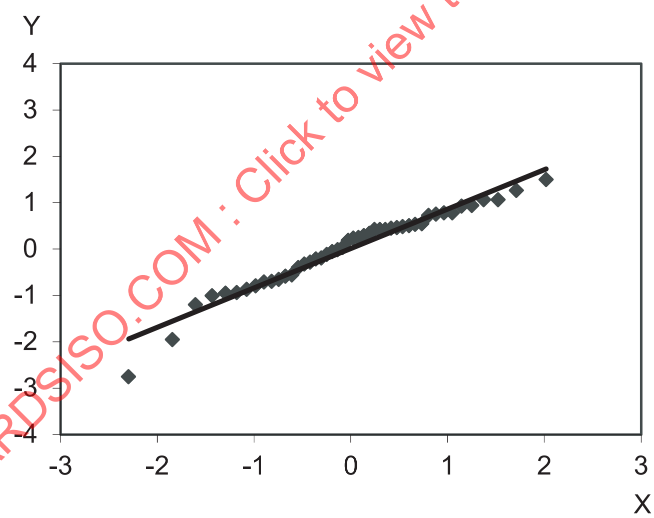


Key

X cumulative normal probability

Y standardized residuals

Figure 9 — Example of a cumulative normal probability plot displaying excellent conformance to normality

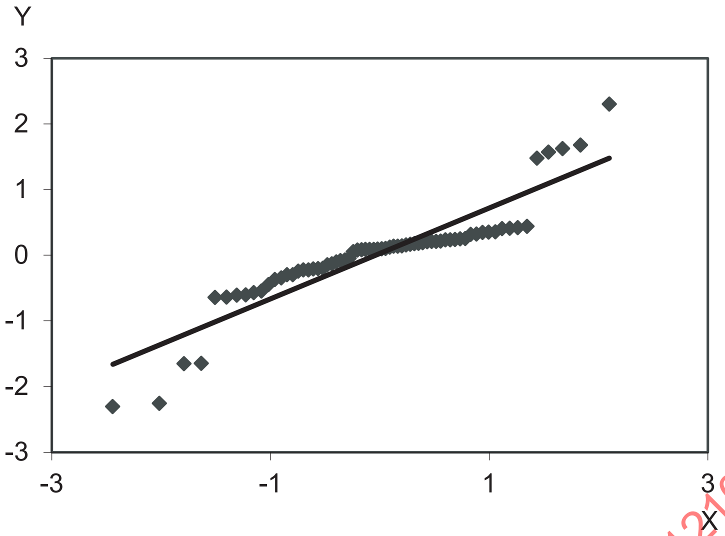


Key

X cumulative normal probability

Y standardized residuals

Figure 10 — Example of a cumulative normal probability plot displaying acceptable conformance to normality



Key

X cumulative normal probability

Y standardized residuals

Figure 11 — Example of a cumulative normal probability plot displaying marginally acceptable conformance to normality

8.3.4 Quantitative measures of normality for the residuals

Should the analyst wish to quantitatively evaluate the assumption of normality, the procedures of Reference [9] (as noted in 6.4) can be applied. For regression analysis, however, it is the residuals that are evaluated as opposed to the actual observations.

8.3.5 Remedial measures

Should the results for the tools evaluating the quality of the fit indicate there might be some issues, the following measures are recommended:

— check for invalid data;

— determine if a more complicated model (quadratic) is required;

— examine the S-N curve and the diagnostic plots to see if there are two or more populations within the data requiring segregation into separate data sets. Unique curves for each population may be required.

8.3.6 Determination of the need for a quadratic model

Should the S-N plot suggest a lack of fit and the residuals plot have the appearance of Figure 7, then application of the so-called "general linear test" [12] can quantitatively evaluate the significance of using this relationship. To perform this test, first analyse the results with both the linear and quadratic models. The general linear test uses the sum of squares error, R_{SSE}, and the corresponding degrees of freedom for each of the fitted linear and quadratic expressions.

