Thesis for the Degree of Licentiate of Philosophy

# Estimating fatigue limits and inclusion sizes using staircase tests and finite lives

Sara Lorén

CHALMERS | GÖTEBORG UNIVERSITY

Department of Mathematical Statistics
Chalmers University of Technology and Göteborg University
SE - 412 96 Göteborg, Sweden

Göteborg, August 2001

Estimating fatigue limits and inclusion sizes using staircase tests and finite lives
Sara Lorén

© Sara Lorén, 2001

ISSN 0347-2809/NO 2001:50
Department of Mathematical Statistics
Chalmers University of Technology and Göteborg University
SE - 412 96 Göteborg
Sweden
Telephone +46 (0)31-772 1000

Chalmers University of Technology
Göteborg, August 2001

**Abstract**

An important design property of steel is the fatigue limit, the load level where the specimens have infinite live. For the material it can also be define endurance limit. This is the stress level at which the specimen has a certain life it will not break before $n_o$ cycles. Hence it is of interest to estimate these limits. The maximum inclusion size in clean steels influences fatigue limit behaviour. Therefore it is also interesting to estimate the distribution of the inclusion size which causes the failure. In the two papers included in the present thesis the results from staircase tests are used to estimate the fatigue limit, endurance limit and the inclusion size distribution.

Two models are compared in the first paper, one that estimates the fatigue limit distribution, and one that estimates the endurance limit distribution. The model that estimates the endurance limit uses the load level for the broken and unbroken specimens. To estimate the fatigue limit distribution the finite life for the broken specimens also has to be used. It is shown that it is possible to estimate the distribution of the fatigue limit although the fatigue limit itself is impossible to observe. When the two models give very different estimates this is an indication that the chosen run-out level is too low.

Mainly two different ways to estimate the inclusion size distribution are described in the second paper. The first way uses the load levels for both the broken and unbroken specimens. The second way uses the load levels for the unbroken specimens, and for the broken ones the size of the inclusion which causes the failures is used. It is shown that from a staircase test with measured inclusion size on the broken specimens it is better to use both load levels and inclusion sizes to estimate the inclusion size distribution instead of using just the different load levels. The models using the inclusions sizes are also used to verify or reject possible fatigue limit models.

**Keywords:** Endurance limit, Censoring, Maximum likelihood, Staircase test

**Acknowledgements**

First of all, I would like to thank my supervisor Professor Jacques de Maré for his ideas and comments. The support from ITM, Sandvik Steel AB and ABB Stal AB is gratefully acknowledged. I would also thank Dr Thomas Svensson (Swedish National testing and Research Institute), Dr Boel Wadman (IVF) and Dr Pär Johannesson for their help throughout this work.

Sara Lorén
Göteborg, August 2001

# 1 Introduction

The fatigue limit is often an important characteristic in fatigue design for metallic material. When the specimen is exposed to a load level below the fatigue limit it will never break. It will have an infinite life. Because one can never test a specimen for infinite time one often uses a load level which one calls endurance limit and below this limit the specimens will never break before some specified finite number of cycles. In the literature some use endurance limit where other use fatigue limit. Here the fatigue limit corresponds to infinite life and endurance limit to some specified finite life.

Fatigue properties of hard steels are strongly influenced by inclusions in the material. Inclusions can be microscopic particles of foreign material, for instance oxides. The endurance limit can also be determined by the inclusions in the specimens. According to [1] the endurance limit is a function of the largest inclusion and the hardness of the material, the $\sqrt{area_{max}}$-model. Therefore it is also interesting to estimate the inclusion size distribution because then one can estimate the endurance limit distribution.

One test to estimate the fatigue limit distribution is the staircase test (or the up-and-down method) [2]. A start load level is chosen. If the specimens fail prior to infinite life (say $n_o$ cycles) the next specimen has to be tested at a lower load level. If the specimen does not fail within $n_o$ cycles, the next specimen is tested at a higher load level. Therefore, each test is dependent on the previous test results. The usual result from a staircase test is the number of failure and survivors at each load level. For the broken specimens one also gets the number of cycles to failure. Under the assumption that there is an inclusion that caused the failure, the size of this inclusion can also be measured for the broken specimens. In the present works the test results from a staircase test are analysed in different ways to estimate the endurance limit, fatigue limit and the inclusion size distribution.

One way to visualise the result of the staircase test is to plot a S-N curve (or Wöhler curve) where S is the load level and N the number of cycles to failure or $n_o$ for the unbroken specimens. For some material the S-N curve has an asymptote at some load level. When this asymptote exists this level is the theoretical fatigue limit.

When estimating the fatigue limit it is tempting to use not only the fact that a specimen has broken but also the actual finite life. This has been done in [1], [2] and [3].

When estimating the fatigue limit using finite lives it is important to think of the connection between the fatigue limit and the finite life. In [3] the finite life depends on the difference of the load and the fatigue limit. In Paper A it is assumed that the load itself determines the finite life through crack propagation. On the other hand if the life is infinite this is because the load is below the fatigue limit. Hence it is assumed that there are two different mechanisms, one which determines whether crack propagation occurs and one which determines the finite life given that crack propagation occurs. In this case the two mechanisms are uncoupled in contrast to the case considered in [3]. The rational for the case with uncoupled mechanisms is inspired by the work in [5]. They claim that finite life occurs for a given load if there is a large enough inclusion. The actual life is not determined entirely by the inclusion size but also by the shape and chemical content of the inclusion, and the most important quantity determining the infinite life is the applied load according to the S-N curve. The endurance limit distribution is estimated with the load levels from the staircase test. The fatigue limit is estimated from the load level on the unbroken specimens and the life of failed specimens.

Direct observations of inclusions sizes are difficult since the inclusions are rare and often occur in the interior of the specimen. A usual way of observing is the following, a number of areas of the same size are observed on polished cross-sections by optical microscopy. The largest inclusion in each area is identified and measured and then the extreme value theory is used to extrapolate the maximum inclusion size distribution of the specimen of interest. A way of doing direct observations of the inclusion in the specimen is to assume that for a staircase test the specimen breaks in the largest inclusion and measure the inclusion size where the specimens have broken. Therefore the staircase test can also be used to estimate the inclusion size distribution.

The inclusion size distribution is estimated in Paper B. In [4] and [5] the distribution of the inclusions was estimated using an estimate of the endurance limit and the relation between the endurance limit and the maximum inclusion size [1]. The maximum inclusion size can of course also be estimated from the size of the inclusions but, as mentioned earlier, it is not easy to measure the inclusion size. The inclusion size distribution in Paper B is estimated by combining the results from a staircase test under different assumptions. Mainly two different ways are used. The first way uses the load level for the broken and unbroken specimens and the relation between the endurance limit and the inclusion size with different assumptions on the inclusion size distribution. The second way also uses the load level on the unbroken specimens but uses the inclusion size on the inclusion that has caused the failure for the broken specimens. This is done under different assumption on the inclusion size distribution and whether or not the largest inclusion caused the failure.

## 2 Paper A

In this paper a model which estimates the endurance limit distribution and one model which estimates the fatigue limit distribution are compared.

The model which estimates the endurance limit is a variant of a model in [4]. Here the endurance limit means the limit where the specimens have a specified finite life. The logarithm of the endurance limit is assumed to be normally distributed. When the specimen breaks one assumes that one is above the endurance limit and if the specimen survives one assumes that one is below the endurance limit.

The other model estimates the fatigue limit, the limit where the specimen has infinite life. The logarithm of the fatigue limit is assumed to follow a normal distribution. For this model one also uses the information when the specimen has broken, i.e. number of cycles to failure.

It is shown that it is possible to estimate the distribution of the fatigue limit, although the fatigue limit itself is impossible to observe. When the two models give very different estimates this is an indication that the chosen run-out level is too low.

## 3 Paper B

In this paper different approaches of estimating the inclusion size distribution from a staircase test are shown. The main assumption for all the models is that for each load level there is a critical inclusion size. This critical size is such that if the maximum inclusion size in the specimen is smaller than the critical size, the specimen will not break and if the maximum inclusion size is bigger than the critical size the specimen will break.

One model uses the number of survivors and failures at each load level and the relation between the endurance limit and the inclusion sizes distribution. The other two models use the inclusion size for the inclusion which has caused the failure and the load level for the unbroken specimens. One of these models uses the assumption that the specimen breaks in the largest inclusion, and the other one is based on the assumption that the specimens break in some inclusion over a critical size, not necessarily the largest one.

It seems as if the model which uses the inclusion size for the broken specimen gives a better estimate of the inclusion size distribution. The advantage of using the load level and observed inclusion size from a staircase test instead of the inclusion sizes for all specimens is that one does not have to search for the inclusions in the specimens. One only measures the inclusions where the specimen breaks.

## References


[1] Y. Murakami, S. Beretta: Small defects and inhomogeneities in fatigue strength: Experiments, Models and Statistical implications, Extremes, Vol. 2, pp. 123-147 (1999).

[2] R.E. Little, E.H. Jebe: Statistical design of fatigue experiments, Applied Science Publishers, Glasgow, UK, (1975).

[3] F. A. Bastenaire: New method for the statistical evaluation of constant stress amplitude fatigue test result, ASTM STP 511, American Society for Testing and Materials, pp. 3-28 (1972).

[4] S. Nishijima: Statistical Analysis of Small Sample Fatigue Data, Statistical Research on Fatigue and Fracture, pp. 1-19 (1987).

[5] F.G. Pascual, W.Q. Meeker: Estimating Fatigue Curves With the Random Fatigue-Limit Model, Technometrics Vol. 41, No 4. pp. 277-302 (1999).

[6] S. Beretta, Y. Murakami: Statistical analysis of defects for fatigue strength prediction and quality control of materials, Fatigue & Fracture of Engineering Materials & Structures pp. 1049-1065 (1998).

[7] H. Bomas, P. Mayr: Inclusion size distribution and endurance limit of a hard steel, Extremes, Vol. 2, pp. 149-164 (1999).

[8] T. Svensson, J. de Maré: Random features of the fatigue limit, Extremes, Vol. 2, pp. 165-176 (1999).

[9] W.J Dixon, A. M. Mood: A method for obtaining and analyzing sensitivity data, Journal of the American Statistical Association, Vol. 43 (1948).

# Fatigue limit estimated using finite lives

**Abstract**

For metallic material one is interested in estimating the fatigue limit although it is impossible to test if a specimen has infinite life. In this paper two models are compared. One model estimates the fatigue limit and another estimates the endurance limit. The model which estimates the endurance limit is a model which only uses the information that a specimen has broken or not broken at a certain stress level. The model which estimates the fatigue limit uses also the information about when the specimen has broken i.e. after how many cycles the specimen breaks. The two models are compared both on real and simulated staircase tests.

It is shown that it is possible to estimate the distribution of the fatigue limit although the fatigue limit itself is impossible to observe. When the two models give very different estimates it indicates that the chosen run-out level is too low.

KEY WORDS: Endurance limit; Maximum likelihood; Censoring; Information matrix; Staircase test.

## 1 Introduction

The S-N curve describes relations between a given stress amplitude and the number of cycles until a specimen has broken. To get this curve each specimen is exposed to a stress amplitude at some given level, and the number of cycles until the specimen breaks is observed. The test is interrupted if the specimen has not broken before some specified number of cycles $n_o$ (often $2 \cdot 10^6$). Here $n_o$ denotes the endurance life.

For some materials the S-N curve has an asymptote at some stress level. When this asymptote exists the corresponding stress level is called the fatigue limit and this implies that if the specimen is exposed at a lower stress level it will have an infinite life. For the material it can also be define endurance limit. This is the stress level at which the specimen has a certain life it will not break before $n_o$ cycles.

When estimating the fatigue limit it is tempting to use not only the fact that a specimen has broken but also the actual finite life. This has been done in [1], [2] and [3].

Two models, called Model 1 and Model 2, are compared. Model 2, which is a variant of [4], estimates the endurance limit and only uses the information that a specimen has broken or not broken at a certain stress level. Model 1 estimates the fatigue limit. This model also uses information from the finite life.

When estimating fatigue limit using finite lives it is important to think of the connection between the fatigue limit and the finite life. In [3] the finite life depends on the difference of the load and the fatigue limit.

In this paper it is assumed that the load itself determines the finite life through crack propagation. On the other hand if the life is infinite this is because the load is below the fatigue limit. Hence it is assumed that there are two different mechanisms, one which determines whether crack propagation occurs or not and one which determines the finite life given that crack propagation occurs. In this case the two mechanisms are uncoupled in contrast to the case considered in [3]. The rational for the case with uncoupled mechanisms is inspired by the work in [5]. They claim that finite life occurs for a given load if there is a big enough inclusion. The actual life is not determined entirely by the inclusion size but also by the shape and chemical content of the inclusion, and the most important quantity determining the infinite life is the applied load according to the S-N curve.

On the other hand a physical argument for the approached by [3] is to assume that the Fatigue limit act as a threshold and only the part of the load above the threshold drives the crack. This may be justified by the Elber crack closure concept [6].

The question studied in this paper is if the actual finite life carries any information on the fatigue limit although the fatigue limit and the finite life are governed by two different mechanisms. The answer to that question is here that it depends on the result of the staircase test. If the finite life is close to the specified number of cycles where the test is interrupted, it contains information of the actual fatigue limit.

## 2 Model 1

Here the fatigue limit means that if the specimen expose to a lower load level than the fatigue limit it will have infinite life. To estimate the distribution of the fatigue limit the following assumptions are made. The fatigue limit $S_e$ is a random variable whose logarithm is normally distributed with the mean $\mu_{xe}$ and the variance $\sigma_{xe}^2$. The logarithm of the finite life $N$ is also a random variable which is normally distributed with an expectation $\nu(x)$ which is a function of the stress value $S$ through its logarithm $x = ln(S)$. The variance of the finite logarithmic life is not constant. Instead it increases, the closer the stress level is to the fatigue limit and denoted by $\sigma_\nu^2/w^2$, $w$ depends on the stress value.

An interpretation of the model is that there will be no crack propagation at all with a certain probability, $\mathrm{P}(S_e > S)$ on the applied stress $S$, but with the complementary probability $\mathrm{P}(S_e \leq S)$ crack propagation will take place. In the latter case the conditional expected logarithmic life is $\nu(x)$, where as above $x = ln(S)$. The graph of $\nu(x)$ is assumed to be a straight line with a negative slope or more generally a decreasing continuous function of the stress values.

One way of summarizing the situation is to assume that we have a latent variable $Y(x)$ which is either observed or censored. If there are $n$ observations write

$$Y(x_i) = ln(N_i) \sim N(\nu(x_i), \frac{\sigma_\nu^2}{w_i^2}) \qquad i = 1, \dots, n$$
$$X_{e_i} = ln(S_{e_i}) \sim N(\mu_{xe}, \sigma_{xe}^2) \qquad i = 1, \dots, n$$

There are two ways in which $Y(x_i)$ can be censored. The first one is used, if the specimen has survived the test, i.e. the specified number of cycles, $n_o$, but actually has a finite life. The second way the specimen can be censored is when the stress level is below the fatigue limit. In both cases the value of the observation $Z(x_i)$ is put to $y_o$.

$$\begin{aligned} Z(x_i) &= Y(x_i) 1_{\{X_{e_i} < x_i\}} 1_{\{Y(x_i) \leq y_o\}} + y_o(1 - 1_{\{X_{e_i} < x_i\}} 1_{\{Y(x_i) \leq y_o\}}) \\ &= \begin{cases} y_o & \text{if censored} \\ Y(x_i) & \text{otherwise} \end{cases} \end{aligned}$$

where $1_{\{\cdot\}}$ is the indicator function and $y_o$=ln($n_o$)

The probability that $Y(x_i)$ is censored is

$$\mathrm{P}\Big(Y(x_i) \text{ is censored }\Big) = \mathrm{P}\Big(Z(x_i) = y_o\Big) =$$

$$\mathrm{P}\Big(Z(x_i) = y_o | X_e \leq x_i\Big)\mathrm{P}\Big(X_e \leq x_i\Big) + \mathrm{P}\Big(Z(x_i) = y_o | X_e > x_i\Big)\mathrm{P}\Big(X_e > x_i\Big) =$$

$$\Big(1 - \Phi\Big(\frac{y_o - \nu(x_i)}{\sigma_\nu / w_i}\Big)\Big)\Phi\Big(\frac{x_i - \mu_{xe}}{\sigma_{xe}}\Big) + 1 \cdot \Big(1 - \Phi\Big(\frac{x_i - \mu_{xe}}{\sigma_{xe}}\Big)\Big) =$$

$$1 - \Phi\Big(\frac{y_o - \nu(x_i)}{\sigma_\nu / w_i}\Big)\Phi\Big(\frac{x_i - \mu_{xe}}{\sigma_{xe}}\Big)$$

and the likelihood for an uncensored observation is

$$f\Big(y_i; (\sigma_\nu, \nu(x_i))\Big)\Phi\Big(\frac{x_i - \mu_{xe}}{\sigma_{xe}}\Big)$$

where

$$f\Big(y_i; (\sigma_\nu, \nu(x_i))\Big) = \frac{w_i}{\sqrt{2\pi}\sigma_\nu} e^{-(y_i - \nu(x_i))^2 w_i^2 / 2\sigma_\nu^2}$$

and $\Phi(\cdot)$ standard normal cumulative distribution.

## 2.1 Likelihood function

To estimate unknown parameters we maximize the likelihood function for our model. This is done by splitting the observations into two disjoint sets, U corresponding to the uncensored observations and C to the censored ones. The likelihood function is then

$$lik(\boldsymbol{\theta}) = \prod_{i \in U} f\Big(y_i; (\sigma_\nu, \nu(x_i))\Big)\Phi\Big(\frac{x_i - \mu_{xe}}{\sigma_{xe}}\Big) \prod_{i \in C} \mathrm{P}\Big(Z(x_i) = y_o\Big) =$$

$$\prod_{i \in U} f\Big(y_i; (\sigma_\nu, \nu(x_i)\Big)\Phi(\frac{x_i - \mu_{xe}}{\sigma_{xe}}) \prod_{i \in C} \Big(1 - \Phi(\frac{y_o - \nu(x_i)}{\sigma_\nu / w_i})\Phi(\frac{x_i - \mu_{xe}}{\sigma_{xe}})\Big) \quad (1)$$

where

$$\boldsymbol{\theta} = (\sigma_\nu, \nu(\cdot), \sigma_{xe}, \mu_{xe})$$

### 2.2 Variance estimation

The maximum likelihood estimator for the parameter $\boldsymbol{\theta}$ is denote by $\hat{\boldsymbol{\theta}}$. One approximative variance estimates of $\hat{\boldsymbol{\theta}}$ is obtained from the inverse of the observed information matrix $I$

$$I_{jk} \approx -\frac{\partial^2}{\partial\theta'_j \partial\theta'_k} log\Big(lik(\boldsymbol{\theta})\Big)\Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}; \quad j=1,\ldots,5 \quad k=1,\ldots,5.$$

These derivates are calculated by a formula manipulating software (Mathematica).

The approximate variance estimates of $\hat{\mu}_{xe}$ and $\hat{\sigma}_{xe}$ are $\widehat{Var}(\hat{\mu}_{xe}) \approx I^{-1}(5,5)$, $\widehat{Var}(\hat{\sigma}_{xe}) \approx I^{-1}(4,4)$ and $\widehat{Cov}(\hat{\mu}_{xe}, \hat{\sigma}_{xe}) \approx I^{-1}(4,5)$, where $I^{-1}(j,k)$ denotes the $(j,k)$ element of the inverse of the information matrix.

## 3 Model 2

Here the endurance limit means if the specimens is expose to a lower load level than endurance limit it will have a certain life, $n_o$. One way of estimating the endurance limit is to use a model which only uses the information that a specimen has broken or has not broken at a certain stress level. The only assumption which is needed for this model is that the endurance limit $S_{el}$ is a random variable whose logarithm is normally distributed with the mean $\mu_{xel}$ and the variance $\sigma^2_{xel}$.

$$X_{el_i} = ln(S_{el_i}) \sim N(\mu_{xel}, \sigma^2_{xel}) \qquad i=1,\ldots,n$$

The specimen in this model can only be censured in one way and this when the stress level is below the endurance limit. For this model the probability that $Y(x_i)$ is censored is $1-\Phi\Big(\frac{x_i-\mu_{xel}}{\sigma_{xel}}\Big)$.

### 3.1 Likelihood function

The likelihood function for this model is

$$lik(\sigma_{xel}, \mu_{xel}) = \prod_{i \in U} \Phi\Big(\frac{x_i-\mu_{xel}}{\sigma_{xel}}\Big) \prod_{i \in C} \Big(1-\Phi\Big(\frac{x_i-\mu_{xel}}{\sigma_{xel}}\Big)\Big) \qquad (2)$$

### 3.2 Variance estimation

For this model the informations matrix $I$ is a two times two matrix the $(j,k)$ element of the observed information matrix is approximative

$$I_{jk} \approx -\frac{\partial^2}{\partial\theta''_j \partial\theta''_k} log\Big(lik(\boldsymbol{\theta''}, \boldsymbol{x})\Big)\Big|_{\boldsymbol{\theta''}=\hat{\boldsymbol{\theta}}''} \quad j=1,2 \quad k=1,2$$

where $\boldsymbol{\theta''} = (\mu_{xel}, \sigma_{xel})$.

The inverse of the approximative information matrix is again used to get the approximative variance estimates of $\mu_{xel}$ and $\sigma_{xel}$

$\widehat{Var}(\hat{\mu}_{xel}) \approx I^{-1}(1,1)$, $\widehat{Var}(\hat{\sigma}_{xel}) \approx I^{-1}(2,2)$ and $\widehat{Cov}(\hat{\mu}_{xel}, \hat{\sigma}_{xel}) \approx I^{-1}(1,2)$.

## 4 Model 1 and Model 2 under general assumptions

The two models can be described in a more general form. Assume that the logarithm of the fatigue limit is distributed according to a certain distribution $F_{xe}\Big(x;(\sigma_{xe},\mu_{xe})\Big)$, the endurance limit is given as $F_{xel}\Big(x;(\sigma_{xel},\mu_{xel})\Big)$ and the finite life as $F_\nu\Big(y;(\sigma_\nu,\nu(x))\Big)$ with derivative $f_\nu\Big(y;(\sigma_\nu,\nu(x))\Big)$.

The likelihood for Model 1 under the general distribution assumptions is

$$lik(\boldsymbol{\theta}) = \prod_{i\in U} f_\nu\Big(y_i;(\sigma_\nu,\nu(x_i))\Big) F_{xe}\Big(x_i;(\sigma_{xe},\mu_{xe})\Big) \cdot$$

$$\prod_{i\in C}\Big(1 - F_\nu\Big(y_i;(\sigma_\nu,\nu(x_i))\Big) F_{xe}\Big(x_i;(\sigma_{xe},\mu_{xe})\Big)\Big)$$

and for Model 2

$$lik(\sigma_{xel},\mu_{xel}) = \prod_{i\in U} F_{xel}\Big(x_i;(\sigma_{xel},\mu_{xel})\Big) \prod_{i\in C}\Big(1 - F_{xel}\Big(x_i;(\sigma_{xel},\mu_{xel})\Big)\Big)$$

# 5 Example 1

In this example the estimations are with the data in Appendix Table 5, the data are from a real staircase test. For the rest of this example $\nu(x)$ is assumed to be a straight line of the stress values, and the different distributions are all assumed to be normal. With the weight $(w_i)$ equal to one for the lowest stress level and with different weight increase for every 20 MPa increase in stress level. The result of maximizing the likelihood function numerically is in Table 1. The estimates of the endurance limit from Model 2 are $\hat{\sigma}_{xel} = 0.0571$ and $\hat{\mu}_{xel} = 6.3695$.

In this case the two parameters for the fatigue limit and the endurance limit almost coincide, the fatigue limit is somewhat lower than the endurance limit as it should be according to its definition. Both the median of the distribution of the limits are approximately $e^{6.369} = 583.47$ (MPa) and their coefficient of variance is approximately 0.06.

The reason why the two limits coincide here is that the $\Phi\left(\frac{y_o - \nu(x_i)}{\sigma_\nu / w_i}\right)$ factor in (1) is approximately one, i.e. the probability that the logarithm of the lives is below $y_o$ is approximately one and the difference between the two models disappears. This occurs when the finite lives are far below the endurance life $n_o$.

In Figure 1 the result of the example is shown.

The approximative variance estimates in Model 1 are in Table 2 and for Model 2 the corresponding estimates are $\widehat{Var}(\hat{\mu}_{xel}) = 0.0002060$ and $\widehat{Var}(\hat{\sigma}_{xel}) = 0.0006787$, $\widehat{Cov}(\hat{\mu}_{xel}, \hat{\sigma}_{xel}) = 0.00004633$.

Table 1: ML results for Example 1 with Model 1

| weight increase | $\hat{\sigma}_\nu$ | $\hat{\nu}(x)$ | $\hat{\sigma}_{xe}$ | $\hat{\mu}_{xe}$ |
|---|---|---|---|---|
| 0 | 0.7920 | $42.6965 - 4.7865 \cdot x$ | 0.0571 | 6.3693 |
| 0.4 | 1.0943 | $52.6207 - 6.3390 \cdot x$ | 0.0577 | 6.3691 |
| 0.8 | 1.3305 | $57.7412 - 7.1390 \cdot x$ | 0.0582 | 6.3690 |

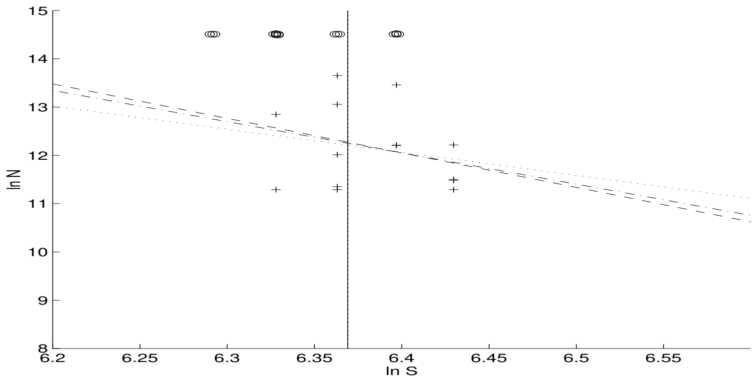


Figure 1: *The right vertical line is the endurance limit and the left vertical lines are the fatigue limit for different weights , the other lines are the $\hat{\nu}(x)$ for different weights, + denotes an uncensored specimen and o denotes a censored one.*

Table 2: Approximative variance estimates in Model 1 for Example 1

| weights increase | $\widehat{Var}(\hat{\mu}_{xe})$ | $\widehat{Var}(\hat{\sigma}_{xe})$ | $\widehat{Cov}(\hat{\mu}_{xe}, \hat{\sigma}_{xe})$ |
|---|---|---|---|
| 0 | 0.0002020 | 0.0005184 | 0.0000347 |
| 0.4 | 0.0001941 | 0.0005120 | 0.0000539 |
| 0.8 | 0.0001893 | 0.0005014 | 0.0000702 |

## 6 Example 2

In this example the estimations are with the data in Appendix Table 6, the data are from a real staircase test. The weights, $w_i$, and the line, $\nu(x)$, in this example are the same as in the previous example. The estimated values for the likelihood function (1) are in Table 3. The estimates for the endurance limit from Model 2 are $\hat{\sigma}_{xel} = 0.08286$ and $\hat{\mu}_{xel} = 6.4219$.

The approximative variance estimates in Model 1 are in Table 4 and for Model 2 the corresponding estimates are $\widehat{Var}(\hat{\mu}_{xel}) = 0.0005272$ and $\widehat{Var}(\hat{\sigma}_{xel}) = 0.002192$, $\widehat{Cov}(\hat{\mu}_{xel}, \hat{\sigma}_{xel}) = 0.0002935$.

In this example the median of the endurance limit is approximately $e^{6.422} = 615$ (MPa) and the median of the fatigue limit is approximately $e^{6.379} = 589$ (MPa) so they do not coincide as in the previous example. But the two limits are not the same and in this example the endurance limit is really the limit at $n_o$ and not at infinity as the fatigue limit is. Here the $\Phi\left(\frac{y_o-\nu(x_i)}{\sigma_\nu/w_i}\right)$ factor in (1) is not approximately one. Therefore the $\Phi\left(\frac{x_i-\mu_{xe}}{\sigma_{xe}}\right)$ factor in (1) increases and consequently $\mu_{xe}$ decreases and there is a difference between the two models . This occurs when some of the finite lives are close to the endurance life.

In Figure 2 the results of the example are shown.