Specifically, this is:

$$F^* = \frac{\left(R_{\text{SSE 1}} - R_{\text{SSE 2}}\right)}{R_{\text{SSE 1}}} \div \frac{v_2}{\left(v_1 - v_2\right)} \tag{27}$$

where

$R_{\text{SSE 1}} = R_{\text{SSE 2}}$	for the simpler model;
$R_{\text{SSE 2}} = R_{\text{SSE}}$	for the candidate (quadratic) model;
v_1	is the degrees of freedom for the simpler model;
v_1	is the degrees of freedom for the candidate model.

If $F^* > F_{\alpha,\, v_1, v_2}$ conclude the quadratic model is significantly improving the fit.

F is the value obtained from the F distribution table (or integration of the F distribution) at the p, number of parameters in the candidate model, the v_2 degrees of freedom and the choice of the α level.

If the results of this test deem that the quadratic relationship is significant, the curve generated from this model should be plotted against the data to verify the relationship yields reasonable S-N character. On occasion, it has been observed that significant quadratic relationships demonstrate inconsistent behaviour with accepted fatigue response. For example, properties can curve downward at low stresses/strains or occasionally curve over at the highest levels of these stimuli. That is, the curve can suggest life improvements as stress or strain increases. Please note, this is within the range of the data and is not referring to any extrapolation. The reason that a significant quadratic model can demonstrate uncharacteristic fatigue behaviour is attributable to several possibilities.

One or more invalid observations exist in the data. Some can become particularly problematic if they are at the lowest or highest levels of stress or strain.

1) The data were not developed according to the recommended procedure presented in 8.4 and biases resulted.

2) More than one failure distribution is contained within the data.

3) The remedial action for the first issue involves suitable identification of invalid results. Then once the outlying data have been removed, the analysis process should be repeated.

For the second issue, either additional data must be appropriately generated or one must accept the linear model and recognize it is not performing an optimal analysis.

For the final issue, the data should be segregated into homogenous populations. If this is a not an option, then other techniques, beyond the scope of this International Standard, are necessary.

Given the S-N response is with the quadratic relationship is acceptable, then the diagnostic residuals and normal probability plots should be developed to confirm the model is adequately analysing the data.

8.4 Calculation of the lower tolerance limit

Estimation of the lower limit of the S-N curve life at a given probability of failure, assuming a normal distribution, at the confidence level, 1 - α follows that given in 6.5 is given by:

$$\hat{Y}_{\text{TL}} = \hat{Y} - k_{(P,\,1-\alpha,\nu)}\hat{\sigma}\left[1 + \boldsymbol{X}'_{\text{H}}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}_{\text{H}}\right]^{\frac{1}{2}} \tag{28}$$

where

$\hat{Y}_{\text{TL}}$	is the lower tolerance estimate;
$k_{(P,\,1-\alpha,\nu)}$	is the factor for a one-sided tolerance limit at a probability of P, a confidence of $1-\alpha$ and ν degrees of freedom;
$\boldsymbol{X}'_{\text{H}}$	is the inverse of the matrix of specific values for which a tolerance estimate is desired; this is also known as the inverse of the “hat matrix”;
$\boldsymbol{X}_{\text{H}}$	is the matrix of specific x values for which a tolerance estimate is desired; this is the “hat matrix”.

For the linear case, this simplifies to

$$\hat{Y}_{\text{TL}} = \hat{Y} - k_{(P,1-\alpha,\,\nu)}\hat{\sigma}\left[1 + \frac{1}{n} + \frac{\left(X - \bar{X}\right)^2}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}\right]^{\frac{1}{2}} \tag{29}$$

The coefficient k is the one-sided tolerance limit for a normal distribution, as given in Table B.1. These values are generated by integration of the non-central t distribution with non-centrality parameter:

$$\delta = \sqrt{n} \tag{30}$$

The number of degrees of freedom, ν, is the same number used in estimating the standard deviation. Please note this is a simplifying assumption, to make calculation of the lower tolerance limit more straightforward.