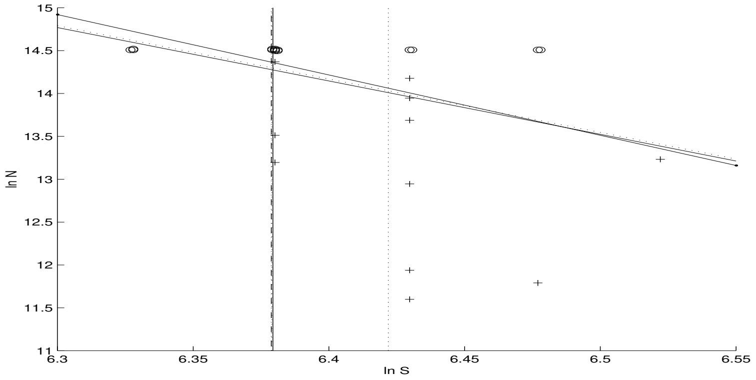


Figure 2: *The right vertical line is the endurance limit and the left vertical lines are the fatigue limit for different weights , the other lines are the $\hat{\nu}(x)$ for different weights, + denotes an uncensored specimen and o denotes a censored one.*

Table 3: ML results for Example 2 with Model 1

| weight increase | $\hat{\sigma}_\nu$ | $\hat{\nu}(x)$ | $\hat{\sigma}_{xe}$ | $\hat{\mu}_{xe}$ |
|---|---|---|---|---|
| 0 | 1.3905 | $53.9795 - 6.2239 \cdot x$ | 0.0100 | 6.3794 |
| 0.4 | 1.9162 | $58.9594 - 6.9925 \cdot x$ | 0.0094 | 6.3788 |
| 0.8 | 2.3231 | $60.3158 - 7.2016 \cdot x$ | 0.0092 | 6.3786 |

Table 4: Approximative variance estimates in Model 1 for Example 2

| weights increase | $\widehat{Var}(\hat{\mu}_{xe})$ | $\widehat{Var}(\hat{\sigma}_{xe})$ | $\widehat{Cov}(\hat{\mu}_{xe}, \hat{\sigma}_{xe})$ |
|---|---|---|---|
| 0 | 0.0012 | 0.2016 | -0.0141 |
| 0.4 | 0.0200 | 1.0096 | -0.1412 |
| 0.8 | 0.0363 | 1.4048 | -0.2249 |

# 7 Simulations

To see how good the estimated parameters are one can simulate a large number of staircase tests and see how much the estimated parameters deviate from the true ones.

The procedure which was used to simulate the staircase tests was as follows:
1. Choose a stress start level $s_1$, step length $s_{step}$ and a stress value $n_o$ where the test is interrupted.
2. Choose $\ln(S_e)$ randomly from a normal distribution with the expectation $\mu_{xe}$ and the variance $\sigma_{xe}^2$.
3. If $s_i < S_e$ then the specimen is censored and the logarithm of the lives length is put to $y_o$. Otherwise choose the logarithm of the lives length, $y_i$, from a normal distribution with the expectation $\alpha + \beta x_i$ and the variance $\sigma_\nu^2 / w_i^2$.
4. If $y_i > y_o$, then set $y_i = y_o$.
5. Let the next $s_{i+1}$ be equal to $s_i + s_{step}$ if the $y_i$ is censored, that is equal to $y_o$. Otherwise let the next $s_{i+1}$ be equal to $s_i - s_{step}$.
6. Restart the procedure from number 2.

Using the method above, 1000 tests were simulated as the procedure described above with the parameters chosen as $s_1$=540, $s_{step}$=20, $n_o$=$2\cdot10^6$, $\mu_{xe} = 6.37$, $\sigma_{xe} = 0.057$, $\nu(x)$ =43-5$\cdot x$, $\sigma_\nu$=0.8 and all $w_i$=1.

In the 1000 simulated tests there were eight tests which are not included in the following results. This is because they are not a realistic staircase test in the sense that they do not overlap. Overlapping does not occur if there is a stress level which separates the failures from the run-outs.

The estimates in the first model are $\hat{\sigma}_\nu$=0.73432, $\hat{\nu}(x)$=50.39832−6.15845⋅ $x$, $\hat{\sigma}_{xe}$=0.04971 and $\hat{\mu}_{xe}$=6.36982. In the second model they are $\hat{\sigma}_{xel}$= 0.04971 and $\hat{\mu}_{xel}$=6.36985.

The mean of the approximative variance estimators in the two models are the same

$$\widehat{\overline{Var}}\begin{pmatrix}\hat{\mu}_{xel}\\ \hat{\sigma}_{xel}\end{pmatrix} = \widehat{\overline{Var}}\begin{pmatrix}\hat{\mu}_{xel}\\ \hat{\sigma}_{xel}\end{pmatrix} = 10^{-4}\begin{pmatrix}1.85 & 0.44\\ 0.44 & 5.62\end{pmatrix}$$

where diagonal elements are variances and off diagonal covariances. The approximative variance estimators are the same since the endurance and fatigue limit are almost the same.

The empirically calculated variances for $\mu_{xe}$ and $\sigma_{xe}$ are 0.000203 and 0.000480 respective, and the covariance is 0.000031. The corresponding results for the endurance limit are 0.000202 and 0.000480 respective, and the covariance is 0.000030. The empirically calculated variances and the approximate calculated variances are of the same size and thus the approximate variance estimates seem to be useful.

# 8 Confidence interval for lower percentile of the fatigue limit and the endurance limit

One is interested in estimating the lower percentile of the fatigue limit which will be denoted by, $S_{e_p}$, where $p$ is the probability of the lower tail.

$$X_{e_p} = ln(S_{e_p}) = \mu_{xe} - \lambda_p \cdot \sigma_{xe} \text{ and } P(S_e \leq S_{e_p}) = p$$

The asymptotic distribution of

$$\frac{\hat{\mu}_{xe} - \lambda_p \cdot \hat{\sigma}_{xe} - (\mu_{xe} - \lambda_p \cdot \sigma_{xe})}{\sqrt{\widehat{Var}(\hat{\mu}_{xe} - \lambda_p \cdot \hat{\sigma}_{xe})}}$$

is standard normal distribution. Using this a one sided $(1-\alpha)100\%$ confidence interval for $X_{e_p}$ is

$$I = [\hat{\mu}_{xe} - \lambda_p \cdot \hat{\sigma}_{xe} - \lambda_\alpha \sqrt{\widehat{Var}(\hat{\mu}_{xe} - \lambda_p \cdot \hat{\sigma}_{xe})}; \infty]$$

where

$$\widehat{Var}(\hat{\mu}_{xe} - \lambda_p \cdot \hat{\sigma}_{xe}) = \widehat{Var}(\hat{\mu}_{xe}) + \lambda_p^2 \widehat{Var}(\hat{\sigma}_{xe}) - 2 \cdot \lambda_p \widehat{Cov}(\hat{\mu}_{xe}, \hat{\sigma}_{xe})$$

The lower estimate of $S_{e_p}$ is denoted by $\underline{S}_{e_p}$ and

$$\underline{X}_{e_p} = ln(\underline{S}_{e_p}) = \hat{\mu}_{xe} - \lambda_p \cdot \hat{\sigma}_{xe} - \lambda_\alpha \sqrt{\widehat{Var}(\hat{\mu}_{xe} - \lambda_p \cdot \hat{\sigma}_{xe})}$$

The lower percentile for the endurance limit, $S_{el_p}$, is calculated in the same way but with $\mu_{ex}$ and $\sigma_{xe}$ replaced by $\mu_{exl}$ and $\sigma_{xel}$.

In Appendix the above calculations are done with different values on $p$ and $\alpha$ in both examples and both models. In Example 1 $S_{e_p}$ and $S_{el_p}$ are the same, but $\underline{S}_{e_p}$ and $\underline{S}_{el_p}$ are different. The reason for this is that $\hat{\mu}_{xe}$ and $\hat{\mu}_{xel}$ respective $\hat{\sigma}_{xe}$ and $\hat{\sigma}_{xel}$ are almost the same but the approximative variance estimations are different. One can also see that $p$ and $\alpha$ influence each other. A higher value on $p$ can be compensated by a lower value on $\alpha$ to get the same value on $\underline{S}_{e_p}$. In Example 2 there is a big difference in $S_{e_p}$ and $S_{el_p}$ and an even bigger difference in $\underline{S}_{e_p}$ and $\underline{S}_{el_p}$. The value on $\underline{S}_{e_p}$ is too low to be used in practice. The reason for this low value is the big approximate variance estimators.

In the previous case of calculating the confidence interval two different numbers the percentile, $p$, and the confidence, $1-\alpha$, are used. One can get a lower bound $\underline{X}^*_{e_p}$ which will be exceeded with probability $1-q_1$, by using the following approximation. Decide that the probability that $S_e$ is greater than $\underline{S}^*_{e_p}$ has to be at least, $1-q_1$. Then use the following

$$P(S_e > \underline{S}^*_{e_p}) \geq P(S_e > S_{e_p} \cap S_{e_p} > \underline{S}^*_{e_p}) = P(S_e > S_{e_p}) P(S_{e_p} > \underline{S}^*_{e_p})$$

$$= (1-p) \cdot (1-\alpha) = 1 - q_1 \qquad (3)$$

The results of the above calculations and the corresponding for $S^*_{el_p}$ are in Appendix. For a given product $1-q_1$ one usually obtains a higher lower bound for a small $\alpha$ than for a small $p$.

One can also do a less approximate calculation for a lower bound $\underline{X}^*_{e_q} = \hat{\mu}_{xe} - a_e \hat{\sigma}_{xe}$ of $X_e$. By calculate $P(S_e > \underline{S}^*_{e_q})$ and let this probability be equal to $1-q_2$. This is the same thing as

$$P(X_e - \underline{X}^*_{e_q} > 0) \approx \Phi\left(\frac{E[X_e - \underline{X}^*_{e_q}]}{\sqrt{Var(X_e - \underline{X}^*_{e_q})}}\right) = 1 - q_2$$

where $E[X_e - \underline{X}^*_{e_q}] \approx a_e \sigma_{xe}$ and

$$Var(X_e - \underline{X}^*_{e_q}) \approx \sigma^2_{xe} + I^{-1}(1,1) + a_e^2 I^{-1}(2,2) - 2a_e I^{-1}(2,1)$$

Decide the exceedence probability $q_2$ one is interested in and solve Equation 4 with respect to $a_e$ and calculate $\underline{X}^*_{e_q}$.

$$\Phi\left(\frac{a_e \hat{\sigma}_{xe}}{\sqrt{\sigma^2_{xe} + I^{-1}(1,1) + a_e^2 I^{-1}(2,2) - 2a_e I^{-1}(2,1)}}\right) = 1 - q_2 \qquad (4)$$

The above calculations and the corresponding for $\underline{X}^*_{el_q}$ are shown in Appendix for different values on $q_2$. For Example 1 the two models gives almost the same result. In the case then $\hat{\sigma}_{xe}$ is smaller than $I^{-1}(2,2)$ equation 4 can only be solved for small values on $1-q_2$. This occurs for Example 2 with the estimates of Model 1.

To check the confidence interval 1000 staircase tests were simulated similar to the procedure described before with the parameters chosen as $s_1$=540, $s_{step}$=20, $n_o$=2·10$^6$, $\mu_{xe}$=6.3693, $\sigma_{xe}$=0.0571, $\nu(x)$=42.6965-4.7865·$x$, $\sigma_\nu$ =0.7920 and all $w_i$=1. In the 1000 simulated tests there were seven tests which are not included in the result because they does not overlap.

The number of times the confidence interval contains $X_{e_p}$ is counted. This number is denoted by $l_e$. The simulations show that the coverage for $X_{e_p}$ is too low i.e. the estimate of $X_{e_p}$ is too high, probably due to the bias in $\hat{\sigma}_{xe}$. The results above and the corresponding results for Model 2 are shown in Appendix.

The theoretical p-value, $P(S_e \leq S_{e_p}) = p$, was compared to the mean of the simulated $\Phi(\frac{\hat{X}_{ep} - \mu_{xe}}{\sigma_{xe}})$. This shows also that the estimated $X_{e_p}$ is too high. The results above and the corresponding results for $X_{el_p}$ are in Appendix.

The mean of the simulated $1-\Phi(\frac{\hat{\underline{X}}^*_{e_p}-\mu_{xe}}{\sigma_{xe}})$ is compare to $P(S_e > \underline{S}^*_{e_p}) \geq (1-p)\cdot(1-\alpha)$ and is bigger than $(1-p)\cdot(1-\alpha)$ as it should be. These calculations and the corresponding results for the endurance limit are shown in Appendix.

$P(S_e > \underline{S}^*_{e_q}) \approx 1-q_2$ and $P(S_e > \underline{S}^*_{el_q}) \approx 1-q_2$ are checked in a similar way. These calculations shows that the lower bound of the fatigue limit $\underline{X}^*_{e_q}$ and $\underline{X}^*_{el_q}$ are estimated too high. The results of these calculations are also shown in Appendix.

# 9 Logistic distribution for the logarithm of the fatigue limit and the endurance limit

The estimates for $\mu_{xe}$, $\mu_{xel}$, $\sigma_{xe}$ and $\sigma_{xel}$ almost seems to be the same if the distribution of the logarithm of the fatigue and the endurance limit are logistic distribution instead of normal distribution.

# 10 Conclusions

As we have shown it is possible to estimate the distribution of the fatigue limit although the fatigue limit itself is impossible to observe. One never knows if a failure will occur eventually. In order to make the estimate possible finite lives have to be considered.

The finite lives seem to give information in some cases and in some cases just the information that they are finite. According to this it is useful to use the finite lives in some cases and in some cases one does not seem to gain anything. But the two models do not estimate the same, Model 1 estimate the fatigue limit and not the endurance limit.

The finite lives give information when some specimens break near the endurance life $n_o$. This is because the $\Phi\left(\frac{y_o-\nu(x_i)}{\sigma_\nu/\sqrt{w_i}}\right)$ factor in (1) is not approximately equal to one, and accordingly functions (1) and (2) do not coincide. In this case the fatigue limit estimate is lower than the endurance limit estimate. The reason for this is that in this case the endurance limit is really the limit at $n_o$ cycles and not at infinity. The fatigue limit, on the other hand, corresponds to infinite life. If interest is in the fatigue limit rather than in endurance limit at $n_o$ cycles the staircase test was carried out at too low run-out level.

The estimate of the lower percentile for the endurance limit and the fatigue limit seems to be estimated too high according to the simulated confidence intervals. To get a lower bound which will be exceeded with some probability it is therefore better to use the approximations in (3) than to use a confidence interval for a specific percentile.

## 11 Acknowledgements

The support from ITM, Sandvik Steel AB and ABB Stal AB is gratefully acknowledged. I would also like to thank Professor Jacques de Maré (Mathematical Statistics, Chalmers University of Technology), Dr Thomas Svensson (Swedish National testing and Research Institute) and Dr Boel Wadman (IVF) for their help throughout this work.

# A Appendix

In Table 5 is the data for exampel one and in Table 6 are the corresponding for exampel two. In Table 7, 8 the calculations are done with the estimates from example 1. In Table 9, 10 are the corresponding for the estimates from example 2. In Table 11 to 13 are the different confidence intervals for the simulated data.

Table 5: Data for Example 1, $n_o = 2 \cdot 10^6$

| specimen No | S | N $\cdot 10^3$ | specimen No | S | N $\cdot 10^3$ |
|---|---|---|---|---|---|
| 1 | 540 | $n_o$ | 16 | 600 | $n_o$ |
| 2 | 560 | 80 | 17 | 620 | 202 |
| 3 | 540 | $n_o$ | 18 | 600 | $n_o$ |
| 4 | 560 | 380 | 19 | 620 | 80 |
| 5 | 540 | $n_o$ | 20 | 600 | $n_o$ |
| 6 | 560 | $n_o$ | 21 | 620 | 99 |
| 7 | 580 | $n_o$ | 22 | 600 | 201 |
| 8 | 600 | 700 | 23 | 580 | $n_o$ |
| 9 | 580 | 80 | 24 | 600 | $n_o$ |
| 10 | 560 | $n_o$ | 25 | 620 | 97 |
| 11 | 580 | 468 | 26 | 600 | 200 |
| 12 | 560 | $n_o$ | 27 | 580 | 85 |
| 13 | 580 | 165 | 28 | 560 | $n_o$ |
| 14 | 560 | $n_o$ | 29 | 580 | 850 |
| 15 | 580 | $n_o$ | 30 | 560 | $n_o$ |

Table 6: Data for Example 2, $n_o = 2 \cdot 10^6$

| specimen No | S | N $\cdot 10^3$ | specimen No | S | N $\cdot 10^3$ |
|---|---|---|---|---|---|
| 1 | 590 | $n_o$ | 14 | 560 | $n_o$ |
| 2 | 620 | 1140 | 15 | 590 | $n_o$ |
| 3 | 590 | 1740 | 16 | 620 | $n_o$ |
| 4 | 560 | $n_o$ | 17 | 650 | 132 |
| 5 | 590 | $n_o$ | 18 | 620 | 109 |
| 6 | 620 | 419 | 19 | 590 | $n_o$ |
| 7 | 590 | $n_o$ | 20 | 620 | 1436 |
| 8 | 620 | 880 | 21 | 590 | $n_o$ |
| 9 | 590 | 538 | 22 | 620 | $n_o$ |
| 10 | 560 | $n_o$ | 23 | 650 | $n_o$ |
| 11 | 590 | $n_o$ | 24 | 680 | 538 |
| 12 | 620 | 153 | 25 | 650 | $n_o$ |
| 13 | 590 | 739 | | | |

Table 7: Confidence interval and lower bounds with estimates from Model 1

| p | $\alpha$ | $\underline{S}_{e_p}$ | $S_{e_p}$ | $1-q_1$ | $\underline{S}^*_{e_p}$ | $1-q_2$ | $\underline{S}^*_{e_q}$ |
|---|---|---|---|---|---|---|---|
| 0.01 | 0.01 | 451 | 511 | 0.98 | 451 | 0.98 | 476 |
| 0.01 | 0.05 | 468 | 511 | 0.94 | 468 | 0.94 | 521 |
| 0.05 | 0.01 | 486 | 531 | 0.94 | 486 | | |
| 0.05 | 0.05 | 499 | 531 | 0.90 | 499 | 0.90 | 535 |

Table 8: Confidence interval and lower bounds with estimates from Model 2

| p | $\alpha$ | $\underline{S}_{el_p}$ | $S_{el_p}$ | $1-q_1$ | $\underline{S}^*_{el_p}$ | $1-q_2$ | $\underline{S}^*_{el_q}$ |
|---|---|---|---|---|---|---|---|
| 0.01 | 0.01 | 444 | 511 | 0.98 | 444 | 0.98 | 424 |
| 0.01 | 0.05 | 463 | 511 | 0.94 | 463 | 0.94 | 515 |
| 0.05 | 0.01 | 480 | 531 | 0.94 | 480 | | |
| 0.05 | 0.05 | 495 | 531 | 0.90 | 495 | 0.90 | 533 |

Table 9: Confidence interval and lower bounds with estimates from Model 1

| p | $\alpha$ | $\underline{S}_{e_p}$ | $S_{e_p}$ | $1-q_1$ | $\underline{S}^*_{e_p}$ |
|---|---|---|---|---|---|
| 0.01 | 0.01 | 47 | 576 | 0.98 | 47 |
| 0.01 | 0.05 | 97 | 576 | 0.94 | 97 |
| 0.05 | 0.01 | 98 | 580 | 0.94 | 98 |
| 0.05 | 0.05 | 163 | 580 | 0.90 | 163 |

Table 10: Confidence interval and lower bounds with estimates from Model 2

| p | $\alpha$ | $\underline{S}_{el_p}$ | $S_{el_p}$ | $1-q_1$ | $\underline{S}^*_{el_p}$ | $1-q_2$ | $\underline{S}^*_{el_q}$ |
|---|---|---|---|---|---|---|---|
| 0.01 | 0.01 | 397 | 507 | 0.98 | 397 | | |
| 0.01 | 0.05 | 427 | 507 | 0.94 | 427 | 0.94 | 482 |
| 0.05 | 0.01 | 452 | 537 | 0.94 | 452 | | |
| 0.05 | 0.05 | 475 | 537 | 0.90 | 475 | 0.90 | 531 |

Table 11: The results of the simulations to check the confidence interval of $X_{e_p}$ and the lower bound $\underline{X}^*_{e_p}$

| $1-\alpha$ | $\frac{l_e}{993}$ | p | Mean of $\Phi(\frac{\hat{X}_{ep}-\mu_{xe}}{\sigma_{xe}})$ | $(1-q_1)$ | Mean of $1-\Phi(\frac{\hat{\underline{X}}^*{}_{ep}-\mu_{xe}}{\sigma_{xe}})$ |
|---|---|---|---|---|---|
| 0.99 | 0.8359 | 0.01 | 0.0536 | 0.9801 | 0.9930 |
| 0.95 | 0.7533 | 0.01 | 0.0536 | 0.9405 | 0.9876 |
| 0.99 | 0.8510 | 0.05 | 0.1094 | 0.9405 | 0.9768 |
| 0.95 | 0.7503 | 0.05 | 0.1094 | 0.9025 | 0.9640 |

Table 12: The results of the simulations to check the confidence interval of $X_{el_p}$ and the lower bound $\underline{X}^*_{el_p}$

| $1-\alpha$ | $\frac{l_{el}}{993}$ | p | Mean of $\Phi(\frac{\hat{X}_{elp}-\mu_{xel}}{\sigma_{xel}})$ | $(1-q_1)$ | Mean of $1-\Phi(\frac{\hat{\underline{X}}^*{}_{elp}-\mu_{xel}}{\sigma_{xel}})$ |
|---|---|---|---|---|---|
| 0.99 | 0.8328 | 0.01 | 0.0535 | 0.9801 | 0.9928 |
| 0.95 | 0.7492 | 0.01 | 0.0535 | 0.9405 | 0.9873 |
| 0.99 | 0.8499 | 0.05 | 0.1099 | 0.9405 | 0.9762 |
| 0.95 | 0.7422 | 0.05 | 0.1099 | 0.9025 | 0.9632 |

Table 13: The results of the simulations to check the lower limits $\underline{X}^*_{e_q}$ and $\underline{X}^*_{el_q}$

| $a_e$ | $1-q_2$ | Mean of $1-\Phi(\frac{\hat{\underline{X}}^*_{e_q}-\mu_{xe}}{\sigma_{xe}})$ | $a_{el}$ | $1-q_2$ | Mean of $1-\Phi(\frac{\hat{\underline{X}}^*_{el_q}-\mu_{xel}}{\sigma_{xel}})$ |
|---|---|---|---|---|---|
| 2.0537 | 0.98 | 0.9286 | 2.0537 | 0.98 | 0.9287 |
| 1.5548 | 0.94 | 0.8792 | 1.5548 | 0.94 | 0.8798 |
| 1.2816 | 0.90 | 0.8397 | 1.2816 | 0.90 | 0.8407 |

## References


[1] F. A. Bastenaire: New method for the statistical evaluation of constant stress amplitude fatigue test result, ASTM STP 511, American Society for Testing and Materials, pp. 3-28 (1972).

[2] S. Nishijima: Statistical Analysis of Small Sample Fatigue Data, Statistical Research on Fatigue and Fracture, pp. 1-19 (1987).

[3] F.G. Pascual, W.Q. Meeker: Estimating Fatigue Curves With the Random Fatigue-Limit Model, Technometrics Vol. 41, No 4. pp. 277-302 (1999).

[4] W.J Dixon, A. M. Mood: A method for obtaining and analyzing sensitivity data, Journal of the American Statistical Association, Vol. 43 (1948).

[5] S. Beretta, Y. Murakami: Statistical analysis of defects for fatigue strength prediction and quality control of materials, Fatigue & Fracture of Engineering Materials & Structures pp. 1049-1065 (1998).

[6] W. Elber: The significance of fatigue crack closure. In Damage tolerance in aircraft structures ASTM STP 486, American Society of Testing and Materials, Philadelphia, USA, pp. 230-242 (1971).

# Estimating inclusion distributions of hard metal using fatigue tests

**Abstract**

The maximum inclusion size in clean steels influences fatigue behaviour and other mechanical properties. In this paper different approaches to estimating the distribution of the inclusion sizes which cause the failure are compared. The data used are from a staircase test. The load levels on the broken and unbroken specimens as well as the size of the inclusions which cause the failures are recorded. One model uses only the load levels and the other two models also use the recorded inclusion size on the broken specimens. The difference between the models using the inclusion size is that one is based on the assumption that the specimen breaks in the largest inclusion, and the other one on the assumption that the specimen breaks in some inclusion over a critical size, not necessarily the largest one.

It is shown that from a staircase test with measured inclusion size on the broken specimens it is better to use both load levels and inclusion sizes to estimate the inclusion size distribution instead of using just the different load levels. The two models using the inclusion sizes almost coincide for big inclusions. These models are also used to verify or reject possible fatigue limit models.

**Keywords:** Maximum likelihood, Extreme value, Staircase test

## 1 Introduction

The mechanical properties of hard and clean metals are in part governed by the inclusion size distribution, e.g. the fatigue limit is determined by the size of the maximum projected area on the plan perpendicular to the load [1].

Direct observations of inclusions sizes are difficult since the inclusions are rare and often occur in the interior of the specimen. A usual way of observing is the following, a number of areas of the same size are observed on polished cross-sections by optical microscopy. The largest inclusion in each area is identified and measured and then the extreme value theory is used to extrapolate the maximum inclusion size distribution of the specimen of interest. In [2] the method of estimating the inclusion size distribution is based on measuring the maximum size in randomly chosen areas or volumes. In [3] the method is based on measuring the size of inclusions larger than a chosen size in a small volume.

The result from a usual staircase test is the number of failures and survivors of different load levels. In [4] and [5] the distribution of the inclusions was estimated using an estimate of the endurance limit and the relation between the endurance limit and the inclusion size [1].

The inclusion size distribution can also be estimated by direct observations on the maximal inclusion size in the specimens or by the load levels from a staircase test. Here the inclusion size and the load level are combined through a staircase test. The load levels on the unbroken specimens are used. For the broken specimens the inclusion which caused the failure is used. This is possible since if the fatigue failure starts at the maximal projected area perpendicular to the load [1], direct observations of this inclusion are possible to carry out on the failed specimens by measuring the inclusion where the specimen breaks. The inclusion sizes are used under two different assumptions, the specimens breaks in the largest inclusion or in some of the inclusions over some critical size.

According to general statistical theory maximal sizes are approximately distributed according to one extreme value distribution. Here the Gumbel and the Frechet extreme value distribution are applied. As a comparison the consequence of assuming lognormal fatigue strength is also calculated. This method is often used in practice [6].

## 2 Inclusion size distribution estimate based on fatigue test data

According to [1] one can use relation (1) below, the so called ($\sqrt{area_{max}}$-model) to determine the fatigue limit, where $HV$ is the Vickers hardness in Hv, $R$ is the stress ratio $S_{min}/S_{max}$ and $area_{max}$ is the largest inclusion area projected onto a plan perpendicular to the applied load in $\mu m^2$ in the volume of the material that is subjected to the maximum stress. The constant $K$ is depending on the location of the inclusion, $K = 1.43$ for surface inclusion and $K = 1.56$ for internal inclusion. The exponent $\alpha$ is given as $\alpha = (HV \cdot 10^{-4} + 0.226)$. The fatigue limit $S_e$ is given by

$$S_e = K(HV + 120)\sqrt{area_{max}}^{-1/6}\left(\frac{1-R}{2}\right)^{\alpha} \quad (1)$$

Equation (1) can be written as $S_e = h(A) = qA^{-1/6}$ where $q$ is a function of $HV$ and $R$, and $A$ is the square root of the largest projected inclusion area perpendicular to the load, $\sqrt{area_{max}}$.

The aim is to estimate the inclusion size distribution from observations on the load levels of the specimens and equation (1), using the information that if the specimen is broken the load level is above the fatigue limit and vice versa for the unbroken specimens. This is the same as $h^{-1}(s)$ is less than the inclusion sizes for the broken specimens and bigger than the inclusion sizes for the unbroken specimens on the stress level $s$. Hence the relation between the cumulative probability distribution function of $S_e$ and $A$ is

$$F_{S_e}(s) = \mathrm{P}(qA^{-1/6} \leq s) = 1 - F_A((s/q)^{-6}) = 1 - F_A(h^{-1}(s)) \qquad (2)$$

Note that for each load level there is a critical inclusion size for which inclusion sizes below this size will not cause failure, $h^{-1}(s)$. This size will also be denoted $a_c$.