The actual non-centrality parameter for a regression model is more accurately given by:

$$\delta = \frac{1}{\left[\boldsymbol{X}'_{H}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}_{H}\right]^{\frac{1}{2}}} \tag{31}$$

However, this requires integration of the non-central t distribution for every specific X value and is quite tedious. It is for this reason, the non-centrality parameter is assumed to be the $\sqrt{n}$. This permits tabular values. These are presented in Annex B for many commonly used values for P, 1-α, and ν.

8.5 Experimental plan for the development of S-N curves

Data generation to support the development of design curves should be performed by first determining the appropriate sample size. For a regression model, a minimum of 10 observations is necessary for exploratory work, but 30 are necessary for a curve intended for design or reliability pursuits. Regardless of the intention, the stress/strain range over, which the data are to be generated, must be identified. Replicate data are not necessary and are in general discouraged. Then specimen test conditions should be allocated in equal increments of stress or strain. Data are generated in this fashion because

— it clearly helps to define if the response is linear or curvilinear, and

— it mitigates against any biases since the data are uniformly distributed.

Fatigue results tend to demonstrate more scatter as the stress or strain diminishes. The uncertainty along the length of the curve increases proportionately. To address this behaviour, it is recommended that specimens be allocated in increments using the double logarithm of strain or stress. This tends to locate more specimens in the higher life regimes.

Worked examples are given in A.3. and A.4.

9 Test report

9.1 Presentation of test results

The test report shall include the following information as appropriate to the type of test.

a) the observed fatigue life at a given stress;

b) the test stress level and the estimated mean fatigue life, plus the estimated standard deviation of the logarithm of the fatigue life. The number of test specimens shall be indicated;

c) a compilation of the experimental fatigue life data obtained for each specimen, the observations such as the mode of failure or non-failure, and indicting the test stress or strain. The R ratio (ratio of minimum to maximum stress or strain), the type of test and the test frequency shall also be reported. If the test was performed in strain control, indicate the method of extensometry and if the test was performed at constant frequency or strain rate. Report the strain rate if test were performed using a constant stain rate. Conversely, if the test frequency was held constant, this value should be reported. Finally, report if the test specimen was smoothed or notched. If notched specimens were used, the corresponding stress concentration factor (K_t) should be reported;

d) a plot of the experimental data on probability coordinates with the line best fitting the results shall be reported. No extrapolation beyond the range of the data shall be presented;

e) the estimated tolerance limit at the selected confidence and reliability level should be reported.

9.2 Fatigue strength at a given life

The test report shall include the following information as appropriate to the type of test:

a) the estimated mean fatigue strength and the estimated standard. Include the number of specimens tested. Report the method used to estimate these parameters, such as the staircase method;

b) a list of the experimental at each stress level and the number of cycles to which each specimen was subjected, with observations on failure or non-failure in the order of the test;

c) The estimated lower limit of the fatigue strength at the selected probability, when necessary. No extrapolation of the probability curve is permissible beyond the limits of the data.

9.3 *S*-*N* curve

The test report shall include the following information as appropriate to the type of test:

a) the estimated mean *S*-*N* curve, showing plots of the experimental data. No extrapolation beyond the limits of the data are permissible;

b) a list of experimental data including the stress or strain level and the number of cycles applied to each specimen. Each should be identified as a failure or non-failure, as appropriate;

c) the estimated lower limit of the of the *S*-*N* curve at the selected probability of failure;

d) plots of residuals and normal probability plots shall be included. Adequate justification for the choice of model, including the standard deviations for the linear or quadratic models, R^2 values, and the results of the general linear test calculation shall be presented. The reviewer should be able to clearly understand the process for final model selection.

Annex A
(informative)

Examples of applications

A.1 Example of statistical estimation of fatigue life

A set of seven data items is given in Table A.1 as an example. Calculate the average and standard deviation using Formulae (3) and (4).

$$\hat{\mu} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{A.1}$$

where

$\hat{\mu}$ is the sample mean;

X_i is the ith observed sample value;

n is the number of data points.

In this case, x_i is the logarithm, base 10, of each observation.