From the staircase test one has $k$ different load levels and the number of survivors and failures at each level. Let $F_A(a)$ denote the cumulative distribution for the maximum inclusion size in the specimen, $n_i$ the number of failures and $m_i$ the number of survivors at load level $s_i$. The likelihood of the experiment in terms of the inclusion size distribution is therefore

$$lik = \prod_{i=1}^{k} \left(1 - F_A(h^{-1}(s_i))\right)^{n_i} \left(F_A(h^{-1}(s_i))\right)^{m_i} \qquad (3)$$

Here two different distributions for the inclusion size are considered, the Gumbel and the Frechet distributions with threshold zero. The likelihood function when using the Gumbel assumption is

$$lik(\mu, \sigma) = \prod_{i=1}^{k} \left(1 - e^{-e^{-\frac{h^{-1}(s_i)-\mu}{\sigma}}}\right)^{n_i} \left(e^{-e^{-\frac{h^{-1}(s_i)-\mu}{\sigma}}}\right)^{m_i} \qquad (4)$$

and the likelihood of the Frechet assumption is

$$lik(\beta,\lambda)=\prod_{i=1}^{k}\left(1-e^{-(\frac{h^{-1}(s_i)}{\beta})^{-\lambda}}\right)^{n_i}\left(e^{-(\frac{h^{-1}(s_i)}{\beta})^{-\lambda}}\right)^{m_i} \tag{5}$$

As a comparison one can estimate the fatigue limit distribution and then use relation (1) to estimate the inclusion size distribution. One way of estimating the fatigue limit distribution is to assume that the logarithm of the fatigue strength follows a normal distribution $X_e = ln(S_e) \sim N(\mu_e, \sigma_e)$ and that the specimen breaks if the load level is above the fatigue limit [6]. The likelihood for estimating the fatigue limit distribution is then

$$lik(\mu_e,\sigma_e)=\prod_{i=1}^{k}\left(F_{X_e}(ln(s_i))\right)^{n_i}\left(1-F_{X_e}(ln(s_i))\right)^{m_i} \tag{6}$$

The difference between this way and the former ways of estimating the inclusion size distribution is that the former are based on the assumption that the inclusion sizes have some specified distribution and the latter method is based on the assumption that the fatigue limit has some specified distribution.

# 3 Inclusion size distribution estimate based on fatigue test data augmented with inclusion size measurements

These models estimate the inclusions size distribution from the load levels of the unbroken specimens, inclusion sizes of the broken specimens and equation (1) under different assumptions.

The size of the inclusion where the specimen breaks is measured and is assumed to be the largest inclusion in that specimen. The test data used in this model are also $k$ different load levels and the number of survivors and failures at each load level. For the broken specimens there are also the inclusions sizes $a_{i1}, \cdots, a_{in_i}$. The inclusion sizes for the unbroken specimens are all assumed to be below $h^{-1}(s_i)$, the critical size $i = 1, \cdots, k$.

Let again $F_A(a)$ be the cumulative distribution function for the inclusion size and $f_A(a)$ is the corresponding density function. The likelihood for the experiment under these assumptions is

$$lik = \prod_{i=1}^{k} \Big( \prod_{j=1}^{n_i} f_A(a_{ij}) \Big) \Big( F_A(h^{-1}(s_i)) \Big)^{m_i} \tag{7}$$

Assuming that the inclusion size follows a Gumbel distribution one gets

$$lik = \prod_{i=1}^{k} \Big( \prod_{j=1}^{n_i} \frac{1}{\sigma} e^{-\frac{a_{ij}-\mu}{\sigma}} e^{-e^{-\frac{a_{ij}-\mu}{\sigma}}} \Big) \Big( e^{-e^{-\frac{h^{-1}(s_i)-\mu}{\sigma}}} \Big)^{m_i} \tag{8}$$

If instead one assumes that the inclusion size follows a Frechet distribution with threshold zero, one gets

$$lik = \prod_{i=1}^{k} \Big( \prod_{j=1}^{n_i} \frac{\lambda}{\beta} (\frac{a_{ij}}{\beta})^{-\lambda-1} e^{-(\frac{a_{ij}}{\beta})^{-\lambda}} \Big) \Big( e^{-(\frac{h^{-1}(s_i)}{\beta})^{-\lambda}} \Big)^{m_i} \tag{9}$$

In this model the specimens are assumed to break in the largest inclusion. Assume now that there is more than one inclusion over the critical size $a_c$ and that the specimen breaks in one of these with the same probability for each of them. To model this the following assumptions are reasonable (see [3]). Assume that the number of inclusions in the specimen that are greater then $a_c$ are Poisson distributed with the parameter $\nu$. The conditional probability that an inclusion is not greater than $a$, given that it is at least $a_c$, follows a Generalized Pareto distribution with the cumulative distribution function

$$F_{A|A \geq a_c}(a) = 1 - \Big( 1 - \frac{\xi(a - a_c)}{\sigma} \Big)^{1/\xi}$$

where $\sigma > 0$, $0 < a - a_c < \frac{\sigma}{\xi}$ if $\xi > 0$ and $0 < a - a_c < \infty$ if $\xi \leq 0$.

Then the inclusion size over some threshold is Generalized Pareto distributed and the number of inclusions $N$ over that threshold is Poisson distributed. When $\xi = 0$, the maximum inclusion size follows a Gumbel distribution [7]. For $\xi < 0$, the maximum inclusion size follows a Frechet distribution and for $\xi > 0$ it has a Weibull distribution. Therefore this assumption does not contradict the previous assumptions.

The unconditional distribution for an inclusion size $A$, is

$$
\begin{aligned}
\mathrm{P}(A \le a) &= \sum_{n=0}^{\infty} \mathrm{P}(A \le a | N = n)\mathrm{P}(N = n) = \\
1 \cdot \mathrm{P}(N = 0) &+ \sum_{n=1}^{\infty} \mathrm{P}(A \le a | N = n)\mathrm{P}(N = n) = \\
\mathrm{P}(N = 0) &+ \sum_{n=1}^{\infty} \mathrm{P}(A_1 \le a)\mathrm{P}(N = n) = \\
\mathrm{P}(N = 0) &+ \mathrm{P}(A_1 \le a)\Big(1 - \mathrm{P}(N = 0)\Big) = \\
e^{-\nu} &+ \Big(1 - \Big(1 - \frac{\xi(a - a_c)}{\sigma}\Big)^{1/\xi}\Big)(1 - e^{-\nu}) = \\
1 - (1 - e^{-\nu})&\Big(1 - \frac{\xi(a - a_c)}{\sigma}\Big)^{1/\xi}, \qquad a > a_c.
\end{aligned}
$$

Note that $\mathrm{P}(A_1 \le a_c) = e^{-\nu} > 0$. Here $A_1$ follows conditionally a Generalized Pareto distribution and denotes the size of an arbitary inclusion above the threshold.

For $a > a_c$ the density function of the inclusion size is

$$
f_A(a) = \frac{1}{\sigma}(1 - e^{-\nu})\Big(1 - \frac{\xi(a - a_c)}{\sigma}\Big)^{1/\xi - 1}.
$$

For $\xi = 0$ the previous density function becomes exponential above the threshold, i.e.

$$
f_A(a) = \frac{1}{\sigma}(1 - e^{-\nu})(e^{-\frac{(a - a_c)}{\sigma}}), \qquad a > a_c.
$$

The critical size $a_{ci}$, the expected number of inclusions $\nu_i$ over that size and the scale parameter $\sigma_i$, for the inclusion size distribution depend on the load level $s_i$. The density function corresponding to the load level $s_i$ is

$$f_A(a) = \frac{1}{\sigma_i}(1 - e^{-\nu_i})\Big(1 - \frac{\xi(a - a_{ci})}{\sigma_i}\Big)^{1/\xi - 1}, \qquad a > a_{ci}, \qquad i = 1, \cdots, k.$$

The critical size for each level is determined by the load level, $a_{ci} = h^{-1}(s_i)$. Denote the smallest critical size as $a_{c0}$ corresponding to the highest load level. The different scale parameters can be related to the scale parameter which corresponds to the lowest critical size in the following way $\sigma_i = \sigma_0 - \xi(a_{ci} - a_{c0})$ for $i = 1, \cdots, k$, where $\sigma_0$ is the scale parameter corresponding to the lowest critical size, i.e. the largest load level.

There is a relation between the different $\nu_i$ values $i = 1, \cdots, k$. Each $\nu_i$ value can be related to $\nu_0$, the expected number of inclusions over $a_{c0}$. The expected number of inclusions over $a_{ci}$ is $\nu_i = \nu_0 \mathrm{P}(A > a_{ci} | A > a_{c0})$ for $i = 1, \cdots, k$ [7].

Hence

$$\nu_i = \nu_0\Big(1 - \frac{\xi(h^{-1}(s_i) - a_{c0})}{\sigma_0}\Big)^{1/\xi}, \qquad h^{-1}(s_i) \geq a_{c0} \qquad i = 1, \cdots, k.$$

The likelihood for this model is

$$\prod_{i=1}^{k}\Big(\prod_{j=1}^{n_i} f_A(a_{ij})\Big)\Big(\mathrm{P}(N_i = 0)\Big)^{m_i} =$$

$$\prod_{i=1}^{k}\Big(\prod_{j=1}^{n_i}\frac{1}{\sigma_i}(1 - e^{-\nu_i})\Big(1 - \frac{\xi(a_{ij} - h^{-1}(s_i))}{\sigma_i}\Big)^{1/\xi - 1}\Big)(e^{-\nu_i})^{m_i} \tag{10}$$

The likelihood in (10) holds when all inclusion sizes are bigger than their corresponding critical value.

For $\xi = 0$ all scale parameters are the same and the expected number of inclusions over the smallest critical sizes $a_o$ is

$$\nu_i = \nu_0 e^{-\frac{(h^{-1}(s_i) - a_{c0})}{\sigma}}, \qquad h^{-1}(s_i) \geq a_{c0} \qquad i = 1, \cdots, k.$$

The likelihood when $\xi$ is zero is therefore

$$\prod_{i=1}^{k}\Big(\prod_{j=1}^{n_i}\frac{1}{\sigma}\Big(1-e^{-\nu_0 e^{-\frac{(h^{-1}(s_i)-a_{c0})}{\sigma}}}\Big)\Big(e^{-\frac{(a_{ij}-h^{-1}(s_i))}{\sigma}}\Big)\Big)\Big(e^{-\nu_0 e^{-\frac{(h^{-1}(s_i)-a_{c0})}{\sigma}}}\Big)^{m_i} \tag{11}$$

Note that the likelihood in (10) uses one more parameter to describe the inclusion size distribution than the other likelihoods. But later in the estimation procedure we will fix $\xi = 0$ and just fit two parameters, $\nu_0$ and $\sigma$.

Note that in order to apply the described likelihoods, the observed inclusion size $a_{ij}$ of the failure number $j$ on the load level $s_i$ has to satisfy the consistency condition that $a_{ij}$ is bigger than the critical size $a_{ci} = h^{-1}(s_i)$.

# 4 Simulation

To study the systematic and random error when estimating the 95 % quantile of the distribution of the inclusion size, a simulation study of a staircase test with 30 specimens is carried out. The staircase test was simulated under four different assumptions:
1. Size distribution is assumed to be Gumbel.
2. Size distribution is assumed to be Frechet with $\lambda = 3$.
3. Size distribution is assumed to be exponential above the threshold, i.e. Generalized Pareto with $\xi = 0$.
4. Size distribution is assumed to be Generalized Pareto with $\xi = -0.1$.

For each of the four size distributions the following models are fitted:
1. Size distribution is assumed to be Gumbel with and without measured inclusion sizes, i.e. likelihood (8) and (4) respectively.
2. Size distribution is assumed to be Frechet with and without measured inclusion sizes, i.e. likelihood (9) and (5) respectively.
3. Size distribution is assumed to be Generalized Pareto, with $\xi = 0$, i.e. likelihood (11).
4. The Logarithm of the fatigue limit follows a normal distribution, i.e. likelihood (6).

The procedure which was used to simulate the staircase test when the inclusion sizes are assumed to be Gumbel or Frechet distribution was as follows:
1. Choose a start level $s_1$, step length $s_{step}$ and a constant $q$.
2. Choose $A$ randomly from a Frechet or a Gumbel distribution.
3. If $s_i < qA^{-1/6}$ the specimen is unbroken and $s_{i+1}$ is equal to $s_i + s_{step}$.
4. If $s_i > qA^{-1/6}$ the specimen is broken and $s_{i+1}$ is equal to $s_i - s_{step}$ and save the inclusion size $A$.
5. Restart the procedure from number 2.

Using the method above, 1000 staircase tests were simulated as the procedure described above with Frechet distribution and 1000 staircase tests with Gumbel distribution. The parameters were chosen so that the simulated data in some sense correspond to the data sets in Appendix C and D. The parameter values were chosen as $s_1 = 800$, $s_{step} = 20$ and $q = 1350$. For the Gumbel distribution the scale parameter $\sigma = 7$ and the location parameter $\mu = 20$. For the Frechet distribution the location parameter $\alpha = 0$, the scale parameter $\beta = 15$ and the shape parameter $\lambda = 3$. Under the Gumbel assumption the expectation and the variance for the inclusion size are 24 and 81 respectively and for the Frechet assumption the corresponding are 20 and 190 respectively.

To simulate the staircase test when the inclusion size over some threshold follows a Generalized Pareto distribution, the following procedure was used:
1. Choose a start level $s_1$, step length $s_{step}$, a constant $q$, expected number of inclusions over a threshold $\nu_o$ and the parameters in the Generalized Pareto distribution.
2. Calculate $\nu_i$, the expected number of inclusions over $a_{ci}$.
3. Choose the number of inclusions over the critical size $N$ randomly from a Poisson distribution with parameter $\nu_i$.
4. If $N = 0$ the specimen is unbroken and $s_{i+1}$ is equal to $s_i + s_{step}$.
5. If $N > 0$ the specimen is broken and $s_{i+1}$ is equal to $s_i - s_{step}$ and choose $A$ randomly from a Generalized Pareto distribution with threshold $a_{ci}$.
6. Restart the procedure from number 2.

Using the method above, 1000 staircase tests were simulated according to the procedure described above. The parameter values were chosen as $s_1 = 800$, $s_{step} = 20$, $q = 1350$, $\sigma = 8$, $\xi = 0$ and $\nu_0 = 10$. The expected number of inclusions over the threshold $a_0$ corresponds to a load level at 1000 MPa.

A simulation was also carried out with the same parameters value except that $\xi$ was -0.1.

In Tables 1 to 4 the mean and standard deviation for the 95 % quantile of the inclusion size distribution are shown for all six models on all four data sets. In Table 1 the simulated staircase test is from the method of assuming that the inclusion size follows a Gumbel distribution, Table 2 the inclusion size follows a Frechet distribution, Table 3 the inclusion size follows a Generalized Pareto distribution with $\xi = 0$ and Table 4 the inclusion size follows a Generalized Pareto distribution with $\xi = -0.1$. The model which uses the direct measured inclusion sizes likelihood (7) is denoted by 'With' in the Tables. The model which just uses the load levels likelihood (3) is denoted by 'Without'. Frechet and Gumbel in the Tables stand for which assumption one has on the inclusion size. The model which uses the Generalized Pareto distribution likelihood (11) is denoted by 'Exponential'. In the calculations $\xi$ is assumed to be zero all the time the distribution is also exponential. The model which assumes that the logarithm of the fatigue limit is normally distributed, i.e. likelihood (6), is denoted by 'Normal'.

According to Tables 1 to 4 it seems better to use the models which use the measured inclusion sizes. If it is better to use the exponential model or the model which assumes that the specimens break in the largest inclusion depends on the distribution on the inclusion size. The exponential model is not always the best one but it seems as if the estimate does not get too wrong even if $\xi$ is always assumed to be zero. The model which uses the inclusion size under Gumbel and Frechet distribution is more dependent on which inclusion size distribution which is the right one.

Table 1: Gumbel distribution with 95 % quantile 40.79

| Distribution assumption | mean | std |
|---|---|---|
| A $\sim$ Gumbel with | 40.47 | 4.24 |
| A $\sim$ Gumbel without | 37.81 | 9.17 |
| A $\sim$ Frechet with | 43.25 | 5.48 |
| A $\sim$ Frechet without | 48.70 | 47.75 |
| A $\sim$ Exponential | 38.86 | 4.40 |
| log(S) $\sim$ Normal | 38.83 | 19.50 |

Table 2: Frechet distribution with 95 % quantile 40.37

| Distribution assumption | mean | std |
|---|---|---|
| A $\sim$ Gumbel with | 38.86 | 9.03 |
| A $\sim$ Gumbel without | 29.55 | 7.99 |
| A $\sim$ Frechet with | 40.03 | 8.53 |
| A $\sim$ Frechet without | 37.25 | 39.23 |
| A $\sim$ Exponential | 37.79 | 9.23 |
| log(S) $\sim$ Normal | 30.55 | 16.68 |

Table 3: Exponential distribution 95 % quantile 42.13

| Distribution assumption | mean | std |
|---|---|---|
| A $\sim$ Gumbel with | 44.59 | 4.98 |
| A $\sim$ Gumbel without | 53.28 | 15.03 |
| A $\sim$ Frechet with | 47.83 | 6.30 |
| A $\sim$ Frechet without | 128.66 | 1005.4 |
| A $\sim$ Exponential | 42.69 | 5.30 |
| log(S) $\sim$ Normal | 62.24 | 93.01 |

Table 4: Pareto distribution 95 % quantile 42.16

| Distribution assumption | mean | std |
|---|---|---|
| A $\sim$ Gumbel with | 50.15 | 6.83 |
| A $\sim$ Gumbel without | 53.38 | 16.38 |
| A $\sim$ Frechet with | 54.00 | 8.45 |
| A $\sim$ Frechet without | 121.33 | 628.57 |
| A $\sim$ Exponential | 48.61 | 7.28 |
| log(S) $\sim$ Normal | 64.04 | 107.88 |

The different estimated inclusion size distributions from two of the simulated staircase tests under Gumbel assumption on the inclusion size are compared to the estimated empirical distribution function $F_n$ and are shown in Figures 1 to 4. To estimate $F_n(i)$ the Kaplan-Meier method [8] is used. Since the data are left censored instead of right censored, the order number is changed to fit to the Kaplan-Meier method. Multiply all observations by minus one and order them $a_{(1)} < a_{(2)} < \cdots < a_{(n)}$. Let $J_a$ denote the set of all indices $j$ where $a_{(j)} \leq -a$ and $a_{(j)}$ represent an observed inclusion size. Let $n_{(j)}$ denote the number of specimens, with observed or possible inclusion sizes less than $a_{(j)}$, $j = 1, \cdots, n$. Then the Kaplan-Meier estimate of F(a) is defined as $\hat{F}(a) = \prod_{j \in J_a} \frac{n_j - 1}{n_j}$.

In Figures 1 and 3 the inclusion size distribution is estimated without the observed inclusion size (3) from two of the simulated staircase tests under Gumbel assumption on the inclusion size, denoted by 'one' and 'two'. The solid line is under Gumbel assumption on the inclusion size, i.e. right distribution assumption. The dotted line is under Frechet assumption on the inclusion size and the dash-dot line is under the assumption that the logarithm of the fatigue limit is normally distributed. The dashed line is the correct inclusion size distribution. The estimated empirical distribution is denoted by crosses.

In Figures 2 and 4 the inclusion size distribution is estimated with the observed inclusion size (7) from the same staircase tests as in Figures 1 and 3. The solid line is under Gumbel assumption on the inclusion size. The dotted line is under Frechet assumption on the inclusion size and the dash-dot line is under the assumption that the logarithm of the fatigue limit is normally distributed the same as before. The dashed line is the correct inclusion size distribution. The estimated empirical distribution is denoted by crosses. The shorter solid line is from using the Generalized Pareto distribution, with $\xi = 0$ (11), i.e. exponential.

Looking at Figures 1 to 4 one sees that the estimated inclusion size distribution from the models using the inclusion sizes best fits the true inclusion size distribution. In the simulation number two the models using the inclusion sizes are much better than the models that just use the different load levels. In simulation number one the difference is not that large.

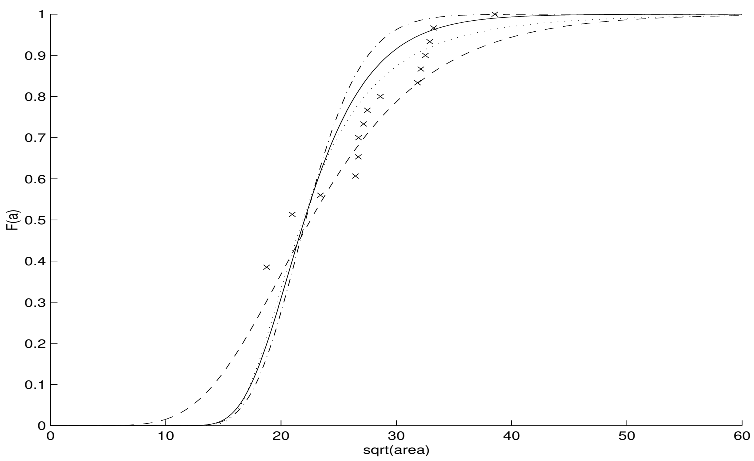


Figure 1: *Cumulative distribution functions for the inclusion size from simulated staircase test sizes unobserved, realization one. Estimated Gumbel —, Frechet ···, Lognormal $- \cdot -$, Correct Gumbel - - -.*

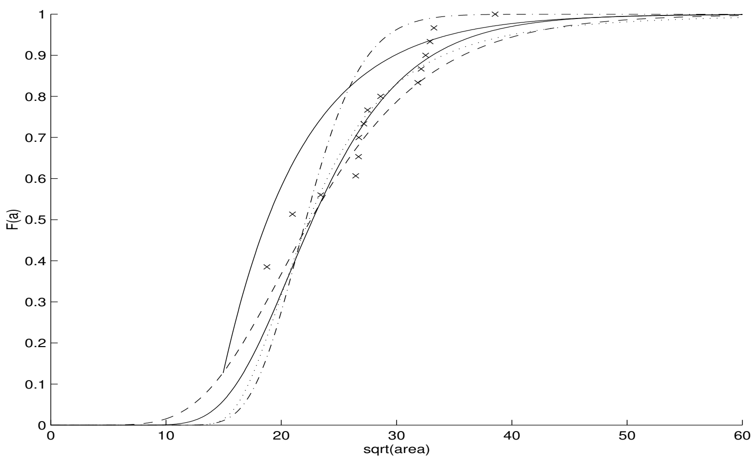


Figure 2: *Cumulative distribution functions for the inclusion size from simulated staircase test sizes observed, realization one. Estimated Gumbel —, Frechet ···, Lognormal $- \cdot -$, Correct Gumbel - - -, Exponential — to the left.*

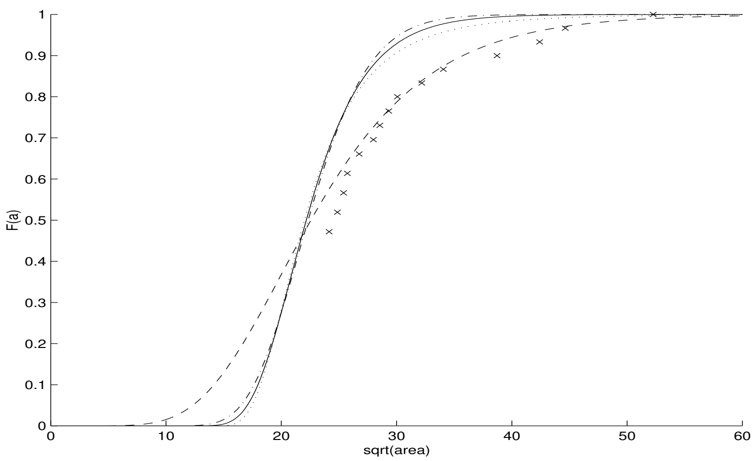


Figure 3: *Cumulative distribution functions for the inclusion size from simulated staircase test sizes unobserved, realization two. Estimated Gumbel —, Frechet $\cdots$, Lognormal $- \cdot -$, Correct Gumbel - - -.*

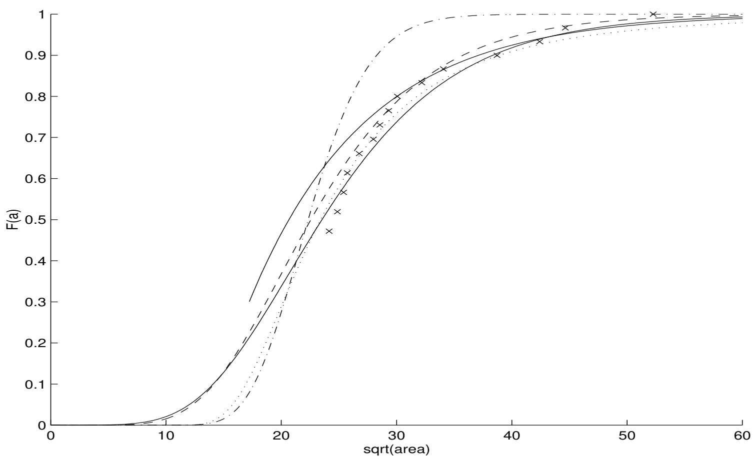


Figure 4: *Cumulative distribution functions for the inclusion size from simulated staircase test sizes observed, realization two. Estimated Gumbel —, Frechet $\cdots$ , Lognormal $- \cdot -$, Correct Gumbel - - -, Exponential — to the left.*

# 5 Example

The different ways of estimating the inclusion size distributions are used to estimate the inclusion size distribution on four different data sets. The data sets are from three different steels and are described in Appendices A to D. These different data will be denoted by data set A, B, C and D. One problem with these data sets is that only for data set A the inclusion sizes are greater than the critical size $a_c$, for all inclusions. Especially for the data set in data set B this criterion is not fulfilled.

The different estimated inclusion size distribution is compared to the estimated empirical distribution function, $F_n$. To estimate $F_n(i)$ the Kaplan-Meier method [8] is used as before.

In Figures 5 to 8 the inclusion size distribution is estimated without the observed inclusion size (3). The solid line is under Gumbel assumption on the inclusion size. The dotted line is under Frechet assumption on the inclusion size and the dash-dot line is under the assumption that the logarithm of the fatigue limit is normally distributed. The estimated empirical distribution is denoted by crosses.

The estimated distribution function for the inclusion size distribution does not fit the estimated empirical distribution too well. This is particularly the case for the estimates of the data from data set B. For these data the inclusions are not inclusions but pores in material and a great number of specimens have broken although they should not have done that. This may be a reason why the empirical distribution does not fit the estimated distribution. When the probability for negative inclusion sizes is small in the estimated Gumbel distribution, the estimated inclusion size distribution under the Gumbel and the lognormal assumption is essentially the same. For the data set in data set C there is one specimen which has broken although it should not have done that and for data set D there are three specimens. One possible explanation for this is a measurement error on the inclusion size or the load level.