Hence,

$$\hat{\mu} = \frac{\left(\log_{10}\left(6,05 \times 10^4\right) + \log_{10}\left(6,31 \times 10^4\right) + \ldots\right)}{7} = 4,905$$

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \hat{\mu}\right)^2}{n-1}} \tag{A.2}$$

Hence,

$$\sqrt{\frac{\left(\log_{10}\left(6,05 \times 10^4\right) - 4,905\right)^2 + \left(\log_{10}\left(6,31 \times 10^4\right) - 4,905\right)^2 + \ldots}{7-1}} = 0,121$$

The lower limit of the fatigue life for a 10 % probability of failure, at a confidence level of 95 %, is estimated from Formula (5), taking $k(0,1;\ 0,95;\ 6)$ as 2,755 as given in Table B.1:

$$\begin{aligned}\hat{x}_{(10)} &= 4,905 - (2,755 \times 0,121)\\ &= 4,572\end{aligned}$$

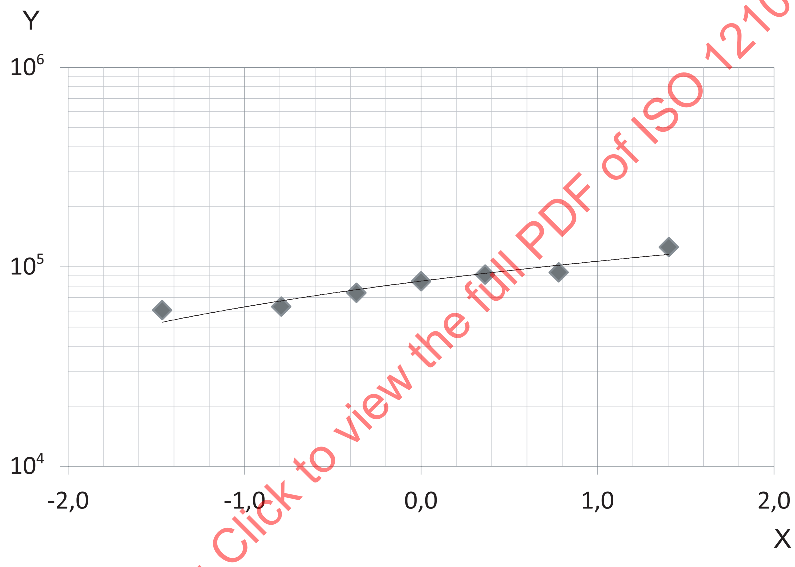
or

$$\begin{aligned}\hat{N}_{(10)} &= 10^{4,572}\\ &= 3,73 \times 10^4 \text{ cycles}\end{aligned}$$

Finally, the median life is $10^{4,905} = 80\ 352$ cycles

Table A.1 — Example of fatigue life data

Specimen number i	Fatigue life N_i, cycles	Log of fatigue life $x_i = \log N_i$	Rank $(i-0{,}5)/N$	Number of standard deviations z
1	$6{,}05 \times 10^4$	4,782	0,071 4	-1,465 50
2	$6{,}31 \times 10^4$	4,800	0,214 3	-0,791 43
3	$7{,}39 \times 10^4$	4,869	0,357 1	-0,365 66
4	$8{,}46 \times 10^4$	4,927	0,500 0	0,000 00
5	$9{,}11 \times 10^4$	4,960	0,642 9	0,363 78
6	$9{,}37 \times 10^4$	4,972	0,785 7	0,780 53
7	$1{,}25 \times 10^5$	5,098	0,928 6	1,404 00



Key

X cumulative normal probability, Z

Y life cycles

Figure A.1 — Example of a cumulative normal probability plot for the fatigue life data given in Table A.1

A.2 Examples of statistical estimation of fatigue strength

A.2.1 Staircase method

When using the staircase method, specimens are tested sequentially under increasing stresses until a failure occurs. An example of a set of data are given in Table A.2. From the beginning, the last non-failure in terms of stress is the first valid data which is 500 MPa in Table A.2. In this test, there are seven failures and eight non-failures. The failure event is therefore the one considered in the analysis.