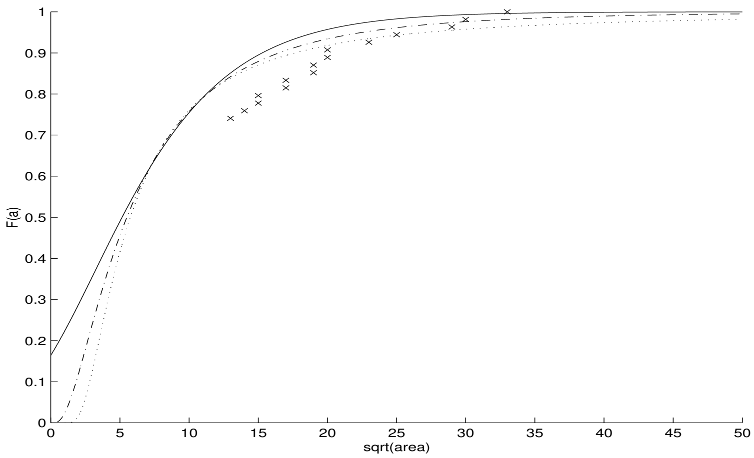


Figure 5: *Data from data set A sizes unobserved. Estimated Gumbel —, Frechet $\cdots$, Lognormal $- \cdot -$.*

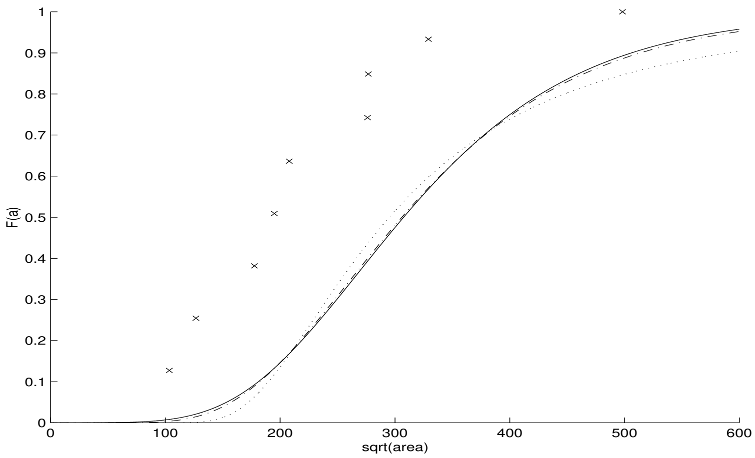


Figure 6: *Data from data set B sizes unobserved. Estimated Gumbel —, Frechet $\cdots$, Lognormal $- \cdot -$.*

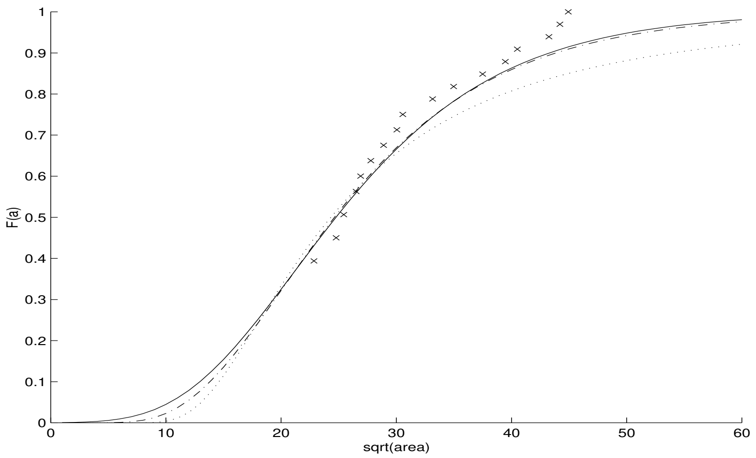


Figure 7: *Data from data set C sizes unobserved. Estimated Gumbel —, Frechet $\cdots$, Lognormal $- \cdot -$.*

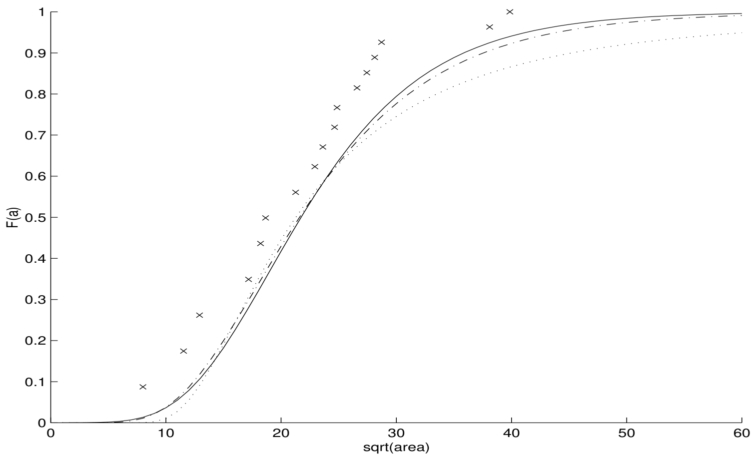


Figure 8: *Data from data set D sizes unobserved. Estimated Gumbel —, Frechet $\cdots$, Lognormal $- \cdot -$.*

In Tables 5 to 7 the 95 % quantile, the expected value and the standard deviation for the inclusion size under the different distribution assumptions on the inclusion size are shown. Under the Frechet assumption on the inclusion size the variance is $\beta^2(\Gamma(1-2/\lambda)-\Gamma(1-1/\lambda)^2)$ if $\lambda > 2$ and else infinite. Since $\hat{\lambda} < 2$ when using the dataset in data set A the standard deviation is not computed for this data set.

Table 5: The 95 % quantile of the inclusion size

| Distribution assumption | Data set A | Data set B | Data set C | Data set D |
|---|---|---|---|---|
| A $\sim$ Gumbel | 19.16 | 581.87 | 50.37 | 41.27 |
| A $\sim$ Frechet | 27.09 | 768.12 | 73.23 | 60.39 |
| log(S) $\sim$ Normal | 22.50 | 593.21 | 51.55 | 43.86 |

Table 6: The expected value of the inclusion size

| Distribution assumption | Data set A | Data set B | Data set C | Data set D |
|---|---|---|---|---|
| A $\sim$ Gumbel | 0.08 | 208.10 | 15.48 | 14.66 |
| A $\sim$ Frechet | 10.13 | 365.97 | 32.17 | 27.38 |
| log(S) $\sim$ Normal | 7.94 | 331.01 | 27.21 | 23.66 |

Table 7: The standard deviation of the inclusion size

| Distribution assumption | Data set A | Data set B | Data set C | Data set D |
|---|---|---|---|---|
| A $\sim$ Gumbel | 6.9 | 135.1 | 12.6 | 9.6 |
| A $\sim$ Frechet | | 304.5 | 40.1 | 28.5 |
| log(S) $\sim$ Normal | 8.3 | 139.5 | 12.9 | 10.7 |

In Figures 9 to 12 the dash-dot lines are under the assumption that the logarithm of the fatigue limit is normally distributed, the same as before. The other two estimated distribution functions are from using the inclusion size of the broken specimens, (7). The solid line is under Gumbel assumption on the inclusion size, dotted line is under Frechet assumption on the inclusion size. The crosses are the estimated distribution function with the

Kaplan-Meier method.

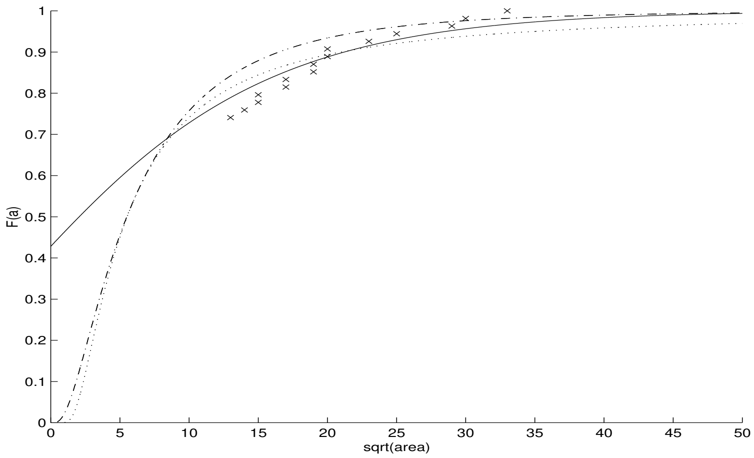


Figure 9: *Estimated cumulative distribution functions for the inclusion sizes on data from data set A inclusion observed. Estimated Gumbel —, Frechet ···, Lognormal − · −.*

The estimated distributions seem to fit quite well the empirical distribution under the Gumbel and Frechet assumption. For the estimate from data set A under the Gumbel assumption the distribution is somewhat strange, since there is quite a big probability that the inclusion size is less than zero. This is a problem with the Gumbel distribution since it is defined for negative sizes as well. If the inclusions are big this is no problem. This material is a very clean steel with few and small inclusions so this can be the reason why the probability for negative inclusions is quite big.

Tables 8 to 10 below correspond to Tables 5 to 7 above but with the likelihood which uses the inclusions size on the broken specimens (7). The standard deviation under Frechet assumption for data set A is not computed for the same reason as before.

The model which uses the Generalized Pareto distribution is only tested on data set A, because this is the only data set which is consistent with the assumptions of the model.

The estimates for the Generalized Pareto distribution are made when $\xi$ is assumed to be zero and are obtained as $\hat{\sigma} = 11.3727$ and $\hat{\nu}_0 = 0.4654$.

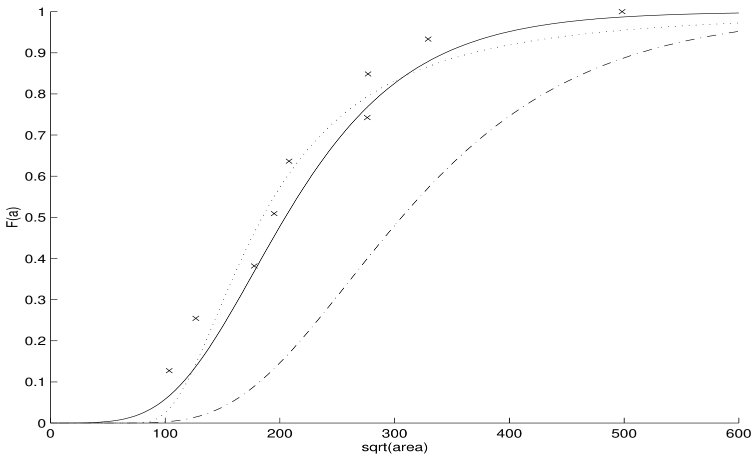


Figure 10: *Estimated cumulative distribution functions for the inclusion sizes on data from data set B inclusion observed. Estimated Gumbel —, Frechet $\cdots$, Lognormal $- \cdot -$.*

Table 8: The 95 % quantile of the inclusion size

| Distribution assumption | Data set A | Data set B | Data set C | Data set D |
|---|---|---|---|---|
| A $\sim$ Gumbel | 28.55 | 397.52 | 45.90 | 36.60 |
| A $\sim$ Frechet | 34.96 | 479.81 | 51.91 | 49.95 |
| log(S) $\sim$ Normal | 22.50 | 593.21 | 51.55 | 43.86 |

The lowest critical inclusion size is then $a_{c0} = 5.3040$. The expected value of the inclusion size is 6.21 and variance 113.11.

In Figure 13 the dash-dot line shows the estimated inclusion size distribution using (11) and the solid line shows the corresponding inclusion size distribution using (8). The crosses are again the estimated distribution function from the Kaplan-Meier method. The distribution function from (11) and that from (8) almost seem to be the same for large inclusion sizes. On the basis of the estimate from (11) there seem to be few inclusions over the critical size. Therefore the assumption for the previous model that the specimen breaks in the largest inclusions seems reasonable.

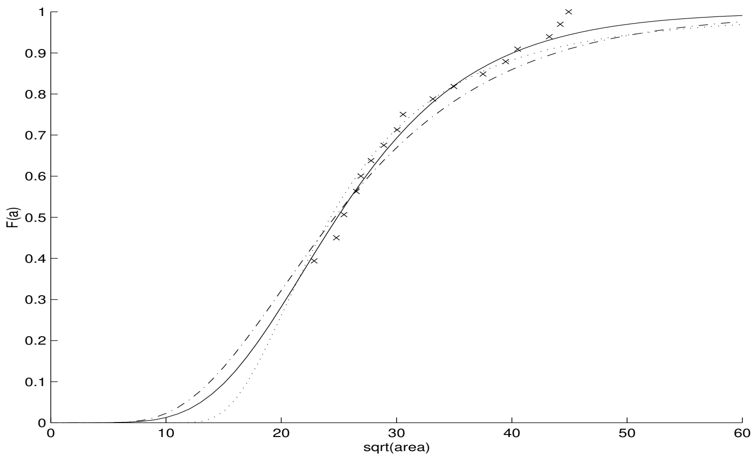


Figure 11: *Estimated cumulative distribution functions for the inclusion sizes on data from data set C inclusion observed. Estimated Gumbel —, Frechet $\cdots$, Lognormal $- \cdot -$.*

Table 9: The expected value of the inclusion size

| Distribution assumption | Data set A | Data set B | Data set C | Data set D |
|---|---|---|---|---|
| A $\sim$ Gumbel | -7.55 | 134.69 | 17.24 | 13.03 |
| A $\sim$ Frechet | 13.14 | 228.81 | 28.03 | 22.72 |
| log(S) $\sim$ Normal | 7.94 | 331.01 | 27.21 | 23.66 |

Table 10: The standard deviation of the inclusion size

| Distribution assumption | Data set A | Data set B | Data set C | Data set D |
|---|---|---|---|---|
| A $\sim$ Gumbel | 13.1 | 95.0 | 10.4 | 8.5 |
| A $\sim$ Frechet | | 189.7 | 15.1 | 23.3 |
| log(S) $\sim$ Normal | 8.3 | 139.5 | 12.9 | 10.7 |

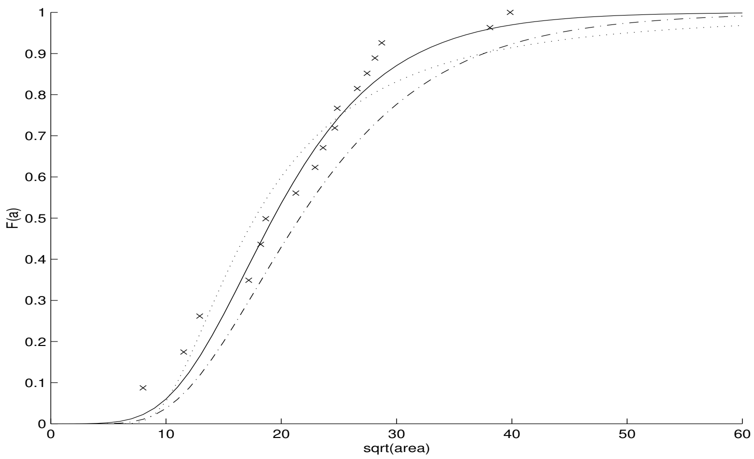


Figure 12: *Estimated cumulative distribution functions for the inclusion sizes on data from data set D inclusion observed. Estimated Gumbel —, Frechet $\cdots$, Lognormal $-\cdot-$.*

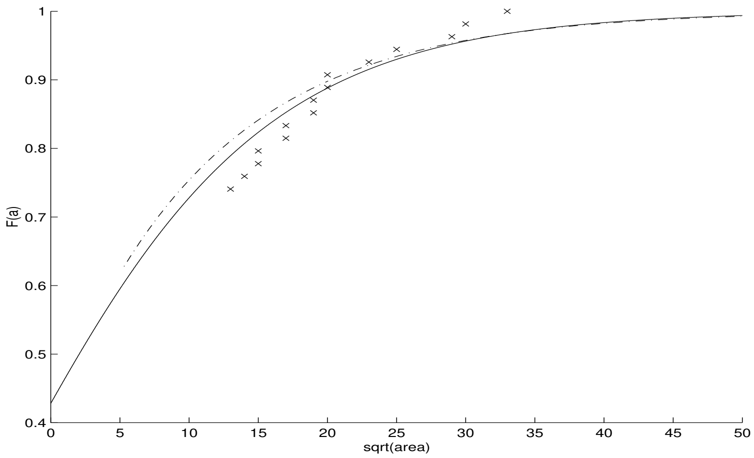


Figure 13: *Estimated cumulative distribution functions for the inclusion sizes on data from data set A. Estimated Gumbel —, Exponential $-\cdot-$.*

## 5.1 Comparison for estimation with and without direct size measurements.

The inclusion size distribution has been estimated by different approaches with and without direct size measurement of the inclusion sizes. All estimated distribution functions of the data from data set A are shown in Figure 14 except the one which uses the assumption that the specimen does not necessarily break in the largest inclusion. In Figure 15 the corresponding estimates are from the data in data set B and in Figure 16 the estimates are from data set C and in Figure 17 from data set D.

When the specimens have broken the fatigue limit is below the load level, $s$, using (2) one gets that the inclusion sizes is above $h^{-1}(s)$. The only data that this condition holds for are those in data set A. For the rest of the data sets some specimens have broken although they should not have done that. This is particularly true of the data in data set B, the specimens have broken at too small inclusions. The results presented from now on in this section do not concern the data in data set B.

Looking at Figures 14 to 17 the model using the inclusion sizes seems to fit the estimated empirical distribution better. For most of the data sets according to Tables 5 to 10 the model which uses only the different load levels has a higher value for the 95 % quantile, the variance and the expected value of the inclusion size.

For the fatigue strength the big inclusions are the interesting ones. Looking at a large inclusion $a$, the different $F_A(a)$ values do not differ too much according to the figures. If instead one takes a large value on $F_A(a)$ there is a big difference in the inclusion sizes for the different models. How much the different estimated inclusions size distributions differ from each other depends on the way the comparsion is made.

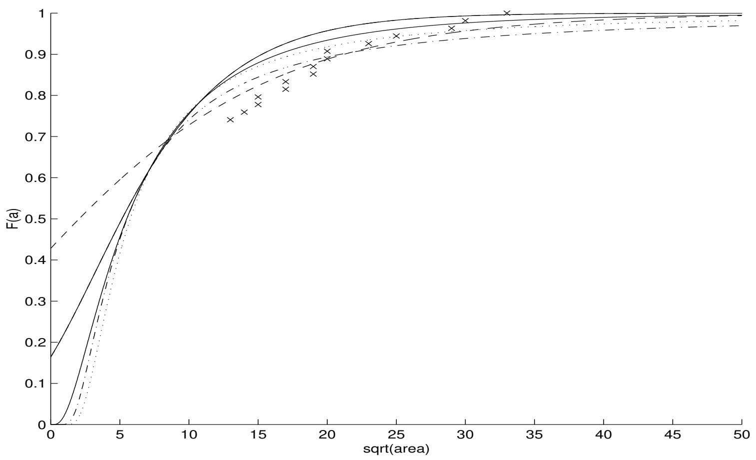


Figure 14: *Cumulative distribution functions for the inclusion size. This figure contains Figure 5 together with 9 , data set A.*

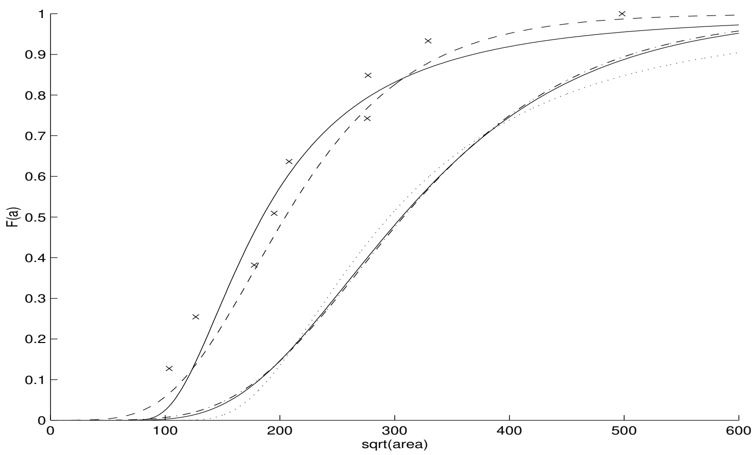


Figure 15: *Cumulative distribution functions for the inclusion size. This figure contains Figure 6 together with 10, data set B.*

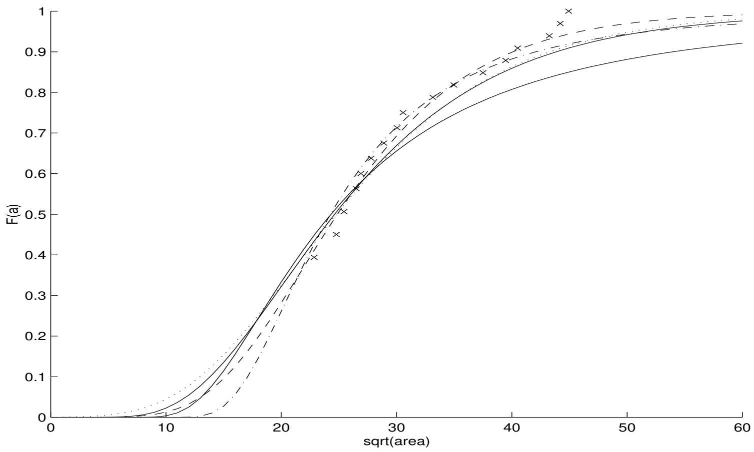


Figure 16: *Cumulative distribution functions for the inclusion size. This figure contains Figure 7 together with 11, data set C.*

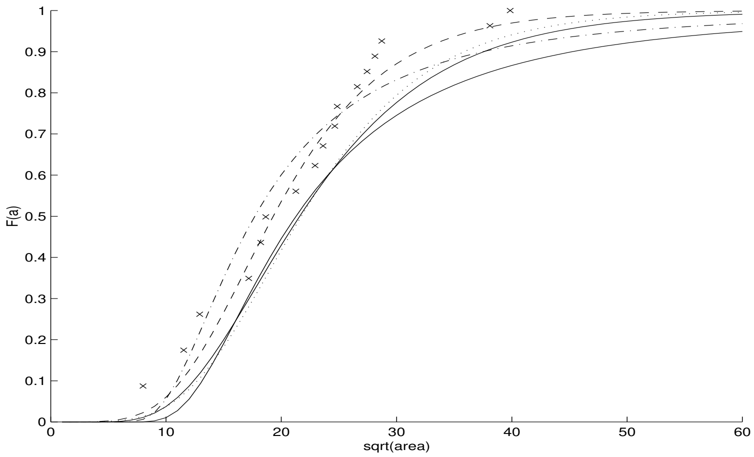


Figure 17: *Cumulative distribution functions for the inclusion size. This figure contains Figure 8 together with 12, data set D.*

## 6 Conclusions

The distribution of the inclusion size where the failure occurs has been estimated using different approaches. It seems as if, according to the simulations, the models which use the inclusion sizes for the broken specimens better estimate the wanted distribution. When estimating the size distribution and using the models above, there is no information in the load level for the broken specimens when the actual size is observed.

As mentioned above, it is better to use the models which use the measured inclusions sizes. Which is the best model depends on the distribution assumption on the inclusion sizes. If one has a correct distribution assumption, the model which uses Gumbel or Frechet distribution seems to be better than the model which uses Generalized Pareto distribution. The Generalized Pareto model seems to be least sensitive to a wrong distribution assumption on the inclusion sizes. This model also gives an estimate of how many inclusions there are over the critical size. When there are few inclusions over that size the difference between the models which use the inclusion size is small. If there is only one inclusion over the critical size and the specimen breaks it has to break in that inclusion.

The sizes of the inclusions are hard to measure as mentioned above. From a staircase test the size of the inclusion which caused the failure can be measured. Hence the staircase method can be used to find big inclusions.

From the measured inclusion size on the broken specimens it can also be concluded whether or not the $\sqrt{area_{max}}$-model is valid. This can be done by comparing the size of the inclusion causing the failure to the critical inclusion size on the load level in question. For the studied data sets one data set clearly violated this condition.

## 7 Acknowledgements

I would like to thank Professor Jacques de Maré (Mathematical Statistics, Chalmers University of Technology), Dr Thomas Svensson (Swedish National testing and Research Institute) and Dr Pär Johannesson (Mathematical Statistics, Chalmers University of Technology) for their help throughout this work. This paper was initiated at the Stochastic Centre workshop at Smögen August 14-25, 2000.

## References

[1] Y. Murakami, S. Beretta: Small defects and inhomogeneities in fatigue strength: Experiments, Models and Statistical implications, Extremes, Vol. 2, pp. 123-147 (1999).

[2] Y. Murakami: Inclusion rating by statistics of extreme values and its application to fatigue strength prediction and quality control of materials, Journal of Research of the National Institute of Standards and Technology, Vol. 99, pp. 345-351 (1994).

[3] C.W. Anderson, H.V. Atkinson, C.M. Sellars: The precision of methods using the statistics of extremes for the estimation of the maximum size of inclusions in clean steels, Acta Materialia, Vol. 48, pp. 4235-4246 (2000).

[4] H. Bomas, P. Mayr: Inclusion size distribution and endurance limit of a hard steel, Extremes, Vol. 2, pp. 149-164 (1999).

[5] T. Svensson, J. de Maré: Random features of the fatigue limit, Extremes, Vol. 2, pp. 165-176 (1999).

[6] W.J Dixon, A. M. Mood: A method for obtaining and analyzing sensitivity data, Journal of the American Statistical Association, Vol. 43 (1948).

[7] G. Shi, V. Atkinson, C. M. Sellars, C. W. Anderson: Application of the generalized pareto distribution to the estimation of the size of the maximum inclusion in clean steels, Acta Materialia Vol. 47, pp. 1455-1468 (1999).

[8] A. Høyland, M. Rausand: System reliability theory, (1994).

# A Appendix

## A.1 Data set A

These data have been obtained from staircase tests on a material with a hardness of 561 Hv, $R = -0.5$ with surface inclusions and hence $K = 1.43$. The constant $q$ is 898. The units for the load levels are MPa and for the square root of the inclusion size it is $\mu m$. In Table 11, the load levels and the square root of the inclusion size for the broken specimen are given. They are denoted by $s$ and $\sqrt{area}$ respectively. In Table 12 the number of unbroken specimens $n$ at each load level $s$, are presented.

Table 11:

| $s$ | $\sqrt{area}$ | $s$ | $\sqrt{area}$ | $s$ | $\sqrt{area}$ |
|---|---|---|---|---|---|
| 620 | 20 | 680 | 15 | 600 | 25 |
| 640 | 17 | 640 | 19 | 600 | 33 |
| 640 | 17 | 620 | 30 | 600 | 15 |
| 640 | 23 | 640 | 29 | 600 | 19 |
| 660 | 13 | 620 | 14 | 600 | 20 |

Table 12:

| $s$ | $n$ |
|---|---|
| 600 | 18 |
| 620 | 10 |
| 640 | 8 |
| 660 | 3 |

## A.2 Data set B

These data are from staircase tests on a material with a hardness of 276 Hv, $R = -1$ with surface inclusions and hence $K = 1.43$. The constant $q$ is 566. The units for the load level is MPa and for the square root of the inclusion size it is $\mu m$ . In Table 13, the load levels and the square root of the inclusion sizes for the broken specimen are given. In Table 14 the load level for the unbroken specimens and the number of broken at each load level, are shown.

Table 13:

| $s$ | $\sqrt{area}$ | $s$ | $\sqrt{area}$ | $s$ | $\sqrt{area}$ |
|---|---|---|---|---|---|
| 210 | 195 | 240 | 127 | 210 | 276 |
| 220 | 329 | 230 | 208 | 250 | 103 |
| 230 | 277 | 220 | 498 | 240 | 178 |

Table 14:

| $s$ | $n$ |
|---|---|
| 200 | 2 |
| 210 | 3 |
| 220 | 2 |
| 230 | 1 |

### A.3 Data set C

These data are from staircase tests on a material with a hardness of 745 Hv, $R = -1$ with surface inclusions and hence $K = 1.43$. The constant $q$ is 1349. The units for the load level is MPa and for the square root of the inclusion size it is $\mu m$. In Table 15, the load levels and the square root of the inclusion sizes for the broken specimen are given. In Table 16 the load level for the unbroken specimens and the number of broken at each load level are shown.

Table 15:

| s | $\sqrt{area}$ | s | $\sqrt{area}$ | s | $\sqrt{area}$ |
|---|---|---|---|---|---|
| 830 | 35.0 | 800 | 24.8 | 780 | 40.5 |
| 830 | 26.5 | 800 | 33.2 | 780 | 44.2 |
| 830 | 39.5 | 800 | 44.9 | 780 | 26.9 |
| 830 | 43.3 | 800 | 30.0 | 780 | 22.8 |
| 830 | 25.4 | 800 | 30.6 | 750 | 37.5 |
| 830 | 27.8 | 800 | 28.9 | | |

Table 16:

| $s$ | $n$ |
|---|---|
| 830 | 3 |
| 800 | 3 |
| 780 | 5 |
| 750 | 5 |

### A.4 Data set D

These data are from staircase tests on a material with a hardness of 745 Hv, $R = -1$ with surface inclusions and hence $K = 1.43$. The constant $q$ is 1349. The units for the load level is MPa and for the square root of the inclusion size it is $\mu m$. In Table 17, the load level and the square root of the inclusion size for the broken specimen are given. In Table 18 the load level for the unbroken specimens and the number of broken at each load level are shown.

Table 17:

| $s$ | $\sqrt{area}$ | $s$ | $\sqrt{area}$ | $s$ | $\sqrt{area}$ |
|---|---|---|---|---|---|
| 1200 | 11.5 | 1000 | 21.3 | 800 | 18.2 |
| 1200 | 38.1 | 1000 | 22.9 | 800 | 28.1 |
| 1200 | 24.6 | 1000 | 17.2 | 800 | 12.9 |
| 1200 | 8.0 | 830 | 28.7 | 780 | 39.9 |
| 1000 | 23.6 | 830 | 27.4 | 780 | 24.9 |
| 1000 | 26.6 | 830 | 18.7 | | |

Table 18:

| $s$ | $n$ |
|---|---|
| 830 | 2 |
| 800 | 3 |
| 780 | 5 |