Only three stress levels are considered in the analysis, as shown in Table A.3, with $S_0 = 500$ MPa and stress step $d = 20$ MPa. The number of the relevant event, f_i, is given in the third column of the table. The values of A, B, C and D are as follows:

$A = 7$; $B = 11$; $C = 7$; $D = 0{,}571$

The mean and the standard deviation of the fatigue strength are calculated from Formulae (9) and (10), as follows:

$$\hat{\mu}_y = 500 + 20\left(7/7 - 1/2\right) = 510 \text{ MPa}$$

$$\hat{\sigma}_y = 1{,}62 \times 20\left(0{,}571 + 0{,}029\right) = 19{,}4 \text{ MPa}$$

Table A.2 — Example of staircase test data

Stress	Sequence number of specimen																
S_i MPa			1				5					10					15
540									X								X
520				X				O		X		X				O	
500			O		X		O				O		X		O		
480		O*				O								O			
460	O*																

X for failure

O for non-failure

* not counted (see the discussion in the first paragraph of A.2).

Table A.3 — Analysis of the data in Table A.2

Stress S_i MPa	Level I	Values f_i	 if_i	 $i^2 f_i$
540	2	2	4	8
520	1	3	3	3
500	0	2	0	0
Sum	—	7	7	11

The lower limit of the fatigue strength for a probability of failure of 10 % is calculated from Formula (11) at a confidence level of 95 %. The value of the appropriate coefficient, $k(0{,}1;\ 0{,}95;\ 6)$, taken from Table B.1, is 2,755.

$$\hat{y}_{(10)} = 510 - \left(2{,}755 \times 19{,}4\right)$$
$$= 456 \text{ MPa}$$

In this example, the stress step d is close enough to the estimated standard deviation and D is greater than 0,3.

A.2.2 Modified staircase method

This example is based on the same fatigue test data as in A.2.1, but only for sequence numbers 1 to 6. The set of data used is given in Table A.4. The standard deviation of the fatigue strength is 19,4 MPa with a number of degrees of freedom of 6, as calculated above. The test was conducted with a stress step of 20 MPa which is close enough to the standard deviation.

The mean fatigue strength is calculated from the data, using Formula (12), as follows:

$$\hat{\mu}_y = (520 + 500 + 480 + 500 + 520 + 540)/6$$
$$= 510 \text{ MPa}$$

The lower limit of the fatigue strength for a probability of failure of 10 % is calculated from Formula (11), at a confidence level of 95 % and taking a value for $k(0,1; 0,95; 6)$ of 2,755 from Table B.1, as follows:

$$\hat{y}_{(10)} = 510 - (2{,}755 \times 19{,}4)$$
$$= 456 \text{ MPa}$$

Table A.4 — Example of modified staircase test data

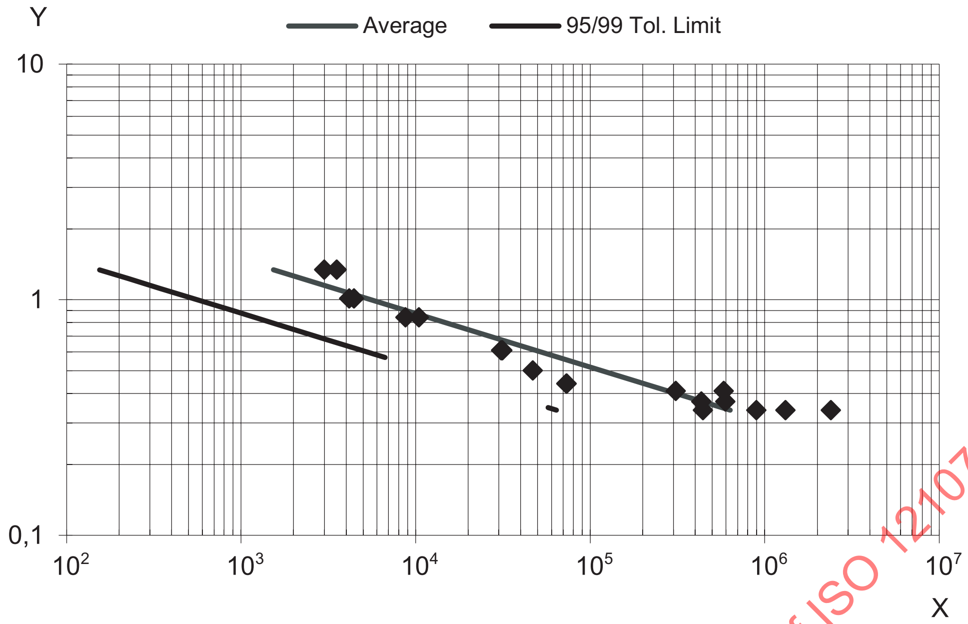
Parameter	Test sequence						
i	1	2	3	4	5	6	7
S_i, MPa	500	520	500	480	500	520	540
Event	O	X	X	O	O	O	[a]

[a] Test not actually carried out (stress level calculated from previous value).

A.3 Statistical estimation of S-N curve

Table A.5 — Strain-controlled low cycle fatigue results

Specimen number	Strain range %	Cycles to failure
1	1,34	3 534
2	1,34	3 002
3	1,01	4 174
4	1,01	4 442
5	0,84	10 477
6	0,84	8 758
7	0,61	31 476
8	0,61	30 990
9	0,50	47 000
10	0,44	73 387
11	0,44	73 280
12	0,41	309 600
13	0,41	583 300
14	0,37	595 800
15	0,37	433 620
16	0,34	2 400 800
17	0,34	1 315 800
18	0,34	895 280
19	0,34	443 930

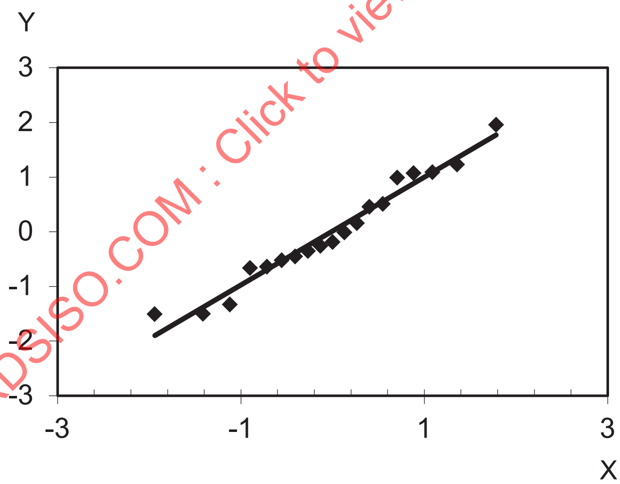


Key

X cycles to failure

Y strain range, %

Figure A.2 — Results for the linear model

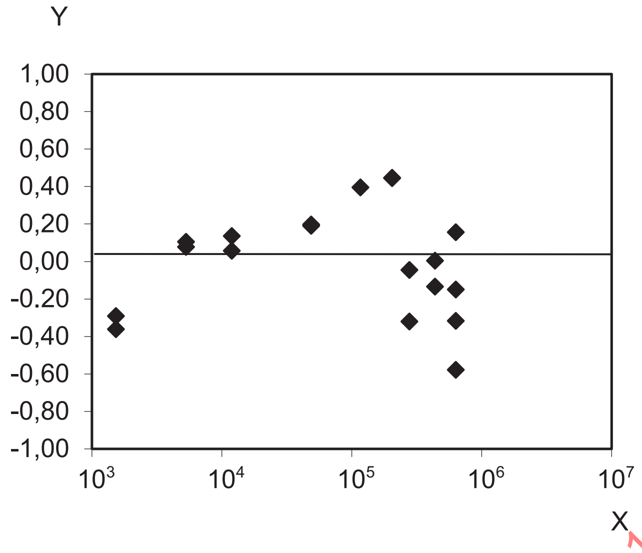


Key

X cumulative normal probability, Z

Y standardized residuals

Figure A.3 — Cumulative normal probability plot for the linear model



Key

X predicted life

Y standardized residuals

Figure A.4 — Residuals plot for the linear model

